



**HAL**  
open science

# Quasi-stationary distribution for strongly Feller Markov processes by Lyapunov functions and applications to hypoelliptic Hamiltonian systems

Arnaud Guillin , Boris Nectoux , Liming Wu

## ► To cite this version:

Arnaud Guillin , Boris Nectoux , Liming Wu. Quasi-stationary distribution for strongly Feller Markov processes by Lyapunov functions and applications to hypoelliptic Hamiltonian systems. 2020. hal-03068461v1

**HAL Id: hal-03068461**

**<https://hal.science/hal-03068461v1>**

Preprint submitted on 15 Dec 2020 (v1), last revised 4 Jul 2023 (v2)

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# QUASI-STATIONARY DISTRIBUTION FOR STRONGLY FELLER MARKOV PROCESSES BY LYAPUNOV FUNCTIONS AND APPLICATIONS TO HYPOELLIPTIC HAMILTONIAN SYSTEMS

ARNAUD GUILLIN  $\diamond$ , BORIS NECTOUX  $\diamond$ , AND LIMING WU  $\diamond$

ABSTRACT. In this paper, we establish a general result for the existence and the uniqueness of the quasi-stationary distribution  $\mu_{\mathcal{D}}$  of a strongly Feller Markov process  $(X_t, t \geq 0)$  killed when it exits a domain  $\mathcal{D}$ , under some Lyapunov function condition. Our result covers the case of hypoelliptic damped Hamiltonian systems. Our method is based on the characterization of the essential spectral radius by means of Lyapunov functions and measures of non-compactness.

*AMS 2010 Subject classifications.* 74A25, 37A60, 60B10, 60F99, 60J25, 60J60, 37A30.

*Key words and Phrases.* Quasi-stationary distribution, strong Feller property, Lyapunov functions, essential spectral radius, hypoelliptic diffusions.

## 1. INTRODUCTION

**1.1. Setting and literature.** The notion of quasi-stationary distribution is a central object in the study of population processes or more generally of models derived from biological systems, see for instance [19, 49, 20, 10, 13, 17] and references therein. More recently, the notion of quasi-stationary distribution has attracted a lot of attention in the mathematical justification of the very efficient accelerated dynamics algorithms [58, 65, 64, 53] (see also [54]) which are widely used in practice and aim at simulating the atomistic evolution of statistical systems over long time scales (by accelerating the sampling of the exit event from a metastable macroscopic state  $\mathcal{D}$ ). Let us be more precise on this. A typical process used in simulation in statistical physics to model the evolution of the positions of the particles of a system is a (stochastic) hypoelliptic damped Hamiltonian systems  $((x_t, v_t), t \geq 0)$  on  $\mathbb{R}^d \times \mathbb{R}^d$  (see (6.1)), where  $x_t \in \mathbb{R}^d$  gathers the positions of the particles of the system and  $v_t \in \mathbb{R}^d$  their velocities at time  $t \geq 0$ , ( $d = 3\mathbf{N}$ ,  $\mathbf{N}$  being the number of particles). In most applications of the algorithms mentioned above, the macroscopic state is typically associated with a subdomain  $\mathcal{D}$  of  $\mathbb{R}^{2d}$  of the form  $\mathcal{D} = \mathbf{O} \times \mathbb{R}^d$ , where  $\mathbf{O}$  is a subdomain of  $\mathbb{R}^d$ . The set  $\mathbf{O}$  is defined in practice as a neighborhood of some local minimum of the potential energy function  $V : \mathbb{R}^d \rightarrow \mathbb{R}$  (the interatomic potential function), where the process  $(x_t, t \geq 0)$  can spend a huge amount of time before leaving it <sup>1</sup> (in this case,

---

*Date:* December 15, 2020.

<sup>1</sup>Because of the presence of energetic barriers.

$\mathcal{D}$  is called a metastable region). For such domains  $\mathcal{D}$ , it is thus expected that the conditional distribution of  $(x_t, v_t)$  before leaving  $\mathcal{D}$  is close to a local equilibrium inside  $\mathcal{D}$ . This local equilibrium inside  $\mathcal{D}$  is described by a quasi-stationary distribution (see (2.3)). The exit event from  $\mathcal{D}$  can thus be studied starting from this quasi-stationary distribution [41, 24, 1, 46] which is at the heart of the mathematical analysis of the accelerated dynamics algorithms mentioned above (see [43, 24, 26, 42, 25, 1] when the considered process is the overdamped Langevin process, an elliptic diffusion). In this work, we consider the existence and the uniqueness of the quasi-stationary distribution for such hypoelliptic processes on such domains  $\mathcal{D}$  (see (6.1) and Theorem 6.8 below).

Many different criteria have been given to ensure the existence and uniqueness of a quasi-stationary distribution for different Markov processes, see [14, 32, 15, 35, 16, 55, 39, 59, 61, 69, 36]. In particular, for elliptic diffusions killed when exiting a bounded subdomain of  $\mathbb{R}^d$ , the existence and uniqueness of a quasi-stationary distribution are well known, see for instance [37, 16, 14, 41, 55, 35]. When considering unbounded domains, it is known that it might exist many quasi-stationary distributions, see [47]. We finally refer to [28, 27] for the study of quasi-stationary distributions on a finite state space (see also [18] when considering a discrete state space) and to [21, 63, 8, 31, 49], and references therein, for the approximation of the quasi-stationary distribution using interacting particle system in different settings (see also [6, 7]).

**1.2. Purpose of this work.** The purpose of this work is to propose a quite general framework (see more precisely **(C1)**-**(C5)** in the next section) in which we can establish for general strong Feller Markov processes  $(X_t, t \geq 0)$  valued in a Polish space  $\mathcal{S}$ :

- (1) the existence of quasi-stationary distribution  $\mu_D$  of the process  $(X_t, t \geq 0)$  inside  $\mathcal{D}$  (see (2.3) for definition);
- (2) the uniqueness of a quasi-stationary distribution  $\mu_D$  satisfying  $\mu_D(\mathbf{W}^{1/p}) < +\infty$ , where  $\mathbf{W}$  is the Lyapunov functional appearing in **(C3)**;
- (3) the exponential convergence of the conditional distribution  $\mathbb{P}_\nu(X_t \in \cdot | t < \sigma_D)$  towards  $\mu_D$ , for any given initial distribution  $\nu$  such that  $\nu(\mathbf{W}^{1/p}) < +\infty$ .

In particular, when the Lyapunov functional  $\mathbf{W}$  is bounded, the uniqueness of the quasi-stationary distribution holds (see (b) in Theorem 2.2 below). Note that the use of Lyapunov type conditions to study quasi-stationarity is also present in the recent works [37, 16, 14], a well known tool to assert the existence of a spectral gap for the semigroup of the non killed Markov process (see for example [29] in weighted spaces or [3] in  $L^2$ ). Our conditions on the semigroup of the killed Markov processes are quite general for strongly Feller Markov processes. This is shown by the applications to hypoelliptic damped Hamiltonian systems.

The main theoretical result of this work is Theorem 2.2. This theorem is then applied to a wide range of hypoelliptic damped Hamiltonian systems with drifts of regularity  $\mathcal{C}^0$  and when  $\mathcal{D} = \mathbf{O} \times \mathbb{R}^d$  ( $\mathbf{O}$  being a subdomain, not necessarily bounded, of  $\mathbb{R}^d$ ), see (6.1) and Theorem 6.8.

**1.3. Organization.** This paper is organised as follows. In the next section we present the general framework and the main theoretical result Theorem 2.2 for the quasi-stationary distribution. To prove the main result, a first key point is the existence of the spectral gap for the semigroup of the killed Markov process. For obtaining it we will use the measures of non-weak-compactness of a positive kernel introduced in [67] to establish the formula of the essential spectral radius by means of the Lyapunov function for non killed Markov processes, which generalizes [67] from discrete time to continuous time case. The use of measures of non-compactness allows us to obtain the existence of a spectral gap for the semigroup of the killed Markov process. That is the content of Theorem 3.5 in Section 3.

The second key ingredient for the main result is a Perron-Frobenius type theorem (see Theorem 4.1) for a general Feller kernel, when the well known Krein-Rutman theorem cannot be applied (that is the case here). This is the purpose of Section 4.

With those preparations which should have independent interest, we prove the main result in Section 5. Finally the applications to hypoelliptic damped Hamiltonian systems are developed in Section 6.

## 2. MAIN RESULT

**2.1. Framework: notations and conditions.** Let  $(X_t, t \geq 0)$  be a time homogeneous Markov process valued in a metric complete separable (say Polish) space  $\mathcal{S}$ , with càdlàg paths and satisfying the strong Markov property, defined on the filtered probability space  $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t \geq 0}, (\mathbb{P}_x)_{x \in \mathcal{S}})$  where  $\mathbb{P}_x(X_0 = x) = 1$  for all  $x \in \mathcal{S}$ . Its transition probability semigroup is denoted by  $(P_t, t \geq 0)$ . Given an initial distribution  $\nu$  on  $\mathcal{S}$ , we write  $\mathbb{P}_\nu(\cdot) = \int_{\mathcal{S}} \mathbb{P}_x(\cdot) \nu(dx)$ . Under  $\mathbb{P}_\nu$ , the distribution of  $X_0$  is  $\nu$ .

Let  $\mathcal{B}(\mathcal{S})$  be the Borel  $\sigma$ -algebra of  $\mathcal{S}$ ,  $b\mathcal{B}(\mathcal{S})$  the space of all bounded and Borel measurable functions  $f$  on  $\mathcal{S}$  (its norm will be denoted by  $f \in b\mathcal{B}(\mathcal{S}) \mapsto \|f\|_{b\mathcal{B}(\mathcal{S})} = \sup_{x \in \mathcal{S}} |f(x)|$ ). The space  $\mathbb{D}([0, T], \mathcal{S})$  of  $\mathcal{S}$ -valued càdlàg paths defined on  $[0, T]$  is equipped with the Skorokhod topology.

We suppose that

- (C1) **(strong Feller property)** There exists  $t_0 > 0$  such that for each  $t \geq t_0$ ,  $P_t$  is strong Feller, i.e.  $P_t f$  is continuous on  $\mathcal{S}$  for any  $f \in b\mathcal{B}(\mathcal{S})$ .
- (C2) **(trajectory Feller property)** For every  $T > 0$ ,  $x \rightarrow \mathbb{P}_x(X_{[0, T]} \in \cdot)$  (the law of  $X_{[0, T]} := (X_t)_{t \in [0, T]}$ ) is continuous from  $\mathcal{S}$  to the space  $\mathcal{M}_1(\mathbb{D}([0, T], \mathcal{S}))$  of probability measures on  $\mathbb{D}([0, T], \mathcal{S})$ , equipped with the weak convergence topology.

Now let  $\mathcal{D}$  be a non-empty open domain of  $\mathcal{S}$ , different from  $\mathcal{S}$ . Consider the first exit time of  $\mathcal{D}$

$$\sigma_{\mathcal{D}} := \inf\{t \geq 0, X_t \in \mathcal{D}^c\} \tag{2.1}$$

where  $\mathcal{D}^c = \mathcal{S} \setminus \mathcal{D}$  is the complement of  $\mathcal{D}$ . The transition semigroup of the killed process  $(X_t, 0 \leq t < \sigma_{\mathcal{D}})$  is for  $t \geq 0$  and  $x \in \mathcal{D}$ ,

$$P_t^{\mathcal{D}} f(x) = \mathbb{E}_x[\mathbf{1}_{t < \sigma_{\mathcal{D}}} f(X_t)], \tag{2.2}$$

for  $f \in b\mathcal{B}(\mathcal{D})$ . Let us now recall the definition of a quasi-stationary distribution.

**Definition 2.1.** A quasi-stationary distribution (QSD in short) of the Markov process  $(X_t, t \geq 0)$  in the domain  $\mathcal{D}$  is a probability measure on  $\mathcal{D}$  such that

$$\mu_{\mathcal{D}}(A) = \mathbb{P}_{\mu_{\mathcal{D}}}(X_t \in A | t < \sigma_{\mathcal{D}}) = \frac{\mathbb{P}_{\mu_{\mathcal{D}}}(X_t \in A, t < \sigma_{\mathcal{D}})}{\mathbb{P}_{\mu_{\mathcal{D}}}(t < \sigma_{\mathcal{D}})}, \quad \forall t > 0, A \in \mathcal{B}(\mathcal{D}). \quad (2.3)$$

where  $\mathcal{B}(\mathcal{D}) := \{A \cap \mathcal{D}; A \in \mathcal{B}(\mathcal{S})\}$ .

**2.2. Main general result.** For a continuous time Markov process, often what is given is its generator  $\mathcal{L}$ , not its transition semigroup  $(P_t, t \geq 0)$ , which is unknown in general. We say that a continuous function  $f$  belongs to the extended domain  $\mathbb{D}_e(\mathcal{L})$  of  $\mathcal{L}$ , if there is some measurable function  $g$  on  $\mathcal{S}$  such that  $\int_0^t |g|(X_s) ds < +\infty, \mathbb{P}_{\mathbf{x}} - a.e.$  for all  $\mathbf{x} \in \mathcal{S}$ , and

$$M_t(f) := f(X_t) - f(X_0) - \int_0^t g(X_s) ds \quad (2.4)$$

is a  $\mathbb{P}_{\mathbf{x}}$ -local martingale for all  $\mathbf{x}$ . Such a function  $g$ , denoted by  $\mathcal{L}f$ , is not unique in general. But it is unique up to the equivalence of *quasi-everywhere* (q.e.): two functions  $g_1, g_2$  are said to be equal q.e., if  $g_1 = g_2$  almost everywhere in the (resolvent) measure  $R_1(\mathbf{x}, \cdot) = \int_0^{+\infty} e^{-t} P_t(\mathbf{x}, \cdot) dt$  for **every**  $\mathbf{x} \in \mathcal{S}$ .

Let us introduce the Lyapunov function condition:

**(C3)** There exist a continuous function  $W : \mathcal{S} \rightarrow [1, +\infty[$ , belonging to the extended domain  $\mathbb{D}_e(\mathcal{L})$ , two sequences of positive constants  $(r_n)$  and  $(b_n)$  where  $r_n \rightarrow +\infty$ , and an increasing sequence of compact subsets  $(K_n)$  of  $\mathcal{S}$ , such that

$$-\mathcal{L}W(\mathbf{x}) \geq r_n W(\mathbf{x}) - b_n \mathbf{1}_{K_n}(\mathbf{x}), \quad q.e.,$$

where  $\mathbf{1}_{K_n}$  is the indicator function of  $K_n$ .

We say that a class  $\mathcal{A}$  of bounded continuous functions on  $\mathcal{D}$  is *measure-separable*, if for any bounded (signed) measure  $\nu$  on  $\mathcal{D}$ , if  $\nu(f) = 0$  for all  $f \in \mathcal{A}$ , then  $\nu = 0$ .

**Theorem 2.2.** Assume that **(C1)**, **(C2)**, and **(C3)** hold. Suppose that the killed process  $(X_t, 0 \leq t < \sigma_{\mathcal{D}})$  satisfies:

**(C4)** for  $t \geq 0$ ,  $P_t^{\mathcal{D}}$  is weakly Feller, i.e. for a measure-separable class  $\mathcal{A}$  of continuous bounded functions with support contained in  $\mathcal{D}$ ,  $P_t^{\mathcal{D}} f$  is continuous on  $\mathcal{D}$ .

**(C5)** there exists  $t_0 > 0$  such that for all  $t \geq t_0$ , for all  $\mathbf{x} \in \mathcal{D}$  and non-empty open subsets  $O$  of  $\mathcal{D}$ ,

$$P_t^{\mathcal{D}}(\mathbf{x}, O) > 0$$

(we can assume this  $t_0 > 0$  is the same as the one in **(C1)**), and there is some  $\mathbf{x}_0 \in \mathcal{D}$  such that  $\mathbb{P}_{\mathbf{x}_0}(\sigma_{\mathcal{D}} < +\infty) > 0$ .

Then, it holds:

(a) there is only one QSD  $\mu_{\mathcal{D}}$  satisfying

$$\mu_{\mathcal{D}}(W^{1/p}) := \int_{\mathcal{D}} W^{1/p}(\mathbf{x}) \mu_{\mathcal{D}}(d\mathbf{x}) < +\infty$$

for some  $p \in ]1, +\infty[$ .

(b) in particular if  $W$  is bounded over  $\mathcal{D}$ , the QSD inside  $\mathcal{D}$  is unique.

(c) the spectral radius of  $P_t^{\mathcal{D}}$  on  $b_{\mathbb{W}^{1/p}}\mathcal{B}(\mathcal{D})$  equals to  $e^{-\lambda_{\mathcal{D}}t}$  for all  $t \geq 0$  where  $0 < \lambda_{\mathcal{D}} < +\infty$  (which is often called least Dirichlet eigenvalue of the killed Markov process), and there is a unique continuous function  $\varphi$  bounded by  $c\mathbb{W}^{1/p}$  and positive everywhere on  $\mathcal{D}$  such that  $\mu_{\mathcal{D}}(\varphi) = 1$  and

$$\mu_{\mathcal{D}}P_t^{\mathcal{D}} = e^{-\lambda_{\mathcal{D}}t}\mu_{\mathcal{D}}, \quad P_t^{\mathcal{D}}\varphi = e^{-\lambda_{\mathcal{D}}t}\varphi \text{ on } \mathcal{D}, \forall t \geq 0. \quad (2.5)$$

Here  $b_{\mathbb{W}^{1/p}}\mathcal{B}(\mathcal{D})$  is the Banach space of all  $\mathcal{B}(\mathcal{D})$ -measurable functions on  $\mathcal{D}$  so that its norm

$$\|f\|_{b_{\mathbb{W}^{1/p}}\mathcal{B}(\mathcal{D})} := \sup_{\mathbf{x} \in \mathcal{D}} \frac{|f(\mathbf{x})|}{\mathbb{W}^{1/p}(\mathbf{x})} < +\infty.$$

(d) for any  $p \in ]1, +\infty[$  fixed, there are some constants  $\delta > 0$  and  $C \geq 1$  such that for any initial distribution  $\nu$  on  $\mathcal{D}$  with  $\nu(\mathbb{W}^{1/p}) < +\infty$ ,

$$|\mathbb{P}_{\nu}(X_t \in A | t < \sigma_{\mathcal{D}}) - \mu_{\mathcal{D}}(A)| \leq Ce^{-\delta t} \frac{\nu(\mathbb{W}^{1/p})}{\nu(\varphi)}, \quad \forall A \in \mathcal{B}(\mathcal{D}), t > 0. \quad (2.6)$$

(e)  $\mathbb{P}_{\mathbf{x}}(\sigma_{\mathcal{D}} < +\infty) = 1$  for every  $\mathbf{x} \in \mathcal{D}$ ,  $X_{\sigma_{\mathcal{D}}}$  and  $\sigma_{\mathcal{D}}$  are  $\mathbb{P}_{\mu_{\mathcal{D}}}$ -independent and

$$\mathbb{P}_{\mu_{\mathcal{D}}}(t < \sigma_{\mathcal{D}}) = e^{-\lambda_{\mathcal{D}}t}.$$

Notice that the set of initial distributions  $\nu$  on  $\mathcal{D}$  with  $\nu(\mathbb{W}^{1/p}) < +\infty$  includes any initial distribution  $\nu$  with compact support in  $\mathcal{D}$ , and thus includes in particular Dirac measures  $\delta_{\mathbf{x}}$  ( $\mathbf{x} \in \mathcal{D}$ ), which is for instance of interest to analyse the mathematical foundations of the accelerated algorithms we mentioned in the introduction<sup>2</sup>. Theorem 2.2 is applied in Section 6 to hypoelliptic damped Hamiltonian systems, see (6.1) and Theorem 6.8. Let us mention that under the assumptions of Theorem 2.2, from Corollary 3.6 and (5.3), for each  $t > 2t_0$  (see **(C1)**):

$$P_t^{\mathcal{D}} : b_{\mathbb{W}^{1/p}}\mathcal{B}(\mathcal{D}) \rightarrow b_{\mathbb{W}^{1/p}}\mathcal{B}(\mathcal{D}) \text{ is compact.}$$

**Remark 2.3.** From Definition (2.3),  $\mu_{\mathcal{D}}$  is a QSD if and only if

$$\mu_{\mathcal{D}}P_t^{\mathcal{D}} = \lambda(t)\mu_{\mathcal{D}}, \quad \lambda(t) = \mathbb{P}_{\mu_{\mathcal{D}}}(t < \sigma_{\mathcal{D}})$$

in other words,  $\mu_{\mathcal{D}}$  must be a common positive left-eigenvector of  $P_t^{\mathcal{D}}$ . The part (c) above says that  $\lambda(t) = e^{-\lambda_{\mathcal{D}}t}$  is exactly the spectral radius of  $P_t^{\mathcal{D}}$  on  $b_{\mathbb{W}^{1/p}}\mathcal{B}(\mathcal{D})$ .

**Remark 2.4.** If we replace in **(C3)**, the assumption that  $r_n \rightarrow +\infty$  as  $n \rightarrow +\infty$  by  $r_n \rightarrow r_{\infty}$  as  $n \rightarrow +\infty$  with  $r_{\infty} > \lambda_{\mathcal{D}}$ , then all the statements of Theorem 2.2 remain valid. It is the condition used by [16]. Such a criterion is in practice quite hard to check since one does not know the Dirichlet eigenvalue  $\lambda_{\mathcal{D}}$ . Note that **(C3)** is known to imply various functional inequalities, depending on the growth of  $r_n$  to infinity such as logarithmic Sobolev inequalities or various  $F$ -Sobolev type inequalities, see [4, 12, 11]. If  $r_n$  goes to infinity slowly one may show that the  $F$ -Sobolev implied by **(C3)** is weaker than ultracontractivity, condition which was behind the "coming down from infinity" property used in [10] for example.

Let us also comment on the other assumptions imposed in [16]. The authors introduce

<sup>2</sup>Which is based on the comparison of the exit event from  $\mathcal{D}$  when  $X_0 = \mathbf{x} \in \mathcal{D}$  and when  $X_0$  is distributed according to a quasi-stationary distribution of the process  $(X_t, t \geq 0)$  in  $\mathcal{D}$ .

condition (F) which includes in particular the variant of **(C3)**, (F2) in their work, but also a local Harnack inequality (F3) which seems quite hard to verify in non elliptic case. We prefer to use our conditions **(C4)** and **(C5)**, easily verified for multidimensional elliptic diffusion processes thus recovering the extent of their results [16, Sect. 4.], but which will be also useful in hypoelliptic cases.

**Remark 2.5.** The key, like in the current literatures, consists in proving that  $P_t^{\mathcal{D}}$  has a spectral gap at its spectral radius  $r_{sp}(P_t^{\mathcal{D}})$  acting on some Banach lattice space  $\mathbb{B}$  of functions. We will work on  $\mathbb{B} = b_W\mathcal{B}(\mathcal{S}), \mathcal{C}_{bW}(\mathcal{S})$  (introduced in Section 3.3 below), which are well adapted to our Lyapunov function condition **(C3)**. The problem is:  $P_t^{\mathcal{D}}$  is not strongly continuous on such Banach spaces, and the domain  $\mathbb{D}_{\mathbb{B}}(\mathcal{L})$  of the generator is not dense in  $\mathbb{B}$ . So we cannot use the spectral theory of strongly continuous semigroups in Functional Analysis.

### 3. ESSENTIAL SPECTRAL RADIUS OF $P_t$

The purpose of this section is to prove Theorem 3.5 which aims at giving a lower bound on the essential spectral radius of  $P_t, t > 0$ . To this end, we first recall a characterization of the essential spectral radius of a semigroup of transition kernels (see Theorem 3.4) obtained in [67].

**3.1. Essential spectral radius.** Let  $\mathbb{B}$  be a real Banach lattice and  $P$  a nonnegative, linear and bounded operator on  $\mathbb{B}$ . A complex number  $\lambda \in \mathbb{C}$  is said to be in the resolvent set  $\rho(P)$  of  $P$  if the inverse  $(\lambda I - P)^{-1}$  on the complexified Banach space  $\mathbb{B}_{\mathbb{C}}$  of  $\mathbb{B}$  exists and is bounded, by definition. The complementary  $\sigma(P) := \mathbb{C} \setminus \rho(P)$  is the spectrum of  $P$  on  $\mathbb{B}$ . The spectral radius of  $P$  is given by Gelfand's formula

$$r_{sp}(P|_{\mathbb{B}}) := \sup \{ |\lambda|; \lambda \in \sigma(P) \} = \lim_{n \rightarrow +\infty} (\|P^n\|_{\mathbb{B} \rightarrow \mathbb{B}})^{1/n}.$$

A complex number  $\lambda \in \mathbb{C}$  does not belong to the (Wolf) essential spectrum  $\sigma_{ess}(P|_{\mathbb{B}})$  of  $P|_{\mathbb{B}}$ , iff  $\lambda I - P$  is a Fredholm operator on the complexified Banach space  $\mathbb{B}_{\mathbb{C}}$  of  $\mathbb{B}$ , by definition. For a point  $\lambda_0$  in the spectrum  $\sigma(P|_{\mathbb{B}})$ ,  $\lambda_0 \notin \sigma_{ess}(P|_{\mathbb{B}})$  iff  $\lambda_0$  is isolated in  $\sigma(P|_{\mathbb{B}})$  and the associated eigen-projection

$$E_{\lambda_0} := \frac{1}{2\pi i} \int_{\Gamma} (\lambda I - P)^{-1} d\lambda$$

(Dunford integral in the counter-clockwise way) is finite dimensional, where  $\Gamma$  is a circumference of sufficiently small radius:  $|\lambda - \lambda_0| = \delta$  such that the disk  $|\lambda - \lambda_0| \leq \delta$  contains no other spectral point than  $\lambda_0$ . Let us recall that by definition, the algebraic multiplicity of  $\lambda_0 \notin \sigma_{ess}(P|_{\mathbb{B}})$  is the dimension of the range of  $E_{\lambda_0}$ . Let us finally mention that since  $P$  is bounded,  $\lambda_0 \notin \sigma_{ess}(P|_{\mathbb{B}})$  is a pole of the resolvent [68, Theorem 4 in §8 in Chap. VIII].

**Definition 3.1.** *The essential spectral radius of  $P$  on  $\mathbb{B}$  is defined by*

$$r_{ess}(P|_{\mathbb{B}}) = \sup \{ |\lambda|; \lambda \in \sigma_{ess}(P|_{\mathbb{B}}) \}.$$

**3.2. Two parameters of non weak compactness and the formulas of the essential spectral radius.** Our state space  $\mathcal{S}$  is Polish with a compatible metric  $\mathbf{d}$  (i.e.,  $(\mathcal{S}, \mathbf{d})$  is complete and separable), whose Borel  $\sigma$ -field is denoted by  $\mathcal{B}(\mathcal{S})$ . The notation “ $K \subset\subset \mathcal{S}$ ” means that  $K$  is compact in  $\mathcal{S}$ . Let  $\mathcal{M}_b(\mathcal{S})$  (resp.  $\mathcal{M}_b^+(\mathcal{S})$ ,  $\mathcal{M}_1(\mathcal{S})$ ) be the space of all  $\sigma$ -additive (resp. and nonnegative; probability) measures of bounded variation on  $(\mathcal{S}, \mathcal{B})$ . The pair relation between  $\nu \in \mathcal{M}_b(\mathcal{S})$  and  $f \in b\mathcal{B}(\mathcal{S})$  is

$$\langle \nu, f \rangle := \nu(f) := \int_{\mathcal{S}} f(\mathbf{x}) d\nu(\mathbf{x}).$$

Using the pair above,  $\mathcal{M}_b(\mathcal{S})$  is a subspace of the dual Banach space  $(b\mathcal{B}(\mathcal{S}))^*$ . For a nonnegative kernel  $P(\mathbf{x}, d\mathbf{y})$ , bounded on  $b\mathcal{B}(\mathcal{S})$ , its adjoint operator  $P^*$  on  $(b\mathcal{B}(\mathcal{S}))^*$  keeps  $\mathcal{M}_b(\mathcal{S})$  stable, i.e., for each  $\nu \in \mathcal{M}_b(\mathcal{S})$ ,

$$P^*\nu(\cdot) = (\nu P)(\cdot) := \int_{\mathcal{S}} \nu(d\mathbf{x}) P(\mathbf{x}, \cdot) \in \mathcal{M}_b(\mathcal{S}).$$

Besides the variation norm  $\|\nu\|_{TV}$ -topology, we shall also consider the following two weak topologies on  $\mathcal{M}_b(\mathcal{S})$ . The weak topology  $\sigma(\mathcal{M}_b(\mathcal{S}), b\mathcal{B}(\mathcal{S}))$  (i.e., the weakest topology on  $\mathcal{M}_b(\mathcal{S})$  for which  $\nu \mapsto \nu(f)$  is continuous for all  $f \in b\mathcal{B}(\mathcal{S})$ ), according to the usual language, will be called  $\tau$ -topology, denoted simply by  $\tau$ . And the weak topology  $\sigma(\mathcal{M}_b(\mathcal{S}), \mathcal{C}_b(\mathcal{S}))$  (the most often used *weak convergence* topology) will be denoted by “ $w$ ”. The space  $\mathcal{C}_b(\mathcal{S})$  is the space of all functions  $f \in b\mathcal{B}(\mathcal{S})$  such that  $f$  is continuous on  $\mathcal{S}$  (its norm is the one of  $b\mathcal{B}(\mathcal{S})$  but we will sometimes write it  $\|\cdot\|_{\mathcal{C}_b(\mathcal{S})}$  when we want to emphasize that we work on  $\mathcal{C}_b(\mathcal{S})$ ).

The following measures of non weak compactness of a positive (i.e. nonnegative and non-zero) kernel  $P(\mathbf{x}, d\mathbf{y})$  were introduced in the third author’s [67].

**Definition 3.2.** (a) For a bounded sub-family  $\mathcal{M}$  of  $\mathcal{M}_b^+(\mathcal{S})$ , define

$$\begin{aligned} \beta_w(\mathcal{M}) &:= \inf_{K \subset\subset \mathcal{S}} \sup_{\nu \in \mathcal{M}} \nu(K^c) \\ \beta_\tau(\mathcal{M}) &:= \sup_{(A_n)_{n \rightarrow \infty}} \lim_{n \rightarrow \infty} \sup_{\nu \in \mathcal{M}} \nu(A_n) \end{aligned} \tag{3.1}$$

where  $\sup_{(A_n)}$  is taken over all sequences  $(A_n)_n \subset \mathcal{B}(\mathcal{S})$  decreasing to  $\emptyset$ .

(b) Let  $P(\mathbf{x}, d\mathbf{y})$  be a nonnegative kernel on  $\mathcal{S}$  such that  $\sup_{\mathbf{x} \in E} P(\mathbf{x}, \mathcal{S}) = \|P1\|_{b\mathcal{B}(\mathcal{S})} < +\infty$  (i.e., the boundedness of kernel  $P$ ). We call

$$\beta_w(P) := \beta_w(\mathcal{M}); \quad \beta_\tau(P) := \beta_\tau(\mathcal{M}) \tag{3.2}$$

where  $\mathcal{M} = \{P(\mathbf{x}, \cdot); \mathbf{x} \in \mathcal{S}\}$ , measure of non- $\tau$ -compactness and measure of non-“ $w$ ”-compactness of  $P$ , respectively.

Here and in the following,  $\mathbf{1}$  will denote the constant function equal to 1 on  $\mathcal{S}$ . Introduce the following assumption

$$\beta_w(\mathbf{1}_K P) = 0 \text{ and } \exists N \geq 1 : \beta_\tau(\mathbf{1}_K P^N) = 0, \forall K \subset\subset \mathcal{S}. \tag{A1}$$

**Remark 3.3.** If  $P$  is Feller and  $P^k$  is strong Feller on  $\mathcal{S}$ , i.e.  $P^k(b\mathcal{B}(\mathcal{S})) \subset \mathcal{C}_b(\mathcal{S})$ , then (A1) is satisfied with  $N = k$ .

In fact for any sequence  $A_n \downarrow \emptyset$  in  $\mathcal{B}(\mathcal{S})$ ,  $f_n(x) = P^k(x, A_n)$  is continuous and converges



to zero for every  $x \in \mathcal{S}$ . Then  $\varphi_n(\nu) := \nu(f_n)$  is a sequence of continuous functionals on  $\mathcal{M}_b^+(\mathcal{S})$  decreasing to 0 for each  $\nu$ . By Dini's monotone convergence theorem,  $\varphi_n$  converge to zero uniformly over compacts of  $\mathcal{M}_b^+(\mathcal{S})$ . Then for each  $K \subset\subset \mathcal{S}$ ,  $\sup_{x \in \mathcal{S}} \mathbf{1}_K(x) P^k(x, A_n) \rightarrow 0$ . That yields  $\beta_r(\mathbf{1}_K P^k) = 0$ .

When  $P$  is Feller, the fact that  $\beta_w(\mathbf{1}_K P) = 0$  for all  $K \subset\subset \mathcal{S}$  is proved similarly using Prokorov's theorem (see for instance [67, (a.iii) in Lemma 3.1]).

**Theorem 3.4.** ([67, Theorem 3.5]) *Let  $P$  be a bounded nonnegative kernel on  $\mathcal{S}$  satisfying (A1). Then, it holds:*

$$r_{\text{ess}}(P|_{b\mathcal{B}(\mathcal{S})}) = \lim_{n \rightarrow \infty} [\beta_w(P^n)]^{1/n}. \quad (3.3)$$

### 3.3. Lyapunov function criterion for the essential spectral radius of $(P_t, t \geq 0)$ .

The main objective of this section is to apply the results of Theorem 3.4 to  $P_t$ , where we recall that  $(P_t, t \geq 0)$  is the semigroup of the (non killed) process  $(X_t, t \geq 0)$ . Let us first introduce some notation. Given a continuous function  $W : \mathcal{S} \rightarrow [1, +\infty[$  (weight function), let

$$b_W\mathcal{B}(\mathcal{S}) := \left\{ f : \mathcal{S} \rightarrow \mathbb{R} \text{ measurable such that } \|f\|_{b_W\mathcal{B}(\mathcal{S})} := \sup_{x \in \mathcal{S}} \frac{|f(x)|}{W(x)} < +\infty \right\},$$

and

$$\mathcal{C}_{b_W}(\mathcal{S}) = \{f \in b_W\mathcal{B}(\mathcal{S}) \text{ such that } f : \mathcal{S} \rightarrow \mathbb{R} \text{ is continuous}\}$$

which are Banach spaces with norm  $\|\cdot\|_{b_W\mathcal{B}(\mathcal{S})}$ . Notice that when  $W = 1$ ,  $b_W\mathcal{B}(\mathcal{S}) = b\mathcal{B}(\mathcal{S})$  and  $\mathcal{C}_{b_W}(\mathcal{S}) = \mathcal{C}_b(\mathcal{S})$ , where  $b\mathcal{B}(\mathcal{S})$  and  $\mathcal{C}_b(\mathcal{S})$  are introduced above. The space of measures

$$\mathcal{M}_{b_W}(\mathcal{S}) = \{\nu \in \mathcal{M}_b(\mathcal{S}) \text{ such that } W(x)\nu(dx) \in \mathcal{M}_b(\mathcal{S})\}$$

is a subspace of the dual Banach space  $(b_W\mathcal{B}(\mathcal{S}))^*$  by regarding each  $\nu \in \mathcal{M}_{b_W}(\mathcal{S})$  as a bounded linear form  $f \rightarrow \nu(f)$  on  $b_W\mathcal{B}(\mathcal{S})$ . We now turn to the main result of this section.

**Theorem 3.5.** *Assume (C1) and (C2). Assume that there is some continuous Lyapunov function  $W : \mathcal{S} \rightarrow [1, +\infty[$  such that for some  $K \subset\subset \mathcal{S}$ ,  $r > 0$ , and  $b > 0$ , it holds:*

$$-\frac{\mathcal{L}W}{W} \geq r\mathbf{1}_{K^c} - b\mathbf{1}_K \quad (3.4)$$

and for some  $p > 1$ ,

$$\mathcal{L}W^p \leq bW^p. \quad (3.5)$$

Then for every  $t > 0$ ,

$$\beta_w(P_{t,W}) \leq e^{-rt} \text{ and in particular } r_{\text{ess}}(P_t|_{b_W\mathcal{B}(\mathcal{S})}) \leq e^{-rt} \quad (3.6)$$

where for  $x, y \in \mathcal{S}$ , we set

$$P_{t,W}(x, dy) = \frac{W(y)}{W(x)} P_t(x, dy). \quad (3.7)$$

*Proof.* Let  $t > 0$  be fixed. Consider the isomorphism  $M_W : f \rightarrow Wf$  from  $b\mathcal{B}(\mathcal{S})$  to  $b_W\mathcal{B}(\mathcal{S})$ . We have, for  $t \geq 0$ ,  $P_{t,W} = M_W^{-1}P_tM_W$ . Then

$$r_{ess}(P_t|_{b_W\mathcal{B}(\mathcal{S})}) = r_{ess}(P_{t,W}|_{b\mathcal{B}(\mathcal{S})}).$$

Then,  $(P_{t,W}, t \geq 0)$  is again a semigroup of transition kernels, but it is not Markov in general.

**Step 1.** Let us check that  $Q = P_{t,W}$  satisfies **(A1)** for  $t > 0$ . At first, from (3.5),  $(e^{-bt}W^p(X_t), t \geq 0)$  is a supermartingale, and then for any  $x \in \mathcal{S}$ ,  $t \geq 0$ ,

$$P_tW^p(x) = e^{bt}\mathbb{E}_x[e^{-bt}W^p(X_t)] \leq e^{bt}W^p(x).$$

Therefore, for any compact  $K \subset\subset \mathcal{S}$ , letting  $q = p/(p-1)$ , we have using in addition Hölder's inequality,

$$\begin{aligned} \beta_w(\mathbf{1}_K Q) &= \inf_{K' \subset\subset \mathcal{S}} \sup_{x \in K} Q(x, \mathcal{S} \setminus K') \\ &\leq \inf_{K' \subset\subset \mathcal{S}} \sup_{x \in K} \frac{((P_tW^p)(x))^{1/p}}{W(x)} P_t(x, \mathcal{S} \setminus K')^{1/q} \\ &\leq e^{bt/p} \left( \inf_{K' \subset\subset \mathcal{S}} \sup_{x \in K} P_t(x, \mathcal{S} \setminus K') \right)^{1/q} \\ &= 0 \end{aligned}$$

by the Feller property of  $P_t$  (guaranteed by **(C2)**), see Remark 3.3.

Let us check the second condition in **(A1)** with some  $N$  so that  $Nt \geq t_0$  (see **(C1)**).

$$\begin{aligned} \beta_\tau(\mathbf{1}_K Q^N) &= \sup_{(A_n)} \limsup_{n \rightarrow \infty} \sup_{x \in K} Q^N(x, A_n) \\ &\leq \sup_{(A_n)} \limsup_{n \rightarrow \infty} \sup_{x \in K} \frac{((P_{Nt}W^p)(x))^{1/p}}{W(x)} P_{Nt}(x, A_n)^{1/q} \\ &\leq e^{Nbt/p} \left( \sup_{(A_n)} \limsup_{n \rightarrow \infty} \sup_{x \in K} P_{Nt}(x, A_n) \right)^{1/q} \end{aligned}$$

where the sup above is taken over all sequences  $(A_n)_n$  in  $\mathcal{B}(\mathcal{S})$  decreasing to  $\emptyset$ . The last factor above, being  $\beta_\tau(\mathbf{1}_K P_{Nt})$ , is equal to zero by the strong Feller property of  $P_{Nt}$  and Remark 3.3.

**Step 2.** Let us prove  $\beta_w(Q) \leq e^{-rt}$ , which yields  $\beta_w(Q^n) \leq \beta_w(Q)^n$  ([67, Proposition 3.2.(e)]) for all  $n$  and then the desired result by Theorem 3.4.

In fact let  $\tau_K := \inf\{s \geq 0; X_s \in K\}$  be the first hitting time to the compact  $K$  (where  $K$  is the compact set appearing in the Lyapunov condition (3.4)). We have

$$\begin{aligned} \beta_w(Q) &= \inf_{K' \subset\subset \mathcal{S}} \sup_{x \in \mathcal{S}} Q(x, \mathcal{S} \setminus K') \\ &= \inf_{K' \subset\subset \mathcal{S}} \sup_{x \in \mathcal{S}} \frac{1}{W(x)} \mathbb{E}_x[W(X_t) \mathbf{1}_{X_t \notin K'}]. \end{aligned} \tag{3.8}$$

Notice that for  $\mathbf{x} \in \mathcal{S}$ , we have

$$\begin{aligned} \frac{1}{\mathbf{W}(\mathbf{x})} \mathbb{E}_x[\mathbf{W}(X_t) \mathbf{1}_{X_t \notin K'}] &\leq \frac{1}{\mathbf{W}(\mathbf{x})} \mathbb{E}_x[\mathbf{W}(X_t) \mathbf{1}_{X_t \notin K', \tau_K \leq t}] + \frac{1}{\mathbf{W}(\mathbf{x})} \mathbb{E}_x[\mathbf{W}(X_t) \mathbf{1}_{\tau_K > t}] \\ &\leq e^{bt/p} \left( \sup_{y \in K} \mathbb{P}_y(\exists s \in [0, t], X_t \notin K') \right)^{1/q} + \frac{1}{\mathbf{W}(\mathbf{x})} \mathbb{E}_x[\mathbf{W}(X_t) \mathbf{1}_{\tau_K > t}] \end{aligned} \quad (3.9)$$

where the second inequality follows by Holder's inequality and the strong Markov property.

Let us first deal with the first term in the r.h.s. of (3.9). By condition **(C2)**, for any  $\varepsilon > 0$ , there is some compact subset  $A_\varepsilon$  in  $\mathbb{D}([0, t], \mathcal{S})$  such that

$$\sup_{y \in K} \mathbb{P}_y(X_{[0, t]} \notin A_\varepsilon^c) < \varepsilon.$$

By the well known property of the Skorokhod topology ([30]), the set

$$B_\varepsilon := \bigcup_{s \in [0, t]} \{\gamma(s); \gamma \in A_\varepsilon\}$$

is relatively compact in  $\mathcal{S}$ . Thus

$$\begin{aligned} \inf_{K' \subset \subset \mathcal{S}} \sup_{y \in K} \mathbb{P}_y(\exists s \in [0, t], X_t \notin K') &\leq \sup_{y \in K} \mathbb{P}_y(\exists s \in [0, t], X_t \notin \overline{B_\varepsilon}) \\ &\leq \sup_{y \in K} \mathbb{P}_y(X_{[0, t]} \notin A_\varepsilon^c) < \varepsilon. \end{aligned}$$

As  $\varepsilon > 0$  is arbitrary,  $\inf_{K' \subset \subset \mathcal{S}} \sup_{y \in K} \mathbb{P}_y(\exists s \in [0, t], X_t \notin K') = 0$ . Substituting it into (3.9), we see from (3.8) that it remains to show that

$$\frac{1}{\mathbf{W}(\mathbf{x})} \mathbb{E}_x[\mathbf{W}(X_t) \mathbf{1}_{\tau_K > t}] \leq e^{-rt}, \quad \forall \mathbf{x} \in \mathcal{S}. \quad (3.10)$$

This is the purpose of the next step.

**Step 3.** Let us deal with (3.10). To this end, introduce for  $t \geq 0$ ,

$$M_t := \frac{\mathbf{W}(X_t)}{\mathbf{W}(X_0)} \exp\left(-\int_0^t \frac{\mathcal{L}\mathbf{W}}{\mathbf{W}}(X_s) ds\right).$$

The key ingredient is the fact that  $(M_t, t \geq 0)$  is a local  $\mathbb{P}_x$ -martingale (for every  $\mathbf{x}$ ), by Ito's formula. Thus,  $(M_t, t \geq 0)$  is then a supermartingale by Fatou's lemma. Then, by the Lyapunov condition (3.4),

$$e^{rt} \frac{1}{\mathbf{W}(\mathbf{x})} \mathbb{E}_x[\mathbf{W}(X_t) \mathbf{1}_{\tau_K > t}] \leq \mathbb{E}_x[M_t] \leq M_0 = 1.$$

This is (3.10). Therefore, the proof of Theorem 3.5 is completed.  $\square$

**Corollary 3.6.** *Assume that **(C1)**, **(C2)**, and **(C3)** are satisfied. If (3.5) holds, then, for each  $t > 2t_0$  (see **(C1)**),  $P_t^{\mathcal{D}} : b_{\mathbf{W}}\mathcal{B}(\mathcal{D}) \rightarrow b_{\mathbf{W}}\mathcal{B}(\mathcal{D})$  is compact.*

*Proof.* From the first step of the proof of Theorem 3.5,  $P_{t,W}$  satisfies **(A1)** with  $N$  such that  $Nt \geq t_0$ . For  $t \geq 0$ , one has

$$P_{t,W}^{\mathcal{D}} = \frac{W(y)}{W(x)} P_t^{\mathcal{D}}(x, dy) \leq \frac{W(y)}{W(x)} P_t(x, dy) = P_{t,W}(x, dy).$$

Let us show that for  $t \geq t_0$ ,  $P_{t,W}$  is strongly Feller like  $P_t$ , i.e. for any  $f \in b\mathcal{B}(\mathcal{S})$ ,  $P_t(Wf)$  is continuous. Let us consider, for any  $n \geq 1$ , let

$$f_n := \frac{W \wedge n}{W} f.$$

Since  $Wf_n$  is bounded,  $P_t(Wf_n)$  is continuous by the strong Feller property of  $P_t$ . Now for any compact  $K \subset\subset \mathcal{S}$ , we have

$$\sup_{x \in K} |P_t(Wf)(x) - P_t(Wf_n)(x)| \leq \sup_{x \in K} [(P_t W^p)(x)]^{1/p} \sup_{x \in K} [(P_t |f_n - f|^q)(x)]^{1/q}.$$

As  $|f_n - f| \leq \|f\|_{b\mathcal{B}(\mathcal{S})}$  and  $|f_n - f| \downarrow 0$  pointwisely on  $\mathcal{S}$ , and since  $\{P_t(x, \cdot), x \in K\}$  is compact in the  $\tau$ -topology, we have

$$\sup_{x \in K} (P_t |f_n - f|^q)(x) \rightarrow 0.$$

Thus, for  $t \geq t_0$ ,  $P_t(Wf)$  is continuous.

From Theorem 3.5 (with  $r = r_n \rightarrow +\infty$  by **(C3)**), we obtain

$$\beta_w(P_{t,W}^{\mathcal{D}}) \leq \beta_w(P_{t,W}) = 0, \text{ for each } t > 0. \quad (3.11)$$

Because for each  $t \geq t_0$ ,  $P_{t,W}$  is strongly Feller on  $\mathcal{S}$ , we have for all  $K \subset\subset \mathcal{D}$ :

$$\beta_\tau(\mathbf{1}_K P_{t,W}^{\mathcal{D}}) \leq \beta_\tau(\mathbf{1}_K P_{t,W}) = 0, \text{ for each } t \geq t_0. \quad (3.12)$$

Therefore, by [67, (f) in Proposition 3.2], it holds for each  $s > 0$ :

$$\beta_\tau(P_{s+t_0,W}^{\mathcal{D}}) \leq \beta_w(P_{s,W}^{\mathcal{D}}) \beta_\tau(P_{t_0,W}^{\mathcal{D}}) = 0.$$

Finally, applying [67, (g) in Proposition 3.2],  $P_{s+2t_0,W}^{\mathcal{D}} : b\mathcal{B}(\mathcal{D}) \rightarrow b\mathcal{B}(\mathcal{D})$  is compact. This concludes the proof of Corollary 3.6.  $\square$

#### 4. A PERRON-FROBENIUS TYPE THEOREM ON $b_W\mathcal{B}$ : A REFINEMENT OF KREIN-RUTMAN'S THEOREM

We now present a version of Perron-Frobenius' theorem we will need for Feller kernels on  $b_W\mathcal{B}(\mathcal{S})$  or  $\mathcal{C}_{bW}(\mathcal{S})$ , which generalizes the Krein-Rutman theorem and is of independent interest.

**Theorem 4.1.** *Let  $Q = Q(x, dy)$  be a positive bounded kernel on  $\mathcal{S}$  and  $W \geq 1$  a continuous weight function on  $\mathcal{S}$ . Assume that:*

- (1) *There exists  $N_1 \geq 1$  such that  $Q^k$  is Feller for all  $k \geq N_1$ , i.e.  $Q^k f \in \mathcal{C}_b(\mathcal{S})$  if  $f \in \mathcal{C}_b(\mathcal{S})$ .*
- (2) *There exists  $N_2 \geq 1$  such that for any  $x \in \mathcal{S}$  and non-empty open subset  $O$  of  $\mathcal{S}$ ,*

$$Q^{N_2}(x, O) > 0.$$

(3) For some  $p > 1$  and constant  $C > 0$ , it holds:

$$QW^p \leq CW^p.$$

Notice that this implies that  $Q$  is well defined and bounded on  $b_W\mathcal{B}(\mathcal{S})$ .

(4)  $Q$  has a spectral gap in  $b_W\mathcal{B}(\mathcal{S})$ ,

$$r_{ess}(Q|_{b_W\mathcal{B}(\mathcal{S})}) < r_{sp}(Q|_{b_W\mathcal{B}(\mathcal{S})}). \quad (4.1)$$

Then, there exist a unique couple  $(\mu, \varphi)$  where  $\mu$  is a probability measure on  $\mathcal{S}$  with  $\mu(W) < +\infty$ ,  $\varphi \in \mathcal{C}_{b_W}(\mathcal{S})$  is positive everywhere on  $\mathcal{S}$  with  $\mu(\varphi) = 1$ , and constants  $r \in ]0, 1[$ ,  $C \geq 1$ , such that

$$\mu Q = r_{sp}(Q|_{b_W\mathcal{B}(\mathcal{S})})\mu, \quad Q\varphi = r_{sp}(Q|_{b_W\mathcal{B}(\mathcal{S})})\varphi \quad (4.2)$$

and

$$\left\| \frac{1}{r_{sp}(Q|_{b_W\mathcal{B}(\mathcal{S})})^n} Q^n f - \varphi \mu(f) \right\|_{b_W\mathcal{B}(\mathcal{S})} \leq Cr^n \|f\|_{b_W\mathcal{B}(\mathcal{S})}, \quad \forall f \in b_W\mathcal{B}(\mathcal{S}). \quad (4.3)$$

In particular

- (a) If a signed measure  $\nu$  satisfies for some  $\lambda \in \mathbb{R}$ ,  $\nu Q = \lambda \nu$  and  $\nu(\varphi) \neq 0$ , then  $\lambda = r_{sp}(Q|_{b_W\mathcal{B}(\mathcal{S})})$  and  $\nu = c\mu$  for some constant  $c$ .
- (b) If  $f \in b_W\mathcal{B}(\mathcal{S})$  satisfies  $\lambda \in \mathbb{R}$ ,  $Qf = \lambda f$  and  $\mu(f) \neq 0$ , then  $\lambda = r_{sp}(Q|_{b_W\mathcal{B}(\mathcal{S})})$  and  $f = c\varphi$  for some constant  $c$ .

Let us start with the following lemma.

**Lemma 4.2.** Let  $Q$  be a bounded (resp. and Feller) kernel with  $r_{sp}(Q|_{b\mathcal{B}(\mathcal{S})}) = 1$ . Then,

(a) For any  $\lambda$  in the resolvent set  $\rho(Q|_{b\mathcal{B}(\mathcal{S})})$  with  $|\lambda| > r_{ess}(Q|_{b\mathcal{B}(\mathcal{S})})$ ,

$$R(\lambda) := (\lambda I - Q)^{-1}$$

is a bounded (resp. and Feller) kernel.

(b) If  $Q$  is Feller,  $r_{sp}(Q|_{\mathcal{C}_b(\mathcal{S})}) = r_{sp}(Q|_{b\mathcal{B}(\mathcal{S})})$ , and  $r_{ess}(Q|_{\mathcal{C}_b(\mathcal{S})}) = r_{ess}(Q|_{b\mathcal{B}(\mathcal{S})})$ .

*Proof.* **(a).** At first for  $\lambda > r_{sp}(Q|_{b\mathcal{B}(\mathcal{S})})$ ,

$$R(\lambda) = \sum_{n=0}^{+\infty} \frac{1}{\lambda^{n+1}} Q^n$$

is a bounded (resp. Feller) kernel.

Now for any  $\lambda \in \rho(Q|_{b\mathcal{B}(\mathcal{S})})$  with  $|\lambda| > r_{ess}(Q|_{b\mathcal{B}(\mathcal{S})})$ , there is a  $\mathcal{C}^1$ -curve  $t \in [0, 1] \mapsto \gamma(t) \in \mathbb{C}$  such that  $\lambda = \gamma(1)$  and  $\gamma(0) > r_{sp}(Q|_{b\mathcal{B}(\mathcal{S})})$  and such that  $\text{Ran}(\gamma) \subset \rho(Q|_{b\mathcal{B}(\mathcal{S})})$ . It is enough to show that there is some (common)  $\delta > 0$  such that for any  $t_0 \in [0, 1]$  such that  $R(\gamma(t_0))$  is a bounded (resp. Feller) kernel, so is  $R(\gamma(t))$  once if  $|t - t_0| < \delta$ .

To this end, let  $M = \sup_{t \in [0, 1]} \|R(\gamma(t))\|_{b\mathcal{B}(\mathcal{S})}$  which is finite (where  $\|R(\gamma(t))\|_{b\mathcal{B}(\mathcal{S})}$  is the operator norm of  $R(\gamma(t))$  on  $b\mathcal{B}(\mathcal{S})$ ). Let  $t \in [0, 1]$  such that  $|t - t_0| \leq \frac{1}{2M(|\gamma'|_{L^\infty} + 1)}$ , so that  $|\gamma(t_0) - \gamma(t)| \leq \frac{1}{2M}$ . Then, for such  $t$ , we have

$$R(\gamma(t)) = \sum_{n=0}^{+\infty} (\gamma(t_0) - \gamma(t))^n R(\gamma(t_0))^{n+1}$$

Thus  $R(\gamma(t))$  is a bounded (resp. and Feller) kernel.

(b). The fact that  $r_{sp}(Q|_{\mathcal{C}_b(\mathcal{S})}) = r_{sp}(Q|_{b\mathcal{B}(\mathcal{S})})$  follows by Gelfand's formula for the spectral radius and the fact that  $\|Q^n\|_{b\mathcal{B}(\mathcal{S})} = \sup_{x \in \mathcal{S}} Q^n(x, \mathcal{S}) = \|Q^n\|_{\mathcal{C}_b(\mathcal{S})}$  (for all  $n \geq 0$ ). The fact that the essential spectral radius of  $Q$  on  $b\mathcal{B}(\mathcal{S})$  and on  $\mathcal{C}_b(\mathcal{S})$  is the same is proved in [67, Proposition 4.7].  $\square$

*Proof of Theorem 4.1.* The proof of Theorem 4.1 is divided into several steps.

**Step 1:** reduction to  $W = 1$ .

Let us consider, for  $x \in \mathcal{S}$

$$Q_W(x, dy) := \frac{W(y)}{W(x)} Q(x, dy).$$

Since by Holder's inequality (see also item (3) in Theorem 4.1), we have for  $x \in \mathcal{S}$ ,

$$QW(x) \leq [Q1(x)]^{1/q} [QW^p(x)]^{1/p} \leq \|Q1\|_{b\mathcal{B}(\mathcal{S})}^{1/q} C^{1/p} W(x)$$

where  $q = p/(p-1)$ , we obtain that  $Q_W 1 \leq \|Q1\|_{b\mathcal{B}(\mathcal{S})}^{1/q} C^{1/p}$ , i.e.  $Q_W$  is a bounded positive kernel on  $\mathcal{S}$ .

Let us prove that  $Q_W^k$  is again Feller for  $k \geq N_1$ , that is, for any  $f \in \mathcal{C}_b(\mathcal{S})$ ,  $Q^k(Wf)$  is continuous (notice that  $Wf$  is continuous over  $\mathcal{S}$  but not necessarily bounded on  $\mathcal{S}$ ). To this end, let us introduce for any  $n \geq 1$  and  $f \in \mathcal{C}_b(\mathcal{S})$ ,

$$f_n := \frac{W \wedge n}{W} f.$$

The function  $Q^k(Wf_n)$  is continuous by the Feller property of  $Q^k$ . Now for any compact  $K \subset\subset \mathcal{S}$ , we have

$$\sup_{x \in K} |Q^k(Wf)(x) - Q^k(Wf_n)(x)| \leq \sup_{x \in K} [(Q^k W^p)(x)]^{1/p} \sup_{x \in K} [(Q^k |f_n - f|^q)(x)]^{1/q}.$$

By assumption (3) in Theorem 4.1,  $\sup_{x \in K} [(Q^k W^p)(x)]^{1/p} \leq C^{k/p} \sup_{x \in K} W(x)$ . Since  $|f_n - f| \leq \|f\|_{b\mathcal{B}(\mathcal{S})}$  and  $f_n \rightarrow f$  uniformly over compacts on  $\mathcal{S}$ , by the tightness of  $\{Q^k(x, dy); x \in K\}$ , we have

$$\sup_{x \in K} (Q^k |f_n - f|^q)(x) \rightarrow 0.$$

Thus  $Q^k(Wf)$  is continuous.

Finally letting  $M_W f = Wf$  which is an isomorphism from  $b\mathcal{B}(\mathcal{S})$  to  $b_W \mathcal{B}(\mathcal{S})$ , we have  $Q_W = M_W^{-1} Q M_W$ , i.e.  $Q|_{b_W \mathcal{B}(\mathcal{S})}$  is similar to  $Q_W|_{b\mathcal{B}(\mathcal{S})}$ . Hence for this theorem, it is enough to prove it for  $Q_W$  on  $b\mathcal{B}(\mathcal{S})$ .

From now on, we assume without loss of generality that  $W = 1$  and  $r_{sp}(Q|_{b\mathcal{B}(\mathcal{S})}) = 1$  (otherwise consider  $Q/r_{sp}(Q|_{b\mathcal{B}(\mathcal{S})})$ ).

**Step 2:** existence of positive eigenfunction and eigen probability measure.

The fact that  $r_{sp}(Q|_{b\mathcal{B}(\mathcal{S})}) (= 1$  by assumption) is in the spectrum of  $Q|_{b\mathcal{B}(\mathcal{S})}$  is well known (see [57, Chap.V, Proposition 4.1]). In addition, by condition (4), 1 is isolated in the spectrum of  $Q|_{b\mathcal{B}(\mathcal{S})}$  and it is a pole (of finite order) of the resolvent of  $Q|_{b\mathcal{B}(\mathcal{S})}$ .

Finally by [57, Chap. V, Theorem 5.5] (cyclic property of the peripheral spectrum), there exists  $m \geq \max\{N_1, N_2\}$  such that

$$\text{for any } \lambda \in \sigma(Q|_{b\mathcal{B}(\mathcal{S})}) \text{ with } |\lambda| = 1, \lambda^m = 1. \quad (4.4)$$

For such a  $m$ , we have:

- (1)  $r_{ess}(Q^m|_{b\mathcal{B}(\mathcal{S})}) < r_{sp}(Q^m|_{b\mathcal{B}(\mathcal{S})}) = 1$  (which follows from the fact that for all  $k \geq 1$ ,  $r_{sp}(Q^k|_{b\mathcal{B}(\mathcal{S})})^{1/k} = r_{sp}(Q|_{b\mathcal{B}(\mathcal{S})}) = 1$  and  $r_{ess}(Q^k|_{b\mathcal{B}(\mathcal{S})})^{1/k} = r_{ess}(Q|_{b\mathcal{B}(\mathcal{S})}) < 1$ ).
- (2)  $1 = r_{sp}(Q|_{b\mathcal{B}(\mathcal{S})}^m) \in \sigma(Q|_{b\mathcal{B}(\mathcal{S})}^m)$ . In particular 1 is an isolated eigenvalue of  $Q^m$  and is a pole of the resolvent of  $Q^m|_{b\mathcal{B}(\mathcal{S})}$ .
- (3) the peripheral spectrum of  $Q^m|_{b\mathcal{B}(\mathcal{S})}$  is reduced to  $\{1\}$  (by (4.4) and the fact that  $\sigma(Q^m) = \sigma(Q)^m$ ), that is:

$$\{\lambda \in \mathbb{C}, |\lambda| = 1\} \cap \sigma(Q^m|_{b\mathcal{B}(\mathcal{S})}) = \{1\}.$$

Let  $\Gamma := \{\lambda \in \mathbb{C}; |\lambda - 1| = \delta\}$  where  $\delta > 0$  is such that

$$\{\lambda \in \mathbb{C}; 0 < |\lambda - 1| \leq \delta\} \subset \rho(Q^m|_{b\mathcal{B}(\mathcal{S})}) \cap \{\lambda, |\lambda| > r_{ess}(Q^m|_{b\mathcal{B}(\mathcal{S})})\}. \quad (4.5)$$

Let us denote by

$$\Pi = \frac{1}{2\pi i} \int_{\Gamma} (\lambda I - Q^m|_{b\mathcal{B}(\mathcal{S})})^{-1} d\lambda. \quad (4.6)$$

Using the Riesz decomposition theorem,  $Q^m|_{b\mathcal{B}(\mathcal{S})} = Q^m|_{b\mathcal{B}(\mathcal{S})}\Pi + Q^m|_{b\mathcal{B}(\mathcal{S})}(I - \Pi)$  where  $I - \Pi$  is the Riesz projector associated with the spectrum of  $Q^m|_{b\mathcal{B}(\mathcal{S})}$  in  $\{\lambda \in \mathbb{C}, |\lambda| < 1\}$ ,  $\Pi(I - \Pi) = (I - \Pi)\Pi = 0$ ,  $Q^m|_{b\mathcal{B}(\mathcal{S})}$  commutes with both  $\Pi$  and  $(I - \Pi)$ , and  $r_{sp}(Q^m|_{b\mathcal{B}(\mathcal{S})}(I - \Pi)) < 1$ , i.e:

$$r_{sp}(Q^m|_{b\mathcal{B}(\mathcal{S})} - Q^m|_{b\mathcal{B}(\mathcal{S})}\Pi|_{b\mathcal{B}(\mathcal{S})}) < 1. \quad (4.7)$$

Notice that condition (2) still holds for all  $k > N_2$ . Indeed condition (2) implies that  $Q(\mathbf{x}, \cdot)$  is positive measure for every  $\mathbf{x} \in \mathcal{S}$  (otherwise, if  $Q(\mathbf{x}_1, \mathcal{S}) = 0$  for some  $\mathbf{x}_1$ , then,  $Q^{N_2}(\mathbf{x}_1, \mathcal{S}) = 0$ , which contradicts condition (2) in Theorem 4.1). Consequently for any non-empty open subset  $O$  of  $\mathcal{S}$  and for any  $\mathbf{x} \in \mathcal{S}$ ,

$$Q^k(\mathbf{x}, O) = \int_{\mathcal{S}} Q^{N_2}(\mathbf{y}, O) Q^{k-N_2}(\mathbf{x}, d\mathbf{y}) > 0.$$

By applying [50, Theorem 4.1.4 and its note]<sup>3</sup> to  $Q^m|_{\mathcal{C}_b(\mathcal{S})}$  there are some nonnegative  $\varphi \in \mathcal{C}_b(\mathcal{S})$  (with  $\varphi \neq 0$ ) and some nonnegative  $\psi \in (\mathcal{C}_b(\mathcal{S}))^*$  (with  $\psi \neq 0$ ) such that

$$Q^m \varphi = \varphi \text{ and } (Q^m)^* \psi = \psi.$$

<sup>3</sup>Because  $1 = r_{sp}(Q^m|_{\mathcal{C}_b(\mathcal{S})})$  is a pole of the resolvent of  $Q^m|_{\mathcal{C}_b(\mathcal{S})}$ . Indeed, by (b) in Lemma 4.2,  $1 = r_{sp}(Q^m|_{b\mathcal{B}(\mathcal{S})}) = r_{sp}(Q^m|_{\mathcal{C}_b(\mathcal{S})})$ . In addition,  $r_{sp}(Q^m|_{\mathcal{C}_b(\mathcal{S})}) \in \sigma(Q^m|_{\mathcal{C}_b(\mathcal{S})})$  (see [57, Chap.V, Proposition 4.1]). Finally, 1 is a pole of the resolvent of  $Q^m|_{\mathcal{C}_b(\mathcal{S})}$  because  $r_{ess}(Q^m|_{\mathcal{C}_b(\mathcal{S})}) = r_{ess}(Q^m|_{b\mathcal{B}(\mathcal{S})}) < r_{sp}(Q^m|_{b\mathcal{B}(\mathcal{S})}) = r_{sp}(Q^m|_{\mathcal{C}_b(\mathcal{S})}) = 1$ .

By [67, Proposition 4.3],  $\psi$  is a positive bounded measure  $\mu$  on  $\mathcal{S}$ . We may assume that  $\mu$  is a probability measure. We claim that  $\mu$  charges all non-empty open subsets  $O$  of  $\mathcal{S}$ . Indeed, as  $\mu Q^m = (Q^m)^* \mu = \mu$ ,

$$\mu(O) = \int_{\mathcal{S}} Q^m(\mathbf{x}, O) \mu(d\mathbf{x}) > 0$$

since  $Q^m(\mathbf{x}, O) > 0$  everywhere on  $\mathcal{S}$ , proved before.

In the same way, for any  $\mathbf{x} \in \mathcal{S}$ , since  $\varphi \neq 0$  is continuous, one has

$$\varphi(\mathbf{x}) = Q^m \varphi(\mathbf{x}) = \int_{\mathcal{S}} \varphi(\mathbf{y}) Q^m(\mathbf{x}, d\mathbf{y}) > 0,$$

i.e.  $\varphi$  is everywhere positive on  $\mathcal{S}$ .

By considering  $\varphi/\mu(\varphi)$  if necessary, we may assume without loss of generality that  $\mu(\varphi) = 1$ .

**Step 3:** proof that the eigenspace  $\text{Ker}(I - Q^m) \cap \mathcal{C}_b(\mathcal{S})$  is one-dimensional, i.e., generated by  $\varphi$ .

Let  $f \in \text{Ker}(I - Q^m) \cap \mathcal{C}_b(\mathcal{S})$ , i.e.  $f \in \mathcal{C}_b(\mathcal{S})$  and  $Q^m f = f$ . Then  $Q^m |f| \geq |f|$ . Since

$$\mu(Q^m |f|) = \mu(|f|),$$

the function  $Q^m |f| - |f|$  is non negative and continuous over  $\mathcal{S}$ , and  $\mu$  charges all non-empty open subsets of  $\mathcal{S}$ , one deduces that  $Q^m |f| = |f|$  everywhere on  $\mathcal{S}$ . In other words  $|f| \in \text{Ker}(I - Q^m) \cap \mathcal{C}_b(\mathcal{S})$ , that is  $\text{Ker}(I - Q^m) \cap \mathcal{C}_b(\mathcal{S})$  is a lattice.

If  $0 \neq f \in \text{Ker}(I - Q^m) \cap \mathcal{C}_b(\mathcal{S})$  were linearly independent of  $\varphi$ , as  $f/\varphi$  is not constant, we can find  $c \in \mathbb{R}$  such that the open sets

$$O_+ = \{f > c\varphi\} \quad \text{and} \quad O_- := \{f < c\varphi\}$$

are both non-empty. Since  $(f - c\varphi)^+ \in \text{Ker}(I - Q^m)$ , we obtain for  $\mathbf{x} \in O_-$ ,

$$0 = (f - c\varphi)^+(\mathbf{x}) = \int_{\mathcal{S}} (f - c\varphi)^+(\mathbf{y}) Q^m(\mathbf{x}, d\mathbf{y}) > 0.$$

This contradiction shows that  $\text{Ker}(I - Q^m) \cap \mathcal{C}_b(\mathcal{S})$  is generated by  $\varphi$ .

**Step 4:** proof that the algebraic multiplicity and the geometric multiplicity of 1 of  $Q^m|_{b\mathcal{B}(\mathcal{S})}$  coincide.

To this end consider the Laurent series of  $(\lambda I - Q^m|_{b\mathcal{B}(\mathcal{S})})^{-1}$  in a neighborhood of 1 in  $\mathbb{C}$ ,

$$(\lambda I - Q^m|_{b\mathcal{B}(\mathcal{S})})^{-1} = A_{-l}(\lambda - 1)^{-l} + \dots + A_{-1}(\lambda - 1)^{-1} + \sum_{k=0}^{\infty} A_k(\lambda - 1)^k$$

where (see (4.6))

$$A_{-1} = \Pi$$

and  $A_{-k-1} = (Q^m|_{b\mathcal{B}(\mathcal{S})} - I)^k \Pi$  ([68], Chap. VIII, §8). Notice that  $\Pi$  is a Feller kernel by Lemma 4.2 and its definition, and thus, so are  $A_{-2}, \dots, A_{-l}$ .



We must prove that  $l = 1$ . For any bounded and measurable function  $f$  over  $\mathcal{S}$  such that  $|f| \leq c\varphi$  for some  $c > 0$ , we have for any  $\lambda > 1$ ,

$$\begin{aligned} |(\lambda - 1)(\lambda I - Q^m|_{b\mathcal{B}(\mathcal{S})})^{-1}f| &= \left| (\lambda - 1) \sum_{n=0}^{+\infty} \frac{1}{\lambda^{n+1}} Q^{mn} f \right| \\ &\leq (\lambda - 1) \sum_{n=0}^{+\infty} \frac{1}{\lambda^{n+1}} Q^{mn} |f| \\ &\leq c(\lambda - 1) \sum_{n=0}^{+\infty} \frac{1}{\lambda^{n+1}} Q^{mn} \varphi = c\varphi \end{aligned}$$

i.e.  $\{(\lambda - 1)(\lambda I - Q^m|_{b\mathcal{B}(\mathcal{S})})^{-1}f, \lambda > 1\}$  is uniformly bounded. Letting  $\lambda \rightarrow 1^+$ , we obtain  $A_{-k}f = 0$  for any  $k \geq 2$ . Because  $A_{-k}$  is a Feller kernel and  $A_{-k}f(x) = 0$  for all  $x \in \mathcal{S}$  and such  $f$ , then  $A_{-k} = 0$  for all  $k \geq 2$ .

**Step 5:** proof of (4.3).

By (b) in Lemma 4.2 and (4.7),

$$r_{sp}((Q^m - Q^m\Pi)|_{\mathcal{C}_b(\mathcal{S})}) < 1.$$

Since  $Q^m\Pi|_{\mathcal{C}_b(\mathcal{S})} = \Pi|_{\mathcal{C}_b(\mathcal{S})}$  by Step 4 (because  $A_{-2} = 0$  implies  $Q^m\Pi|_{b\mathcal{B}(\mathcal{S})} = \Pi|_{b\mathcal{B}(\mathcal{S})}$ ), one has  $r_{sp}((Q^m - \Pi)|_{\mathcal{C}_b(\mathcal{S})}) < 1$  and for all  $n \geq 1$ ,  $Q^{mn}|_{\mathcal{C}_b(\mathcal{S})} - \Pi|_{\mathcal{C}_b(\mathcal{S})} = (Q^m - \Pi)^n|_{\mathcal{C}_b(\mathcal{S})}$ . Therefore, by Gelfand's formula, there exist  $C \geq 1$  and  $r \in ]0, 1[$  such that

$$\|Q^{mn} - \Pi\|_{\mathcal{C}_b(\mathcal{S})} = \|(Q^m - \Pi)^n\|_{\mathcal{C}_b(\mathcal{S})} \leq Cr^n, \quad \forall n \geq 1.$$

Then  $\Pi$  is a nonnegative (Feller) kernel and  $\mu\Pi = \mu$ .

As  $\Pi|_{\mathcal{C}_b(\mathcal{S})}$  is a one-dimensional projection to  $\{c\varphi; c \in \mathbb{R}\}$  by Step 3 and Step 4, there is some  $\psi \in (\mathcal{C}_b(\mathcal{S}))^*$  such that for any  $f \in \mathcal{C}_b(\mathcal{S})$ ,

$$\Pi f = \psi(f)\varphi.$$

Integrating it w.r.t.  $\mu$  and since  $\mu(\varphi) = 1$ , we obtain  $\mu(f) = \psi(f)$ , i.e.  $\psi = \mu$ . In other words,

$$\Pi(x, dy) = \varphi(x)\mu(dy).$$

For the passage to  $b\mathcal{B}(\mathcal{S})$ , we note that for any  $f \in b\mathcal{B}(\mathcal{S})$  with  $|f| \leq 1$ ,

$$\sup_{x \in \mathcal{S}} |Q^{mn}f(x) - \varphi(x)\mu(f)| \leq \sup_{x \in \mathcal{S}} \sup_{g \in \mathcal{C}_b(\mathcal{S}); |g| \leq 1} |Q^{mn}g(x) - \varphi(x)\mu(g)| \leq Cr^n. \quad (4.8)$$

This implies in particular that  $\text{Ker}(I - Q^m|_{b\mathcal{B}(\mathcal{S})}) = \{c\varphi; c \in \mathbb{R}\}$  and  $\text{Ker}(I - (Q^m)^*|_{M_b(\mathcal{S})}) = \{c\mu; c \in \mathbb{R}\}$ .

Now for any eigenfunction  $f$  of  $Q$  in  $b\mathcal{B}(\mathcal{S})$  associated with 1, we have  $Q^m f = f$ , then  $f = c\varphi$ . Thus  $Q\varphi = \varphi$ . Thus  $Q\Pi = \Pi$  on  $b\mathcal{B}(\mathcal{S})$ .

Finally the desired geometric convergence (4.3) follows by (4.8), because for  $0 \leq k \leq m - 1$ ,

$$\|Q^{mn+k} - \Pi\|_{b\mathcal{B}(\mathcal{S})} = \|Q^k(Q^{mn} - \Pi)\|_{b\mathcal{B}(\mathcal{S})} \leq \max_{k \leq m-1} \|Q^k\|_{b\mathcal{B}(\mathcal{S})} \cdot \|(Q^{mn} - \Pi)\|_{b\mathcal{B}(\mathcal{S})}.$$

**Step 6:** Proofs of (a) and (b).

By (4.3), if  $\nu Q = \lambda \nu$  and  $\nu(\varphi) \neq 0$ , we have

$$\|\nu Q^n - \nu(\varphi)\mu\|_{TV} = \|\lambda^n \nu - \nu(\varphi)\mu\|_{TV} \rightarrow 0$$

as  $n \rightarrow +\infty$ . As  $\nu(\varphi) \neq 0$ ,  $\lambda = 1$  and  $\nu = \nu(\varphi) \cdot \mu$ . That is part (a). In the same way we get part (b). This concludes the proof of Theorem 4.1.  $\square$

## 5. PROOF OF THEOREM 2.2

In this section, one proves Theorem 2.2 (see Section 5.2). We first start with preliminary results.

**5.1. Preliminary results.** Let us start with the following proposition.

**Proposition 5.1.** *If  $a < f \in \mathbb{D}_e(\mathcal{L})$  where  $a \in [-\infty, +\infty[$ , then for any  $\mathcal{C}^2$ -concave function  $\varphi : ]a, +\infty[ \rightarrow \mathbb{R}$ ,  $\varphi \circ f \in \mathbb{D}_e(\mathcal{L})$  and*

$$\mathcal{L}\varphi \circ f \leq \varphi'(f)\mathcal{L}f. \quad (5.1)$$

*Proof.* For  $t \geq 0$ , let

$$M_t = f(X_t) - f(X_0) - \int_0^t \mathcal{L}f(X_s)ds$$

which is a local martingale. By Ito's formula (Dellacherie-Meyer [23, p350, Théorème 27]),  $\varphi \circ f \in \mathbb{D}_e(\mathcal{L})$  and

$$d\varphi \circ f(X_t) = \varphi'(f)(X_{t-})[\mathcal{L}f(X_{t-})dt + dM_t] + \frac{1}{2}\varphi''(X_{t-})d[M_c, M_c]_t + dS_t$$

where  $M_c$  is the continuous martingale part of  $M$ , and

$$S_t = \sum_{0 < s \leq t} (\varphi \circ f(X_s) - \varphi \circ f(X_{s-}) - \varphi'(f)(X_{s-})[f(X_s) - f(X_{s-})])$$

By the concavity of  $\varphi$ ,  $d\varphi \circ f(X_t) \leq \varphi'(f)(X_{t-})[\mathcal{L}f(X_t)dt + dM_t]$ . Thus (5.1) holds.  $\square$

The following lemma establishes the strong Feller property of  $P_t^{\mathcal{D}}$  on  $\mathcal{D}$  for  $t \geq t_0$ .

**Lemma 5.2.** *Under (C1) and (C4),  $P_t^{\mathcal{D}}$  is strong Feller on  $\mathcal{D}$  for all  $t \geq t_0$ .*

*Proof.* Let  $(x_n)_{n \geq 0}$  be a sequence of points in  $\mathcal{D}$  converging to  $\mathbf{x} \in \mathcal{D}$ . Let us prove that  $P_t^{\mathcal{D}}f(x_n) \rightarrow P_t^{\mathcal{D}}f(\mathbf{x})$  for any  $f \in b\mathcal{B}(\mathcal{D})$  and  $t \geq t_0$  fixed. Let  $K = \{\mathbf{x}, x_n; n \geq 0\}$ . One has:

$$\beta_\tau(\mathbf{1}_K P_t^{\mathcal{D}}) \leq \beta_\tau(\mathbf{1}_K P_t) = 0$$

by the strong Feller property of  $(P_t, t \geq 0)$  on  $\mathcal{S}$ , see (C1). In other words the family  $\{P_t^{\mathcal{D}}(x_n, \cdot); n \geq 0\}$  is relatively compact in the  $\tau$ -topology. That is equivalent to say that

$$\left\{ h_n := \frac{dP_t^{\mathcal{D}}(x_n, \cdot)}{dm}; n \geq 0 \right\}$$

is relatively compact in the weak topology  $\sigma(L^1(m), L^\infty(m))$ , where  $m$  is the reference measure given by

$$m(\cdot) = P_t^{\mathcal{D}}(\mathbf{x}, \cdot) + \sum_{n=0}^{+\infty} \frac{1}{2^n} P_t^{\mathcal{D}}(x_n, \cdot).$$

By the well known equivalence of the relative compactness and the sequential compactness in  $\sigma(L^1, L^\infty)$  (by the Dunford-Pettis theorem [22, Théorème 25 p. 43]), we have only to prove that the limit point in the  $\tau$ -topology of  $P_t^{\mathcal{D}}(x_n, \cdot)$  is unique and coincides with  $P_t^{\mathcal{D}}(\mathbf{x}, \cdot)$ , i.e. if  $P_t^{\mathcal{D}}(x_{n_k}, \cdot) \rightarrow \nu$  in  $\tau$ -topology for a subsequence  $(n_k)$ , then  $\nu = P_t^{\mathcal{D}}(\mathbf{x}, \cdot)$ . By **(C4)**, we have for any  $f \in \mathcal{A}$ ,

$$\nu(f) = \lim_{k \rightarrow \infty} P_t^{\mathcal{D}} f(x_{n_k}) = P_t^{\mathcal{D}} f(\mathbf{x})$$

By the measure-separability of  $\mathcal{A}$ ,  $\nu = P_t^{\mathcal{D}}(\mathbf{x}, \cdot)$ .  $\square$

**5.2. Proof of the main result.** We will formulate a weaker version of Theorem 2.2.

**Theorem 5.3.** *Assume that **(C1)**, **(C2)**, **(C3)**, **(C4)**, and **(C5)** hold. Suppose moreover that for some  $p > 1$  and  $M > 0$ ,*

$$\mathcal{LW}^p \leq MW^p. \quad (5.2)$$

*Then, all claims in Theorem 2.2 hold with  $W^{1/p}$  replaced by  $W$ .*

Admitting this result, we give at first the proof of Theorem 2.2.

*Proof of Theorem 2.2.* For any  $p > 1$  fixed, by Proposition 5.1 and **(C3)**, it holds:

$$\mathcal{LW}^{1/p} \leq \frac{1}{p} W^{1/p-1} \mathcal{LW} \leq -\frac{r_n}{p} W^{1/p} - \frac{b_n}{p} \mathbf{1}_{K_n}.$$

In other words,  $\tilde{W} = W^{1/p}$  satisfies **(C3)**. Furthermore for each  $n$ ,

$$\mathcal{L}(\tilde{W})^p = \mathcal{LW} \leq b_n \mathbf{1}_{K_n} \leq b_n W = b_n (\tilde{W})^p. \quad (5.3)$$

Thus applying Theorem 5.3 to  $\tilde{W}$ , we obtain Theorem 2.2.  $\square$

Let us now prove Theorem 5.3.

*Proof of Theorem 5.3.* The proof of Theorem 5.3 is divided into several steps.

**Step 1.** Let  $t > 0$  be arbitrary but fixed. Recall that

$$P_{t,W}^{\mathcal{D}} = \frac{W(\mathbf{y})}{W(\mathbf{x})} P_t^{\mathcal{D}}(\mathbf{x}, d\mathbf{y}) \leq \frac{W(\mathbf{y})}{W(\mathbf{x})} P_t(\mathbf{x}, d\mathbf{y}) = P_{t,W}(\mathbf{x}, d\mathbf{y})$$

and that  $P_{t,W}^{\mathcal{D}}$  also satisfies **(A1)** (by (3.11) and (3.12)). Then by Theorem 3.4, and in view of (3.11)

$$r_{ess}(P_t^{\mathcal{D}}|_{b_W \mathcal{B}(\mathcal{D})}) = \lim_{n \rightarrow \infty} \left[ \beta_w(P_{nt,W}^{\mathcal{D}}) \right]^{1/n} = 0.$$

Furthermore by Lemma 5.2,  $P_t^{\mathcal{D}}$  is strong Feller on  $\mathcal{D}$  for all  $t \geq t_0$ . This together with the topological transitivity in **(C5)** implies that for any  $t > 0$ ,  $P_t^{\mathcal{D}}$  is  $m_1$ -irreducible where  $m_1 = \int_{t_0}^{+\infty} e^{-s} P_s^{\mathcal{D}}(\mathbf{x}_1, \cdot) ds$ , for some  $\mathbf{x}_1 \in \mathcal{D}$ . Indeed, let  $A \in \mathcal{B}(\mathcal{D})$  such that  $m_1(A) > 0$ . The function  $g_1(\mathbf{x}) := \int_{t_0}^{+\infty} e^{-s} P_s^{\mathcal{D}}(\mathbf{x}_1, A) ds$  is continuous (since  $P_t^{\mathcal{D}}$  is strong Feller on  $\mathcal{D}$  for all  $t \geq t_0$ ) and positive at  $\mathbf{x}_1$  (by choice of  $A$ ). Then, by **(C5)**, we have

$$P_{Nt}^{\mathcal{D}} g_1(\mathbf{x}) > 0, \quad \forall \mathbf{x} \in \mathcal{D}$$

once if  $Nt \geq t_0$ . By Nummelin [52, Theorem 3.2],

$$r_{sp}(P_t^{\mathcal{D}}|_{b_{\mathbb{W}}\mathcal{B}(\mathcal{D})}) = \lim_{n \rightarrow +\infty} \left( \sup_{x \in \mathcal{D}} \frac{\mathbb{E}_x[\mathbb{W}(X_t) \mathbf{1}_{nt < \sigma_{\mathcal{D}}}]}{\mathbb{W}(x)} \right)^{1/n} > 0.$$

**Step 2.** Let  $\lambda_{\mathcal{D}} := -\log r_{sp}(P_1^{\mathcal{D}}|_{b_{\mathbb{W}}\mathcal{B}(\mathcal{D})})$ . Applying Theorem 4.1 to  $Q = P_1^{\mathcal{D}}$  on  $b_{\mathbb{W}}\mathcal{B}(\mathcal{D})$ , there is a unique couple  $(\mu_{\mathcal{D}}, \varphi)$  where  $\mu_{\mathcal{D}}$  is a probability measure on  $\mathcal{D}$  with  $\mu_{\mathcal{D}}(\mathbb{W}) < +\infty$ ,  $\varphi \in \mathcal{C}_{b_{\mathbb{W}}}(\mathcal{S})$  positive everywhere on  $\mathcal{D}$ ,  $\mu_{\mathcal{D}}(\varphi) = 1$  and

$$\mu_{\mathcal{D}} P_1^{\mathcal{D}} = e^{-\lambda_{\mathcal{D}}} \mu_{\mathcal{D}}, \quad P_1^{\mathcal{D}} \varphi = e^{-\lambda_{\mathcal{D}}} \varphi,$$

and for all  $f \in b_{\mathbb{W}}\mathcal{B}(\mathcal{D})$  and  $n \geq 1$

$$\|e^{n\lambda_{\mathcal{D}}} P_n^{\mathcal{D}} f - \mu_{\mathcal{D}}(f)\varphi\|_{b_{\mathbb{W}}\mathcal{B}(\mathcal{D})} \leq C e^{-\delta n} \|f\|_{b_{\mathbb{W}}\mathcal{B}(\mathcal{D})}. \quad (5.4)$$

where  $C \geq 1$  and  $\delta > 0$  are independent of  $f$  and  $n$ . In addition for any  $t > 0$ , since  $(\mu_{\mathcal{D}} P_t^{\mathcal{D}}) P_1^{\mathcal{D}} = (\mu_{\mathcal{D}} P_1^{\mathcal{D}}) P_t^{\mathcal{D}} = e^{-\lambda_{\mathcal{D}}} \mu_{\mathcal{D}} P_t^{\mathcal{D}}$  and  $\text{Ker}(e^{-\lambda_{\mathcal{D}}} I - (P_1^{\mathcal{D}})^*)$  in  $M_{b_{\mathbb{W}}}(\mathcal{S})$  is one dimensional, one deduces that  $\mu_{\mathcal{D}} P_t^{\mathcal{D}} = \lambda(t) \mu_{\mathcal{D}}$ . By the semigroup property,  $\lambda(t+s) = \lambda(t) \cdot \lambda(s)$ . As  $\lambda(1) = e^{-\lambda_{\mathcal{D}}}$ , one obtains that:

$$\lambda(t) = e^{-\lambda_{\mathcal{D}} t}, \quad t \geq 0.$$

By (a) in Theorem 4.1,

$$r_{sp}(P_t^{\mathcal{D}}|_{b_{\mathbb{W}}\mathcal{B}(\mathcal{D})}) = \lambda(t) = e^{-\lambda_{\mathcal{D}} t}.$$

Since

$$\mathbb{P}_{\mu_{\mathcal{D}}}(t < \tau_{\mathcal{D}}) = \mu_{\mathcal{D}} P_t^{\mathcal{D}} \mathbf{1} = e^{-\lambda_{\mathcal{D}} t} \mu_{\mathcal{D}}(\mathbf{1}) = e^{-\lambda_{\mathcal{D}} t}$$

then  $\lambda_{\mathcal{D}} \geq 0$  and

$$\mathbb{P}_{\mu_{\mathcal{D}}}(X_t \in \cdot | t < \tau_{\mathcal{D}}) = e^{\lambda_{\mathcal{D}} t} \mu_{\mathcal{D}} P_t^{\mathcal{D}}(\cdot) = \mu_{\mathcal{D}},$$

i.e.  $\mu_{\mathcal{D}}$  is a QSD. Let us now prove the uniqueness of the QSD of  $(X_t, t \geq 0)$  in  $\mathcal{D}$  in the set of measures  $\nu$  such that  $\nu(\mathbb{W}) < +\infty$ . To this end, let us consider  $\nu_{\mathcal{D}}$  a QSD satisfying  $\nu_{\mathcal{D}}(\mathbb{W}) < +\infty$ . Then for all  $t \geq 0$

$$\nu_{\mathcal{D}} P_t^{\mathcal{D}} = \lambda_t \nu_{\mathcal{D}}, \quad \lambda_t = \mathbb{P}_{\nu_{\mathcal{D}}}(t < \sigma_{\mathcal{D}}).$$

By (a) in Theorem 4.1, this implies necessarily that

$$\lambda_t = r_{sp}(P_t^{\mathcal{D}}|_{b_{\mathbb{W}}\mathcal{B}(\mathcal{D})}) = e^{-\lambda_{\mathcal{D}} t} \text{ and } \nu_{\mathcal{D}} = \mu_{\mathcal{D}},$$

which concludes the proof of the uniqueness. Finally for any  $t = n + s$  with  $s \in [0, 1[$ , by (5.4) we have for all  $f \in b_{\mathbb{W}}\mathcal{B}(\mathcal{D})$

$$\|e^{(n+s)\lambda_{\mathcal{D}}} P_{n+s}^{\mathcal{D}} f - e^{s\lambda_{\mathcal{D}}} \mu_{\mathcal{D}}(P_s^{\mathcal{D}} f)\varphi\|_{b_{\mathbb{W}}\mathcal{B}(\mathcal{D})} \leq C r^n e^{s\lambda_{\mathcal{D}}} \|P_s^{\mathcal{D}} f\|_{b_{\mathbb{W}}\mathcal{B}(\mathcal{D})}.$$

As  $e^{s\lambda_{\mathcal{D}}} \mu_{\mathcal{D}}(P_s^{\mathcal{D}} f) = \mu_{\mathcal{D}}(f)$  and  $\sup_{s \in [0, 1]} \|P_s^{\mathcal{D}}\|_{b_{\mathbb{W}}\mathcal{B}(\mathcal{D})} \leq \sup_{s \in [0, 1]} \|P_s \mathbb{W}\|_{b_{\mathbb{W}}\mathcal{B}(\mathcal{D})} \leq e^{b_1}$  (by the proof in Step 1 of Theorem 3.5 and **(C3)**), we obtain for all  $f \in b_{\mathbb{W}}\mathcal{B}(\mathcal{D})$ ,

$$\|e^{t\lambda_{\mathcal{D}}} P_t^{\mathcal{D}} f - \mu_{\mathcal{D}}(f)\varphi\|_{b_{\mathbb{W}}\mathcal{B}(\mathcal{D})} \leq C' e^{-\delta t} \|f\|_{b_{\mathbb{W}}\mathcal{B}(\mathcal{D})}, \quad \delta := -\log r, \quad C' = C e^{b_1} e^{\lambda_{\mathcal{D}}}. \quad (5.5)$$

Thus, one has for all  $f \in b_{\mathbb{W}}\mathcal{B}(\mathcal{D})$ :

$$\begin{aligned} |\mathbb{E}_{\nu}[f(X_t)|t < \sigma_{\mathcal{D}}] - \mu_{\mathcal{D}}(f)| &= \left| \frac{e^{\lambda_{\mathcal{D}}t}\nu(P_t^{\mathcal{D}}f)}{e^{\lambda_{\mathcal{D}}t}\nu(P_t^{\mathcal{D}}\mathbf{1})} - \mu_{\mathcal{D}}(f) \right| \\ &= \left| \frac{\mu_{\mathcal{D}}(f)\nu(\varphi) + O_1(e^{-\delta t})\nu(\mathbb{W})\|f\|_{b_{\mathbb{W}}\mathcal{B}(\mathcal{D})}}{\nu(\varphi) + O_2(e^{-\delta t})\nu(\mathbb{W})\|f\|_{b_{\mathbb{W}}\mathcal{B}(\mathcal{D})}} - \mu_{\mathcal{D}}(f) \right| \\ &= \left| \frac{\mu_{\mathcal{D}}(f) + O_1(e^{-\delta t})\frac{\nu(\mathbb{W})}{\nu(\varphi)}\|f\|_{b_{\mathbb{W}}\mathcal{B}(\mathcal{D})}}{1 + O_2(e^{-\delta t})\frac{\nu(\mathbb{W})}{\nu(\varphi)}\|f\|_{b_{\mathbb{W}}\mathcal{B}(\mathcal{D})}} - \mu_{\mathcal{D}}(f) \right| \\ &\leq O_3(t)\frac{\nu(\mathbb{W})}{\nu(\varphi)}\|f\|_{b_{\mathbb{W}}\mathcal{B}(\mathcal{D})}, \end{aligned}$$

where for all  $k = 1, 2, 3$  and for all  $t \geq 0$ ,  $|O_k(t)| \leq Ce^{-\delta t}$ , for some constant  $C$  independent of  $\nu$  and  $f$ . That yields:

$$|\mathbb{E}_{\nu}[f(X_t)|t < \sigma_{\mathcal{D}}] - \mu_{\mathcal{D}}(f)| \leq Ce^{-\delta t}\frac{\nu(\mathbb{W})}{\nu(\varphi)}\|f\|_{b_{\mathbb{W}}\mathcal{B}(\mathcal{D})}, \forall f \in b_{\mathbb{W}}\mathcal{B}(\mathcal{D}), t > 0.$$

**Step 3.** We have proved that  $\lambda_{\mathcal{D}} \in [0 + \infty)$ . Let us now prove that  $\lambda_{\mathcal{D}} > 0$ . If in contrary  $\lambda_{\mathcal{D}} = 0$ , then for all  $t \geq 0$ ,  $\mu_{\mathcal{D}}(P_t^{\mathcal{D}}\mathbf{1}) = \mu_{\mathcal{D}}(\mathbf{1}) = 1$ . This implies that  $P_t^{\mathcal{D}}\mathbf{1}(\mathbf{x}) = 1$  for all  $\mathbf{x} \in \mathcal{D}$  and  $t > 0$ , due to the fact that the function  $1 - P_t^{\mathcal{D}}\mathbf{1}$  is nonnegative and continuous over  $\mathcal{D}$  (by the Feller property of  $P_t^{\mathcal{D}}$ ) and that  $\mu_{\mathcal{D}}$  charges all non-empty open subsets of  $\mathcal{D}$ . That is in contradiction with the second assumption in **(C5)**. Then  $\lambda_{\mathcal{D}} > 0$ . Now for every  $\mathbf{x} \in \mathcal{D}$ , it holds by (5.5) with  $f = \mathbf{1}$ :

$$\mathbb{P}_{\mathbf{x}}(\sigma_{\mathcal{D}} = +\infty) = \lim_{t \rightarrow +\infty} e^{-\lambda_{\mathcal{D}}t} e^{\lambda_{\mathcal{D}}t} \mathbb{P}_{\mathbf{x}}(t < \sigma_{\mathcal{D}}) = 0 \times \varphi(\mathbf{x})\mu_{\mathcal{D}}(\mathbf{1}) = 0.$$

Next  $\mathbb{P}_{\mu_{\mathcal{D}}}(t < \sigma_{\mathcal{D}}) = \mu_{\mathcal{D}}(P_t^{\mathcal{D}}\mathbf{1}) = e^{-\lambda_{\mathcal{D}}t}$ . It remains to prove the independence of  $X_{\sigma_{\mathcal{D}}}$  and  $\sigma_{\mathcal{D}}$ , under  $\mathbb{P}_{\mu_{\mathcal{D}}}$ . For any  $f \in b\mathcal{B}(\partial\mathcal{D})$ , letting for  $x \in \mathcal{D}$ ,

$$u(\mathbf{x}) = \mathbb{E}_{\mathbf{x}}[f(X_{\sigma_{\mathcal{D}}})],$$

we have by the strong Markov property

$$\mathbb{E}_{\mu_{\mathcal{D}}}[f(X_{\sigma_{\mathcal{D}}})\mathbf{1}_{t < \sigma_{\mathcal{D}}}] = \mu_{\mathcal{D}}(P_t^{\mathcal{D}}u) = e^{-\lambda_{\mathcal{D}}t}\mu_{\mathcal{D}}(u) = e^{-\lambda_{\mathcal{D}}t}\mathbb{E}_{\mu_{\mathcal{D}}}[f(X_{\sigma_{\mathcal{D}}})], t \geq 0,$$

which is the desired independence. This concludes the proof of Theorem 5.3.  $\square$

## 6. APPLICATION TO HYPOELLIPTIC DAMPED HAMILTONIAN SYSTEMS

In this section, we apply Theorem 2.2 to hypoelliptic damped Hamiltonian systems on  $\mathbb{R}^{2d}$  (see (6.1)) when  $\mathcal{D} = \mathbf{O} \times \mathbb{R}^d$  ( $\mathbf{O} \subset \mathbb{R}^d$ , see more precisely (6.25)). To this end, we first define the setting we consider and then, we check that the assumptions required to apply Theorem 2.2 are satisfied for such processes (namely **(C1)**-**(C5)**): these are the purposes of Sections 6.1 and 6.2 respectively. Finally, in Section 6.3, we state the main result of this section which is Theorem 6.8.

**6.1. Framework and assumptions.** Let  $d \geq 1$ . Let  $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t \geq 0}, \mathbb{P})$  be a filtered probability space. Let  $(X_t = (x_t, v_t), t \geq 0)$  be the solution of the following hypoelliptic stochastic differential equation on  $\mathbb{R}^{2d}$ :

$$\begin{cases} dx_t = v_t dt, \\ dv_t = -\nabla V(x_t) dt - c(x_t, v_t) v_t dt + \Sigma(x_t, v_t) dB_t, \end{cases} \quad (6.1)$$

where  $(B_t, t \geq 0)$  is a standard  $d$ -dimensional Brownian motion on  $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t \geq 0}, \mathbb{P})$ . Here the phase space is  $\mathcal{S} = \mathbb{R}^{2d}$ . Equation (6.1) describes a system of  $\mathbf{N}$  particles (in this case  $d = 3\mathbf{N}$ ) moving under interaction forces which are subject to random collisions. The function  $c$  is the damping (or friction) coefficient and  $V$  is the particle interaction potential function. We refer for instance to [62, 66, 48], and to the review of the literature [46] for the study of such processes in  $\mathbb{R}^{2d}$ . Let us define the following assumptions on  $V$  and  $c$ :

**(Av1)**  $V : \mathbb{R}^d \rightarrow \mathbb{R}$  is  $\mathcal{C}^1$  and  $V$  is lower bounded on  $\mathbb{R}^d$ .

**(Ac1)**  $c : \mathbb{R}^{2d} \rightarrow \mathbb{R}^{d \times d}$  is continuous. In addition, there exist  $\eta > 0$  and  $L > 0$ , such that

$$\forall v \in \mathbb{R}^d, |x| \geq L : \frac{1}{2} [c(x, v) + c^T(x, v)] \geq \eta I_{\mathbb{R}^d}.$$

Finally, for all  $N > 0$ ,

$$\sup_{|x| \leq N, v \in \mathbb{R}^d} \|c(x, v)\|_{\text{H.S}} < +\infty,$$

where  $\|c(x, v)\|_{\text{H.S}}$  is the Hilbert-Schmidt norm of matrix and where  $c^T$  is the the transpose matrix of  $c$ .

**(A $\Sigma$ )**  $\Sigma : \mathbb{R}^{2d} \rightarrow \mathbb{R}$  a  $\mathcal{C}^\infty$  function, uniformly Lipschitz over  $\mathbb{R}^{2d}$ , and such that for some  $\Sigma_0 > 0$  and  $\Sigma_\infty > 0$ ,

$$\forall x \in \mathbb{R}^{2d}, \Sigma_0 \leq \Sigma(x) \leq \Sigma_\infty.$$

For some results below, **(Ac1)** can be replaced by the less stringent assumption:

**(Ac0)**  $c : \mathbb{R}^{2d} \rightarrow \mathbb{R}^{d \times d}$  is continuous and

$$\exists A \geq 0, \forall x, v \in \mathbb{R}^d : \frac{1}{2} [c(x, v) + c^T(x, v)] \geq -A I_{\mathbb{R}^d}.$$

This will allow us to consider in particular the hypoelliptic damped Hamiltonian systems with unbounded  $v$ -dependent damping coefficient:

$$\text{for some } \ell_0 > 0, \forall x, v \in \mathbb{R}^d, c(x, v) = |v|^{\ell_0} \quad (6.2)$$

and fast growing potential, in the sense that there exist  $\mathbf{n}_0 > 2$ ,  $r_0 > 0$ , and  $r > 0$ , for all  $|x| \geq r_0$ ,

$$V \text{ satisfies } \mathbf{(Av1)}, r^{-1}|x|^{\mathbf{n}_0} \leq V(x) \leq r|x|^{\mathbf{n}_0} \text{ and } r^{-1}|x|^{\mathbf{n}_0} \leq x \cdot \nabla V(x). \quad (6.3)$$

Notice that when (6.2) holds, Assumption **(Ac1)** is not satisfied but **(Ac0)** is satisfied. The condition that  $\mathbf{n}_0 > 2$  is justified in item (2) in Proposition 6.4 below.

**Remark 6.1.** Condition (6.3) is satisfied for instance for a  $\mathcal{C}^1$  function  $V$  over  $\mathbb{R}^d$  which equal a polynomial function with leading terms  $a|x|^{2n}$ , with  $n \geq 2$  and  $a > 0$ , outside a compact set of  $\mathbb{R}^d$ .

When  $V$ ,  $c$ , and  $\Sigma$  satisfy respectively **(Av1)**, **(Ac0)**, and **(A $\Sigma$ )**, there is a unique weak solution to (6.1) by [66, Lemma 1.1]. We will thus always assume at least **(A $\Sigma$ )**, **(Av1)** and **(Ac0)** in what follows.

For  $t \geq 0$ , we recall that  $(P_t, t \geq 0)$  denotes the semigroup of the process  $(X_t, t \geq 0)$ , that is  $P_t(\mathbf{x}, A) = \mathbb{P}_{\mathbf{x}}(X_t \in A)$ , where  $A \in \mathcal{B}(\mathbb{R}^{2d})$  and  $\mathbf{x} = (x, v) \in \mathbb{R}^{2d}$ . In the following, we denote by  $(X_t^0(\mathbf{x}), t \geq 0)$  the process  $(X_t, t \geq 0)$  when  $X_0 = \mathbf{x}$ . Let us also denote by

$$\mathcal{L}_0 = \frac{\Sigma^2(x, v)}{2} \Delta_v + v \cdot \nabla_x - \nabla V(x) \cdot \nabla_v - c(x, v) v \cdot \nabla_v, \quad (6.4)$$

the infinitesimal generator of the diffusion (6.1). Let us recall that  $\mathbb{D}_e(\mathcal{L})$  denotes the extended domain of the generator of the semigroup  $(P_t, t \geq 0)$  of the process (6.1) (see (2.4) for the definition).

Let us check that the assumptions required to apply Theorem 2.2 are satisfied for the process (6.1) when  $\mathcal{D} = \mathbf{O} \times \mathbb{R}^d$  ( $\mathbf{O} \subset \mathbb{R}^d$ , see more precisely (6.25)), by prescribing if necessary, more assumptions on  $V$ ,  $c$ , and  $\ell_0$ .

## 6.2. On the assumptions **(C1)**-**(C5)**.

6.2.1. *On the assumptions **(C1)** and **(C2)**.* One has the following result from [66].

**Lemma 6.2.** *Assume that  $V$ ,  $c$ , and  $\Sigma$  satisfy respectively **(Av1)**, **(Ac0)**, and **(A $\Sigma$ )**. Then, **(C1)** and **(C2)** are satisfied for the process (6.1).*

*Proof.* Let us first prove that  $(P_t, t \geq 0)$  satisfies **(C1)**. Introduce the process  $(X_t^0 = (x_t^0, v_t^0), t \geq 0)$  solution (in the strong sense) to the stochastic differential equation over  $\mathbb{R}^{2d}$ :

$$\begin{cases} dx_t^0 = v_t^0 dt, \\ dv_t^0 = \Sigma(x_t^0, v_t^0) dB_t. \end{cases} \quad (6.5)$$

That is  $(X_t^0, t \geq 0)$  is the process (6.1) when  $V = 0$  and  $c = 0$ . Let  $(P_t^0, t \geq 0)$  be the semigroup of process (6.5). Under **(A $\Sigma$ )**, for  $t > 0$ ,  $P_t^0$  has a smooth density  $(\mathbf{x}, \mathbf{y}) \in \mathbb{R}^{2d} \mapsto p_t^0(\mathbf{x}, \mathbf{y})$  with respect to the Lebesgue measure  $d\mathbf{y}$  by the Hörmander's theorem. Therefore, when  $(\mathbf{x}_n)_n$  converges to  $\mathbf{x} \in \mathbb{R}^{2d}$  as  $n \rightarrow +\infty$ ,  $p_t^0(\mathbf{x}_n, \mathbf{y}) \rightarrow p_t^0(\mathbf{x}, \mathbf{y})$ . Since for all  $n$ ,  $\int_{\mathbb{R}^{2d}} p_t^0(\mathbf{x}_n, \mathbf{y}) d\mathbf{y} = \int_{\mathbb{R}^{2d}} p_t^0(\mathbf{x}, \mathbf{y}) d\mathbf{y} = 1$ , it follows by Scheffé's Lemma that  $p_t^0(\mathbf{x}_n, \mathbf{y}) \rightarrow p_t^0(\mathbf{x}, \mathbf{y})$  in  $L^1(\mathbb{R}^{2d}, d\mathbf{y})$  as  $n \rightarrow +\infty$ . Hence for any  $f \in b\mathcal{B}(\mathbb{R}^{2d})$ ,  $P_t^0 f(\mathbf{x}_n) = \int_{\mathbb{R}^{2d}} f(\mathbf{y}) p_t^0(\mathbf{x}_n, \mathbf{y}) d\mathbf{y} \rightarrow \int_{\mathbb{R}^{2d}} f(\mathbf{y}) p_t^0(\mathbf{x}, \mathbf{y}) d\mathbf{y} = P_t^0 f(\mathbf{x})$ . That is,  $P_t^0$  is strong Feller for  $t > 0$ .

When  $V$  and  $c$  satisfy **(Av1)** and **(Ac0)**, using the same arguments as in the proof of [66, Proposition 1.2], we deduce that, for  $t > 0$ ,  $P_t$  is strong Feller and thus satisfies **(C1)**.

Moreover, for any  $T > 0$ , the mapping

$$\mathbf{x} \in \mathbb{R}^{2d} \mapsto \mathbb{P}_{\mathbf{x}}(X_{[0, T]} \in \cdot) \in \mathcal{M}_1(\mathcal{C}^0([0, T], \mathbb{R}^{2d}))$$

is continuous with respect to the weak convergence of measures on  $\mathcal{C}^0([0, T], \mathbb{R}^{2d})$  (equipped with the uniform convergence topology). Indeed, the weak limit of the weak solutions of (6.1) with starting point  $\mathbf{x}_n$  converges to the weak solution of the martingale problem with starting point  $\mathbf{x}$ , as  $\mathbf{x}_n \rightarrow \mathbf{x} \in \mathbb{R}^{2d}$ , by the continuity of the coefficients

in (6.1) and the uniqueness of the weak solution of (6.1) (see [66, Lemma 1.1]). Thus **(C2)** is satisfied for the process (6.1).  $\square$

6.2.2. *On the assumption (C3).* Let us define the following last assumptions on  $V$  and  $c$ .

**(Av2)** There exists a  $\mathcal{C}^1$  function  $G : \mathbb{R}^d \rightarrow \mathbb{R}^d$  such that  $G$  and  $\nabla G$  are bounded over  $\mathbb{R}^d$ , and such that

$$\nabla V(x) \cdot G(x) \rightarrow +\infty \text{ as } |x| \rightarrow +\infty.$$

**(Ac2)** There exists some  $\mathcal{C}^2$  lower bounded function  $U : \mathbb{R}^d \rightarrow \mathbb{R}$  such that

$$\sup_{x,v \in \mathbb{R}^d} |c^T(x,v)G(x) - \nabla U(x)| < +\infty.$$

**Remark 6.3.** Let us recall some examples of functions  $V$  and  $c$  satisfying **(Av1)**, **(Av2)**, **(Ac1)**, and **(Ac2)** ([66, Remark 3.2]):

(1) If the damping coefficient  $c$  satisfies **(Ac1)** with moreover

$$\sup_{x,v \in \mathbb{R}^d} \|c(x,v)\|_{\text{H.S}} < +\infty,$$

then **(Ac2)** is satisfied with  $U = 0$ .

(2) Assume that  $V$  satisfies **(Av1)** and that:

(a)  $\lim_{|x| \rightarrow +\infty} \frac{x \cdot \nabla V(x)}{|x|} = +\infty$ . Then, **(Av2)** is satisfied with

$$x \mapsto G(x) = \frac{x}{|x|}(1 - \chi),$$

where  $\chi : \mathbb{R}^d \rightarrow \mathbb{R}$  is  $\mathcal{C}^\infty$ , has compact support,  $\chi = 1$  in a neighborhood of 0 in  $\mathbb{R}^d$ .

In particular, **(Av2)** and **(Ac2)** are satisfied when  $V : \mathbb{R}^d \rightarrow \mathbb{R}$  is a  $\mathcal{C}^1$  function such that  $\nabla V(x) \cdot x \geq c_0|x|^{2k}$  ( $k \in \mathbb{N}^*$ ,  $c_0 > 0$ ) outside a compact subset of  $\mathbb{R}^d$  and  $c(x,v) = c_1|x|^{2q}$  ( $q \in \mathbb{N}^*$ ,  $c_1 > 0$ ) on  $\mathbb{R}^{2d}$  (choose indeed  $G$  as above and  $U(x) = c_1(1 - \chi)|x|^{2q+1}/(2q+1)$ ).

(b) There exists  $r > 0$ ,  $|x| > r \mapsto \mathbf{e}_V(x) = \nabla V(x)/|\nabla V(x)|$  is  $\mathcal{C}^1$ , bounded, and with bounded derivatives, and  $\lim_{|x| \rightarrow +\infty} |\nabla V(x)| = +\infty$ . Then, **(Av2)** is satisfied with

$$x \mapsto G(x) = \mathbf{e}_V(x)(1 - \chi),$$

where  $\chi : \mathbb{R}^d \rightarrow \mathbb{R}$  is  $\mathcal{C}^\infty$ , has compact support, and  $\chi = 1$  on  $B(0, r+1)$ .

Notice that when  $d = 1$ ,  $\lim_{|x| \rightarrow +\infty} \frac{x \cdot \nabla V(x)}{|x|} = +\infty$ , **(Ac2)**, and  $\lim_{|x| \rightarrow +\infty} |\nabla V(x)| = +\infty$  are equivalent (under **(Av1)**).

(3) When  $d = 1$ , the case when there exist  $c_1, c_2, w_0 > 0$ , such that

$$\forall x, v \in \mathbb{R}, c(x,v) = c_1x^2 - c_2 \text{ and } V(x) = \frac{1}{2}w_0^2x^2, \quad (6.6)$$

corresponds to the the noisy Van Der Pol oscillator. Then, **(Av1)**, **(Av2)**, **(Ac1)**, and **(Ac2)** are satisfied with  $G(x) = x(1 - \chi)/|x|$ , and  $U(x) = [c_1|x|^3/3 - c_2|x|](1 - \chi)$ , where  $\chi : \mathbb{R} \rightarrow \mathbb{R}$  is  $\mathcal{C}^\infty$ , has compact support, and  $\chi = 1$  in a neighborhood of 0 (see [66, Section 5.3]).



The Hamiltonian of the process (6.1) is, for  $x, v \in \mathbb{R}^d$ ,

$$H(x, v) = V(x) + \frac{1}{2}|v|^2.$$

Assume that **(A $\Sigma$ )**, **(Av1)**, **(Av2)**, **(Ac1)**, and **(Ac2)** hold. Let us introduce for  $(x, v) \in \mathbb{R}^{2d}$ , the modified Hamiltonian [66, Eq. (3.3)],

$$F_1(x, v) = aH(x, v) + v \cdot (bG(x) + \nabla w(x)) + bU(x) \quad (6.7)$$

where  $G, U$  are as **(Av2)** and **(Ac2)**,  $a > 0, b > 0$ , and  $w : \mathbb{R}^d \rightarrow \mathbb{R}$  is a compactly supported  $\mathcal{C}^2$  function. Define, for all  $x, v \in \mathbb{R}^d$ :

$$W_1(x, v) = \exp \left[ F_1(x, v) - \inf_{\mathbb{R}^{2d}} F_1 \right] \geq 1. \quad (6.8)$$

Let us mention that because  $W_1 \in \mathcal{C}^{1,2}(\mathbb{R}^d \times \mathbb{R}^d)$  (i.e.  $W_1$  is  $\mathcal{C}^1$  in the variable  $x$  and  $\mathcal{C}^2$  in the variable  $v$ ),  $W_1 \in \mathbb{D}_e(\mathcal{L})$  and  $\mathcal{L}W_1 = \mathcal{L}_0W_1$  quasi-everywhere (see (2.4)).

Let us now check that **(C3)** is satisfied for  $(P_t, t \geq 0)$  under the above assumptions on  $V$  and  $c$ . This is the purpose of the next proposition.

**Proposition 6.4.** *Assume that  $\Sigma$  satisfies **(A $\Sigma$ )**. Assume that:*

- (1) *The functions  $V$  and  $c$  satisfy **(Av1)**, **(Av2)**, **(Ac1)**, and **(Ac2)**. Then, for a well chosen function  $w \in C_c^2(\mathbb{R}^d, \mathbb{R})$ , and well chosen constants  $a > 0$  and  $b > 0$  (see [66, Eq. (3.4)  $\rightarrow$  Eq. (3.9)] for explicit conditions), Assumption **(C3)** is satisfied for the process (6.1) with the function  $W_1$  defined in (6.8).*
- (2) *The functions  $c$  and  $V$  satisfy respectively (6.2) and (6.3). Then, Assumption **(C3)** is satisfied for the process (6.1) if*

$$\ell_0 < n_0 - 2 \quad (6.9)$$

*with the continuous **bounded** Lyapunov function  $W_2 : \mathbb{R}^{2d} \rightarrow \mathbb{R}$  defined below in (6.12).*

Let us mention that  $W_1$  (see (6.8)) and  $W_2$  (see (6.12)) in Proposition 6.4 are not unique by construction (see indeed [66, Eq. (3.4)  $\rightarrow$  Eq. (3.9)] and the proof of item (2) of Proposition 6.4 below). Moreover, let us notice that  $W_1$  is not bounded over  $\mathbb{R}^{2d}$ . Concerning item (1) in Proposition 6.4, we also refer to [66, Section 5] for other Lyapunov functions in explicit examples like the noisy Van Der Pol oscillator.

*Proof.* Item (1) in Proposition 6.4 is proved in [66, Section 3] (see more precisely Eq. (3.9) there). Let us thus prove item (2) in Proposition 6.4. Assume that  $c$  and  $V$  satisfy respectively (6.2) and (6.3). Recall that the Hamiltonian of (6.1) is

$$(x, v) \in \mathbb{R}^{2d} \mapsto \frac{|v|^2}{2} + V(x).$$

The infinitesimal generator of the process (6.1) is in this case (see (6.4), (6.2), and (6.3)):

$$\mathcal{L}_0 = \frac{\Sigma^2(x, v)}{2} \Delta_v + v \cdot \nabla_x - [\nabla V(x) + |v|^{\ell_0} v] \cdot \nabla_v.$$

Let us construct a Lyapunov function for such a process. To avoid any problem of regularity at 0 in the upcoming computations, let us actually consider

$$(x, v) \in \mathbb{R}^{2d} \mapsto \mathbf{H}_2(x, v) = \frac{|v|^2}{2} + V(x) + \mathbf{k}_0,$$

where  $\mathbf{k}_0 > 0$  is such that  $V(x) + \mathbf{k}_0 \geq 1$  for all  $x \in \mathbb{R}^d$  (see (6.3)). Let

$$a > 0, \alpha > 0, b > 0, \text{ and } \beta > 0.$$

be positive real numbers. Assume that (recall  $\mathbf{n}_0 > 2$ ),

$$0 < \beta - \alpha \leq \frac{1}{2} - \frac{1}{\mathbf{n}_0}, \quad (6.10)$$

so that the function

$$(x, v) \in \mathbb{R}^{2d} \mapsto \mathbf{F}_2(x, v) = -a\mathbf{H}_2^{-\alpha}(x, v) + b x \cdot v \mathbf{H}_2^{\beta-\alpha-1}(x, v),$$

is bounded. Indeed,  $a\mathbf{H}_2^{-\alpha}$  is a bounded function over  $\mathbb{R}^{2d}$ . For the other term, set  $\lambda = \beta - \alpha$ , and use Young's inequality with  $q = 2(1 - \lambda) > 1$  ( $\lambda < 1/2$ , see (6.10)) and  $p = q/(q - 1) = 2(1 - \lambda)/(1 - 2\lambda)$ , to get

$$\frac{|x \cdot v|}{\mathbf{H}_2(x, v)^{1-\lambda}} \leq \frac{x^p}{p\mathbf{H}_2(x, v)^{1-\lambda}} + \frac{v^q}{q\mathbf{H}_2(x, v)^{1-\lambda}}. \quad (6.11)$$

The function  $x, v \mapsto |x \cdot v| \mathbf{H}_2(x, v)^{\lambda-1}$  is thus bounded if  $p \leq \mathbf{n}_0(1 - \lambda)$  (see (6.3)), that is  $2(1 - \lambda)/(1 - 2\lambda) \leq \mathbf{n}_0(1 - \lambda)$  which writes  $2 \leq \mathbf{n}_0 - 2\lambda\mathbf{n}_0$  which is precisely (6.10). Then set, for  $x, v \in \mathbb{R}^d$ ,

$$(x, v) \in \mathbb{R}^{2d} \mapsto \mathbf{W}_2(x, v) = \exp\left(\mathbf{F}_2(x, v) - \inf_{\mathbb{R}^{2d}} \mathbf{F}_2\right), \quad (6.12)$$

which belongs to  $\mathcal{C}^{1,2}(\mathbb{R}^d \times \mathbb{R}^d)$  (thus  $\mathbf{W}_2 \in \mathbb{D}_e(\mathcal{L})$ ). For ease of notation, in the following, we will simply denote  $\mathbf{F}_2$  (resp.  $\mathbf{H}_2$ ,  $\mathbf{W}_2$ ) by  $\mathbf{F}$  (resp.  $\mathbf{H}$ ,  $\mathbf{W}$ ). It holds,  $\frac{\partial_x \mathbf{W}}{\mathbf{W}} = \nabla_x \mathbf{F}$ ,  $\frac{\partial_v \mathbf{W}}{\mathbf{W}} = \nabla_v \mathbf{F}$ , and  $\frac{\Delta_v \mathbf{W}}{\mathbf{W}} = \Delta_v \mathbf{F} + |\nabla_v \mathbf{F}|^2$ . Thus, one has:

$$\frac{\mathcal{L}_0 \mathbf{W}}{\mathbf{W}} = \frac{1}{2} \Sigma^2 \Delta_v \mathbf{F} + \frac{1}{2} \Sigma^2 |\nabla_v \mathbf{F}|^2 + v \cdot \nabla_x \mathbf{F} - [\nabla V + |v|^{\ell_0} v] \nabla_v \mathbf{F}, \quad (6.13)$$

where we recall that  $\Sigma : \mathbb{R}^{2d} \rightarrow \mathbb{R}$  is smooth and bounded, by assumption. One has,

$$\nabla_x \mathbf{F} = a\alpha \nabla V \mathbf{H}^{-\alpha-1} + b v \mathbf{H}^{\beta-\alpha-1} - b(\alpha + 1 - \beta) x \cdot v \nabla V \mathbf{H}^{\beta-\alpha-2},$$

and

$$\nabla_v \mathbf{F} = a\alpha v \mathbf{H}^{-\alpha-1} + b x \mathbf{H}^{\beta-\alpha-1} - b(\alpha + 1 - \beta) x \cdot v v \mathbf{H}^{\beta-\alpha-2}.$$

We then have

$$\begin{aligned} v \cdot \nabla_x \mathbf{F} - [\nabla V + |v|^{\ell_0} v] \nabla_v \mathbf{F} &= \mathbf{H}^{-\alpha-1} \left[ -a\alpha |v|^{\ell_0+2} + b|v|^2 \mathbf{H}^\beta - b x \cdot \nabla V \mathbf{H}^\beta \right. \\ &\quad \left. - b|v|^{\ell_0} x \cdot v \mathbf{H}^\beta + b(\alpha + 1 - \beta) x \cdot v |v|^{\ell_0+2} \mathbf{H}^{\beta-1} \right]. \end{aligned} \quad (6.14)$$

Moreover, it holds:

$$\begin{aligned} \Delta_v \mathbf{F} &= a\alpha \mathbf{H}^{-\alpha-1} - a\alpha |v|^2 (\alpha + 1) \mathbf{H}^{-\alpha-2} + b(\beta - \alpha - 1) x \cdot v \mathbf{H}^{\beta-\alpha-2} \\ &\quad - (d + 1)b(\alpha + 1 - \beta) x \cdot v \mathbf{H}^{\beta-\alpha-2} - b(\alpha + 1 - \beta)(\beta - \alpha - 2) x \cdot v |v|^2 \mathbf{H}^{\beta-\alpha-3}. \end{aligned}$$

The functions  $(x, v) \mapsto a\alpha\mathbf{H}^{-\alpha-1}$  and  $(x, v) \mapsto |v|^2\mathbf{H}^{-\alpha-2}$  are clearly bounded over  $\mathbb{R}^{2d}$ . In addition, since  $\mathbf{H}^{-1}$  is bounded over  $\mathbb{R}^{2d}$ , there exist  $C > 0$ , such that

$$|x \cdot v|\mathbf{H}^{\beta-\alpha-2} \leq C|x \cdot v|\mathbf{H}^{\beta-\alpha-1} = C|x \cdot v|\mathbf{H}^{\lambda-1},$$

and from the analysis led in (6.11) above (see (6.10) and recall that  $\lambda = \beta - \alpha$ ),  $(x, v) \mapsto |x \cdot v|\mathbf{H}^{\beta-\alpha-2}$  is bounded over  $\mathbb{R}^{2d}$ . Finally,  $(x, v) \mapsto |x \cdot v||v|^2\mathbf{H}^{\beta-\alpha-3}$  is also bounded over  $\mathbb{R}^{2d}$  since (see also (6.11) and (6.10)),  $|x \cdot v||v|^2\mathbf{H}^{\beta-\alpha-3} = |x \cdot v|\mathbf{H}^{\lambda-1} \times |v|^2\mathbf{H}^{-2}$ . One then obtains that

$$\Sigma^2\Delta_v\mathbf{F} \text{ is bounded over } \mathbb{R}^{2d}.$$

Similarly, the functions  $|v|\mathbf{H}^{-\alpha-1}$ ,  $|x|\mathbf{H}^{\beta-\alpha-1} = |x|\mathbf{H}^{\lambda-1}$  (recall that  $\lambda < (n_0 - 1)/n_0$  so that  $n_0(1 - \lambda) > 1$ , see (6.10)), and  $|x \cdot v||v|\mathbf{H}^{\beta-\alpha-2} = |x \cdot v|\mathbf{H}^{\lambda-1} \times |v|\mathbf{H}^{-1}$ , are bounded, and then deduces that

$$\Sigma^2|\nabla_v\mathbf{F}|^2 \text{ is bounded over } \mathbb{R}^{2d}.$$

Consequently, from (6.13) and (6.15), one has, for some  $C > 0$  independent of  $x$  and  $v$ ,

$$\begin{aligned} \frac{\mathcal{L}_0\mathbf{W}}{\mathbf{W}} &\leq C + \mathbf{H}^{-\alpha-1} \left[ -a\alpha|v|^{\ell_0+2} + b|v|^2\mathbf{H}^\beta - bx \cdot \nabla V\mathbf{H}^\beta \right. \\ &\quad \left. - b|v|^{\ell_0}x \cdot v\mathbf{H}^\beta + b(\alpha + 1 - \beta)x \cdot v|v|^{\ell_0+2}\mathbf{H}^{\beta-1} \right]. \end{aligned} \quad (6.15)$$

Let us now give a lower bound on the term inside the bracket in (6.15), that we denote by  $\mathbf{M}$ . Let us assume that

$$\beta < 1.$$

Then, it holds, for all  $s \geq 0$  and  $t \geq 0$ ,  $2^{\beta-1}(s^\beta + t^\beta) \leq (s+t)^\beta \leq s^\beta + t^\beta$ . Since there exists  $r_1 > 0$  such that so that for all  $x, v \in \mathbb{R}$ :

$$2^{\beta-1}V_0^\beta(x) + \frac{v^{2\beta}}{2} \leq \mathbf{H}^\beta(x, v) \leq V_0^\beta(x) + \frac{v^{2\beta}}{2^\beta},$$

where for all  $x \in \mathbb{R}^d$  we set:

$$V_0(x) = V(x) + \mathbf{k}_0,$$

which satisfies (see (6.3)), for some  $r > 0$  and  $C > 0$ , and all  $|x| > r$ ,

$$C^{-1}|x|^{n_0} \leq V_0(x) \leq C|x|^{n_0} \text{ and } C^{-1}|x|^{n_0} \leq x \cdot \nabla V_0(x). \quad (6.16)$$

Therefore, since  $b(\alpha + 1 - \beta) > 0$ ,

$$\begin{aligned} \mathbf{M} &\leq -a\alpha|v|^{\ell_0+2} + b|v|^2V_0^\beta + \frac{b}{2^\beta}|v|^{2+2\beta} - 2^{\beta-1}b|x \cdot \nabla V_0|V_0^\beta - \frac{b}{2}|x \cdot \nabla V_0|v^{2\beta} \\ &\quad + \mathbf{1}_{x \cdot v \leq 0}b|x|V_0^\beta|v|^{\ell_0+1} + \frac{\mathbf{1}_{x \cdot v \leq 0}}{2^\beta}b|x||v|^{\ell_0+1+2\beta} \\ &\quad + \mathbf{1}_{x \cdot v \geq 0}b(\alpha + 1 - \beta)\frac{|x|V_0^\beta|v|^{\ell_0+3}}{V_0 + v^2/2} + \frac{\mathbf{1}_{x \cdot v \geq 0}}{2^\beta}b(\alpha + 1 - \beta)\frac{|x||v|^{3+\ell_0+2\beta}}{V_0 + v^2/2}. \end{aligned} \quad (6.17)$$

Let us now find conditions such that  $-a\alpha|v|^{\ell_0+2}$  and  $-2^{\beta-1}bx \cdot \nabla V_0V_0^\beta$  are dominant in the right hand side of (6.17). From (6.16), for  $|x| \geq r$ ,

$$C^{-1}|x|^{n_0+n_0\beta} \leq |x \cdot \nabla V_0(x)|V_0^\beta(x) = V_0^\beta(x)x \cdot \nabla V_0(x), \quad (6.18)$$

for some  $C > 0$  independent of  $x$ . Assume that

$$\beta < \ell_0/2, \quad (6.19)$$

so that  $|v|^{2+2\beta} = o(|v|^{\ell_0+2})$  as  $|v| \rightarrow +\infty$ . In addition, for  $0 < \epsilon < \lambda_0$ , using Young's inequality with  $p_\epsilon = (\ell_0 + 2 - \epsilon)/2 > 1$  and  $q_\epsilon = p_\epsilon/(p_\epsilon - 1) = 1 + 2/(\ell_0 - \epsilon)$ :

$$|v|^2 V_0^\beta \leq p_\epsilon^{-1} |v|^{\ell_0+2-\epsilon} + q_\epsilon^{-1} V_0^{\beta(1+2/(\ell_0-\epsilon))}.$$

From (6.19), for  $\epsilon > 0$  small enough  $\beta < (\ell_0 - \epsilon)/2$  and thus  $\mathbf{n}_0\beta(1 + 2/(\ell_0 - \epsilon)) < \mathbf{n}_0\beta + \mathbf{n}_0$ . Then, for such  $\epsilon > 0$ ,  $|v|^{\ell_0+2-\epsilon} = o(|v|^{\ell_0+2})$  as  $|v| \rightarrow +\infty$  and  $V_0^{\beta(1+2/(\ell_0-\epsilon))} = o(|x|^{\mathbf{n}_0+\beta\mathbf{n}_0}) \rightarrow 0$  as  $|x| \rightarrow +\infty$  (see (6.16)). For  $\epsilon \ll 1$ , using again Young's inequality with  $p_\epsilon = (\ell_0 + 2 - \epsilon)/(\ell_0 + 1) = 1 + (1 - \epsilon)/(\ell_0 + 1) > 1$  and  $q_\epsilon = p_\epsilon/(p_\epsilon - 1) = (\ell_0 + 2 - \epsilon)/(1 - \epsilon) = 1 + (\ell_0 + 1)/(1 - \epsilon)$ :

$$|v|^{\ell_0+1} |x| V_0^\beta \leq p_\epsilon^{-1} |v|^{\ell_0+2-\epsilon} + q_\epsilon^{-1} |x| V_0^{\beta[1+(\ell_0+1)/(1-\epsilon)]}.$$

Let us check that  $(1 + \mathbf{n}_0\beta)[1 + (\ell_0 + 1)/(1 - \epsilon)] < \mathbf{n}_0\beta + \mathbf{n}_0$ . This is equivalent to  $\mathbf{n}_0\beta < (\mathbf{n}_0 - 1)(1 - \epsilon)/(\ell_0 + 1) - 1$ . Notice that from (6.9),  $(\mathbf{n}_0 - 1)/(\ell_0 + 1) > 1$  and thus for  $\epsilon > 0$  small enough  $(\mathbf{n}_0 - 1)(1 - \epsilon)/(\ell_0 + 1) > 1$ . Then, assume that

$$\mathbf{n}_0\beta < (\mathbf{n}_0 - 1)/(\ell_0 + 1) - 1 \quad (6.20)$$

so that for  $\epsilon > 0$  small enough,  $|x| V_0^{\beta[1+(\ell_0+1)/(1-\epsilon)]} = o(|x|^{\mathbf{n}_0+\mathbf{n}_0\beta}) \rightarrow 0$  as  $|x| \rightarrow +\infty$ . Assume also that

$$\beta < 1/2 \quad (6.21)$$

so that, for  $\epsilon > 0$  small enough,  $p_\epsilon = (\ell_0 + 2 - \epsilon)/(\ell_0 + 1 + 2\beta) > 1$ . Then,

$$|x| |v|^{\ell_0+1+2\beta} \leq p_\epsilon^{-1} |v|^{\ell_0+2-\epsilon} + q_\epsilon^{-1} |x|^{\ell_0+2-\epsilon/(1-\epsilon-2\beta)}.$$

Assume that

$$\ell_0 + 2 < (\mathbf{n}_0\beta + \mathbf{n}_0)(1 - 2\beta),$$

which is satisfied if  $\beta > 0$  is small enough since  $\ell_0 + 2 < \mathbf{n}_0$  (see (6.3)). Then for  $\epsilon > 0$  small enough,  $(\ell_0+2-\epsilon)/(1-\epsilon-2\beta) < \mathbf{n}_0\beta + \mathbf{n}_0$  and  $|x|^{\ell_0+2-\epsilon/(1-\epsilon-2\beta)} = o(|x|^{\mathbf{n}_0\beta + \mathbf{n}_0}) \rightarrow 0$  as  $|x| \rightarrow +\infty$ .

Moreover, for  $\beta > 0$  small enough, it holds  $\mathbf{n}_0/(1 + \mathbf{n}_0\beta) > 1$  and thus,  $V_0 + v^2/2 \geq C^{-1} V_0^{1/\mathbf{n}_0 + \beta} v^{3/2 - 2\beta}$ , for some  $C > 0$ . Then, because  $|x| V_0^{-1/\mathbf{n}_0}$  is a bounded function (see (6.16)), it holds for some  $M > 0$  independent of  $x$  and  $v$ ,

$$\frac{|x|^{1+\mathbf{n}_0\beta} |v|^{\ell_0+3}}{V_0 + v^2/2} \leq M |v|^{\ell_0+3/2+2\beta}. \quad (6.22)$$

If  $\beta < 1/4$ , then the left hand side of (6.22) is equal to  $o(|v|^{\ell_0+2})$  as  $|v| \rightarrow +\infty$ . Finally, since  $V_0 + |v|^2/2 \geq C^{-1} V_0^{1/\mathbf{n}_0} |v|^{2(\mathbf{n}_0-1)/\mathbf{n}_0}$ , for some  $M > 0$  independent of  $x$  and  $v$ ,

$$\frac{|x| |v|^{\ell_0+3+2\beta}}{V_0 + v^2/2} \leq M |v|^{\ell_0+1+2\beta+2/\mathbf{n}_0}.$$

Taking  $\beta > 0$  such that  $2\beta + 2/\mathbf{n}_0 < 1$  (this is possible because  $2/\mathbf{n}_0 < 1$  by assumption), the left hand side of the previous inequality is equal to  $o(|v|^{\ell_0+2})$  as  $|v| \rightarrow +\infty$ . In

conclusion, all the previous estimates together with (6.18) imply that there exists  $\eta$  (depending on  $\ell_0$  and  $\mathbf{n}_0$ ) such that if

$$0 < \beta < \eta, \quad (6.23)$$

then (recalling also that  $0 < \alpha < \beta$  and  $\eta \leq \ell_0$ , see (6.10) and (6.19)), there exists  $C > 0$  and a continuous function  $x, v \in \mathbb{R}^d \mapsto \mathbf{K}(x, v)$  such that (see (6.15))

$$\frac{\mathcal{L}_0 W}{W} \leq C - \mathbf{K}(x, v), \quad (6.24)$$

with  $\mathbf{K}(x, v) \rightarrow +\infty$  if  $|x| + |v| \rightarrow +\infty$ . This ends the proof of Proposition 6.4.  $\square$

Let us now consider  $\mathcal{O}$ , a nonempty subdomain of  $\mathbb{R}^d$  (not necessarily bounded), that is  $\mathcal{O}$  is a connected open subset of  $\mathbb{R}^d$ . As explained in the introduction, we are interested in the existence of quasi-stationary distributions for the processes (6.1) in

$$\mathcal{D} = \mathcal{O} \times \mathbb{R}^d. \quad (6.25)$$

Of course, other domains might be considered with our techniques. Recall that  $\sigma_{\mathcal{D}}$  (see (2.1)) is the first exit time from  $\mathcal{D}$  for the process (6.1):

$$\sigma_{\mathcal{D}}(\mathbf{x}) = \inf \{t \geq 0, X_t(\mathbf{x}) \notin \mathcal{D}\} = \inf \{t \geq 0, x_t(\mathbf{x}) \notin \mathcal{O}\}, \quad (6.26)$$

where we recall that  $(X_t(\mathbf{x}), t \geq 0)$  stands for the process  $(X_t, t \geq 0)$  when  $X_0 = \mathbf{x} \in \mathbb{R}^{2d}$ . Let us now check the other assumptions on  $(P_t^{\mathcal{D}}, t \geq 0)$  (the semigroup of the process (6.1) killed when exiting  $\mathcal{D}$ , see (2.2) and (6.26)) needed to apply Theorem 2.2.

**6.2.3. The semigroup  $(P_t^{\mathcal{D}}, t \geq 0)$  is topologically irreducible.**

**Lemma 6.5.** *Assume that  $V$ ,  $c$  and  $\Sigma$  satisfy **(Av1)**, **(Ac0)**, and **(A $\Sigma$ )**. Then,  $(P_t^{\mathcal{D}}, t \geq 0)$  is topologically irreducible. If the open set  $\mathbb{R}^d \setminus \overline{\mathcal{O}}$  is not empty, then for all  $\mathbf{x} \in \mathcal{D}$  and  $t > 0$ ,*

$$\mathbb{P}_{\mathbf{x}}(\sigma_{\mathcal{D}} < t) > 0,$$

*which implies in particular that  $\mathbb{P}_{\mathbf{x}}(\sigma_{\mathcal{D}} < +\infty) > 0$  (thus **(C5)** is satisfied for the process (6.1) when  $\mathcal{D} = \mathcal{O} \times \mathbb{R}^d$ ).*

*Proof.* We will apply the Stroock-Varadhan support theorem.

**Step 1:** the case when  $V = 0$  and  $c = 0$ .

Recall that the process  $(X_t^0 = (x_t^0, v_t^0), t \geq 0)$  is the solution (in the strong sense) to the stochastic differential equation (6.5). Let  $(P_t^{\mathcal{D},0}, t \geq 0)$  denote the semigroup of the process (6.5) killed when exiting  $\mathcal{D}$ . Denote by  $\sigma_{\mathcal{D}}^0$  the first time the process  $(X_t^0 = (x_t^0, v_t^0), t \geq 0)$  exits  $\mathcal{D}$  (see (6.26)). The stochastic differential equation (6.5) writes in the Stratonovich form

$$\begin{cases} dx_t^0 = v_t^0 dt, \\ dv_t^0 = -\frac{1}{2}\Sigma(x_t^0, v_t^0)\nabla_v \Sigma(x_t^0, v_t^0) + \Sigma(x_t^0, v_t^0) \circ dB_t. \end{cases}$$

Let  $\mathcal{O}_1$  be a nonempty open subset of  $\mathcal{O}$  and  $\mathcal{O}_2$  be a nonempty open subset of  $\mathbb{R}^d$ . Consider  $\mathbf{x}_0 = (x_0, v_0) \in \mathcal{D}$  and  $\mathbf{x}_1 = (x_1, v_1) \in \mathcal{O}_1 \times \mathcal{O}_2$ . Let  $t > 0$  and  $\gamma : [0, t] \rightarrow \mathcal{O}$  be a  $\mathcal{C}^1$  and piecewise  $\mathcal{C}^2$  curve such that  $\gamma(0) = x_0$ ,  $\dot{\gamma}(0) = v_0$ ,  $\gamma(t) = x_1$ , and  $\dot{\gamma}(t) = v_1$ .

The construction of such a  $\gamma$  can be done by a local cubic interpolation in time as it is done for instance in [44, Lemma 4.2] (their arguments also hold though  $\mathcal{O}$  is not bounded<sup>4</sup>). For  $s \in [0, t]$ , set

$$\mathbf{Y}(s) = \begin{pmatrix} \mathbf{Y}_1(s) \\ \mathbf{Y}_2(s) \end{pmatrix} \text{ where } \mathbf{Y}_1(s) = \gamma(s) \text{ and } \mathbf{Y}_2(s) = \dot{\gamma}(s).$$

Then, define the piecewise continuous function  $h : [0, t] \rightarrow \mathbb{R}^d$  by:

$$\mathbf{h}(s) = \frac{1}{\Sigma(\gamma(s), \dot{\gamma}(s))} \left[ \ddot{\gamma}(s) + \frac{1}{2} \Sigma(\gamma(s), \dot{\gamma}(s)) \nabla_v \Sigma(\gamma(s), \dot{\gamma}(s)) \right], \quad s \in [0, t].$$

Clearly,  $\mathbf{h} \in L^2([0, t], \mathbb{R}^d)$ ,  $\mathbf{Y}(0) = \mathbf{x}_0$ , and for all  $s \in [0, t]$ ,

$$\begin{cases} \dot{\mathbf{Y}}_1(s) = \mathbf{Y}_2(s), \\ \dot{\mathbf{Y}}_2(s) = -\frac{1}{2} \Sigma(\mathbf{Y}_1(s), \mathbf{Y}_2(s)) \nabla_v \Sigma(\mathbf{Y}_1(s), \mathbf{Y}_2(s)) + \Sigma(\mathbf{Y}_1(s), \mathbf{Y}_2(s)) \mathbf{h}(s). \end{cases}$$

By the Stroock-Varadhan support theorem [60] (see also [5, Theorem 4]), for all  $\varepsilon > 0$  and  $t > 0$ ,

$$\mathbb{P}_{\mathbf{x}_0} \left( \sup_{s \in [0, t]} |X_s^0 - \mathbf{Y}(s)| < \varepsilon \right) > 0.$$

Since for all  $s \in [0, t]$ ,  $\mathbf{Y}(s) = (\gamma^T(s), \dot{\gamma}^T(s))^T \in \mathcal{O} \times \mathbb{R}^d = \mathcal{D}$  and  $\mathcal{O}$  is open, for  $\varepsilon > 0$  small enough, if  $\sup_{s \in [0, t]} |X_s^0 - \mathbf{Y}(s)| < \varepsilon$ , then  $X_s^0 = (x_s, v_s) \in \mathcal{D}$  for all  $s \in [0, t]$  (in particular  $t < \sigma_{\mathcal{D}}^0$  by continuity of the trajectories), and  $x_t \in B(x_1, 2\varepsilon) \subset \mathcal{O}_1$  and  $v_t \in B(v_1, 2\varepsilon) \subset \mathcal{O}_2$ . Thus

$$P_t^{D,0}(\mathbf{x}_0, \mathcal{O}_1 \times \mathcal{O}_2) = \mathbb{P}_{\mathbf{x}_0}(X_t^0 \in \mathcal{O}_1 \times \mathcal{O}_2, t < \sigma_{\mathcal{D}}^0) \geq \mathbb{P}_{\mathbf{x}_0} \left( \sup_{s \in [0, t]} |X_s^0 - \mathbf{Y}(s)| < \varepsilon \right) > 0,$$

which is precisely the topological irreducibility of  $(P_t^{D,0}, t \geq 0)$ . If the open set  $\mathbb{R}^d \setminus \overline{\mathcal{O}}$  is not empty, then choosing  $\mathcal{O}_1$  such that  $\mathcal{O}_1 \subset \mathbb{R}^d \setminus \overline{\mathcal{O}}$ , one deduces with the same arguments as above<sup>5</sup> that for all  $\mathbf{x}_0 \in \mathcal{D}$  and  $t > 0$ ,

$$\mathbb{P}_{\mathbf{x}_0}(\sigma_{\mathcal{D}}^0 < +\infty) \geq \mathbb{P}_{\mathbf{x}_0}(\sigma_{\mathcal{D}}^0 < t) \geq \mathbb{P}_{\mathbf{x}_0}(X_t \in \mathcal{O}_1 \times \mathbb{R}^d) > 0.$$

**Step 2:** the case when  $V \neq 0$  and  $c \neq 0$ .

Let us now come back to the case when  $V \neq 0$  and  $c \neq 0$ . Pick  $f \in b\mathcal{B}(\mathcal{D})$ . Since  $V$  and  $c$  satisfy **(Av1)** and **(Ac0)**, from [66, Lemma 1.1] (and the remark on page 7 there) and by uniqueness in distribution of the weak solution of (6.5), the law of  $(X_t, t \geq 0)$

<sup>4</sup>Let  $\eta : [0, t] \rightarrow \mathcal{O}$  be a  $\mathcal{C}^0$  curve such that  $\eta(0) = x_0$  and  $\eta(t) = x_1$  ( $\mathcal{O}$  is path-connected because it is connected and locally path-connected, since it is open). Then, apply [44, Lemma 4.2] with (using the notation there)  $K = \{x_0, x_1\}$ ,  $\delta_K = \frac{1}{2} \text{dist}_{\mathbb{R}^d}(\text{Ran } \eta, \partial\mathcal{O}) > 0$ ,  $K' = \{(x, v) \in \mathbb{R}^{2d}, \text{dist}_{\mathbb{R}^d}(x, \text{Ran } \eta) \leq \delta_K/2, |v| \leq \max(|v_0|, |v_1|)\}$  (which is a compact and connected subset of  $\mathcal{D}$ ), and  $\varepsilon = \delta_K/2$ .

<sup>5</sup>In this case, consider  $\gamma : [0, t] \rightarrow \mathbb{R}^d$  be a smooth curve such that  $\gamma(0) = x_0$ ,  $\dot{\gamma}(0) = v_0$ ,  $\gamma(t) = x_1$ , and  $\dot{\gamma}(t) = v_1$ . Such a curve can be simply constructed by a (global) cubic interpolation in time.

under  $\mathbb{P}_{x_0}$  is the law of  $(X_t^0, t \geq 0)$  under the probability  $\mathbf{M}_t(x_0)d\mathbb{P}_{x_0}$ , where  $\mathbf{M}_t$  is the following martingale

$$\begin{aligned} \mathbf{M}_t = \exp & \left[ \left( -\frac{1}{\sqrt{2}} \int_0^t \Sigma^{-1}(x_s^0, v_s^0) (c(x_s^0, v_s^0)v_s^0 + \nabla V(x_s^0)) d\mathbf{W}_s \right. \right. \\ & \left. \left. - \frac{1}{4} \int_0^t |\Sigma^{-1}(x_s^0, v_s^0) [c(x_s^0, v_s^0)v_s^0 + \nabla V(x_s^0)]|^2 ds \right) \right], \end{aligned}$$

and where  $(\mathbf{W}_t, t \geq 0)$  is a standard Wiener process under the law of the process  $(X_t^0, t \geq 0)$ . Thus, it holds for  $x \in \mathcal{D}$  and  $t > 0$ :

$$\mathbb{E}_x[f(X_t)\mathbf{1}_{t < \sigma_{\mathcal{D}}}] = \mathbb{E}^{\mathbb{P}}[f(X_t^0(x))\mathbf{1}_{t < \sigma_{\mathcal{D}}^0(x)}\mathbf{M}_t(x)]. \quad (6.27)$$

In particular, for all  $t > 0, x \in \mathcal{D}$ ,

$$\mathbb{P}_x(X_t \in \mathcal{O}_1 \times \mathcal{O}_2, t < \sigma_{\mathcal{D}}) > 0 \text{ if and only if } \mathbb{P}_x(X_t^0 \in \mathcal{O}_1 \times \mathcal{O}_2, t < \sigma_{\mathcal{D}}^0) > 0,$$

and

$$\mathbb{P}_x(\sigma_{\mathcal{D}} < t) > 0 \text{ if and only if } \mathbb{P}_x(\sigma_{\mathcal{D}}^0 < t) > 0.$$

This ends the proof of Lemma 6.5.  $\square$

#### 6.2.4. Weak Feller property of $P_t^{\mathcal{D}}$ .

**Proposition 6.6.** *Assume that  $V, c,$  and  $\Sigma$  satisfy respectively the assumptions **(Av1)**, **(Ac0)**, and **(A $\Sigma$ )**. Assume that  $\mathcal{O}$  is a  $\mathcal{C}^2$  subdomain of  $\mathbb{R}^d$  such that  $\mathbb{R}^d \setminus \overline{\mathcal{O}}$  is non-empty. Then, for  $t > 0, P_t^{\mathcal{D}}$  is strong Feller on  $\mathcal{D}$  (and thus weakly Feller on  $\mathcal{D}$ ). Thus, assumption **(C4)** is satisfied for  $P_t^{\mathcal{D}}$ .*

*Proof.* The proof of Proposition 6.6 is divided into several steps.

**Step 1:** properties of the process  $(X_t^0, t \geq 0)$ , see (6.5).

**Step 1a:** Proof of (6.29).

In this step, we prove that, for  $y = (x_y, v_y) \in \partial\mathcal{D}$ , if

$$\mathbf{n}(x_y) \cdot v_y \geq 0, \quad (6.28)$$

then for all  $t > 0$ , there exists almost surely  $u \in (0, t]$ , such that

$$x_u^0(y) \in \mathbb{R}^d \setminus \overline{\mathcal{O}}. \quad (6.29)$$

Equation (6.29) has been very recently proved in [44, (i) in Proposition 2.8] for the process (6.1) when  $\Sigma$  is constant (that is,  $\Sigma$  is independent of  $x$  and  $v$ ). The proof of (6.29) requires further analysis when  $\Sigma$  is not constant. When  $v_y \cdot \mathbf{n}(x_y) > 0$ , the proof of (6.29) is straightforward. Indeed, because  $\partial\mathcal{O}$  is  $\mathcal{C}^2$ , in a neighborhood  $\mathbf{U}$  of  $x_y \in \partial\mathcal{O}$  in  $\mathbb{R}^d$ ,  $\mathcal{O}$  is given by  $\{\Psi < 0\}$  for some  $\mathcal{C}^2$  function  $\Psi : \mathbb{R}^d \rightarrow \mathbb{R}$  such that  $\mathbf{n}(x_y) = \nabla\Psi(x_y)$  and  $\partial\mathcal{O}$  is given by  $\{\Psi = 0\}$ . Then, for  $t \geq 0$  (sufficiently small, say  $t \leq t^*(y)$ ), so that  $x_t^0(y) \in \mathbf{U}$  for all  $t \in [0, t^*(y)]$ ,

$$\Psi(x_t^0(y)) = \int_0^t \nabla\Psi(x_s^0(y)) \cdot v_s^0(y) ds.$$

In addition, since

$$\nabla\Psi(x_0^0(\mathbf{y})) \cdot v_0^0(\mathbf{y}) = v_{\mathbf{y}} \cdot \mathbf{n}(x_{\mathbf{y}}) > 0$$

and because  $s \geq 0 \mapsto \nabla\Psi(x_s^0(\mathbf{x})) \cdot v_s^0(\mathbf{x})$  is continuous almost surely, one deduces that for all  $t > 0$  small enough,

$$\Psi(x_t^0(\mathbf{y})) > 0,$$

which concludes the proof of (6.29) when  $v_{\mathbf{y}} \cdot \mathbf{n}(x_{\mathbf{y}}) > 0$ . Let us now prove (6.29) when  $v_{\mathbf{y}} \cdot \mathbf{n}(x_{\mathbf{y}}) = 0$ . One has uniformly in  $x \in \mathbf{U}$  (recall that  $\nabla\Psi$  is of regularity  $\mathcal{C}^1$ ),

$$\nabla\Psi(x) = \mathbf{n}(x_{\mathbf{y}}) + O(|x - x_{\mathbf{y}}|),$$

so that, using in addition that  $v_s^0(\mathbf{y}) \cdot \mathbf{n}(x_0^0(\mathbf{y})) = v_{\mathbf{y}} \cdot \mathbf{n}(x_{\mathbf{y}}) = 0$ , for  $t \in [0, t^*(\mathbf{y})]$ ,

$$\begin{aligned} \Psi(x_t^0(\mathbf{y})) &= \int_0^t \left[ v_{\mathbf{y}} + \int_0^s \Sigma((x_u^0(\mathbf{y}), v_u^0(\mathbf{y}))) dB_u \right] \cdot \left[ \mathbf{n}(x_{\mathbf{y}}) + O(\underbrace{|x_s^0(\mathbf{y}) - x_{\mathbf{y}}|}_{=|\int_0^s v_u^0(\mathbf{y}) du|}) \right] ds \\ &= \int_0^t M_s ds + O(t^2) \sup_{s \in [0, t]} |v_s^0(\mathbf{y})|^2 \end{aligned} \tag{6.30}$$

where we set for  $s \geq 0$ ,

$$M_s = \int_0^s \Sigma(x_u^0(\mathbf{y}), v_u^0(\mathbf{y})) d\omega_u$$

and where  $(\omega_u, u \geq 0)$  is a standard one dimensional Brownian motion ( $\omega_u = B_u \cdot \mathbf{n}(x_{\mathbf{y}})$ ). Then, to prove (6.29) and in view of the previous estimate, we have to study the sign of  $\int_0^t M_s ds$  (for small  $t > 0$ ). To this end, it is sufficient to show that

$$\limsup_{t \rightarrow 0^+} \frac{\int_0^t M_s ds}{\mathbf{L}(T^{-1}(t))} > 0 \text{ almost surely,} \tag{6.31}$$

where  $\mathbf{L}(r) = \sqrt{2/3} r^{3/2} \sqrt{\log \log(1/r)}$  (for  $r > 0$ ), and where for  $s \geq 0$ ,

$$T(s) = \int_0^s \Sigma^2(x_u^0(\mathbf{y}), v_u^0(\mathbf{y})) du,$$

which is (almost surely) strictly increasing and continuously differentiable function from  $\mathbb{R}^+ \rightarrow \mathbb{R}^+$  with  $T(s) = \int_0^s \Sigma^2(x_u^0(\mathbf{y})) du \geq \Sigma_0^2 \times s \rightarrow +\infty$  as  $s \rightarrow +\infty$ . Notice that because for all  $s \geq 0$ ,  $\Sigma_0 s \leq T(s) \leq \Sigma_{\infty} s$ , it holds for all  $u \geq 0$ ,

$$\Sigma_{\infty}^{-1} u \leq T^{-1}(u) \leq \Sigma_0^{-1} u.$$

Thus for some  $C > 0$ , it holds, for  $t \geq 0$  large enough,

$$C^{-1} \mathbf{L}(t) \leq \mathbf{L}(T^{-1}(t)) \leq C \mathbf{L}(t).$$

Consequently,  $t^2/\mathbf{L}(T^{-1}(t)) \rightarrow 0$  as  $t \rightarrow 0$  and thus, in view of (6.30), (6.31) implies (6.29). Thus, let us prove (6.31). By assumption on  $\Sigma$ ,  $(M_s, s \geq 0)$  is a continuous martingale and  $[M]_s = T(s) \rightarrow +\infty$  as  $s \rightarrow +\infty$ . Then, using the Dambis-Dubins-Schwarz theorem, there exists a standard one dimensional Brownian motion  $(V_t, t \geq 0)$  such that for all  $s \geq 0$ ,

$$M_s = V_{[M]_s}.$$



Since  $(T^{-1})'(u) = 1/T'(T^{-1}(u)) = \Sigma^{-2}(x_{T^{-1}(u)}^0, v_{T^{-1}(u)}^0)$ ,

$$\int_0^t M_s ds = \int_0^t V_{T(s)} ds = \int_0^{T^{-1}(t)} V_u \Sigma^{-2}(x_{T^{-1}(u)}^0(y), v_{T^{-1}(u)}^0(y)) du. \quad (6.32)$$

Then, setting  $\kappa = T^{-1}(t)$  ( $t > 0$ ), (6.31) is equivalent to

$$\limsup_{\kappa \rightarrow 0^+} \frac{\int_0^\kappa V_u \Sigma^{-2}(x_{T^{-1}(u)}^0(y), v_{T^{-1}(u)}^0(y)) du}{\mathbf{L}(\kappa)} > 0 \text{ almost surely.} \quad (6.33)$$

The proof of (6.33) is the purpose of the next step.

**Step 1b:** Proof of (6.33).

By assumption on  $\Sigma$  and since  $\nabla \Sigma^{-2} = -2(\nabla \Sigma) \Sigma^{-3}$ , uniformly on  $\mathbf{x}, \mathbf{z} \in \mathbb{R}^{2d}$ , one has:

$$\Sigma^{-2}(\mathbf{x}) = \Sigma^{-2}(\mathbf{z}) + O(|\mathbf{x} - \mathbf{z}|).$$

Thus, using also that  $x_{T^{-1}(u)}^0(y) = x_y + \int_0^{T^{-1}(u)} v_s^0(y) ds$  and  $T^{-1}(u) \leq T^{-1}(t) = \kappa$  (for  $0 \leq u \leq t$ ), one has:

$$\begin{aligned} \int_0^\kappa V_u \Sigma^{-2}(X_{T^{-1}(u)}^0) du &= \Sigma^{-2}(x_y, v_y) \int_0^\kappa V_u du \\ &\quad + O\left(\int_0^\kappa |V_u| [ |x_{T^{-1}(u)}^0(y) - x_y| + |v_{T^{-1}(u)}^0 - v_y| ] du\right) \\ &= \Sigma^{-2}(x_y, v_y) \int_0^\kappa V_u du \\ &\quad + \sup_{u \in [0, \kappa]} |V_u| O\left[\int_0^\kappa |x_{T^{-1}(u)}^0(y) - x_y| du + \kappa \sup_{u \in [0, \kappa]} |v_u^0(y) - v_y|\right] \\ &= \Sigma^{-2}(x_y, v_y) \int_0^\kappa V_u du \\ &\quad + \sup_{u \in [0, \kappa]} |V_u| O\left[\int_0^\kappa u \sup_{u \in [0, \kappa]} |v_u^0(y)| du + \kappa \sup_{u \in [0, \kappa]} |v_u^0(y) - v_y|\right] \\ &= \Sigma^{-2}(x_y, v_y) \int_0^\kappa V_u du \\ &\quad + \sup_{u \in [0, \kappa]} |V_u| \left[ \sup_{u \in [0, \kappa]} |v_u^0(y)| O(\kappa^2) + \sup_{u \in [0, \kappa]} |v_u^0(y) - v_y| O(\kappa) \right], \end{aligned}$$

where we have used that  $T^{-1}(u) \leq \Sigma_0^{-1}u$ . Using the law of the iterated logarithm for  $\int_0^\kappa V_u du$  given in [40, (2) in Theorem 1], it holds:

$$\limsup_{\kappa \rightarrow 0^+} \frac{\Sigma^{-2}(x_y, v_y) \int_0^\kappa V_u du}{\mathbf{L}(\kappa)} = \Sigma^{-2}(x_y, v_y) > 0 \text{ almost surely.}$$

Using the law of the iterated logarithm for  $\sup_{u \in [0, \kappa]} |V_u|$  given in [34], it holds:

$$\limsup_{\kappa \rightarrow 0^+} \frac{\sup_{u \in [0, \kappa]} |V_u|}{\mathbf{P}(\kappa)} = 1 \text{ almost surely,}$$

where  $\mathbf{P}(\kappa) = \sqrt{2\kappa}\sqrt{\log \log(1/\kappa)}$ . Thus, since  $\frac{\mathbf{P}(\kappa)}{\mathbf{L}(\kappa)} \sim \sqrt{3}/\kappa$ , one has almost surely as  $\kappa \rightarrow 0^+$ :

$$\kappa \sup_{u \in [0, \kappa]} |v_u^0(\mathbf{y}) - v_{\mathbf{y}}| \frac{\sup_{u \in [0, \kappa]} |V_u|}{\mathbf{L}(\kappa)} = \frac{\sup_{u \in [0, \kappa]} |V_u|}{\mathbf{P}(\kappa)} \frac{\kappa \mathbf{P}(\kappa)}{\mathbf{L}(\kappa)} \sup_{u \in [0, \kappa]} |v_u^0(\mathbf{y}) - v_{\mathbf{y}}| \rightarrow 0,$$

because almost surely,  $\sup_{u \in [0, \kappa]} |v_u^0(\mathbf{y}) - v_{\mathbf{y}}| \rightarrow 0$  as  $\kappa \rightarrow 0^+$ . This concludes the proof of (6.33) (with more precisely  $\limsup_{\kappa \rightarrow 0^+} \int_0^\kappa V_u \Sigma^{-2}(X_{T^{-1}(u)}^0) du / \mathbf{L}(\kappa) = \Sigma^{-2}(x_{\mathbf{y}}, v_{\mathbf{y}})$  almost surely), and thus of (6.31) and then of (6.29).

**Step 1c:** proof of (6.34).

Let  $(\mathbf{x}_n)_n$  be sequence of elements of  $\mathcal{D}$  such that  $\mathbf{x}_n \rightarrow \mathbf{x} \in \mathcal{D}$  as  $n \rightarrow +\infty$ . Let us prove that for  $t > 0$ ,

$$\text{as } n \rightarrow +\infty, \mathbf{1}_{t < \sigma_{\mathcal{D}}^0(\mathbf{x}_n)} \rightarrow \mathbf{1}_{t < \sigma_{\mathcal{D}}^0(\mathbf{x})} \text{ in } \mathbb{P}\text{-probability.} \quad (6.34)$$

Let us prove (6.34). Let us pick  $t > 0$  and a sequence  $\mathbf{x}_n = (x_n, v_n)$  converging to some  $\mathbf{x} = (x, v) \in \mathcal{D}$ . First of all, recall that under  $(\mathbf{A}\Sigma)$ , for  $t > 0$ ,  $P_t^0$  has a density  $p_t^0$  over  $\mathbb{R}^{2d}$  with respect to the Lebesgue measure. It then holds:

$$\mathbb{P}_{\mathbf{x}}(\sigma_{\mathcal{D}} = t) \leq \mathbb{P}_{\mathbf{x}}(x_t^0 \in \partial \mathcal{O}) = \int_{\partial \mathcal{O} \times \mathbb{R}^d} p_t^0(\mathbf{x}, x, v) dx dv = 0, \quad (6.35)$$

since  $\partial \mathcal{O}$  has Lebesgue measure 0. Indeed, because  $\partial \mathcal{O}$  is  $\mathcal{C}^1$ , for any  $x \in \partial \mathcal{O}$ , there exists  $\varepsilon_x > 0$ , such that the open subset  $\partial \mathcal{O} \cap B(x, \varepsilon_x)$  of  $\partial \mathcal{O}$  has Lebesgue measure 0. Moreover these open subsets of  $\partial \mathcal{O}$  clearly cover  $\partial \mathcal{O}$  and because  $\partial \mathcal{O}$  is Lindelöf (due to the fact that  $\mathbb{R}^d$  is Lindelöf and  $\partial \mathcal{O} = \overline{\mathcal{O}} \cap (\mathbb{R}^d \setminus \mathcal{O})$  is closed),  $\partial \mathcal{O} \subset \cup_{i \in \mathbb{N}} B(x_i, \varepsilon_{x_i}) \cap \partial \mathcal{O}$ . Thus  $\partial \mathcal{O}$  has Lebesgue measure 0. Let  $\eta \in (0, 1)$  and  $t > 0$ . Then, using (6.35), for  $n \geq 1$ , it holds:

$$\begin{aligned} \mathbb{P}(|\mathbf{1}_{t < \sigma_{\mathcal{D}}^0(\mathbf{x}_n)} - \mathbf{1}_{t < \sigma_{\mathcal{D}}^0(\mathbf{x})}| \geq \eta) &= \mathbb{P}(|\mathbf{1}_{t < \sigma_{\mathcal{D}}^0(\mathbf{x}_n)} - \mathbf{1}_{t < \sigma_{\mathcal{D}}^0(\mathbf{x})}| \geq \eta, t < \sigma_{\mathcal{D}}^0(\mathbf{x})) \\ &\quad + \mathbb{P}(|\mathbf{1}_{t < \sigma_{\mathcal{D}}^0(\mathbf{x}_n)} - \mathbf{1}_{t < \sigma_{\mathcal{D}}^0(\mathbf{x})}| \geq \eta, t > \sigma_{\mathcal{D}}^0(\mathbf{x})) \\ &= \mathbb{P}(|\mathbf{1}_{t < \sigma_{\mathcal{D}}^0(\mathbf{x}_n)} - 1| \geq \eta, t < \sigma_{\mathcal{D}}^0(\mathbf{x})) \\ &\quad + \mathbb{P}(|\mathbf{1}_{t < \sigma_{\mathcal{D}}^0(\mathbf{x}_n)}| \geq \eta, t > \sigma_{\mathcal{D}}^0(\mathbf{x})) \\ &= \mathbb{P}(\mathbf{1}_{t < \sigma_{\mathcal{D}}^0(\mathbf{x}_n)} = 0, t < \sigma_{\mathcal{D}}^0(\mathbf{x})) + \mathbb{P}(\mathbf{1}_{t < \sigma_{\mathcal{D}}^0(\mathbf{x}_n)} = 1, t > \sigma_{\mathcal{D}}^0(\mathbf{x})). \end{aligned}$$

Because  $\mathcal{O}$  is open, if  $t < \sigma_{\mathcal{D}}^0(\mathbf{x})$  (i.e.  $x_s^0(\mathbf{x}) \in \mathcal{O}$  for all  $s \in [0, t]$ ), there exists  $\beta > 0$  such that for any continuous curve  $\gamma : [0, t] \rightarrow \mathbb{R}^d$ , with  $\gamma(0) = x$  such that

$$\max \{|\gamma(s) - x_s^0(\mathbf{x})|, s \in [0, t]\} \leq \beta, \quad (6.36)$$

implies that

$$\{\gamma(s), s \in [0, t]\} \subset \mathcal{O}.$$

For  $n \geq 0$ , denote by  $\mathbf{d}_n(t) = \max \{|X_s^0(\mathbf{x}_n) - X_s^0(\mathbf{x})|, s \in [0, t]\}$  and pick  $\beta > 0$  such that (6.36) holds (and thus if  $\mathbf{d}_n(t) < \beta$ , then  $t < \sigma_{\mathcal{D}}^0(\mathbf{x}_n)$ ). Then, one has

$$\begin{aligned} \mathbb{P}(\mathbf{1}_{t < \sigma_{\mathcal{D}}^0(\mathbf{x}_n)} = 0, t < \sigma_{\mathcal{D}}^0(\mathbf{x})) &= \mathbb{P}(\mathbf{1}_{t < \sigma_{\mathcal{D}}^0(\mathbf{x}_n)} = 0, t < \sigma_{\mathcal{D}}^0(\mathbf{x}), \mathbf{d}_n(t) < \beta) \\ &\quad + \mathbb{P}(\mathbf{1}_{t < \sigma_{\mathcal{D}}^0(\mathbf{x}_n)} = 0, t < \sigma_{\mathcal{D}}^0(\mathbf{x}), \mathbf{d}_n(t) \geq \beta) \\ &= 0 + \mathbb{P}(\mathbf{1}_{t < \sigma_{\mathcal{D}}^0(\mathbf{x}_n)} = 0, t < \sigma_{\mathcal{D}}^0(\mathbf{x}), \mathbf{d}_n(t) \geq \beta) \\ &\leq \mathbb{P}(\mathbf{d}_n(t) \geq \beta) \\ &\leq \frac{\mathbb{E}[\mathbf{d}_n(t)^2]}{\beta^2} \leq \frac{C|\mathbf{x}_n - \mathbf{x}|^2}{\beta^2} \rightarrow 0 \text{ as } n \rightarrow +\infty, \end{aligned}$$

where we used [33, Lemma 3.3] (Assumption (A) there is satisfied for the process (6.1), see indeed (A $\Sigma$ )). Let us now deal with  $\mathbb{P}(\mathbf{1}_{t < \sigma_{\mathcal{D}}^0(\mathbf{x}_n)} = 1, t > \sigma_{\mathcal{D}}^0(\mathbf{x}))$ . Assume that  $t > \sigma_{\mathcal{D}}^0(\mathbf{x})$ . Let  $\alpha > 0$  be such that  $\sigma_{\mathcal{D}}^0(\mathbf{x}) + \alpha < t$ . From (6.28) and (6.29), and by the strong Markov property of the process (6.1), there exist almost surely  $u \in (0, \alpha)$  and  $r_u > 0$  such that

$$B(x_{\sigma_{\mathcal{D}}^0(\mathbf{x})+u}^0(\mathbf{x}), r_u) \subset \mathbb{R}^d \setminus \overline{\mathcal{O}}.$$

In particular, if  $\mathbf{d}_n(t) \leq r_u/2$  then  $\sigma_{\mathcal{D}}^0(\mathbf{x}_n) < \sigma_{\mathcal{D}}^0(\mathbf{x}) + u < t$  (and thus  $\mathbf{1}_{t < \sigma_{\mathcal{D}}^0(\mathbf{x}_n)} = 0$ ). One has

$$\begin{aligned} \mathbb{P}(\mathbf{1}_{t < \sigma_{\mathcal{D}}^0(\mathbf{x}_n)} = 1, t > \sigma_{\mathcal{D}}^0(\mathbf{x})) &= \mathbb{P}(\mathbf{1}_{t < \sigma_{\mathcal{D}}^0(\mathbf{x}_n)} = 1, t > \sigma_{\mathcal{D}}^0(\mathbf{x}), \mathbf{d}_n(t) \leq r_u/2) \\ &\quad + \mathbb{P}(\mathbf{1}_{t < \sigma_{\mathcal{D}}^0(\mathbf{x}_n)} = 1, t > \sigma_{\mathcal{D}}^0(\mathbf{x}), \mathbf{d}_n(t) > r_u/2) \\ &= 0 + \mathbb{P}(\mathbf{1}_{t < \sigma_{\mathcal{D}}^0(\mathbf{x}_n)} = 1, t > \sigma_{\mathcal{D}}^0(\mathbf{x}), \mathbf{d}_n(t) > r_u/2) \\ &\leq \frac{\mathbb{E}[\mathbf{d}_n(t)^2]}{(r_u/2)^2} \leq \frac{C|\mathbf{x}_n - \mathbf{x}|^2}{(r_u/2)^2} \rightarrow 0 \text{ as } n \rightarrow +\infty. \end{aligned}$$

This concludes the proof of (6.34).

**Step 2:** end of the proof of Proposition 6.6.

Pick  $f : \mathcal{D} \rightarrow \mathbb{R}$  a measurable and bounded function (i.e.  $f \in b\mathcal{B}(\mathcal{D})$ ). Extend  $f$  by 0 outside  $\mathcal{D}$ , so that  $f \in b\mathcal{B}(\mathbb{R}^{2d})$ . Let  $(\mathbf{x}_n)_n$  be sequence of elements of  $\mathcal{D}$  such that  $\mathbf{x}_n \rightarrow \mathbf{x} \in \mathcal{D}$  as  $n \rightarrow +\infty$ . From (6.27), it holds for  $\mathbf{x} \in \mathcal{D}$  and  $t > 0$ :

$$\mathbb{E}_{\mathbf{x}}[f(X_t)\mathbf{1}_{t < \sigma_{\mathcal{D}}} ] = \mathbb{E}^{\mathbb{P}}[f(X_t^0(\mathbf{x}))\mathbf{1}_{t < \sigma_{\mathcal{D}}^0(\mathbf{x})} \mathbf{M}_t(\mathbf{x})].$$

Let  $(\mathbf{x}_n)_n$  be sequence of elements of  $\mathcal{D}$  such that  $\mathbf{x}_n \rightarrow \mathbf{x} \in \mathcal{D}$  as  $n \rightarrow +\infty$ . Then, from the proof of [66, Proposition 1.2],  $f(X_t^0(\mathbf{x}_n)) \rightarrow f(X_t^0(\mathbf{x}))$  in  $\mathbb{P}$ -probability and that  $\mathbf{M}_t(\mathbf{x}_n) \rightarrow \mathbf{M}_t(\mathbf{x})$  in  $L^1(\Omega, \mathbb{P})$ , as  $n \rightarrow +\infty$ . Therefore, using (6.34), in  $\mathbb{P}$ -probability,

$$f(X_t^0(\mathbf{x}_n))\mathbf{1}_{t < \sigma_{\mathcal{D}}^0(\mathbf{x}_n)} \rightarrow \mathbf{1}_{t < \sigma_{\mathcal{D}}^0(\mathbf{x})}f(X_t^0(\mathbf{x})) \text{ as } n \rightarrow +\infty.$$

Thus,  $\mathbb{E}_{\mathbf{x}_n}[f(X_t)\mathbf{1}_{t < \sigma_{\mathcal{D}}}] \rightarrow \mathbb{E}_{\mathbf{x}}[f(X_t)\mathbf{1}_{t < \sigma_{\mathcal{D}}}]$  as  $n \rightarrow +\infty$ , that is  $P_t^{\mathcal{D}}$  is strong Feller for  $t > 0$ . This ends the proof of Proposition 6.6.  $\square$

**Remark 6.7.** If  $\mathcal{D} = \mathcal{O} \times \mathcal{V}$ , where  $\mathcal{V}$  is a smooth bounded subdomain of  $\mathbb{R}^d$ , we refer to [9] for the strong Feller property of  $P_t^{\mathcal{D}}$ .

**6.3. Quasi-stationary distributions for hypoelliptic damped Hamiltonian systems (6.1).** With all the previous results (see Lemma 6.2, Proposition 6.4, Lemma 6.5, and Proposition 6.6), one deduces from Theorem 2.2, the following theorem for the existence and uniqueness of the quasi-stationary distribution of the process (6.1) in  $\mathcal{D} = \mathcal{O} \times \mathbb{R}^d$ .

**Theorem 6.8.** *Assume that  $\Sigma$  satisfies  $(A\Sigma)$ . Let  $\mathcal{O}$  be a  $\mathcal{C}^2$  subdomain of  $\mathbb{R}^d$ , such that  $\mathbb{R}^d \setminus \overline{\mathcal{O}}$  is non-empty. Assume that*

- (1) *The functions  $V$  and  $c$  satisfy  $(Av1)$ ,  $(Av2)$ ,  $(Ac1)$ , and  $(Ac2)$ . Then, Theorem 2.2 is valid for the process (6.1) with  $\mathcal{D} = \mathcal{O} \times \mathbb{R}^d$  and with the Lyapunov function  $W_1$  defined in (6.8) (with parameters  $\mathbf{w} \in \mathcal{C}_c^2(\mathbb{R}^d, \mathbb{R})$ ,  $a > 0$ , and  $b > 0$  well chosen, see [66, Eq. (3.4)  $\rightarrow$  Eq. (3.9)] for explicit conditions on  $\mathbf{w}$ ,  $a$ , and  $b$ ). We refer to Remark 6.3 for concrete examples of functions  $V$  and  $c$  satisfying these assumptions.*
- (2) *The functions  $c$  and  $V$  satisfy respectively (6.2) and (6.3), and (6.9) holds. Then, Theorem 2.2 is valid for the process (6.1) with  $\mathcal{D} = \mathcal{O} \times \mathbb{R}^d$  and with the bounded Lyapunov function  $W_2$  defined in (6.12). Let us emphasize that since  $W_2$  is bounded (see item (2) in Proposition 6.4), item (b) in Theorem 2.2 holds and item (d) in Theorem 2.2 is satisfied for any initial distribution  $\nu$  in  $\mathcal{D}$ .*

For the uniqueness of the quasi-stationary distribution in the whole space of measures on  $\mathcal{D}$  in the specific case where  $\mathcal{O}$  is bounded,  $V$  is  $\mathcal{C}^\infty$  on  $\overline{\mathcal{O}}$ , and both  $\Sigma$  and  $c$  are constant, we refer to the very recent works [45, 44] (see also [56]). Let us mention that their approach<sup>6</sup> is different from the one we developed in this work. We finally refer to [51] for a spectral study of the kinetic Fokker-Planck operator on  $L^2(\mathcal{O} \times \mathbb{R}^d)$  when  $\mathcal{O}$  is bounded with several boundary conditions on  $\partial\mathcal{O} \times \mathbb{R}^d$  (see also [2, 38] and references therein).

**Acknowledgement.**

This work has been (partially) supported by the Project EFI ANR-17-CE40-0030 of the French National Research Agency. The authors thank P. Cattiaux, T. Lelièvre, M. Ramil, and Y. Xiao for fruitful exchanges during the preparation of this work.

REFERENCES

- [1] D. Aristoff and T. Lelièvre. Mathematical analysis of temperature accelerated dynamics. *Multiscale Modeling and Simulation*, 12(1):290–317, 2014.
- [2] S. Armstrong and J.C. Mourrat. Variational methods for the kinetic Fokker-Planck equation. *Preprint arXiv:1902.04037*, 2019.
- [3] D. Bakry, F. Barthe, P. Cattiaux, and A. Guillin. A simple proof of the Poincaré inequality for a large class of probability measures including the log-concave case. *Electron. Commun. Probab.*, 13:60–66, 2008.

---

<sup>6</sup>The study of the semigroup of the killed process is based on a precise analysis of the Cauchy problem  $\partial_t u = (\Delta_v + v \cdot \nabla_x - \nabla V \cdot \nabla_v - \gamma v \cdot \nabla_v)u$  with suitable boundary conditions on a subset of  $\partial\mathcal{O} \times \mathbb{R}^d$  and the use of Feynmann-Kac formula (compactness of the semigroup being obtained by Gaussian upper-bounds).

- [4] D. Bakry, P. Cattiaux, and A. Guillin. Rate of convergence for ergodic continuous Markov processes: Lyapunov versus Poincaré. *J. Funct. Anal.*, 254(3):727–759, 2008.
- [5] G. Ben Arous, M. Gradinaru, and M. Ledoux. Hölder norms and the support theorem for diffusions. In *Annales de l’IHP Probabilités et statistiques*, volume 30, pages 415–436, 1994.
- [6] M. Benaïm and B. Cloez. A stochastic approximation approach to quasi-stationary distributions on finite spaces. *Electronic Communications in Probability*, 20, 2015.
- [7] M. Benaïm, B. Cloez, and F. Panloup. Stochastic approximation of quasi-stationary distributions on compact spaces and applications. *Annals of Applied Probability*, 28(4):2370–2416, 2018.
- [8] A. Binder, T. Lelièvre, and G. Simpson. A generalized parallel replica dynamics. *Journal of Computational Physics*, 284:595–616, 2015.
- [9] P. Cattiaux. Calcul stochastique et opérateurs dégénérés du second ordre. II: Problème de Dirichlet. *Bulletin des sciences mathématiques (Paris. 1885)*, 115(1):81–122, 1991.
- [10] P. Cattiaux, P. Collet, A. Lambert, S. Martínez, S. Méléard, and J. San Martín. Quasi-stationary distributions and diffusion models in population dynamics. *Annals of Probability*, 37(5):1926–1969, 2009.
- [11] P. Cattiaux and A. Guillin. Hitting times, functional inequalities, Lyapunov conditions and uniform ergodicity. *J. Funct. Anal.*, 272(6):2361–2391, 2017.
- [12] P. Cattiaux, A. Guillin, F-Y. Wang, and L. Wu. Lyapunov conditions for super Poincaré inequalities. *J. Funct. Anal.*, 256(6):1821–1841, 2009.
- [13] P. Cattiaux and S. Méléard. Competitive or weak cooperative stochastic lotka–volterra systems conditioned on non-extinction. *Journal of Mathematical Biology*, 60(6):797–829, 2010.
- [14] N. Champagnat, K. A. Coulibaly-Pasquier, and D. Villemonais. Criteria for exponential convergence to quasi-stationary distributions and applications to multi-dimensional diffusions. In *Séminaire de Probabilités XLIX*, pages 165–182. Springer, 2018.
- [15] N. Champagnat and D. Villemonais. Exponential convergence to quasi-stationary distribution and Q-process. *Probability Theory and Related Fields*, 164(1-2):243–283, 2016.
- [16] N. Champagnat and D. Villemonais. General criteria for the study of quasi-stationarity. *Preprint arXiv:1712.08092*, 2017.
- [17] J-R. Chazottes, P. Collet, and S. Méléard. Sharp asymptotics for the quasi-stationary distribution of birth-and-death processes. *Probability Theory and Related Fields*, 164(1-2):285–332, 2016.
- [18] B. Cloez and M.N. Thai. Quantitative results for the Fleming–viot particle system and quasi-stationary distributions in discrete space. *Stochastic Processes and their Applications*, 126(3):680–702, 2016.
- [19] P. Collet, S. Martínez, S. Méléard, and J. San Martín. Quasi-stationary distributions for structured birth and death processes with mutations. *Probability Theory and Related Fields*, 151(1-2):191–231, 2011.
- [20] P. Collet, S. Martínez, and J. San Martín. *Quasi-stationary distributions: Markov chains, diffusions and dynamical systems*. Springer Science & Business Media, 2012.
- [21] P. Del Moral and L. Miclo. Particle approximations of lyapunov exponents connected to schrödinger operators and feynman–kac semigroups. *ESAIM: Probability and Statistics*, 7:171–208, 2003.
- [22] C. Dellacherie and P-A. Meyer. *Probabilités et potentiel*. Hermann, Paris, 1975. Chapitres I à IV, Édition entièrement refondue, Publications de l’Institut de Mathématique de l’Université de Strasbourg, No. XV, Actualités Scientifiques et Industrielles, No. 1372.
- [23] C. Dellacherie and P-A. Meyer. *Probabilités et Potentiel, chap. V-VIII(vol. 2): Théorie des Martingales*, Hermann, Paris, 1980.
- [24] G. Di Gesù, T. Lelièvre, D. Le Peutrec, and B. Nectoux. Jump markov models and transition state theory: the quasi-stationary distribution approach. *Faraday Discussions*, 195:469–495, 2017.
- [25] G. Di Gesù, T. Lelièvre, D. Le Peutrec, and B. Nectoux. Sharp asymptotics of the first exit point density. *Annals of PDE*, 5(2), 2019.

- [26] G. Di Gesù, T. Lelièvre, D. Le Peutrec, and B. Nectoux. The exit from a metastable state: concentration of the exit point distribution on the low energy saddle points, part 1. *Journal de Mathématiques Pures et Appliquées*, 138:242–306, 2020.
- [27] P. Diaconis and L. Miclo. On times to quasi-stationarity for birth and death processes. *Journal of Theoretical Probability*, 22(3):558–586, 2009.
- [28] P. Diaconis and L. Miclo. On quantitative convergence to quasi-stationarity. In *Annales de la Faculté des sciences de Toulouse: Mathématiques*, volume 24, pages 973–1016, 2015.
- [29] D. Down, S. P. Meyn, and R. L. Tweedie. Exponential and uniform ergodicity of Markov processes. *Ann. Probab.*, 23(4):1671–1691, 1995.
- [30] S. N. Ethier and T.G. Kurtz. *Markov processes: characterization and convergence*. John Wiley & Sons, 1986.
- [31] P. Ferrari and N. Maric. Quasi stationary distributions and Fleming-Viot processes in countable spaces. *Electronic Journal of Probability*, 12:684–702, 2007.
- [32] P.A. Ferrari, H. Kesten, S. Martinez, and P. Picco. Existence of quasi-stationary distributions. A renewal dynamical approach. *Annals of Probability*, 23(2):511–521, 1995.
- [33] A. Friedman. *Stochastic differential equations and applications*, volume 1. Academic Press, New York- London, 1975. Probability and Mathematical Statistics, Vol. 28.
- [34] N. Gantert. An inversion of Strassen’s law of the iterated logarithm for small time. *The Annals of Probability*, pages 1045–1049, 1993.
- [35] G. Gong, M. Qian, and Z. Zhao. Killed diffusions and their conditioning. *Probability Theory and Related Fields*, 80(1):151–167, 1988.
- [36] G. He, G. Yang, and Y. Zhu. Some conditional limiting theorems for symmetric Markov processes with tightness property. *Electronic Communications in Probability*, 24, 2019.
- [37] A. Hening, Q. Qi, Z. Shen, and Y. Li. Quasistationary distributions of multidimensional diffusion processes. *Preprint <https://sites.tufts.edu/hening/files/2020/10/QSD.pdf>*, 2020.
- [38] H.J. Hwang, J. Jang, and J.J.L. Velázquez. Nonuniqueness for the kinetic Fokker–Planck equation with inelastic boundary conditions. *Archive for Rational Mechanics and Analysis*, 231(3):1309–1400, 2019.
- [39] M. Kolb and D. Steinsaltz. Quasilimiting behavior for one-dimensional diffusions with killing. *Annals of Probability*, 40(1):162–212, 2012.
- [40] A. Lachal. Local asymptotic classes for the successive primitives of Brownian motion. *The Annals of Probability*, pages 1712–1734, 1997.
- [41] C. Le Bris, T. Lelièvre, M. Luskin, and D. Perez. A mathematical formalization of the parallel replica dynamics. *Monte Carlo Methods and Applications*, 18(2):119–146, 2012.
- [42] D. Le Peutrec and B. Nectoux. Repartition of the quasi-stationary distribution and first exit point density for a double-well potential. *SIAM Journal on Mathematical Analysis*, 52(1):581–604, 2020.
- [43] T. Lelièvre and F. Nier. Low temperature asymptotics for quasistationary distributions in a bounded domain. *Analysis & PDE*, 8(3):561–628, 2015.
- [44] T. Lelièvre, M. Ramil, and J. Reygner. A probabilistic study of the kinetic Fokker-Planck equation in cylindrical domains. *Preprint [arXiv:2010.10157](https://arxiv.org/abs/2010.10157)*, 2020.
- [45] T. Lelièvre, M. Ramil, and J. Reygner. Study of the quasi-distribution distribution for the Langevin diffusion process in cylindrical domains. *In preparation*, 2020.
- [46] T. Lelièvre and G. Stoltz. Partial differential equations and stochastic methods in molecular dynamics. *Acta Numerica*, 25:681–880, 2016.
- [47] M. Lladser and J. San Martín. Domain of attraction of the quasi-stationary distributions for the Ornstein-Uhlenbeck process. *Journal of Applied Probability*, 37(2):511–520, 2016.
- [48] J.C. Mattingly, A.M. Stuart, and D.J. Higham. Ergodicity for SDEs and approximations: locally Lipschitz vector fields and degenerate noise. *Stochastic processes and their applications*, 101(2):185–232, 2002.
- [49] S. Méléard and D. Villemonais. Quasi-stationary distributions and population processes. *Probability Surveys*, 9:340–410, 2012.

- [50] P. Meyer-Nieberg. *Banach Lattices*. Springer-Verlag, 1991.
- [51] F. Nier. *Boundary conditions and subelliptic estimates for geometric Kramers-Fokker-Planck operators on manifolds with boundaries*, volume 252. Memoirs of the American Mathematical Society, 2018.
- [52] E. Nummelin. *General irreducible Markov chains and non-negative operators*, volume 83. Cambridge University Press, 2004.
- [53] D. Perez, E. D Cubuk, A. Waterland, E. Kaxiras, and A.F. Voter. Long-time dynamics through parallel trajectory splicing. *Journal of chemical theory and computation*, 12(1):18–28, 2016.
- [54] D. Perez, B.P. Uberuaga, Y. Shim, J.G. Amar, and A.F. Voter. Accelerated molecular dynamics methods: introduction and recent developments. *Annual Reports in computational chemistry*, 5:79–98, 2009.
- [55] R.G. Pinsky. On the convergence of diffusion processes conditioned to remain in a bounded region for large time to limiting positive recurrent diffusion processes. *Annals of Probability*, pages 363–378, 1985.
- [56] M. Ramil. *Processus cinétiques dans les domaines bord et quasi-stationnarité*. PhD thesis, Ecole des Ponts ParisTech, 2020.
- [57] H.H. Schaefer. Banach lattices. In *Banach Lattices and Positive Operators*, pages 46–153. Springer, 1974.
- [58] M.R. Sorensen and A.F. Voter. Temperature-accelerated dynamics for simulation of infrequent events. *Journal of Chemical Physics*, 112(21):9599–9606, 2000.
- [59] D. Steinsaltz and S.N. Evans. Quasi-stationary distributions for one-dimensional diffusions with killing. *Transactions of the American Mathematical Society*, 359(3):1285–1324, 2007.
- [60] D.W. Stroock and S.R.S. Varadhan. On the support of diffusion processes with applications to the strong maximum principle. In *Proceedings of the Sixth Berkeley Symposium on Mathematical Statistics and Probability (Univ. California, Berkeley, Calif., 1970/1971)*, volume 3, pages 333–359, 1972.
- [61] M. Takeda. Existence and uniqueness of quasi-stationary distributions for symmetric markov processes with tightness property. *Journal of Theoretical Probability*, 32(4):2006–2019, 2019.
- [62] D. Talay. Stochastic Hamiltonian systems: exponential convergence to the invariant measure, and discretization by the implicit Euler scheme. *Markov Processes and Related Fields*, 8(2):163–198, 2002.
- [63] D. Villemonais. General approximation method for the distribution of markov processes conditioned not to be killed. *ESAIM: Probability and Statistics*, 18:441–467, 2014.
- [64] A.F. Voter. A method for accelerating the molecular dynamics simulation of infrequent events. *Journal of Chemical Physics*, 106(11):4665–4677, 1997.
- [65] A.F. Voter. Parallel replica method for dynamics of infrequent events. *Physical Review B*, 57(22):R13 985, 1998.
- [66] L. Wu. Large and moderate deviations and exponential convergence for stochastic damping Hamiltonian systems. *Stochastic Processes and their Applications*, 91(2):205–238, 2001.
- [67] L. Wu. Essential spectral radius for Markov semigroups (I): discrete time case. *Probability Theory and Related Fields*, 128(2):255–321, 2004.
- [68] K. Yosida. *Functional analysis*, 1980. Springer-Verlag, New York/Berlin, 1971.
- [69] J. Zhang, S. Li, and R. Song. Quasi-stationarity and quasi-ergodicity of general Markov processes. *Science China Mathematics*, 57(10):2013–2024, 2014.

**Arnaud Guillin.** UNIVERSITÉ CLERMONT AUVERGNE, CNRS, LMBP, F-63000 CLERMONT-FERRAND, FRANCE

*E-mail address:* `arnaud.guillin@uca.fr`

**Boris Nectoux.** UNIVERSITÉ CLERMONT AUVERGNE, CNRS, LMBP, F-63000 CLERMONT-FERRAND, FRANCE

*E-mail address:* `boris.nectoux@uca.fr`

**Liming Wu.** UNIVERSITÉ CLERMONT AUVERGNE, CNRS, LMBP, F-63000 CLERMONT-FERRAND, FRANCE

*E-mail address:* `Li-Ming.Wu@uca.fr`