Bernoulli **26**(2), 2020, 1473–1503
https://doi.org/10.3150/19-BEJ1168

Limit theorems for long-memory flows on Wiener chaos

SHUYANG BAI[1] and MURAD S. TAQQU[2]

[1]*Department of Statistics, University of Georgia, Athens, GA 30606, USA. E-mail: bsy9142@uga.edu*
[2]*Department of Mathematics and Statistics, Boston University, Boston, MA 02215, USA.*
E-mail: murad@bu.edu

We consider a long-memory stationary process, defined not through a moving average type structure, but by a flow generated by a measure-preserving transform and by a multiple Wiener–Itô integral. The flow is described using a notion of mixing for infinite-measure spaces introduced by Krickeberg (In *Proc. Fifth Berkeley Sympos. Math. Statist. and Probability (Berkeley, Calif., 1965/66), Vol. II: Contributions to Probability Theory, Part 2* (1967) 431–446 Univ. California Press). Depending on the interplay between the spreading rate of the flow and the order of the multiple integral, one can recover known central or non-central limit theorems, and also obtain joint convergence of multiple integrals of different orders.

Keywords: conservative flows; ergodic theory; Fourth Moment Theorem; long memory; long-range dependence

1. Introduction

A stationary sequence $\{X(n)\}$ is said to have *long memory*, or *long-range dependence*, if a quantity measuring dependence decays slowly as the time lag increases so that some nonstandard scaling behavior arises. Typically, long memory is associated with a power decay of the covariance $\mathrm{Cov}[X(n), X(0)]$, like $n^{\beta-1}$ as $n \to \infty$ with $\beta \in (0, 1)$, so that the sum $\sum_{n=1}^{N} X(n)$ scales in the order $N^{1/2+\beta/2}$ as $N \to \infty$. Note that the magnitude of the fluctuations of the sum $\sum_{n=1}^{N} X(n)$ is larger than the standard order $N^{1/2}$ under weak dependence. The phenomenon of long memory has attracted a lot of attention in probability and statistics. On the probability side, long memory gives rise to various non-standard limit theorems leading to interesting scaling limits; on the statistics side, time series data exhibiting long memory is often found in real-life contexts whose analysis typically requires procedures quite different from the weakly dependent cases. For more information about long memory, we refer to the recent monographs Giraitis et al. [14], Beran et al. [6], Samorodnitsky [43], Pipiras and Taqqu [40] and the references therein.

There are various models which exhibit long memory. To generate long memory, a typical way is to introduce some *moving-average* structure with coefficients decaying slowly, like a power law. For example, one of the most popular long-memory model considered in statistics (e.g., Granger and Joyeux [16]), is the long-memory linear process, given by

$$X(n) = \sum_{j=-\infty}^{\infty} f_{n+j}\epsilon_j,$$

1350-7265

where $\{\epsilon_j\}$ is a centered i.i.d. random sequence, and f_j decays slowly as $j \to \pm\infty$. One may consider a similar model (e.g., Brockwell and Marquardt [9])

$$X(n) = \int_{\mathbb{R}} f(n+x) M(dx), \tag{1}$$

where $M(\cdot)$ is a homogeneous independently scattered random measure (or say a Lévy process), and $f(x)$ is a function that decays slowly as $x \to \pm\infty$.

There has been recently some interest in long-memory models where the power-law decay of dependence is generated in a different manner, one involving infinite ergodic theory (Aaronson [1]). To motivate, we rewrite (1) as

$$X(n) = \int_{\mathbb{R}} f\bigl(T^n x\bigr) M(dx), \tag{2}$$

where $T : \mathbb{R} \to \mathbb{R}$ is the shift operator $x \to x+1$, which preserves the Lebesgue measure. The flow $\{T^n\}$ is totally dissipative on $\mathbb{R}$, since any point leaving a subset of $\mathbb{R}$ will not return to this subset again. Note that in this setting, the long memory of $\{X(n)\}$ is due to a function f which decays slowly on a "large" part of the space. It is not due to the flow $\{T^n\}$ since this flow does not return and thus has no "memory".

We now consider a setup which *reverses* the preceding roles of f and $\{T^n\}$, that is, f will now be supported on a "small" piece of the space, while $\{T^n\}$ will generate the memory using a suitable mode of return behavior. In general, let $(E, \mathcal{E}, \mu)$ be a measure space, where μ is a σ-finite measure on the σ-field $\mathcal{E}$. We make the important assumption:

$$\mu(E) = +\infty. \tag{3}$$

Let $T \colon E \to E$ be a measurable transformation which preserves the measure μ, that is, $\mu T^{-1} = \mu$. Consider

$$X(n) = \int_E f\bigl(T^n x\bigr) M(dx), \tag{4}$$

where M is an independently scattered random measure with control measure μ, and where f is a suitable function. We do not have a moving average structure as in (1). To generate an infinitely-lasting memory, we require $\{T^n\}$ to be conservative, that is, the total return count to a set $A \in \mathcal{E}$ with $\mu(A) > 0$, is infinite, namely

$$\sum_{n=1}^{\infty} 1_A\bigl(T^n x\bigr) = \infty$$

for a.e. $x \in A$. On the other hand, we also want the memory to vanish as $n \to \infty$ eventually. This is where the infinite measure assumption (3) plays a key role. To see this, suppose a starting point x is chosen at random using the probability measure $P_A(\cdot) := \mu(\cdot \cap A)/\mu(A)$. Consider the first return time $\tau_A(x) = \inf\{n \in \mathbb{Z}_+,\ x \in A,\ T^n x \in A\}$ and assume that T is ergodic. Then by Kac's formula (Item 1.1.5 of Aaronson [1]), the expected return time $\mathbb{E}\tau_A$ is infinite. In a probabilistic language, the flow $\{T^n\}$ is null-recurrent, namely, while the returns happen infinitely often, they are so sparse that the expected return time is infinite.

Concrete examples of the aforementioned dynamical system $(E, \mathcal{E}, \mu, T)$ can be found in Example 4.1 and 4.2 below. In Example 4.1, the space is $E = [0, 1]$ and μ has an explicit density (22) with respect to the Lebesgue measure. The density has a power-law behavior around the origin, which is the fixed point of T. In Example 4.2, the space E is the path space of a null-recurrent Markov chain and T is the shift operator. The transition probability of the Markov chain has a power-law decay.

We also assume that the function f is supported on a set $G \in \mathcal{E}$ with $0 < \mu(G) < \infty$. The set G has a small size in contrast to $\mu(E) = \infty$. In addition, the set G needs to satisfy some condition to ensure the desired specific long memory property for $\{X(n)\}$. For example, G can be the *Darling-Kac set* (p. 123 of Aaronson [1]) or the *uniform set* considered in Owada and Samorodnitsky [35], or the *uniformly returnning set* considered in Owada [34]. In all these setups, the strength of the memory is characterized by a parameter $\beta \in (0, 1)$. The precise descriptions of these sets are technical, while loosely speaking, they imply that the return count $\sum_{n=1}^{N} 1_A(T^n x)$ averaged over $x \in G$ diverges to infinity like N^β as $N \to \infty$.

Some recent studies considered limit theorems for $\{X(n)\}$ in (4) where the random measure $M(\cdot)$ has infinite variance. Owada and Samorodnitsky [35] and Jung et al. [20] established limit theorems for the sum $\sum_{n=1}^{N} X(n)$. Owada and Samorodnitsky [36], Lacaux and Samorodnitsky [24] and Samorodnitsky and Wang [44] considered limit theorems for the extreme $\max_{n=1}^{N} X(n)$. Owada [34] considered a limit theorem for the sample autocovariance of $X(n)$. See also Chapter 9 of Samorodnitsky [43] for some introductory discussion on these topics.

So far the studies have focused on limit theorems for the linear integral model in (4). We consider here a nonlinear analog involving a multiple stochastic integral:

$$X(n) = \int_{E^k}' f\big(T^n x_1, \dots, T^n x_k\big) M(dx_1) \dots M(dx_k), \tag{5}$$

where $M(\cdot)$ is a Gaussian measure, and the prime $'$ on the integral sign excludes integration on the diagonals $x_p = x_q$, $p \neq q$. In this case, $X(n)$ belongs to the so-called Wiener chaos of order k (see, e.g., Peccati and Taqqu [37]). This choice makes sense in view of the fruitful line of research relating long memory to multiple stochastic integrals, starting from Rosenblatt [41], Taqqu [46,47], Dobrushin and Major [11], etc.

The goal is to recover limit theorems for the sum $\sum_{n=1}^{N} X_n$ using the *mixing* framework of Krickeberg [23]. We shall show that depending on the interplay between the parameter β that characterizes the "spreading rate" of T, and the order of integrals m in (5), we can recover limit theorems involving multiple Wiener–Itô integrals with long memory, namely, obtain as a scaling limit either Brownian motion or a Hermite process. A joint convergence involving both Brownian motion and Hermite processes is also established. We mention that a recent work Bai et al. [3] considered a limit theorem for the sum $\sum_{n=1}^{N} X_n$ when the random measure $M(\cdot)$ in (5) has infinite variance.

The argument of this paper is developed using three main steps:

Step 1: We develop in Section 2 the relevant ergodic-theoretic framework and use it to relate the kernel f in (5) to simple functions with special structures. This step is of special interest: a mixing condition on an infinite-measure space due to Krickeberg [23] is introduced to generate data with long memory.

Step 2: We show in Lemma 5.4 that the desired limit theorems hold if they hold for these special simple functions.

Step 3: We show in Proposition 5.5 that the limit theorems do hold for these special simple functions.

The paper is organized as follows: In Section 2, we introduce the mixing framework of Krickeberg, where we define the space $(E, \mathcal{E}, \mu)$ and the transform $T : E \to E$. The main results involving the multiple integrals are stated in Section 3. Some examples of Krickeberg mixing are given in Section 4. The proofs are found in Section 5.

2. Krickeberg mixing

In this section, we give some background to the main results of Section 3, focusing mainly on the mixing framework in Krickeberg [23]. The section is organized as follows. We start with some motivating discussion and then state the key assumption, namely Assumption 2.1, which involves the mixing condition (7) stated below. Equivalent formulations of Assumption 2.1 are given in Proposition 2.5. The extension to the product measure space is given in Proposition 2.6. Lemma 2.9 then provides the key approximation tool used in the proof of the main results in Section 3.

2.1. Motivation

Let $(E, \mathcal{E}, \mu)$ be a measure space, where μ a measure on the σ-field $\mathcal{E}$. Let T: $E \to E$ be a measurable mapping which preserves the measure μ, namely, $\mu T^{-1} = \mu$. When $\mu(E) < \infty$, the transform T is said to be *mixing* if

$$\lim_{n\to\infty} \mu\bigl(A \cap T^{-n}B\bigr) = \mu(A)\mu(B), \quad \text{for all } A, B \in \mathcal{E}. \tag{6}$$

When $\mu(E) = +\infty$, however, even the following weaker requirement: there exists a fixed sequence $\rho_n \in (0, \infty)$ termed *spreading rate*, typically tending to infinity as $n \to \infty$, such that

$$\lim_{n\to\infty} \rho_n \mu\bigl(A \cap T^{-n}B\bigr) = \mu(A)\mu(B) \tag{7}$$

holds for all $A, B \in \mathcal{E}$, is usually too stringent. This is because there are sets A and B for which it does not hold. This is due to the existence of the *weakly wandering sets* when T is ergodic (i.e., $T^{-1}A = A \Rightarrow \mu(A) = 0$ or $\mu(A^c) = 0$). A set $A \in E$ is said to be *weakly wandering* if there is a subsequence $\{n_k\}$ with $0 = n_0 < n_1 < n_2 < \cdots$ such that $T^{-n_k}A$ are disjoint, $k = 0, 1, 2, \ldots$. The existence of such a set A with positive measure invalidates (7). Indeed, take $B = A$, and we have $\mu(A \cap T^{-n_k}B) = \mu(A \cap T^{-n_k}A) = \mu(\varnothing) = 0$, $k = 1, 2, \ldots$. Hence, (7) cannot hold. By Hajian and Kakutani [17] (see also Eigen et al. [13]), a weakly wandering set with positive measure always exists when T is ergodic and $\mu(E) = +\infty$.

On the other hand, extending an early example of Hopf [18], Krickeberg [23] formulated the mixing relation (7) by introducing a topological structure on $(E, \mathcal{E})$. This bears a resemblance to the necessity of introducing a topological context when developing the theory of weak convergence of measures. Krickeberg [23] then restricted the relation (7) to sets A and B whose boundaries have zero μ measures.

2.2. Basic assumptions

We will use a variant of the setup of Krickeberg [23]. For convenience, we work with a topological assumption, that is, Polish space, which is stronger, but more common, than that in Krickeberg [23], while still retaining all the concrete examples.

Recall that a Polish space is a topological space whose topology can be induced by a complete (Cauchy sequence converges) and separable (contains a countable dense subset) metric. Below for a subset A of a Polish space, $\bar{A}$ denotes its closure, $\mathring{A}$ denotes its interior, and $\partial A = \bar{A} \setminus \mathring{A}$ denotes its boundary. A set A is said to be a *μ-continuity set* if $\mu(\partial A) = 0$, or equivalently, 1_A is μ-a.e. continuous, that is, the set of discontinuity points of 1_A has zero μ measure. The notation $a_n \sim b_n$ means $\lim_{n\to\infty} a_n/b_n = 1$ in what follows.

Assumption 2.1. Let $(E, \mathcal{E}, \mu)$ be a measure space, where μ is an atomless σ-finite measure on the σ-field $\mathcal{E}$ with $\mu(E) = +\infty$. Let $T\colon E \to E$ be a measurable mapping which preserves the measure μ, namely, $\mu T^{-1} = \mu$. Suppose that there exists $G \in \mathcal{E}$ with $\mu(G) \in (0, \infty)$, so that G is a Polish space with $\mathcal{E}_G := \mathcal{E} \cap G$ being the associated Borel σ-field, such that the mixing relation (7) holds for any μ-continuity sets $A, B \in \mathcal{E}_G$ with a positive sequence

$$\rho_n \sim c_T n^{1-\beta}, \quad \beta \in (0, 1), \tag{8}$$

where $c_T > 0$ is a constant which does not depend on A or B.

A number of remarks and consequences of these assumptions are given below.

Remark 2.2. In Assumption 2.1, relation (7) involves μ-continuity sets A and B in $\mathcal{E}_G$. Note that the μ-continuity is with respect to the Polish topology of the subspace G, and strictly speaking, we are referring to μ_G-continuity with μ_G being the restriction of μ to $\mathcal{E}_G$. To simplify notation, we shall still use μ in place of μ_G. Note in particular that under Assumption 2.1, the relation (7) holds if $A = B = G$ since the whole space G is both open and closed and is thus always a μ-continuity set.

Remark 2.3. Some examples satisfying Assumption 2.1 will be provided in Section 4. See also Krickeberg [23], Thaler [48], Kesseböhmer and Slassi [22] Gouëzel [15] and Melbourne and Terhesiu [26].

Remark 2.4. A more general case where c_T in (8) is replaced by a slowly varying function (see Bingham et al. [7]), for example, a logarithm, can be considered as well. We focus on the case (8) for simplicity and also because it is typically found in examples.

Throughout the paper we shall use the following convention: *any function f defined on a subset (e.g., G) will be extended to the full space (e.g., E) by assuming zero value on the complement, whenever this is necessary to make sense of a statement.*

The next proposition provides some equivalent descriptions of the mixing relation (7).

Proposition 2.5. *Let G, $\mathcal{E}_G$ and (ρ_n) be as in Assumption 2.1, where G is Polish with Borel σ-field $\mathcal{E}_G$. Then the following three statements are equivalent*:

(1) *The mixing relation (7) holds for any μ-continuity sets $A, B \in \mathcal{E}_G$;*
(2) *For any bounded and μ-a.e. continuous functions f_1, f_2 on G, we have*

$$\lim_{n\to\infty} \rho_n \int_E f_1 \cdot (f_2 \circ T^n)\, d\mu = \mu(f_1)\mu(f_2); \tag{9}$$

(3) *There exists a collection $\mathcal{C} \subset \mathcal{E}_G$ satisfying the following properties*:
   (a) *$\mathcal{C}$ is a π-system (namely, $A, B \in \mathcal{C} \Rightarrow A \cap B \in \mathcal{C}$) containing G.*
   (b) *$\mathcal{C}$ generates the Polish topology of G in the sense that for any open $U \subset G$ and any $x \in U$, there exists $A \in \mathcal{C}$ so that $x \in \mathring{A} \subset A \subset U$.*
   (c) *Any $A \in \mathcal{C}$ is a μ-continuity set.*
   (d) *For any $A, B \in \mathcal{C}$, the mixing relation (7) holds with the spreading rate (ρ_n).*

The proof is similar to the arguments sketched on P.435 of Krickeberg [23]. We include a proof in Section A.1 below for the sake of completeness.

The next step is to consider a product space. Recall that G is Polish. Let $G^k = G \times \cdots \times G$ be the product space associated with the product topology, which is also a Polish space. Let $\mathcal{E}_G^k$ be the product σ-field of $\mathcal{E}_G$, which is also the Borel σ-field generated by the product subspace topology on G^k (e.g., Lemma 6.4.2 of Bogachev [8]). Let μ^k be the product measure defined on $\mathcal{E}^k$. Define (Cartesian) product transformation

$$T_k(x_1, \dots, x_k) := (Tx_1, \dots, Tx_k). \tag{10}$$

Proposition 2.6. *If Assumption 2.1 holds for $(E, \mathcal{E}, \mu, T, G, \rho_n)$, then it also holds if the former is replaced by $(E^k, \mathcal{E}^k, \mu^k, T_k, G^k, \rho_n^k)$.*

The proof of Proposition 2.6 is included in Section A.2 below.

In view of Propositions 2.6 and 2.5(2), under Assumption 2.1 we have the following product space version of (9):

$$\lim_{n\to\infty} \rho_n^k \int_{E^k} f_1 \cdot (f_2 \circ T_k^n)\, d\mu^k = \mu^k(f_1)\mu^k(f_2), \tag{11}$$

where f_1, f_2 are bounded μ^k-a.e. continuous functions on G^k.

2.3. Approximations

We now introduce a class of approximation functions. It is used in the proof of Proposition 2.5 in Section A.1 below, and also in the proof of the key reduction Lemma 5.4 in Section 5.1.

Definition 2.7. A function $g : G^k \to \mathbb{R}$ is said to be an *elementary function*, if it is a finite linear combination of indicators of k-products of μ-continuity sets in $\mathcal{E}_G$, namely,

$$g(x_1, \dots, x_k) = \sum_{j=1}^{J} a_j 1_{A_{1,j} \times \cdots \times A_{k,j}}(x_1, \dots, x_k) = \sum_{j=1}^{J} a_j 1_{A_{1,j}}(x_1) \times \cdots \times 1_{A_{k,j}}(x_k), \tag{12}$$

where $J \in \mathbb{Z}_+$, a_j's are some real constants and $A_{i,j} \in \mathcal{E}_G$ with $\mu(\partial A_{i,j}) = 0$. A set $A \in \mathcal{E}_G^k$ is said to be an *elementary set*, if 1_A is an elementary function.

Remark 2.8. Any elementary function is μ^k-a.e. continuous in view of

$$\partial(B_1 \times \cdots \times B_k) = (\partial B_1 \times \bar{B}_2 \times \cdots \times \bar{B}_k) \cup \cdots \cup (\bar{B}_1 \times \cdots \times \bar{B}_{k-1} \times \partial B_k). \tag{13}$$

Moreover, the class of elementary functions forms a vector space, and in addition, if f and g are elementary, then so is $f \cdot g$. Using this one can check that if $A, B \in \mathcal{E}_G^k$ are elementary sets, then so are $A \cap B$, $A \setminus B$ and $A \cup B$.

Lemma 2.9. *Let f be a bounded μ-a.e. continuous function on G^k. Then for any $\epsilon > 0$, there exist elementary functions g_1, g_2 in the sense of Definition 2.7, such that $g_1 \le f \le g_2$ and $|\mu^k(f) - \mu^k(g_i)| < \epsilon$, $i = 1, 2$.*

The proof of Lemma 2.9 is included in Section A.3 below.

For a real-valued function f defined on a space E, the notation $\|f\|_\infty = \sup\{|f(x)| : x \in E\}$ denotes its supremum norm. The proof of Lemma 2.9 above also yields the following consequence.

Corollary 2.10. *Given a bounded and μ-a.e. continuous f on G^k, for any $\epsilon > 0$, there exists μ^k-continuity set $G_\epsilon \in \mathcal{E}_G^k$, and an elementary function f_ϵ on G^k, such that*

$$\big\|(f - f_\epsilon)1_{G_\epsilon}\big\|_\infty < \epsilon, \qquad \|f_\epsilon\|_\infty \le \|f\|_\infty, \quad \text{and} \quad \mu^k\big(G^k \setminus G_\epsilon\big) < \epsilon. \tag{14}$$

The proof of Corollary 2.10 can be found in Section A.4 below.

3. Limit theorems

We start with a brief introduction to multiple Wiener–Itô integrals. Let $(E, \mathcal{E}, \mu)$ be the atomless σ-finite measure space. Let $W(\cdot)$ be a Gaussian random measure on $(E, \mathcal{E})$ with control measure μ, which satisfies

$$\mathbb{E}W(A)W(B) = \mu(A \cap B)$$

for any $A, B \in \mathcal{E}$. Let $k, p, q \in \mathbb{Z}_+$. For a function $f \in L^2(\mu^k) = L^2(E^k, \mathcal{E}^k, \mu^k)$, the space of square-integrable real-valued functions defined on the product measure space $(E^k, \mathcal{E}^k, \mu^k)$, one can define the multiple Wiener–Itô integral

$$I_k(f) = \int_{E^k}' f(x_1, \dots, x_k) W(dx_1) \dots W(dx_k),$$

where the prime $'$ indicates the exclusion of the diagonals $x_p = x_q$, $p \neq q$. Due to the off-diagonal nature of the multiple integral, we have $\mathbb{E} I_k(f) = 0$. In addition, one can symmetrize f without changing $I_k(f)$, and hence the symmetry of f is often assumed without loss of generality. If symmetric $f \in L^2(\mu^p)$, $g \in L^2(\mu^q)$, then the covariance

$$\mathbb{E} I_p(f) I_q(g) = \begin{cases} k!\langle f, g\rangle_{L^2(\mu^k)}, & \text{if } p = q = k; \\ 0 & \text{if } p \neq q. \end{cases} \tag{15}$$

The collection of random variables $\{I_k(f) : f \in L^2(\mu^k)\}$ forms an L^2 subspace and is called the k-th Wiener chaos with respect to $W(\cdot)$. For more details about multiple Wiener–Itô integrals, we refer the reader to Major [25] and Peccati and Taqqu [37].

In the case where $E = \mathbb{R}$, $\mathcal{E}$ is Borel and μ is Lebesgue, one can consider the following process defined by a multiple Wiener–Itô integral (Dobrushin [10], Taqqu [47]):

$$Z_{k,H}(t) = c_{k,H} \int_{\mathbb{R}^k}' \int_0^t \prod_{i=1}^k (s - x_i)_+^{(H-1)/k-1/2} \, ds \, W(dx_1) \dots W(dx_k), \tag{16}$$

where $c_{k,H}$ is a constant to ensure $\mathrm{Var}[Z_{k,H}(1)] = 1$. We call $Z_{k,H}(t)$ the standard Hermite process. The process $Z_{k,H}(t)$ also admits other representations (Pipiras and Taqqu [39]).

From now on, we adopt the setup in Section 2, in particular, we shall suppose Assumption 2.1, which involves the measure-preserving dynamic system $(E, \mathcal{E}, \mu, T)$ and the finite-measure Polish subspace G. Let f be a symmetric bounded μ^k-a.e. continuous function on G^k (following the convention made in Section 2, it is extended to $E^k \setminus G^k$ by assuming zero value). Then by the measure-preserving property of T, we have $f \circ T_k^n \in L^2(\mu^k)$, $n \in \mathbb{Z}_+$, where T_k is as in (10). Then one can define a strictly stationary sequence on the k-th Wiener chaos as follows:

$$X(n) = I_k\bigl(f \circ T_k^n\bigr) = \int_{E^k}' f\bigl(T^n x_1, \dots, T^n x_k\bigr) W(dx_1) \dots W(dx_k), \tag{17}$$

where $W(\cdot)$ has control measure μ. The stationarity of $(X(n))_{n \geq 1}$ is shown in Lemma 5.1 below.

The second-order structure of $X(n)$ can be readily clarified. In particular, $\mathbb{E} X(n) = 0$, and by (15),

$$\begin{aligned} \mathbb{E} X(n) X(0) &= k! \int_{G^k} f\bigl(T^n x_1, \dots, T^n x_k\bigr) f(x_1, \dots, x_k) \mu(dx_1) \dots \mu(dx_k) \\ &\sim k! \mu^k(f)^2 \rho_n^{-k} \sim k! c_T^{-k} \mu^k(f)^2 n^{k(\beta-1)}, \end{aligned} \tag{18}$$

where the the first asymptotic equivalence in (18) follows from Proposition 2.6, and the second follows from (8). The goal is to establish limit theorems of the form

$$S_N(t) = \frac{1}{A(N)} \sum_{n=1}^{[Nt]} X(n) \Rightarrow Z(t), \quad t \in [0,1], \tag{19}$$

as $N \to \infty$, where $\Rightarrow$ stands for weak convergence in $D[0,1]$ under the uniform metric (our limits admit continuous sample paths), and $[x]$ stands for the integer part of x.

In view of (18), one sees that $\sum_n \mathbb{E}X(n)X(0)$ is summable if and only if $k(1-\beta) > 1$. The summability of the covariance is typically viewed as an indication of short memory, whereas the absence of summability is typically viewed as an indication of long memory. This explains the case division in our main result Theorem 3.1 below.

Theorem 3.1. *Under Assumption 2.1, suppose that f in (17) is a symmetric bounded and μ-a.e. continuous function on G^k (setting $f = 0$ on $E^k \setminus G^k$), where G is as in Assumption 2.1. Then we have the following three results:*

(a) *If $k(1-\beta) > 1$ and $\sigma^2 = \sum_n \mathbb{E}X(n)X(0) \neq 0$, then (19) holds with*

$$Z(t) = B(t), \qquad A(N) = \sigma N^{1/2},$$

where $B(t)$ is the standard Brownian motion.

(b) *If $k(1-\beta) = 1$ and $\mu^k(f) = \int_{G^k} f \, d\mu^k \neq 0$, then (19) holds with*

$$Z(t) = B(t), \qquad A(N) = 2^{-1/2} \mu^k(f) c_T^{-k/2} \bigl(k! N \ln(N)\bigr)^{1/2},$$

where $B(t)$ is the standard Brownian motion.

(c) *If $0 < k(1-\beta) < 1$, and $\mu^k(f) = \int_{G^k} f \, d\mu^k \neq 0$ then (19) holds with*

$$Z(t) = Z_{k,H}(t), \qquad A(N) = \mu^k(f) c_T^{-k/2} \bigl(k! H(2H-1)\bigr)^{1/2} N^H,$$

where $Z_{k,H}(t)$ is the standard Hermite process in (16) with Hurst index $H = 1 - k(1-\beta)/2$.

Theorem 3.1 is proved in Section 5.2.

Remark 3.2. Note that in case (c), when $k = 1$, one has $H > 1/2$ if $\beta > 0$, and $H < 1$ if $\beta < 1$. In fact, $2H - 2 = \beta - 1$. More generally, for any $k \geq 1$, $0 < k(1-\beta) < 1$ corresponds to $k(2H - 2) + 1 > 0$, which corresponds to a classical condition involving Hermite rank for convergence to the Hermite processes $Z_{k,H}(t)$ (Taqqu [47]).

Remark 3.3. As pointed out by an anonymous referee, it is possible to establish the convergence rate of the marginal distribution of $S_N(1)$ to a standard normal distribution in cases (a) and (b) of Theorem 3.1. See Nourdin and Peccati [30]. This involves an analysis of the third and the fourth moments of $S_N(1)$, which will be considered in a future study.

We can also obtain a joint convergence involving all the cases in Theorem 3.1. In particular, let

$$X_j(n)=\int'_{E^{k_j}} f_j\big(T^n x_1,\dots,T^n x_{k_j}\big)W(dx_1)\dots W(dx_{k_j}),\quad j=1,\dots,J. \tag{20}$$

In (20), although possibly different orders k_j's and different symmetric bounded a.e. continuous functions f_j's on G^{k_j} are involved in different X_j's, we have the same underlying measure space $(E,\mathcal{E},\mu)$, the same transformation T (with the same values of β), the same Gaussian random measure $W(\cdot)$ and the same finite-measure subspace G.

Theorem 3.4. *Suppose $S_{N,j}(t)$ is as in (19) corresponding to $X_j(n)$ with the normalization as given in Theorem 3.1. Assume that*

$$J_1=\{j:(1-\beta)k_j>1\},\qquad J_2=\{j:(1-\beta)k_j=1\}\quad\text{and}\quad J_3=\{j:0<(1-\beta)k_j<1\}$$

(J_2 and J_3 can be empty), and denote the vector process $\mathbf{S}_{N,J_s}(t)=(S_{N,j}(t),\,j\in J_s)$, $s=1,2,3$. Then we have the joint convergence in $D[0,1]^{|J_1|+|J_2|+|J_3|}$,

$$\big(\mathbf{S}_{N,J_1}(t),\mathbf{S}_{N,J_2}(t),\mathbf{S}_{N,J_3}(t)\big)\quad\Rightarrow\quad\big(\mathbf{B}_{J_1}(t),\mathbf{B}_{J_2}(t),\mathbf{Z}_{J_3}(t)\big),$$

where the three limit components $\mathbf{B}_{J_1}(t)$, $\mathbf{B}_{J_2}(t)$, $\mathbf{Z}_{J_3}(t)$ are independent. $\mathbf{B}_{J_1}(t)$ and $\mathbf{B}_{J_2}(t)$ are each vector Brownian motions with the covariance between two components $i,j\in J_1$ or $i,j\in J_2$ given by

$$\mathbb{E}B_i(s)B_j(t)=(s\wedge t)1_{\{k_i=k_j\}}, \tag{21}$$

and $\mathbf{Z}_{J_3}(t)$ is a vector of Hermite processes in (16) with orders k_j and $H_j=1-k_j(1-\beta)/2$, $j\in J_3$, which are defined by the same Gaussian random measure $W(\cdot)$.

Theorem 3.4 is proved in Section 5.5.

Remark 3.5. Note that because of the multiplicative constants in the normalization $A(N)$, every component of the limit has variance 1 at $t=1$. The covariance structure (21) implies that within the case (a) or within case (b), the limit Brownian motions are identical for integrals of the same order, and independent between integrals of different orders.

4. Examples

We first present an example from Thaler [48], p. 108, which satisfies Assumption 2.1 in Section 2.

Example 4.1. Let $(E=[0,1],\mathcal{E}=\mathcal{B})$, where $\mathcal{B}$ denotes the Borel σ-field. Fix a number $p>1$. Let

$$h_p(x)=\frac{1}{x^p}+\frac{1}{(1+x)^p},\quad x\in(0,1], \tag{22}$$

and define a measure μ on $\mathcal{E}$ by

$$\mu(dx) = h_p(x)\,dx.$$

The measure μ has thus a power behavior around the origin. Since the density $h_p(x) > 0$, the measure μ is equivalent to the Lebesgue measure. On the other hand, we have $\mu(E) = \int_0^1 h_p(x)\,dx = \infty$ since $p > 1$ so that the integrability of $h_p(x)$ breaks down near $x = 0$.

Now we introduce the transformation T. The particular choice of μ will ensure that T is measure preserving. Indeed, consider the following function defined on [0, 1]:

$$g_p(x) = x\left(1 + \left(\frac{x}{1+x}\right)^{p-1} - x^{p-1}\right)^{1/(1-p)},$$

which is related to the density function h_p through the differential equation:

$$\frac{d}{dx} g_p(x) = g_p(x)^p h_p(x), \quad x \in (0, 1). \tag{23}$$

Note that g_p maps [0, 1] monotonically to [0, 2] since $\frac{d}{dx} g_p$ is positive on $(0, 1)$. Define now the transformation $T = T_p$ on [0, 1] by

$$T_p(x) := g_p(x) \pmod 1. \tag{24}$$

The density $h_p(x)$ after the transformation (24), because of mod 1 in (24), is

$$\left[h_p\bigl(g_p^{-1}(x)\bigr)\frac{d}{dx} g_p^{-1}(x)\right] + \left[h_p\bigl(g_p^{-1}(1+x)\bigr)\frac{d}{dx} g_p^{-1}(1+x)\right]. \tag{25}$$

Since $\frac{d}{dx} g_p^{-1}(x) = 1/g_p'(g_p^{-1}(x))$, the relations (22) and (23) imply that the expression (25), which involves a change of variable, is exactly equal to $h_p(x)$. Hence, the density after the transformation is still $h_p(x)$, which means that the transformation T_p is μ-preserving.

To understand the nature of the dynamics of the system, note that the point $x = 0$ is known as an *indifferent fixed point* of the system, namely, $T_p(0) = g_p(0) = 0$ and $T_p'(0+) = g_p'(0+) = 1$. Hence starting at a point near $x = 0$, the trajectory of $\{T^n\}$ tends to trapped near $x = 0$ for a while. So the successive visits to a set away from $x = 0$ tends to be sparse (see Figure 1).

Now choose $G = [\epsilon, 1]$, for some $\epsilon \in (0, 1)$. By Corollary 1 on p. 115 of Thaler [48] and our Proposition 2.5, Assumption 2.1 in Section 2 holds with

$$\beta = p^{-1} \in (0, 1).$$

Therefore under this setup, any symmetric Riemann integrable function f on $G^k = [\epsilon, 1]^k$ is bounded and λ^k-a.e. (or equivalently μ^k-a.e.) continuous. Hence, such a function satisfies the condition in Theorem 3.1.

Finally, we mention that the explicit example presented above falls within the class of *dynamical systems* satisfying assumptions (i)–(iv) in Thaler [48]. In addition, such a class of systems on [0, 1] further belongs to a well-studied class called AFN systems (Zweimüller [49,50]), which has multiple indifferent fixed points in general. See Gouëzel [15] and Melbourne and Terhesiu [26]

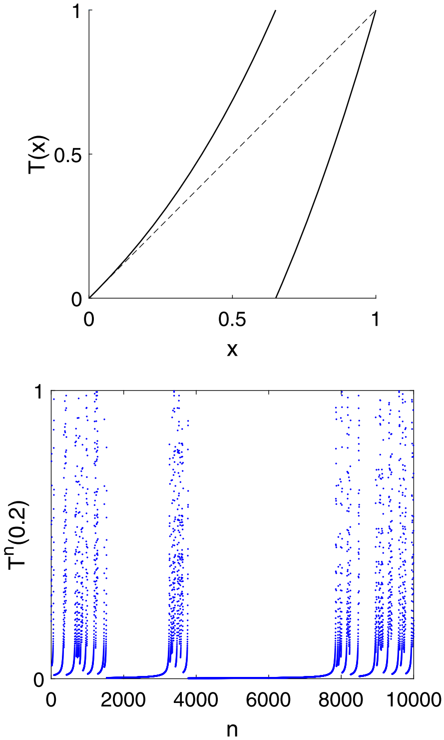

Figure 1. Example 4 with $\beta = 0.75$. Top: graph of x versus $T(x)$; Bottom: trajectory of $\{T^n(0.2)\}$.

for sufficient conditions for an AFN system to satisfy Assumption 2.1. For a general AFN system, G may be chosen to be a disjoint union of closed intervals which are located away from the indifferent fixed points, and f in Theorem 3.1 again can be chosen as a Riemann integrable function on G^k which is a union of hyperrectangles in $[0, 1]^k$.

The next example concerns Markov chains. We refer to, for example, Durrett [12] Chapter 6 for all the common notions and facts regarding Markov chains that we will use below.

Example 4.2. Consider $E = S^{\mathbb{N}}$, the space of S-valued functions defined on $\mathbb{N} = \{0, 1, 2, \ldots\}$, where S is an infinite countable set. We equip S with the discrete topology. Then the countable product space E is a Polish space known as the Baire space, where a topological base of E can be formed by cylinder sets each of the form

$$A = \{x \in S^{\mathbb{N}} : x(n_1) = i_1, \ldots, x(n_m) = i_m\}, \quad 0 \le n_1 < \cdots < n_m, i_l \in S, m \in \mathbb{N}. \tag{26}$$

See Moschovakis [27], Section 1A. The Borel σ-field $\mathcal{E}$ on E thus coincides with the cylindrical σ-field on $S^{\mathbb{N}}$. Define the shift operator $T : E \to E$, $T(x(n))_{n\in\mathbb{N}} = (x(n+1))_{n\in\mathbb{N}}$.

Now suppose that we have an aperiodic, irreducible and null-recurrent Markov chain $X = (X_n)_{n\geq 0}$ with transition probabilities $(p(i,j))_{i,j\in S}$. Let π be the unique invariant measure of the chain, so that $\sum_{i\in S}\pi(i)p(i,j)=\pi(j)$, $\pi(j)>0$ for all $j\in S$ and satisfies the normalization condition $\pi(o)=1$, where $o\in S$ is a distinguished state. The null-recurrence assumption entails that $\pi(S)=\sum_{i\in S}\pi(i)=\infty$.

Now we define a measure on $\mathcal{E}$ as:

$$\mu(\cdot)=\sum_{i\in S}\pi(i)P^i(\cdot),$$

where for each i, $P^i(\cdot)$ is a probability measure defined on $\mathcal{E}$ of a Markov chain starting at i at time zero. Then μ is invariant with respect to the shift operator T. Note that $\mu(E)=\infty$ since $\sum_{i\in S}\pi(i)P^i(E)=\sum_{i\in S}\pi(i)=\pi(S)=\infty$.

The assumptions on the chain imply that $\mu(\{x\})=0$ for any $x\in E$, namely, μ has no atoms. Indeed, suppose without loss of generality that $x(0)=o$. By recurrence, one can assume that the path x visits o infinitely often. Let $p(k)=P(X_k=o,\ X_i\neq o,\ i=1,\dots,k-1|X_0=o)$. Note that $\sum_{k=1}^{\infty}p(k)=1$ by recurrence, and $\sup_{k\geq 1}p(k)<1$ since the chain is irreducible and aperiodic. Let the successive visit times of o by x be $k_1, k_1+k_2, k_1+k_2+k_3,\dots$, where $k_i, i\in\mathbb{Z}_+$. Then using the Markov property,

$$\begin{aligned}\mu\bigl(\{x\}\bigr)&\leq P^o(X_1\neq o,\dots,X_{k_1-1}\neq o, X_{k_1}=o, X_{k_1+1}\neq o,\dots X_{k_1+k_2-1}\neq o, X_{k_1+k_2}=o,\dots)\\&=\prod_{i=1}^{\infty}p(k_i)=0.\end{aligned}$$

Now let $G=\{x\in E: x(0)=o\}$, which as a subspace of E, is again Polish with Borel σ-field $\mathcal{E}_G$. Let

$$\mathcal{C}=\{A\cap G:\ A \text{ is a cylinder set as in (26)}\}\cup\{\varnothing\}.$$

We shall show that the Assumption 2.1 holds by using the paradigm in Proposition 2.5(3). Again in view of Moschovakis [27], Section 1A, the π-system $\mathcal{C}$ is a topological base for G. So (a) and (b) of Proposition 2.5(3) hold, since each set in $\mathcal{C}$ is open. In addition, in view of the discrete topology, each set in $\mathcal{C}$ is closed as well. Hence the boundaries of sets in $\mathcal{C}$ are empty. So (c) in Proposition 2.5(3) holds. For (d), let $A,B\in\mathcal{C}$ be respectively $A=\{x\in S^{\mathbb{N}}: x(0)=o,\ x(n_1)=i_1,\dots,\ x(n_r)=i_r\}$ and $B=\{x\in S^{\mathbb{N}}: x(0)=o,\ x(m_1)=j_1,\dots,\ x(m_s)=j_s\}$, where $0<n_1<\cdots<n_r$, $0<m_1<\cdots<m_s$, $i_\ell,\ j_\ell\in S$. Since $x(0)=o$, the cylinder sets A and B are in $\mathcal{C}$. When n is large enough so that $n>n_r$, we have by the Markov property that

$$\begin{aligned}\mu\bigl(A\cap T^{-n}B\bigr)&=p^{(n_1)}(o,i_1)\left[\prod_{\ell=1}^{r-1}p^{(n_{\ell+1}-n_\ell)}(i_\ell,i_{\ell+1})\right]\\&\quad\times p^{(n-n_r)}(i_r,o)p^{(m_1)}(o,j_1)\left[\prod_{\ell=1}^{s-1}p^{(m_{\ell+1}-m_\ell)}(j_\ell,j_{\ell+1})\right]\\&=\mu(A)p^{(n-n_r)}(i_r,o)\mu(B),\end{aligned}\tag{27}$$

where $p^{(k)}(\cdot,\cdot)$ is the k-step transition probability. Let us now focus on the factor $p^{(n-n_r)}(i_r,o)$. Assume for some constant $c>0$ that

$$p^{(n)}(o,o)\sim cn^{\beta-1}, \tag{28}$$

as $n\to\infty$. Then

$$\lim_n \frac{p^{(n-1)}(o,o)}{p^{(n)}(o,o)}=1.$$

In view of Orey [33], we obtain the so-called *strong ratio limit property*:

$$\lim_n \frac{p^{(n-m)}(s,o)}{p^{(n)}(o,o)}=1 \tag{29}$$

for any $s\in S$ and $m\in\mathbb{N}$. Then (27) implies

$$\lim_n \frac{1}{p^{(n)}(o,o)}\mu\bigl(A\cap T^{-n}B\bigr)=\mu(A)\mu(B).$$

We conclude that (d) of Proposition 2.5(3) holds with

$$\rho_n=\frac{1}{p^{(n)}(o,o)}$$

when n is large enough so that $p^{(n)}(o,o)>0$.

We can now specify the function f in Theorem 3.1 which must be symmetric bounded and μ^k-a.e. continuous on G^k. We note that any function f on G^k which depends only on a finite number of coordinates of E^k is continuous. This is because in this case, for any open $U\in\mathbb{R}$, $f^{-1}U$ can be expressed as a union of sets in $\mathcal{C}^k$, where $\mathcal{C}^k$ forms a topological base for G^k. On the other hand, it is also possible to choose a symmetric bounded continuous function on G^k depending on infinitely many coordinates, for example,

$$f(x_1,\dots,x_k)=1_{\{x_1(0)=\cdots=x_k(0)=o\}}\sum_{n=1}^{\infty}\frac{\sum_{j=1}^k 1_{\{x_j(n)=o\}}}{2^n}.$$

5. Proofs of the main results

First, we build some intermediate steps towards Theorem 3.1. The notation and setup will follow those in Section 2. In particular, we consider the infinite-measure space $(E,\mathcal{E},\mu)$ with a subspace G as in Assumption 2.1. Below c and c_i will denote generic constants whose values may change from line to line.

5.1. Reduction

Lemma 5.1. *The sequence $X(n)$ defined in (17) is strictly stationary.*

Proof. For every $\epsilon > 0$, we choose f_ϵ and G_ϵ as in Corollary 2.10 satisfying (14). Let

$$R_\epsilon = G^k \setminus G_\epsilon,$$

which satisfies $\mu^k(R_\epsilon) < \epsilon$. Recall the original process $X(n)$ is given in (17), and we define another stationary sequence using the same random measure $W(\cdot)$ as

$$X_\epsilon(n) = \int_{E^k}' f_\epsilon\bigl(T^n x_1, \dots, T^n x_k\bigr) W(dx_1) \dots W(dx_k). \tag{30}$$

Then $X_\epsilon(n)$ is stationary since by Theorem 7.26 and 3.4 of Janson [19], $X_\epsilon(n)$ is a measurable function of a multivariate stationary Gaussian process as in (43) below. Consider the difference $Y_\epsilon(n) = X(n) - X_\epsilon(n)$. Then using (15), the fact $G_\epsilon \subset G^k$, as well as (14), we have

$$\begin{aligned}\mathbb{E}Y_\epsilon(n)^2 &= k!\|f - f_\epsilon\|^2_{L^2(\mu^k)} \le 3k!\bigl(\|(f - f_\epsilon)1_{G_\epsilon}\|^2_{L^2(\mu^k)} + \|f 1_{R_\epsilon}\|^2_{L^2(\mu^k)} + \|f_\epsilon 1_{R_\epsilon}\|^2_{L^2(\mu^k)}\bigr)\\ &\le 3k!\bigl(\epsilon^2 \mu^k\bigl(G^k\bigr) + 2\epsilon\|f\|^2_\infty\bigr) \to 0\end{aligned}$$

as $\epsilon \to 0$. So the stationarity of $X(n)$ follows from that of $X_\epsilon(n)$. □

We claim that the tightness of $\{S_N(t)\}$ in (19) immediately follows if the convergence of the finite-dimensional distributions of $S_N(t)$ has been established. To prove this, we shall apply Lemma 2.1 of Taqqu [46] recalled below.

Lemma 5.2 (Lemma 2.1 of Taqqu [46]). *Let $X(n)$ be a strictly stationary sequence and let $A(N) = N^H L(N)$, where $H \in (0, 1]$, and $L(N)$ is a slowly varying function (cf. Bingham et al. [7]) as $N \to \infty$. Let $S_N(t) = \frac{1}{A(N)} \sum_{n=1}^{[Nt]} X(n)$. Then the tightness of $\{S_N(t)\}$ in $D[0, 1]$ under the uniform metric follows from the following conditions: as $N \to \infty$*

$$\mathbb{E}S_N(1)^2 = O(1), \tag{31}$$

and for some $p > 1/2H$,

$$\mathbb{E}\bigl|S_N(1)\bigr|^{2p} = O\bigl(\bigl(\mathbb{E}S_N(1)^2\bigr)^p\bigr). \tag{32}$$

Lemma 5.3. *Suppose in either of the three cases of Theorem 3.1, the finite-dimensional distributions of normalized sum process $S_N(t)$ in (19) converge to those of the limit process $Z(t)$. Then the tightness of $\{S_N(t)\}$ in $D[0, 1]$ under the uniform metric holds.*

Proof. The strict stationarity of $X(n)$ follows from Lemma 5.4 below. We want to apply Lemma 5.2, for which we need to check conditions (31) and (32).

First, note that $S_N(t)$ also belongs to a kth Wiener chaos by linearity. The convergence in finite-dimensional distributions assumed implies that the sequence of random variables $\{S_N(1)\}$ is tight. By Lemma 2.1 (b) of Nourdin and Rosiński [31], one has

$$\sup_N \mathbb{E} S_N(1)^2 < \infty.$$

Hence, (31) holds. In addition, for arbitrarily large $p > 0$, we have by Theorem 3.50 of Janson [19] that

$$\mathbb{E}\big|S_N(1)\big|^{2p} \le C_p\big(\mathbb{E}\big|S_N(1)\big|^2\big)^p$$

for some constant C_p which depends only on p. Hence (32) also holds. The proof is complete. □

The following is the key reduction lemma.

Lemma 5.4. *Suppose the function f defining the stationary sequence $X(n)$ in (17) is restricted to the class of elementary functions in Definition 2.7, and for any such function f, the weak convergence in either of the three cases in Theorem 3.1 holds. Then Theorem 3.1 also holds in full generality, namely, the weak convergence in either of the three cases holds for any symmetric bounded and μ-a.e. continuous function f on G^k.*

Proof. In view of Lemma 5.3, we only need to consider the convergence of finite-dimensional distributions. Let $X_\epsilon(n)$ and $Y_\epsilon(n)$ be as in the proof of Lemma 5.1 above. Then

$$\frac{1}{A(N)}\sum_{n=1}^N Y_\epsilon(n) = \frac{1}{A(N)}\sum_{n=1}^N X(n) - \frac{1}{A(N)}\sum_{n=1}^N X_\epsilon(n).$$

A standard computation using stationarity leads to

$$\mathbb{E}\left|\frac{1}{A(N)}\sum_{n=1}^N Y_\epsilon(n)\right|^2 = \frac{1}{A(N)^2}\sum_{|n|<N}\big(N-|n|\big)\mathbb{E}Y_\epsilon(n)Y(0). \tag{33}$$

Next by (14), we have

$$\begin{aligned}\big|\mu^k(f_\epsilon)-\mu^k(f)\big| &\le \mu^k\big(|f_\epsilon - f|\big) \le \mu^k\big(|f-f_\epsilon|1_{G_\epsilon}\big) + \mu^k\big(|f-f_\epsilon|1_{R_\epsilon}\big)\\ &\le \epsilon\mu^k\big(G^k\big) + 2\|f\|_\infty \mu^k(R_\epsilon) \to 0 \end{aligned}\tag{34}$$

as $\epsilon \to 0$. So in view of a triangular approximation argument (e.g., Lemma 4.2.1 of Giraitis et al. [14]), if one shows that

$$\lim_{\epsilon\to 0}\limsup_{N\to\infty}\mathbb{E}\left|\frac{1}{A(N)}\sum_{n=1}^N Y_\epsilon(n)\right|^2 \to 0, \tag{35}$$

then the proof is concluded.

• We first treat the case (a): $k(1-\beta) > 1$. In this case, $A(N) = \sigma^{-1}N^{1/2}$ and thus from (33) we have

$$\mathbb{E}\left|\frac{1}{A(N)}\sum_{n=1}^{N} Y_\epsilon(n)\right|^2 \le c \sum_{n=-\infty}^{\infty} \left|\mathbb{E}Y_\epsilon(n)Y_\epsilon(0)\right|,$$

where the constant c does not depend on ϵ. We shall use the Dominated Convergence theorem to show that the right-hand side above tends to zero as $\epsilon \to 0$. Since both f and f_ϵ have supports within G^k, then for $n \ge 0$, using (15) we have

$$\begin{aligned}\left|\mathbb{E}Y_\epsilon(n)Y_\epsilon(0)\right| &\le k! \int_{G^k} \left[\left(|f-f_\epsilon|\right)\left(T_k^n x\right)\right]\left[\left(|f-f_\epsilon|\right)(x)\right]\mu^k(dx)\\ &= k! \int_{G_\epsilon} \left[\left(|f-f_\epsilon|\right)\left(T_k^n x\right)\right]\left[\left(|f-f_\epsilon|\right)(x)\right]\mu^k(dx)\\ &\quad + k! \int_{R_\epsilon} \left[\left(|f-f_\epsilon|\right)\left(T_k^n x\right)\right]\left[\left(|f-f_\epsilon|\right)(x)\right]\mu^k(dx). \end{aligned} \tag{36}$$

In view of (14), applying the bound $|f-f_\epsilon|1_{G_\epsilon} < \epsilon 1_{G_\epsilon}$ to the first term in (36), and applying the bound $|f-f_\epsilon|1_{R_\epsilon} \le 2\|f\|_\infty 1_{R_\epsilon}$ to the second term in (36), we have

$$\left|\mathbb{E}Y_\epsilon(n)Y_\epsilon(0)\right| \le k!\left[\left(\epsilon^2\mu^k\left(G_\epsilon \cap T_k^{-n}G_\epsilon\right) + 4\|f\|_\infty^2\mu^k\left(R_\epsilon \cap T_k^{-n}R_\epsilon\right)\right]\right. \tag{37}$$

$$\le k!\left[\epsilon^2 + 4\|f\|_\infty^2\right)\mu^k\left(G^k \cap T_k^{-n}G^k\right)\left.\right] \tag{38}$$

$$\le c\mu^k\left(G^k \cap T_k^{-n}G^k\right), \tag{39}$$

where the constant $c > 0$ does not depend on ϵ or n, and in the inequality (38), we have used $R_\epsilon \subset G^k$ and $G_\epsilon \subset G^k$. Then by Assumption 2.1 and in particular (7) (see also Remark 2.2), we have as $n \to \infty$ that $\mu^k(G^k \cap T_k^{-n}G^k) \sim \rho_n^{-k}\mu^k(G^k)^2$, which is summable over $n \ge 0$ since $\rho_n^{-k} \sim c_T^{-k}n^{k(\beta-1)}$ by (8) and $k(\beta-1) < -1$. In addition, since $G_\epsilon \subset G^k$ and $\mu^k(R_\epsilon) \to 0$ as $\epsilon \to 0$, the right-hand side of (37) converges to 0 as $\epsilon \to 0$. So by the Dominated Convergence theorem with the summable bound (39), one has as $\epsilon \to 0$ that

$$\sum_{n=-\infty}^{\infty} \left|\mathbb{E}Y_\epsilon(n)Y_\epsilon(0)\right| \to 0.$$

Therefore, the relation (35) holds in case (a) of Theorem 3.1.

• We now treat cases (b) and (c): $k(1-\beta) \le 1$. Since $f - f_\epsilon$ is again bounded and a.e. continuous on G^k, by (11) we have

$$\begin{aligned}\left|\mathbb{E}Y_\epsilon(n)Y_\epsilon(0)\right| &\le k! \int_{G^k} \left[\left(|f-f_\epsilon|\right)\left(T_k^n x\right)\right]\left[\left(|f-f_\epsilon|\right)(x)\right]\mu^k(dx)\\ &\sim \rho_n^{-k}k!\mu^k\left(|f-f_\epsilon|\right)^2, \end{aligned} \tag{40}$$

where $\delta_\epsilon := \mu^k(|f-f_\epsilon|)^2 \to 0$ as $\epsilon \to 0$ by (34).

In the case (b): $k(1-\beta)=-1$, $\rho_n^{-k}=c_T^{-k}n^{-1}$. So for some constants c_i's that do not depend on ϵ or N, we have by (33) and (40) that

$$\begin{aligned}\limsup_{N\to\infty}\mathbb{E}\left|\frac{1}{A(N)}\sum_{n=1}^{N}Y_\epsilon(n)\right|^2 &\le c_1\limsup_{N\to\infty}(\ln N)^{-1}\sum_{|n|<N}\left|\mathbb{E}Y_\epsilon(n)Y_\epsilon(0)\right| \\ &\le c_2\delta_\epsilon\lim_{N\to\infty}(\ln N)^{-1}\sum_{n=1}^{N}n^{-1}=c_3\delta_\epsilon,\end{aligned}$$

which tends to zero as $\epsilon\to 0$. Hence, (35) holds in this case.

Case (c) where $0<k(1-\beta)<1$ can be shown similarly, namely, we have

$$\limsup_{N\to\infty}\mathbb{E}\left|\frac{1}{A(N)}\sum_{n=1}^{N}Y_\epsilon(n)\right|^2\le c_1\delta_\epsilon\lim_{N\to\infty}N^{k(1-\beta)-1}\sum_{n=1}^{N}n^{k(\beta-1)}=c_2\delta_\epsilon.$$

Note that the arguments for case (b) and (c) rely on the divergence of $\sum_{n=1}^{N}\rho_n^{-k}$ and they cannot be applied to case (a). $\square$

5.2. Proof of Theorem 3.1

After the reduction lemmas of this section, Theorem 3.1 reduces to the following proposition, which will be proved in the subsections below. Let $\overset{f.d.d.}{\longrightarrow}$ denote convergence of finite-dimensional distributions.

Proposition 5.5. *Let $X(n)$ be as in (17) where f is an elementary function (Definition 2.7) on G^k which is also symmetric.*

(a) *If $k(1-\beta)>1$, then as $N\to\infty$,*

$$\frac{1}{N^{1/2}}\sum_{n=1}^{[Nt]}X(n)\overset{f.d.d.}{\longrightarrow}\sigma B(t),$$

where $B(t)$ is the standard Brownian motion and $\sigma^2=\sum_n\mathbb{E}X(n)X(0)$.

(b) *If $k(1-\beta)=1$, then as $N\to\infty$,*

$$\frac{c_T^{k/2}}{(k!N\ln(N))^{1/2}}\sum_{n=1}^{[Nt]}X(n)\overset{f.d.d.}{\longrightarrow}\sqrt{2}\mu^k(f)B(t),$$

where $B(t)$ is the standard Brownian motion and $\mu^k(f)=\int_{E^k}f\,d\mu^k$.

(c) *If* $k(1-\beta)<1$, *then as* $N\to\infty$,

$$\frac{c_T^{k/2}}{(k!H(2H-1))^{1/2}N^H}\sum_{n=1}^{[Nt]}X(n)\xrightarrow{f.d.d.}\mu^k(f)Z_{k,H}(t)$$

where $Z_{k,H}(t)$ *is the standard Hermite process with Hurst index* $H=1-k(1-\beta)/2$.

The proposition is proved in the following two subsections.

5.3. Proof of Proposition 5.5(a) and (c)

The proof of Proposition 5.5 can be facilitated by the Fourth Moment theorem developed by Nualart and Peccati [32] and Peccati and Tudor [38]. We state the version involving multivariate convergence and contraction based on Theorem 5.2.7 and Theorem 6.2.3 of Nourdin and Peccati [29]. First some notation: recall that $I_k(f)$ denotes a multiple Wiener–Itô integral, where $f\in L^2(\mu^k)$. For symmetric $f\in L^2(\mu^p)$ and $g\in L^2(\mu^q)$, $p,q\geq 1$, we define the r-contraction, $0\leq r\leq \min(p,q)$ as

$$\begin{aligned}&(f\otimes_r g)(x_1,\dots,x_{p+q-2r})\\&\quad=\int_{E^r}f(y_1,\dots,y_r,x_1,\dots,x_{p-r})g(y_1,\dots,y_r,x_{p-r+1},\dots,x_{p+q-2r})\mu(dy_1)\dots\mu(dy_r),\end{aligned}$$

which if $r=0$ the notation $\otimes_0$ simply denotes the tensor product.

Proposition 5.6. *Let* $k_1,\dots,k_m\geq 1$. *Suppose that* $f_{j,n}\in L^2(\mu^{k_j})$ *are symmetric*, $j=1,\dots,m$ *and* $m,n\in\mathbb{Z}_+$. *Let* $\mathbf{U}_n=(I_{k_1}(f_{1,n}),\dots,I_{k_m}(f_{m,n}))$, *where* $I_{k_i}(\cdot)$*'s denote multiple Wiener–Itô integrals defined by the same Gaussian measure on the same* σ*-finite atomless measure space. Assume as* $n\to\infty$, *we have*

1. $\mathbb{E}I_{k_i}(f_{i,n})I_{k_j}(f_{j,n})\to\Sigma_{ij}$ *for some non-negative definite matrix* $\Sigma=(\Sigma_{i,j})$;
2. $\|f_{j,n}\otimes_r f_{j,n}\|_{L^2(\mu^{2k_j-2r})}\to 0$, $r=1,\dots,k_j-1$, *for all* $k_j\geq 2$, $j=1,\dots,m$.

Then $\mathbf{U}_n\xrightarrow{d}N(\mathbf{0},\Sigma)$ *as* $n\to\infty$.

In addition, under condition 1 *above, the following three statements are equivalent*: (*i*) *condition* 2 *above holds*; (*ii*) *the marginal convergence* $I_{k_i}(f_{i,n})\xrightarrow{d}N(0,\Sigma_{ii})$ *holds as* $n\to\infty$, $i=1,\dots,m$; (*iii*) *the multivariate convergence* $\mathbf{U}_n\xrightarrow{d}N(\mathbf{0},\Sigma)$ *as* $n\to\infty$ *holds.*

We also need the notion of *Hermite rank* of a multivariate function. Let $\mathbf{Z}=(Z_j)_{1\leq j\leq d}\in\mathbb{R}^d$ be a d-variate centered Gaussian vector, $d\in\mathbb{Z}_+$. Let $G(\cdot):\mathbb{R}^d\to\mathbb{R}$ be a measurable function such that $\mathbb{E}G(\mathbf{Z})^2<\infty$ and $\mathbb{E}G(\mathbf{Z})=0$. As with (2.2) of Arcones [2], the Hermite rank k of G with respect to the distribution of $\mathbf{Z}$ can be defined in either of the following two equivalent ways:

$$k = \inf\left\{\tau \in \mathbb{Z}_+ : \exists\, l_j \text{ with } \sum_{j=1}^{d} \ell_j = \tau \text{ and } \mathbb{E}\left(G(\mathbf{Z}) \prod_{j=1}^{d} H_{l_j}(Z_j)\right) \neq 0\right\} \tag{41}$$

$$= \inf\bigl\{\tau \in \mathbb{Z}_+ : \exists\, d\text{-variate polynomial } P \text{ of degree } \tau \text{ with } \mathbb{E}\bigl(G(\mathbf{Z})P(\mathbf{Z})\bigr) \neq 0\bigr\}, \tag{42}$$

where $H_l(x) = (-1)^n e^{x^2/2} \frac{d^l}{dx^l} e^{-x^2/2}$ is the lth order Hermite polynomial. Note that if $L : \mathbb{R}^d \to \mathbb{R}^d$ is a one-to-one linear transform, then the Hermite rank of $G \circ L^{-1}$ with respect to the Gaussian vector $L\mathbf{Z}$ has the same Hermite rank as the Hermite rank of G with respect to $\mathbf{Z}$. This can be seen from (42) since $\{P \circ L : P \text{ is a } d\text{-variate polynomial of degree } \tau\} = \{d\text{-variate polynomials of degree } \tau\}$.

Proof of Proposition 5.5 (a) and (c). Since f is elementary, it is of the form (12). Applying Theorem 7.26 and 3.4 of Janson [19], we have

$$\begin{aligned} X(n) &= \int_{E^k}' \sum_{j=1}^{J} a_j 1_{A_{1,j}\times\cdots\times A_{k,j}}\bigl(T^n x_1, \ldots, T^n x_k\bigr) W(dx_1)\ldots W(dx_k) \\ &= \sum_{j=1}^{J} a_j \int_{E^k}' 1_{A_{1,j}}\bigl(T^n x_1\bigr)\ldots 1_{A_{k,j}}\bigl(T^n x_k\bigr) W(dx_1)\ldots W(dx_k) \\ &= \sum_{j=1}^{J} a_j F_j\bigl(Z_{1,j}(n)\ldots Z_{k,j}(n)\bigr), \end{aligned} \tag{43}$$

where $\mathbf{Z}(n) := (Z_{i,j}(n)) := (\int_E 1_{A_{i,j}}(T^n x) W(dx))_{1\le i\le k, 1\le j\le J}$ is a kJ-variate stationary Gaussian process, and $F_j : \mathbb{R}^k \to \mathbb{R}$ is a k-variate polynomial, and

$$F(\cdot) := \sum_{j=1}^{J} a_j F_j(\cdot) : \mathbb{R}^{kJ} \to \mathbb{R}$$

has Hermite rank k (cf. (41) or (42)) with respect to the distribution of $\mathbf{Z}(n)$. We are thus in the setup of De Naranjo [45] and Arcones [2]. In addition, in view of (7) in the context of Assumption 2.1, as $n \to \infty$, one has

$$\begin{aligned} \mathbb{E} Z_{i,j}(n) Z_{p,q}(0) &= \mu\bigl(T^{-n} A_{i,j} \cap A_{p,q}\bigr) \sim \rho_n^{-1} \mu(A_{i,j})\mu(A_{p,q}) \\ &\sim c_T^{-1} n^{\beta-1} \mu(A_{i,j})\mu(A_{p,q}), \end{aligned} \tag{44}$$

and the same asymptotic relation holds for $\mathbb{E} Z_{i,j}(-n) Z_{p,q}(0)$ as $n \to \infty$.

In the case of (a) where $k(1-\beta) > 1$, note that $\sum_{n=-\infty}^{+\infty} |E Z_{i,j}(n) Z_{p,q}(0)| \le c \times \sum_{n=1}^{+\infty} n^{-k(1-\beta)} < \infty$. Then by Theorem 4 of Arcones [2], at any single fixed $t > 0$ the marginal distributions of $S_N(t) := \frac{1}{N^{1/2}} \sum_{n=1}^{[Nt]} X(n)$ converges weakly to the marginal distribution $\sigma B(t)$ as $N \to \infty$. To extend the marginal convergence to the joint convergence of finite-dimensional distributions, in view of the last part of Proposition 5.6 above, it suffices to show the convergence of covariance structure:

$$\lim_{N\to\infty} \mathbb{E} S_N(s) S_N(t) = \sigma^2 (s \wedge t).$$

By Lemma 4.1 of Bai and Taqqu [4], the preceding line will follow from

$$\left|\mathbb{E} X(n) X(0)\right| \le c n^{-k(1-\beta)}, \quad n \ge 1, \tag{45}$$

where the right-hand side above is summable. Indeed, (45) is a consequence of (18).

Now we turn to case (c) where $k(1-\beta) < 1$. By a Gram–Schmidt orthonormalizing linear transformation $L : \mathbb{R}^{kJ} \to \mathbb{R}^{kJ}$ and setting $\widetilde{\mathbf{Z}}(n) := (\widetilde{Z}_{ij}(n))_{1\le i\le k, 1\le j\le J} := L\mathbf{Z}(n)$, one can ensure that $\mathbb{E}\widetilde{Z}_{i,j}(n)\widetilde{Z}_{p,q}(n) = 1\{(i,j) = (p,q)\}$ for any n (recall stationarity). Then $G := F \circ L^{-1}$ still has Hermite rank k with respect to the distribution of $\widetilde{\mathbf{Z}}(n)$ (see the discussion following (42)) and we have

$$X(n) = G\big(\widetilde{\mathbf{Z}}(n)\big). \tag{46}$$

Suppose $\widetilde{Z}_{ij}(n) = \sum_{u,v} \ell_{ij}(u,v) Z_{uv}(n)$ for some real coefficients $\ell_{ij}(u,v)$, $1 \le i, u \le k$, $1 \le j, v \le J$. By (44) and linearity we have as $|n| \to \infty$ that

$$\mathbb{E}\widetilde{Z}_{i,j}(n)\widetilde{Z}_{p,q}(0) \sim c_T^{-1} |n|^{\beta-1} \left(\sum_{u,v} \ell_{ij}(u,v)\mu(A_{u,v})\right)\left(\sum_{u,v} \ell_{pq}(u,v)\mu(A_{u,v})\right). \tag{47}$$

We shall now apply Theorem 2 on p. 238 of De Naranjo [45] (see also Theorem 6 of Arcones [2]). Note that, in general, De Naranjo [45] allows a Hermite-like process in the limit whose multiple integral representation has possibly different Gaussian random measures. Only when these Gaussian random measures are identical, we get the Hermite process. To show this, in view of (H3.1.1) and (H1.1.1) of De Naranjo [45], we set

$$\begin{aligned}
L_{ij,pq}(n) &= n^{1-\beta}\mathbb{E}\widetilde{Z}_{i,j}(n)\widetilde{Z}_{p,q}(0),\\
L_{ij}(n) &= n^{1-\beta}\mathbb{E}\widetilde{Z}_{i,j}(n)\widetilde{Z}_{i,j}(0),\\
L_{pq}(n) &= n^{1-\beta}\mathbb{E}\widetilde{Z}_{p,q}(n)\widetilde{Z}_{p,q}(0),
\end{aligned}$$

which, in view of (47), have limits

$$\begin{aligned}
&c_T^{-1}\left(\sum_{u,v} \ell_{ij}(u,v)\mu(A_{u,v})\right)\left(\sum_{u,v} \ell_{pq}(u,v)\mu(A_{u,v})\right),\\
&c_T^{-1}\left(\sum_{u,v} \ell_{ij}(u,v)\mu(A_{u,v})\right)\left(\sum_{u,v} \ell_{ij}(u,v)\mu(A_{u,v})\right)
\end{aligned}$$

and

$$c_T^{-1}\Big(\sum_{u,v}\ell_{pq}(u,v)\mu(A_{u,v})\Big)\Big(\sum_{u,v}\ell_{pq}(u,v)\mu(A_{u,v})\Big)$$

respectively as $n\to\infty$. Thus the limit in (H3.1.1) of De Naranjo [45] is

$$K_{ij,pq}=\lim_{n\to\infty}\frac{L_{ij,pq}(n)}{\sqrt{L_{ij}(n)L_{pq}(n)}}=1.$$

Hence, in view of (3.2.1) of De Naranjo [45] and the formula before Lemma 3.3 in De Naranjo [45], all the cross-component control (spectral) measures in (3.3.1) of De Naranjo [45][1] satisfying the scaling property (3.3.2) there are identical to each other. This entails that the Gaussian random measures are identical in the multiple integral representation of the limit. At last, note that in view of (18) above and Proposition 2.2.5 of Pipiras and Taqqu [40], the particular constant coefficient chosen in the normalization ensures that

$$\operatorname{Var}\Bigg[\frac{c_T^{k/2}}{\mu^k(f)(k!H(2H-1))^{1/2}N^H}\sum_{n=1}^{N}X(n)\Bigg]\to 1. \qquad \square$$

5.4. Proof of Proposition 5.5 (b)

Because part (b) of Proposition 5.5 involves the critical case, the proof is particularly delicate.

Proof of Proposition 5.5 (b). We shall apply Proposition 5.6. In view of condition 1 of Proposition 5.6, we first need to show that the covariance structure of

$$S_N(t):=\frac{c_T^{k/2}}{(k!N\ln(N))^{1/2}}\sum_{n=1}^{[Nt]}X(n)$$

converges to that of $\sqrt{2}\mu^k(f)B(t)$. Observe that by polarization, one has for fixed $0\le s\le t$ that

$$\begin{aligned}\mathbb{E}S_N(t)S_N(s)&=\frac{1}{2}\big[\mathbb{E}S_N(t)^2+\mathbb{E}S_N(s)^2-\mathbb{E}\big(S_N(t)-S_N(s)\big)^2\big]\\&\sim\frac{1}{2}\big[\mathbb{E}S_N(t)^2+\mathbb{E}S_N(s)^2-\mathbb{E}S_N(t-s)^2\big]\end{aligned}$$

as $N\to\infty$, where the asymptotic equivalence can be justified by the fact that $|[Nt]-[Ns]-[N(t-s)]|\le 3$, $X(n)$ is stationary and $N\ln(N)\to\infty$. So it suffices to show that

$$\mathbb{E}S_N(t)^2\to 2t\mu^k(f)^2,\quad t>0.$$

[1] G_i and G_j in (3.3.1) of De Naranjo [45] should instead be G_i^0 and G_j^0 respectively.

Indeed, by (18), since $k(1-\beta) = -1$,

$$\gamma(n) := \mathbb{E}X(n)X(0) \sim k!\mu^k(f)^2\rho_n^{-k} \sim k!c_T^{-k}\mu^k(f)^2 n^{-1}$$

as $n \to \infty$. Then as with (33), we have

$$\mathbb{E}\left[\sum_{n=1}^{m} X(n)\right]^2 = m \sum_{|n|<m} \gamma(n) - \sum_{|n|<m} |n|\gamma(n).$$

For $a_n \sim n^{-1}$ as $n \to \infty$, we have $\sum_{n=1}^{m} a_n \sim \ln m$ as $m \to \infty$. So as $m \to \infty$ we have

$$m \sum_{|n|<m} a_n \sim 2m \ln m, \qquad \sum_{|n|<m} |n| a_n \sim 2m.$$

Thus

$$\mathbb{E}S_N(t)^2 \sim \frac{2k!c_T^{-k}[Nt]\ln[Nt]\mu^k(f)^2}{c_T^{-k}k!N\ln N} \to 2t\mu^k(f)^2. \tag{48}$$

We are left to check the contraction condition 2 in Proposition 5.6. For simplicity we take $t = 1$, and the argument is similar otherwise. Then ignoring some multiplicative constant we consider

$$\begin{aligned}
S_N &:= \frac{1}{\sqrt{N\ln N}} \sum_{n=1}^{N} X(n) \\
&= \int_{E^k}' \frac{1}{\sqrt{N\ln N}} \sum_{n=1}^{N} f\big(T^n x_1, \dots, T^n x_k\big) W(dx_1) \dots W(dx_k) \\
&=: I_k(G_N).
\end{aligned}$$

Let $\mathcal{R} = \{A \in \mathcal{E}_G : \mu(\partial A) = 0\}$, the collection of μ-continuity sets in $\mathcal{E}_G$. Note that $G \in \mathcal{R}$. To check

$$\|G_N \otimes_r G_N\|_{L^2} \to 0 \quad \text{as } N \to \infty, r = 1, \dots, k-1,$$

by (12) and a triangle inequality, it suffices to check for any $A_1, \dots, A_k \in \mathcal{R}$ and $B_1, \dots, B_k \in \mathcal{R}$, as $N \to \infty$, we have

$$\begin{aligned}
&\int_{E^{2k-2r}} \mu(dx_1) \dots \mu(dx_{2k-2r}) \\
&\quad \times \left[\int_{E^r} \mu(dy_1) \dots \mu(dy_r)\right. \\
&\quad \times \left(\frac{1}{\sqrt{N\ln N}} \sum_{n_1=1}^{N} 1_{A_1 \times \dots \times A_k}\big(T^{n_1}y_1, \dots, T^{n_1}y_r, T^{n_1}x_1, \dots, T^{n_1}x_{k-r}\big)\right)
\end{aligned}$$

$$\times\left(\frac{1}{\sqrt{N\ln N}}\sum_{n_2=1}^{N}1_{B_1\times\cdots\times B_k}\big(T^{n_2}y_1,\ldots,T^{n_2}y_r,T^{n_2}x_{k-r+1},\ldots,T^{n_2}x_{2k-2r}\big)\right)\Bigg]^2$$
$$\to 0. \tag{49}$$

Using the relation $A_i, B_i \subset G$ (see Assumption 2.1), $i = 1, \ldots, k$, it suffices to show (49) with A_i's and B_i's replaced by G. By the measure-preserving property of T, for $m \geq n$,

$$\mu\big(T^{-n}G\cap T^{-m}G\big)=\mu\big(T^{-n}\big(G\cap T^{n-m}G\big)\big)=\mu\big(G\cap T^{n-m}G\big).$$

Define for $k \in \mathbb{Z}$ that

$$\gamma(k):=\mu\big(G\cap T^{-|k|}G\big).$$

In view of the mixing condition (7) and (8), we have

$$\gamma(k)\sim\mu(G)^2\rho_{|k|}^{-1}\sim\mu(G)^2c_T^{-1}|k|^{\beta-1} \tag{50}$$

as $|k| \to \infty$. After expanding the square in (49) where A_i's and B_i's have been replaced by G, and then applying (50), the expression in (49) is bounded by

$$\begin{aligned}
&\frac{1}{(N\ln N)^2}\sum_{n_1,n_2,n_3,n_4=1}^{N}\gamma(n_1-n_2)^r\gamma(n_3-n_4)^r\gamma(n_1-n_3)^{k-r}\gamma(n_2-n_4)^{k-r}\\
&\quad\leq\frac{c}{(N\ln N)^2}\sum_{n_1,n_2,n_3,n_4=1}^{N}|n_1-n_2|_0^{r(\beta-1)}|n_3-n_4|_0^{r(\beta-1)}\\
&\qquad\times|n_1-n_3|_0^{(k-r)(\beta-1)}|n_2-n_4|_0^{(k-r)(\beta-1)}\\
&\quad=:\frac{c}{(N\ln N)^2}\sum_{n_1,n_2,n_3,n_4=1}^{N}g(n_1,n_2,n_3,n_4)
\end{aligned} \tag{51}$$

where $|n|_0 := |n|$ if $n \neq 0$ and $|0|_0 := 1$. Note that since $0 < r < k$, and $k(\beta - 1) = -1$, we have $r(\beta - 1) \in (-1, 0)$ and $(k - r)(\beta - 1) \in (-1, 0)$.

The goal is to show the bound (51) vanishes as $N \to \infty$. The rest of the proof is similar to the proof of Theorem 3.3 of Bai and Taqqu [5] and we only provide a sketch. In particular, when the sum in (51) is over distinct $n_1, \ldots, n_4$, we have

$$\begin{aligned}
&\sum_{1\leq n_1,n_2,n_3,n_4\leq N,\ n_i\neq n_j\text{ for }i\neq j}g(n_1,n_2,n_3,n_4)\\
&\quad=N^{4+2k(\beta-1)}\\
&\qquad\times\sum_{1\leq n_1,n_2,n_3,n_4\leq N,\ n_i\neq n_j\text{ for }i\neq j}g(n_1/N,n_2/N,n_3/N,n_4/N)\frac{1}{N^4}
\end{aligned}$$

$$\leq cN^2 \int_{[0,1]^4} dx_1\, dx_2\, dx_3\, dx_4 |x_1 - x_2|^{r(\beta-1)} |x_3 - x_4|^{r(\beta-1)} \\ \times |x_1 - x_3|^{(k-r)(\beta-1)} |x_2 - x_4|^{(k-r)(\beta-1)},$$

where for the last inequality we have used the fact $k(\beta - 1) = -1$ and applied an integral approximation of the sum (see the arguments below (3.8) of Bai and Taqqu [5] for details). The integrability of the last integral holds by Lemma 3.9 of Bai and Taqqu [5]. Hence, the additional logarithmic factor makes (51) vanish when the sum is over distinct n_1, n_2, n_3, n_4.

Similarly, when the sum in (51) is over n_1, n_2, n_3, n_4 with only three of them distinct, for example, if $n_1 = n_2$ and $n_i \neq n_j$ for $2 \leq i \neq j \leq 4$, then

$$\sum_{1 \leq n_2, n_3, n_4 \leq N,\ n_i \neq n_j \text{ for } i \neq j} g(n_2, n_2, n_3, n_4)$$

$$= N^{3+(2k-r)(\beta-1)} \sum_{1 \leq n_2, n_3, n_4 \leq N,\ n_i \neq n_j \text{ for } i \neq j} g(n_2/N, n_2/N, n_3/N, n_4/N) \frac{1}{N^3}$$

$$\leq cN^{1+r(1-\beta)} \int_{[0,1]^3} dx_2\, dx_3\, dx_4 |x_3 - x_4|^{r(\beta-1)} |x_2 - x_3|^{(k-r)(\beta-1)} |x_2 - x_4|^{(k-r)(\beta-1)},$$

where $r(1 - \beta) < 1$. So the sum (51) vanishes in this case as well.

When the sum in (51) is over n_1, n_2, n_3, n_4 with only two or less of them distinct, one can simply bound g by 1 and then the summation yields N^2 or N respectively, and (51) obviously vanishes as $N \to \infty$ in either of these cases. $\square$

5.5. Proof of Theorem 3.4

A crucial ingredient of the proof is the following result.

Proposition 5.7. *Let $p_1, \dots, p_r, q_1, \dots, q_s$ be positive integers such that $p_i \geq q_j$ for any $i = 1, \dots, r$ and $j = 1, \dots, s$. Suppose*

$$(\mathbf{X}_n, \mathbf{Y}_n) := (X_{1,n} \dots, X_{r,n}, Y_{1,n}, \dots, Y_{s,n}) \\ := \big(I_{p_1}(f_{1,n}), \dots, I_{p_r}(f_{r,n}), I_{q_1}(g_{1,n}), \dots, I_{q_s}(g_{s,n})\big),$$

where $I_{p_i}(\cdot)$'s and $I_{q_j}(\cdot)$'s denote multiple Wiener–Itô integrals defined by the same Gaussian measure on the same σ-finite atomless measure space. Suppose that as $n \to \infty$, $\mathbf{X}_n \xrightarrow{d} \mathbf{X}$ for some Gaussian *random vector $\mathbf{X}$, and $\mathbf{Y}_n \xrightarrow{d} \mathbf{Y}$ for some random vector $\mathbf{Y}$. In addition, $\lim_n \mathbb{E}[X_{i,n} Y_{j,n}] = 0$, $i = 1, \dots, r$, $j = 1, \dots, s$ (which trivially holds if $p_i > q_j$). Then the joint convergence $(\mathbf{X}_n, \mathbf{Y}_n) \xrightarrow{d} (\mathbf{X}, \mathbf{Y})$ holds with independent $\mathbf{X}$ and $\mathbf{Y}$.*

Remark 5.8. This is Theorem 4.7 of Nourdin and Rosiński [31], but with the moment determinacy restriction on the limit removed due to Nourdin et al. [28].

Proof of Theorem 3.4. The tightness in $D[0,1]^J$ follows from tightness of each component in $D[0,1]$ concluded by Theorem 3.1 (see, e.g., Lemma 1 of Bai and Taqqu [4]). We only need to establish convergence of finite-dimensional distributions. We shall only provide an outline.

First, note that going from case (a) to (b) in Theorem 3.1, the order k of the multiple integral must strictly decrease since β is fixed. We thus claim that Proposition 5.7 allows us to treat the joint convergence involving case (a), (b) and (c) of Theorem 3.1 separately. Indeed, assuming the joint convergence within each case (a), (b) and (c), one can first use Proposition 5.7 to conclude joint convergence of (b) (Gaussian limit) and (c) with asymptotic independence, and then similarly joint convergence of (a) (Gaussian limit) with (b) and (c).

For the joint convergence within case (a) or within case (b) where the limit is Gaussian, according to Proposition 5.6, the conclusion follows from the convergence of finite-dimensional distributions of each component to those of a Brownian motion, which holds by Theorem 3.1, and the convergence of the covariance structure cross different components. The latter can be shown by some routine computation using (11). See, for example, the proof of Lemma 2 of Bai and Taqqu [4] for case (a), and the derivation of (48) in the proof of Proposition 5.5 for case (b). One gets identical Brownian motions if the orders are the same and independent Brownian motions if they are different, as indicated in (21).

For the joint convergence within the non-central limit case (c), again by a reduction argument as Lemma 5.4, it suffices to focus on the case where each f_j in (20) is an elementary function in the sense of Definition 2.7. Then one follows the argument in the proof of Proposition 5.5 (c) to relate to limit theorems for functions of multivariate Gaussian processes considered in De Naranjo [45]. The difference is that in (46), one has instead a vector-valued function G now, and hence needs an extension of De Naranjo [45] to joint convergence of the normalized sums of all the vector components. Such an extension can be achieved by similar arguments as in Bai and Taqqu [4]. The key is an extension of Lemma 4.1 of De Naranjo [45] to the case involving multiple integrals of different orders as was done in the proof of Lemma 6 in Bai and Taqqu [4]. □

Appendix: Proofs of results in Section 2

The appendix section includes the proofs of the technical results in Section 2.

A.1. Proof of Proposition 2.5

Proof. First, we show (3) $\Rightarrow$ (1). Let $\mathcal{D}$ be the minimal class of subsets of G containing $\mathcal{C}$ which is closed under (i) finite unions of disjoint sets and (ii) proper set differences. Then by a variant of Dynkin's theorem, the class $\mathcal{D}$ is the field generated by $\mathcal{C}$. Since $\mathcal{D}$ is minimal, we have $\mathcal{D} \subset \mathcal{E}_G$. In addition, the class of μ-continuity subsets of G also forms a field containing $\mathcal{C}$. Then we have both $\mathcal{D} \subset \mathcal{E}_G$ and any set in $\mathcal{D}$ is μ-continuous. Next, one can verify directly that the set operations (i) and (ii) preserve (7). This by minimality of $\mathcal{D}$, implies that the relation (7) holds for any $A, B \in \mathcal{D}$. Next, since $\mu(\cdot)$ restricted on the Polish subspace G is tight (Kallenberg [21], Theorem 16.3), for any $A \in \mathcal{E}_G$ and any $\epsilon > 0$, there exists a compact $K \subset A$, such that

$\mu(A \setminus K) < \epsilon$. Due to the compactness and (b), there exists $D_1 \in \mathcal{D}$ which is a finite union of sets in $\mathcal{C}$, so that $K \subset D_1 \subset \mathring{A}$. This, together with a similar argument with A replaced by $G \setminus A$, entails the existence of $D_1, D_2 \in \mathcal{D}$, such that $D_1 \subset A \subset D_2$ and $|\mu(D_i) - \mu(A)| < \epsilon$, $i = 1, 2$. Similarly for another μ-continuity set $B \in \mathcal{E}_G$, one can find the corresponding $D_1', D_2' \in \mathcal{D}$, so that $D_1' \subset B \subset D_2'$ and $|\mu(D_i') - \mu(B)| < \epsilon$, $i = 1, 2$. Then the relation (7) for A, B can be deduced by the already established relation (7) for D_1, D_1' and that for D_2, D_2' through letting $n \to \infty$ in

$$\rho_n \mu\big(D_1 \cap T^{-n} D_1'\big) \le \rho_n \mu\big(A \cap T^{-n} B\big) \le \rho_n \mu\big(D_2 \cap T^{-n} D_2'\big), \tag{A.1}$$

and then letting $\epsilon \to 0$. Hence, the mixing relation (7) not only holds for $A, B \in \mathcal{C}$ as stated in part (d), but also for μ-continuity $A, B \in \mathcal{E}_G$ as stated in Assumption 2.1.

For (1) $\Rightarrow$ (3), it suffices to choose $\mathcal{C}$ in (3) to consist of all the μ-continuity sets in $\mathcal{E}_G$. In particular to verify (3)(b), one can choose A to be an open and μ-continuous ball under a metrization.

To show (1) $\Rightarrow$ (2), observe that one can express (7) as (9) by using indicator functions 1_A and 1_B, and hence by linearity, the relation holds for f_1, f_2 which are finite linear combinations of indicators of μ-continuity sets in $\mathcal{E}_G$. Then it extends to general non-negative bounded a.e. continuous f_1, f_2 by an approximation similar to (A.1) via Lemma 2.9. Indeed, we have

$$\rho_n \int g_1 \cdot \big(h_1 \circ T^n\big)\, d\mu \le \rho_n \int f_1 \cdot \big(f_2 \circ T^n\big)\, d\mu \le \rho_n \int g_2 \cdot \big(h_2 \circ T^n\big)\, d\mu,$$

where the functions g_1 and g_2 are chosen as in Lemma 2.9 (in the case $k = 1$) which satisfy $g_1 \le f_1 \le g_2$, and similarly for h_1 and h_2 which satisfy $h_1 \le f_2 \le h_2$. At last, it extends to f_1 and f_2 with general signs by linearity, since each bounded a.e. continuous function f on G can be written as a difference of two non-negative bounded a.e. continuous functions, for example, by $f = (f + \|f\|_\infty) - \|f\|_\infty$.

The implication (2) $\Rightarrow$ (1) follows by letting $f_1 = 1_A$ and $f_2 = 1_B$. □

A.2. Proof of Proposition 2.6

Proof. The claim can be verified via Proposition 2.5(3) with the choice of π-system $\mathcal{C}^p := \{B_1 \times \cdots \times B_k : B_i \in \mathcal{E}_G,\ \mu(\partial B_i) = 0\}$. In particular, (a) and (c) of Proposition 2.5(3) can be verified using (13) and the fact $\partial(A \cap B) \subset \partial(A \cap B) \subset \partial A \cup \partial B$. (b) follows from the fact that $\mathring{\mathcal{C}}^p := \{B_1 \times \cdots \times B_k \in \mathcal{C}^p :\ B_i$ is open$\}$ forms a basis for the topology of G^p (see the proof of Lemma 2.9 below). (d) can be directly verified on $\mathcal{C}^p$. □

A.3. Proof of Lemma 2.9

Proof. Suppose the topology of G is induced by a metric d and set $B(x, \delta) = \{y \in G : d(x, y) < \delta\}$, $\delta > 0$. For $x = (x(1), \ldots, x(k)) \in G^k$ and $\delta > 0$, define the neighborhood

$$B_k(x, \delta) = B\big(x(1), \delta\big) \times \cdots \times B\big(x(k), \delta\big).$$

Let $C \subset G^k$ be the set of continuity points of f on G^k, and fix $\epsilon > 0$. Then for every $x \in C$, since f is continuous at x, there exists $\delta(x) > 0$ such that

$$\omega\big(x, \delta(x)\big) := \sup\big\{\big|f(y_1) - f(y_2)\big| : y_1, y_2 \in B_k\big(x, \delta(x)\big)\big\} \le \epsilon.$$

Furthermore, since $\partial B(x(i), \delta) = \{y \in G : d(x, y) = \delta\}$ are disjoint for different δ's and $\mu(G) < \infty$, the set $\Delta_i := \{\delta > 0 : \ \mu(\partial B(x(i), \delta)) > 0\}$ is at most countable, $i = 1, \dots, k$. So by further adjusting $\delta(x)$ to a smaller value within $\cap_{i=1}^k \Delta_i^c$ if necessary, we can assume that each $B(x(i), \delta(x))$ is μ-continuous, $i = 1, \dots, k$, namely, $B_k(x, \delta(x))$ is elementary.

Next, note that the product space G^k is also separable. Recall that a separable metric space is second-countable (see, e.g., Rudin [42], Exercise 2.23) and thus Lindelöf (every open cover has a countable subcover). Since $\{\omega(x, \delta(x)) : \ x \in C\}$ forms an open cover of C, hence there exist $x_n \in C$ and $\delta_n > 0$, such that $\bigcup_{n=1}^\infty B_k(x_n, \delta_n) \supset C$, where each $B_k(x_n, \delta_n)$ is elementary and $\omega(x_n, \delta_n) < \epsilon$. Let $G_N = \bigcup_{n=1}^N B_k(x_n, \delta_n)$ and choose N large enough so that

$$\mu^k\big(G^k \setminus G_N\big) = \mu^k(C \setminus G_N) < \epsilon. \tag{A.2}$$

Decompose D_n by forming $D_n = B_k(x_n, \delta_n) \setminus (\cup_{i=1}^{n-1} B_k(x_i, \delta_i))$. Observe that each D_n remains an elementary set (see Remark 2.8). Then define

$$g_1(x) = \begin{cases} \displaystyle\sum_{n=1}^N \inf\big\{f(x) : x \in D_n\big\} 1_{D_n} & \text{if } x \in G_N, \\ \inf\big\{f(x) : \ x \in G^k\big\} & \text{if } x \in G^k \setminus G_N, \end{cases} \tag{A.3}$$

so that $g_1 \le f_1$, and define similarly g_2 by replacing inf's with sup's above. Then g_1 and g_2 are elementary functions satisfying $g_1 \le f \le g_2$. Since $D_n \subset B_k(x_n, \delta_n)$ and $\omega(x_n, \delta_n) < \epsilon$, we have

$$\big\|(f - g_1)1_{G_N}\big\|_\infty = \sup_{x \in G_N} \big(f(x) - g_1(x)\big) \le \epsilon \tag{A.4}$$

and thus

$$\mu^k\big((f - g_1)1_{G_N}\big) \le \epsilon\mu^k(G_N) \le \epsilon\mu^k\big(G^k\big). \tag{A.5}$$

On the other hand, for $x \in G^k \setminus G_N$, we have

$$\big\|f(x) - g_1(x)\big\|_\infty \le \|f\|_\infty + \|g_1\|_\infty \le 2\|f\|_\infty$$

so that

$$\mu^k\big((f - g_1)1_{G^k \setminus G_N}\big) \le 2\|f\|_\infty \mu^k\big(G^k \setminus G_N\big) \le 2\|f\|_\infty \epsilon. \tag{A.6}$$

Therefore combining (A.5) and (A.6), we have

$$\mu^k(f - g_1) \le \epsilon\big(\mu^k\big(G^k\big) + 2\|f\|_\infty\big).$$

Similarly, the same inequality holds for $\mu^k(g_2 - f)$. $\square$

A.4. Proof of Corollary 2.10

Proof. Let f_ϵ and G_ϵ be g_1 and G_N respectively in (A.3). The first inequality follows from (A.4). The second inequality follows from the definition of g_1 in (A.3). The last inequality follows from (A.2). □

Acknowledgment. We would like to thank the anonymous referees for their useful suggestions. This work was partially supported by the NSF grant DMS-1309009 at Boston University and by the Simons Foundation grant 569118.

References

[1] Aaronson, J. (1997). *An Introduction to Infinite Ergodic Theory*. *Mathematical Surveys and Monographs* **50**. Providence, RI: Amer. Math. Soc. MR1450400 https://doi.org/10.1090/surv/050

[2] Arcones, M.A. (1994). Limit theorems for nonlinear functionals of a stationary Gaussian sequence of vectors. *Ann. Probab.* **22** 2242–2274. MR1331224

[3] Bai, S., Owada, T. and Wang, Y. (2019). A functional non-central limit theorem for multiple-stable processes with long-range dependence. Preprint. Available at arXiv:1902.00628.

[4] Bai, S. and Taqqu, M.S. (2013). Multivariate limit theorems in the context of long-range dependence. *J. Time Series Anal.* **34** 717–743. MR3127215 https://doi.org/10.1111/jtsa.12046

[5] Bai, S. and Taqqu, M.S. (2016). The universality of homogeneous polynomial forms and critical limits. *J. Theoret. Probab.* **29** 1710–1727. MR3571261 https://doi.org/10.1007/s10959-015-0613-0

[6] Beran, J., Feng, Y., Ghosh, S. and Kulik, R. (2013). *Long-Memory Processes*. Heidelberg: Springer. Probabilistic properties and statistical methods. MR3075595 https://doi.org/10.1007/978-3-642-35512-7

[7] Bingham, N.H., Goldie, C.M. and Teugels, J.L. (1987). *Regular Variation*. *Encyclopedia of Mathematics and Its Applications* **27**. Cambridge: Cambridge Univ. Press. MR0898871 https://doi.org/10.1017/CBO9780511721434

[8] Bogachev, V.I. (2007). *Measure Theory*. *Vol. I*, *II*. Berlin: Springer. MR2267655 https://doi.org/10.1007/978-3-540-34514-5

[9] Brockwell, P. and Marquardt, T. (2005). Lévy-driven and fractionally integrated ARMA processes with continuous time parameter. *Statist. Sinica* **15** 477–494. MR2190215

[10] Dobrushin, R.L. (1979). Gaussian and their subordinated self-similar random generalized fields. *Ann. Probab.* **7** 1–28. MR0515810

[11] 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 https://doi.org/10.1007/BF00535673

[12] Durrett, R. (2010). *Probability*: *Theory and Examples*, 4th ed. *Cambridge Series in Statistical and Probabilistic Mathematics* **31**. Cambridge: Cambridge Univ. Press. MR2722836 https://doi.org/10.1017/CBO9780511779398

[13] Eigen, S., Hajian, A., Ito, Y. and Prasad, V. (2014). *Weakly Wandering Sequences in Ergodic Theory*. *Springer Monographs in Mathematics*. Tokyo: Springer. MR3308698 https://doi.org/10.1007/978-4-431-55108-9

[14] Giraitis, L., Koul, H.L. and Surgailis, D. (2012). *Large Sample Inference for Long Memory Processes*. London: Imperial College Press. MR2977317 https://doi.org/10.1142/p591

[15] Gouëzel, S. (2011). Correlation asymptotics from large deviations in dynamical systems with infinite measure. *Colloq. Math.* **125** 193–212. MR2871313 https://doi.org/10.4064/cm125-2-5

[16] Granger, C.W.J. and Joyeux, R. (1980). An introduction to long-memory time series models and fractional differencing. *J. Time Series Anal.* **1** 15–29. MR0605572 https://doi.org/10.1111/j.1467-9892.1980.tb00297.x
[17] Hajian, A.B. and Kakutani, S. (1964). Weakly wandering sets and invariant measures. *Trans. Amer. Math. Soc.* **110** 136–151. MR0154961 https://doi.org/10.2307/1993640
[18] Hopf, E. (1937). *Ergodentheorie. Ergebnisse der Mathematik und Ihrer Grenzgebiete*. Berlin: Springer.
[19] Janson, S. (1997). *Gaussian Hilbert Spaces. Cambridge Tracts in Mathematics* **129**. Cambridge: Cambridge Univ. Press. MR1474726 https://doi.org/10.1017/CBO9780511526169
[20] Jung, P., Owada, T. and Samorodnitsky, G. (2017). Functional central limit theorem for a class of negatively dependent heavy-tailed stationary infinitely divisible processes generated by conservative flows. *Ann. Probab.* **45** 2087–2130. MR3693958 https://doi.org/10.1214/16-AOP1107
[21] Kallenberg, O. (2002). *Foundations of Modern Probability*, 2nd ed. *Probability and Its Applications* (*New York*). New York: Springer. MR1876169 https://doi.org/10.1007/978-1-4757-4015-8
[22] Kesseböhmer, M. and Slassi, M. (2007). Limit laws for distorted critical return time processes in infinite ergodic theory. *Stoch. Dyn.* **7** 103–121. MR2303796 https://doi.org/10.1142/S0219493707001962
[23] Krickeberg, K. (1967). Strong mixing properties of Markov chains with infinite invariant measure. In *Proc. Fifth Berkeley Sympos. Math. Statist. and Probability* (*Berkeley*, *Calif.*, 1965/66), *Vol. II*: *Contributions to Probability Theory*, *Part* 2 431–446. Berkeley, Calif.: Univ. California Press. MR0219140
[24] Lacaux, C. and Samorodnitsky, G. (2016). Time-changed extremal process as a random sup measure. *Bernoulli* **22** 1979–2000. MR3498020 https://doi.org/10.3150/15-BEJ717
[25] Major, P. (1981). *Multiple Wiener–Itô Integrals. Lecture Notes in Math.* **849**. Berlin: Springer. With applications to limit theorems. MR0611334
[26] Melbourne, I. and Terhesiu, D. (2012). Operator renewal theory and mixing rates for dynamical systems with infinite measure. *Invent. Math.* **189** 61–110. MR2929083 https://doi.org/10.1007/s00222-011-0361-4
[27] Moschovakis, Y.N. (2009). *Descriptive Set Theory*, 2nd ed. *Mathematical Surveys and Monographs* **155**. Providence, RI: Amer. Math. Soc. MR2526093 https://doi.org/10.1090/surv/155
[28] Nourdin, I., Nualart, D. and Peccati, G. (2016). Strong asymptotic independence on Wiener chaos. *Proc. Amer. Math. Soc.* **144** 875–886. MR3430861 https://doi.org/10.1090/proc12769
[29] Nourdin, I. and Peccati, G. (2012). *Normal Approximations with Malliavin Calculus. Cambridge Tracts in Mathematics* **192**. Cambridge: Cambridge Univ. Press. From Stein's method to universality. MR2962301 https://doi.org/10.1017/CBO9781139084659
[30] Nourdin, I. and Peccati, G. (2015). The optimal fourth moment theorem. *Proc. Amer. Math. Soc.* **143** 3123–3133. MR3336636 https://doi.org/10.1090/S0002-9939-2015-12417-3
[31] Nourdin, I. and Rosiński, J. (2014). Asymptotic independence of multiple Wiener–Itô integrals and the resulting limit laws. *Ann. Probab.* **42** 497–526. MR3178465 https://doi.org/10.1214/12-AOP826
[32] Nualart, D. and Peccati, G. (2005). Central limit theorems for sequences of multiple stochastic integrals. *Ann. Probab.* **33** 177–193. MR2118863 https://doi.org/10.1214/009117904000000621
[33] Orey, S. (1961). Strong ratio limit property. *Bull. Amer. Math. Soc.* **67** 571–574. MR0132600 https://doi.org/10.1090/S0002-9904-1961-10694-0
[34] Owada, T. (2016). Limit theory for the sample autocovariance for heavy-tailed stationary infinitely divisible processes generated by conservative flows. *J. Theoret. Probab.* **29** 63–95. MR3463078 https://doi.org/10.1007/s10959-014-0565-9
[35] Owada, T. and Samorodnitsky, G. (2015). Functional central limit theorem for heavy tailed stationary infinitely divisible processes generated by conservative flows. *Ann. Probab.* **43** 240–285. MR3298473 https://doi.org/10.1214/13-AOP899

[36] Owada, T. and Samorodnitsky, G. (2015). Maxima of long memory stationary symmetric α-stable processes, and self-similar processes with stationary max-increments. *Bernoulli* **21** 1575–1599. MR3352054 https://doi.org/10.3150/14-BEJ614

[37] Peccati, G. and Taqqu, M.S. (2011). *Wiener Chaos*: *Moments*, *Cumulants and Diagrams*. *Bocconi & Springer Series* **1**. Milan: Springer; Milan: Bocconi Univ. Press. A survey with computer implementation, Supplementary material available online. MR2791919 https://doi.org/10.1007/978-88-470-1679-8

[38] Peccati, G. and Tudor, C.A. (2005). Gaussian limits for vector-valued multiple stochastic integrals. In *Séminaire de Probabilités XXXVIII*. *Lecture Notes in Math*. **1857** 247–262. Berlin: Springer. MR2126978 https://doi.org/10.1007/978-3-540-31449-3_17

[39] Pipiras, V. and Taqqu, M.S. (2010). Regularization and integral representations of Hermite processes. *Statist*. *Probab. Lett*. **80** 2014–2023. MR2734275 https://doi.org/10.1016/j.spl.2010.09.008

[40] Pipiras, V. and Taqqu, M.S. (2017). *Long-Range Dependence and Self-Similarity*. *Cambridge Series in Statistical and Probabilistic Mathematics*, **45**. Cambridge: Cambridge Univ. Press. MR3729426

[41] Rosenblatt, M. (1961). Independence and dependence. In *Proc*. 4*th Berkeley Sympos*. *Math*. *Statist*. *and Prob*., *Vol*. *II* 431–443. Berkeley, Calif.: Univ. California Press. MR0133863

[42] Rudin, W. (1976). *Principles of Mathematical Analysis*, 3rd ed. New York–Auckland–Düsseldorf: McGraw-Hill. International Series in Pure and Applied Mathematics. MR0385023

[43] Samorodnitsky, G. (2016). *Stochastic Processes and Long Range Dependence*. *Springer Series in Operations Research and Financial Engineering*. Cham: Springer. MR3561100 https://doi.org/10.1007/978-3-319-45575-4

[44] Samorodnitsky, G. and Wang, Y. (2019). Extremal theory for long range dependent infinitely divisible processes. *Ann*. *Probab*. **47** 2529–2562. MR3980927 https://doi.org/10.1214/18-AOP1318

[45] Sánchez de Naranjo, M.V. (1993). Non-central limit theorems for nonlinear functionals of k Gaussian fields. *J*. *Multivariate Anal*. **44** 227–255. MR1219205 https://doi.org/10.1006/jmva.1993.1013

[46] Taqqu, M.S. (1974/75). Weak convergence to fractional Brownian motion and to the Rosenblatt process. *Z*. *Wahrsch*. *Verw*. *Gebiete* **31** 287–302. MR0400329 https://doi.org/10.1007/BF00532868

[47] Taqqu, M.S. (1979). Convergence of integrated processes of arbitrary Hermite rank. *Z*. *Wahrsch*. *Verw*. *Gebiete* **50** 53–83. MR0550123 https://doi.org/10.1007/BF00535674

[48] Thaler, M. (2000). The asymptotics of the Perron–Frobenius operator of a class of interval maps preserving infinite measures. *Studia Math*. **143** 103–119. MR1813362 https://doi.org/10.4064/sm-143-2-103-119

[49] Zweimüller, R. (1998). Ergodic structure and invariant densities of non-Markovian interval maps with indifferent fixed points. *Nonlinearity* **11** 1263–1276. MR1644385 https://doi.org/10.1088/0951-7715/11/5/005

[50] Zweimüller, R. (2000). Ergodic properties of infinite measure-preserving interval maps with indifferent fixed points. *Ergodic Theory Dynam*. *Systems* **20** 1519–1549. MR1786727 https://doi.org/10.1017/S0143385700000821

Received November 2018 and revised October 2019