Explicit control of subgeometric ergodicity

Christophe Andrieu* Gersende Fort†

Key words Quantitative bounds of convergence, Markov chains, Subgeometric ergodicity, drift conditions, coupling time.

1 Introduction

This paper discusses quantitative bounds on the convergence rates of Markov chains with general state space, under conditions implying subgeometric ergodicity. Our conditions make the present results of great interest for real applications, such as those requiring uniform-in-θ ergodicity for a family of kernels $\{P_\theta : \theta \in \Theta\}$. This question arises in different fields of applied probability. In recent works, such a uniform-in-θ behavior was assumed to study convergence of adaptive Markov Chain Monte-Carlo algorithms (see [1] (resp. [14]) where quantitative bounds for geometric ergodicity (resp. geometric and subgeometric ergodicity) are required).

Let $\{\Phi_n \ : n \geq 0\}$ be a Markov chain on a general state space X equipped with a countably generated σ-field $\mathcal{B}(\mathsf{X})$. We consider convergence of the iterates of the transition kernel P to a stationary distribution, when convergence is in total variation norm and in f-norm as well. For a measurable function $f : \mathsf{X} \to [1, \infty)$, define the f-norm of a (signed) measure ν as $\sup_{g, |g| \leq f} |\nu(g)|$. The total variation norm corresponds to the f-norm when f is the constant function $f = \mathbf{1}$ and is denoted by $\| \cdot \|_{\mathrm{TV}}$. A ψ-irreducible and aperiodic transition kernel that possesses an invariant probability distribution π is ergodic :

$$\forall x \in \mathsf{X}, \qquad \lim_n \|P^n(x, \cdot) - \pi(\cdot)\|_{\mathrm{TV}} = 0.$$

When $\pi(f) < \infty$ for some measurable function $f \geq 1$, then convergence occurs in f-norm and the kernel is said f-ergodic. If this convergence occurs at a rate $r = \{r(n) \ : n \geq 0\}$ for some non-decreasing sequence of positive real numbers, the kernel is (f, r)-ergodic. Two classes of rate functions are considered in the literature : the first one concerns the geometric sequence $r(n) = \kappa^n$ for some $\kappa > 1$; the second one concerns the subgeometric sequence such that $0 < \liminf_n r(n)/\tilde{r}(n) \leq \limsup_n r(n)/\tilde{r}(n) < \infty$ where $\tilde{r}$ is a rate sequence with the property

*University of Bristol, School of Mathematics, University Walk, BSS 1TW, UK.
†CNRS/ENST, Dpt TSI, 46, rue Barrault, 75634 Paris Cedex 13, France.

$\log \tilde{r}(n)/n \downarrow 0$. This class includes polynomial sequences $r(n) \propto n^\kappa$, $\kappa \geq 0$, logarithmic sequences $r(n) \sim \log(n)^\alpha$ and sub-exponential sequences $\ln r(n) \propto n^\kappa$ for some $0 < \kappa < 1$ (see [19] for a precise definition of these rates and a list of elementary properties).
Sufficient conditions were proposed to determine the rate and the norm of convergence of a transition kernel to its stationary distribution. Most of them are related to the recurrence of special sets called *small set* : $\mathcal{C}$ is a m*-small set* (or simply *small set*) for the kernel P if there exist $m \geq 1$, $\epsilon > 0$ and a probability measure ν on X such that

$$P^m(x, \cdot) \geq \epsilon \nu(\cdot), \qquad \forall x \in \mathcal{C}.$$

A first condition is couched in terms of existence of (f, r)-modulated moments of the return-times to small sets[20], a condition called (f, r)*-regularity*. This condition is quite difficult to check in practice and Tuominen and Tweedie proposed powerful equivalent tools, based on drift inequalities on the form $PV \leq V - \chi + b\mathbf{1}_\mathcal{C}$ where $1 \leq \chi \leq V$ are measurable functions and b is a finite constant [20]. They show that this inequality yields the control of modulated moments, and conversely, that modulated moments are the minimal solutions (in a sense to precise, see [20, 12]). More recently, it was proved that the way in which χ and V are related determines the rate of convergence r and the f-norm in which convergence occurs : if χ is the constant function $\mathbf{1}$, then the chain is ergodic [12, Chater 14]; $\chi = \lambda V$ for some $\lambda < 1$ yields geometric rates [12, Chater 15] in V-norm and $\chi = \phi(V)$ for some concave function ϕ yields subgeometric ergodicity [3] at a rate r that has to be balanced with the function f (the larger the rate, the weaker the norm, and conversely).

Quantitative bounds of convergence for geometrically ergodic kernels were extensively addressed : in that case, it provides an explicit expression of the rate $\kappa > 1$ and makes explicit the dependence of the bounds in the initial value of the chain. Different approaches were proposed based on renewal theory, coupling theory, $\cdots$ (see [11, 16, 15, 6, 7, 4, 2]). Quantitative bounds of convergence for subgeometrically ergodic kernels were first addressed by Fort [6] (see also Fort and Moulines [8] for the polynomial case); two approaches were considered : the renewal one and the coupling one. New developments on subgeometric ergodicity [3] incited new interest in explicit control of ergodicity. [5] build explicit bounds by using the coupling approach. In the present work, we synthesize and (slightly) improve the work by Fort [6] by taking into account recent results [3]. Our approach is based on coupling. Comparison of the present contribution to that of [5] is part of this paper.

We provide an expression of a rate function $r = \{r(n) \ : n \geq 0\}$ and a bound $B_1(r, f; \lambda, \lambda')$ such that

$$\sum_n r(n) \int \lambda(dx)\lambda'(dx') \, \|P^n(x, \cdot) - P^n(x', \cdot)\|_f \leq B_1(r, f; \lambda, \lambda'), \tag{1.1}$$

where λ, λ' are probability measures on X. We also exhibit a rate function $r = \{r(n) \ : n \geq 0\}$, a function $B_2 : \mathsf{X} \times \mathsf{X} \to \mathbb{R}^+$ and a finite constant R such that

$$\sup_n \sup_{(x,x') \in \mathsf{X} \times \mathsf{X}} B_2(x, x')^{-1} \, r(n) \, \|P^n(x, \cdot) - P^n(x', \cdot)\|_f \leq R$$

and

$$\lim_n r(n) \quad \int \lambda(dx)\lambda'(dx') \, \|P^n(x,\cdot) - P^n(x',\cdot)\|_f = 0 \tag{1.2}$$

for any probability measures λ, λ' such that $\int \lambda(dx)\lambda'(dx')B_2(x,x') < \infty$. To that goal, we formulate sufficient conditions for subgeometric ergodicity couched in terms of a drift inequality on the form

$$PV(x) \leq V(x) - \phi \circ V(x) + b\mathbf{1}_{\mathcal{C}}(x), \tag{1.3}$$

where $\mathcal{C}$ is a m-small set and ϕ is a concave function. We express B_1, B_2, the rate r and the constant R in terms of the quantities appearing in the assumptions. These assumptions are more general than what is generally done in the literature since the drift condition is relative to P while the minorization condition is relative to P^m, $m \geq 1$. Our results improve and complete earlier works [6]. In [6], two techniques are considered : the first one relies on the renewal theory and the upper bounds $B_1(r,f;\lambda,\lambda')$ and $B_2(x,x')$ are expressed in terms of modulated moments of the return-time to $\mathcal{C}$, up to a multiplicative constant the construction of which is omitted [6, Chapter 3]. The second approach relies on the construction of a coupling time; quantitative bounds are derived under nested drift conditions that make the exposition quite obscure [6, Chapter 4]. The single drift condition (1.3) now allows an easier and more pedagogical construction of these bounds. For this second approach, a minorization condition of $\phi \circ V$ outside $\mathcal{C}$ was required [6, Condition H4] and [7, Eq. (50)]: this condition can be relaxed as observed by G.O. Roberts and J.S. Rosenthal and this is done in the present work.
The present work confirms earlier results ([20, Theorem 4.2], [3, Theorem 2.8]) : these contributions claim that for a ψ-irreducible aperiodic transition kernel that satisfies the drift inequality (1.3) outside a small set $\mathcal{C}$, the series (1.1) is finite and the limit (1.2) holds for any couple (r,f) where $r = \Psi_1(r_\phi)$ and $f = \Psi_2(\phi \circ V)$; (Ψ_1, Ψ_2) is a pair of inverse Young functions and the rate r_ϕ depends on ϕ. The present work covers this family of couple (r,f) and in that sense, confirms earlier work [6, 7]. It is more precise since it provides explicit control of this convergence ; our computable bounds improve and/or generalize earlier works on that topic [6, 7, 5].

The originality and the main over-value of the present computations is that we compute a bound by assuming *(a)* a drift condition verified for the kernel P and *(b)* a minorization condition for some m-iterated of the kernel. Usually (Cf. most of the references above for the geometric case and [5] for the subgeometric case), the drift and the minorization conditions are both with respect to the same m-iterated of the kernel, while extensions to the general case are left to the interested reader. The present form of the assumptions is directly related to what we are able to prove in practice (see the examples in [19, 18, 3] for example). The construction of the bounds relies on the coupling technique : the rate of convergence in f-norm is related to the existence of moments for the coupling time T of a bivariate process and to the $(f,\mathbf{1})$-modulated moments of T. We provide three kinds of bounds : we first express B_1 and B_2 in terms of modulated moments of the coupling time of a bivariate process (Section 4.2.2). We then express these controls in terms of the first-time the bivariate process enters some coupling set (Section 4.2.1). These moments are easily controlled by a drift condition with respect to the bivariate process ; we show how to obtain this "bivariate" drift condition from a "single" drift

condition on the kernel P. This yields the third level of explicit control (Section 3.1). While explicit, these controls are quite unappealing : for pedagogical purposes, we only list in Section 3 the terms on which the constants depend. The exact definition of the bounds in terms of these constants can be retrieved from Section 4.
In this contribution, the coupling time is related to the first time two independent copies of the Markov Chain enter the same set. More general coupling time can be built, as done in [4] to derive explicit bounds for geometric ergodicity. Nevertheless, the technique presently derived can be easily adapted to the coupling of any bivariate process ; the details are thus omitted. In addition, a drift condition for a (general) bivariate process is quite difficult to check in practice, except when this bivariate process is related to two independent copies (see the examples in [15, 4] for the geometric case) thus justifying that we restrict our attention to the case of independent copies.

The paper is organized as follows : we first set the conditions under which the bounds are derived. The main results are stated in Section 3 : polynomial ergodicity is first addressed for pedagogical purposes (Section 3.2), and general rates are then considered (Section 3.3) ; in Section 3.4, we compare our results to earlier works on that topic. The proofs are postponed in Section 4.

2 Hypotheses and definitions

Let $\{\Phi_n \ : n \geq 0\}$ be a X-valued Markov chain with transition kernel P on $(\mathsf{X}, \mathcal{B}(\mathsf{X}))$. We assume that X is a topological metric space and $\mathcal{B}(\mathsf{X})$ stands for the Borel σ-field.
Denote by $\mathbb{P}_\lambda$ (resp. $\mathbb{E}_\lambda$) the probability (resp. the expectation) on the canonical space for the chain with initial distribution λ; here $\mathbb{P}_x, \mathbb{E}_x$ stands for $\mathbb{P}_{\delta_x}, \mathbb{E}_{\delta_x}$, where δ_x is the point mass in x. Let $\mathcal{F} = \{\mathcal{F}_n \ : n \geq 0\}$ be the natural filtration of the process.

The bounds are derived assuming the following conditions are satisfied :

A1 some subgeometric drift condition outside a small set $\mathcal{C}$:

(i) There exist a measurable function $V : \mathsf{X} \to [1, +\infty)$, a concave monotone nondecreasing differentiable function $\phi : [1, +\infty) \to \mathbb{R}^+$ such that $\lim_{t \to \infty} \phi'(t) = 0$, a set $\mathcal{C} \in \mathcal{B}(\mathsf{X})$ and a non negative constant b such that

$$PV(x) \leq V(x) - \phi \circ V(x) + b \mathbf{1}_{\mathcal{C}}(x), \qquad \forall x \in \mathsf{X}, \tag{2.1}$$

and $\sup_{t \geq 1} \phi(t) > b$.

(ii) $\mathcal{C}$ is a ι-small set with minorizing constant ϵ_ι and minorizing probability distribution ν.

A2 some boundedness conditions : $\sup_{\mathcal{C}} V < \infty$ and $\nu(V) < \infty$.

A3 the transition kernel P is aperiodic.

The drift condition (2.1) introduced by [3] is related to f-ergodicity at a subgeometric rate, for a large range of functions $1 \leq f \leq \phi \circ V$ [3, Theorem 2.8.]. It generalizes the Jarner and Roberts drift condition where $\phi(v) \propto cv^{1-\alpha}$ for some $0 < \alpha \leq 1$ that guarantees f-ergodicity at a polynomial rate ([9]; see also [8]). As commented in [3], the differentiability of ϕ is assumed for convenience but it can be relaxed since a concave function has non-increasing left and right derivatives everywhere. $t \mapsto \phi'(t)$ is non-increasing and converges to c ; the case $c > 0$ corresponds to ergodicity at a geometric rate, thus explaining our condition $\lim_{t \to \infty} \phi'(t) = 0$ [3, Remark 1].

Since ϕ is concave, we have for all $1 \leq a < b$,

$$\phi(a + b) - \phi(a) - \phi(b) \leq \frac{ab}{b - a} \left(\frac{\phi(b)}{b} - \frac{\phi(a)}{a} \right) \leq 0. \tag{2.2}$$

This results from the inequality $\phi(a + b) \leq \left[\phi(a) + (t - a) \frac{\phi(b) - \phi(a)}{b - a} \right]_{t = a + b}$. In addition, since $\lim_t \phi'_t \downarrow 0$, the function $t \mapsto \phi(t)/t$ is decreasing and tends to zero as t tends to infinity. Therefore, there exist positive real numbers v_1, v_2 satisfying the following conditions :

- $v_1 \geq \sup_{\mathcal{C}} V$ and $\phi(v_1) - b > 0$ where b is given by A1(i).
- $\frac{v_1 v_2}{v_2 - v_1}\left(\frac{\phi(v_2)}{v_2} - \frac{\phi(v_1)}{v_1}\right) + b < 0$.

Define $\mathcal{D} = \{x \in \mathsf{X} \ : V(x) \leq v_2\}$, a set which contains $\mathcal{C}$. As proved in the following paragraph, under the stated assumptions, $\mathcal{D}$ is an accessible m-small set with minorizing probability distribution ν (the same measure ν as in A 1(ii))

$$\exists\, m \geq 1, \epsilon > 0, \qquad \forall x \in \mathcal{D}, \qquad\qquad P^m(x, \cdot) \geq \epsilon\, \nu(\cdot).$$

We will give an explicit control of the f-ergodicity in terms of ϕ, b, m, ϵ, $\nu(V)$ and $\sup_{\mathcal{D}} V = v_2$. Before going further, we comment the consequences of the assumptions A1 to A3 and give a sufficient condition for establishing A3 when A1-A2 holds.

We first prove that A1-A2 imply ψ-irreducibility and positive Harris-recurrence of the transition kernel. We then show that combined with a minorization condition of $\phi \circ V$ on $\mathcal{C}$, A1-A2 imply A3. We finally establish the smallness property of the level sets $\{V \leq v\}$ of the drift function V.

Proposition 2.1. *Under A1-2, $\mathcal{C}$ and $\mathcal{D}$ are accessible, P is ψ-irreducible, positive Harris-recurrent and possesses an invariant measure π such that $\pi(\phi \circ V) < \infty$.*

Proof. Define the return-time to $\mathcal{C}$, $\tau_{\mathcal{C}}$ by

$$\tau_{\mathcal{C}} = \inf\{n \geq 1 \ : \Phi_n \in \mathcal{C}\}.$$

From [12, Proposition 11.3.4], $\mathbb{P}_x(\tau_{\mathcal{C}} < \infty) = 1$ for all x in the set $\{V < \infty\}$. Since $\{V < \infty\} = \mathsf{X}$, $\mathcal{C}$ is accessible. Since $\mathcal{C} \subset \mathcal{D}$, $\mathcal{D}$ is accessible. Let $\mathcal{A} \in \mathcal{B}(\mathsf{X})$ be such that $\nu(\mathcal{A}) > 0$;

$$\begin{aligned}\mathbb{P}_x(\tau_{\mathcal{A}} < \infty) &= \mathbb{P}_x(\tau_{\mathcal{A}} \circ \theta^{\tau_{\mathcal{C}}} < \infty, \tau_{\mathcal{C}} < \infty) \\ &\geq \inf_{x \in \mathcal{C}} \mathbb{P}_x(\tau_{\mathcal{A}} < \infty)\mathbb{P}_x(\tau_{\mathcal{C}} < \infty) \geq \epsilon_\iota \nu(\mathcal{A})\mathbb{P}_x(\tau_{\mathcal{C}} < \infty) > 0,\end{aligned}$$

thus establishing that ν is an irreducibility measure for the kernel. Since $\sup_{\mathcal{C}} V < \infty$ and $\mathcal{C}$ is an accessible small set, P is positive Harris-recurrent [12, Theorem 11.3.4.], and $\pi(\phi \circ V) < \infty$ [12, Theorem 14.3.7]. $\square$

Aperiodicity of the transition kernel is required for a positive Harris-recurrent kernel to be ergodic. To prove A3, one can show that the chain is strongly aperiodic i.e. the chain possesses a ν_1-small set $\mathcal{A}$ such that $\nu_1(\mathcal{A}) > 0$. We also have the following sufficient condition.

Lemma 2.2. *Assume A1-2. If there exists $0 < a \leq 1$ such that $\phi \circ V \geq (1-a)^{-1} b$ on $\mathcal{C}^c$, the transition kernel is aperiodic.*

Proof. From A1(i) and the definition of a, we have

$$PV(x) + PV(x') \leq V(x) + V(x') - a\left(\phi \circ V(x) + \phi \circ V(x')\right) + 2b\mathbf{1}_{\mathcal{C}\times\mathcal{C}}(x, x'), \quad (2.3)$$
$$\leq V(x) + V(x') - 2a\phi(1) + 2b\mathbf{1}_{\mathcal{C}\times\mathcal{C}}(x, x'); \quad (2.4)$$

the proof is along the same lines as the proof of [8, Proposition 7] and is omitted. Since $\{V < +\infty\} = \mathsf{X}$, we deduce from (2.4) and [12, Proposition 11.3.4] that the time T_0 when a chain with transition kernel $P \otimes P$ (i.e. two independent copies of a chain with kernel P) started at $(x, x') \in \mathsf{X} \times \mathsf{X}$ enters $\mathcal{C} \times \mathcal{C}$ is finite almost-surely. The proof is by contradiction : assume that the kernel P is periodic of period $d > 1$ and let $D_1, \cdots, D_d$ be the d-cycle [12, Theorem 5.4.4]. ν is non-trivial and there exists an element, say D_d such that $\nu(D_d) > 0$: the smallness property and the definition of a d-cycle both imply that up to a ψ-null set, $\mathcal{C} \subset D_{d-\iota \bmod d}$. Hence, the bivariate process started from $(x, x') \in D_1 \times D_d$ never reach $\mathcal{C} \times \mathcal{C}$, which contradicts the property that T_0 is finite almost-surely, whatever the initial values of the bivariate process. □

In many applications, one is able to prove the drift condition A1(i) outside a level set $\mathcal{C} = \{V \leq c\}$, with a drift function V such that its level sets are small. Hence, the drift inequality still holds by substituting $\mathcal{C}$ for the level set $\{V \leq c + v\}$ whatever $v \geq 0$ (V, ϕ and b are unchanged). In that case, Lemma 2.2 implies the condition A3.
Under A1-2, aperiodicity implies that the level sets of V are petite, so that V is said unbounded off petite set. $\mathcal{C} \in \mathcal{B}(\mathsf{X})$ is said $\nu_{\mathbf{b}}$-petite if there exist a distribution $\mathbf{b} = \{\mathbf{b}(n) \ : n \geq 0\}$ on $\mathbb{N}$, a probability measure $\nu_{\mathbf{b}}$ on $\mathcal{B}(\mathsf{X})$ and a constant $\epsilon > 0$ such that

$$\sum_n \mathbf{b}(n)\ P^n(x, \cdot) \geq \epsilon \nu_{\mathbf{b}}(\cdot), \qquad \forall x \in \mathcal{C}. \quad (2.5)$$

Proposition 2.3. *Assume A1-A2. Then V is unbounded off petite set. If in addition A3 holds, the level set $\{V \leq v\}$ is small with minorizing probability measure ν.*

Proof. Consider the level set $\{x \in \mathsf{X} \ : V(x) \leq v\}$ for some $v \geq 1$. For all N large enough, it holds

$$\sum_n \mathbf{a}_N * \delta_\iota(n)\ P^n(x, \cdot) \geq \frac{\epsilon}{N}\ (1 - \frac{v}{N})\ \nu(\cdot), \qquad \forall x \in \{V \leq v\}, \quad (2.6)$$

where $\mathbf{a} * \mathbf{b}$ denotes the convolution of the two sequences $\mathbf{a}$, $\mathbf{b}$, and $\mathbf{a}_N$ is the uniform distribution on $\{1, \cdots, N\}$. The proof is similar to that of [12, Lemma 11.3.7] and is omitted.
The smallness property of the level set and the expression of the minorizing probability measure result from [12, Theorem 5.5.5 and Theorem 5.5.7]. □

3 Main results

The construction of the explicit upper bounds relies on coupling theory : we need to control modulated moments of the coupling time of a bivariate process where each component is a copy

- in a sense to precise - of the original chain. We claim the main results in this section and postpone the proofs to Section 4. The definition of the bounds thus requires some notations related to the behavior of two independent copies of a Markov chain with transition kernel P (Section 3.1).
For pedagogical purposes, we start with the polynomial case (Section 3.2) and then provide explicit controls for general subgeometric rate (Section 3.3).

3.1 Notations

Our bounds are related to the (f,r)-modulated moments of the time T_0 when two independent copies of a Markov chain with transition kernel P enters $\mathcal{D} \times \mathcal{D}$. Proposition 3.1 below is crucial since it makes explicit the dependence of these moments on the constants appearing in the assumptions.

Let $\bar{\mathbb{E}}_{x,x'}$ be the expectation on the canonical space associated to two independent copies $\{(\Phi_n, \Phi'_n) \ : n \geq 0\}$ of a Markov chain with transition kernel P and initial distribution (x, x'). Define

$$\Delta = \mathcal{D} \times \mathcal{D} \qquad \text{and} \qquad T_0 = \inf\{n \geq 0 \ : (\Phi_n, \Phi'_n) \in \Delta\};$$

set

$$r_\phi(n) = \phi \circ H^{-1}(n) = (H^{-1})'(n), \ n \geq 0, \qquad \text{where} \qquad H(t) = \int_1^t \frac{ds}{\phi(s)}.$$

The assumptions on ϕ (see A1) imply that r_ϕ belongs to the class of the subgeometric rate functions ([3, Lemma 2.3]).

Proposition 3.1. *Assume A1(i) and A2. Then*

$$PV(x) + PV(x') \leq V(x) + V(x') - \phi\left(V(x) + V(x')\right) + 2b\mathbf{1}_{\mathcal{D}\times\mathcal{D}}(x, x'). \tag{3.1}$$

Therefore, for all $(x, x') \notin \Delta$,

$$\bar{\mathbb{E}}_{x,x'}\left[\sum_{k=0}^{T_0-1} r_\phi(k)\right] \leq V(x) + V(x').$$

$$\bar{\mathbb{E}}_{x,x'}\left[\sum_{k=0}^{T_0-1} \phi\{V(\Phi_k) + V(\Phi'_k)\}\right] \leq V(x) + V(x').$$

Proof. The drift inequality yields

$$PV(x) + PV(x') \leq V(x) + V(x') - \phi\left(V(x) + V(x')\right) + 2b\mathbf{1}_{\mathcal{D}\times\mathcal{D}}(x, x') + \Xi(x, x')$$

where

$$\Xi(x, x') = \phi\left(V(x) + V(x')\right) - \phi \circ V(x) - \phi \circ V(x') + b\mathbf{1}_{\mathcal{C}\times\mathcal{D}^c}(x, x') + b\mathbf{1}_{\mathcal{D}^c\times\mathcal{C}}(x, x').$$

By (2.2), for $(x,x') \notin \mathcal{D}\times\mathcal{D}$, $\Xi(x,x') \leq 0$ which proves the lemma for $(x,x') \in [\mathcal{D}\times\mathcal{D}]^c$. Let $(x,x') \in \mathcal{C}\times\mathcal{D}^c$: $\Xi(x,x') \leq \frac{v_1 v_2}{v_2 - v_1}\left(\frac{\phi(v_2)}{v_2} - \frac{\phi(v_1)}{v_1}\right) + b$ which is negative by definition of v_1 and v_2. This concludes the proof of (3.1). The control of the modulated moments results from (3.1), [3, Proposition 2.1] and [12, Proposition 11.3.2]. $\square$

Define the residual kernel

$$R_m(x,dy) = (1-\epsilon)^{-1}\left(P^m(x,dy) - \epsilon\nu_m(dy)\mathbf{1}_{\mathcal{D}}(x)\right), \tag{3.2}$$

which is, under the smallness property of $\mathcal{D}$, a transition kernel on $(\mathsf{X},\mathcal{B}(\mathsf{X}))$.

3.2 Explicit bounds for polynomial ergodicity

In this section, we assume that $\phi(v) = cv^{1-\alpha}$ for some $c>0$ and $0<\alpha\leq 1$. Hence, (2.1) gets

$$PV(x) \leq V(x) - cV(x)^{1-\alpha} + b\mathbf{1}_C(x), \qquad \forall x \in \mathsf{X}, \tag{3.3}$$

This drift condition has been first proposed by Jarner and Roberts [9]; the authors proved that any ψ-irreducible and aperiodic transition kernel satisfying (3.3) with respect to a small set $\mathcal{C}$, possesses an invariant probability measure π such that $\pi(V^{1-\alpha}) < \infty$ and for all $1 \leq \kappa \leq 1/\alpha$, $x \in \mathsf{X}$,

$$\lim_n \,(n+1)^{\kappa-1}\,\|P^n(x,\cdot) - \pi(\cdot)\|_{V^{1-\kappa\alpha}} = 0. \tag{3.4}$$

They also established that this allows the control of modulated moments of the return-time to the set $\mathcal{C}$ ([9, Proof of Theorem 3.6]) so that, by applying the results by Tuominen and Tweedie [20, Theorem 4.2], it holds

$$\sum_{n\geq 0}(n+1)^{\kappa-1}\int \lambda(dx)\lambda'(dx')\,\|P^n(x,\cdot) - P^n(x',\cdot)\|_{V^{1-\kappa\alpha}} < \infty, \tag{3.5}$$

for any probability measures (λ,λ') such that

$$\mathbb{E}_\lambda\left[\sum_{k=0}^{\tau_{\mathcal{C}}-1}(k+1)^{\kappa-1}V^{1-\kappa\alpha}(\Phi_k)\right] + \mathbb{E}_{\lambda'}\left[\sum_{k=0}^{\tau_{\mathcal{C}}-1}(k+1)^{\kappa-1}V^{1-\kappa\alpha}(\Phi_k)\right] < \infty.$$

We are able to make these results more precise : in (3.4), we explicit the dependence upon the initial point x and in (3.5), we give the dependence of an upper bound in terms of the constants and functions appearing in the assumptions A1-A2.

Theorem 3.2. *Assume A1-A3 with $\phi(v) = cv^{1-\alpha}$ for some $c>0$ and $0<\alpha\leq 1$. For all $1 \leq \kappa \leq 1/\alpha$, there exists a finite constant R_κ depends upon $\alpha, c, b, \epsilon, m, \nu(V)$, $\sup_{\mathcal{D}} V$ and κ*

such that for all $(x, x') \in \mathsf{X} \times \mathsf{X}$

$$\sum_{k \geq 0} (k+1)^{\kappa-1} \, \|P^k(x, \cdot) - P^k(x', \cdot)\|_{V^{1-\kappa\alpha}} \leq R_\kappa + \bar{\mathbb{E}}_{x,x'} \left[\sum_{k=0}^{T_0-1} \phi\{V(\Phi_k) + V(\Phi'_k)\} \right] + \bar{\mathbb{E}}_{x,x'} \left[\sum_{k=0}^{T_0-1} (k+1)^{\alpha^{-1}-1} \right].$$

An exact expression of the constant R_κ can be tracked from the proof postponed in Section 4. By Proposition 3.1, we thus have the following corollary ;

Corollary 3.3. *Assume A1-A3 with $\phi(v) = cv^{1-\alpha}$ for some $c > 0$ and $0 < \alpha \leq 1$. Let $1 \leq \kappa \leq 1/\alpha$. There exists a finite constant R_κ depending upon $\alpha, c, b, \epsilon, m, \nu(V), \sup_{\mathcal{D}} V$ and κ such that*

$$\sup_{(x,x') \in \mathsf{X} \times \mathsf{X}} \left(V(x) + V(x')\right)^{-1} \sum_{k \geq 0} (k+1)^{\kappa-1} \, \|P^k(x, \cdot) - P^k(x', \cdot)\|_{V^{1-\kappa\alpha}} \leq R_\kappa.$$

Theorem 3.2 provides optimal controls : for some $V^{1-\kappa\alpha}$-norm, it exhibits the largest rate such that the series can be controlled by $(V(x)+V(x'))$. Nevertheless, the dependence upon the initial values weakens when considering lower rates. Jarner and Roberts established that if A1 holds with $\phi \circ V \propto V^{1-\alpha}$, then a continuum of conditions on the form (2.1) are verified. More precisely, for all $0 \leq \eta < 1$,

$$PV^\eta(x) \leq V^\eta(x) - \eta c V^{\eta-\alpha}(x) + b^\eta \mathbf{1}_{\mathcal{C}}(x),$$

(see [9, Lemma 3.5]). The next proposition can be proved along the same lines as Theorem 3.2, by replacing V with V^η, c with $c\eta$ and b by b^η for $\eta = 1 - (\kappa - 1)\alpha$.

Proposition 3.4. *Assume A1-A3 with $\phi(v) = cv^{1-\alpha}$ for some $0 < \alpha \leq 1$ and some positive constant c. For all $1 \leq \kappa \leq \alpha^{-1}$ and all $0 \leq l \leq (\kappa - 1)$, there exists a finite constant $R_{\kappa,l}$ depending only on $\alpha, c, b, \epsilon, m, \nu(V), \sup_{\mathcal{D}} V$ and κ, l such that*

$$\sup_{(x,x') \in \mathsf{X} \times \mathsf{X}} \left(V^{1-l\alpha}(x) + V^{1-l\alpha}(x')\right)^{-1} \sum_{k \geq 0} (k+1)^{\kappa-1-l} \, \|P^k(x, \cdot) - P^k(x', \cdot)\|_{V^{1-\kappa\alpha}} \leq R_{\kappa,l}.$$

We conclude this section by considering controls of the limit (3.4).

Theorem 3.5. *Assume A1-A3 with $\phi(v) = cv^{1-\alpha}$ for some $0 < \alpha \leq 1$ and some positive constant c. For all $1 \leq \kappa \leq \alpha^{-1}$ and all $1 \leq l \leq \kappa$,*

$$\lim_{k \to \infty} (k+1)^{\kappa-l} \int \lambda(dx)\lambda'(dx') \, \|P^k(x, \cdot) - P^k(x', \cdot)\|_{V^{1-\kappa\alpha}} = 0,$$

for any probability measures (λ, λ') on X such that $\lambda(V^{1-l\alpha}) + \lambda'(V^{1-l\alpha}) < \infty$. Furthermore, there exists a finite constant $R_{\kappa,l}$ depending only on $\alpha, c, b, \epsilon, m, \nu(V), \sup_{\mathcal{D}} V$ and κ, l such that

$$\sup_{k \geq 0} \sup_{(x,x') \in \mathsf{X} \times \mathsf{X}} \left(V^{1-l\alpha}(x) + V^{1-l\alpha}(x')\right)^{-1} (k+1)^{\kappa-l} \, \|P^k(x, \cdot) - P^k(x', \cdot)\|_{V^{1-\kappa\alpha}} \leq R_{\kappa,l}.$$

Under the stated assumptions $\pi(\phi \circ V) < \infty$ (Proposition 2.1) and Theorem 3.5 claims that for all λ such that $\lambda(\phi \circ V) < \infty$, $\lim_k (k+1)^{\kappa-1} \|\lambda P^k - \pi\|_{V^{1-\kappa\alpha}} = 0$, thus confirming the known result (3.4).

3.3 Explicit bounds for subgeometric ergodicity

Let $\mathcal{I}$ be the set of pairs of inverse Young functions, augmented with the pairs $(\mathrm{Id}, \mathbf{1})$ and $(\mathbf{1}, \mathrm{Id})$ where Id stands for the identity function and $\mathbf{1}$ denotes the constant function equal to one. For all $(\Psi_1, \Psi_2) \in \mathcal{I}$, and all (x, y),

$$\Psi_1(x)\Psi_2(x) \leq x + y, \tag{3.6}$$

(see [10, Chapter 1]). This inequality will prove crucial in following development : the idea is that if one is able to control moments of the form $\mathbb{E}_x\left[\sum_{k=0}^{\tau} r_\phi(k)\right]$ and $\mathbb{E}_x\left[\sum_{k=0}^{\tau} \phi \circ V(\Phi_k)\right]$ then we control modulated moments

$$\mathbb{E}_x\left[\sum_{k=0}^{\tau} \Psi_1(r_\phi(k))\ \Psi_2(\phi \circ V)(\Phi_k)\right],$$

where $\Psi_1 \circ r_\phi$ is a subgeometric rate function provided $\lim_t \phi'(t) = 0$ ([3, Lemma 2.3]). [3] proved that any ψ-irreducible and aperiodic transition kernel satisfying (2.1) with respect to a small set $\mathcal{C}$, possesses an invariant probability measure π such that $\pi(\phi(V)) < \infty$ and for any pair $(\Psi_1, \Psi_2) \in \mathcal{I}$

$$\lim_n\ \Psi_1(r_\phi(n))\ \|P^n(x, \cdot) - \pi(\cdot)\|_{\Psi_2(\phi \circ V)} = 0. \tag{3.7}$$

They also established

$$\sum_{n \geq 0} \Psi_1(r_\phi(n)) \int \lambda(dx)\lambda'(dx')\ \|P^n(x, \cdot) - P^n(x', \cdot)\|_{\Psi_2(\phi \circ V)} < \infty, \tag{3.8}$$

for any probability measures (λ, λ') such that $\lambda(V) + \lambda'(V) < \infty$ [3, Theorem 2.8]. Here again, we are able to make these results more precise.

Theorem 3.6. *Assume A1-A3. There exists a finite constant R depending upon $\phi, b, \epsilon, m, \nu(V)$, and $\sup_{\mathcal{D}} V$ such that for all $(\Psi_1, \Psi_2) \in \mathcal{I}$ and all $(x, x') \in \mathsf{X} \times \mathsf{X}$,*

$$\sum_{k \geq 0} \Psi_1(r_\phi(k))\ \|P^k(x, \cdot) - P^k(x', \cdot)\|_{\Psi_2(\phi \circ V)}$$
$$\leq R + \bar{\mathbb{E}}_{x,x'}\left[\sum_{k=0}^{T_0 - 1} \phi\{V(\Phi_k) + V(\Phi'_k)\}\right] + \bar{\mathbb{E}}_{x,x'}\left[\sum_{k=0}^{T_0-1} r_\phi(k)\right].$$

An explicit expression of R can be tracked from the proof of the theorem postponed in Section 4. Theorem 3.6 corroborates [20, Theorem 4.2] and [3, Proposition 2.5 and Theorem

2.8] and gives an explicit control of the convergence. By Proposition 3.1, there exists a constant R depending upon $\phi, b, \epsilon, m, \nu(V)$ and $\sup_{\mathcal{D}} V$ such that

$$\sum_{k\geq 0} \Psi_1\left(r_\phi(k)\right) \; \|P^k(x,\cdot) - P^k(x',\cdot)\|_{\Psi_2(\phi\circ V)} \leq R(V(x) + V(x')).$$

As in the polynomial case, a simple drift inequality provides an infinity of drift conditions : for a concave and differentiable function $\tilde{\phi} \; : [1,\infty) \to \mathbb{R}^+$, the Jensen inequality implies

$$P(\tilde{\phi}\circ V)(x) \leq \tilde{\phi}\circ V(x) - \tilde{\phi}'\circ V(x) \;\; \phi\circ V(x) + \tilde{\phi}(b)\mathbf{1}_{\mathcal{C}}(x).$$

This inequality can in turn be plugged in the above analysis, thus providing new explicit controls with weaker dependence in the initial conditions. We do not pursue further the details of the computations since they can be derived following the same steps as in the polynomial case. We conclude this section by the analogous of Theorem 3.5 for general subgeometric rate.

Theorem 3.7. *Assume A1-A3. Then for all* $(\Psi_1, \Psi_2) \in \mathcal{I}$

$$\lim_{k\to\infty} \Psi_1\left(r_\phi(k)\right) \; \|\lambda P^k - \lambda' P^k\|_{\Psi_2(\phi\circ V)} = 0$$

for any probability measures (λ, λ') *on* X *such that* $\int \lambda(dx)\lambda'(dx')\phi\left(V(x) + V(x')\right) < \infty$*. Furthermore, there exists a finite constant* R *depending only on* $\phi, b, \epsilon, m, \nu(V), \sup_{\mathcal{D}} V$ *such that*

$$\sup_{k\geq 0} \sup_{(x,x')\in\mathsf{X}\times\mathsf{X}} \left\{\phi\left(V(x)+V(x')\right)\right\}^{-1} \; \Psi_1\left(r_\phi(k)\right) \; \|P^k(x,\cdot) - P^k(x',\cdot)\|_{\Psi_2(\phi\circ V)} \leq R.$$

3.4 Earlier works on computational bounds for subgeometric ergodicity

Under the stated assumptions, [3, Proposition 2.6] implies that the chain is $(\Psi_1(r_\phi), \Psi_2(\phi\circ V))$-regular and any probability measure λ such that $\lambda(V) < \infty$ is $(\Psi_1(r_\phi), \Psi_2(\phi\circ V))$-regular ([20]).

- *Tuominen and Tweedie (1994):* In [20], it is proved that for all x in a set that contains $\{x\in\mathsf{X}, V(x) < \infty\}$, $\lim_k \Psi_1\left(r_\phi(k)\right) \; \|P^k(x,\cdot) - \pi\|_{\Psi_2(\phi\circ V)} = 0$ and

$$\sum_k \Psi_1\left(r_\phi(k)\right) \; \|\lambda P^k - \lambda' P^k\|_{\Psi_2(\phi\circ V)} < \infty \tag{3.9}$$

for any probability measures λ, λ' such that $\lambda(V) + \lambda'(V) < \infty$. Our theorems 3.6 and 3.7 imply these results.

- *Fort (2001); Fort and Moulines (2003):* In [6, Chapter 3], the dependence in λ, λ' in the right hand side of (3.9) is given ; it is equal to

$$\mathbb{E}_\lambda\left[\sum_{k=0}^{\tau_{\mathcal{C}}-1} \Psi_1\left(r_\phi(k)\right)\Psi_2(\phi\circ V)(\Phi_k)\right] + \mathbb{E}_{\lambda'}\left[\sum_{k=0}^{\tau_{\mathcal{C}}-1} \Psi_1\left(r_\phi(k)\right)\Psi_2(\phi\circ V)(\Phi_k)\right]$$

which is in turn upper-bounded by

$$\mathbb{E}_\lambda \left[\sum_{k=0}^{\tau_{\mathcal{C}}-1} r_\phi(k) \right] + \mathbb{E}_{\lambda'} \left[\sum_{k=0}^{\tau_{\mathcal{C}}-1} r_\phi(k) \right] + \mathbb{E}_\lambda \left[\sum_{k=0}^{\tau_{\mathcal{C}}-1} (\phi \circ V)(\Phi_k) \right] + \mathbb{E}_{\lambda'} \left[\sum_{k=0}^{\tau_{\mathcal{C}}-1} (\phi \circ V)(\Phi_k) \right],$$

(due to the Young's inequality). The Comparison Theorem [12] and [3, Proposition 2.1] establish that these moments are upper bounded by $V(x) + V(x')$. Hence Theorem 3.6 yields a larger dependence in the initial values than the one in [6, Theorems 3.1 and 3.2]. We think that these bounds are equivalent *i.e.* we postulate that V can be assumed to be equal to $\mathbb{E}_x \left[\sum_{k=0}^{\tau_{\mathcal{C}}-1} \Psi_1(r_\phi(k)) \, \Psi_2(\phi \circ V)(\Phi_k) \right]$ on $\mathcal{C}^c$. This could be the case if one is able to prove that, given $\mathcal{C}$ and ϕ, the minimal pointwise solution to the drift inequality (2.1) is the function U given by $U(x) = \mathbb{E}_x \left[\sum_{k=0}^{\tau_{\mathcal{C}}-1} \Psi_1(r_\phi(k)) \, \Psi_2(\phi \circ U)(\Phi_k) \right]$ on $\mathcal{C}^c$ and $U = 1$ on $\mathcal{C}$. Unfortunately, this is, to our best knowledge, an open question.
In [6, Chapter 4] and in [8], explicit bounds for subgeometric ergodicity are derived (with a special emphasis on the polynomial case). These bounds are derived under a drift condition relative to the transition kernel P, a minorization condition relative to the m-iterated P^m, and a condition on the behavior of the functions appearing in the drift inequality, outside some small set. This last condition makes the set of conditions by [6, 8] more restrictive than what we assume in the present contribution. The results we obtain here slightly improve the results by [6] in terms of the f-norm and the rate of convergence we are able to control (see for example the polynomial case above and the results detailed in [6, Section 4.3.7]). Finally, the technique used in this contribution is far more explicit and simple: hence, explicit bounds can be retrieved from the computations detailed in Section 4, in a far easier way that what is proposed in [6, 8].

- *Douc et al. (2003)* Their work generalizes the coupling construction as done in [4]. In that sense, their work is more general. Their assumptions are more restrictive than what we assume in this paper, since the minorization condition and the drift condition are both relative to the 1-iterated of the kernel, thus limiting the range of applications. Finally, the results given by Theorems 3.6 and 3.7 are better than the results given by [5, Theorem 4.4., Theorem 4.5]: for a given rate $\Psi_1(\rho r_\phi(k))$, where $0 \leq \rho \leq 1$, they are able to control the series and the limits in f-norm where $f \sim \Psi_2((1-\rho)(\phi \circ V))$. We can compare the results as follows : for a given control in $V(x) + V(x')$ for the series, and in $\lambda(V) + \lambda'(V)$ for the limit, they provide a control of convergence in total variation norm and in f-norm at a rate $\Psi_1(r_\phi(k))$ for any function $f \sim \tilde{\Psi}_2((1-\rho)\phi \circ V)$ where $\tilde{\Psi}_1(\rho r_\phi(k)) \sim \Psi_1(r_\phi(k))$; since $\rho < 1$ and $(1-\rho) < 1$, $\tilde{\Psi}_2((1-\rho)\phi \circ V) << \Psi_2(\phi \circ V)$. We provide a control of convergence in $\Psi_2(\phi \circ V)$-norm at a rate $\Psi_1(r_\phi(k))$.

4 Proofs

4.1 The coupling technique

The coupling technique is a powerful tool to derive quantitative bounds for ergodicity. The existence of a small set $\mathcal{D}$ allows the construction of a $(\mathsf{X}\times\mathsf{X}\times\{0,1\})$-valued Markovian process $\{(\Phi_n,\Phi'_n,d_n)\ : n\geq 0\}$ on some canonical probability space endowed with the probability $\check{\mathbb{P}}_{x,x',0}$. Roughly speaking, the process is built as follows : run two independent copies (Φ_n,Φ'_n) of the Markov chain, till the first time T_0 they are in Δ and set the bell variable d_n equal to 0. When the bivariate process is in Δ,

- with probability ϵ, set $d_{n+1}=1$ and draw $\Phi_{n+1}=\Phi'_{n+1}\sim\nu(\cdot)$; for the future, set $d_n=1$ and force the coupling of the bivariate process by setting $\Phi_{n+1}=\Phi'_{n+1}\sim P(\Phi_n;\cdot)$;
- with probability $(1-\epsilon)$, set $d_{n+1}=0$ and draw independently $\Phi_{n+1}\sim R_m(\Phi_n,\cdot)$ and $\Phi'_{n+1}\sim R_m(\Phi'_n,\cdot)$; for the future, repeat the mechanism above : run two independent Markov chains till they hit Δ and couple or not the bivariate process with probability ϵ.

When $m=1$, this process satisfies the key property

$$\check{\mathbb{E}}_{x,x',0}\left[f(\Phi_n)\right]=P^nf(x),\qquad \check{\mathbb{E}}_{x,x',0}\left[f(\Phi'_n)\right]=P^nf(x') \tag{4.1}$$

for all $n\geq 0$, $(x,x')\in\mathsf{X}$, $f\geq 0$; and $\Phi_n=\Phi'_n$ on the set $\{d_n=1\}$. When $m>1$, this property no longer holds due to the gap of size m inserted each time the bivariate process enters Δ. The coupling construction has to be adapted and a companion process has to be introduced as described in [8, page 86] (see also [7]). The resulting process $\{(\Phi_n,\Phi'_n,d_n)\ : n\geq 0\}$ with probability (resp. expectation) denoted by $\tilde{\mathbb{P}}_{x,x',0}$ (resp. $\tilde{\mathbb{E}}_{x,x',0}$) on the canonical space is no longer a Markov process but the key property (4.1) holds. By construction, the coupling time which is materialized by the first time the bell variable is set to 1, is equal to T_j+m for some (random) j, where T_j corresponds to an hitting time on Δ; more precisely, define the stopping-times $\{T_j\ : j\geq 0\}$ - with respect to the natural filtration $\tilde{\mathcal{F}}=\{\tilde{\mathcal{F}}_n\ : n\geq 0\}$ of the companion process as follows :

$$T_0=\inf\{n\geq 0,(\Phi_n,\Phi'_n)\in\Delta\},\qquad T_j=T_0\circ\theta^{T_{j-1}+m}+T_{j-1}, j\geq 1, \tag{4.2}$$

where θ is the shift operator. Then the coupling time T is defined by $T=\inf\{n\geq 0, d_n=1\}$ and by construction of the bell variable, we have,

$$\tilde{\mathbb{E}}_{x,x',0}\left[\{f(\Phi_n)-f(\Phi'_n)\}\mathbf{1}_{T\leq n}\right]=0. \tag{4.3}$$

Before embarking on the proofs, we list some properties of the process on which the proofs below are based : $\tilde{\mathbb{P}}_{x,x',0}$-a.s.,

$$\{d_q=0\}=\{d_l=0,l\leq q\},\qquad\qquad \{d_q=1\}=\{d_l=1,l\geq q\}; \tag{4.4}$$

and, on the set $\{d_{T_j}=0\}$, $j\geq 0$

$$\tilde{\mathbb{P}}_{x,x',0}\left(T=T_j+m|\tilde{\mathcal{F}}_{T_j}\right)=\tilde{\mathbb{P}}_{x,x',0}\left(d_{T_j+m}=1|\tilde{\mathcal{F}}_{T_j}\right)=\epsilon, \tag{4.5}$$

$$\tilde{\mathbb{E}}_{x,x',0}\left[f(\Phi_{T_j+k},\Phi'_{T_j+k})|\tilde{\mathcal{F}}_{T_j}\right]=\int P^k(\Phi_{T_j},dy)P^k(\Phi'_{T_j},dy')f(y,y'),\qquad \forall 1\leq k\leq m-1, \tag{4.6}$$

$$\tilde{\mathbb{E}}_{x,x',0}\left[f(\Phi_{T_j+m},\Phi'_{T_j+m})\mathbf{1}_{d_{T_j+m}=0}|\tilde{\mathcal{F}}_{T_j}\right]=(1-\epsilon)\int R_m(\Phi_{T_j},dy)R_m(\Phi'_{T_j},dy')\ f(y,y'), \tag{4.7}$$

$$\tilde{\mathbb{E}}_{x,x',0}\left[f\circ\theta^{T_j+m}|\tilde{\mathcal{F}}_{T_j+m}\right]=\tilde{\mathbb{E}}_{\Phi_{T_j+m},\Phi'_{T_j+m},0}\left[f\right]. \tag{4.8}$$

Finally, from time 1 to T_0 and from time T_j+m+1 to T_{j+1}, the uncoupled bivariate process is drawn as two independent copies of P; hence, for all measurable functions $f_k\geq 0$, we have on the set $\{d_{T_j+m}=0\}$, $j\geq -1$,

$$\tilde{\mathbb{E}}_{x,x',0}\left[\prod_{k=1}^{T_{j+1}}f_k(\Phi_{k+T_j+m},\Phi'_{k+T_j+m})|\tilde{\mathcal{F}}_{T_j+m}\right]=\bar{\mathbb{E}}_{\Phi_{T_j+m},\Phi'_{T_j+m}}\left[\prod_{k=1}^{T_0}f_k(\Phi_k,\Phi'_k)\right], \tag{4.9}$$

where by convention $T_{-1}=-m$.
We conclude this short exposition by showing that under the stated assumptions, the random times $\{T_j\ :j\geq 0\}$ and the coupling time T are finite almost surely.

Proposition 4.1. *Assume A1. Then for all $(x,x')\in\mathsf{X}\times\mathsf{X}$, $j\geq 0$, $\tilde{\mathbb{P}}_{x,x',0}\left(T_j<\infty\right)=1$ and $\tilde{\mathbb{P}}_{x,x',0}(T<\infty)=1$.*

Proof. The proof is by induction. By Proposition 3.1 and (4.9),

$$\tilde{\mathbb{P}}_{x,x',0}\left(T_0<\infty\right)=\bar{\mathbb{P}}_{x,x',0}\left(T_0<\infty\right)=1.$$

Assume that for $j\geq 0$, $\tilde{\mathbb{P}}_{x,x',0}\left(T_j<\infty\right)=1$; by (4.2), (4.8) and Proposition 3.1,

$$\begin{aligned}\tilde{\mathbb{P}}_{x,x',0}\left(T_{j+1}<\infty\right)&=\tilde{\mathbb{P}}_{x,x',0}\left(T_0\circ\theta^{T_j+m}<\infty,T_j<\infty\right)\\&=\tilde{\mathbb{P}}_{x,x',0}\left(\tilde{\mathbb{P}}_{\Phi_{T_j+m},\Phi'_{T_j+m},0}\left(T_0<\infty\right)\right)=1.\end{aligned}$$

By (4.5), $\tilde{\mathbb{P}}_{x,x',0}(T<\infty)=\sum_{j\geq 0}\tilde{\mathbb{P}}_{x,x',0}(T=T_j+m)=\epsilon\sum_{j\geq 0}(1-\epsilon)^j=1$. □

4.2 Proofs of Theorems 3.2 to 3.7

4.2.1 Preliminary lemmas

Define

$$M=(1-\epsilon)\sup_{(x,x')\in\Delta}\int R_m(x,dy)R_m(x',dy')\tilde{\mathbb{E}}_{y,y',0}\left[\sum_{k=0}^{T_0+m}\tilde{r}_\phi(k)\right],$$

which is finite under A1-3 (see Remark 1 below).
r_ϕ is a subgeometric sequence : there exist finite constants $\bar{c}_\phi, \underline{c}_\phi$ such that

$$\underline{c}_\phi \tilde{r}_\phi(n) \leq r_\phi(n) \leq \bar{c}_\phi \tilde{r}_\phi(n), \qquad \forall n \geq 0,$$

where $\tilde{r}_\phi$ is a non-decreasing positive sequence, $\tilde{r}_\phi \geq 2$ and $\log \tilde{r}_\phi(n)/n$ is non-increasing and tends to zero as n tends to infinity. By [17, Lemmas 1, 2], there exist finite positive constants δ, N, γ such that

$$\begin{aligned}
&\rho := (1+\delta)(1-\epsilon) + M\delta < 1, \\
&\tilde{r}_\phi(n) \leq \delta \sum_{k=0}^{n} \tilde{r}_\phi(k), \text{ for all } n \geq N \\
&\sum_{k=0}^{n+n'} \tilde{r}_\phi(k) \leq (1+\delta) \sum_{k=0}^{n'} \tilde{r}_\phi(k) + \gamma, \text{ for all } n \leq N \text{ and } n' \geq 0.
\end{aligned}$$

Lemma 4.2. *Assume A1-3. Then*

$$\tilde{\mathbb{E}}_{x,x',0}\left[\sum_{n=0}^{T-1} r_\phi(n)\right] \leq \bar{c}_\phi \tilde{\mathbb{E}}_{x,x',0}\left[\sum_{k=0}^{T_0+m-1} \tilde{r}_\phi(k)\right] + R$$

where

$$R = \bar{c}_\phi \frac{\tilde{r}_\phi(0)}{\epsilon}\left(1 - \epsilon + (1+\epsilon)\frac{M\rho}{1-\rho}\right) + \bar{c}_\phi \frac{1+\epsilon}{\epsilon^2} M\gamma.$$

Furthermore, there exist a finite constant $\tilde{R}$ depending only on ϕ such that

$$\tilde{\mathbb{E}}_{x,x',0}\left[r_\phi(T)\right] \leq \tilde{R} \;\; \phi\left(\tilde{\mathbb{E}}_{x,x',0}\left[\sum_{n=0}^{T-1} r_\phi(n)\right]\right). \tag{4.10}$$

Remark 1. It is easily verified that for $(x,x') \in \Delta$,

$$\tilde{\mathbb{E}}_{x,x',0}\left[\sum_{k=0}^{T_0+m-1} \tilde{r}_\phi(k)\right] \leq \underline{c}_\phi^{-1} \sum_{k=0}^{m-1} r_\phi(k)$$

while for $(x,x') \in \Delta^c$, by using the properties of the sequences $\tilde{r}_\phi$, Eq. (4.9) and Proposition 3.1,

$$\begin{aligned}
\tilde{\mathbb{E}}_{x,x',0}\left[\sum_{k=0}^{T_0+m-1} \tilde{r}_\phi(k)\right] &\leq \underline{c}_\phi^{-1}\left\{1 + \underline{c}_\phi^{-1}\sum_{k=1}^{m} r_\phi(k)\right\} \tilde{\mathbb{E}}_{x,x',0}\left[\sum_{k=0}^{T_0-1} \tilde{r}_\phi(k)\right] \\
&\leq \frac{\bar{c}_\phi}{\underline{c}_\phi}\left\{1 + \underline{c}_\phi^{-1}\sum_{k=1}^{m} r_\phi(k)\right\}\left\{V(x) + V(x')\right\}.
\end{aligned}$$

Define

$$\tilde{M} = (1-\epsilon) \sup_{(x,x')\in\Delta} \int R_m(x,dy)R_m(x',dy')\tilde{\mathbb{E}}_{x,x',0}\left[\sum_{n=0}^{T_0+m-1} \phi\{V(\Phi_n)+V(\Phi'_n)\}\right].$$

which is finite under A1-3 (see Remark 2 below).

Lemma 4.3. *Assume A1-3. Then for all $(x,x') \in \mathsf{X}\times\mathsf{X}$*

$$\tilde{\mathbb{E}}_{x,x',0}\left[\sum_{n=0}^{T-1} \phi\{V(\Phi_n)+V(\Phi'_n)\}\right] \le \tilde{\mathbb{E}}_{x,x',0}\left[\sum_{n=0}^{T_0+m-1} \phi\{V(\Phi_n)+V(\Phi'_n)\}\right] + \epsilon^{-1}\tilde{M}.$$

Remark 2. It is easily verified by (4.6) and [12, Proposition 11.3.2] that for $(x,x') \in \Delta$,

$$\tilde{\mathbb{E}}_{x,x',0}\left[\sum_{n=0}^{T_0+m-1} \phi\{V(\Phi_n)+V(\Phi'_n)\}\right] \le V(x)+V(x')+2mb,$$

and for $(x,x') \in \Delta^c$, by (4.6), (4.9), Proposition 3.1 and [12, Proposition 11.3.2]

$$\begin{aligned}\tilde{\mathbb{E}}_{x,x',0}\left[\sum_{n=0}^{T_0+m-1} \phi\{V(\Phi_n)+V(\Phi'_n)\}\right] &\le \bar{\mathbb{E}}_{x,x',0}\left[\sum_{n=0}^{T_0-1} \phi\{V(\Phi_n)+V(\Phi'_n)\}\right] + 2(\sup_{\mathcal{D}} V + mb) \\ &\le V(x)+V(x')+2(\sup_{\mathcal{D}} V + mb).\end{aligned}$$

Define $\bar{M} = \bar{M}_1 \vee \bar{M}_2 \vee \bar{M}_3$ where

$$\begin{aligned}\bar{M}_1 &= \sup_{(x,x')\in\Delta} \phi\{V(x)+V(x')\}, \\ \bar{M}_2 &= \sup_{(x,x')\in\Delta} \int P^k(x,dy)P^k(x',dy')\phi\{V(y)+V(y')\}, \\ \bar{M}_3 &= (1-\epsilon) \sup_{(x,x')\in\Delta} \int R_m(x,dy)R_m(x',dy')\phi\left(V(y)+V(y')\right),\end{aligned}$$

which is finite under A1-3 (see Remark 3 below).

Lemma 4.4. *Assume A1-3. For all $(x,x') \in \mathsf{X}\times\mathsf{X}$,*

$$\tilde{\mathbb{E}}_{x,x',0}\left[\phi\{V(\Phi_n)+V(\Phi'_n)\}\mathbf{1}_{T>n}\right] \le \tilde{\mathbb{E}}_{x,x',0}\left[\phi\{V(\Phi_n)+V(\Phi'_n)\}\mathbf{1}_{T_0>n}\right] + \epsilon^{-1}\bar{M}. \tag{4.11}$$

For any probability measures (λ,λ') on X such that $\int \lambda(dx)\lambda'(dx')\phi\left(V(x)+V(x')\right) < \infty$,

$$\lim_n \tilde{\mathbb{E}}_{\lambda,\lambda',0}\left[\phi\{V(\Phi_n)+V(\Phi'_n)\}\mathbf{1}_{T>n}\right] = 0. \tag{4.12}$$

Remark 3. For $(x,x') \in \Delta$, the first term in the right hand side of (4.11) is zero. For $(x,x') \in \Delta^c$, by Jensen inequality,

$$\begin{aligned}\tilde{\mathbb{E}}_{x,x',0}\left[\phi\{V(\Phi_n)+V(\Phi'_n)\}\mathbf{1}_{T_0>n}\right] &\le \tilde{\mathbb{E}}_{x,x',0}\left[\phi\{(V(\Phi_n)+V(\Phi'_n))\mathbf{1}_{T_0>n} \vee 1\}\right] \\ &\le \phi\{\tilde{\mathbb{E}}_{x,x',0}\left[V(\Phi_n)+V(\Phi'_n)\mathbf{1}_{T_0>n}+1\right]\} \le \phi\{V(x)+V(x')\}+\phi(1).\end{aligned}$$

4.2.2 Conclusion

Theorem 3.2 (resp. Theorem 3.5) deduces from Theorem 3.6 (resp. Theorem 3.7) by choosing $\Psi_1(x) = (x/p)^p$ and $\Psi_2(x) = (x/(1-p))^{1-p}$ for some $0 < p < 1$, or $(\Psi_1, \Psi_2) = (\mathrm{Id}, \mathbf{1})$ or $(\Psi_1, \Psi_2) = (\mathbf{1}, \mathrm{Id})$.
For all $n \geq 0$, $(x, x') \in \mathsf{X}$ and $|f| \leq \Psi_2(\phi \circ V)$,

$$\begin{aligned}
\Psi_1(r_\phi(n)) \, |P^n f(x) - P^n f(x')| &= \Psi_1(r_\phi(n)) \left| \tilde{\mathbb{E}}_{x,x',0} \left[\{ f(\Phi_n) - f(\Phi'_n) \} \right] \right| \\
&\leq \Psi_1(r_\phi(n)) \, \tilde{\mathbb{E}}_{x,x',0} \left[\left| \{ f(\Phi_n) - f(\Phi'_n) \} \right| \mathbf{1}_{T>n} \right] \\
&\leq \Psi_1(r_\phi(n)) \, \tilde{\mathbb{E}}_{x,x',0} \left[\{ \Psi_2(\phi \circ V)(\Phi_n) + \Psi_2(\phi \circ V)(\Phi'_n) \} \mathbf{1}_{T>n} \right] \\
&\leq 2\tilde{\mathbb{E}}_{x,x',0} \left[r_\phi(n) \mathbf{1}_{T>n} \right] + 2\tilde{\mathbb{E}}_{x,x'} \left[\phi\{ V(\Phi_n) + V(\Phi'_n) \} \mathbf{1}_{T>n} \right], \qquad (4.13)
\end{aligned}$$

by using (4.1), (4.3) and (3.6). This yields

$$\sum_{n \geq 0} \Psi_1(r_\phi(n)) \, |P^n f(x) - P^n f(x')| \\
\leq 2 \left\{ \tilde{\mathbb{E}}_{x,x',0} \left[\sum_{n=0}^{T-1} r_\phi(n) \right] + \tilde{\mathbb{E}}_{x,x',0} \left[\sum_{n=0}^{T-1} \phi\{ V(\Phi_n) + V(\Phi'_n) \} \right] \right\}. \qquad (4.14)$$

Theorem 3.6 now follows from Lemmas 4.2 and 4.3, and Remarks 1-2 where the two terms of the right hand side in (4.14) are considered in turn.
For Theorem 3.7, we consider in turn the two terms in the right hand side of (4.13). By Lemma 4.2, Remark 3 and the dominated convergence theorem, $\lim_n \tilde{\mathbb{E}}_{x,x',0} \left[r_\phi(n) \, \mathbf{1}_{T>n} \right]$ for all (x, x'), and by (4.10),

$$\sup_n \sup_{(x,x')} \left\{ \phi \left(V(x) + V(x') \right) \right\}^{-1} \, \tilde{\mathbb{E}}_{x,x',0} \left[r_\phi(n) \, \mathbf{1}_{T>n} \right] < \infty.$$

Hence, using again the dominated convergence theorem, $\lim_n \tilde{\mathbb{E}}_{\lambda,\lambda',0} \left[r_\phi(n) \, \mathbf{1}_{T>n} \right] = 0$ for all (λ, λ') such that $\int \lambda(dx) \lambda'(dx') \phi \left(V(x) + V(x') \right) < \infty$. The second term in the right hand side of (4.13) tends to zero by Lemma 4.4.

4.2.3 Proof of the lemmas

Proof of Lemma 4.2 *(i)* Define $R_\phi(n) = \sum_{k=0}^{n} \tilde{r}_\phi(k)$, $n \geq 0$. By [19, Lemma 1], $\tilde{r}_\phi(n+m) \leq \tilde{r}_\phi(n) \tilde{r}_\phi(m)$ so that $R_\phi(n+m) \leq R_\phi(n) + \tilde{r}_\phi(n) R_\phi(m)$. We write

$$\tilde{\mathbb{E}}_{x,x',0} \left[R_\phi(T-1) \right] = \tilde{\mathbb{E}}_{x,x',0} \left[R_\phi(T_0 + m - 1) \mathbf{1}_{T = T_0 + m} \right] \\
+ \sum_{n \geq 1} \tilde{\mathbb{E}}_{x,x',0} \left[R_\phi(T_n + m - 1) \mathbf{1}_{T = T_n + m} \right]$$

Consider the general term of the sum : by (4.2), (4.7) and (4.8), for $n \geq 1$,

$$
\begin{aligned}
\tilde{\mathbb{E}}_{x,x',0}\left[R_\phi(T_n+m-1)\mathbf{1}_{T=T_n+m}\right] &\leq \tilde{\mathbb{E}}_{x,x',0}\left[R_\phi(T_{n-1}+m)\mathbf{1}_{T=T_n+m}\right] \\
&+ \tilde{\mathbb{E}}_{x,x',0}\left[\tilde{r}_\phi(T_{n-1}+m)R_\phi(m-1+T_0\circ\theta^{T_{n-1}+m})\mathbf{1}_{T=T_n+m}\right] \\
&\leq \tilde{\mathbb{E}}_{x,x',0}\left[R_\phi(T_{n-1}+m)\mathbf{1}_{T\geq T_n+m}\right] + M\,\tilde{\mathbb{E}}_{x,x',0}\left[\tilde{r}_\phi(T_{n-1}+m)\mathbf{1}_{T\geq T_{n-1}+m}\right].
\end{aligned}
$$

Define for $n \geq 1$,

$$
\begin{aligned}
a_{x,x'}(n) &= \tilde{\mathbb{E}}_{x,x',0}\left[R_\phi(T_{n-1}+m)\mathbf{1}_{T\geq T_n+m}\right], \\
b_{x,x'}(n) &= \tilde{\mathbb{E}}_{x,x',0}\left[\tilde{r}_\phi(T_{n-1}+m)\mathbf{1}_{T\geq T_{n-1}+m}\right].
\end{aligned}
$$

We have (the proof is detailed in [13, Lemma 3.1] and is omitted for brevity)

$$
\sum_{n\geq 1} b_{x,x'}(n) \leq \frac{\rho}{1-\rho}\tilde{r}_\phi(0) + \frac{\gamma}{\epsilon}, \quad \sum_{n\geq 1} a_{x,x'}(n) \leq \frac{1-\epsilon}{\epsilon}\tilde{r}_\phi(0) + \frac{M}{\epsilon}\sum_{n\geq 1} b_{x,x'}(n).
$$

Hence,

$$
\tilde{\mathbb{E}}_{x,x',0}\left[R_\phi(T-1)\right] \leq \tilde{\mathbb{E}}_{x,x',0}\left[R_\phi(T_0+m-1)\right] + \frac{\tilde{r}_\phi(0)}{\epsilon}\left(1-\epsilon+(1+\epsilon)\frac{M\rho}{1-\rho}\right) + \frac{1+\epsilon}{\epsilon^2}M\gamma.
$$

(ii) By definition of r_ϕ, there exists $\tilde{R}$ such that $\sup_n r_\phi(n)\{\phi(\sum_{k=0}^{n-1} r_\phi(k))\}^{-1} \leq \tilde{R}$. The result follows from *(i)*.

Proof of Lemma 4.3 We write

$$
\begin{aligned}
\tilde{\mathbb{E}}_{x,x',0}\left[\sum_{n=0}^{T-1}\phi\{V(\Phi_n)+V(\Phi'_n)\}\right] &= \tilde{\mathbb{E}}_{x,x',0}\left[\sum_{n=0}^{T_0+m-1}\phi\{V(\Phi_n)+V(\Phi'_n)\}\right] \\
&+ \sum_{j\geq 0}\tilde{\mathbb{E}}_{x,x',0}\left[\sum_{n=T_j+m}^{T_{j+1}+m-1}\phi\{V(\Phi_n)+V(\Phi'_n)\}\mathbf{1}_{T\geq T_{j+1}+m}\right].
\end{aligned}
$$

We consider the general term of the sum, hereafter denoted by $s_{x,x'}(j)$: for $j \geq 0$, by (4.8)

$$
\begin{aligned}
s_{x,x'}(j) &= \tilde{\mathbb{E}}_{x,x',0}\left[\tilde{\mathbb{E}}_{x,x',0}\left[\sum_{n=T_j+m}^{T_{j+1}+m-1}\phi\{V(\Phi_n)+V(\Phi'_n)\}|\tilde{\mathcal{F}}_{T_j+m}\right]\mathbf{1}_{d_{T_j+m}=0}\right] \\
&= \tilde{\mathbb{E}}_{x,x',0}\left[\tilde{\mathbb{E}}_{\Phi_{T_j+m},\Phi'_{T_j+m},0}\left[\sum_{n=0}^{T_0+m-1}\phi\{V(\Phi_n)+V(\Phi'_n)\}\right]\mathbf{1}_{d_{T_j+m}=0}\right].
\end{aligned}
$$

Eq. (4.7) yields $s_{x,x'}(j) \leq \tilde{\mathbb{P}}_{x,x',0}(d_{T_j=0})\,\tilde{M} = (1-\epsilon)^j\tilde{M}$. Hence, $\sum_{j\geq 0} s_{x,x'}(j) \leq \epsilon^{-1}\tilde{M}$.

Proof of Lemma 4.4 (i) We use the decomposition

$$\{T > n\} = \{n < T_0\} \cup \bigcup_{j \geq 0} \{T_j \leq n < T_{j+1}, d_{T_j+m} = 0\} .$$

so that

$$\tilde{\mathbb{E}}_{x,x',0}\left[\phi\{V(\Phi_n) + V(\Phi'_n)\}\mathbf{1}_{T>n}\right] = \tilde{\mathbb{E}}_{x,x',0}\left[\phi\{V(\Phi_n) + V(\Phi'_n)\}\mathbf{1}_{0<n<T_0}\right] \\ + \sum_{j\geq 0} \tilde{\mathbb{E}}_{x,x',0}\left[\phi\{V(\Phi_n) + V(\Phi'_n)\}\mathbf{1}_{T_j \leq n < T_{j+1}, T>n}\right] .$$

Hereafter, denote by $s_{x,x'}(j)$ the general term of the series. We prove that for all $j \geq 0$, $s_{x,x'}(j) \leq \bar{M}(1-\epsilon)^j$. Let $j \geq 0$; by definition of T, $\{T_j \leq n < T_j+m, n < T\} = \{T_j \leq n < T_j+m, d_{T_j} = 0\}$ so that

$$\tilde{\mathbb{E}}_{x,x'}\left[\phi\{V(\Phi_n) + V(\Phi'_n)\}\mathbf{1}_{T_j=n,T>n}\right] \leq \bar{M}_1 \, \tilde{\mathbb{P}}_{x,x',0}\left(d_{T_j} = 0\right) \leq \bar{M} \, (1-\epsilon)^j. \tag{4.15}$$

By (4.6),

$$\tilde{\mathbb{E}}_{x,x',0}\left[\phi\{V(\Phi_n) + V(\Phi'_n)\}\mathbf{1}_{T_j+1\leq n<T_j+m,T>n}\right] \leq \bar{M}_2 \, \tilde{\mathbb{P}}_{x,x',0}\left(d_{T_j} = 0\right) \leq \bar{M} \, (1-\epsilon)^j. \tag{4.16}$$

Upon noting that $\{n = T_j + m, n < T\} = \{n = T_j + m, d_{T_j+m} = 0\}$, (4.4) and (4.7) yield

$$\tilde{\mathbb{E}}_{x,x',0}\left[\phi\{V(\Phi_n) + V(\Phi'_n)\}\mathbf{1}_{n=T_j+m,T>n}\right] \leq \bar{M}_3 \, \tilde{\mathbb{P}}_{x,x',0}\left(d_{T_j} = 0\right) \leq \bar{M} \, (1-\epsilon)^j.$$

Finally, since $\{T_j + m < n < T_{j+1}, n < T\} = \{T_j + m < n < T_{j+1}, d_{T_j+m} = 0\}$, we have by using (4.8), Jensen inequality and the drift inequality (3.1)

$$\tilde{\mathbb{E}}_{x,x',0}\left[\phi\{V(\Phi_n) + V(\Phi'_n)\}\mathbf{1}_{T_j+m<n<T_{j+1},T>n}\right] \\ \leq \tilde{\mathbb{E}}_{x,x',0}\left[\phi\left(\tilde{\mathbb{E}}_{x,x',0}\left[\{V(\Phi_n) + V(\Phi'_n)\}\mathbf{1}_{T_j+m<n<T_{j+1}}|\tilde{\mathcal{F}}_{T_j+m}\right] \vee 1\right)\mathbf{1}_{d_{T_j+m}=0}\right] \\ \leq \tilde{\mathbb{E}}_{x,x',0}\left[\phi\left(V(\Phi_{T_j+m}) + V(\Phi'_{T_j+m})\right)\mathbf{1}_{d_{T_j+m}=0}\right] \leq \bar{M}_3 \, \tilde{\mathbb{P}}_{x,x',0}\left(d_{T_j} = 0\right) = \bar{M} \, (1-\epsilon)^j.$$

(ii) For the proof of (4.12), we use the decomposition

$$\{T > n\} = \{n < T_0 + m\} \cup \bigcup_{j\geq 0} \{T_j + m \leq n < T_{j+1} + m, d_{T_j+m} = 0\} .$$

Define for all $n, j \geq 0$,

$$a_{\lambda,\lambda'}(n,j) = (1-\epsilon)^{-j}\tilde{\mathbb{E}}_{\lambda,\lambda',0}\left[\phi\{V(\Phi_n) + V(\Phi'_n)\}\mathbf{1}_{T_{j-1}+m\leq n<T_j+m,T>n}\right]$$

where $T_{-1} + m = 0$; we have to prove that $\lim_n \sum_{j\geq 0}(1-\epsilon)^j a_{\lambda,\lambda'}(n,j) = 0$ which will be done by repeated applications of the dominated convergence theorem.
We first establish that $\lim_n a_{\lambda,\lambda'}(n,0) = 0$. From Proposition 3.1 and (4.6), we have for all $(x,x') \in \mathsf{X} \times \mathsf{X}$, $\sum_{n\geq 0} a_{x,x'}(n,0) < \infty$ which implies that $\lim_n a_{x,x'}(n,0) = 0$. In addition, from

(4.15)-(4.16), $a_{x,x'}(n,0) \leq \phi\left(V(x)+V(x')\right)$ which is $\lambda \otimes \lambda'$-integrable by assumptions. By the dominated convergence theorem, it follows that $\lim_n a_{\lambda,\lambda'}(n,0) = 0$.
Let $j \geq 1$; following the same lines as in the proof of (4.11), we have $\sup_{j\geq 1} \sup_n a_{\lambda,\lambda'}(n,j) \leq \bar{M}$. Furthermore,

$$a_{\lambda,\lambda'}(n,j) = (1-\epsilon)^{-j} \tilde{\mathbb{E}}_{\lambda,\lambda',0} \left[\tilde{\mathbb{E}}_{\Phi_{T_{j+m}},\Phi'_{T_{j+m}},0} \left[\phi\{V(\Phi_n) + V(\Phi'_n)\} \mathbf{1}_{n<T_0+m, T>n} \right] \mathbf{1}_{d_{T_{j+m}}=0} \right]$$

and as done above for $a_{\lambda,\lambda'}(n,0)$, the dominated convergence theorem gives $\lim_n a_{\lambda,\lambda'}(n,j) = 0$. We write $\sum_{j\geq 1}(1-\epsilon)^j \; a_{\lambda,\lambda'}(n,j) = (1-\epsilon)\epsilon^{-1}\mathbb{E}[a_{\lambda,\lambda'}(n,\upsilon)]$ where the expectation is with respect to υ, a geometric random variable with success probability ϵ and independent of the process $\{(\Phi_n, \Phi'_n, d_n) \;\; : n \geq 0\}$. Using again the dominated convergence theorem, we have $\lim_n \mathbb{E}[a_{\lambda,\lambda'}(n,\upsilon)] = 0$, and this concludes the proof.

References

References

[1] C. Andrieu and E. Moulines. On the ergodicity Properties of some Adaptive MCMC Algorithms. Technical report, School of Mathematics, University of Bristol, 2005. Available at :http://www.stats.bris.ac.uk/ maxca/.

[2] P.H. Baxendale. Renewal theory and computable convergence rates for geometrically ergodic Markov chains. *Ann. Appl. Probab.*, 15(1):700–738, 2005.

[3] R. Douc, G. Fort, E. Moulines, and P. Soulier. Practical drift conditions for subgeometric rates of convergence. *Ann. Appl. Probab.*, 14(3):1353–1377, 2004.

[4] R. Douc, E. Moulines, and J.S. Rosenthal. Quantitative bounds on convergence of time-inhomogeneous Markov chains. *Ann. Appl. Probab.*, (4), 2004.

[5] R. Douc, E. Moulines, and P. Soulier. Computable bounds for subgeometric ergodicity. Technical report, Prepublication 186, University Evry-Val d'Essonne, 2003. Available at http://www.tsi.enst.fr/∼soulier/paper.html.

[6] G. Fort. *Contrôle explicite d'ergodicité de chaîne de Markov: Applications à l'analyse de convergence de l'algorithme Monte-Carlo EM.* PhD thesis, Université Pierre et Marie Curie, Paris, 2001. Available at http://www.tsi.enst.fr/∼gfort/soutenance.html

[7] G. Fort. Computable bounds for V-geometric ergodicity of Markov transition kernels. Technical report, LMC, University Joseph Fourier, Grenoble., 2002.

[8] G. Fort and E. Moulines. Polynomial ergodicity of Markov transition kernels. *Stoc. Proc. Appl.*, 103(1):57–99, 2003.

[9] S. Jarner and G.O. Roberts. Polynomial convergence rates of Markov Chains. *Ann. Appl. Prob.*, 12(1):224–247, 2002.

[10] M.A. Krasnosel'skii and Y.B. Rutickii. *Convex functions and Orlicz spaces.* Noordhoff, Groningen, 1961.

[11] S. Meyn and R. Tweedie. Computable bounds for geometric convergence rates of Markov chains. *Ann. Appl. Probab.*, 4:981–1011, 1994.

[12] S. P. Meyn and R. L. Tweedie. *Markov chains and stochastic stability.* Springer-Verlag London Ltd., London, 1993.

[13] E. Nummelin and P. Tuominen. The rate of convergence in Orey's theorem for Harris recurrent Markov chains with applications to renewal theory. *Stoc. Proc. Appl.*, 15:295–311, 1983.

[14] G.O. Roberts and J.S. Rosenthal. Coupling and Ergodicity of Adaptive MCMC. Technical report, Dpt of Statistics, University of Toronto, 2005. Available at http://probability.ca/jeff/.

[15] G.O. Roberts and R.L. Tweedie. Bounds on regeneration times and convergence rates for Markov Chains. *Stoc. Proc. Appl.*, 80(2):211–229, 1999.

[16] J.S. Rosenthal. Minorization conditions and convergence rates for Markov chain Monte Carlo. *J. Amer. Statist. Assoc.*, 90(430):558–566, 1995.

[17] C. Stone and S. Wainger. One-sided error estimates in renewal theory. *J. Analyse Math.*, 20:325–352, 1967.

[18] A. Tanikawa. On the rate of convergence of borovkov's multidimensional ergodic markov chains. *J. Appl. Prob.*, 38:328–339, 1997.

[19] H. Thorisson. The queue GI/G/1: finite moments of the cycle variables and uniform rates of convergence. *Stoch. Proc. Appl.*, 19:83–99, 1985.

[20] P. Tuominen and R.L. Tweedie. Subgeometric rates of convergence of f-ergodic Markov Chains. *Adv. in Appl. Prob.*, 26(3):775–798, 1994.