Electronic Journal of Statistics
Vol. 9 (2015) 2035–2057
ISSN: 1935-7524
DOI: 10.1214/15-EJS1070

Stochastic integral convergence: A white noise calculus approach

Chi Tim Ng

Department of Statistics
Chonnam National University
Gwangju 500-757, Korea
e-mail: easterlyng@gmail.com

and

Ngai Hang Chan

School of Statistics
Southwestern University of Finance and Economics
Chengdu, Sichuan, PRC 61130
and
Department of Statistics
Chinese University of Hong Kong
Shatin, New Territories, Hong Kong
e-mail: nhchan@sta.cuhk.edu.hk

Abstract: By virtue of long-memory time series, it is illustrated in this paper that white noise calculus can be used to handle subtle issues of stochastic integral convergence that often arise in the asymptotic theory of time series. A main difficulty of such an issue is that the limiting stochastic integral cannot be defined path-wise in general. As a result, continuous mapping theorem cannot be directly applied to deduce the convergence of stochastic integrals $\int_0^1 H_n(s)\,dZ_n(s)$ to $\int_0^1 H(s)\,dZ(s)$ based on the convergence of (H_n, Z_n) to (H, Z) in distribution. The white noise calculus, in particular the technique of $\mathcal{S}$-transform, allows one to establish the asymptotic results directly.

MSC 2010 subject classifications: Primary 62M10; secondary 62P20.
Keywords and phrases: Convergence, fractional Dickey-Fuller statistic, $\mathcal{S}$-transform, stochastic integral, white noise calculus.

Received February 2015.

1. Introduction

Stochastic integrals are widely used in the asymptotic theories of time series problems. For example, functionals of Brownian motion are employed in [5] to derive the asymptotic distributions of the least squares estimates of the autoregressive processes in the presence of unit roots. To test the long-memoryness of a non-stationary time series, [9] extend the results of [7] and develop a fractional Dickey-Fuller test that is based on stochastic integrals involving both Brownian motions and fractional Brownian motions. Functionals of fractional Brownian motion are also considered in [26] in the unit-root problems. In [4], the fractional Dickey-Fuller statistic is modified to test the fractional cointegration of multivariate non-stationary time series. See [3, 8, 25] and [27] for other existing results of the fractional cointegration. The convergence to stochastic integral is essential to all these asymptotic theories.

Theory of stochastic integral is summarized in [17, 24], and [10]. In particular, [10] provides a general framework so that stochastic integrals can even be defined over a Lie-group. This enhances many applications in physics. However, the stochastic integral convergence is mainly used in the construction of stochastic integrals. For example, in Section 2.3 of [10], the convergence theory is developed to approximate the rough paths of integrand and integrator by smooth paths. Convergence theory is also needed to define stochastic integrals in Ito's sense and in Stratonovich's sense. Going beyond the construction problem of stochastic integrals, it is unclear if such convergence theories can be applied directly to establish asymptotic results in statistics.

The general theory of stochastic integral convergence is discussed in [20]. Subsequent works include [12] and [6], giving the theoretical foundations of the unit-root test method proposed in [9]. It is worth noting that the convergence results of [20] are established under certain conditions that require justifications. In this paper, questions are raised related to the use of continuous mapping theorem and functional central limit theorem in the above-mentioned works. To circumvent such difficulties, an alternative approach that based on the white noise calculus is considered in this paper. In particular, the technique of $\mathcal{S}$-transform is used. For the details of the white noise calculus, one may refer to [13, 19, 18, 14], and [15].

To illustrate the ideas, the fractional Dickey-Fuller test statistic in [9] is revisited in particular. The asymptotic results of the fractional Dickey-Fuller statistic can be generalized to test the fractional cointegration of bivariate time series. It is a future research direction to explore further applications of white noise calculus in statistics involving higher dimensional data. Let $X_{n1}, X_{n2}, \ldots, X_{nn}$ be the observed time series. Suppose that $X_{ni} = \sum_{j=1}^{i} \epsilon_{nj}$, for $i = 1, 2, \ldots$, where $\epsilon_{n1}, \epsilon_{n2}, \ldots, \epsilon_{nn}$ are independent and identically distributed random variables with distribution function $F(\cdot)$, zero mean, unit variance, and finite fourth moment. Let $d \in [0, 1)$. Define $\Delta^d X_{ni} = \sum_{j=0}^{i-1} \pi_j(d) X_{n,i-j}$, where $\pi_j(d)$ are the coefficients in the Taylor series of $(1-z)^d$. The following test statistics are used in [9],

$$\hat{\phi}_{ols} = \frac{\sum_{i=2}^{n} \Delta X_{ni} \Delta^d X_{n,i-1}}{\sum_{i=2}^{n} (\Delta^d X_{n,i-1})^2}, \tag{1.1}$$

$$S_T^2 = n^{-1} \sum_{i=2}^{n} (\Delta X_{ni} - \hat{\phi}_{ols} \Delta^d X_{n,i-1})^2, \tag{1.2}$$

$$t_{\hat{\phi}_{ols}} = \frac{\sum_{i=2}^{n} \Delta X_{ni} \Delta^d X_{n,i-1}}{S_T \left(\sum_{i=2}^{n} (\Delta^d X_{n,i-1})^2\right)^{1/2}}. \tag{1.3}$$

The asymptotic distributions of (1.1)–(1.3) are taken as examples in this paper to demonstrate the application of $\mathcal{S}$-transform in asymptotic theory. In the case $d \in (1/2, 1)$, the sequence $\Delta^d X_{n,i-1}$ can be approximated by a stationary time series. Therefore, the denominators in (1.1) and (1.3) can be handled using ergodic theorem. Moreover, the numerators can be approximated by a Normal distribution according to the martingale central limit theorem. The difficult part lies in dealing with the case $d \in [0, 1/2)$. The following theorem will be established.

Theorem 1.1. *(See [9]) Suppose that $d \in [0, 1/2)$. Let*

$$U_n^{(1)} = \sum_{i=2}^{n} \Delta X_{ni} \Delta^d X_{n,i-1}\,, \tag{1.4}$$

$$U_n^{(2)} = \sum_{i=2}^{n} (\Delta^d X_{n,i-1})^2\,, \tag{1.5}$$

$$U_n^{(3)} = \sum_{i=2}^{n} (\Delta X_{n,i-1})^2\,, \tag{1.6}$$

$$U^{(1)} = \frac{1}{(1-2d)^{1/2}\Gamma(1-d)} \int_0^1 W_d(t)\, dB(t)\,, \tag{1.7}$$

$$U^{(2)} = \frac{1}{(1-2d)\Gamma^2(1-d)} \int_0^1 W_d^2(t)\, dt\,, \tag{1.8}$$

$$U^{(3)} = 1\,, \tag{1.9}$$

where $B(t)$ is a standard Brownian motion and $W_d(t)$ is the Type II fractional Brownian motion,

$$W_d(t) = (1-2d)^{1/2} \int_0^t (t-s)^{-d}\, dB(s)\,, \quad \text{for } t \geq 0\,,$$

see [22] and [9]. Then

$$(n^{-(1-d)} U_n^{(1)}, n^{-2(1-d)} U_n^{(2)}, n^{-1} U_n^{(3)})$$

converges in distribution to $(U^{(1)}, U^{(2)}, U^{(3)})$.

In Section 2, the concepts in white-noise calculus useful to the stochastic integral convergence are discussed. In Section 3, white-noise calculus, in particular, the technique of $\mathcal{S}$-transform is used to furnish the proof of Theorem 1.1. In Section 4, the extension of the results in Section 3 to the fractional cointegration test is discussed. The technical lemmas are given in the Appendix.

2. Stochastic integral convergence: Theoretical background

In this section, we examine the feasibility of using (i) functional central limit theorem and (ii) white-noise calculus to establish the convergence of stochastic integrals. Let (H_n, Z_n), $n = 1, 2, \ldots$ be a sequence of cadlag stochastic processes defined over probability measure spaces $(\Omega_n, \mathcal{F}_n, \mu_n)$.

2.1. Functional central limit theorem approach

The stochastic integral convergence theory in [20] relies on the continuous mapping theorem and functional central limit theorem. By assuming that (H_n, Z_n) converges in distribution to some stochastic process (H, Z) under the Skorohod topology of the cadlag function space $D[0,\infty)$, [20] establish the convergence of $J_n = \int_0^1 H_n(s)\,dZ_n(s)$ in distribution to $J = \int_0^1 H(s)\,dZ(s)$ using continuous mapping theorem. Indeed, the convergence in distribution can be established if the following hold:

(i) There exists a functional $\mathcal{J}$ operating on *all* $(h,z) \in D^2[0,\infty)$ so that for *all* $\omega \in \Omega_n$, $J_n(\omega) = \mathcal{J}(H_n(\omega), Z_n(\omega))$ and $J(\omega) = \mathcal{J}(H(\omega), Z(\omega))$ for *all* $\omega \in \Omega$, and

(ii) the functional $\mathcal{J}$ is continuous in the Skorokhod topology at *all* points $(h,z) \in D^2[0,\infty)$.

It should be noted that in both (i) and (ii), the qualifier "*all*" is crucial. It is not guaranteed that the continuous mapping theorem holds if "*all*" is replaced by "*almost surely*" or "*with probability going to one*". Observe that the operator $(h,z) \mapsto \lim_{\Delta\to 0}\sum h(s_i)[z(t_{i+1}-t_i)]$ cannot be defined for all $h \in D[0,\infty)$ unless $z(s)$ has finite variation, see [24]. Here Δ is the mesh size. It is also a well-known fact that there exists $z(s)$ with infinite variation in $D[0,\infty)$. Therefore, there are "holes" in the functional. To overcome such a difficulty, [20] consider an operation $\mathcal{I}^\delta(h)$ that approximates h by stepwise function, where $\delta > 0$ is chosen arbitrarily small. Though results of the continuity of $\mathcal{I}^\delta$ under the Skorokhod topology for each δ has been obtained in their Lemma 6.1, such continuity is only guaranteed almost surely. The continuity of the functional $(h,z) \mapsto \int_0^1 (\mathcal{I}^\delta(h))(s)\,dz(s)$ entails Lemma 6.1, equations (1.12), and (1.13) in [20] and therefore holds only almost surely.

2.2. White noise calculus approach

To circumvent the difficulties related to the use of continuous mapping theorem and functional central limit theorem, an alternative approach that based on the white noise calculus is considered in this paper. The crucial idea is that random variables and stochastic process can be "characterized" by the so-called $\mathcal{S}$-transform that is deterministic, see [23] and [21]. That means that there is a one-one correspondence between L^2 random variables and their $\mathcal{S}$-transforms. As a result, limits and integrals can be defined indirectly through the $\mathcal{S}$-transform. If $(\Omega_n, \mathcal{F}_n, \mu_n)$ and $(\Omega, \mathcal{F}, \mu)$ are all the same and are Gaussian measures, then it is shown in [1] that the convergence of $\mathcal{S}$-transform pointwise is equivalent to the convergence in L^2. [23] also give similar results of equivalence, however, the convergence is defined in a topology that is coarser than the L^2 topology. Therefore, the results of [1] is more relevant to the applications in statistics.

Let X_{nt}, $t = 1,2,\dots,n$, $n = 1,2,3,\dots$ be an array of random variables such that $(X_{n1}, X_{n2}, \dots, X_{nn})$ is defined over probability spaces $(\Omega_n, \mathcal{F}_n, \mu_n)$.

To establish the results of L^2 convergence, new random variables $\tilde{X}_{ni}$, $i = 1, 2, \ldots, n$, $n = 1, 2, 3, \ldots$ are constructed over a common probability measure space $(\Omega, \mathcal{F}, \mu)$ such that $(\tilde{X}_{n1}, \tilde{X}_{n2}, \ldots, \tilde{X}_{nn})$ has the same distribution as $(X_{n1}, X_{n2}, \ldots, X_{nn})$. Here, we maintain that establishing L^2 convergence can be simpler than establishing the convergence in distribution in certain situations. However, L^2 convergence requires that all random variables are defined over the same probability space. Therefore, in this subsection, $\tilde{X}_{ni}$, $i = 1, 2, \ldots, n$, $n = 1, 2, 3, \ldots$ are constructed over a common probability space $(\Omega, \mathcal{F}, \mu)$ such that $(\tilde{X}_{n1}, \tilde{X}_{n2}, \ldots, \tilde{X}_{nn})$ has the same law as $(X_{n1}, X_{n2}, \ldots, X_{nn})$. This allows us to study the convergence in distribution indirectly through L^2 convergence.

Suppose that $(\Omega, \mathcal{F}, \mu)$ is constructed so that $B(t;\omega)$, $t \in R$ is a standard Brownian motion. The detailed methods of constructing $(\Omega, \mathcal{F}, \mu)$ can be found in [17] and [14]. Equip $(\Omega, \mathcal{F}, \mu)$ with the filtration $\mathcal{F}_t$ generated by $B(t)$. To construct $(\tilde{X}_{n1}, \tilde{X}_{n2}, \ldots, \tilde{X}_{nn})$, consider the following Skorohod embedding scheme. Let $0 = \tau_{n0} \le \tau_{n1} \le \tau_{n2} \le \ldots \le \tau_{nn}$ be the stopping times as prescribed on p.516–518 of [2] so that $\tilde{\epsilon}_{ni} = n^{1/2}(B_{\tau_{ni}} - B_{\tau_{n,i-1}})$ has the same distribution as ϵ_{ni}, $\mathrm{E}(\tau_{ni} - \tau_{n,i-1}) = n^{-1}$, and $\mathrm{E}(\tau_{ni} - \tau_{n,i-1})^2 < 4n^{-2}$. Define $\tilde{X}_{ni} = \sum_{j=1}^{i} \tilde{\epsilon}_{nj}$.

The $\mathcal{S}$-transform is applied to derive the asymptotic distribution of the statistics (1.1)–(1.3). The definition of $\mathcal{S}$-transform is given below.

Definition 2.1. Let $S(R)$ be the Schwarz space, i.e. the space of rapidly decreasing functions on R. The $\mathcal{S}$-transform of a random variable $U \in L^2$ is defined as the functional that maps $\eta \in S(R)$ to

$$\mathcal{S}U(\eta) = E\left\{U \exp\left[\int_R \eta(t)\, dB(t) - \frac{1}{2}\int_R \eta^2(t)\, dt\right]\right\}.$$

Theorem 2.2 of [1] suggests that $U \in L^2$ and the $\mathcal{S}$-transform can be uniquely determined by each other. Moreover, Theorem 2.3 of the same paper establishes the equivalence between

1. $\mathrm{E}(\tilde{U}_n - U)^2 \to 0$ and
2. both $\mathrm{E}\tilde{U}_n^2 \to \mathrm{E}U^2$ and $\mathcal{S}\tilde{U}_n(\eta) \to \mathcal{S}U(\eta)$ for all $\eta \in S(R)$.

The $\mathcal{S}$-transform of the stochastic integrals $J_B = \int_0^1 H(t)\, dB(t)$ and $J_L = \int_0^1 H(t)\, dt$ can be obtained via

$$\begin{aligned} \mathcal{S}J_B(\eta) &= \int_0^1 \eta(t)\mathcal{S}(H(t))(\eta)\, dt, \\ \mathcal{S}J_L(\eta) &= \int_0^1 \mathcal{S}(H(t))(\eta)\, dt. \end{aligned} \tag{2.1}$$

There are at least two merits of using the $\mathcal{S}$-transform approach. The first one is that for each $\eta \in S(R)$, both $\mathcal{S}\tilde{U}_n(\eta)$ and $\mathcal{S}U(\eta)$ are deterministic real-valued scalars. This makes it easy to establish convergence results. The second one is that joint convergence of $(\tilde{U}_n, \tilde{V}_n)$ in distribution to (U, V) can be established component by component. This is not true in general unless all variables are defined on the same probability space and the convergence is in the L^2 sense. If $\tilde{U}_n$ converges in L^2 to U and $\tilde{V}_n$ converges in L^2 to V, Markov inequality suggests that for any $\delta > 0$,

$$P((\tilde{U}_n - U)^2 + (\tilde{V}_n - V)^2 > \delta^2) \le \delta^{-2} E((\tilde{U}_n - U)^2 + (\tilde{V}_n - V)^2) \to 0\,.$$

Therefore the Euclidean distance between $(\tilde{U}_n, \tilde{V}_n)$ and (U, V) goes to zero in probability. Slutsky's lemma suggests that $(\tilde{U}_n, \tilde{V}_n) = (U, V) + (\tilde{U}_n - U, \tilde{V}_n - V)$ converges in distribution to (U, V).

3. $\mathcal{S}$-transform and fractional Dickey-Fuller statistic

In this section, the technique of $\mathcal{S}$-transform described in Section 2 is used to study the asymptotic behavior of fractional Dickey-Fuller statistic. Theorem 1.1 is a direct consequence of the following proposition.

Proposition 3.1. *Let $U^{(1)}$, $U^{(2)}$, and $U^{(3)}$ be defined in (1.7), (1.8), and (1.9) respectively and*

$$\tilde{U}_n^{(1)} = \sum_{i=2}^{n} \Delta \tilde{X}_{ni} \Delta^d \tilde{X}_{n,i-1}\,,$$
$$\tilde{U}_n^{(2)} = \sum_{i=2}^{n} (\Delta^d \tilde{X}_{n,i-1})^2\,,$$
$$\tilde{U}_n^{(3)}) = \sum_{i=2}^{n} (\Delta \tilde{X}_{n,i-1})^2\,.$$

If $d \in [0, 1/2)$, then

$$(n^{-(1-d)} \tilde{U}_n^{(1)}, n^{-2(1-d)} \tilde{U}_n^{(2)}, n^{-1} \tilde{U}_n^{(3)})$$

converges in distribution to $(U^{(1)}, U^{(2)}, U^{(3)})$.

Remark 3.1. In Proposition 3.1, $\tilde{X}$ can further be replaced by X since $(\tilde{X}_{n1}, \tilde{X}_{n2}, \ldots, \tilde{X}_{nn})$ has the same distribution as $(X_{n1}, X_{n2}, \ldots, X_{nn})$. Theorem 1.1 then follows immediately.

New proof based on $\mathcal{S}$-transform. Throughout the paper, the notation $a \wedge b$ and $a \vee b$ refer to $\min\{a, b\}$ and $\max\{a, b\}$ respectively.

From the discussion following Definition 2.1, we see that the limiting distribution of the triple $(\tilde{U}_n^{(1)}, \tilde{U}_n^{(2)}, \tilde{U}_n^{(3)})$ can be obtained component by component provided that they all converge in L^2 to random variables defined on the same probability space. Throughout the proof, if no confusion is made, π_i refers to $\pi_i(d-1)$, $i = 1, 2, \ldots, n$. The coefficients π_j can be approximated by Stirling's formula as

$$\pi_j(d-1) = \frac{\Gamma(j+1-d)}{j!\Gamma(1-d)} \approx \frac{j^{-d}}{\Gamma(1-d)}\,, \tag{3.1}$$

see [16].

Rewrite

$$\tilde{U}_n^{(1)} = \sum_{i=2}^{n} \tilde{\epsilon}_{ni} \sum_{j=0}^{i-2} \pi_j \tilde{\epsilon}_{n,i-j-1}\,, \tag{3.2}$$

$$\begin{aligned}\tilde{U}_n^{(2)} &= \sum_{i=2}^{n}\sum_{j=0}^{i-2} \pi_j^2 \tilde{\epsilon}_{n,i-j-1}^2 + 2\sum_{i=3}^{n}\sum_{j=0}^{i-3}\sum_{\ell=j+1}^{i-2} \pi_j\pi_\ell \tilde{\epsilon}_{n,i-j-1}\tilde{\epsilon}_{n,i-\ell-1} \\ &= \tilde{U}_n^{(2,1)} + 2\tilde{U}_n^{(2,2)}\,. \end{aligned} \tag{3.3}$$

The quantities $\tilde{V}_n^{(1)}$ and $\tilde{V}_n^{(2)}$ defined below will be used in the approximation to $\tilde{U}_n^{(1)}(\eta)$ and $\tilde{U}_n^{(2)}(\eta)$ respectively,

$$\tilde{V}_n^{(1)} = n\sum_{i=2}^{n}\sum_{j=0}^{i-2} \pi_j \left[B\left(\frac{i}{n}\right) - B\left(\frac{i-1}{n}\right)\right] \tag{3.4}$$

$$\cdot \left[B\left(\frac{i-j-1}{n}\right) - B\left(\frac{i-j-2}{n}\right)\right], \tag{3.5}$$

$$\begin{aligned}\tilde{V}_n^{(2)} &= \sum_{i=2}^{n}\sum_{j=0}^{i-2} \pi_j^2 \left[B\left(\frac{i-j-1}{n}\right) - B\left(\frac{i-j-2}{n}\right)\right]^2 \\ &\quad + 2\sum_{i=2}^{n}\sum_{j=0}^{i-3}\sum_{\ell=j+1}^{i-2} \pi_j\pi_\ell \left[B\left(\frac{i-j-1}{n}\right) - B\left(\frac{i-j-2}{n}\right)\right] \\ &\quad \cdot \left[B\left(\frac{i-\ell-1}{n}\right) - B\left(\frac{i-\ell-2}{n}\right)\right] \\ &= \tilde{V}_n^{(2,1)} + 2\tilde{V}_n^{(2,2)}\,. \end{aligned} \tag{3.6}$$

□

Proof of (1.9). Clearly, from the law of large number, $\tilde{U}_n^{(3)}$ converges in probability to $\mathrm{E}\epsilon^2$. □

Proof of (1.7). Lemma A.3 suggest that $n^{-2(1-d)}\mathrm{E}[\tilde{U}_n^{(1)}]^2 \to \mathrm{E}[U^{(1)}]^2$. Next, we show that $n^{-(1-d)}\mathcal{S}\tilde{V}_n^{(1)}(\eta) \to \mathcal{S}\tilde{U}^{(1)}(\eta)$. The $\mathcal{S}$-transform of $\tilde{V}_n^{(1)}(\eta)$ can be obtained using Lemma A.1 as follows,

$$\begin{aligned}\mathcal{S}\tilde{V}_n^{(1)}(\eta) &= n\exp\left(-\frac{1}{2}\int_R \eta^2(t)\,dt\right)\cdot \sum_{i=2}^{n}\sum_{j=0}^{i-2}\pi_j \mathrm{E}\exp\left(\int_{i/n}^{\infty}\eta(t)\,dB(t)\right) \\ &\quad \cdot \mathrm{E}\left\{\exp\left(\int_{(i-1)/n}^{i/n}\eta(t)\,dB(t)\right)\cdot \left[B\left(\frac{i}{n}\right) - B\left(\frac{i-1}{n}\right)\right]\right\} \\ &\quad \cdot \mathrm{E}\left\{\exp\left(\int_{-\infty}^{(i-1)/n}\eta(t)\,dB(t)\right)\cdot \left[B\left(\frac{i-j-1}{n}\right) - B\left(\frac{i-j-2}{n}\right)\right]\right\}. \\ &= n\sum_{i=2}^{n}\sum_{j=0}^{i-2}\pi_j \int_{(i-1)/n}^{i/n}\eta(t)\,dt \int_{(i-j-2)/n}^{(i-j-1)/n}\eta(t)\,dt\,. \end{aligned}$$

Since all rapidly-decreasing functions are bounded, the integrals on the right-hand side of the above expression are all $O(n^{-1})$. Using (3.1), standard arguments can then be used to show that as $n \to \infty$,

$$n^{-(1-d)}\mathcal{S}\tilde{V}_n^{(1)}(\eta) \quad \to \quad \frac{1}{\Gamma(1-d)}\int_0^1 \eta(t)\int_0^t (t-s)^{-d}\eta(s)\,ds\,. \tag{3.7}$$

The integral on the right-hand side exists for $d \in [0,1/2)$. To show that this limit is the same as $\mathcal{S}U^{(1)}$, consider the formula (2.1) and Lemma A.1,

$$\begin{aligned}\mathcal{S}U^{(1)}(\eta) &= \frac{1}{\Gamma(1-d)}\exp\left(-\frac{1}{2}\int_R \eta^2(t)\,dt\right)\\ &\quad \cdot \int_0^1 \eta(t)\mathrm{E}\left\{\exp\left(\int_{-\infty}^{\infty}\eta(s)\,dB(s)\right)\cdot\int_0^t (t-s)^{-d}\,dB(s)\right\}dt\\ &= \frac{1}{\Gamma(1-d)}\int_0^1 \eta(t)\int_0^t (t-s)^{-d}\eta(s)\,ds\,dt\,,\end{aligned}$$

which is the same as the limit (3.7).

Below, we show that the error $\mathcal{S}[\tilde{U}_n^{(1)} - \tilde{V}_n^{(1)}](\eta) = o(n^{1-d})$ and therefore is negligible. Define

$$M_{n,k,j} = \frac{1}{2}[B(\tau_{n,k+j}) - B(\tau_{nk})]^2 - \frac{1}{2}\sum_{\ell=1}^{j}[B(\tau_{n,k+\ell}) - B(\tau_{n,k+\ell-1})]^2\,, \tag{3.8}$$

$$\begin{aligned}N_{n,k,j} &= \frac{1}{2}\left[B\left(\frac{k+j}{n}\right) - B\left(\frac{k}{n}\right)\right]^2\\ &\quad - \frac{1}{2}\sum_{\ell=1}^{j}\left[B\left(\frac{k+\ell}{n}\right) - B\left(\frac{k+\ell-1}{n}\right)\right]^2\,, \end{aligned}\tag{3.9}$$

$$L_{n,k,j} = \mathcal{S}[M_{n,k,j} - N_{n,k,j}](\eta)\,. \tag{3.10}$$

In Lemma A.2, choose $0 < \delta < 1$. Then,

$$\begin{aligned}L_{n,k,j} &= \frac{1}{2}\mathcal{S}\bigg\{\Big[B(\tau_{n,k+j}) - B(\tau_{nk})\Big]^2 - \left[B\left(\frac{k+j}{n}\right) - B\left(\frac{k}{n}\right)\right]^2\\ &\qquad -\Big[\tau_{n,k+j} - \tau_{nk}\Big] + \frac{j}{n}\bigg\}(\eta)\\ &\quad - \frac{1}{2}\sum_{\ell=1}^{j}\mathcal{S}\bigg\{\Big[B(\tau_{n,k+\ell}) - B(\tau_{n,k+\ell-1})\Big]^2\\ &\qquad -\left[B\left(\frac{k+\ell}{n}\right) - B\left(\frac{k+\ell-1}{n}\right)\right]^2 - \Big[\tau_{n,k+\ell} - \tau_{n,k+\ell-1}\Big] + \frac{1}{n}\bigg\}(\eta)\\ &= O([j/n]^{(3-\delta)/2} \wedge j^{1/2}n^{-(2-\delta)/2}) + O(j[1/n]^{(3-\delta)/2} \wedge jn^{-(2-\delta)/2})\\ &= O([j/n]^{(3-\delta)/2} \wedge j^{1/2}n^{-(2-\delta)/2})\,.\end{aligned}\tag{3.11}$$

Employing summation by parts,

$$\begin{aligned}\tilde{U}_n^{(1)} &= n\sum_{i=2}^{n}\pi_{i-2}B(\tau_{n,i-1})\cdot[B(\tau_{ni})-B(\tau_{n,i-1})]\\ &\quad - n\sum_{i=3}^{n}\sum_{j=1}^{i-2}[B(\tau_{n,i-1})-B(\tau_{n,i-1-j})]\cdot[B(\tau_{ni})-B(\tau_{n,i-1})]\cdot[\pi_j-\pi_{j-1}]\\ &= n\pi_{n-2}M_{n,0,n} - n\sum_{i=3}^{n}M_{n,0,i-1}\cdot[\pi_{i-2}-\pi_{i-3}]\\ &\quad - n\sum_{k=1}^{n-2}M_{n,k,n-k}\cdot[\pi_{n-k-1}-\pi_{n-k-2}]\\ &\quad + n\sum_{k=1}^{n-2}\sum_{j=2}^{n-k-1}M_{n,k,j}\cdot[\pi_j-2\pi_{j-1}+\pi_{j-2}]\end{aligned}$$

Similarly,

$$\begin{aligned}\tilde{V}_n^{(1)} &= n\pi_{n-2}N_{n,0,n} - n\sum_{i=2}^{n-1}N_{n,0,i}\cdot[\pi_{i-1}-\pi_{i-2}]\\ &\quad - n\sum_{k=1}^{n-2}N_{n,k,n-k}\cdot[\pi_{n-k-1}-\pi_{n-k-2}]\\ &\quad + n\sum_{k=1}^{n-2}\sum_{j=2}^{n-k-1}N_{n,k,j}\cdot[\pi_j-2\pi_{j-1}+\pi_{j-2}]\,.\end{aligned}$$

From (3.11), $L_{n,0,n} \le O(n^{-(1-\delta)/2})$, $L_{n,0,i} \le O(i^{1/2}n^{-(2-\delta)/2})$, $L_{n,k,n-k} \le O((n-k)^{1/2}n^{-(2-\delta)/2})$, $L_{n,k,j} \le O((j/n)^{(3-\delta)/2})$ for $j < \sqrt{n}$, and $L_{n,k,j} \le O(j^{1/2}n^{-(2-\delta)/2})$ for $j \ge \sqrt{n}$. Using Stirling's formula (3.1) and the facts that $\pi_j - \pi_{j-1} = O(j^{-d-1})$ and $\pi_j - 2\pi_{j-1} + \pi_{j-2} = O(j^{-d-2})$, it can be checked that if $0 < \delta < 1$ is chosen, $n\pi_{n-2}L_{n,0,n}$, $n\sum_{i=2}^{n-1} L_{n,0,i}\cdot[\pi_{i-1}-\pi_{i-2}]$, and $n\sum_{k=1}^{n-2} L_{n,k,n-k}[\pi_{n-k-1}-\pi_{n-k-2}]$ are all $o(n^{1-d})$. In addition, if $0 < \delta < 1/2 - d$ is chosen, then

$$\begin{aligned}& n\sum_{k=1}^{n-2}\sum_{j=2}^{n-k-1}L_{n,k,j}\cdot[\pi_j-2\pi_{j-1}+\pi_{j-2}]\\ &\quad = n\left(\sum_{k=1}^{n-\sqrt{n}-1}\sum_{j=2}^{\sqrt{n}-1} + \sum_{k=1}^{n-\sqrt{n}-1}\sum_{j=\sqrt{n}}^{n-k-1} + \sum_{k=n-\sqrt{n}}^{n-2}\sum_{j=2}^{n-k-1}\right)\\ &\qquad L_{n,k,j}\cdot[\pi_j-2\pi_{j-1}+\pi_{j-2}]\\ &\quad \le O\left(n^{(1+\delta)/2}\sum_{j=2}^{\sqrt{n}}j^{-d-(1+\delta)/2}\right) + O\left(n^{(2+\delta)/2}\sum_{j=\sqrt{n}}^{\infty}j^{-d-3/2}\right)\end{aligned}$$

$$
\begin{aligned}
&+ O\left(n^{\delta/2} \sum_{j=2}^{\sqrt{n}} j^{-d-(1+\delta)/2} \right) \\
&= O(n^{-d/2+(3+\delta)/4}) + O(n^{-d/2+(3+2\delta)/2}) + O(n^{-d/2+(1-\delta)/4}) \\
&= o(n^{1-d}) \,.
\end{aligned}
$$

Therefore, the error $\mathcal{S}[\tilde{U}_n^{(1)} - \tilde{V}_n^{(1)}](\eta)$ is negligible. □

Proof of (1.8). From Lemma A.3, $n^{4(1-d)}\mathrm{E}[\tilde{U}_n^{(2)}]^2 \to \mathrm{E}[U^{(2)}]^2$. Consider the convergence of $\mathcal{S}$-transform. Using Lemma A.1,

$$
\begin{aligned}
\mathcal{S}\tilde{V}_n^{(2,1)}(\eta) &= n \cdot \exp\left(-\frac{1}{2} \int_R \eta^2(t)\, dt \right) \cdot \sum_{i=2}^{n} \sum_{j=0}^{i-2} \pi_j^2 \\
&\quad \mathrm{E}\left\{ \exp\left(\int_R \eta(t)\, dB(t) \right) \cdot \left[B\left(\frac{i-j-1}{n} \right) - B\left(\frac{i-j-2}{n} \right) \right]^2 \right\} \\
&= n \sum_{i=2}^{n} \sum_{j=0}^{i-2} \pi_j^2 \mathrm{E}\left\{ \frac{1}{n} + \left(\int_{(i-j-2)/n}^{(i-j-1)/n} \eta(t)\, dt \right)^2 \right\} .
\end{aligned}
$$

Since rapidly-decreasing functions must be bounded, the terms $|\int_{(i-j-2)/n}^{(i-j-1)/n} \eta(t) dt|$ are uniformly bounded by $O(n^{-1})$ quantities and therefore are negligible. Then, the approximation formula (3.1) yields

$$
n^{-2(1-d)} \mathcal{S}\tilde{V}_n^{(2,1)}(\eta) \ \to\ \frac{1}{\Gamma^2(1-d)} \int_0^1 \int_0^t s^{-2d}\, ds\, dt \,.
$$

Similarly,

$$
\begin{aligned}
&\mathcal{S}\tilde{V}_n^{(2,2)}(\eta) \\
&\quad = n \cdot \exp\left(-\frac{1}{2} \int_R \eta^2(t)\, dt \right) \cdot \sum_{i=2}^{n} \sum_{j=1}^{i-2} \sum_{\ell=j+1}^{i-2} \pi_j \pi_\ell \\
&\qquad \cdot \mathrm{E} \exp\left(\int_{(i-\ell-1)/n}^{\infty} \eta(t)\, dB(t) \right) \\
&\qquad \cdot \mathrm{E}\left\{ \exp\left(\int_{(i-\ell-2)/n}^{(i-\ell-1)/n} \eta(t)\, dB(t) \right) \left[B\left(\frac{i-\ell-1}{n} \right) - B\left(\frac{i-\ell-2}{n} \right) \right]^2 \right\} \\
&\qquad \cdot \mathrm{E}\left\{ \exp\left(\int_{-\infty}^{(i-\ell-2)/n} \eta(t)\, dB(t) \right) \right. \\
&\qquad \left. \cdot \left[B\left(\frac{i-j-1}{n} \right) - B\left(\frac{i-j-2}{n} \right) \right]^2 \right\} . \\
&\quad = n \sum_{i=2}^{n} \sum_{j=1}^{i-2} \sum_{\ell=j+1}^{i-2} \pi_j \pi_\ell \cdot \left(\int_{(i-j-2)/n}^{(i-j-1)/n} \eta(t)\, dt \right) \cdot \left(\int_{(i-\ell-2)/n}^{(i-\ell-1)/n} \eta(t)\, dt \right) .
\end{aligned}
$$

By virtue of the approximation formula (3.1) and symmetry of the function $s^{-d}u^{-d}\eta(t-s)\eta(t-u)$, we have

$$\begin{aligned}n^{-2(1-d)}\mathcal{S}\tilde{V}_n^{(2,2)}(\eta) &\to \frac{1}{\Gamma^2(1-d)}\int_0^1\int_0^t\int_s^t s^{-d}u^{-d}\eta(t-s)\eta(t-u)\,du\,ds\,dt\\ &= \frac{1}{2\Gamma^2(1-d)}\int_0^1\left(\int_0^t (t-s)^{-d}\eta(s)\,ds\right)^2 dt\,.\end{aligned}$$

$\mathcal{S}U^{(2)}(\eta)$ can be obtained using the formula (2.1) and Lemma A.1,

$$\begin{aligned}&(1-2d)^{-1}\mathcal{S}\left\{\int_0^1 W_d^2(t)\,dB(t)\right\}(\eta)\\ &\quad=\exp\left(-\frac{1}{2}\int_R \eta^2(t)\,dt\right)\\ &\qquad\cdot\int_0^1 \mathrm{E}\left\{\exp\left(\int_{-\infty}^{\infty}\eta(s)\,dB(s)\right)\cdot\left(\int_0^t (t-s)^{-d}\,dB(s)\right)^2\right\}dt\\ &\quad=\int_0^1\left\{\left(\int_0^t (t-s)^{-d}\eta(s)\,ds\right)^2+\int_0^t (t-s)^{-2d}ds\right\}dt\,.\end{aligned}$$

It follows that $\mathcal{S}U^{(2)}(\eta)$ is the same as $\lim_{n\to\infty}\mathcal{S}\tilde{V}_n^{(2)}(\eta)$.

Next, we show that the error

$$\mathcal{S}[\tilde{U}_n^{(2)}-\tilde{V}_n^{(2)}](\eta)=[\mathcal{S}\tilde{U}_n^{(2,1)}(\eta)-\mathcal{S}\tilde{V}_n^{(2,1)}(\eta)]+2[\mathcal{S}\tilde{U}_n^{(2,2)}(\eta)-\mathcal{S}\tilde{V}_n^{(2,2)}(\eta)]$$

is $o(n^{2(1-d)})$ and therefore is negligible. Define the quadratic variation processes

$$\begin{aligned}Q_{ni}&=\sum_{j=1}^{i}[B(\tau_{nj})-B(\tau_{n,j-1})]^2\,,\\ R_{ni}&=\sum_{j=1}^{i}[B(j/n)-B((j-1)/n)]^2\,.\end{aligned}$$

Then,

$$\tilde{U}_n^{(2,1)}=n\sum_{j=0}^{n-2}\pi_j^2 Q_{n,n-j-1}\,.$$

Applying summation by parts,

$$\begin{aligned}\tilde{U}_n^{(2,2)}&=n\sum_{i=3}^{n}\pi_{i-2}\sum_{j=0}^{i-3}\pi_j B(\tau_{n,i-j-2})\cdot[B(\tau_{n,i-j-1})-B(\tau_{n,i-j-2})]\\ &\quad-n\sum_{i=3}^{n}\sum_{j=0}^{i-4}\sum_{\ell=j+2}^{i-2}\pi_j[B(\tau_{n,i-j-2})-B(\tau_{n,i-\ell-1})]\\ &\qquad\cdot[B(\tau_{n,i-j-1})-B(\tau_{n,i-j-2})]\cdot[\pi_\ell-\pi_{\ell-1}]\end{aligned}$$

$$
\begin{aligned}
&= n\sum_{i=3}^{n}\pi_{i-2}M_{n,0,i-1} + n\sum_{i=3}^{n}\pi_{i-2}\sum_{j=1}^{i-3}M_{n,0,i-j-1}\cdot[\pi_j-\pi_{j-1}] \\
&\quad - n\sum_{i=3}^{n}\sum_{\ell=0}^{i-2}M_{n,i-\ell-1,\ell-1}\cdot[\pi_\ell-\pi_{\ell-1}] \\
&\quad - n\sum_{i=3}^{n}\sum_{\ell=2}^{i-2}\sum_{j=1}^{\ell-2}M_{n,i-\ell-1,\ell-j-1}\cdot[\pi_j-\pi_{j-1}]\cdot[\pi_\ell-\pi_{\ell-1}]\,.
\end{aligned}
$$

Similarly, using the notation defined in (3.8) to (3.10),

$$
\begin{aligned}
\tilde V_n^{(2,1)} &= n\sum_{j=0}^{n-2}\pi_j^2 R_{n-j-1} \\
\tilde V_n^{(2,2)} &= n\sum_{i=3}^{n}\pi_{i-2}N_{n,0,i-1} + n\sum_{i=3}^{n}\pi_{i-2}\sum_{j=1}^{i-3}N_{n,0,i-j-1}\cdot[\pi_j-\pi_{j-1}] \\
&\quad - n\sum_{i=3}^{n}\sum_{\ell=0}^{i-2}N_{n,i-\ell-1,\ell-1}\cdot[\pi_\ell-\pi_{\ell-1}] \\
&\quad - n\sum_{i=3}^{n}\sum_{\ell=2}^{i-2}\sum_{j=1}^{\ell-2}N_{n,i-\ell-1,\ell-j-1}\cdot[\pi_j-\pi_{j-1}]\cdot[\pi_\ell-\pi_{\ell-1}]\,.
\end{aligned}
$$

Using Stirling's formula (3.1) and Lemma A.2, if $0<\delta<1$ is chosen, then

$$
\begin{aligned}
\mathcal{S}[\tilde U_n^{(2,1)} - \tilde V_n^{(2,1)}](\eta) &\approx n\sum_{j=0}^{n-2} j^{-2d}\mathcal{S}[Q_{n,n-j-1}-R_{n,n-j-1}](\eta) \\
&= O(n^{(3+\delta)/2-2d}) \\
&= o(n^{2-2d})\,.
\end{aligned}
$$

From (3.11),

$$
\begin{aligned}
L_{n,0,i-1} &\le O(i^{1/2}n^{-(2-\delta)/2})\,, \\
L_{n,0,i-j-1} &\le O(i^{1/2}n^{-(2-\delta)/2})\,, \\
L_{n,i-\ell-1,\ell-1} &\le O(\ell^{1/2}n^{-(2-\delta)/2})\,, \\
L_{n,i-\ell-1,\ell-j-1} &\le O(\ell^{1/2}n^{-(2-\delta)/2})\,.
\end{aligned}
$$

In addition, $\sum_{j=1}^{\infty} j^{-d-1} = O(1)$. Consequently, if δ is chosen so that $0<\delta<1-2d$, $\mathcal{S}[\tilde U_n^{(2,2)} - \tilde V_n^{(2,2)}](\eta) = o(n^{2-2d})$. □

4. Beyond the fractional Dickey-Fuller unit root test

The application of Theorem 1.1 is not limited to the unit root test. In this section, we illustrate that Theorem 1.1 can be used to establish the asymptotic theory of residual-based test for fractional cointegration for a bivariate time series. The purpose of this section is to provide further examples of white noise calculus for a bivariate time series.

Consider the following model. Let $(Y_{n1}, Z_{n1}), (Y_{n2}, Z_{n2}), \ldots, (Y_{nn}, Z_{nn})$ be the observed time series and $\vartheta_{n1}, \vartheta_{n2}, \ldots, \vartheta_{nn}$ and $\epsilon_{n1}, \epsilon_{n2}, \ldots, \epsilon_{nn}$ are independent and identically distributed random variables with zero mean, unit variance, and finite fourth moment. Without loss of generality, assume that $\mathrm{Var}(\epsilon_{ni}) = 1$. Suppose that $Z_{ni} = \Delta^{-d_0}\vartheta_{ni}$ and $Y_{ni} = \beta Z_{ni} + X_{ni}$, and $X_{ni} = \Delta^{-d_1}\epsilon_{ni}$, where $1/2 < d_0 \le 1$ and $0 \le d_1 < d_0 - 1/2$ are unknown parameters. We are interested in testing H_0: $(d_0, d_1) = (d_0^*, d_0^*)$ against H_1: $(d_0, d_1) = (d_0^*, d_1^*)$. To construct the test statistic, β can be estimated by either (i) regressing $\Delta^{d_0^*} Y_{ni}$ against $\Delta^{d_0^*} Z_{ni}$ or (ii) regressing Y_{ni} against Z_{ni}. These two cases are considered in subsections 4.1 and 4.2 respectively.

4.1. Regression with fractional differencing

Let

$$\hat{X}_{ni} = Y_{ni} - \hat{\beta} Z_{ni}\,, \tag{4.1}$$

$$\hat{\beta} = \frac{\sum_{i=1}^n \Delta^{d_0^*} Y_{ni} \Delta^{d_0^*} Z_{ni}}{\sum_{i=1}^n (\Delta^{d_0^*} Z_{ni})^2}\,, \tag{4.2}$$

$$\hat{\phi} = \frac{\sum_{i=2}^n \Delta^{d_0^*} \hat{X}_{ni} \Delta^{d_1^*} \hat{X}_{n,i-1}}{\sum_{i=2}^n (\Delta^{d_1^*} \hat{X}_{n,i-1})^2}\,, \tag{4.3}$$

$$S_T^2 = n^{-1} \sum_{i=2}^n (\Delta^{d_0^*} \hat{X}_{ni} - \hat{\phi}_{ols} \Delta^{d_1^*} \hat{X}_{n,i-1})^2\,, \tag{4.4}$$

$$t_{\hat{\phi}} = \frac{\sum_{i=2}^n \Delta^{d_0^*} \hat{X}_{ni} \Delta^{d_1^*} \hat{X}_{n,i-1}}{S_T \left(\sum_{i=2}^n (\Delta^{d_1^*} \hat{X}_{n,i-1})^2 \right)^{1/2}}\,. \tag{4.5}$$

The special case $d_0^* = 1$ was considered in [4]. However, in view of the discussions given in Section 2.1, the asymptotic results were not completely satisfactory as they were derived using the arguments of [9]. The following theorem is now established rigorously.

Theorem 4.1. *Let $d = 1 - d_0^* + d_1^*$ and*

$$\begin{aligned}
&(U_n^{(1)}, U_n^{(2)}, U_n^{(3)}) \\
&\quad = \left(\sum_{i=2}^n \Delta^{d_0^*} \hat{X}_{ni} \Delta^{d_1^*} \hat{X}_{n,i-1}, \quad \sum_{i=2}^n (\Delta^{d_1^*} \hat{X}_{n,i-1})^2, \quad \sum_{i=2}^n (\Delta^{d_0^*} \hat{X}_{n,i-1})^2 \right).
\end{aligned}$$

If $1/2 < d_0^* \le 1$ and $0 \le d_1^* < d_0^* - 1/2$, then under the null hypothesis $(d_0, d_1) = (d_0^*, d_0^*)$,

$$(n^{-(1-d)} U_n^{(1)}, n^{-2(1-d)} U_n^{(2)}, n^{-1} U_n^{(3)})$$

converges in distribution to $(U^{(1)}, U^{(2)}, U^{(3)})$, where $U^{(1)}$, $U^{(2)}$, and $U^{(3)}$ are defined in (1.7), (1.8), and (1.9) respectively.

Proof. Suppose that the null hypothesis $(d_0, d_1) = (d_0^*, d_0^*)$ is true. Using the fact that $\hat{X}_{ni} = X_{ni} - (\hat{\beta} - \beta) Z_{ni}$,

$$\begin{aligned}
&\sum_{i=2}^{n} \Delta^{d_0^*} \hat{X}_{ni} \Delta^{d_1^*} \hat{X}_{n,i-1} \\
&\quad = \sum_{i=2}^{n} \Delta^{d_0^*} X_{ni} \Delta^{d_1^*} X_{n,i-1} - (\hat{\beta} - \beta) \sum_{i=2}^{n} \Delta^{d_0^*} Z_{ni} \Delta^{d_1^*} X_{n,i-1} \\
&\qquad - (\hat{\beta} - \beta) \sum_{i=2}^{n} \Delta^{d_1^*} X_{ni} \Delta^{d_0^*} Z_{n,i-1} + (\hat{\beta} - \beta)^2 \sum_{i=2}^{n} \Delta^{d_0^*} Z_{ni} \Delta^{d_1^*} Z_{n,i-1} .
\end{aligned}$$

The first term can be rewritten as

$$\sum_{i=2}^{n} \Delta(\Delta^{-1} \epsilon_{ni}) \Delta^{d} (\Delta^{-1} \epsilon_{n,i-1}) .$$

Since $\Delta^{-1}\epsilon_{ni}$ is an $I(0)$ process, Theorem 1.1 suggests that the asymptotic distribution (after normalization) is the same as that of $U^{(1)}$. From Lemma A.3, $\Delta^{d_0^*} \hat{Z}_{ni} \Delta^{d_1^*} \hat{X}_{n,i-1}$, $\Delta^{d_0^*} \hat{X}_{ni} \Delta^{d_1^*} \hat{Z}_{n,i-1}$, and $\Delta^{d_0^*} \hat{Z}_{ni} \Delta^{d_1^*} \hat{Z}_{n,i-1}$ are all $O_p(n^{1-d})$. To establish the convergence result of $U_n^{(1)}$, it suffices to show that $\hat{\beta} - \beta = o_p(1)$. Clearly, under the null hypothesis,

$$\hat{\beta} - \beta = \frac{\sum_{i=1}^{n} \epsilon_{ni} \vartheta_{ni}}{\sum_{i=1}^{n} \epsilon_{ni}^2} = O_p(n^{-1/2}) .$$

The results of $U_n^{(2)}$ and $U_n^{(3)}$ can be established similarly. □

Remark. It is possible to generalize this theorem to higher dimensional time series. The proofs, however, become substantially more technical and tedious, and are not directly related to the main theme, white noise calculus, of this paper. For this reason, such kind of generalizations will not be pursued in this paper. They will be dealt with in a future research project under a different context.

4.2. Regressing without Fractional Differencing

Let

$$\hat{X}_{ni} = Y_{ni} - \hat{\beta}_{ols} Z_{ni} , \tag{4.6}$$

$$\hat{\beta}_{ols} = \frac{\sum_{i=1}^{n} Y_{ni} Z_{ni}}{\sum_{i=1}^{n} Z_{ni}^2} . \tag{4.7}$$

For simplicity, suppose that $\vartheta_{n1}, \vartheta_{n2}, \ldots, \vartheta_{nn}$ and $\epsilon_{n1}, \epsilon_{n2}, \ldots, \epsilon_{nn}$ are independent $\mathrm{N}(0,1)$ random variables so that following simple embedding scheme can be used. Let $B(t)$ be a standard Brownian motion. Define $\vartheta_{ni} = n^{1/2}(B(i/n) -$

$B((i-1)/n))$ and $\epsilon_{ni} = n^{1/2}(B(1+i/n) - B(1+(i-1)/n))$, $i = 1, 2, \ldots, n$. With such random variables defined over the same probability space, $\mathcal{S}$-transform can be used to establish the following two propositions. The proofs are very similar to that of Theorem 1.1 and are omitted here.

Proposition 4.1. *Let $d = 1 - d_0^*$ and*

$$U_n^{YZ} = \sum_{i=2}^{n} Y_{ni} Z_{ni} \,, \tag{4.8}$$

$$U_n^{ZZ} = \sum_{i=2}^{n} Z_{ni}^2 \,, \tag{4.9}$$

$$U^{YZ} = \frac{1}{(1-2d)\Gamma^2(1-d)} \int_0^1 W_d^{\vartheta}(t) W_d^{\epsilon}(t)\, dt \,, \tag{4.10}$$

$$U^{ZZ} = \frac{1}{(1-2d)\Gamma^2(1-d)} \int_0^1 [W_d^{\vartheta}(t)]^2\, dt \,, \tag{4.11}$$

where $B(t)$ is a standard Brownian motion and $W_d^{\vartheta}(t)$ and $W_d^{\epsilon}(t)$ are the Type II fractional Brownian motions,

$$W_d^{\vartheta}(t) = (1-2d)^{1/2} \int_0^t (t-s)^{-d}\, dB(s) \,, \quad \text{for } 0 \le t < 1 \,,$$
$$W_d^{\epsilon}(t) = (1-2d)^{1/2} \int_1^{1+t} (1+t-s)^{-d}\, dB(s) \,, \quad \text{for } 0 \le t < 1 \,.$$

Then $(n^{-2(1-d)} U_n^{YZ}, n^{-2(1-d)} U_n^{ZZ})$ converges in distribution to (U^{YZ}, U^{ZZ}) .

Proposition 4.2. *Let $d = 1 - d_0^* + d_1^*$ and*

$$U_n^{(1,XZ)} = \sum_{i=2}^{n} \Delta^{d_1^*} X_{ni} \Delta^{d_0^*} Z_{n,i-1} \,, \tag{4.12}$$

$$U_n^{(2,XZ)} = \sum_{i=2}^{n} \Delta^{d_1^*} X_{n,i-1} \Delta^{d_1^*} Z_{n,i-1} \,, \tag{4.13}$$

$$U_n^{(3,XZ)} = \sum_{i=2}^{n} \Delta^{d_0^*} X_{n,i-1} \Delta^{d_0^*} Z_{n,i-1} \,, \tag{4.14}$$

$$U^{(1,XZ)} = \frac{1}{(1-2d)^{1/2}\Gamma(1-d)} \int_0^1 W_d^{\epsilon}(t)\, dB(t) \,, \tag{4.15}$$

$$U^{(2,XZ)} = \frac{1}{(1-2d)\Gamma^2(1-d)} \int_0^1 W_d^{\vartheta}(t) W_d^{\epsilon}(t)\, dt \,, \tag{4.16}$$

$$U^{(3,XZ)} = 0 \,, \tag{4.17}$$

where $B(t)$ is a standard Brownian motion and $W_d^{\vartheta}(t)$ and $W_d^{\epsilon}(t)$ are the Type II fractional Brownian motions,

$$W_d^{\vartheta}(t) = (1-2d)^{1/2}\int_0^t (t-s)^{-d}\,dB(s)\,, \quad \textit{for } 0 \le t < 1\,,$$
$$W_d^{\epsilon}(t) = (1-2d)^{1/2}\int_1^{1+t} (1+t-s)^{-d}\,dB(s)\,, \quad \textit{for } 0 \le t < 1\,.$$

Then $(n^{-(1-d)}U_n^{(1,XZ)}, n^{-2(1-d)}U_n^{(2,XZ)}, n^{-1}U_n^{(3,XZ)})$ *converges in distribution to* $(U^{(1,XZ)}, U^{(2,XZ)}, U^{(3,XZ)})$.

The asymptotic distributions can therefore be obtained as in the proof of Theorem 4.1 using for example,

$$\begin{aligned}
&\sum_{i=2}^{n} \Delta^{d_0^*}\hat{X}_{ni}\Delta^{d_1^*}\hat{X}_{n,i-1} \\
&\quad = \sum_{i=2}^{n} \Delta^{d_0^*}X_{ni}\Delta^{d_1^*}X_{n,i-1} - (\hat{\beta}_{ols}-\beta)\sum_{i=2}^{n} \Delta^{d_0^*}Z_{ni}\Delta^{d_1^*}X_{n,i-1} \\
&\qquad - (\hat{\beta}_{ols}-\beta)\sum_{i=2}^{n} \Delta^{d_1^*}X_{ni}\Delta^{d_0^*}Z_{n,i-1} + (\hat{\beta}_{ols}-\beta)^2\sum_{i=2}^{n} \Delta^{d_0^*}Z_{ni}\Delta^{d_1^*}Z_{n,i-1}\,.
\end{aligned}$$

Here, we see that the test statistics in subsection 4.1 allow simpler desmmcription of the asymptotic distributions.

Appendix A: Technical lemmas

Lemma A.1. *Let C be a 2×2 symmetric positive-definite matrix and $g(x)$ be a function so that the integral*

$$I(g) = \frac{1}{2\pi|C|^{1/2}}\int_{R^2} g(x)\exp\left(y - \frac{1}{2}(x,y)C^{-1}(x,y)^T\right)dx\,dy$$

exists. Then,

$$I(g) = \frac{\exp(C_{22}/2)}{\sqrt{2\pi C_{11}}}\int_R g(x+C_{12})\exp\left(-\frac{1}{2C_{11}}x^2\right)dx\,.$$

Proof. This can be shown easily using completing squares techniques. □

Lemma A.2. *For any integers $1 \le k < i \le n$, real number $0 < \delta < 2$, and rapidly-decreasing function $\eta(\cdot)$,*

$$\mathcal{S}[\tau_{ni} - \tau_{nk} - (i-k)/n](\eta) = O(n^{-1}(i-k)^{1/2})\,, \tag{A.1}$$

$$\begin{aligned}
&\mathcal{S}\left\{[B(\tau_{ni}) - B(\tau_{nk})]^2 - [B(i/n) - B(k/n)]^2 - [\tau_{ni} - \tau_{nk}] + (i-k)/n\right\}(\eta) \\
&\quad = O([(i-k)/n]^{(3-\delta)/2} \wedge (i-k)^{1/2}n^{-(2-\delta)/2})\,.
\end{aligned} \tag{A.2}$$

Proof of (A.1). It follows from Cauchy-Schwarz inequality and the fact that

$$\Phi = \exp\left(\int_R \eta(t)\,dB(t) - \frac{1}{2}\int_R \eta^2(t)\,dt\right)$$

has moments of any orders. □

Proof of (A.2). Let $\Theta(t) = \int_{-\infty}^t \eta(t)\,dt$, $\tilde{B}(t) = B(t) - \Theta(t)$,

$$\Phi(T) = \exp\left(\int_0^T \eta(t)\,dB(t) - \frac{1}{2}\int_0^T \eta^2(t)\,dt\right).$$

For any $T > 1$,

$$\begin{aligned}
&\mathrm{E}\left\{\Phi(T)\left([B(\tau_{ni}\wedge T) - B(\tau_{nk}\wedge T)]^2 - [(\tau_{ni}\wedge T) - (\tau_{nk}\wedge T)]\right)\right\}\\
&\quad= \mathrm{E}\left\{\Phi(T)\left(\left[\tilde{B}(\tau_{ni}\wedge T) - \tilde{B}(\tau_{nk}\wedge T)\right]^2 - [(\tau_{ni}\wedge T) - (\tau_{nk}\wedge T)]\right)\right\}\\
&\qquad+ 2\left[\Theta\left(\frac{i}{n}\right) - \Theta\left(\frac{k}{n}\right)\right]\cdot \mathrm{E}\left\{\Phi(T)\left[\tilde{B}(\tau_{ni}\wedge T) - \tilde{B}(\tau_{nk}\wedge T)\right]\right\}\\
&\qquad+ \mathrm{E}\left\{\Phi(T)\left[\Theta\left(\frac{i}{n}\right) - \Theta\left(\frac{k}{n}\right)\right]^2\right\}\\
&\qquad+ 2\mathrm{E}\left\{\Phi(T)\left[\tilde{B}(\tau_{ni}\wedge T) - \tilde{B}(\tau_{nk}\wedge T)\right]\right.\\
&\qquad\quad\left.\cdot\left[\Theta(\tau_{ni}\wedge T) - \Theta(\tau_{nk}\wedge T) - \Theta\left(\frac{i}{n}\right) + \Theta\left(\frac{k}{n}\right)\right]\right\}\\
&\qquad+ \mathrm{E}\left\{\Phi(T)\,[\Theta(\tau_{ni}\wedge T) - \Theta(\tau_{nk}\wedge T)]\right.\\
&\qquad\quad\left.\cdot\left[\Theta(\tau_{ni}\wedge T) - \Theta(\tau_{nk}\wedge T) - \Theta\left(\frac{i}{n}\right) + \Theta\left(\frac{k}{n}\right)\right]\right\}\\
&\qquad+ \left[\Theta\left(\frac{i}{n}\right) - \Theta\left(\frac{k}{n}\right)\right]\\
&\qquad\quad\cdot \mathrm{E}\left\{\Phi(T)\left[\Theta(\tau_{ni}\wedge T) - \Theta(\tau_{nk}\wedge T) - \Theta\left(\frac{i}{n}\right) + \Theta\left(\frac{k}{n}\right)\right]\right\}\\
&\quad= I_1(T) + I_2(T) + I_3(T) + I_4(T) + I_5(T) + I_6(T).
\end{aligned}$$

It can be checked from Lemma A.1 that

$$I_3(T) \quad=\quad \mathrm{E}\left\{\Phi(T)\left(\left[B\left(\frac{i}{n}\right) - B\left(\frac{k}{n}\right)\right]^2 - \frac{i-k}{n}\right)\right\}.$$

Since both $B(t)$ and $B^2(t) - t$ are martingales, $I_1(T) = I_2(T) = 0$ by virtue of Girsanov's theorem and Doob's optional stopping theorem. Note that for any rapidly-decreasing function $\eta(\cdot)$,

$$\sup_{t\geq 0}|\Theta(t)| \leq \int_R |\eta(s)|ds < \infty\,.$$

Moreover, both $\tilde{B}(\tau_{ni})$ and $\tilde{B}(\tau_{nk})$ have finite second moments by definition. Then, taking $T\to\infty$, bounded convergence theorem guarantees that

$$\begin{aligned}
&\mathcal{S}\left\{[B(\tau_{ni})-B(\tau_{nk})]^2 - [B(i/n)-B(k/n)]^2 - [\tau_{ni}-\tau_{nk}] + (i-k)/n\right\}(\eta) \\
&= 2\mathrm{E}\left\{\Phi\left[\tilde{B}(\tau_{ni})-\tilde{B}(\tau_{nk})\right]\cdot\left[\Theta(\tau_{ni})-\Theta(\tau_{nk})-\Theta\left(\frac{i}{n}\right)+\Theta\left(\frac{k}{n}\right)\right]\right\} \\
&\quad + \mathrm{E}\left\{\Phi\left[\Theta(\tau_{ni})-\Theta(\tau_{nk})\right]\cdot\left[\Theta(\tau_{ni})-\Theta(\tau_{nk})-\Theta\left(\frac{i}{n}\right)+\Theta\left(\frac{k}{n}\right)\right]\right\} \\
&\quad + \left[\Theta\left(\frac{i}{n}\right)-\Theta\left(\frac{k}{n}\right)\right]\cdot\mathrm{E}\left\{\Phi\left[\Theta(\tau_{ni})-\Theta(\tau_{nk})-\Theta\left(\frac{i}{n}\right)+\Theta\left(\frac{k}{n}\right)\right]\right\} \\
&= I_4+I_5+I_6\,.
\end{aligned}$$

In addition,

$$\begin{aligned}
\mathrm{E}\{\Phi[\tilde{B}(\tau_{ni})-\tilde{B}(\tau_{nk})]^2\} &= \mathrm{E}\left\{\Phi\left[\tau_{ni}-\tau_{nk}\right]\right\} \leq (\mathrm{E}\Phi^2)^{1/2}\cdot[\mathrm{E}(\tau_{ni}-\tau_{nk})^2]^{1/2} \\
&= O((i-k)/n)\,. \qquad \text{(A.3)}
\end{aligned}$$

Note that for any rapidly-decreasing function $\eta(\cdot)$,

$$\sup_{t\geq 0}|\Theta(t)| \leq \int_R |\eta(s)|ds < \infty \quad \text{and} \quad \sup_{t\geq 0}|\eta(t)| < \infty\,.$$

Then, for any real number $0<\delta<2$,

$$\begin{aligned}
&\mathrm{E}\left\{\Phi\left[\Theta(\tau_{ni})-\Theta(\tau_{nk})\right]^2\right\} \\
&\leq \left\{2\sup_{t\geq 0}|\Theta(t)|\right\}^{\delta}\mathrm{E}\left\{\Phi\left[\Theta(\tau_{ni})-\Theta(\tau_{nk})\right]^{2-\delta}\right\} \\
&\leq \left\{2\sup_{t\geq 0}|\Theta(t)|\right\}^{\delta}\left\{\sup_{t\geq 0}|\eta(t)|\right\}^{2-\delta}\mathrm{E}\left\{\Phi\left[\tau_{ni}-\tau_{nk}\right]^{2-\delta}\right\} \\
&\leq \left\{2\sup_{t\geq 0}|\Theta(t)|\right\}^{\delta}\left\{\sup_{t\geq 0}|\eta(t)|\right\}^{2-\delta}(\mathrm{E}\Phi^{2/\delta})^{\delta/2}\cdot\left\{\mathrm{E}\left[\tau_{ni}-\tau_{nk}\right]^2\right\}^{1-\delta/2} \\
&= O([(i-k)/n]^{2-\delta}) \qquad \text{(A.4)}
\end{aligned}$$

and similarly,

$$\begin{aligned}
&\mathrm{E}\left\{\Phi\left[\Theta(\tau_{ni})-\Theta\left(\frac{i}{n}\right)\right]^2\right\} \\
&= \left\{2\sup_{t\geq 0}|\Theta(t)|\right\}^{2\delta}\left\{\sup_{t\geq 0}|\eta(t)|\right\}^{2-2\delta}(\mathrm{E}\Phi^{1/\delta})^{\delta}\cdot\left\{\mathrm{E}\left[\tau_{ni}-\frac{i}{n}\right]^2\right\}^{1-\delta} \\
&= O(n^{-(1-\delta)})\,. \qquad \text{(A.5)}
\end{aligned}$$

Using the bounds (A.4) and (A.5), it can be seen that

$$\begin{aligned}
&\mathrm{E}\left\{\Phi\left[\Theta(\tau_{ni})-\Theta(\tau_{nk})-\Theta\left(\frac{i}{n}\right)+\Theta\left(\frac{k}{n}\right)\right]^2\right\}\\
&\quad\le 2\mathrm{E}\left\{\Phi\left[\Theta(\tau_{ni})-\Theta\left(\frac{i}{n}\right)\right]^2\right\}+2\mathrm{E}\left\{\Phi\left[\Theta(\tau_{nk})-\Theta\left(\frac{k}{n}\right)\right]^2\right\}\\
&\quad\le O(n^{-(1-\delta)})
\end{aligned}$$

and

$$\begin{aligned}
&\mathrm{E}\left\{\Phi\left[\Theta(\tau_{ni})-\Theta(\tau_{nk})-\Theta\left(\frac{i}{n}\right)+\Theta\left(\frac{k}{n}\right)\right]^2\right\}\\
&\quad\le 2\mathrm{E}\left\{\Phi\left[\Theta(\tau_{ni})-\Theta(\tau_{nk})\right]^2\right\}+2\mathrm{E}\left\{\Phi\left[-\Theta\left(\frac{i}{n}\right)-\Theta\left(\frac{k}{n}\right)\right]^2\right\}\\
&\quad\le O([(i-k)/n]^{2-\delta})\,.
\end{aligned}$$

Equivalently,

$$\begin{aligned}
&\mathrm{E}\left\{\Phi\left[\Theta(\tau_{ni})-\Theta(\tau_{nk})-\Theta\left(\frac{i}{n}\right)+\Theta\left(\frac{k}{n}\right)\right]^2\right\}\\
&\quad\le O([(i-k)/n]^{2-\delta}\wedge n^{-(1-\delta)})\,. \qquad \text{(A.6)}
\end{aligned}$$

The bounds (A.3) to (A.6), together with Cauchy-Schwarz inequality suggest that

$$I_4+I_5+I_6=O([(i-k)/n]^{(3-\delta)/2}\wedge(i-k)^{1/2}n^{-(2-\delta)/2})\,. \qquad \square$$

Lemma A.3. *Let $U^{(1)}$ and $U^{(2)}$ be defined in (1.7)–(1.8). Define*

$$\begin{aligned}
E_1&=\frac{1}{\Gamma^2(1-d)}\int_0^1\int_t^1(t-s)^{-2d}\,ds\,dt\,,\\
E_{21}&=\frac{2}{\Gamma^4(1-d)}\int_0^1\int_{t_1}^1\int_0^{t_1}\int_0^{t_2}(t_1-u)^{-2d}(t_2-v)^{-2d}\,dv\,du\,dt_2\,dt_1\,,\\
E_{22}&=\frac{2}{\Gamma^4(1-d)}\int_0^1\int_{t_1}^1\int_0^{t_1}\int_0^{t_1}(t_1-u)^{-d}(t_1-v)^{-d}\\
&\qquad\cdot(t_2-u)^{-d}(t_2-v)^{-d}dv\,du\,dt_2\,dt_1\,.
\end{aligned}$$

Then,

$$E[U^{(1)}]^2=E_1\,, \qquad \text{(A.7)}$$

$$E[U^{(2)}]^2=E_{21}+E_{22}\,, \qquad \text{(A.8)}$$

Let $(\vartheta_{n1},\epsilon_{n1}),(\vartheta_{n2},\epsilon_{n2}),\ldots,(\vartheta_{nn},\epsilon_{nn})$ be independent and identically distributed with zero mean and finite fourth moment. Then,

$$n^{-2(1-d)}E\left\{\sum_{i=2}^n\vartheta_{ni}\sum_{j=0}^{i-2}\pi_j\epsilon_{n,i-j-1}\right\}^2\to E\epsilon^2\cdot E\vartheta^2\cdot E_1\,, \qquad \text{(A.9)}$$

$$n^{-4(1-d)}E\left\{\sum_{i=2}^{n}\left(\sum_{j=0}^{i-2}\pi_j\vartheta_{n,i-j-1}\right)\cdot\left(\sum_{j=0}^{i-2}\pi_j\epsilon_{n,i-j-1}\right)\right\}^2$$
$$\to E\vartheta^2\cdot E\epsilon^2\cdot E_{22}+(E\vartheta\epsilon)^2\cdot E_{21}\,. \tag{A.10}$$

Proof. Identity (A.7) follows from the Itô's isometry. Identity (A.9) is shown as follows,

$$\begin{aligned}
&n^{-2(1-d)}\mathrm{E}\left\{\sum_{i=2}^{n}\vartheta_{ni}\sum_{j=0}^{i-2}\pi_j\epsilon_{n,i-j-1}\right\}^2\\
&=n^{-2(1-d)}\mathrm{E}\epsilon^2\cdot\mathrm{E}\vartheta^2\cdot\sum_{k=1}^{n-1}\sum_{i=k+1}^{n}\pi_{i-k-1}^2\\
&\to\frac{1}{\Gamma^2(1-d)}\mathrm{E}\epsilon^2\cdot\mathrm{E}\vartheta^2\cdot\int_0^1\int_t^1(t-s)^{-2d}\,ds\,dt\,.
\end{aligned}$$

Next, consider Identity (A.10). For any integers $2\le i_1\le i_2\le n$,

$$\begin{aligned}
&\mathrm{E}\left\{\sum_{j=0}^{i_1-2}\pi_{j_1}\vartheta_{n,i_1-j-1}\right\}\cdot\left\{\sum_{j=0}^{i_1-2}\pi_{j_1}\epsilon_{n,i_1-j-1}\right\}\\
&\cdot\left\{\sum_{j=0}^{i_2-2}\pi_{j}\vartheta_{n,i_2-j-1}\right\}\cdot\left\{\sum_{j=0}^{i_2-2}\pi_{j}\epsilon_{n,i_2-j-1}\right\}\\
&\quad=\mathrm{E}\left\{\sum_{a=1}^{i_1-1}\sum_{b=1}^{i_1-1}\pi_{i_1-a-1}\pi_{i_1-b-1}\vartheta_{na}\epsilon_{nb}\right\}\cdot\left\{\sum_{a=1}^{i_2-1}\sum_{b=1}^{i_2-1}\pi_{i_2-a-1}\pi_{i_2-b-1}\vartheta_{na}\epsilon_{nb}\right\}\\
&\quad=[\mathrm{E}\vartheta^2]\cdot[\mathrm{E}\epsilon^2]\cdot\sum_{a=1}^{i_1-1}\sum_{b=1}^{i_1-1}\pi_{i_1-a-1}\pi_{i_2-a-1}\pi_{i_1-b-1}\pi_{i_2-b-1}\\
&\qquad+(\mathrm{E}\vartheta\epsilon)^2\cdot\sum_{a=1}^{i_1-1}\sum_{b=1}^{i_2-1}\pi_{i_1-a-1}^2\pi_{i_2-b-1}^2\\
&\qquad+\{\mathrm{E}\vartheta^2\epsilon^2-2(\mathrm{E}\vartheta\epsilon)^2\}\cdot\sum_{a=1}^{i_1-1}\pi_{i_1-a-1}^2\pi_{i_2-a-1}^2\,.
\end{aligned}$$

Using Stirling's approximation (3.1), as $i_1\to\infty$, the last sum on the right-hand side is

$$\begin{aligned}
\sum_{a=1}^{i_1-1}\pi_{i_1-a-1}^2\pi_{i_2-a-1}^2&\approx\frac{1}{\Gamma^4(1-d)}\sum_{a=1}^{i_1-1}(i_1-a)^{-2d}(i_2-a)^{-2d}\\
&\le\frac{1}{\Gamma^4(1-d)}\sum_{a=1}^{i_1-1}(i_1-a)^{-4d}\\
&=O(i_1^{(1-4d)\vee 0})\,.
\end{aligned}$$

The second sum

$$\sum_{a=1}^{i_1-1}\sum_{b=1}^{i_2-1}\pi_{i_1-a-1}^2\pi_{i_2-b-1}^2 = O(i_1^{1-2d}\cdot i_2^{1-2d})$$

dominates the last sum if $d\in[0,1/2)$. Then, standard arguments show that

$$\begin{aligned}
&\mathrm{E}\left\{\sum_{i=2}^{n}\left(\sum_{j=0}^{i-2}\pi_j\vartheta_{n,i-j-1}\right)\cdot\left(\sum_{j=0}^{i-2}\pi_j\epsilon_{n,i-j-1}\right)\right\}^2\\
&\quad\approx 2\mathrm{E}\vartheta^2\cdot\mathrm{E}\epsilon^2\cdot\sum_{i=2}^{n}\sum_{a=1}^{i-1}\sum_{b=1}^{i-1}\pi_{i-a-1}^2\pi_{i-b-1}^2\\
&\qquad+2(\mathrm{E}\vartheta\epsilon)^2\cdot\sum_{i_1=2}^{n-1}\sum_{i_2=i_1+1}^{n}\sum_{a=1}^{i_1-1}\sum_{b=1}^{i_2-1}\pi_{i_1-a-1}^2\pi_{i_2-b-1}^2\\
&\qquad+2\mathrm{E}\vartheta^2\cdot\mathrm{E}\epsilon^2\cdot\sum_{i_1=2}^{n-1}\sum_{i_2=i_1+1}^{n}\sum_{a=1}^{i_1-1}\sum_{b=1}^{i_1-1}\pi_{i_1-a-1}\pi_{i_2-a-1}\pi_{i_1-b-1}\pi_{i_2-b-1}\\
&\quad\approx\frac{2n^{4(1-d)}}{\Gamma^4(1-d)}\mathrm{E}\vartheta^2\cdot\mathrm{E}\epsilon^2\cdot\int_0^1\int_{t_1}^1\int_0^{t_1}\int_0^{t_2}(t_1-u)^{-2d}(t_2-v)^{-2d}\,dv\,du\,dt_2\,dt_1\\
&\qquad+\frac{2n^{4(1-d)}}{\Gamma^4(1-d)}(\mathrm{E}\vartheta\epsilon)^2\\
&\qquad\cdot\int_0^1\int_{t_1}^1\int_0^{t_1}\int_0^{t_1}(t_1-u)^{-d}(t_1-v)^{-d}(t_2-u)^{-d}(t_2-v)^{-d}\,dv\,du\,dt_2\,dt_1.
\end{aligned}$$

Similar arguments show that

$$\mathrm{E}\left\{\sum_{i=2}^{n}\left\{\sum_{a=1}^{i-1}\left(\frac{i-a-1}{n}\right)^{-d}\left[B\left(\frac{a}{n}\right)-B\left(\frac{a-1}{n}\right)\right]\right\}^2\right\}^2$$

yields the same limit except for the multiplier $n^{4(1-d)}$. This gives Identity (A.8). □

Acknowledgements

We would like to thank the Editor and two anonymous referees for helpful comments. C. T. Ng's research is supported in part by the National Research Foundation of Korea (NRF) grant funded by the government (MSIP) (No. 2011-0030810) and the 2013 Chonnam National University Research Program grant (No. 2013-2299). N. H. Chan's research is supported in part by grants from HKSAR-RGC-GRF Nos. 400313 and 14300514 and HKSAR-RGC-CRF No. CityU8/CRG/12G.

References

[1] Bender, C. (2003). An S-transform approach to integration with respect to a fractional Brownian motion. *Bernoulli* **9**, 955–983. MR2046814

[2] Billingsley, P. (1995) *Probability and Measure, 3rd Edition*. Wiley, New York. MR1324786

[3] Booth, G. G. and Tse, Y. (1995). Long memory in interest rate futures markets: A fractional cointegration analysis. *Journal of Future Markets* **5**, 573–584.

[4] Chan, C. H. (2004). Residual-based test for fractional cointegration, Master Dissertation, Chinese University of Hong Kong.

[5] Chan, N. H. and Wei, C. Z. (1988). Limiting distributions of least squares estimates of unstable autoregressive processes. *The Annals of Statistics* **16**, 367–401. MR0924877

[6] Davidson, J. and De Jong, R. M. (2000) The functional central limit theorem and weak convergence to stochastic integrals II: Fractionally integrated processes, *Econometric Theory* **16**, 643–666. MR1802836

[7] Dickey, D. A. and Fuller, W. A. (1981). Likelihood ratio tests for autoregressive time series with a unit root. *Econometrica* **49**, 1057–1072. MR0625773

[8] Dittmann, I. (2000). Residual-based tests for fractional cointegration: A Monte Carlo study. *Journal of Time Series Analysis* **25**, 615–647.

[9] Dolado, J. J., Gonzalo, J., and Mayoral, L. (2002). A fractional Dickey-Fuller test for unit roots. *Econometrica* **70**, 1963–2006. MR1925162

[10] Friz, P. and Hairer, M. (2014). *A Course on Rough Paths: With an Introduction to Regularity Structures*. Springer Universitext. MR3289027

[11] Gourieroux, C., Maurel, F., and Monfort, A. (1989). Least squares and fractionally integrated regressors, Document de Travail No. 8913, INSEE.

[12] Hansen, B. E. (1992). Convergence to stochastic integrals for dependent heterogeneous processes. *Econometric Theory* **8**, 489–500. MR1202326

[13] Hida, T. (1975). *Analysis of Brownian Functionals. Carleton Mathematical Lecture Notes* **13**, Carleton, Canada MR0451429

[14] Hølden, H., Øksendal, B., Ubøe, J., and Zhang, T. (1996). *Stochastic Partial Differential Equations*. Birkhäuser, Boston. MR1408433

[15] Huang, Z. and Yan, J. (2000). *Introduction to Infinite Dimensional Stochastic Analysis.* Science Press/Kluwer Academic Publishers. MR1851117

[16] Hosking, J. R. M. (1981). Fractional differencing. *Biometrika* **68**, 165–176. MR0614953

[17] Karatzas, I. and Shreve, S. E. (1991) *Brownian Motion and Stochastic Calculus.* Springer Verlag, New York. MR1121940

[18] Kuo, H. H. (1996). *White Noise Distribution Theory.* CRC Press, Boca Raton, Florida. MR1387829

[19] Kuo, H. H. (1975). *Gaussian Measures in Banach spaces.* Lecture notes in mathematics **463**, Springer-Verlag, Berlin/Heidelberg/New York. MR0461643

[20] Kurtz, T. and Protter, P. (1991). Weak limit theorems for stochastic integrals and stochastic differential equations. *The Annals of Probability* **19**, 1035–1070. MR1112406

[21] Lee, Y. J. (1991). Analytic version of test functionals, Fourier transform, and a characterization of measures in white noise calculus. *Journal of Functional Analysis* **100**, 359–380. MR1125230

[22] Marinucci, D., and P. M. Robinson (1999). Alternative forms of Brownian motion. *Journal of Statistical Planning and Inference* **80**, 111–122. MR1713794

[23] Potthoff, J. and Streit, L. (1991). A characterization of Hida distributions. *Journal of Functional Analysis* **101**, 212–229. MR1132316

[24] Protter, P. E. (2004) *Stochastic Integration and Differential Equations, 2nd Edition*. Springer, New York. MR2020294

[25] Robinson, P. M. (2008). Multiple local Whittle estimation in stationary system. *Annals of Statistics* **36**, 2508–2530. MR2458196

[26] Tanaka, K. (1999). The nonstationary fractional unit root. *Econometric Theory* **15**, 549–582. MR1717967

[27] Wang, B., Wang, M., and Chan, N. H. (2015). Residual-based test for fractional cointegration. *Economics Letters* **126**, 43–46.