ELSEVIER

Journal of Econometrics 73 (1996) 101–149

JOURNAL OF
Econometrics

Long memory continuous time models

F. Comte, E. Renault*

Institut d'Economie Industrielle, Université de Sciences Sociales, 31042 Toulouse Cedex, France

Abstract

This paper presents a new family of long memory models: the continuous time moving average fractional process. The continuous time framework allows to reconcile two competitive types of modelling: fractional integration of ARMA processes and fractional Brownian Motion. A comparison with usual discrete time ARFIMA models is lead. Some well-known empirical evidence on macroeconomic and financial time series, such as variability of forward rates, aggregation of responses across heterogeneous agents, are well-captured by this continuous time modelling. Moreover, the usual statistical tools for long memory series and for Stochastic Differential Equations can be jointly applied in this setting.

Key words: Long memory; Continuous time models
JEL classification: C22

1. Introduction

It is well-documented that some macroeconomic time series like real output growth or consumption prices can exhibit long range dependence; see, for instance, Granger and Joyeux (1980) for food prices and Haubrich and Lo (1991) and Sowell (1992) for output growth. After Mandelbrot (1971), several authors have considered the issue of long memory components in financial time series like asset returns or interest rates. Recently, Backus and Zin (1995) have outlined the ability of the fractional difference model to mimic some of the features of the

*Corresponding author.

F. Comte is affliated with CREST and the University of Paris I and E. Renault with GREMAQ, University of Toulouse I, and the Institut Universitaire de France. The authors thank J.M. Dufour, C. Gouriéroux, R.T. Baillie, and two anonymous referees for useful comments, and B. Biais for helpful information.

SSDI 0304-4076(95)01735-V

term structure of interest rates, namely the variability of long yields. Unfortunately, since they have not at their disposal continuous time long memory processes, they are obliged to introduce a discrete time bond pricing model, which is not mainstream with respect to the modern continuous time finance.

The first motivation of this paper is to show that the class of continuous time stochastic processes most commonly employed in finance, namely Stochastic Differential Equations (SDE), can be extended to encompass long memory models. We prove not only that this extension is possible, but also that it is the natural one in order to get variations (of prices or rates) which have an instantaneous variance of order less than two (but not necessarily integer), the usual short memory case (diffusion processes) corresponding to the order one. This property is fundamental in the modern continuous time finance theory (see, for instance, Merton, 1990, Ch. 1) and corresponds to some kind of 'instantaneous unpredictability' of asset prices in the sense of Sims (1984).

Moreover, usual macroeconomic interpretation of long memory models in discrete time like aggregation across agents with heterogeneous beliefs (see Granger, 1980) or persistent variability in forward rates (see Backus and Zin, 1995) can easily be adapted to this continuous time setting.

The second focus of the paper is more statistical. Since continuous time long memory processes are able to mimic some important macroeconomic and financial phenomena, it remains to verify that they allow to define some parametric statistical models whose parameters have both nice structural interpretation and tractable estimators. Until now, the only fully parametric model in discrete time of long range dependence is provided by the so-called fractionally integrated ARMA processes (Granger and Joyeux, 1983). The main drawback of this parametrization is that it is not robust with respect to temporal aggregation, and therefore cannot be explicitly linked to the fractional Brownian Motion first introduced in continuous time by Mandelbrot and Van Ness (1968).

In this paper, we first consider the class of continuous time fractional ARMA processes, and we prove that they reconcile the two competitive paradigms: fractionally integrated ARMA processes and ARMA processes with respect to a fractional noise. Moreover, contrary to the findings of Geweke and Porter-Hudak (1983) in discrete time, the ARMA parameters in continuous time are invariant with respect to the operator of derivation/integration.

We also address the issue of estimating our continuous time long memory models from discrete sampling. We do not try to provide an asymptotically efficient procedure of estimation, but we prove that some usual statistical tools for long memory processes can provide satisfactory estimation of continuous time fractional ARMA processes. In particular, Robinson's (1992) improvement of the usual Geweke and Porter-Hudak procedure (based on the log-periodogram) can be used. Finally, once the degree of integration has been consistently estimated, we can compute the differentiated series for which the usual tools of continuous time ARMA models (see, for instance, Bergstrom, 1990) can be employed.

Fractional integration and fractional derivation in continuous time are presented in Section 2. We first define a moving average representation of continuous time processes that can be extended to a moving average with respect to fractional Brownian Motion. Something like a generalized semimartingale decomposition is set forth in order to check the property of 'instantaneous variance of order less than two'. Fractional ARMA processes in continuous time are then characterized, and we check the long memory properties both in the time and the frequency domain.

Section 3 presents our extension of 'Stochastic Differential Equations' to the case of fractional integration. This leads us to stress the invariance of the ARMA parametrization in continuous time with respect to fractional integration/derivation and to sketch the estimation strategies. We also compare our continuous time filter with the usual operator $(1-L)^d$ for ARFIMA models. Then, we present in Section 4 some applications for asset pricing and macroeconomics, namely the term structure of interest rates and the aggregation across individuals of heterogeneous beliefs. Lastly, we give some concluding remarks about a generalization of our multivariate continuous time models to the case of different orders of integration of the components of the process.

2. Fractional integration and fractional derivation

2.1. Moving average representation of continuous time processes

Just as in a discrete time framework linear representations (i.e., moving average representations of possibly infinite order) can describe all stationary regular processes, a continuous time Wold theorem of representation (see Rozanov, 1968, pp. 116–119) ensures that any linearly regular stationary process Y can be written as

$$Y(t) = m + \int_{-\infty}^{t} A(t-s)\,\mathrm{d}\xi(s), \tag{1}$$

with $\mathrm{d}\xi$ a uncorrelated random measure and A a square matricial function,

$$\int_{0}^{+\infty} A(x)\,{}^{T}\!A(x)\,\mathrm{d}x < +\infty, \tag{2}$$

where ${}^{T}\!A(x)$ denotes the transpose of $A(x)$.

This paper is particularly concerned with Gaussian continuous time processes,

$$\mathrm{d}\xi(t) = \mathrm{d}W(t), \tag{3}$$

where W is a standard Wiener process. Moreover, we can assume without loss of generality that $m = 0$ and that the process $X(t)$ is only defined for $t \geqslant 0$, starting from $t = 0$,

$$X(t) = \int_0^t A(t-s)\,\mathrm{d}W(s). \tag{4}$$

Any stationary process Y defined by (1), (2), and (3) is asymptotically equivalent to the process $m + X$, in the sense that

$$\lim_{t\to\infty} \mathrm{E}[Y(t) - m - X(t)]^T[Y(t - m - X(t)] = 0.$$

In the representation (4), the symbol $\int_0^t$ is analogous to the symbol $\sum_0^t$ used for a discrete time process initialized at time $t = 0$. In a continuous time framework, it is however useful to associate to the integral representation (4) a differential representation to show how some innovation terms appear. In order to do this in a classical way (i.e., by using the canonical decomposition of the semimartingale X), we need to introduce the values for $x = 0$ of the matrix function $A(x)$ and of its derivative $A'(x)$. The expression (2) does not assume that the function A is defined for $x = 0$; the integral may indeed be a generalized Riemann integral for a function A which is only defined on $\mathbb{R}_*^+$.

If the matrix function A is of class C^1 (i.e., differentiable with continuous derivative) on $\mathbb{R}^+$, the definition (4) can also be written as

$$\begin{aligned}
X(t) &= A(0)W(t) + \int_0^t [A(t-s) - A(0)]\,\mathrm{d}W(s) \\
&= A(0)W(t) + \int_0^t \left[\int_0^{t-s} A'(u)\,\mathrm{d}u\right]\mathrm{d}W(s) \\
&= A(0)W(t) + \int_0^t \left[\int_s^t A'(v-s)\,\mathrm{d}v\right]\mathrm{d}W(s),
\end{aligned}$$

so that, by applying Fubini's theorem for stochastic integrals (see Protter, 1990),

$$X(t) = A(0)W(t) + \int_0^t \left[\int_0^v A'(v-s)\,\mathrm{d}W(s)\right]\mathrm{d}v. \tag{5}$$

This provides the differential representation of the X process defined by (4), if the *deterministic function A is of class C^1 on $\mathbb{R}^+$*,

$$\mathrm{d}X(t) = A(0)\,\mathrm{d}W(t) + \left[\int_0^t A'(t-s)\,\mathrm{d}W(s)\right]\mathrm{d}t. \tag{6}$$

The basic idea of fractional derivation is to consider more generally some A functions of class C^1 on $\mathbb{R}_*^+$, but which present a singularity near zero because either $A(0) = 0$ but $A'(0)$ does not exist, or $A(0)$ itself does not exist. Such singularities are introduced by considering $A(x)$ of the following form:

$$A(x) = \frac{x^d \tilde{A}(x)}{\Gamma(d+1)}, \qquad d < 1, \tag{7}$$

where $\tilde{A}$ is of class C^1 on $\mathbb{R}^+$. Let us notice that the constant $\Gamma(d+1)$ is only introduced in order to provide a more symmetric expression of the dual operations of fractional derivation and integration (see Section 2.2).

Two cases have to be distinguished:

- If $d < 0$, $A(0)$ does not exist. In order to obtain well-defined generalized integrals $\int_0^t A(x)\,{}^T\!A(x)\,\mathrm{d}x$ for a constant function $\tilde{A}$, we shall nevertheless always assume that $d > -\frac{1}{2}$.
- If $0 < d < 1$, $A(0)$ does exist [$A(0) = 0$], but $A'(0)$ does not. We shall see that the case $\frac{1}{2} < d < 1$ can be reduced to the case $-\frac{1}{2} < d < 0$ by one derivation. So we shall consider as specific only the case $0 < d < \frac{1}{2}$.

These two cases exactly correspond to the usual definition of continuous time fractional Brownian motion, as described in the seminal paper by Mandelbrot and Van Ness (1968), in the case of a constant function $\tilde{A}$:

Definition 1. A continuous time fractional Brownian motion of order d, $-\frac{1}{2} < d < \frac{1}{2}$, and of dimension n is a process W_d defined by

$$W_d(t) = \int_0^t \frac{(t-s)^d}{\Gamma(d+1)}\,\mathrm{d}W(s),^1$$

where W is an n-dimensional Brownian motion.

For $d = 0$, W_d is the usual Brownian motion W, while for $d \neq 0$, W_d is a process whose increments are not stationary independent but present the well-known property of H self-similarity, where $H = d + \frac{1}{2}$; this means:

The probability distributions of $W_d(ut)/u^H$ and $W_d(t)$ are identical.

H is often called the Hurst coefficient.

---

[1] The process $W_d(t)$ is generally not asymptotically covariance stationary [condition (2) is not fulfilled for $0 < d < \frac{1}{2}$] and the integral $\int_{-\infty}^t \{(t-s)^d/\Gamma(d+1)\}\,\mathrm{d}W(s)$ cannot generally be defined. This is the reason why the fractional Brownian motion B_H ($H = d + \frac{1}{2}$) is generally defined by its increments:

$$B_H(t) - B_H(0) = \frac{1}{\Gamma(H+\frac{1}{2})} \int_{-\infty}^{t} (t-s)^{H-1/2}\,\mathrm{d}W(s) - \frac{1}{\Gamma(H+\frac{1}{2})} \int_{-\infty}^{0} (-s)^{H-1/2}\,\mathrm{d}W(s),$$

where the substraction of the two integrals takes place before the lower limit is allowed to go to $-\infty$. We have slightly changed this usual definition by giving in Definition 1 an initial value of $W_d(0) = 0$.

More generally, we introduce the following definition:

Definition 2. A continuous time fractionally integrated process of order d, $-\frac{1}{2} < d < \frac{1}{2}$, *and of dimension n is a process X defined by*

$$X(t) = \int_0^t \frac{(t-s)^d}{\Gamma(d+1)} \tilde{A}(t-s)\, \mathrm{d}W(s), \qquad t \in [0, T],$$

with W n-dimensional Brownian motion and $\tilde{A}$ *of class* C^1 *on* $[0, T]$.

The representation of X is moreover said to be canonical if the matrix $\tilde{A}(0)$ *is lower triangular.*[2]

In particular, a sufficient condition for such a process to be asymptotically stationary (for any defined value, d, $-\frac{1}{2} < d < \frac{1}{2}$) is that $\tilde{A}$ should be of class C^1 on $\mathbb{R}^+$, with

$$\lim_{x \to +\infty} x\tilde{A}(x) = A_\infty, \tag{8}$$

where A_∞ is a $n \times n$ constant square matrix, since the condition (2) is then trivially fulfilled.

The main advantage of the particular form of the singularity introduced by Definition 2 is that it allows to generalize the canonical decomposition (5) and (6) by hiding the singular part t^d inside the Brownian term by replacing W by W_d. For that purpose, we must mathematically define stochastic integration of a function w.r.t. W_d, and show that a process written as $\int_0^t D(t-s)\,\mathrm{d}W_d(s)$ does admit a decomposition of type (5). This is done in the following lemma:

Lemma 1. *Let* $X(t) = \int_0^t A(t-s)\,\mathrm{d}W(s)$. *Then* (i) $Y(t) = \int_0^t D(t-s)\,\mathrm{d}X(s)$ *is defined as*

$$Y(t) = \int_0^t D(t-s)\,\mathrm{d}X(s) := \frac{\mathrm{d}}{\mathrm{d}t}\left[\int_0^t D(t-s)X(s)\,\mathrm{d}s\right],$$

provided that

(C1) $\int\int_{[0,t]^2} (D(t-s)A(s-u)) \times {}^T(D(t-s)A(s-u))\chi_{[0,s]}(u)\,\mathrm{d}u\,\mathrm{d}s < \infty, \quad \forall t.$[3]

(C2) $D * A(x) = \int_0^x D(x-u)A(u)\,\mathrm{d}u$

admits a.e. on $[0, T]$ *a square integrable derivative.*

---

[2] We already noticed in Comte and Renault (1992) that choosing a canonical representation is always possible if $\tilde{A}(0)$ is invertible. This is due to a Gramm–Schmidt orthonormalization of $\tilde{A}(0)$ ${}'\tilde{A}(0) = T'T$, with T lower triangular, and similar transformation of the Brownian Motion to $\tilde{W} = T^{-1}\tilde{A}(0)W$.

[3] $\chi_{[a,b]}$ is the usual characteristic function: $\chi_{[a,b]}(u) = 1$ if $a \leqslant u \leqslant b$ and 0 otherwise.

Under (C1) and (C2), we also have

$$Y(t) = \int_0^t (D * A)'(t-s)\,\mathrm{d}W(s).$$

(ii) *If moreover* $\int_0^t D'(t-s)\,\mathrm{d}X(s)$ *is well-defined in the previous sense, then* Y *has the decomposition*

$$Y(t) = D(0)X(t) + \int_0^t \left(\int_0^s D'(s-u)\,\mathrm{d}X(u) \right) \mathrm{d}s.^4$$

Thus, if

$$Y(t) = \int_0^t D(t-s)\,\mathrm{d}W_d(s),$$

with D of class C^1 on $[0, T]$, since the conditions (C1) and (C2) are clearly fulfilled when $X = W_d$, we have

$$Y(t) = \int_0^t (D * A)'(t-s)\,\mathrm{d}W(s),$$

with

$$A(t-s) = \frac{(t-s)^d}{\Gamma(d+1)},$$

$$D * A(x) = \int_0^x D(x-u)\frac{u^d}{\Gamma(d+1)}\,\mathrm{d}u,$$

$$(D * A)'(x) = D(0)\frac{x^d}{\Gamma(d+1)} + \int_0^x D'(x-u)\frac{u^d}{\Gamma(d+1)}\,\mathrm{d}u,$$

so that, if we want to write

$$Y(t) = \int_0^t \frac{(t-s)^d}{\Gamma(d+1)}\tilde{A}(t-s)\,\mathrm{d}W(s),$$

an identification gives

$$x^d \tilde{A}(x) = D(0)x^d + \int_0^x D'(x-u)u^d\,\mathrm{d}u,$$

[4] We use equations as $d(D(t-s)X(s)) = D(t-s)\mathrm{d}X(s) - D'(t-s)X(s)\mathrm{d}s$, which gives $(\mathrm{d}/\mathrm{d}t)(\int_0^t D(t-s)X(s)\mathrm{d}s) = \int_0^t D'(t-s)X(s)\mathrm{d}s + D(0)X(t)$.

i.e.,

$$\tilde{A}(x) = D(0) + \int_0^x D'(u)\left(1 - \frac{u}{x}\right)^d \mathrm{d}u.$$

This proves that any process that can be written $Y(t) = \int_0^t D(t-s)\,\mathrm{d}W_d(s)$, with D of class C^1 on $[0, T]$ is a fractionally integrated process of order d. The reciprocal is proved in the Appendix, so that we have the following result:

Proposition 1. If X is a fractionally integrated process of order d, $-\frac{1}{2} < d < \frac{1}{2}$, defined by

$$X(t) = \int_0^t \frac{(t-s)^d}{\Gamma(d+1)} \tilde{A}(t-s)\,\mathrm{d}W(s), \qquad t \in [0, T], \tag{9}$$

with $\tilde{A}$ C^1 on $[0, T]$, then X can be written as

$$X(t) = \int_0^t C(t-s)\,\mathrm{d}W_d(s), \qquad t \in [0, T], \tag{10}$$

with C continuous on $[0, T]$, where

$$C(x) = \frac{1}{\Gamma(1-d)\Gamma(1+d)} \frac{\mathrm{d}}{\mathrm{d}x}\left(\int_0^x (x-s)^{-d} s^d \tilde{A}(s)\,\mathrm{d}s\right). \tag{11}$$

The reciprocal is true if C is supposed C^1, and then the resulting $\tilde{A}$ function is continuous and[5]

$$\tilde{A}(x) = C(0) + \int_0^x C'(u)\left(1 - \frac{u}{x}\right)^d \mathrm{d}u. \tag{12}$$

We can notice that, for d nonzero, the $\tilde{A}$ and C functions are distinct (but in one-to-one relation); this proves that we cannot straightforwardly replace $\{(t-s)^d/\Gamma(d+1)\}\,\mathrm{d}W(s)$ by $\mathrm{d}W_d(s)$. Moreover, formula (10) and (ii) of Lemma 1 give the decomposition for a fractionally integrated process of order d:

$$X(t) = C(0)\,W_d(t) + \int_0^t \left[\int_0^v C'(v-s)\,\mathrm{d}W_d(s)\right]\mathrm{d}v. \tag{13}$$

The decomposition (13) is a real generalization of (5), if we take into account that W has been replaced by W_d in order to hide the singularity $\{(t-s)^d/\Gamma(d+1)\}$ and that the C function appears, instead of $\tilde{A}$. It is then natural to give also a generalization of (6) through writing this relation in a differential way:

$$\mathrm{d}X(t) = C(0)\,\mathrm{d}W_d(t) + \left[\int_0^t C'(t-s)\,\mathrm{d}W_d(s)\right]\mathrm{d}t. \tag{14}$$

[5] As $\tilde{A}(0) = C(0)$, if the representation is canonical, $C(0)$ is also lower triangular.

But it must be noticed that if d is nonzero, the first term $C(0)\,\mathrm{d}W_d(t)$ of decomposition (14) is not a true 'innovation' since it is generally correlated with the second term (the process W_d is not with independent increments for $d \neq 0$). If we define anyway the information set $\mathscr{I}(t)$, relevant at time t, as the natural filtration (see Protter, 1990, p. 3) associated to W_d:

$$\mathscr{I}(t) = \sigma(W_d(\tau), 0 \leqslant \tau \leqslant t),$$

we can see that (14), for an infinitely small h, gives a decomposition of the variation of X, $X(t+h) - X(t)$, into two parts:

- a part $\int_t^{t+h} [\int_0^v C'(v-s)\,\mathrm{d}W_d(s)]\,\mathrm{d}v$, which conditionally to $\mathscr{I}(t)$ has an infinitely small of order h^2 [denoted $O(h^2)$ in the following] variance,
- a part $C(0)[W_d(t+h) - W_d(t)]$, which conditionally to $\mathscr{I}(t)$ has a variance of order $O(h^{2d+1})$. Indeed, if we denote with an index t the conditionment with respect to $\mathscr{I}(t)$, we have

$$V_t(W_d(t+h) - W_d(t)) = V_t(W_d(t+h)) = V_t\left[\int_0^{t+h} \frac{(t+h-s)^d}{\Gamma(d+1)}\,\mathrm{d}W(s)\right]$$

$$= \left[\int_t^{t+h} \frac{(t+h-s)^{2d}}{\Gamma(d+1)^2}\,\mathrm{d}s\right] I_n = \frac{h^{2d+1}}{(2d+1)\Gamma(d+1)^2} I_n,$$

where I_n is the $n \times n$ identity matrix.

This decomposition shows that the class of fractionally integrated processes is a natural extension of the class of diffusion processes and may be particularly useful to give representations of financial variables (stock prices, yields, interest rates, change rates, ...). We find here again the fundamental idea of the modern continuous time finance theory (see Merton, 1990, Ch. 1) that the variations (of prices or rates) have *an instantaneous variance of order less than 2:* $-\frac{1}{2} < d < \frac{1}{2} \Leftrightarrow 0 < 2d+1 < 2$.

The case $d = 0$ corresponds to the usual case of diffusion processes, whose instantaneous variance is of order 1, which is the idea of independent increments. The case of fractionally integrated processes is just the natural generalization since it makes possible the whole 'allowed' interval (0, 2) for the order of the instantaneous variance (and $d = 0$ gives the only integer order for $2d+1$ in this interval!). The idea of independent increments is then replaced by the more general idea of self-similarity [of order $H = d + \frac{1}{2} \in (0, 1)$].

2.2. FRACIMA processes

Given the analogy between discrete and continuous time frameworks, it is natural to look for an autoregressive representation of the fractionally integrated process X defined by a moving average representation (9) and/or (10). The solution of this problem is well-known in the usual case $d = 0$. More precisely, we have detailed in Comte and Renault (1992) the proof of the following proposition:

Proposition 2. Let X admit a representation $X(t) = \int_0^t A(t-s)\,\mathrm{d}W(s)$ *for* $t \in [0, T]$ *with* A C^1 *on* $[0, T]$, *and* $A(0)$ *regular matrix. Then there is an unique matrix function* B, C^1 *on* $[0, T]$, *so that*

$$W(t) = \int_0^t B(t-s)\,\mathrm{d}X(s). \tag{15}$$

B is the unique matrix function solution of the convolution equation:

$$B * A(x) = \int_0^x B(x-u)\,A(u)\,\mathrm{d}u = xI_n, \qquad \forall x \in [0, T].$$

In particular, $B(0) = A(0)^{-1}$.

A process which fulfills the regularity conditions of Proposition 2 was called in Comte and Renault (1992) a Continuous time Invertible Moving Average (CIMA) process.

Unfortunately, the fractionally integrated processes of nonzero order d are precisely processes which don't verify the regularity conditions: 'A C^1 on $[0, T]$ and $A(0)$ regular matrix' [see formula (7)]. This is the reason why the natural generalization of Proposition 2 is obtained by considering 'MA and AR representations' with respect to the fractional Brownian Motion W_d instead of the standard Wiener process W.

Proposition 3. Let X be a fractionally integrated process of order d, $-\frac{1}{2} < d < \frac{1}{2}$, *defined as in (10) by*

$$\begin{cases} X(t) = \int_0^t C(t-s)\,\mathrm{d}W_d(s), & t \in [0, T], \\ C \ C^1 \ \text{on} \ [0, T], & C(0) \ \text{regular matrix.} \end{cases}$$

Then, there is an unique matrix function D, C^1 *on* $[0, T]$, *so that*

$$W_d(t) = \int_0^t D(t-s)\,\mathrm{d}X(s).$$

D is the unique matrix function solution of the convolution equation: $D * C(x) = xI_n$, $\forall x \in [0, T]$. *In particular,* $D(0) = C(0)^{-1}$.

A fractional process which fulfills the regularity conditions of Proposition 3 will be called a FRACIMA process in order to extend the class of CIMA processes to fractional ones.

It is worth noting that the regularity conditions about the behaviour of C in the neighbourhood of zero are directly related to the properties of the matrix function $\tilde{A}$ which defines the representation (9) of the fractional process. In particular, it is straightforward from (12) that $C(0) = \tilde{A}(0)$. Moreover, while the matrix function

$$A(x) = \frac{x^d}{\Gamma(d+1)} \tilde{A}(x)$$

of the MA representation of X with respect to W exhibits some singularities in 0, the regularity of the matrix function C allows us to consider it as the MA coefficients function of a CIMA process $X^{(d)}$, derivative of order d of the process X:

Proposition 4. If X is a fractionally integrated process of order d, $-\frac{1}{2} < d < \frac{1}{2}$,

$$X(t) = \int_0^t \frac{(t-s)^d}{\Gamma(d+1)} \tilde{A}(t-s)\,\mathrm{d}W(s), \qquad t \in [0, T].$$

Then:

$$X^d(t) = \frac{\mathrm{d}}{\mathrm{d}t}\left[\int_0^t \frac{(t-s)^{-d}}{\Gamma(1-d)} X(s)\,\mathrm{d}s\right] = \int_0^t \frac{(t-s)^{-d}}{\Gamma(1-d)}\,\mathrm{d}X(s)$$

is well-defined and m.s. (mean square) continuous. If, moreover, $\tilde{A}(0)$ is invertible and $\tilde{A}$ C^2 on $[0, T]$, then $X^{(d)}$ admits the CIMA representation:

$$X^{(d)}(t) = \int_0^t C(t-s)\,\mathrm{d}W(s),$$

where $\tilde{A}$ and C are one-to-one related by (12) and (11).

Let us emphasize that the operator which provides $X^{(d)}$ from X is strictly speaking a derivation operator only if d is positive. In order to see this, it is useful to keep in mind the following properties of this operator:

(i) For every d, $-\frac{1}{2} < d < \frac{1}{2}$, $[X^{(d)}]^{(-d)} = X$, by a straightforward application of the Fubini theorem and the following well-known identity (see Abramowitz and Stegun, 1972):

$$\int_0^t (1-x)^d x^{-d}\,\mathrm{d}x = \Gamma(1+d)\,\Gamma(1-d).$$

(ii) For every nonnegative integer n,

$$\frac{\mathrm{d}^n}{\mathrm{d}t^n}\left[\int_0^t \frac{(t-s)^n}{\Gamma(n+1)}\,\mathrm{d}X(s)\right] = X(t),$$

and this elementary property can easily be extended to the general case $n = -d, -\frac{1}{2} < d < 0$.

In other words, negative values of d define a primitive operator rather than a derivative one. Symmetrically, for d positive, X appears like an integration of a CIMA process $X^{(d)}$. This justifies the terminology 'fractionally integrated' which naturally generalizes the usual integrated processes. For $-\frac{1}{2} < d < 0$, X can be seen as the differentiation (of order 1) of a fractionally integrated process of order $\beta = d + 1$.

2.3. Long memory properties

The so-called long memory property is usually defined from the autocovariances of the process (or equivalently from the spectral density) by saying that, when h grows to infinity, $\operatorname{cov}(X(t), X(t+h))$ tends towards zero at hyperbolic rate. Such a rate of decrease is much slower than the usual exponential rate associated with the standard ARMA process. This long range dependence can be found in our framework when we consider truly fractionally integrated processes, that is to say in the case $0 < d < \frac{1}{2}$. Of course, we are interested here in asymptotically stationary processes in order to give a sense to $\lim_{h \to +\infty} \operatorname{cov}[X(t), X(t+h)]$ independently of t. For $0 < d < \frac{1}{2}$, the asymptotic stationarity of the fractionally integrated process of order d,

$$X(t) = \int_0^t \frac{(t-s)^d}{\Gamma(d+1)} \tilde{A}(t-s)\,\mathrm{d}W(s),$$

is ensured if we assume (8): $\lim_{x \to +\infty} x\tilde{A}(x) = A_\infty$, where A_∞ is a given nonzero matrix. Indeed, the assumption (8) implies that for every d, $0 < d < \frac{1}{2}$, the process

$$Y(t) = \int_{-\infty}^t \frac{(t-s)^d}{\Gamma(d+1)} \tilde{A}(t-s)\,\mathrm{d}W(s)$$

is a well-defined stationary process [condition (2) is fulfilled]. Then, we obtain the following long-memory property:

Proposition 5. Let $X(t) = \int_0^t \{(t-s)^d/\Gamma(d+1)\} \tilde{A}(t-s)\,\mathrm{d}W(s)$ *be a fractionally integrated process of order d with*

$$0 < d < \tfrac{1}{2} \quad \text{and} \quad \lim_{x \to +\infty} x\tilde{A}(x) = A_\infty \neq 0. \tag{16}$$

Then $Y(t) = \int_{-\infty}^t \{(t-s)^d/\Gamma(d+1)\} \tilde{A}(t-s)\,\mathrm{d}W(s)$ *is a second order stationary process, asymptotically equivalent to X, which verifies*

$$\lim_{h \to +\infty} h^{1-2d} \gamma_Y(h) = \frac{\Gamma(1-2d)\Gamma(d)}{\Gamma(1-d)\Gamma(d+1)^2} A_\infty{}^T A_\infty,$$

where $\gamma_Y(h) = \operatorname{cov}[Y(t), Y(t+h)]$ *is the autocovariance function of X.*

Hence the covariance between $X(t)$ and $X(t+h)$ decreases towards zero when h grows to infinity at the same rate as h^{2d-1}. This is the reason why we shall call in the following *continuous time long memory processes of order d*, processes X verifying (16).

By analogy with discrete time usual long memory properties, this result confirms the interpretation of d, $0 < d < \frac{1}{2}$, as a fractional degree of integration. Moreover, this analogy can be extended to the frequency domain.

In the frequency domain indeed, we can define the spectral density of the continuous time stationary process Y equivalent to X (following the notations of Proposition 5) as the Fourier transform of the autocovariance function:

$$f_X(\lambda) = \int_{-\infty}^{+\infty} e^{-i\lambda h} \gamma_Y(h)\, dh, \tag{17}$$

for λ nonzero real.

In all the following, by convention we shall call '*spectral density of X*' the function f_X defined by (17). Of course, we need to check that the generalized Riemann integral (17) is well-defined for any long memory continuous time process. For that purpose, we prove in the Appendix:

Proposition 6. Let $X(t) = \int_0^t A(t-s)\, dW(s)$ be a continuous time long memory process and $\gamma_Y(h) = \operatorname{cov}[Y(t), Y(t+h)]$ the autocovariance function of the stationary process $Y(t) = \int_{-\infty}^t A(t-s)\, dW(s)$. Then, for every nonzero real λ, the spectral density of X (and Y) is given by

$$f_X(\lambda) = \int_{-\infty}^{+\infty} e^{-i\lambda h} \gamma_Y(h)\, dh = \hat{A}(\lambda)\hat{A}(\lambda)^*,$$

where

$$\hat{A}(\lambda) = \int_{-\infty}^{+\infty} e^{-i\lambda x} A(x)\, dx \tag{18}$$

is the Fourier transform of the matrix function A and $\hat{A}(\lambda)^*$ is the conjugate of the transpose of $\hat{A}(\lambda)$. The integral giving $\hat{A}$ is by (16) a convergent generalized Riemann integral.

The long memory property can then as usual be characterized by the behaviour of the spectral density in the neighbourhood of zero:

Proposition 7. Let $X(t) = \int_0^t A(t-s)\, dW(s)$ be a continuous time long memory process of order d $(0 < d < \frac{1}{2})$ with $A(x) = x^d \tilde{A}(x)/\Gamma(d+1)$ and $A_\infty = \lim_{x \to +\infty} x\tilde{A}(x)$. Then the spectral density f_X of X satisfies

$$\lim_{\lambda \to 0} \lambda^{2d} f_X(\lambda) = c A_\infty{}^T A_\infty, \tag{19}$$

where c is a positive scalar.

The property (19) looks like the usual notion of long memory as it was defined in discrete time by Granger and Joyeux (1980). Moreover, Robinson (1992) emphasizes that a semiparametric characterization of long memory property as (19) is sufficient to build some consistent estimators of the fractional degree of integration d. It is nevertheless important to keep in mind that these usual characterizations are defined from the singularity of the spectral density of a *discrete time process*. But it is easy to give the explicit link between the spectral density of the continuous time process $X(t)$, $t \in \mathbb{R}^+$, and the one of an associated discrete time process $X(k\Delta t)$, $k \in \mathbb{N}$, for a fixed $\Delta t > 0$. If we denote f_X^D the latter, we know by the folding formula (Bergström, 1990, p. 83) and the Poisson formula (Schwartz, 1966, Ch. VII, 7, 5, p. 254) that

$$f_X^D(\lambda) = \frac{1}{\Delta t} \sum_{n \in Z} f_X\left(\frac{\lambda + 2\pi n}{\Delta t}\right). \tag{20}$$

But, for a continuous time long memory process, it is clear that when λ tends towards zero, the dominating term in the sum (20) is the one associated with $n = 0$. Thus, as $\lambda \to 0^+$ and $\Delta t = 1$,

$$f_X^D(\lambda) \sim f_X(\lambda) \sim cA_\infty {}^T A_\infty \lambda^{-2d}, \tag{21}$$

which is exactly the usual definition of long memory in discrete time.

3. Stochastic differential equations and fractional integration

3.1. The invariance property

Stochastic Differential Equations are the continuous time analogue of finite order ARMA processes. Let us first recall how fractional integration was characterized in discrete time for infinite order ARMA models in a paper by Geweke and Porter-Hudak (1983) (GPH) and why the continuous time framework is much more suitable for such a characterization.

On the one hand, fractional differencing introduced by Hosking (1981) and Granger and Joyeux (1980) shows that by a convenient definition of the operator $(1 - L)^d$, a discrete time process $X(t)$ has long memory properties, for $0 < d < \frac{1}{2}$, if $(1 - L)^d X(t)$ admits a stationary invertible ARMA representation:

$$\Phi(L)(1 - L)^d X(t) = \Theta(L)\varepsilon(t), \tag{22}$$

where $\varepsilon(t)$ is a Gaussian discrete time white noise and $LX(t) = X(t - 1)$. The corresponding AR(∞) and MA(∞) representations will be denoted by $\Pi(L)(1 - L)^d X(t) = \varepsilon(t)$ and

$$(1 - L)^d X(t) = H(L)\varepsilon(t). \tag{23}$$

On the other hand, we can define, as GPH (1983), a 'simple fractional Gaussian noise' by

$$\varepsilon_d(t) = W_d(t) - W_d(t-1), \tag{24}$$

so that a Gaussian discrete time white noise appears to be associated to $d = 0$. GPH define then a 'general fractional Gaussian noise' by a generalized ARMA representation:

$$\Phi(L)X(t) = \Theta(L)\varepsilon_d(t). \tag{25}$$

The corresponding AR(∞) and MA(∞) representations are: $\Pi(L)X(t) = \varepsilon_d(t)$ and

$$X(t) = H(L)\varepsilon_d(t). \tag{26}$$

The main result of GPH (see Theorem 1, p. 223) establishes that the set of 'general integrated series' [in the sense of (22)] coincides with the set of 'general fractional Gaussian noises'. Unfortunately, due to the unnatural discretization (24) of the fractional Brownian Motion, it is never proved that the link between the two competitive representations (22) and (25) could be simply derived by a natural inversion of the differencing operator $(1-L)^d$. In other words, the parametric model of the autoregressive and moving average polynomials is not invariant with respect to the transformation from (22) to (25).

On the opposite, the main interest of Propositions 1, 2, 3, and 4 is to provide continuous time MA and AR representations which are invariant with respect to the operator derivation/integration of order d for continuous time processes. More precisely, if we consider as in (10):

$$X(t) = \int_0^t C(t-s)\,\mathrm{d}W_d(s), \tag{27}$$

i.e., a MA with respect to W_d which is the continuous time analogue of (26), we know that the d-derivative $X^{(d)}$ of X [the continuous time analogue of $(1-L)^d X(t)$] can be written, with Proposition 4,

$$X^{(d)}(t) = \int_0^t C(t-s)\,\mathrm{d}W(s) \quad \text{with the } \textit{same } C \text{ function.} \tag{28}$$

i.e., a MA representation of the differentiated process which is the continuous time analogue of (23). But, now, the moving average coefficients are preservated by the derivation/integration operator which relates (27) to (28). Of course, because of Propositions 2 and 3, the same invariance property is fulfilled by the AR representations.

This *invariance property* is the main advantage of the continuous time framework. Moreover, it is maintained when we consider the associated parametric models.

3.2. Parametric statistical models with fractional integration

To address this issue, let us first recall that, just as in discrete time framework, we get parametric statistical models by considering ARMA (p, q) models as close as we want to general MA(∞) or AR(∞) ones. Indeed, SDE models provide a fully parametrized class of CIMA processes which is quite general since the AR representation (15) can always be approximated with a polynomial function B:

Proposition 8. Let X be the solution of

$$\mathrm{d}D^{p-1}X(t) = [M(t) + B_0X(t) + B_1DX(t) + \cdots + B_{p-1}D^{p-1}X(t)]\,\mathrm{d}t + \Sigma\mathrm{d}W(t), \tag{29}$$

where $D^iX(t)$ is the m.s. derivative of order i of X, with initial conditions $X(0) = DX(0) = \cdots = D^{p-1}X(0) = 0$ and M a continuous deterministic vector function. Then $X(t) = \int_0^t [\exp((t-s)\bar{B})]_{1,p}(\Sigma\mathrm{d}W(s) + M(s)\mathrm{d}s)$ and if Σ is invertible, $D^{p-1}X(t)$ admits the CIMA representation:

$$D^{p-1}X(t) = \int_0^t A^{(p-1)}(t-s)(\Sigma\mathrm{d}W(s) + M(s)\mathrm{d}s),$$

$$A^{(p-1)}(t-s) = [\exp((t-s)\bar{B})]_{p,p}, \tag{30}$$

$$W(t) = K(t) + \int_0^t B(t-s)\mathrm{d}D^{p-1}X(s),$$

$$B(x) = \Sigma^{-1}\left(Id - \sum_{i=1}^{p} \frac{B_{p-i}}{i!}x^i\right), \tag{31}$$

where $\bar{B}$ is a $np \times np$ matrix defined by blocks from its p^2 $n \times n$ submatrices $\bar{B} = (\bar{B}_{i,j})_{1 \leq i,j \leq p}$ with $\bar{B}_{i,j} = B_{i-1}$ if $j = p$, $\bar{B}_{i,i+1} = I_n$ if $i = 1, \ldots, p$, and $\bar{B}_{i,j} = 0$ else, and $K(t) = -\int_0^t B(t-s)\, d(\int_0^s [e^{(s-u)\bar{B}}]_{p,p} M(u)\mathrm{d}u)$.

Moreover, this particular set of CIMA processes exactly provides in discrete time a process $X(k\Delta t)$, $k \in Z$, which is a VARMA $(p, p-1)$. This general result proved by Bergström (1984) is easy to understand in the particular case $p = 1$. The solution of the first order [Eq. (29) for $p = 1$] can be written $X(t) = \int_0^t \exp[(t-s)B_0]\,\Sigma\mathrm{d}W(s)$ for $M \equiv 0$, which provides in discrete time a VAR(1) process: $X(t + \Delta t) = \exp[B_0\Delta t]\,X(t) + \varepsilon(t + \Delta t)$, where $\varepsilon(t + \Delta t) = \int_t^{t+\Delta t} \exp[(t + \Delta t - s)B_0]\,\Sigma\,\mathrm{d}W(s)$ is a Gaussian white noise.

More generally, the solution of the higher-order Eq. (29) provides in discrete time a VARMA $(p, p-1)$ process *without any approximation in the discretization.* This is the reason why we claimed that SDE of order p are continuous time analogues of VARMA $(p, p-1)$ processes.

But, while the ARMA representation (22) is not robust to fractional integration [it does not provide the representation (25) with the same polynomials], the SDE model is maintained through fractional derivation/integration:

Proposition 9. A process $X(t)$, $t \in [0, T]$, satisfies a generalized SDE:

$$\mathrm{d}D^{p-1}X(t) = [M + B_0 X(t) + B_1 DX(t) + \cdots + B_{p-1}D^{p-1}X(t)]\,\mathrm{d}t + \Sigma\,\mathrm{d}W_d(t), \quad -\tfrac{1}{2} < d < \tfrac{1}{2}, \tag{32}$$

if and only if its derivative of order d $X^{(d)}$ satisfies the usual SDE:

$$\mathrm{d}D^{p-1}X^{(d)}(t) = \left[\frac{t^{-d}}{\Gamma(1-d)}M + B_0 X^{(d)}(t) + B_1 DX^{(d)}(t) + \cdots + B_{p-1}D^{p-1}X^{(d)}(t)\right]\mathrm{d}t + \Sigma\,\mathrm{d}W(t),$$

where M is a constant vector.

The proof of the result is thus straightforward owing to the invariance property of Proposition 4 and to the one-to-one correspondence between the coefficients of the SDE and the MA representation given in Proposition 8.

Of course, the differential notation used in (32) can always be rigorously defined from the corresponding integral notation; this type of convention was already used in (14).

For statistical inference purpose, the parametric model (32) in the long memory case [$0 < d < \frac{1}{2}$ and regularity condition (16)] can be characterized in terms of spectral density.

We can expect that the generalized spectral density will be given in this case by

$$f_X(\lambda) = \frac{1}{\lambda^{2d}}\left[(i\lambda)^p Id - \sum_{k=0}^{p-1} B_k (i\lambda)^k\right]^{-1} \Sigma^T\Sigma \left[(i\lambda)^p Id - \sum_{k=0}^{p-1} B_k (i\lambda)^k\right]^{-1*}. \tag{33}$$

Indeed this property is known in the case $d = 0$ (see Bergström, 1990, p. 65) and is rigorously proved in the general case in the appendix.

We emphasize that the expression (33) would allow to generalize to continuous time the maximum likelihood inference for long memory as already studied in univariate discrete time framework by Fox and Taqqu (1986) and Dahlhaus (1989). Indeed, the estimation in the frequency domain of continuous time processes from discrete time observations is a well-known problem [with the folding formula (20)] and may likely be extended to the long-memory multivariate framework. It has been empirically studied and illustrated by Comte (1993), in the scalar case through Monte Carlo experiments.

Moreover, if we are first interested in the estimation of the degree d of fractional differentiation, we can use the Robinson's (1992) semiparametric approach. This approach is an improvement of the GPH's classical one which is founded on the periodogramm. Indeed, because of (21), we know that we are in the Robinson's framework. Moreover, we are able to check that the class of long memory processes we have introduced (as discrete time samples of continuous time processes defined by Proposition 5) satisfies the assumptions used by Robinson (1992) in order to build a consistent asymptotically normal estimator of d and of the intercepts of the log-spectral density (the logarithms of the diagonal terms of $cA_\infty{}^TA_\infty$, see (19)):

Proposition 10. If $X(t) = \int_0^t \{(t-s)^d/\Gamma(d+1)\}\,\tilde{A}(t-s)\,\mathrm{d}W(s)$ *is a continuous time long memory process with* $\lim_{x\to+\infty} x\tilde{A}(x) = A_\infty$ *regular matrix, then Assumptions 1 to 5 of Robinson (1992) are fulfilled. Thus, Robinson's consistent asymptotically normal estimation procedure of d and* $\mathrm{diag}(cA_\infty{}^TA_\infty)$ *can be applied from a discrete time sample,* $X(t_i)_{(1\le i\le N)}$, $t_i - t_{i-1} = \Delta t$, $i = 1,\dots,N$, $N \to +\infty$.

We even prove this proposition in the multivariate case of several fractional orders:

$$A(x) = \left(\tilde{a}_{i,j}(x)\frac{x^{d_i}}{\Gamma(1+d_i)}\right)_{1\le i,j\le n};$$

we recall Robinson's assumptions in the proof of Proposition 10 and his result in the Appendix.

Despite the efficiency of Robinson's estimator is not proved, it could be useful to estimate the parametric model (32) since the computation of $X^{(d)}(t)$ (or of a proxy) may allow us to estimate the parameters of the SDE by the usual VARMA technics (see, e.g., Bergström, 1994). This method is also more precisely studied and illustrated in Comte (1993); it is proved that it works quite well thanks to a suitable filter to compute approximated realizations of the short memory process $X^{(d)}(t)$ when $X(t)$ is observed in discrete time.

The main idea is the following. Let us assume that we have at our disposal some regularly sampled discrete time observations of X, say $X(k\Delta t)$, $k = 0, 1, \dots, N$. Even if we knew the true value of d, we would not be able to exactly compute a discrete time sample path of the process $X^{(d)}(t)$, but only some proxies $\hat{X}^{(d)}(k\Delta t)$, $k = 0, 1, \dots, N$, of $X^{(d)}(k\Delta t)$ by applying the definition of the Ito stochastic integral:

$$X^{(d)}(k\Delta t) = \int_0^{k\Delta t} \frac{(k\Delta t - s)^{-d}}{\Gamma(1-d)}\,\mathrm{d}X(s) = \int_0^{k\Delta t} \frac{s^{-d}}{\Gamma(1-d)}\,\mathrm{d}X(k\Delta t - s)$$
$$\sim \sum_{j=0}^{k} \frac{(j\Delta t)^{-d}}{\Gamma(1-d)}[X((k-j)\Delta t) - X((k-j+1)\Delta t)],$$
$$\hat{X}^{(d)}(k\Delta t) = \sum_{j=0}^{k-1} \frac{[(j\Delta t)^{-d} - ((j+1)\Delta t)^{-d}]}{\Gamma(1-d)} X((k-j)\Delta t), \tag{34}$$

since $X(0) = 0$.

This type of approximation is intuitively all the more accurate when the time sampling interval Δt is smaller and k is bigger. Moreover, for $\Delta t = 1$ in (34), it may work much better than the more standard fractional discrete time differencing $(1 - L)^d$:

$$(1 - L)^d X(t) = \sum_{j=0}^{\infty} \frac{\Gamma(j-d)}{\Gamma(-d)} \frac{X(t-j)}{j!}.$$

In other words, we want to argue that the discrete time ARFIMA model, although more familiar to econometricians, is not a well-suited statistical model to study long memory time series whose generating process is a continuous time one in the sense of (32). A more systematic comparison between the suggested discretizations [e.g., (34)] of our continuous time model and the ARFIMA model is performed in the following subsection, in the case $p = 1$ and $n = 1$ for simplicity.

3.3. Comparison with the ARFIMA model

3.3.1. Discretization of the continuous time model

We deal here with the case of a first order and scalar fractional SDE (29) that is written as

$$\mathrm{d}x(t) = -kx(t)\mathrm{d}t + \sigma \mathrm{d}w_d(t), \qquad k > 0, \quad d \in (0, \tfrac{1}{2}). \tag{35}$$

Then we know two integral expressions of $x(t)$:

$$x(t) = \int_0^t \frac{(t-s)^d}{\Gamma(1+d)} \mathrm{d}x^{(d)}(s) = \int_0^t a(t-s)\,\mathrm{d}w(s),$$

where $a(t - s)$ is given with (12) by

$$a(x) = \frac{\sigma}{\Gamma(1+d)} \frac{\mathrm{d}}{\mathrm{d}x}\left[\int_0^x \mathrm{e}^{-ku}(x-u)^d \,\mathrm{d}u\right]$$
$$= \frac{\sigma}{\Gamma(1+d)}\left(x^d - k\mathrm{e}^{-kx}\int_0^x \mathrm{e}^{ku}u^d\,\mathrm{d}u\right).$$

A discrete time approximation of the x process is a formula to numerically evaluate these integrals using only the values of the involved processes $x^{(d)}(s)$ and $w(s)$ on a discrete partition of $[0, t]$:

$$\frac{j}{n}, \quad j = 0, 1, \ldots, [nt].^6$$

[6] $[z]$ is the integer k such that $k \leqslant z < k + 1$.

A natural way to obtain such approximations (see Comte, 1993) is to approximate the integrands by step functions, which gives the following proxy processes:

$$x_{n,1}(t) = \int_0^t \frac{\left(t - \frac{[ns]}{n}\right)^d}{\Gamma(1+d)} \mathrm{d}x^{(d)}(s)$$

$$= \sum_{j=1}^{[nt]} \frac{\left(t - \frac{j-1}{n}\right)^d}{\Gamma(1+d)} \Delta x^{(d)}\left(\frac{j}{n}\right) + \frac{\left(t - \frac{[nt]}{n}\right)^d}{\Gamma(1+d)} \times \left(x^{(d)}(t) - x^{(d)}\left(\frac{[nt]}{n}\right)\right), \tag{36}$$

$$x_{n,2}(t) = \int_0^t a\left(t - \frac{[ns]}{n}\right) \mathrm{d}w(s)$$

$$= \sum_{j=1}^{[nt]} a\left(t - \frac{j-1}{n}\right) \Delta w\left(\frac{j}{n}\right) + a\left(t - \frac{[nt]}{n}\right)\left(w(t) - w\left(\frac{[nt]}{n}\right)\right), \tag{37}$$

where we use the following notations:

$$\Delta x^{(d)}\left(\frac{j}{n}\right) = x^{(d)}\left(\frac{j}{n}\right) - x^{(d)}\left(\frac{j-1}{n}\right),$$

$$\Delta w\left(\frac{j}{n}\right) = w\left(\frac{j}{n}\right) - w\left(\frac{j-1}{n}\right).$$

Of course, the last terms of (36) and (37) respectively can be neglected for great values of n. So, useful proxies are

$$\hat{x}_n(t) = \sum_{j=1}^{[nt]} \frac{\left(t - \frac{j-1}{n}\right)^d}{\Gamma(1+d)} \Delta x^{(d)}\left(\frac{j}{n}\right), \tag{38}$$

$$\tilde{x}_n(t) = \sum_{j=1}^{[nt]} a\left(t - \frac{j-1}{n}\right) \Delta w\left(\frac{j}{n}\right). \tag{39}$$

Indeed, all these approximations converge towards the x process in the functional sense, on compact sets (i.e., for the supremum norm on compact sets and not only pointwise); this convergence is denoted by $\Rightarrow$. This result is proved in Comte (1993):

Proposition 11. $x_{n,1} \Rightarrow x$, $x_{n,2} \Rightarrow x$, $\hat{x}_n \Rightarrow x$, *and* $\tilde{x}_n \Rightarrow x$ *when n goes to infinity.*

Of course, the approximations of integrals (36) and (37) could be numerically improved by more sophisticated numerical methodologies (trapeze, Simpson, ...). For sake of simplicity, we shall be interested here only in the two processes $\hat{x}_n$ and $\tilde{x}_n$. The proxy $\hat{x}_n$ is the most useful to compare our model with the standard discrete time models, whereas the most tractable for mathematical work is $\tilde{x}_n$.

3.3.2. FRACIMA versus ARFIMA

Eq. (38) provides a proxy $\hat{x}_n$ of x in function of the process $x^{(d)}(j/n)$, $j = 0, 1, \dots, [nt]$, which is an AR(1) process associated to an innovation process $u(j/n), j = 0, 1, \dots, [nt]$. Let us denote by

$$(1 - \rho_n L_n) x^{(d)}\left(\frac{j}{n}\right) = u\left(\frac{j}{n}\right)$$

the representation of this process where L_n is the lag operator corresponding to the sampling scheme $j/n, j = 0, 1, \dots,$

$$L_n Y\left(\frac{j}{n}\right) = Y\left(\frac{j-1}{n}\right),$$

and $\rho_n = \mathrm{e}^{-k/n}$ is the correlation coefficient for the time interval $1/n$.

Since the process $x^{(d)}$ is asymptotically stationary, we can assume, without loss of generality, that its initial values are zero:

$$x^{(d)}\left(\frac{j}{n}\right) = 0 \quad \text{for} \quad j \leqslant 0, \tag{40}$$

which of course implies $u(j/n) = 0$ for $j \leqslant 0$.

Because of (38) and (40), we can write

$$\begin{aligned}\hat{x}_n\left(\frac{j}{n}\right) &= \sum_{i=1}^{j} \frac{(j-i+1)^d}{n^d \Gamma(1+d)} \left[x^{(d)}\left(\frac{i}{n}\right) - x^{(d)}\left(\frac{i-1}{n}\right)\right] \\ &= \left[\sum_{i=0}^{j-1} \frac{(i+1)^d - i^d}{n^d \Gamma(1+d)} L_n^i\right] x^{(d)}\left(\frac{j}{n}\right).\end{aligned}$$

Thus,

$$\hat{x}_n\left(\frac{j}{n}\right) = \left[\sum_{i=0}^{j-1} \frac{(i+1)^d - i^d}{n^d \Gamma(1+d)} L_n^i\right] (1 - \rho_n L_n)^{-1} u\left(\frac{j}{n}\right). \tag{41}$$

Eq. (41) gives a parametrization of the process in two parts:

- a long memory part which corresponds to the filter $\sum_{i=0}^{+\infty} a_i (L_n^i / n^d)$ with

$$a_i = \frac{(i+1)^d - i^d}{\Gamma(1+d)}, \tag{42}$$

- and a short memory part which is characterized by the AR(1) process: $(1-\rho_n L_n)^{-1} u(j/n)$.

Indeed, we can show that the long memory filter is 'long term equivalent' to the usual discrete time long memory filter, $(1-L)^{-d} = \sum_{i=0}^{+\infty} b_i L^i$, with

$$b_i = \frac{\Gamma(i+d)}{\Gamma(i+1)\Gamma(d)}, \tag{43}$$

in the sense that there is a long term relationship (a cointegration relation) between the two types of processes.

In order to see this, let us compare two long memory processes:

$$Y_t = \sum_{i=0}^{+\infty} a_i u_{t-i} \quad \text{and} \quad Z_t = \sum_{i=0}^{+\infty} b_i u_{t-i},$$

where a_i and b_i are defined by (42) and (43) respectively and u_t is any short term memory stationary process.

We can then effectively show (see Appendix) that Y_t and Z_t are cointegrated (in the general meaning of cointegration for long memory processes) since $Y_t - Z_t$ is short memory, i.e.,

$$\sum_{i=0}^{+\infty} |a_i - b_i| < +\infty, \tag{44}$$

whereas

$$\sum_{i=0}^{+\infty} a_i = \sum_{i=0}^{+\infty} b_i = +\infty.$$

It is important to notice that this long term equivalence between our long memory filter and the usual discrete time one $(1-L)^{-d}$ *does not imply* that the standard parametrization ARFIMA $(1, d, 0)$ is well-suited in our framework. Indeed, as far as we are concerned with short memory characteristics, they may be hidden by the short term difference between the two filters (42) and (43). In other words, not only

$$(1-\rho_n L_n)(n(1-L_n))^d \hat{x}_n\left(\frac{j}{n}\right)^7$$

is not in general a white noise, but we are not even sure that

$$(n(1-L_n))^d \hat{x}_n\left(\frac{j}{n}\right)$$

[7] The fractional differencing operator $(1-L)^d$ has to be modified into $(n(1-L_n))^d$ in order to correctly normalize the unit root with respect to the unit period of time.

is an AR(1) process (even though we know that it is a short memory stationary process). In other words, the usual discrete time filter $(1 - L)^d$ introduces some mixing between long and short term characteristics whereas the parsimonious parametric model of these characteristics is the continuous time model.

Moreover, this last claim can be easily checked through Monte Carlo experiments as in Comte (1993). When we simulate a FRACIMA process x along (35) and apply to x the two filters suggested respectively by (42) and (43) we observe, through the partial autocorrelogramm, that, whereas the x process filtered along (42) does appear like an AR(1) process, the same x process filtered along (43) entails significant partial autocorrelations at higher orders.

4. Applications

4.1. Application to asset pricing

Long memory properties of financial time series have been recently well-documented: for instance, Ding, Granger, and Engle (1993) investigate a kind of long memory property of stock market returns that leads to a generalization of ARCH models. New GARCH models with long memory are also studied by Baillie, Bollerslev, and Mikkelsen (1993), as another answer to this study.

Recently also Backus and Zin (1995) have proved the ability of the fractional difference model to mimic some stylized facts about the term structure of interest rates.

Unfortunately, since they have not at their disposal continuous time long memory processes, they are obliged to introduce a discrete time bond pricing model which is not mainstream with respect to the modern continuous time finance. In particular, discrete time modelling leads to some assumptions about risk premia which are not consistent with respect to temporal aggregation (see Cox, Ingersoll, and Ross, 1981).

We want to show in this section that the most popular theories of bond pricing can be extended to long memory continuous time processes, and to reconsider Backus and Zin's conjectures in this framework.

The modern theories of term structure of interest rates assume that this structure depends on a restricted number of 'state variables', the absence of arbitrage allowing to specify the joint distribution of the different term rates processes in function of these state variables. Moreover, thanks to Harrison and Kreps (1979), it is well understood that, under only technical regularity conditions, the absence of arbitrage is equivalent to the existence of an equivalent martingale measure. This means that there is a probability measure Q, equivalent to the probability $\mathbb{P}$ which defines the data generating process, under which the price processes of all securities are Q martingales after normalization at each time t by the value exp $[\int_0^t r(s)\mathrm{d}s]$ of continual re-investment of interests brought by one unit of account held from time 0 at the short rate $r(s)$, $s \in [0, t]$. In other words, if a k-dimensional process $X(t)$, $t \in [0, T]$, defines the k state variables (or 'factors') of the model, $r(s)$ is a deterministic function $R(X(s))$ and the price $B(t, T)$ at time t of a bond with maturity T is given by

$$B(t, T) = \mathrm{E}_t\left[\exp\left(-\int_t^T R(X(s))\,\mathrm{d}s\right)\right], \tag{45}$$

where the expectation E_t is computed with respect to the conditional probability distribution of $(X(s))_{s\in[t,T]}$ given $(X(\tau))_{\tau\in[0,t]}$ associated to the probability measure Q. We shall consider in this illustrative example the simplest case of a one-factor model of interest rates, which leads to assume that

$$X(t) = R(X(t)) = r(t). \tag{46}$$

Leaving for further research a more comprehensive theory of asset pricing in a long memory framework, we only assume here that (45) and (46) can be used in a context where the dynamics under Q of the short term interest rate are given by a generalized first order SDE with respect to a fractional noise:

$$\mathrm{d}r(t) = a(b - r(t))\,\mathrm{d}t + \sigma\,\mathrm{d}w_d(t). \tag{47}$$

This is indeed the natural extension to a long memory setting of the seminal model by Vasicek (1977) defined by (45), (46), and (47) with $d = 0$. As a matter of fact, we know from Proposition 9 that in the general case (47) the d-derivative $r^{(d)}(t)$ is the usual Ornstein–Uhlenbeck process, $\mathrm{d}r^{(d)}(t) = a(b - r^{(d)}(t))\mathrm{d}t + \sigma\mathrm{d}w(t)$, whose dynamics is given by

$$r^{(d)}(t) = r^{(d)}(0) + b(1 - \mathrm{e}^{-at}) + \int_0^t \sigma \mathrm{e}^{a(s-t)}\,\mathrm{d}w(s). \tag{48}$$

With a straightforward generalization of Proposition 4, we can deduce from (48) that

$$r(t) = g(t) + \int_0^t \frac{(t-s)^d}{\Gamma(1+d)}\,\tilde{a}(t-s)\,\mathrm{d}w(s),$$

where the deterministic functions $\tilde{a}$ and g are defined by [see (12) with $C(x) = \sigma\mathrm{e}^{-ax}$]:

$$\begin{cases} \tilde{a}(x) = \sigma - \displaystyle\int_0^x a\sigma \mathrm{e}^{-au}\left(1 - \frac{u}{x}\right)^d \mathrm{d}u, \\ r^{(d)}(0) + b(1 - \mathrm{e}^{-at}) = \displaystyle\int_0^t \frac{(t-s)^{-d}}{\Gamma(1-d)}\,g'(s)\,\mathrm{d}s, \qquad g(0) = r(0). \end{cases}$$

We do obtain a long memory model of short term interest rate in the case $0 < d < \frac{1}{2}$ since in this case we can verify (see Appendix) that

$$\lim_{x \to +\infty} x\tilde{a}(x) = \frac{\sigma d}{a}. \tag{49}$$

Moreover, we can check that, as conjectured by Backus and Zin (1995), the order of fractional differentiation is important for the rate of convergence to zero (when the maturity T goes to infinity) of the variance of the yield $-[1/(T-t)]\ln B(t,T)$ of a long bond $B(t,T)$ which is priced by (45). Nevertheless, contrary to what they say, the rate of convergence is hyperbolic in the case of long memory as well as in the case of short memory:

Proposition 12. Under assumptions (45), (46), and (47) with $0 \leqslant d < \frac{1}{2}$, for a fixed t and $T \to +\infty$, $\operatorname{var}[-[1/(T-t)]B(t,T)]$ converges towards zero at the hyperbolic rate of order $2d-2$ (i.e., the variance is of order T^{2d-2}).

This is thus a generalization of the well-known result of the Vasicek model (corresponding to $d = 0$) stating that the long yield can be explicitly computed with $\gamma(\tau) = (1 - e^{-a\tau})/a$ by

$$-\frac{1}{T-t}\ln B(t,T) = b - \frac{\sigma^2}{2a} - \frac{\gamma(T-t)}{T-t}\left(b - \frac{\sigma^2}{2a} - r(t)\right) - \frac{\sigma^2}{4a}\frac{\gamma^2(T-t)}{T-t}. \tag{50}$$

In fact, the interesting effect that makes a part of Backus and Zin's conjecture true does not appear through looking at the long yield but through looking at the forward rate $f(t,T)$ defined by

$$f(t,T) = -\frac{\partial}{\partial T}\ln B(t,T).$$

Eq. (50) implies indeed that in the case of short memory, the variance of the forward rate has the same order as $(\partial\gamma(T-t)/\partial T)^2 = e^{-2a(T-t)}$, which does mean an *exponential* decrease. But the long memory model that we have built allows the setting of more varying forward rates, as suggested by empirical evidence; indeed, their variance has only an hyperbolic rate of decrease:

Proposition 13. Under assumptions (45), (46), and (47) with $0 < d < \frac{1}{2}$, for a fixed t and $T \to +\infty$, $\operatorname{var}[f(T,t)]$ converges towards zero at hyperbolic rate of order $2d-2$.

The continuous time model that gives a correct solution to the problem of temporal aggregation allows thus to have simultaneously forward rates with both hyperbolic or exponential variability, depending on the fact that the memory is short or long, whereas the long yields which are the aggregation of different short term spot and forward rates have always hyperbolic variability.

4.2. Application to macroeconomics

It is now a well-documented evidence that some macroeconomic time series like real output growth or consumption prices entail some long memory features; see, for instance, Granger and Joyeux (1980) for food prices and Haubrich and Lo (1991) and Sowell (1992) for output growth.

Here we want to explain how temporal aggregation across agents with heterogeneous endowments, risk aversion, and beliefs can produce different orders of fractional integration in macroeconomic time series of prices or interest rates. In his seminal paper, Granger (1980) gives some results about aggregation across individuals of independent short memory autoregressive series in discrete time: the aggregated series have different properties than the individual ones, and if a Beta distribution is set on some of the parameters, fractional processes can be obtained from the aggregation. The idea has been illustrated in another way with dependent series by Gonçalvès and Gouriéroux (1988).

We want to show here that these methodologies of aggregation can be extended to our continuous time framework and that the degree of fractional differencing is thus directly linked to the distribution of heterogeneity. We consider for an illustrative example that, for each individual j, a microeconomic process of interest $x_j(t)$, $t \in [0, T]$, is governed by a usual SDE:

$$\mathrm{d}x_j(t) = -\varphi_j x_j(t)\,\mathrm{d}t + \sigma_j\,\mathrm{d}w_j(t)$$

where the individual parameters (φ_j, σ_j) are independently drawn in a given probability distributions of heterogeneity; the independence assumption is maintained across individuals but not necessary between the two processes φ and σ (whose support is included in $\mathbb{R}_*^+$).

In the line of the Granger's methodology, we can assume that the individual processes w_j are independent Brownian Motions and that we are interested in the aggregated time series $S_N = \sum_{j=1}^{N} x_j$, while in Gonçalvès and Gouriéroux's spirit we can consider that $w_j = w^* + \tilde{w}_j$, where w^* and the $\tilde{w}_j$ are independent Brownian Motions and that we are interested in the aggregated time series $\bar{x}_N = (1/N)\sum_{j=1}^{N} x_j$. In any case, the underlying idea is that there is an individual distribution of heterogeneity which is reflected at the macroeconomic level through a summation or average operator. Of course, the random drawings of the individual heterogeneity are assumed to be stochastically independent of the Brownian Motion of interest.

For the Granger's methodology we assume that φ_j and σ_j are independent random variables and that the distribution of heterogeneity of the process φ is given by a Gamma(b) probability distribution, for a real parameter, b, $0 < b < 2$.

Let $f_j(\lambda)$ be the spectral density of x_j and $\bar{f}_N$ the spectral density of $S_N = \sum_{j=1}^N x_j$. With independence of the x_j, we know

$$\bar{f}_N(\lambda) = \sum_{j=1}^{N} f_j(\lambda),$$

$$f_j(\lambda) = |\hat{a}_j(\lambda)|^2 = \left| \int_0^{+\infty} \mathrm{e}^{-\varphi_j x} \mathrm{e}^{-i\lambda x} \sigma_j \mathrm{d}x \right|^2 = \frac{\sigma_j^2}{\varphi_j^2 + \lambda^2}.$$

Hence $(1/N)\, \bar{f}_N(\lambda)$ converges towards

$$\bar{f}(\lambda) = \mathrm{E}\sigma^2 \int \frac{\mathrm{d}F(\varphi)}{\lambda^2 + \varphi^2}$$

$$= \mathrm{E}\sigma^2 \int_0^{+\infty} \frac{\mathrm{e}^{-t} t^{b-1}}{\Gamma(b)(\lambda^2 + t^2)} \mathrm{d}t$$

$$= \mathrm{E}\sigma^2 \frac{\lambda^{b-2}}{\Gamma(b)} \int_0^{+\infty} \frac{\mathrm{e}^{-\lambda u} u^{b-1}}{u^2 + 1} \mathrm{d}u,$$

and thus:

$$\bar{f}(\lambda) \sim C\, \lambda^{-2d},$$

with

$$C = \frac{\mathrm{E}\sigma^2}{\Gamma(b)} \int_0^{+\infty} \frac{u^{b-1}}{u^2 + 1} \mathrm{d}u, \qquad d = -\frac{b-2}{2} \in (0, 1),$$

with straightforward application of Lebesgue's theorem. This property for $d \in (0, \frac{1}{2})$ is then characteristic of long memory processes. This is a first way to see how aggregation can generate fractional processes and long memory properties.

On the other hand, the Gonçalvès and Gouriéroux's assumptions are perhaps more plausible for time series of individual expectations (about interest rates for instance) since they do not imply the mutual independence of the x_j, taking into account a common component w^* among the w_j. They do not imply either that φ_j is independent of σ_j and allow to prove the following result:

Proposition 14. *Let $\bar{x}_n(t) = (1/n) \sum_{j=1}^n x_j(t)$ with x_j as defined above. Then:*

(i) *$\bar{x}_n(t) \to_P \bar{x}(t)$ for $n \to +\infty$, where $\bar{x}$ admits a representation $\bar{x}(t) = f(t) + \int_0^t \bar{a}(t-s)\,\mathrm{d}w^*(s)$ with f deterministic function.*

(ii) *Aggregation increases correlations, for the stationary processes asymptotically equivalent to x_j and $\bar{x}$: if $\varrho(h) = \mathrm{corr}(x_j(t+h), x_j(t))$, then $\bar{\varrho}(h) = \mathrm{corr}(\bar{x}(t+h), \bar{x}(t)) \geqslant \mathrm{E}[\varrho(h)]$.*

(iii) *If $s_j = \mathrm{e}^{-\varphi_j}\sigma_j$ is independent of φ_j and $\varphi \sim Gamma(d)$, then $\bar{x}(t)$ is a fractional process of order $-d$ (and thus long memory is generated by random* *variables without finite mean or with a prolongation of the gamma law to negative noninteger orders).*

5. Concluding remarks

We did not exclude multivariate processes in our representation of FRACIMA processes (9)/(10), but we implicitly imposed to the n components $X_i(t)$, $i = 1, \dots, n$, of the process of interest to have the same order d of integration, since by Proposition 4 each component has to be 'differentiated d times' in order to obtain a CIMA process. On the other hand, Section 4 has provided theoretical and empirical evidence for multivariate macroeconomic modelling where each component $X_i(t)$ of the process $X(t)$ may have a specific order d_i, and thus instead of (7), we would have $A(x) = \Delta(x)\,\tilde{A}(x)$, with $\tilde{A}$ C^1 on $[0, T]$, and

$$\Delta(x) = \operatorname{diag}\left(\frac{x^{d_i}}{\Gamma(1+d_i)}\right)_{1 \leqslant i \leqslant n}, \qquad |d_i| < \tfrac{1}{2}.$$

Indeed, we could have studied such a class of processes from the point of view of the representations;[8] we stressed anyway that the estimation procedure theoretically studied by Robinson (1992) could be applied to such a class of models. Moreover, it is of some interest to compare those processes, that could be called 'fractional processes of vectorial order' $d = (d_1, \dots, d_n)$, with the class obtained with $A(x) = \tilde{A}(x)\Delta(x)$, Δ as previously, that appears as a class of MA processes w.r.t. a fractional noise of vectorial order.

Proofs

Proof of Lemma 1

Let us consider first $Z(t) = \int_0^t D(t-s)X(s)\,\mathrm{d}s$, i.e.,

$$Z(t) = \int_0^t D(t-s)\left(\int_0^s A(s-u)\,\mathrm{d}W(u)\right)\mathrm{d}s.$$

We assume for simplicity that processes are scalar, but the result is obviously valid for n-dimensional processes. With the Fubini's theorem for stochastic

[8] Such a study was in previous versions of the paper and has been eliminated to shorten it; but it is available from the authors on request.

integrals, as stated in Protter (1990), the interversion of integrals can be done if

$$\int_0^t \int_0^s (D(t-s)A(s-u))^2 \, du \, ds < +\infty,$$

which is condition (C1). Under (C1) we have

$$Z(t) = \int_0^t \left(\int_0^{t-u} D(t-u-s)A(s)\, ds \right) dW(u) = \int_0^t D * A(t-u)\, dW(u),$$

which can be differentiated if $D * A(0) = 0$, which is true, and $D * A$ admits a.e. on $[0, T]$ a square integrable derivative, which is condition (C2). The last representation is then the expression of the derivative of Z since $Y = Z'$. For the proof of (ii), if $\int_0^t D'(t-s)\, dX(s)$ is well-defined, then

$$\begin{aligned} \int_0^t \left[\int_0^s D'(s-u)\, dX(u) \right] ds &= \int_0^t \frac{d}{ds} \left(\int_0^s D'(s-u) X(u)\, du \right) ds \\ &= \int_0^t D'(t-u) X(u)\, du \\ &= Y(t) - D(0)X(t), \end{aligned}$$

since in the case where $\int_0^t D'(t-s)\, dX(s)$ does exist an integration by part gives the result using $d(D(t-s)X(s)) = D(t-s)\, dX(s) - D'(t-s)X(s)\, ds$.

Proof of Proposition 1

If $X(t) = \int_0^t C(t-s)\, dW_d(s)$, we have from Lemma 1 that

$$X(t) = \int_0^t \left(C(0) \frac{(t-s)^d}{\Gamma(d+1)} + \int_0^{t-s} C'(t-s-u) \frac{u^d}{\Gamma(d+1)} du \right) dW(s).$$

This implies

$$x^d \tilde{A}(x) = C(0)x^d + \int_0^t C'(x-u)u^d\, du = \frac{d}{dx} \int_0^x C(x-u)u^d\, du,$$

$$\tilde{A}(x) = C(0) + x \int_0^1 C'(ux)(1-u)^d\, du,$$

as announced by (12).

If conversely we can compute C from $\tilde{A}$, we shall be able to go from a formulation to the other, that is, we shall have the equivalence. But if

$$x^d \tilde{A}(x) = \frac{d}{dx} \int_0^x C(x-s) s^d\, ds,$$

and if we set

$$H(s) = \int_0^s C(s-u)u^d \mathrm{d}u,$$

then we have

$$\int_0^x (x-s)^{-d} s^d \tilde{A}(s)\,\mathrm{d}s = \int_0^x (x-s)^{-d} H'(s)\,\mathrm{d}s = \int_0^x s^{-d} H'(x-s)\,\mathrm{d}s$$
$$= \frac{\mathrm{d}}{\mathrm{d}x}\left[\int_0^x s^{-d} H(x-s)\,\mathrm{d}s\right],$$

as $H(0) = 0$,

$$\int_0^x (x-s)^{-d} s^d \tilde{A}(s)\,\mathrm{d}s = \frac{\mathrm{d}}{\mathrm{d}x}\left[\int_0^x (x-s)^{-d}\left(\int_0^s C(u)(s-u)^d \mathrm{d}u\right)\mathrm{d}s\right]$$
$$= \frac{\mathrm{d}}{\mathrm{d}x}\left[\int_0^x C(u)\left(\int_0^{x-u} (x-u-s)^{-d} s^d \mathrm{d}s\right)\mathrm{d}u\right]$$
$$= \frac{\mathrm{d}}{\mathrm{d}x}\left[\int_0^x C(u)(x-u)\Gamma(1+d)\Gamma(1-d)\,\mathrm{d}u\right]$$
$$= \Gamma(1+d)\Gamma(1-d)\int_0^x C(u)\,\mathrm{d}u,$$

from which we can see that

$$\frac{\mathrm{d}}{\mathrm{d}x}\left[\int_0^x (x-s)^{-d} s^d \tilde{A}(s)\,\mathrm{d}s\right] = \Gamma(1+d)\Gamma(1-d)C(x),$$

which gives the announced formula (11). The resulting regularity of C or $\tilde{A}$ is then obvious.

Proof of Proposition 3

Let

$$X(t) = \int_0^t C(t-s)\,\mathrm{d}W_d(s)$$

and D the solution of

$$D * C(x) = \int_0^x D(u)C(x-u)\,\mathrm{d}u = xId.$$

Proposition 2 ensures the existence and unicity of D as soon as $C(0)$ is invertible and C is C^1, $D(0) = C(0)^{-1}$, D of class C^1. We can write

$$\int_0^t D(t-s)X(s)\,\mathrm{d}s = \int_0^t D(t-s)\left(\int_0^s \tilde{A}(s-u)\frac{(s-u)^d}{\Gamma(d+1)}\mathrm{d}W(u)\right)\mathrm{d}s,$$

where $\tilde{A}(x)x^d = F'(x)$, $F(x) = \int_0^x C(x-s)s^d\,\mathrm{d}s$. The integrals can be interverted if

$$\int_0^t\int_0^s (D(t-s)\tilde{A}(s-u)(s-u)^d)^2\,\mathrm{d}u\,\mathrm{d}s < +\infty,$$

which is true. This implies

$$\int_0^t D(t-s)X(s)\,\mathrm{d}s = \int_0^t\left(\int_0^{t-u} D(t-u-s)\tilde{A}(s)\frac{s^d}{\Gamma(d+1)}\mathrm{d}s\right)\mathrm{d}W(u). \tag{51}$$

Now we study the function that appears in (51):

$$\begin{aligned}
\int_0^x D(x-s)F'(s)\,\mathrm{d}s &= \int_0^x D(s)F'(x-s)\,\mathrm{d}s \\
&= \frac{\mathrm{d}}{\mathrm{d}x}\left[\int_0^x D(s)F(x-s)\,\mathrm{d}s\right] \text{ as } F(0)=0,
\end{aligned}$$

$$\begin{aligned}
\int_0^x D(x-s)F'(s)\,\mathrm{d}s &= \frac{\mathrm{d}}{\mathrm{d}x}\left[\int_0^x D(x-s)F(s)\,\mathrm{d}s\right] \\
&= \frac{\mathrm{d}}{\mathrm{d}x}\left[\int_0^x D(x-s)\left(\int_0^s C(s-v)\,v^d\,\mathrm{d}v\right)\mathrm{d}s\right] \\
&= \frac{\mathrm{d}}{\mathrm{d}x}\left[\int_0^x v^d\left(\int_0^{x-v} D(x-v-s)C(s)\,\mathrm{d}s\right)\mathrm{d}v\right],
\end{aligned}$$

and using the convolution equation $D * C(x) = xI_n$:

$$\begin{aligned}
\int_0^x D(x-s)F'(s)\,\mathrm{d}s &= \frac{\mathrm{d}}{\mathrm{d}x}\left[\int_0^x v^d \times (x-v)I_n\,\mathrm{d}v\right] \\
&= \left(\int_0^x v^d\,\mathrm{d}v\right)I_n = \frac{x^{d+1}}{d+1}I_n.
\end{aligned}$$

We derive

$$\int_0^t D(t-s)\,X(s)\,\mathrm{d}s = \int_0^t \frac{(t-u)^{d+1}}{(d+1)\Gamma(d+1)}\,\mathrm{d}W(u),$$

and finally

$$\int_0^t D(t-s)\,dX(s) = \frac{d}{dt}\int_0^t D(t-s)X(s)\,ds = \int_0^t \frac{(t-u)^d}{\Gamma(d+1)}\,dW(u) = W_d(t).$$

Proof of Proposition 4

We just apply the results of Lemma 1, i.e., we check (C1) and (C2); (C1) is satisfied since

$$\int_0^t\int_0^s ((t-s)^{-d}(s-u)^d \tilde{A}(s-u))^2\,du\,ds < +\infty. \tag{52}$$

Indeed

$$\int_0^s (s-u)^{2d}\,\tilde{A}(s-u)^2\,du = \int_0^s u^{2d}\tilde{A}^2(u)\,du = O(s^{2d+1})$$

with the regularity of $\tilde{A}$, and

$$\int_0^t (t-s)^{-2d}s^{2d+1}\,ds < +\infty$$

with the assumptions on d which imply $-2d+1>0$ and $2d+2>0$; thus the condition (52) is fulfilled. Then we have

$$X^{(d)}(t) = \int_0^t F'(t-u)\,dW(u),$$

with

$$F(x) = \frac{1}{\Gamma(1-d)\Gamma(1+d)}\int_0^x (x-s)^{-d}s^d\tilde{A}(s)\,ds, \tag{53}$$

provided that (C2) is satisfied, i.e., $F' = C$ as defined by (11). This gives the announced expression for $X^{(d)}$. This implies that we just have to check that F is C^2 and that $F'(0)$ is invertible. For that purpose, we notice that F can also be written as

$$F(x) = \frac{x}{\Gamma(1-d)\Gamma(1+d)}\int_0^1 (1-u)^{-d}u^d\tilde{A}(ux)\,du.$$

Then with Lebesgue's theorem,

$$\lim_{x\to 0}\frac{F(x)}{x} = \frac{1}{\Gamma(1-d)\Gamma(1+d)}\int_0^1 (1-u)^{-d}u^d\tilde{A}(0)\,du = \tilde{A}(0),$$

because

$$\int_0^1 (1-u)^{-d}u^d\,du = \frac{\Gamma(1-d)\Gamma(1+d)}{\Gamma(2)} = \Gamma(1-d)\,\Gamma(1+d),$$

with a well-known relation on the beta function.[9] This proves that $F'(0) = C(0) = \tilde{A}(0)$ is invertible as $\tilde{A}(0)$ is assumed to be invertible. Moreover, $\forall x \in [0, T]$,

$$\left\| \frac{\partial}{\partial x}(1-u)^{-d}u^d \tilde{A}(ux) \right\| = \|(1-u)^{-d}u^{d+1}\tilde{A}'(ux)\|$$
$$\leqslant M(1-u)^{-d}u^{d+1} \in L^1([0, T]),$$

where $M = \sup_{t \in [0,T]} \|\tilde{A}'(t)\|$. Lebesgue's theorem of derivation under the integral ensures then that F is differentiable, with derivative $C(x)$ where $\Gamma(1-d)\Gamma(1+d)\,C(x)$ is

$$\int_0^1 (1-u)^{-d}u^d \tilde{A}(ux)\mathrm{d}u + x\int_0^1 (1-u)^{-d}u^{d+1}\tilde{A}'(ux)\mathrm{d}u.$$

In the same way, as A is C^2, F is C^1, with derivative $C'(x)$ and $\Gamma(1-d)\times \Gamma(1+d)\,C'(x)$ is

$$2\int_0^1 (1-u)^{-d}u^{d+1}\tilde{A}'(ux)\mathrm{d}u + x\int_0^1 (1-u)^{-d}u^{d+2}\tilde{A}''(ux)\mathrm{d}u,$$

which is obviously continuous on $[0, T]$.

Proof of Proposition 5

$$\gamma_Y(h) = \operatorname{cov}[Y(t+h), Y(t)] = \frac{1}{\Gamma(d+1)^2}\int_0^{+\infty} (x+h)^d x^d \tilde{A}(x+h)^T \tilde{A}(x)\mathrm{d}x$$

can also be written:

$$\Gamma(1+d)^2\gamma(h) = h^{2d+1}\int_0^{+\infty} u^d(u+1)^d \tilde{A}((u+1)h)^T\tilde{A}(uh)\mathrm{d}u,$$

$$\frac{\Gamma(1+d)^2\gamma(h)}{h^{2d-1}} = \int_0^{+\infty} u^{d-1}(u+1)^{d-1}[(u+1)h\tilde{A}((u+1)h)\times uh^T\tilde{A}(uh)]\mathrm{d}u.$$

Then the assumption (16), $\tilde{A}(x) \sim_{x\to\infty} A_\infty/x$, can be written:

$$\forall \varepsilon > 0, \quad \exists M > 0,$$

so that

$$x, y > M \Rightarrow \|x\tilde{A}(x)y\tilde{A}(y) - A_\infty{}^T A_\infty\| < \varepsilon,$$

where the norm here means the coefficient by coefficient absolute value, i.e., we check that each coefficient of the first matrix tends towards the corresponding

---

[9] The function beta is defined (see Abramowitz and Stegun, 1972) by $B(z, w) = \int_0^1 u^{z-1}(1-u)^{w-1}\mathrm{d}u$ and is known to be equal to $\Gamma(z)\Gamma(w)/\Gamma(z+w)$.

coefficient of the second one. Let ε be given. M is then fixed, and $uh > M \Rightarrow (u+1)h > M$. Then

$$\left\| \frac{\Gamma(1+d)^2 \gamma(h)}{h^{2d-1}} - \left(\int_0^{+\infty} u^{d-1}(u+1)^{d-1}\,\mathrm{d}u \right) A_\infty {}^T A_\infty \right\|$$

$$= \left\| \int_0^{+\infty} u^{d-1}(u+1)^{d-1}\,[(u+1)h\tilde{A}((u+1)h) \times uh^T\tilde{A}(uh) - A_\infty {}^T A_\infty]\mathrm{d}u \right\|$$

$$\leqslant \left\| \int_0^{M/h} u^{d-1}(u+1)^{d-1}[(u+1)h\tilde{A}((u+1)h) \times uh^T\tilde{A}(uh) - A_\infty {}^T A_\infty]\mathrm{d}u \right\|$$

$$+ \varepsilon \underbrace{\int_0^{+\infty} u^{d-1}(u+1)^{d-1}\,\mathrm{d}u}_{\text{constant} < \infty \text{ as } d > 0}$$

We just have to deal with the first term:

$$\left\| \int_0^{M/h} u^{d-1}(u+1)^{d-1}\,[(u+1)h\tilde{A}((u+1)h) \times uh^T\tilde{A}(uh) - A_\infty {}^T A_\infty]\mathrm{d}u \right\|$$

$$\leqslant M^2 \int_0^{M/h} u^{d-1}(u+1)^{d-1}\,\| \tilde{A}((u+1)h)^T\tilde{A}(uh)\|\,\mathrm{d}u$$

$$+ \int_0^{M/h} u^{d-1}(u+1)^{d-1}\,\mathrm{d}u \,\| A_\infty {}^T A_\infty \|$$

$$\leqslant \left(M^2 \times \sup_R \|\tilde{A}\|^2 + \| A_\infty {}^T A_\infty \| \right) \left[\int_0^{M/h} u^{d-1}(u+1)^{d-1}\,\mathrm{d}u \right].$$

The convergence of the integral implies that as soon as $M/h < \eta$, i.e., $h > M/\eta$, $\int_0^{M/h} u^{d-1}(u+1)^{d-1}\,\mathrm{d}u \leqslant \varepsilon_1$ i.e., we can find $B\ (= M/\eta)$ so that $h > B$ implies

$$\left\| \frac{\Gamma(1+d)^2 \gamma(h)}{h^{2d-1}} - \left(\int_0^{+\infty} u^{d-1}(u+1)^{d-1}\,\mathrm{d}u \right) A_\infty {}^T A_\infty \right\| \leqslant C\varepsilon + C_1\varepsilon_1,$$

which can be written as

$$\lim_{h\to\infty} \frac{\Gamma(1+d)^2 \gamma(h)}{h^{2d-1}} = \left(\int_0^{+\infty} u^{d-1}(u+1)^{d-1}\,\mathrm{d}u \right) A_\infty {}^T A_\infty,$$

where the integral exists for $d \in (-\frac{1}{2}, \frac{1}{2})$. Moreover,

$$\int_0^{+\infty} u^{d-1}(u+1)^{d-1}\,\mathrm{d}u = \int_1^{+\infty} (v-1)^{d-1} v^{d-1}\,\mathrm{d}v \; (v = u+1)$$

$$= \int_0^1 \left(\frac{1}{x} - 1\right)^{d-1} \left(\frac{1}{x}\right)^{d-1} \frac{\mathrm{d}x}{x^2} \left(x = \frac{1}{v}\right)$$

$$= \int_0^1 x^{-2d}(1-x)^{d-1}\,\mathrm{d}x$$

$$= \frac{\Gamma(1-2d)\Gamma(d)}{\Gamma(1-d)},$$

which gives for $\gamma(h)$ the announced equivalent.

Proof of Proposition 6

If $X(t) = \int_{-\infty}^{t} A(t-s)\,\mathrm{d}W(s)$, then $f_X(\lambda) = \int_{-\infty}^{+\infty} \mathrm{e}^{-i\lambda h}\gamma(h)\,\mathrm{d}h$, where

$$\forall h \in \mathbb{R}, \quad \gamma(h) = \int_0^{+\infty} A(x+h)^T A(x)\,\mathrm{d}x$$

$$= \int_{\mathbb{R}} A(x+h)^T A(x)\,\mathrm{d}x,$$

with A prolongated by $A(x) = 0$ if $x < 0$, implies

$$f_X(\lambda) = \int_{\mathbb{R}} \mathrm{e}^{-i\lambda h}\gamma(h)\,\mathrm{d}h$$

$$= \int_{\mathbb{R}} \mathrm{e}^{-i\lambda h}\left(\int_{\mathbb{R}} A(x+h)^T A(x)\,\mathrm{d}x\right)\mathrm{d}h$$

$$= \int_{\mathbb{R}} \left(\int_{\mathbb{R}} \mathrm{e}^{-i\lambda(y-x)} A(y)\,\mathrm{d}y\right)^T A(x)\,\mathrm{d}x$$

$$= \int_{\mathbb{R}} \mathrm{e}^{-i\lambda y} A(y)\,\mathrm{d}y \int_{\mathbb{R}} \mathrm{e}^{i\lambda x\,T} A(x)\,\mathrm{d}x.$$

This gives the announced formula: $f_X(\lambda) = \hat{A}(\lambda)\hat{A}(\lambda)^*$. Now we have to check that the integral $\int_0^{+\infty} \mathrm{e}^{-i\lambda x} A(x)\,\mathrm{d}x$ is convergent. $\mathrm{e}^{-i\lambda x} A(x) \sim_0 (x^d/\Gamma(d+1))\,\tilde{A}(0)$ gives the convergence for $x \to 0^+$ as soon as $d + 1 > 0$, a condition which is *a fortiori* fulfilled in our case ($|d| < \frac{1}{2}$). Near $+\infty$, (16) implies that $x^d \tilde{A}(x) \sim x^{d-1} A_\infty$, so that we just have to check the convergence of $\int_1^B \mathrm{e}^{-i\lambda x} x^{d-1}\,\mathrm{d}x$ for $B \to +\infty$; the result is obvious for $d < 0$, and for $0 < d < \frac{1}{2}$, we have by integration by parts:

$$\int_1^B x^{d-1}\mathrm{e}^{-i\lambda x}\,\mathrm{d}x = \left[\frac{-\mathrm{e}^{-i\lambda x}}{i\lambda} x^{d-1}\right]_1^B + \frac{d-1}{i\lambda}\int_1^B x^{d-2}\mathrm{e}^{-i\lambda x}\,\mathrm{d}x.$$

The first term admits then a limit for $B \to +\infty$ and $d < 1$, and the second term is an absolutely convergent integral for $d < 1$.

Proof of Proposition 7

$$\hat{A}(\lambda) = \int_0^{+\infty} e^{-i\lambda x} x^d \tilde{A}(x)\,dx = \frac{1}{\lambda^{d+1}} \int_0^{+\infty} e^{-iu} u^d \tilde{A}\left(\frac{u}{\lambda}\right) du.$$

Now, (16) implies again that for $\lambda \to 0^+$, $\tilde{A}(u/\lambda) \sim (\lambda/u) A_\infty$, so that with the same technics as in Proposition 5 we have

$$\lim_{\lambda \to 0} \lambda^d \hat{A}(\lambda) = \left(\int_0^{+\infty} e^{-iu} u^{d-1}\,du \right) A_\infty,$$

which with $f_X(\lambda) = \hat{A}(\lambda)\hat{A}(\lambda)^*$ gives (19), since the integral $\int_0^{+\infty} e^{-iu} u^{d-1}\,du$ is convergent near zero for $d > 0$ and near $+\infty$ by the integration by part:

$$\int_1^{+\infty} e^{-iu} u^{d-1}\,du = \left[\frac{-1}{i} e^{-iu} u^{d-1} \right]_1^{+\infty} + \frac{d-1}{i} \int_1^{+\infty} e^{-iu} u^{d-2}\,du.$$

The first term is finite since $d - 1 < 0$, and the second one is a convergent integral as soon as $d < 1$. We can notice that we found that c in (19) is

$$c = \left| \int_0^{+\infty} e^{-iu} u^{d-1}\,du \right|^2.$$

Proof of Proposition 8

We suppose $M \equiv 0$ for simplicity. We set

$$\bar{B} = \begin{bmatrix} 0 & I & 0 & \cdots & 0 \\ 0 & 0 & I & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & I \\ B_0 & B_1 & B_2 & \cdots & B_{p-1} \end{bmatrix}, \quad Y(t) = \begin{bmatrix} X(t) \\ DX(t) \\ \vdots \\ D^{p-1}X(t) \end{bmatrix}, \quad W = \begin{bmatrix} 0 \\ \vdots \\ 0 \\ W \end{bmatrix},$$

and

$$\bar{\Sigma} = \mathrm{diag}(0, \ldots, 0, \Sigma).$$

If $X(t)$ is solution of (29), then $Y(t)$ satisfies the first-order SDE, $dY(t) = \bar{B}Y(t)dt + \bar{\Sigma}\,d\bar{W}(t)$, whose solution is known to be $Y(t) = \int_0^t \exp(\bar{B}(t-s))\,\bar{\Sigma}\,dW(s)$. With the form of Y, $\bar{W}$, and $\bar{\Sigma}$, we obtain back the moving average form of X and formula (30) for $D^{p-1}X$, Moreover by integration by parts, we have

$$\int_0^t B(t-s)\,\mathrm{d}D^{p-1}X(s)$$

$$=\int_0^t [d[B(t-s)D^{p-1}X(s)] + B'(t-s)D^{p-1}X(s)\,\mathrm{d}s]$$

$$=B(0)^{-1}D^{p-1}X(t)+\int_0^t B'(t-s)D^{p-1}X(s)\,\mathrm{d}s \quad \text{as} \quad D^{p-1}X(0)=0.$$

With another integration by part, we have

$$\int_0^t B(t-s)\,\mathrm{d}D^{p-1}X(s)=\Sigma^{-1}D^{p-1}X(t)-\Sigma^{-1}B_{p-1}D^{p-2}X(t)$$
$$+\int_0^t B''(t-s)D^{p-2}X(s)\,\mathrm{d}s,$$

and with immediate recurrence,

$$\int_0^t B(t-s)\,\mathrm{d}D^{p-1}X(s)=\Sigma^{-1}D^{p-1}X(t)-\Sigma^{-1}\sum_{k=1}^{p-1}B_{p-k}D^{p-k-1}X(t)$$
$$+\int_0^t B^{(p)}(t-s)X(s)\,\mathrm{d}s.$$

Eq. (29) fulfilled by X can then be written as $d\int_0^t B(t-s)\,\mathrm{d}D^{p-1}X(s)=\mathrm{d}W(t)$, which gives the representation (31).

Proof of Eq. (33)

First, if X admits CIMA representations (30) and (31), then

$$f_X(\lambda)=\left[(i\lambda)^p Id-\sum_{k=0}^{p-1}B_k(i\lambda)^k\right]^{-1}\Sigma^T\Sigma\left[(i\lambda)^p Id-\sum_{k=0}^{p-1}B_k(i\lambda)^k\right]^{-1*}, \quad (54)$$

where $f_X(\lambda)$ denotes the spectral density of the continuous time process X. Indeed, $f_X(\lambda)=\hat{A}(\lambda)\times\hat{A}(\lambda)^*$, so that we have to compute $\hat{A}(\lambda)=\int_0^{+\infty}\mathrm{e}^{i\lambda x}A(x)\,\mathrm{d}x$.

As $B*A^{(p-1)}(x)=\int_0^x B(u)A^{(p-1)}(x-u)\,\mathrm{d}u=xI_n$, i.e., $(d^{p-1}/\mathrm{d}x^{(p-1)})\,(B*A(x))=xI_n$ since $A^{(i)}(0)=0$, $0\leqslant i\leqslant p-2$, we have $B*A(x)=(x^p/p!)I_n$ as $B*A^{(i)}(0)=0, 0\leqslant i\leqslant p-2$. Let $z\in C$, $\mathrm{Re}(z)>0$, and let L denote the Laplace Transform;[10] then

$$L(B*A)(z)=LB(z)\times LA(z)=\frac{I_n}{z^{(p+1)}}$$

[10] I.e., $Lh(z)=\int_0^{+\infty}\mathrm{e}^{-tz}h(t)\,\mathrm{d}t$.

as

$$L(x^p)(z) = \int_{\mathbb{R}^+} e^{-tz} t^p \, dt = \frac{\Gamma(p+1)}{z^{p+1}}.$$

Using $L(x^k)(z) = \Gamma(k+1)/z^{(k+1)}$, $\forall k$, we find

$$LB(z) = \Sigma^{-1} \left[\frac{1}{z} Id - \sum_{k=1}^{p} \frac{B_{p-k}}{z^{k+1}} \right],$$

$$LA(z) = \frac{LB(z)^{-1}}{z^{p+1}} = \left[z^p \, Id - \sum_{k=0}^{p-1} B_k z^k \right]^{-1} \Sigma.$$

Since $A(x) = 0$ for $x < 0$ and the above formula still has sense for $\operatorname{Re}(z) \to 0$, we can write

$$\hat{A}(\lambda) = \left[(i\lambda)^p \, Id - \sum_{k=0}^{p-1} B_k (i\lambda)^k \right]^{-1} \Sigma,$$

which with (18) gives (54).

Now it can be seen from Lemma 1 that $X(t) = \int_0^t C(t-s)\,dW_d(s)$ is $p-1$ times differentiable if and only if C is, with $C(0) = C'(0) = \cdots = C^{(p-2)}(0) = 0$. We know moreover from (18) that $f_X(\lambda) = \hat{A}(\lambda)\hat{A}(\lambda)^*$, where $A(x) = (x^d/\Gamma(d+1))\tilde{A}(x)$ here. With (12) and $Lv'(t) = tLv(t)$ if $v(0) = 0$ and $\lim_{h \to +\infty} v(x) e^{-tx} = 0$, then for any $z \in C$ so that $\operatorname{Re}(z) > 0$,

$$L(x^d \tilde{A}(x))(z) = zLC(z)L(s^d)(z) = zLC(z)\frac{\Gamma(d+1)}{z^{d+1}} = \frac{\Gamma(d+1)}{z^d} LC(z),$$

since $\tilde{A}$ is bounded in $\mathbb{R}$. LC is known from (54). This implies

$$LA(z)LA(z)^* = \frac{1}{|z|^{2d}} LC(z)LC(z)^*,$$

and as all expressions have sense for $\operatorname{Re}(z) \to 0$ and $D * C^{(p-1)}(x) = xI_n$, we have (33).

Proof of Proposition 10

Assumption 1: $\exists C_j \in (0, +\infty)$, $\exists d_j \in (-\frac{1}{2}, \frac{1}{2})$, $\exists a \in (0, 2]$ so that $f_{j,j}^D(\lambda) = C_j \lambda^{-2d_j} + O(\lambda^{a-2d_j})$ for $\lambda = 1, \ldots, n$ as $j \to 0^+$.

Assumption 2: In a neighbourhood $(0, \varepsilon)$ of the origin, $f_{j,j}^D$ is differentiable and $(d/d\lambda) \ln f_{jj}^D(\lambda) = O(1/\lambda)$ as $\lambda \to 0^+, j = 1, \ldots, n$.

Assumption 3: For some $\beta \in [1, 2]$, $R_{j,k}(\lambda) - R_{j,k}(0) = O(\lambda^\beta)$ as $\lambda \to 0^+$, for $j = 1, \ldots, n$, where $R_{j,k}$ is the coherency between X_{jt} and X_{kt}: $R_{j,k}(\lambda) = f_{j,k}^D(\lambda)/\sqrt{f_{jj}^D(\lambda) f_{k,k}^D(\lambda)}$.

Assumption 4: $R(0)$ is nonsingular.

Assumption 5: X_t, $t = 1, 2, \dots$, is a Gaussian process

is obviously fulfilled since W is an n-dimensional Brownian motion.

Assumption 6: As $n \to +\infty$,

$$l/m + m^{1/2}\log(m)/l + (\log(n))^2/m + m^{1+1/2\min(\alpha,\beta)}/n \to 0,$$

but the integers m and l are 'user-chosen positive integers, which both tend to infinity with n, but more slowly, with $l/m \to 0$ also'.

In the case, $A(x) = \mathrm{diag}(x^{d_i}/\Gamma(1+d_i))\,\tilde{A}(x)$, we have $f_X(\lambda) = \hat{A}(\lambda)\hat{A}(\lambda)^*$ and

$$LA(z)LA(z)^* = \mathrm{diag}\left(\frac{1}{z^{d_i}}\right)LC(z)LC(z)^*\,(1/\bar{z}^{d_i}),$$

with

$$C(x) = \frac{\mathrm{d}}{\mathrm{d}x}\left[\int_0^x \mathrm{diag}\left(\frac{(x-s)^{-d_i}s^{d_i}}{\Gamma(1+d_i)\Gamma(1-d_i)}\right)\tilde{A}(s)\,\mathrm{d}s\right].$$

Then provided that the principal complex determination of the logarithm is chosen: $\ln(i) = i\pi/2$, $i^{d_j} = \exp(d_j \ln(i)) = \mathrm{e}^{id_j\pi/2}$ is well defined and $\overline{i^{d_j}} = \mathrm{e}^{-id_j\pi/2}$, we can write

$$f_X(\lambda) = \mathrm{diag}\left(\frac{\mathrm{e}^{-id_j\pi/2}}{\lambda^{d_j}}\right)\hat{C}(\lambda)\hat{C}(\lambda)^*\,\mathrm{diag}\left(\frac{\mathrm{e}^{id_j\pi/2}}{\lambda^{d_j}}\right). \tag{55}$$

In particular,

$$f_{X_i}(\lambda) = \frac{1}{\lambda^{2d_i}}\sum_{k=1}^{n}|\hat{C}_{i,k}(\lambda)|^2,$$

which illustrates that the spectral density of a component of the process X_i is associated with a specific order d_i. This is the reason why the semiparametric results of Robinson for the estimation of a vector of orders can be applied. A parametrization of C derived from the case of generalized SDE can then be set through choosing C as the inverse by convolution of the usual B polynomial function (i.e., B is polynomial and C is solution of $B * C(x) = xI_n$, $\forall x$).

We check that the previous spectral density satisfies the Assumptions 1 to 4 since Assumption 5 is obviously satisfied and Assumption 6 is a technical assumption that just has to be done:

Assumption 1: Let us recall that $f(\lambda)$ is given by (55) and $f^D(\lambda)$ by (20) that we write for $\Delta t = 1$. Thus,

$$f_{j,j}(\lambda) - \frac{1}{\lambda^{2d_j}}[\hat{C}(0)\hat{C}(0)^*]_{j,j} = \frac{1}{\lambda^{2d_j}}([\hat{C}(\lambda)\hat{C}(\lambda)^* - \hat{C}(0)\hat{C}(0)^*]_{j,j}),$$

and $\hat{C}(\lambda)$ is defined and C^1 in 0 so that $\hat{C}(\lambda) = \hat{C}(0) + \lambda \hat{C}'(\zeta)$ with $\zeta \in (0, \lambda)$ which implies $f_{j,j}(\lambda) = (1/\lambda^{2d_j}) [\hat{C}(0)\hat{C}(0)^*]_{j,j} + O(\lambda^{1-2d_j})$. With the constant terms $\sum_{n \in Z^*} f(2n\pi)$ which are of order $O(1)$, we have

$$f_{j,j}^D(\lambda) = \frac{C_j}{\lambda^{2d_j}} + O(\lambda^{a-2d_j}),$$

with

$$C_j = [\hat{C}(0)\hat{C}(0)^*]_{j,j}, \qquad a = 2d_j \in (0, 2] \quad \text{for} \quad d_j > 0,$$

since the sum of a term of order $O(1)$ and a term of order $O(\lambda^{1-2d_j})$ is of order $O(1)$.

Assumption 2: In the same way as previously, we have $f'_{j,j}(\lambda) \sim (-2d_j/\lambda^{2d_j+1}) \times [\hat{C}(0)\hat{C}(0)^*]_{j,j}$, which implies that $f'^D_{j,j}(\lambda)$ admits the same equivalent, and thus that $f'^D_{j,j}(\lambda)/f^D_{j,j}(\lambda) = O(1/\lambda)$ for $\lambda \to 0^+$.

Assumption 3: Choosing $\beta = 1$, Assumption 3 will be fulfilled if $R_{j,k}$ admits a bounded derivative near zero, which is what we check.

$$f_{j,k}^D(\lambda) = \frac{C_{j,k}}{\lambda^{d_j+d_k}} + O(1)$$

since $1 - (d_j + d_k) > 0$ and

$$f'^D_{j,k}(\lambda) = \frac{-(d_j + d_k) C_{j,k}}{\lambda^{d_j+d_k+1}} + O\left(\frac{1}{\lambda^{d_j+d_k}}\right),$$

where $C_{j,k} = e^{-i\pi/2(d_j-d_k)}[\hat{C}(0)\hat{C}(0)^*]_{j,k}$. Thus:

$$R'_{j,k}(\lambda) = \frac{2f'^D_{j,k}(\lambda)f^D_{j,j}(\lambda)f^D_{k,k}(\lambda) - f^D_{j,k}(\lambda)f'^D_{j,j}(\lambda)f^D_{k,k}(\lambda) - f^D_{j,k}(\lambda)f^D_{j,j}(\lambda)f'^D_{k,k}(\lambda)}{2(f^D_{j,j}(\lambda)f^D_{k,k}(\lambda))^{3/2}},$$

$$R'_{j,k}(\lambda) = \frac{O(\lambda^{-3(d_j+d_k)})}{((C_j\lambda^{-2d_j} + O(1))(C_k\lambda^{-2d_k} + O(1)))^{3/2}}$$
$$= \frac{O(1)}{(C_jC_k)^{3/2}(1 + O(\lambda^{2\max(d_j,d_k)}))^{3/2}},$$

which gives the existence and thus boundedness of the derivative of $R_{j,j}$ in zero.

Assumption 4: Let us assume first that $\hat{C}(0)$ exists and is regular. The definition of $R_{j,k}$ implies

$$R(\lambda) = \operatorname{diag}\left(\frac{1}{\sqrt{f_j^D(\lambda)}}\right) f^D(\lambda) \operatorname{diag}\left(\frac{1}{\sqrt{f_j^D(\lambda)}}\right);$$

moreover for $\lambda \to 0^+$, we have

$$f_j(\lambda) \sim \frac{1}{\lambda^{2d_j}} \sum_{k=0}^{n} |\hat{C}_{j,k}(0)|^2,$$

since we assumed the existence of $\hat{C}(0)$, and $\sum_{k=0}^{n} |\hat{C}_{j,k}(0)|^2 \neq 0$ with the regularity of $\hat{C}(0)$. Let us set

$$C_j = \left(\sum_{k=0}^{n} |\hat{C}_{j,k}(0)|^2 \right)^{-1/2}.$$

Then using that $f^D(\lambda) \sim f(\lambda)$ for $\lambda \to 0^+$ and formula (55), we find that $R(0) = \operatorname{diag}(C_j)\hat{C}(0)\hat{C}(0)^* \operatorname{diag}(C_j)$. Since $C_j \neq 0$, $\forall j = 1, \dots, n$, $R(0)$ is thus regular if and only if $\hat{C}(0)$ is.

Now we have to check that the existence and regularity of $\hat{C}(0)$ is implied by the existence and the regularity of the limit, $\lim_{x \to +\infty} x\tilde{A}(x) = A_\infty$, which is the true assumption of the proposition. The expression of C in function of $\tilde{A}$ implies that

$$\begin{aligned}
\hat{C}(0) &= \int_0^{+\infty} C(x)\,\mathrm{d}x \\
&= \lim_{x \to +\infty} \int_0^x \operatorname{diag}\left(\frac{(x-s)^{d_i} s^{d_i}}{\Gamma(1-d_i)\Gamma(1+d_i)} \right) \tilde{A}(s)\,\mathrm{d}s \\
&= \lim_{x \to +\infty} \int_0^1 \operatorname{diag}\left(\frac{(1-s)^{d_i - 1} s^{d_i}}{\Gamma(1-d_i)\Gamma(1+d_i)} \right) (xs)\,\tilde{A}(xs)\,\mathrm{d}s.
\end{aligned}$$

Since $u\tilde{A}(u)$ is bounded coefficient by coefficient for $u \to +\infty$, null for $u = 0$, defined and continuous on $\mathbb{R}^+$; this function is thus coefficient by coefficient bounded on $\mathbb{R}^+$, i.e., there is a matrix $M = (m_{i,j})$ satisfying $|u\tilde{A}_{i,j}(u)| \leqslant m_{i,j}$, $\forall u \in \mathbb{R}^+$. The Lebesgue's theorem gives thus:

$$\begin{aligned}
&\lim_{x \to +\infty} \int_0^1 \operatorname{diag}\left(\frac{(1-s)^{d_i - 1} s^{d_i}}{\Gamma(1-d_i)\Gamma(1+d_i)} \right) (xs)\,\tilde{A}(xs)\,\mathrm{d}s \\
&\quad = \operatorname{diag}\left(\int_0^1 \frac{(1-s)^{d_i - 1} s^{d_i}}{\Gamma(1-d_i)\Gamma(1+d_i)}\,\mathrm{d}s \right) A_\infty,
\end{aligned}$$

so that

$$\hat{C}(0) = \operatorname{diag}\left(\frac{1}{d_i} \right) A_\infty,$$

which gives the result.

Proof of the cointegration formula (44)

We consider the development of a_k as given by (42), for great values of k:

$$a_k = \frac{1}{\Gamma(d)}\left(k^{d-1} - \frac{(d-1)}{2}k^{d-2} + \mathrm{o}(k^{d-2})\right),$$

and we use the Stirling formula:

$$\Gamma(z) = \mathrm{e}^{-z}z^{z-1/2}(2\pi)^{1/2}\left(1 + \frac{1}{12z} + \mathrm{o}\left(\frac{1}{z}\right)\right), \qquad z \to +\infty,$$

as given in Abramowitz and Stegun (1972, 6.1.37, p. 257) to compute the development of b_k as given by (43), for great values of k:

$$b_k = \frac{1}{\Gamma(d)}\left(k^{d-1} + \frac{d(d-1)}{2}k^{d-2} + \mathrm{o}(k^{d-2})\right).$$

Then $a_k - b_k$ is of order $\mathrm{O}(k^{d-2})$, and thus the terms are summable since $d-2<-1$.

Proof of Eq. (49)

We know that $C(x) = \sigma\mathrm{e}^{-ax}$; then

$$a(x) = \frac{\mathrm{d}}{\mathrm{d}x}\left[\int_0^x (x-u)^d C(u)\,\mathrm{d}u\right]$$

implies

$$a(x) = \sigma x^d - \sigma a\int_0^x u^d \mathrm{e}^{-a(x-u)}\,\mathrm{d}u.$$

Now let d be positive, $d \in (0, \frac{1}{2})$, then $\tilde{a}(x) = a(x)/x^d$ can be computed as

$$\tilde{a}(x) = \sigma\frac{d}{x^d}\int_0^x (x-s)^{d-1}\mathrm{e}^{-as}\,\mathrm{d}s,$$

and thus:

$$x\tilde{a}(x) = \sigma d\left[\int_0^{x/2}\left(1-\frac{s}{x}\right)^{d-1}\mathrm{e}^{-as}\,\mathrm{d}s + \int_{x/2}^{x}\left(1-\frac{s}{x}\right)^{d-1}\mathrm{e}^{-as}\,\mathrm{d}s\right].$$

Lebesgue's theorem can be applied to the first right-hand side term since $(1-s/x)^{d-1} \leqslant (\frac{1}{2})^{d-1}$ for $s \in (0, x/2)$, so that its limit for $x \to +\infty$ is $\int_0^{+\infty}\mathrm{e}^{-ax}\,\mathrm{d}x = 1/a$. The second right-hand side term is $\leqslant x\mathrm{e}^{-ax}/d2^d$, whose limit is zero for $x \to +\infty$. This implies that for $d > 0$, $\lim_{x\to+\infty} x\tilde{a}(x) = \sigma(d/a)$ [and thus $a(x) \sim x^{d-1}a_\infty$ for $x \to +\infty$, with $a_\infty = \sigma d/a$].

Proof of Proposition 12

First we compute $B(t, T)$ with assumptions (45), (46), and (47):

$$B(t,T) = \mathrm{E}^*\left[\mathrm{e}^{-\int_t^T r(s)\mathrm{d}s} | \mathscr{F}_t\right] = \mathrm{E}^*\left[\mathrm{e}^{-\int_t^T g(s)\mathrm{d}s - \int_t^T\left(\int_0^s a(s-u)\mathrm{d}\tilde{W}(u)\right)\mathrm{d}s} | \mathscr{F}_t\right],$$

and Fubini's theorem for stochastic integrals allows us to write:

$$\int_t^T\left(\int_0^s a(s-u)\mathrm{d}\tilde{W}(u)\right)\mathrm{d}s$$

$$= \int_0^T\left(\int_0^s a(s-u)\mathrm{d}\tilde{W}(u)\right)\mathrm{d}s - \int_0^t\left(\int_0^s a(s-u)\mathrm{d}\tilde{W}(u)\right)\mathrm{d}s$$

$$= \int_0^T\left(\int_u^T a(s-u)\mathrm{d}s\right)\mathrm{d}\tilde{W}(u) - \int_0^t\left(\int_u^t a(s-u)\mathrm{d}s\right)\mathrm{d}\tilde{W}(u)$$

$$= \int_t^T\left(\int_0^{T-u} a(s)\mathrm{d}s\right)\mathrm{d}\tilde{W}(u) + \int_0^t\left(\int_0^{T-u} a(s)\mathrm{d}s - \int_0^{t-u} a(s)\mathrm{d}s\right)\mathrm{d}\tilde{W}(u)$$

$$= \int_t^T\left(\int_0^{T-u} a(s)\mathrm{d}s\right)\mathrm{d}\tilde{W}(u) + \int_0^t\left(\int_{t-u}^{T-u} a(s)\mathrm{d}s\right)\mathrm{d}\tilde{W}(u).$$

This implies that B can be written:

$$B(t,T) = \mathrm{e}^{-\int_t^T g(s)\mathrm{d}s}$$

$$\times\,\mathrm{E}^*\left[\exp\left(-\underbrace{\int_t^T\left(\int_0^{T-u} a(s)\mathrm{d}s\right)\mathrm{d}\tilde{W}(u)}_{\underline{\in}\mathscr{F}_t} - \underbrace{\int_0^t\left(\int_{t-u}^{T-u} a(s)\mathrm{d}s\right)\mathrm{d}\tilde{W}(u)}_{\perp\mathscr{F}_t}\right)\Bigg|\mathscr{F}_t\right],$$

$$B(t,T) = \mathrm{e}^{-\int_t^T g(s)\mathrm{d}s}\,\mathrm{E}^*\left[\mathrm{e}^{-x-\int_t^T\left(\int_0^{T-u} a(s)\mathrm{d}s\right)\mathrm{d}\tilde{W}(u)} | \mathscr{F}_t\right]\Bigg| x = \int_0^t\left(\int_{t-u}^{T-u} a(s)\mathrm{d}s\right)\mathrm{d}\tilde{W}(u),$$

$$B(t,T) = \mathrm{e}^{-\int_t^T g(s)\mathrm{d}s}\left[\mathrm{e}^{-\mathrm{E}^*U + 1/2\operatorname{var}U}\right]\Big| x = \int_0^t\left(\int_{t-u}^{T-u} a(s)\mathrm{d}s\right)\mathrm{d}\tilde{W}(u),$$

where

$$U = x + \int_t^T \left(\int_0^{T-u} a(s)\,\mathrm{d}s \right) \mathrm{d}\tilde{W}(u).$$

This implies

$$\mathrm{E}^* U = x$$

and

$$\operatorname{var} U = \int_t^T \left(\int_0^{T-u} a(s)\,\mathrm{d}s \right)^2 \mathrm{d}u = \int_0^{T-t} \left(\int_0^{T-t-u} a(s)\,\mathrm{d}s \right)^2 \mathrm{d}u$$

$$= \int_0^{T-t} \left(\int_0^{u} a(s)\,\mathrm{d}s \right)^2 \mathrm{d}u.$$

This implies thus:

$$B(t, T) = \exp\left(-\int_t^T g(s)\,\mathrm{d}s + \tfrac{1}{2} \int_0^{T-t} \left(\int_0^{u} a(s)\,\mathrm{d}s \right)^2 \mathrm{d}u \right.$$

$$\left. - \int_0^t \left(\int_{t-u}^{T-u} a(s)\,\mathrm{d}s \right) \mathrm{d}\tilde{W}(u) \right).$$

Now let

$$y(t, T) = -\frac{1}{T-t} \ln B(t, T).$$

Then:

$$\operatorname{var}(y(t, T)) = \frac{1}{(T-t)^2} \int_0^t \left(\int_{t-u}^{T-u} a(s)\,\mathrm{d}s \right)^2 \mathrm{d}u$$

$$= \frac{1}{(T-t)^2} \int_0^t \left(\int_0^{T-t} a(s+v)\,\mathrm{d}s \right)^2 \mathrm{d}v.$$

Then for $T \to +\infty$, and using $a(s+v) \sim a_\infty (s+v)^d$, we have

$$V(y(t, T)) \sim \frac{1}{(T-t)^2} \int_0^t \left(\int_A^{T-t} a(s+v)\,\mathrm{d}s \right)^2 \mathrm{d}v,$$

$$V(y(t, T)) \sim \frac{a_\infty^2}{d^2 (T-t)^2} \int_0^t (T-t+v)^{2d}\,\mathrm{d}v,$$

$$V(y(t, T)) \sim \frac{a_\infty^2}{d^2 T^2} \frac{1}{2d+1} (T^{2d+1} - (T-t)^{2d+1}),$$

and the very last term can be written

$$T^{2d+1}\left(1-\left(1-\frac{t}{T}\right)^{2d+1}\right) \sim T^{2d+1}(2d+1)\frac{t}{T}.$$

Lastly we have:

$$V(y(t,T)) \sim \frac{a_\infty^2 t}{d^2} T^{2d-2}.$$

Proof of Proposition 13

The result is quite straightforward since we have with formula (49):

$$Vf(t,T) = \int_0^t a(T-u)^2 \,\mathrm{d}u = \int_{T-t}^{T} a^2(u)\,\mathrm{d}u$$

and for great values of T,

$$\int_{T-t}^{T} a^2(u)\,\mathrm{d}u \sim a_\infty \int_{T-t}^{T} x^{2(d-1)}\,\mathrm{d}x$$

and

$$\int_{T-t}^{T} x^{2(d-1)}\,\mathrm{d}x = \frac{T^{2d-1}}{2d-1}\left(1-\left(1-\frac{t}{T}\right)^{2d-1}\right) \sim tT^{2d-2}.$$

which gives the polynomial rate of convergence and the result.

Proof of Proposition 14

(i) $$\bar{x}_n(t) = \frac{1}{n}\left[\sum_{j=1}^{n} (\mathrm{e}^{-\varphi_j t} x_j(0) + \int_0^t \mathrm{e}^{-\varphi_j(t-s)}\sigma_j \mathrm{d}w_j(s)\right]$$

can also be written:

$$\bar{x}_n(t) = \frac{1}{n}\sum_{j=1}^{n} \mathrm{e}^{-\varphi_j t} x_j(0) + \int_0^t \frac{1}{n}\sum_{j=1}^{n} (\mathrm{e}^{-\varphi_j(t-s)}\sigma_j)\,\mathrm{d}w^*(s)$$
$$+\frac{1}{n}\sum_{j=1}^{n}\int_0^t \mathrm{e}^{-\varphi_j(t-s)}\,\mathrm{d}\tilde{w}_j(s).$$

The random variables $\mathrm{e}^{-\varphi_j t}x_j(0)$ and $\int_0^t \mathrm{e}^{-\varphi_j(t-s)}\mathrm{d}\tilde{w}_j(s)$ are i.i.d. so that their empirical mean converges towards their expectation a.s. with the strong law of large numbers, i.e., their limits for $n \to +\infty$ are respectively $\mathrm{E}[\mathrm{e}^{-\varphi(t-s)}x(0)]$ and 0 since all variables are zero-mean. Moreover, $(1/n)\sum_{j=1}^{n} \mathrm{e}^{-\varphi_j(t-s)}\sigma_j$ goes to $\mathrm{E}[\mathrm{e}^{-\varphi(t-s)}\sigma]$ and Lebesgue's theorem for stochastic integrals [if $X_n(s)$ are bounded, $X_n(s) \to X(s)$ a.s., then $\int_0^t X_n(s)\,\mathrm{d}W(s) \to \int_0^t X(s)\,\mathrm{d}W(s)$ in probability, uniformly on all compacts] implies, since all variables are positive so that we have boundedness, that

$$\int_0^t \frac{1}{n}\sum_{j=1}^{n} (\mathrm{e}^{-\varphi_j(t-s)}\sigma_j)\,\mathrm{d}w^*(s) \to \int_0^t \mathrm{E}[\mathrm{e}^{-\varphi(t-s)}\sigma]\,\mathrm{d}w^*(s)$$

in probability. Thus:

$$\bar{x}(t) = \mathrm{E}[\mathrm{e}^{-\varphi t}x(0)] + \int_0^t \mathrm{E}[\mathrm{e}^{-\varphi(t-s)}\sigma]\,\mathrm{d}w^*(s),$$

which ends the proof of (i) with explicit functions $f(t)$ and $\bar{a}(t-s)$.

(ii) To compute correlations depending only on h, we work of course with the stationary equivalent of the processes $x_j(t) = \int_{-\infty}^t \mathrm{e}^{-\varphi_j(t-s)}\sigma_j\,\mathrm{d}w_j(s)$ and $\bar{x}(t) = \int_{-\infty}^t \mathrm{E}[\mathrm{e}^{-\varphi(t-s)}\sigma]\,\mathrm{d}w^*(s)$, so that

$$\begin{aligned}\operatorname{cov}(x_j(t+h), x_j(t)) &= \int_{-\infty}^{t} \mathrm{e}^{-\varphi_j(t+h-s)}\mathrm{e}^{-\varphi_j(t-s)}\mathrm{d}s \\ &= \mathrm{e}^{-\varphi_j h}\int_0^{+\infty} \mathrm{e}^{-2\varphi_j x}\,\mathrm{d}x = \frac{\mathrm{e}^{-\varphi_j h}}{2\varphi_j}.\end{aligned}$$

Thus, $\varrho(h) = \mathrm{e}^{-\varphi_j h}$ since $\operatorname{var}(x_j) = 1/2\varphi_j$ and $\mathrm{E}[\varrho(h)] = \mathrm{E}[\mathrm{e}^{-\varphi h}]$. For the correlations of the aggregated process, we have in the same way:

$$\bar{\varrho}(h) = \frac{\int_0^{+\infty} \mathrm{E}[\mathrm{e}^{-\varphi(x+h)}\sigma]\,\mathrm{E}[\mathrm{e}^{-\varphi x}\sigma]\,\mathrm{d}x}{\int_0^{+\infty} \mathrm{E}[\mathrm{e}^{-\varphi x}\sigma]^2\,\mathrm{d}x}.$$

Then we have the inequality $\bar{\varrho}(h) \geqslant \mathrm{E}[\varrho(h)]$ if we can prove that

$$\mathrm{E}[\mathrm{e}^{-\varphi(x+h)}\sigma] \geqslant \mathrm{E}[\mathrm{e}^{-\varphi x}\sigma]\,\mathrm{E}[\mathrm{e}^{-\varphi h}\sigma],$$

i.e., $\operatorname{cov}(\mathrm{e}^{-\varphi h}\sigma, \mathrm{e}^{-\varphi x}\sigma) \geqslant 0$ for any positive x, h, σ, and φ. This result is a consequence of the lemma:

Lemma 2. Let X be a random variable. For any f, g monotonic functions, $\operatorname{cov}(f(X), g(X)) \geqslant 0$,

which is in Gouriéroux and Monfort (1990, p. 544, ex. 11.1) and can be used with $f_x(X) = \mathrm{e}^{-Xx}$ and $g_h(X) = \mathrm{e}^{-hX}$.

(iii) If $s_j = \mathrm{e}^{-\varphi_j}\sigma_j$ is independent of φ_j, then:

$$\bar{x}(t) = \mathrm{E}[\mathrm{e}^{-\varphi(t-s)}x(0)] + \int_0^t \mathrm{E}[\mathrm{e}^{-\varphi(t-s-1)}]\,\mathrm{E}[s]\,\mathrm{d}w^*(s).$$

Moreover we know that if φ is Gamma (d), then $\mathrm{E}[\mathrm{e}^{-\varphi x}] = (1/(x+1))^d$ and $\mathrm{E}[\mathrm{e}^{-\varphi(t-s-1)}] = (t-s)^{-d}$. $\bar{x}$ is thus a fractional process of order $-d$.

Appendix

Robinson (1992) proves the following theorem:

Theorem 1. Under Assumptions 1–6,

$$\begin{bmatrix} \dfrac{m^{1/2}}{\log(n)}(\tilde{c}^J - c^J) \\ 2m^{1/2}(\tilde{d}^J - d) \end{bmatrix} \xrightarrow{d}_{N\to+\infty} \mathcal{N}\left(0, J\begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix} \otimes \Omega^{(J)}\right),$$

and the covariance matrix in the limiting distribution is consistently estimated by

$$J\begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix} \otimes \tilde{\Omega}^{(J)},$$

where Assumptions 1–6 have been already given, and the notations are as follows:

- *J is a positive integer and $Y^{(J)}_{j,k} = \log(\sum_{l=1}^{J} I_j(\lambda_{k+l-J})), j = 1,\dots,n, k = l + J, l + 2J,\dots,m$, where l and m are user-chosen positive integers satisfying Assumption 6 and I_j is the periodogram of X_{jt}, $t = 1,\dots,N$: $I_j(\lambda) = (1/2\pi N)|\sum_{t=1}^{N} X_{jt}e^{it\lambda}|^2, j = 1,\dots,n$.*
- *The unobservable random variables $U_{j,k}$ are then defined by $Y^{(J)}_{j,k} = c^J_j - d_j(2\log(\lambda_k)) + U^{(J)}_{j,k}$, $j = 1,\dots,n$, $k = l + J, l + 2J,\dots,m$, $c^J_j = \log(C - j) + \psi(J)$,* where ψ is the digamma function, $\psi(z) = (\mathrm{d}/\mathrm{d}z)\log(\Gamma(z))$ *and $U^{(J)}_k = (U^{(J)}_{1,k},\dots,U^{(J)}_{n,k})'$.*

Now

$$\begin{cases} c^J = (c^J_1,\dots,c^J_n)' \\ d = (d_1,\dots,d_n)' \end{cases}.$$

The O.L.S. estimators $\tilde{c}^J$ and $\tilde{d}^J$ of c^J and d are given by

$$\begin{bmatrix} \tilde{c}^J \\ \tilde{d}^J \end{bmatrix} = \mathrm{vec}(Y^{(J)}, Z^{(J)}, (Z^{(J)\prime} Z^{(J)})^{-1}),$$

where

$$\begin{cases} Z^{(J)} = (Z_{l+J}, Z_{l+2J},\dots,Z_m)' \\ Y^{(J)} = (Y^{(J)}_1,\dots,Y^{(J)}_n) \end{cases},$$

$$Z_k = (1, -2\log(\lambda_k))', \qquad Y^{(J)}_j = (Y^{(J)}_{j,l+J},\dots,Y^{(J)}_{j,m})'.$$

The O.L.S. residuals are denoted by $\tilde{U}_k^{(J)} = Y_k^{(J)} - \tilde{c}^J - \tilde{d}^J\ (2\log(\lambda_k))$ for $k = l + J, l + 2J, \ldots, m$, and the matrix of sample variances and covariances is

$$\tilde{\Omega}^{(J)} = \frac{1}{m-l} \sum_{k=l+J,\ldots,m} \tilde{U}_k^{(J)} \tilde{U}_k^{(J)\prime}.$$

References

Abramowitz, M. and I.A. Stegun, 1972, Handbook of mathematical functions (Dover Publications, New York), NY).

Backus, D.K. and S.E. Zin, 1995, Long-memory inflation uncertainty: Evidence from the term structure of interest rates, Journal of Money, Credit and Banking 25, 681–700.

Baillie, R.T., T. Bollerslev, and H.O.A. Mikkelsen, 1993, Fractionally integrated generalized autoregressive conditional heteroskedasticity, Journal of Econometrics, forthcoming.

Bergstrom, A.R., 1990, Continuous time econometric modelling, in: C.W.J. Granger and G. Mizon, eds., Advanced text in econometrics (Oxford University Press, Oxford).

Cheung, Y.W. and K.S. Lai, 1993, A fractional cointegration analysis, Journal of Business and Economic Statistics 11, 103–112.

Comte, F., 1993, Simulation and estimation of long memory continuous time models, Journal of Time Series Analysis, forthcoming.

Comte, F. and E. Renault, 1992, Noncausality in continuous time VARMA models, Econometric Theory, forthcoming.

Cox, J., J. Ingersoll, and S. Ross, 1981, A re-examination of traditional hypotheses about the term structure of interest rates, Journal of Finance 36, 769–799.

Dahlhaus, R., 1989, Efficient parameter estimation for self-similar processes, Annals of Statistics 17, 1749–1766.

Ding, Z., C.W.J. Granger, and R.F. Engle, 1993, A long memory property of stock market returns and a new model, Journal of Empirical Finance 1, 83–108.

Fox, R. and M.S. Taqqu, 1986, Large sample properties of parameter estimates for strongly dependent stationary time series, Annals of Statistics 14, 517–532.

Geweke, J. and S. Porter-Hudak, 1983, The estimation and application of long memory time series models, Journal of Time Series Analysis 4, 221–238.

Goncalves, E. and C. Gourieroux, 1988, Agrégation de processus autorégressifs d'ordre 1, Annales d'Economie et de Statistisques 12, 127–149.

Gourieroux, C. and A. Monfort, 1990, Séries temporelles et modèles dynamiques (Economica, Paris).

Granger, C.W.J., 1980, Long memory relationships and the aggregation of dynamic models, Journal of Econometrics, 14, 227–238.

Granger, C.W.J. and R. Joyeux, 1980, An introduction to long memory time series models and fractional differencing, Journal of Time Series Analysis 1, 15–29.

Harrison, J. and D. Kreps, 1979, Martingale and arbitrage in multiperiods securities markets, Journal of Financial Economic Theory 20, 380–408.

Haubrich, J. and A. Lo, 1991, The sources and nature of long-term memory in the business cycle, Unpublished manuscript (Federal Reserve Bank of Cleveland, Cleveland, OH).

Hosking, J.M.R., 1981, Fractional differencing, Biometrica 68, 165–176.

Kunsch,H., 1987, Discrimination between monotonic trends and long range dependence, Journal of Applied Probabilities 23, 1025–1030.

Lawrance, A.J. and N.T. Kottegoda, 1977, Stochastic modelling of Riverflow times series, Journal of the Royal Statistical Society A 140, Part 1, 1–47.

Lo, A.W., 1991, Long term memory in stock market prices, Econometrica 59, 1279–1313.

Mandelbrot, B.B. and Van Ness, 1968, Fractional Brownian motions, fractional noises and applications, SIAM Review 10, 422–437.

Mandelbrot, B.B., 1971, When can price be arbitraged efficiently? A limit to the validity of the random walk and martingale models, Review of Economics and Statistics 53, 225–236.

Merton, R.C., 1990, Continuous-time finance (Blackwell, Oxford).

Protter, P., 1990, Stochastic integration and differential equations (Springer-Verlag, New York, NY).

Robinson, P.M., 1992, Log-periodogram regression of time series with long range dependence, Working paper, London School of Economics Invited Session of the European Congress of the Econometric Society, Brussels, August (London School of Economics, London).

Rozanov, Yu A., 1968, Stationary ramdon processes (Holden-Day, San Francisco, CA).

Schwartz, L., 1966, Théorie des distributions (Hermann, Paris).

SIMS, C.A., 1984, Martingale-like behavior of asset prices and interest rates, Discussion paper 205 (Department of Economics, University of Minnesota, Minneapolis, MN).

Sowell, F., 1992, Modeling long-run behavior with the fractional ARMA model, Journal of Monetary Economics 29, 277–302.

Vasicek, O., 1977, An equilibrium characterization of the term structure, Journal of Financial Economics 5, 177–188.