APPROXIMATING SOME VOLTERRA TYPE STOCHASTIC INTEGRALS WITH APPLICATIONS TO PARAMETER ESTIMATION

HENRIK HULT

ABSTRACT. We use a general representation of continuous Gaussian processes as the limit of a sequence of processes in the associated reproducing kernel Hilbert space, to Gaussian processes represented as Volterra type stochastic integrals with respect to Brownian motion, including the fractional Brownian motion. As special cases of this representation we obtain for example, the Karhunen-Loéve decomposition for standard Brownian motion and a wavelet representation for fractional Brownian motion. We also show how the representation can be used to estimate parameters. In particular we derive an estimator for the mean-reverting parameter in an Ornstein-Uhlenbeck process driven by a fractional Brownian motion.

1. INTRODUCTION

Gaussian processes admitting representation as a Volterra type stochastic integral with respect to the standard Brownian motion such as the fractional Brownian motion are used in several fields including telecommunications, subsurface hydrology and mathematical finance. An important problem in such areas is to estimate the parameters within a given family of models, for example in a stochastic differential equation driven by the fractional Brownian motion. In this paper we will consider the fractional Brownian motion and related processes. Since the fractional Brownian motion is quite complicated, one generally tries to find an approximation which is easier to handle. Different approximations have been studied in the literature. In Comte (1996) [3], and Comte and Renault (1996) [4], Gaussian processes of the form,

$$X_t = \int_0^t a(t-s)dB_s, \tag{1.1}$$

where $\{B_t\}_{t\geq 0}$ is standard Brownian motion are studied. In particular the authors show how to estimate parameters in some special cases when $a(\cdot)$ depend on an unknown parameter.

A natural tool when studying approximations is to use wavelets. We can represent the process in a wavelet basis, $\{\psi_{j,k}\}$, as a Mercer-type representation,

$$X_t = \sum_{j,k} \psi_{j,k}(t)\xi_{j,k}, \tag{1.2}$$

where the wavelet coefficients, $\{\xi_{j,k}\}$, is a sequence of Gaussian variables, see Abry, Flandrin, Taqqu and Veitch (2000) [1], and the references therein. The correlations between the wavelet coefficients depend on the number of vanishing moments of the wavelet used. Ideal would be to have independent coefficients. In Meyer, Sellan and Taqqu (1999) [15], the authors obtain a wavelet decomposition for the fractional

1991 *Mathematics Subject Classification.* Primary: 60H07, 28C20; Secondary: 60H10, 60H20.
Key words and phrases. fractional Brownian motion, reproducing kernel Hilbert space, Gaussian process, likelihood function.

Brownian motion using a generalization of the midpoint displacement technique. They show that it has the representation

$$B^H(t) = \sum_k S_k \Phi_k^H(t) + \sum_{j,k} \xi_{j,k} \Psi_{j,k}^H(t), \tag{1.3}$$

where convergence holds almost surely uniformly on compact intervals, $\{S_k\}$ is a fractional ARIMA process and $\{\xi_{j,k}\}$ are is a sequence of i.i.d. $N(0,1)$ random variables and independent of $\{S_k\}$. The sequence $\left\{\Phi_k^H, \Psi_{j,k}^H\right\}$ is constructed from a suitable wavelet basis, for instance the Meyer wavelets with vanishing moments of all orders. As we shall see below, in Example 3.3, the representation (1.3) can be derived as a special case of a more general representation. Let us also mention the work by Carmona, Coutin and Montseny (2000) [2] and Norros, Mannersalo and Wang (1999) [17], on further approximations and problems concerning simulation of the fractional Brownian motion.

In this paper we study centered Gaussian processes $X = \{X_t\}_{t\in T}$, where T is an index set, admitting a representation of the form:

$$X_t = \int_T V(t,r)dB_r, \quad t \in T, \tag{1.4}$$

where $\{B_t\}_{t\in T}$ denotes standard Brownian motion and V is a kernel satisfying some conditions, see Section 2. The conditions are such that X has a continuous version and includes the fractional Brownian motion. In the case where the kernel V depends on some unknown parameter θ we suggest a general procedure for estimating θ. In particular, we show explicitly how to estimate the drift parameter, θ, in the process

$$Y(t) = \theta a(t) + B^H(t),$$

where $a(\cdot)$ is a deterministic function and $\left\{B_t^H, t \in T\right\}$ is the fractional Brownian motion. We prove consistency and normality of the estimator. These results extends previous work by Norros, Valkeila and Virtamo (1999) [18]. We also derive an estimator for the mean reverting parameter in a fractional Ornstein-Uhlenbeck process. The estimators are based on a finite number of 'spectral components' but to compute these components accurately we need continuous observations of the process. The main tool used to derive these results is the following representation of Gaussian processes:

$$X_t = \sum_j \Psi_j(t)\xi_j, \quad t \in T, \tag{1.5}$$

where $\{\xi_j\}_1^\infty$ is a sequence of i.i.d. $N(0,1)$ random variables and $\{\Psi_j\}_1^\infty$ is an orthonormal basis in the reproducing kernel Hilbert space associated with X. If T is compact and X is a.s. continuous the sum converges a.s. uniformly. To get an explicit representation (1.5) we need to find an orthonormal basis in the reproducing kernel Hilbert space. For processes of the form (1.4) this is not difficult since, as we will see, the reproducing kernel Hilbert space is the image of $L^2(T)$ under the integral transform V:

$$(Vf)(t) = \int_T V(t,s)f(s)ds, \quad f \in L^2(T),$$

equipped with the inner product,

$$\langle Vf, Vg\rangle_{V(L^2(T))} = \langle f, g\rangle_{L^2(T)}.$$

An orthonormal basis for the reproducing kernel Hilbert space is then obtained from an orthonormal basis in $L^2(T)$ by applying the integral transform V on each basis function. That is, if $\{\psi_j\}_1^\infty$ is an orthonormal basis in $L^2(T)$ then $\{\Psi_j\}_1^\infty$, where $\Psi_j = V\psi_j$, is an orthonormal basis in $V(L^2(T))$. It should be noted that the representation (1.5) is also useful for simulations.

The paper is organized as follows. In Section 2 we give some preliminaries on fractional calculus, define Volterra type processes and introduce the reproducing kernel Hilbert space associated with a Gaussian process. In Section 3 we state the general representation (1.5) for Volterra type processes (processes of the form (1.4)) and obtain as special cases, the Karhunen-Loéve representation for the standard Brownian motion and a wavelet representation (1.3) for the fractional Brownian motion. We also give an approximation of Skorohod-type stochastic integrals with respect to Volterra type processes with deterministic as well as stochastic integrands. In Section 4 we apply the results to parameter estimation and give some examples including estimation of the mean-reverting parameter in an Ornstein-Uhlenbeck process driven by a fractional Brownian motion. We also discuss consistency of the derived estimators.

2. Preliminaries

We begin by reviewing some results on fractional calculus to be used later in the paper. The main reference on fractional calculus is the book by Samko, Kilbas and Marichev (1987) [20]. The definition of the fractional integral on an interval that we will use is the Liouville fractional integral and is given by:

Definition 2.1. For $f \in L^1([a,b])$ and $\alpha > 0$, the integrals

$$\begin{aligned}(I^\alpha_{a^+}f)(t) &\triangleq \frac{1}{\Gamma(\alpha)}\int_a^t (t-s)^{\alpha-1}f(s)ds, \quad t \geq a,\\ (I^\alpha_{b^-}f)(t) &\triangleq \frac{1}{\Gamma(\alpha)}\int_t^b (s-t)^{\alpha-1}f(s)ds, \quad t \leq b,\end{aligned}$$

is called the *right-* and *left fractional integral* of order α, respectively.

One can also define fractional differentiation as follows.

Definition 2.2. For functions f given in the interval $[a,b]$, the *left-handed* and *right-handed fractional derivative* of order $\alpha > 0$, is defined by,

$$\begin{aligned}(D^\alpha_{a^+}f)(t) &\triangleq \left(\frac{d}{dt}\right)^{[\alpha]+1} I^{1-\{\alpha\}}_{a^+}f(t),\\ (D^\alpha_{b^-}f)(t) &\triangleq \left(-\frac{d}{dt}\right)^{[\alpha]+1} I^{1-\{\alpha\}}_{b^-}f(t),\end{aligned}$$

respectively where $[\alpha]$ denotes the integer part of α and $\{\alpha\} = \alpha - [\alpha]$.

The connection between fractional integration and differentiation is given by the following theorem, see Samko et.al. [20], Theorem 2.4, p. 44.

Theorem 2.1. *For $\alpha > 0$ we have,*

$$D^\alpha_{a^+}I^\alpha_{a^+}f = f, \quad \text{for } f \in L^1([a,b]), \tag{2.1}$$

$$I^\alpha_{a^+}D^\alpha_{a^+}f = f, \quad \text{for } f \in I^\alpha_{a^+}(L^1([a,b])). \tag{2.2}$$

Because of this theorem we will sometimes write, $I^{-\alpha}_{a^+}$ for $D^\alpha_{a^+}$. By Corollary 2 on p. 46 in [20], we have the following integration by parts formula:

$$\int_a^b f(t)(D^\alpha_{a^+}g)(t)dt = \int_a^b g(t)(D^\alpha_{b^-}f)(t)dt, \tag{2.3}$$

for $0 < \alpha < 1$ and $f \in I^\alpha_{b^-}(L^p), g \in I^\alpha_{a^+}(L^q), 1/p + 1/q \leq 1 + \alpha$.

We can similarly define fractional integration and differentiation on the real line by analogous expressions:

$$
\begin{aligned}
(I_+^\alpha f)(t) &\triangleq \frac{1}{\Gamma(\alpha)}\int_{-\infty}^{t}(t-s)^{\alpha-1}f(s)ds, \quad t\in\mathbb{R}\\
(I_-^\alpha f)(t) &\triangleq \frac{1}{\Gamma(\alpha)}\int_{t}^{\infty}(s-t)^{\alpha-1}f(s)ds, \quad t\in\mathbb{R},
\end{aligned}
$$

We refer to Samko et.al. [20], Chapter 2, for further details.

2.1. **Gaussian processes of Volterra type.** Let us denote by T an index set which will be a compact interval or the whole of $\mathbb{R}$. We will often use the unit interval and therefore we introduce the notation $I=[0,1]\subset\mathbb{R}$. We may think of $t\in T$ as time. Consider a deterministic function $V: T\times T\to\mathbb{R}_+$ satisfying the following hypothesis:

$$
(H)\begin{cases}
(1)\ V(0,s)=0 \text{ for all } s\in T \text{ and } V(t,s)=0 \text{ for } s>t.\\
(2)\ \text{There are constants } C,\alpha>0 \text{ such that for all } s,t\in T\\
\qquad \int_T (V(t,r)-V(s,r))^2\,dr \le C|t-s|^\alpha.
\end{cases}
$$

Let $B=\{B_t\}_{t\in T}$ denote the standard Brownian motion on T. The standard Brownian motion on $T=\mathbb{R}$ is two independent Brownian motions starting at time 0 running in opposite time directions. Let $X=\{X_t\}_{t\in T}$ be the process defined by

$$
X_t=\int_T V(t,s)dB_s, \quad t\in T. \tag{2.4}
$$

Since (H2) implies that $V(t,\cdot)\in L^2(T)$ for each $t\in T$, the process X is well defined. Clearly X is Gaussian and has covariance function

$$
\rho(t,s)=\mathbb{E}(X_tX_s)=\int_T V(t,r)V(s,r)dr. \tag{2.5}
$$

Definition 2.3. A process $X=\{X_t\}_{t\in T}$ with representation (2.4) and kernel V satisfying (H) is called a *Volterra type process.*

The hypothesis (H) guarantees the existence of a continuous version of the process X. Indeed,

$$
\begin{aligned}
\mathbb{E}(|X_t-X_s|^p) &= \mathbb{E}\left(\left|\int_T V(t,r)dB_r-\int_T V(s,r)dB_r\right|^p\right)\\
&\le \left(\mathbb{E}\left|\int_T V(t,r)dB_r-\int_T V(s,r)dB_r\right|^2\right)^{p/2}\\
&\le C\left(\int_T (V(t,r)-V(s,r))^2\,dr\right)^{p/2}\\
&\le C|t-s|^{\alpha p/2}\\
&\le C|t-s|^{1+\delta},
\end{aligned}
$$

for some $\delta>0$ and $p>2/\alpha$. Hence, Kolmogorov's condition is satisfied and a continuous version exist. Therefore, for any compact set $\Lambda\subset\mathbb{R}$, we can define a measure $\mathbb{P}$ on $\Omega=C(\Lambda)$, the space of continuous functions on Λ vanishing at zero, with X as coordinate process. We also note that V induces an integral transform of functions in $L^2(T)$,

$$
(Vf)(t)=\int_T V(t,s)f(s)ds.
$$

The image of $L^2(T)$ under this integral transform consist of continuous functions. Clearly, by (H2) and Hölder's inequality the function $(Vf)(t)$, $f \in L^2(T)$, is continuous. Finally, we remark that by (H1) X is adapted to the natural filtration generated by B. The following are our primary examples.

Example 2.1 (Brownian motion). Let $T = I$ and

$$V(t,s) = \mathbf{1}_{[0,t]}(s).$$

Then V satisfies (H) and X is the standard Brownian motion with covariance function $\rho(t,s) = t \wedge s$. We will denote standard Brownian motion by $\{B_t\}_{t\in T}$. Note that, as an integral operator acting on functions in $L^2(T)$, V is simply the integration operator:

$$(Vf)(t) = \int_T V(t,s)f(s)ds = \int_0^t f(s)ds =: (I^1_{0+}f)(t).$$

Example 2.2 (Ornstein-Uhlenbeck process). Let $T = I$ and

$$V(t,s) = e^{\theta(t-s)}\mathbf{1}_{[0,t]}(s).$$

Then V satisfies (H) and X is the Ornstein-Uhlenbeck process. That is, the solution to the stochastic differential equation

$$dX_t = \theta X_t dt + dB_t, \quad X_0 = 0.$$

This can easily be verified using Itô's formula.

Example 2.3 (Fractional Brownian motion). For $H \in (0,1)$, let $T = I$ and $V(t,s) = K_H(t,s)$ where,

$$K_H(t,s) \triangleq \frac{1}{\sqrt{V_H}\Gamma(H+\frac{1}{2})}(t-s)^{H-\frac{1}{2}}{}_1F_2(H-\frac{1}{2},\frac{1}{2}-H,H+\frac{1}{2},1-\frac{t}{s})\mathbf{1}_{[0,t)}(s),$$

with ${}_1F_2$ the Gauss hypergeometric function and

$$V_H \triangleq \frac{\Gamma(2-2H)\cos(\pi H)}{\pi H(1-2H)} \tag{2.6}$$

a normalizing constant which makes $\mathbb{E}\left[X(1)^2\right] = 1$. Then V satisfies (H) and X is the fractional Brownian motion with index H, see Decreusefond and Üstünel (1999) [8]. Note that, as an integral operator acting on functions in $L^2(I)$, V is the operator:

$$(Vf)(t) = \int_I V(t,s)f(s)ds = \int_0^t K_H(t,s)f(s)ds = (K_Hf)(t),$$

which is an isomorphism from $L^2(I)$ onto $I^{H+1/2}_{0+}(L^2(I))$ and

$$\sqrt{V_H}K_H f = \begin{cases} I^{2H}_{0+}t^{1/2-H}I^{1/2-H}_{0+}t^{H-1/2}f, & \text{for } H \le 1/2, \\ I^1_{0+}t^{H-1/2}I^{H-1/2}_{0+}t^{1/2-H}f, & \text{for } H \ge 1/2, \end{cases}$$

see Samko et.al. [20] p. 187. The fBm has stationary increments and covariance function

$$\rho(t,s) = \mathbb{E}(X_tX_s) = \frac{1}{2}\left(t^{2H} + s^{2H} - |t-s|^{2H}\right).$$

We will denote fBm by $\left\{B^H_t\right\}_{t\in I}$. One of the most important properties of the fractional Brownian motion is self-similarity in the sense that for any $c > 0$, $\left\{B^H(ct), t \in I\right\} \overset{d}{=} \left\{c^H B^H(t), t \in I\right\}$.

One can also define the fractional Brownian directly on the whole real line. This moving average representation is often used in the literature. For $H \in (0,1)$ let

$$M_H(t,s) = \frac{1}{C_1(H)}\left((t-s)^{H-1/2}_+ - (-s)^{H-1/2}_+\right), \quad s,t \in \mathbb{R},$$

where $u_+ = \max(0,u)$ and

$$C_1(H) = \left\{ \int_0^\infty \left((1+s)^{H-1/2} - s^{H-1/2} \right)^2 ds + \frac{1}{2H} \right\}^{1/2}.$$

Then X is the fractional Brownian motion, see Samorodnitsky and Taqqu (1994) [21] p. 321. Note that, as an integral operator acting on functions in $L^2(\mathbb{R})$, M_H can be written as

$$(M_H f)(t) = \int_{-\infty}^{t\vee 0} \left((t-s)_+^{H-1/2} - (-s)_+^{H-1/2} \right) f(s) ds,$$

This operator can be expressed in terms of the Marchaud fractional derivative, see Pipiras and Taqqu (2000) [19] and Samko et.al. [20]. It should be noted that for $t<0$, M_H is not a Volterra type operator.

Since the kernel of the fractional Brownian motion is fairly complicated one sometimes modifies it to simplify computations. One such modification is given in the next example where we simply remove the Gauss hypergeometric function (and its singularity at zero). This process is sometimes referred to as fractional Brownian motion of type II, fBm(II).

Example 2.4 (Fractional Brownian motion of type II). For $H \in (0,1)$, let $T = I$ and $V(t,s) = J_H(t,s)$ where,

$$J_H(t,s) \triangleq \frac{1}{\Gamma(H+\frac{1}{2})} (t-s)^{H-\frac{1}{2}} \mathbf{1}_{[0,t)}(s).$$

Then V satisfies (H) and X is fractional Brownian motion of type II, see Feyel and la Pradelle (1999) [9]. It is convenient to use the parameterization $\alpha = H + 1/2$ when working with fBm(II). The fBm(II) has non-stationary increments and covariance function

$$\rho(t,s) = \mathbb{E}(X_t X_s) = \frac{1}{\Gamma(\alpha)^2} \int_0^{t\wedge s} (t-r)^{\alpha-1}(s-r)^{\alpha-1} dr.$$

We will denote fBm(II) by $\{W_t^H\}_{t\in I}$. Note that, as an integral operator acting on functions in $L^2(I)$, V is the Liouville fractional integration operator:

$$(Vf)(t) = \int_I V(t,s) f(s) ds = \int_0^t J_H(t,s) f(s) ds = (I_{0+}^{\alpha} f)(t),$$

see Feyel and la Pradelle [9].

2.2. **The reproducing kernel Hilbert space.** The general definition of a reproducing kernel is:

Definition 2.4. A function ρ defined on $T \times T$ is said to be a *reproducing kernel* for the Hilbert space $\mathcal{H}$ of functions on T if

(1) $\rho(t,\cdot) \in \mathcal{H}$ for each $t \in T$
(2) $\langle G, \rho(t,\cdot)\rangle_{\mathcal{H}} = G(t)$, for each $G \in \mathcal{H}$.

For Gaussian processes we can construct the reproducing kernel Hilbert space with reproducing kernel $\rho(t,s) = \mathbb{E}[X_t X_s]$ as follows. Let H be the closure in $L^2(\Omega)$ of the space spanned by $\{X_t\}_{t\in T}$ equipped with the inner product $\langle \xi, \zeta\rangle_H = \mathbb{E}(\xi\zeta)$, for $\xi, \zeta \in H$. The reproducing kernel Hilbert space $\mathcal{H}$ associated with X is the space $R(H) = \{R(\xi) : \xi \in H\}$ where for any $\zeta \in H$, $R(\zeta)$ is the function $R(\zeta)(t) = \langle \zeta, X_t\rangle_H = \mathbb{E}(\zeta X_t)$. $\mathcal{H}$ has inner product $\langle F, G\rangle_{\mathcal{H}} = \langle R^{-1}F, R^{-1}G\rangle_H$. Since $\rho(s,\cdot) = R(X_s)$, we clearly have $\rho(s,\cdot) \in \mathcal{H}$. Furthermore, (2) is satisfied since for $\xi \in H$, $\langle R(\xi), \rho(s,\cdot)\rangle_{\mathcal{H}} = \mathbb{E}(\xi X_s) = R(\xi)(s)$. For more details on reproducing kernel Hilbert spaces we refer to Grenander (1981) [10] or Janson (1997) [11].

For Volterra type processes we have the expression (2.5) of the covariance function. Then the reproducing kernel Hilbert space can also be represented as the image of $L^2(T)$ under the integral transform V, $\mathcal{H} = V(L^2(T))$ equipped with the inner product, $\langle F, G\rangle_{\mathcal{H}} = \langle V^{-1}F, V^{-1}G\rangle_{L^2(T)}$. Indeed, it is clear that $V(L^2(T))$ is a Hilbert space with $\rho(t,\cdot) \in V(L^2(T))$, proving (1) and for any $F = Vf$ we have

$$\langle F, \rho(t,\cdot)\rangle_{\mathcal{H}} = \langle f, V(t,\cdot)\rangle_{L^2(T)} = \int_T V(t,u)f(u)ds = F(t).$$

which proves (2).

We also note that by Theorem 3 in Kallianpur (1971) [12] we have that, for compact T, the closure of $\mathcal{H}$ in $C(T)$ is equal to the support of $\mathbb{P}$. This will be an important property in the applications to parameter estimation in Section 4.

Example 2.5 (Brownian motion). For the standard Brownian motion on $T = I$ we saw that V is just the integration operator. Therefore the reproducing kernel Hilbert space is simply the Sobolev space $W_0^{1,2}$ of differentiable functions, vanishing at zero, with first derivative in $L^2(T)$.

Example 2.6 (Fractional Brownian motion on I). For the fractional Brownian motion on $T = I$ we have $V = K_H$. Since K_H is an isomorphism from $L^2(I)$ onto $I_{0+}^{H+1/2}(L^2(I))$ the reproducing kernel Hilbert space is $I_{0+}^{H+1/2}(L^2(I))$ with the inner product $\langle F, G\rangle_{\mathcal{H}} = \langle K_H^{-1}F, K_H^{-1}G\rangle_{L^2(I)}$.

Example 2.7 (Fractional Brownian motion on $\mathbb{R}$). For the fractional Brownian motion on $T = \mathbb{R}$ we have the representation $X_t = \int_{-\infty}^{t\vee 0} M_H(t,s)dB_s$ which is not a Volterra type representation for $t < 0$. However, it is easy to check that the same arguments as in the case of Volterra processes hold to deduce that the reproducing kernel Hilbert space is the space $M_H(L^2(\mathbb{R}))$. It can be further characterized using fractional integration and differentiation operators, see Pipiras and Taqqu [19].

Example 2.8 (Fractional Brownian motion of type II). For $V = I_{0+}^{\alpha}$, i.e. in the case of fractional Brownian motion of type II on $T = I$, the reproducing kernel Hilbert space is the space $I_{0+}^{\alpha}(L^2(T))$ with the inner product $\langle F, G\rangle_{\mathcal{H}} = \langle I_{0+}^{-\alpha}F, I_{0+}^{-\alpha}G\rangle_{L^2(T)}$. That is, the fractional Sobolev space $W_0^{\alpha,2}$.

3. Representation of Volterra type stochastic integrals

In this section we apply a general representation theorem for Gaussian processes to Volterra type processes. We give as examples, how to obtain the Karhunen-Loéve decomposition, the classical Lévy construction of standard Brownian motion and also a wavelet representation for the fractional Brownian motion studied in Meyer et.al. [15]. In the next section we will use this representation for parameter estimation in a fractional Brownian motion with drift and a fractional Ornstein-Uhlenbeck process. The general representation theorem mentioned above is the following, which is proved in Janson (1997) [11], Theorem 8.22.

Theorem 3.1. *Suppose that $\{X_t\}_{t\in\Lambda}$ is a Gaussian process on some index set Λ and $\{\Psi_j\}_1^\infty$ is a countable orthonormal basis in the associated reproducing kernel Hilbert space $\mathcal{H}$. Then there exist independent standard normal random variables $\{\xi_j\}_1^\infty$ such that for each $t \in \Lambda$,*

$$X_t = \sum_j \Psi_j(t)\xi_j \tag{3.1}$$

with the sum converging in $L^2(\Omega)$ and almost surely.

Conversely, for any sequence $\{\xi_j\}_1^\infty$ of independent standard normal random variables, the sum converges almost surely for each $t \in \Lambda$ and defines a centered Gaussian process with the same distribution as $\{X_t\}_{t\in\Lambda}$.

As indicated in Remark 8.24 in [11] we can prove the following corollary.

Corollary 3.1. *If X is a.s. continuous and $T \subset \Lambda$ is compact then the sum in (3.1) converges almost surely uniformly on T.*

Proof. Since X is a.s. continuous, $\mathcal{H}$ consist of continuous functions on Λ. We consider the compact subset $T \subset \Lambda$ and we may take $\Omega = C(T)$. We have in particular that each $\Psi_j(\cdot)$ restricted to T is in $C(T)$ so the partial sums,

$$S_n(\cdot) = \sum_{j=1}^{n} \Psi_j(\cdot)\xi_j$$

form a random sequence of elements in $C(T)$. Now, since T is compact, $C(T)^* \cong M(T)$, where $M(T)$ is the space of finite signed regular Borel measures on T. On Λ the function $\rho(t,\cdot)$ can be expanded in the basis $\{\Psi_j\}_1^\infty$ and we find that

$$\rho(t,\cdot) = \sum_j \Psi_j(\cdot)\langle \Psi_j(\cdot), \rho(t,\cdot)\rangle_{\mathcal{H}} = \sum_j \Psi_j(\cdot)\Psi_j(t).$$

Furthermore, $\mathbb{E}\left[S_n(t)^2\right] = \sum_{j=1}^{n} \Psi_j(t)^2 \leq \sum_{j=1}^{\infty} \Psi_j(t)^2 = \rho(t,t)$ and $\mathbb{E}\left[X(t)^2\right] = \rho(t,t)$. Since $S_n(t) \to X(t)$ in $L^2(\Omega)$ for each t, and $\rho(t,t)$ is bounded on T, the bounded convergence theorem implies that for $\mu \in M(T)$,

$$\int_T S_n(t)d\mu(t) \to \int_T X(t)d\mu(t), \quad \text{in } L^2(\Omega).$$

The statement now follows from the Lévy-Itô-Nisio Theorem, see Ledoux and Talagrand (1991) [13], Theorem 2.4. □

In our quest for finding an explicit representation we are left with the problem of finding an orthonormal basis in $\mathcal{H}$. This is however very simple because $\mathcal{H} = V(L^2(T))$. Simply take any orthonormal basis $\{\psi_j\}_1^\infty$ in $L^2(T)$ and apply the integral transform V on each function. Then $\{\Psi_j\}_1^\infty$ with $\Psi_j(\cdot) = (V\psi_j)(\cdot)$, is an orthonormal basis in $V(L^2(T))$. Indeed,

$$\langle \Psi_j, \Psi_k\rangle_{V(L^2(T))} = \langle V^{-1}\Psi_j, V^{-1}\Psi_k\rangle_{L^2(T)} = \langle \psi_j, \psi_k\rangle_{L^2(T)} = 0, \quad j \neq k$$

and

$$\|\Psi_j\|_{V(L^2(T))} = \|\psi_j\|_{L^2(T)} = 1.$$

Example 3.1 (Brownian motion). Let $\{\psi_j\}_0^\infty$ be the orthonormal basis of $L^2(I)$ defined by

$$\psi_j(t) = \sqrt{2}\cos(t/\lambda_j),$$

where $\lambda_j = 1/\pi(j+1/2)$. Then the functions

$$\Psi_j(t) = (I_{0+}^1\psi_j)(t) = \int_0^t \sqrt{2}\cos(r/\lambda_j)dr = \lambda_j\sqrt{2}\sin(t/\lambda_j),$$

form an orthonormal basis in the reproducing kernel Hilbert space $W_0^{1,2}$ and

$$B_t = \sum_j \lambda_j\sqrt{2}\sin(t/\lambda_j)\xi_j,$$

where $\{\xi_j\}_0^\infty$ is a sequence of i.i.d. $N(0,1)$ random variables. This is the standard Karhunen-Loéve expansion for Brownian motion.

Example 3.2 (Brownian motion). Let $\{\psi_n\}_0^\infty$ be the Haar basis of $L^2(I)$ defined by,

$$\begin{aligned}
\psi_0(t) &= \mathbf{1}_{[0,1]}(t), \\
\psi_n(t) &= 2^{j/2}\psi\left(2^j t - k\right), \quad n = 2^j + k,\ j \geq 0 \text{ and } 0 \leq k < 2^j.
\end{aligned}$$

where $\psi(t) = 1$ on $[0, 1/2)$ and $\psi(t) = -1$ on $[1/2, 1]$. Integrating gives us the basis $\{\Psi_n\}_0^\infty$ of the reproducing kernel Hilbert space,

$$\begin{aligned} \Psi_0(t) &= t, \\ \Psi_n(t) &= 2^{-j/2}\Psi\left(2^j t - k\right), \qquad n = 2^j + k, \end{aligned}$$

where $\Psi(\cdot)$ is the primitive of $\psi(\cdot)$,

$$\Psi(t) = \frac{1}{2}\max(0, 1 - |2t-1|).$$

We obtain the representation,

$$B_t = \sum_n \Psi_n(t)\xi_n,$$

where $\{\xi_n\}_0^\infty$ is an i.i.d. $N(0,1)$ sequence. This is a classical construction of Brownian motion by Lévy, see Steele (2001) [22] for a nice exposition.

Remark 3.1. We see that any orthonormal $L^2(T)$-basis, $\{\psi_j\}_1^\infty$, yields a representation of Brownian motion as,

$$B_t = \sum_j \Psi_j(t)\xi_j,$$

where, $\Psi_j(t) = \int_0^t \psi(s)ds$ and $\{\xi_j\}_1^\infty$ is a sequence of i.i.d. $N(0,1)$ random variables.

Example 3.3 (Fractional Brownian motion). Consider the fractional Brownian motion on $T = \mathbb{R}$ with the kernel $M_H(t,s)$. Let $\{\phi, \psi\}$ be a wavelet pair such that $\{\phi_k, \psi_{j,k}\}$, where,

$$\begin{aligned} \phi_k(t) &= \phi(t-k) \quad k = 0, \pm 1, \pm 2, \ldots \\ \psi_{j,k}(t) &= 2^{j/2}\psi(2^j t - k), \quad j \geq 0,\ k = 0, \pm 1, \pm 2, \ldots \end{aligned}$$

form a basis in $L^2(\mathbb{R})$. For instance we may take the Meyer wavelets, see Meyer et.al. [15] and references therein. We will derive the wavelet representation of fractional Brownian motion given in the recent paper by Meyer et.al. [15] as a special case of (3.1). Our goal is therefore to obtain the representation,

$$B^H(t) = \sum_{k=-\infty}^{\infty} S_k^H \Phi_H(t-k) + \sum_{j\geq 0}\sum_{k=-\infty}^{\infty} 2^{-jH}\Psi_H(2^j t - k)\epsilon_{j,k} - b_0, \qquad (3.2)$$

where convergence holds uniformly on compact intervals, $\{S_k\}_{-\infty}^{\infty}$ is a fractional ARIMA and $\{\epsilon_{j,l}\}$ is a sequence of i.i.d. $N(0,1)$ random variables, independent of $\{S_k\}$. The random variable b_0 is a correction such that $B^H(0) = 0$.

Since B^H has representation (3.1) and $\{\phi_k, \psi_{j,k}\}$ is an orthonormal basis in $L^2(\mathbb{R})$, there exists independent sequences $\{\eta_k\}$ and $\{\xi_{j,k}\}$ of i.i.d. $N(0,1)$ random variables such that (3.1) can be written as,

$$B^H(t) = \sum_k \eta_k \Phi_k(t) + \sum_{j,k}\xi_{j,k}\Psi_{j,k}(t), \qquad (3.3)$$

where $\Phi_k(t) = (M_H\phi_k)(t)$, $\Psi_{j,k}(t) = (M_H\psi_{j,k})(t)$ and M_H is the integral transform in Example 2.3. By Corollary 3.1, convergence holds almost surely uniformly on compact intervals. It is easy to verify that $\Psi_{j,k}(t)$ coincides with $2^{-j/2}\Psi_H(2^j t - k)$ in [15]. If we define $\tilde{\Phi}(t-k) \triangleq \Phi_k(t) - \Phi_k(t-1)$ and use Lemma 10 in Meyer et.al. [15], we see that the first sum in (3.3) can be rewritten as,

$$\sum_k \eta_k \Phi_k(t) = \sum_k S_k\tilde{\Phi}(t-k) - \sum_k S_k\tilde{\Phi}(0-k),$$

where,

$$\begin{aligned} S_k &= \eta_1 + \eta_2 + \ldots + \eta_k, \quad k \geq 1 \\ S_0 &= 0 \\ S_k &= -\eta_0 - \eta_1 - \ldots - \eta_{k+1}, \quad k \leq -1. \end{aligned}$$

This is not quite the representation in [15]. However, if we express the function $\tilde{\Phi}$ in the function Φ_H by, $\tilde{\Phi} = \langle \tilde{\Phi}, \Phi_H \rangle_{\mathcal{H}} \Phi_H$ we get the desired representation.

3.1. **Representation of stochastic integrals.** In this section we show how to approximate stochastic integrals w.r.t. a Volterra type process X.

There are essentially two different ways to define a stochastic integral of deterministic functions with respect to the Volterra type process X, see Decreusefond (2000) [6]. For our purposes it is convenient to define the stochastic integral as the extension of the isometry:

$$\mathcal{I}_X : L^2(T) \longrightarrow L^2(\Omega) \tag{3.4}$$

$$V(t, \cdot) \longrightarrow X(t). \tag{3.5}$$

One advantage with this approach is that orthogonality relations become straightforward

$$\mathbb{E}[X_t | X_u, u \leq s] = \mathcal{I}_X(V(t, \cdot)\mathbf{1}_{[0,s]}),$$

but it is *not* the limit of Riemann sums. This might indicate that it is difficult to approximate this stochastic integral since we can not approximate it by Riemann sums. However, if we note that the isometry is

$$\mathcal{I}_X = R^{-1} \circ V, \tag{3.6}$$

we can approximate it, using the representation (3.1), see Proposition 3.1 below. We will now also define a stochastic integral stochastic integrands with respect to X. It turns out, the natural stochastic integral in our context is the Skorohod integral, which can be defined for any Gaussian process. The machinery needed to deal with these stochastic integrals is quite extensive and we refer to Nualart (1995) [16] for a more comprehensive introduction on Malliavin calculus. Let the index set T be compact and put $\Omega = C(T)$. The dual space Ω^* of Ω is the space $M(T)$ of finite signed regular Borel measures on T. Let $F : \Omega \to \mathbb{R}$ be a functional of the form

$$F(\omega) = f(\langle \omega, \eta_1 \rangle_{\Omega,\Omega^*}, \ldots, \langle \omega, \eta_n \rangle_{\Omega,\Omega^*}),$$

where $\eta_i \in \Omega^*$, $i = 1, 2, \ldots, n$. If f belongs to the space $C_p^\infty(\mathbb{R}^n)$ of infinitely continuously differentiable functions with all its partial derivatives having polynomial growth, we call F a *smooth cylindrical functional* and denote by $\mathcal{S}$ the space of all smooth cylindrical functionals. Let R^* and V^* denote the adjoints of R and V, respectively. We introduce the derivative of $F \in \mathcal{S}$ as,

$$\nabla F(\omega) = \sum_{j=1}^{n} \partial_j f(\langle \omega, \eta_1 \rangle_{\Omega,\Omega^*}, \ldots, \langle \omega, \eta_n \rangle_{\Omega,\Omega^*}) \rho(\eta_j),$$

where $\rho : \Omega^* \to \mathcal{H}$ is the mapping $R \circ R^*$, or equivalently $V \circ V^*$. We denote by δ the divergence operator, i.e. the adjoint of ∇. It is characterized by the relation

$$\mathbb{E}[F\delta(U)] = \mathbb{E}[\langle \nabla F, U \rangle_{\mathcal{H}}], \tag{3.7}$$

for $U \in L^2(\Omega; \mathcal{H})$ and $F \in \mathcal{S}$.

Definition 3.1. For a process $\{u_t\}_{t \in T}$ such that $\mathbb{E}\|u\|_{L^2(T)}^2 < \infty$, the *stochastic integral* with respect to X is defined as,

$$\int_T u_s \delta X_s \triangleq \delta(Vu).$$

For adapted processes $\{u_t\}_{t\in T}$ this definition is nothing but the Ito integral w.r.t. the standard Brownian motion $\{B_t\}_{t\in T}$ in the representation (2.4) of the Volterra process, see Decreusefond and Üstünel [8] Theorem 4.8.

Using the decomposition (3.1) we can approximate these stochastic integrals for integrals as the following proposition shows.

Proposition 3.1. *Let $\{\psi_j\}_1^\infty$ be an orthonormal basis in $L^2(T)$ and X a Volterra type process with kernel V. Then there exist a sequence $\{\xi_j\}_1^\infty$ of i.i.d. $N(0,1)$ random variables such that the following statements hold.*

(1) *If $f \in L^2(T)$ is a deterministic function and $f_j = \langle f, \psi_j\rangle_{L^2(T)}$, then*

$$\mathcal{I}_X(f) = \sum_{j=1}^{\infty} f_j \xi_j, \quad a.s.$$

(2) *If $u \in L^2(T;\mathcal{H})$ is an adapted process and $u_j = \langle u, \psi_j\rangle_{L^2(T)}$, then*

$$\delta(Vu) = \sum_{j=1}^{\infty} u_j \xi_j, \quad a.s.$$

(3) *Let $u \in L^2(T;\mathcal{H})$ be any process and $u_j = \langle u, \psi_j\rangle_{L^2(T)}$. If ∇u_j exists for all $j \geq 1$ and $\mathbb{E}[(\sum_{j=1}^\infty |u_j|)^2] < \infty$, then*

$$\delta(Vu) = \sum_{j=1}^{\infty} u_j \xi_j - \sum_{j=1}^{\infty} \langle \nabla u_j, \Psi_j\rangle_{\mathcal{H}}, \quad a.s.$$

Proof. (1) Since $\{\psi_j\}_1^\infty$ is an orthonormal basis in $L^2(T)$ it follows that $\Psi_j = V\psi_j$ form an orthonormal basis of the reproducing kernel Hilbert space and therefore $\{\xi_j\}_1^\infty$, where $\xi_j \triangleq R^{-1}\circ V\psi_j$, is an orthonormal basis in H. Hence, $\{\xi_j\}_1^\infty$ is a sequence of i.i.d. $N(0,1)$ random variables. Writing the function f as, $f(t) = \sum_{j=1}^\infty f_j\psi_j(t)$, where $f_j = \langle f, \psi_j\rangle_{L^2(T)}$ and using (3.6) yields,

$$\mathcal{I}_X(f) = R^{-1}\circ Vf = R^{-1}\circ V\left(\sum_{j=1}^{\infty} f_j\psi_j\right) = R^{-1}\left(\sum_{j=1}^{\infty} f_j\Psi_j\right) = \sum_{j=1}^{\infty} f_j\xi_j.$$

Since $f\in L^2(T)$ we have that the sequence $\{f_j\}_1^\infty$ is in $l^2(\mathbb{N})$ and $S_n \triangleq \sum_{j=1}^n f_j\xi_j$, is a square integrable martingale so convergence follows from the martingale convergence theorem. This proves (1).

(2) First we show the result for an elementary adapted process u^n and then we use a limit argument for general adapted processes u. Suppose that u is an elementary adapted process of the form $u(t) = \sum_{i=1}^n F_i\mathbf{1}_{(t_i,t_{i+1}]}(t)$, where $F_i \in \mathcal{S}$ and F_i is $\mathcal{F}_{t_i}$-measurable, where $\{\mathcal{F}_t\}_{t\in T}$ denotes the natural filtration generated by X. By (3.7) and the product rule for the derivative operator we have for any $G\in\mathcal{S}$,

$$\begin{aligned}
\mathbb{E}\Big[G\delta(V\sum_{i=1}^n F_i\mathbf{1}_{(t_i,t_{i+1}]})\Big] &= \mathbb{E}\Big[\sum_{i=1}^n \big(\langle\nabla G, \sum_{i=1}^n F_i V\mathbf{1}_{(t_i,t_{i+1}]}\rangle_{\mathcal{H}}\big)\Big] \\
&= \mathbb{E}\Big[\sum_{i=1}^n \big(\langle\nabla(GF_i), V\mathbf{1}_{(t_i,t_{i+1}]}\rangle_{\mathcal{H}} - \langle G\nabla F_i, V\mathbf{1}_{(t_i,t_{i+1}]}\rangle_{\mathcal{H}}\big)\Big] \\
&= \mathbb{E}\Big[\sum_{i=1}^n \big(GF_i\delta(V\mathbf{1}_{(t_i,t_{i+1}]}) - G\langle\nabla F_i, V\mathbf{1}_{(t_i,t_{i+1}]}\rangle_{\mathcal{H}}\big)\Big] \\
&= \mathbb{E}\Big[\sum_{i=1}^n \big(GF_i\delta(V\mathbf{1}_{(t_i,t_{i+1}]}) - G\int_T V^{-1}(\nabla F_i)(t)\mathbf{1}_{(t_i,t_{i+1}]}(t)dt\big)\Big] \\
&= \mathbb{E}\Big[G\sum_{i=1}^n F_i\delta(V\mathbf{1}_{(t_i,t_{i+1}]})\Big],
\end{aligned}$$

where the last equality holds since $(\nabla F_i)(t) = 0$ on $(t_i, T]$ by Proposition 1.2.4 in Nualart [16] and this implies that $(V^{-1}\nabla F_i)(t) = 0$ on $(t_i, T]$. Hence,

$$\delta(V \sum_{i=1}^{n} F_i \mathbf{1}_{(t_i,t_{i+1}]}) = \sum_{i=1}^{n} F_i \delta(V \mathbf{1}_{(t_i,t_{i+1}]}).$$

By (1), $\delta(V\mathbf{1}_{(t_i,t_{i+1}]}) = \sum_{j=1}^{\infty} f_j \xi_j$ where $f_j = \int_T \sum_{i=1}^{n} \mathbf{1}_{(t_i,t_{i+1}]}(t)\psi_j(t)dt$ so we get,

$$\delta(V \sum_{i=1}^{n} F_i \mathbf{1}_{(t_i,t_{i+1}]}) = \sum_{j=1}^{\infty}\sum_{i=1}^{n} \xi_j \int_T \sum_{i=1}^{n} F_i \mathbf{1}_{(t_i,t_{i+1}]}(t)\psi_j(t)dt = \sum_{j=1}^{\infty} \xi_j u_j.$$

Any adapted process in $L^2(\Omega \times T)$ can be approximated by a sequence u^n of simple adapted processes in the norm of $L^2(\Omega \times T)$. Furthermore, $u^n \to u$ in $L^2(\Omega \times T)$ implies that $u_j^n \triangleq \langle u^n, \psi_j \rangle_{L^2(T)} \to u_j$ in $L^2(\Omega)$. Since δ is closed (see Nualart [16] Section 1.3) and V is continuous $\delta(Vu^n) \to \delta(Vu)$ in $L^2(\Omega)$ which completes the proof of (2).

(3) Observe that for deterministic functions $f \in L^2(T)$ the integrals $\delta(Vf)$ and $\mathcal{I}_X(f)$ coincide. We expand Vu in the basis functions $\{\Psi_j\}_1^{\infty}$ of $\mathcal{H}$, $(Vu)(\cdot) = \sum_{j=1}^{\infty} u_j \Psi_j(\cdot)$, where $u_j = \langle Vu, \Psi_j \rangle_{\mathcal{H}} = \langle u, \psi_j \rangle_{L^2(T)}$ and convergence holds a.s. in $\mathcal{H}$. By (3.7) and the product rule for the derivative operator we have for any $F \in \mathcal{S}$,

$$\begin{aligned}
\mathbb{E}[F\delta(u_j\Psi_j)] &= \mathbb{E}[\langle \nabla F, u_j \Psi_j \rangle_{\mathcal{H}}] \\
&= \mathbb{E}[\langle \nabla(F u_j), \Psi_j \rangle_{\mathcal{H}} - \langle F \nabla u_j, \Psi_j \rangle_{\mathcal{H}}] \\
&= \mathbb{E}[F u_j \delta(\Psi_j) - F\langle \nabla u_j, \Psi_j \rangle_{\mathcal{H}}] \\
&= \mathbb{E}[F(u_j \delta(\Psi_j) - \langle \nabla u_j, \Psi_j \rangle_{\mathcal{H}})].
\end{aligned}$$

Hence, $\delta(u_j\Psi_j) = u_j\Psi_j - \langle \nabla u_j, \Psi_j \rangle_{\mathcal{H}}$. The stochastic integral of Vu with respect to X can now be written as,

$$\begin{aligned}
\delta(Vu) &= \delta(V \sum_{j=1}^{\infty} u_j \psi_j) = \delta(\sum_{j=1}^{\infty} u_j V\psi_j) = \sum_{j=1}^{\infty} u_j \delta(\Psi_j) - \sum_{j=1}^{\infty} \langle \nabla u_j, \Psi_j \rangle_{\mathcal{H}} \\
&= \sum_{j=1}^{\infty} u_j \mathcal{I}_X(\psi_j) - \sum_{j=1}^{\infty} \langle \nabla u_j, \Psi_j \rangle_{\mathcal{H}}.
\end{aligned}$$

The interchange of the sum and the Skorohod integral in the third equaltiy is motivated by the following argument.

$$\begin{aligned}
&\mathbb{E}[F\delta(\sum_{j=1}^{\infty} u_j\Psi_j)] = \mathbb{E}[\langle \nabla F, \sum_{j=1}^{\infty} u_j \Psi_j \rangle_{\mathcal{H}}] = \mathbb{E}[\langle V^{-1}\nabla F, \sum_{j=1}^{\infty} u_j \psi_j \rangle_{L^2(T)}] \\
&= \mathbb{E}[\int_T (V^{-1}\nabla F)(t) \sum_{j=1}^{\infty} u_j \psi_j(t) dt] = \mathbb{E}[\int_T \sum_{j=1}^{\infty} (V^{-1}\nabla F)(t) u_j \psi_j(t) dt] \\
&= \mathbb{E}[\sum_{j=1}^{\infty} \int_T (V^{-1}\nabla F)(t) u_j \psi_j(t) dt] = \mathbb{E}[\sum_{j=1}^{\infty} \langle V^{-1}\nabla F, u_j \psi_j \rangle_{L^2(T)}] \\
&= \mathbb{E}[F \sum_{j=1}^{\infty} \delta(u_j \Psi_j)]
\end{aligned}$$

where the 5:th equality holds if

$$\mathbb{E}[\sum_{j=1}^{\infty} \int_T |(V^{-1}\nabla F)(t) u_j \psi_j(t)| dt] < \infty.$$

By Cauchy-Schwartz inequality,

$$\begin{aligned}
\mathbb{E}\Big[\sum_{j=1}^{\infty}\int_T |(V^{-1}\nabla F)(t)u_j\psi_j(t)|dt\Big] &\le \mathbb{E}\Big[\sum_{j=1}^{\infty}\|V^{-1}\nabla F\|_{L^2(T)}\|u_j\psi_j\|_{L^2(T)}\Big] \\
&= \mathbb{E}\Big[\sum_{j=1}^{\infty}\|\nabla F\|_{\mathcal{H}}|u_j|\Big] = \mathbb{E}\Big[\|\nabla F\|_{\mathcal{H}}\sum_{j=1}^{\infty}|u_j|\Big] \\
&\le \Big(\mathbb{E}[\|\nabla F\|_{\mathcal{H}}^2]\Big)^{1/2}\Big(\mathbb{E}[(\sum_{j=1}^{\infty}|u_j|)^2]\Big)^{1/2} < \infty.
\end{aligned}$$

The first factor is finite since $F \in \mathcal{S}$ and the second factor is finite by assumption. □

4. Applications to parameter estimation

We will now show how the representation (3.1) can be used to estimate parameters in some stochastic differential equations driven by a Volterra type process. In this section we assume that the index set T is a compact interval $[0, T]$.

Suppose that the Volterra type process $\{X_t\}_{t\in T}$ has representation (2.4) and that the kernel V depends on some unknown parameter θ, $V(t,s) = V(t,s;\theta)$, that we want to estimate. We write $V_\theta(t,s)$ to indicate the dependence on θ.

Example 4.1. A simple example of the situation described above is the Ornstein-Uhlenbeck process, see Example 2.2, where

$$X_t = \int_0^t e^{\theta(t-r)}dB_r, \quad t \in T.$$

Example 4.2. Another example is when $\theta = \sigma$ is the (unknown) variance of the process, $V(t,s;\sigma) = \sigma V(t,s)$. This example will be studied in detail in Example 4.3 in the case of fractional Brownian motion.

Assume that we have continuous observations of X on T. If we can compute the coefficients ξ_j in the representation (3.1) and if the dependence on the unknown parameter θ is simple we should be able to construct an estimator of θ based on the first n coefficients, say. This will be illustrated in several examples below. In order to compute the coefficients ξ_j we would like to compute the inner product $\langle X(\cdot), \Psi_j(\cdot)\rangle_{\mathcal{H}}$, where as usual $\{\Psi_j\}_1^\infty$ is an orthonormal basis of $\mathcal{H}$. However, since X is typically not in $\mathcal{H}$ the inner product $\langle X(\cdot), \Psi_j(\cdot)\rangle_{\mathcal{H}}$ does not make sense. To avoid this problem we may take an approximation X^ε of X such that $X^\varepsilon \in \mathcal{H}$. Since $\mathcal{H}$ is dense in the support of $\mathbb{P}$, see Theorem 3 in Kallianpur [12], for each $\varepsilon > 0$, we can find a $X^\varepsilon \in \mathcal{H}$ such that

$$\sup_{t\in T}|X^\varepsilon(t) - X(t)| < \varepsilon. \tag{4.1}$$

The approximation X^ε can be represented in the orthonormal basis $\{\Psi_j\}_1^\infty$ as

$$X^\varepsilon(t) = \sum_j \Psi_j(t)\xi_j^\varepsilon, \quad t \in T,$$

with the sum converging in $\mathcal{H}$ and in $C(T)$. We would expect the coefficient ξ_j^ε to be a good approximation of ξ_j for $j = 1, \dots, n$ if ε is close to zero. The problem is now reduced to finding the coefficients of the approximation X^ε. Using an integration by parts formula for V we can for sufficiently regular basis functions $\psi_j \in L^2(T)$ compute the coefficients ξ_j without having to find an explicit approximation X^ε. To explain this we need the integration by parts formula for the operator V^{-1} and its adjoint $V^{*^{-1}}$ with respect to the $L^2(T)$-inner product, defined by the relation

$$\int_T f(t)(V^{-1}g)(t)dt = \int_T g(t)(V^{*^{-1}}f)(t)dt, \tag{4.2}$$

for $f \in V^*(L^2(T))$ and $g \in V(L^2(T))$. An example is the integration by parts formula (2.3) for fractional differentiation on the interval $[a,b]$ where $V^{-1} = D^{\alpha}_{a+}$ and $\tilde{V}^{-1} = D^{\alpha}_{b-}$. For $\psi_j \in V^*(L^2(T))$, we can compute the coefficients ξ_j^ε as

$$\begin{aligned}\xi_j^\varepsilon &= \langle X^\varepsilon(\cdot), \Psi_j(\cdot)\rangle_{V(L^2(T))} = \langle V^{-1}X^\varepsilon(\cdot), V^{-1}\Psi_j(\cdot)\rangle_{L^2(T)} \\ &= \int_T (V^{-1}X^\varepsilon)(t)V^{-1}(\Psi_j)(t)dt = \int_T (V^{-1}X^\varepsilon)(t)\psi_j(t)dt \\ &= \int_T X^\varepsilon(t)(V^{*^{-1}}\psi_j)(t)dt,\end{aligned}$$

where the last equality follows from the integration by parts formula (4.2). We may rewrite the last expression as,

$$\xi_j^\varepsilon = \int_T X(t)(V^{*^{-1}}\psi_j)(t)dt + \int_T (X^\varepsilon(t) - X(t))\,(V^{*^{-1}}\psi_j)(t)dt.$$

Since ε is arbitrary we can let $\varepsilon \to 0$ and since $X^\varepsilon \to X$ in $C(T)$ as $\varepsilon \to 0$ we find

$$\xi_j = \lim_{\varepsilon\to 0} \xi_j^\varepsilon = \int_0^1 X(t)(V^{*^{-1}}\psi_j)(t)dt.$$

For basis functions $\{\psi_j\}_1^\infty$, such that $\psi_j \in V^*(L^2(T))$, we can compute $(V^{*^{-1}}\psi)(t)$ and the ξ_j's are obtained. This heuristic argument gives us a candidate for ξ_j. The next proposition proves that this candidate is indeed the correct one.

Proposition 4.1. *Let $\{X_t\}_{t\in T}$ be a Volterra type process. Assume that $\{\psi_j\}_1^\infty$ is an orthonormal basis in $L^2(T)$ and the integration by parts formula (4.2) holds. Furthermore suppose that $\psi_j \in V^*(L^2(T))$, $j \geq 1$.*

Then, the coefficients $\{\xi_j\}_1^\infty$ in the representation (3.1) are given by

$$\xi_j = \int_T X(t)(V^{*^{-1}}\psi_j)(t)dt, \quad j = 1, 2, \ldots \tag{4.3}$$

Proof. By the proof of Theorem 3.1 we have that $\xi_j = R^{-1}\Psi_j(t)$, where $\{\Psi_j\}_1^\infty$ is an orthonormal basis in the reproducing kernel Hilbert space. For processes of the form (2.4) we have that the reproducing kernel Hilbert space is $V(L^2(T))$ and hence $\Psi_j = V\psi_j$ is an orthonormal basis in this space. We need to prove that $R(\xi_j)(t) = \Psi_j(t)$ for ξ_j given by (4.3). Indeed, using integration by parts,

$$\begin{aligned}R(\xi_j)(t) &= \mathbb{E}(\xi_j X(t)) = \mathbb{E}\left(X(t)\int_T X(s)(V^{*^{-1}}\psi_j)(s)ds\right) \\ &= \int_T \rho(t,s)(V^{*^{-1}}\psi_j)(s)ds = \int_T (V^{-1}\rho(t,\cdot))(s)\psi_j(s)ds \\ &= \int_T V(t,s)\psi_j(s)ds = (V\psi_j)(t) = \Psi_j(t).\end{aligned}$$

□

We will soon give several examples where this proposition can be applied, but first we show that there is such an integration by parts formula for the kernel of the fractional Brownian motion.

Lemma 4.1. *The operator $K_H^{-1} : K_H(L^2(I)) \to L^2(I)$ defined by*

$$K_H^{-1} f = \begin{cases} t^{1/2-H} D_{0+}^{1/2-H} t^{H-1/2} D_{0+}^{2H} f, & \text{for } H \leq 1/2, \\ t^{H-1/2} D_{0+}^{H-1/2} t^{1/2-H} D_{0+}^{1} f, & \text{for } H \geq 1/2, \end{cases} \tag{4.4}$$

is the inverse operator to K_H and satisfies the integration by parts formula

$$\int_0^1 f(t)(K_H^{-1}g)(t)dt = \int_0^1 g(t)(K_H^{*^{-1}} f)(t)dt,$$

where $K_H^{*^{-1}}$ *is defined by*

$$K_H^{*^{-1}} f = \begin{cases} D_{1^-}^{2H} t^{H-1/2} D_{1^-}^{1/2-H} t^{1/2-H} f, & \text{for } H \le 1/2, \\ D_{1^-}^{1} t^{1/2-H} D_{1^-}^{H-1/2} t^{H-1/2} f, & \text{for } H \ge 1/2. \end{cases}$$

Proof. To prove that K_H^{-1} is the inverse of K_H we need to show that for $f \in K_H(L^2(I))$, $K_H K_H^{-1} f = f$ and that for $g \in L^2(I)$, $K_H^{-1} K_H g = g$. This follows from repeated use of (2.1) and (2.2). The second statement follows from the integration by parts formula (2.3). □

4.1. **Examples.** Next, we give several examples when using the reproducing kernel Hilbert space structure, and the representation (3.1) in particular, becomes effective.

Example 4.3 (Estimating the variance in fBm). In this example we consider the fractional Brownian motion $X = \{X_t\}_{t\in I}$ with unknown volatility σ,

$$X_t = \sigma B_t^H, \quad t \in I.$$

We want to estimate σ from continuous observations of X on I. In this case we have,

$$V(t,s;\sigma) = \sigma K_H(t,s).$$

To derive an estimator, $\hat{\sigma}$, of σ, we use Proposition 4.1 to compute the coefficients ξ_j in the representation (3.1). To apply this proposition we have to use the integration by parts formula (4.2). As an $L^2(T)$ basis we may take the trigonometric basis,

$$\psi_j(t) = \sqrt{2}\cos(t/\lambda_j), \quad \lambda_j = 1/(\pi(j+1/2)), \quad j = 0,1,\ldots$$

We have, $\xi_j = x_j/\sigma$, where

$$x_j = \int_0^1 X(t)(K_H^{*^{-1}}\psi_j)(t)dt.$$

Since the ξ_j's are i.i.d. $N(0,1)$ we find that the x_j's are i.i.d. $N(0,\sigma^2)$ and the likelihood function based on the first n coefficients is,

$$L_n(\sigma) = \prod_{j=1}^n \frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{1}{2}(x_j/\sigma)^2\right).$$

Maximizing w.r.t. σ yields the maximum likelihood estimator $\hat{\sigma}_n^2$ of σ^2 as,

$$\hat{\sigma}_n^2 = \frac{1}{n}\sum_{j=1}^n x_j^2,$$

which has the same distribution as $(\sigma^2/n)\cdot Z$ where Z has Chi-squared distribution with n degrees of freedom. This estimator is unbiased since,

$$\mathbb{E}\left(\hat{\sigma}_n^2 - \sigma^2\right) = \mathbb{E}\left(\frac{\sigma^2}{n}Z - \sigma^2\right) = 0,$$

and consistent as $n \to \infty$,

$$\hat{\sigma}_n^2 - \sigma^2 = \sigma^2 \frac{\sum_{j=1}^n (\xi_j^2 - 1)}{n} \to 0, \quad a.s.$$

by the strong law of large numbers.

The next example illustrates a situation where we have a deterministic drift with an unknown parameter.

Example 4.4 (Estimating the drift parameter). Let $\{Y_t\}_{t\in T}$ satisfy,

$$Y(t) = \theta a(t) + X(t), \quad Y_0 = 0, \tag{4.5}$$

where $a(\cdot)$ is some known deterministic function, θ is an unknown parameter to be estimated and $\{X_t\}_{t\in T}$ is a Volterra type process with kernel V satisfying the integration by parts formula (4.2). We assume that $a(\cdot)$ belongs to $\mathcal{H}$. If we would only have $a(\cdot)$ in $C(T)$ we can at least approximate it by an element in $\mathcal{H}$ since $\mathcal{H}$ is dense in the support of $\mathbb{P}$. We expand $a(\cdot)$ in an orthonormal basis $\{\Psi_j\}_1^\infty$ of $\mathcal{H}$ as, $\sum_{j=1}^\infty a_j \Psi_j(t)$, $t \in T$, where,

$$\begin{aligned} a_j &= \langle a(\cdot), \Psi_j(\cdot)\rangle_{\mathcal{H}} = \langle V^{-1}a(\cdot), V^{-1}\Psi_j(\cdot)\rangle_{L^2(T)} \\ &= \int_T (V^{-1}a)(t)\psi_j(t)dt = \int_T a(t)(V^{*^{-1}}\psi_j)(t)dt, \end{aligned}$$

and the last step follows from the integration by parts formula (4.2). By (3.1) for X we find that Y can be written as,

$$Y(t) = \theta a(t) + X(t) = \theta \sum_j \Psi_j(t)a_j + \sum_j \Psi_j(t)\xi_j.$$

If we put $y_j = \theta a_j + \xi_j$ we get $Y(t) = \sum_j \Psi_j(t)y_j$, where convergence holds a.s. in $C(T)$. To compute the coefficients y_j we use Proposition 4.1 and as in the previous example we may take $\{\psi_j\}_0^\infty$ to be the trigonometric basis. Since $\xi_j = y_j - \theta a_j$ and the ξ_j's are i.i.d. $N(0,1)$, the likelihood function based on the first n coefficients is

$$L_n(\theta) = \prod_{j=1}^n \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{1}{2}(y_j - \theta a_j)^2\right).$$

Maximizing w.r.t. θ yields the maximum likelihood estimator, $\hat{\theta}_n$, of θ as,

$$\hat{\theta}_n = \frac{\sum_{j=1}^n a_j y_j}{\sum_{j=1}^n (a_j)^2}.$$

The estimator $\hat{\theta}_n$ has normal distribution with mean θ and variance $1/\sum_{j=1}^n (a_j)^2$.

In Norros et.al. [18] this example is studied for the fractional Brownian motion with linear drift $a(t) = t$. We will now show that our approach leads to analogous results in this special case.

Example 4.5 (Fractional Brownian motion). Let $T = I$ and consider the process $\{Y_t\}_{t\in I}$ satisfying,

$$Y(t) = \theta t + B^H(t), \quad Y_0 = 0, \tag{4.6}$$

where $\{B_t^H\}_{t\in I}$ is the fractional Brownian motion. By Example 4.4, with $a(t) = t$, the estimator $\hat{\theta}_n$ has normal distribution with mean θ and variance $1/\sum_{j=1}^n a_j^2$. We can explicitly compute the variance as $n \to \infty$. First note that

$$\sum_j a_j^2 = \langle a(\cdot), a(\cdot)\rangle_{\mathcal{H}} = \int_0^1 |(K_H)^{-1}a)(t)|^2 dt. \tag{4.7}$$

We begin to compute $(K_H^{-1}a)(t)$. Using Table 9.1 in Samko et.al. [20] we find that

$$(K_H^{-1}a)(t) = \sqrt{V_H}\frac{\Gamma(3/2-H)}{\Gamma(2-2H)}t^{1/2-H},$$

with V_H as in (2.6). Plugging in this expression in (4.7) yields the variance as $n \to \infty$,

$$\frac{1}{\sum_{j=1}^\infty a_j^2} = \frac{1}{V_H}\frac{(2-2H)\Gamma(2-2H)^2}{\Gamma(3/2-H)^2} = \frac{\pi H(1-2H)(2-2H)\Gamma(2-2H)}{\cos(\pi H)\Gamma(3/2-H)^2}$$

We can compare this estimator with the simple mean $Y_T/T = Y_1$ which is unbiased and has variance 1 for all $H \in (0,1)$. The variance of the estimator $\hat{\theta}_\infty$, as a function of H, is illustrated in figure 1. We conclude from this figure that the estimator $\hat{\theta}_\infty$ has significantly lower variance than the simple mean for small values of H, has the same variance for $H = 1/2$ and has negligible lower variance for $H \in (1/2, 1)$.

Remark 4.1. The results obtained in Example 4.5 are analogous to the results in Norros et.al. [18], where the authors derive and apply a Girsanov transformation. The explanation is that we have in fact a Girsanov type formula of Y in terms of the coefficients in the representation (3.1). Define the measure $\mathbb{Q}$ by

$$\frac{d\mathbb{Q}}{d\mathbb{P}} = \exp\left(\theta \sum_{j=1}^{\infty} a_j \xi_j - \frac{\theta^2}{2} \sum_{j=1}^{\infty} a_j^2\right).$$

Since $a(\cdot) \in V(L^2(T))$ the sequence $\{a_j\}_1^\infty \in l^2(\mathbb{N})$ and by the Cameron-Martin Theorem, see Theorem 7.7, p. 21 in Malliavin (1997) [14] the y_j's are i.i.d. $N(0,1)$ under $\mathbb{Q}$. Hence, the process Y is the fractional Brownian motion under $\mathbb{Q}$. We note that this holds for all drift functions $a(\cdot) \in V(L^2(T))$ and generalizes the results in Norros et.al. [18]. Another way to see this is to note that by Proposition 3.1

$$\sum_{j=1}^{\infty} a_j \xi_j = \mathcal{I}_{B^H}(V^{-1}a) = \delta(a), \text{ and } \sum_{j=1}^{\infty} a_j^2 = \|a\|_{\mathcal{H}}^2,$$

and apply the Girsanov transformation in Decreusefond and Üstünel [8].

Example 4.6 (Stochastic differential equations). Let X be a Volterra type process with kernel V satisfying the integration by parts formula (4.2) and denote the associated reproducing kernel Hilbert space by $\mathcal{H}$, as usual. Suppose that u is an adapted process such that $\mathbb{E}\|u\|_{L^2(T)}^2 < \infty$ and that Vu is observable. Let Y be the process defined by

$$Y(t) = \theta \int_0^t V(t,s)u(s)ds + X(t), \quad t \in T. \tag{4.8}$$

For questions concerning existence of solutions of SDE's driven by fractional Brownian motion we refer to Coutin and Decreusefond (1998) [5]. Since Vu takes values in $\mathcal{H}$ it may be written as, $(Vu)(t) = \sum_{j=1}^{\infty} u_j \Psi_j(t)$, $t \in T$, where, $u_j = \langle u(\cdot), \psi_j(\cdot)\rangle_{L^2(T)}$, and $\{\psi_j\}_1^\infty$ an orthonormal basis in $L^2(T)$. Similar to Example 4.4 we can now derive the maximum likelihood estimator, $\hat{\theta}_n$, of θ as,

$$\hat{\theta}_n = \frac{\sum_{j=1}^n u_j y_j}{\sum_{j=1}^n u_j^2}.$$

The final example before considering consistency of the estimators is estimation of the mean reverting parameter in an Ornstein-Uhlenbeck process driven by the fractional Brownian motion.

Example 4.7 (Estimation of the mean reverting parameter). Let the process $Y = \{Y_t\}_{t\in T}$ satisfy the equation,

$$Y_t = \theta \int_0^t a(Y_s)ds + B_t^H, \quad Y_0 = 0, \tag{4.9}$$

where θ is an unknown parameter to be estimated and $\{B_t^H\}_{t\in T}$ is the fractional Brownian motion. We assume that $a(\cdot)$ is such that there exists a solution to (4.9) with Y having almost surely continuous sample paths and such that $A(t) \triangleq$

$\int_0^t a(Y_s)ds$ belongs to $\mathcal{H}$. By this assumption A may be represented as, $A(t) = \sum_{j=1}^{\infty} A_j \Psi_j(t)$, $t \in T$, where,

$$A_j = \langle A(\cdot), \Psi_j(\cdot) \rangle_{\mathcal{H}} = \langle V^{-1}A(\cdot), V^{-1}\Psi_j(\cdot) \rangle_{L^2(T)}$$
$$= \int_T (K_H^{-1}A)(t)\psi_j(t)dt = \int_T A(t)(K_H^{*^{-1}}\psi_j)(t)dt.$$

Similar to Example 4.4 we can derive the maximum likelihood estimator, $\hat{\theta}_n$, of θ as,

$$\hat{\theta}_n = \frac{\sum_{j=1}^n A_j y_j}{\sum_{j=1}^n A_j^2}.$$

4.2. **Consistency of estimators.** In this last part of the paper we discuss consistency of the estimators in the different examples in the previous section. In this section we assume that all functions and processes are defined on $\mathbb{R}_+$ and the subscript T indicates the restriction to the interval $[0, T]$.

Let X be a Volterra type process with kernel V satisfying the integration by parts formula (4.2) and having $\mathcal{H}_T$ as the associated reproducing kernel Hilbert space on the interval $[0, T]$. We denote by $\{\psi_j^T\}_1^\infty$ an orthonormal basis of $L^2([0, T])$ and $\{\Psi_j^T\}_1^\infty$ the associated basis in $\mathcal{H}_T$. Suppose that $A \in L^2(\Omega_T; \mathcal{H}_T)$ for each $T > 0$, and that A is adapted. We define the process Y by

$$Y(t) = \theta A(t) + X(t), \quad Y_0 = 0.$$

From the examples in the previous section we know that we can derive the maximum likelihood estimator, $\hat{\theta}_\infty^T$, of θ as,

$$\hat{\theta}_\infty^T = \frac{\sum_{j=1}^\infty A_j^T y_j^T}{\sum_{j=1}^\infty (A_j^T)^2}.$$

The following theorem gives a sufficient condition for consistency of the estimator $\hat{\theta}_\infty^T$, as $T \to \infty$.

Theorem 4.1. *Suppose that X, Y, A and $\hat{\theta}_\infty^T$ are as above. If $\sum_{j=1}^\infty (A_j^T)^2 \to \infty$ a.s. as $T \to \infty$, then $\hat{\theta}_\infty^T \to \theta$ a.s.*

Proof. Since $y_j^T = A_j^T + \xi_j^T$ we have that

$$\hat{\theta}_\infty^T = \frac{\sum_{j=1}^\infty A_j^T y_j^T}{\sum_{j=1}^\infty (A_j^T)^2} = \theta + \frac{\sum_{j=1}^\infty A_j^T \xi_j^T}{\sum_{j=1}^\infty (A_j^T)^2}.$$

By Proposition 3.1, $M_T \triangleq \sum_{j=1}^\infty A_j^T \xi_j^T = \delta(A\mathbf{1}_{[0,T]})$ and by a straightforward generalization of Theorem 4.7 in Decreusefond and Üstünel [8] to Volterra type processes, M_T is a martingale with associated Doob-Meyer process

$$\langle M \rangle_T = \int_0^T \left[(V^{-1}A)(s)\right]^2 ds = \|A\|_{\mathcal{H}_T}^2 = \sum_{j=1}^\infty (A_j^T)^2,$$

which tends to infinity almost surely as $T \to \infty$, by assumption. From the law of large numbers for martingales it follows that, as $T \to \infty$,

$$\frac{M_T}{\langle M \rangle_T} \to 0, \quad \text{a.s. on } \{\langle M \rangle_\infty = \infty\}.$$

Hence, $\hat{\theta}_\infty^T \to \theta$ a.s. □

We will now show how to apply this theorem in some of the examples in the previous section.

Example 4.5 (continued). In the case where $A(t) = t$ we computed the sum

$\sum_{j=1}^{\infty} A_j^T = \sum_{j=1}^{\infty} a_j^T$ for $T = 1$. To find the corresponding expression for arbitrary T we proceed as follows. Assume that $\{\psi_j\}_1^{\infty}$ is an orthonormal basis of $L^2(I)$. Then we can construct an orthonormal basis of $L^2(T)$, $\{\psi_j^T\}_1^{\infty}$ by stretching and rescaling. Simply take

$$\psi_j^T(s) = \frac{1}{\sqrt{T}}\psi_j(s/T), \quad j = 1, 2, \ldots$$

Indeed,

$$\begin{aligned}\int_0^T \psi_i^T(s)\psi_j^T(s)ds &= \int_0^1 \psi_i^T(tT)\psi_j(tT)T dt = \int_0^1 \frac{1}{\sqrt{T}}\psi_i(t)\frac{1}{\sqrt{T}}\psi_j(t)T dt \\ &= \int_0^1 \psi_i(t)\psi_j(t)dt = \begin{cases} 1 & \text{if } i = j \\ 0 & \text{if } i \neq j \end{cases}\end{aligned}$$

The ξ_j^T's are i.i.d. $N(0,1)$ and the coefficients a_j^T can be related to a_j as,

$$\begin{aligned}a_j^T &= \int_0^T (K_H^{-1}a)(s)\psi_j^T(s)ds = \int_0^T Cs^{1/2-H}\psi_j^T(s)ds = C\int_0^1 (tT)^{1/2-H}\psi_j^T(tT)T dt \\ &= C\int_0^1 (tT)^{1/2-H}\frac{1}{\sqrt{T}}\psi_j(t)T dt = T^{1-H}\int_0^1 Ct^{1/2-H}\psi_j(t)dt = T^{1-H}a_j.\end{aligned}$$

Therefore, $\sum_{j=1}^{\infty}(a_j^T)^2 = T^{2-2H}\sum_{j=1}^{\infty} a_j^2 \to \infty$, as $T \to \infty$. Hence the condition in Theorem 4.1 is satisfied and the estimator $\hat{\theta}_{\infty}^T$ is consistent.

Example 4.6 (continued). It is straightforward to apply Theorem 4.1 in the case of a stochastic differential equation of the form (4.8). To obtain consistency of the estimator it is sufficient to verify that $\int_0^T u(t)^2 dt \to \infty$ a.s. as $T \to \infty$.

Example 4.7 (continued). To derive consistency in the case of a fractional Ornstein-Uhlenbeck process by applying Theorem 4.1 we need to verify that, as $T \to \infty$,

$$\int_0^T \left[(V^{-1}A)(t)\right]^2 dt \to \infty.$$

Unfortunately, we have not been able to verify this condition and a proof can not be included in this paper.

5. Acknowledgements

The author wishes to express his gratitude to his supervisor Boualem Djehiche for suggesting the topic of this paper and for his continuous encouragement.

References

[1] P. Abry, P. Flandrin, M. Taqqu, and D. Veitch, Wavelets for the analysis, estimation and synthesis of scaling data, in: K. Park and W. Willinger, eds., Self-Similar Network Traffic and Performance Evaluation (Wiley, New York 2000) pp 39-88.

[2] P. Carmona, L. Coutin and G. Montseny, Approximation of some Gaussian processes, Statistical Inference for Stochastic Processes 3 (2000) 161-171.

[3] F. Comte, Simulation and estimation of long memory continuous time models, Journal of Time Series Analysis 17 No. 1 (1996) 19-36.

[4] F. Comte and E. Renault, Long memory continuous time models, Journal of Econometrics 73 (1996) 101-149.

[5] L. Coutin and L, Decreusefond, Stochastic Differential Equations Driven by a Fractional Brownian Motion, Preprint (1998).

[6] L. Decreusefond, A Skorohod-Stratonovich integral for the fractional Brownian motion, Proceedings of the 7-th Workshop on Stochastic Analysis and Related Fields (2000).

[7] L. Decreusefond, Regularity properties of some stochastic Volterra integrals with singular kernel, to appear in Potential Analysis (2000).

[8] L. Decreusefond and A.S. Üstünel, Stochastic Analysis of the Fractional Brownian Motion, Potential Analysis 10 (1999) 177-214.
[9] D. Feyel and A. la Paradelle, On Fractional Brownian Processes, Potential Analysis 10 (1999) 273-288.
[10] U. Grenander, Abstract Inference, John Wiley & Sons (1981).
[11] S. Janson, Gaussian Hilbert Spaces, Cambridge University Press (1997).
[12] G. Kallianpur, Abstract Wiener Processes and Their Reproducing Kernel Hilbert Spaces, Z. Wahrscheinlichkeitstheorie verw. Geb. 17 (1971) 113-123.
[13] M. Ledoux and M. Talagrand, Probability in Banach Spaces, Springer (1991).
[14] P. Malliavin, Stochastic Analysis, Springer (1997).
[15] Y. Meyer, F. Sellan and M. Taqqu, Wavelets, Generalized White Noise and Fractional Integration: The Synthesis of Fractional Brownian Motion, Journal of Fourier Analysis and Applications 5 No. 5 (1999) 465-494.
[16] D. Nualart, The Malliavin Calculus and Related Topics, Springer (1995).
[17] I. Norros, P. Mannersalo and J.L. Wang, Simulation of Fractional Brownian Motion with Conditionalized Random Midpoint Displacement, Advanced Performance Analysis 2 No. 1 (1999) 77-101.
[18] I. Norros, E. Valkeila and J. Virtamo, An elementary approach to a Girsanov formula and other analytical results on fractional Brownian motions, Bernoulli 5 No. 4 (1999) 571-588.
[19] V. Pipiras and M. Taqqu, Integration questions related to fractional Brownian motion, Probability Theory and Related Fields 118 (2000) 251-291.
[20] S.G. Samko, A.A. Kilbas and O.I. Marichev, Fractional Integrals and Derivatives, Theory and Applications, Gordon and Breach Science Publishers (1987).
[21] G. Samorodnitsky and M. Taqqu, Stable Non-Gaussian Random Processes, Chapman and Hall (1994).
[22] J.M. Steele, Stochastic Calculus and Financial Applications, Springer (2001).

(H. Hult) Department of Mathematics, KTH, S-100 44 Stockholm, Sweden
E-mail address: `henrikh@math.kth.se`
URL: `http://www.math.kth.se/~henrikh/`

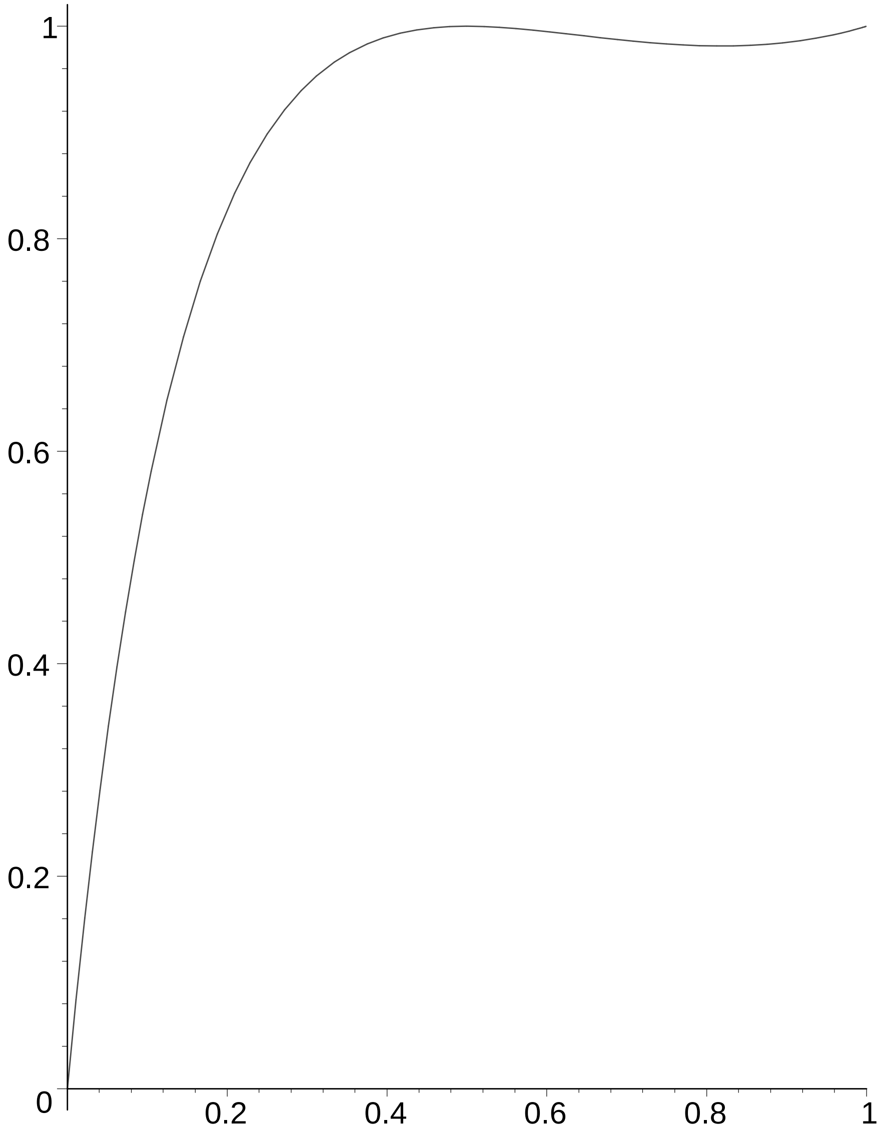

FIGURE 1. Variance of the estimator $\hat{\theta}_\infty$ in Example 4.5.