Asymptotic preserving schemes for SDEs driven by fractional Brownian motion in the averaging regime

Charles-Edouard Bréhier

Asymptotic preserving schemes for SDEs driven by fractional Brownian motion in the averaging regime

Charles-Edouard Bréhier

Abstract. We design numerical schemes for a class of slow-fast systems of stochastic differential equations, where the fast component is an Ornstein-Uhlenbeck process and the slow component is driven by a fractional Brownian motion with Hurst index $H > 1/2$. We establish the asymptotic preserving property of the proposed scheme: when the time-scale parameter goes to 0, a limiting scheme which is consistent with the averaged equation is obtained. With this numerical analysis point of view, we thus illustrate the recently proved averaging result for the considered SDE systems and the main differences with the standard Wiener case.

1. Introduction

Multiscale and stochastic systems propose theoretical and computational challenges in all fields of science, including for instance fluid dynamics, biology, finance and engineering. Many models are driven by the standard Wiener process, however the fractional Brownian motion [**8**] is another popular model, especially in finance and turbulence modelling.

In this article, we consider stochastic differential equations (SDEs) of the following type:

$$\left\{\begin{array}{l} dX^\epsilon(t) = g(X^\epsilon(t), m^\epsilon(t))d\beta^H(t), \\ dm^\epsilon(t) = -\dfrac{1}{\epsilon}m^\epsilon(t)dt + \dfrac{\sqrt{2}}{\sqrt{\epsilon}}dB(t), \end{array}\right. \tag{1}$$

with $X^\epsilon(0) = x_0$, $m^\epsilon(0) = m_0^\epsilon$, where ϵ is the time-scale separation parameter, $\big(B(t)\big)_{t\geqslant 0}$ is a standard Wiener process, $g : \mathbb{R}^2 \to \mathbb{R}$ is a sufficiently regular mapping (see Section 2.3 for more details), and $\big(\beta^H(t)\big)_{t\geqslant 0}$ is a fractional Brownian motion with Hurst index $H > 1/2$, such that B and β^H are independent. Note that m^ϵ is a standard real-valued Ornstein-Uhlenbeck process. Since $H > 1/2$, the evolution equation for X^ϵ, *i.e.* the first equation in (1), is understood in the sense of Young, see for instance [**10, 11**].

The objective of this article is to build and study the behavior of numerical schemes in the regime $\epsilon \to 0$. In that regime, Hairer and Li proved recently [**3**] that the following averaging principle holds (see Section 2.4 below): the slow component X^ϵ converges in probability to

1991 *Mathematics Subject Classification.* 60H35,60H10;60G22;65C30.

Key words and phrases. fractional Brownian motion; averaging principle; Euler schemes; asymptotic preserving schemes.

the solution $\overline{X}^{(H)}$ of the averaged equation

$$d\overline{X}^{(H)}(t) = \overline{g}(\overline{X}^{(H)}(t))d\beta^H(t), \tag{2}$$

with initial condition $\overline{X}^{(H)}(0) = x_0$, where the averaged coefficient is defined by $\overline{g}(x) = \mathbb{E}_{m\sim\mathcal{N}(0,1)}[\overline{g}(x,m)]$. See also **[2, 7]** for other recent contributions dealing with the behavior of multiscale systems driven by fractional Brownian motion. It is worth mentioning that the expression of the averaging principle differs from the case of SDEs driven by standard Wiener processes in two directions: first, if $H = 1/2$, the averaged equation is given by

$$d\overline{X}^{(1/2)}(t) = \big(\overline{g^2}(\overline{X}^{(1/2)}(t))\big)^{\frac{1}{2}} d\beta^{1/2}(t),$$

second, if $H = 1/2$, the convergence only holds in distribution in general. We refer for instance to **[4, 13]** and to **[12**, Chapter 17] for seminal references on the averaging principle for SDEs driven by standard Wiener processes.

In this article, we revisit the averaging principle from **[3]** with a numerical analysis point of view: precisely we consider the notion of asymptotic preserving schemes, see the recent contribution **[1]** for stochastic systems and references therein. We retrieve the same differences between the standard and fractional Brownian motion cases at the discrete time level. Let us now describe the scheme studied in this article: let Δt denote the time-step size, then the scheme is given by

$$\begin{cases} X^\epsilon_{n+1} = X^\epsilon_n + g(X^\epsilon_n, m^\epsilon_{n+1})\big(\beta^H(t_{n+1}) - \beta^H(t_n)\big) \\ m^\epsilon_{n+1} = e^{-\frac{\Delta t}{\epsilon}} m^\epsilon_n + \sqrt{1 - e^{-\frac{2\Delta t}{\epsilon}}}\gamma_n, \end{cases} \tag{3}$$

where $\big(\gamma_n\big)_{n\geqslant 0}$ are independent standard Gaussian random variables. Note that the Ornstein-Uhlenbeck process m^ϵ is discretized exactly in distribution: m^ϵ_n and $m^\epsilon(n\Delta t)$ are equal in distribution for all $n \geqslant 0$. The slow component X^ϵ is discretized using a standard Euler scheme (with an implicit treatment of the fast component). It is well-known that the scheme (3) is consistent with the system (1) when $\Delta t \to 0$, when the parameter ϵ is fixed, see for instance **[5, 6, 9]** for the study of the Euler schemes and variants for SDEs driven by fractional Brownian motion with Hurst index $H > 1/2$.

When the time-scale separation parameter vanishes, *i.e.* $\epsilon \to 0$, it is straightforward to check that for all $n \geqslant 1$, one has $m^\epsilon_n \to \gamma_n$, and $X^\epsilon_n \to X^0_n$, where the limiting scheme satisfies

$$X^0_{n+1} = X^0_n + g(X^0_n, \gamma_n)\big(\beta^H(t_{n+1}) - \beta^H(t_n)\big). \tag{4}$$

The main result of this article is the consistence of the limiting scheme (4) with the averaged equation (2), see Theorem 2.2 below for a rigorous statement: when $\Delta t \to 0$, X^0_N converges in probability to $\overline{X}(T)$, where $T = N\Delta t$. This result means that the scheme (3) is asymptotic preserving: the following diagram commutes when $\Delta t, \epsilon \to 0$

$$\begin{array}{ccc} X^\epsilon_N & \xrightarrow{\Delta t\to 0} & X^\epsilon(T) \\ \big\downarrow{\scriptstyle \epsilon\to 0} & & \big\downarrow{\scriptstyle \epsilon\to 0} \\ X^0_N & \xrightarrow{\Delta t\to 0} & \overline{X}(T). \end{array} \tag{5}$$

In practice, this property means that the time-step size Δt can be chosen independently of the time-scale separation parameter ϵ, and that the scheme is both consistent with the model

when ϵ is fixed and able to capture the limiting averaged equation. Asymptotic preserving schemes for SDEs driven by standard Wiener processes have been introduced and studied in the recent preprint [**1**]. To the best of our knowledge, they have not been studied in the case of SDEs driven by fractional Brownian motion. Our study reveals that the differences seen in the averaging principle at the continuous-time level also appear for the discretization: in the fractional Brownian motion case, the convergence holds in probability (instead of in distribution) and in the limiting equation the average of g (instead of g^2) is computed.

Observe that proposing an asymptotic preserving scheme in a stochastic context is not trivial. Like in [**1**], if the Ornstein-Uhlenbeck component was discretized using a implicit Euler scheme

$$m^\epsilon_{n+1} = \frac{1}{1+\frac{\Delta t}{\epsilon}}\big(m^\epsilon_n + \sqrt{\frac{\Delta t}{\epsilon}}\gamma_n\big),$$

the associated limiting scheme would be given by

$$X_{n+1} = X_n + g(X_n,0)\big(\beta^H(t_{n+1}) - \beta^H(t_n)\big)$$

instead of (4), which is not consistent with (2) in general: it is consistent only if $\overline{g}(x) = g(x,0)$ for all x. In addition, if the Ornstein-Uhlenbeck process $\big(m^\epsilon(t)\big)_{t\geqslant 0}$ is replaced by an arbitrary ergodic process, the averaging principle still holds (with an appropriate modification of the definition of $\overline{g}$), however there is no known construction of an asymptotic preserving scheme in this general case, even in the standard Brownian motion case, to the best of our knowledge.

The main result of this article is proved first for a simplified case, assuming that $g(x,m) = g(m)$ only depends on the variable m – in that case $\overline{g}$ is a constant. The proof in the simplified case is elementary, and it is provided in order to exhibit the main ideas and the main differences compared with the standard Brownian motion case. The main result is then proved in the general case, using more technical arguments.

This article is organized as follows. Section 2 presents notation and preliminary results: in particular assumptions concerning the multiscale SDE system are given in Section 2.3, the averaging principle from [**3**] is presented in Section 2.4, numerical schemes are presented in Section 2.5, and the main result, Theorem 2.2, is stated and discussed in Section 2.6. The simplified case ($g(x,m) = g(m)$) is studied in Section 3, in particular the comparison with the standard Brownian motion case is performed in Section 3.2. The general case is studied in Section 4. Section 5 is devoted to concluding remarks and perspectives for future works.

2. General setting

To simplify notation, in this article we consider real-valued processes. Up to straightforward modifications the results are generalized to higher dimension.

2.1. Notation. Let $H \in (\frac{1}{2},1)$, let $\big(\beta^H(t)\big)_{t\geqslant 0}$ be a real-valued fractional Brownian motion with Hurst index H, and let $\big(B(t)\big)_{t\geqslant 0}$ be a standard real-valued Brownian motion. It is assumed that β^H and B are independent. Let $\mathcal{F}^H = \sigma\{\beta^H(t);\ t\geqslant 0\}$ be the σ-field generated by the fractional Brownian motion β^H. The conditional expectation operator $\mathbb{E}[\cdot|\mathcal{F}^H]$ is denoted by $\mathbb{E}^H[\cdot]$ in the sequel.

The time-scale separation parameter is denoted by ϵ, without loss of generality $\epsilon \in (0,1)$. The slow variable is denoted by X^ϵ, whereas the fast variable is denoted by m^ϵ.

2.2. A criterion for convergence in probability. In this article, convergence of random variables is understood as convergence in probability – except in Section 3.2 where convergence in distribution needs to be considered.

Let us state an elementary criterion to express convergence in probability in terms of convergence of averages. We use the following convention throughout this article: a mapping $\varphi : \mathbb{R} \to \mathbb{R}$ is said to be of class $\mathcal{C}_b^K$ for some $K \in \mathbb{N}$ if it is bounded and K times continuously differentiable, and if its derivatives of order $1, \dots, K$ are bounded.

LEMMA 2.1. *Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space and $\mathcal{G} \subset \mathcal{F}$ be a σ-field. Let $K \in \mathbb{N}$.*

Let X be a $\mathcal{G}$-measurable real-valued random variable, and let $\big(X_N\big)_{N\in\mathbb{N}}$ be a sequence of real-valued random variables.

The following statements are equivalent.

(i) X_N converges to X in probability when $N \to \infty$: for all $\eta \in (0,1)$,

$$\mathbb{P}(|X_N - X| > \eta) \underset{N\to\infty}{\to} 0.$$

(ii) For any function $\varphi : \mathbb{R} \to \mathbb{R}$ of class $\mathcal{C}_b^K$, one has

$$\mathbb{E}\big[\big|\mathbb{E}[\varphi(X_N)|\mathcal{G}] - \varphi(X)\big|\big] \underset{N\to\infty}{\to} 0. \tag{6}$$

The proof of Lemma 2.1 is given in the appendix.

The expression of convergence in probability in the form of (6) is convenient for several reasons. First, it allows us to provide proofs using Taylor expansion arguments. Second, in the simplified case (g does not depend on x) it allows us to provide an elementary proof and a comparison with the standard Brownian case. Finally, expression (6) may be appropriate to exhibit a speed of convergence, however this question goes beyond the scope of this article and is left open for future works.

2.3. The multiscale stochastic system. In this article, we study multiscale stochastic systems of the type

$$\begin{cases} dX^\epsilon(t) = g(X^\epsilon(t), m^\epsilon(t))d\beta^H(t), \\ dm^\epsilon(t) = -\dfrac{1}{\epsilon}m^\epsilon(t)dt + \dfrac{\sqrt{2}}{\sqrt{\epsilon}}dB(t), \end{cases} \tag{7}$$

with initial conditions $X^\epsilon(0) = x_0$ and $m^\epsilon(0) = m_0$, which are assumed to be deterministic and independent of the parameter ϵ, for simplicity. Assume that $g : \mathbb{R}^2 \to \mathbb{R}$ is of class $\mathcal{C}^3$, with bounded derivatives of order $1, 2, 3$. Then the system (7) admits a unique solution, such that for all $t \geqslant 0$ one has

$$\begin{cases} X^\epsilon(t) = x_0 + \displaystyle\int_0^t g(X^\epsilon(s), m^\epsilon(s))d\beta^H(s), \\ m^\epsilon(t) = e^{-\frac{t}{\epsilon}}m_0 + \sqrt{2}\displaystyle\int_0^t e^{-\frac{t-s}{\epsilon}}dB(s). \end{cases}$$

Note that m^ϵ is an Ornstein-Uhlenbeck process, for any value of $\epsilon \in (0,1)$. The stochastic integral in the X^ϵ component is interpreted as a Young integral.

In Section 3, we study a simplified case, where $g(x,m) = g(m)$ for all $x, m \in \mathbb{R}^2$. The general case is studied in Section 4

2.4. The averaging principle. The goal of this section is to state the averaging principle result from [**3**].

Define the averaged coefficient $\overline{g} : \mathbb{R} \to \mathbb{R}$ as follows

$$\overline{g}(x) = \int g(x,m)d\nu(m), \quad x \in \mathbb{R},$$

where $\nu = \mathcal{N}(0,1)$ is the standard Gaussian distribution. Note that ν is the unique invariant distribution of the Ornstein-Uhlenbeck process m^ϵ, for all $\epsilon \in (0,1)$, and for every fixed $t \in (0,\infty)$ and $x \in \mathbb{R}$, one has

$$\mathbb{E}[g(x,m^\epsilon(t))] \underset{\epsilon\to 0}{\to} \overline{g}(x).$$

The mapping $\overline{g}$ inherits the regularity of the mapping g with respect to the x-variable: $\overline{g}$ is of class $\mathcal{C}^3$, with bounded derivatives of order $1,2,3$.

Let $\big(\overline{X}(t)\big)_{t\geqslant 0}$ be the unique solution of the averaged equation

$$d\overline{X}(t) = \overline{g}(\overline{X}(t))d\beta^H(t), \tag{8}$$

with initial condition $\overline{X}(0) = x_0$. One has for all $t \geqslant 0$

$$\overline{X}(t) = x_0 + \int_0^t \overline{g}(\overline{X}(s))d\beta^H(s),$$

where the stochastic integral is interpreted in the sense Young sense.

The averaging principle from [**3**] states that X^ϵ converges in probability to $\overline{X}$ when $\epsilon \to 0$. In this article, we consider a weaker version: for all $T \geqslant 0$, $X^\epsilon(T)$ converges in distribution to $\overline{X}(T)$. Owing to Lemma 2.1, one has the following formulation:

$$\lim_{\epsilon\to 0} \mathbb{E}\big[\big|\mathbb{E}^H[\varphi(X^\epsilon(T))] - \varphi(\overline{X}(T))\big|\big] = 0 \tag{9}$$

for every function φ of class $\mathcal{C}^3_b$.

2.5. Numerical schemes. Let us introduce the numerical scheme studied in this article. Let $T \in (0,\infty)$ and let Δt denote the time-step size. We assume that $T = N\Delta t$ where $N \in \mathbb{N}$ is an integer. Set $t_n = n\Delta t$ and $\gamma_n = \Delta t^{-1/2}\big(B(t_{n+1}) - B(t_n)\big)$ for all $n \geqslant 0$. For any values of the time-scale separation parameter ϵ and of the time-step size Δt, the numerical scheme is defined by the recursion

$$\begin{cases} X^\epsilon_{n+1} = X^\epsilon_n + g(X^\epsilon_n, m^\epsilon_{n+1})\big(\beta^H(t_{n+1}) - \beta^H(t_n)\big) \\ m^\epsilon_{n+1} = e^{-\frac{\Delta t}{\epsilon}} m^\epsilon_n + \sqrt{1 - e^{-\frac{2\Delta t}{\epsilon}}}\gamma_n, \end{cases} \tag{10}$$

with $x^\epsilon_0 = x_0$ and $m^\epsilon_0 = m_0$. To simplify notation, the convention $\delta\beta^H_n = \beta^H(t_{n+1}) - \beta^H(t_n)$ is used below.

When $\epsilon \to 0$, with fixed time-step size Δt, it is straightforward to prove that $X^\epsilon_n \to X^0_n$ (in probability), for all $n \geqslant 0$, where

$$X^0_{n+1} = X^0_n + g(X^0_n, \gamma_n)\big(\beta^H(t_{n+1}) - \beta^H(t_n)\big), \tag{11}$$

with $X^0_0 = x_0$.

Let us finally introduce the auxiliary scheme

$$\overline{X}_{n+1} = \overline{X}_n + \overline{g}(\overline{X}_n)\big(\beta^H(t_{n+1}) - \beta^H(t_n)\big), \tag{12}$$

for all $n \geqslant 0$, with $\overline{X}_0 = x_0$. Note that the auxiliary scheme (12) is the standard Euler scheme with time-step size Δt applied to the averaged equation (8).

2.6. Asymptotic preserving property. We are now in position to state the main result of this article.

THEOREM 2.2. *The scheme* (10) *is asymptotic preserving: the following diagram commutes*

$$\begin{array}{ccc} X_N^\epsilon & \xrightarrow{\Delta t \to 0} & X^\epsilon(T) \\ \Big\downarrow{\scriptstyle \epsilon \to 0} & & \Big\downarrow{\scriptstyle \epsilon \to 0} \\ X_N^0 & \xrightarrow{\Delta t \to 0} & \overline{X}(T) \end{array} \tag{13}$$

where convergence is understood as convergence in probability, and $T = N\Delta t$, with arbitrary fixed $T \in (0, \infty)$.

The asymptotic preserving property can be rewritten as follows: for any real-valued mapping φ of class $\mathcal{C}_b^3$, one has

$$\lim_{\Delta t \to 0} \lim_{\epsilon \to 0} \mathbb{E}\Big[\big|\mathbb{E}^H[\varphi(X_N^\epsilon)] - \varphi(\overline{X}(T))\big|\Big] = \lim_{\epsilon \to 0} \lim_{\Delta t \to 0} \mathbb{E}\Big[\big|\mathbb{E}^H[\varphi(X_N^\epsilon)] - \varphi(\overline{X}(T))\big|\Big] = 0. \tag{14}$$

The reformulation (14) is due to the criterion for convergence in probability given by Lemma 2.1. In order to prove Theorem 2.2, we only need to prove that the limiting scheme is consistent with the averaged equation, *i.e.* that the following result holds.

PROPOSITION 2.3. *Let $T \in (0, \infty)$, and let the time-step size Δt satisfy $T = N\Delta t$, with $N \in \mathbb{N}$.*

Let $\left(X_n^0\right)_{n \geqslant 0}$ be given by the limiting scheme (11)*, and let $\left(\overline{X}_n\right)_{n \geqslant 0}$ be given by the auxiliary scheme* (12)*, with $X_0^0 = \overline{X}_0 = x_0$.*

For any real-valued mapping φ of class $\mathcal{C}_b^3$, one has

$$\lim_{\Delta t \to 0} \mathbb{E}\Big[\big|\mathbb{E}^H[\varphi(X_N^0)] - \varphi(\overline{X}_N)\big|\Big] = 0.$$

REMARK 2.4. *In the simplified case (Section 3), it is sufficient to assume that the functions φ are of class $\mathcal{C}_b^2$.*

Let us provide the proof of Theorem 2.2, assuming that Proposition 2.3 holds.

PROOF OF THEOREM 2.2. On the one hand, for fixed Δt, X_N^ϵ converges in probability to X_N^0 when $\epsilon \to 0$, by construction of the scheme. Note that the auxiliary scheme (12) is consistent with the averaged equation, see for instance **[5, 6, 9]**: when $\Delta t \to 0$, $\overline{X}_N$ converges in probability to $\overline{X}(T)$. Owing to Proposition 2.3 and to Lemma 2.1, we deduce that X_N^0 converges to $\overline{X}(T)$ in probability when $\Delta t \to 0$.

On the other hand, for fixed ϵ, the scheme (10) is consistent with (7) when $\Delta t \to 0$, in particular X_N^ϵ converges in probability to $X^\epsilon(T)$, see for instance **[5, 6, 9]**. Owing to the averaging principle, one has (9), which means that $X^\epsilon(T)$ converges to $\overline{X}(T)$ in probability when $\epsilon \to 0$.

We thus conclude that the diagram (13) commutes, where convergence is understood as convergence in probability, thus the scheme is asymptotic preserving. □

It only remains to prove Proposition 2.3. The proof is given first in the simplified case (g does not depend on x) in Section 3, then in the general case in Section 4. The proof in the simplified case is elementary and is given for pedagogical reasons and to illustrate the differences with the standard Brownian motion case. The analysis of the general case requires more technical arguments.

3. Study of the simplified problem

In this section, we assume that the mapping g only depends on m: one has $g(x,m) = g(m)$ for all $(x,m) \in \mathbb{R}^2$. Then the averaged quantity $\overline{g}$ is a constant:

$$\overline{g} = \int g(m) d\nu(m).$$

In addition, note that one has $\overline{X}_n = \overline{X}(t_n) = x_0 + \overline{g}\beta^H(t_n)$ for all $n \geqslant 0$.

REMARK 3.1. *In the simplified case, it is sufficient to assume that g is globally Lipschitz continuous.*

Below, first we provide the proof of Proposition 2.3 in this case, second we provide a comparison with the case where the fractional Brownian motion β^H with $H > 1/2$ is replaced by a standard Brownian motion β. We illustrate the two main fundamental differences: in the latter case the convergence is understood as convergence in distribution, and the averaged equation is not given by averaging g – one needs to average g^2.

The standard Brownian motion case is already well-understood (see for instance [**12**, Chapter 17] for the averaging principle and [**1**] for the design and analysis of asymptotic preserving schemes), however we provide details for pedagocial reasons – and the presentation differs from [**1**].

3.1. Proof of Proposition 2.3 in the simplified case.

PROOF OF PROPOSITION 2.3 IN THE SIMPLIFIED CASE. Let us introduce a family of auxiliary random variables: for all $n \in \{0, \dots, N\}$, set

$$X_N^{(n)} = x_0 + \sum_{k=0}^{n-1} \overline{g}\delta\beta_k^H + \sum_{k=n}^{N-1} g(\gamma_k)\delta\beta_k^H.$$

Note that by construction, one has $X_N^0 = X_N^{(0)}$ and $\overline{X}_N = X_N^{(N)}$. In addition, for all $n \in \{0, \dots, N-1\}$, set $S_N^{(n)} = x_0 + \sum_{k=0}^{n-1} \overline{g}\delta\beta_k^H + \sum_{k=n+1}^{N-1} g(\gamma_k)\delta\beta_k^H$. Then one has

$$\begin{aligned} X_N^{(n)} &= S_N^{(n)} + g(\gamma_n)\delta\beta_n^H \\ X_N^{(n+1)} &= S_N^{(n)} + \overline{g}\delta\beta_n^H. \end{aligned}$$

for all $n \in \{0, \dots, N-1\}$.

Let φ be of class $\mathcal{C}_b^2$ (we take $K = 2$ when applying Lemma 2.1 in the simplified case). Observe that $\overline{X}_N$ is $\mathcal{F}^H$-measurable, thus one has $\mathbb{E}^H[\varphi(\overline{X}_N)] = \varphi(\overline{X}_N)$. Using a telescoping

sum argument and a second-order Taylor expansion, one obtains

$$\begin{aligned}\mathbb{E}^H[\varphi(X_N^0)] - \varphi(\overline{X}_N) &= \mathbb{E}^H[\varphi(X_N^{(0)})] - \mathbb{E}^H[\varphi(X_N^{(N)})] \\ &= \sum_{n=0}^{N-1} \big(\mathbb{E}^H[\varphi(X_N^{(n)})] - \mathbb{E}^H[\varphi(X_N^{(n+1)})]\big) \\ &= \sum_{n=0}^{N-1} \big(\mathbb{E}^H[\varphi(S_N^{(n)} + g(\gamma_n)\delta\beta_n^H)] - \mathbb{E}^H[\varphi(S_N^{(n)} + \overline{g}\delta\beta_n^H)]\big) \\ &= \sum_{n=0}^{N-1} \mathbb{E}^H[\varphi'(S_N^{(n)})(g(\gamma_n) - \overline{g})]\delta\beta_n^H + \sum_{n=0}^{N-1} \mathrm{O}(|\delta\beta_n^H|^2).\end{aligned}$$

On the one hand, since the random variables $\big(\gamma_n\big)_{0\leqslant n\leqslant N-1}$ are independent, and are independent of β^H, one has

$$\mathbb{E}^H[\varphi'(S_N^{(n)})(g(\gamma_n) - \overline{g})] = \mathbb{E}^H[\varphi'(S_N^{(n)})]\mathbb{E}[g(\gamma_n) - \overline{g}] = 0$$

by definition of $\overline{g}$. Indeed, observe that $S_N^{(n)}$ only depends on γ_k with $k \neq n$, and on β^H.

On the other hand, since the Hurst index satisfies $H > 1/2$, one has

$$\sum_{n=0}^{N-1} \mathbb{E}[|\delta\beta_n^H|^2] = \mathrm{O}\big(\Delta t^{2H-1}\big) \underset{\Delta t\to 0}{\to} 0.$$

Gathering the estimates then gives the required convergence result: for all functions φ of class $\mathcal{C}_b^2$, one has

$$\mathbb{E}\big[\big|\mathbb{E}^H[\varphi(X_N^0)] - \varphi(\overline{X}_N)\big|\big] \underset{\Delta t\to 0}{\to} 0$$

which concludes the proof of Proposition 2.3 in the simplified case. □

REMARK 3.2. *The proof above provides a rate of convergence $2H-1$, which is consistent with the rate of convergence of the standard Euler scheme for SDEs driven by fractional Brownian motion.*

3.2. Comparison with the standard Brownian Motion case. The objective of this section is to provide a comparison with the situation where the fractional Brownian motion β^H is replaced by a standard Brownian motion β (independent of B). We thus consider the system

$$\begin{cases} dX^\epsilon(t) = g(m^\epsilon(t))d\beta(t), \\ dm^\epsilon(t) = -\dfrac{1}{\epsilon}m^\epsilon(t)dt + \dfrac{\sqrt{2}}{\sqrt{\epsilon}}dB(t). \end{cases} \tag{15}$$

The associated numerical scheme is defined by

$$\begin{cases} X_{n+1}^\epsilon = X_n^\epsilon + g(m_{n+1}^\epsilon)\big(\beta(t_{n+1}) - \beta(t_n)\big) \\ m_{n+1}^\epsilon = e^{-\frac{\Delta t}{\epsilon}}m_n^\epsilon + \sqrt{1 - e^{-\frac{2\Delta t}{\epsilon}}}\gamma_n, \end{cases} \tag{16}$$

with initial conditions $X^\epsilon(0) = X_0^\epsilon = x_0$ and $m^\epsilon(0) = m_0^\epsilon = m_0$.

On the one hand, in that setting the averaging principle holds as follows: X^ϵ converges in distribution to $\overline{X}$ defined by

$$\overline{X}(t) = x_0 + \big(\overline{g^2}\big)^{\frac{1}{2}}\beta(t)$$

with $\overline{g^2} = \int g^2 d\nu$. Note that in general $\overline{g^2} > \overline{g}^2$.

On the other hand, for fixed $\Delta t > 0$, one has the convergence result (in probability) $X_n^\epsilon \to X_n^0$ for all $n \in \{0, \dots, N\}$, where

$$X_{n+1}^0 = X_n^0 + g(\gamma_n)\big(\beta(t_{n+1}) - \beta(t_n)\big).$$

The scheme (16) is asymptotic preserving, when convergence is understood in distribution.

PROPOSITION 3.3. *The limiting scheme is consistent, for convergence in distribution, with the averaged equation. More precisely, let $T \in (0, \infty)$ and let φ be of class $\mathcal{C}_b^3$. Then*

$$\lim_{\Delta t \to 0} \lim_{\epsilon \to 0} \mathbb{E}[\varphi(X_N^\epsilon)] = \lim_{\epsilon \to 0} \lim_{\Delta t \to 0} \mathbb{E}[\varphi(X_N^\epsilon)] = \mathbb{E}[\varphi(\overline{X}(T))].$$

The result above is an immediate consequence of [**1**, Theorem 3.7]. However, for pedagogical reasons, we provide a proof of the consistency of the limiting scheme with the averaged equation, using the same approach as in the proof of Proposition 2.3 above in the simplified case. This allows us to give a comparison of the standard and fractional Brownian motion cases.

PROOF. As in the proof of Proposition 2.3 (in the simplified case) above, introduce the following family of random variables: for all $n \in \{0, \dots, N\}$ define

$$X_N^{(n)} = x_0 + \sum_{k=0}^{n-1} \big(\overline{g^2}\big)^{\frac{1}{2}} \delta\beta_k + \sum_{k=n}^{N-1} g(\gamma_k)\delta\beta_k.$$

Note that by construction, one has $X_N^0 = X_N^{(0)}$ and $\overline{X}(t_N) = X_N^{(N)}$. In addition, for all $n \in \{0, \dots, N-1\}$, set $S_N^{(n)} = x_0 + \sum_{k=0}^{n-1} \big(\overline{g^2}\big)^{\frac{1}{2}} \delta\beta_k + \sum_{k=n+1}^{N-1} g(\gamma_k)\delta\beta_k$. Then one has

$$\begin{aligned} X_N^{(n)} &= S_N^{(n)} + g(\gamma_n)\delta\beta_n \\ X_N^{(n+1)} &= S_N^{(n)} + \big(\overline{g^2}\big)^{\frac{1}{2}} \delta\beta_n. \end{aligned}$$

Let φ be of class $\mathcal{C}_b^3$. Using a telescoping sum argument and a third-order Taylor expansion, one obtains

$$\begin{aligned}
\mathbb{E}^H[\varphi(X_N^0)] - \varphi(\overline{X}(t_N)) &= \mathbb{E}^H[\varphi(X_N^{(0)})] - \mathbb{E}^H[\varphi(X_N^{(N)})] \\
&= \sum_{n=0}^{N-1} \big(\mathbb{E}^H[\varphi(X_N^{(n)})] - \mathbb{E}^H[\varphi(X_N^{(n+1)})]\big) \\
&= \sum_{n=0}^{N-1} \big(\mathbb{E}^H[\varphi(S_N^{(n)} + g(\gamma_n)\delta\beta_n)] - \mathbb{E}^H[\varphi(S_N^{(n)} + \big(\overline{g^2}\big)^{\frac{1}{2}}\delta\beta_n)]\big) \\
&= \sum_{n=0}^{N-1} \mathbb{E}^H[\varphi'(S_N^{(n)})\big(g(\gamma_n) - \big(\overline{g^2}\big)^{\frac{1}{2}}\big)\delta\beta_n] \\
&\quad + \frac{1}{2}\sum_{n=0}^{N-1} \mathbb{E}^H[\varphi'(S_N^{(n)})(g^2(\gamma_n) - \overline{g^2})\delta\beta_n^2] \\
&\quad + \sum_{n=0}^{N-1} \mathrm{O}(|\delta\beta_n|^3).
\end{aligned}$$

The first order term vanishes in expectation, since the increments of the standard Brownian motion β are independent. Indeed, observe that the random variables $S_N^{(n)}$, γ_n and $\delta\beta_n$ are independent: for all $n \in \{0, \dots, N-1\}$

$$\mathbb{E}\big[\mathbb{E}^H[\varphi'(S_N^{(n)})(g(\gamma_n) - \big(\overline{g^2}\big)^{\frac{1}{2}})\delta\beta_n]\big] = \mathbb{E}[\varphi'(S_N^{(n)})]\mathbb{E}[g(\gamma_n) - \big(\overline{g^2}\big)^{\frac{1}{2}}]\mathbb{E}[\delta\beta_n] = 0,$$

using $\mathbb{E}[\delta\beta_n] = 0$.

The second order term vanishes almost surely by the definition of the averaged coefficient $\overline{g^2}$: for all $n \in \{0, \dots, N-1\}$, using the independence of the increments of the standard Brownian motion β one has

$$\mathbb{E}^H[\varphi'(S_N^{(n)})(g^2(\gamma_n) - \overline{g^2})\delta\beta_n^2] = \mathbb{E}^H[\varphi'(S_N^{(n)})]\mathbb{E}[g^2(\gamma_n) - \overline{g^2}]\Delta t = 0.$$

Finally, the last term satisfies $\sum_{n=0}^{N-1} \mathbb{E}|\delta\beta_n|^3 = \mathrm{O}\big(\Delta t^{\frac{1}{2}}\big) \underset{\Delta t \to 0}{\to} 0$.

Gathering the estimates gives the required convergence result

$$\lim_{\Delta t \to 0} \big|\mathbb{E}[\varphi(X_N^0)] - \mathbb{E}[\varphi(\overline{X}(T))]\big| = 0$$

and concludes the proof. □

In the proof above, we can exhibit the two main differences between the fractional and standard Brownian motion case. In the latter case, the first order terms of the Taylor expansion only vanishes in expectation, hence the need to consider convergence in distribution. In addition, the averaging procedure is only visible in the second order terms of the Taylor expansion, hence a different expression of the averaged coefficient, whereas it is visible in the first order term in the fractional Brownian motion case.

4. Study of the general case

The analysis of the general case requires more involved techniques, compared with the simplified case. We first state the useful auxiliary results in Section 4.1, and give the proof

of Proposition 2.3 in Section 4.2. The proof of an auxiliary result stated in Section 4.1 is given in Section 4.3.

4.1. Auxiliary results. Let us state the main auxiliary results which are used in Section 4.2 below to prove Proposition 2.3.

Let $T \in (0,\infty)$ be fixed. For all $\alpha \in (0,1)$, if $f : [0,T] \to \mathbb{R}$ is a real-valued function, set

$$\|f\|_\alpha = \sup_{0 \leqslant t_1 < t_2 \leqslant T} \frac{|f(t_2) - f(t_1)|}{|t_2 - t_1|}.$$

Recall that f is α-Hölder continuous if $\|f\|_\alpha < \infty$.

First, the trajectories of a fractional Brownian motion $\big(\beta^H(t)\big)_{0 \leqslant t \leqslant T}$ with Hurst index H are α-Hölder continuous, for all $\alpha \in (0,H)$. More precisely, for all $\alpha \in (0,H)$, there exists an almost surely finite random variable C_α, such that

$$\|\beta^H\|_\alpha \leqslant C_\alpha, \tag{17}$$

moreover $\mathbb{E}[C_\alpha^m] < \infty$ for all $m \in \mathbb{N}$. The property (17) is a consequence of the Kolmogorov regularity criterion, using the equality $\mathbb{E}[|\beta^H(t_2) - \beta^H(t_1)|^2] = |t_2 - t_1|^{2H}$ and the fact that β^H is a Gaussian process.

Next, in order to study discrete-time processes, it is convenient to introduce the following variant of the Hölder semi-norms. Let Δt denote the time-step size, with the condition $T = N\Delta t$ for some $N \in \mathbb{N}$. For every $t \in [0,T]$, let $\ell(t) \in [0,T]$ be defined by $\ell(t) = n\Delta t$ for all $t \in [n\Delta t, (n+1)\Delta t)$, $n \in \{0,\dots,N-1\}$ and $\ell(T) = T$. To simplify notation, we omit the dependence of ℓ with respect to Δt. For all $\alpha \in (0,1)$ and $0 \leqslant a \leqslant b \leqslant T$, if $f : [0,T] \to \mathbb{R}$ is a real-valued function, set

$$\|f\|_{a,b,\alpha,\Delta t} = \sup_{a \leqslant t_1 < t_2 \leqslant b ; \ell(t_1) = t_1} \frac{|f(t_2) - f(t_1)|}{|t_2 - t_1|}.$$

Observe that the only change in the definition is the condition $\ell(t_1) = t_1$. The dependence with respect to the left and right hand points of the interval is also made explicit (we only need $a = 0$ and $b = T$ for the standard version $\|\cdot\|_\alpha$).

In the sequel, we employ the following result to estimate Young integrals.

LEMMA 4.1. *Let $z = \big(z(t)\big)_{0 \leqslant t \leqslant T}$ be a α-Hölder continuous real-valued function with $\alpha \in (0,1)$. Let $M \in \mathbb{N}$, and let $F : \mathbb{R}^M \to \mathbb{R}$ be a continuously differentiable function.*

Assume that real-valued mappings $y_1,\dots,y_M$ satisfy $\|y_m\|_{\alpha',\Delta t} < \infty$ for all $m = 1,\dots,M$, for some $\alpha' \in (0,1)$ such that $\alpha + \alpha' > 1$.

Then there exists $C \in (0,\infty)$, which does not depend on Δt, such that the following holds: for all $s,t \in [0,T]$ such that $\ell(s) = s$, one has

$$\begin{aligned} \Big| \int_s^t F(y(\ell(r))dz(r) \Big| &\leqslant C \sup_{s \leqslant r \leqslant t} |F(y(\ell(r))| \|z\|_\alpha (t-s)^\alpha \\ &\quad + \sum_{m=1}^M \sup_{s \leqslant r \leqslant t} |\partial_{y_m} F(y(\ell(r))| \|y_m\|_{s,t,\alpha',\Delta t} \|z\|_\alpha (t-s)^{\alpha+\alpha'}, \end{aligned} \tag{18}$$

where to simplify notation $F(y(\ell(r)) = F(y_1(\ell(r)),\dots,y_M(\ell(r)))$.

Lemma 4.1 is a variant of [**6**, Lemma A.1], where the dependence with respect to $\|y_m\|_{s,t\alpha',\Delta t}$ when m varies is made more explicit. This is instrumental in the proof of Lemma 4.2 below, where bounds for derivatives are proved succesively. We refer to [**6**] for the proof. Note that the standard inequalities for Young integrals, considering the standard Hölder semi-norm $\|\cdot\|_{\alpha'}$ cannot be applied, since $r \mapsto y(\ell(r))$ is piecewise constant, and thus in general is not (Hölder) continuous.

Let us now introduce a generalized version of the auxiliary scheme (12): for all $n \in \{0,\dots,N\}$, $k \in \{n,\dots,N-1\}$ and $x \in \mathbb{R}$, set

$$\begin{aligned} \overline{X}_{n,k+1}(x) &= \overline{X}_{n,k}(x) + \overline{g}(\overline{X}_{n,k}(x))\delta\beta_k^H \\ \overline{X}_{n,n}(x) &= x. \end{aligned} \tag{19}$$

The definition above is indeed a generalization of (12): one has $\overline{X}_n = \overline{X}_{0,n}(x_0)$ for all $n \in \{0,\dots,N\}$.

Finally, let us introduce the auxiliary functions u_n, for $n \in \{0,\dots,N\}$, as follows. Given a real-valued mapping $\varphi : \mathbb{R} \to \mathbb{R}$ of class $\mathcal{C}_b^3$, define

$$u_n(x) = \varphi(\overline{X}_{n,N}(x)) \tag{20}$$

for all $n \in \{0,\dots,N\}$ and $x \in \mathbb{R}$. To simplify notation, the dependence of u_n with respect to the time-step size Δt is omitted. Note that $u_N = \varphi$. Observe that the u_n's are random functions. They satisfy the following property: for all $n \in \{0,\dots,N-1\}$ and all $x \in \mathbb{R}$, one has

$$u_n(x) = u_{n+1}\big(x + \overline{g}(x)\delta\beta_n^H\big). \tag{21}$$

Indeed, by construction $\overline{X}_{n+1,N}\big(x + \overline{g}(x)\delta\beta_n^H\big) = \overline{X}_{n,N}(x)$.

We are now in position to state the main auxiliary result of this section.

LEMMA 4.2. *Assume that φ is of class $\mathcal{C}_b^3$. There exist an almost surely finite random variable Λ, and an almost surely positive random variable $\Delta t_0 \leqslant T$, such that for all $\Delta t \in (0, \Delta t_0)$ one has*

$$\sup_{0\leqslant n\leqslant N} \sup_{x\in\mathbb{R}} \Big(|u_n(x)| + |u_n'(x)| + |u_n''(x)|\Big) \leqslant \Lambda. \tag{22}$$

The proof of Lemma 4.2 is technical and is postponed to Section 4.3. The arguments are similar to those used in [**6**] to prove boundedness of the solutions of Euler or Milstein schemes applied to SDEs driven by fractional Brownian motion. Indeed, the proof consists in first expressing the first and second order derivatives of u_n using solutions $\eta_{n,N}$ and $\zeta_{n,N}$ of variation equations (see equations (24) and (25) below), second in proving appropriate bounds.

REMARK 4.3. *In the standard Brownian motion case (Section 3.2), the role of u_n would be played by a function defined as $\mathbb{E}[\varphi(\overline{X}_{n,N}(x))]$. The Markov property would play a key role in the analysis, and by homogeneity it would be sufficient to look at the properties of the mapping $\mathbb{E}[\varphi(\overline{X}_{0,n}(x))]$.*

In the fractional Brownian motion case, the Markov property is not satisfied, and taking expectation is not relevant. Instead of the Markov property, the flow property (21) *is satisfied. Observe that the auxiliary functions u_n need to be random. In addition, u_n needs to be defined in terms of $\overline{X}_{n,N}$, instead of $\overline{X}_{0,N-n}$, since the process is not time homogeneous.*

Even if the increments of the fractional Brownian motion $\big(\delta\beta_n^H\big)_{0\leqslant n\leqslant N-1}$ are not independent, the definition (20) *of the random functions u_n above makes sense. In the property* (21)*, this lack of independence results in the following property: u_{n+1} is not independent of $\delta\beta_n^H$.*

4.2. Proof of Proposition 2.3 in the general case. The objective of this section is to provide the proof of Proposition 2.3. The guideline of the proof is similar to the approach in the simplified case, except for the first step:

- the error is written in terms of the functions u_n defined by (20) with a telescoping sum argument
- then the property (21) and a second order Taylor expansion argument are used
- the first order term vanishes by definition of the averaged coefficient $\overline{g}$
- the second order term is handled using the boundedness of u_n, uniformly in Δt, obtained in Lemma 4.2, using the condition $H > 1/2$ for the Hurst index.

Proof of Proposition 2.3 in the general case. Let φ be of class $\mathcal{C}_b^3$, and let u_n be defined by (20), for $n = 0, \dots, N$. Recall that $\mathbb{E}^H[\cdot]$ denotes the conditional expectation with respect to the σ-field $\mathcal{F}^H$ generated by the fractional Brownian motion β^H. On the one hand, by definition of u_0 one has $\varphi(\overline{X}_N) = u_0(x_0) = \mathbb{E}^H[u_0(x_0)]$, where $x_0 = \overline{X}_0 = X_0^0$ (u_0 is a $\mathcal{F}^H$-measurable random variable). On the other hand, one has $\varphi(X_N^0) = u_N(X_N^0)$. Using a telescoping sum argument, one then obtains

$$\begin{aligned}\mathbb{E}^H[\varphi(X_N^0)] - \varphi(\overline{X}_N) &= \mathbb{E}^H[u_N(X_N^0)] - \mathbb{E}^H[u_0(X_0^0)] \\ &= \sum_{n=0}^{N-1} \big(\mathbb{E}^H[u_{n+1}(X_{n+1}^0)] - \mathbb{E}^H[u_n(X_n^0)]\big) \\ &= \sum_{n=0}^{N-1} \mathbb{E}^H\big[u_{n+1}\big(X_n^0 + g(X_n^0,\gamma_n)\delta\beta_n^H\big) - u_{n+1}\big(X_n^0 + \overline{g}(X_n^0)\delta\beta_n^H\big)\big],\end{aligned}$$

where in the last line we have used the definition (11) of the limiting scheme, and the property (21) of the functions u_n.

A Taylor expansion argument then gives, for all $n \in \{0, \dots, N-1\}$

$$\begin{aligned}&\mathbb{E}^H\big[u_{n+1}\big(X_n^0 + g(X_n^0,\gamma_n)\delta\beta_n^H\big) - u_{n+1}\big(X_n^0 + \overline{g}(X_n^0)\delta\beta_n^H\big)\big] \\ &\qquad = \mathbb{E}^H[u'_{n+1}(X_n^0)\big(g(X_n^0,\gamma_n) - \overline{g}(X_n^0)\big)]\delta\beta_n^H + R_n,\end{aligned}$$

with $|R_n| \leqslant C \sup|u''_{n+1}(\cdot)|(\delta\beta_n^H)^2$, since g is assumed to be bounded.

Assume that Δt satisfies the condition $\Delta t \leqslant \Delta t_0$ where the random variable Δt_0 is given in Lemma 4.2. On the one hand, using the Hölder continuity property (17) of the fractional Brownian motion, with $\alpha \in (\frac{1}{2}, H)$ and the bound for the second order derivative of u_n stated in Lemma 4.2, there exists an almost surely finite constant $\mathbf{\Lambda}$, which does not depend on Δt, such that

$$|R_n| \leqslant \mathbf{\Lambda}\Delta t^{2\alpha}$$

for all $n \in \{0, \dots, N-1\}$, if $\Delta t \leqslant \Delta t_0$.

On the other hand, using the definition of $\overline{g}$ and conditioning with respect to X_n^0, one obtains the key property

$$\mathbb{E}^H[u'_{n+1}(X_n^0)\big(g(X_n^0,\gamma_n) - \overline{g}(X_n^0)\big)] = 0,$$

which means that the first-order term vanishes for all $n \in \{0, \dots, N-1\}$.

Finally, one obtains

$$\left|\mathbb{E}^H[\varphi(X_N^0)] - \varphi(\overline{X}_N)\right| \leqslant \sum_{n=0}^{N-1} |R_n| \leqslant \mathbf{\Lambda} T \Delta t^{2\alpha-1} \underset{\Delta t \to 0}{\to} 0,$$

almost surely, since α is chosen such that $2\alpha > 1$.

Since φ is assumed to be bounded, by the dominated convergence theorem the almost sure convergence implies

$$\mathbb{E}\big[\left|\mathbb{E}^H[\varphi(X_N^0)] - \varphi(\overline{X}_N)\right|\big] \underset{\Delta t \to 0}{\to} 0.$$

This holds for all functions φ of class $\mathcal{C}_b^3$.

This concludes the proof of Proposition 2.3 in the general case. □

REMARK 4.4. *The proof above also handles the simplified case treated in Section 3, with a slightly different point of view. In that case $u_n(x) = \varphi(x + \overline{g}\delta\beta_n^H)$, and the proof of Lemma 4.2 is straightforward in this simplified case.*

4.3. Proof of Lemma 4.2. It remains to give the proof of Lemma 4.2, which first requires to introduce additional notation. By the definition (20) of the functions u_n, one has the following expressions for $u_n'(x)$ and $u_n''(x)$:

$$\begin{aligned} u_n'(x) &= \varphi'(\overline{X}_{n,N}(x))\eta_{n,N}(x) \\ u_n''(x) &= \varphi'(\overline{X}_{n,N}(x))\zeta_{n,N}(x) + \varphi''(x)(\overline{X}_{n,N}(x))(\eta_{n,N}(x))^2, \end{aligned} \tag{23}$$

where for all $n \in \{0, \dots, N-1\}$, $k \in \{n, \dots, N-1\}$ and $x \in \mathbb{R}$, one has

$$\eta_{n,k+1}(x) = \eta_{n,k}(x) + \overline{g}'(\overline{X}_{n,N}(x))\eta_{n,k}(x)\delta\beta_k^H \tag{24}$$

and

$$\zeta_{n,k+1}(x) = \zeta_{n,k}(x) + \overline{g}'(\overline{X}_{n,N}(x))\zeta_{n,k}(x)\delta\beta_k^H + \overline{g}''(\overline{X}_{n,N}(x))(\eta_{n,k}(x))^2\delta\beta_k^H \tag{25}$$

with initial conditions $\eta_{n,n}(x) = 1$ and $\zeta_{n,n}(x) = 0$.

The expressions in (23) are obtained by recursion arguments. In the sequel, to simplify notation, we let $\overline{X}_{n,k} = \overline{X}_{n,k}(x)$, $\eta_{n,k} = \eta_{n,k}(x)$ and $\zeta_{n,k} = \zeta_{n,k}(x)$.

Let us introduce auxiliary continuous-time processes $\tilde{X}_n, \tilde{\eta}_n$ and $\tilde{\zeta}_n$, defined on the interval $[t_n, T]$ (with $t_n = n\Delta t$), for all $n \in \{0, \dots, N-1\}$: for all $k \in \{n, \dots, N-1\}$, if $t \in [t_k, t_{k+1})$, then

$$\begin{aligned} \tilde{X}_n(t) &= \overline{X}_{n,k} + \overline{g}(\overline{X}_{n,k}(\beta^H(t) - \beta_H(t_k)) \\ \tilde{\eta}_n(t) &= \eta_{n,k} + \overline{g}'(\overline{X}_{n,k})\eta_{n,k}(\beta^H(t) - \beta^H(t_k)) \\ \tilde{\zeta}_n(t) &= \zeta_{n,k} + \overline{g}'(\overline{X}_{n,k})\zeta_{n,k}(\beta^H(t) - \beta^H(t_k)) + \overline{g}''(\overline{X}_{n,k})(\eta_{n,k})^2(\beta^H(t) - \beta^H(t_k)). \end{aligned}$$

One has $\tilde{X}_n(t_k) = \overline{X}_{n,k}$, $\tilde{\eta}_n(t_k) = \eta_{n,k}$ and $\tilde{\zeta}_n(t_k) = \zeta_{n,k}$ for all $n \leqslant k \leqslant N$.

We are now in position to prove Lemma 4.2.

PROOF OF LEMMA 4.2. The proof is divided into three steps, where bounds for $\tilde{X}_n$, $\tilde{\eta}_n$ and $\tilde{\zeta}_n$ are proved successively.

Step 1. The auxiliary process $\tilde{X}_n$ satisfies the following property: for all $t_n \leqslant s \leqslant t \leqslant T$, such that $s = \ell(s)$, one has

$$\tilde{X}_n(t) - \tilde{X}_n(s) = \int_s^t F_0(\tilde{X}_n(\ell(r)))d\beta^H(r),$$

where $F_0 = \overline{g}$.

Using Lemma 4.1 with $\alpha = \alpha' \in (\frac{1}{2}, H)$, and using the boundedness of $\overline{g}$ and $\overline{g}'$, one obtains the inequality

$$\|\tilde{X}_n\|_{s,t,\alpha,\Delta t} \leqslant C_0(g, \|\beta^H\|_\alpha)\Big(1 + (t-s)^\alpha \|\tilde{X}_n\|_{s,t,\alpha,\Delta t}\Big).$$

Let τ_0 be a positive random variable, chosen such that

$$C_0(g, \|\beta^H\|_\alpha)\tau_0^\alpha \leqslant \frac{1}{2},$$

and $\tau_0 \leqslant T$.

Then if $s = \ell(s) \leqslant t$ satisfy $t - s \leqslant \tau_0$, one obtains

$$\|\tilde{X}_n\|_{s,t,\alpha,\Delta t} \leqslant 2C_0(g, \|\beta^H\|_\alpha),$$

which gives

$$|\tilde{X}_n(t)| \leqslant |\tilde{X}_n(s)| + 2\tau_0^\alpha C_0(g, \|\beta^H\|_\alpha).$$

If the time step size Δt satisfies $\Delta t \leqslant \tau_0$, one thus obtains

$$|\overline{X}_{n,k_2}| \leqslant |\overline{X}_{n,k_1}| + 2\tau_0^\alpha C_0(g, \|\beta^H\|_\alpha)$$

for integers $k_1 < k_2$ such that $(k_2 - k_1)\Delta t \leqslant \tau_0$, and iterating the argument and using the condition $\overline{X}_{n,n} = x$, one obtains

$$|\overline{X}_{n,k}| \leqslant |x| + 2N_0\tau_0^\alpha C_0(g, \|\beta^H\|_\alpha)$$

where the integer N_0 is chosen such that $N_0\tau_0 \geqslant T$, for all $k \in \{n, \dots, N\}$, if $\Delta t \leqslant \tau_0$.

Step 2. The auxiliary process $\tilde{\eta}_n$ satisfies the following property: for all $t_n \leqslant s \leqslant t \leqslant T$, such that $s = \ell(s)$, one has

$$\tilde{\eta}_n(t) - \tilde{\eta}_n(s) = \int_s^t F_1(\tilde{X}_n(\ell(r)), \tilde{\eta}_n(\ell(r)))d\beta^H(r),$$

where $F_1(x, \eta) = \overline{g}'(x)\eta$. Note that $|F_1(x,\eta)| + |\partial_x F_1(x,\eta)| \leqslant C|\eta|$ and $|\partial_\eta F_1(x,\eta)| \leqslant C$.

Using Lemma 4.1 with $\alpha = \alpha' \in (\frac{1}{2}, H)$, and using the boundedness of $\overline{g}$, $\overline{g}'$ and $\overline{g}''$, one obtains the inequality

$$\begin{aligned}\|\tilde{\eta}_n\|_{s,t,\alpha,\Delta t} &\leqslant C_1(g)\|\beta^H\|_\alpha \|\tilde{\eta}_n\|_{s,t,\infty} \\ &\quad + C_1(g)\|\beta^H\|_\alpha (t-s)^\alpha \Big(\|\tilde{\eta}_n\|_{s,t,\infty}\|\tilde{X}_n\|_{s,t,\alpha,\Delta t} + \|\tilde{\eta}_n\|_{s,t,\alpha,\Delta t}\Big),\end{aligned}$$

where $\|\tilde{\eta}_n\|_{s,t,\infty} = \sup_{r\in[s,t]} |\tilde{\eta}_n(r)|$.

Owing to Step 1, if $t - s \leqslant \tau_0$, one has $\|\tilde{X}_n\|_{s,t,\alpha,\Delta t} \leqslant 2C_0(g, \|\beta^H\|_\alpha)$. In addition, one has $\|\tilde{\eta}_n\|_{s,t,\infty} \leqslant |\tilde{\eta}_n(s)| + (t-s)^\alpha \|\eta\|_{s,t,\alpha,\Delta t}$.

One thus obtains an inequality of the type

$$\|\tilde{\eta}_n\|_{s,t,\alpha,\Delta t} \leqslant C_1(g, \|\beta\|_\alpha)\Big(|\tilde{\eta}_n(s)| + (t-s)^\alpha \|\tilde{\eta}_n\|_{s,t,\alpha,\Delta t}\Big)$$

if $t \geqslant s = \ell(s)$ and $t - s \leqslant \tau_0$. Let τ_1 be a positive random variable, chosen such that

$$C_1(g, \|\beta\|_\alpha)\tau_1^\alpha \leqslant \frac{1}{2}$$

and $\tau_1 \leqslant \tau_0$.

Then if $s = \ell(s) \leqslant t$ satisfy $t - s \leqslant \tau_1$, one obtains

$$\|\tilde{\eta}_n\|_{s,t,\alpha,\Delta t} \leqslant 2C_1(g, \|\beta\|_\alpha)|\tilde{\eta}_n(s)|,$$

which gives

$$|\tilde{\eta}_n(t)| \leqslant \big(1 + 2C_1(g, \|\beta\|_\alpha)\tau_1^\alpha\big)|\tilde{\eta}_n(s)|.$$

If the time step size satisfies $\Delta t \leqslant \tau_1$, one thus obtains

$$|\eta_{n,k_2}| \leqslant \big(1 + 2C_1(g, \|\beta\|_\alpha)\tau_1^\alpha\big)|\eta_{n,k_1}|$$

for integers $k_1 < k_2$ such that $(k_2 - k_1)\Delta t \leqslant \tau_1$, and iterating the argument and using the condition $\eta_{n,n} = 1$ one obtains

$$|\eta_{n,k}| \leqslant \big(1 + 2C_1(g, \|\beta\|_\alpha)\tau_1^\alpha\big)^{N_1}$$

where the integer N_1 is chosen such that $N_1\tau_1 \geqslant T$, for all $k \in \{n, \dots, N\}$, if $\Delta t \leqslant \tau_1$. This implies the uniform bound

$$\|\tilde{\eta}_n\|_{t_n,T,\infty} \leqslant C_1'(g, \|\beta^H\|_\alpha), \tag{26}$$

which holds for all $\Delta t \leqslant \tau_1$ and all $n \in \{0, \dots, N\}$.

Step 3. The auxiliary process $\tilde{\zeta}_n$ satisfies the following property: for all $t_n \leqslant s \leqslant t \leqslant T$, such that $s = \ell(s)$, one has

$$\tilde{\zeta}_n(t) - \tilde{\zeta}_n(s) = \int_s^t F_2(\tilde{X}_n(\ell(r)), \tilde{\eta}_n(\ell(r)), \tilde{\eta}_n(\ell(r)))d\beta^H(r),$$

where $F_2(x, \eta, \zeta) = \overline{g}'(x)\zeta + \overline{g}''(x)\eta^2$. Note that $|F_2(x, \eta, \zeta)| + |\partial_x F_2(x, \eta, \zeta)| \leqslant C|\zeta| + C|\eta|^2$, $|\partial_\eta F_2(x, \eta, \zeta)| \leqslant C|\eta|$ and $|\partial_\zeta F_2(x, \eta, \zeta)| \leqslant C$.

Using Lemma 4.1 with $\alpha = \alpha' \in (\frac{1}{2}, H)$, and using the boundedness of $\overline{g}$, $\overline{g}'$, $\overline{g}''$ and $\overline{g}^{(3)}$, one obtains the inequality

$$\begin{aligned}\|\tilde{\zeta}_n\|_{s,t,\alpha,\Delta t} &\leqslant C_1(g)\|\beta^H\|_\alpha(\|\tilde{\zeta}_n\|_{s,t,\infty} + \|\tilde{\eta}_n\|_{s,t,\infty}^2) \\ &+ C_2(g)\|\beta^H\|_\alpha(t-s)^\alpha\Big(\big(\|\tilde{\zeta}_n\|_{s,t,\infty} + \|\tilde{\eta}_n\|_{s,t,\infty}^2\big)\|\tilde{X}_n\|_{s,t,\alpha,\Delta t} + \|\tilde{\eta}_n\|_{s,t,\infty}\|\tilde{\eta}_n\|_{s,t,\alpha,\Delta t} + \|\tilde{\zeta}_n\|_{s,t,\alpha,\Delta t}\Big),\end{aligned}$$

where $\|\tilde{\eta}_n\|_{s,t,\infty} = \sup_{r\in[s,t]} |\tilde{\eta}_n(r)|$ and $\|\tilde{\zeta}_n\|_{s,t,\infty} = \sup_{r\in[s,t]} |\tilde{\zeta}_n(r)|$.

Owing to Steps 1 and 2, if $t - s \leqslant \tau_1$, one has $\|\tilde{X}_n\|_{s,t,\alpha,\Delta t} \leqslant C_0'(g, \|\beta^H\|_\alpha)$ and $\|\tilde{\eta}_n\|_{s,t,\infty} \leqslant C_1'(g, \|\beta^H\|_\alpha)$, and one has the uniform bound $\|\tilde{\eta}_n\|_{t_n,T,\infty} \leqslant C_1'(g, \beta^H\|_\alpha)$ if $\Delta t \leqslant \tau_1$. In addition, one has $\|\tilde{\zeta}_n\|_{s,t,\infty} \leqslant |\tilde{\zeta}_n(s)| + (t-s)^\alpha\|\zeta\|_{s,t,\alpha,\Delta t}$.

One thus obtains an inequality of the type

$$\|\tilde{\zeta}_n\|_{s,t,\alpha,\Delta t} \leqslant C_2(g, \|\beta\|_\alpha)\Big(1 + |\tilde{\zeta}_n(s)| + (t-s)^\alpha\|\tilde{\zeta}_n\|_{s,t,\alpha,\Delta t}\Big)$$

if $t \geqslant s = \ell(s)$ and $t - s \leqslant \tau_0$. Let τ_2 be a positive random variable, chosen such that

$$C_2(g, \|\beta\|_\alpha)\tau_2^\alpha \leqslant \frac{1}{2}$$

and $\tau_2 \leqslant \tau_1$.

Then if $s = \ell(s) \leqslant t$ satisfy $t - s \leqslant \tau_1$, one obtains

$$\|\tilde{\zeta}_n\|_{s,t,\alpha,\Delta t} \leqslant 2C_2(g, \|\beta\|_\alpha)(1 + |\tilde{\zeta}_n(s)|),$$

which gives

$$|\tilde{\zeta}_n(t)| \leqslant \big(1 + 2C_2(g, \|\beta\|_\alpha)\tau_2^\alpha\big)|\tilde{\zeta}_n(s)| + 2C_2(g, \|\beta\|_\alpha)\tau_2^\alpha.$$

If the time step size satisfies $\Delta t \leqslant \tau_2$, one thus obtains

$$|\zeta_{n,k_2}| \leqslant \big(1 + 2C_2(g, \|\beta\|_\alpha)\tau_2^\alpha\big)|\zeta_{n,k_1}| + 2C_2(g, \|\beta\|_\alpha)\tau_2^\alpha$$

for integers $k_1 < k_2$ such that $(k_2 - k_1)\Delta t \leqslant \tau_1$, and iterating the argument and using the condition $\zeta_{n,n} = 0$ one obtains

$$|\zeta_{n,k}| \leqslant \big(1 + 2C_2(g, \|\beta\|_\alpha)\tau_2^\alpha\big)^{N_2}$$

where the integer N_2 is chosen such that $N_2\tau_2 \geqslant T$, for all $k \in \{n, \dots, N\}$, if $\Delta t \leqslant \tau_2$. This implies the uniform bound

$$\|\tilde{\zeta}_n\|_{t_n,T,\infty} \leqslant C_2'(g, \|\beta^H\|_\alpha), \tag{27}$$

which holds for all $\Delta t \leqslant \tau_1$ and all $n \in \{0, \dots, N\}$.

Conclusion

Owing to the expressions (23) for $u_n'(x)$ and $u_n''(x)$, and to the bounds (26) and (27) for $\eta_{n,N} = \tilde{\eta}_n(T)$ and $\zeta_{n,N} = \tilde{\zeta}_n(T)$, one obtains

$$\sup_{x\in\mathbb{R}} \big(|u_n(x)| + |u_n'(x)| + |u_n''(x)|\big) \leqslant \Lambda$$

for all $n \in \{0, \dots, N\}$, where Λ is an almost surely finite random variable, if $\Delta t \leqslant \Delta t_0 = \tau_2$ where τ_2 is the positive random variable constructed in Step 3 above.

This concludes the proof of (22).

□

5. Discussion

In this article, we have studied a class of Euler schemes (3) for slow-fast systems of stochastic differential equations (1). The slow component X^ϵ is driven by a fractional Brownian motion with Hurst index $H > 1/2$, and converges in probability to a process $\overline{X}$, owing to the averaging principle recently proved in **[3]**. We have proved that well-chosen numerical schemes are able to reproduce a version of the averaging principle at the discrete-time level: they satisfy the asymptotic preserving property stated in Theorem 2.2. In particular, the time-step size Δt can be chosen independently of the time scale separation parameter ϵ. We have illustrated the differences with the standard Brownian motion case treated in the recent work **[1]**.

We have left open the important question of proving error estimates: is it possible to prove a uniform accuracy property (as in **[1]** in the standard Brownian case), *i.e.* an error estimate depending on Δt, uniform with respect to ϵ? Studying this question may require more involved techniques (such as the ones employed in **[6]** and references therein) than those used in this manuscript. More precisely, it is challenging to prove error estimates for

$$\mathbb{E}\big[\big|\mathbb{E}^H[\varphi(X_N^\epsilon)] - \mathbb{E}^H[\varphi(\overline{X}_N^0)]\big|\big]$$

when $\epsilon \to 0$, for fixed $\Delta t > 0$, which are uniform with respect to Δt.

Note that obtaining nice error estimates when $\Delta t \to 0$ and/or $\epsilon \to 0$ may provide an alternative proof of the averaging principle from **[3]** by a temporal discretization technique (similar to the one proposed in the seminal article **[4]** in the standard Brownian motion case). In this article, we have assumed that the Hurst index satisfies $H > 1/2$. To the best of our knowledge, the validity and the expression of the averaging principle for SDEs driven by a fractional Brownian motion with Hurst index $H < 1/2$ is not known. The construction of well-chosen numerical schemes, with associated nice error estimates, may provide a strategy to generalize the averaging principle to the case $H < 1/2$. The scheme would then be asymptotic preserving – but stating this property needs to identify the limit at the theoretical level. We leave the challenging question of the generalization for $H < 1/2$ for future works.

As mentioned at the beginning of Section 2, it is straightforward to generalize the results of this manuscript to higher dimensional situation. It would also be straightforward to consider systems of the type

$$\begin{cases} dX^\epsilon(t) = f(X^\epsilon, m^\epsilon)dt + \displaystyle\sum_{j=1}^m g_j(X^\epsilon(t), m^\epsilon(t))d\beta_j^H(t), \\ \\ dm^\epsilon(t) = -\dfrac{1}{\epsilon}m^\epsilon(t)dt + \dfrac{\sqrt{2}}{\sqrt{\epsilon}}\sigma(X^\epsilon(t))dB(t), \end{cases}$$

where $X^\epsilon(t) \in \mathbb{R}^d$, $\sigma_j : \mathbb{R}^d \times \mathbb{R} \to \mathbb{R}^d$, and β_j^H are independent fractional Brownian motions. In particular the treatment of the drift term f has been performed in **[1]**, this is why assuming that $f(x, m) = 0$ in the present article is legitimate, in order to focus on the main features due to the fractional Brownian motion. Assuming that $\sigma(x) = 1$ also simplifies the presentation, however it is straightforward to check that the proof of Proposition 2.3 remains unchanged (up to modifying notation), thus Theorem 2.2 also holds in this case.

Instead of using the standard Euler scheme (3) to discretize the slow component X^ϵ of (1), in order to increase the performance it would be tempting to employ the scheme studied in **[6]**, of the type

$$X_{n+1}^\epsilon = X_n^\epsilon + g(X_n^\epsilon, m_{n+1}^\epsilon)\big(\beta^H(t_{n+1}) - \beta^H(t_n)\big) + \frac{1}{2}\big(\nabla_x g g\big)(X_n^\epsilon, m_{n+1}^\epsilon)\Delta t^{2H}.$$

The definition of m_n^ϵ is not modified from (3). When $\epsilon \to 0$, one obtains the limiting scheme

$$X_{n+1}^0 = X_n^0 + g(X_n^0, \gamma_n)\big(\beta^H(t_{n+1}) - \beta^H(t_n)\big) + \frac{1}{2}\big(\nabla_x g g\big)(X_n^\epsilon, \gamma_n)\Delta t^{2H}.$$

This scheme is consistent (when $\Delta t \to 0$) with the solution of the averaged equation, however it is not clear that the good performance of the modified scheme is preserved when $\epsilon = 0$: we have $\mathbb{E}[g(\cdot, \gamma_n)] = \overline{g}(\cdot)$, however in general one may have $\mathbb{E}[\big(\nabla_x g g\big)(\cdot, \gamma_n)] \neq \big(\nabla_x \overline{g g}\big)(\cdot)$ – however observe that the equality holds if $d = 1$. The construction of more efficient methods than the standard Euler scheme when both $\epsilon > 0$ and $\epsilon = 0$ may be investigated in future works.

Acknowledgments

The author would like to thank Martin Hairer for discussions when this work was initiated, and Shmuel Rakotonirina-Ricquebourg for discussions concerning the proof of Lemma 2.1.

Appendix A. Proof of Lemma 2.1

PROOF. *Proof that (i)⇒(ii).* Assume that $X_N \to X$ in probability. For any function φ of class $\mathcal{C}_b^K$, φ is bounded and Lipschitz continuous, thus there exists $C(\varphi) \in (0,\infty)$ such that

$$\begin{aligned}\mathbb{E}\big[\big|\mathbb{E}[\varphi(X_N)|\mathcal{G}] - \varphi(X)\big|\big] &\leqslant \mathbb{E}\big[\big|\varphi(X_N) - \varphi(X)\big|\big] \\ &\leqslant C(\varphi)\mathbb{E}\big[\min(|X_N - X|, 1)\big] \underset{N\to\infty}{\to} 0,\end{aligned}$$

where the last step is a consequence of convergence in probability. As a consequence (i) implies (ii).

Proof that (ii)⇒(i). Introduce an auxiliary function $\varphi : \mathbb{R} \to \mathbb{R}$ be such that

- φ is of class $\mathcal{C}^\infty$
- $\varphi(x) = 0$ if $|x| \leqslant \frac{1}{2}$
- $\varphi(x) = 0$ if $|x| \geqslant 1$
- $0 \leqslant \varphi(x) \leqslant 1$ if $\frac{1}{2} \leqslant |x| \leqslant 1$.

For every $k \in \mathbb{Z}$ and $\eta \in (0,1)$, introduce the interval $I_{k,\eta} = [\frac{k\eta}{2}, \frac{(k+1)\eta}{2})$, $m_{k,\eta} = \frac{2k\eta+1}{4}$ and the function

$$\varphi_{k,\eta} = \varphi\big(\eta^{-1}(\cdot - m_{k,\eta})\big).$$

Then $\varphi_{k,\eta}$ is of class $\mathcal{C}_b^K$ (where K is an arbitrary integer).

For all $\eta \in (0,1)$, one has

$$\mathbb{P}(|X_N - X| > \eta) = \sum_{k\in\mathbb{Z}} \mathbb{E}\big[\mathbb{1}_{|X_N - X|>\eta}\mathbb{1}_{X\in I_{k,\eta}}\big].$$

Owing to the dominated convergence theorem – one has $\mathbb{E}\big[\mathbb{1}_{|X_N - X|>\eta}\mathbb{1}_{X\in I_{k,\eta}}\big] \leqslant \mathbb{P}(X \in I_{k,\eta})$ with $\sum_{k\in\mathbb{Z}} \mathbb{P}(X \in I_{k,\eta}) = 1$ – it suffices to prove that for all $k \in \mathbb{Z}$ one has

$$\mathbb{E}\big[\mathbb{1}_{|X_N - X|>\eta}\mathbb{1}_{X\in I_{k,\eta}}\big] \underset{N\to\infty}{\to} 0.$$

Note that combining the conditions $|X_N - X| > \eta$ and $X \in I_{k,\eta}$ implies that $\varphi_{k,\eta}(X_N) = 1$ and $\varphi_{k,\eta}(X) = 0$. Using the fact that X is $\mathcal{G}$-measurable, one thus obtains

$$\begin{aligned}\mathbb{E}\big[\mathbb{1}_{|X_N - X|>\eta}\mathbb{1}_{X\in I_{k,\eta}}\big] &\leqslant \mathbb{E}\big[\varphi_{k,\eta}(X_N)\mathbb{1}_{X\in I_{k,\eta}}\big] \\ &\leqslant \mathbb{E}\big[\mathbb{E}[\varphi_{k,\eta}(X_N)|\mathcal{G}]\mathbb{1}_{X\in I_{k,\eta}}\big] \\ &\leqslant \mathbb{E}\big[\big(\mathbb{E}[\varphi_{k,\eta}(X_N)|\mathcal{G}] - \varphi_{k,\eta}(X)\big)\mathbb{1}_{X\in I_{k,\eta}}\big] \\ &\leqslant \mathbb{E}\big[\big|\mathbb{E}[\varphi_{k,\eta}(X_N)|\mathcal{G}] - \varphi_{k,\eta}(X)\big|\big] \underset{n\to\infty}{\to} 0,\end{aligned}$$

using the assumption that (ii) holds, with $\varphi = \varphi_{k,\eta}$.

Applying the dominated convergence theorem then gives

$$\mathbb{P}(|X_N - X| > \eta) \underset{N\to\infty}{\to} 0,$$

for all $\eta \in (0,1)$. As a consequence X_N converges to X in probability when $N \to \infty$, and (ii) implies (i).

This concludes the proof. □

References

[1] C.-E. Bréhier and S. Rakotonirina-Ricquebourg. On asymptotic preserving schemes for a class of stochastic differential equations in averaging and diffusion approximation regimes. *Preprint*, 2020.

[2] J. Gehringer and X.-M. Li. Rough homogenisation with fractional dynamics. *Preprint*, 2020.

[3] M. Hairer and X.-M. Li. Averaging dynamics driven by fractional Brownian motion. *Ann. Probab.*, 48(4):1826–1860, 2020.

[4] R. Z. Hasminskii. On the principle of averaging the Itô's stochastic differential equations. *Kybernetika (Prague)*, 4:260–279, 1968.

[5] J. Hong, C. Huang, and X. Wang. Strong convergence rate of Runge–Kutta methods and simplified step-N Euler schemes for SDEs driven by fractional Brownian motions. *Preprint*, 2017.

[6] Y. Hu, Y. Liu, and D. Nualart. Rate of convergence and asymptotic error distribution of Euler approximation schemes for fractional diffusions. *Ann. Appl. Probab.*, 26(2):1147–1207, 2016.

[7] X.-M. Li and J. Sieber. Slow-fast systems with fractional environment and dynamics. *Preprint*, 2020.

[8] B. B. Mandelbrot and J. W. Van Ness. Fractional Brownian motions, fractional noises and applications. *SIAM Rev.*, 10:422–437, 1968.

[9] Y. Mishura and G. Shevchenko. The rate of convergence for Euler approximations of solutions of stochastic differential equations driven by fractional Brownian motion. *Stochastics*, 80(5):489–511, 2008.

[10] D. Nualart. *The Malliavin calculus and related topics*. Probability and its Applications (New York). Springer-Verlag, Berlin, second edition, 2006.

[11] D. Nualart and A. Răşcanu. Differential equations driven by fractional Brownian motion. *Collect. Math.*, 53(1):55–81, 2002.

[12] G. A. Pavliotis and A. M. Stuart. *Multiscale methods*, volume 53 of *Texts in Applied Mathematics*. Springer, New York, 2008. Averaging and homogenization.

[13] A. Y. Veretennikov. On an averaging principle for systems of stochastic differential equations. *Mat. Sb.*, 181(2):256–268, 1990.

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