The Functional Central Limit Theorem and Weak Convergence to Stochastic Integrals II: Fractionally Integrated Processes[1]

James Davidson
Cardiff University

Robert M. de Jong
Michigan State University

Final Revision, May 1999

Abstract

This paper derives a functional central limit theorem for the partial sums of fractionally integrated processes, otherwise known as $I(d)$ processes for $|d| < 1/2$. Such processes have long memory and the limit distribution is the so-called fractional Brownian motion, having correlated increments even asymptotically. The underlying shock variables may themselves exhibit quite general weak dependence, by being near-epoch dependent functions of mixing processes. Several weak convergence results for stochastic integrals having fractional integrands and weakly dependent integrators are also obtained. Taken together, these results permit $I(p+d)$ integrands for any integer $p \geq 1$.

1. Introduction

In De Jong and Davidson (2000), we obtain functional limit results for a broad class of serially dependent and heterogeneously distributed vector processes, which, however, are weakly dependent, otherwise characterised as 'short memory'. The defining feature of such processes is that their normalised partial sums converge to processes having independent Gaussian increments, specifically, Brownian motion in the case where the variances are uniformly bounded away from infinity and zero.

The present paper extends these results by allowing the processes to exhibit long memory. Specifically, we consider fractionally integrated processes, otherwise known as $I(d)$ processes. The chief interest in this class of processes is that they define a continuum linking the stationary short-memory case ($d = 0$) with the integrated or unit root case ($d = 1$). The limit processes for the partial sums of these variables for $-1/2 < d < 1/2$ are the so-called fractional Brownian motions, which differ from ordinary Brownian motion in having correlated

[1] We thank Bruce Hansen, Peter Phillips and two anonymous referees for their comments on earlier versions of this paper. Any errors are ours alone.

increments. For similar results, see among other references Davydov (1970), Taqqu (1975), Chan and Terrin (1995), Csörgő and Mielniczuk (1995), Chung (1997), Robinson and Marinucci (1998). The main novel feature of the present results is that the shock variables (fractional differences) are permitted to be near-epoch dependent on a mixing process, a very general form of weak dependence allowing various forms of nonlinear dynamics (see Davidson 2000). We also prove stochastic integral convergence for fractionally integrated processes with respect to weakly dependent integrator processes. These results are useful in, for example, the analysis of cointegrated regressions where the observed variables are $I(1+d)$ processes, whereas the residuals are short memory processes. They are straightforwardly extended to cases where the variables are $I(p+d)$ for any integer $p \geq 1$.

The paper is organised as follows. Section 2 reviews the main properties of the fractional model. Section 3 derives some technical results and gives the FCLT for a real-valued fractional process, and then extends the result to the multivariate case. Section 4 gives the results on stochastic integral convergence with fractional integrands. Different approaches are needed for the cases of negative and positive values of d. Section 5 concludes the paper, and the proofs are gathered in Appendices A-C. Extensive use will be made of the results in De Jong and Davidson (2000), and for convenience we will refer to that paper below as WD (for 'weak dependence') and will also refer to theorems and equations of the paper by attaching the prefix WD to the reference.

2. Fractionally Integrated Processes

The class of processes we consider are customarily written in the form

$$x_t = (1-L)^{-d} u_t, \tag{2.1}$$

where L is the lag operator, $-1/2 < d < 1/2$, and u_t, the fractional difference of x_t, is a stationary, weakly dependent process to be specified. By the obvious binomial expansion, they have the MA representation

$$x_t = \sum_{j=0}^{\infty} b_j u_{t-j} \tag{2.2}$$

where

$$b_j = \frac{\Gamma(j+d)}{\Gamma(d)\Gamma(j+1)}. \tag{2.3}$$

Granger and Joyeux (1980) and Hosking (1981) are standard references on these processes. Stirling's approximation formula for the gamma function yields the well-known property that the MA coefficients decline hyperbolically to zero, and letting $x_j \sim y_j$ denote that $x_j/y_j \to 1$ as $j \to \infty$, we can write

$$b_j \sim \frac{1}{\Gamma(d)} j^{d-1}. \tag{2.4}$$

The b_j are therefore square-summable for $d < 1/2$, which is the condition necessary for the processes to be stationary with finite variance, whereas $d > -1/2$ is necessary for the process to have an invertible MA representation—see Hosking (1981) for details. The partial sums, defined as $\sum_{t=1}^{s} x_t$ for $s > 0$, represent a useful class of models for nonstationary series, containing the popular unit root case corresponding to $d = 0$ in this setup. They are of particular interest when $-1/2 < d < 0$ since, although nonstationary with infinite variance, they have the property of eventual independence of initial conditions.[2] It is best to think of the corresponding x_t series as being generated as the simple differences of these nonstationary series, and so to exhibit a generalised form of over-differencing.

Defining $\sigma_n^2 = E(\sum_{t=1}^n x_t)^2$, consider the scaled partial sum process

$$X_n(\xi) = \sigma_n^{-1} \sum_{t=1}^{[n\xi]} x_t, \tag{2.5}$$

where x_t is defined by (2.1) with $|d| < 1/2$. We show that under appropriate conditions, $X_n \xrightarrow{d} X$, where X is a fractional Brownian motion, defined for $d \in (-1/2, 1/2)$ by

$$\begin{aligned} X(\xi) &= \frac{1}{\Gamma(d+1)V_d^{1/2}} \left(\int_0^\xi (\xi - s)^d dB(s) \right. \\ &\qquad \left. + \int_{-\infty}^0 \left[(\xi - s)^d - (-s)^d \right] dB(s) \right) \end{aligned} \tag{2.6}$$

for $0 \le \xi \le 1$. Here, B is standard Brownian motion and

$$V_d = \frac{1}{\Gamma(d+1)^2} \left(\frac{1}{2d+1} + \int_0^\infty ((1+\tau)^d - \tau^d)^2 d\tau \right), \tag{2.7}$$

this scale constant being chosen to make $EX(1)^2 = 1$. See Mandelbrot and van Ness (1968) for additional details.[3] Note that $X = B$ when $d = 0$. For the other cases, these processes have correlated increments, positively correlated when $d > 0$ and negatively correlated otherwise. Thus, it is easily verified from the definition in (2.6) that

$$E(X(\xi + \delta) - X(\xi))^2 = \delta^{2d+1} \tag{2.8}$$

for $\xi \in [0, 1)$ and $0 < \delta < 1 - \xi$, and hence that, for example,

$$E(X(\xi)(X(\xi + \delta) - X(\xi))) = \frac{1}{2} \left((\xi + \delta)^{2d+1} - \xi^{2d+1} - \delta^{2d+1} \right). \tag{2.9}$$

There are a number of extensions implicit in our results, which we mention here, although the details must be left to future work. Stationarity of the fractional differences ensures that the limit process is (2.6), and we will maintain this assumption in the sequel, but it is not a prerequisite for weak convergence as such. In WD we present results under weak dependence which allow global nonstationarity, with trends in the variances of the process. The limit processes in such cases are not Brownian motion, but transformations of Brownian motion involving some squeezing and stretching of the time domain. Combining global nonstationarity and strong dependence could yield an enlarged class of a.s. continuous, Gaussian limit processes, extending (2.6).

An important feature of the results is that for the case $0 < d < 1/2$, they generalise in principle beyond the fractional model. The general class of long memory MA processes, in which the coefficients merely satisfy

$$b_j \sim L(j) j^{d-1} \tag{2.10}$$

where $L(j)$ denotes any slowly varying component,[4] have essentially the same asymptotic prop-

---

[2] For a recent application of this type of model see Byers, Davidson and Peel (1997)

[3] Our formula differs from Mandelbrot and van Ness (1968) equation (2.1) since they do not normalise the variance of $X(1)$ to unity. We also implicitly impose the condition $X(0) = 0$ a.s.

[4] A sequence is said to be slowly varying at ∞ if it satisfies $L(xj)/L(j) \to 1$ as $j \to \infty$, for any $x > 0$.

erties as the fractional process. The modifications to the proofs to include a slowly varying component are straightforward in principle. Moreover, given a nonstationary process of this general MA form with $1/2 < d < 1$, its simple difference has the same asymptotic properties as the fractional model with $-1/2 < d < 0$, including, of course, that the partial sums converge to fractional Brownian motion. However, be equally careful to note that the negative fractional model behaves differently from the general case of (2.10) with $-1/2 < d < 0$, since it has the distinctive 'over-differencing' property that the MA coefficients sum to zero. The latter property is not a feature by the general class of MA processes with summable coefficients, and if the partial sums of these processes converge, it is to ordinary Brownian motion. Limit results of the latter sort may be obtained as corollaries of the theorems of this paper, but we likewise do not pursue these extensions here.

3. An FCLT for Fractionally Integrated Processes

Following the approach of Davydov (1970), we note that after substituting $x_t = \sum_{j=0}^{\infty} b_j u_{t-j}$ in (2.5) and summing the terms in a different order, we are able to write

$$X_n(\xi) - X_n(\xi') = \sigma_n^{-1} \sum_{t=-\infty}^{[n\xi]} a_{nt}(\xi, \xi') u_t \tag{3.1}$$

for $\xi > \xi'$, where

$$a_{nt}(\xi, \xi') = \sum_{j=\max\{0,[n\xi']-t+1\}}^{[n\xi]-t} b_j. \tag{3.2}$$

When b_j has the form in (2.4), this decomposition has the following properties.

Lemma 3.1

(a) *If* $\xi' < x \le \xi$ *then*

$$a_{n,[nx]}(\xi, \xi') \sim \frac{1}{\Gamma(d+1)}([n\xi] - [nx])^d,$$

and if $-\infty < x \le \xi'$ *then*

$$a_{n,[nx]}(\xi, \xi') \sim \frac{1}{\Gamma(d+1)}([n\xi] - [nx])^d - ([n\xi'] - [nx])^d;$$

(b) *hence,*

$$\sum_{t=-\infty}^{[n\xi]} a_{nt}(\xi, \xi')^2 \sim V_d(n(\xi - \xi'))^{2d+1}. \tag{3.3}$$

(See Appendix B for proofs for this section.)

The following assumption forms the basis of our functional limit results.

Assumption 1 *The sequence* $\{u_t, -\infty < t < \infty\}$

(a) *has zero mean,*

(b) *is uniformly* L_r*-bounded for* $r > 2$*,*

(c) *is* L_2*-NED of size* $-1/2$ *on* V_t *with* $d_t = 1$*, where* V_t *is either an* α*-mixing sequence of size*

$-r/(r-2)$, *or a* ϕ*-mixing sequence of size* $-r/(2(r-1))$,

(d) *is covariance stationary, and* $0 < \sigma_u^2 < \infty$ *where*

$$\sigma_u^2 = \lim_{n\to\infty} n^{-1} \sum_{t=1}^{n}\sum_{s=1}^{n} E(u_t u_s). \tag{3.4}$$

The condition $\sigma_u^2 < \infty$ actually follows from Assumption 1(c), according to Lemma A.2(a) of the Appendix. Also note that under Assumption 1(b) it is possible to set the NED and other magnitude indices to unity without loss of generality, and we do this without comment in the sequel. Compare Assumption WD.1, and note that the latter assumption is satisfied by the array $\sigma_u^{-1} n^{-1/2} u_t$ when u_t obeys Assumption 1.

For this model we have the following basic property.

Lemma 3.2 *When* x_t *is defined by (2.1) and Assumption 1 holds,* $\sigma_n^2 \sim \sigma_u^2 V_d n^{2d+1}$.

The following result is then immediate on repeating the argument of Lemma 3.2 with $[n\xi]+1$ and $[n(\xi+\delta)]$ substituted for 1 and n as the limits of the sum which defines σ_n^2, and modifying the application of Lemma 3.1 as appropriate:

Corollary 3.1 *If* X_n *is defined by (2.5) and (2.1), and Assumption 1 holds,*

$$E(X_n(\xi+\delta) - X_n(\xi))^2 \to \delta^{2d+1} \tag{3.5}$$

as $n \to \infty$, *for* $\xi \in [0,1)$ *and* $0 < \delta < 1 - \xi$.

In view of this last result, which shows that the limit distribution of X_n has the same covariance structure as fractional Brownian motion, it remains to show that the limit is Gaussian and has a.s. continuous sample paths.

Theorem 3.1 *If* X_n *is defined by (2.5) and (2.1) where* $|d| < 1/2$, *and Assumption 1 holds, then* $X_n \xrightarrow{d} X$, *where* X *is fractional Brownian motion.*

Note that Davydov (1970) has given this result for the case where u_t is i.i.d. and $0 < d < 1/2$.

It is of interest to note that our proof of Theorem 3.1 requires us to show that condition (WD.3.3) holds for the relevant array of constants – see (B-36). This is indeed the case for $d > -1/2$. Wooldridge and White's (1988) FCLT, for example, cannot be adapted to the present problem, since condition (WD.3.6) does not hold for the case $d < 0$. The approach to the FCLT developed in WD turns out to be indispensable for the result. However, note how the condition fails when $d = -1/2$. The limiting finite dimensional distributions are still Gaussian in this case since Assumption WD.1 still holds, but every increment has unit variance in the limit, and hence the limiting sample paths are not continuous. There are difficulties with the multivariate extension of the FCLT for limit processes which are not a.s. continuous, since the generalisation of the Cramér-Wold theorem is not guaranteed to hold without this restriction (see Davidson 1994 Theorem 29.16). A modified weak convergence result does evidently hold for this case, but we will not consider that problem here.

We next consider the extension to a vector of $I(d)$ processes where d may differ between elements, and may be zero for some elements. Define

$$x_t = \Delta(L)^{-1} u_t \quad (m \times 1) \tag{3.6}$$

where u_t is a $m \times 1$ vector of weakly dependent processes, and $\Delta(L)$ is a diagonal $m \times m$ lag polynomial matrix having diagonal elements $(1-L)^{d_i}$, $|d_i| < 1/2$, for $i = 1, ..., m$. Define

$$\hat{D}_n = \text{diag}(n^{d_1+1/2}, ..., n^{d_m+1/2}), \tag{3.7}$$

and $X_n(\xi) = \sum_{t=1}^{[n\xi]} X_{nt}$ where $X_{nt} = \hat{D}_n^{-1} x_t$ $(m \times 1)$. Summing the terms in a different order, as before, we obtain as the generalisation of (3.1),

$$X_n(\xi) - X_n(\xi') = \sum_{t=-\infty}^{[n\xi]} \hat{D}_n^{-1} \hat{A}_{nt}(\xi, \xi') u_t, \tag{3.8}$$

where

$$\hat{A}_{nt}(\xi, \xi') = \text{diag}(a_{1nt}(\xi, \xi'), ..., a_{mnt}(\xi, \xi')) \tag{3.9}$$

and $a_{jnt}(\xi, \xi')$ represents the element of the form (3.2) for the case $d = d_j$, for $j = 1, ..., m$. We introduce the following generalisation of Assumption 1.

Assumption 2 *Each element of the m-vector-valued sequence $\{u_t, -\infty < t < \infty\}$ satisfies Assumption 1, and*

$$\Omega_u = \lim_{n \to \infty} n^{-1} \sum_{t=1}^{n} \sum_{s=1}^{n} E(u_t u_s') \tag{3.10}$$

is finite and positive definite.

We then have the following result, by a straightforward extension of the arguments of Lemmas 3.1 and 3.2.

Lemma 3.3 *When Assumption 2 holds,*

$$E(X_n(\xi+\delta) - X_n(\xi))(X_n(\xi+\delta) - X_n(\xi))' \to \hat{K}(\delta)\Psi\hat{K}(\delta), \tag{3.11}$$

for $\xi \in [0,1)$ and $0 < \delta < 1 - \xi$, where $\hat{K}(\delta) = \text{diag}(\delta^{d_1+1/2}, ..., \delta^{d_m+1/2})$, and, letting ω_{ij} denote the (i,j)th element of Ω_u, the elements of the matrix Ψ ($m \times m$) are defined by

$$\psi_{ij} = \frac{\omega_{ij}}{\Gamma(d_i+1)\Gamma(d_j+1)} \times$$

$$\left(\frac{1}{d_i + d_j + 1} + \int_0^\infty \left((1+\tau)^{d_i} - \tau^{d_i}\right)\left((1+\tau)^{d_j} - \tau^{d_j}\right) d\tau \right), \tag{3.12}$$

for $i, j = 1, ..., m$.

With this result, we can give a multivariate extension of Theorem 3.1 as follows.

Theorem 3.2 *If u_t satisfies Assumption 2, then $X_n \xrightarrow{d} X$, where X is a vector whose ith element is a fractional Brownian motion with parameter d_i, and $EX(1)X(1)' = \Psi$.*

4. Stochastic Integrals

Let $X_{nt} = \hat{D}_n^{-1} x_t$, where x_t $(p \times 1)$ is defined by (3.6), and $W_{nt} = n^{-1/2} w_t$ $(q \times 1)$, and let $X_n(\xi) = \sum_{t=1}^{[n\xi]} X_{nt}$ and $W_n(\xi) = \sum_{t=1}^{[n\xi]} W_{nt}$. Let

$$G_n = \sum_{t=1}^{n-1} \sum_{s=1}^{t} X_{ns} W_{n,t+1}' \; (p \times q) \tag{4.1}$$

and $\Lambda_n^{XW} = E(G_n)$. In this section we derive sufficient conditions for

$$\left(X_n, W_n, G_n - \Lambda_n^{XW}\right) \xrightarrow{d} \left(X, W, \int_0^1 X dW'\right) \tag{4.2}$$

where W is Brownian motion and X is fractional Brownian motion with $X(0) = 0$. In WD we showed this convergence for weakly dependent processes under essentially the same best conditions as for the joint FCLT. We are able to extend the same approach to the 'negative fractional' case, as follows, with a minor additional restriction on the dependence.

Theorem 4.1 *Suppose* $-1/2 < d_i < 0$ *for* $i = 1, ..., p$. *Let Assumption 2 hold for the vector* (u_t, w_t), *and also suppose that the NED numbers for these variables,* $\nu^U(m)$ *and* $\nu^W(m)$, *are of size* $d - 1/2$, *where* $d = \min_{1\leq i\leq p} d_i$. *Then (4.2) holds.*

(See Appendix C for proofs for this section.)

However, we have not been able to adapt this method of proof to the case $d > 0$. The obstacle to success is the need to show (for $p = 1$) that $E \max_{1\leq t\leq n} \left(\sum_{s=1}^t x_s\right)^2 = O(n^{2d+1})$.[5] Given this, the rest of the proof could be applied with only minor amendments, but in default of this result we must consider a different approach. We give two theorems, establishing the convergence under slightly different conditions. The first follows Hansen (1992), in making use of one of the range of weak convergence theorems for semimartingale integrator processes given by Kurtz and Protter (1991). Some dependence restrictions additional to Assumption 2 must again be imposed, and we are not able to express these simply in terms of near-epoch dependence on a mixing process. Let E_s^t and E_t respectively represent the expectations conditional on the σ-fields $\mathcal{H}_s^t = \sigma(V_s, ..., V_t)$ and $\mathcal{H}_t$, where for convenience of notation we write $\mathcal{H}_t$ for $\mathcal{H}_{-\infty}^t$.

Theorem 4.2 *Suppose* $0 < d_i < 1/2$ *for* $i = 1, ..., p$. *In addition to Assumption 2 for the vector* (u_t, w_t), *assume that this vector is adapted to* $\mathcal{H}_t$. *Then, (4.2) holds in each of the following cases, where* $d = \max_{1\leq i\leq p} d_i$:

1. (w_t, u_t) *is strong mixing of size* $-2r/(r-2)$.
2. (w_t, u_t) *is uniform mixing of size* -2.
3. $$\left\| E_t w_{t+j} - E_{t-m}^{t+m} E_t w_{t+j} \right\|_2 \leq B(j, m), \tag{4.3}$$
   where $\sum_{j=0}^{\infty} B(j, m) = \zeta(m) < \infty$ *defines a sequence of size* $-d$.[6]

Note that no restrictions additional to Assumption 2 and the adaptation need hold for u_t, under condition 3.

The novel condition here is (4.3). The algebra of conditional expectations does not permit us to derive this condition merely from the fact that w_t is near-epoch dependent on a mixing process, but we can show that it holds in some leading cases. For example, it is clearly sufficient for $(w_t, \mathcal{H}_t)$ to be a martingale difference. Also consider the linear process case, say

$$w_t = \sum_{i=0}^{\infty} \Theta_i V_{1,t-i} \tag{4.4}$$

[5] Non-summable series arise in the application of McLeish's (1975a) maximal inequality to this case, see Lemma A.3 of the Appendix. For the case $d < 0$ the lemma establishes that the expected maximum is $O(n)$, which suffices at the cost of strengthening the NED sizes from $-1/2$ to $d - 1/2$, as in Theorem 4.1.

[6] For a vector a, $\|a\|_p$ here denotes $\sum_i \|a_i\|_p$. Also, for a matrix A, $\|A\|_p$ denotes $\sum_i \sum_j \|A_{ij}\|_p$.

where V_{1t} $(s \times 1)$ is a i.i.d. sub-vector of V_t, and the Θ_i $(q \times s)$ are matrices of coefficients. Letting $\theta_i = \max_{j,k} |(\Theta_i)_{jk}|$, suppose that $\theta_i = O(i^{-3/2-d-\varepsilon})$ for $\varepsilon > 0$. Since $w_t - E_{t-m}^{t+m} w_t = \sum_{i=m+1}^{\infty} \Theta_i V_{1,t-i}$ in this case, we obtain

$$\begin{aligned} \left\| w_t - E_{t-m}^{t+m} w_t \right\|_2 &\leq C \left(\sum_{i=m+1}^{\infty} \theta_i^2 \right)^{1/2} \\ &= O(m^{-1-d-\varepsilon}) \end{aligned} \tag{4.5}$$

for $C = qs \max_{1 \leq k \leq s} \|v_{1tk}\|_2$, so that w_t is L_2-NED of size $-1-d$ on $\{V_{1t}\}$, and therefore a L_2-mixingale of size $-1-d$, by Theorem 17.6 of Davidson (1994). Moreover, since $E_t w_{t+j} = \sum_{i=0}^{\infty} \Theta_{i+j} V_{1,t-i}$, we have

$$\begin{aligned} \left\| E_t w_{t+j} - E_{t-m}^{t+m} E_t w_{t+j} \right\|_2 &\leq C \left(\sum_{i=m+j+1}^{\infty} \theta_i^2 \right)^{1/2} \\ &= O((m+j)^{-1-d-\varepsilon}). \end{aligned} \tag{4.6}$$

Observe that

$$\sum_{j=0}^{\infty} \left(\sum_{i=m+j+1}^{\infty} \theta_i^2 \right)^{1/2} = O(m^{-d-\varepsilon}), \tag{4.7}$$

so that (4.3) holds with $\zeta(m)$ given by (4.7).

A point of interest about Theorem 4.2 is that the proof is not valid for negative d. This is apparent on comparing expressions (C-20) and (C-23) of the proof. Both this result and Theorem 4.1 are needed to cover the complete spectrum of cases. There is no apparent obstacle to showing joint convergence for a vector containing both positive and negative fractionals, provided the process (u_t, w_t) satisfies the conditions of both results. However, we forego the details here.

A possibly undesirable feature of Theorem 4.2 is the adaptation requirement, which is imposed under the theorem of Kurtz and Protter (1991) on which it is based. This implies that the processes u_t and w_t must not depend on future values of the underlying process V_t. While this is quite a normal feature of dynamic econometric models, so that the assumption need not be restrictive in practice, it should be noted that it is not essential. In other words, while the limiting process W must be a martingale with respect to a filtration to which X must also be adapted, this need not imply that X_{nt} and W_{nt} are $\mathcal{H}_t$-measurable, for finite n.

One can avoid the latter condition by obtaining the limiting martingale through a blocking argument, as used, following Chan and Wei (1988), in Theorems WD.4.1 and 4.1. However, to combine the Chan-Wei approach with the Hansen-type argument requires an a.s. boundedness restriction on w_t, which for practical applications may be less attractive than adaptation. We therefore give this result as an alternative to Theorem 4.2.

Theorem 4.3 *Let the assumptions of Theorem 4.2 hold, except that w_t and u_t need not be adapted to $\mathcal{H}_t$, but $|w_t| < \infty$ and the sequence $\{E_t w_{t+j}, j \geq 1\}$ is summable, with probability 1. Then (4.2) holds.*

Finally, we point out that these results can be extended to cases where the integrand is $I(1+p+d)$ for all positive integers p. Noting that $n^{-1} \sum_{t=1}^{[n\xi]} \sum_{s=1}^{t} X_{ns} \xrightarrow{d} \int_0^{\xi} X(s)ds$ by the continuous mapping theorem, we have the following result, which follows by Theorem 2.4(i) of Chan and Wei (1988). Solely for ease of exposition, we give the scalar case of the result, as follows.

Theorem 4.4 *Let the assumptions of Theorem 4.1 hold in the case* $d < 0$*, and those of either Theorem 4.2 or Theorem 4.3 in the case* $d > 0$*. Let* $W_{nt} = n^{-1/2}w_t$ *and let*

$$X_{nt} = \frac{1}{n^{p+d+1/2}} \sum_{s_p=1}^{t} \cdots \sum_{s_0=1}^{s_1} \sum_{l=0}^{\infty} b_l u_{s_0-l} \tag{4.8}$$

for $p \in \mathbb{N}$*. Then, if* X_n*,* W_n *and* G_n *are defined as before,*

$$\left(X_n, W_n, G_n - \Lambda_n^{XW}\right) \xrightarrow{d} \left(X, W, \int_0^1 X dW\right) \tag{4.9}$$

where

$$X(\xi_p) = \int_0^{\xi_p} \dots \int_0^{\xi_1} Z(\xi_0) d\xi_0 \dots d\xi_{p-1} \tag{4.10}$$

is the p*-fold integral of a fractional Brownian motion* Z*, and*

$$\Lambda_n^{XW} = \frac{1}{n^{p+d+1}} \sum_{t=1}^{n-1} \sum_{s_p=1}^{t} \cdots \sum_{s_0=1}^{s_1} \sum_{l=0}^{\infty} b_l E(u_{s_0-l} w_{t+1}). \tag{4.11}$$

5. Conclusion

Applying techniques developed in De Jong and Davidson (2000), in this paper we obtain weak convergence results for the fractionally integrated class of long memory processes, in which the limit processes are fractional Brownian motion. The functional central limit theorem is proved assuming that the fractional differences of the process are near-epoch dependent on a mixing process. We also prove the weak convergence of stochastic integrals having these processes as integrands, under a range of, collectively, mild dependence restrictions. By combining our results with the continuous mapping theorem applied to partial sums, it is possible to characterise these weak limits for normalised $I(1 + d)$ processes for any $d > -1/2$.

Appendix A. Some Technical Lemmas

Lemma A.1 *If* $a \leq b < 0$*, then*

$$\sum_{k=1}^{j-1} k^a (j-k)^b = \begin{cases} O(j^{a+b+1}) & a > -1 \\ O(j^b \log j) & a = -1 \\ O(j^b) & a < -1 \end{cases}$$

Proof. For the case $a > -1$ this follows by an integral approximation using the Beta function, and for the other cases by elementary summability arguments. ∎

The following lemmas extend Lemmas WD.A.3 and WD.A.4. Lemma A.3 is stated for arbitrary coefficients b_k, and has an application to the fractional model in the case $d < 0$. Note that for $d > 0$ the sum in (A-1) diverges.

Lemma A.2 *Let $\{X_{nt}, \mathcal{G}_{nt}\}$ and $\{Y_{nt}, \mathcal{G}_{nt}\}$ be triangular L_2-mixingale arrays of size $-1/2$ with mixingale magnitude indices a_{nt}^X and a_{nt}^Y respectively. Then*
(a)

$$\sum_{t=-\infty}^{n}\sum_{s=-\infty}^{n}|E(X_{nt}Y_{ns})| \le C\left(\sum_{t=-\infty}^{n}(a_{nt}^X)^2\right)^{1/2}\left(\sum_{t=-\infty}^{n}(a_{nt}^Y)^2\right)^{1/2}$$

for $C > 0$.
(b) *If $\sum_{t=-\infty}^{n}(a_{nt}^X)^2 = O(1)$ and $\sum_{t=-\infty}^{n}(a_{nt}^Y)^2 = O(1)$, and $\gamma_n \ge 1$ is an increasing integer-valued function of n with $\gamma_n \to \infty$ as $n \to \infty$, then*

$$\lim_{n\to\infty}\sum_{t=-\infty}^{n}\sum_{s=-\infty}^{n}|E(X_{nt}Y_{ns})|I(|t-s| > \gamma_n) = 0.$$

Proof. Similarly to Lemma WD.A.3, this is by analogy with Lemma 4 of De Jong (1997). ∎

Lemma A.3 *Let $\{Y_t, \mathcal{G}_t\}$ be a stationary L_2-mixingale with mixingale numbers $\psi^Y(j)$. If $Z_t = \sum_{k=0}^{\infty} b_k Y_{t-k}$, then*

$$E\max_{1\le t\le n}\left(\sum_{s=1}^{t}Z_s\right)^2 \le n\sum_{j=1}^{\infty}(\log j)^2\left(C_1\sum_{k=0}^{j}b_k\psi^Y(j-k) + C_2\left(\sum_{k=j+1}^{\infty}b_k^2\right)^{1/2}\right)^2. \quad \text{(A-1)}$$

for $C_1, C_2 > 0$.

Proof. The argument from McLeish (1975) Thm 1.6 yields the inequality

$$\begin{aligned} E\max_{1\le t\le n}\left(\sum_{s=1}^{t}Z_s\right)^2 &\le C_3\sum_{j=1}^{\infty}(\log j)^2\sum_{t=1}^{n}EE(Z_t|\mathcal{G}_{t-j})^2. \\ &+C_4\sum_{t=1}^{n}EE(Z_t|\mathcal{G}_t)^2 \\ &+C_5\sum_{j=1}^{\infty}(\log j)^2\sum_{t=1}^{n}E(Z_t - E(Z_t|\mathcal{G}_{t+j}))^2 \end{aligned} \quad \text{(A-2)}$$

for constants $C_3, C_4, C_5 > 0$. But for $j \ge 0$,

$$\|E(Z_t|\mathcal{G}_{t-j})\|_2 \le \sum_{k=0}^{j}b_k\left\|E\left(Y_{s-k}|\mathcal{G}_{t-j}\right)\right\|_2 + \left\|\sum_{k=j+1}^{\infty}b_kY_{s-k}\right\|_2 \quad \text{(A-3)}$$

using the Minkowski and Jensen inequalities, and similarly,

$$\|Z_t - E(Z_t|\mathcal{G}_{t+j})\|_2 \le \sum_{k=0}^{\infty}b_k\left\|Y_{s-k} - E\left(Y_{s-k}|\mathcal{G}_{t+j}\right)\right\|_2. \quad \text{(A-4)}$$

To bound the second majorant term in (A-3), apply Lemma A.2(a) putting $Y_{ns} = X_{ns} = b_{t-s}Y_s$, so that the b coefficients play the role of the mixingale magnitude indices in this case. Finally, apply the mixingale definition, and note that the first majorant term in (A-2) dominates the second and third terms to obtain (A-1). ∎

Lemma A.4 *Let $\{X_t, \mathcal{G}_t\}$ and $\{Y_t, \mathcal{G}_t\}$ be stationary L_2-mixingales with mixingale numbers $\psi^X(j)$ and $\psi^Y(j)$. If $Z_t = \sum_{k=0}^{\infty} b_k Y_{t-k}$, then*

$$\left\| \sum_{t=1}^{n} \sum_{s=1}^{t} X_t Z_s \right\|_1 \leq n \left(\sum_{j=1}^{\infty} (\log j)^2 \psi^X(j)^2 \right)^{1/2}$$

$$\times \left(\sum_{j=1}^{\infty} (\log j)^2 \left(C_1 \sum_{k=0}^{j} b_k \psi^Y(j-k) + C_2 \left(\sum_{k=j+1}^{\infty} b_k^2 \right)^{1/2} \right)^2 \right)^{1/2}. \qquad \text{(A-5)}$$

Proof. This follows straightforwardly by combining the arguments of Lemmas WD.A.4 and A.3. ∎

Appendix B. Proofs for Section 3

B.1 Proof of Lemma 3.1

To prove part (a) of the lemma, we approximate sums by integrals. Consider first the case $0 < d < 1/2$. If $\xi' < x \leq \xi$, then using (2.4),

$$\begin{aligned} a_{n,[nx]}(\xi, \xi') &= \sum_{k=0}^{[n\xi]-[nx]} b_k \\ &\sim \frac{1}{\Gamma(d)} \int_0^{[n\xi]-[nx]} y^{d-1} dy \\ &\sim \frac{1}{\Gamma(d+1)} ([n\xi] - [nx])^d, \end{aligned} \qquad \text{(B-1)}$$

and if similarly if $x \leq \xi'$ then

$$\begin{aligned} a_{n,[nx]}(\xi, \xi') &= \sum_{k=[n\xi']+1-[nx]}^{[n\xi]-[nx]} b_k \\ &\sim \frac{1}{\Gamma(d)} \int_{[n\xi']}^{[n\xi]} (y - [nx])^{d-1} dy \\ &\sim \frac{1}{\Gamma(d+1)} \left(([n\xi] - [nx])^d - ([n\xi'] - [nx])^d \right). \end{aligned} \qquad \text{(B-2)}$$

Since $\sum_{j=0}^{\infty} b_j z^j = (1-z)^{-d}$, note that $\sum_{j=0}^{\infty} b_j = 0$ for $-1/2 < d < 0$. Hence, if $\xi' < x \leq \xi$ then

$$\begin{aligned} a_{n,[nx]}(\xi, \xi') &= -\sum_{k=[n\xi]+1-[nx]}^{\infty} b_k \\ &\sim \frac{-1}{\Gamma(d)} \int_{[n\xi]-[nx]}^{\infty} y^{d-1} dy \end{aligned}$$

$$\sim \frac{1}{\Gamma(d+1)}([n\xi]-[nx])^d, \tag{B-3}$$

and if $x \leq \xi'$ then similarly,

$$\begin{aligned} a_{n,[nx]}(\xi,\xi') &= \sum_{k=[n\xi']+1-[nx]}^{\infty} b_k - \sum_{k=[n\xi]+1-[nx]}^{\infty} b_k \\ &\sim \frac{1}{\Gamma(d+1)}\left(([n\xi]-[nx])^d - ([n\xi']-[nx])^d\right). \end{aligned} \tag{B-4}$$

To prove part (b), write

$$\begin{aligned} \sum_{t=-\infty}^{[n\xi]} a_{nt}(\xi,\xi')^2 &= \sum_{t=[n\xi']+1}^{[n\xi]} a_{nt}(\xi,\xi')^2 + \sum_{t=-\infty}^{[n\xi']} a_{nt}(\xi,\xi')^2 \\ &= M_{1n} + M_{2n}. \end{aligned} \tag{B-5}$$

Letting θ_{nt} denote any positive constant array, note that $\sum_{t=[n\xi']+1}^{[n\xi]} a_{nt}(\xi,\xi')^2 \sim \sum_{t=[n\xi']+1}^{[n\xi]} \theta_{nt}$ if $a_{n,[nx]}(\xi,\xi')^2 \sim \theta_{n,[nx]}$ for every x in the interval $(\xi',\xi]$. This holds in view of the fact that the latter convergence is equivalent to $g_{n,[nx]} \to 0$ as $n \to \infty$ where

$$g_{n,[nx]} = \frac{|a_{n,[nx]}(\xi,\xi')^2 - \theta_{n,[nx]}|}{\theta_{n,[nx]}}. \tag{B-6}$$

Letting $\sum_t(.)$ represent the sum over the specified indices, we have

$$\frac{|\sum_t a_{nt}(\xi,\xi')^2 - \sum_t \theta_{nt}|}{\sum_t \theta_{nt}} \leq \frac{\sum_t g_{nt}\theta_{nt}}{\sum_t \theta_{nt}} \leq \max_{[n\xi']<t\leq[n\xi]} g_{nt}. \tag{B-7}$$

We may argue similarly for the interval $(-\infty,\xi']$, and so approximate M_{1n} and M_{2n} by the sums of squares of the expressions derived in part (a). Further integrations yield

$$M_{1n} \sim \frac{1}{(2d+1)\Gamma(d+1)^2}(n(\xi-\xi'))^{2d+1}, \tag{B-8}$$

and

$$M_{2n} \sim \frac{\int_0^\infty((1+\tau)^d-\tau^d)^2 d\tau}{\Gamma(d+1)^2}(n(\xi-\xi'))^{2d+1}, \tag{B-9}$$

where in the second case we made the change of variable $\tau = ([n\xi']-t)/([n\xi]-[n\xi'])$. ∎

B.2 Proof of Lemma 3.2

Consider Lemma 3.1 in the case $\xi = 1$ and $\xi' = 0$. For brevity, we write a_{nt} to denote $a_{nt}(1,0)$. Let B_n be an increasing integer sequence with the property $B_n \to \infty$ and $B_n/n \to 0$, and define $r_n = n/B_n$. For the purposes of the argument, assume that n increases through a sequence of values such that r_n is always an integer. Write

$$\sum_{t=1}^{n} x_t = \sum_{t=-\infty}^{n} a_{nt}u_t$$

$$
\begin{aligned}
&= \sum_{i=-\infty}^{r_n}\left(\sum_{t=(i-1)B_n+1}^{iB_n} a_{nt}u_t\right) \\
&= B_n^{1/2}\sum_{i=-\infty}^{r_n} a_{n,iB_n}S_{ni} + B_n^{3/2}\sum_{i=-\infty}^{r_n} g_{ni}S^*_{ni} \\
&= A_{1n} + A_{2n} \qquad \text{(B-10)}
\end{aligned}
$$

(say), where

$$
S_{ni} = \frac{1}{B_n^{1/2}}\sum_{t=(i-1)B_n+1}^{iB_n} u_t, \qquad \text{(B-11)}
$$

$$
S^*_{ni} = \sum_{t=(i-1)B_n+1}^{iB_n} \frac{a_{nt} - a_{n,iB_n}}{B_n^{3/2}g_{ni}}u_t, \qquad \text{(B-12)}
$$

and $g_{ni} > 0$ is to be chosen. We show that

(i) $(\sum_{t=-\infty}^{n} a_{nt}^2)^{-1}E(A_{1n}^2) \to \sigma_u^2$, and

(ii) $(\sum_{t=-\infty}^{n} a_{nt}^2)^{-1}E(A_{2n}^2) \to 0$ for a suitable definition of g_{ni}.

This is sufficient to prove the lemma, in view of the Cauchy-Schwarz inequality.

Under Assumption 1 the sequences $\{E(S_{ni}^2),\ n > |i|B_n\}$ converge to σ_u^2 for each fixed integer i, such that

$$
\sup_{-\infty<i\le r_n} |E(S_{ni}^2) - \sigma_u^2| \to 0 \text{ as } n \to \infty. \qquad \text{(B-13)}
$$

Exploiting the mixingale property (Lemma WD.A.1), this follows by Lemma A.2(a) with $Y_{nt} = X_{nt} = B_n^{-1/2}u_{t+(i-1)B_n}$ since the covariances of the terms are both absolutely summable and independent of t. It follows similarly, using Lemma A.2(b), that

$$
\sup_{-\infty<i\le r_n} \sum_{m=1}^{\infty} |E(S_{ni}S_{n,i+m})| = o(1) \text{ as } n \to \infty. \qquad \text{(B-14)}
$$

Next define

$$
W_{ni} = \left(\frac{B_n a_{n,iB_n}^2}{\sum_{t=-\infty}^{n} a_{nt}^2}\right)^{1/2} \text{ if } i \le r_n,\ 0 \text{ otherwise.} \qquad \text{(B-15)}
$$

It is easily verified using the arguments of Lemma 3.1 that $\sum_{i=-\infty}^{r_n} W_{ni}^2 \to 1$. We may write

$$
\begin{aligned}
\left(\sum_{t=-\infty}^{n} a_{nt}^2\right)^{-1} A_{1n}^2 &= \sum_{i=-\infty}^{r_n} W_{ni}^2 S_{ni}^2 + 2\sum_{i=-\infty}^{r_n-1}\sum_{m=1}^{r_n-i} W_{ni}W_{n,i+m}S_{ni}S_{n,i+m} \\
&= T_{11n} + 2T_{12n} \qquad \text{(B-16)}
\end{aligned}
$$

(say). Using (B-13),

$$
\begin{aligned}
\left|T_{11n}^2 - \sigma_u^2\right| &\le \sup_{-\infty<i\le r_n} |E(S_{ni}^2) - \sigma_u^2| \sum_{i=-\infty}^{r_n} W_{ni}^2 + \sigma_u^2\left|\sum_{i=-\infty}^{r_n} W_{ni}^2 - 1\right| \\
&\to 0 \text{ as } n \to \infty. \qquad \text{(B-17)}
\end{aligned}
$$

Also, (B-14) implies that

$$\begin{aligned} |T_{12n}^2| &\le \sum_{m=1}^{\infty}\left(\sum_{i=-\infty}^{r_n}|W_{ni}W_{n,i+m}||E(S_{ni}S_{n,i+m})|\right) \\ &\to 0 \text{ as } n\to\infty, \end{aligned} \tag{B-18}$$

noting that the weights $|W_{ni}W_{n,i+m}|$ in the majorant have a sum over i of at most 1. This completes the proof of (i).

To prove (ii), note that, for $-\infty < t < s \le n$,

$$\begin{aligned} |a_{nt}-a_{ns}| &\le \left|\sum_{j=n-s+1}^{n-t} b_j + I(t\le 0)\sum_{j=\max\{0,1-s\}}^{-t} b_j\right| \\ &\le (s-t)\left(|b_{n-s+1}| + |b_{\max\{0,1-s\}}|I(t\le 0)\right). \end{aligned} \tag{B-19}$$

Therefore, define

$$g_{ni} = |b_{n-B_n i+1}| + I(i\le 0)|b_{\max\{0,1-B_n i\}}| \tag{B-20}$$

such that

$$|a_{nt} - a_{n,iB_n}| \le B_n g_{ni} \tag{B-21}$$

for $(i-1)B_n < t \le iB_n$. In view of (B-21) and previous arguments we can say that

$$\sup_{-\infty<i\le r_n} E(S_{ni}^{*2}) = O(1) \text{ as } n\to\infty, \tag{B-22}$$

and that

$$\sup_{-\infty<i\le r_n}\sum_{m=1}^{\infty} E(S_{ni}^{*}S_{n,i+m}^{*}) = o(1) \text{ as } n\to\infty. \tag{B-23}$$

Now, defining

$$W_{ni}^{*} = \left(\frac{B_n^3 g_{ni}^2}{\sum_{t=-\infty}^{n} a_{nt}^2}\right)^{1/2}, i\le r_n, \tag{B-24}$$

note from (B-20) and the properties of b_j that

$$\begin{aligned} \sum_{i=-\infty}^{r_n-1} B_n^3 g_{ni}^2 &= O\left(B_n^{2d+1}\left(\sum_{i=-\infty}^{r_n}(r_n-i)^{2d-2} + \sum_{i=-\infty}^{0}|i|^{2d-2}\right)\right) \\ &= O(B_n^{2d+1}), \end{aligned} \tag{B-25}$$

and hence

$$\sum_{i=-\infty}^{r_n-1} W_{ni}^{*2} = O(r_n^{-2d-1}). \tag{B-26}$$

In view of the previous arguments relating to A_{1n}, it is clear that

$$\begin{aligned} \left(\sum_{t=-\infty}^{n} a_{nt}^2\right)^{-1} A_{2n}^2 &= \sum_{i=-\infty}^{r_n} W_{ni}^{*2}S_{ni}^{*2} + 2\sum_{i=-\infty}^{r_n-1} W_{ni}^{*}\sum_{m=1}^{r_n-i} W_{n,i+m}^{*}S_{ni}^{*}S_{n,i+m}^{*} \\ &= o(1), \end{aligned} \tag{B-27}$$

which completes the proof. ∎

B.3 Proof of Theorem 3.1

The proof follows the lines of Theorem WD.3.1, by establishing that conditions (WD.3.4) and (WD.3.5) hold for $X_n(\xi)$, where the limit in distribution specified in (WD.3.4) is the fractional Brownian motion, rescaled such that $EX_n(1)^2 = 1$.

To determine the finite dimensional distributions, apply Theorem 2 of De Jong (1997) to the arrays $\{\sigma_n^{-1}a_{nt}(\xi,0)u_t, t = 1-N_n, ..., [n\xi], n \geq 1\}$, where N_n is a sequence of natural numbers such that

$$R_n(\xi,\xi') = \sigma_n^{-1}\sum_{t=-\infty}^{-N_n} a_{nt}(\xi,\xi')u_t \xrightarrow{L_2} 0 \text{ as } n \to \infty \tag{B-28}$$

for any choice of ξ and ξ'. Such a sequence exists, given Corollary 3.1. Consider, for each ξ, the arrays $\sigma_n^{-1}a_{nt}(\xi,0)u_t$. Under Assumption 1, these satisfy Assumption WD.1 with respect to the constant arrays $c_{nt} = \sigma_n^{-1}a_{nt}(\xi,0)$. The style of the latter assumption can be adapted to the present case by considering the blocks of terms corresponding to $t > 0$ and $t \leq 0$ separately, with $K_n = [n\xi]$ and $K_n = N_n$ respectively. The argument is easily extended to the joint distribution of $(X_n(\xi_1), ..., X_n(\xi_p))$, for any finite set of p coordinates using the Cramér-Wold Theorem (Davidson 1994, Theorem 25.5) since the random variable $\sum_{j=1}^p \tau_j X_n(\xi_j)$, with arbitrary weights $\tau_1, ..., \tau_p$, is the partial sum of the array $\{\sigma_n^{-1}\sum_{j=1}^p \tau_j a_{nt}(\xi_j,0)u_t\}$, which similarly satisfies Assumption WD.1 as required.

The second part of the proof is to establish stochastic equicontinuity, and we adapt the proof of Theorem WD.3.1 as follows. Note first that (WD.B.2) holds as before since $\nu_n^2(.)$ is an arbitrary sequence in that expression. It remains to show that Y_n in (WD.B.3) is uniformly square integrable, for an appropriate choice of ν_n.

Write

$$X_n(\xi) - X_n(\xi') = \sigma_n^{-1}\sum_{t=-N_n}^{[n\xi]} a_{nt}(\xi,\xi')u_t + R_n(\xi,\xi'), \tag{B-29}$$

and define

$$\nu_n^2(\xi,\delta) = \sigma_u^2\sigma_n^{-2}\sum_{t=-\infty}^{[n\min\{\xi+\delta,1\}]} a_{nt}(\xi+\delta,\xi)^2, \tag{B-30}$$

Now consider the sequence corresponding to (WD.B.3),

$$Y_n(\delta,\xi') = \sup_{\{\xi:\xi-\xi'<\delta\}}\left|\sigma_n^{-1}\sum_{t=[n\xi']+1}^{[n\xi]} x_t\right|\nu_n^{-1}(\xi',\delta). \tag{B-31}$$

Observe that

$$Y_n(\delta,\xi') \leq Y_{1n}(\delta,\xi') + Y_{2n}(\delta,\xi') \tag{B-32}$$

where

$$Y_{1n}(\delta,\xi') = \sup_{\{\xi:\xi-\xi'<\delta\}}\left|\sum_{t=-N_n}^{[n\xi]} a_{nt}(\xi,\xi'+\delta)u_t\right|\nu_n^{-1}(\xi',\delta) \tag{B-33}$$

and

$$Y_{2n}(\delta,\xi') = \sup_{\{\xi:\xi-\xi'<\delta\}}|R_n(\xi,\xi')|\,\nu_n^{-1}(\xi',\delta). \tag{B-34}$$

The inequality in (B-32) must hold, because it holds (by the triangle inequality) if the value of ξ defined by the sup in (B-31) also appears in (B-33) and (B-34). In that case the sums have the same terms merely added up in a different order, according to (B-29). Taking the individual sups in (B-33) and (B-34) can only increase the majorant side.

Considering $Y_{1n}(\delta,\xi')$ first, note that this sequence is uniformly square integrable by Theorem 16.14 of Davidson (1994), if u_t^2 is uniformly integrable, which holds by Assumption 1. Moreover $Y_{2n}(\delta,\xi')\to 0$ in L_2 norm, so that uniform square integrability holds trivially for this term. Therefore, subject to

$$\lim_{\delta\to 0}\sup_{\xi\in[0,1-\delta]}\limsup_{n\to\infty}\nu_n^2(\xi,\delta)=0, \tag{B-35}$$

the FCLT is proved. But by Lemma 3.1,

$$\sup_{\xi\in[0,1-\delta]}\limsup_{n\to\infty}\sigma_u^2\sigma_n^{-2}\sum_{t=-\infty}^{[n\min\{\xi+\delta,1\}]}a_{nt}(\xi+\delta,\xi)^2=\delta^{2d+1}, \tag{B-36}$$

which converges to 0 with δ for $d>-1/2$. This completes the proof. ∎

B.4 Proof of Theorem 3.2

We appeal, as in the case of Theorem WD.3.2, to Theorem 29.16 of Davidson (1994). The result follows if $\lambda' X_n$ converges in distribution to an a.s. continuous Gaussian limit, $\lambda' X$, for all m-vectors λ of unit length. These limits are not fractional Brownian motions in general, but as a consequence of Lemma 3.3, the cases $\lambda=e_j$, $j=1...,m$ (the columns of the identity matrix of order m) yield the pure fractional BMs with parameters d_j.

To establish convergence of the finite sample distributions, we can apply Theorem 2 of de Jong (1997) as previously, since by Assumption 2, Assumption WD.1 holds for the array $\lambda'\hat{D}_n^{-1}\hat{A}_{nt}(\xi,0)u_t$ with respect to constants

$$c_{nt}=\left(\lambda'\hat{D}_n^{-1}\hat{A}_{nt}(\xi,0)\Omega_u\hat{A}_{nt}(\xi,0)\hat{D}_n^{-1}\lambda\right)^{1/2}. \tag{B-37}$$

The argument now proceeds in a similar way to that of Theorem 3.1. Stochastic equicontinuity is established in the same manner as before, by considering the random variables $\lambda'(X_n(\xi)-X_n(\xi'))$ defined according to (3.8), and also defining the sequences

$$\nu_n^2(\xi,\delta)=\sum_{t=-\infty}^{[n\min(\xi+\delta),1]}\lambda'\hat{D}_n^{-1}\hat{A}_{nt}(\xi+\delta,\xi)\Omega_u\hat{A}_{nt}(\xi+\delta,\xi)\hat{D}_n^{-1}\lambda \tag{B-38}$$

by analogy with (B-30). Note that in this case, using Lemma 3.3, we have

$$\sup_{\xi\in[0,1-\delta]}\limsup_{n\to\infty}\nu_n^2(\xi,\delta)=\lambda'\hat{K}(\delta)\Psi\hat{K}(\delta)\lambda \tag{B-39}$$

which converges to 0 as $\delta\to 0$ provided $d_j>-1/2$ for each j. ∎

Appendix C. Proofs for Section 4

C.1 Proof of Theorem 4.1

The proof follows that of Theorem WD.4.1 at most points. As there, we assume without loss of generality that X and W are scalars, and can therefore write d in place of d_i. However, the proof of Theorem 30.13 of Davidson (1994) must be further modified, as follows. The number of blocks, k_n, must be chosen to be $O(n^{\varepsilon/(\varepsilon-d)})$ where $\varepsilon > 0$ is defined by the NED size assumptions, such that the NED numbers $\nu^U(m)$ and $\nu^W(m)$, are of $O(m^{d-1/2-\varepsilon})$. The convergence specified in (30.64) of Davidson (1994) is now to the pair (X, W) where X is the process specified in (2.6), and W is ordinary Brownian motion. Since X is a.s. continuous, this fact affects the subsequent argument only at the point of equation (30.78) of Davidson (1994), where the expression in the third member becomes (in our notation)

$$\sum_{j=1}^{k_n} \int_{\xi_{j-1}}^{\xi_j} (\xi - \xi_{j-1})^{2d+1} d\xi \tag{C-1}$$

which, however, converges to 0 as before.

Next, note the modifications to the remainder of the proof of Theorem WD.4.1. We have

$$A_n = \sum_{j=1}^{k_n} \sum_{t=p_j}^{n_j-1} \sum_{s=p_j}^{t} (X_{ns}W_{n,t+1} - EX_{ns}W_{n,t+1}) \tag{C-2}$$

where as before we write p_j for $n_{j-1} + 1$, and similarly,

$$B_n = \sum_{j=1}^{k_n} \sum_{t=n_{j-1}}^{n_j-1} \sum_{s=1}^{n_{j-1}} EX_{ns}W_{n,t+1}. \tag{C-3}$$

To show $B_n \to 0$, it is convenient to rewrite it in the form

$$B_n = \sum_{j=1}^{k_n} \sum_{t=n_{j-1}}^{n_j-1} \sum_{s=-\infty}^{n_{j-1}} a_{ns}(\xi_{j-1}, 0)EU_{ns}W_{n,t+1}. \tag{C-4}$$

where $U_{ns} = n^{-1/2-d}u_s$. The same argument as before now goes through using Lemma A.2(b) and applying Lemma 3.1, noting that the mixingale magnitude index for the random variable $a_{ns}(\xi_{j-1}, 0)U_{ns}$ is $n^{-1/2-d}a_{ns}(\xi_{j-1}, 0)$.

To show that $A_n \xrightarrow{p} 0$, note first that the argument as far as equation (WD.C.14) can be copied after making the substitution

$$\sum_{s=p_j}^{t} X_{ns} = \sum_{s=-\infty}^{t} a_{ns}(t/n, \xi_{j-1})U_{ns}, \tag{C-5}$$

similarly to above. At this point the problem reduces, as before, to showing that $\bar{A}_n = \sum_{j=1}^{k_n} Y_{nj} \xrightarrow{p} 0$. Defining $\bar{X}_{ns}$ similarly to (C-5), but with U_{ns} replaced by the truncated form $\bar{U}_{ns}$ according to (WD.C.9), we can write

$$Y_{nj} = \sum_{t=p_j}^{n_j-1} \sum_{s=p_j}^{t} (\bar{X}_{ns}\bar{W}_{n,t+1} - E\bar{X}_{ns}\bar{W}_{n,t+1}). \tag{C-6}$$

Instead of (WD.C.18), we have, for $m > 0$,

$$\begin{aligned} \| \quad & Y_{nj} - E_{j-m}^{j+m} Y_{nj} \|_1 \\ \leq \quad & \sum_{t=p_j}^{n_j-1} \left\| \bar{W}_{n,t+1} - E_{j-m}^{j+m} \bar{W}_{n,t+1} \right\|_2 \left\| \sum_{s=p_j}^{t} \bar{X}_{ns} \right\|_2 \end{aligned}$$

$$+\sum_{s=p_j}^{n_j-1}\left\|\bar{X}_{ns}-E_{j-m}^{j+m}\bar{X}_{ns}\right\|_2\left(\left\|\sum_{t=p_j}^{n_j-1}\bar{W}_{n,t+1}\right\|_2+\left\|\sum_{t=p_j}^{n_j-1}E_{j-m}^{j+m}\bar{W}_{n,t+1}\right\|_2\right) \tag{C-7}$$

By Lemma A.3 and the assumptions,

$$\left\|\sum_{s=p_j}^{t}\bar{X}_{ns}\right\|_2 \leq C_1 n^{-d}k_n^{-1/2} \tag{C-8}$$

for constant $C_1>0$, and $p_j\leq t\leq n_j-1$. Also note that for $m>0$, and $s-l>n_{j-m-1}$,

$$\begin{aligned}&\left\|\bar{U}_{n,s-l}-E_{j-m}^{j+m}\bar{U}_{n,s-l}\right\|_2\\ \leq\ &\left\|\bar{U}_{n,s-l}-E(\bar{U}_{n,s-l}|\sigma(V_{n,n_{j-m-1}+1},\ldots,V_{n,2(s-l)-n_{j-m-1}})\right\|_2\\ \leq\ & n^{-1/2-d}\nu^U(mn/k_n-l),\end{aligned} \tag{C-9}$$

and hence by the Minkowski and Jensen inequalities,

$$\begin{aligned}&\left\|\bar{X}_{ns}-E_{j-m}^{j+m}\bar{X}_{ns}\right\|_2\\ \leq\ &\sum_{l=0}^{mn/k_n}b_l\left\|\bar{U}_{n,s-l}-E_{j-m}^{j+m}\bar{U}_{n,s-l}\right\|_2+2\left\|\sum_{l=mn/k_n+1}^{\infty}b_l\bar{U}_{n,s-l}\right\|_2\\ \leq\ & C_2n^{-1/2-d}\sum_{l=0}^{mn/k_n}b_l\nu^U(mn/k_n-l)+C_3n^{-1/2-d}\left(\sum_{l=mn/k_n+1}^{\infty}b_l^2\right)^{1/2}\end{aligned} \tag{C-10}$$

for $C_2, C_3>0$. where Lemma A.2(a) has been used to bound the last term. The sum in the first term of the majorant is of order $(mn/k_n)^{d-1/2-\varepsilon}$ by Lemma A.1, and the sum in the second term of the majorant is of order $(mn/k_n)^{2d-1}$. After simplification, we obtain

$$\| Y_{nj}-E_{j-m}^{j+m}Y_{nj}\|_1\leq C_4k_n^{-1}n^{-d}(k_n/n)^{\varepsilon-d}m^{d-1/2}, \tag{C-11}$$

for $C_4>0$, where $n^{-d}(k_n/n)^{\varepsilon-d}=O(1)$ by the choice of k_n, and hence, $\{Y_{nj},\mathcal{F}_{nj}\}$ is a L_1-mixingale of size $-1/2$, with summable magnitude indices. The rest of the proof closely follows Theorem WD.4.1 from (WD.C.22) onwards, except that the substitution corresponding to (C-5) is made in the expressions, and Lemma A.4 is applied in place of Lemma WD.A.4. ∎

C.2 Proof of Theorem 4.2

It is sufficient to show $G_n-\Lambda_n^{XW}\xrightarrow{d}\int_0^1 XdW'$, since the joint convergence follows by the discussion following Theorem WD.4.1.

The first step in the proof is to show that $(w_t,\mathcal{H}_t)$ is a L_2-mixingale of size -1, and also that

$$\left(\sum_{j=1}^{\infty}u_tE_tw'_{t+j}-Eu_tw'_{t+j},\mathcal{H}_t\right) \tag{C-12}$$

is a L_1-mixingale whose ith row is of size $-d_i$. We show that any of Conditions 1, 2 and 3 are sufficient. Under Condition 1, the mixingale property for $(w_t,\mathcal{H}_t)$ follows by Theorem 17.6 of Davidson (1994). For (C-12), apply of Theorem 3.2 of Hansen (1992), element by element to obtain, for constants $C_1, C_2 > 0$,[7]

$$\left\| E_{t-m} \sum_{j=1}^{\infty} u_t E_t w'_{t+j} - E u_t w'_{t+j} \right\|_1 \leq C_1 m \alpha_m^{1-2/r} + C_2 \sum_{k=m}^{\infty} \alpha_k^{1-2/r} = O(m^{-1}), \tag{C-13}$$

where α_m denotes the strong-mixing numbers. Under Condition 2, similar arguments follow after substituting the uniform mixing inequality of Serfling (1968) for the strong mixing inequality.[8] Under Condition 3, the required mixingale property of $(w_t, \mathcal{H}_t)$ is implicit in (4.3). The same condition implies (by the Minkowski inequality) that $\sum_{j=1}^{\infty} E_t w_{t+j}$ is L_2-NED on $\{V_t\}$ of size $-d$. Hence, when Assumption 2 holds for u_t, the mixingale property for (C-12) follows by Theorems 17.9 and 17.6 of Davidson (1994), applied element by element.

In the second step, the argument is applied, in effect, to each pairing of elements of x_t and w_t. We therefore assume without loss of generality that x_t and w_t are scalars. Apply the identity

$$w_t = \epsilon_t + z_{t-1} - z_t \tag{C-14}$$

where

$$\epsilon_t = \sum_{k=0}^{n-t} (E_t w_{t+k} - E_{t-1} w_{t+k}) \tag{C-15}$$

and

$$z_t = \sum_{k=1}^{n-t} E_t w_{t+k}, \tag{C-16}$$

for $t = 1, ..., n-1$, and set $z_n = 0$. Noting that

$$\sum_{t=1}^{n-1} \sum_{s=1}^{t} x_s (z_t - z_{t+1}) = \sum_{t=1}^{n-1} x_t z_t, \tag{C-17}$$

we may write

$$n^{-1-d} \sum_{t=1}^{n-1} \sum_{s=1}^{t} x_s w_{t+1} = A_{1n} + A_{2n} \tag{C-18}$$

where

$$A_{1n} = n^{-1-d} \sum_{t=1}^{n-1} \sum_{s=1}^{t} x_s \epsilon_{t+1} \tag{C-19}$$

and

$$A_{2n} = n^{-1-d} \sum_{t=1}^{n-1} x_t z_t. \tag{C-20}$$

Under the assumptions,

$$A_{1n} \xrightarrow{d} \int_0^1 X dW \tag{C-21}$$

by Theorem 3.1 of Hansen (1992), noting that the application of the strong mixing inequality in Hansen's inequality (A.1) simply makes use of the mixingale property of a centred strong mixing process, and is valid in this case by assumption. Next, consider A_{2n}. Note that for

[7] In Hansen's notation we set $\beta = 2$, permissable under our stationarity assumption, and $p = r$.

[8] In this case set $p = \beta = 2$, in Hansen's notation.

constants $C_1, C_2 > 0$, and $\varepsilon > 0$,

$$\begin{aligned}\|E_{t-m}x_tz_t - Ex_tz_t\|_1 &\leq \sum_{k=0}^{m} b_k \left\| E_{t-m}\sum_{j=1}^{n-t}(u_{t-k}E_tw_{t+j} - Eu_{t-k}w_{t+j})\right\|_1 \\ &\quad +2\left\|\sum_{k=m+1}^{\infty} b_k u_{t-k}\right\|_2 \sum_{j=1}^{n-t}\|E_tw_{t+j}\|_2 \\ &\leq C_1\sum_{k=0}^{m} b_k\psi(m-k) + C_2\left(\sum_{k=m+1}^{\infty} b_k^2\right)^{1/2}\sum_{j=1}^{\infty}\zeta(j) \\ &= O(m^{-\varepsilon}) + O\left(m^{d-1/2}\right), \qquad \text{(C-22)}\end{aligned}$$

where $\zeta(j)$ and $\psi(k)$ are the mixingale numbers for $\{w_t, \mathcal{H}_t\}$ and the process in (C-12), respectively. The first of the inequalities in (C-22) uses the Minkowski, Jensen and Cauchy-Schwarz inequalities. The second inequality applies the mixingale assumptions and Lemma A.2(a), putting $Y_{ns} = X_{ns} = b_{t-s}u_s$. The first order-of-magnitude term is evaluated using Lemma A.1, and the second follows by assumption.

The sequence $\{x_tz_t, \mathcal{H}_t\}$ is therefore a stationary L_1-mixingale, and accordingly, uniformly integrable. It follows by the weak law of large numbers of Davidson (1993) (or see Davidson 1994, Theorem 19.11) that

$$n^{-1}\sum_{t=1}^{n-1}(x_tz_t - Ex_tz_t) \xrightarrow{p} 0. \qquad \text{(C-23)}$$

Noting that $E(A_{2n}) = \Lambda_n^{XW}$ completes the proof. ∎

C.3 Proof of Theorem 4.3

As before, assume x_t and w_t are scalars. The first part of the argument follows the same modification of Theorem WD.4.1 as used in Theorem 4.1. It remains to show that $|G_n - G_n^* - \Lambda_n^{UW}| \xrightarrow{p} 0$, similarly to expression (WD.4.7). Using the decomposition in (C-14), and noting that

$$\sum_{t=n_{j-1}+1}^{n_j-1}\sum_{s=n_{j-1}+1}^{t} x_s(z_t - z_{t+1}) = \sum_{t=n_{j-1}+1}^{n_j-1} x_tz_t - \sum_{t=n_j-1}^{n_j-1} x_tz_{n_j}, \qquad \text{(C-24)}$$

we can write

$$G_n - G_n^* = A_{1n} + A_{2n} - A_{3n} \qquad \text{(C-25)}$$

where

$$A_{1n} = n^{-1-d}\sum_{m=0}^{t-n_{j-1}-1} x_{t-m}\epsilon_{t+1}, \qquad \text{(C-26)}$$

$$A_{2n} = n^{-1-d}\sum_{j=1}^{k_n}\sum_{t=n_{j-1}+1}^{n_j-1} x_tz_t \qquad \text{(C-27)}$$

and

$$A_{3n} = n^{-1-d}\sum_{j=1}^{k_n}\sum_{t=n_j-1}^{n_j-1} x_tz_{n_j}. \qquad \text{(C-28)}$$

Now, by the assumption on w_t, $|\epsilon_t| < \infty$ a.s., and by Hölder's inequality,

$$E\epsilon_{t+1}^2 \left(\sum_{m=0}^{t-n_{j-1}-1} x_{t-m} \right)^2 \leq \|\epsilon_{t+1}\|_\infty^2 \, E \left(\sum_{m=0}^{t-n_{j-1}-1} x_{t-m} \right)^2 . \tag{C-29}$$

Therefore, in view of the martingale difference property of ϵ_t,

$$\begin{aligned} E(A_{1n}^2) &\leq Cn^{-1} \sum_{j=1}^{k_n} \sum_{t=n_{j-1}+1}^{n_j-1} E\left(X_n(t/n) - X_n(\xi_{j-1})\right)^2 \\ &\leq Cn^{-1} \sum_{j=1}^{k_n} n_j (\xi_j - \xi_{j-1})^{2d+1} \\ &= O\left(\max_{1\leq j\leq k_n} (\xi_j - \xi_{j-1})^{2d+1} \right) = o(1). \end{aligned} \tag{C-30}$$

for constant $C > 0$. The proof is completed by arguments similar to those used for Theorem 4.2, noting that the term $A_{2n} - E(A_{2n})$ converges in probability to 0, that $E|A_{3n}| < n^{-d}$, and that $|E(A_{2n}) - \Lambda_n^{XW}| \to 0$. ∎

References

Byers, D., J. Davidson and D. Peel (1997) Modelling political popularity: an analysis of long-range dependence in opinion poll data. *Journal of the Royal Statistical Society, Series A* 160,3 471-90.

Chan, N. H. and N. Terrin (1995) Inference for unstable long memory processes with applications to fractional unit root autoregressions. *Annals of Statistics* 23, 1662-1683.

Chan, N.H. and C.Z. Wei (1988) Limiting distributions of least-squares estimates of unstable autoregressive processes. *Annals of Statistics* 16, 367-401.

Csörgő, S. and J. Mielniczuk (1995) Distant long-range dependent sums and regression estimation. *Stochastic Processes and their Applications* 59, 143-155..

Chung, C-F., (1997) Sample variances, sample covariances and linear regression of stationary multivariate long memory processes. Working paper, National Taiwan University.

Davidson, J (1993) An L_1-convergence theorem for heterogeneous mixingale arrays with trending moments. *Statistics and Probability Letters* 16, 301-4.

Davidson, J. (1994) *Stochastic Limit Theory*. Oxford: Oxford University Press.

Davidson, J. (2000) When is a time series I(0)? Evaluating the memory properties of nonlinear dynamic models. Working Paper (revised).

Davydov, Yu. A., (1970) The invariance principle for stationary processes. *Theory of Probability and its Applications* XV, 3, 487-498.

Granger, C.W.J. and Roselyne Joyeux (1980) An introduction to long memory time series models and fractional differencing. *Journal of Time Series Analysis* 1, 1, 15-29.

Hall, P. and C.C. Heyde (1980) *Martingale limit theory and its applications*. New York: Academic Press.

Hansen, B. E. (1992) Convergence to stochastic integrals for dependent heterogeneous processes. *Econometric Theory* 8, 489-500.

Hosking, J. R. M. (1981) Fractional differencing. *Biometrika* 68,1, 165-76.

De Jong, R.M. (1997) Central limit theorems for dependent heterogeneous random variables. *Econometric Theory*, 13, 353-367.

De Jong, R.M. and J. Davidson (2000) The functional central limit theorem and weak convergence to stochastic integrals I: weakly dependent processes. (This issue).

Kurtz, T. G. and P. Protter (1991) Weak limit theorems for stochastic integrals and stochastic differential equations. *Annals of Probability* 19,3, 1035-1070.

Mandelbrot, B. B. and J. W. van Ness (1968) Fractional Brownian motions, fractional noises and applications. *SIAM Review* 10, 4, 422-437.

McLeish, D.L. (1975a) A maximal inequality and dependent strong laws. *Annals of probability* 3, 829-839.

McLeish, D.L. (1975b) Invariance principles for dependent variables. 33, 165-178.

Robinson, P. M. and D. Marinucci (1998) Weak convergence of multivariate fractional processes, STICERD Econometrics Working PaperEM/98/352, London School of Economics.

Serfling, R. J. (1968) Contributions to central limit theory for dependent variables. *Annals of Mathematical Statistics* 39, 1158-75.

Taqqu, M. S. (1975) Weak convergence to fractional Brownian motion and to the Rosenblatt process. *Z. Wahrscheinlichskeitstheorie Verw. Geb.* 31, 287-302

Wooldridge, J.M. and H. White (1988) Some invariance principles and central limit theorems for dependent heterogeneous processes. *Econometric theory* 4, 210-230.