Mathematical foundations for the Parallel Replica algorithm applied to the underdamped Langevin dynamics

Mouad Ramil*[1], Tony Lelièvre†[1,2], and Julien Reygner‡[1]

[1]CERMICS, Ecole des Ponts, Marne-la-Vallée, France
[2]MATHERIALS, Inria, Paris, France

August 12, 2022

Abstract

Molecular dynamics (MD) methods are used to sample the time evolution of complex molecular systems, namely the transition events between configuration states. When these states are metastable, these transitions correspond to rare events and occur over very long-time scale, thus rendering the use of MD methods inefficient.

The Parallel Replica algorithm, was designed to efficiently sample these rare events relying on a parallelization in time computation of the trajectory. Its mathematical formalization was carried out in the literature using the notion of quasi-stationary distribution (QSD), which can be seen as the long-time distribution of the dynamics trapped inside a state. Its existence was only proven recently for the underdamped Langevin dynamics (ULD) involved in the sampling of thermostated molecular systems.

In this research letter we shall state the existence of a QSD for the ULD, for which we can provide a spectral interpretation. The existence of this QSD is then used to justify the application of Parallel Replica algorithm to the ULD. In addition, we shall extend known results regarding the behaviour of the ULD trajectories when the friction coefficient goes to infinity (overdamped limit). This allows us to provide an explicit expression of the overdamped limit of the previous QSD of the ULD.

The findings of this study remain valid for ULD with non-conservative forces. They shall help for a better understanding of the mathematical framework underlying the Parallel Replica algorithm or related algorithms involving the QSD to the ULD. It also extends our knowledge on the overdamped limiting behaviour of the ULD.

Keywords. mathematics, molecular, simulation, statistics/statistical methods.

1 Introduction

1.1 Motivation

In this short note, we report on recent mathematical results obtained in [17, 21] concerning the metastable behaviour of the underdamped Langevin dynamics (ULD). Some of these results provide rigorous foundations to some accelerated dynamics algorithms (Parallel Replica and ParSplice) which are used to simulate efficiently atomistic models for materials over very long time scales.

Molecular Dynamics (MD) methods are used in various application fields (biology, chemistry, materials science) in order to simulate the evolution of a molecular system, namely interacting particles representing atoms or groups of atoms [13, 24, 15, 22, 1]. When considering a thermostated molecular system, subjected to a fixed temperature, its evolution can be modeled by the following Underdamped Langevin Dynamics (ULD):

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

*E-mail: ramilm@snu.ac.kr
†E-mail: tony.lelievre@enpc.fr
‡E-mail: julien.reygner@enpc.fr

where $(q_t, p_t) \in \mathbb{R}^d \times \mathbb{R}^d$ denotes the positions and momenta of the particles at time $t \geq 0$, M is the mass tensor, 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 the rescaled position coordinate $(q_{\gamma t})_{t\geq 0}$ converges to the solution $(\overline{q}_t)_{t\geq 0}$ of the Overdamped Langevin Dynamics (OLD) (see for instance [14, Section 2.2.4]):

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

In practice, the dimension d of the ambiant space of (1) or (2) is large since the number of atoms is large (from a few thousands to millions depending on the application). Notice that both processes are related by the fact that when the force field F is conservative, that is to say when there exists $V : \mathbb{R}^d \to \mathbb{R}$ such that $F = -\nabla V$, then the stationary distribution of $(\overline{q}_t)_{t\geq 0}$ writes

$$\overline{\nu}(\mathrm{d}q) = \frac{1}{Z}\mathrm{e}^{-\beta V(q)}, \qquad Z = \int_{\mathbb{R}^d} \mathrm{e}^{-\beta V(q)}\mathrm{d}q, \tag{3}$$

while the stationary distribution of $(q_t, p_t)_{t\geq 0}$ has the product structure

$$\nu(\mathrm{d}q\mathrm{d}p) = \overline{\nu}(\mathrm{d}q)\frac{\mathrm{e}^{-\frac{\beta|p|^2}{2}}}{(2\pi\beta^{-1})^{\frac{d}{2}}}\mathrm{d}p. \tag{4}$$

This is no longer true when F is not conservative but becomes true at the overdamped limit when γ goes to infinity as was shown recently in [18].

The objective of molecular simulations is to compute quantities describing the macroscopic evolution of the system, following the dynamics (1) for instance. Such quantities depend on the path $(q_t, p_t)_{t\geq 0}$ and correspond for instance to the average time needed for the dynamics to go from one region of the phase space to another one.

The main difficulty behind these computations comes from the timescale gap between the timestep used in the numerical discretization of the dynamics and the time required to observe such macroscopic evolution. In practice, in many cases of interest this would require to compute a simulation for up to 10^9 to 10^{15} iterations. In fact the timescale at the atomistic level is of order 1 fs (10^{-15} s) whereas transition events often occur over timescales of order 1 μs (10^{-6} s) to seconds, which makes it out of reach using naive simulation.

1.2 Accelerated dynamics methods

The timescale gap between the atomistic and the macroscopic level is often described in the literature as the metastability phenomenon. It justifies the use of new sampling methods such as accelerated dynamics methods [28]. This set of methods takes into account the fact that metastable dynamics remain in regions of the phase space for a long time and focus on sampling the exit event from such regions. The metastability nature of these dynamics is due to two main reasons: the presence of energetic barriers (the dynamics needs to climb above a saddle point of the potential energy) and entropic barriers (the path between two states requires to go through a narrow escape in large dimension). More generally, metastability can correspond to a combination of both effects.

Our focus here is on a particular algorithm used for accelerated dynamics methods, which is called the Parallel Replica algorithm [27, 26, 20]. Parallel Replica was introduced by A.F. Voter and co-workers in the late nineties and is used in computational statistical physics to efficiently sample metastable molecular dynamics trajectories. The mathematical formalization of the Parallel Replica algorithm was carried out in the case of the OLD (2) in [12]. The main motivation of this note is to present extensions to the case of the ULD (1) which is most often used in practice to simulate thermostated molecular systems than OLD.

The idea behind Parallel Replica is that when the trajectory remains trapped inside a metastable state for a sufficiently long time, it reaches a local equilibrium before exiting the state. Once this local equilibrium is reached, 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.

1.2.1 Quasi-stationary distribution

The notion of local equilibrium in Parallel Replica can be mathematically formalized through the notion of quasi-stationary distribution (QSD). This formalization, together with the existence of such

QSD for the OLD were obtained in [12]. Extending the application of Parallel Replica to the ULD (1) requires to state the existence of QSD for the ULD. Given the non-elliptic nature of its infinitesimal generator, this work requires different set of tools and has only been done recently in [17] and also in parallel in [10]. We shall describe here its implication on the Parallel Replica algorithm. Let us emphasize that the existence of a QSD is also useful to develop variants such as [11] and to justify the correctness of related algorithms such as ParSplice [19].

In addition, we shall present new results regarding the overdamped limit of the trajectories of (1), i.e. when taking the coefficient friction γ to infinity. It is known in the literature that the law of the rescaled position of (1) $(q_{\gamma t})_{t\in[0,T]}$ converges to the law of the OLD $(\overline{q}_t)_{t\in[0,T]}$. This result was recently extended in [21] to the convergence of the couple $((q_{\gamma t})_{t\in[0,T]}, p_{\gamma T})$. This shall allow us to obtain an explicit expression of the overdamped limit of the QSD of (1) involving the QSD of (2). Such result was also obtained for the stationary distribution with non-conservative forces F in (1) see [18].

These results provide a new insight on the metastable behaviour of the ULD and further our knowledge on its high-friction behavior, making a deeper and explicit link with its overdamped limit (2).

1.3 Outline

This work is organized as follows. First, the Parallel Replica algorithm is introduced in Section 2. Second, it is justified in Section 3 using the notion of QSD, whose existence is provided in the case of the ULD (1). Finally, in Section 4, we consider the overdamped limit of the ULD trajectories which allow us to identify the overdamped limit of its QSD, making a link with the QSD of (2).

2 Parallel Replica

In this section we shall describe in detail the Parallel Replica algorithm. It shall be mathematically justified in the next section in the case of the ULD.

2.1 Description of the algorithm

We assume here that the state space is partitioned into metastable states and also given two different configurations, we can assert whether they belong to the same metastable state. In most common cases, the metastable states correspond to basins of attraction of a potential such that we can differentiate between metastable states by performing a gradient descent algorithm starting from the given configuration. If two different configurations yield the same local minimum then they belong to the same metastable state.

The idea of Parallel Replica is that if the dynamics (1) remain trapped for a very long time inside a given metastable state D, then the first exit event from D can be sampled much more efficiently than by running the original dynamics until the exit time. Notice that the exit event is fully characterized by two random variables: the exit time τ_∂ and the associated exit point from D, which are defined by considering the first hitting time and point on the boundary of D.

The principle of Parallel Replica is to simulate in parallel many independent replicas following the original dynamics (1), to consider the first exit event among the replicas, and to generate from this first exit event a consistent exit time and exit point. The gain is thus obtained in terms of wall clock time. This algorithm can be seen as a way to parallelize a computation in time, which is not an easy problem in general due to the sequential nature of time evolutions. The ultimate aim of this technique is thus to generate efficiently the so-called state-to-state dynamics with the correct statistical properties, i.e. the chronological sequence of metastable states visited by the trajectory and the time spent in each state. Let us emphasize that the objective is thus to get the correct law on the paths in order to compute dynamical quantities. Let us now go into more detail of the algorithm.

2.2 Iterative scheme

Consider a reference trajectory $(X_t = (q_t, p_t))_{t\geq 0}$ following the ULD (1). Assume that we are given a computer with $N > 1$ CPUs in parallel. The simplest version of Parallel Replica [27, 26] iterates the following three steps:

- **Decorrelation step**: a reference walker is run following the dynamics (1) until it remains trapped for a pre-determined correlation time τ_{corr} in some metastable state D. During this step,

the algorithm thus consists simply in integrating the original dynamics. No error is introduced and there is no computational gain. The objective is then to sample the first exit event from D, i.e. the couple $(\tau_\partial, X_{\tau_\partial})$.

- **Dephasing step**: this is a preparation step during which N independent replicas are sampled starting from the last point of the reference trajectory, each one on a different CPU, during $\tau_{\rm corr}$. This is done in parallel. During this step, the simulation clock is stopped. This step is thus pure overhead. If one of the replicas exits the domain D, it is restarted from its initial point. The aim of this step is to prepare N independent configurations distributed under the local equilibrium in D.

- **Parallel step**: it consists in running the N replicas independently in parallel, and in waiting for the first exit event among the N replicas. The first exit event for the reference trajectory is then sampled by considering only the first exit event among the N replicas:

$$(\tau_\partial, X_{\tau_\partial}) = (N\tau_\partial^{i^*}, X_{\tau_\partial^{i^*}}^{i^*}), \tag{5}$$

where $i^* = \mathrm{argmin}_{1\le i\le N}\tau_\partial^i$. Notice that the simulation clock is updated by adding the effective exit time $N\tau_\partial^{i^*}$. The exit point $X_{\tau_\partial^{i^*}}^{i^*}$ is used as the initial condition of the reference walker for the next decorrelation step. The computational gain of the whole algorithm comes from this step which divides in average the wall clock time to sample the exit event by the number of replicas N.

2.3 Exact asymptotics and extension

Our results allow to prove that the Parallel step is exact in the asymptotic $\tau_{\rm corr} \to \infty$ because the equality in law in (5) is then exact. Moreover, the error in law for a finite $\tau_{\rm corr}$ can be shown to decrease exponentially fast in $\tau_{\rm corr}$. This shall be explained in Propositions 3.2 and 3.4 in the next section. In practice, we shall take $\tau_{\rm corr}$ large enough so that the approximation in (5) is accurate enough.

Notice that Parallel Replica does not require the force F to be conservative, nor do we need to define a partition of the state space into metastable states $D_1, D_2, \dots$ Let us mention the ParSplice algorithm [19], which is a recent extension of the Parallel Replica algorithm based on precomputed segments of the trajectories which start under the QSD and end under the QSD, so that they can be spliced to efficiently generate a full trajectory, based on the same idea of parallel sampling for the first exit event.

3 Mathematical formalization of Parallel Replica for the ULD

In this section, we formalize the Parallel Replica algorithm using the notion of quasi-stationary distribution (QSD). The metastable states often correspond to the basins of attraction of a potential V when the force F in (1) is conservative, i.e. $F = -\nabla V$. Therefore, we shall consider metastable states written as cylinders in the position and momenta coordinates as follows:

$$D = \mathcal{O} \times \mathbb{R}^d, \tag{6}$$

where $\mathcal{O}$ will be a smooth bounded domain of $\mathbb{R}^d$.

In Section 3 we provide the definition of a QSD and detail in Sections 3.1 and 3.2 a few results obtained for the ULD regarding the existence and long-time convergence to the QSD. These results are then used in Section 3.3 to justify the application of Parallel Replica to the ULD.

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

The idea behind Parallel Replica is that when the ULD (1) remains trapped for a long time inside some metastable domain D, it reaches a local equilibrium which corresponds to the QSD of the ULD in D, defined as follows:

Definition 3.1. *A probability measure μ with support in D is called a QSD in D for the dynamics $(q_t, p_t)_{t\ge 0}$ if and only if*

$$\forall t > 0, \forall A \subset D, \qquad \mathbb{P}_\mu((q_t, p_t) \in A | \tau_\partial > t) = \mu(A),$$

where $\tau_\partial = \inf\{t > 0 : q_t \notin \mathcal{O}\}$ *and* $\mathbb{P}_\mu(\mathcal{S})$ *is the probability of the event* $\mathcal{S}$ *when initially* (q_0, p_0) *is distributed according to* μ*.*

In other words, μ is a QSD if, when (q_0, p_0) is distributed according to μ, the law of (q_t, p_t) conditionally to the fact that $(q_s)_{0\leq s\leq t}$ remained in the state $\mathcal{O}$ is still μ, for any time t. We refer to [7] for general properties satisfied by QSD.

The existence of a unique QSD has been shown for the OLD (2) on smooth bounded domains $\mathcal{O} \subset \mathbb{R}^d$ in [9, 12, 6]. It was only recently extended to the ULD (1) (see [17, 10]) on domains $D = \mathcal{O} \times \mathbb{R}^d$ as defined in (6). The main difficulty in this case comes from the non-ellipticity of its infinitesimal generator. However, using appropriate Gaussian upper-bounds for its transition density obtained in [16, Theorem 2.19] we are able to build the QSD using a compactness property satisfied by its semigroup (see [17, Sections 3.2, 3.3] for more details).

In addition, an important fact is that the law of the ULD trapped in D for a time t will converge when t goes to infinity to the QSD. Moreover, this convergence is exponentially fast in time [17, Theorem 2.22].

Proposition 3.2. *There exists a unique QSD* μ *for* $(q_t, p_t)_{t\geq 0}$ (1) *on* D*. Besides, there exists* $\alpha > 0$ *such that for any initial condition* (q_0, p_0) *in* D*, there exists* $C > 0$ *such that for all* $t \geq 0$*,*

$$\|\mathcal{L}((q_t, p_t)|\tau_\partial > t) - \mu\|_{TV} \leq C\mathrm{e}^{-\alpha t}, \tag{7}$$

where $\tau_\partial = \inf\{t > 0 : q_t \notin \mathcal{O}\}$ *and* $\|\cdot\|_{TV}$ *is the total-variation norm.*

The QSD can thus be seen as the longtime limit of the dynamics conditioned to stay inside the domain. This proposition is useful to quantify metastability. 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 (7), namely $1/\alpha \ll \mathbb{E}_\mu[\tau_\partial]$.

3.2 Spectral interpretation and first exit event

The density of the QSD μ can be characterized as the solution to an eigenvalue problem. In fact, let us define the following kinetic Fokker-Planck operator on $\mathbb{R}^d \times \mathbb{R}^d$, associated with the ULD (1):

$$\mathrm{L}^* = -p \cdot \nabla_q - F(q) \cdot \nabla_p + \gamma \mathrm{div}_p(p\cdot) + \gamma\beta^{-1}\Delta_p.$$

Proposition 3.3. *There exists a positive function* $\psi : D \mapsto \mathbb{R}$ *such that*

$$\mathrm{d}\mu(q, p) = \psi(q, p)\mathrm{d}q\mathrm{d}p.$$

Moreover, ψ *is solution to the eigenvalue problem*

$$\begin{cases} \mathrm{L}^*\psi(q, p) = -\lambda\psi(q, p) & (q, p) \in D, \\ \psi(q, p) = 0 & q \in \partial\mathcal{O}, p \cdot n(q) < 0, \end{cases}$$

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

Finally, if the ULD (1) is initially distributed according to the QSD in D, then we can explicitly provide the law of the first exit event from D, which is given by the couple of the first exit time from D and its associated exit point. In the case of the ULD (1), this result taken from [17, Corollary 2.16] is the first to compute the law of the first exit point starting from the QSD.

Proposition 3.4. *Let us assume that* (q_0, p_0) *is distributed according to the QSD* μ *in* D*. Then the law of the couple* $(\tau_\partial, (q_{\tau_\partial}, p_{\tau_\partial}))$ *is fully characterized by the following properties:*

- τ_∂ *is exponentially distributed with parameter* λ *(defined in Proposition 3.3);*
- τ_∂ *is independent of* $(q_{\tau_\partial}, p_{\tau_\partial})$*;*
- *The law of* $(q_{\tau_\partial}, p_{\tau_\partial})$ *is given by: for any bounded function* $f : \partial D \to \mathbb{R}$*,*

$$\mathbb{E}_\mu\left[f(q_{\tau_\partial}, p_{\tau_\partial})\right] = \frac{1}{\lambda}\int_{\partial D} |p \cdot n(q)|\, \psi(q, p) f(q, p)\sigma_{\partial\mathcal{O}}(\mathrm{d}q)\mathrm{d}p,$$

where $\sigma_{\partial\mathcal{O}}$ *denotes the Lebesgue measure on the surface* $\partial\mathcal{O}$*.*

3.3 Justification of Parallel Replica

3.3.1 Justification of the parallel step

Following the reasoning made in [12, Proposition 5] one can then show that if (q_0, p_0) is distributed according to the QSD then the sampling of the first exit event can be done as described in (5). The proof mainly relies on the second property in Proposition 3.4 which allows to consider the laws of the exit time and exit point independently. In addition, a simple consequence of the first property stated in Proposition 3.4 ensures that if all replicas are started according to the QSD then $N \min_{1\leq i\leq N} \tau_\partial^i$ follows the same law as τ_∂. This is an immediate consequence of the fact that N multiplied by the minimum of N independent exponential law of same parameter follows the same exponential law. The independency with the associated exit point then allows us to justify the sampling in (5). It is a general fact for Markov processes that, starting from the QSD, the exit time is exponentially distributed and independent from the exit point. The marginal distribution of the latter is however dependent of the dynamics.

3.3.2 Justification of the Decorrelation and Dephasing steps

The decorrelation and dephasing steps consist in sampling the reference particle or the replicas according to the QSD so that the sampling of the first exit event in the Parallel Step is justified. Sampling according to the QSD is done by running the dynamics inside the domain for a time long enough. In fact, Proposition 3.2 guarantees that if the dynamics remains long enough inside D it is distributed according to the QSD of D. Moreover, it ensures that the law of $(q_{\tau_{\rm corr}}, p_{\tau_{\rm corr}})$ conditioned to staying in D up to time $\tau_{\rm corr}$ is close to μ up to a term of order $\mathrm{e}^{-\alpha\tau_{\rm corr}}$. Therefore, the optimal $\tau_{\rm corr}$ should be of order α^{-1}, where α is often called the spectral gap. In practice, physicists may deduce the order of magnitude of α based on harmonic approximations. Other methods have been designed to sample according to the QSD without such requirement, for instance using the Fleming-Viot particle process, where replicas are sent to the position of an arbitrary replica inside the domain when they exit the domain, see for example [11, 4, 5, 2, 3].

In view of the above results, the errors introduced in the algorithm have two origins. First, in the decorrelation step, $\tau_{\rm corr}$ should be chosen sufficiently large so that at the end of the decorrelation step, the reference walker is indeed distributed according to a probability law sufficiently close to the QSD. Second, in the dephasing step, the sampling algorithm of the QSD should be sufficiently precise in order to obtain i.i.d. samples distributed according to μ. For the rejection algorithm, independence is ensured, and the accuracy is again related to the convergence result of Proposition 3.2.

The existence of the QSD for the ULD and the explicit formula for the law of the first exit point in Proposition 3.4 are the first steps for future works to study the exit event in the small temperature regime by proving that Eyring-Kramers rates can be used to approximate the exit point distribution, in the small temperature regime. This is useful to justify other types of accelerated dynamics algorithms such as Hyperdynamics [25] or Temperature Accelerated Dynamics [23], following the methodology in [8] for the OLD.

4 Overdamped limit of the ULD

4.1 Overdamped limit of the trajectory

The aim of this section is to describe the behavior of the ULD when the friction coefficient γ goes to infinity in (1). This allows us to explicit the overdamped limit of the QSD.

It is well known in the literature that the law of the rescaled position of the ULD converges to the law of the OLD (2):

$$\mathcal{L}((q_{\gamma t})_{t\in[0,T]}) \underset{\gamma\to\infty}{\longrightarrow} \mathcal{L}((\overline{q}_t)_{t\in[0,T]}). \tag{8}$$

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 [21] where the following limit is obtained:

$$\mathcal{L}((q_{\gamma t})_{t\in[0,T]}, p_{\gamma T}) \underset{\gamma\to\infty}{\longrightarrow} \mathcal{L}((\overline{q}_t)_{t\in[0,T]}, Z), \tag{9}$$

where $Z \sim \mathcal{N}_d(0, \beta^{-1} I_d)$ is a Gaussian vector independent of the process $(\overline{q}_t)_{t\in[0,T]}$. The proof of (9) consists in perturbing the noise of the ULD (1) 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 ULD.

4.2 Overdamped limit of the QSD

Given the convergence (8), one may wonder whether the marginal in position of the QSD of the ULD might converge to the QSD of the OLD when γ goes to infinity. Actually, considering the more general overdamped limit (9), we shall identify explicitly the overdamped limit of the QSD for the ULD.

Proposition 4.1. *Let $\mu^{(\gamma)}$ (resp. $\overline{\mu}$) be the QSD of $(q_t, p_t)_{t\geq 0}$ on $D = \mathcal{O} \times \mathbb{R}^d$ (resp. $(\overline{q}_t)_{t\geq 0}$ on $\mathcal{O}$) satisfying (1) (resp. (2)). Then,*

$$\mu^{(\gamma)}(\mathrm{d}q\mathrm{d}p) \underset{\gamma\to\infty}{\longrightarrow} \overline{\mu}(\mathrm{d}q)\frac{\mathrm{e}^{-\frac{\beta|p|^2}{2}}}{(2\pi\beta^{-1})^{\frac{d}{2}}}\mathrm{d}p, \tag{10}$$

in terms of weak convergence of measures.

Furthermore, if $\lambda^{(\gamma)}$ is the γ-dependent eigenvalue in Proposition 3.3 then

$$\lambda^{(\gamma)} \underset{\gamma\to\infty}{\sim} \frac{\overline{\lambda}}{\gamma}, \tag{11}$$

where $\overline{\lambda}$ is the similar eigenvalue obtained in Proposition 3.3 for the OLD.

The overdamped limit of the QSD writes therefore as a tensor product of the QSD of the OLD on the position domain and a Gaussian density in velocity. In particular, if we consider just the marginal in position we obtain the previously expected convergence. Notice that we are also able to exhibit in (11) an overdamped limit of the eigenvalue involved in the defintion of the QSD.

5 Conclusion

This work has furthered our knowledge of the behaviour of the ULD, even with non-conservative forces. Namely, in Section 3 we considered its long-time behavior when it remains trapped in a metastable state. We showed that its law then converges to a local equilibrium distribution corresponding to the QSD. Other associated properties like the law of the first exit point from the domain, starting from the QSD, are obtained in this case. This allows to justify mathematically the Parallel Replica algorithm in Section 2 used for the sampling of metastable thermostated molecular dynamics.

In addition, we consider in Section 4 the behaviour of the ULD when the friction coefficient goes to infinity in (1), corresponding to the overdamped limit. We namely extend the usual overdamped limit known in the literature for the position coordinate to the couple of the position and velocity coordinates. This allows us to make explicit the overdamped limit of the QSD obtained in Section 3 which involves the QSD of the OLD (2) on the position domain. A similar result was obtained recently in [18] for the stationary distribution with an additional speed of convergence. All in all, these results provide a better insight on the ULD behaviour for a large friction coefficient γ.

Declarations

M. Ramil is currently supported by Samsung Science and Technology Foundation. T. Lelièvre is supported by the projects ANR QuAMProcs (ANR-19-CE40-0010), ANR SINEQ (ANR-21-CE40-0006) and the project ERC EMC2 (grant agreement 810367). J. Reygner is supported by the projects ANR EFI (ANR-17-CE40-0030) and ANR QuAMProcs (ANR-19-CE40-0010) from the French National Research Agency. The authors have no relevant financial or non-financial interests to disclose. Data sharing not applicable to this article as no datasets were generated or analysed during the current study.

References

[1] M. Allen and D. Tildesley. Computer simulation of liquids. 1987, vol. 385. *New York: Oxford*, 1989.

[2] A. Asselah, P. A. Ferrari, and P. Groisman. Quasistationary distributions and Fleming-Viot processes in finite spaces. *Journal of Applied Probability*, 48(2):322–332, 2011.

[3] A. Binder, T. Lelièvre, and G. Simpson. A generalized parallel replica dynamics. *Journal of Computational Physics*, 284:595–616, 2015.

[4] K. Burdzy, R. Holyst, D. Ingerman, and P. March. Configurational transition in a Fleming-Viot-type model and probabilistic interpretation of Laplacian eigenfunctions. *Journal of Physics A: Mathematical and General*, 29(11):2633, 1996.

[5] K. Burdzy, R. Hołyst, and P. March. A Fleming–Viot particle representation of the Dirichlet Laplacian. *Communications in Mathematical Physics*, 214(3):679–703, 2000.

[6] 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.

[7] P. Collet, S. Martínez, and J. San Martín. *Quasi-stationary distributions*. Probability and its Applications (New York). Springer, Heidelberg, 2013. Markov chains, diffusions and dynamical systems.

[8] G. Di Gesu, T. Lelièvre, D. Le Peutrec, and B. Nectoux. Sharp asymptotics of the first exit point density. *Annals of PDE*, 5(1):1–174, 2019.

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

[10] A. Guillin, B. Nectoux, and L. Wu. Quasi-stationary distribution for strongly Feller Markov processes by Lyapunov functions and applications to hypoelliptic Hamiltonian systems. *https://hal.archives-ouvertes.fr/hal-03068461/*, 2020.

[11] F. Hédin and T. Lelièvre. gen.parrep: A first implementation of the Generalized Parallel Replica dynamics for the long time simulation of metastable biochemical systems. *Computer Physics Communications*, 239:311–324, 2019.

[12] 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.

[13] B. Leimkuhler and C. Matthews. Numerical methods for stochastic molecular dynamics. In *Molecular Dynamics*, pages 261–328. Springer, 2015.

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

[15] T. Lelièvre and G. Stoltz. Partial differential equations and stochastic methods in molecular dynamics. *Acta Numer.*, 25:681–880, 2016.

[16] T. Lelièvre, M. Ramil, and J. Reygner. A probabilistic study of the kinetic Fokker-Planck equation in cylindrical domains. *Journal of Evolution Equations*, 22:38, 2022.

[17] 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.

[18] P. Monmarché and M. Ramil. Overdamped limit at stationarity for non-equilibrium Langevin diffusions. *Electronic Communications in Probability*, 27:1–8, 2022.

[19] 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.

[20] D. Perez, B.P. Uberuaga, and A.F. Voter. The parallel replica dynamics method–Coming of age. *Computational Materials Science*, 100:90–103, 2015.

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

[22] D. C. Rapaport. *The art of molecular dynamics simulation*. Cambridge university press, 2004.

[23] M. R. Sorensen and A. Voter. Temperature-accelerated dynamics for simulation of infrequent events. *The Journal of Chemical Physics*, 112(21):9599–9606, 2000.

[24] M. Tuckerman. *Statistical mechanics: theory and molecular simulation*. Oxford university press, 2010.

[25] A. Voter. A method for accelerating the molecular dynamics simulation of infrequent events. *The Journal of chemical physics*, 106(11):4665–4677, 1997.

[26] A. F. Voter, F. Montalenti, and T. C. Germann. Extending the time scale in atomistic simulation of materials. *Annual review of materials research*, 32(1):321–346, 2002.

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

[28] R.J. Zamora, D. Perez, E. Martinez, B. P. Uberuaga, and A. F. Voter. Accelerated molecular dynamics methods in a massively parallel world. *Handbook of Materials Modeling: Methods: Theory and Modeling*, pages 745–772, 2020.