Recent advances in the long-time analysis of killed degenerate processes and their particle approximation

Bertrand Cloez, Lucas Journel, Pierre Monmarché, Boris Nectoux, Mouad Ramil

RECENT ADVANCES IN THE LONG-TIME ANALYSIS OF KILLED DEGENERATE PROCESSES AND THEIR PARTICLE APPROXIMATION

Bertrand Cloez[1], Lucas Journel[2], Pierre Monmarche[3], Boris Nectoux[4] and Mouad Ramil[5]

Abstract. We review some recent results of quantitative long-time convergence for the law of a killed Markov process conditioned to survival toward a quasi-stationary distribution, and on the analogous question for the particle systems used in practice to sample these distributions. With respect to the existing literature, one of the novelties of these works is the degeneracy of the underlying process with respect to classical elliptic diffusion, namely it can be a non-elliptic hypoelliptic diffusion, a piecewise deterministic Markov process or an Euler numerical scheme.

Résumé. Nous présentons quelques résultats récents d'estimées quantitatives de convergence en temps long pour la loi de processus de Markov tués, conditionnellement à leur survie, vers une mesure quasi-stationnaire, et sur la question analogue pour les systèmes de particules utilisés en pratique pour échantillonner ces mesures. Par rapport à la littérature antérieure, l'une des nouveautés de ces travaux est le caractère dégénéré du processus de Markov sous-jacent par rapport aux diffusions elliptiques classiques, au sens où ce peut être une diffusion hypoelliptique non-elliptique, un processus déterministe par morceaux, ou un schéma d'Euler.

1. Introduction

1.1. **Killed Markov processes**

A killed process on a space $\mathcal{D}$ is a Markov process $(X_t)_{t \geqslant 0}$ on $\mathcal{D} \cup \{\partial\}$ where $\partial \notin \mathcal{D}$ is an arbitrary cemetery point which is an absorbing point, i.e. $X_t = \partial$ for all $t \geqslant \tau = \inf\{s \geqslant 0, \ X_s = \partial\}$. In many cases of interest, it is obtained from a Markov process $(Y_t)_{t \geqslant 0}$ on a state $\mathcal{S}$ by setting $X_t = Y_t$

[1] MISTEA, Université Montpellier, INRAE, Institut Agro, Montpellier, France

[2] LJLL, Sorbonne Université, Paris

[3] LJLL and LCT, Sorbonne Université, Paris

[4] LMBP, Université Clermont Auvergne

[5] CERMICS, Ecole des Ponts, Marne-la-Vallée, France

for $t < \tau$ and $X_t = \partial$ for $t \geqslant \tau$ where τ is defined either by a so-called soft-killing mechanism, namely, given E a standard exponential random variable independent from Y,

$$\tau = \inf \left\{ t \geqslant 0, \ E \leqslant \int_0^t \lambda(Y_s)\mathrm{d}s \right\} ,$$

with a death rate $\lambda : \mathcal{S} \to \mathbb{R}_+$, in which case $\mathcal{D} = \mathcal{S}$, or a so-called hard-killing mechanism, namely

$$\tau = \inf \{ t \geqslant 0, \ Y_t \notin \mathcal{D} \} , \tag{1}$$

where $\mathcal{D}$ is a given domain of $\mathcal{S}$. Since we are only interested in the process before absorption, in these cases it is equivalent to work with either X or Y.

Some objects of interest are then, for all $t \geqslant 0$, $p_t = \mathbb{P}(\tau \leq t)$ the extinction probability, η_t the law of the killed process and μ_t the law of the process conditioned to survival, given by

$$\eta_t(A) = \mathbb{P}(X_t \in A) , \qquad \mu_t(A) = \mathbb{P}(X_t \in A \mid \tau > t) = \frac{\eta_t(A)}{1 - p_t} \tag{2}$$

for all measurable sets A of $\mathcal{D}$.

In most cases, extinction is almost sure, i.e. $\tau < +\infty$ almost surely. In that case, the only invariant measure of X is the Dirac mass at ∂, which is not interesting. In fact, for killed processes, the relevant analogous to invariant measures is given by the quasi-stationary distributions (QSD), which by definitions are the probability measures ν on $\mathcal{D}$ such that

$$X_0 \sim \nu \quad \Rightarrow \quad \forall t \geqslant 0, \ \mu_t = \nu .$$

Similarly, the classical question of convergence toward equilibrium for Markov processes is replaced, for killed process, by the question of the limit of μ_t toward a QSD as $t \to +\infty$ when the initial distribution is not a QSD.

An important property is that, starting from a QSD, the extinction time follows an exponential distribution. In other words, if ν is a QSD, then there exists $\theta > 0$ such that, if $X_0 \sim \nu$, then $p_t = 1 - e^{-\theta t}$ for all $t \geqslant 0$. Equivalently, $\eta_t = e^{-\theta t}\nu$ for all $t \geqslant 0$, which means that ν is an eigenvector of the killed semi-group. In fact, when $\mathcal{D}$ is a finite set, the existence of a QSD stems from the Perron-Frobenius theorem.

For general and extensive references on QSD, we refer the interested reader to the book [22] or the survey [55].

1.2. **Some motivations**

Among other things, the study of killed processes and quasi-stationary distributions is related to metastability phenomena, i.e. cases where a Markov process stays for very long times in some transient states before reaching stationarity. The convergence to a QSD within a transient state is fast and observed before the slow transitions. This arises in particular in population models and stochastic algorithms in molecular dynamics, as we describe in the following.

In population dynamics, a Markov process $(X_t)_{t\geq0}$ is generally used to model the number of individuals. This can for example be a diffusion process on a bounded set as in the case of the evolution of gene populations in the continuous Wright-Fischer model [10, Page 368] or a continuous-time Markov chain as for a Galton-Watson process [1, Chapter III]. An interesting example is given
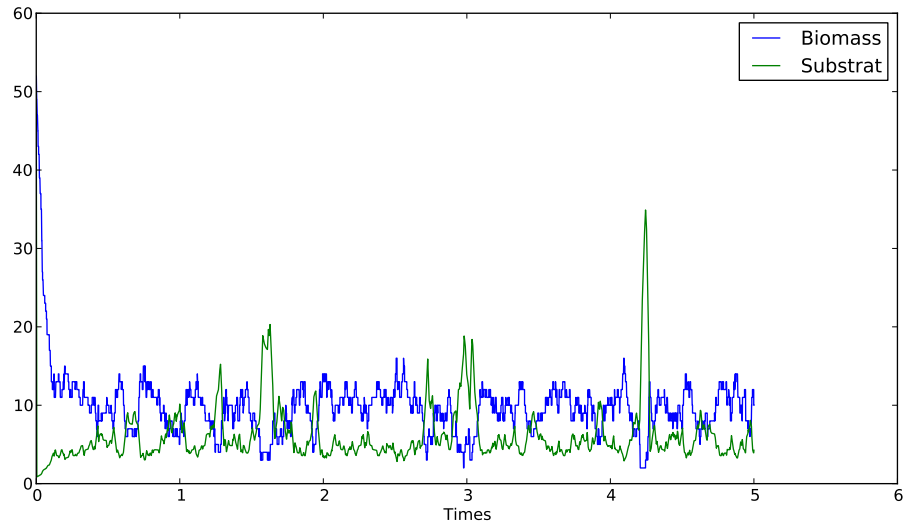
FIGURE 1. A sample path of the chemostat model. Here the blue curve corresponds to the discrete component (N_t) and the green curve corresponds to the continuous component (S_t).

by the chemostat model [17, 21] where $X_t = (N_t, S_t) \in \mathbb{N} \times \mathbb{R}_+$ is a couple of two components: a discrete one N_t and a continuous one S_t. The discrete component represents the number of bacteria in a vessel and evolves as a pure jump process with the addition or removal of an individual at exponential clocks. Parameters of the jump times depend on the continuous component which evolves according an ordinary differential equation whose parameters depend on the discrete component. See [17, 21] for details and Figure 1 for an example of trajectory. This process is not irreducible, belongs to a non-compact space and is hybrid, in the sense that its transition kernel does not admit a density with respect to a simple reference measure. Whatever the parameters, the population of bacteria always hits 0 in finite time (as a consequence of a Borel-Cantelli type arguments since resources are finite), however, when the vessel is large the dynamics is close to a deterministic model having a non-trivial equilibrium. These two antagonistic behaviors, which are common for population models, should be explained by a quasi-stationary behavior.

As we shall see here, the study of QSD can also be useful in molecular dynamics methods for the sampling of the time-evolution of complex particle systems. This sampling is done using numerical schemes which are used in a wide range of application fields (biology, chemistry, materials science for instance). They allow us to compute dynamical quantities of interest for the system without resorting to the actual experiment. For instance, one would like to sample the trajectories of a system which goes from one state to another one, in order to compute the typical time to observe such transitions. This question arises for instance when we want to sample the phase transitions of a thermostated (fixed temperature) molecular system, whose evolution can be modeled by the

Langevin dynamics:

$$\begin{cases} \mathrm{d}x_t = M^{-1}v_t\mathrm{d}t, \\ \mathrm{d}v_t = F(x_t)\mathrm{d}t - \gamma M^{-1}v_t\mathrm{d}t + \sqrt{2\gamma\beta^{-1}}\mathrm{d}B_t, \end{cases} \tag{3}$$

where $(x_t, v_t) \in \mathbb{R}^d \times \mathbb{R}^d$ denotes the positions and momenta of the particles at time $t \geq 0$, M is the mass matrix, F is the interaction force (often F is conservative, i.e. $F = -\nabla V$ for some potential function V), $\gamma > 0$ is a friction parameter, $\beta = (k_B T)^{-1}$ is proportional to the inverse temperature and $(B_t)_{t\geq 0}$ is a standard Brownian motion. When the friction coefficient γ goes to infinity, it is known that its rescaled position coordinate $(x_{\gamma t})_{t\geq 0}$ converges to the solution $(\overline{x}_t)_{t\geq 0}$ of the overdamped Langevin Dynamics (see for instance [48, Section 2.2.4]):

$$\mathrm{d}\overline{x}_t = F(\overline{x}_t)\mathrm{d}t + \sqrt{2\beta^{-1}}\mathrm{d}B_t, \tag{4}$$

Phase transitions are known to be often metastable, which means that the timestep used for the sampling of the dynamics is much smaller than the timescale associated with the transition event. As a practical consequence, the number of iterations needed to sample the transition event will be far too large to be done through naive simulation. The metastability comes from the existence of energetic barriers, namely the trajectory to leave the state requires to climb above a saddle point of the potential energy V.

Several methods have been developed in order to circumvent this issue, among which the Parallel Replica method [62]. The idea behind Parallel Replica is that, as already mentioned, when the system remains trapped inside a metastable state for a sufficiently long time, its distribution becomes close to the QSD of the state. Furthermore, starting from the QSD, the first exit from the state satisfies some well-known properties which allow for the sampling in parallel of the first exit event, thus reducing the wall clock time simulation in Parallel Replica compared to standard simulation.

The main ingredients for the mathematical justification of Parallel Replica are the existence of a QSD in the state considered, as well as a long time convergence of the dynamics trapped inside the state towards the QSD. Such proofs were carried out in the literature in the case of the overdamped Langevin (4) dynamics in [45] but had yet to be extended to the Langevin dynamics (3).

1.3. **Past results and limitations**

As far as continuous-time processes in a continuous space are concerned, there is a wide literature on quasi-stationary distributions for elliptic diffusions like the overdamped Langevin dynamics (4) on smooth bounded domains. We refer for example to [8, 9, 11, 15, 35, 45, 57]. Let us state a theorem to summarize the main results known for the overdamped Langevin.

The infinitesimal generator of the overdamped Langevin dynamics is given by:

$$\overline{\mathcal{L}} = F \cdot \nabla + \beta^{-1}\Delta, \tag{5}$$

with formal adjoint $\overline{\mathcal{L}}^*$ in $\mathrm{L}^2(\mathrm{d}x\mathrm{d}v)$ given by:

$$\overline{\mathcal{L}}^* = -\mathrm{div}(F\cdot) + \beta^{-1}\Delta.$$

The following is based on [11, 35, 41, 45].

Theorem 1 (QSD of the overdamped Langevin process)**.** *Let $\beta > 0$. Assume that $F \in \mathcal{C}^\infty(\mathbb{R}^d, \mathbb{R}^d)$ and let $\mathcal{O}$ be an open $\mathcal{C}^2$ bounded connected set of $\mathbb{R}^d$. Then there exists a unique QSD $\overline{\mu}$ on $\mathcal{O}$ for the killed process in $\mathcal{O}$, $(\overline{X}_t)_{t \geq 0}$, solving* (4)*. Furthermore,*

   (1) *there exists $\overline{\psi} \in \mathcal{C}^2(\mathcal{O}) \cap \mathcal{C}^b(\overline{\mathcal{O}})$ such that $\overline{\mu}(\mathrm{d}q) = \overline{\psi}(q)\mathrm{d}q$, where $\mathrm{d}q$ is the Lebesgue measure on $\mathbb{R}^d$,*

   (2) $\mathrm{Span}(\overline{\psi})$ *is the eigenspace associated with the smallest eigenvalue $\overline{\theta}$ of the operator $-\overline{\mathcal{L}}^*$ with homogeneous Dirichlet boundary conditions on $\partial\mathcal{O}$,*

   (3) *there exist $C > 0$ and $\alpha > 0$ such that for all probability measures ν on $\mathcal{O}$, for all $t \geq 0$,*

$$\left\| \overline{\mu}_t - \overline{\mu} \right\|_{TV} \leq C\mathrm{e}^{-\alpha t},$$

   where $\overline{\mu}_t$ is the law of the process $(\overline{X}_t)_{t \geq 0}$ conditioned to survival, as defined in (2)*, satisfying $\overline{X}_0 \sim \nu$ and $\|\cdot\|_{TV}$ is the total-variation norm on the space of bounded signed measures on $\mathbb{R}^d$.*

Now, additional difficulties arise when the process is not an elliptic diffusion process, which is the case of the processes mentioned in the previous section: the kinetic Langevin dynamics (3) is an hypoelliptic non-elliptic diffusion, and the chemostat model of [17, 21] is a piecewise deterministic Markov process, lacking the regularization properties of diffusions. Moreover, the study of hard-killed process with an unbounded domain (which is the case for the Langevin process since the velocity is unbounded) is also difficult. For instance even for the very simple Ornstein-Uhlenbeck process killed at the origin, there is no uniqueness of the QSD [50].

Multiple approaches were used in the literature to obtain Theorem 1. To name just one, one can apply the Krein-Rutman theorem to the semigroup of (4) killed outside of $\mathcal{O}$, which is compact on $\mathrm{L}^2(\mathcal{O})$, to obtain the QSD. However, such compactness property is not clear once the infinitesimal generator is not elliptic or the domain is unbounded as is the case for (3) on the domain $\mathcal{D} = \mathcal{O} \times \mathbb{R}^d$. Two approaches are presented in Sections 3 and 4 to establish this property in cases that cover the kinetic Langevin process (3). Section 2 presents an alternative method, based on a Doeblin-type minorization condition rather than compactness, to get a result similar to the Krein-Rutman theorem in non-compact space. As can be seen by comparing the general results stated in Theorems 2 and 6, the two methods share some similarities (a Lyapunov function is required in both cases) but each have their interest: the conditions of Theorem 6 are easier to check in practice, but the process is required to be Feller, which is not the case for Theorem 2, while in the latter a non-trivial minorization condition has to be established.

Finally, notice that, when it comes to sample the QSD (for instance as the first step of the Parallel Replica algorithm), a naive Monte Carlo method, consisting of running N independent copies of the process up to a long time t and taking the empirical distribution of the processes which have not been killed, may suffer from a high variance since the size of the sample shrinks with times (possibly, all the processes may have already died at time t). For this reason, systems of interacting particle are used in practice: when a process dies, another process is chosen at random uniformly and is duplicated, which maintains the size of the sample, at the cost of a small bias. The particle system is then a non-killed Markov process, which converges in long time to an invariant measure which is close, for large N, to the law of N independent realizations of the QSD. Although many results are available in various frameworks for the convergence of the system of particles toward the killed process as $N \to \infty$, either at a fixed time t or at stationary (see [16, 23] and references within) obtaining quantitative convergence as $t \to \infty$ uniformly in N is challenging and cannot be obtained directly from the long-time convergence of the killed process. Such a result, in

the case of an elliptic diffusion, is presented in Section 5. The effect of the time-discretization of the continuous-time diffusion, which is a very classical topic for the invariant measure of Markov processes but not for QSD, is also addressed, closing the gap between the theoretical results such as Theorem 1 and the algorithms used in practice to sample the QSD.

2. Harris-type method for non-conservative processes

2.1. Harris theorem type assumption for Krein-Rutman theorem type conclusion

To study (η_t) or (μ_t) defined in (2), it is usual to introduce the family of operators (M_t) defined for all $t \geq 0$ by

$$M_t f(x) = \mathbb{E}\left[f(X_t)\mathbf{1}_{\tau > t} \mid X_0 = x\right],$$

where f is a bounded type function and $x \in \mathcal{D}$. It is easy to show that (M_t) defines a positive semigroup, but in contrast with Markov semigroup, it is not conservative in the sense that

$$M_t \mathbf{1}_{\mathcal{D}} = 1 - p_t \neq 1$$

in general, where $\mathbf{1}$ is the constant function equal to 1.

This particularity is a drawback because there is a rich base of tools to study conservative semigroups as for instance Doeblin contraction or Lyapunov methods. More precisely, a Markov semigroup (P_t) satisfies the Doeblin minorization condition if there exists $t_0, \epsilon > 0$ and a probability measure such that for all $x \in X$

$$\delta_x P_{t_0} \geq \epsilon \nu. \tag{6}$$

This induces uniqueness of an invariant distribution and exponential convergence to it in total variation distance. When the state is not compact Assumption (6) generally does not apply on the whole space. Consequently, it is only supposed to hold on compact sets and an additional assumption is required to prove that the process remains in a compact set for sufficiently long times. This tightness condition is often verified through the existence of a so-called Lyapunov function: there exists a positive function V, such that for a certain $t > 0$,

$$P_t V \leq CV + D,$$

for some constants $C, D > 0$ with $C < 1$. This is often verified by the drift condition:

$$\mathcal{L}V \leq -cV + d, \tag{7}$$

where $\mathcal{L}$ is the generator of (P_t) and $c, d > 0$. With this assumption, Doeblin minorization only need to hold on the sublevel set of V to obtain exponential convergence to a unique invariant measure.

For conservative semigroups, these aspects are well known; see for instance [28, 38, 53, 54].

Several works aimed at generalizing these efficient techniques for non-conservative semi-groups. Hilbert metric and Birkhoff contraction yield another powerful method for the analysis of semigroups, which has been well developed [7, 56, 59]. More recently, let us cite for instance [24, 43, 44] in a general and continuous-time setting or works on discrete-time Feynman-Kac semigroups [23, 26]. Some works associate Doeblin-Harris techniques with Krein-Rutmann theorem and then often need strong Feller properties [31, 36, 46]; See Section 3 and 4. Unfortunately, these results do not apply

or are difficult to apply for piecewise deterministic dynamics as e.g. the chemostat model described in introduction.

In [2, 3, 18], some theoretical results were developed to generalize Doeblin-Harris methods for non-conservative semi-groups mainly in case of branching dynamics (i.e. in case where $M_t \mathbf{1} > \mathbf{1}$). These results are related to [12, 13, 20, 21, 25]. A generalization of Harris Theorem is for instance described in [3, Theorem 2.1] and is given below.

Theorem 2 (Generalized Harris Theorem). *(i) If there exist a couple of positive functions (V, ψ), $t_0 > 0$, $\beta > \alpha > 0$, $C > 0$, $(c, d) \in (0, 1]^2$, a subset K and ν a probability supported by K such that $\sup_K V/\psi < \infty$ and*

(A1) $M_{t_0} V \leq \alpha V + C \mathbf{1}_K \psi$,

(A2) $M_{t_0} \psi \geq \beta \psi$,

(A3) for all positive and measurable function f,

$$\inf_{x \in K} \frac{M_{t_0}(f\psi)(x)}{M_{t_0}\psi(x)} \geq c\,\nu(f)$$

(A4) for all positive integers n

$$\nu\left(\frac{M_{nt_0}\psi}{\psi}\right) \geq d \sup_{x \in K} \frac{M_{nt_0}\psi(x)}{\psi(x)},$$

then, there exists a unique triplet (γ, h, θ) of eigenelements of M with $\gamma(h) = 1$, i.e. satisfying for all $t \geq 0$

$$\gamma M_t = \mathrm{e}^{-\theta t}\gamma \qquad and \qquad M_t h = \mathrm{e}^{-\theta t}h. \tag{8}$$

Moreover, there exist $C, \omega > 0$ such that for all $t \geq 0$ and measure μ,

$$\sup_{|f/V \leq 1| \leq 1} \left| e^{\theta t}\mu M_t f - \mu(h)\gamma(f) \right| \leq C\mu(V)e^{-\omega t}. \tag{9}$$

(ii) Assume that there exist a positive measurable function V, a triplet (γ, h, θ) and constants $C, \omega > 0$ such that (8) and (9) hold. Then, the couple (V, h) satisfies Assumptions (A1),(A2),(A3), and (A4).

By differentiating (8), the triplet (γ, h, θ) is a triplet of eigenelements for the infinitesimal generator $\mathcal{A}$ of $(M_t)_{t \geq 0}$, that is $\gamma \mathcal{A} = -\theta\gamma$ and $\mathcal{A}h = -\theta h$ where the (unbounded) operator $\mathcal{A}$ is defined by $\mathcal{A} = \lim_{t \to 0} \frac{1}{t}(M_t - I)$.

The assumptions of Theorem 2 hold on compact space when (M_t) admits a (upper and lower) bounded density with respect to a fixed measure ν. It is known since the 60's, by Birkhoff [7] for instance, that this type of regularity hypothesis implies the existence of eigen-elements and the convergence. Unfortunately, assuming bounded density does not include simple examples such as the chemostat model.

Assumption (A4) overcomes this problem and allows to treat such an example. This type of condition seems to appear for the first time in [25] without the Lyapunov condition. Article [25] only states a contraction inequality, but this is the main step of the proof for such results. In [12],
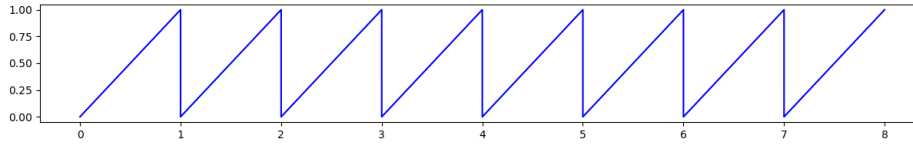
FIGURE 2. A sample path of the process described in Equation 11.

Champagnat and Villemonais highlight these conditions and prove that they are in fact equivalent to the exponential convergence (9). They deduce the existence of eigenelements in the context of absorbed Markov processes (sub-conservative semigroups). In [2], this theorem is extended for general inhomogeneous-time and non-conservative semi-groups. It includes for instance periodic and asymptotically homogeneous ones.

As explained before, requiring a Doeblin type minorization assumption on the whole space is often too strong when the state space is not compact, Assumptions (A1) and (A2) then relax this constraint. This Lyapunov assumption seems to appear for the first time in [20, Theorem 4.2] and in the proof of [21, Theorem 4.1]. However, combining this assumption with the assumptions (A3) and (A4) to obtain uniqueness of eigenelements and exponential convergence comes from [13]. This significant result has been demonstrated in the framework of sub-conservative semigroups. Result [3, Theorem 2.1] generalizes this result for non-conservative semi-groups, and non-bounded functions ψ, and prove that these assumptions are in fact necessary. Let us also point out the follow-up paper [14] of [3, 13] on these aspects. However, Assumption (A4) of Theorem 2 seems simpler to verify that the analogous of [14] and the proof of [3] leads some quantitative estimates. More precisely, we have that $\gamma(V) < +\infty$, $\gamma \gg \nu$ (in the Radon-Nikodym sense),

$$\beta \le e^{-\theta t_0} \le \alpha + C,$$

and

$$c_1 \left(\frac{\psi}{V} \right)^q \psi \le h \le c_2 V$$

where $c_1, c_2 > 0$, $q \in (0,1)$ and the others constants come from Theorem 2.

To conclude this section let us mention that from [18], Condition (A4) is verified when there exist $t_0, \epsilon > 0$ such that

$$\inf_{x,y \in K} \mathbb{P}\left(\exists s \le t_0, X_s = y \mid X_0 = x\right) \ge \epsilon. \tag{10}$$

This condition will be verified in almost all examples of Subsection 2.2

2.2. **Some (very) simple examples**

In this section, we detail several examples where one of the assumptions does not hold in order to highlight their role.

Note however that all assumptions are naturally connected and that one can postpone the difficulty of an assumption by changing the compact K or the Lyapunov functions for instance.
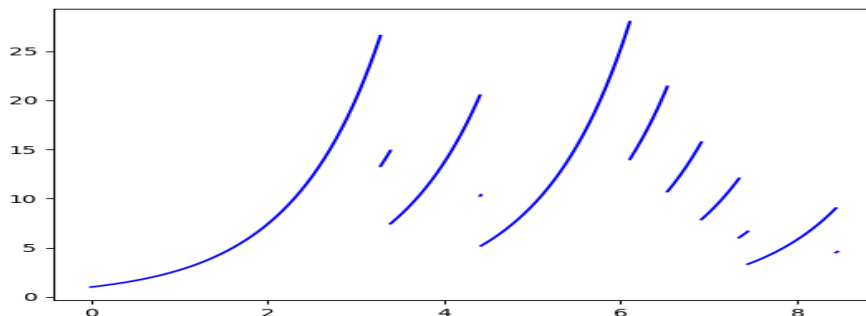
FIGURE 3. A sample path of the process generated by (12).

2.2.1. *Aperiodicity and Assumption (A3)*

In addition to the irreducibly assumption, Assumption (A3) hides an aperiodicity assumption. Indeed the (unabsorbed) Markov process $(X_t)_{t\geq0}$ on $[0,1)$ defined, for all $t \geq 0$, by

$$X_t = X_0 + t \bmod 1 \tag{11}$$

verifies all the assumptions except (A3). Even if the trajectories pass into all the points of $[0,1)$, it does not pass at the same time t_0 uniformly over its stating point. See Figure 2.

A less simple and more interesting example is given by the growth-fragmentation process studied in [6, 34] (in a different formalism). This process grows exponentially at a certain rate α and has multiplicative jumps from x to rx, where $r \in (0,1)$ is a fixed deterministic value, at a rate $B(x)$ depending on its state. We can kill the process at some rate λ as defined in Equation 1. Its generator reads

$$\mathcal{A}f(x) = \alpha x f'(x) + B(x)(f(rx) - f(x)) + \lambda(x)f(x), \tag{12}$$

for smooth functions f and $x > 0$. See Figure 3

In contrast with the previous trivial deterministic process, this growth-fragmentation has a random dynamics but its law is concentrated in fixed time in a discrete space (when the starting distribution is a Dirac mass). It then can not converge to its unique QSD, which has a Lebesgue density.

2.2.2. *Probability of survival and Assumption (A4)*

Let us consider the Markov process which is maybe the simplest one. That is, we consider that $\mathcal{D} = \{1, 2\}$ and X_t jumps from 2 to 1 at rate a and cannot jump from 1 to 2. Similarly, it dies from 1 at rate b and cannot die from 1. It is a pure death process whose transition are represented in Figure 4. Its study is trivial and based on the spectral property of a 2×2 matrix. We have three cases :

- $b < a$: We have exponential convergence to the unique QSD δ_2.
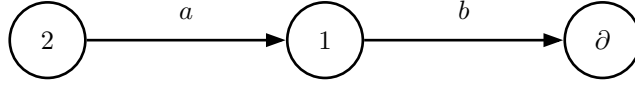- $b = a$: We have a slow convergence to the unique QSD δ_2.

FIGURE 4. The transition diagram of the Markov chain described in Subsection 2.2.2

- $b > a$: We have two QSD:

$$\frac{a}{b}\delta_1 + \frac{b-a}{b}\delta_2, \qquad \delta_2,$$

and, furthermore, the convergence depends on the initial condition : if $\mu_0(\{2\}) = \eta(\{2\}) > 0$ then the conditional law μ_t converges to $\frac{a}{b}\delta_1 + \frac{b-a}{b}\delta_2$.

We see in the third case that $\delta_2 M_t \mathbf{1} \sim e^{-at}$ and $\delta_1 M_t \mathbf{1} \sim e^{-bt}$ are not of the same order and consequently (A4) fails and then the conclusion of Theorem 2 doesn't hold (which is what we naturally and simply observe here).

This process is trivial to study but it is an interesting example to understand the algorithms which approximate the quasi-stationary distribution. For instance the limit in large number of particles and in time of the well used Fleming-Viot algorithm do not commute for this example. Some question on its approximation remains open see [5]. Commutation of limits for QSD approximation, as detailed in Section 5 below is then of prime importance.

A less simple and more interesting example is given by the house-of-card model studied in [19] and references therein. This process belongs to $[0, 1]$ and is redrawn uniformly at random over this interval at rate 1. It is soft-killed at rate $x \mapsto cx^q$, for some $c, q \geq 0$. We can show the following :

- If $q > 1 - 1/c$ then we have exponential convergence to a unique explicit QSD (which has a bounded Lebesgue density). This in particular always verified when $q \geq 1$ or $c < 1$. The convergence can be established in various distances with explicit rates.
- If $c = q+1$ then we have two QSD : one, denoted by ν, has a Lebesgue density proportional to $x \mapsto x^{-q}$ and the second is δ_0. When $\mu_0(\{0\}) = 0$, then we have a (explicit) slow convergence of the conditional laws μ_t to the QSD ν (which has a Lebesgue density proportional to $x \mapsto x^{-q}$). When $\mu_0(\{0\}) > 0$ then $\mu_t \to \delta_0$. Moreover, if $q < 1/2$ then $\eta_t e^{\theta t} \to c_0 \nu$, for some $c_0 > 0$ although if $q \geq 1/2$ then $\eta_t e^{\theta t} \to 0$.
- If $c < q+1$ then the conditional law converges, in mean, to the unique degenerate QSD:

$$\nu(dx) = \left(1 - \frac{c}{q+1}\right)\delta_0(dx) + \frac{x^{-q}}{c}dx$$

For details, additional results and references, see [19]. Note that, in this example, Assumption (10) fails.

2.2.3. *Lyapunov function and Assumption (A1) and (A2)*

The Lyapunov assumptions require to find a kind of sub-solution and super-solution to the eigenvalue problem where the associated sub-eigenvalue β is strictly larger than the super-eigenvalue α. Let us show here a well-studied example of Markov processes where we easily find two Lyapunov functions in such a way that the condition $\alpha = \beta$ becomes a critical situation.

Let us consider the birth and death process on $\mathbb{N}^*$ generated by

$$\mathcal{A}f(n) = b(f(n+1) - f(n)) + d(f(n-1) - f(n)),$$

for every $n \geq 2$ and bounded function f, with $b > d$, and

$$\mathcal{A}f(1) = b_1(f(2) - f(1)) - d_1 f(1),$$

with some $b_1, d_1 > 0$. It is known from [60] that $e^{\theta t}\eta_t$ converge to a unique QSD if and only if

$$(\sqrt{b} - \sqrt{d})^2 + b_1\left(\sqrt{d/b} - 1\right) - d_1 > 0.$$

This result can easily be recovered (and extended) by using Theorem 2 with V and ψ of the form $q \mapsto q^n$ with appropriate $q > 0$. See [3] for details. Note that the condition (10) is trivially satisfied.

3. THE LANGEVIN PROCESS (ROUND 1)

Let $F \in \mathcal{C}^\infty(\mathbb{R}^d, \mathbb{R}^d)$, $\gamma \in \mathbb{R}$, $\sigma > 0$. In this section, we shall consider the following Langevin dynamics:

$$\begin{cases} \mathrm{d}x_t = v_t \mathrm{d}t, \\ \mathrm{d}v_t = F(x_t)\mathrm{d}t - \gamma v_t \mathrm{d}t + \sigma \mathrm{d}B_t, \end{cases} \tag{13}$$

Its infinitesimal generator is given for $(x, v) \in \mathbb{R}^d \times \mathbb{R}^d$ by:

$$\mathcal{L} = v \cdot \nabla_x + F(x) \cdot \nabla_v - \gamma v \cdot \nabla_v + \frac{\sigma^2}{2}\Delta_v, \tag{14}$$

with formal adjoint $\mathcal{L}^*$ in $\mathrm{L}^2(\mathrm{d}x\mathrm{d}v)$:

$$\mathcal{L}^* = -v \cdot \nabla_x - F(x) \cdot \nabla_v + \gamma \mathrm{div}_v(v\cdot) + \frac{\sigma^2}{2}\Delta_v. \tag{15}$$

The metastable states often correspond to the basins of attraction of a potential V when the force F in (3) is conservative, i.e. $F = -\nabla V$. Therefore, we shall consider metastable states written as cylinders in the position and momenta coordinates as follows:

$$\mathcal{D} = \mathcal{O} \times \mathbb{R}^d, \tag{16}$$

where $\mathcal{O}$ is an open $\mathcal{C}^2$ bounded connected set of $\mathbb{R}^d$.

3.1. **Quasi-stationary distribution: existence and long-time convergence**

The justification of the Parallel Replica method mentioned in Section 1.2 for the Langevin dynamics (13) in $\mathcal{D}$ first requires the existence of a QSD on $\mathcal{D}$ for the associated process $(X_t)_{t\geq 0}$ killed outside of $\mathcal{D}$. Such existence can be obtained through the application of the Krein-Rutman theorem to the semigroup of $(X_t)_{t\geq 0}$, which shall provide a left eigenvector of this semigroup corresponding to the QSD density. To apply the Krein-Rutman theorem, this semigroup is required to be compact in a Banach space. To prove the compactness the idea of the proof is to first show

that the semigroup is an integral Hilbert-Schmidt operator in $\mathrm{L}^2(\mathcal{D})$ hence compact in $\mathrm{L}^2(\mathcal{D})$ using appropriate Gaussian upper-bounds for its transition density obtained in [49, Theorem 2.19]. Second, this compactness is propagated across the Banach space $\mathcal{C}^b(\overline{\mathcal{D}})$ of continuous and bounded functions on $\overline{\mathcal{D}}$ using the continuity of the semigroup, leading to the following result in [47, Theorem 2.9].

Theorem 3 (Compactness of the killed semigroup). *For any $t \geq 0$, $p \in [1, +\infty]$ and $f \in \mathrm{L}^p(\mathcal{D})$, the quantity*

$$M_t f : (x, v) \in \overline{\mathcal{D}} \mapsto \mathbb{E}\left[f(X_t) 1_{\tau > t} | X_0 = (x, v)\right] \tag{17}$$

is well-defined. Besides,

(1) the family of operators $(M_t)_{t \geq 0}$ is a semigroup on $\mathrm{L}^p(\mathcal{D})$ and on $\mathcal{C}^b(\overline{\mathcal{D}})$.
(2) For any $t > 0$, the operator M_t is compact from $\mathrm{L}^p(\mathcal{D})$ to $\mathrm{L}^p(\mathcal{D})$, and from $\mathcal{C}^b(\overline{\mathcal{D}})$ to $\mathcal{C}^b(\overline{\mathcal{D}})$.

The application of the Krein-Rutman theorem to the operator M_t for $t > 0$ on the Banach space $\mathcal{C}^b(\overline{\mathcal{D}})$ then yields the existence of a QSD on $\mathcal{D}$ of $(X_t)_{t \geq 0}$, with a density belonging to $\mathcal{C}^b(\overline{\mathcal{D}})$. The uniqueness of this QSD follows from the positivity of the transition density of the semigroup using a Harnack inequality provided in [49]. In addition, one can also obtain a spectral interpretation of this QSD using Itô's formula, see the proofs of [47, Theorems 2.14, 2.17] for more details.

Theorem 4 (Existence and uniqueness of a QSD). *There exists a unique QSD μ on $\mathcal{D}$ for the killed process $(X_t)_{t \geq 0}$. In addition, there exists a smooth positive function ψ such that*

$$\mathrm{d}\mu(x, v) = \psi(x, v)\mathrm{d}x\mathrm{d}v.$$

Moreover, ψ is the unique solution in $\mathcal{C}^2(\mathcal{D}) \cap \mathcal{C}^b(\overline{\mathcal{D}})$ to the eigenvalue problem

$$\begin{cases} \mathcal{L}^*\psi(x, v) = -\theta\psi(x, v) & (x, v) \in \mathcal{D}, \\ \psi(x, v) = 0 & x \in \partial\mathcal{O}, v \cdot n(x) < 0, \end{cases}$$

where $n(x)$ is the unitary outward normal vector at $x \in \partial\mathcal{O}$.

In addition, the compactness of the killed semigroup allows to obtain a spectral decomposition of the killed semigroup. This gives in particular the long-time asymptotics of the law of the dynamics conditioned to remain in $\mathcal{D}$, see [47, Theorem 2.22].

Theorem 5 (Convergence to the QSD in total variation). *There exists $\alpha > 0$ such that for any probability measure ν on $\mathcal{D}$, there exists $C_\nu > 0$ such that for all $t \geq 0$,*

$$\|\mu_t - \mu\|_{TV} \leq C_\nu \mathrm{e}^{-\alpha t}, \tag{18}$$

where μ_t is the law of $(X_t)_{t \geq 0}$ conditioned to survival defined in (2) satisfying $X_0 \sim \nu$, and $\|\cdot\|_{TV}$ denotes the total-variation norm on the space of bounded signed measures on $\mathbb{R}^{2d}$.

The QSD on $\mathcal{D}$ can thus be seen as the longtime limit of the dynamics conditioned to stay inside $\mathcal{D}$. This proposition is useful to understand what is a metastable state. A metastable state is a state such that the exit event from this state happens on a timescale larger than the local equilibration time, namely the time to observe the convergence to the QSD in (18).

Furthermore, if the Langevin dynamics (13) is initially distributed according to the QSD in $\mathcal{D}$, then we can explicitly provide the law of the first exit event from $\mathcal{D}$, which is given by the couple

of the first exit time from D and its associated exit point. As mentioned in the introduction, recall that the first exit time from a domain when starting from a QSD is well known to be distributed according to an exponential law; this is a general property for Markov processes. By contrast, the law of the first exit point is strongly dependent on the dynamics. In the case of the Langevin dynamics (13), this is the first result to compute the law of the first exit point starting from the QSD.

Proposition 1. *Let us assume that X_0 is distributed according to the QSD μ in $\mathcal{D}$. Then the law of the couple (τ, X_τ) is fully characterized by the following properties:*

- *τ follows the exponential law of parameter θ where θ is defined in Theorem 4;*
- *τ is independent of X_τ;*
- *The law of X_τ is given by: for any bounded function $f : \partial \mathcal{D} \mapsto \mathbb{R}$,*

$$\mathbb{E}\left[f(X_\tau)|X_0 \sim \mu\right] = \frac{1}{\theta} \int_{\partial \mathcal{D}} |v \cdot n(x)| \, \psi(x, v) f(x, v) \sigma_{\partial \mathcal{O}}(\mathrm{d}x)\mathrm{d}v,$$

where $n(x)$ is the unitary outward normal vector at $x \in \partial \mathcal{O}$ and $\sigma_{\partial \mathcal{O}}$ denotes the Lebesgue measure on the surface $\partial \mathcal{O}$.

As a result, the existence of a QSD and the long-time convergence to the QSD of the Langevin dynamics trapped in $\mathcal{D}$ ensure that the Parallel Replica method also applies for the Langevin process trapped in $\mathcal{D}$ following the reasoning made in [45, Proposition 5].

3.2. **Quasi-stationary distribution: overdamped limit**

In this section we shall assume that $\gamma > 0$ and $\sigma = \sqrt{2\gamma\beta^{-1}}$ for some $\beta > 0$ in (13). The aim of this section is to describe the behavior of the Langevin dynamics (13) when the friction coefficient γ goes to infinity. This will allow us in particular to explicit the overdamped limit of the QSD.

It is well known in the literature that the law of the rescaled position of the Langevin dynamics converges to the law of the overdamped Langevin dynamics (4), i.e. for all $T > 0$:

$$\mathrm{Law}((x_{\gamma t})_{t \in [0,T]}) \underset{\gamma \to \infty}{\longrightarrow} \mathrm{Law}((\overline{x}_t)_{t \in [0,T]}). \tag{19}$$

However, this result does not provide an overdamped limit for the couple of the position and velocity vectors. Such a limit was provided recently in [58] where the following limit is obtained:

$$\mathrm{Law}((x_{\gamma t})_{t \in [0,T]}, v_{\gamma T}) \underset{\gamma \to \infty}{\longrightarrow} \mathrm{Law}((\overline{x}_t)_{t \in [0,T]}, Z), \tag{20}$$

where $Z \sim \mathcal{N}_d(0, \beta^{-1}I_d)$ is a Gaussian vector independent of the process $(\overline{x}_t)_{t \in [0,T]}$. The proof of (20) consists in perturbing the Brownian noise of the Langevin dynamics (13) by a vanishing term when γ goes to infinity such that the position and velocity become independent. Then it remains to consider the overdamped limit of the position and velocity coordinates separately, which is much easier. In addition, since the perturbation converges to zero when γ goes to infinity, the overdamped limit of the perturbed process is the same as the overdamped limit of the Langevin dynamics. Following this result one can identify explicitly the overdamped limit of the QSD for the Langevin dynamics.

Proposition 2. *Let $\mu^{(\gamma)}$ (resp. $\overline{\mu}$) be the QSD of $(X_t)_{t\geq 0}$ on $D = \mathcal{O} \times \mathbb{R}^d$ (resp. $(\overline{X}_t)_{t\geq 0}$ on $\mathcal{O}$) satisfying* (13) *(resp.* (4)*). Then,*

$$\mu^{(\gamma)}(\mathrm{d}x\mathrm{d}v) \underset{\gamma\to\infty}{\longrightarrow} \overline{\mu}(\mathrm{d}x)\frac{\mathrm{e}^{-\frac{\beta|v|^2}{2}}}{(2\pi\beta^{-1})^{\frac{d}{2}}}\mathrm{d}v, \tag{21}$$

in terms of weak convergence measures.

In particular, if one considers the marginal in position of the Langevin QSD then it converges weakly to the QSD of the overdamped dynamics on $\mathcal{O}$ as a consequence of the above proposition.

4. Langevin processes (round 2)

In this section, we present the main results of [36,37]. These results first aim at giving necessary conditions for existence and uniqueness of a quasi-stationary distribution $\mu_{\mathcal{D}}$ of a strongly Feller Markov process $(X_t, t \geq 0)$ killed when it exits a subdomain $\mathcal{D}$, see Section 4.1. These results are then applied to Langevin processes with continuous or singular potentials, see Sections 4.2 and 4.3.

4.1. **A general result**

4.1.1. *Notations and assumptions*

Let $(X_t, t \geq 0)$ be a time homogeneous Markov process valued in a Polish space $\mathcal{S}$, with càdlàg paths and satisfying the strong Markov property, which is 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 $(P_t, t \geq 0)$.

Let $\mathcal{B}(\mathcal{S})$ be the Borel σ-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 $\mathcal{D}([0,T],\mathcal{S})$ 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 $T > 0$, $x \to \mathbb{P}_x(X_{[0,T]} \in \cdot)$ (the law of $X_{[0,T]} := (X_t)_{t\in[0,T]}$) is continuous from $\mathcal{S}$ to the space $\mathcal{M}_1(\mathcal{D}([0,T],\mathcal{S}))$ of probability measures on $\mathcal{D}([0,T],\mathcal{S})$, equipped with the weak convergence topology.
- **(C3) (Lyapunov function condition)** There exist a continuous function $\mathsf{W} : \mathcal{S} \to [1,+\infty[$, belonging to the extended domain $\mathcal{D}_e(\mathcal{L})$ of the generator of $(X_t, t \geq 0)$, two sequences of positive constants (r_n) and (b_n) where $r_n \to +\infty$, and an increasing sequence of compact subsets (K_n) of $\mathcal{S}$, such that

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

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

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

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

where $\mathcal{D}^c = \mathcal{S} \backslash \mathcal{D}$ is the complement of $\mathcal{D}$. The transition semigroup of the killed process $(X_t, 0 \leq t < \tau)$ is:

$$M_t f(\mathsf{x}) = \mathbb{E}_{\mathsf{x}}[\mathbb{1}_{t < \tau} f(X_t)], \quad t \geq 0, \mathsf{x} \in \mathcal{D}, \, f \in b\mathcal{B}(\mathcal{D}). \tag{22}$$

Assume the following on the killed process $(X_t, 0 \leq t < \tau)$:

(C4) For $t \geq 0$, M_t is Feller, i.e. if f is bounded and continuous over $\mathcal{D}$, so is $M_t f$.

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

$$M_t(\mathsf{x}, O) > 0$$

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

4.1.2. *The general result for existence and uniqueness of the quasi-stationary distribution*

In this section, we provide a slightly less general result than [36, Theorem 2.2]. Indeed, Theorem 6 below is actually still valid when replacing **(C4)** by the less stringent assumption that for $t \geq 0$, M_t is weakly Feller.

Theorem 6. *Assume that* **(C1)**→**(C5)** *hold. Take any $p \in]1, +\infty[$. Then, there exists only one quasi-stationary distribution $\mu_{\mathcal{D}}$ of the process $(X_t, t \geq 0)$ in $\mathcal{D}$ satisfying*

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

In addition, there are some constants $\delta > 0$ and $C \geq 1$ such that for any initial distribution ν on $\mathcal{D}$ with $\nu(\mathsf{W}^{1/p}) < +\infty$,

$$\left| \mathbb{P}_{\nu}(X_t \in A | t < \tau) - \mu_{\mathcal{D}}(A) \right| \leq C e^{-\delta t} \frac{\nu(\mathsf{W}^{1/p})}{\nu(\varphi)}, \quad \forall A \in \mathcal{B}(\mathcal{D}), t > 0, \tag{23}$$

where $\varphi : \mathcal{D} \to \mathbb{R}$ is continuous, $\varphi / \mathsf{W}^{1/p}$ is bounded, and φ is positive within $\mathcal{D}$.

Let us give some remarks on this theorem. If W is bounded over $\mathcal{D}$, the quasi-stationary distribution inside $\mathcal{D}$ is unique. When $\mathcal{S}$ is moreover compact, Theorem 6 is valid without assuming **(C3)** and with $\mathsf{W} = 1$ there (i.e. there exists a unique quasi-stationary distribution in $\mathcal{M}_1(\mathcal{D})$ and (23) holds for any $\nu \in \mathcal{M}_1(\mathcal{D})$). On the contrary, when W is unbounded, Theorem 6 states that there is a unique QSD which integrate $\mathsf{W}^{1/p}$ for all $p > 1$ (in the statement the QSD $\mu_{\mathcal{D}}$ a priori depends on p, but then, clearly, it doesn't, since $\mu_{\mathcal{D}}(\mathsf{W}^{1/p}) < \infty$ for some $p > 1$ implies $\mu_{\mathcal{D}}(\mathsf{W}^{1/q}) < \infty$ for all $q \geqslant p$), but it doesn't prevent the existence of other QSD ν with $\nu(\mathsf{W}^{1/p}) = +\infty$ for all $p > 1$. This is for instance the case for the Ornstein-Uhlenbeck process killed at 0 (see [50]), to which Theorem 6 applies with $\mathsf{W}(\mathsf{x}) = e^{\alpha \mathsf{x}^2}$ with $\alpha < 1/2$.

The intuition behind the assumptions **(C1)**→**(C5)** is the following. It is now well known that a Lyapunov condition provides a spectral gap for a non killed semigroup in some weighted Banach spaces [29,52]. Using results of [64] characterizing the essential spectral radius of a positive operator (combined with the regularity conditions **(C1)** and **(C2)** on the non killed process $(X_t, t \geq 0)$), the Lyapunov type condition **(C3)** is used to get a spectral gap for M_t in $b_{\mathsf{W}^{1/p}}\mathcal{B}(\mathcal{D})$ (the space of measurable functions f over $\mathcal{D}$ such that $f / \mathsf{W}^{1/p}$ is bounded). It is also known that an irreducibility condition and a regularity condition on a non killed semigroup ensure the uniqueness of the invariant

measure. It is similar for killed processes: we show that the regularity condition **(C4)** on M_t and the irreducibility condition **(C5)** imply that M_t admits a principal eigenvalue.

Let us also refer to [3, 8, 12, 13, 32, 42, 61] and references therein for other general criteria for existence and uniqueness of a quasi-stationary distributions.

4.2. **Langevin processes with continuous coefficients**

In this section we apply Theorem 6 to Langevin processes with continuous coefficients (also called hypoelliptic damped Hamiltonian systems). Let us introduce these processes. Let $d \geq 1$ and $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t \geq 0}, \mathbb{P})$ 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 $\mathbb{R}^{2d}$:

$$
\begin{cases}
dx_t = v_t dt, \\
dv_t = -\nabla V(x_t)dt - c(x_t, v_t)v_t dt + \Sigma(x_t, v_t)dB_t,
\end{cases} \tag{24}
$$

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 state space is $\mathcal{S} = \mathbb{R}^{2d}$. Let us define the following assumptions on V and c:

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

(Ac1) $c : \mathbb{R}^{2d} \to \mathrm{M}_d(\mathbb{R})$ (the space of square matrices of size d with real coefficients) is continuous. In addition, there exist $\eta > 0$ and $L > 0$, such that

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

Finally, for all $N > 0$,

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

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

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

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

We also consider the less stringent assumption than **(Ac1)**:

(Ac0) $c : \mathbb{R}^{2d} \to \mathrm{M}_d(\mathbb{R})$ is continuous and

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

This will allow us to consider in particular Langevin processes with unbounded v-dependent damping coefficient:

$$
\text{for some } \ell_0 > 0, \ \forall x, v \in \mathbb{R}^d, \ c(x,v) = |v|^{\ell_0} \tag{25}
$$

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

$$
V \text{ satisfies } \textbf{(Av1)}, \ r^{-1}|x|^{\mathsf{n}_0} \leq V(x) \leq r|x|^{\mathsf{n}_0} \text{ and } r^{-1}|x|^{\mathsf{n}_0} \leq x \cdot \nabla V(x). \tag{26}
$$

Notice that when (25) holds, Assumption **(Ac1)** is not satisfied but **(Ac0)** is satisfied. When V, c, and Σ satisfy respectively **(Av1)**, **(Ac0)**, and **(AΣ)**, there is a unique weak solution to (24) by [63, Lemma 1.1], which is thus a (strong) Markov process. In the following we always assume that **(AΣ)** holds.

As explained in the introduction above (see also [27]), the process $(x_t, t \geq 0)$ is stuck during long periods of time in neighborhoods of the local minima of V. This is due to energetic barriers this process has to cross to visit other regions of $\mathbb{R}^d$. Such neighborhoods are called metastable regions and are thus of the form

$$\mathcal{D} = \mathsf{O} \times \mathbb{R}^d, \tag{27}$$

where $\mathsf{O} \subset \mathbb{R}^d$. We are interested in the existence of quasi-stationary distributions for the processes (24) in domain $\mathcal{D}$ of this form. We assume that O is a $\mathcal{C}^2$ subdomain (not necessarily bounded) of $\mathbb{R}^d$, such that $\mathbb{R}^d \setminus \overline{\mathsf{O}}$ is non-empty.

4.2.1. The case when **(Av1)** and **(Ac1)** hold

Let us define the following last assumptions on V and c.

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

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

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

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

Some examples of functions V and c satisfying **(Av1)**, **(Av2)**, **(Ac1)**, and **(Ac2)** are given in [63, Remark 3.2], see also [36, Remark 6.3]. The Hamiltonian of the process (24) is, for $x, v \in \mathbb{R}^d$,

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

Assume that **(AΣ)**, **(Av1)**, **(Av2)**, **(Ac1)**, and **(Ac2)** hold. Let us introduce for $(x,v) \in \mathbb{R}^{2d}$, the modified Hamiltonian [63, Eq. (3.3)],

$$\mathsf{F}(x,v) = a\,\mathsf{H}(x,v) + v \cdot (b\,G(x) + \nabla\mathsf{w}(x)) + b\,U(x), \tag{28}$$

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

$$\mathsf{W}(x,v) = \exp\left[\mathsf{F}(x,v) - \inf_{\mathbb{R}^{2d}}\mathsf{F}\right] \geq 1. \tag{29}$$

The following asymptotic upper bound on W_1 gives a way to check if a probability measure on ν on $\mathcal{D} \subset \mathbb{R}^{2d}$ satisfies $\nu(\mathsf{W}^{1/p}) < +\infty$ (see the proof of [36, Lemma 2.6]).

Lemma 1. *Assume that V satisfies **(Av1)** and **(Av2)**. Then $\lim_{x \to +\infty} V(x) = +\infty$. Let c be such that **(Ac1)** and **(Ac2)** hold with $\lim_{|x| \to +\infty} U(x)/V(x) = 0$. Then, for any $\epsilon > 0$, there exists $R > 0$ such that if $|x| + |v| \geq R$, $\mathsf{W}(x,v) \leq e^{a(1+\epsilon)\mathsf{H}(x,v)}$.*

We have the following result which ensures existence and uniqueness of the quasi-stationary distribution of the process (24) on $\mathcal{D}$ (see (27)) when V and c satisfy **(Av1)**, **(Av2)**, **(Ac1)**, and **(Ac2)**.

Theorem 7. *Assume that Σ satisfies (AΣ) and that the functions V and c satisfy (Av1), (Av2), (Ac1), and (Ac2). Then, there exist parameters $\mathsf{w} \in \mathcal{C}_c^2(\mathbb{R}^d, \mathbb{R})$, $a > 0$, and $b > 0$ (see [63, Eq. (3.4) $\to$ Eq. (3.9)] for explicit conditions on w, a, and b) such that Theorem 6 is valid for the process (24) with $\mathcal{D} = \mathsf{O} \times \mathbb{R}^d$ and with the Lyapunov function W defined in (29).*

In other words, for all $p > 1$, under (AΣ), (Av1), (Av2), (Ac1), and (Ac2) and when $\mathcal{D} = \mathsf{O} \times \mathbb{R}^d$ (where O is as in Theorem 7), there exists a unique quasi-stationary distribution in $\mathcal{D}$ for the process (24) in the space $\mathscr{M}_p = \{\nu \in \mathcal{M}_1(\mathcal{D}), \nu(\mathsf{W}^{1/p}) < +\infty\}$. In addition, Equation (23) holds for all $\nu \in \mathscr{M}_p$.

To prove Theorem 7 we use Theorem 6. More precisely, we show that (C1), (C2), (C4), and (C5) are satisfied and that, choosing well the function w and taking $a > 0$ and $b > 0$ small enough (see [63, Eq. (3.4) $\to$ Eq. (3.9)]), the function W defined in (29) satisfies (C3).

When O is bounded, V is $\mathcal{C}^\infty$ on $\overline{\mathsf{O}}$, and both Σ and c are constant, the very recent work [47] (see also [49, 58]) establishes the uniqueness of the quasi-stationary distribution in the whole space of measure over $\mathcal{D}$. These results were presented in Section 3 where the authors also considered non gradient vector fields (see indeed (13)). Their approach is different from the one adopted in [36, 37]: it is based on a Feynman-Kac type formula for the killed semigroup, Harnack inequalities, and Gaussian upper bounds. We finally also refer to [4] for the recent study of Hypoelliptic diffusions killed at the boundary of a bounded subdomain of $\mathbb{R}^m$ with non characteristic boundary.

4.2.2. *The case when* (25) *and* (26) *are satisfied*

When (25) and (26) hold, and $\ell_0 < \mathsf{n}_0 - 2$, we are able to construct a bounded Lyapunov function for the process (24) (i.e. a bounded Lyapunov function W over $\mathbb{R}^{2d}$ satisfying (C3)). This implies the following theorem.

Theorem 8. *Assume that Σ satisfies (AΣ) and that the functions V and c satisfy (25) and (26). Assume in addition that*

$$\ell_0 < \mathsf{n}_0 - 2.$$

Then, Theorem 6 is valid with a bounded Lyapunov function. In particular: there exists a unique quasi-stationary distribution in $\mathcal{D} = \mathsf{O} \times \mathbb{R}^d$ for the process (24) in the whole space of probability measures $\mathcal{M}_1(\mathcal{D})$ on $\mathcal{D}$, and in addition, (23) holds for all $\nu \in \mathcal{M}_1(\mathcal{D})$.

4.3. Langevin processes with singular potentials

In this section we consider the existence and uniqueness of a quasi-stationary distribution for the process (24) when the potential V is singular. For ease of exposition, we will just focus here on Lennard-Jones and Coulomb interactions. However much more general singular potentials are considered in [37].

For $d \geq 1$, consider a system of $N \geq 2$ particles in $\mathbb{R}^d$ which cannot collide and let

$$x_t = (x_t^1, \ldots, x_t^N) \in (\mathbb{R}^d)^N \text{ and } v_t = (v_t^1, \ldots, v_t^N) \in (\mathbb{R}^d)^N,$$

denote respectively the positions of the N particles and their velocities, at time $t \geq 0$. The natural space to consider the time evolution of the positions $(x_t, t \geq 0)$ and of the velocities $(v_t, t \geq 0)$ of the N particles is thus

$$\mathcal{S} = \mathcal{O} \times (\mathbb{R}^d)^N, \tag{30}$$

where, if $d = 1$,

$$\mathcal{O} = \left\{ x = (x^1, \ldots, x^N) \in (\mathbb{R})^N, x^1 < x^2 < \ldots < x^N \right\},$$

and if $d \geq 2$,
$$\mathcal{O} = \left\{ x = (x^1, \ldots, x^N) \in (\mathbb{R}^d)^N,\ x^i \neq x^j \text{ for all } i \neq j \right\}.$$
Notice that in both cases, $\mathcal{O}$ is open, path connected, and unbounded. In molecular dynamics, the interatomic potential of the system of N particles is typically of the form, for $x = (x^1, \ldots, x^N) \in \overline{\mathcal{O}}$,

$$V(x) = \sum_{i=1}^{N} V_{\mathrm{c}}(x^i) + \sum_{1 \leq i < j \leq N} V_{\mathrm{I}}(x^i - x^j) \in \overline{\mathbb{R}}, \tag{31}$$

where $V_{\mathrm{c}} : \mathbb{R}^d \to \mathbb{R}$ is the confining potential of the system and $V_{\mathrm{I}} : \Omega \to \mathbb{R}$ (where, if $d = 1$, $\Omega = \{ y < 0 \}$, and if $d \geq 2$, $\Omega = \{ y \neq 0 \}$) is the potential energy modeling the interaction between two particles, the latter becoming infinite when (and only when) $y \in \partial \Omega = \{ 0 \}$ (which prevents from collisions).

Let $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t \geq 0}, \mathbb{P})$ be a filtered probability space. We assume that the evolution of the positions $(x_t, t \geq 0)$ and the velocities $(v_t, t \geq 0)$ of the N particles on $\mathcal{S}$ is described by the following hypoelliptic stochastic differential equation

$$\begin{cases} dx_t = v_t dt, \\ dv_t = -\nabla V(x_t) dt - c(x_t, v_t) v_t dt + \Sigma(x_t, v_t)\, dB_t, \end{cases} \tag{32}$$

where $c : (\mathbb{R}^d)^N \times (\mathbb{R}^d)^N \to \mathrm{M}_{Nd}(\mathbb{R})$ is the friction matrix, $(B_t, t \geq 0)$ is a standard dN-dimensional Brownian motion, and $\Sigma : (\mathbb{R}^d)^N \times (\mathbb{R}^d)^N \to \mathbb{R}$. We set $X_t = (x_t, v_t)$ for $t \geq 0$. Throughout this section, we assume that Σ satisfies $(\mathbf{A\Sigma})$ (with Nd instead of d there) and that c is such that:

(Ac) $c : (\mathbb{R}^d)^N \times (\mathbb{R}^d)^N \to \mathrm{M}_{Nd}(\mathbb{R})$ is a locally Lipschitz function such that:
  (i) there exists $c^* > 0$, $\forall x, v \in (\mathbb{R}^d)^N : \frac{1}{2} \left[c(x, v) + c^T(x, v) \right] \geq c^* I_{(\mathbb{R}^d)^N}$,
  (ii) c is bounded, i.e. $\sup_{x, v \in (\mathbb{R}^d)^N, k, \ell = 1, \ldots, Nd} |c_{k, \ell}(x, v)| < +\infty$.

In the two next sections, we consider existence and uniqueness of a quasi-stationary distribution for the process (32) when V_{I} is the Lennard-Jones potential or the Coulomb potential.

4.3.1. *Quasi-stationary distributions for the Lennard-Jones potential (for any $d \geq 1$) and the Coulomb potential (when $d \geq 3$)*

We recall that the Lennard-Jones potential is defined by, for $y \in \overline{\Omega}$,

$$V_{LJ}(y) = \begin{cases} \dfrac{b}{|y|^{12}} - \dfrac{c}{|y|^6} & \text{if } y \in \Omega \\ +\infty & \text{if } y = 0, \end{cases}$$

where $b, c > 0$. When $d \geq 3$, the Coulomb potential is the function defined on $\overline{\Omega}$ by:

$$V_{co}(y) = \begin{cases} \dfrac{e}{|y|^{d-2}} & \text{if } y \in \Omega \\ +\infty & \text{if } y = 0, \end{cases}$$

where $e > 0$. In this section, the interaction potential V_{I} is either V_{co} or V_{LJ}. The Coulomb potential when $d = 2$ is treated in the next section. We assume that V_{c} is $\mathcal{C}^2$ over $\mathbb{R}^d$ and to

simplify we assume that for some $r > 0$,

$$\forall y \in \mathbb{R}^d, |y| \geq r, \ V_{\mathrm{c}}(y) = A|y|^{\alpha_1}, \tag{33}$$

where $A > 0$ and $\alpha_1 > 1$. We refer to [37] for more general confining potentials. With such confining potentials and interaction potentials, by [37, Proposition 2.3], for any $\mathsf{x} \in \mathcal{S}$, there exists a unique pathwise solution $(X_t = (x_t, v_t), t \geq 0)$ of (32) with $X_0 = \mathsf{x}$, which is moreover non-explosive and remains in $\mathcal{S}$ for all $t \geq 0$.

Recall that the Hamiltonian of the process (24) is, for $(x, v) \in \mathcal{S}$,

$$\mathsf{H}(x, v) = V(x) + \frac{1}{2}|v|^2, \tag{34}$$

where V is given by (31). We have that $\mathsf{H}(x, v) \to +\infty$ (for $(x, v) \in \mathcal{S}$, see (30)) if and only if $x \to \partial\mathcal{O} \cup \{\infty\}$ or $v \to +\infty$.

Let us now consider

$$\eta_1 \in (0, 1].$$

If $\alpha_1 \in (1, 2)$, we assume in addition that $\eta_1 > (2 - \alpha_1)/\alpha_1 \ (\in (0, 1))$. We have the following result (see [37, Theorem 2.4]).

Theorem 9. *Assume that **(Ac)** and **(AΣ)** are satisfied. Assume also that $V_{\mathrm{I}} \in \{V_{co}, V_{LJ}\}$ and that V_{c} is a $\mathcal{C}^2$ potential such that (33) holds. Let O be a subdomain of $\mathcal{O}$ such that $\mathcal{O} \setminus \overline{\mathsf{O}}$ is nonempty and $\partial\mathsf{O} \cap \mathcal{O}$ is $\mathcal{C}^2$. Set $\mathcal{D} = \mathsf{O} \times (\mathbb{R}^d)^N \subset \mathcal{S}$ (see (30)). Choose η_1 as above. Then, Theorem 6 is valid for the process (32) on $\mathcal{D}$ with a continuous and unbounded Lyapunov functional $\mathsf{W} : \mathcal{S} \to [1, +\infty)$ which satisfies, for some $m > 0$, $\mathsf{W} \leq \exp\left[m\mathsf{H}^{\eta_1}\right]$ on $\mathcal{S}$.*

In other words, Theorem 9 states that on $\mathcal{D}$, there exists a unique quasi-stationary distribution for the process (32) in the space $\mathscr{M}_p$ ($\forall p > 1$). In addition, Equation (23) holds for all $\nu \in \mathscr{M}_p$.

We refer to [37, Proposition 2.10] for the explicit construction of W, which is inspired by the previous works [63] and [39]. Notice that O is not necessarily bounded in Theorem 9, and its closure may contain singularities of V, namely some subset of $\partial\mathcal{O}$.

4.3.2. *Quasi-stationary distributions for the Coulomb potential when $d = 1, 2$*

In this section, we consider the interaction potential V_{I} defined by

$$\text{for all } y \in \mathbb{R}^d: \ V_{\mathrm{I}}(y) = -\log|y| \text{ if } y \neq 0, \text{ else } V_{\mathrm{I}}(y) = +\infty. \tag{35}$$

When $d = 2$, V_{I} is the Coulomb potential, and when $d = 1$, V_{I} corresponds to a log singularity pairwise potential. We also assume that the confining potential $V_{\mathrm{c}} : \mathbb{R}^d \to \mathbb{R}$ satisfies: $V_{\mathrm{c}} \in \mathcal{C}^2(\mathbb{R}^d, \mathbb{R})$ and for some $r > 0$, it holds:

$$\forall y \in \mathbb{R}^d, |y| \geq r, \ V_{\mathrm{c}}(y) = A|y|^{\alpha_2}, \tag{36}$$

where $A > 0$ and $\alpha_2 \geq 0$ (we refer to [37] for more general confining potentials).

With such an interaction potential and confining potentials, by [37, Proposition 3.1], for any $\mathsf{x} \in \mathcal{S}$ (see (30)), there exists a unique pathwise solution $(X_t = (x_t, v_t), t \geq 0)$ of (32) with $X_0 = \mathsf{x}$, which is moreover non-explosive and remains in $\mathcal{S}$ (see (30)) for all $t \geq 0$.

Theorem 10. *Let $d \in \{1, 2\}$. Assume that $(\boldsymbol{Ac})$ and $(\boldsymbol{A\Sigma})$ are satisfied. Assume also that V_{I} is given by (35) and that V_{c} is a $\mathcal{C}^2$ potential such that (36) holds. Let O be a subdomain of $\mathcal{O}$ such that $\mathcal{O} \setminus \overline{\mathsf{O}}$ is nonempty and $\partial \mathsf{O} \cap \mathcal{O}$ is $\mathcal{C}^2$. Set $\mathcal{D} = \mathsf{O} \times (\mathbb{R}^d)^N \subset \mathcal{S}$ (see (30)). Take $\eta_2 \in (0, 1]$. Then, Theorem 6 is valid for the process (32) on $\mathcal{D}$ with a continuous and unbounded Lyapunov functional $\mathsf{W} : \mathcal{S} \to [1, +\infty)$ which satisfies, for some $m > 0$, $\mathsf{W} \leq \exp\left[m\mathsf{H}^{\eta_2}\right]$ on $\mathcal{S}$.*

We refer to [37, Proposition 3.3] for the explicit construction of W, which is inspired by [51]. Let us emphasize that there is no restriction in Theorem 10 on η_2 (i.e. one can choose any η_2 in $(0, 1]$). We mention that in [37], we also deduce that large deviation principles hold for the non killed process (i.e. the process $(X_t, t \geq 0)$ on $\mathcal{S}$) with Lennard-Jones and Coulomb potential interactions, as well as the exponential convergence of its law towards its invariant measure. The existence and uniqueness of a quasi-stationary distribution for elliptic diffusions with Lennard-Jones type interactions is also studied there.

5. Numerical approximation with particles

This section is based on [40]. The goal is to introduce a numerical scheme whose aim is to sample the QSD of a killed diffusion, and to prove quantitative long-time convergence rates for this algorithm.

5.1. **The problem**

Given a Markov process X killed at time τ, we introduce a numerical scheme in order to sample its quasi-stationary distribution ν_*, or the law of X_t conditioned on survival. With great generality, one may introduce a particle approximation of the process conditioned on survival, called a Fleming-Viot process in most of the literature. It consists in $N \in \mathbb{N}$ independent processes (X^i), having the same dynamic as X, but when one of them dies, instead of staying at the cemetery point, it chooses uniformly at random another one of the processes, and branches onto it. This gives rise to a mean-field interacting particle system, with Markov semi-group $(P_{N,t})_{t \geqslant 0}$. In the particular case of a softly-killed elliptic SDE on the d-dimensional torus:

$$\mathrm{d}X_t = b(X_t)\mathrm{d}t + \mathrm{d}B_t, \tag{37}$$

where $b \in \mathcal{C}^1(\mathbb{T}^d)$ and B is a d-dimensional Brownian motion, we propose a numerical scheme of this Fleming-Viot process.

Given a random variable $E \in \mathbb{R}_+$ with law $e^{-t}\mathrm{d}t$, we define the soft killing as:

$$\tau = \inf\left\{t \geqslant 0, E \leqslant \int_0^t \lambda(X_s)\mathrm{d}s\right\}, \tag{38}$$

where $\lambda : \mathbb{T}^d \to \mathbb{R}_+$ is a L_λ-Lipschitz death rate, for some $L_\lambda > 0$. As before, write:

$$\mu_t = \mathcal{L}aw(X_t | \tau > t),$$

the law of the process conditioned on survival.

First, we introduce a discretized version of the killed diffusion. We define a Markov kernel K^γ on $\mathbb{T}^d$ for γ small enough as follow: fix $x \in \mathbb{T}^d$, then

$$K^\gamma(x, \cdot) = \mathcal{N}(x + \gamma b(x), \gamma I).$$

We define a Markov chain (X_n) by $X_{n+1} \sim K^\gamma(X_n, \mathrm{d}x)$. Given a sequence of independent and uniform on $[0,1]$ random variable (U_n), we define the death time of the chain (X_n) as:

$$\bar\tau = \inf\{n \geqslant 0, U_n \leqslant p(X_n)\}, \tag{39}$$

where $p(x) = 1 - e^{-\gamma\lambda(x)}$. Write

$$\eta_n = \mathcal{L}aw(X_n | \tau > n).$$

The discretized killing as been defined as such in order to have for all $t \geqslant 0$:

$$\eta_{\lfloor t/\gamma \rfloor} \to \mu_t,$$

as $\gamma \to 0$. This Markov chain also admits a QSD ν_γ.

Now, we want to define two processes: a Markov process (Y_n) whose law is η_n, defined for all time, and a particle approximation of this Markov chain (Y_n). To this end, we define first, given some probability measure $\mu \in \mathcal{M}^1(\mathbb{T}^d)$, a Markov kernel Q_μ such that $Q_\mu(x, \cdot)$ corresponds to the law of some random variable having done a step of discretized diffusion, and upon death, it does a step of discretized diffusion, conditioned on survival, with initial condition μ. More precisely, we construct a random variable of law $Q_\mu(x, \cdot)$ as follow:

(1) Draw $X_0 \sim K^\gamma(x, \cdot)$ and $U_0 \sim \mathcal{U}([0,1])$.
(2) If $U_0 \geqslant p(X_0)$, set $X = X_0$ in $\mathbb{T}^d$ (in that case, we say the particle has moved from x to X_0 without dying).
(3) If $U_0 < p(X_0)$ then set $i = 1$ and, while X is not defined, do:
    (a) Draw a new X_i' distributed according to μ, a new $X_i \sim \mathcal{N}(X_i' + \gamma b(X_i'), \gamma I_d)$ and a new $U_i \sim \mathcal{U}([0,1])$.
    (b) If $U_i \geqslant p(X_i)$, set $X = X_i$ in $\mathbb{T}^d$ (in that case, we say the particle has died, resurrected at X_i', moved to X_i and survived).
    (c) If $U_i < p(X_i)$, set $i \leftarrow i + 1$ (in that case, we say the particle has died, resurrected at X_i', moved to X_i and died again) and go back to step (a).

We introduce the following non-linear process, solution to the problem:

$$\begin{cases} Y_{n+1} \sim Q_{\tilde\eta_n}(Y_n, \cdot) \\ \tilde\eta_n = \mathcal{L}aw(Y_n) \end{cases}. \tag{40}$$

This non-linear Markov chain does at each time n a step of discretized diffusion conditioned on survival. As said, the reason to introduce this Markov chain is the following property:

Proposition 3. *For all $n \in \mathbb{N}$*

$$\tilde\eta_n = \mathcal{L}aw(X_n \mid \tau > n).$$

Now we can define the numerical scheme of the Fleming-Viot process. For $\mathbf{x} = (x_1, \ldots, x_N)$, we write:

$$\pi(\mathbf{x}) := \frac{1}{N} \sum_{i=1}^N \delta_{x_i} \in \mathcal{P}(\mathbb{T}^d) \tag{41}$$

the empirical measure and:

$$R_{\gamma,N}(\mathbf{x}, \cdot) = Q_{\pi(x)}(x_1, \cdot) \otimes \cdots \otimes Q_{\pi(x)}(x_N, \cdot). \tag{42}$$

We then define recursively the process by $\mathbf{X}_{n+1} \sim R_{\gamma,N}(\mathbf{X}_n, \cdot)$. This can be understood as follow: each particles is an Euler-Maruyama scheme of the diffusion, and if one of them dies at time n, then it does a step of K^γ, starting from $\pi(\mathbf{X}_{n-1})$, conditioned on survival. As N goes to infinity, if the initial condition condition of this particle system is N iid random variable of law η_0, then $\pi(\mathbf{X}_n)$ is expected to be close to the common law of the X^i's. This can then be shown to be close to η_n. This is the propagation of chaos phenomenon.

5.2. **Main results**

For this Markov chain, we proved two theorem. The first one is about the long-time behavior of the process:

Theorem 11. *There exists $c_0, \gamma_0 > 0$ that depends only on the drift b and the dimension d such that, if λ is Lipschitz with a constant L_λ such that*

$$L_\lambda e^{\gamma \|\lambda\|_\infty} \quad < \quad c_0, \tag{43}$$

then there exist a distance ρ equivalent to $|\cdot|$ on $\mathbb{T}^d$, such that $\rho_N(x,y) = \sum \rho(x_i, y_i)$ is a distance on $\mathbb{T}^{dN}$, and $\kappa > 0$, such that for all $\gamma \in (0, \gamma_0]$ $N \in \mathbb{N}$, and all $\boldsymbol{x}, \boldsymbol{y} \in \mathbb{T}^{dN}$, there exists a coupling $(\boldsymbol{X}, \boldsymbol{Y})$ of $\delta_{\boldsymbol{x}} R_{N,\gamma}$ and $\delta_{\boldsymbol{y}} R_{N,\gamma}$ such that:

$$\mathbb{E}\left(\rho_N\left(\boldsymbol{X}, \boldsymbol{Y}\right)\right) \leqslant (1 - \gamma\kappa)\rho_N(\boldsymbol{x}, \boldsymbol{y}).$$

This yields the existence of a stationary measure $\mu_{\infty,N,\gamma}$, and exponential convergence of the Markov chain towards this probability measure.

To construct the coupling of theorem 11, we first construct a coupling of $Q_\mu(x, \cdot)$ and $Q_\nu(x, \cdot)$, for any $x, y \in \mathbb{T}$, $\mu, \nu \in \mathcal{P}(\mathbb{T}^d)$. This coupling is done as follow: thanks to [30], we are able couple the step of discretized diffusion to get a coupling (X', Y') of $\delta_x K^\gamma$ and $\delta_y K^\gamma$. The probability that one of them dies and not the other is less then $L_\lambda \mathbb{E}(|X' - Y'|)$. If they die together, then we construct again thanks to [30] a sequence of coupling of μK^γ and νK^γ.

Replacing in this coupling μ and ν by $\pi(\mathbf{x})$ and $\pi(\mathbf{y})$ yields a coupling of $R_{\gamma,N}(x, \cdot)$ and $R_{\gamma,N}(y, \cdot)$.

The second theorem is about the convergence of the empirical measure of the process towards the quasi-stationary distribution of the process (37) killed at time τ, in the limit $\gamma \to 0$ and $n, N \to \infty$.

Theorem 12. *Under the hypothesis of theorem 11, there exists $C > 0$ such that for all $N \in \mathbb{N}$, $\gamma \in (0, \gamma_0]$, $t \geqslant 0$ and $\mu_0 \in \mathcal{P}(\mathbb{T}^{dN})$, if $(\boldsymbol{X}_k)_{k \in \mathbb{N}}$ is a Markov chain with initial distribution μ_0 and transition kernel $R_{N,\gamma}$,*

$$\mathbb{E}\left[\mathcal{W}_1\left(\pi(\boldsymbol{X}_{\lfloor t/\gamma \rfloor}), \nu\right)\right] \leqslant C\left(\sqrt{\gamma} + \alpha(N) + e^{-\kappa t}\right),$$

where

$$\alpha(N) = \begin{cases} N^{-1/2} & \text{if } d = 1, \\ N^{-1/2} \ln(1+N) & \text{if } d = 2, \\ N^{-1/d} & \text{if } d > 2. \end{cases}$$

The second theorem is a consequence of the first one. The rate of convergence in γ, N and t would be the same in the case of N independent diffusion, hence are optimal. The rate of convergence in N comes from [33].

To prove propagation of chaos, we construct a coupling of the discretized Fleming-Viot process, with a system of N independent non-linear process defined in equation (40). This coupling is done is a similare fashion as for the previous coupling of $R_{\gamma,N}(x,\cdot)$ and $R_{\gamma,N}(x,\cdot)$. For the limit $\gamma \to 0$, the coupling of the process is made with a continuous time Fleming-Viot process. Actually, thanks to those couplings, we are able to prove similar convergence when only one or two parameters among t, γ, or N are sent to their limit. We can summarize our results thanks to the diagram of figure 5.
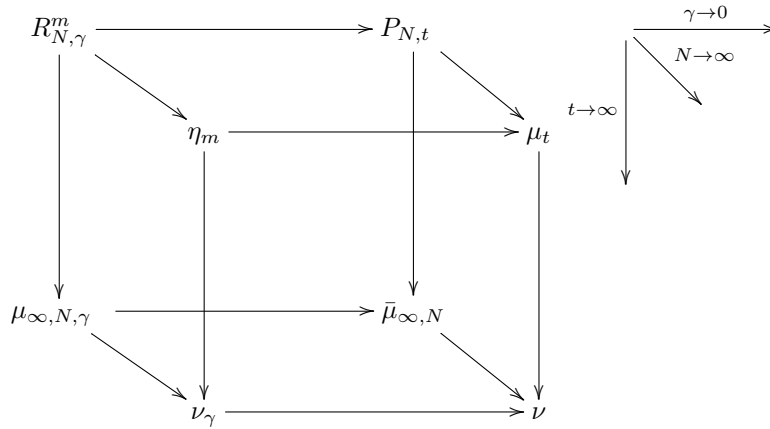

FIGURE 5. Summary of the results

REFERENCES

[1] K. B Athreya and P. E Ney. *Branching processes*. Courier Corporation, 2004.

[2] V. Bansaye, B. Cloez, and P. Gabriel. Ergodic behavior of non-conservative semigroups via generalized Doeblin's conditions. *Acta Applicandae Mathematicae*, pages 1–44, 2017.

[3] V. Bansaye, B. Cloez, P. Gabriel, and A. Marguet. A non-conservative harris ergodic theorem. *Journal of the London Mathematical Society*, n/a(n/a).

[4] M. Benaïm, N. Champagnat, W. Oçafrain, and D. Villemonais. Degenerate processes killed at the boundary of a domain. *Preprint arXiv:2103.08534*, 2021.

[5] M. Benaim, Cloez, B., and F. Panloup. Stochastic approximation of quasi-stationary distributions on compact spaces and applications. *Ann. Appl. Probab.*, 28(4):2370–2416, 2018.

[6] É. Bernard, M. Doumic, and P. Gabriel. Cyclic asymptotic behaviour of a population reproducing by fission into two equal parts. *Kinet. Relat. Models*, 12(3):551–571, 2019.

[7] G. Birkhoff. Extensions of Jentzsch's theorem. *Trans. Amer. Math. Soc.*, 85:219–227, 1957.

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

[9] 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.

[10] D. Chafai and F. Malrieu. *Recueil de modèles aléatoires*, volume 78. Springer, 2016.

[11] 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*, volume 2215 of *Lecture Notes in Math.*, pages 165–182. Springer, Cham, 2018.

[12] N. Champagnat and D. Villemonais. Exponential convergence to quasi-stationary distribution and Q-process. *Probab. Theory Related Fields*, 164(1-2):243–283, 2016.

[13] N. Champagnat and D. Villemonais. General criteria for the study of quasi-stationarity. Preprint, arXiv:1712.08092, 2018.

[14] N. Champagnat and D. Villemonais. Practical criteria for r-positive recurrence of unbounded semigroups. *Electronic Communications in Probability*, 25:1–11, 2020.

[15] N. Champagnat and D. Villemonais. Lyapunov criteria for uniform convergence of conditional distributions of absorbed markov processes. *Stochastic Processes and their Applications*, 135:51–74, 2021.

[16] B. Cloez and J. Corujo. Uniform in time propagation of chaos for a moran model, 2021.

[17] B. Cloez and C. Fritsch. Quasi-stationary behavior for an hybrid model of chemostat: the crump-young model. *arXiv preprint arXiv:2205.05318*, 2022.

[18] B. Cloez and P. Gabriel. On an irreducibility type condition for the ergodicity of nonconservative semigroups. *C. R., Math., Acad. Sci. Paris*, 358(6):733–742, 2020.

[19] B. Cloez and P. Gabriel. Fast, slow convergence, and concentration in the house of cards replicator-mutator model. *arXiv preprint arXiv:2203.07924*, 2022.

[20] P. Collet, S. Martínez, S. Méléard, and J. San Martín. Quasi-stationary distributions for structured birth and death processes with mutations. *Probab. Theory Related Fields*, 151(1-2):191–231, 2011.

[21] P. Collet, S. Martinez, S. Méléard, and J. San Martín. Stochastic models for a chemostat and long-time behavior. *Adv. Appl. Probab.*, 45(3):822–836, 2013.

[22] P. Collet, S. Martinez, and J. San Martín. *Quasi-Stationary Distributions*. Probability and Its Applications. Springer Berlin, Heidelberg.

[23] P. Del Moral. *Mean field simulation for Monte Carlo integration*, volume 126 of *Monographs on Statistics and Applied Probability*. CRC Press, Boca Raton, FL, 2013.

[24] P. Del Moral, E. Horton, and A. Jasra. On the stability of positive semigroups, 2021.

[25] P. Del Moral and L. Miclo. On the stability of nonlinear Feynman-Kac semigroups. *Ann. Fac. Sci. Toulouse Math. (6)*, 11(2):135–175, 2002.

[26] P. Del Moral and L. Miclo. On convergence of chains with occupational self-interactions. *Proc. R. Soc. Lond. Ser. A Math. Phys. Eng. Sci.*, 460(2041):325–346, 2004. Stochastic analysis with applications to mathematical finance.

[27] 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.

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

[29] D. Down, S. P. Meyn, and R. L. Tweedie. Exponential and Uniform Ergodicity of Markov Processes. *The Annals of Probability*, 23(4):1671–1691, 1995.

[30] A. Eberle and M. B Majka. Quantitative contraction rates for markov chains on general state spaces. *Electronic Journal of Probability*, 24:1–36, 2019.

[31] G. Ferré, M. Rousset, and G. Stoltz. More on the long time stability of feynman–kac semigroups. *Stochastics and Partial Differential Equations: Analysis and Computations*, 9(3):630–673, 2021.

[32] G. Ferré, M. Rousset, and G. Stoltz. More on the long time stability of Feynman–Kac semigroups. *Stochastics and Partial Differential Equations: Analysis and Computations*, 9(3):630–673, 2021.

[33] N. Fournier and A. Guillin. On the rate of convergence in Wasserstein distance of the empirical measure. *Probab. Theory Related Fields*, 162(3-4):707–738, 2015.

[34] P. Gabriel and H. Martin. Periodic asymptotic dynamics of the measure solutions to an equal mitosis equation. *Ann. Henri Lebesgue*, 5:275–301, 2022.

[35] G. L. Gong, M. P. Qian, and Z. X. Zhao. Killed diffusions and their conditioning. *Probab. Theory Related Fields*, 80:151–167, 1988.

[36] A. Guillin, B. Nectoux, and L. Wu. Quasi-stationary distribution for strongly Feller Markov processes by Lyapunov functions and applications to hypoelliptic Hamiltonian systems. *Preprint hal-03068461*, 2020.

[37] A. Guillin, B. Nectoux, and L. Wu. Quasi-stationary distribution for Hamiltonian dynamics with singular potentials. *Probability Theory and Related Fields, to appear*, 2021.

[38] M. Hairer and J. C. Mattingly. Yet another look at harris' ergodic theorem for markov chains. In *Seminar on Stochastic Analysis, Random Fields and Applications VI*, pages 109–117. Springer, 2011.

[39] D.P. Herzog and J.C. Mattingly. Ergodicity and Lyapunov functions for Langevin dynamics with singular potentials. *Communications on Pure and Applied Mathematics*, 72(10):2231–2255, 2019.

[40] L. Journel and P. Monmarché. Convergence of a particle approximation for the quasi-stationary distribution of a diffusion process: uniform estimates in a compact soft case, 2020.

[41] R. Knobloch and L. Partzsch. Uniform conditional ergodicity and intrinsic ultracontractivity. *Potential Anal.*, 33(2):107–136, 2010.

[42] M. Kolb and D. Steinsaltz. Quasilimiting behavior for one-dimensional diffusions with killing. *The Annals of Probability*, 40(1):162–212, 2012.

[43] I. Kontoyiannis and S. P. Meyn. Spectral theory and limit theorems for geometrically ergodic Markov processes. *Ann. Appl. Probab.*, 13(1):304–362, 2003.

[44] I. Kontoyiannis and S. P. Meyn. Large deviations asymptotics and the spectral theory of multiplicatively regular Markov processes. *Electron. J. Probab.*, 10:no. 3, 61–123, 2005.

[45] C. Le Bris, T. Lelièvre, M. Luskin, and D. Perez. A mathematical formalization of the parallel replica dynamics. *Monte Carlo Methods Appl.*, 18(2):119–146, 2012.

[46] T. Lelièvre, M. Ramil, and J. Reygner. Quasi-stationary distribution for the langevin process in cylindrical domains, part i: existence, uniqueness and long-time convergence. *Stochastic Processes and their Applications*, 144:173–201, 2022.

[47] T. Lelièvre, M. Ramil, and J. Reygner. Quasi-stationary distribution for the Langevin process in cylindrical domains, part i: existence, uniqueness and long-time convergence. *Stochastic Processes and their Applications*, 144:173–201, 2022.

[48] T. Lelièvre, M. Rousset, and G. Stoltz. *Free energy computations*. Imperial College Press, London, 2010. A mathematical perspective.

[49] T. Lelièvre, M. Ramil, and J. Reygner. A probabilistic study of the kinetic Fokker-Planck equation in cylindrical domains. *Journal of Evolution Equations*, 22(2):1–74, 2022.

[50] 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, 2000.

[51] Y. Lu and J.C. Mattingly. Geometric ergodicity of Langevin dynamics with Coulomb interactions. *Nonlinearity*, 33(2):675, 2019.

[52] S. P. Meyn and R. L. Tweedie. *Markov Chains and Stochastic Stability*. Communications and Control Engineering Series. Springer-Verlag London, 1993.

[53] S. P. Meyn and R. L. Tweedie. Stability of Markovian processes. III. Foster-Lyapunov criteria for continuous-time processes. *Adv. in Appl. Probab.*, 25(3):518–548, 1993.

[54] S. P Meyn and R. L Tweedie. *Markov chains and stochastic stability*. Springer Science & Business Media, 2012.

[55] S. Méléard and D. Villemonais. Quasi-stationary distributions and population processes. *Probability Surveys*, 9(none):340 – 410, 2012.

[56] R. D. Nussbaum. Hilbert's projective metric and iterated nonlinear maps. *Mem. Amer. Math. Soc.*, 75(391):iv+137, 1988.

[57] 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. *Ann. Probab.*, 13(2):363–378, 1985.

[58] M. Ramil. Quasi-stationary distribution for the Langevin process in cylindrical domains, part ii: overdamped limit. *Electronic Journal of Probability*, 27:1–18, 2022.

[59] E. Seneta. *Non-negative matrices and Markov chains*. Springer Series in Statistics. Springer, New York, 2006. Revised reprint of the second (1981) edition [Springer-Verlag, New York; MR0719544].

[60] E. A. van Doorn. Birth–death processes and associated polynomials. *Journal of Computational and Applied Mathematics*, 153(1):497 – 506, 2003. Proceedings of the 6th International Symposium on Orthogonal Polynomials, Special Functions and their Applications, Rome, Italy, 18-22 June 2001.

[61] A. Velleret. Unique quasi-stationary distribution, with a possibly stabilizing extinction. *Preprint arXiv:1802.02409*, 2018.

[62] A. F. Voter. Parallel replica method for dynamics of infrequent events. *Physical Review B*, 57(22):R13985, 1998.

[63] L. Wu. Large and moderate deviations and exponential convergence for stochastic damping Hamiltonian systems. *Stochastic Processes and their Applications*, 91(2):205–238, 2001.

[64] L. Wu. Essential spectral radius for Markov semigroups. I. Discrete time case. *Probability Theory and Related Fields*, 128(2):255–321, 2004.