

2016

Probabilistic and statistical problems related to long-range dependence

<https://hdl.handle.net/2144/17707>

"Downloaded from OpenBU. Boston University's institutional repository."

BOSTON UNIVERSITY
GRADUATE SCHOOL OF ARTS AND SCIENCES

Dissertation

**PROBABILISTIC AND STATISTICAL PROBLEMS RELATED TO
LONG-RANGE DEPENDENCE**

by

SHUYANG BAI

MA, Boston University, 2014

Submitted in partial fulfillment of the
requirements for the degree of
Doctor of Philosophy

2016

© Copyright by
SHUYANG BAI
2016

Approved by

First Reader

Murad S. Taqqu, PhD
Professor of Mathematics and Statistics

Second Reader

Mamikon S. Ginovyan, PhD
Professor of Mathematics and Statistics

Third Reader

Ting Zhang, PhD
Professor of Mathematics and Statistics

Acknowledgments

First, I would like to express my greatest gratitude to my advisor Murad S. Taqqu. He has been guiding my adventure in this intriguing field of long-range dependence ever since my starting year at Boston University. I'm always being inspired by his remarkable dedication to scholarship. The memories of working with him into late nights, over long phone calls and on bumpy airplanes, are still striking me. I've benefited immensely from his demand of precision, his encouragement of independent thinking, as well as his support to travel to numerous conferences.

Second, I would like to thank Mamikon S. Ginovyan and Ting Zhang, with whom I've had fruitful research interactions with. Working with them has greatly expanded my horizons. I would also like to thank the rest of my dissertation committee: Ashis Gangopadhyay and Konstantinos Spiliopoulos. My thanks also goes to other professors who have taught me at Boston University and Beijing Normal University.

Last but not least, I want to thank my family who has been warmly supporting me during the past five years, despite some dark moments they have had to endure. I would especially like to thank my wife Pan Yuan for her company, her patience and her encouragement. Without her love, I would not have been able to cope with the challenges in my PhD journey.

PROBABILISTIC AND STATISTICAL PROBLEMS RELATED TO
LONG-RANGE DEPENDENCE

SHUYANG BAI

Boston University, Graduate School of Arts and Sciences, 2016

Major Professor: Murad S. Taqqu, Professor of Mathematics and Statistics

ABSTRACT

The thesis is made up of a number of studies involving *long-range dependence (LRD)*, that is, a slow power-law decay in the temporal correlation of stochastic models. Such a phenomenon has been frequently observed in practice. The models with LRD often yield non-standard probabilistic and statistical results. The thesis includes in particular the following topics:

- **Multivariate limit theorems.** We consider a vector made of stationary sequences, some components of which have LRD, while the others do not. We show that the joint scaling limits of the vector exhibit an asymptotic independence property.
- **Non-central limit theorems.** We introduce new classes of stationary models with LRD through Volterra-type nonlinear filters of white noise. The scaling limits of the sum lead to a rich class of non-Gaussian stochastic processes defined by multiple stochastic integrals.
- **Limit theorems for quadratic forms.** We consider continuous-time quadratic forms involving continuous-time linear processes with LRD. We show that the scaling limit of such quadratic forms depends on both the strength of LRD and the decaying rate of the quadratic coefficient.
- **Behavior of the generalized Rosenblatt process.** The generalized Rosenblatt process arises from scaling limits under LRD. We study the behavior of this process as its two critical parameters approach the boundaries of the defining region.
- **Inference using self-normalization and resampling.** We introduce a procedure

called “self-normalized block sampling” for the inference of the mean of stationary time series. It provides a unified approach to time series with or without LRD, as well as with or without heavy tails. The asymptotic validity of the procedure is established.

Contents

1	Introduction	1
1.1	Multivariate limit theorems (Chapter 2 and 3)	2
1.2	Non-central limit theorems (Chapter 4 and 5)	3
1.3	Limit theorems for quadratic forms (Chapter 6 and 7)	4
1.4	Behavior of the generalized Rosenblatt process (Chapter 8)	4
1.5	Inference using self-normalization and resampling (Chapter 9)	5
2	Multivariate limit theorems in the context of long-range dependence	6
2.1	Introduction	6
2.2	Review of the univariate results	9
2.3	Multivariate convergence results	13
2.4	Proofs of the multivariate convergence results	23
2.4.1	Proof of Theorem 2.3.1 (SRD case)	23
2.4.2	Proof of Theorem 2.3.3 (LRD case)	29
2.4.3	Proof of Theorem 2.3.6 (SRD and LRD mixed case)	33
2.4.4	Proof of Theorem 2.3.9 (Gaussian linear process case)	34
2.5	Invariance of the joint distribution among different representations of the Hermite process	35
2.6	Asymptotic independence of Wiener-Itô integral vectors	40
3	Multivariate limits of multilinear polynomial-form processes with long memory	48
3.1	Introduction	48

3.2	Preliminaries	51
3.3	Multivariate convergence results	57
3.4	Proofs for the multivariate convergence results	60
3.4.1	Pure SRD case	60
3.4.2	Pure LRD case	62
3.4.3	Mixed SRD and LRD case	62
3.4.4	Weak convergence in $D[0, 1]^J$	69
4	Generalized Hermite processes, discrete chaos and limit theorems	72
4.1	Introduction	72
4.2	Brief review of Hermite processes	75
4.3	Generalized Hermite Processes	79
4.3.1	General kernels	79
4.3.2	Special kernels and examples	86
4.3.3	Fractionally filtered kernels	89
4.4	Discrete chaos processes	94
4.5	Hypercontractivity for infinite discrete chaos	98
4.6	Joint convergence of the discrete chaoses	99
4.6.1	Central limit theorem	100
4.6.2	Non-central limit theorem	102
4.6.3	Non-central limit theorem with fractional filter	105
4.6.4	Mixed multivariate limit theorem	111
5	The universality of homogeneous polynomial forms and critical limits	116
5.1	Introduction	116
5.2	Universality of homogeneous polynomial forms	118
5.3	Application: boundary between short and long memory	123
5.3.1	The setting	123

5.3.2	Statement of the limit theorems	126
5.3.3	Proof of Theorem 5.3.3	127
5.3.4	Proof of Theorem 5.3.5	135
6	Functional Limit Theorems for Toeplitz Quadratic Functionals of Con-	
	tinuous time Gaussian Stationary Processes	138
6.1	Introduction	138
6.2	The Main Results	140
6.3	Preliminaries	146
6.4	Proof of Main Results	155
7	Limit theorems for quadratic forms of Lévy-driven continuous-time linear	
	processes	159
7.1	Introduction	159
7.2	Main results: central and non-central limit theorems	163
7.2.1	Central limit theorems	164
7.2.2	Non-central limit theorems	167
7.3	Preliminaries	171
7.4	Proof of the central limit theorems	175
7.5	Proof of the non-central limit theorems	185
8	Behavior of the generalized Rosenblatt process	
	at extreme critical exponent values	201
8.1	Introduction	201
8.2	Main results	204
8.3	Proof of the main theorems	210
8.3.1	Proof of Theorem 8.2.1	214
8.3.2	Proof of Theorem 8.2.2	219
8.3.3	Proof of Theorem 8.2.5	226

8.3.4	Proof of Theorem 8.2.7	232
8.4	No convergence in $L^2(\Omega)$	236
8.5	Convergence rate of marginal distribution on the boundaries	239
8.6	A constructive proof of Theorem 8.2.2	246
9	A unified approach to self-normalized block sampling	256
9.1	Introduction	256
9.2	Self-Normalized Block Sampling	261
9.3	Asymptotic Theory	266
9.3.1	Results in the Gaussian subordination case	267
9.3.2	Further analysis of Assumption A3	277
9.3.3	Strong mixing case	283
9.4	Examples	286
9.5	Monte Carlo Simulations	290
	Bibliography	296
	Curriculum Vitae	308

List of Tables

9.1	Empirical coverage probabilities of lower and upper (paired in parentheses) one-sided 90% confidence intervals with different combinations of the index d , sample size n and block size $b_n = \lfloor cn^{0.5} \rfloor$ when $a_j = (1 + j)^{d-1}$, $j \geq 0$	294
9.2	Empirical coverage probabilities of lower and upper (paired in parentheses) one-sided 90% confidence intervals with different combinations of the index d , sample size n and block size $b_n = \lfloor cn^{0.5} \rfloor$ when $a_j = (1 + j)^{d-1} \log(1 + j)$, $j \geq 0$	295
9.3	Empirical coverage probabilities of lower and upper (paired in parentheses) one-sided 90% confidence intervals with the TAR model (9.65) for different combinations of sample size n and block size $b_n = \lfloor cn^{0.5} \rfloor$	295

List of Figures

8.1	Region Δ defined in (8.2). The three edges of the triangle are named e_1, e_2 and d (diagonal), while the middle line segment (symmetric axis) is named m	202
8.2	Illustration of limit taking in Theorem 8.2.1	205
8.3	Illustration of limit taking in Theorem 8.2.2	206
8.4	Illustration of limit taking in Theorem 8.2.5	207
8.5	Illustration of limit taking in Theorem 8.2.7	208
8.6	Illustration of the neighborhood $\mathcal{N}(D_\epsilon)$ of D_ϵ in Theorem 8.5.4	241
9.1	Examples of realized \hat{F}_{n,b_n} for models (a)–(c) with $n = 500$, $c = 1$ and different choices of d . The x -axis represents the self-normalized block sums, which have been appropriately centered and scaled.	293

List of Abbreviations and Symbols

FARIMA .	Fractionally integrated autoregressive moving average
LRD	Long-range dependence
SRD	Short-range dependence
a.e./a.s. . . .	Almost surely / almost everywhere
i.i.d.	independent and identically distributed
f.f.d.	finite-dimensional distributions
$E(X)$	Expectation of random variable X
$\text{Cov}(X, Y)$	Covariance between random variables X and Y
$\text{Corr}(X, Y)$	Correlation between random variables X and Y
$D[0, 1]$	Skorohod space, i.e., the functions on $[0, 1]$ which are right-continuous and have left limits
$\mathcal{F}, \mathcal{F}^{-1}$	Fourier and inverse Fourier transform
H -sssi	Self-similar with stationary increments and Hurst index H
$\text{Var}(X)$	Variance of random variable X
\mathbb{R}, \mathbb{R}_+	Real line and positive real line
\mathbb{Z}, \mathbb{Z}_+	Integers and positive integers
$\stackrel{d}{=}$	Equal in distribution
\xrightarrow{d}	Convergence in distribution
\xrightarrow{p}	Convergence in probability
$\xrightarrow{f.d.d.}$	Convergence in finite-dimensional distributions
\Rightarrow	Weak convergence in the Skorohod space $D[0, 1]$
$\ \cdot\ _p$	L^p norm
$\#E$	Cardinality of the set E

Chapter 1

Introduction

In many statistical models, the random noise sequence $\{X_n\}$ is assumed to be independent as the index n varies. For example, this is the case when n indexes different experiments carried out independently. In time series analysis where n indexes the time, treating dependent $\{X_n\}$ is the rule rather than the exception. When the dependence is weak, the large sample theory in statistical inference usually necessitates only a minor modification from the independent case. This is due to the fact that when the dependence is weak (also termed *short-range dependence* (SRD) or *short memory*), one typically has the following central limit theorem describing the scaling behavior of the sample sum:

$$\frac{1}{N^{1/2}} \sum_{n=1}^{[Nt]} X_n \Rightarrow \sigma B(t), \quad (1.1)$$

as $N \rightarrow \infty$, where $B(t)$ is the standard Brownian motion and where

$$\sigma^2 = \sum_{n=-\infty}^{+\infty} \text{Cov}[X_n, X_0], \quad (1.2)$$

the sum of auto-covariance of all orders, the so-called *long-run variance*. Here “ \Rightarrow ” stands for weak convergence in the Skorohod space $D[0, 1]$ (Billingsley [1999]). In the independent case, we just have (1.1) with $\sigma^2 = \text{Var}[X_n]$.

On the other hand, when the dependence is so strong that the covariance function $\text{Cov}[X_n, X_0]$ behaves like $n^{-\alpha}$ as $n \rightarrow \infty$, $\alpha \in (-1, 0)$, the long-run variance in (1.2) becomes infinite, and (1.1) fails to hold. This regime of strong dependence is often addressed

as *long-range dependence* (LRD), also as *long memory*. In practice, long-range dependent data is observed in various fields, e.g., hydrology, finance, internet, biology, etc (see the recent monograph Beran et al. [2013]). Under long-range dependence, it is still possible to establish limit theorems of the type

$$\frac{1}{N^H} \sum_{n=1}^{[Nt]} X_n \Rightarrow Y_H(t), \quad (1.3)$$

where the exponent H , called the *Hurst index*, takes value in the interval $(1/2, 1)$, and where $\{Y_H(t), t \geq 0\}$ is a self-similar process (i.e., $\{Y(ct), t \geq 0\}$ has the same statistical law as $\{c^H Y(t), t \geq 0\}$ for any $c > 0$) with stationary increments. Limit theorems of the type (1.3) are often termed *non-central limit theorems*. The limit process $Y_H(t)$ is typically the fractional Brownian motion, a Gaussian process with dependent increments. But more interestingly, one can as well get convergence to a non-Gaussian limit $Y_H(t)$, which is typically represented by a multiple stochastic integral, e.g., the Hermite processes. See, e.g., Dobrushin and Major [1979], Taqqu [1979], Surgailis [1982], Avram and Taqqu [1987] and Ho and Hsing [1997] for such type of results.

My dissertation focuses on probabilistic and statistical problems related to (1.3). In particular, the dissertation is organized by the following topics:

1.1 Multivariate limit theorems (Chapter 2 and 3)

In Chapter 2, motivated by the needs in statistical inference, we consider multivariate extensions of (1.1) and (1.3) under the Gaussian subordination model. In particular, we consider the vector sequence $(X_{n,1}, \dots, X_{n,J}) = (G_1(Z_n), \dots, G_J(Z_n))$, where $\{Z_n\}$ is a long-range dependent Gaussian sequence and $G_j(\cdot)$'s are different functions. Depending on the choice of $G_j(\cdot)$, the component $\{X_{n,j}\}$ may be short or long-range dependent. We establish multivariate limit theorems for the normalized sum of $(X_{n,1}, \dots, X_{n,J})$, where we find the phenomenon that the short-range dependent components are asymptotically independent of the long-range dependent components, while within each type of the com-

ponents, there is, in general, asymptotic dependence. This part is based on Bai and Taqqu [2013a].

In Chapter 4, we study a similar problem as in Chapter 1, but for a different model of $(X_{n,1}, \dots, X_{n,J})$, where each $X_{n,j}$ is a multilinear moving average of independent and identically distributed (i.i.d.) noise (see (1.4) below). This part is based on Bai and Taqqu [2013b].

1.2 Non-central limit theorems (Chapter 4 and 5)

In Chapter 4, we study limit theorems for the multilinear moving average of the form

$$X_n = \sum'_{i_1, \dots, i_k \geq 0} a(i_1, \dots, i_k) \epsilon_{n-i_1} \dots \epsilon_{n-i_k}, \quad (1.4)$$

where ϵ_i 's are i.i.d. centered random variables with finite variance, the prime ' indicates that the sum excludes the diagonals $i_p = i_q$, $p \neq q$. Depending on the decay of the coefficient $a(\cdot)$, the sequence $\{X_n\}$ can be short or long-range dependent. When it is long-range dependent, the limit $Y_H(t)$, $H > 1/2$, in (1.3) involves the multiple Wiener-Itô integral:

$$Y_H(t) = Z_{H,k}(t) = \int'_{\mathbb{R}^k} \int_0^t g(s - x_1, \dots, s - x_k) ds B(dx_1) \dots B(dx_k), \quad (1.5)$$

where $B(dx)$ is the Brownian random measure, the prime ' indicates the exclusion of the diagonals in the multiple integral, and $g(\cdot)$ is supported on \mathbb{R}_+^k and homogeneous with degree $H - k/2 - 1$. This generalizes the Hermite processes considered in the literature where $g(\cdot)$ is a product of powers. To get $Y_H(t)$ with $H < 1/2$ beyond (1.5), an additional linear filter needs to be applied to $\{X_n\}$ in (1.4). This part is based on Bai and Taqqu [2014a].

In Chapter 5, we consider the case where $\{X_n\}$ is right at the border between short- and long-range dependence. To establish the limit theorem in this delicate case, certain universality result on random multilinear forms involving the Malliavin calculus are used.

This part is based on Bai and Taqqu [2015a].

1.3 Limit theorems for quadratic forms (Chapter 6 and 7)

In Chapter 6 and 7, instead of studying limit theorems for the linear summation functional in (1.3), we consider limit theorems for the Toeplitz type quadratic form

$$Q_T(t) = \int_0^{Tt} \int_0^{Tt} a(s_1 - s_2)X(s_1)X(s_2)ds_1ds_2$$

as $T \rightarrow \infty$ after suitable normalization, where $X(s)$ is a Gaussian process (Chapter 6) or Lévy-driven linear process (Chapter 7), and $a(\cdot)$ is a symmetric coefficient function. The study of $Q_T(t)$ is related to the nonparametric inference of the spectrum of $X(s)$. The type of limit we get depends on the “combined dependence” of $a(\cdot)$ and $X(s)$, that is, it depends on the rate of decay of $a(\cdot)$ as well as the rate of decay of correlation of $X(s)$. When the “combined dependence” is weak, the limit is Brownian motion; when the “combined dependence” is strong, the limit is a non-Gaussian self-similar process represented by a double Wiener-Itô integral. Different representations of this non-Gaussian limit process were also studied. This part is based on Bai et al. [2015] and Bai et al. [2016a].

1.4 Behavior of the generalized Rosenblatt process (Chapter 8)

In Chapter 8, we study a special case of (1.5), that is,

$$R_{\gamma_1, \gamma_2}(t) = \int_{\mathbb{R}^2}' \int_0^t (s - x_1)_+^{\gamma_1} (s - x_2)_+^{\gamma_2} ds B(dx_1)B(dx_2),$$

$$(\gamma_1, \gamma_2) \in \mathbf{\Delta} := \{(\gamma_1, \gamma_2) : \gamma_1, \gamma_2 < -1/2, \gamma_1 + \gamma_2 > -3/2\},$$

called the generalized Rosenblatt process, which was first formally considered in Maejima and Tudor [2012]. In particular, we analyzed the moments of $R_{\gamma_1, \gamma_2}(t)$, based on which we were able to establish interesting distributional behavior of the normalized process $R_{\gamma_1, \gamma_2}(t)$ as (γ_1, γ_2) approaches the boundaries of the triangular region $\mathbf{\Delta}$. On each of the

two symmetric boundaries, the limit is non-Gaussian. On the third diagonal boundary, the limit is Brownian motion. The rates of convergence to these boundaries are also given. The situation is particularly delicate as one approaches the corners of the triangle, because the limit process will depend on how these corners are approached. This part is based on Bai and Taqqu [2015d].

1.5 Inference using self-normalization and resampling (Chapter 9)

The inference procedure for the mean of a stationary time series is usually quite different under various model assumptions because the partial sum process (see (1.3)) behaves differently depending on whether the time series is short or long-range dependent, or whether it has a light or heavy-tailed marginal distribution. These procedures usually involve estimation of additional nuisance parameters. It is often challenging for practitioners to decide which procedure to use given the data, and to know whether their estimation of the nuisance parameters is reliable. A procedure, called self-normalized block sampling, is able to alleviate this challenge by unifying the inference procedure for various aforementioned model assumptions. It avoids the estimation of many nuisance parameters, and requires only the choice of one bandwidth. In Chapter 9, we developed an asymptotic theory for the self-normalized block sampling. Monte Carlo simulations are presented to illustrate its competitive finite-sample performance. The asymptotic consistency of the procedure involves a bound on maximal linear correlation between two blocks of a long-memory time series. This part is based on Bai et al. [2016b].

Chapter 2

Multivariate limit theorems in the context of long-range dependence

We study the limit law of a vector made up of normalized sums of functions of long-range dependent stationary Gaussian series. Depending on the memory parameter of the Gaussian series and on the Hermite ranks of the functions, the resulting limit law may be (a) a multivariate Gaussian process involving dependent Brownian motion marginals, or (b) a multivariate process involving dependent Hermite processes as marginals, or (c) a combination. We treat cases (a), (b) in general and case (c) when the Hermite components involve ranks 1 and 2. We include a conjecture about case (c) when the Hermite ranks are arbitrary, although the conjecture can be resolved in some special cases.

2.1 Introduction

A stationary time series displays long-range dependence if its auto-covariance decays slowly or if its spectral density diverges around the zero frequency. When there is long-range dependence, the asymptotic limits of various estimators are often either Brownian Motion or a Hermite process. The most common Hermite processes are fractional Brownian motion (Hermite process of order 1) and the Rosenblatt process (Hermite process of order 2), but there are Hermite processes of any order. Fractional Brownian motion is the only Gaussian Hermite process.

Most existing limit theorems involve univariate convergence, that is, convergence to a single limit process, for example, Brownian motion or a Hermite process (Breuer and Major

[1983], Dobrushin and Major [1979], Taqqu [1979]). In time series analysis, however, one often needs joint convergence, that is, convergence to a vector of processes. This is because one often needs to consider different statistics of the process jointly. See, for example, Lévy-Leduc et al. [2011], Rooch [2012]. We establish a number of results involving joint convergence, and conclude with a conjecture.

Our setup is as follows. Suppose $\{X_n\}$ is a stationary Gaussian series with mean 0, variance 1 and regularly varying auto-covariance

$$\gamma(n) = L(n)n^{2d-1} \quad (2.1)$$

where

$$0 < d < 1/2,$$

and L is a slowly varying function at infinity. This is often referred to “*long-range dependence*” (*LRD*) or “*long memory*” in the literature, and d is called the *memory parameter*. The higher d , the stronger the dependence. The slow decay (2.1) of $\gamma(n)$ yields

$$\sum_{n=-\infty}^{\infty} |\gamma(n)| = \infty.$$

The case where

$$\sum_{n=-\infty}^{\infty} |\gamma(n)| < \infty,$$

is often referred to “*short-range dependence*” (*SRD*) or “*short memory*”. See Beran [1994], Doukhan et al. [2003], Giraitis et al. [2012] for more details about these notions.

We are interested in the limit behavior of the finite-dimensional distributions (f.d.d.) of the following vector as $N \rightarrow \infty$:

$$\mathbf{V}_N(t) = \left(\frac{1}{A_j(N)} \sum_{n=1}^{\lfloor Nt \rfloor} \left(G_j(X_n) - \mathbb{E}G_j(X_n) \right) \right)_{j=1, \dots, J}, \quad (2.2)$$

where G_j , $j = 1, \dots, J$ are nonlinear functions, $t > 0$ is the time variable, and $A_j(N)$'s are appropriate normalizations which make the variance of each component at $t = 1$ tend to 1. Observe that the same sequence $\{X_n\}$ is involved in each component of \mathbf{V}_N , in contrast to Ho and Sun [1990] who consider the case $J = 2$ and $\{(X_n, Y_n)\}$ is a bivariate Gaussian vector series.

Note also that convergence in f.d.d. implies that our results continue to hold if one replaces the single time variable t in (2.2) with a vector (t_1, \dots, t_J) which would make $\mathbf{V}_N(t_1, \dots, t_J)$ a random field.

Depending on the memory parameter of the Gaussian series and on the Hermite ranks of the functions (Hermite ranks are defined in Section 2.2), the resulting limit law for (2.2) may be:

- (a) a multivariate Gaussian process with dependent Brownian motion marginals,
- (b) or a multivariate process with dependent Hermite processes as marginals,
- (c) or a combination.

We treat cases (a), (b) in general and case (c) when the Hermite components involve ranks 1 and 2 only. To address case (c), we apply a recent asymptotic independence theorem of Nourdin and Rosinski [2014] of Wiener-Itô integral vectors. We include a conjecture about case (c) when the Hermite ranks are arbitrary. This conjecture has been recently resolved by Nourdin et al. [2016]. We also prove that the Hermite processes in the limit are dependent on each other. Thus, in particular, fractional Brownian motion and the Rosenblatt process in the limit are dependent processes even though they are uncorrelated. Although our results are formulated in terms of convergence of f.d.d. , under some additional assumption, they extend to weak convergence in $D[0, 1]^J$ (J-dimensional product space where $D[0, 1]$ is the space of Càdlàg functions on $[0, 1]$ with the uniform metric), as noted in Theorem 2.3.12 at the end of Section 2.3.

The chapter is structured as follows. We review the univariate results in Section 2.2. In Section 2.3, we state the corresponding multivariate results. Section 2.4 contains the

proofs of the theorems in Section 2.3. Section 2.5 shows that the different representations of the Hermite processes are also equivalent in a multivariate setting. Section 2.6 refers to the results of Nourdin and Rosinski [2014] and concerns asymptotic independence of Wiener-Itô integral vectors.

2.2 Review of the univariate results

We review first results involving (2.2) when $J = 1$ in (2.2). Assume that G belongs to $L^2(\phi)$, the set of square-integrable functions with respect to the standard Gaussian measure ϕ . This Hilbert space $L^2(\phi)$ has a complete orthogonal basis $\{H_m(x)\}_{m \geq 0}$, where H_m is the *Hermite polynomial* defined as

$$H_m(x) = (-1)^m \exp\left(\frac{x^2}{2}\right) \frac{d^m}{dx^m} \exp\left(-\frac{x^2}{2}\right),$$

(Nourdin and Peccati [2012], Chapter 1.4). Therefore, every function $G \in L^2(\phi)$ admits the following type of expansion:

$$G = \sum_{m \geq 0} g_m H_m, \tag{2.3}$$

where $g_m = (m!)^{-1} \int_{\mathbb{R}} G(x) H_m(x) d\phi(x)$.

Since $H_0(x) = 1$ and since we always center the series $\{G(X_n)\}$ by subtracting its mean in (2.2), we may always assume $g_0 = \mathbb{E}G(X_n) = 0$. The smallest index $k \geq 1$ for which $g_k \neq 0$ in the expansion (2.3) is called the *Hermite rank* of G .

Since $\{X_n\}$ is a stationary Gaussian series, it has the following spectral representation

$$X_n = \int_{\mathbb{R}} e^{inx} dW(x), \tag{2.4}$$

where W is the complex Hermitian ($W(A) = \overline{W(-A)}$) Gaussian random measure specified by $\mathbb{E}W(A)\overline{W(B)} = F(A \cap B)$. The measure F is called the *spectral distribution* of $\{X_n\}$,

is also called the *control measure* of W , and is defined by

$$\gamma(n) = \mathbb{E}X_n X_0 = \int_{\mathbb{R}} e^{inx} dF(x),$$

(see Lifshits [2012], Chapter 3.2).

Multiple Wiener-Itô integrals (Major [2014])

$$I_m(K) = \int_{\mathbb{R}^m}'' K(x_1, \dots, x_m) dW(x_1) \dots dW(x_m) \quad (2.5)$$

where

$$\int_{\mathbb{R}^m} |K(x_1, \dots, x_m)|^2 dF(x_1) \dots dF(x_m) < \infty,$$

play an important role because of the following connection between Hermite polynomials and multiple Wiener-Itô integrals (Nourdin and Peccati [2012] Theorem 2.7.7):

$$H_m(X_n) = \int_{\mathbb{R}^m}'' e^{in(x_1 + \dots + x_m)} dW(x_1) \dots dW(x_m), \quad (2.6)$$

where the double prime $''$ indicates that one doesn't integrate on the hyper-diagonals $x_j = \pm x_k, j \neq k$. Throughout this chapter, $I_m(\cdot)$ denotes a m -tuple Wiener-Itô integral of the type in (2.5).

We now recall some well-known univariate results:

Theorem 2.2.1. (SRD Case.) *Suppose the memory parameter d and the Hermite rank $k \geq 1$ of G satisfy*

$$0 < d < \frac{1}{2} \left(1 - \frac{1}{k}\right).$$

Then

$$\frac{1}{A(N)} \sum_{n=1}^{[Nt]} G(X_n) \xrightarrow{f.d.d.} B(t),$$

where $B(t)$ is a standard Brownian Motion, " $\xrightarrow{f.d.d.}$ " denotes convergence in finite-dimensional distributions along the time variable $t > 0$, $A(N) \propto N^{1/2}$ is a normalization factor such

that

$$\lim_{N \rightarrow \infty} \text{Var} \left(\frac{1}{A(N)} \sum_{n=1}^N G(X_n) \right) = 1.$$

Remark 2.2.2. It can indeed be shown that in the setting of Theorem 2.2.1,

$$\text{Var} \left(\sum_{n=1}^N G(X_n) \right) \sim \sigma^2 N, \quad (2.7)$$

where

$$\sigma^2 = \sum_{m=k}^{\infty} g_m^2 m! \sum_{n=-\infty}^{\infty} \gamma(n)^m. \quad (2.8)$$

Recall that the g_m 's are the coefficients of the Hermite expansion of G , and γ is the auto-covariance function of $\{X_n\}$.

Remark 2.2.3. The condition $0 < d < \frac{1}{2}(1 - \frac{1}{k})$ can be replaced with a weaker condition

$$\sum_{n=-\infty}^{\infty} |\gamma(n)|^k < \infty,$$

or equivalently, $\sum_{n=-\infty}^{\infty} |\gamma_G(n)| < \infty$, where $\gamma_G(n)$ is the auto-covariance function of $\{G(X_n)\}$. See Theorem 4.6.1 in Giraitis et al. [2012]. If $d = \frac{1}{2}(1 - \frac{1}{k})$ but as $N \rightarrow \infty$,

$$\sum_{n=-N}^N |\gamma(n)|^k = \sum_{-N}^N n^{-1} |L(n)|^k =: L^*(N) \rightarrow \infty$$

is slowly varying, then one still gets convergence to Brownian motion (Theorem 1' of Breuer and Major [1983]), but with the normalization

$$A(N) \propto (NL^*(N))^{1/2}.$$

For example, if the slowly varying function in (2.1) is $L(n) \sim c > 0$, then $A(N) \propto (N \ln N)^{1/2}$.

The original proof of Theorem 2.2.1 (Breuer and Major [1983]) was done by a method of moments using the so-called diagram formulas (Peccati and Taqqu [2011]), which provide

explicit ways to compute the cumulants of Hermite polynomials of Gaussian random variable. Recently, a remarkable technique for establishing central limit theorems of multiple Wiener-Itô integral was found by Nualart and Peccati [2005], Peccati and Tudor [2005], whereby in the multiple Wiener-Itô integral setting, convergence of the fourth moment, or some equivalent easier-to-check condition, implies directly the Gaussian limit. See Theorem 7.2.4 in Nourdin and Peccati [2012] for a proof in the case $t = 1$.

Theorem 2.2.4. (LRD Case.) *Suppose that the memory parameter d and the Hermite rank $k \geq 1$ of G satisfy*

$$\frac{1}{2}\left(1 - \frac{1}{k}\right) < d < \frac{1}{2}.$$

Then

$$\frac{1}{A(N)} \sum_{n=1}^{\lfloor Nt \rfloor} G(X_n) \xrightarrow{f.d.d.} Z_d^{(k)}(t) := I_k(f_{k,d}^{(t)}),$$

where the control measure of $I_k(\cdot)$ is Lebesgue,

$$A(N) \propto N^{1+(d-1/2)k} L(N)^{k/2}$$

is a normalization such that

$$\lim_{N \rightarrow \infty} \text{Var} \left(\frac{1}{A(N)} \sum_{n=1}^N G(X_n) \right) = 1,$$

and

$$f_{k,d}^{(t)}(x_1, \dots, x_k) = b_{k,d} \frac{e^{it(x_1 + \dots + x_k)} - 1}{i(x_1 + \dots + x_k)} |x_1|^{-d} \dots |x_k|^{-d},$$

where

$$b_{k,d} = \left(\frac{(k(d-1/2)+1)(2k(d-1/2)+1)}{k! (2\Gamma(1-2d) \sin(d\pi))^k} \right)^{1/2}$$

is the normalization constant to guarantee unit variance for $Z^{(k)}(1)$.

For a proof, see Dobrushin and Major [1979] and Pipiras and Taqqu [2010]. The process $Z_d^{(k)}(t)$ appearing in the limit is called a *Hermite process*.

Remark 2.2.5. It can indeed be shown that in the setting of Theorem 2.2.4,

$$\text{Var} \left(\sum_{n=1}^N G(X_n) \right) = L_G(N) N^{2d_G+1} \quad (2.9)$$

for some slowly varying function $L_G(N) \propto L(N)^k$ and

$$d_G = (d - 1/2)k + 1/2$$

(see e.g. (3.3.8) in Giraitis et al. [2012]). Since $d < 1/2$, increasing the Hermite rank k decreases the memory parameter d_G , hence decreases the dependence. Note that if $k \geq 2$, then the variance growth of $\{G(X_n)\}$ in (2.9) is slower than the variance growth of $\{X_n\}$,

$$\text{Var} \left(\sum_{n=1}^N X_n \right) = L_0(N) N^{2d+1}$$

for some slowly varying function L_0 , but is always faster than the variance growth $\sigma^2 N$ in the SRD case in (2.7).

The process $Z_d^{(1)}(t)$, $t \geq 0$ is a Gaussian process called *fractional Brownian motion*, and $Z_d^{(2)}(t)$, $t \geq 0$ is a non-Gaussian process called *Rosenblatt process*. The Hermite processes $Z_d^{(k)}(t)$ are all so-called *self-similar processes* (Embrechts and Maejima [2002]).

2.3 Multivariate convergence results

Our aim is to study the limit of (2.2), and in particular, to extend Theorem 2.2.1 (SRD) and Theorem 2.2.4 (LRD) to a multivariate setting.

Suppose that for each $j = 1, \dots, J$, the function G_j in (2.2) belongs to $L^2(\phi)$, has Hermite rank k_j and admits Hermite expansion $\sum_{m=k_j}^{\infty} g_{m,j} H_m$ (see (2.3)).

We start with the pure SRD case where every component $\{G_j(X_n)\}$ of $\mathbf{V}_N(t)$ in (2.2) is SRD.

Theorem 2.3.1. (SRD Case.) *If the memory parameter d is small enough so that all*

$\{G_j(X_n)\}, j = 1, \dots, J$ are SRD, that is,

$$d < \frac{1}{2}\left(1 - \frac{1}{k_j}\right), \quad j = 1, \dots, J,$$

then in (2.2)

$$\mathbf{V}_N(t) \xrightarrow{f.d.d.} \mathbf{B}(t),$$

as $N \rightarrow \infty$, where the normalization $A_j(N) \propto N^{1/2}$ is such that for $j = 1, \dots, J$,

$$\lim_{N \rightarrow \infty} \text{Var} \left(\frac{1}{A_j(N)} \sum_{n=1}^N G_j(X_n) \right) = 1. \quad (2.10)$$

Here

$$\mathbf{B}(t) = (B_1(t), \dots, B_J(t))$$

is a multivariate Gaussian process with standard Brownian motions as marginals, and where the cross-covariance between two components is

$$\begin{aligned} \text{Cov}(B_{j_1}(t_1), B_{j_2}(t_2)) &= \lim_{N \rightarrow \infty} \text{Cov}(V_{N,j_1}(t_1), V_{N,j_2}(t_2)) \\ &= (t_1 \wedge t_2) \left[\frac{1}{\sigma_{j_1} \sigma_{j_2}} \sum_{m=k_{j_1} \vee k_{j_2}}^{\infty} g_{m,j_1} g_{m,j_2} m! \sum_{n=-\infty}^{\infty} \gamma(n)^m \right] \end{aligned} \quad (2.11)$$

where

$$\sigma_j^2 = \sum_{m=k_j}^{\infty} g_{m,j}^2 m! \sum_{n=-\infty}^{\infty} \gamma(n)^m. \quad (2.12)$$

This theorem is proved in Section 2.4.1.

Example 2.3.2. Assume that the auto-covariance function $\gamma(n) \sim n^{2d-1}$, $0 < d < 1/4$, as $n \rightarrow \infty$. Let $J = 2$,

$$G_1(x) = aH_2(x) + bH_3(x) = bx^3 + ax^2 - 3bx - a, \quad G_2(x) = cH_3(x) = cx^3 - 3cx.$$

Then in (2.12),

$$\sigma_1^2 = 2a^2 \sum_{n=-\infty}^{\infty} \gamma(n)^2 + 6b^2 \sum_{n=-\infty}^{\infty} \gamma(n)^3, \quad \sigma_2^2 = 6c^2 \sum_{n=-\infty}^{\infty} \gamma(n)^3,$$

and

$$\left(\frac{1}{N^{1/2}} \sum_{n=1}^{[Nt]} (X_n^2 - 1), \frac{1}{N^{1/2}} \sum_{n=1}^{[Nt]} (X_n^3 - 3X_n) \right) \xrightarrow{f.d.d.} (\sigma_1 B_1(t), \sigma_2 B_2(t)),$$

where the Brownian motions B_1 and B_2 have the covariance structure:

$$\text{Cov}(B_1(t_1), B_2(t_2)) = 6b \frac{t_1 \wedge t_2}{\sigma_1 \sigma_2} \sum_{n=-\infty}^{\infty} \gamma(n)^3.$$

B_1 and B_2 are independent when $b = 0$.

Next we consider the case where every component $\{G_j(X_n)\}$ of $\mathbf{V}_N(t)$ in (2.2) is LRD.

Theorem 2.3.3. (LRD Case.) *If the memory parameter d is large enough so that all $G_j(X_n), j = 1, \dots, J$ are LRD, that is,*

$$d > \frac{1}{2} \left(1 - \frac{1}{k_j}\right), \quad j = 1, \dots, J,$$

then in (2.2),

$$\mathbf{V}_N(t) \xrightarrow{f.d.d.} \mathbf{Z}_d^{\mathbf{k}}(t) := \left(I_{k_1}(f_{k_1,d}^{(t)}), \dots, I_{k_J}(f_{k_J,d}^{(t)}) \right), \quad (2.13)$$

where the normalization $A_j(N) \propto N^{1+(d-1/2)k_j} L(N)^{k_j/2}$ is such that for $j = 1, \dots, J$,

$$\lim_{N \rightarrow \infty} \text{Var} \left(\frac{1}{A_j(N)} \sum_{n=1}^N G_j(X_n) \right) = 1. \quad (2.14)$$

Each component of $\mathbf{Z}_d^{\mathbf{k}}(t) := \left(Z_d^{(k_1)}(t), \dots, Z_d^{(k_J)}(t) \right)$ is a standard Hermite process, and $I_k(\cdot)$ denotes k -tuple Wiener-Itô integral with respect to a common complex Hermitian

Gaussian random measure W with Lebesgue control measure, and

$$f_{k,d}^{(t)}(x_1, \dots, x_k) = b_{k,d} \frac{e^{it(x_1 + \dots + x_k)} - 1}{i(x_1 + \dots + x_k)} |x_1|^{-d} \dots |x_k|^{-d}, \quad (2.15)$$

where $b_{k,d}$'s are the same normalization constants as in Theorem 2.2.4.

This theorem is proved in Section 2.4.2.

Example 2.3.4. Assume that auto-covariance function $\gamma(n) \sim n^{2d-1}$, $1/4 < d < 1/2$, as $n \rightarrow \infty$. Let $J = 2$,

$$G_1(x) = H_1(x) = x, \quad G_2(x) = H_2(x) = x^2 - 1,$$

then

$$\left(\frac{1}{N^{1/2+d}} \sum_{n=1}^{[Nt]} X_n, \frac{1}{N^{2d}} \sum_{n=1}^{[Nt]} (X_n^2 - 1) \right) \xrightarrow{f.d.d.} \left(\frac{1}{d(2d+1)} Z_d^{(1)}(t), \frac{1}{d(4d-1)} Z_d^{(2)}(t) \right),$$

where the standard fractional Brownian motion $Z_d^{(1)}(t)$ and standard Rosenblatt process $Z_d^{(2)}(t)$ share the same random measure in the Wiener-Itô integral representation. The components $Z_d^{(1)}$ and $Z_d^{(2)}$ are uncorrelated but dependent as stated below.

In Theorem 2.3.3, the marginal Hermite processes

$$Z_d^{(k_1)}(t) = I_{k_1}(f_{k_1,d}^{(t)}), \dots, Z_d^{(k_J)}(t) = I_{k_J}(f_{k_J,d}^{(t)})$$

are dependent on each other. To prove this, we use a different representation of the Hermite process, namely, the positive half-axis representation given in (2.45).

Proposition 2.3.5. *The marginal Hermite processes $Z_d^{(k_1)}, \dots, Z_d^{(k_J)}$ involved in Theorem 2.3.3 are dependent.*

Proof. From Ustunel and Zakai [1989], we have the following criterion for the independence of multiple Wiener-Itô integrals: suppose that symmetric $g_1 \in L^2(\mathbb{R}_+^p)$ and $g_2 \in L^2(\mathbb{R}_+^q)$.

Then $I_p(g_1)$ and $I_q(g_2)$ ($p, q \geq 1$) are independent if and only if

$$g_1 \otimes_1 g_2 := \int_{\mathbb{R}_+} g_1(x_1, \dots, x_{p-1}, u) g_2(x_p, \dots, x_{p+q-2}, u) du = 0 \text{ in } L^2(\mathbb{R}_+^{p+q-2}).$$

We shall apply this criterion to the positive half-axis integral representation (2.45) of Hermite processes (see also Pipiras and Taqqu [2010]):

$$\begin{aligned} Z_d^{(k)}(t) &= c_{k,d} I_k \left(g_{k,d}^{(t)}(x_1, \dots, x_k) \right) \\ &:= c_{k,d} \int_{\mathbb{R}_+^k}' \left[\int_0^t \prod_{j=1}^k x_j^{-d} (1 - sx_j)_+^{d-1} ds \right] dB(x_1) \dots dB(x_k), \end{aligned}$$

where B is Brownian motion, the prime $'$ indicates the exclusion of diagonals with $x_j = x_k, j \neq k$ and $c_{k,d}$ is some normalization constant. In fact, for a vector made up of Hermite processes sharing the same random measure in their Wiener-Itô integral representation, the joint distribution does not change when switching from one representation of Hermite process to another. See Section 2.5.

One can then see (let $t = 1$ and thus $g_{k,d} := g_{k,d}^{(1)}$) that for all $(x_1, \dots, x_{p+q-2}) \in \mathbb{R}_+^{p+q-2}$:

$$\begin{aligned} &(g_{p,d} \otimes_1 g_{q,d})(x_1, \dots, x_{p+q-2}) \\ &= \int_{\mathbb{R}_+} \left(\int_0^1 \prod_{j=1}^{p-1} x_j^{-d} (1 - sx_j)_+^{d-1} u^{-d} (1 - su)_+^{d-1} ds \times \right. \\ &\quad \left. \int_0^{1-p+q-2} \prod_{j=p} x_j^{-d} (1 - sx_j)_+^{d-1} u^{-d} (1 - su)_+^{d-1} ds \right) du > 0 \end{aligned}$$

because every term involved in the integrand is positive. \square

Theorem 2.3.1 and Theorem 2.3.3 describe the convergence of $\mathbf{V}_N(t)$ in (2.2) when the $\{G_j(X_n)\}, j = 1, \dots, J$ are all purely SRD or purely LRD. However, when the components in $\mathbf{V}_N(t)$ are mixed, that is, some of them are SRD and some of them are LRD, it is not immediately clear what the limit behavior is and also what the inter-dependence structure

between the SRD and LRD limit components is. We show that the SRD part and LRD part are asymptotically independent so that one could join the limits of Theorem 2.3.1 and Theorem 2.3.3 together, in the case when the G_j 's in the LRD part only involve the two lowest Hermite ranks, namely, $k = 1$ or $k = 2$. This is stated in the next theorem where the letter ‘‘S’’ refers to the SRD part and ‘‘L’’ to the LRD part.

Theorem 2.3.6. (SRD and LRD Mixed Case.) *Separate the SRD and LRD parts of $\mathbf{V}_N(t)$ in (2.2), that is, let $\mathbf{V}_N(t) = (\mathbf{S}_N(t), \mathbf{L}_N(t))$, where*

$$\mathbf{S}_N(t) = \left(\frac{1}{A_{1,S}(N)} \sum_{n=1}^{[Nt]} G_{1,S}(X_n), \dots, \frac{1}{A_{J_S,S}(N)} \sum_{n=1}^{[Nt]} G_{J_S,S}(X_n) \right), \quad (2.16)$$

$$\mathbf{L}_N(t) = \left(\frac{1}{A_{1,L}(N)} \sum_{n=1}^{[Nt]} G_{1,L}(X_n), \dots, \frac{1}{A_{J_L,L}(N)} \sum_{n=1}^{[Nt]} G_{J_L,L}(X_n) \right), \quad (2.17)$$

where $G_{j,S}$ has Hermite rank $k_{j,S}$, and $G_{j,L}$ has Hermite rank $k_{j,L}$,

$$A_{j,S} \propto N^{1/2} \quad \text{and} \quad A_{j,L} \propto N^{1+(d-1/2)k_{j,L}} L(N)^{k_{j,L}/2}$$

are the correct normalization factors such that for $j = 1, \dots, J_S$ and $j = 1, \dots, J_L$ respectively,

$$\lim_{N \rightarrow \infty} \text{Var} \left(\frac{1}{A_{j,S}(N)} \sum_{n=1}^N G_{j,S}(X_n) \right) = 1, \quad \lim_{N \rightarrow \infty} \text{Var} \left(\frac{1}{A_{j,L}(N)} \sum_{n=1}^N G_{j,L}(X_n) \right) = 1. \quad (2.18)$$

In addition,

$$\frac{1}{2} \left(1 - \frac{1}{k_{j_L,L}} \right) < d < \frac{1}{2} \left(1 - \frac{1}{k_{j_S,S}} \right) \quad \text{for all } j_S = 1, \dots, J_S, \quad j_L = 1, \dots, J_L, \quad (2.19)$$

where we allow arbitrary values for $k_{j,S}$ but only $k_{j,L} = 1$ or 2 . (Condition (2.19) makes all $\{G_{j,S}(X_n)\}$ SRD and all $\{G_{j,L}(X_n)\}$ LRD.)

Then we have

$$(\mathbf{S}_N(t), \mathbf{L}_N(t)) \xrightarrow{f.d.d.} (\mathbf{B}(t), \mathbf{Z}_d^{(\mathbf{k}_L)}(t)), \quad (2.20)$$

where the multivariate Gaussian process $\mathbf{B}(t)$ is given in (2.3.1) and the multivariate standard Hermite process $\mathbf{Z}_d^{(\mathbf{k}_L)}(t)$ is given in (2.3.3). Moreover, the vectors $\mathbf{B}(t)$ and $\mathbf{Z}_d^{(\mathbf{k}_L)}(t)$ are independent.

This theorem is proved in Section 2.4.3. Observe that while $\mathbf{B}(t)$ is made up of correlated Brownian motions, it follows from Theorem 2.3.6 that if $\mathbf{Z}_d^{(\mathbf{k}_L)}(t)$ contains fractional Brownian motion as a component, then the fractional Brownian motion will be independent of any Brownian motion component of $\mathbf{B}(t)$.

Example 2.3.7. Assume that the auto-covariance function $\gamma(n) \sim n^{2d-1}$, $1/4 < d < 1/3$, as $n \rightarrow \infty$. Let $J = 2$,

$$G_1(x) = H_2(x) = x^2 - 1, \quad G_2(x) = H_3(x) = x^3 - 3x,$$

then $\sigma^2 = 6 \sum_{n=-\infty}^{\infty} \gamma(n)^3$ and

$$\left(\frac{1}{N^{2d}} \sum_{n=1}^{[Nt]} (X_n^2 - 1), \frac{1}{N^{1/2}} \sum_{n=1}^{[Nt]} (X_n^3 - 3X_n) \right) \xrightarrow{f.d.d.} \left(\frac{1}{d(4d-1)} Z_d^{(2)}(t), \sigma B(t) \right).$$

where the standard Rosenblatt process $Z_d^{(2)}(t)$ and the standard Brownian motion $B(t)$ are independent.

The proof of Theorem 2.3.6 is based a recent result in Nourdin and Rosinski [2014] which characterizes the asymptotic moment-independence of series of multiple Wiener-Itô integral vectors. We also note that in Proposition 5.3 (2) of Nourdin and Rosinski [2014], a special case of Theorem 2.3.6 with $J_S = J_L = 1$ and LRD part involving Hermite rank $k_{1,L} = 2$ is treated. To go from moment-independence to independence, however, requires moment-determinacy of the limit, which we know holds when the Hermite rank $k = 1, 2$, that is, in the Gaussian and Rosenblatt cases. If some other *Hermite distribution*

(marginal distribution of Hermite process) $Z_d^{(k)}$ ($k \geq 3$) is moment-determinate, then we will allow $k_{j,L} = k$ in Theorem 2.3.6. So to this end, the moment-problem of general Hermite distributions is of great interest.

We conjecture the following:

Conjecture 2.3.8. *Theorem 2.3.6 holds without the restriction that $k_{j,L}$ be 1 or 2.*

This conjecture has been recently resolved by Nourdin et al. [2016]. We also show that the conjecture holds in the following special case:

Theorem 2.3.9. (*Gaussian linear process case.*) *Conjecture 2.3.8 holds when*

$$X_n = \sum_{i=1}^{\infty} a_i \epsilon_{n-i},$$

where ϵ_i 's are i.i.d. Gaussian and $\{a_i\}$ is regularly varying as $i \rightarrow \infty$ with exponent $d - 1$, $d \in (0, 1/2)$.

Theorem 2.3.9 is based on the arguments in Bai and Taquu [2013b] and its proof is sketched in Section 2.4.4. In Bai and Taquu [2013b] a different setup is considered: a multilinear polynomial-form process

$$U_n = \sum_{0 < i_1, \dots, i_k < \infty}^{\prime} a_{i_1} \dots a_{i_k} \epsilon_{n-i_1} \dots \epsilon_{n-i_k} \quad (2.21)$$

obtained by applying an off-diagonal multilinear polynomial-form filter to an i.i.d. sequence $\{\epsilon_i\}$, where \prime means exclusion of the diagonals $i_p = i_q$, $p \neq q$, and $\{a_i\}$ is regularly varying. The resulting sequence $\{X(n)\}$ will then display either short or long memory. Now consider a vector of such $X(n)$, whose components are defined through different $\{a_i\}$'s, that is, through different multilinear polynomial-form filters, but using the same $\{\epsilon_i\}$. What is the limit of the normalized partial sums of the vector? It is shown in Bai and Taquu [2013b] that the resulting limit is either a) a multivariate Gaussian process with Brownian motion as marginals, or b) a multivariate Hermite process, or c) a mixture of the two. One has a

similar limit structure as in the present chapter, but also asymptotic independence without restriction on the order k .

Note, however, that the setup (2.21) of Bai and Taqqu [2013b] is different from the case considered in the present chapter even if the ϵ_i 's are Gaussian. This is because, while one can set

$$a(x) = a_{[x]+1} \mathbf{1}_{\{x \geq 0\}}$$

and write

$$X_n = \sum_{i < n} a_{n-i} \epsilon_i \stackrel{d}{=} \int_{\mathbb{R}} a(n-x) W(dx),$$

one does *not* have

$$\sum'_{-\infty < i_1, \dots, i_k < n} a_{n-i_1} \dots a_{n-i_k} \epsilon_{i_1} \dots \epsilon_{i_k} \stackrel{d}{=} \int_{\mathbb{R}} a(n-[x_1]) \dots a(n-[x_k]) W(dx_1) \dots W(dx_k). \quad (2.22)$$

This is because the left-hand side of (2.22) excludes a large interval around the diagonals, which is not the case for the right-hand side. So the result of Bai and Taqqu [2013b] does not apply directly to the right-hand side of (2.22). Observe that this right-hand side falls within our framework because it equals $H_k(X_n)$.

Remark 2.3.10. As mentioned in Remark 2.2.3, the border case $d = \frac{1}{2}(1 - \frac{1}{k_j})$ often leads to convergence to Brownian motion as well. In fact, Theorem 2.3.1 and Theorem 2.3.6 continue to hold if we extend the definition of SRD to the case whenever the limit is Brownian motion regardless of the normalization.

In Theorem 2.3.1, Theorem 2.3.3 and Theorem 2.3.6 we stated the results only in terms of convergence in finite-dimensional distributions, but in fact they hold under weak convergence in $D[0, 1]^J$ (J-dimensional product space where $D[0, 1]$ is the space of Càdlàg functions on $[0, 1]$ with the uniform metric). If one can check that every component of $\mathbf{V}_N(t)$ is tight, then the vector $\mathbf{V}_N(t)$ is tight:

Lemma 2.3.11. *Univariate tightness in $D[0, 1]$ implies multivariate tightness in $D[0, 1]^J$.*

Proof. Suppose every component $X_{j,N}$ (a random element in $S = D[0, 1]$ with uniform metric d) of the J -dimensional random element \mathbf{X}_N is tight, that is, given any $\epsilon > 0$, there exists a compact set K_j in $D[0, 1]$, so that for all N large enough:

$$P(X_{j,N} \in K_j^c) < \epsilon$$

where K_j^c denotes the complement of K_j . If $K = K_1 \times \dots \times K_J$, then K is compact in the product space S^J . We can associate S^J with any compatible metric, e.g., for $\mathbf{X}, \mathbf{Y} \in S^J$,

$$d_m(\mathbf{X}, \mathbf{Y}) := \max_{1 \leq j \leq J} (d(X_j, Y_j)).$$

The sequence \mathbf{X}_N is tight on $D[0, 1]^J$ since

$$P(\mathbf{X}_N \in K^c) = P(\cup_{j=1}^J \{X_{j,N} \in K_j^c\}) \leq \sum_{j=1}^J P(X_{j,N} \in K_j^c) < J\epsilon.$$

□

The univariate tightness is shown in Taqqu [1979] for the LRD case. The tightness for the SRD case was considered in Chambers and Slud [1989] p. 328 and holds under the following additional assumption, that $\{G(X_n)\}$ is SRD, with

$$\sum_{k=1}^{\infty} 3^{k/2} (k!)^{1/2} |g_k| < \infty, \quad (2.23)$$

where g_k is the k -th coefficient of Hermite expansion (2.3) of G . Observe that (2.23) is a strengthening of the basic condition: $\mathbb{E}[G(X_0)^2] = \sum_{k=1}^{\infty} k! g_k^2 < \infty$. Hence we have:

Theorem 2.3.12. *Suppose that condition (2.23) holds for the short-range dependent components. Then the convergence in Theorem 2.3.1, Theorem 2.3.3, Theorem 2.3.6 and Theorem 2.3.9 holds as weak convergence in $D[0, 1]^J$.*

Condition (2.23) is satisfied in the important special case where G is a polynomial of

finite order.

2.4 Proofs of the multivariate convergence results

2.4.1 Proof of Theorem 2.3.1 (SRD case)

We start with a number of lemmas. The first yields the limit covariance structure in (2.11).

Lemma 2.4.1. *Assume that $\sum_n |\gamma(n)|^m < \infty$, then as $N \rightarrow \infty$:*

$$\frac{1}{N} \sum_{n_1=1}^{[Nt_1]} \sum_{n_2=1}^{[Nt_2]} \gamma(n_1 - n_2)^m \rightarrow (t_1 \wedge t_2) \sum_{n=-\infty}^{\infty} \gamma(n)^m. \quad (2.24)$$

Proof. Denote the left-hand side of (2.24) by S_N . Let $a = t_1 \wedge t_2$, and $b = t_1 \vee t_2$, and

$$S_{N,1} = \frac{1}{N} \sum_{n_1=1}^{[Na]} \sum_{n_2=1}^{[Na]} \gamma(n_1 - n_2)^m, \quad S_{N,2} = \frac{1}{N} \sum_{n_1=1}^{[Na]} \sum_{n_2=[Na]+1}^{[Nb]} \gamma(n_1 - n_2)^m,$$

so $S_N = S_{N,1} + S_{N,2}$. We have as $N \rightarrow \infty$,

$$S_{N,1} = a \sum_{n_1=-[Na]+1}^{[Na]-1} \frac{[Na] - |n|}{Na} \gamma(n)^m \rightarrow a \sum_{n=-\infty}^{\infty} \gamma(n)^m.$$

We hence need to show that $S_{N,2} \rightarrow 0$. Let $c(n) = \gamma(n)^m$, then

$$S_{N,2} \leq \frac{1}{N} \sum_{n_1=1}^{[Na]} \sum_{n_2=[Na]+1}^{[Nb]} |c(n_2 - n_1)| = \frac{1}{N} \sum_{n_1=1}^{[Na]} c_{N,n_1} = \int_0^a f_N(u) du,$$

where

$$c_{N,n_1} := \sum_{n_2=[Na]+1}^{[Nb]} |c(n_2 - n_1)| = \sum_{n_2=1}^{[Nb]-[Na]} |c([Na] + n_2 - n_1)|,$$

and for $u \in (0, a)$,

$$f_N(u) := \sum_{n_1=1}^{[Na]} c_{N,n_1} \mathbf{1}_{[\frac{n_1-1}{N}, \frac{n_1}{N})}(u)$$

$$\begin{aligned}
&= \sum_{n_2=1}^{[Nb]-[Na]} \sum_{n_1=1}^{[Na]} |c([Na] + n_2 - n_1)| \mathbf{1}_{\left[\frac{n_1-1}{N}, \frac{n_1}{N}\right)}(u) \\
&= \sum_{n_2=1}^{[Nb]-[Na]} |c([Na] - [Nu] - 1 + n_2)|.
\end{aligned}$$

Now observe that

$$f_N(u) \leq \sum_{n=-\infty}^{\infty} |c(n)| = \sum_{n=-\infty}^{\infty} |\gamma(n)|^m < \infty$$

and that $[Na] - [Nu] \rightarrow \infty$ as $N \rightarrow \infty$. Applying the Dominated Convergence Theorem, we deduce $f_N(u) \rightarrow 0$ on $(0, a)$. Applying the Dominated Convergence Theorem again, we conclude that $S_{N,2} \rightarrow 0$. \square

Now we introduce some notations, setting for $G \in L^2(\phi)$,

$$S_{N,t}(G) := \frac{1}{\sqrt{N}} \sum_{n=1}^{[Nt]} G(X_n). \quad (2.25)$$

The Hermite expansion of each G_j is

$$G_j = \sum_{m=k_j}^{\infty} g_{m,j} H_m \quad (2.26)$$

if G_j has Hermite rank k_j . Since we are in the pure SRD case, we have as in Remark 2.2.3, that the auto-covariance function $\gamma(n)$ of $\{X_n\}$

$$\sum_{n=-\infty}^{\infty} |\gamma(n)|^{k_j} < \infty, \quad \text{for } j = 1, \dots, J.$$

The following lemma states that it suffices to replace a general G_j with a finite linear combination of Hermite polynomials:

Lemma 2.4.2. *If Theorem 2.3.1 holds with a finite linear combination of Hermite polynomials $G_j = \sum_{m=k_j}^M a_{m,j} H_m$ for any $M \geq \max_j(k_j)$ and any $a_{m,j}$, then it also holds for any $G_j \in L^2(\phi)$.*

Proof. First we obtain an L^2 bound for $S_{N,t}(H_m)$. By $\mathbb{E}H_m(X)H_m(Y) = m!\mathbb{E}(XY)^m$ (Proposition 2.2.1 in Nourdin and Peccati [2012]), for $m \geq 1$,

$$\begin{aligned} \mathbb{E}(S_{N,t}(H_m))^2 &= \frac{1}{N} \sum_{n_1, n_2=1}^{[Nt]} \mathbb{E}H_m(X_{n_1})H_m(X_{n_2}) = \frac{m!}{N} \sum_{n_1, n_2=1}^{[Nt]} \gamma(n_1 - n_2)^m \\ &= tm! \sum_{n=1-[Nt]}^{[Nt]-1} \frac{[Nt] - |n|}{Nt} \gamma(n)^m \leq tm! \sum_{n=-\infty}^{\infty} |\gamma(n)|^m. \end{aligned} \quad (2.27)$$

Next, fix any $\epsilon > 0$. By (2.27) and $\|G\|_{L^2(\phi)}^2 = \sum_{m=0}^{\infty} g_m^2 m!$, for $M = M(\epsilon)$ large enough, one has

$$\begin{aligned} &\mathbb{E} \left| S_{N,t}(G_j) - S_{N,t} \left(\sum_{m=k_j}^M g_{m,j} H_m \right) \right|^2 = \mathbb{E} \left| S_{N,t} \left(\sum_{m=M+1}^{\infty} g_{m,j} H_m \right) \right|^2 \\ &= \sum_{m=M+1}^{\infty} g_{m,j}^2 \mathbb{E}(S_{N,t}(H_m))^2 \leq t \sum_{n=-\infty}^{\infty} |\gamma(n)|^{k_j} \sum_{m=M+1}^{\infty} g_{m,j}^2 m! \leq \epsilon t. \end{aligned}$$

Therefore, the J -vector

$$\mathbf{V}_{N,M}(t) = \left(S_{N,t} \left(\sum_{m=k_1}^M g_{m,1} H_m \right), \dots, S_{N,t} \left(\sum_{m=k_J}^M g_{m,J} H_m \right) \right)$$

satisfies $\limsup_N \mathbb{E} \|\mathbf{V}_{N,M}(t) - \mathbf{V}_N(t)\|^2 \leq J\epsilon t$, and thus

$$\lim_M \limsup_N \mathbb{E} \|\mathbf{V}_{N,M}(t) - \mathbf{V}_N(t)\|^2 = 0.$$

By assumption, we have as $N \rightarrow \infty$ $\mathbf{V}_{N,M}(t) \xrightarrow{f.d.d.} \mathbf{B}_M(t) = (B_{M,1}, \dots, B_{M,J})$, where the multivariate Gaussian $\mathbf{B}_M(t)$ has (scaled) Brownian motions as marginals with a covariance structure computed using Lemma 2.4.1 as follows:

$$\mathbb{E}(B_{M,j_1}(t_1)B_{M,j_2}(t_2)) = \lim_{N \rightarrow \infty} \mathbb{E} \left(S_{N,t_1} \left(\sum_{m=k_{j_1}}^M g_{m,j_1} H_m \right) S_{N,t_2} \left(\sum_{m=k_{j_2}}^M g_{m,j_2} H_m \right) \right)$$

$$\begin{aligned}
&= \lim_{N \rightarrow \infty} \sum_{m=k_{j_1} \vee k_{j_2}}^M g_{m,j_1} g_{m,j_2} m! \sum_{n_1=1}^{[Nt_1]} \sum_{n_2=1}^{[Nt_2]} \gamma(n_1 - n_2)^m \\
&= (t_1 \wedge t_2) \sum_{m=k_{j_1} \vee k_{j_2}}^M g_{m,j_1} g_{m,j_2} m! \sum_{n=-\infty}^{\infty} \gamma(n)^m.
\end{aligned}$$

Furthermore, as $M \rightarrow \infty$, $\mathbf{B}_M(t)$ tends in f.d.d. to $\mathbf{B}(t)$, which is a multivariate Gaussian process with the following covariance structure:

$$\mathbb{E}(B_{j_1}(t_1)B_{j_2}(t_2)) = (t_1 \wedge t_2) \sum_{m=k_{j_1} \vee k_{j_2}}^{\infty} g_{m,j_1} g_{m,j_2} m! \sum_{n=-\infty}^{\infty} \gamma(n)^m.$$

Therefore, applying the triangular argument in Billingsley [1999] Theorem 3.2, we have

$$\mathbf{V}_N(t) \xrightarrow{f.d.d.} \mathbf{B}(t).$$

□

The proof of Theorem 2.3.1 about the pure SRD case relies on Nourdin and Peccati [2012] Theorem 6.2.3, which says that for multiple Wiener-Itô integrals, univariate convergence to normal random variables implies joint convergence to a multivariate normal. We state it as follows:

Lemma 2.4.3. *Let $J \geq 2$ and k_1, \dots, k_j be some fixed positive integers. Consider vectors*

$$\mathbf{V}_N = (V_{N,1}, \dots, V_{N,J}) := (I_{k_1}(f_{N,1}), \dots, I_{k_J}(f_{N,J}))$$

with $f_{N,j}$ in $L^2(\mathbb{R}^{k_j})$. Let C be a symmetric non-negative definite matrix such that

$$\mathbb{E}(V_{N,i}V_{N,j}) \rightarrow C(i, j).$$

Then the univariate convergence as $N \rightarrow \infty$

$$V_{N,j} \xrightarrow{d} N(0, C(j, j)) \quad j = 1, \dots, J$$

implies the joint convergence

$$\mathbf{V}_N \xrightarrow{d} N(\mathbf{0}, C).$$

We now prove Theorem 2.3.1.

Proof. Take time points t_1, \dots, t_I , let $\mathbf{V}_N(t)$ be the vector in (2.2) in the context of Theorem 2.3.1, with G_j replaced by a finite linear combination of Hermite polynomials (Lemma 2.4.2). Thus

$$\mathbf{V}_N(t_i) = \left(\sum_{m=k_1}^M \frac{g_{m,1}}{A_1(N)} S_{N,t_i}(H_m), \dots, \sum_{m=k_J}^M \frac{g_{m,J}}{A_J(N)} S_{N,t_i}(H_m) \right). \quad (2.28)$$

We want to show the joint convergence

$$\left(\mathbf{V}_N(t_1), \dots, \mathbf{V}_N(t_I) \right) \xrightarrow{d} \left(\mathbf{B}(t_1), \dots, \mathbf{B}(t_I) \right) \quad (2.29)$$

with $\mathbf{B}(t)$ being the J -dimensional Gaussian process with covariance structure given by (2.11).

By (2.6), and because the term

$$\frac{g_{m,j}}{A_j(N)} S_{N,t_i}(H_m)$$

involves the m -th order Hermite polynomial only, we can represent it as an m -tuple Wiener-Itô integral:

$$\frac{g_{m,j}}{A_j(N)} S_{N,t_i}(H_m) =: I_m(f_{N,m,i,j})$$

for some square-integrable function $f_{N,m,i,j}$. Now

$$\mathbf{V}_N(t_i) = \left(\sum_{m=k_1}^M I_m(f_{N,m,i,1}), \dots, \sum_{m=k_J}^M I_m(f_{N,m,i,J}) \right) \quad (2.30)$$

To show (2.29), one only needs to show that as $N \rightarrow \infty$, $(I_m(f_{N,m,i,j}))_{m,i,j}$ converges jointly to a multivariate normal with the correct covariance structure.

Note by the univariate SRD result, namely, Theorem 2.2.1, each

$$I_m(f_{N,m,i,j}) = \frac{g_{m,j}}{A_j(N)} S_{N,t_i}(H_m)$$

converges to a univariate normal. Therefore, by Lemma 2.4.3, it's sufficient to show the covariance structure of $(I_m(f_{N,m,i,j}))_{m,i,j}$ is consistent with the covariance structure of $(B_j(t_i))_{i,j}$ as $N \rightarrow \infty$.

Note that $A_j(N) = \sigma_j N^{1/2}$ where σ_j is found in (2.12). If $m_1 \neq m_2$,

$$EI_{m_1}(f_{N,m,i_1,j_1})I_{m_2}(f_{N,m,i_2,j_2}) = \frac{g_{m_1,j_1}g_{m_2,j_2}}{\sigma_{j_1}\sigma_{j_2}N} \mathbb{E} \left(S_{N,t_{i_1}}(H_{m_1})S_{N,t_{i_2}}(H_{m_2}) \right) = 0.$$

If $m_1 = m_2 = m$,

$$\begin{aligned} & \mathbb{E}I_m(f_{N,m,i_1,j_1})I_m(f_{N,m,i_2,j_2}) \\ &= \frac{g_{m,j_1}g_{m,j_2}}{\sigma_{j_1}\sigma_{j_2}} \frac{1}{N} \sum_{n_1=1}^{[Nt_{i_1}]} \sum_{n_2=1}^{[Nt_{i_2}]} \mathbb{E}(H_m(X_{n_1})H_m(X_{n_2})) \\ &= \frac{m!g_{m,j_1}g_{m,j_2}}{\sigma_{j_1}\sigma_{j_2}} \frac{1}{N} \sum_{n_1=1}^{[Nt_{i_1}]} \sum_{n_2=1}^{[Nt_{i_2}]} \gamma(n_1 - n_2) \\ &\rightarrow \frac{t_{i_1} \wedge t_{i_2}}{\sigma_{j_1}\sigma_{j_2}} g_{m,j_1}g_{m,j_2} m! \sum_{n=-\infty}^{\infty} \gamma(n)^m \text{ as } N \rightarrow \infty \end{aligned}$$

by Lemma 2.4.1.

Since every component of \mathbf{V}_N in (2.28) is the sum of multiple Wiener-Itô integrals, it

follows that

$$\mathbb{E}V_{N,j_1}(t_{i_1})V_{N,j_2}(t_{i_2}) \rightarrow \frac{t_{i_1} \wedge t_{i_2}}{\sigma_{j_1}\sigma_{j_2}} \sum_{m=k_{j_1} \vee k_{j_2}}^M g_{m,j_1}g_{m,j_2}m! \sum_{n=-\infty}^{\infty} \gamma(n)^m,$$

which is the covariance in (2.11), where here M is finite due to Lemma 2.4.2. \square

2.4.2 Proof of Theorem 2.3.3 (LRD case)

The pure LRD case is proved by extending the proof in Dobrushin and Major [1979] to the multivariate case. Set

$$S_{N,t}(G) = \sum_{n=1}^{[Nt]} G(X_n).$$

The normalization factor which makes the variance at $t = 1$ tend to 1 is

$$A_j(N) = a_j L(N)^{k_j/2} N^{1+k_j(d-1/2)}, \quad (2.31)$$

where the slowly varying function $L(N)$ stems from the auto-covariance function: $\gamma(n) = L(n)n^{2d-1}$ and where a_j is a normalization constant.

The Hermite expansion of each G_j is given in 2.26 The following reduction lemma shows that it suffices to replace G_j 's with corresponding Hermite polynomials.

Lemma 2.4.4. *If the convergence in (2.13) holds with $g_{k_j,j}H_{k_j}$ replacing G_j , then it also holds for G_j , $j = 1, \dots, J$.*

Proof. By the Cramér-Wold device, we want to show for every $(w_1, \dots, w_J) \in \mathbb{R}^J$, the following convergence:

$$\sum_{j=1}^J w_j \frac{S_{N,t}(G_j)}{A_j(N)} \xrightarrow{f.d.d.} \sum_{j=1}^J w_j Z_d^{(k_j)}(t).$$

Let $G_j^* = g_{k_j+1,j}H_{k_j+1} + g_{k_j+2,j}H_{k_j+2} + \dots$, then

$$\sum_{j=1}^J w_j \frac{S_{N,t}(G_j)}{A_j(N)} = \sum_{j=1}^J w_j \frac{S_{N,t}(g_{k_j,j}H_{k_j})}{A_j(N)} + \sum_{j=1}^J w_j \frac{S_{N,t}(G_j^*)}{A_j(N)}.$$

By the assumption of this lemma and by the Cramér-Wold device,

$$\sum_{j=1}^J w_j \frac{S_{N,t}(g_{k_j,j} H_{k_j})}{A_j(N)} \xrightarrow{f.d.d.} \sum_{j=1}^J w_j Z_d^{(k_j)}(t).$$

Hence it suffices to show that for any $t > 0$,

$$\mathbb{E} \left(\sum_{j=1}^J w_j \frac{S_{N,t}(G_j^*)}{A_j(N)} \right)^2 \rightarrow 0.$$

By the elementary inequality: $(\sum_{j=1}^J x_j)^2 \leq J \sum_{j=1}^J x_j^2$, it suffices to show that for each j ,

$$\mathbb{E} \left(\frac{S_{N,t}(G_j^*)}{A_j(N)} \right)^2 \rightarrow 0.$$

This is because the variance growth of G_j^* (see (2.7) and (2.9)) is at most

$$L_j^*([Nt])[Nt]^{(k_j+1)(2d-1)+2}$$

for some slowly varying function L_j^* , while the normalization

$$A_j(N)^2 = a_j^2 L_j(N)^{k_j} N^{k_j(2d-1)+2}$$

tends more rapidly to infinity. □

The following lemma extends Lemma 3 of Dobrushin and Major [1979] to the multivariate case. It states that if Lemma 3 of Dobrushin and Major [1979] holds in the univariate case in each component, then it holds in the multivariate joint case.

Lemma 2.4.5. *Let F_0 and F_N be symmetric locally finite Borel measures without atoms on \mathbb{R} so that $F_N \rightarrow F$ weakly. Let W_{F_N} and W_{F_0} be complex Hermitian Gaussian measures with control measures F_N and F_0 respectively.*

Let $K_{N,j}$ be a series of Hermitian($K(-\mathbf{x}) = \overline{K(\mathbf{x})}$) measurable functions of k_j variables

tending to a continuous function $K_{0,j}$ uniformly in any compact set in \mathbb{R}^{k_j} as $N \rightarrow \infty$.

Moreover, suppose the following uniform integrability type condition holds for every $j = 1, \dots, J$:

$$\lim_{A \rightarrow \infty} \sup_N \int_{\mathbb{R}^{k_j} \setminus [-A, A]^{k_j}} |K_{N,j}(\mathbf{x})|^2 F_N(dx_1), \dots, F_N(dx_{k_j}) = 0. \quad (2.32)$$

Then we have the joint convergence:

$$\left(I_{k_1}^{(N)}(K_{N,1}), \dots, I_{k_J}^{(N)}(K_{N,J}) \right) \xrightarrow{d} \left(I_{k_1}^{(0)}(K_{0,1}), \dots, I_{k_J}^{(0)}(K_{0,J}) \right). \quad (2.33)$$

where $I_k^{(N)}(\cdot)$ denotes a k -tuple Wiener-Itô integral with respect to complex Gaussian random measure W_{F_N} , $N = 0, 1, 2, \dots$

Proof. By the Cramér-Wold device, we need to show that for every $(w_1, \dots, w_J) \in \mathbb{R}^J$ as $N \rightarrow \infty$,

$$X_N := \sum_{j=1}^J w_j I_{k_j}^{(N)}(K_{N,j}) \xrightarrow{d} X_{0,0} := \sum_{j=1}^J w_j I_{k_j}^{(0)}(K_{0,j}). \quad (2.34)$$

We show first that (2.34) holds when replacing all kernels with simple Hermitian functions g_j of the form:

$$g_j(u_1, \dots, u_{k_j}) = \sum_{i_1, \dots, i_{k_j}=1}^n a_{i_1, \dots, i_{k_j}} 1_{A_{i_1,j} \times \dots \times A_{i_{k_j},j}}(u_1, \dots, u_{k_j}),$$

where $A_{i,j}$'s are bounded Borel sets in \mathbb{R} satisfying $F_0(\partial A_{i,j}) = 0$, $a_{i_1, \dots, i_{k_j}} = 0$ if any two of i_1, \dots, i_{k_j} are equal, and $\overline{g(\mathbf{u})} = g(-\mathbf{u})$. We claim that

$$\sum_{j=1}^s w_j I_{k_j}^{(N)}(g_j) \xrightarrow{d} \sum_{j=1}^s w_j I_{k_j}^{(0)}(g_j). \quad (2.35)$$

Indeed, since $F_N \rightarrow F_0$ weakly and $F_0(\partial A_{i,j}) = 0$, we have as $N \rightarrow \infty$:

$$\mathbb{E}W_{F_N}(A_{i,j})W_{F_N}(A_{k,l}) = F_N(A_{i,j} \cap A_{k,l}) \rightarrow F_0(A_{i,j} \cap A_{k,l}) = \mathbb{E}W_{F_0}(A_{i,j})W_{F_N}(A_{k,l}),$$

thus $(W_{F_N}(A_{i,j}))_{i,j} \xrightarrow{d} (W_{F_0}(A_{i,j}))_{i,j}$ jointly. Since $\sum_{j=1}^s w_j I_{k_j}^{(N)}(g_j)$ is a polynomial of $W_{F_N}(A_{i,j})$, (2.35) holds by the Continuous Mapping Theorem.

Next, due to the atomlessness of F_N , the uniform convergence of $K_{N,j}$ to $K_{0,j}$ on any compact set, (2.32) and the continuity of $K_{0,j}$, for any $\epsilon > 0$, there exist simple Hermitian g_j 's $j = 1, \dots, J$ as above, such that for $N = 0$ and $N > N(\epsilon)$ (large enough),

$$\int_{\mathbb{R}^{k_j}} |K_{N,j}(x_1, \dots, x_{k_j}) - g_j(x_1, \dots, x_{k_j})|^2 F_N(dx_1) \dots F_N(dx_{k_j}) < \epsilon. \quad (2.36)$$

By (2.36) for every $j = 1, \dots, J$, we can find a sequence $g_{M,j}$ such that

$$\|I_{k_j}^{(0)}(K_{0,j}) - I_{k_j}^{(0)}(g_{M,j})\|_{L^2} < 1/M, \quad (2.37)$$

$$\|I_{k_j}^N(K_{N,j}) - I_{k_j}^N(g_j)\|_{L^2} < 1/M \text{ for } N > N(M) \text{ (large enough)}, \quad (2.38)$$

hence by (2.37)

$$X_{0,M} := \sum_{j=1}^J w_j I_{k_j}^{(0)}(g_{M,j}) \xrightarrow{d} X_{0,0} := \sum_{j=1}^J w_j I_{k_j}^{(0)}(K_0) \text{ as } M \rightarrow \infty. \quad (2.39)$$

and by (2.38),

$$\begin{aligned} & \lim_M \limsup_N \mathbb{E}|X_N - X_{N,M}|^2 \\ & := \lim_M \limsup_N \mathbb{E} \left| \sum_{j=1}^J w_j I_{k_j}^{(N)}(K_{N,j}) - \sum_{j=1}^J w_j I_{k_j}^{(N)}(g_{M,j}) \right|^2 = 0. \end{aligned} \quad (2.40)$$

Finally, replacing g_j by $g_{M,j}$ in (2.35), we have

$$X_{N,M} \xrightarrow{d} X_{0,M}. \quad (2.41)$$

Thus (2.34), namely, $X_N \xrightarrow{d} X_{0,0}$, follows now from (2.39), (2.40) and (2.41) and Theorem 3.2 of Billingsley [1999]. \square

We can now prove Theorem 2.3.3:

Proof. Since Lemma 2.4.5 involves only univariate assumptions and concludes with the desired multivariate convergence (2.33), one needs to treat only the univariate case. This is done in Dobrushin and Major [1979]. \square

2.4.3 Proof of Theorem 2.3.6 (SRD and LRD mixed case)

The following result from Nourdin and Rosinski [2014] will be used:

Theorem 2.4.6. (*Theorem 4.7 in Nourdin and Rosinski [2014].*) Consider

$$\begin{aligned}\mathbf{S}_N &= \left(I_{k_{1,S}}(f_{1,S,N}), \dots, I_{k_{J_S,S}}(f_{J_S,S,N}) \right), \\ \mathbf{L}_N &= \left(I_{k_{1,L}}(f_{1,L,N}), \dots, I_{k_{J_L,L}}(f_{J_L,L,N}) \right),\end{aligned}$$

where $k_{j_S,S} > k_{j_L,L}$ for all $j_S = 1, \dots, J_S$ and $j_L = 1, \dots, J_L$.

Suppose that as $N \rightarrow \infty$, \mathbf{S}_N converges in distribution to a multivariate normal law, and \mathbf{L}_N converges in distribution to a multivariate law which has moment-determinate components, then there are independent random vectors \mathbf{Z} and \mathbf{H} , such that

$$(\mathbf{S}_N, \mathbf{L}_N) \xrightarrow{d} (\mathbf{Z}, \mathbf{H}).$$

A proof of Theorem 2.4.6 can be found in Section 2.6 (see Theorem 2.6.3).

Proof of Theorem 2.3.6. Using the reduction arguments of Lemma 2.4.2 and Lemma 2.4.4, we can replace $G_{j,S}$ in (2.16) with $\sum_{m=k_{j,S}}^M g_{m,j,S} H_m$, and we can replace $G_{j,L}$ in (2.17) with $g_{k_L,j,L} H_{k_L}$, where $k_{j,S} > k_{j,L} = 1$ or 2 are the corresponding Hermite ranks and $g_{m,j,S}$, $g_{k_L,j,L}$ are the corresponding coefficients of their Hermite expansions.

Fix finite time points t_i , $i = 1 \dots, I$, we need to consider the joint convergence of the following vector:

$$(S_{i,j_S,N}, L_{i,j_L,N})_{i,j_S,j_L} :=$$

$$\left(\frac{1}{A_{j_S,S}} \sum_{m=k_{j_S,S}}^M g_{m,j_S,S} S_{N,t_i}(H_m), \frac{1}{A_{j_L,L}} g_{k_L,j_L,L} S_{N,t_i}(H_{k_L}) \right)_{i,j_S,j_L}, \quad (2.42)$$

where $i = 1, \dots, I$, $j_S = 1, \dots, J_S$, $j_L = 1, \dots, J_L$.

As in the proof of Theorem 2.3.1, using (2.6), we express Hermite polynomials as multiple Wiener-Itô integrals:

$$S_{i,j_S,N} = \sum_{m=k_{j_S,S}}^M I_m(f_{m,i,j_S,N}), \quad L_{i,j_L,N} = \sum_{m=k_{j_L,L}}^M I_m(f_{m,i,j_L,N}),$$

where $f_{m,i,j_S,N}$, $f_{m,i,j_L,N}$ are some symmetric square-integrable functions.

Express the vector in (2.42) as $(\mathbf{S}_N, \mathbf{L}_N)$, where $\mathbf{S}_N := (S_{i,j_S,N})_{i,j_S}$, $\mathbf{L}_N := (L_{i,j_L,N})_{i,j_L}$.

By Theorem 2.3.1, \mathbf{S}_N converges in distribution to some multivariate normal distribution, and by Theorem 2.3.3, \mathbf{L}_N converges to a multivariate distribution with moment-determinate marginals, because by assumption the limits only involve Hermite rank $k = 1$ (normal distribution) and $k = 2$ (Rosenblatt distribution). The normal distribution is moment-determinate. The Rosenblatt distribution is also moment-determinate because it has an analytic characteristic function (Taqqu [1975] p.301).

We can now use Theorem 2.4.6 to conclude the proof. \square

2.4.4 Proof of Theorem 2.3.9 (Gaussian linear process case)

The proof below is a sketch, since the details are close to the proof of Theorem 3.5 of Bai and Taqqu [2013b].

Proof. Firstly, using Lemma 2.4.2 and Lemma 2.4.4, instead of considering the general nonlinear function G_j , it suffices to focus on a) *for the SRD part*: a finite linear combination of Hermite polynomials whose orders are higher or equal to Hermite rank of G_j ; b) *for the LRD part*: the single Hermite polynomial whose order is equal to the Hermite rank of G_j . In addition, it suffices to consider in the SRD component only the m -truncated version:

$$X_n^{(m)} = \sum_{i=1}^m a_i \epsilon_{n-i}.$$

Secondly, one can write $X_n = \int_{\mathbb{R}} a(n - [x])W(dx)$ and hence by Itô's formula

$$H_k(X(n)) = \int_{\mathbb{R}^k}' a(n - [x_1]) \dots a(n - [x_k])W(dx_1) \dots W(dx_k),$$

where $a(x) = a_{[x]+1}1_{\{x \geq 0\}}$, and $W(\cdot)$ is a Brownian random measure.

The sequence $\{H_k(X_n^{(m)}), n \geq 1\}$ with $k \geq 2$ for the SRD component is always uncorrelated with $W(\cdot)$ since they belong to different Wiener chaoses, and since $\{H_k(X_n^{(m)}), n \geq 1\}$ is m -dependent, the Functional Central Limit Theorem applies, yielding a limit Brownian motion independent of $W(\cdot)$. The Non-Central Limit Theorem for the LRD part in this case holds by Theorem 4.7.1 of Giraitis et al. [2012]. The random measure which defines the limit Hermite processes is exactly the same $W(\cdot)$ as above. Thus the limit Brownian motions for the SRD component and the limit Hermite processes for the LRD component are independent. \square

2.5 Invariance of the joint distribution among different representations of the Hermite process

The Hermite process admits four different representations (Pipiras and Taqqu [2010]):

Let $B(\cdot)$ be the real Gaussian random measure and $W(\cdot)$ be the complex Gaussian random measure, as defined in Section 6 of Taqqu [1979]. $H_0 \in (1 - 1/(2k), 1)$.

1. Time domain representation:

$$Z_{H_0}^{(k)}(t) = a_{k,H_0} = \int_{\mathbb{R}^k}' \left(\int_0^t \prod_{j=1}^k (s - x_j)_+^{H_0 - 3/2} ds \right) B(dx_1) \dots B(dx_k) \quad (2.43)$$

2. Spectral domain representation:

$$Z_{H_0}^{(k)}(t) = b_{k,H_0} \int_{\mathbb{R}^k}'' \frac{e^{i(x_1 + \dots + x_k)t} - 1}{i(x_1 + \dots + x_k)} \prod_{j=1}^k |x_j|^{1/2 - H_0} W(dx_1) \dots W(dx_k) \quad (2.44)$$

3. Positive half-axis representation:

$$Z_{H_0}^{(k)}(t) = c_{k,H_0} \int'_{[0,\infty)^k} \left(\int_0^t \prod_{j=1}^k x_j^{1/2-H_0} (1-sx_j)_+^{H_0-3/2} ds \right) B(dx_1) \dots B(dx_k) \quad (2.45)$$

4. Finite interval representation:

$$Z_{H_0}^{(k)}(t) = d_{k,H_0} \int'_{[0,t]^k} \left(\prod_{j=1}^k x_j^{1/2-H_0} \int_0^t x^{k(H_0-1/2)} \prod_{j=1}^k (s-x_j)_+^{H_0-3/2} ds \right) B(dx_1) \dots B(dx_k) \quad (2.46)$$

where $a_{k,H_0}, b_{k,H_0}, c_{k,H_0}, d_{k,H_0}$ are constant coefficients to guarantee that $\text{Var}(Z_{H_0}^{(k)}(1)) = 1$, given in (1.17) and (1.18) of Pipiras and Taquq [2010].

Keep H_0 fixed throughout. We will prove the following:

Theorem 2.5.1. *The joint distribution of a vector made up of Hermite processes of possibly different orders k , but sharing the same random measure $B(\cdot)$ or $W(\cdot)$ in their Wiener-Itô integral representations, remains the same when switching from one of the above representations to another.*

The following notations are used to denote Wiener-Itô integrals with respect to $B(\cdot)$ and $W(\cdot)$ respectively:

$$I(f) := \int'_{\mathbb{R}^k} f(x_1, \dots, x_k) dB(x_1) \dots dB(x_k),$$

$$\tilde{I}(g) := \int''_{\mathbb{R}^k} g(\omega_1, \dots, \omega_k) dW(\omega_1) \dots dW(\omega_k).$$

where $'$ indicates that we don't integrate on $x_i = x_j, i \neq j$, $''$ indicates that we don't integrate on $\omega_i = \pm\omega_j, i \neq j$, f is a symmetric function and g is an Hermitian function ($g(\omega) = \overline{g(-\omega)}$).

The next lemma establishes the equality in joint distribution between time domain representation (2.43) and spectral domain representation (2.44), which is a multivariate extension of Lemma 6.1 in Taqqu [1979].

Lemma 2.5.2. *Suppose that $A_j(x_1, \dots, x_{k_j})$ is a symmetric function in $L^2(\mathbb{R}^{k_j})$, $j = 1, \dots, J$. Let $\tilde{A}_j(x_1, \dots, x_{k_j})$ be its L^2 -Fourier transform:*

$$\tilde{A}_j(\omega_1, \dots, \omega_{k_j}) = \frac{1}{(2\pi)^{k_j/2}} \int_{\mathbb{R}^{k_j}} \exp(i \sum_{n=1}^{k_j} x_n \omega_n) A_j(x_1, \dots, x_{k_j}) dx_1 \dots dx_{k_j}.$$

Then

$$(I_{k_1}(A_1), \dots, I_{k_J}(A_J)) \stackrel{d}{=} (\tilde{I}_{k_1}(\tilde{A}_1), \dots, \tilde{I}_{k_J}(\tilde{A}_J)).$$

Proof. The proof is a slight extension of the proof of Lemma 6.1 of Taqqu [1979]. The idea is to use a complete orthonormal set $\{\psi_i, i \geq 0\}$ in $L^2(\mathbb{R})$ to represent each A_j as an infinite polynomial form of order k_j with respect to ψ_i 's, as is done in (6.3) of Taqqu [1979]. Each $I_{k_j}(A_j)$ can be then written in the form of (6.4) of Taqqu [1979], which is essentially a function of

$$X_i := \int \psi_i(x) dB(x), \quad i \geq 0,$$

denoted

$$I_{k_j}(A_j) = K_j(\mathbf{X}),$$

where $\mathbf{X} = (X_0, X_1, \dots)$. Thus

$$(I_{k_1}(A_1), \dots, I_{k_J}(A_J)) = \mathbf{K}(\mathbf{X}), \tag{2.47}$$

where the vector function $\mathbf{K} = (K_1, \dots, K_J)$.

Now, \tilde{A}_j can also be written as an infinite polynomial form of order k_j with respect to

$\tilde{\psi}_i, i \geq 0$, where

$$\tilde{\psi}_i(\omega) = (2\pi)^{-1/2} \int e^{ix\omega} \psi_i(x) dx$$

is the L^2 -Fourier transform of ψ_i , as is done in (6.5) of Taqqu [1979]. Set

$$Y_j := \int \tilde{\psi}_j(\omega) dW(\omega), \quad j \geq 0.$$

Then, as in (6.6) of Taqqu [1979], we have

$$\tilde{I}_{k_j}(\tilde{A}_j) = K_j(\mathbf{Y}),$$

where K_j 's are the same as above, $\mathbf{Y} = (Y_0, Y_1, \dots)$, and thus

$$\left(\tilde{I}_{k_1}(\tilde{A}_1), \dots, \tilde{I}_{k_J}(\tilde{A}_J) \right) = \mathbf{K}(\mathbf{Y}). \quad (2.48)$$

By (2.47) and (2.48), it suffices to show that $\mathbf{X} \stackrel{d}{=} \mathbf{Y}$. This is true because by Parseval's identity, \mathbf{X} and \mathbf{Y} both consist of i.i.d. normal random variables with mean 0 and identical variance, . For details, see Taqqu [1979]. \square

We now complete the proof of Theorem 2.5.1. We still need to justify the equality in joint distribution between time domain representation (2.43) and positive half-axis representation (2.45) or finite interval representation (2.46).

First let's summarize the arguments of Pipiras and Taqqu [2010] for going from (2.43) to (2.45) or (2.46). The heuristic idea is that by changing the integration order in (2.43), one would have

$$\begin{aligned} Z_{H_0}^{(k)} &= \int_0^t \left(\int_{\mathbb{R}^k} \prod_{j=1}^k (s - x_j)^{H_0-3/2} B(dx_1) \dots B(dx_k) \right) ds \\ &= \int_0^t H_k \left(\int_{\mathbb{R}} (s - x)_+^{H_0-3/2} B(dx) \right) ds, \end{aligned} \quad (2.49)$$

where H_k is k -th Hermite polynomial. But in fact $g(x) := (s - x)_+^{H_0-3/2} \notin L^2(\mathbb{R})$, and

consequently $G(s) := \int_{\mathbb{R}} (s-x)_+^{H_0-3/2} B(dx)$ is not well-defined.

The way to get around this is to do a regularization, that is, to truncate $g(x)$ as $g_\epsilon(x) := g(x)1_{s-x>\epsilon}(x)$ for $\epsilon > 0$. Now the Gaussian process $G_\epsilon(t) := \int_{\mathbb{R}} g_\epsilon(x)B(dx)$ is well-defined. Next, after some change of variables, one gets the new desired representation of $G_\epsilon(t)$, say $G_\epsilon^*(t)$, where $G_\epsilon^*(t) \stackrel{d}{=} G_\epsilon(t)$. Setting $Z_{\epsilon, H_0}^{(k)}(t) = \int_0^t H_k(G_\epsilon(t))dt$ and $Z_{\epsilon, H_0}^{(k)*}(t) = \int_0^t H_k(G_\epsilon^*(t))dt$, yields

$$Z_{\epsilon, H_0}^{(k)}(t) \stackrel{d}{=} Z_{\epsilon, H_0}^{(k)*}(t). \quad (2.50)$$

Finally by letting $\epsilon \rightarrow 0$, one can show that $Z_{\epsilon, H_0}^{(k)}(t)$ converges in $L^2(\Omega)$ to the Hermite process $Z_{H_0}^{(k)}(t)$, while $Z_{\epsilon, H_0}^{(k)*}(t)$ converges in $L^2(\Omega)$ to some $Z_{H_0}^{(k)*}(t)$, which is the desired alternative representation of $Z_{H_0}^{(k)}(t)$.

The above argument relies on the stochastic Fubini theorem (Theorem 2.1 of Pipiras and Taqqu [2010]) which legitimates the change of integration order, that is, for $f(s, \mathbf{x})$ defined on $\mathbb{R} \times \mathbb{R}^k$, if

$$\int_{\mathbb{R}} \|f(s, \cdot)\|_{L^2(\mathbb{R}^k)} ds < \infty$$

(which is the case after regularization), then

$$\int_{\mathbb{R}^k} \int_{\mathbb{R}} f(s, x_1, \dots, x_k) ds B(dx_1) \dots B(dx_k) = \int_{\mathbb{R}} \int_{\mathbb{R}^k} f(s, x_1, \dots, x_k) B(dx_1) \dots B(dx_k) ds$$

almost surely.

Now, consider the multivariate case. Note that we still have equality of the the joint distributions as in (2.50) and the equality is preserved in the $L^2(\Omega)$ limit as $\epsilon \rightarrow 0$. Moreover, the stochastic Fubini theorem (Theorem 2.1 of Pipiras and Taqqu [2010]) extends naturally to the multivariate setting since the change of integration holds as an almost sure equality. Therefore one gets equality in joint distribution when switching from (2.43) to (2.45) or (2.46). \square

2.6 Asymptotic independence of Wiener-Itô integral vectors

We prove here Theorem 2.4.6 by extending a combinatorial proof of Nourdin and Rosinski [2014].

First, some background. In the papers Ustunel and Zakai [1989] and Kallenberg [1991], a criterion for independence between two random variables belonging to Wiener Chaos, say, $I_p(f)$ and $I_p(g)$, is given as

$$f \otimes_1 g = 0 \quad a.s. \quad (2.51)$$

where \otimes_1 means contraction of order 1 and is defined below.

The result of Nourdin and Rosinski [2014] involves the following problem: if one has sequences $\{f_n\}$, $\{g_n\}$, when will asymptotic independence hold between $I_p(f_n)$ and $I_q(g_n)$ as $n \rightarrow \infty$? Motivated by (2.51), one may guess that the criterion is $f_n \otimes_1 g_n \rightarrow 0$ as $n \rightarrow \infty$. This is, however, shown to be false by a counterexample in Nourdin and Rosinski [2014]: set $p = q = 2$, $f_n = g_n$ and assume that $I_2(f_n) \xrightarrow{d} Z \sim N(0, 1)$. One can then show that $f_n \otimes_1 f_n \rightarrow 0$, while obviously $(I_2(f_n), I_2(f_n)) \xrightarrow{d} (Z, Z)$. Let $\|\cdot\|$ denote the L^2 norm in the appropriate dimension and let $\langle \cdot, \cdot \rangle$ denote the corresponding inner product.

We now define contractions. The contraction \otimes_r between two symmetric square integrable functions f and g is defined as

$$(f \otimes_r g)(x_1, \dots, x_{p-r}, y_1, \dots, y_{q-r}) := \int_{\mathbb{R}^r} f(x_1, \dots, x_{p-r}, s_1, \dots, s_r) g(y_1, \dots, y_{q-r}, s_1, \dots, s_r) ds_1 \dots ds_r$$

If $r = 0$, the contraction is just the tensor product:

$$f \otimes_0 g = f \otimes g := f(x_1, \dots, x_p) g(y_1, \dots, y_q). \quad (2.52)$$

The symmetrized contraction $\tilde{\otimes}_r$ involves one more step, namely, the symmetrization of the function obtained from the contraction. This is done by summing over all permutations

of the variables and dividing by the number of permutations. Note that as the contraction is only defined for symmetric functions, replacing \otimes_r with $\tilde{\otimes}_r$ enables one to consider a sequence of symmetrized contractions of the form

$$\left(\dots \left((f_1 \tilde{\otimes}_{r_1} f_2) \tilde{\otimes}_{r_2} f_3 \right) \dots \right) \tilde{\otimes}_{r_{n-1}} f_n.$$

We will use the following product formula (Proposition 6.4.1 of Peccati and Taqqu [2011]) for multiple Wiener-Itô integrals

$$I_p(f)I_q(g) = \sum_{r=0}^{p \wedge q} r! \binom{p}{r} \binom{q}{r} I_{p+q-2r}(f \otimes_r g) \quad p, q \geq 0. \quad (2.53)$$

Because the symmetrization of the integrand doesn't change the multiple Wiener-Itô integral, \otimes_r could be replaced with $\tilde{\otimes}_r$ in the product formula.

For a vector $\mathbf{q} = (q_1, \dots, q_k)$, we denote $|\mathbf{q}| := q_1 + \dots + q_k$. By a suitable iteration of (2.53), we have the following multiple product formula:

$$\prod_{i=1}^k I_{q_i}(f_i) = \sum_{\mathbf{r} \in C(\mathbf{q}, k)} a(\mathbf{q}, k, \mathbf{r}) I_{|\mathbf{q}| - 2|\mathbf{r}|} \left(\dots (f_1 \tilde{\otimes}_{r_1} f_2) \dots \tilde{\otimes}_{r_{k-1}} f_k \right), \quad (2.54)$$

where $\mathbf{q} \in \mathbb{N}^n$, the index set

$$C(\mathbf{q}, k) = \left\{ \mathbf{r} \in \prod_{i=1}^{k-1} \{0, 1, \dots, q_{i+1}\} : r_1 \leq q_1, r_i \leq (q_1 + \dots + q_i) - 2(r_1 + \dots + r_{i-1}), i = 2, \dots, k-1 \right\},$$

and $a(\mathbf{q}, k, \mathbf{r})$ is some integer factor. The following Theorem 2.6.1 is similar to Theorem 3.4 of Nourdin and Rosinski [2014] but the proof is different¹.

Theorem 2.6.1. (Asymptotic Independence of Multiple Wiener-Itô Integral Vec-

¹The present proof of Theorem 2.6.1 is an extension to Wiener-Itô integral vectors of a combinatorial proof for Wiener-Itô integral scalars given in an original version of Nourdin and Rosinski [2014].

tors.) Suppose we have the joint convergence

$$(\mathbf{U}_{1,N}, \dots, \mathbf{U}_{J,N}) \xrightarrow{d} (\mathbf{U}_1, \dots, \mathbf{U}_J),$$

where

$$\mathbf{U}_{j,N} = \left(I_{q_{1,j}}(f_{1,j,N}), \dots, I_{q_{I_j,j}}(f_{I_j,j,N}) \right).$$

Assume

$$\lim_{N \rightarrow \infty} \|f_{i_1,j_1,N} \otimes_r f_{i_2,j_2,N}\| = 0 \quad (2.55)$$

for all $i_1, i_2, j_1 \neq j_2$, and

$$r = 1, \dots, q_{i_1,j_1} \wedge q_{i_2,j_2},$$

where $\|\cdot\|$ denotes the $L^2(\mathbb{R}^k)$ norm for some appropriate dimension k .

Then using the notation $\mathbf{u}^{\mathbf{k}} = u_1^{k_1} \dots u_m^{k_m}$, we have

$$\mathbb{E}[\mathbf{U}_1^{\mathbf{k}_1} \dots \mathbf{U}_J^{\mathbf{k}_J}] = \mathbb{E}[\mathbf{U}_1^{\mathbf{k}_1}] \dots \mathbb{E}[\mathbf{U}_J^{\mathbf{k}_J}] \quad (2.56)$$

for all $\mathbf{k}_j \in \mathbb{N}^{I_j}$

Moreover, if every component of every \mathbf{U}_j is moment-determinate, then $\mathbf{U}_1, \dots, \mathbf{U}_J$ are independent.

Proof. The index $i = 1, \dots, I_j$ refers to the components within the vector $\mathbf{U}_{j,N}$, $j = 1, \dots, J$. For notational simplicity, we let $I_j = I$, that is, each $\mathbf{U}_{j,N}$ has the same number of components.

Let $|\mathbf{k}|$ denote the sum of its components $k_1 + \dots + k_m$. First to show (2.56), it suffices to show

$$\lim_{N \rightarrow \infty} \mathbb{E} \prod_{j=1}^J (\mathbf{U}_{j,N}^{\mathbf{k}_j} - \mathbb{E}[\mathbf{U}_{j,N}^{\mathbf{k}_j}]) = 0$$

for any $|\mathbf{k}_1| > 0, \dots, |\mathbf{k}_J| > 0$. Note that $\mathbf{U}_{j,N}^{\mathbf{k}_j} = U_{1,j,N}^{k_{1,j}} \dots U_{I,j,N}^{k_{I,j}}$ is a scalar.

By (2.54), one gets

$$I_q(f)^k = \sum_{\mathbf{r} \in C_{q,k}} a(q, k, \mathbf{r}) I_{kq-2|\mathbf{r}|} (\dots (f \tilde{\otimes}_{r_1} f) \dots \tilde{\otimes}_{r_{k-1}} f)$$

where $a(q, k, r)$'s are integer factors which don't play an important role, and $C_{q,k}$ is some index set. If

$$\mathbf{U}_{j,N}^{\mathbf{k}_j} = \prod_{i=1}^I I_{q_{i,j}} (f_{i,j,N})^{k_{i,j}},$$

then

$$\begin{aligned} \mathbf{U}_{j,N}^{\mathbf{k}_j} &= \prod_{i=1}^I \sum_{\mathbf{r} \in C_{q_{i,j}, k_{i,j}}} a(q_{i,j}, k_{i,j}, \mathbf{r}) I_{k_{i,j}q_{i,j}-2|\mathbf{r}|} \left(\dots (f_{i,j,N} \tilde{\otimes}_{r_1} f_{i,j,N}) \dots \tilde{\otimes}_{r_{k_{i,j}-1}} f_{i,j,N} \right) \\ &= \sum_{\mathbf{r}^1 \in C_{q_{1,j}, k_{1,j}}} \dots \sum_{\mathbf{r}^I \in C_{q_{I,j}, k_{I,j}}} \prod_{i=1}^I a(q_{i,j}, k_{i,j}, \mathbf{r}^i) I_{k_{i,j}q_{i,j}-2|\mathbf{r}^i|} (h_{i,j,N}) \end{aligned} \quad (2.57)$$

where

$$h_{i,j,N} = \left(\dots (f_{i,j,N} \tilde{\otimes}_{r_1^i} f_{i,j,N}) \dots \tilde{\otimes}_{r_{k_{i,j}-1}^i} f_{i,j,N} \right).$$

If one applies the product formula (2.54) to the product in (2.57), one gets that $\mathbf{U}_{j,N}^{\mathbf{k}_j}$ involves terms of the form $I_{|\mathbf{p}_j|-2|\mathbf{s}_j|}(H_{j,N})$ (\mathbf{p}_j and \mathbf{s}_j run through some suitable index sets), where

$$H_{j,N} = \left(\dots (h_{1,j,N} \tilde{\otimes}_{s_1} h_{2,j,N}) \dots \tilde{\otimes}_{s_{I-1}} h_{I,j,N} \right).$$

Since the expectation of a Wiener-Itô integral of positive order is 0 while a Wiener-Itô integral of zero order is a constant, $\mathbf{U}_{j,N}^{\mathbf{k}_j} - \mathbb{E}[\mathbf{U}_{j,N}^{\mathbf{k}_j}]$ involves $I_{|\mathbf{p}_j|-2|\mathbf{s}_j|}(H_{j,N})$ with $|\mathbf{p}_j| - 2|\mathbf{s}_j| > 0$ only. Therefore, every $H_{j,N}$ involved in the expression of $\mathbf{U}_{j,N}^{\mathbf{k}_j} - \mathbb{E}[\mathbf{U}_{j,N}^{\mathbf{k}_j}]$ has $n_j = |\mathbf{p}_j| - 2|\mathbf{s}_j| > 0$ variables.

Note that there are no products left at this point in the expression of $\mathbf{U}_{j,N}^{\mathbf{k}_j} - \mathbb{E}[\mathbf{U}_{j,N}^{\mathbf{k}_j}]$, only sums. But to compute $\mathbb{E} \prod_{j=1}^J (\mathbf{U}_{j,N}^{\mathbf{k}_j} - \mathbb{E}[\mathbf{U}_{j,N}^{\mathbf{k}_j}])$, one needs to apply the product formula (2.54) again and then compute the expectation. Since Wiener-Itô integrals of positive order have mean 0, taking the expectation involves focusing on the terms of zero

order which are constants. Since $f \otimes_p g = \langle f, g \rangle = EI_p(f)I_p(g)$ for functions f and g both having p variables, $\mathbb{E} \prod_{j=1}^J (\mathbf{U}_{j,N}^{\mathbf{k}_j} - \mathbb{E}[\mathbf{U}_{j,N}^{\mathbf{k}_j}])$ involves only terms of the form:

$$G_N = (\dots (H_{1,N} \tilde{\otimes}_{t_1} H_{2,N}) \dots \tilde{\otimes}_{t_{J-2}} H_{J-1,N}) \tilde{\otimes}_{t_{J-1}} H_{J,N} \quad (2.58)$$

$$= \int_{\mathbb{R}^{n_J}} (H_{1,N} \tilde{\otimes}_{t_1} H_{2,N}) \dots \tilde{\otimes}_{t_{J-2}} H_{J-1,N}) H_{J,N} \, d\mathbf{x} \quad (2.59)$$

where the contraction size vector $\mathbf{t} = (t_1, \dots, t_{J-1})$ runs through some index set. Since these contractions must yield a constant, we have

$$|\mathbf{t}| = \frac{1}{2}(n_1 + \dots + n_J) > 0, \quad (2.60)$$

where n_j is the number of variables of $H_{j,N}$. There is therefore at least one component (call it t) of \mathbf{t} which is strictly positive and thus there is a pair j_1, j_2 with $j_1 \neq j_2$, such that H_{j_1} and H_{j_2} that have at least one common argument.

One now needs to show that G_N in (2.59) tends to 0. This is done by applying the generalized Cauchy-Schwartz inequalities in Lemma 2.3 of Nourdin and Rosinski [2014] successively, through the following steps:

$$\begin{aligned} & \text{for any } j_1 \neq j_2, i_1, i_2 \text{ and } r > 0, \lim_{N \rightarrow \infty} \|f_{i_1, j_1, N} \otimes_r f_{i_2, j_2, N}\| = 0 \\ & \implies \text{for any } j_1 \neq j_2, i_1, i_2 \text{ and } s > 0, \lim_{N \rightarrow \infty} \|h_{i_1, j_1, N} \otimes_s h_{i_2, j_2, N}\| = 0 \\ & \implies \text{for any } j_1 \neq j_2 \text{ and } t > 0, \lim_{N \rightarrow \infty} \|H_{j_1, N} \otimes_t H_{j_2, N}\| = 0 \end{aligned} \quad (2.61)$$

$$\implies \lim_{N \rightarrow \infty} G_N = 0, \quad (2.62)$$

proving (2.56). Here we illustrate some details for going from (2.61) to (2.62), and omit the first two steps which use a similar argument.

Let $C = \{1, 2, \dots, (n_1 + \dots + n_J)/2\}$. Suppose c is a subset of C , then we use the notation \mathbf{z}_c to denote $\{z_{j_1}, \dots, z_{j_{|c|}}\}$ where $\{j_1, \dots, j_{|c|}\} = c$ and $|c|$ is the cardinality of c . When $c = \emptyset$, $\mathbf{z}_c = \emptyset$.

Observe that (2.59) is a sum (due to symmetrization) of terms of the form:

$$\int_{\mathbb{R}^{|C|}} H_{1,N}(\mathbf{z}_{c_1}) \dots H_{J,N}(\mathbf{z}_{c_J}) d\mathbf{z}_C, \quad (2.63)$$

where every c_j , $j = 1, \dots, J$, is a subset of C . Note that since $|t| = t_1 + \dots + t_J > 0$ in (2.60), there must exist $j_1 \neq j_2 \in \{1, \dots, J\}$, such that $c_0 := c_{j_1} \cap c_{j_2} \neq \emptyset$. By the generalized Cauchy Schwartz inequality (Lemma 2.3 in Nourdin and Rosinski [2014]), one gets a bound for (2.63) as:

$$\left| \int_{\mathbb{R}^{|C|}} H_{1,N}(\mathbf{z}_{c_1}) \dots H_{J,N}(\mathbf{z}_{c_J}) d\mathbf{z}_C \right| \leq \|H_{j_1,N} \otimes_{|c_0|} H_{j_2,N}\| \prod_{j \neq j_1, j_2} \|H_{j,N}\|,$$

where $\|H_{j_1,N} \otimes_{|c_0|} H_{j_2,N}\| \rightarrow 0$ as $N \rightarrow \infty$ by (2.61). In addition, $\|f_{i,j,N}\|$, $N \geq 1$ are uniformly bounded due to the tightness of the distribution of $I_{k_{i,j}}(f_{i,j,N})$, $N \geq 1$ (Lemma 2.1 of Nourdin and Rosinski [2014]). This, by the generalized Cauchy-Schwartz inequality (Lemma 2.3 of in Nourdin and Rosinski [2014]), implies that $\|h_{i,j,N}\|$, $N \geq 1$ are uniformly bounded, which further implies the uniform boundedness of $\|H_{j,N}\|$, $N \geq 1$. Hence (2.63) goes to 0 as $N \rightarrow \infty$ and thus (2.62) holds.

Finally, if every component of every \mathbf{U}_j is moment-determinate, then by Theorem 3 of Petersen [1982], the distribution of $\mathbf{U} := (\mathbf{U}_1, \dots, \mathbf{U}_J)$ is determined by its joint moments. But by (2.56), the joint moments of \mathbf{U} are the same as if the \mathbf{U}_j 's were independent. Then the joint moment-determinancy implies independence. \square

Corollary 2.6.2. *With the notation of Theorem 2.6.1, suppose that condition (2.55) is satisfied and that as $N \rightarrow \infty$, each $\mathbf{U}_{j,N}$ converges in distribution to some multivariate law which has moment-determinate components. Then there are independent random vectors $\mathbf{U}_1, \dots, \mathbf{U}_J$ such that*

$$(\mathbf{U}_{1,N}, \dots, \mathbf{U}_{J,N}) \xrightarrow{d} (\mathbf{U}_1, \dots, \mathbf{U}_J). \quad (2.64)$$

Proof. Since each $\mathbf{U}_{j,N}$ converges in distribution, the vector of vectors $(\mathbf{U}_{1,N}, \dots, \mathbf{U}_{J,N})$ is tight in distribution, so any of its subsequence has a further subsequence converging in

distribution to a vector $(\mathbf{U}_1, \dots, \mathbf{U}_J)$. But by Theorem 2.6.1, the \mathbf{U}_j 's are independent. Moreover, the convergence in distribution of each $\mathbf{U}_{j,N}$ implies that $\mathbf{U}_{j,N} \xrightarrow{d} \mathbf{U}_j$, and hence (2.64) holds. \square

Now we are in the position to state the result used in Theorem 2.3.6 in the proof of the SRD and LRD mixed case.

Theorem 2.6.3. *Consider*

$$\begin{aligned}\mathbf{S}_N &= \left(I_{k_{1,S}}(f_{1,S,N}), \dots, I_{k_{J_S,S}}(f_{J_S,S,N}) \right), \\ \mathbf{L}_N &= \left(I_{k_{1,L}}(f_{1,L,N}), \dots, I_{k_{J_L,L}}(f_{J_L,L,N}) \right),\end{aligned}$$

where $k_{j_S,S} > k_{j_L,L}$ for all $j_S = 1, \dots, J_S$ and $j_L = 1, \dots, J_L$.

Suppose that as $N \rightarrow \infty$, \mathbf{S}_N converges in distribution to a multivariate normal law, and \mathbf{L}_N converges in distribution to a multivariate law which has moment-determinate components, then there are independent random vectors \mathbf{Z} and \mathbf{H} , such that

$$(\mathbf{S}_N, \mathbf{L}_N) \xrightarrow{d} (\mathbf{Z}, \mathbf{H}).$$

Proof. By Corollary 2.6.2, we only need to check the contraction condition (2.55). This is done as in the proof of Theorem 4.7 of Nourdin and Rosinski [2014]. For the convenience of the reader, we present the argument here.

Using the identity

$$\|f \otimes_r g\|^2 = \langle f \otimes_{p-r} f, g \otimes_{q-r} g \rangle$$

where $r = 1, \dots, p \wedge q$, f and g have respectively p and q variables, we get for $r = 1, \dots, k_{i,L}$,

$$\begin{aligned}\|f_{i,S,N} \otimes_r f_{j,L,N}\|^2 &= \langle f_{i,S,N} \otimes_{k_{i,S}-r} f_{j,S,N}, f_{j,L,N} \otimes_{k_{j,L}-r} f_{j,L,N} \rangle \\ &\leq \|f_{i,S,N} \otimes_{k_{i,S}-r} f_{j,S,N}\| \|f_{j,L,N} \otimes_{k_{j,L}-r} f_{j,L,N}\| \rightarrow 0\end{aligned}$$

because $\|f_{i,S,N} \otimes_{k_{i,S}-r} f_{j,S,N}\| \rightarrow 0$ by the Nualart-Peccati Central Limit Theorem Nualart

and Peccati [2005], and for the second term, one has by Cauchy-Schwartz inequality,

$$\|f_{j,L,N} \otimes_{k_{j,L}-r} f_{j,L,N}\| \leq \|f_{j,L,N}\|^2$$

(generalized Cauchy-Schwartz inequality in Nourdin and Rosinski [2014] Lemma 2.3), which is bounded due to the tightness of the distribution of $I_{k_{j,L}}(f_{j,L,N})$ (Lemma 2.1 of Nourdin and Rosinski [2014]). Therefore (2.55) holds and the conclusion follows from Corollary 2.6.2. □

Chapter 3

Multivariate limits of multilinear polynomial-form processes with long memory

Consider a vector of multilinear polynomial-form processes with either short or long memory components. The components have possibly different coefficients but same noise elements. We study the limit of the normalized partial sums of the vector and identify the independent components.

3.1 Introduction

A linear process is generated by applying a linear time-invariant filter to i.i.d. random variables. A common model for stationary *long-range dependent* (LRD) (or *long-memory*) time series is a causal linear process with regularly varying coefficients as the lag tends to infinity, namely, $X(n) = \sum_{i=1}^{\infty} a_i \epsilon_{n-i}$, where the ϵ_i 's are i.i.d. with mean 0 and finite variance, and the coefficients satisfy

$$a_i = i^{d-1} L(i) \text{ with } 0 < d < 1/2,$$

and L is a slowly varying function at infinity (i.e., $L(x) > 0$ when x is large enough and $\lim_{x \rightarrow \infty} L(\lambda x)/L(x) = 1 \forall \lambda > 0$). Note that $0 < d < 1/2$ implies $\sum_{i=1}^{\infty} |a_i| = \infty$ but $\sum_{i=1}^{\infty} a_i^2 < \infty$, so $X(n)$ is well-defined in L^2 sense. It is well-known that the autocovariance $\gamma(n)$ of $X(n)$ is regularly varying with power $2d - 1$, and that the partial sum of $X(n)$

when suitably normalized converges to fractional Brownian motion with Hurst index

$$H = d + 1/2.$$

See for example Chapter 4.4 of Giraitis et al. [2012].

A family of processes related to multilinear processes are the so-called *multilinear polynomial-form processes* (or *discrete-chaos processes*), which are defined as

$$X(n) = \sum_{1 \leq i_1 < \dots < i_k < \infty} a_{i_1} \dots a_{i_k} \epsilon_{n-i_1} \dots \epsilon_{n-i_k}, \quad (3.1)$$

where

$$\sum_{i=1}^{\infty} a_i^2 < \infty,$$

and ϵ_i 's are i.i.d., and the $k > 0$ is the *order*. $X(n)$ is also said to belong to a *discrete chaos* of order k . The multilinear polynomial-form process $X(n)$ can be viewed as generated by nonlinear filters applied to i.i.d. random variables when $k > 1$. We call such a nonlinear filter defined in (3.1) a *multilinear polynomial-form filter*. Such a process often arises from considering a polynomial of a linear process (see, e.g., Surgailis [1982]).

If $a_i = i^{d-1}L(i)$ with $0 < d < 1/2$, when $k > 1$, that is, except for linear processes, the partial sum of $X(n)$ when suitably normalized no longer converges to a fractional Brownian motion, but depending on d and k , it either converges to a *Hermite process* if $X(n)$ is still LRD, or it converges to a Brownian motion if $X(n)$ is *short-range dependent* (SRD), that is, when the autocovariance of $X(n)$ is absolutely summable. See Giraitis et al. [2012] for more details.

In Statistics, however, one often needs convergence when $X(n)$ is a vector rather than a scalar. This leads us to the following question: if one applies different multilinear polynomial-form filters to the same i.i.d. sequence $\{\epsilon_i\}$, what is the joint limit behavior of the J -vector of the partial sums? More specifically, assume that $\{\epsilon_i\}$ are i.i.d with

mean 0 and variance 1. Consider the multilinear polynomial-form processes:

$$X_j(n) := \sum_{1 \leq i_1 < \dots < i_{k_j} < \infty} a_{i_1, j} \dots a_{i_{k_j}, j} \epsilon_{n-i_1} \dots \epsilon_{n-i_{k_j}}, \quad j = 1, \dots, J,$$

where k_1, \dots, k_J are orders for $X_1(n), \dots, X_J(n)$ respectively, $\{a_{i, j}\}$ are regularly varying coefficients. Let

$$Y_{j, N}(t) = \frac{1}{A_j(N)} \sum_{n=1}^{[Nt]} X_j(n), \quad t \geq 0, \quad (3.2)$$

where $A_j(N)$ is a normalization factor such that $\lim_{N \rightarrow \infty} \text{Var}[Y_{j, N}(1)] = 1$, $j = 1, \dots, J$.

We want to study the limit of the following vector process as $N \rightarrow \infty$:

$$\mathbf{Y}_N(t) := (Y_{1, N}(t), \dots, Y_{J, N}(t)). \quad (3.3)$$

Depending on $\{a_{i, j}\}$ and k_j , the components of $\mathbf{Y}_N(t)$ can be either purely SRD, or purely LRD, or a mixture of SRD and LRD. In Bai and Taqqu [2013a], a similar type of problem is considered for nonlinear functions of a LRD Gaussian process. We show here that the results for multilinear polynomial-form processes are similar to those in Bai and Taqqu [2013a]. But in the present context, we are able to provide a complete answer to the problem, in contrast to what happens in Bai and Taqqu [2013a], where the mixed SRD and LRD case is stated as a conjecture in some cases.

In addition, we distinguish here between two types of SRD sequences, one involving a linear process ($k = 1$) and one involving higher-order multilinear polynomial-form process ($k \geq 2$). For the first type of process, we get dependence with the LRD limit component, while for the second type, we get independence.

The chapter is organized as follows. In Section 3.2, some properties of multilinear polynomial-form processes are given and the univariate limit theorems under SRD and LRD are reviewed. In Section 3.3, we state the multivariate convergence results in three cases: a) pure SRD case, b) pure LRD case and c) mixed SRD and LRD case. The result of the general mixed case is stated in Theorem 3.3.5. In Section 3.4, we give the proofs of

the results in Section 3.3.

3.2 Preliminaries

In this section, we introduce some facts about multilinear polynomial-form processes as well as the univariate limit theorems for the partial sums.

Suppose that $X(n)$ is the multilinear polynomial-form process in (3.1). Note first, the condition $\sum_{i=1}^{\infty} a_i^2 < \infty$ guarantees that $X(n)$ is well-defined in L^2 , since

$$E[X(n)^2] = \sum_{1 \leq i_1 < \dots < i_k < \infty} a_{i_1}^2 \dots a_{i_k}^2 < \infty.$$

We use throughout a convention $a_i = 0$ for $i \leq 0$. One can compute the autocovariance of $X(n)$ as:

$$\gamma(n) = \sum_{1 \leq i_1 < \dots < i_k < \infty} a_{n+i_1} a_{i_1} \dots a_{n+i_k} a_{i_k}, \quad n \in \mathbb{Z}. \quad (3.4)$$

The following proposition describes the asymptotic behavior of $\gamma(n)$ under the assumption: $a_i = i^{d-1}L(i)$, $i \geq 1$, $0 < d < 1/2$.

Proposition 3.2.1. *Suppose $\gamma(n)$ is defined in (3.4), $a_i = i^{d-1}L(i)$, $i \geq 1$ with $0 < d < 1/2$ where L is slowly varying at infinity. Then*

$$\gamma(n) = L^*(n)n^{2d_X-1},$$

for some slowly varying function L^* and

$$d_X = \frac{1}{2} - k\left(\frac{1}{2} - d\right). \quad (3.5)$$

Proof. First we claim that as $n \rightarrow \infty$,

$$\sum_{i=1}^{\infty} a_{n+i} a_i \sim n^{2d-1} B(d, 1-2d) L(n)^2,$$

where $B(.,.)$ is the beta function. Indeed, one can check by Potter's bound for slowly varying functions (Theorem 1.5.6 in Bingham et al. [1989]) and the Dominated Convergence Theorem that as $n \rightarrow \infty$

$$\begin{aligned} \frac{1}{L(n)^{2n^{2d-1}}} \sum_{i=1}^{\infty} a_{n+i} a_i &= \sum_{i=1}^{\infty} \left(\frac{i}{n}\right)^{d-1} \left(1 + \frac{i}{n}\right)^{d-1} \frac{L(i)}{L(n)} \frac{L(n+i)}{L(n)} \frac{1}{n} \\ &\rightarrow \int_0^{\infty} u^{d-1} (1+u)^{d-1} du = B(d, 1-2d). \end{aligned} \quad (3.6)$$

Then note that as $n \rightarrow \infty$,

$$\gamma(n) \sim (k!)^{-1} \left(\sum_{i=1}^{\infty} a_{n+i} a_i \right)^k,$$

(the diagonal terms with $i_p = i_q$ are negligible as $n \rightarrow \infty$. See also Giraitis et al. [2012] p.109). Now we can deduce that

$$\gamma(n) = n^{k(2d-1)} L^*(n) = n^{2dx-1} L^*(n),$$

where

$$L^*(n) = (k!)^{-1} B(d, 1-2d)^k L(n)^{2k}.$$

□

Remark 3.2.2. According to Proposition 3.2.1, when $d < \frac{1}{2}(1 - \frac{1}{k})$ (or $k(2d-1) < -1$), we have $\sum |\gamma(n)| < \infty$, and when $d > \frac{1}{2}(1 - \frac{1}{k})$, we have $\sum |\gamma(n)| = \infty$. So if we assume $a_i = i^{d-1} L(i)$, $0 < d < 1/2$, the quantity $\frac{1}{2}(1 - \frac{1}{k})$ is the boundary between SRD and LRD.

We now define precisely what SRD and LRD mean for a multilinear polynomial-form process $X(n)$, and from then on we use this definition whenever we talk about SRD or LRD.

Definition 3.2.3. Let $X(n)$ be a multilinear polynomial-form process given in (3.1) with coefficient $\{a_i\}$, autocovariance $\gamma(n)$ and order k . We say that $X(n)$ is

(a) SRD, if for some $d \in (-\infty, \frac{1}{2}(1 - \frac{1}{k}))$ and some constant $c > 0$,

$$|a_i| \leq ci^{d-1}, \quad i \geq 1, \quad \sum_{n=-\infty}^{\infty} \gamma(n) > 0; \quad (3.7)$$

(b) LRD, if for some $d \in (\frac{1}{2}(1 - \frac{1}{k}), \frac{1}{2})$ and some L slowly varying at infinity,

$$a_i = i^{d-1}L(i), \quad i \geq 1, \quad \frac{1}{2}(1 - \frac{1}{k}) < d < 1/2. \quad (3.8)$$

Remark 3.2.4. The d in (3.7) and (3.8) are different. In the SRD case, $\{a_i\}$ is only assumed to decay faster than a power function, which implies

$$\sum_n |\gamma(n)| \leq \sum_n \left(\sum_{i=1}^{\infty} |a_{n+i}a_i| \right)^k < \infty$$

by (3.6), and the particular d chosen will not matter in the limit. While in the LRD case, the regularly varying assumption on $\{a_i\}$ yields a *memory parameter* $d_X = \frac{1}{2} - k(\frac{1}{2} - d)$ given by (3.5), and thus d plays an important role.

Next we consider the cross-covariance between of two multilinear polynomial-form processes obtained by applying two multilinear polynomial-form filters to the same $\{\epsilon_i\}$. In particular, set

$$X_1(n) = \sum_{1 \leq i_1 < \dots < i_p < \infty} a_{i_1} \dots a_{i_p} \epsilon_{n-i_1} \dots \epsilon_{n-i_p}, \quad (3.9)$$

$$X_2(n) = \sum_{1 \leq i_1 < \dots < i_q < \infty} b_{i_1} \dots b_{i_q} \epsilon_{n-i_1} \dots \epsilon_{n-i_q}. \quad (3.10)$$

$X_1(n)$ and $X_2(n)$ share the same $\{\epsilon_i\}$ but the sequences $\{a_i\}$ and $\{b_i\}$ can be different.

Then the cross-covariance is

$$\gamma_{1,2}(n) = \text{Cov}(X_1(n), X_2(0)) = \begin{cases} 0 & p \neq q; \\ \sum_{1 \leq i_1 < \dots < i_k < \infty} a_{i_1} b_{n+i_1} \dots a_{i_k} b_{n+i_k} & p = q = k \end{cases} \quad (3.11)$$

for any $n \in \mathbb{Z}$.

The following result will be used to obtain the asymptotic cross-covariance structure between the SRD components of $\mathbf{Y}_N(t)$ in (3.3).

Proposition 3.2.5. *Let $X_1(n)$ and $X_2(n)$ be given as in (3.9) and (3.10) with $p = q = k$, and are both SRD in the sense of Definition 3.2.3. Then the cross-covariance $\gamma_{1,2}(n) = \text{Cov}(X_1(n), X_2(0))$ is absolutely summable:*

$$\sum_{n=-\infty}^{\infty} |\gamma_{1,2}(n)| < \infty. \quad (3.12)$$

Moreover, (3.12) implies that as $N \rightarrow \infty$,

$$\text{Cov} \left(\frac{1}{\sqrt{N}} \sum_{n=1}^{\lfloor Nt_1 \rfloor} X_1(n), \frac{1}{\sqrt{N}} \sum_{n=1}^{\lfloor Nt_2 \rfloor} X_2(n) \right) \rightarrow (t_1 \wedge t_2) \sum_{n=-\infty}^{\infty} \gamma_{1,2}(n). \quad (3.13)$$

In addition, if $k = 1$, then

$$\sum_{n=-\infty}^{\infty} \gamma_{1,2}(n) = \sigma_1 \sigma_2, \quad (3.14)$$

where

$$\sigma_j^2 = \sum_n \text{Cov}(X_j(n), X_j(0)) = \lim_{N \rightarrow \infty} \text{Var} \left(\frac{1}{\sqrt{N}} \sum_{n=1}^N X_j(n) \right), \quad j = 1, 2.$$

Proof. Suppose that $\{a_i\}$ and $\{b_i\}$ satisfy the bound in (3.7) with $d = d_1$ and $d = d_2$ respectively. Using a similar argument as in the proof of Proposition 3.2.1, one can show that

$$|\gamma_{1,2}(n)| \leq |n|^{k(d_1+d_2-1)} L^*(n)$$

for some function $L^*(n)$ slowly varying at $\pm\infty$. Since by assumption $d_1, d_2 < \frac{1}{2}(1 - \frac{1}{k})$, which implies that $k(d_1 + d_2 - 1) < -1$, so we have $\sum_n |\gamma_{1,2}(n)| < \infty$.

The proof of (3.13) follows from the argument of Lemma 4.1 in Bai and Taqqu [2013a],

after noting that

$$\text{Cov} \left(\sum_{n=1}^{[Nt_1]} X_1(n), \sum_{n=1}^{[Nt_2]} X_2(n) \right) = \sum_{n_1=1}^{[Nt_1]} \sum_{n_2=1}^{[Nt_2]} \gamma_{1,2}(n_1 - n_2).$$

Now let's prove (3.14). When $k = 1$,

$$X_1(n) = \sum_{i=1}^{\infty} a_i \epsilon_{n-i}, \quad X_2(n) = \sum_{i=1}^{\infty} b_i \epsilon_{n-i}.$$

Note that by (3.7) with $k = 1$, we have $\sum_i |a_i| < \infty$ and $\sum_i |b_i| < \infty$. The cross-covariance is

$$\gamma_{1,2}(n) = \text{Cov}(X_1(n), X_2(0)) = \sum_{i=1}^{\infty} a_i b_{i+n}.$$

By Fubini,

$$\sum_{n=-\infty}^{\infty} \gamma_{1,2}(n) = \sum_{n=-\infty}^{\infty} \sum_{i=1}^{\infty} a_i b_{n+i} = \left(\sum_{i=1}^{\infty} a_i \right) \left(\sum_{n=1}^{\infty} b_n \right).$$

Since $(\sum_{i=1}^{\infty} a_i)^2 = \sum_n \gamma_1(n) = \sigma_1^2$, and $(\sum_{i=1}^{\infty} b_i)^2 = \sum_n \gamma_2(n) = \sigma_2^2$, we get Relation (3.14). \square

Let's now review the limit theorems for partial sum of a single multilinear polynomial-form process $X(n)$. Let the notation “ $\xrightarrow{f.d.d.}$ ” denote convergence in finite-dimensional distributions.

Theorem 3.2.6. *Suppose that $X(n)$ defined in (3.1) is SRD. Then*

$$\frac{1}{A(N)} \sum_{n=1}^{[Nt]} X(n) \xrightarrow{f.d.d.} B(t),$$

where $A(N)$ is a normalization factor to guarantee unit asymptotic variance at $t = 1$, and $B(t)$ is the standard Brownian motion. In fact,

$$A(N) \sim \sigma \sqrt{N} \text{ as } N \rightarrow \infty \text{ with } \sigma^2 = \sum_n \gamma(n).$$

Theorem 3.2.7. *Suppose that $X(n)$ defined in (3.1) is LRD. Then*

$$\frac{1}{A(N)} \sum_{n=1}^{[Nt]} X(n) \xrightarrow{f.d.d.} Z_d^{(k)}(t),$$

where $A(N)$ is a normalization factor to guarantee unit asymptotic variance at $t = 1$, and $Z_d^{(k)}(t)$ is the so-called Hermite process defined with the aid of the k -tuple Wiener-Itô stochastic integral denoted by $I_k(\cdot)$ (Major [2014]):

$$Z_d^{(k)}(t) = I_k(f_{k,d}^{(t)}) := \int_{\mathbb{R}^k}' f_{k,d}^{(t)}(x_1, \dots, x_k) W(dx_1) \dots W(dx_k) \quad (3.15)$$

where the prime ' indicates the exclusion of the diagonals $x_i = x_j$ for $i \neq j$, $W(\cdot)$ is Brownian random measure, and

$$f_{k,d}^{(t)}(x_1, \dots, x_k) = c_{k,d} \int_0^t \prod_{j=1}^k (s - x_j)_+^{d-1} ds, \quad (3.16)$$

with

$$c_{k,d} = \left(\frac{(k(d-1/2)+1)(2k(d-1/2)+1)\Gamma(1-d)^k}{k!\Gamma(d)^k\Gamma(1-2d)^k} \right)^{1/2}.$$

(See Pipiras and Taqqu [2010].) In fact,

$$A(N) \sim cN^{1+(d-1/2)k}L(N)^{k/2} \text{ as } N \rightarrow \infty \text{ for some } c > 0.$$

For the proofs of Theorem 3.2.6 and Theorem 3.2.7, we refer the reader to Chapter 4.8 in Giraitis et al. [2012], respectively Theorem 4.8.1 and Theorem 4.8.2¹. One may also compare Theorem 3.2.6 and Theorem 3.2.7 to their counterparts in the context of nonlinear functions of a LRD Gaussian process, stated as Theorem 2.1 and Theorem 2.2 in Bai and Taqqu [2013a].

¹The results of Chapter 4.8 in Giraitis et al. [2012] do not include a slowly varying function, nor convergence of finite-dimensional distributions in the case of Theorem 3.2.6. But they can be easily extended.

3.3 Multivariate convergence results

In this section, we state the multivariate joint convergence results for the vector process $\mathbf{Y}_N(t)$ in (3.3). Recall that \mathbf{Y}_N is normalized so that the asymptotic variance of every component at $t = 1$ equals 1.

Theorem 3.3.1. Pure SRD Case. *If all the components in \mathbf{Y}_N defined in (3.3) are SRD in the sense of (3.7), then*

$$\mathbf{Y}_N(t) \xrightarrow{f.d.d.} \mathbf{B}(t) = (B_1(t), \dots, B_J(t)),$$

where $\mathbf{B}(t)$ is a multivariate Gaussian process with $B_1(t), \dots, B_J(t)$ being standard Brownian motions with

$$\text{Cov}(B_p(s), B_q(t)) = (s \wedge t) \frac{\sigma_{p,q}}{\sigma_p \sigma_q}, \quad (3.17)$$

$$\begin{aligned} \sigma_p^2 &= \sum_{n=-\infty}^{\infty} \gamma_p(n) := \sum_{n=-\infty}^{\infty} \text{Cov}(X_p(n), X_p(0)), \\ \sigma_{p,q} &= \sum_{n=-\infty}^{\infty} \gamma_{p,q}(n) := \sum_{n=-\infty}^{\infty} \text{Cov}(X_p(n), X_q(0)). \end{aligned}$$

The normalization $A_j(N)$ in (3.2) satisfies $A_j(N) \sim \sigma_j \sqrt{N}$ as $N \rightarrow \infty$.

Remark 3.3.2. $\sigma_{p,q}$ is well-defined by Proposition 3.2.5.

Remark 3.3.3. In view of (3.11) and (3.17), if all the components of the $\mathbf{Y}_N(t)$ have different order, then the limit components $B_j(t)$ are uncorrelated and hence independent. Otherwise, they are in general dependent and their covariance is given by (3.17).

Theorem 3.3.4. Pure LRD Case. *If all the components in \mathbf{Y}_N defined in (3.3) are LRD in the sense of (3.8) with $d = d_1, \dots, d_J$ respectively, then*

$$\mathbf{Y}_N(t) \xrightarrow{f.d.d.} \mathbf{Z}_d^{\mathbf{k}}(t) = (Z_{d_1}^{(k_1)}(t), \dots, Z_{d_J}^{(k_J)}(t)),$$

where $Z_{d_j}^{(k_j)}(t)$ are Hermite processes sharing the same random measure $W(\cdot)$ in their Wiener-Itô integral representations. The normalization $A_j(N)$ in (3.2) satisfies

$$A_j(N) \sim c_j N^{1+(d_j-1/2)k_j} L(N)^{k_j/2} \text{ as } N \rightarrow \infty, \text{ for some } c_j > 0.$$

The processes $Z_{d_j}^{(k_j)}$, $j = 1, \dots, J$ are dependent.

We now consider the mixed SRD and LRD case.

Theorem 3.3.5. Mixed SRD and LRD Case. Break \mathbf{Y}_N in (3.3) into 3 parts:

$$\mathbf{Y}_N = (\mathbf{Y}_{N,S_1}, \mathbf{Y}_{N,S_2}, \mathbf{Y}_{N,L}),$$

where within \mathbf{Y}_{N,S_1} (J_{S_1} -dimensional) every component is SRD and has order $k_{j,S_1} = 1$, within \mathbf{Y}_{N,S_2} (J_{S_2} -dimensional) every component is SRD and has order $k_{j,S_2} \geq 2$, and within $\mathbf{Y}_{N,L}$ (J_L -dimensional) every component is LRD. Then

$$\mathbf{Y}_N(t) = (\mathbf{Y}_{N,S_1}(t), \mathbf{Y}_{N,S_2}(t), \mathbf{Y}_{N,L}(t)) \xrightarrow{f.d.d.} (\mathbf{W}(t), \mathbf{B}(t), \mathbf{Z}_{\mathbf{d}_L}^{\mathbf{k}_L}(t)), \quad (3.18)$$

where $\mathbf{B}(t) := (B_1(t), \dots, B_{J_{S_2}}(t))$ is the multivariate Gaussian process appearing in Theorem 3.3.1, $\mathbf{Z}_{\mathbf{d}_L}^{\mathbf{k}_L}(t)$ is the multivariate Hermite process appearing in Theorem 3.3.4,

$$\mathbf{W}(t) = (W(t), \dots, W(t)), \quad (3.19)$$

where $W(t)$ is the Brownian motion integrator for defining $\mathbf{Z}_{\mathbf{d}_L}^{\mathbf{k}_L}(t)$ (see (3.15)), and $\mathbf{B}(t)$ is independent of $(\mathbf{W}(t), \mathbf{Z}_{\mathbf{d}_L}^{\mathbf{k}_L}(t))$.

Remark 3.3.6. To understand heuristically why $\mathbf{B}(t)$ and $(\mathbf{W}(t), \mathbf{Z}_{\mathbf{d}_L}^{\mathbf{k}_L}(t))$ are independent, note that $\mathbf{Y}_{N,S_2}(t)$ belongs to chaos of order ≥ 2 , and is thus uncorrelated with $\mathbf{Y}_{N,S_1}(t)$ which belongs to first-order chaos, and also uncorrelated with the random noise $\{\epsilon_i\}$ which also belongs to the first-order chaos, and which after summing becomes asymptotically the Brownian measure $W(\cdot)$ defining $\mathbf{Z}_{\mathbf{d}_L}^{\mathbf{k}_L}(t)$.

Remark 3.3.7. The independence between $\mathbf{B}(t)$ and $\mathbf{Z}_{\mathbf{d}_L}^{\mathbf{k}_L}(t)$ for $k_{j,L} \geq 3$ (the order in LRD component) is in general only a conjecture in the framework of Bai and Taqqu [2013a]. This conjecture is resolved in the special case of causal linear Gaussian processes (Theorem 3.9 of Bai and Taqqu [2013a]) using arguments similar to the proof of Theorem 3.3.5 of the present chapter.

The convergence results in the above theorems are stated in terms of convergence in finite-dimensional distributions, but one can show that in some cases they extend to weak convergence in $D[0, 1]^J$ (J-dimensional product space where $D[0, 1]$ is the space of Càdlàg functions on $[0, 1]$ with uniform metric).

Theorem 3.3.8. *Weak convergence in $D[0, 1]^J$.*

1. *Theorem 3.3.4 holds with “ $\xrightarrow{f.d.d.}$ ” replaced by weak convergence in $D[0, 1]^J$;*
2. *If the SRD component in Theorem 3.3.1 (or Theorem 3.3.5) satisfies either of the following conditions:*
 - a. *There exists $m \geq 0$, such that the coefficients a_i in (3.1) are zero for all $i > m$;*
 - b. *$\{\epsilon_i\}$ are i.i.d. Gaussian;*
 - c. *The order $k = 1$ and $E(|\epsilon_i|^{2+\delta}) < \infty$ for some $\delta > 0$;*
 - d. *The order $k \geq 2$, $\sum_{i=1}^{\infty} |a_i| < \infty$ and $E(|\epsilon_i|^5) < \infty$;*

then Theorem 3.3.1 (or Theorem 3.3.5) holds with “ $\xrightarrow{f.d.d.}$ ” replaced by weak convergence in $D[0, 1]^J$.

Note that tightness in the SRD case results from an interplay between the dependence structure and the finiteness of the moments.

3.4 Proofs for the multivariate convergence results

3.4.1 Pure SRD case

Proof of Theorem 3.3.1. Following the idea of Giraitis et al. [2012] p.108., we define the truncated multilinear polynomial-form processes:

$$X_j^{(m)}(n) = \sum_{1 \leq i_1 < \dots < i_{k_j} \leq m} a_{i_1, j} \dots a_{i_{k_j}, j} \epsilon_{n-i_1} \dots \epsilon_{n-i_{k_j}}, \quad j = 1, \dots, J, \quad (3.20)$$

where $m > \max_j \{k_j\}$. Note that $X_j^{(m)}(n)$ is m -dependent. Set

$$(\sigma_j^{(m)})^2 = \sum_n \text{Cov} \left(X_j^{(m)}(n), X_j^{(m)}(0) \right)$$

(assume m is large enough so that $\sigma_j^{(m)} > 0$), and

$$\sigma_{p,q}^{(m)} = \sum_n \text{Cov} \left(X_p^{(m)}(n), X_q^{(m)}(0) \right)$$

which is well-defined due to Proposition 3.2.5.

Set

$$Y_{N,j}(t) := \frac{1}{\sigma_j \sqrt{N}} \sum_{n=1}^{[Nt]} X_j(n), \quad Y_{N,j}^{(m)}(t) := \frac{1}{\sigma_j^{(m)} \sqrt{N}} \sum_{n=1}^{[Nt]} X_j^{(m)}(n).$$

Theorem 3.3.1 follows if one shows that as $N \rightarrow \infty$,

$$\mathbf{Y}_N^{(m)}(t) =: \left(Y_{N,1}^{(m)}(t), \dots, Y_{N,J}^{(m)}(t) \right) \xrightarrow{f.d.d.} \mathbf{B}^{(m)}(t) := \left(B_1^{(m)}(t), \dots, B_J^{(m)}(t) \right), \quad (3.21)$$

where $B_j^{(m)}(t)$'s are Brownian motions with cross-covariance structure:

$$\text{Cov}(B_p^{(m)}(t_1), B_q^{(m)}(t_2)) = (t_1 \wedge t_2) \frac{\sigma_{p,q}^{(m)}}{\sigma_p^{(m)} \sigma_q^{(m)}}, \quad p, q = 1, \dots, J, \quad (3.22)$$

and as $m \rightarrow \infty$,

$$\sigma_j^{(m)} \rightarrow \sigma_j, \quad \sigma_{p,q}^{(m)} \rightarrow \sigma_{p,q} \quad (3.23)$$

as well as for any $j = 1, \dots, J$ and $t \geq 0$, as $m \rightarrow \infty$,

$$\text{Var} \left[Y_{N,j}^{(m)}(t) - Y_{N,j}(t) \right] \rightarrow 0 \quad (3.24)$$

uniformly in N . Indeed, combining (3.21), (3.23) and (3.24), one obtains the desired convergence:

$$\mathbf{Y}_N(t) = (Y_{N,1}(t), \dots, Y_{N,J}(t)) \xrightarrow{f.d.d.} \mathbf{B}(t) := (B_1(t), \dots, B_J(t)).$$

Relations (3.23) and (3.24) can be shown using the same type of arguments in Giraitis et al. [2012] p.108. We thus only need to show (3.21) and (3.22). By Crámer-Wold device, it suffices to show that for any $(c_1, \dots, c_J) \in \mathbb{R}^J$,

$$\sum_{j=1}^J c_j Y_{N,j}^{(m)}(t) = \frac{1}{\sqrt{N}} \sum_{n=1}^{[Nt]} \left(\sum_{j=1}^J \frac{c_j}{\sigma_j^{(m)}} X_j^{(m)}(n) \right) \xrightarrow{f.d.d.} \sum_{j=1}^J c_j B_j^{(m)}(t) =: G(t) \quad (3.25)$$

where $G(t)$ is a non-standardized Brownian motion. This follows from the fact that the sequence

$$\left\{ \sum_{j=1}^J \frac{c_j}{\sigma_j^{(m)}} X_j^{(m)}(n), n \geq 1 \right\}$$

is m -dependent and is thus subject to functional central limit theorem (Billingsley [1956] Theorem 5.2), which includes convergence in finite-dimensional distributions. The asymptotic cross-covariance structure (3.22) follows from Proposition 3.2.5. \square

3.4.2 Pure LRD case

Proof of Theorem 3.3.4. The joint convergence is proved by combining Theorem 4.8.2. and Proposition 14.3.3 of Giraitis et al. [2012], and the arguments leading to them.

The dependence between the limit Hermite processes with different orders is shown in Proposition 3.1 in Bai and Taqqu [2013a]. \square

3.4.3 Mixed SRD and LRD case

We prove Theorem 3.3.5 through a number of lemmas, one lemma implying the next.

Lemma 3.4.1. *Follow the notations and assumptions in Theorem 3.3.5. Let $X_{j,S_i}^{(m)}(n)$ be the m -truncated multilinear polynomial-form process (see (3.20)) corresponding to the components of \mathbf{Y}_{N,S_i} ($i = 1, 2$) in Theorem 3.3.5, where the orders satisfy $k_{j,S_1} = 1$ and $k_{j,S_2} \geq 2$. Let*

$$Y_{N,j,i}^{(m)}(t) := \frac{1}{\sigma_{j,S_i}^{(m)} \sqrt{N}} \sum_{n=1}^{[Nt]} X_{j,S_i}^{(m)}(n), \quad j = 1, \dots, J_i, \quad i = 1, 2,$$

where (assuming that m is large enough)

$$0 < (\sigma_{j,S_i}^{(m)})^2 := \sum_n \text{Cov}(X_{j,S_i}^{(m)}(n), X_{j,S_i}^{(m)}(0)) < \infty, \quad i = 1, 2.$$

Let

$$W_N(t) := N^{-1/2} \sum_{n=1}^{[Nt]} \epsilon_n, \quad \text{and} \quad \mathbf{Y}_{N,S_i}^{(m)}(t) = (Y_{N,1,i}^{(m)}(t), \dots, Y_{N,J_{S_i},i}^{(m)}(t)), \quad i = 1, 2.$$

Then

$$\left(\mathbf{Y}_{N,S_1}^{(m)}(t), \mathbf{Y}_{N,S_2}^{(m)}(t), W_N(t) \right) \xrightarrow{f.d.d.} \left(\mathbf{W}(t), \mathbf{B}^{(m)}(t), W(t) \right), \quad (3.26)$$

where $W(t)$ is a standard Brownian motion,

$$\mathbf{W}(t) = (W(t), \dots, W(t))$$

(J_{S_2} -dimensional), $\mathbf{B}^{(m)}(t)$ is as given in (3.21), namely, its components are standard Brownian motions with cross-covariance (3.22), and $\mathbf{B}^{(m)}(t)$ is independent of $(\mathbf{W}(t), W(t))$.

Proof. Fix any $\mathbf{w} = (a_1, \dots, a_{J_{S_1}}, b_1, \dots, b_{J_{S_2}}, c) \in \mathbb{R}^{J_{S_1}+J_{S_2}+1}$. By the Cramér-Wold device, we want to show that

$$\begin{aligned} R_N(t; \mathbf{w}) &:= \sum_j a_j Y_{N,j,1}^{(m)}(t) + \sum_j b_j Y_{N,j,2}^{(m)}(t) + cW_N(t) \\ &\xrightarrow{f.d.d.} \sum_j a_j W(t) + \sum_j b_j B_j^{(m)}(t) + cW(t) =: G(t), \end{aligned}$$

where $G(t)$ is a non-standardized Brownian motion whose marginal variance is the limit of the marginal variance of $R_N(t; \mathbf{w})$. Note that one can write

$$R_N(t; \mathbf{w}) = \frac{1}{\sqrt{N}} \sum_{n=1}^{\lfloor Nt \rfloor} U_{\mathbf{w}}^{(m)}(t),$$

where

$$U_{\mathbf{w}}^{(m)}(n) = \sum_{j=1}^{J_{S_1}} \frac{a_j}{\sigma_{j,S_1}^{(m)}} X_{j,S_1}^{(m)}(n) + \sum_{j=1}^{J_{S_2}} \frac{b_j}{\sigma_{j,S_2}^{(m)}} X_{j,S_2}^{(m)}(n) + ce_n^{(m)}$$

with

$$e_n^{(m)} = \sum_{i=(m-1)n+1}^{mn} \epsilon_i.$$

Since $\{U_{\mathbf{w}}^{(m)}(n)\}_n$ is m -dependent, the classical functional central limit theorem applies (Billingsley [1956]), yielding in the limit a Brownian motion $G(t)$ for $R_N(t; \mathbf{w})$. Now that the joint normality is shown, we only need to identify the asymptotic covariance structure as $N \rightarrow \infty$ of the left-hand side of (3.26) to the covariance structure of the right-hand side

of (3.26).

The independence between $\mathbf{B}^{(m)}(t)$ and $(\mathbf{W}(t), W(t))$ follows from the uncorrelatedness between $\mathbf{Y}_{N,S_2}^{(m)}(t)$ (involving chaos of order ≥ 2) and $(\mathbf{Y}_{N,S_1}^{(m)}(t), W_N(t))$ (involving chaos of order 1 only). The asymptotic covariance structure within $\mathbf{Y}_{N,S_2}^{(m)}(t)$ is given in (3.22) (apply Theorem 3.3.1 to $\mathbf{Y}_{N,S_2}^{(m)}$). Hence we are left to show that the asymptotic covariance structure of $(\mathbf{Y}_{N,S_1}^{(m)}(t), W_N(t))$ is that of $(\mathbf{W}(t), W(t))$. Note that in $(\mathbf{Y}_{N,S_1}^{(m)}(t), W_N(t))$, both $\{X_{j,S_1}^{(m)}(n)\}$ and $\{\epsilon_n\}$ are SRD linear processes. So applying (3.13) and (3.14) in Proposition 3.2.5 with $\sigma_1 = \sigma_2 = 1$, the desired asymptotic covariance structure is obtained. \square

Remark 3.4.2. Lemma 3.4.1 can be rephrased as follows: we define an *empirical random measure* on a finite interval Δ as:

$$W_N(\Delta) := \frac{1}{\sqrt{N}} \sum_{n/N \in \Delta} \epsilon_n.$$

Then the joint convergence in Lemma 3.4.1 still holds with $W(t)$ replaced by $(W_N(\Delta_1), \dots, W_N(\Delta_I))$ where $\Delta_i, i = 1, \dots, I$ are disjoint intervals, and $W(t)$ in the limit replaced by $(W(\Delta_1), \dots, W(\Delta_I))$ where $W(\cdot)$ is the Brownian random measure. Observe that while (3.26) involves convergence in distribution, the limit components $\mathbf{W}(t)$ and $W(t)$ both involve the same Brownian motion $W(t)$.

Now we adopt some notations from Giraitis et al. [2012] Chapter 14.3. Let $S_M(\mathbb{R}^k)$ be the class of simple functions defined on \mathbb{R}^k supported on a finite number of $1/M$ -cubes and vanishing on the diagonals. Suppose that h is a function defined on \mathbb{Z}^k which vanishes on diagonals. Let the polynomial form (or discrete multiple integral) with respect to h be

$$Q_k(h) = \sum_{i_1, \dots, i_k \in \mathbb{Z}} h(i_1, \dots, i_k) \epsilon_{i_1} \dots \epsilon_{i_k}, \quad (3.27)$$

where $\sum_{i_1, \dots, i_k} h(i_1, \dots, i_k)^2 < \infty$. The following lemma plays a key role in the proof of Theorem 3.3.5.

Lemma 3.4.3. Replace $(\mathbf{Y}_{N,S_1}^{(m)}(t), \mathbf{Y}_{N,S_2}^{(m)}(t), \mathbf{W}_N(t))$ in Lemma 3.4.1 by $(\mathbf{Y}_{N,S_1}^{(m)}(t), \mathbf{Y}_{N,S_2}^{(m)}(t), \mathbf{Q}_N)$, where $\mathbf{Q}_N = (Q_{k_1}(h_{1,N}), \dots, Q_{k_{J_L}}(h_{J_L,N}))$ and each $Q_{k_p}(h_{p,N})$, $p = 1, \dots, J_L$, is a polynomial-form defined in (3.27) with the same $\{\epsilon_i\}$ as those defining $\mathbf{Y}_{N,S_1}^{(m)}(t)$ and $\mathbf{Y}_{N,S_2}^{(m)}(t)$. Assume that the “normalized continuous extension” of $h_{p,N}$, that is,

$$\tilde{h}_{p,N}(x_1, \dots, x_{k_p}) := N^{k_p/2} h_{p,N}([Nx_1], \dots, [Nx_{k_p}]) \quad (3.28)$$

satisfy that there exists $f_p \in L^2(\mathbb{R}^{k_p})$ for each $p = 1, \dots, J_L$,

$$\lim_{N \rightarrow \infty} \|\tilde{h}_{p,N} - f_p\|_{L^2(\mathbb{R}^{k_p})} \rightarrow 0. \quad (3.29)$$

Now define the limit vector $(\mathbf{W}(t), \mathbf{B}^{(m)}(t), \mathbf{I})$ as follows: $\mathbf{W}(t)$ and $\mathbf{B}^{(m)}(t)$ are as in (3.26), independent, and

$$\mathbf{I} = (I_{k_p}(f_p))_{p=1, \dots, J_L},$$

where each Wiener-Itô integral $I_{k_p}(\cdot)$ has as Brownian motion integrator $W(\cdot)$ the same as the Brownian motion $W(t)$ defining $\mathbf{W}(t)$. Then as $N \rightarrow \infty$,

$$\left(\mathbf{Y}_{N,S_1}^{(m)}(t), \mathbf{Y}_{N,S_2}^{(m)}(t), \mathbf{Q}_N \right) \xrightarrow{f.d.d.} \left(\mathbf{W}(t), \mathbf{B}^{(m)}(t), \mathbf{I} \right). \quad (3.30)$$

Remark 3.4.4. Observe that $\mathbf{B}^{(m)}$ is independent of (\mathbf{W}, \mathbf{I}) .

Proof. The lemma is proved by combining Lemma 3.4.1 with the proof of Proposition 14.3.2 of Giraitis et al. [2012]. By Cramér-Wold, we need to show that for any $\mathbf{a} \in \mathbb{R}^{J_{S_1}}$, $\mathbf{b} \in \mathbb{R}^{J_{S_2}}$ and $\mathbf{c} \in \mathbb{R}^{J_L}$, as $N \rightarrow \infty$,

$$\langle \mathbf{a}, \mathbf{Y}_{N,S_1}^{(m)}(t) \rangle + \langle \mathbf{b}, \mathbf{Y}_{N,S_2}^{(m)}(t) \rangle + \langle \mathbf{c}, \mathbf{Q}_N \rangle \xrightarrow{f.d.d.} \langle \mathbf{a}, \mathbf{W}(t) \rangle + \langle \mathbf{b}, \mathbf{B}^{(m)}(t) \rangle + \langle \mathbf{c}, \mathbf{I} \rangle, \quad (3.31)$$

where $\langle \cdot, \cdot \rangle$ denotes the Euclidean inner product.

Next following the approximation argument that leads to (14.3.14), (14.3.15) and (14.3.16) in Giraitis et al. [2012], one can show that for any $\epsilon > 0$, there exists $M > 0$ and

simple functions $f_{p,\epsilon} \in S_M(\mathbb{R}^{k_p})$, $p = 1, \dots, J_L$, such that for all $N \geq N_0(\epsilon)$ where $N_0(\epsilon)$ is large enough,

$$\|Q_{k_p}(h_{p,N}) - Q_{k_p}(h_{p,\epsilon,N})\|_{L^2(\Omega)} \leq \epsilon, \quad (3.32)$$

$$Q_{k_p}(h_{p,\epsilon,N}) \xrightarrow{d} I_{k_p}(f_{p,\epsilon}) \quad \text{as } N \rightarrow \infty, \quad (3.33)$$

$$\|I_{k_p}(f_{p,\epsilon}) - I_{k_p}(f_p)\|_{L^2(\Omega)} \leq \epsilon, \quad (3.34)$$

where $\|\cdot\|_{L^2(\Omega)}$ denotes the $L^2(\Omega)$ norm,

$$h_{p,\epsilon,N}(j_1, \dots, j_{k_p}) := N^{-k_p/2} f_{p,\epsilon}\left(\frac{j_1}{N}, \dots, \frac{j_{k_p}}{N}\right).$$

Set

$$\mathbf{Q}_{\epsilon,N} := \left(Q_{k_p}(h_{p,\epsilon,N}) \right)_{p=1,\dots,J_L}$$

and

$$\mathbf{I}_\epsilon := \left(I_{k_p}(f_{p,\epsilon}) \right)_{p=1,\dots,J_L}.$$

Now note that $Q_{k_p}(h_{p,\epsilon,N})$ is a multivariate polynomial (thus is a continuous function) of random variables of the form $W_N(\Delta_i)$ where Δ_i 's are disjoint finite intervals and $W_N(\cdot)$ is the empirical random measure as given in Remark 3.4.2. On the other hand, $I_{k_p}(f_{p,\epsilon})$ is a multivariate polynomial of random variables of the form $W(\Delta_i)$. So by Lemma 3.4.1 (with Remark 3.4.2) and the Continuous Mapping Theorem, we have that as $N \rightarrow \infty$,

$$\langle \mathbf{a}, \mathbf{S}_{N,1}^{(m)}(t) \rangle + \langle \mathbf{b}, \mathbf{S}_{N,2}^{(m)}(t) \rangle + \langle \mathbf{c}, \mathbf{Q}_{\epsilon,N} \rangle \xrightarrow{f.d.d.} \langle \mathbf{a}, \mathbf{W}(t) \rangle + \langle \mathbf{b}, \mathbf{B}^{(m)}(t) \rangle + \langle \mathbf{c}, \mathbf{I}_\epsilon \rangle. \quad (3.35)$$

By (3.32) and the Cauchy-Schwartz inequality, we infer that

$$\| \langle \mathbf{c}, \mathbf{Q}_N - \mathbf{Q}_{\epsilon,N} \rangle \|_{L^2(\Omega)} \leq \| \mathbf{c} \| \| \mathbf{Q}_N - \mathbf{Q}_{\epsilon,N} \|_{L^2(\Omega)} \leq \| \mathbf{c} \| \sqrt{J_L} \epsilon, \quad (3.36)$$

where $\|\cdot\|$ denotes the Euclidean norm. Similarly using (3.34),

$$\|(\langle \mathbf{c}, \mathbf{I} - \mathbf{I}_\epsilon \rangle)\|_{L^2(\Omega)} \leq \|\mathbf{c}\| \|\mathbf{I} - \mathbf{I}_\epsilon\|_{L^2(\Omega)} \leq \|\mathbf{c}\| \sqrt{J_L} \epsilon. \quad (3.37)$$

We now apply a usual triangular approximation argument (e.g., Lemma 4.2.1 of Giraitis et al. [2012]). Let

$$\begin{aligned} U_N^{(m)}(t) &= \langle \mathbf{a}, \mathbf{Y}_{N,S_1}^{(m)}(t) \rangle + \langle \mathbf{b}, \mathbf{Y}_{N,S_2}^{(m)}(t) \rangle + \langle \mathbf{c}, \mathbf{Q}_N \rangle, \\ U_{N,\epsilon}^{(m)}(t) &= \langle \mathbf{a}, \mathbf{Y}_{N,S_1}^{(m)}(t) \rangle + \langle \mathbf{b}, \mathbf{Y}_{N,S_2}^{(m)}(t) \rangle + \langle \mathbf{c}, \mathbf{Q}_{\epsilon,N} \rangle, \\ U_\epsilon^{(m)}(t) &= \langle \mathbf{a}, \mathbf{W}(t) \rangle + \langle \mathbf{b}, \mathbf{B}^{(m)}(t) \rangle + \langle \mathbf{c}, \mathbf{I}_\epsilon \rangle, \\ U^{(m)}(t) &= \langle \mathbf{a}, \mathbf{W}(t) \rangle + \langle \mathbf{b}, \mathbf{B}^{(m)}(t) \rangle + \langle \mathbf{c}, \mathbf{I} \rangle. \end{aligned}$$

By (3.35), (3.37) and (3.36), we have that

$$\begin{aligned} U_{N,\epsilon}^{(m)}(t) &\xrightarrow{f.d.d.} U_\epsilon^{(m)}(t) \text{ as } N \rightarrow \infty, \\ U_\epsilon^{(m)}(t) &\xrightarrow{f.d.d.} U^{(m)}(t) \text{ as } \epsilon \rightarrow 0, \\ \limlimsup_{\substack{\epsilon \rightarrow 0 \\ N \rightarrow \infty}} \|U_N^{(m)}(t) - U_{N,\epsilon}^{(m)}(t)\|_{L^2(\Omega)} &= 0, \quad \forall t \geq 0, \end{aligned}$$

which implies

$$U_N^{(m)}(t) \xrightarrow{f.d.d.} U^{(m)}(t),$$

proving (3.31). □

The next lemma gets rid of the m -truncation.

Lemma 3.4.5. *Lemma 3.4.3 holds with the m -truncated normalized partial sums $\mathbf{Y}_{N,S_i}^{(m)}(t)$, $i = 1, 2$, replaced with the non-truncated ones:*

$$\mathbf{Y}_{N,S_i}(t) = \left(\frac{1}{\sigma_{j,S_i} \sqrt{N}} \sum_{n=1}^{\lfloor Nt \rfloor} X_{j,S_i}(n) \right)_{j=1,\dots,J_i}, \quad i = 1, 2$$

where $X_{j,S_i}(n)$ is the non-truncated multilinear polynomial-form process corresponding to the component of \mathbf{Y}_{N,S_i} in Theorem 3.3.5, $\sigma_{j,S_i} := \sum_n \text{Cov}(X_{j,S_i}(n), X_{j,S_i}(0))$ and the limit $\mathbf{B}^{(m)}(t)$ is replaced by $\mathbf{B}(t)$, that is, as $N \rightarrow \infty$,

$$\left(\mathbf{Y}_{N,S_1}(t), \mathbf{Y}_{N,S_2}(t), \mathbf{Q}_N \right) \xrightarrow{f.d.d.} \left(\mathbf{W}(t), \mathbf{B}(t), \mathbf{I} \right), \quad (3.38)$$

where $\mathbf{W}(t) = (W(t), \dots, W(t))$, $\mathbf{B}(t) = (B_1(t), \dots, B_{J_{S_2}}(t))$ are as given in Theorem 3.3.5.

Proof. We apply again the triangular argument at the end of the proof of Lemma 3.4.3 above, but now with $m \rightarrow \infty$, namely, to show $U_N(t) \xrightarrow{f.d.d.} U(t)$, we show

$$\begin{aligned} U_N^{(m)}(t) &\xrightarrow{f.d.d.} U^{(m)}(t) \text{ as } N \rightarrow \infty, \\ U^{(m)}(t) &\xrightarrow{f.d.d.} U(t) \text{ as } m \rightarrow \infty, \\ \lim_{m \rightarrow \infty} \limsup_{N \rightarrow \infty} \|U_N^{(m)}(t) - U_N(t)\|_{L^2(\Omega)} &= 0, \quad \forall t \geq 0, \end{aligned}$$

The first step follows from Lemma 3.4.3. The second follows from (3.23) since that relation implies that the Gaussian vector $(\mathbf{W}, \mathbf{B}^{(m)}(t))$ converges to $(\mathbf{W}, \mathbf{B}(t))$. For the last step, apply the argument leading to (4.8.7) of Giraitis et al. [2012] and hence for any $t \geq 0$ as $N \rightarrow \infty$,

$$\|Y_{N,j,i}^{(m)}(t) - Y_{N,j,i}(t)\|_{L^2(\Omega)} \rightarrow 0, \quad j = 1, \dots, J_{S_i}, \quad i = 1, 2. \quad (3.39)$$

□

Now we prove Theorem 3.3.5:

Proof of Theorem 3.3.5. In view of Lemma 3.4.5, it is only necessary to verify that the assumption on \mathbf{Q}_N are satisfied, that is, we now focus on the LRD component:

$$\mathbf{Y}_{N,L}(t) = \left(\frac{1}{A_{p,L}(N)} \sum_{n=1}^{[Nt]} X_{p,L}(n) \right)_{p=1, \dots, J_L}$$

in Theorem 3.3.5. Choose as kernels $h_{p,N}$ in (3.28) those obtained from $\mathbf{Y}_{N,L}$, that is,

$$h_{p,N}^{(t)}(s_1, \dots, s_{k_{p,L}}) = c(p, N) N^{-1+k_{p,L}(1/2-d_{p,L})} \sum_{n=1}^{[Nt]} \prod_{i=1}^{k_{p,L}} a_{n-s_i, p},$$

where $c(p, N) > 0$ is some normalization constant. By Theorem 4.8.2 of Giraitis et al. [2012], (3.29) holds and so therefore does Lemma 3.4.5. This concludes the proof of Theorem 3.3.5. \square

3.4.4 Weak convergence in $D[0, 1]^J$

We first state a lemma which will be used to prove case 2d.

Lemma 3.4.6. *Let $Q_k(h)$ be a polynomial form defined in (3.27). If*

$$\sum_{i_1, \dots, i_k} |h(i_1, \dots, i_k)| < \infty, \quad (3.40)$$

and $E(|\epsilon_i|^5) < \infty$, then we have the following hypercontractivity inequality:

$$E(Q_k(h)^4) \leq c E(Q_k(h)^2)^2, \quad (3.41)$$

where $c = (3 + 2E(\epsilon_i^4))^{2k}$.

Proof. Let h_M be the truncated version of h , that is,

$$h_M(i_1, \dots, i_k) = h(i_1, \dots, i_k) \mathbf{1}_{\{i_1 \leq M, \dots, i_k \leq M\}}(i_1, \dots, i_k).$$

By the absolute summability of h , we have

$$E(|Q_k(h_M) - Q_k(h)|) \leq (E|\epsilon_i|)^k \sum_{i_1 > M, \dots, i_k > M} |h(i_1, \dots, i_k)| \rightarrow 0$$

as $M \rightarrow \infty$, and thus

$$Q_k(h_M) \xrightarrow{d} Q_k(h). \quad (3.42)$$

By (11.4.1) of Nourdin and Peccati [2012], we have for $M \geq k$,

$$E(Q_k(h_M)^4) \leq (3 + 2E(\epsilon_i^4))^{2k} E(Q_k(h_M)^2)^2. \quad (3.43)$$

In addition,

$$E(|Q_k(h_M)|^5) \leq A \left(\sum_{i_1, \dots, i_k} |h(i_1, \dots, i_k)| \right)^5 < \infty, \quad (3.44)$$

where $A > 0$ is a constant accounting for the product of absolute moments of $\{\epsilon_i\}$. Note that since h vanishes on the diagonals $i_p = i_q$ when $p \neq q$, there is no moment-order higher than 5 involved there.

Finally, (3.44) implies that $\{Q_k(h_M)^4, M \geq 1\}$ and $\{Q_k(h_M)^2, M \geq 1\}$ are uniformly integrable, and this combined with (3.42) and (3.43) yields (3.41). \square

Proof of Theorem 3.3.8. Convergence in finite-dimensional distributions follows from Theorem 3.3.1, Theorem 3.3.4 and Theorem 3.3.5, so we are left to show tightness in $D[0, 1]^J$. Since univariate tightness implies the multivariate tightness in the product space (Lemma 3.10 of Bai and Taqqu [2013a]), we only need to show that each $\{Y_{j,N}(t), N \geq 1\}$ in (3.2) is tight with respect to the uniform metric. If $X_j(n)$ is LRD, the tightness is shown in Theorem 4.8.2 of Giraitis et al. [2012]. We only need to treat the SRD case.

Suppose that $X(n)$ is a process defined in (3.1) which is SRD.

In case 2a of Theorem 3.3.8, note that X_n is now a stationary m -dependent sequence, so the weak convergence of $S_N(t)$ to Brownian motion, which includes tightness, is classical (Billingsley [1956] Theorem 5.2).

Consider next case 2b. Because ϵ_i are i.i.d. Gaussian, $X(n)$ belongs to the k -th Wiener chaos, or say, can be written as a multiple Wiener-Itô integral of order k (see, e.g., Nourdin and Peccati [2012] Chapter 2.2 and Chapter 2.7). Since the k -th Wiener chaos is a linear space,

$$Y_N(t) := \frac{1}{\sqrt{N}} \sum_{n=1}^{[Nt]} X(n)$$

also belongs to the k -th Wiener chaos, and so does $Y_N(t) - Y_N(s)$ for any $0 \leq s < t$. By the hypercontractivity inequality (Theorem 2.7.2 in Nourdin and Peccati [2012]), we have

$$E[|Y_N(t) - Y_N(s)|^4] \leq cE[|Y_N(t) - Y_N(s)|^2]^2, \quad (3.45)$$

where c is some constant which doesn't depend on s, t or N . Note that $\sum_n |\gamma(n)| < \infty$ due to SRD assumption, we have

$$\begin{aligned} E[|Y_N(t) - Y_N(s)|^2] &= \frac{1}{N} E\left[\left| \sum_{n=1}^{[Nt]-[Ns]} X(n) \right|^2 \right] \\ &= \frac{[Nt] - [Ns]}{N} \sum_{|n| < [Nt] - [Ns]} \left(1 - \frac{|n|}{[Nt] - [Ns]} \right) \gamma(n) \leq \frac{[Nt] - [Ns]}{N} \sum_{n=-\infty}^{\infty} |\gamma(n)|. \end{aligned} \quad (3.46)$$

Combining (3.45) and (3.46), we have for some constant $C > 0$ that

$$E[|Y_N(t) - Y_N(s)|^4] \leq cE[|Y_N(t) - Y_N(s)|^2]^2 \leq C|F_N(t) - F_N(s)|^2,$$

where $F_N(t) = [Nt]/N$. Now by applying Lemma 4.4.1 and Theorem 4.4.1 of Giraitis et al. [2012], we conclude that tightness holds.

Case 2c is shown by Proposition 4.4.4 of Giraitis et al. [2012] with $H = 1/2$.

For case 2d, for $s < t$,

$$\frac{1}{A(N)} \sum_{n=1}^{[Nt]-[Ns]} X(n) = \sum_{1 \leq i_1 < \dots < i_k < \infty} \left(\frac{1}{A(N)} \sum_{n=1}^{[Nt]-[Ns]} a_{n-i_1} \dots a_{n-i_k} \right) \epsilon_{i_1} \dots \epsilon_{i_k}.$$

Thus Lemma 3.4.6 applies with

$$h(i_1, \dots, i_k) = \frac{1}{A(N)} \sum_{n=1}^{[Nt]-[Ns]} a_{n-i_1} \dots a_{n-i_k}$$

since (3.40) holds due to the assumption $\sum_{i \geq 1} |a_i| < \infty$. Tightness then follows by applying the same argument as in case 2b. \square

Chapter 4

Generalized Hermite processes, discrete chaos and limit theorems

We introduce a broad class of self-similar processes $\{Z(t), t \geq 0\}$ called generalized Hermite processes. They have stationary increments, are defined on a Wiener chaos with Hurst index $H \in (1/2, 1)$, and include Hermite processes as a special case. They are defined through a homogeneous kernel g , called “generalized Hermite kernel”, which replaces the product of power functions in the definition of Hermite processes. The generalized Hermite kernels g can also be used to generate long-range dependent stationary sequences forming a discrete chaos process $\{X(n)\}$. In addition, we consider a fractionally-filtered version $Z^\beta(t)$ of $Z(t)$, which allows $H \in (0, 1/2)$. Corresponding non-central limit theorems are established. We also give a multivariate limit theorem which mixes central and non-central limit theorems.

4.1 Introduction

A stochastic process $\{X(t), t \geq 0\}$ with finite variance taking values in \mathbb{R} is said to be *self-similar* if there is a constant called *Hurst coefficient* $H > 0$, such that for any scaling factor $a > 0$, $X(at) \stackrel{f.d.d.}{=} a^H X(t)$, where $\stackrel{f.d.d.}{=}$ means equality in finite-dimensional distributions. If a self-similar process $\{X(t), t \geq 0\}$ has also stationary increments, namely, if for any $h \geq 0$, $\{Y(t) := X(t+h) - X(t), t \geq 0\}$ is a stationary process, then we say that $\{X(t), t \geq 0\}$ is *H-sssi*. The natural range of H is $(0, 1)$, which implies $EX(t) = 0$ for all $t \geq 0$. We refer the reader to Chapter 3 of Embrechts and Maejima [2002] for details.

The fundamental theorem of Lamperti (Lamperti [1962]) states that H -sssi processes are the only possible limit laws of normalized partial sum of stationary sequences, that is, if

$$\frac{1}{A(N)} \sum_{n=1}^{[Nt]} X(n) \xrightarrow{f.d.d.} Y(t)$$

and $A(N) \rightarrow \infty$ as $N \rightarrow \infty$, where $\{X(n)\}$ is stationary, then $\{Y(t), t \geq 0\}$ has to be H -sssi for some $H > 0$, and $A(N)$ has to be regularly varying with exponent H . The notation $\xrightarrow{f.d.d.}$ stands for convergence in finite-dimensional distributions (f.d.d.).

The best known example of Lamperti's fundamental theorem is when $\{X(n)\}$ is i.i.d. or a *short-range dependent* (SRD) sequence, then the limit $Y(t)$ is Brownian motion which is $\frac{1}{2}$ -sssi. If $\{X(n)\}$ has *long-range dependence* (LRD), the limit $Y(t)$ is often H -sssi with $H > 1/2$. The most typical H -sssi process is fractional Brownian motion $B_H(t)$, but there are also non-Gaussian processes, e.g., *Hermite processes* (Taqqu [1979], Dobrushin and Major [1979]). The Hermite process of order 1 is fractional Brownian motion, but when the order is greater than or equal to 2, its law belongs to higher-order Wiener chaos (see, e.g., Peccati and Taqqu [2011]) and is thus non-Gaussian.

The Hermite processes have attracted a lot of attention. The first-order Hermite process, namely fractional Brownian motion, has been studied intensively by numerous researchers since its popularization by Mandelbrot and Van Ness [1968], and we refer the reader to a recent monograph Nourdin [2012] and the references therein. The second-order Hermite process, namely the Rosenblatt process, is also investigated in a number of papers. Recent works include Tudor [2008], Bardet and Tudor [2010], Veillette and Taqqu [2013], Maejima and Tudor [2007, 2013]. Hermite processes frequently appear in statistical inference problems involving LRD, e.g., Lévy-Leduc et al. [2011], Dehling et al. [2013].

It is interesting to note that when the stationary sequence $\{X(n)\}$ is LRD, one can obtain in the limit a much richer class of processes, whereas in the SRD case, one obtains only Brownian motion. The type of limit theorems involving H -sssi processes other than Brownian motion are often called *non-central limit theorems*. While Hermite processes are

the main examples of H -sssi processes obtained as the limit of partial sum of finite-variance LRD sequence, there are very few other limit H -sssi processes which have been considered, with some exceptions Rosenblatt [1979] and Major [1981].

In this chapter, we introduce a broad class of H -sssi ($H > 1/2$) processes $\{Z(t), t \geq 0\}$ with their laws in Wiener chaos, which includes the Hermite processes as a special case. These processes are defined as $Z(t) = I_k(h_t)$, where $I_k(\cdot)$ denotes k -tuple Wiener-Itô integral, and

$$h_t(x_1, \dots, x_k) := \int_0^t g(s - x_1, \dots, s - x_k) 1_{\{s > x_1, \dots, s > x_k\}} ds,$$

with g being some suitable homogeneous function on \mathbb{R}_+^k called *generalized Hermite kernel*.

For example,

$$g(x_1, \dots, x_k) = \max \left(\frac{x_1 \dots x_k}{x_1^{k-\alpha} + \dots + x_k^{k-\alpha}}, x_1^{\alpha/k} \dots x_k^{\alpha/k} \right), \quad \mathbf{x} \in \mathbb{R}_+^k, \quad \alpha \in \left(-\frac{k}{2} - \frac{1}{2}, -\frac{k}{2} \right) \quad (4.1)$$

We call the corresponding H -sssi process $Z(t)$ a *generalized Hermite process*. We then construct a class of *discrete chaos processes* as

$$X(n) = \sum'_{(i_1, \dots, i_k) \in \mathbb{Z}_+^k} g(i_1, \dots, i_k) \epsilon_{n-i_1} \dots \epsilon_{n-i_k},$$

where $\{\epsilon_i\}$ are i.i.d. noise, and the prime ' exclusion of the diagonals $i_p = i_q, p \neq q$. We show that the normalized partial sum of $X(n)$ converges to the generalized Hermite process $Z(t)$ defined by the same g . We also obtain processes with $H \in (0, 1/2)$ by applying an additional fractional filter. The increments of these processes have negative dependence. Finally, we state a multivariate limit theorem which mixes central and non-central limits, including cases where there is an additional fractional filter.

The chapter is organized as follows. In Section 2, we review the Hermite processes. In Section 3, the generalized Hermite processes are introduced. In Section 4, we consider the

discrete chaos processes. In Section 5, we prove a hypercontractivity relation for infinite discrete chaos. In Section 6, we show that the discrete chaos processes converge weakly to the generalized Hermite processes, including situations where $H < 1/2$.

4.2 Brief review of Hermite processes

The Hermite processes are defined with the aid of a multiple stochastic integral called *Wiener-Itô integral*. We give here a brief introduction to this integral. For the proofs of our statements and additional details, we refer the reader to Major [2014] and Nualart [2006], for example. The Wiener-Itô integral is defined for any $f \in L^2(\mathbb{R}^k)$ as

$$I_k(f) := \int'_{\mathbb{R}^k} f(x_1, \dots, x_k) W(dx_1) \dots W(dx_k),$$

where $W(\cdot)$ is Brownian motion viewed as a random integrator, and the prime $'$ indicates that we don't integrate on the diagonals $x_p = x_q$, $p \neq q$. The integral $I_k(\cdot)$ can be defined first for elementary functions $f = \sum_{i=1}^n a_i 1_{A_i}$, where A_i 's are off-diagonal cubes in \mathbb{R}^k . This results in a linear combination of k -fold product of independent centered Gaussian random variables. One then extends this in the usual way to any $f \in L^2(\mathbb{R}^k)$. The random variable $I_k(f)$ is also said to belong to the k -th Wiener chaos \mathcal{H}_k , which is the Hilbert space generated by $I_k(f)$ when f varies in $L^2(\mathbb{R}^k)$. Here we state the following important properties of the Wiener-Itô integral $I_k(\cdot)$:

1. $I_k(\cdot)$ is a linear mapping from $L^2(\mathbb{R}^k)$ to $L^2(\Omega)$.
2. If $f_\sigma(x_1, \dots, x_k) := f(x_{\sigma(1)}, \dots, x_{\sigma(k)})$, where σ is any permutation of $(1, \dots, k)$, then $I_k(f_\sigma) = I_k(f)$. It hence suffices to focus on symmetric integrands (symmetrize f as

$$\tilde{f}(x_1, \dots, x_k) := \frac{1}{k!} \sum_{\sigma} f(x_{\sigma(1)}, \dots, x_{\sigma(k)})$$

when necessary).

3. Suppose $f \in L^2(\mathbb{R}^p)$ and $g \in L^2(\mathbb{R}^q)$, and both are symmetric. Then

$$EI_p(f)I_q(g) = \begin{cases} k! \langle f, g \rangle_{L^2(\mathbb{R}^k)} = k! \int_{\mathbb{R}^k} f(\mathbf{x})g(\mathbf{x})d\mathbf{x}, & \text{if } p = q = k; \\ 0, & \text{if } p \neq q. \end{cases}$$

If $f \in L^2(\mathbb{R}^k)$ is not symmetric, one gets

$$EI_p(f)^2 = \|\tilde{f}\|_{L^2(\mathbb{R}^k)}^2 \leq k! \|f\|_{L^2(\mathbb{R}^k)}^2.$$

An Hermite process of order k is an H -sssi process with $1/2 < H < 1$, which is represented by the following Wiener-Itô integral:

$$Z_H^{(k)}(t) = a_{k,d} \int_{\mathbb{R}^k} \int_0^t \prod_{j=1}^k (s - x_j)_+^{d-1} ds W(dx_1) \dots W(dx_k), \quad (4.2)$$

where

and $a_{k,d}$ is some positive constant that makes $\text{Var}(Z_H^{(k)}(1)) = 1$. We call (4.2) the *time-domain representation*. It is known that Hermite processes admit other representations in terms of Wiener-Itô integrals (see Pipiras and Taqqu [2010]), among which we note the *spectral-domain representation*:

$$Z_H^{(k)}(t) = b_{k,d} \int_{\mathbb{R}^k} \frac{e^{i(u_1 + \dots + u_k)t} - 1}{i(u_1 + \dots + u_k)} |u_1|^{-d} \dots |u_k|^{-d} \widehat{W}(du_1) \dots \widehat{W}(du_k), \quad (4.3)$$

where $\widehat{W}(\cdot)$ is a complex-valued Brownian motion (with real and imaginary parts being independent) viewed as a random integrator (see, e.g., p.22 of Embrechts and Maejima [2002]), the double prime $''$ indicates the exclusion of the hyper-diagonals $u_p = \pm u_q$, $p \neq q$, and $b_{k,d}$ is some positive constant that makes $\text{Var}(Z_H^{(k)}(1)) = 1$. In the sequel, we use $\widehat{I}_k(\cdot)$ to denote a k -tuple Wiener-Itô integral with respect to the complex-valued Brownian motion $\widehat{W}(\cdot)$. In fact, the kernel inside the Wiener-Itô integral in (4.3) is the Fourier transform of the kernel in (4.2) up to some unimportant factors. The connection between

the time-domain and spectral-domain representation is through the following general result:

Proposition 4.2.1. (*Proposition 9.3.1 of Peccati and Taqqu [2011]*) *Let $g_j(\mathbf{x})$ be a real-valued function in $L^2(\mathbb{R}^{k_j})$, $j = 1, \dots, J$. Let*

$$\widehat{g}_j(\mathbf{u}) = \int_{\mathbb{R}^{k_j}} g_j(\mathbf{x}) e^{i\langle \mathbf{u}, \mathbf{x} \rangle} d\mathbf{x}$$

be the Fourier transform. Then

$$\left(I_{k_1}(g_1), \dots, I_{k_J}(g_J) \right) \stackrel{d}{=} \left((2\pi)^{-k_1/2} \widehat{I}_{k_1}(\widehat{g}_1 w^{\otimes k_1}), \dots, (2\pi)^{-k_J/2} \widehat{I}_{k_J}(\widehat{g}_J w^{\otimes k_J}) \right),$$

for any $|w(u)| = 1$ and $w(u) = \overline{w(-u)}$, where $w^{\otimes k}(u_1 \dots u_k) := w(u_1) \dots w(u_k)$.

The factors $w^{\otimes k_j}$ do not change the distributions due to the change-of-variable formula of Wiener-Itô integrals (see, e.g., Proposition 4.2 of Dobrushin [1979]).

The Hermite process of order $k = 1$ is fractional Brownian motion $B_H(t)$, and that of order $k = 2$ is called *Rosenblatt process* whose marginal distribution was discovered by Rosenblatt [1961]. We note that all H -sssi processes with unit variance at $t = 1$ have covariance

$$R(s, t) = \frac{1}{2}(s^{2H} + t^{2H} - |s - t|^{2H}),$$

as is the case for Hermite process of arbitrary order.

Hermite processes arise as limits of partial sum of nonlinear LRD sequences. In the following two theorems, $A(N)$ is a normalization factor guaranteeing unit asymptotic variance for the partial sum process at $t = 1$. We use \Rightarrow to denote weak convergence in the Skorohod space $D[0, 1]$ with the uniform metric.

Theorem 4.2.2. (*Dobrushin and Major [1979], Taqqu [1979].*) *Suppose that $\{X(n)\}$ is a Gaussian stationary sequence with autocovariance*

$$\gamma(n) \sim cn^{2d-1}$$

as $n \rightarrow \infty$ for some constant $c > 0$ and

$$1/2(1 - 1/k) < d < 1/2.$$

Let $H_k(x) := (-1)^k e^{x^2/2} \frac{d^k}{dx^k} e^{-x^2/2}$ be the k -th Hermite polynomial, $k \geq 1$. Then

$$\frac{1}{A(N)} \sum_{n=1}^{[Nt]} H_k(X(n)) \Rightarrow Z_d^{(k)}(t).$$

Theorem 4.2.3. (Surgailis [1982], see also Giraitis et al. [2012] Chapter 4.8.) Let $\{\epsilon_i\}$ be an i.i.d. sequence with mean 0 variance 1,

$$a_n \sim cn^{d-1}$$

as $n \rightarrow \infty$ for some constant $c > 0$ and

$$1/2(1 - 1/k) < d < 1/2.$$

Let

$$X(n) = \sum'_{0 < i_1, \dots, i_k < \infty} a_{i_1} \dots a_{i_k} \epsilon_{n-i_1} \dots \epsilon_{n-i_k},$$

where the prime $'$ indicates that one doesn't sum on the diagonals $i_p = i_q$ $p \neq q$. Then

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

Remark 4.2.4. The Hermite polynomial in Theorem 4.2.2 can be replaced by a general function $G(\cdot)$ such that $EG(X_n) = 0$, $EG(X_n)^2 < \infty$, due to the orthogonal expansion of $G(x)$ with respect to Hermite polynomials, and the fact that only the leading term in the expansion contributes to the limit law. Similarly, the off-diagonal multilinear polynomial-form process $X(n)$ in Theorem 4.2.3 can be replaced by a suitable function of the linear process $Y(n) := \sum_{i \geq 1} a_i \epsilon_{n-i}$. In both of the above theorems $\xrightarrow{f.d.d.}$ can be strengthened to

weak convergence \Rightarrow (Proposition 4.4.2 of Giraitis et al. [2012]).

Remark 4.2.5. The range of the parameter d in both of the theorems guarantees that the summand is LRD in the sense that the autocovariance decays as a power function with an exponent in the range $(-1, 0)$. We note also that the constant $c > 0$ appearing in both theorems can be replaced by a slowly varying function.

4.3 Generalized Hermite Processes

We introduce first some notation, which will be used throughout. $\mathbb{R}_+ = (0, \infty)$, $\mathbb{Z}_+ = \{1, 2, \dots\}$. $\mathbf{x} = (x_1, \dots, x_k) \in \mathbb{R}^k$, $\mathbf{i} = (i_1, \dots, i_k) \in \mathbb{Z}^k$, $\mathbf{0} = (0, \dots, 0)$, $\mathbf{1} = (1, \dots, 1)$. For any real number x , $[x] = \sup\{n \in \mathbb{Z}, n \leq x\}$, and $[\mathbf{x}] = ([x_1], \dots, [x_k])$. We write $\mathbf{x} > \mathbf{y}$ (or \geq) if $x_j > y_j$ (or \geq), $j = 1, \dots, k$. $\langle \mathbf{x}, \mathbf{y} \rangle = \sum_{j=1}^k x_j y_j$, and $\|\mathbf{x}\| = \sqrt{\langle \mathbf{x}, \mathbf{x} \rangle}$, while $\|\cdot\|$ with a subscript is also used to denote the norm of some other space (specified in the subscript). Given a set $A \subset \mathbb{R}$, A^k is the k -fold Cartesian product. $1_A(\cdot)$ is the indicator function of a set A . $L^p(\mathbb{R}^k, \mu)$ denotes the L^p -space on \mathbb{R}^k with measure μ , and μ is omitted if it is Lebesgue measure.

4.3.1 General kernels

The following proposition provides a general way to construct in the time-domain an H -sssi process living in Wiener chaos:

Proposition 4.3.1. *Fix an $H \in (0, 1)$. Suppose that $\{h_t(\cdot), t > 0\}$ is a family of functions defined on \mathbb{R}^k satisfying*

1. $h_t \in L^2(\mathbb{R}^k)$;
2. $\forall \lambda > 0, \exists \beta \neq 0$, such that $h_{\lambda t}(\mathbf{x}) = \lambda^{H+k\beta/2} h_t(\lambda^\beta \mathbf{x})$ for a.e. $\mathbf{x} \in \mathbb{R}^k$ and all $t > 0$;
3. $\forall s > 0, \exists \mathbf{a} \in \mathbb{R}^k$, such that $h_{t+s}(\mathbf{x}) - h_t(\mathbf{x}) = h_s(\mathbf{x} + t\mathbf{a})$ for a.e. $\mathbf{x} \in \mathbb{R}^k$ and all $t > 0$.

Then $Z(t) := I_k(h_t)$ is an H -sssi process.

Condition 1 guarantees that the Wiener-Itô integral is well defined. Condition 2 yields self-similarity, where the term $k\beta/2$ in the exponent compensates for the scaling of the k -tuple Brownian motion integrators. Condition 3 guarantees stationary increments. Self-similarity and stationary increments can be rigorously checked by the change-of-variable formula of Wiener-Itô integrals (Proposition 4.2 of Dobrushin [1979]).

The Hermite process, for instance, which is defined in (4.2) can be obtained following the scheme of Proposition 4.3.1 by letting

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

and

$$g(\mathbf{x}) = \prod_{j=1}^k x_j^{d-1}, \quad x_j > 0. \quad (4.4)$$

It is easy to check that the conditions on h_t in Proposition 4.3.1 are all satisfied with $\beta = -1$ in condition 2 and $H = kd - k/2 + 1$. One can also check that the integrand in the spectral-domain representation in (4.3) also satisfies the first two conditions in Proposition 4.3.1, but with $\beta = 1$ in Condition 2 instead. The third condition, however, must be replaced by $\widehat{h}_{t+s}(\mathbf{u}) - \widehat{h}_t(\mathbf{u}) = e^{-it\langle \mathbf{a}, \mathbf{u} \rangle} \widehat{h}_s(\mathbf{u})$ due to the Fourier-transform relation.

Our first goal is to extend the kernel g in (4.4) to some general class of functions. To do so, we define the following class of functions on \mathbb{R}_+^k , which first appeared in Mori and Oodaira [1986] to study the law of iterated logarithm:

Definition 4.3.2. We say that a nonzero measurable function $g(\mathbf{x})$ defined on \mathbb{R}_+^k is a *generalized Hermite kernel*, if it satisfies

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

One can check that the Hermite kernel g in (4.4) satisfies the above assumptions.

Remark 4.3.3. The range of α in Condition A is non-overlapping for different k , and extends from $-1/2$ to $-\infty$ with all the multiples of $-1/2$ excluded.

Remark 4.3.4. Suppose g_1 and g_2 are generalized Hermite kernels having order k_1, k_2 and homogeneity exponent α_1, α_2 respectively. If in addition, $\alpha_1 + \alpha_2 > -(k_1 + k_2 + 1)/2$, then $g_1 \otimes g_2(\mathbf{x}_1, \mathbf{x}_1) := g_1(\mathbf{x}_1)g_2(\mathbf{x}_2)$ is a generalized Hermite kernel having order $k_1 + k_2$ and homogeneity exponent $\alpha_1 + \alpha_2$.

Theorem 4.3.5. Let $g(\mathbf{x})$ be a generalized Hermite kernel defined in Definition 4.3.2. Then

$$h_t(\mathbf{x}) = \int_0^t g(s\mathbf{1} - \mathbf{x})1_{\{s\mathbf{1} > \mathbf{x}\}} ds$$

is well-defined in $L^2(\mathbb{R}^k)$, $\forall t > 0$, and the process defined by $Z_t := I_k(h_t)$ is an H -sssi process with

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

Proof. To check that $h_t \in L^2(\mathbb{R}^k)$, we write

$$\int_{\mathbb{R}^k} h_t(\mathbf{x})^2 d\mathbf{x} = \int_{\mathbb{R}^k} d\mathbf{x} \int_0^t \int_0^t ds_1 ds_2 g(s_1\mathbf{1} - \mathbf{x})g(s_2\mathbf{1} - \mathbf{x})1_{\{s_1\mathbf{1} > \mathbf{x}\}}1_{\{s_2\mathbf{1} > \mathbf{x}\}}.$$

We want to change the integration order by integrating on \mathbf{x} first. By Fubini, we need to check that the absolute value of the integrand is integrable, that is,

$$\begin{aligned} & 2 \int_0^t ds_1 \int_{s_1}^t ds_2 \int_{\mathbb{R}^k} d\mathbf{x} |g(s_1\mathbf{1} - \mathbf{x})g(s_2\mathbf{1} - \mathbf{x})|1_{\{s_1\mathbf{1} - \mathbf{x} > \mathbf{0}\}} \quad (\text{by symmetry}) \\ &= 2 \int_0^t ds \int_0^{t-s} du \int_{\mathbb{R}_+^k} d\mathbf{w} |g(\mathbf{w})g(u\mathbf{1} + \mathbf{w})| \quad (s = s_1, u = s_2 - s_1, \mathbf{w} = s_1\mathbf{1} - \mathbf{x}) \\ &= 2 \int_0^t ds \int_0^{t-s} du \int_{\mathbb{R}_+^k} u^k d\mathbf{y} |g(u\mathbf{y})g(u + u\mathbf{y})| \\ &= 2 \int_0^t ds \int_0^{t-s} u^{2\alpha+k} du \int_{\mathbb{R}_+^k} d\mathbf{y} |g(\mathbf{y})g(\mathbf{1} + \mathbf{y})| \quad (\text{by Condition A of Definition 4.3.2}), \end{aligned}$$

where the last expression is finite by $2\alpha + k + 1 > 0$ and Condition B. Hence by the same

calculation, but without absolute values,

$$\begin{aligned} \int_{\mathbb{R}^k} h_t(\mathbf{x})^2 d\mathbf{x} &= 2 \int_0^t ds \int_0^{t-s} u^{2\alpha+k} du \int_{\mathbb{R}_+^k} d\mathbf{y} g(\mathbf{y})g(\mathbf{1} + \mathbf{y}) \\ &= \frac{t^{2\alpha+k+2}}{(\alpha + k/2 + 1)(2\alpha + k + 2)} \int_{\mathbb{R}_+^k} d\mathbf{y} g(\mathbf{y})g(\mathbf{1} + \mathbf{y}). \end{aligned}$$

To check self-similarity (Condition 2 of Proposition 4.3.1 with $\beta = -1$),

$$h_{\lambda t}(\mathbf{x}) = \int_0^{\lambda t} g(s\mathbf{1} - \mathbf{x})1_{\{s\mathbf{1} > \mathbf{x}\}} ds = \lambda^{\alpha+1} \int_0^t g(r\mathbf{1} - \lambda^{-1}\mathbf{x})1_{\{r\mathbf{1} > \lambda^{-1}\mathbf{x}\}} \lambda dr = \lambda^{\alpha+1} h_t(\lambda^{-1}\mathbf{x}),$$

where the second equality uses Condition A of Definition 4.3.2. The Hurst coefficient H of $I_k(h_t)$ is obtained from $\alpha + 1 = H - k/2$. To check stationary increments (Condition 3 of Proposition 4.3.1), for any $t, r > 0$,

$$\begin{aligned} h_{t+r}(\mathbf{x}) - h_t(\mathbf{x}) &= \int_t^{t+r} g(s\mathbf{1} - \mathbf{x})1_{\{s\mathbf{1} > \mathbf{x}\}} ds \\ &= \int_0^r g(u\mathbf{1} + t\mathbf{1} - \mathbf{x})1_{\{u\mathbf{1} + t\mathbf{1} > \mathbf{x}\}} du = h_r(\mathbf{x} - t\mathbf{1}). \end{aligned}$$

□

Remark 4.3.6. As a byproduct of the above proof, we obtain that under the conditions of Definition 4.3.2, one has $\int_0^t |g(s\mathbf{1} - \mathbf{x})|1_{\{s\mathbf{1} > \mathbf{x}\}}(s) ds < \infty$ for a.e. $\mathbf{x} \in \mathbb{R}^k$, and

$$\mathbb{E}Z(t)^2(k!)^{-1} \leq \|h_t\|_{L^2(\mathbb{R}^k)}^2 = \frac{t^{2H}}{H(2H-1)} C_g,$$

where $C_g := \int_{\mathbb{R}_+^k} g(\mathbf{x})g(\mathbf{1} + \mathbf{x})d\mathbf{x}$, and the first inequality becomes equality if g and hence h_t is symmetric. Note that $C_g > 0$ must hold, otherwise $h_t(\mathbf{x}) = \int_0^t g(s\mathbf{1} - \mathbf{x})1_{\{s\mathbf{1} > \mathbf{x}\}} ds = 0$ for a.e. $\mathbf{x} \in \mathbb{R}^k$ and any $t > 0$, which implies that g is zero a.e., and thus contradicts the assumption.

Remark 4.3.7. Since $\forall f \in L^2(\mathbb{R}^k)$, $I_k(f) = I_k(\tilde{f})$, where \tilde{f} is the symmetrization of f (Nualart [2006] p.9), it suffices to focus on symmetric generalized Hermite kernels g only.

In the sequel, we will not always assume that g is symmetric for convenience, while being aware that g can always be symmetrized.

Definition 4.3.8. The process

$$Z(t) := \int_{\mathbb{R}^k}' \int_0^t g(s - x_1, \dots, s - x_k) 1_{\{s > x_1, \dots, s > x_k\}} ds W(dx_1) \dots W(dx_k) \quad (4.5)$$

which we simply write $Z(t) = I_k(h_t)$ with $h_t(\mathbf{x}) = \int_0^t g(s\mathbf{1} - \mathbf{x}) 1_{\{s\mathbf{1} > \mathbf{x}\}} ds$, where g is a generalized Hermite kernel defined in Definition 4.3.2, is called a *generalized Hermite process*.

Remark 4.3.9. It is known (see, e.g., Janson [1997] Theorem 6.12) that if a random variable X belongs to the k -th Wiener chaos, then there $\exists a, b, t_0 > 0$ such that for $t \geq t_0$,

$$\exp(-at^{2/k}) \leq P(|X| > t) \leq \exp(-bt^{2/k}).$$

This shows that the generalized Hermite processes of different orders must necessarily have different laws, and the higher the order gets, the heavier the tail of the marginal distribution becomes, while they all have moments of any order.

The generalized Hermite process $Z(t)$ admits a continuous version, which follows from the following general result:

Proposition 4.3.10. *If $\{Z(t), t \geq 0\}$ is an H -sssi process whose marginal distribution satisfies $\mathbb{E}|Z(1)|^\gamma < \infty$ for some $\gamma > H^{-1}$, then $Z(t)$ admits a continuous version.*

Proof. Using stationary increments and self-similarity, we have

$$\mathbb{E}|Z(t) - Z(s)|^\gamma = \mathbb{E}|Z(t-s)|^\gamma = |t-s|^{H\gamma} \mathbb{E}|Z(1)|^\gamma.$$

Since $H\gamma > 1$, Kolmogorov's criterion applies. □

Remark 4.3.11. In Mori and Oodaira [1986], the following laws of iterated logarithm are

obtained for the generalized Hermite process $Z(t)$:

$$\limsup_{n \rightarrow \infty} \frac{Z(n)}{n^H (2 \log_2 n)^{k/2}} = l_1, \quad \liminf_{n \rightarrow \infty} \frac{Z(n)}{n^H (2 \log_2 n)^{k/2}} = l_2 \quad \text{a.s.},$$

where $l_1 = \sup K_h$ and $l_2 = \inf K_h$ with the set

$$K_h := \left\{ \int_{\mathbb{R}^k} h_1(\mathbf{x}) \xi(x_1) \dots \xi(x_k) d\mathbf{x} : \|\xi\|_{L^2(\mathbb{R})} \leq 1 \right\}.$$

In the spirit of (4.3), we can consider the spectral-domain representation of the generalized Hermite processes. Since $h_t(\mathbf{x}) = \int_0^t g(s\mathbf{1} - \mathbf{x}) 1_{\{s\mathbf{1} > \mathbf{x}\}}(s) ds \in L^2(\mathbb{R})$, it always has an L^2 -sense Fourier transform \widehat{h}_t . We give an explicit way to calculate \widehat{h}_t when g is integrable in a neighborhood of the origin. Note that since g is homogeneous, it suffices to assume integrability on the unit cube $(0, 1]^k$.

Proposition 4.3.12. *Suppose that*

$$\int_{(0,1]^k} |g(\mathbf{x})| < \infty. \quad (4.6)$$

Let $g_n(\mathbf{x}) = g(\mathbf{x}) 1_{(0,n]^k}(\mathbf{x})$, and $\widehat{g}_n(\mathbf{u}) := \int_{\mathbb{R}^k} g_n(\mathbf{x}) e^{i\langle \mathbf{u}, \mathbf{x} \rangle} d\mathbf{x}$ be its Fourier transform. Set

$$\widehat{h}_{t,n} := \frac{e^{it\langle \mathbf{u}, \mathbf{1} \rangle} - 1}{i\langle \mathbf{u}, \mathbf{1} \rangle} \widehat{g}_n(-\mathbf{u}),$$

then $\widehat{h}_{t,n}$ converges in $L^2(\mathbb{R}^k)$ to \widehat{h}_t . Moreover, there is a function $\widehat{g}(\mathbf{u})$ defined for a.e. $\mathbf{u} \in \mathbb{R}^k$, such that,

$$\widehat{h}_t(\mathbf{u}) = \frac{e^{it\langle \mathbf{u}, \mathbf{1} \rangle} - 1}{i\langle \mathbf{u}, \mathbf{1} \rangle} \widehat{g}(-\mathbf{u}). \quad (4.7)$$

Proof. Due to (4.6), the Fourier transform of g_n is well-defined pointwise as

$$\widehat{g}_n(\mathbf{u}) = \int_{\mathbb{R}^k} g(\mathbf{x}) 1_{(0,n]^k}(\mathbf{x}) e^{i\langle \mathbf{u}, \mathbf{x} \rangle} d\mathbf{x}. \quad (4.8)$$

Let

$$h_{t,n}(\mathbf{x}) = \int_0^t g_n(s\mathbf{1} - \mathbf{x})1_{\{s\mathbf{1} > \mathbf{x}\}}(s)ds = \int_0^t g(s\mathbf{1} - \mathbf{x})1_{\{\mathbf{x} < s\mathbf{1} \leq \mathbf{x} + n\mathbf{1}\}}(s)ds.$$

Note that $|g_n(\mathbf{x})| \leq |g(\mathbf{x})|$, so by the proof of Theorem 4.3.5, $h_{t,n}(\mathbf{x}) \in L^2(\mathbb{R}^k)$, and by the Dominated Convergence Theorem, $h_{t,n}$ converges to h_t pointwise as $n \rightarrow \infty$. Since $|h_{t,n}| \leq \int_0^t |g(s\mathbf{1} - \mathbf{x})|1_{\{s\mathbf{1} > \mathbf{x}\}}(s)ds$, by the Dominated Convergence Theorem in $L^2(\mathbb{R}^k)$, $h_{t,n}$ converges to h_t in $L^2(\mathbb{R}^k)$. By Plancherel's isometry, $\widehat{h}_{t,n}$, the Fourier transform of $h_{t,n}$, converges in $L^2(\mathbb{R}^k)$ to \widehat{h}_t . But

$$\begin{aligned} \widehat{h}_{t,n}(\mathbf{u}) &:= \int_{\mathbb{R}^k} \int_0^t g(s\mathbf{1} - \mathbf{x})1_{\{\mathbf{x} < s\mathbf{1} \leq \mathbf{x} + n\mathbf{1}\}}(s)ds e^{i\langle \mathbf{u}, \mathbf{x} \rangle} d\mathbf{x} \\ &= \int_0^t \int_{\mathbb{R}^k} e^{i\langle \mathbf{u}, s\mathbf{1} \rangle} g(s\mathbf{1} - \mathbf{x})e^{i\langle -\mathbf{u}, s\mathbf{1} - \mathbf{x} \rangle} 1_{\{\mathbf{0} < s\mathbf{1} - \mathbf{x} \leq n\mathbf{1}\}}(\mathbf{x}) d\mathbf{x} ds \\ &= \int_0^t e^{i\langle \mathbf{u}, s\mathbf{1} \rangle} ds \int_{\mathbb{R}^k} g(\mathbf{y})1_{\{\mathbf{0} < \mathbf{y} \leq n\mathbf{1}\}} e^{i\langle -\mathbf{u}, \mathbf{y} \rangle} d\mathbf{y} \\ &= \frac{e^{it\langle \mathbf{u}, \mathbf{1} \rangle} - 1}{i\langle \mathbf{u}, \mathbf{1} \rangle} \widehat{g}_n(-\mathbf{u}), \end{aligned} \tag{4.9}$$

where the change of integration order is valid because by (4.6),

$$\int_0^t ds \int_{\mathbb{R}^k} d\mathbf{x} |g(s\mathbf{1} - \mathbf{x})|1_{\{\mathbf{x} < s\mathbf{1} \leq \mathbf{x} + n\mathbf{1}\}} = \int_0^t ds \int_{\mathbb{R}^k} |g(\mathbf{y})|1_{\{\mathbf{0} < \mathbf{y} \leq n\mathbf{1}\}} d\mathbf{y} < \infty.$$

We now prove (4.7). The fact that $\widehat{h}_{t,n}$ converges in $L^2(\mathbb{R}^k)$ to \widehat{h}_t implies that \widehat{g}_n is a Cauchy sequence in $L^2(\mathbb{R}^k, \mu_t)$, where μ_t is the measure given by

$$\mu_t(A) = \int_A \left| \frac{e^{it\langle \mathbf{u}, \mathbf{1} \rangle} - 1}{i\langle \mathbf{u}, \mathbf{1} \rangle} \right|^2 d\mathbf{u} = \int_A \frac{2 - 2\cos(t\langle \mathbf{u}, \mathbf{1} \rangle)}{\langle \mathbf{u}, \mathbf{1} \rangle^2} d\mathbf{u}$$

for any measurable set $A \subset \mathbb{R}^k$. Hence there exists a $\widehat{g} \in L^2(\mathbb{R}^k, \mu_t)$ which is the limit of \widehat{g}_n in $L^2(\mathbb{R}^k, \mu_t)$. Since μ_t is equivalent to Lebesgue measure, \widehat{g} is determined a.e. on \mathbb{R}^k , and there exists a subsequence of \widehat{g}_n that converges a.e. to \widehat{g} . So (4.7) holds. \square

Remark 4.3.13. Note that \widehat{g} is not the L^2 -sense Fourier transform of $g1_{\mathbb{R}_+^k}$, since $g \notin$

$L^2(\mathbb{R}_+^k)$. One can, however, evaluate the limit of \widehat{g}_n pointwise as an improper integral, as is done in the Hermite kernel case (4.4) (see Lemma 6.2 of Taqqu [1979]).

The limit \widehat{g} in (4.7) is also a homogeneous function:

Proposition 4.3.14. *The function \widehat{g} defined in Remark 4.3.12 satisfies for any $\lambda > 0$, $g(\lambda \mathbf{u}) = \lambda^{-\alpha-k} \widehat{g}(\mathbf{u})$ for a.e. $\mathbf{u} \in \mathbb{R}^k$.*

Proof. Following (4.8) and using Condition A of Definition 4.3.2, and noting that $\langle \lambda \mathbf{u}, \mathbf{x} \rangle = \langle \mathbf{u}, \lambda \mathbf{x} \rangle$, we have

$$\begin{aligned} \widehat{g}_n(\lambda \mathbf{u}) &= \lambda^{-\alpha} \int_{\mathbb{R}^k} g(\lambda \mathbf{x}) 1_{(0,n]^k}(\mathbf{x}) e^{i\langle \mathbf{u}, \lambda \mathbf{x} \rangle} d\mathbf{x} \\ &= \lambda^{-\alpha-k} \int_{\mathbb{R}^k} g(\mathbf{y}) 1_{(0,\lambda n]^k}(\mathbf{y}) e^{i\langle \mathbf{u}, \mathbf{y} \rangle} d\mathbf{y} = \lambda^{-\alpha-k} \widehat{g}_{n\lambda}(\mathbf{u}). \end{aligned}$$

Then let $n \rightarrow \infty$ through a subsequence so that both sides converge a.e.. □

Remark 4.3.15. The spectral-domain representation of the Hermite process in (4.3) is indeed obtained as $\widehat{g}(\mathbf{u}) = c \prod_{j=1}^k |u_j|^{-d} w(\mathbf{u})$ for some constant $c > 0$, where the function $w(\mathbf{u}) = \prod_{j=1}^k \exp(-\text{sign}(u_j) \frac{i\pi d}{2})$ can be omitted (see Proposition 4.2.1).

4.3.2 Special kernels and examples

We introduce now some subclasses of the generalized Hermite kernels g defined in Definition 4.3.2, which will be of interest later when dealing with limit theorems. Note that the kernel g is determined by its value on the positive unit sphere $\mathcal{S}_+^k := \{\mathbf{x} \in \mathbb{R}_+^k, \|\mathbf{x}\| = 1\}$. Because it is homogeneous, g is always radially continuous and it is decreasing since $\alpha < 0$ in Definition 4.3.2. Thus assuming that g is continuous on \mathcal{S}_+^k a.e. (with respect to the uniform measure on the \mathcal{S}_+^k) is the same as assuming g is continuous a.e. on \mathbb{R}_+^k .

Definition 4.3.16. We say that a generalized Hermite kernel g is of Class (B) (B stands for “boundedness”), if on \mathcal{S}_+^k , it is continuous a.e. and bounded. Consequently,

$$|g(\mathbf{x})| = \|\mathbf{x}\|^\alpha |g(\mathbf{x}/\|\mathbf{x}\|)| \leq c \|\mathbf{x}\|^\alpha$$

for some $c > 0$.

Remark 4.3.17. According to Lemma 7.1 of Mori and Oodaira [1986], Class (B) forms a dense subclass of the class of generalized Hermite kernels in the sense that for any generalized Hermite kernel g and any $\epsilon > 0$, there exists g_ϵ in Class (B), such that $\|h - h_\epsilon\|_{L^2(\mathbb{R}^k)} < \epsilon$, where $h(\mathbf{x}) = \int_0^1 g(s\mathbf{1} - \mathbf{x})1_{\{s\mathbf{1} > \mathbf{x}\}} ds$ and $h_\epsilon(\mathbf{x}) = \int_0^1 g_\epsilon(s\mathbf{1} - \mathbf{x})1_{\{s\mathbf{1} > \mathbf{x}\}} ds$.

Note that Class (B) does not include the original Hermite kernel in (4.4). We now introduce a class of generalized Hermite kernels, called Class (L), which includes generalized Hermite kernels of the form:

$$g(\mathbf{x}) = \prod_{j=1}^k x_j^{\gamma_j}, \quad (4.10)$$

where each $-1 < \gamma_j < -1/2$ and $-k/2 - 1/2 < \sum_j \gamma_j < -k/2$. These particular kernels with $k = 2$ has been considered in Maejima and Tudor [2012] where the resulting process is called non-symmetric Rosenblatt process. We hence call the kernel in (4.10) a *non-symmetric Hermite kernel*. Note that despite the name, one can always symmetrize these kernels. Class (L) will appear in the discrete chaos processes and the limit theorems considered later.

Definition 4.3.18. We say that a generalized Hermite kernel g on \mathbb{R}_+^k having homogeneity exponent α is of Class (L) (L stands for “limit” as in “limit theorems”), if

1. g is continuous a.e. on \mathbb{R}_+^k ;
2. $|g(\mathbf{x})| \leq g^*(\mathbf{x})$ a.e. $\mathbf{x} \in \mathbb{R}_+^k$, where g^* is a finite linear combination of non-symmetric Hermite kernels: $\prod_{j=1}^k x_j^{\gamma_j}$, where $\gamma_j \in (-1, -1/2)$, $j = 1, \dots, k$, and $\sum_{j=1}^k \gamma_j = \alpha \in (-k/2 - 1/2, -k/2)$.

For example, $g^*(\mathbf{x})$ could be $x_1^{-3/4} x_2^{-5/8} + x_1^{-9/16} x_2^{-13/16}$ if $k = 2$. In this case, $\alpha = -11/8$.

Remark 4.3.19. If two functions g_1 and g_2 on \mathbb{R}_+^k satisfy Condition 2 of Definition 4.3.18, then $\int_{\mathbb{R}_+^k} |g_1(\mathbf{x})g_2(\mathbf{1} + \mathbf{x})| d\mathbf{x} < \infty$ automatically holds, which can be seen by using the

following identity: for any $\gamma, \delta \in (-1, -1/2)$,

$$\int_0^\infty x^\gamma (1+x)^\delta dx = B(\gamma+1, -\gamma-\delta-1),$$

where $B(\cdot, \cdot)$ is the beta function. In addition, $\int_{(0,1]^k} |g_1(\mathbf{x})| d\mathbf{x} < \infty$ also holds.

Proposition 4.3.20. *Class (L) contains Class (B).*

Proof. Suppose g is a generalized Hermite kernel of Class (B). Then there exist constants $C_1, C_2 > 0$, such that

$$|g(\mathbf{x})| \leq C_1 \|\mathbf{x}\|^\alpha = C_1 \left(\sum_{j=1}^k x_j^2 \right)^{\alpha/2} \leq C_2 \prod_{j=1}^k x_j^{\alpha/k},$$

where we have used the arithmetic-geometric mean inequality $k^{-1} \sum_{j=1}^k y_j \geq \left(\prod_{j=1}^k y_j \right)^{1/k}$ and $\alpha < 0$. So Condition 2 of Definition 4.3.18 is satisfied with g^* being a single term where $\gamma_1 = \dots = \gamma_k = \alpha/k$. \square

Remark 4.3.21. In view of Remark 4.3.6 and Remark 4.3.19, one can check that Class (B) or Class (L) if adding in the a.e. 0-valued function, with fixed order k and fixed homogeneity component $\alpha \in (-k/2 - 1/2, -k/2)$, forms an inner product space, with the inner product specified as

$$\begin{aligned} \langle g_1, g_2 \rangle &:= \left\langle \int_0^1 g_1(s\mathbf{1} - \cdot) ds, \int_0^1 g_2(s\mathbf{1} - \cdot) ds \right\rangle_{L^2(\mathbb{R}^k)} \\ &= \frac{1}{2H(2H-1)} \int_{\mathbb{R}_+^k} g_1(\mathbf{x})g_2(\mathbf{1} + \mathbf{x}) + g_1(\mathbf{1} + \mathbf{x})g_2(\mathbf{x}) d\mathbf{x}, \end{aligned}$$

where $H = \alpha + k/2 + 1$, which yields the norm

$$\|g\| := \left\| \int_0^1 g(s\mathbf{1} - \cdot) ds \right\|_{L^2(\mathbb{R}^k)} = \left(\frac{1}{H(2H-1)} \int_{\mathbb{R}_+^k} g(\mathbf{x})g(\mathbf{1} + \mathbf{x}) d\mathbf{x} \right)^{1/2}.$$

Here are several examples.

Example 4.3.22. Suppose $g(\mathbf{x}) = \|\mathbf{x}\|^\alpha$, where $\alpha \in (-1/2 - k/2, -k/2)$. This g belongs to Class (B) and thus also Class (L). The pseudo-Fourier transform (Proposition 4.3.12) of g is $\widehat{g}(\mathbf{u}) = c\|\mathbf{u}\|^{-\alpha-k}$ ((25.25) of Samko et al. [1993]) for some constant $c > 0$, which provides the spectral representation by (4.7).

Example 4.3.23. Another example of Class (B):

$$g(\mathbf{x}) = \frac{\prod_{j=1}^k x_j^{a_j}}{\sum_{j=1}^k x_j^b},$$

where $a_j > 0$ and $b > 0$, yielding a homogeneity exponent $\alpha = \sum_{j=1}^k a_j - b \in (-1/2 - k/2, -k/2)$.

Example 4.3.24. We give yet another example of Class (L) but not (B):

$$g(\mathbf{x}) = g_0(\mathbf{x}) \vee \left(\prod_{j=1}^k x_j^{\alpha/k} \right).$$

where $g_0(\mathbf{x}) > 0$ is any generalized Hermite kernel of Class (B) on \mathbb{R}_+^k with homogeneity exponent α .

4.3.3 Fractionally filtered kernels

According to Theorem 4.3.5, the generalized Hermite process introduced above admits a Hurst coefficient $H > 1/2$ only. To obtain an H -sssi process with $0 < H < 1/2$, we consider the following fractionally filtered kernel:

$$h_t^\beta(\mathbf{x}) = \int_{\mathbb{R}} l_t^\beta(s) g(s\mathbf{1} - \mathbf{x}) 1_{\{s\mathbf{1} > \mathbf{x}\}} ds, \quad (4.11)$$

where g is a generalized Hermite kernel defined in Definition 4.3.2 with homogeneity exponent

$$\alpha \in (-k/2 - 1/2, -k/2),$$

and

$$l_t^\beta(s) = \frac{1}{\beta} \left[(t-s)_+^\beta - (-s)_+^\beta \right], \quad \beta \neq 0. \quad (4.12)$$

One can extend it to $\beta = 0$ by writing $l_t^0(s) = 1_{(0,t]}(s)$, but this would lead us back to the generalized Hermite process case. We hence assume throughout that $\beta \neq 0$. The following proposition gives the range of β for which $I_k(h_t^\beta)$ is well-defined.

Proposition 4.3.25. *If*

$$-1 < -\alpha - \frac{k}{2} - 1 < \beta < -\alpha - \frac{k}{2} < \frac{1}{2}, \quad \beta \neq 0 \quad (4.13)$$

then $h_t^\beta \in L^2(\mathbb{R}^k)$.

Proof.

$$\begin{aligned} \int_{\mathbb{R}^k} h_t^\beta(\mathbf{x})^2 d\mathbf{x} &\leq 2 \int_{-\infty}^{\infty} ds_1 \int_{s_1}^{\infty} ds_2 \int_{\mathbb{R}^k} d\mathbf{x} l_t(s_1) l_t(s_2) |g(s_1 \mathbf{1} - \mathbf{x}) g(s_2 \mathbf{1} - \mathbf{x})| 1_{\{s_1 \mathbf{1} > \mathbf{x}\}} \\ &= 2 \int_{-\infty}^{\infty} ds \int_0^{\infty} du \int_{\mathbb{R}_+^k} d\mathbf{w} l_t^\beta(s) l_t^\beta(s+u) |g(\mathbf{w}) g(u \mathbf{1} + \mathbf{w})| \\ &= 2 \int_{-\infty}^{\infty} ds l_t^\beta(s) \int_0^{\infty} l_t^\beta(s+u) u^{2\alpha+k} du \int_{\mathbb{R}_+^k} d\mathbf{y} |g(\mathbf{y}) g(\mathbf{1} + \mathbf{y})|. \end{aligned} \quad (4.14)$$

We thus focus on showing $\int_{-\infty}^{\infty} ds l_t^\beta(s) \int_0^{\infty} l_t^\beta(s+u) u^{2\alpha+k} du < \infty$. Recall that for any $c > 0$, we have

$$\int_0^c (c-s)^{\gamma_1} s^{\gamma_2} ds = c^{\gamma_1+\gamma_2+1} \int_0^1 (1-s)^{\gamma_1} s^{\gamma_2} ds = c^{\gamma_1+\gamma_2+1} \mathbf{B}(\gamma_1+1, \gamma_2+1), \quad \forall \gamma_1, \gamma_2 > -1.$$

So by noting that $\beta > -1$ and $2\alpha+k > -1$, we have

$$\begin{aligned} \int_0^{\infty} l_t^\beta(s+u) u^{2\alpha+k} du &= \frac{1}{\beta} \int_0^{\infty} \left[(t-s-u)_+^\beta - (-s-u)_+^\beta \right] u^{2\alpha+k} du \\ &= \frac{1}{\beta} \left[\int_0^{t-s} (t-s-u)^\beta u^{2\alpha+k} du + \int_0^{-s} (-s-u)^\beta u^{2\alpha+k} du \right] \end{aligned}$$

$$= \frac{B(\beta + 1, 2\alpha + k + 1)}{\beta} \left[(t - s)_+^{\beta + \delta} - (-s)_+^{\beta + \delta} \right],$$

where

$$\delta = 2\alpha + k + 1 \in (0, 1). \quad (4.15)$$

We thus want to determine when the following holds:

$$\int_{\mathbb{R}} \left((t - s)_+^{\beta} - (-s)_+^{\beta} \right) \left((t - s)_+^{\beta + \delta} - (-s)_+^{\beta + \delta} \right) ds < \infty.$$

Suppose $t > 0$. The potential integrability problems appear near $s = -\infty, 0, t$. Near $s = -\infty$, the integrand behaves like $|s|^{2\beta + \delta - 2}$, and thus we need $2\beta + \delta - 2 < -1$; near $s = 0$, the integrand behaves like $|s|^{2\beta + \delta}$, and thus $2\beta + \delta > -1$; near $s = t$, the integrand behaves like $|t - s|^{2\beta + \delta}$, and thus again $2\beta + \delta > -1$. In view of (4.15), these requirements are satisfied by (4.13). \square

Remark 4.3.26. Using (4.14) we obtain as a byproduct of the preceding proof that if β is in the range given in Proposition 4.3.25, then the function $f_{\mathbf{x}, t}(s) := l_t(s)|g(s\mathbf{1} - \mathbf{x})|1_{\{s\mathbf{1} > \mathbf{x}\}}$ is in $L^1(\mathbb{R})$ for any $t > 0$ and a.e. $\mathbf{x} \in \mathbb{R}^k$.

Theorem 4.3.27. *The process defined by $Z^\beta(t) := I_k(h_t^\beta)$ with h_t^β given in (4.11), namely,*

$$Z^\beta(t) = \int_{\mathbb{R}^k} \int_{\mathbb{R}} \frac{1}{\beta} [(t - s)_+^\beta - (-s)_+^\beta] g(s - x_1, \dots, s - x_k) 1_{\{s > x_1, \dots, s > x_k\}} ds W(dx_1) \dots W(dx_k), \quad (4.16)$$

is an H -sssi process with

$$H = \alpha + \beta + k/2 + 1 \in (0, 1).$$

Proof. By (4.12), one has for any $\lambda > 0$, $l_{\lambda t}^\beta(s) = \lambda^\beta l_t^\beta(\frac{s}{\lambda})$, and for any $t, h > 0$, $l_{t+h}^\beta(s) - l_t^\beta(s) = l_h^\beta(s - t)$. In addition, g is homogeneous with exponent α . The conclusion then follows by Proposition 4.3.1. \square

Remark 4.3.28. In the case $\beta > 0$, one is able to write $l_t^\beta(s) = \int_0^t (r - s)_+^{\beta - 1} dr$, and thus

by Fubini

$$h_t^\beta(\mathbf{x}) = \int_0^t dr \int_{\mathbb{R}} ds (r-s)_+^{\beta-1} g(s\mathbf{1} - \mathbf{x}) 1_{\{s\mathbf{1} > \mathbf{x}\}}. \quad (4.17)$$

Remark 4.3.29. To get the anti-persistent case $H < 1/2$, choose

$$\beta \in (-\alpha - k/2 - 1, -\alpha - k/2 - 1/2).$$

We now state an analog of (4.7) for the spectral representation of the process $Z^\beta(t)$:

Proposition 4.3.30. *Suppose that (4.6) holds. Then the L^2 -sense Fourier transform of h_t^β is*

$$\widehat{h}_t^\beta(\mathbf{u}) = (e^{it\langle \mathbf{u}, \mathbf{1} \rangle} - 1)(i\langle \mathbf{u}, \mathbf{1} \rangle)^{-\beta-1} \widehat{g}(-\mathbf{u}) \Gamma(\beta), \quad \text{a.e. } \mathbf{u} \in \mathbb{R}^k, \quad (4.18)$$

where \widehat{g} is defined in Proposition 4.3.12.

Proof. Let $g_n(\mathbf{x}) = g(\mathbf{x}) 1_{(0,n]^k}(\mathbf{x})$, and $l_{t,n}^\beta = \beta^{-1}[(t-s)_+^\beta 1_{\{t-s < n\}} - (-s)_+^\beta 1_{\{-s < n\}}]$. Set

$$h_{t,n}^\beta(\mathbf{x}) = \int_{\mathbb{R}} l_{t,n}(s) g_n(s\mathbf{1} - \mathbf{x}) ds.$$

Similar to the proof of Proposition 4.3.12, one can show that $h_{t,n}^\beta$ converges in $L^2(\mathbb{R}^k)$ to h_t^β as $n \rightarrow \infty$ through the Dominated Convergence Theorem by noting that $|g_n| \leq |g|$ and $|l_{t,n}^\beta| \leq l_t^\beta$.

Since the truncated $l_{t,n}$ and g_n admit L^1 -Fourier transforms $\widehat{l}_{t,n}$ and \widehat{g}_n respectively, one can write the Fourier transform of $h_{t,n}^\beta$ as:

$$\widehat{h}_{t,n}^\beta(\mathbf{u}) = \widehat{l}_{t,n}(\langle \mathbf{u}, \mathbf{1} \rangle) \widehat{g}_n(-\mathbf{u}),$$

(compare with (4.9)). Since $h_{t,n}^\beta$ converges in $L^2(\mathbb{R}^k)$ to h_t^β as $n \rightarrow \infty$, by Plancherel's isometry, $\widehat{h}_{t,n}^\beta$ converges in $L^2(\mathbb{R}^k)$ to \widehat{h}_t^β . One now needs to identify (4.18) with the limit of $\widehat{h}_{t,n}^\beta$.

We first compute $\widehat{l}_{t,n}^\beta$. When $\beta < 0$, one has by change of variable that

$$\begin{aligned} l_{t,n}^\beta(u) &= \beta^{-1} \left(\int_{\mathbb{R}} e^{iux} (t-x)_+^\beta 1_{\{t-x < n\}} dx - \int_{\mathbb{R}} e^{iux} (-x)_+^\beta 1_{\{-x < n\}} dx \right) \\ &= \beta^{-1} (e^{iut} - 1) \int_0^n e^{-ius} s^\beta ds. \end{aligned} \quad (4.19)$$

When $\beta > 0$, one has

$$l_{t,n}^\beta(u) = \int_{\mathbb{R}} 1_{[0,t)}(x) (x-u)_+^{\beta-1} 1_{\{x-u < n\}} dx = (1_{[0,t)} * b_n)(u),$$

where $b_n(x) = (-x)_+^{\beta-1} 1_{\{-x < n\}}$. We have the Fourier transforms $\widehat{1_{[0,t)}}(u) = \frac{e^{iut}-1}{iu}$, and

$$\widehat{b}_n(u) = \int_{\mathbb{R}} e^{-iux} (-x)_+^{\beta-1} 1_{\{-x < n\}} dx = \int_0^n e^{-ius} s^{\beta-1} ds.$$

So

$$\widehat{l}_{t,n}^\beta(u) = \frac{e^{iut}-1}{iu} \int_0^n e^{-ius} s^{\beta-1} ds \quad (4.20)$$

By Gradshteyn and Ryzhik [2007] Formula 3.761.4 and 3.761.9, for $\mu \in (0, 1)$,

$$\begin{aligned} \lim_{n \rightarrow \infty} \int_0^n e^{-ius} s^{\mu-1} ds &= |u|^{-\mu} \Gamma(\mu) \cos\left(\frac{\mu\pi}{2}\right) - i \operatorname{sign}(u) |u|^{-\mu} \Gamma(\mu) \sin\left(\frac{\mu\pi}{2}\right) \\ &= e^{-i \operatorname{sign}(u) \mu\pi/2} |u|^{-\mu} \Gamma(\mu) = (iu)^{-\mu} \Gamma(\mu), \end{aligned}$$

Combining the foregoing limit with (4.19) and (4.20), we deduce

$$\lim_{n \rightarrow \infty} \widehat{l}_{t,n}^\beta = \widehat{l}_t^\beta(u) := (e^{itu} - 1)(iu)^{-\beta-1} \Gamma(\beta).$$

Recall that there exists a subsequence \widehat{g}_{n_k} converges a.e. to the pseudo-Fourier transform \widehat{g} as $k \rightarrow \infty$ (Proposition 4.3.12). So $\widehat{l}_{t,n_k}(\langle \mathbf{u}, \mathbf{1} \rangle) \widehat{g}_{n_k}(-\mathbf{u})$ converges to $\widehat{l}_t(\langle \mathbf{u}, \mathbf{1} \rangle) \widehat{g}(-\mathbf{u})$ for a.e. $\mathbf{u} \in \mathbb{R}^k$. But at the same time $\widehat{l}_{t,n_k}(\langle \mathbf{u}, \mathbf{1} \rangle) \widehat{g}_{n_k}(-\mathbf{u})$ converges in $L^2(\mathbb{R})^k$ to \widehat{h}_t^β . So we identify \widehat{h}_t^β with the expression in (4.18) \square

Remark 4.3.31. By Proposition 4.2.1, we get a spectral representation $Z^\beta(t) \stackrel{f.d.d.}{=} \widehat{I}(\widehat{h}_t^\beta)$. The kernel (4.18) in the spectral-domain has been considered by Major [1981] in the special case where $\widehat{g}(\mathbf{u}) = c \prod_{j=1}^k |u_j|^{-d}$ is the kernel for the spectral representation of Hermite process.

4.4 Discrete chaos processes

In this section, we introduce a class of stationary sequence which converges to a generalized Hermite process of Class (L) as defined in Definition 4.3.18.

First we define the *discrete chaos*, or the *discrete multiple stochastic integral*, $Q_k(\cdot; \epsilon)$ with respect to the i.i.d. noise $\epsilon := (\epsilon_i, i \in \mathbb{Z})$.

Let h be a function defined in \mathbb{Z}^k such that $\sum'_{\mathbf{i} \in \mathbb{Z}^k} h(\mathbf{i})^2 < \infty$, where $'$ indicate the exclusion of the diagonals $i_p = i_q, p \neq q$. The following sum

$$Q_k(h) = Q_k(h, \epsilon) = \sum_{(i_1, \dots, i_k) \in \mathbb{Z}^k} h(i_1, \dots, i_k) \epsilon_{i_1} \dots \epsilon_{i_k} = \sum'_{\mathbf{i} \in \mathbb{Z}^k} h(\mathbf{i}) \prod_{p=1}^k \epsilon_{i_p}, \quad (4.21)$$

is called the *discrete chaos* of order k . It is easy to see that switching the arguments, say i_p and $i_q, p \neq q$, of $h(i_1, \dots, i_k)$, does not change $Q_k(h)$. So if \tilde{h} is the symmetrization h , then $Q_k(h) = Q_k(\tilde{h})$.

The discrete chaos is related to Wiener chaos by a limit theorem. Suppose now we have a sequence of function vectors $\mathbf{h}_n = (h_{1,n}, \dots, h_{j,n})$ where each $h_{j,n} \in L^2(\mathbb{Z}^{k_j}), j = 1, \dots, J$. The following proposition concerns the convergence of the discrete chaos to the Wiener chaos:

Proposition 4.4.1. *Let $\tilde{h}_{j,n}(\mathbf{x}) = n^{k_j/2} h_{j,n}([\mathbf{n}\mathbf{x}] + \mathbf{c}_j), j = 1, \dots, J$, where $\mathbf{c}_j \in \mathbb{Z}^{k_j}$. Suppose that there exists $h_j \in L^2(\mathbb{R}^{k_j})$, such that*

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

as $n \rightarrow \infty$. Then, as $n \rightarrow \infty$,

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

where each $I_{k_j}(\cdot)$, $j = 1, \dots, J$, denotes the k_j -tuple Wiener-Itô integral with respect to the same standard Brownian motion W .

For a proof, we refer the reader to the proof of Proposition 14.3.2 of Giraitis et al. [2012] on the univariate case. The proof for the multivariate case (corresponding to Proposition 14.3.3 of Giraitis et al. [2012]) is similar once the Crámer-Wold Device is applied. The difference between Proposition 4.4.1 and Proposition 14.3.3 of Giraitis et al. [2012] is that we add the shift \mathbf{c}_j for more flexibility. This extension requires only an easy modification to the proof.

The causal *discrete chaos process* of order $k \geq 1$ is a stationary sequence $\{X(n), n \in \mathbb{Z}\}$ defined by:

$$X(n) = \sum'_{0 < i_1, \dots, i_k < \infty} a(i_1, \dots, i_k) \epsilon_{n-i_1} \dots \epsilon_{n-i_k} = \sum'_{-\infty < i_1, \dots, i_k < n} a(n-i_1, \dots, n-i_k) \epsilon_{i_1} \dots \epsilon_{i_k}, \quad (4.22)$$

where ' indicates that the sum excludes the diagonals $i_p = i_q$, $p \neq q$, $\{\epsilon_n\}$ is an i.i.d. sequence with mean 0 and variance 1, $a(\mathbf{i})$ is a function on \mathbb{Z}^k , and we require that it satisfies $\sum'_{\mathbf{i} > \mathbf{0}} a(\mathbf{i})^2 < \infty$, so that $X(n)$ is well-defined in the $L^2(\Omega)$ -sense. Note that when $k = 1$, $X(n)$ is plainly a linear process.

Due to the off-diagonality, the autocovariance of $\{X(n)\}$ is given by the simple formula

$$\gamma(n) := \text{Cov}(X(n), X(0)) = k! \sum'_{\mathbf{i} > \mathbf{0}} \tilde{a}(\mathbf{i}) \tilde{a}(\mathbf{i} + |n|\mathbf{1}), \quad (4.23)$$

where $\tilde{a}(\cdot)$ is the symmetrization of $a(\cdot)$.

We now focus on the following case:

$$a(\mathbf{i}) = g(\mathbf{i})L(\mathbf{i}), \quad (4.24)$$

where g is a generalized Hermite kernel of Class (L) defined in Definition 4.3.18, and L is a bounded function on \mathbb{Z}_+^k which satisfies the following: for any $\mathbf{x} \in \mathbb{R}_+^k$ and for any bounded \mathbb{Z}^k -valued function $\mathbf{B}(\cdot)$ defined on \mathbb{Z}_+ , we have

$$L([n\mathbf{x}] + \mathbf{B}(n)) \rightarrow 1, \text{ as } n \rightarrow \infty. \quad (4.25)$$

Note that $X(n)$ is well-defined in $L^2(\Omega)$ since $\sum_{\mathbf{i} \in \mathbb{Z}_+^k} g^*(\mathbf{i})^2 < \infty$, where g^* is a linear combination of terms of the form $\prod_{j=1}^k x_j^{\gamma_j}$ with every $\gamma_j < -1/2$,

Remark 4.4.2. Note that the boundedness of L and (4.25) are strictly weaker than assuming that $L(\mathbf{i}) \rightarrow 1$ as $\|\mathbf{i}\| \rightarrow \infty$ for some norm $\|\cdot\|$ on \mathbb{R}^k (recall that norms are equivalent in the finite-dimensional space). Indeed, consider

$$L(i_1, i_2) = \begin{cases} 2 & \text{if } i_2 = 1; \\ 1 & \text{otherwise.} \end{cases}$$

Suppose that \mathbf{B} is bounded by M . Then $L([n\mathbf{x}] + \mathbf{B}(n)) = 1$ for large n . On the other hand, consider $\|\mathbf{i}\| = \max(i_1, i_2)$. Then if $(i_1, i_2) = (i_1, 1)$, $i_1 \rightarrow \infty$, we have $\|\mathbf{i}\| = i_1 \rightarrow \infty$ but $L(i_1, i_2) = L(i_1, 1) = 2$.

Remark 4.4.3. In practice, Relation (4.25) implies that for any fixed $\mathbf{x} \in \mathbb{R}_+^k$ and $\mathbf{c} \in \mathbb{Z}_+^k$, $L([n\mathbf{x}] + \mathbf{c}) \rightarrow 1$ as $n \rightarrow \infty$.

The following Proposition shows that one can get long-range dependence if g is of Class (L).

Proposition 4.4.4. *If $a(\mathbf{i})$ is as given in (4.24), where g has homogeneity exponent $\alpha \in (-1/2 - k/2, -k/2)$ (or $2\alpha + k \in (-1, 0)$), then the autocovariance of the discrete chaos*

process $\{X(n)\}$ satisfies

$$\gamma(n) \sim k!C_{\tilde{g}}n^{2H-2}, \text{ as } n \rightarrow \infty, \quad (4.26)$$

where $C_{\tilde{g}} = \int_{\mathbb{R}_+^k} \tilde{g}(\mathbf{x})\tilde{g}(\mathbf{1} + \mathbf{x}) > 0$, $H = \alpha + k/2 + 1 \in (1/2, 1)$, with \tilde{g} being the symmetrization of g . In addition, as $N \rightarrow \infty$,

$$\text{Var}\left[\sum_{n=1}^N X(n)\right] \sim \frac{k!C_{\tilde{g}}}{H(2H-1)}N^{2H}. \quad (4.27)$$

Proof. Assume without loss of generality that g is already symmetric.

$$\begin{aligned} (k!)^{-1}\gamma(n) &= \sum_{\mathbf{i} > \mathbf{0}}^l g(\mathbf{i})g(n\mathbf{1} + \mathbf{i})L(n\mathbf{1} + \mathbf{i})L(\mathbf{i}) \\ &= n^{2\alpha+k} \sum_{\mathbf{i} > \mathbf{0}}^l g\left(\frac{\mathbf{i}}{n}\right)g\left(\mathbf{1} + \frac{\mathbf{i}}{n}\right)L(\mathbf{i})L(n\mathbf{1} + \mathbf{i})\frac{1}{n^k} \\ &= n^{2\alpha+k} \int_{\mathbb{R}_+^k} 1_{D_n^c}(\mathbf{x})g_n(\mathbf{x})g_n(\mathbf{1} + \mathbf{x})d\mathbf{x}, \end{aligned}$$

where $g_n(\mathbf{x}) = g\left(\frac{[n\mathbf{x}] + \mathbf{1}}{n}\right)L([n\mathbf{x}] + \mathbf{1})$, $D_n^c = \{\mathbf{x} \in \mathbb{R}_+^k, [nx_p] \neq [nx_q], p \neq q \in \{1, \dots, k\}\}$. Note that $1_{D_n}(\mathbf{x}) = 1$ as n becomes large enough, for any $\mathbf{x} \in D^c := \{\mathbf{x} \in \mathbb{R}_+^k, x_p \neq x_q, p \neq q \in \{1, \dots, k\}\}$, and that the diagonal set $D := \mathbb{R}_+^k \setminus D^c$ has measure 0. Since g belongs to Class (L), g is continuous a.e., so $g_n(\mathbf{x}) \rightarrow g(\mathbf{x})$ a.e. as $n \rightarrow \infty$. Furthermore, there exists $g^*(\mathbf{x})$ which is a linear combination of the form $\prod_{j=1}^k x_j^{\gamma_j}$ (Condition 2 of Definition 4.3.18), so that for a.e. $\mathbf{x} \in \mathbb{R}_+^k$,

$$|g_n(\mathbf{x})| \leq g^*\left(\frac{[n\mathbf{x}] + \mathbf{1}}{n}\right) \leq g^*(\mathbf{x}),$$

since L is bounded and g^* is decreasing in its every variable. Note that $\int_{\mathbb{R}_+^k} g^*(\mathbf{x})g^*(\mathbf{1} + \mathbf{x})d\mathbf{x} < \infty$, and g is a.e. continuous. So it remains to apply the Dominated Convergence Theorem.

Finally, (4.27) follows by first noting that

$$\text{Var}\left[\sum_{n=1}^N X(n)\right] = \sum_n (N - |n|)\gamma(n) = N \sum_{|n|<N} \gamma(n) - \sum_{|n|<N} |n|\gamma(n),$$

and then using the asymptotics of $\gamma(n)$ just derived. □

4.5 Hypercontractivity for infinite discrete chaos

Let X_M be a finite discrete chaos defined as

$$X_M = \sum_{-M\mathbf{1} \leq \mathbf{i} \leq M\mathbf{1}} h(\mathbf{i}) \epsilon_{i_1} \dots \epsilon_{i_k}, \quad (4.28)$$

where $h(\mathbf{i}) = h(i_1, \dots, i_k)$ is a function on \mathbb{Z}^k , $M \in \mathbb{Z}_+$, and we assume that $\{\epsilon_i\}$ is a sequence of i.i.d. variables with $\mathbb{E}\epsilon_i = 0$, $\mathbb{E}\epsilon_i^2 = 1$. Then we have the following moment-comparison inequality, also called “hypercontractivity inequality”:

Proposition 4.5.1. *Suppose that $\mathbb{E}|\epsilon_i|^p < \infty$ with $p \geq 2$. Then*

$$\mathbb{E}[|X_M|^p]^{1/p} \leq d_{p,k} \mathbb{E}[|X_M|^2]^{1/2}, \quad (4.29)$$

where $d_{p,k}$ is a constant depending only on p and k .

For a proof of (4.29), where M is finite, see Lemma 4.3 of Krakowiak and Szulga [1986], where the so-called MPZ(p) condition (Definition 1.5 of Krakowiak and Szulga [1986]) is trivially satisfied since the ϵ_i ’s are identically distributed.

Now we extend (4.29) to the case $M = \infty$. The result is used in Theorem 4.6.3, 4.6.11 and 4.6.14 below for proving tightness in $D[0, 1]$.

Proposition 4.5.2. *Suppose that $\sum'_{\mathbf{i} \in \mathbb{Z}^k} h(\mathbf{i})^2 < \infty$. Let $X = \sum'_{\mathbf{i} \in \mathbb{Z}^k} h(\mathbf{i}) \prod_{p=1}^k \epsilon_{i_p}$. If for*

some $p' > p > 2$, $E|\epsilon_i|^{p'} < \infty$, then one has

$$E[|X|^p]^{1/p} \leq d_{p,k} E[|X|^2]^{1/2} \quad (4.30)$$

Proof. Let X_M be the truncated finite chaos as in (4.28). The condition on h implies that $X_M \rightarrow X$ in $L^2(\Omega)$. Moreover, one has by (4.29),

$$E[|X_M|^{p'}] \leq d_{p',k}^{p'} E[|X_M|^2]^{p'/2} \leq d_{p',k}^{p'} \left(\sum_{\mathbf{i} \in \mathbb{Z}^k} h(\mathbf{i})^2 \right)^{p'/2}.$$

This implies that $\{|X_M|^p, M \geq 1\}$ and $\{|X_M|^2, M \geq 1\}$ are uniformly integrable, implying convergence of the corresponding moments. So one can then let $M \rightarrow \infty$ on both sides of (4.29) and obtain (4.30). \square

4.6 Joint convergence of the discrete chaoses

Our goal here is to obtain non-central limit theorems for the discrete chaos process introduced in Section 4.4. We shall, in fact, prove both a central limit theorem for the SRD case (getting Brownian motion as limit) and a non-central limit theorem for the LRD case (getting the generalized Hermite process introduced in Section 4.3 as limit). We also consider non-central limit theorems leading to the fractionally filtered generalized Hermite process introduced in Section 4.3.3. Finally, we derive a multivariate limit theorem which mixes central and non-central limit theorems.

We first define here precisely what SRD and LRD stand for in the context of discrete chaos process. Recall that $\tilde{a}(\cdot)$ denotes the symmetrization of $a(\cdot)$.

Definition 4.6.1. We say a discrete chaos process $\{X(n)\}$ given in (4.22) is

- SRD, if $\sum_{n=-\infty}^{\infty} \sum_{\mathbf{i} > \mathbf{0}} |\tilde{a}(\mathbf{i})\tilde{a}(\mathbf{i} + |n|\mathbf{1})| < \infty$ and $\sum_{n=-\infty}^{\infty} \gamma(n) > 0$;
- LRD, if $a(\mathbf{i}) = g(\mathbf{i})L(\mathbf{i})$ as given in (4.24). In particular, g is a generalized Hermite kernel of Class (L).

Remark 4.6.2. The definitions of SRD and LRD in Definition 4.6.1 are distinct. Indeed, the SRD condition implies that $\sum_n |\gamma(n)| < \infty$, while LRD yields $\sum_n |\gamma(n)| = \infty$ by Proposition 4.4.4.

4.6.1 Central limit theorem

Theorem 4.6.3. *If a discrete chaos process $\{X(n)\}$ given in (4.22) is SRD in the sense of Definition 4.6.1, then*

$$\frac{1}{N^{1/2}} \sum_{n=1}^{[Nt]} X(n) \xrightarrow{f.d.d.} \sigma B(t) \quad (4.31)$$

where $B(t)$ is a standard Brownian motion, and $\sigma^2 = \sum_{n=-\infty}^{\infty} \gamma(n)$.

Proof. Assume without loss of generality that $a(\cdot)$ is symmetric. The proof is similar to the proof of Theorem 4.2.3 found on p.108 of Giraitis et al. [2012], so we give only a sketch. The central idea is to introduce the m -truncation of $X(n)$, namely, $X^{(m)}(n) := \sum_{\mathbf{0} < \mathbf{i} \leq \mathbf{m}\mathbf{1}} a(\mathbf{i}) \prod_{j=1}^k \epsilon_{n-i_j}$, and then let $m \rightarrow \infty$. The sequence $\{X^{(m)}(n), n \in \mathbb{Z}\}$ is m -dependent, so the classical invariance principle applies (Billingsley [1956] Theorem 5.2). The long-run variance $\sigma^2 = \sum_n \gamma(n)$ is a standard result. We now check that the $L^2(\Omega)$ approximation is valid as $m \rightarrow \infty$, that is,

$$\lim_{m \rightarrow \infty} \sup_{N \in \mathbb{Z}_+} \text{Var}[Y_N^{(m)}(t) - Y_N(t)] = 0, \quad t > 0, \quad (4.32)$$

where $Y_N^{(m)}(t) = \frac{1}{\sqrt{N}} \sum_{n=1}^{[Nt]} X^{(m)}(n)$ and $Y_N(t) = \frac{1}{\sqrt{N}} \sum_{n=1}^{[Nt]} X(n)$, which is similar to (4.8.7) of Giraitis et al. [2012]. Indeed,

$$\begin{aligned} \text{Var}[Y_N^{(m)}(t) - Y_N(t)] &= \frac{1}{N} \text{Var} \left[\sum_{n=1}^{[Nt]} (X_n^{(m)} - X_n) \right] \\ &= \frac{[Nt]}{N} \sum_{|n| < [Nt]} \gamma_m(n) \left(1 - \frac{|n|}{[Nt]}\right) \leq t \sum_{n=-\infty}^{\infty} |\gamma_m(n)|, \end{aligned} \quad (4.33)$$

where

$$\gamma_m(n) := \mathbb{E}(X_n - X_n^{(m)})(X_0 - X_0^{(m)}) = k! \sum_{\mathbf{i} > m\mathbf{1}} a(\mathbf{i})a(n\mathbf{1} + \mathbf{i}).$$

For a fixed $n \in \mathbb{Z}$, $\gamma_m(n) \rightarrow 0$ as $m \rightarrow \infty$, and $|\gamma_m(n)| \leq \rho(n)$, where

$$\rho(n) = k! \sum_{\mathbf{i} > \mathbf{0}} |a(\mathbf{i})a(\mathbf{i} + n\mathbf{1})|,$$

which satisfies $\sum_n \rho(n) < \infty$ by the SRD assumption in Definition 4.6.1. Since the bound in (4.33) does not depend on N , the Dominated Convergence Theorem applies and thus (4.32) holds. \square

To strengthen the conclusion of Theorem 4.6.3 to weak convergence, we have to make some additional assumptions to prove tightness.

Theorem 4.6.4. *Theorem 4.6.3 holds with $\xrightarrow{f.d.d.}$ replaced by weak convergence \Rightarrow in $D[0, 1]$, if either of the following holds:*

1. *There exists $\delta > 0$, such that $\mathbb{E}(|\epsilon_i|^{2+\delta}) < \infty$;*
2. *There exists an $M > 0$ such that $a(\mathbf{i}) = 0$ whenever $\mathbf{i} > M\mathbf{1}$.*

Proof. Look first at case 1. Let

$$Y_N(t) := \frac{1}{\sqrt{N}} \sum_{n=1}^{\lfloor Nt \rfloor} X(n)$$

Select $p \in (2, 2 + \delta)$. By Proposition 4.5.2, one has

$$\mathbb{E}[|Y_N(t) - Y_N(s)|^p] \leq c\mathbb{E}[|Y_N(t) - Y_N(s)|^2]^{p/2}, \quad (4.34)$$

where c is some constant which doesn't depend on s, t or N . Note that $\sum_n |\gamma(n)| < \infty$

due to SRD assumption, we have

$$\begin{aligned} \mathbb{E} \left[|Y_N(t) - Y_N(s)|^2 \right] &= \frac{1}{N} \mathbb{E} \left[\left| \sum_{n=1}^{\lfloor Nt \rfloor - \lfloor Ns \rfloor} X(n) \right|^2 \right] \\ &= \frac{\lfloor Nt \rfloor - \lfloor Ns \rfloor}{N} \sum_{|n| < \lfloor Nt \rfloor - \lfloor Ns \rfloor} \left(1 - \frac{|n|}{\lfloor Nt \rfloor - \lfloor Ns \rfloor} \right) \gamma(n) \leq \frac{\lfloor Nt \rfloor - \lfloor Ns \rfloor}{N} \sum_{n=-\infty}^{\infty} |\gamma(n)|. \end{aligned} \quad (4.35)$$

Combining (4.34) and (4.35), we have for some constant $C > 0$ that

$$\mathbb{E}[|Y_N(t) - Y_N(s)|^p] \leq cE[|Y_N(t) - Y_N(s)|^2]^{p/2} \leq C|F_N(t) - F_N(s)|^{p/2},$$

where $F_N(t) = \lfloor Nt \rfloor / N$. Now by applying Lemma 4.4.1 and Theorem 4.4.1 of Giraitis et al. [2012], noting that $p/2 > 1$, we conclude that tightness holds.

For case 2, $X(n)$ is M -dependent, so by Theorem 5.2 of Billingsley [1956] tightness holds as well.

□

4.6.2 Non-central limit theorem

The following theorem shows that in the LRD case, the discrete chaos process converges weakly to a generalized Hermite process.

Theorem 4.6.5. *If a discrete chaos process $\{X(n)\}$ given in (4.22) is LRD in the sense of Definition 4.6.1, then*

$$\frac{1}{N^H} \sum_{n=1}^{\lfloor Nt \rfloor} X(n) \Rightarrow Z(t), \quad (4.36)$$

in $D[0, 1]$, where $Z(t)$ is the generalized Hermite process in (4.5), and

$$H = \alpha + k/2 + 1 \in \left(\frac{1}{2}, 1 \right),$$

where $\alpha \in (-1/2 - k/2, -k/2)$ is the homogeneity exponent of g and k is the order of $\{X(n)\}$.

Proof. Tightness in $D[0, 1]$ is standard since $H > 1/2$. We only need to show convergence in finite-dimensional distributions. Assume for simplicity that $a(\mathbf{i}) = g(\mathbf{i})$ or equivalently $L(\mathbf{i}) = 1$. The inclusion of a general L can be done as in the proof of Proposition 4.4.4. We want to show that

$$\frac{1}{N^H} \sum_{n=1}^{[Nt]} X(n) = \sum_{(i_1, \dots, i_k) \in \mathbb{Z}^k} \frac{1}{N^{\alpha+k/2+1}} \sum_{n=1}^{[Nt]} g(n\mathbf{1} - \mathbf{i}) 1_{\{n\mathbf{1} > \mathbf{i}\}} \epsilon_{i_1} \dots \epsilon_{i_k} =: Q_k(h_{t,N}) \xrightarrow{f.d.d.} Z(t), \quad (4.37)$$

where $Q_k(\cdot)$ is defined in (4.21). Now in view of Proposition 4.4.1, we only need to check that

$$\|\tilde{h}_{t,N}(\mathbf{x}) - h_t(\mathbf{x})\|_{L^2(\mathbb{R}^k)} \rightarrow 0, \quad (4.38)$$

where

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

and

$$\begin{aligned} \tilde{h}_{t,N}(\mathbf{x}) &:= N^{k/2} h_{t,N}([N\mathbf{x}] + \mathbf{1}) = \frac{1}{N^{\alpha+1}} \sum_{n=1}^{[Nt]} g(n\mathbf{1} - [N\mathbf{x}] - \mathbf{1}) 1_{\{n\mathbf{1} > [N\mathbf{x}] + \mathbf{1}\}} \\ &= \sum_{n=1}^{[Nt]} g\left(\frac{n\mathbf{1} - [N\mathbf{x}] - \mathbf{1}}{N}\right) 1_{\{n\mathbf{1} > [N\mathbf{x}] + \mathbf{1}\}} \frac{1}{N} \\ &= \int_0^t g\left(\frac{[Ns\mathbf{1}] - [N\mathbf{x}]}{N}\right) 1_{\{[Ns\mathbf{1}] > [N\mathbf{x}]\}} ds - R_N(t, \mathbf{x}). \end{aligned}$$

where

$$R_N(t, \mathbf{x}) = \frac{Nt - [Nt]}{N} g\left(\frac{[Nt\mathbf{1}] - [N\mathbf{x}]}{N}\right) 1_{\{[Nt\mathbf{1}] > [N\mathbf{x}]\}}.$$

Note that we have replaced \mathbf{i} by $[N\mathbf{x}] + \mathbf{1}$ and n by $[Ns] + 1$. By Condition 2 in Definition 4.3.18, there exists a positive generalized Hermite kernel $g^*(\mathbf{x})$ which is a linear combination of the form $\prod_{j=1}^k x_j^{\gamma_j}$, such that $|g(\mathbf{x})| \leq g^*(\mathbf{x})$ for a.e. $\mathbf{x} \in \mathbb{R}_+^k$. We assume without loss of

generality that $g^*(\mathbf{x}) = \prod_{j=1}^k x_j^{\gamma_j}$. Since $[Ns\mathbf{1}] > [N\mathbf{x}]$ implies $s\mathbf{1} > \mathbf{x}$, we have

$$\left| g\left(\frac{[Ns\mathbf{1}] - [N\mathbf{x}]}{N}\right) \right| 1_{\{[Ns\mathbf{1}] > [N\mathbf{x}]\}} \leq \left(\prod_{j=1}^k \left(\frac{[Ns] - [Nx_j]}{N}\right)^{\gamma_j} 1_{\{[Ns] > [Nx_j]\}} \right) 1_{\{s\mathbf{1} > \mathbf{x}\}} \text{ a.e..} \quad (4.39)$$

Moreover, if $0 < [Ns] - [Nx] = k \in \mathbb{Z}_+$, then $Ns - 1 - Nx \leq k$, and hence $s - x \leq \frac{k+1}{N}$.

So we have for any $\gamma < 0$ that

$$\begin{aligned} \sup_{N \geq 1, [Ns] > [Nx]} \left(\frac{[Ns] - [Nx]}{N}\right)^\gamma (s - x)^{-\gamma} &\leq \sup_{N \geq 1, [Ns] - [Nx] = k \geq 1} \left(\frac{k}{N}\right)^\gamma (s - x)^{-\gamma} \\ &\leq \sup_{N \geq 1, k \geq 1} \left(\frac{k}{N}\right)^\gamma \left(\frac{k+1}{N}\right)^{-\gamma} = 2^{-\gamma}. \end{aligned} \quad (4.40)$$

So we have for some constant $C > 0$,

$$\left| g\left(\frac{[Ns\mathbf{1}] - [N\mathbf{x}]}{N}\right) \right| 1_{\{[Ns\mathbf{1}] > [N\mathbf{x}]\}} \leq C g^*(s\mathbf{1} - \mathbf{x}) 1_{\{s\mathbf{1} > \mathbf{x}\}}. \quad (4.41)$$

Since $g(\mathbf{x})$ by assumption of Class (L) is continuous a.e., $g\left(\frac{[Ns\mathbf{1}] - [N\mathbf{x}]}{N}\right) 1_{\{[Ns\mathbf{1}] > [N\mathbf{x}]\}}$ converges a.e. to $g(s\mathbf{1} - \mathbf{x}) 1_{\{s\mathbf{1} > \mathbf{x}\}}$ as $N \rightarrow \infty$. In view of (4.41), and noting that

$$\int_{\mathbb{R}^k} d\mathbf{x} \left(\int_0^t g^*(s\mathbf{1} - \mathbf{x}) 1_{\{s\mathbf{1} > \mathbf{x}\}} ds \right)^2 < \infty$$

because g^* is a generalized Hermite kernel, one then applies the Dominated Convergence Theorem to conclude the L^2 convergence of $\int_0^t g\left(\frac{[Ns\mathbf{1}] - [N\mathbf{x}]}{N}\right) 1_{\{[Ns\mathbf{1}] > [N\mathbf{x}]\}} ds$ to $h_t(\mathbf{x})$. For the remainder term $R_{N,t}(\mathbf{x})$, one has

$$\|R_{N,t}(\mathbf{x})\|_{L^2(\mathbb{R}^k)}^2 = N^{-2H} (Nt - [Nt])^2 \sum_{\mathbf{i} > \mathbf{0}} g(\mathbf{i})^2 \rightarrow 0$$

as $N \rightarrow \infty$. The proof is thus complete. \square

Example 4.6.6. Consider the kernel $g(\mathbf{x})$ defined in (4.1). It belongs to Class (L) by Example 4.3.24. Hence by Theorem 4.6.5, we have the following weak convergence in

$D[0, 1]$:

$$\begin{aligned} & \frac{1}{N^H} \sum_{n=1}^{[Nt]} \sum'_{(i_1, \dots, i_k) \in \mathbb{Z}_+^k} \left(\frac{\prod_{j=1}^k i_j}{\sum_{j=1}^k i_j^{k-\alpha}} \vee \prod_{j=1}^k i_j^{\alpha/k} \right) \epsilon_{n-i_1} \dots \epsilon_{n-i_k} \Rightarrow \\ & \int_{\mathbb{R}^k}' \int_0^t \left(\frac{\prod_{j=1}^k (s-x_j)_+}{\sum_{j=1}^k (s-x_j)_+^{k-\alpha}} \right) \vee \left(\prod_{j=1}^k (s-x_j)_+^{\alpha/k} \right) ds W(dx_1) \dots W(dx_k), \end{aligned}$$

where $H = \alpha + k/2 + 1$.

4.6.3 Non-central limit theorem with fractional filter

In the spirit of Rosenblatt [1979] and Major [1981], we consider here the non-central limit theorem for the fractionally filtered generalized Hermite process introduced in Section 4.3.3. Assume throughout that the generalized Hermite kernel g is of Class (L) (Definition 4.3.18).

Definition 4.6.7. Let $X(n) = \sum'_{\mathbf{i} < n\mathbf{1}} a(n\mathbf{1} - \mathbf{i}) \prod_{j=1}^k \epsilon_{i_j}$ be the same discrete chaos process as in Theorem 4.6.5. We say that a discrete process $U(n)$ is fLRD (fractionally-filtered LRD discrete chaos process) if

$$U(n) = \sum_{m=1}^{\infty} C_m X(n-m) = \sum_{m=-\infty}^{n-1} C_{n-m} \sum'_{\mathbf{i} < m\mathbf{1}} a(m\mathbf{1} - \mathbf{i}) \prod_{j=1}^k \epsilon_{i_j}, \quad (4.42)$$

where $a(\mathbf{i}) = g(\mathbf{i})L(\mathbf{i})$ as in (4.24) with g being a generalized Hermite kernel in Class (L),

$$C_n \sim cn^{\beta-1}$$

as $n \rightarrow \infty$, and where, as in Proposition 4.3.25,

$$\beta \in \left(-\frac{2\alpha + k + 2}{2}, -\frac{2\alpha + k}{2} \right). \quad (4.43)$$

$U(n)$ is well-defined in the $L^2(\Omega)$ sense. Indeed, we have the following:

Lemma 4.6.8. *We have*

$$\sum_{\mathbf{i} \in \mathbb{Z}^k} \left(\sum_{m < n} |C_{n-m} a(m\mathbf{1} - \mathbf{i})| 1_{\{m\mathbf{1} > \mathbf{i}\}} \right)^2 < \infty.$$

Proof. Note that $a(\cdot) = g(\cdot)L(\cdot)$, where g is of Class (L). So by Definition 4.3.18, there exists a $g^*(\mathbf{x}) > 0$ which is a finite linear combination of the form $\prod_{j=1}^k x_j^{\gamma_j}$, such that $|g(\mathbf{x})| < g^*(\mathbf{x})$. Note that L is bounded and $|C_n| \leq cn^{\beta-1}$. Set $n = -1$ without loss of generality due to stationarity. We hence need to show that

$$\sum_{\mathbf{i} \in \mathbb{Z}^k} \left(\sum_{m < -1} (-m)^{\beta-1} g^*(m\mathbf{1} - \mathbf{i}) 1_{\{m\mathbf{1} > \mathbf{i}\}} \right)^2 < \infty. \quad (4.44)$$

It suffices to show this when $\beta > 0$, since for any $\beta' \leq 0$ and $\beta > 0$, $(-m)^{\beta'-1} \leq (-m)^{\beta-1}$ for all $m < -1$. The preceding sum can be rewritten as an integral by replacing m by $[s]$ and \mathbf{i} by $[\mathbf{x}]$:

$$\int_{\mathbb{R}^k} 1_{D^c} d\mathbf{x} \left(\int_{-\infty}^{-1} ds (-[s])^{\beta-1} g^*([s\mathbf{1}] - [\mathbf{x}]) 1_{\{[s\mathbf{1}] > [\mathbf{x}]\}} \right)^2, \quad (4.45)$$

where $D^c = \{\mathbf{x} \in \mathbb{R}^k : [x_p] \neq [x_q], p \neq q\}$. By $[s] \leq s$, $\beta - 1 < 0$, and (4.41), (4.45) is bounded by (up to a constant)

$$\begin{aligned} & \int_{\mathbb{R}^k} d\mathbf{x} \left(\int_{-\infty}^{-1} ds (-s)_+^{\beta-1} g^*(s\mathbf{1} - \mathbf{x}) 1_{\{s\mathbf{1} > \mathbf{x}\}} \right)^2 \\ &= \int_{-\infty}^{-1} ds (-s)^{\beta-1} \int_0^{-s} du (-s-u)^{\beta-1} u^{2\alpha+k} \int_{\mathbb{R}_+^k} dy g^*(\mathbf{y}) g^*(\mathbf{1} + \mathbf{y}) \\ &= \int_1^{\infty} s^{2\alpha+2\beta+k-1} ds B(\beta, 2\alpha+k+1) C_{g^*} < \infty, \end{aligned}$$

where we have used a change of variable similar to the lines below (4.14), and in addition the assumptions $\beta > 0$, $2\alpha + k > -1$, $2\alpha + 2\beta + k < 0$, and g^* is a generalized Hermite kernel.

□

Remark 4.6.9. Lemma 4.6.8 not only shows that $U(n)$ is well-defined in $L^2(\Omega)$, it also allows changing the order of summations, which will be used in proving the non-central limit theorem below.

Next we want to obtain non-central limit theorems, that is, to show that the suitably normalized partial sum of $U(n)$ defined in (4.42) converges to the fractionally-filtered generalized Hermite process introduced in Section 4.3.3. We need to distinguish two cases: $\beta > 0$ (which increases H) and $\beta < 0$ (which decreases H).

We first consider $\beta > 0$:

Theorem 4.6.10. *Let $U(n)$ be as in (4.42) with $\beta \in (0, -\alpha - k/2)$. Then*

$$\frac{1}{NH} \sum_{n=1}^{[Nt]} U(n) \Rightarrow Z^\beta(t),$$

where

$$1/2 < \alpha + k/2 + 1 < H = \alpha + \beta + k/2 + 1 < 1,$$

and $Z^\beta(t)$ is the fractionally-filtered generalized Hermite process defined in Theorem 4.3.27. It is defined using the same g and β as $U(n)$.

Proof. Since $H > 1/2$, tightness in $D[0, 1]$ is standard. We now show convergence in finite-dimensional distributions. Assume for simplicity that $C_m = m^{\beta-1}$ and $L(\mathbf{i}) = 1$. By Lemma 4.6.8, we are able to change the order of the summations to write:

$$\begin{aligned} \frac{1}{NH} \sum_{n=1}^{[Nt]} U(n) &= \sum_{\mathbf{i} \in \mathbb{Z}^k} \frac{1}{NH} \sum_{n=1}^{[Nt]} \sum_{m < n} (n-m)^{\beta-1} g(m\mathbf{1} - \mathbf{i}) \mathbf{1}_{\{m\mathbf{1} > \mathbf{i}\}} \prod_{j=1}^k \epsilon_{i_j} \\ &= \sum_{\mathbf{i} \in \mathbb{Z}^k} h_{t,N}^\beta(\mathbf{i}) \prod_{j=1}^k \epsilon_{i_j} = Q_k(h_{t,N}^\beta), \end{aligned}$$

and by setting $\tilde{h}_{t,N}^\beta(\mathbf{x}) = N^{k/2} h_{t,N}^\beta([N\mathbf{x}] + \mathbf{1})$, we have

$$\tilde{h}_{t,N}^\beta(\mathbf{x}) = \frac{1}{N^{\alpha+\beta+1}} \sum_{n=1}^{[Nt]} \sum_{m < n} (n-m)^{\beta-1} g(m\mathbf{1} - [N\mathbf{x}] - \mathbf{1}) \mathbf{1}_{\{m\mathbf{1} > [N\mathbf{x}] - \mathbf{1}\}}$$

$$\begin{aligned}
&= \sum_{n=1}^{\lfloor Nt \rfloor} \sum_{m < n} \left(\frac{n-m}{N} \right)^{\beta-1} g \left(\frac{m\mathbf{1} - \lfloor N\mathbf{x} \rfloor - \mathbf{1}}{N} \right) \mathbf{1}_{\{m\mathbf{1} > \lfloor N\mathbf{x} \rfloor - \mathbf{1}\}} \frac{1}{N^2} \\
&= \int_0^t ds \int_{\mathbb{R}} dr \left(\frac{\lfloor Ns \rfloor - \lfloor Nr \rfloor}{N} \right)_+^{\beta-1} g \left(\frac{\lfloor Nr\mathbf{1} \rfloor - \lfloor N\mathbf{x} \rfloor}{N} \right) \mathbf{1}_{\{\lfloor Nr\mathbf{1} \rfloor > \lfloor N\mathbf{x} \rfloor\}} - R_{N,t}(\mathbf{x}) \\
&=: \int_0^t ds \int_{\mathbb{R}} dr G_N(s, r, \mathbf{x}) \mathbf{1}_{K_N} - R_{N,t}(\mathbf{x})
\end{aligned}$$

where we associate \mathbf{i} with $\lfloor N\mathbf{x} \rfloor + \mathbf{1}$, n with $\lfloor Ns \rfloor + 1$, and m with $\lfloor Nr \rfloor + 1$,

$$G_N(s, r, \mathbf{x}) := \left(\frac{\lfloor Ns \rfloor - \lfloor Nr \rfloor}{N} \right)_+^{\beta-1} g \left(\frac{\lfloor Nr\mathbf{1} \rfloor - \lfloor N\mathbf{x} \rfloor}{N} \right),$$

$$K_N = \{\lfloor Ns \rfloor > \lfloor Nr \rfloor, \lfloor Nr\mathbf{1} \rfloor > \lfloor N\mathbf{x} \rfloor\} \subset \{s > r, r\mathbf{1} > \mathbf{x}\},$$

and

$$R_{N,t}(\mathbf{x}) = \frac{Nt - \lfloor Nt \rfloor}{N} \int_{\mathbb{R}} dr \left(\frac{\lfloor Nt \rfloor - \lfloor Nr \rfloor}{N} \right)_+^{\beta-1} g \left(\frac{\lfloor Nr\mathbf{1} \rfloor - \lfloor N\mathbf{x} \rfloor}{N} \right) \mathbf{1}_{\{\lfloor Nr\mathbf{1} \rfloor > \lfloor N\mathbf{x} \rfloor\}}.$$

In view of Proposition 4.4.1, we need to show that $\tilde{h}_{t,N}^\beta \rightarrow h_t^\beta$ and $R_{N,t} \rightarrow 0$ in $L^2(\mathbb{R}^k)$, where

$$h_t^\beta(\mathbf{x}) := \int_0^t ds \int_{\mathbb{R}} dr (s-r)_+^{\beta-1} g(r\mathbf{1} - \mathbf{x}) \mathbf{1}_{\{r\mathbf{1} > \mathbf{x}\}}.$$

Using (4.39) and (4.40) (note that $\beta - 1 < 0$) as in the proof of Theorem 4.6.5, we can bound the integrand as

$$|G_N(s, r, \mathbf{x})| \mathbf{1}_{K_N} \leq C(s-r)_+^{\beta-1} g^*(r\mathbf{1} - \mathbf{x}) \mathbf{1}_{\{r\mathbf{1} > \mathbf{x}\}}$$

for some $C > 0$, where $g^*(\mathbf{x})$ is a generalized Hermite kernel from Definition 4.3.18. Because

$$h^*(\mathbf{x}) := (s-r)_+^{\beta-1} g^*(r\mathbf{1} - \mathbf{x}) \mathbf{1}_{\{r\mathbf{1} > \mathbf{x}\}} \in L^2(\mathbb{R}^k)$$

by (4.17) and Proposition 4.3.25, and g is a.e. continuous, it remains to apply the Dominated Convergence Theorem to conclude $\tilde{h}_{t,N}^\beta \rightarrow h_t^\beta$. For the remainder term $R_{N,t}(\mathbf{x})$, one

has

$$\|R_{N,t}(\mathbf{x})\|_{L^2(\mathbb{R}^k)}^2 = N^{-2H} (Nt - [Nt]) \sum_{\mathbf{i} \in \mathbb{Z}^k} \left(\sum_{m < [Nt]} ([Nt] - m)^{\beta-1} g(m\mathbf{1} - \mathbf{i}) 1_{\{m\mathbf{1} > \mathbf{i}\}} \right)^2,$$

which, in view of (4.44), converges to 0 as $N \rightarrow \infty$. The proof is thus complete. \square

We now treat the case $\beta < 0$. This case is more delicate than the case $\beta > 0$ in two ways: a) an additional assumption on the linear-filter response $\{C_n\}$ has to be made; b) if β is chosen such that $H < 1/2$, then tightness of the normalized partial sum process needs also additional assumptions.

When $\beta < 0$, we have

$$\sum_{n=1}^{\infty} |C_n| < \infty.$$

If f_X is the spectral density of $\{X(n)\}$, then the spectral density of $\{U(n)\}$ is

$$f_U(\lambda) = |C(e^{i\lambda})|^2 f_X(\lambda),$$

where $C(z) := \sum_n C_n z^n$, and the transfer function $H(\lambda) := |C(e^{i\lambda})|^2$ is continuous. Since $X(n)$ is LRD (see Proposition 4.4.4), its spectral density blows up at the origin. To dampen it we need to multiply it by an $H(\lambda)$ which converges to 0 as $\lambda \rightarrow 0$. This means that $H(0) = |\sum_{n=1}^{\infty} C_n|^2 = 0$, and hence we need to assume $\sum_{n=1}^{\infty} C_n = 0$.

Theorem 4.6.11. *Let $U(n)$ be as in (4.42) with $\beta \in (-\alpha - k/2 - 1, 0)$, and assume in addition that*

$$\sum_{n=1}^{\infty} C_n = 0. \tag{4.46}$$

Then

$$\frac{1}{NH} \sum_{n=1}^{[Nt]} U(n) \xrightarrow{f.d.d.} Z^\beta(t),$$

where

$$0 < H = \alpha + \beta + k/2 + 1 < \alpha + k/2 + 1 < 1,$$

$Z^\beta(t)$ is the fractionally-filtered generalized Hermite process defined in Theorem 4.3.27. It is defined using the same g and β as $U(n)$.

If in addition, either a) $H > 1/2$, or b) $H < 1/2$ and for some $p > 1/H$, $\mathbb{E}|\epsilon_i|^p < \infty$, then the above $\xrightarrow{f.d.d.}$ can be replaced with weak convergence in $D[0, 1]$.

Proof. Note that by Lemma 4.6.8, we can change the order of summations to write:

$$\begin{aligned} Y_N(t) &:= \frac{1}{NH} \sum_{n=1}^{[Nt]} U(n) = \sum'_{\mathbf{i} \in \mathbb{Z}^k} \frac{1}{NH} \sum_{n=1}^{[Nt]} \sum_{m < n} C_{n-m} \sum'_{\mathbf{i} < m\mathbf{1}} a(m\mathbf{1} - \mathbf{i}) \prod_{j=1}^k \epsilon_{i_j} \\ &= \sum'_{\mathbf{i} \in \mathbb{Z}^k} \frac{1}{NH} \sum_{m \in \mathbb{Z}} a(m\mathbf{1} - \mathbf{i}) 1_{\{m\mathbf{1} > \mathbf{i}\}} \sum_{n=1 \vee (m+1)}^{[Nt]} C_{n-m} \prod_{j=1}^k \epsilon_{i_j} = Q_k(h_{t,N}^\beta), \end{aligned}$$

where

$$h_{t,N}^\beta(\mathbf{i}) = \frac{1}{NH} \sum_{m \in \mathbb{Z}} a(m\mathbf{1} - \mathbf{i}) 1_{\{m\mathbf{1} > \mathbf{i}\}} \sum_{n=1 \vee (m+1)}^{[Nt]} C_{n-m}.$$

Making use of (4.46), and using l to denote a generic function such that $l(i) \rightarrow 1$ as $i \rightarrow \infty$, we have if $m \geq 1$,

$$\sum_{n=1 \vee (m+1)}^{[Nt]} C_{n-m} = \sum_{n=1}^{[Nt]-m} C_n = - \sum_{n=[Nt]-m+1}^{\infty} C_n = \beta^{-1} l([Nt] - m + 1) ([Nt] - m + 1)_+^\beta;$$

and if $m \leq 0$,

$$\begin{aligned} \sum_{n=1 \vee (m+1)}^{[Nt]} C_{n-m} &= \sum_{n=1}^{[Nt]} C_{n-m} = \sum_{n=-m+1}^{[Nt]-m} C_n = \sum_{n=[Nt]-m+1}^{\infty} C_n - \sum_{n=-m+1}^{\infty} C_n \\ &= \beta^{-1} \left[l([Nt] - m + 1) ([Nt] - m + 1)_+^\beta - l(-m) (-m)_+^\beta \right]. \end{aligned}$$

So by letting \mathbf{i} correspond to $[N\mathbf{x}] + \mathbf{1}$ and m to $[Ns] + 1$ (omitting L and l for simplicity),

$$\begin{aligned} \tilde{h}_{t,N}^\beta(\mathbf{x}) &= N^{k/2} h_{t,N}^\beta([N\mathbf{x}] + \mathbf{1}) \\ &= \frac{1}{\beta} \int_{\mathbb{R}} g\left(\frac{[Ns]\mathbf{1} - [N\mathbf{x}]}{N}\right) 1_{\{[Ns]\mathbf{1} > [N\mathbf{x}]\}} \left(\left(\frac{[Nt] - [Ns]}{N}\right)_+^\beta - \left(\frac{-[Ns] - 1}{N}\right)_+^\beta \right) ds. \end{aligned}$$

Using similar arguments as in the proof of Theorem 4.6.5, we can bound the absolute value of the integrand above by $Cg^*(s\mathbf{1} - \mathbf{x})1_{\{s\mathbf{1} > \mathbf{x}\}} \left((t-s)_+^\beta - (-s)_+^\beta \right)$ for some $C > 0$, where g^* is a generalized Hermite kernel from Definition 4.3.18 (for the last term, we use $[Ns] + 1 \geq Ns$). Note that $\beta < 0$ in this case. By applying the Dominated Convergence Theorem, we get the desired f.d.d. convergence using Proposition 4.4.1.

Now we turn to the weak convergence. When $H > 1/2$, the tightness is standard. To show tightness under condition $H < 1/2$ and $E|\epsilon_i|^p < \infty$, Proposition 4.5.2 and the above f.d.d. convergence imply that for some constant $c, C > 0$ free from s, t and N ,

$$E|Y_N(t) - Y_N(s)|^{p'} \leq cE[|Y_N(t) - Y_N(s)|^2]^{p'/2} \leq C|F_N(t) - F_N(s)|^{p'H},$$

where $F_N(t) = [Nt]/N$, $p' < p$ and $p'H > 1$. Now by Lemma 4.4.1 and Theorem 4.4.1 of Giraitis et al. [2012], we conclude that tightness holds. \square

4.6.4 Mixed multivariate limit theorem

In Bai and Taqqu [2013a], a multivariate version of Theorem 4.2.3 is obtained, where both central and non-central convergence appear simultaneously. We will state here a similar theorem.

Suppose that $\mathbf{X}(n) = (X_1(n), \dots, X_J(n))$ is a vector of discrete chaos process defined on the same noise but with different coefficients, that is,

$$X_j(n) = \sum_{0 < i_1, \dots, i_{k_j} < \infty}^l a_j(i_1, \dots, i_{k_j}) \epsilon_{n-i_1} \dots \epsilon_{n-i_{k_j}} = \sum_{\mathbf{i} > \mathbf{0}}^l a_j(\mathbf{i}) \prod_{p=1}^{k_j} \epsilon_{n-i_p}, \quad (4.47)$$

where we assume $\{\epsilon_i\}$ is an i.i.d. random sequence with mean 0 and variance 1. For convenience we let $a_j(i_1, \dots, i_{k_j}) = a_j(\mathbf{i}) = a_j(\mathbf{i})1_{\{\mathbf{i} > \mathbf{0}\}}$, and $\tilde{a}_j(\cdot)$ denotes the symmetrization of $a_j(\cdot)$.

Definition 4.6.12. We say that the vector sequence of discrete chaos processes $\{\mathbf{X}(n)\}$ is

- SRD, if every component $X_j(n)$ is SRD in the sense of Definition 4.6.1, and in addition, for any $p \neq q \in \{1, \dots, J\}$,

$$\sum_{n=-\infty}^{\infty} \sum_{\mathbf{i} > \mathbf{0}} |\tilde{a}_p(\mathbf{i})\tilde{a}_q(n\mathbf{1} + \mathbf{i})| < \infty; \quad (4.48)$$

- LRD, if every component $X_j(n)$ is LRD in the sense of Definition 4.6.1.
- fLRD, if every component $X_j(n)$ is a fractionally-filtered LRD discrete chaos process in the sense of Definition 4.6.7. Note: these components were denoted $U(n)$ in that definition.

Remark 4.6.13. If the vector sequence is SRD, then (4.48) guarantees that the cross-covariance $\gamma_{p,q}(n) := \text{Cov}(X_p(n), X_q(0))$ satisfies $\sum_n |\gamma_{p,q}(n)| < \infty$. As in Proposition 2.5 of Bai and Taqqu [2013a], we have that as $N \rightarrow \infty$,

$$\text{Cov} \left(\frac{1}{\sqrt{N}} \sum_{n=1}^{\lfloor Nt_1 \rfloor} X_p(n), \frac{1}{\sqrt{N}} \sum_{n=1}^{\lfloor Nt_2 \rfloor} X_q(n) \right) \rightarrow (t_1 \wedge t_2) \sum_{n=-\infty}^{\infty} \gamma_{p,q}(n). \quad (4.49)$$

Note that $\gamma_{p,q}(n) = 0$ always if the orders $k_p \neq k_q$.

We will now consider a general case where SRD and LRD and fLRD vectors can all be present in $\mathbf{X}(n)$. We divide $\mathbf{X}(n)$ into four parts

$$\mathbf{X}(n) = (\mathbf{X}_{S_1}(n), \mathbf{X}_{S_2}(n), \mathbf{X}_L(n), \mathbf{X}_F(n))$$

of dimension $J_{S_1}, J_{S_2}, J_L, J_F$ respectively, which are defined as follows:

- (i) all the components of $\mathbf{X}_{S_1}(n) = (X_{1,S_1}(n), \dots, X_{J_{S_1},S_1}(n))$ have order $k = 1$, namely, are all linear processes;
- (ii) every component of $\mathbf{X}_{S_2}(n) = (X_{1,S_2}(n), \dots, X_{J_{S_2},S_2}(n))$ has order $k \geq 2$, and the *combined vector*

$$\mathbf{X}_S(n) = (\mathbf{X}_{S_1}(n), \mathbf{X}_{S_2}(n)) = (X_{1,S}(n), \dots, X_{J_S,S}(n)), \quad J_S = J_{S_1} + J_{S_2},$$

is SRD in the sense of Definition 4.6.12;

- (iii) the vector $\mathbf{X}_L(n) = (X_{1,L}(n), \dots, X_{J_L,L}(n))$ is LRD in the sense of Definition 4.6.12, with correspondingly generalized Hermite kernels $\mathbf{g} = (g_{1,L}, \dots, g_{J_L,L})$;
- (iv) the vector $\mathbf{X}_F(n) = (X_{1,F}(n), \dots, X_{J_F,F}(n))$ is fLRD in the sense of Definition 4.6.12, with correspondingly generalized Hermite kernels $\mathbf{g} = (g_{1,F}, \dots, g_{J_F,F})$ and fractional exponent $\boldsymbol{\beta} = (\beta_1, \dots, \beta_{J_F})$.

We now state the multivariate limit theorem. We use Y_N (with subscript S_1 , S_2 , L or F) to denote the corresponding normalized sum $Y_N(t) := N^{-H} \sum_{n=1}^{\lfloor Nt \rfloor} X(n)$, where $X(n)$ is a component of $\mathbf{X}(n)$, H is such that $\text{Var}(Y_N(1))$ converges to some constant $c > 0$ as $N \rightarrow \infty$.

Theorem 4.6.14. *Following the notation defined above, one has*

$$(\mathbf{Y}_{N,S_1}(t), \mathbf{Y}_{N,S_2}(t), \mathbf{Y}_{N,L}(t), \mathbf{Y}_{N,F}(t)) \xrightarrow{f.d.d.} (\mathbf{B}_1(t), \mathbf{B}_2(t), \mathbf{Z}(t), \mathbf{Z}^\beta(t)), \quad (4.50)$$

where

- (i) $\mathbf{B}_1(t) = \mathbf{W}(t) := (\sigma_1 W(t), \dots, \sigma_{J_{S_1}} W(t))$ defined by the same standard Brownian motion $W(t)$, and

$$\sigma_p = \sum_{n=-\infty}^{\infty} \sum_{i>0} a_{p,S_1}(n) a_{p,S_1}(n+i), \quad p = 1, \dots, J_{S_1}.$$

- (ii) $\mathbf{B}_2(t)$ is a multivariate Brownian motion with the covariance given by (4.49);
- (iii) $\mathbf{Z}(t)$ is a multivariate generalized Hermite process defined as in (4.5) by the kernels $(g_{1,L}, \dots, g_{J_L,L})$ and using the $W(t)$ in Point (i) as Brownian motion integrator.
- (iv) $\mathbf{Z}^\beta(t)$ is a multivariate fractionally-filtered generalized Hermite process defined as in (4.16) by the kernels $(g_{1,F}, \dots, g_{J_F,F})$, fractional exponent $\beta = (\beta_1, \dots, \beta_{J_F})$ and using the $W(t)$ in Point (i) as Brownian motion integrator.

Moreover, $\mathbf{B}_2(t)$ is always independent of $(\mathbf{B}_1(t), \mathbf{Z}(t), \mathbf{Z}^\beta(t))$.

In addition, $\xrightarrow{f.d.d.}$ in (4.50) can be replaced with weak convergence in $D[0, 1]^J$, if every component of \mathbf{X}_{S_1} and \mathbf{X}_{S_2} satisfies the assumption in Theorem 4.6.4, and every component of \mathbf{X}_F satisfies the assumption given at the end of Theorem 4.6.11.

The proof is similar to that of Theorem 3.5 of Bai and Taqqu [2013a]. We only provide some heuristics. The processes $\mathbf{B}_2(t), \mathbf{Z}(t)$ and $\mathbf{Z}^\beta(t)$ involve the same integrator $W(\cdot)$ because they are defined in terms of the same ϵ_i 's. To understand the independence statement, note that the independence between \mathbf{B}_2 and W stems from the uncorrelatedness between X_{S_2} and X_{S_1} , since X_{S_2} belongs to a discrete chaos of order $k \geq 2$, while X_{S_1} belongs to a discrete chaos of order $k = 1$. \mathbf{B}_2 is therefore independent of \mathbf{B}_1 . \mathbf{B}_2 is also independent of \mathbf{Z} and \mathbf{Z}^β , because \mathbf{Z} and \mathbf{Z}^β have W as integrators.

Remark 4.6.15. The pairwise dependence between components of \mathbf{Z} , of \mathbf{Z}^β , and between cross components in Theorem 4.6.14 can be checked using the criterion due to Ustunel and Zakai [1989], that is, if $f \in L^2(\mathbb{R}^p)$ and $g \in L^2(\mathbb{R}^q)$, and both are symmetric, then the multiple Wiener-Itô integrals $I_p(f)$ and $I_q(g)$ are independent, if and only if

$$f \otimes_1 g(x_1, \dots, x_{p+q-2}) := \int_{\mathbb{R}} f(x_1, \dots, x_{p-1}, y) g(x_p, \dots, x_{p+q-2}, y) dy = 0 \quad a.e..$$

For example, suppose that two generalized Hermite kernels g_1 and g_2 on \mathbb{R}_+^p and \mathbb{R}_+^q are symmetric, then the corresponding two generalized Hermite processes are independent if

and only if

$$\int_{\mathbb{R}} \int_0^t g_1(s - x_1, \dots, s - x_{p-1}, s - y) ds \int_0^t g_2(s - x_p, \dots, s - x_{p+q-2}, s - y) ds dy = 0 \quad a.e., \quad (4.51)$$

where we use the abbreviation $g_j(\mathbf{x}) = g_j(\mathbf{x})1_{\{\mathbf{x} > \mathbf{0}\}}$, $j = 1, 2$. Obviously, if g_1 and g_2 are both positive, then the dependence always holds. This is true, for example, for the symmetrized version of the kernels in (4.10).

Chapter 5

The universality of homogeneous polynomial forms and critical limits

Nourdin et al. [2010] established the following universality result: if a sequence of off-diagonal homogeneous polynomial forms in i.i.d. standard normal random variables converges in distribution to a normal, then the convergence also holds if one replaces these i.i.d. standard normal random variables in the polynomial forms by any independent standardized random variables with uniformly bounded third absolute moment. The result, which was stated for polynomial forms with a finite number of terms, can be extended to allow an infinite number of terms in the polynomial forms. Based on a contraction criterion derived from this extended universality result, we prove a central limit theorem for a strongly dependent nonlinear processes, whose memory parameter lies at the boundary between short and long memory.

5.1 Introduction

In Nourdin et al. [2010], a universality result was established for the following off-diagonal homogeneous polynomial form

$$Q_k(N_n, f_n, \mathbf{X}) := \sum_{1 \leq i_1, \dots, i_k \leq N_n} f_n(i_1, \dots, i_k) X_{i_1} \dots X_{i_k}, \quad (5.1)$$

where f_n is a sequence of symmetric functions on \mathbb{Z}_+^k *vanishing on the diagonals* ($f_n(i_1, \dots, i_k) = 0$ if $i_p = i_q$ for some $p \neq q$), and $\mathbf{X} = (X_1, X_2, \dots)$ is a sequence of standardized independent random variables, and N_n is a finite sequence such that $N_n \rightarrow \infty$ as $n \rightarrow \infty$.

The universality result says that if $\mathbf{Z} = (Z_1, Z_2, \dots)$ is an i.i.d. standard normal sequence and $Q_k(N_n, f_n, \mathbf{Z})$ converges weakly to a normal distribution as $n \rightarrow \infty$, then the same weak convergence to normal holds if \mathbf{Z} is replaced by \mathbf{X} , where \mathbf{X} is any standardized independent sequence with some uniform higher moment bound.

It is natural to try to eliminate the finiteness of N_n in the preceding result. This extension was mentioned in Remark 1.13 of Nourdin et al. [2010], but was not explicitly done. One would encounter a number of difficulties if one were to extend the method of proof used for finite N_n to $N_n = \infty$. We will note, however, that this extension can be easily achieved using a simple approximation argument. We find it valuable to have such an extension and the corresponding contraction criterion (Theorem 5.2.6) since it can be directly applied to limit theorems in the context of long memory.

We consider such an application in Section 5.3 where we suppose that

$$f_N(i_1, \dots, i_k) = \frac{1}{A(N)} \sum_{n=1}^N a(n - i_1, \dots, n - i_k) 1_{\{-\infty < i_1 < n, \dots, -\infty < i_k < n\}},$$

and where the function $a(\cdot)$ behaves essentially like a homogeneous function with exponent α . The resulting polynomial form $Q_k(f_N)$ is then the partial sum of a stationary process. The exponent α is chosen in such a way that the corresponding stationary process lives on the boundary between short and long memory. We use the contraction criterion to prove that a central limit theorem holds but with the nonstandard normalization $\sqrt{N \ln N}$. This delicate case seems difficult to treat otherwise.

The chapter is organized as follows. In Section 5.2, we state the and prove the extension of the universality result (Theorem 5.2.1), and as a byproduct, a criterion for asymptotic normality (Theorem 5.2.6). In Section 5.3.1, we state the critical limit theorem obtained

by applying the criterion. In Section 5.3.3 and 5.3.4 we give the proofs.

5.2 Universality of homogeneous polynomial forms

Let $\ell^2(\mathbb{Z}^k)$, $k \geq 1$, denote the space of symmetric square summable functions on \mathbb{Z}^k vanishing on the diagonals equipped with the discrete L^2 norm. Let $\mathbf{X} = (X_1, X_2, \dots)$ be a sequence of independent random variables satisfying $\mathbb{E}X_i = 0$ and $\mathbb{E}X_i^2 = 1$. By modifying the notation (5.1), one defines for $f \in \ell^2(\mathbb{Z}^k)$:

$$Q_k(f, \mathbf{X}) := \sum_{-\infty < i_1, \dots, i_k < \infty} f(i_1, \dots, i_k) X_{i_1} \dots X_{i_k}.$$

One has

$$\mathbb{E}Q_k(f, \mathbf{X}) = 0.$$

Consider now two homogeneous polynomial forms $Q_{k_1}(f_1, \mathbf{X})$ and $Q_{k_2}(f_2, \mathbf{X})$, where $f_1 \in \ell^2(\mathbb{Z}^{k_1})$ and $f_2 \in \ell^2(\mathbb{Z}^{k_2})$. Then the covariance of $Q_{k_1}(f_1, \mathbf{X})$ and $Q_{k_2}(f_2, \mathbf{X})$ is

$$\langle f_1, f_2 \rangle := \mathbb{E}Q_{k_1}(f_1, \mathbf{X})Q_{k_2}(f_2, \mathbf{X}) \tag{5.2}$$

$$= \begin{cases} k! \sum_{-\infty < i_1, \dots, i_k < \infty} f_1(i_1, \dots, i_k) f_2(i_1, \dots, i_k), & \text{if } k_1 = k_2 = k; \\ 0 & \text{if } k_1 \neq k_2. \end{cases} \tag{5.3}$$

We then have the following extension of Nourdin et al. [2010] Theorem 1.2:

Theorem 5.2.1. *For each $j = 1, \dots, m$, suppose that $k_j \geq 2$, and let $f_{n,j}(\cdot)$ be a sequence of functions in $\ell^2(\mathbb{Z}^{k_j})$. Let Σ be an $m \times m$ symmetric non-negative definite matrix whose each diagonal entry is positive. Assume in addition that*

$$\sup_n \sum_{-\infty < i_1, \dots, i_{k_j} < \infty} f_{n,j}(i_1, \dots, i_{k_j})^2 < \infty. \tag{5.4}$$

Then the following two statements are equivalent:

1. For every sequence $\mathbf{X} = (X_1, X_2, \dots)$ where X_1, X_2, \dots are independent random variables satisfying $\mathbb{E}X_i = 0, \mathbb{E}X_i^2 = 1$, and

$$\sup_i \mathbb{E}|X_i|^3 < \infty, \quad (5.5)$$

the following joint weak convergence to a multivariate normal distribution holds:

$$\left(Q_{k_j}(f_{n,j}, \mathbf{X}) \right)_{j=1}^m \xrightarrow{d} N(\mathbf{0}, \Sigma). \quad (5.6)$$

2. For a sequence $\mathbf{Z} = (Z_1, Z_2, \dots)$ of i.i.d. standard normal random variables, the following joint weak convergence to a multivariate normal distribution holds:

$$\left(Q_{k_j}(f_{n,j}, \mathbf{Z}) \right)_{j=1}^m \xrightarrow{d} N(\mathbf{0}, \Sigma). \quad (5.7)$$

Remark 5.2.2. Condition (5.4) can be re-expressed as

$$\sup_n \mathbb{E} Q_{k_j}(f_{n,j}, \mathbf{Z})^2 = k_j! \sup_n \sum_{-\infty < i_1, \dots, i_{k_j} < \infty} f_{n,j}(i_1, \dots, i_{k_j})^2 < \infty. \quad (5.8)$$

Remark 5.2.3. One can recover Nourdin et al. [2010] Theorem 1.2 from Theorem 5.2.1 by replacing $f_{n,j}(i_1, \dots, i_{k_j})$ with $f_{n,j}(i_1, \dots, i_{k_j}) \mathbf{1}_{1 \leq i_1, \dots, i_{k_j} \leq N_n}(i_1, \dots, i_{k_j})$.

Remark 5.2.4. In the one dimensional case: $m = 1$, one can relax the assumption (5.5) by $\sup_i \mathbb{E}|X_i|^{2+\delta} < \infty$ for any $\delta > 0$. See Theorem 1.10 of Nourdin et al. [2010].

Proof of Theorem 5.2.1. We need to prove that (5.7) implies (5.6). Define the N_n -truncated functions

$$\tilde{f}_{n,j}(i_1, \dots, i_{k_j}) = f_{n,j}(i_1, \dots, i_{k_j}) \mathbf{1}_{\{-N_n \leq i_1 \leq N_n, \dots, -N_n \leq i_{k_j} \leq N_n\}}, \quad j = 1, \dots, m.$$

For any $n \in \mathbb{Z}_+$, we can find N_n large enough, so that for all $j = 1, \dots, m$,

$$\begin{aligned} \mathbb{E} \left| Q_{k_j}(f_{n,j}, \mathbf{Z}) - Q_{k_j}(\tilde{f}_{n,j}, \mathbf{Z}) \right|^2 &= \mathbb{E} \left| Q_{k_j}(f_{n,j}, \mathbf{X}) - Q_{k_j}(\tilde{f}_{n,j}, \mathbf{X}) \right|^2 \\ &= k_j! \|\tilde{f}_{n,j} - f_{n,j}\|_{\ell^2(\mathbb{Z}^{k_j})}^2 \leq \frac{1}{n}. \end{aligned} \quad (5.9)$$

Assume without loss of generality that $N_n \rightarrow \infty$ as $n \rightarrow \infty$. By (5.7) and (5.9), one has

$$\left(Q_{k_j}(\tilde{f}_{n,j}, \mathbf{Z}) \right)_{j=1}^m \xrightarrow{d} N(\mathbf{0}, \Sigma).$$

Using the original version of the universality result in Nourdin et al. [2010] Theorem 1.2, one gets

$$\left(Q_{k_j}(\tilde{f}_{n,j}, \mathbf{X}) \right)_{j=1}^m \xrightarrow{d} N(\mathbf{0}, \Sigma). \quad (5.10)$$

The conclusion (5.6) follows from (5.9) and (5.10). \square

Remark 5.2.5. Using the same argument as in the preceding proof, one can eliminate the finiteness of N_n in (5.1) in the following related universality results for homogeneous polynomial forms: (a) Theorem 1.12 of Nourdin et al. [2010] concerning for convergence to a χ^2 distribution; (b) Theorem 3.4 of Peccati and Zheng [2014] which is the counterpart of Theorem 5.2.1 here with Z_i 's being standardized Poisson random variables.

Theorem 5.2.1 gives rise to a practical criterion for the convergence (5.6). We first introduce the discrete contraction operator: for $f \in \ell^p(\mathbb{Z}^p)$ and $g \in \ell^q(\mathbb{Z}^q)$, $p, q \geq 2$, we define

$$\begin{aligned} (f \star_r g)(i_1, \dots, i_{p+q-2r}) &= \\ &\sum_{j_1, \dots, j_r = -\infty}^{\infty} f(j_1, \dots, j_r, i_1, \dots, i_{p-r}) g(j_1, \dots, j_r, i_{p-r+1}, \dots, i_{p+q-2r}) \end{aligned} \quad (5.11)$$

for $r = 0, \dots, p \wedge q$, where in the case $r = 0$ it is understood as the tensor product.

Theorem 5.2.6. Let $\{f_{n,j}(\cdot), n \in \mathbb{Z}_+\}$ be a sequence of functions in $\ell^2(\mathbb{Z}^{k_j})$ satisfying

(5.4), $j = 1, \dots, m$, where $k_j \geq 2$. Let Σ be an $m \times m$ symmetric non-negative definite matrix whose each diagonal entry is positive, such that

$$\Sigma(i, j) = \lim_{n \rightarrow \infty} \langle f_{n,i}, f_{n,j} \rangle, \quad (5.12)$$

where $\langle \cdot, \cdot \rangle$ is defined in (5.3). Then the following are equivalent:

1. For every $\mathbf{X} = (X_1, X_2, \dots)$ with X_i 's being independent random variables satisfying $\mathbb{E}X_i = 0, \mathbb{E}X_i^2 = 1$ and $\sup_i \mathbb{E}|X_i|^3 < \infty$, we have the following joint weak convergence to normal:

$$\left(Q_{k_j}(f_{n,j}, \mathbf{X}) \right)_{j=1}^m \xrightarrow{d} N(\mathbf{0}, \Sigma). \quad (5.13)$$

2. The following contractions are vanishing:

$$\lim_{n \rightarrow \infty} \|f_{n,j} \star_r f_{n,j}\|_{2k_j - 2r} = 0, \quad \text{for all } r = 1, \dots, k_j - 1 \text{ and all } j = 1, \dots, m. \quad (5.14)$$

where $\|\cdot\|_k$ denotes the discrete L^2 norm on $\ell^2(\mathbb{Z}^k)$.

Proof. By Theorem 5.2.1, the statement 1 is equivalent to $\left(Q_{k_j}(f_{n,j}, \mathbf{Z}) \right)_{j=1}^m \xrightarrow{d} N(\mathbf{0}, \Sigma)$, where \mathbf{Z} is a sequence of i.i.d. standard Gaussian variables. Note also that each $Q_{k_j}(f_{n,j}, \mathbf{Z})$ can be expressed as a k_j -tuple Wiener-Itô integral with respect to Brownian motion. For Wiener-Itô integrals, joint convergence to the normal is equivalent to marginal convergence, and marginal convergence is equivalent to the contraction relations. More precisely, by applying Theorem 6.2.3 and 5.2.7 of Nourdin and Peccati [2012], one gets the equivalence to (5.14). See also Theorem 7.5 of Nourdin et al. [2010]. \square

Remark 5.2.7. We shall use the implication “Statement 2 \Rightarrow Statement 1” of the preceding theorem in the sequel. As for the reversed implication, namely, “Statement 1 \Rightarrow Statement 2”, the stipulation “For every” is important here, as well as in Theorem 5.2.1, because there are random variables X_i 's, for example Rademacher, that is $X_i = \pm 1$ with probability 1/2 each, for which one may have convergence in (5.13) even when (5.14) does

not hold (see Nourdin et al. [2010], Section 1.6, p.1956).

Remark 5.2.8. One may wonder if the universality result extends to a continuous setting, namely, when $Q_k(f_n)$ is replaced by a multiple integral on a Borel measure space (A, \mathcal{A}, μ) :

$$I_k(f_n, \xi) = \int_{A^k}' f_n(x_1, \dots, x_k) \xi(dx_1) \dots \xi(dx_k),$$

where $f \in L^2(A^k)$, the prime ' indicates the exclusion of diagonals $x_p = x_q$, $p \neq q$, and $\xi(\cdot)$ is an independently scattered random measure with an atomless control measure $\mu(\cdot)$. Does $I_k(f_n, \xi)$ exhibits a similar universality phenomenon? Namely, if $I_k(f_n, \xi)$ converges in distribution to normal for a Gaussian $\xi(\cdot)$, does the convergence also hold for general class of $\xi(\cdot)$ with the same control measure $\mu(\cdot)$? It is known that the law of $\xi(\cdot)$ has to be infinitely divisible and $\xi(\cdot)$ admits the decomposition:

$$\xi(B) = G(B) + \int_{\mathbb{R}} \int_A u 1_B(x) \widehat{N}(du, dx), \quad (5.15)$$

where $G(\cdot)$ is a Gaussian random measure on A and $\widehat{N}(\cdot)$ is an independent compensated Poisson random measure on $\mathbb{R} \times A$. See Section 5.3 of Peccati and Taqqu [2011] for more details.

One may think of adapting the approximation argument used in the proof of Theorem 5.2.1 to the multiple integral case, which would involve partitioning the space A into subsets of small measure. The problem is that unlike the Gaussian part, the Poisson part does not scale as $\mu(B) \rightarrow 0$. To see this in the simplest situation, take $\xi(B) = \widehat{P}(B)$, where $\widehat{P}(\cdot)$ is a compensated Poisson random measure on A with control measure $\mu(\cdot)$. Note that $\widehat{P}(B) + \mu(B)$ follows a Poisson distribution with mean $\mu(B)$. Since its cumulants are all equal to $\mu(B)$ (see (3.1.5) of Peccati and Taqqu [2011]), and since the third moment of a centered random variable is equal to the third cumulant, one has $E(\widehat{P}(B))^3 = \mu(B)$. This means that although we have the standardization

$$E \left| \widehat{P}(B) / \sqrt{\mu(B)} \right|^2 = 1, \quad (5.16)$$

we also have

$$\begin{aligned} \lim_{\mu(B) \rightarrow 0} \mathbb{E} \left| \widehat{P}(B) / \sqrt{\mu(B)} \right|^3 &= \lim_{\mu(B) \rightarrow 0} \mathbb{E} \left| \widehat{P}(B) \right|^3 \mu(B)^{-3/2} \geq \lim_{\mu(B) \rightarrow 0} \mathbb{E} \widehat{P}(B)^3 \mu(B)^{-3/2} \\ &= \lim_{\mu(B) \rightarrow 0} \mu(B)^{-1/2} = \infty. \end{aligned}$$

This will violate condition (5.5) as the partition of A becomes finer. In fact, one can show that $\widehat{P}(B) / \sqrt{\mu(B)} \rightarrow 0$ in probability as $\mu(B) \rightarrow 0$, which means, in view of (5.16), that the uniform integrability of $|\widehat{P}(B) / \sqrt{\mu(B)}|^2$ fails. For further insights, see Rotar [1979].

5.3 Application: boundary between short and long memory

5.3.1 The setting

Bai and Taqqu [2014a] considered the following *discrete chaos processes*:

$$X(n) = \sum_{-\infty < i_1, \dots, i_k < n} a(n - i_1, \dots, n - i_k) \epsilon_{i_1} \dots \epsilon_{i_k}, \quad (5.17)$$

where $k \geq 2$, $a(\cdot) : \mathbb{Z}_+^k \rightarrow \mathbb{R}$ is symmetric and vanishes on the diagonals, and ϵ_i 's are i.i.d. random variables with mean 0 and variance 1. Note that $\mathbb{E}X(n) = 0$.

In particular, Bai and Taqqu [2014a] studied limit theorems for normalized partial sum process of $X(n)$:

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

where $[\cdot]$ means integer part, and $A(N)$ is a suitable normalization factor. Depending on the behavior of $a(\cdot)$, the stationary process $X(n)$ may exhibit short or long memory.

As shown in Bai and Taqqu [2014a], in the short memory case, namely when the coefficient in (5.17) satisfies the summability condition

$$\sum_{n=1}^{\infty} \sum_{0 < i_1, \dots, i_k < \infty} \left| a(i_1, \dots, i_k) a(i_1 + n, \dots, i_k + n) \right| < \infty, \quad (5.18)$$

and $E|\epsilon_i|^{2+\delta} < \infty$ for some $\delta > 0$, the following central limit convergence as $N \rightarrow \infty$ holds:

$$\frac{1}{N^{1/2}} \sum_{n=1}^{[Nt]} X(n) \Rightarrow \sigma B(t) \quad (5.19)$$

for some $\sigma \geq 0$, where $B(t)$ is a standard Brownian motion.

In the long memory case, assume that

$$a(\cdot) = g(\cdot)L(\cdot)1_{D^c}, \quad (5.20)$$

where

$$D^c := \{(i_1, \dots, i_k) : i_p \neq i_q \text{ for } p \neq q\} \quad (5.21)$$

guarantees that $a(\cdot)$ vanishes on the diagonals. The function $L(\cdot) : \mathbb{Z}_+^k \rightarrow \mathbb{R}$ satisfies¹

$$\lim_{|\mathbf{i}| \rightarrow \infty} L(\mathbf{i}) = 1, \quad (5.22)$$

and $g(\cdot) : \mathbb{R}^k \rightarrow \mathbb{R}$ is the so-called *generalized Hermite kernel of Class (L)*.

Definition 5.3.1. A nonzero a.e. continuous function $g(\cdot) : \mathbb{R}^k \rightarrow \mathbb{R}$ is called a generalized Hermite kernel of Class (L) (GHK(L)) if it satisfies

1. $g(\cdot)$ is homogeneous with exponent α , namely, $g(\lambda \mathbf{x}) = \lambda^\alpha g(\mathbf{x})$, for all $\lambda > 0$, where

$$\alpha \in \left(-\frac{k+1}{2}, -\frac{k}{2} \right); \quad (5.23)$$

2. The function $g(\cdot)$ satisfies the bound

$$|g(\mathbf{x})| \leq g^*(\mathbf{x}) := c \sum_{j=1}^m x_1^{\gamma_{j1}} \dots x_k^{\gamma_{jk}}, \quad (5.24)$$

¹In Bai and Taqqu [2014a] eq. (25), $L(\cdot)$ is assumed to satisfy a slightly weaker condition than (5.22), that is, $\lim_{N \rightarrow \infty} L([N\mathbf{x}] + \mathbf{B}(N)) = 1$ for any $\mathbf{x} \in \mathbb{R}_+^k$ and any bounded sequence $\mathbf{B}(N)$ in \mathbb{Z}_+^k instead of $\lim_{\|\mathbf{x}\| \rightarrow \infty} L(\mathbf{x}) = 0$. Note that $L([N\mathbf{x}] + \mathbf{B}(N))$, $N \rightarrow \infty$, lets the argument increase in a specific band in the first quadrant, whereas $L(\mathbf{x})$, $\|\mathbf{x}\| \rightarrow \infty$, allows \mathbf{x} to increase in an arbitrary way in the first quadrant. Here for simplicity we just assume (5.22), while the results stated here also hold under the weaker condition.

with the constant $c > 0$, $-1 < \gamma_{jl} < -1/2$ and $\sum_{l=1}^k \gamma_{jl} = \alpha$ for all $l = 1, \dots, m$.

If g is a GHK(L), the following constant is well-defined (the integral is absolutely integrable)

$$C_g = \int_{\mathbb{R}_+^k} g(x_1, \dots, x_k) g(1 + x_1, \dots, 1 + x_k) dx_1 \dots dx_k, \quad (5.25)$$

and $C_g > 0$ always (Remark 3.6 of Bai and Taqqu [2014a]). Under this setup, Theorem 6.5 of Bai and Taqqu [2014a] showed that as $N \rightarrow \infty$,

$$\frac{1}{N^H} \sum_{n=1}^{[Nt]} X(n) \Rightarrow \int_{\mathbb{R}}' \int_0^t g(s_1 - x_1, \dots, s_k - x_k) 1_{\{s_1 > x_1, \dots, s_k > x_k\}} W(dx_1) \dots W(dx_k), \quad (5.26)$$

where $W(\cdot)$ is the Brownian random measure, the prime $'$ indicates the exclusion of the diagonals $x_p = x_q$, $p \neq q$, and

$$H = \alpha + \frac{k}{2} + 1.$$

The limit in (5.26) was called a *generalized Hermite process* which generalizes the Hermite process (see, e.g., Dobrushin and Major [1979] and Taqqu [1979]) which corresponds to the special case $g(\mathbf{x}) = x_1^{\alpha/k} \dots x_k^{\alpha/k}$.

There is, however, a boundary case which the limit theorems (5.19) and (5.26) did not cover. This boundary case is as follows: set as in the long memory case

$$a(\cdot) = g(\cdot) L(\cdot) 1_{D^c}, \quad (5.27)$$

where D^c is as in (5.21), $L(\cdot)$ is as in (5.22), and g is a function satisfying the assumptions in Definition 5.3.1 except that instead of assuming (5.23), the homogeneity exponent is set as $\alpha = -\frac{k+1}{2}$.

Remark 5.3.2. Note that if $\alpha < -\frac{k+1}{2}$, we are in the short memory regime. Indeed Proposition 5.4 of Bai and Taqqu [2015b] showed that $\alpha < -\frac{k+1}{2}$ implies (5.18), and thus (5.19) holds. So (5.28) is exactly the boundary case between short and long memory.

5.3.2 Statement of the limit theorems

Let throughout \Rightarrow denote weak convergence in Skorohod space $D[0, 1]$ with uniform metric. We shall show by the criterion formulated in Theorem 5.2.6, that a central limit theorem holds with an extra logarithmic factor in the normalization:

Theorem 5.3.3 (Nonlinear case). *Let*

$$X(n) = \sum_{-\infty < i_1, \dots, i_k < n} a(n - i_1, \dots, n - i_k) \epsilon_{i_1} \dots \epsilon_{i_k}$$

as in (5.17) with $k \geq 2$ and the coefficient $a(\cdot)$ specified as in (5.27) where

$$\alpha = -\frac{k+1}{2}. \quad (5.28)$$

Assume also that $E|\epsilon_i|^3 < \infty$ and $C_g > 0$. Then

$$Y_N(t) := \frac{1}{\sqrt{N \ln N}} \sum_{n=1}^{[Nt]} X(n) \Rightarrow \sigma B(t)$$

where $\sigma = \sqrt{2C_g}$, and $B(t)$ is a standard Brownian motion.

Remark 5.3.4. Theorem 5.3.3 may be compared to a similar boundary case of limit theorems for nonlinear transform of long-memory Gaussian noise first considered in Breuer and Major [1983] Theorem 1'. The proof there was done by a method of moments. See also Breton and Nourdin [2008] who gave an alternative proof using the Malliavin calculus.

Note that to apply Theorem 5.2.6, the process $X(n)$ in (5.17) needs to have order $k \geq 2$. For completeness, we state also the corresponding result for linear process, namely, the case $k = 1$ in Theorem 5.3.3, though the limit theorem for linear process is classical (see, e.g., Davydov [1970]).

Theorem 5.3.5 (Linear case). *Let*

$$X(n) = \sum_{-\infty < i < n} a(n-i)\epsilon_i,$$

where $a(n) = L(n)n^{-1}$ as $n \rightarrow \infty$, and let $L(n) \rightarrow c \neq 0$, and the i.i.d. standardized noise ϵ_i 's satisfy $E|\epsilon_i|^{2+\delta} < \infty$ for some $\delta > 0$. Then as $N \rightarrow \infty$,

$$Y_N(t) := \frac{1}{\sqrt{N} \ln N} \sum_{n=1}^{\lfloor Nt \rfloor} X(n) \Rightarrow \sigma B(t)$$

where $\sigma = \sqrt{2|c|}$, and $B(t)$ is a standard Brownian motion.

5.3.3 Proof of Theorem 5.3.3

We first compute the asymptotic variance of the sum.

Lemma 5.3.6. *Let $X(n)$ be given as in (5.17) with the coefficient specified as in (5.27) and α as in (5.28). Then C_g defined in (5.25) is non-negative. If $C_g > 0$, then as $N \rightarrow \infty$*

$$E \left[\sum_{n=1}^N X(n) \right]^2 \sim 2C_g N \ln N.$$

If $C_g = 0$, then

$$E \left[\sum_{n=1}^N X(n) \right]^2 = o(N \ln N). \quad (5.29)$$

Proof. Assume for simplicity $L(\cdot) = 1$, and it is easy to extend the following arguments to the general case. First, since $g(\cdot)$ is homogeneous with exponent $\alpha = -k/2 - 1/2$ by (5.28), one can write

$$\begin{aligned} \gamma(n) &:= EX(n)X(0) = \sum_{0 < i_1, \dots, i_k < \infty} g(i_1, \dots, i_k) g(i_1 + n, \dots, i_k + n) 1_{D^c}(i_1, \dots, i_k) \\ &= n^{-1} \sum_{0 < i_1, \dots, i_k < \infty} g\left(\frac{i_1}{n}, \dots, \frac{i_k}{n}\right) g\left(\frac{i_1}{n} + 1, \dots, \frac{i_k}{n} + 1\right) 1_{D^c}(i_1, \dots, i_k) n^{-k} \end{aligned}$$

$$\begin{aligned}
&= n^{-1} \int_{\mathbb{R}_+^k} g\left(\frac{[nx_1]+1}{n}, \dots, \frac{[nx_k]+1}{n}\right) g\left(\frac{[nx_1]+1}{n}+1, \dots, \frac{[nx_k]+1}{n}+1\right) \times \\
&\quad 1_{D^c}([nx_1], \dots, [nx_k]) dx_1 \dots dx_k \\
&=: n^{-1} C_n(g).
\end{aligned}$$

Because the bounding function g^* in Definition 5.3.1 is decreasing in every variable, the absolute of the integrand above is bounded by

$$g^*(x_1, \dots, x_k) g^*(x_1+1, \dots, x_k+1) = c^2 \sum_{j_1, j_2=1}^m x_1^{\gamma_{j_1,1}} (x_1+1)^{\gamma_{j_2,1}} \dots x_k^{\gamma_{j_1,k}} (x_k+1)^{\gamma_{j_2,k}}$$

which is integrable on \mathbb{R}_+^k because all $\gamma_{p,q} \in (-1, -1/2)$ and

$$\int_{\mathbb{R}_+} x^\gamma (x+1)^{\gamma'} dx < \infty \quad \text{for any } -1 < \gamma, \gamma' < -1/2.$$

Since g is assumed to be a.e. continuous, by the Dominated Convergence Theorem, as $n \rightarrow \infty$ we have

$$C_n(g) \rightarrow C_g := \int_{\mathbb{R}^k} g(x_1, \dots, x_k) g(x_1+1, \dots, x_k+1) dx_1 \dots dx_k.$$

Hence when $C_n \neq 0$, one has when $n > 0$

$$\gamma(n) \sim n^{-1} C_g,$$

and when $C_n = 0$, one has

$$\gamma(n) = o(n^{-1}).$$

We shall use the fact that if $a_n \sim n^{-1}$ as $n \rightarrow \infty$, then $\sum_{n=1}^N a_n \sim \ln N$ as $N \rightarrow \infty$. So when $C_g \neq 0$, one has

$$\mathbb{E} \left[\sum_{n=1}^N X(n) \right] = \sum_{n_1, n_2=1}^N \gamma(n_1 - n_2) = N \sum_{n=-N+1}^{N-1} \gamma(n) - \sum_{n=-N+1}^{N-1} |n| \gamma(n) \sim 2C_g N \ln N.$$

Note that since $\gamma(n) \sim n^{-1}C_g$, the term $\sum_{n=-N+1}^{N-1} |n|\gamma(n) \sim 2C_g N$ and is thus negligible.

The preceding asymptotic equivalence also shows that if $C_g \neq 0$ then $C_g > 0$ because the variance is non-negative.

If $C_g = 0$, following similar lines of argument, one gets (5.29).

□

Lemma 5.3.7. *Define the mapping $(\cdot, \cdot)_0 : \mathbb{R}^2 \rightarrow \mathbb{R}$ as*

$$(x_1, x_2)_0 = \begin{cases} |x_1 - x_2| & \text{if } x_1 \neq x_2; \\ 1 & \text{if } x_1 = x_2 = x. \end{cases}$$

For $-1 < \gamma_1, \gamma_2 < -1/2$ and $n_1, n_2 \in \{1, 2, \dots\}$, we have for some constant $C > 0$ not depending on n_1, n_2 that

$$\sum_{p \in \mathbb{Z}} (n_1 - p)_+^{\gamma_1} (n_2 - p)_+^{\gamma_2} \leq C(n_1, n_2)_0^{\gamma_1 + \gamma_2 + 1}.$$

Proof. For the case $n_1 = n_2 = n$, choose $C = \sum_{p < n} (n - p)^{\gamma_1 + \gamma_2} < \infty$ since $\gamma_1 + \gamma_2 < -1$.

When $n_1 \neq n_2$, suppose that $n_1 < n_2$. Then

$$\begin{aligned} \sum_{p \in \mathbb{Z}} (n_1 - p)_+^{\gamma_1} (n_2 - p)_+^{\gamma_2} &= \sum_{p=1}^{\infty} p^{\gamma_1} (n_2 - n_1 + p)_2^{\gamma_2} \leq \int_0^{\infty} x^{\gamma_1} (n_2 - n_1 + x)^{\gamma_2} dx \\ &= (n_2 - n_1)^{\gamma_1 + \gamma_2 + 1} \int_0^{\infty} y^{\gamma_1} (1 + y)^{\gamma_2} dy, \end{aligned}$$

where the integral converges.

□

The following simple fact will be used.

Lemma 5.3.8. *Suppose that $\gamma_j < -1/2$ for all $j = 1, \dots, k$, $k \geq 2$, and $\gamma_1 + \dots + \gamma_k \geq -k/2 - 1/2$. Then*

$$-\frac{r}{2} - \frac{1}{2} < \gamma_1 + \dots + \gamma_r < -\frac{r}{2} \quad \text{for all } r = 1, \dots, k-1.$$

In addition, each $\gamma_j > -1$, $j = 1, \dots, k$.

Proof. The inequality $\gamma_1 + \dots + \gamma_r < -\frac{r}{2}$ is obvious. For the other inequality, suppose that $\gamma_1 + \dots + \gamma_r \leq -r/2 - 1/2$ for some $r \in \{1, \dots, k\}$. Because $\gamma_{r+1}, \dots, \gamma_k < -1/2$, we get the contradiction: $\gamma_1 + \dots + \gamma_k < -r/2 - 1/2 - (k-r)/2 = -k/2 - 1/2$.

Then we show by contradiction that each $\gamma_j > -1$. Suppose, e.g., $\gamma_k \leq -1$. By what was just proved, one has $\gamma_1 + \dots + \gamma_{k-1} < -(k-1)/2$. Thus by adding $\gamma_k \leq -1$, one gets $\gamma_1 + \dots + \gamma_k < -k/2 - 1/2$, which contradicts the assumption. \square

We need the following lemma, which is a consequence of Corollary 1.1 (b) of Terrin and Taqqu [1991b].

Lemma 5.3.9. *If $\alpha_1, \dots, \alpha_m$, $m \geq 2$, satisfy*

$$\alpha_1, \dots, \alpha_m > -1, \quad \sum_{i=1}^m \alpha_i + m > 1, \quad (5.30)$$

then for any $c > 0$

$$\int_{[0,c]^m} |x_1 - x_2|^{\alpha_1} |x_2 - x_3|^{\alpha_2} \dots |x_{m-1} - x_m|^{\alpha_{m-1}} |x_m - x_1|^{\alpha_m} dx_1 \dots dx_m < \infty.$$

We need also the following hypercontractivity inequality for proving tightness in $D[0, 1]$ (Proposition 5.2 of Bai and Taqqu [2014a])

Lemma 5.3.10. *Suppose that $h \in \ell^2(\mathbb{Z}^k)$ vanishing on the diagonals. Let*

$$X = \sum_{\mathbf{i} \in \mathbb{Z}^k} h(\mathbf{i}) \prod_{p=1}^k \epsilon_{i_p}, \quad k \geq 1.$$

If for some $p' > p > 2$, $E|\epsilon_i|^{p'} < \infty$, then one has for some constant $c_{p,k} > 0$ which does not depend on h that

$$E[|X|^{p'}]^{1/p'} \leq c_{p,k} E[|X|^2]^{1/2}.$$

Proof of Theorem 5.3.3. Let $C > 0$ be a constant whose value can change from line to line.

We first show that the finite-dimensional distributions of $Y_N(t)$ converges to those of $\sigma B(t)$ using Theorem 5.2.6. First, the convergence of the covariance structure of $Y_N(t)$ to that of $\sigma B(t)$ follows from Lemma 5.3.6 the fact that for $s \leq t$ we have

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

as $N \rightarrow \infty$, since $X(n)$ is stationary. We now check the contraction conditions (5.14). For simplicity we set $L(\cdot) = 1$ and $t = 1$. We can write

$$Y_N(1) = \sum_{-\infty < i_1, \dots, i_k < +\infty} f_N(i_1, \dots, i_k) \epsilon_{i_1} \dots \epsilon_{i_k}$$

where

$$f_N(i_1, \dots, i_k) = \frac{1}{\sqrt{N \ln N}} \sum_{n=1}^N g(n - i_1, \dots, n - i_k) \mathbf{1}_{D^c \cap \{i_1 < n, \dots, i_k < n\}}. \quad (5.31)$$

To simplify notation, we set

$$\mathbf{p} = (p_1, \dots, p_r), \quad \mathbf{q} = (q_1, \dots, q_{k-r}),$$

$$\mathbf{i}_1 = (i_1, \dots, i_{k-r}), \quad \mathbf{i}_2 = (i_{k-r+1}, \dots, i_{2k-2r}), \quad \mathbf{i} = (\mathbf{i}_1, \mathbf{i}_2),$$

and let $\mathbf{1}$ stand for a vector of 1's of suitable dimension. We also use the convention that $\mathbf{x}^{\mathbf{a}} = x_1^{a_1} \dots x_n^{a_n}$ if $\mathbf{x} = (x_1, \dots, x_n)$ and $\mathbf{a} = (a_1, \dots, a_n)$. Let $(\Sigma \mathbf{x}) = x_1 + \dots + x_n$ if $\mathbf{x} = (x_1, \dots, x_n)$.

Set $g^*(\cdot)$ be as in Definition 5.3.1 which we write by splitting $\mathbf{x} = (\mathbf{x}_1, \mathbf{x}_2)$, where $\mathbf{x}_1 \in \mathbb{R}_+^r$ and $\mathbf{x}_2 \in \mathbb{R}_+^{k-r}$:

$$g^*(\mathbf{x}_1, \mathbf{x}_2) = c \sum_{j=1}^m \mathbf{x}_1^{\beta_j} \mathbf{x}_2^{\boldsymbol{\eta}_j}, \quad \beta_j = (\gamma_{j1}, \dots, \gamma_{jr}), \quad \boldsymbol{\eta}_j = (\gamma_{j,r+1}, \dots, \gamma_{jk}), \quad (5.32)$$

so that

$$\sum_{i=1}^r \beta_{ji} + \sum_{i=1}^{k-r} \eta_{ji} = \sum_{i=1}^k \gamma_{ji} = \alpha, \quad (5.33)$$

which we write simply as $\sum \boldsymbol{\beta} + \sum \boldsymbol{\eta} = \sum \boldsymbol{\gamma} = \alpha$. For convenience, if some component x_j of \mathbf{x} is negative, we set $\mathbf{x}^a = 0$ and hence $g^*(\mathbf{x}) = 0$. Then in view of (5.31), (5.11) and (5.24),

$$\begin{aligned} |(f_N \star_r f_N)(\mathbf{i})| &\leq \frac{1}{N \ln N} \sum_{n_1, n_2=1}^N \sum_{\mathbf{p}} g^*(n_1 \mathbf{1} - \mathbf{p}, n_1 \mathbf{1} - \mathbf{i}_1) g^*(n_2 \mathbf{1} - \mathbf{p}, n_2 \mathbf{1} - \mathbf{i}_2) \\ &= \frac{c^2}{N \ln N} \sum_{n_1, n_2=1}^N \sum_{j_1, j_2=1}^m (n_1 \mathbf{1} - \mathbf{i}_1)^{\boldsymbol{\eta}_{j_1}} (n_2 \mathbf{1} - \mathbf{i}_2)^{\boldsymbol{\eta}_{j_2}} \sum_{\mathbf{p}} (n_1 \mathbf{1} - \mathbf{p})^{\boldsymbol{\beta}_{j_1}} (n_2 \mathbf{1} - \mathbf{p})^{\boldsymbol{\beta}_{j_2}}, \end{aligned}$$

by using (5.32). By Lemma 5.3.7, we have for the last sum,

$$\begin{aligned} \sum_{\mathbf{p}} (n_1 \mathbf{1} - \mathbf{p})^{\boldsymbol{\beta}_{j_1}} (n_2 \mathbf{1} - \mathbf{p})^{\boldsymbol{\beta}_{j_2}} &= \\ \sum_{p_1, \dots, p_r} \prod_{u=1}^r (n_1 - p_u)^{\gamma_{j_1, u}} \prod_{v=1}^r (n_2 - p_v)^{\gamma_{j_2, v}} &\leq C (n_1, n_2)_0^{(\Sigma \boldsymbol{\beta}_{j_1}) + (\Sigma \boldsymbol{\beta}_{j_2}) + r}. \end{aligned}$$

Hence

$$\begin{aligned} \|f_N \star_r f_N\|_{2k-2r}^2 &= \sum_{\mathbf{i}} [(f_N \star_r f_N)(\mathbf{i})]^2 \\ &\leq \frac{C}{N^2 (\ln N)^2} \sum_{\mathbf{i}} \left(\sum_{n_1, n_2=1}^N \sum_{j_1, j_2=1}^m (n_1, n_2)_0^{(\Sigma \boldsymbol{\beta}_{j_1}) + (\Sigma \boldsymbol{\beta}_{j_2}) + r} (n_1 \mathbf{1} - \mathbf{i}_1)^{\boldsymbol{\eta}_{j_1}} (n_2 \mathbf{1} - \mathbf{i}_2)^{\boldsymbol{\eta}_{j_2}} \right)^2 \\ &= \frac{C}{N^2 (\ln N)^2} \sum_{j_1, j_2, j_3, j_4=1}^m \sum_{n_1, n_2, n_3, n_4=1}^N (n_1, n_2)_0^{(\Sigma \boldsymbol{\beta}_{j_1}) + (\Sigma \boldsymbol{\beta}_{j_2}) + r} (n_3, n_4)_0^{(\Sigma \boldsymbol{\beta}_{j_3}) + (\Sigma \boldsymbol{\beta}_{j_4}) + r} \\ &\quad \times \sum_{\mathbf{i}_1} (n_1 \mathbf{1} - \mathbf{i}_1)^{\boldsymbol{\eta}_{j_1}} (n_3 \mathbf{1} - \mathbf{i}_1)^{\boldsymbol{\eta}_{j_3}} \sum_{\mathbf{i}_2} (n_2 \mathbf{1} - \mathbf{i}_2)^{\boldsymbol{\eta}_{j_2}} (n_4 \mathbf{1} - \mathbf{i}_2)^{\boldsymbol{\eta}_{j_4}} \\ &\leq \frac{C}{N^2 (\ln N)^2} \sum_{j_1, j_2, j_3, j_4=1}^m \sum_{n_1, n_2, n_3, n_4=1}^N (n_1, n_2)_0^{(\Sigma \boldsymbol{\beta}_{j_1}) + (\Sigma \boldsymbol{\beta}_{j_2}) + r} (n_3, n_4)_0^{(\Sigma \boldsymbol{\beta}_{j_3}) + (\Sigma \boldsymbol{\beta}_{j_4}) + r} \\ &\quad \times (n_1, n_3)_0^{(\Sigma \boldsymbol{\eta}_{j_1}) + (\Sigma \boldsymbol{\eta}_{j_3}) + k - r} (n_2, n_4)_0^{(\Sigma \boldsymbol{\eta}_{j_2}) + (\Sigma \boldsymbol{\eta}_{j_4}) + k - r} \end{aligned} \quad (5.34)$$

where we have applied again Lemma 5.3.7 to get the last inequality. Note that if one adds up the power exponents in the last expression, one gets

$$(\Sigma\boldsymbol{\beta}_{j_1}) + (\Sigma\boldsymbol{\eta}_{j_1}) + (\Sigma\boldsymbol{\beta}_{j_2}) + (\Sigma\boldsymbol{\eta}_{j_2}) + (\Sigma\boldsymbol{\beta}_{j_3}) + (\Sigma\boldsymbol{\eta}_{j_3}) + (\Sigma\boldsymbol{\beta}_{j_4}) + (\Sigma\boldsymbol{\eta}_{j_4}) + 2k = 4\alpha + 2k = -2, \quad (5.35)$$

by (5.33), where the last equality of (5.35) is due to assumption (5.28).

Note also that by Lemma 5.3.8, we have for $r \in \{1, \dots, k-1\}$ that

$$-\frac{r}{2} - \frac{1}{2} < (\Sigma\boldsymbol{\beta}_{j_1}), (\Sigma\boldsymbol{\beta}_{j_2}), (\Sigma\boldsymbol{\beta}_{j_3}), (\Sigma\boldsymbol{\beta}_{j_4}) < -\frac{r}{2},$$

and

$$-\frac{k-r}{2} - \frac{1}{2} < (\Sigma\boldsymbol{\eta}_{j_1}), (\Sigma\boldsymbol{\eta}_{j_3}), (\Sigma\boldsymbol{\eta}_{j_2}), (\Sigma\boldsymbol{\eta}_{j_4}) < -\frac{k-r}{2}.$$

Let $\alpha_1 = (\Sigma\beta_{j_1}) + (\Sigma\beta_{j_2}) + r$ be the exponent of $(n_1, n_2)_0$ in (5.34). Then

$$-1 = -r/2 - 1/2 - r/2 - 1/2 + r < \alpha_1 < -r/2 - r/2 + r = -r + r = 0.$$

Define similarly $\alpha_2, \alpha_3, \alpha_4$ for the other exponents in (5.34), which all lie strictly between -1 and 0 . Hence, the convergence

$$\lim_{N \rightarrow \infty} \|f_N \star_r f_N\|_{2^{k-2r}}^2 = 0, \quad r = 1, \dots, k-1, \quad (5.36)$$

will follow if one shows that

$$\sup_N N^{-2} \sum_{n_1, n_2, n_3, n_4=1}^N (n_1, n_2)_0^{\alpha_1} (n_2, n_3)_0^{\alpha_2} (n_3, n_4)_0^{\alpha_3} (n_4, n_1)_0^{\alpha_4} < \infty, \quad (5.37)$$

where by (5.35)

$$-1 < \alpha_j < 0, \quad j = 1, \dots, 4, \quad \alpha_1 + \alpha_2 + \alpha_3 + \alpha_4 = -2. \quad (5.38)$$

Let's consider first the sum in (5.37) over only distinct n_1, \dots, n_4 (we use the prime ' to indicate that the sum does not include the diagonals). In this case,

$$\begin{aligned} & \sum'_{1 \leq n_1, n_2, n_3, n_4 \leq N} \left| \frac{n_1}{N} - \frac{n_2}{N} \right|^{\alpha_1} \left| \frac{n_2}{N} - \frac{n_3}{N} \right|^{\alpha_2} \left| \frac{n_3}{N} - \frac{n_4}{N} \right|^{\alpha_3} \left| \frac{n_4}{N} - \frac{n_1}{N} \right|^{\alpha_4} N^{-4} \\ &= \int \left| \frac{[Nx_1] - [Nx_2]}{N} \right|^{\alpha_1} \left| \frac{[Nx_2] - [Nx_3]}{N} \right|^{\alpha_2} \left| \frac{[Nx_3] - [Nx_4]}{N} \right|^{\alpha_3} \left| \frac{[Nx_4] - [Nx_1]}{N} \right|^{\alpha_4} \times \\ & \quad I\{N^{-1} \leq x_i \leq 1 + N^{-1}, [Nx_i] \neq [Nx_j], \forall i \neq j\} dx. \end{aligned}$$

Note that for any $x, y > 0$, one has that $|[Nx] - [Ny]| = n$ implies that $|Nx - Ny| \leq n + 1$ which implies $|x - y| \leq (n + 1)/N$, for $n \geq 0$. Then since each $\alpha < 0$, we get

$$\begin{aligned} & \sup_N \left| \frac{[Nx] - [Ny]}{N} \right|^\alpha |x - y|^{-\alpha} I\{[Nx] \neq [Ny]\} \\ & \leq \sup_{|[Nx] - [Ny]| = n, n \in \mathbb{Z}_+} \left(\frac{n}{N} \right)^\alpha \left(\frac{n + 1}{N} \right)^{-\alpha} = \sup_{n \in \mathbb{Z}_+} \left(\frac{n + 1}{n} \right)^{-\alpha} = 2^{-\alpha}. \end{aligned}$$

Hence the the sum in (5.37) over distinct n_1, \dots, n_4 is bounded by

$$C \int_{[0,2]^4} |x_1 - x_2|^{\alpha_1} |x_2 - x_3|^{\alpha_2} |x_3 - x_4|^{\alpha_3} |x_4 - x_1|^{\alpha_4} dx_1 dx_2 dx_3 dx_4,$$

which is finite due to Lemma 5.3.9.

Consider now the the sum in (5.37) over n_1, \dots, n_4 with only three of them distinct. Let, for example, $n_1 = n_4$, and we need to show that the following

$$\begin{aligned} & \sup_N N^{-2} \sum'_{1 \leq n_1, n_2, n_3 \leq N} |n_1 - n_2|^{\alpha_1} |n_2 - n_3|^{\alpha_2} |n_3 - n_1|^{\alpha_3} = \\ & \sup_N N^{1+\alpha_1+\alpha_2+\alpha_3} \sum'_{1 \leq n_1, n_2, n_3 \leq N} \left| \frac{n_1}{N} - \frac{n_2}{N} \right|^{\alpha_1} \left| \frac{n_2}{N} - \frac{n_3}{N} \right|^{\alpha_2} \left| \frac{n_3}{N} - \frac{n_1}{N} \right|^{\alpha_3} N^{-3} < \infty. \end{aligned}$$

Note that (5.38) entails that $-2 < \alpha_1 + \alpha_2 + \alpha_3 < -1$. Then $N^{1+\alpha_1+\alpha_2+\alpha_3} \rightarrow 0$ as $N \rightarrow \infty$, and the boundedness of the multiple sum can be established similarly as above using integral approximation and Lemma 5.3.9.

If the sum in (5.37) is over n_1, \dots, n_4 with only two or less of them distinct, the boundedness is easily established through bounding all the summands by one constant, because we have the factor N^{-2} .

So (5.37) holds and thus (5.36) holds, and the convergence of finite-dimensional distributions is proved.

Now we show tightness. By Lemma 5.3.10, one can choose $p \in (2, 3)$, so that by Lemma 5.3.6 if $0 < s < t < 1$, one has for N large enough,

$$\begin{aligned} \mathbb{E}|Y_N(t) - Y_N(s)|^p &\leq C[\mathbb{E}|Y_N(t) - Y_N(s)|^2]^{p/2} \leq C \left[\frac{[Nt] - [Ns]}{N} \cdot \frac{\ln([Nt] - [Ns])}{\ln N} \right]^{p/2} \\ &\leq C \left[\frac{[Nt] - [Ns]}{N} \right]^{p/2 - \delta}, \end{aligned}$$

where $\delta > 0$ is small enough so that $p/2 - \delta > 1$. The last inequality is true because $\ln x$ is slowly varying as $x \rightarrow \infty$ and so one applies the Potter's bound (see e.g., equation (2.3.6) of Giraitis et al. [2012]). Note that $F_N(t) := [Nt]/N$ is a non-decreasing right continuous function on $[0, 1]$ and that F_N converges uniformly to $F(t) := t$ as $N \rightarrow \infty$. Hence by Lemma 4.4.1 and Theorem 4.4.1 of Giraitis et al. [2012], the tightness in $D[0, 1]$ is proved. \square

5.3.4 Proof of Theorem 5.3.5

Proof. Set for simplicity $L(n) = c$. The covariance $\gamma(n)$ for $n > 0$ is

$$\gamma(n) = \mathbb{E}X(n)X(0) = \sum_{i=1}^{\infty} a_{i+n}a_i = c^2 \sum_{i=1}^{\infty} (i+n)^{-1}i^{-1}.$$

Note that as $n \rightarrow \infty$,

$$\begin{aligned} \sum_{i=2}^{\infty} (i+n)^{-1}i^{-1} &= n^{-1} \sum_{i=2}^{\infty} \left(\frac{i}{n} + 1 \right)^{-1} \left(\frac{i}{n} \right)^{-1} \frac{1}{n} \\ &= n^{-1} \int_{2/n}^{\infty} \left(\frac{[nx]}{n} + 1 \right)^{-1} \left(\frac{[nx]}{n} \right)^{-1} dx \sim n^{-1} \ln n. \end{aligned}$$

The last asymptotic can be seen from:

$$\int_{2/n}^{\infty} (x+1)^{-1} x^{-1} dx \leq \int_{2/n}^{\infty} \left(\frac{[nx]}{n} + 1\right)^{-1} \left(\frac{[nx]}{n}\right)^{-1} dx \leq \int_{1/n}^{\infty} (y+1)^{-1} y^{-1} dy,$$

where we have used the fact $x - 1/n \leq [nx]/n \leq x$, and both the lower and upper bounds are asymptotically equivalent to $\ln n$ as $n \rightarrow \infty$.

Hence

$$\gamma(n) \sim c^2 n^{-1} \ln n \quad \text{as } n \rightarrow \infty. \quad (5.39)$$

So as $N \rightarrow \infty$, one has

$$\begin{aligned} \mathbb{E} \left(\sum_{n=1}^N X(n) \right)^2 &= N \sum_{n=-N+1}^{N-1} \gamma(n) - \sum_{n=-N+1}^{N-1} |n| \gamma(n) \\ &\sim 2c^2 N \sum_{n=1}^N n^{-1} \ln n \sim 2c^2 N \int_1^N x^{-1} \ln x dx \sim 2c^2 N (\ln N)^2. \end{aligned} \quad (5.40)$$

Note that by (5.39) the term $\sum_{n=-N+1}^{N-1} |n| \gamma(n) = O(N \ln N)$ and is thus negligible. Having obtained the asymptotic variance (5.40), the proof is then concluded by applying Davydov [1970] Theorem 2 (though this theorem was stated for a linearly interpolated version of $Y_N(t)$ in the space $C[0, 1]$, it is straightforward to adapt the the proof, which consists of showing convergence of finite-dimensional distributions and establishing tightness by moment estimate, to establish convergence in $D[0, 1]$ with the uniform metric.) \square

Remark 5.3.11. One may wonder if it is possible to get a different normalization in the nonlinear case in Theorem 5.3.3, since the normalization in the linear case in Theorem 5.3.5 has an extra $\sqrt{\ln N}$ factor. This is not possible under our setting where the kernel g is homogeneous with exponent α and is bounded by a linear combination of products of *purely* power functions $x_1^{\gamma_1} \dots x_k^{\gamma_k}$, where each $\gamma_j < -1/2$ and $\gamma_1 + \dots + \gamma_k = \alpha$.

Indeed, if one wanted to get some extra logarithmic factor in the covariance $\gamma(n)$, one would set for example $g(x_1, \dots, x_k) = x_1^{\gamma_1} \dots x_k^{\gamma_k}$ with $\gamma_k = -1$. But this will not achieve the stated goal. Indeed, by Lemma 5.3.8, using contradiction, we have $\alpha = \gamma_1 + \dots + \gamma_k <$

$-k/2 - 1/2$, which falls into the short memory regime (see Remark 5.3.2) and thus the normalization is \sqrt{N} as in (5.19).

Chapter 6

Functional Limit Theorems for Toeplitz Quadratic Functionals of Continuous time Gaussian Stationary Processes

The chapter establishes weak convergence in $C[0,1]$ of normalized stochastic processes, generated by Toeplitz type quadratic functionals of a continuous time Gaussian stationary process, exhibiting long-range dependence. Both central and non-central functional limit theorems are obtained.

6.1 Introduction

Let $\{X(t), t \in \mathbb{R}\}$ be a centered real-valued stationary Gaussian process with spectral density $f(x)$ and covariance function $r(t)$, that is, $r(t) = \widehat{f}(t) = \int_{\mathbb{R}} e^{ixt} f(x) dx$, $t \in \mathbb{R}$. We are interested in describing the limit (as $T \rightarrow \infty$) of the following process, generated by Toeplitz type quadratic functionals of the process $X(t)$:

$$Q_T(t) = \int_0^{Tt} \int_0^{Tt} \widehat{g}(u-v) X(u) X(v) du dv, \quad t \in [0, 1], \quad (6.1)$$

where

$$\widehat{g}(t) = \int_{\mathbb{R}} e^{ixt} g(x) dx, \quad t \in \mathbb{R}, \quad (6.2)$$

is the Fourier transform of some integrable even function $g(x)$, $x \in \mathbb{R}$. We will refer to $g(x)$ and to its Fourier transform $\widehat{g}(t)$ as a generating function and generating kernel for

the process $Q_T(t)$, respectively.

The limit of the process (6.1) is completely determined by the spectral density $f(x)$ (or covariance function $r(t)$) and the generating function $g(x)$ (or generating kernel $\widehat{g}(t)$), and depending on their properties, the limit can be either Gaussian (that is, $Q_T(t)$ with an appropriate normalization obeys a central limit theorem), or non-Gaussian. The following two questions arise naturally:

- (a) Under what conditions on $f(x)$ (resp. $r(t)$) and $g(x)$ (resp. $\widehat{g}(t)$) will the limit be Gaussian?
- (b) Describe the limit process, if it is non-Gaussian.

Similar questions were considered by Fox and Taqqu [1987], Ginovyan and Sahakyan [2005], and Terrin and Taqqu [1990] in the discrete time case.

Here we work in continuous time, and establish weak convergence in $C[0, 1]$ of the process (6.1). The limit processes can be Gaussian or non-Gaussian. The limit non-Gaussian process is identical to that of in the discrete time case, obtained in Terrin and Taqqu [1990].

But first some brief history. The question (a) goes back to the classical monograph by Grenander and Szegö [1958], where the problem was considered for discrete time processes, as an application of the authors' theory of the asymptotic behavior of the trace of products of truncated Toeplitz matrices (see Grenander and Szegö [1958], p. 217-219). Later the question (a) was studied by Ibragimov [1963] and Rosenblatt [1962], in connection to the statistical estimation of the spectral function $F(x)$ and covariance function $r(t)$, respectively. Since 1986, there has been a renewed interest in both questions (a) and (b), related to the statistical inferences for long memory processes (see, e.g., Avram [1988], Fox and Taqqu [1987], Ginovyan and Sahakyan [2005], Ginovyan et al. [2014], Giraitis et al. [2012], Giraitis and Surgailis [1990], Giraitis and Taqqu [2001], Terrin and Taqqu [1991a], Taniguchi and Kakizawa [2012], and references therein). In particular, Avram [1988], Fox and Taqqu [1987], Giraitis and Surgailis [1990], Ginovyan and Sahakyan [2005] have ob-

tained sufficient conditions for the Toeplitz type quadratic forms $Q_T(1)$ to obey the central limit theorem (CLT), when the model $X(t)$ is a discrete time process.

For continuous time processes the question (a) was studied in Ibragimov [1963] (in connection to the statistical estimation of the spectral function), Ginovyan and Sahakyan [2007] and Ginovyan et al. [2014], where sufficient conditions in terms of $f(x)$ and $g(x)$ ensuring central limit theorems for quadratic functionals $Q_T(1)$ have been obtained.

The rest of the chapter is organized as follows. In Section 6.2 we state the main results of this chapter (Theorems 6.2.1 - 6.2.9). In Section 7.3 we prove a number of preliminary lemmas that are used in the proofs of the main results. Section 6.4 contains the proofs of the main results.

Throughout the chapter the letters C and c with or without indices will denote positive constants whose values can change from line to line.

6.2 The Main Results

In this section we state our main results. Throughout the chapter we assume that $f, g \in L^1(\mathbb{R})$, and with no loss of generality, that $g \geq 0$ (see Ginovyan and Sahakyan [2007] and Giraitis and Surgailis [1990]).

We first examine the case of *central limit theorems*, and consider the following standard normalized version of (6.1):

$$\tilde{Q}_T(t) := T^{-1/2} (Q_T(t) - \mathbb{E}[Q_T(t)]), \quad t \in [0, 1]. \quad (6.3)$$

Our first result, which is an extension of Theorem 1 of Ginovyan and Sahakyan [2007], involves the convergence of finite-dimensional distributions of the process $\tilde{Q}_T(t)$ to that of a standard Brownian motion.

Theorem 6.2.1. *Assume that the spectral density $f(x)$ and the generating function $g(x)$*

satisfy the following conditions:

$$f \cdot g \in L^1(\mathbb{R}) \cap L^2(\mathbb{R}) \quad (6.4)$$

and

$$\mathbb{E}[\tilde{Q}_T^2(1)] \rightarrow 16\pi^3 \int_{-\infty}^{\infty} f^2(x)g^2(x)dx \quad \text{as } T \rightarrow \infty. \quad (6.5)$$

Then we have the following convergence of finite-dimensional distributions

$$\tilde{Q}_T(t) \xrightarrow{f.d.d.} \sigma B(t),$$

where $\tilde{Q}_T(t)$ is as in (6.3), $B(t)$ is a standard Brownian motion, and

$$\sigma^2 := 16\pi^3 \int_{-\infty}^{\infty} f^2(x)g^2(x)dx. \quad (6.6)$$

To extend the convergence of finite-dimensional distributions in Theorem 6.2.1 to the weak convergence in the space $C[0, 1]$, we impose an additional condition on the underlying Gaussian process $X(t)$ and on the generating function g . It is convenient to impose this condition in the time domain, that is, on the covariance function $r := \hat{f}$ and the generating kernel $a := \hat{g}$. The following condition is an analog of the assumption in Theorem 2.3 of Giraitis and Taqqu [2001]:

$$r(\cdot) \in L^p(\mathbb{R}), \quad a(\cdot) \in L^q(\mathbb{R}) \quad \text{for some } p, q \geq 1, \quad \frac{1}{p} + \frac{1}{q} \geq \frac{3}{2}. \quad (6.7)$$

Remark 6.2.2. In fact under (6.4), the condition (6.7) is sufficient for the convergence in (6.5). Indeed, let $\bar{p} = p/(p-1)$ be the Hölder conjugate of p and let $\bar{q} = q/(q-1)$ be the Hölder conjugate of q . Since $1 \leq p, q \leq 2$, one has by the Hausdorff-Young inequality and (6.7) that $\|f\|_{\bar{p}} \leq c_p \|r\|_p$, $\|g\|_{\bar{q}} \leq c_q \|a\|_q$, and hence

$$f(\cdot) \in L^{\bar{p}}, \quad g(\cdot) \in L^{\bar{q}}, \quad \frac{1}{\bar{p}} + \frac{1}{\bar{q}} = 2 - \frac{1}{p} - \frac{1}{q} \leq 1/2.$$

Then the convergence in (6.5) follows from the proof of Theorem 3 from Ginovyan and Sahakyan [2007]. Note that a similar assertion in the discrete time case was established in Giraitis and Surgailis [1990].

Remark 6.2.3. Observe that condition (6.7) is fulfilled if the functions $r(t)$ and $a(t)$ satisfy the following: there exist constants $C > 0$, α^* and β^* , such that

$$|r(t)| \leq C(1 \wedge |t|^{\alpha^*-1}), \quad |a(t)| \leq C(1 \wedge |t|^{\beta^*-1}), \quad (6.8)$$

where $0 < \alpha^*, \beta^* < 1/2$ and $\alpha^* + \beta^* < 1/2$. Indeed, to see this, note first that $r(\cdot)$, $a(\cdot) \in L^\infty(\mathbb{R})$. Then one can choose $p, q \geq 1$ such that $p(\alpha^* - 1) < -1$ and $q(\beta^* - 1) < -1$, which entails that $r(\cdot) \in L^p(\mathbb{R})$ and $a(\cdot) \in L^q(\mathbb{R})$. Since $1/p + 1/q < 2 - \alpha^* - \beta^*$ and $2 - \alpha^* - \beta^* > 3/2$, one can further choose p, q to satisfy $1/p + 1/q \geq 3/2$.

The next results, two functional central limit theorems, extend Theorems 1 and 5 of Ginovyan and Sahakyan [2007] to weak convergence in the space $C[0, 1]$ of the stochastic process $\tilde{Q}_T(t)$ to a standard Brownian motion.

Theorem 6.2.4. *Let the spectral density $f(x)$ and the generating function $g(x)$ satisfy condition (6.4). Let the covariance function $r(t)$ and the generating kernel $a(t)$ satisfy condition (6.7). Then we have the following weak convergence in $C[0, 1]$:*

$$\tilde{Q}_T(t) \Rightarrow \sigma B(t),$$

where $\tilde{Q}_T(t)$ is as in (6.3), σ is as in (6.6), and $B(t)$ is a standard Brownian motion.

Recall that a function $u(x)$, $x \in \mathbb{R}$, is called slowly varying at 0 if it is non-negative and for any $t > 0$

$$\lim_{x \rightarrow 0} \frac{u(xt)}{u(x)} \rightarrow 1.$$

Let $SV_0(\mathbb{R})$ be the class of slowly varying at zero functions $u(x)$, $x \in \mathbb{R}$, satisfying the following conditions: for some $a > 0$, $u(x)$ is bounded on $[-a, a]$, $\lim_{x \rightarrow 0} u(x) = 0$, $u(x) =$

$u(-x)$ and $0 < u(x) < u(y)$ for $0 < x < y < a$. An example of a function belonging to $SV_0(\mathbb{R})$ is $u(x) = |\ln|x||^{-\gamma}$ with $\gamma > 0$ and $a = 1$.

Theorem 6.2.5. *Assume that the functions f and g are integrable on \mathbb{R} and bounded outside any neighborhood of the origin, and satisfy for some $a > 0$*

$$f(x) \leq |x|^{-\alpha} L_1(x), \quad |g(x)| \leq |x|^{-\beta} L_2(x), \quad x \in [-a, a] \quad (6.9)$$

for some $\alpha < 1$, $\beta < 1$ with $\alpha + \beta \leq 1/2$, where $L_1(x)$ and $L_2(x)$ are slowly varying at zero functions satisfying

$$L_i \in SV_0(\mathbb{R}), \quad x^{-(\alpha+\beta)} L_i(x) \in L^2[-a, a], \quad i = 1, 2. \quad (6.10)$$

Let, in addition, the covariance function $r(t)$ and the generating kernel $a(t)$ satisfy condition (6.7). Then we have the following weak convergence in $C[0, 1]$:

$$\tilde{Q}_T(t) \Rightarrow \sigma B(t),$$

where $\tilde{Q}_T(t)$ is as in (6.3), σ is as in (6.6), and $B(t)$ is a standard Brownian motion.

Remark 6.2.6. The conditions $\alpha < 1$ and $\beta < 1$ ensure that the Fourier transforms of f and g are well defined. Observe that when $\alpha > 0$ the process $\{X(t), t \in \mathbb{Z}\}$ may exhibit long-range dependence. We also allow here $\alpha + \beta$ to assume the critical value $1/2$.

Remark 6.2.7. The assumptions $f \cdot g \in L^1(\mathbb{R})$, $f, g \in L^\infty(\mathbb{R} \setminus [-a, a])$ and (6.10) imply that $f \cdot g \in L^2(\mathbb{R})$, so that σ^2 in (6.6) is finite.

Remark 6.2.8. One may wonder, why, in Theorem 6.2.5, we suppose that $L_1(x)$ and $L_2(x)$ belong to $SV_0(\mathbb{R})$ instead of merely being slowly varying at zero. This is done in order to deal with the critical case $\alpha + \beta = 1/2$. Suppose that we are away from this critical case, namely, $f(x) = |x|^{-\alpha} l_1(x)$ and $g(x) = |x|^{-\beta} l_2(x)$, where $\alpha + \beta < 1/2$, and $l_1(x)$ and $l_2(x)$ are slowly varying at zero functions. Assume also that $f(x)$ and $g(x)$ are integrable

and bounded on $(-\infty, -a) \cup (a, +\infty)$ for any $a > 0$. We claim that Theorem 6.2.5 applies. Indeed, choose $\alpha' > \alpha$, $\beta' > \beta$ with $\alpha' + \beta' < 1/2$. Write $f(x) = |x|^{-\alpha'} |x|^\delta l_1(x)$, where $\delta = \alpha' - \alpha > 0$. Since $l_1(x)$ is slowly varying, when $|x|$ is small enough, for some $\epsilon \in (0, \delta)$ we have $|x|^\delta l_1(x) \leq |x|^{\delta-\epsilon}$. Then one can bound $|x|^{\delta-\epsilon}$ by $c |\ln |x||^{-1} \in SV_0(\mathbb{R})$ for small $|x| < 1$. Hence one has when $|x| < 1$ is small enough, $f(x) \leq |x|^{-\alpha'} \left(c |\ln |x||^{-1} \right)$. Similarly, when $|x| < 1$ is small enough, one has $g(x) \leq |x|^{-\beta'} \left(c |\ln |x||^{-1} \right)$. All the assumptions in Theorem 6.2.5 are now readily checked with α, β replaced by α' and β' , respectively.

Now we state a *non-central limit theorem* in the continuous time case. Let the spectral density f and the generating function g satisfy

$$f(x) = |x|^{-\alpha} L_1(x) \quad \text{and} \quad g(x) = |x|^{-\beta} L_2(x), \quad x \in \mathbb{R}, \quad \alpha < 1, \beta < 1, \quad (6.11)$$

with slowly varying at zero functions $L_1(x)$ and $L_2(x)$ such that $\int_{\mathbb{R}} |x|^{-\alpha} L_1(x) dx < \infty$ and $\int_{\mathbb{R}} |x|^{-\beta} L_2(x) dx < \infty$. We assume in addition that the functions $L_1(x)$ and $L_2(x)$ satisfy the following condition, called Potter's bound (see Giraitis et al. [2012], formula (2.3.5)): for any $\epsilon > 0$ there exists a constant $C > 0$ so that if T is large enough, then

$$\frac{L_i(u/T)}{L_i(1/T)} \leq C(|u|^\epsilon + |u|^{-\epsilon}), \quad i = 1, 2. \quad (6.12)$$

Note that a sufficient condition for (6.12) to hold is that $L_1(x)$ and $L_2(x)$ are bounded on intervals $[a, \infty)$ for any $a > 0$, which is the case for the slowly varying functions in Theorem 6.2.5.

Now we are interested in the limit process of the following normalized version of the process $Q_T(t)$ given by (6.1), with f and g as in (6.11):

$$Z_T(t) := \frac{1}{T^{\alpha+\beta} L_1(1/T) L_2(1/T)} (Q_T(t) - \mathbf{E}[Q_T(t)]). \quad (6.13)$$

Theorem 6.2.9. *Let f and g be as in (6.11) with $\alpha < 1$, $\beta < 1$ and slowly varying at zero functions $L_1(x)$ and $L_2(x)$ satisfying (6.12), and let $Z_T(t)$ be as in (6.13). Then for*

$\alpha + \beta > 1/2$, we have the following weak convergence in the space $C[0, 1]$:

$$Z_T(t) \Rightarrow Z(t),$$

where the limit process $Z(t)$ is given by

$$Z(t) = \int_{\mathbb{R}^2}'' H_t(x_1, x_2) W(dx_1) W(dx_2), \quad (6.14)$$

with

$$H_t(x_1, x_2) = |x_1 x_2|^{-\alpha/2} \int_{\mathbb{R}} \left[\frac{e^{it(x_1+u)} - 1}{i(x_1 + u)} \right] \cdot \left[\frac{e^{it(x_2-u)} - 1}{i(x_2 - u)} \right] |u|^{-\beta} du, \quad (6.15)$$

where $W(\cdot)$ is a complex Gaussian random measure with Lebesgue control measure, and the double prime in the integral (6.14) indicates that the integration excludes the diagonals $x_1 = \pm x_2$.

Remark 6.2.10. Comparing Theorem 6.2.9 and Theorem 1 of Terrin and Taqqu [1990], we see that the limit process $Z(t)$ is the same both for continuous and discrete time models.

Remark 6.2.11. Denoting by P_T and P the measures generated in $C[0, 1]$ by the processes $Z_T(t)$ and $Z(t)$ given by (6.13) and (6.14), respectively, Theorem 6.2.9 can be restated as follows: under the conditions of Theorem 6.2.9, the measure P_T converges weakly in $C[0, 1]$ to the measure P as $T \rightarrow \infty$. A similar assertion can be stated for Theorems 6.2.4 and 6.2.5.

It is worth noting that although the statement of our Theorem 6.2.9 is similar to that of Theorem 1 of Terrin and Taqqu [1990], the proof is different and simpler, and does not use the hard analysis of Terrin and Taqqu [1990], although some technical results of Terrin and Taqqu [1990] are stated in lemmas and used in the proofs. Our approach in the CLT case (Theorems 6.2.1 - 6.2.5), uses the method developed in Ginovyan and Sahakyan [2007], which itself is based on an approximation of the trace of the product of truncated Toeplitz operators. For the non-CLT case (Theorem 6.2.9), we use the integral representation of

the underlying process and properties of Wiener-Itô integrals.

6.3 Preliminaries

In this section we state a number of lemmas which will be used in the proof of the theorems. The following result extends Lemma 9 of Ginovyan and Sahakyan [2007].

Lemma 6.3.1. *Let $Y(t)$ be a centered stationary Gaussian process with spectral density $f_Y(x) \in L^1(\mathbb{R}) \cap L^2(\mathbb{R})$. Consider the normalized process:*

$$L_T(t) := \frac{1}{T^{1/2}} \left(\int_0^{Tt} Y^2(u) du - \mathbb{E} \left[\int_0^{Tt} Y^2(u) du \right] \right). \quad (6.16)$$

Then we have the following convergence of finite-dimensional distributions:

$$L_T(t) \xrightarrow{f.d.d.} \sigma_Y B(t), \quad \sigma_Y^2 = 4\pi \int_{-\infty}^{\infty} f_Y^2(x) dx, \quad (6.17)$$

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

Remark 6.3.2. Observe that the normalized processes $\tilde{Q}_T(t)$ and $L_T(t)$, given by (6.3) and (6.16), can be expressed by double Wiener-Itô integrals (see, e.g., the proof of Lemma 6.3.10 below). In our proofs we will use the following fact about weak convergence of multiple Wiener-Itô integrals: given the convergence of the covariance, the multivariate convergence to a Gaussian vector is implied by the univariate convergence of each component (see Peccati and Tudor [2005], Proposition 2).

Proof of Lemma 6.3.1. For a fixed t , the univariate convergence in distribution

$$L_T(t) \xrightarrow{d} N(0, t\sigma_Y^2) \quad \text{as } T \rightarrow \infty$$

follows from Lemma 9 of Ginovyan and Sahakyan [2007]. To show (6.17), in view of Remark 6.3.2 and Proposition 2 of Peccati and Tudor [2005], it remains to show that the covariance structure of $L_T(t)$ converges to that of $\sigma_Y B(t)$. Specifically, it suffices to show that for any

$0 < s < t$,

$$\mathbb{E} \left[(L_T(t) - L_T(s))^2 \right] \rightarrow \sigma_Y^2 \cdot (t - s) \quad \text{as } T \rightarrow \infty. \quad (6.18)$$

Indeed, using the fact that for a Gaussian vector (G_1, G_2) we have

$$\text{Cov}(G_1^2, G_2^2) = 2[\text{Cov}(G_1, G_2)]^2,$$

and letting $r_Y(u) = \int_{\mathbb{R}} e^{ixu} f_Y(x) dx$ be the covariance function of $Y(t)$, we can write

$$\mathbb{E} \left[(L_T(t) - L_T(s))^2 \right] = 2(t - s) \int_{-T(t-s)}^{T(t-s)} \left(1 - \frac{|u|}{T(t-s)} \right) r_Y^2(u) du.$$

Since $f_Y(x) \in L^2(\mathbb{R})$, the Fourier transform $r_Y(u) \in L^2(\mathbb{R})$ as well. So by the Dominated Convergence Theorem and Parseval-Plancherel's identity, we have as $T \rightarrow \infty$

$$\mathbb{E} \left[(L_T(t) - L_T(s))^2 \right] \rightarrow 2(t-s) \int_{-\infty}^{\infty} r_Y^2(u) du = 4\pi(t-s) \int_{-\infty}^{\infty} f_Y^2(x) dx = \sigma_Y^2(t-s). \quad (6.19)$$

□

We now discuss some results which allow one to reduce the general quadratic functional in Theorem 6.2.1 to a special quadratic functional introduced in Lemma 6.3.1.

By Theorem 16.7.2 from Ibragimov and Linnik [1971], the underlying process $X(t)$ admits a moving average representation:

$$X(t) = \int_{-\infty}^{\infty} \hat{a}(t-s) B(ds) \quad \text{with} \quad \int_{-\infty}^{\infty} |\hat{a}(t)|^2 dt < \infty, \quad (6.20)$$

where $B(t)$ is a standard Brownian motion, and $\hat{a}(t)$ is such that its inverse Fourier transform $a(x)$ satisfies $f(x) = 2\pi|a(x)|^2$. Assuming the conditions (6.4) and (6.5), we set

$$b(x) = (2\pi)^{1/2} a(x) (g(x))^{1/2},$$

and observe that the function $b(x)$ is then in $L^2(\mathbb{R})$ due to condition (6.4). Consider the

stationary process

$$Y(t) = \int_{-\infty}^{\infty} \hat{b}(t-s)B(ds) \quad (6.21)$$

constructed using the Fourier transform $\hat{b}(t)$ of $b(x)$ and the *same* Brownian motion $B(t)$ as in (6.20). The process $Y(t)$ has spectral density (see Ginovyan and Sahakyan [2007], equation (4.7))

$$f_Y(x) = 2\pi f(x)g(x). \quad (6.22)$$

We have the following approximation result which immediately follows from Lemma 10 of Ginovyan and Sahakyan [2007].

Lemma 6.3.3. *Let $\tilde{Q}_T(t)$ be as in (6.3) and let $L_T(t)$ be as in (6.16) with $Y(t)$ constructed as in (6.21). Then under the conditions (6.4) and (6.5), for any $t > 0$, we have*

$$\lim_{T \rightarrow \infty} \text{Var}[\tilde{Q}_T(t) - L_T(t)] = 0.$$

The following lemma is a straightforward adaptation of Lemma 4.2 of Giraitis and Taqqu [1998] for functions defined on \mathbb{R} .

Lemma 6.3.4. *If $p_j \geq 1$, $j = 1, \dots, k$, where $k \geq 2$ and $\sum_{j=1}^k \frac{1}{p_j} = k - 1$, then*

$$\int_{\mathbb{R}^{k-1}} |f_1(x_1) \dots f_{k-1}(x_{k-1}) f_k(x_1 + \dots + x_{k-1})| dx_1 \dots dx_{k-1} \leq \prod_{j=1}^k \|f_j\|_{p_j}.$$

The following lemma will be used to establish tightness in the space $C[0, 1]$ in Theorem 6.2.4.

Lemma 6.3.5. *Let the covariance function $r(t)$ and the generating kernel $a(t)$ satisfy condition (6.7), and let $\tilde{Q}_T(t)$ be as in (6.3). Then for all $0 \leq s \leq t \leq 1$ and $T > 0$, there exists a constant $C > 0$, such that*

$$\mathbb{E} \left[|\tilde{Q}_T(t) - \tilde{Q}_T(s)|^2 \right] \leq C(t - s). \quad (6.23)$$

Proof. For convenience we use the Wick product notation: $:X(u)X(v): := X(u)X(v) - \mathbb{E}[X(u)X(v)]$. So for $0 \leq s \leq t \leq 1$, we can write

$$\begin{aligned} & \tilde{Q}_T(t) - \tilde{Q}_T(s) \\ &= \frac{1}{\sqrt{T}} \left(\int_0^{Tt} \int_0^{Tt} a(u-v) :X(u)X(v): dudv - \int_0^{Ts} \int_0^{Ts} a(u-v) :X(u)X(v): dudv \right) \\ &= \frac{1}{\sqrt{T}} \int_{Ts}^{Tt} \int_{Ts}^{Tt} a(u-v) :X(u)X(v): dudv + \frac{2}{\sqrt{T}} \int_0^{Ts} \int_{Ts}^{Tt} a(u-v) :X(u)X(v): dudv \\ &:= A(s, t, T) + B(s, t, T). \end{aligned}$$

Now we estimate $B(s, t, T)$ (the function $A(s, t, T)$ can be estimated similarly). We have by Theorem 3.9 of Janson [1997] that

$$\begin{aligned} \mathbb{E}[B^2(s, t, T)] &= \frac{4}{T} \int_0^{Ts} du_1 \int_{Ts}^{Tt} dv_1 \int_0^{Ts} du_2 \int_{Ts}^{Tt} dv_2 \times \\ &\quad a(u_1 - v_1)a(u_2 - v_2)\mathbb{E}(:X(u_1)X(v_1)::X(u_2)X(v_2):) \\ &= \frac{4}{T} \int_0^{Ts} du_1 \int_{Ts}^{Tt} dv_1 \int_0^{Ts} du_2 \int_{Ts}^{Tt} dv_2 a(u_1 - v_1)a(u_2 - v_2) \times \\ &\quad [r(u_1 - u_2)r(v_1 - v_2) + r(u_1 - v_2)r(v_1 - u_2)] \\ &:= B_1(s, t, T) + B_2(s, t, T). \end{aligned}$$

By the change of variables $x_1 = u_1 - v_1$, $x_2 = v_2 - u_2$, $x_3 = u_2 - u_1$, $x_4 = v_2$, and noting that $r(\cdot)$ and $a(\cdot)$ are even functions, we have

$$B_1(s, t, T) \leq \frac{4}{T} \int_{Ts}^{Tt} dx_4 \int_{\mathbb{R}^3} |a(x_1)a(x_2)r(x_3)r(x_1 + x_2 + x_3)| dx_1 dx_2 dx_3.$$

Since $|r(t)| \leq r(0)$, we have $r(\cdot) \in L^\infty(\mathbb{R})$. We also have $r(\cdot) \in L^p(\mathbb{R})$ by condition (6.7), where $1/p + 1/q \geq 3/2$. The L^p -interpolation theorem states that if a function is in L^{p_1} and L^{p_2} with $0 < p_1 \leq p_2 \leq \infty$, then it is in $L^{p'}$, $p_1 \leq p' \leq p_2$. By the L^p -interpolation

theorem, one can choose $p' \geq p$ such that $r(\cdot) \in L^{p'}(\mathbb{R})$ and

$$\frac{1}{p'} + \frac{1}{p'} + \frac{1}{q} + \frac{1}{q} = 3, \quad \text{that is, } \frac{1}{p'} + \frac{1}{q} = \frac{3}{2}.$$

Then by Lemma 6.3.4, one has $B_1(s, t, T) \leq 4\|r\|_{p'}^2 \|a\|_q^2 (t-s)$. Similarly, one can establish the bound $B_2(s, t, T) \leq C(t-s)$, and hence $B(s, t, T) \leq C(t-s)$. So (6.23) is proved. \square

The lemmas that follow will be used in the proof of Theorem 6.2.9.

Lemma 6.3.6. *Define*

$$\Delta_t(x) = \int_0^t e^{isx} ds = \frac{e^{itx} - 1}{ix}, \quad (6.24)$$

Then for any $\delta \in (0, 1)$, there exists a constant $c > 0$ depending only on δ , such that

$$|\Delta_t(x)| \leq c|t|^\delta f_\delta(x), \quad t \in [0, 1], \quad x \in \mathbb{R}, \quad (6.25)$$

where

$$f_\delta(x) = \begin{cases} |x|^{\delta-1} & \text{if } |x| > 1; \\ 1 & \text{if } |x| \leq 1. \end{cases} \quad (6.26)$$

Proof. In view of (6.24), we have $|\Delta_t(x)| \leq \int_0^t |e^{isx}| ds = t$. So under the constraint $t \in [0, 1]$, we have $|\Delta_t(x)| \leq t \leq t^\delta$. On the other hand, from Lemma 2 from Terrin and Taqqu [1990], with some constant $C > 0$, we have $|e^{ix} - 1| \leq C|x|^\delta$, $\delta \in (0, 1)$. So

$$|\Delta_t(x)| \leq \frac{|e^{itx} - 1|}{|x|} \leq C|tx|^\delta |x|^{-1} = Ct^\delta |x|^{\delta-1}.$$

Combining this with (6.26), we obtain (6.25). \square

We quote Lemma 1 of Terrin and Taqqu [1990] in a special case, convenient for our purposes.

Lemma 6.3.7. *Let $\gamma_i < 1$, $\gamma_i + \gamma_{i+1} > 1/2$, and let δ be such that*

$$0 \leq \delta < \frac{\gamma_i + \gamma_{i+1}}{2},$$

where $i = 1, \dots, 4$ (with $\gamma_5 = \gamma_1$). Then

$$\int_{\mathbb{R}^4} f_\delta(y_1 - y_2) f_\delta(y_2 - y_3) f_\delta(y_3 - y_4) f_\delta(y_4 - y_1) |y_1|^{-\gamma_1} |y_2|^{-\gamma_2} |y_3|^{-\gamma_3} |y_4|^{-\gamma_4} d\mathbf{y} < \infty,$$

where $f_\delta(\cdot)$ is as in (6.26).

Lemma 6.3.7 can be used to establish the following result.

Lemma 6.3.8. *The function*

$$H_t^*(x_1, x_2) := |x_1|^{\alpha_1/2} |x_2|^{\alpha_2/2} \int_{\mathbb{R}} |\Delta_t(x_1 + u) \Delta_t(x_2 - u)| |u|^{-\beta} du \quad (6.27)$$

is in $L^2(\mathbb{R}^2)$ for all $(\alpha_1, \alpha_2, \beta)$ in the open region $\{(\alpha_1, \alpha_2, \beta) : \alpha_1, \alpha_2, \beta < 1, \alpha_i + \beta > 1/2, i = 1, 2\}$.

Proof. It suffices focus on the case where $t \in [0, 1]$, otherwise a change of variable can reduce it to this case. We have by suitable change of variables and Lemma 6.3.6 that

$$\begin{aligned} & \|H_t^*\|_{L^2(\mathbb{R}^2)}^2 \\ &= \int_{\mathbb{R}^4} |\Delta_t(y_1 - y_2) \Delta_t(y_2 - y_3) \Delta_t(y_3 - y_4) \Delta_t(y_4 - y_1)| |y_1|^{-\alpha_1} |y_2|^{-\beta} |y_3|^{-\alpha_2} |y_4|^{-\beta} d\mathbf{y} \\ &\leq C \int_{\mathbb{R}^4} f_\delta(y_1 - y_2) f_\delta(y_2 - y_3) f_\delta(y_3 - y_4) f_\delta(y_4 - y_1) |y_1|^{-\alpha_1} |y_2|^{-\beta} |y_3|^{-\alpha_2} |y_4|^{-\beta} d\mathbf{y}. \end{aligned}$$

Then apply Lemma 6.3.7, noting that δ can be chosen arbitrarily small. \square

Lemma 6.3.9. *Define the function*

$$H_{t,T}^*(x_1, x_2) = A_{1,T}(x_1, x_2) |x_1 x_2|^{-\alpha/2} \int_{\mathbb{R}} |\Delta_t(x_1 + u) \Delta_t(x_2 - u)| |u|^{-\beta} A_{2,T}(u) du, \quad (6.28)$$

where

$$A_{1,T}(x_1, x_2) = \sqrt{\frac{L_1(x_1/T) L_1(x_2/T)}{L_1(1/T) L_1(1/T)}}, \quad A_{2,T}(u) = \frac{L_2(u/T)}{L_2(1/T)}. \quad (6.29)$$

Then for large enough T , we have $H_{t,T}^*(x_1, x_2) \in L^2(\mathbb{R}^2)$.

Proof. By (6.12) and (6.29), for any $\epsilon > 0$ there exists $C > 0$, such that for T large enough,

$$|A_{1,T}(x_1, x_2)| \leq C(|x_1|^\epsilon + |x_1|^{-\epsilon})(|x_2|^\epsilon + |x_2|^{-\epsilon}) \quad (6.30)$$

and

$$|A_{2,T}(u)| \leq C(|u|^\epsilon + |u|^{-\epsilon}). \quad (6.31)$$

Hence, with some constant $C > 0$,

$$\begin{aligned} |H_{t,T}^*(x_1, x_2)| &\leq C \int_{\mathbb{R}} |\Delta_t(x_1 + u) \Delta_t(x_2 - u)| |u|^{-\beta} (|u|^\epsilon + |u|^{-\epsilon}) du \times \\ &|x_1 x_2|^{-\alpha/2} (|x_1|^\epsilon + |x_1|^{-\epsilon})(|x_2|^\epsilon + |x_2|^{-\epsilon}). \end{aligned} \quad (6.32)$$

Because by Lemma 6.3.8, the function H_t^* in (6.27) is in $L^2(\mathbb{R}^2)$ for all $(\alpha_1, \alpha_2, \beta)$ in an open region $\{(\alpha, \beta) : \alpha_1, \alpha_2, \beta < 1, \alpha_i + \beta > 1/2, i = 1, 2\}$. By choosing ϵ small enough, we infer that the right-hand side of (6.32) is in $L^2(\mathbb{R}^2)$, and the result follows. \square

Lemma 6.3.10. *Let $Z_T(t)$ be as in (6.13), and let*

$$Z_T'(t) := \int_{\mathbb{R}^2}'' H_{t,T}(x_1, x_2) W(dx_1) W(dx_2), \quad (6.33)$$

where

$$H_{t,T}(x_1, x_2) = A_{1,T}(x_1, x_2) |x_1 x_2|^{-\alpha/2} \left[\int_{\mathbb{R}} \Delta_t(x_1 + u) \Delta_t(x_2 - u) |u|^{-\beta} A_{2,T}(u) du \right]. \quad (6.34)$$

Then $Z_T(t) \stackrel{f.d.d.}{=} Z_T'(t)$, that is, the processes $Z_T(t)$ and $Z_T'(t)$ have the same finite-dimensional distributions.

Proof. Using the spectral representation of $X(t)$ (see, e.g., Doob [1953], Chapter XI, Section 8): $X(t) = \int_{\mathbb{R}} e^{itx} \sqrt{f(x)} W(dx)$, where $W(\cdot)$ is a complex Gaussian measure with Lebesgue control measure, and the diagram formula (see, e.g., Major [2014], Chapter 5), we have

$$X(u)X(v) - \mathbb{E}[X(u)X(v)] = \int_{\mathbb{R}^2} e^{i(ux_1+vx_2)} \sqrt{f(x_1)f(x_2)} W(dx_1)W(dx_2).$$

By a stochastic Fubini Theorem (see Pipiras and Taqqu [2010], Theorem 2.1) and Lemma 6.3.9, one can change the integration order to get (note that by (6.2) we have $\widehat{g}(t) = \int_{\mathbb{R}} e^{itx} g(x) dx$):

$$\begin{aligned} & [T^{\alpha+\beta} L_1(1/T) L_2(1/T)] Z_T(t) \\ &= \int_{\mathbb{R}^2} \sqrt{f(x_1)f(x_2)} \int_0^{Tt} \int_0^{Tt} \int_{\mathbb{R}} e^{i(u-v)w} g(w) dw e^{i(ux_1+vx_2)} dudv W(dx_1)W(dx_2) \\ &= \int_{\mathbb{R}^2} \sqrt{f(x_1)f(x_2)} \int_{\mathbb{R}} \int_0^{Tt} e^{iu(x_1+w)} du \int_0^{Tt} e^{iv(x_2-w)} dv |w|^{-\beta} L(w) dw W(dx_1)W(dx_2) \\ &= \int_{\mathbb{R}^2} \sqrt{f(x_1)f(x_2)} \int_{\mathbb{R}} \Delta_{Tt}(x_1+w) \Delta_{Tt}(x_2-w) |w|^{-\beta} L_2(w) dw W(dx_1)W(dx_2). \end{aligned}$$

Now we use the change of variables $w \rightarrow w/T$, $x_1 \rightarrow x_1/T$, $x_2 \rightarrow x_2/T$, where the latter two change of variables are subject to the rule $W(dx/T) \stackrel{d}{=} T^{-1/2} W(dx)$ (see, e.g., Dobrushin [1979], Proposition 4.2), to obtain

$$\begin{aligned} Z_T(t) \stackrel{f.d.d.}{=} & \frac{1}{T^{\alpha+\beta} L_1(1/T) L_2(1/T)} \times \int_{\mathbb{R}^2} \sqrt{f(x_1/T) f(x_2/T)} \times \\ & \int_{\mathbb{R}} \Delta_t(x_1+u) \Delta_t(x_2-u) |w/T|^{-\beta} L_2(w/T) T dw T^{-1} W(dx_1) W(dx_2). \quad (6.35) \end{aligned}$$

Taking into account the equality $f(x/T) = |x/T|^{-\alpha} L_1(x/T)$ and equations in (6.29), we see that the right hand side of (6.35) coincides with (6.33). This completes the proof. \square

The lemmas that follow will be used to establish tightness in the space $C[0, 1]$ in Theorem 6.2.9.

Lemma 6.3.11. *Let δ be a fixed number within the range $(0, (\alpha + \beta)/2)$, and let $Z_T(t)$ be*

as in (6.13). Then for all $0 \leq s \leq t \leq 1$ and T large enough, there exists a constant $C > 0$, such that

$$\mathbb{E} [|Z_T(t) - Z_T(s)|^2] \leq C(t-s)^{2\delta}. \quad (6.36)$$

The same estimate also holds for the corresponding limiting process $Z(t)$ defined by (6.14), (6.15).

Proof. First, in view of Lemma 6.3.10, we have $\mathbb{E} [|Z_T(t) - Z_T(s)|^2] = \mathbb{E} [|Z'_T(t) - Z'_T(s)|^2]$. Next, using the linearity of the multiple stochastic integral, we can write

$$Z'_T(t) - Z'_T(s) = \int_{\mathbb{R}^2}'' H_{s,t,T}(x_1, x_2) W(dx_1) W(dx_2),$$

where

$$\begin{aligned} H_{s,t,T}(x_1, x_2) &= A_{1,T}(x_1, x_2) |x_1 x_2|^{-\alpha/2} \times \\ &\int_{\mathbb{R}} [\Delta_t(x_1 + u) \Delta_t(x_2 - u) - \Delta_s(x_1 + u) \Delta_s(x_2 - u)] |u|^{-\beta} A_{2,T}(u) du. \end{aligned} \quad (6.37)$$

The term in the brackets of the integrand in (6.37) can be rewritten as follows:

$$\begin{aligned} &\Delta_t(x_1 + u) \Delta_t(x_2 - u) - \Delta_s(x_1 + u) \Delta_s(x_2 - u) \\ &= \int_0^t \int_0^t e^{iw_1(x_1+u)} e^{iw_2(x_2-u)} dw_1 dw_2 - \int_0^s \int_0^s e^{iw_1(x_1+u)} e^{iw_2(x_2-u)} dw_1 dw_2 \\ &= \int_0^s dw_1 \int_s^t dw_2 \dots + \int_s^t dw_1 \int_0^s dw_2 \dots + \int_s^t dw_1 \int_s^t dw_2 \dots \\ &= \Delta_s(x_1 + u) \Delta_{t-s}(x_2 - u) + \Delta_{t-s}(x_1 + u) \Delta_s(x_2 - u) + \Delta_{t-s}(x_1 + u) \Delta_{t-s}(x_2 - u). \end{aligned}$$

Now we apply Lemma 6.3.6 to get

$$\begin{aligned} &|\Delta_t(x_1 + u) \Delta_t(x_2 - u) - \Delta_s(x_1 + u) \Delta_s(x_2 - u)| \\ &\leq C[s^\delta(t-s)^\delta + (t-s)^\delta s^\delta + (t-s)^{2\delta}] f_\delta(x_1 + u) f_\delta(x_2 - u) \end{aligned}$$

$$\leq C(t-s)^\delta f_\delta(x_1+u)f_\delta(x_2-u), \quad (6.38)$$

where the last inequality follows because $0 \leq s^\delta \leq 1$ and $0 \leq (t-s)^\delta \leq 1$.

Next, using formula (4.5') of Major [2014], (6.37) and (6.38), we can write

$$\begin{aligned} \mathbb{E} [|Z_T(t) - Z_T(s)|^2] &= \|H_{s,t,T}\|_{L^2(\mathbb{R}^2)}^2 \leq C|t-s|^{2\delta} \int_{\mathbb{R}^2} dx_1 dx_2 A_{1,T}(x_1, x_2)^2 |x_1 x_2|^{-\alpha} \times \\ &\int_{\mathbb{R}^2} du_1 du_2 f_\delta(x_1+u_1) f_\delta(x_2-u_1) f_\delta(-x_1+u_2) f_\delta(-x_2-u_2) |u_1 u_2|^{-\beta} A_{2,T}(u_1) A_{2,T}(u_2) \\ &\leq C|t-s|^{2\delta} \int_{\mathbb{R}^4} dy_1 dy_2 dy_3 dy_4 A_{1,T}(y_1, y_3)^2 A_{2,T}(y_2) A_{2,T}(y_4) \times \\ & f_\delta(y_1-y_2) f_\delta(y_2-y_3) f_\delta(y_3-y_4) f_\delta(y_4-y_1) |y_1|^{-\alpha} |y_2|^{-\beta} |y_3|^{-\alpha} |y_4|^{-\beta}, \end{aligned} \quad (6.39)$$

where we have applied the change of variables: $y_1 = x_1$, $y_2 = -u_1$, $y_3 = -x_2$, $y_4 = u_2$.

Since by assumption $\alpha < 1$, $\beta < 1$ and $\alpha + \beta > 1/2$, and the exponent ϵ in (6.30) and (6.31) can be chosen arbitrarily small, for a fixed δ satisfying $0 < \delta < (\alpha + \beta)/2$, we can apply Lemma 1 of Terrin and Taqqu [1990] to conclude that the integral

$$\begin{aligned} &\int_{\mathbb{R}^4} A_{1,T}(y_1, y_3)^2 A_{2,T}(y_2) A_{2,T}(y_4) f_\delta(y_1-y_2) f_\delta(y_2-y_3) f_\delta(y_3-y_4) f_\delta(y_4-y_1) \times \\ & |y_1|^{-\alpha} |y_2|^{-\beta} |y_3|^{-\alpha} |y_4|^{-\beta} d\mathbf{y} \end{aligned}$$

is bounded for sufficiently large T , which in view of (6.39) implies (6.36). The proof for $Z_T(t)$ is thus complete. The proof for $Z(t)$ is similar and so we omit the details. \square

6.4 Proof of Main Results

Proof of Theorem 6.2.1. By Lemma 6.3.3, for any $0 \leq t_1 < \dots < t_n$, and constants c_1, \dots, c_n , we have

$$\lim_{T \rightarrow \infty} \text{Var} \left[\sum_{j=1}^n c_j \left(\tilde{Q}_T(t_j) - L_T(t_j) \right) \right] = 0.$$

Therefore the convergence of finite-dimensional distributions of $\tilde{Q}_T(t)$ to that of Brownian motion $\sigma B(t)$ follows from Lemma 6.3.1 with $f_Y(\cdot)$ given in (6.22) and the Cramér-Wold Device. \square

Proof of Theorem 6.2.4. In view of the well-known Prokhorov's Theorem (see, e.g., Billingsley [1999], p. 58), to prove the theorem, we need to show convergence of finite-dimensional distributions and tightness. The former has been established in Theorem 6.2.1. To prove tightness, observe that by Lemma 6.3.5 and the hypercontractivity inequality of the multiple Wiener-Itô integrals (see Major [2014], Corollary 5.6), for any $T > 0$ and $0 \leq s \leq t \leq 1$, there exists a constant $C > 0$ to satisfy

$$\mathbb{E} \left[|\tilde{Q}_T(t) - \tilde{Q}_T(s)|^4 \right] \leq C_2 \left(\mathbb{E} \left[|\tilde{Q}_T(t) - \tilde{Q}_T(s)|^2 \right] \right)^2 \leq C(t-s)^2. \quad (6.40)$$

Now the tightness of the family of measures generated by the processes $\{\tilde{Q}_T(t) : T > 0\}$ in $C[0, 1]$ follows from Lemma 5.1 of Ibragimov [1963]. \square

Proof of Theorem 6.2.5. The convergence of finite-dimensional distributions follows from Theorem 6.2.1. In fact, the assumptions on f and g in Theorem 6.2.5 imply the conditions (6.4) and (6.5) in Theorem 6.2.1 (see the proof of Theorem 5 in Ginovyan and Sahakyan [2007]). The tightness can be shown similarly as in the proof of Theorem 6.2.4. \square

Proof of Theorem 6.2.9. As in the proof of Theorem 6.2.4, we need to show convergence of finite-dimensional distributions and tightness. We first prove the convergence of finite-dimensional distributions, that is, $Z_T(t) \xrightarrow{f.d.d.} Z(t)$ as $T \rightarrow \infty$, where $Z_T(t)$ and $Z(t)$ are defined by (6.13) and (6.14), respectively.

By Lemma 6.3.10, the process $Z_T(t)$ defined in (6.13) has the same finite-dimensional distributions as the process $Z'_T(t)$ defined in (6.33). Therefore, taking into account the linearity of multiple Wiener-Itô integral, and applying Crámer-Wold device, to prove

$Z_T(t) \xrightarrow{f.d.d.} Z(t)$, it is enough to show that as $T \rightarrow \infty$,

$$H_{t,T}(x_1, x_2) \rightarrow H_t(x_1, x_2) \quad \text{in} \quad L^2(\mathbb{R}^2), \quad (6.41)$$

where $H_t(x_1, x_2)$ and $H_{t,T}(x_1, x_2)$ are as in (6.15) and (6.34), respectively.

First, we show pointwise convergence for a.e. $(x_1, x_2) \in \mathbb{R}^2$, that is,

$$H_{t,T}(x_1, x_2) = A_{1,T}(x_1, x_2)|x_1 x_2|^{-\alpha/2} \int_{\mathbb{R}} \Delta_t(x_1 + u) \Delta_t(x_2 - u) |u|^{-\beta} A_{2,T}(u) du \quad (6.42)$$

$$\rightarrow H_t(x_1, x_2) = |x_1 x_2|^{-\alpha/2} \int_{\mathbb{R}} \Delta_t(x_1 + u) \Delta_t(x_2 - u) |u|^{-\beta} du \quad \text{as } T \rightarrow \infty. \quad (6.43)$$

Because $L_1(x)$ is a slowly varying function, we have $A_{1,T}(x_1, x_2) \rightarrow 1$ as $T \rightarrow \infty$, where $A_{1,T}$ is as in (6.29). To show that the integral in (6.42) converges to the integral in (6.43), note first that by (6.29), $A_{2,T}(u) \rightarrow 1$ as $T \rightarrow \infty$ because $L_2(x)$ is a slowly varying function. Hence one only needs to bound the integrand properly and apply the Dominated Convergence Theorem. To this end, observe that by (6.31) for T large enough, we have

$$g_T(u; x_1, x_2) := |\Delta_t(x_1 + u)| |\Delta_t(x_2 - u)| |u|^{-\beta} A_{2,T}(u) \quad (6.44)$$

$$\leq C |\Delta_t(x_1 + u)| |\Delta_t(x_2 - u)| |u|^{-\beta} (|u|^\epsilon + |u|^{-\epsilon}) := g_\epsilon(u; x_1, x_2). \quad (6.45)$$

By choosing ϵ small enough, using Fubini Theorem and Lemma 6.3.8, we conclude that $g_\epsilon(\cdot; x_1, x_2) \in L^1(\mathbb{R})$ for a.e. $(x_1, x_2) \in \mathbb{R}^2$. Now (6.41) follows from (6.32) and the Dominated Convergence Theorem.

To prove tightness, first observe that by the hypercontractivity inequality of the multiple Wiener-Itô integrals (see Major [2014], Corollary 5.6) and Lemma 6.3.11, for T large enough and for any $0 \leq s \leq t \leq 1$, there exists a constant $C > 0$ to satisfy

$$\mathbb{E} [|Z_T(t) - Z_T(s)|^4] \leq C_2 (\mathbb{E} [|Z_T(t) - Z_T(s)|^2])^2 \leq C |t - s|^{4\delta}, \quad (6.46)$$

where δ is a fixed number within the range $0 < 4\delta < 2(\alpha + \beta)$. Since by assumption $\alpha + \beta > 1/2$, we can choose δ to satisfy $4\delta > 1$. Inequalities similar to (6.46) hold also for the limit process $Z(t)$.

In view of (6.46) and a similar inequality for $Z(t)$, it follows from Kolmogorov's criterion (see, e.g., Bass [2011] Theorem 8.1(1)) that the processes $Z_T(t)$ and $Z(t)$ admit continuous versions when T is large enough.

Now the tightness of the family of measures generated by the processes $\{Z_T(t) : T > 0\}$ in $C[0, 1]$ follows from Lemma 5.1 of Ibragimov [1963]. Theorem 6.2.9 is proved. \square

Chapter 7

Limit theorems for quadratic forms of Lévy-driven continuous-time linear processes

We study the asymptotic behavior of a suitable normalized stochastic process $\{Q_T(t), t \in [0, 1]\}$. This stochastic process is generated by a Toeplitz type quadratic functional of a Lévy-driven continuous-time linear process. We show that under some L^p -type conditions imposed on the covariance function of the model and the kernel of the quadratic functional, the process $Q_T(t)$ obeys a central limit theorem, that is, the finite-dimensional distributions of the standard \sqrt{T} normalized process $Q_T(t)$ tend to those of a normalized standard Brownian motion. In contrast, when the covariance function of the model and the kernel of the quadratic functional have a slow power decay, then we have a non-central limit theorem for $Q_T(t)$, that is, the finite-dimensional distributions of the process $Q_T(t)$, normalized by T^γ for some $\gamma > 1/2$, tend to those of a non-Gaussian non-stationary-increment self-similar process which can be represented by a double stochastic Wiener-Itô integral on \mathbb{R}^2 .

7.1 Introduction

Let $\{X(t), t \in \mathbb{R}\}$ be a Lévy-driven continuous-time stationary linear process defined by

$$X(t) = \int_{\mathbb{R}} a(t-s)\xi(ds), \quad (7.1)$$

where $a(\cdot)$ is a function from $L^2(\mathbb{R})$, and $\xi(t)$ is a Lévy process satisfying the conditions:

$$E\xi(t) = 0, E\xi^2(1) = 1 \text{ and } E\xi^4(1) < \infty.$$

A Lévy process, $\{\xi(t), t \in \mathbb{R}\}$ is a process with independent and stationary increments, continuous in probability, with sample-paths which are right-continuous with left limits (càdlàg) and $\xi(0) = \xi(0-) = 0$. The Wiener process $\{B(t), t \geq 0\}$ and the centered Poisson process $\{N(t) - EN(t), t \geq 0\}$ are typical examples of centered Lévy processes. Notice that the covariance function of $X(t)$ is given by

$$r(t) = EX(t)X(0) = \int_{\mathbb{R}} a(t+x)a(x)dx, \quad (7.2)$$

and it possesses the spectral density

$$f(\lambda) = \frac{\sigma^2}{2\pi} |\widehat{a}(\lambda)|^2 = \frac{\sigma^2}{2\pi} \left| \int_{\mathbb{R}} e^{-i\lambda t} a(t) dt \right|^2, \quad \lambda \in \mathbb{R}. \quad (7.3)$$

The function $a(\cdot)$ plays the role of a time-invariant filter.

Processes of the form (7.1) appear in many fields of science (economics, finance, physics, etc.), and cover a large class of popular models in continuous-time time series modeling. For instance, the so-called continuous-time autoregressive moving average (CARMA) models, which are the continuous-time analogs of the classical autoregressive moving average (ARMA) models in discrete-time case, are of the form (7.1) and play a central role in the representation of continuous-time stationary time series. Lévy-driven CARMA processes permit the modelling of heavy-tailed and asymmetric time series and incorporate both distributional and sample-path information (see, e.g., Brockwell [2001, 2014]).

Consider the following Toeplitz type quadratic functional of the process $X(u)$:

$$Q_T := \int_0^T \int_0^T b(u-v)X(u)X(v) du dv, \quad T > 0, \quad (7.4)$$

where

$$b(t) := \widehat{g}(t) = \int_{\mathbb{R}} e^{i\lambda t} g(\lambda) d\lambda, \quad t \in \mathbb{R},$$

is the Fourier transform of some integrable even function $g(\lambda)$, $\lambda \in \mathbb{R}$. We will refer to $g(\lambda)$ and to its Fourier transform $b(t)$ as a generating function and generating kernel for the functional Q_T , respectively.

In this chapter we are interested in the asymptotic behavior as $(T \rightarrow \infty)$ of the stochastic process $\{Q_T(t), t \in [0, 1]\}$, generated by the functional Q_T :

$$Q_T(t) := \int_0^{Tt} \int_0^{Tt} b(u-v) X(u) X(v) du dv, \quad t \in [0, 1]. \quad (7.5)$$

Our goal is to establish functional limit theorems of the form

$$\frac{1}{A(T)} (Q_T(t) - \mathbb{E}Q_T(t)) \xrightarrow{f.d.d.} L(t), \quad (7.6)$$

where $A(T)$ is a normalization factor, $L(t)$ is the limit process, and the symbol $\xrightarrow{f.d.d.}$ stands for convergence of finite-dimensional distributions.

Functionals of the form (7.5) and their discrete counterparts arise naturally in the statistical estimation of the spectrum of stationary processes. Limits such as (7.6) are necessary to establish asymptotic properties of these estimators (see, for example, Fox and Taqqu [1986], Ginovyan [2011], Giraitis et al. [2012], and references therein).

In the case where the underlying model $\{X(u), u \in \mathbb{R}\}$ is a Wiener-driven process, that is, $X(u)$ is a Gaussian process, limit theorems of the form (7.4) were established in Bai et al. [2015], among others, where it was shown that if both the spectral density f of $X(u)$ and the generating function g are regularly varying at the origin of orders α and β , respectively, then it is the sum $\alpha + \beta$ that determines the limiting process $L(t)$. In fact, when

$$\alpha + \beta \leq 1/2,$$

the limit process $L(t)$ is a normalized standard Brownian motion, while when

$$\alpha + \beta > 1/2,$$

the limit $L(t)$ is a non-Gaussian self-similar process, which can be represented as a double Wiener-Itô integral on \mathbb{R}^2 .

In this chapter, we consider the general case where the model $\{X(u), u \in \mathbb{R}\}$ is a continuous-time linear process driven from Lévy noise $\xi(u)$ with time invariant filter $a(\cdot)$. Specifically, we show that under some L^p -type conditions imposed on the filter $a(\cdot)$ and the kernel $b(\cdot)$ of the quadratic functional, the process $Q_T(t)$ obeys a central limit theorem, that is, the finite-dimensional distributions of the standard \sqrt{T} normalized process $Q_T(t)$ tend to those of a normalized standard Brownian motion. In contrast, when the functions $a(\cdot)$ and $b(\cdot)$ have slow power decay, then we have a non-central limit theorem for $Q_T(t)$, that is, the finite-dimensional distributions of the process $Q_T(t)$, normalized by T^γ for some $\gamma > 1/2$, tend to those of a non-Gaussian non-stationary-increment self-similar process which can be represented by a double stochastic Wiener-Itô integral on \mathbb{R}^2 .

We point out that our proofs of the central limit theorems are based on a new approximation approach which reduces the quadratic integral form to a single integral form. This method can also be adapted to the discrete-time case. To prove the non-central limit theorems, we use the spectral representation of the underlying process, the properties of Wiener-Itô integrals, and a continuous analog of a method to establish convergence in distribution of quadratic functionals to double Wiener-Itô integrals, developed by Surgailis [1982] (see also Giraitis et al. [2012]).

Limit theorems for quadratic forms of the type (7.5) have been considered by a number of authors, mostly for discrete-time stationary processes (see, e.g., Grenander and Szegő [1958], Fox and Taqqu [1985, 1987], Giraitis and Surgailis [1990], Terrin and Taqqu [1990], Giraitis and Taqqu [1999], Ginovyan and Sahakyan [2005], and references therein). The continuous-time case where $X(t)$ is Gaussian has been mainly considered in Ginovian

[1994], Ginovyan and Sahakyan [2007], and Bai et al. [2015].

To the best of our knowledge, the only work addressing the quadratic functionals of the Lévy-driven continuous-time linear process $X(t)$ is Avram et al. [2010], where a central limit theorem for the quadratic functional (7.4) was stated (without proof) under some L^p -type conditions imposed on the spectral density $f(\lambda)$ of $X(u)$ and the generating function $g(\lambda)$ (see Remark 7.2.6 below). For a related study of the sample covariances of Lévy-driven moving average processes we refer to the recent papers by Cohen and Lindner [2013], and Spangenberg [2015].

In our setting, where the underlying process $X(t)$ is not necessarily Gaussian, additional complications arise due to the contribution of the random diagonal term in the double stochastic integral with respect to Lévy noise, which is not present in the case of Gaussian noise (see Remark 7.2.3 below).

The chapter is organized as follows. In Section 7.2 we state the main results of the chapter. In Section 7.3 we give a number of preliminary results that are used in the proofs of the main results. Sections 7.4 and 7.5 contain the proofs of the main results.

7.2 Main results: central and non-central limit theorems

In this section, we state our main results, involving central and non-central limit theorems for suitably normalized process $Q_T(t)$ given by (7.5) under short and long-range dependence conditions.

Let $\{X(t), t \in \mathbb{R}\}$ be a centered real-valued linear process given by (7.1) with filter $a(\cdot) \in L^2(\mathbb{R})$ and covariance function $r(\cdot)$ given by (7.2).

Throughout the chapter we will use the following notation. The symbol $*$ will stand for the convolution:

$$(h * g)(u) = \int_{\mathbb{R}} h(u - x)g(x)dx,$$

while the symbol $\bar{*}$ will be used to denote the reversed convolution:

$$(h^{\bar{*}2})(u) = (h\bar{*}h)(u) = \int_{\mathbb{R}} h(u+x)h(x)dx.$$

By \mathcal{F} and \mathcal{F}^{-1} we will denote the Fourier and the inverse Fourier transforms:

$$(\mathcal{F}h)(u) = \widehat{h}(u) = \int_{\mathbb{R}} e^{ixu}h(x)dx, \quad (\mathcal{F}^{-1}h)(u) = \frac{1}{2\pi} \int_{\mathbb{R}} e^{-ixu}h(x)dx.$$

We will use the following well-known identities:

$$\mathcal{F}(h * g) = \mathcal{F}(h) \cdot \mathcal{F}(g) \tag{7.7}$$

and

$$\mathcal{F}(h\bar{*}g) = \mathcal{F}(h) \cdot \overline{\mathcal{F}(g)}. \tag{7.8}$$

7.2.1 Central limit theorems

The theorem that follows contains L^p -type sufficient conditions for $Q_T(t)$ to obey central limit theorem, and is proved in Section 7.4.

Theorem 7.2.1. *Let $X(t)$ be as in (7.1), and let $Q_T(t)$ be as in (7.5). Assume that*

$$a(\cdot) \in L^p(\mathbb{R}) \cap L^2(\mathbb{R}), \quad b(\cdot) \in L^q(\mathbb{R}) \tag{7.9}$$

with

$$1 \leq p, q \leq 2, \quad \frac{2}{p} + \frac{1}{q} \geq \frac{5}{2}. \tag{7.10}$$

Then

$$\tilde{Q}_T(t) := \frac{1}{\sqrt{T}} (Q_T(t) - \mathbb{E}Q_T(t)) \xrightarrow{f.d.d.} \sigma B(t), \tag{7.11}$$

where $B(t)$ is a standard Brownian motion, and

$$\sigma^2 = \int_{\mathbb{R}} [2K_A(v) + \kappa_4 K_B(v)] dv, \quad (7.12)$$

where κ_4 is the fourth cumulant of $\xi(1)$, and

$$K_A(v) = \left((a * b)^{\bar{*}2} \cdot a^{\bar{*}2} \right)(v), \quad K_B(v) = \left((a * b) \cdot a \right)^{\bar{*}2}(v). \quad (7.13)$$

Remark 7.2.2. Young's inequality for convolution (see, e.g., Bogachev [2007], Theorem 3.9.4) states that for any numbers p, p_1, q satisfying $1 \leq p \leq p_1 \leq \infty$ and $\frac{1}{p_1} = \frac{1}{p} + \frac{1}{q} - 1$, and for any functions $f \in L^p(\mathbb{R}), g \in L^q(\mathbb{R})$ the function $f * g$ is defined almost everywhere, $f * g \in L^{p_1}(\mathbb{R})$, and one has

$$\|f * g\|_{p_1} \leq \|f\|_p \|g\|_q. \quad (7.14)$$

Applying this inequality to the convolution in (7.2), we get $\|r\|_{p_1} \leq \|a\|_p^2 < \infty$, where $1 + 1/p_1 = 2/p$. Hence the relations (7.9) and (7.10) imply that

$$r(\cdot) \in L^{p_1}(\mathbb{R}), \quad b(\cdot) \in L^q(\mathbb{R}), \quad \frac{1}{p_1} + \frac{1}{q} = \frac{2}{p} - 1 + \frac{1}{q} \geq \frac{5}{2} - 1 = \frac{3}{2}. \quad (7.15)$$

The condition (7.15) is sufficient for the convergence in Theorem 7.2.1 to hold in the case where $\xi(t)$ is Brownian motion (see Theorem 2.2 of Bai et al. [2015]). In fact, in this case, the convergence in Theorem 7.2.1 holds under even a weaker condition imposed on the generating function $g(\lambda)$ and the spectral density $f(\lambda)$ of $X(t)$ (see Theorem 2.1 of Bai et al. [2015]).

Remark 7.2.3. In contrast to the cases where the model is either a discrete-time linear process (Giraitis and Surgailis [1990]), or a continuous-time Gaussian process (Bai et al. [2015]), it is convenient to impose the time-domain conditions (7.9) and (7.10) on the functions $a(\cdot)$ and $b(\cdot)$, instead of on the spectral density $f(\lambda)$ and the generating function $g(\lambda)$. This allows us to analyze the random diagonal term which arises from the double

stochastic integral with respect to a non-Gaussian Lévy process.

In the discrete-time case the random diagonal term is estimated by the full double sum (see, e.g., Giraitis and Surgailis [1990], relation (2.3)), while in the continuous-time Gaussian case, there is no such random diagonal term. In the continuous-time non-Gaussian case, we have a random diagonal term in the form of a single stochastic integral that cannot be controlled by the double integral, and hence we need to treat it separately (see (7.61) in the proof of Theorem 4.6.5).

Remark 7.2.4. Observe that the long-run variance σ^2 given by (7.12) can be expressed in terms of the spectral density $f(\lambda)$ and the generating function $g(\lambda)$, provided that these functions satisfy some regularity conditions. Indeed, using (7.7), (7.8) and Parseval-Plancherel theorem, under suitable integrability conditions on $a(\cdot)$ and $b(\cdot)$, we can write

$$\begin{aligned} \int_{\mathbb{R}} K_A(v)dv &= \int_{\mathbb{R}} (a * b)^{\bar{*}2}(v) a^{\bar{*}2}(v) dv = \frac{1}{2\pi} \int_{\mathbb{R}} \mathcal{F}((a * b)^{\bar{*}2})(\lambda) \overline{\mathcal{F}(a^{\bar{*}2})(\lambda)} d\lambda = \\ &= \frac{1}{2\pi} \int_{\mathbb{R}} |\mathcal{F}(a * b)(\lambda)|^2 |\mathcal{F}(a)(\lambda)|^2 d\lambda = \frac{1}{2\pi} \int_{\mathbb{R}} |\widehat{a}(\lambda) \widehat{b}(\lambda)|^2 |\widehat{a}(\lambda)|^2 d\lambda \\ &= 8\pi^3 \int_{\mathbb{R}} f(\lambda)^2 g(\lambda)^2 d\lambda, \end{aligned}$$

where in the last equality we used the fact $|\widehat{a}|^2 = 2\pi f$ and $\widehat{b} = 2\pi g$ (because $b(\cdot)$ is an even function). Similarly, we have

$$\begin{aligned} \int_{\mathbb{R}} K_B(v)dv &= \int_{\mathbb{R}} dv \int_{\mathbb{R}} dx \left((a * b) \cdot a \right)(x) \left((a * b) \cdot a \right)(x + v) = \left(\int_{\mathbb{R}} (a * b)(x) a(x) dx \right)^2 \\ &= \frac{1}{4\pi^2} \left(\int_{\mathbb{R}} \widehat{a}(x) \widehat{b}(x) \overline{\widehat{a}(x)} d\lambda \right)^2 = 4\pi^2 \left[\int_{\mathbb{R}} f(\lambda) g(\lambda) d\lambda \right]^2. \end{aligned}$$

So an alternative expression for σ^2 in (7.12) is

$$\sigma^2 = 16\pi^3 \int_{\mathbb{R}} f(\lambda)^2 g(\lambda)^2 d\lambda + \kappa_4 \left[2\pi \int_{\mathbb{R}} f(\lambda) g(\lambda) d\lambda \right]^2, \quad (7.16)$$

which should be compared with Avram et al. [2010] (Theorem 4.1), and Giraitis and Surgailis [1990] for an analogous expression in the discrete-time case.

Remark 7.2.5. The discrete-time analog of Theorem 7.2.1 with $t = 1$ and ξ being Gaussian was established in Giraitis and Surgailis [1990]. A special case of Theorem 7.2.1 with $t = 1$ and ξ being Gaussian was established in Ginovian [1994] and Ginovyan and Sahakyan [2007]. Theorem 7.2.1 for Wiener-driven model ($\kappa_4 = 0$) was proved in Bai et al. [2015].

Remark 7.2.6. For Lévy-driven model with $t = 1$ and σ^2 given by (7.16), a version of Theorem 7.2.1 was stated in Avram et al. [2010] (Theorem 4.1). They impose conditions on the spectral density $f(\cdot)$ and the generating function $g(\cdot)$, and assume the existence of all moments of the driving Lévy process $\xi(t)$. The details of the proof of Theorem 4.1 in Avram et al. [2010] is unfortunately omitted. It is not clear, at least to us, how the omitted details of the method-of-moment proof can be carried out given the complexity of computing the moments of multiple integrals with respect to non-Gaussian Lévy noise (see Peccati and Taqqu [2011], Chapter 7).

The following corollary, proved in Section 7.4, contains sufficient conditions for the assumptions in Theorem 7.2.1 to hold.

Corollary 7.2.7. *The convergence in (7.11) holds if the functions $a(\cdot)$ and $b(\cdot)$ satisfy the following conditions:*

$$a(\cdot), b(\cdot) \in L^\infty(\mathbb{R}), \quad |a(x)| \leq c|x|^{\alpha/2-1}, \quad |b(x)| \leq c|x|^{\beta-1} \quad (7.17)$$

with

$$0 < \alpha, \beta < 1, \quad \alpha + \beta < 1/2.$$

7.2.2 Non-central limit theorems

We now state the non-central limit theorems. We make the following assumptions on the functions $a(\cdot)$, $b(\cdot)$ and on their Fourier transforms $\widehat{a}(\cdot)$ and $\widehat{b}(\cdot)$.

Assumption 1. The Fourier transform $\widehat{a}(\cdot)$ of $a(\cdot) \in L^2(\mathbb{R})$ satisfies

$$\widehat{a}(x) = A(x)|x|^{-\alpha/2}L_1^{1/2}(x),$$

where $L_1(x)$ is an even non-negative function slowly varying at zero and bounded on intervals $[c, \infty)$ for any $c > 0$, and $A(x)$ is a complex-valued function satisfying $|A(x)| = 1$, and $\lim_{x \rightarrow 0^+} A(x) = A_0$ for some A_0 on the complex unit circle (since $\widehat{a}(-x) = \overline{\widehat{a}(x)}$, we also have $\lim_{x \rightarrow 0^-} A(x) = \overline{A_0}$).

Assumption 2. The generating function $\widehat{b}(\cdot) \in L^1(\mathbb{R})$ and satisfies

$$\widehat{b}(x) = |x|^{-\beta} L_2(x),$$

where $L_2(x)$ is an even non-negative function slowly varying at zero and bounded on intervals $[c, \infty)$ for any $c > 0$.

Assumption 3. The parameters α and β above satisfy

$$-1/2 < \alpha < 1, \quad -1/2 < \beta < 1, \quad \alpha + \beta > 1/2. \quad (7.18)$$

Assumption 4. There exist numbers α^* and β^* satisfying

$$0 < \alpha^*, \beta^* < 1 \quad 1 < \alpha^* + \beta^* < \alpha + \beta + 1/2,$$

such that

$$|a(x)| \leq C|x|^{\alpha^*/2-1}, \quad |b(x)| \leq C|x|^{\beta^*-1}.$$

The proof of the following theorem can be found in Section 7.5.

Theorem 7.2.8. *Suppose that Assumptions 1 - 4 hold. Then as $T \rightarrow \infty$*

$$\widetilde{Q}_T(t) := \frac{1}{T^{\alpha+\beta} L_1(1/T) L_2(1/T)} (Q_T(t) - \mathbb{E}Q_T(t)) \xrightarrow{f.d.d.} Z_{\alpha,\beta}(t), \quad (7.19)$$

where

$$Z_{\alpha,\beta}(t) = \frac{1}{2\pi} \int_{\mathbb{R}^2} |x_1 x_2|^{-\alpha/2} \int_{\mathbb{R}} \frac{e^{it(x_1+u)} - 1}{i(x_1+u)} \frac{e^{it(x_2-u)} - 1}{i(x_2-u)} |u|^{-\beta} du W(dx_1) W(dx_2), \quad (7.20)$$

where $W(\cdot)$ is a complex-valued Brownian motion, and the double prime " indicates the exclusion of the hyper-diagonals $u_p = \pm u_q$, $p \neq q$.

Remark 7.2.9. The regular variation conditions on $\widehat{a}(\cdot)$ and $\widehat{b}(\cdot)$ in Assumptions 1 - 3 generally do not follow from the corresponding regular variation conditions imposed on the inverse Fourier transforms $a(\cdot)$ and $b(\cdot)$. This implication only holds under some additional assumptions on the slowly varying factors of $a(\cdot)$ and $b(\cdot)$. For instance, it will hold if we have (see Bingham et al. [1989], formula (4.3.7))

$$a(x) = x^{\alpha/2-1} \ell_1(x) 1_{[0, \infty)}(x), \quad b(x) = |x|^{\beta-1} \ell_2(x), \quad (7.21)$$

where $0 < \alpha < 1$, $0 < \beta < 1$, $\alpha + \beta > 1/2$, and $\ell_1(x)$ and $\ell_2(x)$ are even non-negative functions which are locally bounded, slowly varying at infinity and *quasi-monotone*. Recall that a slowly varying function $l(\cdot)$ is said to be quasi-monotone if it has locally bounded variation, and for all $\delta > 0$, one has (see Bingham et al. [1989], Section 2.7)

$$\int_0^x t^\delta |d\ell(t)| = O(x^\delta l(x)) \quad \text{as } x \rightarrow \infty.$$

A sufficient condition for a slowly varying $\ell(x)$ with locally bounded variation to be quasi-monotone is that $x^\delta \ell(x)$ is increasing and $x^{-\delta} \ell(x)$ is decreasing when x is large enough, for any $\delta > 0$ (see Theorem 1.5.5 and Corollary 2.7.4 in Bingham et al. [1989]).

Notice also that Assumption 4 will be satisfied if (7.21) holds (see Lemma 7.5.6).

Remark 7.2.10. Let the functions $a(\cdot)$ and $b(\cdot)$ be as in (7.21) with $\alpha < 0$ or $\beta < 0$ (by (7.18) only one of α and β can be negative). Assume that $\alpha < 0$ and $\beta > 0$. Then for the corresponding regular variation of $\widehat{a}(\cdot)$ to hold, one needs to impose in addition that $\int_0^\infty a(x) dx = 0$. In this case, one does not need to assume quasi-monotonicity for ℓ_1 (see Corollary 1.40 of Soulier [2009]). Similar considerations hold if $\beta < 0$ and $\alpha > 0$ instead.

Remark 7.2.11. Note that Assumption 1 holds with $\alpha = 0$ if $a(\cdot) \in L^1(\mathbb{R})$ and $\int_0^\infty a(x) dx \neq 0$, and Assumption 2 holds with $\beta = 0$ if $b(\cdot) \in L^1(\mathbb{R})$ and $\int_0^\infty b(x) dx \neq 0$.

The next theorem contains time-domain representations for the limiting process $Z_{\alpha,\beta}(t)$ in (7.20) in the case $\alpha, \beta \geq 0$, which will be proved in Section 7.5.

Theorem 7.2.12. *The limiting process $Z_{\alpha,\beta}(t)$ in (7.20) admits the following time-domain representations:*

(a) *when $\alpha > 0, \beta > 0$:*

$$Z_{\alpha,\beta}(t) \stackrel{f.d.d.}{=} c_{\alpha,\beta} \int_{\mathbb{R}^2}' \int_0^t \int_0^t |u-v|^{\beta-1} (u-x_1)_+^{\alpha/2-1} (v-x_2)_+^{\alpha/2-1} dudv B(dx_1)B(dx_2), \quad (7.22)$$

$$\text{where } c_{\alpha,\beta} = \frac{\Gamma(1-\beta)\sin(\beta\pi/2)}{\pi\Gamma(\alpha/2)^2};$$

(b) *when $\alpha > 1/2, \beta = 0$:*

$$Z_{\alpha,\beta}(t) \stackrel{f.d.d.}{=} c_{\alpha} \int_{\mathbb{R}^2}' \int_0^t (u-x_1)_+^{\alpha/2-1} (u-x_2)_+^{\alpha/2-1} du B(dx_1)B(dx_2), \quad (7.23)$$

$$\text{where } c_{\alpha} = \frac{\sin(\alpha\pi/2)\Gamma(1-\alpha/2)}{\pi\Gamma(\alpha/2)};$$

(c) *when $\alpha = 0, \beta > 1/2$:*

$$Z_{\alpha,\beta}(t) \stackrel{f.d.d.}{=} c_{\beta} \int_{[0,t]^2}' |x_1-x_2|^{\beta-1} B(dx_1)B(dx_2), \quad (7.24)$$

where $c_{\beta} = \frac{\Gamma(1-\beta)\sin(\beta\pi/2)}{\pi}$, $B(\cdot)$ is the real Brownian random measure and $'$ indicates the exclusion of the diagonals.

Remark 7.2.13. In view of (7.5) and (7.21), the representation (7.22) gives an explicit insight of the convergence in Theorem 7.2.8 (see Theorem 7.2.14 below). The process in (7.23) is known as Rosenblatt process (see Taquq [1975]), and the corresponding convergence in Theorem 7.2.8 is the continuous-time analog of the discrete-time case considered in Fox and Taquq [1985]. The representation (7.24) is obtained because for $\alpha = 0$, the underlying process $X(t)$ has short memory and in this case, one expects that in the limit $X(t)dt$ in (7.5) can be replaced by the white noise $B(dt)$.

In the cases where either α or β satisfying (7.18) is negative, we were not able to obtain appropriate elementary expressions for the time-domain representation of the limiting process $Z_{\alpha,\beta}(t)$.

Using the time-domain representation (7.22), one can state a non-central limit theorem in the case where $\alpha, \beta > 0$ without going to the spectral domain. This simplifies the assumptions imposed on the functions $a(\cdot)$ and $b(\cdot)$.

Theorem 7.2.14. *Suppose that the functions $a(\cdot)$ and $b(\cdot)$ are given by (7.21), where $0 < \alpha < 1$, $0 < \beta < 1$, $\alpha + \beta > 1/2$, and $\ell_1(x)$ and $\ell_2(x)$ are even functions slowly varying at infinity and bounded on bounded intervals. Then as $T \rightarrow \infty$,*

$$\frac{1}{T^{\alpha+\beta}\ell_1(T)\ell_2(T)} (Q_T(t) - \mathbb{E}Q_T(t)) \xrightarrow{f.d.d.} \int_{\mathbb{R}^2}' \int_0^t \int_0^t |u-v|^{\beta-1} (u-x_1)_+^{\alpha/2-1} (v-x_2)_+^{\alpha/2-1} dudv B(dx_1)B(dx_2).$$

The theorem is proved in Section 7.5.

7.3 Preliminaries

We first introduce the notion of multiple *off-diagonal* (Itô-type) stochastic integral with respect to Lévy noise, called Lévy-Itô multiple stochastic integral, and briefly discuss its properties. All the claims we shall make below can be found in Peccati and Taqqu [2011] and Farré et al. [2010]. Let f be a function in $L^2(\mathbb{R}^k)$. Then we can define the following off-diagonal multiple stochastic integral:

$$I_k^\xi(f) = \int_{\mathbb{R}^k}' f(x_1, \dots, x_k) \xi(dx_1) \dots \xi(dx_k), \quad (7.25)$$

where $\xi(t)$ is a Lévy process with $\mathbb{E}\xi(t) = 0$ and $\text{Var}[\xi(t)] = \sigma_\xi^2 t$, and the prime $'$ indicates that we do not integrate on the diagonals $x_i = x_j$, $i \neq j$. Indeed, the integral $I_k^\xi(f)$ can be first defined for $f = 1_{A_1 \times \dots \times A_k}$, where A_1, \dots, A_k are disjoint Borel sets, as $I_k^\xi(f) =$

$\xi(A_1) \dots \xi(A_k)$, and then using linearity and L^2 -approximation to define for general $f \in L^2(\mathbb{R}^k)$. The multiple integral $I_k^\xi(\cdot)$ satisfies

$$\|I_k^\xi(f)\|_{L^2(\Omega)}^2 \leq k! \sigma_\xi^{2k} \|f\|_{L^2(\mathbb{R}^k)}^2. \quad (7.26)$$

The inequality in (7.26) becomes equality if f is symmetric:

$$\|I_k^\xi(f)\|_{L^2(\Omega)}^2 = k! \sigma_\xi^{2k} \|f\|_{L^2(\mathbb{R}^k)}^2. \quad (7.27)$$

As before $B(\cdot)$ will stand for the real-valued Brownian motion. Setting $\xi(\cdot) = B(\cdot)$, we get the so-called multiple Wiener-Itô integral (see Itô [1951]):

$$I_k^B(f) = \int_{\mathbb{R}^k}' f(x_1, \dots, x_k) B(dx_1) \dots B(dx_k). \quad (7.28)$$

The Wiener-Itô integral can also be defined with respect to the complex-valued Brownian motion:

$$I_k^W(g) = \int_{\mathbb{R}^k}'' g(u_1, \dots, u_k) W(du_1) \dots W(du_k), \quad (7.29)$$

where $g \in L^2(\mathbb{R}^k)$ is a complex-valued function satisfying $g(-u_1, \dots, -u_k) = \overline{g(u_1, \dots, u_k)}$, and $W(\cdot)$ is a complex-valued Brownian motion (with real and imaginary parts being independent) viewed as a random integrator (see, e.g., Embrechts and Maejima [2002], p.22), and the double prime $''$ indicates the exclusion of the hyper-diagonals $u_p = \pm u_q$, $p \neq q$.

The next result, which can be deduced from Proposition 9.3.1 of Peccati and Taqqu [2011] and Proposition 4.2 of Dobrushin [1979], gives a relationship between the integrals $I_k^B(\cdot)$ and $I_k^W(\cdot)$, defined by (7.28) and (7.29), respectively.

Proposition 7.3.1. *Let $f_j(\cdot)$ be real-valued functions in $L^2(\mathbb{R}^{k_j})$, $j = 1, \dots, J$, and let*

$$\widehat{f}_j(w_1, \dots, w_{k_j}) = \int_{\mathbb{R}^{k_j}} f_j(x_1, \dots, x_{k_j}) e^{i(x_1 w_1 + \dots + x_{k_j} w_{k_j})} dx_1 \dots dx_{k_j}$$

be the L^2 -Fourier transform of $f_j(\cdot)$. Then

$$\left(I_{k_1}^B(f_1), \dots, I_{k_J}^B(f_J) \right) \stackrel{d}{=} \left((2\pi)^{-k_1/2} I_{k_1}^W \left(\widehat{f}_1 A^{\otimes k_1} \right), \dots, (2\pi)^{-k_J/2} I_{k_J}^W \left(\widehat{f}_J A^{\otimes k_J} \right) \right),$$

for any function $A(u) : \mathbb{R} \rightarrow \mathbb{C}$ such that $|A(u)| = 1$ and $A(w) = \overline{A(-w)}$ a.e., where $A^{\otimes k}(w_1, \dots, w_k) := A(w_1) \cdots A(w_k)$.

We also will need a stochastic Fubini's theorem (see Peccati and Taqqu [2011], Theorem 5.12.1).

Lemma 7.3.2. *Let (S, μ) be a measure space with $\mu(S) < \infty$, and let $f(s, x_1, \dots, x_k)$ be a function on $S \times \mathbb{R}^k$ such that*

$$\int_S \int_{\mathbb{R}^k} f(s, x_1, \dots, x_k)^2 dx_1 \dots dx_k \mu(ds) < \infty,$$

then we can change the order of the multiple stochastic integration $I_k^\xi(\cdot)$ and the deterministic integration $\int_S f(s, \cdot) \mu(ds)$:

$$\int_S I_k^\xi(f(s, \cdot)) \mu(ds) = I_k^\xi \left(\int_S f(s, \cdot) \mu(ds) \right).$$

There is a *with-diagonal* (Stratonovich-type) counterpart of the integral $I_k^\xi(f)$, denoted

$$\mathring{I}_k^\xi(f) = \int_{\mathbb{R}^k} f(x_1, \dots, x_k) \xi(dx_1) \dots \xi(dx_k), \quad (7.30)$$

which includes all the diagonals. We refer to Farré et al. [2010] for a comprehensive treatment of Stratonovich-type integrals $\mathring{I}_k^\xi(f)$. For the with-diagonal integral $\mathring{I}_k^\xi(f)$ to be well-defined, the integrand f needs also to be square-integrable on all the diagonals of \mathbb{R}^k . More precisely, it is required that $f \in L^2(\Lambda_n)$, with $\Lambda_n = \sum_{\sigma \in \Pi_n} \lambda_\sigma$, where Π_n denotes all the partitions of $\{1, \dots, n\}$, and λ_σ denotes the Lebesgue measure on the diagonals specified by the partition σ , provided that the variables in the same block of σ are identified. For example, if $\sigma = \{\{1, 2\}, \{3\}\}$, then λ_σ is the two-dimensional Lebesgue

measure on $\{x_1 = x_2, x_3\}$, and

$$\|f\|_{L^2(\lambda_\sigma)}^2 = \int_{\mathbb{R}^3} f^2(x_1, x_2, x_3) d\lambda_\sigma(x_1, x_2, x_3) = \int_{\mathbb{R}^2} f^2(x_1, x_1, x_3) dx_1 dx_3.$$

For with-diagonal integrals, we have the following simple product formula:

$$\mathring{I}_p^\xi(f) \mathring{I}_q^\xi(g) = \mathring{I}_{p+q}^\xi(f \otimes g).$$

The with-diagonal integral $\mathring{I}_k^\xi(f)$ can be expressed by off-diagonal integrals of lower orders using the Hu-Meyer formula (see Farré et al. [2010], Theorem 5.9). We shall only use the special case when $k = 2$, in which case we have

$$\mathring{I}_2^\xi(f) = \int_{\mathbb{R}^2}' f(x_1, x_2) \xi(dx_1) \xi(dx_2) + \int_{\mathbb{R}} f(x, x) \xi_c^{(2)}(dx) + \int_{\mathbb{R}} f(x, x) dx, \quad (7.31)$$

where

$$\xi_c^{(2)}(t) = \xi^{(2)}(t) - \mathbb{E}\xi^{(2)}(t) = \xi^{(2)}(t) - |t| \quad (7.32)$$

and $\xi^{(2)}(t)$ is the quadratic variation of $\xi(t)$, which is non-deterministic if $\xi(t)$ is non-Gaussian (see Farré et al. [2010], equation (10)). The centered process $\xi_c^{(2)}(t)$ is called a *Teugels martingale* (of second order), which is a Lévy process with the same filtration as $\xi(t)$, whose quadratic variation is deterministic:

$$[\xi_c^{(2)}(t), \xi_c^{(2)}(t)] = \kappa_4 t,$$

where κ_4 is the fourth cumulant of $\xi(1)$. For any $f, g \in L^2(\mathbb{R})$, one has (see Farré et al. [2010], page 9),

$$\mathbb{E} \left[\int_{\mathbb{R}} g(x) \xi_c^{(2)}(dx) \int_{\mathbb{R}} h(x) \xi_c^{(2)}(dx) \right] = \kappa_4 \int_{\mathbb{R}} f(x) g(x) dx. \quad (7.33)$$

The decomposition (7.31) implies that

$$EI_k^{\circ\xi}(f) = \int_{\mathbb{R}} f(x, x) dx.$$

Consider now the following integrals, the first of which is an off-diagonal double integral and the second is a single integral with respect to Teugels martingale $\xi_c^{(2)}(t)$:

$$\int_{\mathbb{R}^2}' f(x_1, x_2) \xi(dx_1) \xi(dx_2) \quad \text{and} \quad \int_{\mathbb{R}} g(x) \xi_c^{(2)}(dx). \quad (7.34)$$

Notice that for any $f \in L^2(\mathbb{R}^2)$ and $g \in L^2(\mathbb{R})$ the integrals in (7.34) are uncorrelated. This can easily be verified in the case $f = 1_{A \times B}$, $g = 1_C$ for any disjoint Borel sets A and B and any Borel set C . Indeed, treating $\xi_c^{(2)}(\cdot)$ as a random measure, we have

$$\begin{aligned} & E[\xi(A)\xi(B)\xi_c^{(2)}(C)] = \\ & E \left[\xi(A)\xi(B) \left(\xi_c^{(2)}(C \cap A^c \cap B) + \xi_c^{(2)}(C \cap A \cap B^c) + \xi_c^{(2)}(C \cap A^c \cap B^c) \right) \right] = 0 \end{aligned} \quad (7.35)$$

since, for example, $\xi(A)$ is independent of $\xi_c^{(2)}(C \cap A^c \cap B)$ and $\xi(B)$, and $E\xi(A) = 0$. Using linearity and L^2 -approximation, it can easily be shown that the integrals in (7.34) are uncorrelated for any $f \in L^2(\mathbb{R}^2)$ and $g \in L^2(\mathbb{R})$.

7.4 Proof of the central limit theorems

In this section, we prove the central limit theorems stated in Section 6.2 (Theorem 7.2.1 and Corollary 7.2.7). We first derive some preliminary results. We set

$$R_T(x_1, x_2) = \frac{1}{\sqrt{T}} \int_0^T \int_0^T b(u-v) a(u-x_1) a(v-x_2) du dv, \quad (7.36)$$

and

$$S_T(x_1, x_2) = \frac{1}{\sqrt{T}} \int_0^T [(a * b)(v-x_1)] [a(v-x_2)] dv. \quad (7.37)$$

Lemma 7.4.1. *Let $a(\cdot)$ and $b(\cdot)$ satisfy (7.9) and (7.10), and let $R_T(x_1, x_2)$ and $S_T(x_1, x_2)$ be as in (7.36) and (7.37) with $x_1 \neq x_2$. The following assertions hold.*

(a) *We have*

$$\lim_{T \rightarrow \infty} \|S_T\|_{L^2(\mathbb{R}^2)}^2 = \int_{\mathbb{R}} K_A(u) du, \quad (7.38)$$

where $K_A(\cdot)$ is as in (7.13).

(b) *We have*

$$\lim_{T \rightarrow \infty} \|R_T - S_T\|_{L^2(\mathbb{R}^2)} = 0. \quad (7.39)$$

(c) *For any $M > 0$, there exists a function $c_M(\cdot, \cdot)$ supported on $[-2M, 2M]^2$, so that the function*

$$S_T^M(x_1, x_2) = \frac{1}{\sqrt{T}} \int_0^T c_M(v - x_1, v - x_2) dv,$$

satisfies the relation:

$$\lim_{M \rightarrow \infty} \limsup_{T \rightarrow \infty} \|R_T - S_T^M\|_{L^2(\mathbb{R}^2)} = 0. \quad (7.40)$$

Proof of Lemma 7.4.1. We first prove assertion (a). We will use the following notation: $\|\cdot\|_r$ will denote the $L^r(\mathbb{R})$ norm, and $|a|(x) = |a(x)|$, $|b|(x) = |b(x)|$, $|c|(x) = |c(x)|$.

By (7.9) and (7.10) we have $a(\cdot) \in L^p(\mathbb{R}) \cap L^2(\mathbb{R})$. Hence by the Riesz-Thorin theorem, $a(\cdot) \in L^{p'}(\mathbb{R})$ for any $p \leq p' \leq 2$. Setting $p' = 2$, we get $1 + 1/q \leq 2$, which is less than $5/2$. This implies that there is a number p' such that $2/p' + 1/q = 5/2$. Thus, without loss of generality, we can assume that

$$a(\cdot) \in L^p(\mathbb{R}), \quad b \in L^q(\mathbb{R}), \quad \frac{2}{p} + \frac{1}{q} = \frac{5}{2}. \quad (7.41)$$

Let p and q be as in (7.41). Define the numbers q_1, q_1^*, q_2 to satisfy the following equations:

$$\frac{1}{q_1} + \frac{1}{q_1^*} = 1, \quad 1 + \frac{1}{q_1^*} = \frac{2}{p}, \quad 1 + \frac{1}{q_1} = \frac{2}{q_2}, \quad 1 + \frac{1}{q_2} = \frac{1}{p} + \frac{1}{q}. \quad (7.42)$$

(Going from the last to the first equality in (7.42), one can solve successively for q_2 , q_1^* , q_1 and then verify using (7.41) that the first equality in (7.42) holds.)

Taking into account (7.42), the relation

$$\begin{aligned} |R_T(x_1, x_2)| &\leq \frac{1}{\sqrt{T}} \int_0^T \int_{-\infty}^{+\infty} |b(u-v)a(u-x_1)a(v-x_2)| dudv \\ &= \frac{1}{\sqrt{T}} \int_0^T (|a| * |b|)(v-x_1)|a|(v-x_2)dv, \end{aligned} \quad (7.43)$$

and by using Hölder's inequality and Young's inequality for convolution (see (7.14)), we can write

$$\begin{aligned} &\|R_T\|_{L^2(\mathbb{R}^2)}^2 \\ &\leq \frac{1}{T} \int_{[0,T]^2} dv_1 dv_2 \int_{\mathbb{R}^2} dx_1 dx_2 (|a| * |b|)(v_1 - x_1)|a|(v_1 - x_2)(|a| * |b|)(v_2 - x_1)|a|(v_2 - x_2) \\ &= \frac{1}{T} \int_{[0,T]^2} dv_1 dv_2 \left((|a| * |b|)^{\bar{*}2} \cdot |a|^{\bar{*}2} \right) (v_1 - v_2) \\ &= \int_{-T}^T \left(1 - \frac{|v|}{T} \right) \left((|a| * |b|)^{\bar{*}2} \cdot |a|^{\bar{*}2} \right) (v) dv \leq \int_{\mathbb{R}} \left((|a| * |b|)^{\bar{*}2} \cdot |a|^{\bar{*}2} \right) (v) dv \\ &\stackrel{\text{Hölder}}{\leq} \|(|a| * |b|)^{\bar{*}2}\|_{q_1} \| |a|^{\bar{*}2} \|_{q_1^*} \stackrel{\text{Young}}{\leq} \|(|a| * |b|)^{\bar{*}2}\|_{q_1} \|a\|_p^2 \\ &\stackrel{\text{Young}}{\leq} \| |a| * |b| \|_{q_2}^2 \|a\|_p^2 \stackrel{\text{Young}}{\leq} \|a\|_p^4 \|b\|_q^2. \end{aligned} \quad (7.44)$$

Similarly, we get

$$\|S_T\|_{L^2(\mathbb{R}^2)}^2 \leq \|a\|_p^4 \|b\|_q^2. \quad (7.46)$$

In view of (7.37), (7.46) and Fubini's theorem, we obtain

$$\|S_T\|_{L^2(\mathbb{R}^2)}^2 = \int_{-T}^T \left(1 - \frac{|v|}{T} \right) \left((a * b)^{\bar{*}2} \cdot a^{\bar{*}2} \right) (v) dv,$$

which converges to the limit claimed in (7.38) by the dominated convergence theorem.

Now we proceed to prove assertions (b) and (c).

To this end, for $M > 0$ we set $a_M(x) = a(x)1_{[-M,M]}(x)$, $a_M^-(x) = a(x) - a_M(x)$,

$b_M(x) = b(x)1_{[-M,M]}(x)$ and $b_M^-(x) = b(x) - b_M(x)$, and define

$$R_T^M(x_1, x_2) = \frac{1}{\sqrt{T}} \int_0^T \int_0^T b_M(u-v)a_M(u-x_1)a_M(v-x_2)dudv. \quad (7.47)$$

In view of (7.36), (7.47) and the identity

$$\begin{aligned} baa - b_M a_M a_M &= (baa - b_M a a) + (b_M a a - b_M a_M a) + (b_M a_M a - b_M a_M a_M) \\ &= b_M^- a a + b_M a_M^- a + b_M a_M a_M^-, \end{aligned}$$

we have

$$\begin{aligned} R_T(x_1, x_2) - R_T^M(x_1, x_2) &= \\ \frac{1}{\sqrt{T}} \int_0^T \int_0^T dudv & [b_M^-(u-v)a(u-x_1)a(v-x_2) + \\ b_M(u-v)a_M^-(u-x_1)a(v-x_2) &+ b_M(u-v)a_M(u-x_1)a_M^-(v-x_2)]. \end{aligned}$$

Similar to (7.45), one gets

$$\|R_T - R_T^M\|_{L^2(\mathbb{R}^2)}^2 \leq C(\|b_M^-\|_q^2 \|a\|_p^4 + \|b_M\|_q^2 \|a_M^-\|_p^2 \|a\|_p^2 + \|b_M\|_q^2 \|a_M\|_p^2 \|a_M^-\|_p^2),$$

where the right-hand side does not involve T . Since $\|a_M^-\|_p \rightarrow 0$ and $\|b_M^-\|_q \rightarrow 0$ as $M \rightarrow \infty$, one obtains

$$\lim_{M \rightarrow \infty} \limsup_{T \rightarrow \infty} \|R_T - R_T^M\|_{L^2(\mathbb{R}^2)} = 0. \quad (7.48)$$

Now we set

$$c_M(x_1, x_2) = (a_M * b_M)(x_1)a_M(x_2), \quad (7.49)$$

and define

$$S_T^M(x_1, x_2) = \frac{1}{\sqrt{T}} \int_0^T c_M(v-x_1, v-x_2)du$$

$$= \frac{1}{\sqrt{T}} \int_0^T (a_M * b_M)(v - x_1) a_M(v - x_2) dv.$$

In the same way as we derived (7.48), we have

$$\lim_{M \rightarrow \infty} \limsup_{T \rightarrow \infty} \|S_T - S_T^M\|_{L^2(\mathbb{R}^2)} = 0. \quad (7.50)$$

Observe that

$$S_T^M(x_1, x_2) = \frac{1}{\sqrt{T}} \int_0^T dv \left(\int_{\mathbb{R}} dub_M(v - u) a_M(u - x_1) \right) a_M(v - x_2). \quad (7.51)$$

Suppose that $T > M$. In view of (7.47) and (7.51) and using the fact that $b_M(\cdot)$ is supported on $[-M, M]$, we have

$$\begin{aligned} & S_T^M(x_1, x_2) - R_T^M(x_1, x_2) \\ &= \frac{1}{\sqrt{T}} \int_0^T dv \int_{\mathbb{R} \setminus [0, T]} dub_M(u - v) a_M(u - x_1) a_M(v - x_2) \\ &= \frac{1}{\sqrt{T}} \int_0^T dv \int_T^\infty dub_M(u - v) a_M(u - x_1) a_M(v - x_2) \\ &\quad + \frac{1}{\sqrt{T}} \int_0^T dv \int_{-\infty}^0 dub_M(u - v) a_M(u - x_1) a_M(v - x_2) \\ &= \frac{1}{\sqrt{T}} \int_{T-M}^T dv \int_T^\infty dub_M(u - v) a_M(u - x_1) a_M(v - x_2) \\ &\quad + \frac{1}{\sqrt{T}} \int_0^M dv \int_{-M}^0 dub_M(u - v) a_M(u - x_1) a_M(v - x_2) \\ &=: A_{T,1}^M(x_1, x_2) + A_{T,2}^M(x_1, x_2). \end{aligned}$$

Thus, using the arguments similar to those in (7.43) and (7.45), one has

$$\begin{aligned} \|A_{T,1}^M\|_{L^2(\mathbb{R}^2)}^2 &\leq \frac{1}{T} \int_{[T-M, T]^2} dv_1 dv_2 \left((|a| * |b|)^{\bar{*}2} \cdot |a|^{\bar{*}2} \right) (v_1 - v_2) \\ &= \frac{1}{T} \int_{[0, M]^2} dv_1 dv_2 \left((|a| * |b|)^{\bar{*}2} \cdot |a|^{\bar{*}2} \right) (v_1 - v_2) \\ &\leq \frac{M}{T} \int_{\mathbb{R}} dv \left((|a| * |b|)^{\bar{*}2} \cdot |a|^{\bar{*}2} \right) (v) \rightarrow 0 \text{ as } T \rightarrow \infty, \end{aligned}$$

where $\int_{\mathbb{R}} dv \left((|a| * |b|)^{\bar{*}2} \cdot |a|^{\bar{*}2} \right) (v)$ is finite due to (7.45). Similarly, one can show that

$$\|A_{T,2}^M\|_{L^2(\mathbb{R}^2)}^2 \rightarrow 0 \text{ as } T \rightarrow \infty.$$

Hence

$$\lim_{M \rightarrow \infty} \limsup_{T \rightarrow \infty} \|S_T^M - R_T^M\|_{L^2(\mathbb{R}^2)}^2 = 0. \quad (7.52)$$

Combining (7.48) (7.50) and (7.52), we obtain the desired relations (7.39) and (7.40) with $c_M(\cdot, \cdot)$ as in (7.49). This completes the proof of Lemma 7.4.1. \square

The next result is similar to Lemma 7.4.1, where \mathbb{R}^2 is replaced by \mathbb{R} . We set

$$R_T(x) = R_T(x, x) = \frac{1}{\sqrt{T}} \int_0^T \int_0^T b(u-v)a(u-x)a(v-x)dudv$$

and

$$S_T(x) = S_T(x, x) = \frac{1}{\sqrt{T}} \int_0^T (a * b)(v-x)a(v-x)dv$$

where $R_T(\cdot, \cdot)$ and $S_T(\cdot, \cdot)$ are as in (7.36) and (7.37).

Lemma 7.4.2. *Assume that $a(\cdot)$ and $b(\cdot)$ be as in (7.9), with p and q satisfying*

$$1 \leq p, q \leq 2, \quad \frac{2}{p} + \frac{1}{q} \geq 2. \quad (7.53)$$

Then the following assertions hold.

(a) *We have*

$$\lim_{T \rightarrow \infty} \|S_T\|_{L^2(\mathbb{R})}^2 = \int_{\mathbb{R}} K_B(u)du, \quad (7.54)$$

where $K_B(\cdot)$ is as in (7.13).

(b) *We have*

$$\lim_{T \rightarrow \infty} \|R_T - S_T\|_{L^2(\mathbb{R})} = 0. \quad (7.55)$$

(c) For any $M > 0$, there exists a function $d_M(\cdot)$ supported on $[-2M, 2M]$, so that the function

$$S_T^M(x) = \frac{1}{\sqrt{T}} \int_0^T d_M(v-x) dv,$$

satisfies the relation:

$$\lim_{M \rightarrow \infty} \limsup_{T \rightarrow \infty} \|R_T - S_T^M\|_{L^2(\mathbb{R})} = 0. \quad (7.56)$$

Remark 7.4.3. Obviously the condition (7.53) is implied by condition (7.10).

Proof of Lemma 7.4.2. The proof is similar to that of Lemma 7.4.1. We thus outline the key steps of the proof omitting the details.

As in the proof of Lemma 7.4.1, in view of the Riesz-Thorin theorem one can assume that

$$a(\cdot) \in L^p(\mathbb{R}), \quad b(\cdot) \in L^q(\mathbb{R}), \quad \frac{2}{p} + \frac{1}{q} = 2. \quad (7.57)$$

Let p and q be as in (7.57). Define the number p^* to satisfy the following equations:

$$\frac{1}{p} + \frac{1}{p^*} = 1, \quad 1 + \frac{1}{p^*} = \frac{1}{p} + \frac{1}{q}.$$

Observe that by the equality in (7.57), one has

$$\frac{1}{p} + \frac{1}{p^*} = \frac{1}{p} + \frac{1}{p} + \frac{1}{q} - 1 = 2 - 1 = 1.$$

Then using Hölder's inequality and Young's inequality for convolution (see (7.14)), we can write (note the difference between (7.44) and (7.58))

$$\begin{aligned} & \|R_T\|_{L^2(\mathbb{R})}^2 \\ & \leq \frac{1}{T} \int_{[0, T]^2} dv_1 dv_2 \int_{\mathbb{R}} dx (|a| * |b|)(v_1 - x) |a|(v_1 - x) (|a| * |b|)(v_2 - x) |a|(v_2 - x) \\ & = \frac{1}{T} \int_{[0, T]^2} dv_1 dv_2 \left((|a| * |b|) \cdot |a| \right)^{\bar{*}2}(v_1 - v_2) \\ & = \int_{-T}^T \left(1 - \frac{|v|}{T} \right) \left((|a| * |b|) \cdot |a| \right)^{\bar{*}2}(v) dv \leq \int_{\mathbb{R}} \left((|a| * |b|) \cdot |a| \right)^{\bar{*}2}(v) dv \end{aligned} \quad (7.58)$$

$$\stackrel{\text{Young}}{\leq} \|(|a| * |b|) \cdot |a|\|_1^2 \stackrel{\text{H\"older}}{\leq} \| |a| * |b| \|_{p^*}^2 \|a\|_p^2 \stackrel{\text{Young}}{\leq} \|a\|_p^4 \|b\|_q^2. \quad (7.59)$$

Similarly, we get

$$\|S_T\|_{L^2(\mathbb{R}^2)}^2 \leq \|a\|_p^4 \|b\|_q^2. \quad (7.60)$$

Then the assertion (a) of the lemma follows from (7.60), Fubini's theorem and dominated convergence theorem.

To prove assertions (b) and (c), we set

$$a_M(x) = a(x)1_{[-M,M]}(x)$$

and

$$b_M(x) = b(x)1_{[-M,M]}(x),$$

and consider the functions

$$R_T^M(x) = \frac{1}{\sqrt{T}} \int_0^T \int_0^T b_M(u-v)a_M(u-x)a_M(v-x)dudv,$$

and

$$S_T^M(x) = \frac{1}{\sqrt{T}} \int_0^T d_M(v-x)dv, \quad \text{where } d_M(x) = \left((a_M * b_M) \cdot a_M \right)(x).$$

Then using the arguments of the proof of Lemma 7.4.1 but now with $x_1 = x_2 = x$, it can be shown that

$$\lim_{M \rightarrow \infty} \limsup_{T \rightarrow \infty} \|R_T^M - R_T\|_{L^2(\mathbb{R})}^2 = 0,$$

$$\lim_{M \rightarrow \infty} \limsup_{T \rightarrow \infty} \|S_T^M - R_T^M\|_{L^2(\mathbb{R})}^2 = 0,$$

$$\lim_{M \rightarrow \infty} \limsup_{T \rightarrow \infty} \|S_T^M - S_T\|_{L^2(\mathbb{R})}^2 = 0.$$

Lemma 7.4.2 is proved. □

Proof of Theorem 7.2.1. By (7.31) and Lemma 7.3.2 one can write

$$\tilde{Q}_T(t) = A_T(t) + B_T(t),$$

where

$$A_T(t) = \int_{\mathbb{R}^2}' \frac{1}{\sqrt{T}} \int_0^{Tt} \int_0^{Tt} b(u-v)a(u-x_1)a(v-x_2)dudv \xi(dx_1)\xi(dx_2),$$

and

$$B_T(t) = \int_{\mathbb{R}} \frac{1}{\sqrt{T}} \int_0^{Tt} \int_0^{Tt} b(u-v)a(u-x)a(v-x)dudv \xi_c^{(2)}(dx). \quad (7.61)$$

Choosing $c_M(x_1, x_2)$ as in Lemma 7.4.1 and setting

$$A_T^M(t) = \int_{\mathbb{R}^2}' \frac{1}{\sqrt{T}} \int_0^{Tt} c_M(u-x_1, u-x_2)du \xi(dx_1)\xi(dx_2), \quad (7.62)$$

one has by (7.27) and relation (7.40) of Lemma 7.4.1 that

$$\lim_{M \rightarrow \infty} \limsup_{T \rightarrow \infty} \mathbb{E}|A_T(t) - A_T^M(t)|^2 = 0, \quad \forall t > 0. \quad (7.63)$$

Choosing $d_M(x)$ as in Lemma 7.4.2 and setting

$$B_T^M(t) = \int_{\mathbb{R}} \frac{1}{\sqrt{T}} \int_0^{Tt} d_M(u-x)du \xi_c^{(2)}(dx), \quad (7.64)$$

one has by (7.27) and relation (7.56) of Lemma 7.4.2 that

$$\lim_{M \rightarrow \infty} \limsup_{T \rightarrow \infty} \mathbb{E}|B_T(t) - B_T^M(t)|^2 = 0, \quad \forall t > 0. \quad (7.65)$$

To complete the proof of the theorem, in view of (7.63) and (7.65), it is enough to show that as $T \rightarrow \infty$,

$$\tilde{Q}_T^M(t) := A_T^M(t) + B_T^M(t) \xrightarrow{f.d.d.} \sigma_M B(t) \quad (7.66)$$

with $\sigma_M \geq 0$ satisfying

$$\lim_{M \rightarrow \infty} \sigma_M^2 = \lim_{T \rightarrow \infty} \text{Var}[A_T(1) + B_T(1)] = \sigma^2. \quad (7.67)$$

To this end, observe first that by the stochastic Fubini Lemma 7.3.2, one has

$$\tilde{Q}_T^M(t) = \frac{1}{\sqrt{T}} \int_0^{Tt} Y_M(u) du,$$

where

$$Y_M(u) = \int_{\mathbb{R}^2}' c_M(u - x_1, u - x_2) \xi(dx_1) \xi(dx_2) + \int_{\mathbb{R}} d_M(u - x) \xi_c^{(2)}(dx),$$

and $\xi_c^{(2)}(\cdot)$ is the Teugel martingale defined in (7.32). Note that $Y_M(u)$ is independent of the σ -field generated by $\{\xi(s) : s < u - 2M, s > u + 2M\}$ since $c_M(\cdot, \cdot)$ vanishes outside $[-2M, 2M]^2$ and $d_M(\cdot)$ vanishes outside $[-2M, 2M]$, implying that $Y_M(u)$ is a stationary $4M$ -dependent process. Then the convergence in (7.66) can be deduced from a classical central limit theorem for M -dependent processes by combining the discretization argument in the proof of Theorem 18.7.1 of Ibragimov and Linnik [1971] and Theorem 5.2 of Billingsley [1956].

To show (7.67), it is enough to note that by the arguments before (7.35), the random variables $A_T(1)$ and $B_T(1)$ are uncorrelated. Hence by (7.27) and (7.38) with $k = 2$, we have

$$\text{Var}[A_T(1)] \rightarrow 2 \int_{\mathbb{R}} K_A(u) du,$$

and by (7.27), (7.54) and (7.33) we obtain

$$\text{Var}[B_T(1)] \rightarrow \kappa_4 \int_{\mathbb{R}} K_B(u) du.$$

This completes the proof of Theorem 7.2.1. \square

Proof of Corollary 7.2.7. In view of Theorem 7.2.1, it is enough to verify that the condi-

tions (7.9) and (7.10) are satisfied. First, noting that by assumptions $0 < \alpha, \beta < 1$ and $\alpha + \beta < 1/2$, we can choose $1 \leq p, q \leq 2$ to satisfy

$$p(\alpha/2 - 1) < -1, \quad q(\beta - 1) < -1 \iff \frac{2}{p} < 2 - \alpha, \quad \frac{1}{q} < 1 - \beta, \quad (7.68)$$

implying that

$$\frac{2}{p} + \frac{1}{q} < 3 - \alpha - \beta. \quad (7.69)$$

Next, since $\alpha + \beta < 1/2$, we have $3 - \alpha - \beta > \frac{5}{2}$, and hence in view of (7.69) the numbers p and q can be chosen to satisfy $2/p + 1/q \geq 5/2$. Thus (7.10) is satisfied.

It is easy to see that with the p, q chosen above, in view of (7.17), we have

$$a(\cdot) \in L^p(\mathbb{R}) \cap L^2(\mathbb{R}), \quad b(\cdot) \in L^q(\mathbb{R}),$$

and thus (7.9) is satisfied. □

7.5 Proof of the non-central limit theorems

In this section we prove the non-central limit theorems stated in Section 6.2 (Theorems 7.2.8-7.2.14).

We first state and prove some preliminary lemmas. The following lemma, which is a continuous analog of Propositions 14.3.2 and 14.3.3 of Giraitis et al. [2012], plays a key role in our proofs. It provides conditions for Lévy-Itô multiple stochastic integrals to converge in distribution to Wiener-Itô multiple stochastic integrals.

Lemma 7.5.1. *For $T > 0$ and $f_{j,T}(\cdot) \in L^2(\mathbb{R}^{k_j})$, $j = 1, \dots, J$, we set*

$$h_{j,T}(x_1, \dots, x_{k_j}) := T^{k_j/2} f_{j,T}(Tx_1, \dots, Tx_{k_j}), \quad (7.70)$$

and assume that there exist $f_j \in L^2(\mathbb{R}^{k_j})$ such that as $T \rightarrow \infty$

$$\|h_{j,T} - f_j\|_{L^2(\mathbb{R}^{k_j})} \rightarrow 0, \quad j = 1, \dots, k_j. \quad (7.71)$$

Then for any Lévy process $\xi(\cdot)$ with $E\xi(1) = 0$ and $E\xi^2(1) = 1$, we have the following joint convergence in distribution:

$$\left(I_{k_1}^\xi(f_{1,T}), \dots, I_{k_J}^\xi(f_{J,T}) \right) \xrightarrow{d} \left(I_{k_1}^B(f_1), \dots, I_{k_J}^B(f_J) \right). \quad (7.72)$$

Proof. For simplicity, we prove the result in the case where $J = 1$ and we will drop the index j . In this case the proof is similar to that of Proposition 14.3.2 of Giraitis et al. [2012]. The general case $J > 1$, which corresponds to Proposition 14.3.3 of Giraitis et al. [2012], can be obtained by similar arguments using the Cramér-Wold Device.

Let $S_M(\mathbb{R}^k)$, $M \in \mathbb{Z}_+$, be the class of functions that are piecewise constant on the $1/M$ -grid of $[-M, M]^k$ (each piece of the grid has linear length $1/M$), and vanishing on the diagonals. Set $\mathcal{S}_k = \cup_{M=1}^\infty S_M(\mathbb{R}^k)$, and observe that \mathcal{S}_k is a dense subset of $L^2(\mathbb{R}^k)$. Then in view of (7.26), for any $\epsilon > 0$, there exists $f_\epsilon \in \mathcal{S}_k$ such that

$$\mathbb{E}|I_k^B(f) - I_k^B(f_\epsilon)|^2 \leq k! \|f - f_\epsilon\|_{L^2(\mathbb{R}^k)}^2 \leq \epsilon. \quad (7.73)$$

Define

$$f_{\epsilon,T}(x_1, \dots, x_k) = T^{-k/2} f_\epsilon(x_1/T, \dots, x_k/T), \quad (7.74)$$

and note that

$$\|h_T - f_\epsilon\|_{L^2(\mathbb{R}^k)}^2 \leq 2\|h_T - f\|_{L^2(\mathbb{R}^k)}^2 + 2\|f - f_\epsilon\|_{L^2(\mathbb{R}^k)}^2. \quad (7.75)$$

By (7.71) we have $\lim_{T \rightarrow \infty} \|h_T - f\|_{L^2(\mathbb{R}^k)} = 0$. Hence in view of (7.26), (7.75) and a change of variable, we can write

$$\limsup_{T \rightarrow \infty} \mathbb{E}|I_k^\xi(f_T) - I_k^\xi(f_{\epsilon,T})|^2 \leq k! \limsup_{T \rightarrow \infty} \|f_T - f_{\epsilon,T}\|_{L^2(\mathbb{R}^k)}^2 =$$

$$k! \limsup_{T \rightarrow \infty} \|h_T - f_\epsilon\|_{L^2(\mathbb{R}^k)}^2 \leq 2k! \limsup_{T \rightarrow \infty} \|f - f_\epsilon\|_{L^2(\mathbb{R}^k)}^2 \leq 2\epsilon. \quad (7.76)$$

To complete the proof of the lemma, in view of formulas (7.73) and (7.76), and Theorem 8.6.2 of Resnick [1999], it remains to show that as $T \rightarrow \infty$:

$$I_k^\xi(f_{\epsilon,T}) \xrightarrow{d} I_k^B(f_\epsilon). \quad (7.77)$$

Since $f_\epsilon(\cdot) \in \mathcal{S}_k$, we have

$$f_\epsilon(x_1, \dots, x_k) = \sum_{1 \leq i_1, \dots, i_k \leq N} c(i_1, \dots, i_k) 1_{\Delta_{i_1} \times \dots \times \Delta_{i_k}}(x_1, \dots, x_k),$$

where $N > 0$, $c(i_1, \dots, i_k) \in \mathbb{R}$, Δ_i 's are disjoint intervals so that $\cup_i \Delta_i = [-M, M]$, and the prime ' indicates that the sum does not include the diagonals $i_p = i_q$ for $p \neq q$. Then we have

$$\begin{aligned} I_k^\xi(f_{\epsilon,T}) &= T^{-k/2} \int_{\mathbb{R}^k} \sum_{1 \leq i_1, \dots, i_k \leq N} c(i_1, \dots, i_k) 1_{\Delta_{i_1} \times \dots \times \Delta_{i_k}}(x_1/T, \dots, x_k/T) \xi(dx_1) \dots \xi(dx_k) \\ &= \sum_{1 \leq i_1, \dots, i_k \leq N} c(i_1, \dots, i_k) \xi_T(\Delta_{i_1}) \dots \xi_T(\Delta_{i_k}), \end{aligned} \quad (7.78)$$

where

$$\xi_T(\Delta_i) = \frac{1}{\sqrt{T}} \int_{T\Delta_i} \xi(dx).$$

Combining the discretization argument in the proof of Theorem 18.7.1 of Ibragimov and Linnik [1971] and Theorem 5.2 of Billingsley [1956], it can be shown that the random variables $\xi_T(\Delta_i)$ satisfy central limit theorem. Hence for any $N \geq 1$, we have as $T \rightarrow \infty$

$$\left(\xi_T(\Delta_1), \dots, \xi_T(\Delta_N) \right) \xrightarrow{d} \left(B(\Delta_1), \dots, B(\Delta_N) \right),$$

where $B(\cdot)$ is a Gaussian random measure appearing in the Wiener-Itô integral (7.28). Hence applying continuous mapping theorem, from (7.78) we obtain

$$I_k^\xi(f_{\epsilon,T}) \xrightarrow{d} \sum_{1 \leq i_1, \dots, i_k \leq N} c(i_1, \dots, i_k) B(\Delta_{i_1}) \dots B(\Delta_{i_k}) = I_k^B(f_\epsilon),$$

implying (7.77). Lemma 7.5.1 is proved. \square

From Proposition 7.3.1 and Lemma 7.5.1, we easily infer the following result which is the spectral version of Lemma 7.5.1.

Corollary 7.5.2. *Let $\widehat{f}_{j,T}$ be the L^2 -Fourier transform of $f_{j,T}$. Set*

$$\widehat{h}_{j,T}(x_1, \dots, x_{k_j}) := T^{-k_j/2} \widehat{f}_{j,T}(x_1/T, \dots, x_{k_j}/T), \quad (7.79)$$

and assume that there exist $\widehat{f}_j \in L^2(\mathbb{R}^{k_j})$ such that as $T \rightarrow \infty$,

$$\|\widehat{h}_{j,T} - \widehat{f}_j\|_{L^2(\mathbb{R}^{k_j})} \rightarrow 0, \quad j = 1, \dots, J. \quad (7.80)$$

Then for any Lévy process $\xi(\cdot)$ with $E\xi(1) = 0$ and $E\xi^2(1) = 1$, we have the joint convergence in distribution:

$$\left(I_{k_1}^\xi(f_{1,T}), \dots, I_{k_J}^\xi(f_{J,T}) \right) \xrightarrow{d} \left((2\pi)^{-k_1/2} I_{k_1}^W(\widehat{f}_1 A^{\otimes k_1}), \dots, (2\pi)^{-k_J/2} I_{k_J}^W(\widehat{f}_J A^{\otimes k_J}) \right), \quad (7.81)$$

with any $A(\cdot)$ satisfying the conditions of Proposition 7.3.1.

The following lemma establishes a change of integration order in situations where the Fubini's theorem is not directly applicable.

Lemma 7.5.3. *Let $a(\cdot)$, $b(\cdot)$, $\widehat{a}(\cdot)$ and $\widehat{b}(\cdot)$ be as in Assumptions 1-4 in Section 7.2.2. Set*

$$g_T(x_1, x_2) = \int_0^T \int_0^T b(u-v) a(u-x_1) a(v-x_2) du dv.$$

Then $g_T(\cdot) \in L^2(\mathbb{R}^2)$ and for the L^2 -Fourier transform \widehat{g}_T of g_T , we have

$$\begin{aligned}\widehat{g}_T(w_1, w_2) &:= \int_{\mathbb{R}^2} e^{i(w_1 x_1 + w_2 x_2)} g(x_1, x_2) dx_1 dx_2 \\ &= \frac{1}{2\pi} \widehat{a}(-w_1) \widehat{a}(-w_2) \int_{\mathbb{R}} \frac{e^{iT(w_1+w)} - 1}{i(w_1+w)} \frac{e^{iT(w_1-w)} - 1}{i(w_1-w)} \widehat{b}(w) dw\end{aligned}$$

for a.e. $(w_1, w_2) \in \mathbb{R}^2$.

Proof. First, by the Cauchy-Schwartz inequality and Assumption 4, one has

$$\begin{aligned}\|g_T\|_{L^2(\mathbb{R}^2)}^2 &\leq \int_{\mathbb{R}^2} dx_1 dx_2 \left(\int_0^T \int_0^T |b(u-v)a(u-x_1)a(v-x_2)| dudv \right)^2 \\ &\leq \|a\|_{L^2(\mathbb{R})}^4 \left(\int_0^T \int_0^T |b(u-v)| dudv \right)^2 < \infty.\end{aligned}\tag{7.82}$$

Let $a_M(x) = a(x)1_{[-M, M]}(x)$, and let $\widehat{g}_{T, M}$ be the L^2 -Fourier transform of $g_{T, M}$ given by

$$g_{T, M}(x_1, x_2) = \int_0^T \int_0^T b(u-v)a_M(u-x_1)a_M(v-x_2)dudv.$$

Since, as $M \rightarrow \infty$, $a_M(x)$ and $\widehat{a}_M(w)$ converge in L^2 to $a(x)$ and $\widehat{a}(w)$, respectively, one can find a subsequence $M_n \uparrow \infty$, so that $a_{M_n}(x)$ and $\widehat{a}_{M_n}(w)$ converge a.e. to their limits.

So by (7.82) and the dominated convergence theorem, one has as $n \rightarrow \infty$

$$\begin{aligned}\|g_{T, M_n} - g_T\|_{L^2(\mathbb{R}^2)} &= \int_{\mathbb{R}^2} dx_1 dx_2 \times \\ &\left(\int_0^{Tt} \int_0^{Tt} b(u-v) [a_{M_n}(u-x_1)a_{M_n}(v-x_2) - a(u-x_1)a(v-x_2)] dudv \right)^2 \rightarrow 0.\end{aligned}$$

Therefore, one can choose a subsequence of M_n , still denoted by M_n , so that $g_{T, M_n}(x_1, x_2)$ converges to $g_T(x_1, x_2)$ a.e. $(x_1, x_2) \in \mathbb{R}^2$ as well as in L^2 -norm, and $\widehat{g}_{T, M_n}(w_1, w_2)$ converges to $\widehat{g}_T(w_1, w_2)$ a.e. $(w_1, w_2) \in \mathbb{R}^2$ as well as in L^2 -norm.

Since $b(\cdot)$ is an even function and $\widehat{b}(\cdot) \in L^1(\mathbb{R})$, one has

$$b(x) = \frac{1}{2\pi} \int_{\mathbb{R}} e^{-ixw} \widehat{b}(w) dw = \frac{1}{2\pi} \int_{\mathbb{R}} e^{ixw} \widehat{b}(w) dw.$$

Next, taking into account that $a_{M_n}(\cdot)$ has a finite support, and hence is in $L^1(\mathbb{R})$, one can write

$$\int_{\mathbb{R}} a_{M_n}(u-x)e^{iwx}dx = e^{iwu} \int_{\mathbb{R}} a_{M_n}(u-x)e^{i(-w)(u-x)}dx = e^{iwu}\widehat{a}_{M_n}(-w).$$

Then by Fubini's theorem, one can change the integration order to obtain

$$\begin{aligned} & \widehat{g}_{T,M_n}(w_1, w_2) \\ &= \int_{\mathbb{R}^2} e^{i(w_1x_1+w_2x_2)}dx_1dx_2 \int_0^{Tt} \int_0^{Tt} b(u-v)a_{M_n}(u-x_1)a_{M_n}(v-x_2)dudv \\ &= \frac{1}{2\pi} \int_{\mathbb{R}} dw \int_0^{Tt} \int_0^{Tt} dudv e^{i(u-v)w}\widehat{b}(w) e^{iw_1u}\widehat{a}_{M_n}(-w_1) e^{iw_2v}\widehat{a}_{M_n}(-w_2) \\ &= \frac{1}{2\pi} \widehat{a}_{M_n}(-w_1)\widehat{a}_{M_n}(-w_2) \int_{\mathbb{R}} \frac{e^{iTt(w_1+w)} - 1}{i(w_1+w)} \frac{e^{iTt(w_2-w)} - 1}{i(w_2-w)} \widehat{b}(w)dw. \end{aligned}$$

Finally, as $n \rightarrow \infty$, we have a.e. convergence of \widehat{g}_{T,M_n} to \widehat{g}_T and the a.e. convergence of \widehat{a}_{M_n} to \widehat{a} with M_n chosen above, and the result follows. The proof is then complete. \square

Lemma 7.5.4. *Let α^* and β^* be as in Assumption 4. Then*

$$\begin{aligned} & \int_{[0,1]^4} du_1du_2du_3du_4 \int_{\mathbb{R}} dx \left(|u_1 - u_2|^{\beta^*-1} |u_3 - u_4|^{\beta^*-1} |u_1 - x|^{\alpha^*/2-1} \right. \\ & \quad \left. \times |u_2 - x|^{\alpha^*/2-1} |u_3 - x|^{\alpha^*/2-1} |u_4 - x|^{\alpha^*/2-1} \right) < \infty \quad (7.83) \end{aligned}$$

Proof. Setting $\alpha_0 = \alpha^*/2 - 1 \in (-1, -1/2)$ and $\beta_0 = \beta^* - 1 \in (-1, 0)$, and noting that by the assumption $\alpha^* + \beta^* > 1$, we have

$$2\alpha_0 + \beta_0 > -2. \quad (7.84)$$

By the change of variables $v_1 = u_1 - u_2, v_2 = u_3 - u_4, v_3 = u_1, v_4 = u_1 - x, v_5 = u_3 - x$, and by enlarging the integration region if necessary, we can use the equality $\int_{\mathbb{R}} |v|^{\alpha_0} |x-v|^{\alpha_0} dv =$

$c_{\alpha_0}|x|^{2\alpha_0+1}$ with some constant $c_{\alpha_0} > 0$, to bound the integral in (7.83) as follows:

$$\begin{aligned} & c \int_{[-1,1]^3} dv_1 dv_2 dv_3 |v_1|^{\beta_0} |v_2|^{\beta_0} \int_{\mathbb{R}} |v_4|^{\alpha_0} |v_4 - v_1|^{\alpha_0} dv_4 \int_{\mathbb{R}} |v_5|^{\alpha_0} |v_5 - v_2|^{\alpha_0} dv_5 \\ & = c \int_{[-1,1]^3} dv_1 dv_2 dv_3 |v_1|^{2\alpha_0+\beta_0+1} |v_2|^{2\alpha_0+\beta_0+1}. \end{aligned} \quad (7.85)$$

The last integral in (7.85) is finite because by (7.84) we have $2\alpha_0 + \beta_0 + 1 > -1$. \square

The following lemma, which is a consequence of Corollary 1.1 (b) from Terrin and Taquq [1991b], will be used in the proof of Theorem 7.2.12.

Lemma 7.5.5. *Let $\alpha_1, \dots, \alpha_m$, $m \geq 2$ be real numbers satisfying*

$$\alpha_1, \dots, \alpha_m > -1, \quad \sum_{i=1}^m \alpha_i + m > 1, \quad (7.86)$$

then

$$\int_{[0,1]^m} |x_1 - x_2|^{\alpha_1} |x_2 - x_3|^{\alpha_2} \dots |x_{m-1} - x_m|^{\alpha_{m-1}} |x_m - x_1|^{\alpha_m} dx_1 \dots dx_m < \infty.$$

The next lemma, which provides a bound for slowly varying functions, called Potter's bound (see Giraitis et al. [2012], formula (2.3.6)), will be used in the proof of the main result.

Lemma 7.5.6. *Let $L(\cdot) : (0, \infty) \rightarrow \mathbb{R}$ be a function slowly varying at $u = 0$ and bounded on intervals $[c, \infty)$ for any $c > 0$. Then for any $\epsilon > 0$, there exists a constant $C > 0$, so that if T is large enough, then for any $u \in (0, \infty)$*

$$\frac{L(u/T)}{L(1/T)} \leq C(|u|^\epsilon + |u|^{-\epsilon}). \quad (7.87)$$

We now are ready to prove the non-central limit theorems stated in Section 6.2 (Theorems 7.2.8-7.2.14).

Proof of Theorem 7.2.8. As in the proof of Theorem 7.2.1, one can write

$$\tilde{Q}_T(t) = A_T(t) + B_T(t), \quad (7.88)$$

where now

$$A_T(t) = \int_{\mathbb{R}^2}' \frac{1}{T^{\alpha+\beta} L_1(1/T) L_2(1/T)} \int_0^{Tt} \int_0^{Tt} b(u-v) a(u-x_1) a(v-x_2) dudv \xi(dx_1) \xi(dx_2), \quad (7.89)$$

and

$$B_T(t) = \int_{\mathbb{R}} \frac{1}{T^{\alpha+\beta} L_1(1/T) L_2(1/T)} \int_0^{Tt} \int_0^{Tt} b(u-v) a(u-x) a(v-x) dudv \xi_c^{(2)}(dx). \quad (7.90)$$

In view of (7.88)-(7.90), to prove the theorem, it is enough to show that $A_T(t)$ converges in finite-dimensional distributions to the limit $Z_{\alpha,\beta}(t)$ given by (7.20), and $\lim_{T \rightarrow \infty} EB_T^2(t) = 0$.

We first prove that

$$A_T(t) \xrightarrow{f.d.d.} Z_{\alpha,\beta}(t) \quad \text{as } T \rightarrow \infty. \quad (7.91)$$

The relation (7.91) we deduce from Corollary 7.5.2. To this end, we write $A_T(t) = I_2^\xi(f_{T,t})$, where

$$f_{T,t} = \frac{1}{T^{\alpha+\beta} L_1(1/T) L_2(1/T)} \int_0^{Tt} \int_0^{Tt} b(u-v) a(u-x_1) a(v-x_2) dudv.$$

Denoting by $\hat{f}_{T,t}$ the L^2 -Fourier transform of $f_{T,t}$, and using Lemma 7.5.3, we have

$$\begin{aligned} \hat{f}_{T,t}(w_1, w_2) &= \frac{1}{2\pi T^{\alpha+\beta} L_1(1/T) L_2(1/T)} \hat{a}(-w_1) \hat{a}(-w_2) \\ &\quad \times \int_{\mathbb{R}} \frac{e^{iTt(w_1+w)} - 1}{i(w_1+w)} \frac{e^{iTt(w_1-w)} - 1}{i(w_1-w)} \hat{b}(w) dw. \end{aligned}$$

By changing the variables w_1 , w_2 and w by x_1 , x_2 and u/T , respectively, one has

$$\begin{aligned} \widehat{h}_{T,t}(x_1, x_2) &:= T^{-1} \widehat{f}_{T,t}(x_1/T, x_2/T) = \frac{1}{2\pi} \frac{\widehat{a}(-x_1/T)}{T^{\alpha/2} L_1(1/T)^{1/2}} \frac{\widehat{a}(-x_2/T)}{T^{\alpha/2} L_1(1/T)^{1/2}} \\ &\quad \times \int_{\mathbb{R}} \frac{e^{it(x_1+u)} - 1}{i(x_1+u)} \frac{e^{it(x_1-u)} - 1}{i(x_1-u)} \frac{\widehat{b}(u/T)}{T^\beta L_2(1/T)} du. \end{aligned}$$

Next, by the Assumptions 1 and 2 and the property of slowly varying functions: $\lim_{T \rightarrow \infty} L_i(x/T)L_i(1/T) = 1$, one has

$$\begin{aligned} \widehat{h}_{T,t}(x_1, x_2) &\rightarrow \widehat{f}_t(x_1, x_2) := \frac{1}{2\pi} H(x_1)H(x_2)|x_1x_2|^{-\alpha/2} \\ &\quad \times \int_{\mathbb{R}} \frac{e^{it(x_1+u)} - 1}{i(x_1+u)} \frac{e^{it(x_2-u)} - 1}{i(x_2-u)} |u|^{-\beta} du \quad (7.92) \end{aligned}$$

for a.e. $(x_1, x_2) \in \mathbb{R}^2$, where $H(x) = \bar{A}_0$ if $x \geq 0$ and $H(x) = A_0$ if $x < 0$, with A_0 as in Assumption 1.

By (7.87), when T is large enough, with some constant $C > 0$ one has

$$\begin{aligned} |\widehat{h}_{T,t}(x_1, x_2)| &\leq h_t^*(x_1, x_2) := C|x_1x_2|^{-\alpha/2}(|x_1|^\epsilon + |x_1|^{-\epsilon})(|x_2|^\epsilon + |x_2|^{-\epsilon}) \\ &\quad \times \int_{\mathbb{R}} \left| \frac{e^{it(x_1+u)} - 1}{i(x_1+u)} \frac{e^{it(x_2-u)} - 1}{i(x_2-u)} \right| |u|^{-\beta} (|u|^\epsilon + |u|^{-\epsilon}) du. \quad (7.93) \end{aligned}$$

By Lemma 3.9 of Bai et al. [2015], for small enough ϵ , the bounding function $h_t^*(x_1, x_2)$ in (7.93) belongs to $L^2(\mathbb{R}^2)$.

Therefore, by the dominated convergence theorem, we have

$$\lim_{T \rightarrow \infty} \|\widehat{h}_{T,t} - \widehat{f}_t\|_{L^2(\mathbb{R}^2)} = 0.$$

Now we can apply Corollary 7.5.2 to obtain (7.91). Note that the function $H(x)$ in (7.92) can be omitted since it plays the role of the function $A(\cdot)$ in Proposition 7.3.1. The proof of (7.91) is complete.

Next we prove that

$$\lim_{T \rightarrow \infty} \mathbb{E} B_T^2(t) = 0. \quad (7.94)$$

For simplicity we consider the case $t = 1$ and set

$$D_T(x) = \int_0^T \int_0^T b(u-v)a(u-x)a(v-x)dudv. \quad (7.95)$$

Then by Assumption 4 and the change of variables $u_i \rightarrow Tu_i$ and $v_i \rightarrow Tv_i$, we get

$$\|D_T\|_{L^2(\mathbb{R})}^2 \leq CT^{2\alpha^*+2\beta^*-1}D(\alpha^*, \beta^*), \quad (7.96)$$

where

$$D(\alpha^*, \beta^*) = \int_{[0,1]^4} du_1 du_2 dv_1 dv_2 \int_{\mathbb{R}} dx |u_1 - v_1|^{\beta^*-1} |u_1 - x|^{\alpha^*/2-1} |v_1 - x|^{\alpha^*/2-1} \\ \times |u_2 - v_2|^{\beta^*-1} |u_2 - x|^{\alpha^*/2-1} |v_2 - x|^{\alpha^*/2-1}. \quad (7.97)$$

By Lemma 7.5.4, the last integral is finite. Since L_1 and L_2 in (7.90) are slowly varying, for any $\epsilon > 0$ and for large enough T we have (see Bingham et al. [1989], Proposition 1.3.6)

$$L_1(1/T) \geq T^{-\epsilon/4}, \quad L_2(1/T) \geq T^{-\epsilon/4}. \quad (7.98)$$

Therefore, in view of (7.90) and (7.95)-(7.98), we can write

$$\mathbb{E} B_T^2(t) \leq CT^{2(\alpha^*+\beta^*-\alpha-\beta)+\epsilon-1},$$

where $2(\alpha^* + \beta^* - \alpha - \beta) + \epsilon - 1 < 0$ by Assumption 4 if ϵ is chosen small enough. This completes the proof of (7.94). Theorem 7.2.8 is proved. \square

Proof of Theorem 7.2.12. To prove assertion (a), we start with the corresponding “time

domain" kernel in (7.22), namely,

$$f_t(x_1, x_2) = \int_0^t \int_0^t |u - v|^{\beta-1} (u - x_1)_+^{\alpha/2-1} (v - x_2)_+^{\alpha/2-1} dudv,$$

and observe that $f_t \in L^2(\mathbb{R}^2)$. Indeed, using the equality

$$\int_{\mathbb{R}} (u - x)_+^{\alpha/2-1} (v - x)_+^{\alpha/2-1} dx = C_\alpha |u - v|^{\alpha-1}$$

with $0 < \alpha < 1$ and some $C_\alpha > 0$, and Lemma 7.5.5 with $\alpha + \beta > 1/2$, one has

$$\begin{aligned} & \int_{\mathbb{R}^2} f_t(x_1, x_2)^2 dx_1 dx_2 \\ &= C_\alpha^2 \int_{[0, t]^4} |u_1 - v_1|^{\beta-1} |u_1 - u_2|^{\alpha-1} |u_2 - v_2|^{\beta-1} |v_1 - v_2|^{\alpha-1} du_1 dv_1 du_2 dv_2 < \infty. \end{aligned}$$

To determine the Fourier transform \widehat{f}_t of f_t , we truncate f_t as follows:

$$f_t^A(x_1, x_2) = \int_0^t \int_0^t |u - v|^{\beta-1} (u - x_1)_+^{\alpha/2-1} (v - x_2)_+^{\alpha/2-1} \mathbf{1}_{\{u-x_1 < A, v-x_2 < A\}} dudv.$$

Then by the dominated convergence theorem, one has as $A \rightarrow \infty$

$$f_t^A(x_1, x_2) \rightarrow f_t(x_1, x_2) \quad \text{in } L^2(\mathbb{R}^2).$$

Thus by Parseval-Plancherel isometry, as $A \rightarrow \infty$, for the Fourier transforms, we have

$$\widehat{f}_t^A(x_1, x_2) \rightarrow \widehat{f}_t(x_1, x_2) \quad \text{in } L^2(\mathbb{R}^2).$$

Hence we can let $A \rightarrow \infty$ through a suitable subsequence to get

$$\widehat{f}_t^A(w_1, w_2) \rightarrow \widehat{f}_t(w_1, w_2) \quad a.e. \tag{7.99}$$

Next, we determine \widehat{f}_t explicitly. We apply Fubini's theorem to obtain

$$\begin{aligned}\widehat{f}_t^A(w_1, w_2) &:= \int e^{i(w_1x_1+w_2x_2)} f_t^A(x_1, x_2) dx_1 dx_2 \\ &= \int_0^t \int_0^t |u-v|^{\beta-1} e^{i(w_1u+w_2v)} dudv \\ &\quad \times \int_0^A e^{-iw_1y_1} y_1^{\alpha/2-1} dy_1 \times \int_0^A e^{-iw_2y_2} y_2^{\alpha/2-1} dy_2.\end{aligned}\tag{7.100}$$

We first deal with the first integral on the right-hand side of (7.100), which we rewrite in a convenient form. To this end, observe that by formula 3.761.9 of Jeffrey and Zwillinger [2007]

$$\begin{aligned}\lim_{B \rightarrow \infty} \int_{-B}^B |w|^{-\beta} e^{ixw} dw &= \int_{\mathbb{R}} |w|^{-\beta} e^{ixw} dw \\ &= 2 \int_0^\infty w^{-\beta} \cos(xw) dw = 2|x|^{\beta-1} \Gamma(1-\beta) \sin(\beta\pi/2).\end{aligned}$$

Set

$$M_\beta = \sup_{B>0} \left| \int_{-B}^B |w|^{-\beta} e^{iw} dw \right| < \infty,$$

and make a change of variable $w' = (u-v)w$ to obtain

$$\int_{-B}^B |w|^{-\beta} e^{iw(u-v)} dw \leq M_\beta |u-v|^{\beta-1}.$$

Note also that since $\beta > 0$, we have

$$\int_0^t \int_0^t dudv |u-v|^{\beta-1} < \infty.$$

Hence using the dominated convergence theorem, we can write

$$\begin{aligned}&\int_0^t \int_0^t dudv |u-v|^{\beta-1} e^{i(w_1u+w_2v)} \\ &= \frac{1}{2\Gamma(1-\beta) \sin(\beta\pi/2)} \int_0^t \int_0^t dudv \lim_{B \rightarrow \infty} \int_{-B}^B |w|^{-\beta} e^{iw(u-v)} dw e^{i(w_1u+w_2v)}\end{aligned}$$

$$\begin{aligned}
&= \frac{1}{2\Gamma(1-\beta)\sin(\beta\pi/2)} \lim_{B \rightarrow \infty} \int_{-B}^B \frac{e^{it(w_1+w)} - 1}{i(w_1+w)} \frac{e^{it(w_2-w)} - 1}{i(w_2-w)} |w|^{-\beta} dw \\
&= \frac{1}{2\Gamma(1-\beta)\sin(\beta\pi/2)} \int_{-\infty}^{\infty} \frac{e^{it(w_1+w)} - 1}{i(w_1+w)} \frac{e^{it(w_2-w)} - 1}{i(w_2-w)} |w|^{-\beta} dw, \tag{7.101}
\end{aligned}$$

because in view of (7.93) and the fact that $h_t^*(x_1, x_2) \in L^2(\mathbb{R}^2)$, the last integral converges absolutely.

Next, we focus on the last two integrals in (7.100). By formulas 3.761.4 and 3.761.9 of Jeffrey and Zwillinger [2007], we have

$$\begin{aligned}
\lim_{A \rightarrow \infty} \int_0^A e^{-iwy} y_+^{\alpha/2-1} dy &= \int_0^\infty e^{-iwy} y^{\alpha/2-1} dy = \int_0^\infty [\cos(wy) - i \sin(wy)] y^{\alpha/2-1} dy \\
&= |w|^{-\alpha/2} \Gamma(\alpha/2) [\cos(\alpha\pi/4) - i \operatorname{sign}(w) \sin(\alpha\pi/4)] \\
&= |w|^{-\alpha/2} \Gamma(\alpha/2) \exp[-i \operatorname{sign}(w) \alpha\pi/4]. \tag{7.102}
\end{aligned}$$

Combining (7.99)-(7.102), one gets

$$\begin{aligned}
\widehat{f}_t(w_1, w_2) &= \frac{\Gamma(\alpha/2)^2}{2\Gamma(1-\beta)\sin(\beta\pi/2)} \int_{-\infty}^{\infty} \frac{e^{it(w_1+w)} - 1}{i(w_1+w)} \frac{e^{it(w_2-w)} - 1}{i(w_2-w)} |w|^{-\beta} dw \\
&\quad \times \exp[-i(\operatorname{sign}(w_1) + \operatorname{sign}(w_2))\alpha\pi/4] |w_1 w_2|^{-\alpha/2}.
\end{aligned}$$

The proof of assertion (a) can be concluded using Proposition 7.3.1, and noting that the factor

$$\exp[-i(\operatorname{sign}(w_1) + \operatorname{sign}(w_2))\alpha\pi/4]$$

in the last formula can be omitted as it plays the role as $A^{\otimes 2}(w_1, w_2)$ in Proposition 7.3.1.

To prove assertion (b), we set

$$\widehat{1}_{[0,t]}(x) = \int_{\mathbb{R}} 1_{[0,t]}(w) e^{iwx} dw = \frac{e^{itx} - 1}{ix},$$

and use the property of Fourier transform for convolutions to obtain

$$\begin{aligned} & \int_{\mathbb{R}} \frac{e^{it(x_1+u)} - 1}{i(x_1+u)} \frac{e^{it(x_2-u)} - 1}{i(x_2-u)} du \\ &= \int_{\mathbb{R}} \widehat{1}_{[0,t]}(x_1+u) \widehat{1}_{[0,t]}(x_2-u) du = \left(\widehat{1}_{[0,t]} * \widehat{1}_{[0,t]} \right) (x_1+x_2) \\ &= \widehat{(1_{[0,t]} \cdot 1_{[0,t]})} (x_1+x_2) = \widehat{1}_{[0,t]}(x_1+x_2) = \frac{e^{it(x_1+x_2)} - 1}{i(x_1+x_2)}. \end{aligned}$$

So in view of (7.20), the process $Z_{\alpha,0}$ can be written as follows:

$$Z_{\alpha,0}(t) = \frac{1}{2\pi} \int_{\mathbb{R}^2} |x_1 x_2|^{-\alpha/2} \frac{e^{it(x_1+x_2)} - 1}{i(x_1+x_2)} W(dx_1) W(dx_2),$$

which is the well-known spectral-domain representation of the Rosenblatt process (see Taqqu [1979]). Thus, the time-domain representation stated in (7.23) follows from Theorem 1.1 of Pipiras and Taqqu [2010].

To prove assertion (c), we set

$$f_t(x_1, x_2) = 1_{[0,t] \times [0,t]}(x_1, x_2) |x_1 - x_2|^{\beta-1},$$

where $\beta > 0$, and observe that by (7.101),

$$\widehat{f}_t(w_1, w_2) = \frac{1}{2\Gamma(1-\beta) \sin(\beta\pi/2)} \int_{-\infty}^{\infty} \frac{e^{it(w_1+w)} - 1}{i(w_1+w)} \frac{e^{it(w_2-w)} - 1}{i(w_2-w)} |w|^{-\beta} dw,$$

which, in view of Proposition 7.3.1, implies (7.20). This completes the proof of Theorem 7.2.12. \square

Proof of Theorem 7.2.14. As in the proof of Theorem 7.2.8, we can write

$$\frac{1}{T^{\alpha+\beta} \ell_1(T) \ell_2(T)} (Q_T - \mathbb{E}Q_T) = A_T(t) + B_T(t),$$

where $A_T(t)$ and $B_T(t)$ are given in (7.89) and (7.90), respectively, with $L_j(1/T)$ replaced

by $\ell_j(T)$, $j = 1, 2$.

Since (7.21) implies Assumption 4 of Theorem 7.2.8, as in the proof of Theorem 7.2.8, we get $\lim_{T \rightarrow \infty} EB_T^2(t) = 0$, implying that the term $B_T(t)$ is negligible.

Next, setting

$$f_{t,T}(x_1, x_2) = \frac{1}{T^{\alpha+\beta} \ell_1(T) \ell_2(T)} \int_0^{Tt} \int_0^{Tt} b(u-v) a(u-x_1) a(v-x_2) dudv,$$

we have $A_T(t) = I_2^\xi(f_{t,T})$. Then in view of Lemma 7.5.1, we can write

$$\begin{aligned} h_{t,T}(x_1, x_2) &:= T f_{t,T}(Tx_1, Tx_2) \\ &= \frac{1}{T^{\alpha+\beta-1} \ell_1(T) \ell_2(T)} \int_0^{Tt} \int_0^{Tt} b(u-v) a(u-Tx_1) a(v-Tx_2) dudv \\ &= \int_0^t \int_0^t |u-v|^{\beta-1} (u-x_1)_+^{\alpha/2-1} (v-x_2)_+^{\alpha/2-1} \\ &\quad \times \frac{\ell_1(T(u-x_1))}{\ell_1(T)} \frac{\ell_1(T(v-x_2))}{\ell_1(T)} \frac{\ell_2(T(u-v))}{\ell_2(T)} dudv, \end{aligned} \tag{7.103}$$

where we have applied the change of variables $u \rightarrow uT$ and $v \rightarrow vT$. Let

$$f_t(x_1, x_2) = \int_0^t \int_0^t |u-v|^{\beta-1} (u-x_1)_+^{\alpha/2-1} (v-x_2)_+^{\alpha/2-1} dudv. \tag{7.104}$$

To complete the proof of the theorem, in view of Lemma 7.5.1, it is enough to show that for every $t > 0$,

$$h_{t,T}(x_1, x_2) \rightarrow f_t(x_1, x_2) \quad \text{in } L^2(\mathbb{R}^2) \text{ as } T \rightarrow \infty. \tag{7.105}$$

By the property of slowly varying functions, we have as $T \rightarrow \infty$

$$\frac{\ell_1(T(u-x_1))}{\ell_1(T)}, \frac{\ell_1(T(v-x_2))}{\ell_1(T)}, \frac{\ell_2(T(u-v))}{\ell_2(T)} \rightarrow 1$$

for $u > x_1$, $v > x_2$ and $u \neq v$ (recall that $\ell_2(\cdot)$ is an even function). Next, by Potter's

bound in Lemma 7.5.6, for any $\epsilon > 0$, one has

$$1_{\{u > x_1, v > x_2\}} \left| \frac{\ell_1(T(u - x_1))}{\ell_1(T)} \frac{\ell_1(T(v - x_2))}{\ell_1(T)} \frac{\ell_2(T(u - v))}{\ell_2(T)} \right| \leq CR_\epsilon(x_1, x_2, u, v),$$

where

$$R_\epsilon(x_1, x_2, u, v) = [(u - x_1)_+^\epsilon + (u - x_1)_+^{-\epsilon}][(v - x_2)_+^\epsilon + (v - x_2)_+^{-\epsilon}][|u - v|^\epsilon + |u - v|^{-\epsilon}].$$

Thus the function $h_{t,T}(x_1, x_2)$ in (7.103) is bounded by

$$f_{t,\epsilon}(x_1, x_2) := C \int_0^t \int_0^t |u - v|^{\beta-1} (u - x_1)_+^{\alpha/2-1} (v - x_2)_+^{\alpha/2-1} R_\epsilon(x_1, x_2, u, v) dudv.$$

By choosing $\epsilon > 0$ small enough, as in the proof of Theorem 7.2.12 (a), we can use Lemma 7.5.4 to show that $f_{t,\epsilon}(x_1, x_2) \in L^2(\mathbb{R}^2)$. Then the dominated convergence theorem can be applied to obtain (7.105). Theorem 7.2.14 is proved. \square

Chapter 8

Behavior of the generalized Rosenblatt process at extreme critical exponent values

8.1 Introduction

Maejima and Tudor [2012] considered recently the following process defined through a second-order Wiener-Itô integral:

$$Z_{\gamma_1, \gamma_2}(t) = A \int_{\mathbb{R}^2}' \left[\int_0^t (s - x_1)_+^{\gamma_1} (s - x_2)_+^{\gamma_2} ds \right] B(dx_1) B(dx_2), \quad (8.1)$$

where $A \neq 0$ is a constant, $B(\cdot)$ is a Brownian random measure, the prime $'$ indicates the *exclusion* of the diagonals $x_1 = x_2$ in the double stochastic integral, and the exponents γ_1, γ_2 live in the following open triangular region (see Figure 8.1):

$$\Delta = \{(\gamma_1, \gamma_2) : -1 < \gamma_1 < -1/2, -1 < \gamma_2 < -1/2, \gamma_1 + \gamma_2 > -3/2\}. \quad (8.2)$$

This ensures that the integrand in (8.1) is in $L^2(\mathbb{R}^2)$, and hence the process $Z_{\gamma_1, \gamma_2}(t)$ is well-defined (see Theorem 3.5 and Remark 3.1 of Bai and Taqqu [2014a]).

We shall call $Z_{\gamma_1, \gamma_2}(t)$ a *generalized Rosenblatt process*. The *Rosenblatt process* $Z_\gamma(t)$ (Taqqu [1975]) becomes the special case

$$Z_\gamma(t) = Z_{\gamma, \gamma}(t), \quad -3/4 < \gamma < -1/2. \quad (8.3)$$

Recent studies on the Rosenblatt process $Z_\gamma(t)$ include Tudor and Viens [2009], Bardet

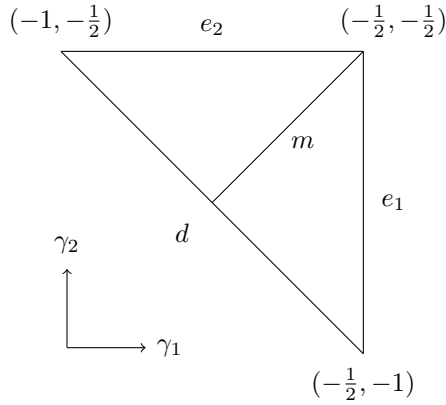


Figure 8.1: Region Δ defined in (8.2).

The three edges of the triangle are named e_1, e_2 and d (diagonal), while the middle line segment (symmetric axis) is named m .

and Tudor [2010], Arras [2013], Maejima and Tudor [2013], Veillette and Taqqu [2013] and Bojdecki et al. [2013]. The Rosenblatt and the generalized Rosenblatt processes are of interest because they are the simplest extension to the non-Gaussian world of the Gaussian fractional Brownian motion.

Fractional Brownian motion $B_H(t)$, $1/2 < H < 1$ is defined through a single Wiener-Itô (or Wiener) integral:

$$B_H(t) = C \int_{\mathbb{R}} \left[\int_0^t (s-x)_+^{H-3/2} ds \right] B(dx),$$

and has covariance

$$EB_H(s)B_H(t) = \frac{C'}{2} (|s|^{2H} + |t|^{2H} - |s-t|^{2H}), \quad (8.4)$$

where C and C' are two related constants. Fractional Brownian motion reduces to Brownian motion if one sets $H = 1/2$ in (8.4). Fractional Brownian motion has stationary increments and, for any $1/2 < H < 1$, these increments have a covariance which decreases slowly as the lag increases. This slow decay is often referred to as *long memory* or *long-range dependence*. Fractional Brownian motion is also self-similar with *self-similarity*

parameter (Hurst index) H , that is, $B_H(\lambda t)$ has the same finite-dimensional distributions as $\lambda^H B_H(t)$ for any $\lambda > 0$. It follows from Bai and Taqqu [2014a] that the generalized Rosenblatt process $Z_{\gamma_1, \gamma_2}(t)$ is also self-similar with stationary increments with self-similarity parameter

$$H = \gamma_1 + \gamma_2 + 2 \in (1/2, 1). \quad (8.5)$$

We get $1/2 < H < 1$ because $\gamma_1, \gamma_2 < -1/2$ imply $H < 1$ and $\gamma_1 + \gamma_2 > -3/2$ implies $H > 1/2$.

Fractional Brownian motion and the generalized Rosenblatt process $Z_{\gamma_1, \gamma_2}(t)$ belong to a broad class of self-similar processes with stationary increments defined on a Wiener chaos called *generalized Hermite processes*. The generalized Hermite processes appear as limits in various types of non-central limit theorems involving Volterra-type nonlinear process. In particular, the generalized Rosenblatt process $Z_{\gamma_1, \gamma_2}(t)$ can arise as limit when considering a quadratic form involving two long-memory linear processes with different memory parameters. See Bai and Taqqu [2014a, 2015b,c] for details.

It will be convenient to express the generalized Rosenblatt process as follows,

$$Z_{\gamma_1, \gamma_2}(t) = \frac{A}{2} \int_{\mathbb{R}^2} \left[\int_0^t [(s - x_1)_+^{\gamma_1} (s - x_2)_+^{\gamma_2} + (s - x_1)_+^{\gamma_2} (s - x_2)_+^{\gamma_1}] ds \right] B(dx_1) B(dx_2), \quad (8.6)$$

where we replaced the kernel $A \int_0^t (s - x_1)_+^{\gamma_1} (s - x_2)_+^{\gamma_2} ds$ by its symmetrized version. The process $Z_{\gamma_1, \gamma_2}(t)$ remains invariant under such a modification.

The goal of this chapter is to study the distributional behavior of the standardized $Z_{\gamma_1, \gamma_2}(t)$ (where A in (8.6) is chosen so that $\text{Var}[Z_{\gamma_1, \gamma_2}(1)] = 1$), as (γ_1, γ_2) approaches the boundaries of the region Δ defined in (8.2).

We show that on the diagonal boundary d , the limit is Brownian motion. On each of the two symmetric boundaries e_1 and e_2 of Δ , the limit is non-Gaussian: it is a fractional Brownian motion times an independent Gaussian random variable. We give two different proofs of this convergence, one based on the method of moments, and one which provides

more intuitive insight. We also give the rate of convergence to the marginal distribution in the preceding two cases.

The situation at the corners is particularly delicate. At the corner $(\gamma_1, \gamma_2) = (-\frac{1}{2}, -\frac{1}{2})$, the limit process is a linear combination of two independent degenerate chi-square processes. At the other two corners, the limit is a linear combination of two processes: a Brownian motion and the product of another Brownian motion times an independent Gaussian random variable. These linear combinations, which depend on the direction at which the critical exponents approach the corners, will be given explicitly.

We also show that the convergences mentioned cannot be strengthened from weak convergence to $L^2(\Omega)$ convergence, nor even to convergence in probability.

The chapter is organized as follows. In Section 8.2, we state the main results with proofs in Section 8.3. In the following three sections, we provide some additional results: showing that $L^2(\Omega)$ convergence cannot hold, establishing the rate of marginal convergence on the boundaries d , e_1 and e_2 , and giving an alternate proof of the convergence on the boundaries e_1 and e_2 .

8.2 Main results

In the following theorems, we let \Rightarrow denote weak convergence in the space $C[0, 1]$ with uniform metric. The multiplicative factor A in (8.6) is chosen so that $\text{Var}[Z_{\gamma_1, \gamma_2}(1)] = 1$. See (8.21) below for an explicit expression.

We focus first on results concerning the behavior of $Z_{\gamma_1, \gamma_2}(t)$ as (γ_1, γ_2) approaches the boundary of $\mathbf{\Delta}$ in (8.2), excluding the corners. Theorem 8.2.1 involves convergence to the diagonal edge d of $\mathbf{\Delta}$, where the limit is Brownian motion. See Figure 8.2.

Theorem 8.2.1. *Let $Z_{\gamma_1, \gamma_2}(t)$, $(\gamma_1, \gamma_2) \in \mathbf{\Delta}$, be defined in (8.6) with $A = A(\gamma_1, \gamma_2)$ in (8.21). When $\gamma_1 + \gamma_2 \rightarrow -3/2$ with $\gamma_1, \gamma_2 > -1 + \epsilon$ for arbitrarily fixed $\epsilon > 0$, we have*

$$Z_{\gamma_1, \gamma_2}(t) \Rightarrow B(t), \tag{8.7}$$

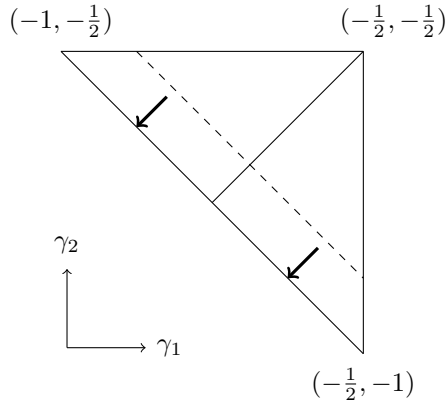


Figure 8.2: Illustration of limit taking in Theorem 8.2.1

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

One has $\gamma_1 + \gamma_2 = -3/2$ all through the diagonal d . The corners of the triangle are excluded by the requirement $\gamma_1, \gamma_2 > -1 + \epsilon$. Convergence to Brownian motion in (8.7) is expected heuristically since the self-similarity parameter $H = \gamma_1 + \gamma_2 + 2 \rightarrow 1/2$ (see (8.5)), and $1/2$ is the self-similarity parameter of Brownian motion.

The next Theorem 8.2.2 involves convergence to either one of the two sides e_1 and e_2 of Δ . The vertical side e_1 and the horizontal side e_2 are parameterized respectively by $(-1/2, \gamma)$ and $(\gamma, -1/2)$ where $-1 < \gamma < -1/2$. See Figure 8.3.

Theorem 8.2.2. *Let $Z_{\gamma_1, \gamma_2}(t)$, $(\gamma_1, \gamma_2) \in \Delta$, be defined in (8.6) with $A = A(\gamma_1, \gamma_2)$ in (8.21). When $(\gamma_1, \gamma_2) \rightarrow (-1/2, \gamma)$ or $(\gamma_1, \gamma_2) \rightarrow (\gamma, -1/2)$, where $-1 < \gamma < -1/2$, we have*

$$Z_{\gamma_1, \gamma_2}(t) \Rightarrow WB_{\gamma+3/2}(t), \quad (8.8)$$

where $B_{\gamma+3/2}(t)$ is a standard fractional Brownian motion with self-similarity parameter $\gamma + 3/2$, and W is a standard normal random variable which is independent of $B_{\gamma+3/2}(t)$.

Remark 8.2.3. The convergence (8.8) is more involved since $WB_{\gamma+3/2}(t)$ is a self-similar process with stationary increments having self-similarity parameter $H = \gamma + 3/2 \in (1/2, 1)$, and hence displays long-range dependence. This convergence may be understood heuristically as follows: $Z_{\gamma_1, \gamma_2}(t)$ in (8.1) can be regarded as an integrated process of a long-range

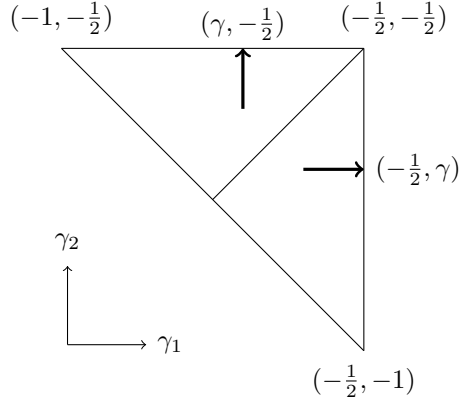


Figure 8.3: Illustration of limit taking in Theorem 8.2.2

dependent bilinear moving average of white noise. This bilinear moving average involves a double summation. As the exponent $\gamma_1 \rightarrow -1/2$, the corresponding summation yields a term which is extremely persistent, so that it behaves like a frozen Gaussian variable which is independent of the fractional noise defined through the other summation.

Remark 8.2.4. Although intuitively the generalized Rosenblatt processes $Z_{\gamma_1, \gamma_2}(t)$ in (8.1) form a richer class than the Rosenblatt process $Z_\gamma(t)$ in (8.3), they are both self-similar with stationary increments, and hence have the same covariance (8.4) when $2\gamma = \gamma_1 + \gamma_2$. To show that they are different processes, one can compare the higher moments, as was done in Bai and Taqqu [2014b]. The convergence (8.8) provides another evidence that there are values of (γ_1, γ_2) for which $Z_{\gamma_1, \gamma_2}(t)$ is different from $Z_\gamma(t)$. Indeed the limit $WB_{\gamma+3/2}(t)$ has a symmetric marginal distribution (the so-called product-normal distribution), while the marginal distribution of the Rosenblatt process $Z_\gamma(t)$ is skewed with a nonzero third cumulant (see (10) and (12) of Veillette and Taqqu [2013], or set $\gamma_1 = \gamma_2 = \gamma$ in (8.20) below).

Note that in Theorem 8.2.1 and 8.2.2, we exclude the three corners $(\gamma_1, \gamma_2) = (-\frac{1}{2}, -\frac{1}{2})$, $(-1, -1/2)$ and $(-1/2, -1)$. It turns out that the limit behavior of $Z_{\gamma_1, \gamma_2}(t)$ at these corners depends on the direction these corners are approached. Due to the symmetry of $Z_{\gamma_1, \gamma_2}(t)$ in (γ_1, γ_2) , it is sufficient to focus on the case $\gamma_1 \geq \gamma_2$, that is, we focus on the subregion

of Δ in (8.2) delimited by line segments e_1, d and m in Figure 8.4.

Consider first the corner $(\gamma_1, \gamma_2) = (-1/2, -1)$. We will approach it through the line

$$\gamma_2 = \frac{1}{\rho - 1}(\gamma_1 + 1/2) - 1,$$

which can also be expressed as

$$\frac{\gamma_1 + \gamma_2 + 3/2}{\gamma_2 + 1} = \rho.$$

The line passes through the corner $(-1/2, -1)$ and has a negative slope of $1/(\rho - 1)$, $0 \leq \rho \leq 1$. See Figure 8.4. When $\rho = 0$, the line coincides with the diagonal edge d of the triangle Δ , which has slope -1 . When $\rho = 1$, the line coincides with the vertical side e_1 of Δ , which has slope $-\infty$.

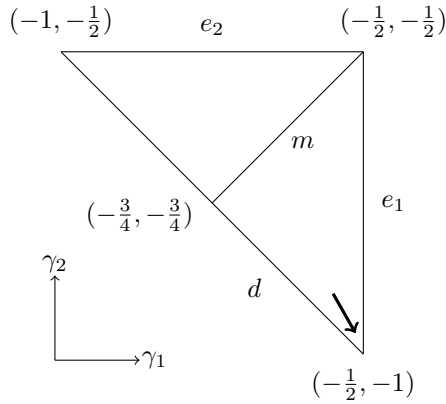


Figure 8.4: Illustration of limit taking in Theorem 8.2.5

Theorem 8.2.5 (The corner $(\gamma_1, \gamma_2) = (-1/2, -1)$).

Let $Z_{\gamma_1, \gamma_2}(t)$, $(\gamma_1, \gamma_2) \in \Delta$, be defined in (8.6) with $A = A(\gamma_1, \gamma_2)$ in (8.21). Suppose that $\gamma_1 \geq \gamma_2$. If $(\gamma_1, \gamma_2) \rightarrow (-1/2, -1)$ in such a way that

$$\frac{\gamma_1 + \gamma_2 + 3/2}{\gamma_2 + 1} = 1 + \frac{\gamma_1 + 1/2}{\gamma_2 + 1} \rightarrow \rho \in [0, 1], \quad (8.9)$$

then

$$Z_{\gamma_1, \gamma_2}(t) \Rightarrow X_\rho(t) := \rho^{1/2}WB(t) + (1 - \rho)^{1/2}B'(t), \quad (8.10)$$

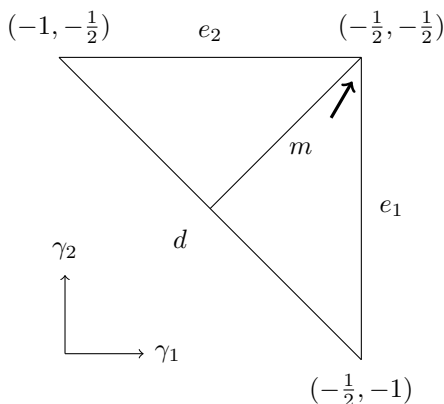


Figure 8.5: Illustration of limit taking in Theorem 8.2.7

where W is a standard normal random variable, $B(t)$ and $B'(t)$ are standard Brownian motions, and W , $B(t)$ and $B'(t)$ are independent.

Remark 8.2.6. In Theorem 8.2.5, the limit $X_\rho(t)$ is an independent linear combination of the two limits obtained in Theorem 8.2.2 and 8.2.1 (edges e_1 and d), after setting $\gamma = -1$ in Theorem 8.2.2. Note that since $\gamma + 3/2 = -1 + 3/2 = 1/2$, the fractional Brownian motion $B_{\gamma+3/2}(t)$ in Theorem 8.2.2 becomes Brownian motion $B(t)$.

Consider now the corner $(\gamma_1, \gamma_2) = (-1/2, -1/2)$. We will approach it through the line

$$\gamma_2 = \frac{1}{\rho}(\gamma_1 + 1/2) - 1/2,$$

which passes through it and has a positive slope of $1/\rho$, $0 \leq \rho \leq 1$. See Figure 8.5. When $\rho = 0$, the line coincides with the vertical side e_1 of Δ , which has slope $+\infty$. When $\rho = 1$, the line coincides with the middle line m , which has slope 1.

Theorem 8.2.7 (The corner $(\gamma_1, \gamma_2) = (-1/2, -1/2)$).

Let $Z_{\gamma_1, \gamma_2}(t)$, $(\gamma_1, \gamma_2) \in \Delta$, be defined in (8.6) with $A = A(\gamma_1, \gamma_2)$ in (8.21). Suppose that $\gamma_1 \geq \gamma_2$. If $(\gamma_1, \gamma_2) \rightarrow (-1/2, -1/2)$ in such a way that

$$\frac{\gamma_1 + 1/2}{\gamma_2 + 1/2} \rightarrow \rho \in [0, 1], \quad (8.11)$$

then

$$\begin{aligned} Z_{\gamma_1, \gamma_2}(t) &\Rightarrow Y_\rho(t) \\ &= t \cdot \left[\frac{(\rho + 1)^{-1} + (2\sqrt{\rho})^{-1}}{\sqrt{2(\rho + 1)^{-2} + (2\rho)^{-1}}} \cdot X_1 + \frac{(\rho + 1)^{-1} - (2\sqrt{\rho})^{-1}}{\sqrt{2(\rho + 1)^{-2} + (2\rho)^{-1}}} \cdot X_2 \right], \end{aligned} \quad (8.12)$$

where X_1 and X_2 two independent standardized chi-squared random variables with one degree of freedom (with mean 0 and variance 1). The case $\rho = 0$ is understood as the limit as $\rho \rightarrow 0$.

Remark 8.2.8. Since by (8.5), the self-similarity parameter H equals $\gamma_1 + \gamma_2 + 2$, we get that H tends to 1 as $(\gamma_1, \gamma_2) \rightarrow (-1/2, -1/2)$. It is known (see e.g., Theorem 3.1.1 of Embrechts and Maejima [2002]) that the only self-similar finite-variance processes with stationary increments having $H = 1$ are degenerate processes. We see this in Theorem 8.2.7, where the limit is a random variable multiplied by t .

Remark 8.2.9. In Theorem 8.2.7, if $\rho = 1$, $Y_\rho(t)$ reduces to tX_1 , where X_1 is a standardized chi-squared random variable with one degree of freedom. Consider now the standardized Rosenblatt process $Z_\gamma(t)$ in (8.3). In this case, $\gamma_1 = \gamma_2 = \gamma$ and thus $\rho = 1$, which corresponds to the middle line m in Figure 8.5. From Theorem 8.2.7, we conclude that if $\gamma \rightarrow -1/2$, then the limit is tX_1 . This is consistent with a previous result of Veillette and Taqqu [2013], that the limit is a standardized chi-squared random variable when $t = 1$.

Remark 8.2.10. If $\rho = 0$, $Y_\rho(t) = \frac{t}{\sqrt{2}}(X_1 - X_2)$, which has the same distribution as $t(WB)$, where W and B are two independent standard normal random variables (see (8.31) below). This is consistent with Theorem 8.2.2, where on the edge e_1 the limit is $WB_{\gamma+3/2}$. This tends, as $\gamma \rightarrow -1/2$, to $W \cdot B_1(t) = W \cdot B \cdot t = t(WB)$, where B is a standard Gaussian random variable.

Remark 8.2.11. Theorems 8.2.1 to 8.2.7 are consistent with Theorem 3.1 of Nourdin and Poly [2012], stating that the limit of a double Wiener-Itô integral can only be a linear combination of a normal and an independent double Wiener-Itô integral.

Remark 8.2.12. Theorem 8.2.5 and 8.2.7 concern the limit behavior of $Z_{\gamma_1, \gamma_2}(t)$ as (γ_1, γ_2) approaches the corners along some straight-line direction. What happens if one does not approach the corners following a straight-line direction? Then, there will be no convergence. To see this, consider the case of Theorem 8.2.5 (a similar argument can be made for Theorem 8.2.7). Let

$$\rho(\gamma_1, \gamma_2) = \frac{\gamma_1 + \gamma_2 + 3/2}{\gamma_2 + 1} \in (0, 1)$$

parameterize the straight-line direction. Suppose that $\rho(\gamma_1, \gamma_2)$ does not converge as (γ_1, γ_2) approaches the corner $(-\frac{1}{2}, -1)$. Then there are two subsequences of (γ_1, γ_2) , such that $\rho(\gamma_1, \gamma_2)$ of the first subsequence converges to ρ_1 and $\rho(\gamma_1, \gamma_2)$ of the second subsequence converges to ρ_2 , with $\rho_1 \neq \rho_2$. By Theorem 8.2.5, the corresponding processes $Z_{\gamma_1, \gamma_2}(t)$ converge to two different limits. Therefore, the original process $Z_{\gamma_1, \gamma_2}(t)$ does not converge if (γ_1, γ_2) does not follow a straight-line direction.

8.3 Proof of the main theorems

Since we will use a method of moments, we state first a cumulant formula for a linear combination of $Z_{\gamma_1, \gamma_2}(t)$ at finite time points. We let $\kappa_m(\cdot)$ denote the m -th cumulant. In the following proposition, the constant A in (8.6) is arbitrary.

Proposition 8.3.1. *The m -th cumulant ($m \geq 2$) of $\sum_{i=1}^n c_i Z_{\gamma_1, \gamma_2}(t_i)$, $c_i \in \mathbb{R}$, $t_i \in [0, \infty)$, equals*

$$\kappa_m \left(\sum_{i=1}^n c_i Z_{\gamma_1, \gamma_2}(t_i) \right) = \frac{1}{2} (m-1)! A^m C_m(\gamma_1, \gamma_2; \mathbf{t}, \mathbf{c}), \quad (8.13)$$

where

$$\begin{aligned} C_m(\gamma_1, \gamma_2; \mathbf{t}, \mathbf{c}) = & \sum_{\sigma \in \{1, 2\}^m} \sum_{i_1, \dots, i_m=1}^n c_{i_1} \dots c_{i_m} \int_0^{t_{i_1}} ds_1 \dots \int_0^{t_{i_m}} ds_m \\ & \prod_{j=1}^m \left[(s_j - s_{j-1})_+^{\gamma_{\sigma_j} + \gamma_{\sigma'_{j-1}} + 1} \mathbb{B}(\gamma_{\sigma'_{j-1}} + 1, -\gamma_{\sigma_j} - \gamma_{\sigma'_{j-1}} - 1) \right. \\ & \left. + (s_{j-1} - s_j)_+^{\gamma_{\sigma_j} + \gamma_{\sigma'_{j-1}} + 1} \mathbb{B}(\gamma_{\sigma_j} + 1, -\gamma_{\sigma_j} - \gamma_{\sigma'_{j-1}} - 1) \right], \quad (8.14) \end{aligned}$$

where

$$B(x, y) = \int_0^1 u^{x-1}(1-u)^{y-1} du = \int_0^\infty w^{x-1}(1+w)^{-x-y} dw, \quad x, y > 0, \quad (8.15)$$

is the beta function, the sum runs over $\sigma = (\sigma_1, \dots, \sigma_m)$ with $\sigma_i = 1$ or 2 , and σ' is the complement of σ , namely, $\sigma'_i = 1$ if $\sigma_i = 2$ and $\sigma'_i = 2$ if $\sigma_i = 1$, $i = 1, \dots, m$. Moreover $\sigma'_0 = \sigma'_m$ and $s_0 = s_m$, $i = 1, \dots, m$.

Proposition 8.3.1 is an extension of Theorem 2.1 of Bai and Taqu [2014b]. We shall use the following cumulant formula for a double Wiener-Itô integral (see, e.g., (8.4.3) of Nourdin and Peccati [2012]):

Lemma 8.3.2. *If f is a symmetric function in $L^2(\mathbb{R}^2)$, then the m -th cumulant of the double Wiener-Itô integral $X = \int_{\mathbb{R}^2}' f(y_1, y_2) B(dy_1) B(dy_2)$ is given by the following circular integral:*

$$\kappa_m(X) = 2^{m-1}(m-1)! \int_{\mathbb{R}^m} f(y_1, y_2) f(y_2, y_3) \dots f(y_{m-1}, y_m) f(y_m, y_1) dy_1 \dots dy_m.$$

Proof of Proposition 8.3.1. Set

$$g(x, y) = \frac{A}{2} (x_+^{\gamma_1} y_+^{\gamma_2} + x_+^{\gamma_2} y_+^{\gamma_1}).$$

Let

$$h_t(x, y) = \int_0^t g(s-x, s-y) ds,$$

and observe that h_t is symmetric. So using the linearity of the Wiener-Itô integral and Lemma 8.3.2, we have

$$\kappa_m \left(\sum_{i=1}^n c_i Z_{\gamma_1, \gamma_2}(t_i) \right)$$

$$\begin{aligned}
&= \kappa_m \left(\int_{\mathbb{R}^2}' \sum_{i=1}^n c_i h_{t_i}(x_1, x_2) B(dx_1) B(dx_2) \right) \\
&= 2^{m-1} (m-1)! \int_{\mathbb{R}^m} d\mathbf{x} \prod_{j=1}^m \left[\sum_{i=1}^n c_i h_{t_i}(x_j, x_{j+1}) \right] \\
&= 2^{m-1} (m-1)! \sum_{i_1, \dots, i_m=1}^n c_{i_1} \dots c_{i_m} \int_{\mathbb{R}^m} d\mathbf{x} \prod_{j=1}^m \int_0^{t_{i_j}} g(s_j - x_j, s_j - x_{j+1}) ds_j,
\end{aligned}$$

and hence

$$\begin{aligned}
\kappa_m \left(\sum_{i=1}^n c_i Z_{\gamma_1, \gamma_2}(t_i) \right) &= \frac{1}{2} (m-1)! A^m \sum_{i_1, \dots, i_m=1}^n c_{i_1} \dots c_{i_m} \times \\
&\int_0^{t_{i_1}} ds_1 \dots \int_0^{t_{i_m}} ds_m \left(\int_{\mathbb{R}^m} \prod_{j=1}^m [(s_j - x_j)_+^{\gamma_1} (s_j - x_{j+1})_+^{\gamma_2} + (s_j - x_j)_+^{\gamma_2} (s_j - x_{j+1})_+^{\gamma_1}] d\mathbf{x} \right),
\end{aligned} \tag{8.16}$$

where we view the index j as modulo m , e.g., $x_{m+1} = x_1$.

Then using the notation in the statement of Proposition 8.3.1, one has

$$\begin{aligned}
I &:= \int_{\mathbb{R}^m} \prod_{j=1}^m [(s_j - x_j)_+^{\gamma_1} (s_j - x_{j+1})_+^{\gamma_2} + (s_j - x_j)_+^{\gamma_2} (s_j - x_{j+1})_+^{\gamma_1}] d\mathbf{x} \\
&= \sum_{\sigma \in \{1,2\}^m} \int_{\mathbb{R}^m} \prod_{j=1}^m (s_j - x_j)_+^{\gamma_{\sigma_j}} (s_j - x_{j+1})_+^{\gamma_{\sigma'_j}} d\mathbf{x} \\
&= \sum_{\sigma \in \{1,2\}^m} \int_{\mathbb{R}^m} \prod_{j=1}^m (s_j - x_j)_+^{\gamma_{\sigma_j}} (s_{j-1} - x_j)_+^{\gamma_{\sigma'_{j-1}}} d\mathbf{x},
\end{aligned}$$

and thus

$$\begin{aligned}
I &= \sum_{\sigma \in \{1,2\}^m} \prod_{j=1}^m \left[(s_j - s_{j-1})_+^{\gamma_{\sigma_j} + \gamma_{\sigma'_{j-1}} + 1} \mathbf{B}(\gamma_{\sigma'_{j-1}} + 1, -\gamma_{\sigma_j} - \gamma_{\sigma'_{j-1}} - 1) \right. \\
&\quad \left. + (s_{j-1} - s_j)_+^{\gamma_{\sigma_j} + \gamma_{\sigma'_{j-1}} + 1} \mathbf{B}(\gamma_{\sigma_j} + 1, -\gamma_{\sigma_j} - \gamma_{\sigma'_{j-1}} - 1) \right],
\end{aligned} \tag{8.17}$$

where we have used the following relation valid for $a, b \in (-1, -1/2)$:

$$\int_{\mathbb{R}} (s_1 - u)_+^a (s_2 - u)_+^b du = (s_2 - s_1)_+^{a+b+1} B(a+1, -a-b-1) + (s_1 - s_2)_+^{a+b+1} \mathcal{B}(b+1, -a-b-1). \quad (8.18)$$

(See Lemma 3.2 of Bai and Taqqu [2014b].) Substituting (8.17) into (8.16), equation (8.13) is obtained. \square

Note that $EZ_{\gamma_1, \gamma_2}(1) = 0$ by the property of Wiener-Itô integral, and hence the second and the third moments coincide with the second and the third cumulants. As two special cases of Proposition 8.3.1, one has the following explicit formulas for the second and the third moment of the generalized Rosenblatt distribution (Bai and Taqqu [2014b], Theorem 2.1):

The second moment of $Z_{\gamma_1, \gamma_2}(1)$ is

$$\begin{aligned} \mu_2(\gamma_1, \gamma_2) &= \frac{A^2}{(\gamma_1 + \gamma_2 + 2)(2(\gamma_1 + \gamma_2) + 3)} \\ &\quad \times \left[\mathcal{B}(\gamma_1 + 1, -\gamma_1 - \gamma_2 - 1) \mathcal{B}(\gamma_2 + 1, -\gamma_1 - \gamma_2 - 1) \right. \\ &\quad \left. + \mathcal{B}(\gamma_1 + 1, -2\gamma_1 - 1) \mathcal{B}(\gamma_2 + 1, -2\gamma_2 - 1) \right], \end{aligned} \quad (8.19)$$

The third moment of $Z_{\gamma_1, \gamma_2}(1)$ is

$$\begin{aligned} \mu_3(\gamma_1, \gamma_2) &= \frac{2A^3}{(\gamma_1 + \gamma_2 + 2)(3(\gamma_1 + \gamma_2) + 5)} \\ &\quad \times \left[\sum_{\sigma \in \{1, 2\}^3} B(\gamma_{\sigma_1} + 1, -\gamma_{\sigma_1} - \gamma_{\sigma'_3} - 1) B(\gamma_{\sigma'_1} + 1, -\gamma_{\sigma'_1} - \gamma_{\sigma_2} - 1) \right. \\ &\quad \left. \times B(\gamma_{\sigma'_2} + 1, -\gamma_{\sigma'_2} - \gamma_{\sigma_3} - 1) B(\gamma_{\sigma'_1} + \gamma_{\sigma_2} + 2, \gamma_{\sigma'_2} + \gamma_{\sigma_3} + 2) \right]. \end{aligned} \quad (8.20)$$

To standardize $Z_{\gamma_1, \gamma_2}(t)$, we set $\mu_2(\gamma_1, \gamma_2) = 1$. By (8.19), this determines the constant A as:

$$A(\gamma_1, \gamma_2) = \left[(\gamma_1 + \gamma_2 + 2)(2(\gamma_1 + \gamma_2) + 3) \right]^{1/2}$$

$$\begin{aligned} & \times \left[\mathcal{B}(\gamma_1 + 1, -\gamma_1 - \gamma_2 - 1) \mathcal{B}(\gamma_2 + 1, -\gamma_1 - \gamma_2 - 1) \right. \\ & \quad \left. + \mathcal{B}(\gamma_1 + 1, -2\gamma_1 - 1) \mathcal{B}(\gamma_2 + 1, -2\gamma_2 - 1) \right]^{-1/2}. \end{aligned} \quad (8.21)$$

8.3.1 Proof of Theorem 8.2.1

We will use a result for bounding integral of powers of linear functions in Euclidean space. First some notation. Let $L_1(\mathbf{s}) = \langle \mathbf{w}_1, \mathbf{s} \rangle, \dots, L_m(\mathbf{s}) = \langle \mathbf{w}_m, \mathbf{s} \rangle$ be linear functions on \mathbb{R}^n , where $\langle \cdot, \cdot \rangle$ denotes the Euclidean inner product. Let

$$P(\mathbf{s}) = \prod_{j=1}^m |L_j(\mathbf{s})|^{\alpha_j}.$$

Set $T = \{\mathbf{w}_1, \dots, \mathbf{w}_m\}$. For any nonempty $W \subset T$, define

$$S(W) = T \cap \text{span}\{W\}, \quad (8.22)$$

where $\text{span}\{W\}$ denotes linear subspace spanned by W , and define the quantity

$$d(P, W) = |W| + \sum_{j: \mathbf{w}_j \in S(W)} \alpha_j,$$

where $|W|$ is the cardinality of the set W . Then we have the following so-called power counting lemma:

Lemma 8.3.3 (Theorem 3.1 of Fox and Taqqu [1987]). *Suppose that*

$$d(P, W) > 0. \quad (8.23)$$

for any $W \subset T$ which consists of linearly independent \mathbf{w}_j 's¹. Then

$$\int_{[0,1]^n} P(\mathbf{s})d\mathbf{s} < \infty.$$

Lemma 8.3.4. *The function*

$$f(\alpha_1, \dots, \alpha_m) := \int_{[0,1]^m} |s_1 - s_m|^{\alpha_1} |s_2 - s_1|^{\alpha_2} \dots |s_m - s_{m-1}|^{\alpha_m} d\mathbf{s} \quad (8.24)$$

is finite and continuous on the domain

$$D = \left\{ (\alpha_1, \dots, \alpha_m) : \alpha_i > -1, \sum_{i=1}^m \alpha_i + m > 1 \right\}. \quad (8.25)$$

Proof. We first show that $f(\alpha_1, \dots, \alpha_m) < \infty$ on D using Lemma 8.3.3. Following the notation introduced for the lemma, we have $L_1(\mathbf{s}) = s_1 - s_m$, $L_2(\mathbf{s}) = s_2 - s_1, \dots$, $L_m(\mathbf{s}) = s_m - s_{m-1}$, and hence $\mathbf{w}_1 = (1, 0, \dots, 0, -1)$, $\mathbf{w}_2 = (-1, 1, 0, \dots, 0)$, \dots , $\mathbf{w}_m = (0, \dots, 0, -1, 1)$ and $T = \{\mathbf{w}_1, \dots, \mathbf{w}_m\}$.

It is easy to see that a subset $W \subset T$ consists of linearly independent \mathbf{w}_j 's if and only if $|W| \leq m - 1$. When $|W| \leq m - 2$, the set $S(W)$ defined in (8.22) is equal to W . The condition (8.23) is satisfied in this case because each $\alpha_j > -1$ and hence

$$D(P, W) = |W| + \sum_{j:\mathbf{w}_j \in S(W)} \alpha_j > |W| + \sum_{j:\mathbf{w}_j \in W} (-1) = |W| - |W| = 0.$$

When $|W| = m - 1$, one has $\text{span}(W) = T$, and hence $S(W) = T$. Thus the condition (8.23) in this case becomes

$$D(P, W) = m - 1 + \sum_{i=1}^m \alpha_i > 0,$$

¹Theorem 3.1 of Fox and Taqqu [1987] states that it is enough to consider $W \subset T$ consisting of linearly independent \mathbf{w}_j 's with *negative* exponent α_j 's. This is because the non-negative exponents α_j cannot make the integral $\int_{[0,1]^n} P(\mathbf{s})d\mathbf{s}$ blow up.

which is satisfied in view of (8.25). Hence the integral $f(\alpha_1, \dots, \alpha_m)$ in (8.24) is finite by Lemma 8.3.3.

To verify the continuity of $f(\alpha_1, \dots, \alpha_m)$, suppose that as $n \rightarrow \infty$, $\alpha_n \rightarrow \alpha := (\alpha_1, \dots, \alpha_m)$. Then for large n , $\alpha_n \geq \alpha_\epsilon := (\alpha_1 - \epsilon, \dots, \alpha_m - \epsilon)$, where the small ϵ is chosen such that $\alpha_\epsilon \in D$. Denote the integrand in (8.24) by $I(\mathbf{s}; \alpha)$, and recall that $I(\mathbf{s}; \alpha)$ is decreasing in every component of α . Hence when n is large, $I(\mathbf{s}; \alpha_n) \leq I(\mathbf{s}; \alpha_\epsilon)$. Since $I(\mathbf{s}; \alpha_\epsilon)$ is integrable, we can apply the Dominated Convergence Theorem to obtain the convergence $f(\alpha_n) \rightarrow f(\alpha)$ as $n \rightarrow \infty$, proving the continuity. \square

In the following corollary, the exponents are supposed to be away from the boundary of the set D defined in (8.25).

Corollary 8.3.5. *Let C_1, C_2 be two fixed constants such that $C_1 > -1$ and $C_2 > 1$. Then the function $f(\alpha_1, \dots, \alpha_m)$ defined in (8.24) is bounded on the domain*

$$D(C_1, C_2) = \left\{ (\alpha_1, \dots, \alpha_m) : \alpha_i \geq C_1, \sum_{i=1}^m \alpha_i + m \geq C_2 \right\}.$$

Proof. Let M be a large positive constant. Define

$$\begin{aligned} D_M(C_1, C_2) &= D(C_1, C_2) \cap (-\infty, M]^m \\ &= \left\{ (\alpha_1, \dots, \alpha_m) : C_1 \leq \alpha_i \leq M, \sum_{i=1}^m \alpha_i + m \geq C_2 \right\}. \end{aligned}$$

Since $D_M(C_1, C_2)$ is a compact subset of D in (8.25), and $f(\alpha_1, \dots, \alpha_m)$ is continuous on D by Lemma 8.3.4, we deduce that f is bounded on $D_M(C_1, C_2)$. The boundedness on $D(C_1, C_2)$ follows since f decreases when any α_i increases. \square

Lemma 8.3.6. *Let $A(\gamma_1, \gamma_2)$ be as in (8.21), where $(\gamma_1, \gamma_2) \in \Delta$ which is defined in (8.2). Then there exists a constant $C > 0$ independent of γ_1 and γ_2 such that*

$$|A(\gamma_1, \gamma_2)| \leq C[2(\gamma_1 + \gamma_2) + 3]^{1/2}.$$

Proof. This is immediate by noting that the beta function $B(x, y)$ defined in (8.15) is decreasing in x and in y . Since in addition Δ is a bounded region, the beta functions in (8.21) are bounded from below, and hence the factor with negative power $-1/2$ in (8.21) is bounded from above. \square

The following hypercontractivity inequality for multiple Wiener-Itô integral (see, e.g., Corollary 5.6 of Major [2014] or Theorem 2.7.2 of Nourdin and Peccati [2012]) is useful:

Lemma 8.3.7. *For any $m \in \mathbb{Z}_+$, there exists a constant $C_m > 0$, such that*

$$\mathbb{E}|I_k(f)|^{2m} \leq C_m (\mathbb{E}|I_k(f)|^2)^m, \quad \text{for all } f \in L^2(\mathbb{R}^k).$$

Tightness of standardized $Z_{\gamma_1, \gamma_2}(t)$ in $C[0, 1]$ will follow from the following lemma:

Lemma 8.3.8. *Let $Z_{\gamma_1, \gamma_2}(t)$ be as in (8.6) with A as in (8.21) and (γ_1, γ_2) in the region Δ defined in (8.2). Then there exists a constant $C > 0$ which does not depend on γ_1, γ_2 , such that for all $0 \leq s \leq t \leq 1$,*

$$\mathbb{E}|Z_{\gamma_1, \gamma_2}(t) - Z_{\gamma_1, \gamma_2}(s)|^4 \leq C(t - s)^2,$$

which implies that the law of $\{Z_{\gamma_1, \gamma_2}(t) : (\gamma_1, \gamma_2) \in \Delta\}$ is tight in $C[0, 1]$.

Proof. Using Lemma 8.3.7, self-similarity and stationary-increment property of $Z_{\gamma_1, \gamma_2}(t)$, one has

$$\begin{aligned} \mathbb{E}|Z_{\gamma_1, \gamma_2}(t) - Z_{\gamma_1, \gamma_2}(s)|^4 &\leq C_2 (\mathbb{E}|Z_{\gamma_1, \gamma_2}(t) - Z_{\gamma_1, \gamma_2}(s)|^2)^2 \\ &= C_2(t - s)^{4H} \leq C_2(t - s)^2, \end{aligned}$$

where $H := \gamma_1 + \gamma_2 + 2 \geq 1/2$ and $0 \leq t - s \leq 1$. So $Z_{\gamma_1, \gamma_2}(t)$ by Kolmogorov's criterion admits a continuous version. Tightness follows from, e.g., Prokhorov [1956] Lemma 2.2. \square

We now prove Theorem 8.2.1. By Lemma 8.3.8, tightness in $C[0, 1]$ holds. We are left

to show convergence of finite-dimensional distributions ($\xrightarrow{f.d.d.}$). From here on, we let C and c denote constants whose values can change from line to line.

Proof of $\xrightarrow{f.d.d.}$ Theorem 8.2.1. Due to self-similarity and stationary increments, the covariance of the standardized $Z_{\gamma_1, \gamma_2}(t)$ is

$$\mathbb{E}Z_{\gamma_1, \gamma_2}(s)Z_{\gamma_1, \gamma_2}(t) = \frac{1}{2} \left(s^{2\gamma_1+2\gamma_2+4} + t^{2\gamma_1+2\gamma_2+4} - |s-t|^{2\gamma_1+2\gamma_2+4} \right), \quad t, s \geq 0,$$

which converges to the Brownian motion covariance $\mathbb{E}B(s)B(t) = s \wedge t = \frac{1}{2}(s+t-|s-t|)$ as $\gamma_1 + \gamma_2 \rightarrow -3/2$. By using the method of moments, it is sufficient to show that

$$\kappa_m \left(\sum_{i=1}^n c_i Z_{\gamma_1, \gamma_2}(t_i) \right) \rightarrow 0, \quad m \geq 3. \quad (8.26)$$

As $\gamma_1 + \gamma_2 \rightarrow -3/2$, the factor $A(\gamma_1, \gamma_2)$ in (8.21) converges to zero by Lemma 8.3.6. It is therefore sufficient to show that for $m \geq 3$, and $\gamma_1, \gamma_2 > -1 + \epsilon$, the factor $C_m(\gamma_1, \gamma_2; \mathbf{t}, \mathbf{c})$ in (8.14) is bounded.

Under the constraints $\gamma_1 + \gamma_2 \geq -3/2$ and $\gamma_1, \gamma_2 > -1 + \epsilon$ (or equivalently $\gamma_1, \gamma_2 < -1/2 - \epsilon$), the factors $\mathbb{B}(\gamma_{\sigma'_j-1} + 1, -\gamma_{\sigma_j} - \gamma_{\sigma'_j-1} - 1)$ and $\mathbb{B}(\gamma_{\sigma_j} + 1, -\gamma_{\sigma_j} - \gamma_{\sigma'_j-1} - 1)$ are bounded by a constant $C > 0$ for any σ and j . This is because the beta function $\mathbb{B}(x, y)$ defined in (8.15) is bounded if both x and y stay away from a neighborhood of 0. Choosing $T \geq \max(t_1, \dots, t_n)$, one then has

$$\begin{aligned} |C_m(\gamma_1, \gamma_2; \mathbf{t}, \mathbf{c})| &\leq C \sum_{\sigma \in \{1, 2\}^m} \int_{[0, T]^m} d\mathbf{s} \prod_{j=1}^m |s_j - s_{j-1}|^{\gamma_{\sigma_j} + \gamma_{\sigma'_j-1} + 1} \\ &\leq C \sum_{\sigma \in \{1, 2\}^m} \int_{[0, 1]^m} d\mathbf{s} \prod_{j=1}^m |s_j - s_{j-1}|^{\gamma_{\sigma_j} + \gamma_{\sigma'_j-1} + 1}, \end{aligned}$$

where the last constant C depends on T , m and ϵ .

We now want to apply Corollary 8.3.5 to establish the boundedness of each of the term

in the preceding sum. Using the notation in Lemma 8.3.4, we set

$$\alpha_j = \gamma_{\sigma_j} + \gamma_{\sigma'_{j-1}} + 1.$$

Recall that γ_{σ_j} and $\gamma_{\sigma'_{j-1}}$ are either γ_1 or γ_2 and $\gamma_{\sigma_j} + \gamma_{\sigma'_j} = \gamma_1 + \gamma_2$. Now since $\gamma_1 + \gamma_2 \geq -3/2$ and $\gamma_j \geq -1 + \epsilon$, we have

$$\alpha_j \geq \begin{cases} 2\gamma_j + 1 \geq -1 + 2\epsilon, & \text{if } \sigma'_{j-1} = \sigma_j; \\ \gamma_1 + \gamma_2 + 1 \geq -3/2 + 1 = -1/2, & \text{if } \sigma'_{j-1} \neq \sigma_j; \end{cases}$$

We get $\alpha_j \geq C_1 := -1 + 2\epsilon > -1$.

On the other hand, when $m \geq 3$,

$$\sum_{i=1}^m \alpha_i + m = m(\gamma_1 + \gamma_2) + 2m \geq m(-3/2) + 2m = \frac{m}{2} \geq C_2 := \frac{3}{2} > 1.$$

So Corollary 8.3.5 can be applied to deduce the boundedness of $|C_m(\gamma_1, \gamma_2; \mathbf{t}, \mathbf{c})|$ when $\gamma_1, \gamma_2 \geq -1 + \epsilon$, and the proof is thus concluded. \square

Remark 8.3.9. Theorem 8.2.1 involves convergence to a Gaussian process. In this case, according to the results of Nualart and Peccati [2005] and Peccati and Tudor [2005], it suffices to show that (8.26) holds for $m = 4$ and $n = 1$. Focusing on the fourth cumulant, the covariance structure, and the one-dimensional distribution, however, does not simplify significantly the proof as can be seen by examining the proof of Theorem 8.2.1.

8.3.2 Proof of Theorem 8.2.2

Lemma 8.3.10. *Suppose that $\alpha > -1$, then for any $t_1, t_2 \in \mathbb{R}$,*

$$\int_0^{t_1} \int_0^{t_2} |x_1 - x_2|^\alpha dx_1 dx_2 = \frac{1}{(\alpha + 1)(\alpha + 2)} (|t_1|^{\alpha+2} + |t_2|^{\alpha+2} - |t_1 - t_2|^{\alpha+2}).$$

Proof. Suppose $0 < t_1 \leq t_2$. The other cases are similar. Then

$$\begin{aligned}
& \int_0^{t_1} \int_0^{t_2} |x_1 - x_2|^\alpha dx_1 dx_2 \\
&= \int_0^{t_1} \int_0^{t_1} |x_1 - x_2|^\alpha dx_1 dx_2 + \int_0^{t_1} \int_{t_1}^{t_2} (x_2 - x_1)^\alpha dx_2 dx_1 \\
&= \frac{2}{(\alpha + 1)(\alpha + 2)} t_1^{\alpha+2} + \frac{1}{(\alpha + 1)(\alpha + 2)} [t_2^{\alpha+2} - t_1^{\alpha+2} - (t_2 - t_1)^{\alpha+2}] \\
&= \frac{1}{(\alpha + 1)(\alpha + 2)} [t_1^{\alpha+2} + t_2^{\alpha+2} - (t_2 - t_1)^{\alpha+2}].
\end{aligned}$$

□

Below the notation $A \sim B$ means asymptotic equivalence, namely, the ratio A/B converges to 1. We include first a fact about the asymptotics of the beta function $\mathcal{B}(\cdot, \cdot)$ when one of the exponents approaches the boundary.

Lemma 8.3.11. *Let $0 < b_0 < b_1 < \infty$. Then as $\alpha \rightarrow 0$, we have*

$$\alpha \mathcal{B}(\alpha, \beta) \rightarrow 1$$

uniformly in $\beta \in [b_0, b_1]$. Since the beta function is symmetric, we also have $\alpha \mathcal{B}(\beta, \alpha) \rightarrow 1$ as $\alpha \rightarrow 0$ uniformly in $\beta \in [b_0, b_1]$.

Proof. Assume without loss of generality that $b_0 \leq 1 \leq b_1$. Fix any small $\epsilon > 0$. Then

$$\mathcal{B}(\alpha, \beta) = \int_0^\epsilon x^{\alpha-1} (1-x)^{\beta-1} dx + \int_\epsilon^1 x^{\alpha-1} (1-x)^{\beta-1} dx =: I_1(\alpha, \beta; \epsilon) + I_2(\alpha, \beta; \epsilon). \quad (8.27)$$

For $I_1(\alpha, \beta; \epsilon)$, we have

$$\begin{aligned}
\alpha^{-1} \epsilon^\alpha (1-\epsilon)^{b_1-1} &= \int_0^\epsilon x^{\alpha-1} dx (1-\epsilon)^{b_1-1} \leq \\
I_1(\alpha, \beta; \epsilon) &\leq \int_0^\epsilon x^{\alpha-1} dx (1-\epsilon)^{b_0-1} = \alpha^{-1} \epsilon^\alpha (1-\epsilon)^{b_0-1}.
\end{aligned}$$

This yields that

$$(1 - \epsilon)^{b_1 - 1} \leq \liminf_{\alpha \rightarrow 0, \beta \in [b_0, b_1]} \alpha I_1(\alpha, \beta, \epsilon) \leq \limsup_{\alpha \rightarrow 0, \beta \in [b_0, b_1]} \alpha I_1(\alpha, \beta, \epsilon) \leq (1 - \epsilon)^{b_0 - 1}. \quad (8.28)$$

For $I_2(\alpha, \beta; \epsilon)$, it is uniformly bounded with respect to $\alpha \leq 1$ and β as follows:

$$I_2(\alpha, \beta; \epsilon) \leq \epsilon^{\alpha - 1} \int_{\epsilon}^1 (1 - x)^{\beta - 1} dx = \epsilon^{\alpha - 1} \beta^{-1} (1 - \epsilon)^{\beta} \leq \epsilon^{-1} b_0^{-1} (1 - \epsilon)^{b_0}. \quad (8.29)$$

Combining (8.27), (8.28) and (8.29), we get

$$(1 - \epsilon)^{b_1 - 1} \leq \liminf_{\alpha \rightarrow 0, \beta \in [b_0, b_1]} \alpha \mathcal{B}(\alpha, \beta) \leq \limsup_{\alpha \rightarrow 0, \beta \in [b_0, b_1]} \alpha \mathcal{B}(\alpha, \beta) \leq (1 - \epsilon)^{b_0 - 1}.$$

Since ϵ is arbitrary, we get that $\alpha \mathcal{B}(\alpha, \beta) \rightarrow 1$ as $\alpha \rightarrow 0$. □

The limit $\alpha \mathcal{B}(\alpha, \beta) \rightarrow 1$ as $\alpha \rightarrow 0$ will be used extensively, mostly in the form

$$\mathcal{B}(\alpha, \beta) \sim \alpha^{-1} \rightarrow \infty.$$

Lemma 8.3.12. *Let $WB_{\gamma+3/2}(t)$ be the process given as Theorem 8.2.2. We also include the case $\gamma = -1$ where $B_{\gamma+3/2}(t) = B_{1/2}(t)$ is Brownian motion. Then the m -th cumulant of the linear combination of $WB_{\gamma+3/2}(t)$ at different time points is given by*

$$\begin{aligned} & \kappa_m \left(\sum_{i=1}^n c_i WB_{\gamma+3/2}(t_i) \right) = \\ & (m-1)! \left[\sum_{i_1, i_2=1}^n \frac{c_{i_1} c_{i_2}}{2} (|t_{i_1}|^{2\gamma+3} + |t_{i_2}|^{2\gamma+3} - |t_{i_1} - t_{i_2}|^{2\gamma+3}) \right]^{m/2} \end{aligned} \quad (8.30)$$

if m is even, and 0 if m is odd.

Proof.

$$\sum_{i=1}^n c_i WB_{\gamma+3/2}(t_i) = W \sum_{i=1}^n c_i B_{\gamma+3/2}(t_i) = \sigma WZ,$$

where Z is a standard normal random variable which is independent of W , and

$$\begin{aligned}\sigma &= \left(\text{Var} \left[\sum_{i=1}^n c_i B_{\gamma+3/2}(t_i) \right] \right)^{1/2} = \left[\text{E} \sum_{i_1, i_2=1}^n c_{i_1} c_{i_2} B_{\gamma+3/2}(t_{i_1}) B_{\gamma+3/2}(t_{i_2}) \right]^{1/2} \\ &= \left[\sum_{i_1, i_2=1}^n \frac{c_{i_1} c_{i_2}}{2} (|t_{i_1}|^{2\gamma+3} + |t_{i_2}|^{2\gamma+3} - |t_{i_1} - t_{i_2}|^{2\gamma+3}) \right]^{1/2},\end{aligned}$$

using the covariance of fractional Brownian motion. Then note that

$$WZ = \frac{1}{2} \left[\left(\frac{W+Z}{\sqrt{2}} \right)^2 - \left(\frac{W-Z}{\sqrt{2}} \right)^2 \right], \quad (8.31)$$

where $Z_1^2 := \left[\frac{W+Z}{\sqrt{2}} \right]^2$ and $Z_2^2 := \left[\frac{W-Z}{\sqrt{2}} \right]^2$ are two independent χ_1^2 (chi-squared random variables with one degree of freedom). The independence is due to the fact that $Z+W$ and $Z-W$ are uncorrelated. Since the m -th cumulant of a χ_1^2 variable is $2^{m-1}(m-1)!$, and using the scaling property and the additive property of cumulant under independence, we have

$$\begin{aligned}\kappa_m(\sigma WZ) &= \left(\frac{\sigma}{2} \right)^m [\kappa_m(Z_1^2) + (-1)^m \kappa_m(Z_2^2)] \\ &= \left(\frac{\sigma}{2} \right)^m [2^{m-1}(m-1)! + (-1)^m 2^{m-1}(m-1)!],\end{aligned}$$

which is equal to 0 if m is odd, and equal to $\sigma^m(m-1)!$ if m is even, proving (8.30). \square

Remark 8.3.13. Starting with the χ_1^2 characteristic function $\phi(t) = (1 - 2it)^{-1/2}$, it is easy to derive using (8.31) that the characteristic function of the standard product-normal distribution WZ is $\varphi(t) = (1 + t^2)^{-1/2}$.

In view of Lemma 8.3.8, we are left to prove the convergence of the finite-dimensional distributions ($\xrightarrow{f.d.d.}$) in Theorem 8.2.2.

Proof of $\xrightarrow{f.d.d.}$ in Theorem 8.2.2. By the Cramér-Wold device, we need to show as $\gamma_1 \rightarrow$

$-1/2$ and $\gamma_2 \rightarrow \gamma \in (-1/2, -1)$ that

$$\sum_{i=1}^n c_i Z_{\gamma_1, \gamma_2}(t_i) \xrightarrow{d} \sum_{i=1}^n c_i W B_{\gamma+3/2}(t_i).$$

Since $\sum_{i=1}^n c_i W B_{\gamma+3/2}(t_i)$ has an analytic characteristic function (Remark 8.3.13), its distribution is moment-determinate. And hence we can apply a method of moments here. In fact, by Theorem 3.4 of Nourdin and Poly [2012], only a finite number of moments are required to prove convergence in distribution.

The cumulant formula of $\sum_{i=1}^n c_i Z_{\gamma_1, \gamma_2}(t_i)$ is given in Proposition 8.3.1, which involves the factors $A(\gamma_1, \gamma_2)$ in (8.21) (recall that Z_{γ_1, γ_2} is standardized) and $C_m(\gamma_1, \gamma_2; \mathbf{t}, \mathbf{c})$ in (8.14). Assume $m \geq 2$ below.

Examining $A(\gamma_1, \gamma_2)$, by Lemma 8.3.11, one can see that as $\gamma_1 \rightarrow -1/2$ and $\gamma_2 \rightarrow \gamma$,

$$\begin{aligned} A(\gamma_1, \gamma_2)^m &\sim [(\gamma + 3/2)(2\gamma + 2)]^{m/2} [\mathcal{B}(1/2, -\gamma - 1/2)\mathcal{B}(\gamma + 1, -\gamma - 1/2) \\ &\quad + \mathcal{B}(1/2, -2\gamma_1 - 1)\mathcal{B}(\gamma + 1, -2\gamma - 1)]^{-m/2}. \end{aligned}$$

The first two and the fourth beta functions are bounded but the third blows up since

$$\mathcal{B}(1/2, -2\gamma_1 - 1) \sim (-2\gamma_1 - 1)^{-1}$$

as $\gamma_1 \rightarrow -1/2$ by Lemma 8.3.11. Hence as $\gamma_1 \rightarrow -1/2$,

$$\begin{aligned} A(\gamma_1, \gamma_2)^m &\sim [(\gamma + 3/2)(2\gamma + 2)]^{m/2} [\mathcal{B}(1/2, -2\gamma_1 - 1)\mathcal{B}(\gamma + 1, -2\gamma - 1)]^{-m/2} \\ &\sim (-2\gamma_1 - 1)^{m/2} (2\gamma + 3)^{m/2} (\gamma + 1)^{m/2} \mathcal{B}(\gamma + 1, -2\gamma - 1)^{-m/2}, \quad (8.32) \end{aligned}$$

which converges to zero.

On the other hand, in the expression of $C_m(\gamma_1, \gamma_2; \mathbf{t}, \mathbf{c})$ in (8.14), the only factors diverging to ∞ as $\gamma_1 \rightarrow -1/2$ and $\gamma_2 \rightarrow \gamma$ are $\mathcal{B}(\gamma_{\sigma'_j-1} + 1, -\gamma_{\sigma_j} - \gamma_{\sigma'_j-1} - 1)$ and $\mathcal{B}(\gamma_{\sigma_j} + 1, -\gamma_{\sigma_j} - \gamma_{\sigma'_j-1} - 1)$ and only when $\sigma_j = \sigma'_j-1 = 1$, because $-\gamma_{\sigma_j} - \gamma_{\sigma'_j-1} - 1 = -2\gamma_1 - 1 \rightarrow 0$

and hence the beta functions each diverge like $(-2\gamma_1 - 1)^{-1}$ by Lemma 8.3.11. To get the highest order of divergence to ∞ , one chooses $\sigma \in \{1, 2\}^m$ such that $\sigma_j = \sigma'_{j-1} = 1$ happens as many times as possible.

In the case m is odd,

$$\max_{\sigma \in \{1, 2\}^m} \#\{j : \sigma_j = \sigma'_{j-1} = 1, j = 1, \dots, m\} = (m-1)/2,$$

because if $\sigma_j = \sigma'_{j-1} = 1$, then $\sigma'_j = 2$, and we therefore cannot have $\sigma_{j+1} = \sigma'_j = 1$. So

$$C_m(\gamma_1, \gamma_2; \mathbf{t}, \mathbf{c}) \sim cB(1/2, -2\gamma_1 - 1)^{(m-1)/2} \sim c(-2\gamma_1 - 1)^{-(m-1)/2}, \quad (8.33)$$

which diverges to ∞ as $\gamma_1 \rightarrow -1/2$. By (8.32) and (8.33), when m is odd,

$$\kappa_m \left(\sum_{i=1}^n c_i Z_{\gamma_1, \gamma_2}(t_i) \right) = \frac{1}{2} (m-1)! A(\gamma_1, \gamma_2)^m C_m(\gamma_1, \gamma_2; \mathbf{t}, \mathbf{c}) \sim c(-2\gamma_1 - 1)^{1/2} \rightarrow 0. \quad (8.34)$$

When m is even, the sequences σ for which one has the greatest number of j 's such that $\sigma_j = \sigma'_{j-1} = 1$ is

$$\operatorname{argmax}_{\sigma \in \{1, 2\}^m} \#\{j : \sigma_j = \sigma'_{j-1} = 1, j = 1, \dots, m\} = (1, 2, 1, 2, \dots, 1, 2) \text{ or } (2, 1, 2, 1, \dots, 2, 1), \quad (8.35)$$

and one gets maximally $m/2$ number of j 's where $\sigma_j = \sigma'_{j-1} = 1$. The product of the $m/2$ contributing beta factors diverge like $(-2\gamma_1 - 1)^{m/2}$. But since the case m even will yield a nonzero limit, we need to keep track of the multiplicative constants. Because $\sigma = (1, 2, 1, 2, \dots, 1, 2)$ and $\sigma = (2, 1, 2, 1, \dots, 2, 1)$ yield the same term, one has as $\gamma_1 \rightarrow -1/2$ and $\gamma_2 \rightarrow \gamma$ that

$$C_m(\gamma_1, \gamma_2; \mathbf{t}, \mathbf{c}) \sim 2(-2\gamma_1 - 1)^{-m/2} \left[\sum_{i_1, \dots, i_n=1}^n c_{i_1} \dots c_{i_m} \mathbb{B}(\gamma + 1, -2\gamma - 1)^{m/2} \right]$$

$$\begin{aligned}
& \times \int_0^{t_{i_1}} \dots \int_0^{t_{i_m}} |s_1 - s_2|^{2\gamma+1} |s_3 - s_4|^{2\gamma+1} \dots |s_{m-1} - s_m|^{2\gamma+1} ds \\
& = 2(-2\gamma_1 - 1)^{-m/2} (2\gamma + 3)^{-m/2} (\gamma + 1)^{-m/2} B(\gamma + 1, -2\gamma - 1)^{m/2} \\
& \times \left[\sum_{i_1, i_2=1}^n \frac{c_{i_1} c_{i_2}}{2} (|t_{i_1}|^{2\gamma+3} + |t_{i_2}|^{2\gamma+3} - |t_{i_1} - t_{i_2}|^{2\gamma+3}) \right]^{m/2}, \quad (8.36)
\end{aligned}$$

where the asymptotic equivalence \sim in the first line can be justified by the Dominated Convergence Theorem, and the last equality is due to Lemma 8.3.10.

Combining (8.13), (8.32) and (8.36), one gets as $\gamma_1 \rightarrow -1/2$ and $\gamma_2 \rightarrow \gamma$ that for m even,

$$\begin{aligned}
& \kappa_m \left(\sum_{i=1}^n c_i Z_{\gamma_1, \gamma_2}(t_i) \right) \rightarrow \\
& (m-1)! \left[\sum_{i_1, i_2=1}^n \frac{c_{i_1} c_{i_2}}{2} (|t_{i_1}|^{2\gamma+3} + |t_{i_2}|^{2\gamma+3} - |t_{i_1} - t_{i_2}|^{2\gamma+3}) \right]^{m/2}. \quad (8.37)
\end{aligned}$$

The proof is concluded by comparing (8.34) and (8.37) with Lemma 8.3.12. \square

We state a byproduct of the preceding proof which will be used in Section 8.5.

Corollary 8.3.14. *Under the condition and the notation of Theorem 8.2.2, when $m \geq 4$ is even, we have*

$$\kappa_m(Z_{\gamma_1, \gamma_2}(1)) = (m-1)! + O(-\gamma_1 - 1/2).$$

Proof. We are focusing here on the marginal distribution and hence $t = 1$, $c = 1$ and $n = 1$ in (8.14). To get the rate of convergence $O(-\gamma_1 - 1/2)$, we need to expand $C_m(\gamma_1, \gamma_2; 1, 1)$ to a higher order than (8.36). Following the preceding proof of Theorem 8.2.2, we need to consider the σ 's with the second most occurrences of $\sigma'_{j-1} = \sigma_j = 1$. These σ 's have $\sigma'_{j-1} = \sigma_j = 1$ occurring $m/2 - 1$ times instead of $m/2$ times as in (8.35). Adding this type of σ 's into (8.36), we have

$$C_m(\gamma_1, \gamma_2; 1, 1) = c_{\gamma, m} (-\gamma_1 - 1/2)^{-m/2} + O\left((- \gamma_1 - 1/2)^{-m/2+1}\right),$$

where $c_{\gamma,m}$ is the constant given by (8.36) with $t = 1$, $c = 1$ and $n = 1$. By Proposition 8.3.1,

$$\kappa_m(Z_{\gamma_1, \gamma_2}(1)) = \frac{1}{2}(m-1)!A(\gamma_1, \gamma_2)^m C_m(\gamma_1, \gamma_2; 1, 1).$$

So the conclusion follows in view of the expression $A(\gamma_1, \gamma_2)^m$ in (8.32). \square

8.3.3 Proof of Theorem 8.2.5

Lemma 8.3.15. *Let $t_1, \dots, t_m > 0$, and $m \geq 4$ be an even integer. Consider the function:*

$$\begin{aligned} f(a, b; \mathbf{t}) = & \int_0^{t_1} \dots \int_0^{t_m} |x_1 - x_m|^a |x_2 - x_1|^b |x_3 - x_2|^a |x_4 - x_3|^b \dots \\ & \times |x_{m-1} - x_{m-2}|^a |x_m - x_{m-1}|^b d\mathbf{x}, \end{aligned} \quad (8.38)$$

where $-1 < a, b < 0$. Then as $(a, b) \rightarrow (0, -1)$, we have that

$$f(a, b; \mathbf{t}) \sim (b+1)^{-m/2} \prod_{i=2,4,\dots,m} (t_i + t_{i-1} - |t_i - t_{i-1}|).$$

Proof. First, assume without loss of generality that $t_1, \dots, t_m < 1$. Otherwise one can scale them by a change of variables.

We first derive a lower bound for $f(a, b; \mathbf{t})$. Since each $|x_i - x_{i-1}|^a \geq 1$, one has by Lemma 8.3.10 that

$$\begin{aligned} f(a, b; \mathbf{t}) \geq f(0, b; \mathbf{t}) &= \prod_{i=2,4,\dots,m} \int_0^{t_i} \int_0^{t_{i-1}} |x_i - x_{i-1}|^b dx_i dx_{i-1} \\ &= (b+1)^{-m/2} (b+2)^{-m/2} \prod_{i=2,4,\dots,m} \left(t_i^{b+2} + t_{i-1}^{b+2} - |t_i - t_{i-1}|^{b+2} \right) \\ &\sim (b+1)^{-m/2} \prod_{i=2,4,\dots,m} (t_i + t_{i-1} - |t_i - t_{i-1}|) \quad \text{as } b \rightarrow -1. \end{aligned} \quad (8.39)$$

To get an upper bound for $f(a, b; \mathbf{t})$, we apply the Cauchy-Schwarz inequality to break the cyclic structure. In particular in (8.38), view $|x_1 - x_m|^a |x_3 - x_2|^a$ as the integrand, and

treat the other factors as the density of measure. We have

$$f(a, b; \mathbf{t}) \leq \sqrt{f_1(a, b; \mathbf{t})f_2(a, b; \mathbf{t})}, \quad (8.40)$$

where

$$\begin{aligned} f_1(a, b; \mathbf{t}) = & \int_0^{t_1} dx_1 \dots \int_0^{t_m} dx_m |x_1 - x_m|^{2a} |x_2 - x_1|^b |x_4 - x_3|^b |x_5 - x_4|^a \dots \\ & \times |x_{m-1} - x_{m-2}|^a |x_m - x_{m-1}|^b, \end{aligned}$$

and

$$\begin{aligned} f_2(a, b; \mathbf{t}) = & \int_0^{t_1} dx_1 \dots \int_0^{t_m} dx_m |x_3 - x_2|^{2a} |x_2 - x_1|^b |x_4 - x_3|^b |x_5 - x_4|^a \dots \\ & \times |x_{m-1} - x_{m-2}|^a |x_m - x_{m-1}|^b. \end{aligned}$$

Set

$$|x|^a = 1 + h_a(x).$$

Then the integrand in f_1 can be rewritten as

$$[1 + h_{2a}(x_1 - x_m)] |x_2 - x_1|^b |x_4 - x_3|^b [1 + h_a(x_5 - x_4)] \dots [1 + h_a(x_{m-1} - x_{m-2})] |x_m - x_{m-1}|^b.$$

Observe that the product of terms involving neither h_a nor h_{2a} equals $f(0, b; \mathbf{t})$. Hence one can write

$$f_1(a, b; \mathbf{t}) = f(0, b; \mathbf{t}) + R(a, b; \mathbf{t}),$$

where the remainder $R(a, b; \mathbf{t})$ is a sum of terms each involving at least one h_a or h_{2a} . We claim that $|R(a, b; \mathbf{t})| = o((b+1)^{-m/2})$. Indeed, let $R_1(a, b; \mathbf{t})$ be the term of $R(a, b; \mathbf{t})$ involving only one h_{2a} and no other h_a . Using the fact that when f is a non-negative

function and $0 < x_1, x_2 < t$, we have

$$\int_0^t f(x_2 - x_1) dx_2 = \int_{-x_1}^{t-x_1} f(x) dx \leq \int_{-1}^1 f(x) dx.$$

Therefore,

$$\begin{aligned} & |R_1(a, b; \mathbf{t})| \\ &= \int_0^{t_1} dx_1 \dots \int_0^{t_m} dx_m h_{2a}(x_1 - x_m) |x_2 - x_1|^b |x_4 - x_3|^b \dots |x_m - x_{m-1}|^b \\ &\leq \int_0^{t_1} dx_1 \int_0^{t_3} dx_3 \dots \int_0^{t_m} dx_m h_{2a}(x_1 - x_m) \int_{-1}^1 |x_2|^b dx_2 |x_4 - x_3|^b \dots |x_m - x_{m-1}|^b \\ &\leq 2(b+1)^{-1} \int_0^{t_3} dx_3 \dots \int_0^{t_m} dx_m \int_{-1}^1 h_{2a}(x_1) dx_1 |x_4 - x_3|^b \dots |x_m - x_{m-1}|^b \\ &\leq 2(b+1)^{-1} \int_0^{t_3} dx_3 \dots \int_0^{t_m} dx_m \int_{-1}^1 (|x_1|^{2a} - 1) dx_1 |x_4 - x_3|^b \dots |x_m - x_{m-1}|^b \\ &= 4[(2a+1)^{-1} - 1](b+1)^{-1} \int_0^{t_3} dx_3 \dots \int_0^{t_m} dx_m |x_4 - x_3|^b |x_6 - x_5|^b \dots |x_m - x_{m-1}|^b \\ &\leq \dots \leq C[(2a+1)^{-1} - 1](b+1)^{-m/2} = o(1)(b+1)^{-m/2}. \end{aligned} \tag{8.41}$$

Similar estimates apply to the other terms of $R(a, b; \mathbf{t})$, which may involve a greater number of h_a or h_{2a} , and end up converging faster to zero as $a \rightarrow 0$. Hence

$$f_1(a, b; \mathbf{t}) \leq f(0, b; \mathbf{t}) + o\left((b+1)^{-m/2}\right) \sim (b+1)^{-m/2} \prod_{i=2,4,\dots,m} (t_i + t_{i-1} - |t_i - t_{i-1}|)$$

using (8.39). The same estimate holds for $f_2(a, b; \mathbf{t})$. Hence by (8.40),

$$f(a, b; \mathbf{t}) \leq f(0, b; \mathbf{t}) + o\left((b+1)^{-m/2}\right) \sim (b+1)^{-m/2} \prod_{i=2,4,\dots,m} (t_i + t_{i-1} - |t_i - t_{i-1}|). \tag{8.42}$$

Combining (8.39) and (8.42) concludes the proof. \square

Lemma 8.3.16. *Let $X_\rho(t)$ be the limit process in (8.10). For $m \geq 3$,*

$$\kappa_m \left(\sum_{i=1}^n c_i X_\rho(t_i) \right) = \begin{cases} \rho^{m/2} (m-1)! \left[\sum_{i,j=1}^n c_i c_j \frac{1}{2} (|t_i| + |t_j| - |t_i - t_j|) \right]^{m/2} & \text{if } m \text{ is even;} \\ 0 & \text{if } m \text{ is odd.} \end{cases}$$

Proof. Then because $B_1(t)$, $B_2(t)$ and W are independent,

$$\kappa_m \left(\sum_{i=1}^n c_i X_\rho(t_i) \right) = \kappa_m \left(\rho^{1/2} \sum_{i=1}^n c_i W B(t_i) \right) + \kappa_m \left((1-\rho)^{1/2} \sum_{i=1}^n c_i B'(t_i) \right).$$

Now note that the second term is Gaussian and thus the cumulants of order higher than 2 is always zero. Applying Lemma 8.3.12 (with $\gamma = -1$) to the first term concludes the proof. \square

Now we proceed to the proof of Theorem 8.2.5. Again by Lemma 8.3.8, tightness always holds. We only need to show the convergence of the finite-dimensional distributions.

Proof of $\xrightarrow{f.d.d.}$ in Theorem 8.2.5. The distribution of $\sum_{i=1}^n c_i X_\rho(t_i)$ is moment-determinate since it is a second-order polynomial in normal random variables (see, e.g., Slud [1993]). One can therefore use a method of moments.

We analyze the asymptotics of the cumulants in (8.13) with $m \geq 3$ and $A(\gamma_1, \gamma_2)$ as given in (8.21) as $(\gamma_1, \gamma_2) \rightarrow (-1/2, -1)$. First, by Lemma 8.3.11,

$$\begin{aligned} & A(\gamma_1, \gamma_2)^m \\ & \sim (\gamma_1 + \gamma_2 + 3/2)^{m/2} [\mathbf{IB}(1/2, 1/2) \mathbf{IB}(\gamma_2 + 1, 1/2) + \mathbf{IB}(1/2, -2\gamma_1 - 1) \mathbf{IB}(\gamma_2 + 1, 1)]^{-m/2} \\ & \sim (\gamma_1 + \gamma_2 + 3/2)^{m/2} [\mathbf{IB}(1/2, -2\gamma_1 - 1) \mathbf{IB}(\gamma_2 + 1, 1)]^{-m/2} \\ & \sim (\gamma_1 + \gamma_2 + 3/2)^{m/2} (-2\gamma_1 - 1)^{m/2} (\gamma_2 + 1)^{m/2}, \end{aligned} \tag{8.43}$$

which converges to 0.

Now we analyze the asymptotics of the terms of $C_m(\gamma_1, \gamma_2; \mathbf{t}, \mathbf{c})$ in (8.14) as σ varies in $\{1, 2\}^m$. When m is even, consider first the two main terms where

$$\sigma = (1, 2, 1, 2, \dots, 1, 2) \text{ and } \sigma = (2, 1, 2, 1, \dots, 2, 1),$$

which correspond to $\#\{j : \sigma_j = \sigma'_{j-1} = 1\} = m/2$. As in the proof of Theorem 8.2.2, the corresponding term when $\sigma = (1, 2, 1, 2, \dots, 1, 2)$ in (8.14) (it is the same for $\sigma = (2, 1, 2, 1, \dots, 2, 1)$) is

$$\begin{aligned} & \sum_{i_1, \dots, i_m=1}^n c_{i_1} \dots c_{i_m} \mathcal{B}(\gamma_1 + 1, -2\gamma_1 - 1)^{m/2} \mathcal{B}(\gamma_2 + 1, -2\gamma_2 - 1)^{m/2} \times \\ & \int_0^{t_{i_1}} ds_1 \dots \int_0^{t_{i_m}} ds_m |s_1 - s_m|^{2\gamma_1+1} |s_2 - s_1|^{2\gamma_2+1} \dots |s_{m-1} - s_{m-2}|^{2\gamma_1+1} |s_m - s_{m-1}|^{2\gamma_2+1} \\ & \sim (-2\gamma_1 - 1)^{-m/2} (\gamma_2 + 1)^{-m} \left[\sum_{i,j=1}^n c_i c_j \frac{1}{2} (|t_i| + |t_j| - |t_i - t_j|) \right]^{m/2}, \end{aligned} \quad (8.44)$$

where the last line is due to Lemma 8.3.11 and Lemma 8.3.15.

Any other σ term in (8.14) is negligible because it is of order $O((-2\gamma_1 - 1)^{-r} (\gamma_2 + 1)^{-m})$, where

$$r = \#\{j : \sigma_j = \sigma'_{j-1} = 1\} = \#\{j : \sigma_j = \sigma'_{j-1} = 2\} < m/2. \quad (8.45)$$

Indeed, let us suppose (8.45) and examine a corresponding σ term in the expansion of the product $\prod_{j=1}^m$ in (8.14). Call this term P_m . In P_m , there are r factors of

$$\mathcal{B}(\gamma_1 + 1, -2\gamma_1 - 1) |s_j - s_{j-1}|^{2\gamma_1+1}, \quad (8.46)$$

and there are r factors of

$$\mathcal{B}(\gamma_2 + 1, -2\gamma_2 - 1) |s_j - s_{j-1}|^{2\gamma_2+1}. \quad (8.47)$$

Since (8.45) implies that $\#\{j : \sigma_j \neq \sigma'_{j-1}\} = m - 2r$, there are also $m - 2r$ factors in P_m ,

which are either

$$(s_j - s_{j-1})_+^{\gamma_1 + \gamma_2 + 1} \mathcal{B}(\gamma_1 + 1, -\gamma_1 - \gamma_2 - 1) + (s_{j-1} - s_j)_+^{\gamma_1 + \gamma_2 + 1} \mathcal{B}(\gamma_2 + 1, -\gamma_1 - \gamma_2 - 1),$$

or

$$(s_j - s_{j-1})_+^{\gamma_1 + \gamma_2 + 1} \mathcal{B}(\gamma_2 + 1, -\gamma_1 - \gamma_2 - 1) + (s_{j-1} - s_j)_+^{\gamma_1 + \gamma_2 + 1} \mathcal{B}(\gamma_1 + 1, -\gamma_1 - \gamma_2 - 1).$$

These last two expressions are both bounded by

$$|s_j - s_{j-1}|^{\gamma_1 + \gamma_2 + 1} [\mathcal{B}(\gamma_2 + 1, -\gamma_1 - \gamma_2 - 1) + \mathcal{B}(\gamma_1 + 1, -\gamma_1 - \gamma_2 - 1)]. \quad (8.48)$$

In view of Lemma 8.3.11, the beta functions in (8.46), (8.47) and (8.48) behave like $(-2\gamma_1 - 1)^{-1}$, $(\gamma_2 + 1)^{-1}$ and $(\gamma_2 + 1)^{-1}$ respectively. Therefore, the beta functions contribute an order

$$(-2\gamma_1 - 1)^{-r} (\gamma_2 + 1)^{-r} (\gamma_2 + 1)^{-(m-2r)} = (-2\gamma_1 - 1)^{-r} (\gamma_2 + 1)^{-(m-r)}.$$

The integrand involving $|s_{j-1} - s_j|^{2\gamma_2 + 1}$ contribute an order $(\gamma_2 + 1)^{-r}$. So the total order is $(-2\gamma_1 - 1)^{-r} (\gamma_2 + 1)^{-m}$. These arguments can be rigorously justified by first applying the Cauchy-Schwartz as in (8.40) to break the cyclic integrand, and then bound as in (8.41). Therefore in view of (8.44), and after also including the case $\sigma = (2, 1, 2, 1, \dots, 2, 1)$, we conclude that

$$C_m(\gamma_1, \gamma_2; \mathbf{t}, \mathbf{c}) \sim 2(-2\gamma_1 - 1)^{-m/2} (\gamma_2 + 1)^{-m} \left[\sum_{i,j=1}^n c_i c_j \frac{1}{2} (|t_i| + |t_j| - |t_i - t_j|) \right]^{m/2}, \quad (8.49)$$

if m is even.

When m is odd, there are at most $(m - 1)/2$ times of $\sigma_j = \sigma'_{j-1} = 1$ or $\sigma_j = \sigma'_{j-1} = 2$.

It can be shown similarly that $C_m(\gamma_1, \gamma_2; \mathbf{t}, \mathbf{c})$ is of the order

$$(-2\gamma_1 - 1)^{-(m-1)/2}(\gamma_2 + 1)^{-m}, \quad (8.50)$$

which is dominated by the order of convergence to 0 of $A(\gamma_1, \gamma_2)^m$ in (8.43). Now combining this fact with (8.9), (8.13), (8.43) and (8.49), we have when m is even,

$$\begin{aligned} & \kappa_m \left(\sum_{i=1}^n c_i Z_{\gamma_1, \gamma_2}(t_i) \right) \\ & \sim \left(\frac{\gamma_1 + \gamma_2 + 3/2}{\gamma_2 + 1} \right)^{m/2} (m-1)! \left[\sum_{i,j=1}^n c_i c_j \frac{1}{2} (|t_i| + |t_j| - |t_i - t_j|) \right]^{m/2} \\ & \rightarrow \rho^{m/2} (m-1)! \left[\sum_{i,j=1}^n c_i c_j \frac{1}{2} (|t_i| + |t_j| - |t_i - t_j|) \right]^{m/2}, \end{aligned} \quad (8.51)$$

and when m is odd,

$$\kappa_m \left(\sum_{i=1}^n c_i Z_{\gamma_1, \gamma_2}(t_i) \right) \rightarrow 0.$$

Now use Lemma 8.3.16 to identify the limit process. □

8.3.4 Proof of Theorem 8.2.7

We state first a combinatorial result.

Lemma 8.3.17. *Let $\sigma = (\sigma_1, \dots, \sigma_m) \in \{1, 2\}^m$. Let $\sigma' = (\sigma'_1, \dots, \sigma'_m)$ be the complement of σ , namely, $\sigma'_i = 1$ if $\sigma_i = 2$ and $\sigma'_i = 2$ if $\sigma_i = 1$, $i = 1, \dots, m$. Let σ_0 be understood as σ_m and let σ'_0 be understood as σ'_m . Then for a fixed integer $0 \leq r \leq m/2$,*

$$\#\left\{ \sigma \in \{1, 2\}^m : \#\{j : \sigma_j = \sigma'_{j-1} = 1\} = r \right\} = 2 \binom{m}{2r}. \quad (8.52)$$

Proof. If $\sigma_{j-1} \neq \sigma_j$, we say that there is an alternation at j . There are $\binom{m}{k}$ ways to place k alternations. The positions of the alternations determine the whole σ up to the replacement of 1's into 2's and vice-versa. Hence there are $2 \binom{m}{k}$ possible σ 's. To relate k

to r , note that the relation $\sigma_j = \sigma'_{j-1}$ holds if and only if $\sigma_{j-1} \neq \sigma_j$. Since

$$r = \#\{j : \sigma_j = \sigma'_{j-1} = 1\} = \#\{j : \sigma_j = \sigma'_{j-1} = 2\},$$

we have

$$k = \#\{j : \sigma_j \neq \sigma_{j-1}\} = \#\{j : \sigma_j = \sigma'_{j-1} = 1\} + \#\{j : \sigma_j = \sigma'_{j-1} = 2\} = 2r.$$

□

Lemma 8.3.18. *Let $Y_\rho(t)$ be the limit process in (8.12). For $m \geq 3$,*

$$\begin{aligned} \kappa_m \left(\sum_{i=1}^n c_i Y_\rho(t_i) \right) &= \frac{[(\rho+1)^{-1} + (2\sqrt{\rho})^{-1}]^m + [(\rho+1)^{-1} - (2\sqrt{\rho})^{-1}]^m}{[(\rho+1)^{-2} + (4\rho)^{-1}]^{m/2}} \\ &\quad \times \left(\sum_{i=1}^n c_i t_i \right)^m \frac{(m-1)!}{2}. \end{aligned} \quad (8.53)$$

Proof. Let

$$a_\rho = \frac{(\rho+1)^{-1} + (2\sqrt{\rho})^{-1}}{\sqrt{2(\rho+1)^{-2} + (2\rho)^{-1}}}, \quad b_\rho = \frac{(\rho+1)^{-1} - (2\sqrt{\rho})^{-1}}{\sqrt{2(\rho+1)^{-2} + (2\rho)^{-1}}}$$

Because X_1 and X_2 are two independent standardized χ_1^2 random variables, we have

$$\begin{aligned} &\kappa_m \left(\sum_{i=1}^n c_i Y_\rho(t_i) \right) \\ &= \kappa_m \left(\sum_{i=1}^n c_i t_i (a_\rho X_1 + b_\rho X_2) \right) = \left(\sum_{i=1}^n c_i t_i \right)^m [\kappa_m(a_\rho X_1) + \kappa_m(b_\rho X_2)] \\ &= \left(\sum_{i=1}^n c_i t_i \right)^m (a_\rho^m + b_\rho^m) \kappa(X_1) = 2^{m/2} (a_\rho^m + b_\rho^m) \left(\sum_{i=1}^n c_i t_i \right)^m \frac{(m-1)!}{2}. \end{aligned}$$

The factor $2^{m/2}(a_\rho^m + b_\rho^m)$ can be rewritten as the first factor in (8.53). □

Note that $a + b \sim A + B$ for $a, b, A, B > 0$, if $a \sim A$, $b \sim B$ and $a/b \sim \lambda$, where λ is a fixed number from 0 to ∞ (can be ∞), as will always be the case under our assumptions.

We now prove Theorem 8.2.7. In view of Lemma 8.3.8, we only need to show the convergence of the finite-dimensional distributions.

Proof of $\xrightarrow{f.d.d.}$ in Theorem 8.2.7. We can use a method of moments again because the limit $\sum_{i=1}^n c_i Y_\rho(t_i)$ is a second-order polynomial in normal random variables. We analyze the asymptotics of the cumulants in (8.13) with $m \geq 3$ and $A(\gamma_1, \gamma_2)$ in (8.21) as $(\gamma_1, \gamma_2) \rightarrow (-1/2, -1/2)$. Lemma 8.3.11 yields

$$A(\gamma_1, \gamma_2)^m \sim [(-\gamma_1 - \gamma_2 - 1)^{-2} + (-2\gamma_1 - 1)^{-1}(-2\gamma_2 - 1)^{-1}]^{-m/2}, \quad (8.54)$$

and C_m in (8.14) satisfies

$$C_m(\gamma_1, \gamma_2; \mathbf{t}, \mathbf{c}) \sim \left(\sum_{i=1}^n c_i t_i \right)^m \sum_{\sigma \in \{1,2\}^m} \prod_{j=1}^m (-\gamma_{\sigma_j} - \gamma_{\sigma'_{j-1}} - 1)^{-1}, \quad (8.55)$$

where we get the term $(\sum_{i=1}^n c_i t_i)^m$ from $\sum_{i_1, \dots, i_m=1}^n c_{i_1} \dots c_{i_m} \int_0^{t_{i_1}} ds_1 \dots \int_0^{t_{i_m}} ds_m$.

Let $r = \#\{j : \sigma_j = \sigma'_{j-1} = 1\} = \#\{j : \sigma_j = \sigma'_{j-1} = 2\}$. Then using Lemma 8.3.17, we can write

$$\begin{aligned} & \sum_{\sigma \in \{1,2\}^m} \prod_{j=1}^m (-\gamma_{\sigma_j} - \gamma_{\sigma'_{j-1}} - 1)^{-1} \\ &= \sum_{0 \leq r \leq m/2} 2 \binom{m}{2r} (-2\gamma_1 - 1)^{-r} (-2\gamma_2 - 1)^{-r} (-\gamma_1 - \gamma_2 - 1)^{-(m-2r)}. \end{aligned} \quad (8.56)$$

Hence by (8.13), (8.54), (8.55) and (8.56), one has

$$\kappa_m \left(\sum_{i=1}^n c_i Z_{\gamma_1, \gamma_2}(t_i) \right) \sim (m-1)! \left(\sum_{i=1}^n c_i t_i \right)^m \sum_{0 \leq r \leq m/2} \binom{m}{2r} U(\gamma_1, \gamma_2; m, r). \quad (8.57)$$

where

$$U(\gamma_1, \gamma_2; m, r) := \frac{(-2\gamma_1 - 1)^{-r} (-2\gamma_2 - 1)^{-r} (-\gamma_1 - \gamma_2 - 1)^{-(m-2r)}}{[(-\gamma_1 - \gamma_2 - 1)^{-2} + (-2\gamma_1 - 1)^{-1}(-2\gamma_2 - 1)^{-1}]^{m/2}}.$$

As $(\gamma_1, \gamma_2) \rightarrow (-1/2, -1/2)$ and $(\gamma_1 + 1/2)/(\gamma_2 + 1/2) \rightarrow \rho \in [0, 1]$, in the case $\rho > 0$, some elementary calculation shows

$$U(\gamma_1, \gamma_2; m, r) \rightarrow \frac{[1/(2\sqrt{\rho})]^{2r} [1/(\rho + 1)]^{m-2r}}{[(\rho + 1)^{-2} + (4\rho)^{-1}]^{m/2}}, \quad (8.58)$$

and in the case $\rho = 0$,

$$U(\gamma_1, \gamma_2; m, r) \rightarrow \begin{cases} 1 & \text{if } r = m/2 \text{ (} m \text{ must be even in this case);} \\ 0 & \text{if } r < m/2. \end{cases} \quad (8.59)$$

This expression (8.59) also coincides with the limit in (8.58) as $\rho \rightarrow 0$. In the argument below we omit the case $\rho = 0$, which can be either treated separately, or obtained by taking the limit as $\rho \rightarrow 0$.

Set $a = 1/(2\sqrt{\rho})$ and $b = 1/(\rho + 1)$. Using the identity $(a + b)^m + (a - b)^m = \sum_{0 \leq r \leq m/2} 2 \binom{m}{2r} a^{2r} b^{m-2r}$, one can write following (8.57) and (8.58) that

$$\kappa_m \left(\sum_{i=1}^n c_i Z_{\gamma_1, \gamma_2}(t_i) \right) \rightarrow \frac{(a + b)^m - (a - b)^m}{(a^2 + b^2)^{m/2}} \left(\sum_{i=1}^n c_i t_i \right)^m \frac{(m-1)!}{2},$$

which is (8.53). Now use Lemma 8.3.18 to identify the limit process, concluding the proof. \square

Additional results

We deal now with the following additional three points:

1. We show that the weak convergence proved in the previous theorems *cannot* be strengthened to convergence in $L^2(\Omega)$ nor even in probability;
2. We apply the results of Nourdin and Peccati [2013] and Eichelsbacher and Thäle [2014] to determine the rate of convergence on the boundaries d and e_1 (or e_2);

3. We include an alternate proof of Theorem 8.2.2 in the spirit of Remark 8.2.3 which provides further insight on the convergence.

8.4 No convergence in $L^2(\Omega)$

The generalized Rosenblatt process $Z_{\gamma_1, \gamma_2}(t)$ was defined in (8.1) (see also (8.6)). We have shown weak convergence (convergence in distribution) for the generalized Rosenblatt process $Z_{\gamma_1, \gamma_2}(t)$ in previous theorems. Is it possible that some of these convergences are actually in a stronger mode, say, in probability? We provide a negative answer here.

Theorem 8.4.1. *In Theorem 8.2.1, 8.2.2, 8.2.5 and 8.2.7, the weak convergence cannot be extended to convergence in $L^2(\Omega)$, nor even to convergence in probability.*

Remark 8.4.2. In fact, it suffices to show that the convergence cannot be extended to convergence in $L^2(\Omega)$. This is because, on a fixed order Wiener chaos, convergence in $L^2(\Omega)$ and convergence in probability are equivalent. See Schreiber [1969]. Alternatively, to verify the equivalence, suppose that X_n is a sequence on a fixed order Wiener chaos, and X_n converges in probability to X . The sequence is therefore tight. Then by, e.g., Lemma 2.1(ii) of Nourdin and Rosinski [2014], $\sup_n \mathbb{E}|X_n|^p < \infty$ for any $p > 0$, which entails uniform integrability and hence convergence in $L^2(\Omega)$.

To prove Theorem 8.4.1, it suffices to show that any sequence of

$$Z_{\gamma_1, \gamma_2} := Z_{\gamma_1, \gamma_2}(1)$$

as (γ_1, γ_2) approach the boundaries is *not* a Cauchy sequence in $L^2(\Omega)$. Let (α_1, α_2) and (γ_1, γ_2) be in the region Δ in (8.2). Then since Z_{γ_1, γ_2} is standardized, we have

$$\mathbb{E}(Z_{\alpha_1, \alpha_2} - Z_{\gamma_1, \gamma_2})^2 = 2 - 2\mathbb{E}Z_{\alpha_1, \alpha_2}Z_{\gamma_1, \gamma_2}. \quad (8.60)$$

If (α_1, α_2) and (γ_1, γ_2) converge to the same point on the boundary, we may expect that $\mathbb{E}Z_{\alpha_1, \alpha_2}Z_{\gamma_1, \gamma_2} \rightarrow 1$ and hence $\mathbb{E}(Z_{\alpha_1, \alpha_2} - Z_{\gamma_1, \gamma_2})^2 \rightarrow 0$, which would prove Cauchy conver-

gence. We will show, however, that

$$\liminf_{(\alpha_1, \alpha_2), (\gamma_1, \gamma_2) \rightarrow \text{boundary point}} \mathbf{E}Z_{\alpha_1, \alpha_2}Z_{\gamma_1, \gamma_2} < 1. \quad (8.61)$$

In other words, we will show that there is no $L^2(\Omega)$ continuity at the boundary.

First we compute the covariance in (8.60).

Lemma 8.4.3.

$$\begin{aligned} \mathbf{E}Z_{\alpha_1, \alpha_2}Z_{\gamma_1, \gamma_2} &= A(\alpha_1, \alpha_2)A(\gamma_1, \gamma_2)(\alpha_1 + \alpha_2 + \gamma_1 + \gamma_2 + 3)^{-1}(\alpha_1 + \alpha_2 + \gamma_1 + \gamma_2 + 4)^{-1} \\ &\times [\mathbf{B}(\alpha_1 + 1, -\alpha_1 - \gamma_1 - 1)\mathbf{B}(\alpha_2 + 1, -\alpha_2 - \gamma_2 - 1) \\ &+ \mathbf{B}(\gamma_1 + 1, -\alpha_1 - \gamma_1 - 1)\mathbf{B}(\gamma_2 + 1, -\alpha_2 - \gamma_2 - 1) \\ &+ \mathbf{B}(\alpha_2 + 1, -\alpha_2 - \gamma_1 - 1)\mathbf{B}(\alpha_1 + 1, -\alpha_1 - \gamma_2 - 1) \\ &+ \mathbf{B}(\gamma_1 + 1, -\alpha_2 - \gamma_1 - 1)\mathbf{B}(\gamma_2 + 1, -\alpha_1 - \gamma_2 - 1)]. \end{aligned} \quad (8.62)$$

Proof. We shall use the representation (8.6) of $Z_{\gamma_1, \gamma_2}(t)$ in order to apply the formula

$$\mathbf{E}I_2(f)I_2(g) = 2\langle f, g \rangle_{L^2(\mathbb{R}^2)}$$

for symmetric functions f and g (see (7.3.39) of Peccati and Taqqu [2011]). Using (8.18), we get

$$\begin{aligned} &2A(\alpha_1, \alpha_2)^{-1}A(\gamma_1, \gamma_2)^{-1}\mathbf{E}Z_{\alpha_1, \alpha_2}Z_{\gamma_1, \gamma_2} \\ &= \int_{[0,1]^2} ds \int_{\mathbb{R}^2} d\mathbf{x} [(s_1 - x_1)_+^{\alpha_1}(s_1 - x_2)_+^{\alpha_2} + (s_1 - x_1)_+^{\alpha_2}(s_1 - x_2)_+^{\alpha_1}] \\ &\quad \times [(s_2 - x_1)_+^{\gamma_1}(s_2 - x_2)_+^{\gamma_2} + (s_2 - x_1)_+^{\gamma_2}(s_2 - x_2)_+^{\gamma_1}] \\ &= 2 \int_{[0,1]^2} ds \left[(s_2 - s_1)_+^{\alpha_1 + \alpha_2 + \gamma_1 + \gamma_2 + 2} \mathbf{B}(\alpha_1 + 1, -\alpha_1 - \gamma_1 - 1) \mathbf{B}(\alpha_2 + 1, \alpha_2 - \gamma_2 - 1) \right. \\ &\quad + (s_1 - s_2)_+^{\alpha_1 + \alpha_2 + \gamma_1 + \gamma_2 + 2} \mathbf{B}(\gamma_1 + 1, -\alpha_1 - \gamma_1 - 1) \mathbf{B}(\gamma_2 + 1, -\alpha_2 - \gamma_2 - 1) \\ &\quad \left. + (s_2 - s_1)_+^{\alpha_1 + \alpha_2 + \gamma_1 + \gamma_2 + 2} \mathbf{B}(\alpha_2 + 1, -\alpha_2 - \gamma_1 - 1) \mathbf{B}(\alpha_1 + 1, -\alpha_1 - \gamma_2 - 1) \right] \end{aligned}$$

$$+ (s_1 - s_2)_+^{\alpha_1 + \alpha_2 + \gamma_1 + \gamma_2 + 2} \mathcal{IB}(\gamma_1 + 1, -\alpha_2 - \gamma_1 - 1) \mathcal{IB}(\gamma_2 + 1, -\alpha_2 - \gamma_2 - 1) \Big]$$

Since $\alpha_1 + \alpha_2 > -3/2$ and $\gamma_1 + \gamma_2 > -3/2$, we have $\alpha_1 + \alpha_2 + \gamma_1 + \gamma_2 + 2 > -1$. Since

$$\int_{[0,1]^2} (s_1 - s_2)_+^u ds = \int_{[0,1]^2} (s_2 - s_1)_+^u ds = (u+1)^{-1}(u+2)^{-1}$$

for $u > -1$, we get (8.62). □

Proof of Theorem 8.4.1.

Case of Theorem 8.2.1. By (8.7), an element of the second chaos converges in distribution to a Gaussian. That this cannot be extended to convergence in $L^2(\Omega)$ follows from the fact that $\{I_2(f) : f \in L^2(\mathbb{R}^2)\}$ is a closed subspace in $L^2(\Omega)$. Hence the $L^2(\Omega)$ limit of a double Wiener-Itô integral must still be a double Wiener-Itô integral, which means that it cannot be Gaussian.

Case of Theorem 8.2.2. Let $(\alpha_1, \alpha_2) \rightarrow (-1/2, \gamma)$ and $(\gamma_1, \gamma_2) \rightarrow (-1/2, \gamma)$, where $\gamma \in (-1, -1/2)$. Assume in addition that the convergence speeds are comparable, that is, $(\alpha_1 + 1/2)/(\gamma_1 + 1/2) \sim r \in (0, 1)$. Then using (8.32) with $m = 1$, Lemma 8.3.11, and (8.62), one has

$$\begin{aligned} \mathbb{E}Z_{\alpha_1, \alpha_2}Z_{\gamma_1, \gamma_2} &\sim (-2\alpha_1 - 1)^{1/2}(-2\gamma_1 - 1)^{1/2}(2\gamma + 3)(\gamma + 1)\mathcal{IB}(\gamma + 1, -2\gamma - 1)^{-1} \\ &\quad \times (2 + 2\gamma)^{-1}(3 + 2\gamma)^{-1}[2\mathcal{IB}(\gamma + 1, -2\gamma - 1)(-\alpha_1 - \gamma_1 - 1)^{-1}] \\ &\sim \frac{(-2\alpha_1 - 1)^{1/2}(-2\gamma_1 - 1)^{1/2}}{(-\alpha_1 - \gamma_1 - 1)} \sim 2r^{1/2}/(1+r) < 1. \end{aligned}$$

Case of Theorem 8.2.5. When $\rho < 1$, the limit in (8.10) involves a Gaussian component, which by the same reason as in “Case of Theorem 8.2.1” implies that $L^2(\Omega)$ convergence cannot hold. We only need to consider the case $\rho = 1$.

We therefore suppose that $(\alpha_1, \alpha_2) \rightarrow (-1/2, -1)$ and $(\gamma_1, \gamma_2) \rightarrow (-1/2, -1)$ and that $\rho = 1$, that is by (8.9), that $(\alpha_1 + 1/2)/(\alpha_2 + 1) \rightarrow 0$ and $(\gamma_1 + 1/2)/(\gamma_2 + 1) \rightarrow 0$. Assume in addition that $(\alpha_1 + 1/2)/(\gamma_1 + 1/2) \sim (\alpha_2 + 1)/(\gamma_2 + 1) \sim r \in (0, 1)$. By (8.43) with $m = 1$, Lemma 8.3.11, and (8.62), we have

$$\begin{aligned} & EZ_{\alpha_1, \alpha_2} Z_{\gamma_1, \gamma_2} \\ & \sim (\alpha_1 + \alpha_2 + 3/2)^{1/2} (-2\alpha_1 - 1)^{1/2} (\alpha_2 + 1)^{1/2} (\gamma_1 + \gamma_2 + 3/2)^{1/2} (-2\gamma_1 - 1)^{1/2} (\gamma_2 + 1)^{1/2} \\ & \quad \times (\alpha_1 + \alpha_2 + \gamma_1 + \gamma_2 + 3)^{-1} (-\alpha_1 - \gamma_1 - 1)^{-1} [(\alpha_2 + 1)^{-1} + (\gamma_2 + 1)^{-1}] \\ & \sim \frac{(\alpha_2 + 1)(-2\alpha_1 - 1)^{1/2} (\gamma_2 + 1)(-2\gamma_1 - 1)^{1/2}}{(\alpha_2 + 1 + \gamma_2 + 1)(-\alpha_1 - \gamma_1 - 1)} [(\alpha_2 + 1)^{-1} + (\gamma_2 + 1)^{-1}] \\ & \sim 2r^{1/2}/(r + 1) < 1. \end{aligned}$$

Case of Theorem 8.2.7. Suppose $(\alpha_1, \alpha_2) \rightarrow (-1/2, -1/2)$ and $(\gamma_1, \gamma_2) \rightarrow (-1/2, -1/2)$ and that $(\alpha_1 + 1/2)/(\alpha_2 + 1/2) \sim (\gamma_1 + 1/2)/(\gamma_2 + 1/2) \sim \rho$, where $\rho \in [0, 1]$. Assume in addition that $(\alpha_1 + 1/2)/(\gamma_1 + 1/2) \sim (\alpha_2 + 1/2)/(\gamma_2 + 1/2) \sim \bar{r} \in (0, 1)$. We apply (8.54) with $m = 1$, (8.62) and Lemma 8.3.11. In this case, all beta functions in (8.62) blow up and we get

$$\begin{aligned} & EZ_{\alpha_1, \alpha_2} Z_{\gamma_1, \gamma_2} \sim [(-\alpha_1 - \alpha_2 - 1)^{-2} + (-2\alpha_1 - 1)^{-1} (-2\alpha_2 - 1)^{-1}]^{-1/2} \\ & \quad \times [(-\gamma_1 - \gamma_2 - 1)^{-2} + (-2\gamma_1 - 1)^{-1} (-2\gamma_2 - 1)^{-1}]^{-1/2} \times \frac{1}{2} \\ & \quad \times [2(-\alpha_1 - \gamma_1 - 1)^{-1} (-\alpha_2 - \gamma_2 - 1)^{-1} + 2(-\alpha_2 - \gamma_1 - 1)^{-1} (-\alpha_1 - \gamma_2 - 1)^{-1}] \\ & \sim \frac{4r}{(r + 1)^2} \left(\frac{(r + \rho)(1 + r\rho) + (r + 1)^2 \rho}{(1 + \rho)^2 + 4\rho} \right) \frac{(1 + \rho)^2}{(r + \rho)(1 + r\rho)}, \end{aligned}$$

which is close to zero if r is small. Thus (8.61) holds. \square

8.5 Convergence rate of marginal distribution on the boundaries

Rates of convergence of the marginal distribution of multiple Wiener-Itô integrals are available when the limit is Gaussian or is a product of independent Gaussians. We can

thus apply these rates when converging to the boundaries of the triangle, with some corners excluded.

First we consider the convergence rate of the marginal distribution in the case of Theorem 8.2.1 and 8.2.5 and the limit being Gaussian. We use the notation $A \asymp B$, where A and B are two nonnegative quantities, to denote that there exist constants $c < C$ independent of A and B such that $cB \leq A \leq CB$. Let $d_{TV}(X, Y)$ denote the total variation distance between the distributions of random variables X and Y , namely

$$d_{TV}(X, Y) = \sup_{S \in \mathcal{B}(\mathbb{R})} |P(X \in S) - P(Y \in S)|,$$

where $\mathcal{B}(\mathbb{R})$ denotes the Borel sets on \mathbb{R} .

In Nourdin and Peccati [2013] Theorem 1.2, the following result was established:

Lemma 8.5.1. *Let $\{F_\gamma : \gamma \in G \subset \mathbb{R}^k\}$ be a family of random variables defined on a fixed-order Wiener chaos satisfying $\mathbb{E}F_\gamma^2 = 1$, where G is an open set of indices. Suppose that the third cumulant $\kappa_3(F_\gamma)$ and the fourth cumulant $\kappa_4(F_\gamma)$ converge uniformly to zero as $\gamma \in G$ approaches a set $E \subset \overline{G}$ (as the distance between the point γ and the set E converges to zero). Then there exists a neighborhood $\mathcal{N}(E)$ of E in \mathbb{R}^k , such that when $\gamma \in \mathcal{N}(E) \cap G$, we have*

$$d_{TV}(F_\gamma, N) \asymp M(F_\gamma), \tag{8.63}$$

where N is a standard normal random variable and

$$M(F_\gamma) = \max(|\mathbb{E}F_\gamma^3|, |\mathbb{E}F_\gamma^4 - 3|) = \max(|\kappa_3(F_\gamma)|, |\kappa_4(F_\gamma)|). \tag{8.64}$$

Remark 8.5.2. Though the theorem was originally stated in Nourdin and Peccati [2013] for a sequence $\{F_n\}$ with a discrete parameter n , examining the proof there one sees that for (8.63) to hold, one only needs $\kappa_3(F_\gamma)$ and $\kappa_4(F_\gamma)$ to converge uniformly to zero, which is implied by our statement of the theorem.

Remark 8.5.3. Earlier in Biermé et al. [2012], the same result (8.63) was established for

the following distributional distance $d_B(\cdot, \cdot)$:

$$d_B(X, Y) = \sup_{h \in \mathcal{U}} \{|Eh(X) - Eh(Y)|\}, \tag{8.65}$$

where \mathcal{U} is the class of functions that are twice differentiable with continuous derivatives satisfying $\|h''\|_\infty < \infty$.

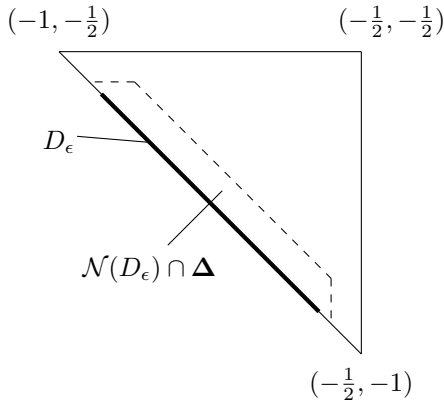


Figure 8.6: Illustration of the neighborhood $\mathcal{N}(D_\epsilon)$ of D_ϵ in Theorem 8.5.4

In the case of Theorem 8.2.1, we considered convergence to the boundary d through the neighborhood $\mathcal{N}(D_\epsilon) \cap \Delta$ illustrated in Figure 8.6. Applying Lemma 8.5.1, we get the following:

Theorem 8.5.4. *Let $Z_{\gamma_1, \gamma_2} = Z_{\gamma_1, \gamma_2}(1)$, and let N be a standard normal random variable. Then under the assumptions of Theorem 8.2.1, there exists a neighborhood $\mathcal{N}(D_\epsilon)$ of the diagonal line segment $D_\epsilon := \{\gamma_1 + \gamma_2 + 3/2 = 0 : \gamma_1, \gamma_2 > -1 + \epsilon\}$, such that when $(\gamma_1, \gamma_2) \in \mathcal{N}(D_\epsilon) \cap \Delta$, we² have*

$$d_{TV}(Z_{\gamma_1, \gamma_2}, N) \asymp (\gamma_1 + \gamma_2 + 3/2)^{3/2}. \tag{8.66}$$

Proof. Since N is Gaussian, we can apply Lemma 8.5.1. To do so, we need to compute the cumulants κ_3 and κ_4 which are given in Proposition 8.3.1. We examine the relation (8.13)

²Since Δ is an open set, $\mathcal{N}(D_\epsilon) \cap \Delta$ does not contain the segment D_ϵ .

of Proposition 8.3.1 with $A = A(\gamma_1, \gamma_2)$ given in (8.21), $m = 1$, $t = 1$, and $c = 1$. The factor $C_m(\gamma_1, \gamma_2, 1, 1)$ in (8.14) is a positive continuous function with respect to (γ_1, γ_2) . This can be shown by the Dominated Convergence Theorem as in Lemma 8.3.4. Under the assumption of Theorem 8.2.1, the parameter (γ_1, γ_2) is restricted away from boundary. So $C_m(\gamma_1, \gamma_2, 1, 1)$ is bounded below away from zero and bounded above away from infinity, and so are the factors in (8.21) except $[2(\gamma_1 + \gamma_2) + 3]^{1/2}$, which goes to zero as $\gamma_1 + \gamma_2 \rightarrow -3/2$. We get

$$\kappa_m(Z_{\gamma_1, \gamma_2}) \asymp A(\gamma_1, \gamma_2)^m \asymp (\gamma_1 + \gamma_2 + 3/2)^{m/2}, \quad m \geq 3. \quad (8.67)$$

The maximum in (8.64) is then $\kappa_3(F_\gamma)$. Combining this with (8.63), we get (8.66). \square

From (8.67) and (8.63), it is the third cumulant that determines the rate of convergence in the case of Theorem 8.2.1. When (γ_1, γ_2) is allowed to be close to the corner $(-1/2, -1)$, that is, in the case of Theorem 8.2.5 when $\rho = 0$, we will show that the fourth cumulant may come into play in the rate of convergence.

Theorem 8.5.5. *Let $Z_{\gamma_1, \gamma_2} = Z_{\gamma_1, \gamma_2}(1)$, and let N be a standard normal random variable. Then under the assumptions of Theorem 8.2.5 when $\rho = 0$, that is when*

$$-\gamma_1 - 1/2 \sim \gamma_2 + 1, \quad (8.68)$$

there exists a neighborhood \mathcal{N} of $(-1/2, -1)$, such that when $(\gamma_1, \gamma_2) \in \mathcal{N} \cap \Delta$, we have³

$$d_{TV}(Z_{\gamma_1, \gamma_2}, N) \asymp (\gamma_1 + \gamma_2 + 3/2)^{3/2} (\gamma_2 + 1)^{-1} (1 + L(\gamma_1, \gamma_2)), \quad (8.69)$$

as $(\gamma_1, \gamma_2) \rightarrow (-1/2, -1)$, where

$$L(\gamma_1, \gamma_2) = \sqrt{(-\gamma_1 - 1/2)^{-1} - (\gamma_2 + 1)^{-1}} = o\left((-\gamma_1 - 1/2)^{-1/2}\right) \text{ or } o\left((\gamma_2 + 1)^{-1/2}\right). \quad (8.70)$$

³As before, since Δ is an open set, $\mathcal{N} \cap \Delta$ does not contain the limit point $(-1/2, -1)$.

Proof. First in view of (8.9) with $\rho = 0$, we have

$$V(\gamma_1, \gamma_2) := (\gamma_1 + \gamma_2 + 3/2)^{3/2}(\gamma_2 + 1)^{-1} \rightarrow 0, \quad \text{as } (\gamma_1, \gamma_2) \rightarrow (-1/2, -1).$$

By (8.13), (8.43), (8.50) with $m = 3$, and (8.68), we get for the third cumulant

$$\kappa_3(Z_{\gamma_1, \gamma_2}) \asymp (-\gamma_1 - 1/2)^{1/2}(\gamma_1 + \gamma_2 + 3/2)^{3/2}(\gamma_2 + 1)^{-3/2} \sim V(\gamma_1, \gamma_2). \quad (8.71)$$

By (8.51) with $m = 4$ and also (8.68), we have for the fourth cumulant

$$\begin{aligned} \kappa_4(Z_{\gamma_1, \gamma_2}) &\asymp \left(\frac{\gamma_1 + \gamma_2 + 3/2}{\gamma_2 + 1} \right)^2 \\ &\sim V(\gamma_1, \gamma_2) \left(\frac{\gamma_1 + \gamma_2 + 3/2}{(-\gamma_1 - 1/2)(\gamma_2 + 1)} \right)^{1/2} = V(\gamma_1, \gamma_2)L(\gamma_1, \gamma_2). \end{aligned} \quad (8.72)$$

Since $\max(x, y) \asymp x + y$ for $x, y \geq 0$, we get

$$\max[\kappa_3(\gamma_1, \gamma_2), \kappa_4(\gamma_1, \gamma_2)] \asymp V(\gamma_1, \gamma_2) [1 + L(\gamma_1, \gamma_2)].$$

We thus apply Lemma 8.5.1 to get (8.69). At last, note that (8.68) entails that

$$L(\gamma_1, \gamma_2) = (-\gamma_1 - 1/2)^{-1/2} \sqrt{1 - \frac{-\gamma_1 - 1/2}{\gamma_2 + 1}} = o\left((-\gamma_1 - 1/2)^{-1/2}\right) \text{ or } o\left((\gamma_2 + 1)^{-1/2}\right).$$

□

Remark 8.5.6. In view of Remark 8.5.3, Theorem 8.5.4 and 8.5.5 also hold if the distance $d_{TV}(\cdot, \cdot)$ is replaced by the distance $d_B(\cdot, \cdot)$ defined by (8.65).

Remark 8.5.7. The rate of convergence to zero in (8.69) is always slower than that of (8.66), which is expected since the corner $(-1/2, -1)$ also belongs to the non-Gaussian boundary.

Remark 8.5.8. From (8.71) and (8.72), one has

$$\frac{\kappa_4(Z_{\gamma_1, \gamma_2})}{\kappa_3(Z_{\gamma_1, \gamma_2})} \asymp \sqrt{(-\gamma_1 - 1/2)^{-1} - (\gamma_2 + 1)^{-1}} = L(\gamma_1, \gamma_2),$$

which is the term (8.70) appearing in (8.69). Note that $(-\gamma_1 - 1/2)^{-1} > (\gamma_2 + 1)^{-1}$ when $(\gamma_1, \gamma_2) \in \mathbf{\Delta}$. Therefore in the case of Theorem 8.2.5, the fourth cumulant plays a role in determining the rate of convergence as follows: if the fourth cumulant converges much slower compared with the third cumulant, that is, if $L(\gamma_1, \gamma_2) \rightarrow \infty$, then this will slow the rate of convergence in (8.69); if $L(\gamma_1, \gamma_2)$ is asymptotically bounded, then both the third and fourth cumulants behave like $V(\gamma_1, \gamma_2)$.

Now we consider the marginal convergence rate in the case of Theorem 8.2.2 (see Figure 8.3). This theorem involves a non-Gaussian limit. For two random variables X and Y we define the Wasserstein distance between their distributions to be

$$d_W(X, Y) = \sup_{h \in \mathcal{L}} \{ |Eh(X) - Eh(Y)| \},$$

where \mathcal{L} is the class of 1-Lipschitz functions ($h \in \mathcal{L}$ if $|h(x) - h(y)| \leq |x - y|$). The following result follows from Eichelsbacher and Thäle [2014].

Lemma 8.5.9. *Let $Y = Z_1 Z_2$ where Z_i 's are two independent standard normal variables and let $F = I_2(f)$ be an element on the second-order Wiener chaos with $EF^2 = 1$. Then there exists a constant $C > 0$ such that*

$$d_W(F, Y) \leq C \left(1 + \frac{1}{6} \kappa_3(F)^2 - \frac{1}{3} \kappa_4(F) + \frac{1}{120} \kappa_6(F) \right)^{1/2}. \quad (8.73)$$

Proof. By Proposition 1.2(iii) of Gaunt [2014], the distribution of $Z_1 Z_2$ is the symmetric Variance-Gamma $VG(1, 0, 1, 0)$, that is, $VG(2r, 0, 1/\lambda, 0)$ with $r = 1/2$ and $\lambda = 1$. Inserting these values of r and λ in Theorem 5.10(b) of Eichelsbacher and Thäle [2014] gives (8.73). □

Using the preceding result, we get the following bound for the convergence rate as (γ_1, γ_2) approaches the boundary e_1 .

Theorem 8.5.10. *Let $Z_{\gamma_1, \gamma_2} = Z_{\gamma_1, \gamma_2}(1)$, and let $Y = Z_1 Z_2$ be as in Lemma 8.5.9. As*

$$(\gamma_1, \gamma_2) \rightarrow (-1/2, \gamma), \quad -1 < \gamma < -1/2,$$

we have

$$d_W(Z_{\gamma_1, \gamma_2}, Y) = O\left((- \gamma_1 - 1/2)^{1/2}\right). \quad (8.74)$$

Proof. Following the proof of Theorem 8.2.2, one has by (8.34) that as $(\gamma_1, \gamma_2) \rightarrow (-1/2, \gamma)$,

$$\kappa_3(Z_{\gamma_1, \gamma_2}) = O\left((- \gamma_1 - 1/2)^{1/2}\right). \quad (8.75)$$

On the other hand by (8.37), we have the convergence $\kappa_m(Z_{\gamma_1, \gamma_2}) \rightarrow (m-1)!$ for m even.

So $\kappa_4(Z_{\gamma_1, \gamma_2}) \rightarrow 6$ and $\kappa_6(Z_{\gamma_1, \gamma_2}) \rightarrow 120$, and hence

$$1 + \frac{1}{6}\kappa_3(Z_{\gamma_1, \gamma_2})^2 - \frac{1}{3}\kappa_4(Z_{\gamma_1, \gamma_2}) + \frac{1}{120}\kappa_6(Z_{\gamma_1, \gamma_2}) \rightarrow 1 + 0 - 2 + 1 = 0.$$

We thus need to study the rate of convergence of the even-order cumulants κ_4 and κ_6 . It follows from Corollary 8.3.14 that

$$\kappa_4(Z_{\gamma_1, \gamma_2}) = 6 + O(-\gamma_1 - 1/2), \quad \kappa_6(Z_{\gamma_1, \gamma_2}) = 120 + O(-\gamma_1 - 1/2). \quad (8.76)$$

The proof is concluded by plugging (8.75) and (8.76) in (8.73). \square

Recently Arras et al. [2016] obtained the rate of convergence when the limit is $\sum_{i=1}^q \alpha_i X_i$ where X_i 's are standardized chi-square random variables with one degree of freedom. Applying this result (Theorem 3.1 of Arras et al. [2016]) to the convergence of $(\gamma_1, \gamma_2) \in \Delta$ to the corner $(-1/2, -1/2)$ in the context of Theorem 8.2.7, they obtained as $\gamma_1 \rightarrow -1/2$

that

$$d_W(Z_{\gamma_1, \gamma_2}, Y_\rho(1)) = O((- \gamma_1 - 1/2)^{1/2}),$$

where $Y_\rho(1)$ is as in Theorem 8.2.7. See Example 3.2 of Arras et al. [2016].

8.6 A constructive proof of Theorem 8.2.2

The method-of-moments proof of Theorem 8.2.2 gives little intuitive insight of the convergence. Motivated by the observation made in Remark 8.2.3, we give an alternate proof of Theorem 8.2.2. The proof is based on discretization which removes the singularities at $s = x_1$ and $s = x_2$ of the integrand in (8.1), so that one is able to interchange the integration orders between $\int_{\mathbb{R}^2}' \cdot B(dx_1)B(dx_2)$ and $\int_0^t \cdot ds$. Then one uses the triangular approximation described at the end of the proof.

The proof is based on several lemmas. We use below the notation $(s, x)_N^\gamma$ to denote:

$$(s, x)_N^\gamma := \left(\frac{[Ns] - [Nx] + 1}{N} \right)^\gamma I\{[Ns] > [Nx]\}, \quad \gamma < 0. \quad (8.77)$$

Define also

$$[s - x]_N^\gamma := (s - x + 2/N)^\gamma I\{s > x + 1/N\} \leq (s, x)_N^\gamma \leq (s - x)^\gamma I\{s > x\} = (s - x)_+^\gamma. \quad (8.78)$$

Let $Z_{\gamma_1, \gamma_2}(t)$ be as in (8.1), and let

$$Z_{\gamma_1, \gamma_2}^N(t) = A_N(\gamma_1, \gamma_2) \int_{\mathbb{R}^2}' \int_0^t (s, x_1)_N^{\gamma_1} (s, x_2)_N^{\gamma_2} ds B(dx_1)B(dx_2), \quad (8.79)$$

where the Brownian measure $B(\cdot)$ is the same as the one defining $Z_{\gamma_1, \gamma_2}(t)$, and where $A_N(\gamma_1, \gamma_2)$ is chosen such that $\mathbb{E}Z_{\gamma_1, \gamma_2}^N(1)^2 = 1$.

Lemma 8.6.1. *For any $t > 0$, we have*

$$\lim_{N \rightarrow \infty} \limsup_{(\gamma_1, \gamma_2) \rightarrow (-1/2, \gamma)} \mathbb{E} |Z_{\gamma_1, \gamma_2}(t) - Z_{\gamma_1, \gamma_2}^N(t)|^2 = 0. \quad (8.80)$$

Proof. We take for simplicity that $t = 1$, while the other cases can be proved similarly.

Note that

$$\mathbb{E} |Z_{\gamma_1, \gamma_2}(1) - Z_{\gamma_1, \gamma_2}^N(1)|^2 = 2 - 2\mathbb{E} Z_{\gamma_1, \gamma_2}(1) Z_{\gamma_1, \gamma_2}^N(1).$$

So we need to show that

$$\lim_{N \rightarrow \infty} \liminf_{(\gamma_1, \gamma_2) \rightarrow (-1/2, \gamma)} \mathbb{E} Z_{\gamma_1, \gamma_2}(1) Z_{\gamma_1, \gamma_2}^N(1) \geq 1. \quad (8.81)$$

Indeed, using the symmetrized kernel in (8.6), we have

$$\begin{aligned} \mathbb{E} Z_{\gamma_1, \gamma_2}(1) Z_{\gamma_1, \gamma_2}^N(1) &= \frac{1}{2} A(\gamma_1, \gamma_2) \frac{1}{2} A_N(\gamma_1, \gamma_2) 2! \int_{\mathbb{R}^2} dx_1 dx_2 \int_0^1 \int_0^1 ds_1 ds_2 \\ &\quad \times [(s_1 - x_1)_+^{\gamma_1} (s_1 - x_2)_+^{\gamma_2} + (s_1 - x_1)_+^{\gamma_2} (s_1 - x_2)_+^{\gamma_1}] \\ &\quad \times [(s_2, x_1)_N^{\gamma_1} (s_2, x_2)_N^{\gamma_2} + (s_2, x_1)_N^{\gamma_2} (s_2, x_2)_N^{\gamma_1}]. \end{aligned} \quad (8.82)$$

By definition,

$$\begin{aligned} A_N(\gamma_1, \gamma_2)^{-2} &= \frac{1}{2} \int_0^1 \int_0^1 ds_1 ds_2 \int_{\mathbb{R}^2} dx_1 dx_2 [(s_1, x_1)_N^{\gamma_1} (s_1, x_2)_N^{\gamma_2} + (s_1, x_1)_N^{\gamma_2} (s_1, x_2)_N^{\gamma_1}] \\ &\quad \times [(s_2, x_1)_N^{\gamma_1} (s_2, x_2)_N^{\gamma_2} + (s_2, x_1)_N^{\gamma_2} (s_2, x_2)_N^{\gamma_1}]. \end{aligned}$$

Applying the second inequality of (8.78) to (8.82), and using the normalization $A_N(\gamma_1, \gamma_2)$,

we have

$$\mathbb{E} Z_{\gamma_1, \gamma_2}(1) Z_{\gamma_1, \gamma_2}^N(1) \geq \frac{1}{2} A(\gamma_1, \gamma_2) A_N(\gamma_1, \gamma_2) 2 A_N(\gamma_1, \gamma_2)^{-2} = \frac{A(\gamma_1, \gamma_2)}{A_N(\gamma_1, \gamma_2)}.$$

So (8.81) follows from the next lemma. \square

Lemma 8.6.2. *Let the normalizations $A(\gamma_1, \gamma_2)$ and $A_N(\gamma_1, \gamma_2)$ be as in (8.21) and (8.79).*

Then

$$\lim_{N \rightarrow \infty} \lim_{(\gamma_1, \gamma_2) \rightarrow (-1/2, \gamma)} \frac{A(\gamma_1, \gamma_2)}{A_N(\gamma_1, \gamma_2)} = 1, \quad (8.83)$$

where $-1 < \gamma_1, \gamma_2 < -1/2$.

Proof. By the second inequality of (8.78), we have

$$A_N(\gamma_1, \gamma_2)^{-2} \leq A(\gamma_1, \gamma_2)^{-2}. \quad (8.84)$$

By the first inequality of (8.78), we have

$$\begin{aligned} A_N(\gamma_1, \gamma_2)^{-2} &\geq \frac{1}{2} \int_0^1 \int_0^1 ds_1 ds_2 \int_{\mathbb{R}^2} dx_1 dx_2 \left([s_1 - x_1]_N^{\gamma_1} [s_1 - x_2]_N^{\gamma_2} + [s_1 - x_1]_N^{\gamma_2} [s_1 - x_2]_N^{\gamma_1} \right) \\ &\quad \times \left([s_2 - x_1]_N^{\gamma_1} [s_2 - x_2]_N^{\gamma_2} + [s_2 - x_1]_N^{\gamma_2} [s_2 - x_2]_N^{\gamma_1} \right) \\ &= P_N(\gamma_1, \gamma_2) + Q_N(\gamma_1, \gamma_2), \end{aligned} \quad (8.85)$$

where

$$P_N(\gamma_1, \gamma_2) = 2 \int_{0 < s_1 < s_2 < 1} ds_1 ds_2 \int_{\mathbb{R}} [s_1 - x_1]_N^{\gamma_1} [s_2 - x_1]_N^{\gamma_1} dx_1 \int_{\mathbb{R}} [s_1 - x_2]_N^{\gamma_2} [s_2 - x_2]_N^{\gamma_2} dx_2,$$

and

$$Q_N(\gamma_1, \gamma_2) = 2 \int_{0 < s_1 < s_2 < 1} ds_1 ds_2 \int_{\mathbb{R}} [s_1 - x_1]_N^{\gamma_1} [s_2 - x_1]_N^{\gamma_2} dx_1 \int_{\mathbb{R}} [s_1 - x_2]_N^{\gamma_2} [s_2 - x_2]_N^{\gamma_1} dx_2.$$

In the integrals over \mathbb{R} , the exponents of Q_N alternate where as those of P_N are the same.

Note that for $\alpha, \beta \in (-1, -1/2)$ and $0 < s_1 < s_2 < 1$, we have

$$\begin{aligned} &\int_{\mathbb{R}} [s_1 - x]_N^\alpha [s_2 - x]_N^\beta dx \\ &= \int_{-\infty}^{s_1 - 1/N} (s_1 - x + 2/N)^\alpha (s_2 - x + 2/N)^\beta dx \\ &= \int_0^\infty (u + 3/N)^\alpha (s_2 - s_1 + u + 3/N)^\beta du \\ &\leq \int_0^\infty u^\alpha (u + s_2 - s_1)^\beta du = (s_2 - s_1)^{\alpha + \beta + 1} \mathbf{B}(\alpha + 1, -\alpha - \beta - 1), \end{aligned} \quad (8.86)$$

after setting $u = s_1 - x - 1/N$. Thus the term Q_N from (8.85) satisfies

$$\begin{aligned} Q_N(\gamma_1, \gamma_2) &\leq 2(2\gamma_1 + 2\gamma_2 + 3)^{-1}(2\gamma_1 + 2\gamma_2 + 4)^{-1} \\ &\quad \times \mathbf{B}(\gamma_1 + 1, -\gamma_1 - \gamma_2 - 1)\mathbf{B}(\gamma_2 + 1, -\gamma_1 - \gamma_2 - 1) = O(1). \end{aligned} \quad (8.87)$$

as $(\gamma_1, \gamma_2) \rightarrow (-1/2, \gamma)$. The other term P_N in view of (8.78) and (8.86) becomes

$$\begin{aligned} P_N(\gamma_1, \gamma_2) &= 2 \int_{0 < s_1 < s_2 < 1} ds_1 ds_2 \int_0^\infty (u + 3/N)^{\gamma_1} (s_2 - s_1 + u + 3/N)^{\gamma_1} du \\ &\quad \times \int_0^\infty (u + 3/N)^{\gamma_2} (s_2 - s_1 + u + 3/N)^{\gamma_2} du. \end{aligned}$$

Now in the second integral, use $(u + 3/N)^{\gamma_2} \geq (s_2 - s_1 + u + 3/N)^{\gamma_2}$, and in the third integral, replace u by $u(s_2 - s_1)$ and then factor $s_2 - s_1$. One gets

$$\begin{aligned} P_N(\gamma_1, \gamma_2) &\geq 2 \int_{0 < s_1 < s_2 < 1} ds_1 ds_2 \int_0^\infty (s_2 - s_1 + u + 3/N)^{2\gamma_1} du \\ &\quad \times (s_2 - s_1)^{2\gamma_2 + 1} \int_0^\infty \left(u + \frac{3}{N(s_2 - s_1)}\right)^{\gamma_2} \left(1 + u + \frac{3}{N(s_2 - s_1)}\right)^{\gamma_2} du \end{aligned}$$

Since $\int_0^\infty (s_2 - s_1 + u + 3/N)^{2\gamma_1} du = (-2\gamma_1 - 1)^{-1}(s_2 - s_1 + 3/N)^{2\gamma_1 + 1}$, one has

$$\begin{aligned} P_N(\gamma_1, \gamma_2) &\geq 2(-2\gamma_1 - 1)^{-1} \int_{0 < s_1 < s_2 < 1} ds_1 ds_2 (s_2 - s_1 + 3/N)^{2\gamma_1 + 1} (s_2 - s_1)^{2\gamma_2 + 1} \\ &\quad \times \int_0^\infty \left(u + \frac{3}{N(s_2 - s_1)}\right)^{\gamma_2} \left(u + \frac{3}{N(s_2 - s_1)} + 1\right)^{\gamma_2} du =: R_N(\gamma_1, \gamma_2). \end{aligned} \quad (8.88)$$

As $(\gamma_1, \gamma_2) \rightarrow (-1/2, \gamma)$, we have

$$\begin{aligned} (-2\gamma_1 - 1)R_N(\gamma_1, \gamma_2) &\rightarrow 2 \int_{0 < s_1 < s_2 < 1} ds_1 ds_2 (s_2 - s_1)^{2\gamma + 1} \\ &\quad \times \int_0^\infty \left(u + \frac{3}{N(s_2 - s_1)}\right)^\gamma \left(u + \frac{3}{N(s_2 - s_1)} + 1\right)^\gamma du. \end{aligned}$$

As $N \rightarrow \infty$, by the Monotone Convergence Theorem, the right-hand side of the preceding

line converges to

$$\begin{aligned} & 2 \int_{0 < s_1 < s_2 < 1} ds_1 ds_2 (s_2 - s_1)^{2\gamma+1} \int_0^\infty u^\gamma (u+1)^\gamma du \\ & = (2\gamma + 3)^{-1} (\gamma + 1)^{-1} \mathbf{B}(\gamma + 1, -2\gamma - 1). \end{aligned}$$

On the other hand, from (8.32) with $m = 2$ we have

$$A(\gamma_1, \gamma_2)^2 \sim (-2\gamma_1 - 1)(2\gamma + 3)(\gamma + 1) \mathbf{B}(\gamma + 1, -2\gamma - 1)^{-1}. \quad (8.89)$$

Hence

$$\lim_{N \rightarrow \infty} \lim_{(\gamma_1, \gamma_2) \rightarrow (-1/2, \gamma)} A(\gamma_1, \gamma_2)^2 R_N(\gamma_1, \gamma_2) = 1 \quad (8.90)$$

Combining (8.85), (8.87), (8.88) and (8.90) yields

$$\liminf_{N \rightarrow \infty} \liminf_{(\gamma_1, \gamma_2) \rightarrow (-1/2, \gamma)} \frac{A(\gamma_1, \gamma_2)^2}{A_N(\gamma_1, \gamma_2)^2} \geq 1,$$

This with (8.84) yields (8.83). □

We will now interchange the integrals $\int_0^t \cdot ds$ and $\int_{\mathbb{R}^2}' \cdot dx_1 dx_2$, and write

$$\begin{aligned} Z_{\gamma_1, \gamma_2}^N(t) &= A_N(\gamma_1, \gamma_2) \int_{\mathbb{R}^2}' \left[\int_0^t (s, x_1)_N^{\gamma_1}(s, x_2)_N^{\gamma_2} B(dx_1) B(dx_2) ds \right] \\ &= A_N(\gamma_1, \gamma_2) \int_0^t \left[\int_{\mathbb{R}^2}' (s, x_1)_N^{\gamma_1}(s, x_2)_N^{\gamma_2} B(dx_1) B(dx_2) \right] ds, \quad \text{a.s.}, \end{aligned} \quad (8.91)$$

by the stochastic Fubini theorem (see Pipiras and Taqqu [2010] Theorem 2.1). It applies since

$$\int_0^t \int_{\mathbb{R}^2} [(s, x_1)_N^{\gamma_1}(s, x_2)_N^{\gamma_2}]^2 dx_1 dx_2 ds < \infty. \quad (8.92)$$

Relation (8.92) follows from the following lemma.

Lemma 8.6.3. For any $\gamma \in (-1, -1/2)$, $t > 0$ and $N \in \mathbb{Z}_+$, we have

$$\sup_{s \in [0, t]} \int_{\mathbb{R}} (s, x)_N^{2\gamma} dx < \infty.$$

Proof. In view of (8.77),

$$\begin{aligned} \int_{\mathbb{R}} (s, x)_N^{2\gamma} dx &= \frac{1}{N} \int_{\mathbb{R}} \left(\frac{[Ns] - [Nx] + 1}{N} \right)^{2\gamma} I_{\{[Ns] > [Nx]\}} d(Nx) \\ &= N^{-2\gamma-1} \sum_{-\infty < i < [Ns]} ([Ns] - i + 1)^{2\gamma} = N^{2\gamma-1} \sum_{k=2}^{\infty} k^{-2\gamma} < \infty \end{aligned}$$

since $\gamma < -1/2$, where we set $k = [Ns] - i + 1$. Since the last expression does not depend on s , the conclusion of the lemma holds. \square

By the product formula of Wiener-Itô integrals (see, e.g., Nourdin and Peccati [2012] Theorem 2.7.10), the process $Z_{\gamma_1, \gamma_2}^N(t)$ in (8.91) can be rewritten as follows:

$$\begin{aligned} Z_{\gamma_1, \gamma_2}^N(t) &= A_N(\gamma_1, \gamma_2) \int_0^t ds \times \\ &\left[\int_{\mathbb{R}} (s, x_1)_N^{\gamma_1} B(dx_1) \int_{\mathbb{R}} (s, x_2)_N^{\gamma_2} B(dx_2) - \mathbb{E} \int_{\mathbb{R}} (s, x_1)_N^{\gamma_1} B(dx_1) \int_{\mathbb{R}} (s, x_2)_N^{\gamma_2} B(dx_2) \right] \end{aligned}$$

Note that by the scaling property of Brownian motion, for $j = 1, 2$,

$$\begin{aligned} X_{\gamma_j}^N(s) &:= \int_{\mathbb{R}} (s, x)_N^{\gamma_j} B(dx) = \int_{\mathbb{R}} \left(\frac{[Ns] - [Nx] + 1}{N} \right)^{\gamma_j} I_{\{[Ns] > [Nx]\}} B(dx) \\ &\stackrel{f.d.d.}{=} N^{-\gamma_j-1/2} \sum_{-\infty < i < [Ns]} ([Ns] - i + 1)^{\gamma_j} \epsilon_i, \end{aligned}$$

where ϵ_i 's are i.i.d. standard normal random variables, and $\stackrel{f.d.d.}{=}$ means equal in finite-dimensional distributions. Hence (recall that the Hurst index $H = \gamma_1 + \gamma_2 + 2$),

$$Z_{\gamma_1, \gamma_2}^N(t) \stackrel{f.d.d.}{=} A_N(\gamma_1, \gamma_2) \int_0^t [X_{\gamma_1}^N(s) X_{\gamma_2}^N(s) - \mathbb{E} X_{\gamma_1}^N(s) X_{\gamma_2}^N(s)] ds$$

$$= A_N(\gamma_1, \gamma_2) N^{-H} \sum_{n=1}^{[Nt]} [Y_{\gamma_1}(n) Y_{\gamma_2}(n) - \mathbf{E} Y_{\gamma_1}(n) Y_{\gamma_2}(n)] + R_N(t, \gamma_1, \gamma_2) \quad (8.93)$$

where

$$Y_{\gamma}(n) = \sum_{-\infty < i < n-1} (n-i)^{\gamma} \epsilon_i = \sum_{i=2}^{\infty} i^{\gamma} \epsilon_{n-i} \quad (8.94)$$

is a linear stationary sequence and

$$\begin{aligned} R_N(t, \gamma_1, \gamma_2) &= A_N(\gamma_1, \gamma_2) N^{-H} (Nt - [Nt]) \\ &\quad \times \left(Y_{\gamma_1}([Nt] + 1) Y_{\gamma_2}([Nt] + 1) - \mathbf{E} Y_{\gamma_1}([Nt] + 1) Y_{\gamma_2}([Nt] + 1) \right). \end{aligned} \quad (8.95)$$

We first show that this preceding remainder term is negligible:

Lemma 8.6.4.

$$\lim_{N \rightarrow \infty} \limsup_{(\gamma_1, \gamma_2) \rightarrow (-1/2, \gamma)} \mathbf{E} R_N(t, \gamma_1, \gamma_2)^2 = 0 \quad (8.96)$$

Proof. Since $Nt - [Nt] \leq 1$ and $Y_{\gamma}(n)$ is stationary, we can write

$$\mathbf{E} R_N(t, \gamma_1, \gamma_2)^2 \leq N^{-2H} A_N(\gamma_1, \gamma_2)^2 \left[\mathbf{E} Y_{\gamma_1}(0)^2 Y_{\gamma_2}(0)^2 - (\mathbf{E} Y_{\gamma_1}(0) Y_{\gamma_2}(0))^2 \right].$$

We have

$$\mathbf{E} Y_{\gamma_1}(0) Y_{\gamma_2}(0) = \sum_{i=2}^{\infty} i^{\gamma_1 + \gamma_2}, \quad \mathbf{E} Y_{\gamma_j}(0)^2 = \sum_{i=2}^{\infty} i^{2\gamma_j}, \quad j = 1, 2. \quad (8.97)$$

By the diagram formula (see, e.g., Janson [1997] Theorem 1.36), we have for jointly centered Gaussian variables (Y_1, Y_2) that $\mathbf{E} Y_1^2 Y_2^2 = 2(\mathbf{E} Y_1 Y_2)^2 + \mathbf{E} Y_1^2 \mathbf{E} Y_2^2$. Expressing this as $\mathbf{E} Y_1^2 Y_2^2 - (\mathbf{E} Y_1 Y_2)^2 = (\mathbf{E} Y_1 Y_2)^2 + \mathbf{E} Y_1^2 \mathbf{E} Y_2^2$, one gets

$$\mathbf{E} R_N(t, \gamma_1, \gamma_2)^2 \leq N^{-2H} A_N(\gamma_1, \gamma_2)^2 \left[\left(\sum_{i=2}^{\infty} i^{\gamma_1 + \gamma_2} \right)^2 + \left(\sum_{i=2}^{\infty} i^{2\gamma_1} \right) \left(\sum_{i=2}^{\infty} i^{2\gamma_2} \right) \right]. \quad (8.98)$$

The first and last sums remain bounded as $(\gamma_1, \gamma_2) \rightarrow (-1/2, \gamma)$, but this is not the case

for the second sum. Since the function $x^{2\gamma_1}$ is decreasing, we have for any integer $k \geq 0$,

$$\begin{aligned} (-2\gamma_1 - 1)^{-1}(k + 2)^{2\gamma_1+1} &= \int_2^\infty (x + k)^{2\gamma_1} dx \leq \int_2^\infty (x + k)^{\gamma_1} x^{\gamma_1} dx \\ &\leq \sum_{i=2}^\infty (i + k)^{\gamma_1} i^{\gamma_1} \leq \sum_{i=2}^\infty i^{2\gamma_1} \leq \int_1^\infty x^{2\gamma_1} dx = (-2\gamma_1 - 1)^{-1}. \end{aligned} \quad (8.99)$$

In particular, $\sum_{i=2}^\infty i^{2\gamma_1}$ explodes like $(-2\gamma_1 - 1)^{-1}$ as $\gamma_1 \rightarrow -1/2$. This, however, will be compensated by $A_N(\gamma_1, \gamma_2)^2$, since by (8.83) and (8.89), we have $A_N(\gamma_1, \gamma_2) \sim A(\gamma_1, \gamma_2) \asymp (-2\gamma_1 - 1)$ as $(\gamma_1, \gamma_2) \rightarrow (-1/2, \gamma)$. Hence (8.98) implies

$$\limsup_{(\gamma_1, \gamma_2) \rightarrow (-1/2, \gamma)} N^{2H} \mathbf{E} R_N(t, \gamma_1, \gamma_2)^2 < \infty,$$

which entails (8.96). □

The following lemma is key:

Lemma 8.6.5. *Let $Y_\gamma(n)$ be as in (8.94). As $(\gamma_1, \gamma_2) \rightarrow (-1/2, \gamma)$, one has the following joint convergence in distribution:*

$$\left(A(\gamma_1, \gamma_2) Y_{\gamma_1}(n), Y_{\gamma_2}(n) \right)_{n=1}^N \xrightarrow{d} \left(\sigma_\gamma W, Y_\gamma(n) \right)_{n=1}^N,$$

for any $N \in \mathbb{Z}_+$, where W is a standard normal random variable which is independent of $Y_\gamma(n)$, and

$$\sigma_\gamma = (2\gamma + 3)^{1/2} (\gamma + 1)^{1/2} \mathbf{B}(\gamma + 1, -2\gamma - 1)^{-1/2}. \quad (8.100)$$

Proof. Since $\left(A(\gamma_1, \gamma_2) Y_{\gamma_1}(n), Y_{\gamma_2}(n) \right)_{n=1}^N$ is always a centered and jointly Gaussian vector, we only need to show that its covariance structure converges to that of $\left(\sigma_\gamma W, Y_\gamma(n) \right)_{n=1}^N$. Let us first compute the covariance of $A(\gamma_1, \gamma_2) Y_{\gamma_1}$. By (8.89) and (8.99), we have for $m \geq n$ (similarly for $m < n$)

$$\mathbf{E} [A(\gamma_1, \gamma_2) Y_{\gamma_1}(n) A(\gamma_1, \gamma_2) Y_{\gamma_1}(m)]$$

$$\begin{aligned}
&= A(\gamma_1, \gamma_2)^2 \mathbf{E}[Y_{\gamma_1}(n)Y_{\gamma_1}(m)] \\
&\sim (2\gamma + 3)(\gamma + 1) \mathcal{B}(\gamma + 1, -2\gamma - 1)^{-1} (-2\gamma_1 - 1) \sum_{i=2}^{\infty} (i + m - n)^{\gamma_1} i^{\gamma_1} \\
&\sim (2\gamma + 3)(\gamma + 1) \mathcal{B}(\gamma + 1, -2\gamma - 1)^{-1} = \sigma_{\gamma}^2.
\end{aligned}$$

Since the limit is independent of n , the limit process is indeed a fixed Gaussian random variable, say $\sigma_{\gamma}W$.

We now focus on the cross-covariance between $A(\gamma_1, \gamma_2)Y_{\gamma_1}$ and Y_{γ_2} . We have for $m \geq n$ (similarly for $m < n$) that

$$\begin{aligned}
&\mathbf{E}[A(\gamma_1, \gamma_2)Y_{\gamma_1}(n)Y_{\gamma_2}(m)] \\
&\sim [(2\gamma + 3)(\gamma + 1) \mathcal{B}(\gamma + 1, -2\gamma - 1)^{-1} (-2\gamma_1 - 1)]^{1/2} \sum_{i=2}^{\infty} (i + m - n)^{\gamma_1} i^{\gamma} \rightarrow 0, \quad (8.101)
\end{aligned}$$

because $\sum_{i=2}^{\infty} i^{-1/2+\gamma} < \infty$. Thus we have asymptotic independence. Finally as $\gamma_2 \rightarrow \gamma$, the covariance structure of the second term Y_{γ_2} converges to that of Y_{γ} . The proof is then complete. \square

The following convergence of normalized sum of long-memory linear process to fractional Brownian motion can be found in Giraitis et al. [2012] Corollary 4.4.1, which was originally due to Davydov [1970].

Lemma 8.6.6. *Let $Y_{\gamma}(n)$ be as in (8.94). Then as $N \rightarrow \infty$*

$$Z_{\gamma}^N(t) := N^{-\gamma-2/3} \sum_{n=1}^{[Nt]} Y_{\gamma}(n) \xrightarrow{f.d.d.} \sigma_{\gamma}^{-1} B_{\gamma+3/2}(t)$$

where σ_{γ} is as in (8.100) and $B_{\gamma+3/2}(t)$ is a standard fractional Brownian motion with Hurst index $\gamma + 3/2$.

We are now ready to combine the last few lemmas into an alternate proof of Theorem 8.2.2.

Proof of Theorem 8.2.2. Tightness still follows from Lemma 8.3.8. To prove the convergence of the finite-dimensional distributions, namely, to prove that

$$Z_{\gamma_1, \gamma_2}(t) \xrightarrow{f.d.d.} WB_{\gamma+3/2} \quad \text{as } (\gamma_1, \gamma_2) \rightarrow (-1/2, \gamma),$$

it is sufficient to show that the following triangular approximation relations hold (see, e.g., Lemma 4.2.1 of Giraitis et al. [2012]):

$$\lim_{N \rightarrow \infty} \limsup_{(\gamma_1, \gamma_2) \rightarrow (-1/2, \gamma)} \mathbb{E} \left| Z_{\gamma_1, \gamma_2}(t) - \frac{A(\gamma_1, \gamma_2)}{A_N(\gamma_1, \gamma_2)} [Z_{\gamma_1, \gamma_2}^N(t) - R_N(t, \gamma_1, \gamma_2)] \right|^2 = 0, \quad (8.102)$$

$$\frac{A(\gamma_1, \gamma_2)}{A_N(\gamma_1, \gamma_2)} [Z_{\gamma_1, \gamma_2}^N(t) - R_N(t, \gamma_1, \gamma_2)] \xrightarrow{f.d.d.} \sigma_\gamma W Z_\gamma^N(t) \quad \text{as } (\gamma_1, \gamma_2) \rightarrow (-1/2, \gamma), \quad (8.103)$$

$$\sigma_\gamma W Z_\gamma^N(t) \xrightarrow{f.d.d.} WB_{\gamma+3/2}(t), \quad \text{as } N \rightarrow \infty. \quad (8.104)$$

The convergence (8.102) follows from Lemma 8.6.1, Lemma 8.6.2 and Lemma 8.6.4. For the convergence (8.103), we have by (8.93), Lemma 8.6.5 and (8.101) that

$$\begin{aligned} & \frac{A(\gamma_1, \gamma_2)}{A_N(\gamma_1, \gamma_2)} [Z_{\gamma_1, \gamma_2}^N(t) - R_N(t, \gamma_1, \gamma_2)] \\ &= N^{-H} \sum_{n=1}^{[Nt]} [A(\gamma_1, \gamma_2) Y_{\gamma_1}(n) Y_{\gamma_2}(n) - \mathbb{E} A(\gamma_1, \gamma_2) Y_{\gamma_1}(n) Y_{\gamma_2}(n)] \\ & \xrightarrow{f.d.d.} N^{-\gamma-3/2} \sum_{n=1}^{[Nt]} [\sigma_\gamma W Y_\gamma(n) - 0] = \sigma_\gamma W Z_\gamma^N(t). \end{aligned}$$

Finally, (8.104) follows from Lemma 8.6.6. \square

Chapter 9

A unified approach to self-normalized block sampling

The inference procedure for the mean of a stationary time series is usually quite different under various model assumptions because the partial sum process behaves differently depending on whether the time series is short or long-range dependent, or whether it has a light or heavy-tailed marginal distribution. In the current chapter, we develop an asymptotic theory for the self-normalized block sampling, and prove that the corresponding block sampling method can provide a unified inference approach for the aforementioned different situations in the sense that it does not require the a priori estimation of auxiliary parameters. Monte Carlo simulations are presented to illustrate its finite-sample performance. The R function implementing the method is available from the authors.

9.1 Introduction

Given samples X_1, \dots, X_n from a stationary process $\{X_i\}_{i \in \mathbb{Z}}$ with mean $\mu = E(X_0)$, the sample average $\bar{X}_n = n^{-1} \sum_{i=1}^n X_i$ serves as a natural estimator for the population mean μ . To conduct statistical inference on the mean μ such as hypothesis testing or the construction of confidence intervals, one needs an asymptotic theory on the sample average for dependent data. The development of such a theory has been an active area of research. Consider first the classical case, where by assuming certain short-range dependence conditions, one

obtains the usual central limit theorem, that is,

$$n^{1/2}(\bar{X}_n - \mu) \xrightarrow{d} N(0, \sigma^2), \quad (9.1)$$

where \xrightarrow{d} denotes the convergence in distribution, and σ^2 is the long-run variance which typically is the sum of autocovariances of all orders. The short-range dependence conditions mentioned above include, but are not limited to, the m -dependence condition of Hoeffding and Robbins [1948], the strong mixing condition of Rosenblatt [1956] and its variants, and the p -stability condition based on functional dependence measures of Wu [2005]; see also Ibragimov and Linnik [1971], Peligrad [1996], Maxwell and Woodroffe [2000], Bradley [2007], Wu [2011] and references therein. Once one has (9.1), an asymptotic $100(1 - \alpha)\%$ confidence interval of μ can be constructed as

$$[\bar{X}_n - n^{-1/2}\sigma q_{1-\alpha/2}, \bar{X}_n + n^{-1/2}\sigma q_{1-\alpha/2}] \quad (9.2)$$

where $q_{1-\alpha/2}$ is the $(1 - \alpha/2)$ -th quantile of the standard normal distribution. However, the implementation of (9.2) requires the estimation of a nuisance parameter σ , which can itself be a challenging problem and often relies on techniques including tapering and thresholding to achieve consistency; see for example Whitney and Kenneth [1987], Flegal and Jones [2010], Politis [2011] and Zhang and Wu [2012] among others.

If the process $(X_i)_{i \in \mathbb{Z}}$ is heavy-tailed (distributional tail behaving like $x^{-\alpha}$ with $\alpha \in (1, 2)$) so that the variance is infinite, one typically has

$$n^{1-1/\alpha} \ell(n)^{-1} (\bar{X}_n - \mu) \xrightarrow{d} S_\alpha(\sigma, \beta, 0), \quad (9.3)$$

where $\ell(n)$ is a slowly varying function satisfying $\lim_{n \rightarrow \infty} \ell(an)/\ell(n) = 1$ for any $a > 0$, and $S_\alpha(\sigma, \beta, 0)$ is the centered α -stable random variable with scale parameter $\sigma > 0$ and skewness parameter $\beta \in [-1, 1]$. We refer the reader to the monographs by Samorodnitsky and Taqqu [1994], Nolan [2015] and Resnick [2007] for an introduction. See also Adler

et al. [1998] for examples of heavy tails from finance, signal processing, networks, etc. Here the use of (9.3) for constructing confidence interval as in (9.2) becomes more difficult due to additional unknown parameters σ , α and β , as well as the unknown $\ell(n)$.

There has been a considerable amount of research focusing on the situation where the short-range dependence condition fails, and processes with *long-range dependence* (also called “long memory” or “strong dependence”) has attracted a lot of attention in various fields including econometrics, finance, hydrology and telecommunication among others; see for example Mandelbrot and Wallis [1968], Ding et al. [1993], Leland et al. [1994] and Baillie [1996]. We also refer the reader to the monographs by Doukhan et al. [2003], Giraitis et al. [2012] and Beran et al. [2013] for an introduction. For long-range dependent processes, it may be established that

$$n^{1-H} \ell(n)^{-1} (\bar{X}_n - \mu) \xrightarrow{d} Y, \quad (9.4)$$

where $H \in (1/2, 1)$ is the Hurst index (or the long memory index), $\ell(n)$ is a slowly varying function, and Y is typically a random variable which can be expressed by a multiple Wiener-Itô integral and is not necessarily Gaussian. The large sample theory of the form (9.4) has been studied by Davydov [1970], Taqqu [1975], Dobrushin and Major [1979], Avram and Taqqu [1987], Ho and Hsing [1997], Wu [2006] and Bai and Taqqu [2014a] among others. Therefore, the asymptotic behavior of the sample average and thus the inference procedure can become very different for long-range dependent processes, and the convergence rate in (9.4) depends critically on the Hurst index H which characterizes the dependence strength. Hence, in order to apply (9.4) for inference, unlike the case with short-range dependence and light tail, one needs to estimate in addition the Hurst index H and possibly the slowly varying function $\ell(n)$, which can be quite nontrivial. Furthermore, the distribution of a non-Gaussian Y (which also depends on H) has not been numerically evaluated in general. For the special case of the Rosenblatt distribution where it is evaluated, see Veillette and Taqqu [2013].

There has recently been a surge of attention in using some random normalizers to avoid,

or reduce the number of nuisance parameters that need to be estimated for statistical inference. For example, McElroy and Politis [2002] considered using the sample standard deviation as the normalizer for inference on the mean of heavy-tailed linear processes that satisfy the strong mixing condition; see also Romano and Wolf [1999] for the use of a similar normalizer for independent observations. Lobato [2001], Shao [2010], Zhou and Shao [2013] and Huang et al. [2015] used a normalization of the type

$$D_n = \left\{ n^{-1} \sum_{k=1}^n \left(\sum_{i=1}^k X_i - \frac{k}{n} \sum_{i=1}^n X_i \right)^2 \right\}^{1/2} \quad (9.5)$$

for finite-variance short-range dependent time series. Fan [2010] used the normalizer D_n for long-range dependent time series with finite variances. Results have also been obtained by McElroy and Politis [2013] using a lag-window normalizer instead of D_n in (9.5). McElroy and Politis [2007], moreover, considered the following non-centered stochastic volatility model $X_i = \mu + \sigma_i Z_i$, $i \geq 1$, where $\{\sigma_i\}$ and $\{Z_i\}$ are independent, $\{\sigma_i\}$ is i.i.d. heavy-tailed and $\{Z_i\}$ is a Gaussian process. They proposed to use a random normalizer involving two terms that account for heavy-tailedness and long memory respectively. The term in their normalizer which accounts for long memory requires the choice of an additional tuning parameter. Therefore, it seems that the specific form of the normalization depends critically on the particular time series that is being considered, and different normalizers have been used in the literature to account for the heavy-tail and/or long-range dependent characteristics of the time series.

The current chapter aims to provide a unified inference procedure by adopting the normalizer D_n in (9.5) and developing an asymptotic theory using self-normalized block sums. As observed by Shao [2011], self-normalization itself is not able to fully avoid the problem of estimating the nuisance parameters, as the asymptotic distribution at least depends on the unknown Hurst index H for long-range dependent processes. In order to provide a unified approach that does not rely on the estimation of any nuisance parameter to determine

the strength of dependence or heavy-tailedness, certain nonparametric techniques such as the block sampling¹ must be utilized to obtain the asymptotic quantiles. However, this requires developing an asymptotic theory on the self-normalized block sums for a general class of processes. This task may be nontrivial if we want it to include processes with long-range dependence and/or heavy-tails. Block sampling has been mainly studied in the literature in the *non-self-normalized* setting, where the normalizer converges in probability to a nonzero constant, thus simplifying the proof; see for example Hall et al. [1998] for nonlinear transforms of Gaussian processes, Nordman and Lahiri [2005] for linear processes, and Zhang et al. [2013] for nonlinear transforms of linear processes. Jach et al. [2012] applied block sampling to the model $X_i = \mu + \sigma_i Z_i$, $i \geq 1$, considered by McElroy and Politis [2007] but with Z_i replaced by $g(Z_i)$ where g is a possibly nonlinear function with Hermite rank one. For more information on block sampling, see Sherman and Carlstein [1996] and Lahiri [2003]. Betken and Wendler [2015] recently obtained interesting results in the context of long-range dependence. They are briefly discussed in Section 9.3.2 (see (9.58) below).

The current chapter considers *self-normalized* block sums using D_n in (9.5) as normalizer. As observed by Fan [2010], the development of an asymptotic theory in this case can be very nontrivial even for Gaussian processes. Developing a rigorous proof is stated as an open problem. The goal of this chapter is to develop such a proof for nonlinear functions of Gaussian processes with either short or long-range dependence, and including heavy-tails.

The remaining of the chapter is organized as follows. Section 9.2 introduces the self-normalized block sampling (SNBS) method, whose asymptotic theory is established in Section 9.3. Section 9.4 contains examples. Monte Carlo simulations are carried out in Section 9.5 to examine the finite-sample performance of the method.

¹The following terms are used interchangeably in the literature: block sampling, subsampling, sampling window method.

9.2 Self-Normalized Block Sampling

Let X_1, \dots, X_n be observations from a stationary process $(X_i)_{i \in \mathbb{Z}}$ with mean $\mu = E(X_0)$, and denote by $S_{j,k} = \sum_{i=j}^k X_i$, $j \leq k$, its partial sums from j to k . Of particular interest is $S_{1,n} = \sum_{i=1}^n X_i$. We propose using the self-normalized quantity

$$T_n^* = \frac{S_{1,n} - n\mu}{D_n} \quad (9.6)$$

for making statistical inference on the mean μ , where D_n , defined in (9.5), can now be written

$$D_n = \left\{ n^{-1} \sum_{k=1}^n \left(S_{1,k} - \frac{k}{n} S_{1,n} \right)^2 \right\}^{1/2}. \quad (9.7)$$

In order to make inference on μ , we need to know the distribution $P(T_n^* \leq x)$.

A first idea is to use the asymptotic distribution of (9.6). This would require knowing the weak limit of the normalized partial sum process, namely,

$$\{n^{-H} \ell(n)^{-1} (S_{[nt]} - n\mu), 0 \leq t \leq 1\} \Rightarrow \{Y(t), 0 \leq t \leq 1\}, \quad (9.8)$$

where $t \in [0, 1]$, $[nt]$ denotes the largest integer not exceeding nt , and \Rightarrow denotes weak convergence in Skorokhod space with suitable topology. By Lamperti [1962], if (9.8) holds, then the process $Y(t)$ is self-similar with stationary increments, with Hurst index² $0 < H < 1$ (H -sssi), and with $\ell(\cdot)$ a slowly varying function. Recall that a process $Y(t)$ is said to be self-similar with Hurst index H if $\{Y(ct), t \geq 0\}$ has the same finite-dimensional distributions as $\{c^H Y(t), t \geq 0\}$, for any $c > 0$.

The most important example of (9.8) is when $(X_i)_{i \in \mathbb{Z}}$ is short-range dependent and admits finite variance, in which case one expects

$$\{n^{-1/2} (S_{[nt]} - n\mu), 0 \leq t \leq 1\} \Rightarrow \{\sigma B(t), 0 \leq t \leq 1\}, \quad (9.9)$$

²We exclude the degenerate case $H = 1$.

where $B(\cdot)$ is the standard Brownian motion, and $\sigma^2 > 0$ is the long-run variance; see for example, the invariance principle of Herrndorf [1984] under strong mixing, and also the strong invariance principle of Wu [2007]. When $\{X_i\}$ is short-range dependent but has infinite variance with distributional tail regularly varying of order $-\alpha$ where $\alpha \in (1, 2)$, one has typically

$$\{n^{-1/\alpha}\ell(n)^{-1}(S_{\lfloor nt \rfloor} - n\mu), 0 \leq t \leq 1\} \Rightarrow \{L_{\alpha,\sigma,\beta}(t), 0 \leq t \leq 1\}, \quad (9.10)$$

where $L_{\alpha,\sigma,\beta}(t)$ is a centered α -stable Lévy process with scale parameter $\sigma > 0$ and skewness parameter $\beta \in [-1, 1]$. See, for example, Skorokhod [1957], Avram and Taqqu [1992], Tyran-Kamińska [2010a], Tyran-Kamińska [2010b] and Basrak et al. [2012] for the specification of the corresponding Skorohod topology.

Under long-range dependence, the limit in (9.8) can be quite complicated. A typical class of convergence in this case is

$$\{n^{-H}\ell(n)^{-1}(S_{\lfloor nt \rfloor} - n\mu), 0 \leq t \leq 1\} \Rightarrow \{cZ_{m,H}(t), 0 \leq t \leq 1\}, \quad (9.11)$$

where $1/2 < H < 1$, $Z_{m,H}(\cdot)$ is the m -th order Hermite process which can be expressed by a multiple Wiener-Itô integral (see, e.g., Dobrushin and Major [1979] and Taqqu [1979]), and c is a constant depending on H , m and $\ell(n)$. A Hermite process $Z_{m,H}(\cdot)$ with $m \geq 2$ is non-Gaussian, and when $m = 1$ it is the Gaussian process called fractional Brownian motion, also denoted by $B_H(\cdot)$. One can also consider the anti-persistent case $H < 1/2$, where the limit can be more complicated than $Z_{m,H}(\cdot)$ (see Major [1981]).

Applying the *same* normalization $n^{-H}\ell(n)^{-1}$ to both the numerator and denominator of T_n^* in (9.6), one can establish as in Lobato [2001], via (9.8) and the Continuous Mapping Theorem that as $n \rightarrow \infty$,

$$T_n^* = \frac{n^{-H}\ell(n)^{-1}(S_{1,n} - n\mu)}{n^{-H}\ell(n)^{-1} \left\{ n^{-1} \sum_{k=1}^n (S_{1,k} - \frac{k}{n}S_{1,n})^2 \right\}^{1/2}} \xrightarrow{d} T := \frac{Y(1)}{D}, \quad (9.12)$$

with

$$D = \left[\int_0^1 \{Y(s) - sY(1)\}^2 ds \right]^{1/2}. \quad (9.13)$$

Note that $D > 0$ almost surely. Indeed, if $P(D = 0) > 0$, then with positive probability $Y(s) = sY(1)$, which has locally bounded variation. This cannot happen by Theorem 3.3 of Vervaat [1985], since we assume $H < 1$.

In particular, in the short-range dependent case (9.9), one gets

$$T_n^* \xrightarrow{d} \frac{B(1)}{\left[\int_0^1 \{B(s) - sB(1)\}^2 ds \right]^{1/2}},$$

where the limit does not depend on any nuisance parameter. However, this nice property no longer holds in the other cases (9.10) and (9.11), since $Y(t)$ in either case involves additional parameters. Therefore, except for short-range dependent light-tailed processes, self-normalization itself is usually not able to fully avoid the problem of estimating the nuisance parameters, and we shall follow here Hall et al. [1998] and consider a block sampling approach. See also Chapter 5 of Politis et al. [1999]. Let

$$T_{i,b_n}^* = \frac{S_{i,i+b_n-1} - b_n\mu}{\sqrt{b_n^{-1} \sum_{k=i}^{i+b_n-1} (S_{i,k} - b_n^{-1}(k-i+1)S_{i,i+b_n-1})^2}} =: \frac{S_{i,i+b_n-1} - b_n\mu}{D_{i,b_n}}, \quad (9.14)$$

$1 \leq i \leq n - b_n + 1$, which is the block version of T_n^* in (9.6) for the subsample X_i, \dots, X_{i+b_n-1} , where b_n denotes the block size. Observe that there is a considerable overlap between successive blocks, since as i increases to $i + 1$, the subsample becomes $X_{i+1}, \dots, X_{i+b_n}$, and thus includes many of the same observations.

We consider using the empirical distribution function

$$\widehat{F}_{n,b_n}^*(x) = \frac{1}{n - b_n + 1} \sum_{i=1}^{n-b_n+1} \mathbf{I}(T_{i,b_n}^* \leq x), \quad (9.15)$$

where $\mathbf{I}(\cdot)$ is the indicator function, to approximate the distribution $P(T_n^* \leq x)$ of T_n^* in (9.6). In practice, the mean μ in (9.14) is unknown and we shall replace it by the average

\bar{X}_n of the whole sample, which turns (9.14) into

$$T_{i,b_n} = \frac{S_{i,i+b_n-1} - b_n \bar{X}_n}{\sqrt{b_n^{-1} \sum_{k=i}^{i+b_n-1} (S_{i,k} - b_n^{-1}(k-i+1)S_{i,i+b_n-1})^2}}, \quad (9.16)$$

whose empirical distribution function is given by

$$\hat{F}_{n,b_n}(x) = \frac{1}{n - b_n + 1} \sum_{i=1}^{n-b_n+1} \mathbf{I}(T_{i,b_n} \leq x). \quad (9.17)$$

The asterisk in T_{i,b_n}^* indicates that the centering involves the unknown population mean μ , in contrast to T_{i,b_n} , where the centering involves instead the sample average \bar{X}_n . We call the above inference procedure involving using $\hat{F}_{n,b_n}(x)$ in (9.17) to approximate the distribution of T_n^* in (9.6), the *self-normalized block sampling (SNBS)* method. One can then construct confidence intervals or test hypotheses for the unknown population mean μ . For instance, to construct a one-sided $100(1 - \alpha)\%$ confidence interval for μ , one gets first the α -th quantile q_α of the empirical distribution $\hat{F}_{n,b_n}(x)$ in (9.17). Since

$$1 - \alpha \approx P(T_n^* \geq q_\alpha) = P\left(\frac{S_{1,n} - n\mu}{D_n} \geq q_\alpha\right) = P(\mu \leq \bar{X}_n - q_\alpha D_n/n),$$

where D_n is defined in (9.7), then the $100(1 - \alpha)\%$ confidence interval is constructed as

$$(-\infty, \bar{X}_n - q_\alpha D_n/n]. \quad (9.18)$$

The idea of using block sampling to approximate distributions of self-normalized quantities is not new, and it has been applied by Fan [2010] and McElroy and Politis [2013] to long-range dependent processes with finite variances. However, the aforementioned papers did not provide a full theoretical justification for their inference procedure based on block sampling, and as commented by Fan [2010] such a task can be very nontrivial even for Gaussian processes and has been stated as an open problem. In addition, the aforementioned papers only considered the situation with finite variances, and therefore it has not

been known whether one could unify the inference procedure for processes with long-range dependence and/or heavy-tails.

Recently, Jach et al. [2012] considered this problem in the setting of stochastic volatility models where the error term can be nicely decomposed into two independent factors, with one being a function of long-range dependent Gaussian processes while the other being i.i.d. heavy-tailed³. But in their paper, the nonlinear function is restricted to have Hermite rank one and the choice of slowly varying functions is also greatly limited as neither $\log n$ nor $\log \log n$ are allowed. In addition, their random normalizer is specifically tailored to the aforementioned stochastic volatility model, and involves two different terms to account for the long-range dependent and heavy-tailed characteristics of the time series. Furthermore, the term in their normalizer that accounts for long-range dependence also requires the choice of an additional tuning parameter as in the estimation of the long-run variance for short-range dependent processes. We also mention that the proof of Jach et al. [2012], which relies on the θ -weak dependence, does not seem to be applicable in the current setting, since using our random normalizer D_n in the denominator makes the self-normalized quantity a non-Lipschitz function of the data.

The current chapter proposes to consider the use of (9.17) to provide a unified inference procedure without the estimation of a nuisance parameter for a wide class of processes, where the limit of the partial sum process can be a Brownian motion, an α -stable Lévy process, a Hermite process or other processes. In Section 9.3, we develop an asymptotic theory for the self-normalized block sums and establish the theoretical consistency of the aforementioned method, namely,

$$|\widehat{F}_{n,b_n}(x) - P(T_n^* \leq x)| \rightarrow 0 \tag{9.19}$$

in probability as $n \rightarrow \infty$.

³As noted in Section 9.4 below, we can recover the consistency result of Jach et al. [2012] by replacing our normalization D_n by the one found in that paper.

9.3 Asymptotic Theory

We establish the asymptotic consistency of self-normalized block sampling for the following two classes of stationary processes: (a) nonlinear transforms of Gaussian stationary processes (called Gaussian subordination), and (b) those satisfying strong mixing conditions. The first allows for long-range dependence and non-central limits, while the second involves short-range dependent processes. Both classes allow for heavy-tails with infinite variance.

Let $D[0, 1]$ be the space of càdlàg (right continuous with left limits) functions defined on $[0, 1]$, endowed with Skorokhod's M_2 topology. The M_2 topology is weaker than the other topologies proposed by Skorokhod [1956], in particular, weaker than the most commonly used J_1 topology. A sequence of function $x_n(t) \in D[0, 1]$ converges to $x(t) \in D[0, 1]$ in M_2 topology as $n \rightarrow \infty$, if and only if $\lim_n \sup_{t_1 \leq t \leq t_2} x_n(t) = \sup_{t_1 \leq t \leq t_2} x(t)$ and $\lim_n \inf_{t_1 \leq t \leq t_2} x_n(t) = \inf_{t_1 \leq t \leq t_2} x(t)$ for any t_1, t_2 at continuity points of $x(t)$ (see statement 2.2.10 of Skorokhod [1956]).

We consider the M_2 topology instead of J_1 since there are known examples in the heavy tailed case where convergence fails under J_1 but holds under M_2 (see Avram and Taqqu [1992], Tyran-Kamińska [2010b] and Basrak et al. [2012]). To apply the continuous mapping argument, we need the following lemma.

Lemma 9.3.1. Integration on $[0, 1]$ is a continuous functional for $D[0, 1]$ under the M_2 topology.

Proof. Suppose that $x_n(t) \rightarrow x(t)$ in the M_2 topology. For any partition $\mathcal{T} = \{0 = t_0 < t_1 < \dots < t_{k-1} < t_k = 1\}$, define $m_{i,n} = \inf_{t_{i-1} \leq t \leq t_i} x_n(t)$, $M_{i,n} = \sup_{t_{i-1} \leq t \leq t_i} x_n(t)$, $m_i = \inf_{t_{i-1} \leq t \leq t_i} x(t)$ and $M_i = \sup_{t_{i-1} \leq t \leq t_i} x(t)$, $i = 1, \dots, k$. Note that

$$\begin{aligned} \sum_{i=1}^k m_{i,n}(t_i - t_{i-1}) &\leq \int_0^1 x_n(t) dt \leq \sum_{i=1}^k M_{i,n}(t_i - t_{i-1}), \\ \sum_{i=1}^k m_i(t_i - t_{i-1}) &\leq \int_0^1 x(t) dt \leq \sum_{i=1}^k M_i(t_i - t_{i-1}). \end{aligned} \tag{9.20}$$

The function $x(t)$ is Riemann integrable since, as an element in $D[0, 1]$, it is a.e. continuous and bounded on $[0, 1]$. Riemann integrability implies that for any $\epsilon > 0$, one can choose a partition \mathcal{T} so that

$$0 \leq \sum_{i=1}^k M_i(t_i - t_{i-1}) - \sum_{i=1}^k m_i(t_i - t_{i-1}) < \epsilon. \quad (9.21)$$

Modify the partition, if necessary, so that all the t_i 's are at continuity points of $x(t)$, without changing (9.21). This is possible since $x(t)$ has at most countable discontinuity points and is bounded. By the characterization of convergence in $D[0, 1]$ with M_2 topology, we have

$$\begin{aligned} \lim_n \sum_{i=1}^k m_{i,n}(t_i - t_{i-1}) &= \sum_{i=1}^k m_i(t_i - t_{i-1}), \\ \lim_n \sum_{i=1}^k M_{i,n}(t_i - t_{i-1}) &= \sum_{i=1}^k M_i(t_i - t_{i-1}). \end{aligned} \quad (9.22)$$

Combining (9.20), (9.21) and (9.22) concludes that $\limsup_n |\int_0^1 x_n(t)dt - \int_0^1 x(t)dt| \leq \epsilon$. \square

9.3.1 Results in the Gaussian subordination case

Let

$$\{\mathbf{Z}_i = (Z_{i,1}, \dots, Z_{i,J}), i \in \mathbb{Z}\} \quad (9.23)$$

be an \mathbb{R}^J -valued Gaussian stationary process satisfying $EZ_{i,j} = 0$ for any i, j . Define

$$\mathbf{Z}_p^q = (\mathbf{Z}_p, \dots, \mathbf{Z}_q). \quad (9.24)$$

We shall view \mathbf{Z}_p^q as a vector of dimension $J \times (q-p+1)$ involving observations from time p to time q . The covariance matrix of \mathbf{Z}_1^m will be written for convenience as a four-dimensional

array involving i_1, i_2, j_2, j_2 :

$$\Sigma_m = \left(\gamma_{j_1, j_2}(i_2 - i_1) := \mathbb{E}Z_{i_1, j_1} Z_{i_2, j_2} \right)_{1 \leq i_1, i_2 \leq m, 1 \leq j_1, j_2 \leq J}. \quad (9.25)$$

We assume throughout that Σ_m is non-singular for every $m \in \mathbb{Z}_+$. The cross-block covariance matrix between \mathbf{Z}_1^m and \mathbf{Z}_{k+1}^{k+m} is

$$\Sigma_{k,m} = \left(\gamma_{j_1, j_2}(i_2 + k - i_1) := \mathbb{E}Z_{i_1, j_1} Z_{i_2+k, j_2} \right)_{1 \leq i_1, i_2 \leq m, 1 \leq j_1, j_2 \leq J}. \quad (9.26)$$

Let $\rho(\cdot, \cdot)$ denote the canonical correlation (maximum correlation coefficient) between $L^2(\Omega)$ random vectors $\mathbf{U} = (U_1, \dots, U_p)$ and $\mathbf{V} = (V_1, \dots, V_q)$. Let $\langle \cdot, \cdot \rangle$ denote the inner product in an Euclidean space of a suitable dimension. Then

$$\rho(\mathbf{U}, \mathbf{V}) = \sup_{\mathbf{x} \in \mathbb{R}^p, \mathbf{y} \in \mathbb{R}^q} \left| \text{Corr}(\langle \mathbf{x}, \mathbf{U} \rangle, \langle \mathbf{y}, \mathbf{V} \rangle) \right|. \quad (9.27)$$

Let $\rho_{k,m}$ be the between-block canonical correlation:

$$\rho_{k,m} = \rho \left(\mathbf{Z}_1^m, \mathbf{Z}_{k+1}^{k+m} \right). \quad (9.28)$$

We now introduce the assumptions for the self-normalized block sampling procedure. $\{X_i\}$ is the stationary process (time series) we observe.

A1. $X_i = G(\mathbf{Z}_i, \dots, \mathbf{Z}_{i-l}) = G(\mathbf{Z}_{i-l}^i)$ with mean $\mu = \mathbb{E}X_i$, where $\{\mathbf{Z}_i\}$ is a vector-valued stationary Gaussian process as in (9.23), and l is a fixed non-negative integer.

A2. We have weak convergence in $D[0, 1]$ endowed with the M_2 topology for the partial sum:

$$\left\{ \frac{1}{n^H \ell(n)} (S_{[nt]} - n\mu), 0 \leq t \leq 1 \right\} \Rightarrow \{Y(t), 0 \leq t \leq 1\},$$

for some nonzero H -sssi process $Y(t)$, where $0 < H < 1$ and $\ell(\cdot)$ is a slowly varying function.

A3. As $n \rightarrow \infty$, the block size $b_n \rightarrow \infty$, $b_n = o(n)$, and satisfies

$$\sum_{k=0}^n \rho_{k, l+b_n} = o(n), \quad (9.29)$$

where $\rho_{k,m}$ is the between-block canonical correlation defined in (9.28).

Remark 9.3.1. The data-generating specification in A1 allows us to get a variety of limits in A2, covering short-range dependence, long-range dependence, and heavy tails. When the covariance function of $X(n)$ is absolutely summable (short-range dependence), one typically gets in A2 convergence to Brownian motion (see, e.g., Breuer and Major [1983], Ho and Sun [1987] and Chambers and Slud [1989]). When the covariance of $X(n)$ is regularly varying of order between -1 and 0 (long-range dependence), one may get in A2 convergence to the Hermite-type processes (see, e.g., Taqqu [1975], Dobrushin and Major [1979], Taqqu [1979] and Arcones [1994]).

Moreover, as shown in Sly and Heyde [2008] in the case $J = 1$, when $G(\cdot)$ is chosen such that $X(n)$ is short-range dependent and heavy-tailed, so that $X(n)$ has infinite variance but finite mean, one can obtain in A2, convergence to an infinite-variance α -stable Lévy process; if $X(n)$ is long-range dependent and heavy-tailed, then the limit may be a finite-variance Hermite process, even though $X(n)$ may have infinite variance. All these situations are allowed under Assumptions A1–A3.

For sufficient conditions for Assumption A3 to hold, see Proposition 9.3.1 and Section 9.3.2.

Since the denominators in (9.12) are nonzero almost surely, Assumption A2, Lemma 9.3.1 and the Continuous Mapping Theorem imply the following (see Kallenberg [2006], Corollary 4.5):

Lemma 9.3.2. T_{i,b_n}^* in (9.14) converges in distribution to T in (9.12).

The following result allows us to relate the correlation of nonlinear functions to the correlation of linear functions.

Lemma 9.3.3. Let $(\mathbf{Z}_i)_{i \in \mathbb{Z}}$ be a centered \mathbb{R}^J -valued Gaussian stationary process as in (9.23), and let \mathbf{Z}_p^q be defined as in (9.24). Let \mathcal{F}_{Jm} be the set of all functions F on \mathbb{R}^{Jm} satisfying $\mathbb{E}F(\mathbf{Z}_1^m)^2 < \infty$. Then for $k \geq m$, one has

$$\sup_{F, G \in \mathcal{F}_{Jm}} \left| \text{Corr}(F(\mathbf{Z}_1^m), G(\mathbf{Z}_{k+1}^{k+m})) \right| = \rho(\mathbf{Z}_1^m, \mathbf{Z}_{k+1}^{k+m}) = \rho_{k,m}. \quad (9.30)$$

Proof. The equality is the well-known Gaussian maximal correlation equality. See, e.g., Theorem 1 of Kolmogorov and Rozanov [1960] or Theorem 10.11 of Janson [1997]. \square

Our goal is to show that (9.19) holds, namely, \widehat{F}_{n,b_n} is a consistent estimator of $P(T_n^* \leq x)$. This will be a consequence of the following theorem.

Theorem 9.3.1. Assume that Assumptions A1–A3 hold. Let $F(x)$ be the CDF (cumulative distribution function) of T in (9.12), and let $\widehat{F}_{n,b_n}(x)$ be as in (9.17). As $n \rightarrow \infty$, we have

$$\widehat{F}_{n,b_n}(x) \xrightarrow{p} F(x), \quad x \in C(F), \quad (9.31)$$

where $C(F)$ denotes the set of continuity points of $F(x)$. If $F(x)$ is continuous, then (9.31) can be strengthened to

$$\sup_x \left| \widehat{F}_{n,b_n}(x) - F(x) \right| \rightarrow 0 \quad \text{in probability.} \quad (9.32)$$

Proof.

Step 1. Let $\widehat{F}_{n,b_n}^*(x)$ be as in (9.15). To prove (9.31), we first show that

$$\widehat{F}_{n,b_n}^*(x) \xrightarrow{p} F(x), \quad x \in C(F), \quad (9.33)$$

where we have replaced $\widehat{F}_{n,b_n}(x)$ by $\widehat{F}_{n,b_n}^*(x)$. A bias-variance decomposition yields:

$$\begin{aligned} & \mathbb{E} \left(\left[\widehat{F}_{n,b_n}^*(x) - F(x) \right]^2 \right) \\ &= [\mathbb{E}\widehat{F}_{n,b_n}^*(x)]^2 - \mathbb{E}[2F(x)\widehat{F}_{n,b_n}^*(x)] + F(x)^2 + \mathbb{E}[\widehat{F}_{n,b_n}^*(x)^2] - [\mathbb{E}\widehat{F}_{n,b_n}^*(x)]^2 \end{aligned}$$

$$\begin{aligned}
&= \left[\mathbb{E} \widehat{F}_{n,b_n}^*(x) - F(x) \right]^2 + \left[\mathbb{E} [\widehat{F}_{n,b_n}^*(x)^2] - [\mathbb{E} \widehat{F}_{n,b_n}^*(x)]^2 \right] \\
&= \left[P(T_{i,b_n}^* \leq x) - P(T \leq x) \right]^2 + \text{Var} [\widehat{F}_{n,b_n}^*(x)].
\end{aligned}$$

By Lemma 9.3.2, the squared bias $[P(T_{i,b_n}^* \leq x) - P(T \leq x)]^2$ converges to zero for $x \in C(F)$ as $b_n \rightarrow \infty$. We thus need to show that $\text{Var}[\widehat{F}_{n,b_n}^*(x)] \rightarrow 0$. By the stationarity of $\{X_i\}$, which implies the stationarity of $\{T_{i,b_n}^*\}$ viewed as a process indexed by i , one has

$$\begin{aligned}
\text{Var}[\widehat{F}_{n,b_n}^*(x)] &= \text{Var} \left[\frac{1}{n-b_n+1} \sum_{i=1}^{n-b_n+1} \mathbb{I}\{T_{i,b_n}^* \leq x\} \right] \\
&= \frac{1}{(n-b_n+1)^2} \sum_{i,j=1}^{n-b_n+1} \text{Cov} [\mathbb{I}\{T_{i,b_n}^* \leq x\}, \mathbb{I}\{T_{j,b_n}^* \leq x\}] \\
&\leq \frac{2}{n-b_n+1} \sum_{k=0}^n |\text{Cov} [\mathbb{I}\{T_{1,b_n}^* \leq x\}, \mathbb{I}\{T_{k+1,b_n}^* \leq x\}]|, \tag{9.34}
\end{aligned}$$

since for any covariance function $\gamma(\cdot)$ of a stationary sequence, we have

$$\sum_{i,j=1}^p |\gamma(i-j)| \leq \sum_{|k|<p} (p-|k|) |\gamma(k)| \leq 2p \sum_{k=0}^p |\gamma(k)|.$$

In view of Assumption A1, X_i depends on $\mathbf{Z}_i, \dots, \mathbf{Z}_{i-l}$. By (9.14), T_{i,b_n}^* is a function of X_i, \dots, X_{i+b_n-1} . Hence T_{1,b_n}^* depends not only on $\mathbf{Z}_1, \dots, \mathbf{Z}_{b_n}$, but also on $\mathbf{Z}_{1-l}, \dots, \mathbf{Z}_0$, and T_{k+1,b_n}^* depends on $\mathbf{Z}_{k+1-l}, \dots, \mathbf{Z}_{k+b_n}$. We shall now apply Lemma 9.3.3 with the same k and $m = l + b_n$. Then when $k \geq l + b_n$, one has

$$|\text{Cov}[\mathbb{I}\{T_{1,b_n}^* \leq x\}, \mathbb{I}\{T_{k+1,b_n}^* \leq x\}]| \leq \frac{1}{4} |\text{Corr}[\mathbb{I}\{T_{1,b_n}^* \leq x\}, \mathbb{I}\{T_{k+1,b_n}^* \leq x\}]| \leq \frac{1}{4} \rho_{k,b_n+l} \tag{9.35}$$

where we have used the following fact⁴: if $0 \leq X \leq 1$, then $\text{Var}[X] \leq 1/4$. We have

$$\text{Var}[\widehat{F}_{n,b_n}^*(x)] \leq \frac{1}{2(n-b_n+1)} \sum_{k=0}^n \rho_{k,b_n+l}, \quad (9.36)$$

which converges to zero because of Assumption A3. Hence $\widehat{F}_{n,b_n}^*(x) \xrightarrow{p} F(x)$ for $x \in C(F)$.

Step 1 of the proof is now complete.

Step 2. We now show that

$$\widehat{F}_{n,b_n}(x) \xrightarrow{p} F(x) \quad \text{for } x \in C(F),$$

that is, we go from (9.33) to (9.31). To do so, we follow the proof of Theorem 11.3.1 of Politis et al. [1999], and express (9.17) as

$$\widehat{F}_{n,b_n}(x) = \frac{1}{n-b_n+1} \sum_{i=1}^{n-b_n+1} \mathbf{I}\{T_{i,b_n}^* \leq x + b_n(\bar{X}_n - \mu)/D_{i,b_n}\}, \quad (9.37)$$

where D_{i,b_n} is as in (9.14). The goal is to show that $b_n(\bar{X}_n - \mu)/D_{i,b_n}$ is negligible. For $\epsilon > 0$, define

$$\begin{aligned} R_n(\epsilon) &= \frac{1}{n-b_n+1} \sum_{i=1}^{n-b_n+1} \mathbf{I}\{b_n(\bar{X}_n - \mu)/D_{i,b_n} \leq \epsilon\} \\ &= \frac{1}{n-b_n+1} \sum_{i=1}^{n-b_n+1} \mathbf{I}\{(b_n^H \ell(b_n))^{-1} D_{i,b_n} \geq \epsilon^{-1} b_n(\bar{X}_n - \mu)(b_n^H \ell(b_n))^{-1}\}. \end{aligned} \quad (9.38)$$

Since $R_n(\epsilon)$ is an average of indicators, we have $R_n(\epsilon) \leq 1$. Our goal is to show that $R_n(\epsilon) \xrightarrow{p} 1$. Note that as $n \rightarrow \infty$,

$$\frac{D_{i,b_n}}{b_n^H \ell(b_n)} = \frac{1}{b_n^H \ell(b_n)} \left(b_n^{-1} \sum_{k=i}^{i+b_n-1} \left(S_{i,k} - b_n^{-1}(k-i-1)S_{i,i+b_n-1} \right)^2 \right)^{1/2}$$

converges in distribution to D in (9.13) by Assumption A2 and continuous mapping. More-

⁴If $0 \leq X \leq 1$, then $\mu = EX \in [0, 1]$, $EX^2 \leq \mu$ and $\text{Var}[X] \leq \mu - \mu^2$ is maximized at $\mu = 1/2$, so that $\text{Var}[X] \leq 1/4$ (for more general results, see Dharmadhikari and Joag-Dev [1989], Lemma 2.2).

over, since $b_n = o(n)$, $H < 1$ and $n(\bar{X}_n - \mu)n^{-H}\ell(n)^{-1}$ converges in distribution to $Y(1)$ by Assumption A2, we have

$$b_n(\bar{X}_n - \mu)(b_n^H \ell(b_n))^{-1} = n(\bar{X}_n - \mu)n^{-H}\ell(n)^{-1} \frac{n^{H-1}\ell(n)}{b_n^{H-1}\ell(b_n)} \xrightarrow{p} 0.$$

Hence for any $\delta > 0$, with probability tending to 1 as $n \rightarrow \infty$, one has

$$1 \geq R_n(\epsilon) \geq \frac{1}{n - b_n + 1} \sum_{i=1}^{n-b_n+1} \mathbf{I}\{(b_n^H \ell(b_n))^{-1} D_{i,b_n} \geq \delta \epsilon^{-1}\}. \quad (9.39)$$

Since as T_{i,b_n}^* in Step 1, D_{i,b_n} is also a function of X_i, \dots, X_{i+b_n-1} , we can follow a same argument as in Step 1, replacing T_{i,b_n}^* by $(b_n^H \ell(b_n))^{-1} D_{i,b_n}$ to obtain a similar result as in (9.33), namely that the empirical distribution of $(b_n^H \ell(b_n))^{-1} D_{i,b_n}$ converges in probability to that of D at all points of continuity of the distribution of D . Therefore

$$\frac{1}{n - b_n + 1} \sum_{i=1}^{n-b_n+1} \mathbf{I}\{(b_n^H \ell(b_n))^{-1} D_{i,b_n} \geq \delta \epsilon^{-1}\} \xrightarrow{p} P(D \geq \delta \epsilon^{-1}) \quad (9.40)$$

for $\delta \epsilon^{-1}$ at continuity point of the CDF of D . Since $P(D > 0) = 1$, we can choose δ small enough to make $P(D \geq \delta \epsilon^{-1})$ as close to 1 as desired. In view of (9.39) and (9.40), we conclude that as $n \rightarrow \infty$,

$$R_n(\epsilon) \xrightarrow{p} 1 \quad (9.41)$$

for any $\epsilon > 0$. Now notice that each summand in the sum (9.37) satisfies

$$\begin{aligned} & \mathbf{I}\{T_{i,b_n}^* \leq x + b_n(\bar{X}_n - \mu)/D_{i,b_n}\} \\ &= \left[\mathbf{I}\{T_{i,b_n}^* \leq x + b_n(\bar{X}_n - \mu)/D_{i,b_n}\} \right] \left[\mathbf{I}\{b_n(\bar{X}_n - \mu)/D_{i,b_n} \leq \epsilon\} + \mathbf{I}\{b_n(\bar{X}_n - \mu)/D_{i,b_n} > \epsilon\} \right] \\ &\leq \mathbf{I}\{T_{i,b_n}^* \leq x + \epsilon\} + \mathbf{I}\{b_n(\bar{X}_n - \mu)/D_{i,b_n} > \epsilon\}, \end{aligned} \quad (9.42)$$

so that by plugging these inequalities in (9.37) and using (9.38), we get

$$\widehat{F}_{n,b_n}(x) \leq \widehat{F}_{n,b_n}^*(x + \epsilon) + 1 - R_n(\epsilon).$$

But by (9.41), $R_n(\epsilon) \xrightarrow{p} 1$. So for any $\gamma > 0$, one has

$$\widehat{F}_{n,b_n}(x) \leq \widehat{F}_{n,b_n}^*(x + \epsilon) + \gamma$$

with probability tending to 1 as $n \rightarrow \infty$. We can now use (9.33) to replace $\widehat{F}_{n,b_n}^*(x + \epsilon)$ by $F(x + \epsilon)$, so that for arbitrary $\gamma' > \gamma$, and for any $x + \epsilon \in C(F)$, one has $\widehat{F}_{n,b_n}(x) \leq F(x + \epsilon) + \gamma'$ with probability tending to 1 as $n \rightarrow \infty$. Now letting $\epsilon \downarrow 0$ through $x + \epsilon \in C(F)$ and using the continuity of $F(\cdot)$ at x , one gets with probability tending to 1 that

$$\widehat{F}_{n,b_n}(x) \leq F(x) + \gamma'', \quad x \in C(F), \quad (9.43)$$

for any $\gamma'' > \gamma'$.

A similar argument, which replaces (9.42) by

$$\mathbb{I}\{T_{i,b_n} \leq x\} \geq \mathbb{I}\{T_{i,b_n}^* \leq x - \epsilon\} - \mathbb{I}\{b_n(\bar{X}_n - \mu)/D_{i,b_n} < -\epsilon\},$$

will show that for any $\gamma'' > 0$, with probability tending to 1,

$$\widehat{F}_{n,b_n}(x) \geq F(x) - \gamma'', \quad x \in C(F). \quad (9.44)$$

Combining (9.43) and (9.44), one gets

$$P(|\widehat{F}_{n,b_n}(x) - F(x)| \leq \gamma'') \rightarrow 1$$

as $n \rightarrow \infty$, and thus (9.31) holds.

Step 3. We now show (9.32). If $F(x)$ is continuous, then by the already established

(9.31), we have $\widehat{F}_{n,b_n}(x) \rightarrow F(x)$ in probability for any $x \in \mathbb{R}$. Let n_i be an arbitrary subsequence, one can then choose a further subsequence of n_i , still denoted as n_i , so that $\widehat{F}_{n_i}(x) \rightarrow F(x)$ almost surely for all rational x by a diagonal subsequence argument. Then by Lemma A9.2 (ii) of Gut [2006], $\sup_{x \in \mathbb{R}} |\widehat{F}_{n_i}(x) - F(x)| \rightarrow 0$ almost surely, and therefore $\sup_{x \in \mathbb{R}} |\widehat{F}_n(x) - F(x)| \rightarrow 0$ in probability. Hence (9.32) is proved. \square

Consistency (9.19) is a simple corollary of Theorem 9.3.1.

Corollary 9.3.1. Assume that Assumptions A1–A3 hold. Then as $n \rightarrow \infty$,

$$|\widehat{F}_{n,b_n}(x) - P(T_n^* \leq x)| \rightarrow 0 \quad \text{in probability.} \quad (9.45)$$

for $x \in C(F)$. If $F(x)$ is continuous, then the preceding convergence can be strengthened to

$$\sup_{x \in \mathbb{R}} |\widehat{F}_{n,b_n}(x) - P(T_n^* \leq x)| \rightarrow 0 \quad \text{in probability.} \quad (9.46)$$

Proof. The first result (9.45) follows directly from the triangle inequality

$$|\widehat{F}_{n,b_n}(x) - P(T_n^* \leq x)| \leq |\widehat{F}_{n,b_n}(x) - F(x)| + |P(T_n^* \leq x) - F(x)|,$$

where $x \in C(F)$ and $F(x) = P(T \leq x)$, by combining Theorem 9.3.2 or 9.3.1 with (9.12). For the second result (9.46), one uses also the fact that (9.12) implies $\sup_{x \in \mathbb{R}} |P(T_n^* \leq x) - F(x)| \rightarrow 0$ as $n \rightarrow \infty$ if $F(x)$ is continuous (see again Lemma A9.2 (ii) of Gut [2006]). \square

Bai and Taqqu [2015e] recently proved the following proposition, showing that the bound (9.29) holds for a large class of models with long-range dependence. Thus, for these models, one has the freedom to choose any $b_n = o(n)$, irrespective of the long-range dependence parameter H .

Proposition 9.3.1 (Bai and Taqqu [2015e], Theorem 2.2 and 2.3). Consider the case $J = 1$.

Suppose that the spectral density of the underlying Gaussian $\{Z_i\}$ is given by

$$f(\lambda) = f_H(\lambda)f_0(\lambda),$$

where $f_H(\lambda) = |1 - e^{i\lambda}|^{-2H+1}$, $1/2 < H < 1$, and $f_0(\lambda)$ is a spectral density which corresponds to a covariance function (or Fourier coefficient) $\gamma_0(n) = \int_{-\pi}^{\pi} f_0(\lambda)e^{in\lambda}d\lambda$. Assume that the following hold:

- (a) There exists $c_0 > 0$ such that $f_0(\lambda) \geq c_0$ for all $\lambda \in (-\pi, \pi]$;
- (b) $\sum_{n=-\infty}^{\infty} |\gamma_0(n)| < \infty$;
- (c) $\gamma_0(n) = o(n^{-1})$.

Then the condition (9.29) in Assumption A3 holds if $b_n = o(n)$. The result extends to the case where the underlying Gaussian $\{\mathbf{Z}_i\}$ is J -dimensional with independent components.

In Proposition 9.3.1, $f_H(\lambda)$ is the spectral density of a FARIMA(0, d , 0) sequence with $d = H - 1/2$, and $f_0(\lambda)$ is the spectral density of a sequence with short-range dependence.

Under the assumptions in Proposition 9.3.1, the spectral density $f(\lambda)$ cannot have a slowly varying factor which diverges to infinity or converges to zero at $\lambda = 0$, because $f_0(\lambda)$ is bounded away from infinity and zero. For $H \in (1/2, 1)$, the FARIMA(p, d, q) model with $d = H - 1/2$ and the fractional Gaussian noise model satisfy the assumptions of Proposition 9.3.1. See Examples 2.1 and 2.2 of Bai and Taqqu [2015e].

We thus have the following result which we formulate for simplicity in the univariate case $J = 1$.

Corollary 9.3.2. Assume that Assumptions A1-A2 hold with $J = 1$, and the underlying Gaussian $\{Z_i\}$ satisfies the assumptions in Proposition 9.3.1. If $b_n \rightarrow \infty$ and $b_n = o(n)$, then the conclusions of Theorem 9.3.1 and Corollary 9.3.1 hold.

9.3.2 Further analysis of Assumption A3

In this section, we discuss the critical Assumption A3, which involves the covariance structure of the underlying Gaussian $\{\mathbf{Z}_i\}$. In particular, we shall give the general bound (9.49) below for the canonical correlation $\rho_{k,m}$ in (9.28), and discuss how it relates to Assumption A3. As noted in Proposition 9.3.1, however, this bound, in the long memory case, can be improved substantially so as to provide more flexibility on the choice of the block size b_n .

To state this general bound, define

$$M_\gamma(k) = \max_{n>k} \max_{1 \leq j_1, j_2 \leq J} |\gamma_{j_1, j_2}(n)|, \quad (9.47)$$

and

$$\lambda_m = \text{the minimum eigenvalue of } \Sigma_m. \quad (9.48)$$

Note that $\lambda_m > 0$ since Σ_m is assumed to be positive definite.

Lemma 9.3.4. Let $\rho_{k,m}$ be as in (9.28), $M_\gamma(k)$ be as in (9.47) and λ_m be as in (9.48). We have the bound

$$\rho_{k,m} \leq \min \left\{ Jm \frac{M(k-m)}{\lambda_m}, 1 \right\}. \quad (9.49)$$

Proof. Let \mathbf{x} and \mathbf{y} be (column) vectors in \mathbb{R}^{Jm} . Note that each $\mathbf{Z}_1^m = (\mathbf{Z}_1, \dots, \mathbf{Z}_m)$ and $\mathbf{Z}_{k+1}^{k+m} = (\mathbf{Z}_{k+1}, \dots, \mathbf{Z}_{k+m})$ are Jm -dimensional Gaussian vectors translated by k units in the time index. Therefore by (9.27),

$$\begin{aligned} \rho_{k,m} &= \rho(\mathbf{Z}_1^m, \mathbf{Z}_{k+1}^{k+m}) = \sup_{\mathbf{x}, \mathbf{y} \in \mathbb{R}^{Jm}} \frac{\mathbb{E}[\langle \mathbf{x}, \mathbf{Z}_1^m \rangle \langle \mathbf{y}, \mathbf{Z}_{k+1}^{k+m} \rangle]}{\left(\text{Var}[\langle \mathbf{x}, \mathbf{Z}_1^m \rangle]\right)^{1/2} \left(\text{Var}[\langle \mathbf{y}, \mathbf{Z}_{k+1}^{k+m} \rangle]\right)^{1/2}} \\ &= \sup_{\mathbf{x}, \mathbf{y} \in \mathbb{R}^{Jm}} \frac{\mathbf{x}^T \Sigma_{k,m} \mathbf{y}}{\sqrt{\mathbf{x}^T \Sigma_m \mathbf{x}} \sqrt{\mathbf{y}^T \Sigma_m \mathbf{y}}}, \end{aligned} \quad (9.50)$$

where Σ_m is as in (9.25), $\Sigma_{k,m}$ is as in (9.26). By relations 6.58(a) and 6.62(a) in Seber

[2008], one has

$$\rho_{k,m} = \sup_{\mathbf{x}, \mathbf{y} \in \mathbb{R}^{Jm}} \frac{|\mathbf{x}^T \Sigma_{k,m} \mathbf{y}|}{\sqrt{\mathbf{x}^T \Sigma_m \mathbf{x}} \sqrt{\mathbf{y}^T \Sigma_m \mathbf{y}}} \leq \sup_{\mathbf{x}, \mathbf{y} \in \mathbb{R}^{Jm}} \frac{1}{\lambda_m} \frac{|\mathbf{x}^T \Sigma_{k,m} \mathbf{y}|}{\|\mathbf{x}\| \|\mathbf{y}\|} \leq \frac{1}{\lambda_m} \sigma_{k,m}, \quad (9.51)$$

where λ_m is the smallest eigenvalue of Σ_m , and $\sigma_{k,m}$ is the maximum singular value⁵ of $\Sigma_{k,m}$. By Seber [2008] 4.66(b) and 4.67(b), $\sigma_{k,m}$ is bounded by the linear size of the matrix $\Sigma_{k,m}$ times the maximum absolute value of all the elements of the matrix. Since the matrix $\Sigma_{k,m}$ has linear size Jm , we have

$$\begin{aligned} \sigma_{k,m} &\leq Jm \max_{1 \leq i_1, i_2 \leq m} \max_{1 \leq j_1, j_2 \leq J} |\gamma_{j_1, j_2}(i_2 + k - i_1)| \\ &\leq Jm \max_{n > k-m} \max_{1 \leq j_1, j_2 \leq J} |\gamma_{j_1, j_2}(n)| = Jm M_\gamma(k - m). \end{aligned}$$

The bound (9.49) is then obtained by noting that $\rho_{k,m} \leq 1$ in view of (9.30). \square

Example 9.3.2. Consider the important scalar case $J = 1$, where $\mathbf{Z}_i = Z_i$. Denote the covariance function of $\{Z_i\}$ by $\gamma(n)$ and its spectral density by $f(\omega)$. In this case, it is known that Σ_m is non-singular for any m if $\lim_{n \rightarrow \infty} \gamma(n) = 0$ (see Proposition 5.1.1 of Brockwell and Davis [1991]), and that the minimum eigenvalue λ_m satisfies

$$\lambda_m \geq 2\pi \operatorname{ess\,inf}_\omega f(\omega), \quad \text{and} \quad \lim_{m \rightarrow \infty} \lambda_m = 2\pi \operatorname{ess\,inf}_\omega f(\omega), \quad (9.52)$$

where “ess inf” denotes the essential infimum with respect to Lebesgue measure on $[-\pi, \pi)$ (see Grenander and Szegö [1958], Chapter 5.2). If $J = 1$, $M_\gamma(k)$ also reduces to

$$M_\gamma(k) = \max_{n > k} |\gamma(n)|. \quad (9.53)$$

Remark 9.3.3. Consider the vector case but suppose that $\{Z_{i,1}\}, \dots, \dots, \{Z_{i,J}\}$ are mu-

⁵Note that $\Sigma_{k,m}$ is not a symmetric matrix. The square of its singular values are the eigenvalues of $\Sigma_{k,m}^T \Sigma_{k,m}$, which is symmetric and non-negative definite.

tually independent, i.e., $\gamma_{j_1, j_2}(n) = \gamma_{j_1, j_2}(n)\mathbf{I}\{j_1 = j_2\}$. Let

$$\Gamma_{m, j} = (\gamma_{j, j}(i_1 - i_2))_{1 \leq i_1, i_2 \leq m}.$$

In this case, we have a block-diagonal $\Sigma_m = \text{diag}(\Gamma_{m, 1}, \dots, \Gamma_{m, J})$. Let

$$\Gamma_{k, m, j} = (\gamma_{j, j}(i_2 + k - i_1))_{1 \leq i_1, i_2 \leq m}.$$

We also have a block-diagonal $\Sigma_{k, m} = \text{diag}(\Gamma_{k, m, 1}, \dots, \Gamma_{k, m, J})$. Let $\rho_{k, m, j}$ be the between-block canonical correlation $\rho(\mathbf{Z}_{1, j}^m, \mathbf{Z}_{k, j}^m)$ in component j , $j = 1, \dots, J$. The block-diagonal structure implies that

$$\rho_{k, m} = \max\{\rho_{k, m, j}, j = 1, \dots, J\}.$$

Proposition 9.3.2. Assumption A3 holds if $b_n = o(n)$ and

$$\sum_{k=0}^n \min \left\{ \frac{b_n}{\lambda_{b_n+l}} M_\gamma(k), 1 \right\} = o(n). \quad (9.54)$$

Proof. In view of Lemma 9.3.4, we have

$$\sum_{k=0}^n \rho_{k, b_n+l} \leq (b_n + l) + \sum_{k=b_n+l}^n \min \left\{ J b_n \frac{M(k - b_n - l)}{\lambda_{b_n+l}}, 1 \right\} = o(n)$$

since $b_n = o(n)$. Hence Assumption A3 holds. \square

Implications of Proposition 9.3.2.

We discuss here the implications of Condition (9.54) in various specific situations. This discussion is restricted to the case $J = 1$ which is of most interest. This discussion can be easily extended to the case of independent components via the observation made in Remark 4.6.15. Let $c, C > 0$ be generic constants whose value can change from expression to expression. The notation $a \asymp b$ means $cb \leq a \leq Cb$ for some $0 < c < C$. Assume throughout that the covariance $\gamma(n) \rightarrow 0$ and $b_n = o(n)$ as $n \rightarrow \infty$. We distinguish two

cases: $\text{ess inf}_\omega f(\omega) > 0$ and $\text{ess inf}_\omega f(\omega) = 0$.

1. Assume first $\text{ess inf}_\omega f(\omega) > 0$.

In view of (9.52), the minimum eigenvalue λ_m is bounded below away from zero, and hence Condition (9.54) holds if

$$b_n \sum_{k=0}^n M_\gamma(k) = o(n), \quad (9.55)$$

where $M_\gamma(k)$ is expressed as (9.53). Consider the case $\sum_{k=0}^{\infty} M_\gamma(k) < \infty$, which implies the typical *short-range dependence* condition: $\sum_{k=1}^{\infty} |\gamma(k)| < \sum_{k=0}^{\infty} M_\gamma(k) < \infty$. Then (9.55) reduces to $b_n = o(n)$. We get in particular:

Corollary 9.3.3. Suppose that $\text{ess inf}_\omega f(\omega) > 0$, and $|\gamma(n)| \leq d_n$, where d_n is non-increasing and summable (typically, $d_n = cn^{-\beta}$ for some constant $c > 0$ and $\beta > 1$). If $b_n = o(n)$, then Assumption A3 holds.

Proof. $|\gamma(k)| \leq d_k$ implies $M_\gamma(k) \leq d_k$, and hence $\sum_{k=0}^{\infty} M_\gamma(k) < \infty$. □

Consider now the situation relevant to *long-range dependence*:

$$\gamma(k) = k^{2H-2}L(k), \quad 1/2 < H < 1, \quad (9.56)$$

where $L(k)$ is a slowly varying function at infinity. By Theorem 1.5.3 of Bingham et al. [1989], Condition (9.56) implies that $M_\gamma(k) \sim k^{2H-2}L(k)$, which entails that $\sum_{k=0}^n M_\gamma(k) \leq cn^{2H-1}L(n)$. Thus (9.55) holds if

$$b_n = o(n^{2-2H}L(n)^{-1}). \quad (9.57)$$

So, the larger H , the smaller the block size b_n .

Corollary 9.3.4. Suppose that $\text{ess inf}_\omega f(\omega) > 0$, and $|\gamma(n)| \leq n^{2H-2}L(n)$, where $1/2 < H < 1$ and L is slowly varying. If $b_n = o(n^{2-2H}L(n)^{-1})$, then Assumption A3 holds.

The case $|\gamma(k)| \leq k^{2H-2}L(k)$ also encompasses the seasonal long memory situations (see, e.g., Haye and Viano [2003]), where $\gamma(k)$ oscillates within a power-law envelope.

In the long-range dependent case, Betken and Wendler [2015] obtained recently a bound for $\rho_{k,m}$ in (9.28) using a result of Adenstedt [1974] under some additional assumptions. Their bound allows (9.29) to hold under the block size condition

$$b_n = o(n^{3/2-H-\epsilon}) \quad (9.58)$$

with arbitrarily small $\epsilon > 0$. The condition (9.58) is better than (9.57) for each H , and $b_n = O(n^{1/2})$ is always allowed.

We have also seen that if the model satisfies the assumptions of Proposition 9.3.1, one can choose

$$b_n = o(n),$$

irrespective of the value of $H \in (1/2, 1)$.

2. Assume now $\text{ess inf}_\omega f(\omega) = 0$.

As mentioned in (9.52), the smallest covariance eigenvalue $\lambda_m \rightarrow \text{ess inf}_\omega f(\omega) = 0$ as $m \rightarrow \infty$. The rate of convergence has been investigated by a number of authors. See, e.g., Kac et al. [1953], Pourahmadi [1988], Serra [1998], Tilli [2003] and Novosel'tsev and Simonenko [2005]. It involves the order of the zeros of $f(\omega)$. We say $f(\omega)$ has a *zero of order* $\nu > 0$ at $\omega = \omega_0$ if $f(\omega) \asymp |\omega - \omega_0|^\nu$. Roughly speaking, the rate at which λ_m converges to zero follows the highest order of the zeros of $f(\omega)$, and the rate of convergence to zero cannot be faster than exponential:

$$\lambda_m \geq e^{-cm} \quad (9.59)$$

for some $c > 0$ (see Pourahmadi [1988] and Tilli [2003]). Let us focus on the situation where $f(\omega)$ has a finite number of zeros of polynomial orders. Specifically, suppose that $f(\omega)$ has zeros of order ν_1, \dots, ν_p at p distinct points $\omega_1, \dots, \omega_p$, and $f(\omega)$ stays positive

outside arbitrary neighborhoods of $\omega_1, \dots, \omega_p$. Then by Theorem 2.2 of Novosel'tsev and Simonenko [2005], one has $\lambda_m \asymp m^{-\nu}$ where

$$\nu = \max(\nu_1, \dots, \nu_p).$$

Therefore,

$$\lambda_{b_n+l} \asymp (b_n + l)^{-\nu} \asymp b_n^{-\nu}$$

and since $M_\gamma(k)$ is non-increasing, we have

$$\sum_{k=0}^n \min \left\{ \frac{b_n}{\lambda_{b_n+l}} M_\gamma(k), 1 \right\} \leq \sum_{k=0}^{p_n} 1 + C b_n^{1+\nu} \sum_{k=p_n+1}^n M_\gamma(k) \leq C (p_n + n b_n^{1+\nu} M_\gamma(p_n)). \quad (9.60)$$

To satisfy (9.54), we need the last expression in (9.60) to be of order $o(n)$. This will be so if as $n \rightarrow \infty$, $p_n = o(n)$, and

$$b_n = o \left([M_\gamma(p_n)]^{-1/(1+\nu)} \right). \quad (9.61)$$

To get the weakest restriction on b_n , let in addition p_n grow fast enough so that $n/p_n = o(n^\delta)$ for any $\delta > 0$ (e.g., choose $n/p_n \asymp \log n$). We have the following two typical cases:

- $M_\gamma(k) = O(e^{-k})$ decays exponentially. In this case, $[M_\gamma(p_n)]^{-1/(1+\nu)} = O(e^{p_n/(1+\nu)})$, so the condition (9.61) is certainly satisfied when $b_n = o(n)$. Hence Assumption A3 holds with $b_n = o(n)$;
- $M_\gamma(k) = O(k^{-\beta})$, $\beta > 0$. In this case, (9.54) holds when

$$b_n = o(n^{\beta/(1+\nu)-\epsilon}) \quad (9.62)$$

for arbitrarily small $\epsilon > 0$. So the worst case is when β is close to 0 and ν is large.

A nice example involving both ν and β is when $Z(n)$ is *anti-persistent* (also called negative memory), e.g., the fractional Gaussian noise (the increments of fractional

Brownian motion) with $H < 1/2$, and FARIMA(p, d, q) with $d = H - 1/2$ so that $-1/2 < d < 0$. In this case, we have $\beta = 2 - 2H$ and $\nu = 1 - 2H$ in (9.62), and hence (9.54) holds with $b_n = o(n^{1-\epsilon})$. Therefore:

Corollary 9.3.5. Suppose that $\{Z_n\}$ is fractional Gaussian noise with $H < 1/2$ or FARIMA(p, d, q) with $-1/2 < d < 0$. If $b_n = o(n^{1-\epsilon})$ for $\epsilon > 0$ arbitrarily small, then Assumption A3 holds.

Remark 9.3.4. We also mention that in Zhang et al. [2013] which studies non-self-normalized block sampling for sample mean, the condition $b_n = o(n^{1-\epsilon})$ for arbitrarily small $\epsilon > 0$ is shown to suffice for consistency. The framework in their paper assumes $\{X_i\}$ to be a univariate nonlinear transform of linear *non-Gaussian* processes. But it is not clear how to adapt their proof to a setting involving the self-normalization considered here.

9.3.3 Strong mixing case

Given a stationary process $\{X_i\}$, let \mathcal{F}_a^b be the σ -field generated by X_a, \dots, X_b , where $-\infty \leq a \leq b \leq +\infty$. Recall that the strong mixing (or α -mixing) coefficient is defined as

$$\alpha(k) = \sup \{ |P(A)P(B) - P(A \cap B)|, A \in \mathcal{F}_{-\infty}^0, B \in \mathcal{F}_k^\infty \}. \quad (9.63)$$

Note that $0 \leq \alpha(k) \leq 1$. The process $\{X_i\}$ is said to be *strong mixing* if

$$\lim_{k \rightarrow +\infty} \alpha(k) = 0.$$

We refer the reader to Bradley [2007] for more details. We shall use the following inequality which can be found in Lemma A.0.2 of Politis et al. [1999].

Lemma 9.3.5. If $U \in \mathcal{F}_{-\infty}^0$ and $V \in \mathcal{F}_k^\infty$, and $0 \leq U, V \leq 1$ almost surely, then

$$|\text{Cov}(U, V)| \leq \alpha(k) \leq 1.$$

We shall assume:

- B1.** $\{X_i\}$ is a strong mixing stationary process with mean $\mu = EX_i$.
- B2.** We have the weak convergence in $D[0, 1]$ endowed with M_2 topology of the partial sum:
- $$\left\{ \frac{1}{n^H \ell(n)} (S_{[nt]} - n\mu), 0 \leq t \leq 1 \right\} \Rightarrow \{Y(t), 0 \leq t \leq 1\},$$
- for some nonzero H -sssi process $Y(t)$, where $0 < H < 1$ and $\ell(\cdot)$ is a slowly varying function.
- B3.** The block size $b_n \rightarrow \infty$ and $b_n = o(n)$ as $n \rightarrow \infty$.

The following theorem establishes the consistency of the self-normalized block sampling under the strong mixing framework.

Theorem 9.3.2. The conclusions of Theorem 9.3.1 and of Corollary 9.3.1 hold under Assumptions B1–B3.

Proof. The structure of the proof and many details are similar to those of Theorem 9.3.1. We only highlight the key differences. See also Politis et al. [1999] or Sherman and Carlstein [1996].

In Step 1, we again need to show (9.33). The term $[P(T_{i,b_n}^* \leq x) - P(T \leq x)]^2 \rightarrow 0$ as before. We need to establish $\text{Var}[\widehat{F}_{n,b_n}^*(x)] \rightarrow 0$. We still have the bound (9.34).

In view of Lemma 9.3.5, one has that,

$$|\text{Cov}[\mathbf{I}\{T_{1,b_n}^* \leq x\}, \mathbf{I}\{T_{k+1,b_n}^* \leq x\}]| \leq \begin{cases} 1 & \text{if } k < b_n, \\ \alpha(k - b_n + 1), & \text{if } k \geq b_n; \end{cases}$$

where $\alpha(\cdot)$ is the mixing coefficient in (9.63). Hence from (9.34), we have

$$\begin{aligned} \text{Var}[\widehat{F}_{n,b_n}^*(x)] &\leq \frac{2}{n - b_n + 1} \left(\sum_{k=0}^{b_n-1} |\text{Cov}[\mathbf{I}\{T_{1,b_n}^* \leq x\}, \mathbf{I}\{T_{k+1,b_n}^* \leq x\}]| \right. \\ &\quad \left. + \sum_{k=b_n}^n |\text{Cov}[\mathbf{I}\{T_{1,b_n}^* \leq x\}, \mathbf{I}\{T_{k+1,b_n}^* \leq x\}]| \right) \end{aligned}$$

$$\begin{aligned}
&\leq \frac{2}{(n - b_n + 1)} \left[b_n + \sum_{k=b_n}^n \alpha(k - b_n + 1) \right] \\
&= \frac{2b_n}{(n - b_n + 1)} + \frac{2}{(n - b_n + 1)} \sum_{k=1}^{n-b_n+1} \alpha(k), \tag{9.64}
\end{aligned}$$

which converges to zero as $n \rightarrow \infty$, because $b_n = o(n)$ by Assumption B3, and $\alpha(k) \rightarrow 0$ as $k \rightarrow \infty$ by Assumption B1 and by applying a Cesàro summation. Hence (9.33) is proved.

Step 2 and 3 proceed exactly as the proof of Theorem 9.3.1. The argument in the proof of Corollary 9.3.1 shows that the conclusion of that corollary continues to hold under Assumptions B1–B3. \square

Remark 9.3.5. In view of Shao [2010], the self-normalized block sampling method considered in this chapter may be extended to more general statistics beyond the sample mean. There are two aspects to consider, self-normalization and block sampling. For the *self-normalization* aspect to work, the general statistics needs to be approximately linear, namely, it admits a functional Taylor expansion in the sense of (2) in Shao [2010]. In this case, Assumption A2 or B2 needs to be replaced by a modified version of Assumption 1 of Shao [2010]. Furthermore, the remainder term in the aforementioned functional Taylor expansion has to satisfy a negligibility condition (see Assumption 2 of Shao [2010] or Assumption II of Shao [2015]). Validating these conditions for particular statistics (e.g., sample quantiles) and particular models (e.g., the Gaussian subordination model in Assumption A1) may be considered in future work. The *block sampling* aspect is likely to continue to be valid, since as shown in the proofs of Theorem 9.3.1 and 9.3.2, the key is to have a bound on the between-block correlation, as the one in Proposition 9.3.1 in the long-memory Gaussian subordination framework, or as in Lemma 9.3.5 in the strong mixing framework.

9.4 Examples

The first two examples of models concern Assumptions A1–A3. They both involve a phase transition.

Example 9.4.1. Suppose that

$$X_i = G(Z_i) = Z_i^2,$$

where $\{Z_i\}$ is a standardized stationary Gaussian process with covariance $\gamma(n) = n^{2d-1}L(n)$, with $d \in (0, 1/2)$, and $L(n)$ is a positive slowly varying function. Then Assumption A1 is satisfied. Moreover, by Taqqu [1975] in the case $d < 1/4$ and Breuer and Major [1983] and Chambers and Slud [1989] in the case $d > 1/4$, Assumption A2 holds with the following dichotomy:

$$\begin{cases} H = 1/2, \ell(n) = 1, Y(t) = \sigma B(t) & \text{if } d < 1/4; \\ H = 2d, \ell(n) = L(n), Y(t) = c_H Z_{2,H}(t) & \text{if } d > 1/4, \end{cases}$$

where $\sigma^2 = \sum_n \text{Cov}[X(n), X(0)]$, c_H is a positive constant, $B(t)$ is the standard Brownian motion and $Z_{2,H}$ is the standard Rosenblatt process (second-order Hermite process). Assume in addition that the assumptions for $\{Z_i\}$ in Proposition 9.3.1 hold. Then one can choose a block size $b_n = o(n)$ to satisfy Assumption A3. Hence Theorem 9.3.1 and Corollary 9.3.1 hold. Without the additional assumptions in Proposition 9.3.1, Assumption A3 is guaranteed at least by the choice $b_n = o(n^{1-2d}L(n)^{-1})$ in view of (9.57).

Example 9.4.2. Let F_α be the cdf of t_α distribution with $1 < \alpha < 2$, so that it has finite mean but infinite variance. Let Φ be the cdf of a standard normal. Suppose that

$$X_i = F_\alpha^{-1}(\Phi(Z_i)),$$

where $\{Z_i\}$ is a standardized stationary Gaussian process with covariance $\gamma(n) = n^{2d-1}L(n)$, $d \in (0, 1/2)$, and $L(n)$ is a positive slowly varying function. The marginal distribution of $\{X_i\}$ is a t_α . Then Assumption A1 is satisfied. By Sly and Heyde (2008), Assumption A2

holds with the following dichotomy (for $0 < d < 1/2$, $1 < \alpha < 2$):

$$\begin{cases} H = 1/\alpha, \ell(n) = 1, Y(t) = c_1 L_\alpha(t) & \text{if } d + 1/2 < 1/\alpha; \\ H = d + 1/2, \ell(n) = L(n), Y(t) = c_2 B_H(t) & \text{if } d + 1/2 > 1/\alpha, \end{cases}$$

where c_1 and c_2 are positive constants, $L_\alpha(t)$ is a symmetric α -stable Lévy process, and $B_H(t)$ is a standard fractional Brownian motion. Assume in addition that the assumptions for $\{Z_i\}$ in Proposition 9.3.1 hold. This will be the case if $\{Z_i\}$ is fractional Gaussian noise or FARIMA(p, d, q). Then $b_n = o(n)$ implies (9.29). Hence Theorem 9.3.1 and Corollary 9.3.1 hold. Without the additional assumptions in Proposition 9.3.1, Assumption A3 is guaranteed at least by the choice $b_n = o(n^{1-2d}L(n)^{-1})$ in view of (9.57).

Example 9.4.3. Consider the following long-memory stochastic duration (LMSD) model (for modeling inter-trade duration, see Deo et al. [2010]):

$$X_i = \xi_i \exp(Z_i),$$

where $\{\xi_i\}$ are i.i.d. positive random variables satisfying $P(\xi_i > x) \sim Ax^{-\alpha}$ as $x \rightarrow \infty$, $A > 0$, $\alpha \in (1, 2)$, Z_i is a Gaussian linear process $Z_i = \sum_{j=1}^{\infty} j^{d-1} l(j) \epsilon_{i-j}$ with $d \in (0, 1/2)$, $l(j)$ a positive and slowly varying function, $\{\epsilon_i\}$ i.i.d. centered Gaussian, and $\{\epsilon_i\}$ is independent of $\{\xi_i\}$. Note that $\mu = EX_i > 0$. The model has the interesting feature that although $EX_i^2 = \infty$, it has the following finite covariance for $h \neq 0$, namely,

$$\text{Cov}[X_i, X_{i+h}] = \text{Cov}[\exp(Z_0), \exp(Z_h)] \mu_\xi^2 \sim ch^{2d-1} l^2(h),$$

as $h \rightarrow \infty$, where $\mu_\xi = E\xi_i$, and we have used the fact that the exponential function has Hermite rank 1 (see Taqqu [1975]). To satisfy Assumption A1, one can rewrite the model as

$$X_i = g(Z'_i) \exp(Z_i),$$

where $\{Z'_i\}$ are i.i.d. standard Gaussian with g chosen such that $g(Z'_i)$ is equal in distribution to ξ_i . This makes the model satisfy Assumption A1 with $J = 2$, $l = 0$, $\mathbf{Z}_i = (Z'_i, Z_i)$ and $G(x_1, x_2) = g(x_1) \exp(x_2)$. By (4.100) and (4.101) of Beran et al. [2013], Assumption A2 holds with the following dichotomy:

$$\begin{cases} H = 1/\alpha, \ell(n) = 1, Y(t) = c_\alpha L_{\alpha,1,1}(t) & \text{if } d + 1/2 < 1/\alpha; \\ H = d + 1/2, \ell(n) = l^2(n), Y(t) = c_d B_H(t) & \text{if } d + 1/2 > 1/\alpha, \end{cases}$$

where c_α, c_d are positive constants, $L_{\alpha,1,1}(t)$ is an α -stable Lévy process with skewness $\beta = 1$ (see (9.10)), and $B_H(t)$ is the standard fractional Brownian motion. If in addition, the assumptions for $\{Z_i\}$ in Proposition 9.3.1 hold, then Assumption A3 is satisfied if $b_n = o(n)$. Hence Theorem 9.3.1 and Corollary 9.3.1 hold. Without the additional assumptions in Proposition 9.3.1, Assumption A3 is at least satisfied if $b_n = o(n^{1-2d}l(n)^{-2})$ (see (9.57) and Remark 4.6.15).

Remark 9.4.4. Consider the non-centered stochastic volatility model $X_i = \sigma_i g(Z_i) + \mu$ in Jach et al. [2012], where σ_i and $g(Z_i)$ are independent, σ_i is i.i.d. with heavy tails and $\{Z_i\}$ is Gaussian with long-range dependence and g has Hermite rank one. This model can be similarly embedded into Assumption A1. However, as far as we know, the functional convergence⁶ needed in Assumption A2 has not been established (only the marginal convergence was established in Jach et al. [2012]). Assumption A2 for this model is, nevertheless, expected to hold in view of its similarity⁷ to the model treated in Kulik and Soulier [2012], Theorem 4.1 (see also Theorem 4.19 of Beran et al. [2013]). Checking Assumption A2 in details is outside the scope of the current chapter. Assumption A3 is dealt with as in Example 9.4.3.

⁶The weak convergence assumed in Assumption A2 allowed us to take advantage of Lemma 9.3.1 in order to establish Lemma 9.3.2.

⁷Both Jach et al. [2012] and Kulik and Soulier [2012] treated stochastic volatility models of the form $X_i = L_i H_i$ (for limit theorems it does not matter whether a level is added or not), where L_i has finite variance and is long-range dependent, while H_i has infinite variance and is i.i.d.. The difference between the two papers is that in Jach et al. [2012] L_i is centered and H_i is not, while in Kulik and Soulier [2012] H_i is centered and L_i is not.

Nevertheless, the consistency of the self-normalized block sampling in Jach et al. [2012] can be shown to hold under our A1 and A3 framework. This is done by adopting the normalization of Jach et al. [2012], with A2 replaced by marginal convergence involving partial sums and sample covariances⁸, and to ensure A3, by assuming $b_n = o(n)$ and that $\{Z_i\}$ is a long-range dependent sequence satisfying the assumptions of Proposition 9.3.1.

We now give two examples with *strong mixing*. The first involves a nonlinear time series and the second involves heavy tails.

Example 9.4.5. Suppose that

$$X_i = \rho|X_{i-1}| + \epsilon_i, \quad 0 < \rho < 1, \quad (9.65)$$

where ϵ_i 's are i.i.d. standard Gaussian. Thus $\{X_i\}$ follows a threshold autoregressive model (Tong [1990]). The Markov process $\{X_i\}$ is strong mixing because it is ergodic⁹ (see Petrucci and Woolford [1984], Theorem 2.1, or Doukhan [1994] p.103), and hence Condition B1 holds. The conditions of Theorem 3(ii) of Wu [2005] are satisfied¹⁰ and therefore Condition B2 holds with $H = 1/2$, $\ell(n) = 1$ and $Y(t) = \sigma B(t)$, where $\sigma^2 = \sum_n \gamma(n) > 0$ and $B(t)$ is standard Brownian motion. Condition B3 holds for any block size $b_n = o(n)$. Therefore, Theorem 9.3.2 holds.

In the following example, both Assumptions A1–A3 and B1–B3 hold.

Example 9.4.6. Consider the MA(1) model

$$X_i = \epsilon_i + a\epsilon_{i-1},$$

where $a \geq 0$ and $\{\epsilon_i\}$ are i.i.d.. Assume that $E\epsilon_i = 0$, $E\epsilon_i^2 = \infty$, and ϵ_i is in the domain of attraction of a stable distribution with an index $\alpha \in (1, 2)$. Let $b_n = o(n)$. By choosing

⁸More precisely, convergence in distribution of a 3-dimensional vector specified in Theorem 3 of Jach et al. [2012].

⁹that is, the Markov chain is irreducible aperiodic and positive recurrent (see Tweedie [1975]).

¹⁰In the terminology of Wu [2005], $R(x, \epsilon) = \rho|x| + \epsilon$, $L_\epsilon = \rho$, $\delta_p(n) = O(n^r)$ for some $0 < r < 1$, so that $\sum_{n=0}^{\infty} n\delta_p(n) < \infty$, implying Theorem 3(ii).

appropriate transforms, we can express ϵ_i as function of Gaussian. Therefore Assumption A1 holds. Assumption B1 holds because $\{X_i\}$ is 2-dependent. By Theorem 2' of Avram and Taqqu [1992], Assumptions A2 and B2 hold with $H = 1/\alpha$, some slowly varying function $\ell(n)$, and $Y(t)$ is an α -stable Lévy process. Also A3 holds with any $b_n = o(n)$ since $\rho_{k,m} = 0$ when $k \geq m + 2$. Therefore, both assumptions A1–A3 and B1–B3 hold in this case.

9.5 Monte Carlo Simulations

We shall carry out here Monte Carlo simulations to examine the finite-sample performance of the self-normalized block sampling (SNBS) method and make a comparison with the recent result of Zhang et al. [2013]. Instead of resorting to self-normalization, the method of Zhang et al. [2013] exploits the regularly varying property of the asymptotic variance to avoid the problem of estimating the nuisance Hurst index. We first consider the case with Gaussian subordination. For this, let

$$X_i = K(Z_i), \quad Z_i = \sum_{j=0}^{\infty} a_j \epsilon_{i-j}, \quad i = 1, \dots, n, \quad (9.66)$$

where $K(\cdot)$ is a possibly nonlinear transformation and $\{\epsilon_k\}$ are i.i.d. standard normal random variables¹¹. We consider the following configurations for (9.66):

- (a) $K(x) = x$ and $a_j = (1 + j)^{d-1}$, $j \geq 0$;
- (b) $K(x) = x^2$ and $a_j = (1 + j)^{d-1}$, $j \geq 0$;
- (c) $K(x) = \Phi_t^{-1}[\Phi_N\{(\sum_{j=0}^{\infty} a_j^2)^{-1/2}x\}]$ and $a_j = (1 + j)^{d-1}$, $j \geq 0$,

where Φ_N is the CDF of the standard normal and Φ_t is the CDF of the Student's t -distribution with degree of freedom 1.5, whose tail probability decays like $|x|^{-3/2}$ as $|x| \rightarrow \infty$ so that it has infinite variance but finite mean.

¹¹To generate the process, we use the approximation $Z_i \approx \sum_{j=0}^{\lfloor n^{3/2} \rfloor - 1} a_j \epsilon_{i-j}$ in our simulation, and the fast Fourier transform (FFT) as mentioned in Wu et al. [2011] is implemented to facilitate the computation. Note that the cutoff $n^{3/2}$ is much greater than the sample size n .

Case (a) represents the Gaussian linear process which has been extensively used in the literature for modeling time series data. It has long-range dependence if $0 < d < 1/2$. We let $d \in \{0.25, -1\}$. The choice $d = 0.25$ corresponds to long-range dependence (LRD) and the choice $d = -1$ corresponds to short-range dependence (SRD).

Case (b) involves an additional nonlinear transformation and now $\{X_i\}$ is LRD if $0.25 < d < 0.5$. We let $d \in \{0.4, 0.2, -1\}$. When $d = 0.4$, both $\{Z_i\}$ and $\{X_i\}$ have LRD (the limit for $\{X_i\}$ is the Rosenblatt process); when $d = 0.2$, $\{Z_i\}$ has LRD and $\{X_i\}$ has SRD (the limit for $\{X_i\}$ is Brownian motion); when $d = -1$, both $\{Z_i\}$ and $\{X_i\}$ have SRD (the limit for $\{X_i\}$ is Brownian motion). See for example Wu [2006] and Zhang et al. [2013].

Case (c) corresponds to a process $\{X_i\}$ with marginal distribution t with 1.5 degrees of freedom and hence with infinite variance. We let $d \in \{0.4, 0.2, -1\}$. When $d = 0.4$ and $d = 0.2$, both $\{Z_i\}$ and $\{X_i\}$ have LRD (the limit for $\{X_i\}$ is the fractional Brownian motion); when $d = -1$, both $\{Z_i\}$ and $\{X_i\}$ have SRD (the limit for $\{X_i\}$ is symmetric (3/2)-stable Lévy motion). See Sly and Heyde [2008] for the boundary between SRD and LRD in the heavy tail case. We also consider the situation with a non-constant slowly varying function, where we let $a_j = (1 + j)^{d-1} \log(1 + j)$, $j \geq 0$, and denote the corresponding cases by (a*), (b*) and (c*), respectively.

We consider the problem of constructing the lower and upper one-sided confidence interval where the nominal level is taken as 90%; see also Nordman and Lahiri [2005] and Zhang et al. [2013] for similar performance assessment of this type. Following Zhang et al. [2013], we use throughout the block sizes $b_n = \lfloor cn^{0.5} \rfloor$, $c \in \{0.5, 1, 2\}$. This does not necessarily represent the optimal choice of b_n , but provides us with a spectrum of reasonable block sizes in our finite-sample simulations. For each realization we compute the self-normalized block sums and its empirical distribution function \hat{F}_{n,b_n} as in (9.17). Examples of realized \hat{F}_{n,b_n} can be found in Figure 9.1 for models (a)–(c) with different choices of d . Let q_α ($\alpha=10\%$) be the 10%-quantile of \hat{F}_{n,b_n} , then the lower 90% one-sided

confidence interval can be constructed as

$$\left(-\infty, \bar{X}_n - n^{-1} \left\{ n^{-1} \sum_{k=1}^n (S_{1,k} - \frac{k}{n} S_{1,n})^2 \right\}^{1/2} q_\alpha \right];$$

Similarly, if $q_{1-\alpha}$ ($1 - \alpha = 90\%$) denotes the 90%-quantile of \hat{F}_{n,b_n} , then the corresponding upper 90% one-sided confidence interval is

$$\left[\bar{X}_n - n^{-1} \left\{ n^{-1} \sum_{k=1}^n (S_{1,k} - \frac{k}{n} S_{1,n})^2 \right\}^{1/2} q_{1-\alpha}, +\infty \right).$$

See (9.18) for details.

In Tables 9.1 and 9.2, we report the empirical coverage probabilities of the constructed confidence intervals based on 5000 realizations for each scenario¹². For example, Table 9.1 displays the following results of simulation. If $d = 0.25$, $c = 0.5$ and $n = 100$, then the self-normalized block sampling (SNBS) simulation yielded the following: the lower 90% confidence interval included the unknown mean μ , 88.3% of the times and the upper 90% confidence interval included the unknown mean μ , 91.1% of the times. We also report the results of the subsampling method of Zhang et al. [2013] for a comparison in the column ZHWW2013. Note that the method of Zhang et al. [2013] does not take advantage of the technique of self-normalization and therefore it requires an additional bandwidth to utilize the regularly varying property of the asymptotic variance.¹³

It can be seen from Tables 9.1 and 9.2 that the method proposed in this chapter performs reasonably well, as most of the empirical coverage probabilities are reasonably close to their nominal level of 90%, except for situations with heavy tails where deviations under small sample sizes are expected. However, the results seem to improve as the sample size increases from $n = 100$ to $n = 500$ and the performance is comparable to the method

¹²When evaluating the empirical coverage probability of the constructed confidence interval, we use the averaged mean of 1000 realizations as an approximation to the true mean.

¹³In Tables 9.1 and 9.2, we let the second bandwidth be $l_n = \lfloor n^{0.9} \rfloor$ when using the method of Zhang et al. [2013]. Many other choices are possible. We also used $l_n = \lfloor 0.5n^{0.9} \rfloor$ and obtained similar results.

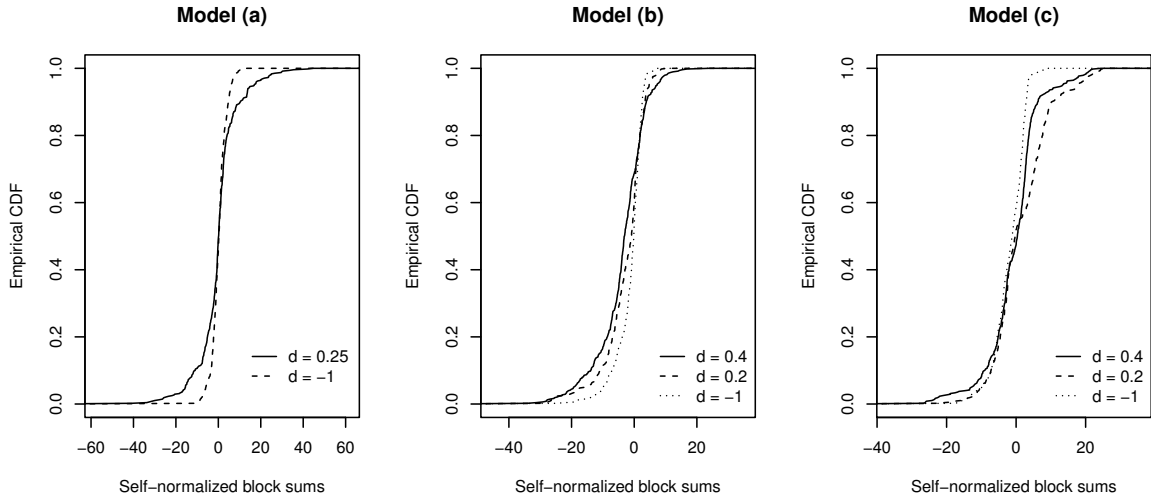


Figure 9.1: Examples of realized \hat{F}_{n,b_n} for models (a)–(c) with $n = 500$, $c = 1$ and different choices of d . The x -axis represents the self-normalized block sums, which have been appropriately centered and scaled.

of Zhang et al. [2013]¹⁴. Note that the choice of sample size $n = 100$ is considered to be challengingly small for inference of long-range dependent processes. Because of self-normalization, our method has the advantage over the one by Zhang et al. [2013] in not requiring the choice of a second bandwidth.

Finally, consider the strong mixing Example 9.4.5, where $X_i = \rho|X_{i-1}| + \epsilon_i$, following the threshold autoregressive model [Tong, 1990]. The ϵ_i 's are i.i.d. Gaussian. The results for $\rho = 0.5$ are summarized in Table 9.3. Observe that the method works quite well in this case as well.

¹⁴The theoretical assumptions in Zhang et al. [2013] do not allow for infinite variance.

d	c	$n = 100$		$n = 500$	
		SNBS	ZHWW2013	SNBS	ZHWW2013
<i>Model (a)</i>					
0.25	0.5	(88.3, 91.1)	(86.8, 90.3)	(92.2, 92.0)	(92.0, 91.5)
	1	(86.1, 86.6)	(85.7, 85.3)	(89.6, 91.2)	(89.3, 91.3)
	2	(82.3, 83.7)	(81.0, 82.2)	(87.5, 87.5)	(87.4, 87.2)
-1	0.5	(93.5, 94.2)	(93.0, 92.9)	(93.2, 93.1)	(92.9, 93.0)
	1	(89.5, 90.7)	(89.0, 90.2)	(91.4, 92.1)	(91.1, 91.7)
	2	(87.1, 86.3)	(86.9, 85.6)	(90.0, 89.0)	(89.9, 89.5)
<i>Model (b)</i>					
0.4	0.5	(90.3, 95.7)	(89.2, 95.2)	(93.2, 96.2)	(92.9, 95.6)
	1	(84.7, 93.6)	(83.8, 92.7)	(88.2, 94.8)	(88.4, 94.9)
	2	(75.9, 91.8)	(75.3, 91.4)	(84.3, 92.8)	(84.0, 92.9)
0.2	0.5	(94.6, 95.8)	(94.0, 94.8)	(95.7, 96.0)	(95.8, 95.6)
	1	(88.8, 93.6)	(88.2, 93.3)	(93.8, 93.6)	(93.7, 93.9)
	2	(81.4, 91.5)	(80.3, 90.8)	(89.4, 92.0)	(89.3, 91.9)
-1	0.5	(97.6, 86.3)	(97.5, 85.5)	(97.0, 86.0)	(97.0, 86.1)
	1	(94.1, 84.2)	(93.5, 83.3)	(94.5, 86.5)	(94.3, 86.5)
	2	(87.2, 84.0)	(86.7, 83.6)	(91.3, 86.6)	(91.2, 86.7)
<i>Model (c)</i>					
0.4	0.5	(74.8, 84.4)	(72.5, 82.9)	(82.2, 78.0)	(81.8, 77.1)
	1	(78.0, 76.9)	(76.5, 75.8)	(77.7, 79.3)	(76.9, 78.9)
	2	(75.5, 73.4)	(74.8, 72.2)	(74.6, 78.6)	(73.8, 78.4)
0.2	0.5	(78.8, 81.4)	(76.7, 79.0)	(80.8, 79.9)	(80.0, 79.6)
	1	(77.0, 80.6)	(75.9, 79.6)	(79.1, 80.8)	(78.7, 80.0)
	2	(77.9, 74.8)	(76.6, 74.1)	(81.1, 77.3)	(80.9, 76.3)
-1	0.5	(82.3, 83.7)	(80.9, 82.2)	(83.6, 85.3)	(83.3, 84.2)
	1	(84.1, 80.0)	(83.2, 79.4)	(81.6, 86.0)	(80.6, 85.5)
	2	(87.4, 71.2)	(86.2, 70.3)	(82.0, 82.9)	(81.7, 82.8)

Table 9.1: Empirical coverage probabilities of lower and upper (paired in parentheses) one-sided 90% confidence intervals with different combinations of the index d , sample size n and block size $b_n = \lfloor cn^{0.5} \rfloor$ when $a_j = (1 + j)^{d-1}$, $j \geq 0$.

d	c	n = 100		n = 500	
		SNBS	ZHWW2013	SNBS	ZHWW2013
<i>Model (a*)</i>					
0.25	0.5	(87.8, 87.9)	(86.4, 86.5)	(91.9, 91.6)	(92.0, 91.6)
	1	(84.0, 84.2)	(82.7, 83.0)	(90.4, 89.4)	(90.1, 88.9)
	2	(78.0, 79.2)	(76.8, 78.6)	(84.7, 85.1)	(84.4, 84.8)
-1	0.5	(93.7, 93.6)	(93.1, 92.5)	(94.0, 94.4)	(93.6, 94.4)
	1	(90.9, 89.8)	(90.1, 88.6)	(93.2, 91.9)	(92.8, 92.0)
	2	(86.4, 86.4)	(85.9, 85.4)	(90.6, 90.3)	(90.2, 90.1)
<i>Model (b*)</i>					
0.4	0.5	(84.7, 95.1)	(83.3, 94.3)	(90.3, 98.0)	(90.4, 97.7)
	1	(80.4, 92.3)	(79.4, 91.8)	(86.2, 96.0)	(86.4, 96.0)
	2	(71.7, 90.2)	(70.5, 89.6)	(79.5, 93.6)	(79.6, 93.9)
0.2	0.5	(89.3, 96.8)	(88.7, 96.3)	(94.3, 97.7)	(94.4, 97.5)
	1	(83.6, 93.9)	(83.0, 93.3)	(90.8, 96.7)	(91.1, 96.7)
	2	(77.7, 91.2)	(76.8, 90.4)	(85.1, 95.6)	(85.1, 95.2)
-1	0.5	(98.3, 86.3)	(97.9, 85.6)	(97.1, 87.0)	(97.1, 87.1)
	1	(93.1, 85.2)	(92.8, 84.6)	(95.1, 85.6)	(94.7, 85.7)
	2	(88.6, 84.1)	(87.9, 83.5)	(92.2, 85.9)	(92.1, 85.5)
<i>Model (c*)</i>					
0.4	0.5	(86.3, 85.8)	(84.6, 83.7)	(92.6, 88.0)	(92.3, 88.0)
	1	(83.4, 77.7)	(82.3, 76.1)	(87.3, 83.3)	(87.7, 83.5)
	2	(74.9, 75.2)	(73.2, 74.0)	(81.1, 81.7)	(81.1, 81.1)
0.2	0.5	(83.0, 85.2)	(80.4, 83.3)	(86.6, 84.6)	(86.8, 84.9)
	1	(80.4, 80.5)	(79.5, 78.7)	(84.4, 81.7)	(84.5, 80.7)
	2	(77.9, 73.5)	(76.9, 72.9)	(80.4, 78.8)	(80.9, 78.3)
-1	0.5	(83.9, 83.1)	(82.3, 81.7)	(88.5, 84.0)	(87.5, 83.0)
	1	(80.6, 83.1)	(80.0, 81.7)	(86.8, 83.4)	(85.9, 82.8)
	2	(83.2, 76.7)	(82.2, 75.8)	(85.8, 82.3)	(85.4, 81.5)

Table 9.2: Empirical coverage probabilities of lower and upper (paired in parentheses) one-sided 90% confidence intervals with different combinations of the index d , sample size n and block size $b_n = \lfloor cn^{0.5} \rfloor$ when $a_j = (1 + j)^{d-1} \log(1 + j)$, $j \geq 0$.

c	n = 100		n = 500	
	SNBS	ZHWW2013	SNBS	ZHWW2013
0.5	(92.1, 94.3)	(91.7, 93.7)	(93.2, 89.6)	(93.0, 89.4)
1	(90.0, 88.9)	(88.8, 88.8)	(91.0, 88.0)	(91.3, 88.5)
2	(86.9, 84.7)	(86.1, 84.0)	(89.9, 87.2)	(90.1, 87.3)

Table 9.3: Empirical coverage probabilities of lower and upper (paired in parentheses) one-sided 90% confidence intervals with the TAR model (9.65) for different combinations of sample size n and block size $b_n = \lfloor cn^{0.5} \rfloor$.

Bibliography

- R.K. Adenstedt. On large-sample estimation for the mean of a stationary random sequence. *The Annals of Statistics*, pages 1095–1107, 1974.
- R.J. Adler, R. Feldman, and Murad S Taqqu, editors. *A Practical Guide to Heavy Tails: Statistical Techniques and Applications*. Birkhäuser, 1998.
- M.A. Arcones. Limit theorems for nonlinear functionals of a stationary Gaussian sequence of vectors. *The Annals of Probability*, pages 2242–2274, 1994.
- B. Arras. A white noise approach to stochastic integration with respect to the Rosenblatt process. *arXiv preprint arXiv:1308.1835*, 2013.
- B. Arras, E. Azmoodeh, G. Poly, and Y. Swan. Stein’s method on the second wiener chaos: 2-wasserstein distance. *arXiv preprint arXiv:1601.03301*, 2016.
- F. Avram. On bilinear forms in Gaussian random variables and Toeplitz matrices. *Probability Theory and Related Fields*, 79(1):37–45, 1988.
- F. Avram and M.S. Taqqu. Noncentral limit theorems and Appell polynomials. *The Annals of Probability*, 15(2):767–775, 1987.
- F. Avram and Murad S Taqqu. Weak convergence of sums of moving averages in the α -stable domain of attraction. *The Annals of Probability*, pages 483–503, 1992.
- F. Avram, N. Leonenko, and L. Sakhno. On a Szegő type limit theorem, the Hölder-Young-Brascamp-Lieb inequality, and the asymptotic theory of integrals and quadratic forms of stationary fields. *ESAIM: Probability and Statistics*, 14:210–255, 2010.
- S. Bai and M.S. Taqqu. Multivariate limit theorems in the context of longrange dependence. *Journal of Time Series Analysis*, 34(6):717–743, 2013a.
- S. Bai and M.S. Taqqu. Multivariate limits of multilinear polynomial-form processes with long memory. *Statistics & Probability Letters*, 83(11):2473–2485, 2013b.
- S. Bai and M.S. Taqqu. Generalized Hermite processes, discrete chaos and limit theorems. *Stochastic Processes and Their Applications*, 124(4):1710–1739, 2014a.
- S. Bai and M.S. Taqqu. Structure of the third moment of the generalized Rosenblatt distribution. *Statistics & Probability Letters*, 94:144–152, 2014b.
- S. Bai and M.S. Taqqu. The universality of homogeneous polynomial forms and critical limits. *To appear in Journal of Theoretical Probability*, 2015a.

- S. Bai and M.S. Taqqu. Convergence of long-memory discrete k-th order Volterra processes. *Stochastic Processes and Their Applications*, 125(5):2026–2053, 2015b.
- S. Bai and M.S. Taqqu. The impact of diagonals of polynomial forms on limit theorems with long memory. *arXiv:1403.7544*. To appear in *Bernoulli*, 2015c.
- S. Bai and M.S. Taqqu. Behavior of the generalized Rosenblatt process at extreme critical exponent values. *The Annals of Probability*. To appear., 2015d.
- S. Bai and Murad S Taqqu. On the validity of resampling methods under long memory. *arXiv Preprint arXiv:1512.00819*, 2015e.
- S. Bai, M.S. Ginovyan, and M.S. Taqqu. Functional limit theorems for Toeplitz quadratic functionals of continuous time Gaussian stationary processes. *Statistics and Probability Letters*, 104, 2015.
- S. Bai, M.S. Ginovyan, and M.S. Taqqu. Limit theorems for quadratic forms of lévy-driven continuous-time linear processes. *Stochastic Processes and their Applications*, 126(4): 1036–1065, 2016a.
- S. Bai, M.S. Taqqu, and T. Zhang. A unified approach to self-normalized block sampling. *Stochastic Processes and Their Applications*. To appear, 2016b. arXiv:1512.00820.
- R.T. Baillie. Long memory processes and fractional integration in econometrics. *Journal of econometrics*, 73(1):5–59, 1996.
- J.M. Bardet and C.A. Tudor. A wavelet analysis of the Rosenblatt process: chaos expansion and estimation of the self-similarity parameter. *Stochastic Processes and their Applications*, 120(12):2331–2362, 2010.
- B. Basrak, D. Krizmanic, and J. Segers. A functional limit theorem for dependent sequences with infinite variance stable limits. *The Annals of Probability*, 40(5):2008–2033, 09 2012.
- R.F. Bass. *Stochastic processes*. Cambridge University Press, 2011.
- J. Beran. *Statistics for Long-Memory Processes*. Chapman and Hall/CRC, 1994.
- J. Beran, Y. Feng, S. Ghosh, and R. Kulik. *Long-Memory Processes: Probabilistic Properties and Statistical Methods*. Springer, 2013.
- A. Betken and M. Wendler. Subsampling for general statistics under long range dependence. *arXiv preprint arXiv:1509.05720*, 2015.
- H. Biermé, A. Bonami, I. Nourdin, and G.i Peccati. Optimal berry-esseen rates on the wiener space: the barrier of third and fourth cumulants. *Latin American Journal of Probability and Mathematical Statistics*, 9(2), 2012.
- P. Billingsley. The invariance principle for dependent random variables. *Transactions of the American Mathematical Society*, 83(1):250–268, 1956.

- P. Billingsley. *Convergence of Probability Measures*. Wiley series in probability and statistics. Wiley, 1999.
- N.H. Bingham, C.M. Goldie, and J.L. Teugels. *Regular Variation*. Encyclopedia of Mathematics and Its Applications. Cambridge University Press, 1989.
- V.I. Bogachev. *Measure Theory*, volume 1. Springer-Verlag, 2007.
- T. Bojdecki, L.G. Gorostiza, and A. Talarczyk. From intersection local time to the Rosenblatt process. *Journal of Theoretical Probability*, pages 1–23, 2013.
- R.C. Bradley. *Introduction to Strong Mixing Conditions*. Kendrick Press, Utah, 2007.
- J.C. Breton and I. Nourdin. Error bounds on the non-normal approximation of hermite power variations of fractional brownian motion. *Electronic Communications in Probability*, 13:482–493, 2008.
- P. Breuer and P. Major. Central limit theorems for non-linear functionals of Gaussian fields. *Journal of Multivariate Analysis*, 13(3):425–441, 1983.
- P.J. Brockwell. Lévy-driven CARMA processes. *Annals of the Institute of Statistical Mathematics*, 53(1):113–124, 2001.
- P.J. Brockwell. Recent results in the theory and applications of CARMA processes. *Annals of the Institute of Statistical Mathematics*, 66(4):647–685, 2014.
- P.J. Brockwell and R.A. Davis. *Time Series: Theory and Methods*. Springer, 1991.
- D. Chambers and E. Slud. Central limit theorems for nonlinear functionals of stationary Gaussian processes. *Probability Theory and Related Fields*, 80(3):323–346, 1989.
- S. Cohen and A. Lindner. A central limit theorem for the sample autocorrelations of a Lévy driven continuous time moving average process. *Journal of Statistical Planning and Inference*, 143(8):1295–1306, 2013.
- Y.A. Davydov. The invariance principle for stationary processes. *Theory of Probability and Its Applications*, 15(3):487–498, 1970.
- H. Dehling, A. Rooch, and M.S. Taqqu. Non-parametric change-point tests for long-range dependent data. *Scandinavian Journal of Statistics*, 40(1):153–173, 2013.
- R. Deo, M. Hsieh, and C.M. Hurvich. Long memory in intertrade durations, counts and realized volatility of NYSE stocks. *Journal of Statistical Planning and Inference*, 140(12):3715–3733, 2010.
- S.W. Dharmadhikari and K. Joag-Dev. Upper bounds for the variances of certain random variables. *Communications in statistics-theory and methods*, 18(9):3235–3247, 1989.
- Z. Ding, C.W.J. Granger, and R.F. Engle. A long memory property of stock market returns and a new model. *Journal of empirical finance*, 1(1):83–106, 1993.

- R.L. Dobrushin. Gaussian and their subordinated self-similar random generalized fields. *The Annals of Probability*, 7(1):1–28, 1979.
- R.L. Dobrushin and P. Major. Non-central limit theorems for non-linear functional of Gaussian fields. *Probability Theory and Related Fields*, 50(1):27–52, 1979.
- J.L. Doob. *Stochastic processes*, volume 101. New York Wiley, 1953.
- P. Doukhan. *Mixing*. Springer, 1994.
- P. Doukhan, G. Oppenheim, and M.S. Taqqu, editors. *Theory and Applications of Long-Range Dependence*. Birkhäuser, 2003.
- P. Eichelsbacher and C. Thäle. Malliavin-Stein method for Variance-Gamma approximation on Wiener space. *arXiv preprint arXiv:1409.5646*, 2014.
- P. Embrechts and M. Maejima. *Selfsimilar Processes*. Princeton University Press, 2002.
- Z. Fan. *Statistical Issues and Developments in Time Series Analysis and Educational Measurement*. Ph.D. thesis, University of Illinois at Urbana-Champaign, 2010.
- M. Farré, M. Jolis, and F. Utzet. Multiple Stratonovich integral and Hu–Meyer formula for Lévy processes. *The Annals of Probability*, 38(6):2136–2169, 2010.
- J.M. Flegal and G.L. Jones. Batch means and spectral variance estimators in Markov chain Monte Carlo. *The Annals of Statistics*, 38:1034–1070, 2010.
- R. Fox and M.S. Taqqu. Noncentral limit theorems for quadratic forms in random variables having long-range dependence. *The Annals of Probability*, pages 428–446, 1985.
- R. Fox and M.S. Taqqu. Large-sample properties of parameter estimates for strongly dependent stationary Gaussian time series. *The Annals of Statistics*, pages 517–532, 1986.
- R. Fox and M.S. Taqqu. Central limit theorems for quadratic forms in random variables having long-range dependence. *Probability Theory and Related Fields*, 74(2):213–240, 1987.
- R.E. Gaunt. Variance-Gamma approximation via Steins method. *Electronic Journal of Probability*, 19(38):1–33, 2014.
- M.S. Ginovian. On Toeplitz type quadratic functionals of stationary Gaussian processes. *Probability Theory and Related Fields*, 100(3):395–406, 1994.
- M.S. Ginovyan. Efficient estimation of spectral functionals for continuous-time stationary models. *Acta Applicandae Mathematicae*, 115(2):233–254, 2011.
- M.S. Ginovyan and A.A. Sahakyan. On the central limit theorem for Toeplitz quadratic forms of stationary sequences. *Theory of Probability & Its Applications*, 49(4):612–628, 2005.

- M.S. Ginovyan and A.A. Sahakyan. Limit theorems for Toeplitz quadratic functionals of continuous-time stationary processes. *Probability Theory and Related Fields*, 138(3-4): 551–579, 2007.
- M.S. Ginovyan, A.A. Sahakyan, and M.S. Taqqu. The trace problem for toeplitz matrices and operators and its impact in probability. *Probability Surveys*, 11:393–440, 2014.
- L. Giraitis and D. Surgailis. A central limit theorem for quadratic forms in strongly dependent linear variables and its application to asymptotical normality of Whittle’s estimate. *Probability Theory and Related Fields*, 86(1):87–104, 1990.
- L. Giraitis and M.S. Taqqu. Central limit theorems for quadratic forms with time-domain conditions. *Annals of Probability*, pages 377–398, 1998.
- L. Giraitis and M.S. Taqqu. Convergence of normalized quadratic forms. *Journal of statistical planning and inference*, 80(1):15–35, 1999.
- L. Giraitis and M.S. Taqqu. Functional non-central and central limit theorems for bivariate Appell polynomials. *Journal of Theoretical Probability*, 14(2):393–426, 2001.
- L. Giraitis, H.L. Koul, and D. Surgailis. *Large Sample Inference for Long Memory Processes*. World Scientific Publishing Company Incorporated, 2012.
- I.S. Gradshteyn and I.M. Ryzhik. *Table of Integrals, Series, and Products*. Table of Integrals, Series, and Products Series. Elsevier, 2007. ISBN 9780080471112.
- U. Grenander and G. Szegő. *Toeplitz forms and their applications*. University of California Press, 1958.
- A. Gut. *Probability: A Graduate Course*. Springer Science & Business Media, 2006.
- P. Hall, B-Y Jing, and S.N. Lahiri. On the sampling window method for long-range dependent data. *Statistica Sinica*, 8(4):1189–1204, 1998.
- M.O. Haye and M.C. Viano. Limit theorems under seasonal long-memory. *Theory and Applications of Long-Range Dependence*, pages 102–110, 2003.
- N. Herrndorf. A functional central limit theorem for weakly dependent sequences of random variables. *The Annals of Probability*, pages 141–153, 1984.
- H. Ho and T. Hsing. Limit theorems for functionals of moving averages. *The Annals of Probability*, 25(4):1636–1669, 1997.
- H. Ho and T. Sun. A central limit theorem for non-instantaneous filters of a stationary gaussian process. *Journal of Multivariate Analysis*, 22(1):144–155, 1987.
- H.C. Ho and T.C. Sun. Limiting distributions of nonlinear vector functions of stationary Gaussian processes. *The Annals of Probability*, pages 1159–1173, 1990.
- W. Hoeffding and H. Robbins. The central limit theorem for dependent random variables. *Duke Mathematical Journal*, 15(3):773–780, 1948.

- Y. Huang, S. Volgushev, and X. Shao. On self-normalization for censored dependent data. *Journal of Time Series Analysis*, 36:109–124, January 2015.
- I.A. Ibragimov. On estimation of the spectral function of a stationary Gaussian process. *Theory of Probability & Its Applications*, 8(4):366–401, 1963.
- I.A. Ibragimov and Y.V. Linnik. *Independent and stationary sequences of random variables*. Wolters-Noordhoff Publishing Groningen, 1971.
- K. Itô. Multiple Wiener integral. *Journal of the Mathematical Society of Japan*, 3(1):157–169, 1951.
- A. Jach, T. McElroy, and D.N. Politis. Subsampling inference for the mean of heavy-tailed long-memory time series. *Journal of Time Series Analysis*, 33(1):96–111, 2012.
- S. Janson. *Gaussian Hilbert Spaces*, volume 129 of *Cambridge Tracts in Mathematics*. Cambridge University Press, 1997.
- A. Jeffrey and D. Zwillinger. *Table of Integrals, Series, and Products*. Table of Integrals, Series, and Products Series. Elsevier Science, 2007. ISBN 9780080471112.
- M. Kac, W.L. Murdock, and G. Szego. On the eigen-values of certain hermitian forms. *Journal of Rational Mechanics and Analysis*, 2(6):767–802, 1953.
- O. Kallenberg. On an independence criterion for multiple wiener integrals. *The Annals of Probability*, 19(2):483–485, 1991.
- O. Kallenberg. *Foundations of Modern Probability*. Springer Science & Business Media, 2006.
- A.N. Kolmogorov and Y.A. Rozanov. On strong mixing conditions for stationary Gaussian processes. *Theory of Probability & Its Applications*, 5(2):204–208, 1960.
- W. Krakowiak and J. Szulga. Random multilinear forms. *The Annals of Probability*, 14(3):955–973, 1986.
- R. Kulik and P. Soulier. Limit theorems for long-memory stochastic volatility models with infinite variance: partial sums and sample covariances. *Advances in Applied Probability*, 44(4):1113–1141, 2012.
- S.N. Lahiri. *Resampling methods for dependent data*. Springer Science & Business Media, 2003.
- J. Lamperti. Semi-stable stochastic processes. *Transactions of the American Mathematical Society*, 104(1):62–78, 1962.
- W.E. Leland, Murad S Taqqu, W. Willinger, and D.V. Wilson. On the self-similar nature of ethernet traffic (extended version). *IEEE/ACM Transactions on Networking*, 2(1):1–15, 1994.

- C. Lévy-Leduc, H. Boistard, E. Moulines, M.S. Taqqu, and V.A. Reisen. Asymptotic properties of U-processes under long-range dependence. *The annals of statistics*, 39(3): 1399–1426, 2011.
- M. Lifshits. *Lectures on Gaussian Processes*. SpringerBriefs in Mathematics. Springer, 2012.
- I.N. Lobato. Testing that a dependent process is uncorrelated. *Journal of the American Statistical Association*, 96(455):1066–1076, 2001.
- M. Maejima and C.A. Tudor. Wiener integrals with respect to the Hermite process and a non-central limit theorem. *Stochastic Analysis and Applications*, 25(5):1043–1056, 2007.
- M. Maejima and C.A. Tudor. Selfsimilar processes with stationary increments in the second Wiener chaos. *Probability and Mathematical Statistics*, 32(1):167–186, 2012.
- M. Maejima and C.A. Tudor. On the distribution of the Rosenblatt process. *Statistics & Probability Letters*, 83(6):1490–1495, 2013.
- P. Major. Limit theorems for non-linear functionals of Gaussian sequences. *Zeitschrift für Wahrscheinlichkeitstheorie und verwandte Gebiete*, 57(1):129–158, 1981.
- P. Major. *Multiple Wiener-Itô Integrals: With Applications to Limit Theorems*. Lecture Notes in Mathematics. Springer, 2 edition, 2014.
- B.B. Mandelbrot and J.W. Van Ness. Fractional Brownian motions, fractional noises and applications. *SIAM review*, 10(4):422–437, 1968.
- B.B. Mandelbrot and J.R. Wallis. Noah, Joseph, and operational hydrology. *Water resources research*, 4(5):909–918, 1968.
- M. Maxwell and M. Woodroffe. Central limit theorems for additive functionals of Markov chains. *The Annals of Probability*, 28(2):713–724, 2000.
- T McElroy and D.N. Politis. Robust inference for the mean in the presence of serial correlation and heavy-tailed distributions. *Econometric Theory*, 18:1019–1039, 2002.
- T. McElroy and D.N. Politis. Self-normalization for heavy-tailed time series with long memory. *Statistica Sinica*, 17:199–220, 2007.
- T. McElroy and D.N. Politis. Distribution theory for the studentized mean for long, short, and negative memory time series. *Journal of Econometrics*, 177:60–74, 2013.
- T. Mori and H. Oodaira. The law of the iterated logarithm for self-similar processes represented by multiple Wiener integrals. *Probability theory and related fields*, 71(3): 367–391, 1986.
- J.P. Nolan. *Stable Distributions - Models for Heavy Tailed Data*. Birkhauser, Boston, 2015. In progress, Chapter 1 online at academic2.american.edu/~jpnolan.

- D.J. Nordman and S.N. Lahiri. Validity of the sampling window method for long-range dependent linear processes. *Econometric Theory*, 21(06):1087–1111, 2005.
- I. Nourdin. *Selected aspects of fractional Brownian motion*. Springer, 2012.
- I. Nourdin and G. Peccati. *Normal Approximations With Malliavin Calculus: From Stein's Method to Universality*. Cambridge Tracts in Mathematics. Cambridge University Press, 2012.
- I. Nourdin and G. Peccati. The optimal fourth moment theorem. *arXiv:1305.1527*, *Proceedings of the American Mathematical Society*, to appear, 2013.
- I. Nourdin and G. Poly. Convergence in law in the second Wiener/Wigner chaos. *Electronic Communications in Probability*, 17:36, 2012.
- I. Nourdin and J. Rosinski. Asymptotic independence of multiple Wiener-Itô integrals and the resulting limit laws. *The Annals of Probability*, 42(2):497–526, 2014.
- I. Nourdin, G. Peccati, and G Reinert. Invariance principles for homogeneous sums: universality of gaussian wiener chaos. *The Annals of Probability*, 38(5):1947–1985, 2010.
- I. Nourdin, D. Nualart, and G. Peccati. Strong asymptotic independence on wiener chaos. *Proceedings of the American Mathematical Society*, 144(2):875–886, 2016.
- A.Y. Novosel'tsev and I. Simonenko. Asymptotic dependence of the extreme eigenvalues of truncated Toeplitz matrices on the rate at which the symbol attains its extremum. *St. Petersburg Mathematical Journal*, 16(4):713–718, 2005.
- D. Nualart. *The Malliavin calculus and related topics*. Springer-Verlag New York, 2006.
- D. Nualart and G. Peccati. Central limit theorems for sequences of multiple stochastic integrals. *The Annals of Probability*, 33(1):177–193, 2005.
- G. Peccati and M.S. Taqqu. *Wiener Chaos: Moments, Cumulants and Diagrams: a Survey With Computer Implementation*. Springer Verlag, 2011.
- G. Peccati and C. Tudor. Gaussian limits for vector-valued multiple stochastic integrals. *Séminaire de Probabilités XXXVIII*, pages 219–245, 2005.
- G. Peccati and C. Zheng. Universal Gaussian fluctuations on the discrete Poisson chaos. *Bernoulli*, 20(2):697–715, 2014.
- M. Peligrad. On the asymptotic normality of sequences of weak dependent random variables. *Journal of Theoretical Probability*, 9:703–715, July 1996.
- L.C. Petersen. On the relation between the multidimensional moment problem and the one-dimensional moment problem. *Mathematica Scandinavica*, 51(2):361–366, 1982.
- J.D. Petrucci and S.W. Woolford. A threshold AR(1) model. *Journal of Applied Probability*, pages 270–286, 1984.

- V. Pipiras and M.S. Taqqu. Regularization and integral representations of Hermite processes. *Statistics and Probability Letters*, 80(23):2014–2023, 2010.
- D.N. Politis. Higher-order accurate, positive semi-definite estimation of large-sample covariance and spectral density matrices. *Econometric Theory*, 27:703–744, August 2011.
- D.N. Politis, J.P. Romano, and M. Wolf. *Subsampling*. Springer Series in Statistics. Springer New York, 1999.
- M. Pourahmadi. Remarks on extreme eigenvalues of Toeplitz matrices. *International Journal of Mathematics and Mathematical Sciences*, 11(1):23–26, 1988.
- Y.V. Prokhorov. Convergence of random processes and limit theorems in probability theory. *Theory of Probability & Its Applications*, 1(2):157–214, 1956.
- S. Resnick. *A Probability Path*. Springer Science & Business Media, 1999.
- S.I. Resnick. *Heavy-Tail Phenomena: Probabilistic and Statistical Modeling*. Springer Science & Business Media, 2007.
- J.P. Romano and M. Wolf. Subsampling inference for the mean in the heavy-tailed case. *Metrika*, 50(1):55–69, 1999.
- A. Rooch. *Change-Point Tests for Long-Range Dependent Data*. PhD thesis, University of Bochum, Germany, 2012.
- M. Rosenblatt. A central limit theorem and a strong mixing condition. *Proceedings of the National Academy of Sciences of the United States of America*, 42(1):43, 1956.
- M. Rosenblatt. Independence and dependence. In *Proc. Fourth Berkeley Symp. Math. Statist. Probab*, volume 2, pages 431–443, 1961.
- M. Rosenblatt. Asymptotic behavior of eigenvalues of Toeplitz forms. *Journal of Mathematics and Mechanics*, 11(6):941–949, 1962.
- M. Rosenblatt. Some limit theorems for partial sums of quadratic forms in stationary Gaussian variables. *Zeitschrift für Wahrscheinlichkeitstheorie und verwandte Gebiete*, 49(2):125–132, 1979.
- V.I. Rotar. Limit theorems for polylinear forms. *Journal of Multivariate analysis*, 9(4): 511–530, 1979.
- S.G. Samko, A.A. Kilbas, and O.I. Marichev. *Fractional Integrals and Derivatives, Theory and Applications*. Gordon and Breach Science Publishers, 1993.
- G. Samorodnitsky and Murad S Taqqu. *Stable Non-Gaussian Random Processes: Stochastic Models with Infinite Variance*, volume 1. CRC Press, 1994.
- M. Schreiber. Fermeture en probabilité de certains sous-espaces d'un espace L^2 . *Zeitschrift für Wahrscheinlichkeitstheorie und Verwandte Gebiete*, 14(1):36–48, 1969.

- G.A.F. Seber. *A Matrix Handbook for Statisticians*, volume 15. John Wiley & Sons, 2008.
- S. Serra. On the extreme eigenvalues of Hermitian (block) Toeplitz matrices. *Linear algebra and its applications*, 270(1):109–129, 1998.
- X. Shao. A self-normalized approach to confidence interval construction in time series. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 72(3):343–366, 2010.
- X. Shao. A simple test of changes in mean in the possible presence of long-range dependence. *Journal of Time Series Analysis*, 32:598–606, November 2011.
- Xiaofeng Shao. Self-normalization for time series: a review of recent developments. *Journal of the American Statistical Association*, page to appear, 2015.
- M. Sherman and E. Carlstein. Replicate histograms. *Journal of the American Statistical Association*, 91(434):566–576, 1996.
- A.V. Skorokhod. Limit theorems for stochastic processes. *Theory of Probability & Its Applications*, 1(3):261–290, 1956.
- A.V. Skorokhod. Limit theorems for stochastic processes with independent increments. *Theory of Probability & Its Applications*, 2(2):138–171, 1957.
- E.V. Slud. The moment problem for polynomial forms in normal random variables. *The Annals of Probability*, pages 2200–2214, 1993.
- A. Sly and C. Heyde. Nonstandard limit theorem for infinite variance functionals. *The Annals of Probability*, 36(2):796–805, 2008.
- P. Soulier. Some applications of regular variation in probability and statistics. *Unpublished Notes*, 2009.
- F. Spangenberg. A central limit theorem for the sample autocovariance of a continuous-time moving average process with long memory. *arXiv preprint arXiv:1502.04851*, 2015.
- D. Surgailis. Zones of attraction of self-similar multiple integrals. *Lithuanian Mathematical Journal*, 22(3):327–340, 1982.
- M. Taniguchi and Y. Kakizawa. *Asymptotic theory of statistical inference for time series*. Springer Science & Business Media, 2012.
- M.S. Taqqu. Weak convergence to fractional Brownian motion and to the Rosenblatt process. *Probability Theory and Related Fields*, 31(4):287–302, 1975.
- M.S. Taqqu. Convergence of integrated processes of arbitrary Hermite rank. *Probability Theory and Related Fields*, 50(1):53–83, 1979.
- N. Terrin and M.S. Taqqu. A noncentral limit theorem for quadratic forms of Gaussian stationary sequences. *Journal of Theoretical Probability*, 3(3):449–475, 1990.

- N. Terrin and M.S. Taqqu. Convergence to a Gaussian limit as the normalization exponent tends to $1/2$. *Statistics & probability letters*, 11(5):419–427, 1991a.
- N. Terrin and M.S. Taqqu. Power counting theorem in Euclidean space. In *Random Walks, Brownian Motion, and Interacting Particle Systems*, pages 425–440. Springer Verlag, 1991b.
- P. Tilli. Universal bounds on the convergence rate of extreme Toeplitz eigenvalues. *Linear algebra and its applications*, 366:403–416, 2003.
- H. Tong. *Non-linear Time Series: A Dynamic System Approach*. Oxford University Press, Oxford, 1990.
- C.A. Tudor. Analysis of the Rosenblatt process. *ESAIM: Probability and Statistics*, 12(1): 230–257, 2008.
- C.A. Tudor and F.G. Viens. Variations and estimators for self-similarity parameters via malliavin calculus. *The Annals of Probability*, 37(6):2093–2134, 2009.
- R.L. Tweedie. Sufficient conditions for ergodicity and recurrence of Markov chains on a general state space. *Stochastic Processes and their Applications*, 3(4):385–403, 1975.
- M. Tyran-Kamińska. Convergence to Lévy stable processes under some weak dependence conditions. *Stochastic Processes and their Applications*, 120(9):1629–1650, 2010a.
- M. Tyran-Kamińska. Functional limit theorems for linear processes in the domain of attraction of stable laws. *Statistics & probability letters*, 80(11):975–981, 2010b.
- A.S. Ustunel and M. Zakai. On independence and conditioning on Wiener space. *The Annals of Probability*, 17(4):1441–1453, 1989.
- M.S. Veillette and M.S. Taqqu. Properties and numerical evaluation of the Rosenblatt distribution. *Bernoulli*, 19(3):982–1005, 2013.
- W. Vervaat. Sample path properties of self-similar processes with stationary increments. *The Annals of Probability*, pages 1–27, 1985.
- K.N. Whitney and D.W. Kenneth. A simple, positive semi-definite, heteroskedasticity and autocorrelation consistent covariance matrix. *Econometrica*, 55(3):703–708, May 1987.
- W.B. Wu. Nonlinear system theory: Another look at dependence. *Proceedings of the National Academy of Sciences of the United States of America*, 102(40):14150–14154, 2005.
- W.B. Wu. Unit root testing for functionals of linear processes. *Econometric Theory*, 22(01):1–14, 2006.
- W.B. Wu. Strong invariance principles for dependent random variables. *The Annals of Probability*, 35(6):2294–2320, 2007.

- W.B. Wu. Asymptotic theory for stationary processes. *Statistics and Its Interface*, 4(2): 207–226, 2011.
- W.B. Wu, G. Michailidis, and D. Zhang. Simulating sample paths of linear fractional stable motion. *IEEE Transactions on Information Theory*, 50(6):1086–1096, June 2011.
- T. Zhang and W.B. Wu. Inference of time-varying regression models. *The Annals of Statistics*, 40(3):1376–1402, 2012.
- T. Zhang, H-C Ho, M. Wendler, and W.B. Wu. Block sampling under strong dependence. *Stochastic Processes and their Applications*, 123(6):2323–2339, 2013.
- Z. Zhou and X. Shao. Inference for linear models with dependent errors. *Journal of the Royal Statistical Society: Series B*, 75:323–343, March 2013.

Curriculum Vitae

- Personal Information* Name: Shuyang Bai
Year of Birth: 1988
Phone: +001-857-204-7830
Email: bsy9142@gmail.com
Department of Statistics, University of Georgia
111 Cummington Street, 101 Cedar Street, GA, 30602, USA
- Research Interest* Probability: Stochastic processes, Self-similar processes, Limit theorems
Statistics: Time series, Long-range dependence, Resampling methods
- Education* Boston University, PhD, Mathematics,
September 2011 – May 2016. Advisor: Murad S. Taqqu
Beijing Normal University, BS, Mathematics and Applied Mathematics,
September 2007 – July 2011.
- Publications*
1. Shuyang Bai, Murad S. Taqqu, Ting Zhang “A unified approach to self-normalized block sampling” (2016). (To appear in *Stochastic Processes and their Applications*).
 2. Shuyang Bai, Murad S. Taqqu. “The behavior of the generalized Rosenblatt process at extreme parameter values”. (2015) (To appear in *The Annals of Probability*).
 3. Shuyang Bai, Murad S. Taqqu “Short-range dependent processes subordinated to the Gaussian may not be strong mixing”. *Statistics & Probability Letters* 110 (2016): 198-200.
 4. Shuyang Bai, Mamikon S. Ginovyan, Murad S. Taqqu “Limit theorems for quadratic forms of Lévy-driven continuous-time linear processes”. *Stochastic Processes and their Applications* 126.4 (2016): 1036-1065.
 5. Shuyang Bai, Murad S. Taqqu. “The universality of homogeneous polynomial forms and critical limits”. (2015) (To appear in *Journal of Theoretical Probability*).
 6. Shuyang Bai, and Murad S. Taqqu. “The impact of diagonals of polynomial forms on limit theorems with long memory”. arXiv:1403.7544 (2014) (To appear in *Bernoulli*).
 7. Shuyang Bai, Mamikon S. Ginovyan, Murad S. Taqqu. “Functional limit theorems for Toeplitz quadratic functionals of continuous-time Gaussian stationary processes”. *Statistics & Probability Letters* 104 (2015): 58-67.

8. Shuyang Bai and Murad S. Taqqu. “Convergence of long-memory discrete k-th order Volterra processes”. *Stochastic Processes and their Applications* 125.5 (2015): 2026-2053.
9. Shuyang Bai and Murad S. Taqqu. “Structure of the third moment of the generalized Rosenblatt distribution”. *Statistics & Probability Letters* 94 (2014): 2473-2485.
10. Shuyang Bai and Murad S. Taqqu. “Generalized Hermite processes, discrete chaos and limit theorems”. *Stochastic Processes and their Applications* 124.4 (2014): 144-152.
11. Shuyang Bai and Murad S. Taqqu. “Multivariate limits of multilinear polynomial-form processes with long memory”. *Statistics & Probability Letters* 83.11 (2013): 2473-2485.
12. Shuyang Bai and Murad S. Taqqu. “Multivariate limit theorems in the context of long-range dependence”. *Journal of Time Series Analysis* 34.6 (2013) 717-743.

*Honors and
Awards*

1. Itô Travel Award, International Mathematical Union, 2015.
2. Financial support, CRM-PIMS Summer School, 2015.
3. Travel award, Conference in Stochastic Analysis and Related Topics, Purdue University, 2015.
4. Travel award, Institute for Mathematics, 2015.
5. Travel award, Banff International Research Station, 2014.
6. Travel award, 7th International Conference on Lévy Processes: Theory and Applications and its satellite summer school, 2013.
7. Dean’s Fellowship, Boston University, 2011.
8. Outstanding Undergraduate Thesis award, Beijing Normal University, 2011.
9. 2010 Meritorious Prize in the US Mathematical Contest in Modeling (MCM).
10. 2007-2010 University Scholarship, Beijing Normal University.