Bernoulli **23**(1), 2017, 710–742
DOI: 10.3150/15-BEJ697

The impact of the diagonals of polynomial forms on limit theorems with long memory

SHUYANG BAI* and MURAD S. TAQQU**

Department of Mathematics and Statistics, Boston University, 111 Cumminton Street, Boston, MA 02215, USA. E-mail: *bsy9142@bu.edu; **bumastat@gmail.com

We start with an i.i.d. sequence and consider the product of two polynomial-forms moving averages based on that sequence. The coefficients of the polynomial forms are asymptotically slowly decaying homogeneous functions so that these processes have long memory. The product of these two polynomial forms is a stationary nonlinear process. Our goal is to obtain limit theorems for the normalized sums of this product process in three cases: exclusion of the diagonal terms of the polynomial form, inclusion, or the mixed case (one polynomial form excludes the diagonals while the other one includes them). In any one of these cases, if the product has long memory, then the limits are given by Wiener chaos. But the limits in each of the cases are quite different. If the diagonals are excluded, then the limit is expressed as in the product formula of two Wiener–Itô integrals. When the diagonals are included, the limit stochastic integrals are typically due to a single factor of the product, namely the one with the strongest memory. In the mixed case, the limit stochastic integral is due to the polynomial form without the diagonals irrespective of the strength of the memory.

Keywords: diagonals; long memory; noncentral limit theorem; self-similar processes; Volterra; Wiener

1. Introduction

Let $X(n)$ be a stationary process with mean 0 and finite variance. We are interested in the following weak convergence of normalized partial sum to a process $Z(t)$:

$$\frac{1}{A(N)}\sum_{n=1}^{[Nt]} X(n) \Rightarrow Z(t) \tag{1}$$

as $N \to \infty$ where $A(N) \to \infty$ is a suitable normalization. The limit $Z(t), t \geq 0$ if it exists, has stationary increments and is self-similar with some index $H > 0$, that is, for any $a > 0$, $\{Z(at), t \geq 0\}$ and $\{a^H Z(t), t \geq 0\}$ have the same finite-dimensional distributions. The parameter H is called the *memory parameter*[1] of the process $X(n)$ and the *Hurst index* or *self-similarity parameter* of the limit process $Z(t)$. The higher the value of H, the stronger the memory of the process $X(n)$.

[1] A precise definition of memory parameter is given in Definition 4.3.

1350-7265

When the dependence in $X(n)$ is weak, one typically ends up in (1) with

$$A(N)=\left(\mathrm{Var}\left[\sum_{n=1}^{N}X(n)\right]\right)^{1/2}\sim cN^{1/2}$$

as $N\to\infty$ for some $c>0$, and $Z(t)$ is the Brownian motion. These types of limit theorems are often called *central limit theorems*.

When, however, the dependence in $X(n)$ is so strong that $\mathrm{Var}[\sum_{n=1}^{N}X(n)]$ grows faster than the linear speed N, and typically as N^{2H} with $H\in(1/2,1)$, the limit process $Z(t)$ in (1) is no longer Brownian motion. $Z(t)$ is in this case a self-similar process with stationary increments which has a *Hurst index* H (see [7]). This type of limit theorems involving non-Brownian limits are often called *noncentral limit theorems*. When the process $X(n)$ is nonlinear and has long memory, the limit $Z(t)$ can be non-Gaussian (e.g., [6,15,16]).

In [1], a noncentral limit theorem is established for an off-diagonal polynomial-form process called kth order *discrete chaos process*:

$$Y'(n)=\sum_{0<i_1,\dots,i_k<\infty}^{\prime}a(i_1,\dots,i_k)\varepsilon_{n-i_1}\cdots\varepsilon_{n-i_k},\tag{2}$$

where the prime $'$ indicates that we do not sum on the diagonals $i_p=i_q$, $p\neq q$, the noise ε_i's are i.i.d. random variables with mean 0 and variance 1, and $a(\cdot)$ is asymptotically some homogeneous function g called *generalized Hermite kernel* (GHK). The limit $Z(t)$, called a *generalized Hermite process*, is expressed by a k-fold *Wiener–Itô integral*:

$$Z(t)=\int_{\mathbb{R}^k}^{\prime}\int_0^t g(s-x_1,\dots,s-x_k)1_{\{s>x_1,\dots,s>x_k\}}\,\mathrm{d}s\,B(\mathrm{d}x_1)\cdots B(\mathrm{d}x_k),\tag{3}$$

where the prime $'$ indicates that we do not integrate on the diagonals $x_p=x_q$, $p\neq q$, and $B(\cdot)$ is Brownian motion. These processes $Z(t)$ include the *Hermite process* considered in [6,16] and [15].

In [2], a noncentral limit theorem is established for a polynomial-form process called kth order *discrete Volterra process*:

$$Y(n)=\sum_{0<i_1,\dots,i_k<\infty}a(i_1,\dots,i_k)\varepsilon_{n-i_1}\cdots\varepsilon_{n-i_k},\tag{4}$$

which differs from $Y'(n)$ in (2) by including the diagonals, and where $a(\cdot)$ is asymptotically $g(\cdot)$, some special type of generalized Hermite kernel called *generalized Hermite kernel of Class (B)* (GHK(B)). The limit $Z(t)$ can be heuristically thought as (3) with diagonals included, and is precisely expressed as a k-fold *centered Wiener–Stratonovich integral*, which is a linear combination of certain Wiener–Itô integrals of orders lower than or equal to k (see [2]).

In this paper, we contrast the effect of two types of stationary sequences in the limit theorem (1). The first stationary sequence is

$$X(n)=Y_1'(n)Y_2'(n),\tag{5}$$

that is, a product of two long memory chaos processes (2) which *exclude* the diagonals. The second stationary sequence is

$$X(n) = Y_1(n)Y_2(n), \tag{6}$$

that is, a product of two long memory processes in (4) which *include* the diagonals. We also consider the mixed case

$$X(n) = Y_1'(n)Y_2(n). \tag{7}$$

Limit theorems for such types of product are of interest, for example, in statistical inference involving long memory processes with different memory parameters ([10], see also Proposition 11.5.6 of [8]), and in the study of covariation of fractional Brownian motions with different Hurst indexes [11]. Typically, the factor processes Y there are assumed to be either linear (or Gaussian) or a transformation of linear process (or Gaussian), which yields in the limit a generalized Rosenblatt processes where $g(x_1, x_2) = x_1^{\gamma_1} x_2^{\gamma_2}$ in (3). By taking the factors Y to be some nonlinear processes as in (5), (6) and (7), one can obtain much richer limit structures, which are briefly described below.

We show that in the case (5), the limit in (1) is expressed as Wiener–Itô integrals which can be obtained by using a rule similar to that used for computing the product of two Wiener–Itô integrals. In fact, if the stationary sequences $Y_1'(n)$ and $Y_2'(n)$ have, respectively, memory parameters $H_1, H_2 \in (1/2, 1)$ with $H_1 + H_2 > 3/2$, then the limit in (1) has Hurst index

$$H = H_1 + H_2 - 1 \in (1/2, 1).$$

In the case (6), in contrast, the limit stochastic integrals are typically due to a single factor $Y_1(n)$ or $Y_2(n)$, namely, the one with the strongest memory parameter. The Hurst index of the limit is then

$$\max(H_1, H_2) \in (1/2, 1)$$

which is always greater than $H_1 + H_2 - 1$. In the case (7), only the off-diagonal factor $Y_1'(n)$ contributes to the limit stochastic integral, irrespective of the strength of the memory.

The paper is organized as follows. Section 2 contains some background. We state the main results in Section 3, namely, Theorem 3.5 for processes without diagonals, Theorem 3.6 for processes with diagonals and Theorem 3.8 for the mixed case. Section 4 provides some preliminary results used in the proofs. Section 5 contains the proofs of the theorems.

2. Background

The following notation will be used throughout: $\mathbf{0}$ denotes the zero vector $(0, 0, \dots, 0)$ and $\mathbf{1} = (1, 1, \dots, 1)$ denotes the vector with ones in every component. For two vectors $\mathbf{x}$ and $\mathbf{y}$ with the same dimension, we write $\mathbf{x} \le \mathbf{y}$ (or $<, \ge, >$) if the inequality holds componentwise. We let

$$[x] = \sup\{n \in \mathbb{Z}: n \le x\}$$

for any real x and for a real vector $\mathbf{x} = (x_1, \dots, x_k)$, we define

$$[\mathbf{x}] = \big([x_1], \dots, [x_k]\big).$$

The notation 1_A denotes the indicator function of a set A. The value of a constant $C > 0$ or $c > 0$ may change from line to line.

In [1], the following classes of functions were introduced.

Definition 2.1. *A measurable function g defined on $\mathbb{R}_+^k$ is called a* generalized Hermite kernel (*GHK*) *with homogeneity exponent*

$$\alpha \in \left(-\frac{k+1}{2}, -\frac{k}{2}\right), \tag{8}$$

if it satisfies

1. $g(\lambda \mathbf{x}) = \lambda^\alpha g(\mathbf{x})$, $\forall \lambda > 0$;
2. $\int_{\mathbb{R}_+^k} |g(\mathbf{1} + \mathbf{x}) g(\mathbf{x})| \,\mathrm{d}\mathbf{x} < \infty$;

A GHK g is said to belong to Class (*B*) [*abbreviated as GHK*(*B*)], *if g is a.e. continuous on $\mathbb{R}_+^k$ and*

$$\big|g(\mathbf{x})\big| \le c\|\mathbf{x}\|^\alpha = c(x_1 + \cdots + x_k)^\alpha$$

($\|\cdot\|$ *is the L^1-norm*) *for some constant $c > 0$.*

Remark 2.2. As it was shown in Theorem 3.5 of [1], if g is a GHK, then

$$\int_0^t \big|g(s\mathbf{1} - \mathbf{x})\big| 1_{\{s\mathbf{1} > \mathbf{x}\}} \,\mathrm{d}s < \infty$$

for a.e. $\mathbf{x} \in \mathbb{R}^k$, and the function

$$h_t(\mathbf{x}) := \int_0^t g(s\mathbf{1} - \mathbf{x}) 1_{\{s\mathbf{1} > \mathbf{x}\}} \,\mathrm{d}s \in L^2\big(\mathbb{R}^k\big).$$

Using a GHK, one can define a self-similar process with stationary increments on a Wiener chaos as follows.

Definition 2.3. *Let g be a GHK on $\mathbb{R}_+^k$ with homogeneity exponent $\alpha \in (-\frac{k+1}{2}, -\frac{k}{2})$, then* (3) *is called a* generalized Hermite process $Z(t)$. *It is self-similar with Hurst index*

$$H = \alpha + k/2 + 1. \tag{9}$$

Example 2.4. If

$$g(\mathbf{x}) = \prod_{j=1}^k x_j^\gamma,$$

where $-1/2 - 1/k < \gamma < -1/2$, then $Z(t)$ in (3) is the *Hermite process* considered in [6] and [16].

Note that GHK(B) does not include the kernel in Example 2.4. We use a GHK(B) because of its boundedness property. The subclass of GHK(B) is, in fact, a dense subset in the whole class of GHK (see Remark 3.17 of [1]).

We now state two limit theorems, the first for the discrete chaos process $Y'(n)$ defined in (2) where the diagonals are excluded, and the second for the Volterra process $Y(n)$ defined in (4) which includes the diagonals.

Suppose that g is a GHK(B) on $\mathbb{R}_+^k$, $L(\cdot)$ is a bounded function defined on $\mathbb{Z}_+^k$ such that

$$\lim_{n\to\infty} L\big([n\mathbf{x}] + \mathbf{B}(n)\big) = 1$$

for any $\mathbf{x} \in \mathbb{R}_+^k$ and any $\mathbb{Z}_+^k$-valued bounded function $\mathbf{B}(n)$, and suppose that the coefficient $a(\cdot)$ in (2) is given by

$$a(\mathbf{i}) = g(\mathbf{i})L(\mathbf{i}). \tag{10}$$

Proposition 2.5 (Theorem 6.5 of [1]). *The following weak convergence holds in* $D[0,1]$:

$$\frac{1}{N^H}\sum_{n=1}^{[Nt]} Y'(n) \Rightarrow Z(t) := I_k(h_t), \tag{11}$$

where $H = \alpha + k/2 + 1 \in (1/2, 1)$,

$$h_t(\mathbf{x}) = \int_0^t g(s\mathbf{1} - \mathbf{x}) 1_{\{s\mathbf{1}>\mathbf{x}\}}\, \mathrm{d}s \tag{12}$$

with g as in (10), *and $I_k(\cdot)$ denotes the k-fold Wiener–Itô integral*, *so that $Z(t)$ is a generalized Hermite process* (3).

We now consider the limit when the diagonals are included. If g is GHK(B) on $\mathbb{R}_+^k$ and is in addition symmetric, we define the following function g_r by identifying r pairs of variables of g and integrating them out, as follows:

$$g_r(\mathbf{x}) = \int_{\mathbb{R}_+^r} g(y_1, y_1, \dots, y_r, y_r, x_1, \dots, x_{k-2r})\, \mathrm{d}\mathbf{y}. \tag{13}$$

In [2], a noncentral limit theorem was established for the Volterra process $Y(n)$ in (4). Let

$$a(\cdot) = g(\cdot)L(\cdot)$$

in (4) be given as in (10) assuming in addition that g is symmetric.

Proposition 2.6 (Theorem 6.2 of [2]). *One has the following weak convergence in* $D[0,1]$:

$$\frac{1}{N^H}\sum_{n=1}^{[Nt]} Y(n) \Rightarrow Z(t) := \sum_{0\le r<k/2} d_{k,r} Z_{k-2r}(t), \tag{14}$$

where $H = \alpha + k/2 + 1 \in (1/2, 1)$,

$$d_{k,r} = \frac{k!}{2^r (k-2r)!r!}, \tag{15}$$

and

$$Z_{k-2r}(t) := \int'_{\mathbb{R}^{k-2r}} \int_0^t g_r(s\mathbf{1} - \mathbf{x}) 1_{\{s\mathbf{1}>\mathbf{x}\}} \, \mathrm{d}s \, B(\mathrm{d}x_1)\cdots B(\mathrm{d}x_k) \tag{16}$$

is a $(k-2r)$*th order generalized Hermite process with GHK given by* g_r *in* (13).

Remark 2.7. The limit process $Z(t)$ in (14) can be simply expressed in terms of a centered Wiener–Stratonovich integral $\mathring{I}_k^c(\cdot)$ as

$$Z(t) = \mathring{I}_k^c(h_t), \tag{17}$$

where h_t is as in (12), and where

$$\mathring{I}_k^c(\cdot) = \sum_{0\le r<k/2} d_{k,r} I_{k-2r}\big(\tau^r \cdot\big).$$

The integral $\mathring{I}_k^c(\cdot)$ differs from the Wiener–Stratonovich integral

$$\mathring{I}_k(\cdot) := \sum_{0\le r\le [k/2]} d_{k,r} I_{k-2r}\big(\tau^r \cdot\big)$$

introduced in [9] by excluding the term $r = k/2$ when k is even. Here, the operator τ^r identifies r pairs of variables of h and integrates them out (see [2]). The operator τ^r is often called a "trace operator."

3. Statement of the main results

We state here the main results, and defer the proofs to Sections 5.1 and 5.2. In the statement of the results, the following expressions are used.

Definition 3.1. *Let* $X(n)$ *be a stationary process with finite variance. We say that*:

1. $X(n)$ *satisfies a* central limit theorem (*CLT*), *if*

$$N^{-1/2}\sum_{n=1}^{[Nt]}\big[X(n) - \mathbb{E}X(n)\big] \Rightarrow \sigma B(t) \tag{18}$$

   in $D[0,1]$, *where* $\sigma^2 = \sum_{n=-\infty}^{\infty} \mathrm{Cov}(X(n), X(0))$;

2. $X(n)$ *satisfies a* noncentral limit theorem (*NCLT*) *with a Hurst index* $H \in (1/2, 1)$ *and limit* $Z(t)$, *if*

$$N^{-H} \sum_{n=1}^{[Nt]} \bigl[X(n) - \mathbb{E}X(n)\bigr] \Rightarrow Z(t) \tag{19}$$

in $D[0, 1]$.

Remark 3.2. In case 1 above, the "long-run variance" σ^2 can be 0. In this case, we understand the limit theorem as degenerate (the normalization $N^{-1/2}$ is too strong). We do not consider here limit theorems involving a Hurst index $H < 1/2$. In case 2, the limit in (19) may be fractional Brownian motion.

We now consider separately the cases where the diagonals of the polynomial forms are excluded (chaos processes) and when they are included (Volterra processes).

3.1. Limit theorem for a product of long-memory chaos processes

Suppose that we have the following two discrete chaos processes (off-diagonal polynomial forms):

$$Y_1'(n) = \sum_{\mathbf{i}\in\mathbb{Z}_+^{k_1}}' a^{(1)}(\mathbf{i})\varepsilon_{n-i_1} \cdots \varepsilon_{n-i_{k_1}}, \qquad Y_2'(n) = \sum_{\mathbf{i}\in\mathbb{Z}_+^{k_2}}' a^{(2)}(\mathbf{i})\varepsilon_{n-i_1} \cdots \varepsilon_{n-i_{k_2}}, \tag{20}$$

where we assume that $a^{(j)} = g^{(j)} L^{(j)}$ as in (10) is symmetric, where $g^{(j)}$ is a symmetric GHK(B) with homogeneity exponent

$$\alpha_j \in (-k_j/2 - 1/2, -k_j/2), \qquad j = 1, 2.$$

Definition 2.3 suggests the following terminology.

Definition 3.3. *The index*

$$H = \alpha + k/2 + 1 \in (1/2, 1) \tag{21}$$

is called the associated Hurst index *of the coefficient* $a(\cdot) = g(\cdot)L(\cdot)$ *in* (10).

Remark 3.4. The associated Hurst indices of the coefficients in $Y_1'(n)$ and $Y_2'(n)$ will determine the Hurst index of the limit process $Z(t)$ in (1).

We want to obtain a limit theorem for the normalized partial sum of the product process:

$$X(n) := Y_1'(n)Y_2'(n). \tag{22}$$

Theorem 3.5. *Let $X(n)$ be the product process in (22). Suppose that H_j is the associated Hurst index of $a^{(j)}(\cdot)$, $j = 1, 2$, and assume that $\mathbb{E}|\varepsilon_i|^{4+\delta} < \infty$ for some $\delta > 0$.*

1. *If $H_1 + H_2 < 3/2$, then $X(n)$ satisfies the CLT (18);*
2. *If $H_1 + H_2 > 3/2$, then $X(n)$ satisfies the NCLT (19) with Hurst index $H = H_1 + H_2 - 1$ and limit*

$$Z(t) = \sum_{r=0}^{k} r! \binom{k_1}{r} \binom{k_2}{r} I_{k_1+k_2-2r}(h_{t,r}), \tag{23}$$

where $k = k_1 \wedge k_2$ if $k_1 \neq k_2$, and $k = k_1 - 1$ if $k_1 = k_2$. The integrand $h_{t,r}$ above is defined as

$$h_{t,r}(\mathbf{x}) = \int_0^t \big(g^{(1)} \otimes_r g^{(2)}\big)(s\mathbf{1} - \mathbf{x}) 1_{\{s\mathbf{1}>\mathbf{x}\}} \, \mathrm{d}s, \tag{24}$$

where

$$\begin{aligned} &g^{(1)} \otimes_r g^{(2)}(\mathbf{x}) \\ &\quad := \int_{\mathbb{R}_+^r} g^{(1)}(y_1, \dots, y_r, x_1, \dots, x_{k_1-r}) g^{(2)}(y_1, \dots, y_r, x_{k_1-r+1}, \dots, x_{k_2+k_2-2r}) \, \mathrm{d}\mathbf{y} \end{aligned} \tag{25}$$

is a GHK, and when $r = 0$, (25) is understood as the tensor product $g^{(1)} \otimes g^{(2)}$. When $r > 0$ in (25), we identify r variables of $g^{(1)}$ and $g^{(2)}$ and integrate over them.

This theorem is proved in Section 5.1.

3.2. Limit theorem for a product of long-memory Volterra processes

Let now

$$X(n) = Y_1(n) Y_2(n), \tag{26}$$

where

$$Y_1(n) = \sum_{\mathbf{i} \in \mathbb{Z}_+^{k_1}} a^{(1)}(\mathbf{i}) \varepsilon_{n-i_1} \cdots \varepsilon_{n-i_{k_1}}, \qquad Y_2(n) = \sum_{\mathbf{i} \in \mathbb{Z}_+^{k_2}} a^{(2)}(\mathbf{i}) \varepsilon_{n-i_1} \cdots \varepsilon_{n-i_{k_2}}. \tag{27}$$

We assume that $a^{(j)} = g^{(j)} L^{(j)}$ in (10) is symmetric, and $g^{(j)}$ is a symmetric GHK(B) with homogeneity exponent $\alpha_j \in (-k_j/2 - 1/2, -k_j/2)$, $j = 1, 2$. In this case, we can write

$$X(n) = \sum_{\mathbf{i} \in \mathbb{Z}_+^{k}} a(\mathbf{i}) \varepsilon_{i_1} \cdots \varepsilon_{i_k},$$

where $k = k_1 + k_2$, and

$$a = a^{(1)} \otimes a^{(2)}. \tag{28}$$

Let $\mathcal{C}_1^2$ to be the collection of partitions of the set $\{1, \dots, k_1\}$ such that each set in the partition contains at least 2 elements, and similarly let $\mathcal{C}_2^2$ be the same thing for $\{k_1 + 1, \dots, k_1 + k_2\}$. Any partition $\pi \in \mathcal{C}_j^2$ can be expressed as $\pi = (P_1, \dots, P_m)$, where P_i, $i = 1, \dots, m$, are subsets ordered according to their smallest elements. For example, if $\pi = \{\{1, 4\}, \{2, 3\}\}$, then $P_1 = \{1, 4\}$ and $P_2 = \{2, 3\}$. Let

$$c_j = \sum_{\pi \in \mathcal{C}_j^2} \sum_{\mathbf{i}>\mathbf{0}}' a_\pi^{(j)}(\mathbf{i}) \mu_\pi, \qquad j = 1, 2, \tag{29}$$

where

$$\mu_\pi = \mu_{p_1} \cdots \mu_{p_m} \qquad \text{with } \mu_p = \mathbb{E}\varepsilon_i^p$$

and $p_i = |P_i| \geq 2$ if $\pi = (P_1, \dots, P_m)$, and where $a_\pi^{(j)}(\cdot)$ denotes $a^{(j)}$ with its variables identified according to the partition π (see (54) below).

The limit theorem for the normalized partial sum of the centered $X(n)$ in (26) includes several cases. We shall use the centered multiple Wiener–Stratonovich integral $\mathring{I}_k^c(\cdot)$ introduced in (17). The theorem states that except for some low-dimensional cases (cases 1–4), the limit is up to some constant the same as the limit for a single factor, namely the one with the highest H_j (cases 5–7).

Theorem 3.6. *Let $X(n)$ be the product process in* (26), *where $a^{(j)}$ has associated Hurst index $H_j = \alpha_j + k_j/2 + 1 \in (1/2, 1)$ (Definition* 3.3*). Assume $\mathbb{E}|\varepsilon_i|^{2k_1+2k_2+\delta} < \infty$ for some $\delta > 0$. Then using the language of Definition* 3.1,

1. *if $k_1 = 1$, $k_2 = 1$, and $H_1 + H_2 < 3/2$, then $X(n)$ satisfies a CLT* (18);
2. *if $k_1 = 1$, $k_2 = 1$, and $H_1 + H_2 > 3/2$, then $X(n)$ satisfies a NCLT* (19) *with Hurst index $H_1 + H_2 - 1$ and limit*

$$Z(t) = \int_{\mathbb{R}^2}' \int_0^t g_1(s - x_1) g_2(s - x_2) 1_{\{s\mathbf{1}>\mathbf{x}\}} \, \mathrm{d}s \, B(\mathrm{d}x_1) B(\mathrm{d}x_2)$$

   (*nonsymmetric Rosenblatt process*);
3. *if $k_1 \geq 2$, $k_2 = 1$, and if c_1 in* (29) *is nonzero, then $X(n)$ satisfies a NCLT* (19) *with Hurst index H_2 and limit*

$$Z(t) = c_1 \int_{\mathbb{R}} \int_0^t g_2(s - x) 1_{\{s>x\}} \, \mathrm{d}s \, B(\mathrm{d}x)$$

   (*fractional Brownian motion*);
4. *if $k_1 = 1$, $k_2 \geq 2$, and if c_2 in* (29) *is nonzero, then $X(n)$ satisfies a NCLT* (19) *with Hurst index H_1 and limit*

$$Z(t) = c_2 \int_{\mathbb{R}} \int_0^t g_1(s - x) 1_{\{s>x\}} \, \mathrm{d}s \, B(\mathrm{d}x)$$

   (*fractional Brownian motion*);

5. *if* $k_1 \geq 2$, $k_2 \geq 2$, $H_1 > H_2$, *and if* c_2 *in* (29) *is nonzero, then* $X(n)$ *satisfies a NCLT* (19) *with Hurst index* H_1, *and the limit*
$$Z(t) = c_2 \mathring{I}^c_{k_1}(h_{t,1}),$$
where $h_{t,1}(\mathbf{x}) = \int_0^t g_1(s\mathbf{1} - \mathbf{x}) 1_{\{s\mathbf{1} > \mathbf{x}\}} \, \mathrm{d}s$;
6. *if* $k_1 \geq 2$, $k_2 \geq 2$, $H_1 < H_2$, *and if* c_1 *in* (29) *is nonzero, then* $X(n)$ *satisfies a NCLT* (19) *with Hurst index* H_2, *and the limit*
$$Z(t) = c_1 \mathring{I}^c_{k_2}(h_{t,2}),$$
where $h_{t,2}(\mathbf{x}) = \int_0^t g_2(s\mathbf{1} - \mathbf{x}) 1_{\{s\mathbf{1} > \mathbf{x}\}} \, \mathrm{d}s$;
7. *if* $k_1 \geq 2$, $k_2 \geq 2$, $H_1 = H_2$, *and if at least one of the* c_j*'s in* (29) *is nonzero, then* $X(n)$ *satisfies a NCLT* (19) *with Hurst index* $H_1 = H_2$, *and the limit*
$$Z(t) = c_1 \mathring{I}^c_{k_2}(h_{t,2}) + c_2 \mathring{I}^c_{k_1}(h_{t,1}).$$

Remark 3.7. These constants c_j's in the theorem are nonzero if, for example, every $a^{(j)}(\mathbf{i}) > 0$, $j = 1, 2$.

The theorem, which is proved in Section 5.2, seems bewildering at first glance. But there is structure into it. The cases 3 and 4 are symmetric, and so are the cases 5 and 6. Case 1 involves short-range dependence, while all the other cases involve long-range dependence. Case 2 involves the nonsymmetric Rosenblatt process, originally introduced by Maejima and Tudor [11]. Cases 3 and 4 involve fractional Brownian motion since one of the orders k equals 1. The typical cases are 5 (and 6). In these cases, quite surprisingly, it is not the orders k_1 or k_2 that matter, but the process $Y_1(n)$ or $Y_2(n)$ in (26) with the highest value of H. In the boundary case 7, where $H_1 = H_2$, they both contribute.

3.3. Limit theorem for the mixed case

Now we consider the mixed case (7), where $Y_1'(n)$ is as in (20) and $Y_2(n)$ is as in (27). Let
$$X(n) = Y_1'(n) Y_2(n). \tag{30}$$

We only state the case which does not overlap Theorem 3.5 and Theorem 3.6, that is, both $Y_1'(n)$ and $Y_2(n)$ are nonlinear: $k_1 \geq 2$ and $k_2 \geq 2$. The limit, up to some constant, turns out to be the same as the limit for the single factor $Y_1'(n)$.

Theorem 3.8. *Let* $X(n)$ *be the product process in* (30), *where* $a^{(j)}$ *has associated Hurst index* $H_j = \alpha_j + k_j/2 + 1 \in (1/2, 1)$ (*Definition* 3.3). *Assume* $k_1 \geq 2$, $k_2 \geq 2$ *and* $\mathbb{E}|\varepsilon_i|^{2+2k_2+\delta} < \infty$ *for some* $\delta > 0$. *Then using the language of Definition* 3.1, *if* c_2 *in* (29) *is nonzero, then* $X(n)$ *satisfies a NCLT* (19) *with Hurst index* $H_1 = \alpha_1 + k_1/2 + 1$, *and the limit is*
$$Z(t) = c_2 I_{k_1}(h_{t,1}),$$
where $h_{t,1}(\mathbf{x}) = \int_0^t g_1(s\mathbf{1} - \mathbf{x}) 1_{\{s\mathbf{1} > \mathbf{x}\}} \, \mathrm{d}s$.

Theorem 3.8 is proved in Section 5.3.

Remark 3.9. If the noises ε_i's are Gaussian, then the normalized partial sum

$$\frac{1}{A(N)} \sum_{n=1}^{[Nt]} X(n)$$

considered in Theorem 3.5 3.6 and 3.8 belongs to a Wiener chaos of finite order. There is a rich literature on obtaining Berry–Esseen type quantitative limit theorems for elements on Wiener chaos. For the case where the limit is Gaussian, see the monograph [12] and the references therein; for the case where the limit belongs to higher-order Wiener chaos, see [4,5] and [14]. The case where ε_i's are non-Gaussian may also be treated using techniques from [13].

The quantitative results mentioned above, however, seem not directly applicable to the limit theorems considered here. This is because, as it will be clear in the proofs of these theorems, $\frac{1}{A(N)} \sum_{n=1}^{[Nt]} X(n)$ does not have a "clean" structure as that considered in the works mentioned above. In particular, the decomposition of $\frac{1}{A(N)} \sum_{n=1}^{[Nt]} X(n)$ yields many terms. Some of the quantitative results mentioned above may be applicable to the terms which contribute to the limit, but there are other terms in the decomposition which converge in $L^2(\Omega)$ to zero. How to deal with these degenerate terms is an open problem.

4. Preliminary results

A central idea in establishing the limit theorems is to involve the *nonsymmetric discrete chaos process* which generalizes the chaos process in (2) by allowing different sequences of noises. We shall now define it. Let $\boldsymbol{\varepsilon}_i = (\varepsilon_i^{(1)}, \dots, \varepsilon_i^{(k)})$ be an i.i.d. vector where each component has mean 0 and finite variance. The components $\varepsilon_i^{(1)}, \dots, \varepsilon_i^{(k)}$ are typically dependent. Introduce the following nonsymmetric discrete chaos process

$$Y'(n) = \sum_{0<i_1,\dots,i_k<\infty}' a(i_1, \dots, i_k) \varepsilon_{n-i_1}^{(1)} \cdots \varepsilon_{n-i_k}^{(k)}, \tag{31}$$

where $\sum_{\mathbf{i} \in \mathbb{Z}_+^k}' a(\mathbf{i})^2 < \infty$ so that $X'(n)$ is well-defined in the $L^2(\Omega)$-sense. Let

$$\Sigma(i, j) = \mathbb{E} \varepsilon_n^{(i)} \varepsilon_n^{(j)}.$$

The autocovariance of $Y'(n)$ is then given by

$$\gamma(n) = \sum_\sigma \sum_{0<i_1,\dots,i_k<\infty}' a(i_1, \dots, i_k) a(i_{\sigma(1)} + n, \dots, i_{\sigma(k)} + n) \Sigma(i_1, i_{\sigma(1)}) \cdots \Sigma(i_k, i_{\sigma(k)}), \tag{32}$$

where in the summation σ runs over all the $k!$ permutations of $\{1, \dots, k\}$. The following lemma is useful for studying the asymptotic properties of the covariance of $X'(n)$.

Lemma 4.1. *Suppose that in* (31), *there exist constant* $c_0 > 0$ *and* $\gamma_j < -1/2$, $j = 1, \dots, k$, *such that*

$$\big|a(i_1, \dots, i_k)\big| \le c_0 i_1^{\gamma_1} \cdots i_k^{\gamma_k}. \tag{33}$$

Let

$$H^* = \alpha + k/2 + 1 \qquad \text{with } \alpha = \sum_{j=1}^{k} \gamma_j. \tag{34}$$

- *If* $H^* < 1/2$, *then* $\sum_{n=-\infty}^{\infty} |\gamma(n)| < \infty$, *and* $\mathrm{Var}[\sum_{n=1}^{N} Y'(n)] \le c_1 N$ *for some* $c_1 > 0$;
- *If* $H^* > 1/2$, *then* $|\gamma(n)| \le c_2 n^{2H^*-2}$ *for some* $c_2 > 0$, *and* $\mathrm{Var}[\sum_{n=1}^{N} Y'(n)] \le c_3 N^{2H^*}$ *for some* $c_3 > 0$.

Proof. The case $H^* < 1/2$ was proved in Proposition 5.4 in [2].

In the case $H^* > 1/2$, let $\widetilde{|a|}$ be the symmetrization of $|a|(\mathbf{i}) := |a(\mathbf{i})|$, then for $n \ge 0$, by (32) and (33),

$$\begin{aligned}
\big|\gamma(n)\big| &\le C_0 \sum_{\mathbf{i} \in \mathbb{Z}_+^k} \widetilde{|a|}(\mathbf{i} + n\mathbf{1}) \widetilde{|a|}(\mathbf{i}) \\
&\le C_1 \sum_{\sigma} \sum_{i_1=1}^{\infty} \cdots \sum_{i_k=1}^{\infty} (i_1 + n)^{\gamma_1} \cdots (i_k + n)^{\gamma_k} i_1^{\gamma_{\sigma(1)}} \cdots i_k^{\gamma_{\sigma(k)}} \\
&\le C_2 \sum_{\sigma} n^{\gamma_1 + \gamma_{\sigma(1)} + 1} \cdots n^{\gamma_k + \gamma_{\sigma(k)} + 1} = C_3 n^{2\alpha + k} = C_3 n^{2H^* - 2},
\end{aligned}$$

where the C_i's are positive constants, and σ in the summation runs over all the permutations of $\{1, \dots, k\}$. $\mathrm{Var}[\sum_{n=1}^{N} Y'(n)] \le c_3 N^{2H^*}$ then follows as a standard result. □

Remark 4.2. In the applications of Lemma 4.1, the inequality (33) is often not seen in this form. For example, the function $a(\cdot)$ defined on $\mathbb{Z}_+^k$ may satisfy

$$\big|a(\mathbf{i})\big| \le C(i_1 + \cdots + i_{k_1})^{\alpha_1} (i_{k_1+1} + \cdots + i_{k_1+k_2})^{\alpha_2},$$

for some $C > 0$, where $k_1 + k_2 = k$, and $\frac{\alpha_j}{k_j} < -\frac{1}{2}$, then it is easily verified by the arithmetic–geometric mean inequality

$$k^{-1} \sum_{j=1}^{k} y_j \ge \left(\prod_{j=1}^{k} y_j \right)^{1/k}$$

for $y_j > 0$, that (33) is satisfied since $\alpha < 0$. It is also verified for a function $a_\pi(\cdot)$ which is $a(\cdot)$ with some of its variables identified.

In general when applying Lemma 4.1, we will omit the verification of (33) which usually can be easily done as indicated above. We will merely count the *total homogeneity exponents* of the bound, which in the preceding example is $\alpha = \alpha_1 + \alpha_2$.

For convenience, we make the following definition.

Definition 4.3. *Let $X(n)$ be a stationary process with mean 0 and finite variance. We say*

- *$X(n)$ has a memory parameter of* at most (*denoted using $\le$*) H, *if*

$$\mathrm{Var}\left[\sum_{n=1}^{N} X(n)\right] \le cN^{2H}$$

  for some $c > 0$;
- *$X(n)$ has a memory parameter (denoted using $=$) H, if*

$$\mathrm{Var}\left[\sum_{n=1}^{N} X(n)\right] \sim cN^{2H}$$

  as $N \to \infty$ for some $c > 0$.

Remark 4.4. In view of the definition above, Lemma 4.1 states that if $Y'(n)$ in (2) satisfies (33), then $Y'(n)$ has a memory parameter of at most $1/2$ if $H^* < 1/2$ and of at most H^* if $H^* > 1/2$.

Proposition 4.5 (Proposition 5.4 of [2]). *Let $Y'(n)$ be given as in (31) with coefficient $a(\cdot)$ satisfying (33) and $H^* < 1/2$ in Lemma 4.1. Then*

$$N^{-1/2}\sum_{n=1}^{[Nt]}\bigl[Y'(n) - \mathbb{E}Y'(n)\bigr] \xrightarrow{f.d.d.} \sigma B(t),$$

where

$$\sigma^2 = \sum_{n=-\infty}^{\infty} \mathrm{Cov}\bigl[Y'(n), Y'(0)\bigr],$$

$B(t)$ is a standard Brownian motion, and $\xrightarrow{f.d.d.}$ stands for convergence of finite-dimensional distributions.

If each $\varepsilon_i^{(1)}, \dots, \varepsilon_i^{(k)}$ has a moment greater than 2, then the tightness of

$$N^{-1/2}\sum_{n=1}^{[Nt]}\bigl[Y'(n) - \mathbb{E}Y'(n)\bigr]$$

in $D[0,1]$ holds and thus $\xrightarrow{f.d.d.}$ can be replaced by weak convergence $\Rightarrow$ in $D[0,1]$.

Remark 4.6. The above $\xrightarrow{f.d.d.}$ or $\Rightarrow$ convergence also holds for a linear combination of different $Y'(n)$'s defined on a common i.i.d. noise vector $\boldsymbol{\varepsilon}_i$, while the $Y'(n)$'s can have different orders and involve different subvectors of $\boldsymbol{\varepsilon}_i$, provided the coefficient of each $Y'(n)$ satisfies (33) with $H^* < 1/2$.

We now state an important result concerning the weak convergence of a discrete chaos to a Wiener chaos. Let h be a function defined on $\mathbb{Z}^k$ such that $\sum'_{\mathbf{i}\in\mathbb{Z}^k_+} h(\mathbf{i})^2 < \infty$, where $'$ indicates the exclusion of the diagonals $i_p = i_q$, $p \neq q$. Let $Q_k(h)$ be defined as follows:

$$Q_k(h) = Q_k(h, \boldsymbol{\varepsilon}) = \sum'_{(i_1,\dots,i_k)\in\mathbb{Z}^k} h(i_1,\dots,i_k)\varepsilon_{i_1}\cdots\varepsilon_{i_k} = \sum'_{\mathbf{i}\in\mathbb{Z}^k} h(\mathbf{i}) \prod_{p=1}^{k} \varepsilon_{i_p}, \tag{35}$$

where ε_i's are i.i.d. noises. Observe that $Q_k(h)$ is invariant under permutation of the arguments of $h(i_1,\dots,i_k)$. So if $\tilde{h}$ is the symmetrization of h, then $Q_k(h) = Q_k(\tilde{h})$.

Suppose now that we have a sequence of function vectors $\mathbf{h}_n = (h_{1,n},\dots,h_{J,n})$ where each $h_{j,n} \in L^2(\mathbb{Z}^{k_j})$, $j = 1,\dots,J$.

Proposition 4.7 (Proposition 4.1 of [1]). *Let*

$$\tilde{h}_{j,n}(\mathbf{x}) = n^{k_j/2} h_{j,n}\big([n\mathbf{x}] + \mathbf{c}_j\big), \qquad j = 1,\dots,J,$$

where $\mathbf{c}_j \in \mathbb{Z}^k$. *Suppose that there exists* $h_j \in L^2(\mathbb{R}^{k_j})$, *such that*

$$\|\tilde{h}_{j,n} - h_j\|_{L^2(\mathbb{R}^{k_j})} \to 0 \tag{36}$$

as $n \to \infty$. *Then, as* $n \to \infty$, *we have the following joint convergence in distribution*:

$$\mathbf{Q} := \big(Q_{k_1}(h_{1,n}),\dots,Q_{k_J}(h_{J,n})\big) \xrightarrow{d} \mathbf{I} := \big(I_{k_1}(h_1),\dots,I_{k_J}(h_J)\big).$$

5. Proofs

5.1. Proof of Theorem 3.5 where diagonals are excluded

We first show that $g^{(1)} \otimes_r g^{(2)}$ in (25) is a GHK.

Lemma 5.1. *Let* $g^{(j)}$ *be a symmetric GHK(B) with homogeneity exponent* α_j *defined on* $\mathbb{R}_+^{k_j}$, $j = 1, 2$. *Suppose in addition that either* $k_1 \geq 2$ *or* $k_2 \geq 2$, *and that*

$$\alpha_1 + \alpha_2 > -(k_1 + k_2 + 1)/2, \tag{37}$$

and set

$$r = \begin{cases} 0,\dots,k_1 \wedge k_2, & \text{if } k_1 \neq k_2, \\ 0,\dots,k_1 - 1, & \text{if } k_1 = k_2. \end{cases}$$

If the function $g^{(1)} \otimes_r g^{(2)}$ *is nonzero, then it is a GHK on* $\mathbb{R}_+^{k_1+k_2-2r}$ *with homogeneity exponent* $\alpha_1 + \alpha_2 + r$.

Proof. When $r = 0$, $g^{(1)} \otimes g^{(2)}$ is a tensor product of two GHK(B)s. It is a GHK because condition 1 of Definition 2.1 is satisfied with homogeneity exponent

$$-(k_1 + k_2 + 1)/2 < \alpha_1 + \alpha_2 < -(k_1 + k_2)/2 \tag{38}$$

[see (8)], and condition 2 of Definition 2.1 is satisfied because

$$\begin{aligned}
&\int_{\mathbb{R}_+^{k_1+k_2}} \left|g^{(1)}(\mathbf{x}_1)g^{(2)}(\mathbf{x}_2)g^{(1)}(\mathbf{1}+\mathbf{x}_1)g^{(2)}(\mathbf{1}+\mathbf{x}_2)\right| \mathrm{d}\mathbf{x}_1\,\mathrm{d}\mathbf{x}_2 \\
&\quad = \int_{\mathbb{R}_+^{k_1}} \left|g^{(1)}(\mathbf{x})g^{(1)}(\mathbf{1}+\mathbf{x})\right| \mathrm{d}\mathbf{x} \int_{\mathbb{R}_+^{k_2}} \left|g^{(2)}(\mathbf{x})g^{(2)}(\mathbf{1}+\mathbf{x})\right| \mathrm{d}\mathbf{x} < \infty.
\end{aligned}$$

We shall now focus on the case $r > 0$.

Consider first $k_1 \geq 2$ and $k_2 = 1$ (the case $k_1 = 1$ and $k_2 \geq 2$ is similar), so that $g^{(2)}(x) = Cx^{\alpha_2}$ for some $C \neq 0$, where $\alpha_2 \in (-1, -1/2)$. Fix an $\mathbf{x} = (x_1, \dots, x_{k-1}) \in \mathbb{R}_+^{k_1-1}$, then

$$\int_0^\infty \left|g^{(1)}(y, \mathbf{x})\right| y^{\alpha_2}\,\mathrm{d}y \leq C \int_0^\infty (y + x_1 \cdots + x_{k_1-1})^{\alpha_1} y^{\alpha_2}\,\mathrm{d}y < \infty,$$

because near $y = 0$ (the other $\mathbf{x} > \mathbf{0}$), the integrand behaves like y^{α_2}, where $\alpha_2 > -1$, while near $y = \infty$, the integrand is like $y^{\alpha_1+\alpha_2}$, where $\alpha_1 < -1$ and $\alpha_2 < -1/2$. Hence, $g^{(1)} \otimes_1 g^{(2)}$ is well-defined in this case. It is easy to check that

$$g^{(1)} \otimes_1 g^{(2)}(\lambda\mathbf{x}) = \lambda^{\alpha_1+\alpha_2+1} g^{(1)} \otimes_1 g^{(2)}(\mathbf{x})$$

for any $\lambda > 0$ by using a change of variable and using the homogeneity of $g^{(j)}$. We are left to show that $g := g^{(1)} \otimes_1 g^{(2)}$ satisfies condition 2 of Definition 2.1. This is true because the function $f(x) := \int_0^\infty (x + y)^{\alpha_1} y^{\alpha_2}\,\mathrm{d}y$ is $f(x) = C_0 x^{\alpha_1+\alpha_2+1}$ for some $C_0 > 0$. So

$$\left|g^{(1)} \otimes_1 g^{(2)}(\mathbf{x})\right| \leq C(x_1 + \cdots + x_{k_1-1})^{\alpha_1+\alpha_2+1} =: g^*(\mathbf{x}) \tag{39}$$

for some $C > 0$. Note that $g^*(\cdot)$ is a GHK(B) on $\mathbb{R}^{k_1-1}$ with

$$-(k_1 - 1)/2 - 1/2 < \alpha_1 + \alpha_2 + 1 < -(k_1 - 1)/2$$

because $\alpha_1 < -1/2$, $\alpha_2 < -k_2/2$ and $\alpha_1 + \alpha_2 > -(1 + k_2 + 1)/2$ by assumption (37). So $g = g^{(1)} \otimes_1 g^{(2)}$ satisfies condition 2 of Definition 2.1 because the dominating function g^* does.

Suppose now that $k_1 \geq 2$ and $k_2 \geq 2$. Consider first the case $1 \leq r \leq (k_1 \wedge k_2) - 1$. Using the bound $g^{(j)}(\mathbf{x}) \leq C\|\mathbf{x}\|^{\alpha_j}$, one has by applying Cauchy–Schwarz and integrating power functions iteratively that

$$\begin{aligned}
&\left|g^{(1)} \otimes_r g^{(2)}(\mathbf{x})\right| \\
&\quad \leq C \int_{\mathbb{R}_+^r} (y_1 + \cdots + y_r + x_1 + \cdots + x_{k_1-r})^{\alpha_1} \\
&\qquad\qquad \times (y_1 + \cdots + y_r + x_{k_1-r+1} + \cdots + x_{k_2+k_2-2r})^{\alpha_2}\,\mathrm{d}y_1 \cdots \mathrm{d}y_r
\end{aligned} \tag{40}$$

$$
\begin{aligned}
&= C \int_{\mathbb{R}_+^{r-1}} \mathrm{d}y_1 \cdots \mathrm{d}y_{r-1} \left(\int_0^\infty (y_1 + \cdots + y_r + x_1 + \cdots + x_{k_1 - r})^{2\alpha_1} \, \mathrm{d}y_r \right)^{1/2} \\
&\qquad \times \left(\int_0^\infty (y_1 + \cdots + y_r + x_{k_1 - r + 1} + \cdots + x_{k_2 + k_2 - 2r})^{2\alpha_2} \, \mathrm{d}y_r \right)^{1/2} \\
&\le C \int_{\mathbb{R}_+^{r-1}} (y_1 + \cdots + y_{r-1} + x_1 + \cdots + x_{k_1 - r})^{\alpha_1 + 1/2} \\
&\qquad \times (y_1 + \cdots + y_{r-1} + x_{k_1 - r + 1} + \cdots + x_{k_2 + k_2 - 2r})^{\alpha_2 + 1/2} \, \mathrm{d}\mathbf{y} \\
&\cdots \\
&\le C (x_1 + \cdots + x_{k_1 - r})^{\alpha_1 + r/2} (x_{k_1 - r + 1} + \cdots + x_{k_1 + k_2 - 2r})^{\alpha_2 + r/2} =: g^*(\mathbf{x}).
\end{aligned}
$$

The dominating function g^* is a GHK because it is a tensor product of two GHK(B)'s on $\mathbb{R}_+^{k_j}$, $j = 1, 2$, and

$$
-\frac{(k_1 - r) + (k_2 - r) + 1}{2} < (\alpha_1 + r/2) + (\alpha_2 + r/2) < -\frac{(k_1 - r) + (k_2 - r)}{2},
$$

as in the inequality (38). Therefore, the bound $g^*(\mathbf{x})$, and hence the kernel $g^{(1)} \otimes_r g^{(2)}$ satisfy condition 2 of Definition 2.1. Moreover, the homogeneity exponent of $g^{(1)} \otimes_r g^{(2)}$ is $\alpha_1 + \alpha_2 + r$ in condition 1 of Definition 2.1. This can be easily verified as above by change of variables and using the homogeneity of $g^{(j)}$.

The only case left is: $k_1 \neq k_2 \ge 2$ and $r = k_1 \wedge k_2$. Suppose $k_1 < k_2$. In this case, condition 2 of Definition 2.1 can be checked by first applying the iterative Cauchy–Schwarz argument leading to (40) until only one variable of $g^{(1)}$ is unintegrated, and then bounding the last fold of integration similarly as in (39). Hence, in this case as well, $g^{(1)} \otimes_r g^{(2)}$ is GHK. □

The following lemma shows a noncentral convergence involving $g^{(1)} \otimes_r g^{(2)}$ appearing in (25).

Lemma 5.2. *Suppose that all the assumptions in Lemma 5.1 hold. Let $a^j(\cdot) = g^{(j)} L^{(j)}$, $j = 1, 2$, be as assumed before. Set*

$$
\begin{aligned}
X'_r(n) := \sum_{(\mathbf{u},\mathbf{i}) > \mathbf{0}}^{\prime} & a^{(1)}(u_1, \ldots, u_r, i_1, \ldots, i_{k_1 - r}) \\
& \times a^{(2)}(u_1, \ldots, u_r, i_{k_1 - r + 1}, \ldots, i_{k_1 + k_2 - 2r}) \varepsilon_{n - i_1} \cdots \varepsilon_{n - i_{k_1 + k_2 - 2r}},
\end{aligned}
$$

where ε_i's are i.i.d. with mean 0 *and variance* 1. *We then have*

$$
\frac{1}{N^H} \sum_{n=1}^{[Nt]} X'_r(n) \overset{f.d.d.}{\longrightarrow} Z_r(t) := I_{k_1 + k_2 - 2r}(h_{t,r})
$$

jointly for all the $r = 0, 1 \dots, k$ *where* k *is as defined in Theorem* 3.5, *and where*

$$H = \alpha_1 + \alpha_2 + (k_1 + k_2)/2 + 1 \in (1/2, 1).$$

Proof. In view of Proposition 4.7, we need only to prove the convergence for a single r and a single $t > 0$, and the joint convergence for different r's and t's follows. We assume for simplicity that $a^{(j)}(\cdot) = g^{(j)}(\cdot)$ (setting $L = 1$), and including a general L in (10) is easy. We focus on the case $r \geq 1$, since the case $r = 0$ follows from Theorem 6.5 of [1], although the proof for case $r = 0$ may be regarded as contained in the proof below with $\mathbf{u}$ being an empty vector.

Let $\mathbf{u} = (u_1, \dots, u_r)$, $\mathbf{i}_1 = (i_1, \dots, i_{k_1-r})$, $\mathbf{i}_2 = (i_{k_1-r+1}, \dots, i_{k_1+k_2-2r})$, and $\mathbf{i} = (\mathbf{i}_1, \mathbf{i}_2)$. We define the sum

$$\sum_{n=1}^{Nt} x_n := \sum_{n=1}^{[Nt]} x_n + \bigl(Nt - [Nt]\bigr) x_{[Nt]+1} = N \int_0^t x_{1+[Ny]} \, \mathrm{d}y.$$

Obviously,

$$\mathbb{E}\left[\frac{1}{N^H} \sum_{n=1}^{[Nt]} X'_r(n) - \frac{1}{N^H} \sum_{n=1}^{Nt} X'_r(n)\right]^2 \to 0$$

as $N \to \infty$. One can thus focus on $\frac{1}{N^H} \sum_{n=1}^{Nt} X'_r(n)$ instead.

$$\begin{aligned}\frac{1}{N^H} \sum_{n=1}^{Nt} X'_r(n) &= \sum_{\mathbf{i} \in \mathbb{Z}^{k_1+k_2-2r}}' \frac{1}{N^H} \sum_{n=1}^{Nt} \sum_{\mathbf{u} \in D(\mathbf{i},n)} g^{(1)}(\mathbf{u}, n\mathbf{1} - \mathbf{i}_1) 1_{\{n\mathbf{1} > \mathbf{i}_1\}} \\ &\qquad \times g^{(2)}(\mathbf{u}, n\mathbf{1} - \mathbf{i}_2) 1_{\{n\mathbf{1} > \mathbf{i}_2\}} \prod_{j=1}^{k_1+k_2-r} \varepsilon_{i_j} \\ &=: Q_{k_1+k_2-2r}(h_{N,t,r}),\end{aligned}$$

using the notation (35), where

$$h_{N,t,r}(\mathbf{i}) := \frac{1}{N^H} \sum_{n=1}^{Nt} \sum_{\mathbf{u} \in D(\mathbf{i},n)} g^{(1)}(\mathbf{u}, n\mathbf{1} - \mathbf{i}_1) g^{(2)}(\mathbf{u}, n\mathbf{1} - \mathbf{i}_2) 1_{\{n\mathbf{1} > \mathbf{i}\}}$$

and

$$D(\mathbf{i}, n) = \bigl\{\mathbf{u} \in \mathbb{Z}_+^r \colon u_p \neq u_q \text{ if } p \neq q; \text{ and } u_p \neq n - i_q \text{ even if } p = q\bigr\}.$$

Set $\mathbf{x}_1 \in \mathbb{R}^{k_1-r}$, $\mathbf{x}_1 \in \mathbb{R}^{k_2-r}$ and $\mathbf{x} = (\mathbf{x}_1, \mathbf{x}_2)$. Define

$$E(\mathbf{x}, N) = \bigl\{\mathbf{u} \in \mathbb{Z}_+^r \colon u_p \neq u_q \text{ if } p \neq q; \text{ and } u_p \neq n - [N x_q] - 1 \text{ even if } p = q\bigr\}.$$

In view of Proposition 4.7 and using the homogeneity of $g^{(j)}$'s, one writes:

$$
\begin{aligned}
\tilde{h}_{N,t,r}(\mathbf{x}) &= N^{(k_1+k_2-2r)/2} h_{N,t}\big([N\mathbf{x}]+\mathbf{1}\big)\\
&= \frac{1}{N^{\alpha_1+\alpha_2+r+1}} \sum_{n=1}^{Nt} \sum_{\mathbf{u}\in E(\mathbf{x},n)} g^{(1)}\big(\mathbf{u}, n\mathbf{1}-[N\mathbf{x}_1]-\mathbf{1}\big) g^{(2)}\big(\mathbf{u}, n\mathbf{1}-[N\mathbf{x}_2]-\mathbf{1}\big) 1_{\{n\mathbf{1}>\mathbf{i}\}}\\
&= \sum_{n=1}^{Nt} \frac{1}{N} \sum_{\mathbf{u}\in E(\mathbf{x},n)} \frac{1}{N^r} g^{(1)}\left(\frac{\mathbf{u}}{N}, \frac{n\mathbf{1}-[N\mathbf{x}_1]-\mathbf{1}}{N}\right) g^{(2)}\left(\frac{\mathbf{u}}{N}, \frac{n\mathbf{1}-[N\mathbf{x}_2]-\mathbf{1}}{N}\right) 1_{\{n\mathbf{1}>\mathbf{i}\}}\\
&= \int_0^t \mathrm{d}s \int_{\mathbb{R}_+^r} \mathrm{d}\mathbf{y}\, g^{(1)}\left(\frac{[N\mathbf{y}]+\mathbf{1}}{N}, \frac{[Ns]\mathbf{1}-[N\mathbf{x}_1]}{N}\right)\\
&\qquad \times g^{(2)}\left(\frac{[N\mathbf{y}]+\mathbf{1}}{N}, \frac{[Ns]\mathbf{1}-[N\mathbf{x}_2]}{N}\right) 1_{\{[Ns]\mathbf{1}>[N\mathbf{x}]\}\cap F(N)},
\end{aligned}
$$

where we correspond $\mathbf{u}$ to $[N\mathbf{y}]+\mathbf{1}$, n to $[Ns]+1$, and

$$
\begin{aligned}
F(N) = \big\{(\mathbf{x},\mathbf{y},s)\colon\ & [Ny_p]\neq[Ny_q], [Nx_p]\neq[Nx_q],\\
& \text{if } p\neq q;\ \text{ and } [Ny_p]\neq[Ns]-[Nx_q] \text{ even if } p=q\big\}.
\end{aligned}
$$

In view of Proposition 4.7, the goal is to show that

$$
\lim_{N\to\infty} \|\tilde{h}_{N,t,r} - h_{t,r}\|_{L^2(\mathbb{R}^{k_1+k_2-2r})} = 0, \tag{41}
$$

where $h_{t,r}$ is given in (24). By the a.e. continuity of $g^{(j)}$'s and the fact that $1_{F(N)} \to 1$ a.e. as $N\to\infty$, one has

$$
\begin{aligned}
& g^{(1)}\left(\frac{[N\mathbf{y}]+\mathbf{1}}{N}, \frac{[Ns]\mathbf{1}-[N\mathbf{x}_1]}{N}\right) g^{(2)}\left(\frac{[N\mathbf{y}]+\mathbf{1}}{N}, \frac{[Ns]\mathbf{1}-[N\mathbf{x}_2]}{N}\right) 1_{\{[Ns]\mathbf{1}>[N\mathbf{x}]\}\cap F(N)}\\
&\quad \to g^{(1)}(\mathbf{y}, s\mathbf{1}-\mathbf{x}_1) g^{(2)}(\mathbf{y}, s\mathbf{1}-\mathbf{x}_2) 1_{\{s\mathbf{1}>\mathbf{x}\}} \qquad \text{for a.e. } (\mathbf{x},\mathbf{y},s).
\end{aligned}
$$

We are left to establish suitable bound to apply the dominated convergence theorem. To this end, since $g^{(j)}(\mathbf{x}) \le C\|\mathbf{x}\|^{\alpha_j} =: g^{(j)*}(\mathbf{x})$ on $\mathbb{R}_+^{k_j}$, we have the following bound:

$$
\begin{aligned}
& \left| g^{(1)}\left(\frac{[N\mathbf{y}]+\mathbf{1}}{N}, \frac{[Ns]\mathbf{1}-[N\mathbf{x}_1]}{N}\right) g^{(2)}\left(\frac{[N\mathbf{y}]+\mathbf{1}}{N}, \frac{[Ns]\mathbf{1}-[N\mathbf{x}_2]}{N}\right)\right| 1_{\{[Ns]\mathbf{1}>[N\mathbf{x}]\}\cap F(N)}\\
&\quad \le g^{(1)*}\left(\frac{[N\mathbf{y}]+\mathbf{1}}{N}, \frac{[Ns]\mathbf{1}-[N\mathbf{x}_1]}{N}\right) g^{(2)*}\left(\frac{[N\mathbf{y}]+\mathbf{1}}{N}, \frac{[Ns]\mathbf{1}-[N\mathbf{x}_2]}{N}\right)\\
&\qquad \times 1_{\{[Ns]\mathbf{1}>[N\mathbf{x}]\}\cap F(N)}\\
&\quad \le C g^{(1)*}(\mathbf{y}, s\mathbf{1}-\mathbf{x}_1) g^{(2)*}(\mathbf{y}, s\mathbf{1}-\mathbf{x}_2) 1_{\{s\mathbf{1}>\mathbf{x}\}},
\end{aligned} \tag{42}
$$

where we have used the following facts: on the set $\{\mathbf{y} > \mathbf{0}, [Ns]\mathbf{1} > [N\mathbf{x}]\}$, we have $([N\mathbf{y}] + 1)/N > \mathbf{y}$, $([Ns] - [Nx_j])/N \geq \frac{1}{2}(s - x_j)$ (see relation (40) in the proof of Theorem 6.5 of [1]) and $g^{(j)*}$ decreases in its every variables, as well as the fact that $\{[Ns]\mathbf{1} > [N\mathbf{x}]\} \subset \{s\mathbf{1} > \mathbf{x}\}$. Note that

$$\begin{aligned}&\int_0^t \mathrm{d}s \int_{\mathbb{R}_+^r} \mathrm{d}\mathbf{y} g^{(1)*}(\mathbf{y}, s\mathbf{1} - \mathbf{x}_1) g^{(2)*}(\mathbf{y}, s\mathbf{1} - \mathbf{x}_2) 1_{\{s\mathbf{1} > \mathbf{x}\}} \\ &\quad = \int_0^t g^{(1)*} \otimes_r g^{(2)*}(s\mathbf{1} - \mathbf{x}) 1_{\{s\mathbf{1} > \mathbf{x}\}} \, \mathrm{d}s.\end{aligned} \tag{43}$$

Since $g^{(1)*}$ and $g^{(2)*}$ are GHK(B)s, so by Lemma 5.1, $g^{(1)*} \otimes_r g^{(2)*}$ is a GHK. This has two consequences. First, by Theorem 3.5 and Remark 3.6 of [1], the integral in $\mathrm{d}s\,\mathrm{d}\mathbf{y}$ on the left-hand side of (43) is finite for a.e. $\mathbf{x} \in \mathbb{R}^{k_1+k_2-2r}$. One can then apply the dominated convergence theorem to conclude that

$$\tilde{h}_{N,t,r}(\mathbf{x}) \to h_{t,r}(\mathbf{x}) \qquad \text{for a.e. } \mathbf{x} \in \mathbb{R}^{k_1+k_2-2r}. \tag{44}$$

But to obtain (41), we need L^2 convergence for the integral in $\mathrm{d}\mathbf{x}$. For this, we use the bound (42):

$$\left|\tilde{h}_{N,t,r}(\mathbf{x})\right| \leq h_{t,r}^*(\mathbf{x}) := C \int_0^t g^{(1)*} \otimes_r g^{(2)*}(s\mathbf{1} - \mathbf{x}) 1_{\{s\mathbf{1} > \mathbf{x}\}} \, \mathrm{d}s.$$

The second consequence of the fact that $g^{(1)*} \otimes_r g^{(2)*}$ is a GHK stems from Remark 2.2, which entails that $h_{t,r}^* \in L^2(\mathbb{R}^{k_1+k_2-2r})$, and hence (41) follows from (44) and the dominated convergence theorem. This concludes the proof of Lemma 5.2. □

We now decompose the product $X(n)$ in (22) in off-diagonal forms (31) as follows: let $\mathbf{u} = (u_1, \ldots, u_r) \in \mathbb{Z}_+^r$, $\mathbf{i}_1 = (i_1, \ldots, i_{k_1-r})$ and $\mathbf{i}_2 = (i_{k_1-r+1}, \ldots, i_{k_1+k_2-2r})$, and $\mathbf{i} = (\mathbf{i}_1, \mathbf{i}_2) \in \mathbb{Z}_+^{k_1+k_2-2r}$, then

$$\begin{aligned}X(n) &= Y_1'(n) Y_2'(n) \\ &= \sum_{r=0}^{k_1 \wedge k_2} r! \binom{k_1}{r} \binom{k_2}{r} \sum_{(\mathbf{u},\mathbf{i}) \in \mathbb{Z}_+^{k_1+k_2-r}}' a^{(1)}(\mathbf{u}, \mathbf{i}_1) a^{(2)}(\mathbf{u}, \mathbf{i}_2) \varepsilon_{n-u_1}^2 \cdots \varepsilon_{n-u_r}^2 \varepsilon_{n-i_1} \cdots \varepsilon_{n-i_{k_1+k_2-2r}},\end{aligned}$$

where we have used the symmetry of $a^{(j)}$'s, while the combinatorial coefficient

$$c(r, k_1, k_2) := r! \binom{k_1}{r} \binom{k_2}{r}$$

is obtained as the number of ways to pair r variables of $a^{(1)}$ to r variables of $a^{(2)}$. We write

$$\varepsilon_{n-i}^2 = 1 + \left(\varepsilon_{n-i}^2 - 1\right) =: A_0(\varepsilon_{n-i}) + A_2(\varepsilon_{n-i}),$$

where $A_0(\varepsilon) = 1$ and $A_2(\varepsilon) = \varepsilon^2 - 1$. These are Appell polynomials which will be introduced in more details in Section 5.2. Set $J_r = \{0, 2\} \times \cdots \times \{0, 2\}$. Then

$$Y_1'(n)Y_2'(n) = \sum_{r=0}^{k_1 \wedge k_2} c(r, k_1, k_2) \sum_{(\mathbf{u},\mathbf{i}) \in \mathbb{Z}_+^{k_1+k_2-r}}' \sum_{\mathbf{j} \in J_r} a^{(1)}(\mathbf{u}, \mathbf{i}_1) a^{(2)}(\mathbf{u}, \mathbf{i}_2) \times A_{j_1}(\varepsilon_{n-u_1}) \cdots A_{j_r}(\varepsilon_{n-u_r}) \varepsilon_{n-i_1} \cdots \varepsilon_{n-i_{k_1+k_2-2r}}.$$

The random variables in each summand are independent because the sum does not include diagonals. Observe that it is only when $k_1 = k_2$, that the mean

$$\mathbb{E} Y_1'(n)Y_2'(n) = k_1! \sum_{\mathbf{u} \in \mathbb{Z}_+^{k_1}}' a^{(1)}(\mathbf{u}) a^{(2)}(\mathbf{u})$$

may possibly be nonzero (this is the case when $r = k_1 = k_2$). Hence, one can use the k defined in Theorem 3.5 to write that

$$X(n) - \mathbb{E}X(n) = \sum_{r=0}^{k} \sum_{\mathbf{j} \in J_r} \sum_{(\mathbf{u},\mathbf{i}) \in \mathbb{Z}_+^{k_1+k_2-r}}' c(r, k_1, k_2) a^{(1)}(\mathbf{u}, \mathbf{i}_1) a^{(2)}(\mathbf{u}, \mathbf{i}_2) \times A_{j_1}(\varepsilon_{n-u_1}) \cdots A_{j_r}(\varepsilon_{n-u_r}) \varepsilon_{n-i_1} \cdots \varepsilon_{n-i_{k_1+k_2-2r}}. \tag{45}$$

A basic term of the preceding decomposition of $X(n) - \mathbb{E}X(n)$ is

$$X_{\mathbf{j}}^r(n) := \sum_{(\mathbf{u},\mathbf{i}) \in \mathbb{Z}_+^{k_1+k_2-r}}' c(r, k_1, k_2) a^{(1)}(\mathbf{u}, \mathbf{i}_1) a^{(2)}(\mathbf{u}, \mathbf{i}_2) \times A_{j_1}(\varepsilon_{n-u_1}) \cdots A_{j_r}(\varepsilon_{n-u_r}) \varepsilon_{n-i_1} \cdots \varepsilon_{n-i_{k_1+k_2-2r}}.$$

Note that $0 \le r \le k_1 \wedge k_2$ if $k_1 \ne k_2$, and $0 \le r \le k_1 - 1$ if $k_1 = k_2$, which implies $k_1 + k_2 - 2r \ge 1$ so that there is at least one i variable. Due to the symmetry of $a^{(j)}$'s, we can suppose without loss of generality that $j_1 = \cdots = j_s = 0$ and $j_{s+1} = \cdots = j_r = 2$, $0 \le s \le r$. One can hence rewrite the basic term as

$$X_{\mathbf{j}}^r(n) = \sum_{(\mathbf{u},\mathbf{i}) \in \mathbb{Z}_+^{k_1+k_2-r}}' c(r, k_1, k_2) a^{(1)}(\mathbf{u}, \mathbf{i}_1, \mathbf{i}_2) a^{(2)}(\mathbf{u}, \mathbf{i}_1, \mathbf{i}_3) \times A_2(\varepsilon_{n-i_1}) \cdots A_2(\varepsilon_{n-i_{r-s}}) \varepsilon_{n-i_{r-s+1}} \cdots \varepsilon_{n-i_{k_1+k_2-r-s}},$$

where

$$\mathbf{u} = (u_1, \ldots, u_s), \qquad \mathbf{i}_1 = (i_1, \ldots, i_{r-s}),$$
$$\mathbf{i}_2 = (i_{r-s+1}, \ldots, i_{k_1-s}), \qquad \mathbf{i}_3 = (i_{k_1-s+1}, \ldots, i_{k_1+k_2-r-s})$$

and $\mathbf{i} = (\mathbf{i}_1, \mathbf{i}_2, \mathbf{i}_3)$. Setting

$$a'(\mathbf{i}) = \sum_{\mathbf{u}\in K(\mathbf{i})} c(r, k_1, k_2) a^{(1)}(\mathbf{u}, \mathbf{i}_1, \mathbf{i}_2) a^{(2)}(\mathbf{u}, \mathbf{i}_1, \mathbf{i}_3), \tag{46}$$

with

$$K(\mathbf{i}) = \{\mathbf{u} > \mathbf{0}\colon\ u_p \neq u_q \text{ if } p \neq q;\ \text{ and } u_p \neq i_q \text{ even if } p = q\},$$

we get

$$X_{\mathbf{j}}^r(n) = \sum_{\mathbf{i}>\mathbf{0}}' a'(\mathbf{i}) A_2(\varepsilon_{n-i_1}) \cdots A_2(\varepsilon_{n-i_{r-s}}) \varepsilon_{n-i_{r-s+1}} \cdots \varepsilon_{n-i_{k_1+k_2-r-s}}. \tag{47}$$

We list here some useful elementary inequalities which will be used many times in the sequel:

Lemma 5.3. *Let $A > 0$, $B > 0$. If $\gamma < -1$, then*

$$\sum_{i=1}^{\infty} (A+i)^{\gamma} \le C A^{\gamma+1}. \tag{48}$$

If $\gamma < 0$, $\beta < -1$, then

$$\sum_{i=1}^{\infty} (A+i)^{\gamma} i^{\beta} \le C A^{\gamma}. \tag{49}$$

If $\gamma < -1/2$, $-1 < \beta < -1/2$, then

$$\sum_{i=1}^{\infty} (A+i)^{\gamma} i^{\beta} \le C A^{\gamma+\beta+1}. \tag{50}$$

If $\gamma < -1/2$, $\beta < -1/2$, then

$$\sum_{i=1}^{\infty} (A+i)^{\gamma} (B+i)^{\beta} \le C A^{\gamma+1/2} B^{\beta+1/2}. \tag{51}$$

Proof. To obtain inequality (48), we have

$$\begin{aligned}\sum_{i=1}^{\infty} (A+i)^{\gamma} &= \sum_{i=1}^{\infty} \int_{i-1}^{i} (A+i)^{\gamma}\,\mathrm{d}x \le \sum_{i=1}^{\infty} \int_{i-1}^{i} (A+x)^{\gamma}\,\mathrm{d}x \\ &= \int_0^{\infty} (A+x)^{\gamma}\,\mathrm{d}x = -(\gamma+1)^{-1} A^{\gamma+1}.\end{aligned}$$

For (49), note that $(A+i)^{\gamma} \le A^{\gamma}$ and $\sum_{i=1}^{\infty} i^{\beta} < \infty$.

For inequality (50), we have

$$\sum_{i=1}^{\infty}(A+i)^{\gamma}i^{\beta}=A^{\gamma+\beta+1}\sum_{i=1}^{\infty}\int_{i-1}^{i}(1+i/A)^{\gamma}(i/A)^{\beta}\,\mathrm{d}(x/A)\leq A^{\gamma+\beta+1}\int_{0}^{\infty}(1+y)^{\gamma}y^{\beta}\,\mathrm{d}y,$$

where the integral is finite since $\beta>-1$ and $\gamma+\beta<-1$.

The last one (51) is obtained by applying Cauchy–Schwarz and (48) as follows:

$$\sum_{i=1}^{\infty}(A+i)^{\gamma}(B+i)^{\beta}\leq\left[\sum_{i=1}^{\infty}(A+i)^{2\gamma}\right]^{1/2}\left[\sum_{i=1}^{\infty}(B+i)^{2\beta}\right]^{1/2}\leq CA^{\gamma+1/2}B^{\beta+1/2}. \qquad \square$$

Remark 5.4. The inequalities (48), (50) and (51) all raise the total power exponent by 1, while inequality (49) kills one of the exponents. These observations are useful in the proof below and also in Section 5.2.

We now state the proof of Theorem 3.5.

Proof of case 1 of Theorem 3.5. We want to apply Proposition 4.5. The condition $\mathbb{E}|\varepsilon_i|^{4+\delta}<\infty$ guarantees that $\mathbb{E}|A_2(\varepsilon)|^{2+\delta'}<\infty$ in (47) holds for some $\delta'>0$ and so the tightness in $D[0,1]$ holds.

We only need to show that $H^*<1/2$ in Lemma 4.1 for each of the basic terms $X_{\mathbf{j}}^r(n)$ in (47).

Suppose without loss of generality that $k_1\leq k_2$. Using the fact $|a^{(j)}(\mathbf{i})|\leq C\|\mathbf{i}\|^{\alpha_j}$ (recall that $\|\cdot\|$ is the L^1-norm), one can bound $a'(\mathbf{i})$ in (46). One has to distinguish two cases. In the first case, where $s<k_1$, one gets

$$\begin{aligned}\left|a'(\mathbf{i})\right| &\leq C\sum_{\mathbf{u}\in\mathbb{Z}_+^s}\left\|(\mathbf{u},\mathbf{i}_1,\mathbf{i}_2)\right\|^{\alpha_1}\left\|(\mathbf{u},\mathbf{i}_1,\mathbf{i}_3)\right\|^{\alpha_2}\\ &\leq C\sum_{\mathbf{u}\in\mathbb{Z}_+^s}\left(u_1+\cdots+u_s+\|\mathbf{i}_1\|+\|\mathbf{i}_2\|\right)^{\alpha_1}\left(u_1+\cdots+u_s+\|\mathbf{i}_1\|+\|\mathbf{i}_3\|\right)^{\alpha_2}\\ &\leq C\left(\|\mathbf{i}_1\|+\|\mathbf{i}_2\|\right)^{\alpha_1+s/2}\left(\|\mathbf{i}_1\|+\|\mathbf{i}_3\|\right)^{\alpha_2+s/2},\end{aligned}$$

after applying (51) to each of the s components of $\mathbf{u}$ iteratively (note: $\mathbf{i}_1$ may not be present). In the second case, where $s=r=k_1$, one gets

$$\begin{aligned}\left|a'(\mathbf{i})\right| &\leq C\sum_{\mathbf{u}\in\mathbb{Z}_+^s}\|\mathbf{u}\|^{\alpha_1}\left\|(\mathbf{u},\mathbf{i}_3)\right\|^{\alpha_2}\\ &\leq C\sum_{\mathbf{u}\in\mathbb{Z}_+^s}(u_1+\cdots+u_s)^{\alpha_1}\left(u_1+\cdots+u_s+\|\mathbf{i}_3\|\right)^{\alpha_2}\\ &\leq C\|\mathbf{i}\|^{\alpha_1+\alpha_2+s},\end{aligned}$$

after applying (51) $s-1$ times, and then (50) to the last component of $\mathbf{u}$. In either case, the total power exponent is raised by s.

According to (34), this yields

$$\begin{aligned} H^* &= \alpha_1+\alpha_2+s+(r-s+k_1-r+k_2-r)/2+1 \\ &= H_1+H_2+(s-r)/2-1 \\ &\le H_1+H_2-1<1/2, \end{aligned} \tag{52}$$

where the last strict inequality is due to the assumption $H_1+H_2<3/2$ of case 1. □

Proof of case 2 of Theorem 3.5. We now suppose that $H_1+H_2>3/2$. As was shown in case 1 above, the off-diagonal chaos coefficient $a'(\cdot)$ in (46) leads to

$$H^* = H_1+H_2+(s-r)/2-1.$$

When $s=r$, we have only factors $A_0(\varepsilon)=1$ in (47). The chaos process $X_r^{\mathbf{j}}(n)$ is up to some constant the process $X_r'(n)$ in Lemma 5.2. Note that Lemma 5.2 concludes a joint convergence for $X_r'(n)$ with different r's. So adding up all the terms corresponding to the case $r=s$ in (45), which yields

$$\sum_{r=0}^{k}\sum_{(\mathbf{u},\mathbf{i})\in\mathbb{Z}_+^{k_1+k_2-r}}^{\prime} r!\binom{k_1}{r}\binom{k_2}{r} a^{(1)}(\mathbf{u},\mathbf{i}_1)a^{(2)}(\mathbf{u},\mathbf{i}_2)\varepsilon_{n-i_1}\cdots\varepsilon_{n-i_{k_1+k_2-2r}},$$

one obtains the noncentral limit claimed in the theorem with a Hurst index $H=H_1+H_2-1>1/2$.

When $s<r$, the corresponding terms are negligible. Indeed,

$$H^* = H_1+H_2+(s-r)/2-1 \le H_1+H_2-1/2-1<1/2.$$

So by Lemma 4.1, the term $X_r^{\mathbf{j}}(n)$ has a memory parameter $H\le 1/2$ in the sense of Definition 4.3. Hence,

$$\lim_{N\to\infty}\mathbb{E}\left[N^{-(H_1+H_2-1)}\sum_{n=1}^{[Nt]}X_r^{\mathbf{j}}(n)\right]^2=0.$$

We have now shown the convergence of finite-dimensional distributions. Tightness in $D[0,1]$ is automatic since $H>1/2$ (see, e.g., Proposition 4.4.2 of [8]). □

5.2. Proof of Theorem 3.6 where diagonals are included

We first recall from [2] the off-diagonal decomposition of a general kth order Volterra process $X(n)$ in (4). The purpose is to decompose $X(n)$ into off-diagonal chaos terms as in (31). To this end, it is convenient to use Appell polynomials. Suppose that ε is a random variable with finite Kth moment. The Appell polynomial with respect to the law of ε is defined through the following recursive relation:

$$\frac{\mathrm{d}}{\mathrm{d}x} A_p(x) = pA_{p-1}, \qquad \mathbb{E}A_p(\varepsilon) = 0, \qquad A_0(x) = 1, \qquad p = 1, \dots, K.$$

We will use the following identity:

$$x^p = \sum_{j=0}^{p} \binom{p}{j} \mu_{p-j} A_j(x), \qquad p = 0, 1, 2, 3, \dots. \tag{53}$$

For more details about Appell polynomials, see for example Chapter 3.3 of [3].

Let $\mathcal{P}_k$ be the collection of all the partitions of $\{1, \dots, k\}$. We further express each partition $\pi \in \mathcal{P}_k$ as $\pi = (P_1, \dots, P_m)$ (so $m = |\pi|$), where the sets P_t's are *ordered* according to their smallest element. If we have a variable $\mathbf{i} \in \mathbb{Z}_+^k$, then $\mathbf{i}_\pi$ denotes a new variable where its components are identified according to π. For example, if $k = 3$, $\pi = (\{1, 2\}, \{3\})$ and $\mathbf{i} = (i_1, i_2, i_3)$, then $\mathbf{i}_\pi = (i_1, i_1, i_2)$. In this case we write $\pi = (P_1, P_2)$ where $P_1 = \{1, 2\}$ and $P_2 = \{3\}$. If $a(\cdot)$ is a function on $\mathbb{Z}_+^k$, then

$$a_\pi(i_1, \dots, i_m) := a(\mathbf{i}_\pi), \tag{54}$$

where $m = |\pi|$. In the preceding example, $a_\pi(\mathbf{i}) = a(i_1, i_2, i_2)$ with $m = 2$. We define a summation operator S'_T as follows: for any $T \subset \{1, \dots, |\pi|\}$, $S'_T(a_\pi)$ is obtained by summing a_π over its variables indicated by T off-diagonally, yielding a function with $|\pi| - |T|$ variables. For instance, if $\pi = (\{1, 5\}, \{2\}, \{3, 4\})$, then $\mathbf{i}_\pi = (i_1, i_2, i_3, i_3, i_1)$ and if $T = \{1, 3\}$, then

$$\left(S'_T a_\pi\right)(i) = \sum_{0<i_1, i_3<\infty}' a(i_1, i, i_3, i_3, i_1),$$

provided that it is well-defined. Note that in this off-diagonal sum, we require also that neither i_1 nor i_3 equals to i. If $T = \varnothing$, S'_T is understood to be the identity operator.

Now, by collecting various diagonal cases and using (53), $X(n)$ in (4) can be decomposed as

$$X(n) = \sum_{\pi \in \mathcal{P}_k} \sum_{\mathbf{i} \in \mathbb{Z}_+^m}' a_\pi(\mathbf{i}) \varepsilon_{n-i_1}^{p_1} \cdots \varepsilon_{n-i_m}^{p_m} = \sum_{\pi \in \mathcal{P}_k} \sum_{\mathbf{j} \in J(\pi)} X_\pi^{\mathbf{j}}(n), \tag{55}$$

where

$$X_\pi^{\mathbf{j}}(n) = \sum_{\mathbf{i} \in \mathbb{Z}_+^m}' a_\pi(\mathbf{i}) c(\mathbf{p}, \mathbf{j}) A_{j_1}(\varepsilon_{n-i_1}) \cdots A_{j_m}(\varepsilon_{n-i_m}), \tag{56}$$

$A_j(\cdot)$ is the jth order Appell polynomial with respect to the law of ε_i, $p_t = |P_t|$, $J(\pi) = \{0, \dots, p_1\} \times \cdots \times \{0, \dots, p_m\}$, and

$$c(\mathbf{p}, \mathbf{j}) = \binom{p_1}{j_1} \cdots \binom{p_m}{j_m} \mu_{p_1-j_1} \cdots \mu_{p_m-j_m}, \qquad \mu_j = \mathbb{E}\varepsilon_i^j. \tag{57}$$

Note that since by assumption $\mu_1 = 0$, when $j_t = 0$, it is only when $p_t \geq 2$ that it is possible to have a nonzero term.

In addition, the expression for the centered $X(n) - \mathbb{E}X(n)$ is the sum in (55) with $J(\pi)$ replaced by $J^+(\pi) := J(\pi) \setminus (0, \dots, 0)$, and

$$\mathbb{E}X(n) = \sum_{\pi \in \mathcal{P}_k} \sum_{\mathbf{i} \in \mathbb{Z}_+^m}' a_\pi(\mathbf{i}) \mu_{p_1} \cdots \mu_{p_m} = \sum_{\pi \in \mathcal{P}_k^2} \sum_{\mathbf{i} \in \mathbb{Z}_+^m}' a_\pi(\mathbf{i}) \mu_{p_1} \cdots \mu_{p_m}, \tag{58}$$

where $\mathcal{P}_k^2$ denotes the collection of partitions of $\{1, \dots, k\}$ such that each set in the partition contains at least 2 elements, namely, $p_t \geq 2$ for all $t = 1, \dots, m$.

So from (55), (56) and the discussion above (58), the summands in the off-diagonal decomposition of $X(n) - \mathbb{E}X(n)$ can be written as

$$X_\pi^{\mathbf{j}}(n) = \sum_{\mathbf{i} \in \mathbb{Z}_+^{k'}}' c(\mathbf{p}, \mathbf{j}) S_T' a_\pi(\mathbf{i}) A_{j_{t_1}}(\varepsilon_{n-i_{t_1}}) \cdots A_{j_{t_{k'}}}(\varepsilon_{n-i_{t_{k'}}}), \tag{59}$$

where $T = \{t = 1, \dots, m\colon\ j_t = 0\}$, and $\{t_1, \dots, t_{k'}\} = \{1, \dots, m\} \setminus T$ (thus $j_{t_1} \geq 1, \dots, j_{t_{k'}} \geq 1$). Note that $T \neq \{1, \dots, m\}$ since $\mathbf{j} \in J^+(\pi)$. In fact, $X_\pi^{\mathbf{j}}(n)$ is of the form (31) with $k = k'$ and $a(\cdot) = c(\mathbf{p}, \mathbf{j}) S_T' a_\pi(\cdot)$.

We now state the proof of Theorem 3.6 case by case. Recall that $C > 0$ denotes a constant whose value can change from line to line.

Proof of case 1. In this case, $g^{(1)}(i) = C_1 i^{\alpha_1}$, and $g^{(2)}(i) = C_2 i^{\alpha_2}$, where C_1 and C_2 are two nonzero constants. The off-diagonal decomposition (55) for the centered $X(n)$ is simply

$$X(n) - \mathbb{E}X(n) = \sum_{0<i_1,i_2<\infty}' a^{(1)}(i_1) a^{(2)}(i_2) \varepsilon_{n-i_1} \varepsilon_{n-i_2} + \sum_{0<i<\infty} a^{(1)}(i) a^{(2)}(i) A_2(\varepsilon_{n-i}), \tag{60}$$

where $A_2(\varepsilon_{n-i}) = \varepsilon_{n-i}^2 - 1$. Note that

$$\left|a^{(1)}(i_1) a^{(2)}(i_2)\right| \leq C i_1^{\alpha_1} i_2^{\alpha_2},$$

so the coefficient of the first term in (60) satisfies (33) with

$$H^* = \alpha_1 + \alpha_2 + (1+1)/2 + 1 = (H_1 - 3/2) + (H_2 - 3/2) + 2 < 1/2$$

by (21), since $H_1 + H_2 < 3/2$. For the second term in (33), one has

$$\left|a^{(1)}(i) a^{(2)}(i)\right| \leq C i^{\alpha_1 + \alpha_2},$$

which yields

$$H^* = \alpha_1 + \alpha_2 + 1/2 + 1 = (H_1 - 3/2) + (H_2 - 3/2) + 3/2 = H_1 + H_2 - 3/2 < 1/2, \tag{61}$$

since $H_1 < 1$ and $H_2 < 1$. Hence, Proposition 4.5 applies. □

Proof of case 2. Now the first term of (60) is subject to Proposition 2.5 with a Hurst index $H = \alpha_1 + \alpha_2 + 2 = H_1 + H_2 - 1 > 1/2$. One can see that for the second term of (60), relation (61) still holds. So by Lemma 4.1, the second term of (60) has a memory parameter $H \le 1/2$ in the sense of Definition 4.3, and hence with the normalization N^{-H}, the normalized partial sum of the second term of (60) converges to 0 in $D[0,1]$. □

Proof of case 3. Recall from (59) that the summands in the off-diagonal decomposition of $X(n) - \mathbb{E}X(n)$ are

$$X_\pi^{\mathbf{j}}(n) = \sum_{\mathbf{i}\in\mathbb{Z}_+^{k'}} c(\mathbf{p},\mathbf{j}) S'_T a_\pi(\mathbf{i}) A_{j_{t_1}}(\varepsilon_{n-i_{t_1}}) \cdots A_{j_{t_{k'}}}(\varepsilon_{n-i_{t_{k'}}}).$$

Consider first the following partition $\pi = (P_1, \dots, P_m)$ of $\{1, \dots, k_1, k_1+1\}$, which we express as

$$\pi = \bigl(P_1, \dots, P_{m_1}, \{k_1+1\}\bigr),$$

with $m_1 = m-1$, $\bigcup_{j=1}^{m_1} P_j = \{1, \dots, k_1\}$, and $P_m = \{k_1+1\}$. Let $T = \{1, \dots, m_1\}$. Recall that to have nonzero $c(\mathbf{p},\mathbf{j})$, one must require $|P_t| \ge 2$ if $t \in T$, and hence $2m_1 \le k_1$. Set $\pi_1 = \{P_1, \dots, P_{m_1}\}$ and let $\mathbf{u} \in \mathbb{Z}_+^{k_1}$. Then applying the off-diagonal summation S'_T, we get

$$\bigl(S'_T a_\pi\bigr)(i) = \sum_{u_p \ne u_q, u_p \ne i} a_{\pi_1}^{(1)}(\mathbf{u}) a^{(2)}(i) = \Bigl(\sum_{u_p \ne u_q} a_{\pi_1}^{(1)}(\mathbf{u})\Bigr) a^{(2)}(i) - R(i), \tag{62}$$

where the difference $R(i)$ includes the terms where some $u_p = i$. Since $|a^{(1)}(\mathbf{i})| \le C(i_1 + \cdots + i_{k_1})^{\alpha_1}$ which implies $|a_\pi^{(1)}(\mathbf{u})| \le C(u_1 + \cdots + u_{m_1})^{\alpha_1}$. Suppose without loss of generality that $u_{m_1} = i$, then by applying (48),

$$\bigl|R(i)\bigr| \le C \sum_{0<u_1,\dots,u_{m_1-1}<\infty} (u_1 + \cdots + u_{m_1-1} + i)^{\alpha_1} i^{\alpha_2} \le C i^{\alpha_2 + (\alpha_1 + m_1 - 1)},$$

where $\alpha_1 + m_1 - 1 < 0$ because $\alpha_1 < -k_1/2 \le -m_1 \le -1$. It follows that $|R(i)| \le C i^{\alpha_2 - \delta}$ for some $\delta > 0$. Since $k_2 = 1$, the term $R(i)$ defines the linear process $\sum_{i>0} R(i)\varepsilon_{n-i}$ but one with smaller memory parameter in the sense of Definition 4.3, than the linear process:

$$\mu_{\pi_1}\Bigl(\sum_{\mathbf{u}>\mathbf{0}}{}' a_{\pi_1}^{(1)}(\mathbf{u})\Bigr) \sum_{i=1}^\infty a^{(2)}(i)\varepsilon_{n-i},$$

resulting from the first term in the right-hand side of (62) (in this case $c(\mathbf{p},\mathbf{j}) = \mu_{\pi_1} := \mu_{p_1} \cdots \mu_{p_{m_1}}$). Collecting all such $\pi_1 \in \mathcal{C}_1^2$, one obtains $c_1 \sum_{i=1}^\infty a^{(2)}(i)\varepsilon_{n-i}$ with c_1 as given in (29). Applying Proposition 2.6 with $k = 1$, we get the noncentral limit in case 3, with a Hurst index

$$H = \alpha_2 + 1/2 + 1 = \alpha_2 + 3/2 = H_2.$$

We now show that in all the other cases, the memory parameter of $X_\pi^{\mathbf{j}}(n)$ is smaller than $H = \alpha_2 + 3/2$, which will conclude the proof. Observe first that

$$\bigl|a(\mathbf{i})\bigr| \le C(i_1 + \cdots + i_{k_1})^{\alpha_1} i_{k_1+1}^{\alpha_2}. \tag{63}$$

Let $\pi = \{P_1, \dots, P_m\}$ is a partition of $\{1, \dots, k_1 + 1\}$, and $T = \{t_1, \dots, t_l\}$, $l \le m - 1$. To bound $|(S'_T a)(\mathbf{i})|$, one can assume without loss of generality that either

(a) $P_j \cap \{k_1 + 1\} = \varnothing$ for $1 \le j \le m - 1$, $P_m = \{k_1 + 1\}$, $T \subset \{1, \dots, m - 1\}$, $\bigcup_{j=1}^l P_{t_j} \ne \{1, \dots, k_1\}$, or

(b) $P_m \cap \{k_1 + 1\} \ne \varnothing$, and $P_m \cap \{1, \dots, k_1\} \ne \varnothing$.

Observe that in the previous case we had $\bigcup_{j=1}^l P_{t_j} = \{1, \dots, k_1\}$ ($l = m_1 = m - 1$) and $P_m = \{k_1 + 1\}$.

In case (a), one has by (63) that

$$\bigl|a_\pi(\mathbf{i})\bigr| \le C(i_1 + \cdots + i_{m-1})^{\alpha_1} i_m^{\alpha_2}.$$

Since in case (a), $\bigcup_{j=1}^l P_{t_j}$ is a strict subset of $\{1, \dots, k_1\}$, we have $l < m - 1$, and thus by applying (48) iteratively, one has that

$$\begin{aligned}\bigl|\bigl(S'_T a_\pi\bigr)(\mathbf{i})\bigr| &\le \sum_{\mathbf{u}>\mathbf{0}} C(u_1 + \cdots + u_l + i_1 + \cdots + i_{m-l-1})^{\alpha_1} i_{m-l}^{\alpha_2} \\ &\le C(i_1 + \cdots + i_{m-l-1})^{\alpha_1+l} i_{m-l}^{\alpha_2},\end{aligned}$$

which results in H^* in (34) equal to

$$\begin{aligned}H^* &= (\alpha_1 + l + \alpha_2) + (m - l)/2 + 1 = \alpha_1 + \alpha_2 + m/2 + l/2 + 1 \\ &< -k_1/2 + \alpha_2 + (k_1 + 1)/2 + 1 = \alpha_2 + 3/2 = H_2\end{aligned}$$

since $\alpha_1 < -k_1/2$, and $m + l = 2l + (m - l) \le k_1 + 1$ (recall that each $|P_t| \ge 2$ if $t \in T$).

In case (b), one can write without loss of generality that

$$\bigl|a_\pi(\mathbf{i})\bigr| \le C(i_1 + \cdots + i_m)^{\alpha_1} i_1^{\alpha_2}$$

since π contains m partitions. If for the above a_π, the summation S'_T includes a sum over the index 1, that is, $1 \in T$, then using (48) and then (50), one has

$$\begin{aligned}\bigl|\bigl(S'_T a_\pi\bigr)(\mathbf{i})\bigr| &\le C \sum_{\mathbf{u}>\mathbf{0}} (u_1 + \cdots + u_l + i_1 + \cdots + i_{m-l})^{\alpha_1} u_1^{\alpha_2} \\ &\le C \sum_{u_1=1}^{\infty} (u_1 + i_1 + \cdots + i_{m-l})^{\alpha_1+l-1} u_1^{\alpha_2} \le C(i_1 + \cdots + i_{m-l})^{\alpha_1+\alpha_2+l}.\end{aligned}$$

Relation (50) does apply because on one hand $\alpha_2 > -1$, and on the other hand, we have $\alpha_1 + l - 1 < -1/2$ since $\alpha_1 < -k_1/2$ and $2(l - 1) + 1 < k_1$ because of $|P_t| \ge 2$ if $t \in T$. This leads to

H^* in (34) equal to

$$H^* = (\alpha_1 + \alpha_2 + l) + (m - l)/2 + 1 = \alpha_1 + \alpha_2 + m/2 + l/2 + 1 < \alpha_2 + 3/2 = H_2.$$

If the summation S'_T does not include the index 1, that is, if $1 \notin T$, one has

$$\begin{aligned}\left|\left(S'_T a_\pi\right)(\mathbf{i})\right| &\le C \sum_{\mathbf{u}>\mathbf{0}} (i_1 + \cdots + i_{m-l} + u_1 + \cdots + u_l)^{\alpha_1} i_1^{\alpha_2} \\ &\le C (i_1 + \cdots + i_{m-l})^{\alpha_1 + l} i_1^{\alpha_2},\end{aligned}$$

by (48), which also yields $H^* < \alpha_2 + 3/2 = H_2$. □

Proof of case 4. Same as case 3. □

Proof of case 5. We consider first in Part 1 all cases of $S'_T a_\pi$ in (59) which contribute to the limit, and in Part 2 negligible cases.

Part 1 of case 5: Suppose that π can be split into π_1 and π_2 which satisfy the following: the subpartition $\pi_1 = \{P_1, \ldots, P_{m_1}\}$ is a partition of $\{1, \ldots, k_1\}$, such that each P_j satisfies $|P_j| \le 2$, and at least one $|P_j| = 1$, $j = 1, \ldots, m_1$.

Thus, suppose without loss of generality that $|P_1| = 2, \ldots, |P_r| = 2$, $0 \le r < m_1$, and $|P_{r+1}| = \cdots = |P_{m_1}| = 1$. Require that the subpartition π_2 belongs to $\mathcal{C}_2^2$, where $\mathcal{C}_2^2$ is the collection of partitions of $\{k_1 + 1, \ldots, k_1 + \cdots + k_2\}$ such that each set in π_2 contains at least 2 elements. $\mathcal{C}_2^2$ is nonempty because $k_2 \ge 2$. Let

$$T = \{1, \ldots, r, m_1 + 1, \ldots, m_1 + m_2\}.$$

Setting $\mathbf{i} = (i_1, \ldots, i_{m_1 - r})$, $\mathbf{u} = (u_1, \ldots, u_r) \in \mathbb{Z}_+^r$ and $\mathbf{v} = (v_1, \ldots, v_{m_2}) \in \mathbb{Z}_+^{m_2}$, one can write

$$\left(S'_T a_\pi\right)(\mathbf{i}) = \sum_{\substack{u_p \ne u_q u_p \ne i_q, u_p \ne v_q, \\ v_p \ne v_q, v_p \ne i_q, \mathbf{u}, \mathbf{v} > \mathbf{0}}} a^{(1)}(u_1, u_1, \ldots, u_r, u_r, i_1, \ldots, i_{m_1 - r}) a^{(2)}_{\pi_2}(\mathbf{v}) \tag{64}$$

$$= \sum_{u_p \ne u_q, u_p \ne v_q, v_p \ne v_q, \mathbf{u}, \mathbf{v} > \mathbf{0}} a^{(1)}(u_1, u_1, \ldots, u_r, u_r, i_1, \ldots, i_{m_1 - r}) a^{(2)}_{\pi_2}(\mathbf{v}) - R_1(\mathbf{i}) \tag{65}$$

$$\begin{aligned} &= \sum_{u_p \ne u_q, \mathbf{u} > \mathbf{0}} a^{(1)}(u_1, u_1, \ldots, u_r, u_r, i_1, \ldots, i_{m_1 - r}) \\ &\quad \times \sum_{v_p \ne v_q, \mathbf{v} > \mathbf{0}} a^{(2)}_{\pi_2}(\mathbf{v}) - R_1(\mathbf{i}) - R_2(\mathbf{i}) \end{aligned} \tag{66}$$

for $i_p \ne i_q$. Relation (66) has the preceding three parts. We shall now apply Proposition 2.6 to the first part. Summing over all possible values of r, one gets a NCLT with Hurst index $H = \alpha_1 + k_1/2 + 1$, where the limit is

$$Z := c_2 \sum_{0 \le r < k_1/2} d_{k,r} Z_{k_1 - 2r},$$

where the process Z_{k_1-2r} is defined in (16) with $g_r = g_r^{(1)}$. Taking into account that in this setting, $c(\mathbf{p},\mathbf{j})$ in (57) and (59) is

$$\binom{p_1}{0}\cdots\binom{p_r}{0}\binom{p_{r+1}}{1}\cdots\binom{p_{m_1}}{1}\binom{p_{m_1+1}}{0}\cdots\binom{p_{m_1+m_2}}{0}(\mu_2)^r\mu_{p_{m_1+1}}\cdots\mu_{p_{m_1+m_2}}$$
$$=: \mu_{\pi_2},$$

since $\mu_2 = 1$, $p_1 = \cdots = p_r = 2$ and $p_{r+1} = \cdots = p_{m_1} = 1$, one gets the nonzero constant c_2 in (29). As in (17), we can express the limit $Z(t)$ as a centered Wiener–Stratonovich integral.

We shall now show that R_1 and R_2 in (66) lead only to terms with Hurst indices strictly less than $H = \alpha_1 + k_1/2 + 1$ in the sense of Definition 4.3, so they are negligible compared to the first term, and hence they do not contribute to the limit.

R_1 in (65) is obtained by taking the difference between the sum in (64) and the sum in (65). Thus R_1 is obtained by identifying some of the u and v variables in the sum in (64) with i variables. Using the fact $a^{(j)}(\mathbf{i}) \le C\|\mathbf{i}\|^{\alpha_j}$, one can see that one of the terms (a coefficient on $\mathbb{Z}_+^{m_1-r}$) in R_1 is bounded by

$$\sum_{u_p\neq u_q, u_p\neq v_q, v_p\neq v_q, \mathbf{u},\mathbf{v}>\mathbf{0}} C\big(\|\mathbf{u}\| + \|\mathbf{i}\|\big)^{\alpha_1}\big(\|\mathbf{v}\| + \big\|\mathbf{i}'\big\|\big)^{\alpha_2}, \tag{67}$$

where $\mathbf{u} = (u_1,\ldots,u_{r-s_1})$, $\mathbf{i} = (i_1,\ldots,i_{m_1-r})$, $\mathbf{v} = (v_1,\ldots,v_{m_2-s_2})$, $\mathbf{i}' = (i_1,\ldots,i_t)$, where

$$0 \le s_1 \le r \wedge (m_1 - r), \qquad 0 \le t \le s_2 \le m_2 \wedge (m_1 - r).$$

If $t = 0$, then $s_2 = 0$, and in addition, either $s_1 > 0$ or $s_2 > 0$. Note that $\mathbf{i}'$ is a subvector of $\mathbf{i}$.

By (48), the term (67) is bounded by

$$\sum_{\mathbf{u},\mathbf{v}>\mathbf{0}} C\big(\|\mathbf{u}\| + \|\mathbf{i}\|\big)^{\alpha_1}\big(\|\mathbf{v}\| + \big\|\mathbf{i}'\big\|\big)^{\alpha_2} \le \begin{cases} C\|\mathbf{i}\|^{\alpha_1+r-s_1} & \text{if } t = 0; \\ C\|\mathbf{i}\|^{\alpha_1+r-s_1}\big\|\mathbf{i}'\big\|^{\alpha_2+m_2-s_2} & \text{if } t > 0. \end{cases}$$

When $t = s_2 = 0$, one must have $s_1 > 0$, and so the term yields

$$H^* = \alpha_1 + r - s_1 + (m_1 - r)/2 + 1 = \alpha_1 + (r + m_1)/2 + 1 - s_1 < \alpha_1 + k_1/2 + 1,$$

because

$$r + m_1 = 2r + (m_1 - r) = k_1.$$

When $s_2 \ge t > 0$, it yields an

$$\begin{aligned} H^* &= \alpha_1 + r - s_1 + \alpha_2 + m_2 - s_2 + (m_1 - r)/2 + 1 \\ &= \alpha_1 + (m_1 + r)/2 + 1 + \alpha_2 + m_2 - s_1 - s_2 \\ &\le \alpha_1 + k_1/2 + 1 + \alpha_2 + k_2/2 - s_1 - s_2 < \alpha_1 + k_1/2 + 1, \end{aligned}$$

since $2m_2 \le k_2$ due to $\pi_2 \in \mathcal{C}_2^2$, and where the last inequality is due to the assumption $\alpha_2 < -k_2/2$.

We now examine R_2 in (66), which is obtained by identifying some of the u variables to the v variables in the first sum in (65). One term of R_2 can be bounded by

$$\sum_{u_p \neq u_q, v_p \neq v_q, \mathbf{u}, \mathbf{v} > \mathbf{0}} C\big(\|\mathbf{u}\| + \|\mathbf{v}_1\| + \|\mathbf{i}\|\big)^{\alpha_1}\big(\|\mathbf{v}_1\| + \|\mathbf{v}_2\|\big)^{\alpha_2},$$

where $\mathbf{u} = (u_1, \ldots, u_{r-s}), \mathbf{v}_1 = (v_1, \ldots, v_s), \mathbf{v}_2 = (v_{s+1}, \ldots, v_{m_2})$ and $\mathbf{i} = (i_1, \ldots, i_{m_1-r})$, where $1 \le s \le (r \wedge m_2)$. By using (48), and then (51) and (49), this term is bounded by

$$\begin{aligned}
&\sum_{\mathbf{u} > \mathbf{0}, \mathbf{v}_1 > \mathbf{0}, \mathbf{v}_2 > \mathbf{0}} C\big(\|\mathbf{u}\| + \|\mathbf{v}_1\| + \|\mathbf{i}\|\big)^{\alpha_1}\big(\|\mathbf{v}_1\| + \|\mathbf{v}_2\|\big)^{\alpha_2} \\
&\quad \le \sum_{\mathbf{v}_1 > \mathbf{0}} C\big(\|\mathbf{v}_1\| + \|\mathbf{i}\|\big)^{\alpha_1 + r - s}\|\mathbf{v}_1\|^{\alpha_2 + m_2 - s} \le C\|\mathbf{i}\|^{\alpha_1 + r - s + (s-1)/2},
\end{aligned}$$

which yields an

$$H^* = \alpha_1 + r - s/2 - 1/2 + (m_1 - r)/2 + 1 = \alpha_1 + (m_1 + r)/2 + 1 - s/2 - 1/2 < \alpha_1 + k_1/2 + 1.$$

So neither R_1 nor R_2 contributes to the limit.

Part 2 of case 5. Suppose now that π and T are *not* as in Part 1. To determine these cases, note that one can always bound $|(S'_T a_\pi)(\mathbf{i})|$ by

$$C\sum_{\mathbf{u} > \mathbf{0}}\big(\|\mathbf{i}_1\| + \|\mathbf{i}_2\| + \|\mathbf{u}_1\| + \|\mathbf{u}_2\|\big)^{\alpha_1}\big(\|\mathbf{i}_1\| + \|\mathbf{i}_3\| + \|\mathbf{u}_1\| + \|\mathbf{u}_3\|\big)^{\alpha_2}, \tag{68}$$

where $\mathbf{i}_j \in \mathbb{Z}_+^{s_j}$, $\mathbf{u}_j \in \mathbb{Z}_+^{t_j}$, $s_j, t_j \ge 0$ and where $s_1 + s_2 + s_3 > 0$ (at least one i variable must remain), and

$$s_1 + s_2 + t_1 + 2t_2 \le k_1, \qquad s_1 + s_3 + t_1 + 2t_3 \le k_2.$$

Thus, the variables in $\mathbf{u}_2$ are at least paired within $a^{(1)}$, and the variables in $\mathbf{u}_3$ are at least paired within $a^{(2)}$.

We note that in Part 1, we had $s_1 = s_3 = t_1 = 0$, and $s_t + 2t_2 = k_1$. Thus, to avoid the situation considered in Part 1, we require

$$\text{if } s_1 = s_3 = t_1 = 0, \qquad \text{then } s_2 + 2t_2 < k_1. \tag{69}$$

As we have dealt with R_1 and R_2 before, by properly applying (48)–(51), the bound in (68) yields

$$H^* < H_1 = \alpha_1/2 + k_1/2 + 1.$$

To check this, we consider the following exhaustive cases:

(a) either $s_1 > 0$, or $s_1 = 0$, $s_2 > 0$, $s_3 > 0$;
(b) $s_1 = s_2 = 0$, $s_3 > 0$;
(c) $s_1 = s_3 = 0$, $s_2 > 0$ but $s_2 + 2t_2 < k_1$.

Note that in case (c), if $s_2 + 2t_2 = k_1$ then $t_1 = 0$, which would contradict (69).

In case (a), for example, if $s_1 > 0$, by applying (48) to the sum over $\mathbf{u}_2$ and $\mathbf{u}_3$, and then (51) on the sum over $\mathbf{u}_1$, we can bound (68) by

$$C\bigl(\|\mathbf{i}_1\| + \|\mathbf{i}_2\|\bigr)^{\alpha_1+t_1/2+t_2}\bigl(\|\mathbf{i}_1\| + \|\mathbf{i}_3\|\bigr)^{\alpha_2+t_1/2+t_3}.$$

This yields

$$\begin{aligned} H^* &= \alpha_1 + \alpha_2 + t_1 + t_2 + t_3 + (s_1 + s_2 + s_3)/2 + 1 \\ &= \alpha_1 + (s_1 + s_2 + t_1 + 2t_2)/2 + 1 + \alpha_2 + (s_3 + t_1 + 2t_3)/2 \\ &\le \alpha_1 + k_1/2 + 1 + \alpha_2 + k_2/2 < \alpha_1 + k_1/2 + 1 = H_1. \end{aligned} \tag{70}$$

In case (b), (68) becomes $C\sum_{\mathbf{u}>\mathbf{0}}(\|\mathbf{u}_1\| + \|\mathbf{u}_2\|)^{\alpha_1}(\|\mathbf{i}_3\| + \|\mathbf{u}_1\| + \|\mathbf{u}_3\|)^{\alpha_2}$ which we can bound by

$$\begin{aligned} &C\sum_{\mathbf{u}_1>\mathbf{0}}\|\mathbf{u}_1\|^{\alpha_1+t_2}\bigl(\|\mathbf{i}_3\| + \|\mathbf{u}_1\|\bigr)^{\alpha_2+t_3} \\ &\quad\le \begin{cases} \|\mathbf{i}_3\|^{\alpha_2+(t_1-1)_+/2+t_3} & \text{if } \alpha_1 + t_1/2 + t_2 < -1/2; \\ \|\mathbf{i}_3\|^{\alpha_1+\alpha_2+t_1+t_2+t_3} & \text{if } -1/2 < \alpha_1 + t_1/2 + t_2 < 0, \end{cases} \end{aligned}$$

where we need to apply first (48), then apply (51) if $t_1 \ge 2$, and finally apply either (49) for the first case or (50) for the second. Note that $\alpha_1 + t_1/2 + t_2 > -1/2$ only if $t_1/2 + t_2 = k_1/2$ since $-k_1/2 - 1/2 < \alpha_1 < -k_1/2$ and $t_1 + 2t_2 \le k_1$. So this yields either an

$$\begin{aligned} H^* &= \alpha_2 + (t_1 - 1)_+/2 + t_3 + s_3/2 + 1 \\ &= \alpha_2 + (s_3 + t_1 + 2t_3)/2 + 1 + (t_1 - 1)_+/2 - t_1/2 \le \alpha_2 + k_2/2 + 1 = H_2 < H_1 \end{aligned} \tag{71}$$

or H^* as in (70).

Similarly, in case (c), (68) is $\sum_{\mathbf{u}>\mathbf{0}} C(\|\mathbf{i}_2\| + \|\mathbf{u}_1\| + \|\mathbf{u}_2\|)^{\alpha_1}(\|\mathbf{u}_1\| + \|\mathbf{u}_3\|)^{\alpha_2}$, which can be bounded by

$$\begin{aligned} &C\sum_{\mathbf{u}_1>\mathbf{0}}\bigl(\|\mathbf{i}_2\| + \|\mathbf{u}_1\|\bigr)^{\alpha_1+t_2}\|\mathbf{u}_1\|^{\alpha_2+t_3} \\ &\quad\le \begin{cases} \|\mathbf{i}_2\|^{\alpha_1+(t_1-1)_+/2+t_2} & \text{if } \alpha_2 + t_1/2 + t_3 < -1/2; \\ \|\mathbf{i}_2\|^{\alpha_1+\alpha_2+t_1+t_2+t_3} & \text{if } -1/2 < \alpha_1 + t_1/2 + t_2 < 0. \end{cases} \end{aligned}$$

So it yields either an

$$\begin{aligned} H^* &= \alpha_1 + (t_1 - 1)_+/2 + t_2 + s_2/2 + 1 \\ &= \alpha_1 + (s_2 + t_1 + 2t_2)/2 + 1 + (t_1 - 1)_+/2 - t_1/2 < \alpha_1 + k_1/2 + 1 = H_1, \end{aligned} \tag{72}$$

or H^* as in (70). To get the strict inequality in (72), we use (69) when $t_1 = 0$, and use $(t_1 - 1)_+/2 < t_1/2$ when $t_1 > 0$. $\square$

Proof of case 6. Same as case 5. □

Proof of case 7. Since $H_1 = H_2$, both factors $a^{(1)}$ and $a^{(2)}$ may contribute to the limit. The proof is similar to case 5, while the other term in the limit arises by exchanging of the role of $a^{(1)}$ and $a^{(2)}$ in the proof of case 5. Note that because $H_1 = H_2$, the equality in "$\leq$" in (71) is attained whenever $t_1 = 0$ and $s_3 + 2t_3 = k_2$, a case which would then be included in the NCLT part of the proof. □

5.3. Proof of Theorem 3.8 the mixed case

The proof is similar to case 5 of Theorem 3.6. We thus only give a sketch.

First, following the same notation as Part 1 of case 5 of Theorem 3.6, we look at the contributing case: the partition π can be split into π_1 and π_2, where since now the factor Y_1' in (30) excludes the diagonals, the first partition π_1 is just $\{\{1\}, \dots, \{m_1\}\}$. This means that the component $\mathbf{u}$ in (64) does not appear, namely, $r = 0$. Hence, instead of (66) one gets

$$a^{(1)}(i_1, \dots, i_{m_1}) \sum_{v_p \neq v_q, \mathbf{v} > \mathbf{0}} a^{(2)}_{\pi_2}(\mathbf{v}) - R(\mathbf{i}), \tag{73}$$

where $i_p \neq i_q$ for $p \neq q$ and the residual term $R(\mathbf{i})$ is as $R_1(\mathbf{i})$ in (66) (there is no R_2 due to absence of $\mathbf{u}$). The first term leads to the noncentral limit $c_2 I_{k_1}(h_{t,1})$ with Hurst index $H_1 = \alpha_1 + k_1/2 + 1$ claimed in Theorem 3.8 by Proposition 2.5. Then treating $R(\mathbf{i})$ in the same way as $R_1(\mathbf{i})$ is treated there, one can show that $R(\mathbf{i})$ leads to terms with Hurst index strictly less than $H_1 = \alpha_1 + k_1/2 + 1$. Since H_1 is used in the normalization, all these terms are negligible.

Next, one follows Part 2 of case 5 of the proof of Theorem 3.6 to show that all other cases of π yield terms with Hurst indices strictly less than $H_1 = \alpha_1 + k_1/2 + 1$. Due to the off-diagonality of Y_1', for the bound (68), we have the following additional restrictions involving the dimensions of the vectors in (68): $t_2 = 0$ ($\mathbf{u}_2$ does not appear), and thus

$$s_1 + s_2 + t_1 = k_1. \tag{74}$$

The argument in the proof of Theorem 3.6 for cases (a) and (c) continue to hold because the quantity H^* continues to be strictly less than H_1. The only case there involving modification is case (b) where $s_1 = s_2 = 0$, $s_3 > 0$, because the original inequality (71) allows $H^* = H_2$ which can be greater than H_1. But now by (74) we have the restriction $t_1 = k_1 \geq 2$. So (71) is now changed to

$$H^* = \alpha_2 + (s_3 + t_1 + 2t_3)/2 + 1 - 1/2 \leq H_2 - 1/2 < 1/2 < H_1.$$

Since $H^* < H_1$, these terms are also negligible. Then the first term of (73) dominates and provides the limit $c_2 I_{k_1}(h_{t,1})$.

Acknowledgements

We would like to thank an anonymous referee for careful reading and constructive suggestions. This work was partially supported by the NSF Grants DMS-10-07616 and DMS-13-09009 at Boston University.

References

[1] Bai, S. and Taqqu, M.S. (2014). Generalized Hermite processes, discrete chaos and limit theorems. *Stochastic Process. Appl.* **124** 1710–1739. MR3163219

[2] Bai, S. and Taqqu, M.S. (2015). Convergence of long-memory discrete kth order Volterra processes. *Stochastic Process. Appl.* **125** 2026–2053. MR3315622

[3] Beran, J., Feng, Y., Ghosh, S. and Kulik, R. (2013). *Long-Memory Processes: Probabilistic Properties and Statistical Methods*. Heidelberg: Springer. MR3075595

[4] Breton, J.-C. (2006). Convergence in variation of the joint laws of multiple Wiener–Itô integrals. *Statist. Probab. Lett.* **76** 1904–1913. MR2271186

[5] Davydov, Yu.A. and Martynova, G.V. (1989). Limit behavior of distributions of multiple stochastic integrals. In *Statistics and Control of Random Processes* (*Russian*) (*Preila*, 1987) 55–57. Moscow: "Nauka". MR1079335

[6] Dobrushin, R.L. and Major, P. (1979). Non-central limit theorems for nonlinear functionals of Gaussian fields. *Z. Wahrsch. Verw. Gebiete* **50** 27–52. MR0550122

[7] Embrechts, P. and Maejima, M. (2002). *Selfsimilar Processes. Princeton Series in Applied Mathematics*. Princeton, NJ: Princeton Univ. Press. MR1920153

[8] Giraitis, L., Koul, H.L. and Surgailis, D. (2012). *Large Sample Inference for Long Memory Processes*. London: Imperial College Press. MR2977317

[9] Hu, Y.Z. and Meyer, P.-A. (1988). Sur les intégrales multiples de Stratonovitch. In *Séminaire de Probabilités, XXII. Lecture Notes in Math.* **1321** 72–81. Berlin: Springer. MR0960509

[10] Koul, H.L., Baillie, R.T. and Surgailis, D. (2004). Regression model fitting with a long memory covariate process. *Econometric Theory* **20** 485–512. MR2061725

[11] Maejima, M. and Tudor, C.A. (2012). Selfsimilar processes with stationary increments in the second Wiener chaos. *Probab. Math. Statist.* **32** 167–186. MR2959876

[12] Nourdin, I. and Peccati, G. (2012). *Normal Approximations with Malliavin Calculus: From Stein's Method to Universality. Cambridge Tracts in Mathematics* **192**. Cambridge: Cambridge Univ. Press. MR2962301

[13] Nourdin, I., Peccati, G. and Reinert, G. (2010). Invariance principles for homogeneous sums: Universality of Gaussian Wiener chaos. *Ann. Probab.* **38** 1947–1985. MR2722791

[14] Nourdin, I. and Poly, G. (2013). Convergence in total variation on Wiener chaos. *Stochastic Process. Appl.* **123** 651–674. MR3003367

[15] Surgailis, D. (1982). Domains of attraction of self-similar multiple integrals. *Litovsk. Mat. Sb.* **22** 185–201. MR0684472

[16] Taqqu, M.S. (1979). Convergence of integrated processes of arbitrary Hermite rank. *Z. Wahrsch. Verw. Gebiete* **50** 53–83. MR0550123

Received March 2014 and revised November 2014