NONPARAMETRIC ESTIMATION IN FRACTIONAL SDE

FABIENNE COMTE* AND NICOLAS MARIE†

Abstract. This paper deals with the consistency and a rate of convergence for a Nadaraya-Watson estimator of the drift function of a stochastic differential equation driven by an additive fractional noise. The results of this paper are obtained via both some long-time behavior properties of Hairer and some properties of the Skorokhod integral with respect to the fractional Brownian motion. These results are illustrated on the fractional Ornstein-Uhlenbeck process.

Contents

1. Introduction

Consider the stochastic differential equation

$$X(t) = X_0 + \int_0^t b(X(s))ds + \sigma B(t), \tag{1}$$

where B is a fractional Brownian motion of Hurst index $H \in]1/2, 1[$, $b : \mathbb{R} \to \mathbb{R}$ is a continuous map and $\sigma \in \mathbb{R}^*$.

Along the last two decades, many authors studied statistical inference from observations drawn from stochastic differential equations driven by fractional Brownian motion.

Most references on the estimation of the trend component in Equation (1) deals

with parametric estimators. In Kleptsyna and Le Breton [10] and Hu and Nualart [12], estimators of the trend component in Langevin's equation are studied. Kleptsyna and Le Breton [10] provide a maximum likelihood estimator, where the stochastic integral with respect to the solution of Equation (1) returns to an Itô integral. In [26], Tudor and Viens extend this estimator to equations with a drift function depending linearly on the unknown parameter. Hu and Nualart [12] provide a least square estimator, where the stochastic integral with respect to the solution of Equation (1) is taken in the sense of Skorokhod. In [13], Hu, Nualart and Zhou extend this estimator to equations with a drift function depending linearly on the unknown parameter.
In Tindel and Neuenkirch [17], the authors study a least square-type estimator defined by an objective function tailor-maid with respect to the main result of Tudor and Viens [27] on the rate of convergence of the quadratic variation of the fractional Brownian motion. In [4], Chronopoulou and Tindel provide a likelihood based numerical procedure to estimate a parameter involved in both the drift and the volatility functions in a stochastic differential equation with multiplicative fractional noise.
On the nonparametric estimation of the trend component in Equation (1), there are only few references. Saussereau [23] and Mishra and Prakasa Rao [18] study the consistency of some Nadaraya-Watson's-type estimators of the drift function b in Equation (1). On the nonparametric estimation in Itô's calculus framework, the reader is referred to Kutoyants [14].

Let $K : \mathbb{R} \to \mathbb{R}_+$ be a kernel that is a nonnegative function with integral equal to 1. The paper deals with the consistency and a rate of convergence for the Nadaraya-Watson estimator

$$\widehat{b}_{T,h}(x) := \frac{\displaystyle\int_0^T K\left(\frac{X(s)-x}{h}\right)\delta X(s)}{\displaystyle\int_0^T K\left(\frac{X(s)-x}{h}\right)ds} \text{ ; } x \in \mathbb{R} \tag{2}$$

of the drift function b in Equation (1), where the stochastic integral with respect to X is taken in the sense of Skorokhod. Since to compute the Skorokhod integral is a challenge, by denoting by X_{x_0} the solution of Equation (1) with initial condition $x_0 \in \mathbb{R}$, the following estimator is also studied:

$$\begin{aligned}\widehat{b}_{T,h,\varepsilon}(x) := & \frac{\displaystyle\int_0^T K\left(\frac{X_{x_0}(s)-x}{h}\right)dX_{x_0}(s)}{\displaystyle\int_0^T K\left(\frac{X_{x_0}(s)-x}{h}\right)ds} \\ & -\alpha_H\sigma^2\frac{\displaystyle\frac{1}{h}\int_0^T\int_0^u K'\left(\frac{X_{x_0}(u)-x}{h}\right)\frac{X_{x_0+\varepsilon}(u)-X_{x_0}(u)}{X_{x_0+\varepsilon}(v)-X_{x_0}(v)}|u-v|^{2H-2}dvdu}{\displaystyle\int_0^T K\left(\frac{X_{x_0}(s)-x}{h}\right)ds}\end{aligned} \tag{3}$$

with $\varepsilon > 0$ and $x \in \mathbb{R}$. In this second estimator, the stochastic integral is taken pathwise. It depends on H, but an estimator of this parameter is for instance provided in Kubilius and Skorniakov [11].
As detailed in Subsection 2.2, the Skorokhod integral is defined via the divergence operator which is the adjoint of the Malliavin derivative for the fractional Brownian motion. If $H = 1/2$, then the Skorokhod integral coincides with Itô's integral on its domain. When $H \in]1/2,1[$, it is more difficult to compute the Skorokhod integral, but not impossible as explained at the end of Subsection 2.2. Note that,

the pathwise stochastic integral defined in Subsection 2.1 would have been a more natural choice, but unfortunately, it does not provide a consistent estimator (see Proposition 3.3).
Clearly, to be computable, the estimator $\widehat{b}_{T,h,\varepsilon}(x)$ requires an observed path of the solution of Equation (1) for two close but different values of the initial condition. This is not possible in any context, but we have in mind the following application field: if $t \mapsto X_{x_0}(\omega, t)$ denotes the concentration of a drug along time during its elimination by a patient ω with initial dose $x_0 > 0$, $t \mapsto X_{x_0+\varepsilon}(\omega, t)$ could be approximated by replicating the exact same protocol on patient ω, but with initial dose $x_0 + \varepsilon$ after the complete elimination of the previous dose.
We mention that we do not study the additional error which occurs when only discrete time observations with step Δ on $[0, T]$ ($T = n\Delta$) are available. Formula (3) has then to be discretized and a study in the spirit of Saussereau [23] (Section 4.3) must be conducted.
Section 2 deals with some preliminary results on stochastic integrals with respect to the fractional Brownian motion and an ergodic theorem for the solution of Equation (1). The consistency and a rate of convergence of the Nadaraya-Watson estimator studied in this paper are stated in Section 3. Almost all the proofs of the paper are provided in Section 4.

Notations:

(1) The vector space of Lipschitz continuous maps from $\mathbb{R}$ into itself is denoted by $\text{Lip}(\mathbb{R})$ and equipped with the Lipschitz semi-norm $\|.\|_{\text{Lip}}$ defined by
$$\|\varphi\|_{\text{Lip}} := \sup\left\{\frac{|\varphi(y) - \varphi(x)|}{|y - x|}\ ;\ x, y \in \mathbb{R} \text{ and } x \neq y\right\}$$
for every $\varphi \in \text{Lip}(\mathbb{R})$.

(2) For every $m \in \mathbb{N}$,
$$C_b^m(\mathbb{R}) := \left\{\varphi \in C^m(\mathbb{R}) : \max_{k \in [\![0,m]\!]} \|\varphi^{(k)}\|_\infty < \infty\right\}.$$

(3) For every $m \in \mathbb{N}^*$,
$$\text{Lip}_b^m(\mathbb{R}) := \left\{\varphi \in C^m(\mathbb{R}) : \varphi \in \text{Lip}(\mathbb{R}) \text{ and } \max_{k \in [\![1,m]\!]} \|\varphi^{(k)}\|_\infty < \infty\right\}$$
and for every $\varphi \in \text{Lip}_b^m(\mathbb{R})$,
$$\|\varphi\|_{\text{Lip}_b^m} := \|\varphi\|_{\text{Lip}} \vee \max_{k \in [\![1,m]\!]} \|\varphi^{(k)}\|_\infty.$$
The map $\|.\|_{\text{Lip}_b^m}$ is a semi-norm on $\text{Lip}_b^m(\mathbb{R})$.

Note that for every $m \in \mathbb{N}^*$,
$$C_b^m(\mathbb{R}) \subset \text{Lip}_b^m(\mathbb{R}).$$

(4) Consider $n \in \mathbb{N}^*$. The vector space of infinitely continuously differentiable maps $f : \mathbb{R}^n \to \mathbb{R}$ such that f and all its partial derivatives have polynomial growth is denoted by $C_p^\infty(\mathbb{R}^n, \mathbb{R})$.

2. Stochastic integrals with respect to the fractional Brownian motion and an ergodic theorem for fractional SDE

On the one hand, this section presents two different methods to define a stochastic integral with respect to the fractional Brownian motion. The first one is based on the pathwise properties of the fractional Brownian motion. Even if this approach is very natural, it is proved in Proposition 3.3 that the pathwise stochastic integral is not appropriate to get a consistent estimator of the drift function b in Equation (1). Another stochastic integral with respect to the fractional Brownian motion is defined via the Malliavin divergence operator. This stochastic integral is called Skorokhod's integral with respect to B. If $H = 1/2$, which means that B is a Brownian motion, the Skorokhod integral defined via the divergence operator coincides with Itô's integral on its domain. This integral is appropriate for the estimation of the drift function b in Equation (1). On the other hand, an ergodic theorem for the solution of Equation (1) is stated in Subsection 2.3.

2.1. **The pathwise stochastic integral.** This subsection deals with some definitions and basic properties of the pathwise stochastic integral with respect to the fractional Brownian motion of Hurst index greater than 1/2.

Definition 2.1. *Consider x and w two continuous functions from $[0,T]$ into $\mathbb{R}$. Consider a partition $D := (t_k)_{k\in[\![0,m]\!]}$ of $[s,t]$ with $m \in \mathbb{N}^*$ and $s,t \in [0,T]$ such that $s < t$. The Riemann sum of x with respect to w on $[s,t]$ for the partition D is*

$$J_{x,w,D}(s,t) := \sum_{k=0}^{m-1} x(t_k)(w(t_{k+1}) - w(t_k)).$$

Notation. With the notations of Definition 2.1, the mesh of the partition D is

$$\delta(D) := \max_{k\in[\![0,m-1]\!]} |t_{k+1} - t_k|.$$

The following theorem ensures the existence and the uniqueness of Young's integral (see Friz and Victoir [6], Theorem 6.8).

Theorem 2.2. *Let x (resp. w) be a α-Hölder (resp. β-Hölder) continuous map from $[0,T]$ into $\mathbb{R}$ with $\alpha,\beta \in]0,1]$ such that $\alpha + \beta > 1$. There exists a unique continuous map $J_{x,w} : [0,T] \to \mathbb{R}$ such that for every $s,t \in [0,T]$ satisfying $s < t$ and any sequence $(D_n)_{n\in\mathbb{N}}$ of partitions of $[s,t]$ such that $\delta(D_n) \to 0$ as $n \to \infty$,*

$$\lim_{n\to\infty} |J_{x,w}(t) - J_{x,w}(s) - J_{x,w,D_n}(s,t)| = 0.$$

The map $J_{x,w}$ is the Young integral of x with respect to w and $J_{x,w}(t) - J_{x,w}(s)$ is denoted by

$$\int_s^t x(u)dw(u)$$

for every $s,t \in [0,T]$ such that $s < t$.

The following proposition is a change of variable for Young's integral.

Proposition 2.3. *Let x be a α-Hölder continuous map from $[0,T]$ into $\mathbb{R}$ with $\alpha \in]1/2,1[$. For every $\varphi \in \mathrm{Lip}_b^1(\mathbb{R})$ and $s,t \in [0,T]$ such that $s < t$,*

$$\varphi(x(t)) - \varphi(x(s)) = \int_s^t \varphi'(x(u))dx(u).$$

For any $\alpha \in]1/2,H[$, the paths of B are α-Hölder continuous (see Nualart [20], Section 5.1). So, for every process $Y := (Y(t))_{t\in[0,T]}$ with β-Hölder continuous paths from $[0,T]$ into $\mathbb{R}$ such that $\alpha + \beta > 1$, by Theorem 2.2, it is natural to define the pathwise stochastic integral of Y with respect to B by

$$\left(\int_0^t Y(s)dB(s)\right)(\omega) := \int_0^t Y(\omega,s)dB(\omega,s)$$

for every $\omega \in \Omega$ and $t \in [0,T]$.

2.2. **The Skorokhod integral.** This subsection deals with some definitions and results on Malliavin calculus in order to define and to provide a suitable expression of Skorokhod's integral.

Consider the vector space

$$\mathcal{H} := \left\{\varphi : \mathbb{R}_+ \to \mathbb{R} : \int_0^{\infty}\int_0^{\infty} |t-s|^{2H-2}|\varphi(s)|\cdot|\varphi(t)|dsdt < \infty\right\}.$$

Equipped with the scalar product

$$\langle\varphi,\psi\rangle_{\mathcal{H}} := H(2H-1)\int_0^{\infty}\int_0^{\infty}|t-s|^{2H-2}\varphi(s)\psi(t)dsdt \ ; \ \varphi,\psi\in\mathcal{H},$$

$\mathcal{H}$ is the reproducing kernel Hilbert space of B. Let $\mathbf{B}$ be the map defined on $\mathcal{H}$ by

$$\mathbf{B}(h) := \int_0^{\cdot} h(s)dB(s) \ ; \ h \in \mathcal{H}$$

which is the Wiener integral of h with respect to B. The family $(\mathbf{B}(h))_{h\in\mathcal{H}}$ is an isonormal Gaussian process.

Definition 2.4. *The Malliavin derivative of a smooth functional*

$$F = f(\mathbf{B}(h_1),\dots,\mathbf{B}(h_n))$$

where $n\in\mathbb{N}^*$, $f\in C_p^{\infty}(\mathbb{R}^n,\mathbb{R})$ and $h_1,\dots,h_n\in\mathcal{H}$ is the $\mathcal{H}$-valued random variable

$$\mathbf{D}F := \sum_{k=1}^{n}\partial_k f(\mathbf{B}(h_1),\dots,\mathbf{B}(h_n))h_k.$$

Proposition 2.5. *The map $\mathbf{D}$ is closable from $L^2(\Omega,\mathcal{A},\mathbb{P})$ into $L^2(\Omega;\mathcal{H})$. Its domain in $L^2(\Omega,\mathcal{A},\mathbb{P})$ is denoted by $\mathbb{D}^{1,2}$ and is the closure of the smooth functionals space for the norm $\|.\|_{1,2}$ defined by*

$$\|F\|_{1,2}^2 := \mathbb{E}(|F|^2) + \mathbb{E}(\|\mathbf{D}F\|_{\mathcal{H}}^2) < \infty$$

for every $F\in L^2(\Omega,\mathcal{A},\mathbb{P})$.

For a proof, see Nualart [20], Proposition 1.2.1.

Definition 2.6. *The adjoint δ of the Malliavin derivative $\mathbf{D}$ is the divergence operator. The domain of δ is denoted by* $\mathrm{dom}(\delta)$ *and $u\in\mathrm{dom}(\delta)$ if and only if there exists a deterministic constant $c>0$ such that for every $F\in\mathbb{D}^{1,2}$,*

$$|\mathbb{E}(\langle\mathbf{D}F,u\rangle_{\mathcal{H}})| \leqslant c\mathbb{E}(|F|^2)^{1/2}.$$

For every process $Y := (Y(s))_{s\in\mathbb{R}_+}$ and every $t>0$, if $Y\mathbf{1}_{[0,t]}\in\mathrm{dom}(\delta)$, then its Skorokhod integral with respect to B is defined on $[0,t]$ by

$$\int_0^t Y(s)\delta B(s) := \delta(Y\mathbf{1}_{[0,t]}).$$

With the same notations:

$$\int_0^t Y(s)\delta X(s) := \int_0^t Y(s)b(X(s))ds + \sigma\int_0^t Y(s)\delta B(s).$$

The following proposition provides the link between the Skorokhod integral and the pathwise stochastic integral of Subsection 2.1.

Proposition 2.7. *If $b \in \text{Lip}_b^1(\mathbb{R})$, then Equation (1) with initial condition $x \in \mathbb{R}$ has a unique solution X_x with α-Hölder continuous paths for every $\alpha \in]0, H[$. Moreover, for every $\varphi \in \text{Lip}_b^1(\mathbb{R})$,*

$$(4) \quad \int_0^t \varphi(X_x(u))\delta X_x(u) = \int_0^t \varphi(X_x(u))dX_x(u) - \alpha_H\sigma^2 \int_0^t \int_0^u \varphi'(X_x(u))\frac{\partial_x X_x(u)}{\partial_x X_x(v)}|u-v|^{2H-2}dvdu,$$

where $\alpha_H = H(2H-1)$.

Moreover, we can prove the following Corollary, which allows us to propose a computable form for the estimator.

Corollary 2.8. *Assume that $b \in \text{Lip}_b^2(\mathbb{R})$ and there exists a constant $M > 0$ such that*

$$b'(x) \leqslant -M \ ; \forall x \in \mathbb{R}.$$

For every $\varphi \in \text{Lip}_b^1(\mathbb{R})$, $x \in \mathbb{R}$ and $\varepsilon, t > 0$,

$$\left|\int_0^t \varphi(X_x(u))\delta X_x(u) - S_\varphi(x,\varepsilon,t)\right| \leqslant C_\varphi \varepsilon t^{2H-1},$$

where

$$S_\varphi(x,\varepsilon,t) := \int_0^t \varphi(X_x(u))dX_x(u) - \alpha_H\sigma^2 \int_0^t \int_0^u \varphi'(X_x(u))\frac{X_{x+\varepsilon}(u) - X_x(u)}{X_{x+\varepsilon}(v) - X_x(v)}|u-v|^{2H-2}dvdu$$

and

$$C_\varphi := H\sigma^2 \frac{\|b''\|_\infty \|\varphi'\|_\infty}{2M^2}.$$

As mentioned in the Introduction, the formula for $S_\varphi(x,\varepsilon,t)$ can be used if two paths of X can be observed with different but close initial conditions.
Lastly, the following theorem, recently proved by Hu, Nualart and Zhou in [13] (see Proposition 4.4), provides a suitable control of Skorokhod's integral to study its long-time behavior.

Theorem 2.9. *Assume that $b \in \text{Lip}_b^2(\mathbb{R})$ and there exists a constant $M > 0$ such that*

$$b'(x) \leqslant -M \ ; \forall x \in \mathbb{R}.$$

There exists a deterministic constant $C > 0$, not depending on T, such that for every $\varphi \in \text{Lip}_b^1(\mathbb{R})$:

$$\mathbb{E}\left(\left|\int_0^T \varphi(X(s))\delta B(s)\right|^2\right) \leqslant C\left(\left(\int_0^T \mathbb{E}(|\varphi(X(s))|^{1/H})ds\right)^{2H} + \left(\int_0^T \mathbb{E}(|\varphi'(X(s))|^2)^{1/(2H)}ds\right)^{2H}\right) < \infty.$$

2.3. **Ergodic theorem for the solution of a fractional SDE.** On the ergodicity of fractional SDEs, the reader can refer to Hairer [7], Hairer and Ohashi [8] and Hairer and Pillai [9] (see Subsection 4.3 for details).

In the sequel, the map b fulfills the following condition.

Assumption 2.10. *The map b belongs to $\text{Lip}_b^{\infty}(\mathbb{R})$ and there exists a constant $M > 0$ such that*

$$b'(x) \leqslant -M \ ; \forall x \in \mathbb{R}. \tag{5}$$

Remarks:

(1) Since $b \in \text{Lip}_b^1(\mathbb{R})$, Equation (1) has a unique solution.
(2) Under Assumption 2.10, the dissipativity conditions of Hairer [7], Hairer and Ohashi [8] and Hu, Nualart and Zhou [13] are fulfilled by b:

$$(x-y)(b(x)-b(y)) \leqslant -M(x-y)^2 \ ; \forall x, y \in \mathbb{R}$$

and there exists a constant $M' > 0$ such that

$$xb(x) \leqslant M'(1-x^2) \ ; \forall x \in \mathbb{R}.$$

Therefore, Assumption 2.10 is sufficient to apply the results proved in [7], [8] and [13] in the sequel.

Proposition 2.11. *Consider a measurable map $\varphi : \mathbb{R} \to \mathbb{R}_+$ such that there exists a nonempty compact subset C of $\mathbb{R}$ satisfying $\varphi(C) \subset]0, \infty[$. Under Assumption 2.10, there exists a deterministic constant $l(\varphi) > 0$ such that*

$$\frac{1}{T}\int_0^T \varphi(X(t))dt \xrightarrow[T\to\infty]{\text{a.s.}/\text{L}^2} l(\varphi) > 0.$$

3. Convergence of the Nadaraya-Watson estimator of the drift function

This section deals with the consistency and rate of convergence of the Nadaraya-Watson estimator of the drift function b in Equation (1).

In the sequel, the kernel K fulfills the following assumption.

Assumption 3.1. $\text{supp}(K) = [-1, 1]$ *and* $K \in C_b^1(\mathbb{R}, \mathbb{R}_+)$.

3.1. **Why is pathwise integral inadequate.** First of all, let us prove that, even if it seems very natural, the pathwise Nadaraya-Watson estimator

$$\widetilde{b}_{T,h}(x) := \frac{\displaystyle\int_0^T K\left(\frac{X(s)-x}{h}\right) dX(s)}{\displaystyle\int_0^T K\left(\frac{X(s)-x}{h}\right) ds} = \frac{\displaystyle\frac{1}{Th}\int_0^T K\left(\frac{X(s)-x}{h}\right) dX(s)}{\widehat{f}_{T,h}(x)}$$

where

$$\widehat{f}_{T,h}(x) := \frac{1}{Th}\int_0^T K\left(\frac{X(s)-x}{h}\right) ds. \tag{6}$$

is not consistent.

For this, we need the following lemma providing a convergence result for $\widehat{f}_{T,h}(x)$. It will also be used to prove Proposition 3.4.

Lemma 3.2. *Under Assumptions 2.10 and 3.1, there exists a deterministic constant $l_h(x) > 0$ such that*

$$\widehat{f}_{T,h}(x) \xrightarrow[T\to\infty]{\text{a.s.}/\text{L}^2} l_h(x) > 0.$$

Proof. Under Assumption 3.1, the map

$$y \in \mathbb{R} \longmapsto \frac{1}{h}K\left(\frac{y-x}{h}\right)$$

satisfies the condition on φ of Proposition 2.11, which applies thus here and gives the result. □

Now, we state the result proving that $\widetilde{b}_{T,h}(x)$ is not consistent to recover $b(x)$.

Proposition 3.3. *Under Assumptions 2.10 and 3.1:*

$$\widetilde{b}_{T,h}(x) \xrightarrow[T\rightarrow\infty]{\mathbb{P}} 0.$$

Proof. Let $\mathcal{K}$ be a primitive function of K. By the change of variable formula for Young's integral (Proposition 2.3):

$$\begin{aligned}\mathcal{K}\left(\frac{X(T)-x}{h}\right) - \mathcal{K}\left(\frac{X(0)-x}{h}\right) &= \frac{1}{h}\int_0^T K\left(\frac{X(s)-x}{h}\right)dX(s)\\ &= T\widehat{f}_{T,h}(x)\widetilde{b}_{T,h}(x).\end{aligned}$$

Then,

$$\widetilde{b}_{T,h}(x) = \frac{1}{T\widehat{f}_{T,h}(x)}\left(\mathcal{K}\left(\frac{X(T)-x}{h}\right) - \mathcal{K}\left(\frac{X(0)-x}{h}\right)\right).$$

Since $\mathcal{K}$ is differentiable with bounded derivative K:

$$|\widetilde{b}_{T,h}(x)| \leqslant \frac{\|K\|_\infty}{Th\widehat{f}_{T,h}(x)}|X(T)-X(0)|.$$

Finally, as we know by Hairer [7], Proposition 3.12 that

$$t \in \mathbb{R}_+ \longmapsto \mathbb{E}(|X(t)|)$$

is uniformly bounded, and by Lemma 3.2 that $\widehat{f}_{T,h}(x)$ converges almost surely to $l_h(x) > 0$ as $T \rightarrow \infty$, it follows that $\widetilde{b}_{T,h}(x)$ converges to 0 in probability, when $T \rightarrow \infty$. □

This is why the Skorokhod integral replaces the pathwise stochastic integral in $\widehat{b}_{T,h}(x)$.

3.2. **Convergence of the Nadaraya-Watson estimator.** This subsection deals with the consistency and rate of convergence of the estimators.

The Nadaraya-Watson estimator $\widehat{b}_{T,h}(x)$ defined by Equation (2) can be decomposed as follows:

$$\widehat{b}_{T,h}(x) - b(x) = \frac{B_{T,h}(x)}{\widehat{f}_{T,h}(x)} + \frac{S_{T,h}(x)}{\widehat{f}_{T,h}(x)}, \quad (7)$$

where $\widehat{f}_{T,h}(x)$ is defined by (6),

$$B_{T,h}(x) := \frac{1}{Th}\int_0^T K\left(\frac{X(s)-x}{h}\right)(b(X(s))-b(x))ds.$$

and

$$S_{T,h}(x) := \frac{\sigma}{Th}\int_0^T K\left(\frac{X(s)-x}{h}\right)\delta B(s).$$

By using the Lipschitz assumption 2.10 on b together with the technical lemmas proved in Section 2, the estimators $\widehat{b}_{T,h}(x)$ and $\widehat{b}_{T,h,\varepsilon}(x)$ can be studied.

Proposition 3.4. *Under Assumptions 2.10 and 3.1,*

$$|\widehat{b}_{T,h}(x)-b(x)|\leqslant\|b\|_{\text{Lip}}h+\frac{|S_{T,h}(x)|}{\widehat{f}_{T,h}(x)},$$

and there exists a positive constant C such that

$$\mathbb{E}(S_{T,h}(x)^2)\leqslant\frac{C}{h^4T^{2(1-H)}}.$$

As a consequence, for fixed $h>0$, we have

$$T^{\beta}V_{T,h}(x)\xrightarrow[T\to\infty]{\mathbb{P}}0\ ;\ \forall\beta\in[0,1-H[,\ \text{where}\ \ V_{T,h}(x):=\left|\frac{S_{T,h}(x)}{\widehat{f}_{T,h}(x)}\right|. \tag{8}$$

Moreover, for $\widehat{b}_{T,h,\varepsilon}$ defined by (3), $\forall\varepsilon>0$,

$$|\widehat{b}_{T,h,\varepsilon}(x)-\widehat{b}_{T,h}(x)|\leqslant C\frac{\varepsilon h^{-2}T^{2H-2}}{\widehat{f}_{T,h}(x)}. \tag{9}$$

Heuristically, Proposition 3.4 says that the pointwise quadratic risk of the kernel estimator $\widehat{b}_{T,h}(x)$ involves a squared bias of order h^2 and a variance term of order $1/(h^4T^{2(1-H)})$. The best possible rate is thus $T^{-\frac{2}{3}(1-H)}$ with a bandwidth choice of order $T^{-\frac{1}{3}(1-H)}$. A more rigorous formulation of this is stated below.
Note also that it follows from (9) that the rate of $\widehat{b}_{T,h,\varepsilon}(x)$ is preserved for any small ε.
We want to emphasize that no order condition is set on the kernel, and the bias term is not bounded in the usual way for kernel setting (see e.g. Tsybakov [25], Chapter 1). Indeed, we can not refer to the expectation of the numerator as a convolution product, because the existence of a stationary density is not ensured. Would it exist, it would be difficult to set adequate regularity conditions on it.

Now, consider a decreasing function $h:[t_0,\infty[\to]0,1[$ ($t_0\in\mathbb{R}_+$) such that

$$\lim_{T\to\infty}h(T)=0\text{ and }\lim_{T\to\infty}Th(T)=\infty$$

and assume that $\widehat{f}_{T,h(T)}(x)$ fulfills the following assumption.

Assumption 3.5. *There exists $l(x)\in]0,\infty]$ such that $\widehat{f}_{T,h(T)}(x)$ converges to $l(x)$ in probability as $T\to\infty$.*

Subsection 3.3 deals with the special case of fractional SDE with Gaussian solution in order to prove that Assumption 3.5 holds in this setting.
In Proposition 3.6, the result of Proposition 3.4 is extended to the estimator $\widehat{b}_{T,h(T)}(x)$ under Assumption 3.5.

Proposition 3.6. *Under Assumptions 2.10, 3.1 and 3.5:*

(1) *If there exists $\beta\in]0,1-H[$ such that $T^{-\beta}=_{T\to\infty}o(h(T)^2)$, then*

$$\widehat{b}_{T,h(T)}(x)\xrightarrow[T\to\infty]{\mathbb{P}}b(x).$$

(2) *For every $\gamma\in]0,\beta[$ such that*

$$h(T)=_{T\to\infty}o(T^{-\gamma})\ \text{and}\ T^{H-1+\gamma}=_{T\to\infty}o(h(T)^2),$$

then

$$T^{\gamma}|\widehat{b}_{T,h(T)}(x)-b(x)|\xrightarrow[T\to\infty]{\mathbb{P}}0.$$

Example. Consider

$$\beta\in\left]\frac{2}{3}(1-H),1-H\right[\text{ and }h(T):=T^{\frac{H-1}{3}}.$$

- $Th(T) = T^{H+\frac{2}{3}(1-H)} \xrightarrow[T\to\infty]{} \infty$.
- $T^{-\beta}/h(T)^2 = T^{-\beta+\frac{2}{3}(1-H)} \xrightarrow[T\to\infty]{} 0$.
- For every $\gamma \in]0,(1-H)/3[$, $h(T)/T^{-\gamma} = T^{\frac{H-1}{3}+\gamma} \xrightarrow[T\to\infty]{} 0$.
- For every $\gamma \in]0,(1-H)/3[$, $T^{H-1+\gamma}/h(T)^2 = T^{\frac{H-1}{3}+\gamma} \xrightarrow[T\to\infty]{} 0$.

In Corollary 3.7, the result of Proposition 3.6 is extended to $\widehat{b}_{T,h(T),\varepsilon(T)}(x)$ where

$$\lim_{T\to\infty} \varepsilon(T) = 0.$$

Corollary 3.7. *Under Assumptions 2.10, 3.1 and 3.5:*

(1) *If there exists $\beta \in]0,1-H[$ such that*

$$T^{-\beta} =_{T\to\infty} o(h(T)^2) \text{ and } \varepsilon(T) =_{T\to\infty} o(h(T)^{-2}T^{2H-2}),$$

then

$$\widehat{b}_{T,h(T),\varepsilon(T)}(x) \xrightarrow[T\to\infty]{\mathbb{P}} b(x).$$

(2) *For every $\gamma \in]0,\beta[$ such that*

$$\begin{cases} h(T) & =_{T\to\infty} o(T^{-\gamma}), \\ T^{H-1+\gamma} & =_{T\to\infty} o(h(T)^2) \\ \varepsilon(T) & =_{T\to\infty} o(h(T)^{-2}T^{2H-2+\gamma}) \end{cases},$$

then

$$T^{\gamma}|\widehat{b}_{T,h(T),\varepsilon(T)}(x) - b(x)| \xrightarrow[T\to\infty]{\mathbb{P}} 0.$$

Example. One can take $\varepsilon(T) := h(T)^2$.

3.3. **Special case of fractional SDE with Gaussian solution.** The purpose of this subsection is to show that Assumption 3.5 holds when the drift function in Equation (1) is linear with a negative slope. Note also that if $H = 1/2$, then $\widehat{f}_{T,h(T)}$ is a consistent estimator of the stationary density for Equation (1) (see Kutoyants [14], Section 4.2).

Assume that Equation (1) has a centered Gaussian stationary solution X and consider the normalized process $Y := X/\sigma_0$ where $\sigma_0 := \sqrt{\text{var}(X_0)}$.

Throughout this subsection, ν is the standard normal density and the autocorrelation function ρ of Y fulfills the following assumption.

Assumption 3.8. $\displaystyle\int_0^T \int_0^T |\rho(v-u)|dvdu =_{T\to\infty} O(T^{2H})$.

The following proposition ensures that under Assumption 3.8, $\widehat{f}_{T,h(T)}$ fulfills Assumption 3.5 for every $x \in \mathbb{R}^*$.

Proposition 3.9. *Under Assumptions 2.10 and 3.1, if Equation (1) has a centered, Gaussian, stationary solution X, the autocorrelation function ρ of $Y := X/\sigma_0$ satisfies Assumption 3.8 and $T^{2H-2} =_{T\to\infty} o(h(T))$, then*

$$\widehat{f}_{T,h(T)}(x) \xrightarrow[T\to\infty]{\mathbb{P}} \frac{1}{\sigma_0}\nu\left(\frac{x}{\sigma_0}\right) > 0 \tag{10}$$

for every $x \in \mathbb{R}^*$.

Now, consider the fractional Langevin equation

$$X(t) = X_0 - \lambda \int_0^t X(s)ds + \sigma B(t), \tag{11}$$

where $\lambda, \sigma > 0$. Equation (11) has a unique solution called Ornstein-Uhlenbeck's process.

On the one hand, the drift function of Equation (11) fulfills Assumption 2.10. So, under Assumption 3.1, by Proposition 3.4,

$$|\widehat{b}_{T,h}(x) + \lambda x| \leqslant \|b\|_{\text{Lip}} h + V_{T,h}(x),$$

where

$$T^\beta V_{T,h}(x) \xrightarrow[T\rightarrow\infty]{\mathbb{P}} 0 \ ; \ \forall \beta \in [0, 2H-1[.$$

On the other hand, by Cheridito et al. [3], Section 2, Equation (11) has a centered, Gaussian, stationary solution X such that:

$$X(t) = \sigma \int_{-\infty}^t e^{-\lambda(t-u)} dB(u) \ ; \ \forall t \in \mathbb{R}_+.$$

Moreover, by Cheridito et al. [3], Theorem 2.3, the autocorrelation function ρ of $Y := X/\sigma_0$ satisfies

$$\rho(T) =_{T\rightarrow\infty} O(T^{2H-2}).$$

So, ρ fulfills Assumption 3.8.

Consider $\beta \in]0, 2H-1[$ and $\gamma \in]0, 2H-1-\beta[$ such that

$$h(T) =_{T\rightarrow\infty} o(T^{-\gamma}) \text{ and } T^{H-1+\gamma} =_{T\rightarrow\infty} o(h(T)^2).$$

Then,

$$\lim_{T\rightarrow\infty} \frac{T^{2H-2}}{h(T)} = \lim_{T\rightarrow\infty} h(T)T^{H-1-\gamma}\frac{T^{H-1+\gamma}}{h(T)^2} = 0.$$

Therefore, by Proposition 3.6 and Proposition 3.9:

$$T^\gamma |\widehat{b}_{T,h(T)}(x) + \lambda x| \xrightarrow[T\rightarrow\infty]{\mathbb{P}} 0 \ ; \ \forall x \in \mathbb{R}^*.$$

4. Proofs

4.1. **Proof of Proposition 2.7.** On the existence, uniqueness and regularity of the paths of the solution of Equation (1), see Lejay [15].

Now, let us prove (4).

Let X_x be the solution of Equation (2.7) with initial condition $x \in \mathbb{R}$. Consider also $\varphi \in \text{Lip}_b^1(\mathbb{R})$ and $t > 0$. By Nualart [20], Proposition 5.2.3:

$$\begin{aligned}\int_0^t \varphi(X_x(u))\delta X_x(u) &= \int_0^t \varphi(X_x(u))b(X_x(u))du + \sigma \int_0^t \varphi(X_x(u))\delta B(u) \\ &= \int_0^t \varphi(X_x(u))dX_x(u) \\ &\quad -\alpha_H \sigma \int_0^t \int_0^t \varphi'(X_x(u))\mathbf{D}_v X_x(u)|u-v|^{2H-2}dvdu.\end{aligned}$$

Consider $u, v \in [0, t]$. On the one hand,

$$\mathbf{D}_v X_x(u) = \sigma \mathbf{1}_{[0,u]}(v) + \int_0^u b'(X_x(r))\mathbf{D}_v X_x(r)dr.$$

Then,

$$\mathbf{D}_v X_x(u) = \sigma \mathbf{1}_{[0,u]}(v) \exp\left(\int_v^u b'(X_x(r))dr\right).$$

On the other hand,

$$\partial_x X_x(u) = 1 + \int_0^u b'(X_x(r))\partial_x X_x(r)dr.$$

Then,

$$\partial_x X_x(u) = \exp\left(\int_0^u b'(X_x(r))dr\right).$$

Therefore,

$$\mathbf{D}_v X_x(u) = \sigma \mathbf{1}_{[0,u]}(v)\frac{\partial_x X_x(u)}{\partial_x X_x(v)}$$

and

$$\begin{aligned}\int_0^t \varphi(X_x(u))\delta X_x(u) &= \int_0^t \varphi(X_x(u))dX_x(u)\\ &\quad -\alpha_H\sigma^2\int_0^t\int_0^u \varphi'(X_x(u))\frac{\partial_x X_x(u)}{\partial_x X_x(v)}|u-v|^{2H-2}dvdu.\end{aligned}$$

4.2. **Proof of Corollary 2.8.** Consider $x \in \mathbb{R}$ and $\varepsilon, t > 0$. For every $s \in [0,t]$,

$$\partial_x X_x(s) = 1 + \int_0^s b'(X_x(r))\partial_x X_x(r)dr$$

and, by Taylor's formula,

$$X_{x+\varepsilon}(s) - X_x(s) = \varepsilon + \int_0^s (X_{x+\varepsilon}(r) - X_x(r))\int_0^1 b'(X_x(r) + \theta(X_{x+\varepsilon}(r) - X_x(r)))d\theta dr.$$

So, for every $(u,v) \in [0,t]^2$ such that $v < u$,

$$\frac{\partial_x X_x(u)}{\partial_x X_x(v)} = \exp\left(\int_v^u b'(X_x(r))dr\right)$$

and

$$\frac{X_{x+\varepsilon}(u) - X_x(u)}{X_{x+\varepsilon}(v) - X_x(v)} = \exp\left(\int_v^u\int_0^1 b'(X_x(r) + \theta(X_{x+\varepsilon}(r) - X_x(r)))d\theta dr\right).$$

For a given $\varphi \in \text{Lip}_b^1(\mathbb{R})$, by Proposition 2.7,

$$\Delta_\varphi^S(x,\varepsilon,t) \leqslant \alpha_H\sigma^2\int_0^t\int_0^u |\varphi'(X_x(u))|\Delta_\varphi(x,\varepsilon,u,v)(u-v)^{2H-2}dvdu,$$

where

$$\Delta_\varphi^S(x,\varepsilon,t) := \left|\int_0^t \varphi(X_x(u))\delta X_x(u) - S_\varphi(x,\varepsilon,t)\right|$$

and, for every $(u,v) \in [0,t]^2$ such that $v < u$,

$$\Delta_\varphi(x,\varepsilon,u,v) := \left|\frac{\partial_x X_x(u)}{\partial_x X_x(v)} - \frac{X_{x+\varepsilon}(u) - X_x(u)}{X_{x+\varepsilon}(v) - X_x(v)}\right|.$$

Since $b'(\mathbb{R}) \subset]-\infty,0]$ and b is two times continuously differentiable,

$$\begin{aligned}
\Delta_{\varphi}(x,\varepsilon,u,v) &= \left|\exp\left(\int_v^u b'(X_x(r))dr\right)\right.\\
&\quad \left. - \exp\left(\int_v^u\int_0^1 b'(X_x(r)+\theta(X_{x+\varepsilon}(r)-X_x(r)))d\theta dr\right)\right|\\
&\leqslant \sup_{z\in b'(\mathbb{R})} e^z\\
&\quad \times \int_v^u \left|b'(X_x(r)) - \int_0^1 b'(X_x(r)+\theta(X_{x+\varepsilon}(r)-X_x(r)))d\theta\right| dr\\
&\leqslant \int_v^u\int_0^1 |b'(X_x(r)) - b'(X_x(r)+\theta(X_{x+\varepsilon}(r)-X_x(r)))|d\theta dr\\
&\leqslant \frac{\|b''\|_\infty}{2}\int_v^u |X_{x+\varepsilon}(r)-X_x(r)|dr.
\end{aligned}$$

Consider $s\in\mathbb{R}_+$. By Equation (1):

$$\begin{aligned}
(X_{x+\varepsilon}(s)-X_x(s))^2 &= \varepsilon^2 + 2\int_0^s (X_{x+\varepsilon}(r)-X_x(r))d(X_{x+\varepsilon}-X_x)(r)\\
&= \varepsilon^2 + 2\int_0^s (X_{x+\varepsilon}(r)-X_x(r))(b(X_{x+\varepsilon}(r))-b(X_x(r)))dr.
\end{aligned}$$

By the mean-value theorem, there exists $x_s\in\mathbb{R}$ such that

$$\begin{aligned}
\partial_s(X_{x+\varepsilon}(s)-X_x(s))^2 &= 2(X_{x+\varepsilon}(s)-X_x(s))^2\frac{b(X_{x+\varepsilon}(s))-b(X_x(s))}{X_{x+\varepsilon}(s)-X_x(s)}\\
&= 2(X_{x+\varepsilon}(s)-X_x(s))^2 b'(x_s) \leqslant -2M(X_{x+\varepsilon}(s)-X_x(s))^2
\end{aligned}$$

and then,

$$|X_{x+\varepsilon}(s)-X_x(s)| \leqslant \varepsilon e^{-Ms}.$$

Therefore,

$$\begin{aligned}
\Delta_{\varphi}(x,\varepsilon,u,v) &\leqslant \frac{\|b''\|_\infty}{2}\varepsilon\int_v^u e^{-Mr}dr\\
&= \frac{\|b''\|_\infty}{2M}\varepsilon(e^{-Mv}-e^{-Mu}) \leqslant \frac{\|b''\|_\infty}{2M}\varepsilon e^{-Mv}.
\end{aligned}$$

Finally, using the above bounds, and in a second stage, the integration by parts formula, we get:

$$\begin{aligned}
\Delta_{\varphi}^S(x,\varepsilon,t) &\leqslant \alpha_H\sigma^2\frac{\|b''\|_\infty}{2M}\varepsilon\int_0^t\int_0^u |\varphi'(X_x(u))|e^{-Mv}(u-v)^{2H-2}dvdu\\
&\leqslant \alpha_H\sigma^2\frac{\|b''\|_\infty\|\varphi'\|_\infty}{2M}\varepsilon\int_0^t e^{-Mv}\int_v^t (u-v)^{2H-2}dudv\\
&= \alpha_H\sigma^2\frac{\|b''\|_\infty\|\varphi'\|_\infty}{2M(2H-1)}\varepsilon\int_0^t e^{-Mv}(t-v)^{2H-1}dv\\
&= \alpha_H\sigma^2\frac{\|b''\|_\infty\|\varphi'\|_\infty}{2M^2}\varepsilon\left(\frac{t^{2H-1}}{2H-1} - \int_0^t e^{-Mv}(t-v)^{2H-2}dv\right)\\
&\leqslant \alpha_H\sigma^2\frac{\|b''\|_\infty\|\varphi'\|_\infty}{2M^2(2H-1)}\varepsilon t^{2H-1} = C_\varphi\varepsilon t^{2H-1}.
\end{aligned}$$

4.3. **Proof of Proposition 2.11.** Consider $\gamma\in]1/2,H[$, $\delta\in]H-\gamma,1-\gamma[$ and $\Omega := \Omega_-\times\Omega_+$, where Ω_- (resp. Ω_+) is the completion of $C_0^\infty(\mathbb{R}_-,\mathbb{R})$ (resp. $C_0^\infty(\mathbb{R}_+,\mathbb{R})$) with respect to the norm $\|.\|_-$ (resp. $\|.\|_+$) defined by

$$\|\omega_-\|_- := \sup_{s<t\leqslant 0}\frac{|\omega_-(t)-\omega_-(s)|}{|t-s|^\gamma(1+|s|+|t|)^\delta}\ ;\ \forall\omega_-\in\Omega_-$$

(resp.

$$\|\omega_+\|_+ := \sup_{0\leqslant s<t} \frac{|\omega_+(t)-\omega_+(s)|}{|t-s|^{\gamma}(1+|s|+|t|)^{\delta}} \ ; \ \forall \omega_+ \in \Omega_+).$$

By Hairer [7], Section 3 or more clearly by Hairer and Ohashi [8], Lemmas 4.1 and 4.2, there exist a Borel probability measure $\mathbb{P}$ on Ω and a transition kernel P from Ω_- to Ω_+ such that:

- The process generated by $(\Omega,\mathbb{P})$ is a two-sided fractional Brownian motion $\widetilde{B}$.
- For every Borel set U (resp. V) of Ω_- (resp. Ω_+),

$$\mathbb{P}(U\times V) = \int_U P(\omega_-,V)\mathbb{P}_-(d\omega_-)$$

  where $\mathbb{P}_-$ is the probability distribution of $(\widetilde{B}(t))_{t\in\mathbb{R}_-}$.

Let $I : \mathbb{R}\times\Omega_+ \rightarrow C^0(\mathbb{R}_+,\mathbb{R})$ be the Itô (solution) map for Equation (1). In general, $I(x,.)$ with $x\in\mathbb{R}$ is not a Markov process. However, the solution of Equation (1) can be coupled with the past of the driving signal in order to bypass this difficulty. In other words, consider the enhanced Itô map $\mathfrak{I} : \mathbb{R}\times\Omega \rightarrow C^0(\mathbb{R}_+,\mathbb{R}\times\Omega_-)$ such that for every $(x,\omega_-,\omega_+)\in\mathbb{R}\times\Omega$ and $t\in\mathbb{R}_+$,

$$\mathfrak{I}(x,\omega_-,\omega_+)(t) := (I(x,\omega_+)(t), p_{\Omega_-}(\theta(\omega_-,\omega_+)(t)))$$

where p_{Ω_-} is the projection from Ω onto Ω_-,

$$\theta(\omega_-,\omega_+)(t) := (\omega_-\sqcup\omega_+)(t+\cdot) - (\omega_-\sqcup\omega_+)(\cdot)$$

and $\omega_-\sqcup\omega_+$ is the concatenation of ω_- and ω_+. By Hairer [7], Lemma 2.12, the process $\mathfrak{I}(x,.)$ is Markovian and has a Feller transition semigroup $(Q(t))_{t\in\mathbb{R}_+}$ such that for every $t\in\mathbb{R}_+$, $(x,\omega_-)\in\mathbb{R}\times\Omega_-$ and every Borel set U (resp. V) of $\mathbb{R}$ (resp. Ω_-),

$$Q(t;(x,\omega_-),U\times V) = \int_V \delta_{I(x,\omega_+)(t)}(U)P(t;\omega_-,d\omega_+)$$

where δ_y is the delta measure located at $y\in\mathbb{R}$ and $P(t;\omega_-,.)$ is the pushforward measure of $P(\omega_-,.)$ by $\theta(\omega_-,.)(t)$.

In order to prove Proposition 2.11, let us first state the following result from Hairer [7] and Hairer and Ohashi [8].

Theorem 4.1. *Under Assumption 2.10:*

(1) *(Irreducibility) There exists* $\tau\in]0,\infty[$ *such that for every* $(x,\omega_-)\in\mathbb{R}\times\Omega_-$ *and every nonempty open set* $U\subset\mathbb{R}$,

$$Q(\tau;(x,\omega_-),U\times\Omega_-) > 0.$$

(2) *There exists a unique probability measure* μ *on* $\mathbb{R}\times\Omega_-$ *such that* $\mu(p_{\Omega_-}\in\cdot) = \mathbb{P}_-$ *and*

$$Q(t)\mu = \mu \ ; \ \forall t\in\mathbb{R}_+.$$

For a proof of Theorem 4.1.(1), see Hairer and Ohashi [8], Proposition 5.8. For a proof of Theorem 4.1.(2), see Hairer [7], Theorem 6.1 which is a consequence of Proposition 2.18, Lemma 2.20 and Proposition 3.12.

Since the Feller transition semigroup Q has exactly one invariant measure μ by

Theorem 4.1, μ is ergodic, and since the first component of the process generated by Q is a solution of Equation (1), by the ergodic theorem for Markov processes:

$$\frac{1}{T}\int_0^T \varphi(X(t))dt = \frac{1}{T}\int_0^T (\varphi\circ p_{\mathbb{R}})(\mathfrak{I}(X_0,.)(t))dt \xrightarrow[T\to\infty]{\text{a.s.}/\mathrm{L}^2} \mu(\varphi\circ p_{\mathbb{R}}).$$

Moreover, $\mu = Q(\tau)\mu$. So,

$$\begin{aligned}\mu(\varphi\circ p_{\mathbb{R}}) &= \int_{\mathbb{R}\times\Omega_-}(\varphi\circ p_{\mathbb{R}})(x,\omega_-)(Q(\tau)\mu)(dx,d\omega_-)\\ &= \int_{\mathbb{R}\times\Omega_-}\varphi(x)\int_{\mathbb{R}\times\Omega_-}Q(\tau;(\bar{x},\bar{\omega}_-),(dx,d\omega_-))\mu(d\bar{x},d\bar{\omega}_-)\\ &\geqslant \min_{x\in C}\varphi(x)\cdot\int_{C\times\Omega_-}\int_{C\times\Omega_-}Q(\tau;(\bar{x},\bar{\omega}_-),(dx,d\omega_-))\mu(d\bar{x},d\bar{\omega}_-)\\ &\geqslant \min_{x\in C}\varphi(x)\cdot\int_{C\times\Omega_-}Q(\tau;(\bar{x},\bar{\omega}_-),\mathrm{int}(C)\times\Omega_-)\mu(d\bar{x},d\bar{\omega}_-).\end{aligned}$$

Since

$$Q(\tau;(\bar{x},\bar{\omega}_-),\mathrm{int}(C)\times\Omega_-) > 0\ ;\ \forall(\bar{x},\bar{\omega}_-)\in\mathbb{R}\times\Omega_-$$

by Theorem 4.1.(1), then

$$\int_{C\times\Omega_-}Q(\tau;(\bar{x},\bar{\omega}_-),\mathrm{int}(C)\times\Omega_-)\mu(d\bar{x},d\bar{\omega}_-) > 0.$$

Therefore, $\mu(\varphi\circ p_{\mathbb{R}}) > 0$.

4.4. **Proof of Proposition 3.4.** First write that, under Assumption 2.10, for any $s\in[0,T]$ such that $X(s)\in[x-h,x+h]$,

$$|b(X(s))-b(x)| \leqslant \|b\|_{\mathrm{Lip}}h.$$

So,

$$\left|\frac{B_{T,h}(x)}{\widehat{f}_{T,h}(x)}\right| \leqslant \|b\|_{\mathrm{Lip}}h. \tag{12}$$

Next, the following Lemma provides a suitable control of $\mathbb{E}(|S_{T,h}(x)|^2)$.

Lemma 4.2. *Under Assumptions 2.10 and 3.1, there exists a deterministic constant $C>0$, not depending on h and T, such that:*

$$\mathbb{E}(|S_{T,h}(x)|^2) \leqslant CT^{2(H-1)}h^{-4}.$$

Proof. Since K belongs to $C_b^1(\mathbb{R},\mathbb{R}_+)$, the map

$$\varphi_h : y\in\mathbb{R}\longmapsto\varphi_h(y) := K\left(\frac{y-x}{h}\right)$$

belongs to $\mathrm{Lip}_b^1(\mathbb{R})$. Moreover, since K and K' are continuous with bounded support $[-1,1]$,

$$\left(\int_0^T\mathbb{E}(|\varphi_h(X(s))|^{1/H})ds\right)^{2H} \leqslant \|K\|_\infty^2T^{2H}$$

and

$$\left(\int_0^T\mathbb{E}(|\varphi_h'(X(s))|^2)^{1/(2H)}ds\right)^{2H} \leqslant \|K'\|_\infty^2T^{2H}h^{-2}.$$

Therefore, by Theorem 2.9, there exists a deterministic constant $C > 0$, not depending on h and T, such that:

$$\begin{aligned}\mathbb{E}(|S_{T,h}(x)|^2) &= \frac{1}{T^2h^2}\mathbb{E}\left(\left|\int_0^T \varphi_h(X(s))\delta B(s)\right|^2\right)\\ &\leqslant CT^{2(H-1)}h^{-4}.\end{aligned}$$

□

First, by Inequality (12) and Equation (7),

$$|\widehat{b}_{T,h}(x) - b(x)| \leqslant \|b\|_{\rm Lip}h + V_{T,h}(x)$$

where $V_{T,h}(x)$ is defined by (8). Consider $\beta \in [0, 1-H[$. By Lemma 4.2:

$$T^{2\beta}\mathbb{E}(|S_{T,h}(x)|^2) \leqslant CT^{2(H-1+\beta)}h^{-4} \xrightarrow[T\to\infty]{} 0.$$

So,

$$T^{\beta}|S_{T,h}(x)| \xrightarrow[T\to\infty]{\mathbb{P}} 0.$$

Moreover, by Lemma 3.2:

$$\frac{1}{\widehat{f}_{T,h}(x)} \xrightarrow[T\to\infty]{\mathcal{D}} \frac{1}{l_h(x)} > 0.$$

Therefore, by Slutsky's lemma:

$$T^{\beta}V_{T,h}(x) \xrightarrow[T\to\infty]{\mathbb{P}} 0.$$

Lastly, the bound (9) follows from the following Lemma.

Lemma 4.3. *Under Assumptions 2.10 and 3.1, there exists a deterministic constant $C > 0$, not depending on ε, h and T, such that:*

$$|\widehat{b}_{T,h,\varepsilon}(x) - \widehat{b}_{T,h}(x)| \leqslant C\frac{\varepsilon h^{-2}T^{2H-2}}{\widehat{f}_{T,h}(x)}.$$

Proof. Since K belongs to $C_b^1(\mathbb{R},\mathbb{R}_+)$, the map

$$\varphi_h : y \in \mathbb{R} \longmapsto \varphi_h(y) := K\left(\frac{y-x}{h}\right)$$

belongs to $\text{Lip}_b^1(\mathbb{R})$. Consider

$$\begin{aligned}S_h(x_0,\varepsilon,T) :=& \int_0^T \varphi_h(X_{x_0}(u))dX_{x_0}(u)\\ &-\alpha_H\sigma^2\int_0^T\int_0^u \varphi_h'(X_{x_0}(u))\frac{X_{x_0+\varepsilon}(u)-X_{x_0}(u)}{X_{x_0+\varepsilon}(v)-X_{x_0}(v)}|u-v|^{2H-2}dvdu.\end{aligned}$$

By Corollary 2.8:

$$\begin{aligned}\left|\int_0^T \varphi_h(X_{x_0}(u))\delta X_{x_0}(u) - S_h(x_0,\varepsilon,T)\right| &\leqslant H\sigma^2\frac{\|b''\|_\infty\|\varphi_h'\|_\infty}{M^2}\varepsilon T^{2H-1}\\ &\leqslant C\varepsilon h^{-1}T^{2H-1},\end{aligned}$$

where

$$C := \frac{H\sigma^2\|b''\|_\infty\|K'\|_\infty}{M^2}.$$

Therefore,

$$|\widehat{b}_{T,h,\varepsilon}(x) - \widehat{b}_{T,h}(x)| \leqslant C\frac{\varepsilon h^{-2}T^{2H-2}}{\widehat{f}_{T,h}(x)}.$$

□

4.5. **Proof of Proposition 3.6.** On the one hand, assume that there exists $\beta \in]0, 1-H[$ such that

$$T^{-\beta} =_{T\rightarrow\infty} o(h(T)^2)$$

in order to show the consistency of the estimator $\widehat{b}_{T,h(T)}(x)$. First, let us prove that

$$\frac{S_{T,h(T)}(x)}{\widehat{f}_{T,h(T)}(x)} \xrightarrow[T\rightarrow\infty]{\mathbb{P}} 0. \tag{13}$$

For $\varepsilon > 0$ arbitrarily chosen:

$$\mathbb{P}\left(\left|\frac{S_{T,h(T)}(x)}{\widehat{f}_{T,h(T)}(x)}\right| \geqslant \varepsilon\right) \leqslant \mathbb{P}(|S_{T,h(T)}(x)| \geqslant \varepsilon T^{H+\beta-1}) + \mathbb{P}(\widehat{f}_{T,h(T)}(x) < T^{H+\beta-1}).$$

By Lemma 4.2:

$$\mathbb{P}(|S_{T,h(T)}(x)| \geqslant \varepsilon T^{H+\beta-1}) \leqslant C\varepsilon^{-2}|h(T)^{-2}T^{-\beta}|^2 \xrightarrow[T\rightarrow\infty]{} 0.$$

So, since

$$\widehat{f}_{T,h(T)}(x) \xrightarrow[T\rightarrow\infty]{\mathbb{P}} l(x) \in]0,\infty],$$

the convergence result (13) is true.

Moreover, by Inequality (12):

$$\frac{B_{T,h(T)}(x)}{\widehat{f}_{T,h(T)}(x)} \xrightarrow[T\rightarrow\infty]{\text{a.s.}} 0. \tag{14}$$

Therefore, by the convergence results (13) and (14) together with Equation (7):

$$\widehat{b}_{T,h(T)}(x) \xrightarrow[T\rightarrow\infty]{\mathbb{P}} b(x).$$

On the other hand, let $\gamma \in]0, \beta[$ be arbitrarily chosen such that

$$h(T) =_{T\rightarrow\infty} o(T^{-\gamma}) \text{ and } T^{H-1+\gamma} =_{T\rightarrow\infty} o(h(T)^2)$$

in order to show that

$$T^{\gamma}|\widehat{b}_{T,h(T)}(x) - b(x)| \xrightarrow[T\rightarrow\infty]{\mathcal{D}} 0. \tag{15}$$

First, by Inequality (12) and Equation (7):

$$T^{\gamma}|\widehat{b}_{T,h(T)}(x) - b(x)| \leqslant \|b\|_{\text{Lip}} T^{\gamma} h(T) + T^{\gamma} V_{T,h(T)}(x). \tag{16}$$

By Lemma 4.2:

$$T^{2\gamma}\mathbb{E}(|S_{T,h(T)}(x)|^2) \leqslant C|h(T)^{-2}T^{H-1+\gamma}|^2 \xrightarrow[T\rightarrow\infty]{} 0.$$

So, since

$$\widehat{f}_{T,h(T)}(x) \xrightarrow[T\rightarrow\infty]{\mathbb{P}} l(x) \in]0,\infty],$$

by Slutsky's lemma:

$$T^{\gamma} V_{T,h(T)}(x) \xrightarrow[T\rightarrow\infty]{\mathcal{D}} 0.$$

Finally, since $h(T) =_{T\rightarrow\infty} o(T^{-\gamma})$, by Equation (16), the convergence result (15) is true.

4.6. **Proof of Corollary 3.7.** In order to establish a rate of convergence for $\widehat{b}_{T,h,\varepsilon}(x)$, Lemma 4.2 and Lemma 4.3 provide a suitable control.
Indeed, by Lemma 4.3, there exists a deterministic constant $C > 0$ such that:

$$\begin{aligned}|\widehat{b}_{T,h(T),\varepsilon(T)}(x) - b(x)| &\leqslant |\widehat{b}_{T,h(T),\varepsilon(T)}(x) - \widehat{b}_{T,h(T)}(x)| + |\widehat{b}_{T,h(T)}(x) - b(x)| \\ &\leqslant C\frac{\varepsilon(T)h(T)^{-2}T^{2H-2}}{\widehat{f}_{T,h(T)}(x)} + |\widehat{b}_{T,h(T)}(x) - b(x)|.\end{aligned}$$

Proposition 3.6 allows to conclude.

4.7. **Proof of Proposition 3.9.** Consider a random variable $U \rightsquigarrow \mathcal{N}(0,1)$ and

$$\mathcal{G} := \{G : \mathbb{R} \rightarrow \mathbb{R} : \mathbb{E}(G(U)) = 0 \text{ and } \mathbb{E}(G(U)^2) < \infty\},$$

which is a subset of $L^2(\mathbb{R}, \nu(y)dy)$.

The Hermite polynomials

$$H_q(y) := (-1)^q e^{y^2/2}\frac{d^q}{dy^q}e^{-y^2/2} \ ; \ y \in \mathbb{R},\ q \in \mathbb{N}$$

form a complet orthogonal system of functions of $L^2(\mathbb{R}, \nu(y)dy)$ such that

$$\mathbb{E}(H_q(U)H_p(U)) = q!\delta_{p,q} \ ; \ \forall p, q \in \mathbb{N}.$$

By Taqqu [24] (see p. 291) and Puig et al. [21], Lemma 3.3:

(1) For any $G \in \mathcal{G}$ and $y \in \mathbb{R}$,

$$G(y) = \sum_{q=m(G)}^{\infty}\frac{J(q)}{q!}H_q(y) \tag{17}$$

in $L^2(\mathbb{R}, \nu(y)dy)$, where

$$J(q) := \mathbb{E}(G(U)H_q(U)) \ ; \ \forall q \in \mathbb{N}$$

and

$$m(G) := \inf\{q \in \mathbb{N} : J(q) \neq 0\}.$$

(2) (Mehler's formula) For any centered, normalized and stationary Gaussian process Z of autocorrelation function R:

$$\mathbb{E}(H_q(Z(u))H_p(Z(v))) = q!R(v-u)^q\delta_{p,q} \ ; \ \forall u, v \in \mathbb{R}_+,\ \forall p, q \in \mathbb{N}. \tag{18}$$

Consider the map $K_T : \mathbb{R} \rightarrow \mathbb{R}$ defined by:

$$K_T(y) := \frac{1}{h(T)}K\left(\frac{y}{h(T)}\right) \ ; \ \forall y \in \mathbb{R}.$$

In order to use (17) and (18) to prove the convergence result (10), note that $\widehat{f}_{T,h(T)}(x)$ can be rewritten as

$$\widehat{f}_{T,h(T)}(x) = \frac{1}{T}\int_0^T G_{T,x}(Y(s))ds - R_{T,x},$$

where

$$R_{T,x} := \frac{1}{\sigma_0}\left(K_T * \nu\left(\frac{\cdot}{\sigma_0}\right)\right)(x) \ ; \ \forall y \in \mathbb{R}$$

and

$$G_{T,x}(y) := K_T(\sigma_0 y - x) - R_{T,x}.$$

Lemma 4.4. *The map $G_{T,x}$ belongs to $\mathcal{G}$ and there exists $T_x > 0$ such that*

$$m(G_{T,x}) = 1 \ ; \ \forall T > T_x.$$

Proof. On the one hand, since K_T is continuous and its support is compact, $G_{T,x} \in L^2(\mathbb{R}, \nu(y)dy)$. Moreover,

$$\begin{aligned}\mathbb{E}(G_{T,x}(U)) &= \int_{-\infty}^{\infty} G_{T,x}(y)\nu(y)dy \\ &= \int_{-\infty}^{\infty} K_T(\sigma_0 y - x)\nu(y)dy - R_{T,x} = 0.\end{aligned}$$

So, $G_{T,x} \in \mathcal{G}$.

On the other hand, for every $q \in \mathbb{N}$, by putting $J_{T,x}(q) := \mathbb{E}(G_{T,x}(U)H_q(U))$,

$$\begin{aligned}J_{T,x}(1) &= \int_{-\infty}^{\infty} G_{T,x}(y)H_1(y)\nu(y)dy \\ &= \int_{(x-h(T))/\sigma_0}^{(x+h(T))/\sigma_0} K_T(\sigma_0 y - x)\nu(y)ydy - R_{T,x}\int_{-\infty}^{\infty} H_0(y)H_1(y)\nu(y)dy \\ &= \int_{(x-h(T))/\sigma_0}^{(x+h(T))/\sigma_0} K_T(\sigma_0 y - x)\nu(y)ydy.\end{aligned}$$

For any $x > 0$, there exists $T_x^+ > 0$ such that for every $T > T_x^+$,

$$I_{T,x} := \left[\frac{x-h(T)}{\sigma_0}; \frac{x+h(T)}{\sigma_0}\right] \subset]0, \infty[.$$

For every $T > T_x^+$, since $y \mapsto K_T(\sigma_0 y - x)$, ν and $\mathrm{Id}_{\mathbb{R}}$ are continuous and strictly positive on $I_{T,x}^{\circ}$, $J_{T,x}(1) > 0$. Symmetrically, for every $x < 0$, there exists $T_x^- > 0$ such that for every $T > T_x^-$, $J_{T,x}(1) < 0$. This concludes the proof. □

Lemma 4.5. *For every* $x \in \mathbb{R}^*$,

$$\sum_{q=1}^{\infty} \frac{J_{T,x}(q)^2}{q!} =_{T\to\infty} O\left(\frac{1}{h(T)}\right).$$

Proof. Since $G_{T,x} \in L^2(\mathbb{R}, \nu(y)dy)$, by Parseval's inequality:

$$\begin{aligned}\sum_{q=1}^{\infty} \frac{J_{T,x}(q)^2}{q!} &= \mathbb{E}(G_{T,x}(U)^2) \\ &= \int_{-\infty}^{\infty} (K_T(\sigma_0 y - x) - R_{T,x})^2\nu(y)dy \\ &\leqslant 2\int_{-\infty}^{\infty} K_T(\sigma_0 y - x)^2\nu(y)dy + 2R_{T,x}^2.\end{aligned}$$

On the one hand,

$$R_{T,x} \xrightarrow[T\to\infty]{} \frac{1}{\sigma_0}\nu\left(\frac{x}{\sigma_0}\right).$$

So,

$$R_{T,x}^2 =_{T\to\infty} O(1).$$

On the other hand,

$$\begin{aligned}\int_{-\infty}^{\infty} K_T(\sigma_0 y - x)^2\nu(y)dy &= \frac{1}{\sigma_0 h(T)}\int_{-1}^{1} K(y)^2\nu\left(\frac{h(T)y + x}{\sigma_0}\right)dy \\ &\leqslant \frac{2\|K\|_\infty^2\|\nu\|_\infty}{\sigma_0 h(T)}.\end{aligned}$$

Therefore,

$$\sum_{q=2}^{\infty}\frac{J_{T,x}(q)^2}{q!} =_{T\to\infty} O\left(\frac{1}{h(T)}\right).$$

□

In order to prove the convergence result (10), since

$$R_{T,x} \xrightarrow[T\to\infty]{} \frac{1}{\sigma_0}\nu\left(\frac{x}{\sigma_0}\right),$$

let us prove that

$$\left|\frac{1}{T}\int_0^T G_{T,x}(Y(s))ds\right| \xrightarrow[T\to\infty]{\mathbb{L}^2} 0. \tag{19}$$

By the decomposition (17) and Mehler's formula (18) applied to $G_{T,x}$ and Y, for every $u,v\in[0,T]$,

$$\mathbb{E}(G_{T,x}(Y(u))G_{T,x}(Y(v))) = \sum_{q=1}^{\infty}\frac{J_{T,x}(q)^2}{q!}\rho(v-u)^q.$$

So, since ρ is a $[-1,1]$-valued function,

$$\begin{aligned}\mathbb{E}\left(\left|\int_0^T G_{T,x}(Y(s))ds\right|^2\right) &= \int_0^T\int_0^T|\mathbb{E}(G_{T,x}(Y(u))G_{T,x}(Y(v)))|dudv\\ &\leqslant \sum_{q=1}^{\infty}\frac{J_{T,x}(q)^2}{q!}\int_0^T\int_0^T|\rho(v-u)|^q dudv\\ &\leqslant \left(\int_0^T\int_0^T|\rho(v-u)|dudv\right)\sum_{q=1}^{\infty}\frac{J_{T,x}(q)^2}{q!}.\end{aligned}$$

Then, by Assumption 3.8 and Lemma 4.5:

$$\lim_{T\to\infty}\mathbb{E}\left(\left|\frac{1}{T}\int_0^T G_{T,x}(Y(s))ds\right|^2\right) = \lim_{T\to\infty}\frac{T^{2H-2}}{h(T)} = 0.$$

Therefore, the convergence result (19) is true.

References

[1] S. Bajja, K. Es-Sebaiy and L. Viitasaari. *Least Square Estimator of Fractional Ornstein-Uhlenbeck Processes with Periodic Mean.* Journal of the Korean Statistical Society 36, 4, 608-622, 2017.

[2] D. Bosq. *Nonparametric Statistics for Stochastic Processes: Estimation and Prediction.* Springer, 1996.

[3] P. Cheridito, H. Kawaguchi and M. Maejima. *Fractional Ornstein-Uhlenbeck Processes.* Electronic Journal of Porbability 8, 3, 1-14, 2003.

[4] A. Chronopoulou and S. Tindel. *On Inference for Fractional Differential Equations.* Stat. Inference Stoch. Process. 16, 1, 29-61, 2013.

[5] P. Friz and M. Hairer. *A Course on Rough Paths.* Springer, 2014.

[6] P. Friz and N. Victoir. *Multidimensional Stochastic Processes as Rough Paths: Theory and Applications.* Cambridge Studies in Applied Mathematics 120, Cambridge University Press, 2010.

[7] M. Hairer. *Ergodicity of Stochastic Differential Equations Driven by Fractional Brownian Motion.* The Annals of Probability 33, 3, 703-758, 2005.

[8] M. Hairer and A. Ohashi. *Ergodic Theory for SDEs with Extrinsic Memory.* The Annals of Probability 35, 5, 1950-1977, 2007.

[9] M. Hairer and N.S. Pillai. *Ergodicity of Hypoelliptic SDEs Driven by Fractional Brownian Motion.* Annales de l'IHP 47, 4, 2544-2598, 2013.

[10] M.L. Kleptsyna and A. Le Breton. *Some Explicit Statistical Results about Elementary Fractional Type Models.* Nonlinear Analysis 47, 4783-4794, 2001.
[11] K. Kubilius and V. Skorniakov. *On Some Estimators of the Hurst Index of the Solution of SDE Driven by a Fractional Brownian Motion.* Statistics and Probability Letters 109, 159-167, 2016.
[12] Y. Hu and D. Nualart. *Parameter Estimation for Fractional Ornstein-Uhlenbeck Processes.* Statistics and Probability Letters 80, 1030-1038, 2010.
[13] Y. Hu, D. Nualart and H. Zhou. *Drift Parameter Estimation for Nonlinear Stochastic Differential Equations Driven by Fractional Brownian Motion.* arXiv:1803.01032v1.
[14] Y. Kutoyants. *Statistical Inference for Ergodic Diffusion Processes.* Springer, 2004.
[15] A. Lejay. *Controlled Differential Equations as Young Integrals: A Simple Approach.* Journal of Differential Equations 249, 1777-1798, 2010.
[16] G. Lindgren. *Lectures on Stationary Stochastic Processes.* PhD course of Lund's University, 2006.
[17] A. Neuenkirch and S. Tindel. *A Least Square-Type Procedure for Parameter Estimation in Stochastic Differential Equations with Additive Fractional Noise.* Stat. Inference Stoch. Process 17, 1, 99-120, 2014.
[18] M.N. Mishra and B.L.S. Prakasa Rao. *Nonparameteric Estimation of Trend for Stochastic Differential Equations Driven by Fractional Brownian Motion.* Stat. Inference. Stoch. Process. 14, 2, 101-109, 2011.
[19] Y. Mishura and K. Ralchenko. *On Drift Parameter Estimation in Models with Fractional Brownian Motion by Discrete Observations.* Austrian Journal of Statistics 43, 3-4, 217-228, 2014.
[20] D. Nualart. *The Malliavin Calculus and Related Topics.* Springer, 2006.
[21] B. Puig, F. Poirion and C. Soize. *Non-Gaussian Simulation Using Hermite Polynomial Expansion: Convergences and Algorithms.* Probabilistic Engineering Mechanics 17, 253-264, 2002.
[22] D. Revuz and M. Yor. *Continuous Martingales and Brownian Motion. Third Edition.* A Series of Comprehensive Studies in Mathematics 293, Springer, 1999.
[23] B. Saussereau. *Nonparametric Inference for Fractional Diffusion.* Bernoulli 20, 2, 878-918, 2014.
[24] M.S. Taqqu. *Weak Convergence to Fractional Brownian Motion and to the Rosenblatt Process.* Z. Warscheinlichkeitstheorie verw. Gebiete 31, 287-302, 1975.
[25] A. B. Tsybakov. *Introduction to nonparametric estimation.* Revised and extended from the 2004 French original. Translated by Vladimir Zaiats. Springer Series in Statistics. Springer, New York, 2009.
[26] C.A. Tudor and F. Viens. *Statistical Aspects of the Fractional Stochastic Calculus.* The Annals of Statistics 35, 3, 1183-1212, 2007.
[27] C.A. Tudor and F. Viens. *Variations and Estimators for Self-Similarity Parameters via Malliavin Calculus.* The Annals of Probability 37, 6, 2093-2134, 2009.

*Laboratoire MAP5, Université Paris Descartes, Paris, France
E-mail address: fabienne.comte@parisdescartes.fr

†Laboratoire Modal'X, Université Paris Nanterre, Nanterre, France
E-mail address: nmarie@parisnanterre.fr

†ESME Sudria, Paris, France
E-mail address: nicolas.marie@esme.fr