Markov kernels under minorization and modulated drift conditions

Loïc Hervé, James Ledoux

Markov kernels under minorization and modulated drift conditions

Loïc HERVÉ, and James LEDOUX *

version: Thursday 10^{th} July, 2025 – 11:40

Abstract

The Markov kernels on general measurable space are studied under a first-order minorization condition and a modulated drift condition. The following issues are addressed: Existence and uniqueness of invariant measures, recurrence/transience properties including Harris-recurrence property, convergence in total variation of iterates, Poisson's equation, perturbation schemes and rate of convergence of iterates including the so-called geometric ergodicity. Extensions under higher order minorization conditions are also discussed with a focus on solutions to Poisson's equation. All the material on Markov kernels provided here is based on a residual kernel approach. This is a simple and efficient way to deal with all mentioned issues. It turns out that this document is essentially self-contained.

AMS subject classification : 60J05, 47B34

Keywords : Small set/function; Minorization condition; Modulated drift condition; Invariant probability measure; Recurrence; Harris-recurrence; Poisson's equation; Rate of convergence; Perturbed Markov kernels

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

Contents

1 Introduction

The purpose of this work is to study Markov kernels on a general measurable space under the so-called Minorization and modulated Drift conditions, generically denoted here by M & D conditions. The following issues are addressed: Existence and uniqueness of invariant measures, recurrence/transience properties including Harris-recurrence property, convergence in total variation of iterates of the Markov kernel in the aperiodic and periodic cases, Poisson's equation, perturbation schemes, and finally rates of convergence in weighted total variation norms of iterates including the so-called geometric ergodicity. In this document, the focus is on non-negative kernels, adopting in this sense the point of view in Seneta's book [Sen06] where discrete Markov chains are studied via non-negative matrices. It can also be thought of as a tribute to Nummelin's book [Num84] from which the idea of the treatment of Markov kernels via a residual kernel approach is borrowed. However, we decide here to keep a total focus on this kernel framework from the beginning to the end. This turns out to be a simple and efficient way to deal with all mentioned issues on Markov kernels.

The M & D conditions are nowadays well known, widely illustrated and used in the literature on Markov chains via the splitting technique for extending the materials on atomic Markov chains to the non-atomic case, or via the coupling technique to derive convergence rates. Both techniques are based on a minorization condition. The reference books on this topic are [Num84, MT09] and more recently [DMPS18]. Here we use neither the splitting technique, nor the coupling construction. This also implies that no regeneration type-method is used here. Actually, with the exception of Sections 6 and 9 which contain a few fairly elementary spectral theory arguments to deal with geometric ergodicity, the only prerequisite for this work is the handling of non-negative kernels. Indeed, the choice to consider Markov kernels satisfying a minorization condition allows us to work immediately with the residual kernel, from which the issues on invariant measures, recurrence/transience including Harris-recurrence and convergence of iterates, can be treated simply. Then additional modulated drift conditions enable us to investigate series of residual kernel iterates, from which solutions to Poisson's equation and the perturbation issue as a by-product are easily deduced. Also mention that the recent book [BH22] proposes a relevant and interesting study under additional weak topological conditions, such as the weak Feller condition. For an approach based on ergodic theory, the reader can consult for example the book [HLL03] as well as the manuscript [Hai06]. These points of view are not addressed in our work.

The theory in [Num84, MT09, DMPS18] is developed under general minorization conditions involving, either the so-called definition of small-set (or small-function), or the more general definition of petite sets. Both definitions are based on some $n-$th iterate of the transition kernel. In our work the focus is on the first-order minorization condition with small-function, which corresponds to the definition [Num84, Def. 2.3] at first order ($n := 1$). This choice provides a relatively simple, straightforward, homogeneous and self-contained presentation, dealing first with the residual kernel, then with the Markov kernel. Note that using small-functions instead of small-sets requires here no additional effort. The choice of the order one for small-functions or small-sets is also motivated by the fact that most of classical examples of Markov chains verifying a minorization condition satisfy it at the first order. Therefore, we found relevant to emphasise the order one, as long as the results are complete and the first-order minorization condition does not need to be strengthened by artificial assumptions. Further M & D conditions including the case of higher order small-functions are introduced in the final section of this work, showing that series of residual kernel iterates may have an interest at least for the study of Poisson's equation.

All the results in this work apply to any discrete-time homogeneous Markov chain, provided that the M & D conditions are fulfilled. For such examples, readers can consult the reference books [Num84, MT09, DMPS18, BH22], as well as the following more specialized works: [FM00, FM03b, AF10, 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, [Mey08] for Markov models in control. Classical instances of $V-$geometrically ergodic Markov chains can be found in e.g [MT09, RR04, DMPS18]. To make concrete checking M & D conditions, two specific examples are discussed in Section 10.

Although our method differs substantially from the splitting or coupling based methods, the conditions sometimes added to the M & D assumptions are related to the classic ones (e.g. irreducibility, period). Here these additional assumptions can be directly introduced under their simplified form, i.e. expressed with the small-function. Other conditions, such as reversibility, only concern the form of the Markov kernel and correspond to standard assumptions. Finally, as previously quoted, the central point is that a non-negative kernel approach is used for deriving all the proposed material. All the needed prerequisites are recalled in Subsection 2.1. The few probabilistic material you need (see Subsection 2.2) is applying well-known formulas on the marginal laws of the Markov chain and on the iterates of its transition kernel to deal with Harris-recurrence in Subsection 4.2. Of course, most of statements expressed in terms of Markov kernels in this work can be translated into a purely probabilistic form for discrete-time homogeneous Markov chains with general state space. To facilitate a comparative reading with the statements in reference probabilistic works as [Num84, MT09, DMPS18], the probabilistic interpretation of the main quantities used in this paper is reported in Appendix A. Further discussions are included in bibliographical comments at the end of each section.

Without using of modulated drift conditions, the first chapters of the reference books [MT09, DMPS18] present numerous sufficient conditions for P to be recurrent or transient, positive recurrent or Harris recurrent. Here the characterizations of these properties presented in Sections 3-4 only focus on two objects linked to the minorization condition: the positive measure μ_R and the function h_R^∞ introduced in Section 3. Then the modulated drift condition introduced in Section 5 allows us to directly and simply apply the results of Sections 3-4 to obtain that P admits an invariant probability measure and is Harris recurrent. Since modulated drift conditions are involved from Section 5 onwards, Sections 6-9 as well as the examples in Section 10 only concern Harris recurrent kernels with invariant probability measure.

The approach in Sections 3-5 is inspired by a part of [Num84] devoted to the so-called potential theory, which was fully developed in the seventies, e.g. see [Nev72, Twe74a, Twe74b, Rev75]. Chapter 2 of Revuz's book [Rev84] provides a full development of this theory, with additional detailed historical commentary and numerous references to early works in discrete state space case. As usually pointed out, the degree of generality and abstraction of the potential theory is high in the works cited above, which makes their access difficult. We have tried to show that the potential theory, i.e. the use of power series associated with a non-negative kernel, is in fact simple and effective when applied to the residual kernel of a transition kernel satisfying a minorization condition. Thus, the potential theory here only concerns the residual kernel, with a direct application to classical probabilistic statements, e.g. see Formula (28) used to prove the recurrence/transience dichotomy, and the definitions (40) used to study the convergence in total variation of the iterates. Similarly, in Section 6 based on spectral theory, and even in Section 9 where geometric ergodicity is introduced on a general Banach space (including the $\mathbb{L}^2-$case), only power series with the residual kernel are used, see (74).

Starting from Section 5, in which the modulated drift condition is introduced, a large part of the results are derived from the papers [HL23a, HL24, HL25a, HL25b]. However, they are revisited and sometimes upgraded/completed to take into account the new material of Subsection 3.3, Section 4, Subsection 8.4, Section 9 (excepted Subsection 9.2), Section 11. Each section ends with a detailed bibliographic discussion, excepted Section 10 on examples where some references are given directly in the development of the models and computations. Most of statements presented in Subsection 3.4 and 5.4 serve to prepare the bibliographic discussion of Sections 3 and 5. The appendices contain the proofs of expected extensions (e.g., convergence in periodic case) and technical complements concerning the truncation procedure illustrating the perturbation results.

This document is expected to offer an interesting alternative to the numerous works devoted to the asymptotic study of Markov kernels, in particular thanks to the direct and self-contained approach provided by the residual kernel. Indeed, ignoring bibliographic comments and the specific Section 11, all the topics addressed from Section 3 to Section 9 are covered in less than 80 pages, namely: Existence and uniqueness of invariant measures; Recurrence/Transience; Harris-recurrence; Convergence in total variation of the iterates; Poisson's equation; $V-$geometric ergodicity; Perturbation schemes; Polynomial ergodicity; Geometric ergodicity on a Banach space including the $\mathbb{L}^2-$case.

2 Main notations and prerequisites

The main notations and definitions used throughout this document are gathered in this section. Most of them are concerned with non-negative kernel calculus. They are standard and the material of this section can be omitted in a first reading.

Let $(\mathbb{X}, \mathcal{X})$ be a measurable space and $\mathcal{X}^* := \mathcal{X} \setminus \{\emptyset\}$ be the subset of non-trivial elements of $\mathcal{X}$. For any $A \in \mathcal{X}^*$, we denote by 1_A the indicator function of A defined by $1_A(x) := 1$ if $x \in A$, and $1_A(x) := 0$ if $x \in A^c$, where $A^c := \mathbb{X} \setminus A$.

2.1 Measures and kernels

- We denote by $\mathcal{B}$ the sets of bounded measurable real-valued functions on $(\mathbb{X}, \mathcal{X})$. The subset of non-zero and non-negative functions in $\mathcal{B}$ is denoted by $\mathcal{B}_+^*$.

- **Non-negative measures on** $(\mathbb{X}, \mathcal{X})$**.** We denote by $\mathcal{M}_+$ (resp. $\mathcal{M}_{+,b}^*$) the set of non-negative (resp. finite positive) measures on $(\mathbb{X}, \mathcal{X})$. For any $\mu \in \mathcal{M}_+$ and any μ-integrable function $g : \mathbb{X} \to \mathbb{R}$, $\mu(g)$ denotes the integral $\int_{\mathbb{X}} g(x)\mu(dx)$. Let μ be a positive measure on $(\mathbb{X}, \mathcal{X})$. Then a set $A \in \mathcal{X}$ is said to be $\mu-$full if $\mu(1_{A^c}) = 0$.

  For $\mu \in \mathcal{M}_+$ and any non-negative measurable function f, we denote by $f \cdot \mu$ the non-negative measure on $(\mathbb{X}, \mathcal{X})$ defined by: $\forall A \in \mathcal{X},\ (f \cdot \mu)(1_A) := \int_{\mathbb{X}} 1_A(x) f(x)\mu(dx)$.

- **Non-negative kernel on** $(\mathbb{X}, \mathcal{X})$**.** A non-negative kernel K on $(\mathbb{X}, \mathcal{X})$ is a map $K : \mathbb{X} \times \mathcal{X} \to [0, +\infty]$ satisfying the two following properties:

  (i) For every $A \in \mathcal{X}$, the function $x \mapsto K(x, A)$ from $\mathbb{X}$ into $[0, +\infty]$ is a measurable function on $(\mathbb{X}, \mathcal{X})$,

  (ii) For every $x \in \mathbb{X}$, the set function $A \mapsto K(x, A)$ from $\mathcal{X}$ into $[0, +\infty]$ is a non-negative measure on $(\mathbb{X}, \mathcal{X})$, denoted by $K(x, dy)$ or $K(x, \cdot)$.

  The set of non-negative kernels on $(\mathbb{X}, \mathcal{X})$ is denoted by $\mathcal{K}_+$. An element $K \in \mathcal{K}_+$ is said to be bounded if the function $x \mapsto K(x, \mathbb{X})$ is bounded on $\mathbb{X}$.

- **Product of two non-negative kernels.** If K_1 and K_2 are in $\mathcal{K}_+$, then K_2K_1 is the element of $\mathcal{K}_+$ defined by

$$\forall x \in \mathbb{X},\ \forall A \in \mathcal{X}, \quad (K_2K_1)(x, A) := \int_{\mathbb{X}} K_1(y, A)\, K_2(x, dy). \tag{1}$$

  The above term $(K_2K_1)(x, A)$ is well-defined in $[0, +\infty]$: indeed $y \mapsto K_1(y, A)$ is a measurable function from $\mathbb{X}$ into $[0, +\infty]$, and its integral is then computed w.r.t. the non-negative measure $K_2(x, dy)$. If K_1 and K_2 are both bounded, then so is K_2K_1.

- **Product of a non-negative measure by a non-negative measurable function.** For any $\mu \in \mathcal{M}_+$ and any measurable function $f : \mathbb{X} \to [0, +\infty]$, we define the following non-negative kernel, denoted by $f \otimes \mu$,

$$\forall x \in \mathbb{X},\ \forall A \in \mathcal{X}, \quad (f \otimes \mu)(x, A) := f(x)\, \mu(1_A). \tag{2}$$

- **Product of a non-negative kernel by a non-negative measure.** Any $\mu \in \mathcal{M}_+$ may be obviously considered as a non-negative kernel (i.e. $\forall x \in \mathbb{X},\ \mu(x, A) := \mu(1_A)$). If $\mu \in \mathcal{M}_+$ and $K \in \mathcal{K}_+$, then the product μK is given as a special case of Definition (1), that is

$$\forall x \in \mathbb{X},\ \forall A \in \mathcal{X}, \quad (\mu K)(x, A) := \int_{\mathbb{X}} K(y, A)\, \mu(dy). \tag{3}$$

  Note that $\mu K \in \mathcal{M}_+$ since it does not depend on $x \in \mathbb{X}$. The measure μ is said to be $K-$invariant if $\mu K = \mu$.

- **Iterates of a non-negative kernel.** Let $K \in \mathcal{K}_+$. For every $n \geq 1$ the $n-$th iterate kernel of K, denoted by K^n, is the element of $\mathcal{K}_+$ defined by induction using the above formula (1). By convention K^0 is defined by: $\forall x \in \mathbb{X},\ \forall A \in \mathcal{X},\ K^0(x, A) = 1_A(x)$ (i.e. $K^0(x, \cdot)$ is the Dirac measure at x).

- **Functional action of a non-negative kernel.** Let $K \in \mathcal{K}_+$. We also denote by K its functional action defined by

$$\forall x \in \mathbb{X}, \quad (Kg)(x) := \int_{\mathbb{X}} g(y)\, K(x, dy), \tag{4}$$

  where $g : \mathbb{X} \to \mathbb{R}$ is any measurable function assumed to be $K(x, \cdot)-$integrable for every $x \in \mathbb{X}$. For such a function g, we have

$$|Kg| \leq K|g|, \quad \text{i.e. } \forall x \in \mathbb{X},\ |(Kg)(x)| \leq (K|g|)(x), \tag{5}$$

where $|g|$ denotes the absolute value of g (or its modulus if g is $\mathbb{C}-$valued). Obviously K is a linear action.

If $K_1, K_2 \in \mathcal{K}_+$ and if $g : \mathbb{X} \to \mathbb{R}$ is a measurable function such that $g_1 := K_1 g$ is well-defined as well as $K_2 g_1$, then

$$(K_2K_1)(g) = (K_2 \circ K_1)(g)$$

where the first term $(K_2K_1)(g)$ denotes the functional action on g of the product kernel K_2K_1 given in (1), while $K_2 \circ K_1$ denotes the usual composition of maps. In particular, for every $n \geq 1$, the functional action of the $n-$th iterate kernel of K^n of K is the $n-$th iterate for composition of the functional action of K. Finally note that the functional action of the kernel K^0 is the identity map I (i.e. $(K^0g)(x) = g(x)$ for any $x \in \mathbb{X}$), which corresponds to the standard convention for linear operators.

Most questions involving a non-negative kernel can be addressed through its functional action, and this is the choice that will generally be made in this document. In particular Inequality *(5)* will be used repeatedly in this work.

- **Functional action of a non-negative measure.** If $\mu \in \mathcal{M}_+$ (thus $\mu \in \mathcal{K}_+$), then its functional action (see (4)) is given by

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

that is $\mu g := \mu(g) 1_{\mathbb{X}}$, provided that g is $\mu-$integrable.

- **Order relation for non-negative kernels.** If K_1 and K_2 are in $\mathcal{K}_+$, the inequality $K_1 \leq K_2$ means that

$$\forall g : \mathbb{X} \to [0, +\infty) \text{ measurable}, \quad 0 \leq K_1 g \leq K_2 g$$

provided that K_1g and K_2g are well-defined (if not, this inequality still holds but in $[0, +\infty]$). In particular, this implies that

$$\forall x \in \mathbb{X}, \quad K_1(x, dy) \leq K_2(x, dy), \quad \text{i.e. } \forall x \in \mathbb{X},\ \forall A \in \mathcal{X},\ K_1(x, 1_A) \leq K_2(x, 1_A).$$

In connection with this order relation, we shall often write $K \geq 0$ for recalling that $K \in \mathcal{K}_+$. When K_1, K_2 are bounded non-negative kernels, the inequality $K_1 \leq K_2$ holds true if, and only if, $K := K_2 - K_1$ is a non-negative kernel, where K is defined by $K(x, A) := K_2(x, A) - K_1(x, A)$ for any $x \in \mathbb{X}$ and $A \in \mathcal{X}$.

Recall that

$$K_1, K_2 \in \mathcal{K}_+ \implies K_1K_2 \in \mathcal{K}_+ \text{ and } K_2K_1 \in \mathcal{K}_+$$

from the definition of the products of two elements of $\mathcal{K}_+$ (see (1)). From this, the following expected rules for sum and product can be easily deduced for any K, K_1, K_2, K_1', K_2' in $\mathcal{K}_+$ (i.e. each element in (6a)-(6c) is a non-negative kernel):

$$K_1 \leq K_2,\ K_1' \leq K_2' \implies K_1 + K_1' \leq K_2 + K_2' \tag{6a}$$

$$K_1 \leq K_2,\ K \in \mathcal{K}_+ \implies KK_1 \leq KK_2 \text{ and } K_1K \leq K_2K \tag{6b}$$

$$K_1 \leq K_2 \implies \forall n \geq 0,\ K_1^n \leq K_2^n. \tag{6c}$$

Properties *(6a)–(6c)* will be used repeatedly hereafter, mainly through the functional action of the involved non-negative kernels.

- **Series of kernels.** For any $(K_i)_{i\in I} \in \mathcal{K}_+^I$ where I is any countable set I, the element $K := \sum_{i\in I} K_i$ is defined in $\mathcal{K}_+$ by

$$\forall x \in \mathbb{X},\ \forall A \in \mathcal{X}, \quad K(x, A) := \sum_{i\in I} K_i(x, A).$$

The following formula holds for all sequences $(K_n)_{n\geq 0} \in \mathcal{K}_+^{\mathbb{N}}$ and $(K'_n)_{n\geq 0} \in \mathcal{K}_+^{\mathbb{N}}$:

$$\sum_{k,n=0}^{+\infty} K_n K'_k = K\, K' \quad \text{with} \quad K := \sum_{n=0}^{+\infty} K_n \quad \text{and} \quad K' := \sum_{k=0}^{+\infty} K'_k. \tag{7}$$

Since this formula is repeatedly used in this work, let us give a proof. Let $x \in \mathbb{X}$ and $A \in \mathcal{X}$. Then (7) is obtained from the following equalities in $[0, +\infty]$:

$$\begin{aligned}
\sum_{k,n=0}^{+\infty} (K_n K'_k)(x, A) &= \sum_{k,n=0}^{+\infty} \int_{\mathbb{X}} K'_k(y, A)\, K_n(x, dy) \\
&= \sum_{n=0}^{+\infty} \left(\sum_{k=0}^{+\infty} \int_{\mathbb{X}} K'_k(y, A)\, K_n(x, dy) \right) \\
&= \sum_{n=0}^{+\infty} \int_{\mathbb{X}} \left(\sum_{k=0}^{+\infty} K'_k(y, A) \right) K_n(x, dy) \\
&= \sum_{n=0}^{+\infty} \int_{\mathbb{X}} K'(y, A) K_n(x, dy) = \int_{\mathbb{X}} K'(y, A) K(x, dy).
\end{aligned}$$

Indeed the first equality is just the definition of $K_n K'_k$, the second one is due to Fubini's theorem for double series of non-negative real numbers, the third one follows from the monotone convergence theorem w.r.t. each non-negative measure $K_n(x, dy)$, and finally the fourth and fifth ones are due to the definition of $K'(y, A)$ and $K(x, dy)$ respectively.

- **Markov and submarkov kernels.** A non-negative kernel K is said to be Markov (respectively submarkov) if $K(x, \mathbb{X}) = 1$ (respectively $K(x, \mathbb{X}) \leq 1$) for any $x \in \mathbb{X}$. In both cases, K is obviously a bounded kernel.

  If K is a Markov kernel, then an element $A \in \mathcal{X}$ is said to be $K-$absorbing if $K(x, A) = 1$ for any $x \in A$. An element $A \in \mathcal{X}$ is said to be an atom for K if the following condition holds: $\forall (x_1, x_2) \in A \times A,\ K(x_1, dy) = K(x_2, dy)$ (such a set is sometimes called a proper atom too, e.g. see [Num84, Def. 4.3]).

  If K is a submarkov kernel, then $K(\mathcal{B}) \subset \mathcal{B}$. A function $g \in \mathcal{B}$ is said to be $K-$harmonic if $Kg = g$ on $\mathbb{X}$. When K is Markov, then the function $1_{\mathbb{X}}$ is always $K-$harmonic.

- **Restriction of functions, measures ans kernels to a subset.** For any $E \in \mathcal{X}$ we denote by $\mathcal{X}_E$ the $\sigma-$algebra induced by $\mathcal{X}$ on the set E, i.e. $\mathcal{X}_E := \{A \cap E, A \in \mathcal{X}\}$. For any $g \in \mathcal{B}$, the restriction g_E to E of g is the bounded $\mathcal{X}_E-$measurable function defined on E by: $\forall x \in E,\ g_E(x) = g(x)$. If $\eta \in \mathcal{M}_+$, then the restriction η_E to E of η is the non-negative measure on $(E, \mathcal{X}_E)$ defined by: $\forall A' \in \mathcal{X}_E,\ \eta_E(1_{A'}) = \eta(1_{A\cap E})$ where A is any element in $\mathcal{X}$ such that $A' = A \cap E$. If $K \in \mathcal{K}_+$, then the restriction K_E of K to E is the non-negative kernel on $(E, \mathcal{X}_E)$ defined by: $\forall x \in E,\ \forall A' \in \mathcal{X}_E,\ K_E(x, A') =$

$K(x, A\cap E)$ where A is any element in $\mathcal{X}$ such that $A' = A\cap E$. When the notation of the function/measure/kernel on $\mathbb{X}$ involves an index, the restriction to E is denoted by $\cdot_{|E}$ to avoid confusion (for instance, if $\eta_i \in \mathcal{M}_+$, the restriction of η_i to E is denoted by $\eta_{i|E}$). Finally observe that, if K is Markov on $(\mathbb{X},\mathcal{X})$ and E is $K-$absorbing, then K_E is a Markov kernel on $(E,\mathcal{X}_E)$.

- **$V-$weighted space and $V-$weighted total variation norm.** Let $V:\mathbb{X}\to(0,+\infty)$ be any measurable function. For every measurable function $g:\mathbb{X}\to\mathbb{R}$, we set

$$\|g\|_V := \sup_{x\in\mathbb{X}} \frac{|g(x)|}{V(x)} \in [0,+\infty],$$

and we define the $V-$weighted space

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

Note that $\mathcal{B}_{1_{\mathbb{X}}} = \mathcal{B}$. The following obvious fact will be repeatedly used hereafter:

$$\forall g\in\mathcal{B}_V,\quad |g| \le \|g\|_V\, V \quad (\text{i.e. } \forall x\in\mathbb{X},\ |g(x)| \le \|g\|_V V(x)).$$

If $(\mu_1,\mu_2)\in(\mathcal{M}^*_{+,b})^2$ is such that $\mu_i(V)<\infty, i=1,2$, then the V-weighted total variation norm $\|\mu_1-\mu_2\|'_V$ is defined by

$$\|\mu_1-\mu_2\|'_V := \sup_{\|g\|_V\le 1} \big|\mu_1(g)-\mu_2(g)\big|. \tag{8}$$

If $V=1_{\mathbb{X}}$, then $\|\cdot\|'_{1_{\mathbb{X}}} = \|\cdot\|_{TV}$ is the standard total variation norm.

- **The Lebesgue space $\mathcal{L}^p(\eta)$ and $\mathbb{L}^p(\eta)$.** Let η be a positive measure on $(\mathbb{X},\mathcal{X})$. For $p\in[1,+\infty)$ we denote by $\mathcal{L}^p(\eta)$ the space of all the measurable complex-valued functions on $\mathbb{X}$ such that $\eta(|f|^p)<\infty$. Moreover $(\mathbb{L}^p(\eta),\|\cdot\|_p)$ denotes the standard Banach space composed of the classes modulo η of the functions in $\mathcal{L}^p(\eta)$ with norm defined by

$$\|f\|_p \equiv \|f\|_{p,\eta} := (\eta(|f|^p))^{1/p}.$$

As usual the space $(\mathbb{L}^\infty(\eta),\|\cdot\|_\infty)$ is the Banach space composed of the classes modulo η of complex-valued measurable functions f on $\mathbb{X}$ such that $\|f\|_\infty<\infty$ where

$$\|f\|_\infty \equiv \|f\|_{\infty,\eta} := \inf\big\{c\in[0,+\infty) : |f|\le c \ \ \eta\text{-a.e. on } \mathbb{X}\big\}. \tag{9}$$

2.2 Markov chain

A Markov chain $(X_n)_{n\ge0}$ on the state space $\mathbb{X}$ with transition/Markov kernel P is a family of random variables (r.v.) on a probability space $(\Omega,\mathcal{F},\mathbb{P})$ such that

$$\forall f\in\mathcal{B},\quad \mathbb{E}[f(X_{n+1}) \mid \sigma(X_0,\ldots,X_n)] = (Pf)(X_n)$$

where $\sigma(X_0,\ldots,X_n)$ is the sub-$\sigma-$algebra of $\mathcal{F}$ generated by the r.v's $X_0,\ldots,X_n$. In particular, for any $A\in\mathcal{X}$,

$$\mathbb{E}[1_A(X_{n+1}) \mid \sigma(X_0,\ldots,X_n)] = (P1_A)(X_n) = \int_A P(x,dy) = P(x,A).$$

Assertions a)-b) below are relevant to link iterated kernels and the Markov chain. The classical statements c)-d) are prerequisites on occupation and hitting times of a set A, which are only used in Subsection 4.2 to study the Harris-recurrence property.

a) We have for any $k \geq 0$, $\mathbb{E}[f(X_{n+k}) \mid \sigma(X_0, \ldots, X_n)] = (P^k f)(X_n)$.

b) The probability $\mathbb{P}$ when $\mathbb{P}\{X_0 = x\} = 1$, is denoted by $\mathbb{P}_x$, and $\mathbb{E}_x$ is the expectation under $\mathbb{P}_x$.

c) Let $A \in \mathcal{X}$. Then the function defined by

$$\forall x \in \mathbb{X}, \quad g_A^\infty(x) := \mathbb{P}_x\bigg\{ \sum_{n=1}^{+\infty} 1_{\{X_n \in A\}} = +\infty \bigg\} \tag{10}$$

is bounded on $\mathbb{X}$ and $P-$harmonic, see Appendix A.

d) Let $A \in \mathcal{X}$ and let g_A be the function on $\mathbb{X}$ defined by

$$\forall x \in \mathbb{X}, \quad g_A(x) = \mathbb{P}_x\{T_A < \infty\} \tag{11}$$

where $T_A := \inf\{n \geq 0 : X_n \in A\}$ is the hitting time of the set A. Then g_A is superharmonic, i.e. $Pg_A \leq g_A$, and we have (see Appendix A):

$$g_A^\infty = \lim_{n \to +\infty} \searrow P^n g_A. \tag{12}$$

3 Minorization condition, invariant measure and recurrence

In this section a standard first-order minorization condition on the Markov kernel P is introduced: $P \geq \psi \otimes \nu$ where $\nu \in \mathcal{M}_{+,b}^*$ and $\psi \in \mathcal{B}_+^*$. This allows us to decompose P as the sum of two submarkovian kernels $R := P - \psi \otimes \nu$, called the residual kernel, and $\psi \otimes \nu$. Two quantities of interest are defined from the residual kernel and its iterates: first the positive measure $\mu_R := \sum_{k=0}^{+\infty} \nu R^k$, second the $R-$harmonic function $h_R^\infty := \lim_n R^n 1_{\mathbb{X}}$. Then the existence of a $P-$invariant positive measure and the classical recurrence/transience dichotomy are studied according that $\mu_R(\psi) = 1$ or not (equivalently $\nu(h_R^\infty) = 0$ or not). The elements μ_R and h_R^∞ defined quite simply from the minorization condition, may seem abstract at first glance. They turn out to be extremely effective tools for proving, in this section and the next one, classical properties on the transition kernel P. Of course, even though the statements under the ad hoc assumptions on μ_R and h_R^∞ have their own interest, this approach would be unattractive without the possibility of deriving results under more standard assumptions. This is for example done in Theorem 3.14 and Subsection 4.5, and definitively accomplished from Section 5 with the introduction of the modulated drift condition.

3.1 The minorization condition $(\boldsymbol{M}_{\nu,\psi})$ and the residual kernel

Recall that $\mathcal{B}_+^*$ is the set of non-negative and non-zero measurable bounded functions on $\mathbb{X}$ and that $\mathcal{M}_{+,b}^*$ is the set of finite positive measures on $(\mathbb{X}, \mathcal{X})$. Let P be a Markov kernel on $(\mathbb{X}, \mathcal{X})$. Let us introduce the *minorization condition* which is in force throughout this document:

$$\exists (\nu, \psi) \in \mathcal{M}_{+,b}^* \times \mathcal{B}_+^* : \quad P \geq \psi \otimes \nu \ \ (\text{i.e. } \forall x \in \mathbb{X},\ P(x, dy) \geq \psi(x)\, \nu(dy)). \tag{$\boldsymbol{M}_{\nu,\psi}$}$$

The function ψ is called a first-order *small-function* in the literature on the topic of Markov chains. That the non-negative function ψ in $(\boldsymbol{M}_{\nu,\psi})$ is bounded is required since $\psi(x)\, \nu(1_{\mathbb{X}}) \leq$

$P(x,\mathbb{X}) = 1$ for any $x \in \mathbb{X}$ and $\nu(1_{\mathbb{X}}) > 0$. Moreover for any $(\psi,\phi) \in \mathcal{B}_+^* \times \mathcal{B}_+^*$ such that $\psi \geq \phi$, if $(\boldsymbol{M}_{\nu,\psi})$ is satisfied then so is $(\boldsymbol{M}_{\nu,\phi})$.

Under $(\boldsymbol{M}_{\nu,\psi})$, let us introduce the following submarkov kernel, called the *residual kernel*, which is central in the analysis here of the Markov kernel P:

$$R \equiv R_{\nu,\psi} := P - \psi \otimes \nu \quad (\text{i.e. } \forall x \in \mathbb{X},\ R(x,dy) := P(x,dy) - \psi(x)\nu(dy)). \tag{13}$$

The most classical instance of minorization condition is when $\psi := 1_S$ for some $S \in \mathcal{X}^*$, that is

$$\exists(\nu,S) \in \mathcal{M}_{+,b}^* \times \mathcal{X}^* : \quad P \geq 1_S \otimes \nu \ \ (\text{i.e. } \forall x \in \mathbb{X},\ P(x,dy) \geq 1_S(x)\,\nu(dy)), \tag{$\boldsymbol{M}_{\nu,1_S}$}$$

in which case the residual kernel is:

$$R \equiv R_{\nu,1_S} := P - 1_S \otimes \nu.$$

Such a set S is called a first-order *small-set*.

The following statement provides a general framework for Condition $(\boldsymbol{M}_{\nu,\psi})$ to hold. Moreover this proposition shows that, even if the minorizing measure ν is defined from $(\boldsymbol{M}_{\nu,1_S})$ with some set S, this condition $(\boldsymbol{M}_{\nu,1_S})$ is not the only one possible.

Proposition 3.1 *Assume that*

$$\forall x \in \mathbb{X}, \quad P(x,dy) \geq q(x,y)\,\lambda(dy) \tag{14}$$

where $q(\cdot,\cdot)$ is a non-negative measurable function on $\mathbb{X}^2$ and λ is a positive measure on $\mathbb{X}$. Let $S \in \mathcal{X}^*$ be such that the measurable non-negative function q_S defined by

$$\forall y \in \mathbb{X}, \quad q_S(y) := \inf_{x\in S} q(x,y)$$

is not $\lambda-$null, that is: $\lambda(1_A) > 0$ where $A := \{y \in \mathbb{X} : q_S(y) > 0\}$. Let $\nu \in \mathcal{M}_{+,b}^*$ and $\psi_S \geq 1_S$ be defined by

$$\nu(dy) := q_S(y)\lambda(dy) \quad \text{and} \quad \forall x \in \mathbb{X},\ \psi_S(x) := 1_S(x) \inf_{y\in A} \frac{q(x,y)}{q_S(y)}. \tag{15}$$

Then P satisfies Condition $(\boldsymbol{M}_{\nu,\psi_S})$ and so $(\boldsymbol{M}_{\nu,1_S})$.

Proof. For any fixed $x \in S$, we have $\nu(1_{\mathbb{X}}) \leq \int_{\mathbb{X}} q(x,y)\lambda(dy) \leq P(x,\mathbb{X}) = 1$ from the definition of ν, q_S and from (14). Thus ν is finite and $\nu(1_A) > 0$, so that $\nu \in \mathcal{M}_{+,b}^*$. Next, from the definition of ψ_S we obtain the following property: $\forall (x,y) \in S \times A,\ q(x,y) \geq q_S(y)\,\psi_S(x)$. In fact this inequality holds for every $(x,y) \in \mathbb{X}^2$ since $q(x,y) \geq 0$. Finally it follows from (14) that, for every $x \in \mathbb{X}$, we have $P(x,dy) \geq \psi_S(x)q_S(y)\lambda(dy)$, i.e. P satisfies $(\boldsymbol{M}_{\nu,\psi_S})$. Note that $\psi_S \geq 1_S$ from the definition of the function q_S, so that $(\boldsymbol{M}_{\nu,1_S})$ is satisfied. □

The next kernel identity (17) is the first key formula of this work. Recall that the residual kernel $R := P - \psi \otimes \nu$ is a submarkov kernel, so that the $n-$th iterate kernel R^n of R defined by induction using Formula (1) is a submarkov kernel too. Also recall that by convention $R^0(x,\cdot)$ is the Dirac measure at x. Finally note that, for every $k \geq 1$, we have $\nu\, R^k \in \mathcal{M}_{+,b}$ (see (3)).

Lemma 3.2 *Let P satisfy Condition $(\boldsymbol{M}_{\nu,\psi})$. Then we have*

$$\forall n \geq 1, \quad 0 \leq R^n \leq P^n, \tag{16}$$

$$P^n = R^n + \sum_{k=1}^{n} P^{n-k}\psi \otimes \nu R^{k-1}, \tag{17}$$

and the kernel identity

$$\sum_{n=0}^{+\infty} P^n = \sum_{n=0}^{+\infty} R^n + \Big(\sum_{n=0}^{+\infty} P^n\psi\Big) \otimes \Big(\sum_{k=0}^{+\infty} \nu R^k\Big). \tag{18}$$

Proof. We have $0 \leq R \leq P$, thus $0 \leq R^n \leq P^n$ using (6c). Set $T_0 := 0$ and $T_n := P^n - R^n$ for $n \geq 1$. Note that Property (17) is equivalent to

$$\forall n \geq 1, \quad T_n = \sum_{k=1}^{n} P^{n-k}\psi \otimes \nu R^{k-1}. \tag{19}$$

Equality (19) is clear for $n = 1$ since $T_1 = P - R = \psi \otimes \nu$. Next we have for any $n \geq 2$

$$R^n = R^{n-1}R = (P^{n-1} - T_{n-1})(P - T_1) = P^n - P^{n-1}T_1 - T_{n-1}R,$$

so that $T_n = P^{n-1}T_1 + T_{n-1}R$. Then (19) holds for $n \geq 2$ by an easy induction based on the previous equality for T_n: For instance use the functional action of kernels to check that, for every $g \in \mathcal{B}$, if $T_{n-1}g = \sum_{k=1}^{n-1} \nu(R^{k-1}g)P^{n-1-k}\psi$, then $T_n g = \sum_{k=1}^{n} \nu(R^{k-1}g)P^{n-k}\psi$.

From (17) and the convention for $P^0 = R^0$ we obtain that (see (7))

$$\begin{aligned}\sum_{n=0}^{+\infty} P^n &= \sum_{n=0}^{+\infty} R^n + \sum_{n=1}^{+\infty}\sum_{k=1}^{n} P^{n-k}\psi \otimes \nu R^{k-1} = \sum_{n=0}^{+\infty} R^n + \sum_{k=1}^{+\infty}\sum_{n=k}^{+\infty} P^{n-k}\psi \otimes \nu R^{k-1} \\ &= \sum_{n=0}^{+\infty} R^n + \Big(\sum_{n=0}^{+\infty} P^n\psi\Big) \otimes \Big(\sum_{k=0}^{+\infty} \nu R^k\Big)\end{aligned}$$

Thus (18) holds and the proof of Lemma 3.2 is complete. □

Under Condition $(\boldsymbol{M}_{\nu,\psi})$, we have $0 \leq R1_{\mathbb{X}} \leq 1_{\mathbb{X}}$. Since R is a non-negative kernel, we get $0 \leq R^{n+1}1_{\mathbb{X}} \leq R^n 1_{\mathbb{X}}$ for any $n \geq 0$. Thus the sequence $(R^n 1_{\mathbb{X}})_{n\geq 0}$ is non-increasing so that it converges point-wise. Consequently we can define the following measurable function $h_R^\infty : \mathbb{X} \to [0, 1]$:

$$h_R^\infty := \lim_n \searrow R^n 1_{\mathbb{X}}. \tag{20}$$

Note that h_R^∞ is $R-$harmonic: indeed, for every $x \in \mathbb{X}$, we have $(R^{n+1}h_R^\infty)(x) = (RR^n h_R^\infty)(x)$, so that $h_R^\infty(x) = (Rh_R^\infty)(x)$ from Lebesgue's theorem applied to the finite non-negative measure $R(x, dy)$ observing that $R^n h_R^\infty \leq R^n 1_{\mathbb{X}} \leq 1_{\mathbb{X}}$.

Under Condition $(\boldsymbol{M}_{\nu,\psi})$ let μ_R denote the positive measure on $(\mathbb{X}, \mathcal{X})$ (not necessarily finite) defined by

$$\mu_R := \sum_{k=0}^{+\infty} \nu R^k. \tag{21}$$

Note that the measure μ_R is positive from $\mu_R(1_{\mathbb{X}}) \geq \nu(1_{\mathbb{X}}) > 0$. The measure μ_R as well as the function h_R^∞ are used throughout this section.

3.2 $P-$invariant measure

First prove the following simple lemma.

Lemma 3.3 *Assume that P satisfies Conditions $(\boldsymbol{M}_{\nu,\psi})$. Let g be a $P-$harmonic function. Then we have*

$$\forall n \geq 0, \quad \nu(g) \sum_{k=0}^{n} R^k \psi = g - R^{n+1} g. \tag{22}$$

In particular we have

$$\forall n \geq 0, \quad 0 \leq \nu(1_{\mathbb{X}}) \sum_{k=0}^{n} R^k \psi = 1_{\mathbb{X}} - R^{n+1} 1_{\mathbb{X}} \leq 1_{\mathbb{X}}. \tag{23}$$

Proof. Let $g \in \mathcal{B}$ be such that $Pg = g$. We have $\nu(g)\psi = (I - R)g$ from the definition (13) of R. Then Property (22) follows from

$$\forall n \geq 0, \quad \nu(g) \sum_{k=0}^{n} R^k \psi = \bigg(\sum_{k=0}^{n} R^k \bigg) (I - R) g = \sum_{k=0}^{n} R^k g - \sum_{k=1}^{n+1} R^k g = g - R^{n+1} g.$$

Since $P1_{\mathbb{X}} = 1_{\mathbb{X}}$, Property (22) with $g := 1_{\mathbb{X}}$ is nothing else than (23). □

Recall that the positive measure ν in $(\boldsymbol{M}_{\nu,\psi})$ is finite (i.e. $\nu(1_{\mathbb{X}}) < \infty$).

Proposition 3.4 *Let P satisfy Condition $(\boldsymbol{M}_{\nu,\psi})$. Then the function series $\sum_{k=0}^{+\infty} R^k \psi$ point-wise converges and is bounded on $\mathbb{X}$. More precisely we have*

$$0 \leq \nu(1_{\mathbb{X}}) \sum_{k=0}^{+\infty} R^k \psi = 1_{\mathbb{X}} - h_R^{\infty} \leq 1_{\mathbb{X}}. \tag{24}$$

Moreover we have $\mu_R(\psi) = \sum_{k=0}^{+\infty} \nu(R^k \psi) \in [0, 1]$, and the following equivalences hold

$$\mu_R(\psi) = 1 \iff \nu(h_R^{\infty}) = 0 \iff \mu_R(h_R^{\infty}) = 0. \tag{25}$$

Note that the property $\mu_R(\psi) \leq 1$ implies that there exists $A \in \mathcal{X}^*$ such that $\mu_R(1_A) < \infty$.

Proof. It follows from (23) that the series of non-negative functions $\sum_{k=0}^{+\infty} R^k \psi$ point-wise converges. When n growths to $+\infty$ in (23), we get the equality in (24) from the definition (20) of h_R^{∞}.

Next integrate w.r.t. the measure ν in (24) and apply the monotone convergence theorem to get $0 \leq \nu(1_{\mathbb{X}})\mu_R(\psi) = \nu(1_{\mathbb{X}}) - \nu(h_R^{\infty}) \leq \nu(1_{\mathbb{X}})$. Since $\nu(1_{\mathbb{X}}) > 0$, it follows that $\mu_R(\psi) \in [0, 1]$ and the first equivalence in (25) holds. Since $Rh_R^{\infty} = h_R^{\infty}$, we have from (21) that $\nu(h_R^{\infty}) = 0$ implies that $\mu_R(h_R^{\infty}) = 0$. Finally, we have $\mu_R(h_R^{\infty}) \geq \nu(h_R^{\infty}) \geq 0$ from the definition (21) of μ_R so that $\mu_R(h_R^{\infty}) = 0$ implies that $\nu(h_R^{\infty}) = 0$. The proof of the second equivalence in (25) is complete. □

Theorem 3.5 ($P-$invariant positive measure) *Assume that P satisfies Condition $(\boldsymbol{M}_{\nu,\psi})$. Then the following assertions hold.*

1. *If $\mu_R(\psi) = 1$ (or equivalently $\nu(h_R^{\infty}) = 0$), then μ_R is a $P-$invariant positive measure.*

2. *If there exists $\zeta \in \mathcal{B}_+^*$ such that $\nu(\zeta) > 0$ and $\mu_R(P\zeta) = \mu_R(\zeta) < \infty$, then we have $\mu_R(\psi) = 1$.*

In particular, if $\nu(\psi) > 0$, then

$$\mu_R \text{ is } P-\text{invariant} \Longleftrightarrow \mu_R(\psi) = 1 \Longleftrightarrow \nu(h_R^\infty) = 0 \Longleftrightarrow \mu_R(h_R^\infty) = 0.$$

Recall that the condition $\nu(\psi) > 0$ is the so-called *strong aperiodicity* property.

Proof. From the definitions (13) of R and (21) of μ_R, the following equalities hold in $[0, +\infty]$:

$$\forall A \in \mathcal{X}, \quad \mu_R(P1_A) = \mu_R(R1_A) + \nu(1_A)\mu_R(\psi) = \mu_R(1_A) + \nu(1_A)\big(\mu_R(\psi) - 1\big)$$

since we have $\mu_R(R1_A) = \mu_R(1_A) - \nu(1_A)$ in $[0, +\infty]$. Consequently, if $\mu_R(\psi) = 1$, then μ_R is a $P-$invariant positive measure and Assertion 1. is proved. Next, if $\zeta \in \mathcal{B}_+^*$ satisfies the assumptions in Assertion 2., then we deduce from $\mu_R(\zeta) = \mu_R(P\zeta) = \mu_R(\zeta) + \nu(\zeta)\big(\mu_R(\psi) - 1\big)$ that $\mu_R(\psi) = 1$. In the last assertion, that $\mu_R(\psi) = 1$ implies the $P-$invariance of μ_R is just Assertion 1. Next, if $\nu(\psi) > 0$ and μ_R is $P-$invariant, then Assertion 2. can be applied to $\zeta := \psi$ since we know that $\mu_R(\psi) < \infty$ from Proposition 3.4, so that we have $\mu_R(\psi) = 1$. The two last equivalences are (25). □

Theorem 3.6 ($P-$invariant probability measure) *If P satisfies Condition $(\boldsymbol{M}_{\nu,\psi})$, then the following assertions are equivalent.*

1. *There exists a $P-$invariant probability measure η on $(\mathbb{X}, \mathcal{X})$ such that $\eta(\psi) > 0$.*
2. $\mu_R(1_{\mathbb{X}}) = \sum_{k=0}^{+\infty} \nu(R^k 1_{\mathbb{X}}) < \infty$.
3. *There exists a probability measure σ on $(\mathbb{X}, \mathcal{X})$ such that $\liminf_{n\to+\infty} \sigma(P^n\psi) > 0$.*

Under any of these three conditions, the following probability measure on $(\mathbb{X}, \mathcal{X})$

$$\pi_R := \mu_R(1_{\mathbb{X}})^{-1}\, \mu_R \quad \textit{with} \quad \mu_R := \sum_{k=0}^{+\infty} \nu R^k \in \mathcal{M}_{*,b}^+ \tag{26}$$

is $P-$invariant with $\mu_R(\psi) = 1$ and $\pi_R(\psi) = \mu_R(1_{\mathbb{X}})^{-1} > 0$.

Proof. We prove the implications: 2. ⇒ 1. ⇒3. ⇒2. If Assertion 2. holds, then Assertion 2. of Theorem 3.5 can be applied with $\zeta := 1_{\mathbb{X}}$. Indeed, $\nu(1_{\mathbb{X}}) > 0$ and $\mu_R(P1_{\mathbb{X}}) = \mu_R(1_{\mathbb{X}}) < \infty$ since P is Markov. Hence we have $\mu_R(\psi) = 1$, so that μ_R is $P-$invariant from Assertion 1. of Theorem 3.5. Thus $\pi_R := \mu_R(1_{\mathbb{X}})^{-1}\, \mu_R$ is a $P-$invariant probability measure such that $\pi_R(\psi) = \mu_R(1_{\mathbb{X}})^{-1} > 0$. The implication 2. ⇒ 1. is proved. Next, if Assertion 1. is fulfilled, then Assertion 3. obviously holds with $\sigma := \eta$. Finally assume that Assertion 3. holds. Then apply Formula (17) to $1_{\mathbb{X}}$ and integrate w.r.t. the probability measure σ to get

$$\forall n \geq 1, \quad \sum_{k=1}^{+\infty} \nu(R^{k-1}1_{\mathbb{X}})\sigma_n(k) \leq \sigma(1_{\mathbb{X}}) = 1 \quad \text{with} \quad \sigma_n(k) := \sigma(P^{n-k}\psi)\, 1_{[1,n]}(k)$$

from $\sigma(R^n 1_{\mathbb{X}}) \geq 0$. Since for every $k \geq 1$ we have $m := \liminf_j \sigma(P^j\psi) = \liminf_n \sigma_n(k)$ and $m > 0$ by hypothesis, we deduce from Fatou's lemma w.r.t. discrete measure that $m \sum_{k=1}^{+\infty} \nu(R^{k-1}1_{\mathbb{X}}) \leq 1$. Thus $\mu_R(1_{\mathbb{X}}) \leq 1/m < \infty$, i.e. Assertion 2. holds true. □

The following standard example of uniform ergodicity illustrates Theorem 3.6. Moreover, the well-known rate of convergence below of $\|P^n(x,\cdot) - \pi_R(\cdot)\|_{TV}$ is obtained from Formula (17).

Example 3.7 (Uniform ergodicity) *Let P satisfy Condition $(\boldsymbol{M}_{\nu,1_{\mathbb{X}}})$, that is there exists $\nu \in \mathcal{M}^*_{+,b}$ such that $P \geq 1_{\mathbb{X}} \otimes \nu$. In other words the whole state space $\mathbb{X}$ is a first-order small-set for P. Then Condition 2. of Theorem 3.6 holds and we have*

$$\forall n \geq 1,\ \forall x \in \mathbb{X}, \quad \|P^n(x,\cdot) - \pi_R(\cdot)\|_{TV} \leq 2(1-\nu(1_{\mathbb{X}}))^n$$

where π_R is the $P-$invariant probability measure given by (26). Indeed the residual kernel $R \equiv R_{\nu,1_{\mathbb{X}}}$ is here $R = P - 1_{\mathbb{X}} \otimes \nu$ so that we have $R1_{\mathbb{X}} = (1-\nu(1_{\mathbb{X}}))1_{\mathbb{X}}$. Consequently we obtain that

$$\forall n \geq 1, \quad R^n 1_{\mathbb{X}} = (1-\nu(1_{\mathbb{X}}))^n 1_{\mathbb{X}}.$$

Thus $\mu_R(1_{\mathbb{X}}) = \sum_{k=0}^{+\infty} \nu(R^k 1_{\mathbb{X}}) = 1$, and it follows from Theorem 3.6 that the probability measure π_R given in (26) is $P-$invariant ($\pi_R = \mu_R$ here). Moreover Formula (17) gives

$$\forall n \geq 1, \quad P^n = R^n + 1_{\mathbb{X}} \otimes \mu_n \quad \textit{with} \quad \mu_n := \sum_{k=1}^{n} \nu R^{k-1}.$$

Consequently we have

$$\forall n \geq 1, \quad P^n - 1_{\mathbb{X}} \otimes \pi_R = R^n - 1_{\mathbb{X}} \otimes \sum_{k=n+1}^{+\infty} \nu R^{k-1},$$

from which we derive that

$$\begin{aligned}
\forall n \geq 1,\ \forall x \in \mathbb{X}, \quad \|P^n(x,\cdot) - \pi_R\|_{TV} &\leq \|R^n(x,\cdot)\|_{TV} + \left\| \sum_{k=n+1}^{+\infty} \nu R^{k-1} \right\|_{TV} \\
&= R^n(x,1_{\mathbb{X}}) + \sum_{k=n+1}^{+\infty} \nu(R^{k-1}1_{\mathbb{X}}) \\
&= 2(1-\nu(1_{\mathbb{X}}))^n.
\end{aligned}$$

3.3 Recurrence/Transience

If P satisfies Condition $(\boldsymbol{M}_{\nu,\psi})$, then P is said to be *recurrent* if the following condition holds:

$$\forall A \in \mathcal{X}: \ \mu_R(1_A) > 0 \Longrightarrow \sum_{k=0}^{+\infty} P^k 1_A = +\infty \text{ on } \mathbb{X} \text{ (i.e. } \forall x \in \mathbb{X},\ \sum_{k=0}^{+\infty} P^k(x,A) = +\infty), \quad (27)$$

where μ_R is the positive measure on $(\mathbb{X}, \mathcal{X})$ defined in (21). Note that if $A \in \mathcal{X}$ is such that $\nu(1_A) > 0$ then $\mu_R(1_A) > 0$. Moreover observe that Equality (18) reads as

$$\sum_{n=0}^{+\infty} P^n = \sum_{n=0}^{+\infty} R^n + \left(\sum_{n=0}^{+\infty} P^n \psi\right) \otimes \mu_R. \quad (28)$$

This potential-type formula is here the key point to obtain the Recurrence/Transience dichotomy. It also shows that μ_R may be thought of as the ideal measure for studying this issue under the minorization condition $(\boldsymbol{M}_{\nu,\psi})$. To get a complete picture of recurrence/transience property for P satisfying Condition $(\boldsymbol{M}_{\nu,\psi})$ in the next statement, let us introduce the following definition. The Markov kernel P is said to be *irreducible* if

$$\sum_{n=1}^{+\infty} P^n\psi > 0 \text{ on } \mathbb{X}, \text{ i.e. } \forall x \in \mathbb{X},\ \exists q \equiv q(x) \geq 1, \quad (P^q\psi)(x) > 0. \tag{29}$$

Recall that under $(\boldsymbol{M}_{\nu,\psi})$, we have $\mu_R(\psi) \in [0,1]$ from Proposition 3.4, and that μ_R is a $P-$invariant positive measure when $\mu_R(\psi) = 1$, or equivalently $\nu(h_R^\infty) = 0$ (see (25)), from Theorem 3.5. Finally, recall that $\|\cdot\|_{1_\mathbb{X}}$ denotes the supremum norm on $\mathcal{B}$ (i.e. $\|g\|_{1_\mathbb{X}} := \sup_{x\in\mathbb{X}} |g(x)|$).

Theorem 3.8 *Let P satisfy Condition $(\boldsymbol{M}_{\nu,\psi})$. Then the following assertions hold.*

1. *Case $\mu_R(\psi) = 1$ (or equivalently $\nu(h_R^\infty) = 0$). The Markov kernel P is recurrent if and only if P is irreducible (see (29)). When P is recurrent, μ_R is the unique $P-$invariant positive measure η (up to a multiplicative positive constant) such that $\eta(\psi) < \infty$, and μ_R is $\sigma-$finite.*

2. *Case $\mu_R(\psi) < 1$ (or equivalently $\nu(h_R^\infty) > 0$). The non-negative function series $\sum_{k=0}^{+\infty} P^k\psi$ is bounded from above by $\nu(h_R^\infty)^{-1}$ on $\mathbb{X}$. If P is irreducible, then P is not recurrent, more precisely P is said to be transient in the following sense: Defining for every integer $k \geq 1$ the set $A_k := \{x \in \mathbb{X} : \sum_{j=0}^{k}(R^j\psi)(x) \geq 1/k\}$ we have*

$$\mathbb{X} = \cup_{k=1}^{+\infty} A_k \quad \textit{and} \quad \forall k \geq 1,\ c_k := \|\sum_{n=0}^{+\infty} P^n 1_{A_k}\|_{1_\mathbb{X}} < \infty.$$

   Actually we have: $\forall k \geq 1,\ c_k \leq k(k+1)(\nu(1_\mathbb{X})^{-1} + \nu(h_R^\infty)^{-1})$.

When P is irreducible, we have the following characterization of recurrence.

Corollary 3.9 *Assume that P satisfies Conditions $(\boldsymbol{M}_{\nu,\psi})$ and is irreducible. Then*

$$P \textit{ is recurrent} \iff \mu_R(\psi) = 1 \iff \nu(h_R^\infty) = 0 \iff \mu_R(h_R^\infty) = 0.$$

Proof. Assume that $\mu_R(\psi) \in [0,1)$. Then P is not recurrent from the second assertion of Theorem 3.8. This proves the first direct implication. The converse one follows from the first assertion of Theorem 3.8. The two last equivalences are (25). □

The proof of Theorem 3.8 is based on the the two following lemmas.

Lemma 3.10 *Let P satisfy Condition $(\boldsymbol{M}_{\nu,\psi})$. If P is irreducible then the following statements hold:*

1. *$\sum_{n=0}^{+\infty} R^n\psi > 0$ on $\mathbb{X}$ and $\mu_R(\psi) > 0$.*

2. *If $\mu_R(\psi) = 1$ (or equivalently $\nu(h_R^\infty) = 0$) then $\sum_{n=0}^{+\infty} P^n\psi = +\infty$ on $\mathbb{X}$.*

Proof. We prove Assertion 1. by contradiction. Assume that there exists $x \in \mathbb{X}$ such that $\sum_{n=0}^{+\infty}(R^n\psi)(x) = 0$. Then we have $h_R^\infty(x) = 1$ from (24). From the definition of $h_R^\infty(x)$ and $R^n 1_{\mathbb{X}} \le 1$, it then follows that: $\forall n \ge 1,\ (R^n 1_{\mathbb{X}})(x) = 1$. Hence we deduce from Formula (17) and $(P^n 1_{\mathbb{X}})(x) = 1$ that

$$\forall n \ge 1, \quad \sum_{k=1}^{n}(P^{n-k}\psi)(x)\,\nu(R^{k-1}1_{\mathbb{X}}) = 0.$$

In particular the first term of this sum of non-negative real numbers is zero, that is we have: $\forall n \ge 1,\ (P^{n-1}\psi)(x)\,\nu(1_{\mathbb{X}}) = 0$. Since P is irreducible (see (29)), we know that there exists $q \equiv q(x) \ge 1$ such that $(P^q\psi)(x) > 0$. Then the previous equality with $n = q+1$ implies that $\nu(1_{\mathbb{X}}) = 0$: Contradiction. This proves the first part of Assertion 1. Next, since $\mu_R(\psi) = \sum_{n=0}^{+\infty}\nu(R^n\psi) = \nu(\sum_{n=0}^{+\infty}R^n\psi)$ from monotone convergence theorem, we have $\mu_R(\psi) > 0$. Assertion 1. is proved. Next, if $\mu_R(\psi) = 1$, then Equality (28) applied to ψ and Assertion 1. imply that $\sum_{n=0}^{+\infty}P^n\psi = +\infty$ on $\mathbb{X}$. □

Lemma 3.11 *Let P satisfy Condition $(\boldsymbol{M}_{\nu,\psi})$ with $\mu_R(\psi) > 0$. If P is recurrent, then $\sum_{k=0}^{+\infty}P^k\psi = +\infty$ on $\mathbb{X}$.*

Proof. Since $\mu_R(\psi) > 0$, there exists $\varepsilon > 0$ such that the set $F_\varepsilon := \{x \in \mathbb{X} : \psi(x) \ge \varepsilon\}$ satisfies $\mu_R(1_{F_\varepsilon}) > 0$ (otherwise we would have $\mu_R(\{x \in \mathbb{X} : \psi(x) > 0\}) = 0$, thus $\mu_R(\psi) = 0$). From recurrence and $1_{F_\varepsilon} \le \psi/\varepsilon$, we obtain that $\sum_{n=0}^{+\infty}P^n\psi = +\infty$ on $\mathbb{X}$. □

Now, let us provide a proof of Theorem 3.8.

Proof of Theorem 3.8. Assume that $\mu_R(\psi) = 1$. If P is irreducible, then $\sum_{k=0}^{+\infty}P^k\psi = +\infty$ on $\mathbb{X}$ from Assertion 2. of Lemma 3.10. It follows from (28) applied to 1_A that $\sum_{k=0}^{+\infty}P^k 1_A = +\infty$ for every $A \in \mathcal{X}$ such that $\mu_R(1_A) > 0$, i.e. P is recurrent. Conversely, if P is recurrent, then it follows from $\mu_R(\psi) = 1$ and Lemma 3.11 that $\sum_{n=0}^{+\infty}P^n\psi = +\infty$ on $\mathbb{X}$. Thus P satisfies (29), i.e. P is irreducible. Now assume that P is recurrent, thus irreducible. Let η be a $P-$invariant positive measure on $(\mathbb{X}, \mathcal{X})$ such that $\eta(\psi) < \infty$. Then η is $\sigma-$finite due to the following well-known argument. Let $Q := \sum_{n=0}^{+\infty}2^{-(n+1)}P^n$ be the Markov resolvent kernel associated with P. Then $Q\psi > 0$ on $\mathbb{X}$ from (29). Hence we have $\mathbb{X} = \{Q\psi > 0\} = \cup_{k\ge1}E_k$ with $E_k := \{Q\psi \ge 1/k\}$, and $\eta(1_{E_k}) \le k\,\eta(Q\psi) = k\,\eta(\psi) < \infty$ from Markov's inequality. Thus η is $\sigma-$finite. Next prove by contradiction that $\eta(\psi) > 0$. Assume that $\eta(\psi) = 0$. Then we obtain that $\eta(1_{E_k}) = 0$ for any $k \ge 1$ from the last inequality above, so that $\eta(1_{\mathbb{X}}) = 0$ since $\mathbb{X} = \cup_{k\ge1}E_k$: This is impossible since η is a positive measure on $(\mathbb{X}, \mathcal{X})$. Now recall that μ_R is $P-$invariant under the assumption $\mu_R(\psi) = 1$ due to Theorem 3.5, and prove that $\eta = \eta(\psi)\mu_R$. From (17) and the $P-$invariance of η we obtain that: $\forall n \ge 1,\ \eta \ge \eta(\psi)\sum_{k=1}^{n}\nu R^{k-1}$. Thus $\eta \ge \eta(\psi)\mu_R$ from the definition (21) of μ_R. Next, since both η and μ_R are $\sigma-$finite from the above, it follows from the Radon-Nikodym theorem that there exists a measurable function v on $\mathbb{X}$ such that $\eta(\psi)\mu_R = v \cdot \eta$ with $0 \le v \le 1_{\mathbb{X}}$ $\eta-$a.e.. Let λ be the non-negative measure on $(\mathbb{X}, \mathcal{X})$ defined by: $\lambda := (1_{\mathbb{X}} - v)\cdot\eta$. Since $\eta(Q\psi) = \eta(\psi) < \infty$ by hypothesis with Q defined above, we obtain that the function $v \times (Q\psi)$ is $\eta-$integrable too, so that

$$\lambda(Q\psi) = \int_{\mathbb{X}}(Q\psi)(x)\,\eta(dx) - \int_{\mathbb{X}}(Q\psi)(x)\,v(x)\,\eta(dx) = \eta(Q\psi) - \eta(\psi)\mu_R(Q\psi) = 0$$

from the $P-$invariance of both η and μ_R and from the assumption $\mu_R(\psi) = 1$. It follows that $\lambda = 0$ since $Q\psi > 0$ on $\mathbb{X}$. Thus we have $v = 1_{\mathbb{X}}$ $\eta-$a.e., so that $\eta(\psi)\mu_R = \eta$. Assertion 1. of Theorem 3.8 is proved.

Now assume that $\mu_R(\psi) < 1$. Thus we have $\nu(h_R^\infty) > 0$ from (25). Recall that $Rh_R^\infty = h_R^\infty$. Then, Formula (17) applied to h_R^∞ and the equality $Rh_R^\infty = h_R^\infty$ give

$$\forall n \geq 1, \quad P^n h_R^\infty = h_R^\infty + \nu(h_R^\infty) \sum_{k=0}^{n-1} P^k \psi,$$

from which we deduce that: $\forall n \geq 1$, $\sum_{k=0}^{n-1} P^k\psi \leq \nu(h_R^\infty)^{-1} 1_{\mathbb{X}}$ since $h_R^\infty \geq 0$ and $P^n h_R^\infty \leq 1_{\mathbb{X}}$ from $h_R^\infty \leq 1_{\mathbb{X}}$. Consequently we have the following inequality on $\mathbb{X}$:

$$0 \leq \sum_{k=0}^{+\infty} P^k \psi \leq \nu(h_R^\infty)^{-1} 1_{\mathbb{X}}.$$

Now assume that P is irreducible. Recall that $\mu_R(\psi) > 0$ from Lemma 3.10. Thus, as in the proof of Lemma 3.11, there exists $\varepsilon > 0$ and a set F_ε such that $\mu_R(1_{F_\varepsilon}) > 0$ and $1_{F_\varepsilon} \leq \psi/\varepsilon$. We deduce that $\sum_{n=0}^{+\infty} P^n 1_{F_\varepsilon}$ is bounded on $\mathbb{X}$. Consequently P is not recurrent. Next let us prove that P is transient as defined in Theorem 3.8. We have $\mathbb{X} = \cup_{k=1}^{+\infty} A_k$. Indeed, otherwise there would exist $x \in \mathbb{X}$ such that: $\forall k \geq 1$, $\sum_{j=0}^{k} (R^j\psi)(x) < 1/k$, so that $\sum_{j=0}^{+\infty} (R^j\psi)(x) = 0$: This contradicts Lemma 3.10. Finally let $k \geq 1$. Observing that $1_{A_k} \leq k \sum_{j=0}^{k} R^j\psi$, we obtain that (see (7))

$$\begin{aligned} \sum_{n=0}^{+\infty} R^n 1_{A_k} &\leq k \sum_{n=0}^{+\infty} R^n \Big(\sum_{j=0}^{k} R^j \psi \Big) = k \sum_{j=0}^{k} R^j \Big(\sum_{n=0}^{+\infty} R^n \psi \Big) \\ &\leq k\, \nu(1_{\mathbb{X}})^{-1} \sum_{j=0}^{k} R^j 1_{\mathbb{X}} \leq k(k+1)\nu(1_{\mathbb{X}})^{-1} 1_{\mathbb{X}} \text{ (using (24) and } R1_{\mathbb{X}} \leq 1_{\mathbb{X}}). \end{aligned}$$

Moreover, integrating the previous inequality w.r.t the positive measure ν, it follows from the monotone convergence theorem that $\mu_R(1_{A_k}) \leq k(k+1)$. Then the last inequalities combined with Formula (28) applied to 1_{A_k} provide

$$\sum_{n=0}^{+\infty} P^n 1_{A_k} \leq k(k+1)\big[\nu(1_{\mathbb{X}})^{-1} + \nu(h_R^\infty)^{-1}\big] 1_{\mathbb{X}}.$$

The proof of Theorem 3.8 is complete. □

Recall that P is irreducible (see (29)) if, and only if, the function series $\sum_{k=0}^{+\infty} P^k\psi$ takes its values in $(0, +\infty]$. Thus, when P is irreducible, the recurrence/transience dichotomy can also be addressed focusing solely on this function series.

Corollary 3.12 *Assume that P satisfies Condition $(\boldsymbol{M}_{\nu,\psi})$ and is irreducible. Then the following alternative holds:*

1. *There exists some $x \in \mathbb{X}$ such that $\sum_{k=0}^{+\infty} (P^k\psi)(x) = +\infty$: In this case P is recurrent, and μ_R is the unique $P-$invariant positive measure η (up to a multiplicative positive constant) such that $\eta(\psi) < \infty$. Moreover we actually have $\sum_{k=0}^{+\infty} P^k\psi = +\infty$ on $\mathbb{X}$. This corresponds to the case $\mu_R(\psi) = 1$ of Theorem 3.8.*

2. *There exists $x \in \mathbb{X}$ such that $\sum_{k=0}^{+\infty}(P^k\psi)(x) < \infty$: In this case the non-negative function series $\sum_{k=0}^{+\infty} P^k\psi$ is bounded from above on $\mathbb{X}$, and P is transient in the sense given in Assertion 2. of Theorem 3.8.*

Proof. Recall that $\mu_R(\psi) \in (0, 1]$ from Proposition 3.4 and Lemma 3.10. In Case 1., the function series $\sum_{k=0}^{+\infty} P^k\psi$ is not bounded on $\mathbb{X}$, so that P satisfies Case 1. of Theorem 3.8. It follows from Lemma 3.11 that $\sum_{k=0}^{+\infty} P^k\psi = +\infty$ on $\mathbb{X}$. In Case 2., P is not recurrent from Lemma 3.11, so that Case 2. of Theorem 3.8 applies. □

When the positive measure μ_R is finite (i.e. $\mu_R(1_{\mathbb{X}}) < \infty$), then we have $\mu_R(\psi) = 1$ from Theorem 3.6. Moreover any $P-$invariant probability measure π is such that $\pi(\psi) < \infty$ since ψ is bounded. Therefore, the following statement is a direct consequence of Assertion 1. of Theorem 3.8.

Corollary 3.13 *Assume that P satisfies Condition $(\boldsymbol{M}_{\nu,\psi})$ with $\mu_R(1_{\mathbb{X}}) < \infty$ and is irreducible. Then P is recurrent, and the probability measure π_R given in (26) is the unique $P-$invariant probability measure.*

In conclusion, we can now present a statement synthesizing the results of Theorem 3.6 and Corollary 3.13. It provides a classical recurrence criterion that involves neither the positive measure μ_R nor the function h_R^∞. To this end, recall that, depending on the nature of the state space $\mathbb{X}$ and the particular form of the Markov kernel P, many classical results guarantee the existence of a $P-$invariant probability measure (see Subsection 3.5). This is even the starting point in Markov chain Monte Carlo algorithms. Formula (17) again plays a crucial role in the proof of the following theorem.

Theorem 3.14 *Assume that P satisfies Condition $(\boldsymbol{M}_{\nu,\psi})$ and is irreducible. If P admits an invariant probability measure η, then $\eta(\psi) > 0$. Moreover η is the unique $P-$ invariant probability measure, it is equal to π_R given in (26), and P is recurrent.*

Proof. Let η be a $P-$invariant probability measure. If $\eta(\psi) = 0$ then for every $n \geq 1$ we have $\eta(R^n 1_{\mathbb{X}}) = 1$ using (17) applied to $1_{\mathbb{X}}$ and integrating w.r.t. the $P-$invariant probability measure η. Hence it follows from Lebesgue's theorem w.r.t. η that $\eta(h_R^\infty) = 1$ with h_R^∞ given in (20). Thus $\eta(h_R^\infty) = \eta(1_{\mathbb{X}})$, from which we deduce that $h_R^\infty = 1_{\mathbb{X}}$ $\eta-$a.s. since $h_R^\infty \leq 1_{\mathbb{X}}$. Hence there exists $x \in \mathbb{X}$ such that $h_R^\infty(x) = 1$. This provides $\sum_{k=0}^{+\infty}(R^k\psi)(x) = 0$ from (24), which contradicts Assertion 1. of Lemma 3.10. We have proved that $\eta(\psi) > 0$, so that $\mu_R(1_{\mathbb{X}}) < \infty$ from Theorem 3.6. Then the recurrence of P and Equality $\eta = \pi_R$ follow from Corollary 3.13. □

3.4 Further statements

The first proposition concerns the $P-$absorbing sets and is used in the next section, as well as in the proof of Propositions 5.11 and 5.12 related to discussion on drift conditions. The second one concerns the condition $\mu_R(1_A) > 0$ and is used in the bibliographic discussions of Subsection 3.5. The last proposition facilitates the transition to the next section, where the central assumption is $h_R^\infty = 0$.

Proposition 3.15 *If P satisfies Condition $(\boldsymbol{M}_{\nu,\psi})$ and is irreducible, then every non-empty $P-$absorbing set is μ_R-full.*

Proof. Let $B \in \mathcal{X}^*$ be a $P-$absorbing set, i.e.: $\forall n \geq 1,\ \forall x \in B,\ P^n(x, B^c) = 0$. Let $x \in B$. Then it follows from (28) applied to 1_{B^c} and from $P^0(x, B^c) = R^0(x, B^c)$ that $\mu_R(1_{B^c}) = 0$ since $\sum_{n=0}^{+\infty}(P^n\psi)(x) > 0$ from the irreducibility condition (29). □

Proposition 3.16 *Let P satisfy Condition $(\boldsymbol{M}_{\nu,\psi})$. Then we have*

$$\forall A \in \mathcal{X}: \quad \mu_R(1_A) = 0 \implies \sum_{n=0}^{+\infty} P^n 1_A = 0 \ \ \nu-a.s.\ on\ \mathbb{X}. \tag{30}$$

Moreover the Markov kernel P is irreducible (see (29)) if, and only if, $\mu_R(\psi) > 0$ and

$$\forall A \in \mathcal{X}: \quad \mu_R(1_A) > 0 \implies \sum_{n=1}^{+\infty} P^n 1_A > 0 \ \ on\ \mathbb{X}. \tag{31}$$

From this proposition, it follows that, if P satisfies Condition $(\boldsymbol{M}_{\nu,\psi})$ and is irreducible, then the following property holds

$$\forall A \in \mathcal{X}: \quad \mu_R(1_A) > 0 \iff \sum_{n=1}^{+\infty} P^n 1_A > 0 \ \ \text{on } \mathbb{X}. \tag{32}$$

Proof. Let $A \in \mathcal{X}$ be such that $\mu_R(1_A) = 0$, i.e. $\nu(R^k 1_A) = 0$ for any $k \geq 0$ from the definition of μ_R. From (17) we have: $\forall n \geq 0,\ P^n 1_A = R^n 1_A$. Thus

$$\nu\Big(\sum_{n=0}^{+\infty} P^n 1_A\Big) = \mu_R(1_A) = 0$$

from the monotone convergence theorem and the definition of μ_R. This proves (30). Now assume that P is irreducible (see (29)). Then we have $\mu_R(\psi) > 0$ from Lemma 3.10. Moreover Equality (28) reads also as $\sum_{n=1}^{+\infty} P^n = \sum_{n=1}^{+\infty} R^n + \big(\sum_{n=0}^{+\infty} P^n\psi\big) \otimes \mu_R$ since $P^0 = R^0$. Thus, we have

$$\forall A \in \mathcal{X},\ \forall x \in \mathbb{X}, \quad \sum_{n=1}^{+\infty} P^n(x, A) \geq \mu_R(1_A) \sum_{n=0}^{+\infty} (P^n\psi)(x),$$

from which we deduce that (31) holds true. Conversely assume that $\mu_R(\psi) > 0$ and Condition (31) is satisfied. Since there exists $\varepsilon > 0$ such that $\mu_R(1_{\{\psi \geq \varepsilon\}}) > 0$ from $\mu_R(\psi) > 0$, it follows from (31) that $\sum_{n=1}^{+\infty} P^n\psi \geq \varepsilon \sum_{n=1}^{+\infty} P^n 1_{\{\psi \geq \varepsilon\}} > 0$ on $\mathbb{X}$, i.e. (29) holds. □

The main assumption in the next Section 4 is $h_R^\infty = 0$, which obviously implies that each of the equivalent conditions in (25) holds. Simple arguments can be used for specific Markov models to check Condition $h_R^\infty = 0$. This is illustrated in the following proposition for discrete state space case. Recall that in this case the minorization Condition $(\boldsymbol{M}_{\nu,\psi})$ is natural considering each state $s \in \mathbb{X}$ as an atom: $\psi := 1_S$ with $S := \{s\}$ and $\nu := P(s, \cdot)$. Another simple illustration of Property $h_R^\infty = 0$ is given in Subsection 10.2.1 for random walk Metropolis-Hastings Markov kernels on a non-discrete state space $\mathbb{X}$.

Proposition 3.17 *Let P be a Markov kernel on a discrete state space $\mathbb{X}$. If P admits a unique invariant probability measure π which is positive on $\mathbb{X}$ (i.e. $\forall x \in \mathbb{X},\ \pi(1_{\{x\}}) > 0$), then the function h_R^∞ in (20) is zero on $\mathbb{X}$, whatever the minorization Condition $(\boldsymbol{M}_{\nu,\psi})$ considered for P.*

Proof. From Condition $(\boldsymbol{M}_{\nu,\psi})$ and Theorem 3.6 we deduce that $\mu_R(\psi) = 1$ since $\pi(\psi) > 0$ using the positivity of π. Thus $\mu_R(h_R^\infty) = 0$ from (25). Since $\pi = \mu_R(1_{\mathbb{X}})^{-1}\mu_R$ (i.e. $\pi = \pi_R$) from Theorem 3.6 and uniqueness of π, we have $\pi(h_R^\infty) = 0$, so that $h_R^\infty = 0$ using again the positivity of π. □

3.5 Further comments and bibliographic discussion

Here we discuss point by point the definitions and results concerning the classical concepts of this section, i.e. irreducibility, recurrence/transience properties and invariant measures, in link with the books [Num84, MT09, DMPS18]. A detailed historical background on these properties can be found in [Num84, pp. 141-144], [MT09, Sec. 4.5, 8.6,10.6] and [DMPS18, Sec. 9.6,10.4,11.6]. In discrete state space, we refer for example to [Nor97, Bré99, Gra14] (see also [Mey08, App. A] for an overview on Markov chains in modern terms).

A) *Small-set and small-functions.* Let $\ell \geq 1$. Recall that a set $S_\ell \in \mathcal{X}^*$ is said to be a $\ell-$order small-set for P in the standard literature on the topic of Markov chains (e.g. see [Num84, MT09, DMPS18]), if the following condition holds

$$\exists \nu_\ell \in \mathcal{M}_{+,b}^* : \quad P^\ell \geq 1_{S_\ell} \otimes \nu_\ell \ \ \text{(i.e. } \forall x \in \mathbb{X},\ P^\ell(x, dy) \geq 1_{S_\ell}(x)\, \nu_\ell(dy)). \tag{33}$$

The existence of small-sets under the irreducibility condition (see Item C)) was proved in [JJ67]. The extension to $\ell-$order small-functions writes as (see [Num84, Def. 2.3, p. 15])

$$\exists (\nu_\ell, \psi_\ell) \in \mathcal{M}_{+,b}^* \times \mathcal{B}_*^+ : \quad P^\ell \geq \psi_\ell \otimes \nu_\ell \ \ \text{(i.e. } \forall x \in \mathbb{X},\ P^\ell(x, dy) \geq \psi_\ell(x)\, \nu_\ell(dy)). \tag{34}$$

Our minorization condition $(\boldsymbol{M}_{\nu,\psi})$ is nothing other than [Num84, Def. 2.3] with order one. Finally recall that $S \in \mathcal{X}^*$ is said to be petite (e.g. see [MT92]) if it is a small-set of order one for the Markov resolvent kernel $\sum_{n=0}^{+\infty} a_n P^n$ for some $(a_n)_n \in [0, +\infty)^{\mathbb{N}}$ such that $\sum_{n=0}^{+\infty} a_n = 1$. The notion of petite sets is not used in this work. The specific resolvent kernel $\sum_{n=0}^{+\infty} 2^{-(n+1)} P^n$ is only used to prove that μ_R is $\sigma-$finite in Assertion 1. of Theorem 3.8 and in Corollary 4.20 to provide a sufficient condition for having $h_R^\infty = 0$.

B) *Residual kernels and invariant measure.* The representation (21) of $P-$invariant measure via the residual kernel was introduced in [Num78, Th. 3] under the Harris recurrence condition and extended to the recurrent case in [Num84, Th. 5.2, Cor. 5.2], so that the positive measure μ_R necessarily satisfies $\mu_R(\psi) = 1$ there. The $P-$invariance of μ_R under the sole Condition $(\boldsymbol{M}_{\nu,\psi})$ was proved in [HL23a] in the specific case when $\mu_R(1_{\mathbb{X}}) < \infty$: This corresponds to Theorem 3.6. This result is extended to the general case in Theorem 3.5, that is: under the sole minorization Condition $(\boldsymbol{M}_{\nu,\psi})$, the $P-$invariance of μ_R is actually guaranteed when $\mu_R(\psi) = 1$, and is even equivalent to this condition under the additional strong aperiodicity assumption $\nu(\psi) > 0$. Consequently, contrary to the statement [Num84, Th. 5.2, Cor. 5.2, p. 73-74], the $P-$invariance of μ_R is here related directly to the condition $\mu_R(\psi) = 1$, which makes it possible to carry out this study independently of the recurrence property, and even independently of any irreducibility condition on P. Recall that the key point in the proof of Theorem 3.5 is the kernel identity (17).

C) *Accessibility and irreducibility conditions.* Recall that if P satisfies Condition $(\boldsymbol{M}_{\nu,1_S})$ then the set S is said to be a first-order small set. Let us comment Condition (29) in case $\psi := 1_S$. This condition then means that the set S is accessible according to [DMPS18, Def. 3.5.1, Lem. 3.5.2]. On the other hand recall that a Markov kernel P is said to be irreducible according to [DMPS18, Def. 9.2.1] if it admits an accessible small set. Thus our definition (29) of irreducibility for a Markov kernel P satisfying Condition ($\boldsymbol{M}_{\nu,1_S}$) coincides with that of [DMPS18] in case of a first-order small set. Now, thanks to Proposition 3.16, let us recall the link with the irreducibility notion used in [Num84, MT09]. First, in connection with the condition $\mu_R(1_S) = 0$ which is not addressed in Proposition 3.16, observe that this condition implies the transience of P from Theorem 3.8. Moreover this condition cannot hold under Condition (29) from Assertion 1. of Lemma 3.10. Finally, nor can this condition be satisfied under the strong aperiodicity condition $\nu(1_S) > 0$ since $\mu_R \geq \nu$. Thus the discussion may be conducted assuming that P satisfies Condition ($\boldsymbol{M}_{\nu,1_S}$) with $\mu_R(1_S) > 0$ (i.e. $\exists k \geq 0,\ \nu(R^k 1_S) \neq 0$). Then it follows from Proposition 3.16 that our definition of P irreducible (see (29)) is equivalent to the μ_R-irreducibility of P as defined in [Num84, p. 11] and [MT09, p. 82], that is (31). Finally, if P satisfies Condition ($\boldsymbol{M}_{\nu,1_S}$) and is irreducible (see (29)), then Equivalence (32) reads as follows for every $A \in \mathcal{X}$: A is accessible if, and only if, $\mu_R(1_A) > 0$.

D) *Maximal irreducibility measures.* Although the notion of maximal irreducibility measures is not explicitly addressed in this work, it has to be discussed since it plays an important role in [Num84, MT09, DMPS18]. First note that, if P satisfies Conditions ($\boldsymbol{M}_{\nu,1_S}$) and (29), then μ_R is an irreducibility measure using the classical terminology in [MT09, DMPS18] (see Item C)). Actually, from the above remark on accessible sets, μ_R is a maximal irreducibility measure according to the definition [DMPS18, Def. 9.2.2], that is: Every accessible set $A \in \mathcal{X}$ is such that $\mu_R(1_A) > 0$. Of course Conditions ($\boldsymbol{M}_{\nu,1_S}$) and (29) also ensure that the minorizing measure ν is an irreducibility measure since $\nu(1_A) > 0$ implies that $\mu_R(1_A) > 0$. However ν is not maximal a priori. As is well known, any irreducibility measure η is absolutely continuous w.r.t. the maximal irreducibility measure μ_R. Indeed the condition $\eta(1_A) > 0$ implies that $\sum_{n=1}^{+\infty} P^n 1_A > 0$ on $\mathbb{X}$ from the definition of $\eta-$irreducibility, so that $\mu_R(1_A) > 0$ from (32) (thus: $\mu_R(1_A) = 0 \Rightarrow \eta(1_A) = 0$).

E) *Recurrence/transience and uniqueness of invariant measure in recurrence case.* Our definition (27) of recurrence corresponds to that in [Num84, pp. 27-28] and [MT09, p. 180] with μ_R as maximal irreducibility measure. From the discussion in Item C), this also corresponds to the recurrence definition [DMPS18, Def. 10.1.1]. The transience property used in Theorem 3.8 is that provided in [MT09, p. 171 and 180] and [DMPS18, Def. 10.1.3]. The Recurrence/Transience dichotomy stated in Theorem 3.8 is a well-known result for irreducible Markov chains, e.g. see [Num84, Th. 3.2, p. 28], [MT09, Th. 8.0.1] and [DMPS18, Th. 10.1.5]. The novelty in Theorem 3.8 is that this dichotomy can be simply declined according to whether $\mu_R(\psi) = 1$ or $\mu_R(\psi) \in [0, 1)$ under the minorization condition ($\boldsymbol{M}_{\nu,\psi}$). Under the minorization condition ($\boldsymbol{M}_{\nu,\psi}$) and the irreducibility assumption, the equivalence between recurrence and Condition $\mu_R(\psi) = 1$ (or $\nu(h_R^\infty) = 0$) was already highlighted in [Rio00, Prop. 9.2].

As indicated in Item B), the existence of $P-$invariant positive measures is obtained in our work under the minorization Condition ($\boldsymbol{M}_{\nu,\psi}$) and independently of any irreducibility condition on P (Theorem 3.5). Existence of $P-$invariant positive measures is classically proved under the recurrence assumption. In fact this is usually done together with the uniqueness issue. Under the recurrence assumption the existence and uniqueness (up to a positive multiplicative constant) of a $P-$invariant positive measure is obtained in [Num84, Th. 5.2, Cor. 5.2, p. 73-74] using the representation (21). This result is proved in [MT09, Th. 10.4.9] and [DMPS18, Th. 11.2.5] via splitting techniques, providing the classical regeneration-type representation of $P-$invariant positive measures.

Note that Theorem 3.14 does not extend to infinite invariant measures, as illustrated in [DMPS18, Ex. 9.2.17] where the irreducible Markov kernel of a random walk on $\mathbb{X} := \mathbb{Z}$ (the set of integers) is shown to admit at least two infinite and not proportional invariant positive measures. Such a Markov kernel is transient: Otherwise, Case 1 of Theorem 3.8 would apply, and irreducibility property would imply uniqueness for invariant measures (up to a multiplicative positive constant).

F) *On Condition 3. of Theorem 3.6.* Under Condition $(\boldsymbol{M}_{\nu,\psi})$ the following equivalence is proved in Theorem 3.6: There exists a $P-$invariant probability measure η such that $\eta(\psi) > 0$ if, and only if, $\liminf_n \sigma(P^n\psi) > 0$ for some probability measure σ. To the best of our knowledge, this result is new. Similar but stronger conditions were introduced in [Sza03, Th. 2.1] under further topological assumptions on $\mathbb{X}$ and P, also see [MS04, Hor04]. In [HLL03, Th. 10.5.1] the condition $\liminf_n (P^n f_0)(x) > 0$ for some $x \in \mathbb{X}$ and function $f_0 : \mathbb{X} \to (0, +\infty)$ is prove to be a necessary and sufficient condition for the existence of an invariant probability measure: There, P is a weak-feller Markov kernel on a locally compact separable space $\mathbb{X}$, and the function f_0 is not a small-function but some strictly positive continuous function vanishing at infinity. Note that the condition $\liminf_n \sigma(P^n\psi) > 0$ implies that $\sum_{n\geq 0} \sigma(P^n\psi) = +\infty$, so that the series function $\sum_{n\geq 0} P^n\psi$ cannot be bounded. This brings us back to the well-known fact that the existence of a $P-$invariant probability measure implies recurrence. In case $\psi := 1_S$, i.e. P satisfies $(\boldsymbol{M}_{\nu,1_S})$, Condition 3. of Theorem 3.6 is equivalent to the existence of a probability measure σ such that $\limsup_n \sigma(P^n 1_{S^c}) < 1$. From Markov's inequality it can be easily checked that this condition holds if $M := \sup_{n\geq 1} \sigma(P^n W) < \infty$ for some measurable function $W : \mathbb{X} \to [0, +\infty)$ and if there exists some $a > M$ such that $S_a := \{W \leq a\}$ is a first-order small-set for P. When $\mathbb{X}$ is a separable metric space, recall that the condition $\sup_{n\geq 1} \sigma(P^n W) < \infty$ for some probability measure σ, or refinements with $\sigma = \delta_x$ for some $x \in \mathbb{X}$ as $\liminf_k (P^k W)(x) < \infty$ in [MT09, Th. 12.1.3], are classically used to obtain the tightness of the family $(\sigma_n)_{n\geq 1}$ of probability measures defined by

$$\forall n \geq 1,\ \forall A \in \mathcal{X}, \quad \sigma_n(1_A) := \sum_{k=0}^{n-1} \sigma(P^k 1_A),$$

and then to extract a subsequence $(\sigma_{n_k})_{k\geq 1}$ weakly converging to some probability measure σ_∞. Further topological assumptions on P, as Feller property, are however necessary to conclude that σ_∞ is $P-$invariant, e.g. see [MT09, DMPS18, Chap. 12]. Refinements on this topic can be found in [HLL95], where a necessary and sufficient condition of existence of an invariant probability measure for weak Feller Markov kernels is provided, and similarly in [HLL96] concerning the existence of a bounded invariant probability density function. In Theorem 3.6 the minorization condition $(\boldsymbol{M}_{\nu,\psi})$ is assumed, but no topological conditions on $\mathbb{X}$ and P are required.

G) *Strong aperiodicity condition $\nu(\psi) > 0$.* This condition is a particular case of the aperiodicity condition introduced in Subsection 4.3.

H) *The splitting construction.* To conclude this bibliographic discussion, it is worth remembering that the concept of small-set has a natural and crucial probabilistic interest in splitting or coupling techniques: This is the thread and backbone of the books [Num84, MT09, DMPS18]. Here this probabilistic aspect is not addressed. More precisely, the minorization condition ($\boldsymbol{M}_{\nu,\psi}$) allows us to write the Markov kernel P as the sum of two non-negative kernels: the residual kernel $R := P - \psi \otimes \nu$ and the rank-one kernel $\psi \otimes \nu$. That R is non-negative is the crucial point to define all the quantities related to R in this section, especially the positive measure μ_R (see (21)) and the function h_R^∞ (see (20)). Actually one of the key points of the present section and of the next ones is the kernel identity (17). This formula is already present in Nummelin's book [Num84, Eq. (4.12)]. It seems that the sole way to obtain a probabilistic sense of this formula is to use the split Markov chain introduced in [Num78] and [AN78]. The idea is to introduce an appropriate enlargement of the state space of the original Markov chain in order to obtain a new Markov chain - the split chain - which has an atom. Then most of statements on the original chain are derived from applying results (obtained for example by the regeneration method) on atomic chains to this split chain. Thus, using the splitting construction requires switching from the original chain to the split chain for assumptions, and vice versa for results. The enlargement of the state space consists roughly in tagging the transitions of the original chain according to the occurrence of a $\psi-$based tossing coin in order to reflect the decomposition $R + \psi \otimes \nu$ of P in two submarkovian kernels. We refer to [Num84, Sec. 4.4], [CMR05, Sec. 14.2], [MT09, Chap. 5] for details. See also [Num02] for a readable survey on this topic in the case of Markov chain Monte Carlo (MCMC) kernels. Finally mention the recent work [DG15] (also see [DMPS18, Chap. 23]), where the splitting construction is used to obtain subgaussian concentration inequalities for geometrically ergodic Markov chains. Here, the kernel-based point of view allows us to study the general Markov chains in a single step. There is no need to resort to an intermediate class of Markov chains, e.g. atomic chains, before dealing with the general case via what may appear to be a technical device, e.g. the split chain. To turn back to our key formula (17), [Num84, Eq. (4.24)] provides a probabilistic interpretation from the splitting construction. What is new here is that we are exploiting Formula (17) solely as a kernel identity, from which we derive in particular Equality (28) which is the key potential-type formula in this section. The price to pay for this presentation is that we only consider Markov kernels satisfying a first-order minorization condition.

Appendix A gives the probabilistic interpretation of the main quantities used in this document. This should facilitate the comparative reading with the statements in reference probabilistic works as [Num84, MT09, DMPS18]. And, as for formula (17), all these probabilistic formulas are obtained from the split chain.

To conclude this section, recall that all the results of this work can be translated or combined in order to obtain probabilistic statements on Markov chains. As an instance, observing that Condition ($\boldsymbol{M}_{\nu,\psi}$) with $\psi := 1_S$ and the irreducibility condition (29) read as $(*)$ below, Theorem 3.14 provides the following statement:

Theorem $\mathcal{A}$. *Let $(X_n)_{n\geq 0}$ be a Markov chain on $(\mathbb{X}, \mathcal{X})$. Assume that there exist a non-empty set $S \in \mathcal{X}$ and a finite positive measure ν on $(\mathbb{X}, \mathcal{X})$ such that*

$$\forall x \in S,\ \mathbb{P}_x(X_1 \in A) \geq \nu(1_A) \quad \textit{and} \quad \forall x \in \mathbb{X},\ \exists n \geq 1,\ \mathbb{P}_x(X_n \in S) > 0. \tag{$*$}$$

If $(X_n)_{n\geq 0}$ admits a stationary distribution π, then it is the unique one, and the following *recurrence property holds: for every set $A \in \mathcal{X}$ with $\pi-$positive measure, the total number of visits to A, $N_A = \sum_{k=0}^{+\infty} 1_A(X_k)$, is a random variable with infinite expected value, i.e.:*

$$\pi(1_A) > 0 \implies \mathbb{E}_x[N_A] = +\infty.$$

4 Harris recurrence and convergence of the iterates

Assume that the Markov kernel P satisfies the first-order minorization condition $(\boldsymbol{M}_{\nu,\psi})$ and recall that $h_R^\infty := \lim_n R^n 1_{\mathbb{X}}$ (point-wise convergence, see (20)), where $R \equiv R_{\nu,\psi}$ is the residual kernel given in (13). Condition $h_R^\infty = 0$, which is stronger than $\nu(h_R^\infty) = 0$ (see (25)), is central throughout this section. Under this condition, the results of the previous section are revisited in Theorem 4.1 below with an additional result on the $P-$harmonic functions. Next, the Markov kernel P is shown to be Harris-recurrent, and the convergence in total variation norm of the iterates of P to its unique invariant probability measure is obtained when $\mu_R(1_{\mathbb{X}}) < \infty$ and P satisfies an aperiodicity condition. Still under condition $\mu_R(1_{\mathbb{X}}) < \infty$ the periodic case is addressed in Subsection 4.4. If Condition $h_R^\infty = 0$ is dropped, then all these results remain true for the restriction of P to the set $H := \{h_R^\infty = 0\}$: This is relevant because this set is proved to be μ_R-full and $P-$absorbing in Lemma 4.6. Thus, for the above-mentioned statements to hold on the whole state space $\mathbb{X}$, Condition $h_R^\infty = 0$ is indeed required. Introducing a drift inequality on P, a sufficient condition for $h_R^\infty = 0$ is presented in Subsection 4.5. For specific Markov models, simpler arguments can be used to check the condition $h_R^\infty = 0$, as illustrated in Proposition 3.17 in discrete state space case and in Subsection 10.2 in the context of Markov chain Monte Carlo methods. Finally recall that the condition $\mu_R(1_{\mathbb{X}}) < \infty$, which is required in the convergence results of this section, is satisfied if, and only if, there exists a $P-$invariant probability measure w.r.t. which the small function ψ has a positive integral (see Theorem 3.6).

4.1 Preliminaries on Condition $h_R^\infty = 0$

Theorem 4.1 *Let P satisfy Condition $(\boldsymbol{M}_{\nu,\psi})$. If $h_R^\infty = 0$, then the following assertions hold.*

1. *The $P-$harmonic functions are constant on $\mathbb{X}$.*
2. *P is irreducible and recurrent.*
3. *The positive measure $\mu_R := \sum_{k=0}^{+\infty} \nu R^k$ (see (21)) satisfies $\mu_R(\psi) = 1$, and is the unique $P-$invariant positive measure η (up to a multiplicative constant) such that $\eta(\psi) < \infty$. If $\mu_R(1_{\mathbb{X}}) < \infty$, then $\pi_R := \mu_R(1_{\mathbb{X}})^{-1}\mu_R$ (see (26)) is the unique $P-$invariant probability measure on $(\mathbb{X}, \mathcal{X})$.*

From Theorem 4.1 it can be seen that Condition $h_R^\infty = 0$ provides relevant properties on the Markov kernel P. The converse holds too. Indeed, if P satisfies Condition $(\boldsymbol{M}_{\nu,\psi})$ with $\mu_R(1_{\mathbb{X}}) < \infty$ and if $\lim_n P^n\psi = \pi_R(\psi)$ (point-wise convergence), then $h_R^\infty = 0$. This directly follows from Formula (17) applied to the function $1_{\mathbb{X}}$, using Lebesgue's theorem for discrete measure and finally formula $\pi_R(\psi)\mu_R(1_{\mathbb{X}}) = 1$ derived from Theorem 3.6. This is implemented in Corollary 4.8.

Proof. It follows from (24) and $h_R^\infty = 0$ that

$$\sum_{k=0}^{+\infty} R^k\psi = \nu(1_{\mathbb{X}})^{-1} 1_{\mathbb{X}}. \tag{35}$$

Let $g \in \mathcal{B}$ be such that $Pg = g$. Recall that, for every $n \geq 0$, we have $\nu(g)\sum_{k=0}^n R^k\psi = g - R^{n+1}g$ from (22). Moreover we have $\lim_n R^n g = 0$ (point-wise convergence) since $|R^n g| \leq R^n|g| \leq \|g\|_{1_{\mathbb{X}}} R^n 1_{\mathbb{X}}$ and $h_R^\infty = 0$. Thus $g = \nu(g)\sum_{k=0}^{+\infty} R^k\psi$. We have proved that g is proportional to $1_{\mathbb{X}}$. This proves Assertion 1.

For Assertion 2., apply the kernel identity (28) to ψ to get

$$\sum_{n=0}^{+\infty} P^n\psi = \sum_{n=0}^{+\infty} R^n\psi + \mu_R(\psi)\sum_{n=0}^{+\infty} P^n\psi.$$

We have $\mu_R(\psi) = 1$ since $h_R^\infty = 0$ (see (25)). Then, we deduce from (35) and the previous equality that $\sum_{k=0}^{+\infty} P^k\psi = +\infty$. Thus the irreducibility property (29) holds, as well as the recurrence property from Theorem 3.8.

The first part of Assertion 3. is a direct consequence of Assertion 1. of Theorem 3.8 using that $\nu(h_R^\infty) = 0$ (i.e. $\mu_R(\psi) = 1$) and that P is recurrent. The second part of Assertion 3. is Corollary 3.13. The proof of Theorem 4.1 is complete. □

To conclude this preliminary discussion about Condition $h_R^\infty = 0$, let us present some properties on the restriction of the function h_R^∞ to a $P-$absorbing set. The notations concerning restriction to a set $E \in \mathcal{X}$ of functions, measures ans kernels are provided in Section 2. Lemma 4.2 is repeatedly used in this section.

Lemma 4.2 *Assume that P satisfies Condition $(\boldsymbol{M}_{\nu,\psi})$ with $\mu_R(\psi) > 0$, where R is the residual kernel given in* (13)*. Let $E \in \mathcal{X}$ be any μ_R-full $P-$absorbing set. Then the Markov kernel P_E on $(E, \mathcal{X}_E)$ satisfies Condition $(\boldsymbol{M}_{\nu_E,\psi_E})$. Moreover the associated residual kernel $P_E - \psi_E \otimes \nu_E$ is the restriction R_E to E of R, and the following equalities hold*

$$\forall x \in E,\ h_{R_E}^\infty(x) := \lim_n R_E^n(x, E) = h_R^\infty(x) \quad \textit{and} \quad \forall n \geq 0,\ \nu_E(R_E^n\psi_E) = \nu(R^n\psi).$$

Proof. Since $\mu_R(\psi) > 0$ and E is μ_R-full, we have $\mu_R(1_E\psi) = \mu_R(\psi) > 0$, thus ψ_E is non-zero. Moreover we have $\nu(1_E) = \nu(1_{\mathbb{X}}) > 0$ since $\mu_R(1_{E^c}) = 0$ implies that $\nu(1_{E^c}) = 0$ from the definition of μ_R. Then Condition $(\boldsymbol{M}_{\nu_E,\psi_E})$ for the Markov kernel P_E on $(E, \mathcal{X}_E)$ is deduced from the minorization condition $(\boldsymbol{M}_{\nu,\psi})$ for P since for every $A' \in \mathcal{X}_E$ and any $A \in \mathcal{X}$ such that $A' = A \cap E$ we have

$$\forall x \in E, \quad P_E(x, A') = P(x, A \cap E) \geq \nu(A \cap E)\psi(x) = \nu_E(A')\psi_E(x).$$

That $P_E - \psi_E \otimes \nu_E$ is the restriction of R to the set E is obvious. It follows that

$$\forall x \in E,\ \forall n \geq 1, \quad R_E^n(x, E) = R^n(x, E) = R^n(x, \mathbb{X})$$

since $R^n(x, E^c) = 0$ from $0 \leq R^n(x, E^c) \leq P^n(x, E^c) = 0$. Consequently we have for every $x \in E$: $\lim_n R_E^n(x, E) = h_R^\infty(x)$. Finally we have: $\forall n \geq 0, \forall x \in E,\ (R_E^n\psi_E)(x) = (R^n\psi)(x)$. Thus $\nu_E(R_E^n\psi_E) = \nu(R^n\psi)$ since $\nu(1_{E^c}) = 0$. □

4.2 Harris-recurrence

Let us present a first application of Theorem 4.1 to the so-called Harris-recurrence property. Let $(X_n)_{n\geq 0}$ be a Markov chain with transition kernel P. If P satisfies Condition $(\boldsymbol{M}_{\nu,\psi})$ and if $h_R^\infty = 0$, we know that P is recurrent from Theorem 4.1, that is (see (27))

$$\forall A \in \mathcal{X}: \ \mu_R(1_A) > 0 \Longrightarrow \forall x \in \mathbb{X},\ \mathbb{E}_x\Big[\sum_{k=0}^{+\infty} 1_{\{X_k\in A\}}\Big] = +\infty.$$

This recurrence property for P is proved below to be reinforced in

$$\forall A \in \mathcal{X}: \ \mu_R(1_A) > 0 \Longrightarrow \ \forall x \in \mathbb{X}, \quad \mathbb{P}_x\Big\{\sum_{n=1}^{+\infty} 1_{\{X_n\in A\}} = +\infty\Big\} = 1. \tag{36}$$

Such a transition kernel P is said to be *Harris-recurrent*.

Theorem 4.3 *Let P satisfy Conditions $(\boldsymbol{M}_{\nu,\psi})$ and $h_R^\infty = 0$. Then the Markov chain $(X_n)_{n\geq 0}$ with transition kernel P is Harris-recurrent.*

First prove the following lemma.

Lemma 4.4 *Let P satisfy Conditions $(\boldsymbol{M}_{\nu,\psi})$ and $\mu_R(\psi) = 1$. If $g \in \mathcal{B}$ is such that $Pg \leq g$, then the non-negative function $g - Pg$ is μ_R-integrable and we have $\mu_R(g - Pg) = 0$.*

Lemma 4.4, which is used below in the proof of Theorem 4.3, has its own interest. Indeed, from the $P-$invariance of μ_R the conclusion of Lemma 4.4 is straightforward under the assumption $\mu_R(1_{\mathbb{X}}) < \infty$ since, for every $g \in \mathcal{B}$, the functions g and Pg are μ_R-integrable and $\mu_R(Pg) = \mu_R(g)$. However, if μ_R is not finite, the conclusion of Lemma 4.4 is no longer obvious.

Proof of Lemma 4.4. For every $n \geq 1$, it follows from $Pg = Rg + \nu(g)\psi$ that

$$\begin{aligned}\sum_{k=0}^{n} \nu\big(R^k(g - Pg)\big) &= \sum_{k=0}^{n} \nu(R^k g) - \sum_{k=0}^{n} \nu(R^{k+1} g) - \nu(g)\sum_{k=0}^{n} \nu(R^k\psi) \\ &= \nu(g)\Big(1 - \sum_{k=0}^{n} \nu(R^k\psi)\Big) - \nu(R^{n+1}g) \\ &\leq 2\|g\|_{1_{\mathbb{X}}}\nu(1_{\mathbb{X}}) < \infty\end{aligned} \tag{37}$$

using $0 \leq \sum_{k=0}^{n} \nu(R^k\psi) \leq \mu_R(\psi) = 1$ and $|g| \leq \|g\|_{1_{\mathbb{X}}} 1_{\mathbb{X}}$. Thus the series $\sum_{k=0}^{+\infty} \nu(R^k(g - Pg))$ of non-negative terms converges, that is $g - Pg$ is μ_R-integrable. Since $\mu_R(\psi) = 1$ (i.e. $\lim_n \sum_{k=0}^{n} \nu(R^k\psi) = 1$ from the definition of μ_R), we know that $\nu(h_R^\infty) = 0$ from (25). Moreover we have $|\nu(R^{n+1}g)| \leq \|g\|_{1_{\mathbb{X}}}\nu(R^{n+1}1_{\mathbb{X}})$ with $\lim_n \nu(R^{n+1}1_{\mathbb{X}}) = \nu(h_R^\infty) = 0$ from the definition of h_R^∞ and Lebesgue's theorem. Thus the property $\mu_R(g - Pg) = 0$ follows from (37). The proof of Lemma 4.4 is complete. □

Proof of Theorem 4.3. Let $A \in \mathcal{X}$ be such that $\mu_R(1_A) > 0$. Recall that the function defined by $g_A^\infty(x) := \mathbb{P}_x\big\{\sum_{n=1}^{+\infty} 1_{\{X_n\in A\}} = +\infty\big\}$ for any $x \in \mathbb{X}$ is a $P-$harmonic function, see Appendix A for details. Thus, under Condition $h_R^\infty = 0$, we know that g_A^∞ is constant on $\mathbb{X}$ from Theorem 4.1. We have to prove that $g_A^\infty = 1_{\mathbb{X}}$, namely that $g_A^\infty(x) = 1$ for at least one $x \in \mathbb{X}$.

Let g_A be defined by: $\forall x \in \mathbb{X},\ g_A(x) := \mathbb{P}_x\{T_A < \infty\}$ where $T_A := \inf\{n \geq 0 : X_n \in A\}$ is the hitting time of the set A. Recall that g_A is superharmonic, i.e. $Pg_A \leq g_A$, and that $g_A^\infty = \lim_n \searrow P^n g_A$, see Appendix A for details. Let $n \geq 0$. It follows from $P(P^n g_A) \leq P^n g_A$ and Lemma 4.4 applies to $P^n g_A$ that the non-negative function $P^n g_A - P^{n+1} g_A$ is such that $\mu_R(P^n g_A - P^{n+1} g_A) = 0$. Thus there exists $E_n \in \mathcal{X}$ such that $\mu_R(1_{E_n^c}) = 0$ and $P^n g_A = P^{n+1} g_A$ on E_n. Now let $E := \cap_{n\geq 0} E_n$. Then we have $\mu_R(1_{E^c}) = 0$ and

$$\forall x \in E,\ \forall n \geq 0, \quad g_A(x) = (P^{n+1} g_A)(x).$$

Passing to the limit when $n \to +\infty$ we obtain that every $x \in E$ satisfies $g_A^\infty(x) = g_A(x)$. Finally we get from $\mu_R(1_{E^c}) = 0$ that $\mu_R(1_{A\cap E}) = \mu_R(1_A) > 0$, and we know that $g_A = 1$ on A from the definition of g_A. Therefore there exists a $x \in \mathbb{X}$ such that $g_A^\infty(x) = 1$. Thus $g_A^\infty = 1_{\mathbb{X}}$ since g_A^∞ is constant on $\mathbb{X}$. The proof of Theorem 4.3 is complete. □

Corollary 4.5 *If P satisfies Condition $(\boldsymbol{M}_{\nu,\psi})$, is irreducible and recurrent, then the restriction P_H of P to the μ_R−full P−absorbing set $H := \{h_R^\infty = 0\}$ is Harris-recurrent.*

The proof of Corollary 4.5 is based on Lemma 4.2 and on the following lemma.

Lemma 4.6 *Assume that P satisfies Condition $(\boldsymbol{M}_{\nu,\psi})$ and is irreducible. If $\nu(h_R^\infty) = 0$, then the set $H := \{h_R^\infty = 0\}$ is P−absorbing and μ_R−full.*

Proof. Since $\nu(h_R^\infty) = 0$ the set H is non-empty. Moreover it follows from $\nu(h_R^\infty) = 0$ and $Rh_R^\infty = h_R^\infty$ that $Ph_R^\infty = h_R^\infty$. Then we have

$$\forall x \in H, \quad 0 = h_R^\infty(x) = (Ph_R^\infty)(x) = \int_{\mathbb{X}} h_R^\infty(y) P(x, dy) \ = \int_{H^c} h_R^\infty(y) P(x, dy)$$

hence $P(x, H^c) = 0$, i.e. $P(x, H) = 1$, for any $x \in H$. Thus H is P−absorbing. That H is μ_R−full follows from Proposition 3.15. □

Proof of Corollary 4.5. We have $\nu(h_R^\infty) = 0$ and $\mu_R(\psi) = 1$ from Corollary 3.9. It follows from Lemma 4.6 that $H := \{h_R^\infty = 0\}$ is P−absorbing and μ_R−full. From Lemma 4.2 applied to the set H, we know that P_H satisfies Condition $(\boldsymbol{M}_{\nu_H,\psi_H})$ and that $h_{R_H}^\infty = 0$ on H from the definition of H. Consequently the last assertion of Corollary 4.5 follows from Theorem 4.3 applied to the Markov kernel P_H on $(H, \mathcal{X}_H)$. □

4.3 Convergence of iterates: the aperiodic case

Set $\overline{D} := \{z \in \mathbb{C} : |z| \leq 1\}$. If P satisfies Condition $(\boldsymbol{M}_{\nu,\psi})$, then the following power series

$$\rho(z) := \sum_{n=1}^{+\infty} \nu(R^{n-1}\psi)\, z^n \tag{38}$$

absolutely converges for every $z \in \overline{D}$ since $\mu_R(\psi) = \sum_{k=0}^{+\infty} \nu(R^k\psi) < \infty$ from Proposition 3.4. If moreover P is irreducible, then this power series ρ is non-zero since $\rho(1) = \mu_R(\psi) > 0$ from Assertion 1. of Lemma 3.10.

If P satisfies Condition $(\boldsymbol{M}_{\nu,\psi})$ and is irreducible, then P is said to be *aperiodic* if $\rho(z)$ defined in (38) is not a power series in z^q for any integer $q \geq 2$. From above the set

$\{n \geq 1 : \nu(R^{n-1}\psi) > 0\}$ is non-empty. Then, using the notation g.c.d. for *greatest common divisor*, this aperiodicity condition is equivalent to

$$\text{g.c.d.}\left\{n \geq 1 : \nu(R^{n-1}\psi) > 0\right\} = 1. \tag{39}$$

This condition obviously holds when P is strongly aperiodic, i.e. $\nu(\psi) > 0$. In Subsection 4.4, under Conditions $(\boldsymbol{M}_{\nu,\psi})$ and $h_R^\infty = 0$, various equivalent conditions for aperiodicity are provided by Theorem 4.14. Actually, Assertion (b) of Theorem 4.14 shows that the aperiodicity condition does not depend on the choice of the couple (ν, ψ) in Condition $(\boldsymbol{M}_{\nu,\psi})$. Moreover, Assertion (c) of Theorem 4.14 shows that aperiodicity condition is equivalent to the non-existence of d-cycle sets for P with $d \geq 2$.

When P satisfies Condition $(\boldsymbol{M}_{\nu,\psi})$ with $\mu_R(1_{\mathbb{X}}) < \infty$, is irreducible and aperiodic, the convergence of probability distributions $(\delta_x P^n)_{n\geq 0}$ to π_R in total variation norm is shown to be equivalent to the property $h_R^\infty = 0$ in the following theorem. As a corollary, the convergence of the probability distributions $(\delta_x P^n)_{n\geq 0}$ to π_R holds for π_R-almost every $x \in \mathbb{X}$. Recall that under these assumptions, π_R is the unique $P-$invariant probability measure from Assertion 3. of Theorem 4.1.

Theorem 4.7 *Let P satisfy Condition $(\boldsymbol{M}_{\nu,\psi})$ with $\mu_R(1_{\mathbb{X}}) < \infty$. If P is irreducible and aperiodic, then the following equivalence holds:*

$$h_R^\infty = 0 \iff \forall x \in \mathbb{X},\ \lim_{n\to+\infty} \|\delta_x P^n - \pi_R\|_{TV} = 0.$$

Corollary 4.8 *Let P satisfy Condition $(\boldsymbol{M}_{\nu,\psi})$ with $\mu_R(1_{\mathbb{X}}) < \infty$. If P is irreducible and aperiodic, then*

$$\lim_{n\to+\infty} \|\delta_x P^n - \pi_R\|_{TV} = 0 \quad \textit{for } \pi_R\textit{-almost every } x \in \mathbb{X}.$$

More precisely the $P-$absorbing and π_R-full set $H := \{h_R^\infty = 0\}$ satisfies

$$H := \left\{x \in \mathbb{X} : \lim_{n\to+\infty} \|\delta_x P^n - \pi_R\|_{TV} = 0\right\}.$$

As expected, Corollary 4.8 follows from Theorem 4.7 applied to the restriction of P to the $P-$absorbing and μ_R-full set $H := \{h_R^\infty = 0\}$ (Lemma 4.6). The proof of Corollary 4.8 is detailed at the end of this subsection. Just observe here that, from the end of this proof, the following fact holds under the sole assumption $(\boldsymbol{M}_{\nu,\psi})$ with $\mu_R(1_{\mathbb{X}}) < \infty$: The convergence $\lim_n (P^n\psi)(x) = \pi_R(\psi)$ for some $x \in \mathbb{X}$ implies that $h_R^\infty(x) = 0$.

Proof of Theorem 4.7. The proof follows from the two next lemmas. Indeed assume that $h_R^\infty = 0$. Then $\lim_n P^n\psi = \pi_R(\psi)1_{\mathbb{X}}$ (point-wise convergence) from Lemma 4.9, thus the desired convergence in total variation norm holds from Lemma 4.11. Conversely assume that, for every $x \in \mathbb{X}$, we have $\lim_{n\to+\infty} \|\delta_x P^n - \pi_R\|_{TV} = 0$. Then it follows from the definition of $\|\cdot\|_{TV}$ that $\lim_{n\to+\infty}(P^n\psi)(x) = \pi_R(\psi)$ since ψ is bounded. Thus $h_R^\infty = 0$ from Lemma 4.9. □

Lemma 4.9 *Let P satisfy Condition $(\boldsymbol{M}_{\nu,\psi})$ with $\mu_R(1_{\mathbb{X}}) < \infty$. If P is irreducible and aperiodic, then*

$$h_R^\infty = 0 \iff \lim_{n\to+\infty} P^n\psi = \pi_R(\psi)1_{\mathbb{X}} \quad (\textit{point-wise convergence on } \mathbb{X}).$$

Proof. Let $D = \{z \in \mathbb{C} : |z| < 1\}$. The following power series

$$\mathcal{P}(z) := \sum_{n=0}^{+\infty} z^n P^n \psi \quad \text{and} \quad \mathcal{R}(z) := \sum_{n=0}^{+\infty} z^n R^n \psi \tag{40}$$

are well-defined on D since ψ is bounded. Note that $\mathcal{P}(z)$ and $\mathcal{R}(z)$ are function series. From the kernel identity (17) applied to ψ it follows that

$$\begin{aligned}\forall z \in D, \quad \mathcal{P}(z) = \sum_{n=0}^{+\infty} z^n P^n \psi &= \sum_{n=0}^{+\infty} z^n R^n \psi + \sum_{n=1}^{+\infty} z^n \sum_{k=1}^{n} \nu(R^{k-1}\psi) P^{n-k}\psi \\ &= \mathcal{R}(z) + \rho(z)\mathcal{P}(z).\end{aligned}$$

where $\rho(z)$ is the power series defined in (38). Using $\mu_R(\psi) = \sum_{k=1}^{+\infty} \nu(R^{k-1}\psi) = 1$ from Theorem 3.6, we have: $\forall z \in D,\ |\rho(z)| < 1$. Thus

$$\forall z \in D, \quad \mathcal{P}(z) = \mathcal{R}(z)\, U(z) \quad \text{with} \quad U(z) := \frac{1}{1-\rho(z)}. \tag{41}$$

Next, for any $k \geq 1$, we have $\nu(R^k 1_{\mathbb{X}}) = \nu(R^{k-1}1_{\mathbb{X}}) - \nu(1_{\mathbb{X}})\nu(R^{k-1}\psi)$ from $R1_{\mathbb{X}} = 1_{\mathbb{X}} - \nu(1_{\mathbb{X}})\psi$. Thus,

$$\forall k \geq 1, \quad \nu(1_{\mathbb{X}})\nu(R^{k-1}\psi) = \nu(R^{k-1}1_{\mathbb{X}}) - \nu(R^k 1_{\mathbb{X}})$$

and

$$\begin{aligned}\forall n \geq 1, \quad \nu(1_{\mathbb{X}}) \sum_{k=1}^{n} k\, \nu(R^{k-1}\psi) &= \sum_{k=1}^{n} k \big[\nu(R^{k-1}1_{\mathbb{X}}) - \nu(R^k 1_{\mathbb{X}})\big] \\ &= \sum_{k=1}^{n} k\, \nu(R^{k-1}1_{\mathbb{X}}) - \sum_{k=2}^{n+1} (k-1)\, \nu(R^{k-1}1_{\mathbb{X}}) \\ &= \sum_{k=1}^{n} \nu(R^{k-1}1_{\mathbb{X}}) - n\, \nu(R^n 1_{\mathbb{X}}).\end{aligned}$$

Hence $m := \sum_{k=1}^{+\infty} k\, \nu(R^{k-1}\psi) \leq \mu_R(1_{\mathbb{X}})\nu(1_{\mathbb{X}})^{-1} < \infty$. Now recall that $\sum_{k=1}^{+\infty} \nu(R^{k-1}\psi) = 1$ and that $\rho(z)$ is not a power series in z^q for any integer $q \geq 2$ since P is assumed to be aperiodic. Consequently the Erdös-Feller-Pollard renewal theorem [EFP49] provides the following property for the power series $U(z) = \sum_{k=0}^{+\infty} u_k z^k$ in (41):

$$\lim_{k \to +\infty} u_k = \frac{1}{m}.$$

Let $x \in \mathbb{X}$. Identifying the coefficients of the power series in Equation (41) (Cauchy product), we obtain that for every $n \geq 0$

$$(P^n\psi)(x) = \sum_{k=0}^{n} u_{n-k} (R^k\psi)(x) = \sum_{k=0}^{+\infty} v_n(k) (R^k\psi)(x) \quad \textit{with} \quad \forall k \geq 0,\ v_n(k) := u_{n-k} 1_{[0,n]}(k).$$

For every $k \geq 1$, we have $\lim_n v_n(k) = 1/m$, and $|v_n(k)| \leq \sup_j |u_j| < \infty$. Moreover recall that $\sum_{k=0}^{+\infty} (R^k\psi)(x) < \infty$ from Proposition 3.4. Then it follows from Lebesgue theorem w.r.t. discrete measure that

$$\forall x \in \mathbb{X}, \quad \lim_n (P^n\psi)(x) = \frac{1}{m} \sum_{k=0}^{+\infty} (R^k\psi)(x). \tag{42}$$

Now we can prove Lemma 4.9. If $h_R^\infty = 0$, then we have $\sum_{k=0}^{+\infty}(R^k\psi)(x) = \nu(1_{\mathbb{X}})^{-1}$ from (35). Hence (42) provides: $\forall x \in \mathbb{X}$, $\lim_n (P^n\psi)(x) = (m\nu(1_{\mathbb{X}}))^{-1}$. Actually the constant $(m\nu(1_{\mathbb{X}}))^{-1}$ equals to $\pi_R(\psi)$ from Lebesgue theorem w.r.t. the $P-$invariant probability measure π_R. The direct implication in Lemma 4.9 is proved. Conversely, assume that $\lim_n P^n\psi = \pi_R(\psi)1_{\mathbb{X}}$ (point-wise convergence). Then we deduce from (42) that $\sum_{k=0}^{+\infty} R^k\psi = c\,1_{\mathbb{X}}$ with $c := m\pi_R(\psi)$. Thus $h_R^\infty = d\,1_{\mathbb{X}}$ with $d = 1 - c\nu(1_{\mathbb{X}})$ from (24). Finally recall that $\mu_R(\psi) = 1$, thus $\nu(h_R^\infty) = 0$ from (25). Hence $d\,\nu(1_{\mathbb{X}}) = 0$, from which we deduce that $h_R^\infty = 0$. □

Remark 4.10 *From the proof of Lemma 4.9 we deduce the following facts. If P satisfies Condition $(\boldsymbol{M}_{\nu,\psi})$ with $\mu_R(1_{\mathbb{X}}) < \infty$, then $m := \sum_{k=1}^{+\infty} k\,\nu(R^{k-1}\psi) < \infty$. If moreover P is irreducible and aperiodic and if $h_R^\infty = 0$, then $m = (\pi_R(\psi)\nu(1_{\mathbb{X}}))^{-1}$.*

Lemma 4.11 *Assume that P satisfies Condition $(\boldsymbol{M}_{\nu,\psi})$ and $\mu_R(1_{\mathbb{X}}) < \infty$. If $h_R^\infty = 0$ and $\lim_n P^n\psi = \pi_R(\psi)1_{\mathbb{X}}$ (point-wise convergence on $\mathbb{X}$), then $\lim_n \|\delta_x P^n - \pi_R\|_{TV} = 0$ for every $x \in \mathbb{X}$.*

Proof. Using (17) and $\pi_R = \pi_R(\psi)\sum_{k=1}^{+\infty} \nu R^{k-1}$ (see (26)), we have for every $n \geq 1$ and $g \in \mathcal{B}$

$$P^n g - \pi_R(g)1_{\mathbb{X}} = R^n g + \sum_{k=1}^{n} \nu(R^{k-1}g)\big(P^{n-k}\psi - \pi_R(\psi)1_{\mathbb{X}}\big) - \pi_R(\psi)\bigg(\sum_{k=n+1}^{+\infty} \nu(R^{k-1}g)\bigg)1_{\mathbb{X}}.$$

Thus

$$\|\delta_x P^n - \pi_R\|_{TV} \leq (R^n 1_{\mathbb{X}})(x) + \sum_{k=1}^{n} \nu(R^{k-1}1_{\mathbb{X}})\big|(P^{n-k}\psi)(x) - \pi_R(\psi)\big| + \pi_R(\psi)\sum_{k=n+1}^{+\infty} \nu(R^{k-1}1_{\mathbb{X}}).$$

We have $\lim_n (R^n 1_{\mathbb{X}})(x) = 0$ from $h_R^\infty = 0$. The term $\sum_{k=n+1}^{+\infty} \nu(R^{k-1}1_{\mathbb{X}})$ also converges to zero when $n \to +\infty$ since $\sum_{k=0}^{+\infty} \nu(R^k 1_{\mathbb{X}}) = \mu_R(1_{\mathbb{X}}) < \infty$. Next note that

$$\sum_{k=1}^{n} \nu(R^{k-1}1_{\mathbb{X}})\big|(P^{n-k}\psi)(x) - \pi_R(\psi)\big| = \sum_{k=1}^{+\infty} \nu(R^{k-1}1_{\mathbb{X}}) f_n(k)$$

with $f_n(k) := |(P^{n-k}\psi)(x) - \pi_R(\psi)|1_{[1,n]}(k)$. Then, using $\sum_{k=1}^{+\infty} \nu(R^{k-1}1_{\mathbb{X}}) < \infty$, the above sum converges to zero when $n \to +\infty$ from Lebesgue's theorem w.r.t. discrete measure since, for every $k \geq 1$, we have $f_n(k) \leq 2\|\psi\|_{1_{\mathbb{X}}}$ and $\lim_n f_n(k) = 0$ by hypothesis. Lemma 4.11 is proved. □

Proof of Corollary 4.8. From Theorem 3.6 we have $\mu_R(\psi) = 1$, so that $\nu(h_R^\infty) = 0$ from (25). Then we know from Lemma 4.6 that the set $H := \{h_R^\infty = 0\}$ is $P-$absorbing and μ_R-full. From Lemma 4.2 applied to $E := H$, it follows that P_H satisfies Condition $(\boldsymbol{M}_{\nu_H,\psi_H})$ with $h_{R_H}^\infty = 0$ from the definition of H, and that g.c.d. $\{n \geq 1 : \nu_H(R_H^{n-1}\psi_H) > 0\} = 1$ since $\nu_H(R_H^{n-1}\psi_H) = \nu(R^{n-1}\psi)$. Thus P_H is irreducible from Theorem 4.1 applied to P_H, and P_H is aperiodic too. Finally note that the positive measure $\sum_{k=0}^{+\infty} \nu_H R_H^k$ is the restriction $\mu_{R|H}$ of μ_R to the set H, so that $\mu_{R|H}(\psi_H) = 1$ since $\mu_R(\psi) = 1$ and H is μ_R-full. Moreover the restriction $\pi_{R|H}$ of π_R to H is a P_H-invariant probability measure on $(H, \mathcal{X}_H)$. Hence Theorem 4.7 applied to P_H shows that, for every $x \in H$, we have $\lim_n \|\delta_x P_H^n - \pi_{R|H}\|_{TV} = 0$. Finally, since we have for every $x \in H$ and $A \in \mathcal{X}$

$$P^n(x, A) - \pi_R(1_A) = P^n(x, A \cap H) - \pi_R(1_{A\cap H}) = P_H^n(x, A \cap H) - \pi_{R|H}(1_{A\cap H})$$

we obtain that: $\forall x \in H,\ \lim_n \|\delta_x P^n - \pi_R\|_{TV} = 0$. This provides the first assertion of Corollary 4.8 since $\pi_R(1_H) = 1$ from $\mu_R(1_{H^c}) = 0$.

It follows from this first assertion that, to obtain the last equality of Corollary 4.8, we just have to prove that, if $x \in \mathbb{X}$ is such that $\lim_n \|\delta_x P^n - \pi_R\|_{TV} = 0$, then $x \in H$, i.e. $h_R^\infty(x) = 0$. Let $x \in \mathbb{X}$ satisfy the previous condition. Then we have $\lim_n (P^n\psi)(x) = \pi_R(\psi)$ since ψ is bounded. Then, passing to the limit when $n \to +\infty$ in (17) applied to the function $1_{\mathbb{X}}$ at this point x, we have

$$1 = h_R^\infty(x) + \pi_R(\psi) \sum_{k=1}^{+\infty} \nu(R^{k-1}1_{\mathbb{X}}) = h_R^\infty(x) + \pi_R(\psi)\mu_R(1_{\mathbb{X}}).$$

The first equality follows from the definition (20) of h_R^∞ and from $\lim_n (P^n\psi)(x) = \pi_R(\psi)$ using Lebesgue's theorem for discrete measure; The second equality follows from the definition (21) of μ_R. Next, since $\pi_R(\psi) = \mu_R(1_{\mathbb{X}})^{-1}$ from Theorem 3.6, we obtain that $h_R^\infty(x) = 0$. The proof of Corollary 4.8 is complete. □

4.4 Convergence of iterates: the periodic case

Assume that P satisfies Condition $(\boldsymbol{M}_{\nu,\psi})$ and is irreducible. Recall that the power series $\rho(z)$ given in (38), namely

$$\rho(z) := \sum_{n=1}^{+\infty} \nu(R^{n-1}\psi)\, z^n$$

is defined on $\overline{D} = \{z \in \mathbb{C} : |z| \leq 1\}$ and is non-zero (see the beginning of Subsection 4.3). Define

$$d := \text{g.c.d.}\, \{n \geq 1 : \nu(R^{n-1}\psi) > 0\} \tag{43}$$

where g.c.d. stands for greatest common divisor computed on a non-empty set.

Under the assumptions of Theorem 4.12 below, i.e. P satisfies $(\boldsymbol{M}_{\nu,\psi})$, $h_R^\infty = 0$ and $\mu_R(1_{\mathbb{X}}) < \infty$, it follows from Theorem 4.1 that P is irreducible, and that π_R is the unique $P-$invariant probability measure. The convergence in total variation norm of the probability measures $(1/d)\sum_{r=0}^{d-1} \delta_x P^{nd+r}$ to π_R is obtained in Theorem 4.12. Under the assumptions of Corollary 4.13 below, i.e. P satisfies $(\boldsymbol{M}_{\nu,\psi})$ with $\mu_R(1_{\mathbb{X}}) < \infty$ and is irreducible, it follows from Corollary 3.13 that π_R is the unique $P-$invariant probability measure. Under these conditions, the above-mentioned convergence in total variation norm is proved to hold π_R-a.s. in Corollary 4.13. In fact these two statements are the natural extensions to the periodic case of Theorem 4.7 and Corollary 4.8.

In the context of this section, various equivalent characterizations of Integer d in (43) are presented in Theorem 4.14 below. It turns out that, under the assumptions of both Theorem 4.12 and Corollary 4.13, the value of d does not depend on the choice of the couple (ν, ψ): This is proved in Remark 4.18 at the end of this subsection. Therefore, Integer d in (43) can be called the *period* of P without any ambiguity. If $d = 1$, then P is aperiodic according to the definition of Subsection 4.3. If $d \geq 2$, then P is said to be *periodic*.

Theorem 4.12 *Let P satisfy Condition $(\boldsymbol{M}_{\nu,\psi})$ with $\mu_R(1_{\mathbb{X}}) < \infty$ and $h_R^\infty = 0$. If P is* *periodic with period $d \geq 2$ (see (43)), then the following convergence holds:*

$$\forall x \in \mathbb{X}, \qquad \lim_{n \to +\infty} \big\| \pi_R - \frac{1}{d} \sum_{r=0}^{d-1} \delta_x P^{nd+r} \big\|_{TV} = 0.$$

The proof of Theorem 4.12 is similar to that of the direct implication of Theorem 4.7 (where $d = 1$). When $d \geq 2$, the proof is just a little more technical, since we have to work with the sums $(1/d) \sum_{r=0}^{d-1} \delta_x P^{nd+r}$. This proof is postponed in Appendix B.

Corollary 4.13 *Assume that P satisfies Condition $(\boldsymbol{M}_{\nu,\psi})$ with $\mu_R(1_{\mathbb{X}}) < \infty$, and is irreducible. If P is periodic with $d \geq 2$ in (43), then the following convergence holds :*

$$\lim_{n \to +\infty} \big\| \pi_R - \frac{1}{d} \sum_{r=0}^{d-1} \delta_x P^{nd+r} \big\|_{TV} = 0 \quad \text{for } \pi_R\text{−almost every } x \in \mathbb{X}.$$

Proof. Using the restriction P_H of P to the μ_R−full P−absorbing set $H := \{h_R^\infty = 0\}$ from Lemma 4.6, Corollary 4.13 is deduced from Theorem 4.12 proceeding as for Corollary 4.8: Use g.c.d. $\{n \geq 1 : \nu_H(R_H^{n-1}\psi_H) > 0\} = d$ from $\nu_H(R_H^{n-1}\psi_H) = \nu(R^{n-1}\psi)$, and apply Theorem 4.12 to obtain that

$$\forall x \in H, \quad \big\| \pi_R - \frac{1}{d} \sum_{r=0}^{d-1} \delta_x P^{nd+r} \big\|_{TV} = 0.$$

□

In the next statement the space $\mathcal{B} = \mathcal{B}_{1_{\mathbb{X}}}$ is extended to complex-valued functions, i.e.:

$$\mathcal{B}(\mathbb{C}) := \big\{ g : \mathbb{X} \to \mathbb{C}, \text{measurable such that } \|g\|_{1_X} := \sup_{x \in \mathbb{X}} |g(x)| < \infty \big\}$$

where $|\cdot|$ stands here for the modulus in $\mathbb{C}$. Recall that $z \in \mathbb{C}$ is said to be an eigenvalue of P on $\mathcal{B}(\mathbb{C})$ if there exists a non-zero function $g \in \mathcal{B}(\mathbb{C})$ such that $Pg = zg$. Finally recall that, under Conditions $(\boldsymbol{M}_{\nu,\psi})$ and $h_R^\infty = 0$, the positive integer $d = \text{g.c.d.}\,\{n \geq 1 : \nu(R^{n-1}\psi) > 0\}$ in (43) is well-defined in the next statement.

Theorem 4.14 *Assume that P satisfies Condition $(\boldsymbol{M}_{\nu,\psi})$ and $h_R^\infty = 0$. Let $\rho(z)$ be the power series given in* (38), *and let $d :=$ g.c.d. $\{n \geq 1 : \nu(R^{n-1}\psi) > 0\}$. Then the following assertions are satisfied:*

(*a*) *The complex numbers z of modulus one satisfying $\rho(z) = 1$ are the d−th roots of unity.*

(*b*) *The eigenvalues of modulus one of P on $\mathcal{B}(\mathbb{C})$ are the d−th roots of unity.*

(*c*) *There exist a μ_R−full P−absorbing set $E \in \mathcal{X}$ and d disjoints sets $C_0, \ldots, C_{d-1}$ in $\mathcal{X}$ such that*

$$E = \bigsqcup_{\ell=0}^{d-1} C_\ell \quad \text{with} \quad \forall \ell = 0, \ldots, d-1,\ \forall x \in C_\ell, \quad P(x, C_{\ell+1}) = 1 \tag{44}$$

using the convention $C_d = C_0$. Moreover d is the greatest positive integer for which such a cycle property holds.

Any of these three conditions characterizes the integer $d := g.c.d.\{n \geq 1 : \nu(R^{n-1}\psi) > 0\}$.

The proof of Theorem 4.14 is based on the following three lemmas.

Lemma 4.15 *Let P satisfy Condition $(\boldsymbol{M}_{\nu,\psi})$ and $h_R^\infty = 0$. Let $d := g.c.d.\{n \geq 1 : \nu(R^{n-1}\psi) > 0\}$. Let $z \in \mathbb{C}$ be such that $|z| = 1$. Then we have $\rho(z) = 1$ if, and only if, z is a $d-$th root of unity.*

Proof. Recall that $\mu_R(\psi) = \sum_{n=1}^{+\infty} \nu(R^{n-1}\psi) = 1$ from Theorem 4.1. Let $z \in \mathbb{C}$ be such that $|z| = 1$ and $\rho(z) = 1$. Then

$$\sum_{n=1}^{+\infty} \nu(R^{n-1}\psi)\, z^n = 1 = \sum_{n=1}^{+\infty} \nu(R^{n-1}\psi).$$

Writing $z = e^{i\theta}$ with $\theta \in [0, 2\pi)$ we obtain that $\sum_{n=1}^{+\infty} \big(1 - \cos(n\theta)\big)\nu(R^{n-1}\psi) = 0$. Define the set $\mathcal{N} := \{n \geq 1 : \nu(R^{n-1}\psi) > 0\}$. Then $n \in \mathcal{N}$ implies that $\cos(n\theta) = 1$. Thus we have: $\forall n \in \mathcal{N}, z^n = 1$. Next from the definition of d, for p large enough there exists $\{n_j\}_{j=1}^p \in \mathcal{N}^p$ such that $d = \sum_{j=1}^p k_j n_j$ for some $\{k_j\}_{j=1}^p \in \mathbb{Z}^p$ (Bézout identity). Thus we have $z^d = \prod_{j=1}^p z^{k_j n_j} = 1$ since $z^{n_j} = 1$. Hence z is a $d-$th root of unity.

Conversely, let z be a $d-$th root of unity, i.e. $z^d = 1$. From the definition of d it then follows that $\rho(z) = \sum_{k=0}^{+\infty} \nu(R^{kd-1}\psi)\, z^{kd} = \mu_R(\psi) = 1$. □

Lemma 4.16 *Let P satisfy Condition $(\boldsymbol{M}_{\nu,\psi})$ and $h_R^\infty = 0$. Let $z \in \mathbb{C}$ be such that $|z| = 1$. Then z is an eigenvalue of P on $\mathcal{B}(\mathbb{C})$ if, and only if, we have $\rho(z) = 1$. Moreover, if any of these two equivalent conditions holds, then*

$$E_z := \{g \in \mathcal{B}(\mathbb{C}) : Pg = zg\} = \mathbb{C} \cdot \widetilde{\psi}_z \quad \textit{with} \quad \widetilde{\psi}_z := \sum_{k=0}^{+\infty} z^{-(k+1)} R^k \psi.$$

Proof. First note that, for any $z \in \mathbb{C}$ such that $|z| = 1$, the above function $\widetilde{\psi}_z$ is well-defined and belongs to $\mathcal{B}(\mathbb{C})$ from Proposition 3.4. Moreover observe that

$$\nu(\widetilde{\psi}_z) = \sum_{k=0}^{+\infty} z^{-(k+1)} \nu(R^k\psi) = \rho(z^{-1}), \tag{45}$$

the exchange between series and $\nu-$integral being valid since $\sum_{k=0}^{+\infty} \nu(R^k\psi) < \infty$ from Proposition 3.4. Now, let $z \in \mathbb{C}$, $|z| = 1$, and let $g \in \mathcal{B}(\mathbb{C})$, $g \neq 0$, be such that $Pg = zg$. Thus we have $\nu(g)\psi = (zI - R)g$ from $P = R + \psi \otimes \nu$. Then we have for every $n \geq 0$

$$\begin{aligned} \nu(g) \sum_{k=0}^{n} z^{-(k+1)} R^k \psi = \Big(\sum_{k=0}^{n} z^{-(k+1)} R^k\Big)(zI - R)g \;&=\; \sum_{k=0}^{n} z^{-k} R^k g - \sum_{k=0}^{n} z^{-(k+1)} R^{k+1} g \\ &=\; g - z^{-(n+1)} R^{n+1} g. \end{aligned} \tag{46}$$

Moreover we have $|R^n g| \leq \|g\|_{1_{\mathbb{X}}} R^n 1_{\mathbb{X}}$, so $\lim_n R^n g = 0$ (point-wise convergence) from Condition $h_R^\infty = 0$. Hence $g = \nu(g)\widetilde{\psi}_z$, with $\nu(g) \neq 0$ since $g \neq 0$ by hypothesis. From (45) it follows that $\nu(g) = \nu(g)\rho(z^{-1})$, thus $\rho(z^{-1}) = 1$, or equivalently $\rho(z) = 1$ from $z^{-1} = \overline{z}$

(the conjugate of z) since $|z| = 1$ and the coefficients of the power series $\rho(\cdot)$ are real (even non-negative).

Conversely let $z \in \mathbb{C}$, $|z| = 1$, be such that $\rho(z) = 1$, thus $\rho(z^{-1}) = 1$. From (45) we have $\nu(\widetilde{\psi}_z) = 1$. Using $P = R + \psi \otimes \nu$ and Lebesgue's theorem w.r.t. $R(x, dy)$ for each $x \in \mathbb{X}$ we obtain that

$$P\widetilde{\psi}_z = z \sum_{k=0}^{+\infty} z^{-(k+2)} R^{k+1}\psi + \nu(\widetilde{\psi}_z)\psi = z\big(\widetilde{\psi}_z - z^{-1}\psi\big) + \psi \ = \ z\widetilde{\psi}_z. \tag{47}$$

Thus z is an eigenvalue of P on $\mathcal{B}(\mathbb{C})$ since $\widetilde{\psi}_z \neq 0$ from $\nu(\widetilde{\psi}_z) = 1$. The claimed equivalence in Lemma 4.16 is proved. The last assertion follows from the first part of the proof, where we obtained that any $g \in \mathcal{B}(\mathbb{C})$ such that $Pg = zg$ with $|z| = 1$ satisfies $g = \nu(g)\widetilde{\psi}_z$. □

Lemma 4.17 *Let P satisfy Condition $(\boldsymbol{M}_{\nu,\psi})$ and $h_R^\infty = 0$. Assume that, for some integer $d_1 \geq 1$, there exist a μ_R−full P−absorbing set $E \in \mathcal{X}$ and d_1 disjoints sets $C_0, \ldots, C_{d_1-1}$ in $\mathcal{X}$ such that*

$$E = \bigsqcup_{\ell=0}^{d_1-1} C_\ell \quad \textit{with} \quad \forall \ell \in \{0, \ldots, d_1 - 1\},\ \forall x \in C_\ell, \quad P(x, C_{\ell+1}) = 1$$

using the convention $C_{d_1} = C_0$. Then every d_1−th root of unity is solution to equation $\rho(z) = 1$. Moreover d_1 divides d with $d :=$ g.c.d. $\{n \geq 1 : \nu(R^{n-1}\psi) > 0\}$.

Proof. Let z be any d_1−th root of unity and define $g_E : E \to \mathbb{C}$ by

$$\forall \ell \in \{0, \ldots, d_1 - 1\},\ \forall x \in C_\ell, \quad g_E(x) = z^\ell.$$

Then we have for every $\ell \in \{0, \ldots, d_1 - 1\}$, and $x \in C_\ell$

$$(P_E g_E)(x) = \int_E g_E(y)P(x, dy) = \int_{C_{\ell+1}} g_E(y)P(x, dy) = z^{\ell+1} = z\, g_E(x)$$

since $P(x, C_{\ell+1}) = 1$ and $g_E(x) = z^\ell$, recalling moreover for the case $\ell = d_1 - 1$ that $C_{d_1} = C_0$ by convention and that $1 = z^{d_1}$. Thus $P_E g_E = z g_E$. Next recall that $\mu_R(\psi) = 1$ from Theorem 4.1. It then follows from Lemma 4.2 that P_E satisfies Condition $(\boldsymbol{M}_{\nu_E,\psi_E})$ on $(E, \mathcal{X}_E)$, that $h_{R_E}^\infty = 0$ on E from the assumption $h_R^\infty = 0$, and finally that

$$\forall z \in \overline{D}, \quad \rho_E(z) := \sum_{n=1}^{+\infty} \nu_E(R_E^{n-1}\psi_E)\, z^n = \rho(z).$$

We can now conclude. Since z is an eigenvalue of P_E, Lemma 4.16 applied to P_E ensures that $\rho_E(z) = 1$, so $\rho(z) = 1$. This proves the first assertion. That d_1 divides d follows from Lemma 4.15. □

Now we prove Theorem 4.14.

Proof. Assertion (a) holds with d given in (43) from Lemma 4.15. Thus so is for Assertion (b) from Lemma 4.16. Now prove that Assertion (c) holds with d given in (43). Let $z_d = e^{2i\pi/d}$,

$\widetilde{\psi}_d := \sum_{k=0}^{+\infty} z_d^{-(k+1)} R^k\psi$, and let $\widetilde{\psi}_{d,0}$ (resp. $\widetilde{\psi}_{d,1}$) denote the real (resp. imaginary) part of the function $\widetilde{\psi}_d$. Then it follows from (35) that

$$\widetilde{\psi}_{d,0} \leq |\widetilde{\psi}_d| \leq \sum_{k=0}^{+\infty} R^k\psi = \nu(1_{\mathbb{X}})^{-1}1_{\mathbb{X}}.$$

We have $\rho(z_d^{-1}) = 1$ from Assertion (a), thus $\nu(\widetilde{\psi}_d) = 1$ from (45). Then we have $\nu(\widetilde{\psi}_{d,0}) = 1 = \nu(\nu(1_{\mathbb{X}})^{-1}1_{\mathbb{X}})$, so that the following equalities hold $\nu-$a.e. on $\mathbb{X}$: $\widetilde{\psi}_{d,0} = \nu(1_{\mathbb{X}})^{-1}1_{\mathbb{X}}$ and $\widetilde{\psi}_{d,1} = 0$. Now define $g_d := \nu(1_{\mathbb{X}})\widetilde{\psi}_d$. From the above we know that $|g_d| \leq 1_{\mathbb{X}}$ and that the set $C_0 := \{g_d = 1\}$ is non-empty. Moreover we have $Pg_d = z_d g_d$ from Lemma 4.16. Let $x \in C_0$. Then

$$1 = g_d(x) = \frac{(Pg_d)(x)}{z_d} = \int_{\mathbb{X}} \frac{g_d(y)}{z_d} P(x, dy)$$

with $|g_d(y)/z_d| \leq 1$ for every $y \in \mathbb{X}$ since $|z_d| = 1$. It follows that $P(x, C_1) = 1$ where $C_1 := \{x \in \mathbb{X} : g_d(x) = z_d\}$. Replacing the set C_0 with C_1, we can similarly prove that, for every $x \in C_1$, we have $P(x, C_2) = 1$ where $C_2 := \{x \in \mathbb{X} : g_d(x) = {z_d}^2\}$. Repeating this arguments provides the existence of sets $C_0, \ldots, C_{d-1}$ in $\mathcal{X}$ satisfying the desired cycle property: $\forall \ell = 0, \ldots, d-1$, $\forall x \in C_\ell$, $P(x, C_{\ell+1}) = 1$. These sets are obviously disjoint. Finally define $E := \bigsqcup_{\ell=0}^{d-1} C_\ell$. This set is $P-$absorbing since, for every $x \in E$, there exists a (unique) $\ell \in \{0, \ldots, d-1\}$ such that $x \in C_\ell$, so that $1 = P(x, C_{\ell+1}) \leq P(x, E) \leq 1$, thus $P(x, E) = 1$. Since P is irreducible from Theorem 4.1, the set E is μ_R-full from Proposition 3.15. We have proved that P satisfies the $d-$cycle property (44) with d defined in (43). The fact that d is the greatest integer for which such a cycle property holds then follows from Lemma 4.17. □

Remark 4.18 *The fact that, under the assumptions of Theorem 4.12, the value d in (43) does not depend on the choice of the couple (ν, ψ) directly follows from Assertion (b) of Theorem 4.14. Now let us prove that the same conclusion holds under the assumptions of Corollary 4.13. That is, assume that P satisfies the conditions of Corollary 4.13 w.r.t. two couples $(\nu, \psi) \in \mathcal{M}_{+,b}^* \times \mathcal{B}_+^*$ and $(\nu', \psi') \in \mathcal{M}_{+,b}^* \times \mathcal{B}_+^*$, and prove that the respective integers d and d' given by (43) are equal. Recall that P admits a unique invariant probability measure, say π, from Corollary 3.13. In particular we have $\pi_R = \pi_{R'} = \pi$. Moreover we have $\mu_R(\psi) = 1$ and $\mu_{R'}(\psi') = 1$, equivalently $\nu(h_R^\infty) = 0$ and $\nu'(h_{R'}^\infty) = 0$, from Theorem 3.6 (see (25)). It then follows from Lemma 4.6 that the sets $H := \{h_R^\infty = 0\}$ and $H' := \{h_{R'}^\infty = 0\}$ are $\pi-$full and $P-$absorbing.*

Now let us consider any $d-$cycle partition E given by Assertion c) of Theorem 4.14 when applied to the restriction P_H of P to H. Note that E is then a $\pi-$full and $P-$absorbing subset of H. Restricting this $d-$cycle partition to the set $E' := E \cap H'$ provides a $d-$cycle partition for $P_{H'}$ since $E' := E \cap H'$ is a $\pi-$full and $P-$absorbing subset of H'. Then Lemma 4.17 applied to $P_{H'}$ shows that d divides d'. Exchanging the role of (ν, ψ) and (ν', ψ') we obtain that d' divides d. Thus $d = d'$.

4.5 Drift criteria for $h_R^\infty = 0$

Now, we introduce a drift condition to have the property $h_R^\infty := \lim_n R^n 1_{\mathbb{X}} = 0$, the relevance of which has been highlighted in Theorems 4.1, 4.3, 4.7, 4.12 and 4.14. Actually, under a drift inequality w.r.t. some measurable function $W : \mathbb{X} \to [0,+\infty)$, the property $h_R^\infty = 0$ is characterized in Proposition 4.19 by a control of h_R^∞ or $\sum_{k=0}^{+\infty} R^k\psi$ on any level set $\mathcal{W}_r := \{x \in \mathbb{X} : W(x) \leq r\}$ of W. Finally, a condition ensuring this control is provided by Corollary 4.20.

Proposition 4.19 *Let P satisfy Condition $(\boldsymbol{M}_{\nu,\psi})$ and the following drift condition for some measurable function $W : \mathbb{X} \to [0,+\infty)$:*

$$\exists b > 0, \quad PW \leq W + b\,\psi. \tag{48}$$

For any $r > 0$ let $\mathcal{W}_r$ denote the level set of order r defined by: $\mathcal{W}_r := \{x \in \mathbb{X} : W(x) \leq r\}$. Then we have the following equivalences

$$h_R^\infty = 0 \iff \forall r > 0,\ \sup_{x\in\mathcal{W}_r} h_R^\infty(x) < 1 \iff \forall r > 0,\ \inf_{x\in\mathcal{W}_r} \sum_{k=0}^{+\infty} (R^k\psi)(x) > 0. \tag{49}$$

Proof. The second equivalence in (49) follows from (24). That $h_R^\infty = 0$ implies the second condition in (49) is obvious. It remains to prove that the second condition in (49), or equivalently the third one, implies that $h_R^\infty = 0$.

In the sequel, the third condition in (49) is assumed to hold. First prove that we have the following point-wise convergence on $\mathbb{X}$

$$\forall \rho > 0, \quad \lim_n R^n 1_{\mathcal{W}_\rho} = 0. \tag{50}$$

Let $\rho > 0$ and define $a \equiv a_\rho := \inf_{x\in\mathcal{W}_\rho} \sum_{k=0}^{+\infty} (R^k\psi)(x)$. By hypothesis we have $a > 0$ and $1_{\mathcal{W}_\rho} \leq a^{-1} \sum_{k=0}^{+\infty} R^k\psi$, from which we deduce that

$$\forall n \geq 1, \quad 0 \leq R^n 1_{\mathcal{W}_\rho} \leq a^{-1} \sum_{k=n}^{+\infty} R^k\psi$$

from the monotone convergence theorem w.r.t. $R^n(x, dy)$ for each $x \in \mathbb{X}$. Property (50) then holds since the series $\sum_{k=0}^{+\infty} R^k\psi$ converges point-wise from Proposition 3.4.

Next note that $\nu(W)\psi \leq PW$ everywhere on $\mathbb{X}$ from $(\boldsymbol{M}_{\nu,\psi})$, so that $\nu(W) < \infty$ and RW is well-defined. Let $d := \max(0\,, (b - \nu(W))/\nu(1_{\mathbb{X}}))$ and prove that

$$RW_d \leq W_d \quad \text{where } W_d := W + d1_{\mathbb{X}}. \tag{51}$$

Note that $\nu(W_d) = \nu(W) + d\nu(1_{\mathbb{X}}) < \infty$ and that $PW_d = PW + d1_{\mathbb{X}}$. It then follows from $RW_d = PW_d - \nu(W_d)\psi$ and from the drift inequality (48) that

$$RW_d \leq W + b\psi + d1_{\mathbb{X}} - \big(\nu(W) + d\,\nu(1_{\mathbb{X}})\big)\psi \leq W_d + \big(b - \nu(W) - d\,\nu(1_{\mathbb{X}})\big)\psi$$

so that $RW_d \leq W_d$ from the definition of d.

Now let us deduce from (50) and (51) that $h_R^\infty = 0$. Let $r > d$ with d given by (51). We have

$$1_{\mathbb{X}} = 1_{\{x\in\mathbb{X}:W_d(x)>r\}} + 1_{\{x\in\mathbb{X}:W_d(x)\leq r\}} \leq \frac{W_d}{r} + 1_{\mathcal{W}_{r-d}}.$$

Thus we get

$$\forall n \geq 1, \quad R^n 1_{\mathbb{X}} \leq \frac{R^n W_d}{r} + R^n 1_{\mathcal{W}_{r-d}} \leq \frac{W_d}{r} + R^n 1_{\mathcal{W}_{r-d}}$$

from the non-negativity of R and from $R^n W_d \leq W_d$ using (51) and an immediate induction. Let $x \in \mathbb{X}$, $\varepsilon > 0$, and fix $r > d$ large enough so that $W_d(x)/r < \varepsilon/2$. From (50) applied to $\rho = r - d$, there exists $N \geq 1$ such that, for every $n \geq N$, we have $0 \leq (R^n 1_{\mathcal{W}_{r-d}})(x) < \varepsilon/2$. Thus: $\forall n \geq N,\ 0 \leq (R^n 1_{\mathbb{X}})(x) < \varepsilon$. This proves that $h_R^\infty = 0$.

□

We conclude this section providing an alternative sufficient condition for $h_R^\infty = 0$. Let us consider the Markov resolvent kernel Q defined by

$$Q := \sum_{n=0}^{+\infty} 2^{-(n+1)} P^n.$$

Corollary 4.20 *Let P satisfy Condition $(\boldsymbol{M}_{\nu,\psi})$ and the drift condition (48) for some measurable function $W : \mathbb{X} \to [0, +\infty)$. If the following condition holds*

$$\forall r > 0, \ \inf_{x \in \mathcal{W}_r} (Q\psi)(x) > 0, \tag{52}$$

then $h_R^\infty = 0$.

Proof. Below we prove that the third condition in (49) is fulfilled. The claimed conclusion then follows from Proposition 4.19. Recall that $\psi \in \mathcal{B}_+^*$, so that $Q\psi$ and the series $\sum_{n=0}^{+\infty} 2^{-(n+1)} R^n \psi$ are well-defined. From (17) we obtain that

$$\begin{aligned} Q\psi &= \sum_{n=0}^{+\infty} 2^{-(n+1)} R^n \psi + \sum_{n=1}^{+\infty} 2^{-(n+1)} \sum_{k=1}^{n} \nu(R^{k-1}\psi) P^{n-k}\psi \\ &= \sum_{n=0}^{+\infty} 2^{-(n+1)} R^n \psi + \left(\sum_{k=1}^{+\infty} 2^{-k} \nu(R^{k-1}\psi)\right)\left(\sum_{n=0}^{+\infty} 2^{-(n+1)} P^n \psi\right) \\ &= \sum_{n=0}^{+\infty} 2^{-(n+1)} R^n \psi + \alpha\, Q\psi \end{aligned}$$

where $\alpha := \sum_{k=1}^{+\infty} 2^{-k} \nu(R^{k-1}\psi)$. Note that, either $\alpha = 0$, or $\alpha < \mu_R(\psi) \leq 1$ from Proposition 3.4, so that

$$\sum_{n=0}^{+\infty} 2^{-(n+1)} R^n \psi = (1-\alpha)\, Q\psi \quad \text{with } 1 - \alpha > 0.$$

Now let $r > 0$ and $a \equiv a_r := \inf_{x \in \mathcal{W}_r} (Q\psi)(x)$. We have $a > 0$ from (52), and

$$\forall x \in \mathcal{W}_r, \quad \sum_{k=0}^{+\infty} (R^k \psi)(x) \geq \sum_{k=0}^{+\infty} 2^{-(k+1)} (R^k \psi)(x) = (1-\alpha)\,(Q\psi)(x) \geq (1-\alpha)a > 0.$$

The third condition in (49) is proved. □

Condition (52) on Q is obviously satisfied under the following stronger condition

$$\forall r > 0, \ \exists q \equiv q(r) \geq 1, \ \inf_{x \in \mathcal{W}_r} (P^q \psi)(x) > 0. \tag{53}$$

Note that requiring Condition (53) means requiring that the irreducibility property for P (see (29)) holds uniformly on each level set $\mathcal{W}_r$. This condition is relevant only for unbounded function W. Indeed, otherwise, the set $\mathcal{W}_r$ is the whole space $\mathbb{X}$ for r large enough, and in this case Condition (53) is restrictive since it requires that $\inf_{x\in\mathbb{X}}(P^q\psi)(x) > 0$ for some $q \geq 1$. If $\mathbb{X}$ is discrete (say $\mathbb{X} = \mathbb{N}$) and $W = (W(n))_{n\in\mathbb{N}}$ is an unbounded increasing sequence, then the sets $\mathcal{W}_r$ are finite: In this case, Condition (53) holds if, and only if,

$$\forall s \in \mathbb{N},\ \exists q \equiv q(s) \geq 1,\ \forall i \in \{0, \dots, s\}, \quad (P^q\psi)(i) > 0.$$

If X is a non-discrete topological space, then a natural assumption for Condition (53) to be fulfilled is that, for every $r > 0$, the set $\mathcal{W}_r$ is compact. However this is not sufficient. An additional natural assumption is that P is weakly Feller (i.e. if $g \in \mathcal{B}$ is continuous on $\mathbb{X}$, then so is Pg). Under these two assumptions, Condition (53) actually holds provided that there exists a bounded and continuous function ψ_0 such that $0 \leq \psi_0 \leq \psi$ and

$$\forall r > 0,\ \exists q \equiv q(r) \geq 1,\ \forall x \in \mathcal{W}_r, \quad (P^q\psi_0)(x) > 0.$$

Indeed the continuous function $P^q\psi_0$ then reaches its lower bound on the compact set $\mathcal{W}_r$, and this lower bound is thus positive under the previous condition.

4.6 Further comments and bibliographic discussion

In the present bibliographic discussion we assume that P is irreducible. The uniqueness of $1_{\mathbb{X}}$ (up to a multiplicative constant) as $P-$harmonic functions is classically studied in link with the Harris-recurrence property, the concept of which was introduced in [Har56]. A nice and comprehensive account of what is Harris recurrence in probabilistic terms is presented in [Bax11]. The study of $P-$harmonic functions is done in [Num84, Th. 3.8, p. 44], [MT09, Th. 17.1.5] and [DMPS18, Th. 10.2.11], essentially using the fact that, for a Markov chain $(X_n)_{n\geq 0}$ on $\mathbb{X}$ and for every $A \in \mathcal{X}$, the function $g_A^\infty : x \mapsto \mathbb{P}_x\{X_k \in A \text{ i.o.}\}$ is a $P-$harmonic function, where i.o. stands for infinitely often. Similarly, under the aperiodicity condition, the Harris-recurrence assumption is classically used to prove the convergence in total variation of the iterates of P to its (unique) invariant probability measure π (i.e. $\forall x \in \mathbb{X},\ \lim_n \|\delta_x P^n - \pi\|_{TV} = 0$). This is proved in [MT09, Ths. 13.0.1, 13.3.5] and [DMPS18, Th. 11.3.1] via renewal theory and splitting construction, also see [AN78] for a proof based on the random renewal time approach and [RR04, Th. 4] for a proof based on coupling method.

In this section, assuming that P satisfies the minorization condition $(\boldsymbol{M}_{\nu,\psi})$, we choose a different approach, first focusing on function $h_R^\infty := \lim_n R^n 1_{\mathbb{X}}$ introduced in the previous section. Indeed the condition $h_R^\infty = 0$ enables us to prove the above conclusion on $P-$harmonic functions (Theorem 4.1), from which the Harris-recurrent property can be derived in Theorem 4.3 using the fact that for every $A \in \mathcal{X}$ the function $x \mapsto \mathbb{P}_x\{X_k \in A \text{ i.o.}\}$ is $P-$harmonic (no surprise there). In the case when measure μ_R is finite and P is aperiodic, the condition $h_R^\infty = 0$ is proved to be equivalent to the above mentioned iterate convergence in total variation (Theorem 4.7). So, to put it simply, the presentation in this section and the resulting statements focus on the condition $h_R^\infty = 0$ depending on the residual kernel R, rather than on the Harris-recurrence property. However note that the proof of Theorem 4.7 is original: Actually Property (24) and the power series formula (41) simply derived from the key equality (17) allow us to directly apply the renewal theorem proved in the seminal paper [EFP49] by Erdös, Feller and Pollard, to the power series $\rho(z)$ in (38) used to define the aperiodicity condition. Finally mention that, for the direct implication in the equivalence of Lemma 4.9, the renewal theorem in [Fel67, Th 1, p. 330] can be directly applied too.

If P is recurrent, then the $P-$harmonic functions are still constant, but up to a negligible set w.r.t. to some maximal irreducibility measure, e.g. see [Num84, Prop. 3.13, p. 44]. In the same way, if P admits an invariant probability measure π, so that P is recurrent from a classical result (e.g. see [DMPS18, Th. 10.1.6]), then the property $\lim_n \|\delta_x P^n - \pi_R\|_{TV} = 0$ is known to hold for $\pi-$almost every $x \in \mathbb{X}$, e.g. see [DMPS18, Th. 11.3.1] and [RR04, pp. 32-33]. This is here highlighted using the explicit set $H := \{h_R^\infty = 0\}$ which is $P-$absorbing and μ_R-full under the recurrence condition (see Corollary 4.5 and the proof of Corollary 4.8). Complements using splitting construction can be found in [Num84, Cor. 5.1, p. 71].

Under the irreducibility condition, the d-cycle property for P stated in Assertion (c) of Theorem 4.14 is the standard definition of the period of P, see [MT09, p. 114] and [DMPS18, Def. 9.3.5]. In our work, under the minorization Condition ($\boldsymbol{M}_{\nu,\psi}$) and irreducibility condition, Integer d is defined by $d :=$ g.c.d.$\{n \geq 1 : \nu(R^{n-1}\psi) > 0\}$. Then the alternative characterizations of this integer d, in particular the d-cycle property for P, are proved under the condition $h_R^\infty = 0$ in Theorem 4.14. The convergence in total variation norm stated in Theorem 4.12 corresponds to the standard statements [MT09, Th. 13.3.4] and [DMPS18, Cor. 11.3.2], except that the condition $h_R^\infty = 0$ is used here in Theorem 4.12 instead of the Harris-recurrence condition in [MT09, DMPS18]. In the same way the π_R-a.e. convergence in total variation norm obtained in Corollary 4.13 corresponds to the standard results in [MT09, Th. 13.3.4] and [DMPS18, Cor. 11.3.2]. Again the direct use of the μ_R-full $P-$absorbing set $H := \{h_R^\infty = 0\}$ provides a short proof of Corollary 4.13. The proofs in [MT09, Th. 13.3.4] and [DMPS18, Cor. 11.3.2] are based on the $d-$cyclic decomposition. The proof given in Appendix B does not rely on the $d-$cycles property: it adapts the arguments of the direct implication of Theorem 4.7 to the periodic case, thus directly giving the conclusion of Theorem 4.14.

The sufficient condition provided in Proposition 4.19 for the condition $h_R^\infty = 0$ to hold is the analogue of the standard statements ensuring that P is recurrent or Harris-recurrent under drift condition, e.g. see [Num84, Prop. 5.10, p. 77], [MT09, Th. 8.4.3, Th. 9.1.8], [DMPS18, Th. 10.2.13]. As recently proved in [XZZ18] under $\phi-$irreducibility condition, such a drift condition (up to a $\phi-$null set) is even a necessary and sufficient condition for recurrence. The drift inequality (48) in Proposition 4.19 is the same as in the previously cited works. Moreover Condition (49) in Proposition 4.19 replaces the classical assumption that W is unbounded off petite set (i.e. each level set $\mathcal{W}_r := \{W \leq r\}$ is a petite set). This last condition means that, for every $r > 0$, there exists $a := (a_n)_{n\geq 0} \in [0,1]^{\mathbb{N}}$ with $\sum_{n=0}^{+\infty} a_n = 1$ and a positive measure $\nu_{r,a}$ such that $Q_a \geq 1_{\mathcal{W}_r} \otimes \nu_{r,a}$ where $Q_a := \sum_{n=0}^{+\infty} a_n P^n$. Expressed with $a_n := 2^{-(n+1)}$, this assumption is clearly stronger than Condition (52) in Corollary 4.20, which only focusses on the lower bound of the function $Q\psi$ on $\mathcal{W}_r$ (no minorizing measure is involved in (52)).

The notion of modulated drift condition is introduced in the next section, where additional bibliographic discussion is provided (see Subsection 5.5). Before diving into the details of these modulated drift conditions, let us present some comment on the probabilistic meaning of the simpler drift condition (48). Let $(X_n)_{n\geq 0}$ be a Markov chain with state space $\mathbb{X}$ and transition kernel P. Let $W : \mathbb{X} \to [0, +\infty)$ be measurable. For any $r > 0$ the set $\mathcal{W}_r = \{x \in \mathbb{X} : W(x) \leq r\}$ must be thought of as the level set of order r in $\mathbb{X}$ w.r.t. the function W. Since $(PW)(x) = \mathbb{E}_x\big[W(X_1)\big]$ for any $x \in \mathbb{X}$, the Markov kernel P satisfies

Condition (48) with $\psi := 1_{\mathcal{W}_s}$ for some $s > 0$ if, and only if,

$$\sup_{x\in\mathcal{W}_s} \mathbb{E}_x\big[W(X_1)\big] < \infty \quad \text{and} \quad \forall x \in \mathbb{X}\setminus\mathcal{W}_s, \quad \mathbb{E}_x\big[W(X_1)\big] \le W(x). \tag{54}$$

The second condition in (54) means that, for any $r > s$, each point $x \in \mathbb{X}$ such that $W(x) = r$ transits in mean into $\mathcal{W}_r$. If $\mathbb{X} := \mathbb{R}^d$ is equipped with some norm $\|\cdot\|$, then W may be of the form $W = v(\|\cdot\|)$ with unbounded increasing function $v : [0, +\infty) \to [0, +\infty)$. In particular, if $W = \|\cdot\|$, then the second condition in (54) means that, starting from $x \in \mathbb{R}^d$ far enough from the origin, the state visited after a first transition of the Markov chain admits in mean a norm less than $\|x\|$, namely is closer to the origin. For a random walk on $\mathbb{N}$, it means that, for i large enough, the steps of the walker starting from i are in mean more to the left than to the right, namely it tends to go back towards 0. In case $\mathbb{X} := \mathbb{Z}$ and $W(x) := |x|$, a typical illustration of the explicit computations needed for obtaining the drift inequality (48) can be found in [MT09, Sect. 8.4.3] for random walks with bounded range and zero mean increment. If $(\mathbb{X}, d)$ is a metric space and $W(x) := d(x, x_0)$, level sets are the balls centred at x_0. However the possibility of considering other level functions more suited to the transition kernel (i.e. possibly considering level sets other than balls) offers flexibility for the validity of Conditions (54) or of the modulated drift condition involved in the next sections.

To illustrate what can be deduced for Markov chains from this section, let us complete the probabilistic Theorem $\mathcal{A}$ stated at the end of Section 3. The following statement follows from Corollary 4.8 and Theorem 4.14 restricted to the absorbing and μ_R−full set $H := \{h_R^\infty = 0\}$:

Theorem $\mathcal{B}$. *Let $(X_n)_{n\ge0}$ be a Markov chain on $(\mathbb{X}, \mathcal{X})$ with stationary distribution π. Assume that Assumption $(*)$ of Theorem $\mathcal{A}$ holds, and that $(X_n)_{n\ge0}$ is aperiodic, i.e.: There do not exist an integer $d \ge 2$, a π−full absorbing set $E \in \mathcal{X}$ and sets $C_0, \dots, C_{d-1}$ in $\mathcal{X}$ such that (with the convention $C_d = C_0$)*

$$E = \bigsqcup_{\ell=0}^{d-1} C_\ell \quad \text{with} \quad \forall \ell = 0, \dots, d-1,\ \forall x \in C_\ell, \quad \mathbb{P}_x\big(X_1 \in C_{\ell+1}\big) = 1.$$

Then the following convergence in total-variation distance holds for π−almost every $x \in \mathbb{X}$:

$$\lim_{n\to+\infty} \sup_{A\in\mathcal{X}} \big|\mathbb{P}_x(X_n \in A) - \pi(1_A)\big| = 0.$$

5 Modulated drift condition and Poisson's equation

Throughout this section, the Markov kernel P is assumed to satisfy the first-order minorization condition $(\boldsymbol{M}_{\nu,\psi})$. Then, the following V_1−modulated drift condition is introduced: $PV_0 \le V_0 - V_1 + b\psi$ with some measurable function $V_0 : \mathbb{X} \to [1, +\infty)$ and the so-called modulated measurable function $V_1 : \mathbb{X} \to [1, +\infty)$. The minorization condition is the first pillar in this work, this modulated drift condition is the second one. Note that the modulated drift condition is a re-enforcement of the drift inequality (48) of Proposition 4.19.

Under the minorization Condition $(\boldsymbol{M}_{\nu,\psi})$ and the V_1−modulated drift condition, the convergence of the series $\sum_{k=0}^{+\infty} R^k V_1$ is proved in Theorem 5.3. Then the series $\sum_{k=0}^{+\infty} R^k 1_{\mathbb{X}}$ converges point-wise since $1_{\mathbb{X}} \le V_1$, so that the function $h_R^\infty := \lim_n R^n 1_{\mathbb{X}}$ (see (20)) is zero on $\mathbb{X}$. Under the same assumptions it is also shown in Theorem 5.3 that the positive measure μ_R given in (21) is finite, i.e. $\mu_R(1_{\mathbb{X}}) < \infty$. Accordingly, when Condition $(\boldsymbol{M}_{\nu,\psi})$ and the

V_1-modulated drift condition are assumed to hold, all the conclusions of Theorems 4.1, 4.3, and Theorem 4.7 or 4.12 hold true, that is:

(i) *The $P-$harmonic functions are constant on $\mathbb{X}$.*

(ii) *P is irreducible (see (29)) and recurrent (see (27)).*

(iii) *The positive measure μ_R (see (21)) is finite (i.e. $\mu_R(1_{\mathbb{X}}) < \infty$) and satisfies $\mu_R(\psi) = 1$. Moreover it is the unique (up to a positive multiplicative constant) $P-$invariant positive measure η such that $\eta(\psi) < \infty$.*

(iv) *$\pi_R := \mu_R(1_{\mathbb{X}})^{-1}\mu_R$ (see (26)) is the unique $P-$invariant probability measure on $(\mathbb{X}, \mathcal{X})$, we have $\pi_R(\psi) > 0$, and P is Harris-recurrent (see (36)).*

(v) *The convergence in total variation of Theorem 4.7 or Theorem 4.12, depending on whether P is aperiodic or periodic, holds.*

Actually the convergence of the series $\sum_{k=1}^{+\infty} R^k V_1$ gives more, in particular it naturally provides solutions to the so-called Poisson's equation (Theorem 5.4). This is the main motivation of this section.

5.1 Modulated drift condition $\boldsymbol{D}_\psi(V_0, V_1)$

Let us introduce the following condition for any couple (V_0, V_1) of measurable functions from $\mathbb{X}$ to $[1, +\infty)$:

$$\exists \psi \in \mathcal{B}_+^*,\ \exists b_0 \equiv b_0(V_0, V_1, \psi) > 0: \quad PV_0 \leq V_0 - V_1 + b_0\psi. \tag{$\boldsymbol{D}_\psi(V_0, V_1)$}$$

This condition is said to be a V_1-*modulated drift condition* for P, and V_0 and V_1 in $\boldsymbol{D}_\psi(V_0, V_1)$ are called *Lyapunov functions* for P. The functions V_0, V_1, ψ are assumed to be everywhere finite, so the function PV_0 is too. It is worth noticing that the modulated function V_1 must be larger than one for the results of this section to hold. In fact, it is only required that V_0 is non-negative and V_1 is uniformly bounded from below by a positive constant. Indeed, if $PV_0' \leq V_0' - V_1' + b'\psi$ for some positive constant b' and some measurable functions $V_0' \geq 0$ and $V_1' \geq c1_{\mathbb{X}}$ with $c > 0$, then Condition $\boldsymbol{D}_\psi(V_0, V_1)$ holds with $V_1 := V_1'/c \geq 1_{\mathbb{X}}$, $V_0 := 1_{\mathbb{X}} + V_0'/c \geq 1_{\mathbb{X}}$ and $b_0 := b'/c > 0$. Moreover observe that if Conditions $\boldsymbol{D}_\phi(V_0, V_1)$ for some $\phi \in \mathcal{B}_+^*$ is satisfied then $\boldsymbol{D}_\psi(V_0, V_1)$ holds for any $\psi \in \mathcal{B}_+^*$ such that $\psi \geq \phi$ (using any constant $b_0(V_0, V_1, \psi)$ larger than $b_0(V_0, V_1, \phi)$).

In the special case $\psi := 1_S$ for some $S \in \mathcal{X}^*$, the above condition writes as

$$\exists S \in \mathcal{X}^*,\ \exists b_0 \equiv b_0(V_0, V_1, 1_S) > 0: \quad PV_0 \leq V_0 - V_1 + b_0 1_S. \tag{$\boldsymbol{D}_{1_S}(V_0, V_1)$}$$

Note that Condition $\boldsymbol{D}_{1_S}(V_0, V_1)$ implies that $V_0 \geq V_1$ on S^c. In fact Condition $\boldsymbol{D}_{1_S}(V_0, V_1)$ is equivalent to : There exists $S \in \mathcal{X}^*$ such that $\sup_{x \in S^c} \Gamma(x) \leq 0$ and $\sup_{x \in S} \Gamma(x) < \infty$ with the measurable finite function $\Gamma(x) := (PV_0)(x) - V_0(x) + V_1(x)$. Thus, if Condition $\boldsymbol{D}_{1_S}(V_0, V_1)$ holds, then any constant $b_0(V_0, V_1, 1_S) \geq \sup_{x \in S} \Gamma(x)$ may be chosen. Finally recall that Conditions $(\boldsymbol{M}_{\nu, 1_S})$ and $\boldsymbol{D}_{1_S}(V_0, V_1)$ are the most classical minorization/drift assumptions in the literature.

Let us return to Markov kernel P satisfying the assumptions of Proposition 3.1. Then both Conditions $(\boldsymbol{M}_{\nu, 1_S})$ and $(\boldsymbol{M}_{\nu, \psi_S})$ hold with $\nu \in \mathcal{M}_{+,b}^*$ and $\psi_S \geq 1_S$ given in (15). Moreover, if

P satisfies $\boldsymbol{D}_{1_S}(V_0, V_1)$, then Condition $\boldsymbol{D}_{\psi_S}(V_0, V_1)$ holds since $\psi_S \geq 1_S$. The next statement ensures that the constant $b_0(V_0, V_1, \psi_S)$ may be chosen smaller than $b_0(V_0, V_1, 1_S)$.

Proposition 5.1 *Let P satisfy the assumptions of Proposition 3.1 and Condition $\boldsymbol{D}_{1_S}(V_0, V_1)$ for some couple (V_0, V_1) of Lyapunov functions on $\mathbb{X}$. Then P satisfies Condition $\boldsymbol{D}_{\psi_S}(V_0, V_1)$ with $\psi_S \geq 1_S$ given in (15), and we can choose*

$$b_0(V_0, V_1, \psi_S) \leq b_0(V_0, V_1, 1_S). \tag{55}$$

Proof. Since ψ_S defined in (15) is such that $\psi_S \geq 1_S$ we already quoted that P also satisfies Condition $\boldsymbol{D}_{\psi_S}(V_0, V_1)$. Next, set

$$b_0(V_0, V_1, \psi_S) := \sup_{x \in S} \frac{\Gamma(x)}{\psi_S(x)} \quad \text{with } \Gamma(x) := (PV_0)(x) - V_0(x) + V_1(x).$$

Since $\psi_S \geq 1_S$, we have $b_0(V_0, V_1, \psi_S) \leq \sup_{x \in S} \Gamma(x) \leq b_0(V_0, V_1, 1_S)$. □

Example 5.2 (Geometric drift condition) *Let us introduce the following so-called $V-$geometric drift condition (to be discussed in Section 6):*

$$\exists \psi \in \mathcal{B}_+^*,\ \exists \delta \in (0,1),\ \exists b \in (0, +\infty): \quad PV \leq \delta V + b\,\psi \tag{$\boldsymbol{G}_\psi(\delta, V)$}$$

where $V : \mathbb{X} \to [1, +\infty)$ is a measurable function. Again recall that the most classical case is when $\psi := 1_S$ for some $S \in \mathcal{X}^*$, that is

$$\exists S \in \mathcal{X}^*,\ \exists \delta \in (0,1),\ \exists b \in (0, +\infty): \quad PV \leq \delta V + b\,1_S. \tag{$\boldsymbol{G}_{1_S}(\delta, V)$}$$

Observe that $\boldsymbol{G}_\psi(\delta, V)$ implies that $PV \leq V - (1-\delta)V + b\,\psi$, so that P satisfies the V_1-modulated drift Condition $\boldsymbol{D}_\psi(V_0, V_1)$ with $V_0 := V/(1-\delta)$, $V_1 := V$ and $b_0 := b/(1-\delta)$.

5.2 Series of the residual kernel iterates

Under Conditions $(\boldsymbol{M}_{\nu,\psi})$–$\boldsymbol{D}_\psi(V_0, V_1)$ the following theorem provides relevant properties on the non-negative kernel $\sum_{k=0}^{+\infty} R^k$ involving the residual kernel R, from which further statements on P and π_R are obtained. Moreover the bounds (57a)-(57b) below are crucial for the study of Poisson's equation in the next subsection. The constant used under Conditions $(\boldsymbol{M}_{\nu,\psi})$-$\boldsymbol{D}_\psi(V_0, V_1)$ in Theorems 5.3 and 5.4 is the following one:

$$d_0 := \max\left(0, \frac{b_0 - \nu(V_0)}{\nu(1_{\mathbb{X}})}\right). \tag{56}$$

Theorem 5.3 *Let P satisfy Conditions $(\boldsymbol{M}_{\nu,\psi})$–$\boldsymbol{D}_\psi(V_0, V_1)$. Then the following non-negative function series and their integral w.r.t. the measure ν satisfy*

$$0 \;\leq\; \sum_{k=0}^{+\infty} R^k 1_{\mathbb{X}} \leq \sum_{k=0}^{+\infty} R^k V_1 \;\leq V_0 + d_0 1_{\mathbb{X}} \;\leq (1 + d_0) V_0 \tag{57a}$$

$$0 \;\leq\; \sum_{k=0}^{+\infty} \nu\big(R^k 1_{\mathbb{X}}\big) \leq \sum_{k=0}^{+\infty} \nu\big(R^k V_1\big) \;\leq \nu(V_0) + d_0 \nu(1_{\mathbb{X}}) \;\leq (1 + d_0)\, \nu(V_0) < \infty. \tag{57b}$$

Moreover the conclusions (i)-(v) *provided at the beginning of this section hold true, as well as the following additional assertions:*

(vi) *The unique $P-$invariant probability measure π_R is such that $\pi_R(V_1) < \infty$.*

(vii) *If $\pi_R(V_0) < \infty$, then $\pi_R(V_1) \le b_0\, \pi_R(\psi) \le b_0$ where b_0 is the constant in $\boldsymbol{D}_\psi(V_0, V_1)$.*

(viii) *if PV_1/V_1 is bounded on $\mathbb{X}$, i.e. $P\mathcal{B}_{V_1} \subset \mathcal{B}_{V_1}$, then the $P-$harmonic functions in $\mathcal{B}_{V_1}$ (i.e. $g \in \mathcal{B}_{V_1}$ such that $Pg = g$) are constant on $\mathbb{X}$.*

Inequalities (57a)-(57b), thus the non-negative constant d_0 in (56), will play a crucial role for the bounds of solutions to Poisson equation in Subsection 5.3 and for the rates of convergence in Section 8. Recall that the constant d_0 depends on the minorizing measure ν in $(\boldsymbol{M}_{\nu,\psi})$ and on the constant $b_0(V_0, V_1, \psi)$ in $\boldsymbol{D}_\psi(V_0, V_1)$.

Proof. From Condition $\boldsymbol{D}_\psi(V_0, V_1)$ we obtain that

$$RV_0 = PV_0 - \nu(V_0)\psi \le V_0 - V_1 + \big(b_0 - \nu(V_0)\big)\psi,$$

$$\text{equivalently} \quad V_1 \le V_0 - RV_0 + (b_0 - \nu(V_0))\psi,$$

from which we derive that

$$\begin{aligned} \forall n \ge 1, \quad 0 \ &\le\ \sum_{k=0}^{n} R^k V_1 \le \sum_{k=0}^{n} R^k V_0 - \sum_{k=1}^{n+1} R^k V_0 + \big(b_0 - \nu(V_0)\big) \sum_{k=0}^{n} R^k \psi \\ &\le\ V_0 + d_0 1_{\mathbb{X}} \end{aligned} \tag{58}$$

using the definition and the non-negativity of d_0 and finally (24). This provides Inequalities (57a) observing that $V \le (1 + d_0)V_0$, and (57b) is then obtained using the monotone convergence theorem. Next, the point-wise convergence of the first series in (57a) proves that $h_R^\infty := \lim_n R^n 1_{\mathbb{X}} = 0$ (see (20)), while the convergence of the first series in (57b) reads as $\mu_R(1_{\mathbb{X}}) = \sum_{k=0}^{+\infty} \nu(R^k 1_{\mathbb{X}}) < \infty$ (see (21)). Recall that the conclusions (i)-(v) provided at the beginning of this section then follows from Theorems 4.1, 4.3, 4.7 and 4.12. Now prove the additional assertions (vi)-(viii). That $\pi_R(V_1) < \infty$ follows from the definition of π_R and from the second inequality in (57b) which provides $\mu_R(V_1) < \infty$. To prove (vii), note that

$$\pi_R(PV_0) = \pi_R(V_0) \le \pi_R(V_0) - \pi_R(V_1) + b_0 \pi_R(\psi)$$

from the $P-$invariance of π_R and $\boldsymbol{D}_\psi(V_0, V_1)$. Finally the proof of (viii) follows the same lines as for Assertion 1. of Theorem 4.1, replacing the function $1_{\mathbb{X}}$ with V_1 and observing that $P(\mathcal{B}_{V_1}) \subset \mathcal{B}_{V_1}$, thus $R(\mathcal{B}_{V_1}) \subset \mathcal{B}_{V_1}$, when PV_1/V_1 is bounded on $\mathbb{X}$. Indeed, first recall that $\widetilde{\psi} := \sum_{k=0}^{+\infty} R^k \psi = \nu(1_{\mathbb{X}})^{-1} 1_{\mathbb{X}}$ from (35) since $h_R^\infty = 0$. Now let $g \in \mathcal{B}_{V_1}$ be such that $Pg = g$. Using $R(\mathcal{B}_{V_1}) \subset \mathcal{B}_{V_1}$ and proceeding as in Lemma 3.3, we obtained that $\nu(g) \sum_{k=0}^{n} R^k \psi = g - R^{n+1} g$ for every $n \ge 1$. Moreover we have $\lim_n R^n g = 0$ since $|R^n g| \le R^n |g| \le \|g\|_{V_1} R^n V_1$ and $\lim_n R^n V_1 = 0$ from (57a). Thus $g = \nu(g)\widetilde{\psi}$, from which it follows that g is constant. □

5.3 Poisson's equation

When P satisfies Conditions $(\boldsymbol{M}_{\nu,\psi})$ and $\boldsymbol{D}_\psi(V_0, V_1)$, recall that π_R given in (26) is the unique $P-$invariant probability measure on $(\mathbb{X}, \mathcal{X})$.

Theorem 5.4 *Let P satisfy Conditions $(\boldsymbol{M}_{\nu,\psi})$–$\boldsymbol{D}_\psi(V_0, V_1)$ and let $R \equiv R_{\nu,\psi}$ be the associated residual kernel given in* (13)*. Then the following assertions hold.*

1. *For any $g \in \mathcal{B}_{V_1}$, the function series $\widetilde{g} := \sum_{k=0}^{+\infty} R^k g$ absolutely point-wise converges on $\mathbb{X}$. Moreover we have $\widetilde{g} \in \mathcal{B}_{V_0}$, more precisely we have*

$$|\widetilde{g}| \leq \|g\|_{V_1}\big(V_0 + d_0 1_{\mathbb{X}}\big) \tag{59a}$$
$$\text{and} \quad \|\widetilde{g}\|_{V_0} \leq (1+d_0)\|g\|_{V_1}. \tag{59b}$$

where the non-negative constant d_0 is defined in (56).

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

$$(I-P)\widetilde{g} = g. \tag{60}$$

Let $h \in \mathcal{B}_{V_1}$ and let $h_0 := h - \pi_R(h)1_{\mathbb{X}}$ be the associated π_R-centred function. Then the function $\widetilde{h}_0 = \sum_{k=0}^{+\infty} R^k h_0$ belongs to $\mathcal{B}_{V_0}$ and satisfies Poisson's equation $(I-P)\widetilde{h}_0 = h_0$. Moreover the following bounds hold

$$|\widetilde{h}_0| \leq \|h\|_{V_1}\big(1+\pi_R(V_1)\big)\big(V_0 + d_0 1_{\mathbb{X}}\big) \leq \|h\|_{V_1}\big(1+\pi_R(V_1)\big)\big(1+d_0\big)V_0.$$

This follows from Theorem 5.4 applied to $g := h_0$, observing that we have from the triangular inequality

$$\|h_0\|_{V_1} \leq \|h\|_{V_1}\big(1+\pi_R(V_1)\big).$$

Proof. Let $g \in \mathcal{B}_{V_1}$. Using $|g| \leq \|g\|_{V_1} V_1$ and $|R^k g| \leq R^k|g| \leq \|g\|_{V_1} R^k V_1$ for $k \geq 1$, Assertion 1. follows from (57a). Next, note that $\pi_R(|g|) < \infty$ since $\pi_R(V_1) < \infty$ from Assertion (vi) of Theorem 5.3. Now define

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

Then, using $P = R + \psi \otimes \nu$ we have

$$\widetilde{g}_n - P\widetilde{g}_n = \widetilde{g}_n - R\widetilde{g}_n - \nu(\widetilde{g}_n)\psi = g - R^{n+1}g - \nu(\widetilde{g}_n)\psi. \tag{61}$$

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 $\nu(\widetilde{g}_n) = \sum_{k=0}^{n} \nu(R^k g)$ and $\mu_R(V_1) < \infty$, we obtain that $\lim_{n\to+\infty} \nu(\widetilde{g}_n) = \mu_R(g)$ from Lebesgue's theorem w.r.t. the measure ν. 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 $\lim_n \widetilde{g}_n = \widetilde{g}$, $|\widetilde{g}_n| \leq \|g\|_{V_1} V_0$ (from Assertion 1.) and $(PV_0)(x) < \infty$. Taking the limit when n goes to infinity in (61), we get that

$$(I-P)\widetilde{g} = g - \mu_R(g)\psi. \tag{62}$$

Next, if we assume that $\pi_R(g) = 0$, then Equality (62) rewrites as $(I-P)\widetilde{g} = g$ since $\mu_R(g) = \pi_R(g)/\pi_R(\psi) = 0$ from (26). Theorem 5.4 is proved. □

For $g \in \mathcal{B}_{V_1}$ such that $\pi_R(g) = 0$, the solution $\widetilde{g} := \sum_{k=0}^{+\infty} R^k g$ in $\mathcal{B}_{V_0}$ to Poisson's equation $(I-P)\widetilde{g} = g$ in Theorem 5.4 is not π_R-centred a priori, i.e. $\pi_R(\widetilde{g}) \neq 0$. The natural way to get a π_R-centred solution is to define $\widehat{g} = \widetilde{g} - \pi_R(\widetilde{g})1_{\mathbb{X}}$, but we then need to assume that $\widetilde{g}$ is π_R-integrable. Accordingly, to obtain such a π_R-centred solution to Poisson's equation in general terms, the assumption $\pi_R(V_0) < \infty$ must be made.

Corollary 5.5 *Let P satisfy Conditions $(\boldsymbol{M}_{\nu,\psi})$–$\boldsymbol{D}_\psi(V_0,V_1)$ with $\pi_R(V_0)<\infty$. For any $g\in\mathcal{B}_{V_1}$ such that $\pi_R(g)=0$, set $\widetilde{g}:=\sum_{k=0}^{+\infty}R^k g$. Then the function $\widehat{g}=\widetilde{g}-\pi_R(\widetilde{g})1_{\mathbb{X}}$ is a π_R-centered solution on $\mathcal{B}_{V_0}$ to Poisson's equation $(I-P)\widehat{g}=g$. Moreover we have*

$$\|\widehat{g}\|_{V_0}\leq(1+d_0)\big(1+\pi_R(V_0)\big)\,\|g\|_{V_1} \tag{63}$$

where the non-negative constant d_0 is given in (56).

Proof. Let $g\in\mathcal{B}_{V_1}$ be such that $\pi_R(g)=0$. Obviously we have $\widehat{g}\in\mathcal{B}_{V_0}$ and $\pi_R(\widehat{g})=0$. Moreover we obtain that $(I-P)\widehat{g}=(I-P)\widetilde{g}=g$ from Theorem 5.4 and $(I-P)1_{\mathbb{X}}=0$. Finally we have

$$\|\widehat{g}\|_{V_0}\leq\big(1+\pi_R(V_0)\,\|1_{\mathbb{X}}\|_{V_0}\big)\|\widetilde{g}\|_{V_0}\leq(1+d_0)\big(1+\pi_R(V_0)\big)\|g\|_{V_1} \tag{64}$$

using the definition of $\widehat{g}$, the triangular inequality and $|\widetilde{g}|\leq\|\widetilde{g}\|_{V_0}V_0$ for the first inequality, and the bound (59b) applied to $\widetilde{g}$ for the second one. □

Let $g\in\mathcal{B}_{V_1}$ be such that $\pi_R(g)=0$. Under the assumptions of Corollary 5.5, when a π_R-centred solution $\mathfrak{g}\in\mathcal{B}_{V_0}$ to Poisson's equation $(I-P)\mathfrak{g}=g$ is known, and when two solutions to Poisson's equation in $\mathcal{B}_{V_0}$ differ from an additive constant, then we have $\mathfrak{g}=\widehat{g}$, so that the bound (63) applies to $\mathfrak{g}$. Of course such a solution $\mathfrak{g}$ may be obtained independently of the function $\widetilde{g}$. For instance it can be given by $\mathfrak{g}=\sum_{k=0}^{+\infty}P^k g$ provided that this series point-wise converges and defines a function of $\mathcal{B}_{V_0}$. Note that the choice of the minorizing measure ν and of the function ψ used in Conditions $(\boldsymbol{M}_{\nu,\psi})$ and $\boldsymbol{D}_\psi(V_0,V_1)$ of Corollary 5.5 naturally has an impact on the constant d_0 in (63).

Remark 5.6 *Recall that, under Conditions $(\boldsymbol{M}_{\nu,\psi})$–$\boldsymbol{D}_\psi(V_0,V_1)$, the function $h_R^\infty:=\lim_n R^n 1_{\mathbb{X}}$ (see (20)) is zero from the convergence of the first series in (57a), so that $\widetilde{\psi}:=\sum_{k=0}^{+\infty}R^k\psi=\nu(1_{\mathbb{X}})^{-1}1_{\mathbb{X}}$ from* (35). *So the presence of the term $\nu(1_{\mathbb{X}})^{-1}$ in the general bound* (59b) *is quite natural (it is not due to the proof of Theorem 5.4). This does not mean that the bound of the V_0- norm of solutions to Poisson's equation could not be improved. But in fact this last question is not well formulated since solutions to Poisson's equation are not unique, and the solutions given in Theorem 5.4 are very specific: they are defined from the residual kernel R, in particular they are not π_R-centred (see Corollary 5.5).*

Remark 5.7 *Assume that P satisfies Conditions $(\boldsymbol{M}_{\nu,1_S})$–$\boldsymbol{D}_{1_S}(V_0,V_1)$ with $V_0\geq V_1$ and* $\inf V_0=1$. *Then we have $d_0=0$ in the bound* (59b) *of Theorem 5.4 if, and only if, S is an atom, i.e. $\forall a\in S,\ \nu(dy)=P(a,dy)$. Indeed, if S is an atom, then P satisfies $\boldsymbol{D}_{1_S}(V_0,V_1)$ with $b_0=\nu(V_0)$ since $V_0\geq V_1$. Thus $d_0=0$. To prove the converse implication, note that*

$$\nu(1_{\mathbb{X}})^{-1}=\nu(1_{\mathbb{X}})^{-1}\|1_{\mathbb{X}}\|_{V_0}\leq(1+d_0)\|1_S\|_{V_1}\leq(1+d_0)$$

from (59b) *applied to $g:=1_S$ and* (35) *with here $\psi:=1_S$. Hence, if $d_0=0$, then $\nu(1_{\mathbb{X}})\geq 1$. Thus S is an atom since, for every $a\in S$, the non-negative measure $\eta_a(dy)=P(a,dy)-\nu(dy)$ satisfies $\eta_a(1_{\mathbb{X}})\leq 0$, so that $\eta_a=0$.*

5.4 Further statements

Under Conditions $(\boldsymbol{M}_{\nu,\psi})$ and for any couple (V,W) of measurable functions from $\mathbb{X}$ to $[1,+\infty)$ such that $\nu(V)<\infty$, let us introduce the following residual-type modulated drift condition involving the residual kernel $R \equiv R_{\nu,\psi}$ given in (13):

$$RV \le V - W. \qquad (\boldsymbol{R}_{\nu,\psi}(V,W))$$

Note that Condition $\boldsymbol{R}_{\nu,\psi}(V,W)$ rewrites as $PV \le V - W + \nu(V)\psi$, which is a specific instance of Condition $\boldsymbol{D}_\psi(V,W)$ with $b_0 = \nu(V)$. The next simple lemma shows that $\boldsymbol{D}_\psi(V_0,V_1)$ generates a residual-type modulated drift condition up to slightly modify V_0. This has been already observed under the weaker drift condition (48) in the proof of Proposition 4.19.

Lemma 5.8 *If P satisfies Conditions $(\boldsymbol{M}_{\nu,\psi})$ and $\boldsymbol{D}_\psi(V_0,V_1)$, then we have $\nu(V_0) < \infty$, and the residual kernel $R \equiv R_{\nu,\psi}$ given in (13) satisfies Condition $\boldsymbol{R}_{\nu,\psi}(V_{0,d_0},V_1)$ with $V_{0,d_0} := V_0 + d_0 1_{\mathbb{X}} \ge V_0$ where $d_0 := \max\big(0, (b_0 - \nu(V_0))/\nu(1_{\mathbb{X}})\big)$.*

Note that d_0 in Lemma 5.8 and also in Lemma 5.9 below is the non-negative constant already given in (56).

Proof. We already quoted that PV_0 is everywhere finite under Condition $\boldsymbol{D}_\psi(V_0,V_1)$, so that $0 \le \nu(V_0)\psi(x) \le (PV_0)(x)$ for every $x \in \mathbb{X}$ from $(\boldsymbol{M}_{\nu,\psi})$. Then it follows that the function RV_0 is well-defined and is everywhere finite. Note that $\nu(V_{0,d_0}) = \nu(V_0) + d_0\nu(1_{\mathbb{X}}) < \infty$ and that $PV_{0,d_0} = PV_0 + d_0 1_{\mathbb{X}}$. We get from the definitions of R and V_{0,d_0}

$$\begin{aligned} RV_{0,d_0} &= PV_{0,d_0} - \nu(V_{0,d_0})\psi = PV_0 + d_0 1_{\mathbb{X}} - \big(\nu(V_0) + d_0\,\nu(1_{\mathbb{X}})\big)\psi \\ &\le V_0 - V_1 + b_0\psi + d_0 1_{\mathbb{X}} - \big(\nu(V_0) + d_0\,\nu(1_{\mathbb{X}})\big)\psi \quad \text{(from Assumption } \boldsymbol{D}_\psi(V_0,V_1)) \\ &\quad = V_{0,d_0} - V_1 + \big(b_0 - \nu(V_0) - d_0\,\nu(1_{\mathbb{X}})\big)\psi \\ &\le V_{0,d} - V_1 \quad \text{(from the definitions of } d_0). \end{aligned}$$

Hence the proof is complete. □

Under Conditions $(\boldsymbol{M}_{\nu,\psi})$-$\boldsymbol{D}_\psi(V_0,V_1)$ and the additional condition $\pi_R(V_0) < \infty$, the sequence $(P^nV_0)_{n\ge1}$ is shown to be bounded in $(\mathcal{B}_{V_0}, \|\cdot\|_{V_0})$ in the following lemma.

Lemma 5.9 *Let P satisfy Conditions $(\boldsymbol{M}_{\nu,\psi})$–$\boldsymbol{D}_\psi(V_0,V_1)$ with $\pi_R(V_0) < \infty$. Then we have for every $n \ge 1$:*

$$P^nV_0 \le V_0 + \frac{\|\psi\|_{1_{\mathbb{X}}}\big(\pi_R(V_0) + d_0\big)}{\pi_R(\psi)} 1_{\mathbb{X}} \quad \textit{with} \quad \|\psi\|_{1_{\mathbb{X}}} := \sup_{x\in\mathbb{X}} \psi(x),\ d_0 := \max\left(0, \frac{b_0 - \nu(V_0)}{\nu(1_{\mathbb{X}})}\right).$$

Proof. It follows from Lemma 5.8 that $RV_{0,d_0} \le V_{0,d_0}$ with $V_{0,d_0} := V_0 + d_0 1_{\mathbb{X}}$ and $R \equiv R_{\nu,\psi}$ in (13). Using the non-negativity of R and iterating this inequality gives: $\forall n \ge 1,\ R^nV_{0,d_0} \le V_{0,d_0}$. From Formula (17) and $0 \le P^k\psi \le \|\psi\|_{1_{\mathbb{X}}} 1_{\mathbb{X}}$, we obtain that

$$\forall n \ge 1, \quad P^nV_{0,d_0} = R^nV_{0,d_0} + \sum_{k=1}^{n} \nu(R^{k-1}V_{0,d_0})\, P^{n-k}\psi \le V_{0,d_0} + \|\psi\|_{1_{\mathbb{X}}}\, \mu_R(V_{0,d_0}) 1_{\mathbb{X}}$$

with $\mu_R = \pi_R/\pi_R(\psi)$ given in (26). This provides the desired inequality using the definition of $V_{0,d_0}, P1_{\mathbb{X}} = 1_{\mathbb{X}}$ and $\pi_R(V_0) < \infty$. □

Now, given any measurable function $V_1 : \mathbb{X} \to [1, +\infty)$, we present a necessary and sufficient condition for P to satisfy a V_1−modulated drift condition.

Proposition 5.10 *Assume that P satisfies Condition $(\boldsymbol{M}_{\nu,\psi})$. Let $V_1 : \mathbb{X} \to [1,+\infty)$ be any measurable function. Then there exists a measurable function $V_0 : \mathbb{X} \to [1,+\infty)$ such that P satisfies $\boldsymbol{D}_{\psi}(V_0, V_1)$ if and only if*

$$\forall x \in \mathbb{X}, \quad \widetilde{V_1}(x) := \sum_{k=0}^{+\infty} (R^k V_1)(x) < \infty \quad \textit{and} \quad \nu(\widetilde{V_1}) < \infty \tag{65}$$

where $R \equiv R_{\nu,\psi}$ is the residual kernel in (13).

Proof. If P satisfies Condition $\boldsymbol{D}_{\psi}(V_0, V_1)$ for some Lyapunov function V_0, then (65) holds true from Theorem 5.3 (in fact we know that $\widetilde{V_1} \leq c\, V_0$ for some positive constant c). Conversely, if V_1 satisfies (65) with $R \equiv R_{\nu,\psi}$ in (13), then we have $(R\widetilde{V_1})(x) = \widetilde{V_1}(x) - V_1(x)$ for every $x \in \mathbb{X}$ from the monotone convergence theorem w.r.t. the measure $R(x, dy)$. Hence Condition $\boldsymbol{R}_{\nu,\psi}(\widetilde{V_1}, V_1)$ holds. Then Condition $\boldsymbol{D}_{\psi}(\widetilde{V_1}, V_1)$ holds with $b_0 := \nu(\widetilde{V_1})$. □

The next statement completes Theorem 3.6.

Proposition 5.11 *Assume that P satisfies Condition $(\boldsymbol{M}_{\nu,\psi})$ and is irreducible. Then the three equivalent conditions 1., 2. and 3. of Theorem 3.6 are also equivalent to the following one: There exists a $P-$absorbing and μ_R-full set $A \in \mathcal{X}$ such that the restriction of P to A satisfies the modulated drift condition $\boldsymbol{D}_{\psi_A}(V_A, 1_A)$ for some measurable function $V_A : A \to [1,+\infty)$, where ψ_A is the restriction of ψ to A.*

If P satisfies the minorization condition $(\boldsymbol{M}_{\nu,\psi})$, is irreducible and admits an invariant probability measure η, then we have $\eta = \pi_R$ from Theorem 3.14, and all the conclusions of Theorem 5.3 then hold on some $P-$absorbing and π_R-full set thanks to Proposition 5.11.

Proof. Under Condition $(\boldsymbol{M}_{\nu,\psi})$, let $R \equiv R_{\nu,R}$ be the residual kernel defined in (13). Assume that Condition 2. of Theorem 3.6 holds, i.e. $\mu_R(1_{\mathbb{X}}) < \infty$. Define on $\mathbb{X}$ the function $V := \sum_{k=0}^{+\infty} R^k 1_{\mathbb{X}}$ taking its value in $[0, +\infty]$ a priori. Since $\nu(V) = \mu_R(1_{\mathbb{X}}) < \infty$, the set

$$A := \big\{x \in \mathbb{X} : V(x) < \infty\big\}$$

is non-empty. Moreover, if $x \in A$, then we have $(RV)(x) < \infty$ since $(RV)(x) = V(x) - 1$ from the monotone convergence theorem w.r.t. the measure $R(x, dy)$. We then obtain that $(PV)(x) = (RV)(x) + \nu(V)\psi(x) = V(x) - 1 + \nu(V)\psi(x) < \infty$. This proves that A is $P-$absorbing. Since P is irreducible, A is μ_R-full from Proposition 3.15. Furthermore, the previous equality proves that the restriction of P to A satisfies the modulated drift condition $\boldsymbol{D}_{\psi_A}(V_A, 1_A)$ where V_A is the restriction of V to the set A.

Conversely assume that the condition provided in Proposition 5.11 holds. Using the fact that A is $P-$absorbing and proceeding as in the proof of Corollary 4.5, it can be proved that the restriction P_A of P to A satisfies on A the minorization condition $(\boldsymbol{M}_{\nu_A,\psi_A})$ with small-function ψ_A and minorizing measure ν_A defined as the restriction of ν to A. Then it follows from Theorem 5.3 applied to the Markov kernel P_A that there exists a unique P_A-invariant probability measure η_A on A and that $\eta_A(\psi_A) > 0$ (apply Assertion (iv) to P_A). Next let us define the following positive measure on $(\mathbb{X}, \mathcal{X})$: $\forall B \in \mathcal{X},\ \eta(1_B) := \eta_A(1_{A\cap B})$. Since A is $P-$absorbing, η is a P-invariant probability measure, and we have $\eta(\psi) = \eta_A(\psi_A) > 0$. Consequently Condition 1. of Theorem 3.6 holds for P and Proposition 5.11 is proved. □

Under Conditions $(\boldsymbol{M}_{\nu,\psi})$–$\boldsymbol{D}_\psi(V_0, V_1)$, the next statement provides a necessary and sufficient condition for the (unique) $P-$invariant probability measure π_R given in (26) to satisfy $\pi_R(V_0) < \infty$.

Proposition 5.12 *Let P satisfy Conditions $(\boldsymbol{M}_{\nu,\psi})$–$\boldsymbol{D}_\psi(V_0, V_1)$. Then the two following conditions are equivalent:*

1. $\pi_R(V_0) < \infty$.
2. *There exists a $P-$absorbing and π_R-full set $A \in \mathcal{X}$ and a measurable function $L \geq V_0$ on A such that the restriction P_A of P to A satisfies the modulated drift condition $\boldsymbol{D}_{\psi_A}(L, V_{0_{|A}})$, where $V_{0_{|A}}$ (resp. ψ_A) is the restriction of V_0 (resp. of ψ) to A.*

Proof. The proof follows the same limes as for Proposition 5.11. Let $R \equiv R_{\nu,R}$ be the residual kernel given in (13). Assume that $\pi_R(V_0) < \infty$ and define on $\mathbb{X}$ the $[0, +\infty]-$valued function $\widetilde{V_0} := \sum_{k=0}^{+\infty} R^k V_0$. Then $\widetilde{V_0} \geq V_0$, and the following equality holds in $[0, +\infty]$: $R\widetilde{V_0} = \widetilde{V_0} - V_0$. Note that there exists $x \in \mathbb{X}$ such that $\widetilde{V_0}(x) < \infty$ since $\nu(\widetilde{V_0}) = \mu_R(V_0) < \infty$ from $\pi_R(V_0) < \infty$, where $\mu_R := \sum_{k=0}^{+\infty} \nu R^k$ (see (26)). Now define the non-empty set $A := \{x \in \mathbb{X} : \widetilde{V_0}(x) < \infty\} \in \mathcal{X}$. Let $x \in A$. Then we have $(R\widetilde{V_0})(x) < \infty$ from $(R\widetilde{V_0})(x) = \widetilde{V_0}(x) - V_0(x)$, so that $(P\widetilde{V_0})(x) = (R\widetilde{V_0})(x) + \nu(\widetilde{V_0})\psi(x) < \infty$. Thus $P(x, A) = 1$. This proves that A is $P-$absorbing. Since P is irreducible from Theorem 5.3, A is π_R-full from Proposition 3.15. Moreover the restriction $L := \widetilde{V_0}_{|A}$ of $\widetilde{V_0}$ to A is a measurable function on A satisfying $RL = L - V_0$ on A, so that the restriction P_A of P to A satisfies the modulated drift condition $\boldsymbol{D}_{\psi_A}(L, V_{0_{|A}})$ as stated in Assertion 2. of Proposition 5.12.

Conversely assume that P satisfies Assertion 2. Then, proceeding as in the proof of Corollary 4.5, we know that P_A satisfies on A the minorization condition $(\boldsymbol{M}_{\nu_A,\psi_A})$ where ν_A is the restriction of the minorizing measure ν to A. Thus it follows from Assertion (vi) of Theorem 5.3 applied to P_A under Condition $(\boldsymbol{M}_{\nu_A,\psi_A})$ and $\boldsymbol{D}_{\psi_A}(L, V_{0_{|A}})$ that the unique P_A-invariant probability measure, say π_A, is such that $\pi_A(V_{0_{|A}}) < \infty$. Using the fact that π_R is the unique $P-$ invariant probability measure, we then obtained that π_A is the restriction of π_R to A and that $\pi_R(V_0) = \pi_A(V_{0_{|A}}) < \infty$ since A is $P-$absorbing and π_R-full. □

We conclude this subsection proving that the bound $\pi_R(V_1) \leq b_0\, \pi_R(\psi)$ in Assertion (vii) of Theorem 5.3 holds even when V_0 is not π_R-integrable.

Proposition 5.13 *Let P satisfy Conditions $(\boldsymbol{M}_{\nu,\psi})$–$\boldsymbol{D}_\psi(V_0, V_1)$. Then $\pi_R(V_1) \leq b_0\, \pi_R(\psi)$.*

Proof. Let $W_1 := V_0 - PV_0 + b_0\psi$. Note that $W_1 \geq V_1$. Since P obviously satisfies Condition $\boldsymbol{D}_\psi(V_0, W_1)$ we know from Assertion (vi) of Theorem 5.3 that $\pi_R(W_1) < \infty$. Thus the function $V_0 - PV_0$ is π_R-integrable. Since $V_0 - RV_0 = V_0 - PV_0 + \nu(V_0)\psi$, we obtain that $V_0 - RV_0$ is π_R-integrable too. Moreover we know from Lemma 5.8 that $RV_{0,d_0} \leq V_{0,d_0}$ with $V_{0,d_0} := V_0 + d_0 1_{\mathbb{X}}$ and $d_0 := \max(0, (b_0 - \nu(V_0))/\nu(1_{\mathbb{X}}))$. Iterating this inequality provides: $\forall k \geq 1,\ R^k V_{0,d_0} \leq V_{0,d_0}$, thus $R^k V_0 \leq V_0 + d_0 1_{\mathbb{X}}$. This shows that $\nu(R^k V_0) < \infty$ for every $k \geq 0$. Next note that $V_0 - PV_0$ and $V_0 - RV_0$ are also μ_R-integrable since $\mu_R := \mu_R(1_{\mathbb{X}})\pi_R$ (see (26)). Therefore it follows from the definition of μ_R that the series $\sum_{k=0}^{+\infty} \nu(R^k(V_0 - RV_0))$

converges, so that

$$\begin{aligned}\mu_R(V_0 - RV_0) := \lim_{n\to+\infty}\sum_{k=0}^{n}\nu\big(R^k(V_0 - RV_0)\big) &= \lim_{n\to+\infty}\sum_{k=0}^{n}\big[\nu\big(R^kV_0\big) - \nu(R^{k+1}V_0)\big]\\ &= \nu(V_0) - \lim_{n\to+\infty}\nu\big(R^{n+1}V_0\big)\end{aligned}$$

from which we deduce that

$$\mu_R(V_0 - PV_0) = \mu_R(V_0 - RV_0) - \nu(V_0)\mu_R(\psi) = -\lim_{n\to+\infty}\nu\big(R^nV_0\big) \le 0$$

since $\mu_R(\psi) = 1$ and $\nu(R^nV_0) \ge 0$. Finally we obtain from the definition of W_1 that

$$\pi_R(V_1) \le \pi_R(W_1) = \pi_R(V_0 - PV_0) + b_0\pi_R(\psi) \le b_0\pi_R(\psi)$$

since $\pi_R(V_0 - PV_0) = \mu_R(1_{\mathbb{X}})^{-1}\mu_R(V_0 - PV_0) \le 0$. □

5.5 Further comments and bibliographic discussion

Condition $\boldsymbol{D}_{1_S}(V_0, V_1)$, extended here to $\boldsymbol{D}_\psi(V_0, V_1)$, is the so-called V_1-modulated drift condition, e.g. see Condition (V3) in [MT09, p. 343]. Although the functions V_0, V_1 in $\boldsymbol{D}_\psi(V_0, V_1)$ satisfy $V_0 \ge V_1$ in general, this condition is not useful in this section. Such drift conditions was first introduced for infinite stochastic matrices in [Fos53] to study the return times to a set. Still in discrete case, further developments based on drift conditions were proposed in [Ken51, Mau57, Pak69] with the aim of studying ergodicity or recurrence properties. The use of modulated drift conditions was extended to general state space in [Twe75]. An historical background on this subject is provided in [MT09, p. 198] and [DMPS18, p. 96, 164, 337]. The main results of this section have been presented in [HL25a]: Here the proof of Theorem 5.3 is slightly simplified using (24), moreover an intermediate bound is added in (59a) to facilitate bibliographic comments of Section 11. Again note that the non-negativity of the residual kernel R plays a crucial role in Theorem 5.3 since the point-wise convergence of the series in (57a) is simply obtained bounding the partial sums (see (58)).

Under the V_1-modulated drift condition $\boldsymbol{D}_{1_S}(V_0, V_1)$ w.r.t. some petite set $S \in \mathcal{X}$, the existence of a solution $\xi \in \mathcal{B}_{V_0}$ to Poisson's equation $(I - P)\xi = g$ was proved in [GM96, Th. 2.3] for every π_R-centred function $g \in \mathcal{B}_{V_1}$, together with the bound $\|\xi\|_{V_0} \le c_0\,\|g\|_{V_1}$ for some positive constant c_0 (independent of g). When S is an atom, the solution ξ in [GM96, Th. 2.3] can be expressed in terms of the first hitting time in S, and the non-atomic case is solved in [GM96] via the splitting method. Under the irreducibility and aperiodicity conditions, Glynn-Meyn's theorem is related to point-wise convergence of the series $\sum_{k=0}^{+\infty} P^kg$, see [MT09, Th. 14.0.1]. With regard to the above two representations of solutions to Poisson's equation, the reader may consult the recent article [GI24]. To the best of our knowledge the constant c_0 in [GM96, Th. 2.3] was unknown. In [Num91] the link between solutions to Poisson's equation $(I - P)\xi = g$ and the residual potential series $\sum_{k=0}^{+\infty} R^kg$ was highlighted in an abstract framework involving harmonic functions for both R and P. No modulated drift condition is used in [Num91], so that the convergence of the previous series must be assumed to hold and no bound for this series is provided there. The study of Poisson's equation via taboo potential theory has been developed in an even more abstract context in [Nev72, Rev84]. In our work this potential approach is restricted to the residual kernel under the first-order minorization condition $(\boldsymbol{M}_{\nu,\psi})$ and the modulated condition $\boldsymbol{D}_\psi(V_0, V_1)$.

The works [HL25a] and more recently [GLL25] seem to be the first ones providing an explicit bound in Poisson's equation, which was previously known only in the atomic case: see [LL18, Prop. 1] for a discrete state-space $\mathbb{X}$. Thus, the novelty of Theorem 5.4 and Corollary 5.5 already proved in [HL25a] is to provide a simple and explicit bound in Poisson's equation in the non-atomic case. Under the first-order minorization condition, the constants in the analogue of Bound (59a) in [GLL25] are very close to ours, and even possibly slightly tighter in particular cases, see [GLL25, Sec. 4] where the bounds are compared for the random walks on the half line. The paper [GLL25] based on a randomized stopping time also provides an explicit and simple bound of solutions to Poisson's equation under higher-order minorization condition (in place of Condition $(\boldsymbol{M}_{\nu,\psi})$). Such an extension is proved in Subsection 11.2 (see Theorem 11.12) using again the series of the iterates of a residual kernel. However the bound obtained in Theorem 11.12 is not as simple and accurate as in [GLL25], see Subsection 11.3. Still with the aim of studying Poisson's equation, an alternative approach based on the use of several first-order small-functions is proposed in Subsection 11.1.

Let us briefly discuss the Central Limit Theorem (C.L.T.), which is a standard topic where Poisson's equation is useful. If $(X_n)_{n\in\mathbb{N}}$ is a Markov chain with state space $\mathbb{X}$ and invariant distribution π, then a measurable $\pi-$centred real-valued function g on $\mathbb{X}$ is said to satisfy the C.L.T. under $\mathbb{P}_\eta$ for some initial probability measure η (i.e. η is the probability distribution of X_0) when the asymptotic distribution of $n^{-1/2}S_n(g)$ with $S_n(g) = \sum_{k=0}^{n-1} g(X_k)$ is the Gaussian distribution $\mathcal{N}(0, \sigma_g^2)$ for some positive constant σ_g^2, called the asymptotic variance of g. We refer to [Jon04] for a general overview on the Markovian C.L.T. and its relationship with drift and mixing conditions. In our context, [DMPS18, Chap. 21] is a nice and comprehensive account on the C.L.T. and the classical approach via Poisson's equation. Here, in link with Corollary 5.5, we just recall the following classical C.L.T. proved in [GM96] for Markov chains satisfying a modulated drift condition:

> **Glynn-Meyn's C.L.T.** [GM96]: *If the transition kernel P of the Markov chain $(X_n)_{n\in\mathbb{N}}$ satisfies Conditions $(\boldsymbol{M}_{\nu,\psi})$–$\boldsymbol{D}_\psi(V_0, V_1)$ with $V_1 \le V_0$, $\pi_R(V_0^2) < \infty$, and if η is any initial probability measure, then every π_R-centred function $g \in \mathcal{B}_{V_1}$ satisfies the C.L.T. under $\mathbb{P}_\eta$ with asymptotic variance given by $\sigma_g^2 = 2\pi_R(g\widehat{g}) - \pi_R(g^2)$, where $\widehat{g} \in \mathcal{B}_{V_0}$ is the solution to Poisson's equation $(I - P)\widehat{g} = g$ provided by Corollary 5.5.*

The condition $\pi_R(V_0^2) < \infty$ is required for the function $\widehat{g}$ to be square π_R-integrable in order to apply the Markovian C.L.T. [DMPS18, Th. 21.2.5] under $\mathbb{P}_{\pi_R}$, where π_R is the unique $P-$invariant probability measure from Theorem 5.3. The extension to any initial probability measure follows from [DMPS18, Cor. 21.1.6] since P is Harris recurrent under the assumptions of Corollary 5.5 from Theorem 5.3. A fairly comprehensive overview on the asymptotic variance in Markov chain C.L.T. is provided in [HR07]. Moreover recall that the finiteness of the asymptotic variance may be relevant beyond the C.L.T., e.g. see [Atc16]. Note that the asymptotic variance σ_g^2 can be upper bounded using the bound (63) (see [HL25a, Cor. 2.7]).

To conclude this section let us make a few additional comments on the modulated drift condition, which is the main assumption of this work together with the minorization condition. If $(X_n)_{n\ge 0}$ is a Markov chain with state space $\mathbb{X}$ and transition kernel P, then the modulated drift condition has the following form when the modulated function V_1 is constant and $\psi = 1_{\mathcal{V}_s}$

for some $s > 0$ where $\mathcal{V}_s = \{x \in \mathbb{X} : V_0(x) \leq s\}$ is the level set of order s w.r.t. the function V_0:

$$\sup_{x\in\mathcal{V}_s} \mathbb{E}_x\big[V_0(X_1)\big] < \infty \quad \text{and} \quad \exists a > 0,\ \forall x \in \mathbb{X} \setminus \mathcal{V}_s, \quad \mathbb{E}_x\big[V_0(X_1)\big] \leq V_0(x) - a. \tag{66}$$

The second condition in (66) means that, for any $r > s$, each point $x \in \mathbb{X}$ such that $V_0(x) = r$ transits in mean to a point of the level set $\mathcal{V}_{r-a}$. For a random walk on $\mathbb{N}$, it means that, for i large enough, the steps of the walker starting from i are in mean strictly more to the left than to the right, the gap being controlled by a fixed additive constant $a > 0$. Recall that the weaker drift condition (54) was introduced in Proposition 4.19 to obtain $\lim_k R^k 1_{\mathbb{X}} = 0$. The additive reduction by the positive constant a in (66) is the sole difference with (54), but it is crucial for obtaining the convergence of the series $\sum_{k=0}^{+\infty} R^k 1_{\mathbb{X}}$ in Theorem 5.3. The general modulated drift condition $\boldsymbol{D}_\psi(V_0, V_1)$ corresponds to (66) with a positive term $V_1(x)$ depending on x instead of the positive constant a.

Under the minorization condition $(\boldsymbol{M}_{\nu,\psi})$, Theorem 3.14 and Proposition 5.11 show that, if P is irreducible and admits an invariant probability measure π, then P satisfies a modulated drift condition with $V_1(x) = 1$ on some absorbing and $\pi-$full set. Hence modulated drift condition is a natural assumption. In the discrete state space, any Markov kernel P satisfying the standard communication property [MT09, p. 78] and admitting an invariant probability measure π satisfies all the conclusions of Theorems 5.3, 5.4 and Corollary 5.5. Indeed $S = \{x\}$ for some state x may be chosen such that $\pi(1_{\{x\}}) > 0$, and $S = \{x\}$ is obviously a first-order small-set. We have $\pi = \pi_R$ from Theorem 3.14. Next, it follows from Proposition 5.11 that P satisfies all the conclusions of Theorem 5.3 on a $P-$absorbing and $\pi-$full set $A \in \mathcal{X}$. In fact we have $A = \mathbb{X}$ here: Indeed, otherwise any $x \in A$ would satisfy $P^n(x, A^c) = 0$ for every $n \geq 1$ with $A^c \neq \emptyset$, which contradicts the communication property between any two states. Various examples of discrete Markov models are presented in [Nor97, Bré99, Gra14]. In fact, many of the above conclusions are milestones in Markov theory. In particular, Forster's criterion as a necessary and sufficient condition of existence of a $P-$invariant probability measure (or for positive recurrence) for irreducible Markov kernels, is nothing else that a $1-$modulated drift condition. This explains why the minorization and drift conditions are so popular for studying Markov models.

Note, however, that Proposition 5.11, as well as Proposition 5.10, are only of theoretical interest. In practice the form of the Markov kernel P is directly taken into account to find explicit functions V_0 and V_1 satisfying Condition $\boldsymbol{D}_\psi(V_0, V_1)$. Finally, as shown for instance for random walks on the half line in [JT03], recall that the condition $\pi_R(V_0) < \infty$ is not automatically fulfilled under Condition $\boldsymbol{D}_\psi(V_0, V_1)$. In fact, as proved in Proposition 5.12, this additional condition $\pi_R(V_0) < \infty$ is closely related to an extra V_0-modulated drift condition. Finally recall that, under the sole condition $PV_0 \leq V_0 - V_1 + b_0 1_{\mathbb{X}}$ for some Lyapunov functions V_0, V_1 and positive constant b_0, we have $\pi_R(V_1) \leq b_0$ from [GZ08, Cor. 4]. This bound is proved in Proposition 5.13 under the stronger Conditions $(\boldsymbol{M}_{\nu,\psi})$–$\boldsymbol{D}_\psi(V_0, V_1)$. Also see [Hai21, Prop. 1.4] which provides, under the same assumptions as in [GZ08], a simple proof of the property $\pi_R(V_1) < \infty$ (without bound).

To conclude this section let us just illustrate properties (iv) and (v) stated at the beginning of this section, by simply converting them into a probabilistic version:

Theorem $\mathcal{C}$. *Let $(X_n)_{n\geq 0}$ be a Markov chain on $(\mathbb{X}, \mathcal{X})$. Assume that there exists a non-empty set $S \in \mathcal{X}$ such that*

$$\forall x \in S,\ \mathbb{P}_x(X_1 \in A) \geq \nu(1_A) \quad \textit{and} \quad \forall x \in S^c,\ \mathbb{E}_x[V_0(X_1)] \leq V_0(x) - V_1(x)$$

for some finite positive measure ν and measurable functions $V_0, V_1 : \mathbb{X} \to [1, +\infty)$ such that V_1 and $x \mapsto \mathbb{E}_x[V_0(X_1)]$ are bounded from above on S. Then $(X_n)_{n\geq 0}$ is Harris-recurrent with unique stationary distribution π. Moreover, if $(X_n)_{n\geq 0}$ is aperiodic, then the following convergence in total-variation distance holds:

$$\forall x \in \mathbb{X}, \qquad \lim_{n\to+\infty} \sup_{A\in\mathcal{X}} \big|\mathbb{P}_x(X_n \in A) - \pi(1_A)\big| = 0.$$

6 $V-$geometric ergodicity

Let $V : \mathbb{X} \to [1, +\infty)$ be measurable. Recall that the $V-$geometric drift condition for P is

$$\exists \psi \in \mathcal{B}_+^*,\ \exists \delta \in (0,1),\ \exists b \in (0, +\infty) : \quad PV \leq \delta V + b\,\psi \qquad (\boldsymbol{G}_\psi(\delta, V))$$

and that this condition provides the modulated drift Condition $\boldsymbol{D}_\psi(V_0, V_1)$ with

$$V_0 := V/(1-\delta), \quad V_1 := V \quad \text{and} \quad b_0 := b/(1-\delta) \tag{67}$$

(see Example 5.2). Now in this section, let us assume that P satisfies the first-order minorization condition $(\boldsymbol{M}_{\nu,\psi})$ and the geometric drift condition $\boldsymbol{G}_\psi(\delta, V)$. It follows from Theorem 5.3 and Condition $\boldsymbol{D}_\psi(V_0, V_1)$ with V_0, V_1 and b_0 given in (67) that the residual kernel $R \equiv R_{\nu,\psi}$ given in (13) fulfils the following properties

$$0 \leq \sum_{k=0}^{+\infty} R^k 1_{\mathbb{X}} \leq \sum_{k=0}^{+\infty} R^k V \leq \frac{1+d_0}{1-\delta}\, V \quad \text{with} \quad d_0 := \max\left(0, \frac{b-\nu(V)}{\nu(1_{\mathbb{X}})(1-\delta)}\right) \tag{68a}$$

$$0 \leq \sum_{k=0}^{+\infty} \nu\big(R^k 1_{\mathbb{X}}\big) \leq \sum_{k=0}^{+\infty} \nu\big(R^k V\big) \leq \frac{(1+d_0)\nu(V)}{1-\delta} < \infty, \tag{68b}$$

so that $h_R^\infty = 0$ and $\pi_R := \mu_R(1_{\mathbb{X}})^{-1}\mu_R$ (see (26)) is the unique $P-$invariant probability measure on $(\mathbb{X}, \mathcal{X})$. Moreover we have from Conclusions (iii) and (vi) of Theorem 5.3 that

$$\mu_R(\psi) = 1 \quad \text{and} \quad \pi_R(V) = \pi_R(V_1) < \infty. \tag{69}$$

Below a direct application of Theorem 5.4 and Corollary 5.5 for Poisson's equation provides Corollary 6.1. Then, assuming further the aperiodicity condition (39), the so-called $V-$geometric ergodicity is obtained in Subsection 6.2 using elementary spectral theory.

6.1 Poisson's equation under the geometric drift condition

Corollary 6.1 *Let P satisfy Conditions $(\boldsymbol{M}_{\nu,\psi})$–$\boldsymbol{G}_\psi(\delta, V)$ and let $R \equiv R_{\nu,\psi}$ be the associated residual kernel given in* (13)*. Then:*

1. *For any $g \in \mathcal{B}_V$, the function series $\widetilde{g} := \sum_{k=0}^{+\infty} R^k g$ absolutely converges on $\mathbb{X}$ (point-wise convergence). Moreover we have $\widetilde{g} \in \mathcal{B}_V$ and*

$$\|\widetilde{g}\|_V \leq \frac{1+d_0}{1-\delta}\, \|g\|_V \quad \textit{with} \quad d_0 := \max\left(0\,, \frac{b-\nu(V)}{\nu(1_{\mathbb{X}})(1-\delta)}\right) \tag{70}$$

where δ, b are the constants given in $\boldsymbol{G}_\psi(\delta, V)$.

2. For every $g \in \mathcal{B}_V$ such that $\pi_R(g) = 0$, the function $\widehat{g} := \widetilde{g} - \pi_R(\widetilde{g})1_{\mathbb{X}}$ is the unique π_R-centered function in $\mathcal{B}_V$ solution to Poisson's equation $(I-P)\widehat{g} = g$, and we have

$$\|\widehat{g}\|_V \;\leq\; \frac{(1+d_0)\,(1+\pi_R(V))}{1-\delta}\,\|g\|_V. \tag{71}$$

For the sake of simplicity this statement is directly deduced below from Theorem 5.4 and Corollary 5.5. A self-contained proof of Corollary 6.1 could be also developed starting from (68a) and mimicking the proofs of Theorem 5.4 and Corollary 5.5.

Proof. Using the modulated drift condition $\boldsymbol{D}_\psi(V_0, V_1)$ with V_0, V_1, b_0 given in (67), it follows from Assertion 1. of Theorem 5.4 that

$$\forall g \in \mathcal{B}_V, \quad \|\widetilde{g}\|_{V_0} \leq\; (1+d_0)\|g\|_V \quad \text{with} \quad d_0 := \max\left(0, \frac{b-\nu(V)}{\nu(1_{\mathbb{X}})(1-\delta)}\right)$$

from which we deduce (70) since $\|\cdot\|_{V_0} = (1-\delta)\|\cdot\|_V$. Now, apply Corollary 5.5 to prove Assertion 2. First note that $\pi_R(V_0) < \infty$ since $V_0 = V/(1-\delta)$ and $\pi_R(V) < \infty$ (see (69)). Next we know from Corollary 5.5 that $\widehat{g} = \widetilde{g} - \pi_R(\widetilde{g})1_{\mathbb{X}}$ is a π_R-centered function in $\mathcal{B}_V$ solution to Poisson's equation $(I-P)\widehat{g} = g$. Moreover observe that $\pi_R(V_0)\,\|1_{\mathbb{X}}\|_{V_0} = \pi_R(V)\,\|1_{\mathbb{X}}\|_V \leq \pi_R(V)$. From the first inequality in (64) and again $\|\cdot\|_{V_0} = (1-\delta)\|\cdot\|_V$, we obtained that

$$\|\widehat{g}\|_V \leq \big(1+\pi_R(V)\big)\,\|\widetilde{g}\|_V$$

from which we deduce (71) using (70).

Finally it follows from Condition $\boldsymbol{G}_\psi(\delta, V)$ that PV/V is bounded on $\mathbb{X}$, i.e. $P\mathcal{B}_V \subset \mathcal{B}_V$, since the small-function ψ is bounded and $1_{\mathbb{X}} \leq V$. Then Assertion (viii) of Theorem 5.3 ensures that $E_1 := \{g \in \mathcal{B}_V : Pg = g\} = \mathbb{R}\cdot 1_{\mathbb{X}}$. Hence two solutions to Poisson's equation in $\mathcal{B}_V$ differ from an additive constant. Consequently $\widehat{g}$ is the unique π_R-centered function in $\mathcal{B}_V$ solution to Poisson's equation $(I-P)\widehat{g} = g$. □

6.2 $V-$geometric ergodicity

Recall that, under Conditions $(\boldsymbol{M}_{\nu,\psi})$–$\boldsymbol{G}_\psi(\delta, V)$, we have $h_R^\infty = 0$, so that the aperiodicity condition (39) corresponds to the case $d = 1$ in Theorem 4.14. Now, under Conditions $(\boldsymbol{M}_{\nu,\psi})$–$\boldsymbol{G}_\psi(\delta, V)$ and (39), the so-called $V-$geometric ergodicity of P is proved below. The proof is based on Inequalities (68a)–(68b), Corollary 6.1 and elementary spectral theory. This requires to extend the definition of $\mathcal{B}_V$ to complex-valued functions, that is: For every measurable function $g : \mathbb{X} \to \mathbb{C}$, set $\|g\|_V := \sup_{x\in\mathbb{X}} |g(x)|/V(x) \in [0, +\infty]$ where $|\cdot|$ stands here for the modulus in $\mathbb{C}$, and let us define

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

Note that, under Condition $\boldsymbol{G}_\psi(\delta, V)$, P defines a bounded linear operator on $\mathcal{B}_V$. Since every function g in $\mathcal{B}_V(\mathbb{C})$ writes as $g = g_1 + ig_2$ with $g_1, g_2 \in \mathcal{B}_V$, Pg is simply defined by $Pg = Pg_1 + iPg_2$, so that P obviously defines a bounded linear operator on the Banach space $(\mathcal{B}_V(\mathbb{C}), \|\cdot\|_V)$ too.

Theorem 6.2 *Assume that P satisfies Conditions $(\boldsymbol{M}_{\nu,\psi})$–$\boldsymbol{G}_\psi(\delta, V)$* and is aperiodic (see (39)). Then P is $V-$geometrically ergodic, that is

$$\exists \rho \in (0,1),\ \exists c_\rho > 0,\ \forall g \in \mathcal{B}_V(\mathbb{C}),\ \forall n \geq 1, \quad \|P^n g - \pi_R(g)1_{\mathbb{X}}\|_V \leq c_\rho\, \rho^n \|g\|_V. \tag{72}$$

Note that the geometric rate of convergence in the case of uniform ergodicity (see Example 3.7) corresponds to the $1_{\mathbb{X}}-$geometric ergodicity.

Let $g \in \mathcal{B}_V$ be such that $\pi_R(g) = 0$. It follows from Property (72) that

$$\sum_{k=0}^{+\infty} \|P^k g\|_V \le c(1-\rho)^{-1}\|g\|_V < \infty.$$

Consequently the function series $\mathfrak{g} := \sum_{k=0}^{+\infty} P^k g$ absolutely converges in $(\mathcal{B}_V, \|\cdot\|_V)$ and

$$\|\mathfrak{g}\|_V \le c(1-\rho)^{-1}\|g\|_V.$$

Note that $\mathfrak{g}$ is π_R-centred and satisfies Poisson's equation $(I-P)\mathfrak{g} = g$, so that $\mathfrak{g}$ equals to the function $\widehat{g}$ of Corollary 6.1. Inequality (71) then provides the following alternative bound:

$$\|\mathfrak{g}\|_V \le \frac{(1+d_0)\,(1+\pi_R(V))}{1-\delta}\,\|g\|_V.$$

Now, the needed prerequisites in spectral theory are listed. Let L be a bounded linear operator on a Banach space $(\mathcal{L}, \|\cdot\|)$:

(S1) The spectrum $\sigma(L)$ of L: $\sigma(L) := \{z \in \mathbb{C} : zI - L \text{ is not invertible}\}$ where I denotes the identity map on $\mathcal{L}$. Recall that $\sigma(L)$ is a compact subset of $\mathbb{C}$.

(S2) The operator-norm of L, still denoted by $\|L\|$: $\|L\| := \sup\{\|Lf\| : f \in \mathcal{L}, \|f\| \le 1\}$.

(S3) The spectral radius $r(L)$ of L: $r(L) := \max\{|z| : z \in \sigma(L)\}$,
and Gelfand's formula: $r(L) = \lim_n \|L^n\|^{1/n}$.

Under the assumptions of Theorem 6.2, Lemmas 6.3–6.4 below show that, for any $z \in \mathbb{C}$ such that $|z| = 1$ and $z \neq 1$, the bounded linear operator $zI - P$ on $\mathcal{B}_V(\mathbb{C})$ is invertible.

Lemma 6.3 *If P satisfies Conditions $(\boldsymbol{M}_{\nu,\psi})$–$\boldsymbol{G}_\psi(\delta, V)$ and is aperiodic, then for any $z \in \mathbb{C}$ such that $|z| = 1$ and $z \neq 1$ the bounded linear operator $zI - P$ on $\mathcal{B}_V(\mathbb{C})$ is one-to-one.*

Proof. Let $z \in \mathbb{C}$ be such that $|z| = 1$ and assume that $zI - P$ on $\mathcal{B}_V(\mathbb{C})$ is not one-to-one, that is: there exists $g \in \mathcal{B}_V(\mathbb{C})$, $g \neq 0$, such that $(zI - P)g = 0$. Below this is proved to be only possible for $z = 1$, which provides the desired result. Let $g \in \mathcal{B}_V(\mathbb{C})$, $g \neq 0$, be such that $(zI - P)g = 0$. Since P, thus R, defines a bounded linear operator on the Banach space $(\mathcal{B}_V(\mathbb{C}), \|\cdot\|_V)$, Equality (46) of Lemma 4.16 can be proved similarly, that is we have:

$$\forall n \ge 0, \quad \nu(g)\sum_{k=0}^{n} z^{-(k+1)}R^k\psi = g - z^{-(n+1)}R^{n+1}g.$$

Moreover we know from Assertion 1. of Corollary 6.1 that the series $\widetilde{g} := \sum_{k=0}^{+\infty} R^k g$ point-wise converges on $\mathbb{X}$, thus: $\lim_k R^k g = 0$ (point-wise convergence). Hence we have $g = \nu(g)\widetilde{\psi}_z$, with $\widetilde{\psi}_z := \sum_{k=0}^{+\infty} z^{-(k+1)}R^k\psi$. Recall that $\widetilde{\psi}_z$ is bounded on $\mathbb{X}$ from Proposition 3.4. Thus g is bounded on $\mathbb{X}$, so that z is an eigenvalue of P on $\mathcal{B}(\mathbb{C})$ and $\rho(z) = 1$ from Lemma 4.16, where $\rho(\cdot)$ is defined (38). Since the aperiodicity condition corresponds to the case $d = 1$ in Theorem 4.14, it follows that $z = 1$ from Assertion (a) of Theorem 4.14. □

Lemma 6.4 *If P satisfies Conditions $(\boldsymbol{M}_{\nu,\psi})$-$\boldsymbol{G}_\psi(\delta, V)$ and is aperiodic, then for every $z \in \mathbb{C}$ such that $|z| = 1$ and $z \neq 1$ the bounded linear operator $zI - P$ on $\mathcal{B}_V(\mathbb{C})$ is surjective.*

Proof. Let $z \in \mathbb{C}$ be such that $|z| = 1$ and $g \in \mathcal{B}_V$. Define

$$\forall n \geq 1, \quad \widetilde{g}_{n,z} := \sum_{k=0}^{n} z^{-(k+1)} R^k g.$$

Using $P = R + \psi \otimes \nu$ we obtain that

$$z\widetilde{g}_{n,z} - P\widetilde{g}_{n,z} = z\,\widetilde{g}_{n,z} - R\widetilde{g}_{n,z} - \nu(\widetilde{g}_{n,z})\psi = g - z^{-(n+1)} R^{n+1} g - \nu(\widetilde{g}_{n,z})\psi. \tag{73}$$

Moreover we have

$$\lim_{n \to +\infty} \widetilde{g}_{n,z} = \widetilde{g}_z := \sum_{k=0}^{+\infty} z^{-(k+1)} R^k g \quad \text{(point-wise convergence)} \tag{74}$$

with $\widetilde{g}_z \in \mathcal{B}_V(\mathbb{C})$ since

$$\sum_{k=0}^{+\infty} |z^{-(k+1)} R^k g| \leq \|g\|_V \sum_{k=0}^{+\infty} R^k V \leq c\,V \quad \text{with} \quad c = (1 + d_0)(1 - \delta)^{-1}$$

from the second inequality in (68a). Also note that, for any $x \in \mathbb{X}$, we have $(PV)(x) < \infty$ from Condition $\boldsymbol{D}_\psi(V_0, V_1)$, and that $|\widetilde{g}_{n,z}| \leq c\,V$. It then follows from Lebesgue's theorem w.r.t. the probability measure $P(x, dy)$ that $\lim_n (P\widetilde{g}_{n,z})(x) = (P\widetilde{g}_z)(x)$. Finally we have

$$\lim_{n \to +\infty} \nu(\widetilde{g}_{n,z}) = \lim_{n \to +\infty} \sum_{k=0}^{n} z^{-(k+1)} \nu(R^k g) = \mu_z(g) := \sum_{k=0}^{+\infty} z^{-(k+1)} \nu(R^k g)$$

since the last series converges from $|z^{-(k+1)} \nu(R^k g)| \leq \|g\|_V\, \nu(R^k V)$ and (68b). Then, when n growths to $+\infty$ in Equality (73) (point-wise convergence on $\mathbb{X}$), we obtain that $(zI - P)\widetilde{g}_z = g - \mu_z(g)\psi$. With $g := \psi$ this provides $(zI - P)\widetilde{\psi}_z = \big(1 - \mu_z(\psi)\big)\psi$ with

$$\widetilde{\psi}_z := \sum_{k=0}^{+\infty} z^{-(k+1)} R^k \psi \;\in \mathcal{B}_V(\mathbb{C}) \quad \text{and} \quad \mu_z(\psi) = \sum_{k=0}^{+\infty} z^{-(k+1)} \nu(R^k \psi) = \rho(z^{-1})$$

where $\rho(\cdot)$ is defined (38). Since $z \neq 1$ and $d = 1$ (aperiodicity condition), we know from Assertion (a) of Theorem 4.14 that $\rho(z^{-1}) \neq 1$. Thus

$$(zI - P)\left(\widetilde{g}_z + \frac{\mu_z(g)}{1 - \mu_z(\psi)} \widetilde{\psi}_z\right) = g,$$

from which we deduce that $zI - P$ is surjective. □

Proof of Theorem 6.2. Recall that $\pi_R(V) < \infty$ under the assumptions of Theorem 6.2 (see (69)). Thus π_R defines a bounded linear form on $\mathcal{B}_V(\mathbb{C})$, so that $\mathcal{B}_0 := \{g \in \mathcal{B}_V(\mathbb{C}) : \pi_R(g) = 0\}$ is a closed subspace of $\mathcal{B}_V(\mathbb{C})$. Note that $\mathcal{B}_0$ is $P-$stable (i.e. $P(\mathcal{B}_0) \subset \mathcal{B}_0$) from the $P-$invariance of π_R. Let P_0 be the restriction of P to $\mathcal{B}_0$. Assertion 2. of Corollary 6.1 shows that $I - P_0$ is invertible on $\mathcal{B}_0$. Next let $z \in \mathbb{C}$ be such that $|z| = 1$, $z \neq 1$. It follows from Lemma 6.3 that $zI - P_0$ is one-to-one. Now, let $g \in \mathcal{B}_0$. From Lemma 6.4 there exists $h \in \mathcal{B}_V(\mathbb{C})$ such that $(zI - P)h = g$. We have $(z-1)\pi_R(h) = \pi_R(g) = 0$, thus $\pi_R(h) = 0$ (i.e. $h \in \mathcal{B}_0$) since $z \neq 1$. Hence $zI - P_0$ is surjective.

We have proved that, for every $z \in \mathbb{C}$ such that $|z| = 1$, the bounded linear operator $zI - P_0$ is invertible on $\mathcal{B}_0$. Let $r(P)$ denote the spectral radius of P on $\mathcal{B}_V(\mathbb{C})$. Recall that $r(P) = \lim_n(\|P^n\|_V)^{1/n}$ from Gelfand's formula, where $\|\cdot\|_V$ denotes here the operator norm on $\mathcal{B}_V(\mathbb{C})$. We know that $r(P) \le 1$ from Lemma 5.9 (in fact we have $r(P) = 1$ since $P1_{\mathbb{X}} = 1_{\mathbb{X}}$). Hence the spectral radius $r_0 = r(P_0)$ of P_0 on $\mathcal{B}_0$ is less than one too. In fact we have $r_0 < 1$ since the spectrum $\sigma(P_0)$ of P_0 is a compact subset of $\mathbb{C}$ which, according to the above, is contained in the unit disk of $\mathbb{C}$ and does not contain any complex number of modulus one.

Let $\rho \in (r_0, 1)$. Since $r_0 = \lim_n(\|P_0^n\|_0)^{1/n}$ from Gelfand's formula where $\|\cdot\|_0$ denotes the operator norm on $\mathcal{B}_0$, there exists a positive constant c_ρ such that: $\|P_0^n\|_0 \le c_\rho\, \rho^n$. Thus

$$\begin{aligned}\forall n \ge 1, \forall g \in \mathcal{B}_V(\mathbb{C}),\ \|P^n g - \pi_R(g)1_{\mathbb{X}}\|_V &= \|P^n(g - \pi_R(g)1_{\mathbb{X}})\|_V \ \text{(from } P^n 1_{\mathbb{X}} = 1_{\mathbb{X}}) \\ &= \|P_0^n(g - \pi_R(g)1_{\mathbb{X}})\|_V \ \text{(since } g - \pi_R(g)1_{\mathbb{X}} \in \mathcal{B}_0) \\ &\le c_\rho\, \rho^n\, \|g - \pi_R(g)1_{\mathbb{X}}\|_V \\ &\le c_\rho(1 + \pi_R(V))\, \rho^n\, \|g\|_V \end{aligned} \tag{75}$$

from triangular inequality and $\pi_R(|g|) \le \pi_R(V)\|g\|_V$. This proves (72). □

6.3 Further comments and bibliographic discussion

A detailed and comprehensive history of geometric ergodicity, from the pioneering papers [Mar06, Doe37, Ken59, VJ62] to modern works, can be found in [MT09, Sec. 15.6, 16.6] and [DMPS18, Sec. 15.5]. Theorem 6.2 corresponds to the statement [MT09, Th. 16.1.2] and [DMPS18, Th. 15.2.4], except that it is stated here with a first-order small-function instead of a petite set. We have adopted the modern form of this statement using the weighted-normed space $\mathcal{B}_V$, which was first proposed in this context by [Spi91, HS92] for discrete Markov kernels. The proof of Theorem 6.2 is based on Poisson's equation (Corollary 6.1), combined with the well-known and elementary prerequisites (S1)-(S3) (p. 57) of spectral theory, which can be found for example in [RS80, Yos95, HL99]. The $V-$geometric ergodicity is fully addressed in [MT09, DMPS18] using renewal theory and Nummelin's splitting construction. Alternative proofs can be found in [RR04] based on coupling arguments, in [Bax05] based on renewal theory, in [HM11] based on an elegant idea using Wasserstein distance, in the recent paper [CnM23] based on the dual version of the geometric drift inequality, and finally in [Wu04, Hen06, HL14a, Del17, HL20] based on spectral theory (quasi-compactness) whose first founding ideas are already present in [DF37]. Note that the use of Wasserstein distance in [HM11] requires the condition $\pi_R(1_S) > 1/2$ on the set S in $(\boldsymbol{M}_{\nu,1_S})$ (see Subsection 8.5-F for details). The link between geometric ergodicity for P (in a less advanced form than the current version), and the residual potential kernel, was highlighted in [NT78, NT82]. We refer to the recent paper [GHLR24] where 27 conditions for geometric ergodicity are discussed. We also mention here that, although the statement [HL20, Prop. A.1.] concerning the existence of a $P-$invariant probability measure under the geometric drift condition is true, its proof presented in [HL20, App. A] is only valid for a weak Feller Markov kernel P.

Since the pioneer work [MT94] much effort has been made to find explicit constant c and rate of convergence ρ in Inequality (72). Under Assumptions $(\boldsymbol{M}_{\nu,\psi})$–$\boldsymbol{G}_\psi(\delta, V)$ and the strong aperiodicity condition, such an issue is fully addressed in [Bax05] via renewal theory. Alternative computable upper bounds of the rate of convergence ρ can be found in [LT96, RT99, RT00, Ros02] using splitting or coupling methods, and in [HL14b, HL24] using spectral theory. We refer to [Qin24] for a recent review on various methods for deriving convergence bounds for MCMC. Recall that any methods based on Hairer and Mattinglsy's result [HM11] are faced to the condition $\pi_R(1_S) > 1/2$ for the small-set S. In fact, extra conditions on $\pi_R(1_S)$ appear in others works related to geometric or polynomial rates of convergence (see Section 8). For example the first part in the proof of [RR04, Th. 9] provides a quantitative control on $V-$geometric rate of convergence under some additional condition on the data in Assumptions $(\boldsymbol{M}_{\nu,1_S})$–$\boldsymbol{G}_{1_S}(\delta, V)$: this condition actually requires that $\pi_R(1_S)$ is bounded from below by some explicit positive constant. Without this extra condition, the convergence rate in [RR04, Th. 9] is no longer quantitative. Finally recall that converting bounds on Wasserstein's distance into (weighted) total variation bounds are generally based on [MS10, Th. 12] which requires that the probability measures $P(x, dy)$ have a density with respect to some reference measure (see also [QH22]). In Section 9 the geometric rate of convergence of the iterates of P is addressed. A theoretical result for P acting on a general Banach space $\mathfrak{B}$ is provided, and then applied to the cases $\mathfrak{B} := \mathcal{B}_V$, $\mathfrak{B} := \mathbb{L}^2(\pi_R)$ and $\mathfrak{B} := \mathcal{B}_{V^\alpha}$ for some suitable $\alpha \in (0, 1]$, under Conditions $(\boldsymbol{M}_{\nu,\psi})$–$\boldsymbol{G}_\psi(\delta, V)$. This result depends on the spectral radius $r_{\mathfrak{B}}$ of R on $\mathfrak{B}$ and on the possible solutions to Equation $\rho(z^{-1}) = 1$ in the complex annulus $\{z \in \mathbb{C} : r_{\mathfrak{B}} < |z| < 1\}$, where $\rho(\cdot)$ is the power series introduced in (38).

Poisson's equation for V-geometrically ergodic Markov models is classically studied starting from Inequality (72), which ensures that, for every $g \in \mathcal{B}_V$ such that $\pi_R(g) = 0$, the function $\mathfrak{g} := \sum_{k=0}^{+\infty} P^k g$ in $\mathcal{B}_V$ is the unique π_R-centred solution to Poisson's equation $(I - P)\mathfrak{g} = g$. A quite different development is proposed in this section: Indeed Poisson's equation is first solved in Corollary 6.1 as a by-product of the modulated drift Condition $\boldsymbol{D}_\psi(V_0, V_1)$ (see (67)). Next this study is used for proving the $V-$geometric ergodicity: Indeed note that this prior study of Poisson's equation plays a crucial role at the beginning of the proof of Theorem 6.2 and that the convergent series in (68a)-(68b) are repeatedly used in the proof of Lemmas 6.3-6.4. A standard use of Poisson's equation is to prove a central limit theorem (C.L.T.). Let P be a Markov kernel satisfying Conditions $(\boldsymbol{M}_{\nu,\psi})$ and the $V-$geometric drift condition $\boldsymbol{G}_\psi(\delta, V)$. Then P satisfies Condition $\boldsymbol{D}_\psi(V_0, V_1)$ with V_0, V_1, b_0 given in (67). Consequently, if $\pi(V^2) < \infty$, then the conclusions of Glynn-Meyn's C.L.T., recalled page 53, hold true (note that $\mathcal{B}_{V_1} = \mathcal{B}_V$ here). Mention that the residual kernel R and its iterates have been considered in [KM03] to investigate the eigenvectors belonging to the dominated eigenvalue of the Laplace kernels associated with $V-$geometrically ergodic Markov kernel P. This issue called "multiplicative Poisson equation" in [KM03] is used to prove limit theorems for the underlying Markov chain (also see [KM05]). This question is not addressed in our work.

To conclude this section, let us give a probabilistic form of Theorem 6.2, considering the case $\psi := 1_S$ to simplify:

Theorem $\mathcal{D}$. *Let $(X_n)_{n\geq 0}$ be a Markov chain on $(\mathbb{X}, \mathcal{X})$. Assume that there exists a non-empty set $S \in \mathcal{X}$ such that*

$$\forall x \in S,\ \mathbb{P}_x(X_1 \in A) \geq \nu(1_A) \quad \text{and} \quad \exists \delta \in (0, 1),\ \forall x \in S^c,\ \mathbb{E}_x[V(X_1)] \leq \delta\, V(x)$$

for some finite positive measure ν and Lyapunov function V such that $x \mapsto \mathbb{E}_x[V(X_1)]$ is *bounded from above on S. Then $(X_n)_{n\geq 0}$ admits a unique stationary distribution π, and we have $\pi(V) < \infty$. Moreover, if $(X_n)_{n\geq 0}$ is aperiodic, then there exist $\rho \in (0,1)$ and $c_\rho \in (0,+\infty)$ such that*

$$\forall g \in \mathcal{B}_V(\mathbb{C}),\ \forall n \geq 1,\ \forall x \in \mathbb{X}, \quad \big|\mathbb{E}_x[g(X_n)] - \pi(g)\big| \leq c_\rho\, \rho^n \|g\|_V\, V(x).$$

7 Perturbation results

The main objective of this section is the control of the deviation between the invariant probability measure of a reference Markov kernel and the invariant probability measure of some Markov kernel which is thought of as a perturbation of the reference one. Thus the bounds on the gap on the invariant probability measures are expected to be expressed in function of that on the Markov kernels. To be consistent, such a bound must converge to 0 when the perturbed kernel converges (in some sense) to the reference one. Throughout this section, the reference Markov kernel is assumed to satisfy the first-order minorization condition ($\boldsymbol{M}_{\nu,\psi}$) and the V_1-modulated drift condition $\boldsymbol{D}_\psi(V_0, V_1)$. The control of the gap on the invariant probability measures is in norm $\|\cdot\|'_{V_1}$ and $\|\cdot\|_{\mathrm{TV}}$ (see (8)). The basic tools are: First the fact that, for two Markov kernels P and K with respective invariant probability measures π and κ, we have

$$\forall g \in \mathcal{B}_{V_1}, \quad \kappa(g) - \pi(g) = \kappa((K-P)\xi)$$

where the function ξ is any solution to Poisson's equation $(I-P)\xi = g - \pi(g)1_{\mathbb{X}}$; Second the control of the solution to Poisson's equation provided by Theorem 5.4. Recall that any Markov kernel satisfying both minorization and modulated drift conditions has a unique invariant probability measure (see the introducing part of Section 5 for a list of properties satisfied by such a Markov kernel).

7.1 Main results

First, let us present a statement based on Theorem 5.4 on Poisson's equation. It gives an estimate in norm $\|\cdot\|'_{V_1}$ and $\|\cdot\|_{\mathrm{TV}}$ of the gap between the invariant probability of a Markov kernel P satisfying Conditions ($\boldsymbol{M}_{\nu,\psi}$)–$\boldsymbol{D}_\psi(V_0, V_1)$ and the invariant probability measure κ of any Markov kernel K on $(\mathbb{X}, \mathcal{X})$ satisfying $\|KV_0\|_{V_0} < \infty$ and $\kappa(V_0) < \infty$.

Proposition 7.1 *Assume that P satisfies Conditions ($\boldsymbol{M}_{\nu,\psi}$)–$\boldsymbol{D}_\psi(V_0, V_1)$, with $P-$invariant probability measure denoted by π. Let K be a Markov kernel on $(\mathbb{X}, \mathcal{X})$ with (any) invariant probability measure κ such that $\|KV_0\|_{V_0} < \infty$ and $\kappa(V_0) < \infty$. Assume that the non-negative function Δ_{V_0} defined on $\mathbb{X}$ by*

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

is $\mathcal{X}-$measurable. Then

$$\|\kappa - \pi\|'_{V_1} \leq (1+d_0)(1+\pi(V_1)\|1_{\mathbb{X}}\|_{V_1})\, \kappa(\Delta_{V_0}) \tag{76}$$

where $d_0 := \max(0, (b_0 - \nu(V_0))/\nu(1_{\mathbb{X}}))$ and $\pi(V_1) < \infty$.

The function Δ_{V_0} on $\mathbb{X}$ quantifying the gap between the two Markov kernels is assumed to be $\mathcal{X}-$measurable in Proposition 7.1. In the other statements of this subsection (Proposition 7.2, Theorem 7.3), such a measurability assumption on the corresponding "gap function" is also introduced. It turns out that, when $\mathcal{X}$ is countably generated, the "gap function" is $\mathcal{X}-$measurable. We refer to Subsection 7.4 for some details.

Proof. Recall that $\|PV_0\|_{V_0} < \infty$ from $\boldsymbol{D}_\psi(V_0, V_1)$, so that Δ_{V_0} and $\kappa(\Delta_{V_0})$ are well-defined under the assumptions of Proposition 7.1.

Let $g \in \mathcal{B}_{V_1}$ be such that $\|g\|_{V_1} \le 1$. Since $\pi(V_1) < \infty$ from Assertion (vi) of Theorem 5.3, $\pi(g)$ is well-defined. Introduce $g_0 := g - \pi(g)1_{\mathbb{X}}$ and the residual kernel $R := P - \psi \otimes \nu$. Let $\widetilde{g_0} := \sum_{k=0}^{+\infty} R^k g_0$ be the function in $\mathcal{B}_{V_0}$ provided by Therorem 5.4. Then we have

$$\kappa\big((K-P)\widetilde{g_0}\big) = \kappa(\widetilde{g_0}) - \kappa(\widetilde{g_0} - g_0) = \kappa(g_0) = \kappa(g) - \pi(g) \tag{77}$$

using the $K-$invariance of κ, the Poisson equation $(I-P)\widetilde{g_0} = g_0$ from Theorem 5.4, and finally the definition of g_0. It follows from the definition of the $\mathcal{X}-$measurable function Δ_{V_0} that

$$|\kappa(g) - \pi(g)| \le \int_{\mathbb{X}} \big|(K\widetilde{g_0})(x) - (P\widetilde{g_0})(x)\big|\, \kappa(dx) \le \|\widetilde{g_0}\|_{V_0} \int_{\mathbb{X}} \Delta_{V_0}(x)\, \kappa(dx) = \|\widetilde{g_0}\|_{V_0} \kappa(\Delta_{V_0}).$$

Finally we know from Theorem 5.4 that $\|\widetilde{g}\|_{V_0} \le (1+d_0)\|g_0\|_{V_1}$ with d_0 defined in (59b), so that

$$\|\widetilde{g_0}\|_{V_0} \le (1+d_0)\, \|g - \pi(g)1_{\mathbb{X}}\|_{V_1} \le (1+d_0)\big(1 + \pi(V_1)\|1_{\mathbb{X}}\|_{V_1}\big)$$

from which we deduce (76). □

Now let $\{P_\theta\}_{\theta\in\Theta}$ be a family of transition kernels on $(\mathbb{X}, \mathcal{X})$, where Θ is an open subset of some metric space. 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 which converges (in a certain sense) to P_{θ_0} when $\theta \to \theta_0$. To that effect, when P_{θ_0} satisfies Conditions $(\boldsymbol{M}_{\nu,\psi})$–$\boldsymbol{D}_\psi(V_0, V_1)$ and $\|P_\theta V_0\|_{V_0} < \infty$ for any $\theta \in \Theta \setminus \{\theta_0\}$, we can define

$$\forall \theta \in \Theta,\ \forall x \in \mathbb{X}, \quad \Delta_{\theta,V_0}(x) := \|P_{\theta_0}(x,\cdot) - P_\theta(x,\cdot)\|'_{V_0}. \tag{78}$$

As a direct consequence of Proposition 7.1, we obtain the following perturbation result.

Proposition 7.2 *Assume that the Markov kernel P_{θ_0} satisfies Conditions $(\boldsymbol{M}_{\nu,\psi})$–$\boldsymbol{D}_\psi(V_0, V_1)$, and let π_{θ_0} be the $P_{\theta_0}-$invariant probability measure. Suppose 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 such that $\pi_\theta(V_0) < \infty$. Finally assume that the non-negative function Δ_{θ,V_0} defined in (78) is $\mathcal{X}-$measurable for any $\theta \in \Theta$. Then we have the two following bounds*

$$\|\pi_\theta - \pi_{\theta_0}\|'_{V_1} \le (1+d_0)\, c_{\theta_0}\, \pi_\theta(\Delta_{\theta,V_0}) \tag{79a}$$

$$\|\pi_\theta - \pi_{\theta_0}\|_{\mathrm{TV}} \le 2\,(1+d_0)\, \pi_\theta(\Delta_{\theta,V_0}) \tag{79b}$$

with $d_0 := \max\big(0, (b_0 - \nu(V_0))/\nu(1_{\mathbb{X}})\big)$ and $c_{\theta_0} := 1 + \pi_{\theta_0}(V_1)\|1_{\mathbb{X}}\|_{V_1} < \infty$. If $\pi_{\theta_0}(V_0) < \infty$ then $c_{\theta_0} \le 1 + b_0\|1_{\mathbb{X}}\|_{V_1}$.

Proof. Under these assumptions, the bound in (79a) directly follows from Proposition 7.1 applied to $(P, K) := (P_{\theta_0}, P_\theta)$ with $\theta \neq \theta_0$. If $\pi_{\theta_0}(V_0) < \infty$ then $c_{\theta_0} \le 1 + b_0\|1_{\mathbb{X}}\|_{V_1}$ from Assertion (vii) of Theorem 5.3.

When Condition $\boldsymbol{D}_\psi(V_0, V_1)$ is satisfied, so is Condition $\boldsymbol{D}_\psi(V_0, 1_{\mathbb{X}})$ since $V_1 \geq 1_{\mathbb{X}}$. Thus, the bound (79a) also holds with $V_1 := 1_{\mathbb{X}}$ and then provides the control of the total variation error since $\|\pi_\theta - \pi_{\theta_0}\|_{\mathrm{TV}} = \|\pi_\theta - \pi_{\theta_0}\|'_{1_{\mathbb{X}}}$. Then, using $\pi_{\theta_0}(1_{\mathbb{X}}) = 1$, $\|1_{\mathbb{X}}\|_{1_{\mathbb{X}}} = 1$, so that $c_{\theta_0} = 2$, we obtain the estimate for $\|\pi_\theta - \pi_{\theta_0}\|_{\mathrm{TV}}$ in (79b). □

Note that the bounds in (79a)–(79b) are of interest only when the term $\pi_\theta(\Delta_{\theta,V_0})$ is computable and can be proved to converge to 0 when $\theta \to \theta_0$. Now, the objective is to propose fair assumptions under which the convergence of the deviation between π_θ and π_{θ_0} to zero can be derived from the following natural condition of closeness between P_θ and P_{θ_0}: $\lim_{\theta\to\theta_0} \Delta_{\theta,V_0}(x) = 0$ for any $x \in \mathbb{X}$. A way is to reinforce the knowledge on the Markov kernel P_θ for $\theta \neq \theta_0$. It turns out that, in many perturbation problems, not only does P_{θ_0} satisfies minorization and modulated drift conditions, but so all other transition kernels in the family $\{P_\theta\}_{\theta\in\Theta}$. Such instances are provided by the standard perturbation schemes of Subsection 7.3. Thus, let us introduce the following minorization and modulated drift conditions w.r.t. the family $\{P_\theta\}_{\theta\in\Theta}$: for every $\theta \in \Theta$

$$\exists \psi_\theta \in \mathcal{B}_+^*,\ \exists \nu_\theta \in \mathcal{M}_{+,b}^*,\ P_\theta \geq \psi_\theta \otimes \nu_\theta, \tag{$\boldsymbol{M_\theta}$}$$

and there exists a couple (V_0, V_1) of Lyapunov functions such that, for every $\theta \in \Theta$

$$\exists b_\theta > 0, \quad P_\theta V_0 \leq V_0 - V_1 + b_\theta\, \psi_\theta. \tag{$\boldsymbol{D_\theta}(V_0,V_1)$}$$

Under Condition $\boldsymbol{D_\theta}(V_0, V_1)$, we have $P_\theta V_0 \leq (1 + b_\theta)V_0$ so that the function Δ_{θ,V_0} defined in (78) is well-defined for any $\theta \in \Theta$. Finally, under the additional conditions $\sup_{\theta\in\Theta} b_\theta < \infty$ and $\inf_{\theta\in\Theta} \nu_\theta(1_{\mathbb{X}}) > 0$, let us introduce the following positive constant

$$d := \max\left(0, \sup_{\theta\in\Theta} \frac{b_\theta - \nu_\theta(V_0)}{\nu_\theta(1_{\mathbb{X}})}\right). \tag{80}$$

In Theorem 7.3 below, each Markov kernel P_θ is assumed to satisfy Conditions ($\boldsymbol{M_\theta}$)–$\boldsymbol{D_\theta}(V_0, V_1)$. Thus the $P_\theta-$invariant probability measure denoted by π_θ in these two statements is given by (26) with $\nu := \nu_\theta$ and $R_\theta := P_\theta - \psi_\theta \otimes \nu_\theta$.

Theorem 7.3 *Assume that, for every $\theta \in \Theta$, P_θ satisfies Conditions ($\boldsymbol{M_\theta}$)–$\boldsymbol{D_\theta}(V_0, V_1)$ and that $b := \sup_{\theta\in\Theta} b_\theta < \infty$ and $\inf_{\theta\in\Theta} \nu_\theta(1_{\mathbb{X}}) > 0$. For any $\theta \in \Theta$, the $P_\theta-$invariant probability measure π_θ is assumed to satisfy $\pi_\theta(V_0) < \infty$. Finally, the non-negative function Δ_{θ,V_0} defined in (78) is assumed to be $\mathcal{X}-$measurable.*

Then we have

$$\begin{aligned} \forall \theta \in \Theta, \qquad \|\pi_{\theta_0} - \pi_\theta\|'_{V_1} &\leq (1 + d) \min\left\{c_{\theta_0}\, \pi_\theta(\Delta_{\theta,V_0}),\, c_\theta\, \pi_{\theta_0}(\Delta_{\theta,V_0})\right\} && (81a)\\ \|\pi_\theta - \pi_{\theta_0}\|_{\mathrm{TV}} &\leq 2\,(1 + d)\, \min\left\{\pi_\theta(\Delta_{\theta,V_0}),\, \pi_{\theta_0}(\Delta_{\theta,V_0})\right\} && (81b)\end{aligned}$$

with d defined in (80) and with

$$c_\theta := 1 + \pi_\theta(V_1)\|1_{\mathbb{X}}\|_{V_1} \leq 1 + b\,\|1_{\mathbb{X}}\|_{V_1}. \tag{82}$$

Moreover, if the following convergence 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

$$\lim_{\theta\to\theta_0} \|\pi_\theta - \pi_{\theta_0}\|'_{V_1} = 0 \quad \textit{and} \quad \lim_{\theta\to\theta_0} \|\pi_\theta - \pi_{\theta_0}\|_{\mathrm{TV}} = 0.$$

Proof. Let $\theta \in \Theta$. Recall that $\|P_\theta V_0\|_{V_0} < \infty$ from $\boldsymbol{D_\theta}(V_0, V_1)$. It is assumed that $\pi_\theta(V_0) < \infty$ and that the function Δ_{θ,V_0} is $\mathcal{X}-$measurable. Thus Proposition 7.1 can be applied to $(P, K) := (P_{\theta_0}, P_\theta)$ and to $(P, K) := (P_\theta, P_{\theta_0})$, which provides Inequality (81a). The bounds in (81b) are derived from (81a) as in Proposition 7.2. The assumption $\pi_\theta(V_0) < \infty$ allows us to obtain as in Proposition 7.2 that $c_\theta \le 1 + b_\theta \|1_{\mathbb{X}}\|_{V_1}$. Thus (82) holds with $b := \sup_{\theta\in\Theta} b_\theta < \infty$.

Next, we have

$$\lim_{\theta\to\theta_0} \pi_{\theta_0}(\Delta_{\theta,V_0}) = \lim_{\theta\to\theta_0} \int_{\mathbb{X}} \Delta_{\theta,V_0}(x)\pi_{\theta_0}(dx) = 0 \tag{83}$$

from Lebesgue's theorem using $\Delta_{\theta,V_0} \le 2(1+b)V_0$, $\pi_{\theta_0}(V_0) < \infty$ and Assumption $(\boldsymbol{\Delta_{V_0}})$. Then we obtain that $\lim_{\theta\to\theta_0} \|\pi_\theta - \pi_{\theta_0}\|'_{V_1} = 0$ and $\lim_{\theta\to\theta_0} \|\pi_\theta - \pi_{\theta_0}\|_{\mathrm{TV}} = 0$ from the second bound in (81a)-(81b) and from the inequality (82). □

Let us stress that, in our perturbation context, π_{θ_0} is (generally) unknown and π_θ is expected to be known, so $\pi_\theta(\Delta_{\theta,V_0})$ to be computable. Thus, the bounds of interest in (81a)-(81b) are the following ones

$$\begin{aligned} \|\pi_\theta - \pi_{\theta_0}\|'_{V_1} &\le (1+d)\, c_{\theta_0}\, \pi_\theta(\Delta_{\theta,V_0}) \le (1+d)(1+b\|1_{\mathbb{X}}\|_{V_1})\, \pi_\theta(\Delta_{\theta,V_0}) \\ \|\pi_\theta - \pi_{\theta_0}\|_{\mathrm{TV}} &\le 2\,(1+d)\ \pi_\theta(\Delta_{\theta,V_0}). \end{aligned}$$

The convergence of $\pi_{\theta_0}(\Delta_{\theta,V_0})$ to 0 when $\theta\to\theta_0$ in (83) is of theoretical interest here. It is used to prove that $\lim_{\theta\to\theta_0} \|\pi_\theta - \pi_{\theta_0}\|'_{V_1} = \lim_{\theta\to\theta_0} \|\pi_\theta - \pi_{\theta_0}\|_{\mathrm{TV}} = 0$ in Theorem 7.3.

7.2 Examples

Let us illustrate the results of Theorem 7.3 through the two following examples where the set of parameters Θ is assumed to be some open metric space.

7.2.1 Geometric drift conditions

In the perturbation context, under Condition $(\boldsymbol{M_\theta})$ for any $\theta \in \Theta$, the standard geometric drift conditions for some Lyapunov function V are the following ones (see $\boldsymbol{G}_\psi(\delta, V)$ in Example 5.2):

$$\forall\theta \in \Theta,\ \exists\delta_\theta \in (0,1),\ \exists C_\theta > 0, \quad P_\theta V \le \delta_\theta\, V + C_\theta\, \psi_\theta. \tag{84}$$

Moreover suppose that $C := \sup_{\theta\in\Theta} C_\theta < \infty$ and $\delta := \sup_{\theta\in\Theta} \delta_\theta \in (0,1)$. Since $P_\theta V \le \delta\, V + C\, \psi_\theta$ for any $\theta \in \Theta$, we know from Example 5.2 that

$$\forall\theta \in \Theta, \quad P_\theta V_0 \le V_0 - V_1 + b\, \psi_\theta$$

with $V_0 := V/(1-\delta)$, $V_1 := V$ and $b := C/(1-\delta)$, that is Condition $\boldsymbol{D_\theta}(V_0, V_1)$ is satisfied for any $\theta \in \Theta$. Thus, we know from Theorem 5.3 that the unique $P_\theta-$invariant probability π_θ is such that $\pi_\theta(V_1) = \pi_\theta(V) < \infty$ for any $\theta \in \Theta$. Let $\theta_0 \in \Theta$ be fixed. Assume that the non-negative function Δ_{θ,V_0} is $\mathcal{X}-$measurable for any $\theta \in \Theta$. Consequently, if $\inf_{\theta\in\Theta} \nu_\theta(1_{\mathbb{X}}) > 0$ where $\nu_\theta \in \mathcal{M}^*_{+,b}$ is given in *$(\boldsymbol{M_\theta})$*, then the family $\{P_\theta\}_{\theta\in\Theta}$ satisfies the assumptions of Theorem 7.3 which provides a control of $\|\pi_\theta - \pi_{\theta_0}\|'_V$ and $\|\pi_\theta - \pi_{\theta_0}\|_{\mathrm{TV}}$. Finally, we have $\lim_{\theta\to\theta_0} \|\pi_\theta - \pi_{\theta_0}\|'_V = 0$ and $\lim_{\theta\to\theta_0} \|\pi_\theta - \pi_{\theta_0}\|_{\mathrm{TV}} = 0$, provided that Condition $(\boldsymbol{\Delta}_V)$ is satisfied.

7.2.2 Random walk on the half line

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

$$X_0^{(\theta)} \in \mathbb{X} \quad \text{and} \quad \forall n \geq 1,\ X_n^{(\theta)} := \max\big(0, X_{n-1}^{(\theta)} + \varepsilon_n^{(\theta)}\big) \tag{85}$$

where $\{\varepsilon_n^{(\theta)}\}_{n\geq 1}$ is a sequence of independent and identically distributed $\mathbb{R}$-valued random variables assumed to be independent of $X_0^{(\theta)}$ and to have a 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

$$\forall x \in \mathbb{X},\ \forall A \in \mathcal{X}, \quad P_\theta(x, A) = 1_A(0) \int_{-\infty}^{-x} \mathfrak{p}_\theta(y)\, dy + \int_{-x}^{+\infty} 1_A(x+y)\, \mathfrak{p}_\theta(y)\, dy. \tag{86}$$

Next define the following Lyapunov functions on $\mathbb{X}$:

$$\forall x \in \mathbb{X}, \quad W'(x) = (1+x)^2, \quad V_0'(x) = 1 + x \quad \text{and} \quad V_1(x) = 1.$$

Assume that

$$m_2 := \sup_{\theta\in\Theta} \mathbb{E}\big[|\varepsilon_1^{(\theta)}|^2\big] < \infty \quad \text{and} \quad \exists x_0 > 0, \quad \sup_{\theta\in\Theta} \int_{-x_0}^{+\infty} y\, \mathfrak{p}_\theta(y)\, dy < 0. \tag{87}$$

Let $\theta_0 \in \Theta$ be fixed. Here the state space is $\mathbb{X} := [0, +\infty)$ equipped with its Borel $\sigma-$algebra $\mathcal{X}$ which is countably generated. Therefore for any Lyapunov function on $\mathbb{X}$, say V, and for any $\theta \in \Theta$, the non-negative function on $\mathbb{X}$, $x \mapsto \Delta_{\theta,V}(x) := \|P_\theta(x,\cdot) - P_{\theta_0}(x,\cdot)\|_V'$, is $\mathcal{X}-$measurable. Next, we have for every $x \in \mathbb{X}$

$$\begin{aligned}
(P_\theta V_0')(x) - V_0'(x) &= \int_{-\infty}^{-x} \mathfrak{p}_\theta(y)\, dy + \int_{-x}^{+\infty} (1+x+y)\, \mathfrak{p}_\theta(y)\, dy - (1+x) \\
&= -x \int_{-\infty}^{-x} \mathfrak{p}_\theta(y)\, dy + \int_{-x}^{+\infty} y\, \mathfrak{p}_\theta(y)\, dy \\
&\leq \int_{-x}^{+\infty} y\, \mathfrak{p}_\theta(y)\, dy.
\end{aligned} \tag{88}$$

Let us introduce from *(87)*

$$c_0' := -\sup_{\theta\in\Theta} \int_{-x_0}^{+\infty} y\, \mathfrak{p}_\theta(y)\, dy > 0.$$

Then we obtain from *(87)* and *(88)*

$$\begin{aligned}
\forall x > x_0, \quad & (P_\theta V_0')(x) - V_0'(x) \leq -c_0' V_1(x) \\
\text{and } \forall x \in [0, x_0], \quad & (P_\theta V_0')(x) - V_0'(x) + c_0' V_1(x) \leq \sqrt{m_2} + c_0' V_1(x) = \sqrt{m_2} + c_0',
\end{aligned}$$

that is

$$P_\theta V_0' \leq V_0' - c_0' V_1 + (c_0' + \sqrt{m_2})\, 1_{[0,x_0]}. \tag{89}$$

Next, we get in a similar way that, for any $x \in \mathbb{X}$,

$$
\begin{aligned}
& (P_\theta W')(x) - W'(x) \\
&= \int_{-\infty}^{-x} \mathfrak{p}_\theta(y)\,dy + \int_{-x}^{+\infty} (1+x+y)^2\, \mathfrak{p}_\theta(y)\,dy - (1+x)^2 \\
&= \big(1-(1+x)^2\big) \int_{-\infty}^{-x} \mathfrak{p}_\theta(y)\,dy + 2(1+x)\int_{-x}^{+\infty} y\,\mathfrak{p}_\theta(y)\,dy + \int_{-x}^{+\infty} y^2\,\mathfrak{p}_\theta(y)\,dy \\
&\leq 2\,(1+x)\int_{-x}^{+\infty} y\,\mathfrak{p}_\theta(y)\,dy + \int_{-x}^{+\infty} y^2\,\mathfrak{p}_\theta(y)\,dy. \qquad (90)
\end{aligned}
$$

Using the above constants m_2, c_0' and x_0, we obtain

$$\forall x > x_0, \quad (P_\theta W')(x) - W'(x) \leq -2\,c_0'\,V_0'(x) + m_2.$$

Then it follows from this inequality and from *(90)* that there exists $x_1 > 0$, which only depends on m_2, c_0' such that

$$
\begin{aligned}
\forall x > s := \max(x_0,x_1), \quad (P_\theta W')(x) - W'(x) &\leq -c_0'\,V_0'(x) \\
\text{and } \forall x \in [0,s], \quad (P_\theta W')(x) - W'(x) + c_0'\,V_0'(x) &\leq 2\sqrt{m_2}\,V_0'(x) + m_2 + c_0' V_0'(x) \\
&\leq (2\sqrt{m_2} + c_0')(1+s) + m_2,
\end{aligned}
$$

that is

$$P_\theta W' \leq W' - c_0' V_0' + \big((1+s)(c_0' + 2\sqrt{m_2}) + m_2\big) 1_{[0,s]}. \qquad (91a)$$

Since $s \geq x_0$, we can use in *(89)* the same compact set $[0,s]$ so that

$$P_\theta V_0' \leq V_0' - c_0' V_1 + (c_0' + \sqrt{m_2})\, 1_{[0,s]}. \qquad (91b)$$

It follows from *(91b)* that P_θ, for any $\theta \in \Theta$, satisfies Condition $\boldsymbol{D_\theta}(V_0, V_1)$ with $\psi_\theta := 1_{[0,s]}$, with Lyapunov functions $V_1 := 1_{\mathbb{X}}$ and $V_0 := V_0'/c'$ for $c' := \min(1, c_0')$, and finally with $b_0 := \sup_{\theta\in\Theta} b_\theta \leq (\sqrt{m_2} + c_0')/c'$. Set $S := [0,s]$. Next assume that the following non-negative function

$$\forall y \in \mathbb{R}, \quad \mathfrak{p}_S(y) := \inf_{\theta\in\Theta} \inf_{x\in S} \mathfrak{p}_\theta(y-x)$$

is positive on some open interval of $\mathbb{R}$. Then, for every $\theta \in \Theta$, P_θ satisfies Condition *($\boldsymbol{M_\theta}$)* with $\psi_\theta := 1_S$ and $\nu_\theta := \nu$, where ν is the positive measure on $\mathbb{R}$ defined by

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

(see Proposition 3.1 for details). Note that both ψ_θ and ν_θ do not depend on θ here. Thus, for every $\theta \in \Theta$, P_θ satisfies Conditions *($\boldsymbol{M_\theta}$)*–$\boldsymbol{D_\theta}(V_0, V_1)$ w.r.t. the Lyapunov functions V_0 and V_1 defined above, with $b_0 := \sup_{\theta\in\Theta} b_\theta < \infty$ and $\inf_{\theta\in\Theta} \nu_\theta(1_{\mathbb{X}}) = \nu(1_{\mathbb{X}}) > 0$. Moreover any P_θ has a unique invariant probability measure denoted by π_θ (see Assertion *(iv)* at the beginning of Section 5).

To apply Theorem 7.3, it remains to prove that $\pi_\theta(V_0) < \infty$, for every $\theta \in \Theta$. We have from *(91a)* that P_θ satisfies Conditions *($\boldsymbol{M_\theta}$)*–$\boldsymbol{D_\theta}(W, V_0')$ with $S_\theta := S$ and with Lyapunov functions $V_0'(x) = 1 + x$ and $W(x) = W'(x)/c$. It follows Assertion *(vi)* of Theorem 5.3 that $\pi_\theta(V_0') < \infty$ so that $\pi_\theta(V_0) < \infty$ from $V_0 = V_0'/c$.

Thus, we have proved that Theorem 7.3 applies under Assumptions *(87)* on the noise process $\{\varepsilon_n^{(\theta)}\}_{n\geq 1}$. However, for these statements to be relevant, we have to investigate the function Δ_{θ,V_0} and the quantity $\pi_\theta(\Delta_{\theta,V_0})$. To that effect, recall that $\mathfrak{p}_\theta$ denotes the probability density function of the noise. Now fix some $\theta_0 \in \Theta$ and define

$$\forall \theta \in \Theta, \quad \forall y \in \mathbb{R}, \quad \rho_\theta(y) := |\mathfrak{p}_\theta(y) - \mathfrak{p}_{\theta_0}(y)|,$$
$$\delta_\theta := \int_{\mathbb{R}} \rho_\theta(y)\, dy \quad \text{and} \quad m_{1,\theta} := \int_{\mathbb{R}} |y|\, \rho_\theta(y)\, dy.$$

Note that $\delta_\theta \leq 2$. Let $g \in \mathcal{B}_{V_0}$ be such that $|g| \leq V_0$. Then we have

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

Thus

$$\forall x \in \mathbb{X}, \quad \Delta_{\theta,V_0}(x) \leq \frac{\delta_\theta(1 + c'\, V_0(x)) + m_{1,\theta}}{c'}.$$

Therefore Condition *($\boldsymbol{\Delta}_{\boldsymbol{V_0}}$)* in Theorem 7.3 holds provided that

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

This is a natural assumption on the noise in our perturbation context, that is: When $\theta \to \theta_0$, the distribution of the perturbed noise converges to that of the unperturbed one in total variation distance, as well as in weighted total variation norm.

Finally we have

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

Hence the following bound (see *(81b)*)

$$\|\pi_\theta - \pi_{\theta_0}\|_{\mathrm{TV}} \leq 2\,(1+d)\; \pi_\theta(\Delta_{\theta,V_0}) \qquad \text{with } d := \max\left(0, \frac{b_0 - \nu(V_0)}{\nu(1_{\mathbb{X}})}\right) \tag{92}$$

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

Note that, for this specific model, it follows from [JT03, Prop. 3.5] that

$$\forall \gamma \in [2, +\infty), \quad \mathbb{E}\big[(\max(0, \varepsilon_1^{(\theta)}))^\gamma\big] < \infty \iff \int_{\mathbb{R}} |x|^{\gamma-1}\, \pi_\theta(dx) < \infty.$$

Therefore, under Conditions *(87)*, the Lyapunov function V_0 is expected to be the greatest possible one providing Condition $\boldsymbol{D_\theta}(V_0, 1_{\mathbb{X}})$ with $\pi_\theta(V_0) < \infty$ for any $\theta \in \Theta$.

7.3 Application to standard perturbation schemes

In the two following perturbation schemes – the truncation of infinite stochastic matrices and a state space discretization procedure of non-discrete models – the unperturbed Markov kernel $P := P_{\theta_0}$ satisfies Conditions $(\boldsymbol{M}_{\nu,1_S})$–$\boldsymbol{D}_{1_S}(V_0, V_1)$, that is the minorization and modulated drift conditions for $\psi_{\theta_0} := 1_S$ for some $S \in \mathcal{X}$. Then it turns out that P_θ satisfies Conditions $(\boldsymbol{M}_{\nu,1_S})$–$\boldsymbol{D}_{1_S}(V_0, V_1)$ for any $\theta \in \Theta$. 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 7.3 are straightforward. Finally, note that the $\sigma-$algebra $\mathcal{X}$ associated with the state spaces $\mathbb{X}$ involved in this subsection is countably generated. As previously quoted, it follows that for any $\theta \in \Theta$, the function Δ_{θ,V_0} quantifying the gap between perturbed and unperturbed Markov kernels in Theorem 7.3, is $\mathcal{X}-$measurable. We will therefore no longer refer to this hypothesis here.

7.3.1 Application to truncation-augmentation of discrete Markov kernels

Let $P := (P(x, y))_{(x,y)\in\mathbb{N}^2}$ be a Markov kernel on the discrete set $\mathbb{X} := \mathbb{N}$. Assume that P satisfies Conditions $(\boldsymbol{M}_{\nu,1_S})$ and $\boldsymbol{D}_{1_S}(V_0, V_1)$

$$P \geq 1_S \otimes \nu \quad \text{and} \quad \exists b_0 > 0,\ PV_0 \leq V_0 - V_1 + b_0 1_S$$

with S, ν and V_0 such that:

- S is a finite subset of $\mathbb{N}$ and the support $\mathrm{Supp}(\nu)$ of $\nu \in \mathcal{M}^*_{+,b}$ is a finite subset of $\mathbb{N}$,
- $V_0 := (V(x))_{x\in\mathbb{N}}$ is an unbounded and non-decreasing sequence with $V(0) \geq 1$.

Thus P has a unique invariant probability measure denoted by π.

For any $k \geq 1$, let $B_k := \{0, \ldots, k\}$ and ${B_k}^c := \mathbb{N} \setminus B_k$. Recall that the k-th truncated and arbitrary augmented matrix P_k of the $(k + 1) \times (k + 1)$ north-west corner truncation of P is defined by

$$\forall (x, y) \in {B_k}^2, \quad P_k(x, y) := P(x, y) + P(x, {B_k}^c)\, \kappa_{x,k}(\{y\}) \tag{93}$$

where $\kappa_{x,k}$ is some probability measure on B_k. A linear augmentation corresponds to the case where $\kappa_{x,k} \equiv \kappa_k$ only depends on k. The so-called first or last column linear augmentation corresponds to the case when κ_k is the Dirac distribution at 0 and at k respectively. The goal here is to prove that the $P-$invariant probability measure π can be approximated by the P_k-invariant probability measure π_k, with an explicit error control in function of the integer k. Since P is an infinite matrix, first define the following extended Markov kernel $\widehat{P}_k$ of P_k on $\mathbb{N}$:

$$\forall (x, y) \in \mathbb{N}^2, \quad \widehat{P}_k(x, y) := P_k(x, y)\, 1_{B_k\times B_k}(x, y) + 1_{{B_k}^c\times\{0\}}(x, y).$$

Similarly, if π_k is a P_k-invariant probability measure on B_k, then we define the extended probability measure $\widehat{\pi}_k$ on $\mathbb{N}$ by

$$\forall x \in \mathbb{N}, \quad \widehat{\pi}_k(1_{\{x\}}) := \pi_k(1_{\{x\}})\, 1_{B_k}(x). \tag{94}$$

The next lemma provides the expected results that $\widehat{\pi}_k$ is a $\widehat{P}_k-$invariant probability measure, which is the unique one provided that π_k is the unique P_k-invariant probability measure. The proof is postponed to Appendix C.

Lemma 7.4 *Let P be a Markov kernel on $\mathbb{N}$, and, for any $k \geq 1$, let P_k be the stochastic matrix P_k given in* (93). *If π_k is a P_k-invariant probability measure on B_k, then $\widehat{\pi}_k$ defined in* (94) *is a $\widehat{P}_k-$invariant probability measure on $\mathbb{X}$. If P_k has a unique invariant probability measure, then so is for $\widehat{P}_k$.*

Next, let $k_0 \in \mathbb{N}$ be the smallest integer such that

$$S \subset B_{k_0} \quad \text{and} \quad \mathrm{Supp}(\nu) \subset B_{k_0}. \tag{95}$$

Let us introduce the following family $\{P_\theta\}_{\theta\in\Theta}$ of Markov kernels with $\theta_0 := +\infty$

$$\Theta := \{k \in \mathbb{N} : k \geq k_0\} \cup \{+\infty\}, \quad P_{+\infty} := P, \quad \forall \theta \in \{k \in \mathbb{N} : k \geq k_0\} : P_\theta := \widehat{P}_k. \tag{96}$$

The next proposition provides assumptions under which the family $\{P_\theta\}_{\theta\in\Theta}$ satisfies all the assumptions of Theorem 7.3, so that all the conclusions of this theorem hold in the present truncation context.

Proposition 7.5 *Let P satisfy Conditions $(\boldsymbol{M}_{\nu,1_S})$–$\boldsymbol{D}_{1_S}(V_0, V_1)$ with $P-$invariant probability measure π such that $\pi(V_0) < \infty$. Then, the family $\{P_\theta\}_{\theta\in\Theta}$ defined in* (96) *satisfies all the assumptions of Theorem 7.3 including $(\boldsymbol{\Delta}_{\boldsymbol{V_0}})$.*

The proof of Proposition 7.5 is based on the following Lemmas 7.6-7.7.

Lemma 7.6 *If P satisfies the conditions $(\boldsymbol{M}_{\nu,1_S})$–$\boldsymbol{D}_{1_S}(V_0, V_1)$, then for every integer $k \geq k_0$, the Markov kernel $\widehat{P}_k$ satisfies the same conditions $(\boldsymbol{M}_{\nu,1_S})$–$\boldsymbol{D}_{1_S}(V_0, V_1)$. Thus, for any $k \geq k_0$, $\widehat{P}_k$ and P_k have a unique invariant probability measure $\widehat{\pi}_k$ and π_k.*

Proof. Let $k \geq k_0$. For every $x \in S$ and every $A \subset \mathbb{N}$ we have

$$\widehat{P}_k(x, A) \geq \sum_{y\in A\cap B_k} \widehat{P}_k(x, y) \geq \sum_{y\in A\cap B_k} P(x, y) = P(x, A \cap B_k) \geq \nu(1_{A\cap B_k}) = \nu(1_A)$$

using successively $x \in S \subset B_{k_0} \subset B_k$ and the definitions of $\widehat{P}_k$ and P_k, Assumption $(\boldsymbol{M}_{\nu,1_S})$, and finally $\mathrm{Supp}(\nu) \subset B_{k_0} \subset B_k$. This proves that $\widehat{P}_k$ satisfies Condition $(\boldsymbol{M}_{\nu,1_S})$ with the same S, ν as for P.

Now let us prove that $\widehat{P}_k$ satisfies Condition $\boldsymbol{D}_{1_S}(V_0, V_1)$ for any integer $k \geq 1$. From $\boldsymbol{D}_{1_S}(V_0, V_1)$ for P, it is sufficient to prove that $\widehat{P}_k V_0 \leq P V_0$. Recall that $V_0 := (V_0(x))_{x\in\mathbb{N}}$ is a non-decreasing sequence with $V(0) \geq 1$. Let $k \geq 1$. We have from the definition of $\widehat{P}_k$

$$\begin{aligned}
\forall x \in B_k, \quad (\widehat{P}_k V_0)(x) &= \sum_{y\in B_k} P(x, y)\, V_0(y) + P(x, {B_k}^c) \sum_{y\in B_k} \kappa_{x,k}(y) V_0(y) \\
&\leq \sum_{y\in B_k} P(x, y) V_0(y) + P(x, {B_k}^c) \Big[V_0(k) \sum_{y\in B_k} \kappa_{x,k}(y)\Big] \\
&= \sum_{y\in B_k} P(x, y) V_0(y) + \sum_{y\in {B_k}^c} P(x, y)\, V_0(k) \\
&\leq \sum_{y\in\mathbb{N}} P(x, y) V_0(y) = (P V_0)(x)
\end{aligned} \tag{97}$$

since for any $(y,z) \in B_k \times {B_k}^c$, $V_0(y) \leq V_0(k) \leq V_0(z)$ and since $\kappa_{x,k}(\cdot)$ is a probability measure on B_k. Next, using the definition of $\widehat{P}_k$, we have for any $k \geq 1$

$$\forall x \in {B_k}^c, \quad (\widehat{P}_k V_0)(x) = V_0(0).$$

Note that $V_0(0)1_{\mathbb{X}} \leq V_0$ since V_0 is non-decreasing. Then $V_0(0)P1_{\mathbb{X}} = V_0(0)1_{\mathbb{X}} \leq PV_0$ since P is a non-negative kernel. Therefore, we have that $(\widehat{P}_k V_0)(x) = V_0(0) \leq (PV_0)(x)$ for any $x \in {B_k}^c$. This proves that $\widehat{P}_k$ satisfies $\boldsymbol{D}_{1_S}(V_0, V_1)$. □

The next lemma states that Condition $(\boldsymbol{\Delta_{V_0}})$ holds when P satisfies $(\boldsymbol{M}_{\nu,1_S})$–$\boldsymbol{D}_{1_S}(V_0,V_1)$.

Lemma 7.7 *If P satisfies Conditions $(\boldsymbol{M}_{\nu,1_S})$–$\boldsymbol{D}_{1_S}(V_0,V_1)$, then Condition $(\boldsymbol{\Delta_{V_0}})$ holds true.*

Proof. From the definition of $\widehat{P}_k$ and (93), we have for every $x \in B_k$

$$\begin{aligned}
\Delta_{k,V_0}(x) &= \sum_{y\in\mathbb{N}} \big|P(x,y) - \widehat{P}_k(x,y)\big| V_0(y) \\
&= P(x,{B_k}^c) \sum_{y\in B_k} \kappa_{x,k}(y)V_0(y) + \sum_{y\in {B_k}^c} P(x,y)V_0(y) \\
&\leq P(x,{B_k}^c)\, V_0(k) + \sum_{y\in {B_k}^c} P(x,y)V_0(y) \\
&\leq \sum_{z\in {B_k}^c} P(x,z)V_0(z) + \sum_{y\in {B_k}^c} P(x,y)V_0(y) \ \leq 2 \sum_{y\in {B_k}^c} P(x,y)V_0(y) \qquad (98)
\end{aligned}$$

since V_0 is non-decreasing and $\kappa_{x,k}(B_k) = 1$. Now fix $x \in \mathbb{N}$. Then it follows from (98) applied to any $k > x$ that $\lim_k \Delta_{k,V_0}(x) = 0$ since $\sum_{y\in\mathbb{N}} P(x,y)V_0(y) = (PV_0)(x) < \infty$ from $\boldsymbol{D}_{1_S}(V_0,V_1)$. Thus Condition $(\boldsymbol{\Delta_{V_0}})$ holds true. □

Finally, for the family $\{P_\theta\}_{\theta\in\Theta}$ defined in (96), note that the $P_\theta-$invariant probability measure π_θ for any $\theta \neq \theta_0$, is finitely supported so that $\pi_\theta(V_0) < \infty$. Since the $P_{\theta_0}-$invariant probability measure π_{θ_0} is assumed to satisfy $\pi_{\theta_0}(V_0) < \infty$ in Proposition 7.5, it follows from Lemmas 7.6-7.7 that all the assumptions of Theorem 7.3 hold true. The proof of Proposition 7.5 is complete.

7.3.2 Application to state space discretization

Assume that $(\mathbb{X}, d)$ is a separable metric space equipped with its Borel σ-algebra $\mathcal{X}$, and that P is a Markov kernel on $(\mathbb{X}, \mathcal{X})$ of the form

$$\forall x \in \mathbb{X}, \quad P(x,dy) = p(x,y)\,\lambda(dy), \qquad (99)$$

where $p : \mathbb{X}^2 \to [0,+\infty)$ is a measurable function and λ is a positive measure on $\mathbb{X}$. Typically $\mathbb{X}$ is $\mathbb{R}^d$ and λ is the Lebesgue measure on $\mathbb{R}^d$. Let $x_0 \in \mathbb{X}$ be fixed, and for every integer $k \geq 1$ consider any $\mathbb{X}_k \in \mathcal{X}$ such that

$$\big\{x \in \mathbb{X} : d(x,x_0) < k\big\} \subseteq \mathbb{X}_k \subseteq \big\{x \in \mathbb{X} : d(x,x_0) \leq k\big\}.$$

Now let $(\delta_k)_{k\geq1} \in (0,+\infty)^{\mathbb{N}}$ be such that $\lim_{k\to+\infty} \delta_k = 0$, and for any $k \geq 1$ consider a finite family $\{\mathbb{X}_{j,k}\}_{j\in I_k}$ of disjoint measurable subsets of $\mathbb{X}_k$ such that

$$\mathbb{X}_k = \bigsqcup_{j\in I_k} \mathbb{X}_{j,k} \qquad \text{with} \ \ \forall j \in I_k,\ \operatorname{diam}(\mathbb{X}_{j,k}) \leq \delta_k \qquad (100)$$

where $\mathrm{diam}(\mathbb{X}_{j,k}) := \sup\big\{d(x,x') \,:\, (x,x') \in \mathbb{X}_{j,k}\big\}$. The positive scalar δ_k must be thought of as the mesh of the partition $\{\mathbb{X}_{j,k}\}_{j\in I_k}$ of $\mathbb{X}_k$. Define

$$\forall k \geq 1,\ \forall (x,y) \in \mathbb{X}^2, \quad p_k(x,y) := 1_{\mathbb{X}_k}(y) \sum_{i\in I_k} 1_{\mathbb{X}_{i,k}}(x) \inf_{t\in\mathbb{X}_{i,k}} p(t,y).$$

Observe that $p_k \leq p$. Next define the following submarkovian kernel $\widehat{Q}_k$ on $(\mathbb{X},\mathcal{X})$:

$$\begin{aligned}\forall x \in \mathbb{X},\ \forall A \in \mathcal{X}, \quad \widehat{Q}_k(x,A) &:= \int_{\mathbb{X}} 1_A(y)\, p_k(x,y)\, \lambda(dy) \\ &= \sum_{i\in I_k} \bigg(\int_{\mathbb{X}_k} 1_A(y) \inf_{t\in\mathbb{X}_{i,k}} p(t,y)\, \lambda(dy)\bigg) 1_{\mathbb{X}_{i,k}}(x). \end{aligned} \tag{101}$$

Note that $\widehat{Q}_k(x,\cdot) = 0$ if $x \in {\mathbb{X}_k}^c := \mathbb{X} \setminus \mathbb{X}_k$. Define $\varphi_k := 1_{\mathbb{X}} - \widehat{Q}_k 1_{\mathbb{X}}$. We have $\varphi_k \equiv 1$ on ${\mathbb{X}_k}^c$, and $0 \leq \varphi_k \leq 1_{\mathbb{X}}$ since $0 \leq \widehat{Q}_k 1_{\mathbb{X}} \leq P1_{\mathbb{X}} = 1_{\mathbb{X}}$. Then the kernel $\widehat{P}_k$ defined on $(\mathbb{X},\mathcal{X})$ by

$$\forall x \in \mathbb{X},\ \forall A \in \mathcal{X}, \quad \widehat{P}_k(x,A) := \widehat{Q}_k(x,A) + 1_A(x_0)\, \varphi_k(x) \tag{102}$$

is a Markov kernel. Let $b_k := 1_{{\mathbb{X}_k}^c}$ and let $\mathcal{F}_k$ be the finite-dimensional space spanned by the system of functions $\mathcal{C}_k := \big\{1_{\mathbb{X}_{i,k}},\ i \in I_k\big\} \cup \{b_k\}$ which forms a basis of $\mathcal{F}_k$. For every measurable function $f : \mathbb{X} \to \mathbb{R}$ such that $(\widehat{P}_k|f|)(x) < \infty$ for any $x \in \mathbb{X}$, we have $\widehat{P}_k f \in \mathcal{F}_k$. Define the linear map $P_k : \mathcal{F}_k \to \mathcal{F}_k$ as the restriction of $\widehat{P}_k$ to $\mathcal{F}_k$. Let $N_k := \dim \mathcal{F}_k = \mathrm{Card}\,(I_k) + 1$, and let B_k be the $N_k \times N_k-$matrix defined as the matrix of P_k with respect to the basis $\mathcal{C}_k$ of $\mathcal{F}_k$. The next lemmas states that B_k is a stochastic matrix and that a $\widehat{P}_k-$invariant probability measure can be derived from any invariant probability measure of the finite stochastic matrix B_k. Their proofs are postponed in Appendix C.

Lemma 7.8 *For any $k \geq 1$, the matrix B_k is a stochastic matrix.*

Thus, for any $k \geq 1$, there exists a stochastic row-vector $\pi_k \in [0,+\infty)^{N_k}$ such that

$$\pi_k\, B_k = \pi_k. \tag{103}$$

Note that $P_k b_k = P_k 1_{{\mathbb{X}_k}^c} = \widehat{P}_k 1_{{\mathbb{X}_k}^c} = \widehat{Q}_k 1_{{\mathbb{X}_k}^c} + 1_{{\mathbb{X}_k}^c}(x_0)\, \varphi_k = 0$ (see (102)) so that the last component of π_k is zero. The component of π_k associated with the element $1_{\mathbb{X}_{i,k}}$ of the basis $\mathcal{C}_k$ is denoted by $\pi_{i,k}$, so that $\pi_k \equiv (\{\pi_{i,k}\}_{i\in I_k}\,, 0)$. For every $k \geq 1$, set

$$\widehat{\pi}_k(f) := \pi_k\, F_k \tag{104}$$

where $F_k \equiv F_k(f)$ is the coordinate vector of $\widehat{P}_k f$ in the basis $\mathcal{C}_k$.

Lemma 7.9 *For any $k \geq 1$, let π_k be a B_k-invariant probability measure. Then $\widehat{\pi}_k$ defined in* (104) *is a $\widehat{P}_k-$invariant probability measure and can be written as*

$$\widehat{\pi}_k(dy) = \mathfrak{p}_k(y)\, \lambda(dy) + \bigg(1 - \int_{\mathbb{X}} \mathfrak{p}_k(y)\, \lambda(dy)\bigg)\, \delta_{x_0}, \tag{105a}$$

where δ_{x_0} is the Dirac distribution at x_0 and $\mathfrak{p}_k$ is the non-negative function defined by

$$\forall y \in \mathbb{X}, \quad \mathfrak{p}_k(y) := 1_{\mathbb{X}_k}(y) \sum_{i\in I_k} \pi_{i,k} \inf_{t\in\mathbb{X}_{i,k}} p(t,y). \tag{105b}$$

Next, assume that there exist a positive integer k_0 and $s \in (0, +\infty)$ such that the function

$$y \mapsto g_{k_0,s}(y) := \inf_{x \in S} p_{k_0}(x, y) \quad \text{with} \quad S := \{x \in \mathbb{X},\ d(x, x_0) \leq s\} \tag{106a}$$

is positive on a subset $D \in \mathcal{X}$ such that $\lambda(1_D) > 0$. Then, define $\nu \in \mathcal{M}_{+,b}^*$ by

$$\forall A \subset \mathcal{X}, \quad \nu(1_A) := \int_{\mathbb{X}} 1_A(y)\, g_{k_0,s}(y)\, \lambda(dy). \tag{106b}$$

The Markov kernels P and $\{\widehat{P}_k\}_{k \geq k_0}$ satisfy Condition $(\boldsymbol{M}_{\nu,1_S})$ w.r.t. the above set S and positive measure ν, i.e.

$$P(x, A) \geq \nu(1_A)\, 1_S(x) \qquad \text{and} \quad \forall k \geq k_0,\ \widehat{P}_k(x, A) \geq \nu(1_A)\, 1_S(x) \tag{107}$$

since

$$\forall k \geq k_0,\ \forall (x, y) \in S \times \mathbb{X}, \quad p(x, y) \geq p_k(x, y) \geq p_{k_0}(x, y) \geq g_{k_0,s}(y).$$

Let us introduce the following family of Markov kernels $\{P_\theta\}_{\theta \in \Theta}$ with $\theta_0 := +\infty$ and

$$\Theta := \{k \in \mathbb{N} : k \geq k_0\} \cup \{+\infty\},\ P_{+\infty} := P,\ \forall \theta \in \{k \in \mathbb{N} : k \geq k_0\}, P_\theta := \widehat{P}_k. \tag{108}$$

The next proposition provides assumptions under which this family $\{P_\theta\}_{\theta \in \Theta}$ satisfies all the assumptions of Theorem 7.3, so that all the conclusions of this theorem hold true in the present context of state space discretization.

Proposition 7.10 *Let P be the Markov kernel defined in* (99) *with a function $p(\cdot, \cdot)$ assumed to be such that $x \mapsto p(x, y)$ is continuous on $\mathbb{X}$ for every $y \in \mathbb{X}$. Assume that P satisfies Condition $\boldsymbol{D}_{1_S}(V_0, V_1)$ with respect to S and ν given in* (106a)–(106b) *and to Lyapunov functions $V_i, i = 0, 1$ on $\mathbb{X}$ of the form $V_i(\cdot) := v_i\big(d(\cdot, x_0)\big)$ for some non-decreasing function $v_i : [0, +\infty) \to [1, +\infty)$. Moreover, assume that the $P-$invariant probability measure π is such that $\pi(V_0) < \infty$.*

Then the family $\{P_\theta\}_{\theta \in \Theta}$ defined in (108) *satisfies all the assumptions of Theorem 7.3 including Condition $(\boldsymbol{\Delta}_{\boldsymbol{V_0}})$.*

Recall that, from (107), the family $\{P_\theta\}_{\theta \in \Theta}$ satisfies Condition $(\boldsymbol{M}_{\nu,1_S})$ with S and ν given in (106a)–(106b). The proof of Proposition 7.10 is complete using the two following lemmas. The first one shows that if the unperturbed Markov kernel $P_{\theta_0} := P$ satisfies Condition $\boldsymbol{D}_{1_S}(V_0, V_1)$, then for any $\theta \in \Theta \backslash \{\theta_0\}$, P_θ satisfies the same condition. The second lemma shows that, under the continuity assumption on $p(\cdot, \cdot)$ in Proposition 7.10, Condition $(\boldsymbol{\Delta}_{\boldsymbol{V_0}})$ holds true.

Lemma 7.11 *If P satisfies Condition $\boldsymbol{D}_{1_S}(V_0, V_1)$ then, for any integer $k \geq k_0$, the Markov kernel $\widehat{P}_k$ satisfies the same Condition $\boldsymbol{D}_{1_S}(V_0, V_1)$.*

Proof. Since P satisfies Condition $\boldsymbol{D}_{1_S}(V_0, V_1)$, it is sufficient to show that

$$\widehat{P}_k V_0 \leq P V_0 \tag{109}$$

to prove the first statement. If $x \in \mathbb{X}_k{}^c$, then $(\widehat{P}_k V_0)(x) = V_0(x_0)\,\varphi_k(x) \le V_0(x_0)$ from (102), $\widehat{Q}_k(x,\cdot) = 0$ for $x \in \mathbb{X}_k{}^c$ and $\varphi_k \le 1_{\mathbb{X}}$. Note that $v_0(0)1_{\mathbb{X}} = V_0(x_0)1_{\mathbb{X}} \le V_0$ since v_0 is non-decreasing, so that $V_0(x_0)1_{\mathbb{X}} \le PV_0$ since P is a Markov kernel. Now, let $x \in \mathbb{X}_k$. Then

$$\begin{aligned}
(\widehat{P}_k V_0)(x) &= (\widehat{Q}_k V_0)(x) + V_0(x_0)\big(1 - (\widehat{Q}_k 1_{\mathbb{X}})(x)\big) \quad \text{(from (102))}\\
&= V_0(x_0) + \big(\widehat{Q}_k(V_0 - V_0(x_0)1_{\mathbb{X}})\big)(x)\\
&= V_0(x_0) + \sum_{i\in I_k}\bigg(\int_{\mathbb{X}_k}\big(V_0(y) - V_0(x_0)\big)\inf_{t\in\mathbb{X}_{i,k}} p(t,y)\,\lambda(dy)\bigg)1_{\mathbb{X}_{i,k}}(x) \ \text{(from(101))}\\
&\le V_0(x_0) + \sum_{i\in I_k}\bigg(\int_{\mathbb{X}}\big(V_0(y) - V_0(x_0)\big)\,p(x,y)\,\lambda(dy)\bigg)1_{\mathbb{X}_{i,k}}(x)\\
&\le V_0(x_0) + (PV_0)(x) - V_0(x_0) \quad \Big(\text{since} \sum_{i\in I_k} 1_{\mathbb{X}_{i,k}}(x) = 1_{\mathbb{X}_k}(x) = 1\Big).
\end{aligned}$$

This proves (109). □

Lemma 7.12 *Let $p(\cdot,\cdot)$ in (99) be such that, for every $y \in \mathbb{X}$, the function $x \mapsto p(x,y)$ is continuous on $\mathbb{X}$. Then the following assertion holds:*

$$\forall x \in \mathbb{X}, \quad \lim_k \|P(x,\cdot) - \widehat{P}_k(x,\cdot)\|'_{V_0} = 0.$$

Proof. Let $x \in \mathbb{X}$ be fixed. Observe that

$$\|P(x,\cdot) - \widehat{Q}_k(x,\cdot)\|'_{V_0} \le \int_{\mathbb{X}} V_0(y)\,\big|p(x,y) - p_k(x,y)\big|\,\lambda(dy).$$

From the continuity assumption on the function $p(\cdot,\cdot)$ we have $\lim_k p_k(x,y) = p(x,y)$ for any $y \in \mathbb{X}$, and we know that $|p(x,y) - p_k(x,y)| \le 2p(x,y)$. From Lebesgue's theorem it follows that $\lim_k \|P(x,\cdot) - \widehat{Q}_k(x,\cdot)\|'_{V_0} = 0$ since $(PV_0)(x) < \infty$. Finally note that

$$\begin{aligned}
\|P(x,\cdot) - \widehat{P}_k(x,\cdot)\|'_{V_0} &\le \|P(x,\cdot) - \widehat{Q}_k(x,\cdot)\|'_{V_0} + V_0(x_0)\,\varphi_k(x)\\
&\le \|P(x,\cdot) - \widehat{Q}_k(x,\cdot)\|'_{V_0} + V_0(x_0)\,\|P(x,\cdot) - \widehat{Q}_k(x,\cdot)\|'_{V_0}
\end{aligned}$$

from (102), $\varphi_k(x) := 1 - (\widehat{Q}_k 1_{\mathbb{X}})(x) = (P1_{\mathbb{X}})(x) - (\widehat{Q}_k 1_{\mathbb{X}})(x)$, $1_{\mathbb{X}} \le V_0$ and the definition of $\|\cdot\|_{V_0}$. The proof of the convergence of $\widehat{P}_k(x,\cdot)$ to $P(x,\cdot)$ in V_0-norm is complete. □

Finally, for the family $\{P_\theta\}_{\theta\in\Theta}$ defined in (108), note that the $P_\theta-$invariant probability measure π_θ for any $\theta \ne \theta_0$, is finitely supported so that $\pi_\theta(V_0) < \infty$. Thus, since the $P_{\theta_0}-$invariant probability measure π_{θ_0} is assumed to satisfy $\pi_{\theta_0}(V_0) < \infty$, Theorem 7.3 applies.

7.4 Further comments and bibliographic discussion

A) *Markovian perturbation issue.* The perturbation theory for Markov chains has been widely developed in the last decades, see e.g. [Sch68, Kar86, Sen93, GM96, SS00, AANQ04, Mit05, MA10, FHL13, HL14a, RS18, Mou21, NR21, HL25a, and references therein]. The perturbation material in Section 7 is based on [HL25a]. Moreover here, in Subsection 7.3, two standard issues are analysed as a perturbation problem: truncation in discrete state space case, and discretization of a non-discrete state space. The central Formula (77) was introduced in [Sch68] for finite irreducible stochastic matrices. This formula can be subsequently used in any problem which can be thought of as a perturbation problem of Markov kernels (e.g. see [Sen93, GM96, LL18] and [MT09, Section 17.7]). The investigation of uniformly ergodic perturbed Markov chains is done in [Mit05, MA10, AFEB16, JMMD15]. Perturbation of reversible transition kernels is studied in [MALR16, NR21]. These two specific issues are not addressed here.

B) *On the Conditions* $(\boldsymbol{M_\theta})$–$\boldsymbol{D_\theta}(V_0, V_1)$ *w.r.t. the family* $\{P_\theta\}_{\theta\in\Theta}$. As illustrated in Subsection 7.3, it turns out that, in standard perturbation problems, the perturbed Markov kernels $\{P_\theta\}_{\theta\neq\theta_0}$ inherit the minorization and drift conditions of the unperturbed one P_{θ_0} (e.g. see also [RRS98], [MALR16, Section 4] or [LL18, Section 2.2]). Similarly, if only two Markov kernels P and $\widetilde{P}$ are involved, then both kernels are assumed or are proved to satisfy the same minorization and drift condition under appropriated conditions, e.g. see [RS18, Cor. 4.2], [MARS20, Th. 9]. Mention that Conditions (DRI) in [DDA21, p. 1589] correspond to $(\boldsymbol{M_\theta})$–$\boldsymbol{D_\theta}(V_0, V_1)$ w.r.t. a family of Markov kernels $\{P_\theta\}_{\theta\in\Theta}$ with $\psi_\theta := 1_S$ for some small-set S, minorizing measure $\nu_\theta \equiv \nu$ and constant $b_\theta \equiv b$, and finally with $V_1 = \phi(V_0)$, where the function $\phi(\cdot)$ satisfies specific conditions (see [DFMS04] for details on these modulated drift conditions). Conditions (DRI) are used to study the convergence of stochastic algorithms where Markov dynamic takes place. This concerns Metropolis-Hastings algorithms for which some parameters must be estimated, and also stochastic gradient descent, stochastic Expectation-Maximisation algorithms, or stochastic algorithms in reinforcement learning. Actually Conditions (DRI) are used in [DDA21] to ensure that the specific Assumption [AMP05, (A3)] for convergence of the stochastic algorithm is satisfied when the underlying Markov dynamic is subgeometrically ergodic. Similar conditions were introduced in [AMP05, (DRI1), Section 6] to deal with geometrically ergodic underlying Markov kernels.

C) *On the condition* $\pi(V_0) < \infty$. For a Markov kernel P with invariant probability measure π, the condition $\pi(V_0) < \infty$ is in force in this section. When P satisfies Conditions $(\boldsymbol{M}_{\nu,\psi})$–$\boldsymbol{D}_\psi(V_0, V_1)$, we have $\pi(V_1) < \infty$ from Theorem 5.3, but recall that the condition $\pi(V_0) < \infty$ does not hold automatically. It is in fact satisfied provided that P satisfies $(\boldsymbol{M}_{\nu,\psi})$ and any preliminary V_0-modulated drift condition $\boldsymbol{D}_\psi(L, V_0)$ for some Lyapunov function L. We refer to Proposition 5.12 for a general statement and to the example of Subsection 7.2.2 for a specific situation. Such nested modulated drift conditions $\boldsymbol{D}_\psi(L, V_0)$ and $\boldsymbol{D}_\psi(V_0, V_1)$ occur in most of the analysis of polynomial or subgeometric convergence rate of Markov models, e.g. see [JR02, FM03b, DFMS04, AFV15, DMPS18].

D) *On the measurability of the function* Δ_V. Let P and K be two Markov kernels on $(\mathbb{X}, \mathcal{X})$ and let V be a Lyapunov function such that $\|PV\|_V < \infty$ and $\|KV\|_V < \infty$. Assume that the $\sigma-$algebra $\mathcal{X}$ is countably generated. Then the function on $\mathbb{X}$, $x \mapsto \Delta_V(x) := \|P(x,\cdot) - K(x,\cdot)\|'_V$, is $\mathcal{X}-$measurable. Indeed, for every $x \in \mathbb{X}$ we have $\|P(x,\cdot) - P'(x,\cdot)\|_V = |\eta_x|(V)$ where $|\eta_x|$ is the total variation measure of the finite signed measure $\eta_x = P(x,\cdot) - K(x,\cdot)$. Moreover the map $x \mapsto |\eta_x|(V)$ is $\mathcal{X}-$measurable since so is $x \mapsto \eta_x(V)$, see [DF64].

E) *On the Condition* $(\boldsymbol{\Delta_V})$. As introduced in [Twe98] for discrete set $\mathbb{X}$, Condition $(\boldsymbol{\Delta_V})$

$$\forall x \in \mathbb{X}, \quad \lim_{\theta\to\theta_0} \Delta_{\theta,V}(x) = \lim_{\theta\to\theta_0} \|P_{\theta_0}(x,\cdot) - P_\theta(x,\cdot)\|'_V = 0,$$

is the expected continuity assumption in order to study the convergence to 0 of the $V-$weighted total variation distance between π_θ and π_{θ_0}. Let us discuss Condition ($\boldsymbol{\Delta_V}$) and alternative assumptions used in prior works.

- The standard operator-norm continuity assumption introduced in [Kar86] writes as $\lim_{\theta\to\theta_0} \|P_\theta - P_{\theta_0}\|_V = 0$, namely

$$\lim_{\theta\to\theta_0} \sup_{x\in\mathbb{X}} \frac{\Delta_{\theta,V}(x)}{V(x)} = 0.$$

  This condition is clearly much more restrictive than Condition ($\boldsymbol{\Delta_V}$). Such a condition 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].

- The weak operator-norm continuity assumptions, based on Keller's approach for perturbed dynamical systems [Kel82], require that

$$\lim_{\theta\to\theta_0} \sup_{x\in\mathbb{X}} \frac{\Delta_{\theta,1_{\mathbb{X}}}(x)}{V(x)} = \lim_{\theta\to\theta_0} \sup_{x\in\mathbb{X}} \frac{\|P_\theta(x,\cdot) - P_{\theta_0}(x,\cdot)\|'_{TV}}{V(x)} = 0. \tag{110}$$

  To understand the difference between Conditions ($\boldsymbol{\Delta_V}$) and (110), consider the following simple example derived from perturbed linear autoregressive models (see [FHL13, Ex. 1] for some details on this perturbed model):

$$\forall\theta \in (0,1),\ \forall x \in \mathbb{X} := \mathbb{R},\ \forall A \in \mathcal{X}, \quad P_\theta(x,A) := \int_{\mathbb{R}} 1_A(y)\, \mathfrak{p}(y-\theta x)\, dy,$$

  where $\mathcal{X}$ is here the Borel $\sigma-$algebra on $\mathbb{R}$ and where $\mathfrak{p}$ is some probability density function with respect to Lebesgue's measure on $\mathbb{R}$. Let $\theta_0 \in (0,1)$ be fixed. Condition ($\boldsymbol{\Delta_V}$) writes as follows

$$\forall x \in \mathbb{R}, \quad \lim_{\theta\to\theta_0} \int_{\mathbb{X}} V(y)\big|\mathfrak{p}(y-\theta x) - \mathfrak{p}(y-\theta_0 x)\big|dy = 0, \tag{111}$$

  while Condition (110) is:

$$\lim_{\theta\to\theta_0} \sup_{x\in\mathbb{R}} \frac{\int_{\mathbb{X}} \big|\mathfrak{p}(z-\theta x) - \mathfrak{p}(z-\theta_0 x)\big|dz}{V(x)} = 0. \tag{112}$$

  Conditions (111) and (112) are quite different. In (111) the convergence is simply point-wise, but the presence of $V(y)$ in the integral may be problematic. In (112) the absence of the function V in the integral is of course an advantage, but the uniform convergence in $x \in \mathbb{R}$ may be problematic (even though it can actually only be proved w.r.t. every compact of $\mathbb{R}$ thanks to the division by $V(x)$).

  The weak continuity assumption (110) has been adapted to V-geometrically ergodic Markov models, either using the Keller-Liverani perturbation theorem from [KL99] (see [FHL13, HL14a, HL23b]), or using [HM11] based on Wasserstein distance as in [SS00] or in [RS18, MARS20]. In the next item, the perturbation bound obtained in [HL14a] and [RS18] under this condition (110) is compared with the bound of Theorem 7.3.

F) *Geometric ergodicity case.* If $\{P_\theta\}_{\theta\in\Theta}$ satisfies the assumptions of the example in Subsection 7.2.1, then the bound (81b) of Theorem 7.3 gives

$$\|\pi_\theta - \pi_{\theta_0}\|_{\mathrm{TV}} \le \frac{2\,(1+\tilde{d})}{1-\delta}\;\pi_\theta(\Delta_{\theta,V}) \quad \text{with } \tilde{d} = \frac{1}{1-\delta}\max\left(0, \frac{C}{m}\right) \tag{113}$$

where $m := \inf_{\theta\in\Theta} \nu_\theta(1_{\mathbb{X}}) > 0$. The focus here is on the comparison of the error bound (113) with that obtained in [HL14a, Prop. 2.1] and [RS18, Eq. (3.19)] (see also [HL23b] for the iterated function systems), that is

$$\|\pi_\theta - \pi_{\theta_0}\|_{\mathrm{TV}} \le c\,\gamma_\theta\,\big|\ln\gamma_\theta\big| \quad \text{with} \quad \gamma_\theta := \sup_{x\in\mathbb{X}} \frac{\Delta_{\theta,1_{\mathbb{X}}}(x)}{V(x)} \tag{114}$$

where the positive constant c depends on the constants δ, C given in Subsection 7.2.1 and on the $V-$geometric rate of convergence of the iterates P_θ^n to the invariant distribution π_θ (i.e. Property (72) w.r.t. P_θ). The interest of the bound (114) may be the use of $\Delta_{\theta,1_{\mathbb{X}}}(x)$ rather than $\Delta_{\theta,V}(x)$ in (113). The drawback of (114) is that it involves a logarithm term, but above all that the constant c in (114) depends on the $V-$geometric rate of convergence of P_θ^n to π_θ, which is unknown in general (or badly estimated).

G) *Approximation by truncation.* Here we focus on the approximation by a truncation of the state space $\mathbb{X}$. Specifically we are interested in the so-called truncation-augmentation technique and essentially in the study of convergence of the truncated invariant probability measure $\widehat{\pi}_n$ to π. We refer to [Wol80, Sen06, GS87a, GS87b, KR90, Hey91, Sim95, Twe98, Liu10, Mas16, LL18, and references therein] for countable set $\mathbb{X}$ and [IGL22, IG22, HL25a] for a continuous state space. Note that the stochastic monotonicity property is widely used in the statements of most of these references. Various points related to the results of Subsection 7.3.1 are discussed below, keeping in mind that truncation scheme is considered as a perturbation issue.

- *Convergence of* $\{\widehat{\pi}_k\}_{n\ge 0}$ *to* π. For V-geometrically ergodic discrete Markov chains, the convergence in V-weighted total variation norm is proved to take place in [Twe98, Th 3.2] for the first-column linear augmentation (see (93) with $\kappa_{x,k}$ is a Dirac distribution at 0). Using regeneration methods, such a convergence is extended to V-geometrically or polynomially ergodic Markov chains with continuous state space in [IG22, Th 2] for a specific linear augmentation. The weak convergence in the case of general augmentation of continuous state space Markov chains has been recently addressed in [IGL22]. Note that in such context, the weak convergence does not provide the convergence in the total variation norm.
- *Rate of Convergence of* $\{\widehat{\pi}_k\}_{n\ge 0}$ *to* π. The bound of Theorem 7.3 for a $V-$geometrically ergodic Markov kernel P and $\psi := 1_S$ for some set S (see also Proposition 7.5) then provides a generalization of the bound (10) in [LL18, Th. 2] to a general state-space $\mathbb{X}$ without assuming the existence of an atom. Similarly the bound of Theorem 7.3 extends the bound (16) in [LL18, Th. 3] (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 $\beta < 1$ (see Condition 3 in [LL18, Th. 3]).

H) *Approximation though numerical computations.* The discretization procedure of the general state-space $\mathbb{X}$ in Subsection 7.3.2 can be used to numerically approximate the

$P-$invariant probability measure. This has been proposed in [HL21] in the specific context of a $V-$geometrically ergodic Markov chain. We refer to [HL21] for various illustrations, in particular for autoregressive models. Here, the procedure has been adapted to a general context in Proposition 7.10, where the geometric drift condition is replaced by any modulated drift condition. Fine discretizations of continuous state-space models used on computers introduce round off errors, and therefore produce bias in the results of computations. Thus, it is of interest to show that such a bias is negligible under fair conditions. There, using perturbation techniques may prove relevant, as highlighted for example in [RRS98, BRR01]. Such an issue was discussed in [HL23b] for a more general mechanism of round-off than in [RRS98, BRR01] and for iterated function systems of Lipschitz maps. It should be noted that the problem addressed in [SS00] fits naturally into the current discussion on the use of perturbation techniques for analysing the effect of numerical approximation on the calculation of stationary characteristics. We refer to [RSQ24, RSQ24, CDJT24, and references therein] for such a study in MCMC computations with respect to weighted total variation, Wasserstein and $\chi-$metrics.

8 Polynomial convergence rates

The definitions of the space $\mathcal{B}_V$ and the $V-$weighted total variation norm are recalled in Section 2, see (8). Throughout this section, P is a Markov kernel on $(\mathbb{X}, \mathcal{X})$ satisfying the first-order minorization condition $(\boldsymbol{M}_{\nu,\psi})$ for some $(\nu, \psi) \in \mathcal{M}^*_{+,b} \times \mathcal{B}^*_+$, as well as the following well-known nested modulated drift conditions: There exists a collection $\{V_i\}_{i=0}^m$ of Lyapunov functions with $m \geq 1$ such that

$$\forall i \in \{0, \ldots, m-1\}, \quad V_{i+1} \leq V_i \quad \text{and} \quad \exists b_i > 0,\ PV_i \leq V_i - V_{i+1} + b_i\, \psi. \quad (\boldsymbol{D}_\psi(V_0 : V_m))$$

Note that $V_m \leq \cdots \leq V_i \leq \cdots \leq V_0$. Since Condition $\boldsymbol{D}_\psi(V_0, V_1)$ is contained in $\boldsymbol{D}_\psi(V_0 : V_m)$ it follows from Theorem 5.3 that π_R given in (26) is the unique $P-$invariant probability measure on $(\mathbb{X}, \mathcal{X})$ and that

$$\pi_R(\psi) > 0, \quad \pi_R(V_1) < \infty, \text{ so } \pi_R(V_i) < \infty \text{ for } i \in \{1, \ldots, m\}. \tag{115}$$

Note that the condition $\pi_R(V_0) < \infty$ is not guaranteed a priori. When $m \geq 2$ the purpose of the next Subsection 8.1 is to obtain explicit bounds for the following quantities:

$$\begin{aligned}
\forall g \in \mathcal{B}_{V_m},\ \forall x \in \mathbb{X}, \qquad & \mathcal{S}_{m-2}(g, x) := \sum_{n=0}^{+\infty} (n+1)^{m-2} \big|(P^n g)(x) - \pi_R(g)\big| \\
\forall x \in \mathbb{X}, \qquad & \mathcal{S}_{m-2}(x) := \sum_{n=0}^{+\infty} (n+1)^{m-2} \|P^n(x, \cdot) - \pi_R\|'_{V_m} \\
\text{and} \quad \forall k \geq 0, \quad \forall x \in \mathbb{X}, \qquad & \big\|P^k(x, \cdot) - \pi_R\big\|_{\mathrm{TV}}.
\end{aligned}$$

To this end the technical condition $\pi_R(|\psi - \pi_R(\psi) 1_{\mathbb{X}}|)/\pi_R(\psi) < 1$ (see (121)) is assumed. When it does not hold, replacing this condition by the strong aperiodicity condition $\nu(1_S) > 0$, it is proved in Subsection 8.4 that the results of Subsection 8.1 then extend to some iterate of P.

8.1 The main statements

Let P satisfy Conditions $(\boldsymbol{M}_{\nu,\psi})$ and $\boldsymbol{D}_\psi(V_0 : V_m)$. Set

$$\forall i \in \{0, \ldots, m-1\}, \quad d_i := \max\left(0, \frac{b_i - \nu(V_i)}{\nu(1_{\mathbb{X}})}\right) \tag{117}$$

with constants b_i given in $\boldsymbol{D}_\psi(V_0 : V_m)$. Moreover define $(D_\ell)_{\ell=0}^{m-1} \in (0+\infty)^m$ by

$$D_0 := 1 + d_0 \quad \text{and} \quad \forall \ell \in \{1, \ldots, m-1\},\ D_\ell := (1+d_\ell) \sum_{j=0}^{\ell-1} \binom{\ell}{j} D_j \tag{118}$$

where $\binom{\ell}{j} = \ell!/j!(\ell-j)!$ denotes the standard binomial coefficient.

Proposition 8.1 *Let P satisfy Conditions $(\boldsymbol{M}_{\nu,\psi})$–$\boldsymbol{D}_\psi(V_0 : V_m)$ for some collection $\{V_i\}_{i=0}^m$ of Lyapunov functions with $m \geq 1$. Let $R \equiv R_{\nu,\psi}$ be the residual kernel given in* (13)*. Then*

$$\forall i \in \{1, \ldots, m\}, \qquad \sum_{n=0}^{+\infty} (n+1)^{i-1} R^n V_i \;\leq\; D_{i-1} V_0, \tag{119a}$$

$$\sum_{n=0}^{+\infty} (n+1)^{i-1} \nu\big(R^n V_i\big) \;\leq\; D_{i-1}\, \nu(V_0) < \infty. \tag{119b}$$

Proof. Let us prove Inequalities (119a) by an induction on m. For $m := 1$, Inequality (119a) corresponds to the second inequality of (57a) in Theorem 5.3. Now suppose that Inequalities (119a) are proved for some $m \geq 1$ and that Conditions $\boldsymbol{D}_\psi(V_0 : V_{m+1})$ hold. Then it follows from Lemma 5.8 used under Condition $\boldsymbol{D}_\psi(V_m, V_{m+1})$ that $RV_{m,d_m} \leq V_{m,d_m} - V_{m+1}$ with $V_{m,d_m} := V_m + d_m 1_{\mathbb{X}} \geq V_m$, where $d_m := \max(0, \nu(1_{\mathbb{X}})^{-1}(b_m - \nu(V_m)))$. Equivalently we have $V_{m+1} \leq V_{m,d_m} - RV_{m,d_m}$, so that we obtain for every $N \geq 1$

$$\begin{aligned}
\sum_{n=0}^{N} (n+1)^m R^n V_{m+1} \;&\leq\; \sum_{n=0}^{N} (n+1)^m R^n V_{m,d_m} - \sum_{n=0}^{N+1} n^m R^n V_{m,d_m} \\
&\leq\; \sum_{n=0}^{N} \big[(n+1)^m - n^m\big] R^n V_{m,d_m} = \sum_{j=0}^{m-1} \binom{m}{j} \sum_{n=0}^{N} n^j R^n V_{m,d_m} \\
&\leq\; \big(1+d_m\big) \sum_{j=0}^{m-1} \binom{m}{j} \sum_{n=0}^{N} n^j R^n V_{j+1} \\
&\leq\; \big(1+d_m\big) \bigg(\sum_{j=0}^{m-1} \binom{m}{j} D_j \bigg) V_0 \;= D_m V_0
\end{aligned}$$

using the binomial expansion, $V_{m,d_m} \leq (1+d_m)V_m \leq (1+d_m)V_{j+1}$ for $j = 0, \ldots, m-1$, the induction hypothesis, and using finally the definition of D_m. This gives Inequalities (119a) at order $m+1$. Finally (119b) follows from (119a) since, for some $x \in S$, we have from Assumption $(\boldsymbol{M}_{\nu,\psi})$: $\nu(V_0) \leq (PV_0)(x) \leq V_0(x) - V_1(x) + b_0 < \infty$. □

Now recall that, under Conditions $(\boldsymbol{M}_{\nu,\psi})$ and $\boldsymbol{D}_\psi(V_0 : V_m)$, the positive measure $\mu_R := \sum_{k=0}^{+\infty} \nu R^k$ satisfies $\mu_R(1_{\mathbb{X}}) < \infty$ and $\mu_R = \pi_R/\pi_R(\psi)$ from Theorem 5.3. Let us introduce the following functions $\Phi_i : \mathbb{X} \to [0, +\infty]$ for $i \in \{0, \ldots, m-2\}$:

$$\Phi_i := \sum_{n=0}^{+\infty} (n+1)^i \left|P^n \phi\right| \quad \text{where} \quad \phi \equiv \phi_\psi := \psi - \pi_R(\psi) 1_{\mathbb{X}} \tag{120}$$

with ψ given in Assumptions $(\boldsymbol{M}_{\nu,\psi})$ and $\boldsymbol{D}_\psi(V_0 : V_m)$, as well as the following condition

$$\mu_R(|\phi|) = \frac{\pi_R(|\phi|)}{\pi_R(\psi)} < 1. \tag{121}$$

Recall that, for every $m \geq 2$, there exists $\{a_{j,m}\}_{j=1}^{m-1} \in \mathbb{R}^{m-1}$ such that

$$\forall k \geq 1, \quad \Sigma_k^{m-2} := \sum_{n=1}^{k} n^{m-2} = \sum_{j=1}^{m-1} a_{j,m}\, k^j, \tag{122}$$

and that the real numbers $\{a_{j,m}\}_{j=1}^{m-1}$ can be computed by induction on m using binomial expansion (e.g. see Subsection 8.3.1 in cases $m := 2, 3$). Next, using D_j's in (118), define the following positive constants

$$\forall \ell \in \{1, \ldots, m-1\}, \quad E_\ell := \sum_{j=1}^{\ell} a_{j,\ell+1}\, D_j. \tag{123}$$

Theorem 8.2 *Let P satisfy Conditions $(\boldsymbol{M}_{\nu,\psi})$–$\boldsymbol{D}_\psi(V_0 : V_m)$ for some collection $\{V_i\}_{i=0}^m$ of Lyapunov functions with $m \geq 2$. Then the following inequalities hold in $[0, +\infty]$:*

$$\begin{aligned} \forall g \in \mathcal{B}_{V_m},\ \forall x \in \mathbb{X}, \quad \mathcal{S}_{m-2}(g, x) &:= \sum_{n=0}^{+\infty} (n+1)^{m-2} \big|(P^n g)(x) - \pi_R(g)\big| \\ &\leq \|g - \pi_R(g) 1_{\mathbb{X}}\|_{V_m}\, W_m(x) \qquad (124) \\ \text{and} \quad \forall x \in \mathbb{X}, \quad \mathcal{S}_{m-2}(x) &:= \sum_{n=0}^{+\infty} (n+1)^{m-2} \big\|P^n(x, \cdot) - \pi_R\big\|'_{V_m} \\ &\leq \theta_m\, W_m(x) \qquad (125) \end{aligned}$$

where $\theta_m := 1 + \pi_R(V_m)\|1_{\mathbb{X}}\|_{V_m}$ and the function W_m is

$$W_m = D_{m-2}\, V_0 + \nu(V_0) \Big[\sum_{j=0}^{m-2} \binom{m-2}{j} D_j\, \Phi_{m-2-j} + \pi_R(\psi)\, E_{m-1}\, 1_{\mathbb{X}} \Big]. \tag{126}$$

If Condition (121) holds, then for every $i \in \{0, \ldots, m-2\}$ we have $\Phi_i \in \mathcal{B}_{V_0}$ and

$$\Phi_i \;\leq\; C_\phi \|\phi\|_{1_{\mathbb{X}}} \Big(D_i V_0 + \nu(V_0) \sum_{j=1}^{i} \binom{i}{j} D_j \Phi_{i-j} + \pi_R(\psi)\nu(V_0) E_{i+1} 1_{\mathbb{X}} \Big) \tag{127}$$

with the convention $\sum_{j=1}^{0} = 0$ and with $C_\phi := (1 - \mu_R(|\phi|))^{-1}$ where $\mu_R(|\phi|)$ is given in (121).

It follows from (126) that

$$\|W_m\|_{V_0} \le D_{m-2} + \nu(V_0) \sum_{j=0}^{m-2} \binom{m-2}{j} D_j \, \|\Phi_{m-2-j}\|_{V_0} + \pi_R(\psi)\, \nu(V_0) E_{m-1} \|1_{\mathbb{X}}\|_{V_0}. \quad (128)$$

Moreover, if Condition (121) holds, then the norms $(\|\Phi_i\|_{V_0})_{i=0}^{m-2}$ can be recursively bounded from (127) by

$$\|\Phi_i\|_{V_0} \le C_\phi \Big(D_i + \nu(V_0) \big[\sum_{j=1}^{i} \binom{i}{j} D_j \, \|\Phi_{i-j}\|_{V_0} + \pi_R(\psi)\, E_{i+1} \|1_{\mathbb{X}}\|_{V_0} \big] \Big). \quad (129)$$

Consequently, for every $x \in \mathbb{X}$, Theorem 8.2 provides explicit bounds for $\mathcal{S}_{m-2}(g,x)$ with $g \in \mathcal{B}_{V_m}$ and for $\mathcal{S}_{m-2}(x)$. In the atomic case (more generally if $b_i \le \nu(V_i)$), the d_i's in (117) are zero (see Remark 5.7), so that the constants D_i defined in (118) and used in the previous estimates simply depend on the integer m. Finally note that $\|1_{\mathbb{X}}\|_{V_m} \le 1$ since $V_m \ge 1_{\mathbb{X}}$ and that $\pi_R(V_m) \le b_{m-1}\pi_R(\psi)$ applying Assertion (vii) of Theorem 5.3 under Condition $\boldsymbol{D}_{\boldsymbol{\psi}}(V_{m-1}, V_m)$. Thus the positive constant θ_m of Theorem 8.2 satisfies

$$\theta_m \le 1 + b_{m-1}\pi_R(\psi).$$

As a by-product of Theorem 8.2 we obtain the following statement on the total variation norm of $P^k(x,\cdot) - \pi_R$.

Corollary 8.3 *Let P satisfy Conditions $(\boldsymbol{M}_{\nu,\psi})$–$\boldsymbol{D}_\psi(V_0 : V_m)$ for some collection $\{V_i\}_{i=0}^m$ of Lyapunov functions with $m \ge 2$. If Condition* (121) *holds then*

$$\forall x \in \mathbb{X},\ \forall k \ge 0, \quad \big\|P^k(x,\cdot) - \pi_R\big\|_{\mathrm{TV}} \le \frac{2^m}{k^{m-1}}\, W_m(x) \quad (130)$$

with W_m given in Theorem 8.2.

Proof. Note that V_m in $\boldsymbol{D}_\psi(V_0 : V_m)$ can be replaced with the function $1_{\mathbb{X}}$ since $V_m \ge 1_{\mathbb{X}}$. Let $x \in \mathbb{X}$. Recall that the sequence $(\|P^n(x,\cdot) - \pi_R\|_{\mathrm{TV}})_{n\ge 0}$ is non-increasing. Let $j \ge 0$. Then we deduce from (125) that

$$(j+1)^{m-1} \, \big\|P^{2j}(x,\cdot) - \pi_R\big\|_{\mathrm{TV}} \le \sum_{n=j}^{2j} (n+1)^{m-2} \big\|P^n(x,\cdot) - \pi_R\big\|_{\mathrm{TV}} \le \theta_m \, W_m(x)$$

with $\theta_m := 1 + \pi_R(V_m)\|1_{\mathbb{X}}\|_{V_m} = 2$ since $V_m = 1_{\mathbb{X}}$ here. Thus

$$\big\|P^{2j}(x,\cdot) - \pi_R\big\|_{\mathrm{TV}} \le \frac{2^m}{(2j)^{m-1}}\, W_m(x).$$

Next, using the sum $\sum_{n=j+1}^{2j+1}$, we obtain the same inequality for $\|P^{2j+1}(x,\cdot) - \pi_R\|_{\mathrm{TV}}$ replacing $(2j)^{m-1}$ with $(2j+1)^{m-1}$. This proves (130). □

In the case $\psi := 1_S$, Condition (121) is proved to be equivalent to $\pi_R(1_S) > 1/2$ in Proposition 8.5. This condition $\pi_R(1_S) > 1/2$ has been used in others works and proved to imply the aperiodicity condition (see the bibliographic comments in Subsection 8.5). This last fact still holds for general first-order small-function ψ, and this is easily seen via Corollary 8.3 when P satisfies Conditions $(\boldsymbol{M}_{\nu,\psi})$–$\boldsymbol{D}_\psi(V_0 : V_m)$ for some $m \geq 2$. Indeed, under these assumptions, the function $h_R^\infty := \lim_n R^n 1_{\mathbb{X}}$ is zero on $\mathbb{X}$ from Theorem 5.3. Now, if Condition (121) is moreover assumed, then it follows from Corollary 8.3 that $\lim_k \|P^k(x,\cdot) - \pi_R\|_{\mathrm{TV}} = 0$ for every $x \in \mathbb{X}$, so that $z := 1$ is the only eigenvalue of modulus one for P on $\mathcal{B}(\mathbb{C}) \equiv \mathcal{B}_{1_{\mathbb{X}}}(\mathbb{C})$: Theorem 4.14 then implies that P is aperiodic.

8.2 Proof of Theorem 8.2

To prove Theorem 8.2 recall that $\phi := \psi - \pi_R(\psi)1_{\mathbb{X}}$ and let us introduce the following functions under Conditions $(\boldsymbol{M}_{\nu,\psi})$–$\boldsymbol{D}_\psi(V_0 : V_m)$:

$$\forall i \in \{0,\ldots,m-2\},\ \forall N \geq 1, \quad \Phi_{i,N} := \sum_{n=0}^{N} (n+1)^i \left|P^n\phi\right|. \tag{131}$$

The following lemma plays a crucial role in the proof of Theorem 8.2.

Lemma 8.4 *Let P satisfy Conditions $(\boldsymbol{M}_{\nu,\psi})$–$\boldsymbol{D}_\psi(V_0 : V_\ell)$ for some collection $\{V_i\}_{i=0}^{\ell}$ of Lyapunov functions with $\ell \geq 2$. Let $(g_n)_{n\geq 0} \in \mathcal{B}_{V_\ell}^{\mathbb{N}}$ and $\zeta \in \mathcal{B}_{V_\ell}$ such that $|g_n| \leq \zeta \leq V_\ell$ and $\pi_R(g_n) = 0$ for every $n \geq 0$. Then we have for every $N \geq 1$ (with the convention $\sum_{j=1}^{0} = 0$)*

$$\begin{aligned}\sum_{n=0}^{N} (n+1)^{\ell-2}\left|P^n g_n\right| \quad &\leq \quad D_{\ell-2}\, V_0 + \left(\sum_{k=1}^{+\infty} \nu(R^{k-1}\zeta)\right) \Phi_{\ell-2,N} \\ &\qquad + \nu(V_0)\left[\sum_{j=1}^{\ell-2} \binom{\ell-2}{j} D_j\, \Phi_{\ell-2-j,N} + \pi_R(\psi)\, E_{\ell-1}\, 1_{\mathbb{X}}\right]. \end{aligned} \tag{132}$$

Let us admit Lemma 8.4 for the moment and prove Theorem 8.2.

Proof of Theorem 8.2. Note that $\Phi_{i,N} \leq \Phi_i$ for every $N \geq 1$, with Φ_i given in (120). If $g \in \mathcal{B}_{V_m}$ is such that $\|g\|_{V_m} \leq 1$ and $\pi_R(g) = 0$, then Inequality (124) in $[0, +\infty]$ with W_m given in (126) directly follows from Inequality (132) applied to $\ell := m$, $g_n := g, \zeta := V_m$, and from

$$\sum_{k=1}^{+\infty} \nu(R^{k-1}V_m) \leq \sum_{k=1}^{+\infty} \nu(R^{k-1}V_1) \leq D_0\, \nu(V_0) \tag{133}$$

thanks to (119b) applied with $i := 1$. The formulation in (124) given for any $g \in \mathcal{B}_{V_m}$ is then easily deduced considering the function $(g - \pi_R(g)1_{\mathbb{X}})/\|g - \pi_R(g)1_{\mathbb{X}}\|_{V_m}$.

Next, to prove Inequality (125), recall that $\theta_m = 1 + \pi_R(V_m)\|1_{\mathbb{X}}\|_{V_m}$, and first note that

$$\forall h \in \mathcal{B}_{V_m}, \quad \|h - \pi_R(h)1_{\mathbb{X}}\|_{V_m} \leq \theta_m \|h\|_{V_m}.$$

Now let $(h_n)_{n\geq 0} \in \mathcal{B}_{V_m}^{\mathbb{N}}$ be such that $\|h_n\|_{V_m} \leq 1$ and set $f_n := h_n - \pi_R(h_n)1_{\mathbb{X}}$. For any $n \geq 0$, we have $\|f_n\|_{V_m} \leq \theta_m$, so that $g_n := f_n/\theta_m$ is such that $|g_n| \leq V_m$ and $\pi_R(g_n) = 0$. Then, applying Inequality (132) to $\ell := m$, $\zeta := V_m$, we obtain that

$$\forall x \in \mathbb{X},\ \forall N \geq 1, \quad \sum_{n=0}^{N} (n+1)^{m-2}\left|(P^n h_n)(x) - \pi_R(h_n)\right| \leq \theta_m\, W_m(x)$$

using again (133). Taking the suppremum bound over the functions $h_0, \ldots, h_N$, we obtain that

$$\forall x \in \mathbb{X},\ \forall N \geq 1, \quad \sum_{n=0}^{N} (n+1)^{m-2} \|P^n(x,\cdot) - \pi_R\|'_{V_m} \leq \theta_m\, W_m(x)$$

from which we deduce (125).

Now observe that Assumptions $\boldsymbol{D}_\psi(V_0 : V_m)$ obviously imply that, for every $i = 0, \ldots, m-2$, Assumptions $\boldsymbol{D}_\psi(V_0 : V_{i+2})$ hold too. Therefore, for any $i = 0, \ldots, m-2$, it follows from Inequality (132) with $\ell := i+2$ applied to $g_n := \phi/\|\phi\|_{1_{\mathbb{X}}}$, $\zeta := |\phi|/\|\phi\|_{1_{\mathbb{X}}}$ and from Condition (121) that

$$\frac{1-\mu_R(|\phi|)}{\|\phi\|_{1_{\mathbb{X}}}}\, \Phi_{i,N} \leq D_i\, V_0 + \nu(V_0)\Big[\sum_{j=1}^{i} \binom{i}{j} D_j\, \Phi_{i-j,N} + \pi_R(\psi)\, E_{i+1}\, 1_{\mathbb{X}}\Big]$$

since $\sum_{k=1}^{+\infty} \nu(R^{k-1}|\phi|) = \mu_R(|\phi|)$. Note that the above functions g_n and ζ satisfy the assumptions of Lemma 8.4 since $|g_n| \leq \zeta \leq 1_{\mathbb{X}} \leq V_\ell$. Recall that $\sum_{j=1}^{0} = 0$ by convention in (132). When $N \to +\infty$, the previous inequality for $i = 0$ shows that the series Φ_0 is convergent and satisfies (127) for $i := 0$. Next this inequality for $i \in \{1, \ldots, m-2\}$ ensures that the series Φ_i is convergent from the convergence of the $(\Phi_j)_{j=0}^{i-1}$, and that Φ_i satisfies Inequality (127). The proof of Theorem 8.2 is complete, provided that Lemma 8.4 is proved. □

Proof of Lemma 8.4. Let $(g_n)_{n\geq 0} \in \mathcal{B}_{V_\ell}^{\mathbb{N}}$ and $\zeta \in \mathcal{B}_{V_\ell}$ such that $|g_n| \leq \zeta \leq V_\ell$ and $\pi_R(g_n) = 0$ for every $n \geq 0$. Note that $\mu_R(g_n) := \sum_{k=1}^{+\infty} \nu(R^{k-1} g_n) = 0$ since $\pi_R(g_n) = 0$. Then, from Formula (17) and $\sum_{k=1}^{n} \nu(R^{k-1} g_n) = -\sum_{k=n+1}^{+\infty} \nu(R^{k-1} g_n)$ with the convention $\sum_{k=1}^{0} = 0$, we obtain that

$$\begin{aligned}
\forall n \geq 0, \quad P^n g_n &= R^n g_n + \sum_{k=1}^{n} \nu(R^{k-1} g_n)\, P^{n-k}\psi \\
&= R^n g_n + \sum_{k=1}^{n} \nu(R^{k-1} g_n)\, P^{n-k}\phi \,-\, \pi_R(\psi)\Big(\sum_{k=n+1}^{+\infty} \nu(R^{k-1} g_n)\Big) 1_{\mathbb{X}}. \quad (134)
\end{aligned}$$

First, using the non-negativity of R and $|g_n| \leq V_\ell \leq V_{\ell-1}$, it follows from (119a) with $i = \ell-1$ that

$$A_N := \sum_{n=0}^{N} (n+1)^{\ell-2}\, |R^n g_n| \leq \sum_{n=0}^{+\infty} (n+1)^{\ell-2}\, R^n |g_n| \leq \sum_{n=0}^{+\infty} (n+1)^{\ell-2}\, R^n V_{\ell-1} \leq D_{\ell-2}\, V_0. \quad (135)$$

Second, using again the convention $\sum_{k=1}^{0} = 0$ and the inequality $|g_n| \le \zeta$, we have

$$\begin{aligned} B_N := \sum_{n=0}^{N} (n+1)^{\ell-2} \Big| \sum_{k=1}^{n} \nu(R^{k-1} g_n)\, P^{n-k}\phi \Big| \ &\le \ \sum_{n=0}^{N} (n+1)^{\ell-2} \sum_{k=1}^{n} \nu(R^{k-1}|g_n|)\, \big|P^{n-k}\phi\big| \\ &= \sum_{k=1}^{N} \nu(R^{k-1}|g_n|) \sum_{n=k}^{N} (n+1)^{\ell-2} \big|P^{n-k}\phi\big| \\ &\le \ \sum_{k=1}^{N} \nu(R^{k-1}\zeta) \sum_{n=0}^{N} (n+1+k)^{\ell-2} \big|P^{n}\phi\big| \\ &= \sum_{j=0}^{\ell-2} \binom{\ell-2}{j} \Big(\sum_{k=1}^{N} k^j\, \nu(R^{k-1}\zeta) \Big) \Phi_{\ell-2-j,N} \\ &\le \ \sum_{j=0}^{\ell-2} \binom{\ell-2}{j} \Big(\sum_{k=1}^{+\infty} k^j\, \nu(R^{k-1}\zeta) \Big) \Phi_{\ell-2-j,N} \end{aligned}$$

where the $\Phi_{i,N}$'s are defined in (131). Then, separating the term for $j = 0$ in the last sum and using $\zeta \le V_\ell \le V_{j+1}$ for $j = 1, \ldots, \ell-2$, it follows from (119b) that

$$B_N \le \Big(\sum_{k=1}^{+\infty} \nu(R^{k-1}\zeta) \Big) \Phi_{\ell-2,N} + \nu(V_0) \sum_{j=1}^{\ell-2} \binom{\ell-2}{j} D_j\, \Phi_{\ell-2-j,N}. \tag{136}$$

Third, recall that, for any $k \ge 1$, $\Sigma_k^{\ell-2} := \sum_{n=1}^{k} n^{\ell-2} = \sum_{j=1}^{\ell-1} a_{j,\ell}\, k^j$ from (122). Then

$$\begin{aligned} C_N &:= \pi_R(\psi) \Big(\sum_{n=0}^{N} (n+1)^{\ell-2} \Big| \sum_{k=n+1}^{+\infty} \nu(R^{k-1} g_n) \Big| \Big) 1_{\mathbb{X}} \\ &\le \pi_R(\psi) \Big(\sum_{n=0}^{+\infty} (n+1)^{\ell-2} \sum_{k=n+1}^{+\infty} \nu(R^{k-1}|g_n|) \Big) 1_{\mathbb{X}} \\ &\le \pi_R(\psi) \Big(\sum_{n=0}^{+\infty} (n+1)^{\ell-2} \sum_{k=n+1}^{+\infty} \nu(R^{k-1}V_\ell) \Big) 1_{\mathbb{X}} \ = \pi_R(\psi) \Big(\sum_{k=1}^{+\infty} \nu(R^{k-1}V_\ell) \sum_{n=1}^{k} n^{\ell-2} \Big) 1_{\mathbb{X}} \\ &\le \pi_R(\psi) \Big(\sum_{j=1}^{\ell-1} a_{j,\ell} \sum_{k=1}^{+\infty} k^j\, \nu(R^{k-1}V_\ell) \Big) 1_{\mathbb{X}} \\ &\le \pi_R(\psi) \nu(V_0) \Big(\sum_{j=1}^{\ell-1} a_{j,\ell} D_j \Big) 1_{\mathbb{X}} = \pi_R(\psi)\, \nu(V_0) E_{\ell-1} 1_{\mathbb{X}} \end{aligned} \tag{137}$$

using (119b) (note that $|g_n| \le V_\ell \le V_{j+1}$ for $j = 1, \ldots, \ell-1$) and the definition of $E_{\ell-1}$ in (123).

From the triangular inequality applied to (134), we obtain that

$$\sum_{n=0}^{N} (n+1)^{\ell-2} |P^n g_n| \le A_N + B_N + C_N.$$

Therefore Inequality (132) follows from (135)–(137). The proof of Lemma 8.4 is complete. □

8.3 Examples

8.3.1 Case $\psi := 1_S$

When $\psi := 1_S$ for some $S \in \mathcal{X}^*$, Condition $\boldsymbol{D}_{1_S}(V_0 : V_m)$ is: There exists a collection $\{V_i\}_{i=0}^m$ of Lyapunov functions with $m \geq 1$ such that

$$\forall i \in \{0, \ldots, m-1\}, \quad V_{i+1} \leq V_i \;\text{ and }\; \exists b_i > 0,\; PV_i \leq V_i - V_{i+1} + b_i\, 1_S. \qquad (\boldsymbol{D}_{1_S}(V_0 : V_m))$$

Proposition 8.5 *Let P satisfy Condition $(\boldsymbol{M}_{\nu,1_S})$ for some couple $(\nu, S) \in \mathcal{M}_{+,b}^* \times \mathcal{X}^*$. Then Condition* (121) *is equivalent to $\pi_R(1_S) > 1/2$. Moreover, under this condition, the constant $C_\phi := (1 - \mu_R(|\phi|))^{-1}$ involved in* (127) *and* (129) *is given by*

$$C_\phi = (2\pi_R(1_S) - 1)^{-1}.$$

Proof. If $\psi := 1_S$, then $\phi = 1_S - \pi_R(1_S)1_{\mathbb{X}}$, so that $|\phi| = \pi_R(1_{S^c})1_S + \pi_R(1_S)1_{S^c}$. Thus $\pi_R(|\phi|) = 2\pi_R(1_{S^c})\pi_R(1_S)$. Hence we have $\mu_R(|\phi|) = \pi_R(|\phi|)/\pi_R(1_S) = 2\pi_R(1_{S^c})$, from which the desired statements are easily deduced. □

Using Proposition 8.5, let us detail the conclusions of Theorem 8.2 under Conditions $(\boldsymbol{M}_{\nu,1_S})$–$\boldsymbol{D}_{1_S}(V_0 : V_m)$ with $(\nu, S) \in \mathcal{M}_{+,b}^* \times \mathcal{X}$ in the cases $m := 2$ and $m := 3$. Recall that the positive constants d_i (see (117)) are

$$\forall i \in \{0, \ldots, m-1\}, \quad d_i := \max\left(0, \frac{b_i - \nu(V_i)}{\nu(1_{\mathbb{X}})}\right)$$

with constants b_i given in $\boldsymbol{D}_{1_S}(V_0 : V_m)$. Moreover note that, in case $\psi := 1_S$, we have $\|\phi\|_{1_{\mathbb{X}}} \leq \max(\pi_R(1_S), 1 - \pi_R(1_S)) \leq 1$.

Case $m := 2$

Let P satisfy Condition $(\boldsymbol{M}_{\nu,1_S})$ with $\pi_R(1_S) > 1/2$ and Conditions $\boldsymbol{D}_{1_S}(V_0 : V_2)$ for some Lyapunov functions V_0, V_1, V_2. Note that $\Sigma_k^0 := k$, i.e. $a_{1,2} = 1$ in (122). Moreover we have $D_0 := 1 + d_0, D_1 := (1 + d_0)(1 + d_1)$ from (118) and $E_1 = D_1$ from (123). Consequently it follows from (124) and (126) applied with $m := 2$ that

$$W_2 = (1 + d_0)\, V_0 + \nu(V_0)\big[\, (1 + d_0)\, \Phi_0 + \pi(1_S)\, (1 + d_0)(1 + d_1)\, 1_{\mathbb{X}}\big] \qquad (138)$$

and we have the following estimate from (127) with $i := 0$:

$$\Phi_0 \;\leq\; \frac{(1 + d_0)}{2\pi(1_S) - 1}\, V_0 + \frac{\pi(1_S)\nu(V_0)(1 + d_0)(1 + d_1)}{2\pi(1_S) - 1}\, 1_{\mathbb{X}}.$$

It follows that $W_2 \leq c_0 V_0 + \pi(1_S)c_1 1_{\mathbb{X}} \leq c_0 V_0 + c_1 1_{\mathbb{X}}$ with the constants c_0, c_1 defined by

$$c_0 := (1 + d_0)\left(1 + \frac{\nu(V_0)\,(1 + d_0)}{2\pi_R(1_S) - 1}\right) \quad c_1 := \nu(V_0)(1 + d_0)(1 + d_1)\left(\frac{\nu(V_0)\,(1 + d_0)}{2\pi_R(1_S) - 1} + 1\right).$$

Apply (124) with $m := 2$ to get

$$\forall g \in \mathcal{B}_{V_2},\; \forall x \in \mathbb{X}, \quad \mathcal{S}_0(g, x) = \sum_{n=0}^{+\infty} \big|(P^n g)(x) - \pi_R(g)\big| \leq \|g - \pi_R(g)1_{\mathbb{X}}\|_{V_2}\, W_2(x)$$
$$\leq \|g - \pi_R(g)1_{\mathbb{X}}\|_{V_2}\, \widehat{c}_2\, V_0(x) \quad (139)$$

where $\widehat{c}_2 := c_0 + c_1\|1_{\mathbb{X}}\|_{V_0}$. Similarly Inequalities (125), where $\theta_2 \leq 1 + b_1\pi_R(\psi)$, and (130) hold with W_2 defined in (138).

Case $m := 3$

Let P satisfy Condition $(\boldsymbol{M}_{\nu,1_S})$ with $\pi_R(1_S) > 1/2$ and Conditions $\boldsymbol{D}_{1_S}(V_0 : V_3)$ for some Lyapunov functions V_0, V_1, V_2, V_3. Here we have $\Sigma_k^1 = k(k+1)/2$, i.e. $a_{1,3} = a_{2,3} = 1/2$ from (122). Thus we get from (118) and (123)

$$i = 0, 1,\ D_i = \prod_{j=0}^{i}(1 + d_i),\ D_2 = (1 + d_2)(D_0 + 2D_1),\ E_1 = D_1,\ E_2 = \frac{D_1 + D_2}{2}.$$

As in the Case $m := 2$, we get from (126) with $m := 3$

$$W_3 := D_1\, V_0 + \nu(V_0)\big[\, D_0\, \Phi_1 + D_1\, \Phi_0 + \pi(1_S)\, E_2\, 1_{\mathbb{X}}\big] \tag{140}$$

and from (127) with $i := 0, 1$

$$\Phi_0 \leq \frac{D_0}{2\pi(1_S) - 1} V_0 + \frac{\pi(1_S)\nu(V_0)E_1}{2\pi(1_S) - 1} 1_{\mathbb{X}},\ \Phi_1 \leq \frac{D_1}{2\pi(1_S) - 1} V_0 + \frac{\nu(V_0)\, D_1}{2\pi(1_S) - 1}\Phi_0 + \frac{\pi(1_S)\nu(V_0)E_2}{2\pi(1_S) - 1} 1_{\mathbb{X}}.$$

Thus, we obtain $W_3 \leq c_0 V_0 + c_1 1_{\mathbb{X}}$ where

$$c_0 := D_1\left[1 + \frac{\nu(V_0)D_0}{2\pi_R(1_S) - 1}\right]^2 c_1 := \nu(V_0)\left[E_2 + \frac{\nu(V_0){D_1}^2 + D_0E_2\nu(V_0)}{2\pi_R(1_S) - 1} + \frac{\nu(V_0)^2 D_0{D_1}^2}{(2\pi_R(1_S) - 1)^2}\right].$$

Consequently it follows from (124) with $m := 3$ that

$$\begin{aligned} \forall g \in \mathcal{B}_{V_3},\ \forall x \in \mathbb{X},\quad \sum_{n=0}^{+\infty}(n+1)\big|(P^n g)(x) - \pi_R(g)\big| &\leq \|g - \pi_R(g)1_{\mathbb{X}}\|_{V_3}\, W_3(x) \\ &\leq \|g - \pi_R(g)1_{\mathbb{X}}\|_{V_3}\, \widehat{c}_3\, V_0(x) \qquad (141) \\ &\quad \text{with } \widehat{c}_3 := c_0 + c_1\|1_{\mathbb{X}}\|_{V_0}. \qquad (142) \end{aligned}$$

Similarly Inequalities (125), where $\theta_3 \leq 1 + b_2\pi_R(\psi)$, and (130) hold with W_3 defined in (140).

8.3.2 Jarner-Roberts's drift conditions

Recall that Jarner-Roberts's drift condition is the following: There exists a Lyapunov function V such that

$$\exists S \in \mathcal{X}^*,\ \exists \alpha \in [0, 1),\ \exists b, c > 0,\quad PV \leq V - c\, V^{\alpha} + b\, 1_S. \tag{143}$$

This is the most classical one leading to nested modulated drift conditions $\boldsymbol{D}_{1_S}(V_0 : V_m)$. The details are recalled in the next proposition. Let $\lfloor\cdot\rfloor$ denote the integer part function on $\mathbb{R}$.

Proposition 8.6 *If P satisfies Condition* (143)*, then P satisfies $\boldsymbol{D}_{1_S}(V_0 : V_m)$ with Integer $m \equiv m(\alpha) := \lfloor(1-\alpha)^{-1}\rfloor \geq 1$ and the following Lyapunov functions*

$$V_m := 1_{\mathbb{X}} \leq V_{m-1} := a_{m-1}V^{\alpha_{m-1}} \leq \cdots \leq V_1 := a_1 V^{\alpha_1} \leq V_0 := a_0 V \tag{144}$$

with $\alpha_1 := 1 - 1/m \in [0, 1)$ and when $m \geq 2$

$$\forall i = 2, \ldots, m-1,\quad \alpha_i = (\alpha_1 - 1)\, i + 1,$$

where the a_i's are explicit constants strictly larger than one.

Assume that the set S in (143) is a first-order small-set (i.e. P satisfies Condition $(\boldsymbol{M}_{\nu,1_S})$) and that $\pi_R(1_S) > 1/2$, so that Condition (121) holds from Proposition 8.5. If $m(\alpha) \geq 2$ (i.e. $\alpha \geq 1/2$), then for any measurable and bounded function $g : \mathbb{X} \to \mathbb{R}$, i.e. $g \in \mathcal{B}_{1_{\mathbb{X}}}$, and for any $x \in \mathbb{X}$, Theorem 8.2 combined with Proposition 8.6 provides an explicit bound for $\sum_{n=0}^{+\infty}(n+1)^{m(\alpha)-2}|(P^n g)(x) - \pi_R(g)|$. For instance the bounds (139) in case $m(\alpha) := 2$ (i.e. $\alpha \in [1/2, 2/3)$), or the bounds (141) in case $m(\alpha) := 3$ (i.e. $\alpha \in [2/3, 3/4)$), apply.

Proof. The construction of the Lyapunov functions V_i is based on the following fact. If W is a measurable function from $\mathbb{X}$ to $[1, +\infty)$ and if $0 < \theta_2 < \theta_1 < 1$ are such that

$$\exists b, c > 0, \quad PW^{\theta_1} \leq W^{\theta_1} - c\, W^{\theta_2} + b\, 1_S,$$

$$\text{then} \quad \exists b', c' > 0, \quad PW^{\theta_2} \leq W^{\theta_2} - c'\, W^{\theta_3} + b'\, 1_S \qquad \text{with} \quad \theta_3 := 2\theta_2 - \theta_1. \tag{145}$$

Indeed we know from [JR02, Lem. 3.5] that

$$\forall \eta \in (0, 1],\ \exists b_\eta, c_\eta > 0, \quad PW^{\eta\theta_1} \leq W^{\eta\theta_1} - c_\eta\, (W^{\theta_1})^{\theta_2/\theta_1 + \eta - 1} + b_\eta 1_S.$$

Then (145) is obtained with $\eta := \theta_2/\theta_1 < 1$. Next note that $\alpha_1 = 1 - 1/m \leq \alpha$, so that

$$PV \leq V - c\, V^{\alpha_1} + b\, 1_S \tag{146}$$

from (143). Of course we can replace c with $c_1 < 1$. Recall that $m := \lfloor (1-\alpha)^{-1} \rfloor$. Then:

- If $\alpha_1 = 0$, i.e. $m = 1$ or $\alpha \in [0, 1/2)$, then $\boldsymbol{D}(V_0 : V_1)$ holds with $V_0 := c_1^{-1} V \geq V_1 := 1_{\mathbb{X}}$.
- If $\alpha_1 = 1/2$, i.e. $m = 2$ or $\alpha \in [1/2, 2/3)$, then we deduce from (146) and Property (145) applied to $W := V, \theta_1 = 1, \theta_2 = \alpha_1$ that

  $$\exists b_1, c_2 > 0, \quad PV^{\alpha_1} \leq V^{\alpha_1} - c_2\, V^{\alpha_2} + b_1 1_S \tag{147}$$

  with $\alpha_2 := 2\alpha_1 - 1 = 0$. Again note that we can choose $c_2 < 1$. Then the procedure stops, and Conditions $\boldsymbol{D}(V_0 : V_2)$ hold with $V_0 := c_1^{-1} c_2^{-1} V \geq V_1 := c_2^{-1} V^{\alpha_1} \geq V_2 := 1_{\mathbb{X}}$.
- If $\alpha_1 > 1/2$, then Property (145) can be used recursively to provide inequalities of the form $PV^{\alpha_{i-1}} \leq V^{\alpha_{i-1}} - c_i\, V^{\alpha_i} + b_{i-1} 1_S$ with $c_i < 1$ and $\alpha_i = 2\alpha_{i-1} - \alpha_{i-2} = (\alpha_1 - 1)\, i + 1$. Actually (145) can only be used until the value $i = m$ since $\alpha_m = 0$ and $\alpha_i < 0$ for $i > m$. Then Conditions $\boldsymbol{D}(V_0 : V_m)$ hold with V_i given in (144), where $a_i = [\prod_{k=i+1}^m c_k]^{-1}$.

□

8.3.3 Application to $V-$geometric rate of convergence

Assume that P satisfies Condition $(\boldsymbol{M}_{\nu,\psi})$ and the $V-$geometric drift condition $\boldsymbol{G}_\psi(\delta, V)$ (see Example 5.2). Then, for every $m \geq 1$, P satisfies Conditions $\boldsymbol{D}_\psi(V_0 : V_m)$ with

$$V_m := V \quad \text{and} \quad \forall i \in \{0, \ldots, m-1\}, \quad V_i = \frac{V}{(1-\delta)^{m-i}}, \quad b_i := \frac{b}{(1-\delta)^{m-i}}. \tag{148}$$

Recall that, if P satisfies $(\boldsymbol{M}_{\nu,1_S})$–$\boldsymbol{G}_{1_S}(\delta, V)$ and is aperiodic, then P is $V-$geometically ergodic from Theorem 6.2. In the next statement, strengthening the aperiodicity assumption with the condition $\pi(1_S) > 1/2$, we present a simpler proof of the $V-$geometric ergodicity of P, with moreover an explicit control of the rate of convergence.

Corollary 8.7 *Assume that P satisfies Conditions $(\boldsymbol{M}_{\nu,1_S})$–$\boldsymbol{G}_{1_S}(\delta, V)$ with $\pi_R(1_S) > 1/2$. Then P is $V-$geometrically ergodic and for every $\tau \in (0,1)$, we have*

$$\forall g \in \mathcal{B}_V,\ \forall n \geq 0, \quad \|P^n g - \pi_R(g)1_{\mathbb{X}}\|_V \leq \frac{\widehat{c}_3(1+\pi(V)\|1_{\mathbb{X}}\|_V)}{\tau\,(1-\delta)^3}\,\rho^n\,\|g\|_V \quad \textit{with} \quad \rho := \tau^{1/n_0}$$

where $\widehat{c}_3$ is provided in (142) *using $\boldsymbol{D}_{1_S}(V_0 : V_m)$ and V_i's given in* (148) *with $m := 3$, and n_0 is the smallest positive integer number such that $\widehat{c}_3(1+\pi(V)\|1_{\mathbb{X}}\|_V) \leq (n_0+1)\tau(1-\delta)^3$.*

Proof. Using here Condition $\boldsymbol{D}_{1_S}(V_0 : V_3)$ and V_0, V_1, V_2, V_3 given (148), it follows from (141) that

$$\forall n \geq 1,\ \forall g \in \mathcal{B}_V,\ \forall x \in \mathbb{X}, \quad \frac{|(P^n g)(x) - \pi_R(g)|}{V(x)} \leq \frac{c\,\|g\|_V}{n+1} \quad \text{with} \quad c := \frac{\widehat{c}_3(1+\pi(V)\|1_{\mathbb{X}}\|_V)}{(1-\delta)^3}.$$

Recall that the operator-norm of any bounded linear operator L on $(\mathcal{B}_V, \|\cdot\|_V)$ is defined by: $\|L\|_V := \sup\{\|Lg\|_V : g \in \mathcal{B}_V, \|g\|_V \leq 1\}$. From the above inequality we then obtain that $\|P^n - \Pi\|_V \leq c/(n+1)$ with $\Pi := 1_{\mathbb{X}} \otimes \pi_R$. Let $\tau \in (0,1)$ and $n_0 \equiv n_0(\tau)$ be the smallest positive integer such that $c/(n_0+1) \leq \tau$. Then, writing $n = qn_0 + r$ with $r \in \{0, \ldots, n_0-1\}$, we deduce that

$$\forall n \geq 1, \quad \|P^n - \Pi\|_V \leq \|(P-\Pi)^r\|_V \times \big(\|(P-\Pi)^{n_0}\|_V\big)^q \leq \frac{c}{\tau}\,\rho^n \quad \text{with} \quad \rho := \tau^{1/n_0}$$

since $\|(P-\Pi)^r\|_V \leq c$ and $\tau^{-r/n_0} \leq \tau^{-1}$. □

8.4 Complements on Condition (121) using some iterate of P

Let P satisfy Conditions $(\boldsymbol{M}_{\nu,\psi})$–$\boldsymbol{D}_\psi(V_0 : V_m)$ with $(\nu, \psi) \in \mathcal{M}^*_{+,b} \times \mathcal{B}^*_+$. Recall that Condition (121) is

$$\mu_R(|\phi|) = \frac{\pi_R(|\phi|)}{\pi_R(\psi)} < 1 \qquad \text{where } \phi \equiv \phi_\psi := \psi - \pi_R(\psi)1_{\mathbb{X}}.$$

When this condition does not hold, Theorem 8.2 and Corollary 8.3 may not be relevant: Indeed recall that Condition (121) ensures that Inequalities (127) for the Φ_i's hold, from which the V_0-weighted norm of W_m can be deduced (see (128) and Cases $m := 2, 3$ in Subsection 8.3.1). To overcome this problem, considering some iterate P^ℓ instead of P may be of interest. To that effect recall that the minorization condition of order ℓ for some integer $\ell \geq 1$ is the following one:

$$\exists(\nu_\ell, \psi_\ell) \in \mathcal{M}^*_{+,b} \times \mathcal{B}^*_+, \quad P^\ell \geq \psi_\ell \otimes \nu_\ell. \tag{149}$$

This condition is nothing else but Condition $(\boldsymbol{M}_{\nu_\ell,\psi_\ell})$ for the Markov kernel P^ℓ, and ψ_ℓ is called a $\ell-$order small function for P. The following statement is then obvious.

Corollary 8.8 *Let us assume that, for some $\ell \geq 1$, the Markov kernel P^ℓ satisfies Conditions $(\boldsymbol{M}_{\nu_\ell,\psi_\ell})$ and $\boldsymbol{D}_{\psi_\ell}(V_0 : V_m)$ for some $(\nu_\ell, \psi_\ell) \in \mathcal{M}^*_{+,b} \times \mathcal{B}^*_+$ and for some collection $\{V_i\}_{i=0}^m$ of Lyapunov functions. Then Theorem 8.2 and Corollary 8.3 apply to P^ℓ.*

Of course, for Corollary 8.8 to be relevant, Condition (121) for P^ℓ must be satisfied. This requires in particular being able to calculate the iterate P^ℓ, which is in any case also necessary to check that P^ℓ satisfies Conditions $(\boldsymbol{M}_{\nu_\ell,\psi_\ell})$ and $\boldsymbol{D}_{\psi_\ell}(V_0 : V_m)$. This problem is circumvented in the following statement, in which the minorization and nested modulated drift conditions are initially assumed for P. More precisely, let P satisfy Conditions $(\boldsymbol{M}_{\nu,\psi})$–$\boldsymbol{D}_\psi(V_0 : V_m)$ with $\mu_R(|\phi|) \geq 1$, so that Condition (121) does not hold. The next theorem states that, under the strong aperiodicity condition $\nu(\psi) > 0$, there exists an integer $\ell \geq 2$ such that the assumptions of Corollary 8.8 hold with some small-function ψ_ℓ for P^ℓ satisfying Condition (121), so that all the conclusions of Theorem 8.2 and Corollary 8.3 apply to P^ℓ.

Theorem 8.9 *Let P satisfy Conditions $(\boldsymbol{M}_{\nu,\psi})$–$\boldsymbol{D}_\psi(V_0 : V_m)$ with $m \geq 2$ for some $(\nu, \psi) \in \mathcal{M}^*_{+,b} \times \mathcal{B}^*_+$ such that $\nu(\psi) > 0$. Then*

1. *π_R is the unique $P^\ell-$invariant probability measure for each $\ell \geq 1$.*
2. *There exists an integer $\ell_0 \geq 1$ such that the following assertions hold for every $\ell \geq \ell_0$:*
   - *(a) the Markov kernel P^ℓ satisfies $(\boldsymbol{M}_{\nu,\psi_\ell})$–$\boldsymbol{D}_{\psi_\ell}(V_0 : V_m)$ with $\psi_\ell := P^{\ell-1}\psi$;*
   - *(b) Condition (121) holds for the small-function $\psi_\ell := P^{\ell-1}\psi$ and rewrites as*

$$\mu_R(|\phi_\ell|) := \frac{\pi_R(|\phi_\ell|)}{\pi_R(\psi_\ell)} < 1 \qquad \text{with } \phi_\ell := \psi_\ell - \pi_R(\psi_\ell)1_{\mathbb{X}}. \tag{150}$$

Using these properties, all the conclusions of Theorem 8.2 and Corollary 8.3 apply to the Markov kernel P^ℓ replacing the constant C_ϕ by $(1 - \mu_R(|\phi_\ell|))^{-1}$.

In (150) we have $\pi_R(\psi_\ell) = \pi_R(\psi)$ from the $P-$invariance of π_R. Thus (150) reads as follows

$$\mu_R(|\phi_\ell|) = \mu_R(|P^{\ell-1}\psi - \pi_R(\psi)1_{\mathbb{X}}|) = \frac{\pi_R(|P^{\ell-1}\psi - \pi_R(\psi)1_{\mathbb{X}}|)}{\pi_R(\psi)} < 1.$$

When $\ell := 1$, this is nothing else than $\mu_R(|\phi|) < 1$. When $\mu_R(|\phi|) \geq 1$, Theorem 8.9 states that the condition $\mu_R(|\phi_\ell|) < 1$ is fulfilled for ℓ large enough. The proof of Theorem 8.9 is based on the following lemma.

Lemma 8.10 *Let P satisfy Condition $(\boldsymbol{M}_{\nu,\psi})$ for some $(\nu, \psi) \in \mathcal{M}^*_{+,b} \times \mathcal{B}^*_+$ such that $\nu(\psi) > 0$, and Condition $\boldsymbol{D}_\psi(V, W)$ for some couple (V, W) of Lyapunov functions on $\mathbb{X}$. For every $\ell \geq 1$, set $\psi_\ell := P^{\ell-1}\psi$. Then P^ℓ satisfies Conditions $(\boldsymbol{M}_{\nu,\psi_\ell})$–$\boldsymbol{D}_{\psi_\ell}(V, W)$, that is*

$$P^\ell \;\geq\; \psi_\ell \otimes \nu \tag{151a}$$

$$\exists b_\ell > 0, \quad P^\ell V \;\leq\; V - W + b_\ell\, \psi_\ell. \tag{151b}$$

Proof. For $\ell := 1$, Inequalities (151a)–(151b) are just $(\boldsymbol{M}_{\nu,\psi})$–$\boldsymbol{D}_\psi(V, W)$. Let $\ell \geq 2$ be fixed. Inequality (151a) follows from $(\boldsymbol{M}_{\nu,\psi})$ and the non-negativity of the kernel $P^{\ell-1}$. To obtain (151b), let $V_d := V + d1_{\mathbb{X}}$ with $d := \max(0, (b - \nu(V))/\nu(1_{\mathbb{X}}))$ where b is the constant in $\boldsymbol{D}_\psi(V, W)$, so that P satisfies $\boldsymbol{R}(V_d, W)$ from Lemma 5.8 used under Condition $\boldsymbol{D}_\psi(V, W)$. Namely we have: $RV_d \leq V_d - W$. Iterating this inequality shows that $R^\ell V_d \leq V_d - W$. Then, under $(\boldsymbol{M}_{\nu,\psi})$, it follows from Formula (17) in Lemma 3.2 applied to V_d that

$$P^\ell V_d = R^\ell V_d + \sum_{k=1}^{\ell} \nu(R^{k-1}V_d)\, P^{\ell-k}\psi \leq V_d - W + \sum_{k=1}^{\ell} \nu(R^{k-1}V_d)\, P^{\ell-k}\psi$$

which is equivalent to

$$P^\ell V \le V - W + \sum_{k=1}^{\ell} \nu(R^{k-1}V_d)\, P^{\ell-k}\psi$$

since $P^\ell V_d = P^\ell V + d1_{\mathbb{X}}$ using the definition of V_d and P^ℓ is a Markov kernel. It remains to prove that the last term in the previous inequality is bounded from above by $b_\ell\,\psi_\ell = b_\ell P^{\ell-1}\psi$ for some $b_\ell > 0$. To do this, prove that, for any $k \in \{1,\dots,\ell\}$, $P^{\ell-k}\psi \le c_{\ell,k}P^{\ell-1}\psi$ for some $c_{\ell,k} > 0$. In fact, we have $P^{\ell-1}\psi \ge \nu(\psi)^{k-1}P^{\ell-k}\psi$ for every $k \in \{1,\dots,\ell\}$ so that we can set $c_k \equiv c_{\ell,k} := \nu(\psi)^{-(k-1)}$. Indeed, this is trivial for $k := 1$ and for $k \in \{2,\dots,\ell\}$, this follows from the minorization conditions (151a) applied to ψ, i.e. the relation $P^j\psi \ge \nu(\psi)P^{j-1}\psi$ for $j \ge 1$. This proves the desired inequality. □

Proof of Theorem 8.9. We know from Theorem 5.3 that $\mu_R(1_{\mathbb{X}}) < \infty$ and that the function $h_R^\infty := \lim_n R^n 1_{\mathbb{X}}$ is zero on $\mathbb{X}$. Moreover it follows from the strong aperiodicity condition $\nu(\psi) > 0$ and Theorem 4.7 that $\lim_{n\to+\infty} \|\delta_x P^n - \pi_R\|_{\mathrm{TV}} = 0$, from which we deduce the two following facts: first π_R is the unique invariant probability measure for each iterate P^ℓ; second there exists $\ell_0 \ge 1$ such that

$$\mu_R(|P^{\ell_0-1}\psi - \pi_R(\psi)1_{\mathbb{X}}|) = \frac{\pi_R(|P^{\ell_0-1}\psi - \pi_R(\psi)1_{\mathbb{X}}|)}{\pi_R(\psi)} < 1$$

since Lebesgue's theorem w.r.t. π_R ensures that $\lim_n \pi_R(|P^n\psi - \pi_R(\psi)1_{\mathbb{X}}|) = 0$. As previously quoted, this condition is nothing else than Property (150) for $\ell := \ell_0$, that is $\mu_R(|\phi_{\ell_0}|) < 1$. Let $\ell \ge \ell_0$. Using Conditions $(\boldsymbol{M}_{\nu,\psi})$–$\boldsymbol{D}_\psi(V_0 : V_m)$ for P we deduce from Lemma 8.10 that P^ℓ satisfies $(\boldsymbol{M}_{\nu,\psi_\ell})$–$\boldsymbol{D}_{\psi_\ell}(V_0 : V_m)$ with $\psi_\ell := P^{\ell-1}\psi$. Moreover Theorem 5.3 applied to P^ℓ and the fact that π_R is the unique $P^\ell-$invariant probability measure ensure that

$$\pi_R = \pi_{R_\ell} := \mu_{R_\ell}(1_{\mathbb{X}})^{-1}\mu_{R_\ell} = \pi_R(\psi_\ell)\mu_{R_\ell} = \pi_R(\psi)\mu_{R_\ell} \quad \text{with} \quad \mu_{R_\ell} := \sum_{k=1}^{+\infty} \nu R_\ell^{k-1} \in \mathcal{M}_*^+$$

where $R_\ell := P^\ell - \psi_\ell \otimes \nu$ is the residual kernel associated with P^ℓ under the minorization condition $(\boldsymbol{M}_{\nu,\psi_\ell})$. In particular note that $\mu_{R_\ell} = \mu_R$. Consequently Condition (121) for P^ℓ under $(\boldsymbol{M}_{\nu,\psi_\ell})$–$\boldsymbol{D}_{\psi_\ell}(V_0 : V_m)$ writes as

$$\mu_{R_\ell}(|\phi_\ell|) = \mu_R(|\phi_\ell|) = \frac{\pi_R(|\phi_\ell|)}{\pi_R(\psi_\ell)} < 1 \quad \text{where} \quad \phi_\ell := \psi_\ell - \pi_R(\psi_\ell)1_{\mathbb{X}},$$

which is exactly (150). □

The proof of Lemma 8.10 ensures that

$$P^\ell V \le V - W + b_\ell\,\psi_\ell \quad \text{with} \quad b_\ell := \sum_{k=1}^{\ell} c_{\ell,k}\,\nu(R^{k-1}V_d) \tag{152}$$

where $V_d := V + d1_{\mathbb{X}}$ with $d := \max(0, \nu(1_{\mathbb{X}})^{-1}(b - \nu(V)))$ and where $c_{\ell,k}$ are any positive constant such that we have: $\forall k \in \{1,\dots,\ell\},\ P^{\ell-k}\psi \le c_{\ell,k}P^{\ell-1}\psi$. Such constants $c_{\ell,k}$ exist, for instance $c_{\ell,k} = \nu(\psi)^{-(k-1)}$ (see the proof of Lemma 8.10). If P^ℓ is computable, then the constant b_ℓ in (151b) can be computed using directly the Markov kernel P^ℓ rather than using formula in (152). Actually the general minorization and nested modulated drift conditions for P^ℓ in Corollary 8.8 are the following ones with $\psi_\ell := P^{\ell-1}\psi$:

$$\exists\, \nu_\ell \in \mathcal{M}_{+,b}^*,\ \forall x \in \mathbb{X}, \quad P^\ell \geq P^{\ell-1}\psi \otimes \nu_\ell$$
$$\forall i \in \{0,\dots,m-1\},\ \exists b_{\ell,i} > 0, \quad P^\ell V_i \leq V_i - V_{i+1} + b_{\ell,i}\, P^{\ell-1}\psi.$$

Under the assumptions of Theorem 8.9, the two previous conditions are satisfied for every $\ell \geq 1$, and Property (150) holds for ℓ large enough.

8.5 Further comments and bibliographic discussion

Theorem 8.2 and Corollary 8.3 may be relevant whenever explicit modulated drift conditions are known: for such examples stated with $\psi := 1_S$, e.g. see [FM00, FM03b, DFM16] in the context of Metropolis algorithm, [LH07, LH12] for queueing systems, [JT02] for Markov chains associated with the mean of Dirichlet processes. The main classical results on the rate of convergence of iterates in the non-geometric case are now recalled. The condition $\pi(1_S) > 1/2$ is also discussed at the end of this subsection.

A) *On the subgeometric convergence rates.* In continuation of the pioneering works [Pit74] based on recurrence times moments and [NT83, TT94] using spitting techniques, the polynomial rates of convergence were first addressed in [JR02] under the nested modulated drift conditions $\boldsymbol{D}_{1_S}(V_0 : V_m)$ w.r.t. petite sets, of which Jarner-Roberts's drift condition (143) is a special case (see Subsection 3.5-A for the definition of a petite set). Then explicit bounds for $\|P^n(x,\cdot) - \pi\|_{\mathrm{TV}}$ have been proposed in [FM03b, DMS07] thanks to coupling methods under the sub-geometric drift condition $PV \leq V - \phi \circ V + b1_S$ (where ϕ is essentially some non-negative concave increasing differentiable function on $[1,+\infty)$). This so-called sub-geometric drift condition encompasses Jarner-Roberts's. Actually, whatever the form of the starting single drift condition (e.g. the sub-geometric one or Jarner-Roberts's), the nested modulated drift conditions $\boldsymbol{D}_\psi(V_0 : V_m)$ must be implemented in practice anyway, see [FM03b, Rem. 3]. This is recalled for Jarner-Roberts's drift condition in Proposition 8.6 based on [JR02, proof of Th. 3.6]. Readers can also consult [DFMS04] for various statements and examples on different rates of convergence, [But14, DFM16] for rates of convergence in Wasserstein distance, and finally [DMPS18, Sec. 17.3] for further bibliographical complements. An alternative operator-type approach is presented in [Del17] inspired by the work of Yosida-Kakutani [YK41]. The works [JR02, AFV15] are discussed below.

B) *Comments on Jarner-Roberts's paper* [JR02]. Let P satisfy Conditions $(\boldsymbol{M}_{\nu,1_S})$ and $\boldsymbol{D}_{1_S}(V_0 : V_m)$ with $m \geq 1$ and some petite set S_i in each modulated drift condition. Let π denote the $P-$invariant probability measure. It is proved in [JR02, Th. 3.2] that

$$\forall x \in \mathbb{X}, \quad \lim_{n\to+\infty} (n+1)^{m-1} \big\|P^n(x,\cdot) - \pi\big\|'_{V_m} = 0, \tag{154}$$

provided that P also satisfies the standard $\eta-$irreducibility and aperiodicity assumptions w.r.t. some positive measure η (also see [FM03b, Th. 1]). The polynomial asymptotics (154) ensures that $\|P^n(x,\cdot) - \pi\|_{\mathrm{TV}} \leq c(x)/n^{m-1}$ for every $x \in \mathbb{X}$, but with unknown constant $c(x)$. In particular the explicit bounds of $\|P^k(x,\cdot)-\pi\|_{\mathrm{TV}}$ in [FM03b, Th. 2] and [DMS07, Th. 2.1] do not seem to provide any information on the quantitative polynomial rate of convergence in (154) (see also [CF09, Pro. 2.2]). Estimate (130) in Corollary 8.3 provides $c(x) = cV_0(x)$ with an explicit constant c.

C) *On the results in* [AFV15] *under Jarner-Roberts's drift condition.* Using a coupling construction in the context of subgeometric Markov chains, rates of convergence are addressed in [AFV15, Th. 1] in terms of series of the form $\sum_{n=0}^{+\infty} r(n)|(P^n g)(x) - (P^n g)(x')|$ where $(r(n))_{n\geq 0}$ is some sequence of positive real numbers related to a subgeometric drift condition. In particular their results apply to Markov kernels satisfying $(\boldsymbol{M}_{\nu,1_S})$ and Jarner-Roberts's drift condition (143). Under these conditions (and some additional minor assumptions), it is proved in [AFV15, Cor. 1, homogeneous case with $\xi = 1$] that there exists a constant $C > 0$ such that for any $(x, x') \in \mathbb{X}^2$ and any $g \in \mathcal{B}_{1_{\mathbb{X}}}$

$$\sum_{n=0}^{+\infty} (n+1)^{m-1} |(P^n g)(x) - (P^n g)(x')| \leq C \, \|g\|_{1_{\mathbb{X}}} \big(V(x) + V(x') - 1\big)$$

with $m := \lfloor (1-\alpha)^{-1}$, where α is the constant in (143) (as in Proposition 8.6). Thus, if $\pi(V) < \infty$, then $\mathcal{S}_{m-1}(g,x) \leq C \, \|g\|_{1_{\mathbb{X}}} (V(x) + \pi(V) - 1)$. The fact that $\mathcal{S}_{m-1}(g,x)$ can be estimated in [AFV15, Cor. 1], while Theorem 8.2 only provides an estimate for $\mathcal{S}_{m-2}(g,x)$, is due to the additional condition $\pi(V) < \infty$ which is imposed in [AFV15, Cor. 1], but not in Theorem 8.2. Indeed recall that the only moment condition guaranteed under Assumption (143) is $\pi(V^\alpha) < \infty$, and that the assumption $\pi(V) < \infty$ actually generates an additional modulated drift condition (see Proposition 5.12).

D) *Comments on* [MT09, Th. 14.0.1]. If P satisfies the assumptions of Theorem 8.2 with $m := 2$ (thus requiring two nested modulated drift conditions), then

$$\forall x \in \mathbb{X}, \quad \sum_{n=0}^{+\infty} \big\| P^n(x,\cdot) - \pi \big\|'_{V_2} \leq \widehat{c}_2 \; V_0(x)$$

with explicit constant $\widehat{c}_2$ (see (139)). This statement may be surprising on first reading compared with the classical result [MT09, Th. 14.0.1]. Indeed, we know from [MT09, Th. 14.0.1] that, if P satisfies Condition $(\boldsymbol{M}_{\nu,1_S})$ and the single modulated drift condition $\boldsymbol{D}_{1_S}(V,W)$ for some Lyapunov functions V and W such that $\pi(V) < \infty$, then there exist a $P-$absorbing set $A \in \mathbb{X}$ and a (non-explicit) constant $c > 0$ such that

$$\forall x \in A, \quad \sum_{n=0}^{+\infty} \big\| P^n(x,\cdot) - \pi \big\|'_{W} \leq c \; V(x) \tag{155}$$

provided that P is irreducible and aperiodic. But again, note that the assumption $\pi(V) < \infty$ in [MT09, Th. 14.0.1] for Lyapunov function V in Condition $\boldsymbol{D}_{1_S}(V,W)$ generates another modulated drift condition $\boldsymbol{D}_{1_S}(L,V)$ for some Lyapunov function $L \geq V$ (see Proposition 5.12). Hence the assumptions of [MT09, Th. 14.0.1] actually involve two nested modulated drift conditions too.

E) *On the constant c in* (155). Using some refinements on the modulated drift condition, the authors in [FM03a, Prop. 13] present an explicit bound in [MT09, Th. 14.0.1], i.e. an explicit constant c in (155) (consider $\lambda = \delta_x$ and $\mu = \pi$ in [FM03a, Prop. 13]). These statements imply that the inequality $\pi(1_D) > 1/2$ for some small-set D holds: This inequality is nothing else than Condition (121) from Proposition 8.5. Indeed the Lyapunov function W in the modulated drift condition $\boldsymbol{D}_{1_S}(V,W)$ considered in [FM03a, Prop. 13] satisfies $W \geq b/(1-a)$ on D^c for some $a \in (0,1)$ and some small-set $D \in \mathcal{X}^*$ containing the small-set S of $\boldsymbol{D}_{1_S}(V,W)$, and where b denotes here the constant in $\boldsymbol{D}_{1_S}(V,W)$. Thus we have $\pi(1_{D^c}) \leq \pi(W)(1-a)/b$. Since the condition $\pi(V) < \infty$ is required in [FM03a, Prop. 13] for obtaining the bound (155), it follows from $\boldsymbol{D}_{1_S}(V,W)$ that $\pi(W) \leq b\pi(1_S)$. Thus we have $\pi(1_{D^c}) \leq (1-a)\pi(1_S)$, from which we deduce that

$$\pi(1_D) \geq \pi(1_S) \geq \frac{\pi(1_{D^c})}{1-a} = \frac{1-\pi(1_D)}{1-a}.$$

Hence we obtain that $\pi(1_D) \geq 1/(2-a)$. Thus, as claimed, the condition $\pi(1_D) > 1/2$ is assumed in [FM03a, Prop. 13].

F) *Again on Condition* (121). As in [FM03a] (see the previous point), the assumption $\pi(1_D) > 1/2$ for some small-set D occurs in the nested modulated drift conditions in [FM03b, p. 78] introduced for the study of polynomial ergodicity (see [FM03b, Eq. (50)] and apply the arguments of the previous point). Similarly the assumption $\pi(1_S) > 1/2$ is present in the geometric rate of convergence obtained in [Ros95, Th. 12], see [Jer16] and [QH21, Prop. 17]. In fact, technical conditions linking the set S to the data V, δ and b of Condition $\boldsymbol{G}_{1_S}(\delta,V)$ always lead to impose a restrictive assumption on $\pi(1_S)$. Such a technical condition is assumed on P in [HM11], and the extension to the general case requires the use of averaged Markov kernels $(\sum_{k=0}^{N} P^k)/(N+1)$ for N large enough. Finally, let us mention that conditions on S, V, δ and b occur in reversibility case too. Indeed, considering a reversible Markov kernel P w.r.t. invariant probability measure π, the authors in [TM22, Prop. 1] provide a $\mathbb{L}^2(\pi)-$rate of convergence when P satisfies Conditions $(\boldsymbol{M}_{\nu,1_S})$–$\boldsymbol{G}_{1_S}(\delta,V)$ with a small-set S satisfying the condition of [HM11], that is $S = \{V \leq s\}$ for $s > 2b/(1-\delta)$. This condition actually implies that $\pi(1_S) > 2/3$. Indeed

$$\pi(1_{S^c}) = \pi(1_{\{V>s\}}) \leq \frac{\pi(V)}{s} \leq \frac{b\pi(1_S)}{s(1-\delta)} < \frac{\pi(1_S)}{2}$$

from Markov inequality and $(1-\delta)\pi(V) \leq b\pi(1_S)$ derived from $\boldsymbol{G}_{1_S}(\delta,V)$ and the $P-$invariance of π. Thus $\pi(1_S) > 2/3$. Accordingly the discussion in [QH21, QH22] concerning the trade-off that must be made in [Ros95, Th. 12] between, on the one hand, the condition $\pi(1_S) > 1/2$ requiring a sufficiently large small-set S and, on the other hand, the total mass $\nu(1_{\mathbb{X}})$ requiring S not to be too large, generalizes to all the papers cited above and also to the framework of Theorem 8.2. To overcome this problem in the geometric case, the authors of [YR23] have introduced the notion of large-sets and a generalized geometric drift condition, see in particular [YR23, Th. 2.6] where a basic part of the proof is a modification of arguments used in [Ros95]. In the polynomial case, Theorem 8.9 provides an alternative, using a $\ell-$order small-function with $\ell \geq 2$.

9 Geometric rate of convergence of the iterates

In Subsection 9.1 the geometric rate of convergence of the iterates of P is studied on some general Banach space $\mathfrak{B}$ by introducing the spectral radius of the residual kernel R on $\mathfrak{B}$. This general framework is then applied under the first-order minorization condition $(\boldsymbol{M}_{\nu,\psi})$ and the geometric drift condition $\boldsymbol{G}_\psi(\delta, V)$ to obtain the rate of convergence, first for $V-$weighted norm in Subsection 9.2 to complete Theorem 6.2, second for $\mathbb{L}^2(\pi_R)-$norm in Subsection 9.3 with the specific reversible case in Subsection 9.4, and finally for $V^\alpha-$weighted norm in Subsection 9.5 for α belonging to some interval $\mathcal{A} \subset (0,1]$. Further statements on the reversible and positive reversible cases are provided in Subsection 9.6. The spaces $\mathcal{L}^1(\pi_R)$ and $\mathcal{L}^2(\pi_R)$, as well as the standard Lebesgue spaces $(\mathbb{L}^1(\pi_R), \|\cdot\|_1)$, $(\mathbb{L}^2(\pi_R), \|\cdot\|_2)$ and $(\mathbb{L}^\infty(\pi_R), \|\cdot\|_\infty)$ w.r.t. the probability measure π_R, are defined in Section 2. Finally, when L is a bounded linear operator on a Banach space $\mathfrak{B}$, we write $L \in \mathcal{L}(\mathfrak{B})$ for short. If $\mathfrak{B}$ is composed of complex-valued measurable functions on $\mathbb{X}$, or of classes modulo π_R of such functions, then for any non-negative kernel K on $\mathbb{X}$ we simply write $K \in \mathcal{L}(\mathfrak{B})$ to express that the functional action of K on $\mathfrak{B}$ defines a bounded linear operator on $\mathfrak{B}$. The prerequisites in spectral theory are those given by (S1)-(S3) in Subsection 6.2 (see page 57).

9.1 Geometric rate of convergence on a Banach space

Let P satisfy Condition $(\boldsymbol{M}_{\nu,\psi})$, as well as the two conditions $h_R^\infty = 0$ and $\mu_R(1_{\mathbb{X}}) < \infty$ which are satisfied for example under the modulated drift condition $\boldsymbol{D}_\psi(V_0, V_1)$. Under these conditions, all the conclusions of Theorem 4.1 hold true: The $P-$harmonic functions are constant on $\mathbb{X}$; P is irreducible and recurrent; The positive measure μ_R satisfies $\mu_R(\psi) = 1$ and is the unique $P-$invariant positive measure η (up to a positive multiplicative constant) such that $\eta(\psi) < \infty$; Finally $\pi_R := \mu_R(1_{\mathbb{X}})^{-1}\mu_R$ (see (26)) is the unique $P-$invariant probability measure on $(\mathbb{X}, \mathcal{X})$. Let $(\mathfrak{B}, \|\cdot\|)$ be a Banach space satisfying the following assumptions:

Assumptions (B). *Either the set $\mathfrak{B}$ is composed of $\mathbb{C}$-valued measurable functions on $\mathbb{X}$ and $\mathcal{B}_{1_{\mathbb{X}}} \subset \mathfrak{B} \subset \mathcal{L}^1(\pi_R)$; or $\mathfrak{B}$ is composed of classes modulo π_R of $\mathbb{C}$-valued measurable functions on $\mathbb{X}$ and $\mathbb{L}^\infty(\pi_R) \subset \mathfrak{B} \subset \mathbb{L}^1(\pi_R)$. Moreover, in both cases, the norm $\|\cdot\|$ on $\mathfrak{B}$ is assumed to satisfy the following condition:*

$$\exists c > 0,\ \forall g \in \mathfrak{B}, \quad \pi_R(|g|) \le c\|g\|. \tag{156}$$

If $P \in \mathcal{L}(\mathfrak{B})$, then P is said to be geometrically ergodic on $(\mathfrak{B}, \|\cdot\|)$ if

$$\exists \rho \in (0,1),\ \exists c_\rho > 0,\ \forall g \in \mathfrak{B},\ \forall n \ge 1, \quad \|P^n g - \pi_R(g)1_{\mathbb{X}}\| \le c_\rho\, \rho^n \|g\|. \tag{157}$$

In this case we define the following real number $\varrho_{\mathfrak{B}} \in (0,1)$

$$\varrho_{\mathfrak{B}} \equiv \varrho_{\mathfrak{B}}(P) := \inf\big\{\rho \in (0,1) \text{ such that Property (157) holds}\big\}. \tag{158}$$

The power series $\rho(z)$ used below is that introduced to define the aperiodicity condition (see (38)-(39)). Finally, when the residual kernel R belongs to $\mathcal{L}(\mathfrak{B})$, we denote by $r_{\mathfrak{B}}$ the spectral radius of R on $(\mathfrak{B}, \|\cdot\|)$.

Theorem 9.1 *Assume that P satisfies $(\boldsymbol{M}_{\nu,\psi})$ with $h_R^\infty = 0$, $\mu_R(1_{\mathbb{X}}) < \infty$, and is aperiodic. Let $(\mathfrak{B}, \|\cdot\|)$ be a Banach space satisfying Assumptions* **(B)** *and assume that $P \in \mathcal{L}(\mathfrak{B})$. Then $R \in \mathcal{L}(\mathfrak{B})$. Moreover, if $r_{\mathfrak{B}} < 1$, then P is geometrically ergodic on $(\mathfrak{B}, \|\cdot\|)$. More precisely the radius of convergence of the power series $\rho(z) := \sum_{n=1}^{+\infty} \nu(R^{n-1}\psi)\, z^n$ is larger than $1/r_{\mathfrak{B}}$, and the following alternative holds:*

(a) If Equation $\rho(z^{-1}) = 1$ has no solution $z \in \mathbb{C}$ such that $r_{\mathfrak{B}} < |z| < 1$, then $\varrho_{\mathfrak{B}} \le r_{\mathfrak{B}}$.

(b) Otherwise, we have $\varrho_{\mathfrak{B}} = \max\big\{|z| : z \in \mathbb{C},\ \rho(z^{-1}) = 1,\ r_{\mathfrak{B}} < |z| < 1\big\}$.

Based on the definition of the spectral radius $r_{\mathfrak{B}}$ of R on $\mathfrak{B}$, the following simple lemma is the first key point to prove Theorem 9.1.

Lemma 9.2 *Let us assume that P satisfies Condition $(\boldsymbol{M}_{\nu,\psi})$ with $h_R^\infty = 0$ and $\mu_R(1_{\mathbb{X}}) < \infty$, and that $P \in \mathcal{L}(\mathfrak{B})$ where $(\mathfrak{B}, \|\cdot\|)$ is a Banach space satisfying Assumptions* **(B)**. *Then $R \in \mathcal{L}(\mathfrak{B})$, and the following assertions hold:*

1. *For every $z \in \mathbb{C}$ such that $|z| > r_{\mathfrak{B}}$ and for every $g \in \mathfrak{B}$, the series $\widetilde{g}_z := \sum_{k=0}^{+\infty} z^{-(k+1)} R^k g$ absolutely converges in $\mathfrak{B}$.*
2. *The radius of convergence of $\rho(z) := \sum_{n=1}^{+\infty} \nu(R^{n-1}\psi)\, z^n$ is larger than $1/r_{\mathfrak{B}}$.*

Proof. Recall that the operator-norm of any $L \in \mathcal{L}(\mathfrak{B})$ is denoted by $\|L\|$ for simplicity. From $(\boldsymbol{M}_{\nu,\psi})$ and the $P-$invariance of π_R we know that $\pi_R \geq \pi_R(\psi)\nu$ with $\pi_R(\psi) > 0$ (see Theorem 3.6). Thus

$$\forall g \in \mathfrak{B}, \quad \nu(|g|) \leq \pi_R(\psi)^{-1}\pi_R(|g|) \leq c\, \pi_R(\psi)^{-1}\|g\| \tag{159}$$

due to (156). From the definition of R and (159), we obtain that, for every $g \in \mathfrak{B}$, the function Rg (or its class modulo π_R) belongs to $\mathfrak{B}$ with

$$\|Rg\| \leq \|Pg\| + \nu(|g|)\|\psi\| \leq \big(\|P\| + c\, \pi_R(\psi)^{-1}\|\psi\|\big)\|g\|$$

where $\|P\|$ is the operator-norm of P. Note that $\|\psi\|$ is well-defined since ψ is bounded, so that ψ (or its class) belongs to $\mathfrak{B}$. Thus $R \in \mathcal{L}(\mathfrak{B})$. Now prove Assertion 1. From the definition of $r_{\mathcal{B}}$ we know that

$$\forall \gamma \in (r_{\mathfrak{B}}, +\infty),\ \exists c_\gamma > 0,\ \forall g \in \mathfrak{B},\ \forall n \geq 1, \quad \|R^n g\| \leq c_\gamma\, \gamma^n\, \|g\|. \tag{160}$$

Let $z \in \mathbb{C}$ be such that $|z| > r_{\mathfrak{B}}$ and let $\gamma \in (r_{\mathfrak{B}}, |z|)$. Then for every $g \in \mathfrak{B}$ we have

$$|z|^{-(k+1)}\|R^k g\| \leq |z|^{-1} c_\gamma\, (\gamma/|z|)^k\, \|g\|,$$

from which we deduce that $\sum_{k=0}^{+\infty} |z|^{-(k+1)}\|R^k g\| < \infty$. This provides Assertion 1. since $(\mathfrak{B}, \|\cdot\|)$ is a Banach space. Now prove Assertion 2.. Let $\gamma > r_{\mathfrak{B}}$. From (159) and (160) we obtain that

$$0 \leq \nu(R^k\psi) \leq c\, \pi_R(\psi)^{-1}\|R^k\psi\| \leq c\, \pi_R(\psi)^{-1}\, c_\gamma\, \gamma^k\, \|\psi\|$$

so that the series $\sum_{n=1}^{+\infty} \nu(R^{n-1}\psi)\, z^n$ converges for every $z \in \mathbb{C}$ such that $|z| < 1/\gamma$. Hence the radius of convergence of the power series $\rho(z)$ is larger than $1/\gamma$, thus larger than $1/r_{\mathfrak{B}}$ since γ is any real number in $(r_{\mathfrak{B}}, +\infty)$. □

Recall that, in case $\mathfrak{B} := \mathcal{B}_V(\mathbb{C})$, the series involved in Lemma 9.2 are those used in Section 6.2 to study the invertibility of the operator $zI - P$ for $z \in \mathbb{C}$ of modulus one, see Lemmas 6.3-6.4. From these lemmas and the compactness of the spectrum, the geometric ergodicity on $\mathcal{B}_V(\mathbb{C})$ was then easily deduced in Theorem 6.2, i.e. $\varrho_{\mathfrak{B}} < 1$, but without control of the rate of convergence because Lemmas 6.3-6.4 only focus on the complex numbers of modulus one. Using Lemma 9.2 and repeating on the general space $\mathfrak{B}$ the arguments of Section 6.2, the proof of Theorem 9.1 as a whole is therefore a refinement, often even a simple copy, of that of Theorem 6.2. Indeed it can be similarly shown that, for any $z \in \mathbb{C}$ such that $|z| > r_{\mathfrak{B}}$, the operator $zI - P$ is invertible on $\mathfrak{B}$ if, and only if, $\rho(z^{-1}) \neq 1$. Then the alternative (a)-(b) of Theorem 9.1 is obtained noticing that $\varrho_{\mathfrak{B}}$ is nothing else but the spectral radius of the restriction P_0 of P to the subspace $\mathfrak{B}_0 := \{g \in \mathfrak{B} : \pi_R(g) = 0\}$ of $\mathfrak{B}$. For the reader's convenience, the proof of Theorem 9.1 is postponed to Appendix D, where the following additional statements are also obtained in Case (b) of Theorem 9.1.

Proposition 9.3 *Let P satisfy the assumptions of Theorem 9.1 with $r_{\mathcal{B}} < 1$. Then the following properties hold in Case (b) of Theorem 9.1. For every $r \in (r_{\mathfrak{B}}, 1)$ the set*

$$\mathcal{S}_r := \{z \in \mathbb{C},\ \rho(z^{-1}) = 1,\ r \leq |z| < 1\}$$

is finite, and it is non-empty for $r \in (r_{\mathfrak{B}}, 1)$ sufficiently close to $r_{\mathfrak{B}}$. Moreover every $z \in \mathcal{S}_r$ is an eigenvalue of P on $\mathfrak{B}$ with

$$E_z := \{g \in \mathfrak{B} : Pg = zg\} = \mathbb{C} \cdot \widetilde{\psi}_z$$

where $\widetilde{\psi}_z \in \mathfrak{B}$ is non-zero and is defined by $\widetilde{\psi}_z := \sum_{k=0}^{+\infty} z^{-(k+1)} R^k \psi$.

9.2 Rate of convergence in $V-$geometric ergodicity

When P is $V-$geometrically ergodic for some Lyapunov function V (see (72) in Theorem 6.2), we define the following real number $\varrho_V \in (0, 1)$

$$\varrho_V \equiv \varrho_V(P) := \inf\big\{\rho \in (0, 1) \text{ such that Property (72) holds}\big\}. \tag{161}$$

In other words ϱ_V is nothing else but $\varrho_{\mathfrak{B}}$ with $\mathfrak{B} := \mathcal{B}_V(\mathbb{C})$. To apply Theorem 9.1 in the case $\mathfrak{B} := \mathcal{B}_V(\mathbb{C})$, we first prove the following statement, in which r_V denotes for short the spectral radius of the residual kernel R on $\mathcal{B}_V(\mathbb{C})$ (i.e. $r_V \equiv r_{\mathcal{B}_V(\mathbb{C})}$ with the notation of Theorem 9.1).

Proposition 9.4 *Let P satisfy $(\boldsymbol{M}_{\nu,\psi})$–$\boldsymbol{G}_\psi(\delta, V)$. Then $r_V := \lim_n \|R^n\|_V^{1/n}$ satisfies*

$$\forall n \geq 1, \quad r_V \leq \|R^n\|_V^{1/n} = \|R^n V\|_V^{1/n} \tag{162}$$

and

$$r_V = \lim_n \|R^n V\|_V^{1/n} < 1.$$

Proof. Under $(\boldsymbol{M}_{\nu,\psi})$, the operator-norm $\|R^n\|_V$ equals to $\|R^n V\|_V$ from the non-negativity of R^n, and

$$\forall n \geq 1, \quad r_V = \lim_{k \to +\infty} \|R^{kn}\|_V^{\frac{1}{kn}} \leq \|R^n\|_V^{\frac{1}{n}}$$

from Gelfand's formula and $\|R^{kn}\|_V \leq \|R^n\|_V^k$. Then (162) holds true. Now, using Condition $\boldsymbol{D}_\psi(V_0 : V_2)$ with $V_2 := V$ and V_0, V_1 given in (148) (case $m = 2$), it follows from Proposition 8.1 that $\sum_{n=0}^{+\infty} (n+1)\, R^n V_2 \leq D_1\, V_0$, thus:

$$\forall n \geq 1, \quad R^n V = R^n V_2 \leq \frac{D_1}{n+1}\, V_0 = \frac{D_1}{(1-\delta)^2 (n+1)}\, V$$

since $V_0 = V/(1-\delta)^2$. From $\|R^n\|_V = \|R^n V\|_V$, there exists $n_0 \geq 1$ such that $\|R^{n_0}\|_V \leq 2/3$ (for instance). Next, writing $n = kn_0 + r$ with $r \in \{0, \ldots, n_0 - 1\}$ and setting $\theta := (2/3)^{1/n_0}$, $A := \max(1, \|R\|_V)^{n_0}$, we obtain that

$$\|R^n\|_V = \|R^r (R^{n_0})^k\|_V \leq \|R\|_V^r \left(\frac{2}{3}\right)^k \leq \max(1, \|R\|_V)^{n_0} \left(\frac{2}{3}\right)^{(n-r)/n_0} \leq \frac{3\,A}{2}\,\theta^n$$

thus $r_V \leq \theta < 1$ using Gelfand's formula. □

Under Conditions $(\boldsymbol{M}_{\nu,\psi})$–$\boldsymbol{G}_\psi(\delta, V)$ we have $h_R^\infty = 0$, $\mu_R(1_{\mathbb{X}}) < \infty$ and $\pi_R(V) < \infty$ (see the beginning of Section 6). Moreover the Banach space $(\mathcal{B}_V(\mathbb{C}), \|\cdot\|_V)$ satisfies Assumptions **(B)** since $1_{\mathbb{X}} \leq V$ and

$$\forall g \in \mathcal{B}_V(\mathbb{C}), \quad \pi_R(|g|) \leq \pi_R(V)\,\|g\|_V.$$

When P satisfies $(\boldsymbol{M}_{\nu,\psi})$–$\boldsymbol{G}_\psi(\delta, V)$ and is aperiodic, we know from Theorem 6.2 that P is $V-$geometrically ergodic, i.e. $\varrho_V < 1$. Corollary 9.5 below is thus a refinement of Theorem 6.2 since it provides a bound (even the exact value in Case (b)) of the real number ϱ_V. Corollary 9.5 is a direct consequence of Proposition 9.4 and Theorem 9.1.

Corollary 9.5 *Assume that P satisfies $(\boldsymbol{M}_{\nu,\psi})$–$\boldsymbol{G}_\psi(\delta, V)$ and is aperiodic. Then the radius of convergence of the power series $\rho(z) := \sum_{n=1}^{+\infty} \nu(R^{n-1}\psi)\, z^n$ is larger than $1/r_V$. Moreover the alternative $(a) - (b)$ of Theorem 9.1 and the additional statements of Proposition 9.3 hold with $\mathfrak{B} := \mathcal{B}_V(\mathbb{C})$, $\varrho_{\mathfrak{B}} := \varrho_V$ and $r_{\mathfrak{B}} := r_V$.*

9.3 Geometric ergodicity on $\mathbb{L}^2(\pi_R)$

Let P satisfy Conditions $(\boldsymbol{M}_{\nu,\psi})$–$\boldsymbol{G}_\psi(\delta, V)$, so that π_R is the unique $P-$invariant probability measure. The operator-norm on $(\mathbb{L}^2(\pi_R), \|\cdot\|_2)$ is also denoted by $\|\cdot\|_2$. Recall that $P \in \mathcal{L}(\mathbb{L}^2(\pi_R))$, more precisely $\|P\|_2 = 1$. Indeed we have $P1_{\mathbb{X}} = 1_{\mathbb{X}}$ and $\|Pg\|_2 \leq \|g\|_2$ for every $g \in \mathbb{L}^2(\pi_R)$ since

$$\|Pg\|_2{}^2 = \int_{\mathbb{X}} \left| \int_{\mathbb{X}} g(y) P(x, dy) \right|^2 \pi_R(dx) \leq \int_{\mathbb{X}} \int_{\mathbb{X}} |g(y)|^2 P(x, dy) \pi_R(dx) = \int_{\mathbb{X}} |g(x)|^2\, \pi_R(dx)$$

from the Cauchy-Schwarz inequality w.r.t. the probability measure $P(x, dy)$ and from the $P-$invariance of π_R. If P is geometrically ergodic on $\mathbb{L}^2(\pi_R)$, i.e. when (157) holds with $(\mathfrak{B}, \|\cdot\|) := (\mathbb{L}^2(\pi_R), \|\cdot\|_2)$, then the corresponding real number $\varrho_{\mathbb{L}^2(\pi_R)}(P)$ in (158) is denoted for short by ϱ_2. Recall that, if $L \in \mathcal{L}(\mathbb{L}^2(\pi_R))$, then its adjoint $L^* \in \mathcal{L}(\mathbb{L}^2(\pi_R))$ is defined by:

$$\forall (f, g) \in \mathbb{L}^2(\pi_R) \times \mathbb{L}^2(\pi_R), \quad \int_{\mathbb{X}} (Lf)(x)\, \overline{g(x)}\, \pi_R(dx) = \int_{\mathbb{X}} f(x)\, \overline{(L^*g)(x)}\, \pi_R(dx). \qquad (163)$$

The residual kernel R is also a bounded linear operator on $(\mathbb{L}^2(\pi_R), \|\cdot\|_2)$: in fact it is a contraction on $\mathbb{L}^2(\pi_R)$, i.e. $\|R\|_2 \leq 1$, since $0 \leq R \leq P$. Let R^* be the adjoint operator of R on $\mathbb{L}^2(\pi_R)$, and define the following $[0, +\infty]-$valued quantity

$$\vartheta_V := \limsup_{n \to +\infty} \left\| \frac{R^{*n} V}{V} \right\|_\infty^{1/n}, \qquad (164)$$

where $\|\cdot\|_\infty \equiv \|\cdot\|_{\infty,\pi_R}$ is defined in (9). Recall that the spectral radius r_V of R on $\mathcal{B}_V(\mathbb{C})$ satisfies $r_V < 1$ from Proposition 9.4. We simply denote by r_2 the spectral radius of R on $\mathbb{L}^2(\pi_R)$ (i.e. $r_2 \equiv r_{\mathbb{L}^2(\pi_R)}$ with the notation of Theorem 9.1). Note that $r_2 \leq 1$ from Gelfand's formula since R is a contraction on $\mathbb{L}^2(\pi_R)$.

Theorem 9.6 *Assume that P satisfies $(\boldsymbol{M}_{\nu,\psi})$–$\boldsymbol{G}_\psi(\delta, V)$ with $\pi_R(V^2) < \infty$ and is aperiodic. If $\vartheta_V < \infty$, then $r_2 \leq (r_V \vartheta_V)^{1/2}$. Next, if $\vartheta_V < 1/r_V$, then $r_2 < 1$ and P is geometrically ergodic on $\mathbb{L}^2(\pi_R)$. More precisely the radius of convergence of the power series $\rho(z) := \sum_{n=1}^{+\infty} \nu(R^{n-1}\psi)\, z^n$ is larger than $1/r_2$. Moreover the alternative (a)-(b) of Theorem 9.1 and the additional statements of Proposition 9.3 hold with $\mathfrak{B} := \mathbb{L}^2(\pi_R)$, $\varrho_{\mathfrak{B}} := \varrho_2$ and $r_{\mathfrak{B}} := r_2$.*

In the proof below we use the following well-known fact. Let $L \in \mathcal{L}(\mathfrak{B})$ for some Banach space $(\mathfrak{B}, \|\cdot\|)$ and assume that there exists a dense subset $\mathcal{D}$ in $\mathfrak{B}$ and a positive constant d such that: $\forall h \in \mathcal{D},\ \|Lh\| \leq d\|h\|$. Then the operator-norm $\|L\|$ of L on $(\mathfrak{B}, \|\cdot\|)$ is less than d. Indeed, let $g \in \mathfrak{B}$ and $(h_n)_n \in \mathcal{D}^n$ be such that $\lim_n \|g - h_n\| = 0$. Then

$$\|Lg\| \leq \|L(g-h_n)\| + \|Lh_n\| \leq \|L\|\,\|g-h_n\| + d\,\|h_n\|.$$

When $n \to +\infty$ this provides $\|Lg\| \leq d\,\|g\|$ since $\lim_n \|h_n\| = \|g\|$.

Proof of Theorem 9.6. Assume that $\vartheta_V < \infty$, and let $(\vartheta, r) \in (\vartheta_V, +\infty) \times (r_V, +\infty)$. From the definition of ϑ_V and r_V we know that

$$\exists n_0 \geq 1,\ \forall n \geq n_0,\ R^{*n}V \leq \vartheta^n V \quad \pi_R\text{–a.s.} \quad \text{and} \quad \exists d > 0,\ \forall n \geq 1,\ R^n V \leq d\, r^n V. \tag{165}$$

Let $g \in \mathcal{B}_{1_{\mathbb{X}}}(\mathbb{C})$ (i.e. $g : \mathbb{X} \to \mathbb{C}$ is bounded and measurable). We have for every $n \geq n_0$

$$\begin{aligned}
\|R^n g\|_2^{\,2} &= \int_{\mathbb{X}} \bigg(\int_{\mathbb{X}} \frac{g(y)}{V(y)^{1/2}} V(y)^{1/2}\, R^n(x, dy) \bigg)^2 \pi_R(dx) \\
&\leq \int_{\mathbb{X}} \bigg(\int_{\mathbb{X}} \frac{|g(y)|^2}{V(y)} R^n(x, dy) \bigg) (R^n V)(x)\, \pi_R(dx) \\
&\leq d\, r^n \int_{\mathbb{X}} (R^n \frac{|g|^2}{V})(x)\, V(x)\, \pi_R(dx) \\
&= d\, r^n \int_{\mathbb{X}} \frac{|g(x)|^2}{V(x)}\, (R^{*n}V)(x)\, \pi_R(dx) \\
&\leq d\,(r\vartheta)^n \int_{\mathbb{X}} |g(x)|^2\, \pi_R(dx)
\end{aligned}$$

using successively the Cauchy-Schwarz inequality w.r.t. the non-negative measure $R^n(x, dy)$, the second inequality in (165), the definition of the adjoint operator R^{*n} of R^n noticing that $|g|^2/V$ and V belong to $\mathbb{L}^2(\pi_R)$ since $g \in \mathcal{B}_{1_{\mathbb{X}}}(\mathbb{C})$, $V \geq 1$ and $\pi_R(V^2) < \infty$, and finally using the first inequality in (165). We have proved that

$$\forall g \in \mathcal{B}_{1_{\mathbb{X}}}(\mathbb{C}), \quad \|R^n g\|_2 \leq d^{1/2} (r\vartheta)^{n/2}\, \|g\|_2.$$

From the density of $\mathcal{B}_{1_{\mathbb{X}}}(\mathbb{C})$ in $\mathbb{L}^2(\pi_R)$ it follows that the operator-norm $\|R^n\|_2$ of R^n on $\mathbb{L}^2(\pi_R)$ satisfies $\|R^n\|_2 \leq d^{1/2}(r\vartheta)^{n/2}$, from which we deduce that $r_2 \leq (r\vartheta)^{1/2}$ from Gelfand's formula. This provides $r_2 \leq (r_V \vartheta_V)^{1/2}$ since r and ϑ are arbitrarily close to r_V and ϑ_V respectively. Next, if $\vartheta_V < 1/r_V$, then $r_2 < 1$ and the other assertions of Theorem 9.6 follows from Theorem 9.1 applied with $(\mathfrak{B}, \|\cdot\|) := (\mathbb{L}^2(\pi_R), \|\cdot\|_2)$, observing that this Banach space obviously satisfies Assumptions **(B).** □

9.4 Geometric ergodicity on $\mathbb{L}^2(\pi_R)$ in the reversible case

Again P is assumed to satisfy $(\boldsymbol{M}_{\nu,\psi})$–$\boldsymbol{G}_\psi(\delta, V)$. Recall that P is said to be reversible with respect to its (unique) invariant probability measure π_R if

$$\pi_R(dx)P(x,dy) = \pi_R(dy)P(y,dx).$$

This is equivalent to the condition $P^* = P$ where P^* is the adjoint operator of P on $\mathbb{L}^2(\pi_R)$. In other words P is reversible if, and only if, P is self-adjoint, that is:

$$\forall (f,g) \in \mathbb{L}^2(\pi_R) \times \mathbb{L}^2(\pi_R), \quad \int_{\mathbb{X}} (Pf)(x)\, \overline{g(x)}\, \pi_R(dx) = \int_{\mathbb{X}} f(x)\, \overline{(Pg)(x)}\, \pi_R(dx). \tag{166}$$

Geometric ergodicity on $\mathbb{L}^2(\pi_R)$ (case $\mathfrak{B} := \mathbb{L}^2(\pi_R)$) in the reversible case is particularly interesting since not only can the value $\rho := \varrho_2 \equiv \varrho_{\mathbb{L}^2(\pi_R)}(P) \in (0,1)$ be considered in Property (157), but also the corresponding constant c_{ϱ_2} is equal to one. Namely:

Lemma 9.7 *Assume that P is reversible and is geometrically ergodic on $\mathbb{L}^2(\pi)$ for some $P-$invariant probability measure π. Then*

$$\forall g \in \mathbb{L}^2(\pi),\ \forall n \geq 1, \quad \|P^n g - \pi(g)1_{\mathbb{X}}\|_2 \leq \varrho_2^n \|g\|_2 \tag{167}$$

where $\varrho_2 \equiv \varrho_{\mathbb{L}^2(\pi)}(P) \in (0,1)$ is given in (158).

Proof. To obtain Property (167) note that ϱ_2 is the spectral radius of the operator $P - \Pi$ where $\Pi := 1_{\mathbb{X}} \otimes \pi$: This follows from the definition of ϱ_2 and Equality $P^n - \Pi = (P-\Pi)^n$ due to the $P-$invariance of π. Moreover, since $P - \Pi$ is self-adjoint from the reversibility of P, we know that ϱ_2 equals to the operator-norm $\|P - \Pi\|_2$. Thus

$$\forall n \geq 1, \quad \|P^n - \Pi\|_2 = \|(P-\Pi)^n\|_2 \leq {\|P-\Pi\|_2}^n = {\varrho_2}^n$$

from which we deduce (167). □

Recall that r_V denotes the spectral radius of the residual kernel R on $\mathcal{B}_V(\mathbb{C})$ and that ϱ_V is defined in (161). Under the assumptions of the following theorem we know that $r_V < 1$ and $\varrho_V < 1$ from Proposition 9.4 and Corollary 9.5. Finally recall that r_2 denotes the spectral radius of R on $\mathbb{L}^2(\pi_R)$.

Theorem 9.8 *Let P satisfy Conditions $(\boldsymbol{M}_{\nu,\psi})$–$\boldsymbol{G}_\psi(\delta, V)$ with $\pi_R(V^2) < \infty$. If P is reversible and aperiodic, then*

$$r_2 \leq \big(r_V \max(r_V, \varrho_V)\big)^{1/2} < 1 \tag{168}$$

and P is geometrically ergodic on $\mathbb{L}^2(\pi_R)$. More precisely the radius of convergence of the power series $\rho(z) := \sum_{n=1}^{+\infty} \nu(R^{n-1}\psi)\, z^n$ is larger than $1/r_2$, and Property (167) *holds with ϱ_2 satisfying the following alternative:*

(a) If Equation $\rho(x^{-1}) = 1$ has no solution in the interval $(-1, -r_2)$, then $\varrho_2 \leq r_2$.

(b) Otherwise, we have $\varrho_2 = \max\big\{|x| : \rho(x^{-1}) = 1,\, x \in (-1, -r_2)\big\}$.

Moreover the additional statements of Proposition 9.3 hold with $\mathfrak{B} := \mathbb{L}^2(\pi_R)$, $\varrho_{\mathfrak{B}} := \varrho_2$, $r_{\mathfrak{B}} := r_2$, and with set $\mathcal{S}_r$ for $r \in (r_2, 1)$ given here by: $\mathcal{S}_r := \{x \in (-1, -r_2),\, \rho(x^{-1}) = 1\}$.

The proof of Theorem 9.8 is based on the following proposition.

Proposition 9.9 *If P satisfies Conditions $(\boldsymbol{M}_{\nu,\psi})$–$\boldsymbol{G}_{\psi}(\delta, V)$ with $\pi_R(V^2) < \infty$ and is reversible, then we have $\vartheta_V \leq \max(\varrho_V, r_V)$ where ϑ_V is defined in* (164).

To prove Proposition 9.9 we use the two following lemmas. Recall that, for any non-negative measurable function f, we denote by $f \cdot \pi_R$ the non-negative measure defined on $(\mathbb{X}, \mathcal{X})$ by $(f \cdot \pi_R)(1_A) := \int_{\mathbb{X}} 1_A(x) f(x) \pi_R(dx)$ for every $A \in \mathcal{X}$.

Lemma 9.10 *Let P satisfy Conditions $(\boldsymbol{M}_{\nu,\psi})$–$\boldsymbol{G}_{\psi}(\delta, V)$. Then there exists $\zeta \in \mathcal{B}_+^*$ such that $\nu = \zeta \cdot \pi_R$. Moreover $T := \psi \otimes \nu$ defines a bounded linear operator on $\mathbb{L}^2(\pi_R)$, and its adjoint operator T^* on $\mathbb{L}^2(\pi_R)$ is defined by:*

$$T^* = \zeta \otimes (\psi \cdot \pi_R). \tag{169}$$

Proof. From $(\boldsymbol{M}_{\nu,\psi})$ and the $P-$invariance of π_R we have $\pi_R \geq \pi_R(\psi)\nu$, so that ν is absolutely continuous w.r.t. π_R, i.e.: there exists a non-negative π_R-integrable function ζ_0 such that $\nu = \zeta_0 \cdot \pi_R$. Thus we have $\pi_R \geq \pi_R(\psi)(\zeta_0 \cdot \pi_R)$, so that

$$\forall A \in \mathcal{X}, \quad \int_A \big(1_{\mathbb{X}} - \pi_R(\psi)\zeta_0\big)\, d\pi_R \geq 0.$$

Therefore the set $A_0 = \{x \in \mathbb{X} : \zeta_0(x) > \pi_R(\psi)^{-1}\}$ is such that $\pi_R(A_0) = 0$. Then, defining $\zeta(x) = 0$ for $x \in A_0$ and $\zeta(x) = \zeta_0(x)$ for $x \in \mathbb{X} \setminus A_0$, we obtain that $\nu = \zeta \cdot \pi_R$ with ζ bounded by $\pi_R(\psi)^{-1}$ on $\mathbb{X}$. This proves the first assertion. Next we have from $T = \psi \otimes (\zeta \cdot \pi_R)$

$$\begin{aligned}
\forall (f,g) \in \mathbb{L}^2(\pi_R)^2, \quad \int_{\mathbb{X}} (Tf)(x)\, \overline{g(x)}\, \pi_R(dx) &= \int_{\mathbb{X}} (\zeta \cdot \pi_R)(f) \psi(x)\, \overline{g(x)}\, \pi_R(dx) \\
&= \int_{\mathbb{X}} \int_{\mathbb{X}} f(y) \zeta(y) \pi_R(dy) \psi(x) \overline{g(x)}\, \pi_R(dx) \\
&= \int_{\mathbb{X}} f(y) \int_{\mathbb{X}} \psi(x) \overline{g(x)}\, \pi_R(dx)\, \zeta(y) \pi_R(dy) \\
&= \int_{\mathbb{X}} f(y)\, \overline{(\psi \cdot \pi_R)(g) \zeta(y)}\, \pi_R(dy)
\end{aligned}$$

from which we deduce that $T^* = \zeta \otimes (\psi \cdot \pi_R)$. □

Lemma 9.11 *Assume that P satisfies $(\boldsymbol{M}_{\nu,\psi})$–$\boldsymbol{G}_{\psi}(\delta, V)$ and is reversible. Let $\zeta \in \mathcal{B}_+^*$ be given in Lemma 9.10. Then the following equalities of linear operators on $\mathbb{L}^2(\pi_R)$ hold*

$$\forall n \geq 1, \qquad P^n = R^{*n} + \sum_{k=1}^{n} P^{n-k} \zeta \otimes (R^{k-1}\psi \cdot \pi_R). \tag{170}$$

Note that Formula (170) is not the adjoint version of (17). However, starting from Equality $P = R^* + T^*$ and using Formula (169), the proof by induction of (170) is identical to that of (17), except that function equalities must be considered here in $\mathbb{L}^2(\pi_R)$. For completeness, a proof of Lemma 9.11 is provided in Appendix E.

Proof of Proposition 9.9. Recall that $\sum_{k=1}^{+\infty} R^{k-1}\psi = \nu(1_{\mathbb{X}})^{-1}1_{\mathbb{X}}$ from (35). Thus

$$\begin{aligned}\forall n\geq 1,\quad \sum_{k=1}^{n}(R^{k-1}\psi\cdot\pi_R)(V) &= \sum_{k=1}^{+\infty}(R^{k-1}\psi\cdot\pi_R)(V) - \sum_{k=n}^{+\infty}(R^{k}\psi\cdot\pi_R)(V)\\ &= \nu(1_{\mathbb{X}})^{-1}\pi_R(V) - \varepsilon_n \quad\text{with}\quad \varepsilon_n := \sum_{k=n}^{+\infty}(R^{k}\psi\cdot\pi_R)(V)\end{aligned}$$

from monotone convergence theorem. Applying (170) with $g := V$ and using the previous equality, we can write that for every $n \geq 1$

$$\begin{aligned}R^{*n}V &= P^nV - \sum_{k=1}^{n}(R^{k-1}\psi\cdot\pi_R)(V)\,P^{n-k}\zeta\\ &= P^nV - \sum_{k=1}^{n}(R^{k-1}\psi\cdot\pi_R)(V)\big(P^{n-k}\zeta - \nu(1_{\mathbb{X}})1_{\mathbb{X}}\big) - \nu(1_{\mathbb{X}})\Big(\sum_{k=1}^{n}(R^{k-1}\psi\cdot\pi_R)(V)\Big)1_{\mathbb{X}}\\ &= P^nV - \pi_R(V)1_{\mathbb{X}} - \sum_{k=1}^{n}(R^{k-1}\psi\cdot\pi_R)(V)\big(P^{n-k}\zeta - \nu(1_{\mathbb{X}})1_{\mathbb{X}}\big) + \nu(1_{\mathbb{X}})\,\varepsilon_n 1_{\mathbb{X}}. \qquad (171)\end{aligned}$$

Let $\gamma > \max(\varrho_V, r_V)$. Note that the series $\widetilde{\psi}_\gamma := \sum_{k=1}^{+\infty}\gamma^{-k}R^{k-1}\psi$ absolutely converges in $\mathcal{B}_V$ from $\gamma > r_V$ and the definition of r_V. Moreover there exists $d_\gamma > 0$ such that: $\forall k \geq 1,\ R^k\psi \leq d_\gamma\|\psi\|_V\gamma^kV$. Set $a_\gamma := d_\gamma\|\psi\|_V\pi_R(V^2)/(1-\gamma)$. Then

$$\forall n\geq 1,\quad \varepsilon_n \leq a_\gamma\gamma^n \quad\text{and}\quad 0 \leq \sum_{k=1}^{n}\gamma^{-k}(R^{k-1}\psi\cdot\pi_R)(V) \leq (\widetilde{\psi}_\gamma\cdot\pi_R)(V)$$

with $(\widetilde{\psi}_\gamma\cdot\pi_R)(V) < \infty$ since $\widetilde{\psi}_\gamma \in \mathcal{B}_V$ and $\pi_R(V^2) < \infty$ by hypothesis. Finally, from the definition of ϱ_V and $\gamma > \varrho_V$, we know that there exists $c_\gamma > 0$ such that:

$$\forall n \geq 1,\ \forall g \in \mathcal{B}_V(\mathbb{C}),\quad |P^ng - \pi_R(g)1_{\mathbb{X}}| \leq c_\gamma\|g\|_V\,\gamma^n\,V.$$

Since V, ζ belong to $\mathcal{B}_V(\mathbb{C})$ and $\nu(1_{\mathbb{X}}) = \pi_R(\zeta)$ from the definition of ζ in Lemma 9.10, the previous inequality can be applied to both V and ζ in (171). We then deduce from the triangular inequality in (171) and the above facts that

$$\begin{aligned}\forall n\geq 1,\quad \frac{R^{*n}V}{V} &\leq c_\gamma\,\gamma^n\|V\|_V + c_\gamma\,\gamma^n\|\zeta\|_V\sum_{k=1}^{n}\gamma^{-k}(R^{k-1}\psi\cdot\pi_R)(V) + \nu(1_{\mathbb{X}})a_\gamma\gamma^n\,\frac{1_{\mathbb{X}}}{V}\\ &\leq \big[c_\gamma + c_\gamma\,\|\zeta\|_V(\widetilde{\psi}_\gamma\cdot\pi_R)(V) + \nu(1_{\mathbb{X}})a_\gamma\big]\,\gamma^n\end{aligned}$$

using $1_{\mathbb{X}} \leq V$. Thus we have $\vartheta_V \leq \gamma$, and finally $\vartheta_V \leq \max(\varrho_V, r_V)$ since γ is arbitrarily close to $\max(\varrho_V, r_V)$. □

Proof of Theorem 9.8. From Theorem 9.6 we know that $r_2 \leq (r_V\vartheta_V)^{1/2}$, so that the bound (168) is deduced from Proposition 9.9. The conclusions of Theorem 9.8 then follow from Property (167) and Theorem 9.1 applied with $\mathfrak{B} = \mathbb{L}^2(\pi_R)$, $\varrho_{\mathfrak{B}} = \varrho_2$, and $r_{\mathfrak{B}} = r_2$ since the following equality holds here:

$$\{z \in \mathbb{C} : r_2 \leq |z| < 1,\ \rho(z^{-1}) = 1\} = \{x \in (-1, -r_2) : \rho(x^{-1}) = 1\}.$$

Indeed, let $z \in \mathbb{C}$ be such that $\rho(z^{-1}) = 1$ and $r_2 < |z| < 1$. Then z is an eigenvalue of P on $\mathbb{L}^2(\pi_R)$ from Proposition 9.3, i.e. $\exists h \in \mathbb{L}^2(\pi_R),\ h \neq 0,\ Ph = zh$. From reversibility we then obtain that $z \in \mathbb{R}$ (apply (166) with $f = g = h$). Moreover Equation $\rho(x^{-1}) = \sum_{n=1}^{+\infty} \nu(R^{n-1}\psi)\, x^{-n} = 1$ has no solution $x \in (r_2, 1)$ since $\rho(1) = \mu_R(\psi) = 1$. The claimed equality is proved. □

Remark 9.12 *If P satisfies the assumptions of Theorem 9.8, then P is also $V-$geometrically ergodic, and the alternative (a)-(b) of Theorem 9.1 holds with $\mathfrak{B} := \mathcal{B}_V(\mathbb{C})$, $\varrho_{\mathfrak{B}} := \varrho_V$ and $r_{\mathfrak{B}} := r_V$ (see Corollary 9.5). In fact, as in Theorem 9.8, it can be deduced from reversibility and $\mu_R(\psi) = 1$ that Equation $\rho(z^{-1}) = 1$ can be only addressed on the interval $(-1, -r_V)$, that is: Under the assumptions of Theorem 9.8 the following alternative holds*

(a) If Equation $\rho(x^{-1}) = 1$ has no solution $x \in (-1, -r_V)$, then $\varrho_V \leq r_V$.

(b) Otherwise, we have $\varrho_V = \max\big\{|x| : x \in (-1, -r_V),\ \rho(x^{-1}) = 1\big\}$.

Indeed the alternative of Corollary 9.5 can be restricted to the interval $(-1, -r_V)$ proceeding as in Theorem 9.8. More precisely, apply Proposition 9.3 with $\mathfrak{B} := \mathcal{B}_V(\mathbb{C})$ and observe that $\mathcal{B}_V(\mathbb{C}) \subset \mathbb{L}^2(\pi_R)$ from $\pi_R(V^2) < \infty$: Then the arguments used at the end of the proof of Theorem 9.8 can be repeated.

9.5 From $V-$geometric ergodicity to $V^\alpha-$geometric ergodicity

Recall that the modulated drift condition $\boldsymbol{D}_\psi(V_0 : V_2)$ derived from $\boldsymbol{G}_\psi(\delta, V)$ with $V_2 := V$ and $V_0 := V/(1-\delta)^2$ plays a central role in Proposition 9.4 to obtain $r_V < 1$. Here we present an alternative approach using Lyapunov function V^α for suitable exponents $\alpha \in (0, 1]$, in order to obtain a simple upper bound of the spectral radius of R on $\mathcal{B}_{V^\alpha}(\mathbb{C})$. More specifically we restrict this study to the case when P satisfies Conditions $(\boldsymbol{M}_{\nu,1_S})$–$\boldsymbol{G}_{1_S}(\delta, V)$ for some $S \in \mathcal{X}^*$, and we define the following set associated with the residual kernel $R := P - 1_S \otimes \nu$:

$$\mathcal{A} := \big\{\alpha \in (0, 1] : RV^\alpha \leq \delta^\alpha\, V^\alpha\big\}. \tag{172}$$

Note that $RV^\alpha = PV^\alpha - \nu(V^\alpha)1_S$.

Proposition 9.13 *Let P satisfy Conditions $(\boldsymbol{M}_{\nu,1_S})$–$\boldsymbol{G}_{1_S}(\delta, V)$ for some $S \in \mathcal{X}^*$ such that $K := \sup_{x\in S}(PV)(x) < \infty$. Then the set $\mathcal{A}$ is non-empty and reduces to*

$$\mathcal{A} = \big\{\alpha \in (0, 1] : \forall x \in S,\ (RV^\alpha)(x) \leq \delta^\alpha\, V(x)^\alpha\big\}. \tag{173}$$

Moreover we have $\mathcal{A} = (0, \widehat{\alpha}_0]$ with $\widehat{\alpha}_0 := \sup \mathcal{A}$, $\widehat{\alpha}_0 \in (0, 1]$, and

$$\forall \alpha \in \mathcal{A}, \quad r_{V^\alpha} \leq \|R\|_{V^\alpha} \leq \delta^\alpha \tag{174}$$

where $\|R\|_{V^\alpha}$ (resp. r_{V^α}) denotes the operator norm (resp. the spectral radius) of R on $\mathcal{B}_{V^\alpha}(\mathbb{C})$. Finally, if S is an atom, then $\widehat{\alpha}_0 = 1$.

Under Assumptions $(\boldsymbol{M}_{\nu,1_S})$–$\boldsymbol{G}_{1_S}(\delta, V)$ with V bounded on S, we have $K < \infty$. That K is finite is necessary to obtain (175) in the following proof.

Proof. Let $\alpha \in (0,1]$. If $x \in \mathbb{X} \setminus S$, then we have $(PV^\alpha)(x) \le \delta^\alpha\, V(x)^\alpha$ from $\boldsymbol{G}_{1_S}(\delta, V)$ and Jensen's inequality w.r.t. $P(x,dy)$. Hence the definitions (172) and (173) of the set $\mathcal{A}$ are equivalent. Next, if $x \in S$, then we have $(PV^\alpha)(x) \le K^\alpha$ from Jensen's inequality, thus

$$\forall \alpha \in (0,1],\ \forall x \in S, \quad (PV^\alpha)(x) - \delta^\alpha\, V(x)^\alpha - \nu(V^\alpha) \le K^\alpha - \delta^\alpha - \nu(1_{\mathbb{X}})$$

using $1_{\mathbb{X}} \le V$. Moreover we have

$$\lim_{\alpha \to 0} \big(K^\alpha - \delta^\alpha - \nu(1_{\mathbb{X}})\big) = -\,\nu(1_{\mathbb{X}}) \tag{175}$$

with $\nu(1_{\mathbb{X}}) > 0$. Thus the left hand side of the above inequality is negative for every $x \in S$ provided that $\alpha \in (0,1]$ is small enough. We have proved that, for $\alpha \in (0,1]$ small enough, we have $RV^\alpha \le \delta^\alpha\, V^\alpha$. This shows that $\mathcal{A} \neq \emptyset$. Now prove that $\widehat{\alpha}_0 := \sup \mathcal{A} \in \mathcal{A}$. Let $(\alpha_n)_{n\in\mathbb{N}} \in \mathcal{A}^{\mathbb{N}}$ be such that $\lim_n \nearrow \alpha_n = \widehat{\alpha}_0$. Let $x \in \mathbb{X}$. We have $\lim_n V(x)^{\alpha_n} = V(x)^{\widehat{\alpha}_0}$. Moreover we deduce from Lebesgue's theorem w.r.t. $P(x,dy)$ and $\nu(dy)$ that $\lim_n (PV^{\alpha_n})(x) = (PV^{\widehat{\alpha}_0})(x)$ and $\lim_n \nu(V^{\alpha_n}) = \nu(V^{\widehat{\alpha}_0})$ (use $V^{\alpha_n} \le V$, $(PV)(x) < \infty$ and $\nu(V) < \infty$). Since $\alpha_n \in \mathcal{A}$ for any n, this easily implies that $\widehat{\alpha}_0 \in \mathcal{A}$. If S is an atom (i.e. $\nu(\cdot) := P(a_0, \cdot)$ for some $a_0 \in S$), then we have

$$\forall \alpha \in (0,1],\ \forall x \in S, \quad PV^\alpha(x) - \delta^\alpha\, V^\alpha(x) - \nu(V^\alpha) = -\delta^\alpha\, V^\alpha(x) \le 0,$$

so that Inequality $RV^\alpha \le \delta^\alpha V^\alpha$ holds on the set S. Thus, in atomic case, we have $\mathcal{A} = (0,1]$ from the definition (173) of $\mathcal{A}$. Now assume that S is not an atom and prove that $(0, \widehat{\alpha}_0) \subset \mathcal{A}$. Let $x \in S$. Note that $\sigma_x(\cdot) := P(x,\cdot) - \nu(\cdot)$ is a positive measure on $(\mathbb{X}, \mathcal{X})$ from Condition $(\boldsymbol{M}_{\nu,1_S})$: In fact $\sigma := \sigma_x(1_{\mathbb{X}}) = 1 - \nu(1_{\mathbb{X}})$ does not depend on x and is positive since S is not an atom. Thus the following probability measures are well-defined on $(\mathbb{X}, \mathcal{X})$:

$$\forall x \in S, \quad \widehat{\sigma}_x(dy) := \frac{1}{\sigma}\, \sigma_x(dy) = \frac{1}{\sigma}\big(P(x,dy) - \nu(dy)\big). \tag{176}$$

Let $\alpha \in (0, \widehat{\alpha}_0)$. We deduce from Jensen's inequality and from $\widehat{\alpha}_0 \in \mathcal{A}$ that for every $x \in S$

$$\begin{aligned}(PV^\alpha)(x) - \nu(V^\alpha) = \sigma \widehat{\sigma}_x\big((V^{\widehat{\alpha}_0})^{\alpha/\widehat{\alpha}_0}\big) \le \sigma\big(\widehat{\sigma}_x(V^{\widehat{\alpha}_0})\big)^{\alpha/\widehat{\alpha}_0} &= \sigma^{1-\alpha/\widehat{\alpha}_0}\big((PV^{\widehat{\alpha}_0})(x) - \nu(V^{\widehat{\alpha}_0})\big)^{\alpha/\widehat{\alpha}_0} \\ &\le \sigma^{1-\alpha/\widehat{\alpha}_0}\, \delta^\alpha\, V(x)^\alpha.\end{aligned}$$

This gives: $\forall x \in S,\ (RV^\alpha)(x) \le \sigma^{1-\alpha/\widehat{\alpha}_0}\, \delta^\alpha\, V(x)^\alpha \le \delta^\alpha\, V(x)^\alpha$ since $\sigma \le 1$ and $\alpha < \widehat{\alpha}_0$. Hence $\alpha \in \mathcal{A}$ from (173). We have proved that $(0, \widehat{\alpha}_0) \subset \mathcal{A}$. Thus $\mathcal{A} = (0, \widehat{\alpha}_0)$.

It remains to prove (174). Let $\alpha \in \mathcal{A}$. Inequality $RV^\alpha \le \delta^\alpha V^\alpha$ implies that $\|R\|_{V^\alpha} \le \delta^\alpha$ since $\|R\|_{V^\alpha} = \|RV^\alpha\|_{V^\alpha}$ from the non-negativity of R. This proves the second inequality in (174). The first one is obvious from Gelfand's formula. □

According to the notation (161), for every $\alpha \in (0,1]$ the real number $\varrho_{V^\alpha} \equiv \varrho_{V^\alpha}(P)$ stands for the lower bound of all the positive real number ρ such that $\|P^n - 1_{\mathbb{X}} \otimes \pi_R\|_{V^\alpha} = \mathrm{O}(\rho^n)$, where $\|\cdot\|_{V^\alpha}$ denotes here the operator norm on $\mathcal{B}_{V^\alpha}(\mathbb{C})$. Thus P is V^α−geometrically ergodic if, and only if, $\varrho_{V^\alpha} < 1$.

Corollary 9.14 *Let P satisfy Conditions $(\boldsymbol{M}_{\nu,1_S})$–$\boldsymbol{G}_{1_S}(\delta, V)$ for some $S \in \mathcal{X}^*$ such that $K := \sup_{x\in S}(PV)(x) < \infty$. If P is aperiodic, then the following assertions hold.*

1. *For every $\alpha \in (0,1]$, P is V^α−geometrically ergodic (i.e. $\varrho_{V^\alpha} < 1$).*

2. *For every $\alpha \in \mathcal{A}$ the following alternative holds:*

   (a) *If Equation $\rho(z^{-1}) = 1$ has no solution $z \in \mathbb{C}$, $\delta^\alpha < |z| < 1$, then $\varrho_{V^\alpha} \le \delta^\alpha$.*

   (b) *Otherwise, we have $\varrho_{V^\alpha} = \max\big\{|z| : z \in \mathbb{C},\ \rho(z^{-1}) = 1,\ \delta^\alpha < |z| < 1\big\}$.*

In Case 2.(b) Proposition 9.3 applies with $\mathfrak{B} := \mathcal{B}_{V^\alpha}(\mathbb{C})$ and any $r \in (\delta^\alpha, 1)$.

Proof. Let $\alpha \in (0, 1]$. If $x \in \mathbb{X} \setminus S$, then we have $(PV^\alpha)(x) \le \delta^\alpha V(x)^\alpha$ from $\boldsymbol{G}_{1_S}(\delta, V)$ and Jensen's inequality. Moreover, for every $x \in S$, we have $(PV^\alpha)(x) \le K^\alpha$ again from Jensen's inequality. Consequently P satisfies Conditions $(\boldsymbol{M}_{\nu,1_S})$ and $\boldsymbol{G}_{1_S}(\delta^\alpha, V^\alpha)$, so that P is $V^\alpha-$geometrically ergodic from Theorem 6.2 applied with the Lyapunov function V^α. Moreover, if $\alpha \in \mathcal{A}$, then the real number ϱ_{V^α} satisfies the claimed alternative applying Corollary 9.5 with the Lyapunov function V^α and using the upper bound δ^α of r_{V^α} provided by (174). □

Let us now specify the alternative of Corollary 9.14 for $\alpha \in \mathcal{A} = (0, \widehat{\alpha}_0]$ according to whether Case 2.(a) or 2.(b) holds for the specific value $\widehat{\alpha}_0$.

Corollary 9.15 *Let P satisfy the assumptions of Corollary 9.14. Then the following assertions hold.*

(i) *If Case 2.(a) of Corollary 9.14 is fulfilled for $\widehat{\alpha}_0$, then we have: $\forall \alpha \in (0, \widehat{\alpha}_0],\ \varrho_{V^\alpha} \le \delta^\alpha$.*

(ii) *If Case 2.(b) is fulfilled for $\widehat{\alpha}_0$, then there exists a unique $\widehat{\alpha} \in (0, \widehat{\alpha}_0)$ such that $\delta^{\widehat{\alpha}} = \varrho_{V^{\widehat{\alpha}_0}}$, and*

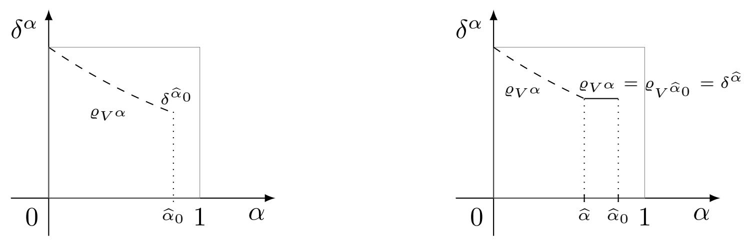
$$\forall \alpha \in (\widehat{\alpha}, \widehat{\alpha}_0],\ \varrho_{V^\alpha} = \varrho_{V^{\widehat{\alpha}_0}}, \qquad \forall \alpha \in (0, \widehat{\alpha}],\ \varrho_{V^\alpha} \le \delta^\alpha.$$

Proof. Case (i) means that there is no solution $z \in \mathbb{C}$ of Equation $\rho(z^{-1}) = 1$ such that $\delta^{\widehat{\alpha}_0} < |z| < 1$, so that the same holds when $\delta^\alpha < |z| < 1$ for $\alpha \in (0, \widehat{\alpha}_0]$, thus $\varrho_{V^\alpha} \le \delta^\alpha$ from Corollary 9.14. Case (ii) means that there exists a solution $z_0 \in \mathbb{C}$ of Equation $\rho(z^{-1}) = 1$ such that $\delta^{\widehat{\alpha}_0} < |z_0| = \varrho_{V^{\widehat{\alpha}_0}} < 1$, and that this equation has no solution $z \in \mathbb{C}$ such that $\varrho_{V^{\widehat{\alpha}_0}} < |z| < 1$. The existence and uniqueness of $\widehat{\alpha} \in (0, \widehat{\alpha}_0)$ such that $\delta^{\widehat{\alpha}} = \varrho_{V^{\widehat{\alpha}_0}}$ hold since $\alpha \mapsto \delta^\alpha$ is bijective from $(0, \widehat{\alpha}_0)$ into $(\delta^{\widehat{\alpha}_0}, 1)$. From Corollary 9.14 we obtain that $\varrho_{V^\alpha} = \varrho_{V^{\widehat{\alpha}_0}}$ for every $\alpha \in (\widehat{\alpha}, \widehat{\alpha}_0]$ since z_0 satisfies $\delta^\alpha < |z_0| < 1$ from $\delta^\alpha < \delta^{\widehat{\alpha}} = \varrho_{V^{\widehat{\alpha}_0}} = |z_0|$. On the other hand, again from Corollary 9.14 we have $\varrho_{V^\alpha} \le \delta^\alpha$ for every $\alpha \in (0, \widehat{\alpha}]$ since there is no solution $z \in \mathbb{C}$ of Equation $\rho(z^{-1}) = 1$ such that $\delta^\alpha < |z| < 1$ from $\varrho_{V^{\widehat{\alpha}_0}} = \delta^{\widehat{\alpha}} \le \delta^\alpha$. □

Figure 1 helps to get a picture of the status of the value δ^α w.r.t. the convergence rate ϱ_{V^α} in the alternative of Corollary 9.15. Note that the upper bound of ϱ_{V^α} degrades when $\alpha \to 0$, which is consistent with $\lim_{\alpha \to 0} V^\alpha = 1_{\mathbb{X}}$ and the fact that P is not $1_{\mathbb{X}}$-geometrically ergodic in general (i.e. P is not uniformly ergodic in general, see Example 3.7).

Recall that $\mathcal{A} = (0, \widehat{\alpha}_0]$ with $\widehat{\alpha}_0 \in (0, 1]$ from Proposition 9.13, and that $\mathcal{A} = (0, 1]$ when S is an atom. In the non-atomic case a positive lower bound of $\widehat{\alpha}_0$ can be obtained using (175) (i.e. consider $\alpha \in (0, 1]$ such that $K^\alpha - \delta^\alpha \le \nu(1_{\mathbb{X}})$). The next statement provides a more accurate estimate of $\widehat{\alpha}_0$.

Proposition 9.16 *Let P satisfy Conditions $(\boldsymbol{M}_{\nu,1_S})$–$\boldsymbol{G}_{1_S}(\delta, V)$ for some $S \in \mathcal{X}^*$ which is not an atom. Assume that $K := \sup_{x \in S}(PV)(x) < \infty$ and define $M := K - \nu(V)$, $\sigma := 1 - \nu(1_{\mathbb{X}}) \in (0, 1)$. Then there exists $\alpha_0 \in (0, 1]$ such that $M^{\alpha_0}\sigma^{1-\alpha_0} \le \delta^{\alpha_0}$, and such an α_0 belongs to $\mathcal{A}$, i.e. $(0, \alpha_0] \subset \mathcal{A}$.*



Case (i) of Corollary 9.15 Case (ii) of Corollary 9.15

Figure 1: Status of the value δ^α w.r.t. ϱ_{V^α} for $\alpha \in \mathcal{A} = (0, \widehat{\alpha}_0]$ according to Cases (i) or (ii) in Corollary 9.15: upper bound in dashed-line, exact value in full-line.

Proof. Recall that $\sigma \in (0,1)$ since S is not an atom. For any $x \in S$ let $\widehat{\sigma}_x$ be the probability measure defined in (176). It follows from Jensen's inequality that

$$\forall \alpha \in (0,1],\ \forall x \in S, \quad (PV^\alpha)(x) - \nu(V^\alpha) = \sigma\widehat{\sigma}_x(V^\alpha) \leq \sigma\big(\widehat{\sigma}_x(V)\big)^\alpha = \sigma^{1-\alpha}\big((PV)(x) - \nu(V)\big)^\alpha,$$

from which we deduce that

$$\forall \alpha \in (0,1],\ \forall x \in S, \quad (PV^\alpha)(x) - \nu(V^\alpha) - \delta^\alpha V(x)^\alpha \leq \sigma^{1-\alpha}\, M^\alpha - \delta^\alpha$$

since $V \geq 1_{\mathbb{X}}$. The claimed conclusion then follows from $\lim_{\alpha\to 0} \sigma^{1-\alpha}\, M^\alpha - \delta^\alpha = \sigma - 1 < 0$. Hence there exists $\alpha_0 \in (0,1]$ such that $M^{\alpha_0}\sigma^{1-\alpha_0} \leq \delta^{\alpha_0}$, and such an α_0 belongs to $\mathcal{A}$ from the definition (173) of $\mathcal{A}$. □

9.6 Further results in the reversible and positive reversible cases

In the particular case when P is reversible and R is self-adjoint on $\mathbb{L}^2(\pi_R)$ too, the proof of Theorem 9.8 is simpler using Inequality $r_2 \leq r_V$ and Proposition 9.4. This is detailed in the following proposition.

Proposition 9.17 *Assume that P satisfies Conditions $(\boldsymbol{M}_{\nu,\psi})$–$\boldsymbol{G}_\psi(\delta, V)$ with $\pi_R(V^2) < \infty$, and that P is reversible and aperiodic. Let $\zeta \in \mathcal{B}^*_+$ be given in Lemma 9.10. Then the residual kernel R is self-adjoint on $\mathbb{L}^2(\pi_R)$ if, and only if, $\zeta = c\,\psi$ for some positive constant c. Moreover, in this case, P is geometrically ergodic on $\mathbb{L}^2(\pi_R)$ and the last assertion of Theorem 9.8 holds with $r_2 = \|R\|_2 \leq r_V < 1$.*

Proof. Since P is reversible, R is self-adjoint on $\mathbb{L}^2(\pi_R)$ if and only if $T := \psi \otimes \nu$ is self-adjoint on $\mathbb{L}^2(\pi_R)$. Thus, the first assertion is obvious from Lemma 9.10. Next, assume that R is self-adjoint on $\mathbb{L}^2(\pi_R)$. Then we know that $r_2 = \|R\|_2$. Moreover recall that $r_2 \leq (r_V\vartheta_V)^{1/2}$ from Theorem 9.6. Thus we have $r_2 \leq r_V$ since $\vartheta_V \leq r_V$ from $R^* = R$ and the definitions of ϑ_V and r_V. That $r_V < 1$ is proved in Proposition 9.4. Hence we have $r_2 < 1$, and the others assertions of Proposition 9.17 follow from Theorem 9.1 applied with $\mathfrak{B} := \mathbb{L}^2(\pi_R)$. □

Although the residual kernel R is not necessarily self-adjoint when P is reversible (see Proposition 9.17), this scenario is not unrealistic, as illustrated in the following proposition.

Proposition 9.18 *Let P satisfy Conditions $(\boldsymbol{M}_{\nu,1_S})$–$\boldsymbol{G}_{1_S}(\delta, V)$ for some $(\nu, S) \in \mathcal{M}^*_{+,b} \times \mathcal{X}^*$ such that $\nu(1_{S^c}) = 0$. If the function $\zeta \in \mathcal{B}^*_+$ in Lemma 9.10 is such that $d :=$*

$\inf_{x\in S}\zeta(x) > 0$, *then P also satisfies Conditions* $(\boldsymbol{M}_{\nu_1,\psi_1})$-$\boldsymbol{G}_{\psi_1}(\delta, V)$ *with* $\psi_1 := \sqrt{c}\,\zeta$ *and* $\nu_1 := \sqrt{c}\,\nu$ *where* $c = (\sup_{x\in S}\zeta(x))^{-1}$. *If moreover P is reversible, then the residual kernel* $R_1 := P - \psi_1 \otimes \nu_1$ *is self-adjoint on* $\mathbb{L}^2(\pi_R)$.

Proof. We have $\nu = \zeta \cdot \pi_R$ from Lemma 9.10, with here $\zeta = 0$ on S^c since $\nu(1_{S^c}) = 0$. Thus

$$P \geq 1_S \otimes \nu \geq c\,\zeta \otimes \nu.$$

Hence $P \geq \psi_1 \otimes \nu_1$ with $\psi_1 := \sqrt{c}\,\zeta$ and $\nu_1 := \sqrt{c}\,\nu = \psi_1 \cdot \pi_R$. Moreover we deduce from $\boldsymbol{G}_{1_S}(\delta, V)$ that

$$PV \leq \delta V + b\,1_S \leq \delta V + b\,d^{-1}\zeta = \delta V + b\,d^{-1}c^{-1/2}\psi_1,$$

thus P satisfies $\boldsymbol{G}_{\psi_1}(\delta, V)$. Finally, under Conditions $(\boldsymbol{M}_{\nu_1,\psi_1})$–$\boldsymbol{G}_{\psi_1}(\delta, V)$, Lemma 9.10 implies that $T_1 := \psi_1 \otimes \nu_1 = \psi_1 \otimes (\psi_1 \cdot \pi_R)$ is self-adjoint on $\mathbb{L}^2(\pi_R)$. Consequently $R_1 := P - \psi_1 \otimes \nu_1$ is self-adjoint when P is reversible. □

If the spectral radius r_V of R on $\mathcal{B}_V(\mathbb{C})$ is easier to compute or to estimate than the spectral radius r_2 of R on $\mathbb{L}^2(\pi_R)$ (using for instance (162)), then the alternative of Theorem 9.8 in reversible case can be replaced with the following one.

Proposition 9.19 *Assume that P satisfies Conditions* $(\boldsymbol{M}_{\nu,\psi})$–$\boldsymbol{G}_{\psi}(\delta, V)$ *with* $\pi_R(V^2) < \infty$, *and that P is reversible and aperiodic. Let* $\Pi_R := 1_{\mathbb{X}} \otimes \pi_R$. *Then the following alternative holds.*

(a) If Equation $\rho(x^{-1}) = 1$ *has no real solution* $x \in (-1, -r_V)$, *then we have* $\varrho_2 \leq r_V$, *thus :* $\forall n \geq 1,\ \|P^n - \Pi_R\|_2 \leq r_V^n$.

(b) Otherwise, we have $\varrho_2 = \varrho_V$, *thus:* $\forall n \geq 1,\ \|P^n - \Pi_R\|_2 \leq \varrho_V^n$.

The proof of Proposition 9.19 below is obtained by combining the results of both Theorem 9.8 and Corollary 9.5. Similarly, combining Theorem 9.8 and Corollary 9.14, it can be proved that the alternative $(a)-(b)$ of Proposition 9.19 holds true with Lyapunov function V^α for $\alpha \in \mathcal{A}$ (in place of V) and with the explicit upper bound δ^α of r_{V^α} (in place of r_V).

Proof. Under the assumptions of Proposition 9.19, Equation $\rho(x^{-1}) = 1$ in the alternative $(a)-(b)$ of Theorem 9.8 only focusses on real numbers $x \in (-1, -r_2)$. The same restriction holds in $V-$geometric ergodicity w.r.t. to the interval $(-1, -r_V)$ as observed in Remark 9.12.

Assume that Equation $\rho(x^{-1}) = 1$ has no real solution $x \in (-1, -r_V)$. Then we have $\varrho_V \leq r_V$ from Remark 9.12. Thus $r_2 \leq (r_V \max(r_V, \varrho_V))^{1/2} \leq r_V$ from Theorem 9.8. This inequality $r_2 \leq r_V$ combined with Theorem 9.8 provides the following alternative. If Equation $\rho(x^{-1}) = 1$ has no solution $x \in (-1, -r_2)$, then we have $\varrho_2 \leq r_2 \leq r_V$. If Equation $\rho(x^{-1}) = 1$ has solutions $x \in (-1, -r_2)$, then these solutions necessarily belong to $[-r_V, -r_2)$, so that $\varrho_2 \leq r_V$ from Theorem 9.8. Each case of the previous alternative provides $\varrho_2 \leq r_V$, thus Case (a) of Proposition 9.19 is proved.

Now prove Case (b). Assume that Equation $\rho(x^{-1}) = 1$ has solutions in $(-1, -r_V)$. Then

$$\varrho_V = \max\big\{|x| : x \in (-1, -r_V),\ \rho(x^{-1}) = 1\big\}$$

from Remark 9.12, in particular we have $\varrho_V > r_V$. It then follows from Theorem 9.8 that $r_2 \leq (r_V \max(r_V, \varrho_V))^{1/2} < \varrho_V$ and that

$$\varrho_2 = \max\big\{|x| : x \in (-1, -r_2),\ \rho(x^{-1}) = 1\big\} = \varrho_V.$$

This proves Case (b). □

Finally recall that a reversible Markov kernel P is said to be positive if the following condition holds

$$\forall g \in \mathbb{L}^2(\pi_R), \quad \int_{\mathbb{X}} (Pg)(x)\,\overline{g(x)}\,\pi_R(dx) \geq 0. \tag{177}$$

The relevant fact to apply Theorem 9.8 to the positive reversible case is that any eigenvalue $z \in \mathbb{C}$ of P (i.e. $\exists h \in \mathbb{L}^2(\pi_R), h \neq 0, Ph = zh$) is in fact a non-negative real number. Indeed we know that $z \in \mathbb{R}$ from reversibility. Moreover Condition (177) applied to h implies that $z\,\pi_R(h^2) = \pi_R(Ph \cdot h) \geq 0$ with $\pi_R(h^2) > 0$ since $h \neq 0$ in $\mathbb{L}^2(\pi_R)$. Thus $z \geq 0$.

Proposition 9.3 and the previous fact then imply that Case (a) of Theorem 9.8 holds when P is positive reversible,that is:

Corollary 9.20 *Let P satisfy Conditions $(\boldsymbol{M}_{\nu,\psi})$–$\boldsymbol{G}_{\psi}(\delta, V)$ with $\pi_R(V^2) < \infty$. If P is aperiodic and positive reversible, then P is geometrically ergodic on $\mathbb{L}^2(\pi_R)$ with $\varrho_2 \leq r_2$, where $r_2 \in (0,1)$ is the spectral radius of the residual kernel R on $\mathbb{L}^2(\pi_R)$.*

If P is reversible, then P^2 is reversible too, and it is positive since

$$\forall g \in \mathbb{L}^2(\pi_R), \quad \int_{\mathbb{X}} (P^2 g)(x) \cdot \overline{g(x)}\pi_R(dx) = \int_{\mathbb{X}} (Pg)(x) \cdot \overline{Pg(x)}\pi_R(dx) \geq 0.$$

Then the following statement can be deduced from Corollary 9.20.

Corollary 9.21 *Assume that P satisfies Condition $(\boldsymbol{M}_{\nu,\psi})$ with $\mu_R(1_{\mathbb{X}}) < \infty$, and is irreducible, aperiodic and reversible. Moreover assume that P^2 satisfies Conditions $(\boldsymbol{M}_{\nu_2,\psi_2})$-$\boldsymbol{G}_{\psi_2}(\delta_2, V)$ for some $(\nu_2, \psi_2) \in \mathcal{M}^*_{+,b} \times \mathcal{B}^*_+$, $\delta_2 \in (0,1)$ and Lyapunov function V such that $\pi_R(V^2) < \infty$. Then P is geometrically ergodic on $\mathbb{L}^2(\pi_R)$ and we have $\varrho_2 \leq \sqrt{r_2(R_2)}$, where $r_2(R_2)$ is the spectral radius of $R_2 := P^2 - \psi_2 \otimes \nu_2$ on $\mathbb{L}^2(\pi_R)$.*

Proof of Corollary 9.21. Recall that π_R is the unique $P-$invariant probability measure under the assumptions on P (see Corollary 3.13). Next, we know from the assumptions on P^2 that P^2 admits a unique invariant probability measure which is given by π_{R_2}. Since π_R is also P^2-invariant, it follows that $\pi_{R_2} = \pi_R$. We deduce from Corollary 9.20 applied to P^2 under Conditions $(\boldsymbol{M}_{\nu_2,\psi_2})$–$\boldsymbol{G}_{\psi_2}(\delta_2, V)$ that P^2 is geometrically ergodic on $\mathbb{L}^2(\pi_R)$ with $\varrho_2(P^2) \leq r_2(R_2)$ with $R_2 := P^2 - \psi_2 \otimes \nu_2$. Now, writing any integer $n \geq 1$ as $n = 2k + r$ with $r \in \{0,1\}$ and defining $\Pi_R := 1_{\mathbb{X}} \otimes \pi_R$, we obtain that

$$P^n - \Pi_R = (P - \Pi_R)^{2k+r} = (P - \Pi_R)^r \left((P^2)^k - \Pi_R\right)$$

from which we easily deduce that $\varrho_2(P) \leq \sqrt{\varrho_2(P^2)} \leq \sqrt{r_2(R_2)}$. □

Let us finally complete Corollary 9.21 proving the following statement.

Corollary 9.22 *If P satisfies Conditions $(\boldsymbol{M}_{\nu,\psi})$–$\boldsymbol{G}_{\psi}(\delta, V)$ with $(\nu, \psi) \in \mathcal{M}^*_{+,b} \times \mathcal{B}^*_+$, and is strongly aperiodic (i.e. $\nu(\psi) > 0$), then P^2 satisfies Conditions $(\boldsymbol{M}_{\nu,P\psi})$-$\boldsymbol{G}_{P\psi}(\delta^2, V)$.*

Using the conditions $(\boldsymbol{M}_{\nu,P\psi}) - \boldsymbol{G}_{P\psi}(\delta^2, V)$ for P^2, the conclusions of Corollary 9.21 apply to $R_2 = P^2 - P\psi \otimes \nu = PR$, where $R := P - \psi \otimes \nu$ is the residual kernel of P w.r.t the minorization condition $(\boldsymbol{M}_{\nu,\psi})$.

Proof. It follows from $(\boldsymbol{M}_{\nu,\psi})$ and one iteration of $\boldsymbol{G}_{\psi}(\delta, V)$ that

$$P^2 \geq P\psi \otimes \nu \quad \text{and} \quad P^2 V \leq \delta^2 V + \delta b \psi + bP\psi \leq \delta^2 V + \big(\delta b \nu(\psi)^{-1} + b\big) P\psi$$

using the non-negativity of P and $P\psi \geq \nu(\psi)\psi$ due to $(\boldsymbol{M}_{\nu,\psi})$ for the last inequality. □

9.7 Further comments and bibliographic discussion

A bibliographic discussion on the V−geometric rate of convergence is presented in Subsection 6.3. The general presentation in Theorem 9.1 based on the condition $r_{\mathfrak{B}} < 1$ is new to the best of our knowledge. Actually Theorem 9.1 is the extended version of [HL24], which focused solely on V−geometric ergodicity. Here the case $\mathfrak{B} = \mathcal{B}_V(\mathbb{C})$ is obtained in Subsection 9.2 under Conditions $(\boldsymbol{M}_{\nu,\psi})$–$\boldsymbol{G}_\psi(\delta, V)$ as a by-product of Theorem 9.1. It should be noted that the condition $r_V < 1$ is here obtained in a much more elementary way than in [HL24] thanks to Proposition 8.1, see Proposition 9.4. More generally, all the arguments used in this section, including those in Appendix D, are based solely on the spectral theory prerequisites (S1)-(S3) presented in Subsection 6.2 (page 57).

The rate of convergence in $\mathbb{L}^2(\pi_R)$ is classically studied for reversible Markov kernels. Here the first application of Theorem 9.1 to the case $\mathfrak{B} := \mathbb{L}^2(\pi_R)$ is addressed in Theorem 9.6 for general Markov kernels, introducing the quantity ϑ_V linked to the adjoint operator of R on $\mathbb{L}^2(\pi_R)$, see (164). To our knowledge this result is new. The computation used for bounding $\|R^n g\|_2{}^2$ in the proof of Theorem 9.6 is inspired by [TM22]. The second application in Theorem 9.8 concerns the reversible case. It is in fact a weak version of the classical result in [RR97], stating that an aperiodic and reversible Markov kernel satisfying Conditions $(\boldsymbol{M}_{\nu,1_S})$–$\boldsymbol{G}_{1_S}(\delta, V)$ with $\pi_R(V^2) < \infty$ is geometrically ergodic on $\mathbb{L}^2(\pi_R)$ with $\varrho_2 \leq \varrho_V$, see also [RT01, Bax05, KM12, DMPS18]. The proof in [RR97] is based on an argument involving spectral measures. Explicit rates of convergence are obtained in [Bax05, TM22] under minorization and geometric drift conditions. In Theorem 9.8 the geometric ergodicity on $\mathbb{L}^2(\pi_R)$ is proved, but Inequality $\varrho_2 \leq \varrho_V$ is only obtained when $\max(r_V, r_2) < \varrho_V$, in which case we actually have $\varrho_2 = \varrho_V$ according to the alternative stated in both Corollary 9.5 and Theorem 9.8. Complements and examples for reversible Markov kernels, in particular in connection with MCMC algorithms, can be found in [RR04] and [DMPS18, Chap. 2 and 22]. The positive reversible assumption addressed in Corollaries 9.20-9.21 is detailed in [DMPS18, Def. 22.4.6 and examples therein]. Finally recall that the geometric ergodicity of P on $\mathbb{L}^2(\pi_R)$ implies the geometric ergodicity on $\mathbb{L}^p(\pi_R)$ for every $p \in (1, +\infty)$ from the Riesz-Thorin interpolation theorem, e.g. see [Ros71, DMPS18]. For a general study of positive operators on Lebesgue's space $\mathbb{L}^p$ with applications to Markov kernels, the reader can consult [Hin00, Hin02, Wu04, GW06].

The drift inequality $RV^\alpha \leq \delta^\alpha V^\alpha$ for some suitable exponents $\alpha \in (0, 1]$ was introduced in [HL24] to study Poisson's equation and the V^α−geometric ergodicity under Conditions $(\boldsymbol{M}_{\nu,1_S})$–$\boldsymbol{G}_{1_S}(\delta, V)$. Here the focus is only on the V^α−geometric ergodicity. The fact that such exponents form an interval $\mathcal{A} \subset [0, 1)$ completes this study (see Proposition 9.13). Recall that we have $\mathcal{A} = (0, 1]$ in atomic case. In fact this equality $\mathcal{A} = (0, 1]$ may also occur for non-atomic small-set S, even in the case of a continuous state space $\mathbb{X}$, see [HL24, Ex. 5.1]. Using the function series $\widetilde{g}$ of Corollary 6.1, the inequality $r_{V^\alpha} \leq \delta^\alpha$ proved in Proposition 9.13 can be used to obtain a simple bound for the V^α−weighted norm of solutions to Poisson's equation. This bound detailed in [HL24] involves the constant $(1 - \delta^\alpha)^{-1}$, which is large when the drift inequality $RV^\alpha \leq \delta^\alpha V^\alpha$ is only satisfied for α close to zero. In such a case, the bounds (70) and (71) for the V−weighted norm of solutions to Poisson's equation are more relevant.

Finally, we emphasize the following point which is important in practice and not addressed in our work: What is called rate of convergence in this section only concerns the real number

$\varrho_{\mathfrak{B}}$ in (158). The constant c_ρ in (157) is not investigated here (see the references given in Subsection 6.3 on this topic). We simply recall that the most favourable case is reversibility, since in this case ϱ_2 can be considered in (157) (case $\mathfrak{B} := \mathbb{L}^2(\pi_R)$) with associated constant $c_{\varrho_2} = 1$ (see (167)).

10 Examples

In this section, the M & D conditions are discussed for two classes of Markov kernels. The first class is associated with the Markov kernel of the mean of the Dirichlet process. The second one concerns the Metropolis-Hastings (MH) Markov kernels with a focus on two specific instances: the so-called symetric random walk MH Markov kernel and the independent MH Markov kernel. Such an analysis allows us to illustrate a large part of the material provided in this document. Each example is accompanied by brief bibliographic comments to introduce the framework involving such Markov models, but also to help the reader go further.

These two classes of Markov models lead to classical applications in statistics and applied probability. These applications require a number of numerical controls for which the results of the previous sections may be relevant. It is important to point out that there are a large collection of Markov models in time series analysis, in system control, in stochastic operations research, for which such material has been developed. It is beyond the scope of this section to be exhaustive. We refer the reader to books such as [Mey08, MT09, DMS14, DMPS18, Mey22] for an overview of the applicability of the M & D conditions to the asymptotic control of a Markov kernel. Note that checking M & D conditions are often somewhat easy. The random walks on half line in Subsection 7.2.2 (considered with a single fixed theta) is a typical instance of the calculations providing a drift inequality. However the two examples discussed here show that this is far from always the case, in particular when the state space $\mathbb{X}$ is multi-dimensional. Thus, some technical passages concerning the proof of these drift conditions are simply summarized and a precise reference is provided allowing the reader to easily find the needed technical complements.

10.1 Markov chain for the mean of the Dirichlet process

Let $\mu \in \mathcal{M}_{+,b}^*$ be any finite positive measure on $(\mathbb{X} := \mathbb{R}, \mathcal{X})$ where $\mathcal{X}$ is the Borel $\sigma-$algebra on $\mathbb{R}$. Let $\mu_0 := \mu/\mu(1_{\mathbb{R}})$ be the probability measure associated with μ. Throughout this subsection, the probability measure μ_0 is assumed not to be a Dirac measure. Let us introduce the following sequence of r.v. $(X_n)_{n\geq 0}$

$$X_0 := x \in \mathbb{R}, \quad \forall n \geq 1,\ X_n := f(X_{n-1}, (U_n, W_n)) \tag{178}$$

where $f\big(x, (u, w)\big) := ux + (1-u)w$, $(U_n, W_n)_{n\geq 1}$ is a sequence of i.i.d. random vectors, such that U_n and W_n are independent, $W_n \sim \mu_0$, and finally U_n has a Beta probability distribution with parameters $(\mu(1_{\mathbb{R}}), 1)$, that is with the probability density function with respect to the Lebesgue measure on $(\mathbb{R}, \mathcal{X})$: $x \mapsto \mu(1_{\mathbb{R}})\, x^{\mu(1_{\mathbb{R}})-1} 1_{(0,1)}(x)$. Then $(X_n)_{n\geq 0}$ is a homogeneous Markov chain with state space $\mathbb{X} = \mathbb{R}$ and transition kernel P given by

$$\forall x \in \mathbb{R}, \forall A \in \mathcal{X}, \quad P(x, A) = \mathbb{P}_x(X_1 \in A) = \mathbb{P}(f(x, (U_1, W_1)) \in A). \tag{179}$$

Such a homogeneous Markov chain has been introduced in [FT89] to analyse the probability distribution of the mean of a Dirichlet process. This is the basic tool for an MCMC method for sampling this probability distribution (see comments in Subsection 10.1.5).

In Subsection 10.1.1, the minorization condition ($\boldsymbol{M}_{\nu,1_S}$) is shown to hold true for any compact set S. Next various drift conditions are provided. Since any compact set is a first-order small set, these conditions are essentially derived from the control of the difference

$$PV_0(x) - V_0(x) = \mathbb{E}_x[V_0(X_1)] - V_0(x) = \mathbb{E}\big[V_0(xU_1 + (1-U_1)W_1)\big] - V_0(x) \tag{180}$$

for x large enough for some Lyapunov function V_0. An appropriated moment condition on the probability measure μ_0 is required for the function PV_0 to be well-defined. We essentially follow [GT01, JT02].

10.1.1 Minorization condition

Let us somewhat explicit the transition kernel P in (179) according to [GT01, p. 579, (15)]:

$$\begin{aligned}
&\forall x \in \mathbb{R},\ \forall A \in \mathcal{B}(\mathbb{R}),\\
P(x,A) &= \mathbb{P}(W_1 \in A, W_1 = x) + \mathbb{P}(W_1 \in A, U_1(x - W_1) + W_1 \in A)\\
&= \mu_0(\{x\})\delta_x(A) + (1-\mu_0(\{x\}))\mathbb{P}(W_1 \in A, U_1(x-W_1) + W_1 \in A \mid W_1 \neq x)\\
&\geq c\,\mathbb{P}(W_1 \in A, U_1(x-W_1) + W_1 \in A \mid W_1 \neq x) =: c\int_A p(x,y)dy
\end{aligned} \tag{181}$$

where $c := 1 - \sup_{x \in D_0} \mu_0(\{x\})$ with D_0 defined as the set of discontinuity points of the distribution function of μ_0, and finally where $p(x,\cdot)$ denotes the density probability function with respect to the Lebesgue measure on $\mathbb{R}$ associated with the conditional probability distribution of X_1 given $X_0 = x, X_1 \neq x$. Note that $c > 0$ since the probability measure μ_0 is assumed not to be a Dirac measure. Using a similar way to Proposition 3.1, the minorizing measure $\nu \in \mathcal{M}^*_{+,b}$ is obtained in assessing $\inf_{x\in S} p(x,y)$ with the compact set $S = [-\kappa,\kappa]$ for some $\kappa > 0$. Specifically, it is obtained in [GT01, (17)] that

$$\forall x \in \mathbb{R}, \forall A \in \mathcal{B}(\mathbb{R}),\ P(x,A) \geq 1_S(x)\,\nu(1_A) \quad \text{where } \nu(dy) := c\,\min(p(-\kappa,y), p(\kappa,y))dy \tag{182}$$

using that $x \mapsto p(x,y)$ is non-decreasing for $x < y$ and non-increasing for $y < x$. If $[a,b]$ with $-\infty \leq a < b \leq +\infty$ is the smallest interval containing the support of the probability measure μ_0, then the function $y \mapsto \min(p(-\kappa,y), p(\kappa,y))$ is shown in [GT01, p. 580] to be positive on the interval $[a,b]$. In other words, the transition kernel P satisfies the minorization condition ($\boldsymbol{M}_{\nu,1_S}$) for any compact set $S := [-\kappa,\kappa]$. We refer to [GT01, (15)-(16)-(17)] for details.

Finally, assume that the compact set S is large enough so that $S \cap [a,b] \neq \emptyset$. Then $\nu(1_S) \geq \nu(1_{S\cap[a,b]}) > 0$. Hence the transition kernel is strongly aperiodic.

Note that if $x \in [a,b]$, then $X_1 := f(x,(U_1,W_1)) = U_1 x + (1-U_1)W_1 \in [a,b]$ from $U_1 \in [0,1]$, $W_1 \in \operatorname{supp}(\mu_0) \subset [a,b]$. Thus $X_n \in [a,b]$ for every $n \geq 1$ and the interval $[a,b]$ is $P-$absorbing. Let us assume that $\operatorname{supp}(\mu_0)$ is bounded. Then $[a,b]$ is compact and if $X_0 := x \in [a,b]$ then $[a,b]$ may be considered as the state space $(X_n)_{n\geq 0}$. Since we know from Subsection 10.1.1 that $[a,b]$ is a first-order small set, it follows that the homogeneous Markov chain $(X_n)_{n\geq 0}$ is uniformly ergodic. More precisely, we know from Example 3.7 that

$$\forall n \geq 1,\ \forall x \in [a,b], \quad \|P^n(x,\cdot) - q_\mu\|_{TV} \leq 2(1-\nu(1_{\mathbb{R}}))^n.$$

10.1.2 A basic modulated drift condition

Let us introduce the following moment condition on the probability measure μ_0:

$$\int_{\mathbb{R}} \ln(1+|x|)\,\mu_0(dx) < \infty. \tag{183}$$

Consider the Lyapunov function $x \mapsto V_0(x) := 1 + \ln(1+|x|)$ on $\mathbb{R}$. Note that the moment condition (183) is equivalent to $\mathbb{E}[V_0(W_1)] < \infty$. Without lost of generality it can be assumed that $\mu(1_{\mathbb{X}}) < 2$. Indeed, if $\mu(1_X) \geq 2$, it suffices to replace V_0 with the Lyapunov function $x \mapsto \mu(1_{\mathbb{R}})(1+\ln(1+|x|))/2$ in the following computations. We get from (180) that

$$\begin{aligned} \forall x \in \mathbb{R}, \quad (PV_0)(x) - V_0(x) &= \mathbb{E}\left[\ln \frac{|xU_1 + (1-U_1)W_1|}{1+|x|}\right] \leq \mathbb{E}[Z(x)] \\ &\qquad \text{with } Z(x) := \ln \frac{|x|U_1 + (1-U_1)|W_1|}{1+|x|}. \end{aligned} \tag{184}$$

We have

$$\lim_{|x|\to+\infty} Z(x) \overset{\mathbb{P}-a.s.}{=} \ln U_1 \tag{185}$$

and, for any $x \in \mathbb{R}$, $Z(x) \leq \ln(1+|W_1|)$ using $0 \leq U_1 \leq 1$. Since the r.v. $\ln(1+|W_1|)$ is integrable from (183), the dominated convergence theorem provides

$$\lim_{|x|\to+\infty} \mathbb{E}[Z(x)] = \mathbb{E}[\ln U_1] = -\frac{1}{\mu(1_{\mathbb{R}})} < -1 \tag{186}$$

since $\mu(1_{\mathbb{R}}) < 2$. It follows that there exists $\kappa > 0$ such that

$$\forall x \in [-\kappa,\kappa]^c, \quad (PV_0)(x) - V_0(x) \leq -1.$$

Since the function $x \mapsto (PV_0)(x)$ is bounded on the compact $[-\kappa,\kappa]$, there exists $b_0 > 0$ such that

$$\forall x \in \mathbb{R}, \quad (PV_0)(x) \leq V_0(x) - 1_{\mathbb{R}} + b_0\, 1_S$$

where $S := [-\kappa,\kappa]$. Thus, the transition kernel P satisfies the modulated drift condition $\boldsymbol{D}_{1_S}(V_0, 1_{\mathbb{R}})$, Since it also satisfies Condition $(\boldsymbol{M}_{\nu,1_S})$ from Subsection 10.1.1, the following assertions follow from Section 5:

(i) The $P-$harmonic functions are constant on $\mathbb{R}$

(ii) P is irreducible and recurrent.

(iii) $\pi_R := \mu_R(1_{\mathbb{X}})^{-1}\mu_R$ (see (26)) is the unique $P-$invariant probability measure on $(\mathbb{R}, \mathcal{X})$, we have $\pi_R(1_S) > 0$, and P is Harris-recurrent.

Furthermore, it follows fom [FT89, Th. 1] that π_R is the probability distribution of the mean of a Dirichlet process with (measure) parameter μ. Such a context is briefly discussed in Subsection 10.1.5. Accordingly, the probability measure π_R is denoted by q_μ in the sequel.

(iv) The following convergence in total variation of Theorem 4.7 holds

$$\forall x \in \mathbb{R}, \quad \lim_{n\to+\infty} \|\delta_x P^n - q_\mu\|_{TV} = 0.$$

(v) Let $R := P - 1_S \otimes \nu$ be the residual kernel associated with P. For any $g \in \mathcal{B}$, we have $\widetilde{g} := \sum_{k=0}^{+\infty} R^k g \in \mathcal{B}_{V_0}$ and

$$\|\widetilde{g}\|_{V_0} \leq (1+d_0)\|g\|_{1_\mathbb{R}} \quad \text{with} \quad d_0 := \max\left(0\,,\, \frac{b_0 - \nu(V_0)}{\nu(1_\mathbb{X})}\right)$$

where b_0 is the positive constant given in $\boldsymbol{D}_{1_S}(V_0, 1_\mathbb{R})$. For any $g \in \mathcal{B}$ such that $q_\mu(g) = 0$, the function $\widetilde{g}$ satisfies Poisson's equation

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

10.1.3 Polynomial ergodicity

Let us introduce the following moment condition on the probability measure μ_0:

$$\exists s > 1, \qquad \int_\mathbb{R} \ln^s(1+|x|)\,\mu_0(dx) < \infty. \tag{187}$$

Consider the following Lyapunov function $x \mapsto V(x) := 1 + \ln^s(1+|x|)$ on $\mathbb{R}$. The moment condition (187) reads as $\mathbb{E}[V(W_1)] < \infty$. Next we get from (180)

$$\forall x \in \mathbb{R}, \quad (PV)(x) - V(x) \leq \mathbb{E}\left[\big(\ln(1+|x|) + Z(x)\big)^s - \big(\ln(1+|x|)\big)^s\right] \tag{188}$$

with $Z(x)$ defined in (184) and satisfying the convergences (185) and (186). The objective is to show that there exists $c > 0$ such that, for $|x|$ large enough,

$$(PV)(x) - V(x) \leq -c\,V(x)^{\frac{s-1}{s}}.$$

From (188), it is sufficient to prove that there exists $a > 0$ such that

$$\lim_{|x|\to+\infty} \mathbb{E}\left[\frac{\big(\ln(1+|x|) + Z(x)\big)^s - \big(\ln(1+|x|)\big)^s}{V(x)^{\frac{s-1}{s}}}\right] = -a.$$

Since $V(x)^{\frac{s-1}{s}} \sim \ln^{s-1}(1+|x|))$ as $|x| \to +\infty$, it is equivalent to prove that there exists $b > 0$ such that

$$\lim_{|x|\to+\infty} \mathbb{E}\left[\frac{\big(\ln(1+|x|) + Z(x)\big)^s - \big(\ln(1+|x|)\big)^s}{\ln^{s-1}(1+|x|)}\right] = -b. \tag{189}$$

Following the same way as in Subsection 10.1.2, it is proved below that

$$\lim_{|x|\to+\infty} \frac{\big(\ln(1+|x|) + Z(x)\big)^s - \big(\ln(1+|x|)\big)^s}{\ln^{s-1}(1+|x|)} \xrightarrow[|x|\to+\infty]{a.s.} s\,\ln U_1. \tag{190}$$

We have from the mean value theorem that, for every $x \in \mathbb{R}$, there exists $\xi(x)$ between $\ln(1+|x|)$ and $\ln(1+|x|) + Z(x)$ such that

$$\big(\ln(1+|x|) + Z(x)\big)^s - \big(\ln(1+|x|)\big)^s = sZ(x)\,\xi(x)^{s-1}.$$

Note that $\lim_{|x|\to+\infty} Z(x)/\ln(1+|x|) = 0$ from (185). Thus we have

$$\lim_{|x|\to+\infty} \frac{\xi(x)^{s-1}}{\ln^{s-1}(1+|x|)} = 1$$

and the $\mathbb{P}-$almost sure convergence (190) also follows from (185). Next, using again (185) and $\ln U_1 < 0$, we get $|\xi(x)^{s-1}|/\ln^{s-1}(1+|x|) \le 1$ for x large enough. We know from Subsection 10.1.2 that $Z(x) \le \ln(1+|W_1|)$ which is integrable from (187). Then it follows from (190) and the dominated convergence theorem that (189) holds with $b = \mathbb{E}[s \ln U_1] = s/\mu(1_{\mathbb{R}})$.

Since the function $x \mapsto (PV)(x)$ is bounded on any compact, we obtained that there exists $c > 0, b > 0$ such that

$$\forall x \in \mathbb{R}, \quad (PV)(x) \le V(x) - c\, V(x)^{\frac{s-1}{s}} + b\, 1_S$$

where $S := [-\kappa, \kappa]$ for some $\kappa > 0$. We know from Subsection 10.1.1 that S is a first-order small set, so that P satisfies Condition $(\boldsymbol{M}_{\nu,1_S})$. From Proposition 8.6 (see (144)), P also satisfies the nested modulated drift conditions $\boldsymbol{D}_{1_S}(V_0 : V_m)$ with $m := \lfloor s \rfloor \ge 1$, $V_m := 1_{\mathbb{X}}, V_0 := a_0 V$ for some $a_0 > 0$. Note that the set S in $(\boldsymbol{M}_{\nu,1_S})$ and $\boldsymbol{D}_{1_S}(V_0 : V_m)$ may be chosen large enough in order to satisfies $q_\mu(1_S) > 1/2$, so that Condition (121) holds from Proposition 8.5. If $s \ge 2$ then for any measurable and bounded function $g : \mathbb{X} \to \mathbb{R}$ and for any $x \in \mathbb{X}$, Theorem 8.2 provides a bound for

$$\sum_{n=0}^{+\infty} (n+1)^{\lfloor s \rfloor - 2} |(P^n g)(x) - q_\mu(g)|.$$

For instance the bounds (139) in case $s \in [2, 3)$ or the bounds (141) in case $s \ge 3$ can be used.

10.1.4 Geometric ergodicity

Assume that there exists $s > 0$ such that

$$\int_{\mathbb{R}} |v|^s \mu_0(dv) > \infty. \tag{191}$$

Introduce the Lyapunov function $x \mapsto V(x) := 1 + |x|^s$ on $\mathbb{R}$. Note that the moment condition (191) is equivalent to $\mathbb{E}[V(W_1)] < \infty$. Then, it follows from (180) that

$$\forall x \in \mathbb{R}, \quad (PV)(x) - V(x) \quad \le \quad |x|^s \left(\mathbb{E}\big[Y(x)^s\big] - 1\right)$$

where $Y(x) := U_1 + (1 - U_1)|W_1|/|x|$. Since $Y(x) \le 1 + |W_1|$ for every $|x| \ge 1$ and the r.v. $Y(x)$ $\mathbb{P}-$a.s. converges to U_1, it follows from the dominated convergence theorem that

$$\lim_{|x| \to +\infty} \mathbb{E}\big[Y(x)^s\big] = \mathbb{E}\big[U_1{}^s\big] = \frac{\mu(1_{\mathbb{R}})}{\mu(1_{\mathbb{R}}) + s} < 1.$$

Let any $\delta \in (\mathbb{E}[U_1{}^s], 1)$. Then there exits $\kappa > 0$ such that

$$\forall |x| > \kappa, \quad (PV)(x) - V(x) \le |x|^s (\mathbb{E}\big[U_1{}^s\big] - 1) \le (\delta - 1) V(x).$$

Next, since PV is bounded on the compact $S := [-\kappa, \kappa]$, there exists $b > 0$ such that $(PV)(x) \le b$ for any $x \in S$. Thus minorization condition $(\boldsymbol{M}_{\nu,1_S})$ holds for this compact set S. Moreover the above inequality shows that Condition $\boldsymbol{D}_{1_S}(V, V_1)$ holds with $V_1 := (1-\delta)V$. Equivalently, $\boldsymbol{G}_{1_S}(\delta, V)$ holds true.

Since P satisfies Conditions $(\boldsymbol{M}_{\nu,1_S})$-$\boldsymbol{G}_{1_S}(\delta, V)$ and is strongly aperiodic, then the following assertions hold from Sections 6 and 9:

(i) P is $V-$geometrically ergodic, that is

$$\exists \rho \in (0,1),\ \exists c_\rho > 0,\ \forall g \in \mathcal{B}_V(\mathbb{C}),\ \forall n \geq 1, \quad \|P^n g - q_\mu(g) 1_{\mathbb{X}}\|_V \leq c_\rho\, \rho^n \|g\|_V. \quad (192)$$

(ii) Let $g \in \mathcal{B}_V$ be such that $q_\mu(g) = 0$. Then the function series $\mathfrak{g} := \sum_{k=0}^{+\infty} P^k g$ absolutely converges in $(\mathcal{B}_V, \|\cdot\|_V)$ and

$$\|\mathfrak{g}\|_V \leq c_\rho (1-\rho)^{-1} \|g\|_V.$$

Note that $\mathfrak{g}$ is $q_\mu-$centred and satisfies Poisson's equation $(I-P)\mathfrak{g} = g$. From Corollary 6.1, we also have the following alternative bound:

$$\|\mathfrak{g}\|_V \leq \frac{(1+d_0)\,(1+q_\mu(V))}{1-\delta}\, \|g\|_V \qquad \text{with } d_0 := \max\left(0\,,\ \frac{b-\nu(V)}{\nu(1_{\mathbb{X}})(1-\delta)}\right).$$

(iii) The conclusions of Corollary 9.5 apply to

$$\varrho_V := \inf\big\{\rho \in (0,1) \text{ such that Property (192) holds}\big\}$$

10.1.5 Further comments and bibliographic discussion

Let us briefly recall here the context of the present Subsection 10.1. Let $\mu \in \mathcal{M}_{+,b}^*$ be any finite positive measure on $(\mathbb{X} := \mathbb{R}, \mathcal{B}(\mathbb{R}))$ where $\mathcal{B}(\mathbb{R})$ is the Borel $\sigma-$algebra on $\mathbb{R}$. A random probability measure on the set of probability measures on $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$, is called a Dirichlet process and is denoted by $\mathcal{D}_\mu$, if for any $k \geq 1$ and any finite measurable partition $E_1, \ldots, E_k$ of $\mathbb{R}$, the random vector $\big(\mathcal{D}_\mu(1_{E_1}), \mathcal{D}_\mu(1_{E_2}), \ldots, \mathcal{D}_\mu(1_{E_k})\big)$ has a Dirichlet distribution with parameters k and $(\mu(1_{E_1}), \ldots, \mu(1_{E_k}))$ (see [Fer73]). A Dirichlet process is of fundamental importance in Bayesian non-parametric statistics. For an overview, we refer to [GvdV17, Chap. 4] and to [Teh17, Mur23] in a machine learning context. Specifically, we are interested here in the so-called mean of a Dirichlet process, that is in the random variable

$$M_{\mathcal{D}_\mu} := \int_{\mathbb{R}} x\, \mathcal{D}_\mu(dx) \quad (193)$$

with probability distribution

$$\forall A \in \mathcal{B}(\mathbb{R}), \quad q_\mu(1_A) := \mathbb{P}\{M_{\mathcal{D}_\mu} \in A\}.$$

Let us mention that the analysis of the probability distribution of the mean functional $\int_{\mathbb{R}} g(x)\, \mathcal{D}_\mu(dx)$ for a measurable function $g : \mathbb{X} \to \mathbb{R}$ can be reduced to that of the mean (193) since the random variables $\int_{\mathbb{X}} g(x)\, \mathcal{D}_\mu(dx)$ and $\int_{\mathbb{X}} x\, (\mathcal{D}_{\mu\circ g^{-1}})(dx)$ have the same probability distribution (e.g. see [GvdV17, p 83]). We refer to the survey [LP09] on the mean functional of the Dirichlet process and to [Tor23] for a recent contribution on the central limit theorem. As mentioned in [LP09, p. 49], the mean functional of a Dirichlet process also has interested in other topics in mathematics.

Let $\mu_0 := \mu/\mu(1_{\mathbb{R}})$ be the probability measure associated with μ, which is assumed not to be a Dirac measure. Then we know that q_μ has a p.d.f. with respect to the Lebesgue measure on $\mathbb{R}$. But this p.d.f. is of difficult use. We refer to [LP09, Section 2.2] for a discussion on the exact probability distribution of $M_{\mathcal{D}_\mu}$. A natural issue is to design an MCMC algorithm to get (approximated) samples of this probability distribution. This is addressed in [FT89] where the Markov chain defined in (178) and having invariant probability distribution q_μ, is introduced and analyzed. This work is continued in [GT01, JT02], where convergence rates of the algorithm are provided. Below are a few more facts to complete the results presented in the previous subsections.

- The random variable $\int_{\mathbb{R}} |x| \, \mathcal{D}_\mu(dx)$ is $\mathcal{D}_\mu-$a.s. finite if, and only if, Condition (183) holds (see [FT89, Th. 4]).
- There exits $\rho > 1$ such that

$$\forall x \in \mathbb{X}, \quad \rho^n \|P^n(x, \cdot) - q_\mu\|_{TV} \xrightarrow[n \to +\infty]{} 0$$

  if, and only if, Condition (191) holds (see [JT02, Th. 2.3]).
- Note that the two previous results are also obtained in a more general framework of measure-valued Markov chains in [FGW12, Ths 4 and 5].
- Let us introducce the Lyapunov function $x \mapsto V_\beta(x) := \max(\log^\beta(1+|x|), 1)$. If Condition (187) holds, then (see [JT02, Th. 3.1])

$$\forall \beta \in [0, s-1], \quad \forall x \in \mathbb{R}, \quad n^\beta \|P^n(x, \cdot) - q_\mu\|_{V_{s-1-\beta}} \xrightarrow[n \to +\infty]{} 0.$$

  In particular, we have

$$\forall x \in \mathbb{R}, \quad n^{s-1} \|P^n(x, \cdot) - q_\mu\|_{TV} \xrightarrow[n \to +\infty]{} 0.$$

  Some converse statements also hold (see [JT02, Th. 3.1] for details).

10.2 Metropolis-Hasting's Markov chain

Let $\mathbb{X}$ be an open subset of $\mathbb{R}^d$, where $\mathbb{R}^d$ is equipped with the Euclidean norm $\|\cdot\|$. The set $\mathbb{X}$ is assumed to be connected w.r.t. the topology induced on $\mathbb{X}$ by the norm $\|\cdot\|$. Finally let $\mathcal{X}$ be the Borel $\sigma-$algebra on $\mathbb{X}$. First, consider some probability measure π on $(\mathbb{X}, \mathcal{X})$ which has a p.d.f., also denoted by π, w.r.t. the Lebesgue measure on $\mathbb{R}^d$, i.e. $\pi(dx) = \pi(x)dx$. The function π is assumed to be positive on $\mathbb{X}$. Second let K be a transition kernel on $(\mathbb{X}, \mathcal{X})$ such that each probability measure $K(x, dy)$ has a p.d.f. $y \mapsto k(x, y)$ w.r.t. the Lebesgue measure on $\mathbb{R}^d$. It is assumed that K is such that $\{(x,y) \in \mathbb{X}^2, \pi(x)k(x,y) > \pi(y)k(y,x)\}$ is of positive Lebesgue measure. This implies that K is not reversible w.r.t. the probability measure π (see Subsection 9.4). Let us introduce the Metropolis-Hasting (MH) Markov chain with state space $\mathbb{X}$ and transition kernel P given by (see Subsection 10.2.4 for contextual comments):

$$\forall x \in \mathbb{X}, \forall A \in \mathcal{X}, \quad P(x, A) := \int_A p(x,y)dy + r(x)\delta_x(A) \tag{194a}$$

where

$$\forall (x,y) \in \mathbb{X}^2, \quad p(x,y) := \frac{1}{\pi(x)} \min\big(\pi(x)k(x,y), \pi(y)k(y,x)\big) \tag{194b}$$

$$\forall x \in \mathbb{X}, \quad r(x) := 1 - \int_{\mathbb{X}} p(x,y)dy. \tag{194c}$$

Note that $r(x)$ is the probability for staying in x. From (194b), observe that

$$\forall (x,y) \in \mathbb{X}^2, \quad \pi(x)p(x,y) = \min\big(\pi(x)k(x,y), \pi(y)k(y,x)\big) = \pi(y)p(y,x)$$
$$\pi(x)r(x)\delta_x(y) = \pi(y)r(y)\delta_y(x)$$

so that $\pi(dx)P(x,dy) = \pi(dy)P(y,dx)$ and the Markov kernel P is $\pi-$reversible. It follows that π is a $P-$invariant probability measure.

In order to simplify the presentation, the following positivity condition is assumed throughout this subsection:

(P) The functions π and $k(\cdot,\cdot)$ are assumed to be positive on $\mathbb{X}$ and $\mathbb{X}^2$ respectively.

Under Condition (P), the function $p(\cdot,\cdot)$ is positive on $\mathbb{X}^2$ so that we obtain from (194a)-(194b) that, for any set $A \in \mathcal{X}$ with positive Lebesgue measure,

$$\forall x \in \mathbb{X}, \quad P(x,A) \geq \int_A p(x,y)dy > 0. \tag{195}$$

Note that this implies that $r(x) < 1$ for every $x \in \mathbb{X}$.

In Subsection 10.2.1, under Condition (P) and assuming that functions π and $k(\cdot,\cdot)$ are continuous, a minorization Condition $(M_{\nu_S,1_S})$ for some positive measure ν_S is shown to hold for any compact set $S \subset \mathbb{X}$ of positive Lebesgue measure. Next convergence conditions are provided in the two following specific cases:

1. The Independent Metropolis-Hastings (IMH) Markov chain defined by the following condition: $\forall (x,y) \in \mathbb{X}^2, k(x,y) \equiv q(y)$ for some positive measurable function $q(\cdot)$ on $\mathbb{X}$ such that $\int_{\mathbb{X}} q(y)dy = 1$, that is the function $(x,y) \mapsto k(x,y)$ only depends on the second variable.

2. The symmetric Random Walk Metropolis-Hastings (RWMH) Markov chain defined by: $\forall (x,y) \in \mathbb{X}^2, k(x,y) \equiv q(\|x-y\|)$ for some positive measurable function $q(\cdot)$ on $[0,+\infty)$ such that $\int_{\mathbb{X}} q(\|u\|)du = 1$, that is the function $(x,y) \mapsto k(x,y)$ only depends on the distance between x and y.

We essentially follow [Tie94, MT96, RR96, JH00] for the geometric ergodic case and [FM00, JR07] for the polynomial ergodic one.

10.2.1 Minorization condition

Recall that

$$\forall x \in \mathbb{X}, \quad P(x,dy) \geq p(x,y)dy$$

with $p(x,y)$ given in (194b). Let S be any compact set of $\mathbb{X}$ with positive Lebesgue measure. As in [MT96, Lem. 2.1], in addition to Condition (P), let us consider the following continuity conditions on the functions π and $k(\cdot,\cdot)$:

(C) The functions π and $k(\cdot,\cdot)$ are continuous on $\mathbb{X}$ and $\mathbb{X}^2$ respectively.

Then $\pi_S := \max_{x\in S}\pi(x) \in (0, +\infty)$ and $k_S := \min_{(x,y)\in S^2} k(x, y) > 0$. It follows from (194b) that

$$\begin{aligned}\forall x \in S, \forall y \in \mathbb{X}, \quad p(x,y) &= \min\big(k(x,y), k(y,x)\frac{\pi(y)}{\pi(x)}\big) \\ &\geq \min\big(k(x,y), k(y,x)\frac{\pi(y)}{\pi(x)}\big)1_S(y) \\ &\geq k_S\frac{\pi(y)}{\pi_S}\,1_S(y)\end{aligned}$$

since $1 \geq \pi(y)/\pi_S$ for any $y \in S$. Thus, under Conditions (P)-(C), S is a first-order small set for P with minorizing measure $\nu_S(dy) := (k_S/\pi_S)1_S(y)\pi(dy) \in \mathcal{M}^*_{+,b}$. Next consider the associated residual kernel $R := P - 1_S \otimes \nu_S$. Since P satisfies $(M_{\nu_S,1_S})$ and $\pi(dx)$ is a $P-$invariant probability measure such that $\pi(1_S) > 0$, it follows from Theorem 3.6 that

$$\pi_R := \mu_R(1_{\mathbb{X}})^{-1}\,\mu_R \quad \text{with} \quad \mu_R := \sum_{k=0}^{+\infty}\nu_S R^k \in \mathcal{M}^+_{*,b}$$

is $P-$invariant with $\mu_R(1_S) = 1$. We have $P(x, S) > 0$ for every $x \in \mathbb{X}$ from (195), so that P is irreducible (see (29)). It follows from Theorem 3.14 that π is the unique $P-$invariant probability measure, thus $\pi \equiv \pi_R$. Finally note that $\nu_S(1_S) = (k_S/\pi_S)\pi(1_S) > 0$, thus P is (strongly) aperiodic.

Now, let us check that $h^\infty_R := \lim_n \downarrow R^n 1_{\mathbb{X}} = 0$ (see (20)). The function h^∞_R is bounded and satisfies $Rh^\infty_R = h^\infty_R$. Since $\mu_R(1_S) = 1$, we have $\pi(h^\infty_R) = 0$ and $\nu(h^\infty_R) = 0$ from (25). Thus, we have $h^\infty_R(x) = 0$ for $\pi-$almost $x \in \mathbb{X}$ and $Ph^\infty_R = Rh^\infty_R + \nu(h^\infty_R)1_S = h^\infty_R$, i.e. h^∞_R is a bounded $P-$harmonic function. Note that $\{h^\infty_R > 0\}$ is also negligible for Lebesgue's measure on $\mathbb{X}$ since $\pi(dy) = \pi(y)dy$ with positive p.d.f π on $\mathbb{X}$ by hypothesis. Then it follows from (194a) that

$$\forall x \in \mathbb{X}, \quad (Ph^\infty_R)(x) \;=\; \int_{\mathbb{X}} p(x,y)h^\infty_R(y)dy + r(x)h^\infty_R(x) = r(x)h^\infty_R(x).$$

Using $Ph^\infty_R = h^\infty_R$, we get that $(1 - r(x))h^\infty_R(x) = 0$ for every $x \in \mathbb{X}$. Since $r(x) < 1$ for any $x \in \mathbb{X}$, we obtain that $h^\infty_R = 0$.

Thus, under Conditions (P)-(C), the transition kernel P of the Metropolis-Hastings chain defined by (194a)–(194b)–(194c) has the following properties from the results of Section 4:

(i) P is irreducible and aperiodic.

(ii) The probability measure π is the unique $P-$invariant probability measure. Moreover, we have $\pi \equiv \pi_R := \mu_R(1_{\mathbb{X}})^{-1}\mu_R$ (see (26)) with the residual kernel $R := P - 1_S \otimes \nu_S$ where $S \subset \mathbb{X}$ is any compact of positive Lebesgue measure.

(iii) P is Harris-recurrent.

(iv) The $P-$harmonic functions are constant on $\mathbb{X}$.

(v) The following convergence in total variation holds

$$\forall x \in \mathbb{X}, \quad \lim_{n\to+\infty}\|\delta_x P^n - \pi\|_{TV} = 0.$$

If the state space $\mathbb{X}$ is bounded with π and $k(\cdot,\cdot)$ assumed to be continuous on the respective closure $\overline{\mathbb{X}}$ and $\overline{\mathbb{X}}^2$ of $\mathbb{X}$ and $\mathbb{X}^2$, then it is clear that P satisfies the minorization condition $(M_{\nu_{\mathbb{X}},1_{\mathbb{X}}})$ with $\nu_{1_{\mathbb{X}}} := (k_{\mathbb{X}}/\pi_{\mathbb{X}})\pi$ where $\pi_{\mathbb{X}} := \max_{x\in\overline{\mathbb{X}}}\pi(x), k_{\mathbb{X}} := \min_{(x,y)\in\overline{\mathbb{X}}^2} k(x,y)$. Then we know from Example 3.7 that

$$\forall n \geq 1,\ \forall x \in \mathbb{X}, \quad \|P^n(x,\cdot) - \pi\|_{TV} \leq 2(1 - k_{\mathbb{X}}/\pi_{\mathbb{X}})^n.$$

When $\mathbb{X}$ is unbounded, the uniform ergodicity does not hold in general for MH kernels. For instance, the RWMH Markov chain is never uniformly ergodic, see [MT96, Th. 3.1]. However, the following condition is introduced in [Tie94, Cor. 4],[MT96, Th. 2.1] for the IMH Markov chain associated with function $q(\cdot)$ to be uniformly ergodic: there exits a constant $M > 1$ such that

$$\forall x \in \mathbb{X}, \quad \pi(x) \leq M q(x). \tag{196}$$

Indeed, under Condition (196) for the IMH Markov chain, we have from (194b)

$$\forall (x,y) \in \mathbb{X}^2 \quad p(x,y) = \min\big(q(y), \pi(y)q(x)/\pi(x)\big) \geq \frac{1}{M}\,\pi(y).$$

Thus the whole state space $\mathbb{X}$ is a first-order small set with minorizing measure $\nu := \pi/M$. Since $\nu(1_{\mathbb{X}}) = 1/M$, we have from Example 3.7 that

$$\forall n \geq 1,\ \forall x \in \mathbb{X}, \quad \|P^n(x,\cdot) - \pi\|_{TV} \leq 2(1 - 1/M)^n.$$

In fact, it follows from [Wan22, Th. 2] that the previous rate of convergence is exact, that is

$$\forall n \geq 1, \quad \sup_{x\in\mathbb{X}} \|P^n(x,\cdot) - \pi\|_{TV} = 2(1 - 1/M)^n.$$

10.2.2 Geometric ergodicity of an RWMH Markov kernel

The functions $p(\cdot,\cdot), r(\cdot)$ in (194b)-(194c) are as follows for the RWMH Markov kernel

$$\forall (x,y) \in \mathbb{X}^2,\ p(x,y) := a(x,y)\,k(x,y) \ \text{ with } a(x,y) := \frac{1}{\pi(x)}\min\big(\pi(x),\pi(y)\big) \in [0,1] \tag{197a}$$

$$\forall x \in \mathbb{X},\ r(x) := \int_{\mathbb{X}} (1 - a(x,y))\,k(x,y)dy = \int_{{A_x}^c} (1 - a(x,y))\,k(x,y)dy \tag{197b}$$

where $A_x := \{(x,y) \in \mathbb{X}^2 : a(x,y) = 1\} = \{(x,y) \in \mathbb{X}^2 : \pi(y) \geq \pi(x)\}$. The basic assumptions (P)-(C) on π and $k(\cdot,\cdot)$ read as:

- The p.d.f. π is positive and continuous on $\mathbb{X}$.
- For any $(x,y) \in \mathbb{X}^2$, $k(x,y) = q(\|x-y\|)$ with q positive and continuous on $[0,+\infty)$.

Here, the specific case of super-exponential p.d.f. π is considered, that is π satisfies the following additional assumption:

(SE) The p.d.f. π has continuous first derivatives on $\mathbb{X}$ such that

$$\lim_{\|x\|\to+\infty} \langle \frac{x}{\|x\|}, \nabla \ln \pi(x)\rangle = -\infty \tag{198}$$

where ∇ is the gradient operator and $\langle\cdot,\cdot\rangle$ is the scalar product associated with the Euclidean norm $\|\cdot\|$.

Under Conditions (P)-(C)-(SE), the following additional assumption on the Markov kernel K

$$\liminf_{\|x\|\to+\infty} K(x,A_x) = \liminf_{\|x\|\to+\infty} \int_{A_x} k(x,y)dy > 0 \tag{199}$$

is used below to prove that P defined by (194a)–(194b)–(194c) satisfies Conditions ($\boldsymbol{M}_{\nu,1_S}$)–$\boldsymbol{G}_{1_S}(\delta,V)$ for some compact set S of positive Lebesgue measure and some Lyapunov function V.

Let us introduce the following function on $\mathbb{X}$: $V(x) := c\pi(x)^{-1/2}$. Since π satisfies (P)-(C) and vanishes in tails from (198), c may be chosen in order to have $V \geq 1_{\mathbb{X}}$, so that V is a Lyapunov function. Next, we have from (194a) and (197a)-(197b) and the definition of the set A_x that

$$\begin{aligned}\forall x\in\mathbb{X},\quad (PV)(x) &= \int_{\mathbb{X}} a(x,y)\,k(x,y)V(y)dy + V(x)\int_{{A_x}^c}\big(1-a(x,y)\big)k(x,y)dy \\ &= \int_{A_x} k(x,y)V(y)dy + V(x)\int_{{A_x}^c}\left(1-a(x,y)+a(x,y)\frac{V(y)}{V(x)}\right)k(x,y)dy.\end{aligned}$$

Then, it follows from the definitions of the set A_x and of the Lyapunov function V that

$$\forall x\in\mathbb{X},\quad \frac{(PV)(x)}{V(x)} = \int_{A_x}\left(\frac{\pi(x)}{\pi(y)}\right)^{1/2} k(x,y)dy + \int_{{A_x}^c}\left(1-\frac{\pi(y)}{\pi(x)}+\left(\frac{\pi(y)}{\pi(x)}\right)^{1/2}\right)k(x,y)dy. \tag{200}$$

Thus, since $0<\pi(y)<\pi(x)$ on ${A_x}^c$ we have

$$\forall x\in\mathbb{X},\quad \frac{(PV)(x)}{V(x)} \leq 2. \tag{201a}$$

Next, under Conditions (P)-(C)-(SE), it can be shown that

$$\limsup_{\|x\|\to+\infty}\frac{(PV)(x)}{V(x)} = \limsup_{\|x\|\to+\infty}\int_{{A_x}^c} k(x,y)dy = \limsup_{\|x\|\to+\infty} K(x,{A_x}^c) = 1-\liminf_{\|x\|\to+\infty} K(x,A_x) < 1$$

from (199). In fact, this property follows from geometric considerations involving suitable subsets of $\mathbb{X}$, depending on x and constructed from Condition (199). Details of this geometric study and the proof of the above property are given in [JH00, p 350-352]. Therefore, for $R>0$ large enough, there exists $\delta\in(0,1)$ such that

$$\text{for any } x\in\mathbb{X} \text{ such that } \|x\|\geq R,\quad \frac{(PV)(x)}{V(x)}\leq\delta. \tag{201b}$$

Finally, P satisfies Conditions ($\boldsymbol{M}_{\nu_S,1_S}$) since the compact ball $S := \{x\in\mathbb{X} : \|x\|\leq R\}$ is a small-set with respect to some minorizing measure ν_S from Subsection 10.2.1. Moreover, the inequalities (201a)-(201b) show that P satisfies the drift condition $\boldsymbol{G}_{1_S}(\delta,V)$ with positive constant b given by:

$$b := \sup_{\|x\|\leq R}\big[(PV)(x)-\delta V(x)\big].$$

Consequently, P is reversible, satisfies Conditions ($\boldsymbol{M}_{\nu_S,1_S}$)–$\boldsymbol{G}_{1_S}(\delta,V)$ and is strongly aperiodic. Then the following assertions hold true from Sections 6 and 9 (see Theorem 6.2, Corollary 9.5, and Remark 9.12):

(i) P is $V-$geometrically ergodic, that is

$$\exists \rho \in (0,1),\ \exists c_\rho > 0,\ \forall g \in \mathcal{B}_V(\mathbb{C}),\ \forall n \geq 1, \quad \|P^n g - \pi(g)1_{\mathbb{X}}\|_V \leq c_\rho\, \rho^n \|g\|_V.$$

(ii) For any $g \in \mathcal{B}_V$ such that $\pi(g) = 0$, the $\pi-$centred function series $\mathfrak{g} := \sum_{k=0}^{+\infty} P^k g$ absolutely converges in $(\mathcal{B}_V, \|\cdot\|_V)$ with

$$\|\mathfrak{g}\|_V \leq c_\rho(1-\rho)^{-1}\|g\|_V.$$

Note that $\mathfrak{g}$ is $\pi-$centred and satisfies Poisson's equation $(I-P)\mathfrak{g} = g$. From Corollary 6.1, we also have the following bound:

$$\|\mathfrak{g}\|_V \leq \frac{(1+d_0)\,(1+\pi(V))}{1-\delta}\,\|g\|_V \quad \text{with } d_0 := \max\left(0\,,\ \frac{b-\nu(V)}{\nu(1_{\mathbb{X}})(1-\delta)}\right).$$

(iii) The conclusions of Remark 9.12 apply to

$$\varrho_V := \inf\big\{\rho \in (0,1) \text{ such that Property (192) holds}\big\}$$

Likewise, all material of Sections 9.4 and 9.6 on the geometric ergodicity in $\mathbb{L}^2(\pi)$ of reversible Markov kernels is relevant for P.

10.2.3 Polynomial ergodicity of an RWMH Markov kernel

As shown in [JT03], exponential or lighter tails of π is necessary and essentially sufficient for geometric ergodicity of the RWMH Markov kernel. Here polynomial tails for probability measure π are considered, in which case polynomial ergodicity of the MH Markov kernel is the best convergence rate we can expect. In contrast to geometric case, it turns out that the choice of the p.d.f. $q(\|\cdot\|)$ has a direct impact on the polynomial convergence rates (see [JR07, Section 3]). As in [JR07], the discussion is restricted to the case when π is spherically symmetric and the probability measure with p.d.f. $q(\|\cdot\|)$ is heavy-tailed. The assumptions on the set $\mathbb{X} \subset \mathbb{R}^d$ are those presented at the beginning of Subsection 10.2. Recall that a positive function $f(\cdot)$ on $(0,+\infty)$ is said to be a normalized slowly varying function if $f(u) = c\exp\big(\int_a^u (\varepsilon(v)/v)dv\big)$ for $u \geq a$ with some positive constants a, c and $\lim_{v\to+\infty}\varepsilon(v) = 0$ (see [BGT87, p. 15]).

- Let π be a continuous strictly positive and spherically symmetric density function on $\mathbb{X}$ which has the following representation for $\|x\|$ large enough

  $$\pi(x) = \frac{f(\|x\|)}{\|x\|^{d+r}},$$

  where $r > 0$ and f is a normalized slowly varying function such that $\lim_{\|x\|\to+\infty} f(\|x\|) \in (0,+\infty)$.

- The function $q(\cdot)$ is positive on $[0,+\infty)$ and there exists $\eta \in (0,2)$ such that, for $\|x\|$ large enough,

  $$q(\|x\|) = \frac{f_q(\|x\|)}{\|x\|^{d+\eta}},$$

  where f_q is a normalized slowly varying function such that $\lim_{\|x\|\to+\infty} f_q(\|x\|) \in (0,+\infty)$.

Then, for any $s \in (\eta, r+\eta)$, the following Jarner-Roberts's drift condition

$$PV \le V - c\, V^{\frac{s-\eta}{s}} + b\, 1_S \tag{202}$$

holds with $V(x) := \max(\|x\|^s, 1)$, some positive constants c, b and a centred compact ball S (see [JR07, Prop. 6 and p. 811-812] for the details). Thus, from Proposition 8.6 (see (144)), P also satisfies the nested modulated drift conditions $\boldsymbol{D}_{1_S}(V_0 : V_m)$ where $m := \lfloor s/\eta \rfloor \ge 1$, $V_0 = aV$ for some $a > 1$ and $V_m := 1_{\mathbb{X}}$. We know from Subsection 10.2.1 that S is a first-order small set so that the RWMH Markov kernel P satisfies $(\boldsymbol{M}_{\nu,1_S})$–$\boldsymbol{D}_{1_S}(V_0 : V_m)$. Note that the set S may be chosen large enough in order to satisfy $\pi(1_S) > 1/2$, so that Condition (121) holds from Proposition 8.5. If $s \ge 2\eta$, then for any measurable and bounded function $g : \mathbb{X} \to \mathbb{R}$ and for any $x \in \mathbb{X}$, Theorem 8.2 provides a bound for

$$\sum_{n=0}^{+\infty} (n+1)^{\lfloor s/\eta \rfloor - 2} |(P^n g)(x) - \pi(g)|.$$

To be more explicit, consider for instance the case $s \in [2\eta, 3\eta)$ (i.e. $m = 2$). Then we have the following bound in total-variation norm from Corollary 8.3 and the material p. 84:

$$\forall x \in \mathbb{X},\ \forall n \ge 0, \quad \big\|P^n(x,\cdot) - \pi\big\|_{\mathrm{TV}} \le \frac{4(c_0 V_0(x) + c_1 \|1_{\mathbb{X}}\|_{V_0})}{n}$$

with

$$c_0 := (1+d_0)\left(1 + \frac{\nu(V_0)\,(1+d_0)}{2\pi(1_S) - 1}\right) \quad c_1 := \nu(V_0)(1+d_0)(1+d_1)\left(\frac{\nu(V_0)\,(1+d_0)}{2\pi(1_S)-1} + 1\right)$$

$$\forall i \in \{0,1\}, \quad d_i := \max\left(0, \frac{b_i - \nu(V_i)}{\nu(1_{\mathbb{X}})}\right)$$

with constants b_i given in $\boldsymbol{D}_{1_S}(V_0 : V_m)$. In this case, the bounds (139) also hold. Similarly, in case $s \ge 3\eta$, the bounds (141) can be used.

10.2.4 Further comments and bibliographic discussion

The MH Markov kernel defined by (194a)-(194b)-(194c) is associated with the Markov chain generated by the so-called Metropolis-Hastings algorithm. Such a kind of homogeneous Markov chain is in force in Markov Chain Monte Carlo (MCMC) algorithms for sampling the probability distribution π, called the target distribution. A major fact, in particular in Bayesian framework, is that π has only to be known up to a multiplicative constant for P to be well-defined. Here, the rationale underlying the Markov dynamics is: use an easy sampled auxiliary Markov chain $\{Y_n\}_{n\ge 0}$ with transition kernel K, called the proposal kernel, to generate a path of the MH Markov chain $\{X_n\}_{n\ge 0}$ with Markov kernel P which has π as $P-$invariant probability measure and converges in distribution to π. Given that the current state is $X_n = x_n$, get a candidate state y_n from the proposal probability distribution distribution $K(x_n, \cdot)$: Then either accept this candidate as the value of X_{n+1} (i.e. $X_{n+1} := y_n$) with probability $a(x_n, y_n)$, or stay at x_n (i.e. $X_{n+1} := x_n$) with probability $1 - a(x_n, y_n)$. This Markov dynamics corresponds to the definition (194a)-(194b)-(194c) of P, where (194b) can be reformulated in the usual form:

$$\forall (x,y) \in \mathbb{X}^2, \quad p(x,y) = a(x,y)k(x,y) \text{ with } a(x,y) := \min\left(1, \frac{\pi(y)k(y,x)}{\pi(x)k(x,y)}\right) \in [0,1].$$

In the case of an IMH Markov chain, the candidate state is selected according to the probability distribution $q(\cdot)$ irrespective the state x_n. In the case of the RWMH Markov chain, the candidate state is selected according to the probability distribution $q(\|\cdot - x_n\|)$. Using x_n as a realization of the random variable X_n with (approximate) probability distribution π requires that $(X_n)_{n\geq 0}$ converges in distribution to π and that n is large enough. Thus, since a long time ago, such algorithms have been a central support for much research on convergence and convergence rates of Markov chains. There is a plethora of literature on MCMC. We refer for instance to [RR04, RC04, BGJM11] for an overview of the topic.

Note that the basic conditions (P)-(C) can be weakened. We refer for instance to [JH00]. However note that, without conditions (P)-(C), a compact set of positive Lebesgue measure may not be a first-order small-set even if it is always a $\ell-$order small-set for some $\ell \geq 1$. Thus, an analysis may have to consider multiple-step transitions which is highly problematical for MCMC due to the untractability of P^ℓ for $\ell \geq 2$.

When P is an IMH Markov kernel, Condition (196) was introduced to obtain uniform ergodicity. It follows from the definitions in Example 3.7 and Theorem 6.2 that $1_{\mathbb{X}}-$geometric ergodicity is equivalent to uniform ergodicity, which turns to be equivalent to assume Condition (196) from [MT96, Th. 2.1] (see also [Wan22, Th. 1]). Under this condition, the convergence rate is explicit. Note that checking Condition (196) is not easy in the standard multidimensional settings of MCMC. Moreover such a condition makes it possible to use a direct independent sampling of π, the accept-reject method using the instrumental p.d.f. $q(\cdot)$ (see [RC04, Section 2.3]). Polynomial rates of convergence of the IMH Markov kernel are obtained in [JR02] when the condition (196) is violated. The result [JR07, Prop. 9] illustrates how the polynomial rate depends on the relative heaviness of π and $q(\cdot)$.

Recall that, if the RWMH Markov kernel P satisfies Conditions ($\boldsymbol{M}_{\nu,1_S}$)–$\boldsymbol{G}_{1_S}(\delta, V)$, then π is such that $\pi(V) < \infty$ (see the beginning of Section 6). Under these conditions, it is well-known from [JT03, Th. 2.2] that the Lyapunov function V is such $V(x) \geq c\exp(s\|x\|)$ for some positive constants $c, s > 0$ and $\|x\|$ is large enough, so that $\int_{\mathbb{X}} \exp(s\|x\|)\pi(x)dx < \infty$. Thus, the probability measure π must have exponential or lighter tails for P to be geometrically ergodic, irrespective of the proposal p.d.f. $q(\|\cdot\|)$. Note that the family of super-exponential p.d.f. π includes all the examples provided in the standard references [MT96, RR96, JH00] on geometric ergodicity of RWMH Markov kernels. In the MCMC context, Condition (199) means that the acceptance probability is uniformly bounded away from zero. For super-exponential p.d.f. π, Condition (199) is also necessary for P to be geometrically ergodic from [JH00, Th. 4.1]. The following sufficient condition for (199) is introduced in [JH00, Th. 4.3]

$$\limsup_{\|x\|\to+\infty} \langle \frac{x}{\|x\|}, \frac{\nabla\pi(x)}{\|\nabla\pi(x)\|} \rangle < 0. \tag{203}$$

This condition is shown in [JH00, Th. 4.4.] to be stable under translation, addition and multiplication. It is used to show that a very large class of probability measures π satisfies Condition (203) (see [JH00, Th. 4.6 and (46)]), including those with tails at least as light as multivariate Gaussian considered in [RR96].

In contrast to the geometric ergodicity, the order of polynomial ergodicity depends on the tails of both π and $q(\|\cdot\|)$. We refer to [JR07, and references therein] for such discussions in one-dimensional and multi-dimensional state spaces. Note that the Lyapunov function $V(x) = \max(\|x\|^s, 1)$ is used here in the polynomial case instead of the standard

Lyapunov function $V(x) = \pi(x)^{-\alpha}$ which leads to suboptimal convergence rates for spherically asymmetric target distributions π (see [JR07, Subsection 3.3]). In the case discussed in Subsection 10.2.3, we know from [JR02] that the drift inequality (202) provides the rate

$$\lim_{n\to+\infty} n^{\frac{s}{\eta}-1}\big\|P^n(x,\cdot)-\pi\big\|_{\mathrm{TV}} = 0.$$

As recalled in Subsection 10.2.3, explicit polynomial rate of convergence can be deduced from Section 8, but here we only focus on the optimality of the exponent in the polynomial rate. More specifically, when $s \uparrow r+\eta$, we obtain from the above convergence that $\sup_\delta(\lim_n n^\delta\big\|P^n(x,\cdot)-\pi\big\|_{\mathrm{TV}} = 0) \geq r/\eta$. In fact, we have an equality from [JT03, Prop. 4.2]. We refer to [JR07, Subsection 3.3] for results supporting the idea that, for polynomial target density π, using heavy-tailed proposal p.d.f. $q(\|\cdot\|)$ can improve the polynomial convergence rate.

Below are a few more references on various issues connected to RWMH and IMH Markov kernel:

- [Tie94, RR06] for Harris-recurrence.
- [BJ24b] for a recent overview of various methods for obtaining geometric convergence rates, [BJ24a] for a recent contribution on this topic, and finally [DFMS04, Section 3.2 and references therein] for conditions to get subgeometric rates of convergence.
- [RR11, DCWY19] on the mixing time and convergence time.
- When P is $\pi-$reversible, the $V-$geometric ergodicity is shown to be equivalent to the $\mathbb{L}^2(\pi)-$geometric ergodicity in [RR97] (see Subsection 9.7). This result was motivated by an analysis of a specific MCMC simulation algorithm. We refer to [Qin24] for a recent overview on convergence of MCMC and especially for $\mathbb{L}^2-$convergence. Finally, functional inequalities techniques are used in [ALPW24, and references therein] for analysing the $\mathbb{L}^2(\pi)-$spectral gap and the $\mathbb{L}^2(\pi)-$mixing times of MH Markov kernels.

11 Poisson's equation: Beyond first-order small-functions

Recall that the modulated drift condition $\boldsymbol{D}_{1_E}(V_0, V_1)$ for some $E \in \mathcal{X}^*$ and Lyapunov functions V_0 and V_1 is:

$$\exists b_0 > 0: \quad PV_0 \leq V_0 - V_1 + b_0 1_E.$$

To derive such a condition for P, we need to search for a set E and Lyapunov functions V_0 and V_1 such that $\Gamma := PV_0 - V_0 + V_1 \leq 0$ outside the set E. If such elements exist, then all that remains is to check that Γ is bounded from above on E. In general, this last requirement poses no problem (e.g. see Section 10). In order to apply the results of Sections 3–9, the set E in the modulated drift condition $\boldsymbol{D}_{1_E}(V_0, V_1)$ must be a first-order small-set. Unfortunately, this condition is not automatically satisfied, mainly due to the size of the set E. Actually, whatever the method used, the fact that the set E in $\boldsymbol{D}_{1_E}(V_0, V_1)$ is not necessarily a first-order small-set makes the study more complex. This is why higher-order small-sets or petite sets were introduced in the regenerative method, see Subsection 3.5.

If the set E in Condition $\boldsymbol{D}_{1_E}(V_0, V_1)$ is not a first-order small-set, it turns out that E can generally be written as a finite union of first-order small-sets. Moreover, in this case, E is a $\ell-$order small-set for some $\ell \geq 2$ (see Subsection 11.3). Thus, based on these two observations and focussing on the bound of solutions to Poisson's equation under the modulated drift condition, Theorem 5.4 is extended:

- In Subsection 11.1, under the drift Condition $\boldsymbol{D}_\psi(V_0, V_1)$ with $\psi = \sum_{i=1}^s b_i\psi_i$, where the ψ_i's are assumed to be first-order small-functions.
- In Subsection 11.2, under the drift Condition $\boldsymbol{D}_\psi(V_0, V_1)$ with $\psi \in \mathcal{B}_+^*$ assumed to be a small-function of order $\ell \geq 2$.

11.1 M & D conditions with several first-order small-functions

In this subsection the assumptions on P are the following ones: There exists an integer $s \geq 2$ such that

$$\exists\big\{(\nu_i, \psi_i)\big\}_{i=1}^s \in \big(\mathcal{M}_{+,b}^* \times \mathcal{B}_+^*\big)^s,\ \forall i = 1, \ldots, s: \quad P \geq \psi_i \otimes \nu_i. \tag{$\boldsymbol{M}_{1:s}$}$$

$$\exists\{b_i\}_{i=1}^s \in [0, +\infty)^s: \quad PV_0 \leq V_0 - V_1 + \sum_{i=1}^s b_i\psi_i \tag{$\boldsymbol{D}_{1:s}(V_0,V_1)$}$$

for some Lyapunov functions V_0 and V_1. In other words, P satisfies for every $i = 1, \ldots, s$ the first-order minorization condition $(\boldsymbol{M}_{\nu_i,\psi_i})$ with some $(\nu_i, \psi_i) \in \mathcal{M}_{+,b}^* \times \mathcal{B}_+^*$, as well as the V_1-modulated drift condition with some linear combination of the first-order small-functions ψ_i in the last term. In Subsection 11.1.1, some of the results of Section 3 are extended under the minorization conditions $(\boldsymbol{M}_{1:s})$. Next, in Subsection 11.1.2, the results of Section 5 on Poisson's equation are generalized under Conditions $(\boldsymbol{M}_{1:s})$–$\boldsymbol{D}_{1:s}(V_0, V_1)$.

11.1.1 Invariant probability measure

Under Condition $(\boldsymbol{M}_{1:s})$ for some $s \geq 2$, let us define the associated residual kernel

$$R := P - \sum_{i=1}^s \psi_i \otimes \nu_i \tag{204}$$

as well as the following positive measures

$$\forall i = 1, \ldots, s: \quad \mu_i := \sum_{k=0}^{+\infty} \nu_i R^k. \tag{205}$$

Note that μ_i is positive since $\mu_i \geq \nu_i$. It is worth noticing that μ_i is not the positive measure $\mu_{R_i} := \sum_{k=0}^{+\infty} \nu_i R_i^k$ associated with the residual kernel $R_i := P - \psi_i \otimes \nu_i$ defined from the sole Condition $(\boldsymbol{M}_{\nu_i,\psi_i})$.

Proposition 11.1 *Let P satisfy Condition $(\boldsymbol{M}_{1:s})$. Then we have*

$$\forall n \geq 1, \quad 0 \leq R^n \leq P^n \quad \textit{and} \quad P^n = R^n + \sum_{i=1}^s \sum_{k=1}^n P^{n-k}\psi_i \otimes \nu_i R^{k-1}. \tag{206}$$

Moreover, for $i = 1, \ldots, s$, the function series $\sum_{k=0}^{+\infty} R^k\psi_i$ point-wise converges and are bounded on $\mathbb{X}$, and we have

$$0 \leq \sum_{i=1}^s \nu_i(1_{\mathbb{X}}) \sum_{k=0}^{+\infty} R^k\psi_i = 1_{\mathbb{X}} - h_R^\infty \leq 1_{\mathbb{X}} \quad \textit{where} \quad h_R^\infty := \lim_n \searrow R^n 1_{\mathbb{X}}. \tag{207}$$

This proposition is an easy extension of Lemmas 3.2-3.3. The proof is only sketched below.

Proof. The first property in (206) follows from $0 \leq R \leq P$. For the second one, we prove by induction that

$$\forall n \geq 1, \quad T_n := P^n - R^n = \sum_{i=1}^{s} \sum_{k=1}^{n} P^{n-k}\psi_i \otimes \nu_i R^{k-1}. \tag{208}$$

Equality (208) is clear for $n = 1$ from the definition (204) of R. Next we have for any $n \geq 2$

$$R^n = R^{n-1}R = (P^{n-1} - T_{n-1})(P - T_1) = P^n - P^{n-1}T_1 - T_{n-1}R,$$

so that $T_n = P^{n-1}T_1 + T_{n-1}R$. Hence, if Formula (208) holds for T_{n-1} with some $n \geq 2$, then

$$\forall g \in \mathcal{B}, \quad T_n g = P^{n-1}T_1 g + T_{n-1}Rg = \sum_{i=1}^{s} \nu_i(g) P^{n-1}\psi_i + \sum_{i=1}^{s} \sum_{k=2}^{n} \nu_i(R^{k-1}g) P^{n-k}\psi_i$$

which is the desired formula for T_n. Properties in (206) are proved. Now observe that we have $\sum_{i=1}^{s} \nu_i(1_{\mathbb{X}})\psi_i = (I - R)1_{\mathbb{X}}$ from $P1_{\mathbb{X}} = 1_{\mathbb{X}}$ and the definition (204) of R. Thus

$$\forall n \geq 0, \quad \sum_{i=1}^{s} \nu_i(1_{\mathbb{X}}) \sum_{k=0}^{n} R^k \psi_i = \left(\sum_{k=0}^{n} R^k \right) (I - R) 1_{\mathbb{X}} = 1_{\mathbb{X}} - R^{n+1} 1_{\mathbb{X}}.$$

Since $0 \leq R1_{\mathbb{X}} \leq 1_{\mathbb{X}}$ and R is a non-negative kernel, we have $0 \leq R^{n+1}1_{\mathbb{X}} \leq R^n 1_{\mathbb{X}}$ for any $n \geq 0$, so that the sequence $(R^n 1_{\mathbb{X}})_{n \geq 0}$ is non-increasing and converges point-wise. This provides Property (207). □

In Theorem 11.3 below we prove that a suitable linear combination $\sum_{i=1}^{s} a_i \mu_i$ with μ_i defined in (205) is a finite positive $P-$invariant measure provided that each μ_i is finite. This result cannot be derived from Theorem 3.6 under each condition ($\boldsymbol{M}_{\nu_i,\psi_i}$). In other words, the positive measures μ_i in (205) are not $P-$invariant a priori. Indeed, as already observed, μ_i is not the positive measure μ_{R_i}, and anyway the condition $\mu_i(1_{\mathbb{X}}) < \infty$ does not imply that $\mu_{R_i}(1_{\mathbb{X}}) < \infty$. To find the specific linear combination $\sum_{i=1}^{s} a_i \mu_i$ providing a positive $P-$invariant measure, we need to prove Lemma 11.2 below. Let us introduce the following non-negative $s \times s-$matrix M and column vector $\mathfrak{u}_\nu \in (0, +\infty)^s$ defined by:

$$M := \big(\mu_j(\psi_i)\big)_{(i,j)\in\{1,\dots,s\}^2} \quad \text{and} \quad \mathfrak{u}_\nu = \big(\nu_1(1_{\mathbb{X}}), \dots, \nu_s(1_{\mathbb{X}})\big)^\top \tag{209}$$

where $(\cdot)^\top$ denotes the transpose operator.

Lemma 11.2 *Let P satisfy Conditions ($\boldsymbol{M}_{1:s}$). Then the positive measures μ_i in (205) satisfy*

$$\forall i \in \{1, \dots, s\}, \quad \mu_i(\psi_i) \in [0, 1].$$

Moreover, if each positive measure μ_i is finite (i.e. $\sum_{i=1}^{s} \mu_i(1_{\mathbb{X}}) < \infty$), then $M^\top \mathfrak{u}_\nu = \mathfrak{u}_\nu$, and there exists a column vector $\mathfrak{a} := (a_1, \dots, a_s)^\top \in [0, +\infty)^s$ with $\sum_{j=1}^{s} a_j > 0$ such that

$$M\mathfrak{a} = \mathfrak{a}, \ \textit{that is: } \forall i \in \{1, \dots, s\}, \quad \sum_{j=1}^{s} \mu_j(\psi_i)\, a_j = a_i. \tag{210}$$

Proof. Since $R \le R_i := P - \psi_i \otimes \nu_i$ we have

$$\forall i = 1, \ldots, s, \quad 0 \le \mu_i(\psi_i) = \sum_{k=0}^{+\infty} \nu_i(R^k \psi_i) \le \sum_{k=0}^{+\infty} \nu_i(R_i^k \psi_i) = \mu_{R_i}(\psi_i) \le 1$$

from Proposition 3.4 applied under Condition $(\boldsymbol{M}_{\nu_i,\psi_i})$. The first assertion is proved. Now assume that $\sum_{i=1}^s \mu_i(1_{\mathbb{X}}) < \infty$. Recall that h_R^∞ is defined in (207). We have $\nu_j(h_R^\infty) = \lim_n \nu_j(R^n 1_{\mathbb{X}})$ for $j = 1, \ldots, s$ from Lebesgue's theorem, so that $\nu_j(h_R^\infty) = 0$ since $\mu_j(1_{\mathbb{X}}) = \sum_{k=0}^{+\infty} \nu_j(R^k 1_{\mathbb{X}}) < \infty$ by hypothesis. Next, integrating (207) w.r.t. ν_j for $j \in \{1, \ldots, s\}$ provides

$$\forall j \in \{1, \ldots, s\}, \quad \sum_{i=1}^s \mu_j(\psi_i)\, \nu_i(1_{\mathbb{X}}) = \nu_j(1_{\mathbb{X}}), \text{ i.e. } M^\top \mathfrak{u}_\nu = \mathfrak{u}_\nu,$$

from the definition (205) of μ_j and $\nu_j(h_R^\infty) = 0$. Note that $\mathfrak{u}_\nu \ne 0$, more precisely $m := \min\{\nu_i(1_{\mathbb{X}}) : i = 1, \ldots, s\}$ is positive. Thus $M^\top$ is a non-negative matrix with 1 as eigenvalue. Moreover we have $\mathbf{1} \le m^{-1}\mathfrak{u}_\nu$ with $\mathbf{1} = (1, \ldots, 1)^\top$ where we use here the canonical order relation on $\mathbb{R}^s$.

Let $\|\cdot\|_\infty$ denote the supremum norm on $\mathbb{R}^s$. Setting $A := M^\top$ we have for every $n \ge 1$ and every $x \in \mathbb{R}^s$

$$\|A^n x\|_\infty \le \|A^n \mathbf{1}\|_\infty \|x\|_\infty \le m^{-1} \|A^n \mathfrak{u}_\nu\|_\infty \|x\|_\infty \le m^{-1} \|\mathfrak{u}_\nu\|_\infty \|x\|_\infty \tag{211}$$

since A is non-negative and $A^n \mathfrak{u}_\nu = \mathfrak{u}_\nu$. Thus, we have $\|A^n\|_\infty \le m^{-1}\|\mathfrak{u}_\nu\|_\infty$ where $\|A^n\|_\infty$ denotes the matrix-norm of A^n associated with $\|\cdot\|_\infty$, and 1 is an eigenvalue of A. This proves that the spectral radius of $A := M^\top$ is one. Accordingly M is a non-negative matrix with spectral radius one. Then (210) follows from the Perron-Frobenius theorem applied to the matrix M (e.g. see [Sen81, p. 28]) and [BP79, Th. 2.1.1, p. 26]). □

Theorem 11.3 *Assume that P satisfies Condition $(\boldsymbol{M}_{1:s})$ and that every positive measure μ_i in* (205) *is finite, i.e. $\sum_{i=1}^s \mu_i(1_{\mathbb{X}}) < \infty$. Let $\{a_j\}_{j=1}^s \in [0, +\infty)^s$ with $\sum_{j=1}^s a_j > 0$ given in Lemma 11.2. Then*

$$\mu_R := \sum_{i=1}^s a_i \mu_i$$

is a finite positive $P-$invariant measure. Consequently $\pi_R := \mu_R(1_{\mathbb{X}})^{-1}\, \mu_R$ is a $P-$invariant probability measure.

Proof. Note that μ_R is positive since so are the μ_i's and $\sum_{j=1}^s a_j > 0$. From the definitions (204) and (205) of R and μ_i we obtain that

$$\begin{aligned}
\forall A \in \mathcal{X}, \quad \mu_R(P1_A) &= \mu_R(R1_A) + \sum_{i=1}^s \mu_R(\psi_i)\nu_i(1_A) \\
&= \sum_{i=1}^s a_i\mu_i(R1_A) + \sum_{i=1}^s \bigg(\sum_{j=1}^s a_j\mu_j(\psi_i)\bigg)\nu_i(1_A) \\
&= \sum_{i=1}^s a_i\mu_i(1_A) - \sum_{i=1}^s a_i\nu_i(1_A) + \sum_{i=1}^s \bigg(\sum_{j=1}^s a_j\mu_j(\psi_i)\bigg)\nu_i(1_A) \\
&= \mu_R(1_A) + \sum_{i=1}^s \nu_i(1_A)\bigg(- a_i + \sum_{j=1}^s a_j\mu_j(\psi_i)\bigg).
\end{aligned}$$

It follows from (210) that μ_R is $P-$invariant. □

11.1.2 Poisson's equation

Under Conditions $(\boldsymbol{M}_{1:s})$–$\boldsymbol{D}_{1:s}(V_0, V_1)$ we first study the kernel series $\sum_{k=0}^{+\infty} R^k$ where R is the residual kernel defined in (204).

Theorem 11.4 *Let P satisfy Conditions $(\boldsymbol{M}_{1:s})$–$\boldsymbol{D}_{1:s}(V_0, V_1)$ and define*

$$d_{1:s} := \max_{i=1,\dots,s} \max\left(0, \frac{b_i - \nu_i(V_0)}{\nu_i(1_{\mathbb{X}})}\right).$$

Then

$$\sum_{k=0}^{+\infty} R^k 1_{\mathbb{X}} \le \sum_{k=0}^{+\infty} R^k V_1 \quad \le \quad V_0 + d_{1:s} 1_{\mathbb{X}} \le (1 + d_{1:s}) V_0 \tag{212a}$$

$$\forall i \in \{1, \dots, s\}, \ \sum_{k=0}^{+\infty} \nu_i\big(R^k 1_{\mathbb{X}}\big) \quad \le \quad \sum_{k=0}^{+\infty} \nu_i\big(R^k V_1\big) \le (1 + d_{1:s})\, \nu_i(V_0) < \infty. \tag{212b}$$

Proof. To simplify the presentation, set $d := d_{1:s}$, and let $V_{0,d} := V_0 + d 1_{\mathbb{X}}$. Then

$$\begin{aligned}
RV_{0,d} \quad &= \quad PV_{0,d} - \sum_{i=1}^{s} \nu_i(V_{0,d})\psi_i = PV_0 + d1_{\mathbb{X}} - \sum_{i=1}^{s} \big(\nu_i(V_0) + d\,\nu_i(1_{\mathbb{X}})\big)\psi_i \\
&\le \quad V_0 - V_1 + \sum_{i=1}^{s} b_i \psi_i + d1_{\mathbb{X}} - \sum_{i=1}^{s} \big(\nu_i(V_0) + d\,\nu_i(1_{\mathbb{X}})\big)\psi_i \ \text{ from } \boldsymbol{D}_{1:s}(V_0, V_1) \\
&\quad = V_{0,d} - V_1 + \sum_{i=1}^{s} \left(\frac{b_i - \nu_i(V_0)}{\nu_i(1_{\mathbb{X}})} - d\right) \nu_i(1_{\mathbb{X}})\, \psi_i \\
&\le \quad V_{0,d} - V_1
\end{aligned}$$

from the definition of d. Equivalently we have $V_1 \le V_{0,d} - RV_{0,d}$, thus

$$\forall n \ge 1, \quad \sum_{k=0}^{n} R^k V_1 \le \sum_{k=0}^{n} R^k V_{0,d} - \sum_{k=1}^{n+1} R^k V_{0,d} \le V_{0,d} - R^{n+1} V_{0,d} \le V_{0,d}$$

since we have $R^{n+1} V_{0,d} \ge 0$. This provides (212a) using $V_1 \ge 1_{\mathbb{X}}$ and $V_0 \ge 1_{\mathbb{X}}$. Note that $\nu_i(V_0) < \infty$ from $(\boldsymbol{M}_{1:s})$–$\boldsymbol{D}_{1:s}(V_0, V_1)$. Then Inequalities (212b) are deduced from (212a) and the monotone convergence theorem w.r.t. each positive measure ν_i. □

Recall that 1 is an eigenvalue of the matrix $M = (\mu_j(\psi_i))_{i,j=1,\dots,s}$ with associated non-zero and non-negative eigenvector $\mathfrak{a} = (a_1, \dots, a_s)^\top$, see (210). Consequently, 1 is a simple eigenvalue of M if, and only if, $\{x \in \mathbb{R}^s : Mx = x\} = \mathbb{R} \cdot \mathfrak{a}$.

Corollary 11.5 *Assume that P satisfies Conditions $(\boldsymbol{M}_{1:s})$–$\boldsymbol{D}_{1:s}(V_0, V_1)$. Then we have $\sum_{i=1}^{s} \mu_i(V_1) < \infty$, so that the $P-$invariant probability measure $\pi_R := \mu_R(1_{\mathbb{X}})^{-1} \mu_R$ of Theorem 11.3 is well-defined and such that $\pi_R(V_1) < \infty$. Moreover we have*

$$\pi_R \quad = \quad \sum_{i=1}^{s} \pi_R(\psi_i)\mu_i \tag{213a}$$

$$M\mathfrak{p} \quad = \quad \mathfrak{p} \quad \textit{with} \quad \mathfrak{p} := (\pi_R(\psi_1), \dots, \pi_R(\psi_s))^\top. \tag{213b}$$

Finally, if the eigenvalue 1 *of* M *is simple, then* π_R *is the unique* $P-$*invariant probability measure.*

Proof. We have $\sum_{i=1}^{s}\mu_i(V_1) < \infty$ from (212b) and the definition (205) of μ_i. Then we deduce from Theorem 11.3 that the $P-$invariant probability measure $\pi_R := \mu_R(1_{\mathbb{X}})^{-1}\,\mu_R$ is well-defined and satisfies $\pi_R(V_1) < \infty$. Now let η be any finite positive $P-$invariant measure. Then it follows from (206) that

$$\forall n \geq 1,\ \forall A \in \mathcal{X}, \quad \eta(1_A) = \eta(P^n 1_A) = \eta(R^n 1_A) + \sum_{j=1}^{s} \eta(\psi_j) \sum_{k=1}^{n} \nu_j(R^{k-1}1_A).$$

Note that $0 \leq \eta(R^n 1_A) \leq \eta(R^n 1_{\mathbb{X}})$ and that $\lim_n R^n 1_{\mathbb{X}} = 0$ (point-wise) from (212a). Thus $\lim_n \eta(R^n 1_A) = 0$ from Lebesgue's theorem since η is finite by hypothesis (i.e. $\eta(1_{\mathbb{X}}) < \infty$). When $n \to +\infty$ in the above equality we then obtain that

$$\eta = \sum_{j=1}^{s} \eta(\psi_j)\mu_j$$

from the definition (205) of μ_i. It follows from this equality applied to the small-functions ψ_i for $i = 1, \ldots, s$ that $M\mathfrak{b} = \mathfrak{b}$ with $\mathfrak{b} := (\eta(\psi_1), \ldots, \eta(\psi_s))^\top$. Note that the previous facts applied with $\eta := \pi_R$ provides (213a)-(213b). Finally assume that 1 is a simple eigenvalue of M. Then there exists $c > 0$ such that $\mathfrak{b} = c\,\mathfrak{a}$, so that $\eta = c\sum_{j=1}^{s} a_j\mu_j$. This provides the uniqueness of the $P-$invariant probability measure. □

To solve Poisson's equation under Conditions $(\boldsymbol{M}_{1:s})$–$\boldsymbol{D}_{1:s}(V_0, V_1)$ we need to prove the following lemma. We denote by I_s the identity $s \times s-$matrix.

Lemma 11.6 *Let* P *satisfy Condition* $(\boldsymbol{M}_{1:s})$ *with* $\sum_{i=1}^{s}\mu_i(1_{\mathbb{X}}) < \infty$. *For* $i \in \{1, \ldots, s\}$, *let* $\phi_i := \psi_i - \pi_R(\psi_i)1_{\mathbb{X}}$, *and define the following* $s \times s-$*matrix*

$$M_0 := \big(\mu_j(\phi_i)\big)_{(i,j)\in\{1,\ldots,s\}^2}.$$

If the eigenvalue 1 *of* $M = (\mu_j(\psi_i))_{(i,j)\in\{1,\ldots,s\}^2}$ *is simple, then the matrix* $I_s - M_0$ *is invertible.*

Proof. Using $\mu_j(\phi_i) = \mu_j(\psi_i) - \pi_R(\psi_i)\mu_j(1_{\mathbb{X}})$ it follows that

$$M_0 = M - \mathfrak{p}\cdot\mathfrak{m}^\top \quad \text{with} \quad \mathfrak{p} := \big(\pi_R(\psi_1), \ldots, \pi_R(\psi_s)\big)^\top \quad \text{and} \quad \mathfrak{m} := \big(\mu_1(1_{\mathbb{X}}), \ldots, \mu_s(1_{\mathbb{X}})\big)^\top.$$

Thus we have $I_s - M_0 = I_s - M + \mathfrak{p}\cdot\mathfrak{m}^\top$. Next, let $x \in \mathbb{R}^s$ be such that $(I_s - M_0)\cdot x = 0$, that is

$$(I_s - M)\cdot x = -\,\mathfrak{p}\cdot\mathfrak{m}^\top\cdot x = -\bigg(\sum_{i=1}^{s}\mu_i(1_{\mathbb{X}})\,x_i\bigg)\mathfrak{p}. \tag{214}$$

Recall that $M\mathfrak{p} = \mathfrak{p}$ (see (213b)). Hence we have $(I_s - M)^2\cdot x = 0$. Moreover denote by $\|\cdot\|_\infty$ the matrix-norm associated with the supremum norm on $\mathbb{R}^s$. Then we have $\sup_{n\geq1}\|M^n\|_\infty < \infty$ from (211) using the fact that a matrix and its transpose have the same norm. It follows that $\mathrm{Ker}(I_s - M)^2 = \mathrm{Ker}(I_s - M)$, so that the previous equality $(I_s - M)^2\cdot x = 0$ implies that $(I_s - M)\cdot x = 0$. Finally, since 1 is assumed to be a simple eigenvalue of M by hypothesis, we obtain that $\mathrm{Ker}(I_s - M) = \mathbb{R}\cdot\mathfrak{p}$, thus $x = c\,\mathfrak{p}$ for some $c \in \mathbb{R}$. From (214) we deduce that

$$c\sum_{i=1}^{s}\mu_i(1_{\mathbb{X}})\,\pi_R(\psi_i) = 0.$$

Thus we have $c = 0$ (i.e. $x = 0$) since $\sum_{i=1}^{s} \mu_i(1_{\mathbb{X}})\, \pi_R(\psi_i) = \pi_R(1_X) = 1$ from (213a). We have proved that $\mathrm{Ker}(I_s - M_0) = \{0\}$, so that $I_s - M_0$ is invertible. □

Now we study Poisson's equation under Conditions $(\boldsymbol{M}_{1:s})$–$\boldsymbol{D}_{1:s}(V_0, V_1)$.

Theorem 11.7 *Let P satisfy Conditions $(\boldsymbol{M}_{1:s})$–$\boldsymbol{D}_{1:s}(V_0, V_1)$ and let R be defined in (204). Then the following assertions hold.*

1. *For any $g \in \mathcal{B}_{V_1}$, the function series $\widetilde{g} := \sum_{k=0}^{+\infty} R^k g$ absolutely converges on $\mathbb{X}$ (point-wise convergence). Moreover we have $\widetilde{g} \in \mathcal{B}_{V_0}$ and*

$$\|\widetilde{g}\|_{V_0} \leq (1 + d_{1:s})\|g\|_{V_1} \tag{215}$$

with $d_{1:s}$ given in Theorem 11.4.

2. *Assume moreover that the eigenvalue 1 of the matrix $M := (\mu_j(\psi_i))_{(i,j)\in\{1,\ldots,s\}^2}$ is simple, and set $(I_s - M_0)^{-1} := (\alpha_{i,j})_{(i,j)\in\{1,\ldots,s\}^2}$ where the matrix M_0 is defined in Lemma 11.6. Finally let $\phi_i := \psi_i - \pi_R(\psi_i)1_{\mathbb{X}}$ for $i \in \{1, \ldots, s\}$. Then, for any $g \in \mathcal{B}_{V_1}$ such that $\pi_R(g) = 0$, the following function*

$$\widetilde{g}_P := \widetilde{g} + \sum_{i,j=1}^{s} \alpha_{i,j}\, \mu_i(g)\widetilde{\phi}_j \quad \text{with} \quad \widetilde{g} := \sum_{k=0}^{+\infty} R^k g,\ \ \widetilde{\phi}_i := \sum_{k=0}^{+\infty} R^k \phi_i \tag{216}$$

satisfies Poisson's equation $(I - P)\widetilde{g}_P = g$ with the following bounds

$$\|\widetilde{g}_P\|_{V_0} \leq (1 + d_{1:s})\|g\|_{V_1}\bigg(1 + (1 + d_{1:s}) \sum_{i,j=1}^{s} \nu_i(V_0)\, |\alpha_{i,j}|\, \|\phi_j\|_{V_1}\bigg). \tag{217}$$

Proof. Let $g \in \mathcal{B}_{V_1}$. Using $|g| \leq \|g\|_{V_1} V_1$ and $|R^k g| \leq R^k|g| \leq \|g\|_{V_1} R^k V_1$, Assertion 1. follows from (212a). Now, let us prove Assertion 2. Since the eigenvalue 1 of M is assumed to be simple, we know from Corollary 11.5 that π_R is the unique $P-$invariant probability measure and that $\pi_R(|g|) < \infty$ since $\pi_R(V_1) < \infty$. Now for every $n \geq 1$ define $\widetilde{g}_n := \sum_{k=0}^{n} R^k g$. Then, using $P = R + \sum_{i=1}^{s} \psi_i \otimes \nu_i$ we have

$$\widetilde{g}_n - P\widetilde{g}_n = \widetilde{g}_n - R\widetilde{g}_n - \sum_{i=1}^{s} \nu_i(\widetilde{g}_n)\psi_i = g - R^{n+1}g - \sum_{i=1}^{s} \nu_i(\widetilde{g}_n)\psi_i. \tag{218}$$

Repeating the same arguments as in the proof of Assertion 2. of Theorem 5.4 and taking the limit when n goes to infinity in (218) we obtain that

$$(I - P)\widetilde{g} = g - \sum_{i=1}^{s} \mu_i(g)\psi_i. \tag{219}$$

Now assume that $\pi_R(g) = 0$. Using $\phi_i := \psi_i - \pi_R(\psi_i)1_{\mathbb{X}}$ and $0 = \pi_R(g) = \sum_{i=1}^{s} \mu_i(g)\pi_R(\psi_i)$ from (213a) we obtain that

$$(I - P)\widetilde{g} = g - \sum_{i=1}^{s} \mu_i(g)\phi_i. \tag{220}$$

Recall that $M_0 := \big(\mu_j(\phi_i)\big)_{(i,j)\in\{1,\ldots,s\}^2}$. Applying Equality (220) to $g := \phi_i$ for $i \in \{1,\ldots,s\}$, provides the following function system:

$$\Gamma = (I_s - M_0)\Phi \quad \text{with} \quad \Phi := (\phi_1,\ldots,\phi_s)^\top \quad \text{and} \quad \Gamma := \big((I-P)\widetilde{\phi}_1,\ldots,(I-P)\widetilde{\phi}_s\big)^\top.$$

Since the eigenvalue 1 of M is assumed to be simple, we know from Lemma 11.6 that the matrix $I_s - M_0$ is invertible. Hence, we have $\Phi = (I_s - M_0)^{-1}\Gamma$, that is

$$\forall i \in \{1,\ldots,s\} \quad \phi_i = (I-P)\bigg(\sum_{j=1}^{s} \alpha_{i,j}\widetilde{\phi}_j\bigg),$$

where the $\alpha_{i,j}$'s denote the coefficients of the matrix $(I_s - M_0)^{-1}$. From these equalities and (220) it follows that, for every $g \in \mathcal{B}_{V_1}$ such that $\pi_R(g) = 0$, the function $\widetilde{g}_P$ defined in (216) satisfies Poisson's equation $(I-P)\widetilde{g}_P = g$. Finally note that for $i = 1,\ldots,s$ we have $|\mu_i(g)| \le (1+d_{1:s})\nu_i(V_0)\|g\|_{V_1}$ from (212b). The bound (217) for $\|\widetilde{g}_P\|_{V_0}$ is then deduced from (215) applied to g and ϕ_i. □

The following lemma provides sufficient conditions for the matrix M in Theorem 11.7 to have $\lambda := 1$ as simple eigenvalue. Recall that any non-negative $s \times s-$matrix A is said to be irreducible (e.g. see [Sen81, Def. 1.6]) if

$$\forall (i,j) \in \{1,\ldots,s\}^2,\ \exists n \equiv n(i,j) \ge 1, \quad A^n(i,j) > 0 \tag{221}$$

Lemma 11.8 *Let P satisfy Condition ($\boldsymbol{M}_{1:s}$) and M be the $s \times s-$matrix defined in* (209). *Under any of the two following conditions, the eigenvalue $\lambda = 1$ of M is simple and the vector $\mathfrak{a}$ in Lemma 11.2 is positive:*

1. *M satisfies Condition* (221).
2. *The non-negative $s \times s-$matrix $N := \big(\nu_j(\psi_i)\big)_{(i,j)\in\{1,\ldots,s\}^2}$ satisfies Condition* (221).

Note that the measures μ_i are unknown in general, so are M^n cannot be computed. In this case the sufficient condition on the computable $s \times s-$matrix N is relevant.

Proof. That Condition (221) for M is sufficient for $\lambda = 1$ to be a simple eigenvalue of M is standard from the Perron-Frobenius theorem (see [Sen81, Th. 1.5]). Under Condition (221), it still follows from Perron-Frobenius's theorem that the vector $\mathfrak{a}$ is positive as an eigenvector associated with the the eigenvalue $\lambda = 1$. Finally, since $M \ge N$ from $\mu_i \ge \nu_i$, we have $M^n \ge N^n$ for every $n \ge 1$. Hence, if N is irreducible, so is M. □

11.1.3 The specific case of two first-order small-functions

In case $s := 2$, we know from Lemma 11.8 that the irreducibility condition (221) for the matrix $N := (\nu_j(\psi_i))_{(i,j)\in\{1,2\}^2}$, is a sufficient condition for the eigenvalue 1 of matrix M to be simple. Condition (221) for N is equivalent to: $\nu_1(\psi_2) > 0$ and $\nu_2(\psi_1) > 0$. Below, this is weakened to $\nu_1(\psi_2) > 0$ or $\nu_2(\psi_1) > 0$, and the solution to Poisson's equation and its $\|\cdot\|_{V_0}-$norm are specified. The definitions of $\widetilde{g}$, ϕ_i and $\widetilde{\phi}_i$ (for $i = 1,2$ here) used below are those of Theorem 11.7.

Proposition 11.9 *Let P satisfy Conditions* $(\boldsymbol{M}_{1:2})-\boldsymbol{D}_{1:2}(V_0,V_1)$ *and*

$$\nu_1(\psi_2)+\nu_2(\psi_1)>0. \tag{222}$$

Then the following assertions hold.

1. *The eigenvalue $\lambda=1$ of the $2\times 2-$matrix M in (209) is simple.*
2. *The matrix $M_0 := \big(\mu_j(\phi_i)\big)_{i,j=1,2}$ of Lemma 11.6 is such that*

$$\Delta := \det(I_s-M_0) = 2-\mu_1(\psi_1)-\mu_2(\psi_2)>0.$$

3. *For any $g\in\mathcal{B}_{V_1}$ such that $\pi_R(g)=0$, the following function*

$$\widetilde{g}_P := \widetilde{g}+\Delta^{-1}\sum_{i=1}^{2}\mu_i(g)\zeta_i \quad \text{with} \quad \begin{cases} \zeta_1 &:= (1-\mu_2(\phi_2))\widetilde{\phi_1}+\mu_2(\phi_1)\widetilde{\phi_2} \\ \zeta_2 &:= \mu_1(\phi_2)\widetilde{\phi_1}+(1-\mu_1(\phi_1))\widetilde{\phi_2} \end{cases} \tag{223}$$

satisfies Poisson's equation $(I-P)\widetilde{g}_P=g$ with the following bounds

$$\|\widetilde{g}_P\|_{V_0} \le (1+d_{1,2})\|g\|_{V_1}\left(1+(1+d_{1,2})\Delta^{-1}\sum_{i=1}^{2}\nu_i(V_0)M_i\right) \tag{224}$$

$$\text{with} \quad \begin{cases} M_1 &:= |1-\gamma_{22}|(\|\psi_1\|_{V_1}+\pi_R(\psi_1))+|\gamma_{12}|\big(\|\psi_2\|_{V_1}+\pi_R(\psi_2)\big) \\ M_2 &:= |\gamma_{21}|\big(\|\psi_1\|_{V_1}+\pi_R(\psi_1)\big)+|1-\gamma_{11}|(\|\psi_2\|_{V_1}+\pi_R(\psi_2)) \end{cases}$$

where $\gamma_{ij}:=\mu_j(\psi_i)-\pi_R(\psi_i)\mu_j(1_{\mathbb{X}})$ for $(i,j)\in\{1,2\}^2$.

Proof. Assume that the eigenvalue 1 of M is not simple. Then the trace of M is 2, so that we have $\mu_1(\psi_1)+\mu_2(\psi_2)=2$. Thus $\mu_1(\psi_1)=\mu_2(\psi_2)=1$ since $\mu_i(\psi_i)\in[0,1]$ for $i=1,2$ from Lemma 11.2. Recall that $M^\top\mathfrak{u}_\nu=\mathfrak{u}_\nu$ with $\mathfrak{u}_\nu:=(\nu_1(1_{\mathbb{X}}),\nu_2(1_{\mathbb{X}}))^\top$ from Lemma 11.2 so that $\mu_1(\psi_2)=\mu_2(\psi_1)=0$. Since $\mu_i\ge\nu_i$, we have $\nu_1(\psi_2)=\nu_2(\psi_1)=0$: This contradict (222). Assertion 1. is proved.

To prove Assertion 2., let us generically denote by $C1$ and C_2 the first and second column vectors of a $2\times 2-$determinant. From (213a) we know that $\sum_{i=1}^2\pi_R(\psi_i)>0$. Assume that $\pi_R(\psi_1)>0$. Then, replacing C_1 with $\pi_R(\psi_1)C_1+\pi_R(\psi_2)C_2$ in Δ, we obtain that

$$\Delta=\pi_R(\psi_1)^{-1}\begin{vmatrix} \pi_R(\psi_1) & -\mu_2(\phi_1) \\ \pi_R(\psi_2) & 1-\mu_2(\phi_2) \end{vmatrix}$$

using Formula (213a) and $\pi_R(\phi_1)=\pi_R(\phi_2)=0$. Next, using $\phi_i:=\psi_i-\pi_R(\psi_i)1_{\mathbb{X}}$ and replacing C_2 with $C_2-\mu_2(1_{\mathbb{X}})C_1$ in the last determinant provides

$$\Delta=\pi_R(\psi_1)^{-1}\begin{vmatrix} \pi_R(\psi_1) & -\mu_2(\psi_1)+\mu_2(1_{\mathbb{X}})\pi_R(\psi_1) \\ \pi_R(\psi_2) & 1-\mu_2(\psi_2)+\mu_2(1_{\mathbb{X}})\pi_R(\psi_2) \end{vmatrix}=\pi_R(\psi_1)^{-1}\begin{vmatrix} \pi_R(\psi_1) & -\mu_2(\psi_1) \\ \pi_R(\psi_2) & 1-\mu_2(\psi_2) \end{vmatrix}.$$

Finally, from Equality $\pi_R(\psi_2)\mu_2(\psi_1)=\pi_R(\psi_1)-\pi_R(\psi_1)\mu_1(\psi_1)$ which follows from (213a), the desired formula for Δ is easily deduced. If $\pi_R(\psi_1)=0$, then $\pi_R(\psi_2)>0$ and the computation od Δ is similar. The argument for proving Assertion 1. shows that we have $\mu_1(\psi_1)+\mu_2(\psi_2)<2$ under Condition (222). The proof of Assertion 2. is complete.

Finally Assertion 3. follows from Theorem 11.7 and the direct computation of the $2\times 2-$matrix $(I_s-M_0)^{-1}$. □

Condition (222) is quite generic. This is because the support of at least one of the minorizing measures, say ν_2, is a closed set strictly larger than S_2. We therefore have $\nu_2(1_{S_1}) > 0$ when S_1 intersects the support of ν_2. This is illustrated in the following simple example.

Example 11.10 *Let $\mathbb{X} := \mathbb{N}$ and assume that $P = (P(i,j))_{(i,j)\in\mathbb{N}^2}$ satisfies $\boldsymbol{D}_{1_E}(V_0, V_1)$ for some Lyapunov functions V_0, V_1 and $E := \{0, 1, 2\}$. This means that*

$$\forall i \geq 3: \quad \sum_{j=0}^{+\infty} P(i,j)V_0(j) \leq V_0(i) - V_1(i)$$

and that b_0 in $\boldsymbol{D}_{1_E}(V_0, V_1)$ can be defined as

$$b_0 := \max_{i=0,1,2} \Big\{ \sum_{j=0}^{+\infty} P(i,j)V_0(j) - V_0(i) + V_1(i) \Big\} < \infty.$$

Assume that

$$\alpha_0 := \min(P(0,0), P(1,0)) > 0, \quad P(0,2) = 0, \quad P(1,1) = 0, \quad P(2,0) = 0, \quad P(2,1) > 0.$$

Then it can be easily seen that E is not a first-order small-set. Now write $E = S_1 \sqcup S_2$ with $S_1 := \{0, 1\}$ et $S_2 := \{2\}$. Then S_1 is a first-order small-set with associated minorizing measure $\nu_1 := \alpha_0\delta_0$ where δ_0 is the Dirac distribution at 0. Next S_2 is also a first-order small-set, even an atom, with associated minorizing measure $\nu_2 := P(2, \cdot)$. Finally, since $P(2,1) > 0$, we have $\nu_2(1_{S_1}) > 0$, so that Condition (222) holds.

11.2 Poisson's equation under higher-order minorization condition

Recall that Theorem 5.4 provides a bound for the V_0−weighted norm of solutions to Poisson's equation under the V_1−modulated drift Condition $\boldsymbol{D}_\psi(V_0, V_1)$, that is

$$\exists b_0 \equiv b_0(V_0, V_1, \psi) > 0: \quad PV_0 \leq V_0 - V_1 + b_0\psi,$$

where $\psi \in \mathcal{B}_+^*$ satisfies the first-order minorization condition ($\boldsymbol{M}_{\nu,\psi}$) with some $\nu \in \mathcal{M}_{+,b}^*$. Here Theorem 5.4 is extended to Markov kernel P still satisfying $\boldsymbol{D}_\psi(V_0, V_1)$ but assuming now that ψ is a ℓ−order small-function for some integer $\ell \geq 2$, namely

$$\exists \ell \geq 2,\ \exists \nu_\ell \in \mathcal{M}_{+,b}^*: \quad P^\ell \geq \psi \otimes \nu_\ell. \qquad (\boldsymbol{M}^\ell_{\nu_\ell,\psi})$$

Iterating $\boldsymbol{D}_\psi(V_0, V_1)$ we obtain that

$$\forall k = 1, \ldots, \ell, \quad P^k V_0 \leq V_0 - \sum_{j=0}^{k-1} P^j V_1 + b_0 \sum_{j=0}^{k-1} P^j \psi. \qquad (225)$$

It follows from the previous inequality for $k := \ell$ that

$$P^\ell V_0 \quad \leq \quad V_0 - V_1 + b_0\psi + b_0 \sum_{j=1}^{\ell-1} P^j \psi. \qquad (226)$$

Note that Theorem 5.4 does not apply directly to the Markov kernel P^ℓ under Conditions $\boldsymbol{D}_\psi(V_0,V_1)$ and $(\boldsymbol{M}^\ell_{\nu_\ell,\psi})$. Indeed $(\boldsymbol{M}^\ell_{\nu_\ell,\psi})$ is a first-order minorization condition for P^ℓ with minorizing measure ν_ℓ and small-function ψ, but Inequality (226) is not (and does not provide) Condition $\boldsymbol{D}_\psi(V_0,V_1)$ for P^ℓ because of the term $b_0\sum_{j=1}^{\ell-1}P^j\psi$. In fact, Lemma 11.13 below shows that the proof of Theorem 5.4 can be adapted, provided that $b_0\sum_{j=1}^{\ell-1}P^j\psi \le a_\ell P^\ell\psi$ for some $a_\ell\in(0,+\infty)$.

Thus, let us introduce the following condition:

$$\forall j=1,\dots,\ell-1,\ \exists\alpha_j\in(0,+\infty),\quad P^j\psi\le\alpha_j P^\ell\psi. \tag{227}$$

which ensures that $b_0\sum_{j=1}^{\ell-1}P^j\psi\le a_\ell P^\ell\psi$ with $a_\ell:=b_0\sum_{j=1}^{\ell-1}\alpha_j$. A sufficient condition for (227) is

$$\exists a>0,\quad P\psi\le aP^2\psi$$

which is satisfied if $\psi\le aP\psi$ from the non-negativity of P. Indeed, it can be easily seen that Condition $P\psi\le aP^2\psi$ provides (227) with $\alpha_j=a^{\ell-j}$ (smaller constants α_j in (227) can be found when $P^j\psi$ for $j=1,\dots,\ell$ are computable). In particular, under Condition $(\boldsymbol{M}_{\nu,\psi})$ with $\nu(\psi)>0$, we have $\psi\le\nu(\psi)^{-1}P\psi$, so that Condition (227) holds. However Condition (227) is of course much weaker than $(\boldsymbol{M}_{\nu,\psi})$.

The following lemma collects the direct consequences of Inequality (225) recalling that ψ is a non-negative bounded function on $\mathbb{X}$, and of Inequality (226) assuming that Condition (227) holds true.

Lemma 11.11 *Let P satisfy Condition $\boldsymbol{D}_\psi(V_0,V_1)$ for some Lyapunov functions V_0, V_1, and for some $\psi\in\mathcal{B}^*_+$ satisfying $(\boldsymbol{M}^\ell_{\nu_\ell,\psi})$ and assume that Condition (227) holds. Then the following assertions hold:*

$$\text{the function }\ \Gamma_\ell:=\max_{k=1,\dots,\ell-1}P^kV_0-V_0\ \text{ is bounded from above on }\mathbb{X}, \tag{228a}$$

$$\text{and }\ \exists(a_0,a_\ell)\in[0,+\infty)^2,\quad P^\ell V_0\le V_0-V_1+a_0\psi+a_\ell P^\ell\psi. \tag{228b}$$

From the above, under the assumptions of Lemma 11.11, Inequality (228b) holds with $a_0:=b_0$ and $a_\ell:=b_0\sum_{j=1}^{\ell-1}\alpha_j$ where b_0 and the α_j's are given in $\boldsymbol{D}_\psi(V_0,V_1)$ and in Conditions (227) respectively. However, in practice it is of course relevant to search for the smallest possible constants.

Under the $\ell-$order minorization Condition $(\boldsymbol{M}^\ell_{\nu_\ell,\psi})$, we denote by R_ℓ the submarkov kernel defined on $(\mathbb{X},\mathcal{X})$ by

$$R_\ell:=P^\ell-\psi\otimes\nu_\ell.$$

Theorem 11.12 *Let P satisfy Condition $\boldsymbol{D}_\psi(V_0,V_1)$ for some Lyapunov functions V_0, V_1, and for some $\psi\in\mathcal{B}^*_+$ satisfying $(\boldsymbol{M}^\ell_{\nu_\ell,\psi})$.* Assume that Condition (227) *holds. Then the following assertions hold:*

1. *There exists a unique $P-$invariant probability measure π on $(\mathbb{X},\mathcal{X})$, and $\pi(V_1)<\infty$.*

2. *For every $g\in\mathcal{B}_{V_1}$ such that $\pi(g)=0$, the function*

$$\tilde g:=\sum_{k=0}^{\ell-1}P^k\tilde g_\ell\quad\text{with}\quad\tilde g_\ell=\sum_{k=0}^{+\infty}{R_\ell}^kg \tag{229}$$

belongs to $\mathcal{B}_{V_0}$*, satisfies Poisson's equation* $(I-P)\tilde g = g$*, and the following bound holds true:*

$$|\tilde g| \le \|g\|_{V_1}\big[\ell(V_0+\Gamma_\ell+c_\ell 1_{\mathbb{X}}) - a_\ell\,\psi_\ell\big] \quad \textit{with} \quad c_\ell := \frac{\max\big(0, a_0-\nu_\ell(V_0)\big)+a_\ell\big(1+\nu_\ell(\psi)\big)}{\nu_\ell(1_{\mathbb{X}})} \tag{230}$$

where Γ_ℓ *is the upper bounded function given in* (228a)*,* a_0, a_ℓ *are any positive constants such that* (228b) *holds, and finally* $\psi_\ell := \sum_{k=0}^{\ell-1} P^k\psi$.

In Case $\ell = 1$ we have $\Gamma_\ell = 0$ and $a_\ell = 0$, so that the bound (230) corresponds to (59a) in Theorem 5.4.

The proof of Theorem 11.12 is a direct consequence of the next Lemmas 11.13, 11.14, observing moreover that the non-negative function $W_0 := \sum_{k=0}^{\ell-1} P^k V_0$ introduced in Lemma 11.14 satisfies $W_0 \le \ell(V_0+\Gamma_l)$ using (228a).

Lemma 11.13 *Let* P *satisfy Conditions* ($\boldsymbol{M}^{\ell}_{\nu_\ell,\psi}$) *and* (228b) *for some* $\ell \ge 2$ *and non-negative measurable functions* V_0 *and* V_1 *on* $\mathbb{X}$*. Then*

$$0 \le \sum_{k=0}^{+\infty} R_\ell^k V_1 \ \le V_0 + c_\ell 1_{\mathbb{X}} - a_\ell\,\psi \quad \textit{with} \quad c_\ell := \frac{\max\big(0, a_0-\nu_\ell(V_0)\big)+a_\ell\big(1+\nu_\ell(\psi)\big)}{\nu_\ell(1_{\mathbb{X}})}. \tag{231}$$

Proof. From (228b) we obtain that

$$R_\ell V_0 = P^\ell V_0 - \nu_\ell(V_0)\psi \le V_0 - V_1 + \big(a_0-\nu_\ell(V_0)\big)\psi + a_\ell\,P^\ell\psi,$$

$$\text{equivalently:} \quad V_1 \le V_0 - R_\ell V_0 + \big(a_0-\nu_\ell(V_0)\big)\psi + a_\ell\,P^\ell\psi.$$

Moreover observe that $\sum_{k=0}^{+\infty} R_\ell^k\psi \le \nu_\ell(1_{\mathbb{X}})^{-1}1_{\mathbb{X}}$ from Inequality (24) applied to the Markov kernel P^ℓ under the mimorization condition ($\boldsymbol{M}^{\ell}_{\nu_\ell,\psi}$). Thus we have

$$\begin{aligned}
\forall n\ge 1,\quad \sum_{k=0}^{n} R_\ell^k V_1 \ &\le\ \sum_{k=0}^{n} R_\ell^k V_0 - \sum_{k=1}^{n+1} R_\ell^k V_0 + \big(a_0-\nu_\ell(V_0)\big)\sum_{k=0}^{n} R_\ell^k\psi + a_\ell \sum_{k=0}^{n} R_\ell^k P^\ell\psi \\
&\le\ V_0 + \frac{\max\big(0,a_0-\nu_\ell(V_0)\big)}{\nu_\ell(1_{\mathbb{X}})}1_{\mathbb{X}} + a_\ell\sum_{k=0}^{n} R_\ell^k\big(R_\ell\psi+\nu_\ell(\psi)\psi\big) \\
&= V_0 + \frac{\max\big(0,a_0-\nu_\ell(V_0)\big)}{\nu_\ell(1_{\mathbb{X}})}1_{\mathbb{X}} + a_\ell\sum_{k=1}^{n+1} R_\ell^k\psi + a_\ell\,\nu_\ell(\psi)\sum_{k=0}^{n} R_\ell^k\psi \\
&\le\ V_0 + \frac{\max\big(0,a_0-\nu_\ell(V_0)\big)}{\nu_\ell(1_{\mathbb{X}})}1_{\mathbb{X}} + \frac{a_\ell}{\nu_\ell(1_{\mathbb{X}})}1_{\mathbb{X}} - a_\ell\,\psi + \frac{a_\ell\,\nu_\ell(\psi)}{\nu_\ell(1_{\mathbb{X}})}1_{\mathbb{X}} \\
&\le\ V_0 + \frac{\max\big(0,a_0-\nu_\ell(V_0)\big)+a_\ell\big(1+\nu_\ell(\psi)\big)}{\nu_\ell(1_{\mathbb{X}})}1_{\mathbb{X}} - a_\ell\psi.
\end{aligned}$$

This proves (231). □

Lemma 11.14 *Let* P *satisfy Conditions* ($\boldsymbol{M}^{\ell}_{\nu_\ell,\psi}$) *and* (228b) *for some* $\ell \ge 2$ *and measurable functions* $V_0 \ge 0$ *and* $V_1 \ge 1_{\mathbb{X}}$ *on* $\mathbb{X}$*. Then there exists a unique* $P-$*invariant probability* *measure* π *on* $(\mathbb{X}, \mathcal{X})$*, and we have* $\pi(V_1) < \infty$*. Moreover, for every* $g \in \mathcal{B}_{V_1}$ *such that* $\pi(g) = 0$*, the function*

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

satisfies the following bound

$$|\tilde{g}| \le \|g\|_{V_1}\big(W_0 + \ell\, c_\ell 1_{\mathbb{X}} - a_\ell \psi_\ell\big) \quad \textit{with} \quad W_0 := \sum_{k=0}^{\ell-1} P^k V_0 \quad \textit{and} \quad \psi_\ell := \sum_{k=0}^{\ell-1} P^k \psi \tag{232}$$

where the positive constants a_ℓ *and* c_ℓ *are given in* (228b) *and* (231) *respectively. Moreover* $\tilde{g}$ *satisfies Poisson's equation* $(I - P)\tilde{g} = g$*.*

Proof. From $1_{\mathbb{X}} \le V_1$ and (231) we know that $\lim_k R_\ell^k 1_{\mathbb{X}} = 0$ (point-wise convergence) and that $\sum_{k=0}^{+\infty} \nu_\ell(R_\ell^k 1_{\mathbb{X}}) \le \nu_\ell(V_0) + c_\ell \nu_\ell(1_{\mathbb{X}}) < \infty$ from the monotone convergence theorem (use $(\boldsymbol{M}^{\ell}_{\nu_\ell,\psi})$ and (228b) to get $\nu_\ell(V_0) < \infty$). Thus

$$\mu_{R_\ell} := \sum_{k=0}^{+\infty} \nu_\ell R_\ell^k$$

is a finite positive measure on $(\mathbb{X}, \mathcal{X})$. Then it follows from Assertion 3. of Theorem 4.1 applied to the Markov kernel P^ℓ with residual kernel R_ℓ that $\pi_{R_\ell} := \mu_{R_\ell}(1_{\mathbb{X}})^{-1}\mu_{R_\ell}$ is the unique $P^\ell-$invariant probability measure. Moreover we have $\mu_{R_\ell}(V_1) < \infty$ from (231) and the monotone convergence theorem. Next the following probability measure

$$\pi := \frac{1}{\ell}\sum_{k=0}^{\ell-1} \pi_{R_\ell} P^k$$

is $P-$invariant using that $\pi_{R_\ell} P^\ell = \pi_{R_\ell}$. In fact we have $\pi = \pi_{R_\ell}$ since π is also $P^\ell-$invariant and π_{R_ℓ} is the unique $P^\ell-$invariant probability measure. The first assertion of Lemma 11.14 is proved.

Now let $g \in \mathcal{B}_{V_1}$ be such that $\pi(g) = 0$. Let us prove that the function $\widetilde{g}_\ell = \sum_{k=0}^{+\infty} {R_\ell}^k g$ satisfies the following bound

$$|\widetilde{g}_\ell| \le \|g\|_{V_1}\big(V_0 + c_\ell 1_{\mathbb{X}} - a_\ell \psi\big) \tag{233}$$

and that $(I - P^\ell)\widetilde{g}_\ell = g$. We have $|g| \le \|g\|_{V_1} V_1$, and for $k \ge 1$

$$|R_\ell^k g| \le R_\ell^k |g| \le \|g\|_{V_1} R_\ell^k V_1.$$

Then we obtain (233) using (231). Next define: $\forall n \ge 1,\ \widetilde{g}_{\ell,n} := \sum_{k=0}^{n} R_\ell^k g$. Using Equality $P^\ell = R_\ell + \psi \otimes \nu_\ell$, we obtain that

$$\widetilde{g}_{\ell,n} - P^\ell \widetilde{g}_{\ell,n} = \widetilde{g}_{\ell,n} - R_\ell \widetilde{g}_{\ell,n} - \nu_\ell(\widetilde{g}_{\ell,n})\psi = g - R_\ell^{n+1} g - \nu_\ell(\widetilde{g}_{\ell,n})\psi$$

from which we deduce that $(I - P^\ell)\widetilde{g}_\ell = g - \mu_{R_\ell}(g)\psi$ when $n \to +\infty$ (repeating the arguments of the proof of Theorem 5.4). Hence we have $(I - P^\ell)\widetilde{g}_\ell = g$ since $\mu_{R_\ell}(g) = \mu_{R_\ell}(1_{\mathbb{X}})\pi_{R_\ell}(g) = \mu_{R_\ell}(1_{\mathbb{X}})\pi(g) = 0$.

We can now conclude. From (233) it follows that $\tilde{g} := \sum_{k=0}^{\ell-1} P^k \widetilde{g}_\ell$ satisfies the following bound:

$$|\tilde{g}| \le \sum_{k=0}^{\ell-1} P^k |\widetilde{g}_\ell| \le \|g\|_{V_1} \sum_{k=0}^{\ell-1} P^k \big(V_0 + c_\ell 1_{\mathbb{X}} - a_\ell \psi\big) = \|g\|_{V_1} \big(W_0 + \ell\, c_\ell 1_{\mathbb{X}} - a_\ell \psi_\ell\big)$$

with $W_0 := \sum_{k=0}^{\ell-1} P^k V_0$ and $\psi_\ell := \sum_{k=0}^{\ell-1} P^k \psi$. Moreover we have

$$(I-P)\tilde{g} = (I-P)\sum_{k=0}^{\ell-1} P^k \widetilde{g}_\ell = (I-P^\ell)\widetilde{g}_\ell = g.$$

The proof of Lemma 11.14 is complete. □

Now let us present an alternative statement to Theorem 11.12. Assume that P satisfies Conditions $\boldsymbol{D}_\psi(V_0, V_1)$ and (227). Using the following inequality (see (225) with $k := \ell$)

$$P^\ell V_0 \le V_0 - W_1 + b_0 \sum_{j=0}^{\ell-1} P^j \psi \quad \text{with} \quad W_1 := \sum_{j=0}^{\ell-1} P^j V_1,$$

we obtain that

$$\exists (a_0, a_\ell) \in [0, +\infty)^2, \quad P^\ell V_0 \le V_0 - W_1 + a_0 \psi + a_\ell P^\ell \psi. \tag{234}$$

Next, under Conditions (234) and $(\boldsymbol{M}^\ell_{\nu_\ell,\psi})$, the conclusion of Lemma 11.13 obviously extends with W_1 in place of V_1. Next, for $i = 0, 1$, define

$$\widehat{V}_i := \frac{1}{\ell} \sum_{j=0}^{\ell-1} P^j V_i,$$

and observe that $\widehat{V}_i = W_i/\ell$. Then, Lemma 11.14 can be straightforwardly adapted to obtain the following statement.

Proposition 11.15 *Let P satisfy the assumptions of Theorem* 11.12. *Then the following assertions hold:*

1. *There exists a unique $P-$invariant probability measure π on $(\mathbb{X}, \mathcal{X})$, and $\pi(\widehat{V}_1) < \infty$.*
2. *For every $g \in \mathcal{B}_{\widehat{V}_1}$ such that $\pi(g) = 0$, the function $\tilde{g}$ given in* (229) *belongs to $\mathcal{B}_{\widehat{V}_0}$, satisfies Poisson's equation $(I-P)\tilde{g} = g$, and*

$$|\tilde{g}| \le \|g\|_{\widehat{V}_1} \big(\widehat{V}_0 + c_\ell\, 1_{\mathbb{X}} - a_\ell\, \widehat{\psi}_\ell\big) \tag{235}$$

where c_ℓ is defined as in Theorem 11.12 *using here any positive constants a_0, a_ℓ such that* (234) *holds, and where $\widehat{\psi}_\ell := (1/\ell)\sum_{k=0}^{\ell-1} P^k \psi$.*

The interest of the alternative bound (235) is that the order ℓ is no longer a multiplicative factor. Its disadvantage is that this bound uses the modified Lyapunov functions $\widehat{V}_0$ and $\widehat{V}_1$. On this subject recall that $\widehat{V}_0 \le V_0 + \Gamma_\ell$ (see (228a)), and note that $\widehat{V}_1$ in (235) can be replaced with any measurable function U_1 such that $1_{\mathbb{X}} \le U_1 \le \widehat{V}_1$ since $\mathcal{B}_{U_1} \subset \mathcal{B}_{\widehat{V}_1}$ and $\|g\|_{\widehat{V}_1} \le \|g\|_{U_1}$ for every $g \in \mathcal{B}_{U_1}$.

Let us conclude with some comments on Theorem 11.12. First note that the bounds (230) and (235) may be of interest even if ψ is also a first-order small-function: Indeed, when the mass $\nu(1_{\mathbb{X}})$ of the first-order minorizing measure ν in $(\boldsymbol{M}_{\nu,\psi})$ is too small, then the constant d_0 in the bounds (59a)-(59b) of Theorem 5.4 may be too large to be relevant. Of course the mass $\nu_\ell(1_{\mathbb{X}})$ involved in the bound (230) is expected to be greater. Explicit bounds for the constants a_0 and a_ℓ in Inequality (228b) are proposed after Lemma 11.11. However, in practice, when the functions $P^\ell V_0$ and $P^\ell \psi$ are computable, it is more relevant to estimate constants a_0 and a_ℓ in (228b) directly. The same holds for Condition (234). If ψ is also a first-order small-function (i.e. Condition $(\boldsymbol{M}_{\nu,\psi})$ holds) and if $\pi(V_0) < \infty$, then the bound

$$\Gamma_\ell \leq \frac{\|\psi\|_{1_{\mathbb{X}}}\big(\pi(V_0) + d_0\big)}{\pi(\psi)}$$

provided by Lemma 5.9 can be used in (230). Likewise, under the geometric drift condition $\boldsymbol{G}_\psi(\delta, V)$ which provides the V_1-modulated drift Condition $\boldsymbol{D}_\psi(V_0, V_1)$ with $V_0 := V/(1-\delta)$, $V_1 := V$ and $b_0 := b/(1-\delta)$ (see Example 5.2), the bound $\Gamma_\ell \leq b\|\psi\|_{1_{\mathbb{X}}}(1-\delta)^{-2}$ can be used in (230). Indeed an easy iteration of $\boldsymbol{G}_\psi(\delta, V)$ provides

$$\forall k \geq 0, \quad P^k V \leq V + b\|\psi\|_{1_{\mathbb{X}}}(1-\delta)^{-1}1_{\mathbb{X}}.$$

However, even in the previous specific cases, it is preferable to search for an upper bound of Γ_ℓ simply by using the functions $P^k V_0$ for $k = 1, \ldots, \ell - 1$, provided that these functions are computable.

11.3 Further comments and bibliographic discussion

That the first-order minorization condition may be restrictive for the set E involved in drift conditions (e.g. in $PV_0 \leq V_0 - V_1 + b1_E$) is addressed in this section. Actually, when E is not a first-order small-set, it can generally be written as a finite union of first-order small-sets. Thus, a first way to overcome this issue is based on this decomposition of E. The material is proposed in Subsection 11.1 and is new to the best of our knowledge. As an extension of the first-order minorization condition, the results of Subsection 11.1 are interesting, but mainly from a theoretical point of view because the matrix involved in Theorem 11.7 is not computable in general.

The notions of higher-order small-functions (or small-sets) and petite sets have been already presented in Subsection 3.5. When the set E in the modulated drift condition can only be written as a finite union of first-order small-sets, then E is a $\ell-$order small-set for some $\ell \geq 2$, that is $P^\ell \geq 1_E \otimes \nu_\ell$ with some minorizing measure ν_ℓ. This follows from the following classical results (e.g. see [MT09, Prop. 5.5.5, Th. 5.5.7]: First every small-set is obviously petite; Second a finite union of petite sets is petite under irreducibility condition; Third, every petite set is a small-set under the additional aperiodicity condition. Partly motivated by this observation in Subsection 11.2, Theorem 11.12 and Proposition 11.15 provide an extension of Theorem 5.4 on the bound for solutions to Poisson's equation to the case when the function ψ in the modulated drift condition is any higher-order small-function.

It is well-documented in the literature on Markov chains that the standard way to go beyond the first-order minorization condition is to search for a drift condition for P involving an explicit higher order small-set E. In particular, this means finding an explicit positive integer ℓ and a computable minorizing measure ν_ℓ. Note that such ℓ might be too large to search for the explicit minorizing measure ν_ℓ, due to the difficult (or even impossible) computation of the iterate P^ℓ. In many of the papers cited in Subsections 5.5, 6.3 and 8.5 in link with Poisson's equation or convergence rates, the general statements are presented in the context of general small-sets or even petite sets. However, truly explicit bounds for the two above themes are usually only addressed with a first-order small-set (i.e. $\ell = 1$). Indeed, finding such explicit bounds under higher-order minorization conditions (i.e. $\ell \geq 2$) remains a difficult issue. Perhaps the most successful work in this regard is the recent paper [GLL25], where the authors provide a nice and explicit bound for solutions to Poisson's equation under modulated drift conditions involving a $\ell-$order small-set E. The bound obtained in [GLL25] via a randomized stopping time is close to (59b) in case $\ell = 1$, and it remains very simple when $\ell \geq 2$. Such a bound is proposed in Theorem 11.12 and Proposition 11.15, but the additional condition (227) is required and, above all, the bounds (230) and (235) are far from being as simple as that in [GLL25][1]. The main interest of Subsection 11.2 is that the proof of Theorem 11.12 is close to that of Theorem 5.4 (first-order case) and quite elementary compared to the more sophisticated one in [GLL25]. Finally recall that the interest of considering a higher-order small set has been also addressed in Subsection 8.4 to extend polynomial convergence rates under weaker conditions than (121).

[1]We thank the authors of [GLL25] for sharing their work with us. It has motivated the results of Subsection 11.2 on bound for the solutions to Poisson's equation under a higher-order minorization condition.

A Probabilistic complements

The split chain (e.g. see [Num84, Num02]).

Let $(X_n)_{n\geq 0}$ be a Markov chain on the space $(\mathbb{X}, \mathcal{X})$ with kernel transition P satisfying condition $(\boldsymbol{M}_{\nu,\psi})$ with $\nu \in \mathcal{M}^*_{+,b}, \psi \in \mathcal{B}^*_+$, that is

$$R := P - \psi \otimes \nu \geq 0.$$

Let us introduce the bivariate Markov chain $((X_n, Y_n))_{n\geq 0}$ with the state space $\mathbb{X} \times \{0, 1\}$ and the following transition kernel $\widehat{P}$: for every bounded measurable function f on $\mathbb{X} \times \{0, 1\}$

$$\mathbb{E}[f(X_{n+1}, Y_{n+1}) \mid \sigma(X_k, Y_k, k \leq n)] = \mathbb{E}[f(X_{n+1}, Y_{n+1}) \mid \sigma(X_n)] = (\widehat{P}f)(X_n)$$

with

$$\forall A \in \mathcal{X}, \quad \widehat{P}(x, A \times \{0\}) = R(x, A) \quad \widehat{P}(x, A \times \{1\}) = \psi(x)\,\nu(1_A).$$

$((X_n, Y_n))_{n\geq 0}$ is called *the split chain* associated with $(X_n)_{n\geq 0}$. Note that, for any $A \in \mathcal{X}$, $\widehat{P}(x, A \times \{0, 1\}) = \widehat{P}(x, A \times \{0\}) + \widehat{P}(x, A \times \{1\}) = P(x, A)$ so that the marginal process $(X_n)_{n\geq 0}$ is indeed the original Markov with transition kernel P. Next, for any $f \in \mathcal{B}$ and $x \in \mathbb{X}$, $\mathbb{E}[f(X_{n+1}) \mid X_n = x, Y_{n+1} = 1] = \nu(1_{\mathbb{X}})^{-1}\nu(f)$ for every $n \geq 1$. It follows that the set $\mathbb{X} \times \{1\}$ is an atom for the split chain. Let $\sigma_{\{1\}} := \inf\{n \geq 1, Y_n = 1\}$ be the return time to the atom $\mathbb{X} \times \{1\}$ of the split chain or the return time of $(Y_n)_{n\geq 0}$ to state 1. It is a regeneration times of the split chain. Such a material leads to using the so-called regeneration method to analyze the split chain and to deduce, when possible, the properties of the original Markov chain.

Probabilistic counterparts in terms of the split chain of various quantities in the present document.

Let us introduce the probability measure $\widehat{\nu} = \nu(1_{\mathbb{X}})^{-1}\nu$ on $\mathbb{X}$. The probability $\mathbb{P}$ when $\mathbb{X}_0$ has probability distribution η, is denoted by $\mathbb{P}_\eta$ and $\mathbb{E}_\eta$ is the expectation under $\mathbb{P}_\eta$.

$\forall A \in \mathcal{X}$ and $\forall x \in \mathcal{X}$:

- $(R^n 1_A)(x) = R^n(x, A) = \mathbb{P}_x\{X_n \in A, \sigma_{\{1\}} > n\}$;
  $(R^n 1_{\mathbb{X}})(x) = R^n(x, \mathbb{X}) = \mathbb{P}_x\{\sigma_{\{1\}} > n\}$;
  $\sum_{n=0}^{+\infty}(R^n 1_{\mathbb{X}})(x) = \mathbb{E}_x[\sigma_{\{1\}}]$;
- $h^\infty_R(x) := \lim_n (R^n 1_{\mathbb{X}})(x) = \mathbb{P}_x\{\sigma_{\{1\}} = +\infty\}$;
- $(R^{n-1}\psi)(x) = \mathbb{P}_x\{\sigma_{\{1\}} = n\}/\nu(1_{\mathbb{X}})$, $\sum_{k=1}^{n}(R^{k-1}\psi)(x) = \mathbb{P}_x\{\sigma_{\{1\}} \leq n\}/\nu(1_{\mathbb{X}})$;
  $\sum_{n=1}^{+\infty}(R^{n-1}\psi)(x) = \mathbb{P}_x\{\sigma_{\{1\}} < \infty\}/\nu(1_{\mathbb{X}})$;
- $\mu_R(1_A) = \nu(1_{\mathbb{X}})\sum_{n=0}^{+\infty}\mathbb{P}_{\widehat{\nu}}\{X_n \in A, \sigma_{\{1\}} > n\}$, $\mu_R(1_{\mathbb{X}}) = \nu(1_{\mathbb{X}})\mathbb{E}_{\widehat{\nu}}[\sigma_{\{1\}}]$
  $\mu_R(\psi) = \mathbb{P}_{\widehat{\nu}}\{\sigma_{\{1\}} < \infty\}$.
- Formula (17). For any $n \geq 1$, let $L_n := \min\{k = 0, \ldots, n-1 : Y_{n-k} = 1\}$, be the time elapsed since the last visit of $(Y_n)_{n\geq 0}$ to 1 before time n. Then $\{\sigma_{\{1\}} \leq n\} = \sqcup_{k=0}^{n-1}\{L_n = k\}$ and Formula (17) has the following probabilistic meaning
  $$\mathbb{P}_x\{X_n \in A\} = \mathbb{P}_x\{X_n \in A, \sigma_{\{1\}} > n\} + \sum_{k=0}^{n-1}\mathbb{P}_x\{X_n \in A, L_n = k\} .$$

Properties of the functions g_A^∞ and g_A.

Below Properties c)-d) of Subsection 2.2 are proved. Let P be a Markov kernel on $(\mathbb{X}, \mathcal{X})$, and let $(X_n)_{n\geq 0}$ be the canonical Markov chain on $(\mathbb{X}^{\mathbb{N}}, \mathcal{X}^{\otimes\mathbb{N}})$ associated with P, e.g. see [DMPS18, Sec. 3.1]. Denote by $\theta : \mathbb{X}^{\mathbb{N}} \to \mathbb{X}^{\mathbb{N}}$ the shift operator defined by:

$$\forall x = (x_n)_{n\in\mathbb{N}} \in \mathbb{X}^{\mathbb{N}}, \quad \theta(x) := (x_{n+1})_{n\in\mathbb{N}}.$$

Recall that Markov's property states that, for any r.v. Y on $(\mathbb{X}^{\mathbb{N}}, \mathcal{X}^{\otimes\mathbb{N}})$ taking its values in $[0, +\infty]$, we have (e.g. see [DMPS18, Th. 3.3.3])

$$\forall n \geq 1,\ \forall x \in \mathbb{X}, \quad \mathbb{E}_{X_n}[Y] = \mathbb{E}_x\big[Y \circ \theta^n \,|\, \sigma(X_0, \ldots, X_n)\big] \quad \mathbb{P}_x\text{−a.s..}$$

For any $A \in \mathcal{X}$, let $N_A := \sum_{n=1}^{+\infty} 1_{\{X_n\in A\}}$. Recall that: $\forall x \in \mathbb{X},\ g_A^\infty(x) := \mathbb{P}_x\{N_A = +\infty\}$. For every $x \in \mathbb{X}$ we have

$$\begin{aligned}
(Pg_A^\infty)(x) = \mathbb{E}_x[g_A^\infty(X_1)] &= \mathbb{E}_x\big[\mathbb{E}_{X_1}\big[1_{\{N_A=+\infty\}}\big]\big] && \text{(from the definition of } g_A^\infty) \\
&= \mathbb{E}_x\big[\mathbb{E}_x\big[\{1_{\{N_A=+\infty\}} \circ \theta \,|\, \sigma(X_0, X_1)\big]\big] && \text{(from Markov's property)} \\
&= \mathbb{E}_x\big[1_{\{N_A=+\infty\}} \circ \theta\big] = \mathbb{P}_x\{N_A = +\infty\} = g_A^\infty(x)
\end{aligned}$$

where we have used the classical property $\mathbb{E}\big[\mathbb{E}[Z \,|\, \mathcal{F}]\big] = \mathbb{E}[Z]$ of conditional expectation and the fact that the events $\{N_A \circ \theta = +\infty\}$ and $\{N_A = +\infty\}$ are obviously equal. Thus g_A^∞ is P−harmonic.

Recall that $T_A := \inf\{n \geq 0 : X_n \in A\}$ and that g_A is defined by: $\forall x \in \mathbb{X},\ g_A(x) = \mathbb{P}_x\{T_A < \infty\}$. Let $S_A := \inf\{n \geq 1 : X_n \in A\} = 1 + T_A \circ \theta$. Note that $\{S_A < \infty\} = \{T_A \circ \theta\ < \infty\}$. For every $x \in \mathbb{X}$ we have

$$\begin{aligned}
(Pg_A)(x) = \mathbb{E}_x[g_A(X_1)] &= \mathbb{E}_x\big[\mathbb{E}_{X_1}[1_{\{T_A<\infty\}}]\big] && \text{(from the definition of } g_A) \\
&= \mathbb{E}_x\big[\mathbb{E}_x[1_{\{T_A<\infty\}} \circ \theta \,|\, \sigma(X_0, X_1)]\big] && \text{(from Markov's property)} \\
&= \mathbb{E}_x[1_{\{T_A<\infty\}} \circ \theta] \\
&= \mathbb{E}_x[1_{\{S_A<\infty\}}] \\
&\leq \mathbb{E}_x[1_{\{T_A<\infty\}}] = g_A(x) && \text{(since } \{S_A < \infty\} \subset \{T_A < \infty\}).
\end{aligned}$$

Thus g_A is superharmonic. Since the kernel P is non-negative, the sequence of non-negative functions $(P^n g_A)_{n\geq 0}$ is non-increasing so it converges point-wise. Next, for every $x \in \mathbb{X}$, it follows from $(P^n g_A)(x) = \mathbb{E}_x[g_A(X_n)]$ and from the same arguments as above that

$$(P^n g_A)(x) = \mathbb{E}_x\big[\mathbb{E}_{X_n}[1_{\{T_A<\infty\}}]\big] = \mathbb{E}_x\big[\mathbb{E}_x\big[1_{\{T_A<\infty\}} \circ \theta^n \,|\, \sigma(X_0, \ldots, X_n)\big] = \mathbb{E}_x\big[1_{\{T_A<\infty\}} \circ \theta^n\big]$$

and that $(P^n g_A)(x) = \mathbb{P}_x(E_n)$ with $E_n := \cup_{k\geq n}\{X_k \in A\}$. We have $\{X_k \in A \text{ i.o.}\} = \cap_{n\geq 0} E_n = \big\{N_A = +\infty\big\}$, where i.o. stands in short for "infinitely often". Since $(E_n)_{n\geq 0}$ is non-increasing for the inclusion, we obtain that

$$\lim_n \searrow (P^n g_A)(x) = \lim_n \searrow \mathbb{P}_x(E_n) = \mathbb{P}_x\{X_k \in A \text{ i.o.}\} = \mathbb{P}_x\{N_A = +\infty\} = g_A^\infty(x).$$

B Proof of Theorem 4.12

From the definition of d in (43), there exists an integer $\ell_0 \geq 1$ such that the power series $\rho(z) := \sum_{n=1}^{+\infty} \nu(R^{n-1}\psi)\, z^n$ writes as

$$\forall z \in \overline{D}, \quad \rho(z) := \sum_{k=\ell_0}^{+\infty} \nu(R^{kd-1}\psi)\, z^{kd}. \tag{236}$$

The proof of Theorem 4.12 is based on the two following lemmas.

Lemma B.1 *Let P satisfy Condition $(\boldsymbol{M}_{\nu,\psi})$ with $\mu_R(1_{\mathbb{X}}) < \infty$. Then*

$$\lim_{n\to+\infty} P^{dn}\psi = \zeta_\psi := \frac{1}{m_d}\sum_{k=0}^{+\infty} R^{kd}\psi \quad \textit{with} \quad m_d := \sum_{k=\ell_0}^{+\infty} k\,\nu(R^{kd-1}\psi) < \infty. \tag{237}$$

Proof. Using the definition of the integer d, the arguments here are close to those used in the proof of the direct implication in Lemma 4.9. Note that $\sum_{k=0}^{+\infty} R^{dk}\psi$ is a bounded function on $\mathbb{X}$ from Proposition 3.4, and that $m_d < \infty$ from Remark 4.10. Now define

$$\forall z \in D, \quad \mathcal{P}_d(z) := \sum_{n=0}^{+\infty} z^n P^{dn}\psi, \quad \mathcal{R}_d(z) := \sum_{n=0}^{+\infty} z^n R^{dn}\psi, \quad \rho_d(z) := \sum_{k=\ell_0}^{+\infty} \nu(R^{kd-1}\psi)\, z^k.$$

with $D = \{z \in \mathbb{C} : |z| < 1\}$. Note that the power series ρ in (236) satisfies $\rho(z) = \rho_d(z^d)$. Thus $\rho_d(z)$ is not a power series in z^q for any integer $q \geq 2$: Indeed, otherwise we would have $\rho_d(z) := \sum_{\ell=\ell_0'}^{+\infty} \nu(R^{q\ell d-1}\psi)\, z^{q\ell}$ for some integers $\ell_0' \geq 1$ and $q \geq 2$, thus

$$\rho(z) = \sum_{\ell=\ell_0'}^{+\infty} \nu(R^{q\ell d-1}\psi)\, z^{q\ell d},$$

which contradicts the definition (43) of d. Moreover observe that $|\rho_d(z)| < 1$ for every $z \in D$ since $\mu_R(\psi) = \sum_{k=\ell_0}^{+\infty} \nu(R^{kd-1}\psi) = 1$ from Theorem 3.6. Now using (17) applied to ψ and the definition of d (see (236)) it follows that $P^{dn}\psi = R^{dn}\psi$ for every $n \in \{0, \ldots, \ell_0 - 1\}$ and that

$$\forall n \geq \ell_0, \quad P^{dn}\psi = R^{dn}\psi + \sum_{k=\ell_0}^{n} \nu(R^{dk-1}\psi) P^{d(n-k)}\psi.$$

Considering the associated power series and interchanging sums for the last term, we easily obtain that

$$\forall z \in D, \quad \mathcal{P}_d(z) = \mathcal{R}_d(z)\, U_d(z) \quad \text{with} \quad U_d(z) := \frac{1}{1 - \rho_d(z)}. \tag{238}$$

Next, we deduce from the Erdös-Feller-Pollard renewal theorem [EFP49] that the coefficients $u_{d,k}$ of the power series $U_d(z) = \sum_{k=0}^{+\infty} u_{d,k} z^k$ in (238) satisfy: $\lim_k u_{d,k} = 1/m_d$. Then, identifying the coefficients in Equation (238) (Cauchy product), we obtain that $P^{dn}\psi = \sum_{k=0}^{n} u_{d,n-k} R^{dk}\psi$ for every $n \geq 0$. Since $\sum_{k=0}^{+\infty} R^{dk}\psi < \infty$ from Proposition 3.4, Property (237) follows from Lebesgue theorem w.r.t. discrete measure. □

Lemma B.2 *Let P satisfy Condition $(\boldsymbol{M}_{\nu,\psi})$ with $\mu_R(1_{\mathbb{X}}) < \infty$ and $h_R^\infty = 0$. Then there exists a sequence $(\varepsilon_n)_n \in \mathcal{B}^{\mathbb{N}}$ such that $\lim_n \varepsilon_n = 0$ (point-wise convergence) and*

$$\forall h \in \mathcal{B},\ \|h\|_{1_{\mathbb{X}}} \leq 1,\ \exists \xi_h \in \mathcal{B}, \quad |P^{dn}h - \xi_h| \leq \varepsilon_n.$$

Proof. Here, using the definition of the integer d, the arguments are close to those used in the proof of Lemma 4.11. For $r = 0, \ldots, d-1$ set $\zeta_{r,\psi} := P^r\zeta_\psi$ with ζ_ψ given in (237).

Note that $\zeta_{r,\psi} \in \mathcal{B}$, and that $\lim_n P^{dn+r}\psi = \zeta_{r,\psi}$ (point-wise convergence) from Lebesgue's theorem w.r.t. $P^r(x,dy)$ for each $x \in \mathbb{X}$. Now for every $h \in \mathcal{B}$ define $\xi_h \in \mathcal{B}$ by

$$\xi_h := \sum_{r=0}^{d-1}\bigg(\sum_{j=1}^{+\infty}\nu(R^{dj-r-1}h)\bigg)\zeta_{r,\psi}. \tag{239}$$

Then using again (17) and observing that every integer $k = 1,\dots,dn$ writes as $k = dj - r$ for $r = 0,\dots,d-1$ and $j = 1,\dots,n$, we obtain that for every $n \geq 1$

$$P^{dn}h - \xi_h = R^{dn}h + \sum_{r=0}^{d-1}\sum_{j=1}^{n}\nu(R^{dj-r-1}h)\big(P^{d(n-j)+r}\psi - \zeta_{r,\psi}\big) - \sum_{r=0}^{d-1}\bigg(\sum_{j=n+1}^{+\infty}\nu(R^{dj-r-1}h)\bigg)\zeta_{r,\psi}.$$

Thus, if $\|h\|_{1_{\mathbb{X}}} \leq 1$ (i.e. $|h| \leq 1_{\mathbb{X}}$), then we have $|P^{dn}h - \xi_h| \leq \varepsilon_n$ with $\varepsilon_n \in \mathcal{B}$ defined by

$$\varepsilon_n := R^{dn}1_{\mathbb{X}} + \sum_{r=0}^{d-1}\sum_{j=1}^{n}\nu(R^{dj-r-1}1_{\mathbb{X}})\big|P^{d(n-j)+r}\psi - \zeta_{r,\psi}\big| + \sum_{r=0}^{d-1}\|\zeta_{r,\psi}\|_{1_{\mathbb{X}}}\sum_{j=n+1}^{+\infty}\nu(R^{dj-r-1}1_{\mathbb{X}}).$$

We have $\lim_n \varepsilon_n = 0$ (point-wise convergence). Indeed, the last term converges to zero when $n \to +\infty$ since $\sum_{k=0}^{+\infty}\nu(R^k 1_{\mathbb{X}}) = \mu_R(1_{\mathbb{X}}) < \infty$; The second sum above converges to zero when $n \to +\infty$ from Lebesgue's theorem w.r.t. discrete measure recalling that $\lim_n P^{dn+r}\psi = \zeta_{r,\psi}$; Finally $\lim_n R^{dn}1_{\mathbb{X}} = 0$ from $h_R^\infty = 0$.

□

Proof of Theorem 4.12. Let $g \in \mathcal{B}$ be such that $|g| \leq 1_{\mathbb{X}}$. Note that for $r = 0,\dots,d-1$ we have $|P^r g| \leq P^r|g| \leq P^r 1_{\mathbb{X}} = 1_{\mathbb{X}}$. Thus for $r = 0,\dots,d-1$ we can consider $\xi_{r,g} := \xi_{P^r g}$, where $\xi_{P^r g}$ is the function of Lemma B.2 associated to $h = P^r g$. Let $\gamma_g = \frac{1}{d}\sum_{r=0}^{d-1}\xi_{r,g}$. Then

$$\bigg|\gamma_g - \frac{1}{d}\sum_{r=0}^{d-1}P^{nd+r}g\bigg| \leq \frac{1}{d}\sum_{r=0}^{d-1}\big|\xi_{r,g} - P^{nd}(P^r g)\big| \leq \varepsilon_n \tag{240}$$

from Lemma B.2. Thus we have $\gamma_g = \lim_n \frac{1}{d}\sum_{r=0}^{d-1}P^{nd+r}g$ (point-wise convergence). From Lebesgue's theorem w.r.t. $P(x,dy)$ for each $x \in \mathbb{X}$, we then obtain that

$$P\gamma_g = \lim_{n\to+\infty}\frac{1}{d}\sum_{r=1}^{d}P^{nd+r}g = \gamma_g \tag{241}$$

the last equality being obviously deduced from $\lim_{n\to+\infty}P^{nd+d}g = \lim_{n\to+\infty}P^{nd}g$. Thus γ_g is a $P-$harmonic function, so that $\gamma_g = c_g 1_{\mathbb{X}}$ for some constant c_g from Theorem 4.1. Moreover, using the second equality of (241) and applying Lebesgue's theorem w.r.t. the $P-$invariant probability measure π_R, we obtain that $\pi_R(g) = \pi_R(\gamma_g)$, so $\gamma_g = \pi_R(g)1_{\mathbb{X}}$. Finally, applying the function inequality (240) to any fixed $x \in \mathbb{X}$ and taking the supremum on all the functions $g \in \mathcal{B}$ such that $|g| \leq 1_{\mathbb{X}}$, we obtain the desired total variation convergence of Theorem 4.12 since $\lim_n \varepsilon_n(x) = 0$ from Lemma B.2. □

C Proof of Lemmas 7.4, 7.8 and 7.9

Proof of Lemma 7.4. We deduce from the definitions of $\widehat{P}_k$ and $\widehat{\pi}_k$ that

$$\forall y \in {B_k}^c, \quad \sum_{x\in\mathbb{N}}\widehat{P}_k(x,y)\,\widehat{\pi}_k(\{x\}) = 0 = \widehat{\pi}_k(\{y\}).$$

Using successively the above equality, the definitions of $\widehat{\pi}_k$ and $\widehat{P}_k$, the P_k-invariance of π_k, and again the definition of $\widehat{\pi}_k$, we obtain

$$\begin{aligned}\forall y \in B_k, \quad \sum_{x\in\mathbb{N}} \widehat{P}_k(x,y)\,\widehat{\pi}_k(\{x\}) &= \sum_{x\in B_k} \widehat{P}_k(x,y)\,\widehat{\pi}_k(\{x\})\\ &= \sum_{x\in B_k} P_k(x,y)\,\pi_k(\{x\}) = \pi_k(\{y\}) = \widehat{\pi}_k(\{y\}).\end{aligned}$$

Thus $\widehat{\pi}_k$ is a $\widehat{P}_k-$invariant probability measure. To prove the uniqueness, consider any $\widehat{P}_k-$invariant probability measure $\widehat{\eta} = (\widehat{\eta}(\{x\}))_{x\in\mathbb{N}}$. Then

$$\forall y \in {B_k}^c, \quad \widehat{\eta}(\{y\}) = \sum_{x\in\mathbb{N}} \widehat{P}_k(x,y)\,\widehat{\eta}(\{x\}) = 0$$

from the definition of $\widehat{P}_k$. Thus

$$\forall y \in B_k, \quad \widehat{\eta}(\{y\}) = \sum_{x\in\mathbb{N}} \widehat{P}_k(x,y)\,\widehat{\eta}(\{x\}) = \sum_{x\in B_k} \widehat{P}_k(x,y)\,\widehat{\eta}(\{x\}) = \sum_{x\in B_k} P_k(x,y)\,\widehat{\eta}(\{x\})$$

from the definition of $\widehat{P}_k$. Thus $\eta := (\widehat{\eta}(\{x\}))_{x\in B_k}$ is a P_k-invariant probability measure on B_k. This proves that $\widehat{\eta} = \widehat{\pi}_k$. □

Proof of Lemma 7.8. Recall that $b_k := 1_{{\mathbb{X}_k}^c}$ and $\mathcal{F}_k$ is the finite-dimensional space with basis $\mathcal{C}_k := \big\{1_{\mathbb{X}_{i,k}},\ i \in I_k\big\} \cup \{b_k\}$. The $N_k \times N_k-$matrix B_k is defined as the matrix of P_k with respect to $\mathcal{C}_k$ with $N_k := \dim \mathcal{F}_k = \mathrm{Card}\,(I_k) + 1$. Note that

$$P_k b_k = \widehat{P}_k b_k = \widehat{Q}_k b_k + b_k(x_0)\psi_k = 0. \tag{242}$$

Since $g \in \mathcal{F}_k$ writes in the basis $\mathcal{C}_k$ as $g = \sum_{i\in I_k} g(x_{i,k}) + g(\overline{x}_k)b_k$ where $x_{i,k} \in \mathbb{X}_{i,k}$ and $\overline{x}_k \in \mathbb{X} \setminus \mathbb{X}_k$, we can write for every $j \in I_k$

$$\begin{aligned}&P_k 1_{\mathbb{X}_{j,k}}\\ &= \widehat{P}_k 1_{\mathbb{X}_{j,k}} = \sum_{i\in I_k} (\widehat{P}_k 1_{\mathbb{X}_{j,k}})(x_{i,k})\,1_{\mathbb{X}_{i,k}} + (\widehat{P}_k 1_{\mathbb{X}_{j,k}})(\overline{x}_k)\,b_k \qquad (\text{since } P_k 1_{\mathbb{X}_{j,k}} \in \mathcal{F}_k)\\ &= \sum_{i\in I_k} \big[(\widehat{Q}_k 1_{\mathbb{X}_{j,k}})(x_{i,k}) + 1_{\mathbb{X}_{j,k}}(x_0)\,\psi_k(x_{i,k})\big]\,1_{\mathbb{X}_{i,k}} + \big[(\widehat{Q}_k 1_{\mathbb{X}_{j,k}})(\overline{x}_k) + 1_{\mathbb{X}_{j,k}}(x_0)\,\psi_k(\overline{x}_k)\big]\,b_k\\ &= \sum_{i\in I_k} \big[(\widehat{Q}_k 1_{\mathbb{X}_{j,k}})(x_{i,k}) + 1_{\mathbb{X}_{j,k}}(x_0)\,\psi_k(x_{i,k})\big]\,1_{\mathbb{X}_{i,k}} + 1_{\mathbb{X}_{j,k}}(x_0)\,b_k.\end{aligned}$$

The previous equalities show that B_k is a non-negative matrix. Moreover Equality $P_k 1_{\mathbb{X}} = 1_{\mathbb{X}}$ reads as matrix equality $B_k \cdot \mathbf{1}_k = \mathbf{1}_k$ where $\mathbf{1}_k$ is the coordinate vector of $1_{\mathbb{X}}$ in the basis $\mathcal{C}_k$. Thus B_k is a stochastic matrix. □

Proof of Lemma 7.9. Recall that b_k is defined by $b_k = 1_{\mathbb{X}} - \sum_{i\in I_k} 1_{\mathbb{X}_{i,k}}$. From $\psi_k := 1_{\mathbb{X}} - \widehat{Q}_k 1_{\mathbb{X}}$ it follows that $\psi_k = b_k + \sum_{i\in I_k} 1_{\mathbb{X}_{i,k}} - \widehat{Q}_k 1_{\mathbb{X}}$. Define

$$m_{i,k}(f) := \int_{\mathbb{X}_k} f(y) \inf_{t\in\mathbb{X}_{i,k}} p(t,y)\,d\mu(y)$$

and observe that $\widehat{Q}_k f = \sum_{i\in I_k} m_{i,k}(f)\, 1_{\mathbb{X}_{i,k}}$. Then we deduce from (101) and (102) that

$$\begin{aligned}\widehat{P}_k f := (\widehat{Q}_k f) + f(x_0)\,\psi_k &= \sum_{i\in I_k} m_{i,k}(f)\, 1_{\mathbb{X}_{i,k}} + f(x_0)\big(b_k + \sum_{i\in I_k} 1_{\mathbb{X}_{i,k}} - \widehat{Q}_k 1_{\mathbb{X}}\big)\\ &= \sum_{i\in I_k} \big[m_{i,k}(f) + f(x_0) - f(x_0)\, m_{i,k}(1_{\mathbb{X}})\big]\, 1_{\mathbb{X}_{i,k}} + f(x_0) b_k,\end{aligned}$$

so that (104) and $\sum_{i\in I_k} \pi_{i,k} = 1$ give

$$\begin{aligned}\widehat{\pi}_k(f) &:= \sum_{i\in I_k} \pi_{i,k}\, [m_{i,k}(f) + f(x_0) - f(x_0)\, m_{i,k}(1_{\mathbb{X}})]\\ &= \sum_{i\in I_k} \pi_{i,k}\, m_{i,k}(f) + f(x_0)\bigg(1 - \sum_{i\in I_k} \pi_{i,k}\, m_{i,k}(1_{\mathbb{X}})\bigg). \end{aligned} \tag{243}$$

This proves Formula (105a). Now we prove that $\widehat{\pi}_k$ defines a $\widehat{P}_k$-invariant probability measure on $(\mathbb{X}, \mathcal{X})$. Note that

$$\forall i \in I_k, \quad m_{i,k}(1_{\mathbb{X}}) \le \int_{\mathbb{X}} p(x_{i,k}, y)\, d\mu(y) = (P1_{\mathbb{X}})(x_{i,k}) = 1,$$

thus

$$\int_{\mathbb{X}} \mathfrak{p}_k(y)\, d\mu(y) = \sum_{i\in I_k} \pi_{i,k}\, m_{i,k}(1_{\mathbb{X}}) \le 1.$$

It follows from this remark and from (243) that $\widehat{\pi}_k$ is a probability measure on $\mathbb{X}$. Finally $B_k \cdot F_k$ is the coordinate vector of $\widehat{P}_k^2 f$ in $\mathcal{C}_k$ since $\widehat{P}_k f \in \mathcal{F}_k$ and F_k is the coordinate vector of $\widehat{P}_k f$ in $\mathcal{C}_k$. Consequently we deduce from (104) and (103) that

$$\widehat{\pi}_k(\widehat{P}_k f) := \pi_k\, B_k\, F_k = \pi_k\, F_k = \widehat{\pi}_k(f).$$

Thus $\widehat{\pi}_k$ is $\widehat{P}_k$-invariant.

□

D Proof of Theorem 9.1 and Proposition 9.3

Here we assume that P satisfy Condition $(\boldsymbol{M}_{\nu,\psi})$ with $h_R^\infty = 0$ and $\mu_R(1_{\mathbb{X}}) < \infty$, and that $P \in \mathcal{L}(\mathfrak{B})$ where $(\mathfrak{B}, \|\cdot\|)$ is a Banach space satisfying Assumptions **(B)**. The properties of Lemma 9.2 are repeatedly used below, that is: $R \in \mathcal{L}(\mathfrak{B})$, the radius of convergence of the power series $\rho(z) := \sum_{n=1}^{+\infty} \nu(R^{n-1}\psi)\, z^n$ is larger than $1/r_{\mathfrak{B}}$ where $r_{\mathfrak{B}}$ denotes the spectral radius of R on $\mathfrak{B}$, and finally the series $\widetilde{g}_z := \sum_{k=0}^{+\infty} z^{-(k+1)} R^k g$ absolutely converges in $\mathfrak{B}$ for every $z \in \mathbb{C}$ such that $|z| > r_{\mathfrak{B}}$ and for every $g \in \mathfrak{B}$.

Lemma D.1 *If $r_{\mathfrak{B}} < 1$, then the following assertions hold for every $z \in \mathbb{C}$ such that $|z| > r_{\mathfrak{B}}$. The operator $zI - P$ is invertible on $\mathfrak{B}$ if, and only if, we have $\rho(z^{-1}) \neq 1$. Moreover, if $\rho(z^{-1}) = 1$, then z is an eigenvalue of P on $\mathfrak{B}$, and $E_z := \{g \in \mathfrak{B} : Pg = zg\} = \mathbb{C} \cdot \widetilde{\psi}_z$ with $\widetilde{\psi}_z := \sum_{k=0}^{+\infty} z^{-(k+1)} R^k \psi$ being non zero in $\mathfrak{B}$ and satisfying $\nu(\widetilde{\psi}_z) = 1$.*

Proof. Let $z \in \mathbb{C}$ be such that $|z| > r_{\mathfrak{B}}$. Assume that $zI - P$ is not one-to-one on $\mathfrak{B}$, that is: $\exists g \in \mathfrak{B}, g \neq 0, Pg = zg$. Note that $\lim_n |z|^{-n} \|R^n g\| = 0$ using the definition of $r_{\mathfrak{B}}$ and $|z| > r_{\mathfrak{B}}$ (use (160) with $\gamma \in (r_{\mathfrak{B}}, |z|)$). Since $R \in \mathcal{L}(\mathfrak{B})$, Equality (46) of Lemma 4.16 can be proved similarly, that is we have:

$$\forall n \geq 0, \quad \nu(g) \sum_{k=0}^{n} z^{-(k+1)} R^k \psi = g - z^{-(n+1)} R^{n+1} g.$$

Then the following equality holds in $\mathfrak{B}$

$$g = \nu(g) \sum_{k=0}^{+\infty} z^{-(k+1)} R^k \psi$$

and $\nu(g) \neq 0$ since g is assumed to be non-zero. Note that $g \mapsto \nu(g)$ is a continuous linear map from $\mathfrak{B}$ to $\mathbb{C}$ due to (159). Thus, integrating the previous equality w.r.t. ν, we obtain that $\nu(g) = \nu(g)\rho(z^{-1})$, thus $\rho(z^{-1}) = 1$. We have proved by contrapositive that $|z| > r_{\mathfrak{B}}$ and $\rho(z^{-1}) \neq 1$ imply that $zI - P$ is one-to-one. Now prove that $|z| > r_{\mathfrak{B}}$ and $\rho(z^{-1}) \neq 1$ imply that $zI - P$ is surjective on $\mathfrak{B}$. Let $z \in \mathbb{C}$ be such that $|z| > r_{\mathfrak{B}}$, let $g \in \mathfrak{B}$ and define

$$\forall n \geq 1, \quad \widetilde{g}_{n,z} := \sum_{k=0}^{n} z^{-(k+1)} R^k g.$$

Using $P = R + \psi \otimes \nu$ we obtain that

$$z\widetilde{g}_{n,z} - P\widetilde{g}_{n,z} = z\,\widetilde{g}_{n,z} - R\widetilde{g}_{n,z} - \nu(\widetilde{g}_{n,z})\psi = g - z^{-(n+1)} R^{n+1} g - \nu(\widetilde{g}_{n,z})\psi. \tag{244}$$

Next the following convergences hold, in $\mathfrak{B}$ for the first two, in $\mathbb{C}$ for the last one

$$\lim_{n \to +\infty} \widetilde{g}_{n,z} = \widetilde{g}_z := \sum_{k=0}^{+\infty} z^{-(k+1)} R^k g, \quad \lim_n P\widetilde{g}_{n,z} = P\widetilde{g}_z, \quad \lim_{n \to +\infty} \nu(\widetilde{g}_{n,z}) = \nu(\widetilde{g}_z) \tag{245}$$

from Lemma 9.2 (use $P \in \mathcal{L}(\mathfrak{B})$ for the second one). Then, passing to the limit when $n \to +\infty$ in (244) provides the following equality in $\mathfrak{B}$:

$$(zI - P)\widetilde{g}_z = g - \nu(\widetilde{g}_z)\psi. \tag{246}$$

In particular, with $g = \psi$, we obtain that

$$(zI - P)\widetilde{\psi}_z = \big(1 - \rho(z^{-1})\big)\psi \quad \text{with} \quad \widetilde{\psi}_z := \sum_{k=0}^{+\infty} z^{-(k+1)} R^k \psi.$$

since $\nu(\widetilde{\psi}_z) = \rho(z^{-1})$. Consequently, if $\rho(z^{-1}) \neq 1$, then

$$(zI - P)\left(\widetilde{g}_z + \frac{\nu(\widetilde{g}_z)}{1 - \rho(z^{-1})}\widetilde{\psi}_z\right) = g,$$

from which we deduce that $zI - P$ is surjective since $\widetilde{g}_z$ and $\widetilde{\psi}_z$ belong to $\mathfrak{B}$.

We have proved that, if $z \in \mathbb{C}$ is such that $|z| > r_{\mathfrak{B}}$, then $\rho(z^{-1}) \neq 1$ implies that $zI - P$ is invertible on $\mathfrak{B}$. Conversely let $z \in \mathbb{C}$ be such that $|z| > r_{\mathfrak{B}}$ and $\rho(z^{-1}) = 1$. Let us prove that $zI - P$ is not invertible on $\mathfrak{B}$. Recall that the series $\widetilde{\psi}_z := \sum_{k=0}^{+\infty} z^{-(k+1)} R^k\psi$ absolutely converges in $\mathfrak{B}$ and that $\nu(\widetilde{\psi}_z) = \rho(z^{-1}) = 1$ from Lemma 9.2. Moreover we have $\widetilde{\psi}_z \neq 0$ in $\mathfrak{B}$. This is obvious from $\nu(\widetilde{\psi}_z) \neq 0$ if $\mathfrak{B}$ is a space composed of functions. This is also true if $\mathfrak{B}$ is a space composed of classes of functions modulo π_R : Indeed $\widetilde{\psi}_z = 0$ in $\mathfrak{B}$ would imply that $\widetilde{\psi}_z = 0$ π_R−a.s., which is impossible since $\nu(\widetilde{\psi}_z) \neq 0$ and ν is absolutely continuous w.r.t. π_R from the inequality $\nu \leq \pi_R(\psi)^{-1}\pi_R$ derived from the minorization condition $(\boldsymbol{M}_{\nu,\psi})$ and the P−invariance of π_R with $\pi_R(\psi) > 0$. Next the equalities in (47) can be applied to prove Equality $P\widetilde{\psi}_z = z\widetilde{\psi}_z$ in $\mathfrak{B}$. Thus $zI - P$ is not one-to-one on $\mathfrak{B}$, thus is not invertible on $\mathfrak{B}$. Finally, the fact that $E_z = \mathbb{C} \cdot \widetilde{\psi}_z$ follows from the first part of the proof. □

Now let $\mathfrak{B}_0 := \{g \in \mathfrak{B} : \pi_R(g) = 0\}$. Note that $\mathfrak{B}_0$ is a closed subspace of $\mathfrak{B}$ since the linear form $g \mapsto \pi_R(g)$ is continuous from $\mathfrak{B}$ to $\mathbb{C}$ from Condition (156). Thus $(\mathfrak{B}_0, \|\cdot\|)$ is a Banach space. Moreover $\mathfrak{B}_0$ is P−stable (i.e. $P(\mathfrak{B}_0) \subset \mathfrak{B}_0$) from the P−invariance of π_R. Let P_0 denote the restriction of P to $\mathfrak{B}_0$.

Lemma D.2 *If* $r_{\mathfrak{B}} < 1$*, then* $I - P_0$ *is invertible on* $(\mathfrak{B}_0, \|\cdot\|)$.

Proof. From (246) applied to $z = 1$, we obtain that

$$\forall g \in \mathfrak{B}, \quad (I - P)\widetilde{g}_1 = g - \mu_R(1_{\mathbb{X}})\pi_R(g)\psi \quad \text{with} \quad \widetilde{g}_1 := \sum_{k=0}^{+\infty} R^k g \ \in \mathfrak{B}$$

since $\nu(\widetilde{g}_1) = \mu_R(g) = \mu_R(1_{\mathbb{X}})\pi_R(g)$ from (26). Hence, if $\pi_R(g) = 0$, then $\widetilde{g}_1$ is solution to Poisson equation $(I - P)\widetilde{g}_1 = g$. Moreover we know from Lemma D.1 that $E_1 := \{g \in \mathfrak{B} : Pg = g\}$ has dimension one, i.e. $E_1 = \mathbb{C} \cdot 1_{\mathbb{X}}$. Hence two solutions to Poisson's equation in $\mathfrak{B}$ differ from an additive constant. Consequently $\widehat{g}_1 := \widetilde{g}_1 - \pi_R(\widetilde{g}_1)1_{\mathbb{X}}$ is the unique π_R−centered solution in $\mathfrak{B}$ to Poisson's equation $(I - P)\widehat{g} = g$. This proves the claimed statement. □

Proof of Theorem 9.1. Let $z \in \mathbb{C}$ be such that $|z| > r_{\mathfrak{B}}$, $z \neq 1$, and $\rho(z^{-1}) \neq 1$. Then $zI - P$ is invertible on $\mathfrak{B}$ from Lemma D.1. Thus $zI - P_0$ is also one-to-one on $\mathfrak{B}_0$. Now, let $g \in \mathfrak{B}_0$. From Lemma D.1 there exists $h \in \mathfrak{B}$ such that $(zI - P)h = g$, thus $(z - 1)\pi_R(h) = \pi_R(g) = 0$ from the P−invariance of π_R. Hence $\pi_R(h) = 0$ (i.e. $h \in \mathfrak{B}_0$) since $z \neq 1$, and consequently $zI - P_0$ is surjective on $\mathfrak{B}_0$. We have proved that, for any $z \in \mathbb{C}$ such that $|z| > r_{\mathfrak{B}}$, $z \neq 1$, and $\rho(z^{-1}) \neq 1$, the operator $zI - P_0$ is invertible on $\mathfrak{B}_0$. Moreover we know that $I - P_0$ is invertible on $\mathfrak{B}_0$ from Lemma D.2.

Now recall that $\rho(z^{-1}) \neq 1$ for every $z \in \mathbb{C}$ such that $|z| = 1$, $z \neq 1$, from the aperiodicity condition (39) (i.e. $z = 1$ is the only complex number of modulus one solution to $\rho(z^{-1}) = 1$). Moreover, if $z \in \mathbb{C}$ is such that $|z| > 1$, then $\rho(z^{-1}) \neq 1$ since

$$|\rho(z^{-1})| \leq \sum_{n=1}^{+\infty} \nu(R^{n-1}\psi)\, |z|^{-n} < \sum_{n=1}^{+\infty} \nu(R^{n-1}\psi) = \mu_R(\psi) = 1.$$

Let ϱ_0 denote the spectral radius of P_0 on $\mathfrak{B}_0$, and recall that the prerequisites in spectral theory are given by $(S1)$-$(S3)$ in Subsection 6.2. From the above we then obtain that $\varrho_0 < 1$ and that the following alternative holds:

(a') If Equation $\rho(z^{-1}) = 1$ has no solution $z \in \mathbb{C}$ such that $r_{\mathfrak{B}} < |z| < 1$, then $zI - P_0$ is invertible on $\mathfrak{B}_0$ for every $z \in \mathbb{C}$ such that $|z| > r_{\mathfrak{B}}$. Thus $\varrho_0 \leq r_{\mathfrak{B}}$.

(b') Otherwise, we have $\varrho_0 = \max\big\{|z| : z \in \mathbb{C},\ \rho(z^{-1}) = 1,\ r_{\mathfrak{B}} < |z| < 1\big\}$.

Moreover recall that $\varrho_0 = \lim_n(\|P_0^n\|_0)^{1/n}$ from Gelfand's formula, where $\|\cdot\|_0$ denotes the operator norm on $\mathfrak{B}_0$. Let $\rho \in (\varrho_0, 1)$. Then there exists a positive constant c_ρ such that: $\|P_0^n\|_0 \le c_\rho \rho^n$. Thus

$$\begin{aligned}
\forall n \ge 1,\ \forall g \in \mathfrak{B}, \quad \|P^n g - \pi_R(g)1_{\mathbb{X}}\| &= \|P^n(g - \pi_R(g)1_{\mathbb{X}})\| \quad (\text{from } P^n 1_{\mathbb{X}} = 1_{\mathbb{X}}) \\
&= \|P_0^n(g - \pi_R(g)1_{\mathbb{X}})\| \quad (\text{since } g - \pi_R(g)1_{\mathbb{X}} \in \mathcal{B}_0) \\
&\le c_\rho \rho^n \|g - \pi_R(g)1_{\mathbb{X}}\| \quad (\text{from } \|P_0^n\|_0 \le c_\rho \rho^n) \\
&\le c_\rho\big(1 + c\|1_{\mathbb{X}}\|\big)\rho^n \|g\| \quad (\text{from (156)}).
\end{aligned}$$

Using the definition (158) of $\varrho_{\mathfrak{B}}$, we then obtain that $\varrho_{\mathfrak{B}} \le \varrho_0$ since ρ is any real number in $(\varrho_0, 1)$. Hence Case (a) of Theorem 9.1 which corresponds to Case (a') is proved. To prove Case (b) of Theorem 9.1 which corresponds to the above case (b'), consider $z \in \mathbb{C}$ such that $r_{\mathfrak{B}} < |z| < 1$, $\rho(z^{-1}) = 1$ and $|z| = \varrho_0$. Then z is an eigenvalue of P from Lemma D.1, i.e. $\exists g \in \mathfrak{B}, g \ne 0, Pg = zg$. Moreover, from the $P-$invariance of π_R, we have $\pi_R(g) = z\pi_R(g)$, thus $\pi_R(g) = 0$ since $z \ne 1$. Hence we have: $\forall n \ge 1,\ \|P^n g - \pi_R(g)1_{\mathbb{X}}\| = \|P^n g\| = \varrho_0^n \|g\|$. It then follows from the definition of $\varrho_{\mathfrak{B}}$ that $\varrho_{\mathfrak{B}} \ge \varrho_0$. Thus $\varrho_{\mathfrak{B}} = \varrho_0$ in Case (b). Theorem 9.1 is proved. □

Proof of Proposition 9.3. In case (b) we know that, for $r \in (r_{\mathfrak{B}}, 1)$ sufficiently close to $r_{\mathfrak{B}}$, the set $\mathcal{S}_r := \{z \in \mathbb{C},\ \rho(z^{-1}) = 1,\ r \le |z| < 1\}$ is non-empty. Moreover $\mathcal{S}_r$ is finite from the analyticity of the power series $\rho(\cdot)$. The last assertion of Proposition 9.3 is proved in Lemma D.1. □

E Proof of Lemma 9.11

Using $P = R + T$ it follows from Lemma 9.10 that $P = P^* = R_1 + U_1$ with $R_1 = R^*$ and $U_1 = T^*$ defined by: $\forall g \in \mathbb{L}^2(\pi_R),\ U_1 g = \pi_R(\psi g)\zeta$. Now for $n \ge 2$ set $U_n := P^n - R_1^n$. Note that Property (170) is equivalent to

$$\forall n \ge 1,\ \forall g \in \mathbb{L}^2(\pi_R), \quad U_n g = \sum_{k=1}^{n} \pi_R(g \cdot R^{k-1}\psi) P^{n-k}\zeta. \tag{247}$$

Property (247) is obvious for $n = 1$ from the definition of U_1 and R_1. Next we have

$$\forall n \ge 2, \quad P^n - U_n = R_1^n = R_1^{n-1} R_1 = (P^{n-1} - U_{n-1})(P - U_1),$$

so that

$$\forall n \ge 2, \quad U_n = P^{n-1}U_1 + U_{n-1}R_1 = P^{n-1}U_1 + U_{n-1}R^*. \tag{248}$$

Now, if for some $n \ge 2$ we have

$$\forall g \in \mathbb{L}^2(\pi_R), \quad U_{n-1} g = \sum_{k=1}^{n-1} \pi_R(g \cdot R^{k-1}\psi) P^{n-1-k}\zeta,$$

then we deduce from (248) that

$$\begin{aligned}\forall g \in \mathbb{L}^2(\pi_R), \quad U_n g &= \pi_R(\psi g) P^{n-1}\zeta + \sum_{k=1}^{n-1} \pi_R(R^* g \cdot R^{k-1}\psi) P^{n-1-k}\zeta \\ &= \pi_R(\psi g) P^{n-1}\zeta + \sum_{k=1}^{n-1} \pi_R(g \cdot R^{k}\psi) P^{n-1-k}\zeta \\ &= \sum_{k=1}^{n} \pi_R(g \cdot R^{k-1}\psi) P^{n-k}\zeta.\end{aligned}$$

Property (247) is proved by induction.

References

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

[AF10] Y. Atchadé and G. Fort. Limit theorems for some adaptive MCMC algorithms with subgeometric kernels. *Bernoulli*, 16(1):116–154, 2010.

[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.

[ALPW24] C. Andrieu, A. Lee, A. Power, and A. Q. Wang. Explicit convergence bounds for Metropolis Markov chains: Isoperimetry, spectral gaps and profiles. *Ann. Appl. Probab.*, 34(4):4022 – 4071, 2024.

[AMP05] C. Andrieu, É. Moulines, and P. Priouret. Stability of stochastic approximation under verifiable conditions. *SIAM J. Control and Optimization*, 44(1):283–312, 2005.

[AN78] K. B. Athreya and P. Ney. A new approach to the limit theory of recurrent markov chains. *Trans. Amer. Math. Soc.*, 245:493–501, 1978.

[Atc16] Y. F. Atchadé. Markov chain Monte Carlo confidence intervals. *Bernoulli*, 22(3):1808–1838, 2016.

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

[Bax11] P. Baxendale. T. E. Harris's contributions to recurrent Markov processes and stochastic flows. *Ann. Probab.*, 39(2):417–428, 2011.

[BGJM11] S. Brooks, A. Gelman, G. L. Jones, and X.-L. Meng. Handbook of Markov chain Monte Carlo. In *Handbooks of modern statistical methods*. Chapman & Hall/CRC, 2011.

[BGT87] N.H. Bingham, C.M. Goldie, and J.L. Teugels. *Regular Variation*. Cambridge University Press, 1987.

[BH22] M. Benaïm and T. Hurth. *Markov Chains on Metric Spaces—a Short Course*. Springer, Cham, 2022.

[BJ24a] R. Bhattacharya and G. L. Jones. Explicit constraints on the geometric rate of convergence of random walk Metropolis-Hastings. *Bernoulli*, 2024. to appear.

[BJ24b] A. Brown and G. L. Jones. Convergence rates of Metropolis-Hastings algorithms. *WIRE Computational statistics*, 16(5), 2024.

[BP79] A Berman and R. J. Plemmons. *Nonnegative Matrices in the Mathematical Sciences*. Academic Press, 1979.

[Bré99] P. Brémaud. *Markov Chains*. Springer, 1999.

[BRR01] L. Breyer, G. O. Roberts, and J. S. Rosenthal. A note on geometric ergodicity and floating-point roundoff error. *Statist. Probab. Lett.*, 53(2):123–127, 2001.

[But14] O. Butkovsky. Subgeometric rates of convergence of Markov processes in the Wasserstein metric. *Ann. Appl. Probab.*, 24(2):526–552, 2014.

[CDJT24] T. Cui, J. Dong, A. Jasra, and X. T. Tong. Convergence speed and approximation accuracy of numerical MCMC. *Adv. Appl. Probab.*, pages 1–33, 2024.

[CF09] S. B. Connor and G. Fort. State-dependent Foster-Lyapunov criteria for subgeometric convergence of Markov chains. *Stochastic Process. Appl.*, 119(12):4176–4193, 2009.

[CMR05] O. Cappé, E. Moulines, and T. Rydén. *Inference in Hidden Markov Models*. Springer, New York, 2005.

[CnM23] J. A. Cañizo and S. Mischler. Harris-type results on geometric and subgeometric convergence to equilibrium for stochastic semigroups. *J. Functional Analysis*, 284(7):109830, 2023.

[DCWY19] R. Dwivedi, Y. Chen, M. J. Wainwright, and B. Yu. Log-concave sampling: Metropolis-Hastings algorithms are fast. *Journal of Machine Learning Research*, 2019.

[DDA21] V. Debavelaere, S. Durrleman, and S. Allassonnière. On the convergence of stochastic approximations under a subgeometric ergodic Markov dynamic. *Electron. J. Stat.*, 15(1):1583–1609, 2021.

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

[DF37] W. Doeblin and R. Fortet. Sur des chaînes à liaisons complètes. *Bull. Soc. Math. France*, 65:132–148, 1937.

[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.

[DG15] J. Dedecker and S. Gouëzel. Subgaussian concentration inequalities for geometrically ergodic Markov chains. *Electron. Commun. Probab.*, 20:no. 64, 12, 2015.

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

[DMS07] R. Douc, E. Moulines, and P. Soulier. Computable convergence rates for subgeometric ergodic Markov chains. *Bernoulli*, 13(3):831–848, 2007.

[DMS14] R. Douc, E. Moulines, and D. Stoffer. *Nonlinear Time Series*. CRC Press, 2014.

[Doe37] W. Doeblin. Sur les propriété es asymptotiques de mouvement régis par certains types de chaînes simples. *Bull. Math. Soc. Roum. Sci.*, 39(1) and 39(2):57–115 and 3–61, 1937.

[EFP49] P. Erdös, W. Feller, and H. Pollard. A property of power series with positive coefficients. *Bull. Amer. Math. Soc.*, 55(2):201 – 204, 1949.

[Fel67] W. Feller. *An introduction to probability theory and its applications, Vol. I.* John Wiley and Sons, New York, 3th edition, 1967.

[Fer73] T. S. Ferguson. A Bayesian analysis of some nonparametric problems. *Ann. Statist.*, 1(2):209–230, 1973.

[FGW12] S. Favaro, A. Guglielmi, and S. G. Walker. A class of measure-valued Markov chains and Bayesian nonparametrics. *Bernoulli*, 18(3):1002–1030, 2012.

[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.

[FM03a] G. Fort and E. Moulines. Convergence of the Monte Carlo expectation maximization for curved exponential families. *Ann. Statist.*, 31(4):1220–1259, 2003.

[FM03b] G. Fort and E. Moulines. Polynomial ergodicity of Markov transition kernels. *Stochastic Process. Appl.*, 103(1):57–99, 2003.

[Fos53] F. G. Foster. On the stochastic matrices associated with certain queuing processes. *Ann. Math. Statistics*, 24:355–360, 1953.

[FT89] P. D. Feigin and R. L. Tweedie. Linear functionals and Markov chains associated with dirichlet processes. *Mathematical Proceedings of the Cambridge Philosophical Society*, 105(3):579–585, 1989.

[GHLR24] M.A. Gallegos-Herrada, D. Ledvinka, and J.S. Rosenthal. Equivalences of geometric ergodicity of Markov chains. *J Theor. Probab.*, 37:1230–1256, 2024.

[GI24] P. W. Glynn and A. Infanger. Solution representations for Poisson's equation, martingale structure, and the Markov chain central limit theorem. *Stoch. Syst.*, 14(1):47–68, 2024.

[GLL25] P. W. Glynn, N. Lin, and Y. Liu. Computable bounds on the solution to Poisson's equation for general Harris chains. arXiv:2504.01703, 2025.

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

[Gra14] C. Graham. *Markov Chains. Analytic and Monte Carlo Computations*. Hoboken, NJ: John Wiley & Sons; Paris: Dunod, 2014.

[GS87a] D. Gibson and E. Seneta. Augmented truncations of infinite stochastic matrices. *J. Appl. Probab.*, 24(3):600–608, 1987.

[GS87b] D. Gibson and E. Seneta. Monotone infinite stochastic matrices and their augmented truncations. *Stochastic Process. Appl.*, 24(2):287–292, 1987.

[GT01] A. Guglielmi and R. L. Tweedie. Markov chain Monte Carlo estimation of the law of the mean of a Dirichlet process. *Bernoulli*, 7(4):573–592, 2001.

[GvdV17] S. Ghosal and A. van der Vaart. Dirichlet processes. In *Fundamentals of Nonparametric Bayesian Inference*, pages 59–101. Cambridge University Press, 2017.

[GW06] F. Gong and L. Wu. Spectral gap of positive operators and applications. *J. Math. Pures Appl. (9)*, 85(2):151–191, 2006.

[GZ08] P. W. Glynn and A. Zeevi. Bounding stationary expectations of Markov processes. In *Markov processes and related topics: a Festschrift for Thomas G. Kurtz*, volume 4 of *Inst. Math. Stat. (IMS) Collect.*, pages 195–214. Inst. Math. Statist., 2008.

[Hai06] M. Hairer. Ergodic properties of Markov processes. Online version,3 july 2025, 2006.

[Hai21] M. Hairer. Convergence of Markov processes. Online version,1 july 2025, 2021.

[Har56] T. E. Harris. The existence of stationary measures for certain Markov processes. In *Proceedings of the Third Berkeley Symposium on Mathematical Statistics and Probability, 1954–1955, vol. II*, pages 113–124. Univ. California Press, Berkeley-Los Angeles, Calif., 1956.

[Hen06] H. Hennion. Quasi-compactness and absolutely continuous kernels, applications to Markov chains. arxiv-0606680, June 2006.

[Hey91] D. P. Heyman. Approximating the stationary distribution of an infinite stochastic matrix. *J. Appl. Probab.*, 28(1):96–103, 1991.

[Hin00] M. Hino. Exponential decay of positivity preserving semigroups on L^p. *Osaka J. Math.*, 37(3):603–624, 2000.

[Hin02] M. Hino. Correction to: "Exponential decay of positivity preserving semigroups on L^p" [Osaka J. Math. **37** (2000), no. 3, 603–624; MR1789439 (2001k:60109)]. *Osaka J. Math.*, 39(3):771, 2002.

[HL99] F. Hirsch and G. Lacombe. *Elements of Functional Analysis*. Springer, 1999.

[HL14a] L. Hervé and J. Ledoux. Approximating Markov chains and V-geometric ergodicity via weak perturbation theory. *Stochastic Process. Appl.*, 124(1):613–638, 2014.

[HL14b] L. Hervé and J. Ledoux. Spectral analysis of Markov kernels and aplication to the convergence rate of discrete random walks. *Adv. Appl. Probab.*, 46(4):1036–1058, 2014.

[HL20] L. Hervé and J. Ledoux. V-geometrical ergodicity of Markov kernels via finite-rank approximations. *Electron. Commun. Probab.*, 25(23):1–12, March 2020.

[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.

[HL25a] L. Hervé and J. Ledoux. Computable bounds for solutions to Poisson's equation and perturbation of Markov kernels. *Bernoulli*, 31(3):2302–2322, 2025.

[HL25b] L. Hervé and J. Ledoux. Polynomial convergence rates for Markov kernels under nested modulated drift conditions. *J. Theor. Probab.*, 38, 2025. hal-04351652.

[HLL95] O. Hernández-Lerma and J.-B. Lasserre. Invariant probabilities for Feller-Markov chains. *J. Appl. Math. Stochastic Anal.*, 8(4):341–345, 1995.

[HLL96] O. Hernández-Lerma and J.-B. Lasserre. Existence of bounded invariant probability densities for Markov chains. *Statist. Probab. Lett.*, 28(4):359–366, 1996.

[HLL03] O. Hernandez-Lerma and J.B. Lasserre. *Markov Chains and Invariant Probabilities*. Progress in Mathematics. Birkhäuser Basel, 2003.

[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.

[Hor04] K. Horbacz. Random dynamical systems with jumps. *J. Appl. Probab.*, 41(3):890–910, 2004.

[HR07] O. Häggström and J. S. Rosenthal. On variance conditions for Markov chain CLTs. *Electron. Comm. Probab.*, 12:454–464, 2007.

[HS92] A. Hordijk and F. Spieksma. On ergodicity and recurrence properties of a Markov chain with an application to an open Jackson network. *Adv. Appl. Probab.*, 24(2):343–376, 1992.

[IG22] A. Infanger and P. G. Glynn. On convergence of a truncation scheme for approximating stationary distributions of continuous state space Markov chains and processes. arxiv 2206.11738, 2022.

[IGL22] A. Infanger, P. G. Glynn, and Y Liu. On convergence of general truncation-augmentation schemes for approximating stationary distributions of Markov chains. arxiv 2203.15167, 2022.

[Jer16] D. Jerison. *The Drift and Minorization Method for Reversible Markov Chains*. PhD thesis, Stanford University, 2016.

[JH00] S. F. Jarner and E. Hansen. Geometric ergodicity of Metropolis algorithms. *Stochastic Process. Appl.*, 85(2):341–361, 2000.

[JJ67] N. Jain and B. Jamison. Contributions to Doeblin's theory of Markov processes. *Z. Wahrsch. Verw. Gebiete*, 8:19–40, 1967.

[JMMD15] J. E. Johndrow, J. C. Mattingly, S. Mukherjee, and D. Dunson. Optimal approximating Markov chains for Bayesian inference. arXiv:1508.03387v3, 2015.

[Jon04] G. L. Jones. On the Markov chain central limit theorem. *Probability 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.

[JR07] S. F. Jarner and G. O. Roberts. Convergence of heavy-tailed Monte Carlo Markov chain algorithms. *Scand. J. Statist.*, 34(4):781–815, 2007.

[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.

[Ken51] D. G. Kendall. On non-dissipative Markoff chains with an enumerable infinity of states. *Proc. Cambridge Philos. Soc.*, 47:633–634, 1951.

[Ken59] D. G. Kendall. Unitary dilations of Markov transition operators, and the corresponding integral representations for transition-probability matrices. In *Probability and statistics: The Harald Cramér volume (edited by Ulf Grenander)*, pages 139–161. Almqvist & Wiksell, Stockholm, 1959.

[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.

[KM03] I. Kontoyiannis and S. P. Meyn. Spectral theory and limit theorems for geometrically ergodic Markov processes. *Ann. Appl. Probab.*, 13(1):304–362, 2003.

[KM05] I. Kontoyiannis and S. P. Meyn. Large deviations asymptotics and the spectral theory of multiplicatively regular Markov processes. *Electron. J. Probab.*, 10:61–123, 2005.

[KM12] I. Kontoyiannis and S. P. Meyn. Geometric ergodicity and the spectral gap of non-reversible Markov chains. *Probab. Theory Related Fields*, 154(1-2):327–339, 2012.

[KR90] V. V. Kalashnikov and S. T. Rachev. *Mathematical Methods for Construction of Queueing Models*. Wadsworth and Brooks Cole, 1990.

[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.

[Liu10] Y. Liu. Augmented truncation approximations of discrete-time Markov chains. *Oper. Res. Lett.*, 38(3):218–222, 2010.

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

[LP09] A. Lijoi and I. Prünster. Distributional properties of means of random probability measures. *Statistics Surveys*, 3:47 – 95, 2009.

[LT96] R. B. Lund and R. L. Tweedie. Geometric convergence rates for stochastically ordered Markov chains. *Math. Oper. Res.*, 21(1):182–194, 1996.

[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.

[Mar06] A. A. Markov. Extension of the law of large numbers to dependent quantities (in russian). *Izv. Fiz.-Matem. Obsch. Kazan Univ. (2nd Ser.)*, 15:135–156, 1906.

[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.

[Mas16] H. Masuyama. Error bounds for augmented truncations of discrete-time block-monotone Markov chains under subgeometric drift conditions. *SIAM J. Matrix Anal. Appl.*, 37(3):877–910, 2016.

[Mau57] J. G. Mauldon. On non-dissipative Markov chains. *Proc. Cambridge Philos. Soc.*, 53:825–835, 1957.

[Mey08] S. P. Meyn. *Control Techniques for Complex Networks*. Cambridge University Press, Cambridge, 2008.

[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.

[MS04] J. Myjak and T. Szarek. On the existence of an invariant measure for Markov-Feller operators. *J. Math. Anal. Appl.*, 294(1):215–222, 2004.

[MS10] N. Madras and D. Sezer. Quantitative bounds for Markov chain convergence: Wasserstein and total variation distances. *Bernoulli*, 16(3):882–908, 2010.

[MT92] S. P. Meyn and R. L. Tweedie. Stability of Markovian processes. I. Criteria for discrete-time chains. *Adv. Appl. Probab.*, 24(3):542–574, 1992.

[MT94] S. P. Meyn and R. L. Tweedie. Computable bounds for geometric convergence rates of Markov chains. *Ann. Probab.*, 4:981–1011, 1994.

[MT96] K. L. Mengersen and R. L. Tweedie. Rates of convergence of the Hastings and Metropolis algorithms. *Ann. Statist.*, 24(1):101–121, 1996.

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

[Mur23] K. P. Murphy. *Probabilistic Machine Learning: Advanced Topics (supplementary material)*. MIT Press, 2023.

[Nev72] J. Neveu. Potentiel Markovien récurrent des chaînes de Harris. *Ann. Inst. Fourier (Grenoble)*, 22(2):85–130, 1972.

[Nor97] J. R. Norris. *Markov Chains*. Cambridge University Press, 1997.

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

[NT78] E. Nummelin and R. L. Tweedie. Geometric ergodicity and R-positivity for general Markov chains. *Ann. Probability*, 6(3):404–420, 1978.

[NT82] E. Nummelin and P. Tuominen. Geometric ergodicity of Harris recurrent Markov chains with applications to renewal theory. *Stochastic Process. Appl.*, 12(2):187–202, 1982.

[NT83] E. Nummelin and P. Tuominen. The rate of convergence in Orey's theorem for Harris recurrent Markov chains with applications to renewal theory. *Stochastic Process. Appl.*, 15(3):295–311, 1983.

[Num78] E. Nummelin. A splitting technique for Harris recurrent Markov chains. *Z. Wahrsch. Verw. Gebiete*, 43:309–318, 1978.

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

[Num91] E. Nummelin. On the Poisson equation in the potential theory of a single kernel. *Math. Scand.*, 68(1):59–82, 1991.

[Num02] E. Nummelin. MC's for MCMC'ists. *Int. Statist. Rev.*, 70(2):215–240, 2002.

[Pak69] A. G. Pakes. Some conditions for ergodicity and recurrence of Markov chains. *Operations Res.*, 17:1058–1061, 1969.

[Pit74] J. W. Pitman. Uniform rates of convergence for Markov chain transition probabilities. *Z. Wahrsch. Verw. Gebiete*, 29:193–227, 1974.

[QH21] Q. Qin and J. P. Hobert. On the limitations of single-step drift and minorization in Markov chain convergence analysis. *Ann. Appl. Probab.*, 31(4):1633–1659, 2021.

[QH22] Q. Qin and J. P. Hobert. Wasserstein-based methods for convergence complexity analysis of MCMC with applications. *Ann. Appl. Probab.*, 32(1):124–166, 2022.

[Qin24] Q. Qin. Convergence bounds for Monte Carlo Markov Chains. To appear in Handbook of Markov Chain Monte Carlo (2nd edition), arXiv:2409.14656, 2024.

[RC04] C. P. Robert and P. Casella. *Monte Carlo Statistical methods*. Springer, 2004.

[Rev75] D. Revuz. *Markov Chains*. North-Holland Publishing Co., Amsterdam-Oxford; American Elsevier Publishing Co., Inc., New York, 1975. North-Holland Mathematical Library, Vol. 11.

[Rev84] D. Revuz. *Markov chains*, volume 11 of *North-Holland Mathematical Library*. North-Holland Publishing Co., Amsterdam, second edition, 1984.

[Rio00] Emmanuel Rio. *Théorie asymptotique des processus aléatoires faiblement dépendants*, volume 31 of *Mathématiques & Applications (Berlin) [Mathematics & Applications]*. Springer-Verlag, Berlin, 2000.

[Ros71] M. Rosenblatt. *Markov processes. Structure and Asymptotic Behavior*. Springer-Verlag, New-York, 1971.

[Ros95] J. S. Rosenthal. Minorization conditions and convergence rates for Markov chain Monte Carlo. *J. Amer. Statist. Assoc.*, 90(430):558–566, 1995.

[Ros02] J. S. Rosenthal. Quantitative convergence rates of Markov chains: a simple account. *Electron. Comm. Probab.*, 7:123–128, 2002.

[RR96] G. I. Roberts and R. L. Rweedie. Geometric convergence and central limit thoerems for multidimensional Hastings and Metropolis algorithms. *Biometrika*, 83(1):95–110, 1996.

[RR97] G. O. Roberts and J. S. Rosenthal. Geometric ergodicity and hybrid Markov chains. *Electron. Commun. Probab.*, 2:13–25, 1997.

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

[RR06] G. O. Roberts and J. S. Rosenthal. Harris recurrence of Metropolis-within-Gibbs and trans-dimensional Markov chains. *Ann. Appl. Probab.*, 16(4):2123–2139, 2006.

[RR11] G. O. Roberts and J. S. Rosenthal. Quantitative non-geometric convergence bounds for independent samplers. *Methodol. Comput. Appl. Probab.*, 13:391–403, 2011.

[RRS98] G. O. Roberts, J. S. Rosenthal, and P. O. Schwartz. Convergence properties of perturbed Markov chains. *J. Appl. Probab.*, 35(1):1–11, 1998.

[RS80] M. Reed and B. Simon. *Methods of modern mathematical physics. I.* Academic Press, Inc. [Harcourt Brace Jovanovich, Publishers], New York, second edition, 1980. Functional analysis.

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

[RSQ24] D. Rudolf, A. Smith, and M Quiroz. Perturbations of Markov chains. To appear in Handbook of Markov Chain Monte Carlo (2nd edition), arXiv:2404.10251, 2024.

[RT99] G. O. Roberts and R. L. Tweedie. Bounds on regeneration times and convergence rates for Markov chains. *Stochastic Process. Appl.*, 80(2):211–229, 1999.

[RT00] G. O. Roberts and R. L. Tweedie. Rates of convergence of stochastically monotone and continuous time Markov models. *J. Appl. Probab.*, 37(2):359–373, 2000.

[RT01] G. O. Roberts and R. L. Tweedie. Geometric L^2 and L^1 convergence are equivalent for reversible Markov chains. *J. Appl. Probab.*, 38A:37–41, 2001.

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

[Sen81] E. Seneta. *Nonnegative matrices and Markov chains*. Springer-Verlag, New York, 1981.

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

[Sen06] E. Seneta. *Nonnegative matrices and Markov chains*. Springer-Verlag, New York, 2006.

[Sim95] F. Simonot. Sur l'approximation de la distribution stationnaire d'une chaîne de Markov stochastiquement monotone. *Stochastic Process. Appl.*, 56(1):133–149, 1995.

[Spi91] F. Spieksma. *Geometrically Ergodic Markov Chains and the Optimal Control of Queues*. PhD thesis, University of Leiden, 1991.

[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.

[Sza03] T. Szarek. Invariant measures for Markov operators with application to function systems. *Studia Math.*, 154(3):207–222, 2003.

[Teh17] Y. W. Teh. Dirichlet process. In C. Sammut and G. I. Webb, editors, *Encyclopedia of Machine Learning and Data Mining*, pages 361–370. Springer US, Boston, MA, 2017.

[Tie94] L. Tierney. Markov chains for exploring posterior distributions. *Ann. Statist.*, 22(4):1701–1762, 1994.

[TM22] A. Taghvaei and P. G. Mehta. On the Lyapunov Foster criterion and Poincaré inequality for reversible Markov chains. *IEEE Trans. Automatic Control*, 67(5):2605–2609, 2022.

[Tor23] G. L. Torrisi. Quantitative multidimensional central limit theorems for means of the Dirichlet-Ferguson measure. *ALEA*, 20:825–860, 2023.

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

[Twe74a] R. L. Tweedie. R-theory for Markov chains on a general state space. I. Solidarity properties and R-recurrent chains. *Ann. Probability*, 2:840–864, 1974.

[Twe74b] R. L. Tweedie. R-theory for Markov chains on a general state space. II. r-subinvariant measures for r-transient chains. *Ann. Probability*, 2:865–878, 1974.

[Twe75] R. L. Tweedie. Sufficient conditions for ergodicity and recurrence of Markov chains on a general state space. *Stochastic Process. Appl.*, 3(4):385–403, 1975.

[Twe98] R. L. Tweedie. Truncation approximations of invariant measures for Markov chains. *J. Appl. Probab.*, 35(3):517–536, 1998.

[VJ62] D. Vere-Jones. Geometric ergodicity in denumerable Markov chains. *Quart. J. Math. Oxford Ser. (2)*, 13:7–28, 1962.

[Wan22] G. Wang. Exact convergence analysis of the independent Metropolis-Hastings algorithms. *Bernoulli*, 28(3):2012–2033, 2022.

[Wol80] D. Wolf. Approximation of the invariant probability measure of an infinite stochastic matrix. *Adv. Appl. Probab.*, 12(3):710–726, 1980.

[Wu04] L. Wu. Essential spectral radius for Markov semigroups. I. Discrete time case. *Probab. Theory Related Fields*, 128(2):255–321, 2004.

[XZZ18] L. Xu, S. Zhang, and R. Zhang. A note on Foster-Lyapunov drift condition for recurrence of Markov chains on general state spaces. *J. Theor.. Probab.*, 31(4):1923–1928, 2018.

[YK41] K. Yosida and S. Kakutani. Operator-theoretical treatment of Markoff's process and mean ergodic theorem. *Ann. Math.*, 42(1):188–228, 1941.

[Yos95] K. Yosida. *Functional analysis*. Classics in Mathematics. Springer-Verlag, Berlin, 1995. Reprint of the sixth (1980) edition.

[YR23] J. Yang and J. S. Rosenthal. Complexity results for MCMC derived from quantitative bounds. *Ann. Appl. Probab.*, 33(2):1259–1300, 2023.