



HAL
open science

Exploring the behavior of asymptotic preserving schemes for a class of stochastic systems

Charles-Édouard Bréhier

► **To cite this version:**

Charles-Édouard Bréhier. Exploring the behavior of asymptotic preserving schemes for a class of stochastic systems. 2025. <hal-05287175>

HAL Id: hal-05287175

<https://hal.science/hal-05287175v1>

Preprint submitted on 28 Sep 2025

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.



HAL Authorization

EXPLORING THE BEHAVIOR OF ASYMPTOTIC PRESERVING SCHEMES FOR A CLASS OF STOCHASTIC SYSTEMS

CHARLES-EDOUARD BRÉHIER

ABSTRACT. We consider a class of slow-fast stochastic systems, in the regime governed by the averaging principle. The fast component is a Brownian motion on a torus. We first consider a system where the slow component solves an ordinary differential equation: we study a numerical scheme, which is shown to be asymptotic preserving and to satisfy uniform strong error estimates. Numerical experiments illustrate the results. We then provide numerical experiments for variants of the systems where the slow evolution is also driven by a Brownian motion. Finally, we propose and test a candidate higher-order method.

AMS Classification. 60H35, 65C30.

Keywords. Stochastic differential equations. Multiscale systems. Asymptotic preserving schemes.

1. INTRODUCTION

In this article, we consider a class of systems of stochastic differential equations depending on a small parameter ϵ :

$$\begin{cases} dx^\epsilon(t) = f(x^\epsilon(t), y^\epsilon(t)) dt, & t \geq 0; & x^\epsilon(0) = x_0, \\ dy^\epsilon(t) = \frac{d\beta(t)}{\sqrt{\epsilon}}, & t \geq 0; & y^\epsilon(0) = y_0. \end{cases}$$

The fast component y^ϵ is a Brownian motion, and is considered as taking values in the one-dimensional torus $\mathbb{T} = \mathbb{R}/\mathbb{Z}$. The slow component x^ϵ is solution to an ordinary differential equation, where the vector field f depends on both components (and is assumed to be periodic in the second variable).

Numerical integration of multi-scale systems of the type above is not straightforward, since the small parameter ϵ usually imposes stringent and prohibitive conditions on the time-step size Δt . The literature contains many works devoted to the construction of efficient multiscale integrators, see for instance [1, 2, 9, 10, 13, 14, 15, 16, 18]. The list of references is not exhaustive.

Often, the construction of efficient integrators is based on first understanding the behavior of the multiscale system when the parameter ϵ goes to 0. For the class of systems considered in this work, this is governed by the averaging principle, see for instance the monograph [17]: the fast component y^ϵ has an ergodic behavior and can be eliminated from the slow component x^ϵ . As a result when ϵ goes to 0 the slow component x^ϵ converges to the solution X to an ordinary differential equation

$$dX(t) = F(X(t)) dt, \quad t \geq 0; \quad X(0) = x_0,$$

where $F(x)$ is obtained by averaging $f(x, y)$ with respect to the uniform distribution on the torus for y :

$$F(x) = \int_{\mathbb{T}} f(x, y) dy, \quad \forall x \in \mathbb{R}.$$

Details are presented in Section 2. This averaging principle inspires the construction of the numerical integrators using different approaches as described in the references listed above.

In this article, we focus on a particular point of view: we consider asymptotic preserving methods, which have first been constructed for certain deterministic problems [11, 12]. For stochastic systems, asymptotic preserving schemes have been first studied in [8], where two regimes are considered: the averaging and the diffusion approximation regime. Here we only focus on the first regime. Note that in the averaging regime asymptotic preserving schemes have been considered for stochastic differential equations driven by a fractional Brownian motion (in the slow component) in [3] and also for a class of stochastic partial differential equations in [7]. The two articles [5] and [6] presents further results in the diffusion approximation regime, in the Smoluchowski–Kramers and the Wong–Zakai regimes respectively.

The guideline of asymptotic preserving schemes is to construct numerical schemes such that the following diagram is commutative:

$$\begin{array}{ccc} x_N^\epsilon & \xrightarrow{\Delta t \rightarrow 0} & x^\epsilon(T) \\ \downarrow \epsilon \rightarrow 0 & & \downarrow \epsilon \rightarrow 0 \\ X_N & \xrightarrow{\Delta t \rightarrow 0} & X(T), \end{array}$$

where the final time $T = N\Delta t$ is fixed. Moreover, one may obtain error estimates which are uniform with respect to ϵ when Δt goes to 0. Numerical schemes which satisfy such error bounds are referred to as uniformly accurate.

The main contributions of this article are to show how to construct an asymptotic preserving scheme for the considered class of systems and how to prove uniform (strong) error estimates. The behavior of the scheme is explored using numerical experiments and rigorous error bounds are also obtained.

The article is organized as follows. Section 2 presents the problem, the averaging principle, the scheme and some numerical experiments. Section 3 then presents arguments to explain the behavior observed in the numerical experiments, in particular to show that the scheme is asymptotic preserving and uniformly accurate. The proofs are technical and are postponed to the appendix for convenience. Sections 4 and 5 mainly present numerical experiments: first to investigate variants of the systems, second to explore whether more efficient schemes can be designed.

Notation. The torus is defined as $\mathbb{T} = \mathbb{R}/\mathbb{Z}$. The projection of a real number $z \in \mathbb{R}$ onto the torus \mathbb{T} is denoted by $\hat{z} \in \mathbb{T}$.

The systems considered in this article are driven by a Brownian motion $(\beta(t))_{t \geq 0}$, which is defined on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$. The expectation operator is denoted by $\mathbb{E}[\cdot]$.

Given a final time T , the time-step size of the numerical schemes is denoted by $\Delta t = T/N$ where $N \in \mathbb{N}^*$ is an arbitrary integer. The associated grid times are denoted by $t_n = n\Delta t$ for $0 \leq n \leq N$. Moreover, the Brownian increments are denoted by

$$(1) \quad \delta\beta_{n+1} = \beta(t_{n+1}) - \beta(t_n), \quad \forall n \in \{0, \dots, N\}.$$

The time-scale separation parameter ϵ is assumed without loss of generality to take values in $(0, 1)$.

The initial values $x_0 \in \mathbb{R}$ and $y_0 \in \mathbb{T}$ are supposed to be independent of ϵ for simplicity.

Finally, the vector field appearing in the evolution of the slow component x^ϵ is denoted by

$$(2) \quad f : (x, y) \in \mathbb{R} \times \mathbb{T} \mapsto f(x, y) \in \mathbb{R}.$$

For well-posedness of the system and for justifying the convergence of numerical schemes, it is assumed that f is globally Lipschitz continuous. In addition, for simplicity it is assumed that f is bounded. Stronger regularity conditions may be imposed for some results.

Note that the presentation is made in dimension $d = 1$ for simplicity, however the arguments could easily be generalized in arbitrary dimension.

Review of the results from [8]. Before proceeding with the main contributions of this article, let us briefly review the construction and analysis of asymptotic preserving schemes from [8]. This is interesting to highlight the novelties of this work.

In [8], the fast component y^ϵ is an Ornstein–Uhlenbeck process instead of being a Brownian motion on the torus: one considers

$$\begin{cases} dx^\epsilon(t) = f(x^\epsilon(t), y^\epsilon(t)) dt, & t \geq 0; & x^\epsilon(0) = x_0, \\ dy^\epsilon(t) = -\frac{y^\epsilon(t)}{\epsilon} dt + \frac{\sqrt{2}}{\sqrt{\epsilon}} d\beta(t), & t \geq 0; & y^\epsilon(0) = y_0. \end{cases}$$

The averaging principle states that x^ϵ converges to the solution X to the ordinary differential equation

$$dX(t) = F(X(t)) dt, \quad t \geq 0; \quad X(0) = x_0,$$

where F is computed by averaging $f(x, y)$ with respect to a standard Gaussian distribution $\nu = \mathcal{N}(0, 1)$ for y :

$$F(x) = \int_{\mathbb{R}} f(x, y) d\nu(y), \quad \forall x \in \mathbb{R}.$$

The asymptotic preserving scheme proposed in [8] is given as follows:

$$\begin{cases} x_{n+1}^\epsilon = x_n^\epsilon + \Delta t f(x_n^\epsilon, y_{n+1}^\epsilon) \\ y_{n+1}^\epsilon = e^{-\frac{\Delta t}{\epsilon}} y_n^\epsilon + \sqrt{1 - e^{-\frac{2\Delta t}{\epsilon}}} \gamma_{n+1}, \end{cases}$$

where $(\gamma_n)_{n \geq 1}$ is a sequence of independent standard Gaussian random variables.

It is shown that the scheme above is asymptotic preserving, with limiting scheme when $\epsilon \rightarrow 0$ given by

$$X_{n+1} = X_n + \Delta t f(X_n, \gamma_{n+1}).$$

The main reason which justifies that the scheme is asymptotic preserving, more precisely that the limiting scheme is consistent with the limiting equation, is the property on the conditional expectation

$$\mathbb{E}[f(X_n, \gamma_{n+1}) | X_n] = F(X_n).$$

Finally the main result in [8] are the following uniform weak error estimates

$$(3) \quad \sup_{\epsilon \in (0, 1)} |\mathbb{E}[\varphi(x_N^\epsilon)] - \varphi(x^\epsilon(T))| \leq C\sqrt{\Delta t},$$

for sufficiently smooth functions φ . The rate of convergence with respect to Δt is equal to 1/2 when considered uniformly with respect to ϵ , but would be 1 for fixed ϵ (with error constant C_ϵ depending on ϵ). The weak order of convergence 1/2 may be suboptimal. Note that below we will mainly consider strong error estimates instead of weak error estimates, and this has not been considered previously.

2. DESCRIPTION OF THE PROBLEM

2.1. The system and the scheme. In this article, we consider slow-fast systems of stochastic differential equations

$$(4) \quad \begin{cases} dx^\epsilon(t) = f(x^\epsilon(t), y^\epsilon(t)) dt, & t \geq 0, \\ dy^\epsilon(t) = \frac{d\beta(t)}{\sqrt{\epsilon}}, & t \geq 0, \\ x^\epsilon(0) = x_0, & y^\epsilon(0) = y_0. \end{cases}$$

Since the vector field f given by (2) is globally Lipschitz continuous, for any value of $\epsilon \in (0, 1)$ the system (4) admits a unique solution.

When $\epsilon \rightarrow 0$, the behavior of the slow component x^ϵ is governed by the so-called averaging principle. The fast process y^ϵ satisfies

$$y^\epsilon(t) = y_0 + \frac{\beta(t)}{\sqrt{\epsilon}}, \quad \forall t \geq 0.$$

Owing to Lemma 3.4, for any positive time $t \in (0, \infty)$, the random variable $\widehat{y^\epsilon(t)}$ converges in distribution to the uniform distribution on \mathbb{T} when $\epsilon \rightarrow 0$. This suggests that the fast component y^ϵ may be eliminated when $\epsilon \rightarrow 0$. When $\epsilon \rightarrow 0$, an averaging principle holds: one has the convergence of the slow component x^ϵ to the solution X of the ordinary differential equation

$$(5) \quad \begin{cases} dX(t) = F(X(t)) dt, & t \geq 0, \\ X(0) = x_0 \end{cases}$$

where the mapping $F : \mathbb{R} \rightarrow \mathbb{R}$ is defined from the vector field f by averaging with respect to the fast variable y according to the uniform distribution on \mathbb{T} : the averaged coefficient F is defined by

$$(6) \quad F(x) = \int_{\mathbb{T}} f(x, y) dy, \quad \forall x \in \mathbb{R}.$$

It is straightforward to check that F is bounded and globally Lipschitz continuous, therefore the ordinary differential equation (5) admits a unique solution.

For the discretization of the system (4), the following scheme is proposed: given the time-step size $\Delta t = T/N$ and recalling the definition (1) of the Brownian increments, for all $n \in \{0, \dots, N-1\}$ set

$$(7) \quad \begin{cases} x_{n+1}^\epsilon = x_n^\epsilon + \Delta t f(x_n^\epsilon, y_{n+1}^\epsilon), \\ y_{n+1}^\epsilon = y_n^\epsilon + \frac{\delta \beta_{n+1}}{\sqrt{\epsilon}}, \end{cases}$$

with the initial values $x_0^\epsilon = x_0$ and $y_0^\epsilon = y_0$.

2.2. Numerical experiments. Before proceeding with the error analysis, we report the results of numerical experiments concerning the convergence of the strong error

$$\left(\mathbb{E}[|x_N^\epsilon - x^\epsilon(t_N)|^2] \right)^{\frac{1}{2}}$$

when $\Delta t = T/N \rightarrow 0$, depending on the parameter ϵ .

The initial values are $x_0 = 0$ and $y_0 = 0$. The final time is $T = 1$. The vector field f is given by

$$f(x, y) = \cos(4\pi x) \cos^2(2\pi(x + y)).$$

The time-step size Δt takes values in

$$\{2^{-j} \cdot 10^{-2}; -2 \leq j \leq 10\} = \left\{ \frac{1}{25}, \frac{1}{50}, \frac{1}{100}, \dots, \frac{1}{51200}, \frac{1}{102400} \right\}.$$

The strong error is estimated by a Monte Carlo average over 10^2 independent realizations. The reference solution is computed using the scheme (7) with time-step size $\Delta t_{\text{ref}} = 2^{-11} \cdot 10^{-2}$.

Figure 1 displays how the strong error depends on the time-step size Δt as $\Delta t \rightarrow 0$, for different values of the time-scale separation parameter ϵ : on the left one has $\epsilon \in \{10^{-1}, 10^{-2}, 10^{-3}, 10^{-4}\}$ and on the right one has $\epsilon \in \{8 \cdot 10^{-3}, 4 \cdot 10^{-3}, 2 \cdot 10^{-3}, 10^{-3}\}$.

Three reference lines with slopes $1/3$, $1/2$ and 1 are indicated.

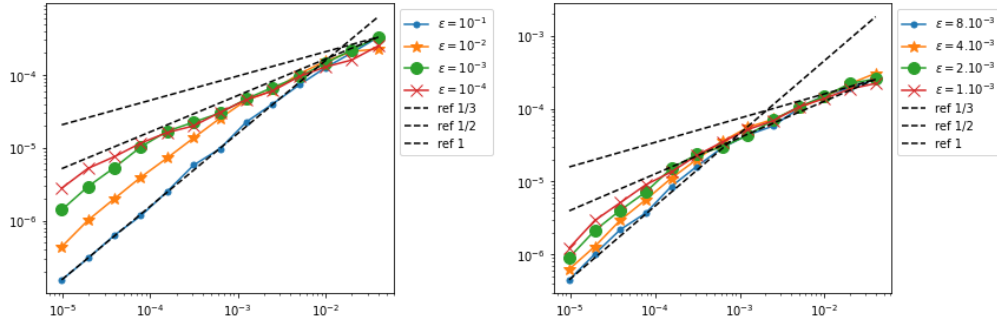


FIGURE 1. Behavior of the strong error for the scheme (7) applied to (5) as a function of the time-step size Δt for different values of ϵ . Left: $\epsilon \in \{10^{-1}, 10^{-2}, 10^{-3}, 10^{-4}\}$. Right: $\epsilon \in \{8 \cdot 10^{-3}, 4 \cdot 10^{-3}, 2 \cdot 10^{-3}, 10^{-3}\}$.

It is observed in Figure 1 that

- for any fixed value of ϵ , the strong error converges to 0 with a rate 1 with respect to Δt , for sufficiently small Δt ,
- the convergence with rate 1 with respect to Δt is not uniform with respect to ϵ ,
- a uniform convergence with respect to ϵ with rate $1/2$ is observed.

The behavior is explained by the results described below:

- for fixed ϵ , the convergence with rate 1 is given in Proposition 3.2, with an error constant depending on ϵ ,
- uniform error estimates with rate of convergence $1/3$ are given in Theorem 3.8.

3. ANALYSIS OF THE SCHEME

3.1. Strong error estimates. In this section, we establish strong error estimates which allow to explain one of the observations associated with Figure 1: for arbitrary fixed ϵ , the rate of convergence is 1 with respect to Δt , with an error constant which depends on ϵ . This is obtained in Proposition 3.2.

Precisely, we state and prove two results, with different orders of convergence with respect to Δt , and also different rates for the dependence with respect to ϵ . Proposition 3.1 states that the error is at most of size $(\Delta t/\epsilon)^{1/2}$, whereas Proposition 3.2 states that the error is at most of size $\Delta t/\epsilon$. What is observed in Figure 1 is associated to Proposition 3.2.

Proposition 3.1. *There exists $C \in (0, \infty)$ such that for all $\epsilon \in (0, 1)$ and all Δt one has*

$$(8) \quad \sup_{0 \leq n \leq N} (\mathbb{E}[|x_n^\epsilon - x^\epsilon(t_n)|^2])^{\frac{1}{2}} \leq C \frac{\sqrt{\Delta t}}{\sqrt{\epsilon}}.$$

Proposition 3.2. *Assume that the mapping $f : \mathbb{R} \times \mathbb{T} \rightarrow \mathbb{R}$ is of class \mathcal{C}^2 , with bounded first and second order derivatives.*

There exists $C \in (0, \infty)$ such that for all $\epsilon \in (0, 1)$ and all Δt one has

$$(9) \quad \sup_{0 \leq n \leq N} (\mathbb{E}[|x_n^\epsilon - x^\epsilon(t_n)|^2])^{\frac{1}{2}} \leq C \frac{\Delta t}{\epsilon}.$$

The proofs of Propositions 3.1 and 3.2 are postponed to Section A.

Remark 3.3. *The values of the error constants C appearing in Propositions 3.1 and 3.2 depend on the final time $T = 1$, on the initial values x_0 and y_0 , and on the mapping f . To simplify the notation the dependence is omitted.*

Note that improving the rate of convergence in Proposition 3.2 compared with Proposition 3.1 requires stronger regularity conditions on the mapping f (Proposition 3.1 only requires the global Lipschitz continuity property). The techniques of the proof for Propositions 3.1 and 3.2 are similar to the proofs of strong error estimates for the Euler–Maruyama scheme for stochastic differential equations driven by additive noise, with rates 1/2 for globally Lipschitz continuous drift and 1 for drifts of class \mathcal{C}^2 .

3.2. Asymptotic preserving property. In this section, we study the behavior of the numerical scheme (7) when $\epsilon \rightarrow 0$ for arbitrary fixed time-step size $\Delta t > 0$. It is proved that, for all $n \in \{0, \dots, N\}$, x_n^ϵ converges in distribution to X_n given by the limiting scheme (10) when $\epsilon \rightarrow 0$, see Lemma 3.5 for a precise statement. It is then shown that the limiting scheme (10) is consistent with the averaged equation, see Lemma 3.7. Combining the result shows that the scheme (7) is asymptotic preserving.

The proofs of the results stated in this section are postponed to Section B.

Let us first provide a key technical result which justifies the convergence in distribution as $\epsilon \rightarrow 0$.

Lemma 3.4. *Let $\gamma \sim \mathcal{N}(0, 1)$ be a standard real-valued Gaussian random variable.*

For all $\epsilon > 0$, let $\gamma^\epsilon = \frac{\gamma}{\sqrt{\epsilon}}$.

Then $\hat{\gamma}^\epsilon$ converges in distribution to the uniform distribution on \mathbb{T} .

Let $(U_n)_{n \geq 1}$ be a sequence of independent random variables which are uniformly distributed on the torus \mathbb{T} . The result of Lemma 3.4 suggests to introduce the following scheme: for all $n \in \{0, \dots, N - 1\}$, set

$$(10) \quad \begin{cases} X_{n+1} = X_n + \Delta t f(X_n, Y_{n+1}), \\ Y_{n+1} = Y_n + U_{n+1}, \end{cases}$$

with initial values $X_0 = x_0$ and $Y_0 = y_0$. Note that the scheme (10) is constructed by replacing the independent random variables $\delta\beta_{n+1}/\sqrt{\epsilon}$ appearing in (7) by the random variables U_{n+1} for all $n \in \{0, \dots, N - 1\}$.

The scheme (10) can be considered as a limiting scheme when $\epsilon \rightarrow 0$, owing to the following result.

Lemma 3.5. *For any $N \in \mathbb{N}^*$ (with associated time-step size $\Delta t = T/N$), when $\epsilon \rightarrow 0$, one has the following convergence in distribution result in $(\mathbb{R} \times \mathbb{T})^{N+1}$:*

$$((x_n^\epsilon, y_n^\epsilon))_{0 \leq n \leq N} \xrightarrow[\epsilon \rightarrow 0]{} ((X_n, Y_n))_{0 \leq n \leq N}.$$

The expression of the limiting scheme (10) is not convenient. Indeed, note that $(X_n)_{0 \leq n \leq N}$ is not a Markov process, since X_{n+1} is given as a function of X_n and of the hidden state Y_{n+1} . It is possible to propose an alternative Markov representation for the limiting scheme: for all $n \in \{0, \dots, N-1\}$, set

$$(11) \quad \tilde{X}_{n+1} = \tilde{X}_n + \Delta t f(\tilde{X}_n, U_{n+1}),$$

with initial value $\tilde{X}_0 = X_0 = x_0$.

Lemma 3.6 shows that X_n and \tilde{X}_n are equal in distribution at all times $n \in \{0, \dots, N\}$.

Lemma 3.6. *Let $N \in \mathbb{N}^*$ (with associated time-step size $\Delta t = T/N$), and let $(X_n)_{0 \leq n \leq N}$ and $(\tilde{X}_n)_{0 \leq n \leq N}$ be defined by (10) and (11) respectively.*

For all $n \in \{0, \dots, N\}$, one has the equality in distribution

$$\tilde{X}_n = X_n.$$

The solutions X_n and \tilde{X}_n to limiting schemes (10) and (11) are random variables, whereas the solution $X(t_n)$ to the averaged equation (5) is deterministic. Lemma 3.7 states that \tilde{X}_n converges in the mean-square sense to $X(t_n)$ when $\Delta t = T/N \rightarrow 0$, meaning that the scheme (7) is asymptotic preserving.

Lemma 3.7. *Let $(X(t))_{t \in [0, T]}$ be given by the averaged ordinary differential equation (5) and let $(\tilde{X}_n)_{0 \leq n \leq N}$ be given by the limiting scheme (11) with time-step size $\Delta t = T/N$. There exists $C \in (0, \infty)$ such that for all $N \in \mathbb{N}^*$ one has*

$$(12) \quad \sup_{0 \leq n \leq N} (\mathbb{E}[|X_n - X(t_n)|^2])^{\frac{1}{2}} = \sup_{0 \leq n \leq N} (\mathbb{E}[|\tilde{X}_n - X(t_n)|^2])^{\frac{1}{2}} \leq C\sqrt{\Delta t}.$$

Note that one can replace \tilde{X}_n by X_n in (12), since for all $n \in \{0, \dots, N\}$ the random variables X_n and \tilde{X}_n are equal in distribution owing to Lemma 3.6.

The results above show that the scheme is asymptotic preserving, meaning that one has a commutative diagram when the parameter ϵ and the time-step size $\Delta t = T/N$ go to 0:

$$(13) \quad \begin{array}{ccc} x_N^\epsilon & \xrightarrow{\Delta t \rightarrow 0} & x^\epsilon(T) \\ \downarrow \epsilon \rightarrow 0 & & \downarrow \epsilon \rightarrow 0 \\ X_N & \xrightarrow{\Delta t \rightarrow 0} & X(T), \end{array}$$

where for simplicity the property is stated only at the final time $T = N\Delta t$. The convergence refers to convergence in distribution when considering the convergence of x_N^ϵ to X_N when $\epsilon \rightarrow 0$.

3.3. Uniform strong error estimates. Propositions 3.1 and 3.2 provide upper bounds on the error $(\mathbb{E}[|x_n^\epsilon - x^\epsilon(T)|^2])^{\frac{1}{2}}$, one obtains order 1/2 and 1 respectively with respect to Δt but the error bound is not uniform with respect to ϵ . This can be observed in Figure 1, which suggests that uniform error estimates with respect to ϵ can be obtained, with order of convergence 1/2 with respect to Δt . The main result of this section, Theorem 3.8, provides such uniform error estimates, but with order of convergence 1/3 with respect to Δt (in the setting of Proposition 3.2). In the setting of Proposition 3.1 one would obtain order of convergence 1/4 with respect to Δt , see Remark C.3.

Theorem 3.8. *Assume that the mapping $f : \mathbb{R} \times \mathbb{T} \rightarrow \mathbb{R}$ is of class C^2 , with bounded first and second order derivatives.*

There exists $C \in (0, \infty)$ such that one has

$$(14) \quad \sup_{\epsilon \in (0, 1)} \sup_{0 \leq n \leq N} (\mathbb{E}[|x_n^\epsilon - x^\epsilon(t_n)|^2])^{\frac{1}{2}} \leq C\Delta t^{\frac{1}{3}}.$$

The proof of Theorem 3.8 is postponed to Section C. It is inspired by the commutative diagram (13): using the triangle inequality, for all $\epsilon \in (0, 1)$, one has

$$\begin{aligned} & \sup_{0 \leq n \leq N} (\mathbb{E}[|x_n^\epsilon - x^\epsilon(t_n)|^2])^{\frac{1}{2}} \\ & \leq \sup_{0 \leq n \leq N} (\mathbb{E}[|x_n^\epsilon - X_n|^2])^{\frac{1}{2}} + \sup_{0 \leq n \leq N} (\mathbb{E}[|X_n - X(t_n)|^2])^{\frac{1}{2}} + \sup_{0 \leq n \leq N} (\mathbb{E}[|X(t_n) - x^\epsilon(t_n)|^2])^{\frac{1}{2}}. \end{aligned}$$

The result given in Theorem 3.8 is consistent with the observations drawn from the numerical experiments in Figure 1, however the observed order of convergence is 1/2 instead of 1/3. Whether the rate of convergence 1/3 obtained in Theorem 3.8 is optimal or whether it is possible to prove higher order of convergence, is left for future investigations.

4. STUDY OF STOCHASTIC VARIANTS

In this section, we consider variants of the system (4) and of the scheme (7), where the slow component x^ϵ is not solution to an ordinary differential equation but to a stochastic differential equation driven by a Wiener process $(W(t))_{t \geq 0}$, which is independent of the Wiener process $(\beta(t))_{t \geq 0}$.

From the numerical experiments, conjectures that would generalize the results stated in Section 3 are proposed.

Similar to (1), the increments of the Wiener process $(W(t))_{t \geq 0}$ are denoted by

$$\delta W_{n+1} = W(t_{n+1}) - W(t_n), \quad \forall n \in \{0, \dots, N\}.$$

4.1. Additive noise perturbation. Let us first consider the case where the slow equation is a stochastic differential equation driven by additive noise:

$$(15) \quad \begin{cases} dx^\epsilon(t) = f(x^\epsilon(t), y^\epsilon(t)) dt + dW(t), & t \geq 0, \\ dy^\epsilon(t) = \frac{d\beta(t)}{\sqrt{\epsilon}}, & t \geq 0, \\ x^\epsilon(0) = x_0, \quad y^\epsilon(0) = y_0. \end{cases}$$

Let us mention that the averaging principle also holds for the class of systems (15): when $\epsilon \rightarrow 0$, the slow component x^ϵ converges to the solution X to the stochastic differential equation driven by additive noise

$$(16) \quad \begin{cases} dX(t) = F(X(t)) dt + dW(t), & t \geq 0, \\ X(0) = x_0, \end{cases}$$

where the drift F is defined by (6).

For the discretization of the system (15), we propose the following numerical scheme, as a variant of (7): given the time-step size $\Delta t = T/N$, for all $n \in \{0, \dots, N-1\}$ set

$$(17) \quad \begin{cases} x_{n+1}^\epsilon = x_n^\epsilon + \Delta t f(x_n^\epsilon, y_{n+1}^\epsilon) + \delta W_{n+1}, \\ y_{n+1}^\epsilon = y_n^\epsilon + \frac{\delta \beta_{n+1}}{\sqrt{\epsilon}}, \end{cases}$$

with the initial values $x_0^\epsilon = x_0$ and $y_0^\epsilon = y_0$.

Adapting the arguments from Section 3.2, one may show that x_N^ϵ converges in distribution when $\epsilon \rightarrow 0$ to X_N given by the limiting scheme

$$\begin{cases} X_{n+1} = X_n + \Delta t f(X_n, Y_{n+1}) + \delta W_{n+1} \\ Y_{n+1} = Y_n + U_{n+1} \end{cases}$$

and to \tilde{X}_n given by the alternative Markov representation

$$\tilde{X}_{n+1} = \tilde{X}_n + \Delta t f(\tilde{X}_n, U_{n+1}) + \delta W_{n+1},$$

with the initial values $X_0 = \tilde{X}_0 = x_0$ and $Y_0 = y_0$, where $(U_n)_{n \geq 1}$ is a sequence of independent random variables which are uniformly distributed in the torus \mathbb{T} . The random variables X_N and \tilde{X}_N are equal in distribution.

Note that one obtains an asymptotic preserving property for the scheme (17), i.e. the commutative diagram (13) is also valid for this variant.

Let us present numerical experiments which illustrate the behavior of the strong error

$$\left(\mathbb{E}[|x_N^\epsilon - x^\epsilon(t_N)|^2] \right)^{\frac{1}{2}}$$

when $\Delta t = T/N \rightarrow 0$, depending on the parameter ϵ .

The parameters of the numerical experiments are the same as in Section 2.2.

Figure 2 displays how the strong error depends on the time-step size Δt as $\Delta t \rightarrow 0$, for different values of the time-scale separation parameter ϵ : on the left one has $\epsilon \in \{10^{-1}, 10^{-2}, 10^{-3}, 10^{-4}\}$ and on the right one has $\epsilon \in \{8 \cdot 10^{-3}, 4 \cdot 10^{-3}, 2 \cdot 10^{-3}, 10^{-3}\}$.

Three reference lines with slopes 1/3, 1/2 and 1 are indicated.

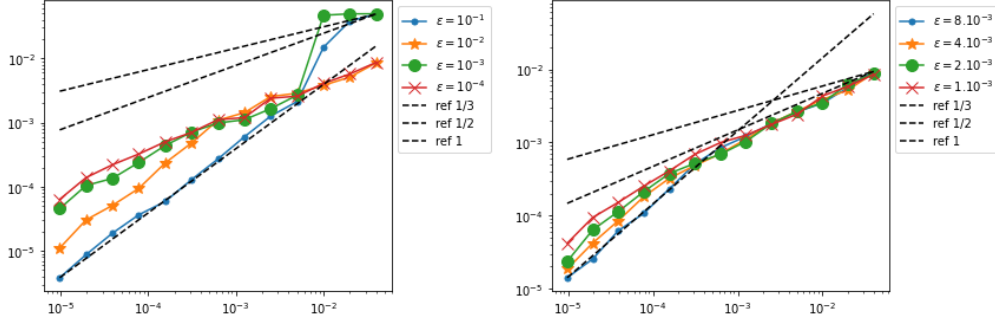


FIGURE 2. Behavior of the strong error for the scheme (17) applied to (15) as a function of the time-step size Δt for different values of ϵ . Left: $\epsilon \in \{10^{-1}, 10^{-2}, 10^{-3}, 10^{-4}\}$. Right: $\epsilon \in \{8 \cdot 10^{-3}, 4 \cdot 10^{-3}, 2 \cdot 10^{-3}, 10^{-3}\}$.

The observations are the same as the ones from those drawn from Figure 1 in Section 2.2: at fixed ϵ the order of convergence with respect to Δt is equal to 1 and one has error estimates which are uniform with respect to ϵ with order of convergence 1/2 with respect to Δt . We conjecture that the results stated in Section 3 can be generalized for systems (15) and the scheme (17) when the slow component is driven by additive noise.

4.2. Multiplicative noise perturbation. Let us now consider another variant of the system (4), where the slow component x^ϵ is solution to a stochastic differential equation (with Itô interpretation of the noise):

$$(18) \quad \begin{cases} dx^\epsilon(t) = f(x^\epsilon(t), y^\epsilon(t))dt + g(x^\epsilon, \Delta t(t))dW(t), & t \geq 0, \\ dy^\epsilon(t) = \frac{d\beta(t)}{\sqrt{\epsilon}}, & t \geq 0, \\ x^\epsilon(0) = x_0, & y^\epsilon(0) = y_0. \end{cases}$$

Contrary to the system (15) considered in Section 4.1, the noise is not additive, it depends on the diffusion mapping $g : \mathbb{R} \times \mathbb{R}$. Note that it is important to assume that $g(x)$ only depends on the slow component x and does not depend on the fast component y . It is assumed that g is bounded and Lipschitz continuous.

The averaging principle also holds for the class of systems (18): when $\epsilon \rightarrow 0$, the slow component x^ϵ converges to the solution X to the stochastic differential equation driven by multiplicative noise

$$(19) \quad \begin{cases} dX(t) = F(X(t)) dt + g(X(t)) dW(t), & t \geq 0, \\ X(0) = x_0, \end{cases}$$

where the drift F is defined by (6). Note that the diffusion coefficient g in (19) is the same as in (18).

For the discretization of the system (18), we first propose the following numerical scheme, where the slow component is approximated using the Euler–Maruyama method: given the time-step size $\Delta t = T/N$, for all $n \in \{0, \dots, N-1\}$ set

$$(20) \quad \begin{cases} x_{n+1}^\epsilon = x_n^\epsilon + \Delta t f(x_n^\epsilon, y_{n+1}^\epsilon) + g(x_n^\epsilon) \delta W_{n+1} \\ y_{n+1}^\epsilon = y_n^\epsilon + \frac{\delta \beta_{n+1}}{\sqrt{\epsilon}}, \end{cases}$$

with the initial values $x_0^\epsilon = x_0$ and $y_0^\epsilon = y_0$.

Adapting the arguments from Section 3.2, one may show that x_N^ϵ converges in distribution when $\epsilon \rightarrow 0$ to X_N given by the limiting scheme

$$\begin{cases} X_{n+1} = X_n + \Delta t f(X_n, Y_{n+1}) + g(X_n) \delta W_{n+1} \\ Y_{n+1} = Y_n + U_{n+1} \end{cases}$$

and to \tilde{X}_n given by the alternative Markov representation

$$\tilde{X}_{n+1} = \tilde{X}_n + \Delta t f(\tilde{X}_n, U_{n+1}) + g(\tilde{X}_n) \delta W_{n+1},$$

with initial values $X_0 = \tilde{X}_0 = x_0$ and $Y_0 = y_0$, where $(U_n)_{n \geq 1}$ is a sequence of independent random variables which are uniformly distributed in the torus \mathbb{T} .

Note that one obtains an asymptotic preserving property for the scheme (17), i.e. the commutative diagram (13) is also valid for this variant.

Let us present numerical experiments which illustrate the behavior of the strong error

$$(\mathbb{E}[|x_N^\epsilon - x^\epsilon(t_N)|^2])^{\frac{1}{2}}$$

when $\Delta t = T/N \rightarrow 0$, depending on the parameter ϵ .

The parameters of the numerical experiments are the same as in Section 2.2, and the diffusion coefficient g is given by

$$g(x) = \frac{1}{4} \cos(2\pi x) + \frac{1}{2}.$$

Figure 3 displays how the strong error depends on the time-step size Δt as $\Delta t \rightarrow 0$, for different values of the time-scale separation parameter ϵ : on the left one has $\epsilon \in \{10^{-1}, 10^{-2}, 10^{-3}, 10^{-4}\}$ and on the right one has $\epsilon \in \{8 \cdot 10^{-3}, 4 \cdot 10^{-3}, 2 \cdot 10^{-3}, 10^{-3}\}$.

Three reference lines with slopes 1/3, 1/2 and 1 are indicated. The behavior observed in Figure 3 is quite

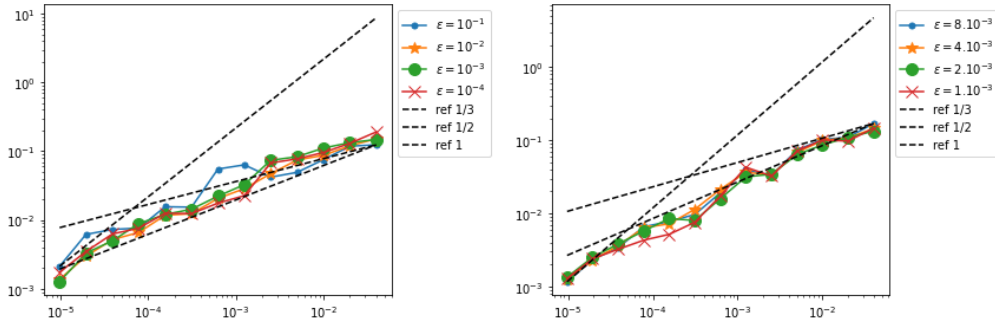


FIGURE 3. Behavior of the strong error for the scheme (20) applied to (18) as a function of the time-step size Δt for different values of ϵ . Left: $\epsilon \in \{10^{-1}, 10^{-2}, 10^{-3}, 10^{-4}\}$. Right: $\epsilon \in \{8 \cdot 10^{-3}, 4 \cdot 10^{-3}, 2 \cdot 10^{-3}, 10^{-3}\}$.

different from what is seen in Figures 1 and 2. This is due to the fact that for fixed ϵ the strong order of convergence is equal to 1/2 instead of 1, i.e. only Proposition 3.1 could be obtained for the scheme (20) associated to the system (18). However, it is still observed in Figure 3 that error estimates which are uniform with respect to ϵ can be obtained, we thus conjecture that a variant of Theorem 3.8 holds, possibly with order of convergence 1/4 (see Remark C.3).

The order of convergence 1/2 with respect to Δt is due to the application of the Euler–Maruyama scheme for the discretization of the slow component x^ϵ . It is natural to study whether applying an higher order scheme would have a substantial impact on the behavior of the error.

Let us thus consider the following numerical scheme, where the slow component is approximated using the Milstein method: given the time-step size $\Delta t = T/N$, for all $n \in \{0, \dots, N-1\}$ set

$$(21) \quad \begin{cases} x_{n+1}^\epsilon = x_n^\epsilon + \Delta t f(x_n^\epsilon, y_{n+1}^\epsilon) + g(x_n^\epsilon) \delta W_{n+1} + \frac{1}{2} (g'g)(x_n^\epsilon) (\delta W_{n+1}^2 - \Delta t) \\ y_{n+1}^\epsilon = y_n^\epsilon + \frac{\delta \beta_{n+1}}{\sqrt{\epsilon}}, \end{cases}$$

with the initial values $x_0^\epsilon = x_0$ and $y_0^\epsilon = y_0$.

Again, one may show that x_N^ϵ converges in distribution when $\epsilon \rightarrow 0$ to X_N given by the limiting scheme

$$\begin{cases} X_{n+1} = X_n + \Delta t f(X_n, Y_{n+1}) + g(X_n) \delta W_{n+1} + \frac{1}{2} (g'g)(X_n) (\delta W_{n+1}^2 - \Delta t) \\ Y_{n+1} = Y_n + U_{n+1} \end{cases}$$

and to \tilde{X}_n given by the alternative Markov representation

$$\tilde{X}_{n+1} = \tilde{X}_n + \Delta t f(\tilde{X}_n, U_{n+1}) + g(\tilde{X}_n) \delta W_{n+1} + \frac{1}{2} (g'g)(\tilde{X}_n) (\delta W_{n+1}^2 - \Delta t),$$

with initial values $X_0 = \tilde{X}_0 = x_0$ and $Y_0 = y_0$, where $(U_n)_{n \geq 1}$ is a sequence of independent random variables which are uniformly distributed in the torus \mathbb{T} .

Again one obtains an asymptotic preserving property for the scheme (17), i.e. the commutative diagram (13) is also valid for this variant.

Figure 4 displays how the strong error depends on the time-step size Δt as $\Delta t \rightarrow 0$, for different values of the time-scale separation parameter ϵ : on the left one has $\epsilon \in \{10^{-1}, 10^{-2}, 10^{-3}, 10^{-4}\}$ and on the right one has $\epsilon \in \{8 \cdot 10^{-3}, 4 \cdot 10^{-3}, 2 \cdot 10^{-3}, 10^{-3}\}$.

All the parameters are the same as in the numerical experiment presented above in Figure 3. Three reference lines with slopes 1/3, 1/2 and 1 are indicated. The sample size for the Monte Carlo experiment is

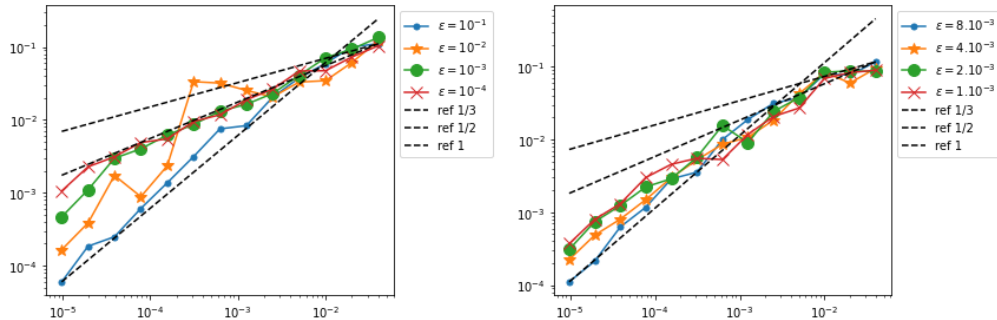


FIGURE 4. Behavior of the strong error for the scheme (21) applied to (18) as a function of the time-step size Δt for different values of ϵ . Left: $\epsilon \in \{10^{-1}, 10^{-2}, 10^{-3}, 10^{-4}\}$. Right: $\epsilon \in \{8 \cdot 10^{-3}, 4 \cdot 10^{-3}, 2 \cdot 10^{-3}, 10^{-3}\}$.

not large so the statistical error is visible in Figure 4.

Comparing Figures 3 and 4 clearly shows that using the Milstein method in the scheme (21) instead of the Euler–Maruyama method in the scheme (20) leads to improved results, which are comparable to those in Figures 1 and 2. Therefore the choice of the integrator is important. We conjecture that the results from Section 3, especially the uniform strong error estimates given by Theorem 3.8, can be generalized for the system (18) when the scheme (21) is applied. Proving such results is out of the scope of this contribution and may be considered in future works.

4.3. A stochastic system. Finally, we consider another variant of the system, where the slow component is solution to a stochastic differential equation (with Itô interpretation of the noise) with no drift term and with diffusion coefficient given by the vector field f :

$$(22) \quad \begin{cases} dx^\epsilon(t) = f(x^\epsilon(t), y^\epsilon(t))dW(t), & t \geq 0, \\ dy^\epsilon(t) = \frac{d\beta(t)}{\sqrt{\epsilon}}, & t \geq 0, \\ x^\epsilon(0) = x_0, & y^\epsilon(0) = y_0. \end{cases}$$

Contrary to the situation described in Section 4.2, the diffusion coefficient f depends on both variables x and y .

For the discretization of the system (22), we propose the following numerical scheme, where the slow component is approximated using the Euler–Maruyama method (like in the scheme (20)): given the time-step size $\Delta t = T/N$, for all $n \in \{0, \dots, N-1\}$ set

$$(23) \quad \begin{cases} x_{n+1}^\epsilon = x_n^\epsilon + g(x_n^\epsilon, m_{n+1}^\epsilon)\delta W_{n+1} \\ m_{n+1}^\epsilon = m_n^\epsilon + \frac{\delta\beta_{n+1}}{\sqrt{\epsilon}} \end{cases}$$

with the initial values $x_0^\epsilon = x_0$ and $y_0^\epsilon = y_0$ and with Wiener increments $\delta W_{n+1} = W(t_{n+1}) - W(t_n)$.

Let us present numerical experiments which illustrate the behavior of the strong error

$$\left(\mathbb{E}[|x_N^\epsilon - x^\epsilon(t_N)|^2]\right)^{\frac{1}{2}}$$

when $\Delta t = T/N \rightarrow 0$, depending on the parameter ϵ .

The parameters of the numerical experiments are the same as in Section 2.2.

Figure 5 displays how the strong error depends on the time-step size Δt as $\Delta t \rightarrow 0$, for different values of the time-scale separation parameter ϵ : on the left one has $\epsilon \in \{10^{-1}, 10^{-2}, 10^{-3}, 10^{-4}\}$ and on the right one has $\epsilon \in \{8 \cdot 10^{-3}, 4 \cdot 10^{-3}, 2 \cdot 10^{-3}, 10^{-3}\}$.

Three reference lines with slopes $1/3$, $1/2$ and 1 are indicated. The behavior observed in Figure 5 is quite

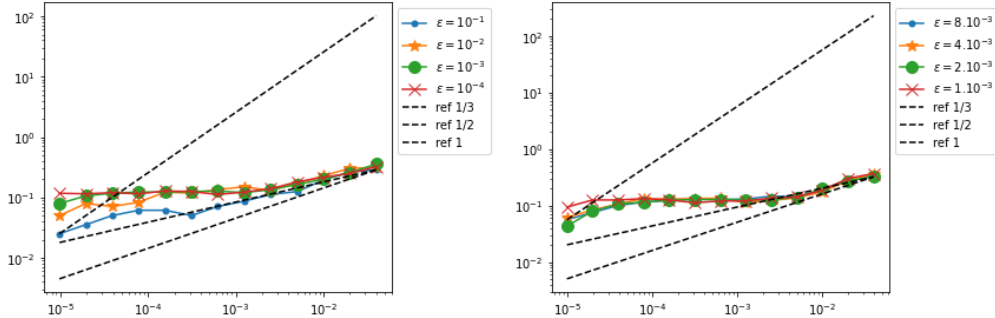


FIGURE 5. Behavior of the strong error for the scheme (23) applied to (22) as a function of the time-step size Δt for different values of ϵ . Left: $\epsilon \in \{10^{-1}, 10^{-2}, 10^{-3}, 10^{-4}\}$. Right: $\epsilon \in \{8 \cdot 10^{-3}, 4 \cdot 10^{-3}, 2 \cdot 10^{-3}, 10^{-3}\}$.

different from what has been observed in all the other numerical experiments: one does not observe good convergence of the scheme when Δt goes to 0. This result is not due to the choice of the scheme, but to the fact that the averaging principle does not hold in the strong sense, but only in the weak sense.

The correct version of the averaging principle is the following, see for instance the recent work [4]: with the mapping $\sigma : \mathbb{R} \rightarrow \mathbb{R}$ defined by

$$\sigma(x) = \left(\int_{\mathbb{T}} f(x, m)^2 dm\right)^{\frac{1}{2}}, \quad \forall x \in \mathbb{R},$$

x^ϵ converges in distribution to the solution X to the stochastic differential equation

$$(24) \quad \begin{cases} dX(t) = \sigma(X(t)) dW(t), & \forall t \geq 0, \\ X(0) = x_0. \end{cases}$$

This implies that for any bounded and continuous function $\varphi : \mathbb{R} \rightarrow \mathbb{R}$, one has

$$\mathbb{E}[\varphi(x^\epsilon(T))] \xrightarrow{\epsilon \rightarrow 0} \mathbb{E}[\varphi(X(T))].$$

However, the convergence is not strong, meaning that $\mathbb{E}[|x^\epsilon(T) - X(T)|^2]$ does not converge to 0 when $\epsilon \rightarrow 0$.

Since the averaging principle does not hold in the strong sense, it is not surprising to observe that the scheme (23) is not asymptotic preserving in the strong sense. This justifies the observations following Figure 5.

The averaging principle holds in the weak sense, i.e. in the sense of convergence in distribution. It is thus natural to investigate the behavior of the weak error

$$|\mathbb{E}[\varphi(x_N^\epsilon)] - \mathbb{E}[\varphi(x^\epsilon(t_N))]|$$

when $\Delta t = T/N \rightarrow 0$, depending on the parameter ϵ , where $\varphi : \mathbb{R} \rightarrow \mathbb{R}$ is a function of class \mathcal{C}^∞ with bounded derivatives. In the numerical experiments, we choose

$$\varphi(x) = \sin^3(2\pi x) - \cos(3\pi x).$$

As in other numerical experiments, the initial values are $x_0 = 0$ and $y_0 = 0$. The final time is $T = 1$. The vector field f is given by

$$f(x, y) = \cos(4\pi x) \cos^2(2\pi(x + y)).$$

For the numerical experiments on the weak error, the time-step size Δt takes values in

$$\{2^{-j} \cdot 10^{-2}; -2 \leq j \leq 3\} = \left\{ \frac{1}{25}, \frac{1}{50}, \frac{1}{100}, \dots, \frac{1}{400}, \frac{1}{800} \right\}.$$

The weak error is estimated by a Monte Carlo average over 10^4 independent realizations. The reference solution is computed using the scheme (7) with time-step size $\Delta t_{\text{ref}} = 2^{-4} \cdot 10^{-2}$.

Figure 6 displays how the weak error depends on the time-step size Δt as $\Delta t \rightarrow 0$, for different values of the time-scale separation parameter ϵ : on the left one has $\epsilon \in \{10^{-1}, 10^{-2}, 10^{-3}, 10^{-4}\}$ and on the right one has $\epsilon \in \{8 \cdot 10^{-3}, 4 \cdot 10^{-3}, 2 \cdot 10^{-3}, 10^{-3}\}$.

Two reference lines with slopes $1/2$ and 1 are indicated.

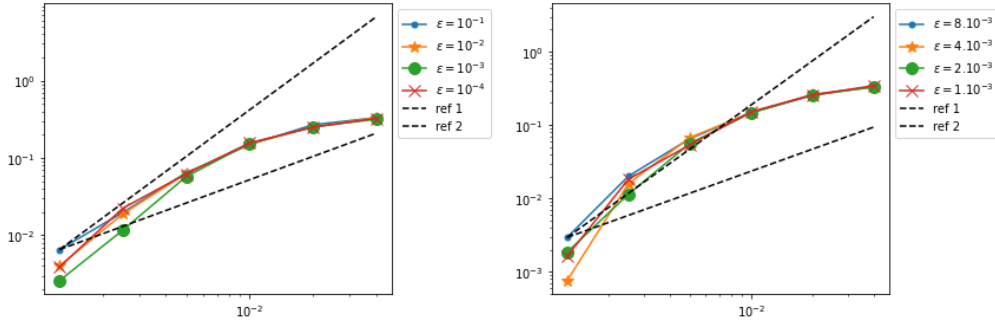


FIGURE 6. Behavior of the weak error for the scheme (23) applied to (22) as a function of the time-step size Δt for different values of ϵ . Left: $\epsilon \in \{10^{-1}, 10^{-2}, 10^{-3}, 10^{-4}\}$. Right: $\epsilon \in \{8 \cdot 10^{-3}, 4 \cdot 10^{-3}, 2 \cdot 10^{-3}, 10^{-3}\}$.

Compared with Figure 5, better convergence properties are observed in Figure 6 for the weak error. It would be possible to establish that the scheme is asymptotic preserving in the weak sense, i.e. in the sense of convergence in distribution. In addition, it would be interesting to investigate whether one could prove weak error estimates which are uniform with respect to ϵ . These questions are out of the scope of this contribution and may be investigated in future works.

5. STUDY OF AN HIGHER ORDER METHOD

In this section, we consider the system (4) and we investigate whether one is able to construct higher order methods. The presentation is heuristic and less rigorous than in Sections 2 and 3.

The averaging principle states that x^ϵ converges when ϵ goes to 0 to the solution to the ordinary differential equation (5). For the scheme (7), the limiting scheme (10) (or equivalently (11)) obtained when ϵ goes to 0 is an order one method for (5): indeed, using the formulation (11), one has

$$\tilde{X}_{n+1} = \tilde{X}_n + \Delta t F_n$$

where $F_n = f(\tilde{X}_n, U_{n+1})$ satisfies

$$\mathbb{E}[F_n | \tilde{X}_n] = F(\tilde{X}_n).$$

Having a higher order method associated to the ordinary differential equation (5) would require to have an expansion

$$\tilde{X}_{n+1} = \tilde{X}_n + \Delta t F_n^{(1)} + \Delta t^2 F_n^{(2)} + O(\Delta t^3),$$

where $F_n^{(1)}$ and $F_n^{(2)}$ satisfy

$$\mathbb{E}[F_n^{(1)} | \tilde{X}_n] = F(\tilde{X}_n), \quad \mathbb{E}[F_n^{(2)} | \tilde{X}_n] = \frac{1}{2} F'(\tilde{X}_n) F(\tilde{X}_n).$$

It is possible to construct a method which satisfies these properties using a variant of Heun's integrator. However, some care is needed, and the numerical experiments below show that the proposed method seems to be more efficient than the scheme (7) only in the weak sense, not in the strong sense.

To define the new method, some additional notation is needed. For all $n \in \{0, \dots, N-1\}$, intermediate grid points are denoted by $t_{n+1/2} = (t_n + t_{n+1})/2 = t_n + \Delta t/2$. Associated Brownian increments on the intervals $[t_n, t_{n+1/2}]$ and $[t_{n+1/2}, t_{n+1}]$ are respectively $\beta(t_{n+1/2}) - \beta(t_n)$ and $\beta(t_{n+1}) - \beta(t_{n+1/2})$. Recall that these Gaussian random variables are independent due to the properties of Brownian motion.

We propose the following numerical scheme: given the time-step size $\Delta t = T/N$, for all $n \in \{0, \dots, N-1\}$ set

$$(25) \quad \begin{cases} x_{n+1/2}^\epsilon = x_n^\epsilon + \frac{\Delta t}{2} f(x_n^\epsilon, y_{n+1/2}^\epsilon), \\ y_{n+1/2}^\epsilon = y_n^\epsilon + \frac{\beta(t_{n+1/2}) - \beta(t_n)}{\sqrt{\epsilon}}, \\ x_{n+1}^\epsilon = x_n^\epsilon + \Delta t f(x_{n+1/2}^\epsilon, y_{n+1}^\epsilon), \\ y_{n+1}^\epsilon = y_{n+1/2}^\epsilon + \frac{\beta(t_{n+1}) - \beta(t_{n+1/2})}{\sqrt{\epsilon}}, \end{cases}$$

with the initial values $x_0^\epsilon = x_0$ and $y_0^\epsilon = y_0$.

Let us briefly justify why the proposed scheme (25) is a good candidate for higher order approximation. Using the same arguments as in Section 3.2, one may show that x_n^ϵ converges in distribution to X_n given by the limiting scheme

$$(26) \quad \begin{cases} X_{n+1/2} = X_n + \frac{\Delta t}{2} f(X_n, Y_{n+1/2}), \\ Y_{n+1/2} = M_n + U_{n+1/2} \\ X_{n+1} = X_n + \Delta t f(X_{n+1/2}, Y_{n+1}), \\ Y_{n+1} = M_{n+1/2} + U_{n+1}, \end{cases}$$

with the initial values $X_0 = x_0$ and $Y_0 = y_0$, where $(U_n)_{n \geq 1}$ and $(U_{n+1/2})_{n \geq 0}$ are two independent sequences of independent random variables which are uniformly distributed in the torus \mathbb{T} . An alternative Markov representation of the limiting scheme, such that one has the equality in distribution $\tilde{X}_n = X_n$ for all $n \in \{0, \dots, N\}$, is given by

$$(27) \quad \begin{cases} \tilde{X}_{n+1/2} = \tilde{X}_n + \frac{\Delta t}{2} f(\tilde{X}_n, U_{n+1/2}), \\ \tilde{X}_{n+1} = \tilde{X}_n + \Delta t f(\tilde{X}_{n+1/2}, U_{n+1/2} + U_{n+1}), \end{cases}$$

with the initial value $\tilde{X}_0 = x_0$.

Given the expression above, under appropriate regularity conditions a Taylor expansion yields the expression

$$\tilde{X}_{n+1} = \tilde{X}_n + \Delta t f(\tilde{X}_n, U_{n+\frac{1}{2}} + U_{n+1}) + \frac{\Delta t^2}{2} \partial_x f(\tilde{X}_n, U_{n+\frac{1}{2}} + U_{n+1}) f(\tilde{X}_n, U_{n+\frac{1}{2}}) + O(\Delta t^3).$$

It is then possible to check that the requirements mentioned above for having a higher order method are fulfilled: taking into account the independence of the random variables $U_{n+\frac{1}{2}}$ and U_{n+1} and using the tower property of conditional expectation, one obtains for the first order term $F_n^{(1)} = f(\tilde{X}_n, U_{n+\frac{1}{2}} + U_{n+1})$

$$\begin{aligned} \mathbb{E}[f(\tilde{X}_n, U_{n+\frac{1}{2}} + U_{n+1}) | \tilde{X}_n] &= \mathbb{E}[\mathbb{E}[f(\tilde{X}_n, U_{n+\frac{1}{2}} + U_{n+1}) | \tilde{X}_n, U_{n+\frac{1}{2}}] | \tilde{X}_n] \\ &= \mathbb{E}[F(\tilde{X}_n) | \tilde{X}_n] \\ &= F(\tilde{X}_n) \end{aligned}$$

and one obtains for the second order term $F_n^{(2)} = \frac{1}{2} \partial_x f(\tilde{X}_n, U_{n+\frac{1}{2}} + U_{n+1}) f(\tilde{X}_n^{0, \Delta t}, U_{n+\frac{1}{2}})$

$$\begin{aligned} \frac{1}{2} \mathbb{E}[\partial_x f(\tilde{X}_n, U_{n+\frac{1}{2}} + U_{n+1}) f(\tilde{X}_n, U_{n+\frac{1}{2}}) | \tilde{X}_n] &= \frac{1}{2} \mathbb{E}[\mathbb{E}[\partial_x f(\tilde{X}_n, U_{n+\frac{1}{2}} + U_{n+1}) f(\tilde{X}_n, U_{n+\frac{1}{2}}) | \tilde{X}_n, U_{n+\frac{1}{2}}] | \tilde{X}_n] \\ &= \frac{1}{2} \mathbb{E}[F'(\tilde{X}_n) f(\tilde{X}_n, U_{n+\frac{1}{2}}) | \tilde{X}_n] \\ &= \frac{1}{2} F'(\tilde{X}_n) F(\tilde{X}_n). \end{aligned}$$

Let us report first numerical experiments on the behavior of the strong error. The numerical parameters are the same as in Section 2.2.

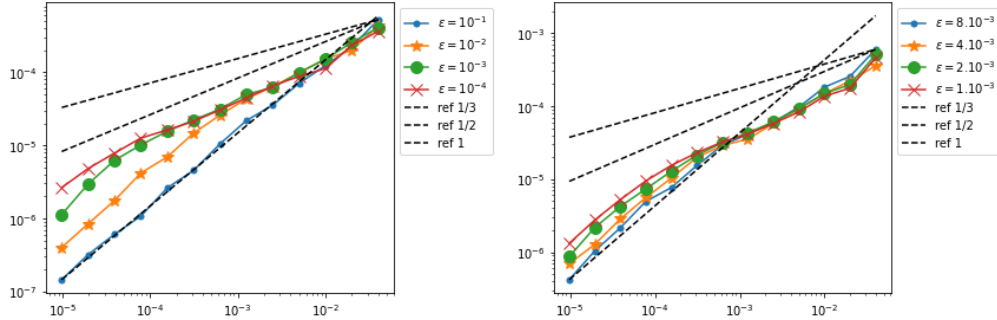


FIGURE 7. Behavior of the strong error for the scheme (25) applied to (5) as a function of the time-step size Δt for different values of ϵ . Left: $\epsilon \in \{10^{-1}, 10^{-2}, 10^{-3}, 10^{-4}\}$. Right: $\epsilon \in \{8.10^{-3}, 4.10^{-3}, 2.10^{-3}, 1.10^{-3}\}$.

The behavior observed in Figure 7 is the same as for the standard scheme (7) seen in Figure 1, in terms of orders of convergence. Therefore it does not seem that the proposed scheme (25) is a higher order method in the strong sense.

When one considers weak errors, the schemes (7) and (25) have different behaviors. This observation suggests that one may prove better results for the proposed scheme (25), this question is left for future works.

Figure 8 displays how the weak error

$$|\mathbb{E}[\varphi(x_N^\epsilon)] - \mathbb{E}[\varphi(x^\epsilon(t_N))]|$$

behaves when $\Delta t = T/N \rightarrow 0$, depending on the parameter ϵ . The function φ is given by

$$\varphi(x) = \sin^3(2\pi x) - \cos(3\pi x).$$

As in other numerical experiments, the initial values are $x_0 = 0$ and $y_0 = 0$. The final time is $T = 1$. The vector field f is given by

$$f(x, y) = \cos(4\pi x) \cos^2(2\pi(x + y)).$$

For the numerical experiments on the weak error, the time-step size Δt takes values in

$$\{2^{-j} \cdot 10^{-2}; -2 \leq j \leq 3\} = \left\{ \frac{1}{25}, \frac{1}{50}, \frac{1}{100}, \dots, \frac{1}{400}, \frac{1}{800} \right\}.$$

The parameter ϵ takes values in $\epsilon \in \{10^{-1}, 10^{-2}, 10^{-3}, 10^{-4}\}$.

The weak error is estimated by a Monte Carlo average over $4 \cdot 10^4$ independent realizations. The reference solution is computed using the scheme (7) with time-step size $\Delta t_{\text{ref}} = 2^{-4} \cdot 10^{-2}$.

Since the behavior of the weak error for the scheme (7) has not been reported earlier, this is presented in the left part of Figure 8, whereas the right part of Figure 8 presents the results for the scheme (25). One conjectures that for the standard scheme (4), one may prove weak error estimates which are uniform with respect to ϵ , with rate of convergence $1/2$ with respect to Δt . This behavior is consistent with what is observed for the strong error in Figure 1. For the scheme (25), at least in some regimes the weak order of convergence with respect to Δt seems to be equal to 1. To confirm this, refined numerical experiments and analysis may be needed.

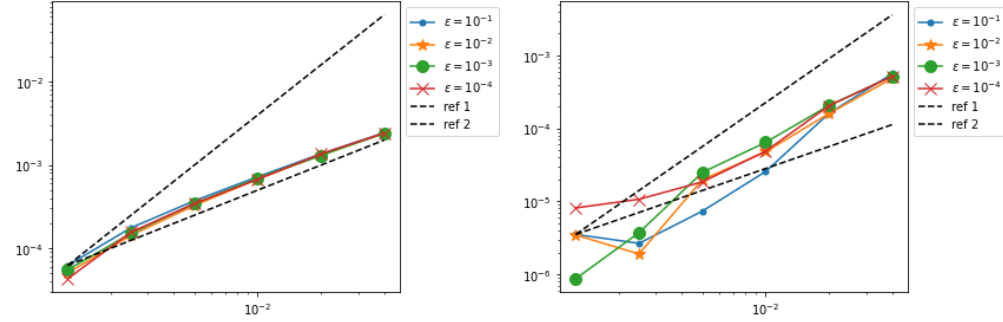


FIGURE 8. Behavior of the weak error for the scheme (7) (left) and for the scheme (25) (right) applied to (5) as a function of the time-step size Δt for different values of ϵ .

Remark 5.1. *Instead of the scheme (25), one may propose the simpler version*

$$\begin{cases} x_{n+\frac{1}{2}}^\epsilon = x_n^\epsilon + \frac{\Delta t}{2} f(x_n^\epsilon, y_{n+1}^\epsilon) \\ x_{n+1}^\epsilon = x_n^\epsilon + \Delta t f(x_n^\epsilon, y_{n+1}^\epsilon) \\ y_{n+1}^\epsilon = y_n^\epsilon + \frac{\beta(t_{n+1}) - \beta(t_n)}{\sqrt{\epsilon}}, \end{cases}$$

which does not require to introduce the Brownian increments on a refined grid as in (25). This method would not lead to higher-order of convergence, indeed one would obtain the limiting scheme

$$\begin{cases} X_{n+\frac{1}{2}} = X_n + \frac{\Delta t}{2} f(X_n, Y_{n+1}) \\ X_{n+1} = X_n + \Delta t f(X_n, Y_{n+1}) \\ Y_{n+1} = Y_n + U_n, \end{cases}$$

with its associated alternative Markov representation:

$$(28) \quad \begin{cases} \tilde{X}_{n+\frac{1}{2}} = \tilde{X}_n + \frac{\Delta t}{2} f(\tilde{X}_n, U_{n+1}) \\ \tilde{X}_{n+1} = \tilde{X}_n + \Delta t f(\tilde{X}_{n+\frac{1}{2}}, U_{n+1}). \end{cases}$$

Performing a Taylor expansion yields

$$\tilde{X}_{n+1} = \tilde{X}_n + \Delta t f(\tilde{X}_n, U_{n+1}) + \frac{\Delta t^2}{2} \partial_x f(\tilde{X}_n, U_{n+1}) f(\tilde{X}_n, U_{n+1}) + O(\Delta t^3),$$

where for the first and second order terms one has

$$\begin{aligned} \mathbb{E}[f(\tilde{X}_n, U_{n+\frac{1}{2}} + U_{n+1}) | \tilde{X}_n] &= F(\tilde{X}_n) \\ \mathbb{E}[\partial_x f(\tilde{X}_n, U_{n+1}) f(\tilde{X}_n, U_{n+1}) | \tilde{X}_n] &\neq F'(\tilde{X}_n) F(\tilde{X}_n). \end{aligned}$$

The choice of decomposing the Brownian increments on the subintervals $[t_n, t_{n+\frac{1}{2}}]$ and $[t_{n+\frac{1}{2}}, t_{n+1}]$ is fundamental for the good performance of the scheme (25). When considering the limit $\epsilon \rightarrow 0$, this yields two independent uniform random variables $U_{n+\frac{1}{2}}$ and U_{n+1} , and this allows to obtain the correct second order term.

ACKNOWLEDGMENTS

The author would like to thank the organizers of the BIT65 conference for giving the opportunity to present results on the topic of this article.

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] Ibrahim Almuslimani, Philippe Chartier, Mohammed Lemou, and Florian Méhats. Uniformly accurate schemes for drift-oscillatory stochastic differential equations. *Appl. Numer. Math.*, 181:468–482, 2022.
- [3] Charles-Edouard Bréhier. Asymptotic preserving schemes for SDEs driven by fractional Brownian motion in the averaging regime. *J. Math. Anal. Appl.*, 509(1):Paper No. 125940, 20, 2022.
- [4] Charles-Edouard Bréhier. The averaging principle for stochastic differential equations driven by a Wiener process revisited. *C. R. Math. Acad. Sci. Paris*, 360:265–273, 2022.
- [5] Charles-Edouard Bréhier. Uniform strong and weak error estimates for numerical schemes applied to multiscale SDEs in a Smoluchowski-Kramers diffusion approximation regime. *J. Comput. Dyn.*, 10(3):387–424, 2023.
- [6] Charles-Edouard Bréhier. Uniform error bounds for numerical schemes applied to multiscale SDEs in a Wong-Zakai diffusion approximation regime. *J. Comput. Dyn.*, 11(2):189–219, 2024.
- [7] Charles-Edouard Bréhier. Uniform weak error estimates for an asymptotic preserving scheme applied to a class of slow-fast parabolic semilinear SPDEs. *SMAI J. Comput. Math.*, 10:175–228, 2024.
- [8] Charles-Edouard Bréhier and Shmuel 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.
- [9] Weinan E, Di Liu, and Eric Vanden-Eijnden. Analysis of multiscale methods for stochastic differential equations. *Comm. Pure Appl. Math.*, 58(11):1544–1585, 2005.
- [10] Dror Givon, Ioannis G. Kevrekidis, and Raz Kupferman. Strong convergence of projective integration schemes for singularly perturbed stochastic differential systems. *Commun. Math. Sci.*, 4(4):707–729, 2006.
- [11] Shi Jin. Efficient asymptotic-preserving (AP) schemes for some multiscale kinetic equations. *SIAM J. Sci. Comput.*, 21(2):441–454, 1999.
- [12] Shi Jin. Asymptotic preserving (AP) schemes for multiscale kinetic and hyperbolic equations: a review. *Riv. Math. Univ. Parma (N.S.)*, 3(2):177–216, 2012.
- [13] Ioannis G. Kevrekidis, C. William Gear, James M. Hyman, Panagiotis G. Kevrekidis, Olof Runborg, and Constantinos Theodoropoulos. Equation-free, coarse-grained multiscale computation: enabling microscopic simulators to perform system-level analysis. *Commun. Math. Sci.*, 1(4):715–762, 2003.
- [14] Adrien Laurent. A uniformly accurate scheme for the numerical integration of penalized Langevin dynamics. *SIAM J. Sci. Comput.*, 44(5):A3217–A3243, 2022.
- [15] Frédéric Legoll, Tony Lelièvre, Keith Myerscough, and Giovanni Samaey. Parareal computation of stochastic differential equations with time-scale separation: a numerical convergence study. *Comput. Vis. Sci.*, 23(1-4):9, 2020.
- [16] Tiejun Li, Assyr Abdulle, and Weinan E. Effectiveness of implicit methods for stiff stochastic differential equations. *Commun. Comput. Phys.*, 3(2):295–307, 2008.
- [17] Grigorios A. Pavliotis and Andrew M. Stuart. *Multiscale methods*, volume 53 of *Texts in Applied Mathematics*. Springer, New York, 2008. Averaging and homogenization.
- [18] Hannes Vandecasteele, Przemyslaw Zieliński, and Giovanni Samaey. Efficiency of a micro-macro acceleration method for scale-separated stochastic differential equations. *Multiscale Model. Simul.*, 18(3):1272–1298, 2020.

STATEMENTS AND DECLARATIONS

Competing Interests. The author has no competing interests to declare that are relevant to the content of this article.

APPENDIX A. PROOF OF STRONG ERROR ESTIMATES

The objective of this section is to prove the strong error estimates from Propositions 3.1 and 3.2.

Note that for any parameter $\epsilon \in (0, 1)$ and any time-step size Δt , the solution of (4) at time $t = t_n = n\Delta t$ is given by

$$(29) \quad x^\epsilon(t_n) = x_0 + \int_0^{t_n} f(x^\epsilon(s), \frac{\beta(s)}{\sqrt{\epsilon}}) ds$$

whereas the solution of the scheme (7) at iteration n is given by

$$(30) \quad x_n^\epsilon = x_0 + \Delta t \sum_{k=0}^{n-1} f(x_k^\epsilon, \frac{\beta(t_{k+1})}{\sqrt{\epsilon}}).$$

A.1. Proof of Proposition 3.1. Combining the identities (29) and (30), the error $x^\epsilon(t_n) - x_n^\epsilon$ is decomposed as follows: for all $n \in \{0, \dots, N\}$, one has

$$\begin{aligned} x^\epsilon(t_n) - x_n^\epsilon &= \sum_{k=0}^{n-1} \int_{t_k}^{t_{k+1}} \left[f(x^\epsilon(s), \frac{\beta(s)}{\sqrt{\epsilon}}) - f(x^\epsilon(s), \frac{\beta(t_{k+1})}{\sqrt{\epsilon}}) \right] ds \\ &\quad + \sum_{k=0}^{n-1} \int_{t_k}^{t_{k+1}} \left[f(x^\epsilon(s), \frac{\beta(t_{k+1})}{\sqrt{\epsilon}}) - f(x^\epsilon(t_k), \frac{\beta(t_{k+1})}{\sqrt{\epsilon}}) \right] ds \\ &\quad + \Delta t \sum_{k=0}^{n-1} \left[f(x^\epsilon(t_k), \frac{\beta(t_{k+1})}{\sqrt{\epsilon}}) - f(x_k^\epsilon, \frac{\beta(t_{k+1})}{\sqrt{\epsilon}}) \right]. \end{aligned}$$

The mapping $f : \mathbb{R} \times \mathbb{T} \rightarrow \mathbb{R}$ is assumed to be globally Lipschitz continuous, therefore applying Minkowski's inequality one obtains the following upper bound: there exists $C \in (0, \infty)$ such that for all $n \in \{0, \dots, N\}$ one has

$$\begin{aligned} (\mathbb{E}[\|x^\epsilon(t_n) - x_n^\epsilon\|^2])^{\frac{1}{2}} &\leq \frac{C}{\sqrt{\epsilon}} \sum_{k=0}^{n-1} \int_{t_k}^{t_{k+1}} (\mathbb{E}[|\beta(s) - \beta(t_{k+1})|^2])^{\frac{1}{2}} ds \\ &\quad + C \sum_{k=0}^{n-1} \int_{t_k}^{t_{k+1}} (\mathbb{E}[\|x^\epsilon(s) - x^\epsilon(t_k)\|^2])^{\frac{1}{2}} ds \\ &\quad + C\Delta t \sum_{k=0}^{n-1} (\mathbb{E}[\|x^\epsilon(t_k) - x_k^\epsilon\|^2])^{\frac{1}{2}}. \end{aligned}$$

For all $k \in \{0, \dots, N-1\}$ and $s \in [t_k, t_{k+1}]$, one has

$$\mathbb{E}[|\beta(s) - \beta(t_{k+1})|^2] = (t_{k+1} - s) \leq \Delta t,$$

owing to the standard properties of Brownian motion. Moreover, recalling that the mapping f is assumed to be bounded, for all $k \in \{0, \dots, N-1\}$ and all $s \in [t_k, t_{k+1}]$, one has

$$|x^\epsilon(s) - x^\epsilon(t_k)| = \left| \int_{t_k}^s f(x^\epsilon(r), \frac{\beta(r)}{\sqrt{\epsilon}}) dr \right| \leq \int_{t_k}^s |f(x^\epsilon(r), \frac{\beta(r)}{\sqrt{\epsilon}})| dr \leq C(s - t_k) \leq C\Delta t.$$

Therefore, for all $n \in \{0, \dots, N\}$ one has

$$(\mathbb{E}[\|x^\epsilon(t_n) - x_n^\epsilon\|^2])^{\frac{1}{2}} \leq C \left(\frac{\sqrt{\Delta t}}{\sqrt{\epsilon}} + \Delta t \right) + C\Delta t \sum_{k=0}^{n-1} (\mathbb{E}[\|x^\epsilon(t_k) - x_k^\epsilon\|^2])^{\frac{1}{2}}$$

Applying the discrete Grönwall inequality, one obtains the following upper bound:

$$\sup_{0 \leq n \leq N} (\mathbb{E}[\|X^\epsilon(t_n) - X_n^\epsilon\|^2])^{\frac{1}{2}} \leq C \frac{\sqrt{\Delta t}}{\sqrt{\epsilon}} + C\Delta t.$$

Note that for all $\epsilon \in (0, 1)$, one has the upper bound

$$\Delta t \leq \sqrt{T} \sqrt{\Delta t} \leq \sqrt{T} \frac{\sqrt{\Delta t}}{\sqrt{\epsilon}}.$$

This yields the error estimates (8) and the proof of Proposition 3.1 is thus completed.

A.2. Proof of Proposition 3.2. Combining the identities (29) and (30), the error $x^\epsilon(t_n) - x_n^\epsilon$ is decomposed as follows: for all $n \in \{0, \dots, N\}$, one has

$$x^\epsilon(t_n) - x_n^\epsilon = \delta_n^{\epsilon,1} + \delta_n^{\epsilon,2} + \delta_n^{\epsilon,3},$$

with auxiliary error terms $\delta_n^1, \delta_n^2, \delta_n^3$ defined by

$$\begin{aligned} \delta_n^{\epsilon,1} &= \sum_{k=0}^{n-1} \int_{t_k}^{t_{k+1}} \left[f(x^\epsilon(s), \frac{\beta(s)}{\sqrt{\epsilon}}) - f(x^\epsilon(t_k), \frac{\beta(s)}{\sqrt{\epsilon}}) \right] ds, \\ \delta_n^{\epsilon,2} &= \sum_{k=0}^{n-1} \int_{t_k}^{t_{k+1}} \left[f(x^\epsilon(t_k), \frac{\beta(s)}{\sqrt{\epsilon}}) - f(x_k^\epsilon, \frac{\beta(s)}{\sqrt{\epsilon}}) \right] ds, \\ \delta_n^{\epsilon,3} &= \sum_{k=0}^{n-1} \int_{t_k}^{t_{k+1}} \left[f(x_k^\epsilon, \frac{\beta(s)}{\sqrt{\epsilon}}) - f(x_k^\epsilon, \frac{\beta(t_{k+1})}{\sqrt{\epsilon}}) \right] ds. \end{aligned}$$

For the first error term δ_n^1 , recalling that the mapping f is globally Lipschitz continuous and applying Minkowski's inequality, for all $n \in \{0, \dots, N\}$ one has

$$(\mathbb{E}[|\delta_n^{\epsilon,1}|^2])^{\frac{1}{2}} \leq C \sum_{k=0}^{n-1} \int_{t_k}^{t_{k+1}} (\mathbb{E}[|x^\epsilon(s) - x^\epsilon(t_k)|^2])^{\frac{1}{2}} ds.$$

Moreover, as in the proof of Proposition 3.1, since the mapping f is assumed to be bounded, for all $k \in \{0, \dots, N-1\}$ and all $s \in [t_k, t_{k+1}]$, one has

$$|x^\epsilon(s) - x^\epsilon(t_k)| \leq C(s - t_k) \leq C\Delta t.$$

Therefore for all $n \in \{0, \dots, N\}$ one has

$$(\mathbb{E}[|\delta_n^{\epsilon,1}|^2])^{\frac{1}{2}} \leq C\Delta t.$$

For the second error term δ_n^2 , recalling that the mapping f is globally Lipschitz continuous and applying Minkowski's inequality, for all $n \in \{0, \dots, N\}$ one has

$$(\mathbb{E}[|\delta_n^{\epsilon,2}|^2])^{\frac{1}{2}} \leq C\Delta t \sum_{k=0}^{n-1} (\mathbb{E}[|x^\epsilon(t_k) - x_k^\epsilon|^2])^{\frac{1}{2}}.$$

The treatment of the third term δ_n^3 requires more attention. Applying Itô's formula, for all $k \in \{0, \dots, N-1\}$ and all $s \in [t_k, t_{k+1}]$ one has

$$f(x_k^\epsilon, \frac{\beta(s)}{\sqrt{\epsilon}}) - f(x_k^\epsilon, \frac{\beta(t_{k+1})}{\sqrt{\epsilon}}) = \int_s^{t_{k+1}} \partial_y f(x_k^\epsilon, \frac{\beta(r)}{\sqrt{\epsilon}}) \frac{d\beta(r)}{\sqrt{\epsilon}} + \int_s^{t_{k+1}} \partial_y^2 f(x_k^\epsilon, \frac{\beta(r)}{\sqrt{\epsilon}}) \frac{dr}{2\epsilon}.$$

Integrating with respect to $s \in [t_k, t_{k+1}]$ and applying stochastic and deterministic versions of Fubini's theorem, for all $k \in \{0, \dots, N-1\}$, one obtains

$$\begin{aligned}
& \int_{t_k}^{t_{k+1}} \left[f\left(x_k^\epsilon, \frac{\beta(s)}{\sqrt{\epsilon}}\right) - f\left(x_k^\epsilon, \frac{\beta(t_{k+1})}{\sqrt{\epsilon}}\right) \right] ds \\
&= \int_{t_k}^{t_{k+1}} \int_s^{t_{k+1}} \partial_y f\left(x_k^\epsilon, \frac{\beta(r)}{\sqrt{\epsilon}}\right) \frac{d\beta(r)}{\sqrt{\epsilon}} ds + \int_{t_k}^{t_{k+1}} \int_s^{t_{k+1}} \partial_y^2 f\left(x_k^\epsilon, \frac{\beta(r)}{\sqrt{\epsilon}}\right) \frac{dr}{2\epsilon} ds \\
&= \int_{t_k}^{t_{k+1}} \int_{t_k}^r \partial_y f\left(x_k^\epsilon, \frac{\beta(r)}{\sqrt{\epsilon}}\right) ds \frac{d\beta(r)}{\sqrt{\epsilon}} + \int_{t_k}^{t_{k+1}} \int_{t_k}^r \partial_y^2 f\left(x_k^\epsilon, \frac{\beta(r)}{\sqrt{\epsilon}}\right) ds \frac{dr}{2\epsilon} \\
&= \int_{t_k}^{t_{k+1}} (r - t_k) \partial_y f\left(x_k^\epsilon, \frac{\beta(r)}{\sqrt{\epsilon}}\right) \frac{d\beta(r)}{\sqrt{\epsilon}} + \int_{t_k}^{t_{k+1}} (r - t_k) \partial_y^2 f\left(x_k^\epsilon, \frac{\beta(r)}{\sqrt{\epsilon}}\right) \frac{dr}{2\epsilon}.
\end{aligned}$$

Therefore for all $n \in \{0, \dots, N\}$, the error term $\delta_n^{\epsilon,3}$ is decomposed as

$$\delta_n^{\epsilon,3} = \delta_n^{\epsilon,3,1} + \delta_n^{\epsilon,3,2},$$

where the auxiliary error terms $\delta_n^{\epsilon,3,1}$ and $\delta_n^{\epsilon,3,2}$ are defined by

$$\begin{aligned}
\delta_n^{\epsilon,3,1} &= \sum_{k=0}^{n-1} \int_{t_k}^{t_{k+1}} (r - t_k) \partial_y f\left(x_k^\epsilon, \frac{\beta(r)}{\sqrt{\epsilon}}\right) \frac{d\beta(r)}{\sqrt{\epsilon}}, \\
\delta_n^{\epsilon,3,2} &= \sum_{k=0}^{n-1} \int_{t_k}^{t_{k+1}} (r - t_k) \partial_y^2 f\left(x_k^\epsilon, \frac{\beta(r)}{\sqrt{\epsilon}}\right) \frac{dr}{2\epsilon}.
\end{aligned}$$

Applying the Itô isometry property and recalling that the first order partial derivative $\partial_y f$ is bounded, for the error term $\delta_n^{\epsilon,3,1}$, for all $n \in \{0, \dots, N-1\}$, one has the upper bound

$$\mathbb{E}[|\delta_n^{\epsilon,3,1}|^2] = \frac{1}{\epsilon} \sum_{k=0}^{n-1} \int_{t_k}^{t_{k+1}} \left| (r - t_k) \partial_y f\left(x_k^\epsilon, \frac{\beta(r)}{\sqrt{\epsilon}}\right) \right|^2 dr \leq C \frac{\Delta t^2}{\epsilon}.$$

Applying Minkowski's inequality and recalling that the second order partial derivative $\partial_y^2 f$ is bounded, for the error term $\delta_n^{\epsilon,3,2}$, for all $n \in \{0, \dots, N-1\}$, one has the upper bound

$$(\mathbb{E}[|\delta_n^{\epsilon,3,2}|^2])^{\frac{1}{2}} \leq \frac{1}{2\epsilon} \sum_{k=0}^{n-1} \int_{t_k}^{t_{k+1}} (r - t_k) (\mathbb{E}[|\partial_y^2 f\left(x_k^\epsilon, \frac{\beta(r)}{\sqrt{\epsilon}}\right)|^2])^{\frac{1}{2}} dr \leq C \frac{\Delta t}{\epsilon}.$$

Gathering the upper bounds on the error terms $\delta_n^{\epsilon,3,1}$ and $\delta_n^{\epsilon,3,2}$ above, and using the upper bound $\epsilon^{-\frac{1}{2}} \leq \epsilon^{-1}$ for all $\epsilon \in (0, 1)$, for the error term $\delta_n^{\epsilon,3}$, for all $n \in \{0, \dots, N\}$ one obtains

$$(\mathbb{E}[|\delta_n^{\epsilon,3}|^2])^{\frac{1}{2}} \leq \frac{C\Delta t}{\epsilon}.$$

Finally, gathering the estimates for the error terms $\delta_n^{\epsilon,1}$, $\delta_n^{\epsilon,2}$ and $\delta_n^{\epsilon,3}$, and using the upper bound $1 \leq \epsilon^{-1}$ for all $\epsilon \in (0, 1)$, for all $n \in \{0, \dots, N\}$ one obtains

$$(\mathbb{E}[|x^\epsilon(t_n) - x_n^\epsilon|^2])^{\frac{1}{2}} \leq C\Delta t \sum_{k=0}^{n-1} (\mathbb{E}[|x^\epsilon(t_k) - x_k^\epsilon|^2])^{\frac{1}{2}} + C \frac{\Delta t}{\epsilon}.$$

Applying the discrete Grönwall inequality, one obtains

$$\sup_{0 \leq n \leq N} (\mathbb{E}[|x^\epsilon(t_n) - x_n^\epsilon|^2])^{\frac{1}{2}} \leq C \frac{\Delta t}{\epsilon}$$

This yields the error estimate (9) and the proof of Proposition 3.2 is thus completed.

APPENDIX B. PROOF OF THE ASYMPTOTIC PRESERVING PROPERTY

Proof of Lemma 3.4. Let $\gamma \sim \mathcal{N}(0, 1)$ be a standard real-valued Gaussian random variable and $U \sim \mathcal{U}(\mathbb{T})$ be a uniformly distributed random variable on the torus \mathbb{T} .

To prove the convergence in distribution $\widehat{\gamma}^\epsilon \rightarrow U$ when $\epsilon \rightarrow 0$, it suffices to show that for all $n \in \mathbb{Z}$,

$$\mathbb{E}[e^{i2\pi n \widehat{\gamma}^\epsilon}] \xrightarrow{\epsilon \rightarrow 0} \mathbb{E}[e^{i2\pi n U}] = \mathbb{1}_{n=0}.$$

Let us consider $n \in \mathbb{Z} \setminus \{0\}$ (the case $n = 0$ is trivial). One has

$$\mathbb{E}[e^{i2\pi n \widehat{\gamma}^\epsilon}] = \mathbb{E}[e^{i2\pi n \gamma^\epsilon}] = e^{-\frac{(2\pi)^2 n^2}{2\epsilon}} \xrightarrow{\epsilon \rightarrow 0} 0.$$

The proof is completed. \square

Proof of Lemma 3.5. For all $n \in \{1, \dots, N-1\}$, let $\delta\beta_{n+1}^\epsilon = \frac{\delta\beta_{n+1}}{\sqrt{\epsilon}}$.

The scheme (7) can be written recursively as follows: for all $n \in \{0, \dots, N-1\}$, one has

$$(x_{n+1}^\epsilon, y_{n+1}^\epsilon) = \Phi_{\Delta t}(x_n^\epsilon, y_n^\epsilon, \widehat{\delta\beta_{n+1}^\epsilon})$$

where the mapping $\Phi_{\Delta t} : \mathbb{R} \times \mathbb{T}^2 \rightarrow \mathbb{R} \times \mathbb{T}$ is defined by

$$\Phi_{\Delta t}(x, y, \widehat{\gamma}) = (x + \Delta t f(x, y), y + \widehat{\gamma}).$$

It is then straightforward to show that for all $n \in \{1, \dots, N\}$ one has

$$(x_n^\epsilon, y_n^\epsilon) = \Phi_{\Delta t}^{(n)}(x_0, y_0, \widehat{\delta\beta_1^\epsilon}, \dots, \widehat{\delta\beta_n^\epsilon})$$

for some mapping $\Phi_{\Delta t}^{(n)} : \mathbb{R} \times \mathbb{T}^{n+1} \rightarrow \mathbb{R} \times \mathbb{T}$. As a result, one has

$$((x_n^\epsilon, y_n^\epsilon))_{0 \leq n \leq N} = \Psi_{\Delta t}(x_0, y_0, \widehat{\delta\beta_1^\epsilon}, \dots, \widehat{\delta\beta_N^\epsilon}),$$

for some mapping $\Psi_{\Delta t} : \mathbb{R} \times \mathbb{T}^{N+1} \rightarrow (\mathbb{R} \times \mathbb{T})^{N+1}$.

The mapping $\Psi_{\Delta t}$ is continuous. Moreover, the random variables $\widehat{\delta\beta_1^\epsilon}, \dots, \widehat{\delta\beta_N^\epsilon}$ are independent. Owing to Lemma 3.4, one then obtains the convergence in distribution

$$((x_n^\epsilon, y_n^\epsilon))_{0 \leq n \leq N} \xrightarrow{\epsilon \rightarrow 0} ((X_n, Y_n))_{0 \leq n \leq N} = \Psi_{\Delta t}(x_0, y_0, U_1, \dots, U_N)$$

where U_1, \dots, U_N are independent random variables uniformly distributed on \mathbb{T} , such that $\widehat{\delta\beta_n^\epsilon} \rightarrow U_n$ in distribution when $\epsilon \rightarrow 0$, for all $n \in \{1, \dots, N\}$.

By construction of the mappings $\Psi_{\Delta t}$ and $\Phi_{\Delta t}^{(n)}$, for all $n \in \{0, \dots, N-1\}$, one obtains

$$(X_n, Y_n) = \Phi_{\Delta t}^{(n)}(x_0, y_0, U_1, \dots, U_n)$$

$$(X_{n+1}, Y_{n+1}) = \Phi_{\Delta t}(X_n, Y_n, U_{n+1}),$$

which finally shows that the limit in distribution $((X_n, Y_n))_{0 \leq n \leq N} = \lim_{\epsilon \rightarrow 0} ((x_n^\epsilon, y_n^\epsilon))_{0 \leq n \leq N}$ is given by the scheme (10).

This concludes the proof of Lemma 3.5. \square

Proof of Lemma 3.6. The objective is to prove that for all $n \in \{0, \dots, N\}$ the random variables X_n and \tilde{X}_n , given by (10) and (11) respectively, are equal in distribution. This is established recursively. When $n = 0$, the equality is trivial since $\tilde{X}_0 = X_0 = x_0$.

Let $n \in \{0, \dots, N-1\}$ and assume that X_n and \tilde{X}_n are equal in distribution.

Let $\varphi : \mathbb{R} \rightarrow \mathbb{R}$ be a bounded continuous function. Then one has

$$\begin{aligned} \mathbb{E}[\varphi(X_{n+1})] &= \mathbb{E}[\varphi(X_n + \Delta f(X_n, Y_{n+1}))] \\ &= \mathbb{E}[\varphi(X_n + \Delta t f(X_n, Y_n + U_{n+1}))], \end{aligned}$$

$$\mathbb{E}[\varphi(\tilde{X}_{n+1})] = \mathbb{E}[\varphi(\tilde{X}_n + \Delta t f(\tilde{X}_n + U_{n+1}))]$$

The random variables (X_n, Y_n) and U_{n+1} are independent. Similarly, the random variables \tilde{X}_n and U_{n+1} are independent. As a result, one obtains

$$\begin{aligned}\mathbb{E}[\varphi(X_n + \Delta t f(X_n, Y_n + U_{n+1}))] &= \mathbb{E}[\bar{\varphi}(X_n, Y_n)], \\ \mathbb{E}[\varphi(\tilde{X}_n + \Delta t f(\tilde{X}_n, U_{n+1}))] &= \mathbb{E}[\bar{\varphi}(\tilde{X}_n, 0)],\end{aligned}$$

where the auxiliary function $\bar{\varphi}$ is defined by

$$\bar{\varphi}(x, y) = \int_{\mathbb{T}} \varphi(x + \Delta t f(x, y + u)) \, du = \int_{\mathbb{T}} \varphi(x + \Delta t f(x, u)) \, du.$$

From the above (which is due to translational invariance of the uniform distribution on the torus \mathbb{T}), one has $\bar{\varphi}(x, y) = \bar{\varphi}(x, 0)$ for all $y \in \mathbb{T}$. The function $\bar{\varphi}(\cdot, 0)$ is continuous.

As a result, taking into account the assumption that X_n and \tilde{X}_n are equal in distribution, one obtains

$$\mathbb{E}[\varphi(X_{n+1})] = \mathbb{E}[\bar{\varphi}(X_n, 0)] = \mathbb{E}[\bar{\varphi}(\tilde{X}_n, 0)] = \mathbb{E}[\varphi(\tilde{X}_{n+1})].$$

Since the identity above holds for arbitrary bounded continuous function φ , if the random variables X_n and \tilde{X}_n are equal in distribution, then the random variables X_{n+1} and \tilde{X}_{n+1} .

The equality in distribution of X_n and \tilde{X}_n for all $n \in \{0, \dots, N\}$ is then obtained by recursion. The proof of Lemma 3.6 is then completed. \square

Proof of Lemma 3.7. For all $n \in \{0, \dots, N\}$, one has

$$\begin{aligned}X(t_n) &= x_0 + \int_0^{t_n} f(X(s)) \, ds = x_0 + \sum_{k=0}^{n-1} \int_{t_k}^{t_{k+1}} F(X(s)) \, ds, \\ \tilde{X}_n &= x_0 + \Delta t \sum_{k=0}^{n-1} f(\tilde{X}_k, U_{k+1}).\end{aligned}$$

Therefore, the error $X(t_n) - \tilde{X}_n$ can be decomposed as

$$X(t_n) - \tilde{X}_n = \delta_n^1 + \delta_n^2 + \delta_n^3,$$

with auxiliary error terms $\delta_n^1, \delta_n^2, \delta_n^3$ defined by

$$\begin{aligned}\delta_n^1 &= \sum_{k=0}^{n-1} \int_{t_k}^{t_{k+1}} [F(X(s)) - F(X(t_k))] \, ds, \\ \delta_n^2 &= \Delta t \sum_{k=0}^{n-1} [F(X(t_k)) - F(\tilde{X}_k)], \\ \delta_n^3 &= \Delta t \sum_{k=0}^{n-1} [F(\tilde{X}_k) - f(\tilde{X}_k, U_{k+1})].\end{aligned}$$

Recalling that f is assumed to be bounded and globally Lipschitz continuous, the mapping F defined by (6) is also bounded and globally Lipschitz continuous. Therefore the solution X to the ordinary differential equation (5) satisfies the property

$$|X(s) - X(t_k)| \leq C|s - t_k| \leq C\Delta t, \quad \forall s \in [t_k, t_{k+1}], \quad \forall k \in \{0, \dots, N-1\},$$

for some $C \in (0, \infty)$. By the Lipschitz continuity property of F , one obtains the upper bound

$$\sup_{0 \leq n \leq N} |\delta_n^1| \leq C\Delta t.$$

For the second error term δ_n^2 , owing to Minkowski's inequality, using the Lipschitz continuity property of F , one has

$$(\mathbb{E}[|\delta_n^2|^2])^{\frac{1}{2}} \leq C\Delta t \sum_{k=0}^{n-1} (\mathbb{E}[|X(t_k) - \tilde{X}_k|^2])^{\frac{1}{2}}.$$

To deal with the third error term δ_n^3 , one has

$$\begin{aligned} |\delta_n^3|^2 &= \Delta t^2 \sum_{k=0}^{n-1} |F(\tilde{X}_k) - f(\tilde{X}_k, U_{k+1})|^2 \\ &\quad + 2\Delta t^2 \sum_{k=0}^{n-1} \sum_{\ell=k+1}^{n-1} (F(\tilde{X}_k) - f(\tilde{X}_k, U_{k+1})) (F(\tilde{X}_\ell) - f(\tilde{X}_\ell, U_{\ell+1})) \end{aligned}$$

Since the mappings f and F are bounded, for the first term in the right-hand side above, one obtains the upper bound

$$\mathbb{E} \left[\left(\Delta t^2 \sum_{k=0}^{n-1} |F(\tilde{X}_k) - f(\tilde{X}_k, U_{k+1})|^2 \right) \right] \leq C\Delta t.$$

For the second error term, note that for all $k, \ell \in \{0, \dots, n-1\}$ such that $\ell \geq k+1$, the random variables \tilde{X}_k, U_{k+1} and \tilde{X}_ℓ are independent of $U_{\ell+1}$. As a result,

$$\mathbb{E} \left[(F(\tilde{X}_k) - f(\tilde{X}_k, U_{k+1})) (F(\tilde{X}_\ell) - f(\tilde{X}_\ell, U_{\ell+1})) \right] = 0.$$

Therefore, one obtains the upper bound

$$\sup_{0 \leq n \leq N} (\mathbb{E}[|\delta_n^3|^2])^{\frac{1}{2}} \leq C\Delta t^{\frac{1}{2}}.$$

Combining the upper bounds, for all $n \in \{0, \dots, N\}$, one has

$$(\mathbb{E}[|X(t_n) - \tilde{X}_n|^2])^{\frac{1}{2}} \leq C\Delta t^{\frac{1}{2}} + C\Delta t \sum_{k=0}^{n-1} (\mathbb{E}[|X(t_k) - \tilde{X}_k|^2])^{\frac{1}{2}}.$$

Applying the discrete Grönwall inequality, one obtains the inequality

$$\sup_{0 \leq n \leq N} (\mathbb{E}[|X(t_n) - \tilde{X}_n|^2])^{\frac{1}{2}} \leq C\Delta t^{\frac{1}{2}}$$

This yields the error estimate (12) and the proof of Lemma 3.7 is thus completed. \square

APPENDIX C. PROOF OF THE UNIFORM STRONG ERROR ESTIMATES

This section is devoted to stating and proving the auxiliary results required in Section 3.3 to prove Theorem 3.8.

Given the time-step size $\Delta t = T/N$, let us introduce an auxiliary scheme: for all $n \in \{0, \dots, N-1\}$, set

$$\bar{X}_{n+1} = \bar{X}_n + \Delta t F(\bar{X}_n),$$

with initial value $\bar{X}_0 = x_0$. This is the standard (forward) Euler scheme applied to the ordinary differential equation (5), and the following error bounds are standard: there exists $C \in (0, \infty)$ such that for all $N \in \mathbb{N}^*$ one has

$$(31) \quad \sup_{0 \leq n \leq N} |\bar{X}_n - X(t_n)| \leq C\Delta t.$$

Lemma C.1 is related to the averaging principle presented in Section 2, and gives an upper bound for the error between $x^\epsilon(t)$ and $X(t) = \lim_{\epsilon \rightarrow 0} x^\epsilon(t)$. The (strong) order of convergence with respect to ϵ is equal to $1/2$, which is standard in the literature. For completeness a sketch of proof is provided.

Lemma C.1. *There exists $C \in (0, \infty)$ such that for all $\epsilon \in (0, 1)$ one has*

$$(32) \quad \sup_{t \in [0, T]} (\mathbb{E}[|x^\epsilon(t) - X(t)|^2])^{\frac{1}{2}} \leq C\sqrt{\epsilon}.$$

Proof of Lemma C.1. Let $g : \mathbb{R} \times \mathbb{T} \rightarrow \mathbb{R}$ be defined such that for all $x \in \mathbb{R}$ the mapping $g(x, \cdot)$ is solution to the Poisson equation

$$\begin{cases} -\frac{1}{2}\partial_y^2 g(x, y) = f(x, y) - F(x), & \forall y \in \mathbb{T}, \\ \int_{\mathbb{T}} g(x, y) dy = 0. \end{cases}$$

The function g is given by

$$g(x, y) = \int_0^{+\infty} (\mathbb{E}[f(x, y + \beta(t))] - F(x)) dt, \quad \forall x \in \mathbb{R}, \forall y \in \mathbb{T}.$$

By standard arguments, one can show that the mapping g and its partial derivatives satisfies the following upper bounds:

$$\sup_{x \in \mathbb{R}, y \in \mathbb{T}} (|g(x, y)| + |\partial_x g(x, y)| + |\partial_y g(x, y)| + |\partial_y^2 g(x, y)|) < \infty.$$

The details are omitted.

For all $t \in [0, T]$, one has

$$\begin{aligned} x^\epsilon(t) - X(t) &= \int_0^t f(x^\epsilon(s), y^\epsilon(s)) ds - \int_0^t F(X(s)) ds \\ &= \int_0^t [f(x^\epsilon(s), y^\epsilon(s)) - F(x^\epsilon(s))] ds + \int_0^t [F(x^\epsilon(s)) - F(X(s))] ds. \end{aligned}$$

Applying the Itô formula to the solution $(x^\epsilon(t), y^\epsilon(t))_{t \in [0, T]}$ to (4), for all $t \in [0, T]$ one obtains

$$\begin{aligned} g(x^\epsilon(t), y^\epsilon(t)) - g(x_0, y_0) &= \int_0^t f(x^\epsilon(s), y^\epsilon(s)) \partial_x g(x^\epsilon(s), y^\epsilon(s)) ds \\ &\quad + \frac{1}{\sqrt{\epsilon}} \int_0^t \partial_y g(x^\epsilon(s), y^\epsilon(s)) d\beta(s) + \frac{1}{2\epsilon} \int_0^t \partial_y^2 g(x^\epsilon(s), y^\epsilon(s)) ds. \end{aligned}$$

Owing to the definition of the mapping g , one then obtains

$$\begin{aligned} \int_0^t [f(x^\epsilon(s), y^\epsilon(s)) - F(x^\epsilon(s))] ds &= -\frac{1}{2} \int_0^t \partial_y^2 g(x^\epsilon(s), y^\epsilon(s)) ds \\ &= \epsilon [g(x_0, y_0) - g(x^\epsilon(t), y^\epsilon(t))] \\ &\quad + \epsilon \int_0^t f(x^\epsilon(s), y^\epsilon(s)) \partial_x g(x^\epsilon(s), y^\epsilon(s)) ds + \sqrt{\epsilon} \int_0^t \partial_y g(x^\epsilon(s), y^\epsilon(s)) d\beta(s). \end{aligned}$$

Due to the bounds on g and its derivatives given above, there exists $C \in (0, \text{inf}ty)$ such that for all $t \in [0, T]$ and all $\epsilon \in (0, 1)$ one has

$$(\mathbb{E}[|g(x_0, y_0) - g(x^\epsilon(t), y^\epsilon(t))|^2])^{\frac{1}{2}} + (\mathbb{E}[\int_0^t f(x^\epsilon(s), y^\epsilon(s)) \partial_x g(x^\epsilon(s), y^\epsilon(s)) ds]^2)^{\frac{1}{2}} \leq C,$$

and applying the Itô isometry formula one has

$$(\mathbb{E}[\int_0^t \partial_y g(x^\epsilon(s), y^\epsilon(s)) d\beta(s)]^2)^{\frac{1}{2}} = (\int_0^t \mathbb{E}[|\partial_y g(x^\epsilon(s), y^\epsilon(s))|^2] ds)^{\frac{1}{2}} \leq C.$$

As a consequence, there exists $C \in (0, 1)$ such that for all $\epsilon \in (0, 1)$ one has

$$\sup_{0 \leq t \leq T} (\mathbb{E}[\int_0^t [f(x^\epsilon(s), y^\epsilon(s)) - F(x^\epsilon(s))] ds]^2)^{\frac{1}{2}} \leq C\sqrt{\epsilon}.$$

To conclude the proof, it suffices to note that due to the decomposition of $x^\epsilon(t) - X(t)$ above and taking into account the Lipschitz continuity property of F , one obtains for all $t \in [0, T]$

$$(\mathbb{E}[|x^\epsilon(t) - X(t)|^2])^{\frac{1}{2}} \leq C\sqrt{\epsilon} + C \int_0^t (\mathbb{E}[|x^\epsilon(s) - X(s)|^2])^{\frac{1}{2}} ds.$$

Applying Grönwall's inequality then provides the error estimates (32) and concludes the proof of Lemma C.1. \square

A version of Lemma C.1 in the discrete time setting, to deal with the behavior of x_n^ϵ when $\epsilon \rightarrow 0$, is given below. Note that Lemma 3.5 would suggest to consider the error $x_n^\epsilon - X_n$ or $x_n^\epsilon - \bar{X}_n$. Instead, the error $x_n^\epsilon - \bar{X}_n$ is considered. By the triangle inequality, combining the error estimates (12) (31) give

$$\sup_{0 \leq n \leq N} (\mathbb{E}[|\bar{X}_n - X_n|^2])^{\frac{1}{2}} \leq C\sqrt{\Delta t}.$$

Lemma C.2. *There exists $C \in (1, \infty)$ such that for all $\epsilon \in (0, 1)$ and all $N \in \mathbb{N}^*$, one has*

$$(33) \quad \sup_{0 \leq n \leq N} (\mathbb{E}[|x_n^\epsilon - \bar{X}_n|^2])^{\frac{1}{2}} \leq C \left(\sqrt{\Delta t} + \frac{\sqrt{\epsilon}}{\sqrt{1 - e^{-\frac{\Delta t}{C\epsilon}}}} \right).$$

Proof of Lemma C.2. For all $n \in \{0, \dots, N\}$, one has

$$\begin{aligned} x_n^\epsilon - \bar{X}_n &= \Delta t \sum_{k=0}^{n-1} \left(f(x_k^\epsilon, \frac{\beta(t_{k+1})}{\sqrt{\epsilon}}) - F(\bar{X}_k) \right) \\ &= \Delta t \sum_{k=0}^{n-1} \left(f(x_k^\epsilon, \frac{\beta(t_{k+1})}{\sqrt{\epsilon}}) - f(\bar{X}_k, \frac{\beta(t_{k+1})}{\sqrt{\epsilon}}) \right) + \Delta t \sum_{k=0}^{n-1} \left(f(\bar{X}_k, \frac{\beta(t_{k+1})}{\sqrt{\epsilon}}) - F(\bar{X}_k) \right). \end{aligned}$$

Applying Minkowski's inequality, one obtains the upper bound

$$(\mathbb{E}[|x_n^\epsilon - \bar{X}_n|^2])^{\frac{1}{2}} \leq C\Delta t \sum_{k=0}^{n-1} (\mathbb{E}[|x_k^\epsilon - \bar{X}_k|^2])^{\frac{1}{2}} + \Delta t (\mathbb{E}[|\sum_{k=0}^{n-1} (f(\bar{X}_k, \frac{\beta(t_{k+1})}{\sqrt{\epsilon}}) - F(\bar{X}_k))|^2])^{\frac{1}{2}}.$$

To deal with the second term in the right-hand side above, one has almost surely

$$\begin{aligned} \left| \sum_{k=0}^{n-1} \left(f(\bar{X}_k, \frac{\beta(t_{k+1})}{\sqrt{\epsilon}}) - F(\bar{X}_k) \right) \right|^2 &= \sum_{k=0}^{n-1} \left| f(\bar{X}_k, \frac{\beta(t_{k+1})}{\sqrt{\epsilon}}) - F(\bar{X}_k) \right|^2 \\ &\quad + 2 \sum_{0 \leq k < \ell \leq n-1} \left(f(\bar{X}_k, \frac{\beta(t_{k+1})}{\sqrt{\epsilon}}) - F(\bar{X}_k) \right) \left(f(\bar{X}_\ell, \frac{\beta(t_{\ell+1})}{\sqrt{\epsilon}}) - F(\bar{X}_\ell) \right). \end{aligned}$$

One the one hand, since the mappings f and F are bounded, one has for all $n \in \{0, \dots, N\}$

$$\mathbb{E} \left[\sum_{k=0}^{n-1} \left| f(\bar{X}_k, \frac{\beta(t_{k+1})}{\sqrt{\epsilon}}) - F(\bar{X}_k) \right|^2 \right] \leq Cn \leq \frac{C}{\Delta t}.$$

On the other hand, using the independence of increments of Brownian motion and a conditional expectation argument, for all $k, \ell \in \{0, \dots, n-1\}$ such that $\ell > k$, one has

$$\mathbb{E} \left[\left(f(\bar{X}_k, \frac{\beta(t_{k+1})}{\sqrt{\epsilon}}) - F(\bar{X}_k) \right) \left(f(\bar{X}_\ell, \frac{\beta(t_{\ell+1})}{\sqrt{\epsilon}}) - F(\bar{X}_\ell) \right) \right] = \mathbb{E} \left[h_{t_{k+1}, t_{\ell+1}, \epsilon} \left(\bar{X}_k, \bar{X}_\ell, \frac{\beta(t_{k+1})}{\sqrt{\epsilon}} \right) \right]$$

with the auxiliary mapping $h_{t_{k+1}, t_{\ell+1}, \epsilon}$ defined by

$$\begin{aligned} h(x_1, x_2, \gamma) &= (f(x_1, \gamma) - F(x_1)) \mathbb{E} \left[\left(f(x_2, \gamma + \frac{\beta(t_{\ell+1}) - \beta(t_{k+1})}{\sqrt{\epsilon}}) - F(x_2) \right) \right] \\ &= (f(x_1, \gamma) - F(x_1)) \mathbb{E} \left[\left(f(x_2, \gamma + \frac{\beta(t_{\ell-k})}{\sqrt{\epsilon}}) - F(x_2) \right) \right]. \end{aligned}$$

The mappings f and F are bounded. Moreover, there exists $C \in (1, \infty)$ such that for all $x, \gamma \in \mathbb{R}$ and all $t \geq 0$, one has

$$|\mathbb{E}[f(x, \gamma + \beta(t))] - F(x)| \leq Ce^{-\frac{t}{C}}.$$

As a result, one obtains the following upper bounds: for all $k, \ell \in \{0, \dots, n-1\}$ such that $\ell > k$, one has

$$|\mathbb{E}[h_{t_{k+1}, t_{\ell+1}, \epsilon}(\bar{X}_k, \bar{X}_\ell, \frac{\beta(t_{k+1})}{\sqrt{\epsilon}})]| \leq Ce^{-\frac{(\ell-k)\Delta t}{C\epsilon}}.$$

As a result, one has

$$\begin{aligned}
\left| \mathbb{E} \left[\sum_{0 \leq k < \ell \leq n-1} \left(f(\bar{X}_k, \frac{\beta(t_{k+1})}{\sqrt{\epsilon}}) - F(\bar{X}_k) \right) \left(f(\bar{X}_\ell, \frac{\beta(t_{\ell+1})}{\sqrt{\epsilon}}) - F(\bar{X}_\ell) \right) \right] \right| &\leq C \sum_{0 \leq k < \ell \leq n-1} e^{-\frac{(\ell-k)\Delta t}{C\epsilon}} \\
&\leq C \sum_{k=0}^{n-1} \frac{e^{-\frac{\Delta t}{C\epsilon}}}{1 - e^{-\frac{\Delta t}{C\epsilon}}} \\
&\leq \frac{C e^{-\frac{\Delta t}{C\epsilon}}}{\Delta t (1 - e^{-\frac{\Delta t}{C\epsilon}})}
\end{aligned}$$

Combining the upper bounds obtained above and using the upper bound $\sup_{x \geq 0} \sqrt{x} e^{-x} < \infty$, one obtains for all $n \in \{0, \dots, N\}$

$$\begin{aligned}
\Delta t (\mathbb{E} [| \sum_{k=0}^{n-1} (f(\bar{X}_k, \frac{\beta(t_{k+1})}{\sqrt{\epsilon}}) - F(\bar{X}_k)) |^2])^{\frac{1}{2}} &\leq C \sqrt{\Delta t} + C \frac{C \sqrt{\Delta t} e^{-\frac{\Delta t}{2C\epsilon}}}{\sqrt{1 - e^{-\frac{\Delta t}{C\epsilon}}}} \\
&\leq C \sqrt{\Delta t} + C \frac{C \sqrt{\epsilon}}{\sqrt{1 - e^{-\frac{\Delta t}{C\epsilon}}}}.
\end{aligned}$$

Therefore, for all $n \in \{0, \dots, N\}$, one has

$$(\mathbb{E} [\|x_n^\epsilon - \bar{X}_n\|^2])^{\frac{1}{2}} \leq C \Delta t \sum_{k=0}^{n-1} (\mathbb{E} [\|x_k^\epsilon - \bar{X}_k\|^2])^{\frac{1}{2}} + C \sqrt{\Delta t} + C \frac{C \sqrt{\epsilon}}{\sqrt{1 - e^{-\frac{\Delta t}{C\epsilon}}}}.$$

Applying the discrete Grönwall inequality then yields the error estimates (33) and concludes the proof of Lemma C.2. \square

Proof of Theorem 3.8. Let $\alpha \in [0, 1]$ be an auxiliary parameter.

On the one hand, if the parameters Δt and ϵ satisfy the condition $\Delta t^\alpha \leq \epsilon$, then applying Proposition 3.2 yields the upper bound

$$(34) \quad \sup_{0 \leq n \leq N} (\mathbb{E} [|x_n^\epsilon - x^\epsilon(t_n)|^2])^{\frac{1}{2}} \leq C \frac{\Delta t}{\epsilon} = C \Delta t^{1-\alpha} \frac{\Delta t^\alpha}{\epsilon} \leq C \Delta t^{1-\alpha}.$$

On the other hand, if the parameters Δt and ϵ satisfy the condition $\Delta t^\alpha \geq \epsilon$, applying Minkowski's inequality, one has

$$\begin{aligned}
&\sup_{0 \leq n \leq N} (\mathbb{E} [|x_n^\epsilon - x^\epsilon(t_n)|^2])^{\frac{1}{2}} \\
&\leq \sup_{0 \leq n \leq N} (\mathbb{E} [|x_n^\epsilon - \bar{X}_n|^2])^{\frac{1}{2}} + \sup_{0 \leq n \leq N} (\mathbb{E} [|\bar{X}_n - X(t_n)|^2])^{\frac{1}{2}} + \sup_{0 \leq n \leq N} (\mathbb{E} [|X(t_n) - x^\epsilon(t_n)|^2])^{\frac{1}{2}}.
\end{aligned}$$

Owing to the error estimates (33) (from Lemma C.2, (31) and (32) (from Lemma C.2), there exists $C \in (1, \infty)$ such that one has

$$\sup_{0 \leq n \leq N} (\mathbb{E} [|x_n^\epsilon - x^\epsilon(t_n)|^2])^{\frac{1}{2}} \leq C \sqrt{\Delta t} + C \frac{C \sqrt{\epsilon}}{\sqrt{1 - e^{-\frac{\Delta t}{C\epsilon}}}} + C \Delta t + C \sqrt{\epsilon}.$$

Under the condition $\Delta t^\alpha \geq \epsilon$, one has

$$1 - \exp\left(-\frac{\Delta t}{C\epsilon}\right) \geq 1 - \exp\left(-\frac{1}{C\epsilon^{\frac{1}{\alpha}-1}}\right) \geq 1 - \exp\left(-\frac{1}{C}\right).$$

As a result, one obtains the upper bound

$$(35) \quad \sup_{0 \leq n \leq N} (\mathbb{E} [|x_n^\epsilon - x^\epsilon(t_n)|^2])^{\frac{1}{2}} \leq C(\sqrt{\Delta t} + \sqrt{\epsilon}) \leq C \Delta t^{\frac{\alpha}{2}}.$$

Choosing $\alpha = 2/3$, one has $1 - \alpha = \frac{\alpha}{2} = \frac{1}{3}$, and combining the error estimates (34) and (35) one obtains for all $\epsilon \in (0, 1)$ and $N \in \mathbb{N}^*$ (with $\Delta t = T/N$)

$$\sup_{0 \leq n \leq N} (\mathbb{E}[|x_n^\epsilon - x^\epsilon(t_n)|^2])^{\frac{1}{2}} \leq C\Delta t^{\frac{1}{3}}.$$

The above estimate does not require any conditions on $\epsilon \in (0, 1)$ and Δt . This yields the uniform strong error estimates (14). The proof of Theorem 3.8 is completed. \square

Remark C.3. *If one applies Proposition 3.1 instead of Proposition 3.2 in the first argument of the proof above (in the case $\Delta t^\alpha \leq \epsilon$), the inequality (34) is replaced by*

$$(36) \quad \sup_{0 \leq n \leq N} (\mathbb{E}[|x_n^\epsilon - x^\epsilon(t_n)|^2])^{\frac{1}{2}} \leq C \frac{\sqrt{\Delta t}}{\sqrt{\epsilon}} = C\Delta t^{\frac{1-\alpha}{2}} \frac{\Delta t^{\frac{\alpha}{2}}}{\epsilon^{\frac{1}{2}}} \leq C\Delta t^{\frac{1-\alpha}{2}}.$$

Choosing $\alpha = 1/2$, one has $\frac{1-\alpha}{2} = \frac{\alpha}{2} = \frac{1}{4}$, and combining the error estimates (36) and (35) one obtains uniform strong error estimates

$$\sup_{0 \leq n \leq N} (\mathbb{E}[|x_n^\epsilon - x^\epsilon(t_n)|^2])^{\frac{1}{2}} \leq C\Delta t^{\frac{1}{4}}$$

with order of convergence $1/4$ instead of $1/3$ in Theorem 3.8, as claimed in Section 3.3.

UNIVERSITE DE PAU ET DES PAYS DE L'ADOUR, E2S UPPA, CNRS, LMAP, PAU, FRANCE
 Email address: charles-edouard.brehier@univ-pau.fr