

COMPUTABLE BOUNDS FOR V -GEOMETRIC ERGODICITY OF MARKOV TRANSITION KERNELS

GERSENDE FORT

SUMMARY

This paper discusses quantitative bounds for the fluctuations of the n -step transition law of a Markov kernel P on a general state space $\|P^n(x, \cdot) - P^n(x', \cdot)\|$. The fluctuations are successively measured in total variation norm and in V -norm. The basic assumptions are couched in terms of (a) a minorization condition on some m -iterated of the kernel ($m \geq 1$), and (b) a Foster-Lyapunov drift condition on the (one-iterated) kernel P , which are conditions implying a geometrical decay, in n , of the fluctuations.

1991 Mathematics Subject classification: 60J10, 60J22, 65C40.

Keywords: V -geometric ergodicity, rates of convergence, computable bounds.

Acknowledgements: This work was initiated during my Ph.D. I would like to thank Eric Moulines for introducing me to the topic of computable bounds for ergodicity of Markov kernels.

Affiliations: G. Fort, LMC-IMAG, 51, rue des Mathématiques, BP 53, 38041 Grenoble Cedex 9, France, *email:* Gersende.Fort@imag.fr

1. INTRODUCTION

We consider a Markov chain $\Phi = (\Omega, \mathcal{A}, \mathcal{F}_n, \{\phi_n\}, P_x)$ on a general state space $\mathcal{X}$ endowed with a countably generated σ -field $\mathcal{B}(\mathcal{X})$. Let P be the transition kernel and denote by $P^n(x, \cdot)$ its n -step transition law, $P^n(x, \cdot) := P(\phi_n \in \cdot | \phi_0 = x)$. For some probability distribution λ and some positive measurable function f , we denote by $\lambda P^n f$ the conditional expectation of $f(\Phi_n)$ given the initial distribution λ

$$\lambda P^n f := \int \lambda(dx) P^n(x, dy) f(y);$$

the notation $P^n f(x)$ stands for $\delta_x P^n f$. We assume throughout the paper that P is ψ -irreducible and aperiodic (see Meyn and Tweedie (1993) for relevant definitions on Markov chains). Our aim is to give an explicit upper-bound B for the fluctuations of the iterated of P *i.e.*

$$\|P^n(x, \cdot) - P^n(x', \cdot)\| \leq B(x, x'; n), \quad (1)$$

where $\|\cdot\|$ will be successively the total variation norm and the V -norm. Remind that the total variation norm of a signed measure μ is given by

$$\|\mu\|_{\text{TV}} := \sup_{A \in \mathcal{B}(\mathcal{X})} \mu(A) - \inf_{A \in \mathcal{B}(\mathcal{X})} \mu(A) = \sup_{\{f, |f| \leq 1\}} |\mu(f)|,$$

and for a positive measurable function V , the V -norm

$$\|\mu\|_V := \sup_{\{f, |f| \leq V\}} |\mu(f)|,$$

so that $\|\mu\|_1 = \|\mu\|_{\text{TV}}$. Markov chains are widely used in statistics and this necessitates explicit control of the convergence. For example, the bound (1) yields quantitative control for V -ergodicity of the Markov chains and thus is of practical interest to monitor convergence of Markov Chain Monte Carlo samplers. We improve and extend earlier results for control in total variation norm and in V -norm (Baxendale (1993); Fort (2001); Kolassa (2000); Lund and Tweedie (1996); Mengersen and Tweedie (1996); Meyn and Tweedie (1994); Roberts and Tweedie (1999); Rosenthal (1995, 2002)). To that goal, we formulate basic assumptions, couched in terms of a minorization condition on some m -iterated kernel, coupled with a drift condition on the one-iterated kernel. For the control in total variation norm and in V -norm, we give the expression of a rate $\beta > 1$ such that $\limsup_n \beta^n B(x, x'; n) < \infty$, expression in terms of the constants appearing in the minorization condition and in the drift condition. This rate improves previous ones and coincides with the rate exhibited in Rosenthal (2002) in a contemporaneous work restricted to the control of the fluctuations (1) in total variation norm.

For countable Markov chain, the basic condition involved for such a scope is the existence of a state a such that $\sum_n \beta^n P_a(\tau_a \geq n)$ is finite, where τ_a is the return-time to $\{a\}$, $\tau_a := \inf\{n \geq 1, \phi_n = a\}$. For general Markov chain, the corresponding condition has to be satisfied for some ν_m -small set $C \in \mathcal{B}(\mathcal{X})$ *i.e.* a set such that there exist a positive integer m , a constant $\epsilon > 0$ and a probability distribution ν_m and

$$\forall x \in C, \quad P^m(x, \cdot) \geq \epsilon \nu_m(\cdot). \quad (2)$$

It is well-known that if there exists a small set C which is a Kendall set *i.e.* $\sup_{x \in C} \mathbb{E}_x[\kappa^{\tau_C}] < \infty$ for some $\kappa > 1$, then there exist some constants $\beta > 1$ and $R < \infty$ such that

$$\sum_n \beta^n \|P^n(x, \cdot) - P^n(x', \cdot)\|_{\text{TV}} \leq R (\mathbb{E}_x[\kappa^{\tau_C}] + \mathbb{E}_{x'}[\kappa^{\tau_C}]). \quad (3)$$

For control in V -norm, the small set has to be a V -Kendall set *i.e.* $\sup_{x \in C} \mathbb{E}_x \left[\sum_{k=0}^{\tau_C-1} \kappa^k V(\Phi_k) \right] < \infty$ for some $\kappa > 1$. In that case,

$$\sum_n \beta^n \|P^n(x, \cdot) - P^n(x', \cdot)\|_V \leq R \left\{ \mathbb{E}_x \left[\sum_{k=0}^{\tau_C-1} \kappa^k V(\Phi_k) \right] + \mathbb{E}_{x'} \left[\sum_{k=0}^{\tau_C-1} \kappa^k V(\Phi_k) \right] \right\}, \quad (4)$$

for some constants $\beta > 1$ and $R < \infty$ (Nummelin and Tuominen, 1982; Meyn and Tweedie, 1993, Theorem 15.0.1.). In this contribution, the rate β and the upper bound are explicitly expressed, and we will see that the given bound is closely related to the modulated moment as in (3) and (4). Most of earlier works on the computation of the bound B in (1) are based on the coupling construction, and we do in this manner. We assume throughout this paper that

A1 there exist some measurable set $C \subseteq D$ of $\mathcal{B}(\mathcal{X})$, a positive integer m , a constant $\epsilon > 0$ and a transition kernel ν on $\mathcal{X} \times \mathcal{X}$ to $\mathcal{X}$ such that

$$\forall (x, x') \in \Delta, \quad P^m(x, \cdot) \wedge P^m(x', \cdot) \geq \epsilon \nu(x, x'; \cdot),$$

where $\Delta := (C \times D) \cup (D \times C)$.

For example, one may choose

$$\nu(x, x'; A) \propto \int_A \left\{ \frac{dP^m(x, \cdot)}{dP^m(x', \cdot)} \wedge 1 \right\} P^m(x', dy),$$

where $dP^m(x, \cdot)/dP^m(x', \cdot)$ denotes the derivative of the absolutely continuous part of $P^m(x, \cdot)$ w.r.t. $P^m(x', \cdot)$; in that case, assumption A1 is known in the literature as the local Doeblin condition. It generalizes the ν_m -smallness assumption. Small sets always exist for ψ -irreducible

Markov kernel so that assumption A1 is always verified with a ν_m -small set C and $\Delta := C \times C$. This assumption allows the definition of a residual kernel R on $\mathcal{X} \times \mathcal{X}$ to $\mathcal{X} \times \mathcal{X}$,

$$R(x, x'; dy \times dy') := \left(\frac{P^m(x, dy) - \epsilon \mathbb{I}_\Delta(x, x') \nu(x, x'; dy)}{1 - \mathbb{I}_\Delta(x, x') \epsilon} \right) \left(\frac{P^m(x, dy') - \epsilon \mathbb{I}_\Delta(x, x') \nu(x, x'; dy')}{1 - \mathbb{I}_\Delta(x, x') \epsilon} \right). \quad (5)$$

For a function U on $\mathcal{X}^2$, set $RU(x, x') := \int R(x, x'; dy \times dy') U(y, y')$. The geometrical decreasing rate of the fluctuations, measured either in total variation norm or in V -norm, is related to the geometrical decreasing rate of the tails of some random-time T , that in words is defined as follows. Let σ be the hitting-time on Δ of two independent copies of a Markov chain with transition kernel P . Define

$$U_*(x, x') := \mathbb{E}_{x, x'}[\beta_*^\sigma] \quad \beta_* > 1, \quad (6)$$

some geometrical moment for the associated canonical probability given the process starts from (x, x') . **(a)** Run two independent copies (ϕ_n, ϕ'_n) of the Markov chain, with initial value (x, x') , till σ when they enter the set Δ . **(b)** With probability ϵ , set $T := m + \sigma$ and otherwise, that is with probability $1 - \epsilon$ run two independent copies with initial distribution $R(\phi_\sigma, \phi'_\sigma; \cdot)$ till they enter the set Δ . Repeat step (b) until a successful Bernoulli trial. Roughly speaking, as the use of the residual kernel introduces a gap of size m , $T = m + T'$ where T' is defined as the combination of (a) a hitting-time on Δ of two independent copies of a Markov chain with kernel P and initial distribution $R(x, x'; \cdot)$ for some $(x, x') \in \Delta$ and (b) a successful Bernoulli trial with parameter ϵ , independent of the Markov chains. This description stresses that the tails of T (and thus the upper bound B) are geometrically decreasing at a rate β_{opt} , depending on ϵ and $\beta_*^m \sup_\Delta RU_*$. We will see that while the rate does not depend upon the norm, the "maximal" control function V is related to the moment U_* . This first result is of theoretical interest except when U_* can be computed in closed form. Nevertheless, we verify that a similar conclusion holds by substituting U_* for an upper-estimate $\hat{U}_* \geq U_*$. It is well-known that any function $U \geq 1$ solving the drift condition $P \otimes P(U) \leq \beta_*^{-1} U$ on Δ^c is a candidate for $\hat{U}_*$. From a practical point of view, one would prefer to solve a drift condition for the kernel P . We establish that condition A2 enables the construction of such a function $\hat{U}_*$.

A2 there exist a measurable function $V \geq 1$ and some positive constants $\lambda < 1$, $b < \infty$ such that $\sup_D P^m V < \infty$ and

$$PV \leq \lambda V + b \mathbb{I}_C \quad (7)$$

$$\beta_*^{-1} := \lambda + b \left\{ \inf_C V + \inf_{D^c} V \right\}^{-1} < 1. \quad (8)$$

Observe that by iterating (7), $\sup_D P^m V < \infty$ if $\sup_D V < \infty$. The inequality (7) is known in the literature as the Foster-Lyapunov drift condition. If (7) holds for some set C then it still holds for some $D \supseteq C$, (λ, b) remaining unchanged. If C is small and V goes to infinity outside C , then the set D satisfying (8) may be chosen as a level set of V and thus inherits of the smallness property. This discussion evidences that

if there exist a small set C , a measurable function $V \geq 1$ that goes to infinity outside C and some constants $\lambda < 1$, $b < \infty$ such that $PV \leq \lambda V + b \mathbb{I}_C$, then there exists a small set $D \supseteq C$ such that A1-2 are verified.

In the literature, explicit control of the fluctuations in total-variation norm are derived under a minorization condition and a drift condition both on the original kernel P . The work by Meyn and Tweedie (1994) for control in V -norm is developed under the same set assumptions. This is a restrictive framework as in many applications, one is able to establish the drift condition A2 for some set D , which is ν_m -small, $m > 1$. In that case, it is basically suggested to consider the m -skeleton $Q = P^m$, the one-step of which satisfies a minorization condition on D . This yields a control of the fluctuations for the initial kernel P upon noting that $P^n = Q^{\lfloor n/m \rfloor} \circ Q^l$ for some $0 \leq l \leq m - 1$, where $\lfloor r \rfloor$ is the largest integer that does not exceed r . Nevertheless, this approach leads to cumbersome computations (see e.g. (Fort, 2001, Chapter 3)), and it is quite difficult to give an explicit expression of the bound in terms of the quantities appearing in A1 and A2. For these reasons, we think that our "un-coupled" assumptions are of interest since the present method provides an alternative to a method that has never been carried out.

In Sections 2.1 and 2.2, we set out the obtained result. Section 2.1 is devoted to the control of the fluctuations (1) in total variation norm. To our best knowledge, very few is known on explicit computations of a bound for the fluctuations (1) in V -norm, the only results available in this direction are in Meyn and Tweedie (1994) and Fort (2001). In Section 2.2, we extend and improve these previous contributions. Examples are worked out in Section 3. The proofs are postponed in Section 4 and 5. As mentioned above, they rely on the coupling method, which

consists in defining a process $\tilde{Z} = (\Omega, \mathcal{A}, \mathcal{F}_n^{\tilde{Z}}, \{\tilde{Z}_n := (\tilde{X}_n, \tilde{X}'_n, \tilde{d}_n)\}, \tilde{\mathbb{P}}_{x,x',i})$ on $\mathcal{Z} := \mathcal{X} \times \mathcal{X} \times \{0, 1\}$ such that

$$\tilde{\mathbb{P}}_{x,x',i}(\tilde{X}_n \in \cdot) = P^n(x, \cdot) \quad \tilde{\mathbb{P}}_{x,x',i}(\tilde{X}'_n \in \cdot) = P^n(x', \cdot), \quad (9)$$

and for some coupling time T , $\tilde{X}_n \mathbb{1}_{T \leq n} = \tilde{X}'_n \mathbb{1}_{T \leq n}$. When $m = 1$, our approach is nothing else than the Nummelin's splitting (see Nummelin (1978)). For $m > 1$, we propose an unusual construction which is adapted from the procedure in Kalashnikov (1994) initially proposed to embed the original Markov process in a regenerative one. This approach differs from the usually suggested sub-sampling technique. In Section 4, we address the case when A1 is verified with $m = 1$. The case $m > 1$ is more technical and is presented in Section 5.

2. STATEMENTS OF THE RESULTS

2.1. Rate of convergence in total variation norm. Let β_* be the largest value such that $\mathbb{E}_{x,x'}[\beta^\sigma] < \infty$ at least at some point in Δ^c . Then the fluctuations $\|P^n(x, \cdot) - P^n(x', \cdot)\|_{\text{TV}}$ are geometrically decreasing with rate β_* whenever $(1 - \epsilon)\beta_*^m \sup_{\Delta} RU_* \leq 1$. Otherwise the largest rate that can be achieved is given by the solution to

$$1 < \beta \leq \beta_* \quad \text{and} \quad (1 - \epsilon) \beta^m \sup_{(x,x') \in \Delta} \int R(x, x'; dy \times dy') \mathbb{E}_{y,y'}[\beta^\sigma] = 1. \quad (10)$$

When the state space is ν_m -small *i.e.* (2) holds with $C := \mathcal{X}$, then $\sigma = 0$ w.p.1 and the maximal rate is β_* such that $1 = (1 - \epsilon)\beta_*^m$ *i.e.* $\beta_* := (1 - \epsilon)^{-1/m}$. We establish in Section 5.3 Remark 12 that in that case,

$$\forall(x, x') \quad \|P^n(x, \cdot) - P^n(x', \cdot)\|_{\text{TV}} \leq 2(1 - \epsilon)^{\lfloor n/m \rfloor}, \quad n \geq m - 1,$$

which is a well-known result (Meyn and Tweedie, 1993, Theorem 16.0.2.).

By using the Jensen's inequality, the solution of (10) may be approximated and lower-bounded by β such that

$$1 < \beta \leq \beta_* \quad \text{and} \quad (1 - \epsilon) \{\beta_*^m \sup_{\Delta} RU_*\}^{\ln \beta / \ln \beta_*} = 1. \quad (11)$$

This yields the following theorem, proved in Section 4.2 for $m = 1$ and Section 5.3 for $m > 1$.

Theorem 1. *Assume A1. Assume in addition that*

A3 *there exists $\beta_* > 1$ such that $\sup_{\Delta} RU_* < \infty$.*

For any (x, x') such that $U_*(x, x') < \infty$ and any $0 \leq \alpha \leq 1$,

$$\|P^n(x, \cdot) - P^n(x', \cdot)\|_{\text{TV}} \leq 2 \beta_*^{-\alpha\tau n} \beta_*^{\alpha\tau(m-1)} U_*^{\alpha\tau}(x, x'), \quad n \geq m-1, \quad (12)$$

where

$$\tau := \min\{1; -\ln(1-\epsilon)/\ln[\beta_*^m \sup_{\Delta} RU_*]\}. \quad (13)$$

When applied with $\alpha = 1$, Theorem 1 asserts that

$$\sup_n \beta_*^{\tau n} \|P^n(x, \cdot) - P^n(x', \cdot)\|_{\text{TV}} \leq 2 \beta_*^{\tau(m-1)} U_*^{\tau}(x, x').$$

If $\epsilon \rightarrow 0$, the probability that some hitting-time on Δ and a successful Bernoulli trial would coincide goes to zero. The tails of the random-time T get heavier and thus the obtained rate of convergence degenerates ($\tau \rightarrow 0$ and $\beta_*^{\tau} \rightarrow 1$). In that case, the right-hand side of (12) tends to 2 and inequality (12) is trivial. Conversely, if $\epsilon \rightarrow 1$, it is very likely that the first hitting-time on Δ and the successful Bernoulli trial would coincide. In that case, $\tau \rightarrow 1$ and the obtained rate is β_* .

The dependence upon the initial value is "collected" in U_* ; it is indeed related to the first time two independent copies of the Markov kernel P reach the set Δ . Observe that the greater the rate, the stronger the dependence upon the initial value. Finally, Theorem 1 implies that for any $1 \leq \beta < \beta_*^{\tau}$,

$$\sum_n \beta^n \|P^n(x, \cdot) - P^n(x', \cdot)\|_{\text{TV}} \leq 2 \beta_*^{\tau(m-1)} (1 - \beta\beta_*^{-\tau})^{-1} U_*^{\tau}(x, x').$$

The same radius of convergence is obtained in (Fort, 2001, Proposition 4.17) under the same assumptions A1 and A2.

In many applications, U_* can not be computed in closed form. Nevertheless, if there exists $\hat{U}_* \geq U_*$, then the above conclusion still holds with U_* replaced by $\hat{U}_*$. Two methods are proposed to obtain such an upper-estimate $\hat{U}_*$, both based on a Foster-Lyapunov drift condition. It is well-known that the minimal pointwise solution to

$$\int P(x, dy) P(x', dy') U(y, y') \leq \beta_*^{-1} U(x, x') \quad (x, x') \in \Delta^c, \quad U \geq 1 \quad \text{on} \quad \Delta, \quad (14)$$

is $\mathbb{E}_{x, x'} [\beta_*^{\sigma}]$ (see e.g. (Meyn and Tweedie, 1993, Chapter 15); a proof is also given in Section 4.4 for the sake of completeness). Any function U solving (14) is a candidate for $\hat{U}_*$. When the expression of $P \otimes P$ is complex, one may prefer to verify assumption A2, couched in terms of a drift condition on the one-iterated kernel P . In that case, we prove in Section 4.4 that $\hat{U}_*(x, x') \sim (V(x) + V(x'))/2$.

Proposition 2. *Assume A1-2. Then the function*

$$\hat{U}_*(x, x') := \frac{1}{2} (V(x) + V(x')) \quad (15)$$

verifies the conditions (14) and $\hat{U}_ \geq U_*$.*

Theorem 1 extends as follows. The proof is given in Section 4.3 for $m = 1$ and in Section 5.4 for $m > 1$.

Proposition 3. *Assume A1, A3.*

(i) *For any function $\hat{U}_* \geq U_*$ such that $\sup_{\Delta} R\hat{U}_* < \infty$ and any $0 \leq \alpha \leq 1$,*

$$\|P^n(x, \cdot) - P^n(x', \cdot)\|_{\text{TV}} \leq 2 \beta_*^{-\alpha \bar{\tau} n} \beta_*^{\alpha \bar{\tau} (m-1)} \hat{U}_*^{\alpha \bar{\tau}}(x, x'), \quad n \geq m-1, \quad (16)$$

where

$$\bar{\tau} := \min\{1; -\ln(1-\epsilon)/\ln[\beta_*^m \sup_{\Delta} R\hat{U}_*]\}. \quad (17)$$

(ii) *If in addition $\hat{U}_*$ solves the drift condition (14) then (16) holds with $\bar{\tau}$ replaced by*

$$\hat{\tau} := \min\{1; -\ln(1-\epsilon)/\ln[\beta_*^m \sup_{\Delta} R\hat{U}_*/\hat{U}_*]\}. \quad (18)$$

From Propositions 2 and 3, we are able to derive under the assumptions A1-2, an upper bound for the fluctuations (1) in total variation norm. Define

$$J := \sup_{(x, x') \in \Delta} (V(x) + V(x'))^{-1} \int R(x, x'; dy, dy') (V(y) + V(y')). \quad (19)$$

Then,

$$\begin{aligned} 2(1-\epsilon) J &\leq \sup_{(x, x') \in \Delta} \left[(P^m V(x) + P^m V(x')) - 2\epsilon \int \nu(x, x'; dy) V(y) \right], \\ J &\leq (1-\epsilon)^{-1} \sup_{(x, x') \in \Delta} [(P^m V(x) + P^m V(x')) / 2 - \epsilon], \end{aligned}$$

and J is finite since $\sup_D P^m V$ is finite. Hence,

Corollary 4. *Assume A1-2. For any $(x, x') \in \{V < \infty\}^2$, any $0 \leq \alpha \leq 1$, and $n \geq m-1$,*

$$\|P^n(x, \cdot) - P^n(x', \cdot)\|_{\text{TV}} \leq 2^{1-\alpha \hat{\tau}} \beta_*^{-\alpha \hat{\tau} n} \beta_*^{\alpha \hat{\tau} (m-1)} (V(x) + V(x'))^{\alpha \hat{\tau}},$$

where $\hat{\tau} := \min\{1; -\ln(1-\epsilon)/\ln[\beta_^m J]\}$.*

Corollary 4 corroborates (and improves) assertion (3). It is known that the minimal pointwise solution to the drift inequality (7) is the function $\mathbb{E}_x [\lambda^{-\sigma_C}]$ where $\sigma_C := \inf\{n \geq 0, \Phi_n \in C\}$ is the hitting-time on C of a chain with kernel P . In addition, any solution V to the drift condition is equivalent to the minimal one (Meyn and Tweedie, 1993, Theorem 15.0.1.) *i.e.* there exist constants c_0, c_1 such that

$$c_0 V(x) \leq \mathbb{E}_x [\beta_*^{\sigma_C}] \leq c_1 V(x).$$

Hence, the set of assumptions A1-2 is similar to the assumption that there exists a Kendall small set, and the dependence upon the initial value exhibited in the Corollary is equivalent to the dependence given in (3).

When A2 holds with $C = D$ and C is a ν_1 -small set, Roberts and Tweedie (1999, 2001) gives an explicit expression of a bound $B(x, x'; \beta)$ and a rate β_{RT} such that

$$\limsup_n \beta^n \|P^n(x, \cdot) - P^n(x', \cdot)\|_{\text{TV}} \leq B(x, x'; \beta) < \infty,$$

for all $\beta < \beta_{\text{RT}}$, where β_{RT} coincides with $\beta_*^{\hat{\tau}}$ given in Corollary 4. They establish (Roberts and Tweedie, 1999, Theorem 3.1.) that the rate β_{RT} improves previous rates due to Meyn and Tweedie (1994); Rosenthal (1995). Corollary 4 claims that this maximal rate can be achieved. When $\hat{\tau} = 1$, a similar assertion can be deduced from a recent work by (Rosenthal, 2002, Theorem 1); under a minorization condition on the original chain P and the assumption that there exists a function h that solves (14), it is proved that if $\hat{\tau} = 1$, then $\limsup_n \beta_*^n \|P^n(x, \cdot) - P^n(x', \cdot)\|_{\text{TV}} \leq B(x, x')$ for some explicitly defined function $B(x, x')$. In all of these contributions, the computation of the bound is in two steps, similar to those of the present work. First, a bound is derived under the assumption that there exists a solution $\hat{U}_*$ to the drift condition (14) outside a set $C \times C$ where C has to be a ν_1 -small set for P . The second step is to compute $\hat{U}_*$ from a solution V to the drift condition (7) on the original chain P ; this is done by assuming that $\sup_C V$ is sufficiently large (see e.g. (Roberts and Tweedie, 1999, Eq.(49)) or (Rosenthal, 1995, Theorem 12)). When this condition is violated, it is suggested to increase C which may be no longer ν_1 -small; the stated result can not be directly applied and the calculations have to be adapted in a non-trivial way. To our best knowledge, these extensions have never been carried out. The present work is thus an alternative since the sets C, D are allowed to be ν_m -small, $m > 1$. A numerical illustration is developed in Section 3.1.

While derived under weaker conditions, the rate β_K computed in Kolassa (2000) also suffers from that smallness condition. Nevertheless, when available, β_K may in some cases improve β_{RT} and thus $\beta_*^{\hat{\tau}}$. This is numerically illustrated in Section 3.1.

Remark 5. Computable bounds for geometrical ergodicity As V is finite at some point and P is ψ -irreducible, the drift condition (7) implies that the set $\{V < \infty\}$ is absorbing and thus V is finite ψ a.e. (Meyn and Tweedie, 1993, Proposition 4.2.3.). If there exists an invariant probability measure π for the kernel P such that $\pi(V^{\hat{\tau}}) < \infty$ then $\psi \prec \pi$ and $\pi(\{V < \infty\}) = 1$ (Meyn and Tweedie, 1993, Proposition 10.1.2.). For example, when (7) holds for some $\nu_{m'}$ -small set C , then the chain is positive-recurrent with invariant probability measure π such that $\pi(V) < \infty$ (Meyn and Tweedie, 1993, Theorem 14.0.1.). We thus deduce from Corollary 4 applied with $\alpha = 1$ that for $n \geq m - 1$,

$$\text{for } \pi \text{ a.a. } x, \quad \|P^n(x, \cdot) - \pi(\cdot)\|_{\text{TV}} \leq 2^{1-\hat{\tau}} \beta_*^{-\hat{\tau}n} \beta_*^{\hat{\tau}(m-1)} \int \pi(dx') \left(V(x) + V(x')\right)^{\hat{\tau}} < \infty, \quad (20)$$

where $\hat{\tau}$ is given in Corollary 4.

2.2. Rate of convergence in V -norm. Let V, f be real-valued functions on X . Define the V -norm of a (signed) function f by

$$\|f\|_V := \sup_X \left| \frac{f}{V} \right|.$$

Let β_* be the largest value such that $\mathbb{E}_{x,x'}[\beta^\sigma] < \infty$. If $(1 - \epsilon)\beta_*^m \sup_{\Delta} RU_* \leq 1$, the fluctuations $\|P^n(x, \cdot) - P^n(x', \cdot)\|_V$ are geometrically decreasing with rate β_* where V is any positive function such that

$$\|V(x) + V(x')\|_{U_*} < \infty, \quad \text{and} \quad \sup_{k \in \{1, \dots, m-1\}} \sup_{(x,x') \in \Delta} \{P^k V(x) + P^k V(x')\} < \infty.$$

If $m = 1$, the condition gets into $\|V(x) + V(x')\|_{U_*} < \infty$. If $(1 - \epsilon)\beta_*^m \sup_{\Delta} RU_* > 1$, β_* have to be substituted for β_{opt} defined as solving

$$1 < \beta < \beta_* \quad \text{and} \quad (1 - \epsilon) \beta^m \sup_{(x,x') \in \Delta} \int R(x, x'; dy \times dy') \mathbb{E}_{y,y'}[\beta^\sigma] = 1.$$

By the Jensen's inequality, β_{opt} is lower-bounded by β satisfying (11). This yields Theorem 6 proved in Section 4.2 for $m = 1$ and in Section 5.3 for $m > 1$. For a function U on $\mathcal{X}^2$ and $\tau \in [0, 1]$, define

$$C_1(U, 1) := 0, \\ \text{if } \tau < 1, \quad C_1(U, \tau) := \sup_{\Delta} U \left((1 - \epsilon) \beta_*^m \sup_{\Delta} \{RU/U\} - 1 \right) (\beta_*^{1-\tau} - 1)^{-1}. \quad (21)$$

Theorem 6. Assume A1, A3. Let V be a positive function such that $C_2(V, 1) < \infty$ where

$$C_2(V, \alpha) := \sup \left(\|V(x) + V(x')\|_{U_*^\alpha}; \sup_{1 \leq k \leq m-1} \sup_{(x, x') \in \Delta} \beta_*^{\alpha k} \{P^k V(x) + P^k V(x')\} \right) < \infty. \quad (22)$$

For all (x, x') such that $U_*(x, x') < \infty$, for all $0 \leq \alpha \leq 1$ such that $C_2(V, \alpha) < \infty$, for all $n \geq m-1$,

$$\|P^n(x, \cdot) - P^n(x', \cdot)\|_V \leq C_2(V, \alpha) \left(\beta_*^{-n} U_*(x, x') + \beta_*^{-\tau n} \beta_*^{(\tau-1)(m-1)} C_1(U_*, \tau) U_*^\tau(x, x') \right)^\alpha,$$

where τ and C_1 are given by (13) and (21).

When applied with $\alpha = 1$, Theorem 6 claims that the V -norm of the fluctuations (1) is geometrically decreasing with maximal rate β_*^τ

$$\limsup_n \beta_*^{\tau n} \|P^n(x, \cdot) - P^n(x', \cdot)\|_V \leq C_2(V, 1) \beta_*^{(\tau-1)(m-1)} \left(\mathbb{1}_{\tau=1} + C_1(U_*, \tau) \mathbb{1}_{\tau < 1} \right) U_*^\tau(x, x').$$

The rate does not depend upon the norm involved and results from a balance between ϵ , m and the largest geometrical moment U_* of the hitting-time σ . For a given V -norm, the upper-bound also stresses that the stronger the rate, the stronger the dependence upon the initial value in the bound. For example, let $0 \leq \rho \leq 1$. Then $C_2(V, 1) < \infty$ implies that $C_2(V^\rho, \alpha) < \infty$ for $\rho \leq \alpha \leq 1$. By Theorem 6, there exists $R < \infty$ such that

$$\limsup_n \beta_*^{\alpha \tau n} \|P^n(x, \cdot) - P^n(x', \cdot)\|_{V^\rho} \leq R U_*^{\alpha \tau}(x, x') < \infty,$$

for all $\rho \leq \alpha \leq 1$ and (x, x') such that $U_*(x, x') < \infty$. A similar balance had already been observed for explicit control of the fluctuations in V -norm when decreasing at a polynomial rate (Fort and Moulines (2002)).

If U_* can not be computed in closed form but there exists $\hat{U}_* \geq U_*$, it is proved in Section 4.3.1 for $m = 1$ and in Section 5.4.1 for $m > 1$ that, under the assumptions of Theorem 6, for any $0 \leq \alpha \leq 1$,

$$\|P^n(x, \cdot) - P^n(x', \cdot)\|_V \leq C_2(V, \alpha) \left(\beta_*^{-n} \hat{U}_*(x, x') + \beta_*^{-\bar{\tau} n} \beta_*^{(\bar{\tau}-1)(m-1)} \bar{C}_1(\hat{U}_*, \bar{\tau}) \hat{U}_*^{\bar{\tau}}(x, x') \right)^\alpha, \quad (23)$$

for $\bar{\tau}$ and C_2 given by (17) and (22), and $\bar{C}_1$ defined as

$$\bar{C}_1(U, 1) = 0 \quad \text{and if } \bar{\tau} < 1, \quad \bar{C}_1(U, \bar{\tau}) := \left((1 - \epsilon) \beta_*^m \sup_{\Delta} RU - 1 \right) (\beta_*^{1-\bar{\tau}} - 1)^{-1}.$$

This result is of minor interest since, in many applications a lower-bound for U_* is not available and thus it is quite difficult to determine a function V such that $C_2(V, 1) < \infty$. Nevertheless, if

the upper-bound $\hat{U}_*$ solves the drift condition (14) then Theorem 6 may be modified as follows. The proof is in Section 4.3.2 for $m = 1$ and Section 5.4.2 for $m > 1$.

Proposition 7. *Assume A1. Assume in addition that there exists a function $\hat{U}_*$ such that (14) holds for some $\beta_* > 1$ and $\sup_{\Delta} R\hat{U}_* < \infty$. Let V be a positive function such that $\hat{C}_2(V, 1) < \infty$ where*

$$\hat{C}_2(V, \alpha) := \sup \left(\|V(x) + V(x')\|_{\hat{U}_*}; \sup_{1 \leq k \leq m-1} \sup_{(x, x') \in \Delta} \beta_*^{\alpha k} \{P^k V(x) + P^k V(x')\} / \hat{U}_*(x, x') \right) < \infty.$$

For any (x, x') such that $\hat{U}_(x, x') < \infty$, any $0 \leq \alpha \leq 1$ such that $\hat{C}_2(V, \alpha) < \infty$ and any $n \geq m - 1$,*

$$\|P^n(x, \cdot) - P^n(x', \cdot)\|_V \leq \hat{C}_2(V, \alpha) \left(\beta_*^{-n} \hat{U}_*(x, x') + \beta_*^{-\hat{\tau}n} \beta_*^{(\hat{\tau}-1)(m-1)} C_1(\hat{U}_*, \hat{\tau}) \hat{U}_*^{\hat{\tau}}(x, x') \right)^{\alpha},$$

where $\hat{\tau}$ and C_1 are given by (18) and (21).

As established in Proposition 2, assumption A2 allows the construction of a function $\hat{U}_*(x, x') \sim (V(x) + V(x'))/2$ that solves (14). By using A2, $PV(x) + PV(x') \leq \beta_*^{-1} (V(x) + V(x')) + 2b' \mathbb{I}_{\Delta}(x, x')$ where

$$b' := b \left(1 - \frac{\inf_C V}{\inf_C V + \inf_{D^c} V} \right).$$

By iterating the drift,

$$\sup_{(x, x') \in \Delta} \beta_*^k \{P^k V(x) + P^k V(x')\} / \{V(x) + V(x')\} \leq 1 + \frac{b'}{\inf_C V} \frac{\beta_*^k - 1}{1 - \beta_*^{-1}},$$

showing that $\hat{C}_2(V, 1) < \infty$. Therefore, set

$$C_1(1) := 0 \quad \text{and if } \hat{\tau} < 1, C_1(\hat{\tau}) := \sup_{(x, x') \in \Delta} \{V(x) + V(x')\} ((1 - \epsilon) \beta_*^m J - 1) \left(\beta_*^{1-\hat{\tau}} - 1 \right)^{-1},$$

for J given by (19) and,

$$C_2 := 1 \vee \sup_{1 \leq k \leq m-1} \sup_{(x, x') \in \Delta} \beta_*^k \{P^k V(x) + P^k V(x')\} / \{V(x) + V(x')\}.$$

Corollary 8. *Assume A1-2. Let $\hat{\tau}$ be given in Corollary 4. For all $(x, x') \in \{V < \infty\}^2$, $n \geq m - 1$,*

$$\|P^n(x, \cdot) - P^n(x', \cdot)\|_V \leq C_2 \left[\beta_*^{-n} (V(x) + V(x')) + 2^{-\hat{\tau}} \beta_*^{-\hat{\tau}n} \beta_*^{(\hat{\tau}-1)(m-1)} C_1(\hat{\tau}) (V(x) + V(x'))^{\hat{\tau}} \right].$$

From the discussion above, $1 \leq C_2 \leq 1 + b'(\inf_C V)^{-1}(\beta_*^{m-1} - 1)(1 - \beta_*^{-1})^{-1}$ and if $m = 1$, $C_2 = 1$.

Hence, if $m = 1$, this asserts that when $1 \leq -\ln(1 - \epsilon)/\ln[\beta_* J]$,

$$\limsup_n \beta_*^n \|P^n(x, \cdot) - P^n(x', \cdot)\|_V \leq (V(x) + V(x')),$$

whereas if $1 > -\ln(1 - \epsilon)/\ln[\beta_* J] =: \hat{\tau}$,

$$\limsup_n \beta_*^{\hat{\tau}n} \|P^n(x, \cdot) - P^n(x', \cdot)\|_V \leq 2^{-\hat{\tau}} C_1(\hat{\tau}) (V(x) + V(x'))^{\hat{\tau}}.$$

Corollary 8 improves previous work by the author, work that gives the expression of a constant B such that

$$\sum_n \beta_*^n \|P^n(x, \cdot) - P^n(x', \cdot)\|_V \leq B(x, x'; \beta)$$

for all $\beta < \beta_*^\tau$ (Fort, 2001, Theorem 4.20.). It extends results by Meyn and Tweedie (1994) that successively address the case when the Foster-Lyapunov drift condition holds for some atom and some ν_1 -small set. The rate given in the present paper has a simple expression in terms of the constants appearing in the minorization condition and in the drift condition. It also (explicitly) covers the case when the drift condition holds for P outside a set C which is not necessarily ν_1 -small.

Finally, Corollary 8 corroborates (and improves) assertion (4). Indeed, if V solves the drift condition (7) then for any $\kappa < \lambda^{-1}$, there exists $R < \infty$ such that (Meyn and Tweedie, 1993, Theorem 15.2.5.)

$$V(x) \leq \mathbb{E}_x \left[\sum_{k=0}^{\tau_C-1} \kappa^k V(\Phi_k) \right] \leq R V(x).$$

This shows that (a) our set of conditions and the assumption that there exists a V -Kendall small-set are similar conditions, and (b) that the dependence upon the initial value exhibited in the Corollary is equivalent to the dependence given in (4).

Remark 9. Computable bound for V -geometrical ergodicity If there exists an invariant probability distribution π such that $\pi(V^{\hat{\tau}}) < \infty$ for $\hat{\tau}$ given in Corollary 4, (see Remark 5 for sufficient conditions), then under the assumptions of Corollary 8, $n \geq m - 1$,

$$\text{for } \pi \text{ a.a. } x, \quad \|P^n(x, \cdot) - \pi(\cdot)\|_V \leq \beta_*^{-\hat{\tau}n} B \int \pi(dx') (V(x) + V(x'))^{\hat{\tau}}. \quad (24)$$

The constant B reads into Corollary 8.

3. EXAMPLES

In that section, we denote by β_F the optimal rate $\beta_{\hat{\tau}}^*$ for $\hat{\tau}$ given by Corollary 4.

3.1. A reflecting random walk. Consider the random walk on $\mathbb{Z}_+$ defined by

$$\begin{aligned} P(0,0) &:= \alpha, & P(0,1) &:= 1 - \alpha, \\ \forall n \geq 1, & P(n,n-1) &:= p > 0.5, & P(n,n+1) &:= 1 - p. \end{aligned}$$

We verify (7) with

$$C := \{0,1\} \quad V(x) := \sqrt{p/(1-p)}^{(x-1) \vee 0} \quad \lambda := 2\sqrt{p(1-p)} \quad b := p - \sqrt{p(1-p)}.$$

Denote by m_D the non-negative integer such that $D := \{0, \dots, m_D\}$. By definition, $\beta_* := \lambda + b/(1 + \sqrt{p/(1-p)}^{m_D \vee 0})$ and we choose $m_D \geq m_D^{(\min)} \vee 1$ where

$$m_D^{(\min)} := \left\lceil \frac{2 \ln[b(1-\lambda)^{-1} - 1]}{\ln p - \ln(1-p)} \right\rceil;$$

$\lceil r \rceil$ denotes upper integer part of r . In Table 1, we give for different models the values of λ , b , $m_D^{(\min)}$, m_D and β_*^{-1} (quantities that do not depend upon α). The rate β_{RT} derived in Roberts and Tweedie (1999) can be computed when $C = D$ and the set C is ν_1 -small. In that case, as commented above, $\beta_{RT} = \beta_F$. When $m_D^{(\min)} > 1$, which implies $D \supset C$ and D ν_m -small with $m > 1$, our un-coupled conditions A1-A2 is thus an alternative. To check the condition A1, we set $m := m_D$ and we successively consider two minorizing kernels. First, we set

$$\nu(x, x'; \cdot) := \delta_0(\cdot) \quad \epsilon_1 := \inf_{x \in D} P^m(x, 0); \quad (25)$$

Then we set

$$\nu(x, x'; \cdot) := P^m(x', \cdot) \quad \epsilon_2 := \inf_{(x, x') \in \Delta} \sum_{k=0}^{m_D} \{P^m(x, k) \wedge P^m(x', k)\};$$

ϵ_1, ϵ_2 and the corresponding rate β_F are listed in Tables 2 and 3. It is easily checked (the details are omitted) that for the considered values (α, p) , if $m = 1$ then $\epsilon_1 = \epsilon_2 = \alpha \wedge p$, and if $m = 2$, $\epsilon_1 = \epsilon_2 = \alpha p$.

We compare β_F with the rate β_K given in (Kolassa, 2000, Theorem 2.5.). β_K is computed as suggested in (Kolassa, 2000, Section 3.3.) with $\rho := \sqrt{p/(1-p)}$, and with the uniform minorization condition

$$\nu(\cdot) := \delta_0(\cdot) \quad \epsilon := p \wedge \alpha.$$

In the investigated values, β_F may improve β_K except when p is close to $1/2$.

Original kernel or m -skeleton? For that toy-example, one can easily check a drift condition for the m -iterated kernel when $C \subseteq D$ are ν_m -small sets. We can thus compare the method proposed in the present paper with the method usually suggested in such a case.

For $p = 0.9$ and $\alpha = 0.7$, we establish A2 with $C := \{0, 1\}$, $D := \{0, 1, 2\}$ and the constants $\lambda := 0.6$, $b := 0.6$ and $\beta_* := 1.5151$. A1 holds with $m = 2$, $\nu(x, x'; \cdot)$ and ϵ given by (25), so that $\epsilon := 0.6300$. By Corollary 4, $\limsup_n \beta^n \|P^n(x, \cdot) - P^n(x', \cdot)\|_{\text{TV}} < \infty$ for all $\beta \leq 1.3179$. Nevertheless, since $PV \leq cV$ for some $c < \infty$, there exists $R < \infty$ such that

$$\|P^n(x, \cdot) - P^n(x', \cdot)\|_{\text{TV}} \leq R \|P^{m\lfloor n/m \rfloor}(x, \cdot) - P^{m\lfloor n/m \rfloor}(x', \cdot)\|_{\text{TV}},$$

and this yields a second method to determine the largest rate β_{opt} such that $\limsup_n \beta_{\text{opt}}^n \|P^n(x, \cdot) - P^n(x', \cdot)\|_{\text{TV}} < \infty$. One may compute $\bar{\beta}$ such that $\limsup_n \bar{\beta}^{\lfloor n/2 \rfloor} \|Q^{\lfloor n/2 \rfloor}(x, \cdot) - Q^{\lfloor n/2 \rfloor}(x', \cdot)\|_{\text{TV}} < \infty$, where $Q := P^2$, and then set $\beta_{\text{opt}} := \bar{\beta}^{1/2}$. We establish that $QV \leq \bar{\lambda}V + \bar{b}\mathbb{1}_C$ with $\bar{\lambda} := 0.54$, $\bar{b} := 0.54$ and $C := \{0, 1\}$. Then we find $\bar{\beta}_* := 1.6835$ by choosing $D := \{0, 1, 2\}$. Finally, we check

$$\forall (x, x') \in \Delta, \quad Q(x, \cdot) \wedge Q(x', \cdot) \geq \epsilon \nu(x, x'; \cdot),$$

for (ϵ, ν) as in (25). By Corollary 4, $\bar{\beta}^{-1} := 0.6463$ and thus $\beta_{\text{opt}} = 1.2439$. This numerical illustration stresses the interest of the method developed in the present paper when the minorization condition holds for some $m > 1$. Roughly speaking, this may be explained as follows. As said above, the optimal rate is related to the decreasing rate of the tails of the coupling time T ; in turn, T is related to some *constructive* hitting-time on Δ for two independent Markov chains $(\phi_n; \phi'_n)$ with kernel P , and initial distribution $R(x, x'; \cdot)$ for some $(x, x') \in \Delta$. When $m > 1$, the widely used definition of this constructive hitting-time is the first hitting-time on Δ for the sampled process $(\phi_{km}; \phi'_{km})$, $k \in \mathbb{Z}_+$. In the present paper, the constructive time is the hitting-time following iteration $m - 1$. This makes T shorter and then improves the computed maximal rate β_F .

3.2. Contracting normals. We compare the upper bound $B(\lambda, \lambda'; n)$ given in Corollary 4 for the fluctuation $\|\lambda P^n(\cdot) - \lambda' P^n(\cdot)\|_{\text{TV}}$ with the upper bound (45) given in Roberts and Tweedie (1999, 2001). To that goal, let $P(x, \cdot)$ be the $\mathcal{N}(x/2, 3/4)$ gaussian distribution. As established in (Roberts and Rosenthal, 1997, Example 6.1), $PV(x) = V(x)/4 + 3/2$ where $V(x) := 1 + x^2$. Hence, A2 is verified with

$$V(x) := 1 + x^2, \quad C := \{x, |x| \leq \sqrt{3}\}, \quad \lambda := 5/8, \quad b := 9/8.$$

We choose $D := C$ so that $\beta_*^{-1} := 0.85$. Finally, A1 holds with $m = 1$ and $\epsilon = 0.3173$. This yields $\tau = 0.2881$ and $\beta_F = \beta_{RT} = 1.0479$. To apply the results by Roberts and Tweedie (1999), we verify the assumptions of Theorem 5.2. with

$$\lambda_2 := \lambda + b/(1 + \sup_C V) = \beta_*^{-1} \quad b_2 := b \sup_C V/(1 + \sup_C V) = 0.9.$$

We compute the bound for $n = 45$ and $n = 60$ and assume that $\lambda(V) + \lambda'(V) = 2$. By applying Corollary 4, we obtain

$$B(\lambda, \lambda'; 45) = 0.1485 \quad B(\lambda, \lambda'; 60) = 0.0735,$$

which improves the bound by Roberts and Tweedie (1999) that gives

$$B(\lambda, \lambda'; 45) = 0.5291 \quad B(\lambda, \lambda'; 60) = 0.3435.$$

4. PROOFS: CASE $m = 1$

Throughout this section, we assume that A1 holds with $m = 1$. The minorization condition A1 enables the construction of a Markovian process $Z := (\Omega, \mathcal{A}, \mathcal{F}_n^Z, \{Z_n := (X_n, X'_n, d_n)\}, \mathbb{P}_{x, x', i})$ on $\mathcal{Z}$ described by the following algorithm.

4.1. Construction of an extended process. Initialize $Z_0 = (x, x', d)$ and do recursively

- (1) if $d_n = 0$ and
 - (i) $(X_n, X'_n) \notin \Delta$, then set $d_{n+1} := 0$ and sample independently (conditionally to (X_n, X'_n)) the r.v. X_{n+1} and X'_{n+1} from the distributions $P(X_n, \cdot)$ and $P(X'_n, \cdot)$.
 - (ii) $(X_n, X'_n) \in \Delta$, then
 - with probability ϵ , set $d_{n+1} := 1$, and draw $X_{n+1} = X'_{n+1} \sim \nu(X_n, X'_n; \cdot)$.
 - with probability $1 - \epsilon$, set $d_{n+1} := 0$ and sample independently (conditionally to (X_n, X'_n)), $X_{n+1} \sim R(X_n, X'_n; \cdot)$ and $X'_{n+1} \sim R(X_n, X'_n; \cdot)$.
- (2) if $d_n = 1$, set $d_{n+1} := 1$, draw $X_{n+1} = X'_{n+1} \sim P(X_n, \cdot)$.

It is well-known that the fundamental relation (9) is satisfied with $\tilde{Z} := Z$. Setting $T := \inf\{n \geq 0, d_n = 1\}$, with the convention that $\inf \emptyset = +\infty$, the coupling's inequality states that for any probability distributions λ, λ' on $\mathcal{B}(\mathcal{X})$

$$\|\lambda P^n - \lambda' P^n\|_{TV} \leq \int \lambda(dx) \lambda'(dx') \|P^n(x, \cdot) - P^n(x', \cdot)\|_{TV} \leq 2 \mathbb{P}_{\lambda, \lambda', 0}(T > n) \quad (26)$$

and for any positive function V ,

$$\begin{aligned} \|\lambda P^n - \lambda' P^n\|_V &\leq \int \lambda(dx) \lambda'(dx') \|P^n(x, \cdot) - P^n(x', \cdot)\|_V \\ &\leq \sup_{\{f, |f| \leq V\}} \mathbb{E}_{\lambda, \lambda', 0} [\mathbb{1}_{T > n} |f(X_n) - f(X'_n)|] = \mathbb{E}_{\lambda, \lambda', 0} [\mathbb{1}_{T > n} \{V(X_n) + V(X'_n)\}]. \end{aligned} \quad (27)$$

We derive a bound for the fluctuations by controlling the tails of the coupling time T , which is by definition closely related to the control of moments of some hitting-time on Δ of the Markovian process Z . Let

$$T_0 := \inf\{n \geq 0, (X_n, X'_n) \in \Delta\}$$

be the hitting-time on Δ and

$$T_{j+1} := \inf\{n \geq T_j + 1, (X_n, X'_n) \in \Delta\}, \quad j \geq 0,$$

be the successive return-times to Δ . Then $T_{j+1} = T_0 \circ \theta^{T_j+1}$ where θ denotes the usual shift operator. Any times Z enters the set $\Delta \times \{0\}$, (say at T_j) then $T = T_j + 1$ with probability ϵ . Consequently, the tails of T depend both on the minorization constant ϵ and, by using the strong Markov property, on the tails of T_0 . Such a connection is established in Theorem 10.

4.2. A fundamental result. By construction, the transition kernel P_* of the Markov process Z satisfies for all $(x, x') \in \Delta^c$,

$$P_*(x, x', 0; dy \times dy' \times \{i\}) = P(x, dy) P(x', dy') \delta_0(i).$$

Hence, the geometrical moment $U_*(x, x')$ defined by (6) coincides on Δ^c with $\mathbb{E}_{x, x', 0}[\beta_*^{T_0}]$, the geometrical moment of T_0 for the process Z starting from $(x, x', 0)$. On Δ , both functions have the same value 1. Then

$$U_*(x, x') = \mathbb{E}_{x, x', 0}[\beta_*^{T_0}] \quad (x, x') \in \mathcal{X} \times \mathcal{X}.$$

Theorem 10. *Under A1 and A3,*

(i) *for any $0 \leq \alpha \leq 1$,*

$$\mathbb{P}_{x, x', 0}(T > n) \leq \mathbb{E}_{x, x', 0}[U_*^{\alpha \tau}(X_n, X'_n) \mathbb{1}_{d_n=0}] \leq \beta_*^{-\alpha \tau n} U_*^{\alpha \tau}(x, x'), \quad (28)$$

(ii) *and if $\tau < 1$, for any $0 \leq \alpha \leq 1$,*

$$\mathbb{E}_{x, x', 0}[U_*^\alpha(X_n, X'_n) \mathbb{1}_{d_n=0}] \leq (\beta_*^{-n} U_*(x, x') + \beta_*^{-\tau n} C_1(U_*, \tau) U_*^\tau(x, x'))^\alpha$$

where U_* , τ and C_1 are respectively given by (6), (13) and (21).

Theorems 1 and 6 result from Theorem 10 and (26-27).

Proof. (i) For all $1 \leq \beta \leq \beta_*$, $0 \leq \alpha \leq 1$

$$\mathbb{P}_{x,x',0}(T > n) = \mathbb{P}_{x,x',0}(d_n = 0) \leq \mathbb{E}_{x,x',0}[M_n^\alpha] \leq \mathbb{E}_{x,x',0}[M_n]^\alpha,$$

where $M_n := \{U_*(X_n, X'_n)\}^{\ln \beta / \ln \beta_*} \mathbb{1}_{d_n=0}$. Applying the Jensen's inequality,

$$\begin{aligned} \mathbb{E}_{x,x',0}[M_{n+1} | \mathcal{F}_n^Z] &\leq \beta^{-1} \{U_*(X_n, X'_n)\}^{\ln \beta / \ln \beta_*} \mathbb{1}_{\Delta^c \times \{0\}}(Z_n) \\ &\quad + (1 - \epsilon) \{RU_*(X_n, X'_n)\}^{\ln \beta / \ln \beta_*} \mathbb{1}_{\Delta \times \{0\}}(Z_n) \end{aligned}$$

so, upon noting that $U_* = 1$ on Δ ,

$$\begin{aligned} \mathbb{E}_{x,x',0}[M_{n+1} | \mathcal{F}_n^Z] &\leq \beta^{-1} M_n + \beta^{-1} \left(\beta (1 - \epsilon) \{RU_*(X_n, X'_n)\}^{\ln \beta / \ln \beta_*} - 1 \right) \mathbb{1}_{\Delta \times \{0\}}(Z_n) \\ &\leq \beta^{-1} M_n + \beta^{-1} \left(\beta (1 - \epsilon) \{\sup_{\Delta} RU_*\}^{\ln \beta / \ln \beta_*} - 1 \right) \mathbb{1}_{\Delta \times \{0\}}(Z_n). \end{aligned} \quad (29)$$

If $\beta_* (1 - \epsilon) \sup_{\Delta} RU_* \leq 1$ i.e. $\tau = 1$ then (29) applied with $\beta = \beta_*$ yields $\mathbb{E}_{x,x',0}[M_{n+1} | \mathcal{F}_n^Z] \leq \beta_*^{-1} M_n$. Hence,

$$\mathbb{E}_{x,x',0}[M_n] \leq \beta_*^{-n} \mathbb{E}_{x,x',0}[M_0] = \beta_*^{-n} U_*(x, x').$$

Otherwise, choose β such that $\beta (1 - \epsilon) \{\sup_{\Delta} RU_*\}^{\ln \beta / \ln \beta_*} = 1$ that is $\beta := \beta_*^\tau$; then

$$\mathbb{E}_{x,x',0}[M_n] \leq \beta_*^{-\tau n} \mathbb{E}_{x,x',0}[M_0] = \beta_*^{-\tau n} U_*^\tau(x, x').$$

(ii) Set $C' := (1 - \epsilon) \sup_{\Delta} RU_* - \beta_*^{-1}$; as $\tau < 1$ then $C' > 0$ and by applying (29) with $\beta = \beta_*$,

$$\mathbb{E}_{x,x',0}[M_{n+1}] \leq \beta_*^{-1} \mathbb{E}_{x,x',0}[M_n] + C' \mathbb{P}_{x,x',0}(Z_n \in \Delta \times \{0\}) \leq \beta_*^{-1} \mathbb{E}_{x,x',0}[M_n] + C' \mathbb{P}_{x,x',0}(T > n),$$

where $M_n := U_*(X_n, X'_n) \mathbb{1}_{d_n=0}$. Using the first assertion with $\alpha = 1$

$$\mathbb{E}_{x,x',0}[M_{n+1}] \leq \beta_*^{-1} \mathbb{E}_{x,x',0}[M_n] + C' \beta_*^{-\tau n} U_*^\tau(x, x'),$$

and

$$\begin{aligned} \mathbb{E}_{x,x',0}[M_n] &\leq \beta_*^{-n} U_*(x, x') + C' \sum_{k=0}^{n-1} \beta_*^{-k} \beta_*^{-\tau(n-1-k)} U_*^\tau(x, x') \\ &\leq \beta_*^{-n} U_*(x, x') + \beta_*^{-\tau n} \beta_*^\tau C' (1 - \beta_*^{\tau-1})^{-1} U_*^\tau(x, x'). \end{aligned}$$

Hence, $\mathbb{E}_{x,x',0}[U_*^\alpha(X_n, X'_n) \mathbb{1}_{d_n=0}] \leq \left(\beta_*^{-n} U_*(x, x') + \beta_*^{-\tau(n-1)} (1 - \beta_*^{\tau-1})^{-1} U_*^\tau(x, x') \right)^\alpha$. This yields the desired result upon noting that $U_* = 1$ on Δ . $\square$

Two extensions of Theorem 10 may be proposed. The first one addresses the case when U_* is substituted for an upper-bound $\hat{U}_*$. The second one deals with the existence of a function $\hat{U}_*$ satisfying (14) (and thus also verifying $\hat{U}_* \geq U_*$).

4.3. Two extensions of Theorem 10.

4.3.1. *First extension.* (i) If $\hat{U}_* \geq U_*$, (29) may be replaced with

$$\mathbb{E}_{x,x',0} [M_{n+1} | \mathcal{F}_n^Z] \leq \beta^{-1} M_n + \beta^{-1} \left(\beta (1 - \epsilon) \left\{ \sup_{\Delta} R \hat{U}_* \right\}^{\ln \beta / \ln \beta_*} - 1 \right) \mathbb{1}_{\Delta \times \{0\}}(Z_n).$$

Set $\bar{\tau} := \min\{1; -\ln(1 - \epsilon) / \ln(\beta_* \sup_{\Delta} R \hat{U}_*)\}$. If $(1 - \epsilon) \beta_* \sup_{\Delta} R \hat{U}_* \leq 1$, i.e. $\bar{\tau} = 1$, choose $\beta = \beta_*$; otherwise, choose β such that $\beta (1 - \epsilon) \left\{ \sup_{\Delta} R \hat{U}_* \right\}^{\ln \beta / \ln \beta_*} = 1$. In both cases, $\beta = \beta_*^{\bar{\tau}}$ and

$$\mathbb{E}_{x,x',0} [M_n] \leq \beta_*^{-\bar{\tau}n} \mathbb{E}_{x,x',0} [M_0] \leq \beta_*^{-\bar{\tau}n} \hat{U}_*^{\bar{\tau}}(x, x').$$

Hence,

$$\mathbb{P}_{x,x',0} (T > n) \leq \mathbb{E}_{x,x',0} [U_*^{\alpha \bar{\tau}}(X_n, X'_n) \mathbb{1}_{d_n=0}] \leq \beta_*^{-\alpha \bar{\tau}n} \hat{U}_*^{\alpha \bar{\tau}}(x, x')$$

which yields Proposition 3(i).

(ii) If $\bar{\tau} < 1$, then $\bar{C}' := (1 - \epsilon) \sup_{\Delta} R \hat{U}_* - \beta_*^{-1} > 0$ and as in the proof of Theorem 10(ii), it is established that

$$\mathbb{E}_{x,x',0} [M_{n+1}] \leq \beta_*^{-1} \mathbb{E}_{x,x',0} [M_n] + \bar{C}' \beta_*^{-\bar{\tau}n} U_*^{\bar{\tau}}(x, x') \leq \beta_*^{-1} \mathbb{E}_{x,x',0} [M_n] + \bar{C}' \beta_*^{-\bar{\tau}n} \hat{U}_*^{\bar{\tau}}(x, x')$$

where $M_n := U_*(X_n, X'_n) \mathbb{1}_{d_n=0}$. Thus,

$$\mathbb{E}_{x,x',0} [U_*(X_n, X'_n) \mathbb{1}_{d_n=0}] \leq \beta_*^{-n} \hat{U}_*(x, x') + \beta_*^{-\bar{\tau}n} \left((1 - \epsilon) \beta_* \sup_{\Delta} R \hat{U}_* - 1 \right) (\beta_*^{1-\bar{\tau}} - 1)^{-1} \hat{U}_*^{\bar{\tau}}(x, x'),$$

which yields (23).

4.3.2. *Second extension.* (i) If $\hat{U}_* \geq U_*$ and $\hat{U}_*$ checks the drift condition (14) then the proof of Theorem 10 may be altered by setting $M_n := \{\hat{U}_*(X_n, X'_n)\}^{\ln \beta / \ln \beta_*} \mathbb{1}_{d_n=0}$. By applying the drift inequality outside Δ

$$\mathbb{E}_{x,x',0} [M_{n+1} | \mathcal{F}_n^Z] \mathbb{1}_{\Delta^c \times \{0\}}(Z_n) \leq \beta^{-1} M_n \mathbb{1}_{\Delta^c \times \{0\}}(Z_n),$$

and

$$\mathbb{E}_{x,x',0} [M_{n+1} | \mathcal{F}_n^Z] \mathbb{1}_{\Delta \times \{0\}}(Z_n) \leq (1 - \epsilon) \{\hat{U}_*(X_n, X'_n)\}^{\ln \beta / \ln \beta_*} \mathbb{1}_{\Delta \times \{0\}}(Z_n).$$

(29) is thus replaced with

$$\mathbb{E}_{x,x',0} [M_{n+1} | \mathcal{F}_n^Z] \leq \beta^{-1} M_n + \beta^{-1} \left(\beta (1-\epsilon) \left\{ \sup_{\Delta} R \hat{U}_* / \hat{U}_* \right\}^{\ln \beta / \ln \beta_*} - 1 \right) \{ \hat{U}_*(X_n, X'_n) \}^{\ln \beta / \ln \beta_*} \mathbb{1}_{\Delta \times \{0\}}(Z_n).$$

Set $\hat{\tau} := \min\{1; -\ln(1-\epsilon)/\ln[\beta_* \sup_{\Delta} R \hat{U}_* / \hat{U}_*]\}$. As in the proof of Theorem 10, by choosing $\beta = \beta_*^{\hat{\tau}}$, $\mathbb{E}_{x,x',0} [M_n] \leq \beta_*^{-\hat{\tau}n} \hat{U}_*^{\hat{\tau}}(x, x')$. Hence,

$$\mathbb{P}_{x,x',0}(T > n) \leq \mathbb{E}_{x,x',0} [\hat{U}_*^{\alpha \hat{\tau}}(X_n, X'_n) \mathbb{1}_{d_n=0}] \leq \beta_*^{-\alpha \hat{\tau}n} \hat{U}_*^{\alpha \hat{\tau}}(x, x'),$$

which yields Proposition 3(ii).

(ii) If $\hat{\tau} < 1$, then $\hat{C}' := \sup_{\Delta} \hat{U}_* \left((1-\epsilon) \sup_{\Delta} \{R \hat{U}_* / \hat{U}_* - \beta_*^{-1}\} \right) > 0$. As above, it is easily seen that

$$\mathbb{E}_{x,x',0} [\hat{U}_*(X_n, X'_n) \mathbb{1}_{d_n=0}] \leq \beta_*^{-1} \mathbb{E}_{x,x',0} [\hat{U}_*(X_{n-1}, X'_{n-1}) \mathbb{1}_{d_{n-1}=0}] + \beta_*^{-\hat{\tau}n} \hat{C}' \hat{U}_*^{\hat{\tau}}(x, x'),$$

which yields Proposition 7.

4.4. Proof of Proposition 2. Let $(x, x') \in \Delta^c \cap \{V < \infty\}^2$.

$$\begin{aligned} \mathcal{I}(x, x') &:= 2 \int P(x, dy) P(x', dy') \hat{U}_*(y, y') \\ &\leq PV(x) + PV(x') \leq \lambda (V(x) + V(x')) + b (\mathbb{1}_C(x) + \mathbb{1}_C(x')). \end{aligned}$$

As $(x, x') \in \Delta^c$, $x \in C$ implies $x' \in D^c$ and $x' \in C$ implies $x \in D^c$. Then

$$\mathcal{I}(x, x') (V(x) + V(x'))^{-1} \leq \lambda + b \left\{ \mathbb{1}_{C \times D^c}(x, x') / (V(x) + V(x')) + \mathbb{1}_{D^c \times C}(x, x') / (V(x) + V(x')) \right\}.$$

As $V(x) + V(x') \geq \inf_C V + \inf_{D^c} V$ on $(C \times D^c) \cup (D^c \times C)$, then

$$\mathcal{I}(x, x') \leq \beta_*^{-1} (V(x) + V(x')), \quad (x, x') \in \Delta^c \cap \{V < \infty\}^2.$$

This remains true on Δ^c , consequently $(P \otimes P) \hat{U}_* \leq \beta^{-1} \hat{U}_*$ on Δ^c .

This inequality shows that the hitting-time σ is finite $\mathbb{P}_{x,x'}$ -a.s. for all $(x, x') \in \Delta^c \cap \{V < \infty\}^2$ (Meyn and Tweedie, 1993, Proposition 11.3.2.). On Δ , $\hat{U}_* = U_* = 1$ and on Δ^c ,

$$\begin{aligned} \hat{U}_*(x, x') &\geq \beta_* \int_{\Delta} (P \otimes P)(x, x'; dy \times dy') \times 1 + \beta_* \int_{\Delta^c} (P \otimes P)(x, x'; dy \times dy') \hat{U}_*(y, y') \\ &\geq \beta_* \mathbb{P}_{x,x'}(\sigma = 1) + \beta_*^2 \int_{\Delta^c} (P \otimes P)(x, x'; dy \times dy') (P \otimes P) \hat{U}_*(y, y'). \end{aligned}$$

By proceeding iteratively, for all $n \geq 1$,

$$\hat{U}_*(x, x') \geq \sum_{k=1}^n \beta_*^k \mathbb{P}_{x,x'}(\sigma = k) \rightarrow_n U_*(x, x'),$$

for all $(x, x') \in \Delta^c \cap \{V < \infty\}^2$, and therefore on Δ^c .

5. PROOFS: CASE $m > 1$

Throughout this section, it is assumed that A1 holds for some $m > 1$.

In that case, the fundamental property (9) is not satisfied for the process Z defined in Section 4. We have to define a companion process $\tilde{Z} := \{\tilde{Z}_n\}$, which is no longer Markovian but for which the property holds. The algorithm, also described in Fort and Moulines (2002), is included here for the sake of completeness.

5.1. Construction of a companion process. Let $\{Z_n\}$ be the extended process defined as in Section 4.1. In words, the use of the residual kernel introduces a jump of size $(m - 1)$ on the time-scale implying a kind of time distortion that has to be corrected.

Set $z_{m_1:m_2} := (z_{m_1}, \dots, z_{m_2})$ with the convention that $z_{m_1:m_2} = \emptyset$ if $m_1 > m_2$. Set $Q_0 := 0$ and $\tilde{Z}_0 := Z_0$. Do recursively

- (1) if $d_n = 0$,
 - (i) if $(X_n, X'_n) \notin \Delta$, set $Q_{n+1} := Q_n + 1$ and $\tilde{Z}_{Q_{n+1}} := Z_{n+1}$.
 - (ii) if $(X_n, X'_n) \in \Delta$, set $Q_{n+1} := Q_n + m$ and $\tilde{Z}_{Q_{n+1}} := Z_{n+1}$. Sample $\{\tilde{X}_{Q_{n+k}}\}$, for $k \in \{1, \dots, (m-1)\}$ from the distribution of $P(\Phi_1 \in \cdot, \dots, \Phi_{m-1} \in \cdot | \Phi_m = X_{n+1}, \Phi_0 = X_n)$; sample similarly and independently $\{\tilde{X}'_{Q_{n+k}}\}$, $k \in \{1, \dots, (m-1)\}$. Finally, set $\tilde{d}_{Q_{n+k}} := \tilde{d}_{Q_n}$, $k \in \{1, \dots, (m-1)\}$.
- (2) if $d_n = 1$, set $Q_{n+1} := Q_n + 1$ and $\tilde{Z}_{Q_{n+1}} := Z_{n+1}$.

In words, every time $Z \in \Delta \times \{0\}$, a gap of size $(m - 1)$ is inserted and filled in conditionally independently from the conditional distribution of the Markov chain given the initial and final values. Define the coupling-time T by $T := \inf\{n \geq m, \tilde{d}_n = 0\}$. Then

$$\|\lambda P^n - \lambda' P^n\|_{\text{TV}} \leq 2 \tilde{\mathbb{P}}_{\lambda, \lambda', 0}(T > n) \quad (30)$$

and for any positive function V ,

$$\|\lambda P^n - \lambda' P^n\|_V \leq \tilde{\mathbb{E}}_{\lambda, \lambda', 0} \left[\mathbb{1}_{T > n} \{V(\tilde{X}_n) + V(\tilde{X}'_n)\} \right], \quad (31)$$

where $\tilde{\mathbb{P}}_{x,x',0}$ (resp. $\tilde{\mathbb{E}}_{x,x',0}$) denotes the probability (resp. expectation) on the canonical space associated to the process $\tilde{Z}$, given that $\tilde{Z}_0 = (x, x', 0)$. When $m > 1$, the tails of the coupling-time T are closely related to the tails of some "constructive" hitting-time on $\Delta \times \{0\}$. Because of the inserted samples, if $\tilde{Z}_n \in \Delta \times \{0\}$, this does not imply that with probability ϵ , time $n + m$ is the coupling time. The hitting-times *of interest* are given by $T'_j := T_j + j(m - 1)$ (so that $T'_0 = T_0$), and in order to distinguish these constructive random times, we introduce a *random* counter c_n that counts at time n , the elapsed time from the last random time T'_j . Set $T'_{-1} := -m$ and define a $\mathcal{F}_n^{\tilde{Z}}$ -adapted process $\{c_n\}$ by

$$c_n := j \quad \text{if} \quad T'_k \leq T'_k + j = n < T'_{k+1} \quad \text{for some } k \geq -1,$$

so that for any $n \geq 0$, $c_n = 0$ iff $n = T'_k$ for some $k \geq 0$. The coupling-time T is thus related to the hitting-time on the set $\Delta \times \{0\} \times \{0\}$ of the $\mathcal{Z} \times \mathbb{Z}_+$ -valued process $\{\bar{Z}_n := (\tilde{Z}_n, c_n)\}$.

5.2. A fundamental result. Define for $n \geq m - 1$,

$$\mathbb{U}_*(\bar{Z}_{n+1-m:n}) := \beta_*^{m-1} U_*(\tilde{X}_n, \tilde{X}'_n) (\mathbb{1}_{c_n=0} + \mathbb{1}_{c_n \geq m}) + \sum_{k=1}^{m-1} \beta_*^{m-1-k} U_*(\tilde{X}_{n-k}, \tilde{X}'_{n-k}) \mathbb{1}_{c_n=k},$$

which in words, can be interpreted as the geometrical moment of the hitting-time on $\mathcal{X}^2 \times \{0, 1\} \times \{m - 1\}$ for the process $\bar{Z}$ with initial value $\bar{Z}_n$; remind that on $\{c_n = k\}$, $0 \leq k \leq m - 1$, $(\tilde{X}_{n-k}, \tilde{X}'_{n-k}) \in \Delta$ and $U_*(\tilde{X}_{n-k}, \tilde{X}'_{n-k}) = 1$.

Theorem 11. *Assume A1, A3. For all $n \geq m - 1$,*

(i) *and any $0 \leq \alpha \leq 1$,*

$$\tilde{\mathbb{P}}_{x,x',0}(T > n) \leq \tilde{\mathbb{E}}_{x,x',0} \left[\mathbb{U}_*^{\alpha\tau}(\bar{Z}_{n+1-m:n}) \mathbb{1}_{\tilde{d}_n=0} \right] \leq \beta_*^{-\alpha\tau n} \beta_*^{\alpha\tau(m-1)} U_*^{\alpha\tau}(x, x').$$

(ii) *and if $\tau < 1$, for any $0 \leq \alpha \leq 1$,*

$$\tilde{\mathbb{E}}_{x,x',0} \left[\mathbb{U}_*^\alpha(\bar{Z}_{n+1-m:n}) \mathbb{1}_{\tilde{d}_n=0} \right] \leq \left(\beta_*^{-n} \beta_*^{m-1} U_*(x, x') + \beta_*^{-\tau n} \beta_*^{\tau(m-1)} C_1(U_*, \tau) U_*^\tau(x, x') \right)^\alpha,$$

where U_* , τ and C_1 are given by (6), (13) and (21).

Proof. (i) For all $1 \leq \beta \leq \beta_*$, $0 \leq \alpha \leq 1$, $n \geq m - 1$,

$$\tilde{\mathbb{P}}_{x,x',0}(T > n) = \tilde{\mathbb{P}}_{x,x',0}(\tilde{d}_n = 0) \leq \tilde{\mathbb{E}}_{x,x',0}[M_n^\alpha] \leq \tilde{\mathbb{E}}_{x,x',0}[M_n]^\alpha,$$

where $M_n := \{\mathbb{U}_*(\bar{Z}_{n+1-m:n})\}^{\ln \beta / \ln \beta_*} \mathbb{I}_{\tilde{d}_n=0}$. For $n \geq m-1$,

$$\begin{aligned}
\tilde{\mathbb{E}}_{x,x',0}[M_{n+1} \mathbb{I}_{c_n=m-1}] &= \tilde{\mathbb{E}}_{x,x',0}[M_{n+1} \mathbb{I}_{c_n=m-1, c_{n+1} \in \{0,m\}}] \\
&= \beta^{m-1} \tilde{\mathbb{E}}_{x,x',0} \left[\mathbb{I}_{\tilde{d}_{n+1}=0} \{U_*(\tilde{X}_{n+1}, \tilde{X}'_{n+1})\}^{\ln \beta / \ln \beta_*} \mathbb{I}_{c_{n+1-m}=0} \right] \\
&= \beta^{m-1} \tilde{\mathbb{E}}_{x,x',0} \left[\mathbb{I}_{\tilde{d}_{n+1-m}=0, c_{n+1-m}=0} \tilde{\mathbb{E}}_{x,x',0} \left[\{U_*(\tilde{X}_{n+1}, \tilde{X}'_{n+1})\}^{\ln \beta / \ln \beta_*} | \mathcal{F}_{n+1-m}^{\tilde{Z}} \right] \right] \\
&= \beta^{m-1} \tilde{\mathbb{E}}_{x,x',0} \left[\mathbb{I}_{\tilde{d}_{n+1-m}=0, c_{n+1-m}=0} (1-\epsilon) \{RU_*(\tilde{X}_{n+1-m}, \tilde{X}'_{n+1-m})\}^{\ln \beta / \ln \beta_*} \right] \\
&\leq \tilde{\mathbb{E}}_{x,x',0} \left[\mathbb{I}_{\tilde{d}_n=0, c_n=m-1} (1-\epsilon) \beta^{m-1} \{RU_*(\tilde{X}_{n+1-m}, \tilde{X}'_{n+1-m})\}^{\ln \beta / \ln \beta_*} \right]. \quad (32)
\end{aligned}$$

Similarly, upon noting that on $\{c_n \geq m\}$, $(\tilde{X}_n, \tilde{X}'_n) \notin \Delta$,

$$\begin{aligned}
\tilde{\mathbb{E}}_{x,x',0}[M_{n+1} \mathbb{I}_{c_n \geq m}] &= \tilde{\mathbb{E}}_{x,x',0}[M_{n+1} \mathbb{I}_{c_n \geq m, c_{n+1} \in \{0,m\}}] \\
&= \tilde{\mathbb{E}}_{x,x',0} \left[\mathbb{I}_{\tilde{d}_n=0, c_n \geq m} \tilde{\mathbb{E}}_{x,x',0} \left[\beta^{m-1} \{U_*(\tilde{X}_{n+1}, \tilde{X}'_{n+1})\}^{\ln \beta / \ln \beta_*} | \mathcal{F}_n^{\tilde{Z}} \right] \right] \\
&\leq \beta^{-1} \tilde{\mathbb{E}}_{x,x',0} \left[\mathbb{I}_{\tilde{d}_n=0, c_n \geq m} \beta^{m-1} \{U_*(\tilde{X}_n, \tilde{X}'_n)\}^{\ln \beta / \ln \beta_*} \right] = \beta^{-1} \tilde{\mathbb{E}}_{x,x',0}[M_n \mathbb{I}_{c_n \geq m}]. \quad (33)
\end{aligned}$$

For $k \in \{0, \dots, m-2\}$,

$$\begin{aligned}
\tilde{\mathbb{E}}_{x,x',0}[M_{n+1} \mathbb{I}_{c_n=k}] &= \tilde{\mathbb{E}}_{x,x',0}[M_{n+1} \mathbb{I}_{c_n=k, c_{n+1}=k+1}] \\
&= \tilde{\mathbb{E}}_{x,x',0} \left[\mathbb{I}_{\tilde{d}_{n+1}=0, c_n=k} \beta^{m-1-(k+1)} \{U_*(\tilde{X}_{n+1-(k+1)}, \tilde{X}'_{n+1-(k+1)})\}^{\ln \beta / \ln \beta_*} \right] \\
&= \beta^{-1} \tilde{\mathbb{E}}_{x,x',0} \left[\mathbb{I}_{\tilde{d}_n=0, c_n=k} \beta^{m-1-k} \{U_*(\tilde{X}_{n-k}, \tilde{X}'_{n-k})\}^{\ln \beta / \ln \beta_*} \right] = \beta^{-1} \tilde{\mathbb{E}}_{x,x',0}[M_n \mathbb{I}_{c_n=k}]. \quad (34)
\end{aligned}$$

From (32) to (34), for all $n \geq m-1$,

$$\tilde{\mathbb{E}}_{x,x',0}[M_{n+1}] \leq \beta^{-1} \tilde{\mathbb{E}}_{x,x',0}[M_n] + \beta^{-1} \tilde{\mathbb{E}}_{x,x',0} \left[\mathbb{I}_{\tilde{d}_n=0, c_{n+1-m}=0} \rho_{n+1} \right] \quad (35)$$

where, upon noting that $\bar{U}_*(\tilde{X}_{n+1-m}, \tilde{X}'_{n+1-m}) = 1$ on $\{c_{n+1-m} = 0\}$,

$$\begin{aligned}
\rho_{n+1} &:= (1-\epsilon) \beta^m \{RU_*(\tilde{X}_{n+1-m}, \tilde{X}'_{n+1-m})\}^{\ln \beta / \ln \beta_*} - U_*(\tilde{X}_{n+1-m}, \tilde{X}'_{n+1-m})^{\ln \beta / \ln \beta_*} \\
&= (1-\epsilon) \beta^m \{RU_*(\tilde{X}_{n+1-m}, \tilde{X}'_{n+1-m})\}^{\ln \beta / \ln \beta_*} - 1. \quad (36)
\end{aligned}$$

For $n \geq m-1$, by choosing $\beta := \beta_*^\tau$, $\rho_{n+1} \leq 0$. Hence

$$\tilde{\mathbb{E}}_{x,x',0}[M_n] \leq \beta_*^{-\tau n} \beta_*^{\tau(m-1)} \tilde{\mathbb{E}}_{x,x',0}[M_{m-1}].$$

We now prove that for $\beta = \beta_*^\tau$, $\tilde{\mathbb{E}}_{x,x',0}[M_{m-1}] \leq U_*^\tau(x, x')$. By construction of the process $\tilde{Z}$, $\tilde{\mathbb{P}}_{x,x',0}(\tilde{d}_{m-1} = 0) = \tilde{\mathbb{P}}_{x,x',0}(c_0 = 0) = 1$ for all (x, x') . If $(x, x') \in \Delta$, then $\tilde{\mathbb{P}}_{x,x',0}(c_{m-1} = m-1) =$

1 and $\tilde{\mathbb{E}}_{x,x',0}[M_{m-1}] = U_*^\tau(x, x')$ on Δ . Otherwise,

$$\begin{aligned} \tilde{\mathbb{E}}_{x,x',0}[M_{m-1} \mathbb{1}_{c_{m-1} \neq m-1}] &= \beta^{m-1} \tilde{\mathbb{E}}_{x,x',0} \left[\mathbb{1}_{c_{m-2} \geq m} \{U_*(\tilde{X}_{m-1}, \tilde{X}'_{m-1})\}^{\ln \beta / \ln \beta_*} \right] \\ &+ \sum_{k=0}^{m-3} \beta^{m-1-(k+1)} \tilde{\mathbb{E}}_{x,x',0} \left[\mathbb{1}_{c_{m-2}=k} \{U_*(\tilde{X}_{m-1-(k+1)}, \tilde{X}'_{m-1-(k+1)})\}^{\ln \beta / \ln \beta_*} \right] \\ &\leq \beta^{-1} \tilde{\mathbb{E}}_{x,x',0} \left[\mathbb{1}_{c_{m-2} \geq m} \beta^{m-1} \{U_*(\tilde{X}_{m-2}, \tilde{X}'_{m-2})\}^{\ln \beta / \ln \beta_*} \right] \\ &+ \beta^{-1} \sum_{k=0}^{m-3} \tilde{\mathbb{E}}_{x,x',0} \left[\mathbb{1}_{c_{m-2}=k} \beta^{m-1-k} \{U_*(\tilde{X}_{m-2-k}, \tilde{X}'_{m-2-k})\}^{\ln \beta / \ln \beta_*} \right] \\ &\leq \beta^{-1} \tilde{\mathbb{E}}_{x,x',0}[M_{m-2} \mathbb{1}_{c_{m-1} \neq m-1}] = \beta^{-1} \tilde{\mathbb{E}}_{x,x',0}[M_{m-2} \mathbb{1}_{c_{m-2} \neq m-2}]. \end{aligned}$$

By a trivial recursion, as $\beta = \beta_*^\tau$

$$\tilde{\mathbb{E}}_{x,x',0}[M_{m-1} \mathbb{1}_{c_{m-1} \neq m-1}] \leq \beta_*^{-\tau(m-1)} \tilde{\mathbb{E}}_{x,x',0}[M_0 \mathbb{1}_{c_0 \neq 0}] = \mathbb{1}_{\Delta^c}(x, x') U_*^\tau(x, x'),$$

upon noting that $\tilde{\mathbb{P}}_{x,x',0}(c_0 \in \{0, m\}) = 1$. Then

$$\tilde{\mathbb{E}}_{x,x',0}[M_n] \leq \beta_*^{-\tau n} \beta_*^{\tau(m-1)} U_*^\tau(x, x'). \quad (37)$$

(ii) If $\tau < 1$, then $C' := (1 - \epsilon) \beta_*^{m-1} \sup_\Delta R U_* - \beta_*^{-1} > 0$. Hence, by applying (35) with $\beta = \beta_*$ and upon noting that $\beta_*^{-1} \rho_{n+1} \leq C'$,

$$\tilde{\mathbb{E}}_{x,x',0}[M_{n+1}] \leq \beta_*^{-1} \tilde{\mathbb{E}}_{x,x',0}[M_n] + C' \tilde{\mathbb{P}}_{x,x',0}(\tilde{d}_n = 0, c_{n+1-m} = 0)$$

where $M_n := \mathbb{U}_*(\bar{Z}_{n+1-m:n}) \mathbb{1}_{\tilde{d}_n=0}$. By using the first assertion of the Theorem, this yields that for all $n \geq m-1$,

$$\tilde{\mathbb{E}}_{x,x',0}[M_{n+1}] \leq \beta_*^{-1} \tilde{\mathbb{E}}_{x,x',0}[M_n] + C' \tilde{\mathbb{P}}_{x,x',0}(T > n).$$

As above, it may be proved that $\tilde{\mathbb{E}}_{x,x',0}[M_{m-1}] \leq U_*(x, x')$. Hence,

$$\begin{aligned} \tilde{\mathbb{E}}_{x,x',0}[M_n] &\leq \beta_*^{-n+m-1} \tilde{\mathbb{E}}_{x,x',0}[M_{m-1}] + C' \beta_*^{\tau(m-1)} \sum_{k=0}^{n-m} \beta_*^k \beta_*^{-\tau(n-k-1)} U_*^\tau(x, x'), \\ &\leq \beta_*^{-n+m-1} U_*(x, x') + \beta_*^{-\tau(n-1)} C' \beta_*^{\tau(m-1)} (1 - \beta_*^{\tau-1})^{-1} U_*^\tau(x, x'), \end{aligned}$$

which yields the desired result. $\square$

5.3. Proofs of Theorems 1 and 6. For any positive function V , for all $1 \leq k \leq m-1$,

$$\tilde{\mathbb{E}}_{x,x',0} \left[\mathbb{1}_{\tilde{d}_n=0, c_n=k} \{V(\tilde{X}_n) + V(\tilde{X}'_n)\} \right] = \tilde{\mathbb{E}}_{x,x',0} \left[\mathbb{1}_{\tilde{d}_n=0, c_n=k} \{P^k V(\tilde{X}_{n-k}) + P^k V(\tilde{X}'_{n-k})\} \right].$$

Hence, upon noting that $U_*^\alpha(\tilde{X}_{n-k}, \tilde{X}'_{n-k}) = 1$ on $\{c_n = k\}$ for all $0 \leq \alpha \leq 1$,

$$\begin{aligned} & \tilde{\mathbb{E}}_{x,x',0} \left[\mathbb{1}_{\tilde{d}_n=0} \{V(\tilde{X}_n) + V(\tilde{X}'_n)\} \right] \\ & \leq \beta_*^{\alpha(1-m)} \|V(x) + V(x')\|_{U_*^\alpha} \tilde{\mathbb{E}}_{x,x',0} \left[\left(\mathbb{1}_{\tilde{d}_n=0, c_n=0} + \mathbb{1}_{\tilde{d}_n=0, c_n \geq m} \right) \beta_*^{\alpha(m-1)} U_*^\alpha(\tilde{X}_n, \tilde{X}'_n) \right] \\ & + \beta_*^{\alpha(1-m)} \sum_{k=1}^{m-1} \beta_*^{\alpha k} \sup_{\Delta} \{P^k V(x) + P^k V(x')\} \tilde{\mathbb{E}}_{x,x',0} \left[\mathbb{1}_{\tilde{d}_n=0, c_n=k} \beta_*^{\alpha(m-1-k)} U_*^\alpha(\tilde{X}_{n-k}, \tilde{X}'_{n-k}) \right] \end{aligned} \quad (38)$$

and

$$\tilde{\mathbb{E}}_{x,x',0} \left[\mathbb{1}_{\tilde{d}_n=0} \{V(\tilde{X}_n) + V(\tilde{X}'_n)\} \right] \leq \beta_*^{\alpha(1-m)} C_2(V, \alpha) \tilde{\mathbb{E}}_{x,x',0} \left[\mathbb{U}_*^\alpha(\bar{Z}_{n+1-m:n}) \mathbb{1}_{\tilde{d}_n=0} \right].$$

where C_2 is given by (22).

If $V = 1$, by choosing $\alpha = \tau$,

$$\|P^n(x, \cdot) - P^n(x', \cdot)\|_{\text{TV}} \leq \beta_*^{\tau(1-m)} \left(2 \vee 2\beta_*^{\tau(m-1)} \right) \tilde{\mathbb{E}}_{x,x',0} \left[\mathbb{U}_*^\tau(\bar{Z}_{n+1-m:n}) \mathbb{1}_{\tilde{d}_n=0} \right],$$

and by Theorem 11(i), $\|P^n(x, \cdot) - P^n(x', \cdot)\|_{\text{TV}} \leq 2\beta_*^{-\tau n} \beta_*^{\tau(m-1)} U_*^\tau(x, x')$. This establishes Theorem 1.

For any function V ,

$$\|P^n(x, \cdot) - P^n(x', \cdot)\|_V \leq \beta_*^{\alpha(1-m)} C_2(V, \alpha) \tilde{\mathbb{E}}_{x,x',0} \left[\mathbb{U}_*^\alpha(\bar{Z}_{n+1-m:n}) \mathbb{1}_{\tilde{d}_n=0} \right]$$

If $\tau = 1$, by Theorem 11(i), $\|P^n(x, \cdot) - P^n(x', \cdot)\|_V \leq \beta_*^{-\alpha n} C_2(V, \alpha) U_*^\alpha(x, x')$; if $\tau < 1$, by Theorem 11(ii),

$$\|P^n(x, \cdot) - P^n(x', \cdot)\|_V \leq C_2(V, \alpha) \left(\beta_*^{-n} U_*(x, x') + \beta_*^{-\tau n} \beta_*^{(\tau-1)(m-1)} C_1(U_*, \tau) U_*^\tau(x, x') \right)^\alpha.$$

Theorem 6 results from these inequalities.

Remark 12. The upper bound given in Theorem 1 can be improved when the whole state space is small.

By the contraction property, for all $k \leq n$, $\|P^n(x, \cdot) - P^n(x', \cdot)\|_{\text{TV}} \leq \|P^{n-k}(x, \cdot) - P^{n-k}(x', \cdot)\|_{\text{TV}}$.

If the state space is ν_m -small, this implies that (a)

$$\|P^n(x, \cdot) - P^n(x', \cdot)\|_{\text{TV}} \leq \|P^{\lfloor n/m \rfloor m}(x, \cdot) - P^{\lfloor n/m \rfloor m}(x', \cdot)\|_{\text{TV}},$$

and (b) $\tilde{\mathbb{P}}_{x,x',0}(c_{qm} = 0) = 1$ for all $q \in \mathbb{Z}_+$, and for all (x, x') . Hence, in the above calculations (38) applied with $V = 1$ and $n = qm$, the cases $\{c_{qm} = k\}$ for $k \in \{1, \dots, m-1\}$ have to be omitted. This implies that

$$\begin{aligned} \|P^n(x, \cdot) - P^n(x', \cdot)\|_{\text{TV}} &\leq \|P^{\lfloor n/m \rfloor m}(x, \cdot) - P^{\lfloor n/m \rfloor m}(x', \cdot)\|_{\text{TV}} \\ &\leq 2\beta_*^{\tau(1-m)} \tilde{\mathbb{E}}_{x,x',0} \left[\mathbb{U}_*^\tau(\bar{Z}_{\lfloor n/m \rfloor m+1-m:\lfloor n/m \rfloor m}) \mathbb{1}_{\bar{d}_{\lfloor n/m \rfloor m}=0} \right] \\ &\leq 2\beta_*^{\tau(1-m)} \beta_*^{-\tau \lfloor n/m \rfloor m} \beta_*^{\tau(m-1)} U_*^\tau(x, x') \\ &\leq 2 \beta_*^{-\tau \lfloor n/m \rfloor m}. \end{aligned}$$

5.4. Two extensions of Theorem 11.

5.4.1. First extension. If there exists $\hat{U}_* \geq U_*$, then U_* may be substituted for $\hat{U}_*$ in (36) and (37). In that case, τ has to be replaced by $\bar{\tau}$, defined by (17). This yields Proposition 3(i).

The proof of Theorem 11(ii) may be easily modified (the modification is on the same line as in Section 4.3.1 and the details are omitted). Then

$$\tilde{\mathbb{E}}_{x,x',0} \left[\mathbb{U}_*(\tilde{X}_n, \tilde{X}'_n) \mathbb{1}_{\bar{d}_n=0} \right] \leq \beta_*^{-(n+1-m)} \hat{U}_*(x, x') + \bar{C}' \beta_*^{\bar{\tau}(m-1)} \beta_*^{-\bar{\tau}(n-1)} (1 - \beta_*^{\bar{\tau}-1})^{-1} \hat{U}_*^{\bar{\tau}}(x, x'),$$

where $\bar{C}' := ((1 - \epsilon)\beta_*^{m-1} \sup_{\Delta} R\hat{U}_* - \beta_*^{-1}) > 0$. The discussion of Section 5.3 extends easily, and this yields Eq.(23).

5.4.2. Second extension. If $\hat{U}_* \geq U_*$ and $\hat{U}_*$ checks the drift condition (14), the proof of Theorem 11(i) may be altered by substituting U_* for $\hat{U}_*$ in the definition of $\mathbb{U}_*$. Therefore,

$$\tilde{\mathbb{E}}_{x,x',0} \left[\hat{\mathbb{U}}_*(\bar{Z}_{n+1-m:n}) \mathbb{1}_{\bar{d}_n=0} \right] \leq \beta_*^{-\hat{\tau}n} \beta_*^{(m-1)\hat{\tau}} \hat{U}_*^{\hat{\tau}}(x, x'),$$

where $\hat{\tau}$ is given by (18) and

$$\hat{\mathbb{U}}_*(\bar{Z}_{n+1-m:n}) := \beta_*^{m-1} \hat{U}_*(\tilde{X}_n, \tilde{X}'_n) (\mathbb{1}_{c_n=0} + \mathbb{1}_{c_n \geq m}) + \sum_{k=1}^{m-1} \beta_*^{m-1-k} \hat{U}_*(\tilde{X}_{n-k}, \tilde{X}'_{n-k}) \mathbb{1}_{c_n=k}.$$

This yields Proposition 3(ii). Theorem 11(ii) gets into

$$\tilde{\mathbb{E}}_{x,x',0} \left[\hat{\mathbb{U}}_*(\tilde{X}_n, \tilde{X}'_n) \mathbb{1}_{\bar{d}_n=0} \right] \leq \beta_*^{-n+m-1} \hat{U}_*(x, x') + \hat{C}' \beta_*^{\hat{\tau}(m-1)} \beta_*^{-\hat{\tau}(n-1)} (1 - \beta_*^{\hat{\tau}-1})^{-1} \hat{U}_*^{\hat{\tau}}(x, x'),$$

where $\hat{C}' := (\sup_{\Delta} \hat{U}_*) \left((1 - \epsilon)\beta_*^{m-1} \sup_{\Delta} \{R\hat{U}_*/\hat{U}_*\} - \beta_*^{-1} \right) > 0$.

The discussion of Section 5.3 extends easily, and this yields Proposition 7.

6. TABLES

p	λ	b	$(m_D^{(\min)}; m_D)$	β_*^{-1}
0.6	0.9798	0.1101	(8,9)	0.9951
0.7	0.9165	0.2417	(2,3)	0.9695
0.8	0.8000	0.4000	(0,1)	0.9333
0.9	0.6000	0.6000	(0,1)	0.7500
0.95	0.4359	0.7320	(0,1)	0.5725

TABLE 1

p	0.6	0.6	0.6	0.7	0.7	0.7
α	0.05	0.25	0.5	0.05	0.25	0.5
ϵ_1	0.0008	0.0042	0.0084	0.0245	0.1225	0.2450
β_F^{-1}	0.9999	0.9999	0.9999	0.9986	0.9932	0.9865
ϵ_2	0.0680	0.1075	0.1152	0.0577	0.2537	0.4180
β_F^{-1}	0.9997	0.9995	0.9994	0.9964	0.9830	0.9757
ϵ	0.05	0.25	0.5	0.05	0.25	0.5
β_K^{-1}	0.9968	0.9841	0.9798	0.9956	0.9782	0.9573

TABLE 2

p	0.8	0.8	0.8	0.9	0.9	0.9	0.95	0.95	0.95
α	0.05	0.25	0.5	0.05	0.25	0.5	0.05	0.25	0.5
ϵ_2	0.05	0.25	0.5	0.05	0.25	0.5	0.05	0.25	0.5
β_F^{-1}	0.9793	0.9333	0.9333	0.9696	0.8539	0.7500	0.9564	0.7853	0.5814
ϵ	0.05	0.25	0.5	0.05	0.25	0.5	0.05	0.25	0.5
β_K^{-1}	0.9855	0.9302	0.8673	0.9775	0.8993	0.8000	0.9709	0.8616	0.7419

TABLE 3

REFERENCES

BAXENDALE, P. (1993). Uniform estimates for geometric ergodicity of recurrent Markov processes. University of Southern California.

- FORT, G. (2001). *Contrôle explicite d'ergodicité de chaîne de Markov: Applications à l'analyse de convergence de l'algorithme Monte-Carlo EM*. Ph.D. thesis, Université Pierre et Marie Curie, Paris.
- FORT, G. and MOULINES, E. (2002). Computable bounds for subgeometrical ergodicity. To appear in *Stochastic Processes Appl.*
- FORT, G., MOULINES, E., ROBERTS, G. and ROSENTHAL, J. (2001). On the geometric ergodicity of hybrid samplers. Submitted for publication, <http://www-lmc.imag.fr/lmc-sms/Gersende.Fort/biblio.html>.
- KALASHNIKOV, V. V. (1994). Regeneration and general Markov Chains. *J. Appl. Math. Stochastic Anal.* **7** 357–371.
- KOLASSA, J. (2000). Explicit bounds for geometric convergence of markov chains. *J. Appl. Probab.* **37** 642–651.
- LUND, R. B. and TWEEDIE, R. L. (1996). Geometric convergence rates for stochastically ordered Markov Chains. *Math. Oper. Res.* **20** 182–194.
- MENGERSEN, K. and TWEEDIE, R. (1996). Rates of convergence of the Hastings and Metropolis algorithms. *Ann. Statist.* **24** 101–121.
- MEYN, S. and TWEEDIE, R. (1994). Computable bounds for geometric convergence rates of Markov chains. *Ann. Appl. Probab.* **4** 981–1011.
- MEYN, S. P. and TWEEDIE, R. L. (1993). *Markov chains and stochastic stability*. Springer-Verlag London Ltd., London.
- NUMMELIN, E. (1978). A splitting technique for Harris recurrent Markov Chains. *Z. Wahrscheinlichkeitstheorie verw. Gebiete* **43** 309–318.
- NUMMELIN, E. and TUOMINEN, P. (1982). Geometric ergodicity of Harris recurrent Markov chains with applications to renewal theory. *Stochastic Processes Appl.* **2** 187–202.
- ROBERTS, G. and ROSENTHAL, J. (1997). Shift-coupling and convergence rates of ergodic averages. *Stochastic models* **13**.
- ROBERTS, G. and TWEEDIE, R. (1999). Bounds on regeneration times and convergence rates for Markov Chains. *Stochastic Processes Appl.* **80** 211–229.
- ROBERTS, G. and TWEEDIE, R. (2001). Corrigendum to: “Bounds on regeneration times and convergence rates for Markov chains”. *Stochastic Process. Appl.* **91** 337–338.
- ROSENTHAL, J. (2002). Quantitative convergence rates of Markov chains: A simple account. [Http://www.statslab.cam.ac.uk/mcmc/pages/latest.html](http://www.statslab.cam.ac.uk/mcmc/pages/latest.html).

ROSENTHAL, J. S. (1995). Minorization conditions and convergence rates for Markov chain Monte Carlo. *J. Amer. Statist. Assoc.* **90** 558–566.