Weak approximation of a fractional SDE

Ivan Nourdin, Samy Tindel

WEAK APPROXIMATION OF A FRACTIONAL SDE

IVAN NOURDIN AND SAMY TINDEL

Abstract. In this note, a diffusion approximation result is shown for stochastic differential equations driven by a fractional Brownian motion B with Hurst parameter $H > 1/3$. We shall use a Gaussian regular approximation of B for sake of clarity, and our method of proof will rely on the algebraic integration theory introduced in [9].

1. Introduction

After a decade of efforts [2, 5, 9, 15, 16, 20, 21], it can arguably be said that the basis of the stochastic integration theory with respect to a rough path in general, and with respect to a fractional Brownian motion (fBm in the sequel) in particular, has been now settled in a rather simple and secure way. This allows in particular to define rigorously and solve equations on an arbitrary interval $[0, T]$ with $T > 0$, of the form:

$$dy_t = \sigma(y_t)\, dB_t + b(y_t)\, dt, \quad y_0 = a \in \mathbb{R}^n, \tag{1}$$

where $\sigma : \mathbb{R}^n \to \mathbb{R}^{n\times d}$, $b : \mathbb{R}^n \to \mathbb{R}^n$ are two bounded and smooth functions, and B stands for a d-dimensional fBm with Hurst parameter $H > 1/4$. A question which arises naturally in this context is then to try to establish some of the basic properties of the process y defined by (1), and this global program has already been started as far as moments estimates [11], large deviations [14, 18], or properties of the law [4, 19] are concerned.

In the current note, we wish to address another natural problem related to the fractional diffusion process y defined by (1). Indeed, in the case where B is an ordinary Brownian motion, one of the most popular method in order to simulate y is the following: approximate B by a sequence of smooth or piecewise linear functions $(B^\varepsilon)_{\varepsilon>0}$ which converges in law to B, e.g. an interpolated and rescaled random walk. Then see if the process y^ε solution of equation (1) driven by B^ε converges in law, as a process, to y. This kind of result, usually known as diffusion approximation, has been thoroughly studied in the literature (see e.g. [12, 24, 25]), since it also shows that equations like (1) may emerge as the limit of a noisy equation driven by a fast oscillating function. The diffusion approximation program has also been taken up in the fBm case by Marty in [17], with some random wave problems in mind, but only in the cases where $H > 1/2$ or the dimension d of the fBm is 1. We thus propose to go one step further in this direction, and treat a more complex situation in the fBm setting.

Indeed, we will consider in the sequel a d-dimensional fBm with Hurst parameter $H > 1/3$ as the driving process of equation (1). As an approximating sequence of B, we shall

Date: September 6, 2007.
2000 *Mathematics Subject Classification.* 60H10, 60H05.
Key words and phrases. Weak approximation, fractional Brownian motion, rough paths.

choose $(X^\varepsilon)_{\varepsilon>0}$, where $X^{\varepsilon,i}$ is the smooth Gaussian process defined, for any $t \geq 0$ and $1 \leq i \leq d$, by:

$$X^{\varepsilon,i}(t) \triangleq \varepsilon^H m^i(t/\varepsilon), \quad \text{where} \quad m^i(t) \triangleq \int_0^t \left(B^i_{u+1} - B^i_u\right) du. \tag{2}$$

Let us then consider the process y^ε solution to equation (1) driven by X^ε, namely:

$$dy^\varepsilon_t = \sigma\left(y^\varepsilon_t\right) dX^\varepsilon_t + b\left(y^\varepsilon_t\right) dt, \quad y^\varepsilon_0 = a \in \mathbb{R}^n, \quad t \in [0,T]. \tag{3}$$

Then we shall prove the following theorem:

Theorem 1.1. *Let $(y^\varepsilon)_{\varepsilon>0}$ be the family of processes defined by (3), and let $1/3 < \gamma < H$, where H is the Hurst parameter of B. Then, as $\varepsilon \to 0$, y^ε converges in law to the process y obtained as the solution to (1), where the convergence takes place in the Hölder space $\mathcal{C}^\gamma([0,T];\mathbb{R}^n)$.*

This theorem is thus an extension of [17] to the case of a multidimensional fBm with $H > 1/3$. Let us also mention at this point that we could certainly have treated the case $H > 1/4$ as well, but we preferred to stick to the case $H > 1/3$ for computational and notational sake. It is also worth noting that our approximation X^ε of the fBm is quite easy to manipulate, but may be a little clumsy for simulation purposes. We could have used instead some of the natural approximations in law of the process B itself, such as the random walk transformed by the kernel K_H associated to the fBm considered in [23], the Stroock type approximation using a Poisson process studied in [3, 6], or the correlated random walks introduced in [7]. However, these kind of approximation lead to some technical problems for the convergence of finite dimensional distributions that we prefer to treat in a subsequent communication. The aim of this note is thus quite simple: we wish to show that diffusion approximation results are possible in an irregular fractional context.

The general strategy we shall follow in order to get our main result is rather natural in the rough path theory: it is a well-known fact that y is a continuous function of B and of the Lévy area of B (which will be called $\mathbf{B^2}$), considered as elements of some suitable Hölder spaces. Hence, it will be sufficient to check the convergence of the corresponding approximations B^ε and $\mathbf{B^{2,\varepsilon}}$ in their respective Hölder spaces. Then the two main technical problems we will have to solve are the following:

(1) First of all, we shall use the *simplified* version of the rough path formalism, called algebraic integration, introduced by Gubinelli in [9] and also explained in [19], which will be summarized in the next section. In the particular context of weak approximations, this allows us to deal with approximations of B and $\mathbf{B^2}$ directly, without recurring to discretized paths as in [5]. However, the algebraic integration formalism relies on some space $\mathcal{C}^\gamma_k$, where k stands for a number of variables in $[0,T]$, and γ for a Hölder type exponent. Thus, an important step will be to find a suitable tightness criterion in these spaces.
(2) The convergence of finite dimensional distributions (FDD in the sequel) for the Lévy area $\mathbf{B^2}$ will be based on some estimates from Malliavin calculus (see e.g. [21]) and generalized stochastic calculus *à la* Russo-Vallois [22].

These two technical steps will be developed at Section 3.

Here is how our paper is structured: at Section 2, we shall recall the main notions of the algebraic integration theory. Then Section 3 will be devoted to the weak convergence itself, divided into the tightness result (Section 3.1) and the FDD convergence (Section 3.2).

2. Basic facts about algebraic integration and fractional SDEs

This section contains a summary of the algebraic integration introduced in [9], which we also used in [19] in order to solve and analyze fractional SDEs. We recall its main features here, since our approximation result will also be obtained in this setting.

Let x be a Hölder continuous $\mathbb{R}^d$-valued function of order γ, with $1/3 < \gamma \leq 1/2$, and $\sigma : \mathbb{R}^n \to \mathbb{R}^{n\times d}$, $b : \mathbb{R}^n \to \mathbb{R}^n$ be two bounded and smooth functions. We shall consider in the sequel the n-dimensional equation

$$dy_t = \sigma(y_t)\, dx_t + b(y_t)\, dt, \quad y_0 = a \in \mathbb{R}^n, \quad t \in [0,T]. \tag{4}$$

In order to define rigorously and solve this equation, we will need some algebraic and analytic notions which are introduced in the next subsection.

2.1. Increments.

We first present the basic algebraic structures which will allow us to define a pathwise integral with respect to irregular functions. For an arbitrary real number $T > 0$, a vector space V and an integer $k \geq 1$ we denote by $\mathcal{C}_k(V)$ the set of functions $g : [0,T]^k \to V$ such that $g_{t_1\cdots t_k} = 0$ whenever $t_i = t_{i+1}$ for some $i \leq k-1$. Such a function will be called a $(k-1)$-*increment*, and we will set $\mathcal{C}_*(V) = \cup_{k\geq 1}\mathcal{C}_k(V)$. An important elementary operator is defined by

$$\delta : \mathcal{C}_k(V) \to \mathcal{C}_{k+1}(V), \qquad (\delta g)_{t_1\cdots t_{k+1}} = \sum_{i=1}^{k+1} (-1)^{k-i} g_{t_1\cdots \hat{t}_i \cdots t_{k+1}}, \tag{5}$$

where $\hat{t}_i$ means that this particular argument is omitted. A fundamental property of δ, which is easily verified, is that $\delta\delta = 0$, where $\delta\delta$ is considered as an operator from $\mathcal{C}_k(V)$ to $\mathcal{C}_{k+2}(V)$. We will denote $\mathcal{ZC}_k(V) = \mathcal{C}_k(V) \cap \mathrm{Ker}\delta$ and $\mathcal{BC}_k(V) = \mathcal{C}_k(V) \cap \mathrm{Im}\delta$.

Some simple examples of actions of δ are obtained for $g \in \mathcal{C}_1(V)$ and $h \in \mathcal{C}_2(V)$. Then, for any $s, u, t \in [0,T]$, we have

$$(\delta g)_{st} = g_t - g_s, \quad \text{and} \quad (\delta h)_{sut} = h_{st} - h_{su} - h_{ut}. \tag{6}$$

Furthermore, it is easily checked that $\mathcal{ZC}_{k+1}(V) = \mathcal{BC}_k(V)$ for any $k \geq 1$. In particular, the following basic property holds:

Lemma 2.1. *Let $k \geq 1$ and $h \in \mathcal{ZC}_{k+1}(V)$. Then there exists a (non unique) $f \in \mathcal{C}_k(V)$ such that $h = \delta f$.*

Observe that Lemma 2.1 implies that all elements $h \in \mathcal{C}_2(V)$ with $\delta h = 0$ can be written as $h = \delta f$ for some (non unique) $f \in \mathcal{C}_1(V)$. Thus we get a heuristic interpretation of $\delta|_{\mathcal{C}_2(V)}$: it measures how much a given 1-increment is far from being an exact increment of a function, i.e., a finite difference.

Note that our further discussion will mainly rely on k-increments with $k \leq 2$. For the simplicity of the exposition, we will assume from now that $V = \mathbb{R}^d$. We measure the

size of these increments by Hölder norms, which are defined in the following way: for $f \in \mathcal{C}_2(V)$ let

$$\|f\|_\mu = \sup_{s,t\in[0,T]} \frac{|f_{st}|}{|t-s|^\mu}, \quad \text{and} \quad \mathcal{C}_2^\mu(V) = \{f \in \mathcal{C}_2(V);\, \|f\|_\mu < \infty\}\,.$$

Obviously, the usual Hölder spaces $\mathcal{C}_1^\mu(V)$ are determined in the following way: for a continuous function $g \in \mathcal{C}_1(V)$ simply set

$$\|g\|_\mu = \|\delta g\|_\mu, \tag{7}$$

and we will say that $g \in \mathcal{C}_1^\mu(V)$ iff $\|g\|_\mu$ is finite. Note that $\|\cdot\|_\mu$ is only a semi-norm on $\mathcal{C}_1(V)$, but we will work in general on spaces of the type

$$\mathcal{C}_{1,a}^\mu(V) = \{g : [0,T] \to V;\, g_0 = a,\, \|g\|_\mu < \infty\}\,, \tag{8}$$

for a given $a \in V$, on which $\|g\|_\mu$ is a norm. For $h \in \mathcal{C}_3(V)$ set in the same way

$$\begin{aligned} \|h\|_{\gamma,\rho} &= \sup_{s,u,t\in[0,T]} \frac{|h_{sut}|}{|u-s|^\gamma |t-u|^\rho} \qquad (9)\\ \|h\|_\mu &= \inf\left\{\sum_i \|h_i\|_{\rho_i,\mu-\rho_i};\, h = \sum_i h_i,\, 0<\rho_i<\mu\right\}, \end{aligned}$$

where the infimum is taken over all sequences $\{h_i \in \mathcal{C}_3(V)\}$ such that $h = \sum_i h_i$ and for all choices of the numbers $\rho_i \in (0,\mu)$. Then $\|\cdot\|_\mu$ is easily seen to be a norm on $\mathcal{C}_3(V)$, and we set

$$\mathcal{C}_3^\mu(V) := \{h \in \mathcal{C}_3(V);\, \|h\|_\mu < \infty\}\,.$$

Eventually, let $\mathcal{C}_3^{1+}(V) = \cup_{\mu>1}\mathcal{C}_3^\mu(V)$, and note that the same kind of norms can be considered on the spaces $\mathcal{ZC}_3(V)$, leading to the definition of the spaces $\mathcal{ZC}_3^\mu(V)$ and $\mathcal{ZC}_3^{1+}(V)$.

With these notations in mind, the crucial point in the current approach to pathwise integration of irregular paths is that the operator δ can be inverted under mild smoothness assumptions. This inverse is called Λ. The proof of the following proposition may be found in [9], and in a more elementary form in [10]:

Proposition 2.2. *There exists a unique linear map $\Lambda : \mathcal{ZC}_3^{1+}(V) \to \mathcal{C}_2^{1+}(V)$ such that*

$$\delta\Lambda = Id_{\mathcal{ZC}_3^{1+}(V)} \quad \text{and} \quad \Lambda\delta = Id_{\mathcal{C}_2^{1+}(V)}.$$

In other words, for any $h \in \mathcal{C}_3^{1+}(V)$ such that $\delta h = 0$ there exists a unique $g = \Lambda(h) \in \mathcal{C}_2^{1+}(V)$ such that $\delta g = h$. Furthermore, for any $\mu > 1$, the map Λ is continuous from $\mathcal{ZC}_3^\mu(V)$ to $\mathcal{C}_2^\mu(V)$ and we have

$$\|\Lambda h\|_\mu \le \frac{1}{2^\mu - 2}\|h\|_\mu, \qquad h \in \mathcal{ZC}_3^\mu(V). \tag{10}$$

Moreover, Λ has a nice interpretation in terms of generalized Young integrals:

Corollary 2.3. *For any 1-increment $g \in \mathcal{C}_2(V)$ such that $\delta g \in \mathcal{C}_3^{1+}$ set $\delta f = (Id - \Lambda\delta)g$. Then*

$$(\delta f)_{st} = \lim_{|\Pi_{ts}|\to 0} \sum_{i=0}^{n} g_{t_i\, t_{i+1}},$$

where the limit is over any partition $\Pi_{st} = \{t_0 = s, \dots, t_n = t\}$ *of* $[s,t]$*, whose mesh tends to zero. Thus, the 1-increment* δf *is the indefinite integral of the 1-increment* g*.*

2.2. **Weakly controlled paths.** This subsection is devoted to the definition of generalized integrals with respect to a rough path of order 2, and to the resolution of equation (4). Notice that, in the sequel of our paper, we will use both the notations $\int_s^t f dg$ or $\mathcal{J}_{st}(f\, dg)$ for the integral of a function f with respect to a given increment dg on the interval $[s,t]$. The second notation $\mathcal{J}_{st}(f\, dg)$ will be used to avoid some cumbersome notations in our computations. Observe also that the drift term b is generally harmless if one wants to solve the equation (4). See e.g. Remark 3.14 in [19]. Hence, we will simply deal with an equation of the form

$$dy_t = \sigma(y_t)\, dx_t, \quad t \in [0,T], \qquad \text{with} \qquad y_0 = a \tag{11}$$

in the remainder of this section.

Before going into the technical details, let us make some heuristic considerations about the properties that a solution of equation (4) should have. Set $\hat{\sigma}_t = \sigma(y_t)$, and suppose that y is a solution of (11), with $y \in \mathcal{C}_1^\kappa$ for a given $1/3 < \kappa < \gamma$. Then the integral form of our equation can be written as

$$y_t = a + \int_0^t \hat{\sigma}_u dx_u, \qquad t \in [0,T]. \tag{12}$$

Our approach to generalized integrals induces us to work with increments of the form $(\delta y)_{st} = y_t - y_s$ instead of (12). However, it is easily checked that one can decompose (12) into

$$(\delta y)_{st} = \int_s^t \hat{\sigma}_u dx_u = \hat{\sigma}_s (\delta x)_{st} + \rho_{st}, \qquad \text{with} \qquad \rho_{st} = \int_s^t (\hat{\sigma}_u - \hat{\sigma}_s) dx_u,$$

if our integral is linear. We thus have obtained a decomposition of y of the form $\delta y = \hat{\sigma}\delta x + \rho$. Let us see, still at a heuristic level, which regularity we can expect for $\hat{\sigma}$ and r: If σ is a C_b^1-function, we have that $\hat{\sigma}$ is bounded and

$$|\hat{\sigma}_t - \hat{\sigma}_s| \le \|\nabla\sigma\|_\infty \|y\|_\kappa |t-s|^\kappa,$$

where $\|y\|_\kappa$ denotes the Hölder norm of y defined by (7). Hence we have that $\hat{\sigma}$ belongs to $\mathcal{C}_1^\kappa$ and is bounded. As far as ρ is concerned, it should inherit both the regularities of $\delta\hat{\sigma}$ and x, provided that the integral $\int_s^t (\hat{\sigma}_u - \hat{\sigma}_s) dx_u = \int_s^t (\delta\hat{\sigma})_{su} dx_u$ is well defined. Thus, one should expect that $\rho \in \mathcal{C}_2^{2\kappa}$, and even $\rho \in \mathcal{C}_2^{\kappa+\gamma}$. To summarize, we have found that a solution δy of the equation should be decomposable into

$$\delta y = \hat{\sigma}\delta x + \rho, \qquad \text{with} \qquad \hat{\sigma} \in \mathcal{C}_1^\gamma \text{ bounded and } \rho \in \mathcal{C}_2^{2\kappa}. \tag{13}$$

This is precisely the structure we will demand for a possible solution of (11):

Definition 2.4. *Let* z *be a path in* $\mathcal{C}_1^\kappa(\mathbb{R}^k)$ *with* $\kappa \le \gamma$ *and* $2\kappa + \gamma > 1$*. We say that* z *is a controlled path based on* x*, if* $z_0 = a$*, which is a given initial condition in* $\mathbb{R}^k$*, and* $\delta z \in \mathcal{C}_2^\kappa(\mathbb{R}^k)$ *can be decomposed into*

$$\delta z = \zeta \delta x + r, \quad \textit{i.e.} \quad (\delta z)_{st} = \zeta_s (\delta x)_{st} + \rho_{st}, \quad s,t \in [0,T], \tag{14}$$

with $\zeta \in \mathcal{C}_1^\kappa(\mathbb{R}^{k\times d})$ and ρ is a regular part belonging to $\mathcal{C}_2^{2\kappa}(\mathbb{R}^k)$. The space of classical controlled paths will be denoted by $\mathcal{Q}_{\kappa,a}(\mathbb{R}^k)$, and a path $z \in \mathcal{Q}_{\kappa,a}(\mathbb{R}^k)$ should be considered in fact as a couple (z,ζ). The natural semi-norm on $\mathcal{Q}_{\kappa,a}(\mathbb{R}^k)$ is given by

$$\mathcal{N}[z;\mathcal{Q}_{\kappa,a}(\mathbb{R}^k)] = \mathcal{N}[z;\mathcal{C}_1^\kappa(\mathbb{R}^k)] + \mathcal{N}[\zeta;\mathcal{C}_1^b(\mathbb{R}^{k,d})] + \mathcal{N}[\zeta;\mathcal{C}_1^\kappa(\mathbb{R}^{k,d})] + \mathcal{N}[\rho;\mathcal{C}_2^{2\kappa}(\mathbb{R}^k)]$$

with $\mathcal{N}[g;\mathcal{C}_1^\kappa(V)]$ defined by (7) and $\mathcal{N}[\zeta;\mathcal{C}_1^b(V)] = \sup_{0\le s\le T}|\zeta_s|_V$.

Having defined our algebraic and analytic framework, we now can give a sketch of the strategy we used in [19] in order to solve equation (11):

(1) We verified the stability of $\mathcal{Q}_{\kappa,a}(\mathbb{R}^k)$ under a smooth map $\varphi : \mathbb{R}^k \to \mathbb{R}^n$.
(2) We defined rigorously the integral $\int z_u dx_u = \mathcal{J}(zdx)$ for a controlled path z and computed its decomposition (14).
(3) We solved equation (11) in the space $\mathcal{Q}_{\kappa,a}(\mathbb{R}^k)$ by a fixed point argument.

Actually, for the second point we had to assume a priori the following hypothesis on the driving rough path, which is standard in rough path type considerations:

Hypothesis 2.5. *The $\mathbb{R}^d$-valued γ-Hölder path x admits a Lévy area, that is a process $\mathbf{x^2} = \mathcal{J}(dxdx) \in \mathcal{C}_2^{2\gamma}(\mathbb{R}^{d\times d})$ satisfying*

$$\delta\mathbf{x^2} = \delta x \otimes \delta x, \quad \textit{i.e.} \quad \left[(\delta\mathbf{x^2})_{sut}\right](i,j) = [\delta x^i]_{su}[\delta x^j]_{ut}, \quad s,u,t \in [0,T],\ i,j \in \{1,\dots,d\}.$$

Then we proved in [19] the following result, using the strategy sketched above:

Theorem 2.6. *Let x be a process satisfying Hypothesis 2.5 and $\sigma : \mathbb{R}^n \to \mathbb{R}^{n\times d}$ be a C^2 function, which is bounded together with its derivatives. Then*

(1) *Equation (11) admits a unique solution y in $\mathcal{Q}_{\kappa,a}(\mathbb{R}^n)$ for any $\kappa < \gamma$ such that $2\kappa + \gamma > 1$.*
(2) *The mapping $(a, x, \mathbf{x^2}) \mapsto y$ is continuous from $\mathbb{R}^n \times \mathcal{C}_1^\gamma(\mathbb{R}^d) \times \mathcal{C}_2^{2\gamma}(\mathbb{R}^{d\times d})$ to $\mathcal{Q}_{\kappa,a}(\mathbb{R}^n)$.*

We shall see in the next subsection that this general theorem can be applied in the fBm context.

2.3. **Malliavin calculus preliminaries and application to the fBm.** In order to see that Theorem 2.6 can be applied to the fBm, one should check that almost surely, the process B satisfies our assumption 2.5. This will be based on some Malliavin calculus techniques, which we proceed to recall now for further use, following [21]. In fact, we focus here on the case $H \in (\frac{1}{3}, \frac{1}{2})$, the case $H > 1/2$ being easier to treat.

Let $B = (B^1, \dots, B^d)$ be a d-dimensional fBm of Hurst index $H \in (\frac{1}{3}, \frac{1}{2})$. In other words, $B^1, \dots, B^d$ are d independent centered Gaussian processes with covariance function

$$R(t,s) = \frac{1}{2}\big(s^{2H} + t^{2H} - |t-s|^{2H}\big).$$

It can be shown that each B^i can be represented as:

$$B_t^i = \int_0^t K(t,s)dW_s^i,$$

where $W = (W^1, \dots, W^d)$ is a d-dimensional Wiener process, and K is a kernel whose exact expression is:

$$K(t,s) = c_H \left[\left(\frac{s}{t(t-s)} \right)^{1/2-H} + (1/2-H) s^{1/2-H} \int_s^t \frac{du}{u^{3/2-H}(u-s)^{1/2-H}} \right],$$

for a constant $c_H > 0$.

Let $\mathcal{E}$ be the set of step-functions on $[0,T]$ with values in $\mathbb{R}^d$. Consider the Hilbert space $\mathcal{H}$ defined as the closure of $\mathcal{E}$ with respect to the scalar product induced by

$$\left\langle (\mathbf{1}_{[0,t_1]}, \dots, \mathbf{1}_{[0,t_d]}), (\mathbf{1}_{[0,s_1]}, \dots, \mathbf{1}_{[0,s_d]}) \right\rangle_{\mathcal{H}} = \sum_{i=1}^d R(t_i, s_i), \quad s_i, t_i \in [0,T],\ i = 1, \dots, d.$$

Then a natural representation of the scalar product in $\mathcal{H}$ is given via the operator K^* defined from $\mathcal{E}$ to $L^2([0,T])$ by:

$$[K^*\varphi](t) = K(T,t)\varphi(t) + \int_t^T [\varphi(r) - \varphi(t)] \partial_r K(r,t)\, dr,$$

and it can be checked that K^* can be extended as an isometry between $\mathcal{H}$ and the Hilbert space $L^2([0,T]; \mathbb{R}^d)$. Thus the inner product in $\mathcal{H}$ can be defined as:

$$\langle \varphi, \psi \rangle_{\mathcal{H}} \triangleq \langle K^*\varphi, K^*\psi \rangle_{L^2([0,T];\mathbb{R}^d)}.$$

The mapping $(\mathbf{1}_{[0,t_1]}, \dots, \mathbf{1}_{[0,t_d]}) \mapsto \sum_{i=1}^d B_{t_i}^i$ can also be extended into an isometry between $\mathcal{H}$ and the first Gaussian chaos $H_1(B)$ associated with $B = (B^1, \dots, B^d)$. We denote this isometry by $\varphi \mapsto B(\varphi)$. It is shown in [21] that $\mathcal{C}_1^\gamma(\mathbb{R}^d) \subset \mathcal{H}$ whenever $\gamma > 1/2 - H$, which allows to define $B(\varphi)$ for such kind of functions.

Let $\mathcal{S}$ be the set of smooth cylindrical random variables of the form

$$F = f(B(\varphi_1), \dots, B(\varphi_k)), \qquad \varphi_i \in \mathcal{H}, \quad i = 1, \dots, k,$$

where $f \in C^\infty(\mathbb{R}^k, \mathbb{R})$ is bounded with bounded derivatives. The derivative operator D of a smooth cylindrical random variable of the above form is defined as the $\mathcal{H}$-valued random variable

$$DF = \sum_{i=1}^k \frac{\partial f}{\partial x_i}(B(\varphi_1), \dots, B(\varphi_n))\varphi_i.$$

This operator is closable from $L^p(\Omega)$ into $L^p(\Omega; \mathcal{H})$. As usual, $\mathbb{D}^{1,2}$ denotes the closure of the set of smooth random variables with respect to the norm

$$\|F\|_{1,2}^2 = \mathrm{E}|F|^2 + \mathrm{E}\|DF\|_{\mathcal{H}}^2.$$

In particular, if $D^i F$ designates the Malliavin derivative of a functional $F \in \mathbb{D}^{1,2}$ with respect to B^i, we have $D^i B_t^j = \delta_{i,j} \mathbf{1}_{[0,t]}$ for $i, j = 1, \dots, d$, where $\delta_{i,j}$ stands for the Kronecker symbol. The divergence operator δ^{S} is then defined as the adjoint of the derivative operator (observe the unusual notation δ^{S} for this Skorokhod integral, due to the fact that δ has already been used at Section 2). If a random variable $u \in L^2(\Omega; \mathcal{H})$ belongs to $\mathrm{dom}(\delta^{\mathsf{S}})$, the domain of the divergence operator, then $\delta^{\mathsf{S}}(u)$ is defined by the duality relationship

$$\mathrm{E}(F\delta^{\mathsf{S}}(u)) = \mathrm{E}\langle DF, u \rangle_{\mathcal{H}}, \tag{15}$$

for every $F \in \mathbb{D}^{1,2}$. The following are two basic properties of the divergence operator:

i) $\mathbb{D}^{1,2}(\mathcal{H}) \subset \mathrm{Dom}\delta^{\mathfrak{S}}$ and, for any $u \in \mathbb{D}^{1,2}(\mathcal{H})$:

$$E\big(\delta^{\mathfrak{S}}(u)^2\big) = E\big(\|u\|_{\mathcal{H}}^2\big) + E\big(\|Du\|_{\mathcal{H}\otimes\mathcal{H}}^2\big).$$

ii) If $u \in \mathrm{Dom}(\delta^{\mathfrak{S}})$ and $F \in \mathbb{D}^{1,2}$ is such that $Fu \in L^2(\Omega;\mathcal{H})$, then we have the following integration by parts formula:

$$\delta(Fu) = F\delta^{\mathfrak{S}}(u) - \langle DF, u\rangle_{\mathcal{H}}. \tag{16}$$

Let us now define the symmetric integral of a stochastic process according to Russo and Vallois [22]:

Definition 2.7. *Let $u = \{u_t;\, t \in [0,T]\}$ be a $\mathbb{R}^d$-valued continuous stochastic process. Assume that*

$$\int_0^T u_s \cdot d^\circ B_s \triangleq P - \lim_{\varepsilon\to 0} \frac{1}{2\varepsilon}\int_0^T u_s \cdot (B_{s+\varepsilon} - B_{s-\varepsilon})ds$$

exists. Then this limit is called symmetric integral of u with respect to B. *The one-dimensional integral of a process u with respect to any of the B^i can also be defined in the same way.*

It is shown in [21, Proposition 5] that if u is smooth enough in the Malliavin calculus sense, then it is also integrable with respect to the fBm in the sense of Definition 2.7. In particular, if $u_s = B_s^i = B^i(\mathbf{1}_{[0,s]})$, then $u \in \mathbb{D}^{1,2}$, it satisfies the assumptions of [21, Proposition 5], and the Itô-Stratonovich formula for the fBm shows that

$$(B_t^i - B_s^i)^2 = 2\int_s^t B_u^i\, d^\circ B_u^i, \quad \text{for} \quad s,t \in [0,T]. \tag{17}$$

We are now ready to prove that Theorem 2.6 can be applied to the fBm, which amounts to check Hypothesis 2.5. Notice that this proof is included here in order to emphasize the role played by Stratonovich type integrals in our approach, but is also done in [19].

Proposition 2.8. *Let B be a d-dimensional fBm and suppose $H > 1/3$. Then almost all sample paths of B satisfy Hypothesis 2.5, with any Hölder exponent $1/3 < \gamma < H$, and a Lévy area given by*

$$\mathbf{B}^{\mathbf{2}}_{st} = \int_s^t d^\circ B_u \otimes \int_s^u d^\circ B_v, \quad \textit{i. e.} \quad \mathbf{B}^{\mathbf{2}}_{st}(i,j) = \int_s^t d^\circ B_u^i \int_s^u d^\circ B_v^j, \quad i,j \in \{1,\dots,d\},$$

for $0 \le s < t \le T$, where the stochastic integrals are understood in the Stratonovich sense of Definition 2.7.

Proof. First of all, it is a classical fact that $B \in \mathcal{C}_1^\gamma$ for any $1/3 < \gamma < H$, when B is a fBm with $H > 1/3$. Hence, let us show first that $\mathbf{B^2}$ is a well-defined process: indeed, according to (17), we have

$$\mathbf{B}^{\mathbf{2}}_{st}(i,i) = \frac{(B_t^i - B_s^i)^2}{2}, \quad i \in \{1,\dots,d\}, \quad 0 \le s < t \le T. \tag{18}$$

Furthermore, we have already mentioned that $\mathcal{C}_1^\gamma \subset \mathcal{H}$ for any $\gamma > 1/2 - H$. In particular, if $H > \gamma > 1/3$, the condition $\gamma > 1/2 - H$ is satisfied, and conditioning with respect to B^j, $\mathbf{B}^{\mathbf{2}}_{st}(i,j)$ can also be defined as a Wiener integral with respect to B^i, of the form $B^i(\varphi)$ for a well-chosen φ. Hence, $\mathbf{B^2}$ is almost surely a well-defined element of $\mathcal{C}_2$.

Now, the substitution formula for Stratonovich integrals easily yields that $\delta\mathbf{B^2} = \delta B \otimes \delta B$. Furthermore, by stationarity and the scaling property of the fBm, we have that

$$\mathrm{E}\left[|\mathbf{B}^{\mathbf{2}}_{st}(i,j)|^2\right] = |t-s|^{4H}\mathrm{E}\left[|\mathbf{B}^{\mathbf{2}}_{01}(i,j)|^2\right] = c|t-s|^{4H}.$$

Invoking this equality and thanks to the fact that $\mathbf{B^2}$ is a process in the second chaos of the fBm B, on which all L^p norms are equivalent for $p > 1$, we get that

$$\mathrm{E}\left[|\mathbf{B}^{\mathbf{2}}_{st}(i,j)|^p\right] \leq c_p|t-s|^{2pH}.$$

This allows to conclude, thanks to a slight variation of Garsia's lemma, that $\mathbf{B^2} \in \mathcal{C}_2^{2\gamma}(\mathbb{R}^{d,d})$ for any $\gamma < 1/3$, which ends the proof.

□

With all these results in hand, we have obtained a reasonable definition of diffusion processes driven by a fBm, and we can now proceed to their approximation in law.

3. Weak convergence

In this section, we will introduce a natural regularization of B, namely X^ε, which converges in law to B. This will allow to interpret equation (3) in the usual Riemann sense. We will then study the convergence in law of the process y^ε solution to (3) towards the solution y of (1).

As mentioned in the introduction, the approximation of B we shall deal with is defined as follows, for $i = 1, \ldots, d$:

$$X^{\varepsilon,i}(t) \triangleq \varepsilon^H m^i(t/\varepsilon), \quad \text{where} \quad m^i(t) \triangleq \int_0^t \left(B^i_{u+1} - B^i_u\right)du. \tag{19}$$

Furthermore, we have recalled at Theorem 2.6 that the solution y to (1) is a continuous function of $(a, B, \mathbf{B^2})$, considered respectively as elements of $\mathbb{R}^n, \mathcal{C}_1^\gamma$ and $\mathcal{C}_2^{2\gamma}$ for $1/3 < \gamma < H$. Thus our approximation theorem 1.1 can be easily deduced from the following result:

Theorem 3.1. *For any $\varepsilon > 0$, let $\mathbf{X}^{\mathbf{2},\varepsilon} = (\mathbf{X}^{\mathbf{2},\varepsilon}_{st}(i,j))_{s,t\geq 0;\, i,j=1,\ldots,d}$ be the natural Lévy's area associated to X^ε, defined by*

$$\mathbf{X}^{\mathbf{2},\varepsilon}_{st}(i,j) = \int_s^t (X^{j,\varepsilon}_u - X^{j,\varepsilon}_s)dX^{i,\varepsilon}_u,$$

where the integral is understood in the Riemann sense. Then, as $\varepsilon \to 0$,

$$(X^\varepsilon, \mathbf{X}^{\mathbf{2},\varepsilon}) \xrightarrow{\text{Law}} (B, \mathbf{B^2}), \tag{20}$$

where $\mathbf{B^2}$ denotes the Lévy area defined at Proposition 2.8, and where the convergence in law holds in spaces $\mathcal{C}_1^\mu(\mathbb{R}^d) \times \mathcal{C}_2^{2\mu}(\mathbb{R}^d)$, for any $\mu < H$.

The remainder of this section is devoted to the proof of Theorem 3.1. As usual in the context of weak convergence of stochastic processes, we divide the proof into the weak convergence for finite-dimensional distributions and a tightness type result.

Remark 3.2. A natural idea in order to prove Theorem 3.1 could be to try to use the methodology initiated by Kurtz and Protter in [13]. But the problem, here, is that the quantities we are dealing with are not "close enough" to a martingale.

3.1. **Tightness.** We first need a general tightness criterion in the Hölder spaces $\mathcal{C}_1^\mu$ and $\mathcal{C}_2^{2\mu}$.

Lemma 3.3. *Put*

$$\mathscr{E}^\gamma = \left\{(x, \mathbf{x^2}) \in \mathcal{C}_1^\gamma \times \mathcal{C}_2^{2\gamma} \mid x_0 = \mathbf{x}^{\mathbf{2}}_{00} = 0 \text{ and } \mathbf{x^2} \text{ verifies (21)}\right\},$$

where

$$\forall s,t \geq 0, \forall i,j = 1,\dots,d: \quad \mathbf{x}^{\mathbf{2}}_{st}(i,j) = \mathbf{x}^{\mathbf{2}}_{0t}(i,j) - \mathbf{x}^{\mathbf{2}}_{0s}(i,j) - x^i_s(x^j_t - x^j_s). \tag{21}$$

Let μ such that $0 \leq \mu < \gamma$. Then, any bounded subset $\mathscr{Q}$ of $\mathscr{E}^\gamma$ is precompact in $\mathcal{C}_1^\mu \times \mathcal{C}_2^{2\mu}$.

Proof. Let $(x^n, \mathbf{x^{2,n}})$ be a sequence of $\mathscr{Q}$. By assumption, $(x^n, \mathbf{x}^{\mathbf{2,n}}_{0\cdot})$ is bounded and equicontinuous. Then, Ascoli's theorem applies and, at least along a subsequence, which may also be called $(x^n, \mathbf{x}^{\mathbf{2,n}}_{0\cdot})$, it converges uniformly to $(x, \mathbf{x}^{\mathbf{2}}_{0\cdot})$. Using (21), we obtain in fact that $(x^n, \mathbf{x^{2,n}})$ converges uniformly to $(x, \mathbf{x^2})$. Moreover, since we obviously have

$$\|x\|_\mu \leq \liminf_{n\to\infty} \|x^n\|_\mu \quad \text{and} \quad \|\mathbf{x^2}\|_{2\mu} \leq \liminf_{n\to\infty} \|\mathbf{x^{2,n}}\|_{2\mu},$$

we deduce that $(x, \mathbf{x^2}) \in \mathcal{C}_1^\mu \times \mathcal{C}_2^{2\mu}$. Finally, we have

$$\|x - x^n\|_\mu \longrightarrow 0 \quad \text{and} \quad \|\mathbf{x^2} - \mathbf{x^{2,n}}\|_{2\mu} \longrightarrow 0,$$

owing to the fact that

$$\|x - x^n\|_\mu \leq \|x - x^n\|_\gamma \|x - x^n\|_\infty^{1-\frac{\mu}{\gamma}} \leq \big(\|x\|_\gamma + \|x^n\|_\gamma\big)\|x - x^n\|_\infty^{1-\frac{\mu}{\gamma}}$$

and similarly:

$$\|\mathbf{x^2} - \mathbf{x^{2,n}}\|_{2\mu} \leq \big(\|\mathbf{x^2}\|_{2\gamma} + \|\mathbf{x^{2,n}}\|_{2\gamma}\big)\|\mathbf{x^2} - \mathbf{x^{2,n}}\|_\infty^{1-\frac{\mu}{\gamma}}.$$

□

We will use the last result in order to get a reasonable tightness criterion for our approximation processes X^ε and $\mathbf{X^{2,\varepsilon}}$, by means of a slight elaboration of [15, Corollary 6.1]:

Proposition 3.4. *Let X^ε and $\mathbf{X^{2,\varepsilon}}$ be defined as above. If for every $\eta > 0$, there exists $\gamma > \mu$ and $A < \infty$ such that*

$$\sup_{0<\varepsilon\leq 1} P[\|X^\varepsilon\|_\gamma > A] \leq \eta \quad \text{and} \quad \sup_{0<\varepsilon\leq 1} P[\|\mathbf{X^{2,\varepsilon}}\|_{2\gamma} > A] \leq \eta, \tag{22}$$

then $(X^\varepsilon, \mathbf{X^{2,\varepsilon}})$ is tight in $\mathcal{C}_1^\mu(\mathbb{R}^d) \times \mathcal{C}_2^{2\mu}(\mathbb{R}^{d,d})$.

Proof. Recall the Prokhorov theorem relating precompactness of measures *on* a space to compactness of sets *in* the space. This result states that a family M of probability measures on the Borel sets of a complete separable metric space S is weakly precompact if and only if for every $\eta > 0$ there exists a compact set $K_\eta \subset S$ such that

$$\sup_{\mu\in M} \mu(S - K_\eta) \leq \eta.$$

Furthermore, it is readily checked that the couple $(X^\varepsilon, \mathbf{X^{2,\varepsilon}})$ satisfies the assumption (21), which allows to apply Lemma 3.3. Hence, combining this lemma with Prokhorov's theorem, our proposition is easily proved.

□

Finally, in order to get the tightness of our approximating sequence, we will need the following technical lemma:

Lemma 3.5. *Let*

$$\phi(\alpha, w) = \frac{1}{2}\big(2w^\alpha - (w+1)^\alpha - |w-1|^\alpha\big), \quad \textit{for } \alpha \in (0,2) \textit{ and } w \geq 0.$$

Then, for any $\alpha \in (0,2)$, there exists $c_\alpha > 0$ such that

$$|\phi(\alpha, w)| \leq c_\alpha \begin{cases} w^{\alpha-2} & \textit{if } w \geq 2, \\ 1 & \textit{if } 0 \leq w \leq 2. \end{cases}$$

Proof. The case $w \leq 2$ is easy. Thus, assume $w \geq 2$. Then

$$|\phi(\alpha, w)| = \alpha|\alpha - 1| \int_0^1 dx \int_0^1 \frac{dy}{(w+y-x)^{2-\alpha}} = \alpha|\alpha-1| w^{\alpha-2} \int_0^1 dx \int_0^1 \frac{dy}{\big(1 + \frac{y-x}{w}\big)^{2-\alpha}}.$$

But, for $(x,y) \in [0,1]^2$, we have $\frac{|y-x|}{w} \leq \frac{1}{w} \leq \frac{1}{2}$ and the conclusion follows.

□

Let us turn now to the main result of this subsection:

Proposition 3.6. *The sequence $\{(X^\varepsilon, \mathbf{X}^{\mathbf{2},\varepsilon}); \varepsilon > 0\}$ defined at Theorem 3.1 is tight in $\mathcal{C}_1^\mu(\mathbb{R}^d) \times \mathcal{C}_2^{2\mu}(\mathbb{R}^{d,d})$.*

Proof. Thanks to Proposition 3.4, we just have to prove that $(X^\varepsilon, \mathbf{X}^{\mathbf{2},\varepsilon})$ verifies (22). For an arbitrary $\eta \in (0,1)$, we will first deal with the relation

$$\sup_{0<\varepsilon\leq 1} P[\|X^\varepsilon\|_\gamma > A] \leq \eta, \tag{23}$$

for $A = A_\eta$ large enough, and $1/3 < \gamma < H$. To this purpose, let us recall some basic facts about Sobolev spaces, for which we refer to [1]: for $\alpha \in (0,1)$ and $p \geq 1$, the Sobolev space $\mathcal{W}^{\alpha,p}([0,T]; \mathbb{R}^n)$ is induced by the semi-norm

$$\|f\|_{\alpha,p}^p = \int_0^T \int_0^T \frac{|f(t) - f(s)|^p}{|t-s|^{1+\alpha p}}\, dsdt. \tag{24}$$

Then the Sobolev imbedding theorem states that, if $\alpha p > 1$, then $\mathcal{W}^{\alpha,p}([0,T]; \mathbb{R}^n)$ is continuously imbedded in $\mathcal{C}_1^\gamma(\mathbb{R}^n)$ for any $\gamma < \alpha - 1/p$, where the spaces $\mathcal{C}_1^\gamma$ have been defined by relation (7); notice that, in both (7) and (24), the sup part of the usual Hölder or Sobolev norm has been omitted, but can be recovered since we are dealing with fixed initial conditions. In order to prove (23), it is thus sufficient to check that, for any $p \geq 1$ sufficiently large and $\alpha < H$, the following bound holds true:

$$\sup_{0<\varepsilon\leq 1} E\left[\int_0^T \int_0^T \frac{|X^\varepsilon(t) - X^\varepsilon(s)|^p}{|t-s|^{1+\alpha p}}\, dsdt\right] \leq M_{\alpha,p} < \infty. \tag{25}$$

However, invoking Lemma 3.5, where the function ϕ is defined and bounded, we get, for $\varepsilon > 0$ and $t > s \geq 0$:

$$\begin{aligned} E[|X^{\varepsilon,i}(t) - X^{\varepsilon,i}(s)|^2] &= -2\varepsilon^{2H} \int_0^{\frac{t-s}{\varepsilon}} du \int_0^u \phi(2H, w) dw \\ &\leq 2(t-s)\varepsilon^{2H-1} \int_0^{\frac{t-s}{\varepsilon}} |\phi(2H, w)| dw \\ &\leq \begin{cases} 2(t-s)\varepsilon^{2H-1} \left|\phi\big(2H+1, \frac{t-s}{\varepsilon}\big)\right| & \text{if } t-s \geq 2\varepsilon, \\ c_H (t-s)^2 \varepsilon^{2H-2} & \text{if } t-s \leq 2\varepsilon, \end{cases} \\ &\leq c_H (t-s)^{2H}. \end{aligned}$$

Thus, using the fact that X^ε is a Gaussian process, we deduce that (25) holds for any $\alpha < H$ and p large enough, from which (23) is easily seen.

On the other hand, thanks again to Lemma 3.5, we also have

$$\begin{aligned} E[|\mathbf{X}^{\mathbf{2},\varepsilon}_{st}(i,j)|^2] &= \varepsilon^{4H} \int_0^{\frac{t-s}{\varepsilon}} du \int_0^u dw\, \phi(w) \int_{\frac{s}{\varepsilon}}^{u+\frac{s}{\varepsilon}} dv \int_{v-w}^{v-\frac{s}{\varepsilon}} dz\, \phi(z) \\ &\leq \varepsilon^{4H-2} (t-s)^2 \left(\int_0^{\frac{t-s}{\varepsilon}} \phi(w) dw \right)^2 \leq c_H (t-s)^{4H}. \end{aligned}$$

Since $\mathbf{X}^{\mathbf{2},\varepsilon}$ is a process in the second chaos of the fBm B, on which all L^p norms are equivalent for $p > 1$, we get that

$$E[|\mathbf{X}^{\mathbf{2},\varepsilon}_{st}(i,j)|^p] \leq c_{H,p} |t-s|^{2pH} \text{ for any } s, t \geq 0 \text{ and } \varepsilon > 0. \tag{26}$$

In order to conclude that $\mathbf{X}^2$ verifies the second inequality in (22), let us recall the following inequality from [9]: let $g \in \mathcal{C}_2(V)$ for a given Banach space V; then, for any $\kappa > 0$ and $p \geq 1$ we have

$$\|g\|_\kappa \leq c \left(U_{\kappa+2/p;p}(g) + \|\delta g\|_\gamma \right) \quad \text{with} \quad U_{\gamma;p}(g) = \left(\int_0^T \int_0^T \frac{|g_{st}|^p}{|t-s|^{\gamma p}} ds dt \right)^{1/p}. \tag{27}$$

By plugging inequality (26) into (27) and by recalling that $\delta \mathbf{X}^{2,\varepsilon} = \delta X^\varepsilon \otimes \delta X^\varepsilon$ and inequality (25), we obtain easily the second part of (22).

□

3.2. **FDD convergence.** This subsection is devoted to the second part of the proof of Theorem 1.1, namely the convergence of finite dimensional distributions. More specifically, we shall prove the following:

Proposition 3.7. *Let $(X^\varepsilon, \mathbf{X}^{\mathbf{2},\varepsilon})$ be the approximation process defined by (19). Then,*

$$\text{f.d.d.} - \lim_{\varepsilon \to 0} (X^\varepsilon, \mathbf{X}^{\mathbf{2},\varepsilon}) = (B, \mathbf{B}^{\mathbf{2}}), \tag{28}$$

where f.d.d.$-\lim$ *stands for the convergence in law of the finite dimensional distributions. Otherwise stated, for any $k \geq 1$ and any family $\{s_i, t_i;\, i \leq k, 0 \leq s_i < t_i \leq T\}$, we have*

$$\mathcal{L} - \lim_{\varepsilon \to 0} (X^\varepsilon_{t_1}, \mathbf{X}^{\mathbf{2},\varepsilon}_{s_1 t_1}, \dots, X^\varepsilon_{t_k}, \mathbf{X}^{\mathbf{2},\varepsilon}_{s_k t_k}) = (B^\varepsilon_{t_1}, \mathbf{B}^{\mathbf{2},\varepsilon}_{s_1 t_1}, \dots, B^\varepsilon_{t_k}, \mathbf{B}^{\mathbf{2},\varepsilon}_{s_k t_k}).$$

Proof. The proof is divided in several steps.

(i) Reduction of the problem. For simplicity, we assume that $d = 2$ (the general case can be treated along the same lines, up to some cumbersome notations). Thus, in this particular case, the fdd convergence (28) will be a consequence of the following identity:

$$\big(X^{1,\varepsilon}, X^{2,\varepsilon}, \mathbf{X}^{\mathbf{2},\varepsilon}_{0\cdot}(1,2)\big) \overset{\text{f.d.d.}}{\longrightarrow} \big(B^1, B^2, \mathbf{B}^{\mathbf{2}}_{0\cdot}(1,2)\big). \tag{29}$$

Indeed, we have, on the one hand:

$$\mathbf{X}^{\mathbf{2},\varepsilon}_{st}(i,j) = \mathbf{X}^{\mathbf{2},\varepsilon}_{0t}(i,j) - \mathbf{X}^{\mathbf{2},\varepsilon}_{0s}(i,j) - X^{i,\varepsilon}_s\big(X^{j,\varepsilon}_t - X^{j,\varepsilon}_s\big)$$

and, on the other hand:

$$\mathbf{X}^{\mathbf{2},\varepsilon}_{st}(i,j) + \mathbf{X}^{\mathbf{2},\varepsilon}_{st}(j,i) = \big(X^{i,\varepsilon}_t - X^{i,\varepsilon}_s\big)\big(X^{j,\varepsilon}_t - X^{j,\varepsilon}_s\big),$$

which allows to go from (29) to (28) easily. Observe that the couple $(\mathbf{X}^{\mathbf{2},\varepsilon}_{0\cdot}(1,1), \mathbf{X}^{\mathbf{2},\varepsilon}_{0\cdot}(2,2))$ should also be considered. However, thanks to relation (18), the convergence in law of these components can be easily deduced from the convergence of $(X^{1,\varepsilon}, X^{2,\varepsilon})$.

(ii) Rescaling. According to the scaling property for the fBm, namely the fact that for any $\varepsilon > 0$, we have:

$$(B_{t/\varepsilon})_{t\geq 0} = \varepsilon^{-H}(B_t)_{t\geq 0} \quad \text{in law},$$

we obtain that the triple $(X^{1,\varepsilon}, X^{2,\varepsilon}, \mathbf{X}^{\mathbf{2},\varepsilon}_{0\cdot}(1,2))$ has the same law as

$$\left(\int_0^\cdot \frac{B^1_{u+\varepsilon} - B^1_u}{\varepsilon} du, \int_0^\cdot \frac{B^2_{u+\varepsilon} - B^2_u}{\varepsilon} du, \int_0^\cdot \frac{B^1_{u+\varepsilon} - B^1_u}{\varepsilon}\left(\int_0^u \frac{B^2_{v+\varepsilon} - B^2_v}{\varepsilon} du\right) du\right). \tag{30}$$

(iii) Proof of (29). As we will see, the key point is that all the components of (30) converge in fact in L^2. Thus, we can treat each component separately, for a fixed arbitrary $t \geq 0$. Furthermore, for $i = 1, 2$, we have

$$\int_0^t \frac{B^i_{u+\varepsilon} - B^i_u}{\varepsilon} du = \frac{1}{\varepsilon}\int_t^{t+\varepsilon} B^i_u du - \frac{1}{\varepsilon}\int_0^\varepsilon B^i_u du \overset{\mathrm{L}^2}{\longrightarrow} B^i_t - B^i_0 = B^i_t,$$

by Lebesgue theorem, where the L^2-convergence stands for a convergence in $\mathrm{L}^2(\Omega\times[0,T])$. So, it remains to prove that, again for fixed $t \geq 0$,

$$A_\varepsilon \triangleq \int_0^t \frac{B^1_{u+\varepsilon} - B^1_u}{\varepsilon}\left(\int_0^u \frac{B^2_{v+\varepsilon} - B^2_v}{\varepsilon} dv\right) du \overset{\mathrm{L}^2}{\longrightarrow} \int_0^t B^2_u d^\circ B^1_u, \tag{31}$$

where we recall that the Stratonovich differential $d^\circ B^1$ has been introduced at Definition 2.7.

In the sequel, we denote by $c_{H,t}$ a generic constant (depending only of H and t) that can be different from line to line. In order to prove (29), we shall first show that

$$M_\varepsilon \triangleq \int_{[0,t]^2} |E[(B_{u+\varepsilon} - B_u)(B_{v+\varepsilon} - B_v)]| du dv \leq c_{H,t}\,\varepsilon^{2H+1}, \tag{32}$$

for any $u, v \in [0, T]$. To this purpose, for $\varepsilon, x > 0$, set $D_\varepsilon = \{u, v \in [0, T]^2 : |u - v| \leq \varepsilon\}$ and let φ_ε be the function defined on $\mathbb{R}_+$ by $\varphi_\varepsilon(x) = \frac{1}{2}((x+\varepsilon)^{2H} - |x-\varepsilon|^{2H} - 2\,x^{2H})$.

Then, we have:

$$\begin{aligned} M_\varepsilon &= \int_{D_\varepsilon} |\varphi_\varepsilon(|u-v|)|dudv + \int_{[0,t]^2\setminus D_\varepsilon} |\varphi_\varepsilon(|u-v|)|dudv \\ &\leq c_{H,t}\left(\varepsilon^{2H+1} + \left|\int_{[0,t]^2\setminus D_\varepsilon} \varphi_\varepsilon(|u-v|)dudv\right|\right), \end{aligned}$$

invoking the fact that $\varphi_\varepsilon(x) < 0$ for $x > \varepsilon$ and $|\varphi(x)| \leq c_{H,t}\,\varepsilon^{2H}$ for $0 \leq x \leq \varepsilon$. Hence,

$$\begin{aligned} M_\varepsilon &\leq c_{H,t}\left(\varepsilon^{2H+1} + \left|\int_{[0,t]^2} \varphi_\varepsilon(|u-v|)dudv\right|\right) \\ &= c_{H,t}\left(\varepsilon^{2H+1} + \left|\int_{[0,t]^2} E[(B_{u+\varepsilon}-B_u)(B_{v+\varepsilon}-B_v)]dudv\right|\right) \\ &= c_{H,t}\left(\varepsilon^{2H+1} + E\left|\int_0^t (B_{u+\varepsilon}-B_u)du\right|^2\right) \\ &= c_{H,t}\left(\varepsilon^{2H+1} + E\left|\int_t^{t+\varepsilon} B_u du - \int_0^\varepsilon B_u du\right|^2\right) \leq c_{H,t}\,\varepsilon^{2H+1}, \end{aligned}$$

where the last step is obtained thanks to Jensen's inequality. This proves our announced claim (32). Set now, for $\varepsilon > 0$:

$$B_\varepsilon \triangleq (2\varepsilon)^{-1}\int_0^t (B^1_{u+\varepsilon} - B^1_u)(B^2_u + B^2_{u+\varepsilon})du.$$

We shall prove that $\mathrm{L}^2 - \lim_{\varepsilon\to 0}(A_\varepsilon - B_\varepsilon) = 0$. Observe that $A_\varepsilon - B_\varepsilon$ is equal to:

$$\begin{aligned} A_\varepsilon - B_\varepsilon &= \varepsilon^{-1}\int_0^t (B^1_{u+\varepsilon} - B^1_u)\left(\varepsilon^{-1}\int_0^u dv(B^2_{v+\varepsilon} - B^2_v) - \frac{B^2_u + B^2_{u+\varepsilon}}{2}\right)du \\ &= \varepsilon^{-2}\int_0^t (B^1_{u+\varepsilon} - B^1_u)\left(-\int_0^\varepsilon B^2_v dv + \int_u^{u+\varepsilon}\left[B^2_v - \frac{B^2_u + B^2_{u+\varepsilon}}{2}\right]dv\right)du \\ &= -C_\varepsilon + D_\varepsilon, \end{aligned}$$

where

$$\begin{aligned} C_\varepsilon &= \varepsilon^{-2}\left(\int_t^{t+\varepsilon} B^1_u du - \int_0^\varepsilon B^1_u du\right)\int_0^\varepsilon B^2_v dv \\ D_\varepsilon &= \varepsilon^{-2}\int_0^t (B^1_{u+\varepsilon} - B^1_u)\left(\int_u^{u+\varepsilon}\left[B^2_v - \frac{B^2_u + B^2_{u+\varepsilon}}{2}\right]dv\right)du. \end{aligned}$$

Then, on the one hand, we have:

$$\begin{aligned} E\big[|C_\varepsilon|^2\big] &= \varepsilon^{-4}\,E\left[\left(\int_t^{t+\varepsilon} B^1_u du - \int_0^\varepsilon B^1_u du\right)^2\right]E\left[\left(\int_0^\varepsilon B^2_v dv\right)^2\right] \\ &\leq c_{H,t}\,\varepsilon^{-1}\int_0^\varepsilon E\big[|B^2_v|^2\big]dv = c_{H,t}\,\varepsilon^{-1}\int_0^\varepsilon v^{2H}dv = c_{H,t}\,\varepsilon^{2H}. \end{aligned}$$

In order to treat the term D_ε, write $L_\varepsilon(u,w)$ for the process $B^2_w - \frac{1}{2}(B^2_u + B^2_{u+\varepsilon})$. Then, by independence, we have that $E[|D_\varepsilon|^2]$ satisfies:

$$\begin{aligned} E[|D_\varepsilon|^2] &= \varepsilon^{-4}\int_{[0,t]^2} dudv E[(B^1_{u+\varepsilon}-B^1_u)(B^1_{v+\varepsilon}-B^1_v)]\int_u^{u+\varepsilon} dw \int_v^{v+\varepsilon} dz E\left[L_\varepsilon(u,w)L_\varepsilon(v,z)\right] \\ &\leq \varepsilon^{-4}\int_{[0,t]^2} dudv \left|E[(B^1_{u+\varepsilon}-B^1_u)(B^1_{v+\varepsilon}-B^1_v)]\right| \int_u^{u+\varepsilon} dw \int_v^{v+\varepsilon} dz E\left[|L_\varepsilon(u,w)|\,|L_\varepsilon(v,z)|\right] \\ &\leq c_{H,t}\,\varepsilon^{2H-2}\int_{[0,t]^2} dudv \left|E[(B^1_{u+\varepsilon}-B^1_u)(B^1_{v+\varepsilon}-B^1_v)]\right| \leq c_{H,t}\,\varepsilon^{4H-1}, \end{aligned}$$

where the last inequality is due to (32). Thus, since $H > \frac{1}{3}$, we deduce that $\mathrm{L}^2 - \lim_{\varepsilon\to 0} D_\varepsilon = 0$. Recall now that $A_\varepsilon - B_\varepsilon = C_\varepsilon + D_\varepsilon$. Hence, since we have proved that $\mathrm{L}^2 - \lim_{\varepsilon\to 0} C_\varepsilon = \mathrm{L}^2 - \lim_{\varepsilon\to 0} D_\varepsilon = 0$, the convergence (31) holds if and only if the convergence $\mathrm{L}^2 - \lim_{\varepsilon\to 0} B_\varepsilon = \int_0^t B^2_u d^\circ B^1_u$ holds. This is also equivalent, using an obvious change of variable, to prove that

$$E_\varepsilon \triangleq (2\varepsilon)^{-1}\int_0^t B^2_u (B^1_{u+\varepsilon} - B^1_{(u-\varepsilon)\vee 0})du \xrightarrow{\mathrm{L}^2} \int_0^T B^2_u d^\circ B^1_u = \delta^{\mathfrak{S},B^1}(B^2). \tag{33}$$

To this purpose, observe that, since B^1 and B^2 are independent, we have:

$$E_\varepsilon = \delta^{\mathfrak{S},B^1}\left(B^{2,\varepsilon}\right), \quad \text{with} \quad B^{2,\varepsilon} \triangleq (2\varepsilon)^{-1}\int_0^T B^2_u \mathbf{1}_{[(u-\varepsilon)\vee 0, u+\varepsilon]}(\cdot)du,$$

where $\delta^{\mathfrak{S},B^1}(B^{2,\varepsilon})$ has to be interpreted as a Wiener integral with respect to B^1 for a fixed path of B^2 (see Section 2.3). Thus,

$$\begin{aligned} E\left[\left|\delta^{\mathfrak{S},B^1}(B^{2,\varepsilon}) - \delta^{\mathfrak{S},B^1}(B^2)\right|^2\right] &= E\left[\left\|B^{2,\varepsilon} - B^2\right\|^2_{\mathcal{H}}\right] \\ &\leq E\left[\left\|B^{2,\varepsilon} - B^2\right\|^2_{\mathcal{H}}\right] + E\left[\left\|D^2 B^{2,\varepsilon} - D^2 B^2\right\|^2_{\mathcal{H}\otimes\mathcal{H}}\right] \\ &= E\left[\left|\delta^{\mathfrak{S},B^2}(B^{2,\varepsilon}) - \delta^{\mathfrak{S},B^2}(B^2)\right|^2\right]. \end{aligned}$$

Finally, notice that it is proved in [2, Theorem 2] (see also Section 4 of this latter reference) that $\lim_{\varepsilon\to 0} E[|\delta^{\mathfrak{S},B^2}(B^{2,\varepsilon}) - \delta^{\mathfrak{S},B^2}(B^2)|^2] = 0$. Thus (33) holds true, and this finishes the proof of Proposition 3.7. □

References

[1] R. A. Adams (1975): *Sobolev spaces.* Academic Press.

[2] E. Alòs, J. A. Léon and D. Nualart (2001): *Stratonovich stochastic calculus for fractional Brownian motion with Hurst parameter lesser that 1/2.* Taiwanese J. Math. **5** (3), 609-632.

[3] X. Bardina, M. Jolis and C.A. Tudor (2003): *Convergence in law to the multiple fractional integral.* Stochastic Process. Appl. **105** (2), 315-344.

[4] F. Baudoin and L. Coutin (2007): *Operators associated with a stochastic differential equation driven by fractional Brownian motions.* Stochastic Process. Appl. **117** (5), 550-574.

[5] L. Coutin and Z. Qian (2002): *Stochastic rough path analysis and fractional Brownian motion.* Probab. Theory Relat. Fields **122**, 108-140.

[6] R. Delgado and M. Jolis (2000): *Weak approximation for a class of Gaussian processes.* J. Appl. Probab. **37** (2), 400-407.

[7] N. Enriquez (2004): *A simple construction of the fractional Brownian motion.* Stochastic Process. Appl. **109** (2), 203-223

[8] A. Garsia, E. Rodemich and H. Rumsey (1970): *A real variable lemma and the continuity of paths of Gaussian processes.* Indiana U. Math. J. V. **20**, 565-578.
[9] M. Gubinelli (2004): *Controlling rough paths.* J. Funct. Anal. **216**, 86-140.
[10] M. Gubinelli and S. Tindel (2007): *Rough evolution equation.* In preparation.
[11] Y. Hu and D. Nualart (2006): *Rough path analysis via fractional calculus.* Preprint.
[12] J. Jeon, K. Lee and Y.T. Kim (2004): *A Wong-Zakai type approximation for multiple Wiener-Stratonovich integrals.* Stochastic Anal. Appl. **22**, no. 5, 1257-1271.
[13] T. G. Kurtz and P.E. Protter (1996): *Weak convergence of stochastic integrals and differential equations.* Probabilistic models for nonlinear partial differential equations (Montecatini Terme, 1995), Lecture Notes in Math. **1627**, Springer, 1-41.
[14] M. Ledoux, Z. Qian and T. Zhang (2002): *Large deviations and support theorem for diffusion processes via rough paths.* Stochastic Process. Appl. **102** (2), 265-283.
[15] A. Lejay (2003): *An Introduction to Rough Paths.* Séminaire de probabilités XXXVII, vol. **1832** of Lecture Notes in Mathematics, 1-59.
[16] T. Lyons and Z. Qian (2002): *System control and rough paths.* Oxford University Press.
[17] R. Marty (2005): *Asymptotic behavior of differential equations driven by periodic and random processes with slowly decaying correlations.* ESAIM: Probability and Statistics **9**, 165-184.
[18] A. Millet and M. Sanz-Solé (2006): *Large deviations for rough paths of the fractional Brownian motion.* Ann. Inst. H. Poincaré Probab. Statist. **42** (2), 245-271.
[19] A. Neuenkirch, I. Nourdin, A. Rößler and S. Tindel (2006): *Trees and asymptotic developments for fractional diffusion processes.* Preprint.
[20] I. Nourdin and T. Simon (2006): *Correcting Newton-Côtes integrals by Lévy areas.* Bernoulli **13** (3), 695-711.
[21] D. Nualart (2003): *Stochastic calculus with respect to the fractional Brownian motion and applications.* Contemporary Mathematics **336**, 3-39.
[22] F. Russo and P. Vallois (1993): *Forward, backward and symmetric stochastic integration.* Probab. Theory Relat. Fields **97**, 403-421.
[23] T. Sottinen (2001): *Fractional Brownian motion, random walks and binary market models.* Finance Stochast. **5**, 343-355.
[24] G. Tessitore and J. Zabczyk (2006): *Wong-Zakai approximations of stochastic evolution equations.* J. Evol. Equ. **6**, no. 4, 621-655.
[25] E. Wong and M. Zakai (1965): *On the relation between ordinary and stochastic differential equations.* Internat. J. Engrg. Sci. **3**, 213–229.

Ivan Nourdin: Laboratoire de Probabilités et Modèles Aléatoires, Université Pierre et Marie Curie, Boîte courrier 188, 4 Place Jussieu, 75252 Paris Cedex 5, France. *Email:* `nourdin@ccr.jussieu.fr`

Samy Tindel: Institut Élie Cartan Nancy, B.P. 239, 54506 Vandœuvre-lès-Nancy Cedex, France. *Email:* `tindel@iecn.u-nancy.fr`