



**HAL**  
open science

# Uniform strong and weak error estimates for numerical schemes applied to multiscale SDEs in a Smoluchowski-Kramers diffusion approximation regime

Charles-Edouard Bréhier

## ► To cite this version:

Charles-Edouard Bréhier. Uniform strong and weak error estimates for numerical schemes applied to multiscale SDEs in a Smoluchowski-Kramers diffusion approximation regime. *Journal of Computational Dynamics*, 2023, 10 (3), pp.387-424. 10.3934/jcd.2023005 . hal-03741115

**HAL Id: hal-03741115**

**<https://hal.science/hal-03741115>**

Submitted on 31 Jul 2022

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

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

# Uniform strong and weak error estimates for numerical schemes applied to multiscale SDEs in a Smoluchowski–Kramers diffusion approximation regime

Charles-Edouard Bréhier

ABSTRACT. We study a family of numerical schemes applied to a class of multiscale systems of stochastic differential equations. When the time scale separation parameter vanishes, a well-known Smoluchowski–Kramers diffusion approximation result states that the slow component of the considered system converges to the solution of a standard Itô stochastic differential equation. We propose and analyse schemes for strong and weak effective approximation of the slow component. Such schemes satisfy an asymptotic preserving property and generalize the methods proposed in the recent article [4]. We fill a gap in the analysis of these schemes and prove strong and weak error estimates, which are uniform with respect to the time scale separation parameter.

## 1. Introduction

In this article, we consider multiscale systems of stochastic differential equations of the type

$$(1) \quad \begin{cases} dq^\epsilon(t) = \frac{p^\epsilon(t)}{\epsilon} dt \\ dp^\epsilon(t) = -\frac{p^\epsilon(t)}{\epsilon^2} dt + \frac{f(q^\epsilon(t))}{\epsilon} dt + \frac{\sigma(q^\epsilon(t))}{\epsilon} d\beta(t), \end{cases}$$

where  $\epsilon \in (0, \epsilon_0)$  is a time-scale separation parameter. The unknowns  $q^\epsilon(t)$  and  $p^\epsilon(t)$  and the Wiener process  $\beta$  take values in  $\mathbb{R}^d$ , and the mapping  $f$  and  $\sigma$  satisfy appropriate regularity conditions. The objective of this article is to study numerical schemes for the approximation of the component  $Q^\epsilon$ , for arbitrary values of the time-scale separation parameter  $\epsilon$ , in particular when it vanishes. This is not a trivial task since the component  $p^\epsilon$  evolves at the fast time scale  $t/\epsilon^2$ , and a crude discretization would impose stringent conditions on the time-step size  $\Delta t$ .

It is a well-known result in the analysis of multiscale stochastic systems that  $q^\epsilon(t)$  converges, at least in distribution, when  $\epsilon \rightarrow 0$ , to  $q^0(t)$ , for all  $t \geq 0$ , where  $X^0$  is the solution of the stochastic differential equation

$$(2) \quad dq^0(t) = f(q^0(t))dt + \sigma(q^0(t))d\beta(t)$$

where the noise is interpreted in the sense of Itô. We refer for instance to [18, Chapter 11] for a description of this type of convergence result, and see Proposition 2.2 below for a precise statement, where convergence is understood in a stronger sense than convergence in distribution. The convergence result  $q^\epsilon \rightarrow q^0$  is often called a Smoluchowski–Kramers diffusion approximation result in the literature. If  $\sigma$  is constant and equal to the identity, and if  $f = -\nabla V$  for some potential energy function  $V : \mathbb{R}^d \rightarrow \mathbb{R}$ , the SDE system (1) describes the Langevin dynamics, whereas the SDE (2) describes the overdamped Langevin dynamics, see for instance [14, Sections 2.2.3 and 2.2.4], and also the recent article [20] and references therein.

In order to define numerical schemes which perform better than crude methods when  $\epsilon$  varies and may vanish, it is relevant to resort to the notion of asymptotic preserving schemes as studied in the recent article [4]: if  $\Delta t = T/N$  denotes the time-step size with given  $T \in (0, \infty)$  and  $N \in \mathbb{N}$ , one has a commutative

diagram property

$$\begin{array}{ccc} q_N^{\epsilon, \Delta t} & \xrightarrow{N \rightarrow \infty} & q^\epsilon(T) \\ \downarrow \epsilon \rightarrow 0 & & \downarrow \epsilon \rightarrow 0 \\ q_N^{0, \Delta t} & \xrightarrow{N \rightarrow \infty} & q^0(T), \end{array}$$

where  $(q_n^{\epsilon, \Delta t}, p_n^{\epsilon, \Delta t})_{0 \leq n \leq N}$  is the scheme for given values of  $\epsilon$  and  $\Delta t$ , and one needs to check that

- the scheme is consistent for any value of  $\epsilon > 0$  when  $\Delta t \rightarrow 0$ ,
- there exists a limiting scheme  $(q_n^{0, \Delta t})_{0 \leq n \leq N}$  when  $\epsilon \rightarrow 0$  for any value of  $\Delta t > 0$ ,
- the limiting scheme is consistent with the limiting equation when  $\Delta t \rightarrow 0$ .

As explained in [4], the last property may fail to hold for some crude methods. However in the situation considered in this article there are no such subtleties for instance related to the interpretation of the noise.

In this article, we study numerical schemes and obtain strong and weak error estimates which are uniform with respect to the time-scale separation parameter  $\epsilon \in (0, \epsilon_0)$ , meaning that  $q^\epsilon(T)$  can be approximated by  $q_N^{\epsilon, \Delta t}$  with a cost which is independent of  $\epsilon$ . On the one hand, for the numerical scheme

$$(3) \quad \begin{cases} q_{n+1}^{\epsilon, \Delta t} = q_n^{\epsilon, \Delta t} + \frac{\Delta t}{\epsilon} p_{n+1}^{\epsilon, \Delta t} \\ p_{n+1}^{\epsilon, \Delta t} = p_n^{\epsilon, \Delta t} - \frac{\Delta t}{\epsilon^2} p_{n+1}^{\epsilon, \Delta t} + \frac{\Delta t f(q_n^{\epsilon, \Delta t})}{\epsilon} + \frac{\sigma(q_n^{\epsilon, \Delta t})}{\epsilon} \Delta \beta_n, \end{cases}$$

one obtains uniform strong error estimates

$$(4) \quad \sup_{\epsilon \in (0, \epsilon_0)} (\mathbb{E}[|q_N^{\epsilon, \Delta t} - q^\epsilon(N\Delta t)|^2])^{\frac{1}{2}} \leq C(T)\Delta t,$$

see Theorem 3.1 for a precise statement. On the other hand, for the numerical scheme

$$(5) \quad \begin{cases} q_{n+1}^{\epsilon, \Delta t} = q_n^{\epsilon, \Delta t} + \epsilon(1 - e^{-\frac{\Delta t}{\epsilon^2}})p_n^{\epsilon, \Delta t} + (\Delta t - \epsilon^2(1 - e^{-\frac{\Delta t}{\epsilon^2}}))f(q_n^{\epsilon, \Delta t}) \\ \quad + \sigma(q_n^{\epsilon, \Delta t})(\beta(t_{n+1}) - \beta(t_n)) - \sigma(q_n^{\epsilon, \Delta t}) \int_{t_n}^{t_{n+1}} e^{-\frac{t_{n+1}-s}{\epsilon^2}} d\beta(s) \\ p_{n+1}^{\epsilon, \Delta t} = e^{-\frac{\Delta t}{\epsilon^2}} p_n^{\epsilon, \Delta t} + \epsilon(1 - e^{-\frac{\Delta t}{\epsilon^2}})f(q_n^{\epsilon, \Delta t}) + \frac{1}{\epsilon} \sigma(q_n^{\epsilon, \Delta t}) \int_{t_n}^{t_{n+1}} e^{-\frac{t_{n+1}-s}{\epsilon^2}} d\beta(s), \end{cases}$$

and functions  $\varphi : \mathbb{R}^d \rightarrow \mathbb{R}$  of class  $\mathcal{C}^3$ , one obtains weak error estimates

$$(6) \quad |\mathbb{E}[\varphi(q_N^{\epsilon, \Delta t})] - \mathbb{E}[\varphi(q^\epsilon(T))]| \leq C(T, \varphi)(\Delta t + \epsilon),$$

$$(7) \quad \sup_{\epsilon \in (0, \epsilon_0)} |\mathbb{E}[\varphi(q_N^{\epsilon, \Delta t})] - \mathbb{E}[\varphi(q^\epsilon(T))]| \leq C(T, \varphi)\Delta t^{\frac{1}{2}}.$$

See Theorem 3.3 for a precise statement, and Proposition 2.1 for explanations concerning the construction of the numerical scheme (5).

The uniform weak error estimate (7) may be suboptimal: indeed one only obtains order 1/2, and therefore (7) is a straightforward corollary of the uniform strong error estimate (4). The weak error estimate (6) is not uniform with respect to  $\epsilon$  but may be more precise in regimes where  $\epsilon$  is negligible compared with  $\Delta t$ . The weak error estimate (6) is also similar to error bounds which may be obtained for different multiscale numerical schemes, for instance based on the Heterogeneous Multiscale Method. Improving the uniform weak error estimate (7) to obtain order 1 is left for future works.

To the best of our knowledge, the strong error estimates (4) and the weak error estimates (6)–(7) have not been obtained previously and our results thus fill a gap in the literature. These proofs require delicate and non trivial arguments. On the one hand, proving (4) is based on an appropriate change of unknowns and analysis of multiple error terms. On the other hand, proving (6)–(7) is based on the standard approach using solutions of Kolmogorov equations for weak error analysis. Proposition 6.1 gives the required bounds on the derivatives of these solutions, with a careful analysis of the dependence with respect to  $\epsilon$ .

Note that the recent preprint [3] is also concerned with the proof of uniform (strong) error estimates for a class of multiscale SDE systems in a diffusion approximation regime. However, the structure of the

systems, the results and the techniques of proof are substantially different, which justifies to perform the analysis in separate articles.

The analysis of numerical methods for multiscale stochastic differential equations is an active research area. The recent articles [2] and [12] propose uniformly accurate methods for SDE systems which are different from (1) considered in this article. The recent article [4] has introduced a notion of asymptotic preserving schemes which applies to (1), and some uniform error estimates were proved for SDE systems in an averaging regime. We also refer to the PhD thesis [19] for supplementary results and numerical experiments. In this article, as already mentioned, we fill a gap in [4] and prove some uniform error estimates in the diffusion approximation regime, for the schemes (3) and (5) applied to (1). The articles [7] and [15] illustrate why effective numerical approximation of solutions of SDEs may be more subtle than for deterministic problems. Many other techniques have been introduced to design effective methods for the numerical approximation of multiscale SDE systems, let us mention spectral methods [1], heterogeneous multiscale methods [6], projective integration methods [8], equation-free methods [10], parareal methods [13], micro-macro acceleration methods [23] for instance. We refer to the monographs [9, 11, 16, 17] for general results on numerical methods applied to stochastic differential equations.

This article is organized as follows. Section 2 describes the setting, in particular the considered multiscale SDE systems are presented in Section 2.1 and the numerical schemes studied in this work are given in Section 2.2. The main results of this article are stated and discussed in Section 3: uniform strong error estimates are given in Theorem 3.1 (Section 3.1) and weak error estimates are given in Theorem 3.3 (Section 3.2). Moment bounds are stated and proved in Section 4. Theorem 3.1 is proved in Section 5 whereas Theorem 3.3 is proved in Section 6. Auxiliary regularity results for solutions of Kolmogorov equations, with a careful analysis of the dependence with respect to  $\epsilon$ , are stated in Section 6.1 and proved in Section 6.3.

## 2. Setting

Let  $d \in \mathbb{N}$  be an integer. The norm and inner product in the standard Euclidian space  $\mathbb{R}^d$  are denoted by  $|\cdot|$  and  $\langle \cdot, \cdot \rangle$  respectively. The set of  $d \times d$  matrices with real-valued entries is denoted by  $\mathcal{M}_d(\mathbb{R})$ . The same notation is used to denote the norm and inner product in the space  $\mathbb{R}^{2d}$ . Let  $(\beta(t))_{t \geq 0}$  be a  $\mathbb{R}^d$ -valued standard Wiener process, defined on a probability space  $(\Omega, \mathcal{F}, \mathbb{P})$  which satisfies the usual conditions. The expectation operator is denoted by  $\mathbb{E}[\cdot]$ .

The time-scale separation parameter is denoted by  $\epsilon$ . Without loss of generality, it is assumed that  $\epsilon \in (0, \epsilon_0)$ , where  $\epsilon_0$  is an arbitrary positive parameter. The time-step size of the numerical schemes is denoted by  $\Delta t$ . It is assumed that  $\Delta t = T/N$  where  $T \in (0, \infty)$  is an arbitrary positive real number, and  $N \in \mathbb{N}$  is an integer. For all  $n \in \{0, \dots, N\}$ , let  $t_n = n\Delta t$ . Without loss of generality, it is assumed that  $\Delta t \in (0, \Delta t_0)$ , where  $\Delta t_0 = T/N_0$  is an arbitrary positive real number. Equivalently, it is assumed that  $N \geq N_0$ . For all  $n \in \{0, \dots, N-1\}$ , set  $\Delta\beta_n = \beta(t_{n+1}) - \beta(t_n)$ .

If  $\varphi : \mathbb{R}^d \rightarrow \mathbb{R}$  is a mapping of class  $\mathcal{C}^3$ , its first, second and third order derivatives are denoted by  $\nabla\varphi$ ,  $\nabla^2\varphi$  and  $\nabla^3\varphi$  respectively. Set

$$\begin{aligned} \|\varphi\|_1 &= \sup_{x \in \mathbb{R}^d} \sup_{h^1 \in \mathbb{R}^d} |\nabla\varphi(x) \cdot h^1|, \\ \|\varphi\|_2 &= \|\varphi\|_1 + \sup_{x \in \mathbb{R}^d} \sup_{h^1, h^2 \in \mathbb{R}^d} |\nabla^2\varphi(x) \cdot (h^1, h^2)|, \\ \|\varphi\|_3 &= \|\varphi\|_2 + \sup_{x \in \mathbb{R}^d} \sup_{h^1, h^2, h^3 \in \mathbb{R}^d} |\nabla^3\varphi(x) \cdot (h^1, h^2, h^3)|. \end{aligned}$$

If  $\phi : (q, p) \in \mathbb{R}^d \times \mathbb{R}^d \rightarrow \mathbb{R}$  is of class  $\mathcal{C}^1$ ,  $\nabla_q\phi$  and  $\nabla_p\phi$  are the partial derivatives of  $\phi$  with respect to the variables  $q$  and  $p$  respectively. Similar notation is used for higher order derivatives.

**2.1. The multiscale SDE system.** We consider the following class of multiscale SDE systems

$$(8) \quad \begin{cases} dq^\epsilon(t) = \frac{p^\epsilon(t)}{\epsilon} dt \\ dp^\epsilon(t) = -\frac{p^\epsilon(t)}{\epsilon^2} dt + \frac{f(q^\epsilon(t))}{\epsilon} dt + \frac{\sigma(q^\epsilon(t))}{\epsilon} d\beta(t), \end{cases}$$

where  $q^\epsilon(t) \in \mathbb{R}^d$  and  $p^\epsilon(t) \in \mathbb{R}^d$  for all  $t \geq 0$ . The mappings  $f : \mathbb{R}^d \rightarrow \mathbb{R}^d$  and  $\sigma : \mathbb{R}^d \rightarrow \mathcal{M}_d(\mathbb{R})$  satisfy Assumption 1 below.

ASSUMPTION 1. *Let  $f : \mathbb{R}^d \rightarrow \mathbb{R}^d$  and  $\sigma : \mathbb{R}^d \rightarrow \mathcal{M}_d(\mathbb{R})$  be mappings of class  $\mathcal{C}^3$ , with bounded derivatives of order 1, 2, 3. The mapping  $\sigma$  is assumed to be bounded.*

For all  $q \in \mathbb{R}^d$  and  $i, j \in \{1, \dots, d\}$ , set

$$(9) \quad a_{ij}(q) = \sum_{k=1}^d \sigma_{ik}(q) \sigma_{jk}(q).$$

Note that  $a$  is of class  $\mathcal{C}^3$ , and since  $\sigma$  and its derivatives are bounded,  $a$  and its derivatives are bounded. Therefore the mapping  $a$  is globally Lipschitz continuous: there exists  $C \in (0, \infty)$  such that for all  $q_1, q_2 \in \mathbb{R}^d$  one has

$$(10) \quad \sum_{i,j=1}^d |a_{ij}(q_2) - a_{ij}(q_1)| \leq C|q_2 - q_1|.$$

The initial values for the SDE system (8) are given by  $q^\epsilon(0) = q_0^\epsilon$  and  $p^\epsilon(0) = p_0^\epsilon$ , such that Assumption 2 below is satisfied.

ASSUMPTION 2. *There exists  $q_0^0 \in \mathbb{R}^d$  such that*

$$q_0^\epsilon \xrightarrow{\epsilon \rightarrow 0} q_0^0.$$

Moreover, one has the following uniform upper bound:

$$\sup_{\epsilon \in (0, \epsilon_0)} |p_0^\epsilon| < \infty.$$

It is assumed that the initial values  $q_0^\epsilon \in \mathbb{R}^d$  and  $p_0^\epsilon \in \mathbb{R}^d$  are deterministic. The case of random initial values, independent of the Wiener process  $(\beta(t))_{t \geq 0}$ , can be treated by a standard conditioning argument, provided that suitable moment bounds are satisfied. This treatment is omitted in the sequel.

All the estimates below depend on the value of  $\sup_{\epsilon \in (0, \epsilon_0)} |p_0^\epsilon|$ , but this is not indicated explicitly.

Under Assumptions 1 and 2, the SDE system (8) admits a unique solution  $(q^\epsilon(t), p^\epsilon(t))_{t \geq 0}$ , since  $f$  and  $\sigma$  are globally Lipschitz continuous. The solution can be expressed as follows: for all  $t \geq 0$ , one has

$$(11) \quad \begin{cases} q^\epsilon(t) = q_0^\epsilon + \frac{1}{\epsilon} \int_0^t p^\epsilon(s) ds, \\ p^\epsilon(t) = e^{-\frac{t}{\epsilon^2}} p_0^\epsilon + \frac{1}{\epsilon} \int_0^t e^{-\frac{t-s}{\epsilon^2}} f(q^\epsilon(s)) ds + \frac{1}{\epsilon} \int_0^t e^{-\frac{t-s}{\epsilon^2}} d\beta(s). \end{cases}$$

The following change of unknowns is employed below: for all  $t \geq 0$ , set

$$(12) \quad \begin{cases} Q^\epsilon(t) = q^\epsilon(t) + \epsilon p^\epsilon(t) \\ P^\epsilon(t) = \epsilon p^\epsilon(t). \end{cases}$$

The  $\mathbb{R}^{2d}$ -valued process  $(Q^\epsilon(t), P^\epsilon(t))_{t \geq 0}$  is then solution of the SDE system

$$(13) \quad \begin{cases} dQ^\epsilon(t) = f(Q^\epsilon(t) - P^\epsilon(t)) dt + \sigma(Q^\epsilon(t) - P^\epsilon(t)) d\beta(t), \\ dP^\epsilon(t) = -\frac{P^\epsilon(t)}{\epsilon^2} dt + f(Q^\epsilon(t) - P^\epsilon(t)) dt + \sigma(Q^\epsilon(t) - P^\epsilon(t)) d\beta(t). \end{cases}$$

To retrieve properties of  $q^\epsilon(t)$ , note that for all  $t \geq 0$  one has

$$q^\epsilon(t) = Q^\epsilon(t) - P^\epsilon(t).$$

The change of unknowns is instrumental in the proofs of moment bounds and strong error estimates for  $Q^\epsilon$  and  $P^\epsilon$ , however weak error analysis is performed using only the unknowns  $q^\epsilon$  and  $p^\epsilon$ .

**2.2. The numerical schemes.** We introduce two numerical schemes: the first one is used to obtain strong approximation of  $q^\epsilon(t)$ , whereas the second one is used to obtain weak approximation of  $q^\epsilon(t)$ . The same notation is used for the two schemes, since it will always be clear in the statements of the results and in the analysis below which scheme is considered.

The first numerical scheme is defined as follows:

$$(14) \quad \begin{cases} q_{n+1}^{\epsilon, \Delta t} = q_n^{\epsilon, \Delta t} + \frac{\Delta t}{\epsilon} p_{n+1}^{\epsilon, \Delta t} \\ p_{n+1}^{\epsilon, \Delta t} = p_n^{\epsilon, \Delta t} - \frac{\Delta t}{\epsilon^2} p_{n+1}^{\epsilon, \Delta t} + \frac{\Delta t f(q_n^{\epsilon, \Delta t})}{\epsilon} + \frac{\sigma(q_n^{\epsilon, \Delta t})}{\epsilon} \Delta \beta_n \end{cases}$$

with the initial values  $q_0^{\epsilon, \Delta t} = q_0^\epsilon$  and  $p_0^{\epsilon, \Delta t} = p_0^\epsilon$  (given by Assumption 2). In the scheme (14), the  $q$ -component is discretized explicitly, whereas the  $p$ -component is treated implicitly. This choice is made to ensure stability properties (in particular to be able to choose the time-step size  $\Delta t$  independently of the time scale separation parameter  $\epsilon$ ) in the second equation, and to ensure good behavior of  $q_n^{\epsilon, \Delta t}$  when  $\epsilon$  vanishes, in the first equation, as will be explained below. In addition, discretizing the  $q$ -component explicitly is needed to have consistent approximation in the sense of Itô of the contribution of the noise. Note that, in fact, the scheme (14) can be implemented explicitly in practice owing to the following equivalent formulation, computing first  $p_{n+1}^{\epsilon, \Delta t}$  and then  $q_{n+1}^{\epsilon, \Delta t}$ :

$$\begin{cases} q_{n+1}^{\epsilon, \Delta t} = q_n^{\epsilon, \Delta t} + \frac{\Delta t}{\epsilon} p_{n+1}^{\epsilon, \Delta t} \\ p_{n+1}^{\epsilon, \Delta t} = \frac{1}{1 + \frac{\Delta t}{\epsilon^2}} \left( p_n^{\epsilon, \Delta t} + \frac{\Delta t f(q_n^{\epsilon, \Delta t})}{\epsilon} + \frac{\sigma(q_n^{\epsilon, \Delta t})}{\epsilon} \Delta \beta_n \right). \end{cases}$$

The second numerical scheme is defined as follows:

$$(15) \quad \begin{cases} q_{n+1}^{\epsilon, \Delta t} = q_n^{\epsilon, \Delta t} + \epsilon(1 - e^{-\frac{\Delta t}{\epsilon^2}}) p_n^{\epsilon, \Delta t} + (\Delta t - \epsilon^2(1 - e^{-\frac{\Delta t}{\epsilon^2}})) f(q_n^{\epsilon, \Delta t}) \\ \quad + \sigma(q_n^{\epsilon, \Delta t})(\beta(t_{n+1}) - \beta(t_n)) - \sigma(q_n^{\epsilon, \Delta t}) \int_{t_n}^{t_{n+1}} e^{-\frac{t_{n+1}-s}{\epsilon^2}} d\beta(s) \\ p_{n+1}^{\epsilon, \Delta t} = e^{-\frac{\Delta t}{\epsilon^2}} p_n^{\epsilon, \Delta t} + \epsilon(1 - e^{-\frac{\Delta t}{\epsilon^2}}) f(q_n^{\epsilon, \Delta t}) + \frac{1}{\epsilon} \sigma(q_n^{\epsilon, \Delta t}) \int_{t_n}^{t_{n+1}} e^{-\frac{t_{n+1}-s}{\epsilon^2}} d\beta(s). \end{cases}$$

The numerical scheme (15) is appropriate to obtain approximation in distribution of  $q^\epsilon(t_n)$  and  $p^\epsilon(t_n)$ : indeed, it suffices to sample at each iteration a  $\mathbb{R}^{2d}$ -valued centered Gaussian random variable

$$\left( \beta(t_{n+1}) - \beta(t_n), \int_{t_n}^{t_{n+1}} e^{-\frac{t_{n+1}-s}{\epsilon^2}} d\beta(s) \right),$$

or equivalently of a family of  $d$  independent  $\mathbb{R}^2$ -valued centered Gaussian random variables

$$\left( \beta_j(t_{n+1}) - \beta_j(t_n), \int_{t_n}^{t_{n+1}} e^{-\frac{t_{n+1}-s}{\epsilon^2}} d\beta_j(s) \right),$$

with  $j = 1, \dots, d$ , which have the same covariance matrix with entries given by

$$\begin{aligned} \mathbb{E}[(\beta_j(t_{n+1}) - \beta_j(t_n))^2] &= \Delta t \\ \mathbb{E}[(\beta_j(t_{n+1}) - \beta_j(t_n)) \int_{t_n}^{t_{n+1}} e^{-\frac{t_{n+1}-s}{\epsilon^2}} d\beta_j(s)] &= \int_{t_n}^{t_{n+1}} e^{-\frac{t_{n+1}-s}{\epsilon^2}} ds = \epsilon^2(1 - e^{-\frac{\Delta t}{\epsilon^2}}) \\ \mathbb{E}[(\int_{t_n}^{t_{n+1}} e^{-\frac{t_{n+1}-s}{\epsilon^2}} d\beta(s))^2] &= \int_{t_n}^{t_{n+1}} e^{-2\frac{t_{n+1}-s}{\epsilon^2}} ds = \frac{\epsilon^2}{2}(1 - e^{-\frac{2\Delta t}{\epsilon^2}}). \end{aligned}$$

It suffices to compute the square root or the Cholesky decomposition of the covariance matrix to sample the required Gaussian random variables. The construction of the second numerical scheme (15) is motivated by the following result.

PROPOSITION 2.1. Let  $\epsilon \in (0, \epsilon_0)$  and  $\Delta t \in (0, \Delta t_0)$ . Introduce the continuous-time auxiliary process  $(\tilde{q}^{\epsilon, \Delta t}(t), \tilde{p}^{\epsilon, \Delta t}(t))_{t \geq 0}$  defined such that for all  $n \in \{0, \dots, N-1\}$  and  $t \in [t_n, t_{n+1}]$ , one has

$$(16) \quad \begin{cases} d\tilde{q}^{\epsilon, \Delta t}(t) = \frac{\tilde{p}^{\epsilon, \Delta t}(t)}{\epsilon} dt \\ d\tilde{p}^{\epsilon, \Delta t}(t) = -\frac{\tilde{p}^{\epsilon, \Delta t}(t)}{\epsilon^2} dt + \frac{f(\tilde{q}^{\epsilon, \Delta t}(t_n))}{\epsilon} dt + \frac{\sigma(\tilde{q}^{\epsilon, \Delta t}(t_n))}{\epsilon} d\beta(t), \end{cases}$$

with initial values  $\tilde{q}^{\epsilon, \Delta t}(0) = q_0^\epsilon$  and  $\tilde{p}^{\epsilon, \Delta t}(0) = p_0^\epsilon$ , and such that  $t \in [0, T] \mapsto \tilde{q}^{\epsilon, \Delta t}(t)$  and  $t \in [0, T] \mapsto \tilde{p}^{\epsilon, \Delta t}(t)$  are continuous. Then for all  $n \in \{1, \dots, N\}$  one has

$$(q_n^{\epsilon, \Delta t}, p_n^{\epsilon, \Delta t}) = (\tilde{q}^{\epsilon, \Delta t}(t_n), \tilde{p}^{\epsilon, \Delta t}(t_n)).$$

PROOF OF PROPOSITION 2.1. Let  $n \in \{0, \dots, N-1\}$ , then for all  $t \in [t_n, t_{n+1}]$ , one has

$$\tilde{p}^{\epsilon, \Delta t}(t) = e^{-\frac{t-t_n}{\epsilon^2}} \tilde{p}^{\epsilon, \Delta t}(t_n) + \epsilon(1 - e^{-\frac{t-t_n}{\epsilon^2}}) f(\tilde{q}^{\epsilon, \Delta t}(t_n)) + \frac{1}{\epsilon} \sigma(\tilde{q}^{\epsilon, \Delta t}(t_n)) \int_{t_n}^t e^{-\frac{t-s}{\epsilon^2}} d\beta(s).$$

This then gives the equality

$$\begin{aligned} \tilde{q}^{\epsilon, \Delta t}(t_{n+1}) &= \tilde{q}^{\epsilon, \Delta t}(t_n) + \frac{1}{\epsilon} \int_{t_n}^{t_{n+1}} \tilde{p}^{\epsilon, \Delta t}(t) dt \\ &= \tilde{q}^{\epsilon, \Delta t}(t_n) + \epsilon(1 - e^{-\frac{\Delta t}{\epsilon^2}}) \tilde{p}^{\epsilon, \Delta t}(t_n) + (\Delta t - \epsilon^2(1 - e^{-\frac{\Delta t}{\epsilon^2}})) f(\tilde{q}^{\epsilon, \Delta t}(t_n)) \\ &\quad + \frac{1}{\epsilon^2} \sigma(\tilde{q}^{\epsilon, \Delta t}(t_n)) \int_{t_n}^{t_{n+1}} \int_{t_n}^t e^{-\frac{t-s}{\epsilon^2}} d\beta(s) dt, \end{aligned}$$

where using the stochastic Fubini theorem one obtains

$$\begin{aligned} \int_{t_n}^{t_{n+1}} \int_{t_n}^t e^{-\frac{t-s}{\epsilon^2}} d\beta(s) dt &= \int_{t_n}^{t_{n+1}} \int_s^{t_{n+1}} e^{-\frac{t-s}{\epsilon^2}} dt d\beta(s) \\ &= \int_{t_n}^{t_{n+1}} \epsilon^2 (1 - e^{-\frac{t_{n+1}-s}{\epsilon^2}}) d\beta(s) \\ &= \epsilon^2 (\beta(t_{n+1}) - \beta(t_n)) - \epsilon^2 \int_{t_n}^{t_{n+1}} e^{-\frac{t_{n+1}-s}{\epsilon^2}} d\beta(s). \end{aligned}$$

Since  $\tilde{q}^{\epsilon, \Delta t}(0) = q_0^{\epsilon, \Delta t}$  and  $\tilde{p}^{\epsilon, \Delta t}(0) = p_0^{\epsilon, \Delta t}$ , it is then straightforward to check that  $\tilde{q}^{\epsilon, \Delta t}(t_n) = q_n^{\epsilon, \Delta t}$  and  $\tilde{p}^{\epsilon, \Delta t}(t_n) = p_n^{\epsilon, \Delta t}$  for all  $n \in \{1, \dots, N\}$ . The proof of Proposition 2.1 is thus completed.  $\square$

As a consequence of Proposition 2.1, observe that if  $f$  and  $\sigma$  are constant, then the scheme (15) is exact:  $q_n^{\epsilon, \Delta t} = q^\epsilon(t_n)$  and  $p_n^{\epsilon, \Delta t} = p^\epsilon(t_n)$ . Using the implementation of the scheme explained above, this means that one obtains a scheme which is exact in distribution. The numerical scheme (15) is thus constructed by freezing the values of the  $q$ -component on each interval  $[t_n, t_{n+1}]$  when applying the mappings  $f$  and  $\sigma$ , and by computing the exact solution of the SDE depending on  $f(q_n^{\epsilon, \Delta t})$  and  $\sigma(q_n^{\epsilon, \Delta t})$  on each interval.

Note that the continuous auxiliary processes  $(\tilde{q}^{\epsilon, \Delta t}(t), \tilde{p}^{\epsilon, \Delta t}(t))_{t \geq 0}$  play a role below in the proof of the uniform weak error estimates for the numerical scheme (15).

Like in the continuous-time setting (see (12)), it is convenient to introduce auxiliary unknowns

$$(17) \quad \begin{cases} Q_n^{\epsilon, \Delta t} = q_n^{\epsilon, \Delta t} + \epsilon p_n^{\epsilon, \Delta t} \\ P_n^{\epsilon, \Delta t} = \epsilon p_n^{\epsilon, \Delta t}. \end{cases}$$

After proving some properties for the unknowns  $Q_n^{\epsilon, \Delta t}$  and  $P_n^{\epsilon, \Delta t}$ , the identity  $q_n^{\epsilon, \Delta t} = Q_n^{\epsilon, \Delta t} - P_n^{\epsilon, \Delta t}$  is then used to retrieve properties of the unknown  $q_n^{\epsilon, \Delta t}$ .

If the first numerical scheme (14) is used, the system after the change of variables reads

$$(18) \quad \begin{cases} Q_{n+1}^{\epsilon, \Delta t} = Q_n^{\epsilon, \Delta t} + \Delta t f(Q_n^{\epsilon, \Delta t} - P_n^{\epsilon, \Delta t}) + \sigma(Q_n^{\epsilon, \Delta t} - P_n^{\epsilon, \Delta t}) \Delta \beta_n \\ P_{n+1}^{\epsilon, \Delta t} = \frac{1}{1 + \frac{\Delta t}{\epsilon^2}} \left( P_n^{\epsilon, \Delta t} + \Delta t f(Q_n^{\epsilon, \Delta t} - P_n^{\epsilon, \Delta t}) + \sigma(Q_n^{\epsilon, \Delta t} - P_n^{\epsilon, \Delta t}) \Delta \beta_n \right). \end{cases}$$

If the second numerical scheme (15) is used, the system after the change of variables reads

$$(19) \quad \begin{cases} Q_{n+1}^{\epsilon, \Delta t} = Q_n^{\epsilon, \Delta t} + \Delta t f(Q_n^{\epsilon, \Delta t} - P_n^{\epsilon, \Delta t}) + \sigma(Q_n^{\epsilon, \Delta t} - P_n^{\epsilon, \Delta t}) \Delta \beta_n \\ P_{n+1}^{\epsilon, \Delta t} = e^{-\frac{\Delta t}{\epsilon^2}} P_n^{\epsilon, \Delta t} + \epsilon^2 (1 - e^{-\frac{\Delta t}{\epsilon^2}}) f(Q_n^{\epsilon, \Delta t} - P_n^{\epsilon, \Delta t}) + \sigma(Q_n^{\epsilon, \Delta t} - P_n^{\epsilon, \Delta t}) \int_{t_n}^{t_{n+1}} e^{-\frac{t_{n+1}-s}{\epsilon^2}} d\beta(s). \end{cases}$$

It is straightforward to check that the numerical schemes (14) and (15) give consistent strong and weak approximation respectively of the solution of the system (8) when  $\Delta t \rightarrow 0$ , for any fixed value of the time-scale separation parameter  $\epsilon \in (0, \epsilon_0)$ . Since the objective of this article is to prove that these schemes can be run with a cost independent of  $\epsilon$ , it is relevant to study the behavior when  $\epsilon \rightarrow 0$  of  $q^\epsilon(t)$  and of  $q_n^{\epsilon, \Delta t}$ .

**2.3. Asymptotic behavior when the time scale separation parameter vanishes.** Consider the stochastic differential equation

$$(20) \quad dq^0(t) = f(q^0(t))dt + \sigma(q^0(t))d\beta(t),$$

where  $q^0(t) \in \mathbb{R}^d$ , with initial value  $q^0(0) = q_0^0 = \lim_{\epsilon \rightarrow 0} q_0^\epsilon$  (see Assumption 2). Owing to Assumption 1,  $f$  and  $\sigma$  are globally Lipschitz continuous, therefore the SDE (20) admits a unique solution  $(q^0(t))_{t \geq 0}$ . This solution satisfies the identity

$$(21) \quad q^0(t) = q_0^0 + \int_0^t f(q^0(s))ds + \int_0^t \sigma(q^0(s))d\beta(s)$$

for all  $t \geq 0$ .

Consider also the standard Euler–Maruyama scheme applied to the SDE (20), with time-step size  $\Delta t$ : set  $q_0^{0, \Delta t} = q_0^0$  and for all  $n \in \{0, \dots, N-1\}$ , set

$$(22) \quad q_{n+1}^{0, \Delta t} = q_n^{0, \Delta t} + \Delta t f(q_n^{0, \Delta t}) + \sigma(q_n^{0, \Delta t}) \Delta \beta_n,$$

where we recall that  $\Delta \beta_n = \beta(t_{n+1}) - \beta(t_n)$ .

One has the following convergence result when  $\epsilon \rightarrow 0$ .

**PROPOSITION 2.2.** *Let Assumptions 1 and 2 be satisfied. For all  $T \in (0, \infty)$ , there exists  $C(T) \in (0, \infty)$  such that for all  $\epsilon \in (0, \epsilon_0)$ , one has*

$$(23) \quad \sup_{0 \leq t \leq T} \mathbb{E}[|q^\epsilon(t) - q^0(t)|^2] \leq C(T) \left( |q_0^\epsilon - q_0^0|^2 + \epsilon^2 (1 + |q_0^\epsilon|^2 + |p_0^\epsilon|^2) \right) \xrightarrow{\epsilon \rightarrow 0} 0.$$

In addition, for all  $\Delta t \in (0, \Delta t_0)$ , one has

$$(24) \quad \sup_{n=0, \dots, N} \mathbb{E}[|q_n^{\epsilon, \Delta t} - q_n^{0, \Delta t}|^2] \leq C(T) \left( |q_0^\epsilon - q_0^0|^2 + \epsilon^2 (1 + |q_0^\epsilon|^2 + |p_0^\epsilon|^2) \right) \xrightarrow{\epsilon \rightarrow 0} 0,$$

where  $(q_n^{\epsilon, \Delta t}, p_n^{\epsilon, \Delta t})_{n=0, \dots, N}$  is given either by (14) or by (15).

The proof of Proposition 2.2 is postponed to Section 4.4, since it requires moment bounds (uniform with respect to  $\epsilon$ ) which are stated and proved in Section 4.

### 3. Main results

We are now in position to state the main results of this article. First, in Section 3.1, we study strong error estimates when the numerical scheme (14) is used. Second, in Section 3.2, we study strong error estimates when the numerical scheme (15) is used.

**3.1. Uniform strong error estimates.** In this subsection, let us consider the numerical scheme (14). One has the following result concerning the strong error  $\mathbb{E}[|q_n^{\epsilon, \Delta t} - q^\epsilon(n\Delta t)|^2]$  for the  $q$ -component, when  $\Delta t \rightarrow 0$ .

**THEOREM 3.1.** *Let Assumptions 1 and 2 be satisfied, and let  $(q_n^{\epsilon, \Delta t})_{n \geq 0}$  be given by the numerical scheme (14). For all  $T \in (0, \infty)$ , there exists  $C(T) \in (0, \infty)$  such that for all  $\Delta t = T/N \in (0, \Delta t_0)$  and  $n \in \{1, \dots, N\}$ , one has*

$$(25) \quad \sup_{\epsilon \in (0, \epsilon_0)} \mathbb{E}[|q_n^{\epsilon, \Delta t} - q^\epsilon(n\Delta t)|^2] \leq C(T) (1 + |q_0^0|^2) \Delta t + \frac{C(T)}{(n+1)^2}.$$



The remarkable property of (25) is that the error estimate is uniform with respect to the time-scale separation parameter  $\epsilon$ . In fact, the proof provides a more precise error estimate, which is combined with Assumption 2 to obtain (25): one has

$$(26) \quad \mathbb{E}[|q_n^{\epsilon, \Delta t} - q^\epsilon(n\Delta t)|^2] \leq C(T)(1 + |q_0^\epsilon|^2 + |p_0^\epsilon|^2)\Delta t + \frac{C(T)\epsilon^2}{(n+1)^2}|p_0^\epsilon|^2.$$

As a consequence, one obtains the following uniform strong error estimates. First, choosing  $N = n$ , and using the fact that  $1/N^2 = \Delta t^2/T^2$ , one obtains the uniform strong error estimate at the final time  $T = N\Delta t$ :

$$(27) \quad \sup_{\epsilon \in (0, \epsilon_0)} \mathbb{E}[|q_N^{\epsilon, \Delta t} - q^\epsilon(N\Delta t)|^2] \leq C(T)(1 + |q_0^0|^2)\Delta t.$$

Second, if  $p_0^\epsilon = 0$ , one also obtains

$$\sup_{\epsilon \in (0, \epsilon_0)} \sup_{n=0, \dots, N} \mathbb{E}[|q_n^{\epsilon, \Delta t} - q^\epsilon(n\Delta t)|^2] \leq C(T)(1 + |q_0^0|^2)\Delta t,$$

where the error estimate is uniform with respect to both  $\epsilon \in (0, \epsilon_0)$  and  $n \in \{0, \dots, N\}$ .

Observe that by letting  $\epsilon \rightarrow 0$  in the strong error estimate (26) and using Proposition 2.2, one retrieves the standard strong error estimate for the Euler–Maruyama scheme (22) applied to the SDE (20): one has

$$(28) \quad \sup_{n=0, \dots, N} \mathbb{E}[|q_n^{0, \Delta t} - q^0(n\Delta t)|^2] \leq C(T)(1 + |q_0^0|^2)\Delta t.$$

In general, when  $\sigma$  is not constant, the error estimate (28) above is optimal: the Euler–Maruyama scheme has strong order of convergence equal to  $1/2$ , thus the order of convergence in Theorem 3.1 is also optimal in general. If the diffusion coefficient  $\sigma$  is constant, the strong order of convergence of the Euler–Maruyama scheme applied to the SDE (20) driven by additive noise is in fact equal to 1: one can replace  $\Delta t$  by  $\Delta t^2$  in the right-hand side of (28). However, it does not seem possible to improve the order of convergence in Theorem 3.1 when  $\sigma$  is assumed to be constant using the arguments of the proof: in that case Theorem 3.1 may not be optimal.

Let us also provide a strong error estimate for the second numerical scheme (15).

**PROPOSITION 3.2.** *Let Assumptions 1 and 2 be satisfied, and let  $(q_n^{\epsilon, \Delta t})_{n \geq 0}$  be given by the numerical scheme (15). For all  $T \in (0, \infty)$ , there exists  $C(T) \in (0, \infty)$  such that for all  $\Delta t = T/N \in (0, \Delta t_0)$ , one has*

$$(29) \quad \sup_{\epsilon \in (0, \epsilon_0)} \sup_{n=0, \dots, N} \mathbb{E}[|q_n^{\epsilon, \Delta t} - q^\epsilon(n\Delta t)|^2] \leq C(T)(1 + |q_0^0|^2)\Delta t.$$

The proof of Proposition 3.2 would follow the same strategy as the proof of Theorem 3.1. In fact, compared with the proof of Theorem 3.1 several of the error terms vanish, which is due to replacing  $\frac{1}{1 + \frac{\Delta t}{\epsilon^2}}$  by  $e^{-\frac{\Delta t}{\epsilon^2}}$  in the expressions. This explains why the extra term in (26) does not appear in (29). The details are omitted.

**3.2. Uniform weak error estimates.** In this subsection, let us consider the numerical scheme (15). One has the following result concerning the weak error  $|\mathbb{E}[\varphi(q_n^{\epsilon, \Delta t})] - \mathbb{E}[\varphi(q^\epsilon(n\Delta t))]|$  for the  $q$ -component, when  $\Delta t \rightarrow 0$ .

**THEOREM 3.3.** *Let Assumptions 1 and 2 be satisfied, and let  $(q_n^{\epsilon, \Delta t})_{n \geq 0}$  be given by the numerical scheme (15). For all  $T \in (0, \infty)$ , there exists  $C(T) \in (0, \infty)$  such that for all  $\Delta t = T/N \in (0, \Delta t_0)$ , any function  $\varphi : \mathbb{R}^d \rightarrow \mathbb{R}$  of class  $\mathcal{C}^3$  with bounded derivatives of order 1, 2, 3, and all  $\epsilon \in (0, \epsilon_0)$ , one has*

$$(30) \quad \sup_{n=0, \dots, N} |\mathbb{E}[\varphi(q_n^{\epsilon, \Delta t})] - \mathbb{E}[\varphi(q^\epsilon(n\Delta t))]| \leq C(T) \|\varphi\|_3 (1 + |q_0^0|^2) (\Delta t + \mathcal{R}(\epsilon, \Delta t)),$$

where the residual error term  $\mathcal{R}(\epsilon, \Delta t)$  is defined by

$$(31) \quad \mathcal{R}(\epsilon, \Delta t) = \int_0^{\Delta t} \frac{\epsilon}{\Delta t} (1 - e^{-\frac{t}{\epsilon^2}}) dt \geq 0$$

and satisfies the following inequality: there exists  $C \in (0, \infty)$  such that for all  $\Delta t \in (0, \Delta t_0)$  and  $\epsilon \in (0, \epsilon_0)$  one has

$$(32) \quad \mathcal{R}(\epsilon, \Delta t) \leq C \min(\epsilon, \Delta t^{\frac{1}{2}}, \frac{\Delta t}{\epsilon}).$$

The proof of the inequality (32) is straightforward: indeed, one has the inequalities  $0 \leq 1 - e^{-s} \leq 2$  and  $0 \leq 1 - e^{-s} \leq s$ , and therefore  $0 \leq 1 - e^{-s} \leq (2s)^{\frac{1}{2}}$ , for all  $s \geq 0$ . Observe that one obtains the uniform upper bound

$$\sup_{\epsilon \in (0, \epsilon_0)} \mathcal{R}(\epsilon, \Delta t) \leq C \Delta t^{\frac{1}{2}}$$

which is optimal: indeed, choosing  $\epsilon = \sqrt{\Delta t}$ , one has

$$\frac{1}{\Delta t^{\frac{1}{2}}} \mathcal{R}(\sqrt{\Delta t}, \Delta t) = \frac{1}{\Delta t} \int_0^{\Delta t} (1 - e^{-\frac{t}{\Delta t}}) dt = \int_0^1 (1 - e^{-s}) ds = e^{-1}$$

for all  $\Delta t \in (0, \Delta t_0)$ .

Let us state two immediate consequences of Theorem (3.3). On the one hand, one obtains the following uniform weak error estimate

$$(33) \quad \sup_{\epsilon \in (0, \epsilon_0)} \sup_{n=0, \dots, N} |\mathbb{E}[\varphi(q_n^{\epsilon, \Delta t})] - \mathbb{E}[\varphi(q^\epsilon(n\Delta t))]| \leq C(T) \|\varphi\|_3 (1 + |q_0^0|^2) \Delta t^{\frac{1}{2}}$$

where the order of convergence is equal to 1/2 (and this cannot be improved when using (30) owing to the observation above). On the other hand, letting  $\epsilon \rightarrow 0$ , one retrieves the standard weak error estimate with order 1 for the Euler–Maruyama scheme (22) applied to the SDE (20): one has

$$(34) \quad \sup_{n=0, \dots, N} |\mathbb{E}[\varphi(q_n^{0, \Delta t})] - \mathbb{E}[\varphi(q^0(n\Delta t))]| \leq C(T) \|\varphi\|_3 (1 + |q_0^0|^2) \Delta t.$$

For a fixed value of  $\epsilon$ , the weak error estimate (30) also gives order of convergence 1 with respect  $\Delta t$ , but the corresponding error estimate is not uniform with respect to  $\epsilon$ .

Note that it would suffice to apply (29) from Proposition 3.2 above to obtain the uniform weak error estimate (33), but to retrieve (34) one needs the refined error analysis which gives (30). Even if the order of convergence in the uniform weak error estimate (33) is 1/2, one has the following error estimate

$$(35) \quad \sup_{n=0, \dots, N} |\mathbb{E}[\varphi(q_n^{\epsilon, \Delta t})] - \mathbb{E}[\varphi(q^\epsilon(n\Delta t))]| \leq C(T) \|\varphi\|_3 (1 + |q_0^0|^2) (\Delta t + \epsilon),$$

owing to (30) and (32), which is relevant in situations where  $\epsilon$  is negligible compared with  $\Delta t$ .

## 4. Moment bounds

This section is devoted to state and prove moment bounds, which are uniform with respect to  $\epsilon \in (0, \epsilon_0)$  and  $\Delta t \in (0, \Delta t_0)$ , for the solutions of (8), of (14) and of (15). The proofs are based on the changes of unknowns introduced in Sections (2.1) and (2.2). Even if the arguments are mostly elementary, it is worth giving full details for completeness. The proofs of Propositions 4.2 and 4.3 are variants of the proof of Proposition 4.1 in discrete-time situations. After proving the moment bounds, the proof of Proposition 2.2 is provided.

### 4.1. Moment bounds for the SDE system.

PROPOSITION 4.1. *Let Assumption 1 be satisfied. For all  $T \in (0, \infty)$ , there exists  $C(T) \in (0, \infty)$  such that for all  $\epsilon \in (0, \epsilon_0)$ , one has*

$$(36) \quad \sup_{0 \leq t \leq T} \left( \mathbb{E}[|q^\epsilon(t)|^2] + \mathbb{E}[|p^\epsilon(t)|^2] \right) \leq C(T) (1 + |q_0^\epsilon|^2 + |p_0^\epsilon|^2).$$

PROOF OF PROPOSITION 4.1. Let  $(Q^\epsilon(t), P^\epsilon(t))_{t \geq 0}$  be the solution of the SDE system (13). For all  $t \in [0, T]$ , one has

$$(37) \quad \begin{cases} Q^\epsilon(t) = Q^\epsilon(0) + \int_0^t f(Q^\epsilon(s) - P^\epsilon(s)) ds + \int_0^t \sigma(Q^\epsilon(s) - P^\epsilon(s)) d\beta(s) \\ P^\epsilon(t) = e^{-\frac{t}{\epsilon^2}} P^\epsilon(0) + \int_0^t e^{-\frac{t-s}{\epsilon^2}} f(Q^\epsilon(s) - P^\epsilon(s)) ds + \int_0^t e^{-\frac{t-s}{\epsilon^2}} \sigma(Q^\epsilon(s) - P^\epsilon(s)) d\beta(s). \end{cases}$$

Using Itô's isometry formula, and the Lipschitz continuity property of the mappings  $f$  and  $\sigma$  (Assumption 1) one obtains the inequality

$$\mathbb{E}[|Q^\epsilon(t)|^2] + \mathbb{E}[|P^\epsilon(t)|^2] \leq C(T) \left( 1 + \mathbb{E}[|Q^\epsilon(0)|^2] + \mathbb{E}[|P^\epsilon(0)|^2] + \int_0^t (\mathbb{E}[|Q^\epsilon(s)|^2] + \mathbb{E}[|P^\epsilon(s)|^2]) ds \right).$$

for all  $t \in [0, T]$ . Since  $Q^\epsilon(0) = q_0^\epsilon + \epsilon p_0^\epsilon$  and  $P^\epsilon(0) = \epsilon p_0^\epsilon$ , applying Gronwall's lemma yields the moment bound

$$\sup_{0 \leq t \leq T} (\mathbb{E}[|Q^\epsilon(t)|^2] + \mathbb{E}[|P^\epsilon(t)|^2]) \leq C(T) \left( 1 + |q_0^\epsilon|^2 + \|p_0^\epsilon\|^2 \right).$$

Using the identity  $q^\epsilon(t) = Q^\epsilon(t) - P^\epsilon(t)$  then gives

$$\sup_{0 \leq t \leq T} \mathbb{E}[|q^\epsilon(t)|^2] \leq C(T) \left( 1 + |q_0^\epsilon|^2 + |p_0^\epsilon|^2 \right).$$

Using the second identity from (11) and writing  $Q^\epsilon(t) - P^\epsilon(t) = q^\epsilon(t)$ , applying Itô's isometry formula, and using the Lipschitz continuity properties of  $f$  and  $\sigma$  and the moment bounds for  $q^\epsilon(t)$  obtained above, one then obtains, for all  $t \in [0, T]$ ,

$$\begin{aligned} \mathbb{E}[|p^\epsilon(t)|^2] &\leq C|p_0^\epsilon|^2 + \frac{CT}{\epsilon^2} \int_0^t e^{-\frac{2(t-s)}{\epsilon^2}} (1 + \mathbb{E}[|q^\epsilon(s)|^2]) ds \\ &\leq C(T) \left( 1 + |q_0^\epsilon|^2 + |p_0^\epsilon|^2 \right) \left( 1 + \frac{1}{\epsilon^2} \int_0^t e^{-\frac{2(t-s)}{\epsilon^2}} ds \right) \\ &\leq C(T) \left( 1 + |q_0^\epsilon|^2 + |p_0^\epsilon|^2 \right), \end{aligned}$$

which gives the required moment bounds for  $p^\epsilon(t)$ , using the identity

$$\frac{1}{\epsilon^2} \int_0^\infty e^{-\frac{2r}{\epsilon^2}} dr = \frac{1}{2}.$$

This concludes the proof of Proposition 4.1. □

#### 4.2. Moment bounds for the first numerical scheme.

PROPOSITION 4.2. Let  $(q_n^{\epsilon, \Delta t}, p_n^{\epsilon, \Delta t})_{n=0, \dots, N}$  be given by the numerical scheme (14). Let Assumption 1 be satisfied. For all  $T \in (0, \infty)$ , there exists  $C(T) \in (0, \infty)$  such that for all  $\epsilon \in (0, \epsilon_0)$  and  $\Delta t \in (0, \Delta t_0)$ , one has

$$(38) \quad \sup_{n=0, \dots, N} \left( \mathbb{E}[|q_n^{\epsilon, \Delta t}|^2] + \mathbb{E}[|p_n^{\epsilon, \Delta t}|^2] \right) \leq C(T) (1 + |q_0^\epsilon|^2 + |p_0^\epsilon|^2).$$

PROOF OF PROPOSITION 4.2. Let  $(Q_n^{\epsilon, \Delta t}, P_n^{\epsilon, \Delta t})_{n=0, \dots, N}$  be given by (18). One has the identities

$$(39) \quad \begin{cases} Q_n^{\epsilon, \Delta t} = Q_0^{\epsilon, \Delta t} + \Delta t \sum_{k=0}^{n-1} f(Q_k^{\epsilon, \Delta t} - P_k^{\epsilon, \Delta t}) + \sum_{k=0}^{n-1} \sigma(Q_k^{\epsilon, \Delta t} - P_k^{\epsilon, \Delta t}) \Delta \beta_k \\ P_n^{\epsilon, \Delta t} = \frac{1}{(1 + \frac{\Delta t}{\epsilon^2})^n} P_0^{\epsilon, \Delta t} + \Delta t \sum_{k=0}^{n-1} \frac{1}{(1 + \frac{\Delta t}{\epsilon^2})^{n-k}} f(Q_k^{\epsilon, \Delta t} - P_k^{\epsilon, \Delta t}) + \sum_{k=0}^{n-1} \frac{1}{(1 + \frac{\Delta t}{\epsilon^2})^{n-k}} \sigma(Q_k^{\epsilon, \Delta t} - P_k^{\epsilon, \Delta t}) \Delta \beta_k \end{cases}$$

for all  $n \in \{0, \dots, N\}$ . Using Itô's isometry formula and the Lipschitz continuity property of the mappings  $f$  and  $\sigma$  (Assumption 1), one obtains the inequality

$$\mathbb{E}[|Q_n^{\epsilon, \Delta t}|^2] + \mathbb{E}[|P_n^{\epsilon, \Delta t}|^2] \leq C(T) \left( 1 + \mathbb{E}[|Q_0^{\epsilon, \Delta t}|^2] + \mathbb{E}[|P_0^{\epsilon, \Delta t}|^2] + \Delta t \sum_{k=0}^{n-1} (\mathbb{E}[|Q_k^{\epsilon, \Delta t}|^2] + \mathbb{E}[|P_k^{\epsilon, \Delta t}|^2]) \right)$$

for all  $n \in \{0, \dots, N\}$ . Since  $Q_0^{\epsilon, \Delta t} = q_0^\epsilon + \epsilon p_0^\epsilon$  and  $P_0^{\epsilon, \Delta t} = \epsilon p_0^\epsilon$ , applying discrete Gronwall's lemma yields the moment bound

$$\sup_{n=0, \dots, N} \left( \mathbb{E}[|Q_n^{\epsilon, \Delta t}|^2] + \mathbb{E}[|P_n^{\epsilon, \Delta t}|^2] \right) \leq C(T) (1 + |q_0^\epsilon|^2 + |p_0^\epsilon|^2).$$

Using the identity  $q_n^{\epsilon, \Delta t} = Q_n^{\epsilon, \Delta t} - P_n^{\epsilon, \Delta t}$  then gives

$$\sup_{n=0, \dots, N} \mathbb{E}[|q_n^{\epsilon, \Delta t}|^2] \leq C(T) (1 + |q_0^\epsilon|^2 + |p_0^\epsilon|^2).$$

Using the identity  $p_n^{\epsilon, \Delta t} = P_n^{\epsilon, \Delta t} / \epsilon$ , one has

$$p_n^{\epsilon, \Delta t} = \frac{1}{\left(1 + \frac{\Delta t}{\epsilon^2}\right)^n} p_0^\epsilon + \frac{\Delta t}{\epsilon} \sum_{k=0}^{n-1} \frac{1}{\left(1 + \frac{\Delta t}{\epsilon^2}\right)^{n-k}} f(q_k^{\epsilon, \Delta t}) + \frac{1}{\epsilon} \sum_{k=0}^{n-1} \frac{1}{\left(1 + \frac{\Delta t}{\epsilon^2}\right)^{n-k}} \sigma(q_k^{\epsilon, \Delta t}) \Delta \beta_k.$$

Applying Itô's isometry formula, and using the Lipschitz continuity properties of  $f$  and  $\sigma$  and the moment bounds for  $q_n^{\epsilon, \Delta t}$  obtained above, one has, for all  $n \in \{0, \dots, N\}$

$$\begin{aligned} \mathbb{E}[|p_n^{\epsilon, \Delta t}|^2] &\leq C |p_0^\epsilon|^2 + \frac{CT\Delta t}{\epsilon^2} \sum_{k=0}^{n-1} \frac{1}{\left(1 + \frac{\Delta t}{\epsilon^2}\right)^{2(n-k)}} (1 + \mathbb{E}[|q_k^{\epsilon, \Delta t}|^2]) \\ &\leq C(T) (1 + |q_0^\epsilon|^2 + |p_0^\epsilon|^2) \left( 1 + \frac{\Delta t}{\epsilon^2} \sum_{k=0}^{n-1} \frac{1}{\left(1 + \frac{\Delta t}{\epsilon^2}\right)^{2(n-k)}} \right) \\ &\leq C(T) (1 + |q_0^\epsilon|^2 + |p_0^\epsilon|^2), \end{aligned}$$

which gives the required moment bounds for  $p_n^\epsilon$ , using the inequality

$$\tau \sum_{\ell=1}^{\infty} \frac{1}{(1 + \tau)^{2\ell}} = \frac{1}{2 + \tau} \leq \frac{1}{2}$$

with  $\tau = \Delta t / \epsilon^2 \geq 0$ . This concludes the proof of Proposition 4.2.  $\square$

### 4.3. Moment bounds for the second numerical scheme.

**PROPOSITION 4.3.** *Let  $(q_n^{\epsilon, \Delta t}, p_n^{\epsilon, \Delta t})_{n=0, \dots, N}$  be given by the numerical scheme (15). Let Assumption 1 be satisfied. For all  $T \in (0, \infty)$ , there exists  $C(T) \in (0, \infty)$  such that for all  $\epsilon \in (0, \epsilon_0)$  and  $\Delta t \in (0, \Delta t_0)$ , one has*

$$(40) \quad \sup_{n=0, \dots, N} \left( \mathbb{E}[|q_n^{\epsilon, \Delta t}|^2] + \mathbb{E}[|p_n^{\epsilon, \Delta t}|^2] \right) \leq C(T) (1 + |q_0^\epsilon|^2 + |p_0^\epsilon|^2).$$

**PROOF OF PROPOSITION 4.2.** Let  $(Q_n^{\epsilon, \Delta t}, P_n^{\epsilon, \Delta t})_{n=0, \dots, N}$  be given by (19). One has the identities

$$\begin{aligned} Q_n^{\epsilon, \Delta t} &= Q_0^{\epsilon, \Delta t} + \Delta t \sum_{k=0}^{n-1} f(Q_k^{\epsilon, \Delta t} - P_k^{\epsilon, \Delta t}) + \sum_{k=0}^{n-1} \sigma(Q_k^{\epsilon, \Delta t} - P_k^{\epsilon, \Delta t}) \Delta \beta_k \\ P_n^{\epsilon, \Delta t} &= e^{-\frac{n\Delta t}{\epsilon^2}} P_0^{\epsilon, \Delta t} + \Delta t \sum_{k=0}^{n-1} e^{-\frac{(n-k)\Delta t}{\epsilon^2}} f(Q_k^{\epsilon, \Delta t} - P_k^{\epsilon, \Delta t}) + \sum_{k=0}^{n-1} e^{-\frac{(n-k)\Delta t}{\epsilon^2}} \sigma(Q_k^{\epsilon, \Delta t} - P_k^{\epsilon, \Delta t}) \Delta \beta_k \end{aligned}$$

for all  $n \in \{0, \dots, N\}$ . Using Itô's isometry formula and the Lipschitz continuity property of the mappings  $f$  and  $\sigma$  (Assumption 1), one obtains the inequality

$$\mathbb{E}[|Q_n^{\epsilon, \Delta t}|^2] + \mathbb{E}[|P_n^{\epsilon, \Delta t}|^2] \leq C(T) \left( 1 + \mathbb{E}[|Q_0^{\epsilon, \Delta t}|^2] + \mathbb{E}[|P_0^{\epsilon, \Delta t}|^2] + \Delta t \sum_{k=0}^{n-1} (\mathbb{E}[|Q_k^{\epsilon, \Delta t}|^2] + \mathbb{E}[|P_k^{\epsilon, \Delta t}|^2]) \right)$$

for all  $n \in \{0, \dots, N\}$ . Since  $Q_0^{\epsilon, \Delta t} = q_0^\epsilon + \epsilon p_0^\epsilon$  and  $P_0^{\epsilon, \Delta t} = \epsilon p_0^\epsilon$ , applying discrete Gronwall's lemma yields the moment bound

$$\sup_{n=0, \dots, N} \left( \mathbb{E}[|Q_n^{\epsilon, \Delta t}|^2] + \mathbb{E}[|P_n^{\epsilon, \Delta t}|^2] \right) \leq C(T)(1 + |q_0^\epsilon|^2 + |p_0^\epsilon|^2).$$

Using the identity  $q_n^{\epsilon, \Delta t} = Q_n^{\epsilon, \Delta t} - P_n^{\epsilon, \Delta t}$  then gives

$$\sup_{n=0, \dots, N} \mathbb{E}[|q_n^{\epsilon, \Delta t}|^2] \leq C(T)(1 + |q_0^\epsilon|^2 + |p_0^\epsilon|^2).$$

Using the identity  $p_n^{\epsilon, \Delta t} = P_n^{\epsilon, \Delta t}/\epsilon$ , one has

$$p_n^{\epsilon, \Delta t} = e^{-\frac{n\Delta t}{\epsilon^2}} p_0^\epsilon + \frac{\Delta t}{\epsilon} \sum_{k=0}^{n-1} e^{-\frac{(n-k)\Delta t}{\epsilon^2}} f(q_k^{\epsilon, \Delta t}) + \frac{1}{\epsilon} \sum_{k=0}^{n-1} e^{-\frac{(n-k)\Delta t}{\epsilon^2}} \sigma(q_k^{\epsilon, \Delta t}) \Delta \beta_k.$$

Applying Itô's isometry formula, and using the Lipschitz continuity properties of  $f$  and  $\sigma$  and the moment bounds for  $q_n^{\epsilon, \Delta t}$  obtained above, one has, for all  $n \in \{0, \dots, N\}$

$$\begin{aligned} \mathbb{E}[|p_n^{\epsilon, \Delta t}|^2] &\leq C|p_0^\epsilon|^2 + \frac{CT\Delta t}{\epsilon^2} \sum_{k=0}^{n-1} e^{-\frac{2(n-k)\Delta t}{\epsilon^2}} (1 + \mathbb{E}[|q_k^{\epsilon, \Delta t}|^2]) \\ &\leq C(T)(1 + |q_0^\epsilon|^2 + |p_0^\epsilon|^2) \left(1 + \frac{\Delta t}{\epsilon^2} \sum_{k=0}^{n-1} e^{-\frac{2(n-k)\Delta t}{\epsilon^2}}\right) \\ &\leq C(T)(1 + |q_0^\epsilon|^2 + |p_0^\epsilon|^2), \end{aligned}$$

which gives the required moment bounds for  $p_n^\epsilon$ , using the inequality

$$\tau \sum_{\ell=1}^{\infty} e^{-2\tau\ell} = \frac{\tau}{e^{2\tau} - 1} \leq C < \infty$$

with  $\tau = \Delta t/\epsilon^2 \geq 0$ . This concludes the proof of Proposition 4.2.  $\square$

**4.4. Proof of Proposition 2.2.** As a consequence of Proposition 4.1, we are now in position to provide the proof of the inequalities (23) and (24) from Proposition 2.2.

PROOF OF THE INEQUALITY (23). For all  $t \in [0, T]$ , the error is decomposed as

$$\begin{aligned} q^\epsilon(t) - q^0(t) &= Q^\epsilon(t) - q^0(t) - P^\epsilon(t) \\ &= q_0^\epsilon - q_0^0 + \int_0^t (f(q^\epsilon(s)) - f(q^0(s))) ds + \int_0^t (\sigma(q^\epsilon(s)) - \sigma(q^0(s))) d\beta(s) - \epsilon p^\epsilon(t). \end{aligned}$$

Applying Itô's isometry formula, using the Lipschitz continuity property of  $f$  and  $\sigma$  and the moment bound (36), one obtains, for all  $t \in [0, T]$ ,

$$\mathbb{E}[|q^\epsilon(t) - q^0(t)|^2] \leq C(T) \left( |q_0^\epsilon - q_0^0|^2 + \int_0^t \mathbb{E}[|q^\epsilon(s) - q^0(s)|^2] ds + \epsilon^2 (1 + |q_0^\epsilon|^2 + |p_0^\epsilon|^2) \right).$$

Applying Gronwall's lemma and using Assumption 2, the proof of the inequality (23) is completed.  $\square$

The proof of the inequality (24) is identical for the two choices of numerical schemes (14) and (15), and is a variant of the proof of the inequality (23) in a discrete time situation.

PROOF OF THE INEQUALITY (24). For all  $n \in \{0, \dots, N-1\}$ , the error is decomposed as

$$\begin{aligned} q_n^{\epsilon, \Delta t} - q_n^0, \Delta t &= Q_n^{\epsilon, \Delta t} - q_n^0, \Delta t - P_n^{\epsilon, \Delta t} \\ &= q_0^\epsilon - q_0^0 + \Delta t \sum_{k=0}^{n-1} (f(q_k^{\epsilon, \Delta t}) - f(q_k^0, \Delta t)) + \sum_{k=0}^{n-1} (\sigma(q_k^{\epsilon, \Delta t}) - \sigma(q_k^0, \Delta t)) - \epsilon p_n^{\epsilon, \Delta t}. \end{aligned}$$

Applying Itô's isometry formula, using the Lipschitz continuity property of  $f$  and  $\sigma$  and the moment bound (38), one obtains, for all  $n \in \{0, \dots, N-1\}$ ,

$$\mathbb{E}[|q_n^{\epsilon, \Delta t} - q_n^{0, \Delta t}|^2] \leq C(T) \left( |q_0^\epsilon - q_0^0|^2 + \Delta t \sum_{k=0}^{n-1} \mathbb{E}[|q_k^{\epsilon, \Delta t} - q_k^{0, \Delta t}|^2] + \epsilon^2 (1 + |q_0^\epsilon|^2 + |p_0^\epsilon|^2) \right).$$

Applying Gronwall's lemma and using Assumption 2, the proof of the inequality (24) is completed.  $\square$

## 5. Proof of the strong error estimates

Before proceeding with the proof of Theorem 3.1, let us state two useful inequalities:

$$(41) \quad \sup_{\tau \in [0, \infty)} \frac{|1 - e^{-\tau}|}{\tau^{\frac{1}{2}}} < \infty,$$

and

$$(42) \quad \sup_{n \in \mathbb{N}} \sup_{\tau \in (0, \infty)} (n+1) \left( \frac{1}{(1+\tau)^n} - e^{-n\tau} \right) < \infty.$$

The proof of the inequality (41) is straightforward: the mapping  $\tau \in [0, \infty) \rightarrow e^{-\tau}$  is bounded and Lipschitz continuous. To prove the inequality (42), it suffices to check that the maximum of the function

$$\tau \in [0, \infty) \mapsto \frac{1}{(1+\tau)^n} - e^{-n\tau} \in [0, \infty)$$

is attained for a real number  $\tau = \tau_n$  satisfying  $e^{-n\tau_n} = \frac{1}{(1+\tau_n)^{n+1}}$ : as a consequence

$$\sup_{\tau \in (0, \infty)} \left( \frac{1}{(1+\tau)^n} - e^{-n\tau} \right) = \frac{1}{(1+\tau_n)^n} - e^{-n\tau_n} = \frac{\tau_n}{(1+\tau_n)^{n+1}} \leq \frac{\tau_n}{(n+1)\tau_n} \leq \frac{1}{n+1}.$$

Before proceeding with the proof, let us also state and prove the following auxiliary result, concerning temporal regularity of the processes  $Q^\epsilon$  and  $P^\epsilon$ , uniformly with respect to  $\epsilon$ .

LEMMA 5.1. *Let Assumption 1 be satisfied. For all  $T \in (0, \infty)$ , there exists  $C(T) \in (0, \infty)$  such that for all  $\epsilon \in (0, \epsilon_0)$  and all  $s_1, s_2 \in [0, T]$ , one has*

$$(43) \quad \mathbb{E}[|Q^\epsilon(s_2) - Q^\epsilon(s_1)|^2] + \mathbb{E}[|P^\epsilon(s_2) - P^\epsilon(s_1)|^2] \leq C(T) (1 + |q_0^\epsilon|^2 + |p_0^\epsilon|^2) |s_2 - s_1|.$$

PROOF OF LEMMA 5.1. On the one hand, for all  $s_1, s_2 \in [0, T]$ , with  $s_1 \leq s_2$ , one has

$$Q^\epsilon(s_2) - Q^\epsilon(s_1) = \int_{s_1}^{s_2} f(q^\epsilon(s)) ds + \int_{s_1}^{s_2} \sigma(q^\epsilon(s)) d\beta(s),$$

therefore using Itô's isometry formula, the Lipschitz continuity property of the mappings  $f$  and  $\sigma$  (Assumption 1) and the moment bounds (36) from Proposition 4.1, one obtains

$$\mathbb{E}[|Q^\epsilon(s_2) - Q^\epsilon(s_1)|^2] \leq C(T) \int_{s_1}^{s_2} (1 + \mathbb{E}[|q^\epsilon(s)|^2]) ds \leq C(T) (1 + |q_0^\epsilon|^2 + |p_0^\epsilon|^2) |s_2 - s_1|.$$

On the one hand, for all  $s_1, s_2 \in [0, T]$ , with  $s_1 \leq s_2$ , one has

$$P^\epsilon(s_2) - P^\epsilon(s_1) = \left( e^{-\frac{s_2 - s_1}{\epsilon^2}} - 1 \right) P^\epsilon(s_1) + \int_{s_1}^{s_2} e^{-\frac{s_2 - s}{\epsilon^2}} f(q^\epsilon(s)) ds + \int_{s_1}^{s_2} e^{-\frac{s_2 - s}{\epsilon^2}} \sigma(q^\epsilon(s)) d\beta(s).$$

Using the inequality (41), the identity  $P^\epsilon(s_1) = \epsilon P^\epsilon(s_1)$ , Itô's isometry formula, the Lipschitz continuity property of the mappings  $f$  and  $\sigma$  (Assumption 1) and the moment bounds (36) from Proposition 4.1, one obtains

$$\begin{aligned} \mathbb{E}[|P^\epsilon(s_2) - P^\epsilon(s_1)|^2] &\leq C |s_2 - s_1| \mathbb{E}[|p^\epsilon(s_1)|^2] + C(T) \int_{s_1}^{s_2} (1 + \mathbb{E}[|q^\epsilon(s)|^2]) ds \\ &\leq C(T) (1 + |q_0^\epsilon|^2 + |p_0^\epsilon|^2) |s_2 - s_1|. \end{aligned}$$

Gathering the estimates then concludes the proof of Lemma 5.1.  $\square$

We are now in position to provide the proof of Theorem 3.1

PROOF OF THEOREM 3.1. Let us first describe the decomposition of the error. Using the identities

$$q^\epsilon(t_n) = Q^\epsilon(t_n) - P^\epsilon(t_n), \quad q_n^{\epsilon, \Delta t} = Q_n^{\epsilon, \Delta t} - P_n^{\epsilon, \Delta t},$$

the mean-square error is bounded as follows:

$$\mathbb{E}[|q_n^{\epsilon, \Delta t} - q^\epsilon(t_n)|^2] \leq 2\mathbb{E}[|Q_n^{\epsilon, \Delta t} - Q^\epsilon(t_n)|^2] + 2\mathbb{E}[|P_n^{\epsilon, \Delta t} - P^\epsilon(t_n)|^2].$$

Let  $E_n^{\epsilon, \Delta t} = \mathbb{E}[|Q_n^{\epsilon, \Delta t} - Q^\epsilon(t_n)|^2] + \mathbb{E}[|P_n^{\epsilon, \Delta t} - P^\epsilon(t_n)|^2]$ . Recall the expressions (37) for  $Q^\epsilon(t_n)$  and  $P^\epsilon(t_n)$ , and the expressions (39) for  $q_n^\epsilon$  and  $p_n^\epsilon$ .

On the one hand, one has

$$\begin{aligned} Q_n^{\epsilon, \Delta t} - Q^\epsilon(t_n) &= Q_0^{\epsilon, \Delta t} - Q^\epsilon(0) \\ &\quad + \Delta t \sum_{k=0}^{n-1} f(Q_k^{\epsilon, \Delta t} - P_k^{\epsilon, \Delta t}) - \int_0^{t_n} f(Q^\epsilon(t) - P^\epsilon(t)) dt \\ &\quad + \sum_{k=0}^{n-1} \sigma(Q_k^{\epsilon, \Delta t} - P_k^{\epsilon, \Delta t}) \Delta \beta_k - \int_0^{t_n} \sigma(Q^\epsilon(t) - P^\epsilon(t)) d\beta(t) \\ &= e_{n, Q, 1, 1}^{\epsilon, \Delta t} + e_{n, Q, 1, 2}^{\epsilon, \Delta t} + e_{n, Q, 2, 1}^{\epsilon, \Delta t} + e_{n, Q, 2, 2}^{\epsilon, \Delta t}, \end{aligned}$$

where one sets

$$\begin{aligned} e_{n, Q, 1, 1}^{\epsilon, \Delta t} &= \sum_{k=0}^{n-1} \int_{t_k}^{t_{k+1}} (f(Q^\epsilon(t_k) - P^\epsilon(t_k)) - f(Q^\epsilon(t) - P^\epsilon(t))) dt \\ e_{n, Q, 1, 2}^{\epsilon, \Delta t} &= \Delta t \sum_{k=0}^{n-1} (f(Q_k^{\epsilon, \Delta t} - P_k^{\epsilon, \Delta t}) - f(Q^\epsilon(t_k) - P^\epsilon(t_k))) \\ e_{n, Q, 2, 1}^{\epsilon, \Delta t} &= \sum_{k=0}^{n-1} \int_{t_k}^{t_{k+1}} (\sigma(Q^\epsilon(t_k) - P^\epsilon(t_k)) - \sigma(Q^\epsilon(t) - P^\epsilon(t))) d\beta(t) \\ e_{n, Q, 2, 2}^{\epsilon, \Delta t} &= \sum_{k=0}^{n-1} (\sigma(Q_k^{\epsilon, \Delta t} - P_k^{\epsilon, \Delta t}) - \sigma(Q^\epsilon(t_k) - P^\epsilon(t_k))) \Delta \beta_k. \end{aligned}$$

On the other hand, one has

$$\begin{aligned} P_n^{\epsilon, \Delta t} - P^\epsilon(t_n) &= \frac{1}{(1 + \frac{\Delta t}{\epsilon^2})^n} P_0^{\epsilon, \Delta t} - e^{-\frac{t_n}{\epsilon^2}} P_0^\epsilon \\ &\quad + \Delta t \sum_{k=0}^{n-1} \frac{1}{(1 + \frac{\Delta t}{\epsilon^2})^{n-k}} f(Q_k^{\epsilon, \Delta t} - P_k^{\epsilon, \Delta t}) - \int_0^{t_n} e^{-\frac{t_n-t}{\epsilon^2}} f(Q^\epsilon(t) - P^\epsilon(t)) dt \\ &\quad + \sum_{k=0}^{n-1} \frac{1}{(1 + \frac{\Delta t}{\epsilon^2})^{n-k}} \sigma(Q_k^{\epsilon, \Delta t} - P_k^{\epsilon, \Delta t}) \Delta \beta_k - \int_0^{t_n} e^{-\frac{t_n-t}{\epsilon^2}} \sigma(Q^\epsilon(t) - P^\epsilon(t)) d\beta(t) \\ &= e_{n, P, 0}^{\epsilon, \Delta t} \\ &\quad + e_{n, P, 1, 1}^{\epsilon, \Delta t} + e_{n, P, 1, 2}^{\epsilon, \Delta t} + e_{n, P, 1, 3}^{\epsilon, \Delta t} + e_{n, P, 1, 4}^{\epsilon, \Delta t} \\ &\quad + e_{n, P, 2, 1}^{\epsilon, \Delta t} + e_{n, P, 2, 2}^{\epsilon, \Delta t} + e_{n, P, 2, 3}^{\epsilon, \Delta t} + e_{n, P, 2, 4}^{\epsilon, \Delta t} \end{aligned}$$

where one sets

$$e_{n, P, 0}^{\epsilon, \Delta t} = \left( \frac{1}{(1 + \frac{\Delta t}{\epsilon^2})^n} - e^{-\frac{t_n}{\epsilon^2}} \right) P_0^{\epsilon, \Delta t} = \epsilon \left( \frac{1}{(1 + \frac{\Delta t}{\epsilon^2})^n} - e^{-\frac{t_n}{\epsilon^2}} \right) p_0^\epsilon,$$

then

$$\begin{aligned}
e_{n,P,1,1}^{\epsilon,\Delta t} &= \sum_{k=0}^{n-1} \int_{t_k}^{t_{k+1}} e^{-\frac{t_n-t}{\epsilon^2}} (f(Q^\epsilon(t_k) - P^\epsilon(t_k)) - f(Q^\epsilon(t) - P^\epsilon(t))) dt \\
e_{n,P,1,2}^{\epsilon,\Delta t} &= \sum_{k=0}^{n-1} \int_{t_k}^{t_{k+1}} e^{-\frac{t_n-t}{\epsilon^2}} (f(Q_k^{\epsilon,\Delta t} - P_k^{\epsilon,\Delta t}) - f(Q^\epsilon(t_k) - P^\epsilon(t_k))) dt \\
e_{n,P,1,3}^{\epsilon,\Delta t} &= \sum_{k=0}^{n-1} \int_{t_k}^{t_{k+1}} (e^{-\frac{t_n-t_k}{\epsilon^2}} - e^{-\frac{t_n-t}{\epsilon^2}}) f(Q_k^{\epsilon,\Delta t} - P_k^{\epsilon,\Delta t}) dt \\
e_{n,P,1,4}^{\epsilon,\Delta t} &= \Delta t \sum_{k=0}^{n-1} \left( \frac{1}{(1 + \frac{\Delta t}{\epsilon^2})^{n-k}} - e^{-\frac{t_n-t_k}{\epsilon^2}} \right) f(Q_k^{\epsilon,\Delta t} - P_k^{\epsilon,\Delta t})
\end{aligned}$$

and finally

$$\begin{aligned}
e_{n,P,2,1}^{\epsilon,\Delta t} &= \sum_{k=0}^{n-1} \int_{t_k}^{t_{k+1}} e^{-\frac{t_n-t}{\epsilon^2}} (\sigma(Q^\epsilon(t_k) - P^\epsilon(t_k)) - \sigma(Q^\epsilon(t) - P^\epsilon(t))) d\beta(t) \\
e_{n,P,2,2}^{\epsilon,\Delta t} &= \sum_{k=0}^{n-1} \int_{t_k}^{t_{k+1}} e^{-\frac{t_n-t}{\epsilon^2}} (\sigma(Q_k^{\epsilon,\Delta t} - P_k^{\epsilon,\Delta t}) - \sigma(Q^\epsilon(t_k) - P^\epsilon(t_k))) d\beta(t) \\
e_{n,P,2,3}^{\epsilon,\Delta t} &= \sum_{k=0}^{n-1} \int_{t_k}^{t_{k+1}} (e^{-\frac{t_n-t_k}{\epsilon^2}} - e^{-\frac{t_n-t}{\epsilon^2}}) \sigma(Q_k^{\epsilon,\Delta t} - P_k^{\epsilon,\Delta t}) d\beta(t) \\
e_{n,P,2,4}^{\epsilon,\Delta t} &= \sum_{k=0}^{n-1} \left( \frac{1}{(1 + \frac{\Delta t}{\epsilon^2})^{n-k}} - e^{-\frac{t_n-t_k}{\epsilon^2}} \right) \sigma(Q_k^{\epsilon,\Delta t} - P_k^{\epsilon,\Delta t}) \Delta\beta_k.
\end{aligned}$$

Let us prove error estimates for each of the terms defined above. Let us start with the terms appearing in the right-hand side of the expression of the error term  $Q_n^{\epsilon,\Delta t} - Q^\epsilon(t_n)$  above.

Using the Cauchy–Schwarz inequality, the Lipschitz continuity of  $f$  (Assumption 1) and the inequality (43) from Lemma 5.1, one obtains

$$\begin{aligned}
\mathbb{E}[|e_{n,Q,1,1}^{\epsilon,\Delta t}|^2] &\leq CT \sum_{k=0}^{n-1} \int_{t_k}^{t_{k+1}} (\mathbb{E}[|Q^\epsilon(t_k) - Q^\epsilon(t)|^2] + \mathbb{E}[|P^\epsilon(t_k) - P^\epsilon(t)|^2]) dt \\
&\leq C(T)(1 + |q_0^\epsilon|^2 + |p_0^\epsilon|^2) \Delta t.
\end{aligned}$$

Using the Cauchy–Schwarz inequality and the Lipschitz continuity of  $f$ , one obtains

$$\mathbb{E}[|e_{n,Q,1,2}^{\epsilon,\Delta t}|^2] \leq CT \Delta t \sum_{k=0}^{n-1} E_k^{\epsilon,\Delta t}.$$

Using Itô's isometry formula, the Lipschitz continuity of  $\sigma$  and Lemma (43), one obtains

$$\begin{aligned}
\mathbb{E}[|e_{n,Q,2,1}^{\epsilon,\Delta t}|^2] &\leq C \sum_{k=0}^{n-1} \int_{t_k}^{t_{k+1}} (\mathbb{E}[|Q^\epsilon(t_k) - Q^\epsilon(t)|^2] + \mathbb{E}[|P^\epsilon(t_k) - P^\epsilon(t)|^2]) dt \\
&\leq C(T)(1 + |q_0^\epsilon|^2 + |p_0^\epsilon|^2) \Delta t.
\end{aligned}$$

Using Itô's isometry formula and the Lipschitz continuity of  $\sigma$ , one obtains

$$\mathbb{E}[|e_{n,Q,2,2}^{\epsilon,\Delta t}|^2] \leq CT \Delta t \sum_{k=0}^{n-1} E_k^{\epsilon,\Delta t}.$$

Gathering the estimates, for all  $n \in \{1, \dots, N\}$  one obtains the upper bound

$$\begin{aligned}
(44) \quad \mathbb{E}[|Q_n^{\epsilon,\Delta t} - Q^\epsilon(t_n)|^2] &\leq 4 \left( \mathbb{E}[|e_{n,Q,1,1}^{\epsilon,\Delta t}|^2] + \mathbb{E}[|e_{n,Q,1,2}^{\epsilon,\Delta t}|^2] + \mathbb{E}[|e_{n,Q,2,1}^{\epsilon,\Delta t}|^2] + \mathbb{E}[|e_{n,Q,2,2}^{\epsilon,\Delta t}|^2] \right) \\
&\leq CT \Delta t \sum_{k=0}^{n-1} E_k^{\epsilon,\Delta t} + C(T)(1 + |q_0^\epsilon|^2 + |p_0^\epsilon|^2) \Delta t.
\end{aligned}$$



Let us now treat the terms appearing in the right-hand side of the expression of the error term  $P_n^{\epsilon, \Delta t} - P^\epsilon(t_n)$  above.

Using the inequality (42), one obtains

$$\mathbb{E}[|e_{n,P,0}^{\epsilon, \Delta t}|^2] \leq \frac{C\epsilon^2}{(n+1)^2} |p_0^\epsilon|^2.$$

Using the Cauchy–Schwarz inequality, the Lipschitz continuity of  $f$  and Lemma (43), one obtains

$$\begin{aligned} \mathbb{E}[|e_{n,P,1,1}^{\epsilon, \Delta t}|^2] &\leq CT \sum_{k=0}^{n-1} \int_{t_k}^{t_{k+1}} (\mathbb{E}[|Q^\epsilon(t_k) - Q^\epsilon(t)|^2] + \mathbb{E}[|P^\epsilon(t_k) - P^\epsilon(t)|^2]) dt \\ &\leq C(T)(1 + |q_0^\epsilon|^2 + |p_0^\epsilon|^2) \Delta t. \end{aligned}$$

Using the Cauchy–Schwarz inequality and the Lipschitz continuity of  $f$ , one obtains

$$\mathbb{E}[|e_{n,P,1,2}^{\epsilon, \Delta t}|^2] \leq CT \Delta t \sum_{k=0}^{n-1} E_k^{\epsilon, \Delta t}.$$

Using the Cauchy–Schwarz inequality, the Lipschitz continuity of  $f$  and the inequality (41), one obtains

$$\begin{aligned} \mathbb{E}[|e_{n,P,1,3}^{\epsilon, \Delta t}|^2] &\leq C(T)(1 + |q_0^\epsilon|^2 + |p_0^\epsilon|^2) \sum_{k=0}^{n-1} \int_{t_k}^{t_{k+1}} (e^{-\frac{t-t_k}{\epsilon^2}} - 1)^2 e^{-2\frac{t_n-t}{\epsilon^2}} dt \\ &\leq C(T)(1 + |q_0^\epsilon|^2 + |p_0^\epsilon|^2) \frac{\Delta t}{\epsilon^2} \int_0^{t_n} e^{-2\frac{t_n-t}{\epsilon^2}} dt \\ &\leq C(T)(1 + |q_0^\epsilon|^2 + |p_0^\epsilon|^2) \Delta t. \end{aligned}$$

Using the Cauchy–Schwarz inequality, the Lipschitz continuity of  $f$  and the inequality (42), one obtains

$$\begin{aligned} \mathbb{E}[|e_{n,P,1,4}^{\epsilon, \Delta t}|^2] &\leq C(T)(1 + |q_0^\epsilon|^2 + |p_0^\epsilon|^2) \sum_{k=0}^{n-1} \frac{\Delta t}{(n-k)^2} \\ &\leq C(T)(1 + |q_0^\epsilon|^2 + |p_0^\epsilon|^2) \Delta t. \end{aligned}$$

Using Itô's isometry formula, the Lipschitz continuity of  $\sigma$  and Lemma (43), one obtains

$$\begin{aligned} \mathbb{E}[|e_{n,P,2,1}^{\epsilon, \Delta t}|^2] &\leq CT \sum_{k=0}^{n-1} \int_{t_k}^{t_{k+1}} (\mathbb{E}[|Q^\epsilon(t_k) - Q^\epsilon(t)|^2] + \mathbb{E}[|P^\epsilon(t_k) - P^\epsilon(t)|^2]) dt \\ &\leq C(T)(1 + |q_0^\epsilon|^2 + |p_0^\epsilon|^2) \Delta t. \end{aligned}$$

Using Itô's isometry formula and the Lipschitz continuity of  $\sigma$ , one obtains

$$\mathbb{E}[|e_{n,P,2,2}^{\epsilon, \Delta t}|^2] \leq CT \Delta t \sum_{k=0}^{n-1} E_k^{\epsilon, \Delta t}.$$

Using Itô's isometry formula, the Lipschitz continuity of  $\sigma$  and the inequality (41), one obtains

$$\begin{aligned} \mathbb{E}[|e_{n,P,2,3}^{\epsilon, \Delta t}|^2] &\leq C(T)(1 + |q_0^\epsilon|^2 + |p_0^\epsilon|^2) \sum_{k=0}^{n-1} \int_{t_k}^{t_{k+1}} (e^{-\frac{t-t_k}{\epsilon^2}} - 1)^2 e^{-2\frac{t_n-t}{\epsilon^2}} dt \\ &\leq C(T)(1 + |q_0^\epsilon|^2 + |p_0^\epsilon|^2) \frac{\Delta t}{\epsilon^2} \int_0^{t_n} e^{-2\frac{t_n-t}{\epsilon^2}} dt \\ &\leq C(T)(1 + |q_0^\epsilon|^2 + |p_0^\epsilon|^2) \Delta t. \end{aligned}$$

Using Itô's isometry formula, the Lipschitz continuity of  $\sigma$  and the inequality (42), one obtains

$$\begin{aligned} \mathbb{E}[|e_{n,P,2,4}^{\epsilon, \Delta t}|^2] &\leq C(T)(1 + |q_0^\epsilon|^2 + |p_0^\epsilon|^2) \sum_{k=0}^{n-1} \frac{\Delta t}{(n-k)^2} \\ &\leq C(T)(1 + |q_0^\epsilon|^2 + |p_0^\epsilon|^2) \Delta t. \end{aligned}$$

Gathering the estimates, for all  $n \in \{1, \dots, N\}$  one obtains the upper bound

$$\begin{aligned}
(45) \quad \mathbb{E}[|P_n^{\epsilon, \Delta t} - P^\epsilon(t_n)|^2] &\leq 9\mathbb{E}[|e_{n,P,0}^{\epsilon, \Delta t}|^2] \\
&+ 9\left(\mathbb{E}[|e_{n,P,1,1}^{\epsilon, \Delta t}|^2] + \mathbb{E}[|e_{n,P,1,2}^{\epsilon, \Delta t}|^2] + \mathbb{E}[|e_{n,P,1,3}^{\epsilon, \Delta t}|^2] + \mathbb{E}[|e_{n,P,1,4}^{\epsilon, \Delta t}|^2]\right) \\
&+ 9\left(\mathbb{E}[|e_{n,P,2,1}^{\epsilon, \Delta t}|^2] + \mathbb{E}[|e_{n,P,2,2}^{\epsilon, \Delta t}|^2] + \mathbb{E}[|e_{n,P,2,3}^{\epsilon, \Delta t}|^2] + \mathbb{E}[|e_{n,P,2,4}^{\epsilon, \Delta t}|^2]\right) \\
&\leq C\Delta t \sum_{k=1}^{n-1} E_k^{\epsilon, \Delta t} + \frac{C\epsilon^2}{(n+1)^2} |p_0^\epsilon|^2 + C(T)(1 + |q_0^\epsilon|^2 + |p_0^\epsilon|^2)\Delta t.
\end{aligned}$$

Let us now conclude the proof of the strong error estimates. Owing to the upper bounds (44) and (45), for all  $n \in \{0, \dots, N\}$ , one obtains the upper bound

$$\begin{aligned}
E_n^{\epsilon, \Delta t} &= \mathbb{E}[|Q_n^{\epsilon, \Delta t} - Q^\epsilon(t_n)|^2] + \mathbb{E}[|P_n^{\epsilon, \Delta t} - P^\epsilon(t_n)|^2] \\
&\leq C\Delta t \sum_{k=1}^{n-1} E_k^{\epsilon, \Delta t} + \frac{C\epsilon^2}{(n+1)^2} |p_0^\epsilon|^2 + C(T)(1 + |q_0^\epsilon|^2 + |p_0^\epsilon|^2)\Delta t,
\end{aligned}$$

with  $E_0^{\epsilon, \Delta t} = 0$ .

For all  $n \in \{1, \dots, N\}$ , set

$$\hat{E}_n^{\epsilon, \Delta t} = E_n^{\epsilon, \Delta t} - \frac{C\epsilon^2}{(n+1)^2} |p_0^\epsilon|^2,$$

then one obtains the inequality

$$\hat{E}_n^{\epsilon, \Delta t} \leq C(T)\Delta t \sum_{k=1}^{n-1} \hat{E}_k^{\epsilon, \Delta t} + C(T)(1 + |q_0^\epsilon|^2 + |p_0^\epsilon|^2)\Delta t,$$

using the fact that  $\sum_{k=1}^{\infty} 1/k^2 < \infty$ . Applying discrete Gronwall's lemma yields the inequality

$$\sup_{n=1, \dots, N} \hat{E}_n^{\epsilon, \Delta t} \leq C(T)(1 + |q_0^\epsilon|^2 + |p_0^\epsilon|^2)\Delta t,$$

therefore one obtains for all  $n \in \{1, \dots, N\}$  the upper bound

$$E_n^{\epsilon, \Delta t} \leq \frac{C\epsilon^2}{(n+1)^2} |p_0^\epsilon|^2 + C(T)(1 + |q_0^\epsilon|^2 + |p_0^\epsilon|^2)\Delta t.$$

Using Assumption 2 then concludes the proof of Theorem 3.1.  $\square$

## 6. Proof of the weak error estimates

The objective of this section is to provide the proof of Theorem 3.3. The most important

**6.1. Regularity estimates for solutions of Kolmogorov equations.** Let  $\varphi : \mathbb{R}^d \rightarrow \mathbb{R}$  be a mapping of class  $\mathcal{C}^3$ , with bounded derivatives of order 1, 2, 3. For all  $(q, p) \in \mathbb{R}^{2d}$  and  $t \geq 0$ , let

$$(46) \quad u^\epsilon(t, q, p) = \mathbb{E}_{q,p}[\varphi(q^\epsilon(t))]$$

where the subscript means that the process  $(q^\epsilon(t), p^\epsilon(t))_{t \geq 0}$  is solution of the SDE (8) with initial values  $q^\epsilon(0) = q$  and  $p^\epsilon(0) = p$ .

As a consequence of the definition (46) of the function  $u$ , the weak error may be written as

$$(47) \quad \mathbb{E}[\varphi(q_n^{\epsilon, \Delta t})] - \mathbb{E}[\varphi(q^\epsilon(t_n))] = \mathbb{E}[u^\epsilon(0, q_n^{\epsilon, \Delta t}, p_n^{\epsilon, \Delta t})] - \mathbb{E}[u^\epsilon(t_n, q_0^{\epsilon, \Delta t}, p_0^{\epsilon, \Delta t})],$$

for all  $n \in \{0, \dots, N\}$ , where  $q_n^{\epsilon, \Delta t}$  and  $p_n^{\epsilon, \Delta t}$  are given by (15), see for instance [21, 22].

Since the mappings  $\varphi$ ,  $f$  and  $\sigma$  are of class  $\mathcal{C}^3$  with bounded derivatives, the mapping

$$(t, q, p) \in [0, \infty) \times \mathbb{R}^{2d} \mapsto u^\epsilon(t, q, p)$$

is of class  $C_t^1 C_{q,p}^3$ , moreover it is the solution of the Kolmogorov equation (see [5] for instance)

$$(48) \quad \begin{cases} \partial_t u^\epsilon(t, q, p) = \frac{1}{\epsilon} \left( \nabla_q u^\epsilon(t, q, p) \cdot p + \nabla_p u^\epsilon(t, q, p) \cdot f(q) \right) \\ \quad + \frac{1}{\epsilon^2} \left( -\nabla_p u^\epsilon(t, q, p) \cdot p + \nabla_p^2 u^\epsilon(t, q, p) : a(q) \right), \\ u^\epsilon(0, q, p) = \varphi(q), \end{cases}$$

where the mapping  $a$  is given by (9) and the following notation is used:

$$\nabla_p^2 u^\epsilon(t, q, p) : a(q) = \sum_{i,j=1}^d \partial_i \partial_j u^\epsilon(t, q, p) a_{ij}(q).$$

Let us state upper bounds on the first, second and third order spatial derivatives of  $u^\epsilon$ , with a careful analysis of the dependence with respect to the parameter  $\epsilon$ .

**PROPOSITION 6.1.** *For all  $T \in (0, \infty)$ , there exists  $C(T) \in (0, \infty)$  such that for any function  $\varphi : \mathbb{R}^d \rightarrow \mathbb{R}$  of class  $C^3$  with bounded derivatives of order 1, 2, 3 and for all  $h^1, h^2, h^3 \in \mathbb{R}^d$ , one has*

$$(49) \quad \sup_{\epsilon \in (0, \epsilon_0)} \sup_{(t, q, p) \in [0, T] \times \mathbb{R}^{2d}} \frac{|\nabla_p u^\epsilon(t, q, p) \cdot h^1|}{\epsilon} \leq C(T) \|\varphi\|_1 |h^1|.$$

$$(50) \quad \sup_{\epsilon \in (0, \epsilon_0)} \sup_{(t, q, p) \in [0, T] \times \mathbb{R}^{2d}} \left( \frac{|\nabla_q \nabla_p u^\epsilon(t, q, p) \cdot (h^1, h^2)|}{\epsilon} + \frac{|\nabla_p^2 u^\epsilon(t, q, p) \cdot (h^1, h^2)|}{\epsilon^2} \right) \leq C(T) \|\varphi\|_2 |h^1| |h^2|.$$

$$(51) \quad \sup_{\epsilon \in (0, \epsilon_0)} \sup_{(t, q, p) \in [0, T] \times \mathbb{R}^{2d}} \left( \frac{|\nabla_q \nabla_p^2 u^\epsilon(t, q, p) \cdot (h^1, h^2, h^3)|}{\epsilon^2} + \frac{|\nabla_p^3 u^\epsilon(t, q, p) \cdot (h^1, h^2, h^3)|}{\epsilon^3} \right) \leq C(T) \|\varphi\|_3 |h^1| |h^2| |h^3|.$$

The proof of Proposition 6.1 is postponed to Section 6.3.

**6.2. Proof of Theorem 3.3.** This section is devoted to the proof of Theorem 3.3. Many of the arguments are standard, the novelty is to deal carefully with the dependence with respect to  $\epsilon$  using the regularity estimates from Proposition 6.1 above.

Before proceeding with the proof of Theorem 3.3, let us state and prove an auxiliary result, which is a variant of Lemma 5.1 for the auxiliary processes  $\tilde{q}^{\epsilon, \Delta t}$  and  $\tilde{p}^{\epsilon, \Delta t}$ .

**LEMMA 6.2.** *Let Assumption 1 be satisfied. For all  $T \in (0, \infty)$ , there exists  $C(T) \in (0, \infty)$  such that for all  $\epsilon \in (0, \epsilon_0)$ , all  $\Delta t \in (0, \Delta t_0)$ , all  $n \in \{0, \dots, N-1\}$  and all  $t \in [t_n, t_{n+1}]$ , one has*

$$(52) \quad \mathbb{E}[|\tilde{q}^{\epsilon, \Delta t}(t) - \tilde{q}^{\epsilon, \Delta t}(t_n)|^2] + \epsilon^2 \mathbb{E}[|\tilde{p}^{\epsilon, \Delta t}(t) - \tilde{p}^{\epsilon, \Delta t}(t_n)|^2] \leq C(T)(1 + |q_0^\epsilon|^2 + |p_0^\epsilon|^2) \Delta t.$$

**PROOF OF LEMMA 6.2.** Introduce the auxiliary variables  $\tilde{Q}^{\epsilon, \Delta t}(t) = \tilde{q}^{\epsilon, \Delta t}(t) + \epsilon \tilde{p}^{\epsilon, \Delta t}(t)$  and  $\tilde{P}^{\epsilon, \Delta t}(t) = \epsilon \tilde{p}^{\epsilon, \Delta t}(t)$ . The inequality (52) is a straightforward consequence of the following claim: for all  $T \in (0, \infty)$ , there exists  $C(T) \in (0, \infty)$  such that for all  $\epsilon \in (0, \epsilon_0)$ , all  $\Delta t \in (0, \Delta t_0)$  and all  $n \in \{0, \dots, N-1\}$  and all  $t \in [t_n, t_{n+1}]$ , one has

$$(53) \quad \mathbb{E}[|\tilde{Q}^{\epsilon, \Delta t}(t) - \tilde{Q}^{\epsilon, \Delta t}(t_n)|^2] + \mathbb{E}[|\tilde{P}^{\epsilon, \Delta t}(t) - \tilde{P}^{\epsilon, \Delta t}(t_n)|^2] \leq C(T)(1 + |q_0^\epsilon|^2 + |p_0^\epsilon|^2) \Delta t.$$

Let us establish the claim (53). Recall that  $\tilde{q}^{\epsilon, \Delta t}(t_n) = q_n^{\epsilon, \Delta t}$  and  $\tilde{p}^{\epsilon, \Delta t}(t_n) = p_n^{\epsilon, \Delta t}$  for all  $n \in \{0, \dots, N-1\}$ , owing to Proposition 2.1.

On the one hand, for all  $t \in [t_n, t_{n+1}]$ , one has

$$\tilde{Q}^{\epsilon, \Delta t}(t) - \tilde{Q}^{\epsilon, \Delta t}(t_n) = \int_{s_1}^{s_2} f(q_n^{\epsilon, \Delta t}) ds + \int_{s_1}^{s_2} \sigma(q_n^{\epsilon, \Delta t}) d\beta(s),$$

therefore using Itô's isometry formula, the Lipschitz continuity property of the mappings  $f$  and  $\sigma$  (Assumption 1) and the moment bounds (40) from Proposition 4.3, one obtains

$$\mathbb{E}[|\tilde{Q}^{\epsilon, \Delta t}(t) - \tilde{Q}^{\epsilon, \Delta t}(t_n)|^2] \leq C(T) \int_{t_n}^t (1 + \mathbb{E}[|q_n^{\epsilon, \Delta t}|^2]) ds \leq C(T)(1 + |q_0^\epsilon|^2 + |p_0^\epsilon|^2) \Delta t.$$

On the other hand, for all  $t \in [t_n, t_{n+1}]$ , one has

$$\tilde{P}^{\epsilon, \Delta t}(t) - \tilde{P}^{\epsilon, \Delta t}(t_n) = \left(e^{-\frac{t-t_n}{\epsilon^2}} - 1\right) \tilde{P}^{\epsilon, \Delta t}(t_n) + \int_{t_n}^t e^{-\frac{t-s}{\epsilon^2}} f(q_n^{\epsilon, \Delta t}) ds + \int_{t_n}^t e^{-\frac{t-s}{\epsilon^2}} \sigma(q_n^{\epsilon, \Delta t}) d\beta(s).$$

Using the inequality (41), the identity  $\tilde{P}^{\epsilon, \Delta t}(t_n) = \epsilon p_n^{\epsilon, \Delta t}$ , Itô's isometry formula, the Lipschitz continuity property of the mappings  $f$  and  $\sigma$  (Assumption 1) and the moment bounds (40) from Proposition 4.3, one obtains

$$\begin{aligned} \mathbb{E}[|\tilde{P}^{\epsilon, \Delta t}(t) - \tilde{P}^{\epsilon, \Delta t}(t_n)|^2] &\leq C \Delta t \mathbb{E}[|p_n^{\epsilon, \Delta t}|^2] + C(T) \int_{t_n}^t (1 + \mathbb{E}[|q_n^{\epsilon, \Delta t}|^2]) ds \\ &\leq C(T)(1 + |q_0^\epsilon|^2 + |p_0^\epsilon|^2) \Delta t. \end{aligned}$$

Gathering the estimates yields the claim (53) and then concludes the proof of Lemma 5.1.  $\square$

PROOF OF THEOREM 3.3. Owing to the expression (47) of the weak error using the function  $u^\epsilon$  defined by (46), the auxiliary process defined by (16) and a standard telescoping sum argument, one has the following decomposition of the weak error: for all  $n \in \{0, \dots, N\}$ ,

$$\begin{aligned} (54) \quad \mathbb{E}[\varphi(q_n^{\epsilon, \Delta t})] - \mathbb{E}[\varphi(q^\epsilon(t_n))] &= \mathbb{E}[u^\epsilon(0, q_n^{\epsilon, \Delta t}, p_n^{\epsilon, \Delta t})] - \mathbb{E}[u^\epsilon(t_n, q_0^{\epsilon, \Delta t}, p_0^{\epsilon, \Delta t})] \\ &= \sum_{m=0}^{n-1} (\mathbb{E}[u^\epsilon(t_n - t_{m+1}, q_{m+1}^{\epsilon, \Delta t}, p_{m+1}^{\epsilon, \Delta t})] - \mathbb{E}[u^\epsilon(t_n - t_m, q_m^{\epsilon, \Delta t}, p_m^{\epsilon, \Delta t})]) \\ &= \sum_{m=0}^{n-1} (\mathbb{E}[u^\epsilon(t_n - t_{m+1}, \tilde{q}^{\epsilon, \Delta t}(t_{m+1}), \tilde{p}^{\epsilon, \Delta t}(t_{m+1}))] - \mathbb{E}[u^\epsilon(t_n - t_m, \tilde{q}^{\epsilon, \Delta t}(t_m), \tilde{p}^{\epsilon, \Delta t}(t_m))]). \end{aligned}$$

Using Itô's formula, the fact that  $u^\epsilon$  is solution of the Kolmogorov equation (48) and the fact that the process  $(\tilde{q}^{\epsilon, \Delta t}(t), \tilde{p}^{\epsilon, \Delta t}(t))_{t \geq 0}$  is solution of the SDE (16), for all  $m \in \{0, \dots, n-1\}$ , one obtains

$$\mathbb{E}[u^\epsilon(t_n - t_{m+1}, \tilde{q}^{\epsilon, \Delta t}(t_{m+1}), \tilde{p}^{\epsilon, \Delta t}(t_{m+1}))] - \mathbb{E}[u^\epsilon(t_n - t_m, \tilde{q}^{\epsilon, \Delta t}(t_m), \tilde{p}^{\epsilon, \Delta t}(t_m))] = \mathcal{E}_{m,n,1}^{\epsilon, \Delta t} + \mathcal{E}_{m,n,2}^{\epsilon, \Delta t}$$

where the error terms in the right-hand side above are defined for all  $0 \leq m \leq n-1$  by

$$\begin{aligned} \mathcal{E}_{m,n,1}^{\epsilon, \Delta t} &= \frac{1}{\epsilon} \int_{t_m}^{t_{m+1}} \mathbb{E}[\nabla_p u^\epsilon(t_n - t, \tilde{q}^{\epsilon, \Delta t}(t), \tilde{p}^{\epsilon, \Delta t}(t)) \cdot (f(\tilde{q}^{\epsilon, \Delta t}(t_m)) - f(\tilde{q}^{\epsilon, \Delta t}(t)))] dt \\ \mathcal{E}_{m,n,2}^{\epsilon, \Delta t} &= \frac{1}{\epsilon^2} \int_{t_m}^{t_{m+1}} \mathbb{E}[\nabla_p^2 u^\epsilon(t_n - t, \tilde{q}^{\epsilon, \Delta t}(t), \tilde{p}^{\epsilon, \Delta t}(t)) : (a(\tilde{q}^{\epsilon, \Delta t}(t_m)) - a(\tilde{q}^{\epsilon, \Delta t}(t)))] dt. \end{aligned}$$

Observe that the error term  $\mathcal{E}_{m,n,1}^{\epsilon, \Delta t}$ , resp.  $\mathcal{E}_{m,n,2}^{\epsilon, \Delta t}$ , vanishes if the mapping  $f$  is constant, resp. if the mapping  $\sigma$  is constant. This is due to the construction of the numerical scheme (15).

The error terms  $\mathcal{E}_{m,n,1}^{\epsilon, \Delta t}$  and  $\mathcal{E}_{m,n,2}^{\epsilon, \Delta t}$  are then decomposed as follows: set

$$\begin{aligned} \mathcal{E}_{m,n,1,1}^{\epsilon, \Delta t} &= \frac{1}{\epsilon} \int_{t_m}^{t_{m+1}} \mathbb{E}[\nabla_p u^\epsilon(t_n - t, \tilde{q}^{\epsilon, \Delta t}(t_m), \tilde{p}^{\epsilon, \Delta t}(t_m)) \cdot (f(\tilde{q}^{\epsilon, \Delta t}(t_m)) - f(\tilde{q}^{\epsilon, \Delta t}(t)))] dt \\ \mathcal{E}_{m,n,1,2}^{\epsilon, \Delta t} &= \mathcal{E}_{m,n,1}^{\epsilon, \Delta t} - \mathcal{E}_{m,n,1,1}^{\epsilon, \Delta t}, \\ \mathcal{E}_{m,n,2,1}^{\epsilon, \Delta t} &= \frac{1}{\epsilon^2} \int_{t_m}^{t_{m+1}} \mathbb{E}[\nabla_p^2 u^\epsilon(t_n - t, \tilde{q}^{\epsilon, \Delta t}(t_m), \tilde{p}^{\epsilon, \Delta t}(t_m)) : (a(\tilde{q}^{\epsilon, \Delta t}(t_m)) - a(\tilde{q}^{\epsilon, \Delta t}(t)))] dt \\ \mathcal{E}_{m,n,2,2}^{\epsilon, \Delta t} &= \mathcal{E}_{m,n,2}^{\epsilon, \Delta t} - \mathcal{E}_{m,n,2,1}^{\epsilon, \Delta t}. \end{aligned}$$

It remains to obtain upper bounds for the four error terms.

• Since the mapping  $f$  is of class  $\mathcal{C}^2$  with bounded second-order derivative, a Taylor expansion argument gives

$$\mathcal{E}_{m,n,1,1}^{\epsilon, \Delta t} = \mathcal{E}_{m,n,1,1,1}^{\epsilon, \Delta t} + \mathcal{E}_{m,n,1,1,2}^{\epsilon, \Delta t},$$

where  $\mathcal{E}_{m,n,1,1,1}^{\epsilon, \Delta t}$  is defined by

$$\mathcal{E}_{m,n,1,1,1}^{\epsilon, \Delta t} = \frac{1}{\epsilon} \int_{t_m}^{t_{m+1}} \mathbb{E}[\nabla_p u^\epsilon(t_n - t, \tilde{q}^{\epsilon, \Delta t}(t_m), \tilde{p}^{\epsilon, \Delta t}(t_m)) \cdot (Df(\tilde{q}^{\epsilon, \Delta t}(t_m)) \cdot (\tilde{q}^{\epsilon, \Delta t}(t_m) - \tilde{q}^{\epsilon, \Delta t}(t)))] dt$$

and where  $\mathcal{E}_{m,n,1,1,2}^{\epsilon,\Delta t}$  satisfies

$$\begin{aligned} |\mathcal{E}_{m,n,1,1,2}^{\epsilon,\Delta t}| &\leq C(T) \|\varphi\|_1 \int_{t_m}^{t_{m+1}} \mathbb{E}[|\tilde{q}^{\epsilon,\Delta t}(t_m) - \tilde{q}^{\epsilon,\Delta t}(t)|^2] dt \\ &\leq C(T) \|\varphi\|_1 (1 + |q_0^0|^2) \Delta t^2 \end{aligned}$$

using first the inequality (49) from Proposition 6.1 and second the inequality (52) from Lemma 6.2 and Assumption 2.

To deal with the error term  $\mathcal{E}_{m,n,1,1,1}^{\epsilon,\Delta t}$ , a conditional expectation argument is used. For all  $t \in [t_m, t_{m+1}]$ , solving the auxiliary stochastic differential equation (16), one has

$$(55) \quad \begin{cases} \tilde{q}^{\epsilon,\Delta t}(t) - \tilde{q}^{\epsilon,\Delta t}(t_m) = \frac{1}{\epsilon} \int_{t_m}^t \tilde{p}^{\epsilon,\Delta t}(s) ds = \epsilon(\tilde{p}^{\epsilon,\Delta t}(t_m) - \tilde{p}^{\epsilon,\Delta t}(t)) + \int_{t_m}^t f(q_m^{\epsilon,\Delta t}) ds + \int_{t_m}^t \sigma(q_m^{\epsilon,\Delta t}) d\beta(s) \\ \tilde{p}^{\epsilon,\Delta t}(t) - \tilde{p}^{\epsilon,\Delta t}(t_m) = (e^{-\frac{t-t_m}{\epsilon^2}} - 1)\tilde{p}^{\epsilon,\Delta t}(t_m) + \frac{1}{\epsilon} \int_{t_m}^t e^{-\frac{t-s}{\epsilon^2}} f(q_m^{\epsilon,\Delta t}) ds + \frac{1}{\epsilon} \int_{t_m}^t e^{-\frac{t-s}{\epsilon^2}} \sigma(q_m^{\epsilon,\Delta t}) d\beta(s). \end{cases}$$

Therefore, the error term  $\mathcal{E}_{m,n,1,1,1}^{\epsilon,\Delta t}$  can be rewritten as

$$\begin{aligned} \mathcal{E}_{m,n,1,1,1}^{\epsilon,\Delta t} &= \int_{t_m}^{t_{m+1}} \mathbb{E}[\nabla_p u^\epsilon(t_n - t, \tilde{q}^{\epsilon,\Delta t}(t_m), \tilde{p}^{\epsilon,\Delta t}(t_m)) \cdot (Df(\tilde{q}^{\epsilon,\Delta t}(t_m)) \cdot (\tilde{p}^{\epsilon,\Delta t}(t) - \tilde{p}^{\epsilon,\Delta t}(t_m)))] dt \\ &\quad - \frac{1}{\epsilon} \int_{t_m}^{t_{m+1}} \mathbb{E}[\nabla_p u^\epsilon(t_n - t, \tilde{q}^{\epsilon,\Delta t}(t_m), \tilde{p}^{\epsilon,\Delta t}(t_m)) \cdot (Df(\tilde{q}^{\epsilon,\Delta t}(t_m)) \cdot \int_{t_m}^t f(q_m^{\epsilon,\Delta t}) ds)] dt \\ &= \int_{t_m}^{t_{m+1}} \mathbb{E}[\nabla_p u^\epsilon(t_n - t, \tilde{q}^{\epsilon,\Delta t}(t_m), \tilde{p}^{\epsilon,\Delta t}(t_m)) \cdot (Df(\tilde{q}^{\epsilon,\Delta t}(t_m)) \cdot ((e^{-\frac{t-t_m}{\epsilon^2}} - 1)\tilde{p}^{\epsilon,\Delta t}(t_m)))] dt \\ &\quad + \frac{1}{\epsilon} \int_{t_m}^{t_{m+1}} \mathbb{E}[\nabla_p u^\epsilon(t_n - t, \tilde{q}^{\epsilon,\Delta t}(t_m), \tilde{p}^{\epsilon,\Delta t}(t_m)) \cdot (Df(\tilde{q}^{\epsilon,\Delta t}(t_m)) \cdot \int_{t_m}^t (e^{-\frac{t-s}{\epsilon^2}} - 1) f(q_m^{\epsilon,\Delta t}) ds)] dt, \end{aligned}$$

since the terms with  $\int_{t_m}^t \sigma(q_m^{\epsilon,\Delta t}) d\beta(s)$  and  $\int_{t_m}^t e^{-\frac{t-s}{\epsilon^2}} \sigma(q_m^{\epsilon,\Delta t}) d\beta(s)$  vanish in expectation.

Using the inequality (49) from Proposition 6.1, the boundedness of  $Df$ , the moment bounds (40) from Proposition 4.3 and Assumption 2, one obtains

$$\begin{aligned} |\mathcal{E}_{m,n,1,1,1}^{\epsilon,\Delta t}| &\leq C(T) \|\varphi\|_1 \int_{t_m}^{t_{m+1}} \epsilon(1 - e^{-\frac{t-t_m}{\epsilon^2}}) dt \mathbb{E}[|\tilde{p}^{\epsilon,\Delta t}(t_m)|] \\ &\quad + C(T) \|\varphi\|_1 \Delta t^2 (1 + \mathbb{E}[|\tilde{q}^{\epsilon,\Delta t}(t_m)|]) \\ &\leq C(T) \|\varphi\|_1 \left( \int_0^{\Delta t} \epsilon(1 - e^{-\frac{t}{\epsilon^2}}) dt + \Delta t^2 \right) (1 + |q_0^\epsilon|^2 + |p_0^\epsilon|^2) \\ &\leq C(T) \Delta t \|\varphi\|_1 (\mathcal{R}(\epsilon, \Delta t) + \Delta t) (1 + |q_0^0|^2), \end{aligned}$$

where we recall that  $\mathcal{R}(\epsilon, \Delta t)$  is defined by (31).

• Owing to the inequality (50) from Proposition 6.1, for all  $m \in \{0, \dots, n-1\}$  and using the Lipschitz continuity of  $f$ , one obtains

$$\begin{aligned} |\mathcal{E}_{m,n,1,2}^{\epsilon,\Delta t}| &\leq C(T) \|\varphi\|_2 \int_{t_m}^{t_{m+1}} \mathbb{E}[|\tilde{q}^{\epsilon,\Delta t}(t) - \tilde{q}^{\epsilon,\Delta t}(t_m)|^2] dt \\ &\quad + C(T) \epsilon \|\varphi\|_2 \int_{t_m}^{t_{m+1}} \mathbb{E}[|\tilde{q}^{\epsilon,\Delta t}(t) - \tilde{q}^{\epsilon,\Delta t}(t_m)| |\tilde{p}^{\epsilon,\Delta t}(t) - \tilde{p}^{\epsilon,\Delta t}(t_m)|] dt. \end{aligned}$$

Using the inequality (52) from Lemma 6.2 and Assumption 2, one then obtains the upper bound

$$|\mathcal{E}_{m,n,1,2}^{\epsilon,\Delta t}| \leq C(T) \|\varphi\|_2 (1 + |q_0^0|^2) \Delta t^2.$$

• Since the mapping  $a$  is of class  $\mathcal{C}^2$  with bounded second-order derivative, a Taylor expansion argument gives

$$\mathcal{E}_{m,n,2,1}^{\epsilon,\Delta t} = \mathcal{E}_{m,n,2,1,1}^{\epsilon,\Delta t} + \mathcal{E}_{m,n,2,1,2}^{\epsilon,\Delta t},$$

where  $\mathcal{E}_{m,n,2,1,1}^{\epsilon,\Delta t}$  is defined by

$$\mathcal{E}_{m,n,2,1,1}^{\epsilon,\Delta t} = \frac{1}{\epsilon^2} \int_{t_m}^{t_{m+1}} \mathbb{E}[\nabla_p^2 u^\epsilon(t_n - t, \tilde{q}^{\epsilon,\Delta t}(t_m), \tilde{p}^{\epsilon,\Delta t}(t_m)) : (Da(\tilde{q}^{\epsilon,\Delta t}(t_m)) \cdot (\tilde{q}^{\epsilon,\Delta t}(t_m) - \tilde{q}^{\epsilon,\Delta t}(t))) )] dt$$

and where  $\mathcal{E}_{m,n,2,1,2}^{\epsilon,\Delta t}$  satisfies

$$\begin{aligned} |\mathcal{E}_{m,n,2,1,2}^{\epsilon,\Delta t}| &\leq C(T) \|\varphi\|_2 \int_{t_m}^{t_{m+1}} \mathbb{E}[|\tilde{q}^{\epsilon,\Delta t}(t_m) - \tilde{q}^{\epsilon,\Delta t}(t)|^2] dt \\ &\leq C(T) \|\varphi\|_2 (1 + |q_0^0|^2) \Delta t^2 \end{aligned}$$

using first the inequality (50) from Proposition 6.1 and second the inequality (52) from Lemma 6.2 and Assumption 2.

Like in the treatment of the error term  $\mathcal{E}_{m,n,1,1,1}^{\epsilon,\Delta t}$  above, a conditional expectation argument is used to deal with the error term  $\mathcal{E}_{m,n,2,1,1}^{\epsilon,\Delta t}$ . The expressions (55) of  $\tilde{q}^{\epsilon,\Delta t}(t) - \tilde{q}^{\epsilon,\Delta t}(t_m)$  and  $\tilde{p}^{\epsilon,\Delta t}(t) - \tilde{p}^{\epsilon,\Delta t}(t_m)$  give

$$\begin{aligned} \mathcal{E}_{m,n,2,1,1}^{\epsilon,\Delta t} &= \frac{1}{\epsilon} \int_{t_m}^{t_{m+1}} \mathbb{E}[\nabla_p^2 u^\epsilon(t_n - t, \tilde{q}^{\epsilon,\Delta t}(t_m), \tilde{p}^{\epsilon,\Delta t}(t_m)) : (Da(\tilde{q}^{\epsilon,\Delta t}(t_m)) \cdot (\tilde{p}^{\epsilon,\Delta t}(t) - \tilde{p}^{\epsilon,\Delta t}(t_m))) )] dt \\ &\quad - \frac{1}{\epsilon^2} \int_{t_m}^{t_{m+1}} \mathbb{E}[\nabla_p^2 u^\epsilon(t_n - t, \tilde{q}^{\epsilon,\Delta t}(t_m), \tilde{p}^{\epsilon,\Delta t}(t_m)) : (Da(\tilde{q}^{\epsilon,\Delta t}(t_m)) \cdot \int_{t_m}^t f(q_m^{\epsilon,\Delta t}) ds)] dt \\ &= \frac{1}{\epsilon} \int_{t_m}^{t_{m+1}} \mathbb{E}[\nabla_p^2 u^\epsilon(t_n - t, \tilde{q}^{\epsilon,\Delta t}(t_m), \tilde{p}^{\epsilon,\Delta t}(t_m)) : (Da(\tilde{q}^{\epsilon,\Delta t}(t_m)) \cdot ((e^{-\frac{t-t_m}{\epsilon^2}} - 1) \tilde{p}^{\epsilon,\Delta t}(t_m))) )] dt \\ &\quad + \frac{1}{\epsilon^2} \int_{t_m}^{t_{m+1}} \mathbb{E}[\nabla_p^2 u^\epsilon(t_n - t, \tilde{q}^{\epsilon,\Delta t}(t_m), \tilde{p}^{\epsilon,\Delta t}(t_m)) : (Da(\tilde{q}^{\epsilon,\Delta t}(t_m)) \cdot \int_{t_m}^t (e^{-\frac{t-s}{\epsilon^2}} - 1) f(q_m^{\epsilon,\Delta t}) ds)] dt. \end{aligned}$$

Using the inequality (50) from Proposition 6.1, the boundedness of  $Da$ , and the moment bounds (40) from Proposition 4.3, one obtains

$$\begin{aligned} |\mathcal{E}_{m,n,2,1,1}^{\epsilon,\Delta t}| &\leq C(T) \|\varphi\|_2 \int_{t_m}^{t_{m+1}} \epsilon (1 - e^{-\frac{t-t_m}{\epsilon^2}}) dt \mathbb{E}[|\tilde{p}^{\epsilon,\Delta t}(t_m)|] \\ &\quad + C(T) \|\varphi\|_2 \Delta t^2 (1 + \mathbb{E}[|\tilde{q}^{\epsilon,\Delta t}(t_m)|]) \\ &\leq C(T) \|\varphi\|_2 \left( \int_0^{\Delta t} \epsilon (1 - e^{-\frac{t}{\epsilon^2}}) dt + \Delta t^2 \right) (1 + |q_0^0|^2 + |p_0^0|^2) \\ &\leq C(T) \Delta t \|\varphi\|_2 \left( \mathcal{R}(\epsilon, \Delta t) + \Delta t \right) (1 + |q_0^0|^2). \end{aligned}$$

• Owing to the inequality (51) from Proposition 6.1, for all  $m \in \{0, \dots, n-1\}$  and using the Lipschitz continuity property (10) of  $a$ , one obtains

$$\begin{aligned} |\mathcal{E}_{m,n,2,2}^{\epsilon,\Delta t}| &\leq C(T) \|\varphi\|_3 \int_{t_m}^{t_{m+1}} \mathbb{E}[|\tilde{q}^{\epsilon,\Delta t}(t) - \tilde{q}^{\epsilon,\Delta t}(t_m)|^2] dt \\ &\quad + C(T) \epsilon \|\varphi\|_3 \int_{t_m}^{t_{m+1}} \mathbb{E}[|\tilde{q}^{\epsilon,\Delta t}(t) - \tilde{q}^{\epsilon,\Delta t}(t_m)| |\tilde{p}^{\epsilon,\Delta t}(t) - \tilde{p}^{\epsilon,\Delta t}(t_m)|] dt. \end{aligned}$$

Using the inequality (52) from Lemma 6.2 and Assumption 2, one then obtains the upper bound

$$|\mathcal{E}_{m,n,2,2}^{\epsilon,\Delta t}| \leq C(T) \|\varphi\|_3 (1 + |q_0^0|^2) \Delta t^2.$$

• Gathering the estimates, one finally obtains

$$\begin{aligned} |\mathbb{E}[\varphi(q_n^{\epsilon,\Delta t})] - \mathbb{E}[\varphi(q^\epsilon(t_n))]| &\leq \sum_{m=0}^{n-1} (|\mathcal{E}_{m,n,1}^{\epsilon,\Delta t}| + |\mathcal{E}_{m,n,2}^{\epsilon,\Delta t}|) \\ &\leq C(T) \|\varphi\|_3 \left( \mathcal{R}(\epsilon, \Delta t) + \Delta t \right) (1 + |q_0^0|^2) \end{aligned}$$

for all  $n \in \{0, \dots, N-1\}$ , where  $\mathcal{R}(\epsilon, \Delta t)$  is defined by (31). This concludes the proof of the weak error estimate (30) and the proof of Theorem 3.3.  $\square$

### 6.3. Proof of Proposition 6.1.

PROOF OF THE INEQUALITY (49). For all  $\mathbf{h} = (h_q, h_p) \in \mathbb{R}^{2d}$ , one has

$$\nabla_q u^\epsilon(t, q, p) \cdot h_q + \nabla_p u^\epsilon(t, q, p) \cdot h_p = \nabla_{(q,p)} u^\epsilon(t, q, p) \cdot \mathbf{h} = \mathbb{E}_{q,p}[\nabla_q \varphi(q^\epsilon(t)) \cdot \eta_q^{\epsilon, \mathbf{h}}(t)],$$

where  $t \mapsto \eta^{\epsilon, \mathbf{h}}(t) = (\eta_q^{\epsilon, \mathbf{h}}(t), \eta_p^{\epsilon, \mathbf{h}}(t)) \in \mathbb{R}^{2d}$  is solution of the stochastic differential equation

$$\begin{cases} d\eta_q^{\epsilon, \mathbf{h}}(t) = \frac{\eta_p^{\epsilon, \mathbf{h}}(t)}{\epsilon} dt \\ d\eta_p^{\epsilon, \mathbf{h}}(t) = -\frac{\eta_p^{\epsilon, \mathbf{h}}(t)}{\epsilon^2} dt + \frac{1}{\epsilon} Df(q^\epsilon(t)) \cdot \eta_q^{\epsilon, \mathbf{h}}(t) dt + \frac{1}{\epsilon} D\sigma(q^\epsilon(t)) \cdot \eta_q^{\epsilon, \mathbf{h}}(t) d\beta(t), \end{cases}$$

with initial value  $\eta^{\epsilon, \mathbf{h}}(0) = \mathbf{h}$ , equivalently  $\eta_q^{\epsilon, \mathbf{h}}(0) = h_q$  and  $\eta_p^{\epsilon, \mathbf{h}}(0) = h_p$ .

The inequality (49) is a straightforward consequence of the following claim: for all  $m \in \mathbb{N}$ , there exists  $C_m(T) \in (0, \infty)$  such that

$$(56) \quad \sup_{(t,q,p) \in [0,T] \times \mathbb{R}^{2d}} \mathbb{E}_{q,p}[|\eta_q^{\epsilon, \mathbf{h}}(t)|^{2m}] \leq C_m(T) (|h_q|^{2m} + \epsilon^{2m} |h_p|^{2m}).$$

Indeed, it suffices to choose  $m = 1$  and to set  $\mathbf{h} = (0, h)$  to obtain (49). It thus remains to prove the claim (56). First, observe that, for all  $t \geq 0$ , one has

$$\eta_p^{\epsilon, \mathbf{h}}(t) = e^{-\frac{t}{\epsilon^2}} h_p + \frac{1}{\epsilon} \int_0^t e^{-\frac{t-s}{\epsilon^2}} Df(q^\epsilon(s)) \cdot \eta_q^{\epsilon, \mathbf{h}}(s) ds + \frac{1}{\epsilon} \int_0^t e^{-\frac{t-s}{\epsilon^2}} D\sigma(q^\epsilon(s)) \cdot \eta_q^{\epsilon, \mathbf{h}}(s) d\beta(s).$$

In addition, for all  $t \geq 0$ , one has

$$\begin{aligned} \eta_q^{\epsilon, \mathbf{h}}(t) &= h_q + \frac{1}{\epsilon} \int_0^t \eta_p^{\epsilon, \mathbf{h}}(s) ds \\ &= h_q + \int_0^t Df(q^\epsilon(s)) \cdot \eta_q^{\epsilon, \mathbf{h}}(s) ds + \int_0^t D\sigma(q^\epsilon(s)) \cdot \eta_q^{\epsilon, \mathbf{h}}(s) d\beta(s) + \epsilon(\eta_p^{\epsilon, \mathbf{h}}(0) - \eta_p^{\epsilon, \mathbf{h}}(t)). \end{aligned}$$

On the one hand, using the Hölder and Burkholder–Davis–Gundy inequalities, one obtains the following upper bound for  $\eta_p^{\epsilon, \mathbf{h}}(t)$ : for all  $t \in [0, T]$  one has

$$\mathbb{E}[|\eta_p^{\epsilon, \mathbf{h}}(t)|^{2m}] \leq C_m(T) \left( |h_p|^{2m} + \frac{1}{\epsilon^{2m}} \int_0^t \mathbb{E}[|\eta_q^{\epsilon, \mathbf{h}}(s)|^{2m}] ds \right).$$

On the other hand, using the Hölder and Burkholder–Davis–Gundy inequalities and the upper bound above for  $\eta_p^{\epsilon, \mathbf{h}}(t)$ , one obtains the following upper bound for  $\eta_q^{\epsilon, \mathbf{h}}(t)$ : for all  $t \in [0, T]$  one has

$$\mathbb{E}[|\eta_q^{\epsilon, \mathbf{h}}(t)|^{2m}] \leq C_m(T) \left( |h_q|^{2m} + \int_0^t \mathbb{E}[|\eta_q^{\epsilon, \mathbf{h}}(s)|^{2m}] ds + \epsilon^{2m} |h_p|^{2m} + \epsilon^{2m} \mathbb{E}[|\eta_p^{\epsilon, \mathbf{h}}(t)|^{2m}] \right).$$

Combining the two upper bounds gives the inequality

$$\mathbb{E}[|\eta_q^{\epsilon, \mathbf{h}}(t)|^{2m}] \leq C_m(T) \left( |h_q|^{2m} + \epsilon^{2m} |h_p|^{2m} + \int_0^t \mathbb{E}[|\eta_q^{\epsilon, \mathbf{h}}(s)|^{2m}] ds \right)$$

for all  $t \in [0, T]$ . Applying Gronwall's inequality then yields the claim (56). This concludes the proof of the inequality (49).  $\square$

PROOF OF THE INEQUALITY (50). For all  $\mathbf{h}^1 = (h_q^1, h_p^1) \in \mathbb{R}^{2d}$  and  $\mathbf{h}^2 = (h_q^2, h_p^2) \in \mathbb{R}^{2d}$ , one has

$$\begin{aligned} \nabla_{q,p}^2 u^\epsilon(t, q, p) \cdot (\mathbf{h}^1, \mathbf{h}^2) &= \nabla_q \nabla_q u^\epsilon(t, q, p) \cdot (h_q^1, h_q^2) + \nabla_p \nabla_p u^\epsilon(t, q, p) \cdot (h_p^1, h_p^2) \\ &\quad + \nabla_q \nabla_p u^\epsilon(t, q, p) \cdot (h_q^1, h_p^2) + \nabla_p \nabla_q u^\epsilon(t, q, p) \cdot (h_p^1, h_q^2) \\ &= \mathbb{E}[\nabla_q \varphi(q^\epsilon(t)) \cdot \zeta_q^{\epsilon, \mathbf{h}^1, \mathbf{h}^2}(t)] + \mathbb{E}[\nabla_q^2 \varphi(q^\epsilon(t)) \cdot (\eta_q^{\epsilon, \mathbf{h}^1}(t), \eta_q^{\epsilon, \mathbf{h}^2}(t))], \end{aligned}$$

where  $t \mapsto \zeta^{\epsilon, \mathbf{h}^1, \mathbf{h}^2}(t) = (\zeta_q^{\epsilon, \mathbf{h}^1, \mathbf{h}^2}(t), \zeta_p^{\epsilon, \mathbf{h}^1, \mathbf{h}^2}(t)) \in \mathbb{R}^{2d}$  is solution of the stochastic differential equation

$$\begin{cases} d\zeta_q^{\epsilon, \mathbf{h}^1, \mathbf{h}^2}(t) = \frac{\zeta_p^{\epsilon, \mathbf{h}^1, \mathbf{h}^2}(t)}{\epsilon} dt \\ d\zeta_p^{\epsilon, \mathbf{h}^1, \mathbf{h}^2}(t) = -\frac{1}{\epsilon^2} \zeta_p^{\epsilon, \mathbf{h}^1, \mathbf{h}^2}(t) dt + \frac{1}{\epsilon} Df(q^\epsilon(t)) \cdot \zeta_q^{\epsilon, \mathbf{h}^1, \mathbf{h}^2}(t) dt + \frac{1}{\epsilon} D\sigma(q^\epsilon(t)) \cdot \zeta_q^{\epsilon, \mathbf{h}^1, \mathbf{h}^2}(t) d\beta(t) \\ \quad + \frac{1}{\epsilon} D^2 f(q^\epsilon(t)) \cdot (\eta_q^{\epsilon, \mathbf{h}^1}(t), \eta_q^{\epsilon, \mathbf{h}^2}(t)) dt + \frac{1}{\epsilon} D^2 \sigma(q^\epsilon(t)) \cdot (\eta_q^{\epsilon, \mathbf{h}^1}(t), \eta_q^{\epsilon, \mathbf{h}^2}(t)) d\beta(t), \end{cases}$$

with initial value  $\zeta^{\epsilon, \mathbf{h}^1, \mathbf{h}^2}(0) = 0$ , equivalently  $\zeta_q^{\epsilon, \mathbf{h}^1, \mathbf{h}^2}(0) = \zeta_p^{\epsilon, \mathbf{h}^1, \mathbf{h}^2}(0) = 0$ .

The inequality (50) is a straightforward consequence of the inequality (49) and of the following claim: for all  $m \in \mathbb{N}$ , there exists  $C_m(T) \in (0, \infty)$  such that

$$(57) \quad \sup_{(t, q, p) \in [0, T] \times \mathbb{R}^{2d}} \mathbb{E}_{q, p} [|\zeta_q^{\epsilon, \mathbf{h}^1, \mathbf{h}^2}(t)|^{2m}] \leq C_m(T) (|h_q^1|^{2m} + \epsilon^{2m} |h_p^1|^{2m}) (|h_q^2|^{2m} + \epsilon^{2m} |h_p^2|^{2m}).$$

Indeed, it suffices to apply the inequality (56) with  $m = 1$  combined with the Cauchy–Schwarz inequality, and the inequality (57) with  $m = 1$ , and to set either  $\mathbf{h}^1 = (h^1, 0)$ ,  $\mathbf{h}^2 = (0, h^2)$ , or  $\mathbf{h}^1 = (0, h^1)$ ,  $\mathbf{h}^2 = (0, h^2)$ , to obtain (50). It thus remains to prove the claim (57). First, observe that, for all  $t \geq 0$ , one has

$$\begin{aligned} \zeta_p^{\epsilon, \mathbf{h}^1, \mathbf{h}^2}(t) &= \frac{1}{\epsilon} \int_0^t e^{-\frac{t-s}{\epsilon^2}} Df(q^\epsilon(s)) \cdot \zeta_q^{\epsilon, \mathbf{h}^1, \mathbf{h}^2}(s) ds + \frac{1}{\epsilon} \int_0^t e^{-\frac{t-s}{\epsilon^2}} D\sigma(q^\epsilon(s)) \cdot \zeta_q^{\epsilon, \mathbf{h}^1, \mathbf{h}^2}(s) d\beta(s) \\ &\quad + \frac{1}{\epsilon} \int_0^t e^{-\frac{t-s}{\epsilon^2}} D^2 f(q^\epsilon(s)) \cdot (\eta_q^{\epsilon, \mathbf{h}^1}(s), \eta_q^{\epsilon, \mathbf{h}^2}(s)) ds + \frac{1}{\epsilon} \int_0^t e^{-\frac{t-s}{\epsilon^2}} D^2 \sigma(q^\epsilon(s)) \cdot (\eta_q^{\epsilon, \mathbf{h}^1}(s), \eta_q^{\epsilon, \mathbf{h}^2}(s)) d\beta(s). \end{aligned}$$

In addition, for all  $t \geq 0$ , one has

$$\begin{aligned} \zeta_q^{\epsilon, \mathbf{h}^1, \mathbf{h}^2}(t) &= \frac{1}{\epsilon} \int_0^t \zeta_p^{\epsilon, \mathbf{h}^1, \mathbf{h}^2}(s) ds = -\epsilon \zeta_p^{\epsilon, \mathbf{h}^1, \mathbf{h}^2}(t) \\ &\quad + \int_0^t Df(q^\epsilon(s)) \cdot \zeta_q^{\epsilon, \mathbf{h}^1, \mathbf{h}^2}(s) ds + \int_0^t D\sigma(q^\epsilon(s)) \cdot \zeta_q^{\epsilon, \mathbf{h}^1, \mathbf{h}^2}(s) d\beta(s) \\ &\quad + \int_0^t D^2 f(q^\epsilon(s)) \cdot (\eta_q^{\epsilon, \mathbf{h}^1}(s), \eta_q^{\epsilon, \mathbf{h}^2}(s)) ds + \int_0^t D^2 \sigma(q^\epsilon(s)) \cdot (\eta_q^{\epsilon, \mathbf{h}^1}(s), \eta_q^{\epsilon, \mathbf{h}^2}(s)) d\beta(s). \end{aligned}$$

On the one hand, using the Hölder and Burkholder–Davis–Gundy inequalities, one obtains the following upper bound for  $\zeta_p^{\epsilon, \mathbf{h}^1, \mathbf{h}^2}(t)$ : for all  $t \in [0, T]$  one has

$$\begin{aligned} \mathbb{E}[|\zeta_p^{\epsilon, \mathbf{h}^1, \mathbf{h}^2}(t)|^{2m}] &\leq \frac{C_m(T)}{\epsilon^{2m}} \left( \int_0^t \mathbb{E}[|\zeta_q^{\epsilon, \mathbf{h}^1, \mathbf{h}^2}(s)|^{2m}] ds + \int_0^t \mathbb{E}[|\eta_q^{\epsilon, \mathbf{h}^1}(s)|^{2m} |\eta_q^{\epsilon, \mathbf{h}^2}(s)|^{2m}] ds \right) \\ &\leq \frac{C_m(T)}{\epsilon^{2m}} \left( \int_0^t \mathbb{E}[|\zeta_q^{\epsilon, \mathbf{h}^1, \mathbf{h}^2}(s)|^{2m}] ds + (|h_q^1|^{2m} + \epsilon^{2m} |h_p^1|^{2m}) (|h_q^2|^{2m} + \epsilon^{2m} |h_p^2|^{2m}) \right), \end{aligned}$$

where the inequality (56) is used in the second step above.

On the other hand, using the Hölder and Burkholder–Davis–Gundy inequalities, one obtains the following upper bound for  $\zeta_q^{\epsilon, \mathbf{h}^1, \mathbf{h}^2}(t)$ : for all  $t \in [0, T]$  one has

$$\begin{aligned} \mathbb{E}[|\zeta_q^{\epsilon, \mathbf{h}^1, \mathbf{h}^2}(t)|^{2m}] &\leq C_m(T) \epsilon^{2m} \mathbb{E}[|\zeta_p^{\epsilon, \mathbf{h}^1, \mathbf{h}^2}(t)|^{2m}] \\ &\quad + C_m(T) \int_0^t \mathbb{E}[|\zeta_q^{\epsilon, \mathbf{h}^1, \mathbf{h}^2}(s)|^{2m}] ds + C_m(T) \int_0^t \mathbb{E}[|\eta_q^{\epsilon, \mathbf{h}^1}(s)|^{2m} |\eta_q^{\epsilon, \mathbf{h}^2}(s)|^{2m}] ds \\ &\leq C_m(T) \epsilon^{2m} \mathbb{E}[|\zeta_p^{\epsilon, \mathbf{h}^1, \mathbf{h}^2}(t)|^{2m}] \\ &\quad + C_m(T) \left( \int_0^t \mathbb{E}[|\zeta_q^{\epsilon, \mathbf{h}^1, \mathbf{h}^2}(s)|^{2m}] ds + (|h_q^1|^{2m} + \epsilon^{2m} |h_p^1|^{2m}) (|h_q^2|^{2m} + \epsilon^{2m} |h_p^2|^{2m}) \right), \end{aligned}$$

where the inequality (56) is used in the second step above.

Combining the two upper bounds gives the inequality

$$\mathbb{E}[|\zeta_q^{\epsilon, \mathbf{h}^1, \mathbf{h}^2}(t)|^{2m}] \leq C_m(T) \left( \int_0^t \mathbb{E}[|\zeta_q^{\epsilon, \mathbf{h}^1, \mathbf{h}^2}(s)|^{2m}] ds + (|h_q^1|^{2m} + \epsilon^{2m} |h_p^1|^{2m}) (|h_q^2|^{2m} + \epsilon^{2m} |h_p^2|^{2m}) \right)$$



for all  $t \in [0, T]$ . Applying Gronwall's inequality then yields the claim (57). This concludes the proof of the inequality (50).  $\square$

PROOF OF THE INEQUALITY (51). For all  $\mathbf{h}^1 = (h_q^1, h_p^1) \in \mathbb{R}^{2d}$ ,  $\mathbf{h}^2 = (h_q^2, h_p^2) \in \mathbb{R}^{2d}$  and  $\mathbf{h}^3 = (h_q^3, h_p^3) \in \mathbb{R}^{2d}$ , one has

$$\begin{aligned} \nabla_{q,p}^3 u^\epsilon(t, q, p) \cdot (\mathbf{h}^1, \mathbf{h}^2, \mathbf{h}^3) &= \mathbb{E}[\nabla_q \varphi(q^\epsilon(t)) \cdot \xi_q^{\epsilon, \mathbf{h}^1, \mathbf{h}^2, \mathbf{h}^3}(t)] \\ &+ \mathbb{E}[\nabla_q^2 \varphi(q^\epsilon(t)) \cdot (\eta_q^{\epsilon, \mathbf{h}^1}(t), \zeta_q^{\epsilon, \mathbf{h}^2, \mathbf{h}^3}(t))] \\ &+ \mathbb{E}[\nabla_q^2 \varphi(q^\epsilon(t)) \cdot (\eta_q^{\epsilon, \mathbf{h}^2}(t), \zeta_q^{\epsilon, \mathbf{h}^3, \mathbf{h}^1}(t))] \\ &+ \mathbb{E}[\nabla_q^2 \varphi(q^\epsilon(t)) \cdot (\eta_q^{\epsilon, \mathbf{h}^3}(t), \zeta_q^{\epsilon, \mathbf{h}^1, \mathbf{h}^2}(t))] \\ &+ \mathbb{E}[\nabla_q^3 \varphi(q^\epsilon(t)) \cdot (\eta_q^{\epsilon, \mathbf{h}^1}(t), \eta_q^{\epsilon, \mathbf{h}^2}(t), \eta_q^{\epsilon, \mathbf{h}^3}(t))] \end{aligned}$$

where  $t \mapsto \xi_q^{\epsilon, \mathbf{h}^1, \mathbf{h}^2, \mathbf{h}^3}(t) = (\zeta_q^{\epsilon, \mathbf{h}^1, \mathbf{h}^2, \mathbf{h}^3}(t), \zeta_p^{\epsilon, \mathbf{h}^1, \mathbf{h}^2, \mathbf{h}^3}(t)) \in \mathbb{R}^{2d}$  is solution of the stochastic differential equation

$$\left\{ \begin{aligned} d\zeta_q^{\epsilon, \mathbf{h}^1, \mathbf{h}^2, \mathbf{h}^3}(t) &= \frac{\zeta_p^{\epsilon, \mathbf{h}^1, \mathbf{h}^2, \mathbf{h}^3}(t)}{\epsilon} dt \\ d\zeta_p^{\epsilon, \mathbf{h}^1, \mathbf{h}^2, \mathbf{h}^3}(t) &= -\frac{1}{\epsilon^2} \zeta_p^{\epsilon, \mathbf{h}^1, \mathbf{h}^2, \mathbf{h}^3}(t) dt + \frac{1}{\epsilon} Df(q^\epsilon(t)) \cdot \xi_q^{\epsilon, \mathbf{h}^1, \mathbf{h}^2, \mathbf{h}^3}(t) dt + \frac{1}{\epsilon} D\sigma(q^\epsilon(t)) \cdot \xi_q^{\epsilon, \mathbf{h}^1, \mathbf{h}^2, \mathbf{h}^3} d\beta(t) \\ &+ \frac{1}{\epsilon} D^2 f(q^\epsilon(t)) \cdot (\eta_q^{\epsilon, \mathbf{h}^1}(t), \zeta_q^{\epsilon, \mathbf{h}^2, \mathbf{h}^3}(t)) dt + \frac{1}{\epsilon} D^2 \sigma(q^\epsilon(t)) \cdot (\eta_q^{\epsilon, \mathbf{h}^1}(t), \zeta_q^{\epsilon, \mathbf{h}^2, \mathbf{h}^3}(t)) d\beta(t) \\ &+ \frac{1}{\epsilon} D^2 f(q^\epsilon(t)) \cdot (\eta_q^{\epsilon, \mathbf{h}^2}(t), \zeta_q^{\epsilon, \mathbf{h}^3, \mathbf{h}^1}(t)) dt + \frac{1}{\epsilon} D^2 \sigma(q^\epsilon(t)) \cdot (\eta_q^{\epsilon, \mathbf{h}^2}(t), \zeta_q^{\epsilon, \mathbf{h}^2, \mathbf{h}^3}(t)) d\beta(t) \\ &+ \frac{1}{\epsilon} D^2 f(q^\epsilon(t)) \cdot (\eta_q^{\epsilon, \mathbf{h}^3}(t), \eta_q^{\epsilon, \mathbf{h}^1, \mathbf{h}^2}(t)) dt + \frac{1}{\epsilon} D^2 \sigma(q^\epsilon(t)) \cdot (\eta_q^{\epsilon, \mathbf{h}^3}(t), \zeta_q^{\epsilon, \mathbf{h}^1, \mathbf{h}^2}(t)) d\beta(t) \\ &+ \frac{1}{\epsilon} D^3 f(q^\epsilon(t)) \cdot (\eta_q^{\epsilon, \mathbf{h}^1}(t), \eta_q^{\epsilon, \mathbf{h}^2}(t), \eta_q^{\epsilon, \mathbf{h}^3}(t)) dt + \frac{1}{\epsilon} D^3 \sigma(q^\epsilon(t)) \cdot (\eta_q^{\epsilon, \mathbf{h}^1}(t), \eta_q^{\epsilon, \mathbf{h}^2}(t), \eta_q^{\epsilon, \mathbf{h}^3}(t)) d\beta(t) \end{aligned} \right.$$

with initial value  $\xi_q^{\epsilon, \mathbf{h}^1, \mathbf{h}^2, \mathbf{h}^3}(0) = 0$ , equivalently  $\zeta_q^{\epsilon, \mathbf{h}^1, \mathbf{h}^2, \mathbf{h}^3}(0) = \zeta_p^{\epsilon, \mathbf{h}^1, \mathbf{h}^2, \mathbf{h}^3}(0) = 0$ .

The inequality (51) is a straightforward consequence of the inequality (49) and of the following claim: for all  $m \in \mathbb{N}$ , there exists  $C_m(T) \in (0, \infty)$  such that

$$(58) \quad \sup_{(t, q, p) \in [0, T] \times \mathbb{R}^{2d}} \mathbb{E}_{q, p} [|\xi_q^{\epsilon, \mathbf{h}^1, \mathbf{h}^2, \mathbf{h}^3}(t)|^{2m}] \leq C_m(T) (|h_q^1|^{2m} + \epsilon^{2m} |h_p^1|^{2m}) (|h_q^2|^{2m} + \epsilon^{2m} |h_p^2|^{2m}) (|h_q^3|^{2m} + \epsilon^{2m} |h_p^3|^{2m}).$$

Indeed, it suffices to apply the inequality (56) with  $m = 2$  combined with the Cauchy–Schwarz inequality, the inequality (57) with  $m = 1$  and the inequality (58) with  $m = 1$ , and to set either  $\mathbf{h}^1 = (h^1, 0)$ ,  $\mathbf{h}^2 = (0, h^2)$ ,  $\mathbf{h}^3 = (0, h^3)$  or  $\mathbf{h}^1 = (0, h^1)$ ,  $\mathbf{h}^2 = (0, h^2)$ ,  $\mathbf{h}^3 = (0, h^3)$ . It thus remains to prove the claim (57).

First, observe that, for all  $t \geq 0$ , one has

$$\begin{aligned} \xi_q^{\epsilon, \mathbf{h}^1, \mathbf{h}^2, \mathbf{h}^3}(t) &= \frac{1}{\epsilon} \int_0^t e^{-\frac{t-s}{\epsilon^2}} Df(q^\epsilon(s)) \cdot \xi_q^{\epsilon, \mathbf{h}^1, \mathbf{h}^2, \mathbf{h}^3}(s) ds + \frac{1}{\epsilon} \int_0^t e^{-\frac{t-s}{\epsilon^2}} D\sigma(q^\epsilon(s)) \cdot \xi_q^{\epsilon, \mathbf{h}^1, \mathbf{h}^2, \mathbf{h}^3}(s) d\beta(s) \\ &+ \frac{1}{\epsilon} \int_0^t e^{-\frac{t-s}{\epsilon^2}} D^2 f(q^\epsilon(s)) \cdot (\eta_q^{\epsilon, \mathbf{h}^1}(s), \zeta_q^{\epsilon, \mathbf{h}^2, \mathbf{h}^3}(s)) ds + \frac{1}{\epsilon} \int_0^t e^{-\frac{t-s}{\epsilon^2}} D^2 \sigma(q^\epsilon(s)) \cdot (\eta_q^{\epsilon, \mathbf{h}^1}(s), \zeta_q^{\epsilon, \mathbf{h}^2, \mathbf{h}^3}(s)) d\beta(s) \\ &+ \frac{1}{\epsilon} \int_0^t e^{-\frac{t-s}{\epsilon^2}} D^2 f(q^\epsilon(s)) \cdot (\eta_q^{\epsilon, \mathbf{h}^2}(s), \eta_q^{\epsilon, \mathbf{h}^3, \mathbf{h}^1}(s)) ds + \frac{1}{\epsilon} \int_0^t e^{-\frac{t-s}{\epsilon^2}} D^2 \sigma(q^\epsilon(s)) \cdot (\eta_q^{\epsilon, \mathbf{h}^2}(s), \zeta_q^{\epsilon, \mathbf{h}^3, \mathbf{h}^1}(s)) d\beta(s) \\ &+ \frac{1}{\epsilon} \int_0^t e^{-\frac{t-s}{\epsilon^2}} D^2 f(q^\epsilon(s)) \cdot (\eta_q^{\epsilon, \mathbf{h}^3}(s), \eta_q^{\epsilon, \mathbf{h}^1, \mathbf{h}^2}(s)) ds + \frac{1}{\epsilon} \int_0^t e^{-\frac{t-s}{\epsilon^2}} D^2 \sigma(q^\epsilon(s)) \cdot (\eta_q^{\epsilon, \mathbf{h}^3}(s), \eta_q^{\epsilon, \mathbf{h}^1, \mathbf{h}^2}(s)) d\beta(s) \\ &+ \frac{1}{\epsilon} \int_0^t e^{-\frac{t-s}{\epsilon^2}} D^3 f(q^\epsilon(s)) \cdot (\eta_q^{\epsilon, \mathbf{h}^1}(s), \eta_q^{\epsilon, \mathbf{h}^2}(s), \eta_q^{\epsilon, \mathbf{h}^3}(s)) ds \\ &+ \frac{1}{\epsilon} \int_0^t e^{-\frac{t-s}{\epsilon^2}} D^3 \sigma(q^\epsilon(s)) \cdot (\eta_q^{\epsilon, \mathbf{h}^1}(s), \eta_q^{\epsilon, \mathbf{h}^2}(s), \eta_q^{\epsilon, \mathbf{h}^3}(s)) d\beta(s). \end{aligned}$$

In addition, for all  $t \geq 0$ , one has

$$\begin{aligned}
\xi_q^{\epsilon, \mathbf{h}^1, \mathbf{h}^2, \mathbf{h}^3}(t) &= \frac{1}{\epsilon} \int_0^t \xi_p^{\epsilon, \mathbf{h}^1, \mathbf{h}^2, \mathbf{h}^3}(s) ds = -\epsilon \xi_p^{\epsilon, \mathbf{h}^1, \mathbf{h}^2, \mathbf{h}^3}(t) \\
&+ \int_0^t Df(q^\epsilon(s)) \cdot \xi_q^{\epsilon, \mathbf{h}^1, \mathbf{h}^2, \mathbf{h}^3}(s) ds + \int_0^t D\sigma(q^\epsilon(s)) \cdot \xi_q^{\epsilon, \mathbf{h}^1, \mathbf{h}^2, \mathbf{h}^3}(s) d\beta(s) \\
&+ \int_0^t D^2 f(q^\epsilon(s)) \cdot (\eta_q^{\epsilon, \mathbf{h}^1}(s), \zeta_q^{\epsilon, \mathbf{h}^2, \mathbf{h}^3}(s)) ds + \int_0^t D^2 \sigma(q^\epsilon(s)) \cdot (\eta_q^{\epsilon, \mathbf{h}^1}(s), \zeta_q^{\epsilon, \mathbf{h}^2, \mathbf{h}^3}(s)) d\beta(s) \\
&+ \int_0^t D^2 f(q^\epsilon(s)) \cdot (\eta_q^{\epsilon, \mathbf{h}^2}(s), \eta_q^{\epsilon, \mathbf{h}^3, \mathbf{h}^1}(s)) ds + \int_0^t D^2 \sigma(q^\epsilon(s)) \cdot (\eta_q^{\epsilon, \mathbf{h}^2}(s), \zeta_q^{\epsilon, \mathbf{h}^3, \mathbf{h}^1}(s)) d\beta(s) \\
&+ \int_0^t D^2 f(q^\epsilon(s)) \cdot (\eta_q^{\epsilon, \mathbf{h}^3}(s), \eta_q^{\epsilon, \mathbf{h}^1, \mathbf{h}^2}(s)) ds + \int_0^t D^2 \sigma(q^\epsilon(s)) \cdot (\eta_q^{\epsilon, \mathbf{h}^3}(s), \eta_q^{\epsilon, \mathbf{h}^1, \mathbf{h}^2}(s)) d\beta(s) \\
&+ \int_0^t D^3 f(q^\epsilon(s)) \cdot (\eta_q^{\epsilon, \mathbf{h}^1}(s), \eta_q^{\epsilon, \mathbf{h}^2}(s), \eta_q^{\epsilon, \mathbf{h}^3}(s)) ds \\
&+ \int_0^t D^3 \sigma(q^\epsilon(s)) \cdot (\eta_q^{\epsilon, \mathbf{h}^1}(s), \eta_q^{\epsilon, \mathbf{h}^2}(s), \eta_q^{\epsilon, \mathbf{h}^3}(s)) d\beta(s).
\end{aligned}$$

On the one hand, using the Hölder and Burkholder–Davis–Gundy inequalities, one obtains the following upper bound for  $\xi_p^{\epsilon, \mathbf{h}^1, \mathbf{h}^2, \mathbf{h}^3}(t)$ : for all  $t \in [0, T]$  one has

$$\begin{aligned}
\mathbb{E}[|\xi_p^{\epsilon, \mathbf{h}^1, \mathbf{h}^2, \mathbf{h}^3}(t)|^{2m}] &\leq \frac{C_m(T)}{\epsilon^{2m}} \int_0^t \mathbb{E}[|\xi_q^{\epsilon, \mathbf{h}^1, \mathbf{h}^2, \mathbf{h}^3}(s)|^{2m}] ds \\
&+ \frac{C_m(T)}{\epsilon^{2m}} \int_0^t (\mathbb{E}[|\eta_q^{\epsilon, \mathbf{h}^1}(s)|^{2m} |\zeta_q^{\epsilon, \mathbf{h}^2, \mathbf{h}^3}(s)|^{2m}] + \mathbb{E}[|\eta_q^{\epsilon, \mathbf{h}^2}(s)|^{2m} |\zeta_q^{\epsilon, \mathbf{h}^3, \mathbf{h}^1}(s)|^{2m}] + \mathbb{E}[|\eta_q^{\epsilon, \mathbf{h}^3}(s)|^{2m} |\zeta_q^{\epsilon, \mathbf{h}^1, \mathbf{h}^2}(s)|^{2m}]) ds \\
&+ \frac{C_m(T)}{\epsilon^{2m}} \int_0^t \mathbb{E}[|\eta_q^{\epsilon, \mathbf{h}^1}(s)|^{2m} |\eta_q^{\epsilon, \mathbf{h}^2}(s)|^{2m} |\eta_q^{\epsilon, \mathbf{h}^3}(s)|^{2m}] ds \\
&\leq \frac{C_m(T)}{\epsilon^{2m}} \left( \int_0^t \mathbb{E}[|\xi_q^{\epsilon, \mathbf{h}^1, \mathbf{h}^2, \mathbf{h}^3}(s)|^{2m}] ds + (|h_q^1|^{2m} + \epsilon^{2m} |h_p^1|^{2m}) (|h_q^2|^{2m} + \epsilon^{2m} |h_p^2|^{2m}) (|h_q^3|^{2m} + \epsilon^{2m} |h_p^3|^{2m}) \right),
\end{aligned}$$

where the inequalities (56) and (57) are used in the second step above.

On the other hand, using the Hölder and Burkholder–Davis–Gundy inequalities, one obtains the following upper bound for  $\xi_q^{\epsilon, \mathbf{h}^1, \mathbf{h}^2, \mathbf{h}^3}(t)$ : for all  $t \in [0, T]$  one has

$$\begin{aligned}
\mathbb{E}[|\xi_q^{\epsilon, \mathbf{h}^1, \mathbf{h}^2, \mathbf{h}^3}(t)|^{2m}] &\leq C_m(T) \epsilon^{2m} \mathbb{E}[|\xi_p^{\epsilon, \mathbf{h}^1, \mathbf{h}^2, \mathbf{h}^3}(t)|^{2m}] \\
&+ C_m(T) \int_0^t \mathbb{E}[|\xi_q^{\epsilon, \mathbf{h}^1, \mathbf{h}^2, \mathbf{h}^3}(s)|^{2m}] ds \\
&+ C_m(T) \int_0^t \mathbb{E}[|\eta_q^{\epsilon, \mathbf{h}^1}(s)|^{2m} |\zeta_q^{\epsilon, \mathbf{h}^2, \mathbf{h}^3}(s)|^{2m}] ds \\
&+ C_m(T) \int_0^t \mathbb{E}[|\eta_q^{\epsilon, \mathbf{h}^2}(s)|^{2m} |\zeta_q^{\epsilon, \mathbf{h}^3, \mathbf{h}^1}(s)|^{2m}] ds \\
&+ C_m(T) \int_0^t \mathbb{E}[|\eta_q^{\epsilon, \mathbf{h}^3}(s)|^{2m} |\zeta_q^{\epsilon, \mathbf{h}^1, \mathbf{h}^2}(s)|^{2m}] ds \\
&+ C_m(T) \int_0^t \mathbb{E}[|\eta_q^{\epsilon, \mathbf{h}^1}(s)|^{2m} |\eta_q^{\epsilon, \mathbf{h}^2}(s)|^{2m} |\eta_q^{\epsilon, \mathbf{h}^3}(s)|^{2m}] ds \\
&\leq C_m(T) \epsilon^{2m} \mathbb{E}[|\xi_p^{\epsilon, \mathbf{h}^1, \mathbf{h}^2, \mathbf{h}^3}(t)|^{2m}] \\
&+ C_m(T) \int_0^t \mathbb{E}[|\xi_q^{\epsilon, \mathbf{h}^1, \mathbf{h}^2, \mathbf{h}^3}(s)|^{2m}] ds \\
&+ (|h_q^1|^{2m} + \epsilon^{2m} |h_p^1|^{2m}) (|h_q^2|^{2m} + \epsilon^{2m} |h_p^2|^{2m}) (|h_q^3|^{2m} + \epsilon^{2m} |h_p^3|^{2m}),
\end{aligned}$$

where the inequalities (56) and (57) are used in the second step above.

Combining the two upper bounds gives the inequality

$$\begin{aligned} \mathbb{E}[|\xi_q^{\epsilon, \mathbf{h}^1, \mathbf{h}^2, \mathbf{h}^3}(t)|^{2m}] &\leq C_m(T) \int_0^t \mathbb{E}[|\xi_q^{\epsilon, \mathbf{h}^1, \mathbf{h}^2, \mathbf{h}^3}(s)|^{2m}] ds \\ &\quad + C_m(T) \left( (|h_q^1|^{2m} + \epsilon^{2m} |h_p^1|^{2m}) (|h_q^2|^{2m} + \epsilon^{2m} |h_p^2|^{2m}) (|h_q^3|^{2m} + \epsilon^{2m} |h_p^3|^{2m}) \right) \end{aligned}$$

for all  $t \in [0, T]$ . Applying Gronwall's inequality then yields the claim (58). This concludes the proof of the inequality (51). □

### Acknowledgements

This work is partially supported by the projects ADA (ANR-19-CE40-0019-02) and SIMALIN (ANR-19-CE40-0016) operated by the French National Research Agency.

### References

- [1] A. Abdulle, G. A. Pavliotis, and U. Vaes. Spectral methods for multiscale stochastic differential equations. *SIAM/ASA J. Uncertain. Quantif.*, 5(1):720–761, 2017.
- [2] I. Almuslimani, P. Chartier, M. Lemou, and F. Méhats. Uniformly accurate schemes for drift-oscillatory stochastic differential equations. *Appl. Numer. Math.*, 181:468–482, 2022.
- [3] C.-E. Bréhier. Uniform error bounds for numerical schemes applied to multiscale SDEs in a Wong–Zakai diffusion approximation regime. *arXiv preprint*, 2022.
- [4] C.-E. Bréhier and S. Rakotonirina-Ricquebourg. On asymptotic preserving schemes for a class of stochastic differential equations in averaging and diffusion approximation regimes. *Multiscale Model. Simul.*, 20(1):118–163, 2022.
- [5] S. Cerrai. *Second order PDE's in finite and infinite dimension*, volume 1762 of *Lecture Notes in Mathematics*. Springer-Verlag, Berlin, 2001. A probabilistic approach.
- [6] W. E, D. Liu, and E. Vanden-Eijnden. Analysis of multiscale methods for stochastic differential equations. *Comm. Pure Appl. Math.*, 58(11):1544–1585, 2005.
- [7] J. Frank and G. A. Gottwald. A note on statistical consistency of numerical integrators for multiscale dynamics. *Multiscale Model. Simul.*, 16(2):1017–1033, 2018.
- [8] D. Givon, I. G. Kevrekidis, and R. Kupferman. Strong convergence of projective integration schemes for singularly perturbed stochastic differential systems. *Commun. Math. Sci.*, 4(4):707–729, 2006.
- [9] E. Gobet. *Monte-Carlo methods and stochastic processes*. CRC Press, Boca Raton, FL, 2016. From linear to non-linear.
- [10] I. G. Kevrekidis, C. W. Gear, J. M. Hyman, P. G. Kevrekidis, O. Rumborg, and C. Theodoropoulos. Equation-free, coarse-grained multiscale computation: enabling microscopic simulators to perform system-level analysis. *Commun. Math. Sci.*, 1(4):715–762, 2003.
- [11] P. E. Kloeden and E. Platen. *Numerical solution of stochastic differential equations*, volume 23 of *Applications of Mathematics (New York)*. Springer-Verlag, Berlin, 1992.
- [12] A. Laurent. A uniformly accurate scheme for the numerical integration of penalized Langevin dynamics. *SIAM J. Sci. Comput.*, to appear.
- [13] F. Legoll, T. Lelièvre, K. Myerscough, and G. Samaey. Parareal computation of stochastic differential equations with time-scale separation: a numerical convergence study. *Comput. Vis. Sci.*, 23(1-4):9, 2020.
- [14] T. Lelièvre, M. Rousset, and G. Stoltz. *Free energy computations*. Imperial College Press, London, 2010. A mathematical perspective.
- [15] T. Li, A. Abdulle, and W. E. Effectiveness of implicit methods for stiff stochastic differential equations. *Commun. Comput. Phys.*, 3(2):295–307, 2008.
- [16] G. N. Milstein and M. V. Tretyakov. *Stochastic numerics for mathematical physics*. Scientific Computation. Springer, Cham, 2021.
- [17] G. Pagès. *Numerical probability*. Universitext. Springer, Cham, 2018. An introduction with applications to finance.
- [18] G. A. Pavliotis and A. M. Stuart. *Multiscale methods*, volume 53 of *Texts in Applied Mathematics*. Springer, New York, 2008. Averaging and homogenization.
- [19] S. Rakotonirina-Ricquebourg. *Etude théorique et numérique d'équations cinétiques stochastiques multi-échelles*. PhD thesis.
- [20] M. Rousset, Y. Xu, and P.-A. Zitt. A weak overdamped limit theorem for Langevin processes. *ALEA Lat. Am. J. Probab. Math. Stat.*, 17(1):1–21, 2020.
- [21] D. Talay. Discrétisation d'une équation différentielle stochastique et calcul approché d'espérances de fonctionnelles de la solution. *RAIRO Modél. Math. Anal. Numér.*, 20(1):141–179, 1986.
- [22] D. Talay and L. Tubaro. Expansion of the global error for numerical schemes solving stochastic differential equations. *Stochastic Anal. Appl.*, 8(4):483–509 (1991), 1990.
- [23] H. Vandecasteele, P. a. Zieliński, and G. Samaey. Efficiency of a micro-macro acceleration method for scale-separated stochastic differential equations. *Multiscale Model. Simul.*, 18(3):1272–1298, 2020.

UNIV LYON, UNIVERSITÉ CLAUDE BERNARD LYON 1, CNRS UMR 5208, INSTITUT CAMILLE JORDAN, 43 BLVD. DU 11  
NOVEMBRE 1918, F-69622 VILLEURBANNE CEDEX, FRANCE  
*E-mail address:* `brehier@math.univ-lyon1.fr`