Computable bounds for solutions to Poisson's equation and perturbation of Markov kernels

Loïc Hervé, James Ledoux

Computable bounds for solutions to Poisson's equation and perturbation of Markov kernels

Loïc HERVÉ, and James LEDOUX *

version: Saturday 21$^{\text{st}}$ June, 2025 – 13:38

Abstract

We consider a Markov kernel on a measurable space, satisfying a minorization condition and a modulated drift condition. Then we show that there exists a solution to the so-called Poisson equation whose norm can be bounded from above using the modulated drift condition. This new bound is very simple and can be easily computed. This result is obtained using the submarkov residual kernel given by the minorization condition. Such a bound allows us to provide new control on the weighted total variation norms of the deviation between the invariant probability measure π_{θ_0} of a Markov kernel P_{θ_0} and the invariant probability measure π_θ of some perturbation P_θ of P_{θ_0}. From the standard connexion between Poisson's equation and the central limit theorem, a simple and computable bound on the asymptotic variance is also derived.

AMS subject classification :

Keywords : Asymptotic variance; drift conditions; Invariant probability measure; perturbed Markov kernels; Poisson's equation;

1 Introduction

Let $(\mathbb{X},\mathcal{X})$ be a measurable space. We denote by $\mathcal{M}^+$ the set of finite non-negative measures on $(\mathbb{X},\mathcal{X})$. For any $\mu \in \mathcal{M}^+$ and any μ-integrable function $g : \mathbb{X}\to\mathbb{R}$, we set $\mu(g) = \int_{\mathbb{X}} g d\mu$. Let $\mathcal{M}^+_*$ be the set of positive measures, i.e. $\mathcal{M}^+_* := \{\mu \in \mathcal{M}^+ : \mu(1_{\mathbb{X}}) > 0\}$. If $V : \mathbb{X}\to[1,+\infty)$ is measurable, then for every measurable function $g : \mathbb{X}\to\mathbb{R}$, we define $\|g\|_V := \sup_{x\in\mathbb{X}} |g(x)|/V(x) \in [0,+\infty]$, and the space

$$\mathcal{B}_V := \{g : \mathbb{X}\to\mathbb{R}, \text{measurable such that } \|g\|_V < \infty\}.$$

Recall that a non-negative kernel $P(x,dy) \in \mathcal{M}^+$, $x \in \mathbb{X}$ is said to be a Markov (respectively submarkov) kernel if $P(x,\mathbb{X}) = 1$ (respectively $P(x,\mathbb{X}) \le 1$) for any $x \in \mathbb{X}$. We denote by P its functional action defined by

$$\forall x \in \mathbb{X}, \quad (Pg)(x) := \int_{\mathbb{X}} g(y)\, P(x,dy),$$

*Univ Rennes, INSA Rennes, CNRS, IRMAR-UMR 6625, F-35000, France. Loic.Herve@insa-rennes.fr, James.Ledoux@insa-rennes.fr

where $g : \mathbb{X} \to \mathbb{R}$ is any $P(x,\cdot)-$integrable function. Let P be a Markov kernel on $(\mathbb{X}, \mathcal{X})$ satisfying the following standard minorization and modulated drift conditions

$$\exists S \in \mathcal{X},\ \exists \nu \in \mathcal{M}_*^+, \quad \forall x \in \mathbb{X},\ \forall A \in \mathcal{X}, \quad P(x,A) \geq \nu(1_A)\, 1_S(x) \tag{$\boldsymbol{S}$}$$

$$\exists b > 0, \quad PV_0 \leq V_0 - V_1 + b1_S \tag{$\boldsymbol{D}(V_0,V_1)$}$$

where S in $\boldsymbol{D}(V_0,V_1)$ is the set given in $(\boldsymbol{S})$, and where V_0 and V_1 are measurable functions from $\mathbb{X}$ to $[1,+\infty)$ which are usually called Lyapunov functions for P. Under these conditions it was proved in Theorem 2.3 from [GM96] that there exists a $P-$invariant probability measure π such that $\pi(V_1) < \infty$, and that there exists a positive constant c_0 such that, for any $g \in \mathcal{B}_{V_1}$ satisfying $\pi(g) = 0$, the Poisson equation

$$(I - P)\widehat{g} = g \tag{1}$$

admits a solution $\widehat{g} \in \mathcal{B}_{V_0}$ such that $\pi(\widehat{g}) = 0$ and

$$\|\widehat{g}\|_{V_0} \leq c_0\, \|g\|_{V_1}. \tag{2}$$

Note that the function g is not assumed to be $\pi-$centred in the original Glynn-Meyn's statement. Throughout our paper, the condition $\pi(g) = 0$ will be used to simplify the statements. Simply apply the results to the function $g - \pi(g)1_{\mathbb{X}}$ to restore the general context. Under the aperiodicity condition, Glynn-Meyn's theorem is related to point-wise convergence of the series $\sum_{k=0}^{+\infty} P^k g$, see Theorem 14.0.1 from [MT09]. We point out that the constant c_0 in (2) is unknown in general. In Section 2, the following theorem is proved (Theorem 2.4).

Theorem 1 *Assume that P satisfies Conditions $(\boldsymbol{S})$-$\boldsymbol{D}(V_0,V_1)$. Then P admits an invariant probability measure π such that $\pi(1_S) > 0$ and $\pi(V_1) < \infty$. Moreover let us introduce the submarkov residual kernel $R := P - \nu(\cdot)1_S$. Then, for every $g \in \mathcal{B}_{V_1}$ such that $\pi(g) = 0$, the function $\widetilde{g} := \sum_{k=0}^{+\infty} R^k g$ belongs to $\mathcal{B}_{V_0}$ and satisfies the Poisson equation (1), that is $(I - P)\widetilde{g} = g$, with*

$$\|\widetilde{g}\|_{V_0} \leq a\|g\|_{V_1} \quad \textit{where} \quad a := 1 + \max\left(0\,,\, \frac{b - \nu(V_0)}{\nu(1_{\mathbb{X}})}\right). \tag{3}$$

Let us comment on the conclusions of Theorem 1 and the main ideas of its proof. The $P-$invariant probability measure π satisfying $\pi(1_S) > 0$ has a Nummelin-type representation, that is: $\pi = \mu(1_{\mathbb{X}})^{-1}\mu$ with $\mu := \sum_{k=0}^{+\infty} \nu R^k \in \mathcal{M}_*^+$. This representation is classical under various hypotheses, see Theorem 5.2 and Corollary 5.2 from [Num84]. Here we use the version recently proved in [HL23a] under the sole minorization condition $(\boldsymbol{S})$, see Recall 2.1. Next the original trick in the present work is that, under Assumption $\boldsymbol{D}(V_0,V_1)$, the submarkov residual kernel $R := P - \nu(\cdot)1_S$ satisfies the following drift condition

$$RV_{0,d} \leq V_{0,d} - V_1 \tag{4}$$

where $V_{0,d} := V_0 + d1_{\mathbb{X}}$ with $d := \max\{0, (b - \nu(V_0))/\nu(1_{\mathbb{X}})\}$, see Lemma 2.3. Then the residual-type drift condition (4) enables us to define the function $\widetilde{g}$ and to obtain the bound (3), while the above representation of π is proved to be crucial here to obtain that $\widetilde{g}$ is a solution to the Poisson equation. The innovative point in Theorem 1 is that the bound (3) is simple and explicit.

Beyond Theorem 1 which has its own interest, we are also interested in the following four applications of the bound (3).

- *Bound for a $\pi-$centred solution to Poisson's equation (Corollary 2.6).* Let $g \in \mathcal{B}_{V_1}$ be such that $\pi(g) = 0$. Then $\widehat{g} = \widetilde{g} - \pi(\widetilde{g})1_{\mathbb{X}}$ is a $\pi-$centred function in $\mathcal{B}_{V_0}$ solution to Poisson's equation $(I - P)\widehat{g} = g$, and it satisfies $\|\widehat{g}\|_{V_0} \leq a\,(1 + \pi(V_0))\,\|g\|_{V_1}$ from (3). The condition $\pi(V_0) < \infty$ is here required. This result is particularly relevant in the case when two solutions to Poisson's equation in $\mathcal{B}_{V_0}$ are known to differ from an additive constant. Indeed, in this case, for any $\pi-$centred function $\xi \in \mathcal{B}_{V_0}$ solution to Poisson equation $(I - P)\xi = g$, we have $\xi = \widehat{g}$, so that the previous bound applies to ξ. Of course such a solution ξ may be obtained independently of the function $\widetilde{g}$. For instance use ξ given by Theorem 2.3 from [GM96], in particular $\xi = \sum_{k=0}^{+\infty} P^k g$ whenever this series point-wise converges and defines a function of $\mathcal{B}_{V_0}$.

- *Bound for the asymptotic variance (Corollary 2.7).* If $g \in \mathcal{B}_{V_1}$ is such that $\pi(g) = 0$, then the so-called asymptotic variance $\gamma_g^2 = \pi((\widetilde{g})^2 - (P\widetilde{g})^2)$, which is involved in the central limit theorem for Markov chains, satisfies $\gamma_g^2 \leq 2\,a^2\,\pi(V_0^2)\,\|g\|_{V_1}^2$ thanks to (3). Here the condition $\pi(V_0^2) < \infty$ is assumed to hold.

- *The geometrically ergodic case (Corollary 2.8).* The so-called $V-$geometrical ergodicity property is based on the geometric drift condition $PV \leq \delta V + K1_S$ for some $\delta \in (0, 1)$, $K \in (0, +\infty)$. In this specific geometric case, the bounds obtained in [HL24] for solutions to Poisson's equation involve the constant $(1 - \delta^{\alpha_0})^{-1}$ for some $\alpha_0 \in (0, 1]$. This constant derived from the geometric drift condition and the spectral theory is very large when α_0 is close to zero. Since the geometric drift condition is a special case of modulated drift condition $\boldsymbol{D}(V_0, V_1)$, Property (3) can then be used to obtain an alternative bound for solutions to Poisson's equation in such a case.

- *Bound in perturbation issues (Section 3).* Here the weighted total variation norm is used, that is: If $(\mu_1, \mu_2) \in (\mathcal{M}^+)^2$ is such that $\mu_i(V) < \infty, i = 1, 2$ for some measurable function $V : \mathbb{X} \to [1, +\infty)$, then the V-weighted total variation norm $\|\mu_1 - \mu_2\|_V$ is defined by

  $$\|\mu_1 - \mu_2\|_V := \sup_{|g| \leq V} \big|\mu_1(g) - \mu_2(g)\big|. \tag{5}$$

  When P and P' are two Markov kernels on $(\mathbb{X}, \mathcal{X})$ with respective invariant probability measures π and π', the following formula is of interest to control $\pi'(g) - \pi(g)$:

  $$\pi'(g) - \pi(g) = \pi'\big((P' - P)\xi\,\big) \tag{6}$$

  where the function ξ is any solution to Poisson equation $(I - P)\xi = g - \pi(g)1_{\mathbb{X}}$. Accordingly, using in Formula (6) the solution $\xi := \sum_{k=0}^{+\infty} R^k(g - \pi(g)1_{\mathbb{X}})$ provided by Theorem 1, the explicit bound (3) is of great interest. This perturbation issue is addressed in Section 3 for a general family $\{P_\theta\}_{\theta\in\Theta}$ of Markov kernels, each of them satisfying a minorization condition and a modulated drift condition w.r.t. some Lyapunov functions V_0 and V_1 (independent of θ). Thus, denoting by π_θ the $P_\theta-$invariant probability measure provided by Theorem 1, and fixing some $\theta_0 \in \Theta$, it is proved in Theorem 3.2 that $\|\pi_\theta - \pi_{\theta_0}\|_{V_1} \to 0$ when $\theta \to \theta_0$, provided that for every $x \in \mathbb{X}$ we have $\Delta_{\theta,V_0}(x) := \|P_\theta(x, \cdot) - P_{\theta_0}(x, \cdot)\|_{V_0} \to 0$ when $\theta \to \theta_0$. Moreover a bound for $\|\pi_\theta - \pi_{\theta_0}\|_{V_1}$ is given in terms of the quantity $\pi_\theta(\Delta_{\theta,V_0})$. Here P_{θ_0} may be considered as the Markov kernel of interest, and the P_θ's for $\theta \neq \theta_0$ must be thought of some perturbed Markov kernels which are more tractable than P_{θ_0}. In particular the term $\pi_\theta(\Delta_{\theta,V_0})$ is expected to be known or at least computable for $\theta \neq \theta_0$.

Condition $\boldsymbol{D}(V_0, V_1)$ is the so-called V_1-modulated drift condition, e.g. see Condition (V3) in Chapter 14 of [MT09]. Although the functions V_0, V_1 in $\boldsymbol{D}(V_0, V_1)$ satisfy $V_0 \geq V_1$ in general, this condition is not useful in the present work. Condition $\boldsymbol{D}(V_0, V_1)$ has been widely used to analyse the geometric or sub-geometric rate of convergence in total variation norms of the Markov chain to its invariant probability measure π (e.g. see Chapters 16, 17 of [DMPS18] for an overview and various examples, and [Del17] for an alternative operator-type approach). To the best of our knowledge, an estimate of the constant c_0 in (2) is only provided in Proposition 1 in [LL18] for a discrete state-space $\mathbb{X}$ and in [Mas19] for a continuous-time Markov chain with a general state-space $\mathbb{X}$. In both [LL18] and [Mas19] the existence of an atom is assumed, and standard regeneration approach is then applied under the V_1-modulated drift condition to obtain the existence and a bound of a $\pi-$centred solution to Poisson's equation. Here, we use a quite different approach that does not require the existence of an atom. The perturbation theory for Markov chains has been widely developed in the last decades. Formula (6) was first used in [Sch68] for finite irreducible stochastic matrices, see also [Sen93]. This formula can be subsequently used in any problem which can be thought of as a perturbation problem of Markov kernels (e.g. see [GM96, LL18] and Section 17.7 in [MT09]). Recall that the strong continuity assumption introduced in [Kar86] is suitable when $P_\theta = P_{\theta_0} + \theta D$ where $\theta \in \mathbb{R}$ and D is a real-valued kernel satisfying $D(x, 1_{\mathbb{X}}) = 0$ for every $x \in \mathbb{X}$, e.g. see [AANQ04, Mou21]. Note that neither the specific investigation of uniformly ergodic Markov chains as in [Mit05, MA10, AFEB16, JMMD], nor that of reversible transition kernels as in [MALR16, NR21], are addressed here. Recently the approach from [Kel82] for perturbed dynamical systems involving a weak continuity assumption has been adapted to V-geometrically ergodic Markov models, either using the Keller-Liverani perturbation theorem from [KL99] (see [FHL13, HL14, HL23b]), or using an elegant idea of [HM11] based on Wasserstein distance as in [SS00] or in [RS18, MARS20]. These works only concern the geometrically ergodic case.

To the best of our knowledge, the bound (3) for solution to Poisson's equation, as well as the results obtained in the four applications above, are new. As in the recent work [HL23a] providing a $P-$invariant probability measure under Condition $(\boldsymbol{S})$, the proof of Theorem 1 is self-contained. In particular there is no need to study the atomic case first and then to apply the splitting technique to encompass the general case, i.e. to introduce an appropriate enlargement of the state space in order to get a new Markov kernel which has an atom, see Section 4.4 in [Num84] for details. All the bounds in this work apply whenever explicit modulated drift condition $\boldsymbol{D}(V_0, V_1)$ is known: for such examples, e.g. see [FM00, FM03, DFM16] in the context of the Metropolis algorithm, [TT94, DFM16] for autoregressive models, [LH07, LH12] for queueing systems, [JT02] for Markov chains associated with the mean of Dirichlet processes. Classical instances of $V-$geometrically ergodic Markov chains can be found in [MT09, RR04, DMPS18].

Theorem 1 is detailed and proved in Section 2, together with Corollaries 2.6–2.8 which concern the first three applications above. Then Section 3 is devoted to the last application above (perturbation issue). In the specific context of $V-$geometrically ergodic Markov kernels, the perturbation results of Section 3 can be compared with those in [HL14] and [RS18], see Example 3.7. Finally, in Example 3.8, our results are made explicit for perturbed random walks on the half line.

2 The drift conditions and Poisson's equation

Let P be a Markov kernel on $(\mathbb{X}, \mathcal{X})$. The following statement proved in [HL23a] provides a $P-$invariant probability measure under the minorization Condition ($\boldsymbol{S}$) (i.e. when P admits a small set, e.g. see [MT09]). The positive measure ν in ($\boldsymbol{S}$) is often written in the literature as $\nu = \varepsilon \mathsf{p}$ for some probability measure p on $(\mathbb{X}, \mathcal{X})$ and some $\varepsilon \in (0,1]$. This formulation is not used here (note that the positive real number $\nu(1_{\mathbb{X}})$ used in some bounds below equals to ε). Actually, the main property derived from Condition ($\boldsymbol{S}$) here is that the following so-called residual kernel R is a submarkov kernel

$$\forall x \in \mathbb{X}, \quad R(x,\cdot) := P(x,\cdot) - \nu(\cdot) 1_S(x). \tag{7}$$

For any non-negative kernel $K(x,dy) \in \mathcal{M}^+$, $x \in \mathbb{X}$, recall that the $n-$th iterate kernel of $K(x,dy)$ with $n \geq 1$ is denoted by $K^n(x,dy)$, $x \in \mathbb{X}$, and K^n stands for its functional action. As usual K^0 is the identity map I by convention.

Under Condition ($\boldsymbol{S}$), the following statement from [HL23a] provides a simple characterization for P to have an invariant probability measure π such that $\pi(1_S) > 0$. For convenience of the reader, the analytic proof from [HL23a] is reported in Annex A.

Recall 2.1 *If P satisfies the minorization condition ($\boldsymbol{S}$), then the following assertions are equivalent.*

1. *There exists a $P-$invariant probability measure π on $(\mathbb{X}, \mathcal{X})$ such that $\pi(1_S) > 0$.*

2. $\displaystyle\sum_{k=1}^{+\infty} \nu(R^{k-1}1_{\mathbb{X}}) < \infty$ *with R given in (7).*

Under any of these two conditions

$$\pi \equiv \pi_{\nu,R} := \mu(1_{\mathbb{X}})^{-1}\,\mu \quad \textit{with} \quad \mu := \sum_{k=1}^{+\infty} \nu R^{k-1} \in \mathcal{M}_*^+ \tag{8}$$

is a $P-$invariant probability measure on $(\mathbb{X}, \mathcal{X})$ with $\mu(1_S) = 1$ and $\pi(1_S) = \mu(1_{\mathbb{X}})^{-1} > 0$.

Assume that P satisfies Condition ($\boldsymbol{S}$) and that $V_0 : \mathbb{X} \to [1, +\infty)$ is a measurable function such that the function PV_0 is everywhere finite, i.e. $\forall x \in \mathbb{X}$, $(PV_0)(x) < \infty$. Then we have $\nu(V_0) \leq (PV_0)(x) < \infty$ for any $x \in S$ from ($\boldsymbol{S}$), so that the non-negative function RV_0 is well-defined. Now, given another measurable function $V_1 : \mathbb{X} \to [1, +\infty)$, let us introduce the following residual-type drift condition:

$$RV_0 \leq V_0 - V_1. \tag{$\boldsymbol{R}(V_0,V_1)$}$$

Note that Condition $\boldsymbol{D}(V_0, V_1)$ when $b = \nu(V_0)$ reduces to $\boldsymbol{R}(V_0, V_1)$. Moreover Condition $\boldsymbol{R}(V_0, V_1)$ implies that $V_1 \leq V_0$ since $RV_0 \geq 0$, and that $PV_0 \leq (1 + \nu(V_0))V_0$, hence $\|PV_0\|_{V_0} < \infty$.

Assuming the residual-type drift condition $\boldsymbol{R}(V_0, V_1)$ and using Formula (8) for the $P-$invariant probability measure π, we can derive the next statement which will be central for obtaining the bound (3) under the general V_1-modulated drift condition $\boldsymbol{D}(V_0, V_1)$ (see Theorem 2.4). It states that, for any $g \in \mathcal{B}_{V_1}$ the series $\sum_{k=0}^{+\infty} R^k g$ pointwise converges in $\mathbb{X}$ and defines a function $\widetilde{g}$ in $\mathcal{B}_{V_0}$ satisfying the nice bound $\|\widetilde{g}\|_{V_0} \leq \|g\|_{V_1}$. Moreover $\widetilde{g}$ is a solution to Poisson's equation when $\pi(g) = 0$.

Proposition 2.2 *Assume that P satisfies Condition ($\boldsymbol{S}$) and that $V_0 : \mathbb{X} \to [1, +\infty)$ is a measurable function such that PV_0 is everywhere finite. If the residual kernel R given in (7) satisfies the drift condition $\boldsymbol{R}(V_0, V_1)$ for some measurable function $V_1 : \mathbb{X} \to [1, +\infty)$, then the following assertions hold.*

1. *For any $g \in \mathcal{B}_{V_1}$, the function $\widetilde{g} := \sum_{k=0}^{+\infty} R^k g$ is well-defined on $\mathbb{X}$ and $\widetilde{g} \in \mathcal{B}_{V_0}$ with*

$$\|\widetilde{g}\|_{V_0} \leq \|g\|_{V_1}. \tag{9}$$

2. *The $P-$invariant probability measure $\pi \equiv \pi_{\nu,R}$ in (8) is well-defined and satisfies $\pi(V_1) < \infty$.*

3. *For any $g \in \mathcal{B}_{V_1}$ such that $\pi(g) = 0$, the function $\widetilde{g}$ satisfies Poisson's equation*

$$(I - P)\widetilde{g} = g. \tag{10}$$

Proof. Let $x \in \mathbb{X}$. From $\boldsymbol{R}(V_0, V_1)$, we derive that $V_1 \leq V_0 - RV_0$ and we obtain

$$\forall n \geq 1, \quad \sum_{k=0}^{n} (R^k V_1)(x) \leq V_0(x), \tag{11}$$

so that $\sum_{k=0}^{+\infty} (R^k V_1)(x) \leq V_0(x)$. Now let $g \in \mathcal{B}_{V_1}$. Using $|g| \leq \|g\|_{V_1} V_1$, it follows that

$$\sum_{k=0}^{+\infty} \big|(R^k g)(x)\big| \leq \|g\|_{V_1} V_0(x).$$

This proves Assertion 1. Next it follows from $1_{\mathbb{X}} \leq V_1$ and $\sum_{k=0}^{+\infty} R^k V_1 \leq V_0$ that

$$\sum_{k=0}^{+\infty} \nu(R^k 1_{\mathbb{X}}) \leq \sum_{k=0}^{+\infty} \nu(R^k V_1) \leq \nu(V_0) < \infty.$$

Hence the positive measure $\mu := \sum_{k=0}^{+\infty} \nu R^k$ is such that

$$0 < \nu(1_{\mathbb{X}}) \leq \mu(1_{\mathbb{X}}) \leq \mu(V_1) \leq \nu(V_0) < \infty.$$

Then Assertion 2 follows from Recall 2.1.

Let $g \in \mathcal{B}_{V_1}$. Since $\pi(V_1) = \mu(1_{\mathbb{X}})^{-1} \mu(V_1) < \infty$, we have $\pi(|g|) < \infty$. Now define

$$\forall n \geq 1, \quad \widetilde{g}_n := \sum_{k=0}^{n} R^k g.$$

Then, using $P = R + \nu(\cdot) 1_S$ and setting $\mu_n(g) := \nu(\widetilde{g}_n) = \sum_{k=0}^{n} \nu(R^k g)$ we have

$$(I - P)\widetilde{g}_n = \widetilde{g}_n - R\widetilde{g}_n - \mu_n(g) 1_S = g - R^{n+1} g - \mu_n(g) 1_S. \tag{12}$$

We know that $\lim_n R^{n+1}g = 0$ (pointwise convergence) from the convergence of the series $\sum_{k=0}^{+\infty} R^k g$. Moreover, using $\mu(V_1) < \infty$, we obtain that $\lim_{n\to+\infty} \mu_n(g) = \mu(g)$. Finally, for every $x \in \mathbb{X}$, we have $\lim_n (P\widetilde{g}_n)(x) = (P\widetilde{g})(x)$ from Lebesgue's theorem applied to the sequence $(\widetilde{g}_n)_n$ w.r.t. the probability measure $P(x, dy)$ since $\lim_n \widetilde{g}_n = \widetilde{g}$, $|\widetilde{g}_n| \le \|g\|_{V_1} V_0$ and $(PV_0)(x) < \infty$. Taking the limit when n goes to infinity in (12), we obtain

$$(I - P)\widetilde{g} = g - \mu(g)1_S. \tag{13}$$

Next, if we assume that $\pi(g) = 0$, then Equality (13) rewrites as Equality (10) since $\mu(g) = \pi(g)/\pi(1_S) = 0$ from the representation of π. The proof of Proposition 2.2 is complete. □

Now assume that P satisfies the minorization condition ($\boldsymbol{S}$) and the V_1-modulated drift condition $\boldsymbol{D}(V_0, V_1)$ for some couple (V_0, V_1) of Lyapunov functions. This implies that the function PV_0 is everywhere finite. Then we have $\nu(V_0) < \infty$ from ($\boldsymbol{S}$). Thus the non-negative function RV_0 is well-defined where R is the residual kernel defined in (7). If Condition $\boldsymbol{D}(V_0, V_1)$ holds with an atom S (i.e. $\forall x \in S,\ P(x, \cdot) = \nu(\cdot)$) and with $V_0 \ge V_1$ on S, then $b = \nu(V_0)$ may be chosen in $\boldsymbol{D}(V_0, V_1)$, so that Condition $\boldsymbol{R}(V_0, V_1)$ holds too. In the non-atomic case, the drift condition $\boldsymbol{D}(V_0, V_1)$ on P may not directly provide the residual-type condition $\boldsymbol{R}(V_0, V_1)$ since the constant b may be strictly larger than $\nu(V_0)$. However, starting from Assumption $\boldsymbol{D}(V_0, V_1)$, the next lemma shows that the slight change of the Lyapunov function V_0 into $V_{0,d} = V_0 + d1_{\mathbb{X}}$, with some suitable positive constant d, does provide the residual-type drift condition $\boldsymbol{R}(V_{0,d}, V_1)$.

Lemma 2.3 *Assume that P satisfies Conditions ($\boldsymbol{S}$) and $\boldsymbol{D}(V_0, V_1)$ w.r.t. some couple (V_0, V_1) of Lyapunov functions. Let $c \ge (b - \nu(V_0))/\nu(1_{\mathbb{X}})$. Then the residual kernel R defined in (7) satisfies Condition $\boldsymbol{R}(V_{0,d}, V_1)$, where $V_{0,d} := V_0 + d1_{\mathbb{X}} \ge V_0$ with $d = \max(0, c)$.*

Proof. We already know that the function RV_0 is well-defined and is finite from Assumptions $\boldsymbol{D}(V_0, V_1)$ and ($\boldsymbol{S}$). Set $d := \max(0, c)$ and $V_{0,d} := V_0 + d1_{\mathbb{X}}$. Note that $\nu(V_{0,d}) = \nu(V_0) + d\nu(1_{\mathbb{X}}) < \infty$ and that $PV_{0,d} = PV_0 + d1_{\mathbb{X}} < \infty$. We have

$$\begin{aligned} RV_{0,d} = PV_{0,d} - \nu(V_{0,d})1_S &= PV_0 + d1_{\mathbb{X}} - \big(\nu(V_0) + d\,\nu(1_{\mathbb{X}})\big)1_S \\ &\le V_0 - V_1 + b1_S + d1_{\mathbb{X}} - \big(\nu(V_0) + d\,\nu(1_{\mathbb{X}})\big)1_S \\ &\le V_{0,d} - V_1 + \big(b - \nu(V_0) - d\,\nu(1_{\mathbb{X}})\big)1_S \end{aligned}$$

from the definitions of R and $V_{0,d}$, and from Assumption $\boldsymbol{D}(V_0, V_1)$. The proof is complete. □

Under the standard V_1-modulated drift condition $\boldsymbol{D}(V_0, V_1)$ on P, the following theorem is derived from Lemma 2.3 and Proposition 2.2. It can be thought of as an extension of Theorem 2.3 in [GM96] (and Theorem 17.7.1 in [MT09]) in that it provides an explicit and simple bound on the V_0-norm of a solution to Poisson's equation. To the best of our knowledge, the joint use of Lemma 2.3 and Proposition 2.2, as well as the bound (14) below, are new.

Theorem 2.4 *Assume that P satisfies the minorization Condition ($\boldsymbol{S}$) and the V_1-modulated drift condition $\boldsymbol{D}(V_0, V_1)$ w.r.t. some couple (V_0, V_1) of Lyapunov functions.*

Then the conclusions stated in Assertions 1 − 3 of Proposition 2.2 hold true with the following bound in place of (9)

$$\forall g \in \mathcal{B}_{V_1}, \quad \|\widetilde{g}\|_{V_0} \le a\|g\|_{V_1} \quad \textit{with} \quad a := 1 + \max\left(0,\, \frac{b - \nu(V_0)}{\nu(1_{\mathbb{X}})}\right) \tag{14}$$

where $\widetilde{g} := \sum_{k=0}^{+\infty} R^k g$, $\nu \in \mathcal{M}_*^+$ *is given in* $(\boldsymbol{S})$ *and* b *is the positive constant given in* $\boldsymbol{D}(V_0, V_1)$.

Recall that, if the set S in Conditions $(\boldsymbol{S})$-$\boldsymbol{D}(V_0, V_1)$ is an atom and if $V_0 \geq V_1$ on S, then Condition $\boldsymbol{D}(V_0, V_1)$ holds with $b = \nu(V_0)$, so that $a = 1$. Moreover mention that, when $g \in \mathcal{B}_{V_1}$ is such that $\pi(g) \neq 0$, the centred function $g_0 := g - \pi(g)1_{\mathbb{X}}$ is such that $\widetilde{g}_0 := \sum_{k=0}^{+\infty} R^k g_0$ is a solution in $\mathcal{B}_{V_0}$ to Poisson's equation $(I-P)\widetilde{g}_0 = g_0 = g - \pi(g)1_{\mathbb{X}}$. The bound (14) is similar to those in Proposition 1 in [LL18] and in [Mas19], which have been obtained under V_1-modulated drift conditions and assuming the existence of an atom. In this work, we do not assume the existence of an atom. Nor do we use the splitting method for passing from the atomic case to the general one. Actually the atomic case is encompassed by the assumptions of Proposition 2.2.

Proof. Let $V_{0,d} := V_0 + d1_{\mathbb{X}}$ where $d := \max(0, \widehat{c})$ with $\widehat{c} := (b - \nu(V_0))/\nu(1_{\mathbb{X}})$. Note that V_0 and $V_{0,d}$ are equivalent functions in the sense that $V_0 \leq V_{0,d} \leq (1+d)V_0$. Then Proposition 2.2 applied with the drift condition of Lemma 2.3 shows that, for any $g \in \mathcal{B}_{V_1}$, the function $\widetilde{g} := \sum_{k=0}^{+\infty} R^k g$ belongs to $\mathcal{B}_{V_{0,d}}$ with

$$\|\widetilde{g}\|_{V_{0,d}} \leq \|g\|_{V_1},$$

and that $\widetilde{g}$ satisfies $(I-P)\widetilde{g} = g$ when $\pi(g) = 0$. Next (14) holds since $\|\cdot\|_{V_0} \leq (1+d)\|\cdot\|_{V_{0,d}}$ from the inequality $V_{0,d} \leq (1+d)V_0$. □

Remark 2.5 *Assume that Conditions* $(\boldsymbol{S})$ *and* $\boldsymbol{D}(V_0, V_1)$ *are satisfied for the Markov kernel* P^ℓ *with some* $\ell \geq 2$. *Moreover assume that* π *is the unique invariant probability measure for both* P *and* P^ℓ. *Recall that* $M := \|PV_0\|_{V_0} < \infty$ *from* $\boldsymbol{D}(V_0, V_1)$. *Set* $R_\ell := P^\ell - \nu(\cdot)1_S$. *Then, for every* $g \in \mathcal{B}_{V_1}$ *such that* $\pi(g) = 0$, *the function*

$$\tilde{g} := \sum_{k=0}^{\ell-1} P^k \widetilde{g}_\ell \quad \textit{with} \quad \widetilde{g}_\ell = \sum_{k=0}^{+\infty} {R_\ell}^k g$$

belongs to $\mathcal{B}_{V_0}$ *and satisfies the Poisson equation* $(I-P)\tilde{g} = g$. *Moreover we have*

$$\|\tilde{g}\|_{V_0} \leq \frac{a\big(M^\ell - 1\big)}{M-1}\, \|g\|_{V_1} \quad \textit{with} \quad a := 1 + \max\left(0\,,\, \frac{b - \nu(V_0)}{\nu(1_{\mathbb{X}})}\right)$$

where $\nu \in \mathcal{M}_*^+$ *and* b *are here given in Conditions* $(\boldsymbol{S})$-$\boldsymbol{D}(V_0, V_1)$ *related to* P^ℓ. *Indeed Theorem 2.4 applied to the Markov kernel* P^ℓ *shows that, for every* $g \in \mathcal{B}_{V_1}$ *such that* $\pi(g) = 0$, *the function* $\widetilde{g}_\ell$ *belongs to* $\mathcal{B}_{V_0}$ *and satisfies* $(I - P^\ell)\widetilde{g}_\ell = g$, *with moreover*

$$\|\widetilde{g}_\ell\|_{V_0} \leq a\, \|g\|_{V_1}.$$

The claimed statements then follow from $(I-P)\tilde{g} = (I-P^\ell)\widetilde{g}_\ell = g$ *and from the inequality* $\|\tilde{g}\|_{V_0} \leq \|\widetilde{g}_\ell\|_{V_0}(M^\ell - 1)/(M-1)$. *The interest of this remark is essentially theoretical because the constant* M^ℓ *degrades when* ℓ *is large.*

Note that the invariant probability measure $\pi \equiv \pi_{\nu,R}$ in (8), which is involved in Proposition 2.2 and Theorem 2.4, only satisfies the moment condition $\pi(V_1) < \infty$, and there is no guarantee that $\pi(V_0) < \infty$. For a Markov model satisfying Assumption $\boldsymbol{D}(V_0, V_1)$, it is worth noticing that the condition $\pi(V_0) < \infty$ holds provided that P satisfies any preliminary V_0-modulated drift condition, that is: $PW \leq W - V_0 + b1_S$ for some Lyapunov function W (apply Theorem 2.4 to the couple (W, V_0)). Recall that such nested modulated drift conditions $\boldsymbol{D}(W, V_0)$ and $\boldsymbol{D}(V_0, V_1)$ occur in most of the analysis of polynomial or subgeometric convergence rate of Markov models, e.g. see [JR02, FM03, AFV15] and in particular Lemma 3.5 in [JR02] in the polynomial case and [DFMS04, DMPS18] in the subgeometric case.

In Corollary 2.6 below, we prove that, if the invariant probability measure $\pi \equiv \pi_{\nu,R}$ in (8) is such that $\pi(V_0) < \infty$, then it is the unique one integrating V_0. This statement is suitable to the perturbation results of the next Section 3, in which the moment condition $\pi(V_0) < \infty$ is involved.

Corollary 2.6 *Let P satisfying the assumptions of Theorem 2.4. Assume that the invariant probability measure $\pi \equiv \pi_{\nu,R}$ in (8) is such that $\pi(V_0) < \infty$. Then*

1. *π is the unique $P-$invariant probability measure which integrates V_0.*
2. *For any $g \in \mathcal{B}_{V_1}$ such that $\pi(g) = 0$, let $\widetilde{g} := \sum_{k=0}^{+\infty} R^k g$. Then the function $\widehat{g} = \widetilde{g} - \pi(\widetilde{g})1_{\mathbb{X}}$ is a $\pi-$centered solution on $\mathcal{B}_{V_0}$ to Poisson's equation $(I-P)\widehat{g} = g$. Moreover we have*

$$\|\widehat{g}\|_{V_0} \leq a\,(1 + \pi(V_0))\,\|g\|_{V_1} \tag{15}$$

where the positive constant a is given in (14).

Note that, when Poisson's equation has a unique solution up to an additive constant, Inequality (15) gives a bound for the norm of the solution in Glynn-Meyn's theorem.

Proof. Let $g \in \mathcal{B}_{V_1}$. We know from Theorem 2.4 and Equality (13) that the associated function $\widetilde{g} := \sum_{k=0}^{+\infty} R^k g$ is in $\mathcal{B}_{V_0}$ and satisfies Equation $(I-P)\widetilde{g} = g - \mu(g)1_S$ with $\mu := \sum_{k=0}^{+\infty} \nu R^k \in \mathcal{M}_*^+$. Recall that $\pi = \mu(1_{\mathbb{X}})^{-1}\mu$. Consequently, if η is a $P-$invariant positive measure on $\mathbb{X}$ such that $\eta(V_0) < \infty$, then we have $\eta((I-P)\widetilde{g}) = 0 = \eta(g) - \mu(g)\eta(1_S)$, thus $\eta = \eta(1_S)\mu = \eta(1_S)\mu(1_{\mathbb{X}})\pi$. This proves the first assertion of Corollary 2.6.

To prove the second one, first note that $\widehat{g} \in \mathcal{B}_{V_0}$ and the property $\pi(\widehat{g}) = 0$ (under $\pi(V_0) < \infty$) are obvious. Moreover, if g is such that $\pi(g) = 0$, then we have $(I-P)\widehat{g} = (I-P)\widetilde{g} = g$ from Theorem 2.4 and $(I-P)1_{\mathbb{X}} = 0$. Finally we have

$$\|\widehat{g}\|_{V_0} \leq \big(1 + \pi(V_0)\,\|1_{\mathbb{X}}\|_{V_0}\big)\|\widetilde{g}\|_{V_0} \leq a\,(1 + \pi(V_0))\|g\|_{V_1}$$

using the definition of $\widehat{g}$, the triangular inequality and $|\widetilde{g}| \leq \|\widetilde{g}\|_{V_0} V_0$ for the first inequality, and finally $\|1_{\mathbb{X}}\|_{V_0} \leq 1$ and the bound (14) applied to $\widetilde{g}$ for the second one. □

For Markov kernels satisfying a modulated drift condition, the existence and uniqueness of the $P-$invariant probability measure is investigated in many works under various hypothesis, e.g. see [MT09, DMPS18] and Theorem 1 in [FM03]. For instance, if P is $\psi-$irreducible for some positive measure ψ on $(\mathbb{X}, \mathcal{X})$ and satisfies Condition $\boldsymbol{D}(V_0, V_1)$, then P has a unique invariant probability measure, e.g. see Theorem 6.12 in [Mey22].

The next proposition provides a computable bound for the so-called asymptotic variance involved in the central limit theorem for Markov chains, e.g. see Chapter 17 in [MT09], Chapter 21 in [DMPS18] and [Jon04]. To the best of our knowledge, this bound is new, and this is achieved thanks to the simple bound (14) of Theorem 2.4, since the asymptotic variance is known to be closely related to Poisson's equation.

Corollary 2.7 *Assume that P satisfies Conditions ($\boldsymbol{S}$)-$\boldsymbol{D}(V_0, V_1)$ and that the invariant probability measure $\pi \equiv \pi_{\nu,R}$ given in (8) is such that $\pi(V_0^2) < \infty$. For any $g \in \mathcal{B}_{V_1}$ such that $\pi(g) = 0$, set $\gamma_g^2 = \pi((\widetilde{g})^2 - (P\widetilde{g})^2)$ where $\widetilde{g}$ is the solution to Poisson's equation $(I - P)\widetilde{g} = g$ provided by Theorem 2.4. Then we have*

$$\gamma_g^2 \leq 2\, a^2\, \pi(V_0^2)\, \|g\|_{V_1}^2$$

where a is the positive constant given in (14).

Proof. From Theorem 2.4, we obtain that

$$\gamma_g^2 \leq \pi(\widetilde{g}^2) + \pi((P\widetilde{g})^2) \leq 2\pi(\widetilde{g}^2) \leq 2\, a^2 \pi(V_0^2)\, \|g\|_{V_1}^2$$

using successively the Cauchy-Schwarz inequality $(P\widetilde{g})^2 \leq P\,\widetilde{g}^2$, the $P-$invariance of π and finally $|\widetilde{g}| \leq a\|g\|_{V_1} V_0$ from (14). □

To conclude this section let us apply all the previous statements to the case when P satisfies Condition ($\boldsymbol{S}$) and the following so-called $V-$geometric drift condition

$$\exists \delta \in (0,1),\ \exists K \in (0,+\infty), \quad PV \leq \delta V + K 1_S \qquad (\boldsymbol{G}(\delta, V))$$

for some Lyapunov function V, where $S \in \mathcal{X}$ is the set in ($\boldsymbol{S}$). Then rewriting Condition $\boldsymbol{G}(\delta, V)$ as $PV \leq V - (1-\delta)V + K1_S$, we obtain that P satisfies the following Condition $\boldsymbol{D}(V_0, V)$

$$PV_0 \leq V_0 - V + b1_S \quad \text{with} \quad V_0 := \frac{V}{1-\delta} \quad \text{and} \quad b := \frac{K}{1-\delta}. \tag{16}$$

In addition to Conditions ($\boldsymbol{S}$) and $\boldsymbol{G}(\delta, V)$, we assume as in [Bax05] that $\nu(1_S) > 0$ (strong aperiodicity condition), so that P is $V-$geometrically ergodic. In particular we know that $\pi(V) < \infty$ and that two solutions to Poisson's equation in $\mathcal{B}_V$ differ by a constant. Observing that $\|\cdot\|_{V_0} = (1-\delta)\|\cdot\|_V$ and that $\pi(V_0)\,\|1_{\mathbb{X}}\|_{V_0} = \pi(V)\,\|1_{\mathbb{X}}\|_V$, the next statements are easily deduced from Theorem 2.4, Corollary 2.6 and Corollary 2.7.

Corollary 2.8 *Assume that P satisfies the minorization Condition ($\boldsymbol{S}$) with $\nu(1_S) > 0$ and the $V-$geometric drift condition $\boldsymbol{G}(\delta, V)$ w.r.t. some Lyapunov function V. Then*

1. *The conclusions stated in Assertions $1-3$ of Proposition 2.2 hold true with the following bound in place of (9):*

$$\forall g \in \mathcal{B}_V, \quad \|\widetilde{g}\|_{V_0} \leq a\|g\|_V \quad \textit{with here} \quad a := 1 + \max\left(0\,,\ \frac{K - \nu(V)}{\nu(1_{\mathbb{X}})(1-\delta)}\right)$$

   where $\widetilde{g} := \sum_{k=0}^{+\infty} R^k g$, so that

$$\forall g \in \mathcal{B}_V, \quad \|\widetilde{g}\|_V \leq \frac{a}{1-\delta}\,\|g\|_V \tag{17}$$

   where $\nu \in \mathcal{M}_*^+$ is given in ($\boldsymbol{S}$) and δ, K are the constants given in $\boldsymbol{G}(\delta, V)$.

2. *For every $g \in \mathcal{B}_V$ such that $\pi(g) = 0$, the function $\widehat{g} = \sum_{k=0}^{+\infty} P^k g$ is the unique $\pi-$centered function in $\mathcal{B}_V$ solution to Poisson's equation $(I-P)\widehat{g} = g$, with*

$$\|\widehat{g}\|_V \quad \leq \quad \frac{a\,(1+\pi(V))}{1-\delta}\,\|g\|_V. \tag{18}$$

3. *If* $\pi(V^2) < \infty$ *then, for any* $g \in \mathcal{B}_V$ *such that* $\pi(g) = 0$*, the asymptotic variance* $\gamma_g^2 = \pi((\widetilde{g})^2 - (P\widetilde{g})^2)$ *with* $\widetilde{g} := \sum_{k=}^{+\infty} R^k g$ *solution of Poisson's equation* $(I-P)\widetilde{g} = g$*, satisfies*

$$\gamma_g^2 \leq \frac{2\,a^2\pi(V^2)}{(1-\delta)^2}\,\|g\|_V^2.$$

For every $g \in \mathcal{B}_V$ such that $\pi(g) = 0$, the $V-$weighted norm of the unique $\pi-$centered solution $\widehat{g} = \sum_{k=0}^{+\infty} P^k g$ to Poisson equation $(I-P)\widehat{g} = g$ can be directly bounded using the $V-$geometric ergodicity, i.e. $\exists C > 0$, $\exists \rho \in (0,1)$, $\|P^k g\|_V \leq C\rho^k \|g\|_V$. However the constants C and ρ are often unknown or badly estimated. By contrast the bound (17) is explicit. In this geometric ergodicity context, bounds (17) and (18) are similar to those in Equations (35) and (36a) in [HL24] obtained for the norm $\|\cdot\|_{V^{\alpha_0}}$ for some $\alpha_0 \in (0,1]$. Actually the method from [HL24] consists in converting the $V-$geometric drift condition $\boldsymbol{G}(\delta, V)$ into the following residual-type geometric drift condition $RV^{\alpha_0} \leq \delta^{\alpha_0} V^{\alpha_0}$. When the positive constant K in $\boldsymbol{G}(\delta, V)$ is such that $K \leq \nu(V)$ (in particular in the atomic case), the previous residual-type drift condition holds with $\alpha_0 = 1$: in this case the bounds obtained in Equations (35) and (36a) in [HL24] for $\|\widetilde{g}\|_V$ and $\|\widehat{g}\|_V$ are exactly (17) and (18) with $a = 1$. By contrast, if α_0 is close to zero (i.e. δ^{α_0} is close to one), then the bounds in [HL24] degrade since they depend on $(1-\delta^{\alpha_0})^{-1}$. In this case, (17) and (18) are alternative bounds for $\|\widetilde{g}\|_V$ and $\|\widehat{g}\|_V$.

3 General perturbation results

In this section we deal with the quantitative control of the deviation between the invariant probability measures of Markov kernels. Let us first present a preliminary statement based on Theorem 2.4.

Proposition 3.1 *Assume that* P *satisfies Conditions* $(\boldsymbol{S})$-$\boldsymbol{D}(V_0, V_1)$*, and let* $\pi \equiv \pi_{\nu,R}$ *be the* $P-$*invariant probability measure given in (8). Let* P' *be another Markov kernel on* $(\mathbb{X}, \mathcal{X})$ *with some invariant probability measure* π' *such that* $\|P'V_0\|_{V_0} < \infty$ *and* $\pi'(V_0) < \infty$*. Finally assume that the non-negative function* Δ_{V_0} *defined by*

$$\forall x \in \mathbb{X}, \quad \Delta_{V_0}(x) := \|P(x,\cdot) - P'(x,\cdot)\|_{V_0}$$

is measurable on $(\mathbb{X}, \mathcal{X})$*. Then*

$$\|\pi' - \pi\|_{V_1} \leq a(1 + \pi(V_1))\,\pi'(\Delta_{V_0}) \tag{19}$$

where the positive constant a *is defined in (14).*

Under the assumptions of Proposition 3.1, the function Δ_{V_0} is well-defined and everywhere finite on $\mathbb{X}$ since so are PV_0 and $P'V_0$. Moreover, if the $\sigma-$algebra $\mathcal{X}$ is countably generated, then Δ_{V_0} is measurable on $(\mathbb{X}, \mathcal{X})$. Indeed, for every $x \in \mathbb{X}$ we have $\|P(x,\cdot) - P'(x,\cdot)\|_{V_0} = |\eta_x|(V_0)$ where $|\eta_x|$ is the total variation measure of the finite signed measure $\eta_x = P(x,\cdot) - P'(x,\cdot)$. Moreover the map $x \mapsto |\eta_x|(V_0)$ is measurable since so is $x \mapsto \eta_x(V_0)$, see [DF64].

Proof. Recall that $\|PV_0\|_{V_0} < \infty$ from $\boldsymbol{D}(V_0, V_1)$, so that Δ_{V_0} and $\pi'(\Delta_{V_0})$ are well-defined under the assumptions of Proposition 3.1.

Let $g \in \mathcal{B}_{V_1}$ such that $\|g\|_{V_1} \leq 1$. Since $\pi(V_1) < \infty$ from Theorem 2.4, $\pi(g)$ is well-defined. Define $g_0 = g - \pi(g)1_{\mathbb{X}}$ and $\widetilde{g_0} := \sum_{k=0}^{+\infty} R^k g_0$ with the residual kernel $R := P - \nu(\cdot)1_S$. Then we have

$$\pi'\big((P' - P)\widetilde{g_0}\big) = \pi'(\widetilde{g_0}) - \pi'(\widetilde{g_0} - g_0) = \pi'(g_0) = \pi'(g) - \pi(g) \tag{20}$$

using the P'−invariance of π', the Poisson equation $(I - P)\widetilde{g_0} = g_0$ from Theorem 2.4, and finally the definition of g_0. It follows from the definition of Δ_{V_0} that

$$|\pi'(g) - \pi(g)| \leq \int_{\mathbb{X}} \big|(P'\widetilde{g_0})(x) - (P\widetilde{g_0})(x)\big|\, d\pi'(x) \leq \|\widetilde{g_0}\|_{V_0} \int_{\mathbb{X}} \Delta_{V_0}(x)\, d\pi'(x).$$

Finally we know from Theorem 2.4 that $\|\widetilde{g_0}\|_{V_0} \leq a\|g_0\|_{V_1}$ with a defined in (14), so that

$$\|\widetilde{g_0}\|_{V_0} \leq a\, \|g - \pi(g)1_{\mathbb{X}}\|_{V_1} \leq a\big(1 + \pi(V_1)\|1_{\mathbb{X}}\|_{V_1}\big)$$

from which we deduce (19) since $\|1_{\mathbb{X}}\|_{V_1} \leq 1$. □

Now let $\{P_\theta\}_{\theta\in\Theta}$ denote a family of transition kernels on $(\mathbb{X}, \mathcal{X})$, where Θ is an open subset of some metric space. Let us introduce the following minorization and modulated drift conditions w.r.t. this family $\{P_\theta\}_{\theta\in\Theta}$:

$$\forall\theta \in \Theta,\ \exists S_\theta \in \mathcal{X},\ \exists\nu_\theta \in \mathcal{M}_*^+,\ \forall x \in \mathbb{X},\ \forall A \in \mathcal{X}, \quad P_\theta(x, A) \geq \nu_\theta(1_A)\, 1_{S_\theta}(x), \tag{$\boldsymbol{S_\Theta}$}$$

and there exists a couple (V_0, V_1) of Lyapunov functions such that

$$\forall\theta \in \Theta,\ \exists b_\theta > 0, \quad P_\theta V_0 \leq V_0 - V_1 + b_\theta 1_{S_\theta}. \tag{$\boldsymbol{D_\Theta}(V_0, V_1)$}$$

Let us fix some $\theta_0 \in \Theta$. The family $\{P_\theta, \theta \in \Theta \setminus \{\theta_0\}\}$ must be thought of as a family of transition kernels which are perturbations of P_{θ_0} and converge (in a certain sense) to P_{θ_0} when $\theta \to \theta_0$. To that effect, under the Conditions ($\boldsymbol{S_\Theta}$)-$\boldsymbol{D_\Theta}(V_0, V_1)$ we define

$$\forall\theta \in \Theta,\ \forall x \in \mathbb{X}, \quad \Delta_{\theta,V_0}(x) := \|P_\theta(x, \cdot) - P_{\theta_0}(x, \cdot)\|_{V_0}. \tag{21}$$

Finally, under the additional conditions $\sup_{\theta\in\Theta} b_\theta < \infty$ and $\inf_{\theta\in\Theta} \nu_\theta(1_{\mathbb{X}}) > 0$, we define the following positive constant

$$a := 1 + \max\big(0, c\big) \quad \text{with} \quad c := \sup_{\theta\in\Theta} \frac{b_\theta - \nu_\theta(V_0)}{\nu_\theta(1_{\mathbb{X}})}. \tag{22}$$

In Theorem 3.2 and Corollary 3.3 below, each Markov kernel P_θ is assumed to satisfy the assumptions of Theorem 2.4. Accordingly, the P_θ−invariant probability measure denoted by π_θ in these two statements is $\pi_\theta \equiv \pi_{\theta,\nu,R}$, i.e. the probability measure given by (8) with $\nu = \nu_\theta$ and $R_\theta := P_\theta - \nu_\theta(\cdot)1_{S_\theta}$). Since π_θ is assumed below to satisfy $\pi_\theta(V_0) < \infty$, we know from the first assertion of Corollary 2.6 that there is no ambiguity about what π_θ is in Theorem 3.2 and Corollary 3.3. To avoid any measurability problems for the functions Δ_{θ,V_0} in the next theorem, the σ−algebra $\mathcal{X}$ is assumed to be countably generated.

Theorem 3.2 *Assume that the family $\{P_\theta\}_{\theta\in\Theta}$ satisfies Conditions ($\boldsymbol{S_\Theta}$)-$\boldsymbol{D_\Theta}(V_0, V_1)$ with $b := \sup_{\theta\in\Theta} b_\theta < \infty$ and $\inf_{\theta\in\Theta} \nu_\theta(1_{\mathbb{X}}) > 0$. Moreover assume that, for every $\theta \in \Theta$, the P_θ−invariant probability measure $\pi_\theta \equiv \pi_{\theta,\nu,R}$ provided by (8) satisfies $\pi_\theta(V_0) < \infty$.*
Then we have

$$\|\pi_\theta - \pi_{\theta_0}\|_{V_1} \leq a \min\big\{c_{\theta_0}\, \pi_\theta(\Delta_{\theta,V_0})\, ,\, c_\theta\, \pi_{\theta_0}(\Delta_{\theta,V_0})\big\} \tag{23}$$

with a defined in (22) and with

$$\forall \theta \in \Theta, \quad c_\theta := 1 + \pi_\theta(V_1)\|1_{\mathbb{X}}\|_{V_1} \leq 1 + b. \tag{24}$$

If the following additional assumption holds

$$\forall x \in \mathbb{X}, \quad \lim_{\theta \to \theta_0} \Delta_{\theta,V_0}(x) = 0, \tag{$\boldsymbol{\Delta_{V_0}}$}$$

then we have

$$\lim_{\theta \to \theta_0} \|\pi_\theta - \pi_{\theta_0}\|_{V_1} = 0.$$

It may seem surprising to obtain such a statement without any assumption on the rate of convergence of the iterates of P_θ. But actually, under the conditions of Theorem 3.2 combined with the standard aperiodicity and irreducibility assumptions, it is well-known from Theorem 14.0.1 in [MT09] that, for every $\theta \in \Theta$ and for every $x \in \mathbb{X}$, the series $\sum_{k\geq 0} \|P_\theta^k(x,\cdot) - \pi\|_{V_1}$ converges. However note that this result is not used in the proof of Theorem 3.2.

Proof. Let $\theta \in \Theta$. We have $P_\theta V_0 \leq (1+b)V_0$ from $\boldsymbol{D_\Theta}(V_0, V_1)$ and the definition of the positive constant b. Thus Proposition 3.1 can be applied to $(P, P') := (P_{\theta_0}, P_\theta)$ and to $(P, P') := (P_\theta, P_{\theta_0})$. This provides Inequality (23). Also observe that the bound in (24) follows from the inequality $\pi_\theta(V_1) \leq b_\theta \leq b$ which is easily deduced from $\boldsymbol{D_\Theta}(V_0, V_1)$ using $\pi_\theta(P_\theta V_0) = \pi_\theta(V_0)$ (recall that $\pi_\theta(V_0) < \infty$ by hypothesis). Next we have

$$\lim_{\theta \to \theta_0} \pi_{\theta_0}(\Delta_{\theta,V_0}) = \lim_{\theta \to \theta_0} \int_{\mathbb{X}} \Delta_{\theta,V_0}(x) d\pi_{\theta_0}(x) = 0 \tag{25}$$

from Lebesgue's theorem using $\Delta_{\theta,V_0} \leq 2(1+b)V_0$, $\pi_{\theta_0}(V_0) < \infty$ and Assumption $(\boldsymbol{\Delta_{V_0}})$. Then we obtain that $\lim_{\theta \to \theta_0} \|\pi_\theta - \pi_{\theta_0}\|_{V_1} = 0$ from the second bound in (23) and from the inequality (24). □

When Condition $\boldsymbol{D_\Theta}(V_0, V_1)$ is satisfied, so is Condition $\boldsymbol{D_\Theta}(V_0, 1_{\mathbb{X}})$ since $V_1 \geq 1_{\mathbb{X}}$. Thus, when Theorem 3.2 applies, then it also applies with $V_1 := 1_{\mathbb{X}}$ and then provides the control of the total variation error since $\|\pi_\theta - \pi_{\theta_0}\|_{TV} = \|\pi_\theta - \pi_{\theta_0}\|_{1_{\mathbb{X}}}$. Using $\pi_\theta(1_{\mathbb{X}}) = 1$, $\|1_{\mathbb{X}}\|_{1_{\mathbb{X}}} = 1$, so that we have here $c_\theta := 1 + \pi_\theta(1_{\mathbb{X}})\|1_{\mathbb{X}}\|_{1_{\mathbb{X}}} = 2$, we obtain the following estimate for $\|\pi_\theta - \pi_{\theta_0}\|_{TV}$.

Corollary 3.3 *Under the assumptions of Theorem 3.2 we have*

$$\|\pi_\theta - \pi_{\theta_0}\|_{TV} \leq 2\,a\,\min\left\{\pi_\theta(\Delta_{\theta,V_0})\,,\, \pi_{\theta_0}(\Delta_{\theta,V_0})\right\} \tag{26}$$

with a defined in (22). If moreover $\{P_\theta\}_{\theta\in\Theta}$ satisfies Assumption $(\boldsymbol{\Delta_{V_0}})$, then we have $\lim_{\theta \to \theta_0} \|\pi_\theta - \pi_{\theta_0}\|_{TV} = 0$.

The convergence of $\pi_{\theta_0}(\Delta_{\theta,V_0})$ to 0 when $\theta \to \theta_0$ in (25) is of theoretical interest: it is used to prove that $\lim_{\theta \to \theta_0} \|\pi_\theta - \pi_{\theta_0}\|_{V_1} = 0$ in Theorem 3.2. Indeed it is worth noticing that this term $\pi_{\theta_0}(\Delta_{\theta,V_0})$ in the bounds (23) and (26) is not computable in practice since the probability measure π_{θ_0} may be considered as unknown in our perturbation context. By contrast, the value of $\pi_\theta(\Delta_{\theta,V_0})$ in bounds (23) and (26) is expected to be known or at least computable for $\theta \neq \theta_0$, so that the bounds of interest in (23) and (26) are

$$\|\pi_\theta - \pi_{\theta_0}\|_{V_1} \leq a\,c_{\theta_0}\,\pi_\theta(\Delta_{\theta,V_0}) \quad \text{and} \quad \|\pi_\theta - \pi_{\theta_0}\|_{TV} \leq 2\,a\,\pi_\theta(\Delta_{\theta,V_0}) \tag{27}$$

with c_{θ_0} given in (24). However the bounds in (27) are relevant only if $\lim_{\theta\to\theta_0}\pi_\theta(\Delta_{\theta,V_0})=0$, which is not guaranteed under the conditions of Theorem 3.2. For this purpose note that, in the proof of Theorem 3.2, the conditions ($\boldsymbol{S_\Theta}$) and $\boldsymbol{D_\Theta}(V_0,V_1)$ for P_θ with $\theta\neq\theta_0$ are only used for obtaining the inequality $\|\pi_\theta-\pi_{\theta_0}\|_{V_1}\le ac_\theta\,\pi_{\theta_0}(\Delta_{\theta,V_0})$ of (23). Consequently, if we are only interested in the two bounds in (27), then the assumptions of Theorem 3.2 can be relaxed as follows.

Proposition 3.4 *Assume that the (unperturbed) Markov kernel $P:=P_{\theta_0}$ satisfies Conditions* ($\boldsymbol{S}$) and $\boldsymbol{D}(V_0,V_1)$*. Moreover assume that, for every $\theta\in\Theta\setminus\{\theta_0\}$, we have $\|P_\theta V_0\|_{V_0}<\infty$ and that there exists a $P_\theta-$invariant probability measure π_θ on $(\mathbb{X},\mathcal{X})$ such that $\pi_\theta(V_0)<\infty$. Then the two bounds in (27) hold.*

Indeed, under the assumptions of Proposition 3.4, the first bound in (27) directly follows from Proposition 3.1 applied to $(P,P'):=(P_{\theta_0},P_\theta)$ with $\theta\neq\theta_0$. The second bound in (27) is obtained by replacing V_1 with $1_{\mathbb{X}}$. In Proposition 3.4, the existence of a $P_\theta-$invariant probability measure π_θ is required when $\theta\neq\theta_0$ since we do not assume that P_θ satisfies minorization and modulated drift condition for $\theta\neq\theta_0$. Actually, π_θ may be any $P_\theta-$invariant probability measure when $\theta\neq\theta_0$, while π_{θ_0} is the $P_{\theta_0}-$invariant probability measure given by (8). In any case the assumption $\pi_\theta(V_0)<\infty$ is required for every $\theta\in\Theta\setminus\{\theta_0\}$. Finally let's stress once again that the bounds in (27) are of interest only when the term $\pi_\theta(\Delta_{\theta,V_0})$ is computable and can be proved to converge to 0 when $\theta\to\theta_0$. To simply illustrate what the function Δ_{θ,V_0} and Condition ($\boldsymbol{\Delta_{V_0}}$) are in a concrete case, let us consider the standard issue of truncation of an infinite stochastic matrix $P=(P(x,y))_{x,y\in\mathbb{N}}$. For any $k\ge1$ let $B_k:=\{0,\ldots,k\}$ and ${B_k}^c:=\mathbb{N}\setminus B_k$. Recall that the k-th truncated and arbitrary augmented stochastic matrix P_k of the $(k+1)\times(k+1)$ north-west corner truncation of P is defined by (e.g. see [LL18]):

$$\forall(x,y)\in B_k\times B_k,\quad P_k(x,y):=P(x,y)+P(x,{B_k}^c)\,\psi_{x,k}(y)$$

where $\psi_{x,k}(\cdot)$ is some probability measure on B_k. Define the following extended Markov kernel $\widehat{P}_k$ of P_k on $\mathbb{N}$:

$$\forall(x,y)\in\mathbb{N}\times\mathbb{N},\quad \widehat{P}_k(x,y):=P_k(x,y)\,1_{B_k\times B_k}(x,y)+1_{B_k^c\times\{0\}}(x,y).$$

The problem is to approximate the invariant probability measure of P by that of the stochastic matrix $\widehat{P}_k$, that is roughly speaking, by that of the finite stochastic matrix P_k. This can be thought of as a perturbation issue introducing the family $\{P_\theta\}_{\theta\in\Theta}$ of Markov kernels with $\theta_0=+\infty$: $\Theta:=\big(\mathbb{N}\setminus\{0\}\big)\cup\{+\infty\}$, $P_{+\infty}:=P$ and $\forall\theta\ge1$, $P_\theta:=\widehat{P}_k$. Let us specify the quantity $\Delta_{k,V_0}(x)$ defined in (21) for some function $V_0\ge1$. From the definitions of P_k and $\widehat{P}_k$, we have for every $x\in B_k$

$$\Delta_{k,V_0}(x) = \sum_{y\in\mathbb{N}}\big|P(x,y)-\widehat{P}_k(x,y)\big|V_0(y)=P(x,{B_k}^c)\sum_{y\in B_k}\psi_{x,k}(y)V_0(y)+\sum_{y\in {B_k}^c}P(x,y)V_0(y).$$

When V_0 is non-decreasing, the following control of $\Delta_{k,V_0}(x)$ for $x\in B_k$ is then easily obtained:

$$\begin{aligned}\forall x\in B_k,\quad \Delta_{k,V_0}(x) &\le P(x,{B_k}^c)\,V_0(k)+\sum_{y\in {B_k}^c}P(x,y)V_0(y)\\ &\le \sum_{z\in {B_k}^c}P(x,z)V_0(z)+\sum_{y\in {B_k}^c}P(x,y)V_0(y)\ \le 2\sum_{y\in {B_k}^c}P(x,y)V_0(y).\end{aligned}$$

Let $x \in \mathbb{N}$ be fixed. Using this control for any $k > x$, we obtain the convergence $\lim_{k\to+\infty} \Delta_{k,V_0}(x) = 0$ from $(PV_0)(x) = \sum_{y\in\mathbb{N}} P(x,y)V_0(y) < \infty$, i.e. Condition $(\boldsymbol{\Delta_{V_0}})$ holds true. Finally note that the $\widehat{P}_k-$probability measure π_k is supported in B_k and computable using the finite matrix P_k, so that the term $\pi_\theta(\Delta_{\theta,V_0}) = \sum_{x\in B_k} \pi_k(x)\Delta_{k,V_0}(x)$ used in (27) can be easily bounded using the previous inequality on $\Delta_{k,V_0}(x)$.

Remark 3.5 (Stability issue) *In some classical perturbation schemes, as the standard truncations of infinite stochastic matrices or the state space discretization procedure of non-discrete models, the whole family* $\{P_\theta\}_{\theta\in\Theta}$ *satisfies Conditions* $(\boldsymbol{S_\Theta})$-$\boldsymbol{D_\Theta}(V_0, V_1)$ *provided that the unperturbed Markov kernel* $P := P_{\theta_0}$ *satisfies Conditions* $(\boldsymbol{S})$-$\boldsymbol{D}(V_0, V_1)$*. Moreover the set* S *and the constant* b *involved for* $P := P_{\theta_0}$ *in* $(\boldsymbol{S})$-$\boldsymbol{D}(V_0, V_1)$ *can often be used for the perturbed Markov kernels* P_θ*. In this case the conditions* $b := \sup_{\theta\in\Theta} b_\theta < \infty$ *and* $\inf_{\theta\in\Theta} \nu_\theta(1_\mathbb{X}) > 0$ *of Theorem 3.2 and Corollary 3.3 are straightforward. In the context of geometric drift conditions, the previous facts are proved to hold in many papers for truncations of infinite stochastic matrices (e.g. see [LL18, HL14] and references therein), in [HL21] for the state space discretization procedure. The case of Markov models satisfying modulated drift conditions can be addressed similarly.*

Remark 3.6 *If* P_θ *is replaced with iterate* P_θ^ℓ *for some* $\ell \geq 2$ *in Conditions* $(\boldsymbol{S_\Theta})$-$\boldsymbol{D_\Theta}(V_0, V_1)$ *and if for every* $\theta \in \Theta$ *both* P_θ *and* P_θ^ℓ *admit a unique invariant probability measure* π_θ*, then all the previous perturbation results still hold replacing* $\pi_\theta(\Delta_{\theta,V_0})$ *with* $\pi_\theta(\Delta_{\ell,\theta,V_0})$*, where*

$$\forall x \in \mathbb{X}, \quad \Delta_{\ell,\theta,V_0}(x) := \|P_\theta^\ell(x,\cdot) - P_{\theta_0}^\ell(x,\cdot)\|_{V_0}.$$

Indeed, under the previous assumptions, Theorem 3.2 and Corollary 3.3 obviously apply to the family $\{P_\theta^\ell\}_{\theta\in\Theta}$*. The same remark is valid in Proposition 3.4 when* $P_{\theta_0}^\ell$ *satisfies Conditions* $(\boldsymbol{S})$ *and* $\boldsymbol{D}(V_0, V_1)$ *for some* $\ell \geq 2$*.*

Our perturbation results are discussed through the two following examples.

Example 3.7 (Geometric drift conditions) *In the perturbation context, under Condition* $(\boldsymbol{S_\Theta})$*, the standard geometric drift conditions for some Lyapunov function* V *are the following ones:*

$$\forall \theta \in \Theta, \quad K_\theta := \sup_{x\in S_\theta} (P_\theta V)(x) < \infty \quad \text{and} \quad \delta_\theta := \sup_{x\in S_\theta^c} \frac{(P_\theta V)(x)}{V(x)} \in (0,1). \tag{28}$$

In addition to Assumptions $(\boldsymbol{S_\Theta})$ *and (28), we assume that, for every* $\theta \in \Theta$*, we have* $\nu_\theta(1_{S_\theta}) > 0$ *where* $(S_\theta, \nu_\theta) \in \mathcal{X} \times \mathcal{M}_*^+$ *is given in* $(\boldsymbol{S_\Theta})$*, so that each* P_θ *is* $V-$*geometrically ergodic, with unique* $P_\theta-$*invariant probability measure denoted by* π_θ *satisfying* $\pi_\theta(V) < \infty$ *(e.g. see [Bax05]). Moreover assume that* $K := \sup_{\theta\in\Theta} K_\theta < \infty$ *and* $\delta := \sup_{\theta\in\Theta} \delta_\theta < 1$*. Then*

$$\forall \theta \in \Theta, \quad P_\theta V \leq \delta V + K 1_{S_\theta} \leq V - (1-\delta)V + K1_{S_\theta}.$$

Note that the second inequality reads as the Condition $\boldsymbol{D}(V_0, V)$*,* $P_\theta V_0 \leq V_0 - V + b1_S$*, with* $V_0 = V/(1-\delta)$*,* $V_1 = V$ *and* $b = K/(1-\delta)$*, so that Theorem 3.2 could be applied here to control* $\|\pi_\theta - \pi_{\theta_0}\|_V$*. Mention that the bound of Theorem 3.2 then provides a generalization of the bound (10) in [LL18] to the truncation of a transition kernel defined on a general* *state-space $\mathbb{X}$ without assuming the existence of an atom. Similarly the bound of Theorem 3.2 extends the bound (16) in [LL18] (with $m=1$) to a general state-space $\mathbb{X}$ without assuming that the residual kernel is a contraction on $\mathcal{B}_V$, i.e. $RV \le \beta V$ for some $\beta < 1$.*

The focus here is on the comparison of our results with Proposition 2.1 in [HL14] and Equation (3.19) in [RS18], so that it only concerns the geometric case. We only apply Corollary 3.3 in order to control the total variation norm $\|\pi_\theta - \pi_{\theta_0}\|_{TV}$. Hence, we only use the following Condition $\boldsymbol{D}(V_0, 1_{\mathbb{X}})$ derived from $\boldsymbol{D}(V_0, V)$ using $V \ge 1_{\mathbb{X}}$:

$$\forall \theta \in \Theta, \quad P_\theta V_0 \le V_0 - 1_{\mathbb{X}} + b\, 1_{S_\theta} \quad \text{with} \quad V_0 = \frac{V}{1-\delta} \quad \text{and} \quad b := \frac{K}{1-\delta}.$$

Therefore, if $m := \inf_{\theta\in\Theta} \nu_\theta(1_{\mathbb{X}}) > 0$, then $\{P_\theta\}_{\theta\in\Theta}$ satisfies the assumptions of Theorem 3.2 and we have from Corollary 3.3

$$\|\pi_\theta - \pi_{\theta_0}\|_{TV} \le \frac{2\,a}{1-\delta}\, \min\left\{\pi_\theta(\Delta_{\theta,V})\,,\, \pi_{\theta_0}(\Delta_{\theta,V})\right\} \quad \text{with} \quad a = 1 + \max\left(0, \frac{b}{m}\right) \tag{29}$$

using the fact that $\Delta_{\theta,V_0}(x) = \Delta_{\theta,V}(x)/(1-\delta)$. Moreover we have $\lim_{\theta\to\theta_0} \|\pi_\theta - \pi_{\theta_0}\|_{TV} = 0$, provided that Condition ($\boldsymbol{\Delta_{V_0}}$) is satisfied here with $V_0 := V$ (see Corollary 3.3). Recall that, if the term $\pi_\theta(\Delta_{\theta,V})$ can be computed and is proved to converge to 0 when $\theta \to \theta_0$, then the bound of interest in (29) is

$$\|\pi_\theta - \pi_{\theta_0}\|_{TV} \le \frac{2\,a}{1-\delta}\; \pi_\theta(\Delta_{\theta,V}) \tag{30}$$

and that (30) can be obtained under less restrictive assumptions focussing on P_{θ_0} by using Proposition 3.4 (see also Remark 3.5).

Now, let us compare Inequality (30) with the bound obtained in Proposition 2.1 in [HL14] and in Equation (3.19) in [RS18] (see also [HL23b] for the iterated function systems), that is

$$\|\pi_\theta - \pi_{\theta_0}\|_{TV} \le C\, \gamma_\theta \,\big|\ln \gamma_\theta\big| \quad \text{with} \quad \gamma_\theta := \sup_{x\in\mathbb{X}} \frac{\Delta_{\theta,1_{\mathbb{X}}}(x)}{V(x)} \tag{31}$$

where the positive constant C depends on the above constants δ, K and on the $V-$geometric rate of convergence of the iterates P_θ^n to the invariant distribution π_θ.

- *The interest of the bound (31) is that it uses $\Delta_{\theta,1_{\mathbb{X}}}(x)$ rather than $\Delta_{\theta,V}(x)$ in (30). Note that the supremum bound over $x \in \mathbb{X}$ in the definition of γ_θ only requires to consider this supremum on a level set $\{x \in \mathbb{X} : V(x) \le c\}$, observing that the supremum on the complementary set is arbitrarily small when c is large enough (use $\Delta_{\theta,1_{\mathbb{X}}}(x)/V(x) \le 2/c$ when $V(x) > c$).*
- *The drawback of (31) is that it involves a logarithm term, but above all that the constant C in (31) depends on the $V-$geometric rate of convergence of P_θ^n to π_θ, which is unknown in general or badly estimated.*

In conclusion, to prove that $\lim_{\theta\to\theta_0} \|\pi_\theta - \pi_{\theta_0}\|_{TV} = 0$, it is more relevant to use the results in [HL14, RS18]. However, if the term $\pi_\theta(\Delta_{\theta,V})$ can be computed for $\theta \ne \theta_0$ and if $\pi_\theta(\Delta_{\theta,V})$ converges to 0 when $\theta \to \theta_0$, then the bound (30) is much more relevant than (31) since the multiplicative constant in (30) is simple and easily computable, in contrast to that in (31).

Example 3.8 (Perturbed random walk on the half line) *Let Θ be some open metric space. For any $\theta \in \Theta$, let us consider the random walk $\{X_n^{(\theta)}\}_{n\in\mathbb{N}}$ on the half line $\mathbb{X} := [0, +\infty)$ given by*

$$X_0^{(\theta)} \in \mathbb{X} \quad \text{and} \quad \forall n \geq 1,\ X_n^{(\theta)} := \max\big(0, X_{n-1}^{(\theta)} + \varepsilon_n^{(\theta)}\big) \tag{32}$$

where $\{\varepsilon_n^{(\theta)}\}_{n\geq 1}$ is a sequence of i.i.d. $\mathbb{R}$-valued random variables assumed to be independent of $X_0^{(\theta)}$, and to have a common parametric probability density function $\mathfrak{p}_\theta$ w.r.t. the Lebesgue measure on $\mathbb{R}$. The transition kernel associated with $\{X_n^{\theta}\}_{n\in\mathbb{N}}$ is given by

$$\forall x \in \mathbb{X},\ \forall A \in \mathcal{X}, \quad P_\theta(x, A) = 1_A(0)\int_{-\infty}^{-x} \mathfrak{p}_\theta(y)\,dy + \int_{-x}^{+\infty} 1_A(x+y)\,\mathfrak{p}_\theta(y)\,dy. \tag{33}$$

Assume that

$$m_1 := \sup_{\theta\in\Theta} \mathbb{E}\big[|\varepsilon_1^{(\theta)}|\big] < \infty, \qquad \forall \theta \in \Theta,\ m_{2,\theta} := \mathbb{E}\big[|\varepsilon_1^{(\theta)}|^2\big] < \infty,\ \ \mathbb{E}\big[\varepsilon_1^{(\theta)}\big] < 0, \tag{34a}$$

and

$$\exists x_0 > 0 \text{ such that } c_0 := -\sup_{\theta\in\Theta} \int_{-x_0}^{+\infty} y\,\mathfrak{p}_\theta(y)\,dy > 0. \tag{34b}$$

Under the negative moment condition $\mathbb{E}\big[\varepsilon_1^{(\theta)}\big] < 0$, there always exists a positive scalar $x_0(\theta) > 0$ such that $\int_{-x_0(\theta)}^{+\infty} y\,\mathfrak{p}_\theta(y)\,dy < 0$. Thus Assumption (34b) means that a scalar $x_0 > 0$, uniform w.r.t. $\theta \in \Theta$, can be chosen. Moreover, under the second order moment condition in (34a), we know from Proposition 3.5 in [JT03] that P_θ has a unique invariant probability measure π_θ and that $\pi_\theta(L_0) < \infty$ where $L_0(x) := 1 + x$ for any $x \in \mathbb{X}$. Let us introduce the following functions on $\mathbb{X}$:

$$\forall x \in \mathbb{X}, \quad V_0(x) := \frac{L_0(x)}{c} = \frac{1+x}{c} \quad \text{and} \quad V_1(x) := 1$$

where $c := \min(1, c_0)$. Next, assume that the function defined by

$$\forall y \in \mathbb{R}, \quad h(y) := \inf_{\theta\in\Theta} \inf_{x\in[0,x_0]} \mathfrak{p}_\theta(y - x)$$

with x_0 given by (34b), is positive on some open interval of $\mathbb{R}$. Then,

1. *Condition ($\boldsymbol{S_\Theta}$) is satisfied with $S_\theta := [0, x_0]$ and $\nu_\theta := \nu$, where ν is the positive measure on $\mathbb{R}$ defined by*

$$\forall A \in \mathcal{X}, \quad \nu(1_A) := \int_{\mathbb{X}} 1_A(y)\,h(y)\,dy.$$

2. *Condition $\boldsymbol{D_\Theta}(V_0, V_1)$ holds with $S_\theta = [0, x_0]$ and $b_0 = (m_1 + c_0)/c$. In particular it follows from Theorem 2.4 that the unique $P_\theta-$invariant probability measure π_θ derived from [JT03] has the representation (8).*

3. Let us fix $\theta_0 \in \Theta$. Define

$$\forall \theta \in \Theta, \quad \forall y \in \mathbb{R}, \quad \rho_\theta(y) := |\mathfrak{p}_\theta(y) - \mathfrak{p}_{\theta_0}(y)|, \tag{35a}$$

$$e_\theta := \int_{\mathbb{R}} \rho_\theta(y)\, dy \quad \text{and} \quad e_{1,\theta} := \int_{\mathbb{R}} |y|\, \rho_\theta(y)\, dy. \tag{35b}$$

Then Condition $(\boldsymbol{\Delta_{V_0}})$ of Corollary 3.3 holds provided that

$$\lim_{\theta \to \theta_0} (e_\theta + e_{1,\theta}) = 0$$

and we have

$$\forall \theta \in \Theta, \quad \pi_\theta(\Delta_{\theta,V_0}) \leq \frac{e_\theta(1 + c\, \pi_\theta(V_0)) + e_{1,\theta}}{c}. \tag{36}$$

The details of the checking are postponed to Appendix B. Thus, Corollary 3.3 apply under Assumptions (34a)-(34b) on the noise process $\{\varepsilon_n^{(\theta)}\}_{n\geq 1}$, and the following bound (see (27))

$$\|\pi_\theta - \pi_{\theta_0}\|_{TV} \leq 2\,(1+d)\ \pi_\theta(\Delta_{\theta,V_0}) \qquad \text{with } d = \max\left(0, \frac{b_0 - \nu(V_0)}{\nu(1_{\mathbb{X}})}\right)$$

is of interest, provided that the quantities e_θ, $e_{1,\theta}$ and $\pi_\theta(V_0)$ are computable for $\theta \neq \theta_0$ and that both e_θ and $e_{1,\theta}$ converge to 0 when $\theta \to \theta_0$.

Let $m \geq 2$. Under conditions (34a) expressed in terms of moments of order $(m-1)$ and m on the noise process and under Condition (34b), it can be shown that Conditions $(\boldsymbol{S_\Theta})$-$\boldsymbol{D_\Theta}(V_0, V_1)$ also hold for an appropriate set S and the functions $V_1(x) = (1+x)^{m-2}$, $V_0(x) = L_0(x)/c$ where $L_0(x) = (1+x)^{m-1}$. The proof follows similar lines than that for $m = 2$ in Appendix B. From such extensions, it can be proved that the term $\pi_\theta(V_0)$ in (36) satisfies $\sup_{\theta\in\Theta} \pi_\theta(V_0) < \infty$, provided that the condition $m_1 < \infty$ in (34a) is replaced with the stronger one: $m_2 := \sup_{\theta\in\Theta} \mathbb{E}[|\varepsilon_1^{(\theta)}|^2] < \infty$.

A $P-$invariant probability measure under Condition $(\boldsymbol{S})$

The analytic proof of Recall 2.1 from [HL23a] is reported below. Note that it does not need to introduce the concepts of irreducibility, recurrence, atom or splitted chain associated with the Markov kernel.

Let P satisfy Condition $(\boldsymbol{S})$ and T be the following kernel

$$\forall x \in \mathbb{X},\ \forall A \in \mathcal{X}, \quad T(x, A) := \nu(1_A)\, 1_S(x)$$

so that $R = P - T$. Note that, for every $k \geq 1$, we have $\nu\, R^{k-1} \in \mathcal{M}^+$. Recall that for two nonnegative kernels K_1 and K_2, the inequality $K_1 \leq K_2$ means that for any measurable nonnegative function g, $K_1 g \leq K_2 g$. Set $T_0 := 0$ and $T_n := P^n - R^n$ for $n \geq 1$. Then

$$\forall n \geq 1,\ 0 \leq T_n \leq P^n,\ T_n - T_{n-1}P = (P^{n-1} - T_{n-1})T \text{ and } T_n = \sum_{k=1}^{n} \nu(R^{k-1}\cdot)P^{n-k}1_S. \tag{37}$$

The first property follows from $0 \leq R \leq P$. The second one is deduced from $P^n - T_n = (P^{n-1} - T_{n-1})(P - T)$. Finally, the last one is clear for $n = 1$ and it holds for $n \geq 2$ by an easy induction based on $T_n = P^{n-1}T + T_{n-1}R$.

Now, let us prove Recall 2.1. Assume that Assertion 1 holds. We deduce from (37) that $0 \leq \pi\big((P^n - T_n)1_{\mathbb{X}}\big) = 1 - \pi(T_n 1_{\mathbb{X}}) = 1 - \pi(1_S)\sum_{k=1}^{n} \nu(R^{k-1}1_{\mathbb{X}})$ from which it follows that $\sum_{k=1}^{+\infty} \nu(R^{k-1}1_{\mathbb{X}}) \leq \pi(1_S)^{-1} < \infty$ since $\pi(1_S) > 0$ by hypothesis. This gives Assertion 2. Conversely, if Assertion 2 holds, then $\mu := \sum_{k=1}^{+\infty} \nu R^{k-1} \in \mathcal{M}_*^+$ since $\mu(1_{\mathbb{X}}) \geq \nu(1_{\mathbb{X}}) > 0$. Moreover we have

$$\begin{aligned}
\forall A \in \mathcal{X}, \quad \mu(P1_A) &= \sum_{k=1}^{+\infty} \nu\big(P^k 1_A - T_{k-1}P1_A\big) \quad \text{from } R^{k-1} = P^{k-1} - T_{k-1} \\
&= \sum_{k=1}^{+\infty} \nu\big(P^k 1_A - T_k 1_A\big) + \sum_{k=1}^{+\infty} \nu\big(P^{k-1}T1_A - T_{k-1}T1_A\big) \quad \text{from (37)} \\
&= \mu(1_A) + \mu(T1_A) - \nu(1_A) \\
&= \mu(1_A) + \nu(1_A)\mu(1_S) - \nu(1_A) \qquad \text{from the definition of } T.
\end{aligned}$$

With $A = \mathbb{X}$ we obtain that $0 = \nu(1_{\mathbb{X}})\mu(1_S) - \nu(1_{\mathbb{X}})$, thus $\mu(1_S) = 1$ since $\nu(1_{\mathbb{X}}) > 0$. Consequently μ is $P-$invariant, so that $\pi := \mu(1_{\mathbb{X}})^{-1}\mu$ is an $P-$invariant distribution such that $\pi(1_S) = \mu(1_{\mathbb{X}})^{-1} > 0$.

B Conditions $(S_\Theta) - D_\Theta(V_0, V_1)$ for perturbed random walks on the half line

Let us assume that Conditions (34a)-(34b) hold, and that the function $h(\cdot)$ defined by

$$\forall y \in \mathbb{R}, \quad h(y) := \inf_{\theta \in \Theta} \inf_{x \in [0, x_0]} \mathfrak{p}_\theta(y - x)$$

with x_0 given by (34b), is positive on some open interval of $\mathbb{X} = [0, +\infty)$. Then, using the definition (33) of the kernel P_θ , we can write

$$\begin{aligned}
\forall \theta \in \Theta, \forall x \in [0, x_0],\ \forall A \in \mathcal{X}, \quad P_\theta(x, A) &= 1_A(0)\int_{-\infty}^{-x} \mathfrak{p}_\theta(y)\, dy + \int_0^{+\infty} 1_A(y)\, \mathfrak{p}_\theta(y - x)\, dy \\
&\geq \int_0^{+\infty} 1_A(y)\, \mathfrak{p}_\theta(y - x)\, dy \geq \int_0^{+\infty} 1_A(y)\, h(y)\, dy
\end{aligned}$$

so that Condition $(\boldsymbol{S_\Theta})$ is satisfied with $S_\theta \equiv [0, x_0]$ and the positive measure defined by $\nu(1_A) := \int_{\mathbb{X}} 1_A(y)\, h(y)\, dy$ for any $A \in \mathcal{X}$.

Next, recall that the functions L_0, V_0 and V_1 are defined as : for any $x \in \mathbb{X}$, $L_0(x) := 1 + x, V_0(x) := L_0(x)/c$ with $c := \min(1, c_0)$ (see (34b)) and $V_1(x) = 1$. Then we have from the definition (33) of P_θ:

$$\begin{aligned}
\forall x \in \mathbb{X}, \quad (P_\theta L_0)(x) - L_0(x) &= \int_{-\infty}^{-x} \mathfrak{p}_\theta(y)\, dy + \int_{-x}^{+\infty} (1 + x + y)\, \mathfrak{p}_\theta(y)\, dy - (1 + x) \\
&= -x \int_{-\infty}^{-x} \mathfrak{p}_\theta(y)\, dy + \int_{-x}^{+\infty} y\, \mathfrak{p}_\theta(y)\, dy \\
&\leq \int_{-x}^{+\infty} y\, \mathfrak{p}_\theta(y)\, dy. \qquad (38)
\end{aligned}$$

Then we obtain from (38) and (34b)

$$\begin{aligned} \forall x > x_0, &\quad (P_\theta L_0)(x) - L_0(x) \leq -c_0 V_1(x) \\ \text{and } \forall x \in [0, x_0], &\quad (P_\theta L_0)(x) - L_0(x) + c_0 V_1(x) \leq c_0 V_1(x) + m_1 = c_0 + m_1, \end{aligned}$$

that is

$$P_\theta L_0 \leq L_0 - c_0 V_1 + (c_0 + m_1)\, 1_{[0,x_0]}$$

or

$$P_\theta V_0 \leq V_0 - V_1 + b_0 1_{[0,x_0]} \tag{39}$$

with $b_0 := (c_0 + m_1)/c$. Thus the familly of kernels $\{P_\theta\}_{\theta\in\Theta}$ satisfies Condition $\boldsymbol{D}_{\boldsymbol{\Theta}}(V_0, V_1)$.

Next, we investigate the function $x \mapsto \Delta_{\theta,V_0}(x)$ and the quantity $\pi_\theta(\Delta_{\theta,V_0})$. To that effect, fix some $\theta_0 \in \Theta$. Here the state space is $\mathbb{X} := [0, +\infty)$ equipped with its Borel $\sigma-$algebra $\mathcal{X}$ which is countably generated. Therefore for any $\theta \in \Theta$, the non-negative function on $\mathbb{X}$, $x \mapsto \Delta_{\theta,V_0}(x) := \|P_\theta(x, \cdot) - P_{\theta_0}(x, \cdot)\|_{V_0}$, is $\mathcal{X}-$measurable. We use the quantities $\rho_\theta, e_\theta, e_{1,\theta}$ in (35a)-(35b). Note that $e_\theta \leq 2$. Let $g \in \mathcal{B}_{V_0}$ be such that $|g| \leq V_0$. Then we have

$$\begin{aligned} \forall x \in \mathbb{X}, \quad \big|(P_\theta g)(x) - (P_{\theta_0} g)(x)\big| &\leq V_0(0) \int_{-\infty}^{-x} \rho_\theta(y)\, dy + \int_{-x}^{+\infty} V_0(x+y)\, \rho_\theta(y)\, dy \\ &\leq \frac{e_\theta}{c} + \frac{1}{c} \int_{\mathbb{R}} \big(1 + x + |y|\big)\, \rho_\theta(y)\, dy \\ &\leq \frac{e_\theta}{c} + e_\theta V_0(x) + \frac{e_{1,\theta}}{c}. \end{aligned}$$

Thus

$$\forall x \in \mathbb{X}, \quad \Delta_{\theta,V_0}(x) \leq \frac{e_\theta(1 + c\, V_0(x)) + e_{1,\theta}}{c}. \tag{40}$$

Therefore Condition $(\boldsymbol{\Delta}_{\boldsymbol{V_0}})$ of Corollary 3.3 holds provided that

$$\lim_{\theta \to \theta_0} \big(e_\theta + e_{1,\theta}\big) = 0.$$

Such a condition ensures that $\lim_{\theta\to\theta_0} \|\pi_\theta - \pi_{\theta_0}\|_{TV} = 0$. Finally we have from (40)

$$\forall \theta \in \Theta, \quad \pi_\theta(\Delta_{\theta,V_0}) \leq \frac{e_\theta(1 + c\, \pi_\theta(V_0)) + e_{1,\theta}}{c}.$$

Recall that $\pi_\theta(V_0) < \infty$ under Assumptions (34a)-(34b) for the noise process $\{\varepsilon_n^{(\theta)}\}_{\theta\in\Theta}$.

References

[AANQ04] E. Altman, K. E. Avrachenkov, and R. Núñez-Queija. Perturbation analysis for denumerable Markov chains with application to queueing models. *Adv. in Appl. Probab.*, 36(3):839–853, 2004.

[AFEB16] P. Alquier, N. Friel, R. Everitt, and A. Boland. Noisy Monte Carlo: convergence of Markov chains with approximate transition kernels. *Stat. Comput.*, 26(1-2):29–47, 2016.

[AFV15] C. Andrieu, G. Fort, and M. Vihola. Quantitative convergence rates for subgeometric Markov chains. *J. Appl. Probab.*, 52(2):391–404, 2015.

[Bax05] P. H. Baxendale. Renewal theory and computable convergence rates for geometrically ergodic Markov chains. *Ann. Appl. Probab.*, 15(1B):700–738, 2005.

[Del17] B. Delyon. Convergence rate of the powers of an operator. Applications to stochastic systems. *Bernoulli*, 23(4A):2129–2180, 2017.

[DF64] L. Dubins and D. Freedman. Measurable sets of measures. *Pacific J. Math.*, 14:1211–1222, 1964.

[DFM16] A. Durmus, G. Fort, and E. Moulines. Subgeometric rates of convergence in Wasserstein distance for Markov chains. *Ann. Inst. Henri Poincaré Probab. Stat.*, 52(4):1799–1822, 2016.

[DFMS04] 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.

[DMPS18] R. Douc, E. Moulines, P. Priouret, and P. Soulier. *Markov Chains.* Springer, 2018.

[FHL13] D. Ferré, L. Hervé, and J. Ledoux. Regular perturbation of V-geometrically ergodic Markov chains. *J. Appl. Probab.*, 50(1):184–194, 2013.

[FM00] G. Fort and E. Moulines. V-subgeometric ergodicity for a Hastings-Metropolis algorithm. *Statist. Probab. Lett.*, 49(4):401–410, 2000.

[FM03] G. Fort and E. Moulines. Polynomial ergodicity of Markov transition kernels. *Stochastic Process. Appl.*, 103(1):57–99, 2003.

[GM96] P. W. Glynn and S. P. Meyn. A Lyapounov bound for solutions of the Poisson equation. *Ann. Probab.*, 24(2):916–931, 1996.

[HL14] L. Hervé and J. Ledoux. Approximating Markov chains and V-geometric ergodicity via weak perturbation theory. *Stochastic Process. Appl.*, 124(1):613–638, 2014.

[HL21] L. Hervé and J. Ledoux. State-discretization of V-geometrically ergodic Markov chains and convergence to the stationary distribution. *Methodol. Comput. Appl. Probab.*, 22(3):905–925, 2021.

[HL23a] L. Hervé and J. Ledoux. Rate of convergence of Nummelin-type representation of the invariant distribution of a Markov chain via the residual kernel. *Electron. Commun. Probab.*, 28:1–13, 2023.

[HL23b] L. Hervé and J. Ledoux. Robustness of iterated function systems of Lipschitz maps. *J. Appl. Probab.*, 60(3):921–944, 2023.

[HL24] L. Hervé and J. Ledoux. Explicit bounds for spectral theory of geometrically ergodic Markov kernels and applications. *Bernoulli*, 30(1):581–609, 2024.

[HM11] M. Hairer and J. C. Mattingly. Yet another look at Harris' ergodic theorem for Markov chains. In Robert Dalang, Marco Dozzi, and Francesco Russo, editors, *Seminar on Stochastic Analysis, Random Fields and Applications VI*, pages 109–117, Basel, 2011. Springer Basel.

[JMMD] J. E. Johndrow, J. C. Mattingly, S. Mukherjee, and D. Dunson. Optimal approximating markov chains for bayesian inference. arXiv:1508.03387v3.

[Jon04] G. L. Jones. On the Markov chain central limit theorem. *Probab. Surveys*, 1:299–320, 2004.

[JR02] S. F. Jarner and G. O. Roberts. Polynomial convergence rates of Markov chains. *Ann. Appl. Probab.*, 12(1):224–247, 2002.

[JT02] S. F. Jarner and R. L. Tweedie. Convergence rates and moments of Markov chains associated with the mean of Dirichlet processes. *Stochastic Process. Appl.*, 101(2):257–271, 2002.

[JT03] S. F. Jarner and R. L. Tweedie. Necessary conditions for geometric and polynomial ergodicity of random-walk-type. *Bernoulli*, 9(4):559–578, 2003.

[Kar86] N. V. Kartashov. Inequalities in theorems of ergodicity and stability for Markov chains with common phase space, parts I and II. *Theor. Probab. Appl.*, 30:pp. 247–259 and pp. 505–515, 1986.

[Kel82] G. Keller. Stochastic stability in some chaotic dynamical systems. *Monatsh. Math.*, 94(4):313–333, 1982.

[KL99] G. Keller and C. Liverani. Stability of the spectrum for transfer operators. *Annali della Scuola Normale Superiore di Pisa - Classe di Scienze (4)*, XXVIII:141–152, 1999.

[LH07] Y. Y. Liu and Z. T. Hou. Explicit convergence rates of the embedded $M/G/1$ queue. *Acta Math. Sin. (Engl. Ser.)*, 23(7):1289–1296, 2007.

[LH12] X. Li and Z. Hou. Subgeometric rates of convergence of the $GI/G/1$ queueing system. *Acta Math. Sci. Ser. B (Engl. Ed.)*, 32(5):1983–1996, 2012.

[LL18] Y. Liu and W. Li. Error bounds for augmented truncation approximations of Markov chains via the perturbation method. *Adv. in Appl. Probab.*, 50(2):645–669, 2018.

[MA10] Z. Mouhoubi and D. Aïssani. New perturbation bounds for denumerable Markov chains. *Linear Algebra Appl.*, 432(7):1627–1649, 2010.

[MALR16] F. J. Medina-Aguayo, A. Lee, and G. O. Roberts. Stability of noisy Metropolis-Hastings. *Stat. Comput.*, 26(6):1187–1211, 2016.

[MARS20] F. J. Medina-Aguayo, D. Rudolf, and N. Schweizer. Perturbation bounds for Monte Carlo within Metropolis via restricted approximations. *Stochastic Process. Appl.*, 130(4):2200–2227, 2020.

[Mas19] H. Masuyama. Specific bounds for a probabilistically interpretable solution of the Poisson equation for general state-space Markov chains with queueing applications. arXiv : 1909.07567, 2019.

[Mey22] S. P. Meyn. *Control Systems and Reinforcement Learning*. Cambridge University Press, 2022.

[Mit05] A. Y. Mitrophanov. Sensitivity and convergence of uniformly ergodic Markov chains. *J. Appl. Probab.*, 42(4):1003–1014, 2005.

[Mou21] Z. Mouhoubi. Perturbation and stability bounds for ergodic general state Markov chains with respect to various norms. *Matematiche (Catania)*, 76(1):243–276, 2021.

[MT09] S. P. Meyn and R. L. Tweedie. *Markov Chains and Stochastic Stability*. Cambridge University Press, 2th edition, 2009.

[NR21] J. Negrea and J. S. Rosenthal. Error bounds for approximations of geometrically ergodic reversible Markov chains. *Adv. in Appl. Probab.*, 53(4):981–1022, 2021.

[Num84] E. Nummelin. *General Irreducible Markov Chains and Nonnegative Operators*. Cambridge University Press, Cambridge, 1984.

[RR04] G. O. Roberts and J. S. Rosenthal. General state space Markov chains and MCMC algorithms. *Probab. Surveys*, 1:20–71, 2004.

[RS18] D. Rudolf and N. Schweizer. Perturbation theory for Markov chains via Wasserstein distance. *Bernoulli*, 24(4A):2610–2639, 2018.

[Sch68] P. J. Schweitzer. Perturbation theory and finite Markov chains. *J. Appl. Probab.*, 5:401–413, 1968.

[Sen93] E. Seneta. Sensitivity of finite Markov chains under perturbation. *Statist. Probab. Lett.*, 17(2):163–168, 1993.

[SS00] T. Shardlow and A. M. Stuart. A perturbation theory for ergodic Markov chains and application to numerical approximations. *SIAM J. Numer. Anal.*, 37(4):1120–1137, 2000.

[TT94] P. Tuominen and R. L. Tweedie. Subgeometric rates of convergence of f-ergodic Markov chains. *Adv. in Appl. Probab.*, 26(3):775–798, 1994.