Temporal aggregation of stationary and nonstationary continuous-time processes

(SHORT RUNNING TITLE: AGGREGATION OF CONTINUOUS-TIME PROCESSES) *†

Henghsiu Tsai ‡and K. S. Chan
Academia Sinica and University of Iowa

November 25, 2003

We study the autocorrelation structure of aggregates from a continuous-time process. The underlying continuous-time process or some of its higher derivative is assumed to be a stationary Continuous-time Auto-Regressive Fractionally Integrated Moving-Average (CARFIMA) process with Hurst parameter H. We derive closed-form expressions for the limiting autocorrelation function and the normalized spectral density of the aggregates, as the extent of aggregation increases to infinity. The limiting model of the aggregates, after appropriate number of differencing, is shown to be some functional of the standard fractional Brownian motion with the same Hurst parameter of the continuous-time process from which the aggregates are measured. These results are then used to assess the loss of forecasting efficiency due to aggregation.

---

*AMS 2000 subject classifications. Primary 62M10; Secondary 60G18

†Key words and phrases. Asymptotic efficiency of prediction, Autocorrelation, CARFIMA models, Fractional Brownian motion, Long memory, Spectral density.

‡Supported by Academia Sinica and the National Science Council (NSC 92-2118-M-001-018), R.O.C.

1 Introduction

Continuous-time processes are often measured over regular or irregular time intervals. Two common measurement models are (i) instantaneous sampling where the (continuous-time) process is sampled instantaneously at the sampling epochs, e.g. instantaneous river-flow measured at a particular date in each year, and (ii) aggregation which integrates the process over the sampling intervals, e.g. annual tree ring growth. The discrete-time data so obtained are also referred to as the stock and flow variables, see Harvey (1990, p. 506). Here, we study the correlation structure of the aggregates from a (centered) continuous-time process.

With a fixed, finite extent of aggregation, the correlation structure of the aggregates would depend on the exact correlation structure of the continuous-time process. However, with increasing aggregation, a simpler limiting model for the aggregates may emerge. For example, for continuous-time process of short-memory, the Central Limit Theorem implies that the aggregates become white noises asymptotically with increasing aggregation. However, the situation is rather unclear if the underlying process is non-stationary or of long-memory, which is our focus here. In the discrete-time setting, similar problems have been studied by Tiao (1972), Beran & Ocker (working paper, 2000, available at *http : //netec.mcc.aac.uk/WoPEc/data/knzcofedp.html*) and Man & Tiao (working paper, 2001).

A unified framework for studying non-stationary and/or long-memory continuous-time process $\{Y_t\}$ is to assume that its r-th derivative process, denoted by $\{Y_t^{(r)}\}$, is a stationary Continuous-time Auto-Regressive Fractionally Integrated Moving-Average (CARFIMA) process with Hurst parameter H. See section 2 for an overview of CARFIMA models. Note that the zero-th derivative of $\{Y_t\}$ is $\{Y_t\}$ itself. If $\{Y_t^{(r)}\}$ has the Hurst parameter $0 < H < 1$, the sum $r + H$ is referred to as its (fractional) integration order. The process is stationary if and only if the integration order is less than one. In the stationary case, i.e., $r = 0$, the process is of long memory for $1/2 < H < 1$, of short memory if $H = 1/2$, and anti-persistent for $0 < H < 1/2$.

The aggregates are obtained by integration via the formula $\int_{(n-1)\Delta}^{n\Delta} Y_u du$, $n \in Z$, where Δ is the aggregation interval length. Here, we consider the correlation structure of the aggregates when $\Delta \longrightarrow \infty$. For $r \geq 1$, the aggregates require r-th differencing to achieve stationarity. In section 3 we derive for the r-th differenced aggregates the limiting normalized spectral density function, which preserves the Hurst parameter but is independent of the short-memory component of the underlying continuous-time process. Using the self-similarity property of the fractional Brownian process, we furthermore express the limiting model of the aggregates as some functional of a fractional Brownian process with the same Hurst parameter of the underlying process. This limiting process representation enables us to derive in section 3 a closed-form solution of the limiting auto-correlation function. These results will be useful for modeling aggregates from continuous-time processes.

Finally, in section 4, we consider the forecasting of aggregates. If the complete past trajectory of the underlying process is available, the future aggregates can be forecasted more accurately than just using past aggregates. However, it is often expensive to measure all the past trajectory of a continuous-time process while aggregated data are more readily available. A natural question arises as to how much forecasting efficiency can be gained by using the basic continuous-time process relative to that based on aggregated data. For one-step prediction, we show that, for the stationary case, the loss of forecasting efficiency due to aggregation is generally less than 10% and at most 20%. However, the loss of forecasting efficiency increases drastically with the (fractional) integration order in the non-stationary case.

2 A brief review of the CARFIMA models

We now briefly review CARFIMA processes. For further details, see Tsai & Chan (working paper, 2003, available at http://www.stat.uiowa.edu/techrep). See also Chambers (1996, 1998), Comte (1996), Comte & Renault (1996), Kleptsyna *et al.* (2000), and Viano *et al.* (1994) for earlier related works. Heuristically, a centered CARFIMA(p,H,q) process $\{Y_t\}$ is defined as the solution of a p-th order stochastic differential equation with suitable initial conditions and driven by a standard fractional Brownian motion with Hurst parameter H and its derivatives up to and including order $0 \leq q < p$. Specifically, for $t \geq 0$,

$$Y_t^{(p)} - \alpha_p Y_t^{(p-1)} - \cdots - \alpha_1 Y_t = \sigma\{B_{t,H}^{(1)} + \beta_1 B_{t,H}^{(2)} + \cdots + \beta_q B_{t,H}^{(q+1)}\}, \qquad (1)$$

where $0 \leq q < p$ and $\{B_{t,H} = B_t^H, t \geq 0\}$ is the standard fractional Brownian motion with Hurst parameter $0 < H < 1$; recall the superscript $^{(j)}$ denotes j-fold differentiation with respect to t. We assume that $\sigma > 0$ and $\beta_q \neq 0$, $dY_t^{(j-1)} = Y_t^{(j)} dt, j = 1, ..., p-1$. Equation (1) can be succinctly represented as $\alpha(D)Y_t = \sigma\beta(D)DB_t^H$, where $\alpha(z) = z^p - \alpha_p z^{p-1} - \cdots - \alpha_1$, $\beta(z) = 1 + \beta_1 z + \beta_2 z^2 + \cdots + \beta_q z^q$.

We now recall the definition of fractional Brownian motions. Let $0 < H < 1$ be a fixed number. It is well known (see, e.g., Duncan *et al.*, 2000) that there exists a Gaussian stochastic process $\{B_t^H, t \geq 0\}$ satisfying the following three properties, namely (i) with the initial condition $B_0^H = 0$, (ii) of zero mean, i.e. $E(B_t^H) = 0$ for all $t \geq 0$, and (iii) with the covariance kernel defined as

$$E(B_t^H B_s^H) = \frac{1}{2}\{|t|^{2H} + |s|^{2H} - |t-s|^{2H}\}, \tag{2}$$

for all $s, t \geq 0$. The Gaussian process $\{B_t^H\}$ is called the (standard) fractional Brownian motion with Hurst parameter H. The standard fractional Brownian motion with $H = 1/2$ equals the standard Brownian motion.

The fractional Brownian motion is nowhere differentiable (Mandelbrot & Van Ness, 1968), so the stochastic equation (1) is interpreted as some stochastic integral equation, specifically in terms of the *observation* and *state* equations (see Tsai and Chan, 2003 and Brockwell, 1993):

$$Y_t = \beta^T X_t, \qquad t \geq 0, \tag{3}$$

$$dX_t = AX_t dt + \sigma\delta_p dB_t^H, \tag{4}$$

where the superscript ^{T} denotes taking transpose,

$$A = \begin{bmatrix} 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & 1 \\ \alpha_1 & \alpha_2 & \alpha_3 & \cdots & \alpha_p \end{bmatrix}, \quad X_t = \begin{bmatrix} X_t^{(0)} \\ X_t^{(1)} \\ \vdots \\ X_t^{(p-2)} \\ X_t^{(p-1)} \end{bmatrix}, \quad \delta_p = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ 1 \end{bmatrix}, \quad \beta = \begin{bmatrix} 1 \\ \beta_1 \\ \vdots \\ \beta_{p-2} \\ \beta_{p-1} \end{bmatrix},$$

and $\beta_j = 0$ for $j > q$.

The process $\{Y_t, t \geq 0\}$ is said to be a CARFIMA(p,H,q) process with parameter $(\theta, \sigma) = (\alpha_1, ..., \alpha_p, \beta_1, \cdots, \beta_q, H, \sigma)$ if $Y_t = \beta^T X_t$, where X_t is the solution of (4) with the initial condition X_0. The state equation has the unique solution

$$X_t = e^{At}X_0 + \sigma\int_0^t e^{A(t-u)}\delta_p dB_u^H, \tag{5}$$

where $e^{At} = I + \sum_{n=1}^{\infty}\{(At)^n(n!)^{-1}\}$, and I is the identity matrix. Equation (4) with a deterministic initial condition admits an asymptotically weakly stationary solution if and only if all the eigenvalues of A have negative real parts. In the stationary case, $\{X_t\}$ and hence $\{Y_t\}$ can be extended to be stationary processes that are defined for all real t. Specifically, for $t \in R$,

$$X_t = \sigma \int_{-\infty}^{t} e^{A(t-u)}\delta_p dB_u^H; \tag{6}$$

for a construction of $\{B_t^H, t \in R\}$, see Mandelbrot & Van Ness (1968) and Norros *et al.* (1999).

The spectral density function of $\{Y_t\}$ can be shown to be

$$f_Y(w) \quad = \quad \frac{\sigma^2}{2\pi}\Gamma(2H+1)\sin(\pi H)|w|^{1-2H}\frac{|\beta(iw)|^2}{|\alpha(iw)|^2}, \tag{7}$$

where $\Gamma()$ is the Gamma function. Let $g(w) = \sigma^2(2\pi)^{-1}\Gamma(2H+1)\sin(\pi H)|\beta(iw)|^2/|\alpha(iw)|^2$. Then, $f_Y(w) = g(w)|w|^{1-2H}$, so that, as $|w| \to 0$, $f_Y(w) \sim g(0)|w|^{1-2H}$, i.e. the spectral density is asymptotically linear with slope $1-2H$ on the log-log scale for $|w|$ small.

3 Limiting behavior of aggregates of a continuous time process

Let $\{Y_t, t \in R\}$ be a continuous-time process. We assume that its r-th derivative process is a stationary CARFIMA(p, H, q) model, i.e.,

$$\alpha(D)D^r Y_t = \sigma\beta(D)DB_t^H, \qquad t \in R. \tag{8}$$

Note that $\{Y_t\}$ is a stationary process if and only if $r = 0$. Consider the case that the regular continuous-time process is digitalized by aggregation over intervals of length Δ, i.e.,

$$Y_n^{\Delta} \quad = \quad \int_{(n-1)\Delta}^{n\Delta} Y_u du. \tag{9}$$

Then $\{Y_n^{\Delta}, n \in Z\}$ is stationary if and only if $r = 0$. For $r > 0$, it can be readily checked that the aggregates can be transformed to stationarity by r-fold differencing. Here, we study the correlation structure of the r-differenced series

$\{\nabla^r Y_n^{\Delta}\}$, where ∇ is the differencing operator. For conciseness, we write $Y_n^{r,\Delta}$ for $\nabla^r Y_n^{\Delta}$ below. Write the lag-h auto-correlation of the differenced aggregates by $\rho_\Delta(h)$, and $\rho_\infty(h) = \lim_{\Delta\to\infty} \rho_\Delta(h)$. Our main results derive a closed-formed expression for the limiting autocorrelation function $\rho_\infty(h)$ and give a stochastic representation that realizes the limiting autocorrelation function.

Due to the central limit effect, the short-memory structure of the CARFIMA model may be expected to vanish with increasing aggregation, i.e., $\Delta \longrightarrow \infty$. Indeed, this is true by examining the limiting form of the spectral density function of the Δ-aggregates, as is shown in the following theorem.

THEOREM 1 *Assume that $\{Y_t^{(r)}\}$ is a stationary solution of (1).*
(a) For $r \geq 0$, the spectral density function of $\{Y_n^{r,\Delta}\}$ is given by

$$f_{r,\Delta}(w) \quad = \quad \frac{1}{\Delta}\{2(1-\cos w)\}^{r+1} \sum_{k=-\infty}^{\infty} \left|\frac{w+2k\pi}{\Delta}\right|^{-2r-2H-1} g\left(\frac{w+2k\pi}{\Delta}\right), \quad (10)$$

where $-\pi < w < \pi$ and $g(w)$ is defined below (7).
(b) As $\Delta \to \infty$, the normalized spectral density function of $\{Y_n^{r,\Delta}\}$ converges to

$$f_r(w) \quad = \quad K\{2(1-\cos w)\}^{r+1} \sum_{k=-\infty}^{\infty} |w+2k\pi|^{-2r-2H-1}, \quad (11)$$

where K is the normalization constant ensuring that $\int_{-\pi}^{\pi} f_r(w)dw = 1$.

Note (i) both $f_{r,\Delta}$ and f_r are of $O(w^{1-2H})$ for $|w| \longrightarrow 0$, so that the limit of the Δ-aggregates has the same Hurst parameter as the underlying CARFIMA process, and (ii) the limiting normalized spectral density function is independent of the short-memory parameter, confirming the central limit effect. It can be seen from the proof of Theorem 1 that part (b) of Theorem 1 actually holds for any continuous-time process whose spectral density function is of the form $w^{1-2H}g(w)$, where g is a bounded, integrable function that is continuous at $w=0$, with $g(0) > 0$. The convergence of the normalized spectral density functions of the Δ-aggregates ensures the convergence of the corresponding autocorrelation functions, and hence after suitable re-scaling the Δ-aggregates converge to some limiting Gaussian process, in finite-finite-dimensional distributions; the re-scaling is a function of Δ and is proportional to Δ^{-r-H}, as can be seen from the proofs of the following theorem.

For the case $r = 0$, the limiting normalized spectral density function coincides with that of the fractional Gaussian noise with the Hurst parameter H, see Beran (1994, p. 53). In other words, for large Δ, and after appropriate scaling, the Δ-aggregates are asymptotically distributed as $\{B_n^H - B_{n-1}^H\}$. In the general case when $r \geq 1$, we can make use of the self-similarity property of the fractional Gaussian process to derive the form of the limiting process as some functional of the fractional Gaussian process. Furthermore, in the following theorem, we derive the autocorrelation function of the limit of the r-th differenced Δ-aggregates of a CARFIMA$(p, r+H, q)$ process.

THEOREM 2 *Assume* $0 < H < 1$. *(a) For* $r \geq 0$, *and in terms of distributions,*

$$Y_n^{r,\Delta} = \sigma\beta' \sum_{k=0}^{r+1}\sum_{j=0}^{r} \binom{r+1}{k} \frac{(-1)^{k+1}}{(r-j)!} \Delta^{H+r-j} A^{-j-1}\delta_p \int_0^{(n-k)} (n-k-v)^{r-j} dB_v^H. \quad (12)$$

(b) When Δ *is large, the limiting model for* $\{-\Delta^{-H-r}\alpha_1 Y_n^{r,\Delta}/\sigma, n \in Z\}$ *is distributed as that of*

$$\left\{\sum_{j=0}^{r}\int_{n-j-1}^{n-j}\sum_{k=0}^{j}\binom{r+1}{k}\frac{(-1)^{k+1}}{r!}(n-k-u)^r dB_u^H, n \in Z\right\}, \quad (13)$$

or equivalently, the limiting model equals

$$\left\{\int_{n-r-1}^{n} f_{n,r}(u) dB_u^H, n \in Z\right\}, \quad (14)$$

where $f_{n,r}(u)$ *is the probability density function of* $U_{n,r} = n - r - 1 + \sum_{i=1}^{r+1} U_i$, *where* $\{U_i\}_{i=1}^{r+1}$ *is a sequence of independently and identically distributed uniform (0,1) random variables.*

(c) For $0 < H < 1$, *the limiting autocorrelation function of* $\{Y_n^{r,\Delta}\}$ *is given by*

$$\begin{aligned}\rho_\infty(h) &= K \sum_{k=-r-1}^{r+1} (-1)^{r+k+1}\binom{2r+2}{k+r+1}|h+k|^{2H+2r} \\ &= K\sum_{k=0}^{2r+2}(-1)^k\binom{2r+2}{k}|r+1-h-k|^{2H+2r}, \end{aligned} \quad (15)$$

where K *is a normalization constant ensuring that* $\rho_\infty(0) = 1$.

(d) For $H = 1/2$, *the limiting model for* $\{-\Delta^{-H-r}\alpha_1 Y_n^{r,\Delta}/\sigma, n \in Z\}$ *is the ARIMA(0,r,r) model. The corresponding limiting autocorrelation function given by equation (15) can be simplified to*

$$\rho_\infty(h) = \frac{K}{(2r+1)!}\sum_{k=0}^{r-h}(-1)^k\binom{2r+2}{k}(r+1-h-k)^{2r+1}, \quad \textit{if } |h| \leq r, \quad (16)$$

and 0 *otherwise, where* K *is a normalization constant ensuring that* $\rho_\infty(0) = 1$.

Equation (16) is a special case of equation (15), by Equation (0.154.6) of Gradshteyn & Ryzhik (1994); namely, for any integers N, p, real number α and $N \geq p \geq 1$,

$$\sum_{k=0}^{N}(-1)^k\binom{N}{k}(\alpha+k)^{p-1}=0. \tag{17}$$

Equations (15) and (16) together provide an asymptotic framework for studying the long-memory properties of aggregate time series data. Below, we give some examples illustrating the preceding theorem.

Example 1: The limiting model derived in Theorem 2 (b) becomes the fractional Gaussian noise, i.e., $\{B_n^H - B_{n-1}^H\}$, for $r=0$. If $r=1$, the limiting model simplifies to $\{-\int_{n-2}^{n-1}(n-u-2)dB_u^H + \int_{n-1}^{n}(n-u)dB_u^H\}$.

Example 2: The limiting autocorrelation function of the aggregates takes the form $\rho_\infty(h) = K\{|h-1|^{2H} - 2|h|^{2H} + |h+1|^{2H}\}$, for $r=0$. For $r=1$, $\rho_\infty(h) = K\{|h-2|^{2H+2} - 4|h-1|^{2H+2} + 6|h|^{2H+2} - 4|h+1|^{2H+2} + |h+2|^{2H+2}\}$. For r=2, $\rho_\infty(h) = K\{|h-3|^{2H+4} - 6|h-2|^{2H+4} + 15|h-1|^{2H+4} - 20|h|^{2H+4} + 15|h+1|^{2H+4} - 6|h+2|^{2H+4} + |h+3|^{2H+4}\}$.

4 Forecasting efficiency of aggregate series

Suppose we are interested in doing one step ahead prediction of the future aggregate observation Y_{n+1}^Δ. This can be done by using the aggregate series $\{Y_k^\Delta, k \leq n\}$. In principle, more accurate prediction can be obtained from the basic continuous-time process $\{Y_t, t \leq n\Delta\}$ if it is available. This raises the issue of what is the possible gain one may obtain in the prediction using the basic series as compared to just using past and current aggregates. Let $\sigma_{1,r,H}^2(\Delta)$ be the one-step mean square prediction variance of Y_{n+1}^Δ by using the aggregate series $\{Y_k^\Delta, k \leq n\}$ and $\sigma_{2,r,H}^2(\Delta)$ be the corresponding mean square prediction variance obtained by using the basic continuous-time process. Following Tiao (1972), we define the limiting efficiency of the one-step ahead prediction using the basic series as compared to that of using the temporally aggregated series to be the limiting variance ratio $\xi_1(r,H) = \lim_{\Delta\longrightarrow\infty}\{\sigma_{1,r,H}^2(\Delta)/\sigma_{2,r,H}^2(\Delta)\}$.

THEOREM 3 *For any non-negative integer r and $0 < H < 1$, the limiting* *prediction variance ratio equals*

$$\xi_1(r,H) = 2^{r+1}(2r+2H)\Gamma^2(r+H+1/2) \times \exp\left\{ \frac{r+1}{2\pi}\int_{-\pi}^{\pi} \log(1-\cos\lambda)d\lambda + \frac{1}{2\pi}\int_{-\pi}^{\pi} \log \sum_{k=-\infty}^{\infty} |w+2k\pi|^{-2r-2H-1}dw \right\}.$$

The computation of $\xi_1(r,H)$ requires evaluating the sum $\sum_{k=-\infty}^{\infty}|w+2k\pi|^{-2r-2H-1}$. If $H=1/2$, it is possible to evaluate this infinite sum using well-known methods, see Chambers (1996). For example, by equations (822) and (824) of Jolley (1961), we have the following two simplifications:

$$\xi_1(0,1/2)=1; \quad \xi_1(1,1/2)=\exp\left\{\frac{1}{2\pi}\int_{-\pi}^{\pi}\log\left\{3-2\sin^2\left(\frac{w}{2}\right)\right\}dw\right\}.$$

But with non-integer values of the exponent, some approximation method is required to compute the series. Here we adopt the method of Chambers (1996) as follows. First note that

$$\frac{1}{2\pi}\int_{-\pi}^{\pi}\log\sum_{k=-\infty}^{\infty}|w+2k\pi|^{-2r-2H-1}dw = (-2r-2H-1)\log(2\pi)+2\int_0^{1/2}\log\sum_{k=-\infty}^{\infty}|y+k|^{-2r-2H-1}dy.$$

The series $\sum_{k=-\infty}^{\infty}|y+k|^{-2r-2H-1}$ can be approximated by $\sum_{j=1}^{M}(j-y)^{-2r-2H-1} + \sum_{j=0}^{M}(j+y)^{-2r-2H-1} + (2r+2H)^{-1}\{(M-y)^{-2r-2H}+(M+y)^{-2r-2H}\}$ for some large M. The results summarized in Table 1 are based on $M=10,000$. Indeed, the results based on $M=100$ are essentially the same, suggesting that setting $M=10,000$ provides an adequate approximation to the series.

Table 1 shows the variance ratio $\xi_1(r,H)$ for various H and r. Note that the limiting variance ratios for $r+H=0.5,1.5,2.5,3.5,4.5$ are almost identical to those for $d=0,1,2,3,4$ of Tiao (1972), respectively. Indeed, the fractional differencing parameter d in a discrete-time ARFIMA model relates to the fractional integration order in a continuous-time CARFIMA model by the equation $d=r+H-1/2$. The values of $\xi_1(r,H)$ for $r+H=0.9999, 1.9999, 2.9999, 3.9999$ are close to but larger than those for $d=0.5,1.5,2.5,3.5$ of Man & Tiao (working paper, 2001), respectively. In our case, we are not able to compute the ratio for $H+r\in Z$ because D_H in Theorem 3 is undefined for $H=1$. Indeed, for $H=1$, the fractional Brownian process becomes deterministic, see Beran (1994, p. 53). It

Table 1: *Values of the one-step prediction variance ratio* $\xi_1(r, H)$ *for various* H *and* r.

$H\backslash r$	0	1	2	3	4
0.0001	1.001	1.181	3.721	25.508	307.970
0.1000	1.087	1.267	4.374	31.974	405.347
0.2000	1.058	1.375	5.182	40.306	535.958
0.3000	1.027	1.507	6.182	51.080	711.646
0.4000	1.007	1.669	7.427	65.070	948.832
0.5000	**1.000**	**1.866**	**8.982**	**83.311**	**1270.197**
0.6000	1.007	2.106	10.932	107.190	1707.161
0.7000	1.028	2.398	13.390	138.580	2303.387
0.8000	1.063	2.753	16.501	180.005	3119.724
0.9000	1.114	3.188	20.455	234.888	4241.231
0.9999	1.181	3.718	25.490	307.717	5783.730

appears that the variance ratio $\xi_1(r, H)$ increases with $r + H$, the integration order, for $r + H \geq 1/2$.

Table 1 indicates that for the stationary case, i.e., $0 < H + r < 1$, the gain of forecasting efficiency by using the basic series as compared to just using the aggregates is at most 20% and generally less than 10% for $0 < H + r <= 0.80$. However, the efficiency gain increase rapidly with $H+r$ for nonstationary processes.

Acknowledgement

We thank Academia Sinica, the University of Iowa and the National Science Council (NSC 92-2118-M-001-018), R.O.C. for partial support.

Appendix 1

Proof of Theorem 1

Proof of (a) First note that $Y_t^{(r)}$ admits the following spectral representation (Priestley, 1981, Theorem 4.11.1): $Y_t^{(r)} = \int_{-\infty}^{\infty} e^{itw} dZ(w)$ with $E|dZ(w)|^2 = f_Y(w)dw$ where $f_Y()$ is given by (7). Routine calculus shows that, for $r \geq 1$ and $t \geq 0$, $Y_t = p(t) + \int_0^t (t-u)^{r-1} Y_u^{(r)} du/(r-1)!$, for $t < 0$, $Y_t = p(t) + (-1)^r \int_t^0 (u -$

$t)^{r-1}Y_u^{(r)}du/(r-1)!$ where $p(t) = \sum_{j=0}^{r-1} Y_0^{(j)} t^j/j!$. Because $p(t)$ vanishes upon r-th differencing, we can assume $Y_0^{(j)} = 0$ for all $0 \le j \le r-1$, without loss of generality. Thus, for $r \ge 0$ and $n \ge 1$, $Y_n^{\Delta} = Q_n - Q_{n-1}$, where

$$\begin{aligned} Q_n &= \frac{1}{r!}\int_0^{n\Delta} (n\Delta - u)^r Y_u^{(r)} du \qquad \text{(A1)} \\ &= \frac{1}{r!}\int_{-\infty}^{\infty}\int_0^{n\Delta} (n\Delta - u)^r e^{iuw} du dZ(w) \\ &= \int_{-\infty}^{\infty} \left\{ \frac{1}{(iw)^{r+1}} e^{in\Delta w} - \sum_{j=1}^{r} \frac{(n\Delta)^j}{j!(iw)^{r-j+1}} \right\} dZ(w). \end{aligned}$$

Therefore, for $n \ge r+1$,

$$\begin{aligned} Y_n^{r,\Delta} &= \nabla^{r+1} Q_n \\ &= \sum_{k=0}^{r+1} \binom{r+1}{k} (-1)^k Q_{n-k} \qquad \text{(A2)} \\ &= \int_{-\infty}^{\infty} \frac{1}{(iw)^{r+1}} \sum_{k=0}^{r+1} \binom{r+1}{k} (-1)^k e^{i(n-k)\Delta w} dZ(w) \\ &\quad - \int_{-\infty}^{\infty} \sum_{j=1}^{r}\sum_{k=0}^{r+1} \binom{r+1}{k} (-1)^k \frac{((n-k)\Delta)^j}{j!(iw)^{r-j+1}} dZ(w) \\ &= I - \Pi. \end{aligned}$$

Equation (17) implies that $\Pi = 0$, hence

$$\begin{aligned} Y_n^{r,\Delta} &= \int_{-\infty}^{\infty} \frac{1}{(iw)^{r+1}} \sum_{k=0}^{r+1} \binom{r+1}{k} (-1)^k e^{i(n-k)\Delta w} dZ(w) \\ &= \sum_{k=-\infty}^{\infty} \int_{(2k-1)\pi/\Delta}^{(2k+1)\pi/\Delta} (iw)^{-r-1} e^{in\Delta w} (1 - e^{-i\Delta w})^{r+1} dZ(w) \\ &= \sum_{k=-\infty}^{\infty} \int_{-\pi}^{\pi} \left\{ i\left(\frac{w+2k\pi}{\Delta}\right)\right\}^{-r-1} e^{inw}(1-e^{-iw})^{r+1} dZ\left(\frac{w+2k\pi}{\Delta}\right) \\ &= \int_{-\pi}^{\pi} e^{inw} dZ_\Delta(w), \end{aligned}$$

where

$$dZ_\Delta(w) = \sum_{k=-\infty}^{\infty} \left\{ i\left(\frac{w+2k\pi}{\Delta}\right)\right\}^{-r-1} (1-e^{-iw})^{r+1} dZ\left(\frac{w+2k\pi}{\Delta}\right), \qquad |w| \le \pi.$$

Validity of the above equalities for all integer n follows from stationarity. Thus, by (4.11.19) of Priestley (1981),

$$\begin{aligned} f_{r,\Delta}(w)dw &= E\left(|dZ_\Delta(w)|^2\right) \\ &= \frac{1}{\Delta}\{2(1-\cos w)\}^{r+1} \sum_{k=-\infty}^{\infty} \left|\frac{w+2k\pi}{\Delta}\right|^{-2r-2H-1} g\left(\frac{w+2k\pi}{\Delta}\right) dw, \end{aligned}$$

$-\pi < w < \pi$. This proves the result.

Proof of (b) Consider

$$\Delta^{-2r-2H} f_{r,\Delta}(w) = \{2(1-\cos(w)\}^{r+1} \sum_{k=-\infty}^{\infty} |w+2k\pi|^{-2r-2H-1} g\left(\frac{w+2k\pi}{\Delta}\right),$$

which tends to $\{2(1-\cos(w)\}^{r+1} \sum_{k=-\infty}^{\infty} |w+2k\pi|^{-2r-2H-1} g(0)$ by the dominated convergence theorem, owing to boundedness of g and the fact that $\sum_{k=-\infty}^{\infty} |w+2k\pi|^{-2r-2H-1} < \infty$ for $r \geq 0$ and $0 < H < 1$. The convergence of the normalization constants of $f_{r,\Delta}$ to K follows along similar arguments.

Proof of Theorem 2

Proof of (a) It follows from equations (A1) and (A2) that

$$Y_n^{r,\Delta} = \frac{1}{r!} \sum_{k=0}^{r+1} \binom{r+1}{k} (-1)^k \int_0^{(n-k)\Delta} ((n-k)\Delta - u)^r Y_u^{(r)} du.$$

However, $Y_u^{(r)}$ can be written as $Y_u^{(r)} = \beta^T X_u$, where X_u is given by equation (5). With no loss of generality, we assume $X_0 = 0$ as it has negligible effects for large t, then, for $n \geq r+1$,

$$\begin{aligned}
& Y_n^{r,\Delta} \\
= \ & \frac{\sigma}{r!} \beta^T \sum_{k=0}^{r+1} \binom{r+1}{k} (-1)^k \int_0^{(n-k)\Delta} ((n-k)\Delta - u)^r \int_0^u e^{A(u-v)} \delta_p dB_v^H du \\
= \ & \frac{\sigma}{r!} \beta^T \sum_{k=0}^{r+1} \binom{r+1}{k} (-1)^k \int_0^{(n-k)\Delta} \int_v^{(n-k)\Delta} ((n-k)\Delta - u)^r e^{A(u-v)} \delta_p du dB_v^H \\
= \ & -\frac{\sigma}{r!} \beta^T \sum_{k=0}^{r+1} \binom{r+1}{k} (-1)^k \sum_{j=0}^{r} \frac{r!}{(r-j)!} \int_0^{(n-k)\Delta} ((n-k)\Delta - v)^{r-j} A^{-j-1} \delta_p dB_v^H.
\end{aligned}$$

Thus, by the self similarity property of the fractional Brownian motion, $Y_n^{r,\Delta}$ has the same distribution as

$$\sigma \beta^T \sum_{k=0}^{r+1} \sum_{j=0}^{r} \binom{r+1}{k} \frac{(-1)^{k+1}}{(r-j)!} \Delta^{H+r-j} A^{-j-1} \delta_p \int_0^{n-k} (n-k-v)^{r-j} dB_v^H.$$

This proves the result.

Proof of (b) Note that $\beta^T A^{-1} \delta_p = \alpha_1^{-1}$ and the dominating term of $Y_n^{r,\Delta}$ is

$$\Delta^{H+r} \sigma \beta^T A^{-1} \delta_p \sum_{k=0}^{r+1} \binom{r+1}{k} \frac{(-1)^{k+1}}{r!} \int_0^{n-k} (n-k-u)^r dB_u^H$$

$$
\begin{aligned}
&= \frac{\sigma}{\alpha_1}\Delta^{H+r}\sum_{k=0}^{r+1}\binom{r+1}{k}\frac{(-1)^{k+1}}{r!}\int_0^{n-r-1}(n-k-u)^r dB_u^H \\
&\quad +\frac{\sigma}{\alpha_1}\Delta^{H+r}\sum_{k=0}^{r+1}\binom{r+1}{k}\frac{(-1)^{k+1}}{r!}\sum_{j=n-r-1}^{n-k-1}\int_j^{j+1}(n-k-u)^r dB_u^H \\
&= I+\Pi,
\end{aligned}
$$

where $I=0$ by equation (17). The result follows by stationarity and the fact that

$$
\Pi = \frac{\sigma}{\alpha_1}\Delta^{H+r}\sum_{j=n-r-1}^{n-1}\int_j^{j+1}\sum_{k=0}^{n-j-1}\binom{r+1}{k}\frac{(-1)^{k+1}}{r!}(n-k-u)^r dB_u^H.
$$

The alternative representation (14) follows from the following lemma; see Stuart & Ord (1994, p. 387) for a proof of the Lemma.

Lemma 4 *Let $Y=n-r-1+\sum_{i=1}^{r+1}U_i$, where $\{U_i\}_{i=1}^{r+1}$ is a sequence of independent and identically distributed uniform (0,1) random variables, then the probability density function of Y is given by*

$$
f_{n,r}(y)=\frac{1}{r!}\sum_{k=0}^{j}\binom{r+1}{k}(-1)^k(n-y-k)^r, \tag{A3}
$$

for $n-j-1\le y\le n-j$ and $j=0,1,\cdots,r$, and zero elsewhere.

Proof of (c) Consider the case that $1/2<H<1$. For any integer h,

$$
\begin{aligned}
\rho_\infty(h) &= \mathrm{Cov}\left(\int_{n+h-r-1}^{n+h} f_{n+h,r}(u)dB_u^H, \int_{n-r-1}^{n} f_{n,r}(v)dB_v^H\right) \\
&= C_H\int_{n-r-1}^{n}\int_{n+h-r-1}^{n+h} f_{n+h,r}(u)f_{n,r}(v)|u-v|^{2H-2}dudv \\
&= C_H\mathrm{E}\left[|U_{n+h,r}-U_{n,r}|^{2H-2}\right] \\
&= C_H\mathrm{E}\left[\left|h-r-1+\sum_{i=1}^{2r+2}U_i\right|^{2H-2}\right],
\end{aligned} \tag{A4}
$$

where throughout this appendix $\{U_i\}$ denotes a sequence of independent uniform (0,1) random variables. Let $W_{r,h}=h+\sum_{i=1}^{2r}U_i$. Routine calculus show that for $H>-1$, any integer h and any non-negative integer r,

$$
\mathrm{E}\left(|W_{r,h}|^H\right)=\left\{\prod_{k=1}^{2r}(H+k)\right\}^{-1}\sum_{k=0}^{2r}(-1)^k\binom{2r}{k}|h+k|^{H+2r}. \tag{A5}
$$

The proof of (c) for $1/2<H<1$ now follows from equations (A4) and (A5). The proof for the case that $0<H<1/2$ is similar and hence omitted. For a proof for $H=1/2$, see part (d) below.

Proof of (d) This follows from lemma 4 and the fact that for $H = 1/2$,

$$\begin{aligned}\gamma_\infty(h) &= \operatorname{cov}\left(\int_{n+h-r-1}^{n+h} f_{n+h,r}(u)dB_u^H, \int_{n-r-1}^{n} f_{n,r}(v)dB_v^H\right) \\ &= \int_{n+h-r-1}^{n} f_{n+h,r}(u)f_{n,r}(u)du \\ &= \int_{n+h-r-1}^{n} f_{n,r}(u-h)f_{n,r}(u)du \\ &= f_X(-h),\end{aligned}$$

where $f_X(\cdot)$ is the probability density function of $X = \sum_{i=1}^{2r+2} U_i - r - 1$.

Proof of Theorem 3

The theorem follows readily from the expressions of $\sigma^2_{i,r,H}(\Delta)$, $i = 1, 2$, to be derived below. Routine algebra shows that, for $r \geq 1$, the one-step prediction error of $Y^{r,\Delta}_{n+1}$ equals that of Y^{Δ}_{n+1}. Hence, the one-step mean square prediction variance of Y^{Δ}_{n+1} by using the aggregate series $\{Y^{\Delta}_k, k \leq n\}$ can be computed by Kolmogorov's formula (Theorem 5.8.1. of Brockwell & Davis, 1991):

$$\begin{aligned}\sigma^2_{1,r,H}(\Delta) &= 2\pi \exp\left\{\frac{1}{2\pi}\int_{-\pi}^{\pi} \log f_{r,\Delta}(\lambda)d\lambda\right\} \\ &= \Delta^{2r+2H}2^{r+1}\sigma^2\Gamma(2H+1)\sin(\pi H) \\ &\quad \times \exp\left\{\frac{r+1}{2\pi}\int_{-\pi}^{\pi}\log(1-\cos\lambda)d\lambda + \frac{1}{2\pi}\int_{-\pi}^{\pi}\log g_\Delta(\lambda)d\lambda\right\},\end{aligned}$$

where

$$g_\Delta(w) = \sum_{k=-\infty}^{\infty} |w+2k\pi|^{-2r-2H-1}|\beta(i(w+2k\pi)/\Delta)|^2/|\alpha(i(w+2k\pi)/\Delta)|^2.$$

Consequently,

$$\begin{aligned}\sigma^2_{1,r,H}(\Delta) &= \Delta^{2r+2H}2^{r+1}\sigma^2\Gamma(2H+1)\sin(\pi H)/\alpha_1^2 \\ &\quad \times \exp\left\{\frac{r+1}{2\pi}\int_{-\pi}^{\pi}\log(1-\cos\lambda)d\lambda + \frac{1}{2\pi}\int_{-\pi}^{\pi}\log\sum_{k=-\infty}^{\infty}|w+2k\pi|^{-2r-2H-1}dw\right\} \\ &\quad +o(\Delta^{2r+2H}).\end{aligned} \tag{A6}$$

On the other hand, we claim that for $0 < r + H < 1/2$,

$$\sigma^2_{2,H,r}(\Delta) = \frac{\sigma^2 D_H^2}{2H\alpha_1^2}\Delta^{2H} + o(\Delta^{2H}), \tag{A7}$$

for $r + H \geq 1/2$ and $0 < H < 1$,

$$\sigma^2_{2,H,r}(\Delta) = \frac{\sigma^2 D_H^2 \Delta^{2r+2H}}{\alpha_1^2(2H+2r)}\left[\prod_{k=1}^{r}(H+k-0.5)\right]^{-2} + O(\Delta^{2H+2r-1}), \tag{A8}$$

which are shown below. The theorem follows from equations (A6), (A7), (A8) and the following two formulas of Abramowitz & Stegun (1965); namely (i) the reflection formula: $\Gamma(z)\Gamma(1-z) = \pi\csc(\pi z)$, where $0 < z < 1$, and (ii) the duplication formula: $\Gamma(2z) = (2\pi)^{-1/2}2^{2z-1/2}\Gamma(z)\Gamma(z+1/2)$.

Now, we prove equation (A8) as follows. First consider the case for $1/2 < r + H < 1$. Without loss of generality, we consider positive n and hence for for large Δ it suffices to consider the case of large t. Equation (5) shows that for large t, X_0 has neglibible effect on X_t (indeed of order $O(\exp(A\Delta)$ for $t \geq \Delta$)) and hence for simplicity, we assume $X_0 = 0$. Then integration by parts (Norros *et al.*, 1999) yields

$$\begin{aligned} Y_t &= \sigma \int_0^t \beta' e^{A(t-u)} \delta_p dB_u^H \\ &= \sigma\beta'\delta_p B_t^H + \sigma\beta' A \int_0^t e^{A(t-u)}\delta_p B_u^H du. \end{aligned} \tag{A9}$$

By equation (1.1) of Norros *et al.* (1999),

$$B_t^H = D_H \int_0^t (t-u)^{H-1/2} dB_u + D_H \int_{-\infty}^0 \left\{(t-u)^{H-1/2} - (-u)^{H-1/2}\right\} dB_u,$$

where B_u is the standard Brownian motion and D_H is given in the theorem. Thus,

$$Y_t = \sigma D_H \int_0^t f(t,v) dB_v + \sigma D_H \int_{-\infty}^0 g(t,v) dB_v, \tag{A10}$$

where

$$f(t,v) = \beta'\delta_p(t-v)^{H-1/2} + \beta' A \int_v^t e^{A(t-u)}\delta_p(u-v)^{H-1/2} du$$

and

$$g(t,v) = \beta'\delta_p\{(t-v)^{H-1/2} - (-v)^{H-1/2}\} + \int_0^t \beta' A e^{A(t-u)}\delta_p \left\{(u-v)^{H-1/2} - (-v)^{H-1/2}\right\} du.$$

So,

$$\begin{aligned} Y_n^\Delta &= \int_{(n-1)\Delta}^{n\Delta} Y_s ds \\ &= \int_{(n-1)\Delta}^{n\Delta} \sigma D_H \int_0^s f(s,v) dB_v ds + \int_{(n-1)\Delta}^{n\Delta} \sigma D_H \int_{-\infty}^0 g(s,v) dB_v ds \\ &= \sigma D_H \int_0^{(n-1)\Delta} \int_{(n-1)\Delta}^{n\Delta} f(u,v) du dB_v + \sigma D_H \int_{(n-1)\Delta}^{n\Delta} \int_v^{n\Delta} f(u,v) du dB_v \\ &\quad + \sigma D_H \int_{-\infty}^0 \int_{(n-1)\Delta}^{n\Delta} g(u,v) du dB_v. \end{aligned}$$

It follows from equation (A10) that, for $u \geq (n-1)\Delta$,

$$E\left(Y_u | Y_v, -\infty < v \leq (n-1)\Delta\right) = \sigma D_H \int_0^{(n-1)\Delta} f(u,v) dB_v + \sigma D_H \int_{-\infty}^0 g(u,v) dB_v;$$

see Grenander and Rosenblatt (1984, p. 82). Let $\tilde{Y}_n^\Delta$ be the first-step predictor of Y_n^Δ. Now,

$$\begin{aligned} \tilde{Y}_n^\Delta &= \int_{(n-1)\Delta}^{n\Delta} E\left(Y_u | Y_v, v \leq (n-1)\Delta\right) du \\ &= \sigma D_H \int_0^{(n-1)\Delta} \int_{(n-1)\Delta}^{n\Delta} f(u,v) du dB_v + \sigma D_H \int_{-\infty}^0 \int_{(n-1)\Delta}^{n\Delta} g(u,v) du dB_v, \end{aligned}$$

demonstrating that the prediction error equals $Y_n^\Delta - \tilde{Y}_n^\Delta = \sigma D_H \int_{(n-1)\Delta}^{n\Delta} \int_v^{n\Delta} f(u,v) du dB_v$, hence the prediction variance equals $\sigma^2_{2,H,r}(\Delta) = \sigma^2 D_H^2 \int_{(n-1)\Delta}^{n\Delta} (\int_v^{n\Delta} f(u,v) du)^2 dv$, where

$$\begin{aligned} & \int_v^{n\Delta} f(t,v) dt \\ =\ & \int_v^{n\Delta} \beta' \delta_p (t-v)^{H-1/2} dt + \beta' A \int_v^{n\Delta} \int_v^t e^{A(t-u)} \delta_p (u-v)^{H-1/2} du dt \\ =\ & \int_v^{n\Delta} \beta' \delta_p (t-v)^{H-1/2} dt + \beta' A \int_v^{n\Delta} \int_u^{n\Delta} e^{A(t-u)} \delta_p (u-v)^{H-1/2} dt du \\ =\ & \beta' \int_v^{n\Delta} e^{A(n\Delta - u)} \delta_p (u-v)^{H-1/2} du \\ =\ & \beta' \int_0^{n\Delta - v} e^{A(n\Delta - v - w)} \delta_p w^{H-1/2} dw \qquad \text{(A11)} \\ =\ & -\beta' A^{-1} \delta_p (n\Delta - v)^{H-1/2} + O((n\Delta - v)^{H-3/2}) \\ =\ & -\frac{1}{\alpha_1} (n\Delta - v)^{H-1/2} + O((n\Delta - v)^{H-3/2}). \end{aligned}$$

Therefore,

$$\begin{aligned} \sigma^2_{2,H,r}(\Delta) &= \frac{\sigma^2 D_H^2}{\alpha_1^2} \int_{(n-1)\Delta}^{n\Delta} \left\{ (n\Delta - v)^{2H-1} + O\left((n\Delta - v)^{2H-2} \right) \right\} dv \\ &= \frac{\sigma^2 D_H^2}{2H\alpha_1^2} \Delta^{2H} + O(\Delta^{2H-1}). \end{aligned}$$

This proves equation (A8) for $1/2 < r + H < 1$. The proofs for the other cases of (A8) and equation (A7) are similar and hence omitted.

References

[1] Abramowitz, M. & Stegun , I. (1965). *Handbook of Mathematical Functions: with Formulas, Graphs, and Mathematical Tables.* New York: Dover Publications.

[2] Beran, J. (1994). *Statistics for Long-Memory Presses.* New York: Chapman and Hall.

[3] Beran, J. & Ocker. D. (2000). Temporal aggregation of stationary and non-stationary FARIMA(p,d,0) models. Working paper.

[4] Brockwell, P. J. (1993). Threshold ARMA processes in continuous time. In *Dimension Estimation and Models*, Ed. H. Tong, pp. 170-90. River Edge: World Scientific.

[5] Brockwell, P. J. & Davis, R. A. (1991). *Time Series: Theory and Methods.* New York: Springer-Verlag.

[6] Chambers, M. J., (1996). The estimation of continuous parameter long-memory time series models. *Econometric Theory* **12**, 374-90.

[7] Chambers, M. J., (1998). Long memory and aggregation in macroeconomic time series. *International economic review* **39**, no. 4, 1053-72.

[8] Comte, F. (1996). Simulation and estimation of long memory continuous time models. *J. Time Ser. Anal.* **17**, 19-36.

[9] Comte, F. & Renault, E. (1996). Long memory continuous time models. *J. Econometrics* **73**, 101-49.

[10] Duncan, Tyrone E., Hu, Yaozhong & Pasik-Duncan Bozenna (2000). Stochastic calculus for fractional Brownian motion. I. Theory. *SIAM J. Control Optim.* **38**, 582–612.

[11] Gradshteyn, I. S. & Ryzhik ,I.M. (1994). *Table of integrals, series, and products.* Boston: Academic Press.

[12] Grenander, U. and Rosenblatt, M. (1984). *Statistical analysis of stationary time series.* New York: Chelsea Publishing.

[13] Harvey, A. C. (1990). *Forecasting, structural time series models and the Kalman filter.* New York: Cambridge University Press.

[14] Jolley, L.B.W. (1961). *Summation of Series.* New York: Dover.

[15] Kleptsyna, M. L., Le Breton, A. & Roubaud, M.-C. (2000). Parameter estimation and optimal filtering for fractional type stochastic systems. *Statistical inference for stochastic processes* **3**, 173-82.

[16] Man, K.S. & Tiao, G.C. (2001). Limiting behavior of temporal aggregates of fractionally differenced processes. Working paper.

[17] Mandelbrot, B.B. & Van Ness, J. W. (1968). Fractional Brownian motions, fractional noises and applications. *SIAM Review* **10**, 422-37.

[18] Norros, I., Valkeila E. & Tirtamo, J. (1999). An elementary approach to Girsanov formula and other analytical results on fractional Brownian motions. *Bernoulli*, **5**, no. 4, 571-587.

[19] Priestley, M. B. (1981). *Spectral analysis and time series.* Academic Press.

[20] Stuart, A & Ord, K. (1994). *Kendall's advanced theory of statistics. Volume 1, Distribution theory.* London: Edward Arnold.

[21] Tiao, G. C. (1972). Asymptotic behavior of temporal aggregates of time series. *Biometrika.* **59**, No. 3, 525-531.

[22] Tsai, H. & Chan, K. S. (2003). Maximum likelihood estimation of a class of continuous-time long-memory processes. Technical Report, Department of Statistics & Actuarial Science, University of Iowa.

[23] Viano, M. C., Deniau, C. & Oppenheim, G. (1994). Continuous-time fractional ARMA processes. *Statistics & probability letters.* **21**, 323-336.

INSTITUTE OF STATISTICAL SCIENCE
ACADEMIA SINICA
TAIPEI, TAIWAN 115

E-MAIL: htsai@stat.stat.sinica.edu.tw

DEPARTMENT OF STATISTICS
AND ACTUARIAL SCIENCE
UNIVERSITY OF IOWA
IOWA CITY, IA 52242, USA

E-MAIL: kung-sik-chan@uiowa.edu

Proofs of equations (5), (6), (A5), (A7), (A8) and Theorem 2 (c), included for the perusal of the referees, and not part of the paper:

Proof of equation (5)

It suffices to show that X_t defined by the right side of equation (5) satisfies the integral equation implied by (4), which we now verify. Now, $\int_0^t AX_s ds + \int_0^t \sigma\delta_p dB_s^H = I + II$ where $I = \int_0^t A\{e^{As}X_0 + \sigma\int_0^s e^{A(s-u)}\delta_p dB_u^H\}ds$. It follows from routine calculations that $I = (e^{At} - I)X_0 + \{\sigma\int_0^t e^{A(t-u)}\delta_p dB_u^H - \sigma\delta_p B_t^H\}$; the last expression enclosed by the curly brackets follows from the equality that $\int_0^t \sigma\int_0^s Ae^{A(s-u)}\delta_p dB_u^H ds = \sigma\int_0^t\int_u^t Ae^{A(s-u)}ds\delta_p dB_u^H = \sigma\int_0^t\{e^{A(t-u)} - I\}\delta_p dB_u^H$, the interchange of the order of integration being justified by equation (2.2) in Norros et al. (1999) and the discussion there. Altogether, $I + II$ equals $(e^{At} - I)X_0 + \sigma\int_0^t e^{A(t-u)}\delta_p dB_u^H$. On the other hand if X_t is defined by the right side of (5), $X_t - X_0 = (e^{At} - I)X_0 + \sigma\int_0^t e^{A(t-u)}\delta_p dB_u^H$, showing that $\{X_t\}$ defined by (5) is a solution of (4). The unicity of the solution can be proved as in Tsai and Chan (working paper, 2003), and hence omitted.

Proof of equation (6)

The proof is similar to that of equation (5) and hence omitted.

Proof of equation (A5)

First note that it is trivial to see that equation (A5) holds for $r = 0$. Now suppose equation (A5) holds for $r = 1, \ldots, s$, note that $\mathrm{E}\left(|W_{s+1,h}|^H\right) = \mathrm{E}\left[|W_{s,h} + U + V|^H\right]$, where U and V are two independent uniform (0,1) random variables independent of $W_{s,h}$. Note also that the probability density function of $U + V$ is

$$f_{U+V}(x) = \begin{cases} x, & \text{for} \quad 0 < x < 1, \\ 2 - x, & \text{for} \quad 1 < x < 2. \end{cases}$$

Thus,

$$\begin{aligned}
& \mathrm{E}\left(|W_{s+1,h}|^H\right) \\
= & \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} |w + x|^H f_{W_{s,h}}(w) f_{U+V}(x) dw dx \\
= & \int_{-\infty}^{\infty}\int_0^1 x|w + x|^H f_{W_{s,h}}(w) dx dw + \int_{-\infty}^{\infty}\int_1^2 (2 - x)|w + x|^H f_{W_{s,h}}(w) dx dw \\
= & \int_{-\infty}^{\infty}\int_w^{w+1} |y|^H (y - w) f_{W_{s,h}}(w) dy dw + \int_{-\infty}^{\infty}\int_{w+1}^{w+2} |y|^H (2 - y + w) f_{W_{s,h}}(w) dy dw \\
= & \frac{1}{(H+1)(H+2)}\mathrm{E}\left[|W_{s,h}|^{H+2} - 2|W_{s,h+1}|^{H+2} + |W_{s,h+2}|^{H+2}\right], \qquad \text{(A12)}
\end{aligned}$$

where the last equality follows from routine calculus. By using equation (A5), equation (A12) becomes

$$
\begin{aligned}
&(H+1)(H+2)\prod_{k=1}^{2s+2}(H+2+k)\mathrm{E}\left(|W_{s+1,h}|^H\right) \\
= &\sum_{k=0}^{2s}(-1)^k\binom{2s}{k}|h+k|^{H+2s+2} - 2\sum_{k=0}^{2s}(-1)^k\binom{2s}{k}|h+k+1|^{H+2s+2} \\
&+\sum_{k=0}^{2s}(-1)^k\binom{2s}{k}|h+k+2|^{H+2s+2} \\
= &\sum_{k=0}^{2s}(-1)^k\binom{2s}{k}|h+k|^{H+2s+2} + 2\sum_{k=1}^{2s+1}(-1)^k\binom{2s}{k-1}|h+k|^{H+2s+2} \\
&+\sum_{k=2}^{2s+2}(-1)^k\binom{2s}{k-2}|h+k|^{H+2s+2} \\
= &\sum_{k=0}^{2s+2}(-1)^k\binom{2s+2}{k}|h+k|^{H+2s+2}.
\end{aligned}
$$

This proves equation (A5).

Proof of Theorem 2 (c) for $0 < H < 1/2$.

For any integer h,

$$
\begin{aligned}
&\gamma_\infty(h) \\
= &\operatorname{Cov}\left(\int_{n+h-r-1}^{n+h} f_{n+h,r}(u)dB_u^H, \int_{n-r-1}^{n} f_{n,r}(v)dB_v^H\right) \\
= &Hf_{n+h,r}(n+h)\int_{n-r-1}^{n}|n+h-v|^{2H-1}\mathrm{sign}(n+h-v)f_{n,r}(v)dv \\
&-Hf_{n+h,r}(n+h-r-1)\int_{n-r-1}^{n}|n+h-r-1-v|^{2H-1}\mathrm{sign}(n+h-r-1-v)f_{n,r}(v)dv \\
&+H\int_{n-r-1}^{n}\int_{n+h-r-1}^{n+h} f'_{n+h,r}(u)f_{n,r}(v)|u-v|^{2H-1}\mathrm{sign}(v-u)dudv,
\end{aligned}
$$

where $\mathrm{sign}(w)$ equals -1, 0 or 1 depending on whether w is negative, zero, or positive. But $f_{n+h,r}(n+h) = f_{n+h,r}(n+h-r-1) = 0$ and

$$
f'_{n+h,r}(u) = \begin{cases} f_{n+h-1,r-1}(u), & n+h-r-1 \le u \le n+h-r \\ f_{n+h-1,r-1}(u) - f_{n+h,r-1}(u), & n+h-r \le u \le n+h-1 \\ -f_{n+h,r-1}(u), & n+h-1 \le u \le n+h. \end{cases}
$$

Therefore,

$$
\begin{aligned}
\gamma_\infty(h) \;=\; &H\int_{n+h-r-1}^{n+h-r}\int_{n-r-1}^{n} f_{n+h-1,r-1}(u)f_{n,r}(v)|u-v|^{2H-1}\mathrm{sign}(v-u)dudv \\
&-H\int_{n+h-r}^{n+h}\int_{n-r-1}^{n} f_{n+h-1,r-1}(u)f_{n,r}(v)|u-v|^{2H-1}\mathrm{sign}(v-u)dudv
\end{aligned}
$$

$$
\begin{aligned}
&= HE\left[|Y_{n+h-1,r-1}-Y_{n,r}|^{2H-1}\operatorname{sign}(Y_{n,r}-Y_{n+h-1,r-1})\right] \\
&\quad -HE\left[|Y_{n+h,r-1}-Y_{n,r}|^{2H-1}\operatorname{sign}(Y_{n,r}-Y_{n+h,r-1})\right] \\
&= -HE\left[\left|\sum_{i=1}^{2r+1}U_i+h-r-1\right|^{2H-1}\operatorname{sign}\left(\sum_{i=1}^{2r+1}U_i+h-r-1\right)\right] \\
&\quad +HE\left[\left|\sum_{i=1}^{2r+1}U_i+h-r\right|^{2H-1}\operatorname{sign}\left(\sum_{i=1}^{2r+1}U_i+h-r\right)\right] \\
&= H(2H-1)E\left[\left|h-r-1+\sum_{i=1}^{2r+2}U_i\right|^{2H-2}\right], \qquad \text{(A13)}
\end{aligned}
$$

where the last equality follows from routine calculus. Equations (A4) and (A13) are identical, which shows that (c) is true for $0 < H < 1/2$.

Proof of equation (A7)

For $0 < r + H < 1/2$, we first note that for $a > -1$,

$$
\lim_{y\longrightarrow\infty}\frac{\int_0^y e^x x^a dx}{e^y y^a}=\lim_{y\longrightarrow\infty}\frac{e^y y^a}{e^y y^a+ae^y y^{a-1}}=\lim_{y\longrightarrow\infty}\frac{1}{1+a/y}=1.
$$

Therefore,

$$
\int_0^y e^{x-y}x^a dx = y^a + o(y^a),
$$

and so equation (A11) becomes

$$
-\beta' A^{-1}\delta_p(n\Delta - v)^{H-1/2} + o((n\Delta - v)^{H-1/2}).
$$

Therefore,

$$
\begin{aligned}
\sigma^2_{2,H,r}(\Delta) &= \frac{\sigma^2 D_H^2}{\alpha_1^2}\int_{(n-1)\Delta}^{n\Delta}\left\{(n\Delta-v)^{2H-1}+o\left((n\Delta-v)^{2H-1}\right)\right\}dv \\
&= \frac{\sigma^2 D_H^2}{2H\alpha_1^2}\Delta^{2H}+o(\Delta^{2H}).
\end{aligned}
$$

This proves the case for $0 < r + H < 1/2$.

Proof of equation (A8) for the cases other than $1/2 < r + H < 1$.

For $r \geq 1$, $0 < H < 1$ and $H \neq 1/2$, we first note that by equation (A10),

$$
\begin{aligned}
Y_u^r &= \frac{1}{(r-1)!}\int_0^u (u-v)^{r-1}Y_v^{(r)}dv \\
&= \frac{\sigma D_H}{(r-1)!}\int_0^u\int_w^u (u-v)^{r-1}f(v,w)dvdB_w \\
&\quad +\frac{\sigma D_H}{(r-1)!}\int_{-\infty}^0\int_0^u (u-v)^{r-1}g(v,w)dvdB_w.
\end{aligned}
$$

Similar to the case $r=0$, we can check that

$$\begin{aligned} Y_n^\Delta - \tilde{Y}_n^\Delta &= \frac{\sigma D_H}{r!}\int_{(n-1)\Delta}^{n\Delta}\int_w^{n\Delta}(n\Delta - v)^r f(v,w)dvdB_w \\ &= \frac{\sigma D_H}{r!}\int_{(n-1)\Delta}^{n\Delta} F(n\Delta, w)dB_w, \end{aligned}$$

where

$$\begin{aligned} & F(n\Delta, v) \\ =\ & \int_v^{n\Delta}(n\Delta - t)^r f(t,v)dt \\ =\ & \beta'\delta_p\int_v^{n\Delta}(n\Delta - t)^r (t-v)^{H-1/2}dt \\ & +\beta' A\int_v^{n\Delta}\int_u^{n\Delta}(n\Delta - t)^r e^{A(t-u)}\delta_p(u-v)^{H-1/2}dtdu \\ =\ & \beta' A^{-r} r!\int_v^{n\Delta}(w-v)^{H-1/2}e^{A(n\Delta - w)}\delta_p dw \\ & -\beta'\sum_{j=1}^{r} A^{-j}\delta_p\frac{r!}{(r-j)!}\int_v^{n\Delta}(w-v)^{H-1/2}(n\Delta - v)^{r-j}dw \\ =\ & -\frac{r!}{\alpha_1}\left[\prod_{k=1}^{r}(H+k-1/2)\right]^{-1}(n\Delta - v)^{H+r-1/2} + O\left((n\Delta - v)^{H+r-3/2}\right). \end{aligned}$$

Thus,

$$\begin{aligned} \sigma^2_{2,H,r}(\Delta) &= \frac{\sigma^2 D_H^2}{(r!)^2}\int_{(n-1)\Delta}^{n\Delta} F^2(n\Delta, w)dw \\ &= \frac{\sigma^2 D_H^2}{\alpha_1^2}\left[\prod_{k=1}^{r}(H+k-1/2)\right]^{-2}\int_{(n-1)\Delta}^{n\Delta}\left\{(n\Delta - v)^{2H+2r-1} + O(\Delta^{2H+2r-2})\right\}dv \\ &= \frac{\sigma^2 D_H^2 \Delta^{2r+2H}}{\alpha_1^2(2H+2r)}\left[\prod_{k=1}^{r}(H+k-1/2)\right]^{-2} + O(\Delta^{2H+2r-1}). \end{aligned}$$

This proves the case for $r \geq 1$, $0 < H < 1$ and $H \neq 1/2$. The case for $r+H = 1/2$ is trivial, so we now prove the case for $r \geq 1$ and $H = 1/2$. It can be shown by similar technique that

$$Y_n^\Delta - \tilde{Y}_n^\Delta = \frac{1}{r!}\int_{(n-1)\Delta}^{n\Delta} H(n\Delta, w)dB_w,$$

where

$$\begin{aligned} H(n\Delta, w) &= \int_w^{n\Delta}(n\Delta - v)^r e^{\alpha_1(v-w)}dv \\ &= \beta' A^{-r-1}e^{A(n\Delta - w)}\delta_p r! - \sum_{j=0}^{r}\beta' A^{-j-1}\delta_p\frac{r!}{(r-j)!}(n\Delta - w)^{r-j}. \end{aligned}$$

Therefore, $\sigma^2_{2,H,r}(\Delta) = \mathrm{var}(Y_n^{r,\Delta} - \tilde{Y}_n^{r,\Delta}) = \frac{\Delta^{2r+1}}{(2r+1)(r!)^2\alpha_1^2} + O(\Delta^{2r})$.