Advanced Review

Fractal (fractional) Brownian motion

Winston C. Chow*

Fractal Brownian motion, also called fractional Brownian motion (fBm), is a class of stochastic processes characterized by a single parameter called the Hurst parameter, which is a real number between zero and one. fBm becomes ordinary standard Brownian motion when the parameter has the value of one-half. In this manner, it generalizes ordinary standard Brownian motion. Here, we precisely define fBm, compare it with Brownian motion, and describe its unique mathematical and statistical properties, including fractal behavior. Ideas of how such properties make these stochastic processes useful models of natural or man-made systems in life are described. We show how to use these processes as unique random noise representations in state equation models of some systems. We finally present statistical state equation estimation techniques where such processes replace traditional Gaussian white noises.

WIREs Comp Stat 2011 3 149–162 DOI: 10.1002/wics.142

Keywords: almost surely; Bayesian estimation; Brownian motion; fractal; fractal Brownian motion; fractional Brownian motion; fractal dimension; fractal Gaussian noise; fractional Gaussian noise; functional; Hausdorff-Besicovitch dimension; Hurst parameter; Kalman filter; loss function; measurable; quadratic mean integral; sample path integral; self-similar; wide-sense stationary

INTRODUCTION

Because of their unique properties, Brownian motion (BM) and its associated Gaussian white noise are traditionally used to model random phenomena. Two of the important properties are that BM is a Gaussian process and has independent increments. Some natural or man-made phenomena exhibit random properties with long-term time dependencies that do not quickly decouple. Such behavior does not fit independent increment processes such a BM. Fractal Brownian motion (fBm) is a stochastic process having long-term dependencies and yet still is Gaussian making it a better mathematical and statistical model than BM for some applications. The fBm stochastic process is the topic of this article.

We define both BM and fBm and explain that the latter actually generalizes the former process. We describe the main properties of fBm. These are the properties that make it a valuable applications tool. We then give examples of applications to help readers get a feel for how this process is useful in the real world. Fractal Brownian motion is most often also called 'fractional Brownian motion' and abbreviated fBm.

In this article, let (Ω, A, P) be the probability space with sample set Ω, σ-algebra A, and probability measure P. So when we speak of a stochastic process sample path, say $X(t)$ (or $X(t,\omega)$) for $t \in U \equiv (-\infty, \infty)$ or $[0,\infty)$, we mean $\{X(t) \equiv X(t,\omega) : t \in U,\ \omega \in \Omega\}$ where $X(t)$ is measurable in A. Notice a stochastic process is a real valued function of two sets, but we only require it to be measurable in the σ-algebra of 1 of them, namely A, as in ordinary random variables. So we define as follows here.

Definition 1 (Ref 1, pp. 44–45) A stochastic process is called *measurable*, if in addition to A measurability (as ordinary random variables), it is also measurable in the σ-algebra generated by Cartesian products of the Lebesgue σ-algebra of subsets of time domain U and σ-algebra A sets with U Lebesgue measurable in $(-\infty, \infty)$.

Stochastic processes can be legitimate but not strictly measurable by Definition 1. Such processes lack some properties. fBm is a measurable stochastic process as will be validated.

BROWNIAN MOTION

We first describe standard BM.

*Correspondence to: winstonchow@gmualumni.org

C/o Dr. Edward J. Wegman, Department of Computational and Data Sciences MS 6A2, George Mason University, 4400 University Drive, Fairfax, VA, USA

DOI: 10.1002/wics.142

Definition 2 (Ref 1) A standard BM $\{B(t) : t \in U\}$ in space (Ω, A, P) such that $U \equiv (-\infty, \infty)$ or $[0,\infty)$ is a Gaussian stochastic process for which $E[B(t)] = 0$ and $E[B(t)B(u)] = (1/2)(|t| + |u| - |t - u|)$ for all $t, u \in U$.

BM can be shown to be not only continuous in probability but also equivalent almost surely (a.s.), that is, with probability 1, to a BM that is continuous (Ref 1, Section 2.2–2.4).

Theorem 1 (Ref 1, Section 2.2–2.4) *BM is equivalent with probability 1 (a.s.) to BM with continuous sample paths on any finite interval.*

For the case of $U \equiv [0, \infty)$, the covariance given in Definition 2 is equivalent to $E[B(t)B(u)] \equiv \min(t, u)$ for all $t, s \in U \equiv [0, \infty)$, which also implies that $\mathrm{var}(B(t)) = t$. Regardless of which U set we choose, this set is referred as 'time' parameters, as it often represents time in real life.

It can be shown B has independent increments, that is, for any t_1, t_2, t_3, $t_4 \in U$ where $t_1 < t_2 < t_3 < t_4$, $B(t_2) - B(t_1)$ and $B(t_4) - B(t_3)$ are stochastically independent. Also, for all $t_1, t_2 \in U$ such that $t_2 \geq t_1$, $B(t_2) - B(t_1)$ is 0 mean Gaussian and $E\{(B(t_2) - B(t_1))^2\} = t_2 - t_1$. Knowing $E[B(t)] = 0$ and $E[B(t)B(u)] = (1/2)(|t| + |u| - |t - u|)$, we also have $B(0) = 0$ a.s.

FRACTIONAL (OR FRACTAL) BROWNIAN MOTION

fBm is defined as a transformation of ordinary BM.

Definition 3 (Refs 2,4) fBm

$$
\begin{aligned}
B_H(t) &= \frac{1}{\Gamma(H+\frac{1}{2})}\left[\int_{-\infty}^{0}\left(|t-s|^{H-\frac{1}{2}} - |s|^{H-\frac{1}{2}}\right)\mathrm{d}B(s)\right.\\
&\quad \left.+ \int_0^t |t-s|^{H-\frac{1}{2}}\mathrm{d}B(s)\right]\\
&\equiv \frac{1}{\Gamma(H+\frac{1}{2})}\left(\int_{-\infty}^{0} M_1(t,s)\mathrm{d}B(s)\right.\\
&\quad \left.+ \int_0^t \mathrm{M}_2(\mathrm{t},\mathrm{s})\mathrm{d}B(s)\right),
\end{aligned}
$$

where $M_1(t,s) \equiv |t-s|^{H-\frac{1}{2}} - |s|^{H-\frac{1}{2}}$ and $\mathrm{M}_2(t,s) \equiv |t-s|^{H-\frac{1}{2}}$, B is BM, and Γ is the Gamma function.

We note

$$\int_0^t M_2(t,s)\mathrm{d}B(s) = -\int_t^0 M_2(t,s)\mathrm{d}B(s),$$

accounting $t < 0$, and will later explain $B_H(0) = 0$ a.s. fBm's dimensionless parameter H, known as the *Hurst parameter*, is such that $0 < H < 1$. fBm generalizes BM because it becomes BM for $H = 1/2$. Stochastic integrals of deterministic functions with respect to stochastic processes are defined to be mean square (q.m.) limits of sequences of sums. For given $\omega \in \Omega$, fBm consists of such Stieltjes type integrals in the q.m. sense, but modified because the first integral is improper and the second is improper for $H - 1/2 < 0$. We first claim these integrals do converge and demonstrate this in the Appendix.

Theorem 2 *The integrals defining fBm exist in the mean square sense and fBm is a measurable stochastic process.*

Because such integrals are limits of sums of products of deterministic step functions and time increments of B, which is 0 mean Gaussian, fBm is also 0 mean Gaussian. Now we give the covariance formula.

Theorem 3 (Ref 2)

$$
\begin{aligned}
\mathrm{Cov}(B_H(t), B_H(s)) &= \frac{V_H}{2}(|t|^{2H} + |s|^{2H} - |t-s|^{2H});\\
V_H &\equiv \frac{-\Gamma(2-2H)\cos(\pi H)}{\pi H(2H-1)}\\
&= \mathrm{var}\ B_H(1).
\end{aligned}
$$

Theorem 4 *fBm has wide-sense stationary increments.*

Using the covariance formula and the fact that fBm has stationary increments, we show fBm behaves as a self-similar process. We first define self-similarity and then show fBm is a special case of it with proof in the Appendix.

Definition 4 (Ref 4) A process X is *self-similar* if there exists $H \geq 0$ such that for all $h > 0$, $X(t_0 + ht) - X(t_0) \cong h^H\{X(t_0 + t) - X(t_0)\}$ where $\cong$ designates equal finite distributions.

Theorem 5 $B_H(t_0 + ht) - B_H(t_0) \cong h^H B_H(t)$.

This makes fBm statistical fractals because time scaling of them have equal finite distributions. Thus fBm is also called fractal Brownian motion. Also, this fractal property makes fBm a.s. nondifferentiable. There is a proof in the Appendix that is simplified from the one given in Ref 4.

Theorem 6 *fBm is a.s. nondifferentiable.*

In studying applications of fBm, we prefer to look at stochastic integrals as a.s. sample path integrals instead of quadratic mean (q.m.) ones, if sample path versions exist, because a.s. sample path integrals seem easier to interpret as models of physical events. We show in the Appendix this does work on fBm.

Theorem 7 *If a stochastic processes sequence* $\{X_k : k = 1, 2, \ldots\}$ *converges in both in q.m. to* X *and a.s. to* Y*, then* $X = Y$ *a.s.*

Corollary 1 *The integral in the definition of fBm can be represented a.s. as a sample path Riemann–Stieltjes integral and thus a.s. equal to its q.m. integral.*

Thus, fBm can a.s. be defined as both a q.m. and a sample path integral and they are a.s. equal.

Lemma 1 *Let* $a, b \in [-\infty, \infty], \alpha, \beta, c \in (-\infty, \infty)$*, and consider functions* $f, g, g_1, g_2 : (-\infty, \infty) \to (-\infty, \infty)$*. If* $\lim_{\alpha \to a} f(\alpha, \beta) = g_2(\beta) \to c < \infty$ *as* $\beta \to$ b *and*

$$\lim_{\beta \to b} f(\alpha, \beta) = g_1(\alpha) < \infty; \ \textit{then} \ \lim_{\alpha \to a} g_1(\alpha) = c.$$

So $\lim_{\alpha \to a} \lim_{\beta \to b} f(\alpha, \beta) = \lim_{\beta \to b} \lim_{\alpha \to a} f(\alpha, \beta)$.

Mandelbrot and Van Ness[4] showed fBm is a.s. continuous on any compact sample set without using BM's continuity property. Our proof of Theorem 8 in the Appendix demonstrates the continuity of fBm based on BM's a.s. continuity.

Theorem 8 *fBm is continuous on any closed and bounded set with probability 1.*

fBm concepts are needed to define fractional Gaussian noise (FGN) (or fractal Gaussian noise). As with defining Gaussian white noise from BM, fractal Gaussian noise is very useful in statistical modeling. Ideally, we like to say FGN is the time derivative of fBm as analogous to wishing to differentiate BM to get Gaussian white noise. Unfortunately, fBm is nondifferentiable a.s., for any $H \in (0, 1)$, including $H = 1/2$ (BM). As with deriving Gaussian white noise from BM, a method for numerical applications is to use incremental processes instead of actual derivatives to 'approximate' FGN processes, but functionals also can be used give a rigorous definition circumventing the need to differentiate. Both ideas are explained here.

To represent FGN as an incremental process, we define the representation as follows:

Definition 5 Given $h > 0$, t, $r \in (-\infty, \infty)$ define $B_{H,h} \equiv [B_H(t + h) - B_H(t)]/h$ as a *computational representation* of fractional (or fractal) Gaussian noise.

This lends itself to two results. They approximate ideas from Ref 2, but we supplement these ideas by stating precise mathematical forms, making minor revisions, and deriving their proofs in the Appendix.

Theorem 9 *Given* $h > 0$; $t, r \in (-\infty, \infty)$ *and* $B_{H,h} \equiv [B_H(t + h) - B_H(t)]/h$*, its covariance is*

$$\begin{aligned} &R_{B_{H,h}}(r) \\ &\quad \equiv \frac{1}{h^2} E\{[B(s + h) - B(s)][B(s + r + h) - B(s + r)]\} \\ &\quad = \frac{V_H h^{2H-2}}{2} \left[\left| \frac{r}{h} + 1 \right|^{2H} - 2 \left| \frac{r}{h} \right|^{2H} + \left| \frac{r}{h} - 1 \right|^{2H} \right]. \end{aligned}$$

Corollary 2 $\lim_{h \to 0} R_{B_{H,h}}(r) = V_H H(2H - 1)|r|^{2H-2} \equiv R_{G_H}(r)$.

Although FGN does not exist due to nondifferentiability and the above is an improvised definition for computation, ideas from functionals give a rigorous definition without numerical approximations. For simplicity, we consider the integral as sample path instead of q.m., as we already showed both exist and are equal.

Definition 6 A *continuous linear functional* is a linear map $T : K \to (-\infty, \infty)$ for a vector space K. We will let $f \in K$ be such that $f \in L^2\{(-\infty, \infty)\}$.

Theorem 10 *Suppose* $G_H : K \to (-\infty, \infty)$ *such that*

$$G_H(f) \equiv \int_{-\infty}^{\infty} f(t) \mathrm{d}B_H(t); f \in L^2\{(-\infty, \infty)\}.$$

Let

$$\begin{aligned} &E[G_H(f)G_H(g)] \\ &\quad = E\left[\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(t)g(s) \mathrm{d}B_H(s) \mathrm{d}B_H(t)\right] < \infty \end{aligned}$$

and

$$\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(t)g(s) \ R_{G_H}(t - s) \mathrm{d}s \mathrm{d}t < \infty.$$

Then

$$\begin{aligned} &E\left[\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(t)g(s) \mathrm{d}B_H(s) \mathrm{d}B_H(t)\right] \\ &\quad = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(t)g(s) R_{G_H}(t - s) \mathrm{d}s \, \mathrm{d}t \\ &\quad \equiv V_H H(2H - 1) \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(t)g(s)|t - s|^{2H-2} \mathrm{d}s \mathrm{d}t \end{aligned}$$

We call G_H 'fractional or fractal Gaussian noise', avoiding the issue that fBm is a.s. nondifferentiable.

Definition 7 For K as previously defined, the linear functional $G_H : K \to (-\infty, \infty)$ such that

$$G_H(f) \equiv \int_{-\infty}^{\infty} f(t)\mathrm{d}B_H(t) < \infty \text{ for } f \in L^2\{(-\infty, \infty)\}$$

is known as FGN and merely symbolized

$$G_H(f) \equiv \int_{-\infty}^{\infty} f(t)G_H(t)\mathrm{d}t \equiv \int_{-\infty}^{\infty} f(t)[\mathrm{d}B_H(t)/\mathrm{d}t]\mathrm{d}t.$$

(As always we have $\omega \in \Omega$.)

Since fBm increments are wide-sense stationary, one can use an orthogonal increments process in frequency domain in the functional, as authors of Ref 2 did, instead of $\mathrm{d}B_H(t)$. Yet our more simple and direct idea of using $\mathrm{d}B_H(t)$ works and avoids needing to describe a new process. Also as fBm is a.s. not differentiable, we use asymptotic (limits) results to prove Theorem 10 in the Appendix.

Barton and Poor[2] give the spectral density function of FGN as

$$S_{G_H}(\omega) = |\omega|^{1-2H};\ 0 \neq \omega \in (-\infty, \infty)$$

as an approximation to the spectral density of $R_{B_{H,b}}(r)$ for $0 < |\omega\delta| \ll 1$. This was in done by taking the Fourier transform and approximating at very small magnitudes of $\omega\delta$. Note $1 - 2H < 0$ for $H > 1/2$. Processes with this type of spectral density functions, for which the frequency variables are in the denominators of fractions, are known as processes with '$1/f$' type spectrum.

APPLICATIONS OF fBm

BM has been conventionally used, often in the form of either its increments or Gaussian white noise (FGN with $H = 1/2$) to model the random component of real world events. This is often desirable because BM increments and Gaussian white noise are time independent and Gaussian. Yet there are phenomena in our world that do not exhibit statistical independence. From the covariance formulas for FGN and increments of fBm for $H \in (1/2, 1)$, we see they exhibit long term slowly decreasing time covariance (autocovariance). Being Gaussian implies they exhibit long term slowly decreasing statistical dependencies. Thus, concepts from fBm are useful as a model for those types of events. Furthermore, fBm with $H \in (0, 1/2)$ are called, according to Mandelbrot,[5] antipersistent processes, meaning that increments of such processes tend to cause the fBm sample path to return to its previous point. (Perhaps we may wish to analogously call fBm with $H \in (1/2, 1)$ as a persistent stochastic process.) Thus, antipersistent processes have tendencies to dissipate more slowly than BM or fBm if $H \in (1/2, 1)$. Phenomena having this tendency are candidates for fBm models with $H \in (0, 1/2)$, but they are less common than long-term dependent systems with $H \in (1/2, 1)$.[5]

According to Mandelbrot[5] as well as Barton and Poor,[2] various physical events have periods of high and low values that are persistent in nature. For example, it is known that the river discharges and flood levels of many rivers are very time persistent.[5] Also errors in communication channels[2] tend to appear in time persistent groups. Also, within the bursts are sub-bursts that are also persistent, seemingly like fractals. These time persistent phenomena are very possibly better modeled with increments of fBm than with increments of BM.

Because fBm is self-similar and thus statistical fractals, mathematicians and statisticians are interested in its fractal dimension. This fact is useful in characterizing fBm that is used as physical models because there is a simple relation between fractal dimensions and Hurst constants. Thus, we can estimate the Hurst parameter, characterizing the fBm by finding its fractal dimension.

Definition 8 (Barnsley[6]) Let A be a compact subset of a complete metric space and given any $\varepsilon > 0$, let $N(A, \varepsilon)$ be the minimum amount of closed balls of ε that covers A. If

$$D = \lim_{\varepsilon \to 0}\{\ln[N(A, \varepsilon)]/[\ln(1/\varepsilon)]\} < \infty,$$

then $D \equiv D(A)$ is called the *fractal dimension* of A. Fractal dimension can be numerically estimated due to the following.

Theorem 11 (Box Counting Theorem[6]) *Let A be a compact set in an n-dimensional Euclidean metric space. Let $N_k(A)$ be the number of closed square boxes with $1/2^k$ side length covering the space (with origin as a vertex) that intersect A.*

$$\text{If} \qquad D = \lim_{k \to \infty}\left\{\frac{\ln[N_k(A)]}{\ln(2^k)}\right\} < \infty,$$

then it is the fractal dimension of A.

Another type of dimension that is given by Barnsley is the Hausdorff–Besicovitch dimension. See Barnsley,[6] Falconer,[7] or Mandelbrot[5] for its definition. The relation between the Hurst parameter of fBm and dimension is as follows:

Theorem 12 (Falconer[7]) *A fBm with Hurst parameter H has both fractal dimension and Hausdorff–Besicovitch dimension of* $2 - H$.

With this relation we can use either the Box Counting Theorem or D's definition to estimate the Hurst parameter of a fBm model. The Box Counting Theorem should be easier although both require meticulous counting. In using the theorem, one can experimentally count to obtain $N_n(A)$. Because

$$D = \lim_{\varepsilon \to 0} \left\{ \frac{\ln[N(A,\varepsilon)]}{\ln(1/\varepsilon)} \right\},$$

if it exists, we can also plot the numerator verses the denominator of the ratio and estimate the slope of the straight line $\ln[N(A,\varepsilon)] = D \ln(1/\varepsilon)$. This can be done by least squares regression via the linear regression model $\ln[N(A,\varepsilon)] = D\ln(1/\varepsilon) + C + \eta$ with regression constant C and model noise η. We can also observe results, study residuals, see whether the regression constant is small, and/or do a test of hypothesis. If analysis and/or a hypothesis test show a poor fit, A may not be well modeled as fBm or any self-similar process. If we get a good fit and assume self-similarity implies being fBm, we can use the above results to get the Hurst parameter. This extends ideas of Ref 6.

A vector of stochastic process of more than one dimension may be used to represent multivariate random phenomena. A random vector with scalar fBm elements or increments can be used to model multidimensional systems with fBm or incremental fBm characteristics. A vector process parameterized by time is embedded in a $n + 1$ Euclidean space, where the random vector ordinate is in n space and we add one extra dimension for the time parameter abscissa. With multidimensional fBm, the fractal dimension is given by a generalization of the above.

Theorem 13 (Schroeder[8], Chapter 5) *If a fBm vector process with independent components in Euclidean n space (thus in* $n + 1$ *space including time) has Hurst parameter H for each scalar component, H is related to its fractal dimension D by* $H + D = n + 1$ *a.s. The result for scalar fBm can be deduced from this by letting* $n + 1 = 2$.

Furthermore, Mandebrot gave a formula for the Hausdorff–Besicovitch dimension for a fBm vector ordinate embedded in n space without the time dimension as follows.

Theorem 14 (Mandelbrot[5]) *Let* $\mathbf{B}_H$ *be a fBm vector embedded in Euclidean n space, whose components are independent. Then its Hausdorf–Besicovitch dimension is* $D_H = 1/H$ *if* $1/H < n$ *and* $D_H = n$ *if* $1/H \geq n$. *(So on planar surfaces,* $D_H = 2$ *if* $1/H \geq 2$.*)*

The relation between D and H can again be used to find H from an estimate of D. We wish to slightly qualify the above three dimension-related theorems by saying they hold a.s. only, as being statistical, there exist events with probability 0 where the relationships fail. Dimensions are useful in characterizing fBm models of the real world systems as fractals. Yet note again they are fractals only in the statistical sense. However, real life phenomena rarely show exact scale factor duplication of their attributes, but only similar attributes regardless to scale. Exact fractals rarely occur in reality. Thus statistical fractals may actually be better models of reality than deterministic fractals.

Physical system states are often modeled by the following linear differential equations,[9] where the terms here may be vectors and matrices as well as scalars: $\mathbf{X}(t) = \mathbf{F}(t)\mathbf{X}(t) + \mathbf{G}(t)\mathbf{W}(t)$ for which $\mathbf{W}$ is a vector or scalar of Gaussian white noise. Because BM is not differentiable, we prefer to write $\mathrm{d}\mathbf{X}(t) = \mathbf{F}(t)\mathbf{X}(t)\mathrm{d}t + \mathbf{G}(t)\mathrm{d}\mathbf{B}(t)$. For processes whose random disturbances are not independent but possess long-term dependencies or antipersistencies, we can use the analogous model, $\mathrm{d}\mathbf{X}(t) = \mathbf{F}(t)\mathbf{X}(t)\mathrm{d}t + \mathbf{G}(t)\mathrm{d}\mathbf{B}_H(t)$; $H \in (0, 1)$. The solution to this stochastic differential equation must be

$$\mathbf{X}(t) = \boldsymbol{\Phi}(t, t_0)\mathbf{X}(t_0) + \int_{t_0}^{t} \boldsymbol{\Phi}(t,\alpha)\mathbf{G}(\alpha)\mathrm{d}\mathbf{B}_H(\alpha),$$

given initial condition $X(t_0)$, as its derivation is analogous to such equation driven by Gaussian white noise, which has the same solution form. It can be shown our so-called 'transition matrix' must be such that $\dot{\boldsymbol{\Phi}}(t, t_0) = \mathbf{F}(t)\boldsymbol{\Phi}(t, t_0)$ where $\boldsymbol{\Phi}(t_0, t_0) = \mathbf{I}$ and replacing t_0 by τ to get $\Phi(t, \tau)$.[10] (Section 2.3). Using this solution form, we can write it in discrete form for times t_0, t_1, t_2, ... as follows: $\mathbf{X}(t_{k+1}) = \boldsymbol{\Phi}(t_{k+1}, t_K)\mathbf{X}(t_k) + \mathbf{N}(t_k)$ for which

$$\mathbf{N}(t_k) \equiv \int_{t_{k-1}}^{t_k} \boldsymbol{\Phi}(t,\alpha)\mathrm{d}\mathbf{B}_H(\alpha).$$

For small time increments, we approximate

$$\int_{t_k}^{t_{k+1}} \boldsymbol{\Phi}(t,\alpha)\mathrm{d}\mathbf{B}_H(\alpha)$$

as $\boldsymbol{\Phi}(t_{k+1}, t_k)[\mathbf{B}_H(t_{k+1}) - \mathbf{B}_H(t_k)] \equiv \boldsymbol{\Phi}(t_{k+1}, t_k)\Delta\mathbf{B}_H(t_{k+1})$. Suppose we are also given the following model to represent measurement of the system state: $\mathbf{Z}(t_{k+1}) = \mathbf{A}(t_{k+1})\mathbf{X}(t_{k+1}) + \mathbf{V}(t_{k+1})$ such that

$$\mathbf{V}(t_k) \equiv \int_{t_k}^{t_{k+1}} \mathbf{S}(t_k)\mathrm{d}\mathbf{B}(\tau)$$

for which the random noise $\mathbf{B}$ is a BM vector and $\mathbf{S}$ is a scalar or positive semidefinite matrix. If the state equation was driven by Gaussian white noise, it is well known that the state, $\mathbf{X}(t_{k+1})$, can be estimated with the Kalman filter.[10] Yet Kalman filters fail for fBm, as the required independence assumption of the driving noise is obviously absent. So we resort to Bayesian estimation to find an optimal estimate, which, as in Kalman filter theory, is optimal with respect to the expectation of a prescribed loss function. Let $\hat{\mathbf{X}}(\mathbf{Z}(t_1),\ldots,\mathbf{Z}(t_r)) \equiv \hat{\mathbf{X}}(\mathbf{Z}^*)$ represent a state estimate given the r vectors of observations. Let actual observations be represented by lower case letters. So let $\mathbf{z}^* \equiv (\mathbf{z}(t_1),\ldots,\mathbf{z}(t_r))^{\mathrm{T}}$ be the set of r observation vectors of the measurement model $\mathbf{Z}^*$. The prescribed loss function is $L(\mathbf{X},\hat{\mathbf{X}}(\mathbf{z}^*)) \equiv (\mathbf{X}-\hat{\mathbf{X}}(\mathbf{z}^*))^{\mathrm{T}}(\mathbf{X}-\hat{\mathbf{X}}(\mathbf{z}^*))$ and so we optimize by finding the $\hat{\mathbf{X}}(\mathbf{z}^*)$ that minimizes $E\{(\mathbf{X}-\hat{\mathbf{X}}(\mathbf{z}^*))^{\mathrm{T}}(\mathbf{X}-\hat{\mathbf{X}}(\mathbf{z}^*))\}$ for a given set of measurement observations, $\mathbf{z}^*$. The following Bayesian estimation theorem is well known in the literature.

Theorem 15 (Meditch[10], Section 5.2) *Given the loss function, $(\mathbf{X}-\hat{\mathbf{X}}(\mathbf{Z}^*))^{\mathrm{T}}(\mathbf{X}-\hat{\mathbf{X}}(\mathbf{Z}^*))$, the estimate $\hat{\mathbf{X}}$ that optimizes its expectation, $E[(\mathbf{X}-\hat{\mathbf{X}}(\mathbf{Z}^*))^{\mathrm{T}}(\mathbf{X}-\hat{\mathbf{X}}(\mathbf{Z}^*))]$, is $E(\mathbf{X}|\mathbf{Z}^*)$, the 'Bayesian estimate'. With actual physically observed measurements, $E(\mathbf{X}|\mathbf{z}^*)$ optimizes $E\{(\mathbf{X}-\hat{\mathbf{X}}(\mathbf{z}^*))^{\mathrm{T}}(\mathbf{X}-\hat{\mathbf{X}}(\mathbf{z}^*))\}$.*

In fact $E(\mathbf{X}|\mathbf{Z}^*)$ is an optimal estimator to a broad class of other loss functions too (Ref 10, Section 5.2). To show a formula for this estimator, we need the following to take the inverse of a covariance matrix.

Lemma 2 (Meditch[10], p. 94) *Let $\mathbf{X}$ and $\mathbf{Z}$ be random vectors of dimensions n and m, respectively, with covariance matrix*

$$\mathbf{P} = \begin{pmatrix} \mathbf{P}_{\mathrm{XX}} & \mathbf{P}_{\mathrm{XZ}} \\ \mathbf{P}_{\mathrm{ZX}} & \mathbf{P}_{\mathrm{ZZ}} \end{pmatrix}$$

such that $\mathbf{P}_{\mathrm{RS}} \equiv \mathrm{cov}(\mathbf{R},\mathbf{S})$ for any random vector $\mathbf{R}$ and $\mathbf{S}$. Then $\mathbf{P}^{-1} = \begin{pmatrix} \mathrm{A} & \mathrm{B} \\ \mathrm{B}^{\mathrm{T}} & \mathrm{C} \end{pmatrix}$ where

$$\mathbf{A} = (\mathbf{P}_{\mathrm{XX}} - \mathbf{P}_{\mathrm{XY}}\mathbf{P}_{\mathrm{YY}}^{-1}\mathbf{P}_{\mathrm{YX}})^{-1} = \mathbf{P}_{\mathrm{XX}}^{-1} + \mathbf{P}_{\mathrm{XX}}^{-1} \times \mathbf{P}_{\mathrm{XZ}}\mathbf{C}\mathbf{P}_{\mathrm{ZX}}\mathbf{P}_{\mathrm{XX}}^{-1};$$

$$\mathbf{B} = -\mathbf{A}\mathbf{P}_{\mathrm{XZ}}\mathbf{P}_{\mathrm{ZZ}}^{-1} = -\mathbf{P}_{\mathrm{XX}}^{-1}\mathbf{P}_{\mathrm{XZ}}\mathbf{C};$$

and

$$\mathbf{C} = (\mathbf{P}_{\mathrm{ZZ}} - \mathbf{P}_{\mathrm{ZX}}\mathbf{P}_{\mathrm{XX}}^{-1}\mathbf{P}_{\mathrm{XZ}})^{-1} = \mathbf{P}_{\mathrm{ZZ}}^{-1} + \mathbf{P}_{\mathrm{ZZ}}^{-1} \times \mathbf{P}_{\mathrm{ZX}}\mathbf{A}\mathbf{P}_{\mathrm{XZ}}\mathbf{P}_{\mathrm{ZZ}}^{-1}.$$

These formulas can be derived by solving for $\mathbf{A}$, $\mathbf{B}$, and $\mathbf{C}$ in $\mathbf{P}^{-1}\mathbf{P} = \mathbf{I}$. This is used in deriving the following optimal estimator, $E(\mathbf{X}|\mathbf{Z}^*)$, for Gaussian processes. See Ref 10, Section 3.6, for another explanation of the formula for the expectation $E(\mathbf{X}|\mathbf{Z})$, which from Theorem 15 forms the actual optimal estimator.

Theorem 16 *Given Gaussian random vectors, $\mathbf{X}$ and $\mathbf{Z}$, the optimal estimator, $\hat{\mathbf{X}}(\mathbf{Z}) = E(\mathbf{X}|\mathbf{Z})$, is given by the formula*

$$E(\mathbf{X}|\mathbf{Z}) = E[\mathbf{X}] + \mathbf{P}_{\mathrm{XZ}}\mathbf{P}_{\mathrm{ZZ}}^{-1}(\mathbf{Z} - E[\mathbf{Z}]) \quad \textit{and}$$

$$\mathrm{cov}(\mathbf{X}|\mathbf{Z}) \equiv E\{(\mathbf{X} - E(\mathbf{X}|\mathbf{Z}))(\mathbf{X} - E(\mathbf{X}|\mathbf{Z}))^{\mathrm{T}}|\mathbf{Z}\} = \mathbf{P}_{\mathrm{XX}} - \mathbf{P}_{\mathrm{XZ}}\mathbf{P}_{\mathrm{ZZ}}^{-1}\mathbf{P}_{\mathrm{ZX}}.$$

Analogously, given observations $\mathbf{z}$ of $\mathbf{Z}$,

$$\hat{\mathbf{x}}(\mathbf{z}) \equiv E(\mathbf{X}|\mathbf{z}) = E[\mathbf{X}] + \mathbf{P}_{\mathrm{XZ}}\mathbf{P}_{\mathrm{ZZ}}^{-1}(\mathbf{z} - E[\mathbf{Z}]) \textit{ and}$$

$$\mathrm{cov}(\mathbf{X}|\mathbf{z}) = \mathbf{P}_{\mathrm{XX}} - \mathbf{P}_{\mathrm{XZ}}\mathbf{P}_{\mathrm{ZZ}}^{-1}\mathbf{P}_{\mathrm{ZX}}.$$

Notice the covariance is actually $\mathrm{cov}(\mathbf{X}|\mathbf{z}) \equiv E\{(\mathbf{X}-\hat{\mathbf{X}}(\mathbf{z}))(\mathbf{X}-\hat{\mathbf{X}}(\mathbf{z}))^{\mathrm{T}}|\mathbf{z}\}$ and so this conditional covariance is actually the covariance of the 'estimation error', given observations $\mathbf{z}$. Therefore this gives us an idea of how 'good' is the optimal estimator. We can even go a step further, as one can show that given Gaussian random vectors, $\mathbf{X}$ and $\mathbf{Z}$, the (also Gaussian) random vector $\mathbf{X} - E(\mathbf{X}|\mathbf{Z})$ is independent of the random vector $\mathbf{Z}$.[10] Thus $E\{(\mathbf{X}-\hat{\mathbf{X}}(\mathbf{z}))(\mathbf{X}-\hat{\mathbf{X}}(\mathbf{z}))^{\mathrm{T}}|\mathbf{z}\} = E\{(\mathbf{X}-\hat{\mathbf{X}}(\mathbf{z}))(\mathbf{X}-\hat{\mathbf{X}}(\mathbf{z}))^{\mathrm{T}}\}$, making this estimation error covariance the same regardless of observations, just as in Kalman filters.

We can use this theorem to find the optimal state estimator of the discrete state model we previously derived, $\mathbf{X}(t_{k+1}) = \mathbf{\Phi}(t_{k+1},t_K)\mathbf{X}(t_k) + \mathbf{N}(t_k)$, with its corresponding measurement model, $\mathbf{Z}(t_{k+1}) = \mathbf{A}(t_{k+1})\mathbf{X}(t_{k+1}) + \mathbf{V}(t_{k+1})$. This is because the noises are stochastic integrals with respect to fBm and BM, respectively, making them Gaussian. So the processes themselves are Gaussian. For simplicity in notation we rewrite this as $\mathbf{X}(k+1) = \mathbf{\Phi}(k+1,k)\mathbf{X}(k) + \mathbf{N}(k)$ and $\mathbf{Z}(k+1) = \mathbf{A}(k+1)\mathbf{X}(k+1) + \mathbf{V}(k+1)$, respectively. Therefore given all the (discrete) observations up to time t_{k+1}, $\mathbf{z}^*(q) \equiv (\mathbf{z}(1),\ldots,\mathbf{z}(q))^{\mathrm{T}}$, the optimal estimator must be the random vector $\hat{\mathbf{x}}(k+1) = E(\mathbf{X}(k+1)) + \mathbf{P}_{\mathrm{XZ}^*}\mathbf{P}_{\mathrm{Z}^*\mathrm{Z}^*}^{-1}(\mathbf{z}^*(q) - E(\mathbf{Z}^*(q)))$ and the error covariance must be

$$\mathrm{cov}(\mathbf{X}(k+1)|\mathbf{z}^*) = E\{[\mathbf{X}(k+1) - E(\mathbf{X}(k{+}1)|\mathbf{z}^*)] \times [\mathbf{X}(k+1) - E(\mathbf{X}|\mathbf{z}^*)]^{\mathrm{T}}\} = \mathbf{P}_{\mathrm{XX}} - \mathbf{P}_{\mathrm{XZ}^*}\mathbf{P}_{\mathrm{Z}^*\mathrm{Z}^*}^{-1}\mathbf{P}_{\mathrm{Z}^*\mathrm{X}}.$$

We emphasize $k+1$ does not necessarily need to be the current state measured state. We assume the state mean at initial time t_0. $k+1$ designates any other time. So t_{k+1} may be the time of prior or current observations. The same theorem gives us the following estimator for all states together:

$$\mathbf{X}^* = (\mathbf{X}_1, \ldots, \mathbf{X}_{k+1}, \ldots \mathbf{X}_q)^{\mathrm{T}} :$$
$$\hat{\mathbf{X}} = E(\mathbf{X}^*) + \mathbf{P}_{\mathbf{X}^*\mathbf{Z}^*}\mathbf{P}^{-1}_{\mathbf{Z}^*\mathbf{Z}^*}[\mathbf{Z}^*(q) - E(\mathbf{Z}^*(q))]$$

with the covariance error

$$\mathbf{cov}[(\mathbf{X}^* - \hat{\mathbf{X}})^{\mathrm{T}}(\mathbf{X}^* - \hat{\mathbf{X}})] = \mathbf{P}_{\mathbf{X}^*\mathbf{X}^*} - \mathbf{P}_{\mathbf{X}^*\mathbf{Z}^*}\mathbf{P}^{-1}_{\mathbf{Z}^*\mathbf{Z}^*}\mathbf{P}_{\mathbf{Z}^*\mathbf{X}^*}.$$

Both $E(\mathbf{X}(k+1))$ and $E(\mathbf{X}^*)$ can be computed by noting that iterating the state model gives

$$\begin{aligned} E[\mathbf{X}(k+1)] &= \mathbf{\Phi}(k+1,k)E[\mathbf{X}(k)] + E[\mathbf{N}(k)] \\ &= \mathbf{\Phi}(k+1,0)E[\mathbf{X}_0] + \sum_{i=0}^{k} \Phi(k+1,\ i+1)E[\mathbf{N}(i)] \\ &= \mathbf{\Phi}(k+1,0)E[\mathbf{X}_0]. \end{aligned}$$

Recall we assume the initial condition, $E[\mathbf{X}_0]$, is given.

To compute the various covariance ($\mathbf{P}_{\mathbf{RS}} \equiv \mathrm{cov}(\mathbf{R},\mathbf{S})$) in both $\hat{\mathbf{X}}$ and the error covariance, first note that by the vector counterpart of our prior derivation,

$$\mathrm{cov}(\mathbf{N}(j),\mathbf{N}(k)) = V_H H(2H-1)\int_{t_j}^{t_{j+1}}\int_{t_k}^{t_{k+1}} \Phi(t_{j+1},\alpha) \times \mathbf{G}(\alpha)\mathbf{G}^{\mathrm{T}}(\beta)\Phi^{\mathrm{T}}(t_{k+1},\beta)|\alpha-\beta|^{2H-2}\mathrm{d}\alpha\mathrm{d}\beta.$$

As BM has orthogonal increments, the properties of stochastic integrals of deterministic functions with respect to BM implies

$$\mathrm{cov}(\mathbf{V}(j),\mathbf{V}(k)) = \int_{t_k}^{t_{j+1}} \mathbf{S}(\alpha)\mathbf{S}^{\mathrm{T}}(\alpha)\mathrm{d}\alpha$$

($\mathbf{S}$ as previously defined). In real world applications, it is reasonable to assume that the system state noise, $\mathbf{N}$, and the measurement noise, $\mathbf{V}$ are statistically independent. With these two formulas for the two types of noise covariance and again noting

$$\mathbf{X}(k+1) = \mathbf{\Phi}(k+1,0)\mathbf{X}_0 + \sum_{i=0}^{k}\Phi(k+1,i+1)\mathbf{N}(i)$$

and given the measurement model $\mathbf{Z}(k+1) = \mathbf{A}(k+1)\mathbf{X}(k+1) + \mathbf{V}(k+1)$, we can compute all the required covariance in the formulas for $\hat{\mathbf{X}}$ and the error covariance by algebraic manipulations.

Our last hurdle is to compute the inverse expression, $\mathbf{P}^{-1}_{\mathbf{Z}^*\mathbf{Z}^*}$. If this inverse is too large to easily manage, we can partition it into four submatrices and use Lemma 2, which requires only finding the upper and lower submatrices containing the diagonal. If these two matrices are still large, we can find its inverse by partitioning further, and so on. Otherwise, we can use Lemma 2 to build up to $\mathbf{P}^{-1}_{\mathbf{Z}^*\mathbf{Z}^*}$ as in the following way. Let

$$\mathbf{P}^{-1}_{\mathbf{Z}^*\mathbf{Z}^*} = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1q} \\ a_{21} & a_{22} & & \\ \vdots & & & \vdots \\ a_{q1} & & \cdots & a_{qq} \end{pmatrix}^{-1}.$$

We begin with finding

$$\mathbf{M}_1 = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}^{-1}.$$

Then we add one more dimension to get

$$\mathbf{M}_2 = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix}$$

and find its inverse by partitioning into $\mathbf{M}_1$,

$$\mathbf{N}_1 \equiv (a_{33}), \mathbf{L}_1 = \begin{pmatrix} a_{31} & a_{32} \end{pmatrix}, \text{ and } \mathbf{R}_1 \equiv \begin{pmatrix} a_{13} \\ a_{23} \end{pmatrix}$$

and using the Lemma, which requires $\mathbf{M}_1^{-1}$ and $\mathbf{N}_1^{-1} = (1/a_{33})$. We can then add another dimension to get

$$\begin{pmatrix} a_{11} & a_{12} & a_{13} & a_{14} \\ a_{21} & a_{22} & a_{23} & a_{24} \\ a_{31} & a_{32} & a_{33} & a_{34} \\ a_{41} & a_{42} & a_{43} & a_{44} \end{pmatrix}$$

and partition to get analogous submatrices, $\mathbf{M}_2, \mathbf{N}_2 \equiv (a_{44}), \mathbf{L}_2$, and $\mathbf{R}_2$, to find the inverse of this 4×4 matrix. We can recursively continue until we compute $\mathbf{P}^{-1}_{\mathbf{Z}^*\mathbf{Z}^*}$.

Now we have all the pieces, giving us a batch optimal estimator of a discrete linear model driven by FGN with respect to the expectation of the quadratic loss function and given a linear observation measurement model with time-independent Gaussian noise. This actually generalizes the Kalman filter, which is a sequential estimator of a linear model driven by Gaussian white noise. The beauty of this method is we found an estimator that no longer requires time-independent increments of Gaussian process models.

Our batch estimator can be modified to a sequential estimator for $\{\hat{\mathbf{X}}(k) : k = 1, 2, \ldots, n < \infty\}$, as in Ref 11, which also gives a matrix inverse needing less computation. The sequential method is much more efficient to compute than the batch. Yet it is less accurate than the batch, as only measurements up to the current state being estimated are used in the sequential method. So only the last estimate, $\hat{\mathbf{X}}(n)$, uses all measurements. In the batch method, all measurements are used on every estimate.

CONCLUSIONS

In this article, we described a generalization of standard BM called fractional Brownian motion (fBm), which is also called fractal Brownian motion due to its fractal properties. We defined it and saw how it is parameterized by a single Hurst parameter and how BM is a special case. We then explained its main properties and proved some of them in the Appendix. We demonstrated in the Appendix how its a.s. sample path continuity can be proven from the a.s. continuity of BM. We finally explained how fBm is used to model real world events. As part of this explanation, we gave relationships between the fractal dimension of fBm and its Hurst parameter. We finally described statistical inference on linear models driven by fractional or fractal Gaussian noise instead of the traditional Gaussian white noise.

APPENDIX: PROOFS OF THEOREMS

Proof of Theorem 2 In the process of deriving the already established finite covariance of fBm (from Barton and Poor[2]), one had to derive $E[B(t)B(s)]$. This implicitly also derived $E(B^2(t))$, meaning it too is finite. Yet this derivation is equivalent to deriving the fact that

$$\infty > E\Big\{\Big(\frac{1}{\Gamma(H+\frac{1}{2})}\Big)^2\Big(\int_{-\infty}^{0} M_1(t,s)\,\mathrm{d}B(s) + \int_0^t \mathrm{M}_2(\mathrm{t},\mathrm{s})\,\mathrm{d}B(s)\Big)\Big(\int_{-\infty}^{0} M_1(t,s)\mathrm{d}B(s) + \int_0^t \mathrm{M}_2(\mathrm{t},\mathrm{s})\mathrm{d}B(s)\Big)\Big\},$$

which via properties of orthogonal increments of BM processes (Ref 1, p. 99, where $F(\mathrm{d}\lambda) = \mathrm{d}\lambda$) equals

$$\Big(\frac{1}{\Gamma(H+\frac{1}{2})}\Big)^2\Big(\int_{-\infty}^{0} M_1^2(t,s)\mathrm{d}s + \int_0^t M_2^2(t,s)\mathrm{d}s\Big).$$

Because this must exist, being used to derive the established variance, the real valued function,

$$B_H(t) \equiv \lim_{\substack{\alpha\to 0\\ R\to\infty}} \Big(\int_{-R}^{-\alpha} M_1(t,s)\mathrm{d}B(s) + \int_0^t \mathrm{M}_2(\mathrm{t},\mathrm{s})\mathrm{d}B(s)\Big),$$

is measurable as a stochastic process (Definition 1) by an orthogonal increments process property in Ref 3, p. 430. This is because Doob's derivation of it holds even on integration interval $(-\infty, 0)$, although the integrals uses mean square limits in α and R. Therefore fBm exists as a real valued function and is measurable as a stochastic process. (We did not say fBm exists to prove it exists. We said because the variance exists, integrals used to derive it must exist. Yet these integrals happen to be ones implying fBm's existence.)

Proof of Theorem 4 $E[(B_H(t) - B_H(s))(B_H(u) - B_H(v))] = E[B_H(t)B_H(u)] - E[B_H(s)B_H(u)] - E[B_H(t)B_H(v)] + E[B_H(s)B_H(v)]$. Using the covariance formula for fBm, this equals

$$\frac{V_H}{2}\{|s-u|^{2H} - |t-u|^{2H} + |t-v|^{2H} - |s-v|^{2H}\}.$$

Thus the covariance of increments are not affected by the time itself but only by differences in times between two incremental processes. These increments also have 0 and thus constant means. So they must be wide-sense stationary.

Proof of Theorem 5 $\mathrm{cov}\{[B_H(t_0+ht) - B_H(t_0)][B_H(t_0+hr) - B_H(t_0)]\} = \mathrm{cov}[(B_H(ht))B_H(hr)]$ because fBm has stationary increments. By covariance of fBm,

$$\begin{aligned}\mathrm{cov}[(B_H(ht))B_H(hr)] &= |h^H|\frac{V_H}{2}(|t^H| + |r^H| - |t-r|^H)\\ &= h^H \mathrm{cov}(B_H(t)B_H(r))\end{aligned}$$

for $h > 0$. Because Gaussian processes are determined by the covariance and mean and the two processes in the theorem both have 0 mean, we must have that the two processes have the same finite distributions.

Proof of Theorem 6 Using the self-similarity formula, $\lim_{h\to 0}\frac{1}{h}[B_H(t_0+ht) - B_H(t_0)] \cong \lim_{h\to 0}\frac{h^H}{h}B_H(t) = \lim_{h\to 0}\frac{1}{h^{1-H}}B_H(t)$. Note $1 - H \in (0,1)$. Because this limit diverges for nonzero fBm, so does $\frac{1}{h}[B_H(t_0+ht) - B_H(t_0)]$ a.s., for it has the same finite distributions. Thus B_H is a.s. nondifferentiable.

Proof of Theorem 7 By known properties, q.m. convergence imply convergence in probability, which imply the existence of a subsequence that converges a.s. to the same limit[1] (p. 20). Thus, there is a subsequence $\{X_{k_j}: k, j = 1, 2, \ldots\}$ of $\{X_k\}$ that converges a.s. to X. Yet $X_k \to Y$ a.s. and so X_{k_j}, being a subsequence of X_k must also converge to Y a.s. Thus $X = Y$ a.s.

Proof of Corollary 1 Let $t > 0$. The integrand,

$$M_1(t, s) \equiv |t-s|^{H-\frac{1}{2}} - |s|^{H-\frac{1}{2}} \text{ of } \int_{-R}^{-\alpha} M_1(t, s)\mathrm{d}B(s);$$

$R, \alpha > 0$ is of bounded variation for any interval $[-R, -\alpha]$. Recall B is a.s. continuous. With a term, M_1, bounded and the other, B continuous, by real analysis, Riemann–Stieltjes integral,

$$\int_{-R}^{-\alpha} M_1(t, s)\mathrm{d}B(s) < \infty,$$

exists a.s. on $\omega \in \Omega$ as a sample path integral. Because limits of same sequence of sums are usable to define both integrals, by Theorem 7, it equals a.s. to its q.m. counterpart for all $\alpha, R > 0$. As a q.m. integral,

$$\infty > \int_{-\infty}^{0} M_1(t, s)\mathrm{d}B(s).$$

So there is a K where

$$\left|\int_{-R}^{-\alpha} M_1(t, s)\mathrm{d}B(s)\right| \leq K$$

for all $\alpha, R > 0$. Because the q.m. integral a.s. equals the sample path integral, it must be also K bounded a.s. as a sample path integral for all $\alpha, R \geq 0$. So

$$\int_{-\infty}^{0} M_1(t, s)\mathrm{d}B(s)$$

must exist a.s. as both sample path and q.m. integrals, and they are a.s. equal. For $t \leq 0$, we get integrals

$$\int_{-R}^{t-\beta} M_1(t, s)\mathrm{d}B(s) + \int_{t-\beta}^{-\alpha} M_1(t, s)\mathrm{d}B(s);$$

$\alpha, \beta, R > 0$ if $H - 1/2 < 0$. Yet, the argument is analogous to the above except for the fact that we must take care of two improper integrals in analogous ways. We can also show

$$\int_{0}^{t} M_2(t, s)\mathrm{d}B(s)\mathrm{d}B_H(s)$$

exists as both q.m. and a.s as sample path integrals in a similar way as above, accounting for possible cases of improper integrals as in the previous.

Proof of Lemma 1 Given $\varepsilon > 0$, for case of $-\infty < a, b < \infty$ there exists δ_1, δ_2, $\delta_3 > 0$ such that for a real number c, $|f(\alpha, \beta) - g_2(\beta)| < \varepsilon, |g_2(\beta) - c| < \varepsilon$, and $|f(\alpha, \beta) - g_1(\alpha)| < \varepsilon$ whenever $|\alpha - a| < \delta_1$, $|\beta - b| < \delta_2$ and $|\beta - b| < \delta_3$. Let $\delta = \min(\delta_1, \delta_2, \delta_3)$. Note $|g_1(\alpha) - c| = |g_1(\alpha) - f(\alpha, \beta) + f(\alpha, \beta) - g_2(\beta) + g_2(\beta) - c| < 3\varepsilon$ when $|\alpha - a| < \delta$ and $|\beta - b| < \delta$ using the 'Triangle Inequality'. Let $|\alpha - a| < \delta$. Because $|g_1(\alpha) - c|$ does not depend on β and the first inequality holds for any β, it must be that $|g_1(\alpha) - c| < 3\varepsilon$. Thus

$$\lim_{\alpha\to a} g_1(\alpha) = c, \text{ giving } \lim_{\alpha\to a}\lim_{\beta\to b} f(\alpha, \beta) = \lim_{\beta\to b}\lim_{\alpha\to a} f(\alpha, \beta).$$

If either a and/or b has infinite magnitude, there exists N or Ns instead of the respective δ or δs, such that the same condition(s) holds for all $n > N$. The argument then follows in an analogous way.

Proof of Theorem 8 Let A be any closed and bounded real numbered set and $\omega \in \Omega$ any sample path where B is sample path continuous. Let $t \in A$.

$$B_H(t+h) \equiv \frac{1}{\Gamma(H+\frac{1}{2})}\left(\int_{-\infty}^{0} M_1(t, s)\mathrm{d}B(s) + \int_{0}^{t} \mathrm{M}_2(\mathrm{t}, \mathrm{s})\mathrm{d}B(s)\right).$$

Because M_1 and M_2 diverge for $H - 1/2 < 0$ and the first integral has $-\infty$ lower limit, we treat integrals as 'improper'. Proper integral have similar arguments without limits that define improper integrals. Also as BM is continuous a.s. only, we choose a $\omega \in \Omega$ such that BM is indeed continuous for this proof.

Case I $t > 0$:

$$\int_{-\infty}^{0} M_1(t, s)\mathrm{d}B(s) \equiv \lim_{\alpha\to 0}\lim_{R\to\infty}\int_{-R}^{-\alpha} M_1(t, s)\mathrm{d}B(s);$$

$\alpha, R > 0$. By Stieltjes Integration by parts,[12]

$$\int_{-R}^{-\alpha} M_1(t, s)\mathrm{d}B(s) = M_1(t, -\alpha)B(-\alpha) - M_1(t, -R)B(-R) - \int_{-R}^{-\alpha} B(s)\frac{\partial}{\partial s}M_1(t, s)\mathrm{d}s.$$

By the formula of $M_1, M_1(t,-\alpha)B(-\alpha) - M_1(t,-R)B(-R)$ is continuous at any $t \in A$. Also, because BM is continuous and, for $t > 0$,

$$B(s)\frac{\partial}{\partial s}M_1(t+h,s)\mathrm{d}s = B(s)\left(H-\frac{1}{2}\right) \times \left\{-(t-s)^{H-\frac{3}{2}} + (-s)^{H-\frac{3}{2}}\right\},$$

this term is continuous at any $t \in A$ and $s \in [-R, 0]$. Thus

$$\int_{-R}^{-\alpha} B(s)\frac{\partial}{\partial s}M_1(t,s)\mathrm{d}s$$

is continuous at any given $t \in A^{12}$ (p. 416). Adding continuous terms makes

$$\int_{-R}^{-\alpha} M_1(t,s)\mathrm{d}B(s)$$

continuous. For $h > 0$,

$$\lim_{h\to 0}\int_{-\infty}^{0} M_1(t,s)\mathrm{d}B(s) \equiv \lim_{h\to 0}\lim_{\alpha\to 0}\lim_{R\to\infty}\int_{-R}^{\alpha} M_1(t,s)\mathrm{d}B(s).$$

By Lemma 1 for any given $t > 0$ and our chosen $\omega \in \Omega$,

$$\lim_{h\to 0}\lim_{\alpha\to 0}\lim_{R\to\infty}\int_{-R}^{0} M_1(t,s)\mathrm{d}B(s) = \lim_{\alpha\to 0}\lim_{R\to\infty}\lim_{h\to 0}\int_{-R}^{-\alpha} M_1(t+h,s)\mathrm{d}B(s).$$

Because we showed

$$\int_{-R}^{-\alpha} M_1(t,s)\mathrm{d}B(s)$$

is continuous,

$$\lim_{h\to 0}\int_{-R}^{-\alpha} M_1(t+h,s)\mathrm{d}B(s) = \int_{-R}^{-\alpha} M_1(t,s)\mathrm{d}B(s).$$

Thus,

$$\lim_{h\to 0}\lim_{\alpha\to 0}\lim_{R\to\infty}\int_{-R}^{0} M_1(t+h,s)\,\mathrm{d}B$$
$$= \lim_{\alpha\to 0}\lim_{R\to\infty}\int_{-R}^{-\alpha} M_1(t,s)\,\mathrm{d}B \equiv \int_{-\infty}^{0} M_1(t,s)\mathrm{d}B$$
$$\Rightarrow \int_{-\infty}^{0} M_1(t,s)\mathrm{d}B$$

is continuous for our chosen ω. Now consider M_2. Integrating by parts,

$$\int_0^t M_2(t,s)\mathrm{d}B(s) \equiv \lim_{\alpha\to 0}\int_0^{t-\alpha} M_2(t,s)\mathrm{d}B(s)$$
$$= \lim_{\alpha\to 0}\left\{M_2(t,t-\alpha)B(t-\alpha) - 0 - \int_0^{t-\alpha} B(s)\frac{\partial}{\partial s}M_2(t,s)\mathrm{d}s\right\}.$$

Because BM is continuous for the chosen $\omega \in \Omega$ and, for $t > 0$,

$$B(s)\frac{\partial}{\partial s}M_2(t+h,s) = -B(s)\left(H-\frac{1}{2}\right)(t-s)^{H-\frac{3}{2}},$$

this term is continuous on any given $0 < t \in A$ and $s \in [-R, 0]$.

$$\text{So}\qquad \int_0^{t-\alpha} B(s)\frac{\partial}{\partial s}M_1(t,s)\mathrm{d}s$$

is continuous at any given $t \in A^{12}$ (p. 417). Also note $M_2(t,t-\alpha)B(t-\alpha)$ is continuous at t because BM is continuous. Adding continuous term makes

$$\int_0^{t-\alpha} M_2(t,s)\mathrm{d}B(s)$$

continuous at t. To show

$$\int_0^t M_2(t,s)\mathrm{d}B(s),$$

itself, is continuous, write

$$\lim_{h\to 0}\int_0^{t+h} M_2(t+h,s)\mathrm{d}B(s) \equiv \lim_{h\to 0}\lim_{\alpha\to 0}\int_0^{t+h-\alpha} M_2(t+h,s)\,\mathrm{d}B(s).$$

By the Lemma,

$$\lim_{h\to 0}\lim_{\alpha\to 0}\int_0^{t+h-\alpha} M_2(t+h,s)\mathrm{d}B(s) = \lim_{\alpha\to 0}\lim_{h\to 0}\int_0^{t+h-\alpha} M_2(t+h,s)\mathrm{d}B(s),$$

and by the continuity of

$$\int_0^{t-\alpha} M_2(t,s)\mathrm{d}B(s),$$

we have

$$\lim_{\alpha\to 0}\lim_{h\to 0}\int_0^{t+h-\alpha} M_2(t+h,s)\mathrm{d}B(s)$$
$$=\lim_{\alpha\to 0}\int_0^{t-\alpha} M_2(t,s)\mathrm{d}B(s) \equiv \int_0^t M_2(t,s)\mathrm{d}B(s)$$

(definition). Thus

$$\lim_{h\to 0}\int_0^{t+h} M_2(t+h,s)\mathrm{d}B(s) = \int_0^t M_2(t,s)\mathrm{d}B(s)$$
$$\Rightarrow \int_0^t M_2(t,s)\mathrm{d}B(s)$$

is continuous at $t \in A$.

Case II $t < 0$: Let $h > 0$. Without loss of generality, let $t+h < 0$.

Note $\lim_{h\to 0} B_H(t+h)$

$$= \lim_{h\to 0}\frac{1}{\Gamma(H+1/2)}\Bigg(\int_{-\infty}^{c} M_1(t+h,s)\mathrm{d}B(s) + \int_c^{t+h} M_1(t+h,s)\mathrm{d}B(s) + \int_{t+h}^{0} M_1(t+h,s)\mathrm{d}B(s) + \int_0^{t+h} M_2(t+h)\mathrm{d}B(s)\Bigg)$$
$$= \lim_{h\to 0}\frac{1}{\Gamma(H+1/2)}\Bigg(\int_{-\infty}^{c} M_1(t+h,s)\mathrm{d}B(s) + \int_c^{t+h} M_1(t+h,s)\mathrm{d}B(s) + \int_{t+h}^{0}[M_1(t+h,s) - M_2(t+h,s)]\mathrm{d}B(s)\Bigg) \text{ for } c \in (-\infty, t+h).$$

To perform derivatives in the Stieltjes Integration by parts formula, note M_1 is continuous for $s \in (-\infty, t+h)$, and $M_1 - M_2$ is continuous for $s \in (t+h, 0)$. By the same argument as the $t > 0$ case in dealing with limits, except we use $t+h+\alpha$ on the last integral, instead of $t+h-\alpha$, that is,

$$\int_{t+h+\alpha}^{0}[M_1(t+h,s) - M_2(t+h,s)]\mathrm{d}B(s),$$

we conclude $\lim_{h\to 0} B_H(t+h) = B_H(t)$.

Case III $t = 0$:

$$\int_{-\infty}^{0} M_1(h,s)\,\mathrm{d}B(s) = 0$$

for $t = 0$. So

$$B_H(0) = \lim_{\alpha,\beta\to 0}\int_\beta^\alpha \mathrm{s}^{\mathrm{H}-1/2}\mathrm{d}B(s)$$
$$\equiv \lim_{\alpha,\beta\to 0}\int_\beta^\alpha \mathrm{M}_2(0,\mathrm{s})\mathrm{d}B(s).$$

By arguments similar to before,

$$\lim_{h\to 0}\lim_{\alpha,\beta\to 0}\int_\beta^{\alpha+h} \mathrm{M}_2(\mathrm{h},\mathrm{s})\mathrm{d}B(s)$$
$$= \lim_{\alpha,\beta\to 0}\int_\beta^{\alpha} \mathrm{M}_2(0,\mathrm{s})\mathrm{d}B(s) \equiv B_H(0);\ \alpha,\beta > 0,$$

implying continuity at zero. (Note this also implies $B_H(0) = 0$ a.s.)

Therefore sample paths of fBm must be continuous. Because our chosen $\omega \in \Omega$ appears with probability one, sample paths of fBm must be continuous a.s.

Proof of Theorem 9

$$R_{B_{H,h}}(r) \equiv \frac{1}{h^2}E\{[B_H(s+h) - B_H(s)][B_H(s+r+h) - B_H(s+r)]\} = E\{[B_H(s+h) - B_H(s)] \times [B_H(s+r+h) - B_H(s+r)]\} = E\{B_H(s+h) \times B_H(s+r+h)\} - E\{B_H(s)B_H(s+r+h)\} - E\{B_H(s+h)B_H(s+r)\} + E\{B_H(s)B_H(s+r)\}.$$

From the covariance formula for fBm,

$$R_{B_{H,h}}(r) = \frac{V_H}{2h^2}\{[|s+h|^{2H} + |s+r+h|^{2H} - |r|^{2H}] - [|s|^{2H} + |s+r+h|^{2H} - |r+h|^{2H}] - [|s+h|^{2H} + |s+r|^{2H} - |r-h|^{2H}]\} + \frac{V_H}{2h^2}\{|s|^{2H} + |s+r|^{2H} - |r|^{2H}\} = \frac{V_H}{2h^2}(|r+h|^{2H} - 2|r|^{2H} + |r-h|^{2H})$$
$$= \frac{V_H h^{2H-2}}{2}\left[\left|\frac{r}{h} + 1\right|^{2H} - 2\left|\frac{r}{h}\right|^{2H} + \left|\frac{r}{h} - 1\right|^{2H}\right].$$

Proof of Corollary 2

$$R_{B_{H,h}}(r)$$
$$= \frac{V_H h^{2H-2}}{2}\left(\left|\frac{r}{h}+1\right|^{2H} - 2\left|\frac{r}{h}\right|^{2H} + \left|\frac{r}{h}-1\right|^{2H}\right)$$
$$= \frac{V_H h^{2H-2}}{2}\left(\left|\frac{r}{h}+\frac{h}{h}\right|^{2H} - 2\left|\frac{r}{h}\right|^{2H} + \left|\frac{r}{h}-\frac{h}{h}\right|^{2H}\right)$$

$$= \frac{V_H h^{2H-2}}{2} h^{-2H}(|r+h|^{2H} - 2|r|^{2H} + |r-h|^{2H})$$
$$= \frac{V_H}{2} \frac{(|r+h|^{2H} - 2|r|^{2H} + |r-h|^{2H})}{h^2}.$$

Because both the numerator and denominator of the fraction $\to 0$ as $h \to 0$, differentiate the numerator and denominator twice with respect to h, using L'Hopital's Rule twice, concluding $\lim_{h\to 0} R_{B_{H,h}}(r) = V_H H(2H-1)|r|^{2H-2}$.

Proof of Theorem 10 Define indefinite integrals as a limit of definite integrals. Using Barton and Poor's definition,[2] but interpreting it as 'sample path', given partitions

$$\pi_{1,n} = \{t_0, \ldots, t_n : -R = t_0 < t_1 < \cdots < t_n = R\}$$

and

$$\pi_{2,m} = \{s_0, \ldots, s_n : -R = s_0 < s_1 < \cdots < s_m = R\}$$

for which $0 < R < \infty$ and $\lim_{n\to\infty} \max_{1\le j\le n}(\Delta t_j) = \lim_{m\to\infty} \max_{1\le k\le m}(\Delta s_k) = 0$.

$$\int_{-R}^{R}\int_{-R}^{R} f(t)g(s)\mathrm{d}B_H(s)\mathrm{d}B_H(t)$$
$$\equiv \lim_{n,m\to\infty} \sum_{k=1}^{n}\sum_{j=1}^{m} f(\alpha_j)g(\sigma_k)\,\Delta B_H(s_j)\,\Delta B_H(t_k);$$

$\alpha_j \in [t_{j-1},\ t_j], \sigma_k \in [s_{k-1}, s_k]$. Given (Ω, A, P),

$$E\left[\int_{-R}^{R}\int_{-R}^{R} f(t)g(s)\mathrm{d}B_H(s)\mathrm{d}B_H(t)\right]$$
$$\equiv \int_{\Omega}\left[\int_{-R}^{R}\int_{-R}^{R} f(t)g(s)\mathrm{d}B_H(s)\mathrm{d}B_H(t)\right]\mathrm{d}P(\omega).$$
$$\equiv \int_{\Omega}\left\{\lim_{n,m\to\infty}\sum_{k=1}^{n}\sum_{j=1}^{m} f(\alpha_j)g(\sigma_k)\Delta B_H(s_j,\omega)\right.$$
$$\left.\times\Delta B_H(t_k,\omega)\right\}\mathrm{d}P(\omega);$$

$\alpha_j \in [t_{j-1},\ t_j],\ \sigma_k \in [s_{k-1},\ s_k]$. Note

$$\int_{-R}^{R}\int_{-R}^{R} f(t)g(s)\mathrm{d}B_H(s)\mathrm{d}B_H(t) < \infty$$

and bounded for $R >$ (some constant). So the magnitude of the limit of the sequence

$$\sum_{k=1}^{n}\sum_{j=1}^{m} f(\alpha_j)g(\sigma_k)\Delta B_H(s_j)\Delta B_H(t_k);$$
$$\times\ \alpha_j \in [t_{j-1}, t_j], \sigma_k \in [s_{k-1}, s_k]$$

is a.s. bounded. So each finite sum term and its limit are bounded. Thus there is a constant greater than each term, which converges a.s. By using the Lebesgue Convergence Theorem[13] and manipulating,

$$\int_{\Omega}\left\{\lim_{n,m\to\infty}\sum_{k=1}^{n}\sum_{j=1}^{m} f(\alpha_j)g(\sigma_k)\Delta B_H(s_j,\omega)\Delta B_H(t_k,\omega);\right.$$
$$\left.\alpha_j \in [t_{j-1}, t_j], \sigma_k \in [s_{k-1}, s_k]\right\}\ \mathrm{d}P(\omega)$$
$$= \lim_{n,m\to\infty}\int_{\Omega}\left\{\sum_{k=1}^{n}\sum_{j=1}^{m} f(\alpha_j)g(\sigma_k)\Delta B_H(s_j,\omega)\right.$$
$$\left.\times[\Delta B_H(t_k,\omega)];\ \alpha_j \in [t_{j-1}, t_j], \sigma_k \in [s_{k-1}, s_k]\right\}\mathrm{d}P(\omega)$$
$$= \lim_{n,m\to\infty}\sum_{k=1}^{n}\sum_{j=1}^{m} f(\alpha_j)g(\sigma_k)\int_{\Omega}\Delta B_H(s_j,\omega)\Delta B_H(t_k,\omega)$$
$$\mathrm{d}P(\omega) = \lim_{n,m\to\infty}\sum_{k=1}^{n}\sum_{j=1}^{m} f(\alpha_j)g(\sigma_k)\int_{\Omega}\frac{1}{\Delta t_j\Delta t_k}$$
$$\times\ \Delta B_H(s_j,\omega)\ \Delta B_H(t_k,\omega)\mathrm{d}P(\omega)\ \Delta t_j\ \Delta s_k$$
$$\equiv \lim_{n,m\to\infty}\sum_{k=1}^{n}\sum_{j=1}^{m} f(\alpha_j)g(\sigma_k)E[\frac{1}{\Delta t_j\Delta t_k}\Delta B_H(s_j,\omega)$$
$$\times\ \Delta B_H(t_k,\omega)]\Delta t_j\Delta s_k = \lim_{n,m\to\infty}\sum_{k=1}^{n}\sum_{j=1}^{m} f(\alpha_j)g(\sigma_k)$$
$$\times\ E[\frac{1}{h_{m,n}^2}\Delta B_H(s_j,\omega)\Delta B_H(t_k,\omega)]h_{m,n}^2$$

by letting $\Delta t_j = \Delta s_k = h_{m,n} > 0$, which leaves the limit unchanged. Now note that

$$\left|\lim_{n,m\to\infty}\left\{\sum_{k=1}^{n}\sum_{j=1}^{m} f(\alpha_j)g(\sigma_k)E\left[\frac{1}{h_{m,n}^2}\Delta B_H(s_j,\omega)\right.\right.\right.$$
$$\left.\times\Delta B_H(t_k,\omega)\right]h_{m,n}^2 - \sum_{k=1}^{n}\sum_{j=1}^{m} f(\alpha_j)g(\sigma_k)R_{G_H}(\alpha_j - \sigma_k)$$

$$\times\, h_{m,n}^2 \Bigg\}\Bigg| = \Bigg|\lim_{n,m\to\infty}\sum_{k=1}^{n}\sum_{j=1}^{m} f(\alpha_j)g(\sigma_k)\Bigg\{E\Bigg[\frac{1}{h_{m,n}^2}$$

$$\times \Delta B_H(s_j,\omega)\ \Delta B_H(t_k,\omega)\Bigg] - R_{G_H}(\alpha_j-\sigma_k)\Bigg\} h_{m,n}^2\Bigg|.$$

By Corollary 2 given $\varepsilon > 0, \exists N$ where

$$\Bigg|E\left[\frac{1}{h_{m,n}^2}\Delta B_H(s_j,\omega)\ \Delta B_H(t_k,\ \omega)\right]$$
$$-R_{G_H}(\alpha_j\ -\ \sigma_k)\Bigg| < \varepsilon$$

for $n, m > N; i = 1, \ldots, m; j = 1, \ldots, n$. So

$$\Bigg|\sum_{k=1}^{n}\sum_{j=1}^{m} f(\alpha_j)g(\sigma_k)\Bigg\{E\left[\frac{1}{h_{m,n}^2}\ \Delta B_H(s_j,\omega)\ \Delta B_H(t_k,\omega)\right]$$

$$-R_{G_H}(\alpha_j - \sigma_k)\Bigg\} h_{m,n}^2\Bigg| < \varepsilon\Bigg|\sum_{k=1}^{n}\sum_{j=1}^{m} f(\alpha_j)g(\sigma_k)h_{m,n}^2\Bigg|$$

$$\to \varepsilon\left|\int_{-R}^{R}\int_{-R}^{R} f(t)g(s)\mathrm{d}s\mathrm{d}t\right| < \infty$$

because knowing

$$f, g \in L^2(R)\left|\int_{-\infty}^{\infty}\ \int_{-\infty}^{\infty} f(t)g(s)\mathrm{d}s\mathrm{d}t\right| < \infty.$$

Thus

$$\left|\sum_{k=1}^{n}\ \sum_{j=1}^{m} f(\alpha_j)g(\sigma_k)h_{m,n}^2\right|$$

is bounded. Because ε is arbitrary, this implies

$$\lim_{n,m\to\infty}\Bigg\{\sum_{k=1}^{n}\sum_{j=1}^{m} f(\alpha_j)g(\sigma_k)E\Bigg[\frac{1}{h_{m,n}^2}\Delta B_H(s_j,\omega)$$

$$\times \Delta B_H(t_k,\omega)\Bigg] h_{m,n}^2 - \sum_{k=1}^{n}\ \sum_{j=1}^{m} f(\alpha_j)g(\sigma_k)R_{G_H}$$

$$\times(\alpha_j-\sigma_k)h_{m,n}^2\Bigg\} = 0.$$

By using Triangle Inequality on absolute values of differences on the following three pairs,

$$E\left[\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(t)g(s)\mathrm{d}B_H(s)\mathrm{d}B_H(t)\right] \text{ and } \sum_{k=1}^{n}\sum_{j=1}^{m} f(\alpha_j)g(\sigma_k)$$

$$\times\, E\left[\frac{1}{h_{m,n}^2}\Delta B_H(s_j,\omega)\Delta B_H(t_k,\omega)\right] h_{m,n}^2;$$

$$\sum_{k=1}^{n}\sum_{j=1}^{m} f(\alpha_j)g(\sigma_k)E\left[\frac{1}{h_{m.n}^2}\Delta B_H(s_j,\omega)\Delta B_H(t_k,\omega)\right] h_{m,n}^2$$

$$\text{and } \sum_{k=1}^{n}\ \sum_{j=1}^{m} f(\alpha_j)g(\sigma_k)R_{G_H}(\alpha_j-\sigma_k)h_{m,n}^2;$$

and also

$$\sum_{k=1}^{n}\sum_{j=1}^{m} f(\alpha_j)g(\sigma_k)R_{G_H}(\alpha_j-\sigma_k)h_{m,n}^2$$

$$\text{and } \int_{-R}^{R}\ \int_{-R}^{R} f(t)g(s)\ R_{G_H}(t-s)\mathrm{d}s\mathrm{d}t,$$

one infers that

$$E\left[\int_{-R}^{R}\int_{-R}^{R} f(t)g(s)\mathrm{d}B_H(s)\mathrm{d}B_H(t)\right]$$
$$= \int_{-R}^{R}\ \int_{-R}^{R} f(t)g(s)\ R_{G_H}(t\ -\ s)\mathrm{d}s\mathrm{d}t$$
$$\equiv V_H H(2H-1)\int_{-R}^{R}\int_{-R}^{R} f(t)g(s)|t-s|^{2H-2}\mathrm{d}s\mathrm{d}t$$

for all $R \in (0, \infty)$. Because we assume the indefinite integrals exist, let $R \to \infty$ to obtain the desired result.

Proof of Theorem 16 (outline) We prove this by noting the conditional probability density function (pdf) is given by

$$p(\mathbf{x}|\mathbf{z}) = \frac{p_{\mathrm{X,Z}}(\mathbf{x},\mathbf{z})}{p_{\mathrm{Z}}(\mathbf{z})},$$

provided the respective pdf's exist. So for Gaussian random vectors, this quotient is well known to be

$$p(\mathbf{x}|\mathbf{z}) = \frac{K_1 \exp\left\{-(1/2)\begin{bmatrix}\mathbf{x}-E(\mathbf{X})\\ \mathbf{z}-E(\mathbf{Z})\end{bmatrix}^T\begin{bmatrix}\mathrm{P_{XX}} & \mathrm{P_{XZ}}\\ \mathrm{P_{ZX}} & \mathrm{P_{ZZ}}\end{bmatrix}^{-1}\begin{bmatrix}\mathbf{x}-E(\mathbf{X})\\ \mathbf{z}-E(\mathbf{Z})\end{bmatrix}\right\}}{K_2\exp\{-(1/2)[\mathbf{z}-E(\mathbf{Z})]^T\mathbf{P}_{\mathrm{ZZ}}^{-1}[\mathbf{z}-E(\mathbf{Z})]\}}$$

where K_1 and K_2 are well known constants in a joint Gaussian pdf. Let $K \equiv K_1/K_2$. By the Lemma,

$$\begin{bmatrix}\mathrm{P_{XX}} & \mathrm{P_{XZ}}\\ \mathrm{P_{ZX}} & \mathrm{P_{ZZ}}\end{bmatrix}^{-1} = \begin{bmatrix}\mathrm{A} & \mathrm{B}\\ \mathrm{B^T} & \mathrm{C}\end{bmatrix}$$

with $\mathbf{A}, \mathbf{B}$, and $\mathbf{C}$ as previously defined. Carrying out the quotient in the $p(\mathbf{x}|\mathbf{z})$ formula by merging exponents in the exponentials and manipulating,

$$p(\mathbf{x}|\mathbf{z}) = K\exp(-1/2)[\mathbf{x} - E(\mathbf{X}) - \mathbf{P}_{XZ}\mathbf{P}_{ZZ}^{-1}(\mathbf{z} - E(\mathbf{Z}))]^{\mathrm{T}}(\mathbf{P}_{XX} - \mathbf{P}_{XZ}\mathbf{P}_{ZZ}^{-1}\mathbf{P}_{ZX})^{-1} \times [\mathbf{x} - E(\mathbf{X}) - \mathbf{P}_{XZ}\mathbf{P}_{ZZ}^{-1}(\mathbf{z} - E(\mathbf{Z}))].$$

Notice this conditional pdf still has the formula of a Gaussian pdf and thus must also be Gaussian. So by inspection from knowing the form of a multivariate Gaussian pdf, we must have that

$$\hat{\mathbf{x}}(\mathbf{z}) \equiv E(\mathbf{X}|\mathbf{z}) = E(\mathbf{X}) + \mathbf{P}_{XZ}\mathbf{P}_{ZZ}^{-1}(\mathbf{z} - E(\mathbf{Z}))$$

and

$$\mathrm{cov}(\mathbf{X}|\mathbf{z}) \equiv E\{(\mathbf{X} - E(\mathbf{X}|\mathbf{z}))(\mathbf{X} - E(\mathbf{X}|\mathbf{z}))^{\mathrm{T}}|\mathbf{z}\} = \mathbf{P}_{XX} - \mathbf{P}_{XZ}\mathbf{P}_{ZZ}^{-1}\mathbf{P}_{ZX},$$

the desired result.

REFERENCES

1. Wong E, Hajek B. *Stochastic Processes in Engineering Systems*. New York: Springer-Verlag; 1985.
2. Barton RJ, Poor HV. Signal detection in fractional Gaussian noise. *IEEE Trans Inform Theory* 1988, 43: 943–959.
3. Doob JL. *Stochastic Processes*. New York: John Wiley & Sons; 1953.
4. Mandelbrot BB, Van Ness JW. Fractional Brownian motions, fractional noises and applications. *SIAM Rev* 1968, 10:422–437.
5. Mandelbrot BB. *The Fractal Geometry of Nature (Updated and Augmented)*. New York: W. H. Freeman; 1983.
6. Barnsley M. *Fractals Everywhere*. San Diego, CA: Academic Press; 1988, Chapter 5.
7. Falconer K. *Fractal Geometry: Mathematical Foundations and Applications*. Chichester: John Wiley & Sons; 1990, Chapter 16.
8. Schroeder M. *Fractals, Chaos, Power Laws: Minutes from an Infinite Paradise*. New York: W. H. Freeman; 1991.
9. Chow WC. *Modeling and Inference with Fractional Brownian motion and Fractional Gaussian Noise*, Ph.D. Dissertation, School of Information Technology and Engineering, George Mason University, Fairfax, VA, Spring, 1994.
10. Meditch JS. *Stochastic Optimal Linear Estimation and Control*. New York: McGraw-Hill; 1969.
11. Chow WC, Schumacher PW Jr. Using fractional Gaussian noise models in orbit determination. *J Guid Cont Dyn* 2003, 25:593–607.
12. Olmsted JMH. *Real Variables*. New York: Appleton-Century-Crofts; 1959.
13. Royden HL. *Real Analysis*. 2nd ed. New York: MacMillan Publishing Co.; 1968.

