

SMAI-JCM

SMAI JOURNAL OF
COMPUTATIONAL MATHEMATICS

The velocity jump Langevin process
and its splitting scheme: long time
convergence and numerical accuracy

NICOLAÏ GOURAUD, LUCAS JOURNAL & PIERRE MONMARCHÉ

Volume 11 (2025), p. 533-584.

<https://doi.org/10.5802/smai-jcm.134>

© The authors, 2025.



*The SMAI Journal of Computational Mathematics is a member
of the Centre Mersenne for Open Scientific Publishing*

<http://www.centre-mersenne.org/>

Submissions at <https://smai-jcm.centre-mersenne.org/ojs/submission>

e-ISSN: 2426-8399



Inria



The velocity jump Langevin process and its splitting scheme: long time convergence and numerical accuracy

NICOLAÏ GOURAUD¹
LUCAS JOURNEL²
PIERRE MONMARCHÉ³

¹ Qubit Pharmaceuticals, Paris, France and Sorbonne Université, LJLL, UMR 7598 CNRS and LCT, UMR 7616 CNRS, Paris, France

E-mail address: nicolai@qubit-pharmaceuticals.com

² Institut de Mathématiques, Université de Neuchâtel, Switzerland

E-mail address: lucas.journal@unine.ch

³ Sorbonne Université, LJLL, UMR 7598 CNRS and LCT, UMR 7616 CNRS, Paris, France and Institut Universitaire de France, Paris, France

E-mail address: pierre.monmarche@sorbonne-universite.fr.

Abstract. The Langevin dynamics is a diffusion process extensively used, in particular in molecular dynamics simulations, to sample Gibbs measures. Some alternatives based on (piecewise deterministic) kinetic velocity jump processes have gained interest over the last decade. One interest of the latter is the possibility to split forces (at the continuous-time level), reducing the numerical cost for sampling the trajectory. Motivated by this, a numerical scheme based on hybrid dynamics combining velocity jumps and Langevin diffusion, numerically more efficient than their classical Langevin counterparts, has been introduced for computational chemistry in [43]. The present work is devoted to the numerical analysis of this scheme. Our main results are, first, the exponential ergodicity of the continuous-time velocity jump Langevin process, second, a Talay–Tubaro expansion of the invariant measure of the numerical scheme on the torus, showing in particular that the scheme is of weak order 2 in the step-size and, third, a bound on the quadratic risk of the corresponding practical MCMC estimator (possibly with Richardson extrapolation). With respect to previous works on the Langevin diffusion, new difficulties arise from the jump operator, which is non-local.

2020 Mathematics Subject Classification. 65C40, 60J76, 60J60.

Keywords. velocity jump process, Langevin diffusion, piecewise deterministic Markov process, molecular dynamics, exponential ergodicity, hypocoercivity, Talay–Tubaro expansion, weak discretization error.

1. Motivations and overview

Molecular dynamics (MD) is a popular numerical tool to infer macroscopic properties of matter from simulations at the microscopic level. In the framework of statistical physics, thermodynamical quantities are seen as average values of certain functions called observables with respect to a probability measure. The primary objective of MD is to efficiently sample these measures, typically by computing long trajectories of stochastic processes. Here, we focus on the so-called canonical ensemble, in which the number of particles, the volume of the system and its temperature are fixed.

Let us consider a system of N interacting particles. Denote by $x \in \mathcal{X}$ their positions and $v \in \mathbb{R}^{3N}$ their velocities, with $\mathcal{X} = \mathbb{R}^{3N}$ or $\mathcal{X} = \mathbb{T}^{3N} = \mathbb{R}^{3N}/\mathbb{Z}^{3N}$ if periodic boundary conditions are enforced. Let $M = \text{diag}(m_1 I_3, \dots, m_N I_3)$ be the mass matrix of the particles, $\beta = 1/(k_B T)$ the

N.G. acknowledges a Ph.D. fellowship from Qubit Pharmaceuticals.

L.J. is supported by the grant n°200029-21991311 from the Swiss National Science Foundation.

The research of P.M. is supported by the projects SWIDIMS (ANR-20-CE40-0022) and CONVIVIALITY (ANR-23-CE40-0003) of the French National Research Agency.

<https://doi.org/10.5802/smai-jcm.134>

© The authors, 2025

inverse temperature of the system (where k_B is the Boltzmann constant), $U : \mathbb{R}^{3N} \rightarrow \mathbb{R}$ the potential energy function encoding the interactions between particles and finally $H(x, v) = U(x) + \frac{1}{2}v^T M v$ the Hamiltonian of the system, corresponding to its total energy. In the canonical ensemble, the statistical distribution of microscopic configurations is described by the Boltzmann–Gibbs measure, defined by

$$d\mu(x, v) = \mathcal{Z}_\mu^{-1} \exp(-\beta H(x, v)) dx dv, \quad (1.1)$$

where $\mathcal{Z}_\mu = \int_{\mathbb{R}^{6N}} \exp(-\beta H(x, v)) dx dv$. In most cases, computing macroscopic quantities of the form $\mathbb{E}_\mu[\varphi]$ analytically or with deterministic numerical quadrature is out of reach, as they are high-dimensional integrals involving an unknown constant $\mathcal{Z}_\mu$. Instead, they can be approximated by simulating a Markov process $(X_t, V_t)_{t \geq 0}$ that is ergodic with respect to μ , which means that for any φ (in a certain class of functions), almost surely:

$$\frac{1}{t} \int_0^t \varphi(X_s, V_s) ds \xrightarrow[t \rightarrow \infty]{} \int_{\mathbb{R}^{6N}} \varphi(x, v) d\mu(x, v) = \mathbb{E}_\mu[\varphi].$$

A popular process that has this property (under mild conditions on the potential U) is the kinetic Langevin diffusion, defined as the solution of the following SDE:

$$\begin{cases} dX_t = V_t dt \\ dV_t = -M^{-1} \nabla U(X_t) dt - \gamma M^{-1} V_t dt - M^{-1} \sqrt{2\gamma\beta^{-1}} dB_t, \end{cases} \quad (1.2)$$

where $(B_t)_{t \geq 0}$ is a standard Brownian motion in $\mathbb{R}^{3N}$ and $\gamma > 0$ is a friction parameter. In practice, the Langevin diffusion cannot be simulated exactly, leading to discretization splitting schemes such as BAOAB [9, 28, 29, 30] where the parts of the dynamics corresponding to the free transport, the forces and the friction/dissipation are simulated separately.

When simulating the dynamics, the most expensive part is the computation of the forces ∇U . A common approach to reduce the cost is to substitute this gradient with a stochastic version, as in the stochastic gradient descent in optimization (ubiquitous in machine learning [8], see also examples in MD in [23, 24, 46]), or to treat various parts of the potentials with different time steps (multi-time-step methods), see [49]. We are interested here in an alternative approach that involves replacing (1.2) with a hybrid model combining a classical Langevin diffusion and a piecewise deterministic Markov process. This hybrid model still samples from μ but can be simulated using a numerical splitting scheme that requires fewer gradient computations per time step.

The idea involves decomposing ∇U as a sum of vector fields $(F_i)_{0 \leq i \leq K}$ for some $K \in \mathbb{N}$, where F_0 denotes the computationally inexpensive component (typically in MD the short-range forces exhibiting fast variation), while the F_i terms, for $i \geq 1$, represent long-range forces, which are computationally more demanding than F_0 (as each atom interacts with all others through these forces, unlike the short-range ones). Then, we define the velocity jump Langevin process $(X_t, V_t)_{t \geq 0}$ as the Markov process which follows the Langevin diffusion process associated with the force F_0 , with the velocity V_t undergoing additional random jumps at rate $\lambda_i(x, v)$ following a jump kernel $q_i(x, v, dv')$ (both defined below in (2.3)), that depend on F_i in a way that ensures that the equilibrium measure of the process is indeed the Gibbs measure μ .

This process can be simulated with a splitting scheme similar to BAOAB, but with an additional part arising from the velocity jumps. The simulation of the jump times is based on the thinning method, see [31, 32]. Suppose that $\lambda_i(x, v) \leq \bar{\lambda}_i$ for $1 \leq i \leq K$, and denote $\bar{\lambda} = \sum_i \bar{\lambda}_i$. A direct computation on the generator shows that the jumps can be exactly simulated this way: starting from (x, v) ,

- (1) Draw $\mathcal{E}$ a standard exponential random variable, and let $T = \mathcal{E}/\bar{\lambda}$ be the next jump time proposal.

- (2) Draw I in $\llbracket 1, K \rrbracket$ such that $\mathbb{P}(I = i) = \bar{\lambda}_i / \bar{\lambda}$. Propose a jump of type I at time T .
- (3) Accept the jump with probability $\lambda_I(x, v) / \bar{\lambda}_I$, in which case the velocity is resampled at time T according to $q_I(x, v, dv')$, otherwise the velocity at time T is simply v (there is no jump).

Unlike a traditional numerical scheme for the Langevin diffusion (like BAOAB), where the gradient ∇U is computed at every time step, here we only need to evaluate F_i when a jump of type i is proposed (at step 3 above), which does not occur at every time step for every i . However, similarly to BAOAB and contrary to multi-time-step splitting methods, there is still a unique step-size in the discretization and no additional parameters to tune: the frequency at which F_i is evaluated is random and is adapted to each force through the bound $\bar{\lambda}_i$. By contrast, when using piecewise deterministic Markov processes such as the Bouncy Particle (BPS) [10, 36, 45] or Zig-Zag [5] samplers (which can be seen as particular cases of the velocity jump Langevin diffusion where $F_0 = 0$ and $\gamma = 0$, so that, between velocity jumps, the deterministic motion is simply the free transport $\dot{x} = v$), thinning is usually employed to sample exactly the full continuous-time process, without any time-discretization. However, when $F_0 \neq 0$, this is only possible when it is a linear function (as in the so-called boomerang sampler [6]), which is not the case we are interested in. Besides, note that the use of discretization schemes even for piecewise deterministic Markov processes have recently gained interest [3, 4, 12, 44].

The idea of the hybrid velocity jump Langevin approach was first introduced in [43], in the particular case where the jump mechanism is the one of the BPS, with a decomposition of the forces based on long-range pairwise interactions. Later, in [21], a generalized version was introduced in which the jump mechanism, originally defined in [42], can be seen as an interpolation between the BPS and the Hamiltonian dynamics, which is shown to be numerically more stable, while preserving dynamical properties, also of interest in MD.

While [21] presents an implementation of the Langevin velocity jump approach on a molecular dynamics code and numerical experiments, the present work is devoted to its mathematical study. We prove a second order weak error expansion *à la Talay–Tubaro* [48] of the discretization bias on the invariant measure, see Theorem 2.2. The proof is based on classical weak backward error analysis arguments [27, 30, 34, 48] and, in particular, crucially relies on the ergodicity properties of the continuous-time process, which is of interest by itself and is in our case the topic of Theorem 2.1.

By comparison with previous works, the main new difficulties arise from the combination of, first, the degenerate diffusion feature of the kinetic Langevin diffusion, second, the non-local jump operator and, third, the necessity to get pointwise bounds on derivatives of the semi-group. Indeed, the exponential L^2 convergence estimates obtained for kinetic piecewise deterministic Markov processes in recent works [1, 42] with the Dolbeault–Mouhot–Schmeiser modified norm [15] are not sufficient to conduct the rest of the argument, and neither are the results in V -norm based on Harris theorem [7, 11, 13, 17, 36] or in Wasserstein distances using coupling [14]. We thus have to work with modified Sobolev norms as originally done in [47], then generalized and applied in numerous works [25, 26, 50, 52]. Thus, one of our main contributions is that we successfully implemented this method in a case with a non-local operator (i.e. a non-diffusive Markov process), leading to Theorem 3.1 (which is an important step in the proof of Theorem 2.1). Although the H^1 modified norm method has mostly been used for diffusion processes, it has already been successfully applied to non-diffusive dynamics in [14, 20, 35, 38]; however, let us emphasize that, contrary to our case, in all these works the dynamics is a contraction of the Wasserstein 2 distance, from which the exponential decay of the H^1 norm is clear, and in fact the computations to prove the decay of the Wasserstein distance and of the gradient part of the H^1 norm are essentially the same (as discussed in [35, 39, 41]). In particular, in these cases, to get the decay of the H^1 norm, the dissipation of the L^2 part of the norm is not exploited, which in the general case (i.e. without a Wasserstein 2 contraction) is a crucial feature of Villani’s method (as in Lemma 3.7 below). Note that this is possible in our case because, although the process undergoes non-local jumps, it also

has a diffusive part, from which the L^2 dissipation is bounded below by a first-order term which would be the dissipation of the usual Langevin diffusion (cf. the *carré du champ* operator (3.8)).

A second important difficulty is the design of a suitable Lyapunov function for the velocity jump Langevin diffusion, given its combination of diffusive, non-local and kinetic features (see Propositions 3.2 for the continuous-time process and Proposition 4.1 for the numerical scheme). The last important ingredient for establishing our main result is a finite-time numerical analysis of the discretization bias, Theorem 4.5, whose proof is close to the proof of Theorem 2.6 of the recent [4] (which is a similar result for piecewise deterministic samplers).

This work is organized as follows. In Section 2, the process and its splitting scheme are defined, our main results are stated, and some comments on the method from the applicable angle are given. Section 3 is devoted to the proof of Theorem 2.1 (the long-time convergence of the continuous-time process). The weak error result, Theorem 2.2, is proven in Section 4.

2. Setting and main results

2.1. Notations

In all the following, $|\cdot|$ denotes the Euclidean norm and $\cdot$ the standard dot product. In the case of multi-indices $\alpha \in \mathbb{N}^d$, we also denote $|\alpha| = \sum_{i=1}^d \alpha_i$. The set of smooth functions with compact support from $\mathbb{R}^{2d}$ to $\mathbb{R}$ is denoted $\mathcal{C}_c^\infty(\mathbb{R}^{2d})$. For any operators A and B , we denote $[A, B] = AB - BA$ their commutator. In the remainder of this work, we will denote by C various constants that may change from line to line.

2.2. Definitions and assumptions

Fix some smooth potential $U \in \mathcal{C}^\infty(\mathcal{X}, \mathbb{R})$ where either $\mathcal{X} = \mathbb{R}^d$ or $\mathcal{X} = \mathbb{T}^d$ for some $d \geq 1$. In the case of particle systems, $d = 3N$. For simplicity, in the remainder of this work, we fix $\beta = 1$, $M = I_d$ and consider the following decomposition of the forces. Assuming that $U = U_0 + U_1$ with U_0, U_1 smooth, let

$$F_0 = \nabla U_0, \quad F_i = \partial_i U_1 e_i, \quad i \in \llbracket 1, d \rrbracket, \quad (2.1)$$

where $(e_i)_{1 \leq i \leq d}$ stands for the canonical basis of $\mathbb{R}^d$. In other words, we decompose ∇U_1 in a similar fashion as in the Zig-Zag process:

$$\nabla U_1(x) = \sum_{i=1}^d \partial_i U_1(x) e_i.$$

Fix a friction parameter $\gamma > 0$, a jump parameter $\rho \in [-1, 1)$ and an activation function $\Psi \in \mathcal{C}^\infty(\mathbb{R}, \mathbb{R}_+)$. The velocity jump Langevin process is defined as the Markov process on $\mathcal{X} \times \mathbb{R}^d$ with infinitesimal generator given by

$$\mathcal{L} = \mathcal{L}_H + \mathcal{L}_D + \mathcal{L}_J, \quad (2.2)$$

where for any function $h \in \mathcal{C}_c^\infty(\mathcal{X} \times \mathbb{R}^d)$:

$$\mathcal{L}_H h(x, v) = v \cdot \nabla_x h(x, v) - \nabla U_0(x) \cdot \nabla_v h(x, v) \quad (\text{Hamiltonian dynamics})$$

$$\mathcal{L}_D h(x, v) = -\gamma v \cdot \nabla_v h(x, v) + \gamma \Delta_v h(x, v) \quad (\text{friction/dissipation})$$

$$\mathcal{L}_J h(x, v) = \sum_{i=1}^d \mathcal{L}_J^i h(x, v) \quad (\text{velocity jumps})$$

with

$$\mathcal{L}_J^i h(x, v) = \lambda_i(x, v) \int_{\mathbb{R}^d} (h(x, v') - h(x, v)) q_i(x, v, dv'),$$

where λ_i, q_i satisfy

$$\lambda_i(x, v) \int_{\mathbb{R}^d} h(x, v') q_i(x, v, dv') = \frac{2}{1-\rho} \mathbb{E} \left[h(x, V^i) \Psi \left(\frac{\partial_i U_1(x)}{2} (v_i - V_i^i) \right) \right], \quad (2.3)$$

and where

$$V_i^i = \rho v_i + \sqrt{1-\rho^2} \xi, \quad V_j^i = v_j \quad \forall j \neq i,$$

and $\xi \sim \mathcal{N}(0, 1)$ is a one-dimensional standard Gaussian variable. In the definition above, $q_i(x, v, \cdot)$ is a probability measure for all $(x, v) \in \mathcal{X} \times \mathbb{R}^d$, in other words λ_i , for all $i \in \llbracket 1, d \rrbracket$, is obtained by taking $h = 1$ in the formula above.

We will work under the following set of assumptions on U and Ψ .

Assumption 1.

- (1) *the Gibbs measure is well defined, i.e. $\int_{\mathcal{X}} e^{-U(x)} dx < \infty$.*
- (2) *The derivatives of U_1 of all order are bounded.*
- (3) *$\nabla^2 U_0$ as well as its derivatives of all order are bounded, and there exist $\kappa > 0$ and a compact set $\mathcal{K} \subset \mathcal{X}$ such that for all $x \notin \mathcal{K}$*

$$-x \cdot \nabla U_0(x) \leq -\kappa |x|^2.$$

- (4) *Ψ has all its derivatives bounded, and for all $s \in \mathbb{R}$,*

$$\Psi(s) - \Psi(-s) = s. \quad (2.4)$$

Under Assumption 1, it is not difficult to see that the process is well-defined, see [42]. Let (X_t, V_t) be a Markov process with generator $\mathcal{L}$, and let $(P_t)_{t \geq 0}$ be the associated semi-group,

$$P_t f(x, v) = \mathbb{E}_{(x, v)} [f(X_t, V_t)],$$

defined for suitable $f : \mathcal{X} \times \mathbb{R}^d \rightarrow \mathbb{R}$. Define as well the set of admissible functions:

$$\mathcal{A} := \left\{ f \in \mathcal{C}^\infty(\mathcal{X} \times \mathbb{R}^d, \mathbb{R}) \mid \forall \alpha \in \mathbb{N}^{2d}, \exists C > 0, 0 < c < 1, |\partial^\alpha f| \leq C e^{cH} \right\}.$$

Let us now introduce a splitting scheme for the process. Starting from (2.2), we further split the Hamiltonian part as $\mathcal{L}_H = \mathcal{L}_A + \mathcal{L}_B$ where $\mathcal{L}_A = v \cdot \nabla_x$ stands for the free transport and $\mathcal{L}_B = -\nabla U_0 \cdot \nabla_v$ for the acceleration. Motivated by the Trotter/Strang formula

$$e^{t\mathcal{L}} = e^{\frac{t}{2}\mathcal{L}_B} e^{\frac{t}{2}\mathcal{L}_J} e^{\frac{t}{2}\mathcal{L}_A} e^{t\mathcal{L}_O} e^{\frac{t}{2}\mathcal{L}_A} e^{\frac{t}{2}\mathcal{L}_J} e^{\frac{t}{2}\mathcal{L}_B} + \mathcal{O}(t^3), \quad (2.5)$$

given a step-size $\delta > 0$, we call BJAOAJB the Markov chain $(\bar{X}_n, \bar{V}_n)$ on $\mathcal{X} \times \mathbb{R}^d$ with transition kernel

$$Q = e^{\frac{\delta}{2}\mathcal{L}_B} e^{\frac{\delta}{2}\mathcal{L}_J} e^{\frac{\delta}{2}\mathcal{L}_A} e^{\delta\mathcal{L}_O} e^{\frac{\delta}{2}\mathcal{L}_A} e^{\frac{\delta}{2}\mathcal{L}_J} e^{\frac{\delta}{2}\mathcal{L}_B}. \quad (2.6)$$

From the definition of $\mathcal{L}_A$, $\mathcal{L}_B$, $\mathcal{L}_J$ and $\mathcal{L}_O$, we have, for $t > 0$,

$$\begin{aligned} e^{t\mathcal{L}_B} f(x, v) &= f(x, v - t\nabla U_0(x)), \\ e^{t\mathcal{L}_A} f(x, v) &= f(x + tv, v), \\ e^{t\mathcal{L}_O} f(x, v) &= \mathbb{E} \left[f(x, e^{-\gamma t} v + \sqrt{1 - e^{-2\gamma t}} \xi) \right] \quad \text{with } \xi \sim \mathcal{N}(0, I_d), \\ e^{t\mathcal{L}_J} f(x, v) &= \mathbb{E} [f(x, W_t)], \end{aligned}$$

where $(W_s)_{s \geq 0}$ is a continuous-time Markov chain on $\mathbb{R}^d$ initialized at $W_0 = v$ and with jump rate $\lambda(x, \cdot)$ and jump kernel $q(x, \cdot, \cdot)$ given by

$$\lambda(x, w) = \sum_{i=1}^d \lambda_i(x, w), \quad \lambda(x, w)q(x, w, dw') = \sum_{i=1}^d \lambda_i(x, w)q_i(x, w, dw').$$

Hence, one transition of the chain $(\bar{X}_n, \bar{V}_n)$ is given by the succession of steps BJAQJB, where:

$$\begin{aligned} (B) : v &\leftarrow v - \frac{\delta}{2} \nabla U_0(x) \\ (J) : v &\leftarrow W_{\delta/2} \\ (A) : x &\leftarrow x + \frac{\delta}{2} v \\ (O) : v &\leftarrow e^{-\gamma\delta} v + \sqrt{1 - e^{-2\gamma\delta}} \xi. \end{aligned}$$

2.3. Main results

Our first main result, established in Section 3, gives point-wise estimates for the exponential long-time convergence of the semi-group, together with all its derivatives.

Theorem 2.1. *Under Assumption 1, for all $f \in \mathcal{A}$ and all multi-index $\alpha \in \mathbb{N}^{2d}$, there exist $C, q > 0$, $b \in (0, 1)$, such that for all $t > 0$, $(x, v) \in \mathbb{R}^{2d}$,*

$$|\partial^\alpha (P_t f - \mu(f))(x, v)| \leq C e^{-qt} e^{bH(x, v)}. \quad (2.7)$$

Note that although $e^{bH(x, v)}$ grows fast as x and v go to infinity, this term is an optimal prefactor depending on the initial conditions of the process. When $t = 0$, the inequality reads $|\partial^\alpha f(x, v) - \mu(f)| \leq C e^{bH(x, v)}$, which corresponds to the assumption $f \in \mathcal{A}$, i.e. that f and its derivatives grow at most exponentially. In particular, the case $\alpha = 0$ in (2.7) corresponds to the weak convergence of the process for all test functions in $\mathcal{A}$. In addition, since $b < 1$, a crucial property of the right-hand side of (2.7) is that it is μ -integrable, which will be used to control the remainder terms in the expansion of the weak error of the numerical scheme. In addition, ergodicity of the continuous-time process is necessary to deduce the expansion of the invariant measure of the numerical scheme from the finite-time error estimates, by letting t go to infinity.

Formally, Equation (2.5) yields that this palindromic form gives rise to a second-order scheme. Our second main result is that, indeed, the discretization bias on the invariant measure is of order δ^2 . Here the analysis is restricted to the compact torus, since this is the main case of interest in practice in MD and otherwise the construction of a Lyapunov function is more intricate for the numerical scheme than the continuous-time process (as in other cases such as [25]).

Theorem 2.2. *In the case $\mathcal{X} = \mathbb{T}^d$, under Assumption 1, there exists $\delta_0 > 0$ such that for all $\delta \in (0, \delta_0]$, the BJAQJB chain with transition kernel Q admits a unique invariant measure μ_δ . Moreover, for all $f \in \mathcal{A}$, there exists $c_f \in \mathbb{R}$ such that*

$$\mu_\delta(f) = \mu(f) + c_f \delta^2 + \mathcal{O}_{\delta \rightarrow 0}(\delta^3).$$

The proof of this theorem is the topic of Section 4. Along the proof of Theorem 2.2, we have to prove that the Markov chain with transition Q is ergodic. In fact we prove a more precise statement, namely that the BJAQJB chain converges to its equilibrium in V -norm exponentially fast with a rate which, as should be expected, is linear in the step-size, as we state now. Given $V : \mathbb{T}^d \times \mathbb{R}^d \rightarrow [1, \infty)$, for $f : \mathbb{T}^d \times \mathbb{R}^d \rightarrow \mathbb{R}$ and a signed measure μ over $\mathbb{T}^d \times \mathbb{R}^d$ we write $\|f\|_V = \|f/V\|_\infty$ and $\|\mu\|_V = \sup\{\mu(f), \|f\|_V \leq 1\}$. The next statement is established in Section 4.3.

Theorem 2.3. *Let $b \in (0, 1/2)$ and $V(x, v) = e^{b|v|^2}$ for $(x, v) \in \mathbb{T}^d \times \mathbb{R}^d$. In the case $\mathcal{X} = \mathbb{T}^d$, under Assumption 1, there exist $\delta_0, \lambda, C > 0$ such that for all $\delta \in (0, \delta_0]$, $(x, v) \in \mathbb{T}^d \times \mathbb{R}^d$,*

$$\|\delta_{(x,v)} Q^n - \mu_\delta\|_V \leq C e^{-\lambda n \delta} V(x, v).$$

Combining Theorems 2.2 and 2.3, classical arguments then allow to bound the quadratic risk of an MCMC estimator based on the BJAOJB chain, possibly with a Richardson extrapolation, as detailed in Section 4.5:

Corollary 2.4. *Let $b \in (0, 1/4)$ and $V(x, v) = e^{b|v|^2}$ for $(x, v) \in \mathbb{T}^d \times \mathbb{R}^d$. In the case $\mathcal{X} = \mathbb{T}^d$, under Assumption 1, there exist $\delta_0 > 0$ such that for all $f \in \mathcal{A}$ with $\|f\|_V < \infty$, there exists $C > 0$ such that for all $n \in \mathbb{N}$ and all $\delta \in (0, \delta_0)$, given $(X_n, V_n)_{n \in \mathbb{N}}$ a BJAOJB chain,*

$$\mathbb{E} \left(\left| \frac{1}{n} \sum_{k=1}^n f(X_k) - \mu(f) \right|^2 \right) \leq C \left(\delta^4 + \frac{1}{n\delta} \mathbb{E} \left(e^{2b|V_0|^2} \right) \right).$$

Moreover, if $(\tilde{X}_n, \tilde{V}_n)_{n \in \mathbb{N}}$ is another independent BJAOJB chain with step-size $\delta/2$,

$$\mathbb{E} \left(\left| \frac{4}{3n} \sum_{k=1}^n f(\tilde{X}_k) - \frac{1}{3n} \sum_{k=1}^n f(X_k) - \mu(f) \right|^2 \right) \leq C \left(\delta^6 + \frac{1}{n\delta} \mathbb{E} \left(e^{2b|V_0|^2} + e^{2b|\tilde{V}_0|^2} \right) \right).$$

Here we state a simple bound but it is easily checked in the proof that in fact the dependency in the initial condition disappears exponentially fast with $n\delta$, i.e. in the bounds above, $\mathbb{E}(e^{2b|V_0|^2})$ can be replaced by $1 + e^{-cn\delta} \mathbb{E}(e^{2b|V_0|^2})$ for some $c > 0$, and similarly with $\tilde{V}_0$.

2.4. Discussion and applications

Decomposition of the forces. Following the construction of the jump mechanism presented in [42], the velocity jump Langevin process could be defined for any decomposition of ∇U_1 into arbitrary vector fields F_i . In the present work, the choice of decomposing ∇U_1 along the canonical basis of $\mathbb{R}^d$ was made to avoid singular terms of the form $\nabla U_1/|\nabla U_1|$ in the hypocoercivity computations. However, each of the F_i in (2.1) could be further decomposed along the vector e_i , and all the results presented in the previous sections would still hold, as the proofs would lead to exactly the same computations.

To give an example, consider for instance the following classical functional form

$$U = U_0 + \sum_{1 \leq i < j \leq d} g(|x_i - x_j|),$$

where U_0 gathers short-range, computationally inexpensive terms, to be treated by the Langevin part of the process, and $U_1 = U - U_0$ is a sum of pairwise long-range terms (for instance, in molecular dynamics, long-range electrostatic and van der Waals interactions), with $g : \mathbb{R} \rightarrow \mathbb{R}$. Here, the term “long-range” refers to the fact that we suppose that there exists a cutoff radius $r_c > 0$ such that $g(r) = 0$ for all $r < r_c$, which implies that there are no singularities when $x_i = x_j$ for any i, j . Following the decomposition (2.1), for all $i \in \llbracket 1, d \rrbracket$, $F_i(x) = \sum_{j \neq i} \partial_i g(|x_i - x_j|) e_i$. We can then define $F_{ij} = \partial_i g(|x_i - x_j|) e_i$ (namely the force of the particle j acting on the particle i), and associate a jump mechanism to each F_{ij} .

This type of decomposition was used in the implementations of the method for MD applications presented in [21, 43], except that the splitting was done at the level of the three-dimensional atoms instead of individual coordinates.

Conditions on U . In general, when using piecewise deterministic samplers such as BPS or Zig-Zag, it is not necessary to assume that ∇U is bounded: jump times can be sampled with thinning using a time-inhomogeneous bound (for instance, if $\nabla^2 U$ is bounded, we can bound $|\nabla U(x + tv)| \leq |\nabla U(x)| + t\|\nabla^2 U\|_\infty|v|$). We could do the same here but this is less convenient because, between jumps, the process follows a random Langevin diffusion and thus we cannot have an almost sure bound on $|\nabla U_1(x_t)|$. Regardless, this is not our cases of interest, which are the practical settings of [21, 43] in MD simulations. In this situation, U has some singular parts (short-range forces) and thus, even if it is possible in principle to use thinning, it is not efficient. The point made in [21, 43] is precisely that thinning is efficient for bounded forces, so that in these works ∇U_1 only gathers long-range interaction forces, which vanish at infinity. This is why, in Assumption 1, we assume ∇U_1 to be bounded. In our theoretical study, this condition is used in the proof of Lemma 3.9.

Note however that when $\mathcal{X} = \mathbb{T}^d$, the condition on U in Assumption 1 is simply that U_0 and U_1 are smooth (taking $\mathcal{K} = \mathbb{T}^d$).

Choice of Ψ . The condition (2.4) on Ψ is here to ensure the invariance of μ . The usual choice, used in the Zig-Zag or BPS samplers, and in the practical implementations in [21, 43] is $\Psi(s) = \max(0, s)$, which indeed satisfies (2.4). However, the computations in Sobolev spaces in our proofs require that the generator has smooth coefficients, which prevents such a Ψ . An admissible example is given by $\Psi(s) = a \log(\exp(s/a) + 1)$ for any $a > 0$.

Role of ρ . In the jump mechanism, the case $\rho = -1$ corresponds to the Bouncy Particle Sampler, namely $\lambda_i(x, v) = \Psi(\partial_i U_1(x)v_i)$, with deterministic jumps: $q_i(x, v, dv') = \delta_{(x, -v_i)}$. This version, introduced and implemented in [43] is the most natural one and simplest to implement. Moreover it yields the least amount of jumps (since λ_i grows with ρ), and therefore corresponds to the best computational speedup.

However, as explored in [21], choosing the BPS as a jump mechanism has also several issues. First, numerical instabilities may arise in MD applications. Then, although the process samples the Gibbs measure (1.1), dynamical properties such as diffusion constants are not well preserved. Indeed, a specificity of MD applications, and contrary to other fields such as Bayesian statistics, is that sampling (1.1) or its position marginal (with density proportional to $e^{-\beta U}$) is not the only objective, and dynamical properties such as transition times or transport coefficients are also quantities of interest. In MD, the Langevin diffusion is in fact a model for the time evolution of molecular systems, and it is usually not possible to replace (1.2) by some other Markov sampler of the Gibbs measure, like the overdamped Langevin diffusion, Hamiltonian Monte Carlo or random walk Metropolis. Additionally, in that case, in (1.2), the friction parameter γ is fixed instead of chosen by the user.

That being said, the generalized jump mechanism (2.3), first introduced in [42] as a general class of kinetic jump processes sampling the Gibbs measure (of which the BPS and Zig-Zag process are particular cases), was integrated to the velocity jump Langevin approach in [21]. An interesting property, proved in [42, Theorem 3.6], is that this process can be tuned to be arbitrarily close to the Langevin diffusion in terms of stochastic trajectories: when $\rho \rightarrow 1$, the distribution of the velocity jump Langevin process converges to the Langevin diffusion in the space of *càdlàg* trajectories endowed with the Skorokhod topology. Therefore, the general expression (2.3) can be seen as an interpolation between the BPS and Hamiltonian dynamics associated to U_1 . As shown numerically in [21], a choice of ρ not too far from 1 makes the velocity jump Langevin process more suitable to estimate the dynamical statistics, and numerically more stable, with, of course, a trade-off between the accuracy of these dynamical properties and the numerical cost of the simulation, as would be the case with any numerical approximation of (1.2). Indeed, as ρ grows, there are more jumps, therefore more pairwise interactions to compute.

In addition, note that in the present work, we take $\rho \in [-1, 1)$ as a constant for simplicity, but it could actually be taken as a function of x or depend on $i \in \llbracket 1, d \rrbracket$.

Simulation of the jumps. As stated in the introduction, when simulating the jump process with generator $\mathcal{L}_J$, jump times can be sampled by using the thinning procedure, provided that a bound on the jump rates λ_i is available. In our case, since Ψ' and ∇U_1 are bounded, for each $i \in \llbracket 1, d \rrbracket$,

$$\begin{aligned} \lambda_i(x, v) &= \frac{2}{1-\rho} \mathbb{E} \left[\Psi \left(\frac{\partial_i U_1(x)}{2} (v_i - V_i^i) \right) \right] \\ &\leq \frac{1}{1-\rho} (2\Psi(0) + \|\Psi'\|_\infty \|\partial_i U_1\|_\infty \mathbb{E}[|v_i - V_i^i|]) \\ &\leq \frac{1}{1-\rho} \left(2\Psi(0) + \|\Psi'\|_\infty \|\partial_i U_1\|_\infty \left((1-\rho)|v_i| + \sqrt{\frac{2(1-\rho^2)}{\pi}} \right) \right) =: \bar{\lambda}_i(|v_i|). \end{aligned}$$

Note that since this bound depends on $|v_i|$, the value of $\bar{\lambda}_i$ changes after each jump, except if $\rho = -1$, where v_i becomes $-v_i$ after a jump of type i . In this case, using the properties of exponential distributions, the simulation of $\mathcal{L}_J$ during a time δ using thinning is simpler: starting from (x, v) ,

- (1) Draw $M \sim \text{Poisson}(\delta \sum_i \bar{\lambda}_i(|v_i|))$, the total number of jump proposals.
- (2) For each $m \in \llbracket 1, M \rrbracket$, draw $I_m \in \llbracket 1, d \rrbracket$ such that $\mathbb{P}(I_m = i) = \bar{\lambda}_i / \sum_{j=1}^d \bar{\lambda}_j$ and with probability $\lambda_{I_m} / \bar{\lambda}_{I_m}$, change the sign of v_{I_m} .

If $\rho > -1$, when a jump of type i occurs, v_i becomes

$$\rho v_i + \sqrt{1-\rho^2} \tilde{\xi},$$

where $\tilde{\xi}$ is a random variable with density proportional to

$$f_{(x,v)}(y) = \Psi \left(\frac{\partial_i U_1(x)}{2} \left((1-\rho)v_i - \sqrt{1-\rho^2}y \right) \right) e^{-y^2/2}.$$

It is shown in [42, Section 5.3], and implemented in [21] in the context of the velocity jump Langevin method, that this random variable can be simulated exactly with rejection sampling, by using either Gamma, exponential, shifted Rayleigh or Gaussian distributions as proposals, depending on the values of (x, v) .

Complexity of the algorithm. Let us discuss the efficiency of the procedure. In a classical discretization scheme of the Langevin diffusion, all the d derivatives of U_1 are computed at each time step, whereas in the BJAOAJB scheme, there are as many computations of terms $\partial_i U_1$ as jump proposals. As we saw, when $\rho = -1$, there are in average $\delta \sum_{i=1}^d \bar{\lambda}_i(|v_i|)$ such proposals per time step. At equilibrium, the velocity v is distributed according to a standard Gaussian, which implies that on average,

$$\mathbb{E}_\mu[\bar{\lambda}_i] = \Psi(0) + \sqrt{\frac{2}{\pi}} \|\Psi'\|_\infty \|\partial_i U_1\|_\infty.$$

Therefore, for $\rho = -1$, we expect the simulation of the velocity jump Langevin using thinning to yield a computational advantage if

$$\delta \left(\Psi(0) + \sqrt{\frac{2}{\pi}} \|\Psi'\|_\infty \|\nabla U_1\|_\infty \right) < 1. \quad (2.8)$$

When ρ increases, the average number of jumps increases, and goes to infinity when $\rho \rightarrow 1$, therefore computational advantage is only to be expected if ρ is not too close to 1. We provide some numerical values in the next paragraph in the case of MD simulations.

Application to molecular dynamics. In MD simulations, the inequality (2.8) is easily satisfied since the commonly used time step is extremely small ($\delta \approx 10^{-15}$ seconds), a choice imposed by the fastest vibrational modes of molecular systems. In [21, 43], the part of the interaction potential to be treated by the jump process is constituted of the long-range electrostatic and van der Waals forces, namely

$$U_1(x) = \sum_{1 \leq i < j \leq N} \chi(r_{ij}) \left(\varepsilon \left[\left(\frac{\sigma}{r_{ij}} \right)^{12} - \left(\frac{\sigma}{r_{ij}} \right)^6 \right] + \frac{q_i q_j}{r_{ij}} \right),$$

where N is the number of the atoms in the system, r_{ij} is the distance between atoms i and j , q_i is the electric charge of the atom i , σ and ε are parameters of the Lennard–Jones potential, and χ is a smooth switching function such that $\chi(r) = 0$ if $r < r_c$, with r_c a given cutoff radius.

In a box of water molecules with the classical SPC/Fw model, by choosing $r_c = 7$ angstroms as a cutoff, the bound on the derivatives of U_1 is of the order of $10^{-2} \text{ kcal.mol}^{-1}.\text{\AA}^{-1}$, and if we consider $\Psi(s) = \log(\exp(s) + 1)$, $\Psi(0) = \log(2)$, $\Psi' < 1$, which implies that (2.8) is largely satisfied, with $\delta = 10^{-3}$ ps in standard MD units, the left hand side of (2.8) being of order 10^{-3} . We refer to [43, Section 4.C] for further theoretical discussion on the speedup provided by the procedure in the Bouncy particle sampler case, when simulating boxes of water molecules.

When $\rho > -1$, the average number of jumps increases. However, as shown extensively in [21] on molecular dynamics simulations, a significant speedup is provided, as long as ρ is not too close to 1, for several reasons. First, thanks to the extremely small time step, the average number of jump proposals per time step remains significantly smaller than the total number of pairwise interactions. Second, the jump procedure is easily parallelizable over all atoms, which make it suitable for parallel architectures such as GPUs. In particular, jumps on pairwise interactions scale better than certain many-body interactions such as the Ewald sum appearing the reciprocal electrostatic force (classically used to compute long range electrostatic interactions, see for instance [19]), and replacing a part of this sum by jumps on pairwise direct electrostatic forces yields a better speedup. Finally, by combining the jump procedure with classical multi-time-step methods, the intrinsically random nature of jumps allows to mitigate some well-known resonance issues of multi-time-stepping, which allows to choose a slightly larger time step. In [21], the choice $\rho = 0$ seems to provide a good compromise between speedup and preservation of dynamical properties, leading for instance to a total acceleration of 19.3 % in simulation speed compared to a state-of-art multi-time-step discretization of Langevin dynamics for a system of 96000 water molecules.

Numerical illustrations. Let us give a numerical illustration of the algorithm with various parameters in the case of two-dimensional Gaussian distributions. We take $d = 2$,

$$U(x, y) = \frac{x^2 + 11y^2}{2},$$

and consider the following splitting

$$U_0(x, y) = \frac{x^2 + y^2}{2}, \quad U_1(x, y) = 5y^2,$$

which means that the Langevin dynamics $\mathcal{L}_H + \mathcal{L}_D$ samples the standard Gaussian distribution, while the jumps $\mathcal{L}_J$ bring asymmetry to the process. We fix a time step ($\delta = 0.01$) and a time of simulation ($n = 10^4$ time steps) and show in Figure 2.1 how the jump parameter ρ and the friction γ influence the

trajectories. When the friction is very small (in Figure 2.1 (a) and (b)), the Langevin part resembles the Hamiltonian dynamics associated to U_0 . In that case, if ρ is close to 1, the process is close to the Hamiltonian dynamics sampling U . If $\rho = -1$, the jump process corresponds to the BPS, there are fewer jumps, but there are more visible since they tend to change the direction of the trajectory to bring asymmetry. When the friction is high (in Figure 2.1 (e) and (f)), as expected, both trajectories (with either $\rho = 1$ or ρ close to 1) resemble an overdamped Langevin diffusion process. The intermediate case (in Figure 2.1 (c) and (d)) correspond to a mix between those extreme situations, with noisy but ballistic dynamics.

For a thorough empirical study of the efficiency of the algorithm on MD applications, we refer to the previous [43] (restricted to the case $\rho = -1$) or the companion paper [21] in the general case.

3. Geometric ergodicity

This section is devoted to the proof of Theorem 2.1, which relies on hypocoercive computations *à la Villani* in Sobolev spaces. We only consider the case $\mathcal{X} = \mathbb{R}^d$, the proof being easily adapted and simpler in the compact periodic case.

3.1. Proof of Theorem 2.1

In this section, we give the key steps of the proof, postponing the proofs of the main intermediary results to the rest of Section 3.

Step 1. Sobolev hypocoercivity. Fix $k \in \mathbb{N}$ and define the Sobolev space of order k by

$$H^k(\mu) = \left\{ f : \mathbb{R}^{2d} \rightarrow \mathbb{R} \text{ measurable, } \|f\|_{H^k}^2 := \sum_{i+j \leq k} \int_{\mathbb{R}^{2d}} |\nabla_x^i \nabla_v^j f|^2 d\mu < \infty \right\},$$

where we denoted

$$|\nabla_x^i \nabla_v^j f|^2 = \sum_{|\alpha_1|=i; |\alpha_2|=j} |\partial_x^{\alpha_1} \partial_v^{\alpha_2} f|^2.$$

We will first establish the exponential convergence of P_t to μ in H^k . The velocity jump Langevin process is hypocoercive, in the sense that the generator of this process does not satisfy

$$\langle f, \mathcal{L}f \rangle_{H^k} \leq -\rho \|f\|_{H^k}^2, \quad \forall f \in H^k,$$

for any $\rho > 0$, but such that there exists a norm equivalent to $\|\cdot\|_{H^k}$ such that the previous inequality holds. To prove exponential convergence to equilibrium, following [25, 26, 52], for $f \in H^k(\mu)$, we introduce

$$N_k(f) = \int_{\mathbb{R}^{2d}} f^2 d\mu + \sum_{p=1}^k \int_{\mathbb{R}^{2d}} \left(\sum_{i=0}^{p-1} \omega_{i,p} |\nabla_x^i \nabla_v^{p-i} f|^2 + \omega_{p,p} |(\nabla_x^p - \nabla_x^{p-1} \nabla_v) f|^2 \right) d\mu, \quad (3.1)$$

where $(\omega_{i,p})_{p \in \mathbb{N}, 1 \leq i \leq p}$ are positive numbers, which is equivalent to $\|f\|_{H^k}^2$. We refer to N_k as the modified Sobolev norm (although strictly speaking it is the square of a norm). With computations similar to [47] (although rather presented in terms of Gamma calculus as in [37]), and proceeding by induction in k , our goal is to design $(\omega_{i,p})_{p \in \mathbb{N}, 1 \leq i \leq p}$ in such a way that, at least informally,

$$\partial_t N_k (P_t f - \mu(f)) \leq -\rho_k N_k (P_t f - \mu(f)),$$

for all $f \in H^k$ for some $\rho_k > 0$. To make rigorous this argument, instead of justifying the time differentiation, we will rely on Lumer–Phillips theorem [51, Chapter IX, p. 250], see in Section 3.5. This leads to the following:

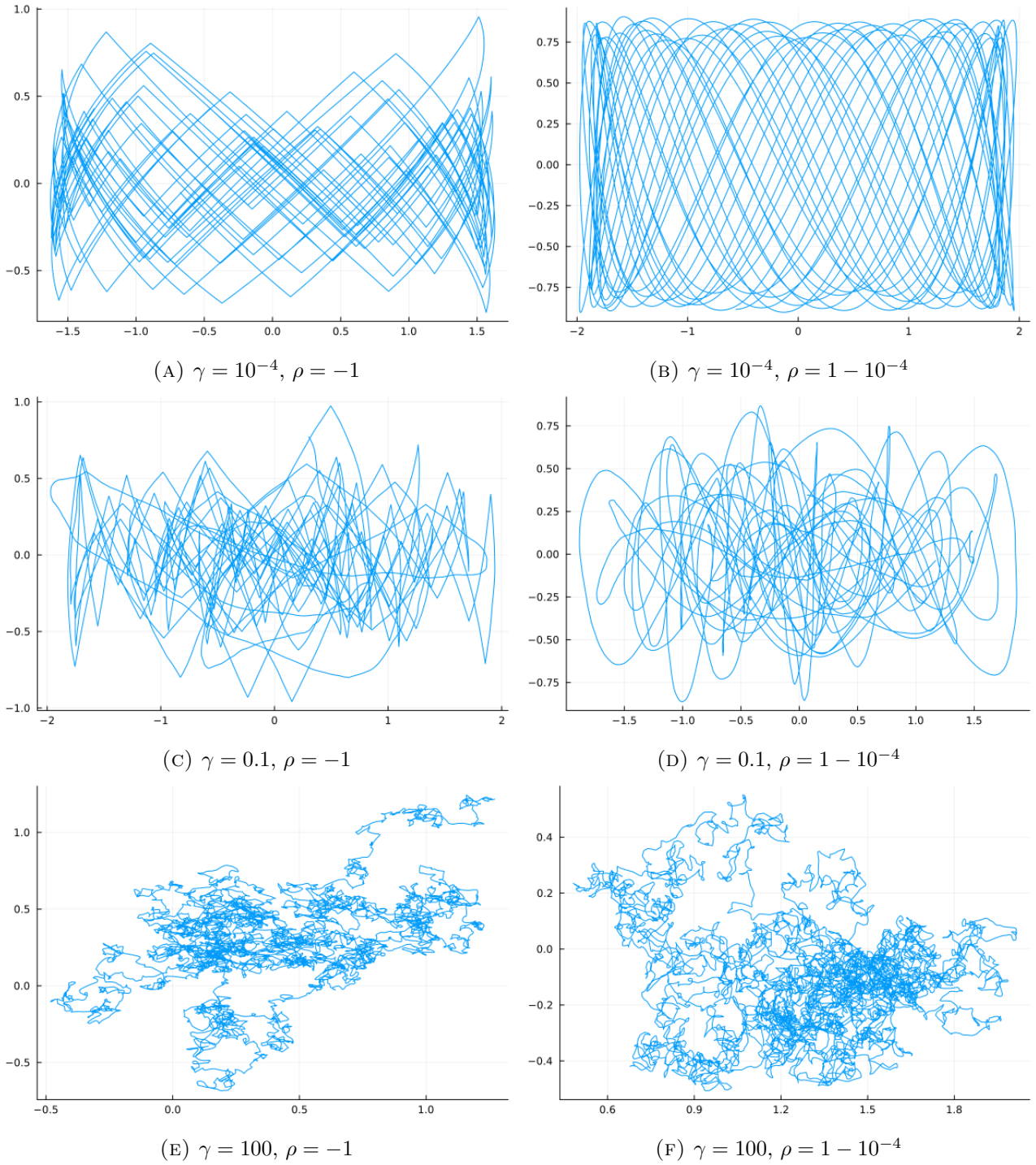


FIGURE 2.1. Trajectories of a splitting scheme for the velocity jump Langevin process

Theorem 3.1. *There exist positive $(\omega_{i,p})_{p \in \mathbb{N}, 1 \leq i \leq p}$ and $(\rho_k)_{k \in \mathbb{N}^*}$ such that for all $k \in \mathbb{N}^*$, $f \in H^k(\mu)$ and $t \geq 0$,*

$$N_k(P_t f - \mu(f)) \leq e^{-\rho_k t} N_k(f - \mu(f)).$$

This is proven below (from Section 3.2 to Section 3.5).

Thanks to Sobolev embeddings, this already gives a result similar to Theorem 2.1 except that in the right-hand-side of (2.7), $b = 1$. This would not be enough for establishing Theorem 2.2 since e^H is not integrable with respect to μ , from which it cannot be a Lyapunov function for P_t , and then neither for the BJAOAJB chain. Hence, we wouldn't be able to control the expectation of the discretization bias in the proof of Theorem 2.2.

Step 2. Lyapunov function and re-weighted Sobolev norms. To solve this issue, we will work with a particular weighted Sobolev norms. The weight $V_b : \mathbb{R}^{2d} \rightarrow \mathbb{R}$ is defined, for any $a, b > 0$, by

$$V_b(x, v) = \exp \left(b \left(U_0(x) + \frac{|v|^2}{2} - a \frac{x \cdot v}{\sqrt{1 + |x|^2}} \right) \right). \quad (3.2)$$

We show in Section 3.6 that it is a Lyapunov function for $\mathcal{L}^*$, the adjoint operator of $\mathcal{L}$ in $L^2(\mu)$ (which is computed in Proposition 3.5):

Proposition 3.2. *There exist $a > 0, b_0 \in (0, 1)$ such that for all $b \in (0, b_0)$, there exist $\eta > 0, C > 0$ such that*

$$\mathcal{L}^* V_b \leq -\eta V_b + C.$$

Combining Theorem 3.1 and Proposition 3.2, following computations as in the proof of Theorem 3.1 with a modified norm similar to (3.1) but with the additional weight V_b , we end up with the following:

Proposition 3.3. *For all $\alpha \in \mathbb{N}^{2d}$ and all $f \in \mathcal{A}$, there exist $b \in (0, 1), \eta > 0, C > 0$ such that for all $t \geq 0$,*

$$\int_{\mathbb{R}^{2d}} |\partial^\alpha (P_t f - \mu(f))|^2 V_b d\mu \leq C e^{-\eta t}.$$

The proof is postponed to Section 3.6.

Step 3. Conclusion. Using Step 1 and 2, we prove Theorem 2.1:

Proof of Theorem 2.1. Let $f \in \mathcal{A}$ and $\alpha \in \mathbb{N}^{2d}$. Without loss of generality we assume that $\mu(f) = 0$. Let $b' \in (0, 1)$ such that the result of Proposition 3.3 holds for $V_{b'}$, i.e.

$$\int_{\mathbb{R}^{2d}} |\partial^\nu P_t f|^2 V_{b'} d\mu \leq C e^{-\eta t},$$

for all $\nu \leq \alpha$, and let $b \in (0, 1)$ and $\varepsilon > 0$ such that $b' > 1 - (1 - \varepsilon)b$. We then have

$$\begin{aligned} \int_{\mathbb{R}^{2d}} |\partial^\alpha (P_t f e^{-bH})|^2 &\leq C \sum_{\nu \leq \alpha} \int_{\mathbb{R}^{2d}} |\partial^\nu P_t f|^2 |\partial^{\alpha-\nu} e^{-bH}|^2 \\ &\leq C \sum_{\nu \leq \alpha} \int_{\mathbb{R}^{2d}} |\partial^\nu P_t f|^2 e^{-(1-\varepsilon)bH} \\ &= C \sum_{\nu \leq \alpha} \int_{\mathbb{R}^{2d}} |\partial^\nu P_t f|^2 e^{(1-(1-\varepsilon)b)H} d\mu \\ &\leq C \sum_{\nu \leq \alpha} \int_{\mathbb{R}^{2d}} |\partial^\nu P_t f|^2 V_{b'} d\mu \\ &\leq C e^{-\rho t}, \end{aligned}$$

for some $\rho > 0$. We conclude with Sobolev embedding. ■

3.2. Preliminary: Gamma calculus

The proof of Theorem 3.1 relies on an induction argument and Gamma calculus, as presented in [37], which in our case is simply a convenient way to compute time derivatives of quantities of the form (3.1) along the semi-group. Fix some smooth $\phi : \mathcal{C}_c^\infty(\mathbb{R}^{2d}) \rightarrow \mathcal{C}_c^\infty(\mathbb{R}^{2d})$ and $L \in \{\mathcal{L}_H, \mathcal{L}_J, \mathcal{L}_D\}$. Define the generalized *carré du champ* for all $h \in \mathcal{C}_c^\infty(\mathbb{R}^{2d})$ by

$$\Gamma_{L,\phi}(h) = \frac{1}{2} (L(\phi(h)) - D_h\phi(h)Lh), \quad (3.3)$$

where $D_h\phi$ denotes the differential operator of ϕ . When $\phi(h) = h^2$ we retrieve the usual carré du champ, which is simply denoted by Γ_L . Formally, using that $\int Lh d\mu = 0$ for any suitable h by invariance of μ (which will be proven in Section 3.3), we have:

$$\partial_t N_k(P_t f - \mu(f)) = - \int_{\mathbb{R}^{2d}} \Gamma(P_t f) d\mu - \sum_{p=1}^k \int_{\mathbb{R}^{2d}} \left(\sum_{i=0}^{p-1} \omega_{i,p} \Gamma_{i,p}(P_t f) + \omega_{p,p} \Gamma_{p,p}(P_t f) \right) d\mu, \quad (3.4)$$

where, for $i < p$,

$$\Gamma_{i,p}(h) = \Gamma_{L, |\nabla_x^i \nabla_v^{p-i}|^2}(h) \quad \text{and} \quad \Gamma_{p,p}(h) = \Gamma_{L, |\nabla_x^p - \nabla_x^{p-1} \nabla_v|^2}(h).$$

If we could show that for all $h \in \mathcal{C}_c^\infty(\mathbb{R}^{2d})$

$$\int_{\mathbb{R}^{2d}} \Gamma(h) d\mu + \sum_{p=1}^k \int_{\mathbb{R}^{2d}} \left(\sum_{i=0}^{p-1} \omega_{i,p} \Gamma_{i,p}(h) + \omega_{p,p} \Gamma_{p,p}(h) \right) d\mu \geq \rho_k N_k(h - \mu(h)),$$

for some $\rho_k > 0$, then Theorem 3.1 would follow, at least informally, in the sense that we haven't justified the time derivatives in (3.4). We will actually prove a slightly stronger result (see Proposition 3.12 below) for the sake of the induction argument. Recall from [37] the proposition that allows for the computations of this generalized *carré du champ* (the proof is straightforward).

Proposition 3.4. *If there exists $A = (A_1, \dots, A_p) : \mathcal{C}^\infty \rightarrow (\mathcal{C}^\infty)^p$ a linear operator such that $\phi(h) = |Ah|^2$, then*

$$\Gamma_{L,\phi}(h) = \Gamma_L(Ah) + Ah \cdot [L, A]h,$$

where $\Gamma_L(Ah) = \sum_{i=1}^p \Gamma_L(A_i h)$ and $[L, A] = ([L, A_1], \dots, [L, A_p])$.

In the remainder of this work, we will denote Γ for Γ_L . For all i, j, i', j' such that $i + j = i' + j'$, the scalar product here has to be understood as

$$\nabla_x^i \nabla_v^j h \cdot \nabla_x^{i'} \nabla_v^{j'} h := \sum_{|\alpha_1|=i} \sum_{|\alpha_2|=j'} \sum_{|\nu|=i'-i} (\partial_x^{\alpha_1} \partial_v^{\alpha_2+\nu} h) (\partial_x^{\alpha_1+\nu} \partial_v^{\alpha_2} h).$$

In order to justify the discussion above, we will start by proving that the Gibbs measure is indeed invariant for the velocity jump Langevin process in Section 3.3. Section 3.4 will treat the case $k = 1$ which is the initialization of the induction argument. Section 3.5 will present the proof of convergence in H^k space, namely Theorem 3.1. Finally, Propositions 3.2 and 3.3 will be proved in Section 3.6.

3.3. Invariance of μ

The use of the velocity jump process for the sampling of μ , as well as the use of hypocoercivity method rely on the fact that μ is indeed an invariant measure of velocity jump Langevin process. Let us prove this property by computing the adjoint of $\mathcal{L}$ in $L^2(\mu)$, denoted as $\mathcal{L}^*$.

Proposition 3.5. *Under Assumption 1, for all $h \in \mathcal{C}_c^\infty(\mathbb{R}^{2d})$:*

$$\mathcal{L}^*h = -v \cdot \nabla_x h + \nabla U_0 \cdot \nabla_v h - \gamma v \cdot \nabla_v h + \gamma \Delta_v h + \frac{2}{1-\rho} \sum_{i=1}^d \mathbb{E} \left[(h(\cdot, V^i) - h) \Psi \left(\frac{\partial_i U_1}{2} (V_i^i - v_i) \right) \right].$$

Proof. Let $g, h \in \mathcal{C}_c^\infty(\mathbb{R}^{2d})$. By integrating by parts, we get

$$\begin{aligned} \int_{\mathbb{R}^{2d}} g \mathcal{L}_H h \, d\mu &= \mathcal{Z}_\mu^{-1} \int_{\mathbb{R}^{2d}} g(v \cdot \nabla_x h - \nabla U_0 \cdot \nabla_v h) e^{-U(x)} e^{-|v|^2/2} \, dx \, dv \\ &= -\mathcal{Z}_\mu^{-1} \int_{\mathbb{R}^{2d}} h v \cdot \nabla_x (g e^{-U(x)}) e^{-|v|^2/2} \, dx \, dv \\ &\quad + \mathcal{Z}_\mu^{-1} \int_{\mathbb{R}^{2d}} h \nabla_v \cdot (g e^{-|v|^2/2} \nabla U_0(x)) e^{-U(x)} \, dx \, dv \\ &= - \int_{\mathbb{R}^{2d}} h v \cdot \nabla_x g \, d\mu + \int_{\mathbb{R}^{2d}} g h v \cdot \nabla U(x) \, d\mu \\ &\quad + \int_{\mathbb{R}^{2d}} h \nabla U_0(x) \cdot \nabla_v g \, d\mu - \int_{\mathbb{R}^{2d}} g h v \cdot \nabla U_0(x) \, d\mu \\ &= \int_{\mathbb{R}^{2d}} h(-v \cdot \nabla_x g + \nabla U_0 \cdot \nabla_v g) \, d\mu + \int_{\mathbb{R}^{2d}} g h v \cdot \nabla U_1 \, d\mu. \end{aligned}$$

We also have

$$\begin{aligned} \int_{\mathbb{R}^{2d}} g \mathcal{L}_D h \, d\mu &= \mathcal{Z}_\mu^{-1} \int_{\mathbb{R}^{2d}} g(-\gamma v \cdot \nabla_v h + \gamma \Delta_v h) e^{-U(x)} e^{-|v|^2/2} \, dx \, dv \\ &= \gamma \mathcal{Z}_\mu^{-1} \left(\int_{\mathbb{R}^{2d}} g \nabla_v (e^{-|v|^2/2}) \cdot \nabla_v h e^{-U(x)} \, dx \, dv + \int_{\mathbb{R}^{2d}} h \Delta_v (g e^{-|v|^2/2}) e^{-U(x)} \, dx \, dv \right) \\ &= \gamma \mathcal{Z}_\mu^{-1} \left(\int_{\mathbb{R}^{2d}} \nabla_v (g e^{-|v|^2/2}) \cdot \nabla_v h e^{-U(x)} \, dx \, dv - \int_{\mathbb{R}^{2d}} \nabla_v g \cdot \nabla_v h e^{-U(x)} e^{-|v|^2/2} \, dx \, dv \right) \\ &\quad + \gamma \mathcal{Z}_\mu^{-1} \int_{\mathbb{R}^{2d}} h \Delta_v (g e^{-|v|^2/2}) e^{-U(x)} \, dx \, dv \\ &= -\gamma \int_{\mathbb{R}^{2d}} \nabla_v g \cdot \nabla_v h \, d\mu = \int_{\mathbb{R}^{2d}} h \mathcal{L}_D g \, d\mu. \end{aligned}$$

For the jump part, denote $k_0(x, v, \cdot)$ the law of the random variable V^i . Then k_0 is reversible with respect to the standard Gaussian measure, in the sense that for all $x \in \mathbb{R}^d$,

$$k_0(x, v, dv') e^{-|v|^2/2} \, dv = k_0(x, v', dv) e^{-|v'|^2/2} \, dv'. \quad (3.5)$$

Indeed, since $V_j^i = v_j$ for $j \neq i$ and

$$V_i^i = \rho v_i + \sqrt{1-\rho^2} \xi,$$

the coordinates of V^i are independent and each transition $v_j \mapsto V_j^i$ is reversible for the standard one-dimensional Gaussian. Using this reversibility, and the fact that $\Psi(s) - \Psi(-s) = s$, we get that

$$\begin{aligned} \int_{\mathbb{R}^{2d}} g \mathcal{L}_J h \, d\mu &= \frac{2}{1-\rho} \sum_{i=1}^d \int_{\mathbb{R}^{2d}} g \mathbb{E} \left[(h(x, V^i) - h(x, v)) \Psi \left(\frac{\partial_i U_1(x)}{2} (v_i - V_i^i) \right) \right] \, d\mu \\ &= \frac{2}{1-\rho} \sum_{i=1}^d \int_{\mathbb{R}^{2d}} h \mathbb{E} \left[(g(x, V^i) - g) \Psi \left(-\frac{\partial_i U_1(x)}{2} (v_i - V_i^i) \right) \right] \\ &\quad + \frac{2}{1-\rho} \sum_{i=1}^d \int_{\mathbb{R}^{2d}} g h \mathbb{E} \left[\Psi \left(-\frac{\partial_i U_1(x)}{2} (v_i - V_i^i) \right) - \Psi \left(\frac{\partial_i U_1(x)}{2} (v_i - V_i^i) \right) \right] \, d\mu \end{aligned}$$

$$\begin{aligned}
 &= \frac{2}{1-\rho} \sum_{i=1}^d \int_{\mathbb{R}^{2d}} h \mathbb{E} \left[(g(x, V^i) - g) \Psi \left(-\frac{\partial_i U_1}{2} (v_i - V_i^i) \right) \right] \\
 &\quad - \frac{1}{1-\rho} \sum_{i=1}^d \int_{\mathbb{R}^{2d}} g h \partial_i U_1 \mathbb{E} \left[(v_i - V_i^i) \right] d\mu \\
 &= \frac{2}{1-\rho} \sum_{i=1}^d \int_{\mathbb{R}^{2d}} h \mathbb{E} \left[(g(x, V^i) - g) \Psi \left(-\frac{\partial_i U_1}{2} (v_i - V_i^i) \right) \right] - \int_{\mathbb{R}^{2d}} g h v \cdot \nabla U_1 d\mu.
 \end{aligned}$$

Therefore, by summing the three terms, we get the result. $\blacksquare$

Corollary 3.6. *Under Assumption 1, for all $h \in \mathcal{C}_c^\infty(\mathbb{R}^{2d})$,*

$$\int_{\mathbb{R}^{2d}} \mathcal{L}h d\mu = 0.$$

Proof. This is a direct consequence of the fact that $\mathcal{L}^*1 = 0$ (although the constant function 1 is not compactly supported, the computations of Proposition 3.5 clearly works when $g = 1$). $\blacksquare$

3.4. Exponential decay in H^1

In this section, we prove convergence of the process in $H^1(\mu)$, which is the initialization step of the induction argument to prove Theorem 3.1. The modified Sobolev norm here reads

$$N_1(h) = \int_{\mathbb{R}^{2d}} h^2 d\mu + \int_{\mathbb{R}^{2d}} \omega_{0,1} |\nabla_v h|^2 + \omega_{1,1} |(\nabla_x - \nabla_v)h|^2 d\mu,$$

and Equation (3.4) can be written:

$$\partial_t N_1(P_t f - \mu(f)) = - \int_{\mathbb{R}^{2d}} \Gamma(P_t f) d\mu - \int_{\mathbb{R}^{2d}} (\omega_{0,1} \Gamma_{0,1}(P_t f) + \omega_{1,1} \Gamma_{1,1}(P_t f)) d\mu.$$

The goal of this section is to prove the following which, in view of (3.4), implies an exponential decay in H^1 as it bounds the norm N_1 in terms of its dissipation along the semi-group.

Lemma 3.7. *Under Assumption 1, there exist $\omega_{0,1}, \omega_{1,1}, \rho_1$ such that for all $h \in \mathcal{C}_c^\infty(\mathbb{R}^{2d})$ satisfying $\mu(h) = 0$:*

$$\int_{\mathbb{R}^{2d}} \Gamma(h) d\mu + \int_{\mathbb{R}^{2d}} (\omega_{0,1} \Gamma_{0,1}(h) + \omega_{1,1} \Gamma_{1,1}(h)) d\mu \geq \rho_1 \left(N_1(h) + \int_{\mathbb{R}^{2d}} |\nabla_x \nabla_v h|^2 + |\nabla_v^2 h|^2 d\mu \right).$$

The terms of order 2 in the right hand side are present to compensate future terms in the induction, but are not required for the convergence in $H^1(\mu)$. This convergence relies on the so-called Poincaré inequality.

Proposition 3.8. *Under Assumption 1, the Gibbs measure (1.1) satisfies a Poincaré inequality: there exists $C_P > 0$ such that*

$$\int_{\mathbb{R}^{2d}} (h - \mu(h))^2 d\mu \leq C_P \int_{\mathbb{R}^{2d}} |\nabla h|^2 d\mu. \quad (3.6)$$

for all $h \in \mathcal{C}_c^\infty(\mathbb{R}^{2d})$.

Proof. Proof of such inequality can be found in [2]. $\blacksquare$

In order to apply Proposition 3.4, we need to express the commutators. For the Hamiltonian part we get:

$$[\mathcal{L}_H, \nabla_x] = \nabla^2 U_0 \nabla_v, \quad [\mathcal{L}_H, \nabla_v] = -\nabla_x.$$

For the diffusion part we get:

$$[\mathcal{L}_D, \nabla_x] = 0, \quad [\mathcal{L}_D, \nabla_v] = \gamma \nabla_v.$$

Regarding the jump part, we have the following.

Lemma 3.9. *Under Assumption 1, there exists $C > 0$ such that for all $h \in \mathcal{C}_c^\infty(\mathbb{R}^{2d})$:*

$$\int_{\mathbb{R}^{2d}} |[\nabla_x, \mathcal{L}_J]h|^2 d\mu + \int_{\mathbb{R}^{2d}} |[\nabla_v, \mathcal{L}_J]h|^2 d\mu \leq C \left(\int_{\mathbb{R}^{2d}} |\nabla_v h|^2 d\mu + \int_{\mathbb{R}^{2d}} |\nabla_v^2 h|^2 d\mu \right).$$

To prove Lemma 3.9, we need two intermediate results:

Lemma 3.10. *For any function $h \in \mathcal{C}_c^\infty(\mathbb{R}^{2d})$ and any $x \in \mathbb{R}^d$,*

$$\int_{\mathbb{R}^d} |v|^2 h^2(x, v) e^{-|v|^2/2} dv \leq 2d \int_{\mathbb{R}^d} h^2(x, v) e^{-|v|^2/2} dv + 4 \int_{\mathbb{R}^d} |\nabla_v h(x, v)|^2 e^{-|v|^2/2} dv.$$

Proof. This is a particular case of the Lemma A.24 of [50]. Notice that $ve^{-|v|^2/2} = -\nabla e^{-|v|^2/2}$. Hence, an integration by parts and Young's inequality yield, for any x ,

$$\begin{aligned} \int_{\mathbb{R}^d} |v|^2 h^2(x, v) e^{-|v|^2/2} dv &= - \int_{\mathbb{R}^d} h^2(x, v) v \cdot \nabla e^{-|v|^2/2} dv \\ &= \int_{\mathbb{R}^d} \nabla \cdot (h^2(x, v) v) e^{-|v|^2/2} dv \\ &= d \int_{\mathbb{R}^d} h^2(x, v) e^{-|v|^2/2} dv + \int_{\mathbb{R}^d} 2h(x, v) \nabla h(x, v) \cdot v e^{-|v|^2/2} dv \\ &\leq d \int_{\mathbb{R}^d} h^2(x, v) e^{-|v|^2/2} dv + \int_{\mathbb{R}^d} 2|\nabla h(x, v)|^2 e^{-|v|^2/2} dv \\ &\quad + \frac{1}{2} \int_{\mathbb{R}^d} |v|^2 h^2(x, v) e^{-|v|^2/2} dv, \end{aligned}$$

and thus the result. ■

Proposition 3.11 (Poincaré for Gaussians). *For any function $h \in \mathcal{C}_c^\infty(\mathbb{R}^d)$,*

$$\int_{\mathbb{R}^d} \left(h(x) - \frac{1}{(2\pi)^{d/2}} \int_{\mathbb{R}^d} h(y) e^{-|y|^2/2} dy \right)^2 e^{-|x|^2/2} dx \leq \int_{\mathbb{R}^d} |\nabla h(x)|^2 e^{-|x|^2/2} dx. \quad (3.7)$$

Proof. Proof of such inequality can be found in [2]. ■

Proof of Lemma 3.9. We have

$$[\nabla_x, \mathcal{L}_J]h = \sum_{i=1}^d [\nabla_x, \mathcal{L}_J^i]h = \frac{1}{1-\rho} \sum_{i=1}^d \mathbb{E} \left[(v_i - V_i^i)(h(x, V^i) - h) \Psi' \left(\frac{\partial U_1}{2}(v_i - V_i^i) \right) \right] \nabla \partial_i U_1,$$

and

$$[\nabla_v, \mathcal{L}_J]h = \sum_{i=1}^d \mathbb{E} \left[\partial_i U_1 (h(x, V^i) - h) \Psi' \left(\frac{\partial U_1}{2}(v_i - V_i^i) \right) - 2\partial_{v_i} h(x, V^i) \Psi \left(\frac{\partial U_1}{2}(v_i - V_i^i) \right) \right] e_i.$$

Using the fact that Ψ' , ∇U_1 and $\nabla^2 U_1$ are bounded, there exists a constant $C > 0$ such that

$$|[\nabla_x, \mathcal{L}_J]h|^2 \leq C \sum_{i=1}^d \mathbb{E} [(v_i - V_i^i)^2 (h(x, V^i) - h)^2],$$

and

$$|[\nabla_v, \mathcal{L}_J]h|^2 \leq C \sum_{i=1}^d \mathbb{E}[(h(x, V^i) - h)^2 + (\partial_{v_i} h(x, V^i))^2 (1 + (v_i - V_i^i)^2)].$$

Recall from (3.5) that the law of the random variable V^i is reversible with respect to the standard Gaussian measure. Therefore,

$$\begin{aligned} \int_{\mathbb{R}^{2d}} \mathbb{E} \left[(1 + (v_i - V_i^i)^2) (\partial_{v_i} h(x, V^i))^2 \right] d\mu &= \int_{\mathbb{R}^{2d}} (\partial_{v_i} h)^2 \left(1 + \mathbb{E} \left[(v_i - V_i^i)^2 \right] \right) d\mu \\ &\leq \int_{\mathbb{R}^{2d}} (\partial_{v_i} h)^2 (2 + v_i^2) d\mu \\ &\leq 4 \int_{\mathbb{R}^{2d}} ((\partial_{v_i} h)^2 + (\partial_{v_i}^2 h)^2) d\mu, \end{aligned}$$

where we used Lemma 3.10 with $d = 1$ in the last inequality. By denoting Π_{v_i} the projection

$$\Pi_{v_i} h(x, v) = \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} h(x, v_1, \dots, v_{i-1}, w_i, v_{i+1}, \dots, v_d) e^{-w_i^2/2} dw_i,$$

we also have

$$\begin{aligned} \int_{\mathbb{R}^{2d}} \mathbb{E} \left[(1 + (v_i - V_i^i)^2) (h(x, V^i) - h)^2 \right] d\mu &\leq 2 \int_{\mathbb{R}^{2d}} \mathbb{E} \left[(1 + (v_i - V_i^i)^2) ((h(x, V^i) - \Pi_{v_i} h)^2 + (h - \Pi_{v_i} h)^2) \right] d\mu \\ &\leq 4 \int_{\mathbb{R}^{2d}} \left(1 + \mathbb{E} \left[(v_i - V_i^i)^2 \right] \right) (h - \Pi_{v_i} h)^2 d\mu \\ &\leq 4 \int_{\mathbb{R}^{2d}} (2 + v_i^2) (h - \Pi_{v_i} h)^2 d\mu \\ &\leq 16 \int_{\mathbb{R}^{2d}} (h - \Pi_{v_i} h)^2 d\mu + 16 \int_{\mathbb{R}^{2d}} (\partial_{v_i} h)^2 d\mu, \end{aligned}$$

where we used again Lemma 3.10 with $d = 1$ in the last inequality. The Poincaré inequality (3.7) can be written

$$\int_{\mathbb{R}^{2d}} (h - \Pi_{v_i} h)^2 d\mu \leq \int_{\mathbb{R}^{2d}} |\partial_{v_i} h|^2 d\mu,$$

and thus

$$\int_{\mathbb{R}^{2d}} \mathbb{E} \left[(1 + (v_i - V_i^i)^2) (h(x, V^i) - h)^2 \right] d\mu \leq 32 \int_{\mathbb{R}^{2d}} (\partial_{v_i} h)^2 d\mu.$$

Therefore, there exists $C > 0$ such that

$$\int_{\mathbb{R}^{2d}} |[\nabla, \mathcal{L}_J]h|^2 d\mu \leq C \sum_{i=1}^d \int_{\mathbb{R}^{2d}} ((\partial_{v_i} h)^2 + |\nabla_v \partial_{v_i} h|^2) d\mu = C \left(\int_{\mathbb{R}^{2d}} |\nabla_v h|^2 d\mu + \int_{\mathbb{R}^{2d}} |\nabla_v^2 h|^2 d\mu \right),$$

which concludes the proof. $\blacksquare$

We now have everything to prove Lemma 3.7.

Proof of Lemma 3.7. Fix $\omega_{0,1} > \omega_{1,1} > 0$ that we are going to choose later. Let's first treat the derivative of the L^2 norm. For all $h \in C_c^\infty(\mathbb{R}^{2d})$, $(x, v) \in \mathbb{R}^{2d}$, we have

$$\Gamma(h)(x, v) = 2\gamma |\nabla_v h(x, v)|^2 + \sum_{i=1}^d \lambda_i(x, v) \int_{\mathbb{R}^d} (h(x, v') - h(x, v))^2 q_i(x, v, dv'). \quad (3.8)$$

This already implies that

$$\Gamma(h) \geq 2\gamma |\nabla_v h|^2.$$

To get a lower bound on $\Gamma_{0,1}(h)$ and $\Gamma_{1,1}(h)$, we use Proposition 3.4. Equation (3.8) yields:

$$\Gamma(\nabla_v h) \geq 2\gamma|\nabla_v^2 h|^2, \quad \Gamma((\nabla_x - \nabla_v)h) \geq 2\gamma|(\nabla_x \nabla_v - \nabla_v^2)h|^2,$$

so that using $\omega_{0,1} > \omega_{1,1}$ and Young inequality:

$$\omega_{0,1}\Gamma(\nabla_v h) + \omega_{1,1}\Gamma((\nabla_x - \nabla_v)h) \geq 2\gamma(\omega_{0,1} - \omega_{1,1})|\nabla_v^2 h|^2 + \omega_{1,1} \left(\gamma|\nabla_v^2 h|^2 + 2\gamma/3|\nabla_x \nabla_v h|^2 \right).$$

Let us now look at the commutators. Fix $\varepsilon > 0$ and write using Young inequality and Lemma 3.9:

$$\begin{aligned} \int_{\mathbb{R}^{2d}} \nabla_v h \cdot [\mathcal{L}, \nabla_v] h \, d\mu &= \int_{\mathbb{R}^{2d}} \nabla_v h \cdot (-\nabla_x h + \gamma \nabla_v + [\mathcal{L}_J, \nabla_v] h) \, d\mu \\ &\geq - \int_{\mathbb{R}^{2d}} \left(\frac{1}{\varepsilon} + 1 \right) |\nabla_v h|^2 + \frac{\varepsilon}{2} |(\nabla_x - \nabla_v)h|^2 + \frac{\varepsilon}{2} C \left(|\nabla_v h|^2 + |\nabla_v^2 h|^2 \right) \, d\mu, \end{aligned}$$

for some $C > 0$. Using again Young inequality and Lemma 3.9, we get that there exists $C > 0$ such that:

$$\begin{aligned} \int_{\mathbb{R}^{2d}} (\nabla_x - \nabla_v)h \cdot [\mathcal{L}, \nabla_x - \nabla_v] h \, d\mu &= \int_{\mathbb{R}^{2d}} (\nabla_x - \nabla_v)h \cdot \left(\nabla^2 U_0 \nabla_v h + \nabla_x h - \gamma \nabla_v h + [\mathcal{L}_J, \nabla_x - \nabla_v] h \right) \, d\mu \\ &\geq \int_{\mathbb{R}^{2d}} \frac{1}{2} |(\nabla_x - \nabla_v)h|^2 - C \left(|\nabla_v h|^2 + |\nabla_v^2 h|^2 \right) \, d\mu. \end{aligned}$$

Now all is left to do is to choose ε , $\omega_{0,1}$ and $\omega_{1,1}$. We choose them such that

$$\frac{\omega_{1,1}}{\omega_{0,1}} \leq \frac{2\gamma - C\varepsilon}{2\gamma + C}, \quad \varepsilon < \min \left(\frac{\omega_{1,1}}{2\omega_{0,1}}, \frac{2\gamma}{C} \right), \quad C\omega_{1,1} + \omega_{0,1} \left(\frac{1}{\varepsilon} + 1 + \frac{C\varepsilon}{2} \right) \leq \gamma$$

so that

$$\begin{aligned} \int_{\mathbb{R}^{2d}} \Gamma(h) \, d\mu + \int_{\mathbb{R}^{2d}} (\omega_{0,1}\Gamma_{0,1}(h) + \omega_{1,1}\Gamma_{1,1}(h)) \, d\mu &\geq \int_{\mathbb{R}^{2d}} \gamma|\nabla_v h|^2 + \frac{\omega_{1,1}}{4}|(\nabla_x - \nabla_v)h|^2 + \omega_{1,1}\gamma \left(|\nabla_v^2 h|^2 + \frac{2}{3}|\nabla_x \nabla_v h|^2 \right) \, d\mu \\ &\geq \tilde{\rho}_1 \int_{\mathbb{R}^{2d}} \omega_{0,1}|\nabla_v h|^2 + \omega_{1,1}|(\nabla_x - \nabla_v)h|^2 + |\nabla_x \nabla_v h|^2 + |\nabla_v^2 h|^2 \, d\mu, \end{aligned}$$

for some $\tilde{\rho}_1 > 0$. Since we consider h such that $\mu(h) = 0$, the Poincaré inequality (3.6) yields that there exists $\rho_1 > 0$ such that

$$\begin{aligned} \tilde{\rho}_1 \int_{\mathbb{R}^{2d}} \omega_{0,1}|\nabla_v h|^2 + \omega_{1,1}|(\nabla_x - \nabla_v)h|^2 + |\nabla_x \nabla_v h|^2 + |\nabla_v^2 h|^2 \, d\mu &\geq \rho_1 \left(\mathcal{N}_1(h) + \int_{\mathbb{R}^{2d}} |\nabla_x \nabla_v h|^2 + |\nabla_v^2 h|^2 \, d\mu \right), \end{aligned}$$

which concludes the proof. $\blacksquare$

3.5. Exponential decay in H^k

Recall that we defined the modified Sobolev norm as

$$N_k(h) := \int_{\mathbb{R}^{2d}} h^2 \, d\mu + \sum_{p=1}^k \int_{\mathbb{R}^{2d}} \left(\sum_{i=0}^{p-1} \omega_{i,p} |\nabla_x^i \nabla_v^{p-i} h|^2 + \omega_{p,p} |(\nabla_x^p - \nabla_x^{p-1} \nabla_v)h|^2 \right) \, d\mu.$$

In order to control its derivative (3.4), we prove by induction on k the following lemma, generalizing Lemma 3.7.

Proposition 3.12. *Under Assumption 1, there exist $(\omega_{i,p})_{1 \leq i \leq p}$ such that for all $k \in \mathbb{N}$, there exists $\rho_k > 0$ such that for all $h \in C_c^\infty(\mathbb{R}^{2d}, \mathbb{R})$ with $\mu(h) = 0$:*

$$\begin{aligned} \int_{\mathbb{R}^{2d}} \Gamma(h) d\mu + \sum_{p=1}^k \int_{\mathbb{R}^{2d}} \left(\sum_{i=0}^{p-1} \omega_{i,p} \Gamma_{i,p}(h) + \omega_{p,p} \Gamma_{p,p}(h) \right) d\mu \\ \geq \rho_k \left(N_k(h) + \int_{\mathbb{R}^{2d}} \sum_{i=0}^k |\nabla_x^i \nabla_v^{k+1-i} h|^2 d\mu \right). \end{aligned} \quad (3.9)$$

To prove this lemma, we will need formulas for partial derivatives of products, provided by the Leibniz formula

Proposition 3.13. *For all $g, h \in C_c^\infty$ and $\alpha_1, \alpha_2 \in \mathbb{N}^d$,*

$$\partial_x^{\alpha_1} \partial_v^{\alpha_2} (gh) = \sum_{\nu_1 \leq \alpha_1, \nu_2 \leq \alpha_2} \binom{\alpha_1}{\nu_1} \binom{\alpha_2}{\nu_2} (\partial_x^{\nu_1} \partial_v^{\nu_2} g) (\partial_x^{\alpha_1 - \nu_1} \partial_v^{\alpha_2 - \nu_2} h)$$

where for $\alpha = (\alpha_1, \dots, \alpha_d)$ and $\nu = (\nu_1, \dots, \nu_d)$,

$$\binom{\alpha}{\nu} = \prod_{i=1}^d \binom{\alpha_i}{\nu_i},$$

and

$$\nu \leq \alpha \iff \forall i \in \llbracket 1, d \rrbracket, \nu_i \leq \alpha_i.$$

For any $\alpha \in \mathbb{N}^d$ we will also denote $|\alpha| = \sum_{i=1}^d \alpha_i$. Let us look at the commutators. As in the H^1 case, we will look separately the parts associated to the Hamiltonian dynamic $\mathcal{L}_H$, the diffusion $\mathcal{L}_D$ and the jump process $\mathcal{L}_J$.

Lemma 3.14 (Hamiltonian commutators). *Let $i \in \llbracket 0, k \rrbracket$. There exists $C > 0$ such that the following holds. If $i < k$, we have*

$$\nabla_x^i \nabla_v^{k+1-i} h \cdot [\mathcal{L}_H, \nabla_x^i \nabla_v^{k+1-i}] h \geq -|\nabla_x^i \nabla_v^{k+1-i} h|^2 - C \left(|\nabla_x^{i+1} \nabla_v^{k-i} h|^2 + \sum_{j=0}^{i-1} |\nabla_x^j \nabla_v^{k+2-i-j} h|^2 \right).$$

In the case $i = k$, for all $\eta > 0$, we have

$$\nabla_x^k \nabla_v h \cdot [\mathcal{L}_H, \nabla_x^k \nabla_v] h \geq -\frac{\eta}{8} |(\nabla_x^{k+1} - \nabla_x^k \nabla_v) h|^2 - C \left(1 + \frac{1}{\eta} \right) |\nabla_x^k \nabla_v h|^2 - C \sum_{j=0}^{k-1} |\nabla_x^j \nabla_v^2 h|^2.$$

Finally, we have

$$(\nabla_x^{k+1} - \nabla_x^k \nabla_v) h \cdot [\mathcal{L}_H, \nabla_x^{k+1} - \nabla_x^k \nabla_v] h \geq \frac{1}{2} |(\nabla_x^{k+1} - \nabla_x^k \nabla_v) h|^2 - C \left(\sum_{j=0}^k |\nabla_x^j \nabla_v h|^2 + \sum_{j=0}^{k-1} |\nabla_x^j \nabla_v^2 h|^2 \right).$$

Proof. For all $\alpha_1, \alpha_2 \in \mathbb{N}^d$ with $|\alpha_1| = i$ and $|\alpha_2| = k+1-i$, where $\alpha_2 = (\alpha_{2,1}, \dots, \alpha_{2,d})$, write

$$\alpha_2^j = (\alpha_{2,1}, \dots, \alpha_{2,j-1}, \alpha_{2,j} - 1, \alpha_{2,j+1}, \dots, \alpha_{2,d}),$$

and we have

$$\begin{aligned}
 \partial_x^{\alpha_1} \partial_v^{\alpha_2} \mathcal{L}_H h &= \sum_{j=1}^d \partial_x^{\alpha_1} \partial_v^{\alpha_2} \left(v_j \partial_{x_j} h - \partial_j U_0 \partial_{v_j} h \right) \\
 &= \sum_{j=1}^d \partial_v^{\alpha_2} (v_j \partial_x^{\alpha_1} \partial_{x_j} h) - \partial_x^{\alpha_1} (\partial_j U_0 \partial_v^{\alpha_2} \partial_{v_j} h) \\
 &= \sum_{j=1}^d \left(\sum_{\nu_2 \leq \alpha_2} \binom{\alpha_2}{\nu_2} \partial_v^{\nu_2} (v_j) \partial_v^{\alpha_2 - \nu_2} \partial_x^{\alpha_1} \partial_{x_j} h - \sum_{\nu_1 \leq \alpha_1} \binom{\alpha_1}{\nu_1} \partial_x^{\nu_1} \partial_j U_0 \partial_x^{\alpha_1 - \nu_1} \partial_v^{\alpha_2} \partial_{v_j} h \right) \\
 &= \sum_{j=1}^d \left(\alpha_{2,j} \partial_v^{\alpha_2} \partial_x^{\alpha_1} \partial_{x_j} h + v_j \partial_v^{\alpha_2} \partial_x^{\alpha_1} \partial_{x_j} h - \sum_{\nu_1 \leq \alpha_1} \binom{\alpha_1}{\nu_1} \partial_x^{\nu_1} \partial_j U_0 \partial_x^{\alpha_1 - \nu_1} \partial_v^{\alpha_2} \partial_{v_j} h \right).
 \end{aligned}$$

Therefore,

$$[\mathcal{L}_H, \partial_x^{\alpha_1} \partial_v^{\alpha_2}] h = \sum_{j=1}^d \left(\sum_{\nu_1 \leq \alpha_1; \nu_1 \neq 0} \binom{\alpha_1}{\nu_1} \partial_x^{\nu_1} \partial_j U_0 \partial_x^{\alpha_1 - \nu_1} \partial_v^{\alpha_2} \partial_{v_j} h - \alpha_{2,j} \partial_v^{\alpha_2} \partial_x^{\alpha_1} \partial_{x_j} h \right).$$

Using the bound on the derivatives of U_0 , this yields that there exists a $C > 0$ such that

$$\nabla_x^i \nabla_v^{k+1-i} h \cdot [\mathcal{L}_H \nabla_x^i \nabla_v^{k+1-i}] h \geq -|\nabla_x^i \nabla_v^{k+1-i} h|^2 - C \left(|\nabla_x^{i+1} \nabla_v^{k-i} h|^2 + \sum_{j=0}^{i-1} |\nabla_x^j \nabla_v^{k+2-i} h|^2 \right).$$

If $i = k$, then for any $\eta > 0$, using the Young inequality

$$\begin{aligned}
 \sum_{\substack{|\alpha_1|=k; |\alpha_2|=1 \\ j \in \llbracket 1, d \rrbracket}} (\partial_x^{\alpha_1} \partial_v^{\alpha_2} h) (\alpha_{2,j} \partial_v^{\alpha_2} \partial_x^{\alpha_1} \partial_{x_j} h) &= \sum_{\substack{|\alpha_1|=k \\ j \in \llbracket 1, d \rrbracket}} (\partial_x^{\alpha_1} \partial_{v_j} h) (\partial_x^{\alpha_1} \partial_{x_j} h) \\
 &= \sum_{\substack{|\alpha_1|=k \\ j \in \llbracket 1, d \rrbracket}} (\partial_x^{\alpha_1} \partial_{v_j} h) (\partial_x^{\alpha_1} \partial_{x_j} - \partial_x^{\alpha_1} \partial_{v_j} + \partial_x^{\alpha_1} \partial_{v_j}) h \\
 &\leq \left(1 + \frac{\eta}{8} \right) |\nabla_x^k \nabla_v h|^2 + \frac{\eta}{8} |(\nabla_x^{k+1} - \nabla_x^k \nabla_v) h|^2.
 \end{aligned}$$

which again implies that

$$\nabla_x^k \nabla_v h \cdot [\mathcal{L}_H, \nabla_x^k \nabla_v] h \geq -C \left(1 + \frac{1}{\eta} \right) |\nabla_x^k \nabla_v h|^2 - \frac{\eta}{8} |(\nabla_x^{k+1} - \nabla_x^k \nabla_v) h|^2 - C \sum_{j=0}^{k-1} |\nabla_x^j \nabla_v^2 h|^2.$$

Similarly, for β a multi-index such that $|\beta| = k$ and $i \in \llbracket 1, d \rrbracket$,

$$\begin{aligned}
 [\mathcal{L}_H, \partial_x^\beta \partial_{x_i} - \partial_x^\beta \partial_{v_i}] h &= \partial_x^\beta \partial_{x_i} h + \sum_{j=1}^d \sum_{\nu \leq \beta} \binom{\beta}{\nu} \partial^\nu \partial_{i,j} U_0 \partial_x^{\beta-\nu} \partial_{v_j} h \\
 &\quad + \sum_{j=1}^d \sum_{\substack{\nu \leq \beta \\ \nu \neq 0}} \binom{\beta}{\nu} \partial^\nu \partial_j U_0 \partial_x^{\beta-\nu} \partial_{x_i} \partial_{v_j} h - \sum_{j=1}^d \sum_{\substack{\nu \leq \beta \\ \nu \neq 0}} \binom{\beta}{\nu} (\partial^\nu \partial_j U_0) (\partial_x^{\beta-\nu} \partial_{v_i} \partial_{v_j} h),
 \end{aligned}$$

and thus, by applying Young inequality,

$$\begin{aligned}
 (\nabla_x^{k+1} - \nabla_x^k \nabla_v) h \cdot [\mathcal{L}_H, \nabla_x^{k+1} - \nabla_x^k \nabla_v] h \\
 \geq \frac{1}{2} |(\nabla_x^{k+1} - \nabla_x^k \nabla_v) h|^2 - C \left(\sum_{j=0}^k |\nabla_x^j \nabla_v h|^2 + \sum_{j=0}^{k-1} |\nabla_x^j \nabla_v^2 h|^2 \right),
 \end{aligned}$$

which concludes the proof. $\blacksquare$

Regarding the diffusion part, we have

Lemma 3.15. *For $i \in \llbracket 0, k \rrbracket$,*

$$\nabla_x^i \nabla_v^{k+1-i} h \cdot [\mathcal{L}_D, \nabla_x^i \nabla_v^{k+1-i}] h = \gamma(k+1-i) |\nabla_x^i \nabla_v^{k+1-i} h|^2,$$

and there exists $C > 0$ such that

$$(\nabla_x^{k+1} - \nabla_x^k \nabla_v) h \cdot [\mathcal{L}_D, \nabla_x^{k+1} - \nabla_x^k \nabla_v] h \geq -\frac{1}{8} |(\nabla_x^{k+1} - \nabla_x^k \nabla_v) h|^2 - C |\nabla_x^k \nabla_v h|^2.$$

Proof. We have, for all $\alpha_1, \alpha_2 \in \mathbb{N}^d$,

$$\begin{aligned} \partial_x^{\alpha_1} \partial_v^{\alpha_2} \mathcal{L}_D h(x, v) &= -\gamma \sum_{j=1}^d \partial_v^{\alpha_2} (v_j \partial_x^{\alpha_1} \partial_{v_j} h) + \gamma \sum_{j=1}^d \partial_x^{\alpha_1} \partial_v^{\alpha_2} (\partial_{v_j^2}^2 h) \\ &= -\gamma \sum_{j=1}^d v_j \partial_v^{\alpha_2} \partial_x^{\alpha_1} \partial_{v_j} h - \gamma \sum_{j=1}^d \alpha_{2,j} \partial_v^{\alpha_2} \partial_x^{\alpha_1} \partial_{v_j} h + \gamma \sum_{j=1}^d \partial_x^{\alpha_1} \partial_v^{\alpha_2} (\partial_{v_j^2}^2 h). \end{aligned}$$

Therefore, if $|\alpha_1| = i$ and $|\alpha_2| = k+1-i$, we have

$$[\mathcal{L}_D, \partial_x^{\alpha_1} \partial_v^{\alpha_2}] h(x, v) = \gamma \sum_{j=1}^d \alpha_{2,j} \partial_v^{\alpha_2} \partial_x^{\alpha_1} h = \gamma(k+1-i) \partial_x^{\alpha_1} \partial_v^{\alpha_2} h,$$

and

$$\nabla_x^i \nabla_v^{k+1-i} h \cdot [\mathcal{L}_D, \nabla_x^i \nabla_v^{k+1-i}] h = \gamma(k+1-i) |\nabla_x^i \nabla_v^{k+1-i} h|^2.$$

Similarly, for all $\beta \in \mathbb{N}^d$ and $i \in \llbracket 1, d \rrbracket$,

$$\begin{aligned} (\partial_x^\beta \partial_{x_i} - \partial_x^\beta \partial_{v_i}) \mathcal{L}_D h(x, v) &= -\gamma \sum_{j=1}^d v_i \partial_x^\beta \partial_{x_i} \partial_{v_j} h + \gamma \sum_{j=1}^d \partial_{v_i} (v_j \partial_x^\beta \partial_{v_j} h) \\ &\quad + \gamma \sum_{j=1}^d \partial_x^\beta \partial_{x_i} \partial_{v_j^2}^2 h - \gamma \sum_{j=1}^d \partial_x^\beta \partial_{v_i} \partial_{v_j^2}^2 h \\ &= -\gamma \sum_{j=1}^d v_j \partial_x^\beta \partial_{x_i} \partial_{v_j} h + \gamma \sum_{j=1}^d v_j \partial_{v_i} \partial_x^\beta \partial_{v_j} h + \gamma \partial_x^\beta \partial_{v_i} h \\ &\quad + \gamma \sum_{j=1}^d \partial_x^\beta \partial_{x_i} \partial_{v_j^2}^2 h - \gamma \sum_{j=1}^d \partial_x^\beta \partial_{v_i} \partial_{v_j^2}^2 h. \end{aligned}$$

Therefore,

$$[\mathcal{L}_D, \partial_x^\beta \partial_{x_i} - \partial_x^\beta \partial_{v_i}] h(x, v) = \gamma \partial_x^\beta \partial_{v_i} h,$$

and finally, by applying Young inequality, we get

$$\begin{aligned} (\nabla_x^{k+1} - \nabla_x^k \nabla_v) h \cdot [\mathcal{L}_D, \nabla_x^{k+1} - \nabla_x^k \nabla_v] h &= \gamma \sum_{\substack{|\beta|=k \\ j \in \llbracket 1, d \rrbracket}} (\partial_x^\beta \partial_{x_j} - \partial_x^\beta \partial_{v_j}) h \partial_x^\beta \partial_{v_j} h \\ &\geq -\frac{1}{8} |(\nabla_x^{k+1} - \nabla_x^k \nabla_v) h|^2 - C |\nabla_x^k \nabla_v h|^2, \end{aligned}$$

for some $C > 0$. ■

Finally, let us look at the jump part.

Lemma 3.16. *For all $k \in \mathbb{N}$, there exists $C > 0$ such that for $i \in \llbracket 0, k \rrbracket$ and $\varepsilon > 0$,*

$$\begin{aligned} & \int_{\mathbb{R}^{2d}} \nabla_x^i \nabla_v^{k+1-i} h \cdot [\mathcal{L}_J, \nabla_x^i \nabla_v^{k+1-i}] h \, d\mu \\ & \geq -\frac{1}{\varepsilon} \int_{\mathbb{R}^{2d}} |\nabla_x^i \nabla_v^{k+1-i} h|^2 \, d\mu - C\varepsilon \int_{\mathbb{R}^{2d}} \left(\sum_{\substack{j \leq i \\ l \leq k+1-j}} |\nabla_x^j \nabla_v^l h|^2 + |\nabla_x^i \nabla_v^{k+2-i} h|^2 \right) d\mu. \end{aligned}$$

and

$$\begin{aligned} & \int_{\mathbb{R}^{2d}} (\nabla_x^{k+1} - \nabla_x^k \nabla_v h) \cdot [\mathcal{L}_J, \nabla_x^{k+1} - \nabla_x^k \nabla_v] h \, d\mu \\ & \geq -\frac{1}{8} \int_{\mathbb{R}^{2d}} |(\nabla_x^{k+1} - \nabla_x^k \nabla_v) h|^2 \, d\mu - C \left(\int_{\mathbb{R}^{2d}} \sum_{\substack{i \leq k \\ j \leq k+1-i}} |\nabla_x^i \nabla_v^j h|^2 \, d\mu + \int_{\mathbb{R}^{2d}} |\nabla_x^k \nabla_v^2 h|^2 \, d\mu \right). \end{aligned}$$

To prove this Lemma, we need a generalization of Proposition 3.10:

Lemma 3.17. *Let $n \in \mathbb{N}^*$. There exists $C > 0$ such that for any $h \in \mathcal{C}_c^\infty(\mathbb{R})$,*

$$\int_{\mathbb{R}} v^{2n} h^2(v) e^{-v^2/2} \, dv \leq C \sum_{k=0}^n \int_{\mathbb{R}} (h^{(k)}(v))^2 e^{-v^2/2} \, dv.$$

Proof. We prove the inequality by induction on n . The case $n = 1$ corresponds to Lemma 3.10. Suppose that the inequality holds for some $n \in \mathbb{N}^*$. Then

$$\begin{aligned} \int_{\mathbb{R}} v^{2n+2} h^2(v) e^{-v^2/2} \, dv &= \int_{\mathbb{R}} v^{2n} (vh(v))^2 e^{-v^2/2} \, dv \\ &\leq C \sum_{k=0}^n \int_{\mathbb{R}} ((vh(v))^{(k)})^2 e^{-v^2/2} \, dv \\ &= C \sum_{k=0}^n \int_{\mathbb{R}} (kh^{(k-1)} + vh^{(k)})^2 e^{-v^2/2} \, dv \\ &\leq C \sum_{k=0}^n 2 \int_{\mathbb{R}} (k^2 (h^{(k-1)})^2 + v^2 (h^{(k)})^2) e^{-v^2/2} \, dv \\ &\leq 2Cn^2 \sum_{k=0}^{n-1} \int_{\mathbb{R}} (h^{(k)}(v))^2 e^{-v^2/2} \, dv + 4C \sum_{k=0}^n \int_{\mathbb{R}} (h^{(k)}(v))^2 e^{-v^2/2} \, dv \\ &\quad + 8C \sum_{k=1}^{n+1} \int_{\mathbb{R}} (h^{(k)})^2 e^{-v^2/2} \, dv \\ &\leq \tilde{C} \sum_{k=0}^{n+1} \int_{\mathbb{R}} (h^{(k)})^2 e^{-v^2/2} \, dv, \end{aligned}$$

and hence it holds for $n + 1$. ■

Proof of Lemma 3.16. For all $i, j \in \llbracket 1, d \rrbracket$,

$$\partial_{v_j}(x, V^i) = (0, e_j + (\rho - 1)e_i \mathbb{1}_{j=i}), \quad \partial_{x_j}(x, V^i) = (e_j, 0), \quad \partial_x^\nu \partial_x^\alpha(x, V^i) = 0 \quad \forall \quad |\nu| + |\alpha| \geq 2.$$

As a consequence, for any $\nu_1, \nu_2 \in \mathbb{N}^d$, we have

$$\partial_x^{\nu_1} \partial_v^{\nu_2}(h(x, V^i)) = \rho^{\nu_2, i} (\partial_x^{\nu_1} \partial_v^{\nu_2} h)(x, V^i).$$

For all $i \in \llbracket 1, d \rrbracket$,

$$\partial_x^\nu \frac{\partial_i U_1}{2}(v_i - V_i^i) = \partial^\nu \partial_i U_1 \frac{v_i - V_i^i}{2}, \quad \partial_x^\nu \partial_{v_i} \frac{\partial_i U_1}{2}(v_i - V_i^i) = \frac{1-\rho}{2} \partial^\nu \partial_i U_1.$$

This yields that the only non-vanishing derivatives of $\Psi(\frac{\partial_i U_1}{2}(v_i - V_i^i))$ are of the type $\partial_x^\alpha \partial_{v_i}^k \Psi(\frac{\partial_i U_1}{2}(v_i - V_i^i))$. For $k \in \mathbb{N}$ and $\alpha \in \mathbb{N}^d$, we have:

$$\partial_{v_i}^k \Psi\left(\frac{\partial_i U_1}{2}(v_i - V_i^i)\right) = \Psi^{(k)}\left(\frac{\partial_i U_1}{2}(v_i - V_i^i)\right) \left(\frac{1-\rho}{2}\right)^k (\partial_i U_1)^k,$$

and

$$\partial_x^\alpha \partial_{v_i}^k \Psi\left(\frac{\partial_i U_1}{2}(v_i - V_i^i)\right) = \left(\frac{1-\rho}{2}\right)^k \sum_{\nu \leq \alpha} \binom{\alpha}{\nu} \partial_x^\nu \Psi^{(k)}\left(\frac{\partial_i U_1}{2}(v_i - V_i^i)\right) \partial_x^{\alpha-\nu} (\partial_i U_1)^k.$$

Therefore, using that $\Psi(s) \leq c + |s|$ for some $c > 0$, and that Ψ' and $\partial_i U_1$ and their derivatives of all order are bounded, we get that for all $k \in \mathbb{N}$ and $\alpha \in \mathbb{N}^d$ there exists $C > 0$ such that:

$$\left| \partial_x^\alpha \partial_{v_i}^k \Psi\left(\frac{\partial_i U_1}{2}(v_i - V_i^i)\right) \right| \leq C \left(1 + |v_i - V_i^i|^{| \alpha | + k}\right).$$

Thus, for $\alpha \in \mathbb{N}^{2d}$, we have

$$\begin{aligned} [\partial^\alpha, \mathcal{L}_J^i]h &= \frac{2}{1-\rho} \mathbb{E} \left[\partial^\alpha \left((h(x, V^i) - h) \Psi\left(\frac{\partial_i U_1}{2}(v_i - V_i^i)\right) \right) \right] \\ &\quad - \frac{2}{1-\rho} \mathbb{E} \left[(\partial^\alpha h(x, V^i) - \partial^\alpha h) \Psi\left(\frac{\partial_i U_1}{2}(v_i - V_i^i)\right) \right] \\ &= \frac{2}{1-\rho} \sum_{\nu < \alpha} \binom{\alpha}{\nu} \mathbb{E} [(\partial^\nu (h(x, V^i)) - \partial^\nu h) \partial^{\alpha-\nu} \left(\Psi\left(\frac{\partial_i U_1}{2}(v_i - V_i^i)\right) \right)] \\ &\quad + \frac{2}{1-\rho} \mathbb{E} \left[(\partial^\alpha (h(x, V^i)) - \partial^\alpha h(x, V^i)) \Psi\left(\frac{\partial_i U_1}{2}(v_i - V_i^i)\right) \right] \\ &= \frac{2}{1-\rho} \sum_{\nu < \alpha} \binom{\alpha}{\nu} \mathbb{E} \left[(\rho^{\nu_i+d} \partial^\nu h(x, V^i) - \partial^\nu h) \partial^{\alpha-\nu} \left(\Psi\left(\frac{\partial_i U_1}{2}(v_i - V_i^i)\right) \right) \right] \\ &\quad + \frac{2}{1-\rho} \mathbb{E} \left[(\rho^{\alpha_i+d} - 1) \partial^\alpha h(x, V^i) \Psi\left(\frac{\partial_i U_1}{2}(v_i - V_i^i)\right) \right]. \end{aligned}$$

As a consequence, there exists a $C > 0$ such that

$$\begin{aligned} &|[\partial^\alpha, \mathcal{L}_J^i]h|^2 \\ &\leq C \left(\mathbb{E}[(\partial^\alpha h(x, V^i))^2 (1 + |v_i - V_i^i|^2)] + \sum_{\nu < \alpha} \mathbb{E}[(\partial^\nu h(x, V^i) - \partial^\nu h)^2 (1 + |v_i - V_i^i|^{2|\alpha-\nu|})] \right). \end{aligned}$$

As in the H^1 case, the integrals with respect to μ of the previous expectations can be bounded using the reversibility of the law of V^i with respect to the Gaussian measure, as well as Lemma 3.17. For any $n \in \mathbb{N}^*$,

$$\mathbb{E}[|v_i - V_i^i|^{2n}] = \mathbb{E}[|(1-\rho)v_i + \sqrt{1-\rho^2}\xi|^{2n}] \leq C(1 + |v_i|^{2n}).$$

Therefore, using (3.5) and Lemma 3.17, for all $\nu \in \mathbb{N}^{2d}$ and $n \in \mathbb{N}$:

$$\begin{aligned} \int_{\mathbb{R}^{2d}} \mathbb{E}[|v_i - V_i^i|^{2n} \partial^\nu h(x, V^i)^2] d\mu &= \int_{\mathbb{R}^{2d}} (\partial^\nu h)^2 \mathbb{E}[|v_i - V_i^i|^{2n}] \\ &\leq C \int_{\mathbb{R}^{2d}} (1 + |v_i|^{2n}) |\partial^\nu h|^2 d\mu \leq C \sum_{k=0}^n \int_{\mathbb{R}^{2d}} |\partial_{v_i}^k \partial^\nu h|^2 d\mu. \end{aligned}$$

Putting everything together, we get that for all $\alpha \in \mathbb{N}^{2d}$, there is $C > 0$ such that

$$\int_{\mathbb{R}^{2d}} |[\partial^\alpha, \mathcal{L}_J^i]h|^2 d\mu \leq C \left(\int_{\mathbb{R}^{2d}} (\partial_{v_i} \partial^\alpha h)^2 d\mu + \sum_{\nu \leq \alpha} \sum_{k=0}^{|\alpha-\nu|} \int_{\mathbb{R}^{2d}} (\partial_{v_i}^k \partial^\nu h)^2 d\mu \right).$$

For any $i \in \llbracket 0, k \rrbracket$ and $\varepsilon > 0$, Young inequality yields

$$\begin{aligned} & \int_{\mathbb{R}^{2d}} \nabla_x^i \nabla_v^{k+1-i} h \cdot [\mathcal{L}_J, \nabla_x^i \nabla_v^{k+1-i}] h d\mu \\ & \geq -\frac{1}{\varepsilon} \int_{\mathbb{R}^{2d}} |\nabla_x^i \nabla_v^{k+1-i} h|^2 d\mu - C\varepsilon \int_{\mathbb{R}^{2d}} \left(\sum_{\substack{j \leq i \\ l \leq k+1-j}} |\nabla_x^j \nabla_v^l h|^2 + |\nabla_x^i \nabla_v^{k+2-i} h|^2 \right) d\mu \end{aligned}$$

To treat the term with derivative $\nabla_x^{k+1} - \nabla_x^k \nabla_v$, write for $\alpha \in \mathbb{N}^d$

$$[\partial_x^\alpha, \mathcal{L}_J^i]h = \frac{2}{1-\rho} \sum_{\nu < \alpha} \binom{\alpha}{\nu} \mathbb{E} \left[\left(\partial_x^\nu h(x, V^i) - \partial_x^\nu h \right) \partial_x^{\alpha-\nu} \left(\Psi \left(\frac{\partial_i U_1}{2} (v_i - V_i^i) \right) \right) \right],$$

so that no derivatives of order $k+2$ appear, nor derivatives of the form ∇_x^{k+1} , and we have

$$\int_{\mathbb{R}^{2d}} |[\partial_x^\alpha, \mathcal{L}_J^i]h|^2 d\mu \leq C \int_{\mathbb{R}^{2d}} \sum_{\nu < \alpha} \sum_{k=0}^{|\alpha-\nu|} (\partial_{v_i}^k \partial_x^\nu h)^2 d\mu.$$

Hence, the only derivatives of order $k+2$ that appear in $[\mathcal{L}_J, \nabla_x^{k+1} - \nabla_x^k \nabla_v]h$ come from the term $[\mathcal{L}_J, \nabla_x^k \nabla_v]h$ that we treated above. Using Young inequality, we get

$$\begin{aligned} & \int_{\mathbb{R}^{2d}} (\nabla_x^{k+1} - \nabla_x^k \nabla_v h) \cdot [\mathcal{L}_J, \nabla_x^{k+1} - \nabla_x^k \nabla_v] h d\mu \\ & \geq -\frac{1}{8} \int_{\mathbb{R}^{2d}} |\nabla_x^{k+1} - \nabla_x^k \nabla_v h|^2 d\mu - C \left(\int_{\mathbb{R}^{2d}} \sum_{\substack{i \leq k \\ j \leq k+1-i}} |\nabla_x^i \nabla_v^j h|^2 d\mu - \int_{\mathbb{R}^{2d}} |\nabla_x^k \nabla_v^2 h|^2 d\mu \right), \end{aligned}$$

which concludes the proof. $\blacksquare$

Combining all the commutator terms, we get that for $i \in \llbracket 0, k-1 \rrbracket$ there exists $C > 0$ such that for all $\varepsilon > 0$ small enough,

$$\begin{aligned} & \int_{\mathbb{R}^{2d}} \nabla_x^i \nabla_v^{k+1-i} h \cdot [\mathcal{L}, \nabla_x^i \nabla_v^{k+1-i}] h d\mu \\ & \geq -\frac{C}{\varepsilon} \int_{\mathbb{R}^{2d}} |\nabla_x^i \nabla_v^{k+1-i} h|^2 d\mu - C \int_{\mathbb{R}^{2d}} |\nabla_x^{i+1} \nabla_v^{k-i} h|^2 d\mu - C \sum_{j=0}^{i-1} \int_{\mathbb{R}^{2d}} |\nabla_x^j \nabla_v^{k+2-i} h|^2 d\mu \\ & \quad - C\varepsilon \int_{\mathbb{R}^{2d}} \left(\sum_{\substack{j \leq i \\ l \leq k+1-j}} |\nabla_x^j \nabla_v^l h|^2 + |\nabla_x^i \nabla_v^{k+2-i} h|^2 \right) d\mu, \end{aligned}$$

as well as

$$\begin{aligned}
 & \int_{\mathbb{R}^{2d}} \nabla_x^k \nabla_v h \cdot [\mathcal{L}, \nabla_x^k \nabla_v] h \, d\mu \\
 & \geq -C \left(\frac{1}{\varepsilon} + \frac{1}{\eta} \right) \int_{\mathbb{R}^{2d}} |\nabla_x^k \nabla_v h|^2 \, d\mu - \frac{\eta}{8} \int_{\mathbb{R}^{2d}} |(\nabla_x^{k+1} - \nabla_x^k \nabla_v) h|^2 \, d\mu - C \sum_{j=0}^{k-1} \int_{\mathbb{R}^{2d}} |\nabla_x^j \nabla_v^2 h|^2 \, d\mu \\
 & \quad - C\varepsilon \int_{\mathbb{R}^{2d}} \left(\sum_{\substack{j \leq k \\ l \leq k+1-j}} |\nabla_x^j \nabla_v^l h|^2 + |\nabla_x^k \nabla_v^2 h|^2 \right) \, d\mu,
 \end{aligned}$$

and finally

$$\begin{aligned}
 & \int_{\mathbb{R}^{2d}} (\nabla_x^{k+1} - \nabla_x^k \nabla_v h) \cdot [\mathcal{L}, \nabla_x^{k+1} - \nabla_x^k \nabla_v] h \, d\mu \\
 & \geq \frac{1}{4} \int_{\mathbb{R}^{2d}} |\nabla_x^{k+1} - \nabla_x^k \nabla_v h|^2 \, d\mu - C \int_{\mathbb{R}^{2d}} \sum_{\substack{i \leq k \\ j \leq k+1-i}} |\nabla_x^i \nabla_v^j h|^2 \, d\mu \\
 & \quad - C \int_{\mathbb{R}^{2d}} \left(\sum_{j=0}^k |\nabla_x^j \nabla_v h|^2 + \sum_{j=0}^k |\nabla_x^j \nabla_v^2 h|^2 \right) \, d\mu.
 \end{aligned}$$

Thanks to those expressions, we are now able to prove Proposition 3.12.

Proof of Proposition 3.12.. The fact that inequality (3.9) holds for $k = 1$ is proven in Lemma 3.7. Let $k \in \mathbb{N}^*$, and suppose that inequality (3.9) holds for such k and some $\rho_k > 0$ and $\omega_{i,p} > 0$, $0 \leq i \leq p \leq k$. Fix some $\omega_{i,k+1} > 0$, $0 \leq i \leq k+1$, $\varepsilon > 0$ small enough so that the previous formulas for the commutators hold, and choose $\eta = \omega_{k+1,k+1}/\omega_{k,k+1}$. By assumption we have

$$\begin{aligned}
 & \int_{\mathbb{R}^{2d}} \Gamma(h) \, d\mu + \sum_{p=1}^{k+1} \int_{\mathbb{R}^{2d}} \left(\sum_{i=0}^{p-1} \omega_{i,p} \Gamma_{i,p}(h) + \omega_{p,p} \Gamma_{p,p}(h) \right) \, d\mu \\
 & \geq \rho_k \left(N_k(h) + \int_{\mathbb{R}^{2d}} \sum_{i=0}^k |\nabla_x^i \nabla_v^{k+1-i} h|^2 \, d\mu \right) + \int_{\mathbb{R}^{2d}} \left(\sum_{i=0}^k \omega_{i,k+1} \Gamma_{i,k+1}(h) + \omega_{k+1,k+1} \Gamma_{k+1,k+1}(h) \right) \, d\mu.
 \end{aligned}$$

By using Proposition 3.4, the bounds on the commutators and the fact that

$$\Gamma(\nabla_x^i \nabla_v^j h) \geq 2\gamma |\nabla_x^i \nabla_v^{j+1} h|^2,$$

$$\Gamma((\nabla_x^{k+1} - \nabla_x^k \nabla_v) h) \geq 2\gamma |(\nabla_x^{k+1} \nabla_v - \nabla_x^k \nabla_v^2) h|^2 \geq \gamma |\nabla_x^{k+1} \nabla_v h|^2 - 2\gamma |\nabla_x^k \nabla_v^2 h|^2,$$

the term of order $k+2$ are bounded below by

$$\begin{aligned}
 & 2\gamma \sum_{i=0}^k \omega_{i,k+1} |\nabla_x^i \nabla_v^{k+2-i} h|^2 + \gamma \omega_{k+1,k+1} |\nabla_x^{k+1} \nabla_v h|^2 \\
 & \quad - (C_1 + 2\gamma) \omega_{k+1,k+1} |\nabla_x^k \nabla_v^2 h|^2 - C_1 \varepsilon \sum_{i=0}^k \omega_{i,k+1} |\nabla_x^i \nabla_v^{k+2-i} h|^2,
 \end{aligned}$$

for some constant $C_1 > 0$. We may bound from below the terms of order $k + 1$ for ε small enough by

$$\begin{aligned} \rho_k \sum_{i=0}^k |\nabla_x^i \nabla_v^{k+1-i} h|^2 - \sup_{0 \leq i \leq k+1} \omega_{i,k+1} \frac{2C_2}{\varepsilon} \sum_{i=0}^k |\nabla_x^i \nabla_v^{k+1-i} h|^2 \\ - C_2 \frac{\omega_{k,k+1}^2}{\omega_{k+1,k+1}} |\nabla_x^k \nabla_v h|^2 + \frac{\omega_{k+1,k+1}}{8} |(\nabla_x^{k+1} - \nabla_x^k \nabla_v) h|^2, \end{aligned}$$

for some constant $C_2 > 0$. Finally, the term of order at most k may be bounded from below by

$$\left(\rho_k \inf_{i+j \leq k} \omega_{i,i+j} - 3C_3 \sup_{0 \leq i \leq k+1} \omega_{i,k+1} \right) \sum_{i+j \leq k} |\nabla_x^i \nabla_v^j h|^2,$$

for some constant $C_3 > 0$. Set $C_4 = \max(C_1, C_2, C_3)$ (which is independent from ε and the weights $\omega_{i,j}$). Now, choose ε and $\omega_{i,k+1}$ for $i \in \llbracket 1, k+1 \rrbracket$ such that

$$\frac{\omega_{k+1,k+1}}{\omega_{k,k+1}} \leq \frac{\gamma - C_4 \varepsilon}{2\gamma + C_4}, \quad \varepsilon < \gamma/C_4,$$

$$\omega_{i,k+1} < \varepsilon \rho_k / 4C_4, \quad 1 \leq i \leq k-1, \quad \omega_{k,k+1} \left(\frac{2}{\varepsilon} + \frac{\omega_{k,k+1}}{\omega_{k+1,k+1}} \right) < \rho_k / (2C_4),$$

and

$$6C_4 \sup_{0 \leq i \leq k+1} \omega_{i,k+1} < \rho_k \inf_{i+j \leq k} \omega_{i,i+j}.$$

We get

$$\begin{aligned} \int_{\mathbb{R}^{2d}} \Gamma(h) d\mu + \sum_{p=1}^{k+1} \int_{\mathbb{R}^{2d}} \left(\sum_{i=0}^{p-1} \omega_{i,p} \Gamma_{i,p}(h) + \omega_{p,p} \Gamma_{p,p}(h) \right) d\mu \\ \geq \tilde{\rho}_{k+1} \left(\sum_{p=1}^{k+1} \int_{\mathbb{R}^{2d}} \left(\sum_{i=0}^{p-1} \omega_{i,p} |\nabla_x^i \nabla_v^{p-i} f|^2 + \omega_{p,p} |(\nabla_x^p - \nabla_x^{p-1} \nabla_v) f|^2 \right) + \sum_{i=0}^k |\nabla_x^i \nabla_v^{k+2-i} h|^2 d\mu \right), \end{aligned}$$

for any

$$\tilde{\rho}_{k+1} < \min \left(\frac{\rho_k}{2}, \frac{\omega_{k+1,k+1}}{8}, \gamma \inf_{i+j \leq k+1} \omega_{i,i+j} \right).$$

As in the $H^1(\mu)$ case, an application of the Poincaré inequality (3.6) yields inequality (3.9), which concludes the induction. $\blacksquare$

Proof of Theorem 3.1. Our goal is to apply Lumer–Phillips theorem to the operator $\mathcal{L} + \rho_k/2I$, where I denotes the identity operator of H^k . This theorem can be stated as follows: an operator A on a Hilbert space generates a contraction semi-group if and only if it is maximally dissipative, see [51, Chapter IX, p. 250]. Fix $k \in \mathbb{N}$. The scalar product

$$\begin{aligned} \langle f, g \rangle_k^2 = \int_{\mathbb{R}^{2d}} f g d\mu \\ + \int_{\mathbb{R}^{2d}} \sum_{p=1}^k \left(\sum_{i=0}^{p-1} \omega_{i,p} \nabla_x^i \nabla_v^{p-i} f \cdot \nabla_x^i \nabla_v^{p-i} g + \omega_{p,p} (\nabla_x^p - \nabla_x^{p-1} \nabla_v) f \cdot (\nabla_x^p - \nabla_x^{p-1} \nabla_v) g \right) d\mu \end{aligned}$$

generates a norm equivalent to the usual norm of H^k . Proposition 3.12 and a density argument yield that the operator $\mathcal{L} + \rho_k/2I$, is dissipative:

$$\forall h \in D(\mathcal{L}), \quad \langle (\mathcal{L} + \rho_k/2I)h, h \rangle_k \leq 0,$$

where $D(\mathcal{L})$ denote the domain of $\mathcal{L}$ in H^k defined by:

$$f \in D(\mathcal{L}), g = \mathcal{L}f \iff f \in H^k, \lim_{t \rightarrow 0} \left\| \frac{P_t f - f}{t} - g \right\|_{H^k} = 0.$$

We are left to show that $\mathcal{L} + \rho I$, for some $\rho < \rho_k/2$, is surjective. Fix such a $\rho < \rho_k/2$. Thanks to Proposition 3.12, we have that

$$\Lambda : (f, g) \mapsto \langle -(\mathcal{L} + \rho I)f, g \rangle_k$$

is coercive and continuous from $(H^k)^2$ to $\mathbb{R}$ for $k \geq 1$. Hence, we may apply Lax–Milgram theorem to get that for all $g \in H^k$, there exists $f \in H^k$ such that for all $h \in H^k$, we have:

$$\langle -(\mathcal{L} + \rho I)f, h \rangle_{H^k} = \langle -g, h \rangle_{H^k},$$

which implies that f is a solution to the equation

$$(\mathcal{L} + \rho I)f = g,$$

and $\mathcal{L} + \rho_k/2I$ is maximally dissipative. Lumer–Phillips Theorem then yields that the semi-group generated by $\mathcal{L} + \rho_k/2$ is a contraction on H^k : for all $f \in H^k$

$$N_k \left(e^{\rho_k t/2} (P_t f - \mu(f)) \right) \leq N_k (f - \mu(f)),$$

which concludes the proof. ■

3.6. Lyapunov function and re-weighted Sobolev norm

The goal of this section is to establish Propositions 3.2 and 3.3, thanks to which the previous exponential decay in H^k is adapted to a weighted-Sobolev norm.

Recall the definition of V_b in (3.2).

Proof of Proposition 3.2. Let $a > 0$ and $b \in (0, 1)$. First, notice that

$$\begin{aligned} \nabla_x V_b &= b \left(\nabla U_0 - a \left(\frac{v}{\sqrt{1+|x|^2}} - \frac{x \cdot v}{(1+|x|^2)^{3/2}} x \right) \right) V_b, \\ \nabla_v V_b &= b \left(v - a \frac{x}{\sqrt{1+|x|^2}} \right) V_b, \\ \Delta_v V_b &= b \left(d + b \left(|v|^2 + a^2 \frac{|x|^2}{1+|x|^2} - 2a \frac{x \cdot v}{\sqrt{1+|x|^2}} \right) \right) V_b. \end{aligned}$$

Let us denote

$$\begin{aligned} \mathcal{L}_1^* f &= -v \cdot \nabla_x f + (\nabla U_0 - \gamma v) \cdot \nabla_v f + \gamma \Delta_v f, \\ \mathcal{L}_2^* f &= \frac{2}{1-\rho} \sum_{i=1}^d \mathbb{E} \left[(f(x, V^i) - f(x, v)) \Psi \left(\frac{\partial_i U_1}{2} (V_i^i - v_i) \right) \right]. \end{aligned}$$

Notice that $\mathcal{L}_1^*$ is the adjoint of the generator of the Langevin diffusion associated to the measure with density proportional to $\exp(-(U_0(x) + |v|^2/2))$, not μ . When we integrate with respect to μ , as we saw in Section 3.3, additional terms appear in $(\mathcal{L}_D + \mathcal{L}_H)^*$ and $\mathcal{L}_J^*$ and then cancel out when they are summed. Therefore, although $\mathcal{L}_1^* \neq (\mathcal{L}_D + \mathcal{L}_H)^*$ and $\mathcal{L}_2^* \neq \mathcal{L}_J^*$, we indeed have $\mathcal{L}^* = \mathcal{L}_1^* + \mathcal{L}_2^*$.

Using that $-\nabla U_0(x) \cdot x \leq -\kappa|x|^2 + C$, and the Young inequality, for all $\varepsilon > 0$,

$$\begin{aligned} \frac{\mathcal{L}_1^* V_b}{bV_b} &= \left(\frac{a}{\sqrt{1+|x|^2}} - \gamma(1-b) \right) |v|^2 - a \frac{(x \cdot v)^2}{(1+|x|^2)^{3/2}} \\ &\quad - a \frac{x \cdot \nabla U_0}{\sqrt{1+|x|^2}} + a\gamma(1-2b) \frac{x \cdot v}{\sqrt{1+|x|^2}} + b\gamma a^2 \frac{|x|^2}{1+|x|^2} + d\gamma \\ &\leq \left(\frac{a}{\sqrt{1+|x|^2}} + \frac{a\gamma(1-2b)}{2\varepsilon\sqrt{1+|x|^2}} - \gamma(1-b) \right) |v|^2 + (\varepsilon\gamma(1-2b) - \kappa) \frac{a|x|^2}{\sqrt{1+|x|^2}} + C + d\gamma + b\gamma a^2. \end{aligned}$$

Then, by taking small enough ε and a , the terms in front of $|v|^2$ and $|x|^2$ are negative. This shows that there exist constants C_1, C_2, C_3 such that

$$\mathcal{L}_1^* V_b \leq (C_1 - C_2|v|^2 - C_3|x|)V_b. \quad (3.10)$$

Now,

$$\mathcal{L}_2^* V_b = CV_b \sum_{i=1}^d \mathbb{E} \left[\left(\exp \left(b \left(\frac{|V_i^i|^2 - |v_i|^2}{2} - a \frac{x_i \cdot (V_i^i - v_i)}{\sqrt{1+|x|^2}} \right) \right) - 1 \right) \Psi \left(\frac{\partial_i U_1}{2} (V_i^i - v_i) \right) \right].$$

The term inside the previous exponential can be expressed as:

$$\begin{aligned} &\frac{b}{2} \left((\rho^2 - 1)v_i^2 + (1 - \rho^2)\xi^2 + 2\rho\sqrt{1 - \rho^2}v_i\xi - \frac{2ax_i}{\sqrt{1+|x|^2}}((\rho - 1)v_i + \sqrt{1 - \rho^2}\xi) \right) \\ &\leq \frac{b}{2} \left((\rho^2 - 1 + \varepsilon a(1 - \rho))v_i^2 + (1 - \rho^2)\xi^2 - \left(\frac{a\sqrt{1 - \rho^2}}{\sqrt{1+|x|^2}}x_i - 2\rho\sqrt{1 - \rho^2}v_i \right) \xi + \frac{a(1 - \rho)}{\varepsilon(1+|x|^2)}x_i^2 \right). \end{aligned}$$

For any $z \in \mathbb{C}$, if $\xi \sim \mathcal{N}(0, 1)$,

$$\mathbb{E}[e^{z\xi}] = e^{z^2/2},$$

and for any $x \in \mathbb{R}$, $\mathbb{E}[e^{x\xi^2}] < \infty$ if and only if $x < 1/2$. Therefore, the previous exponential is integrable if $b < \frac{1}{1-\rho^2}$, which is always the case since $b < 1$ and $\rho^2 < 1$. We have

$$\begin{aligned} &\mathbb{E} \left[\exp \left(b \left(\frac{|V_i^i|^2 - |v_i|^2}{2} - a \frac{x_i \cdot (V_i^i - v_i)}{\sqrt{1+|x|^2}} \right) \right) \right] \\ &\leq \frac{\exp \left(\frac{b}{2} \left((\rho^2 - 1 + \varepsilon a(1 - \rho))v_i^2 + \frac{a(1-\rho)}{\varepsilon(1+|x|^2)}x_i^2 \right) \right)}{\sqrt{1 - b(1 - \rho^2)}} \times \exp \left(\frac{b^2}{4} \left(\frac{a\sqrt{1 - \rho^2}}{\sqrt{1+|x|^2}}x_i - 2\rho\sqrt{1 - \rho^2}v_i \right)^2 \right) \\ &\leq C \exp \left(\frac{b}{2} \left(\rho^2 - 1 + \varepsilon a(1 - \rho) + 4b\rho^2(1 - \rho^2) \right) v_i^2 \right) \times \exp \left(\frac{b^2}{2} \left(\frac{a(1 - \rho)}{\varepsilon} + a^2(1 - \rho^2) \right) \frac{x_i^2}{1+|x|^2} \right). \end{aligned}$$

By taking small enough a and b , the term in front of v_i^2 is negative, and the previous expectation is bounded. Recall that for all $s \in \mathbb{R}$, $\Psi(s) \leq C + |s|$, therefore

$$\begin{aligned} \mathcal{L}_2^* V_b &= CV_b \sum_{i=1}^d \mathbb{E} \left[(V_b(x, V^i)/V_b - 1) \Psi \left(\frac{\partial_i U_1}{2} (V_i^i - v_i) \right) \right] \\ &\leq CV_b \sum_{i=1}^d \sqrt{\mathbb{E}[(V_b(x, V^i)/V_b - 1)^2]} \mathbb{E} \left[\Psi^2 \left(\frac{\partial_i U_1}{2} (V_i^i - v_i) \right) \right] \\ &\leq CV_b \sum_{i=1}^d \sqrt{1 + \mathbb{E}[(V_i^i - v_i)^2]} \leq C(1 + |v|)V_b. \end{aligned}$$

Thus, there exist positive constants C_1, C_2 and C_3 such that

$$\mathcal{L}^* V_b \leq (C_1 - C_2|v|^2 - C_3|x|)V_b,$$

which shows that V_b is indeed a Lyapunov function for $\mathcal{L}^*$. ■

Proof of Proposition 3.3. We define, for all $k \in \mathbb{N}^*$ and $h \in \mathcal{C}_c^\infty$, the function ϕ_k by

$$\phi_k(h) = h^2 + \sum_{p=1}^k \sum_{i=0}^{p-1} \omega_{i,p} |\nabla_x^i \nabla_v^{p-i} h|^2 + \omega_{p,p} |(\nabla_x^p - \nabla_x^{p-1} \nabla_v) h|^2,$$

where the $\omega_{i,j}$ are the same as in the norms N_k . In order to prove the result of the lemma, it is sufficient to show the exponential decay of $\int \phi_k(f_t) V_b d\mu$, with $f_t = P_t f - \mu(f)$, for a certain $b \in (0, 1)$. Following the notations introduced in equation (3.3), and by denoting $\mathcal{L}_0 = \mathcal{L}_H + \mathcal{L}_D$,

$$\begin{aligned} \partial_t \int_{\mathbb{R}^{2d}} \phi_k(f_t) V_b d\mu &= \int_{\mathbb{R}^{2d}} D_h \phi_k(f_t) \mathcal{L} f_t V_b d\mu \\ &= \int_{\mathbb{R}^{2d}} D_h \phi_k(f_t) \mathcal{L}_0 f_t V_b d\mu + \int_{\mathbb{R}^{2d}} D_h \phi_k(f_t) \mathcal{L}_J f_t V_b d\mu \\ &= -2 \int_{\mathbb{R}^{2d}} \Gamma_{\mathcal{L}_0, \phi_k}(f_t) V_b d\mu + \int_{\mathbb{R}^{2d}} \mathcal{L}_0(\phi_k(f_t)) V_b d\mu + \int_{\mathbb{R}^{2d}} D_h \phi_k(f_t) \mathcal{L}_J f_t V_b d\mu \\ &= -2 \int_{\mathbb{R}^{2d}} \Gamma_{\mathcal{L}_0, \phi_k}(f_t) V_b d\mu + \int_{\mathbb{R}^{2d}} \phi_k(f_t) \mathcal{L}_0^* V_b d\mu + \int_{\mathbb{R}^{2d}} D_h \phi_k(f_t) \mathcal{L}_J f_t V_b d\mu. \end{aligned}$$

The adjoint of $\mathcal{L}_0$ in $L^2(\mu)$ is given by

$$\mathcal{L}_0^* h = -v \cdot \nabla_x h + (\nabla U_0 - \gamma v) \cdot \nabla_v h + \gamma \Delta_v h + h v \cdot \nabla U_1.$$

Therefore, using the inequality (3.10) shown in the proof of Proposition 3.2, there exists a $b \in (0, 1)$ such that

$$\begin{aligned} \mathcal{L}_0^* V_b &\leq b(-c_1 |v|^2 - c_2 |x| + C) V_b + C' |v| V_b \\ &\leq (c_0 |v| - c_1 |v|^2 - c_2 |x| + c_3) V_b, \end{aligned}$$

which shows that V_b is a Lyapunov function for $\mathcal{L}_0^*$. Moreover, since Lemmas 3.14 and 3.15 give bounds on the commutators associated to $\mathcal{L}_H$ and $\mathcal{L}_D$ (without having to integrate them with respect to μ as it is the case for the commutators with the jump part), then in the proof of Proposition 3.12, removing the parts associated with the jumps and the integrals with respect to μ shows that $\Gamma_{\mathcal{L}_0, \phi} \geq 0$. As a consequence, there exist $b \in (0, 1)$, $\rho > 0$, $\eta > 0$ and $C > 0$ such that

$$\begin{aligned} \partial_t \int_{\mathbb{R}^{2d}} \phi_k(f_t) V_b d\mu &\leq -\rho \int_{\mathbb{R}^{2d}} \phi_k(f_t) V_b d\mu + C \int_{\mathbb{R}^{2d}} \phi_k(f_t) d\mu + \int_{\mathbb{R}^{2d}} D_h \phi_k(f_t) \mathcal{L}_J f_t V_b d\mu \\ &\leq -\rho \int_{\mathbb{R}^{2d}} \phi_k(f_t) V_b d\mu + C e^{-\eta t} + \int_{\mathbb{R}^{2d}} D_h \phi_k(f_t) \mathcal{L}_J f_t V_b d\mu, \end{aligned}$$

where the last inequality comes from Theorem 3.1, since $N_k(h) = \int \phi_k(h) d\mu$. Let us now look at the term $\int_{\mathbb{R}^{2d}} D_h \phi_k(f_t) \mathcal{L}_J f_t V_b d\mu$. The expression $D_h \phi_k(f_t) \mathcal{L}_J f_t$ is a linear combination of terms of the form $\partial^\alpha f_t \partial^\alpha \mathcal{L}_J f_t$, with $\alpha \in \mathbb{N}^{2d}$. Notice that if we show that for any $\alpha \in \mathbb{N}^{2d}$, there exists $b \in (0, 1)$ such that

$$\int_{\mathbb{R}^{2d}} \partial^\alpha f_t \partial^\alpha \mathcal{L}_J(f_t) V_b d\mu \leq C e^{-\rho t},$$

then we will have

$$\int_{\mathbb{R}^{2d}} D_h \phi_k(f_t) \mathcal{L}_J f_t V_b d\mu \leq C e^{-\rho t},$$

thus,

$$\partial_t \int_{\mathbb{R}^{2d}} \phi_k(f_t) V_b d\mu \leq -\rho \int_{\mathbb{R}^{2d}} \phi_k(f_t) V_b d\mu + C e^{-\eta t},$$

and finally

$$\int_{\mathbb{R}^{2d}} \phi_k(f_t) V_b d\mu \leq C e^{-\eta t}.$$

Let $\alpha \in \mathbb{N}^{2d}$. Successive integrations by parts yield

$$\begin{aligned} \int_{\mathbb{R}^{2d}} \partial^\alpha h \partial^\alpha \mathcal{L}_J(h) V_b d\mu &= (-1)^{|\alpha|} \int_{\mathbb{R}^{2d}} \mathcal{L}_J(h) \partial^\alpha (\partial^\alpha h V_b e^{-H}) dx dv \\ &= (-1)^{|\alpha|} \sum_{\nu \leq \alpha} \binom{\alpha}{\nu} \int_{\mathbb{R}^{2d}} \mathcal{L}_J(h) \partial^{2\alpha-\nu} h \partial^\nu (V_b e^{-H}) d\mu. \end{aligned}$$

Since ∇U_1 and $\nabla^2 U_0$ are bounded (which in particular implies that $|\nabla U_0(x)| \leq C(1 + |x|)$), for any $\nu \in \mathbb{N}^{2d}$, there exist $p > 0$ and $C > 0$ such that

$$|\partial^\nu (V_b(x, v) e^{-H(x, v)})| \leq C(1 + |x|^p + |v|^p) V_b(x, v) e^{-H(x, v)}.$$

Moreover, since $\nabla U_0(x) \cdot x \geq \kappa|x|^2$ outside of a compact set, there exist $C > 0$ such that $U_0(x) \geq \kappa/2|x|^2 - C$. This implies that for any $b < b'$ and any $p \in \mathbb{N}^*$,

$$(1 + |x|^p + |v|^p) V_b(x, v) \leq C V_{b'}.$$

Hence,

$$\begin{aligned} \int_{\mathbb{R}^{2d}} \partial^\alpha h \partial^\alpha \mathcal{L}_J(h) V_b d\mu &\leq C \int_{\mathbb{R}^{2d}} |\mathcal{L}_J(h)| \left(\sum_{\alpha \leq \nu \leq 2\alpha} |\partial^\nu h| \right) V_{2b} d\mu \\ &\leq C \sqrt{N_{2|\alpha|}(h) \int_{\mathbb{R}^{2d}} |\mathcal{L}_J(h)|^2 V_{4b} d\mu}. \end{aligned}$$

Using the assumptions on the function Ψ , we have that

$$\Psi \left(\pm \frac{\partial_i U_1}{2} (v_i - V_i^i) \right) \leq C(1 + |v_i - V_i^i|).$$

Recall that we also have shown in the proof of Proposition 3.2 that

$$\mathbb{E}[V_b(x, V^i)] \leq C V_b,$$

and that, for any $b < b'$,

$$|v|^2 V_b \leq C V_{b'}.$$

Therefore, using the reversibility of the law of V^i with respect to the standard Gaussian measure,

$$\begin{aligned} \int |\mathcal{L}_J h|^2 V_b d\mu &\leq C \sum_{i=1}^d \int \mathbb{E} \left[(h(x, V^i) - h)^2 \Psi^2 \left(\frac{\partial_i U_1}{2} (v_i - V_i^i) \right) \right] V_b d\mu \\ &\leq C \sum_{i=1}^d \int \mathbb{E} \left[(h(x, V^i)^2 + h^2) \Psi^2 \left(\frac{\partial_i U_1}{2} (v_i - V_i^i) \right) \right] V_b d\mu \\ &\leq C \sum_{i=1}^d \int h^2 \left(\mathbb{E} \left[V_b(x, V^i) \Psi^2 \left(\frac{\partial_i U_1}{2} (V_i^i - v_i) \right) \right] + \mathbb{E} \left[\Psi^2 \left(\frac{\partial_i U_1}{2} (v_i - V_i^i) \right) \right] V_b \right) d\mu \\ &\leq C \sum_{i=1}^d \int h^2 (\mathbb{E}[V_b(x, V^i)(1 + |v_i - V_i^i|^2)] + (1 + |v_i|^2) V_b) d\mu \\ &\leq C \int h^2 V_{b'} d\mu. \end{aligned}$$

Finally, notice that

$$\begin{aligned} \partial_t \int f_t^2 V_b d\mu &= 2 \int f_t \mathcal{L} f_t V_b d\mu = -2 \int \Gamma(f_t) V_b d\mu + \int \mathcal{L}(f_t^2) V_b d\mu \\ &\leq \int f_t^2 \mathcal{L}^* V_b d\mu \leq -\eta \int f_t^2 V_b d\mu + C \int f_t^2 d\mu \\ &\leq -\eta \int f_t^2 V_b d\mu + C e^{-\rho t}. \end{aligned}$$

Therefore,

$$\int f_t^2 V_b d\mu \leq C e^{-\rho t}.$$

Using Theorem 3.1, we then have that for b small enough, there exist $C > 0$ and $\rho > 0$, such that

$$\begin{aligned} \int_{\mathbb{R}^{2d}} \partial^\alpha f_t \partial^\alpha \mathcal{L}_J(f_t) V_b d\mu &\leq C \sqrt{\int_{\mathbb{R}^{2d}} N_{2|\alpha|}(f_t) d\mu} \int_{\mathbb{R}^{2d}} f_t^2 V_{5b} d\mu \\ &\leq C e^{-\rho t}, \end{aligned}$$

which concludes the proof. ■

4. Weak error of the numerical scheme

In this section, devoted to the proof of Theorem 2.2, $\mathcal{X} = \mathbb{T}^d$. The structure of the proof is the following. First, in Section 4.1, by a Lyapunov argument, we establish uniform-in-time Gaussian moment bounds for the BJAOAJB chain (this is Proposition 4.1). This is then used in Section 4.2 to provide an upper bound for the weak error of the scheme in a finite-time horizon, leading to Theorem 4.5, which is of interest for itself (the proof essentially relies on Taylor expansions of the semi-groups, the expectation of the local errors being bounded thanks to the Gaussian moments established in Proposition 4.1). Geometric ergodicity for the BJAOAJB chain is established in Section 4.3 (since a Lyapunov function is already available, in order to apply Harris' theorem, it only remains to prove a minorization condition). It is then possible to let time go to infinity in Theorem 4.5 and, following the classical proof of Talay and Tubaro, to conclude the proof of Theorem 2.2, as detailed in Section 4.4. For completeness, the classical proof of Corollary 2.4 is provided in Section 4.5.

4.1. Lyapunov function for the numerical scheme

For all $z = (x, v) \in \mathbb{T}^d \times \mathbb{R}^d$ and $b > 0$, we define the function V_b by

$$V_b(z) = e^{b|v|^2}.$$

In the following, we will denote by $(\bar{Z}_n)_{n \in \mathbb{N}}$ the discrete-time Markov chain with transition kernel Q defined in (2.6), corresponding to the BJAOAJB splitting scheme with time step δ , and by A, B, J and O the kernels of the different steps of the scheme, namely $B = e^{\delta/2 \mathcal{L}_B}, A = e^{\delta/2 \mathcal{L}_A}, J = e^{\delta/2 \mathcal{L}_J}, O = e^{\delta \mathcal{L}_O}$. The goal of this section is to show the following result.

Proposition 4.1. *Let $b < \frac{1}{2}$. There exist $\delta_0, C, C' > 0$ such that for all $\delta \leq \delta_0$,*

$$QV_b(z) \leq (1 - C\delta)V_b(z) + C'\delta.$$

As a consequence, for all $n \in \mathbb{N}$,

$$\mathbb{E}_z[V_b(\bar{Z}_n)] = Q^n V_b(z) \leq e^{-C\delta n} V_b(z) + \frac{C'}{C}.$$

In order to prove this proposition, we will need to bound the terms BV_b, AV_b, OV_b and JV_b . First, since A only acts on positions, $AV_b = V_b$.

Lemma 4.2. *For all $\delta_0, \varepsilon > 0, b \in (0, 1/2)$, there exists $C_\varepsilon > 0$ such that for all $0 < \delta \leq \delta_0$,*

$$BV_b \leq (1 + C_\varepsilon \delta) e^{b(1+\varepsilon\delta)|v|^2}.$$

Proof. For all $z = (x, v) \in \mathbb{T}^d \times \mathbb{R}^d$, using the Young inequality,

$$\begin{aligned} BV_b(z) &= e^{b|v - \frac{\delta}{2}\nabla U_0(x)|^2} = e^{b\left(|v|^2 - \delta\nabla U_0(x) \cdot v + \frac{\delta^2}{4}|\nabla U_0(x)|^2\right)} \\ &\leq e^{b((1+\varepsilon\delta)|v|^2 + \delta(\frac{\delta}{4} + \frac{1}{4\varepsilon})\|\nabla U_0\|_\infty^2)} \leq (1 + C_\varepsilon\delta)e^{b(1+\varepsilon\delta)|v|^2}, \end{aligned}$$

the last inequality being satisfied if δ is small enough. $\blacksquare$

Lemma 4.3. *There exists $\delta_0 > 0$ (independent from b) such that for all $b \in (0, 1/2)$, there exist $C_1, C_2 > 0$ such that for $0 < \delta \leq \delta_0$,*

$$OV_b(z) \leq (1 + C_2\delta)e^{b(1-C_1\delta)|v|^2}.$$

Proof. We start by noticing that for all $a_1, a_2 \in \mathbb{R}$ and $\xi \sim \mathcal{N}(0, 1)$, the random variable $\exp(a_1\xi^2 + a_2\xi)$ is integrable if and only if $a_1 < 1/2$. In that case, by recognizing the density of $\mathcal{N}(a_2/(1 - 2a_1), 1/(1 - 2a_1))$,

$$\begin{aligned} \mathbb{E}[\exp(a_1\xi^2 + a_2\xi)] &= \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} \exp(a_1x^2 + a_2x - x^2/2) dx \\ &= \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} \exp\left(\frac{-(1 - 2a_1)}{2} \left(x - \frac{a_2}{1 - 2a_1}\right)^2 + \frac{a_2^2}{2(1 - 2a_1)}\right) dx \\ &= \frac{1}{\sqrt{1 - 2a_1}} \exp\left(\frac{a_2^2}{2(1 - 2a_1)}\right). \end{aligned}$$

Using this remark,

$$\begin{aligned} OV_b(z) &= \mathbb{E}\left[\exp\left(b|e^{-\gamma\delta}v + \sqrt{1 - e^{-2\gamma\delta}}\xi|^2\right)\right] \\ &= \exp\left(b e^{-2\gamma\delta}|v|^2\right) \mathbb{E}\left[\exp\left(2be^{-\gamma\delta}\sqrt{1 - e^{-2\gamma\delta}}v \cdot \xi + b(1 - e^{-2\gamma\delta})|\xi|^2\right)\right] \\ &= \frac{1}{(1 - 2b(1 - e^{-2\gamma\delta}))^{d/2}} \exp\left(b e^{-2\gamma\delta}|v|^2 + \frac{2b^2 e^{-2\gamma\delta}(1 - e^{-2\gamma\delta})|v|^2}{1 - 2b(1 - e^{-2\gamma\delta})}\right) \\ &= \frac{1}{(1 - 2b(1 - e^{-2\gamma\delta}))^{d/2}} \exp\left(b|v|^2 \frac{e^{-2\gamma\delta}}{1 - 2b(1 - e^{-2\gamma\delta})}\right). \end{aligned}$$

Then, there exist $\delta_0 > 0$ and constants $C_1, C_2 > 0$ such that for all $0 < \delta \leq \delta_0$,

$$\frac{e^{-2\gamma\delta}}{1 - 2b(1 - e^{-2\gamma\delta})} \leq 1 - C_1\delta,$$

and

$$\frac{1}{(1 - 2b(1 - e^{-2\gamma\delta}))^{d/2}} \leq 1 + C_2\delta.$$

Hence

$$OV_b(z) \leq (1 + C_2\delta)e^{b(1-C_1\delta)|v|^2}.$$

Notice that δ_0 can in fact be chosen independently from b . On one side,

$$\frac{1}{(1 - 2b(1 - e^{-2\gamma\delta}))^{d/2}} \leq \frac{1}{(1 - 2\gamma\delta)^{d/2}} \leq 1 + d\gamma\delta + O(\delta^2),$$

and on the other side, the Taylor–Lagrange formula shows that there exists $0 < c < 2\gamma\delta$ such that

$$\frac{e^{-2\gamma\delta}}{1 - 2b(1 - e^{-2\gamma\delta})} = 1 - 2\gamma(1 - 2b)\delta + \left(4b^2 - 3b + \frac{1}{2}\right)c^2.$$

If $1/4 < b < 1/2$, $4b^2 - 3b + \frac{1}{2} < 0$ so we directly have

$$\frac{e^{-2\gamma\delta}}{1 - 2b(1 - e^{-2\gamma\delta})} \leq 1 - 2\gamma(1 - 2b)\delta,$$

without any condition on δ . Now, if $0 < b < 1/4$, $4b^2 - 3b + \frac{1}{2} < 1/2$, and

$$\frac{e^{-2\gamma\delta}}{1 - 2b(1 - e^{-2\gamma\delta})} \leq 1 - \gamma\delta + 2\gamma^2\delta^2,$$

and δ_0 can be chosen independently from b . ■

Lemma 4.4. *For all $0 < b_1 < b_2 < 1/2$, there exist $\delta_0 > 0$ and $K > 0$ such that for all $\delta \leq \delta_0$ and all $b \in [b_1, b_2]$,*

$$JV_b(z) \leq V_b(z)(1 + \delta K) + \delta K.$$

Proof. Let $0 < b_1 < b_2 < 1/2$ and $b \in [b_1, b_2]$. In the proof of this lemma, constants C can depend on b_1 and b_2 , but not on b . We will denote $Z_k = (X_k, V_k)$ the jump process after k jumps, and Z_δ the jump process after a time δ . We start by expanding $JV_b(z)$ according to the number of jumps:

$$\mathbb{E}_z[V_b(Z_\delta)] = \sum_{k=0}^{\infty} \mathbb{E}_z[V_b(Z_\delta) \mathbb{1}_{\{k \text{ jumps}\}}].$$

First,

$$\mathbb{E}_z[V_b(Z_\delta) \mathbb{1}_{\{0 \text{ jumps}\}}] = V_b(z) \mathbb{P}(0 \text{ jumps}) \leq V_b(z).$$

Then, let us show by induction on k that there exists a constant C such that for all $k \geq 1$,

$$\mathbb{E}_z[V_b(Z_\delta) \mathbb{1}_{\{k \text{ jumps}\}}] \leq C^k \delta^k (1 + V_b(z)). \quad (4.1)$$

In order to prove this inequality, we will use some useful bounds. First, recall that there exist a constant C such that for all $z \in \mathbb{T}^d \times \mathbb{R}^d$,

$$\lambda(z) \leq C(1 + |v|). \quad (4.2)$$

As we saw in the proof of Lemma 4.3, if $a_1 < 1/2$, then the random variable $\exp(a_1\xi^2 + a_2\xi)$ is integrable and

$$\mathbb{E}[\exp(a_1\xi^2 + a_2\xi)] = \frac{1}{\sqrt{1 - 2a_1}} \exp\left(\frac{a_2^2}{2(1 - 2a_1)}\right).$$

More generally, by recognizing the density of a Gaussian random variable with mean $a_2/(1 - 2a_1)$ and variance $1/(1 - 2a_1)$,

$$\mathbb{E}[|\xi|^n \exp(a_1\xi^2 + a_2\xi)] = \frac{1}{\sqrt{1 - 2a_1}} \exp\left(\frac{a_2^2}{2(1 - 2a_1)}\right) \mathbb{E}[|X|^n],$$

where $X \sim \mathcal{N}(a_2/(1 - 2a_1), 1/(1 - 2a_1))$. As a consequence, the random variable $V_b(Z_1)$ is integrable if and only if $b(1 - \rho^2) < 1/2$, which is always true since $b \in (0, 1/2)$.

$$\begin{aligned} \mathbb{E}_z[V_b(Z_1)] &\leq \sum_{i=1}^d \mathbb{E}_z \left[\exp(b|\rho v_i + \sqrt{1 - \rho^2}\xi|^2) \Psi \left(\frac{\partial_i U_1}{2} \left((1 - \rho)v_i + \sqrt{1 - \rho^2}\xi \right) \right) \right] \\ &\leq C \sum_{i=1}^d \mathbb{E}_z \left[\exp(b|\rho v_i + \sqrt{1 - \rho^2}\xi|^2) (1 + |(1 - \rho)v_i + \sqrt{1 - \rho^2}\xi|) \right] \\ &\leq C \sum_{i=1}^d \exp(b\rho^2|v_i|^2) \mathbb{E}[\exp(b(1 - \rho^2)\xi^2 + 2b\rho\sqrt{1 - \rho^2}v_i\xi)(1 + 2|v_i| + |\xi|)], \end{aligned}$$

so that

$$\mathbb{E}_z[V_b(Z_1)] \leq C(1 + |v|^2) e^{b\alpha|v|^2}, \quad (4.3)$$

where we denoted $\alpha = \frac{\rho^2}{1-2b(1-\rho^2)}$. Notice that

$$\alpha < 1 \iff \rho^2 < 1 - 2b(1 - \rho^2) \iff b < 1/2,$$

and that $b\alpha = b\rho^2/(1 - 2b(1 - \rho^2))$, seen as a function of b , is continuous and strictly positive on $(0, 1/2)$, so is bounded and reaches its extrema on $[b_1, b_2]$. Since $|v|^2 e^{b\alpha|v|^2} = o(V_b)$ and $|v| = o(V_b)$ when v goes to infinity, let

$$K = \max(R^2, R) \sup_{b \in [b_1, b_2]} e^{b\alpha R^2},$$

where $R > 0$ satisfies

$$|v| \geq R \implies \max(|v|, |v|^2) \leq \inf_{b \in [b_1, b_2]} e^{b(1-\alpha)|v|^2},$$

so that

$$|v|^2 e^{b\alpha|v|^2} \leq V_b(z) + K, \quad (4.4)$$

and

$$|v| \leq V_b(z) + K. \quad (4.5)$$

Furthermore, we also have that

$$\int_0^\delta (\delta - t) \lambda(z) e^{-\lambda(z)t} dt = \frac{\lambda(z)\delta + e^{-\lambda(z)\delta} - 1}{\lambda(z)} \leq \frac{\lambda(z)\delta^2}{2} \leq C\delta^2(1 + |v|). \quad (4.6)$$

We start by showing (4.1) in the case $k = 1$. We suppose that $\lambda(z) > 0$ (otherwise there is no jump and the result is straightforward). By conditioning with respect to the only jump time,

$$\begin{aligned} \mathbb{E}_z[V_b(Z_\delta) \mathbb{1}_{\{1 \text{ jump}\}}] &= \int_0^\delta \lambda(z) e^{-\lambda(z)t} \int_{\mathbb{T}^d \times \mathbb{R}^d} q(z, d\tilde{z}) V_b(\tilde{z}) e^{-(\delta-t)\lambda(\tilde{z})} dt \\ &\leq \int_0^\delta \lambda(z) e^{-\lambda(z)t} \int_{\mathbb{T}^d \times \mathbb{R}^d} q(z, d\tilde{z}) V_b(\tilde{z}) dt \\ &\leq \mathbb{E}_z[V_b(Z_1)] \lambda(z) \delta. \end{aligned}$$

Then, using (4.3), (4.4) and (4.5), we have

$$\mathbb{E}_z[V_b(Z_\delta) \mathbb{1}_{\{1 \text{ jump}\}}] \leq C\delta e^{b\alpha|v|^2} (1 + |v|^2) \leq C\delta (V_b + K) \leq C\delta(1 + V_b),$$

which proves the inequality for $k = 1$. Now let $k \geq 1$ and assume that (4.1) holds for k . By conditioning with respect to the first jump time,

$$\begin{aligned} \mathbb{E}_z[V_b(Z_\delta) \mathbb{1}_{\{k+1 \text{ jumps}\}}] &= \int_0^\delta \lambda(z) e^{-\lambda(z)t} \int_{\mathbb{T}^d \times \mathbb{R}^d} q(z, d\tilde{z}) \mathbb{E}_{\tilde{z}}[V_b(Z_{\delta-t}) \mathbb{1}_{\{k \text{ jumps}\}}] dt \\ &\leq C^k \int_0^\delta (\delta - t)^k \lambda(z) e^{-\lambda(z)t} \int_{\mathbb{T}^d \times \mathbb{R}^d} q(z, d\tilde{z}) (V_b(\tilde{z}) + 1) dt \\ &\leq \delta^{k-1} C^k (\mathbb{E}_z[V_b(Z_1)] + 1) \int_0^\delta (\delta - t) \lambda(z) e^{-\lambda(z)t} dt. \end{aligned}$$

Then, using (4.6),

$$\mathbb{E}_z[V_b(Z_\delta) \mathbb{1}_{\{k+1 \text{ jumps}\}}] \leq \delta^{k+1} C^{k+1} (e^{b\alpha|v|^2} + 1) (1 + |v|^2) \leq \delta^{k+1} C^{k+1} (1 + V_b),$$

which proves the inequality for $k + 1$ and ends the induction. By choosing $\delta < 1/C$, we can sum the inequalities (4.1):

$$\begin{aligned} JV_b(z) = \mathbb{E}_z[V_b(Z_\delta)] &= \sum_{k=0}^{\infty} \mathbb{E}_z[V_b(Z_\delta) \mathbb{1}_{\{k \text{ jumps}\}}] \leq V_b(z) + (1 + V_b(z)) \sum_{k=1}^{\infty} (C\delta)^k \\ &\leq V_b(z) + \delta \frac{C}{1 - C\delta} (1 + V_b(z)), \end{aligned}$$

so there exist δ_0 and $K > 0$ such that for all $\delta < \delta_0$,

$$JV_b \leq V_b(1 + \delta K) + \delta K.$$

■

We are now able to prove Proposition 4.1.

Proof of Proposition 4.1. Let $b \in (0, 1/2)$ and $\varepsilon > 0$, whose value will be specified later. Thanks to Lemma 4.2, for all $\delta_1 > 0$, there exist $C_\varepsilon > 0$ such that for all $\delta < \delta_1$,

$$BV_b \leq (1 + C_\varepsilon \delta) e^{b(1+\varepsilon\delta)|v|^2}.$$

Now let $\delta_2 > 0$ such that $b_2 = b(1 + \varepsilon\delta_2) < 1/2$. Notice that for all $0 < \delta \leq \delta_2$, $b < b' := b(1 + \varepsilon\delta) < b_2$. Thanks to Lemma 4.4, there exist $\delta_3 > 0$ and $K > 0$ (whose value is uniform over $[b, b_2]$) such that for all $\delta < \delta_3$,

$$JV_{b'} \leq V_{b'}(1 + \delta K) + \delta K.$$

Therefore, for all $\delta < \min(\delta_1, \delta_2, \delta_3)$,

$$JBV_b(z) \leq (1 + C_\varepsilon \delta) JV_{b'}(z) \leq (1 + C_\varepsilon \delta) V_{b'}(z)(1 + \delta K) + \delta(1 + C_\varepsilon \delta)K,$$

and

$$AJBV_b(z) \leq (1 + C_\varepsilon \delta) AV_{b'}(z)(1 + \delta K) + \delta(1 + C_\varepsilon \delta)K = (1 + C_\varepsilon \delta) V_{b'}(z)(1 + \delta K) + \delta(1 + C_\varepsilon \delta)K.$$

Lemma 4.3 gives a $\delta_4 > 0$ (whose value is independent from b or b') such that for $\delta < \delta_4$,

$$\begin{aligned} OV_{b'} &\leq (1 + C_2 \delta) \exp(b'(1 - C_1 \delta)|v|^2) \\ &\leq (1 + C_2 \delta) \exp(b(1 + \varepsilon\delta)(1 - C_1 \delta)|v|^2) \\ &\leq (1 + C_2 \delta) \exp(b(1 - (C_1 - \varepsilon)\delta)|v|^2). \end{aligned}$$

We would like to have $\varepsilon < C_1$ (In fact, because of the two applications of the kernel B , we will need to have $\varepsilon < C_1/2$). When looking at the proof of the Lemma 4.3, the value of C_1 can depend on b' (and therefore on ε) if $1/4 < b' < 1/2$, in which case $C_1 = 2\gamma(1 - 2b') = 2\gamma(1 - 2b(1 + \varepsilon\delta))$. The constant ε can always be chosen smaller than half of this value if $\varepsilon < \frac{\gamma(1-2b)}{1+4\gamma b\delta_1}$. In the other cases, C_1 does not depend on b' , so we can also choose $\varepsilon < C_1/2$. We now denote $a = C_1 - \varepsilon$. For all $\delta < \min(\delta_1, \delta_2, \delta_3, \delta_4)$,

$$\begin{aligned} OAJBV_b(z) &\leq (1 + C_\varepsilon \delta) OV_{b'}(z)(1 + \delta K) + \delta(1 + C_\varepsilon \delta)K \\ &\leq (1 + C_\varepsilon \delta)(1 + C_2 \delta) \exp(b(1 - a\delta)|v|^2)(1 + \delta K) + \delta(1 + C_\varepsilon \delta)K, \end{aligned}$$

and

$$AOAJBV_b(z) \leq (1 + C_\varepsilon \delta)(1 + C_2 \delta) \exp(b(1 - a\delta)|v|^2)(1 + \delta K) + \delta(1 + C_\varepsilon \delta)K.$$

From there, in order to re-apply the bounds on the jump kernel, we can let δ_5 such that $b_3 = b(1 - a\delta_5) > 0$, which gives a $\delta_6 > 0$ such that for $0 < \delta \leq \min(\delta_1, \delta_2, \delta_3, \delta_4, \delta_5, \delta_6)$, corresponding to a $b_3 < b(1 - a\delta) < b$,

$$J\left(e^{b(1-a\delta)|v|^2}\right) \leq e^{b(1-a\delta)|v|^2}(1 + \delta K') + \delta K',$$

which yields

$$JAOAJBV_b(z) \leq (1 + C_\varepsilon \delta)(1 + C_2 \delta)(1 + \delta K)((1 + \delta K') \exp(b(1 - a\delta)|v|^2) + \delta K') + \delta(1 + C_\varepsilon \delta)K.$$

Finally, for the last application of the kernel B , we need to have

$$(1 + \varepsilon\delta)(1 - a\delta) = 1 - (a - \varepsilon)\delta - a\varepsilon\delta^2 < 1,$$

which is possible because $a = C_1 - \varepsilon$ and that $\varepsilon < C_1/2$. We then denote $c = a - \varepsilon$. Finally, let $\delta_0 = \min(\delta_1, \delta_2, \delta_3, \delta_4, \delta_5, \delta_6)$. There exists $C' > 0$ such that for all $0 < \delta < \delta_0$,

$$\begin{aligned} BJA O A J B V_b(z) &\leq (1 + C_\varepsilon \delta)(1 + C_2 \delta)(1 + \delta K)((1 + \delta K')B \exp(b(1 - a\delta)|v|^2) + \delta K') + \delta(1 + C_\varepsilon \delta)K \\ &\leq (1 + C_\varepsilon \delta)(1 + C_2 \delta)(1 + \delta K)((1 + \delta K')(1 + C_\varepsilon \delta) \exp(b(1 - c\delta)|v|^2) + \delta K') + \delta(1 + C_\varepsilon \delta)K \\ &\leq V_b(z)(1 + C'\delta)^5 \exp(-bc\delta|v|^2) + \delta C', \end{aligned}$$

and hence

$$BJA O A J B V_b(z) \leq V_b(z) \exp(5C'\delta - bc\delta|v|^2) + \delta C'. \quad (4.7)$$

Now, writing $R_0^2 = 10C'/(bc)$, distinguishing cases, we see that

$$\begin{aligned} e^{5C'\delta - bc\delta|v|^2} &\leq e^{-5C'\delta} \mathbb{1}_{|v| \geq R_0} + e^{5C'\delta} \mathbb{1}_{|v| < R_0} \leq (1 - 5C'e^{-5C'\delta_0} \delta) \mathbb{1}_{|v| \geq R_0} + (1 + 5C'e^{5C'\delta_0} \delta) \mathbb{1}_{|v| < R_0} \\ &\leq 1 - c'\delta + C''\delta \mathbb{1}_{|v| < R_0}, \end{aligned}$$

for some $c', C'' > 0$ independent from δ . Plugging this in (4.7) we obtain that there exist $K, K' > 0$ independent from δ such that for all z ,

$$QV_b(z) \leq (1 - K\delta)V_b(z) + K'\delta,$$

which by induction yields

$$\begin{aligned} Q^n V_b(z) &\leq (1 - K\delta)^n V_b(z) + K'\delta \sum_{k=0}^{n-1} (1 - K\delta)^k \\ &\leq e^{-K\delta n} V_b(z) + \frac{K'}{K}, \end{aligned}$$

concluding the proof. ■

4.2. Finite-time error expansion of the numerical scheme

In this section, $Z_t = (X_t, V_t)$ denotes the continuous-time process with generator $\mathcal{L}$ and with Markov semi-group P_t , $(\bar{Z}_n)_{n \in \mathbb{N}}$ is the discrete-time process corresponding to the BJA O A J B splitting-scheme with time-step δ . We will denote

$$\bar{P}_t = e^{\frac{t}{2}B} e^{\frac{t}{2}J} e^{\frac{t}{2}A} e^{tO} e^{\frac{t}{2}A} e^{\frac{t}{2}J} e^{\frac{t}{2}B},$$

so that $\bar{P}_\delta f = \mathbb{E}[f(\bar{Z}_1)]$. The goal of this section is to show that the finite-time weak error of the numerical scheme is of order 2 in the time-step.

Theorem 4.5. *Let $\delta_0 > 0$ such that Proposition 4.1 holds, $f \in \mathcal{A}$, $z \in \mathbb{T}^d \times \mathbb{R}^d$. There exists a function $t \mapsto C(t, z) > 0$ from $\mathbb{R}_+$ to $\mathbb{R}_+$ such that for all $n \in \mathbb{N}$ and $\delta \in (0, \delta_0]$,*

$$|\mathbb{E}_z[f(Z_{n\delta})] - \mathbb{E}_z[f(\bar{Z}_n)]| \leq C(n\delta, z)\delta^2.$$

More precisely, by denoting S_3 the operator

$$\begin{aligned} S_3 &= \frac{7}{144} (2[L_O, [L_O, L_A]] - [L_A, [L_A, L_O]] \\ &\quad + 2[L_A + L_O, [L_A + L_O, L_J]] - [L_J, [L_J, L_A + L_O]] \\ &\quad + 2[L_A + L_J + L_O, [L_A + L_J + L_O, L_B]] - [L_B, [L_B, L_A + L_J + L_O]]), \end{aligned}$$

then

$$\mathbb{E}_z[f(Z_{n\delta})] - \mathbb{E}_z[f(\overline{Z}_n)] = \delta^3 \sum_{k=0}^{n-1} \mathbb{E}[S_3 P_{(n-k-1)\delta} f(\overline{Z}_k)] + R_{n,z},$$

with

$$\left| \sum_{k=0}^{n-1} \mathbb{E}[S_3 P_{(n-k-1)\delta} f(\overline{Z}_k)] \right| \leq \frac{C}{\delta} \left(V_b(z) e^{-qn\delta} + 1 \right),$$

and

$$|R_{n,z}| \leq C\delta^3 \left(V_b(z) e^{-qn\delta} + 1 \right).$$

In order to prove this theorem, we will need the following lemmas. Recall that given $V : \mathbb{T}^d \times \mathbb{R}^d \rightarrow [1, \infty)$, for $f : \mathbb{T}^d \times \mathbb{R}^d \rightarrow \mathbb{R}$ we write $\|f\|_V = \|f/V\|_\infty$. Let, for all $b \in (0, 1/2)$, $k \in \mathbb{N}^*$ and $f \in \mathcal{A}$,

$$\|f\|_{b,k} = \sum_{|\alpha| \leq k} \|\partial^\alpha f\|_{V_b}.$$

When $\mathcal{X} = \mathbb{T}^d$, the set $\mathcal{A}$ reads

$$\begin{aligned} \mathcal{A} &= \left\{ f \in \mathcal{C}^\infty(\mathbb{T}^d \times \mathbb{R}^d, \mathbb{R}) \mid \forall \alpha \in \mathbb{N}^{2d}, \exists C > 0, b \in (0, 1/2), |\partial^\alpha f| \leq C V_b \right\} \\ &= \left\{ f \in \mathcal{C}^\infty(\mathbb{T}^d \times \mathbb{R}^d, \mathbb{R}) \mid \forall k \in \mathbb{N}, \exists b \in (0, 1/2), \|f\|_{b,k} < \infty \right\}. \end{aligned}$$

Lemma 4.6. *For all $t_0 > 0$, $f \in \mathcal{A}$, $k \geq 2$, $b \in (0, 1/2)$ such that $\|f\|_{b,k} < \infty$ and $b' \in (b, 1/2)$, there exists $C > 0$ such that for $t \in [0, t_0]$ and $T \in \{\mathcal{L}_A, \mathcal{L}_B, \mathcal{L}_J, \mathcal{L}_O, e^{t\mathcal{L}_A}, e^{t\mathcal{L}_B}, e^{t\mathcal{L}_J}, e^{t\mathcal{L}_O}\}$,*

$$\|Tf\|_{b',k'} \leq C\|f\|_{b,k},$$

with $k' = k, k-1$ or $k-2$ (depending on T). In particular, the operator T leaves stable the set $\mathcal{A}$.

In the proof of this lemma, we will need the formula for higher-order derivatives of composite functions [33]:

Proposition 4.7 (Higher derivatives of composite functions). *Let $g : \mathbb{R}^n \rightarrow \mathbb{R}^l$ and $f : \mathbb{R}^l \rightarrow \mathbb{R}$ be smooth functions, and $\alpha = (\alpha_1, \dots, \alpha_n) \in \mathbb{N}^n$ be a multi-index. α is said to be decomposed into $s \in \mathbb{N}^*$ (pairwise distinct) parts $p_1, \dots, p_s$ in $\mathbb{N}^n$ with multiplicities $m_1, \dots, m_s$ in $\mathbb{N}^l$ (the p 's and the m 's are multi-indexes) if*

$$\alpha = |m_1|p_1 + \dots + |m_s|p_s$$

holds and all parts are different. We define the total multiplicity by

$$m = m_1 + \dots + m_s$$

and $\mathcal{D}_\alpha$ as the set of all such decompositions of α . Then,

$$\partial^\alpha (f \circ g)(x) = \alpha! \sum_{(s,p,m) \in \mathcal{D}} (\partial^m f)(g(x)) \prod_{k=1}^s \frac{1}{m_k!} \left[\frac{1}{p_k!} \partial^{p_k} g(x) \right]^{m_k}$$

where $\alpha! = \prod_j \alpha_j!$ and $x^\alpha = \prod_j x_j^{\alpha_j}$. The term $\left[\frac{1}{p_k!} \partial^{p_k} g(x) \right]^{m_k}$ is thus the short form for

$$\prod_{i=1}^l \left[\frac{1}{p_{k,1}! \dots p_{k,n}!} \partial^{p_k} g_i(x) \right]^{m_{k,i}}$$

where $g = (g_1, \dots, g_l)$.

Proof of Lemma 4.6. Let $f \in \mathcal{A}$, $k \in \mathbb{N}^*$, $b \in (0, 1/2)$ such that $\|f\|_{b,k} < \infty$ and $b' \in (b, 1/2)$. As a first step, we prove the result for t_0 small enough.

First, for all $\alpha \in \mathbb{N}^{2d}$ with $|\alpha| \leq k-1$,

$$|\partial^\alpha \mathcal{L}_B f| = |\partial^\alpha (\nabla U_0 \cdot \nabla_v f)| = \sum_{i=1}^d \sum_{\nu \leq \alpha} |\partial^\nu \partial_i U_0 \partial^{\alpha-\nu} \partial_{v_i} f| \leq C V_b(z) \|f\|_{b,k},$$

thus

$$\|\mathcal{L}_B f\|_{b,k-1} \leq C \|f\|_{b,k}.$$

Similarly, for $|\alpha| \leq k-1$,

$$|\partial^\alpha \mathcal{L}_A f| = |\partial^\alpha (v \cdot \partial_x f)| \leq |v| \sum_{i=1}^d \sum_{\nu \leq \alpha} \partial^\nu \partial_{x_i} f \leq C |v| V_b(z) \|f\|_{b,k} \leq C V_{b'}(z) \|f\|_{b,k},$$

which yields

$$\|\mathcal{L}_A f\|_{b',k-1} \leq C \|f\|_{b,k}.$$

If $|\alpha| \leq k-2$,

$$|\partial^\alpha \mathcal{L}_O f| = |-\gamma \partial^\alpha (v \cdot \nabla_v f) + \gamma \partial^\alpha \Delta_v f| \leq C V_{b'}(z) \|f\|_{b,k},$$

and thus

$$\|\mathcal{L}_O f\|_{b',k-2} \leq C \|f\|_{b,k}.$$

Regarding the jump generator, we saw in the proof of Lemma 3.16 that for all $\nu_1, \nu_2 \in \mathbb{N}^d$,

$$\partial_x^{\nu_1} \partial_v^{\nu_2} (f(x, V^i)) = \rho^{\nu_2, i} (\partial_x^{\nu_1} \partial_v^{\nu_2} f)(x, V^i),$$

and that the only non-vanishing derivatives of $\Psi \left(\frac{\partial_i U_1}{2} (v_i - V_i^i) \right)$ are of the type

$$\partial_x^\nu \partial_{v_i}^n \Psi \left(\frac{\partial_i U_1}{2} (v_i - V_i^i) \right)$$

with

$$\left| \partial_x^\nu \partial_{v_i}^n \Psi \left(\frac{\partial_i U_1}{2} (v_i - V_i^i) \right) \right| \leq C (1 + |v_i - V_i^i|^{|n|+n}).$$

Therefore, if $|\alpha| \leq k$,

$$\begin{aligned} \partial^\alpha \mathcal{L}_J f(z) &= \frac{2}{1-\rho} \sum_i \partial^\alpha \mathbb{E} \left[(f(x, V^i) - f(x, v)) \Psi \left(\frac{\partial_i U_1}{2} (v_i - V_i^i) \right) \right] \\ &= \frac{2}{1-\rho} \sum_i \sum_{\nu \leq \alpha} \mathbb{E} \left[\partial^\nu (f(x, V^i) - f(x, v)) \partial^{\alpha-\nu} \Psi \left(\frac{\partial_i U_1}{2} (v_i - V_i^i) \right) \right] \\ &\leq C \frac{2}{1-\rho} \sum_i \sum_{\nu \leq \alpha} \mathbb{E} \left[\left(\rho^{\nu_{d+i}} \partial^\nu f(x, V^i) - \partial^\nu f(x, v) \right) (1 + |v_i - V_i^i|^{\alpha-\nu}) \right] \\ &\leq C \|f\|_{b,k} \sum_i \sum_{\nu \leq \alpha} \mathbb{E} \left[\left(V_b(x, V^i) + V_b(z) \right) (1 + |v_i - V_i^i|^{\alpha-\nu}) \right] \\ &\leq C \|f\|_{b,k} \left(V_{b'}(z) + \mathbb{E}[(1 + |v|^{\alpha} + |\xi|^{\alpha}) V_b(x, V^i)] \right). \end{aligned}$$

In the proof of Lemma 4.4, we saw that

$$\mathbb{E}_z[\exp(b|\rho v_i + \sqrt{1-\rho^2} \xi|^2)] \leq C V_b(z),$$

and that if $a_1 < 1/2$,

$$\mathbb{E}[|\xi|^n \exp(a_1 \xi^2 + a_2 \xi)] = \frac{1}{\sqrt{1-2a_1}} \exp\left(\frac{a_2^2}{2(1-2a_1)}\right) \mathbb{E}[|X|^n],$$

where $X \sim \mathcal{N}(a_2/(1-2a_1), 1/(1-2a_1)) \sim a_2/(1-2a_1) + (1/\sqrt{1-2a_1})\xi$ with $\xi \sim \mathcal{N}(0, 1)$. Therefore,

$$\mathbb{E}[|\xi|^{\alpha} V_b(x, V^i)] \leq C V_{b'}(z),$$

and thus

$$\partial^\alpha Jf(z) \leq C\|f\|_{b,k}V_{b'}(z),$$

which yields

$$\|\mathcal{L}_J\|_{b',k} \leq C\|f\|_{b,k}.$$

Let us now look at the semi-groups. The derivatives of U being bounded (recall that the position space is a torus in this section), we have for all $t < 1$ and $|\alpha| \leq k$,

$$\begin{aligned} \partial^\alpha e^{tB}f(z) &= \partial^\alpha f(x, v - t\nabla U_0(x)) \\ &= \alpha! \sum_{(s,p,m) \in \mathcal{D}} (\partial^m f)(x, v - t\nabla U_0(x)) \prod_{k=1}^s \frac{1}{m_k!} \left[\frac{1}{p_k!} \partial^{p_k}(x, v - t\nabla U_0(x)) \right]^{m_k} \\ &\leq C\|f\|_{b,k}V_b(x, v - t\nabla U_0(x)) \leq C(1 + C_\varepsilon t)\|f\|_{b,k}V_{b(1+\varepsilon t)}(z), \end{aligned}$$

where the last inequality comes from the result of Lemma 4.2. By choosing ε small enough, we have $b(1 + \varepsilon t) < b'$, hence the result:

$$\|e^{tB}f\|_{b',k} \leq C\|f\|_{b,k}.$$

Similarly, for $t < 1$ and $|\alpha| \leq k$,

$$\begin{aligned} \partial^\alpha e^{tA}f(z) &= \partial^\alpha f(x + tv, v) \\ &= \alpha! \sum_{(s,p,m) \in \mathcal{D}} (\partial^m f)(x + tv, v) \prod_{k=1}^s \frac{1}{m_k!} \left[\frac{1}{p_k!} \partial^{p_k}(x + tv, v) \right]^{m_k} \\ &\leq C\|f\|_{b,k}V_b(z), \end{aligned}$$

which yields

$$\|e^{tA}f\|_{b,k} \leq C\|f\|_{b,k}.$$

Again, for $t < 1$ and $|\alpha| \leq k$,

$$\begin{aligned} \partial^\alpha e^{tO}f(z) &= \mathbb{E}[\partial^\alpha f(x, e^{-\gamma t}v + \sqrt{1 - e^{-2\gamma t}}\xi)] \\ &= \alpha! \sum_{(s,p,m) \in \mathcal{D}} \mathbb{E} \left[(\partial^m f)(x, e^{-\gamma t}v + \sqrt{1 - e^{-2\gamma t}}\xi) \prod_{k=1}^s \frac{1}{m_k!} \left[\frac{1}{p_k!} \partial^{p_k}(x, e^{-\gamma t}v + \sqrt{1 - e^{-2\gamma t}}\xi) \right]^{m_k} \right] \\ &\leq C\|f\|_{b,k}\mathbb{E}[V_b(x, e^{-\gamma t}v + \sqrt{1 - e^{-2\gamma t}}\xi)] \leq C\|f\|_{b,k}(1 + C_2 t)|v|^{|\alpha|}V_{b(1-C_1 t)}(z) \leq C\|f\|_{b,k}V_b(z), \end{aligned}$$

where the second to last inequality comes from the result of Lemma 4.3, hence

$$\|e^{tO}f\|_{b,k} \leq C\|f\|_{b,k}.$$

Finally, by denoting (Z_t) the jump process after a time t and Z_n the jump process after n jumps,

$$\partial^\alpha e^{tJ}f(z) = \sum_{n=0}^{\infty} \partial^\alpha \mathbb{E}[f(Z_t)\mathbb{1}_{\{n \text{ jumps}\}}].$$

Recall that we showed in the proof of Lemma 4.4 that for any $b < 1/2$, there exist $C > 0$ such that

$$\mathbb{E}_z[V_b(Z_\delta)\mathbb{1}_{\{n \text{ jumps}\}}] \leq C^n \delta^n (1 + V_b(z)).$$

As we mentioned earlier,

$$\left| \partial_x^\nu \partial_{v_i^n} \Psi \left(\frac{\partial_i U_1}{2}(v_i - V_i^i) \right) \right| \leq C \left(1 + |v_i - V_i^i|^{\nu|+n} \right),$$

therefore

$$\partial^\alpha \lambda(z) \leq C(1 + |v|^{|\alpha|}).$$

As a consequence, for $t < 1$ and $|\alpha| \leq k$,

$$\begin{aligned}
 E_z[\partial^\alpha f(Z_t) \mathbb{1}_{\{0 \text{ jumps}\}}] &= \partial^\alpha (f(z) e^{-\lambda(z)t}) \\
 &= \sum_{\nu \leq \alpha} \binom{\alpha}{\nu} \partial^\nu (e^{-\lambda(z)t}) \partial^{\alpha-\nu} f(z) \\
 &\leq C \|f\|_{b,k} V_b(z) t e^{-\lambda(z)t} \sum_{\nu \leq \alpha} \alpha! \sum_{(s,p,m) \in \mathcal{D}} \prod_{k=1}^s \frac{1}{m_k!} \left[\frac{1}{p_k!} \partial^{p_k} \lambda(z) \right]^{m_k} \\
 &\leq C \|f\|_{b,k} t (1 + |v|^{|\alpha|}) V_b(z) \\
 &\leq C \|f\|_{b,k} t V_{b'}(z).
 \end{aligned}$$

Then, notice that for $k \geq 1$,

$$\begin{aligned}
 \partial^\alpha \mathbb{E}_z[f(Z_\delta) \mathbb{1}_{\{n \text{ jumps}\}}] &= \partial^\alpha \int_0^t \int_{\mathbb{T}^d \times \mathbb{R}^d} \lambda(z) e^{-\lambda(z)s} q(z, \tilde{z}) \mathbb{E}_{\tilde{z}}[f(Z_{t-s}) \mathbb{1}_{\{n-1 \text{ jumps}\}}] d\tilde{z} ds \\
 &\leq C \|f\|_{b,k} \int_0^t \int_{\mathbb{T}^d \times \mathbb{R}^d} \partial^\alpha (\lambda(z) e^{-\lambda(z)s} q(z, \tilde{z})) \mathbb{E}_{\tilde{z}}[V_b(Z_{t-s}) \mathbb{1}_{\{n-1 \text{ jumps}\}}] d\tilde{z} ds \\
 &\leq C^k \|f\|_{b,k} t^{k-1} (1 + V_b(z)) \int_0^t \int_{\mathbb{T}^d \times \mathbb{R}^d} \partial^\alpha (\lambda(z) e^{-\lambda(z)s} q(z, \tilde{z})) d\tilde{z} ds.
 \end{aligned}$$

Let us remind that by denoting $z = (x, v)$ and $\tilde{z} = (\tilde{x}, \tilde{v})$,

$$\lambda(z) q(z, \tilde{z}) \propto \sum_{i=1}^d \Psi \left(\frac{\partial_i U_1}{2} (v_i - \tilde{v}_i) \right) \exp \left(-\frac{(\tilde{v}_i - \rho v_i)^2}{2(1 - \rho^2)} \right),$$

and hence

$$\partial^\alpha (\lambda(z) q(z, \tilde{z})) \propto \sum_{i=1}^d \sum_{\nu \leq \alpha} \partial^\nu \Psi \left(\frac{\partial_i U_1}{2} (v_i - \tilde{v}_i) \right) \partial^{\alpha-\nu} \exp \left(-\frac{(\tilde{v}_i - \rho v_i)^2}{2(1 - \rho^2)} \right).$$

On one hand,

$$\partial^\nu \Psi \left(\frac{\partial_i U_1}{2} (v_i - \tilde{v}_i) \right) \leq C(1 + |v_i|^{|\nu|} + |\tilde{v}_i|^{|\nu|}),$$

and on the other hand

$$\partial^\nu \exp \left(-\frac{(\tilde{v}_i - \rho v_i)^2}{2(1 - \rho^2)} \right) \leq C(|v_i| + |\tilde{v}_i|) \exp \left(-\frac{(\tilde{v}_i - \rho v_i)^2}{2(1 - \rho^2)} \right),$$

hence there exist $n_1, n_2 \in \mathbb{N}$ such that

$$\partial^\alpha (\lambda(z) q(z, \tilde{z})) \leq C(1 + |v|^{n_1} + |\tilde{v}|^{n_2}) \exp \left(-\frac{(\tilde{v}_i - \rho v_i)^2}{2(1 - \rho^2)} \right).$$

Then, since we already showed that

$$\partial^\alpha e^{-\lambda(z)t} \leq C t (1 + |v|^{|\alpha|}),$$

by combining the obtained bounds we get

$$\partial^\alpha (\lambda(z) e^{-\lambda(z)s} q(z, \tilde{z})) \leq C s (1 + |v|^{n_1} + |\tilde{v}|^{n_2}) \exp \left(-\frac{(\tilde{v}_i - \rho v_i)^2}{2(1 - \rho^2)} \right).$$

This last expression is integrable with respect to $\tilde{z}$ with

$$\int |\tilde{v}_i|^m \exp \left(-\frac{(\tilde{v}_i - \rho v_i)^2}{2(1 - \rho^2)} \right) d\tilde{v}_i \leq C(1 + |v|^m),$$

therefore

$$\partial^\alpha \mathbb{E}_z[f(Z_\delta) \mathbb{1}_{\{n \text{ jumps}\}}] \leq \|f\|_{b,k} C^m t^n V_{b'}(z).$$

If $t < 1/C$, we can sum those inequalities over n to get

$$\partial^\alpha e^{tJ} f(z) \leq C \|f\|_{b,k} V_{b'}(z),$$

thus,

$$\|e^{tJ} f\|_{b',k} \leq C \|f\|_{b,k}.$$

This concludes the proof of when $t_0 < 1 \wedge 1/C$.

To get the result with any $t_0 > 0$, we simply use that, for an operator $\mathcal{L}$, $e^{t\mathcal{L}} = (e^{\frac{1}{n}\mathcal{L}})^n$, and thus it is sufficient to iterate the result with n taken such that $t_0/n < 1 \wedge 1/C$. ■

Lemma 4.8. *For all $f \in \mathcal{A}$, all $t \geq 0$ and all $s \leq t$, there exist $0 < b < 1/2$, $C > 0$ and $q > 0$ such that*

$$|S_3 P_t f(z)| + |L^4 P_t f(z)| + |\partial_s^4 \bar{P}_s P_t f(z)| \leq C e^{-qt} V_b(z).$$

Proof. Let $f \in \mathcal{A}$ and $t \geq 0$. Theorem 2.1 implies that $f_t = P_t f - \mu f \in \mathcal{A}$: for all $k \in \mathbb{N}^*$, there exist $C, q > 0$ and $b \in (0, 1/2)$ such that for all $t > 0$,

$$\|f_t\|_{b,k} \leq C e^{-qt}.$$

On one side, S_3 and L_4 consist in sums of several successive applications of $\mathcal{L}_J$, $\mathcal{L}_A$, $\mathcal{L}_B$ and $\mathcal{L}_O$. Since they are Markov generators (that leave at zero the constants), $S_3 P_t f = S_3 f_t$ and $L^4 P_t f = L^4 f_t$, and Lemma 4.6 yields that for $b' \in (b, 1/2)$, there exist $C > 0$ such that

$$\|S_3 f_t\|_{b',k'} \leq C \|f\|_{b,k}$$

$$\|L^4 f_t\|_{b',k'} \leq C \|f\|_{b,k},$$

hence the result when $k' = 0$. Regarding the derivatives of $\bar{P}_t$, we have

$$\partial_t \bar{P}_t = \frac{B}{2} e^{\frac{t}{2}B} e^{\frac{t}{2}J} \dots e^{\frac{t}{2}B} + e^{\frac{t}{2}B} \frac{J}{2} e^{\frac{t}{2}J} \dots e^{\frac{t}{2}B} + \dots + e^{\frac{t}{2}B} e^{\frac{t}{2}J} e^{\frac{t}{2}A} e^{tO} e^{\frac{t}{2}A} e^{\frac{t}{2}J} \frac{B}{2} e^{\frac{t}{2}B},$$

then, by iterating,

$$\partial_t^4 \bar{P}_t = \sum_{k_1, \dots, k_7} \frac{B^{k_1}}{2^{k_1}} e^{\frac{t}{2}B} \frac{J^{k_2}}{2^{k_2}} e^{\frac{t}{2}J} \frac{A^{k_3}}{2^{k_3}} e^{\frac{t}{2}A} O^{k_4} e^{tO} \frac{A^{k_5}}{2^{k_5}} e^{\frac{t}{2}A} \frac{J^{k_6}}{2^{k_6}} e^{\frac{t}{2}J} \frac{B^{k_7}}{2^{k_7}} e^{\frac{t}{2}B},$$

where the sum is over all nonnegative integers $k_1 \dots k_7$ such that $k_1 + \dots + k_7 = 4$. Again, the result follows from Lemma 4.6. ■

Lemma 4.9. *For all $t > 0$, all $f \in \mathcal{A}$ and all $z \in \mathbb{T}^d \times \mathbb{R}^d$,*

$$\bar{P}_t f(z) - P_t f(z) = t^3 S_3 f(z) + \mathcal{O}(t^4).$$

More explicitly, the term of order 4 is given by

$$\bar{P}_t f(z) - P_t f(z) - t^3 S_3 f(z) = \frac{1}{6} \int_0^t (t-s)^3 (\partial_s^4 \bar{P}_s - L^4 P_s) f(z) ds.$$

Proof. We use the same method as in [30]. Let $f \in \mathcal{A}$ and $z \in \mathbb{T}^d \times \mathbb{R}^d$. An order 3 Taylor expansion of $P_t f(z)$ and $\bar{P}_t f(z)$ at $t = 0$ gives

$$P_t f(z) = f(z) + t L f(z) + \frac{t^2}{2} L^2 f(z) + \frac{t^3}{6} L^3 f(z) + \frac{1}{6} \int_0^t (t-s)^3 L^4 P_s f(z) ds,$$

$$\bar{P}_t f(z) = f(z) + t \partial_t \bar{P}_t f(z)|_{t=0} + \frac{t^2}{2} \partial_t^2 \bar{P}_t f(z)|_{t=0} + \frac{t^3}{6} \partial_t^3 \bar{P}_t f(z)|_{t=0} + \frac{1}{6} \int_0^t (t-s)^3 \partial_s^4 \bar{P}_s f(z) ds.$$

We proved in Lemma 4.8 that

$$|L^4 P_t f(z)| \leq C e^{-qt} V_b(z) \quad |\partial_s^4 \bar{P}_s P_t f(z)| \leq C e^{-qt} V_b(z),$$

which shows that the previous integral remainders can be bound by $C(z)t^4$. As we saw in the proof of Lemma 4.8, for all $n \in \llbracket 1, 3 \rrbracket$,

$$\partial_t^n \bar{P}_t = \sum_{k_1, \dots, k_7} \frac{B^{k_1}}{2^{k_1}} e^{\frac{t}{2}B} \frac{J^{k_2}}{2^{k_2}} e^{\frac{t}{2}J} \frac{A^{k_3}}{2^{k_3}} e^{\frac{t}{2}A} O^{k_4} e^{tO} \frac{A^{k_5}}{2^{k_5}} e^{\frac{t}{2}A} \frac{J^{k_6}}{2^{k_6}} e^{\frac{t}{2}J} \frac{B^{k_7}}{2^{k_7}} e^{\frac{t}{2}B},$$

where the sum is over all nonnegative integers $k_1 \dots k_7$ such that $k_1 + \dots + k_7 = n$. In fact, thanks to the particular palindromic form of $\bar{P}_t$, those operators can be computed algebraically using the symmetric Baker–Campbell–Hausdorff (BCH) formula (see for instance [18], Section III.4.2), that states that for any A, B elements of a Lie algebra of a Lie group,

$$\exp\left(\frac{t}{2}A\right) \exp\left(\frac{t}{2}B\right) \exp\left(\frac{t}{2}A\right) = \exp(tS_1 + t^3S_3 + t^5S_5 + \dots)$$

where

$$\begin{aligned} S_1 &= A + B \\ S_3 &= \frac{1}{24}(2[B, [B, A]] - [A, [A, B]]). \end{aligned}$$

Therefore, in our case,

$$e^{\frac{t}{2}A} e^{tO} e^{\frac{t}{2}A} = e^{t\mathcal{B}_1},$$

with $\mathcal{B}_1 = \widehat{S}_1 + t^2\widehat{S}_3 + t^4\widehat{S}_5 + \dots$ with $\widehat{S}_1 = A + O$ and

$$\widehat{S}_3 = \frac{1}{24}(2[O, [O, A]] - [A, [A, O]]).$$

A second application of the formula gives

$$e^{\frac{t}{2}J} e^{t\mathcal{B}_1} e^{\frac{t}{2}J} = e^{t\mathcal{B}_2},$$

where $\mathcal{B}_2 = \widetilde{S}_1 + t^2\widetilde{S}_3 + \dots$ with $\widetilde{S}_1 = \widehat{S}_1 + J = A + O + J$ and

$$\begin{aligned} \widetilde{S}_3 &= \widehat{S}_3 + \frac{1}{24}(2[\widehat{S}_1, [\widehat{S}_1, J]] - [J, [J, \widehat{S}_1]]) \\ &= \frac{1}{24}(2[O, [O, A]] - [A, [A, O]] + 2[A + O, [A + O, J]] - [J, [J, A + O]]). \end{aligned}$$

A third and final application gives

$$\bar{P}_t = e^{\frac{t}{2}B} e^{t\mathcal{B}_2} e^{\frac{t}{2}B} = e^{t\mathcal{B}_3},$$

with $\mathcal{B}_3 = \overline{S}_1 + t^2\overline{S}_3 + t^4\overline{S}_5 + \dots$ where $\overline{S}_1 = \widetilde{S}_1 + B = A + B + J + O = L$ and

$$\begin{aligned} \overline{S}_3 &= \widetilde{S}_3 + \frac{1}{24}(2[\widetilde{S}_1, [\widetilde{S}_1, B]] - [B, [B, \widetilde{S}_1]]) \\ &= \frac{1}{24}(2[O, [O, A]] - [A, [A, O]] + 2[A + O, [A + O, J]] - [J, [J, A + O]] \\ &\quad + 2[A + J + O, [A + J + O, B]] - [B, [B, A + J + O]]). \end{aligned}$$

What we just showed is that

$$\partial_t \bar{P}_t f(z)|_{t=0} = Lf(z), \quad \partial_t^2 \bar{P}_t f(z)|_{t=0} = L^2 f(z), \quad \partial_t^3 \bar{P}_t f(z)|_{t=0} = (L^3 + 7\overline{S}_3)f(z).$$

Finally, set $S_3 = \frac{7}{6}\overline{S}_3$, which yields

$$\bar{P}_t f(z) - P_t f(z) = t^3 S_3 f(z) + \frac{1}{6} \int_0^t (t-s)^3 (\partial_s^4 \bar{P}_s - L^4 P_s) f(z) ds,$$

with

$$\frac{1}{6} \left| \int_0^t (t-s)^3 (\partial_s^4 \bar{P}_s - L^4 P_s) f(z) ds \right| \leq C(z)t^4,$$

which concludes the proof. ■

Proof of Theorem 4.5. We write the Talay–Tubaro expansion of the weak error.

$$\begin{aligned}\mathbb{E}_z[f(Z_{n\delta})] - \mathbb{E}_z[f(\bar{Z}_n)] &= \mathbb{E}[P_{n\delta}f(z) - P_0f(\bar{Z}_n)] \\ &= \sum_{k=0}^{n-1} \mathbb{E}_z[P_{(n-k)\delta}f(\bar{Z}_k) - P_{(n-k-1)\delta}f(\bar{Z}_{k+1})].\end{aligned}$$

By conditioning each term by $\bar{Z}_k$, we have, since $(\bar{Z}_n)$ is a Markov chain,

$$\mathbb{E}_z[f(Z_{n\delta})] - \mathbb{E}_z[f(\bar{Z}_n)] = \sum_{k=0}^{n-1} \mathbb{E}[h_k(\bar{Z}_k)],$$

where, by using the semi-group property of $(P_t)_{t \geq 0}$,

$$\begin{aligned}h_k(z) &= P_{(n-k)\delta}f(z) - \mathbb{E}[P_{(n-k-1)\delta}f(\bar{Z}_1)] \\ &= P_\delta P_{(n-k-1)\delta}f(z) - \bar{P}_\delta P_{(n-k-1)\delta}f(z).\end{aligned}$$

We are now interested in the difference $P_\delta\varphi - \bar{P}_\delta\varphi$ for $\varphi \in \mathcal{A}$. As we saw in Lemma 4.9,

$$\bar{P}_t\varphi(z) - P_t\varphi(z) = t^3 S_3\varphi(z) + \frac{1}{6} \int_0^t (t-s)^3 (\partial_s^4 \bar{P}_s - L^4 P_s)\varphi(z) ds.$$

Now, Lemma 4.8 implies that there exist $b < 1/2$ and $C > 0$ such that

$$|h_k(z)| = |P_\delta P_{(n-k-1)\delta}f(z) - \bar{P}_\delta P_{(n-k-1)\delta}f(z)| \leq C\delta^3 e^{-q(n-k-1)\delta} V_b(z),$$

and then, by using Proposition 4.1,

$$|\mathbb{E}_z[h_k(\bar{Z}_k)]| \leq C\delta^3 e^{-q(n-k-1)\delta} \mathbb{E}_z[V_b(\bar{Z}_k)] \leq C\delta^3 e^{-q(n-k-1)\delta} (e^{-C'\delta k} V_b(z) + 1).$$

By summing those inequalities on k , we have

$$\begin{aligned}|\mathbb{E}_z[f(Z_t)] - \mathbb{E}_z[f(\bar{Z}_n)]| &\leq \sum_{k=0}^{n-1} \mathbb{E}[|h_k(\bar{Z}_k)|] \\ &\leq C\delta^3 \left(V_b(z) e^{-q(n-1)\delta} \sum_{k=0}^{n-1} e^{(q-C')k\delta} + \sum_{k=0}^{n-1} e^{-qk\delta} \right) \\ &\leq C\delta^3 V_b(z) \frac{e^{-qn\delta} - e^{-C'n\delta}}{e^{-q\delta} - e^{-C'\delta}} + C''\delta^2 \leq C\delta^2 (V_b(z) e^{-qn\delta} + 1),\end{aligned}$$

which shows the result, with $C(n\delta, z) = C(V_b(z) e^{-qn\delta} + 1)$. More explicitly,

$$\begin{aligned}\mathbb{E}_z[f(Z_t)] - \mathbb{E}_z[f(\bar{Z}_n)] &= \delta^3 \left(\sum_{k=0}^{n-1} \mathbb{E}[S_3 P_{(n-k-1)\delta}f(\bar{Z}_k)] \right) + \frac{1}{6} \sum_{k=0}^{n-1} \int_0^\delta (\delta-s)^3 \mathbb{E}_z[(\partial_s^4 \bar{P}_s - L^4 P_s)g_k(\bar{Z}_k)] ds,\end{aligned}$$

and the same computations lead to

$$\left| \sum_{k=0}^{n-1} \mathbb{E}[S_3 P_{(n-k-1)\delta}f(\bar{Z}_k)] \right| \leq C \left(V_b(z) \frac{e^{-qn\delta} - e^{-C'n\delta}}{e^{-q\delta} - e^{-C'\delta}} + \frac{1}{1 - e^{-q\delta}} \right) \leq \frac{C}{\delta} (V_b(z) e^{-qn\delta} + 1),$$

and

$$\begin{aligned}\frac{1}{6} \left| \sum_{k=0}^{n-1} \int_0^\delta (\delta-s)^3 \mathbb{E}_z[(\partial_s^4 \bar{P}_s - L^4 P_s)g_k(\bar{Z}_k)] ds \right| &\leq C \sum_{k=0}^{n-1} e^{-q(n-k-1)\delta} \mathbb{E}_z[V_b(\bar{Z}_k)] \int_0^\delta (\delta-s)^3 e^{-qs} ds \\ &\leq C\delta^4 \left(V_b(z) \frac{e^{-qn\delta} - e^{-C'n\delta}}{e^{-q\delta} - e^{-C'\delta}} + \frac{1}{1 - e^{-q\delta}} \right) \\ &\leq C\delta^3 (V_b(z) e^{-qn\delta} + 1).\end{aligned}$$

■

4.3. Ergodicity of the numerical scheme

Proof of Theorem 2.3. The result follows from Harris Theorem (see for instance [22, Theorem 1.2]; the constants are made explicit in [17, Theorem 24]). More precisely, we will prove that, for δ small enough, $Q^{\lfloor 1/\delta \rfloor}$ satisfy the conditions of Harris theorem with constants which are independent from δ . Namely, we need to check, first, a Lyapunov condition

$$Q^{\lfloor 1/\delta \rfloor} V_b \leq \gamma V_b + C, \quad (4.8)$$

where $\gamma \in (0, 1)$ and $C > 0$ are independent from δ , and, second, a local coupling condition: for any compact set K of $\mathbb{T}^d \times \mathbb{R}^d$, there exists $\alpha > 0$ (again, independent from δ) such that for all $(x, v), (x', v') \in K$,

$$\|\delta_{(x,v)} Q^{\lfloor 1/\delta \rfloor} - \delta_{(x',v')} Q^{\lfloor 1/\delta \rfloor}\|_{TV} \leq 2(1 - \alpha). \quad (4.9)$$

When we will have proven these two conditions, we will get by [22, Theorem 1.2] that there exists $C'', \lambda > 0$ (which are independent from $\delta \in (0, \delta_0]$ since they can be expressed in terms of b, γ, C and α where α is given by the coupling condition on the compact set $K = \{V_b \leq 4C/(1 - \gamma)\}$) such that for all $n \in \mathbb{N}$ and all probability measures ν_1, ν_2 over $\mathbb{T}^d \times \mathbb{R}^d$,

$$\|\nu_1 Q^{n\lfloor 1/\delta \rfloor} - \nu_2 Q^{n\lfloor 1/\delta \rfloor}\|_V \leq C'' e^{-\lambda n} \|\nu_1 - \nu_2\|_V.$$

Besides, [22, Theorem 1.2] also gives the existence and uniqueness of an invariant measure μ_δ for $Q^{\lfloor 1/\delta \rfloor}$ (which by standard semi-group argument is thus the unique invariant measure for Q). Before proceeding with the proofs of (4.8) and (4.9), let us first explain the conclusion of the proof of Theorem 2.3 from this. Thanks to Proposition 4.1, for all $k \geq 1$,

$$Q^k V_b \leq (1 - C\delta)^k V_b + C' \frac{1 - (1 - C\delta)^k}{C} \leq \left(1 + \frac{C'}{C}\right) V_b,$$

since $V_b \geq 1$. As noticed in [17, Equation (86)], this implies that for all ν_1, ν_2 ,

$$\|\nu_1 Q^k - \nu_2 Q^k\|_V \leq \left(1 + \frac{C'}{C}\right) \|\nu_1 - \nu_2\|_V.$$

Decomposing any $n \geq 1$ as $n = m\lfloor 1/\delta \rfloor + k$ with $k \in [0, \lfloor 1/\delta \rfloor - 1]$, using the invariance of μ_δ by Q ,

$$\begin{aligned} \|\delta_{(x,v)} Q^n - \mu_\delta\|_V &= \|\delta_{(x,v)} Q^n - \mu_\delta Q^n\|_V \\ &\leq \left(1 + \frac{C'}{C}\right) \|\delta_{(x,v)} Q^{m\lfloor 1/\delta \rfloor} - \mu_\delta Q^{m\lfloor 1/\delta \rfloor}\|_V \\ &\leq C'' e^{-\lambda m} \left(1 + \frac{C'}{C}\right) \|\delta_{(x,v)} - \mu_\delta\|_V \\ &\leq C'' e^{-\lambda(n-1)\delta} \left(1 + \frac{C'}{C}\right) (V_b(x, v) + \mu_\delta(V_b)). \end{aligned}$$

Thanks to Proposition 4.1 and the invariance of μ_δ ,

$$\mu_\delta(V_b) \leq (1 - C\delta)\mu_\delta(V_b) + C'\delta \implies \mu_\delta(V_b) \leq \frac{C'}{C}.$$

Using again that $V_b \geq 1$, we end up with

$$\|\delta_{(x,v)} Q^n - \mu_\delta\|_V \leq C'' e^{-\lambda(n\delta - \delta_0)} \left(1 + \frac{C'}{C}\right)^2 V_b(x, v),$$

with C, λ, C', C'' which are independent from $\delta \in (0, \delta_0)$. This will conclude the proof of Theorem 2.3. It remains to check (4.8) and (4.9). The Lyapunov condition (4.8) straightforwardly follows from Proposition 4.1, using that $(1 - C\delta)^{\lfloor 1/\delta \rfloor}$ goes to $e^{-C} < 1$ as $\delta \rightarrow 0$. The rest of the proof is dedicated to establishing the coupling condition (4.9). This result has been established for the BAOAB chain

in [16]. More precisely, given a compact set K of $\mathbb{T}^d \times \mathbb{R}^d$, [16, Theorem 3] shows that there exists $\delta'_0, \varepsilon > 0$ such that for all $\delta \in (0, \delta'_0)$ and $(x, v), (x', v') \in K$,

$$\|\delta_{(x,v)} \tilde{Q}^{[1/\delta]} - \delta_{(x',v')} \tilde{Q}^{[1/\delta]}\|_{TV} \leq 2(1 - \varepsilon),$$

with $\tilde{Q}$ the transition operator of the BAOAB chain with potential U_0 (i.e. the BJAOAJB chain with $U_1 = 0$). By the coupling characterisation of the total variation norm, it means that, for any $(x, v), (x', v') \in K$, it is possible to define two BAOAB chains $(X_n, V_n)_{n \in \mathbb{N}}$ and $(X'_n, V'_n)_{n \in \mathbb{N}}$ initialized respectively at these points which are equal at time $[1/\delta]$ with probability larger than $\varepsilon > 0$ (to see that we can define such a trajectory, see e.g. the proof of [40, Lemma 3.2]). Denote $\mathcal{A}_1$ the event where these two BAOAB chains have merged before $[1/\delta]$ transitions. In particular, for all $\delta \in (0, \delta'_0)$, $\mathbb{P}(\mathcal{A}_1) \geq \varepsilon$.

Next, let us prove that there exists a compact set K' of $\mathbb{T}^d \times \mathbb{R}^d$, independent from δ small enough and from $(x, v) \in K$, such that the probability that the BAOAB chain $(X_n, V_n)_{n \in \mathbb{N}}$ initialized at (x, v) exits K' before $[1/\delta]$ transitions is smaller than $\varepsilon/4$. This classically follows from the fact V_b is a Lyapunov function for the BAOAB chain. Indeed, notice that Assumption 1 implies that the same assumptions are satisfied when replacing U_1 by 0, which means that Proposition 4.1 applies to BAOAB. Using that $V_b \geq 1$, this implies that $\tilde{Q}V_b \leq (1 + C'\delta)V_b \leq e^{C'\delta}V_b$, i.e. $e^{-C'\delta n}e^{b|V_n|^2}$ is a supermartingale. For $R > 0$, consider the stopping time $\tau = \inf\{n, |V_n| \geq R\}$. Then, letting $n \rightarrow \infty$ in $\mathbb{E}(e^{-C'\delta(n \wedge \tau)}e^{b|V_{n \wedge \tau}|^2}) \leq e^{b|v|^2}$, we get $\mathbb{E}(e^{-C'\delta\tau}) \leq De^{-bR^2}$ with $D = \sup\{e^{b|v_0|^2}, (x_0, v_0) \in K\}$. In particular, $\mathbb{P}(\tau \leq [1/\delta]) \leq e^{C'}De^{-bR^2}$. Taking R large enough (depending on K, b, C' and ε , but not on $\delta \in (0, \delta_0)$), conclusion follows with $K' = \mathbb{T}^d \times [0, R]$.

Considering the two coupled BAOAB chains $(X_n, V_n)_{n \in \mathbb{N}}$ and $(X'_n, V'_n)_{n \in \mathbb{N}}$ as before, denote by $\mathcal{A}_2$ the event where both chains remain in K' up to time $[1/\delta]$. The choice of K' ensures that $\mathbb{P}(\mathcal{A}_2) \geq 1 - \varepsilon/2$, from which $\mathbb{P}(\mathcal{A}_1 \cap \mathcal{A}_2) \geq \varepsilon/2$.

Now, we define two BJAOAJB chains by following the two previous BAOAB chains and adding random jumps. More precisely, let E be a standard exponential variable independent from the BAOAB chains and let $S = \inf\{k \geq 0, E \leq \frac{\delta}{2} \sum_{j=0}^k \lambda(\tilde{X}_j, \tilde{V}_j)\}$, where $(\tilde{X}_j, \tilde{V}_j)_{j \in \mathbb{N}}$ are the intermediary of the chain $(X_n, V_n)_{n \in \mathbb{N}}$ (after the first B step of transition, and then after the BAOA steps; which are the two places where the jumps are performed in BJAOAJB, hence the jump rate λ is evaluated). We define a BJAOAJB chain (Y_n, W_n) which follows (X_n, V_n) up to half-step S (i.e. up to the intermediary half-step J where a jump occurs), at which point a velocity jump is performed, and after that (Y_n, W_n) is evolved independently from the BAOAB chain. Similarly we define a BJAOAJB chain $(Y'_n, W'_n)_{n \in \mathbb{N}}$ initialized at (x', v') from the BAOAB trajectory $(X'_n, V'_n)_{n \in \mathbb{N}}$ (with the same variable E to define the first jump time). Let $\lambda_* = \sup_{K'} \lambda$ and $\mathcal{A}_3 = \{E > \lambda_*\}$, which is independent from the two BAOAB chains and in particular from $\mathcal{A}_1$ and $\mathcal{A}_2$. Under $\mathcal{A}_1 \cap \mathcal{A}_2 \cap \mathcal{A}_3$, the jump rate is bounded by $\lambda_* > 0$ along the two coupled BAOAB trajectories and thus $S > 2[1/\delta]$, which means there is no jump up to time $[1/\delta]$. Hence, under this event, the two BJAOAJB trajectories coincide with the BAOAB ones and thus have merged at time $[1/\delta]$. As a conclusion,

$$\|\delta_{(x,v)} \tilde{Q}^{[1/\delta]} - \delta_{(x',v')} \tilde{Q}^{[1/\delta]}\|_{TV} \leq 2(1 - \mathbb{P}(\mathcal{A}_1 \cap \mathcal{A}_2 \cap \mathcal{A}_3)) \leq 2(1 - e^{-\lambda_*\varepsilon/2}).$$

This concludes the proof of (4.9), hence of Theorem 2.3. $\blacksquare$

4.4. Expansion of the invariant measure of the numerical scheme

Proof of Theorem 2.2. Let $f \in \mathcal{A}$, $\delta_0 > 0$ such that the result of Theorem 2.3 holds and $\delta \in (0, \delta_0]$. The ergodicity of the continuous time process $(Z_t)_{t \geq 0}$ (Theorem 2.1) and the BJAOAJB chain $(\tilde{Z}_n)_{n \geq 0}$

(Theorem 2.3) imply, since $\|f\|_{V_b} < \infty$ for some $b \in (0, 1/2)$, that

$$\mu f = \lim_{n \rightarrow \infty} \frac{1}{n} \sum_{k=1}^n \mathbb{E}_z[f(Z_{k\delta})], \quad \mu_\delta f = \lim_{n \rightarrow \infty} \frac{1}{n} \sum_{k=1}^n \mathbb{E}_z[f(\bar{Z}_k)].$$

Thanks to Theorem 4.5, for any $n \geq 1$,

$$\begin{aligned} \left| \frac{1}{n} \sum_{k=1}^n \mathbb{E}_z[f(Z_{k\delta}) - f(\bar{Z}_k)] \right| &\leq \frac{C\delta^2}{n} \sum_{k=1}^n (V_b(z)e^{-nq\delta} + 1) \\ &\leq C\delta^2 \left(\frac{V_b(z)}{n(1 - e^{-q\delta})} + 1 \right). \end{aligned}$$

Thus, by taking the limit as $n \rightarrow \infty$ in the previous inequality, we get that

$$|\mu f - \mu_\delta f| \leq C\delta^2. \quad (4.10)$$

Let us now look at the error term of order δ^2 . Again, we saw in Theorem 4.5 that for all $k \in \mathbb{N}^*$,

$$\left| \mathbb{E}_z[f(Z_{k\delta}) - f(\bar{Z}_k)] - \delta^3 \sum_{p=0}^{k-1} \mathbb{E}_z[S_3 P_{(k-p-1)\delta} f(\bar{Z}_p)] \right| \leq C\delta^3 (V_b(z)e^{-qn\delta} + 1),$$

thus

$$\left| \frac{1}{n} \sum_{k=1}^n \left(\mathbb{E}_z[f(Z_{k\delta}) - f(\bar{Z}_k)] - \delta^3 \sum_{p=0}^{k-1} \mathbb{E}_z[S_3 P_{(k-p-1)\delta} f(\bar{Z}_p)] \right) \right| \leq C\delta^3 \left(\frac{V_b(z)}{n(1 - e^{-q\delta})} + 1 \right). \quad (4.11)$$

Then,

$$\begin{aligned} \frac{1}{n} \sum_{k=1}^n \sum_{p=0}^{k-1} \mathbb{E}_z[S_3 P_{(k-p-1)\delta} f(\bar{Z}_p)] &= \frac{1}{n} \sum_{k=1}^n \sum_{p=0}^{k-1} \mathbb{E}_z[S_3 P_{p\delta} f(\bar{Z}_{k-p-1})] \\ &= \frac{1}{n} \sum_{p=0}^{n-1} \sum_{k=p+1}^n \mathbb{E}_z[S_3 P_{p\delta} f(\bar{Z}_{k-p-1})] \\ &= \frac{1}{n} \sum_{p=0}^{n-1} \sum_{k=0}^{n-p-1} \mathbb{E}_z[S_3 P_{p\delta} f(\bar{Z}_k)] \\ &= \frac{1}{n} \sum_{p=0}^{n-1} \sum_{k=0}^n \mathbb{E}_z[S_3 P_{p\delta} f(\bar{Z}_k)] - \frac{1}{n} \sum_{p=0}^{n-1} \sum_{k=n-p}^n \mathbb{E}_z[S_3 P_{p\delta} f(\bar{Z}_k)]. \end{aligned}$$

By ergodicity of the numerical scheme, for all $p \in \mathbb{N}$,

$$\frac{1}{n} \sum_{k=0}^n \mathbb{E}_z[S_3 P_{p\delta} f(\bar{Z}_k)] \xrightarrow{n \rightarrow \infty} \mu_\delta S_3 P_{p\delta} f,$$

and thanks to Lemma 4.8 and Proposition 4.1, there exist $b \in (0, 1/2)$, $q > 0$ and $C > 0$ such that

$$\begin{aligned} \frac{1}{n} \left| \sum_{k=0}^n \mathbb{E}_z[S_3 P_{p\delta} f(\bar{Z}_k)] \right| &\leq \frac{C}{n} e^{-pq\delta} \sum_{k=0}^n \mathbb{E}_z[V_b(\bar{Z}_k)] \\ &\leq C e^{-pq\delta} \left(\frac{V_b(z)}{n} \sum_{k=0}^n e^{-C'\delta k} + 1 \right) \\ &\leq C e^{-qp\delta} \left(\frac{V_b(z)}{n(1 - e^{-C'\delta})} + 1 \right). \end{aligned}$$

Since the right term of the last inequality is summable in p , Lebesgue convergence theorem implies that

$$\frac{1}{n} \sum_{p=0}^{n-1} \sum_{k=0}^n \mathbb{E}_z[S_3 P_{p\delta} f(\bar{Z}_k)] \xrightarrow{n \rightarrow \infty} \sum_{p=0}^{\infty} \mu_{\delta} S_3 P_{p\delta} f.$$

Similarly, there exist $b \in (0, 1/2)$, $q > 0$ and $C > 0$ such that

$$\begin{aligned} \frac{1}{n} \sum_{p=0}^{n-1} \sum_{k=n-p}^n \mathbb{E}_z[|S_3 P_{p\delta} f(\bar{Z}_k)|] &\leq \frac{C}{n} \sum_{p=0}^{n-1} e^{-qp\delta} \left(V_b(z) \sum_{k=n-p}^n e^{-C'\delta k} + p \right) \\ &\leq \frac{C}{n} \sum_{p=0}^{n-1} e^{-qp\delta} \left(\frac{V_b(z)}{1 - e^{-C'\delta}} + p \right) \xrightarrow{n \rightarrow \infty} 0, \end{aligned}$$

which shows that

$$\frac{1}{n} \sum_{k=1}^n \sum_{p=0}^{k-1} \mathbb{E}_z[S_3 P_{(k-p-1)\delta} f(\bar{Z}_p)] \xrightarrow{n \rightarrow \infty} \sum_{p=0}^{\infty} \mu_{\delta} S_3 P_{p\delta} f.$$

We may then let $n \rightarrow \infty$ in (4.11) to obtain

$$\left| \mu f - \mu_{\delta} f - \delta^3 \sum_{p=0}^{\infty} \mu_{\delta} S_3 P_{p\delta} f \right| \leq C\delta^3.$$

We have shown that

$$|\mu f - \mu_{\delta} f - \delta^2 \mu_{\delta} g_{\delta}| \leq C\delta^3 \quad \text{with} \quad g_{\delta} = \delta \sum_{p=0}^{\infty} S_3 P_{p\delta} f.$$

Thanks to Theorem 2.1 and Lemma 4.8, for all $t \geq 0$,

$$|\partial_t S_3 P_t f(z)| = |S_3 L P_t f(z)| \leq C^{-qt} V_b(z),$$

for a certain $b < 1/2$. Therefore, for any $n \in \mathbb{N}^*$,

$$\begin{aligned} \left| \int_0^{n\delta} S_3 P_s f \, ds - \delta \sum_{p=0}^{n-1} S_3 P_{p\delta} f \right| &\leq \sum_{p=0}^{n-1} \int_{p\delta}^{(p+1)\delta} |S_3 P_s f - S_3 P_{p\delta} f| \, ds \\ &\leq \sum_{p=0}^{n-1} \sup_{p\delta \leq t \leq (p+1)\delta} |\partial_t S_3 P_t f| \int_{p\delta}^{(p+1)\delta} (s - p\delta) \, ds \\ &\leq \frac{C V_b(z) \delta^2}{2} \sum_{p=0}^{n-1} e^{-pq\delta} \\ &\leq \frac{C V_b(z) \delta}{2}, \end{aligned}$$

which implies, by taking the limit $n \rightarrow \infty$, that

$$\|g_{\delta} - g\|_{V_b} \leq C\delta \quad \text{with} \quad g = \int_0^{\infty} S_3 P_s f \, ds,$$

and thus

$$|\mu_{\delta} g_{\delta} - \mu_{\delta} g| \leq C\delta.$$

Again, $g \in \mathcal{A}$. Indeed, if $f \in \mathcal{A}$, thanks to Lemma 4.6, we know that all spatial derivatives of $S_3 P_t f$ are bounded by a term of the form $C e^{-qt} V_b(z)$ for a certain $b < 1/2$, which is integrable in t on $\mathbb{R}_+$. Therefore, the theorem of differentiation under the integral sign yields that for all $\alpha \in \mathbb{N}^d$,

$$|\partial^{\alpha} g(z)| \leq C V_b(z) \int_0^{\infty} e^{-qs} \, ds \leq C V_b(z).$$

This allows to apply (4.10) to the function g , that is

$$|\mu_\delta g - \mu g| \leq C\delta^2.$$

Finally,

$$\begin{aligned} |\mu f - \mu_\delta f - \delta^2 \mu g| &\leq |\mu f - \mu_\delta f - \delta^2 \mu_\delta g_\delta| + \delta^2 |\mu_\delta g_\delta - \mu_\delta g| + \delta^2 |\mu_\delta g - \mu g| \\ &\leq C\delta^3. \end{aligned}$$

We have therefore explicitated the first term in the expansion of the invariant measure:

$$\mu_\delta f = \mu f + \delta^2 \int_0^\infty \mu S_3 P_s f \, ds + \mathcal{O}(\delta^3),$$

which concludes the proof of the theorem. $\blacksquare$

4.5. Quadratic risk

Proof of Corollary 2.4. Denoting $g = f - \mu(f)$ and ν_0 the distribution of Z_0 ,

$$\begin{aligned} \mathbb{E} \left(\left| \frac{1}{n} \sum_{k=1}^n f(Z_k) - \mu(f) \right|^2 \right) &= \frac{1}{n^2} \sum_{j=1}^n \sum_{k=1}^n \mathbb{E}(g(Z_k)g(Z_j)) \\ &= \frac{1}{n^2} \sum_{j=1}^n \nu_0 Q^j(g^2) + \frac{2}{n^2} \sum_{j=1}^n \sum_{k=j+1}^n \nu_0 Q^j(g Q^{k-j} g). \end{aligned}$$

Bounding $g^2 \leq 4\|f\|_{V_b}^2 V_b^2$ and using Proposition 4.1 (applied to $V_b^2 = V_{2b}$ since $b < 1/4$), we treat the first term as

$$\frac{1}{n^2} \sum_{j=1}^n \nu_0 Q^j(g^2) \leq \frac{4\|f\|_{V_b}^2}{n^2} \sum_{j=1}^n \nu_0 Q^j(V_b^2) \leq \frac{4\|f\|_{V_b}^2}{n^2} \sum_{j=1}^n \left[(1 - C\delta)^j \nu_0(V_b^2) + \frac{C'}{\delta C} \right].$$

For the second term, thanks to Theorem 2.3, we bound $\|Q^k g\|_{V_b} \leq 2Ce^{-\lambda k\delta} \|f\|_{V_b}$, hence $|g Q^k g| \leq 4Ce^{-\lambda k\delta} \|f\|_{V_b}^2 V_b^2$, and then

$$\frac{2}{n^2} \sum_{j=1}^n \sum_{k=j+1}^n \nu_0 Q^j(g Q^{k-j} g) \leq \frac{8C\|f\|_{V_b}^2}{n^2} \sum_{j=1}^n \sum_{k=j+1}^n e^{-\lambda(k-j)\delta} \nu_0 Q^j(V_b^2) \leq \frac{8C\|f\|_{V_b}^2}{n^2(1 - e^{-\lambda\delta})} \sum_{j=1}^n \nu_0 Q^j(V_b^2),$$

and we bound $\sum_{j=1}^n \nu_0 Q^j(V_b^2)$ as before.

For the Richardson extrapolation, we simply bound

$$\begin{aligned} \mathbb{E} \left(\left| \frac{4}{3n} \sum_{k=1}^n f(\tilde{Z}_k) - \frac{1}{3n} \sum_{k=1}^n f(Z_k) - \mu(f) \right|^2 \right) \\ \leq 3\mathbb{E} \left(\left| \frac{4}{3n} \sum_{k=1}^n f(\tilde{Z}_k) - \frac{4}{3} \mu_{\delta/2}(f) \right|^2 \right) \\ + 3\mathbb{E} \left(\left| \frac{1}{3n} \sum_{k=1}^n f(Z_k) - \frac{1}{3} \mu_\delta(f) \right|^2 \right) + 3 \left(\frac{4}{3} \mu_{\delta/2}(f) - \frac{1}{3} \mu_\delta(f) - \mu(f) \right)^2. \end{aligned}$$

The two first terms are bounded by $\frac{C}{n\delta}$ for some $C > 0$ as before, and conclusion follows from Theorem 2.2 (the leading terms of order δ^2 in the asymptotic bias cancelling out). $\blacksquare$

References

- [1] Christophe Andrieu, Alain Durmus, Nikolas Nüsken, and Julien Roussel. Hypocoercivity of piecewise deterministic Markov process-Monte Carlo. *Ann. Appl. Probab.*, 31(5):2478–2517, 2021.
- [2] Dominique Bakry, Ivan Gentil, and Michel Ledoux. *Analysis and geometry of Markov diffusion operators*, volume 348 of *Grundlehren der Mathematischen Wissenschaften*. Springer, 2014.
- [3] Andrea Bertazzi, Joris Bierkens, and Paul Dobson. Approximations of Piecewise Deterministic Markov Processes and their convergence properties. *Stochastic Processes Appl.*, 154:91–153, 2022.
- [4] Andrea Bertazzi, Paul Dobson, and Pierre Monmarché. Splitting schemes for second order approximations of piecewise-deterministic Markov processes. <https://arxiv.org/abs/2301.02537v1>, 2023.
- [5] Joris Bierkens, Paul Fearnhead, and Gareth Roberts. The Zig-Zag process and super-efficient sampling for Bayesian analysis of big data. *Ann. Stat.*, 47(3):1288–1320, 2019.
- [6] Joris Bierkens, Sebastiano Grazi, Kengo Kamatani, and Gareth Roberts. The Boomerang Sampler. In Hal Daumé III and Aarti Singh, editors, *Proceedings of the 37th International Conference on Machine Learning*, volume 119 of *Proceedings of Machine Learning Research*, pages 908–918. PMLR, 2020.
- [7] Joris Bierkens, Gareth Roberts, and Pierre-André Zitt. Ergodicity of the zigzag process. *Ann. Appl. Probab.*, 29(4):2266–2301, 2019.
- [8] Léon Bottou, Frank E. Curtis, and Jorge Nocedal. Optimization methods for large-scale machine learning. *SIAM Rev.*, 60(2):223–311, 2018.
- [9] Nawaf Bou-Rabee. Time Integrators for Molecular Dynamics. *Entropy*, 16:138–162, 2014.
- [10] Alexandre Bouchard-Côté, Sebastian J. Vollmer, and Arnaud Doucet. The bouncy particle sampler: a nonreversible rejection-free Markov chain Monte Carlo method. *J. Am. Stat. Assoc.*, 113(522):855–867, 2018.
- [11] Vincent Calvez, Gaël Raoul, and Christian Schmeiser. Confinement by biased velocity jumps: Aggregation of escherichia coli. *Kinet. Relat. Models*, 8(4):651–666, 2015.
- [12] Alice Corbella, Simon E. F. Spencer, and Gareth Roberts. Automatic Zig-Zag sampling in practice. *Stat. Comput.*, 32(6): article no. 107, 2022.
- [13] George Deligiannidis, Alexandre Bouchard-Côté, and Arnaud Doucet. Exponential ergodicity of the bouncy particle sampler. *Ann. Stat.*, 47(3):1268–1287, 2019.
- [14] George Deligiannidis, Daniel Paulin, Alexandre Bouchard-Côté, and Arnaud Doucet. Randomized Hamiltonian Monte Carlo as scaling limit of the bouncy particle sampler and dimension-free convergence rates. *Ann. Appl. Probab.*, 31(6):2612–2662, 2021.
- [15] Jean Dolbeault, Clément Mouhot, and Christian Schmeiser. Hypocoercivity for linear kinetic equations conserving mass. *Trans. Am. Math. Soc.*, 367(6):3807–3828, 2015.
- [16] Alain Durmus, Aurélien Enfroy, Éric Moulines, and Gabriel Stoltz. Uniform minorization condition and convergence bounds for discretizations of kinetic Langevin dynamics. *Ann. Inst. Henri Poincaré, Probab. Stat.*, 61(1):629–664, 2025.
- [17] Alain Durmus, Aarnaud Guillin, and Pierre Monmarché. Geometric ergodicity of the Bouncy Particle Sampler. *Ann. Appl. Probab.*, 30(5):2069–2098, 2020.
- [18] Gerhard Wanner Ernst Hairer and Christian Lubich. *Geometric Numerical Integration*. Springer, 2nd edition, 2006.
- [19] Ulrich Essmann, Lalith Perera, Max L. Berkowitz, Tom Darden, Hsing Lee, and Lee G. Pedersen. A smooth particle mesh Ewald method. *J. Chem. Phys.*, 103(19):8577–8593, 1995.
- [20] Josephine Evans. Hypocoercivity in Phi-entropy for the linear relaxation Boltzmann equation on the torus. *SIAM J. Math. Anal.*, 53(2):1357–1378, 2021.

- [21] Nicolai Gouraud, Louis Lagardère, Olivier Adjoua, Thomas Plé, Pierre Monmarché, and Jean-Philip Piquemal. Velocity Jumps for Molecular Dynamics. *J. Chem. Theory Comput.*, 21(6):2854–2866, 2025.
- [22] Martin Hairer and Jonathan C. Mattingly. Yet another look at Harris’ ergodic theorem for Markov chains. In *Seminar on stochastic analysis, random fields and applications VI. Centro Stefano Franscini, Ascona (Ticino), Switzerland, May 19–23, 2008.*, pages 109–117. Birkhäuser, 2011.
- [23] Shi Jin, Lei Li, and Jian-Guo Liu. Random Batch Methods (RBM) for interacting particle systems. *J. Comput. Phys.*, 400: article no. 108877, 2020.
- [24] Shi Jin, Lei Li, Zhenli Xu, and Yue Zhao. A random batch Ewald method for particle systems with Coulomb interactions. *SIAM J. Sci. Comput.*, 43(4):b937–b960, 2021.
- [25] Lucas Journal. Weak error expansion of a numerical scheme with rejection for singular Langevin process. *ESAIM, Math. Model. Numer. Anal.*, 58(3):1153–1184, 2024.
- [26] Lucas Journal and Pierre Monmarché. Convergence of the kinetic annealing for general potentials. *Electron. J. Probab.*, 27: article no. 159 (37 pages), 2022.
- [27] Marie Kopec. Weak backward error analysis for Langevin process. *BIT Numer. Math.*, 55(4):1057–1103, 2015.
- [28] Benedict Leimkuhler and Charles Matthews. Rational construction of stochastic numerical methods for molecular sampling. *AMRX, Appl. Math. Res. Express*, 2013(1):34–56, 2013.
- [29] Benedict Leimkuhler and Charles Matthews. Robust and efficient configurational molecular sampling via Langevin dynamics. *J. Chem. Phys.*, 138(17): article no. 174102, 2013.
- [30] Benedict Leimkuhler, Charles Matthews, and Gabriel Stoltz. The computation of averages from equilibrium and nonequilibrium Langevin molecular dynamics. *IMA J. Numer. Anal.*, 36(1):13–79, 2016.
- [31] Vincent Lemaire, Michèle Thieullen, and Nicolas Thomas. Exact simulation of the jump times of a class of piecewise deterministic Markov processes. *J. Sci. Comput.*, 75(3):1776–1807, 2018.
- [32] Peter A. W. Lewis and Gerald S. Shedler. Simulation of nonhomogeneous Poisson processes by thinning. *Nav. Res. Logist. Q.*, 26:403–413, 1979.
- [33] Tsoy-Wo Ma. Higher chain formula proved by combinatorics. *Electron. J. Comb.*, 16(1): article no. N21 (7 pages), 2009.
- [34] Jonathan C. Mattingly, Andrew M. Stuart, and M. V. Tretyakov. Convergence of Numerical Time-Averaging and Stationary Measures via Poisson Equations. *SIAM J. Numer. Anal.*, 48(2):552–577, 2010.
- [35] Pierre Monmarché. On $\mathcal{H}^1$ and entropic convergence for contractive PDMP. *Electron. J. Probab.*, 20: article no. 128 (30 pages), 2015.
- [36] Pierre Monmarché. Piecewise deterministic simulated annealing. *ALEA, Lat. Am. J. Probab. Math. Stat.*, 13(1):357–398, 2016.
- [37] Pierre Monmarché. Generalized Γ calculus and application to interacting particles on a graph. *Potential Anal.*, 50(3):439–466, 2019.
- [38] Pierre Monmarché. A note on Fisher Information hypocoercive decay for the linear Boltzmann equation. *Anal. Math. Phys.*, 11(1): article no. 1 (11 pages), 2021.
- [39] Pierre Monmarché. Almost sure contraction for diffusions on R^d . Application to generalised Langevin diffusions. *Stochastic Processes Appl.*, 161:316–349, 2023.
- [40] Pierre Monmarché. Elementary coupling approach for non-linear perturbation of Markov processes with mean-field jump mechanisms and related problems. *ESAIM, Probab. Stat.*, 27:278–323, 2023.
- [41] Pierre Monmarché. An entropic approach for Hamiltonian Monte Carlo: The idealized case. *Ann. Appl. Probab.*, 34(2):2243–2293, 2024.
- [42] Pierre Monmarché, Mathias Rousset, and Pierre-André Zitt. Exact targeting of gibbs distributions using velocity-jump processes. *Stoch. Partial Differ. Equ., Anal. Comput.*, 11(3):908–947, 2023.

- [43] Pierre Monmarché, Jérémy Weisman, Louis Lagardère, and Jean-Philip Piquemal. Velocity jump processes: An alternative to multi-timestep methods for faster and accurate molecular dynamics simulations. *J. Chem. Phys.*, 153(2): article no. 024101, 2020.
- [44] Filippo Pagani, Augustin Chevallier, Sam Power, Thomas House, and Simon Cotter. NuZZ: Numerical Zig-Zag for general models. *Stat. Comput.*, 34(1): article no. 61, 2024.
- [45] Elias A. J. F. Peters and G. de With. Rejection-free Monte Carlo sampling for general potentials. *Phys. Rev. E*, 85(2): article no. 026703, 2012.
- [46] Pier Paolo Poier, Louis Lagardère, and Jean-Philip Piquemal. O(N) Stochastic Evaluation of Many-Body van der Waals Energies in Large Complex Systems. *J. Chem. Theory Comput.*, 8(3):1633–1645, 2022.
- [47] Denis Talay. Stochastic Hamiltonian systems: Exponential convergence to the invariant measure, and discretization by the implicit Euler scheme. *Markov Process. Relat. Fields*, 8(2):163–198, 2002.
- [48] Denis Talay and Luciano Tubaro. Expansion of the global error for numerical schemes solving stochastic differential equations. *Stochastic Anal. Appl.*, 8(4):483–509, 1990.
- [49] M. Tuckerman, Bruce J. Berne, and Glenn J. Martyna. Reversible multiple time scale molecular dynamics. *J. Chem. Phys.*, 97(3):1990–2001, 1992.
- [50] Cédric Villani. *Hypocoercivity*, volume 950 of *Memoirs of the American Mathematical Society*. American Mathematical Society, 2009.
- [51] Kosaku Yosida. *Functional analysis*. Springer, repr. of the 6th edition, 1994.
- [52] Chaoen Zhang. Hypocoercivity and global hypoellipticity for the kinetic Fokker-Planck equation in H^k spaces. *Kinet. Relat. Models*, 17(3):394–435, 2024.