To appear in *Journal of Applied Probability.*

Linear Estimation of Self-Similar Processes via Lamperti's Transformation[1]

Carl J. Nuzman and H. Vincent Poor

Princeton University

September 27, 1999

Lamperti's transformation, an isometry between self-similar and stationary processes, is used to solve some problems of linear estimation of continuous-time, self-similar processes. These problems include causal whitening and innovations representations on the positive real line, as well as prediction from certain finite and semi-infinite intervals. The method is applied to the specific case of fractional Brownian motion (fBm), yielding alternate derivations of known prediction results, along with some novel whitening and interpolation formulae. Some associated insights into the problem of discrete prediction are also explored. Closed-form expressions for the spectra and spectral factorization of the stationary processes associated with the fBm are obtained as part of this development.

[1] Research supported in part by the U. S. Office of Naval Research under Grant N00014-94-1-0115, and in part by the U. S. Department of Defense NDSEG Fellowship Program

AMS 1991 subject classifications: Primary 60G18; secondary 62M20, 60G25

Key words and phrases: self-similar processes, fractional Brownian motion, linear prediction, whitening, innovations, scale-stationary processes, Lamperti's transformation, Wiener-Kolmogorov filter

Contents

1 Introduction

Beginning with the investigations of Hurst (1951) into river water levels, phenomena fundamentally invariant to changes in scale have been discovered empirically in diverse fields such as medicine, economics, and physics. Mathematically, a related class of random processes were introduced in Lamperti (1962). These processes, referred to as semi-stable processes, arose in the following way: if an infinite sequence of contractions of the time and space scales of a random process yields a limit process, then that limit process must be semi-stable. Mandelbrot referred to the same processes by the now pervasive term self-similar, and was a pioneer in applying self-similar models to understanding scale-invariant phenomena (see e.g. Mandelbrot 1982). Most recently, these processes have received renewed interest due to empirical observations of self-similarity in measurements of ethernet traffic [Leland et al. (1994)] and financial data [Willinger et al. (1999)].

In his original paper, Lamperti pointed out a simple isometry between self-similar processes and stationary processes. Some recent articles demonstrate that this mapping can be an important tool, allowing researchers to apply the extensive body of knowledge about stationary processes to examine the structure of the associated self-similar processes. In Yazici and Kashyap (1997), the authors construct a framework for analyzing second-order self-similar processes and linear scale-invariant systems based on well-known tools for working with second-order stationary processes and time-invariant systems. Similar ideas were also noted briefly in Chapter 7 of Wornell (1991). Burnecki et al (1997) use the transformation to study self-similar α-stable processes, and Albin (1998) provides a connection between the extremal behavior of self-similar processes and that of their stationary counterparts.

In this paper, we show how the connection with stationary processes can be exploited to solve certain problems of linear whitening and prediction of self-similar processes. In all of the problems treated, the index sets are real intervals having the origin as an endpoint; the principle insight is that a self-similar process on such an interval is isometric to a stationary process on a semi-infinite interval. Applying these results to the specific case of fractional Brownian motion (fBm), we obtain innovations representations, whitening filters, and prediction and interpolation formulae which, due to the stationary increments property of the fBm, hold for arbitrary real intervals. The interpolation results and some of the other representations for fBm have not been given explicitly before. In Section 2 we introduce Lamperti's transformation, and highlight pertinent properties of self-similar processes. In Section 3 we lay the groundwork for mean-square integration of such processes and give our main results on prediction and whitening. Section 4 develops the case of fBm in detail, and we end with some conclusions.

2 Self-Similar Random Processes

In this section we examine self-similar random processes from the point of view of Lamperti's transformation, generally following the discussion in Yazici and Kashyap (1997).

2.1 Lamperti's Transformation

Recall that a stochastic process $\{X(t), -\infty < t < \infty\}$ is called *stationary* (or *shift-stationary*) if it satisfies

$$\{X(t-\alpha)\}_{t\in\mathbb{R}} = \{X(t)\}_{t\in\mathbb{R}}$$

for all $\alpha \in \mathbb{R}$. That is, the finite-dimensional probability distributions of the process are unchanged by time shifts.

By analogy, a process $\{Y(t), t > 0\}$ is called *scale-stationary* if its distributions are invariant to scaling of the time axis, i.e. if

$$\{Y(at)\}_{t>0} = \{Y(t)\}_{t>0}, \quad a > 0.$$

Lastly, an *H-self-similar* (or H-ss) process is one whose distributions are almost invariant to scaling of the time axis. More precisely, scaling by a factor $a > 0$ has the same effect as multiplying the process by a factor a^H:

$$\{Y(at)\}_{t>0} = a^H\{Y(t)\}_{t>0}, \quad a > 0, \tag{1}$$

where H is referred to as the *self-similarity parameter* of the process. Note that a scale-stationary process is self-similar with parameter $H = 0$.

We note some properties of self-similar processes; see Vervaat (1987) for a comprehensive discussion. If $Y(t)$ is H-ss, then its second moment

$$E\left\{Y(t)^2\right\} = t^{2H} E\left\{Y(1)^2\right\}$$

is a power of t, if finite. For the important case $H > 0$, $Y(t)$ converges in distribution to zero as $t \to 0$, and hence it makes sense to define $Y(0) = 0$.

If $Y(t)$ is H'-ss, then $t^H Y(t)$ is $(H + H')$-ss. Note also that if $Y(t)$ is 0-ss on $\mathbb{R}^+ = \{\mathrm{t} : \mathrm{t} > 0\}$, then $X(\tau) = Y(e^\tau)$ is a stationary process on $\mathbb{R}$. Combining these observations, we see that for each H there is a one-to-one mapping between H-ss processes and shift-stationary processes.

Definition 2.1 *Let X be a set admitting multiplication by positive real numbers, and $I \subset (0, \infty)$ a positive real index set. For each $H \in \mathbb{R}$, the* Lamperti transformation with parameter H*, denoted L_H, is an invertible map between X-valued functions on I and X-valued functions on $\ln I$. For each $y : I \to X$, the function $L_H y : \ln I \to X$ is given by*

$$(L_H y)(\tau) = e^{-H\tau} y(e^\tau), \quad \tau \in \ln I.$$

For each $x : \ln I \to X$, the inverse transformation is given by

$$(L_H^{-1} x)(t) = t^H x(\ln t), \quad t \in I.$$

As is clear from the definition, the transformation can be applied to ordinary functions. For example, it is easy to see that $L_{-1/p}$ is an isomorphism between $L^p(I)$ and $L^p(\ln I)$. The main utility of this transformation, however, is the property noted in Lamperti (1962) that for an H-ss process $Y(t)$, the process $L_H Y(\tau)$ is stationary. Such a process is referred to as the stationary generator or generating process of Y. Similarly, one can define a second-order notion of self-similarity by applying L_H^{-1} to the wide-sense stationary processes. This leads to the following definition, from Yazici (1997):

Definition 2.2 *A random process $\{Y(t), t \in \mathbb{R}^+\}$ is called* wide-sense H-self-similar *if it satisfies:*

(i) $E\{Y(t)^2\} < \infty$

(ii) $E\{Y(at)\} = a^H E\{Y(t)\}$

(iii) $E\{Y(at_1)Y(at_2)\} = a^{2H} E\{Y(t_1)Y(t_2)\}$

for some $H > 0$ *and each* $a > 0$.

In the rest of this paper, stationarity and self-similarity will be taken to hold in this second-order sense, and we will assume that all processes have zero mean. Equivalently, we focus on properties of self-similar covariance matrices. Of course, Gaussian self-similar processes inherit from their shift-stationary generators the property that the strict and wide-sense definitions are equivalent.

Thus far we have only discussed H-ss processes on $\mathbb{R}^+$. When $H > 0$, we can define $Y(0) = 0$, and consider processes on $\mathbb{R}$ which satisfy (1). Such a process $Y(t)$ is really the concatenation of two H-ss processes on $\mathbb{R}^+$, namely the future process $\{Y(t), t > 0\}$ and the past process $\{Y(-t), t > 0\}$. The processes are *jointly H-ss*, meaning that their joint distributions satisfy the usual scaling law. Applying L_H to the pair produces a pair of jointly stationary processes, say $X_F = L_H Y(t)$ and $X_P = L_H Y(-t)$, which we refer to as the future and past generating processes.

2.2 Fractional Brownian Motion

In modeling many empirical phenomena, it is natural to require that a model have stationary increments. If this condition is applied to an H-self-similar process $Y(t)$, it follows that the increments $Y(s+t) - Y(s)$ form a stationary process in s and an H-ss process in t. To see the latter point, note that

$$\begin{aligned} \{Y(s+at) - Y(s)\}_{t,s\in\mathbb{R}} &= a^H \{Y(s/a+t) - Y(s/a)\}_{t,s\in\mathbb{R}} \\ &= a^H \{Y(s+t) - Y(s)\}_{t,s\in\mathbb{R}}. \end{aligned}$$

The added requirement of stationary increments in fact determines that $0 < H \le 1$, $E\{Y(t)\} = 0$ (unless $H = 1$), and

$$E\{Y(s)Y(t)\} = E\{Y(1)^2\} \cdot \frac{1}{2}\left(|s|^{2H} + |t|^{2H} - |s-t|^{2H}\right).$$

In the degenerate case $H = 1$, this process is a line with random slope $Y(1)$, i.e. $Y(t) = tY(1)$. Excluding this case, and restricting attention to the processes with finite variance, we see that the second-order properties of an H-ss si process are unique, up to a scale factor [Vervaat (1987)].

The Gaussian processes having the specified second-order structure, known collectively as fractional Brownian motion (fBm), are among the most important and widely used H-ss processes. As noted in Mandelbrot and Van Ness (1968), fBm divides naturally into three quite different classes, depending on the value of H. When $H > 1/2$, fBm is said to be *persistent*: disjoint increments are positively correlated. When $H < 1/2$, the increments have negative correlation and are called *anti-persistent*. When $H = 1/2$, fBm becomes the ordinary Brownian motion with independent increments.

If $H \neq 1/2$, the correlation between disjoint increments falls off hyperbolically as the distance between them approaches infinity:

$$E\left\{Y(1)\left(Y(s+1) - Y(s)\right)\right\} \sim 2H(2H-1)s^{2H-2}.$$

Processes exhibiting hyperbolic correlations are typically called *long-memory* or *long-term correlated* processes, although these terms are sometimes reserved for processes whose correlations are unsummable, true for fBm with $H > 1/2$.

3 Linear Operations on Self-Similar Processes

The simple mapping between self-similar and shift-stationary processes suggests the theoretical utility of working with self-similar processes by applying Lamperti's transformation, using stationary methods on the generating processes, and then applying the inverse transformation. In this section, we show how this technique applies in the cases of linear prediction and whitening. We first lay the groundwork by examining mean-square integration of self-similar processes and their stationary generators.

3.1 Mean-Square Integration

The basic tool which we will use for linear operations on wide-sense H-ss processes is the mean-square integral as presented, for example, in [Loève (1955)]. For our purposes, only deterministic integrands are required. The m-s integrals

$$(2) \qquad \int_I f(t)Y(t)dt \quad \text{and} \quad \int_I f(t)dY(t)$$

exist if and only if the deterministic integrals

$$(3) \qquad \int_I \int_I f(t)f(s)R_Y(t,s)dtds \quad \text{and} \quad \int_I f(t)\ d_t \int_I f(s)d_s R_Y(t,s),$$

respectively, exist. Each deterministic integral gives the variance of the corresponding m-s integral.

For fBm, the first requirement of (3) is

$$\frac{1}{2}\int_I \int_I f(t)f(s)\left[|t|^{2H} + |s|^{2H} - |t-s|^{2H}\right] dt\, ds < \infty,$$

satisfied for a wide class of integrands f for $H \in (0,1)$. Mean-square integration with respect to the increments of fBm, however, is defined only for $H \geq 1/2$. The integrands must satisfy

$$(4) \qquad 2H(2H-1)\int_I \int_I f(t)f(s)|t-s|^{2H-2} dt\, ds < \infty, \qquad H > 1/2$$

$$(5) \qquad \int_I f(t)^2\, dt < \infty, \qquad H = 1/2$$

in this case.

Mean-square integrals of self-similar processes are simply related to m-s integrals with respect to their stationary generators. For example, for Y H-ss and $X = L_H Y$,

$$(6) \qquad \int_a^b g(t)Y(t)dt = \int_a^b g(t)t^H X(\ln t)dt = \int_{\ln a}^{\ln b} g(e^\tau)e^{(H+1)\tau}X(\tau)d\tau$$

where the limits of integration a and b are non-negative. The existence of the transformed integral is assured by performing the analogous change of variables in (3).

The following lemma gives the corresponding formula for integrating the increments of a self-similar process.

Lemma 3.1 *Let $\{Y(t),\, t > 0\}$ be H-ss with stationary generator $X(\tau)$. Let $g(t)$ be a function bounded on the finite, positive interval $[a, b]$. Suppose further that the autocorrelation $r_X(t)$ of $X(\tau)$ satisfies the local Lipschitz condition*

$$r_X(0) - r_X(\tau) \leq L|\tau|, \quad |\tau| < \epsilon$$

for some $\epsilon > 0$ and some $L > 0$. Then

$$\int_a^b g(t)dY(t) = \int_{\ln a}^{\ln b} g(e^\tau)e^{H\tau}dX(\tau) + H\int_{\ln a}^{\ln b} g(e^\tau)e^{H\tau}X(\tau)d\tau$$

whenever the above three mean-square integrals exist.

A proof is provided in Appendix A. In essence, a rule analogous to the product rule for differentiation gives

$$dY(e^\tau) = d(e^{H\tau}X(\tau)) = e^{H\tau}\left[dX(\tau) + HX(\tau)d\tau\right]$$

We think of $dX(\tau)+HX(\tau)d\tau$ as the stationary generator of the $(H-1)$-ss increments process $dY(t)$. In general, $Y(t)$ need not be mean-square differentiable, so that the term "increments process" is heuristic. The above lemma extends immediately to the cases $a = 0$ and $b = \infty$ as long as $g(t)$ is bounded for each finite interval as $a \to 0$ and $b \to \infty$.

Stationary processes are commonly modeled as the result of passing other processes through stable, time-invariant linear filters. Let $Z(\tau) = \int_{-\infty}^{\infty} f(\tau - \sigma)X(\sigma)d\sigma$ be a stationary process with spectral density $\phi_Z(\omega) = |F(\omega)|^2\phi_X(\omega)$ obtained by filtering the generator $X(\tau) = L_H Y(\tau)$. Then the 0-ss process $Z(\ln t)$ can be expressed as

$$\begin{aligned} Z(\ln t) &= \int_{-\infty}^{\infty} f(\ln t - \sigma)X(\sigma)d\sigma \\ &= \int_0^{\infty} f(\ln t/s)s^{-H-1}Y(s)ds. \end{aligned} \tag{7}$$

Similarly, if the generalized process $dX(\tau) + HX(\tau)$ is passed through the filter, then $\phi_Z(\omega) = |F(\omega)|^2(H^2 + \omega^2)\phi_X(\omega)$ and

$$\begin{aligned} Z(\ln t) &= \int_{-\infty}^{\infty} f(\ln t - \sigma)\left[dX(\sigma) + HX(\sigma)d\sigma\right] \\ &= \int_0^{\infty} f(\ln t/s)s^{-H}dY(s). \end{aligned} \tag{8}$$

3.2 Linear Prediction of Self-Similar Processes

The problem of linear prediction of any second-order random process is in principle straightforward; the crux of the problem is solving the associated Wiener-Hopf integral equation. When the process is stationary, the special structure of the problem can be exploited to solve the equation in the frequency domain. We show in this section that self-similar processes have structure which can be exploited in exactly the same way, via Lamperti's transformation.

In general, we are interested in linearly predicting a random variable Z after observing an H-ss process $\{Y(t), t \in I\}$. That is, we seek the random variable $\hat{Z}$ in the linear space $L^2(Y, I)$ of Y that minimizes the mean-square distance $E\left\{(Z - \hat{Z})^2\right\}$.

The linear spaces $L^2(Y, I)$ and $L^2(X = L_H Y, \ln I)$ are identical, since for example applying L_H to a finite sum of the form $\sum a_i Y(t_i)$ yields a finite sum of $X(\ln t_i)$. Hence, the optimum linear predictor $\hat{Z}$ based on $\{Y(t), t \in I\}$ is the same as the optimum linear predictor based on $\{X(\tau), \tau \in \ln I\}$. Since $X(\tau)$ is stationary, we can apply well-known results on the prediction of stationary processes.

Example 3.1 : Infinite-interval Prediction

Consider for $H > 0$ a H-ss process $Y(t)$ defined on the whole real line, and suppose that we wish to predict $Y(a)$, $a > 0$ from observations of $Y(t)$ on the semi-infinite interval $I = (-\infty, 0)$. Letting $X_F = L_H Y(t)$ and $X_P = L_H Y(-t)$, we equivalently seek to predict $X_F(\ln a)$ after observing $X_P(\tau)$ on $I' = \mathbb{R}$. That is, given the entire evolution of one stationary process on $\mathbb{R}$, we wish to predict a jointly stationary process. The solution to this problem is given by a non-causal Wiener filter (see e.g. Poor (1994)). We denote by $\phi_P(\omega)$ and $\phi_{FP}(\omega)$ the spectrum of the past process and the cross-spectrum of the forward and past processes, respectively. If the filter response

$$F(\omega) = \frac{\phi_{FP}(\omega)}{\phi_P(\omega)}$$

has inverse Fourier transform $f(\tau)$, then the optimum linear predictor is

$$\hat{X}_F(\ln a) = \int_{-\infty}^{\infty} f(\ln a - \tau) X_P(\tau) d\tau.$$

Rewriting this solution in terms of $Y(t)$ via (7) gives

$$\begin{aligned}
\hat{Y}(a) &= a^H \hat{X}_F(\ln a) \\
&= \int_0^\infty a^H f(\ln(a/t)) t^{-H-1} Y(-t) dt \\
&\stackrel{\text{def}}{=} \int_{-\infty}^0 k_\infty(a,t) Y(-t) dt. \qquad (9)
\end{aligned}$$

The time-invariance of the stationary-domain integral implies self-similarity of the prediction kernel $k_\infty(a,t)$:

$$\begin{aligned}
k_\infty(a,t) &= a^{-1}(a/t)^{H+1} f(\ln(a/t)) \\
&= a^{-1}(1/(t/a))^{H+1} f(\ln(1/(t/a))) \\
&= a^{-1} k_\infty(1, t/a) \qquad (10)
\end{aligned}$$

The minimum mean squared error (MMSE) of linear prediction follows immediately from the well-known result for the Wiener filter. Denoting by

$$D^2(Z) \stackrel{\text{def}}{=} E\left\{(Z - \hat{Z})^2\right\}$$

the MMSE for predicting the random variable Z, we have

$$D^2(Y(a)) = a^{2H} D^2(X_F(\ln a)) = \frac{a^{2H}}{2\pi} \int_{-\infty}^\infty \left[\phi_F(\omega) - \frac{|\phi_{FP}(\omega)|^2}{\phi_P(\omega)}\right] d\omega \qquad (11)$$

Because $D^2(X_F(\alpha))$ is independent of α, $D^2(Y(a))$ is proportional to a^{2H}. Intuitively, one might expect that the ability to predict a process should improve as the time step a gets smaller. This is true here in the sense that the prediction error becomes arbitrarily small as $a \to 0$. However, the *normalized* MMSE

$$d^2(Y(a)) \stackrel{\text{def}}{=} \frac{D^2(Y(a))}{E\{Y(a)^2\}} = \frac{a^{2H} D^2(Y(1))}{a^{2H} E\{Y(1)^2\}} = d^2(Y(1)),$$

is independent of a. *In this relative sense, it is just as difficult to predict the next millisecond of a self-similar process, based on the infinite past, as it is to predict the next millenium.*

Example 3.2 : Finite-interval Prediction

Now consider prediction of $Y(a)$, $a > 0$ as in the previous example, but with observations restricted to the finite past, $I = [-T, 0)$. In terms of the generators, we are predicting $X_F(\ln a)$ from observations of $X_P(\tau)$ on the semi-infinite interval $I' = (-\infty, \ln T]$. The solution to this problem is the non-causal Weiner filter. It is convenient to define $c = a/T$. We get

$$\hat{X}_F(\ln a) = \int_{-\infty}^{\ln a - \ln c} f_c(\ln a - \tau) X_P(\tau) d\tau \tag{12}$$

where $f_c(t)$ is the inverse transform of

$$F_c(\omega) = \frac{1}{\phi_P^+(\omega)} \left[\frac{\phi_{FP}(\omega)}{\phi_P^+(-\omega)} \right]_+^{\ln c}. \tag{13}$$

Here, the function $\phi_P^+(\omega)$ is the causal spectral factorization of $\phi_P(\omega)$ satisfying $\phi_P(\omega) = |\phi_P^+(\omega)|^2$, which exists if $\phi_P(\omega)$ satisfies the Paley-Wiener condition

$$\int_{-\infty}^{\infty} \frac{|\ln \phi_P(\omega)|}{1+\omega^2} d\omega < \infty. \tag{14}$$

If $x(t) \leftrightarrow X(\omega)$ are a Fourier transform pair, the notation $[X(\omega)]_+^a$ denotes the Fourier transform of $x(t)u(t-a)$, where throughout this work $u(t)$ is the indicator function of $\mathbb{R}^+$.

Transforming back to the stationary domain, we get

$$\begin{aligned} \hat{Y}(a) &= (a)^H \hat{X}_F(\ln a) \\ &= (cT)^H \int_0^T f_c(\ln(cT/t))\, t^{-H-1} Y(-t)\, dt \\ &\stackrel{\text{def}}{=} \int_0^T m_T(c,t) Y(-t) dt. \end{aligned} \tag{15}$$

Similar results hold when the observation interval and prediction time are both positive. In that case,

$$F_c(\omega) = \frac{1}{\phi_F^+(\omega)} \left[\phi_F^+(\omega) \right]_+^{\ln c}, \tag{16}$$

and

$$\begin{aligned} \hat{Y}(cT) &= (cT)^H \int_0^T f_c(\ln(cT/t))\, t^{-H-1} Y(t)\, dt \\ &\stackrel{\text{def}}{=} \int_0^T m_T(c,t) Y(t)\, dt \end{aligned} \tag{17}$$

In either of these filtering problems, it may occur that $F_c(\omega)$ has no inverse transform, but that the frequency response $G_c(\omega) = F_c(\omega)/(H + j\omega)$ is invertible. In view of (8), the solution can then be expressed in terms of the increments of $Y(t)$ as $\hat{Y}(a) = \int_0^T M_T(c,t)dY(\pm t)$ with

$$M_T(c,t) = (cT)^H g_c(\ln(cT/t))\, t^{-H},$$

provided that the integral is well-defined, and that the conditions of Lemma 3.1 are satisfied.

From (13) and (16), the impulse response $f_c(\tau)$ depends on a and T only through the ratio $c = a/T$. It follows that the prediction kernels satisfy the self-similarity properties

$$m_T(c,t) = T^{-1} m_1(c, t/T), \quad M_T(c,t) = M_1(c, t/T). \tag{18}$$

It is straightforward to show that the best linear estimate of $Y(a)$ based on $[-T, 0]$ converges in the mean-square sense to the infinite-interval estimate, as $T \to \infty$. Then the infinite-observation kernel is available as the limit

$$k_\infty(a,t) = \lim_{T\to\infty} m_T(a/T, t),$$

in the sense that

$$\int_{-\infty}^0 \int_{-\infty}^0 e(t)e(s)R_Y(t,s)\, dt\, ds = 0$$

where $e(t) = k_\infty(a,t) - \lim_{T\to\infty} m_T(a/T,t)$.

3.3 Whitening and Innovations Representation

In this section we seek to represent causal linear operators connecting the process $Y(t)$ on $[0,\infty)$ with another process $W(t)$ having the same second order structure as the Wiener process. We assume that the stationary generator $X(\tau)$ of $Y(t)$ satisfies the Paley-Wiener condition.

The stationary generator $Z(\tau) = e^{-\tau/2}W(e^\tau)$ of $W(t)$ has autocorrelation function $r_Z(\tau) = e^{-|\tau|/2}$ and spectral density $\phi_Z(\omega) = 1/(1/4 + \omega^2)$. Then $Z(\tau)$ can be obtained by passing $X(\tau)$ through a causal filter with frequency response

$$F_w(\omega) = \frac{\phi_Z^+(\omega)}{\phi_X^+(\omega)} = \frac{1}{(1/2 + i\omega)\phi_X^+(\omega)}.$$

The process $X(\tau)$ can be recovered causally by passing the increment generator $dW(\tau) + 1/2W(\tau)$ through a causal filter

$$G_i(\omega) = \frac{\phi_X^+(\omega)}{(1/2 + i\omega)\phi_Z^+(\omega)} = \phi_X^+(\omega).$$

If $f_w(t)$ and $g_i(t)$ are the corresponding filter responses, we have:

$$W(t) = t^{1/2} \int_0^t f_w(\ln(t/s))s^{-H-1}Y(s)\,ds \tag{19}$$

and

$$Y(t) = t^H \int_0^t g_i(\ln(t/s))s^{-1/2}dW(s). \tag{20}$$

The examples of prediction, whitening, and innovations representation illustrate the utility of Lamperti's transformation, and the significance of the time origin in the structure of self-similar processes. In particular, the information contained in a self-similar process on a positive or negative neighborhood of the origin is equivalent to the information contained in a stationary process on a semi-infinite interval.

A limitation of this approach is that it does not provide solutions for arbitrary observation intervals. For example, a finite observation interval away from the origin maps to a finite interval in the stationary domain, a case which is not handled by the Wiener-Kolmogorov technique. A finite interval including a neighborhood of the origin corresponds to observing semi-infinite intervals of the future *and* past generators. This can be solved by a spectral factorization technique, but it involves the factorization of the spectral matrix

$$\begin{bmatrix} \phi_F(\omega) & \phi_{FP}(\omega) \\ \phi_{FP}(\omega) & \phi_P(\omega) \end{bmatrix},$$

which is in general much more difficult than factorization of $\phi_F(\omega)$ or $\phi_P(\omega)$ only [Wong, (1971), p. 125].

In the case of a self-similar process with stationary increments, the time-origin no longer plays such a crucial role. Observation of the H-ss si process $Y(t)$ on $[T_0, T_0 + T]$, for example, is equivalent to observation of the H-ss si process $Z(t) = Y(t + T_0) - Y(T_0)$ on $[0, T]$. The case of self-similar, stationary increments processes is examined in detail below.

4 Linear Processing of fBm

In this section, we use the techniques described in the above examples to develop explicit whitening, and innovation, prediction, and interpolation formulae for fractional Brownian motion. The prediction and interpolation formulae obtained are also optimal linear filters for non-Gaussian, self-similar, stationary increments processes, and the whitening formulae hold for non-Gaussian processes in the second-order sense. We first give the spectral factorization of the stationary generators of fBm.

4.1 Stationary Generators of fBm

Throughout the rest of this paper, we adopt the common convention that fBm is normalized with $V_H = E\left\{W_H(1)^2\right\} = 1/\left(\sin \pi H \Gamma(2H+1)\right)$.

The stationary generators $X_F(\tau)$ and $X_P(\tau)$ of fBm were introduced in Yazici and Kashyap (1997). They are statistically identical, with autocorrelation

$$r_F(\tau) = r_P(\tau) = V_H \cosh(H\tau) - (V_H/2)\left|2\sinh(\tau/2)\right|^{2H}$$

and cross-correlation

$$r_{FP}(\tau) = V_H \cosh(H\tau) - (V_H/2)\left(2\cosh(\tau/2)\right)^{2H}.$$

For τ near 0, $r_F(\tau)/V_H \sim 1 - 1/2|\tau|^{2H}$; by virtue of Lemma 3.1 the relation

$$dW_H(e^\tau) = e^{H\tau}\left[dX_F(\tau) + HX_F(\tau)\right]$$

is valid for $H \geq 1/2$.

The corresponding power spectral densities have not been given in previous work, and are derived in Appendix B. They are

(21) $$\phi_F(\omega) = \frac{\Gamma(1-H+i\omega)\,\Gamma(1-H-i\omega)}{\Gamma(1/2+i\omega)\,\Gamma(1/2-i\omega)\,(H^2+\omega^2)}$$

and

(22) $$\phi_{FP}(\omega) = \frac{\cos(\pi H)}{\pi}\frac{\Gamma(1-H+i\omega)\,\Gamma(1-H-i\omega)}{H^2+\omega^2}.$$

The spectrum $\phi_F(\omega)$ satisfies the Paley-Wiener condition and is factorable as $\phi_F(\omega) = \phi^+(\omega)\phi^-(\omega)$ where

(23) $$\phi^+(\omega) = \frac{\Gamma(1-H+i\omega)}{\Gamma(1/2+i\omega)\,(H+i\omega)}$$

When $H > 1/2$, the inverse transform of the causal factor is

$$r^+(t) = \frac{1}{\Gamma(H-1/2)} e^{-Ht} B_{1-e^{-t}}(H-1/2, 1-2H)\, u(t). \tag{24}$$

Here and in the following,

$$B_t(a,b) = \int_0^t u^{a-1}(1-u)^{b-1}\, du, \quad 0 \le t \le 1,\, a > 0$$

denotes the incomplete beta functions. When $H < 1/2$, we get

$$\begin{aligned} r^+(t) &= \frac{1}{\Gamma(H+1/2)} \Big[e^{(H-1)t}(1-e^{-t})^{H-1/2} \\ &\quad + (1/2 - H) e^{-Ht} B_{1-e^{-t}}(H+1/2, 1-2H) \Big]\, u(t). \end{aligned} \tag{25}$$

In the special case $H = 1/2$, we have $r(t) = e^{-|t|/2}$, $r^+(t) = e^{-t/2}u(t)$, and $\phi(\omega) = 1/(1/4 + \omega^2)$.

4.2 Whitening Filter and Innovations Representation of fBm

One of the most common representations for fBm is based on fractional integration of white noise from the infinite past, as given in Mandelbrot and Van Ness (1968). Explicitly, we have the causal filter

$$W_H(t) = \frac{1}{\Gamma(H+1/2)} \int_{-\infty}^{\infty} \left[(t-x)^{H-1/2} u(t-x) - (-x)^{H-1/2} u(-x) \right] dW(x)$$

where $W(x)$ is a Wiener process on $\mathbb{R}$.

Using the method of Section 3.3, It is also possible to represent fBm on a a finite interval $[0,T]$ using a finite interval of a white noise process, as the following theorem demonstrates.

Theorem 4.1 *The fBm $W_H(t)$ on $[0,T]$ with $H > 1/2$ can be obtained from a Brownian motion $W(t)$ on $[0,T]$ as*

$$W_H(t) = \frac{t^H}{\Gamma(H-1/2)} \int_0^t (s/t)^H B_{1-s/t}(H-1/2, 1-2H)\; s^{-1/2} dW(s) \tag{26}$$

Conversely $W(t)$ can be obtained from $W_H(t)$ as

$$\begin{aligned} W(t) &= \frac{1}{\Gamma(3/2-H)} \int_0^t (s/t)^{1/2-H} (t-s)^{1/2-H} dW_H(s) \\ &\quad + \frac{(1/2-H)}{\Gamma(3/2-H)} \int_0^t s^{1/2-H} B_{1-s/t}(3/2-H, 0)\, dW_H(s) \end{aligned} \tag{27}$$

When $H < 1/2$, $W_H(t)$ can be represented as

$$W_H(t) = \frac{1}{\Gamma(H+1/2)} \int_0^t (s/t)^{1/2-H} (t-s)^{H-1/2}\, dW(s) \tag{28}$$
$$+ \frac{(1/2-H)}{\Gamma(H+1/2)} \int_0^t s^{H-1/2} B_{1-s/t}(H+1/2, 1-2H)\, dW(s).$$

Conversely, $W(t)$ can be obtained from $W_H(t)$ as

$$W(t) = \frac{1}{\Gamma(1/2-H)} \int_0^t (s/t)^{-H-1/2} (t-s)^{-H-1/2} W_H(s)\, ds \tag{29}$$
$$+ \frac{(H-1/2)}{\Gamma(1/2-H)} \int_0^t s^{-H-1/2} B_{1-s/t}(1/2-H, 0)\, W_H(s)\, ds$$

Although proving this theorem via Lamperti's transformation is a new technique, most of the theorem has been given in some form in previous work. The case $H > 1/2$ was given in Barton and Poor (1988). In Decreusefond and Üstünel (1999), the authors gave (26) and (28) in the comprehensive form

$$W_H(t) = \int_0^t K_H(t,s) dW(s) \tag{30}$$

where

$$K_H(t,s) = \frac{(t-s)^{H-1/2}}{\Gamma(H+1/2)} F\left(1/2-H, H-1/2; H+1/2; 1-t/s\right) 1_{[0,t)}(s)$$

and $F(\cdot,\cdot,\cdot,\cdot)$ is Gauss' hypergeometric function. The authors use this representation extensively in developing a stochastic calculus for fBm. Coutin and Decreusefond (1999) also shows that (30) can be inverted, although (27) and (29) are not given explicitly. Finally, we note that Molchan and Golosov (1969) gave an analogous representation of fBm in terms of a white process $\eta(t)$ having independent but not stationary increments. Because $\eta(t)$ is $(1-H)$-self-similar, and the spectrum of its generating function is simply $1/((1-H)^2+\omega^2)$, it turns out that Molchan and Golosov's representation can also be easily derived using the method of section 3.3. It is also straightforward, for example, to obtain a representation of $W_H(t)$ in terms of $W_{H'}(t)$ for arbitrary H and H'.

Proof of Theorem 4.1: The innovations representations (26) and (28) are immediate after inserting (24) and (25), respectively, into (20) of Example 3.3.

To obtain the whitening filter (29) for $H < 1/2$, note that

$$F_w(\omega) = \frac{\Gamma(1/2+i\omega)\,(H+i\omega)}{\Gamma(1-H+i\omega)\,(1/2+i\omega)} = \frac{\Gamma(1/2+i\omega)}{\Gamma(1-H+i\omega)}\left[1+\frac{(H-1/2)}{1/2+i\omega}\right].$$

The transform pairs (44) and (45) give the inverse transform $f_w(\tau)$ of (19).

$$f_w(\ln t) = \frac{t^{-1/2}}{\Gamma(1/2-H)}\Big[(1-1/t)^{-H-1/2} + (H-1/2)B_{1-1/t}(1/2-H,0)\Big]\,u(\ln t), \tag{31}$$

which we insert into (19) to obtain (29).

For $H > 1/2$, $F_w(\omega)$ has no inverse transform. As in the discussion following (17) we instead invert the freqency response

$$G_w(\omega) = \frac{\Gamma(1/2+i\omega)}{\Gamma(1-H+i\omega)\,(1/2+i\omega)} = \frac{\Gamma(1/2+i\omega)}{\Gamma(2-H+i\omega)}\left[1+\frac{(1/2-H)}{1/2+i\omega}\right],$$

obtaining

$$g_w(\ln t) = \frac{1}{\Gamma(3/2-H)}\Big[t^{-1/2}(1-1/t)^{1/2-H} + (1/2-H)t^{-1/2}B_{1-1/t}(3/2-H,0)\Big]\,u(\ln t). \tag{32}$$

and hence (27) via (8).

4.3 Prediction of fBm

Before giving prediction formulae for fBm, we note that various aspects of the problem have been treated previously by Jaglom (1955), Grigorev (1965), and Molchan (1969), and independently by Gripenburg and Norros (1996). Of particular interest is Molchan (1969), which treats the derivative of fBm as a generalized random process in the sense of Gel'fand. Molchan gives an elegant formula for prediction of this generalized process, from which the prediction formulae collected here can be derived. Gripenberg and Norros give a clear and complete exposition of prediction for fBm with $H > 1/2$. As far as we are aware, the finite-interval prediction formula for $H < 1/2$ has not been given explicitly before, although Grigorev (1965) gives the essential elements. Theorems 4.2, 4.3, and 4.4 collect these prediction results. We sketch proofs based on Examples 3.1 and 3.2, and highlight some ideas suggested by the framework of Lamperti's transformation. Further results on conditional expectations of functionals of fBm, using the stochastic calculus of variations rather than mean-square integration, are given in Decreusefond and Üstünel (1999).

Theorem 4.2 *Let $Y(t)$ be a fractional Brownian motion with parameter $H \in (0,1)$. For each $a > 0$ and $T_0 \in \mathbb{R}$, the predictor $\hat{Y}(T_0+a) = E\{\, Y(T_0+a) \mid Y(t), t < T_0 \,\}$, can be expressed as a m-s integral*

$$\hat{Y}(T_0+a) = Y(T_0) + \frac{\cos(\pi H)}{\pi} a^{H+1/2} \int_0^\infty \frac{Y(T_0-t) - Y(T_0)}{t^{H+1/2}(t+a)} dt. \tag{33}$$

Alternate forms include, for $H < 1/2$

$$\hat{Y}(T_0+a) = \frac{\cos(\pi H)}{\pi} a^{H+1/2} \int_0^\infty \frac{Y(T_0-t)}{t^{H+1/2}(t+a)} dt \tag{34}$$

and for $H > 1/2$

$$\hat{Y}(T_0+a) = \frac{\cos(\pi H)}{\pi} a^{H-1/2} \int_0^\infty \frac{aY(T_0-t) - (t+a)Y(T_0)}{t^{H+1/2}(t+a)} dt \tag{35}$$

$$= Y(T_0) + \frac{\cos(\pi H)}{\pi} \int_0^\infty \int_0^a \frac{s^{H-1/2}}{s+t} ds \; t^{1/2-H} dY(T_0-t). \tag{36}$$

Theorem 4.3 *Let $Y(t)$ be fBm with $H > 1/2$. For each $c > 0$ and $0 < T < \infty$, and $T_0 \in \mathbb{R}$, the predictor $\hat{Y}(cT+T_0) = E\{\, Y(cT+T_0) \mid Y(t), t \in [T_0-T, T_0] \,\}$, can be expressed as $\hat{Y}(T_0+cT) = Y(T_0) + \int_0^T M_T(c,t) dY(T_0-t)$ where*

$$M_T(c,t) = \frac{\cos(\pi H)}{\pi} \left(\frac{t}{T}\right)^{1/2-H} \left(1 - \frac{t}{T}\right)^{1/2-H} \int_0^c \frac{s^{H-1/2}(s+1)^{H-1/2}}{s+t/T} ds. \tag{37}$$

Theorem 4.4 *Let $Y(t)$ be fBm with $H < 1/2$. For each $c > 0$ and $0 < T < \infty$ and $T_0 \in \mathbb{R}$, the predictor $\hat{Y}(T_0+cT) = E\{\, Y(T_0+cT) \mid Y(t), t \in [T_0-T, T_0] \,\}$, can be expressed as an integral $\hat{Y}(T_0+cT) = \int_0^T m_T(c,t) Y(T_0-t) dt$ where*

$$m_T(c,t) = T^{-1}(t/T)^{-H-1/2}(1-t/T)^{-H-1/2} \times \left[(1/2-H)B_{c/(c+1)}(H+1/2, 1-2H) + \frac{c^{H+1/2}(1+c)^{H-1/2}(1-t/T)}{c+t/T}\right]. \tag{38}$$

Proof: The solutions are obtained by examining the increments process

$$Z(t) = Y(T_0 + t) - Y(T_0),$$

using the equations developed in Examples 3.1 and 3.2.

For the infinite observation interval,

$$F(\omega) = \frac{\phi_{FP}(\omega)}{\phi_P(\omega)} = \frac{\cos(\pi H)}{\pi}\Gamma(1/2 + i\omega)\,\Gamma(1/2 - i\omega) = \cos(\pi H)\text{sech}(\pi\omega)$$

which has [Erdélyi (1954), 1.9.1] inverse transform

$$f(t) = \frac{\cos(\pi H)}{2\pi}\text{sech}(t/2).$$

The prediction kernel is

$$\begin{aligned} k_\infty(a, -t) &= a^{-1}(a/t)^{H+1} f(\ln(a/t)) \\ &= \frac{\cos(\pi H)}{2\pi} \frac{2a^H t^{-H-1}}{(a/t)^{1/2} + (t/a)^{1/2}} \\ &= \frac{\cos(\pi H)}{\pi} \frac{a^{H+1/2} t^{-H-1/2}}{a + t}, \end{aligned}$$

leading to (33). Further simplification is possible, due to the identity

$$\int_0^\infty \frac{ds}{s^\alpha (s + a)} = a^{-\alpha} B(1 - \alpha, \alpha) = \frac{a^{-\alpha}\pi}{\sin \pi\alpha}, \quad 0 < \alpha < 1,$$

which can be used with $\alpha = H \pm 1/2$ to obtain expressions (34) and (35). Equation (36) is obtained from (33) using integration by parts, or directly by inverting $G(\omega) = F(\omega)/(H + i\omega)$.

In proving (4.4) and (4.3), we must invert the frequency responses

$$F_c(\omega) = \frac{\cos(\pi H)}{\pi} \frac{\Gamma(1/2 + i\omega)\,(H + i\omega)}{\Gamma(1 - H + i\omega)} \left[\frac{\Gamma(1 - H + i\omega)\,\Gamma(1/2 - i\omega)}{(H + i\omega)}\right]_+^{\ln c}, \quad H < 1/2$$

and $G_c(\omega) = F_c(\omega)/(H + i\omega)$. Both functions can be put into the form

$$Z(\omega) = (a + i\omega) X(\omega) \left[Y(\omega)\right]_+^\gamma$$

where $x(t) \leftrightarrow X(\omega)$ is causal but not differentiable and $y(t) \leftrightarrow Y(\omega)$ is noncausal and differentiable.

The inverse transform of $Z(\omega)$ is

$$\begin{aligned} z(t) &= ax(t) * y(t)u(t-\gamma) + \frac{d}{dt}\left(x(t) * y(t)u(t-\gamma)\right) \\ &= ax(t) * y(t)u(t-\gamma) + y(\gamma)x(t-\gamma) + x(t) * y'(t)u(t-\gamma) \\ &= y(\gamma)x(t-\gamma) + x(t) * (m(t)u(t-\gamma)) \end{aligned}$$

where $m(t) \leftrightarrow (a+i\omega)Y(\omega)$.

Using the transforms in Appendix C and considerable manipulation of integrals, the inverse transforms turn out to be

$$\begin{aligned} f_c(t)\frac{\pi}{\cos(\pi H)} &= \left[(1/2-H)B_{c/(c+1)}(H+1/2, 1-2H) + \frac{c^{H-1/2}(1+c)^{H-1/2}(1-ce^{-t})}{1+ce^{-t}}\right] \\ &\quad \times c^{1/2-H}e^{-t/2}(1-ce^{-t})^{-H-1/2}u(t-\ln c) \end{aligned}$$

and

$$\begin{aligned} g_c(t)\frac{\pi}{\cos(\pi H)} &= c^{1/2-H}e^{-t/2}(1-ce^{-t})^{1/2-H}\int_0^c u^{H-1/2}(1+u)^{H-3/2}\,du\ u(t-\ln c) \\ &\quad + c^{1/2-H}e^{-t/2}(1-ce^{-t})^{1/2-H}\int_0^c \frac{u^{H-1/2}(1+u)^{H-3/2}}{u+ce^{-t}}du\ u(t-\ln c) \\ &= c^{1/2-H}e^{-t/2}(1-ce^{-t})^{1/2-H}\int_0^c \frac{u^{H-1/2}(1+u)^{H-1/2}}{u+ce^{-t}}du\ u(t-\ln c). \end{aligned}$$

Then (37) is immediate from (8), while (38) follows after writing $\hat{Y}(T_0+cT) = Y(T_0)+\int_0^T m_T(c,t)(Y(T_0-t)-Y(T_0))dt$ and noting that in this case $\int_0^T m_T(c,t)\,dt = 1$.

The impulse responses $f_c(t)$ and $g_c(t)$ are plotted in Figure 1 for $H = 0.2$ and $H = 0.8$, respectively, and various values of c. The prediction time a is fixed at unity, so that $c = 1/T$. Note that as $T \to \infty$, $f_{1/T}(t)$ and $g_{1/T}(t)$ approach the corresponding infinite observation impulse responses, which are plotted as solid lines. The agreement quickly becomes very close for $c < 1$. Figure 2 plots examples of the prediction kernels $m_T(1/T,t)$ and $M_T(1/T,t)$ as dashed lines, and the corresponding limits $k_\infty(1,t)$ and $K_\infty(1,t)$ as solid lines.

When $T = \infty$, the normalized MMSE of prediction $d^2(\cdot)$ can be obtained from (11) in Example 3.1. Evaluation of this expression gives

$$d^2\left(W_H(a)\right) = \frac{\sin(\pi H)\Gamma(2H+1)}{2\pi}\int_{-\infty}^{\infty}\frac{|\Gamma(1/2+i\omega)|^2}{|\Gamma(1+H+i\omega)|^2}d\omega$$

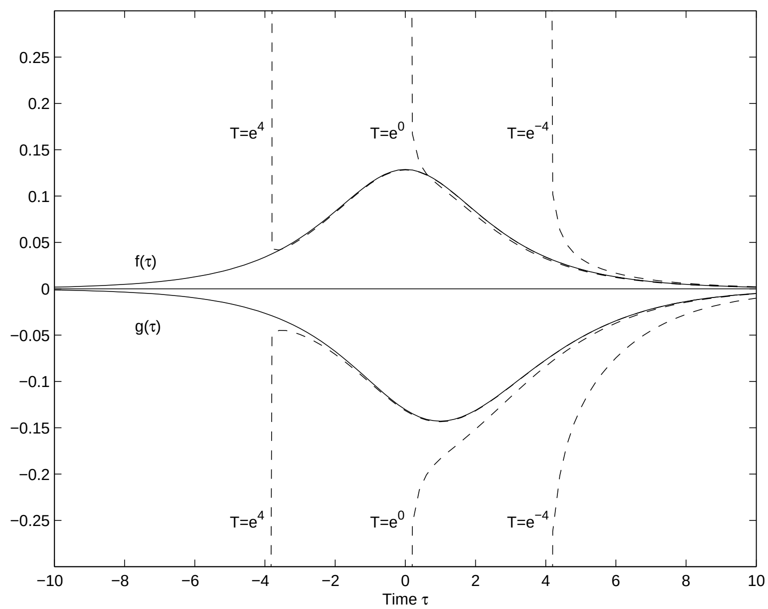

Figure 1: Prediction filter impulse responses for the stationary generator of fBm. The positive dashed functions are $f_c(-\tau)$ for $H = 0.2$, $T = c^{-1} = e^{-4}, e^{0}, e^{4}$. The negative dashed functions are $g_c(-\tau)$ for $H = 0.8$, $T = c^{-1} = e^{-4}, e^{0}, e^{4}$. The solid functions are the corresponding infinite observation responses.

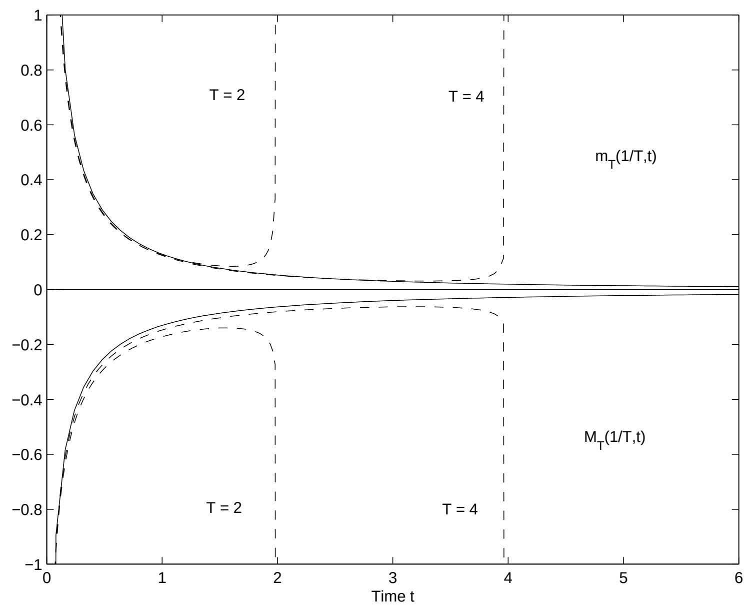

Figure 2: Prediction kernels for unit prediction of fBm. The positive dashed functions depict $m_T(1/T, t)$ for $H = 0.2$, and the negative dashed functions depict $M_T(1/T, t)$ for $H = 0.8$. The solid functions are the corresponding infinite observation kernels $k_\infty(1, t)$ and $K_\infty(1, t)$.

$$
\begin{aligned}
&= \frac{\sin(\pi H)\Gamma(2H)}{\Gamma(1/2+H)^2} \\
&= \frac{\sin(\pi(H-1/2))\Gamma(3/2-H)^2}{\pi(H-1/2)\Gamma(2-2H)} \\
&= \frac{\Gamma(3/2-H)}{\Gamma(2-2H)\,\Gamma(H+1/2)},
\end{aligned}
$$

where the integral is evaluated using [Gradshteyn and Ryzhik (1965), 6.413.2]. The last two equivalent forms were given in Gripenberg and Norros (1996) and Jaglom (1955), respectively.

The normalized MMSE $d^2\left(W_H(a)\right)$ is plotted as the solid line in Figure 3. When $H = 1/2$, of course, prediction is impossible. As fBm becomes increasingly persistent, $H \to 1$, the relative error goes to zero, albeit slowly. Prediction also improves as fBm becomes more anti-persistent; however, in the limit $H \to 0$, observation of the infinite past can only account for half of the variance in the fBm.

4.4 Discrete Prediction of fBm

In practice, discrete samples of fBm would likely be used for prediction. The above discussion of continuous-time prediction sheds light on the discrete problem by providing a lower bound on the achievable MMSE, and by suggesting which discrete samples would be most relevant to a given prediction problem. It turns out that it is possible to approach the lower bound with a small, finite number of discrete samples.

Optimal prediction of $X = W_H(a)$ based on the jointly Gaussian n-vector of observations

$$
Y = \left[\begin{array}{ccc} W_H(t_1) & \cdots & W_H(t_n) \end{array}\right]^T
$$

is solved by the standard formula

$$
\hat{X} = E\{\, X \mid Y = y \,\} = \Sigma_{XY}\Sigma_Y^{-1} y, \tag{39}
$$

with normalized MMSE of prediction given by

$$
d^2\,(X) = 1 - \frac{\Sigma_{XY}\Sigma_Y^{-1}\Sigma_{XY}^T}{\Sigma_X},
$$

where $\Sigma_Y = E\left\{YY^T\right\}$, $\Sigma_X = E\left\{X^2\right\}$, and $\Sigma_{XY} = E\left\{XY^T\right\}$. As the self-similarity of the fBm immediately results in the scale-invariance of the normalized MMSE, we restrict attention to the case of unit prediction, $a = 1$.

If we have the freedom to choose the sampling instants $\{t_i\} \subset I$, where I is some subset of the real line not including a, then we are faced with the problem of maximizing the term $m(\mathbf{t}) \stackrel{\text{def}}{=} \Sigma_{XY} \Sigma_Y^{-1} \Sigma_{YX}^T$, considered as a function of $\mathbf{t} \in I^n$, which can be done by numerical techniques. However, insight into appropriate sampling designs can be obtained by examining the prediction filter in the stationary domain.

Consider the case $I = \{t < 0\}$, where we predict $X_F(\alpha)$ from the entire evolution of $X_P(\tau)$, $t \in \mathbb{R}$. The optimal filter $f(t)$, shown as a solid line in Figure 1, integrates the observation weighted by a hyperbolic secant centered at $X_P(\alpha)$. We clearly see that $X_P(\alpha)$ and nearby observations are of greatest importance in predicting $X_F(\alpha)$; other observations are exponentially damped. The problem is symmetric forward and backward in time in the stationary domain. Note that this symmetry and the special weight given to observations near α are not directly apparent in Figure 2.

Mentally returning to the self-similar domain, we see that $W_H(-a)$ and its logarithmic neighborhood carry the most information about $W_H(a)$, given that all observation times are negative. Equivalently, for observations on $[T_0 - T, T_0]$, the most crucial observation is $W_H(T_0)$. Conditioned on that observation, $W_H(T_0 - a)$ carries the most information about $W_H(T_0 + a)$. This sheds further light on the observation made in Gripenberg and Norros (1996) that roughly speaking, only the last second is needed to predict ahead one second, the last minute to predict the next minute and so on.

Fix $a = 1$ and $I = \{t < 0\}$. If we are allowed only one observation, then it is easy to show that, as expected, the most effective sample is $W_H(-1) = X_P(0)$. The optimizing function in this case is a function of one variable $m(e^\tau) = r_{FP}(\tau)^2$, which has a unique maximum at $\tau = 0$. For multiple observations, the optimum sampling designs depend on H and are difficult to obtain analytically. When the number of observations is sufficiently small, it is a simple matter to numerically seek sampling designs that provide very good prediction. Although the optimal sampling designs depend on H, the dependence on H is not very strong, and suboptimal fixed designs appear to work quite well.

As one would expect in the case of three samples, the best designs in the stationary domain are symmetric and centered at the peak of the corresponding continuous filter impulse response. Back in the ordinary time domain, the samples are spaced

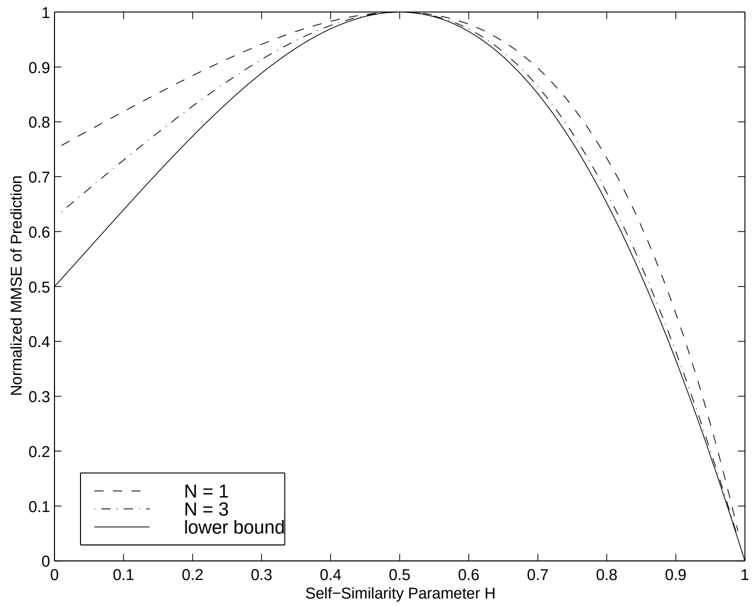

Figure 3: Normalized Minimum Mean-Square Error $d^2\left(W_H(a)\right)$ for prediction of fBm at time $a > 0$, as a function of H. The solid line is obtained using the entire observation interval $(\infty, 0]$, while the dashed lines are obtained with N discrete samples.

exponentially, and geometrically centered about the prediction time. In Figure 3 we compare the normalized MMSE of prediction obtainable by discrete prediction with the continuous prediction lower bound. With just one sample, much of the work is already done. Adding two more samples brings the normalized error very close to its lower bound, particularly for $H > 1/2$. Prediction of fBm is difficult, in the sense that the lower bound is quite high for most values of H. On the other hand, prediction is easy, in the sense that we can achieve nearly optimal performance with extremely low complexity.

4.5 Interpolation of fBm

In addition to being self-similar, the covariance structure of fBm is also time-reversible, so that fBm satisfies $\{W_H(at)\}_{t\in\mathbf{R}} = |a|^H\{W_H(t)\}_{t\in\mathbf{R}}$ for any non-zero a. Equivalently, the cross-correlation between the future and past generators satisfies the symmetry $r_{FP}(t) = r_{FP}(-t)$. When this condition is satisfied, it turns out that interpolation of a self-similar random process on $(0, T)$ based on observations in $\mathbb{R}\backslash(0, \mathrm{T})$ is equivalent to prediction on $(1/T, \infty)$, $T > 0$, based on observations of $(-\infty, 1/T]$. In the stationary domain, the first problem amounts to predicting the future of $X_F(\tau)$ based on the infinite past of $X_F(\tau)$ and the entire evolution of $X_P(\tau)$. The second problem is just the time-reversal of the first, in which we observe $X_P(\tau)$ and the future of $X_F(\tau)$ and predict the past of $X_F(\tau)$. When $r_{FP}(t)$ is symmetric, these linear prediction problems are identical.

Application of these ideas yields interpolation formulae for fBm.

Theorem 4.5 *Let $Y(t)$ be an fBm with parameter $0 < H < 1$. For each $0 < b < T$ the estimate $\hat{Y}(b) = E\{\, Y(b) \mid Y(t), t \in \mathbb{R}\backslash(0, \mathrm{T}) \,\}$ can be expressed by m-s integrals as described below. When $H > 1/2$,*

$$\hat{Y}(b) = (b/T)^{2H} Y(T) + (b/T)^{2H} \frac{\cos(\pi H)}{\pi} \times$$

$$\int_{\mathbb{R}\backslash(0,\mathrm{T})} \left| \frac{t}{T} \right|^{1/2-H} \left| 1 - \frac{t}{T} \right|^{1/2-H} \int_0^{1-b/T} \frac{u^{H-1/2}(1-u)^{-H-1/2}}{|u + t/T - 1|} \, du \, dY(t). \tag{40}$$

When $H < 1/2$,

$$\hat{Y}(b) = \frac{\cos(\pi H)}{\pi} b^{H+1/2} (T-b)^{H+1/2} \int_{\mathbb{R}\backslash(0,\mathrm{T})} \frac{|t|^{-H-1/2} |T-t|^{-H-1/2}}{|t-b|} Y(t) \, dt, \tag{41}$$

and when $H = 1/2$,

$$\hat{Y}(b) = (b/T) Y(T). \tag{42}$$

Proof:

For $H < 1/2$, we set $b = a + T$ and rewrite (34) in terms of self-similar convolutions

$$\hat{Y}(b) = b^H \int_0^\infty f_{b/T}(\ln b/t) t^{-H-1} Y(t) \, dt + b^H \int_0^\infty g_{b/T}(\ln b/t) t^{-H-1} Y(-t) \, dt.$$

Letting $\beta = \ln b$ and $\tau = \ln T$, we have equivalently

$$\hat{X}_F(\beta) = f_{e^{\beta-\tau}}(v) * X_F(v)|_\beta + g_{e^{\beta-\tau}}(v) * X_P(v)|_\beta$$

for the best estimate of $X_F(\beta)$ from $\{X_P(v), v \in \mathbb{R}\}$ and $\{X_F(v), v \leq \tau\}$. Then by the symmetry of $r_F P(t)$, $r_F(t)$, and $r_P(t)$ the predictor of $X_F(\beta)$ from $\{X_P(v), v \in \mathbb{R}\}$

and $\{X_F(v), v \geq \tau\}$ is

$$\hat{X}_F(\beta) = f_{e^{-\beta+\tau}}(-v) * X_F(v)|_\beta + g_{e^{-\beta+\tau}}(-v) * X_P(v)|_\beta.$$

Now applying L_H^{-1} gives

$$\hat{Y}(b) = b^H \int_0^\infty f_{T/b}(\ln t/b) t^{-H-1} Y(t)\, dt + b^H \int_0^\infty g_{T/b}(\ln t/b) t^{-H-1} Y(-t)\, dt.$$

as the predictor of $Y(b)$ from $\mathbb{R}\backslash(0, \mathrm{T})$. Inspection of (34) shows that

$$f_y(\ln x) = \frac{c_H x^{1/2}(y-1)^{H+1/2}}{(x-y)^{H+1/2}(x-1)} u(x-y), \quad y > 1$$

and

$$g_y(\ln x) = \frac{c_H x^{1/2}(y-1)^{H+1/2}}{(x+y)^{H+1/2}(x+1)}$$

where $c_H = \cos(\pi H)/\pi$, from which (41) can be determined.

The case $H > 1/2$ is proved similarly, using (36), and the trivial case $H = 1/2$ is included for comparison.

The optimal mean-squared error of interpolation can also be deduced directly from the prediction results. For prediction from observations up to time T, recall that $D^2(W_H(b)) = M_H(b-T)^{2H}$ where $M_H = \Gamma(3/2-H)/(\Gamma(2-2H)\Gamma(H+1/2))$. Then for $\beta > \tau$ the prediction error for $X_F(\beta)$ given $\{X_F(t), t \leq \tau\}$ and $\{X_P(t), t \in \mathbb{R}\}$ is $D^2(X_F(\beta)) = M_H(1-e^{\tau-\beta})^{2H}$. The time-reversed case, with $\beta < \tau$ gives $D^2(X_F(\beta)) = M_H(1-e^{\beta-\tau})^{2H}$, so that

$$D^2(W_H(b)) = M_H\left(\frac{b(T-b)}{T}\right)^{2H}$$

when the observation set is $\mathbb{R}\backslash(0, \mathrm{T})$.

The normalized interpolation error $D^2(W_H(b))/b^{2H}$ depends on b and T only through their ratio, and is plotted for two values of H in Figure 4. For any value of H, the normalized error begins at M_H and decreases with b/T, although the drop is much swifter for larger values of H. The dashed line in the figure depict the interpolation error when only the endpoints $W_H(0)$ and $W_H(T)$ are observed. Such discrete observations come close to the lower bound when b/T is not too small. In the case of the Weiner process ($H = 1/2$), the solid and dashed lines coincide in the line $1 - b/T$.

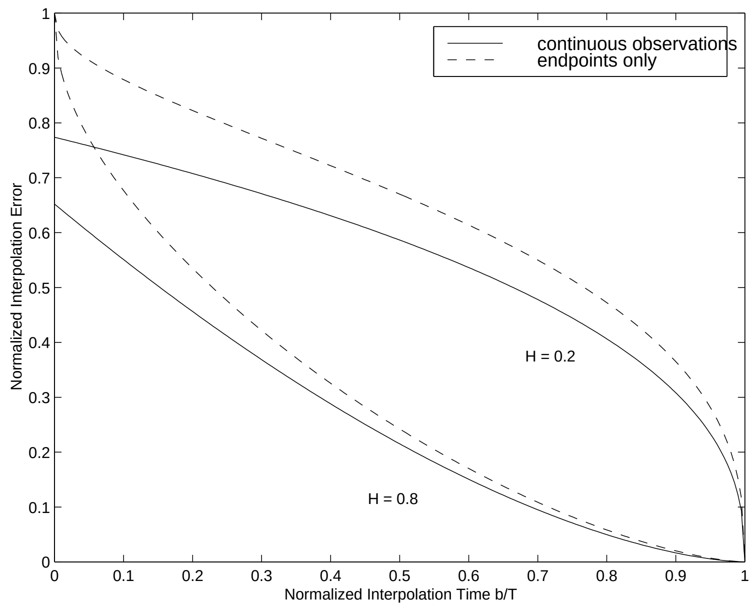

Figure 4: Normalized Minimum Mean-Square Error $d^2\left(W_H(b)\right)$ for interpolation of fBm at time $0 < b < T$. The solid lines are obtained using the entire observation interval $\mathbb{R}\backslash(0,\mathrm{T})$, while the dashed lines are obtained using only $W_H(0)$ and $W_H(T)$.

5 Conclusions

This paper demonstrates the utility of Lamperti's transformation for linear estimation of continuous-time self-similar processes, and for fractional Brownian motion in particular. This approach has the advantage of making available well-known tools and results developed for shift-stationary processes. The approach also naturally brings out many consequences of self-similarity, including self-similarity of the prediction kernels, scale-invariance of the normalized MMSE, and some underlying symmetries.

We expect that are many others problems involving self-similar processes for which Lamperti's transformation could play a useful role. For example, work is currently in progress which uses the approach of this paper to characterize the reproducing kernel Hilbert spaces associated with self-similar random processes.

A Lemmas

Before proving Lemma 3.1, we give sufficient conditions for the product rule of differentiation to hold for mean-square Stieltjes integrals.

Lemma A.1 *Let $G(t)$, $W(t)$, and $Z(t)$ be mutually independent second-order random processes on a finite interval $I = [a,b]$, and let $Y(t) = Z(t)W(t)$. Suppose further that $E\left\{G(t)^2\right\} \leq M$ on I, and that $W(t)$ and $Z(t)$ satisfy the Lipschitz conditions*

$$E\left\{(W(t) - W(s))^2\right\} \leq L|t-s|^{\alpha}$$

and

$$E\left\{(Z(t) - Z(s))^2\right\} \leq L|t-s|^{\beta}$$

for some $L > 0$, $\alpha \geq 1$, and $\beta > 1$, and for all $t, s \in I$. Then

$$\int_I G(t)dY(t) = \int_I G(t)W(t)dZ(t) + \int_I G(t)Z(t)dW(t) \tag{43}$$

whenever any two of the above mean-square integrals exist.

Proof: Let $T_n = \{t_i\}$ be a partition of $[a,b]$. Replacing each of the integrals in (43) with the a finite sum over T_n, the difference between the left and right sides of (43) is

$$R_n = \sum_i G(t_i)\left[W(t_i) - W(t_{i-1})\right]\left[Z(t_i) - Z(t_{i-1})\right].$$

Using the Schwarz inequality, and independence of G,W, and Z, we have

$$E\left\{R_n^2\right\} \leq M\left(\sum_i E\left\{\left[W(t_i) - W(t_{i-1})\right]^2\right\}\right)\left(\sum_i E\left\{\left[Z(t_i) - Z(t_{i-1})\right]^2\right\}\right).$$

The Lipschitz conditions ensure that as the maximum partition interval goes to zero, the sum involving $W(t)$ is bounded and the sum involving $Z(t)$ vanishes, so that $E\left\{R_n^2\right\} \to 0$.

Proof of Lemma 3.1: The change of variables $\tau = \ln t$ gives

$$\int_a^b g(t)dY(t) = \int_{\ln a}^{\ln b} g(e^\tau)dY(e^\tau).$$

We apply Lemma A.1 with $G(\tau) = g(e^\tau)$, $Z(\tau) = e^{H\tau}$, and $W(\tau) = X(\tau)$. The differentiability of $Z(\tau)$ ensures that it satisfies the Lipschitz condition with $\beta = 2$, and since $E\left\{(X(t) - X(s))^2\right\} = 2\left(r_X(0) - r_X(t-s)\right)$, the condition on $W(t)$ is satisfied for $|t-s| < \epsilon$. Since $r_F(0) - r_X(\tau) \leq 2r_X(0)$, choosing $L' = 2\max(L, 2r_X(0)/\epsilon)$ ensures that the condition is satisfied for all t and s.

B Spectra of the Generators of fBm

We wish to find the Fourier transforms of

$$r_F(\tau) = V_H \cosh(H\tau) - (V_H/2)\left|2\sinh(\tau/2)\right|^{2H}$$

and

$$r_{FP}(\tau) = V_H \cosh(H\tau) - (V_H/2)\left(2\cosh(\tau/2)\right)^{2H},$$

and the spectral factorization of $r_F(\tau)$, when $V_H^{-1} = \sin(\pi H)\Gamma(2H+1)$. Both $r_F(t)$ and $r_{FP}(t)$ are bounded above and are $O(e^{-\gamma|t|})$ for large $|t|$, where $\gamma = \min(H, 1-H) > 0$. Thus the region of convergence for the corresponding two-sided Laplace transforms is $-\gamma < s < \gamma$.

We denote the two-sided Laplace transform of $r_F(t)$ by $\phi_F(s)$ and the one-sided transform by $\phi_F^1(s)$. Taking the one-sided Laplace transform of both components of

$r_F(t)$ gives (see Gradshteyn and Ryzhik (1965), 3.541.1):

$$\phi_F^1(s)/V_H = \frac{s}{s^2 - H^2} - \frac{1}{2}B(s-H, 2H+1)\,, \quad s > H$$

which extends analytically to the region of convergence $s > -\gamma$.

The two-sided Laplace transform on $-\gamma < s < \gamma$ is

$$\begin{aligned}
\phi_F(s) &= \phi_F^1(s) + \phi_F^1(-s) \\
&= -(V_H/2)\,[B(s-H, 2H+1) + B(-s-H, 2H+1)] \\
&= \frac{\Gamma(1-H+s)\,\Gamma(1-H-s)}{\Gamma(1/2+s)\,\Gamma(1/2-s)\,(H^2-s^2)}.
\end{aligned}$$

where $B(a,b) = \Gamma(a)\,\Gamma(b)\,/\Gamma(a+b)$ and we have made repeated used of common gamma function identities.

To calculate the cross-spectrum, we first express the one-sided Laplace transform of $g(t) = (2\cosh t/2)^{2H}$ as

$$\begin{aligned}
\hat{g}^1(s) &= \int_0^\infty (2\cosh t/2)^{2H} e^{-st}\,dt \\
&= \int_0^1 (1+\tau)^{2H}\tau^{s-H-1}\,d\tau.
\end{aligned}$$

Then the two-sided transform on $-\gamma < s < \gamma$ is

$$\begin{aligned}
\phi_{FP}(s) &= \phi_{FP}^1(s) + \phi_{FP}^1(-s) \\
&= -\frac{V_H}{2}\int_0^1 (1+t)^{2H} t^{s-H-1}\,dt - \frac{V_H}{2}\int_0^1 (1+t)^{2H} t^{-s-H-1}\,dt \\
&= -\frac{V_H}{2}\int_0^1 (1+t)^{2H} t^{s-H-1}\,dt - \frac{V_H}{2}\int_1^\infty (1+\tau)^{2H} \tau^{s-H-1}\,d\tau \\
&= -(V_H/2)B(s-H, -s-H) \\
&= \frac{\cos(\pi H)\Gamma(1-H+s)\,\Gamma(1-H-s)}{\pi(H^2-s^2)}
\end{aligned}$$

whence (22).

The spectral factor

$$\phi_F^+(\omega) = \frac{\Gamma(1-H+i\omega)}{\Gamma(1/2+i\omega)\,(H+i\omega)} = \frac{\Gamma(1-H+i\omega)}{\Gamma(3/2+i\omega)}\left[1 + \frac{1/2-H}{H+i\omega}\right]$$

clearly satisfies $\phi_F(\omega) = |\phi_F^+(\omega)|^2$. The transform pairs in Appendix C give the the causal inverse transforms (24) and (25).

C Useful Transform Pairs

The following Fourier transform pairs occur frequently while working with the stationary generators of fBm. In each case, we require $c > a > 0$ and $b > 0$. Note that Mellin transform pairs are obained by the change of variables $x = e^{-t}$.

$$\frac{\Gamma(a+i\omega)}{\Gamma(c+i\omega)} \iff \frac{1}{\Gamma(c-a)} e^{-at}(1-e^{-t})^{c-a-1} u(t) \tag{44}$$

$$\frac{\Gamma(a+i\omega)}{\Gamma(c+i\omega)\,(b+i\omega)} \iff \frac{e^{-bt}}{\Gamma(c-a)} B_{1-e^{-t}}(c-a, a-b)\, u(t) \tag{45}$$

$$\Gamma(a+j\omega)\,\Gamma(b-j\omega) \iff \Gamma(a+b)\, e^{-at}(1+e^{-t})^{-a-b} \tag{46}$$

Both (44) and (46) follow from the usual integral representation of the beta function [Gradshteyn and Ryzhik (1965), 8.380.1]. The right side of (45) is the convolution of the right side of (44) by $e^{-bt}u(t)$.

REFERENCES

ABRAMOWITZ, M. A. AND STEGUN, I. A. (1970) *Handbook of Mathematical Functions.* Dover. New York.

ALBIN, J. M. P. (1998) On extremal theory for self-similar processes. *Ann. Prob.* **26,** 743–793.

BARTON, R. J. AND POOR, H. V. (1988) Signal detection in fractional Gaussian noise. *IEEE Trans. Inf. Theory.* **34,** 943–959.

BURNECKI, K., MAEJIMA, M., AND WERON, A. (1997) The Lamperti transformation for self-similar processes. *Yokohama Math. J.* **44,** 25–42.

COUTIN, L. AND DECREUSEFOND, L. Non-linear filtering theory in presence of fractional Brownian motion. *to appear, Annals of Appl. Prob.*

DECREUSEFOND, L. AND ÜSTÜNEL, A. S. Stochastic analysis of the fractional Brownian motion. *Potential Analysis.* **10,** 177–214.

ERDÉRLYI, A. ED. (1954) *Tables of Integral Transforms.* **1** McGraw-Hill. New York.

GRADSHTEYN, I. S. AND RYZHIK, I. M. (1965) *Table of Integrals, Series, and Products.* trans. ed. A. Jeffrey. Academic Press. New York.

GRIGOR'EV, S. V. (1965) Explicit extrapolation formulas for some probability processes. *Kazan. Gos. Univ. Učen. Zap.* **125,** 106–109. Engl. trans. in *Selected Transl. in Math. Statist. and Prob.* **11,** (1973)

GRIPENBERG, G. AND NORROS, I. (1996) On the prediction of fractional Brownian motion. *J. Appl. Prob.* **33,** 400–410.

HURST, H. E. (1951) Long-term Storage Capacity of Reservoirs. *Trans. Amer. Soc. Civil Eng.* **116** 770-808.

JAGLOM, A. M. (1955) Correlation theory of processes with random stationary n-th increments. *Matematika Sbornik.* **37,** 141–196. Engl. trans. in *American Mathematical Society Transl., Series 2* **8,** (1958)

LAMPERTI, J. (1962) Semi-stable stochastic processes. *Trans. Amer. Math. Soc.* **104,** 62–78.

LELAND, W. E., TAQQU, M. S., WILLINGER, W. AND WILSON, D. V. (1994) On the self-similar nature of Ethernet traffic (extended version) *IEEE/ACM*

Trans. Network. **2,** 1–14.

LOÈVE, M. (1955) *Probability Theory.* Van Nostrand. New York.

MANDELBROT, B. B. AND VAN NESS, J. (1968) Fractional Brownian motions, fractional noises, and applications. *SIAM Rev.* **10,** 422–437.

MANDELBROT, B. B. (1982) *The Fractal Geometry of Nature.* W. H. Freeman and Co. San Francisco.

MOLCHAN, G. M. AND GOLOSOV, JU. I. (1969) Gaussian stationary processes with asymptotic power spectrum. *Soviet Mathematics Doklady.* **10,** 134–137.

MOLCHAN, G. M. (1969) Gaussian processes with spectra which are asymptotically equivalent to a power of λ. *Theory of Prob. and its Applications.* **14,** 530–532.

POOR, H. V. (1994) *An Introduction to Signal Detection and Estimation.* Springer-Verlag. New York.

PROTTER, M. H. AND MORREY, C. B. (1977) *A First Course in Real Analysis.* Springer-Verlag. New York.

TAQQU, M. S. (1985) A bibliographic guide to self-similar processes and long-range dependence. *Dependence in Probability and Statistics.* Birkhäuser. Boston.

VERVAAT, W. (1987) Properties of general self-similar processes. *Bull. Int. Stat. Inst.* **52,** 199–216.

WILLINGER, W., TAQQU, M. S., AND TEVEROVSKY, V. (1999) Stock Market Prices and Long-Range Dependence. *Finance and Stochastics.* **3,** 1–13.

WONG, E. (1971) *Stochastic Processes in Information and Dynamical Systems.* McGraw-Hill. New York.

WORNELL, G. (1991) *Synthesis, Analysis, and Processing of Fractal Signals* Ph.D. Thesis. M.I.T. Cambridge, Mass.

YAZICI, B. AND KASHYAP, R. L. (1997) A class of second-order stationary self-similar processes for 1/f phenomena. *IEEE Trans. Signal Proc.* **45,** 396–410.