Quasi-compactness for dominated kernels with application to quasi-stationary distribution theory

D. Villemonais[1,2]

October 23, 2025

Abstract

We establish a domination principle for positive operators, yielding a upper bound on the essential spectral radius and a practical quasi-compactness criterion on weighted supremum spaces. We then apply these results to absorbed Markov processes and show that quasi-compactness of the transition kernel ensures existence and convergence to quasi-stationary distributions in broadly reducible settings, without regularity requirements. In continuous time, we show that measurability plus quasi-compactness at a single time propagates to all times, rules out periodic behavior, and yields convergence to quasi-stationary distributions. Two illustrative cases demonstrate the scope and simplicity of the criteria.

1 Introduction

The fact that compactness/quasi-compactness of a non-negative operator P on a Banach lattice B can be obtained by domination property has been initiated by the remarkable results obtained in [18, Theorem 4.5]: it is proved therein that if B is norm continuous and if P, Q are two non-negative operators with $0 \le P \le Q$ and Q compact, then P is itself a compact operator (we also refer the interested reader to the contemporary article [37], where the author prove by simpler means an important special case). The extension of this domination condition for quasi-compactness has been studied from then, but appears to be false in general: in [15], the authors provide an interesting counter example on some L^2 space, where $0 \le P \le Q$ with Q a quasi-compact operator

[1]IRMA, Université de Strasbourg, France
[2]Institut Universitaire de France
denis.villemonais@unistra.fr

and P not quasi-compact. Quasi-compactness of non-negative operators have nontheless been obtained under particular conditions on B or Q. More precisely, in [6], the author shows that if B is a pre-dual function space and Q is quasi-compact with essential spectral radius 0, then this is also the case for P; in [47], the author shows that if Q satisfies a suitable regularity property and $B = \mathscr{B}_b(E)$, then the essential spectral radius of P is upper bounded by that of Q; several properties entailed by similar domination assumptions are also provided in [34, 45].

Our first main result is that the quasi-compactness of P can be derived from domination properties by positive operators. Assuming that $0 \leq P \leq K + S$, where K is a non-negative compact operator and S a non-negative operator, we show that $r_{ess}(P) \leq r(S)$, where $r_{ess}(P)$ denotes the essential spectral radius of P, as soon as the positive part $(P-S)_+$ of $P-S$ is well defined. Motivated by applications to probability theory, we focus on the setting where B is a weighted function space over a measurable space $(E,\mathscr{E})$ which is countably generated, and where P and Q are kernels on E. In this case, we show that $0 \leq P \leq Q$ does imply that $r_{ess}(P) \leq r_{ess}(Q)$.

Our second main result concerns the problem of existence and convergence to quasi-stationary distributions for discrete time absorbed Markov processes. Let $(X_n)_{n\in\mathbb{N}}$ be a discrete time Markov evolving in a measurable space $(E \cup \{\partial\}, \mathscr{E}^\partial)$, where $\partial \notin E$ is absorbing: for all $n \geq \tau_\partial = \inf\{n \geq 0,\ X_n = \partial\}$, $X_n = \partial$ almost surely. We assume that $\{\partial\} \in \mathscr{E}^\partial$, that the σ-algebra $\mathscr{E}^\partial$ is countably generated and we set $\mathscr{E} = \mathscr{E}^\partial \cap E$. As usual, for any probability measure μ on E, we denote by $\mathbb{P}_\mu$ the law of X with initial distribution μ, and by $\mathbb{E}_\mu$ its associated expectation, and, for all $x \in E \cup \partial$, we set $\mathbb{P}_x := \mathbb{P}_{\delta_x}$ and $\mathbb{E}_x := \mathbb{E}_{\delta_x}$. We are interested in the existence of a probability measure ν_{QS} on E such that

$$\mathbb{P}_{\nu_{QS}}(X_n \in A \mid n < \tau_\partial) = \nu_{QS}(A), \quad \forall n \geq 0 \text{ and } A \in \mathscr{E}.$$

Such a measure is referred to as a *quasi-stationary distribution*. One easily checks (see e.g. [14, 35]) that, if there exists a probability measure μ on E such that

$$\lim_{n\to+\infty} \mathbb{P}_\mu(X_n \in A \mid n < \tau_\partial) = \nu_{QS}(A), \quad \forall A \in \mathscr{E},$$

then ν_{QS} is a quasi-stationary distribution. Among the several approaches that have been devised to prove the existence of quasi-stationary distributions, the study of the spectral properties have been particularly successful for the study of diffusion processes [33, 44, 8, 9, 32, 31, 23, 24] and for birth and death processes [46], often through arguments relying on the auto-adjoint properties of the infinitesimal generator or of an auxiliary generator (see also [11] for a probabilistic approach to the study of these models). Recently, some authors focused specifically on the quasi-compactness of the transition kernel operator, using regularity assumptions such as strong Feller regularity, see e.g. [21, 28, 7, 5, 22], see also the recent [42] with in-depth considerations and practical criteria for Krein-Rutman type results. In this article, we show that the quasi-compactness of the transfer kernel entails strong properties for quasi-stationary distribution theory, extending results from [48, 39, 26] to the non-conservative case. In particular, our condition does not require the semi-group of the process under study to be strongly Feller or even Feller, and can instead leverage the quasi-compactness by domination criterion that we first obtained. We show how this can be done through the use of Lyapunov type criteria.

Our third main result concerns the problem of existence and convergence to quasi-stationary distributions for discrete time absorbed Markov processes. We prove that measurability (in time) of the semi-group and quasi-compactness of the transition kernel at one positive time is enough to deduce the quasi-compactness of the transition kernel at all times t, the aperiodicity of the process, and convergence of the process to quasi-stationary distribution as $t \in [0,+\infty)$ tends to infinity. As far as we know, this result is also new for conservative Markov processes and the study of their stationary distributions.

In Section 2, we state and prove our quasi-compactness by domination criterion. In Section 3, we focus on the study of quasi-stationarity for discrete time processes through quasi-compactness of the transition kernel (Section 3.1) and through Lyapunov type criteria (Section 3.2), and for continous time processes (Section 3.3). We conclude the paper with two applications in discrete (Section 4.1 and continuous time (Section 4.2).

2 Quasi-compactness by domination arguments

Let $(B, \|\cdot\|)$ be a Banach lattice (we refer the reader to [2, Chapter 4] for references and first properties of Banach Lattices), with order $\geq$. A linear operator Q on B is called *positive* if $Qf \geq 0$ for all element $f \in B$ such that $f \geq 0$. Given two linear operators P and Q on B, we say that $0 \leq P \leq Q$ if $0 \leq Pf \leq Qf$ for all element $f \in B$ such that $f \geq 0$. Given a bounded operator Q on B, we denote by $|||Q|||$ the operator norm of Q and by $r(Q)$ its spectral radius (in particular, positive linear operators on a Banach lattice are bounded, according to [2, Theorem 12.3]).

The *essential spectral radius* of an operator Q on B is defined by

$$r_{\text{ess}}(Q) := \lim_{n\to\infty} \inf\left\{|||Q^n - K|||^{1/n} : K \in \mathcal{K}(B)\right\},$$

where $\mathcal{K}(B)$ denotes the set of compact operators on B (we refer the reader to [2, Chapter 5] for definitions and properties of compact operators on Banach lattices). This definition captures the asymptotic behavior of the "non-compact part" of Q under iteration and provides a quantitative measure of how closely the powers of Q can be approximated by compact (or finite rank) operators. The essential spectral radius can also be defined as the common radius of the different essential spectra of Q, as detailed in [20, Section 1.4].

An operator Q on B is said to be *quasi-compact* if $r_{\text{ess}}(Q) < r(Q)$, where $r(Q)$ is the spectral radius of Q.

Remark 1. Following another standard definition, $r_{\text{ess}}(Q)$ can be defined as the infimum of those $\rho \geq 0$ such that B admits a decomposition

$$B = F_\rho \oplus H_\rho,$$

into closed, Q-invariant subspaces with $\dim F_\rho < \infty$, all eigenvalues of $Q_{|F_\rho}$ having modulus $\geq \rho$, and $r(Q_{|H_\rho}) < \rho$ (see e.g. [27]). The bounded linear operator Q is thus quasi-compact if there exists a decomposition

$$B = F \oplus H,$$

into closed, Q-invariant subspaces, where $\dim F < \infty$, all eigenvalues of $Q_{|F}$ have modulus equal to the spectral radius $r(Q)$, and the restriction $Q_{|H}$ satisfies $r(Q_{|H}) < r(Q)$. The link between this alternative definition and the definition we give in terms of compact approximation is a by-product of the properties established in the first chapter of [41].

Finally, given a bounded linear operator Q on B, we say that the *positive part* of Q, noted Q_+, is well defined if, for all $f \geq 0$,

$$Q_+ f := \sup\{Qg,\ g \in B, 0 \leq g \leq f\}$$

defines a positive linear operator on B. In particular, Q_+ is well defined if B is Dedekind complete (see e.g. [2] page 13), which is the case for $L^p(\mu)$ spaces, $1 \leq p \leq \infty$, on a countably generated measured space $(E, \mathscr{E}, \mu)$ with μ σ-finite (see e.g. Example 2.52 p.48 in [4])).

Theorem 1. *Let P be a positive operator on B and assume that there exists two positive operators K and S such that K is compact, $(P - S)_+$ is well defined and $0 \leq P \leq K + S$. Then $r_{ess}(P) \leq r(S)$.*

Theorem 1 immediately entails the following corollary.

Corollary 2. *If, in addition to the assumptions of Theorem 1, we have $r(S) < r(P)$, then P is quasi-compact.*

Proof of Theorem 1. We deduce from $0 \le P \le S + K$ that

$$0 \le (P - S)_+ \le K,$$

where $(P - S)_+$ is a positive kernel defined by

$$\delta_x (P - S)_+ = (\delta_x P - \delta_x S)_+, \ \forall x \in E.$$

We also have

$$P = (P - S)_+ + S \wedge P,$$

where $0 \le S \wedge P \le S$, so that $\|S \wedge P\| \le \|S\|$.

We deduce that, for all $n \ge 0$,

$$P^n = S^n + \sum_{i=1}^{n} S^{i-1}(P - S)_+ S^{n-i} + \sum_{1 \le i < j \le n} S^{i-1}(P - S)_+ S^{j-i}(P - S)_+ S^{n-i-j} + K_n,$$

where K_n is a finite sum of products of operators where $(P - S)_+$ appears at least three times. According to the following lemma, we deduce that K_n is compact.

Lemma 3. *Let $S_1, S_2, S_3, S_4, Q_1, Q_2, Q_3$ be non-negative bounded operators and K_1, K_2, K_3 be non-negative compact operators such that $0 \le Q_i \le K_i$ for all $i \in \{1, 2, 3\}$. Then the operator*

$$S_1 Q_1 S_2 Q_2 S_3 Q_3 S_4$$

is compact.

Proof of Lemma 3. We observe that $0 \le S_1 Q_1 S_2 \le S_1 K_1 S_2$, $0 \le Q_2 S_3 \le K_2 S_3$ and $0 \le Q_3 S_4 \le K_3 S_4$, where the operators $S_1 K_1 S_2$, $K_2 S_3$ and $K_3 S_4$ are compact. Hence, according to Aliprantis–Burkinshaw Theorem 5.14 [Positive Operators, Aliprantis–Burkinshaw], the product of the three operators is a compact operator. □

Considering the first three terms in the above decomposition, we get that

$$\left\| P^n - K_n \right\| \le \|S\|^n + n \|(P - S)_+\| \, \|S\|^{n-1} + \frac{n(n-1)}{2} \|(P - S)_+\|^2 \, \|S\|^{n-2}.$$

We deduce that

$$r_{ess}(P) = \left(r_{ess}(P^n) \right)^{1/n} \le \|S\|^{1-2/n} \left(\|S\|^2 + \|(P - S)_+\| \, \|S\| + \|(P - S)_+\|^2 \right)^{1/n} \xrightarrow[n \to +\infty]{} \|S\|.$$

We thus proved that $r_{ess}(P) \le \|S\|$.

We conclude the proof by proving that this inequality implies that $r_{ess}(P) \le r(S)$. Indeed, we get from $0 \le P \le S + K$ that, for all $n \ge 1$,

$$0 \le P^n \le S^n + K',$$

with K' given by the sum of products of bounded operators where the compact operator K appears at least once. This entails that K' is a compact operator and hence, according to the first part of the proof, that $r_{ess}(P^n) \le \|S^n\|$ and hence that

$$r_{ess}(P) = r_{ess}(P^n)^{1/n} \le \|S^n\|^{1/n} \xrightarrow[n\to+\infty]{} r(S),$$

where we used Gelfand formula for the last inequality. This concludes the proof of Theorem 1 □

We specialize now our results on the case of weighted spaces of measurable functions. More precisely, considering a countably generated measurable space $(E, \mathscr{E})$ and a measurable function $V : E \to (0, +\infty)$, we define the function space

$$L^\infty(V) := \left\{ f : E \to \mathbb{C},\, \exists C > 0,\text{ such that } |f(x)| \le C\, V(x),\ \forall x \in E \right\},$$

endowed with the weighted supremum norm $\|f\|_V = \sup_{x\in E} \frac{|f(x)|}{V(x)}$. We say that a functional $R : E \times \mathscr{E} \to \mathbb{R}$ is a kernel if, for all $x \in E$, $R(x, \cdot)$ is a σ-finite measure, and, for all $A \in \mathscr{E}$, $R(\cdot, A)$ is measurable. A kernel such that $\|R(\cdot, V)\|_V < \infty$ is called a kernel from E to $L^\infty(V)$ and it defines a linear operator on $L^\infty(V)$. In the following result, $r_{ess}(R)$ is then the essential spectral radius of the induced linear operator.

We obtain the following corollary to Theorem 1, which extends [47, Corollary 3.7], by removing the regularity condition on P.

Corollary 4. *Let P and Q be kernels from E to $L^\infty(V)$ such that $0 \le P \le Q$. Then $r_{ess}(P) \le r_{ess}(Q)$.*

Proof of Corollary 4. The proof relies on Theorem 1 and on the following equivalent definition of the positive part of a kernel: given a kernel R from E to $L^\infty(V)$, we have

$$R_+ f(x) := (\delta_x R)_+ f, \tag{1}$$

where $(\delta_x R)_+$ is the positive part of the measure $\delta_x R$. The fact that R_+ is itself a kernel is a consequence of section 2 in [19]. In addition, we have $|||R_+||| \le |||R|||$, since, for all non-negative $f \in L^\infty(V)$ and all $x \in E$,

$$|R_+ f(x)| \le R_+|f|(x) = R|f\mathbf{1}_{A_x}|(x) \le |||R|||\, \|f\mathbf{1}_{A_x}\|_V \le |||R|||\, \|f\|_V,$$

where A_x is the positive set in the Hahn decomposition of the measure $\delta_x R$. Similarly, $r(R_+) \leq r(R)$.

By definition of $r_{ess}(Q)$, for any ε, we can find a finite rank operator K and a (possibly signed) bounded operator S such that $Q = K + S$ and $|||S^n|||^{1/n} \leq r_{ess}(Q) + \varepsilon$ for some $n \geq 1$. Since we also have $0 \leq P^n \leq Q^n = K' + S^n$, with K' a finite rank operator, we can assume without loss of generality that $n = 1$.

Since K is a finite rank operator, its image is a finite subspace of $L^\infty(V)$ and hence there exist $h_1, \dots, h_d$ bounded continuous functionals from $L^\infty(V)$ to $\mathbb{R}$, and $a_1, \dots, a_d \in L^\infty(V)$ such that, for all $x \in E$ and all non-negative $f \in L^\infty(V)$,

$$Kf(x) = \sum_{i=1}^{d} h_i(f) a_i(x) \leq \bar{K} f(x) := \sum_{i=1}^{d} h_{i,+}(f) |a_i|(x)$$

where $h_{i,+}(f) = \sup\{h_i(g), 0 \leq g \leq f\}$. We deduce that $0 \leq P \leq \bar{K} + S_+$, where $\bar{K}$ is a non-negative finite rank operator and $|||S_+||| \leq |||S||| \leq r_{ess}(Q) + \varepsilon$.

Using Theorem 1, we deduce that $r_{ess}(P) \leq r_{ess}(Q) + \varepsilon$, and hence that $r_{ess}(P) \leq r_{ess}(Q)$. □

Remark 2. In L^p spaces, Corollary 4 does not hold true (see the counter example in [15]), precisely because $|||S_+||| \leq |||S|||$ does not hold true in general. We refer the reader to [1] for several examples of disaccording compactness properties of S and S_+ on L^p spaces.

3 Application to quasi-stationary distributions theory

3.1 Quasi-compactness criterion for quasi-stationary distributions in discrete time

In this section, we consider the particular case where P is the submarkov transition kernel of an absorbed Markov process X. More precisely, we set, for all non-negative measurable function $f : E \to \mathbb{R}_+$,

$$Pf(x) := \mathbb{E}_x(f(X_n) \mathbf{1}_{n < \tau_\partial}),$$

with the same assumptions as in the introduction. In particular, $\mathscr{E}$ is assumed to be countably generated throughout the section (see Remark 4 to extend the results outside of this setting). We also assume that we are given a function $V : E \to (0, +\infty)$.

Theorem 5. *Assume that P is a quasi-compact operator on $L^\infty(V)$. Then there exists an integer $d \geq 1$, an integer valued bounded function $j : E \to \mathbb{R}_+$, a finite set I and some probability measures ν_i on E such that $\nu_i(V) < +\infty$ and non-identically zero non-negative functions $\eta_i \in L^\infty(V)$ for each $i \in I$, such that, for all $f \in L^\infty(V)$, all $n \geq 1$, all $k \in \{0,1,\dots,d-1\}$ and all $x \in E$,*

$$\left| \frac{r(P)^{-dn-k}}{(dn+k)^{j(x)}} \mathbb{E}_x(f(X_{dn+k})\mathbf{1}_{nd+k<\tau_\partial}) - \sum_{i\in I} \eta_{i,k}(x)\nu_{i,k}(f) \right| \leq \alpha_{nd+k} V(x)\|f\|_V, \quad (2)$$

where $\nu_{i,k} = \nu_i P^k$, $\eta_{i,k} = r(P)^{-k}\eta_i$ and α_n goes to 0 when $n \to +\infty$.

In addition, there exist disjoint subsets $E_i \subset E$, $i \in I$, such that

1. *$F_i := E_i \cup \{x \in E,\ \mathbb{P}_x(\exists n \geq 0,\ X_{nd} \in E_i) = 0\}$ is a closed subset for $X_{d\cdot}$,*

2. *ν_i is supported by F_i and is a quasi-stationary distribution for $X_{d\cdot}$, with $\nu_i P^d = r(P)^d \nu_i$,*

3. *ν_i restricted to E_i is a quasi-stationary distribution for $X_{d\cdot}$ restricted to E_i,*

4. *on each E_i, $j = 0$, $\eta_i > 0$ and $\eta_{i'} = 0$ for all $i' \neq i$.*

Remark 3. When the process X is conservative, that is $\mathbb{P}_x(\tau_\partial = +\infty) = 1$ for all $x \in E$, it is known that the quasi-compactness of P entails the result [26] (see also the extended preprint version [25]), see also [39] and [48] for quasi-compact Markov kernels on $B_b(E)$ (in the non-conservative case, the treatment of $B = B_b(E)$ and $B = L^\infty(V)$ are equivalent using V-transforms -see Remark 5 below, but this is not the case for the conservative case, since the V transform of a Markov kernel is not, in general, a Markov kernel). When restricting to the particular conservative case, we have $j \equiv 0$ and the novelty of Theorem 5 thus entirely lies in the domination criterion.

Remark 4. In the proof of Theorem 5, we make use of [39, Proposition 3.78, Chapter 6], which requires $\mathscr{E}$ to be countably generated. Besides that, the proof requires only minor modifications to get the same result without the countable generation assumption.

Remark 5. We recall that our result also apply to non-conservative semi-group (non-necessarily sub-Markov) on $L^\infty(V)$, via the classical and straightforward use of V-transform. We refer the reader to [12] for details on this strategy.

Proof of Theorem 5. The proof relies on the decomposition of the state space into subsets $E_1, E'_1, \dots, E_p, E'_p, E'_{p+1}$ such that the semi-group associated to X is well behaved on each E_k, E'_k: each E_k can be decomposed into communication classes where it converges (at times that are multiple of some period) to a one-dimensional operator, while each E'_k is such that the process escapes exponentially fast (compared to $\rho(P)$) from E'_k. This will allow us to apply [13, Theorem 4.1].

The proof is divided in four steps. In the first step, we build E_1 and E'_1. In the second step, we build by induction $E_1, E'_1, \dots, E_p, E'_p$ and E_{p+1}. In the third step, we apply the above cited result to the process at times that are multiples of a given integer d. In the four step, we conclude by extending the convergence to any time.

Step 1. First sets in the state space decomposition. According to Corollary 4, P is a quasi-compact operator on $B = L^\infty(V)$. We deduce from [43] that there exists a non-negative non-zero function $\eta_1 \in B$ such that $P\eta_1 = r(P)\eta_1$. Let $E_1 := \{x \in E,\ \eta_1(x) > 0\}$ and define the Markov kernel T on E by

$$Tf(x) = \frac{1}{r(P)\eta_1} P(\eta_1 f)(x).$$

The Markov kernel T is the so called η_1 transform of P. We emphasize, for later use, that $E \setminus E_1$ is a closed subset for the Markov chain X (meaning that X cannot escape $E \setminus E_1$ before absorption).

We set $V_1 = \frac{V}{\eta_1}$ on E_1. One easily checks that T is a quasi-compact operator on $L^\infty(V_1)$, with underlying state space E_1. We deduce from [25, Corollary IV.3] and the adaptation of [39, Proposition 3.8] (this adaptation is suggested in [25], and one can also adapts the arguments of [48], the main addition being the non-trivial fact that the kernel T is power-bounded) that there exist a finite number $r \geq 1$ of integers d_ρ, $\rho = 1, \dots, r$, and measurable functions $U_{\rho,\delta} \in L^\infty(V_1)$, $\delta = 1, \dots, d_\rho$ such that $\sum_\rho \sum_\delta U_{\rho,\delta} = 1$ and probability measures $m_{\rho,\delta}$ carried by the pairwise disjoint sets $E_{\rho,\delta} = \{U_{\rho,\delta} = 1\}$ such that, if d denotes the least common multiple of the d_ρ, then for every k,

$$\left|\!\left|\!\left| (T)^{nd+k} - \sum_{\rho=1}^{r} \sum_{r=1}^{d_\rho} U_{\rho,\delta-k} \otimes m_{\rho,\delta} \right|\!\right|\!\right| \leq C\rho^n, \tag{3}$$

for some $\rho < 1$ and $C > 0$. In addition, the sets $E_\rho = \cup_\delta E_{\rho,\delta}$ are closed set, and the sets $E_{\rho,\delta}$ and the number d_ρ are the corresponding cyclic classes and period.

In terms of our original sub-Markov process, this entails that the sets $E_{\rho,\delta}$ are absorbing for the Markov chain $(X_{nd})_{n\geq 0}$ and that, for all $f \in L^\infty(V)$ and all $x \in E_{\rho,\delta}$, we have

$$\left| r(P)^{-nd} \mathbb{E}_x \left(f(X_{nd}) \mathbf{1}_{X_n \in E_{1,\rho,\delta}} \right) - \eta_{1,\rho,\delta}(x) \nu_{1,\rho,\delta}(f) \right| \leq C\rho^n V(x) \|f\|_V, \tag{4}$$

where $E_{1,\rho,\delta} := E_{\rho,\delta}$ $\eta_{1,\rho,\delta} := \eta_1 U_{\rho,\delta}$ is positive on $E_{1,\rho,\delta}$ and $\nu_{1,\rho,\delta} = m_{\rho,\delta}(./\eta_1)$. Finally, setting $E'_1 = E_1 \setminus \cup_\rho E_\rho$, we have $m_{\rho,\delta}(E'_1) = 0$ and hence, for all $f \in L^\infty(V)$ and all $x \in E'_1$,

$$\left| r(P)^{-nd} \mathbb{E}_x \left(f(X_{nd}) \mathbf{1}_{X_n \in E'_1} \right) \right| \leq C\rho^n V(x) \|f\|_V. \tag{5}$$

This concludes the first step.

Step 2. Inductive construction of the other subsets. As mentioned above, the subset $E \setminus E_1$ is closed. We denote by $R = P_{|E \setminus E_1}$ the kernel P restricted to $E \setminus E_1$.

If $r(R) < r(P)$, then we set $p = 1$ and $E'_{p+1} = E \setminus E_1$. The definition of the spectral radius entails that (5) holds true with E'_{p+1} instead of E'_1.

Otherwise, we have $r(R) = r(P)$ and, since it is dominated by P, the kernel R is quasi-compact according to Theorem 1. As a consequence, Step 1 applies to R and we deduce that there exists $E_2 \subset E \setminus E_1$ such that (4) and (5) holds true, with the subsets $E_{1,\rho,\delta}$ replaced by the closed subsets $E_{2,\rho,\delta} \subset E_2$ (the range of ρ and δ may of course change, while we can assume without loss of generality that d is the same), the functions $\eta_{1,\rho,\delta}$ replaced by $\eta_{2,\rho,\delta}$ (which are positive on $E_{2,\rho,\delta}$) and the probability measures $\nu_{1,\rho,\delta}$ replaced by the probability measures $\nu_{2,\rho,\delta}$ on $E_{2,\rho,\delta}$, and E'_1 replaced by $E'_2 := E_2 \setminus \cup_{\rho,\delta} E_{2,\rho,\delta}$.

Repeating this procedure, we build iteratively $E_3, E'_3, \ldots, E_p, E'_p, E'_{p+1}$, for some $p \geq 2$. Note that each $\eta_{i,\rho,\delta}$ is an eigenfunction associated to the eigenvalue $r(P)^d$ for the semi-group associated to P^d restricted $E \setminus (E_1 \cup \ldots \cup E_{i-1})$. One easily checks that, to each such eigenfunction, one can build an element $\eta'_{i,\rho,p}$ of $\ker(P^d - r(P)^d I)^i$ on E such that η coincides with $\eta_{i,\rho,\delta}$ on $E_{i,\rho,\delta}$. Since P is quasi-compact, $\sum_{i,d} \ker(P^d - r(P)^d I)^i$ is finite dimensional in B, and hence, since the $\eta'_{i,\rho,p}$ form a free family in B, we deduce that

$$p \leq \dim \left(\sum_{i,d} \ker(P^d - r(P) I)^i \right) < \infty.$$

Step 3. Convergence at times multiples of d. We are now in position to apply [13, Theorem 4.1] to the kernel P^d. Indeed, the state space E can be decomposed in $E_0 = \cup_i E'_i$ and the subsets $E_{i,\rho,\delta}$. Assumption (B1) therein holds true with $W_{i,\rho,\delta} = V_{|E_{i,\rho,\delta}}$ according to (4), Assumption (B2) holds true since communication from the set $E_{i,\rho,\delta}$ to the set $E_{i',\rho',\delta'}$ is only possible (but not necessarily true) if $i < i'$, Assumption (B3) is easily obtained from (5) with $W_0 = V_{|E_0}$ and, finally, Assumption (B4) holds true since $\sum_i W_i = V$ and P is a bounded operator in $L^\infty(V)$. We deduce from [13, Theorem 4.1] that the process $(X_{nd})_{n\geq 0}$ satisfies Assumption (A) in the cited reference: there exists an integer $d \geq 1$, an integer valued function $j : E \to \mathbb{R}_+$, a finite set I and some probability measures

ν_i on E such that $\nu_i(V) < +\infty$ and non-identically zero non-negative functions $\eta_i \in L^\infty(V)$ for each $i \in I$, such that, for all $f \in L^\infty(V)$, all $n \geq 1$ and all $x \in E$,

$$\left| r(P)^{-nd}(nd)^{-j(x)}\mathbb{E}_x(f(X_{dn})\mathbf{1}_{nd<\tau_\partial}) - \sum_{i\in I}\eta_i(x)\nu_i(f)\right| \leq \alpha_n V(x)\|f\|_V, \quad (6)$$

where α_n goes to 0 when $n \to +\infty$. In addition, the fact that $r(P^d) = r(P)^d$ and [13, Theorem 4.1] also entails the last statement of Theorem 5.

Step 4. Conclusion. Let $\nu_{i,k} = r(P)^{-k}\nu_i P^k$, so that $\nu_{i,k+d} = \nu_{i,k}$ for all i, k. Applying (6) to $P^k f$ instead of $P^k f$, we get

$$\left| r(P)^{-nd}(nd)^{-j(x)}\mathbb{E}_x(f(X_{dn+k})\mathbf{1}_{nd+k<\tau_\partial}) - \sum_{i\in I}\eta_i(x)\nu_{i,k}(f)\right| \leq \alpha_n V(x)\|f\|_V.$$

and hence, setting $\eta_{i,k} = r(P)^{-k}\eta_i$,

$$\left| r(P)^{-nd-k}(nd)^{-j(x)}\mathbb{E}_x(f(X_{dn+k})\mathbf{1}_{nd+k<\tau_\partial}) - \sum_{i\in I}\eta_{i,k}(x)\nu_{i,k}(f)\right| \leq r(P)^{-d}\alpha_n V(x)\|f\|_V. \quad (7)$$

In addition, we have

$$\begin{aligned}
r(P)^{-nd-k}&\left|(nd)^{-j(x)} - (nd+k)^{-j(x)}\right|\mathbb{E}_x(|f(X_{dn+k})|\mathbf{1}_{nd+k<\tau_\partial}) \\
&\leq r(P)^{-nd-k}(nd)^{-j(x)}\mathbb{E}_x(|f(X_{dn+k})|\mathbf{1}_{nd+k<\tau_\partial})\left(\frac{(nd+k)^{-j(x)}}{(nd)^{-j(x)}} - 1\right) \\
&\leq \left(\sum_{i\in I}\eta_{i,k}(x)\nu_{i,k}(|f|) + r(P)^{-d}\alpha_n V(x)\|f\|_V\right)\left(\frac{(nd+k)^{-\max j}}{(nd)^{-\max j}} - 1\right) \\
&\leq r(P)^{-d}\left(\sum_{i\in I}\|\eta_i\|_V V(x)\,\nu_{i,k}(V)\|f\|_V + \max_m \alpha_m V(x)\|f\|_V\right)\left(\frac{(nd+k)^{-\max j}}{(nd)^{-\max j}} - 1\right),
\end{aligned}$$

where $\frac{(nd+k)^{-\max j}}{(nd)^{-\max j}} - 1 \to 0$ when $n \to +\infty$. This and (7) allows us to conclude the proof (up to a change in α_n). □

We conclude with two interesting special cases, first when the process is irreducible in the sense of Definition 1 below, and the other one when the state space is topologically irreducible with a lower semi-continuous type property for the semi-group.

Definition 1. We say that the process $(X_n)_{n\in\mathbb{N}}$ is *totally irreducible* if, for all $y \in E$ and $A \in \mathscr{E}$,

$$\exists n \geq 1 \text{ such that } \mathbb{P}_y(X_n \in A) > 0 \Longrightarrow \forall x \in E,\ \exists n \geq 1 \text{ such that } \mathbb{P}_x(X_n \in A) > 0.$$

For a totally irreducible process, the set $E_{1,1,1}$ in the proof of Theorem 5 is equal to E, which immediately entails the following corollary. Note that, if the process is aperiodic, one can choose $d = 1$ in this result.

Corollary 6. *If, in addition to the assumptions of Theorem 5, the process X is totally irreducible, then there exists an integer $d \geq 1$, a probability measure ν on E such that $\nu(V) < +\infty$ and a positive function $\eta \in L^\infty(V)$ such that, for all $f \in L^\infty(V)$, all $n \geq 1$, all $k \in \{0, 1, \dots, d-1\}$ and all $x \in E$,*

$$\left| r(P)^{-dn-k}\mathbb{E}_x(f(X_{dn+k})\mathbf{1}_{nd+k<\tau_\partial}) - \eta(x)\nu_k(f)\right| \leq C\rho^n V(x)\|f\|_V, \qquad (8)$$

where $\nu_k = \nu P^k$, $C > 0$ and $\rho \in (0,1)$.

We focus now on the situation where E is a separable topological space endowed with its Borel σ-field $\mathscr{E}$, X is topologically irreducible (see the definition below) and P has some lower semi-continuous property.

Definition 2. We say that the process $(X_n)_{n\in\mathbb{N}}$ is *topologically irreducible* if, for all open set $O \subset E$, and all $x \in E$, there exists $n \geq 1$ such that $\mathbb{P}_x(X_n \in O) > 0$.

The following result is then an immediate consequence of Corollary 6.

Corollary 7. *If, in addition to the assumptions of Theorem 5, E is a topological space, X is topologically irreducible and, for all $A \in \mathscr{E}$ and all $y \in E$,*

$$\exists n \geq 1 \text{ such that } \mathbb{P}_y(X_n \in A) > 0 \Longrightarrow \liminf_{x\to y} \sup_{n\geq 1} \mathbb{P}_x(X_n \in A) > 0,$$

then the assumptions of Corollary 6 hold true.

3.2 Lyapunov type criterion

Let $(X_n)_{n\in\mathbb{N}}$ be a sub-Markov process on a measurable space $(E, \mathscr{E})$ where $\mathscr{E}$ is countably generated.

We consider the following hypothesis.

Assumption (H). There exist $\theta_1 \in (0, r(P))$, a function $V : E \to (0, +\infty)$ and a measurable subset $E_K \subset E$ such that:

(H1) *(Global Lyapunov criterion).* We have

$$\mathbb{E}_z(V(X_1)\mathbf{1}_{1<\tau_\partial}) \leq \theta_1 V(z), \ \forall z \in E \setminus E_K.$$

(H2) *(Local domination).* There exists a positive compact operator K on $L^\infty(V)$ such that

$$0 \leq Pf(x) \leq Kf(x) + \theta_1 V(x)\|f\|_V, \ \forall x \in E_K, \ f \in L^\infty(V), \ f \geq 0.$$

Under the above assumptions, we have the following result.

Theorem 8. *If Assumption (H) holds true, then the assumptions of Theorem 5 hold true.*

Remark 6. For practical use, we note that $r(P)$ can be bounded from below using Lyapunov type criterion. Following the idea of [11], let us assume that there exists a non-zero non-negative function $\varphi \in L^\infty(V)$ such that $P\varphi \geq \theta\varphi$ for some $\theta > 0$. Then $r(P) \geq \theta$. Indeed, this implies that $P^n\varphi \geq \theta^n\varphi_2$, hence

$$r(P) = \lim_{n\to+\infty} \vert\vert\vert P^n \vert\vert\vert^{1/n} \geq \lim_{n\to+\infty} \frac{\|P^n\varphi\|_V^{1/n}}{\|\varphi\|_V^{1/n}} \geq \theta.$$

Proof of Theorem 8. Let S be the positive kernel defined by

$$Sf(x) = \mathbf{1}_{x\notin E_K} Pf(x) + \mathbf{1}_{x\in E_K}(P-K)_+ f(x),\ \forall x \in E,\ f : L^\infty(V).$$

According to the second line of (H2), we have $\vert\vert\vert S \vert\vert\vert \leq \theta_1 < r(P)$.

Then we have $0 \leq P \leq Q := K + S$ with

$$r_{ess}(Q) \leq r(S) \leq \vert\vert\vert S \vert\vert\vert < r(P).$$

This concludes the proof of Theorem 8. □

It is worthwhile to compare this result with [11], where the authors prove a similar result with a different assumption (H3). More precisely, therein the authors assume that

$$\inf_{n\geq 0} \frac{\inf_{x\in K} \mathbb{P}_x(n < \tau_\partial)}{\sup_{x\in K} \mathbb{P}_x(n < \tau_\partial)}.$$

Despite its apparent simplicity and apparent local aspect, it is usually the most difficult point to check and it involves the control of the trajectories of the process among the whole state space E. The domination property proposed in (H2) can thus be understood as a simpler local alternative.

In practice, there are many ways to prove that an operator K on $L^\infty(V)$ is compact. However, it is often simpler when the base set on which the functions are defined is localized. To deal with this situation, we propose the following localized version of Assumption (H2).

Assumption (H2loc) For all $A > 0$, there exists a compact operator K_A on the function space $L^\infty(\mathbf{1}_{V\leq A})$ such that

$$0 \leq P(f\mathbf{1}_{V\leq A})(x) \leq K_A(f\mathbf{1}_{V\leq A})(x), \quad \forall x \in \{V \leq A\} \text{ and } f \in L^\infty(V).$$

We conclude this section with the following localized result.

Corollary 9. *Assume that (H1) and (H2loc) hold true, and that* $0 < \inf_{E_K} V \leq \sup_{E_K} PV < \infty$. *Then, for any* $p > 1$*, Theorem 5 holds true with* $V^{1/p}$ *instead of* V.

Remark 7. As will be clear in the proof, it is actually sufficient to assume that (H2loc) holds true for one sufficiently small value of $A > 0$, namely such that $A^{p-1} \leq \frac{\sup_{E_K} PV}{\theta_1 \inf_{E_K} V'}$.

Proof of Corollary 9. Let $p > 1$ and set $V' = V^{1/p}$. Let $A > 0$ be such that

$$\frac{1}{A^{p-1}} \frac{\sup_{E_K} PV}{\inf_{E_K} V'} \geq \theta_1,$$

and define the positive operators S, K on $L^\infty(V')$ by

$$Sf(x) = \mathbf{1}_{x \notin E_K} Pf(x) + \mathbf{1}_{x \in E_K} P(f \mathbf{1}_{\{V > A\}})(x).$$

and

$$Kf(x) = \mathbf{1}_{\{V \leq A\}} K_A(f \mathbf{1}_{V \leq A})(x),$$

for all $x \in E$ and $f \geq 0$. We observe that, denoting by $\mathscr{B}$ the unit ball of $L^\infty(V')$, $K(\mathscr{B})$ is compact in $L^\infty(\mathbf{1}_{V \leq A}) = L^\infty(\mathbf{1}_{V' \leq A^{1/p}})$, which is itself continuously embedded in $L^\infty(V')$. We deduce that K is a compact operator on $L^\infty(V')$. In addition, for all $f \in L^\infty(V')$,

$$\begin{aligned}
\left| Sf(x) \right| &\leq \mathbf{1}_{x \notin E_K} PV'(x) \, \|f\|_{V'} + \mathbf{1}_{x \in E_K} P(V' \mathbf{1}_{V > A})(x) \|f\|_{V'} \\
&\leq \mathbf{1}_{x \notin E_K} (PV(x))^{1/p} \, \|f\|_{V'} + \mathbf{1}_{x \in E_K} \frac{PV(x)}{A^{p-1}} \|f\|_{V'} \\
&\leq \theta_1 V'(x) \|f\|_{V'}.
\end{aligned}$$

In particular, we have $0 \leq P \leq K + S$ with K compact on $L^\infty(V')$ and $|||S||| \leq \theta_1$, so that $K + S$ is quasi-comapct with $r_{ess}(Q) \leq \theta_1 < r(P)$. Hence the Assumptions of Theorem 5 are verified. □

Remark 8. There are several ways to prove that a linear operator K is compact on $L^\infty(E')$ for some $E' \subset E$, in particular when this subset is for instance compact or bounded.

1. A simple situation is for instance when there exists a summable sequence of non-negative numbers $(a_k)_{k \geq 0}$ and of probability measures $(\nu_k)_{k \geq 0}$ on E' such that

$$K = \sum_{k \geq 0} a_k \nu_k.$$

2. An other classical situation is when E' is a compact subset (of a topological set E), and $K = R^2$ with R strong Feller, meaning that, for all $f \in L^\infty(E')$, Rf is bounded continuous.

3.3 Quasi-compactness criterion for quasi-stationary distributions in continuous time

Let $(X_t)_{t\geq 0}$ be a continuous time Markov process on a separable topological space $E\cup\{\partial\}$, with E endowed with its Borel σ-field $\mathscr{E}$ and $\{\partial\}\not\subset E$ is measurable and absorbing (see [40, Chapter III], Definition (1.1) and Section 3). We denote the absorption time of X by $\tau_\partial := \inf\{t\geq 0,\ X_t=\partial\}$ and the semi-group of X on E by $(P_t)_{t\geq 0}$, which is defined, for all $x\in E$ and $f\geq 0$, by

$$P_t f(x) := \mathbb{E}_x(f(X_t)\mathbf{1}_{t<\tau_\partial}) = \mathbb{E}_x(f(X_t)\mathbf{1}_{X_t\neq\partial}).$$

Theorem 10. *Assume that, for all bounded measurable function $f\geq 0$, the map $t\in[0,+\infty)\mapsto P_t f(x)$ is measurable and that there exists $T>0$ and a measurable function $V:E\to(0,+\infty)$ such that P_T is a quasi-compact operator on $L^\infty(V)$. Assume in addition that $P_t V\leq C_T V$ for all $t\in[0,T]$ and some constant $C_T>0$. Then*

$$\left|\frac{r(P_T)^{-t}}{(1+t)^{j(x)}}\mathbb{E}_x(f(X_t)\mathbf{1}_{t<\tau_\partial}) - \sum_{i\in I}\eta_i(x)\nu_i(f)\right| \leq \alpha_t V(x)\|f\|_V, \tag{9}$$

where $\alpha_t\to 0$ when $t\to+\infty$, and where the objects η_i,ν_i,j are the same as in the statement of Theorem 5 applied to X_T..

Remark 9. There are two remarkable consequences of the above result on positive continuous time semi-groups $(P_t)_{t\geq 0}$ on an $L^\infty(V)$ space (recall also that this result implicitly applies to general non-conservative semi-groups via V-transforms). First, if P_T is quasi-compact for some $T>0$ and $P_t V\leq C_T V$ for all $t\in[0,T]$ and some constant $C_T>0$, then P_t is quasi-compact for all $t>0$. Second, under theses conditions, $(P_t)_{t\geq 0}$ cannot have a periodic behaviour, contrarily to the discrete time setting, so that, in Theorem 5 applied to P_T, one can take $d=1$.

As far as we know, Thorem 10 and these two consequences are new also in the context of conservative semi-groups.

Remark 10. Since the above theorem relies on the property of the discrete time semi-group $(P_{nT})_{n\geq 0}$, all the discrete time setting results (Corollaries 6 and 7 and the results of Section 3.2), are useful in the context of continuous time semi-groups.

Remark 11. If $(P_t)_{t\in[0,+\infty)}$ is not measurable in t, one can also adapt the discrete time results of the previous sections to the continuous time setting following the classical lines of [36] or, in the context of quasi-stationary distributions, [11]. For instance, following some of the main principles of the above references, the result holds true with I of size 1 if the process satisfies in addition the Doeblin/Lyapunov conditions [11, Assumptions F0, F1, F2]. The treatment of this situation is very similar to this reference, and thus we do not detail it any further. Following this direction, and similarly as in the discrete time Lyapunov type condition, allows to replace the Harnack type condition F3 in [11] with a quasi-compactness condition, which, in some cases and thanks to Theorem 1, can be much easier.

Proof of Theorem 10. Assume without loss of generality that $T = 1$. According to Theorem 5 applied to the T-skeleton chain of X, we deduce that there exist an integer $d \geq 1$, an integer valued bounded function $j : E \to \mathbb{R}_+$, a finite set I and some probability measures ν_i on E such that $\nu_i(V) < +\infty$ and non-identically zero non-negative functions $\eta_i \in L^\infty(V)$ for each $i \in I$, such that, for all $f \in L^\infty(V)$, all $n \geq 1$ and all $x \in E$,

$$\left| \frac{r(P_1)^{-dn}}{(dn)^{j(x)}} \mathbb{E}_x(f(X_{dn})\mathbf{1}_{nd<\tau_\partial}) - \sum_{i\in I} \eta_i(x)\nu_i(f) \right| \leq \bar{\alpha}_{nd} V(x) \|f\|_V, \tag{10}$$

where $\bar{\alpha}_n$ goes to 0 when $n \to +\infty$.

For any fixed $j \in I$, any $h \geq 0$ and $x \in E_j$, we deduce, replacing f by $P_h f$ in the above inequality and using the fact that $\eta_i(x) = 0$ for all $i \neq j$, that

$$\left| \frac{r(P_1)^{-dn}}{(dn)^{j(x)}} \mathbb{E}_x(f(X_{dn+h})\mathbf{1}_{nd<\tau_\partial}) - \eta_j(x)\nu_j(P_h f) \right| \leq \bar{\alpha}_{nd} V(x) \|P_h f\|_V.$$

Similarly, taking as initial distribution $\delta_x P_h$, we obtain

$$\left| \frac{r(P_1)^{-dn}}{(dn)^{j(x)}} \mathbb{E}_x(f(X_{dn+h})\mathbf{1}_{nd<\tau_\partial}) - \sum_{i} \delta_x P_h \eta_i \nu_i(f) \right| \leq \bar{\alpha}_{nd} \delta_x P_h V \|f\|_V.$$

Since $P_h V(x) < \infty$ by assumption, we deduce from the two previous inequalities that, for all $h \geq 0$,

$$\eta_j(x)\nu_j(P_h f) = \sum_{i\in I} \delta_x P_h \eta_i \nu_i(f). \tag{11}$$

In particular, there exists a finite non-negative matrix A_h indexed by $I \times I$ such that

$$r(P_1)^{-h}\nu_j P_h = \sum_{i\in I} [A_h]_{ji}\nu_i.$$

Using the vectorized notation ν for the (column) vector $(\nu_j)_{j\in I}$, we deduce that

$$r(P_1)^{-h}\nu P_h = A_h \nu, \text{ so that } r(P_1)^{-s-t}\nu P_s P_t = r(P_1)^{-t} A_s \nu P_t = A_s A_t \nu.$$

Since $(\nu_j)_{j\in I}$ forms a free family and since $r(P)^{-s-t}\nu P_s P_t = r(P)^{-s-t}\nu P_{s+t} = A_{s+t}\nu$, we observe that A_t is uniquely determined for all $t > 0$ and hence that, for all $s, t \geq 0$,

$$A_s A_t = A_{s+t}.$$

We claim that, denoting $\mathtt{I}$ the identity matrix indexed by $I \times I$,

$$A_t = \mathtt{I}, \quad \forall t \geq 0. \tag{12}$$

Once this is proved, we can conclude: it implies that, for all $h \in [0, d]$, $\nu_j P_h = r(P_1)^h \nu_j$ for all $j \in E$ and hence, from (10) applied to $P_h f$ instead of f, that

$$\left| \frac{r(P_1)^{-dn}}{(dn)^{j(x)}} \mathbb{E}_x(f(X_{dn+h}) \mathbf{1}_{nd<\tau_\partial}) - r(P_1)^h \sum_{i \in I} \eta_i(x) \nu_i(f) \right| \leq \bar{\alpha}_{nd} V(x) \|P_h f\|_V$$
$$\leq C_T \bar{\alpha}_{nd} V(x) \|f\|_V.$$

This entails (9) with

$$\alpha_t := C \bar{\alpha}_{\lfloor t/nd \rfloor nd} + C(1+t)^{-\|j\|_\infty - 1} \mathbf{1}_{\|j\|_\infty \geq 1},$$

for some constant $C > 0$.

It remains to prove (12). Since $r(P)^{-nd} \nu P_{dn} = \nu$ for all $n \geq 0$, we have

$$A_{nd} = \mathtt{I}, \ \forall n \geq 0.$$

As a consequence, for all $t \geq 0$, $A_t A_{\lfloor t/d \rfloor d + d - t} = \mathtt{I}$ and hence A_t is invertible with inverse $A_{\lfloor t/d \rfloor d + d - t}$, which is a matrix with non-negative entries. In particular, we deduce from [17] that, for all $t \geq 0$, A_t is the product of a diagonal matrix D_t and a permutation matrix Q_t. Since this decomposition is unique, we deduce from the properties of $(A_t)_{t \geq 0}$ that, for all $s, t \geq 0$,

$$D_{s+t} = D_s D_t \text{ and } Q_{s+t} = Q_s Q_t.$$

Setting $Q_t := Q_{-t}^{-1}$ for all $t < 0$, the family $(Q_t)_{t \in \mathbb{R}}$ defines a measurable homeomorphism between the topological groups $(\mathbb{R}, +)$ and $(S_I, \times)$ (where S_I denotes the group of permutation matrices indexed by $I \times I$). Since $\mathbb{R}$ has the Baire property and S_I is separable, we deduce from Pettis Theorem, and more precisely from its corollary [30, Theorem 9.10], that Q is continuous in $t \in \mathbb{R}$. Since the set S_I is discrete, we deduce that Q is constant and hence that

$$Q_t = Q_0 = \mathtt{I}, \ \forall t \geq 0.$$

In addition, since $D_{s+t} = D_s D_t$ for all $s, t \geq 0$ and since $t \mapsto D_t$ is bounded over finite intervals (since this is clearly the case for A_t), we deduce that the diagonal elements of D_t are of the form $e^{c_i t}$ for some $c_i \in \mathbb{R}$, $i \in \mathtt{I}$. Since $D_{nd} = \mathtt{I}$ for all $n \geq 0$, we deduce that $c_i = 0$ for all $i \in \mathtt{I}$ and hence that

$$D_t = \mathtt{I}, \ \forall t \geq 0.$$

We have thus proved the claim (12), which concludes the proof of Theorem 10. □

4 Applications

4.1 Markov chain with laziness and killing

Let $(E,\mathscr{E})$ be measurable spaces with $\mathscr{E}$ countably generated and let R be a Markov transition kernel on E, and let $\partial \notin E$. We consider the Markov process X on $E \cup \{\partial\}$ with the following transition dynamics:

$$\mathbb{P}_x(X_1 \in \cdot) = \rho_R(x)R(x,\cdot) + \rho_\delta(x)\,\delta_x + \rho_\partial(x)\delta_\partial,$$

where ρ_R, ρ_δ and ρ_∂ are non-negative measurable functions on $[0,1]$ such that $\rho_R + \rho_\delta + \rho_\partial \equiv 1$, and $\rho_\partial(\partial) = 1$. Said differently, this is a simple Markov process which, when at a point x, evolves according to the transition R with probability $\rho(x)$, is lazy (i.e. it stays in x) with probability $\rho_\delta(x)$, and is sent to the cemetery point ∂ otherwise. Note that ∂ is absorbing.

Corollary 11. *If there is a probability measure μ on E such that*

$$a := \left\| \frac{\mathrm{d}R(x,\cdot)}{\mathrm{d}\mu} \right\|_\infty < \infty$$

and if $\|\rho_\delta\|_\infty + \|\rho_\partial\|_\infty < 1$, then Theorem 5 holds true with $V \equiv 1$.

This example, despite its simplicity, doesn't fit in existing framework for the existence of a quasi-stationary distribution (apart when E is countable). The main reasons are that the family of laws $\mathbb{P}_x(X_1 \in \cdot)$, $x \in E$ is degenerate (because of the δ_x is the definition) and is not regular (since we do not make any assumption on the regularity of the functions ρ_R, ρ_δ and ρ_∂).

Proof of Corollary 11. Let P be the operator on $L^\infty(E)$ associated with X. Then we have $0 \leq P \leq K + S$, where $Kf(x) := a\mu(f)$ and $Sf(x) := \rho_\delta(x)f(x)$ define positive operators such that K is compact (since its range is one-dimensional) and $\rho(S) \leq \|\rho_\delta\|_\infty < 1 - \|\rho_\partial\|_\infty \leq r(P)$. □

4.2 Continuous time processes on $\mathbb{R}^d$ with a locally bounded density and a Lyapunov type function

Let $(\bar{X}_t)_{t\in[0,+\infty)}$ be a continuous time Markov process on $\mathbb{R}^d$, $d \geq 1$. For any $t > 0$, we say that the law of $\bar{X}_t$ admits a locally bounded density if, for all $R > 0$,

$$\sup_{x\in\mathbb{R}^d,\ \|x\|\leq R} \left\| \frac{\mathrm{d}\mathbb{P}(\bar{X}_{t_R} \in \cdot)}{\mathrm{d}\lambda_d} \right\|_\infty < \infty,$$

where λ_d is the Lebesgue measure. This property is known to hold true for many families of continuous time processes since the seminal work of Aronson [3], under mild regularity assumptions, see e.g. the recent works on kinetic stochastic differential equations [16, 29, 38] and references therein.

Let $D \subset \mathbb{R}^d$ and denote by $\tau_D := \inf\{t \geq 0,\ \bar{X}_t \notin D\}$ the exit time from D. We define the process $(X_t)_{t\in[0,+\infty)}$ as

$$X_t := \begin{cases} \bar{X}_t & \text{if } t < \tau_D, \\ \partial & \text{otherwise,} \end{cases}$$

where $\partial \notin \mathbb{R}^d$ is a cemetery point. We denote by $(P_t)_{t\in[0,+\infty)}$ the semi-group associated to the sub-Markov process X on D.

Corollary 12. *Assume that $t \in [0,+\infty) \mapsto P_t$ is measurable and that there exists $T > 0$ such that $\bar{X}_T$ admits a locally bounded density. Assume in addition that there exists a function $V : \mathbb{R}^d \to (0,+\infty)$ which is locally bounded away from 0 and ∞ and such that*

$$r(P_T) > \limsup_{\|x\|\to+\infty} \frac{\mathbb{E}_x(V(\bar{X}_T))}{V(x)}.$$

Then the assumptions of Theorem 10 hold true. In particular, this is the case if, for some $T > 0$, $r(P_T) > 0$ and $\limsup_{\|x\|\to+\infty} \frac{\mathbb{E}_x(V(\bar{X}_T)}{V(x)} = 0$.

Remark 12. In the present setting, we may have $r(P_T) = 0$, which means in our case that $\limsup_{t\to+\infty} \frac{1}{t} \ln \mathbb{E}_x(V(X_t)\mathbf{1}_{t<\tau_\partial}) = -\infty$ for all $x \in D$. This is for instance the case if $d = 2$, with

$$\mathrm{d}\bar{X}^1_t = \bar{X}^2_t\,\mathrm{d}t, \quad \mathrm{d}\bar{X}^2_t = \mathrm{d}B_t,$$

with B a standard one dimensional Brownian motion, and with D a bounded subset of $\mathbb{R} \times (1,+\infty)$, so that the process exit D with probability 1 after a time larger than the position diameter of D. In this case, there may be quasi-stationary distributions, but the conditional law of the process is ill defined for t large enough.

Remark 13. In [22] and [10], under additional structural or regularity assumptions, such as strong Feller continuity of X or Harnack inequality, the case of kinetic diffusions killed at the boundary of a cylinder is considered. Our result is inherently more general and, perhaps even more interestingly, allows for much shorter proofs.

Proof. It is sufficient to prove that P_T is quasi-compact. In order to do so, we aim to apply Theorem 8 to the operator P_T. Let $R > 0$ be large enough so that

$$\theta_1 := \sup_{\|x\|>R} \frac{\mathbb{E}_x(V(\bar{X}_T))}{V(x)} < r(P_T).$$

and let $E_K = \{\|x\| \le R\}$. Then Assumption (H1) is immediate (by definition of θ_1. In addition, for all non-negative $f \in L^\infty(V)$,

$$0 \le P_T f \le \sup_{x\in\mathbb{R}^d,\ \|x\|\le R} \left\| \frac{\mathrm{d}\mathbb{P}(\bar{X}_{t_R} \in \cdot)}{\mathrm{d}\lambda_d} \right\|_\infty \lambda_d(f),$$

which implies (H2) with $K : f \mapsto \lambda_d(f)$. □

References

[1] Y. Abramovich and A. Wickstead. Solutions of several problems in the theory of compact positive operators. *Proceedings of the American Mathematical Society*, 123(10):3021–3026, 1995.

[2] C. D. Aliprantis and O. Burkinshaw. *Positive operators*, volume 119. Springer Science & Business Media, 2006.

[3] D. G. Aronson. Non-negative solutions of linear parabolic equations. *Annali della Scuola Normale Superiore di Pisa-Scienze Fisiche e Matematiche*, 22(4):607–694, 1968.

[4] J. Banasiak, L. Arlotti, et al. *Perturbations of positive semigroups with applications*, volume 17. Springer, 2006.

[5] M. Benaïm, N. Champagnat, W. Oçafrain, and D. Villemonais. Quasi-compactness criterion for strong Feller kernels with an application to quasi-stationary distributions. working paper or preprint, Apr. 2022.

[6] V. Caselles. On the peripheral spectrum of positive operators. *Israel Journal of Mathematics*, 58:144–160, 1987.

[7] M. M. Castro, J. S. W. Lamb, G. O. Méndez, and M. Rasmussen. Existence and uniqueness of quasi-stationary and quasi-ergodic measures for absorbing markov chains: a banach lattice approach, 2021.

[8] P. Cattiaux, P. Collet, A. Lambert, S. Martínez, S. Méléard, and J. San Martín. Quasi-stationary distributions and diffusion models in population dynamics. *Ann. Probab.*, 37(5):1926–1969, 2009.

[9] P. Cattiaux and S. Méléard. Competitive or weak cooperative stochastic Lotka-Volterra systems conditioned to non-extinction. *J. Math. Biol.*, 60(6):797–829, 2010.

[10] N. Champagnat, T. Lelièvre, M. Ramil, J. Reygner, and D. Villemonais. Quasi-stationary distribution for kinetic sdes with low regularity coefficients, 2024.

[11] N. Champagnat and D. Villemonais. General criteria for the study of quasi-stationarity. *ArXiv e-prints*, Dec. 2017.

[12] N. Champagnat and D. Villemonais. Practical criteria for r-positive recurrence of unbounded semigroups. *arXiv preprint arXiv:1904.08619*, 2019.

[13] N. Champagnat and D. Villemonais. Quasi-stationary distributions in reducible state spaces, 2024.

[14] P. Collet, S. Martínez, and J. San Martín. *Quasi-stationary distributions: Markov chains, diffusions and dynamical systems.* Springer Science & Business Media, 2012.

[15] B. De Pagter and A. Schep. Measures of non-compactness of operators in banach lattices. *Journal of functional analysis*, 78(1):31–55, 1988.

[16] P. C. de Raynal, S. Menozzi, A. Pesce, and X. Zhang. Heat kernel and gradient estimates for kinetic sdes with low regularity coefficients. *Bulletin des Sciences Mathématiques*, 183:103229, 2023.

[17] R. DeMarr. Nonnegative matrices with nonnegative inverses. *Proceedings of the American Mathematical Society*, 35(1):307–308, 1972.

[18] P. Dodds and D. Fremlin. Compact operators in banach lattices. *Israel Journal of Mathematics*, 34:287–320, 1979.

[19] L. Dubins and D. Freedman. Measurable sets of measures. *Pacific J. Math.*, 14(4):1211–1222, 1964.

[20] D. E. Edmunds and W. D. Evans. *Spectral theory and differential operators.* Oxford University Press, 2018.

[21] G. Ferré, M. Rousset, and G. Stoltz. More on the long time stability of Feynman-Kac semigroups. *arXiv e-prints*, page arXiv:1807.00390, Jul 2018.

[22] A. Guillin, B. Nectoux, and L. Wu. Quasi-stationary distribution for strongly feller markov processes by lyapunov functions and applications to hypoelliptic hamiltonian systems. *Journal of the European Mathematical Society (EMS Publishing)*, 26(8), 2024.

[23] A. Hening and M. Kolb. Quasistationary distributions for one-dimensional diffusions with singular boundary points. *Stochastic Processes and their Applications*, 129(5):1659 – 1696, 2019.

[24] A. Hening, W. Qi, Z. Shen, and Y. Yi. Quasi-stationary distributions of multi-dimensional diffusion processes. *arXiv preprint arXiv:2102.05785*, 2021.

[25] H. Hennion. Quasi-compactness and absolutely continuous kernels, applications to markov chains. *arXiv preprint math/0606680*, 2006.

[26] H. Hennion. Quasi-compactness and absolutely continuous kernels. *Probability theory and related fields*, 139(3):451–471, 2007.

[27] H. Hennion and L. Hervé. *Limit theorems for Markov chains and stochastic properties of dynamical systems by quasi-compactness.* Springer, 2001.

[28] G. Hinrichs, M. Kolb, and V. Wachtel. Persistence of one-dimensional AR(1)-sequences. *ArXiv e-prints*, Jan. 2018.

[29] H. Hou and X. Zhang. Heat kernel estimates for nonlocal kinetic operators. *arXiv preprint arXiv:2410.18614*, 2024.

[30] A. Kechris. *Classical descriptive set theory*, volume 156. Springer Science & Business Media, 2012.

[31] M. Kolb and D. Steinsaltz. Quasilimiting behavior for one-dimensional diffusions with killing. *Ann. Probab.*, 40(1):162–212, 2012.

[32] J. Littin C. Uniqueness of quasistationary distributions and discrete spectra when ∞ is an entrance boundary and 0 is singular. *J. Appl. Probab.*, 49(3):719–730, 2012.

[33] M. Lladser and J. San Martin. Domain of attraction of the quasi-stationary distributions for the Ornstein-Uhlenbeck process. *J. Appl. Probab.*, 37(2):511–520, 2000.

[34] J. Martínez. The essential spectral radius of dominated positive operators. *Proceedings of the American Mathematical Society*, 118(2):419–426, 1993.

[35] S. Méléard and D. Villemonais. Quasi-stationary distributions and population processes. *Probability Surveys*, 9:340–410, 2012.

[36] S. P. Meyn and R. L. Tweedie. Stability of markovian processes ii: Continuous-time processes and sampled chains. *Advances in Applied Probability*, 25(3):487–517, 1993.

[37] L. D. Pitt. A compactness condition for linear operators on function spaces. *Journal of Operator Theory*, pages 49–54, 1979.

[38] C. Ren and X. Zhang. Heat kernel estimates for kinetic sdes with drifts being unbounded and in kato's class. *Bernoulli*, 31(2):1402–1427, 2025.

[39] D. Revuz. *Markov chains*, volume 11. Elsevier, 2008.

[40] L. C. G. Rogers and D. Williams. *Diffusions, Markov processes, and martingales*, volume 1. Cambridge university press, 2000.

[41] A. F. Ruston. *Fredholm theory in Banach spaces.* Number 86. Cambridge University Press, 2004.

[42] C. F. Sanchez, P. Gabriel, and S. Mischler. On the krein-rutman theorem and beyond. *arXiv preprint arXiv:2305.06652*, 2023.

[43] D. Sasser. Quasi-positive operators. *Pacific Journal of Mathematics*, 14(3):1029–1037, 1964.

[44] D. Steinsaltz and S. Evans. Quasistationary distributions for one-dimensional diffusions with killing. *Transactions of the American Mathematical Society*, 359(3):1285–1324, 2007.

[45] V. G. Troitsky. Measures of non-compactness of operators on banach lattices. *Positivity*, 8(2):165–178, 2004.

[46] E. A. van Doorn. Quasi-stationary distributions and convergence to quasi-stationarity of birth-death processes. *Adv. Appl. Probab.*, 23(4):683–700, 1991.

[47] L. Wu. Essential spectral radius for markov semigroups (i): discrete time case. *Probability theory and related fields*, 128:255–321, 2004.

[48] K. Yosida and S. Kakutani. Operator-theoretical treatment of markoff's process and mean ergodic theorem. *Annals of Mathematics*, 42(1):188–228, 1941.