Self-similarity: Part II— Optimal estimation of fractal processes

Thierry Blu and Michael Unser

The authors are with the Biomedical Imaging Group, Ecole Polytechnique Fédérale de Lausanne (EPFL), CH-1015 Lausanne, Switzerland (e-mail: michael.unser@epfl.ch, phone: +41 21 693 51 75, fax: +41 21 693 37 01)

Abstract

In a companion paper we characterized the class of scale-invariant convolution operators: the generalized fractional derivatives of order γ. We used these operators to specify regularization functionals for a series of Tikhonov-like least squares data fitting problems and proved that the general solution is a fractional spline of twice the order. We investigated the deterministic properties of these smoothing splines and proposed a fast FFT-based implementation.

Here, we present an alternative stochastic formulation to further justify these fractional spline estimators. As suggested by the title, the relevant processes are those that are statistically self-similar; that is, fractional Brownian motion (fBm) and its higher order extensions. To overcome the technical difficulties due to the non-stationary character of fBm, we adopt a distributional formulation due to Gel'fand. This allows us to rigorously specify an innovation model for these fractal processes, which rests on the property that they can be whitened by suitable fractional differentiation. Using the characteristic form of the fBm, we then derive the conditional PDF, $p(B_H(t)|Y)$, where $Y = \{B_H(k) + n[k]\}_{k\in\mathbb{Z}}$ are the noisy samples of the fBm $B_H(t)$ with Hurst exponent H. We find that the conditional mean is a fractional spline of degree $2H$, which proves that this class of functions is indeed optimal for the estimation of fractal-like processes. The result also yields the optimal (MMSE) parameters for the smoothing spline estimator, as well as the connection with kriging and Wiener filtering.

Index Terms

fractional Brownian motion, fractional splines, Wiener filtering, interpolation, self-similar processes, smoothing splines, MMSE estimation

I. Introduction

In the preceding paper [1], we demonstrated the power of the differential formulation of splines by constructing an extended family of fractional splines. These functions are specified in terms of a differential operator L, which, in the present case, is constrained to be scale-invariant (or self-similar). We also investigated an alternative variational formulation which allowed us to recover a subset of these splines (the ones associated with self-adjoint operators) based on the minimization of some scale-invariant "spline energy" involving the same type of operator. We used this deterministic framework to specify a general parametric class of smoothing spline estimators for fitting discrete signals corrupted by noise.

Differential operators also naturally arise in the theory of continuous-time stochastic processes; for instance, it is often possible to specify a process $x(t)$ as the solution of a stochastic differential equation, $\mathrm{L}\{x(t)\} = w(t)$, whose driving term $w(t)$ is white Gaussian noise with variance σ_0^2—this type of representation is often referred to as the innovation model of the process [2], [3]. Now, in the standard case where L is shift-invariant and its inverse is well defined in the L_2-sense (i.e.; $\int_{-\infty}^{+\infty} 1/|L(\omega)|^2 d\omega < \infty$), this

procedure defines a stationary process whose power density is $\Phi_{yy}(\omega) = \sigma_0^2/|L(\omega)|^2$. The interpretation is simply that L is the whitening operator of the process.

Interestingly, there is a perfect parallel between the deterministic differential equations used to define general L-splines, and the stochastic ones just mentioned above. In our previous work, we have taken advantage of this fact to derive an equivalence between spline interpolation and the optimal, continuous-time estimation of stationary processes from their integer samples [4]. In particular, we showed that every continuous-time stationary process with a rational power spectrum has a natural exponential spline space associated with it and that this space contains the optimal solutions of all related minimum mean square error (MMSE) interpolation and estimation problems.

Following this line of thought, it seems quite natural to extend those stochastic results to the classical polynomial splines [5] and their fractional extensions specified in [1]. Unfortunately, this is far less trivial than we would have thought initially because of the lack of correspondence between $1/\omega^\gamma$ spectra and stationary processes. Indeed, the price to pay for self-similary is the zero of order γ in the frequency response of L at $\omega = 0$ which makes the differential system unstable, and substantially complicates the mathematical analysis. Here, as suggested by the title, the relevant stochastic processes are those that are statistically self-similar [6]. These were characterized in 1968 by Mandelbrot and van Ness [7] and named fractional Brownian motion (fBm) because they can be viewed as an extension of Brownian motion, also known as the Wiener process. In this respect, we note that there is an early mention of a link between (thin-plate) splines and fractals in a paper by Szeleski and Terzopoulos in computer graphics, the argument being that both types of entities give rise qualitatively to the same type of frequency behavior [8]. The main difficulty in dealing with fBms is that they are non-stationary[1], meaning that they do not have a well-defined power spectrum. The corollary is that the mathematics of fBm-like processes are much more involved than those of ordinary stationary processes. Yet, it is possible to define some generalized notion of power spectrum for these processes [10]; for fBm, these follow a characteristic $1/\omega^\gamma$ power law, which justifies the common denomination of "$1/f$ noise". There are a few results available on the prediction of such process from their past [9], [11]–[14] and on optimal signal detection in $1/\omega^\gamma$ noise [15]. Also relevant to the topic is the wavelet-based technique for generating fBm-like processes that was proposed by Meyer et al. [16].

[1]It can be shown that there is no mean-square-continuous stationary process whose covariance function is self-similar. However, a careful distributional extension using Gel'fand and Vilenkin's mathematical framework can lead to the definition of Gaussian stationary processes that are self-similar, discontinuous and of infinite power (e.g. white noise) [9]. Qualitatively, these correspond to fBm's with $H < 0$.

The purpose of this paper is to close the gap by showing that the fractional splines are the optimal function spaces for interpolating and estimating fractal processes. A key technical point is the rigorous statement (in the distributional sense) that the fractional derivatives are the whitening operators for fBms and their higher order extensions. This allows us to prove our main theorem which states that the continuous-time Wiener estimator of a generalized fBm, given its (possibly noisy) samples at the integers, is a fractional smoothing spline of order $2\gamma = 2H + 1$ where $H > 0$ is the order of self-similarity (Hurst exponent) of the process.

The paper is organized as follows. In Section II, we introduce the notation while briefly reviewing the key properties of fractional smoothing spline filters. In Section III, we address the issue of the stochastic modeling of self-similar processes. Our contribution is the proposal of an innovation model for the generation of fBm and their extension for Hurst exponents $H > 1$ (cf. Fig. 1). In effect, the fBm is generated by suitable fractional integration of a white noise process. The major difficulty here is of technical nature: it requires giving a precise mathematical meaning to the objects and operations involved (within the context of distribution theory) and searching for an effective way of imposing boundary conditions at the origin that are specific to self-similar processes. To this end, we rely on Gel'fand's theory of generalized stochastic processes (cf. subsection III.B) which allows us to safely manipulate fBms; for instance, compute fractional derivatives and anti-derivatives. In Section IV, we use the proposed distributional characterization of fBm to solve our basic estimation problem. Our main result is an explicit formula for the posterior probability density function $p(x(t_0)|Y)$ for any $t_0 \in \mathbb{R}$ where $Y = \{y[k] = x(k) + n[k]\}_{k\in\mathbb{Z}}$ are the samples of a fBm process $x(t)$ corrupted by additive Gaussian stationary noise n. This automatically yields the optimal spline estimator that includes a rapidly decaying non-stationary component. We then proceed with the derivation of the best "stationary" estimator which turns out to be computationally equivalent to the "deterministic" smoothing spline algorithm investigated in Part I.

II. Fractional spline estimators

In this section, we briefly recall the key features of the fractional smoothing splines that were specified in Part I using a deterministic, variational formulation. In the process, we also adapt the notation slightly to make it more suitable for a statistical formulation.

Our basic estimation problem is to recover an unknown continuous-time function $X = \{x(t)\}_{t\in\mathbb{R}}$ given a sequence of noisy samples $Y = \{y[k] = x(k) + n[k]\}_{k\in\mathbb{Z}}$, where $\{n[k]\}_{k\in\mathbb{Z}}$ is a perturbation signal (noise) whose influence we want to minimize.

The fractional spline estimator of degree α (or, equivalently, of order $2\gamma = \alpha+1$) is a continuous-time function that depends upon the discrete measurements Y and takes the generic form

$$\tilde{x}(t|Y) = \sum_{k\in\mathbb{Z}} c[k]\beta_0^\alpha(t-k) \tag{1}$$

where $\beta_0^\alpha(t)$ is the fractional symmetric B-spline of degree α (cf. first row of Table I). It is uniquely characterized by the sequence of its B-spline coefficients $\{c[k]\}_{k\in\mathbb{Z}}$.

For the class of smoothing spline estimators described in [1], the B-spline coefficients are obtained by appropriate filtering of the input data: $c[k] = (h*y)[k]$. The digital smoothing spline filter h is parametrized by the degree of the spline, a regularization factor $\lambda \geq 0$, and a positive definite convolution kernel $v[k]$ whose discrete Fourier transform is $V(e^{j\omega})$. Here, to simplify the notation, we combine the two latter entities into a single weighting function

$$0 \leq \Lambda(e^{j\omega}) := \frac{\lambda}{V(e^{j\omega})},$$

which we assume to be bounded. Using this formalism, we rewrite the frequency response of the filter (cf. [1, Eq. (20)]) as

$$H(e^{j\omega}) = \frac{1}{|2\sin(\omega/2)|^{\alpha+1}\Lambda(e^{j\omega}) + B_0^\alpha(e^{j\omega})}. \tag{2}$$

where

$$B_0^\alpha(e^{j\omega}) = \sum_{k\in\mathbb{Z}} \beta_0^\alpha(k)e^{-j\omega k}, \tag{3}$$

is the discrete-time Fourier transform of the sampled B-spline basis function.

Putting these elements together, we express the smoothing spline estimator in terms of the input $y[k]$:

$$\tilde{x}(t|Y) = \sum_{k\in\mathbb{Z}} y[k]\varphi_\Lambda(t-k) \tag{4}$$

where

$$\varphi_\Lambda(t) = \sum_{k\in\mathbb{Z}} h[k]\beta_0^\alpha(t-k) \tag{5}$$

is an equivalent fractional spline basis function that represents the "impulse response" of the estimator, keeping in mind that the input Y is discrete, while the output $\tilde{x}$ is analog.

In Part I, we have shown that this spline estimator is optimal in the sense that, for a given input sequence $y[k] \in \ell_2$, it minimizes the regularized error criterion

$$\xi(y,s) = \sum_{k\in\mathbb{Z}} |v*(y[k]-s(k))|^2 + \lambda \int_{-\infty}^{+\infty} |\partial_\tau^\gamma s(t)|^2 \mathrm{d}t,$$

where ∂_τ^γ is a generalized fractional derivative of order $\gamma = \frac{\alpha+1}{2}$ (cf. definition in Table I). Using our present notation, we can write the equivalent frequency domain criterion

$$\xi'(y,s) = \int_0^{2\pi} \frac{\left|Y(e^{j\omega}) - \sum_{n\in\mathbb{Z}} \hat{s}(\omega+2\pi n)\right|^2}{\Lambda(e^{j\omega})} \frac{\mathrm{d}\omega}{2\pi} + \int_{-\infty}^{+\infty} |\omega|^{\alpha+1} |\hat{s}(\omega)|^2 \frac{\mathrm{d}\omega}{2\pi},$$

whose minimum over all functions $s(t) \in L_2$ is achieved by the spline estimator (4).

Since the digital filter h is guaranteed to be BIBO stable (cf. [1, Theorem 2]) and the fractional B-spline basis functions are L_p-stable for $1 \leq p \leq \infty$, there is no compelling reason for restricting ourselves to input signals that are in ℓ_2, other than the fact that the variational criteria listed above may no longer be well defined. In the sequel, we will lift this hypothesis and consider the application of the spline estimator (4) to stochastic signals that are not square summable. It is important to note that this does not change anything from a computational point of view, meaning that the fast smoothing spline algorithm introduced in [1, Section IV] and the corresponding frequency domain analysis remain valid.

III. Stochastic modeling of self-similar processes

To justify the use of the above smoothing spline estimator on statistical grounds, we first need to introduce an appropriate mathematical framework that allows us to characterize fractal-like processes and to apply linear operators to them, including the fractional derivatives listed in Table I. Since this cannot be handled by standard stochastic calculus, we had to turn to Gel'fand and Vilenkin's theory of generalized stochastic processes which constitutes the statistical counterpart of Schwartz's theory of distributions.

A. Self-similar processes: review of standard results

There are a number of technical difficulties with the modeling of self-similar processes, fractional Brownian-motion (fBm) being the most prominent example. This is primarily due to the fact that these processes are non-stationary, meaning that their spectral power density cannot be defined in the conventional sense. Fortunately, the τ-lag increment derived process, $y_\tau(t) = x(t) - x(t-\tau)$, where $x(t)$ denotes the realization of a fBm-like process, is zero-mean, second-order stationary.

In the statistical literature, a process whose increments are stationary is referred to as being intrinsic stationary. These processes are often characterized by their variogram [17],

$$\gamma_x(\tau) := \mathscr{E}\{(x(t) - x(t-\tau))^2\},$$

which measures the variance of the increment process $y_\tau(t)$ as a function of the time-lag.

TABLE I

PRIMARY FUNCTIONS AND OPERATORS FOR SMOOTHING SPLINE ESTIMATION AND THE MODELING OF FRACTAL PROCESSES

	Notation	Defining formula
Spline generating functions:		
Symmetric B-spline of degree α	$\beta_0^\alpha(t)$	$\hat{\beta}_0^\alpha(\omega) = \left\vert \frac{\sin(\omega/2)}{\omega/2} \right\vert^{\alpha+1}$
Generalized fractional B-spline	$\beta_\tau^\alpha(t)$	$\hat{\beta}_\tau^\alpha(\omega) = \left(\frac{1-e^{-j\omega}}{j\omega}\right)^{\frac{\alpha+1}{2}+\tau} \left(\frac{1-e^{j\omega}}{-j\omega}\right)^{\frac{\alpha+1}{2}-\tau}$
Test functions of rapid decay	$\phi(t) \in \mathscr{S}$	$\vert\phi^{(n)}(t)\vert \leq \frac{C_{m,n}}{1+\vert t\vert^m}, \quad \forall, m, n \in \mathbb{N}$
	$\hat{\phi}(\omega) \in \mathscr{S}$	$\hat{\phi}(\omega) = \int_{-\infty}^{+\infty} \phi(t) e^{-j\omega t} \mathrm{d}t$ with $\phi \in \mathscr{S}$
Scale-invariant operators:		
Fractional derivative	$\partial_\tau^\gamma \phi(t)$	$\widehat{\partial_\tau^\gamma \phi}(\omega) = (-j\omega)^{\frac{\gamma}{2}-\tau} (j\omega)^{\frac{\gamma}{2}+\tau} \hat{\phi}(\omega)$
Fractional finite differences	$\Delta_\tau^\gamma \phi(t)$	$\widehat{\Delta_\tau^\gamma \phi}(\omega) = (1-e^{j\omega})^{\frac{\gamma}{2}-\tau} (1-e^{-j\omega})^{\frac{\gamma}{2}+\tau} \hat{\phi}(\omega)$
Left fractional anti-derivative	$__\partial_\tau^\gamma \phi(t)$	$\widehat{__\partial_\tau^\gamma \phi}(\omega) = \frac{\hat{\phi}(\omega) - \sum_{k=0}^{\lceil \gamma - \frac{3}{2} \rceil} \hat{\phi}^{(k)}(0) \frac{\omega^k}{k!}}{(-j\omega)^{\frac{\gamma}{2}-\tau} (j\omega)^{\frac{\gamma}{2}+\tau}}$
Right fractional anti-derivative	$\partial_{__\tau}^\gamma \phi(t)$	$\partial_{__\tau}^\gamma \phi(t) = \int_{-\infty}^{-\infty} \frac{e^{j\omega t} - \sum_{k=0}^{\lceil \gamma - \frac{3}{2} \rceil} \frac{(j\omega t)^k}{k!}}{(-j\omega)^{\frac{\gamma}{2}-\tau} (j\omega)^{\frac{\gamma}{2}+\tau}} \hat{\phi}(\omega) \frac{\mathrm{d}\omega}{2\pi}$

A real valued stochastic process $X = \{x(t)\}_{t\in\mathbb{R}}$ is *self-similar* with index (or Hurst exponent) $H > 0$ if, for any $a > 0$,

$$\{x(at)\}_{t\in\mathbb{R}} \triangleq \{a^H x(t)\}_{t\in\mathbb{R}}$$

where $\triangleq$ denotes the equality of all underlying finite-dimensional distributions.

The fractional Brownian motion (fBm) with Hurst exponent $0 < H < 1$ is a zero-mean Gaussian process that is both self-similar and intrinsically stationary [7]. In particular, this implies that its variogram

is self-similar of order $2H$ (in the sense that $\gamma_x(a\tau) = a^{2H}\gamma_x(\tau)$) and is therefore given by

$$\gamma_x(\tau) = C_H \cdot |\tau|^{2H} \tag{6}$$

where $C_H > 0$ is simply a scaling factor (cf. [1, proof of Proposition 1]). The variance of the fBm has the same self-similar functional form

$$\sigma_x^2(t) = C_H \cdot |t|^{2H}, \tag{7}$$

which is time-dependent, confirming that the process is non-stationary.

If we know both the time-varying variance and the variogram of the process, we easily obtain the autocorrelation function, thanks to the relation

$$c_{xx}(t_1, t_2) := \mathscr{E}\{x(t_1)x(t_2)\} = \frac{1}{2}\left(\mathscr{E}\{x^2(t_1)\} + \mathscr{E}\{x^2(t_2)\} - \gamma_x(\tau)\right).$$

In the case of the fBm process, this yields the following explicit form of the autocorrelation,

$$c_{xx}(t_1, t_2) = \frac{C_H}{2}\left(|t_1|^{2H} + |t_2|^{2H} - |t_2 - t_1|^{2H}\right), \tag{8}$$

which is also self-similar of order $2H$. Conversely, it can be shown that the 2D function defined by (8) is non-negative definite (see Thm. 1 below) and that it is the only possible parametric form of correlation that corresponds to a process that is both self-similar and intrinsically stationary [18].

The notion of intrinsic stationarity can be further generalized by considering processes whose nth-order increments are stationary [12], [19]. Among those, one can also identify the ones that are self-similar, which leads to an extended notion of fractional Brownian motion for larger Hurst exponents H such that $n - 1 < H < n$ where n is the order of the increment. Such processes can be obtained, for example, from the n-fold integration of a conventional fBm with suitable initial conditions [20]. This leads to an autocorrelation function of the form

$$c_{xx}(t_1, t_2) = \frac{C_H}{2} \times \left(-|t_2 - t_1|^{2H} + \sum_{k=0}^{n-1}(-1)^k \binom{2H}{k}\left[\left(\frac{t_1}{t_2}\right)^k |t_2|^{2H} + \left(\frac{t_2}{t_1}\right)^k |t_1|^{2H}\right]\right), \tag{9}$$

where C_H is a constant and $n = \lceil H \rceil$. In the literature, the constant is sometimes expressed as $C_H = \frac{\epsilon_H^2}{\Gamma(2H+1)\sin(\pi H)}$ where ϵ_H^2 is a spectral energy factor; the standard normalized case corresponds to the choice $\epsilon_H^2 = 1$ and $0 < H < 1$.

B. Generalized stochastic processes: Gel'fand-Vilenkin's approach [21]

In order to perform linear operations such as differentiation on random processes, a fruitful approach is to consider them as random distributions and to extend the applicability of the distributional calculus

to these processes. In particular, this formalism provides a rigorous definition of white noise, which plays such as fundamental role in statistical signal processing. A very stimulating and fundamental presentation of this theory can be found in [21].

Although Gel'fand and Vilenkin's approach is a natural extension of the now classical theory of distributions [22], it seems to have been somewhat neglected in the standard literature on random processes, including fBms. One notable exception relating to signal processing is [23]. By contrast, the Itô stochastic calculus and its Stratonovich variant have received a much greater attention; in particular, in statistical physics and financial mathematics [24], [25]. Both types of approaches have their advantages and limitations; for instance, the Itô calculus can handle certain nonlinear operations on random processes that cannot be dealt with the distributional approach. Gel'fand's theory, on the other hand, is ideally suited for performing any kind of linear operations including some, such as fractional derivatives, which are extremely cumbersome to define in a traditional (non-distributional) framework.

Most readers may recall that a distribution $u(t)$ is not defined through its point values (samples), but rather through a series of scalar products (linear functionals) $\langle u, \phi \rangle$ with all test functions $\phi \in \mathscr{S}$ (Schwartz's class). These test functions are indefinitely differentiable and they, as well as all their derivatives, have very rapid decay (i.e., faster than $O(|t|^{-m}), \forall m \in \mathbb{N}^+$). In an analogous fashion, a generalized stochastic process $x(t)$ is not defined by the probability law of its pointwise samples $\{\ldots, x(t_1), x(t_2), \ldots, x(t_N), \ldots\}$, but by the probability law of its scalar products with arbitrary test functions $\phi(t) \in \mathscr{S}$.

Specifically, given $\phi \in \mathscr{S}$, $y = \langle x, \phi \rangle$ is a random variable characterized by a probability density $p_\phi(y)\,\mathrm{d}y$. The characteristic function of this random variable is used to define a functional of ϕ:

$$Z_x(\phi) = \mathscr{E}\{e^{-j\langle x,\phi\rangle}\} = \int e^{-jy} p_\phi(y)\,\mathrm{d}y,$$

where $\mathscr{E}\{\cdot\}$ is the expectation operator. This functional Z_x is called the characteristic form of the process x. It is important to understand that it concentrates all the information available on the generalized stochatic process $x(t)$. For instance, if one wants to access the joint probability of the random variables $y_1 = \langle x, \phi_1 \rangle$, $y_2 = \langle x, \phi_2 \rangle$,... $y_N = \langle x, \phi_N \rangle$, then it suffices to take the inverse Fourier transform of $Z_x(\omega_1\phi_1 + \omega_2\phi_2 + \ldots \omega_N\phi_N)$ with respect to $\omega = \{\omega_1, \omega_2, \ldots \omega_N\}$. This is because the probability density of the set of random variables $\{y_1, y_2, \ldots y_N\}$ is given by

$$p(y_1, y_2, \ldots y_N) = \int \underbrace{\mathscr{E}\Big\{e^{-\sum_{k=1}^{N} j\omega_k\langle x,\phi_k\rangle}\Big\}}_{Z_x\left(\sum_{k=1}^{N}\omega_k\phi_k\right)} e^{\sum_{k=1}^{N} j\omega_k y_k} \frac{\mathrm{d}^N\omega}{(2\pi)^N}$$

Conversely, if $Z(\phi)$ is a continuous form of positive type[2] and satisfies $Z(0) = 1$, then it is the characteristic form of a generalized stochastic process $x(t)$. The continuity of the functional expresses the fact that $Z(\phi_k)$ tends to $Z(\phi)$ when ϕ_k tends to ϕ as $k \to \infty$; it is an essential ingredient that allows the extension of the characteristic form Z_x to potentially larger function spaces than $\mathscr{S}$. For instance, if Z_x is continuous with respect to the L_2-norm then, thanks to the density of $\mathscr{S}$ in L_2, we may extend the functional Z_x to arbitrary functions of L_2. For continuous processes such as fBm, we can even let ϕ_k tend to δ, the Dirac distribution: the continuity property of the characteristic form will ensure the well-definiteness of $Z_x(\delta)$.

Conceptually, this means that the characteristic form can be viewed as the distributional, infinite-dimensional extension of the classical characteristic function. To get a better feeling for this connection, we note that characteristic function (i.e., the Fourier transform of the probability density) of the sample $x(0)$ of a stochastic process $x(t)$ is defined by $\mathscr{E}\{e^{-j\omega x}\}$ which can also be expressed as $Z_x(\omega\delta)$; this corresponds to a one-dimensional analysis of the process with the test functions $\phi(t) = \omega\delta(t)$ parametrized by $\omega \in \mathbb{R}$. The argument obviously also holds in higher dimensions by considering the N-dimensional subspace of test functions $\omega_1\delta(t-t_1) + \cdots \omega_N\delta(t-t_N)$ with $(\omega_1, \cdots, \omega_N) \in \mathbb{R}^N$.

The advantage of working with scalar products instead of point values is that it is possible to exploit duality properties to perform linear operations such as differentiation, Fourier transforms or convolutions. For instance, using the definitions $\langle \partial_\tau^\gamma x, \phi\rangle = \langle x, \partial_{-\tau}^\gamma \phi\rangle$, we are able to compute the fractional derivative of a stochastic process. This can be moved automatically to the characteristic form:

$$Z_{\partial_\tau^\gamma x}(\phi) = Z_x(\partial_{-\tau}^\gamma \phi) \tag{10}$$

More generally, if h is some filter, the characteristic form of $h * x$ is given by $Z_{h*x}(\phi) = Z_x(h^\dagger * \phi)$ where $h^\dagger(x) = h(-x)$.

Similarly, using the definition $\langle \hat{x}, \phi\rangle = \langle x, \hat{\phi}\rangle$, the characteristic form of the Fourier transform of a stochastic process is given by:

$$Z_{\hat{x}}(\phi) = Z_x(\hat{\phi}). \tag{11}$$

The case of generalized, zero-mean Gaussian processes is especially simple to deal with since they are completely defined by their mean and autocorrelation. Specifically, if $c_{xx}(t,t')$ is the correlation form[3]

[2]i.e., if the matrix $\mathbf{Z} = [Z(\phi_k - \phi_l)]_{1\le k,l\le N}$ is positive irrespective of $N \in \mathbb{N}\setminus\{0\}$ and of the choice of $\phi_k \in \mathscr{S}$.

[3]defined by $\langle c_{xx}(t,t'), \phi(t)\psi(t')\rangle = \mathscr{E}\{\langle x,\phi\rangle\langle x,\psi\rangle\}$ for all $\phi, \psi \in \mathscr{S}$.

of the zero-mean process x, then its characteristic form is given by:

$$Z_x(\phi) = \exp\left\{-\frac{1}{2}\int c_{xx}(t,t')\phi(t)\phi(t')\,\mathrm{d}t\mathrm{d}t'\right\}. \tag{12}$$

Conversely, if c_{xx} is a continuous positive distribution of (t,t'), then (12) defines a generalized zero-mean Gaussian process. Stationary processes have a simpler characteristic form:

$$Z_x(\phi) = e^{-\frac{1}{2}\|h_x * \phi\|_{L_2}^2} \tag{13}$$

where $h_x(t)$ is a filter such that $|\hat{h}_x(\omega)|^2$ is the power spectral density of the process $x(t)$. Of particular interest is the case of the normalized Gaussian white noise $W(t)$ which is defined via its characteristic form

$$Z_W(\phi) = e^{-\frac{1}{2}\|\phi\|_{L_2}^2}.$$

Applying Parseval identity and the definition (11) of the Fourier transform of a stochastic process, we easily get the "intuitive" result that the Fourier transform of a white noise is a white noise as well.

However, if we lift the Gaussian hypothesis, other versions of white noise processes can be obtained such as the generalized Poisson process which has the following characteristic form:

$$Z_W(\phi) = \exp\Big(\lambda\int\big(e^{j\phi(t)}-1\big)\mathrm{d}t\Big).$$

In their book, Gel'fand and Vilenkin give the explicit expression of the characteristic form of an even broader class of non-Gaussian white noise-like processes [21].

C. Fractional integrals/anti-derivatives

One of the classical definition of fBm involves a fractional integral of a Wiener (or Brownian motion) process [7], [18]. It is therefore tempting to introduce an extended family of integral operators (fractional anti-derivatives) that are inverse operators for ∂_τ^γ where $\gamma \geq 0$. To this end, we have to make the distinction between left, and right inverse denoted respectively by ${}_-^{\gamma}\!\partial_\tau$ and $\partial_{-\tau}^{\gamma}$. When $\gamma \neq 0$, we propose the following operators

$$\begin{aligned} {}_-^{\gamma}\!\partial_\tau\phi(t) &= \frac{1}{2\pi}\int \frac{\hat{\phi}(\omega) - \sum_{k=0}^{\lceil\gamma-\frac{3}{2}\rceil}\hat{\phi}^{(k)}(0)\frac{\omega^k}{k!}}{(-j\omega)^{\frac{\gamma}{2}-\tau}(j\omega)^{\frac{\gamma}{2}+\tau}}e^{j\omega t}\,\mathrm{d}\omega, \\ \text{and} \\ \partial_{-\tau}^{\gamma}\phi(t) &= \frac{1}{2\pi}\int \frac{e^{j\omega t} - \sum_{k=0}^{\lceil\gamma-\frac{3}{2}\rceil}\frac{(j\omega t)^k}{k!}}{(-j\omega)^{\frac{\gamma}{2}-\tau}(j\omega)^{\frac{\gamma}{2}+\tau}}\hat{\phi}(\omega)\,\mathrm{d}\omega. \end{aligned} \tag{14}$$

When ϕ is regular and decreases fastly enough[4], $\overline{\ }_{-\tau}^{\gamma}\partial\phi$ defines a true function which is either in L_2 when γ is not a half-integer, or is continuous and slowly decreasing (but not in L_2) when γ is a half-integer. It is easy to verify that $\overline{\ }_{-\tau}^{\gamma}\partial$ satisfies $\overline{\ }_{-\tau}^{\gamma}\partial\partial_{\tau}^{\gamma}\phi = \phi$ for every function of $\mathscr{S}$; i.e., $\overline{\ }_{-\tau}^{\gamma}\partial\partial_{\tau}^{\gamma} = \text{Identity}$.

Moreover, it can be checked that $\overline{\ }_{-\tau}^{\gamma}\partial$ and $\partial_{\tau}^{\overline{\ }\gamma}$ are adjoint of one another (in a similar way as $\partial_{-\tau}^{\gamma}$ and ∂_{τ}^{γ} are adjoint); i.e., $\langle\partial_{-\tau}^{\overline{\ }\gamma}\psi,\phi\rangle = \langle\psi,\overline{\ }_{\tau}^{\gamma}\partial\phi\rangle$ when ϕ and ψ are in $\mathscr{S}$. Thanks to duality, we can thus claim that $\langle\partial_{\tau}^{\gamma}\partial_{-\tau}^{\overline{\ }\gamma}\psi,\phi\rangle = \langle\psi,\phi\rangle$; i.e., $\partial_{\tau}^{\gamma}\partial_{-\tau}^{\overline{\ }\gamma} = \text{Identity}$.

This allows to extend the right fractional derivative inverse $\partial_{-\tau}^{\overline{\ }\gamma}$ to a subset of tempered distributions u according to the rule

$$\langle\partial_{-\tau}^{\overline{\ }\gamma}u,\phi\rangle = \langle u,\overline{\ }_{\tau}^{\gamma}\partial\phi\rangle.$$

It is interesting to note that both types of anti-derivative operators are scale-invariant of order $-\gamma$. Intuitively, they may be thought of as (fractional) integrals to which one has imposed special boundary conditions at the origin. This has also the benefit of producing a result that is reasonably localized; i.e., square-integrable in the case when $\phi \in \mathscr{S}$. The left anti-derivative operator, for instance, has a special Dirac distribution annihilation property in the sense that $\overline{\ }_{-\tau}^{\gamma}\partial\delta^{(m)}(t) = 0$, for $m = 0,\cdots,n$, where $n = \lceil\gamma - \frac{3}{2}\rceil$. The right anti-derivative operator, on the other hand, will produce a function (or distribution) $v(t) = \partial_{-\tau}^{\overline{\ }\gamma}u(t)$ that has a $(n+1)$th order zero at the origin: $v^{(m)}(0) = 0$, for $m = 0,\cdots,n$. When we are dealing with a function, this can be achieved by correcting the usual integral with a suitable polynomial that is in the null space of ∂_{τ}^{γ}. In both cases, these are properties that are strictly tied to the origin $t = 0$, indicating that the operators are not shift-invariant.

Thanks to the above distributional relations, we can readily apply these anti-derivative operators to a wide class of generalized stochastic processes x; in particular, the Gaussian stationary ones specified by (26). For instance, by using the definition, $\langle\partial_{-\tau}^{\overline{\ }\gamma}x,\phi\rangle = \langle x,\overline{\ }_{\tau}^{\gamma}\partial\phi\rangle$, we can directly move the right anti-derivative to the characteristic form, which yields

$$Z_{\partial_{-\tau}^{\overline{\ }\gamma}x}(\phi) = Z_x(\overline{\ }_{\tau}^{\gamma}\partial\phi). \tag{15}$$

Of course, the restriction here is that the right hand side of (15) be well defined, which will typically be the case when $\overline{\ }_{\tau}^{\gamma}\partial\phi \in L_2$.

[4]For instance, when γ is not a half-integer, we may wish to ensure that the remainder of the Taylor development of $\hat{\phi}(\omega)$ near $\omega = 0$ be at least $O(|\omega|^{\gamma-1/2+\varepsilon})$ for some positive ε, and that $\int|\hat{\phi}(\omega)|\mathrm{d}\omega < \infty$, which is automatically satisfied when $\phi \in \mathscr{S}$.

D. Distributional characterization of fBm

We now present our first theoretical result on the characterization of fBm.

Theorem 1: The usual ($0 < H < 1$) fractional Brownian motion process $B_H(t)$ is characterized by the form

$$Z_{B_H}(\phi) = \exp\Big(-\frac{\varepsilon_H^2}{4\pi}\int \frac{|\hat{\phi}(\omega)-\hat{\phi}(0)|^2}{|\omega|^{2H+1}}\,\mathrm{d}\omega\Big), \tag{16}$$

where $\varepsilon_H^2 = \Gamma(2H+1)\sin(\pi H)C_H$, the constant C_H being defined according to (6) or (8).

Proof: We already know that a fBm is a Gaussian process with correlation $c_{xx}(t,t') = C_H/2\left(|t|^{2H}+|t'|^{2H}-|t-t'|^{2H}\right)$. We thus only have to prove that

$$\int c_{xx}(t,t')\phi(t)\phi(t')\,\mathrm{d}t\mathrm{d}t' = \varepsilon_H^2\int \frac{|\hat{\phi}(\omega)-\hat{\phi}(0)|^2}{|\omega|^{2H+1}}\,\frac{\mathrm{d}\omega}{2\pi}.$$

In order to do this, we choose $s>0$ and introduce the function $\lambda_s(t) = C_H/2\,|t|^{2H}e^{-s|t|}$ whose Fourier transform is $\hat{\lambda}_s(\omega) = C_H\Gamma(2H+1)\mathrm{Re}\left\{(s+j\omega)^{-2H-1}\right\}$. Obviously, $c_s(t,t') = \lambda_s(t)+\lambda_s(t')-\lambda_s(t-t')$ tends to $c_{xx}(t,t')$ when $s\to 0$. More precisely, Lebesgue dominated[5] convergence theorem ensures that $\int c_s(t,t')\phi(t)\phi(t')\,\mathrm{d}t\mathrm{d}t'$ tends to $\int c_{xx}(t,t')\phi(t)\phi(t')\,\mathrm{d}t\mathrm{d}t'$ when $s\to 0$.

Then, we observe that

$$\int c_s(t,t')\phi(t)\phi(t')\,\mathrm{d}t\mathrm{d}t' = -\int |\hat{\phi}(\omega)-\hat{\phi}(0)|^2\hat{\lambda}_s(\omega)\,\frac{\mathrm{d}\omega}{2\pi}$$

because of the following Fourier equivalences:

- $\displaystyle\int \lambda_s(t)\phi(t)\phi(t')\,\mathrm{d}t\mathrm{d}t' = \hat{\phi}(0)\int \hat{\lambda}_s(\omega)\hat{\phi}(\omega)^*\,\frac{\mathrm{d}\omega}{2\pi} = \hat{\phi}(0)^*\int \hat{\lambda}_s(\omega)\hat{\phi}(\omega)\,\frac{\mathrm{d}\omega}{2\pi}$
- $\displaystyle\int \lambda_s(t-t')\phi(t)\phi(t')\,\mathrm{d}t\mathrm{d}t' = \int \hat{\lambda}_s(\omega)|\hat{\phi}(\omega)|^2\,\frac{\mathrm{d}\omega}{2\pi}$
- $\displaystyle|\hat{\phi}(0)|^2\int \hat{\lambda}_s(\omega)\,\frac{\mathrm{d}\omega}{2\pi} = |\hat{\phi}(0)|^2\lambda_s(0) = 0$

Let us denote by $v(\omega)$ the integrable function $-C_H\Gamma(2H+1)\sin(\pi H)|\hat{\phi}(\omega)-\hat{\phi}(0)|^2|\omega|^{-2H-1}$. We have the limit result $\lim_{s\to 0}|\hat{\phi}(\omega)-\hat{\phi}(0)|^2\hat{\lambda}_s(\omega) = v(\omega)$. Moreover, $|\hat{\phi}(\omega)-\hat{\phi}(0)|^2|\hat{\lambda}_s(\omega)|$ is dominated, up to a constant, by $|v(\omega)|$. Finally, using Lebesgue dominated convergence theorem, we obtain the claimed Fourier expression

$$\lim_{s\to 0} -\int |\hat{\phi}(\omega)-\hat{\phi}(0)|^2\hat{\lambda}_s(\omega)\,\mathrm{d}\omega = C_H\Gamma(2H+1)\sin(\pi H)\int \frac{|\hat{\phi}(\omega)-\hat{\phi}(0)|^2}{|\omega|^{2H+1}}\,\frac{\mathrm{d}\omega}{2\pi}.$$

■

Note that this proof is also a direct method for showing that the expression (8) effectively defines a correlation (i.e., a positive quadratic form).

[5]Notice that $|c_s(t,t')\phi(t)\phi(t')|$ is "dominated" by $(|t|^{2H}+|t'|^{2H}+|t-t'|^{2H})|\phi(t)\phi(t')|$ which is integrable.

It is also possible to extend the expression (16) by making use of the left fractional anti-derivative ${}_{\tau}\partial^{-\frac{1}{2}-H}\phi$ of $\phi(t) \in \mathscr{S}$, defined in subsection III.C. Note that $\tau \in \mathbb{R}$ is an arbitrary free parameter. Thanks to this notation, we rewritte the characteristic form of the fBm as

$$Z_{B_H}(\phi) = e^{-\frac{\varepsilon_H^2}{2}\left\|{}_{\tau}\partial^{-\frac{1}{2}-H}\phi\right\|_{L_2}^2}. \tag{17}$$

The advantage of this formula is that it also yields a natural extension of the fBm for non-integer Hurst exponents $H > 1$. Note that positive integer values of H are excluded because they correspond to anti-derivatives ${}_{\tau}\partial^{-\frac{1}{2}-H}\phi$ that are not necessarily square integrable. Using the same technique as in Thm. 1, it is then possible to compute the autocorrelation of the process defined by (17), resulting in a form that is identical to (9). This also yields a direct relation between the amplitude factor ε_H in (17) and the constant C_H in (9):

$$C_H = \frac{\epsilon_H^2}{\Gamma(2H+1)\sin(\pi H)} \tag{18}$$

This shows that the generalized fBm that is concisely defined by (17) with $H \in \mathbb{R}^+\backslash\mathbb{N}$ is in fact equivalent to the one introduced by Perrin et al. with the help of more traditional techniques [20].

By setting $\varepsilon_H = 1$ in the definition (17), we see that, for $H > 1$, we have $Z_{B_H}(\phi') = Z_{B_{H-1}}(\phi)$. This results in $B_H'(t) = -B_{H-1}(t)$; i.e., the usual derivative of an extended fBm of exponent H is an extended fBm of exponent $H-1$. More generally, we can show that an extended fBm B_H with noninteger Hurst exponent H is $\lfloor H \rfloor$-times continuously differentiable and that $\frac{\mathrm{d}^{\lfloor H \rfloor}}{\mathrm{d}t^{\lfloor H \rfloor}} B_H$ is a usual fractional Brownian motion with Hurst exponent $[H] = H - \lfloor H \rfloor$. In fact, by substituting $\phi(t) = \delta(t)$ in (17), we observe that the Fourier transform of the probability density of $B_H(0) = \langle B_H, \delta \rangle$ equals 1 which means that $B_H(0) = 0$ with probability one. Likewise, we can show that an extended fBm with exponent H is $\lfloor H \rfloor$-times continuously differentiable and that all its derivatives vanish at $t = 0$: $\frac{\mathrm{d}^n}{\mathrm{d}t^n} B_H(0) = 0$ for $n = 0 \ldots \lfloor H \rfloor$.

E. Whitening properties of the fractional derivatives

We now wish to reinterpret the formula (17) that defines both the usual fBm ($0 < H < 1$) and the "extended" one ($H \in \mathbb{R}_+ \setminus \mathbb{N}$). By using the characteristic form $Z_W(\phi) = \exp(-\frac{1}{2}\|\phi\|_{L_2}^2)$ of the normalized Gaussian white noise and the duality between right and left primitive (10), we identify the characteristic form of the fBm with the characteristic form of $\varepsilon_H \, \partial_{-\tau}^{-H-\frac{1}{2}} W$:

$$\begin{aligned} Z_{B_H}(\phi) &= e^{-\frac{\varepsilon_H^2}{2}\left\|{}_{\tau}\partial^{-\frac{1}{2}-H}\phi\right\|_{L_2}^2} \\ &= Z_W(\varepsilon_H \, {}_{\tau}\partial^{-\frac{1}{2}-H}\phi) \\ &= Z_{\varepsilon_H \, \partial_{-\tau}^{-H-\frac{1}{2}} W}(\phi), \end{aligned}$$

which proves the identity between the two processes.

Proposition 1: The extended fractional Brownian motion with arbitrary noninteger positive Hurst exponent H can be expressed as a right $(H+\frac{1}{2})$th anti-derivative of a Gaussian white noise:

$$B_H = \varepsilon_H \, \partial_{-\tau}^{-H-\frac{1}{2}} W. \tag{19}$$

As a corollary, we get that the $(H+\frac{1}{2})$th derivative of an fBm is a white noise:

$$\partial_{\tau}^{H+\frac{1}{2}} B_H = \varepsilon_H \, W, \tag{20}$$

which follows from the right-inverse property $\partial_{\tau}^{H+\frac{1}{2}} \partial_{-\tau}^{-H-\frac{1}{2}} = \text{Identity}$.
Note that, for $\tau = H/2+1/4$, the formula (20) is equivalent to the one that was proposed in [20, eq.(6)].

F. Fractional increment process

As mentioned in subsection III.A, the standard approach for dealing with intrinsically stationary processes is to consider their increment so that the problem reduces to the characterization of a stationary process. Alternatively, when the process is intrinsically stationary of order n, one can apply an nth order differentiator which yields a (generalized) derived process that is stationary as well [12], [19].

Here we propose another possibility that is specifically tailored to the characterization of γth order self-similar processes. It is stochastic transposition of the localization technique discussed in [1, Sections II and III]. Specifically, we chose to apply a discrete operator—the γth order fractional finite difference—that closely approximates the whitening operator of the process (i.e., ∂_{τ}^{γ}). This produces a derived process that is stationary, essentially decorrelated, and yet well-defined in the classical sense (i.e, mean-square continuous) for $\gamma > \frac{1}{2}$.

Proposition 2: Let $B_H(t)$ be an fBm with noninteger Hurst exponent $H > 0$. Then, the derived process $y(t) = \Delta_{\tau}^{H+1/2} B_H(t)$ is zero mean, stationary with covariance function

$$c_{yy}(t_1, t_2) = \varepsilon_H^2 \, \beta_0^{2H}(t_2 - t_1);$$

i.e., $y(t)$ can be expressed as the convolution of a normalized Gaussian white noise W with a fractional B-spline according to

$$\Delta_{\tau}^{H+1/2} B_H(t) = \varepsilon_H \, \beta_{\tau}^{H-\frac{1}{2}} * W.$$

Proof: Using the definition of Δ_{τ}^{γ} (cf. Table 1) and the fact that its adjoint is simply $\Delta_{-\tau}^{\gamma}$, we have that $\langle \Delta_{\tau}^{H+1/2} B_H, \phi \rangle = \langle B_H, \Delta_{-\tau}^{H+1/2} \phi \rangle$. Moreover, expressing B_H as an anti-derivative of white noise according to (20) and using the duality definition for $\partial_{-\tau}^{-H-\frac{1}{2}}$, we find that

$$\langle y, \phi \rangle = \langle W, \varepsilon_H \, {}^{-H-\frac{1}{2}}\partial_{\tau} \Delta_{-\tau}^{H+1/2} \phi \rangle.$$

It is now a simple matter to verify by applying the definition (14) that ${}^{-\gamma}_{\tau}\partial\Delta^{\gamma}_{-\tau}\phi = \beta^{\gamma-1}_{-\tau} * \phi$. The result then follows by noticing that $\{\beta^{\gamma-1}_{-\tau}\}^{\dagger} = \beta^{\gamma-1}_{\tau}$. ■

Practically, this means that we have at our disposal a digital filter that we can apply to $\{x(t)\}_{t\in\mathbb{R}}$ —or to its sampled values $\{x[k]\}_{t\in\mathbb{Z}}$—to produce an output signal (resp., output sequence) that is stationary, with a very short correlation distance and therefore much easier to handle mathematically. The whole concept is schematically represented in Fig. 1.

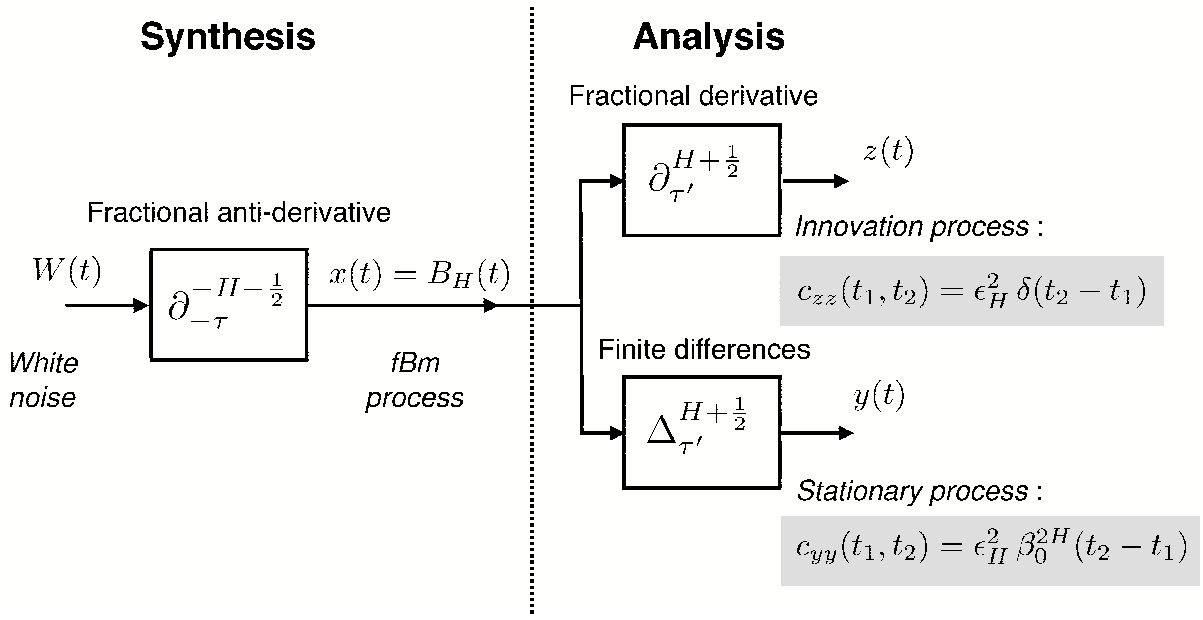

Fig. 1. Synthesis and analysis of a self-similar stochastic process: The fBm process $x(t)$ is generated by fractional integration of white Gaussian noise with suitable initial condition that guarantee that $E\{|x(0)|^2\} = 0$. The whitening operator of the process is a fractional derivative operator $\partial^{\gamma}_{\tau'}$ of corresponding order. Even though $x(t)$ is not necessarily stationary, it can be transformed into a stationary process $y(t)$ through the application of the fractional finite difference operator $\Delta^{\gamma}_{\tau'}$. Since $\Delta^{\gamma}_{\tau'}$ is a "good" discrete approximation of the continuous-time operator $\partial^{\gamma}_{\tau'}$, the correlation function $c_{yy}(\tau)$ is well-defined and concentrated around the origin. In fact, this function may be thought of as a regularized version of the Dirac impulse. Note that τ' can be chosen arbitrarily.

IV. Optimal estimation of fractal-like signals

We are now ready to investigate the problem of the MMSE estimation of fBm signals. To this end, we will first derive the posterior distribution of $x(t_0)$ at a fixed location $t = t_0$ given a series of noisy samples of a fBm with Hurst exponent $0 < H < 1$. In particular, we will establish that the posterior mean

is a fractional spline of degree $2H$ which justifies the use of spline estimators. The next step will be to specify a Wiener-like filtering algorithm that will perform the MMSE estimation of $x(t)$ simultaneously for all $t \in \mathbb{R}$. We will show that this can be achieved via an appropriate tuning of the smoothing spline algorithm described in Section II and that the solution is also applicable for generalized fBms with Hurst exponents greater than 1.

A. Posterior estimation of fBm

Let $x(t)$ be a fBm with Hurst exponent $H < 1$. We suppose that we are observing the signal indirectly through a series of noisy measurements at the integers: $y[k] = x(k) + n[k]$, where $n[k]$ is additive stationary noise that is independent from $x(t)$. The noise is zero-mean and is characterized by its second order statistics: $E\{n[k]n[l]\} = c_{nn}[k-l]$. Our goal is to construct the best estimator of $x(t_0)$ given the measurements Y. To get a complete handle on this problem, we derive the posterior distribution $p(x(t_0)|Y)$ which fully specifies the information about the signal that is contained in the measurements.

Theorem 2: Let $x(t) = B_H(t)$ be a realization of an fBm of noninteger Hurst exponent $1 > H > 0$. Then, the posterior probability density of $x(t_0)$ given the measurements $Y = \{y[k] = x(k) + n[k]\}_{k\in\mathbb{Z}}$, where $n[k]$ is a zero-mean Gaussian stationary noise independent of $x(t)$ with autocorrelation $c_{nn}[k] \in \ell_2$, is the Gaussian density

$$p(x(t_0)|Y) = \frac{1}{\sqrt{2\pi}\sigma(t_0)} \exp\left(-\frac{(x(t_0) - \mu(t_0|Y))^2}{2\sigma^2(t_0|Y)}\right)$$

with time-varying mean

$$\mu(t_0|Y) = \tilde{x}(t_0|Y) - \rho(t_0)\tilde{x}(0|Y), \tag{21}$$

where $\tilde{x}(t_0|Y) = \sum_{k\in\mathbb{Z}} y[k]\varphi_\Lambda(t_0 - k)$ is the fractional smoothing spline of degree $\alpha = 2H$ specified by (4), (5) and (2) in Section II with

$$\Lambda(e^{j\omega}) = \frac{C_{nn}(e^{j\omega})}{\varepsilon_H^2}. \tag{22}$$

The conditional variance is given by

$$\sigma^2(t_0|Y) = \sigma_0^2(t_0) - \rho^2(t_0)\sigma_0^2(0). \tag{23}$$

with

$$\begin{aligned} \sigma_0^2(t_0) &= \frac{C_H}{2}\sum_{l\in\mathbb{Z}} |t_0 - l|^{2H}\varphi_\Lambda(t_0 - l) \\ \rho(t_0) &= \frac{\sum_{l\in\mathbb{Z}} c_{nn}[l]\varphi_\Lambda(t_0 - l)}{\sum_{l\in\mathbb{Z}} c_{nn}[l]\varphi_\Lambda(l)}. \end{aligned} \tag{24}$$

where $\varphi_\Lambda(t_0)$ is the equivalent smoothing spline basis function defined by (5).

Our proof of this result, which uses the characteristic form of the fBm, is given in Appendix I.

As a direct application of Theorem 2, we get the MMSE estimator of the fBm process $x(t)$ which is simply the conditional mean $\mathscr{E}\{x(t)|Y\}$ [3]. The key point for our purpose is that this estimator is a fractional spline, albeit not exactly the smoothing spline solution (4) that we may have wished for initially.

Corollary 1: Consider the noisy samples Y of an fBm process $x(t) = B_H(t)$, as specified in Theorem 2. Then, the MMSE estimator of $x(t)$ given Y is the function $\mu(t|Y)$ defined in Theorem 2, which is a fractional spline of degree $2H$. The corresponding minimum estimation error at location t is $\mathscr{E}\{|x(t) - \mu(t)|^2|Y\} = \sigma^2(t|Y)$ as specified in (23).

The above results call for the following comments.

1) The optimal fBm estimator $\mu(t_0|Y)$ is the sum of two terms that are both fractional splines of degree $2H$. The first component, $\tilde{x}(t_0|Y)$, is precisely the smoothing spline fit of Y, as specified in Section II, with the optimal choice of $\Lambda(e^{j\omega})$ given by (22). The second term, $-\rho(t_0)\tilde{x}(0|Y)$, is a non-stationary correction that ensures that the estimate is zero at $t = 0$, which is consistent with the property that $x(0) = B_H(0)$ is zero with probability one.
2) The variance of the estimator is made up of two terms as well. The first, $\sigma_0^2(t_0)$, is 1-periodic. The second is a correction that expresses the fact that the optimal spline estimate is more accurate near the origin because of the preference that is given to the value zero.
3) Interestingly, the correction function $\rho(t_0)$ is the fractional smoothing spline fit of the autocorrelation of the noise whose value at the origin has been normalized to one. Since $c_{nn}[k] \in \ell_2$ and takes its maximum for $k = 0$, one can expect this function to decay rapidly as one moves away from $t = 0$. For instance, when the noise is white, then $\rho(t_0) = \varphi_\lambda(t_0)/\varphi_\lambda(0)$, which typically decreases like $O(|t|^{-2H-2})$; that is, the rate of decay of the fractional B-splines (cf. [26, Theorem 3.1]).
4) When the measurement noise is zero, we have that $\tilde{x}(0|Y) = 0$ and $\sigma^2(0|Y) = 0$. Then, $\mu(t_0|Y) = \sum_{k\in\mathbb{Z}} y[k]\varphi_{\text{int}}(t-k)$ corresponds to the fractional spline interpolation of the input signal Y. This is equivalent to a smoothing spline estimate with $\Lambda = 0$. As expected, the estimation error $\sigma^2(t_0|Y)$ is zero at the integers because of the interpolation property of the underlying basis function: $\varphi_{\text{int}}(k) = \delta_k$.
5) The optimal estimator $\mu(t_0|Y)$ is undistinguishable from the smoothing spline solution $\tilde{x}(t|Y)$ in the following situations: (a) when the measurement noise is negligible, (b) when $\tilde{x}(0|Y) \approx 0$ which may happen because of (a) or simply by chance, and (c) in all cases for t_0 sufficiently large; that is, as one moves away from the origin which has a very special status due to the self-similarity of

the underlying process.

6) An efficient way to implement the estimator is to apply the Fourier domain algorithm that was presented in [1]. The procedure is to first run the algorithm on $\{y[k]\}$ to obtain the smoothing spline $\tilde{x}(t|Y)$. Second, one applies the same algorithm to the autocorrelation sequence of the noise $\{c_{nn}[k]\}$. One then corrects $\tilde{x}(t|Y)$ by subtracting a reweighted version of the latter spline so as to produce a result that is zero at the origin.
7) The Hurst exponent of the standard Wiener (or Brownian motion) process is $H = \frac{1}{2}$. This corresponds to a simple piecewise linear spline estimator with $\alpha = 1$. In the noise-free case, we recover a classical result by Lévy [11] that states that the optimal estimator of a Brownian motion process is obtained by linearly interpolating the samples. In that case, the estimate is entirely determined by the two neighboring samples. This is not so for other values of H (or when $c_{nn}[k] \neq 0$) because the smoothing spline filter generally has an infinite impulse response which induces coupling.

If one excludes the simpler case of Brownian motion, the present results on the prediction of fractal processes are new to the best of our knowledge. In principle, they should also be generalizable for fBms with $H > 1$, but we expect the formulas to become more complicated. In the sequel, we will investigate the general case as well, but adopt a less frontal approach by searching for the optimal estimator within the slightly more restrictive class of Wiener-filter-like (or stationary) solutions.

B. Wiener filtering of fractal processes

The main point that differentiates the optimal fBm estimator given by (21) from that of a stationary process (cf. [4, Theorem]) is the correction term $-\rho(t_0)\tilde{x}(0|Y)$, which makes the estimator vanish at the origin. Its presence is a consequence of the fBm being non-stationary. This lack of stationarity is obviously a source of complication; it requires the use of an advanced mathematical formulation and makes the task of finding the best estimator much more difficult.

While we have at our disposal a general closed form solution for $H < 1$, it may be justifiable in practice to discard the second, non-stationary part of the MMSE estimator for the following reasons:

1) The exact location of the origin ($t = 0$ in our model) will rarely be known (or controllable) in practice; especially if we are dealing with time series.
2) The Wiener-like estimators that are available for stationary processes have the advantage of computational simplicity. The same can be said for the smoothing spline part of $\mu(t_0|Y)$, which can be implemented by digital filtering of the input data.

3) The estimation error in Corollary 1 does not behave like the estimation error of a stationary process, which is necessarily periodic (cf. Appendix II). Instead, one may wish the wholeness of the nonstationary behavior of the fBm to be captured by the linear estimator, whilst the estimation error would be indistinguishable from the estimation error of a stationary process.
4) There is an alternative "kriging" formulation from the field of geostatistics that yields the best linear unbiased (BLU) estimator of $x(t_0)$ from sampled data (typically, non-uniform and multi-dimensional) [17], [27]. This non-parametric estimator is computed from the variogram (which does not include absolute positional information). Under suitable conditions, this estimator is also known to be the solution of a variational spline problem [28], [29].

Because of these considerations, it makes good sense to search for a suboptimal solution that has the simplicity of a stationary (or Wiener) estimator. In particular, we want to find out whether or not we can improve upon the smoothing spline estimator $\tilde{x}(t|Y)$. The key requirement for such a formulation is expressed by Remark 3) which implies a restriction on the possible linear estimators, as described by the following proposition.

Proposition 3: Let $x(t)$ be an fBm of arbitrary (noninteger) Hurst order H and $y[k] = x(k) + n[k]$ its noisy samples, where $n[k]$ is a zero-mean Gaussian stationary discrete process that is independent of $x(t)$. We build the linear estimator $\tilde{x}(t)$ of $x(t)$:

$$\tilde{x}(t) = \sum_{k\in\mathbb{Z}} y(k)\varphi_k(t). \tag{25}$$

In order for the estimation error $e(t) = x(t) - \tilde{x}(t)$ to behave like the estimation error of a stationary process, it is necessary and sufficient that

$$\sum_{k\in\mathbb{Z}} k^s \varphi_k(t) = t^s, \quad \text{for } s = 0 \dots \lfloor H \rfloor, \tag{26}$$

in the sense of distributions.

The proof is given in Appendix II.

The interesting aspect of this result is that it allows us to evacuate the difficulties associated with the non-stationary character of fractal processes. It is also suggests that the root of the problem lies within the (random) polynomial part of the signal which is included in the null space of the whitening operator $\partial_\tau^{H+\frac{1}{2}}$. In fact, the presence of this null space component is somewhat artificial for it is only here to ensure that the fBm has the correct boundary conditions at the origin. Thus, a possible interpretation of Proposition 3 is that the fBm can be made stationary through the removal of its polynomial trend which is entirely captured by any estimator satisfying (26). Incidentally, this polynomial reproduction property

also plays a fundamental role in wavelet theory [30], [31]. Specifically, when we perform a wavelet analysis of order $N \geq \lfloor H \rfloor$ of a fBm, the polynomial component of the process is entirely projected onto the coarser scale approximation with the consequence that the discrete wavelet coefficients end up being stationary within any given scale. This is the fundamental reason why wavelets act as approximate whitening operators for fractal-like processes [32], [33].

For $H < 1$, we also note that the "stationarizing" hypothesis is in fact equivalent to the unbiasedness constraint that is used for deriving unbiased kriging estimators [17], [27]. In that particular framework, the random process to estimate is expressed as the sum of an unknown constant (the trend) and a random process of known variogram.

Under the "stationarizing" hypothesis (26), we are able to provide the best linear estimator of a fBm. Not so surprisingly, when $0 < H < 1$, this estimator is precisely the first term of (21) that was given in Thm. 2, namely $\tilde{x}(t|Y)$. Note that the result below is valid for values of H larger than 1 as well.

Theorem 3: Let $x(t)$ be an fBm of arbitrary (noninteger) Hurst order H and $y[k] = x(k)+n[k]$ its noisy samples, where $n[k]$ is a zero-mean Gaussian stationary discrete process with autocorrelation $c_{nn}[k] \in \ell_2$, independent of $x(t)$. Then the least mean-square linear estimator (25) satisfying the "stationarizing" conditions (26) is given by

$$\tilde{x}(t) = \sum_{k \in \mathbb{Z}} y[k] \varphi_\Lambda(t-k)$$

where, as in Thm.2, φ_Λ is the smoothing spline of degree $2H$ defined by (5) and (2), with Λ chosen according to (22). The variance of the estimation error is given by

$$\mathscr{E}\{|x(t_0) - \tilde{x}(t_0)|^2\} = \frac{\varepsilon_H^2}{2\Gamma(2H+1)\sin(\pi H)} \sum_{k \in \mathbb{Z}} |t_0 - k|^{2H} \varphi_\Lambda(t_0 - k).$$

For a proof, see Appendix III.

This results closes the loop by showing that the optimal estimator of an fBm is indeed a smoothing spline with matching parameters. In particular, the spectral regularization $\Lambda(e^{j\omega})$ should be set proportional to the power spectrum of the noise.

For instance, when the measurement noise is white with variance σ_n^2, then $\Lambda(e^{j\omega}) = \sigma_n^2/\varepsilon_H^2 = \lambda$ is a scalar that is inversely proportional to the signal-to-noise ratio. This means that the smoothing effect of the spline gets stronger as the power of the noise increases, which is consistent with our expectation. Under those circumstances, the effect of the smoothing spline is analogous to that of a Butterworth filter of order $2H+1$ with a cutoff frequency $\omega_0 = \lambda^{-2H-1}$.

Theorem 3 also gives an explicit expression for the estimation error. By considering the expression for C_H given by (18), we observe that that $\mathscr{E}\{|x(t_0) - \tilde{x}(t_0)|^2\}$ coincides with the primary variance component $\sigma_0^2(t_0)$ in Thm. 2 when $0 < H < 1$. Note that this error is a symmetric 1-periodic function that is minimal at the integers. We expect it to take its maximum at the half integers because these points are the furthest away from the sampling locations (integers). We also expect the variance to increase and to flatten out as the smoothing gets stronger; that is, when $C_{nn}(e^{j\omega})$ is large relative to ε_H^2. In that respect, the second variance term in (23), $\rho^2(t_0)\sigma_0^2(0)$, quantifies the loss of performance of the stationary estimator in Theorem 3 over the optimal one specified in Corollary 1. Here too, in concordance with what has been said earlier in subsection IV-A, this variance bias can be expected to decrease rapidly as one moves away from the origin. This provides some solid, quantitative justification for using the stationary solution (smoothing spline) as a substitute for the optimal one.

Among the proposed solutions, one can single out the smoothing splines of odd degrees $n = 1, 3, \cdots$, which correspond to the optimal solution for the estimation Brownian motion ($n = 1$ (linear splines) for $H = \frac{1}{2}$) and its generalized counterparts (in particular, $n = 3$ (cubic splines) for $H = 1.5$). It is noteworthy that the basic versions of these estimators have a fast recursive implementation [34]. One surprising finding of our investigation is that there is no fractal interpretation for the fractional smoothing splines of even degree (i.e., $H \in \mathbb{N}$), whose building blocks are not piecewise polynomials, but rather "radial basis functions" of the type $|t|^{2n} \log |t|$ [26]. Thus, an open question is whether or not there exist a class of (non-self-similar) processes corresponding to these basis functions, or equivalently, a positive definite form $c_{xx}(t_1, t_2)$ (similar to (9)) that is made up of such building blocks.

Finally, we note that the smoothing spline solution in Theorem 3 for $0 < H < 1$ is equivalent to the BLU estimator that could have been derived using the kriging formalism with compatible hypotheses (cf. [35] for a treatment of the multidimensional case). The key point, however, is that the present cardinal spline framework also yields a fast algorithm that is generally not available for kriging. Since kriging is originally designed for dealing with non-uniform data, the standard approach is to restrict the data to a given number of neighbours of t_0 (which is suboptimal) and to recompute the estimator by solving the normal equations for each position t_0. Clearly, the advantage of using splines is that the optimal solution $\tilde{x}(t|Y)$ can be computed at once for $t \in \mathbb{R}$ using all available data. This is computationally much more efficient, but only possible because we are taking advantage of the shift-invariant structure provided by the uniform grid.

V. CONCLUSION

In this pair of papers, we have established a formal connection between deterministic splines and stochastic fractal processes (fBms). The fundamental, unifying relation that appears in both contexts is the differential equation

$$\partial_\tau^\gamma\{x(t)\} = w(t) \tag{27}$$

that defines a self-similar system with continuous-time input $w(t)$ and output $x(t)$. This characterization is complete in the sense that the fractional derivative, ∂_τ^γ, which is indexed by the order γ and an asymmetry parameter τ, spans the whole class of differential operators that are both shift- and scale-invariant.

We can formally construct all the varieties of deterministic fractional splines identified in [1] by exciting the system with a weighted stream of Dirac impulses $w(t) = \sum_{k\in\mathbb{Z}} a[k]\delta(t-k)$, where the $a[k]$'s are some suitable coefficients. Likewise, we have shown that we could generate all brands of fBm's, including the generalized ones with Hurst exponent $H = \gamma - \frac{1}{2} > 1$, by driving the system with white Gaussian noise.

While the above differential description is appealing conceptually, we must be aware that it is fraught with technical difficulties. Indeed, the price to pay for self-similarity is that the system (27) is unstable with a pole of order γ at $\omega = 0$. In particular, this means that the Green function of ∂_τ^γ—the fundamental building block of the fractional splines—is not in L_1. Likewise, the generalized stochastic processes that are generated by (27) are non-stationary. This calls for a special mathematical (distributional) treatment and also requires the specification of boundary conditions. Another non-trivial aspect is that the fractional operators are typically non-local, in contrast with the classical integer-order derivatives.

The connection also extends nicely beyond the generation process. For instance, we have shown that the MMSE estimator $x(t)$ of a fBm with Hurst exponent H given its—possibly, noisy—samples at the integers lies in a fractional spline space of degree $2H$. We derived the corresponding smoothing spline estimator that can be assimilated to a hybrid version of the Wiener filter for which the input is discrete and the output analog. In particular, we are able to specify the Wiener filter for the Wiener process (with $H = \frac{1}{2}$) bringing together two seemingly incompatible aspects of the research of Norbert Wiener. In that particular case, the optimal estimator happens to be a piecewise-linear smoothing spline estimator that has a fast recursive implementation. As for the estimators for arbitrary values of H, we have shown that they could all be implemented efficiently by means of an FFT-based algorithm.

ACKNOWLEDGMENTS

This work is funded in part by the grant 200020-101821 from the Swiss National Science Foundation.

APPENDIX I
PROOF OF THEOREM 2

We will proceed in three main steps: first, we derive the general expression of the (Gaussian) posterior distribution of $B_H(t_0)$ given a finite number of measurements of B_H; second, we calculate the expression of the posterior expectation of $B_H(t_0)$ given $B_H(k)$ for $k \in \mathbb{Z}$; finally, we evaluate the expression of the posterior standard deviation of $B_H(t_0)$ given $B_H(k)$.

a) Step 1—Posterior probability of $B_H(t_0)$: Taking a more general point of view, we first compute the posterior density probability of $\langle B_H, \psi_{t_0}\rangle$ given K measurements $\langle y, \psi_k\rangle$ where $k = -K, -K+1, \dots K$. We will then set $\psi_{t_0}(t) = \delta(t-t_0)$ and $\psi_k(t) = \delta(t-k)$ to get the desired result.

Using Bayes' rule, we have

$$p\big(\langle B_H,\psi_{t_0}\rangle \big| \langle y,\psi_1\rangle, \langle y,\psi_2\rangle, \dots \langle y,\psi_K\rangle\big) = \frac{p\big(\langle B_H,\psi_{t_0}\rangle, \langle y,\psi_1\rangle, \langle y,\psi_2\rangle, \dots \langle y,\psi_K\rangle\big)}{\displaystyle\int p\big(x, \langle y,\psi_1\rangle, \langle y,\psi_2\rangle, \dots \langle y,\psi_K\rangle\big)\mathrm{d}x}.$$

We thus have to compute the joint probability density of $\big\{\langle B_H,\psi_{t_0}\rangle, \langle y,\psi_1\rangle, \langle y,\psi_2\rangle, \dots \langle y,\psi_K\rangle\big\}$. This will be done through its Fourier transform; i.e., its characteristic function which is expressed using Z_{B_H} and Z_n, the characteristic forms of the processes $B_H(t)$ and $n(t)$:

$$\begin{aligned}
\zeta(\lambda, \lambda_{-K}, \dots \lambda_K) &= \mathscr{E}\Big\{ \exp\big(-j\lambda\langle B_H,\psi_{t_0}\rangle - j\textstyle\sum_k \lambda_k\langle y,\psi_k\rangle\big)\Big\} \\
&= \mathscr{E}\Big\{ \exp\big(-j\langle B_H, \lambda\psi + \textstyle\sum_k \lambda_k\psi_k\rangle - j\langle n, \sum_k \lambda_k\psi_k\rangle\big)\Big\} \\
&= \mathscr{E}\Big\{ \exp\big(-j\langle B_H, \lambda\psi + \textstyle\sum_k \lambda_k\psi_k\rangle\big)\Big\}\mathscr{E}\Big\{ \exp\big(-j\langle n, \sum_k \lambda_k\psi_k\rangle\big)\Big\} \\
&= Z_{B_H}\big(\lambda\psi + \textstyle\sum_k \lambda_k\psi_k\big) Z_n\big(\sum_k \lambda_k\psi_k\big) \\
&= \exp\Big(-\frac{\varepsilon_H^2}{2}\big\|^{-\frac{1}{2}-H}_{\ \ \tau}\partial\big\{\lambda\psi + \textstyle\sum_k \lambda_k\psi_k\big\}\big\|^2_{L_2} - \frac{1}{2}\big\|h_n * \sum_k \lambda_k\psi_k\big\|^2_{L_2}\Big) \\
&= \exp\Big(-\alpha\lambda^2 - 2\lambda\mathbf{U}^\mathsf{T}\mathbf{\Lambda} - \mathbf{\Lambda}^\mathsf{T}\mathbf{A}\mathbf{\Lambda} - \mathbf{\Lambda}^\mathsf{T}\mathbf{S}\mathbf{\Lambda}\Big)
\end{aligned}$$

where we have denoted

$$\begin{aligned}
\alpha &= \frac{\varepsilon_H^2}{2}\big\|^{-\frac{1}{2}-H}_{\ \ \tau}\partial\psi_{t_0}\big\|^2_{L_2} \\
\mathbf{U} &= \Big[\frac{\varepsilon_H^2}{2}\Big\langle {}^{-\frac{1}{2}-H}_{\ \ \tau}\partial\psi_{t_0}, {}^{-\frac{1}{2}-H}_{\ \ \tau}\partial\psi_k\Big\rangle\Big]_{k=-K,-K+1,\dots K} \\
\mathbf{A} &= \Big[\frac{\varepsilon_H^2}{2}\Big\langle {}^{-\frac{1}{2}-H}_{\ \ \tau}\partial\psi_k, {}^{-\frac{1}{2}-H}_{\ \ \tau}\partial\psi_l\Big\rangle\Big]_{k,l=-K,-K+1,\dots K} \\
\mathbf{S} &= \Big[\frac{1}{2}\underbrace{\langle h_n * \psi_k, h_n * \psi_l\rangle}_{=c_{nn}[k-l]}\Big]_{k,l=-K,-K+1,\dots K}
\end{aligned}$$

Performing the inverse Fourier transform of this expression, we obtain the posterior probability density of $x = \langle B_H, \psi_{t_0} \rangle$ given $\mathbf{Y} = [\langle y, \psi_k \rangle]_{k=-K,-K+1,\dots K}$:

$$p(x|\mathbf{Y}) = \text{Const} \times \exp\Big(-\frac{1}{4} \frac{(x - \mathbf{U}^\mathsf{T}(\mathbf{A}+\mathbf{S})^{-1}\mathbf{Y})^2}{\alpha - \mathbf{U}^\mathsf{T}(\mathbf{A}+\mathbf{S})^{-1}\mathbf{U}} \Big).$$

This is a Gaussian density with mean $\mu = \mathbf{V}^\mathsf{T}\mathbf{Y}$ and variance $\sigma^2 = 2\big(\alpha - \mathbf{U}^\mathsf{T}\mathbf{V}\big)$, where $\mathbf{V} = (\mathbf{A}+\mathbf{S})^{-1}\mathbf{U}$. Using the definition of $\mathbf{A}$, $\mathbf{U}$ and $\mathbf{S}$, and denoting by v_l its lth component, $\mathbf{V}$ has to satisfy

$$\frac{\varepsilon_H^2}{2} \Big\langle {}_\tau^{-\frac{1}{2}-H}\partial \psi_{t_0}, {}_\tau^{-\frac{1}{2}-H}\partial \psi_k \Big\rangle = \sum_{l=-K}^{K} \Big(\frac{\varepsilon_H^2}{2} \Big\langle {}_\tau^{-\frac{1}{2}-H}\partial \psi_k, {}_\tau^{-\frac{1}{2}-H}\partial \psi_l \Big\rangle + \frac{1}{2} c_{nn}[k-l] \Big) v_l \tag{28}$$

for all $k = -K, -K+1, \dots K$.

b) Step 2—Computation of $\mu(t_0|Y)$: We let K tend to infinity and observe that finding v_l from the set of equations (28) is still well-posed because the quadratic form characterized by the (infinite) matrix $(\mathbf{A}+\mathbf{S})$ is positive-definite; i.e., if there is a solution—which is not ensured when $K = \infty$—then this solution is unique.

We will reinterpret (28) from an interpolation viewpoint. By inspection of (8), we observe that $c_{xx}(t_0, k) = \varepsilon_H^2 \langle {}_\tau^{-\frac{1}{2}-H}\partial \delta(t-t_0), {}_\tau^{-\frac{1}{2}-H}\partial \delta(t-k) \rangle$ is in the $\text{span}_{l\in\mathbb{Z}}\{\beta_0^{2H}(t_0 - l)\}$. Denoting by $\varphi_{\text{int}}(t) = \varphi_0(t)$ the unique interpolating function that belongs to this span[6], we thus have the interpolation formula

$$\begin{aligned} c_{xx}(t_0, k) &= \sum_{l\in\mathbb{Z}} c_{xx}(l, k) \varphi_{\text{int}}(t_0 - l) \\ &= \sum_{l\in\mathbb{Z}} c_{xx}(l, k) \big(\varphi_{\text{int}}(t_0 - l) - \rho(t_0) \varphi_{\text{int}}(l) \big). \end{aligned}$$

where $\rho(t_0)$ is the function of t_0 specified in (24); the second line above has been obtained by using the interpolation condition $\varphi_{\text{int}}(l) = \delta_l$, and the fact that $c_{xx}(0, k) = 0$. This identity can also be reformulated using a smoothing spline basis $\{\varphi_\Lambda(t_0 - l)\}_{l\in\mathbb{Z}}$: $\varphi_{\text{int}}(t_0) = \varphi_\Lambda(t_0) + \sum_l r_l \varphi_\Lambda(t_0 - l)$ where

$$R(e^{j\omega}) = \frac{1}{\varepsilon_H^2} \frac{|1 - e^{-j\omega}|^{2H+1} C_{nn}(e^{j\omega})}{B_0^{2H}(e^{j\omega})}$$

To do this, we write

$$v_l(t_0) = \varphi_\Lambda(t_0 - l) - \rho(t_0) \varphi_\Lambda(l), \tag{29}$$

which allows to express $\varphi_{\text{int}}(t_0 - l) - \rho(t_0)\varphi_{\text{int}}(l)$ in terms $v_l(t_0)$ as

$$\varphi_{\text{int}}(t_0 - l) - \rho(t_0)\varphi_{\text{int}}(l) = v_l(t_0) + \sum_{l'\in\mathbb{Z}} r_{l'-l} v_{l'}(t_0).$$

[6]In Fourier variables, φ_{int} is obtained by $\hat{\varphi}_{\text{int}}(\omega) = \hat{\beta}_0^{2H}(\omega)/B_0^{2H}(e^{j\omega})$.

Replacing in the interpolation formula, we get

$$c_{xx}(t_0,k) = \sum_{l\in\mathbb{Z}} c_{xx}(l,k)v_l(t_0) + \sum_{l,l'\in\mathbb{Z}} r_{l'-l}\, c_{xx}(l,k)v_{l'}(t_0).$$

The second term on the right hand side is evaluated as

$$\begin{aligned}
\sum_{l,l'\in\mathbb{Z}} r_{l'-l}\, c_{xx}(l,k)v_{l'}(t_0) &= \frac{\varepsilon_H^2}{2\pi}\int \sum_{l,l'\in\mathbb{Z}} r_{l'-l}\frac{(e^{jl\omega}-1)(e^{-jk\omega}-1)}{|\omega|^{2H+1}}v_{l'}(t_0)\,\mathrm{d}\omega \\
&= \frac{\varepsilon_H^2}{2\pi}\int \sum_{l'\in\mathbb{Z}} (e^{-jk\omega}-1)\frac{e^{jl'\omega}R(e^{j\omega})}{|\omega|^{2H+1}}v_{l'}(t_0)\,\mathrm{d}\omega \\
&= \frac{1}{2\pi}\int (e^{-jk\omega}-1)\frac{C_{nn}(e^{j\omega})}{B_0^{2H}(e^{j\omega})}\hat{\beta}_0^{2H}(\omega)\sum_{l'\in\mathbb{Z}} e^{jl'\omega}v_{l'}(t_0)\,\mathrm{d}\omega \\
&= \frac{1}{2\pi}\int_{-\pi}^{\pi} (e^{-jk\omega}-1)C_{nn}(e^{j\omega})\sum_{l'\in\mathbb{Z}} e^{jl'\omega}v_{l'}(t_0)\,\mathrm{d}\omega \\
&= \sum_{l'\in\mathbb{Z}} \big(c_{nn}[l'-k]-c_{nn}[l']\big)v_{l'}(t_0) \\
&= \sum_{l'\in\mathbb{Z}} \underbrace{c_{nn}[l'-k]}_{=c_{nn}[k-l']}v_{l'}(t_0) - \underbrace{\sum_{l'\in\mathbb{Z}} c_{nn}[l']\big(\varphi_\Lambda(t_0-l')-\rho(t_0)\varphi_\Lambda(l')\big)}_{=0 \text{ by definition of } \rho(t_0)}.
\end{aligned}$$

This shows that $v_l(t_0)$ as posed by (29) satisfies $c_{xx}(t_0,k) = \sum_l c_{xx}(l,k)v_l(t_0) + \sum_{l'} c_{nn}[k-l']v_{l'}(t_0)$ which is precisely (28) up to a factor $1/2$. $v_l = v_l(t_0)$ is thus the solution we are looking for, which implies that

$$\mu(t_0|Y) = \underbrace{\sum_{k\in\mathbb{Z}} y_k\varphi_\Lambda(t_0-k)}_{\tilde{x}(t_0|Y)} - \rho(t_0)\underbrace{\sum_{k\in\mathbb{Z}} y_k\varphi_\Lambda(k)}_{\tilde{x}(0|Y)}.$$

The last technical step is to show that the random quantity $\tilde{x}(t|Y)$ is well-defined (convergent) almost surely; this is ensured by the fact that $\varphi_\Lambda(t)$ decreases fast enough ($\propto |k|^{-2H-2}$) to tame the divergence of $\mathscr{E}\{|y_k|\} \propto |k|^H$ as $k \to \pm\infty$.

c) Step 3—Computation of $\sigma^2(t_0|Y)$*:* The variance of the posterior density is given by $\sigma^2 = 2\big(\alpha - \mathbf{U}^\mathsf{T}\mathbf{V}\big)$. Since $2\alpha = c_{xx}(t_0,t_0) = C_H|t_0|^{2H}$, it remains to compute $\mathbf{U}^\mathsf{T}\mathbf{V}$. We have

$$\begin{aligned}
2\mathbf{U}^\mathsf{T}\mathbf{V} &= \sum_{l\in\mathbb{Z}} c_{xx}(t_0,l)v_l(t_0) \\
&= \frac{C_H}{2}\sum_{l\in\mathbb{Z}} \big(|t_0|^{2H}+|l|^{2H}-|t_0-l|^{2H}\big)v_l(t_0) \\
&= \frac{C_H}{2}\Big[|t_0|^{2H}\big(1-\rho(t_0)\big) + \sum_{l\in\mathbb{Z}} |l|^{2H}\varphi_\Lambda(t_0-l) - \rho(t_0)\sum_{l\in\mathbb{Z}} |l|^{2H}\varphi_\Lambda(l) \\
&\qquad - \sum_{l\in\mathbb{Z}} |t_0-l|^{2H}\varphi_\Lambda(t_0-l) + \rho(t_0)\sum_{l\in\mathbb{Z}} |t_0-l|^{2H}\varphi_\Lambda(l)\Big]
\end{aligned}$$

To evaluate this expression we first notice that

$$\sum_{l\in\mathbb{Z}} |l|^{2H}\varphi_\Lambda(t_0-l) = \sum_{l\in\mathbb{Z}} |t_0-l|^{2H}\varphi_\Lambda(l)$$

because $\varphi_\Lambda(t_0-l)$ and $|t_0-l|^{2H}$ span the same space and because the right and the left hand sides coincide at the integers. Secondly, the following sequence of equalities

$$\begin{aligned}
|t_0|^{2H} &= \sum_{l\in\mathbb{Z}} |l|^{2H}\varphi_{\text{int}}(t_0-l) \\
&= \sum_{l\in\mathbb{Z}} |l|^{2H}\varphi_\Lambda(t_0-l) + \sum_{l,l'\in\mathbb{Z}} r_{l'-l}|l|^{2H}\varphi_\Lambda(t_0-l') \\
&= \sum_{l\in\mathbb{Z}} |l|^{2H}\varphi_\Lambda(t_0-l) + \sum_{l,l'\in\mathbb{Z}} r_{l'-l}|l|^{2H}\varphi_\Lambda(t_0-l') \\
&= \sum_{l\in\mathbb{Z}} |l|^{2H}\varphi_\Lambda(t_0-l) + \text{Const}\times \underbrace{\sum_{l'\in\mathbb{Z}} c_{nn}[l']\varphi_\Lambda(t_0-l')}_{=\text{Const}'\times\rho(t_0)}
\end{aligned}$$

shows that

$$\sum_{l\in\mathbb{Z}} |l|^{2H}\varphi_\Lambda(t_0-l) = |t_0|^{2H} + \text{Const}''\times\rho(t_0)$$

where the constant is obtained as $\sum_{l\in\mathbb{Z}}|l|^{2H}\varphi_\Lambda(l)$ by enforcing the equality at $t_0=0$ because $\rho(0)=1$.

Putting things together, we obtain

$$2\mathbf{U}^\mathsf{T}\mathbf{V} = \frac{C_H}{2}\Big[2|t_0|^{2H} + \rho(t_0)^2\sum_{l\in\mathbb{Z}}|l|^{2H}\varphi_\Lambda(l) - \sum_{l\in\mathbb{Z}}|t_0-l|^{2H}\varphi_\Lambda(t_0-l)\Big]$$

which yield the expression for $\sigma^2(t_0|Y)$.

■

APPENDIX II
PROOF OF PROPOSITION 3

We will work with the correlation form of the estimation error $e(t)$. In order to deal with continuous-time processs only, we extrapolate the discrete stationary process $n[k]$ to a continuous stationary process $n(t)$ with correlation function $\eta(t)$, such that $n(k)=n[k]$.

Let $\phi(t)$ be a function of Schwartz' class $\mathscr{S}$. Then, under the hypothesis that φ_k decreases fast enough as $k\to\infty$, we have $\langle e,\phi\rangle = \langle x,\psi_\phi\rangle + \langle n,\psi_\phi-\phi\rangle$ where

$$\psi_\phi(t) \stackrel{\text{def}}{=} \phi(t) - \sum_{k\in\mathbb{Z}}\langle\varphi_k,\phi\rangle\delta(t-k).$$

Since $x(t)$ is an fBm, the correlation form C of $e(t)$ can be expressed as

$$C(\phi,\phi) = \frac{\varepsilon_H^2}{2\pi}\int \frac{1}{|\omega|^{2H+1}}\Big|\hat\psi_\phi(\omega) - \hat\psi_\phi(0) - \hat\psi'_\phi(0)\omega - \ldots \frac{\hat\psi_\phi^{(\lfloor H\rfloor)}(0)}{\lfloor H\rfloor!}\omega^{\lfloor H\rfloor}\Big|^2 \mathrm{d}\omega$$
$$+ \frac{1}{2\pi}\int \hat\eta(\omega)|\hat\psi_\phi(\omega) - \hat\phi(\omega)|^2\mathrm{d}\omega$$

for all $\phi \in \mathscr{S}$ whereas, if $x(t)$ were a stationary process with correlation function $\rho(t)$, the correlation form of $e(t)$ would be given by

$$C_{\text{stationary}}(\phi,\phi) = \frac{1}{2\pi}\int \hat\rho(\omega)|\hat\psi_\phi(\omega)|^2\mathrm{d}\omega + \frac{1}{2\pi}\int \hat\eta(\omega)|\hat\psi_\phi(\omega) - \hat\phi(\omega)|^2\mathrm{d}\omega.$$

We are interested in the conditions on φ_k such that $C(\phi_1,\phi_2) = C_{\text{stationary}}(\phi_1,\phi_2)$ for all $\phi_1,\phi_2 \in \mathscr{S}$.

Let us define the subspace of $\mathscr{S}$

$$\mathscr{S}_0 = \{\phi \in \mathscr{S} \text{ s.t. } \langle \phi(t), t^l - \sum_{k\in\mathbb{Z}} k^l \phi_k(t)\rangle = 0, \quad \text{for } l = 0\ldots\lfloor H\rfloor\} \tag{30}$$

It is easy to check through Parseval identity that these conditions are equivalent to $\hat\psi_\phi^{(l)}(0) = 0$ for $l = 0\ldots\lfloor H\rfloor$.

When $\phi \in \mathscr{S}_0$, the correlation form of the estimation error of an fBm becomes similar to the one of a (hypothetical) stationary process with $\hat\rho(\omega) = \varepsilon_H^2|\omega|^{-2H-1}$. If we enforce the identity $C(\phi,\phi_0) = C_{\text{stationary}}(\phi,\phi_0)$ for all $\phi \in \mathscr{S}$ and $\phi_0 \in \mathscr{S}_0$, then we end up with the following equation

$$\int \hat\rho(\omega)\hat\psi_{\phi_0}(\omega)^*\Big(\hat\psi_\phi(0) + \hat\psi'_\phi(0)\omega + \ldots \frac{\hat\psi_\phi^{(\lfloor H\rfloor)}(0)}{\lfloor H\rfloor!}\omega^{\lfloor H\rfloor}\Big)\mathrm{d}\omega = 0, \quad \forall(\phi,\phi_0) \in \mathscr{S}\times\mathscr{S}_0.$$

Now, because the collection of numbers $\{\int \hat\rho(\omega)\hat\psi_{\phi_0}(\omega)^*\omega^l\mathrm{d}\omega\}_{l=0\ldots\lfloor H\rfloor}$ may assume arbitrary independent values, the only possibility for this identity to hold is that $\hat\psi_\phi^{(l)}(0) = 0$ for $l = 0\ldots\lfloor H\rfloor$ for all $\phi \in \mathscr{S}$. In other words: $\mathscr{S} = \mathscr{S}_0$. Using the definition (30) of $\mathscr{S}_0$, we get that $t^l - \sum_{k\in\mathbb{Z}} k^l\phi_k(t) = 0$ in the sense of distributions. ■

APPENDIX III
PROOF OF THEOREM 3

We follow the same first lines of the proof of Proposition 3 in Appendix II and we note that the autocorrelation of the continuous-time process $n(t)$ is related to the discrete one $n[k]$ through $c_{nn}[k] = \eta(k)$; i.e., $C_{nn}(e^{j\omega}) = \sum_{k\in\mathbb{Z}}\hat\eta(\omega + 2k\pi)$. After replacing (as the result of a limit process) $\phi(t)$ by $\delta(t-t_0)$, we get $\hat\psi_\phi(\omega) = e^{-jt_0\omega} - \Phi_{t_0}(e^{j\omega})$ where

$$\Phi_{t_0}(e^{j\omega}) = \sum_{k\in\mathbb{Z}} \varphi_k(t_0)e^{-jk\omega}.$$

This leads to the variance of the estimation error $e(t_0) = x(t_0) - \tilde{x}(t_0)$ which is obtained as $C(\delta(\cdot - t_0), \delta(\cdot - t_0))$:

$$\begin{aligned}
\mathscr{E}\{|e(t_0)|^2\} &= \frac{\varepsilon_H^2}{2\pi} \int \frac{|e^{-jt_0\omega} - \Phi_{t_0}(e^{j\omega})|^2}{|\omega|^{2H+1}} \mathrm{d}\omega + \frac{1}{2\pi} \int \hat{\eta}(\omega) |\Phi_{t_0}(e^{j\omega})|^2 \mathrm{d}\omega \\
&= \frac{\varepsilon_H^2}{2\pi} \int_0^{2\pi} \left(\sum_{k\in\mathbb{Z}} \frac{|e^{-jt_0(\omega+2k\pi)} - \Phi_{t_0}(e^{j\omega})|^2}{|\omega + 2k\pi|^{2H+1}} \; + \; \Lambda(e^{j\omega}) |\Phi_{t_0}(e^{j\omega})|^2 \right) \mathrm{d}\omega \\
&= \frac{\varepsilon_H^2}{2\pi} \int_0^{2\pi} \Big(\big(R_0(e^{j\omega}) + \Lambda(e^{j\omega})\big) |\Phi_{t_0}(e^{j\omega})|^2 - 2\mathrm{Re}\left\{ R_{t_0}(e^{j\omega})^* \, \Phi_{t_0}(e^{j\omega}) \right\} + R_0(e^{j\omega}) \Big) \mathrm{d}\omega \\
&= \frac{\varepsilon_H^2}{2\pi} \int_0^{2\pi} \big(R_0(e^{j\omega}) + \Lambda(e^{j\omega})\big) \Big| \Phi_{t_0}(e^{j\omega}) - \frac{R_{t_0}(e^{j\omega})}{R_0(e^{j\omega}) + \Lambda(e^{j\omega})} \Big|^2 \mathrm{d}\omega \\
&\qquad + \frac{\varepsilon_H^2}{2\pi} \int_0^{2\pi} \left(R_0(e^{j\omega}) - \frac{|R_{t_0}(e^{j\omega})|^2}{R_0(e^{j\omega}) + \Lambda(e^{j\omega})} \right) \mathrm{d}\omega
\end{aligned}$$

where we have defined $R_{t_0}(e^{j\omega}) = \sum_{k\in\mathbb{Z}} \frac{e^{-jt_0(\omega+2k\pi)}}{|\omega+2k\pi|^{2H+1}}$. Notice that, whenever this expression is finite, we automatically have that $\hat{\psi}_\phi(\omega)|\omega + 2k\pi|^{-H-1/2}$ is square integrable, which also implies that $\hat{\psi}_\phi^{(l)}(0) = 0$ for $l = 0 \ldots \lfloor H \rfloor$. Thus, minimizing the variance of the estimation error with respect to $\varphi_k(t_0)$ subject to the "stationarizing" constraint is simply equivalent to minimizing this same variance without any constraints.

This minimum is obviously given by

$$\begin{aligned}
\Phi_{t_0}(e^{j\omega}) &= \frac{R_{t_0}(e^{j\omega})}{R_0(e^{j\omega}) + \Lambda(e^{j\omega})} \\
&= \frac{\displaystyle\sum_{k\in\mathbb{Z}} \hat{\beta}_0^{2H}(\omega + 2k\pi) e^{-jt_0(\omega+2k\pi)}}{B_0^{2H}(e^{j\omega}) + \Lambda(e^{j\omega}) \Big| 2\sin\left(\frac{\omega}{2}\right) \Big|^{2H+1}} \\
&= \sum_{k\in\mathbb{Z}} \hat{\varphi}_\Lambda(\omega + 2k\pi) e^{-jt_0(\omega+2k\pi)} \\
&= \sum_{k\in\mathbb{Z}} \varphi_\Lambda(t_0 - k) e^{-jk\omega}
\end{aligned}$$

where we have used the fact that $\varphi_\Lambda(t)$ is an even function. Thus, by identifying left and right hand sides, we get that $\varphi_k(t_0) = \varphi_\Lambda(t_0 - k)$. Note that, even if the minimization has been performed over estimators that satisfy $e^{-jt_0\omega} - \Phi_{t_0}(e^{j\omega}) = o(|\omega|^{\lfloor H \rfloor})$, the optimal solution actually satisfies twice as many conditions namely: $e^{-jt_0\omega} - \Phi_{t_0}(e^{j\omega}) = o(|\omega|^{\lfloor 2H+1 \rfloor})$.

As a bonus, we have the expression of the variance of this minimum estimator

$$\begin{aligned}
\mathscr{E}\{|e(t_0)|^2\} &= \frac{\varepsilon_H^2}{2\pi} \int_0^{2\pi} \left(R_0(e^{j\omega}) - R_{t_0}(e^{j\omega})^* \Phi_{t_0}(e^{j\omega}) \right) \mathrm{d}\omega \\
&= \frac{\varepsilon_H^2}{2\pi} \int \frac{1 - e^{jt_0\omega} \Phi_{t_0}(e^{j\omega})}{|\omega|^{2H+1}} \mathrm{d}\omega
\end{aligned}$$

Similar to the proof of Thm. 1, we use the following limit

$$\frac{1}{|\omega|^{2H+1}} = -\frac{1}{2\Gamma(2H+1)\sin(\pi H)} \lim_{s\to 0} \hat{\lambda}_s(\omega)$$

where $\lambda_s(t) = |t|^{2H} e^{-s|t|}$. Using Lebesgue's dominated convergence theorem, we then express the error as

$$\begin{aligned}
\mathscr{E}\{|e(t_0)|^2\} &= -\frac{\varepsilon_H^2}{4\pi\Gamma(2H+1)\sin(\pi H)} \lim_{s\to 0} \int \left(1 - e^{jt_0\omega}\Phi_{t_0}(e^{j\omega})\right)\hat{\lambda}_s(\omega)\mathrm{d}\omega \\
&= \frac{\varepsilon_H^2}{4\pi\Gamma(2H+1)\sin(\pi H)} \lim_{s\to 0} \int e^{jt_0\omega}\Phi_{t_0}(e^{j\omega})\hat{\lambda}_s(\omega)\mathrm{d}\omega \\
&= \frac{\varepsilon_H^2}{2\Gamma(2H+1)\sin(\pi H)} \lim_{s\to 0} \sum_{k\in\mathbb{Z}} \varphi_\Lambda(t_0-k)\lambda_s(t_0-k) \\
&= \frac{\varepsilon_H^2}{2\Gamma(2H+1)\sin(\pi H)} \sum_{k\in\mathbb{Z}} \varphi_\Lambda(t_0-k)|t_0-k|^{2H}
\end{aligned}$$

■

REFERENCES

[1] M. Unser and T. Blu, "Self-similarity: Part I—Splines and operators," submitted.

[2] N. Wiener, *Extrapolation, Interpolation and Smoothing of Stationary Time Series with Engineering Applications*, MIT Press, Cambridge MA, USA, 1964.

[3] A. Papoulis, *Probability, Random Variables, and Stochastic Processes*, McGraw-Hill, New York, 1991.

[4] M. Unser and T. Blu, "Generalized smoothing splines and the optimal discretization of the Wiener filter," *IEEE Trans. Signal Processing*, vol. 53, no. 6, pp. 2146–2159, June 2005.

[5] I.J. Schoenberg, "Contribution to the problem of approximation of equidistant data by analytic functions," *Quart. Appl. Math.*, vol. 4, pp. 45–99, 112–141, 1946.

[6] B. B. Mandelbrot, *Gaussian Self-Affinity and Fractals*, Springer, Berlin, 2001.

[7] B. B. Mandelbrot and J. W. Van Ness, "Fractional Brownian motions fractional noises and applications," *SIAM Review*, vol. 10, no. 4, pp. 422–437, 1968.

[8] R. Szeliski and D. Terzopoulos, "From splines to fractals," *Computer Graphics (Proc. SIGGRAPH'89)*, vol. 23, no. 4, pp. 51–60, 1989.

[9] G.M. Molchan, "Gaussian processes with spectra that are asymptotically equivalent to a power of λ," *Theory of Proba. and its Appl.*, vol. 14, pp. 530532, 1969.

[10] P. Flandrin, "On the spectrum of fractional Brownian motions," *IEEE Transactions on Information Theory*, vol. 35, no. 1, pp. 197–199, 1989.

[11] P. Lévy, *Le mouvement Brownien*, Gauthier-Villars, Paris, France, 1954.

[12] A.M. Yaglom, "Correlation theory of processes with random stationary nth increments," *AMS Transl.*, vol. 2, no. 8, pp. 87141, 1958.

[13] G. Gripenberg and I. Norros, "On the prediction of fractional Brownian motion," *J. Appl. Probab.*, vol. 33, pp. 400–410, 1996.

[14] C.J. Nuzman and H. V. Poor, "Linear estimation of self-similar process via Lamperti's transformation," *J. Appl. Prob.*, vol. 37, no. 2, pp. 429–452, 2000.

[15] R. J. Barton and H. V. Poor, "Signal-detection in fractional Gaussian-noise," *IEEE Transactions on Information Theory*, vol. 34, no. 5, pp. 943–959, 1988.

[16] Y. Meyer, F. Sellan, and M. S. Taqqu, "Wavelets, generalized white noise and fractional integration: The synthesis of fractional Brownian motion," *Journal of Fourier Analysis and Applications*, vol. 5, no. 5, pp. 465–494, 1999.

[17] G. Matheron, "The intrinsic random functions and their applications," *Applied Probability*, vol. 5, no. 12, pp. 439–468, 1973.

[18] M.S. Taqqu, "Fractional Brownian motion and long-range dependence," in *Theory and applications of long-range dependence*, P. Doukhan, G. Oppenheim, and M.S. Taqqu, Eds., pp. 5–38. Birkhauser, Boston, USA, 2003.

[19] A.M. Yaglom, *Correlation theory of stationary and related random functions I: Basic results*, Springer, New York, USA, 1986.

[20] E. Perrin, R. Harba, C. Berzin-Joseph, I. Iribarren, and A. Bonami, "nth-order fractional Brownian motion and fractional Gaussian noises," *IEEE Trans. Signal Processing*, vol. 49, no. 5, pp. 1049–1059, 2001.

[21] I.M. Gelfand and N. Ya. Vilenkin, *Generalized functions. Vol. 4. Applications of harmonic analysis*, Academic press, New York, USA, 1964.

[22] L. Schwartz, *Théorie des Distributions*, Hermann, Paris, 1966.

[23] V. Olshevsky and L. Sakhnovich, "Matched filtering of generalized stationary processes," *IEEE Transactions on Information Theory*, vol. 51, no. 9, pp. 3308–3313, 2005.

[24] K. Itô, *Stochastic Differential Equations*, Memoirs of the American Mathematical Society. American Mathematical Society, Providence, RI, 1951.

[25] C.W. Gardiner, *Handbook of Stochastic Methods*, Springer, Berlin, 2004.

[26] M. Unser and T. Blu, "Fractional splines and wavelets," *SIAM Review*, vol. 42, no. 1, pp. 43–67, 2000.

[27] N. Cressie, "Geostatistics," *American Statistician*, vol. 43, no. 4, pp. 197–202, 1989.

[28] G. Wahba, *Spline models for observational data*, Society for Industrial and Applied Mathematics, Philadelphia, 1990.

[29] D. E. Myers, "Kriging, cokriging, radial basis functions and the role of positive definiteness," *Computers and Mathematics with Applications*, vol. 24, no. 12, pp. 139–148, 1992.

[30] S. Mallat, *A wavelet tour of signal processing*, Academic Press, San Diego, 1998.

[31] M. Unser and T. Blu, "Wavelet theory demystified," *IEEE Trans. Signal Process.*, vol. 51, no. 2, pp. 470–483, February 2003.

[32] P. Flandrin, "Wavelet analysis and synthesis of fractional Brownian-motion," *IEEE Transactions on Information Theory*, vol. 38, no. 2, pp. 910–917, 1992.

[33] G. W. Wornell, "Wavelet-based representations for the $1/f$ family of fractal processes," *Proceedings of the IEEE*, vol. 81, no. 10, pp. 1428–1450, 1993.

[34] M. Unser, A. Aldroubi, and M. Eden, "B-spline signal processing: Part II—Efficient design and applications," *IEEE Trans. Signal Process.*, vol. 41, no. 2, pp. 834–848, February 1993.

[35] S. Tirosh, D. Van de Ville, and M. Unser, "Polyharmonic smoothing splines and the multi-dimensional Wiener filtering of fractal-like signals," *IEEE Transactions on Image Processing*, in press.