Limit theorems for Toeplitz-type quadratic functionals of stationary processes and applications[*]

Mamikon S. Ginovyan[†] and Murad S. Taqqu

Department of Mathematics and Statistics
Boston University,
Boston, MA 02215, USA
e-mail: ginovyan@math.bu.edu; murad@math.bu.edu

Abstract: This is a survey of recent results on central and non-central limit theorems for quadratic functionals of stationary processes. The underlying processes are Gaussian, linear or Lévy-driven linear processes with memory, and are defined either in discrete or continuous time. We focus on limit theorems for Toeplitz and tapered Toeplitz type quadratic functionals of stationary processes with applications in parametric and nonparametric statistical estimation theory. We discuss questions concerning Toeplitz matrices and operators, Fejér-type singular integrals, and Lévy-Itô-type and Stratonovich-type multiple stochastic integrals. These are the main tools for obtaining limit theorems.

Contents

1. Introduction

A significant part of large-sample statistical inference relies on limit theorems of probability theory, which involves sums and quadratic functionals of stationary

observations. Depending on the memory (or dependence) structure of the underlying processes, these functionals, once normalized, can have different limits, and the proofs of such limit theorems generally use different methods. In this paper, we focus on the quadratic functionals. The term 'central limit theorem' (CLT) refers to a statement that a suitably standardized quadratic functional converges in distribution to a Gaussian random variable. Limit theorems where a suitably standardized quadratic functional converges in distribution to a non-Gaussian random variable are termed 'non-central limit theorems' (NCLT).

We present results on central and non-central limit theorems for Toeplitz and tapered Toeplitz type quadratic functionals of stationary processes with applications in parametric and nonparametric statistical estimation theory. The underlying processes are Gaussian, linear or Lévy-driven linear processes with memory, and are defined either in discrete or continuous time. We also discuss some questions concerning Toeplitz matrices and operators, Fejér-type singular integrals, Lévy-Itô-type and Stratonovich-type multiple stochastic integrals, and power counting theorems. These are the main tools for obtaining limit theorems, but they are also of interest in themselves.

1.1. Notation and conventions

The following notation and conventions are used throughout the paper.

The symbol ':=' stands for 'by definition'. c.t.: $=$ continuous-time; d.t.: $=$ discrete-time; s.d.:$=$ spectral density; c.f.:$=$ covariance function; CLT:$=$ central limit theorem; NCLT:$=$ non-central limit theorem. The symbol '$\overset{d}{=}$' stands for equality of the finite-dimensional distributions. The symbol '$\overset{d}{\to}$' stands for convergence in distribution. The symbol '$\overset{f.d.d.}{\longrightarrow}$' stands for convergence of finite-dimensional distributions. The symbol '$\Longrightarrow$' stands for weak convergence. The notation $X_T \overset{d}{\to} \eta \sim N(0, \sigma^2)$ as $T \to \infty$ will mean that the distribution of the random variable X_T tends (as $T \to \infty$) to the centered normal distribution with variance σ^2. $\mathbb{E}[\cdot]$: $=$ expectation operator. $\mathrm{tr}[A]$: $=$ trace of an operator (matrix) A. $\mathbb{I}_A(\cdot)$: $=$ indicator of a set $A \subset \Lambda$. $\mathrm{WN}(0,1)$: $=$ standard white noise.

The standard symbols $\mathbb{N}$, $\mathbb{Z}$ and $\mathbb{R}$ denote the sets of natural, integer and real numbers, respectively. By Λ we denote the frequency domain, that is, $\Lambda := \mathbb{R}$ in the c.t. case, and $\Lambda := [-\pi.\pi]$ in the d.t. case. By $L^p(\mu) := L^p(\Lambda, \mu)$ $(p \geq 1)$ we denote the weighted Lebesgue space with respect to the measure μ, and by $|| \cdot ||_{p,\mu}$ we denote the norm in $L^p(\mu)$. In the special case where $d\mu(\lambda) = d\lambda$, we will use L^p and $|| \cdot ||_p$, respectively. By l^p $(p \geq 1)$ we denote the the space of p-summable sequences. The letters C and c with or without indices are used to denote positive constants, the values of which can vary from line to line.

1.2. The functionals under consideration

Let $\{X(u), \ u \in \mathbb{U}\}$ be a centered real-valued stationary process with spectral density (s.d.) $f(\lambda)$, $\lambda \in \Lambda$ and covariance function (c.f.) $r(t)$, $t \in \mathbb{U}$. We consider

simultaneously the continuous-time (c.t.) case, where $\mathbb{U} = \mathbb{R} := (-\infty, \infty)$, and the discrete-time (d.t.) case, where $\mathbb{U} = \mathbb{Z} := \{0, \pm 1, \pm 2, \ldots\}$. The domain Λ of the frequency variable λ is $\Lambda = \mathbb{R}$ in the c.t. case, and $\Lambda := [-\pi.\pi]$ in the d.t. case.

We first survey the recent results concerning the asymptotic distribution (as $T \to \infty$) of the following Toeplitz type quadratic functionals of the process $X(u)$:

$$Q_T := \begin{cases} \sum_{t=1}^T \sum_{s=1}^T \widehat{g}(t-s)X(t)X(s) & \text{in the d.t. case,} \\ \\ \int_0^T \int_0^T \widehat{g}(t-s)X(t)X(s)\,dt\,ds & \text{in the c.t. case,} \end{cases} \tag{1.1}$$

where

$$\widehat{g}(t) := \int_\Lambda e^{i\lambda t} g(\lambda)\,d\lambda, \quad t \in \mathbb{U} \tag{1.2}$$

is the Fourier transform of some real, even, integrable function $g(\lambda)$, $\lambda \in \Lambda$. We will refer to $g(\lambda)$ and to its Fourier transform $\widehat{g}(t)$ as a *generating function* and *generating kernel* for the functional Q_T, respectively. In the d.t. case the functions $f(\lambda)$ and $g(\lambda)$ are assumed to be 2π-periodic and periodically extended to $\mathbb{R}$. In the c.t. case the process $X(u)$ is assumed measurable and mean square continuous.

The limit distributions of the functionals in (1.1) are completely determined by the spectral density $f(\lambda)$ (or covariance function $r(t)$) and the generating function $g(\lambda)$ (or generating kernel $\widehat{g}(t)$), and depending on their properties, the limits can be either Gaussian (that is, Q_T with an appropriate normalization obeys the central limit theorem), or non-Gaussian.

The following two questions arise naturally:

(a) Under what conditions on $f(\lambda)$ (resp. $r(t)$) and $g(\lambda)$ (resp. $\widehat{g}(t)$) will the limits be Gaussian? (CLT-problem).

(b) Describe the limit distributions, if they are non-Gaussian (NCLT-problem).

We discuss here these questions both for d.t. and c.t. stationary processes, and survey the recent results.

We also survey recent results concerning functional central and non-central limit theorems for the following processes, generated by quadratic functionals in (1.1):

$$Q_T(\tau) := \begin{cases} \sum_{t=1}^{[T\tau]} \sum_{s=1}^{[T\tau]} \widehat{g}(t-s)X(t)X(s) & \text{in the d.t. case,} \\ \\ \int_0^{T\tau} \int_0^{T\tau} \widehat{g}(t-s)X(t)X(s)\,dt\,ds & \text{in the c.t. case,} \end{cases} \tag{1.3}$$

where $\tau \in [0, 1]$ and $[\,\cdot\,]$ stands for the greatest integer.

We say that a functional central limit theorem (FCLT) for $Q_T(\tau)$ holds if the process $Q_T(\tau)$ with an appropriate normalization converges weakly in $C[0, 1]$ in

the c.t. case (and in $D[0, 1]$ in the d.t. case) to Brownian motion. We say that we have a functional non-central limit theorem (FNCLT) for the process $Q_T(\tau)$ if the limit is non-Gaussian or, if Gaussian, it is not Brownian motion.

The tapered case. In the spectral analysis of stationary processes the data are frequently tapered before calculating the statistics of interest. Instead of the original data $\{X(t),\ 0 \le t \le T\}$ the tapered data $\{h(t)X(t),\ 0 \le t \le T\}$ with the data taper $h(t)$ are used for all further calculations. The benefits of tapering the data have been widely reported in the literature (see, e.g., Bloomfield [14], Brillinger [17], Dahlhaus [24, 25, 26], Dahlhaus and Künsch [28], Guyon [69], and references therein). For example, data-tapers are introduced to reduce the so-called 'leakage effects', that is, to obtain better estimation of the spectrum of the model in the case where it contains high peaks. Other application of data-tapers is in situations in which some of the data values are missing. Also, the use of tapers leads to bias reduction, which is especially important when dealing with spatial data. In this case, the tapers can be used to fight the so-called 'edge effects'.

In this case, to establish asymptotic properties of the corresponding estimators, we have to study the asymptotic distribution (as $T \to \infty$) of the following Toeplitz type tapered quadratic functionals of the process $X(u)$:

$$
Q_T^h := \begin{cases} \sum_{t=1}^{T} \sum_{s=1}^{T} \widehat{g}(t-s) h_T(t) h_T(s) X(t) X(s) & \text{in the d.t. case,} \\[2mm] \int_0^T \int_0^T \widehat{g}(t-s) h_T(t) h_T(s) X(t) X(s) \, dt \, ds & \text{in the c.t. case,} \end{cases} \tag{1.4}
$$

where $\widehat{g}(t)$, $t \in \mathbb{U}$ is the Fourier transform of some integrable even function $g(\lambda)$, $\lambda \in \Lambda$, and $h_T(t) := h(t/T)$ with a taper function $h(t)$, $t \in [0, 1]$.

Quadratic functionals of the form (1.1) and (1.4) arise naturally in the context of nonparametric and parametric statistical estimation of the spectrum of stationary processes based on the non-tapered and tapered data, respectively. And their limiting distributions are necessary to establish asymptotic properties of the corresponding estimators. For instance, when we are interested in nonparametric estimation of a linear integral functional $J(f)$ in $L^p(\Lambda)$, $p > 1$, then a natural statistical estimator for $J(f)$ is the linear integral functional of the empirical spectral density (periodogram) of the process $X(t)$, which is a quadratic functional of the form (1.1) in the non-tapered case, and of the form (1.4) in the tapered case. In the case of parametric estimation, for example, the Whittle estimation procedure is based on the minimization of quadratic functionals of the form (1.1) and (1.4).

1.3. A brief history

The problem of describing the asymptotic distribution of Toeplitz type quadratic forms and functionals of stationary processes has a long history, and goes back to the classical monograph by Grenander and Szegő [67], where the CLT-problem for Gaussian processes was considered as an application of the authors' theory of

the asymptotic behavior of the trace of products of truncated Toeplitz matrices and operators.

Later the CLT-problem has been studied by a number of authors. Here we mention only some significant contributions. For the d.t. short memory processes, the problem was studied by Bentkus [10], Ibragimov [75] and Rosenblatt [87], in connection with statistical estimation of the spectral and covariance functions. Since 1986, there has been a renewed interest in this problem, related to the statistical inferences for long memory (long-range dependence) and intermediate memory (anti-persistent) processes (see, e.g., Avram [6], Fox and Taqqu [39], Giraitis and Surgailis [60], Giraitis and Taqqu [62], Has'minskii and Ibragimov [70], Ginovian and Sahakian [52], Terrin and Taqqu [102], and references therein). In particular, Avram [6], Fox and Taqqu [39], Ginovian and Sahakian [52], Giraitis and Surgailis [60], Giraitis and Taqqu [62] have obtained sufficient conditions for quadratic form Q_T to obey the central limit theorem.

In the case of c.t. stationary Gaussian processes the CLT-problem for Toeplitz type quadratic functionals was studied in a number of papers. We cite merely the papers Avram et al. [7], Bai et al. [8, 9], Bryc and Dembo [19], Ginovyan [43, 45, 48], Ginovyan and Sahakyan [53], Ibragimov [75], Leonenko and Sakhno [80], where additional references can be found. The NCLT-problem have been studied in Bai et al. [9], Giraitis and Taqqu [65], and Terrin and Taqqu [101].

Central and non-central limit theorems for tapered quadratic forms of a d.t. long memory Gaussian stationary fields have been proved in Doukhan et al. [32]. A central limit theorem for tapered quadratic functionals Q_T^h, in the case where the underlying model $X(t)$ is a Lévy-driven c.t. stationary linear process has been proved in Ginovyan and Sahakyan [56].

The problem of nonparametric and parametric estimation of the spectrum of the process $X(t)$ both for d.t. and c.t. cases based on the non-tapered data has been considered by many authors (see, e.g., Avram et al. [7], Bentkus [10], Dahlhaus [27], Dzhaparidze [27], Fox and Taqqu [38], Gao et al. [41], Ginovyan [42, 43, 46, 47, 48, 49], Giraitis et al. [59], Giraitis and Surgailis [60], Giraitis and Taqqu [64], Guyon [69], Hasminskii and Ibragimov [70], Ibragimov [75, 76], Leonenko and Sakhno [80], Taniguchi [95], Taniguchi and Kakizawa [96], Taqqu [98], and references therein). The problem in the tapered case was studied in Alomari et al. [1], Dahlhaus [24, 26], Dahlhaus and Künsch [28], Ginovyan [50], Ginovyan and Sahakyan [56], and Ludeña and Lavielle [82].

1.4. Frequency-domain conditions

Conditions that are expressed in terms of the spectral density $f(\lambda)$ and the generating function $g(\lambda)$ will be termed frequency-domain conditions, while conditions that are in terms of the covariance function $r(t)$ and the generating kernel $\widehat{g}(t)$ will be termed time-domain conditions.

There are three sets of frequency-domain conditions for functionals of the form (1.1) to obey the CLT, and these conditions separate the ranges of CLT and NCLT:

   (a) the (L^p, L^q) condition,

   (b) the (α, β) condition,

   (c) the trace condition.

All three are compensation conditions, meaning that the good behavior of one function, say g can compensate for the bad behavior of f and vice versa.

(a) The (L^p, L^q) condition. Let $f \in L^p$ ($p \geq 2$) and $g \in L^q$ ($q \geq 2$). If $1/p + 1/q \leq 1/2$, then the functional Q_T with an appropriate normalization obeys the CLT (see Theorem 3.1(C)), while when $1/p + 1/q > 1/2$, then, in general, Q_T does not obey the CLT. This condition goes back to the classical works of Rosenblatt [86, 87], where when estimating the covariance function $r(t)$ of a d.t. Gaussian process, it was shown that for a sample covariance $\widehat{r}(t)$ (which is a functional of the form (1.1) with $g(\lambda) = \cos(t\lambda)$) to obey the CLT, the spectral density f should satisfy the condition $f \in L^p$ ($p \geq 2$).

(b) The (α, β) condition. If both the spectral density f 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 distribution of Q_T. When $\alpha + \beta \leq 1/2$, then the limiting distribution of Q_T is Gaussian, that is, Q_T with an appropriate normalization obeys the CLT (see Theorem 3.3), while when $\alpha + \beta > 1/2$, then the limiting distribution of Q_T is non-Gaussian (see Theorem 4.6).

(c) The trace condition. This condition, which is more general and implies both conditions (a) and (b), is an implicit condition. It is expressed in terms of traces of products of truncated Toeplitz matrices (in the d.t. case) and operators (in the c.t. case). The idea here is to approximate the traces of products of Toeplitz matrices and operators (which are no longer Toeplitz) by the traces of a Toeplitz matrix and a Toeplitz operator, respectively. Let $A_T(f)$ denote either the $T \times T$ Toeplitz matrix $B_T(f)$ or the T-truncated Toeplitz operator $W_T(f)$ generated by the spectral density f, and let $A_T(g)$ denote either the $T \times T$ Toeplitz matrix $B_T(g)$, or the T-truncated Toeplitz operator $W_T(g)$ generated by the function g (for definitions see Section 3.1, formulas (3.4) and (3.5)). The trace condition is: if

$$fg \in L^2$$

and

$$T^{-1}\big[\mathrm{tr}\big(A_T(f)A_T(g)\big)^2 - \mathrm{tr}\big(A_T(f^2g^2)\big)\big] \to 0 \quad \text{as} \quad T \to \infty,$$

then the quadratic functional Q_T in (1.1) with an appropriate normalization obeys the CLT (see Theorems 3.1(A) and 3.4).

   We will also discuss the time-domain counterparts of (L^p, L^q) and (α, β) conditions.

1.5. Methods and tools

The most commonly used methods to prove central limit theorems for Toeplitz type quadratic forms and functionals are:

(a) the method of cumulants or moments,
(b) the approximation method,
(c) the method of characteristic functions.

To prove the non-central limit theorems for Toeplitz type quadratic functionals, was used:

(a) the spectral representation of the underlying process,
(b) the properties of Lévy-Itô-type and Stratonovich-type multiple stochastic integrals,
(c) the power counting theorem.

Some details of the above methods are described in Section 8.

1.6. The structure of the paper

The paper is structured as follows. In Section 2 we describe the model of interest - a stationary process, and recall some key notions and results from the theory of stationary processes. In Section 3 we present sufficient conditions for Toeplitz type quadratic forms and functionals of the form (1.1) to obey the CLT in the case where the model is either a Gaussian or a linear process. Section 4 contains functional central and noncentral limit theorems for processes generated by Toeplitz type quadratic forms and functionals for Gaussian and linear models. Section 5 is devoted to the functional central and noncentral limit theorems for Lévy-driven linear models. In Section 6 we discuss the case of tapered Toeplitz quadratic functionals, and state central limit theorems. Section 7 contains some applications, involving nonparametric estimation of spectral functionals and Whittle parametric estimation procedure. In Section 8 we briefly discuss the methods and tools, used to prove central and noncentral limit theorems for Toeplitz type quadratic forms and functionals.

2. The model: second-order stationary process

In this section we introduce the model of interest - a second-order stationary process, and recall some key notions and results from the theory of stationary processes.

2.1. Key notions and some basic results

2.1.1. Second-order (wide-sense) stationary process

Let $\{X(u),\ u \in \mathbb{U}\}$ be a centered real-valued second-order (wide-sense) stationary process defined on a probability space $(\Omega, \mathcal{F}, P)$ with covariance function $r(t)$, that is,

$$\mathbb{E}[X(u)] = 0, \quad r(u) = \mathbb{E}[X(t+u)X(t)], \quad u, t \in \mathbb{U},$$

where $\mathbb{E}[\cdot]$ stands for the expectation operator with respect to measure P. We consider simultaneously the c.t. case, where $\mathbb{U} = \mathbb{R} := (-\infty, \infty)$, and the d.t. case, where $\mathbb{U} = \mathbb{Z} := \{0, \pm 1, \pm 2, \ldots\}$. We assume that $X(u)$ is a *non-degenerate process*, that is, $\mathrm{Var}[X(u)] = \mathbb{E}|X(u)|^2 = r(0) > 0$. (Without loss of generality, we assume that $r(0) = 1$). In the continuous-time case, the process $X(u)$ is also assumed mean-square continuous, that is, $\mathbb{E}[X(t) - X(s)]^2 \to 0$ as $t \to s$ $(t, s \in \mathbb{R})$. This assumption is equivalent to that of the covariance function $r(u)$ be continuous at $u = 0$ (see, e.g., Cramér and Leadbetter [23], Section 5.2).

2.1.2. Spectral representations

By the Herglotz theorem in the d.t. case, and the Bochner-Khintchine theorem in the c.t. case (see, e.g., Cramér and Leadbetter [23], Doob [31], Ibragimov and Linnik [77]), there is a finite measure μ on $(\Lambda, \mathfrak{B}(\Lambda))$, where $\Lambda = \mathbb{R}$ in the c.t. case, and $\Lambda = [-\pi.\pi]$ in the d.t. case, and $\mathfrak{B}(\Lambda)$ is the Borel σ-algebra on Λ, such that for any $t \in \mathbb{U}$ the covariance function $r(t)$ admits the following *spectral representation*:

$$r(u) = \int_\Lambda \exp\{i\lambda u\}d\mu(\lambda), \quad u \in \mathbb{U}. \tag{2.1}$$

The measure μ in (2.1) is called the *spectral measure* of the process $X(u)$. The function $F(\lambda) := \mu[-\pi, \lambda]$ in the d.t. case and $F(\lambda) := \mu(-\infty, \lambda]$ in the c.t. case, is called the *spectral function* of the process $X(t)$. If $F(\lambda)$ is absolutely continuous (with respect to the Lebesgue measure), then the function $f(\lambda) := dF(\lambda)/d\lambda$ is called the *spectral density* of the process $X(t)$. Notice that if the spectral density $f(\lambda)$ exists, then $f(\lambda) \geq 0$, $f(\lambda) \in L^1(\Lambda)$, and (2.1) becomes

$$r(u) = \int_\Lambda \exp\{i\lambda u\}f(\lambda)d\lambda, \quad u \in \mathbb{U}. \tag{2.2}$$

Thus, the covariance function $r(u)$ and the spectral function $F(\lambda)$ (resp. the spectral density function $f(\lambda)$) are equivalent specifications of the second order properties for a stationary process $\{X(u), \ u \in \mathbb{U}\}$.

By the well-known Cramér theorem (see, e.g., Cramér and Leadbetter [23]) for any stationary process $\{X(u), \ u \in \mathbb{U}\}$ with spectral measure μ there exists an orthogonal stochastic measure $Z = Z(B)$, $B \in \mathfrak{B}(\Lambda)$, such that for every $u \in \mathbb{U}$ the process $X(u)$ admits the following *spectral representation*:

$$X(u) = \int_\Lambda \exp\{i\lambda u\}dZ(\lambda), \quad u \in \mathbb{U}. \tag{2.3}$$

Moreover, $E\left[|Z(B)|^2\right] = \mu(B)$ for every $B \in \mathfrak{B}(\Lambda)$. For definition and properties of orthogonal stochastic measures and stochastic integral in (2.3) we refer, e.g., Cramér and Leadbetter [23].

2.1.3. Kolmogorov's isometric isomorphism theorem

Given a probability space $(\Omega, \mathcal{F}, P)$, define the L^2-space of random variables $\xi = \xi(\omega)$, $\mathbb{E}[\xi] = 0$:

$$L^2(\Omega) := \left\{ \xi : \ ||\xi||^2 := \int_\Omega |\xi(\omega)|^2 dP(\omega) < \infty \right\}. \tag{2.4}$$

Then $L^2(\Omega)$ becomes a Hilbert space with the following inner product: for $\xi, \eta \in L^2(\Omega)$

$$(\xi, \eta) = \mathbb{E}[\xi\eta] = \int_\Omega \xi(\omega)\eta(\omega)dP(\omega). \tag{2.5}$$

For $a, b \in \mathbb{U}$, $-\infty \le a \le b \le \infty$, we define the space $H_a^b(X)$ to be the closed linear subspace of the space $L^2(\Omega)$ spanned by the random variables $X(u, \omega)$, $u \in [a, b]$:

$$H_a^b(X) := \overline{sp}\{X(u), \ a \le u \le b\}_{L^2(\Omega)}. \tag{2.6}$$

Observe that the subspace $H_a^b(X)$ consists of all finite linear combinations $\sum_{k=1}^n c_k X(u_k)$ $(a \le u_k \le b)$, as well as, their $L^2(\Omega)$-limits.

The space $H(X) := H_{-\infty}^\infty(X)$ is called the *Hilbert space generated by the process* $X(u)$, or the *time-domain* of $X(u)$.

Let μ be the spectral measure of the process $\{X(u), u \in \mathbb{U}\}$. Consider the weighted L^2-space $L^2(\mu) := L^2(\mu, \Lambda)$ of complex-valued functions $\varphi(\lambda)$, $\lambda \in \Lambda$, defined by

$$L^2(\mu) := \left\{ \varphi(\lambda) : \ ||\varphi||_\mu^2 := \int_\Lambda |\varphi(\lambda)|^2 d\mu(\lambda) < \infty \right\}. \tag{2.7}$$

Then $L^2(\mu)$ becomes a Hilbert space with the following inner product: for $\varphi, \psi \in L^2(\mu)$

$$(\varphi, \psi)_\mu = \int_\Lambda \varphi(\lambda)\overline{\psi}(\lambda)d\mu(\lambda). \tag{2.8}$$

The Hilbert space $L^2(\mu, \Lambda)$ is called the *frequency-domain* of the process $X(u)$.

Kolmogorov's isometric isomorphism theorem states that for any stationary process $X(u)$, $u \in \mathbb{U}$, with spectral measure μ there exists a unique isometric isomorphism V between the time- and frequency-domains $H(X)$ and $L^2(\mu)$, such that $V[X(u)] = e^{iu\lambda}$ for any $u \in \mathbb{U}$.

Thus, any linear problem in the time-domain $H(X)$ can be translated into one in the frequency-domain $L^2(\mu)$, and vice versa. This fact allows to study stationary processes using analytic methods.

2.2. Linear processes. Existence of spectral density functions

We will consider here stationary processes possessing spectral density functions. For the following results we refer to Cramér and Leadbetter [23], Doob [31], Ibragimov and Linnik [77].

(a) The spectral function $F(\lambda)$ of a d.t. stationary process $\{X(u), u \in \mathbb{Z}\}$ is absolutely continuous (with respect to the Lebesgue measure), $F(\lambda) = \int_{-\pi}^{\lambda} f(x)dx$, if and only if it can be represented as an infinite moving average:

$$X(u) = \sum_{k=-\infty}^{\infty} a(u-k)\xi(k), \quad \sum_{k=-\infty}^{\infty} |a(k)|^2 < \infty, \qquad (2.9)$$

where $\{\xi(k), k \in \mathbb{Z}\}$ is a standard white noise, that is, a sequence of centered orthonormal random variables.

(b) The covariance function $r(u)$ and the spectral density $f(\lambda)$ of $X(u)$ are given by formulas:

$$r(u) = \mathbb{E}X(u)X(0) = \sum_{k=-\infty}^{\infty} a(u+k)a(k), \qquad (2.10)$$

and

$$f(\lambda) = \frac{1}{2\pi}|\widehat{a}(\lambda)|^2 = \frac{1}{2\pi}\left|\sum_{k=-\infty}^{\infty} a(k)e^{-ik\lambda}\right|^2, \quad \lambda \in \Lambda. \qquad (2.11)$$

(c) In the case where $\{\xi(k), k \in \mathbb{Z}\}$ is a sequence of Gaussian random variables, the process $\{X(u), u \in \mathbb{Z}\}$ is Gaussian.

Similar results hold for c.t. processes. Indeed, the following results hold.

(a) The spectral function $F(\lambda)$ of a c.t. stationary process $\{X(u), u \in \mathbb{R}\}$ is absolutely continuous (with respect to the Lebesgue measure), $F(\lambda) = \int_{-\infty}^{\lambda} f(x)dx$, if and only if it can be represented as an infinite continuous moving average:

$$X(u) = \int_{\mathbb{R}} a(u-t)d\xi(t), \quad \int_{\mathbb{R}} |a(t)|^2 dt < \infty, \qquad (2.12)$$

where $\{\xi(t), t \in \mathbb{R}\}$ is a process with orthogonal increments and $\mathbb{E}|d\xi(t)|^2 = dt$.

(b) The covariance function $r(u)$ and the spectral density $f(\lambda)$ of $X(u)$ are given by formulas:

$$r(u) = \mathbb{E}X(u)X(0) = \int_{\mathbb{R}} a(u+x)a(x)dx, \qquad (2.13)$$

and

$$f(\lambda) = \frac{1}{2\pi}|\widehat{a}(\lambda)|^2 = \frac{1}{2\pi}\left|\int_{\mathbb{R}} e^{-i\lambda t}a(t)dt\right|^2, \quad \lambda \in \mathbb{R}. \qquad (2.14)$$

(c) In the case where $\{\xi(t), t \in \mathbb{R}\}$ is a Gaussian process, the process $\{X(u), u \in \mathbb{Z}\}$ is Gaussian.

2.3. Lévy-driven linear process

We first recall that 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 $\{N(t) - \mathbb{E}N(t), \ t \geq 0\}$ are typical examples of centered Lévy processes. A Lévy-driven linear process $\{X(t), \ t \in \mathbb{R}\}$ is a real-valued c.t. stationary process defined by (2.12), where $\xi(t)$ is a Lévy process satisfying the conditions:

$$\mathbb{E}\xi(t) = 0, \ \mathbb{E}\xi^2(1) = 1 \text{ and } \mathbb{E}\xi^4(1) < \infty.$$

In the case where $\xi(t) = B(t)$, $X(t)$ is a Gaussian process.

The function $a(\cdot)$ in representations (2.9) and (2.12) plays the role of a *time-invariant filter*, and the linear processes defined by (2.9) and (2.12) can be viewed as the output of a linear filter $a(\cdot)$ applied to the process $\{\xi(u), \ t \in \mathbb{U}\}$, called the innovation or driving process of $X(u)$.

Processes of the form (2.12) appear in many fields of science (economics, finance, physics, etc.), and cover a large class of popular models in c.t. time series modeling. For instance, the so-called c.t. autoregressive moving average (CARMA) models, which are the c.t. analogs of the classical autoregressive moving average (ARMA) models in d.t. case, are of the form (2.12) and play a central role in the representations of c.t. stationary time series (see, e.g., Brockwell [18]).

2.4. Dependence (memory) structure of the model

In the frequency domain setting, the statistical and spectral analysis of stationary processes requires *two types of conditions* on the spectral density $f(\lambda)$. The first type controls the *singularities* of $f(\lambda)$, and involves the *dependence (or memory) structure* of the process, while the second type – controls the *smoothness* of $f(\lambda)$.

We will distinguish the following types of stationary models:

(a) short memory (or short-range dependent),

(b) long memory (or long-range dependent),

(c) intermediate memory (or anti-persistent).

The memory structure of a stationary process is essentially a measure of the dependence between all the variables in the process, considering the effect of all correlations simultaneously. Traditionally memory structure has been defined in the time domain in terms of decay rates of the autocorrelations, or in the frequency domain in terms of rates of explosion of low frequency spectra (see, e.g., Beran [11], Beran et al. [12], Giraitis et al. [59], Guégan [68], Robinson [85], and references therein).

It is convenient to characterize the memory structure in terms of the spectral density function.

2.4.1. Short memory models

Much of statistical inference is concerned with *short memory* stationary models, where the spectral density $f(\lambda)$ of the model is bounded away from zero and infinity, that is, there are constants C_1 and C_2 such that

$$0 < C_1 \le f(\lambda) \le C_2 < \infty \quad \text{for all } \lambda \in \Lambda.$$

A typical d.t. short memory model example is the stationary Autoregressive Moving Average $(\text{ARMA})(p, q)$ process $X(t)$ defined to be a stationary solution of the difference equation:

$$\psi_p(B)X(t) = \theta_q(B)\varepsilon(t), \quad t \in \mathbb{Z},$$

where ψ_p and θ_q are polynomials of degrees p and q, respectively, B is the backshift operator defined by $BX(t) = X(t-1)$, and $\{\varepsilon(t), t \in \mathbb{Z}\}$ is a d.t. white noise, that is, a sequence of zero-mean, uncorrelated random variables with variance σ^2. The covariance $r(k)$ of $(\text{ARMA})(p, q)$ process is exponentially bounded:

$$|r(k)| \le Cr^{-k}, \quad k = 1, 2, \ldots; \quad 0 < C < \infty; \ 0 < r < 1,$$

and the spectral density $f(\lambda)$ is a rational function (see, e.g., Brockwell and Davis [19], Section 3.1):

$$f(\lambda) = \frac{\sigma^2}{2\pi} \cdot \frac{|\theta_q(e^{-i\lambda})|^2}{|\psi_p(e^{-i\lambda})|^2}. \tag{2.15}$$

A typical c.t. short-memory model example is the stationary c.t. $\text{ARMA}(p, q)$ processes, denoted by $\text{CARMA}(p, q)$, which is defined to be the solution of a pth order stochastic differential equation with a suitable initial condition and driven by a standard Brownian motion and its derivatives up to and including order $0 \le q < p$. The spectral density function $f(\lambda)$ of the $\text{CARMA}(p, q)$ process is given by the following formula (see, e.g., Brockwell [18]):

$$f(\lambda) = \frac{\sigma^2}{2\pi} \cdot \frac{|\beta(i\lambda)|^2}{|\alpha(i\lambda)|^2}, \tag{2.16}$$

where $\alpha(z)$ and $\beta(z)$ are polynomials of degrees p and q, respectively.

Another important c.t. short memory model is the *Ornstein-Uhlenbeck* process, which is a Gaussian stationary process with covariance function $r(t) = \sigma^2 e^{-\alpha|t|}$ $(t \in \mathbb{R})$, and spectral density

$$f(\lambda) = \frac{\sigma^2}{\pi} \cdot \frac{\alpha^2}{\lambda^2 + \alpha^2}, \quad \alpha > 0, \lambda \in \mathbb{R}. \tag{2.17}$$

2.4.2. Discrete-time long-memory and anti-persistent models

Data in many fields of science (economics, finance, hydrology, etc.), however, is well modeled by stationary processes whose spectral densities are *unbounded* or *vanishing* at some fixed points (see, e.g., Beran [11], Guégan [68], Palma [83], Taqqu [97] and references therein).

A *long-memory* model is defined to be a stationary process with *unbounded* spectral density, and an *anti-persistent* model – a stationary process with *vanishing* (at some fixed points) spectral density.

In the discrete context, a basic long-memory model is the Autoregressive Fractionally Integrated Moving Average (ARFIMA)$(0, d, 0)$) process $X(t)$ defined to be a stationary solution of the difference equation (see, e.g., Brockwell and Davis [19], Section 13.2):

$$(1 - B)^d X(t) = \varepsilon(t), \quad 0 < d < 1/2,$$

where B is the backshift operator and $\varepsilon(t)$ is a d.t. white noise defined above. The spectral density $f(\lambda)$ of $X(t)$ is given by

$$f(\lambda) = |1 - e^{-i\lambda}|^{-2d} = (2\sin(\lambda/2))^{-2d}, \quad 0 < \lambda \le \pi, \quad 0 < d < 1/2. \quad (2.18)$$

Notice that $f(\lambda) \sim c |\lambda|^{-2d}$ as $\lambda \to 0$, that is, $f(\lambda)$ blows up at $\lambda = 0$ like a power function, which is the typical behavior of a long memory model.

A typical example of an *anti-persistent* model is the ARFIMA$(0, d, 0)$ process $X(t)$ with spectral density $f(\lambda) = |1 - e^{-i\lambda}|^{-2d}$ with $d < 0$, which vanishes at $\lambda = 0$. Note that the condition $d < 1/2$ ensures that $\int_{-\pi}^{\pi} f(\lambda)d\lambda < \infty$, implying that the process $X(t)$ is well defined because $E|X(t)|^2 = \int_{-\pi}^{\pi} f(\lambda)d\lambda$.

Data can also occur in the form of a realization of a 'mixed' short-long-intermediate-memory stationary process $X(t)$. A well-known example of such a process, which appears in many applied problems, is an ARFIMA(p, d, q) process $X(t)$ defined to be a stationary solution of the difference equation:

$$\psi_p(B)(1 - B)^d X(t) = \theta_q(B)\varepsilon(t), \quad d < 1/2,$$

where B is the backshift operator, $\varepsilon(t)$ is a d.t. white noise, and ψ_p and θ_q are polynomials of degrees p and q, respectively. The spectral density $f_X(\lambda)$ of $X(t)$ is given by

$$f_X(\lambda) = |1 - e^{-i\lambda}|^{-2d} f(\lambda), \quad d < 1/2, \quad (2.19)$$

where $f(\lambda)$ is the spectral density of an ARMA(p, q) process, given by (2.15). Observe that for $0 < d < 1/2$ the model $X(t)$ specified by (2.19) displays long-memory, for $d < 0$ – intermediate-memory, and for $d = 0$ – short-memory. For $d \ge 1/2$ the function $f_X(\lambda)$ in (2.19) is not integrable, and thus it cannot represent a spectral density of a stationary process. Also, if $d \le -1$, then the series $X(t)$ is not invertible in the sense that it cannot be used to recover a white noise $\varepsilon(t)$ by passing $X(t)$ through a linear filter (see, e.g., Bondon and Palma [16], and Brockweel and Davis [19]).

The ARFIMA(p, d, q) processes, first introduced by Granger and Joyeux [66], and Hosking [73], became very popular due to their ability in providing a good characterization of the long-run properties of many economic and financial time series. They are also very useful for modeling multivariate time series, since they are able to capture a larger number of long term equilibrium relations among economic variables than the traditional multivariate ARIMA models (see, e.g., Guégan [68], and Henry and Zaffaroni [71] for a survey on this topic).

Another important long-memory model is the fractional Gaussian noise (fGn). To define fGn we first consider the *fractional Brownian motion* (fBm) $\{B_H(t), t \in \mathbb{R}\}$ with Hurst index H, $0 < H < 1$, defined to be a centered Gaussian H-self-similar process having stationary increments, that is, $B_H(t)$ satisfies the following conditions:

(a) $B_H(0) = 0$, $\mathbb{E}[B_H(t)] = 0$, $t \in \mathbb{R}$;

(b) $\{B_H(at), t \in \mathbb{R}\} \overset{d}{=} \{a^H B_H(t), t \in \mathbb{R}\}$ for any $a > 0$;

(c) $\{B_H(t + u) - B_H(u), t \in \mathbb{R}\} \overset{d}{=} \{B_H(t), t \in \mathbb{R}\}$ for each fixed $u \in \mathbb{R}$;

(d) the covariance function is given by

$$\mathrm{Cov}(B_H(s), B_H(t)) = \frac{\sigma_0^2}{2} \left[|t|^{2H} - |s|^{2H} - |t - s|^{2H} \right],$$

where the symbol $\overset{d}{=}$ stands for equality of the finite-dimensional distributions, and $\sigma_0^2 = \mathrm{Var} B_H(1)$. Then the increment process

$$\{X(k) := B_H(k + 1) - B_H(k), k \in \mathbb{Z}\},$$

called *fractional Gaussian noise* (fGn), is a d.t. centered Gaussian stationary process with covariance function

$$r(k) = \frac{\sigma_0^2}{2} \left[|k + 1|^{2H} - |k|^{2H} - |k - 1|^{2H} \right], \quad k \in \mathbb{Z}, \tag{2.20}$$

and spectral density function

$$f(\lambda) = c \, |1 - e^{-i\lambda}|^2 \sum_{k=-\infty}^{\infty} |\lambda + 2\pi k|^{-(2H+1)}, \quad -\pi \le \lambda \le \pi, \tag{2.21}$$

where c is a positive constant.

It follows from (2.21) that $f(\lambda) \sim c \, |\lambda|^{1-2H}$ as $\lambda \to 0$, that is, $f(\lambda)$ blows up if $H > 1/2$ and tends to zero if $H < 1/2$. Also, comparing (2.18) and (2.21), we observe that, up to a constant, the spectral density of fGn has the same behavior at the origin as ARFIMA$(0, d, 0)$ with $d = H - 1/2$.

Thus, the fGn $\{X(k), k \in \mathbb{Z}\}$ has long-memory if $1/2 < H < 1$ and is anti-percipient if $0 < H < 1/2$. The variables $X(k)$, $k \in \mathbb{Z}$, are independent if $H = 1/2$. For more details we refer to Samorodnisky and Taqqu [89] and Taqqu [97, 99].

2.4.3. Continuous-time long-memory and anti-persistent models

In the continuous context, a basic process which has commonly been used to model long-range dependence is the fractional Brownian motion (fBm) $\{B_H(t), t \in \mathbb{R}\}$ with Hurst index H, defined above, which can be regarded as a Gaussian process having a 'spectral density':

$$f(\lambda) = c|\lambda|^{-(2H+1)}, \quad c > 0, \ \ 0 < H < 1, \ \ \lambda \in \mathbb{R}. \tag{2.22}$$

The form (2.22) can be understood in a generalized sense (see Yaglom [108], Section 24, Flandrin [36], Solo [91]), since the fBm B_H is a nonstationary process.

A proper stationary model in lieu of fBm is the *fractional Riesz-Bessel motion* (fRBm), introduced in Anh et al. [2], and defined as a c.t. Gaussian process $X(t)$ with spectral density

$$f(\lambda) = c\,|\lambda|^{-2\alpha}(1 + \lambda^2)^{-\beta}, \quad \lambda \in \mathbb{R}, \, 0 < c < \infty, \tag{2.23}$$

where the exponents α and β are such that $0 < \alpha < 1$, $\beta > 0$ and $\alpha + \beta > 1/2$. The exponent α determines the long-range dependence, while the exponent β indicates the second-order intermittency of the process (see, e.g., Anh et al. [3] and Gao et al. [41]).

Notice that the process $X(t)$, specified by (2.23), is stationary if $0 < \alpha < 1/2$ and is non-stationary with stationary increments if $1/2 \le \alpha < 1$. Observe also that the spectral density (2.23) behaves as $O(|\lambda|^{-2\alpha})$ as $|\lambda| \to 0$ and as $O(|\lambda|^{-2(\alpha+\beta)})$ as $|\lambda| \to \infty$. Thus, under the conditions $0 < \alpha < 1/2$, $\beta > 0$ and $\alpha + \beta > 1/2$, the function $f(\lambda)$ in (2.23) is well-defined for both $|\lambda| \to 0$ and $|\lambda| \to \infty$ due to the presence of the component $(1 + \lambda^2)^{-\beta}$, $\beta > 0$, which is the Fourier transform of the Bessel potential.

Comparing (2.22) and (2.23), we observe that the spectral density of fBm is the limiting case as $\beta \to 0$ that of fRBm with Hurst index $H = \alpha - 1/2$.

Another important c.t. long-memory model is the CARFIMA(p, H, q) processes, which is defined to be the solution of a pth order stochastic differential equation with a suitable initial condition and driven by a fractional Brownian motion with Hurst parameter H and its derivatives up to and including order $0 \le q < p$. The spectral density function $f(\lambda)$ of the CARFIMA(p, H, q) process is given by the following formula (see, e.g., Tsai and Chan [104]):

$$f(\lambda) = \frac{\sigma^2}{2\pi}\Gamma(2H + 1)\sin(\pi H)|\lambda|^{1-2H}\frac{|\beta(i\lambda)|^2}{|\alpha(i\lambda)|^2}, \tag{2.24}$$

where $\alpha(z)$ and $\beta(z)$ are polynomials of degrees p and q, respectively. Notice that for $H = 1/2$, the spectral density function given by (2.24) becomes that of the short-memory CARMA(p, q) process, given by (2.16).

3. CLT for Toeplitz type quadratic functionals for Gaussian and linear processes

In this section we present sufficient conditions for quadratic forms and functionals of the form (1.1) to obey the CLT. The processes considered will be d.t. and

c.t. Gaussian or linear processes with memory. The matrix and the operator that characterize the quadratic form and functional are Toeplitz.

As it was mentioned in the introduction, the limit distributions of the functionals in (1.1) are completely determined by the spectral density $f(\lambda)$ (or covariance function $r(t)$) and the generating function $g(\lambda)$ (or generating kernel $\widehat{g}(t)$). Conditions that are in terms of the spectral density $f(\lambda)$ and the generating function $g(\lambda)$ will be called frequency-domain conditions, while conditions that are in terms of the covariance function $r(t)$ and the generating kernel $\widehat{g}(t)$ will be called time-domain conditions.

3.1. Frequency domain conditions

Let $\{X(u),\ u \in \mathbb{U}\}$ be a centered real-valued Gaussian stationary process with spectral density $f(\lambda)$, $\lambda \in \Lambda$ and covariance function $r(t) := \widehat{f}(t)$, $t \in \mathbb{U}$, where $\mathbb{U}$ and Λ are as in Section 2. We are interested in the asymptotic distribution (as $T \to \infty$) of the following Toeplitz type quadratic functionals of the process $X(u)$:

$$Q_T := Q_T(f, g) = \begin{cases} \sum_{t=1}^T \sum_{s=1}^T \widehat{g}(t-s)X(t)X(s) & \text{in the d.t. case,} \\[2mm] \int_0^T \int_0^T \widehat{g}(t-s)X(t)X(s)\, dt\, ds & \text{in the c.t. case,} \end{cases} \tag{3.1}$$

where $\widehat{g}(t)$ is the Fourier transform of some real, even, integrable function $g(\lambda)$, $\lambda \in \Lambda$. In the d.t. case the functions $f(\lambda)$ and $g(\lambda)$ are assumed to be 2π-periodic and periodically extended to $\mathbb{R}$.

Note. We include the function f in the notation $Q_T(f, g)$ to emphasize that the distribution of the quadratic form depends also on the spectral density f. Let Q_T be as in (3.1). By $\widetilde{Q}_T$ we denote the standard normalized quadratic functional:

$$\widetilde{Q}_T := T^{-1/2}\left(Q_T - \mathbb{E}[Q_T]\right). \tag{3.2}$$

As before, the notation

$$\widetilde{Q}_T \xrightarrow{d} \eta \sim N(0, \sigma^2) \quad \text{as} \quad T \to \infty \tag{3.3}$$

will mean that the distribution of the random variable $\widetilde{Q}_T$ tends (as $T \to \infty$) to the centered normal distribution with variance σ^2.

Toeplitz matrices and operators arise naturally in the theory of stationary processes, and serve as tools, to study many topics of the spectral and statistical analysis of d.t. and c.t. stationary processes.

We first define the truncated Toeplitz matrices and operators, generated by integrable real symmetric functions.

Let $\psi(\lambda)$ be an integrable real symmetric function defined on $\Lambda = [-\pi, \pi]$. For $T = 1, 2, \ldots$, the $(T \times T)$-*truncated Toeplitz matrix* generated by $\psi(\lambda)$, denoted by $B_T(\psi)$, is defined by the following equation (see, e.g., Ginovyan and Sahakyan [52], and Grenander and Szegő [67]):

$$B_T(\psi) := \|\widehat{\psi}(t-s)\|_{t,s=1,2,\ldots,T}$$

$$= \begin{pmatrix} \widehat{\psi}(0) & \widehat{\psi}(-1) & \cdots & \widehat{\psi}(1-T) \\ \widehat{\psi}(1) & \widehat{\psi}(0) & \cdots & \widehat{\psi}(2-T) \\ \cdots & \cdots & \cdots & \cdots \\ \widehat{\psi}(T-1) & \widehat{\psi}(T-2) & \cdots & \widehat{\psi}(0) \end{pmatrix}, \qquad (3.4)$$

where $\widehat{\psi}(t) = \int_\Lambda e^{i\lambda t} \psi(\lambda)\, d\lambda$ $(t \in \mathbb{Z})$ are the Fourier coefficients of ψ.

Given a real number $T > 0$ and an integrable real symmetric function $\psi(\lambda)$ defined on $\mathbb{R} := (-\infty, \infty)$, the *T-truncated Toeplitz operator* (also called *Wiener-Hopf operator*) generated by $\psi(\lambda)$, denoted by $W_T(\psi)$, is defined by the following equation (see, e.g., Ginovyan [45], Ginovyan and Sahakyan [53], Grenander and Szegő [67], Ibragimov [75], and Kac [79]):

$$[W_T(\psi)u](t) = \int_0^T \hat{\psi}(t-s)u(s)ds, \quad u(s) \in L^2[0,T], \qquad (3.5)$$

where $\widehat{\psi}(t) = \int_\mathbb{R} e^{i\lambda t} \psi(\lambda)\, d\lambda$ $(t \in \mathbb{R})$ is the Fourier transform of $\psi(\lambda)$.

Let $A_T(f)$ and $A_T(g)$ denote either the T-truncated Toeplitz operators (in the c.t. case), or the $T \times T$ Toeplitz matrices (in the d.t. case) generated by the functions f and g, respectively. Observe that $A_T(f)$ is the covariance matrix (in the d.t. case) and the covariance operator (in the c.t. case) of the process $\{X(u),\ u \in \mathbb{U}\}$.

We assume below that $f, g \in L^1(\Lambda)$, and with no loss of generality, that $g \geq 0$. Also, we set

$$\sigma_0^2 := 16\pi^3 \int_\Lambda f^2(\lambda) g^2(\lambda)\, d\lambda. \qquad (3.6)$$

As usual $\Lambda = [-\pi, \pi]$ in the d.t. case and $\Lambda = \mathbb{R}$ in the c.t. case.

The theorems that follow contain sufficient conditions expressed in terms of $f(\lambda)$ and $g(\lambda)$ to ensure central limit theorems for standard normalized quadratic functionals $\widetilde{Q}_T$ both for d.t. and c.t. Gaussian processes. Some of the assumptions imposed on f allow for long-range dependence $(f(0) = \infty)$, others for discontinuities at other frequencies. Sometimes the good behavior of one function, say g can compensate for the bad behavior of f and vice versa.

Theorem 3.1. *Let f, g, $A_T(f)$, $A_T(g)$, and $\widetilde{Q}_T$ be as above. Each of the following conditions is sufficient for functional Q_T to obey the CLT, that is,*

$$\widetilde{Q}_T \xrightarrow{d} \eta \sim N(0, \sigma_0^2) \quad \text{as} \quad T \to \infty, \qquad (3.7)$$

with σ_0^2 given by (3.6).

(A) $f \cdot g \in L^1(\Lambda) \cap L^2(\Lambda)$ *and*

$$\chi_2(\widetilde{Q}_T) := \frac{2}{T} tr\big[A_T(f)A_T(g)\big]^2 \longrightarrow \sigma_0^2, \qquad (3.8)$$

where $tr[A]$ stands for the trace of the operator (or the matrix) A.

(B) *The function*

$$\varphi(\mathbf{u}) := \varphi(u_1, u_2, u_3) = \int_\Lambda f(\lambda)g(\lambda - u_1)f(\lambda - u_2)g(\lambda - u_3)\,d\lambda \quad (3.9)$$

belongs to $L^2(\Lambda^3)$ and is continuous at $\mathbf{0} = (0,0,0)$.

(C) $f \in L^1(\Lambda) \cap L^p(\Lambda)$ $(p \geq 2)$ *and* $g \in L^1(\Lambda) \cap L^q(\Lambda)$ $(q \geq 2)$ *with* $1/p + 1/q \leq 1/2$.

(D) $f \in L^1(\Lambda) \cap L^2(\Lambda)$, $g \in L^1(\Lambda) \cap L^2(\Lambda)$, $fg \in L^2(\Lambda)$ *and*

$$\int_\Lambda f^2(\lambda)g^2(\lambda - \mu)\,d\lambda \longrightarrow \int_\Lambda f^2(\lambda)g^2(\lambda)\,d\lambda \quad \text{as} \quad \mu \to 0.$$

Remark 3.1. Observe that assertion (A) implies assertions (B) – (D), and assertion (B) implies assertions (C) and (D) (see Giraitis and Surgailis [60], Ginovyan and Sahakyan [52, 53]). For the d.t. case, assertions (A) and (D) were proved in Giraitis and Surgailis [60] (see also Giraitis et al. [59]); assertions (B) was proved in Ginovyan and Sahakyan [52]; assertion (C) for $p = q = \infty$ was first established by Grenander and Szegő ([67], Section 11.7), while the case $p = 2$, $q = \infty$ was proved by Ibragimov [75] and Rosenblatt [87]; in the general d.t. case assertion (C) was proved by Avram [6]. For the c.t. case, assertions (A) – (D) were proved in Ginovyan [45] and in Ginovyan and Sahakyan [53].

The following theorem provides conditions for the CLT to hold when either f or g is a sufficiently smooth function. Given a number α $(0 < \alpha < 1)$. Denote by $\mathrm{Lip}_\alpha(\Lambda)$ the class of Lipschitz functions on Λ. By definition, $\phi \in \mathrm{Lip}_\alpha(\Lambda)$ if there exists a constant $C < \infty$ such that

$$|\phi(x) - \phi(y)| \leq C|x - y|^\alpha \quad \text{for all} \quad x, y \in \Lambda.$$

Theorem 3.2. *Let either* $f \in \mathrm{Lip}_\alpha(\Lambda)$ *or* $g \in \mathrm{Lip}_\alpha(\Lambda)$ *with* $\alpha > 1/2$, *and let* $fg \in L^2(\Lambda)$. *Then* $\widetilde{Q}_T \overset{d}{\to} \eta \sim N(0, \sigma_0^2)$ *as* $T \to \infty$ *with* σ_0^2 *given by (3.6).*

Theorem 3.2 for the d.t. case was proved in Giraitis and Taqqu [61]. For the c.t. case it can be proved similarly.

To state the next theorem, we need to introduce a class of slowly varying functions at zero. Recall that a function $u(\lambda)$, $\lambda \in \mathbb{R}$, is called slowly varying at zero if it is non-negative and for any $t > 0$

$$\lim_{\lambda \to 0} \frac{u(t\lambda)}{u(\lambda)} \to 1.$$

Denote by $SV_0(\Lambda)$ the class of slowly varying functions at zero $u(\lambda)$, $\lambda \in \Lambda$, satisfying the following conditions: for some $a > 0$, $u(\lambda)$ is bounded on $[-a, a]$, $\lim_{\lambda \to 0} u(\lambda) = 0$, $u(\lambda) = u(-\lambda)$ and $0 < u(\lambda) < u(\mu)$ for $0 < \lambda < \mu < a$. An example of a function belonging to $SV_0(\Lambda)$ is $u(\lambda) = |\ln|\lambda||^{-\gamma}$ with $\gamma > 0$ and $a = 1$.

Theorem 3.3. *Assume that the functions f and g are integrable on $\mathbb{R}$ and bounded outside any neighborhood of the origin, and satisfy for some $a > 0$ the following conditions:*

$$f(\lambda) \le |\lambda|^{-\alpha} L_1(\lambda), \quad |g(\lambda)| \le |\lambda|^{-\beta} L_2(\lambda), \quad \lambda \in [-a, a], \qquad (3.10)$$

for some $\alpha < 1$, $\beta < 1$ with $\alpha + \beta \le 1/2$, where $L_1(x)$ and $L_2(x)$ are slowly varying functions at zero satisfying

$$L_i \in SV_0(\mathbb{R}), \quad \lambda^{-(\alpha+\beta)} L_i(\lambda) \in L^2[-a, a], \ i = 1, 2. \qquad (3.11)$$

Then $\widetilde{Q}_T \xrightarrow{d} \eta \sim N(0, \sigma_0^2)$ as $T \to \infty$ with σ_0^2 given by (3.6).

Remark 3.2. For the d.t. case the result of Theorem 3.3 under the condition $\alpha + \beta < 1/2$ was first obtained by Fox and Taqqu [39]. For the general case, including the critical value $\alpha + \beta = 1/2$ it was proved in Ginovyan and Sahakyan [52]. For the c.t. case the result was proved in Ginovyan and Sahakyan [53].

Remark 3.3. The conditions $\alpha < 1$ and $\beta < 1$ in Theorem 3.3 ensure that the Fourier transforms of f and g are well defined. Observe that when $\alpha > 0$ the process $X(t)$ may exhibit long-range dependence. We also allow here $\alpha + \beta$ to assume the critical value $1/2$. The assumptions $f \cdot g \in L^1(\mathbb{R})$, $f, g \in L^\infty(\mathbb{R} \setminus [-a, a])$ and (3.11) imply that $f \cdot g \in L^2(\mathbb{R})$, so that the variance σ_0^2 in (3.6) is finite.

Remark 3.4. In Theorem 3.3, the assumption that $L_1(x)$ and $L_2(x)$ belong to the class $SV_0(\mathbb{R})$ instead of merely being slowly varying at zero is done in order to deal with the critical case $\alpha + \beta = 1/2$.

Remark 3.5. The functions

$$f(\lambda) = |\lambda|^{-\alpha} |\ln|\lambda||^{-\gamma} \quad \text{and} \quad g(\lambda) = |\lambda|^{-\beta} |\ln|\lambda||^{-\gamma}, \qquad (3.12)$$

where $\alpha < 1$, $\beta < 1$, $\alpha + \beta \le 1/2$ and $\gamma > 1/2$, provide examples of spectral density $f(\lambda)$ and generating function $g(\lambda)$ satisfying the conditions of Theorem 3.3 (see Ginovyan and Sahakyan [52]).

Remark 3.6. The slowly varying functions L_1 and L_2 in (3.10) are important because they provide a great flexibility in the choice of spectral density f and generating function g. Theorem 3.3 shows that in the critical case $\alpha + \beta = 1/2$ the limit distribution of the standard normalized quadratic form $\widetilde{Q}_T$ given by (3.2) is Gaussian and depends on the slowly varying factors L_1 and L_2 through f and g.

Remark 3.7. The functions $f(\lambda)$ and $g(\lambda)$ in Theorem 3.3 have singularities at the point $\lambda = 0$, and are bounded in any neighborhood of this point. It can be shown that the choice of the point $\lambda = 0$ is not essential, and instead it can be taken to be any point $\lambda_0 \in [-a, a]$. Using the properties of the products of Toeplitz matrices and operators it can be shown that Theorem 3.3 remains valid when $f(\lambda)$ and $g(\lambda)$ have singularities of the form (3.12) at the same finite number of points of the segment $[-a, a]$ (see Ginovyan and Sahakyan [52]).

Remark 3.8. Assertion (A) of Theorem 3.1 implies Theorem 3.3. On the other hand, for functions $f(\lambda) = \lambda^{-3/4}$ and $g(\lambda) = \lambda^{3/4}$ satisfying the conditions of Theorem 3.3, the function $\varphi(u_1, u_2, u_3)$ in (3.9) is not defined for $u_2 = 0$, $u_1 \neq 0$, $u_3 \neq 0$, showing that assertion (B) of Theorem 3.1 generally does not imply Theorem 3.3 (see Ginovyan and Sahakyan [52]).

Giraitis and Surgailis [60] proved that assertion (A) of Theorem 3.1 remains valid for d.t. linear processes. More precisely, they proved the following theorem.

Theorem 3.4. *Let* $\{X(u), u \in \mathbb{Z}\}$ *be a d.t. stationary linear process of the form* (2.9) *with spectral density* f. *Let* Q_T *be a quadratic form generated by a function* g *given by* (3.1), *and let* $B_T(f)$ *and* $B_T(g)$ *be the* $T \times T$ *Toeplitz matrices generated by the functions* f *and* g, *respectively (defined according to formula* (3.4)). *Assume that*

$$\frac{1}{T} tr\left[B_T(f)B_T(g)\right]^2 \longrightarrow 8\pi^3 \int_{-\pi}^{\pi} f^2(\lambda)g^2(\lambda)d\lambda < \infty, \qquad (3.13)$$

Then the CLT holds for Q_T, *that is,*

$$\widetilde{Q}_T := T^{-1/2}\left(Q_T - \mathbb{E}[Q_T]\right) \xrightarrow{d} \eta \sim N(0, \sigma^2) \quad \text{as} \quad T \to \infty, \qquad (3.14)$$

where

$$\sigma^2 = 16\pi^3 \int_{-\pi}^{\pi} f^2(\lambda)g^2(\lambda)d\lambda + \kappa_4 \left[2\pi \int_{-\pi}^{\pi} f(\lambda)g(\lambda)d\lambda\right]^2, \qquad (3.15)$$

and κ_4 *is the fourth cumulant of* $\xi(0)$.

The next result, which was proved in Giraitis and Taqqu [61], extends Theorem 3.2 to the case of d.t. linear processes.

Theorem 3.5. *Let* $\{X(u), u \in \mathbb{Z}\}$ *be a d.t. stationary linear process of the form* (2.9) *with spectral density* f *and* $\mathbb{E}[\xi^2(0)] = 1$, *and let* Q_T *be a quadratic form generated by a function* g *given by* (3.1). *Suppose either* $f \in \text{Lip}_\alpha(\Lambda)$ *or* $g \in \text{Lip}_\alpha(\Lambda)$ *with* $\alpha > 1/2$, *and* $fg \in L^2(\Lambda)$. *Then* $\widetilde{Q}_T \xrightarrow{d} \eta \sim N(0, \sigma^2)$ *as* $T \to \infty$ *with* σ^2 *given by* (3.15).

Giraitis and Surgailis [60] conjectured that the relation (3.13), and hence, the CLT for Q_T, holds under the single condition that the integral on the right hand side of (3.13) is finite. Ginovyan [44] showed that the finiteness of this integral does not guarantee convergence in (3.13), and conjectured that positiveness and finiteness of the integral in (3.13) must be sufficient for Q_T to obey the CLT.

The next proposition shows that the condition of positiveness and finiteness of the integral in (3.13) also is not sufficient for Q_T to obey the CLT (see Ginovyan and Sahakyan [52]).

Proposition 3.1. *There exist a spectral density* $f(\lambda)$ *and a generating function* $g(\lambda)$ *such that*

$$0 < \int_{-\pi}^{\pi} f^2(\lambda)\,g^2(\lambda)\,d\lambda < \infty \qquad (3.16)$$

and

$$\lim_{T\to\infty} \sup \chi_2(\widetilde{Q}_T) = \lim_{n\to\infty} \sup \frac{2}{T} tr\left(B_T(f)B_T(g)\right)^2 = \infty, \qquad (3.17)$$

where $\chi_2(\widetilde{Q}_T)$ *is the second cumulant of* $\widetilde{Q}_T$. *Thus, the condition (3.16) does not guarantee convergence in (3.13), and hence is not sufficient for* Q_T *to obey the CLT.*

To construct functions $f(\lambda)$ and $g(\lambda)$ satisfying (3.16) and (3.17), for a fixed $p \geq 2$ we choose a number $q > 1$ to satisfy $1/p + 1/q > 1$, and for such p and q we consider the functions $f_0(\lambda)$ and $g_0(\lambda)$ defined by

$$f_0(\lambda) = \begin{cases} \left(2^s/s^2\right)^{1/p}, & \text{if } 2^{-s-1} \leq \lambda \leq 2^{-s}, \ s = 2m \\ 0, & \text{if } 2^{-s-1} \leq \lambda \leq 2^{-s}, \ s = 2m+1, \end{cases} \qquad (3.18)$$

$$g_0(\lambda) = \begin{cases} \left(2^s/s^2\right)^{1/q}, & \text{if } 2^{-s-1} \leq \lambda \leq 2^{-s}, \ s = 2m+1 \\ 0, & \text{if } 2^{-s-1} \leq \lambda \leq 2^{-s}, \ s = 2m, \end{cases} \qquad (3.19)$$

where m is a positive integer. For an arbitrary finite positive constant C we set $g_\pm(\lambda) = g_0(\lambda) \pm C$. Then the functions $f = f_0$ and $g = g_+$ or $g = g_-$ satisfy (3.16) and (3.17) (for details we refer to Ginovyan [44], and Ginovyan and Sahakyan [52]). Consequently, for these functions the quadratic form Q_T does not obey the CLT, and it is of interest to describe the limiting non-Gaussian distribution of Q_T in this special case.

3.2. Time domain conditions

In this subsection we present time-domain sufficient conditions for a quadratic form Q_T of the form (3.1) to obey the CLT. That is, the conditions are in terms of the covariance function $r(t)$ and the generating kernel $\widehat{g}(t)$. The processes considered here will be d.t. Gaussian or linear processes with memory. Observe that the time-domain conditions stated below are, in general, not equivalent to the frequency-domain conditions stated in Theorems 3.1 and 3.3. The method of proof for establishing the CLT is also different under the time-domain conditions. In this case, diagrams are used and the method of moments is applied. This is because, in the frequency-domain, one can use an approximation technique, essentially replacing the possibly unbounded spectral density by a bounded one, which allowed us to approximate the bivariate quadratic forms by univariate sums of m-dependent random variables. Such an approximation technique, however, does not work with time-domain conditions, and one has to deal directly with the covariances, which decrease slowly.

The next results contain sufficient conditions in terms of the covariance function $r(t)$ and the generating kernel $\widehat{g}(t)$ ensuring central limit theorems for standard normalized quadratic form $\widetilde{Q}_T$ for the d.t. Gaussian processes (see Fox and Taqqu [37], and Giraitis and Taqqu [62]).

Theorem 3.6. *If the covariance function $r(t)$ and the generating kernel $\widehat{g}(t)$ satisfy the following condition:*

$$r \in l^p \, (p \geq 1) \quad and \quad \widehat{g} \in l^q \, (q \geq 1) \quad with \quad 1/p + 1/q \geq 3/2,$$

then the CLT holds for the quadratic form Q_T with limiting variance σ_0^2 given by (3.6).

Remark 3.9. In fact, in Giraitis and Taqqu [62] (see also Giraitis and Taqqu [63]) was proved more general result stating that Theorem 3.6 is true for quadratic forms of the form:

$$Q_{T,m,n} = \sum_{t=1}^{T} \sum_{s=1}^{T} \widehat{g}(t-s) P_{m,n}(X(t), X(s)), \quad T \in \mathbb{N}, \tag{3.20}$$

where $P_{m,n}(X(t), X(s))$ is a bivariate Appell polynomial (Wick power) of the linear variables $X(t)$ and $X(s)$, $m, n \geq 0$, $m + n \geq 1$, and $X(t)$ is a d.t. linear process of the form (2.9). Also, observe that $Q_{T,1,1} = Q_T$.

The next theorem, which was proved in Giraitis and Taqqu [62], shows that under some rather restrictive conditions on the covariance function $r(t)$ and the generating kernel $\widehat{g}(t)$, the long-range dependence of the process $X(t)$ can be compensated by the fast decay of the generating kernel $\widehat{g}(t)$ in such a way that the CLT for Q_T holds. These conditions ensure, in fact, that the sufficient conditions in the frequency domain, provided in Theorem 3.3, are satisfied. The theorem involves quasi-monotone sequences: a sequence $\{b(t), t \in \mathbb{Z}\}$ is said to be *quasi-monotonically convergent* to 0 if $b(t) \rightarrow 0$ and $b(t+1) \leq b(t)(1 + c/t)$ as $t \rightarrow 0$ for some $c > 0$. The sequence $b(t)$ has bounded variation if $\sum_{t=1}^{\infty} |b(t+1) - b(t)| < \infty$.

Theorem 3.7. *Assume that $r(t) = |t|^{-\gamma_1} L_1(|t|)$, $\widehat{g}(t) = |t|^{-\gamma_2} L_2(|t|)$ with $0 < \gamma_1, \gamma_2 < 3$ and $\gamma_1 + \gamma_2 > 3/2$. Suppose in addition that both sequences $\{r(t)\}$ and $\{\widehat{g}(t)\}$ have bounded variation and are quasi-monotonically convergent to 0; if $1 < \gamma_1 < 3$, $r(t)$ has the same sign for large t and satisfies $\sum_{t \in \mathbb{Z}} r(t) = 0$; if $1 < \gamma_2 < 3$, $\widehat{g}(t)$ has the same sign for large t and satisfies $\sum_{t \in \mathbb{Z}} \widehat{g}(t) = 0$. Then the CLT holds for the quadratic form Q_T. The limiting variance is expressed by (3.6) provided that $0 < \gamma_1, \gamma_2 < 1$.*

3.3. Operator conditions

In this subsection we assume that the model $X(t)$ is a d.t. Gaussian process defined on a probability space $(\Omega, \mathcal{F}, P)$, and we are interested in asymptotic normality of the quadratic form $Q_T := Q_T(f, g)$ given by (3.1) (d.t. case). Recall the notation (see formula (3.2))

$$\widetilde{Q}_T(f, g) := T^{-1/2} \left(Q_T(f, g) - \mathbb{E}[Q_T(f, g)] \right). \tag{3.21}$$

and (see formula (3.6))

$$\sigma^2(f,g) := 16\pi^3 \int_{-\pi}^{\pi} f^2(\lambda)g^2(\lambda)\,d\lambda. \tag{3.22}$$

We denote by $L^2(dP)$ the L^2-space constructed from a probability measure P such that $X(t) \in L^2(dP)$, $t \in \mathbb{Z}$.

Solev and Gerville-Reache [90] observed that for a fixed spectral density f and a number $T \in \mathbb{N}$ the quadratic form $\widetilde{Q}_T(f,g)$ in (3.21) can be regarded as the value of a linear operator $\widetilde{\mathbf{Q}}_T : g \mapsto \widetilde{Q}_T(f,g)$. It turns out that in order to study the asymptotic normality of the quadratic form $\widetilde{Q}_T(f,g)$, it is enough to understand for what sets B_ε (possibly, depending on f) we have

$$\limsup_{T\to\infty} \sup_{g\in B_\varepsilon} ||\widetilde{Q}_T(f,g)||_{L^2(dP)} \le \varepsilon. \tag{3.23}$$

It can be shown that the function $||\widetilde{Q}_T(f,g)||_{L^2(dP)}$ possesses the following symmetry property regarded as a function of (f,g): if g is a nonnegative function, then

$$||\widetilde{Q}_T(f,g)||^2_{L^2(dP)} = ||\widetilde{Q}_T(g,f)||^2_{L^2(dP)}. \tag{3.24}$$

The next result plays a key role in the proof of the asymptotic normality of $\widetilde{Q}_T(f,g)$ (see Solev and Gerville-Reache [90]).

Theorem 3.8. *Let $\mathcal{G}_0$ and $\mathcal{G}$ ($\mathcal{G}_0 \subset \mathcal{G}$) be linear subsets in the space $L^1[-\pi,\pi]$. Assume that the following conditions hold:*

(a) for every $g_0 \in \mathcal{G}_0$ the limit

$$\lim_{T\to\infty} ||\widetilde{Q}_T(f,g_0)||^2_{L^2(dP)} = \sigma^2(f,g_0) < \infty.$$

exists, and

$$\widetilde{Q}_T(f,g_0) \xrightarrow{d} \eta_0 \sim N(0,\sigma^2(f,g_0)) \quad \text{as} \quad T \to \infty;$$

(b) for every $g \in \mathcal{G}$ and a number $\varepsilon > 0$ there exists a function $g_0 \in \mathcal{G}_0$ such that

$$\limsup_{T\to\infty} ||\widetilde{Q}_T(f,g-g_0||_{L^2(dP)} \le \varepsilon, \quad ||g-g_0)||_{f^2} \le \varepsilon,$$

where $||\cdot||_{L^2(dP)}$ and $||\cdot||_{f^2}$ are the norms in the spaces $L^2(dP)$ and $L^2(f^2,[-\pi,\pi])$, respectively. Then $\sigma^2(f,g) < \infty$, the limit

$$\lim_{T\to\infty} ||\widetilde{Q}_T(f,g)||^2_{L^2(dP)} = \sigma^2(f,g)$$

exists, and

$$\widetilde{Q}_T(f,g) \xrightarrow{d} \eta \sim N(0,\sigma^2(f,g)) \quad \text{as} \quad T \to \infty. \tag{3.25}$$

Thus, if the results holds for every $g_0 \in \mathcal{G}_0$ (part (a)), then it holds for every $g \in \mathcal{G}$ close to g_0 (part (b)).

Given a function $\phi \in L^1[-\pi, \pi]$ and an interval $I \subset [-\pi, \pi]$ of length $|I|$, we set

$$\{\phi\}_I := \frac{1}{|I|} \int_I \phi(x) dx,$$

and define

$$\lambda(\phi) := \sup_I \{\phi\}_I \times \{1/\phi\}_I, \tag{3.26}$$

where the supremum is taken over all intervals $I \subset [-\pi, \pi]$. The condition $\lambda(\phi) < \infty$ implies a type of smoothness of the function $\psi = \log|\phi|$ in the neighborhoods of points where it turns to infinity (for details, see the paper Hunt et al. [74]).

The following result was obtained by Solev and Gerville-Reache [90].

Theorem 3.9. *Assume that the following conditions hold:*

(a) $f \leq f_*$, $|g| \leq g_*$ *and* $\lambda(f_*), \lambda(g_*) < \infty$,
(b) $f_* g_* \in L^2[-\pi, \pi]$.

Then the quadratic form $Q_T(f, g)$ *obeys the CLT, that is, the relation* (3.25) *is satisfied.*

4. Functional limit theorems for Gaussian and linear models

In this section, we establish weak convergence in $C[0, 1]$ for c.t. case and in $D[0, 1]$ for d.t. case of normalized stochastic processes, generated by Toeplitz type quadratic functionals of a Gaussian stationary process, exhibiting long-range dependence. Specifically, we are interested in describing the limiting behavior (as $T \to \infty$) of the following process $Q_T(\tau)$, generated by Toeplitz type quadratic functionals of a Gaussian stationary process $\{X(u), \ u \in \mathbb{U}\}$ with spectral density f:

$$Q_T(\tau) = \begin{cases} \sum_{t=1}^{[T\tau]} \sum_{s=1}^{[T\tau]} \widehat{g}(t-s) X(t) X(s) & \text{in the d.t. case,} \\[2mm] \int_0^{T\tau} \int_0^{T\tau} \widehat{g}(t-s) X(t) X(s) \, dt \, ds & \text{in the c.t. case,} \end{cases} \tag{4.1}$$

where $\tau \in [0, 1]$, $\widehat{g}(t)$, $t \in \mathbb{U}$ is the Fourier transform of some integrable even function $g(\lambda)$, $\lambda \in \Lambda$, and $[\cdot]$ stands for the greatest integer.

The limit of the process (4.1) again is completely determined by the spectral density $f(\lambda)$ (or covariance function $r(t)$) and the generating function $g(\lambda)$ (or generating kernel $\widehat{g}(t)$), and depending on their properties, the limiting process can be Gaussian or not. We say that we have a *functional non-central limit theorem* (FNCLT) for the process $Q_T(\tau)$ if the limit is non-Gaussian or, if Gaussian, it is not Brownian motion. We say that a *functional central limit*

theorem (FCLT) for $Q_T(\tau)$ holds if we have the following weak convergence in $C[0,1]$ in the c.t. case (and in $D[0,1]$ in the d.t. case):

$$\widetilde{Q}_T(\tau) \Longrightarrow \sigma B(\tau), \quad \tau \in [0,1], \tag{4.2}$$

where $\sigma > 0$, $B(\tau)$ is a standard Brownian motion, and $\widetilde{Q}_T(\tau)$ is the standard normalized version of the process $Q_T(\tau)$ in (4.1):

$$\widetilde{Q}_T(\tau) := T^{-1/2} \left(Q_T(\tau) - \mathbb{E}[Q_T(\tau)] \right), \quad \tau \in [0,1]. \tag{4.3}$$

4.1. Functional CLT for $Q_T(\tau)$

We first state functional central limit theorems, when the limit of the normalized process (4.3) is Brownian motion.

In the d.t. case the next result was obtained in Giraitis and Taqqu [65] (see also Fox and Taqqu [38] and Giraitis and Taqqu [61]).

Theorem 4.1. *If the covariance function $r(t)$ and the generating kernel $\widehat{g}(t)$ satisfy the following condition: $r \in l^p$ $(p \geq 1)$ and $\widehat{g} \in l^q$ $(q \geq 1)$ with $1/p + 1/q \geq 3/2$, then for the process $Q_T(\tau)$ the FCLT holds, that is, the convergence (4.2) holds in $D[0,1]$ with limiting variance σ^2 given by*

$$\sigma^2 := \sum_{t,s,v \in \mathbb{Z}} \widehat{g}(t)\widehat{g}(s)\mathrm{Cov}\left(P(X(v), X(v+t)), P(X(0), X(s))\right), \tag{4.4}$$

where $P(X(t), X(s)) := X(t)X(s) - \mathbb{E}[X(t)X(s)]$.

Now we consider the c.t. case. The next result, which is an extension of Theorem 3.1(A), involves the convergence of finite-dimensional distributions of the process $\widetilde{Q}_T(\tau)$ to those of a standard Brownian motion (see Bai et al. [8]).

Theorem 4.2. *Assume that the spectral density $f(\lambda)$ and the generating function $g(\lambda)$ satisfy the following conditions:*

$$f \cdot g \in L^1(\mathbb{R}) \cap L^2(\mathbb{R}) \tag{4.5}$$

and

$$\mathbb{E}[\widetilde{Q}_T^2(1)] \to 16\pi^3 \int_{\mathbb{R}} f^2(x)g^2(x)dx \quad as \ T \to \infty. \tag{4.6}$$

Then we have the following convergence of the finite-dimensional distributions

$$\widetilde{Q}_T(\tau) \overset{f.d.d.}{\longrightarrow} \sigma_0 B(\tau),$$

where $\widetilde{Q}_T(\tau)$ is as in (4.3), $B(\tau)$ is a standard Brownian motion, and (see (3.6)):

$$\sigma_0^2 := 16\pi^3 \int_{\mathbb{R}} f^2(x)g^2(x)dx. \tag{4.7}$$

To extend the convergence of finite-dimensional distributions in Theorem 4.2 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 4.1

$$r(\cdot) \in L^p(\mathbb{R}), \quad a(\cdot) \in L^q(\mathbb{R}) \quad \text{for some} \quad p, q \geq 1, \; \frac{1}{p} + \frac{1}{q} \geq \frac{3}{2}. \qquad (4.8)$$

Remark 4.1. In fact under (4.5), the condition (4.8) is sufficient for the convergence in (4.6). Indeed, let $\bar{p} = p/(p-1)$ and $\bar{q} = q/(q-1)$ be the Hölder conjugates of p and q, respectively. Since $1 \leq p, q \leq 2$, one has by the Hausdorff-Young inequality that $\|f\|_{\bar{p}} \leq c_p \|r\|_p$, $\|g\|_{\bar{q}} \leq c_q \|a\|_q$, and hence, in view if (4.8) we have

$$f(\cdot) \in L^{\bar{p}}, \quad g(\cdot) \in L^{\bar{q}}, \quad \frac{1}{\bar{p}} + \frac{1}{\bar{q}} = 2 - \frac{1}{p} - \frac{1}{q} \leq 1/2. \qquad (4.9)$$

Then the convergence in (4.6) follows from Theorem 3.1(C) (see the proof of Theorem 3 of Ginovyan and Sahakyan [53] in the c.t. case, and Giraitis and Surgailis [60] in the d.t. case).

Remark 4.2. Observe that condition (4.8) is satisfied if the functions $r(t)$ and $a(t)$ satisfy the following: there exist constants $C > 0$, α^* and β^*, such that

$$|r(t)| \leq C(1 \wedge |t|^{\alpha^*-1}), \qquad |a(t)| \leq C(1 \wedge |t|^{\beta^*-1}), \qquad (4.10)$$

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 3.1(A) and 3.3 to weak convergence in the space $C[0,1]$ of the stochastic process $\widetilde{Q}_T(\tau)$ to a standard Brownian motion.

Theorem 4.3. *Let the spectral density f and the generating function g satisfy condition (4.5). Let the covariance function $r(t)$ and the generating kernel $a(t)$ satisfy condition (4.8). Then we have the following weak convergence in $C[0,1]$:*

$$\widetilde{Q}_T(\tau) \Longrightarrow \sigma_0 B(\tau),$$

where $\widetilde{Q}_T(\tau)$ is as in (4.3), σ_0 is as in (4.7), and $B(\tau)$ is a standard Brownian motion.

Recall the class $SV_0(\mathbb{R})$ 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 $[-a, a]$, $\lim_{x \to 0} u(x) = 0$, $u(x) = u(-x)$ and $0 < u(x) < u(y)$ for $0 < x < y < a$.

Theorem 4.4. *Assume that the functions f and g are integrable on $\mathbb{R}$ and bounded outside any neighborhood of the origin, and satisfy for some $a > 0$*

$$f(x) \le |x|^{-\alpha} L_1(x), \quad |g(x)| \le |x|^{-\beta} L_2(x), \quad x \in [-a, a] \qquad (4.11)$$

for some $\alpha < 1$, $\beta < 1$ with $\alpha + \beta \le 1/2$, where $L_1(x)$ and $L_2(x)$ are slowly varying at zero functions satisfying

$$L_i \in SV_0(\mathbb{R}), \quad |x|^{-(\alpha+\beta)} L_i(x) \in L^2[-a, a], \quad i = 1, 2. \qquad (4.12)$$

Let, in addition, the covariance function $r(t)$ and the generating kernel $a(t)$ satisfy condition (4.8). Then we have the following weak convergence in $C[0, 1]$:

$$\widetilde{Q}_T(\tau) \Longrightarrow \sigma_0 B(\tau),$$

where $\widetilde{Q}_T(\tau)$ is as in (4.3), σ_0 is as in (4.7), and $B(\tau)$ is a standard Brownian motion.

Remark 4.3. The proofs of Theorems 4.2 - 4.4, given in Bai et al. [8] uses the method developed in Ginovyan and Sahakyan [53], which itself is based on the approximations of traces of the products of truncated Toeplitz operators (see Section 8.1).

4.2. Functional NCLT for $Q_T(\tau)$

4.2.1. Non-central limit theorems (discrete-time)

For d.t. Gaussian processes the problem of description of the limit distribution of the quadratic form:

$$Q_T := \sum_{t=1}^{T} \sum_{s=1}^{T} \widehat{g}(t - s) X(t) X(s), \quad T \in \mathbb{N}, \qquad (4.13)$$

if it is non-Gaussian, goes back to the classical papers of Rosenblatt [86]–[88]. Later this problem was studied in a series of papers by Fox and Taqqu, Taqqu, and Terrin and Taqqu (see, e.g., [37], [97], [100], [101], [102], and references therein). Suppose that the spectral density $f(\lambda)$ and the generating function $g(\lambda)$ are regularly varying functions at the origin:

$$f(\lambda) = |\lambda|^{-\alpha} L_1(\lambda) \quad \text{and} \quad g(\lambda) = |\lambda|^{-\beta} L_2(\lambda), \quad \alpha < 1, \beta < 1, \qquad (4.14)$$

where $L_1(\lambda)$ and $L_2(\lambda)$ are slowly varying functions at zero, which are bounded on bounded intervals. The conditions $\alpha < 1$ and $\beta < 1$ ensure that the Fourier coefficients of f and g are well defined. When $\alpha > 0$ the model $\{X(t), t \in \mathbb{Z}\}$ exhibits long memory.

It is the sum $\alpha + \beta$ that determines the asymptotic behavior of the quadratic form Q_T. If $\alpha + \beta \le 1/2$, then by Theorem 3.3 the standard normalized quadratic form

$$T^{-1/2} (Q_T - \mathbb{E}[Q_T])$$

converges in distribution to a Gaussian random variable. If $\alpha + \beta > 1/2$, the convergence to a Gaussian distribution fails.

Consider the embedding of the discrete sequence $\{Q_T, T \in \mathbb{N}\}$ into a continuous-time process $\{Q_T(\tau), T \in \mathbb{N}, \tau \in [0,1]\}$ defined by

$$Q_T(\tau) := \sum_{t=1}^{[T\tau]} \sum_{s=1}^{[T\tau]} \widehat{g}(t-s)X(t)X(s), \qquad (4.15)$$

where $[\cdot]$ stands for the greatest integer. The problem of interest is to describe the limit distribution of the following normalized process:

$$\widetilde{Q}_T(\tau) := d_T^{-1} \left(Q_T(\tau) - \mathbb{E}[Q_T(\tau)] \right) \qquad (4.16)$$

where d_T is a suitably chosen normalization factor.

In [86] (see also [88]) Rosenblatt showed that if a d.t. centered Gaussian process $X(t)$ has covariance function $r(t) = (1+t^2)^{\alpha/2-1/2}$ with $1/2 < \alpha < 1$, then the random variable

$$\widetilde{Q}_T := T^{-\alpha} \sum_{t=1}^{T} \left[X^2(t) - 1 \right]$$

has a non-Gaussian limiting distribution, and described this distribution in terms of characteristic functions. This is a special case of (4.16) with $\tau = 1$, $1/2 < \alpha < 1$, $\beta = 0$ and $d_T = T^\alpha$. In [97] (see also [100]) Taqqu extended Rosenblatt's result by showing that the stochastic process

$$\widetilde{Q}_T(\tau) := T^{-\alpha} \sum_{t=1}^{[T\tau]} \left[X^2(t) - 1 \right]$$

converges (as $T \to \infty$) weakly in $D[0,1]$ to a process (called the Rosenblatt process) which has the double Wiener-Itô integral representation

$$Q(\tau) := C_\alpha \int_{\mathbb{R}^2}'' \frac{e^{i\tau(x+y)} - 1}{i(x+y)} |x|^{-\alpha/2} |y|^{-\alpha/2} dZ(x) dZ(y), \qquad (4.17)$$

where $Z(\cdot)$ is a complex-valued Gaussian random measure defined on the Borel σ-algebra $\mathcal{B}(\mathbb{R})$, and satisfying $E[Z(B)] = 0$, $E|Z(B)|^2 = |B|$, and $\overline{Z(-B)} = Z(B)$ for any $B \in \mathcal{B}(\mathbb{R})$. The double prime in the integral (4.17) indicates that the integration excludes the diagonals $x = \pm y$.

Notice that the distribution of the random variable $Q(\tau)$ in (4.17) for $\tau = 1$ is described in Veillette and Taqqu [105].

The next result, proved in Terrin and Taqqu [101] using power counting theorems (see Section 8.6), describes the non-Gaussian limit distribution of the suitable normalized process $Q_T(\tau)$.

Theorem 4.5. *Let $f(\lambda)$ and $g(\lambda)$ be as in (4.14) with $\alpha < 1$, $\beta < 1$ and slowly varying at zero and bounded on bounded intervals factors $L_1(\lambda)$ and $L_2(\lambda)$. Let the process $Q_T(\tau)$ be as in (4.15). Then for $\alpha + \beta > 1/2$ the process:*

$$Z_T(\tau) := \frac{1}{T^{\alpha+\beta}L_1(1/T)L_2(1/T)} \left(Q_T(\tau) - \mathbb{E}[Q_T(\tau)]\right) \qquad (4.18)$$

converges (as $T \to \infty$) weakly in $D[0,1]$ to

$$Z(\tau) := \int_{\mathbb{R}^2}'' K_\tau(x,y)dZ(x)dZ(y), \qquad (4.19)$$

where

$$K_\tau(x,y) = |xy|^{-\alpha/2} \int_{\mathbb{R}} \frac{e^{i\tau(x+u)} - 1}{i(x+u)} \cdot \frac{e^{i\tau(y-u)} - 1}{i(y-u)} |u|^{-\beta}du, \qquad (4.20)$$

The double prime in the integral (4.19) indicates that the integration excludes the diagonals $x = \pm y$ (for the definition of the integral $\int''$ see Section 8.5, formula (8.25)).

Remark 4.4. The limiting process in (4.19) is real-valued, non-Gaussian, and satisfies $\mathbb{E}[Z(\tau)] = 0$ and $\mathbb{E}[Z^2(\tau)] = \int_{\mathbb{R}^2} |K_\tau(x,y)|^2 dxdy$. It is self-similar with parameter $H = \alpha + \beta \in (1/2, 2)$, that is, the processes $\{Z(a\tau), \tau \in [0,1]\}$ and $\{a^H Z(\tau), \tau \in [0,1]\}$ have the same finite dimensional distributions for all $a > 0$.

4.2.2. Non-central limit theorems (continuous-time)

Now we state a *non-central limit theorem* for the process $Q_T(\tau)$ in the c.t. case, that is, when $Q_T(\tau)$ in (4.1) is defined through integrals. Let the spectral density f and the generating function g satisfy

$$f(x) = |x|^{-\alpha}L_1(x) \quad \text{and} \quad g(x) = |x|^{-\beta}L_2(x), \quad x \in \mathbb{R}, \quad \alpha < 1,\ \beta < 1,\ (4.21)$$

with slowly varying at zero functions $L_1(x)$ and $L_2(x)$ such that

$$\int_{\mathbb{R}} |x|^{-\alpha}L_1(x)dx < \infty \quad \text{and} \quad \int_{\mathbb{R}} |x|^{-\beta}L_2(x)dx < \infty.$$

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. [59], formula (2.3.5)): *for any $\epsilon > 0$ there exists a constant $C > 0$ so that if T is large enough, then*

$$\frac{L_i(u/T)}{L_i(1/T)} \le C(|u|^\epsilon + |u|^{-\epsilon}), \quad i = 1, 2. \qquad (4.22)$$

Note that a sufficient condition for (4.22) 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 4.4.

Now we are interested in the limit process of the following normalized version of the process $Q_T(\tau)$ given by (4.1), with f and g as in (4.21):

$$Z_T(\tau) := \frac{1}{T^{\alpha+\beta}L_1(1/T)L_2(1/T)}\left(Q_T(\tau) - \mathbb{E}[Q_T(\tau)]\right). \tag{4.23}$$

Theorem 4.6. *Let f and g be as in (4.21) with $\alpha < 1$, $\beta < 1$ and slowly varying at zero functions $L_1(x)$ and $L_2(x)$ satisfying (4.22), and let $Z_T(\tau)$ be as in (4.23). Then if $\alpha + \beta > 1/2$, we have the following weak convergence in the space $C[0, 1]$:*

$$Z_T(\tau) \Longrightarrow Z(\tau),$$

where the limit process $Z(\tau)$ is given by

$$Z(\tau) = \int_{\mathbb{R}^2}^{\prime\prime} H_\tau(x_1, x_2)W(dx_1)W(dx_2), \tag{4.24}$$

with

$$H_\tau(x_1, x_2) = |x_1 x_2|^{-\alpha/2}\int_{\mathbb{R}}\left[\frac{e^{i\tau(x_1+u)}-1}{i(x_1+u)}\right]\cdot\left[\frac{e^{i\tau(x_2-u)}-1}{i(x_2-u)}\right]|u|^{-\beta}du, \tag{4.25}$$

where $W(\cdot)$ is a complex Gaussian random measure with Lebesgue control measure.

Remark 4.5. Let P_T and P denote the measures generated in $C[0, 1]$ by the processes $Z_T(\tau)$ and $Z(\tau)$ given by (4.23) and (4.24), respectively. Then Theorem 4.6 can be restated as follows: under the conditions of Theorem 4.6, the measure P_T converges weakly in $C[0, 1]$ to the measure P as $T \to \infty$. Similar assertions can be stated for Theorems 4.3 and 4.4.

Remark 4.6. Comparing Theorems 4.5 and 4.6, we see that the limit process $Z(\tau)$ is the same both for continuous and discrete time models. Also, it is worth noting that although the statements of Theorems 4.5 and 4.6 are similar, the proofs are different. The proof of Theorem 4.5, given in Terrin and Taqqu [101], uses hard technical analysis based on power counting theorems, while the proof of Theorem 4.6, given in Bai et al. [9], is simple and uses the spectral representation of the underlying process and properties of Wiener-Itô integrals.

5. Functional limit theorems for Lévy-driven linear models

In this section, we survey results involving functional central and non-central limit theorems for a suitably normalized stochastic process:

$$Q_T(\tau) := \int_0^{T\tau}\int_0^{T\tau}\widehat{g}(t-s)X(t)X(s)dtds, \quad \tau \in [0, 1] \tag{5.1}$$

in the general case where the underlying model $\{X(t),\ t \in \mathbb{R}\}$ is a c.t. linear process driven from *Lévy noise* $\xi(t)$ with time invariant filter $a(\cdot)$, that is,

$$X(t) = \int_{\mathbb{R}}a(t-s)d\xi(s), \quad \int_{\mathbb{R}}|a(s)|^2 ds < \infty. \tag{5.2}$$

These theorems, which were stated and proved in Bai et al. [9], extend the results stated in Section 4 for Wiener-driven processes and show that under some (L^p, L^q)-type conditions imposed on the filter $a(\cdot)$ and the generating kernel $\widehat{g}(\cdot)$ of the quadratic functional, the process $Q_T(\tau)$ obeys a central limit theorem, that is, the finite-dimensional distributions of the standard $\sqrt{T}$ normalized process $Q_T(\tau)$ in (5.1) tend to those of a normalized *standard Brownian motion*. In contrast, when the functions $a(\cdot)$ and $\widehat{g}(\cdot)$ have slow power decay, then we have a non-central limit theorem for $Q_T(\tau)$, that is, the finite-dimensional distributions of the process $Q_T(\tau)$, 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 the proofs of the central limit theorems given in Bai et al. [9] were based on an approximation approach, which reduces the quadratic integral form to a single integral form, while the proofs of the non-central limit theorems, used 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 [94] (see also Giraitis et al. [59]).

It is worth noting that if 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 5.2 below). For this model, in Avram et al. [7], a central limit theorem for the quadratic functional $Q_T(1)$ was stated (without proof) under some (L^p, L^q)-type conditions imposed on the spectral density $f(\lambda)$ and the generating function $g(\lambda)$ (see Remark 5.5 below). For a related study of the sample covariances of Lévy-driven moving average processes we refer to the papers by Cohen and Lindner [22], and Spangenberg [93].

In this section, we follow the paper Bai et al. [9], and will use the following notation.

The symbol $*$ will stand for the convolution:

$$(\phi * \psi)(u) = \int_{\mathbb{R}} \phi(u - x)\psi(x)dx,$$

while the symbol $\bar{*}$ is used to denote the reversed convolution:

$$(\phi^{\bar{*}2})(u) = (\phi\bar{*}\phi)(u) = \int_{\mathbb{R}} \phi(u + x)\phi(x)dx.$$

By $\mathcal{F}$ and $\mathcal{F}^{-1}$ we denote the Fourier and the inverse Fourier transforms:

$$(\mathcal{F}\phi)(u) = \widehat{\phi}(u) = \int_{\mathbb{R}} e^{ixu}\phi(x)dx, \quad (\mathcal{F}^{-1}\phi)(u) = \frac{1}{2\pi}\int_{\mathbb{R}} e^{-ixu}\phi(x)dx.$$

We will use the following well-known identities:

$$\mathcal{F}(\phi * \psi) = \mathcal{F}(\phi) \cdot \mathcal{F}(\psi) \quad \text{and} \quad \mathcal{F}(\phi\bar{*}\psi) = \mathcal{F}(\phi) \cdot \overline{\mathcal{F}(\psi)}. \tag{5.3}$$

5.1. Central limit theorems

Let $\{X(t),\ t \in \mathbb{R}\}$ be a centered real-valued linear process given by (5.2) with filter $a(\cdot) \in L^2(\mathbb{R})$. Recall that the covariance function $r(t)$ and the spectral density $f(\lambda)$ of $X(t)$ are given by formulas (2.13) and (2.14), respectively.

The theorem that follows contains (L^p, L^q)-type time-domain sufficient conditions for the process $Q_T(\tau)$ to obey the central limit theorem (see Bai et al. [9]).

Theorem 5.1. *Let $X(t)$ be as in (5.2), and let $Q_T(\tau)$ be as in (5.1). Assume that*

$$a(\cdot) \in L^p(\mathbb{R}) \cap L^2(\mathbb{R}), \quad \widehat{g}(\cdot) \in L^q(\mathbb{R}) \tag{5.4}$$

with

$$1 \leq p, q \leq 2, \quad \frac{2}{p} + \frac{1}{q} \geq \frac{5}{2}. \tag{5.5}$$

Then

$$\widetilde{Q}_T(\tau) := T^{-1/2} \left(Q_T(\tau) - \mathbb{E}[Q_T(\tau)] \right) \stackrel{f.d.d.}{\longrightarrow} \sigma B(\tau), \tag{5.6}$$

where the symbol $\stackrel{f.d.d.}{\longrightarrow}$ stands for convergence of finite-dimensional distributions, $B(\tau)$ is a standard Brownian motion, and

$$\sigma^2 = \int_{\mathbb{R}} [2K_A(v) + \kappa_4 K_B(v)] \ dv, \tag{5.7}$$

where κ_4 is the fourth cumulant of $\xi(1)$, and

$$K_A(v) = \left((a * b)^{\bar{*}2} \cdot a^{\bar{*}2} \right)(v), \quad K_B(v) = \left((a*b) \cdot a \right)^{\bar{*}2}(v). \tag{5.8}$$

Remark 5.1. Young's convolution inequality (see, e.g., Bogachev [15], Theorem 3.9.4) states that for any numbers p, p_1, q satisfying $1 \leq p, q \leq p_1 \leq \infty$ 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

$$\|f * g\|_{p_1} \leq \|f\|_p \|g\|_q. \tag{5.9}$$

Applying this inequality to the convolution in (2.13), we get $\|r\|_{p_1} \leq \|a\|_p^2 < \infty$, where $1 + 1/p_1 = 2/p$. Hence the relations (5.4) and (5.5) imply that

$$r(\cdot) \in L^{p_1}(\mathbb{R}), \quad b(\cdot) \in L^q(\mathbb{R}), \quad \frac{1}{p_1} + \frac{1}{q} = \frac{2}{p} - 1 + \frac{1}{q} \geq \frac{5}{2} - 1 = \frac{3}{2}. \tag{5.10}$$

The condition (5.10) is sufficient for the convergence in Theorem 5.1 to hold in the case where $\xi(t)$ is Brownian motion (see Theorem 4.3). In fact, in this case, the convergence in Theorem 5.1 holds under even a weaker condition imposed on the generating function $g(\lambda)$ and the spectral density $f(\lambda)$ of $X(t)$ (see Theorem 4.2).

Remark 5.2. In contrast to the cases where the model is either a d.t. linear process (see Giraitis and Surgailis [60]), or a c.t. Gaussian process (see Bai et al. [8]), it is convenient to impose the time-domain conditions (5.4) and (5.5) on the functions $a(\cdot)$ and $\widehat{g}(\cdot)$, instead of on the spectral density $f(\lambda)$ and the generating function $g(\lambda)$. This allows the analysis of the random diagonal term which arises from the double stochastic integral with respect to a non-Gaussian Lévy process. In the d.t. case the random diagonal term is estimated by the full double sum (see Giraitis and Surgailis [60], relation (2.3)), while in the c.t. Gaussian case, there is no such random diagonal term (see Ginovyan and Sahakyan [53]). In the c.t. non-Gaussian case, there is a random diagonal term in the form of a single stochastic integral that cannot be controlled by the double integral, and hence it should be treated separately (see the proof of Theorem 2.1 of Bai et al. [9]).

Remark 5.3. Observe that the long-run variance σ^2 given by (5.7) 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 the equalities in (5.3) and the Parseval-Plancherel theorem, under suitable integrability conditions on $a(\cdot)$ and $b(\cdot)$, we can write

$$
\begin{aligned}
\int_{\mathbb{R}} K_A(v)dv &= \int_{\mathbb{R}} (a*b)^{\bar{*}2}(v) a^{\bar{*}2}(v)dv = \frac{1}{2\pi} \int_{\mathbb{R}} \mathcal{F}\left((a*b)^{\bar{*}2}\right)(\lambda) \overline{\mathcal{F}\left(a^{\bar{*}2}\right)(\lambda)}d\lambda \\
&= \frac{1}{2\pi} \int_{\mathbb{R}} |\mathcal{F}(a*b)(\lambda)|^2 |\mathcal{F}(a)(\lambda)|^2 d\lambda = \frac{1}{2\pi} \int_{\mathbb{R}} |\widehat{a}(\lambda)\widehat{b}(\lambda)|^2 |\widehat{a}(\lambda)|^2 d\lambda \\
&= 8\pi^3 \int_{\mathbb{R}} f(\lambda)^2 g(\lambda)^2 d\lambda,
\end{aligned}
$$

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

$$
\begin{aligned}
\int_{\mathbb{R}} K_B(v)dv &= \int_{\mathbb{R}} dv \int_{\mathbb{R}} dx \Big((a*b)\cdot a\Big)(x)\Big((a*b)\cdot a\Big)(x+v) \\
&= \left(\int_{\mathbb{R}} (a*b)(x)a(x)dx\right)^2 = \frac{1}{4\pi^2} \left(\int_{\mathbb{R}} \widehat{a}(x)\widehat{b}(x)\overline{\widehat{a}(x)}d\lambda\right)^2 \\
&= 4\pi^2 \left[\int_{\mathbb{R}} f(\lambda)g(\lambda)d\lambda\right]^2.
\end{aligned}
$$

So an alternative expression for σ^2 in (5.7) is

$$
\sigma^2 = 16\pi^3 \int_{\mathbb{R}} f(\lambda)^2 g(\lambda)^2 d\lambda + \kappa_4 \left[2\pi \int_{\mathbb{R}} f(\lambda)g(\lambda)d\lambda\right]^2, \qquad (5.11)
$$

which should be compared with Avram et al. [7] (Theorem 4.1), and formula (3.15) for an analogous expression in the d.t. case.

Remark 5.4. The d.t. analog of Theorem 5.1 with $t = 1$ and ξ being Gaussian was established in Giraitis and Surgailis [60]. A special case of Theorem 5.1 with

$t = 1$ and ξ being Gaussian was established in Ginovian [45], and in Ginovian and Sahakyan [53]. Theorem 5.1 for Wiener-driven model ($\kappa_4 = 0$) was proved in Bai et al. [8].

Remark 5.5. For Lévy-driven model with $t = 1$ and σ^2 given by (5.11), a version of Theorem 5.1 was stated in Avram et al. [7] (Theorem 4.1). They impose (L^p, L^q)-type 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. [7] is unfortunately omitted, and it is not clear 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 [84], Chapter 7).

The following corollary contains sufficient conditions for the assumptions in Theorem 5.1 to hold. These conditions involve bounds on the tails of functions $a(\cdot)$ and $b(\cdot)$ by suitable power functions (see Bai et al. [9]).

Corollary 5.1. *The convergence in (5.6) holds if the functions $a(\cdot)$ and $b(\cdot)$ satisfy the following conditions:*

$$a(\cdot), \ b(\cdot) \in L^\infty(\mathbb{R}), \quad |a(x)| \le c|x|^{\alpha/2 - 1}, \quad |b(x)| \le c|x|^{\beta - 1} \tag{5.12}$$

with

$$0 < \alpha, \beta < 1, \quad \alpha + \beta < 1/2.$$

5.2. Non-central limit theorems

We now state the non-central limit theorems. We make the following assumptions on the functions $a(\cdot)$ and $b(\cdot)$, and on their Fourier transforms $\widehat{a}(\cdot)$ and $\widehat{b}(\cdot)$.

Assumption 5.1. 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 $\lim_{x \to 0^+} A(x) = A_0$ for some A_0 on the complex unit circle (since $\widehat{a}(-x) = \overline{\widehat{a}(x)}$, we also have $\lim_{x \to 0^-} A(x) = \overline{A_0}$).

Assumption 5.2. 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 5.3. The parameters α and β above satisfy

$$-1/2 < \alpha < 1, \quad -1/2 < \beta < 1, \quad \alpha + \beta > 1/2. \tag{5.13}$$

Assumption 5.4. There exist numbers α^* and β^* satisfying

$$0 < \alpha^*, \beta^* < 1 \quad 1 < \alpha^* + \beta^* < \alpha + \beta + 1/2,$$

such that

$$|a(x)| \le C|x|^{\alpha^*/2-1}, \quad |b(x)| \le C|x|^{\beta^*-1}.$$

The proof of the following theorem can be found in Bai et al. [9].

Theorem 5.2. *Suppose that Assumptions 5.1 - 5.4 hold. Then as $T \to \infty$ we have*

$$Z_T(\tau) := \frac{1}{T^{\alpha+\beta}L_1(1/T)L_2(1/T)} \left(Q_T(\tau) - \mathbb{E}Q_T(\tau)\right) \xrightarrow{f.d.d.} Z_{\alpha,\beta}(\tau), \tag{5.14}$$

where

$$Z_{\alpha,\beta}(\tau) = \frac{1}{2\pi} \int_{\mathbb{R}^2}'' |x_1 x_2|^{-\alpha/2} \int_{\mathbb{R}} \frac{e^{i\tau(x_1+u)} - 1}{i(x_1+u)} \frac{e^{i\tau(x_2-u)} - 1}{i(x_2-u)} |u|^{-\beta} du$$
$$\times W(dx_1)W(dx_2), \tag{5.15}$$

where $W(\cdot)$ is a complex-valued Brownian motion.

Remark 5.6. The regular variation conditions on $\widehat{a}(\cdot)$ and $\widehat{b}(\cdot)$ in Assumptions 5.1 - 5.3 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 $L_1(\cdot)$ and $L_2(\cdot)$ of $a(\cdot)$ and $b(\cdot)$. For instance, it will hold if we have (see Bingham et al. [13], formula (4.3.7))

$$a(x) = x^{\alpha/2-1}\ell_1(x)1_{[0,\infty)}(x), \quad b(x) = |x|^{\beta-1}\ell_2(x), \tag{5.16}$$

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. [13], Section 2.7):

$$\int_0^x t^\delta |d\ell(t)| = O(x^\delta l(x)) \quad \text{as} \quad x \to \infty.$$

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. [13]).

Notice also that Assumption 5.4 will be satisfied if (5.16) holds.

Remark 5.7. Let the functions $a(\cdot)$ and $b(\cdot)$ be as in (5.16) with $\alpha < 0$ or $\beta < 0$ (by (5.13) 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 [92]). Similar considerations hold if $\beta < 0$ and $\alpha > 0$ instead.

Remark 5.8. Note that Assumption 5.1 holds with $\alpha = 0$ if $a(\cdot) \in L^1(\mathbb{R})$ and $\int_0^\infty a(x) \neq 0$, and Assumption 5.2 holds with $\beta = 0$ if $b(\cdot) \in L^1(\mathbb{R})$ and $\int_0^\infty b(x) \neq 0$.

The next theorem contains time-domain representations for the limiting process $Z_{\alpha,\beta}(\tau)$ in (5.15) in the case $\alpha, \beta \geq 0$ (see Bai et al. [9], Theorem 2.3).

Theorem 5.3. *The limiting process $Z_{\alpha,\beta}(\tau)$ in (5.15) admits the following time-domain representations:*

(a) when $\alpha > 0$, $\beta > 0$:

$$Z_{\alpha,\beta}(\tau) \stackrel{f.d.d.}{=} c_{\alpha,\beta} \int_{\mathbb{R}^2}' \int_0^\tau \int_0^\tau |u - v|^{\beta-1}(u - x_1)_+^{\alpha/2-1}(v - x_2)_+^{\alpha/2-1} du dv$$
$$\times B(dx_1)B(dx_2), \tag{5.17}$$

where

$$c_{\alpha,\beta} = \frac{\Gamma(1 - \beta)\sin(\beta\pi/2)}{\pi\Gamma(\alpha/2)^2};$$

(b) when $\alpha > 1/2$, $\beta = 0$:

$$Z_{\alpha,\beta}(\tau) \stackrel{f.d.d.}{=} c_\alpha \int_{\mathbb{R}^2}' \int_0^\tau (u-x_1)_+^{\alpha/2-1}(u-x_2)_+^{\alpha/2-1} du \, B(dx_1)B(dx_2), \tag{5.18}$$

where

$$c_\alpha = \frac{\sin(\alpha\pi/2)\Gamma(1 - \alpha/2)}{\pi\Gamma(\alpha/2)};$$

(c) when $\alpha = 0$, $\beta > 1/2$:

$$Z_{\alpha,\beta}(\tau) \stackrel{f.d.d.}{=} c_\beta \int_{[0,\tau]^2}' |x_1 - x_2|^{\beta-1} \, B(dx_1)B(dx_2), \tag{5.19}$$

where

$$c_\beta = \frac{\Gamma(1 - \beta)\sin(\beta\pi/2)}{\pi}.$$

Here $B(\cdot)$ is the real Brownian random measure and the prime $'$ in the integrals indicates the exclusion of the diagonals (for the definition of the integral $\int'$ see Section 8.5, formula (8.24)).

Remark 5.9. In view of (5.1) and (5.16), the representation (5.17) gives an explicit insight of the convergence in Theorem 5.2 (see Theorem 5.4 below). The process in (5.18) is known as Rosenblatt process (see Taqqu [97]), and

the corresponding convergence in Theorem 5.2 is the c.t. analog of the d.t. case considered in Fox and Taqqu [37]. The representation (5.19) 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 (5.1) can be replaced by the white noise $B(dt)$.

Remark 5.10. It is of interest to obtain appropriate elementary expressions for the time-domain representation of the limiting process $Z_{\alpha,\beta}(\tau)$, similar to (5.17) - (5.19), in the cases where either α or β satisfying (5.13) is negative.

Using the time-domain representation (5.17), 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 $b(\cdot)$ (see Bai et al. [9]).

Theorem 5.4. *Suppose that the functions $a(\cdot)$ and $b(\cdot)$ are given by (5.16), 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 $T \to \infty$ we have*

$$\frac{1}{T^{\alpha+\beta}\ell_1(T)\ell_2(T)} \left(Q_T(\tau) - \mathbb{E}[Q_T(\tau)]\right) \xrightarrow{f.d.d.}$$
$$\int_{\mathbb{R}^2}' \int_0^\tau \int_0^\tau |u-v|^{\beta-1}(u-x_1)_+^{\alpha/2-1}(v-x_2)_+^{\alpha/2-1}dudv \; B(dx_1)B(dx_2).$$

6. CLT for tapered Toeplitz quadratic functionals

6.1. The problem

In this section we consider a question concerning asymptotic distribution (as $T \to \infty$) of the following *tapered* Toeplitz type quadratic functionals of the centered stationary process $X(u)$, $u \in \mathbb{U}$, with spectral density $f(\lambda)$, $\lambda \in \Lambda$:

$$Q_T^h := \begin{cases} \sum_{t=1}^T \sum_{s=1}^T \widehat{g}(t-s)h_T(t)h_T(s)X(t)X(s) & \text{in the d.t. case,} \\ \\ \int_0^T \int_0^T \widehat{g}(t-s)h_T(t)h_T(s)X(t)X(s)\,dt\,ds & \text{in the c.t. case,} \end{cases} \tag{6.1}$$

where $\widehat{g}(t)$ is the Fourier transform of some integrable even function $g(\lambda)$ and

$$h_T(t) := h(t/T) \tag{6.2}$$

with $h(t)$, $t \in \mathbb{R}$ being a *taper function* to be specified below.
Note. The case $h(t) = \mathbb{I}_{[0,1]}(t)$, where $\mathbb{I}_{[0,1]}(\cdot)$ denotes the indicator of the segment $[0,1]$, will be referred to as the *non-tapered case*.

The limit distribution of the functional (6.1) is completely determined by the functions f, g and h, and depending on their properties it can be either Gaussian (that is, Q_T^h with an appropriate normalization obey central limit theorem), or non-Gaussian.

We discuss here the case where the limit distribution is Gaussian, and present sufficient conditions in terms of functions f, g (and h) ensuring central limit theorems for a standard normalized tapered quadratic functional Q_T^h.

6.2. Statistical motivation

Much of the statistical inferences (parametric and nonparametric estimation, hypotheses testing) about the spectrum or the covariance of a stationary process $\{X(u),\ u \in \mathbb{U}\}$ is based on the sample:

$$\mathbf{X}_T := \left\{ \begin{array}{ll} \{X(t),\, t = 1, \ldots, T\} & \text{in the d.t. case,} \\ \{X(t),\, 0 \le t \le T\} & \text{in the c.t. case.} \end{array} \right. \tag{6.3}$$

In the statistical analysis of stationary processes, however, the data are frequently tapered before calculating the statistics of interest, and the statistical inference procedure, instead of the original data $\mathbf{X}_T$ given by (6.3), is based on the tapered data $\mathbf{X}_T^h$:

$$\mathbf{X}_T^h := \left\{ \begin{array}{ll} \{h_T(t)X(t),\, t = 1, \ldots, T\} & \text{in the d.t. case,} \\ \{h_T(t)X(t),\, 0 \le t \le T\} & \text{in the c.t. case,} \end{array} \right. \tag{6.4}$$

with a taper function $h(t),\ t \in \mathbb{R}$.

The benefits of tapering the data have been widely reported in the literature (see, e.g., Bloomfield [14], Brillinger [17], Dahlhaus [24, 25, 26], Dahlhaus and Künsch [28], Guyon [69], and references therein). For example, data-tapers are introduced to reduce the so-called 'leakage effects', that is, to obtain better estimation of the spectrum of the model in the case where it contains high peaks. Other application of data-tapers is in situations in which some of the data values are missing. Also, the use of tapers leads to bias reduction, which is especially important when dealing with spatial data. In this case, the tapers can be used to fight the so-called 'edge effects'.

Quadratic functionals of the form (6.1) appear both in nonparametric and parametric estimation of the spectrum of the process $X(t)$ based on the tapered data (6.4). For instance, when we are interested in nonparametric estimation of a linear integral functional in $L^p(\Lambda)$, $p > 1$ of the form:

$$J = J(f) := \int_\Lambda f(\lambda)g(\lambda)d\lambda, \tag{6.5}$$

where $g(\lambda) \in L^q(\Lambda)$, $1/p + 1/q = 1$, then a natural statistical estimator for $J(f)$ is the linear integral functional of the *empirical spectral density (periodogram)* of the process $X(t)$. To define this estimator, we first introduce some notation.

For $k \in \mathbb{N} := \{1, 2, \ldots\}$, denote by $H_{k,T}(\lambda)$ the *tapered Dirichlet type kernel*, defined by

$$H_{k,T}(\lambda) := \left\{ \begin{array}{ll} \sum_{t=1}^T h_T^k(t)e^{-i\lambda t} & \text{in the d.t. case,} \\[2ex] \int_0^T h_T^k(t)e^{-i\lambda t}dt & \text{in the c.t. case,} \end{array} \right. \tag{6.6}$$

and put

$$H_{k,T} := H_{k,T}(0). \tag{6.7}$$

Define the finite Fourier transform of the tapered data (6.4):

$$d_T^h(\lambda) := \begin{cases} \sum_{t=0}^T h_T(t)X(t)e^{-i\lambda t} & \text{in the d.t. case,} \\ \\ \int_0^T h_T(t)X(t)e^{-i\lambda t}dt & \text{in the c.t. case,} \end{cases} \tag{6.8}$$

and the tapered periodogram $I_T^h(\lambda)$ of the process $X(t)$:

$$\begin{aligned} I_T^h(\lambda) : \quad &= \quad \frac{1}{C_T}\, d_T^h(\lambda)d_T^h(-\lambda) \\ &= \quad \begin{cases} \frac{1}{C_T}\left|\sum_{t=1}^T h_T(t)X(t)e^{-i\lambda t}\right|^2 & \text{in the d.t. case,} \\ \\ \frac{1}{C_T}\left|\int_0^T h_T(t)X(t)e^{-i\lambda t}dt\right|^2 & \text{in the c.t. case.} \end{cases} \end{aligned} \tag{6.9}$$

where

$$C_T := 2\pi H_{2,T}(0) \neq 0. \tag{6.10}$$

Notice that for non-tapered case $(h(t) = \mathbb{I}_{[0,1]}(t))$, we have $C_T = 2\pi T$.

As an estimator J_T^h for functional $J(f)$, given by (6.5), based on the tapered data (6.4), we consider the averaged tapered periodogram (or a simple 'plug-in' statistic), defined by

$$J_T^h = J(I_T^h, g) := \int_\Lambda I_T^h(\lambda)g(\lambda)d\lambda. \tag{6.11}$$

In view of (6.1), (6.9) and (6.11) we have

$$J_T^h = C_T^{-1}Q_T^h, \tag{6.12}$$

where C_T is as in (6.10).

Thus, to study the asymptotic properties of the estimator J_T^h, we have to study the asymptotic distribution (as $T \to \infty$) of the tapered Toeplitz type quadratic functional Q_T^h given by (6.1) (for details see Section 7).

Throughout the paper, we will assume that the taper function $h(\cdot)$ satisfies the following assumption.

Assumption 6.1. The taper $h : \mathbb{R} \to \mathbb{R}$ is a continuous nonnegative function of bounded variation and of bounded support $[0, 1]$, such that $H_k \neq 0$, where

$$H_k := \lim_{T \to \infty} (1/T)H_{k,T}, \ \ and \ H_{k,T} \ is \ as \ in \ (6.7). \tag{6.13}$$

Observe that in the c.t. case we have $H_k = \int_0^1 h^k(t)dt$.

Remark 6.1. In the d.t. case, an example of a taper function $h(t)$ satisfying Assumption 6.1 is the Tukey-Hanning taper function $h(t) = 0.5(1 - \cos(\pi t))$ for $t \in [0, 1]$. In the c.t. case, a simple example of a taper function $h(t)$ satisfying Assumption 6.1 is the function $h(t) = 1 - t$ for $t \in [0, 1]$ (see, e.g., Anh et al. [5]).

6.3. Central limit theorems for tapered quadratic functional Q_T^h

We will use the following notation. By $\widetilde{Q}_T^h$ we denote the standard normalized quadratic functional:

$$\widetilde{Q}_T^h = T^{-1/2} \left(Q_T^h - \mathbb{E}[Q_T^h] \right). \tag{6.14}$$

Also, we set

$$\sigma_h^2 := 16\pi^3 H_4 \int_\Lambda f^2(\lambda) g^2(\lambda) \, d\lambda, \tag{6.15}$$

where H_4 is as in (6.13). The notation

$$\widetilde{Q}_T^h \xrightarrow{d} \eta \sim N(0, \sigma_h^2) \quad \text{as} \quad T \to \infty \tag{6.16}$$

will mean that the distribution of the random variable $\widetilde{Q}_T^h$ tends (as $T \to \infty$) to the centered normal distribution with variance σ_h^2.

Let $\psi(\lambda)$ be an integrable real symmetric function defined on $[-\pi, \pi]$, and let $h(t)$, $t \in [0, 1]$ be a taper function. For $T = 1, 2, \ldots$, the $(T \times T)$-*truncated tapered Toeplitz matrix* generated by ψ and h, denoted by $B_T^h(\psi)$, is defined by the following equation (see (3.4) for non-tapered case):

$$B_T^h(\psi) := \|\widehat{\psi}(t - s) h_T(t) h_T(s)\|_{t,s=1,2\ldots,T}, \tag{6.17}$$

where $\widehat{\psi}(t)$ $(t \in \mathbb{Z})$ are the Fourier coefficients of ψ.

Given a real number $T > 0$ and an integrable real symmetric function $\psi(\lambda)$ defined on $\mathbb{R}$, the T-*truncated tapered Toeplitz operator* (also called *tapered Wiener-Hopf operator*) generated by ψ and a taper function h, denoted by $W_T^h(\psi)$ is defined as follows (see (3.5) for non-tapered case):

$$[W_T^h(\psi)u](t) = \int_0^T \hat{\psi}(t - s) u(s) h_T(s) ds, \quad u(s) \in L^2([0, T]; h_T), \tag{6.18}$$

where $\hat{\psi}(\cdot)$ is the Fourier transform of $\psi(\cdot)$, and $L^2([0, T]; h_T)$ denotes the weighted L^2-space with respect to the measure $h_T(t) dt$.

Let $A_T^h(f)$ be either the $T \times T$ tapered Toeplitz matrix $B_T^h(f)$, or the T-truncated tapered Toeplitz operator $W_T^h(f)$ generated by the spectral density f and taper h, and let $A_T^h(g)$ denote either the $T \times T$ tapered Toeplitz matrix, or the T-truncated tapered Toeplitz operator generated by the functions g and h.

6.3.1. CLT for Gaussian models

The theorems that follow extend the results of Theorems 3.1 and 3.3 to the tapered case. We assume that the model process $X(t)$ is Gaussian, and with no loss of generality, that $g \geq 0$. The following theorems were proved in Ginovyan and Sahakyan [57].

Theorem 6.1. *Assume that $f \cdot g \in L^1(\Lambda) \cap L^2(\Lambda)$, the taper function h satisfies Assumption 6.1, and for $T \to \infty$*

$$\chi_2(\widetilde{Q}_T^h) = \frac{2}{T} tr\big[A_T^h(f)A_T^h(g)\big]^2 \longrightarrow \sigma_h^2, \qquad (6.19)$$

where σ_h^2 is as in (6.15). Then $\widetilde{Q}_T^h \overset{d}{\to} \eta \sim N(0, \sigma_h^2)$ as $T \to \infty$.

Theorem 6.2. *Assume that the function*

$$\varphi(x_1, x_2, x_3) = \int_\Lambda f(u)g(u - x_1)f(u - x_2)g(u - x_3)\, du \qquad (6.20)$$

belongs to $L^2(\Lambda^3)$ and is continuous at $(0, 0, 0)$, and the taper function h satisfies Assumption 6.1. Then $\widetilde{Q}_T^h \overset{d}{\to} \eta \sim N(0, \sigma_h^2)$ as $T \to \infty$.

Theorem 6.3. *Assume that $f(\lambda) \in L^p(\Lambda)$ $(p \geq 1)$ and $g(\lambda) \in L^q(\Lambda)$ $(q \geq 1)$ with $1/p + 1/q \leq 1/2$, and the taper function h satisfies Assumption 6.1. Then $\widetilde{Q}_T^h \overset{d}{\to} \eta \sim N(0, \sigma_h^2)$ as $T \to \infty$.*

Theorem 6.4. *Let $f \in L^2(\Lambda)$, $g \in L^2(\Lambda)$, $fg \in L^2(\Lambda)$,*

$$\int_\Lambda f^2(\lambda)g^2(\lambda - \mu)\, d\lambda \longrightarrow \int_\Lambda f^2(\lambda)g^2(\lambda)\, d\lambda \quad \text{as} \quad \mu \to 0, \qquad (6.21)$$

and let the taper function h satisfy Assumption 6.1. Then $\widetilde{Q}_T^h \overset{d}{\to} \eta \sim N(0, \sigma_h^2)$ as $T \to \infty$.

To state the next theorem, we recall the class $SV_0(\mathbb{R})$ of slowly varying functions at zero $u(\lambda)$, $\lambda \in \mathbb{R}$, satisfying the following conditions: for some $a > 0$, $u(\lambda)$ is bounded on $[-a, a]$, $\lim_{\lambda \to 0} u(\lambda) = 0$, $u(\lambda) = u(-\lambda)$ and $0 < u(\lambda) < u(\mu)$ for $0 < \lambda < \mu < a$.

Theorem 6.5. *Assume that the functions f and g are integrable on $\mathbb{R}$ and bounded outside any neighborhood of the origin, and satisfy for some $a > 0$*

$$f(\lambda) \leq |\lambda|^{-\alpha} L_1(\lambda), \quad |g(\lambda)| \leq |\lambda|^{-\beta} L_2(\lambda), \quad \lambda \in [-a, a], \qquad (6.22)$$

for some $\alpha < 1$, $\beta < 1$ with $\alpha + \beta \leq 1/2$, where $L_1(x)$ and $L_2(x)$ are slowly varying functions at zero satisfying

$$L_i \in SV_0(\mathbb{R}), \quad \lambda^{-(\alpha+\beta)} L_i(\lambda) \in L^2[-a, a], \ i = 1, 2. \qquad (6.23)$$

Also, let the taper function h satisfy Assumption 6.1. Then $\widetilde{Q}_T^h \overset{d}{\to} \eta \sim N(0, \sigma_h^2)$ as $T \to \infty$.

As in Remark 3.3, the conditions $\alpha < 1$ and $\beta < 1$ in Theorem 6.5 ensure that the Fourier transforms of f and g are well defined. Observe that when $\alpha > 0$ the process $X(t)$ may exhibit long-range dependence. We also allow here $\alpha + \beta$ to assume the critical value $1/2$. The assumptions $f \cdot g \in L^1(\Lambda)$, $f, g \in L^\infty(\Lambda \setminus [-a, a])$ and (6.23) imply that $f \cdot g \in L^2(\Lambda)$, so that the variance σ_h^2 in (6.15) is finite.

6.3.2. CLT for Lévy-driven stationary linear models

Now we assume that the underlying model $X(t)$ is a Lévy-driven stationary linear process defined by (2.12), where $a(\cdot)$ is a filter from $L^2(\mathbb{R})$, and $\xi(t)$ is a Lévy process satisfying the conditions: $\mathbb{E}\xi(t) = 0$, $\mathbb{E}\xi^2(1) = 1$ and $\mathbb{E}\xi^4(1) < \infty$.

The central limit theorem that follow is a tapered counterpart of Theorem 5.1 and was proved in Ginovyan and Sahakyan [56].

Theorem 6.6. *Assume that the filter $a(\cdot)$ and the generating kernel $\widehat{g}(\cdot)$ are such that*

$$a(\cdot) \in L^p(\mathbb{R}) \cap L^2(\mathbb{R}), \quad \widehat{g}(\cdot) \in L^q(\mathbb{R}), \quad 1 \le p, q \le 2, \quad \frac{2}{p} + \frac{1}{q} \ge \frac{5}{2}, \quad (6.24)$$

and the taper h satisfies Assumption 6.1. Then $\widetilde{Q}_T^h \xrightarrow{d} \eta \sim N(0, \sigma_{L,h}^2)$ as $T \to \infty$, where

$$\sigma_{L,h}^2 = 16\pi^3 H_4 \int_{\mathbb{R}} f^2(\lambda) g^2(\lambda) d\lambda + \kappa_4 4\pi^2 H_4 \left[\int_{\mathbb{R}} f(\lambda) g(\lambda) d\lambda \right]^2, \quad (6.25)$$

and H_4 is as in (6.13).

Remark 6.2. Notice that if the underlying process $X(t)$ is Gaussian, then in formula (6.25) we have only the first term and so $\sigma_{L,h}^2 = \sigma_h^2$ (see (6.15)), because in this case $\kappa_4 = 0$. On the other hand, the condition (6.24) is more restrictive than the conditions in Theorems 6.1 - 6.5. Thus, for Gaussian processes Theorems 6.1 - 6.5 improve Theorem 6.6.

Remark 6.3. Central and non-central limit theorems for tapered quadratic forms of a d.t. long memory Gaussian stationary fields have been proved in Doukhan et al. [32].

7. Applications

7.1. Nonparametric estimation of spectral functionals

Suppose we observe a realization $\mathbf{X}_T := \{X(u), 0 \le u \le T$ (or $u = 1, \ldots, T$ in the d.t. case)$\}$ of a centered stationary process $X(t)$ with an *unknown* spectral density function $f(\lambda)$, $\lambda \in \Lambda$. We assume that $f(\lambda)$ belongs to a given (infinite-dimensional) class $\boldsymbol{\Theta} \subset L^p(\Lambda)$ $(p \ge 1)$ of spectral densities possessing some smoothness properties. Let $\Phi(\cdot)$ be some *known* functional, the domain of definition of which contains $\boldsymbol{\Theta}$. The problem is to estimate the value $\Phi(f)$ of the functional $\Phi(\cdot)$ at an unknown point $f \in \boldsymbol{\Theta}$ on the basis of observation $\mathbf{X}_T$, and investigate the asymptotic (as $T \to \infty$) properties of the suggested estimators.

This problem for linear and some nonlinear smooth functionals for d.t. and c.t. stationary models has been extensively discussed in the literature (see, e.g., Dahlhaus and Wefelmeyer [29], Ginovyan [42, 43, 46, 48, 49], Ginovyan and

Sahakyan [56, 57], Has'minskii and Ibragimov [70], Taniguchi [95], Taniguchi and Kakizawa [96], and references therein).

In this section we apply the results of Section 6.3 to show that the statistic J_T^h given by (6.11) is a consistent and asymptotically normal estimator for the linear functional $J(f)$ given by (6.5). We follow the papers Ginovyan and Sahakyan [56, 57]. To state the corresponding results, we first introduce the L^p-Hölder class and set up a set of assumptions.

Given numbers $p \geq 1$, $0 < \alpha < 1$, $r \in \mathbb{N}_0 := \mathbb{N} \cup \{0\}$, we set $\beta = \alpha + r$ and denote by $H_p(\beta)$ the L^p-Hölder class, that is, the class of those functions $\psi(\lambda) \in L^p(\Lambda)$, which have r-th derivatives in $L^p(\Lambda)$ and with some positive constant C satisfy

$$||\psi^{(r)}(\cdot + \lambda) - \psi^{(r)}(\cdot)||_p \leq C|\lambda|^\alpha.$$

Now we list the assumptions.

Assumption 7.1. The filter $a(\cdot)$ and the generating kernel $\widehat{g}(\cdot)$ are such that

$$a(\cdot) \in L^p(\Lambda) \cap L^2(\Lambda), \quad \widehat{g}(\cdot) \in L^q(\Lambda) \tag{7.1}$$

with

$$1 \leq p, q \leq 2, \quad \frac{2}{p} + \frac{1}{q} \geq \frac{5}{2}. \tag{7.2}$$

Assumption 7.2. The spectral density f and the generating function g are such that $f, g \in L^1(\Lambda) \cap L^2(\Lambda)$ and g is of bounded variation.

Assumption 7.3. The spectral density f and the generating function g are such that $f \in H_p(\beta_1)$, $\beta_1 > 0$, $p \geq 1$ and $g(\lambda) \in H_q(\beta_2)$, $\beta_2 > 0$, $q \geq 1$ with $1/p + 1/q = 1$, and one of the conditions a)–d) is satisfied:
 a) $\beta_1 > 1/p$, $\beta_2 > 1/q$
 b) $\beta_1 \leq 1/p$, $\beta_2 \leq 1/q$ and $\beta_1 + \beta_2 > 1/2$
 c) $\beta_1 > 1/p$, $1/q - 1/2 < \beta_2 \leq 1/q$
 d) $\beta_2 > 1/q$, $1/p - 1/2 < \beta_1 \leq 1/p$.

Remark 7.1. In Ginovian [45] it was proved that if Assumption 7.3 is satisfied, then there exist numbers p_1 $(p_1 > p)$ and q_1 $(q_1 > q)$, such that $H_p(\beta_1) \subset L_{p_1}$, $H_q(\beta_2) \subset L_{q_1}$ and $1/p_1 + 1/q_1 \leq 1/2$.

The next theorem controls the bias $E(J_T^h) - J$ and provides sufficient conditions assuring the proper rate of convergence of the bias to zero, which is necessary to obtain the asymptotic normality of the estimator J_T^h. Specifically, we have the following result, which was proved in Ginovyan and Sahakyan [56].

Theorem 7.1. *Let the functionals $J := J(f)$ and $J_T^h := J(I_T^h)$ be defined by (6.5) and (6.11), respectively. Then under Assumptions 7.2 (or 7.3) and 6.1 the following asymptotic relation holds:*

$$T^{1/2} \left[\mathbb{E}(J_T^h) - J \right] \to 0 \quad \text{as} \quad T \to \infty. \tag{7.3}$$

The next theorem, which is an immediate consequence of Theorem 6.6, contains sufficient conditions for functional J_T^h to obey the central limit theorem.

Theorem 7.2. *Let $J := J(f)$ and $J_T^h := J(I_T^h)$ be defined by (6.5) and (6.11), respectively. Then under Assumptions 7.1 and 6.1 the functional J_T^h obeys the central limit theorem. More precisely, we have*

$$T^{1/2}\left[J_T^h - \mathbb{E}(J_T^h)\right] \xrightarrow{d} \eta \quad \text{as} \quad T \to \infty, \tag{7.4}$$

where the symbol $\xrightarrow{d}$ stands for convergence in distribution, and η is a normally distributed random variable with mean zero and variance $\sigma_h^2(J)$ given by

$$\sigma_h^2(J) = 4\pi e(h)\int_{\mathbb{R}} f^2(\lambda)g^2(\lambda)d\lambda + \kappa_4 e(h)\left[\int_{\mathbb{R}} f(\lambda)g(\lambda)d\lambda\right]^2. \tag{7.5}$$

Here κ_4 is the fourth cumulant of $\xi(1)$, and

$$e(h) := \lim_{T\to\infty} \frac{TH_{4,T}}{H_{2,T}^2}. \tag{7.6}$$

Observe that for c.t. case we have

$$e(h) = \lim_{T\to\infty} \frac{TH_{4,T}}{H_{2,T}^2} = \frac{H_4}{H_2^2} = \int_0^1 h^4(t)dt \left(\int_0^1 h^2(t)dt\right)^{-2}.$$

Taking into account the equality

$$T^{1/2}\left[J_T^h - J\right] = T^{1/2}\left[\mathbb{E}(J_T^h) - J\right] + T^{1/2}\left[J_T^h - \mathbb{E}(J_T^h)\right], \tag{7.7}$$

as an immediate consequence of Theorems 7.1 and 7.2, we obtain the next result that contains sufficient conditions for a simple "plug-in" statistic $J(I_T^h)$ to be an asymptotically normal estimator for a linear spectral functional $J(f)$.

Theorem 7.3. *Let the functionals $J := J(f)$ and $J_T^h := J(I_T^h)$ be defined by (6.5) and (6.11), respectively. Then under Assumptions 7.1, 7.2 (or 7.3) and 6.1 the statistic J_T^h is an asymptotically normal estimator for functional J. More precisely, we have*

$$T^{1/2}\left[J_T^h - J\right] \xrightarrow{d} \eta \quad \text{as} \quad T \to \infty, \tag{7.8}$$

where η is as in Theorem 7.2, that is, η is a normally distributed random variable with mean zero and variance $\sigma_h^2(J)$ given by (7.5) and (7.6).

In the Gaussian case we have more accurate result. The next theorem, which was proved in Ginovyan and Sahakyan [57], states that for Gaussian models Assumptions 7.3 and 6.1 are sufficient for the statistic J_T^h to be an asymptotically normal estimator for functional J.

Theorem 7.4. *Let the functionals $J := J(f)$ and $J_T^h := J(I_T^h)$ be defined by (6.5) and (6.11), respectively. Then under Assumptions 7.3 and 6.1 the statistic J_T^h is an asymptotically normal estimator for functional J. More precisely, we have*

$$T^{1/2}\left[J_T^h - J\right] \xrightarrow{d} \eta \quad \text{as} \quad T \to \infty, \tag{7.9}$$

where η is a normally distributed random variable with mean zero and variance $\sigma_h^2(J)$ given by

$$\sigma_h^2(J) = 4\pi e(h) \int_{\mathbb{R}} f^2(\lambda)g^2(\lambda)d\lambda, \qquad (7.10)$$

and $e(h)$ is as in (7.6).

7.2. Parametric estimation: the Whittle procedure

We assume here that the spectral density $f(\lambda)$ belongs to a given parametric family of spectral densities $\mathcal{F} := \{f(\lambda, \theta) : \theta \in \Theta\}$, where $\theta := (\theta_1, \ldots, \theta_p)$ is an unknown parameter and Θ is an open subset in the Euclidean space $\mathbb{R}^p$. The problem of interest is to estimate the unknown parameter θ on the basis of the tapered data (6.4), and investigate the asymptotic (as $T \to \infty$) properties of the suggested estimators, depending on the dependence (memory) structure of the model $X(t)$ and the smoothness of its spectral density f.

There are different methods of estimation: maximum likelihood, Whittle, minimum contrast, etc. Here we focus on the Whittle method.

The Whittle estimation procedure, originally devised for d.t. short memory stationary processes, is based on the smoothed periodogram analysis on a frequency domain, involving approximation of the likelihood function and asymptotic properties of empirical spectral functionals (see Whittle [107]). The Whittle estimation method since its discovery has played a major role in the asymptotic theory of parametric estimation in the frequency domain, and was the focus of interest of many statisticians. Their aim was to weaken the conditions needed to guarantee the validity of the Whittle approximation for d.t. short memory models, to find analogues for long and intermediate memory models, to find conditions under which the Whittle estimator is asymptotically equivalent to the exact maximum likelihood estimator, and to extend the procedure to the c.t. models and random fields.

For the d.t. case, it was shown that for Gaussian and linear stationary models the Whittle approach leads to consistent and asymptotically normal estimators under short, intermediate and long memory assumptions. Moreover, it was shown that in the Gaussian case the Whittle estimator is also asymptotically efficient in the sense of Fisher (see, e. g., Dahlhaus [27], Dzhaparidze [33], Fox and Taqqu [39], Giraitis and Surgailis [60], Guyon [69], Heyde and Gay [72], Taniguchi and Kakizawa [96], Walker [106], and references therein).

For c.t. models, the Whittle estimation procedure has been considered, for example, in Anh et al. [4], Avram et al. [7], Casas and Gao [21], Dzhaparidze [33], Gao [40], Gao et al. [41], Leonenko and Sakhno [80], Tsai and Chan [104], where can also be found additional references. In this case, it was proved that the Whittle estimator is consistent and asymptotically normal.

The Whittle estimation procedure based on the d.t. tapered data has been studied in Alomari et al. [1], Dahlhaus [24], Dahlhaus and Künsch [28], Guyon [69], Ludeña and Lavielle [82]. In the case where the underlying model is a Lévy-driven c.t. linear process with possibly unbounded or vanishing spectral density

function, consistency and asymptotic normality of the Whittle estimator was established in Ginovyan [50].

To explain the idea behind the Whittle estimation procedure, assume for simplicity that the underlying process $X(t)$ is a d.t. Gaussian process, and we want to estimate the parameter θ based on the sample $X_T := \{X(t), t = 1, \ldots, T\}$. A natural approach is to find the maximum likelihood estimator (MLE) $\widehat{\theta}_{T,MLE}$ of θ, that is, to maximize the log-likelihood function $L_T(\theta)$, which in this case takes the form:

$$L_T(\theta) = -\frac{T}{2} \ln 2\pi - \frac{1}{2} \ln \det B_T(f_\theta) - \frac{1}{2} X_T' [B_T(f_\theta)]^{-1} X_T,$$

where $B_T(f_\theta)$ is the Toeplitz matrix generated by f_θ. Unfortunately, the above function is difficult to handle, and no explicit expression for the estimator $\widehat{\theta}_{T,MLE}$ is known (even in the case of simple models). An approach, suggested by P. Whittle, called the Whittle estimation procedure, is to approximate the term $\ln \det B_T(f_\theta)$ by $\frac{T}{2} \int_{-\pi}^{\pi} \ln f_\theta(\lambda) d\lambda$ and the inverse matrix $[B_T(f_\theta)]^{-1}$ by the Toeplitz matrix $B_T(1/f_\theta)$. This leads to the following approximation of the log-likelihood function $L_T(\theta)$, introduced by P. Whittle [107], and called Whittle functional:

$$L_{T,W}(\theta) = -\frac{1}{4\pi} \int_{-\pi}^{\pi} \left[\ln f_\theta(\lambda) + \frac{I_T(\lambda)}{f_\theta(\lambda)} \right] d\lambda,$$

where $I_T(\lambda)$ is the ordinary periodogram of the process $X(t)$.

Now maximizing the Whittle functional $L_{T,W}(\theta)$ with respect to θ, we get the Whittle estimator $\widehat{\theta}_T$ for θ. It can be shown that if

$$T^{-1/2}(L_T(\theta) - L_{T,W}(\theta)) \to 0 \quad \text{as} \quad n \to \infty \quad \text{in probability},$$

then the MLE $\widehat{\theta}_{T,MLE}$ and the Whittle estimator $\widehat{\theta}_T$ are asymptotically equivalent in the sense that $\widehat{\theta}_T$ also is consistent, asymptotically normal and asymptotically Fisher-efficient (see, e.g., Dzhaparidze [33]).

In the continuous context, the Whittle procedure of estimation of a spectral parameter θ based on the sample $X_T := \{X(t), 0 \leq t \leq T\}$ is to choose the estimator $\widehat{\theta}_T$ to minimize the weighted Whittle functional:

$$U_T(\theta) := \frac{1}{4\pi} \int_{\mathbb{R}} \left[\ln f(\lambda, \theta) + \frac{I_T(\lambda)}{f(\lambda, \theta)} \right] \cdot w(\lambda) \, d\lambda, \tag{7.11}$$

where $I_T(\lambda)$ is the continuous periodogram of $X(t)$, and $w(\lambda)$ is a weight function $(w(-\lambda) = w(\lambda), w(\lambda) \geq 0, w(\lambda) \in L^1(\mathbb{R}))$ for which the integral in (7.11) is well defined. An example of common used weight function is $w(\lambda) = 1/(1 + \lambda^2)$.

The Whittle procedure of estimation of a spectral parameter θ based on the tapered sample (6.4) is to choose the estimator $\widehat{\theta}_{T,h}$ to minimize the weighted tapered Whittle functional:

$$U_{T,h}(\theta) := \frac{1}{4\pi} \int_{\Lambda} \left[\log f(\lambda, \theta) + \frac{I_T^h(\lambda)}{f(\lambda, \theta)} \right] \cdot w(\lambda) \, d\lambda, \tag{7.12}$$

where $I_T^h(\lambda)$ is the tapered periodogram of $X(t)$, given by (6.9), and $w(\lambda)$ is a weight function for which the integral in (7.12) is well defined. Thus,

$$\widehat{\theta}_{T,h} := \operatorname*{Arg\,min}_{\theta \in \Theta} U_{T,h}(\theta). \tag{7.13}$$

Here we follow the paper Ginovyan [50]. To state results involving properties of the Whittle estimator, we first introduce the following set of assumptions.

Assumption 7.4. The true value θ_0 of the parameter θ belongs to a compact set Θ, which is contained in an open set S in the p-dimensional Euclidean space $\mathbb{R}^p$, and $f(\lambda, \theta_1) \neq f(\lambda, \theta_2)$ whenever $\theta_1 \neq \theta_2$ almost everywhere in Λ with respect to the Lebesgue measure.

Assumption 7.5. The functions $f(\lambda, \theta)$, $f^{-1}(\lambda, \theta)$ and $\frac{\partial}{\partial \theta_k} f^{-1}(\lambda, \theta)$, $k = 1, \ldots,$ p, are continuous in (λ, θ).

Assumption 7.6. The functions $f := f(\lambda, \theta)$ and $g := w(\lambda) \frac{\partial}{\partial \theta_k} f^{-1}(\lambda, \theta)$ satisfy Assumptions 7.2 or 7.3 for all $k = 1, \ldots, p$ and $\theta \in \Theta$.

Assumption 7.7. The functions $a := a(\lambda, \theta)$ and $b := \widehat{g}$, where g is as in Assumption 7.6, satisfy Assumption 7.1.

Assumption 7.8. The functions $\frac{\partial^2}{\partial \theta_k \partial \theta_j} f^{-1}(\lambda, \theta)$ and $\frac{\partial^3}{\partial \theta_k \partial \theta_j \partial \theta_j} f^{-1}(\lambda, \theta)$ for $k, j, l = 1, \ldots, p$, are continuous in (λ, θ) for $\lambda \in \Lambda$, $\theta \in N_\delta(\theta_0)$, where $N_\delta(\theta_0) := \{\theta : |\theta - \theta_0| < \delta\}$ is some neighborhood of θ_0.

Assumption 7.9. The matrices

$$W(\theta) := \|w_{ij}(\theta)\|, \ A(\theta) := \|a_{ij}(\theta)\|, \ B(\theta) := \|b_{ij}(\theta)\| \quad i, j = 1, \ldots, p \tag{7.14}$$

are positive definite, where

$$w_{ij}(\theta) = \frac{1}{4\pi} \int_\Lambda \frac{\partial}{\partial \theta_i} \ln f(\lambda, \theta) \frac{\partial}{\partial \theta_j} \ln f(\lambda, \theta) w(\lambda) d\lambda, \tag{7.15}$$

$$a_{ij}(\theta) = \frac{1}{4\pi} \int_\Lambda \frac{\partial}{\partial \theta_i} \ln f(\lambda, \theta) \frac{\partial}{\partial \theta_j} \ln f(\lambda, \theta) w^2(\lambda) d\lambda, \tag{7.16}$$

$$b_{ij}(\theta) = \frac{\kappa_4}{16\pi^2} \int_\Lambda \frac{\partial}{\partial \theta_i} \ln f(\lambda, \theta) w(\lambda) d\lambda \int_\mathbb{R} \frac{\partial}{\partial \theta_j} \ln f(\lambda, \theta) w(\lambda) d\lambda, \tag{7.17}$$

and κ_4 is the fourth cumulant of $\xi(1)$.

Consistency of the Whittle estimator. The next theorem contains sufficient conditions for Whittle estimator to be consistent.

Theorem 7.5. *Let $\widehat{\theta}_{T,h}$ be the Whittle estimator defined by (7.13) and let θ_0 be the true value of parameter θ. Then, under Assumptions 7.4–7.7 and 6.1, the statistic $\widehat{\theta}_{T,h}$ is a consistent estimator for θ, that is, $\widehat{\theta}_{T,h} \to \theta_0$ in probability as $T \to \infty$.*

Asymptotic normality of the Whittle estimator. Having established the consistency of the Whittle estimator $\widehat{\theta}_{T,h}$, we can go on to obtain the limiting distribution of $T^{1/2}\left(\widehat{\theta}_{T,h} - \theta_0\right)$ in the usual way by applying the Taylor's formula, the mean value theorem, and Slutsky's arguments. Specifically we have the following result, showing that under the above assumptions, the Whittle estimator $\widehat{\theta}_{T,h}$ is asymptotically normal.

Theorem 7.6. *Suppose that Assumptions 7.4–7.9 and 6.1 are satisfied. Then the Whittle estimator $\widehat{\theta}_{T,h}$ of an unknown spectral parameter θ based on the tapered data (6.4) is asymptotically normal. More precisely, we have*

$$T^{1/2}\left(\widehat{\theta}_{T,h} - \theta_0\right) \xrightarrow{d} N_p\left(0, e(h)\Gamma(\theta_0)\right) \quad \text{as} \quad T \to \infty, \tag{7.18}$$

where $N_p(\cdot, \cdot)$ denotes the p-dimensional normal law, $\xrightarrow{d}$ stands for convergence in distribution,

$$\Gamma(\theta_0) = W^{-1}(\theta_0)\left(A(\theta_0) + B(\theta_0)\right)W^{-1}(\theta_0), \tag{7.19}$$

where the matrices W, A and B are defined in (7.14)-(7.17), and the tapering factor $e(h)$ is given by formula (7.6).

8. Methods and tools

In this section we briefly discuss the methods and tools, used to prove the central and noncentral limit theorems for Toeplitz type quadratic forms and functionals stated in Sections 3–6, as well as the results stated in Section 7.

As mentioned in the introduction, the most commonly used methods to prove central limit theorems are: (a) the method of characteristic functions. (b) the method of cumulants or moments, (c) the approximation method.

To prove the central limit theorems stated in this paper, depending on the underlying model, was used either the method of cumulants or the approximation method. If the underlying model is Gaussian, then the method of cumulants was applied, otherwise the approximation method was applied, which reduces the quadratic integral form to a single integral form. To prove the non-central limit theorems, was used the spectral representation of the underlying process, the properties of Lévy-Itô-type and Stratonovich-type multiple stochastic integrals, and power counting theorems. Some details of the above methods are described below.

8.1. The characteristic functions and cumulant criteria for the CLT

The characteristic functions criterion for the CLT is based on the fact that convergence in distribution is equivalent to the pointwise convergence of the corresponding characteristic functions. The general cumulant criterion for the CLT is based on the following result (see, e.g., Giraitis et al. [59], Corollary 4.2.1).

Proposition 8.1. *Let the random variables $\{X_T, T \in \mathbb{R}\}$ have all moments finite, and let $\mathbb{E}[X_T] \to 0$, $\mathrm{Var}[X_T] \to \sigma^2$, and $\chi_k(X_T) := \mathrm{Cum}_k(X_T) \to 0$ for all $k = 3, 4, \ldots$ as $T \to \infty$. Then $X_T \overset{d}{\to} X \sim N(0, \sigma^2)$ as $T \to \infty$.*

The characteristic functions and cumulant criteria for the CLT for quadratic functionals is based on the following general result (see Ibragimov [75]).

Let ξ be a Gaussian random variable with values in a separable Hilbert space H. In other words, ξ is a random variable with characteristic functional:

$$\varphi(h) = \exp\left\{m(h) - \frac{1}{2}(Rh, h)\right\}, \quad h \in H, \tag{8.1}$$

where $m(h)$ is a continuous linear functional and the correlation operator R is a self-adjoint completely continuous operator with finite trace. We assume, without loss of generality, that $m(h) \equiv 0$. Let A be some linear self-adjoint bounded operator.

The proof of the following result can be found in Ibragimov [75].

Proposition 8.2. *Let the operators R and A be as above. The quadratic form $(A\xi, \xi)$ has the same distribution as the sum $\sum_{k=1}^{\infty} \lambda_k \xi_k^2$, where ξ_k are independent $N(0, 1)$ Gaussian random variables and λ_k are the eigen-values of the operator $B := RA$.*

Remark 8.1. It can easily be shown that the sets of non-zero eigen-values of the operators RA, AR and $R^{1/2}AR^{1/2}$ coincide, where $R^{1/2}$ is the positive definite square root of R.

As mentioned above, Toeplitz matrices and operators arise naturally in the theory of stationary processes, and serve as tools, to study many topics of the spectral and statistical analysis of d.t. and c.t. stationary processes.

Let $A_T(f)$ denote the covariance matrix (or operator) of the process $\{X(u), u \in \mathbb{U}\}$, that is, $A_T(f)$ is either the $T \times T$ Toeplitz matrix $B_T(f)$, or the T-truncated Toeplitz operator $W_T(f)$ generated by the spectral density f, and let $A_T(g)$ denote either the $T \times T$ Toeplitz matrix, or the T-truncated Toeplitz operator generated by the function g (for definitions see formulas (3.4) and (3.5)).

As a consequence of Proposition 8.2, we have the following result, which gives a link between the distribution of the quadratic functional Q_T in (1.1) and the trace problem for Toeplitz matrices and operators (see, e.g., Ginovyan et al. [58], Grenander and Szegő [67], Ibragimov [75]).

1. The quadratic functional Q_T in (1.1) has the same distribution as the sum $\sum_{k=1}^{\infty} \lambda_{k,T} \xi_k^2$ ($\sum_{k=1}^{T} \lambda_{k,T} \xi_k^2$ in the d.t. case), where $\{\xi_k, k \geq 1\}$ are independent $N(0, 1)$ Gaussian random variables and $\{\lambda_{k,T}, k \geq 1\}$ are the eigenvalues of the operator $A_T(f)A_T(g)$.
2. The characteristic function $\varphi_T(t)$ of Q_T is given by

$$\varphi_T(t) = \prod_{k=1}^{\infty} |1 - 2it\lambda_{k,T}|^{-1/2}. \tag{8.2}$$

3. The k–th order cumulant $\chi_k(\cdot)$ of Q_T is given by

$$\chi_k(Q_T) = 2^{k-1}(k-1)! \sum_{j=1}^{\infty} \lambda_{j,T}^k = 2^{k-1}(k-1)! \operatorname{tr}[A_T(f)A_T(g)]^k, \quad (8.3)$$

where $\operatorname{tr}[A]$ stands for the trace of an operator A.

The product in (8.2) and the sum in (8.3) are over $j = 1, \ldots, T$ in the d.t. case.

The tapered case. To study the asymptotic distribution (as $T \to \infty$) of the tapered functional Q_T^h, given by (6.1), we use the method of cumulants, the frequency-domain approach, and the technique of truncated tapered Toeplitz matrices and operators.

Let $A_T^h(f)$ be either the $T \times T$ tapered Toeplitz matrix $B_T^h(f)$, or the T-truncated tapered Toeplitz operator $W_T^h(f)$ generated by the spectral density f and taper h, and let $A_T^h(g)$ denote either the $T \times T$ tapered Toeplitz matrix, or the T-truncated tapered Toeplitz operator generated by the functions g and h (for definitions see formulas (6.17) and (6.18)). Similar to the non-tapered case, we have the following results (cf. Ginovyan et al. [58], Grenander and Szegő [67], Ibragimov [75]).

1. The quadratic functional Q_T^h in (6.1) has the same distribution as the sum $\sum_{j=1}^{\infty} \lambda_{j,T}\xi_j^2$, where $\{\xi_j, j \geq 1\}$ are independent $N(0,1)$ Gaussian random variables and $\{\lambda_{j,T}, j \geq 1\}$ are the eigenvalues of the operator $A_T^h(f)\,A_T^h(g)$.
2. The characteristic function $\varphi(t)$ of Q_T^h is given by formula:

$$\varphi(t) = \prod_{j=1}^{\infty} |1 - 2it\lambda_{j,T}|^{-1/2}. \qquad (8.4)$$

3. The k–th order cumulant $\chi_k(Q_T^h)$ of Q_T^h is given by formula:

$$\chi_k(Q_T) = 2^{k-1}(k-1)! \sum_{j=1}^{\infty} \lambda_{j,T}^k = 2^{k-1}(k-1)! \operatorname{tr}[A_T^h(f)\,A_T^h(g)]^k. \quad (8.5)$$

Thus, to describe the asymptotic distribution of the quadratic functional Q_T^h, we have to control the traces and eigenvalues of the products of truncated tapered Toeplitz operators and matrices.

8.2. Approximation of traces of products of Toeplitz matrices and operators

The trace approximation problem for truncated Toeplitz operators and matrices has been discussed in detail in the survey paper Ginovyan et al. [58] for non-tapered case. Here we present some important results, which were used to prove CLT for quadratic functionals Q_T^h.

Let h be a taper function satisfying Assumption 6.1. Let $A_T^h(\psi)$ be either the $T \times T$ tapered Toeplitz matrix $B_T^h(\psi)$, or the T-truncated tapered Toeplitz operator $W_T^h(\psi)$ generated by a function ψ (for definitions see formulas (6.17) and (6.18))

Observe that, in view of (6.10), (6.13), (6.17) and (6.18), we have

$$\frac{1}{T}\operatorname{tr}\left[A_T^h(\psi)\right] = \frac{1}{T} \cdot \widehat{\psi}(0) \cdot \int_0^T h_T^2(t)dt = 2\pi H_2 \int_\Lambda \psi(\lambda)d\lambda. \tag{8.6}$$

What happens to the relation (8.6) when $A_T^h(\psi)$ is replaced by a product of Toeplitz matrices (or operators)? Observe that the product of Toeplitz matrices (resp. operators) is not a Toeplitz matrix (resp. operator).

The idea is to approximate the trace of the product of Toeplitz matrices (resp. operators) by the trace of a Toeplitz matrix (resp. operator) generated by the product of the corresponding generating functions. More precisely, let $\{\psi_1, \psi_2, \dots, \psi_m\}$ be a collection of integrable real symmetric functions defined on Λ. Let $A_T^h(\psi_i)$ be either the $T \times T$ tapered Toeplitz matrix $B_T^h(\psi_i)$, or the T-truncated tapered Toeplitz operator $W_T^h(\psi_i)$ generated by a function ψ_i and a taper function h. Define

$$S_{A,\mathcal{H},h}(T) := \frac{1}{T}\operatorname{tr}\left[\prod_{i=1}^m A_T^h(\psi_i)\right], \quad M_{\Lambda,\mathcal{H},h} := (2\pi)^{m-1} H_m \int_\Lambda \left[\prod_{i=1}^m \psi_i(\lambda)\right] d\lambda,$$

where H_m is as in (6.13), and let

$$\Delta(T) := \Delta_{A,\Lambda,\mathcal{H},h}(T) = |S_{A,\mathcal{H},h}(T) - M_{\Lambda,\mathcal{H},h}|. \tag{8.7}$$

Proposition 8.3. *Let $\Delta(T)$ be as in (8.7). Each of the following conditions is sufficient for*

$$\Delta(T) = o(1) \quad \text{as} \quad T \to \infty. \tag{8.8}$$

(C1) $\psi_i \in L^1(\Lambda) \cap L^{p_i}(\Lambda)$, $p_i > 1$, $i = 1, 2, \dots, m$, *with* $1/p_1 + \cdots + 1/p_m \le 1$.
(C2) *The function $\varphi(\mathbf{u})$ defined by*

$$\varphi(\mathbf{u}) := \int_\Lambda \psi_1(\lambda)\psi_2(\lambda - u_1)\psi_3(\lambda - u_2)\cdots\psi_m(\lambda - u_{m-1})\, d\lambda, \tag{8.9}$$

where $\mathbf{u} = (u_1, u_2, \dots, u_{m-1}) \in \Lambda^{m-1}$, belongs to $L^{m-2}(\Lambda^{m-1})$ and is continuous at $\mathbf{0} = (0, 0, \dots, 0) \in \Lambda^{m-1}$.

Remark 8.2. In the nontapered case, Proposition 8.3 was proved in Ginovyan et al. [58], in the tapered case, it was proved in Ginovyan [51]. Proposition 8.3 was used to prove parts (B) and (C) of Theorem 3.1 (in the nontapered case) and Theorems 6.2 and 6.3 (in the tapered case). In the special case $m = 4$, $\psi_1 = \psi_3 := f$ and $\psi_2 = \psi_4 := g$, in Ginovyan and Sahakyan [52] (in the d.t. case) and in Ginovyan and Sahakyan [53] (in the c.t. case) it was proved that the conditions of Theorem 3.1(D) and Theorem 3.3 are also sufficient for (8.8).

Remark 8.3. More results concerning the trace approximation problem for truncated Toeplitz operators and matrices can be found in Ginovyan and Sahakyan [54, 55], Ginovyan et al. [58] and Lieberman and Phillips [81].

8.3. Approximation method for the CLT

The approximation method for CLT for quadratic functionals is based on approximation of quadratic functional Q_T by a univariate sum (integral) of m-dependent random variables, and then using the CLT for such variables (see, e.g., Giraitis et al. [59], Section 4.5, and Giraitis and Surgailis [60]).

Let $X_i(u)$ $(i = 1, 2)$ be two linear processes of the form (2.9), subordinated to the same orthonormal sequence $\{\xi(k), \, k \in \mathbb{Z}\}$, with square summable covariance functions $r_i(u)$ $(i = 1, 2)$, that is,

$$X_i(u) = \sum_{k=-\infty}^{\infty} a_i(u - k)\xi(k), \qquad \sum_{k=-\infty}^{\infty} |a_i(k)|^2 < \infty, \quad i = 1, 2, \qquad (8.10)$$

and

$$\sum_{u=-\infty}^{\infty} |r_i(u)|^2 < \infty, \quad i = 1, 2. \qquad (8.11)$$

Denote

$$S_T := \sum_{u=1}^{T} X_1(u)X_2(u). \qquad (8.12)$$

Proposition 8.4. *Let $X_i(u)$ $(i = 1, 2)$ and S_T be as above, and let the quadratic form Q_T be as in (1.1). Then the following assertions hold.*

(a) *The distribution of $T^{-1/2}(S_T - \mathbb{E}[S_T])$ tends to the centered normal distribution with variance:*

$$\sigma^2 := \sum_{u=-\infty}^{\infty} r_1(u)r_2(u) + \kappa_4 r_{1,2}^2, \qquad (8.13)$$

where κ_4 is the fourth cumulant of $\xi(0)$, and $r_{1,2} = \mathbb{E}[X_1(0)X_2(0)]$.

(b) $\mathrm{Var}(Q_T - S_T) = o(T)$ *as $T \to \infty$.*

A similar result is true for c.t. linear processes of the form (2.12), where now we have (see Ginovyan and Sahakyan [53]):

$$S_T := \int_0^T X_1(u)X_2(u)du.$$

8.4. Fejér-type singular integrals

We define Fejér-type kernels and singular integrals, and state some of their properties that were used to prove the limit theorems stated in Section 7.

For a number k $(k = 2, 3, \ldots)$ and a taper function h satisfying Assumption 6.1 consider the following Fejér-type 'tapered' kernel function:

$$\Phi_{k,T}^h(\mathbf{u}) := \frac{H_T(\mathbf{u})}{(2\pi)^{k-1} H_{k,T}(0)}, \quad \mathbf{u} = (u_1, \ldots, u_{k-1}) \in \mathbb{R}^{k-1}, \qquad (8.14)$$

where

$$H_T(\mathbf{u}) := H_{1,T}(u_1)\cdots H_{1,T}(u_{k-1})H_{1,T}\left(-\sum_{j=1}^{k-1} u_j\right), \qquad (8.15)$$

and the function $H_{k,T}(\cdot)$ is defined by (6.6) with $H_{k,T}(0) = H_{k,T} \neq 0$ (see (6.7)).

The proofs of propositions that follow can be found in Ginovyan and Sahakyan [56]. The next result shows that, similar to the classical Fejér kernel, the 'tapered' kernel $\Phi_{k,T}^h(\mathbf{u})$ is an approximation identity (see Ginovyan and Sahakyan [56], Lemma 3.4).

Proposition 8.5. *For any $k = 2, 3, \ldots$ and a taper function h satisfying Assumption 6.1 the kernel $\Phi_{k,T}^h(\mathbf{u})$, $\mathbf{u} = (u_1, \ldots, u_{k-1}) \in \mathbb{R}^{k-1}$, possesses the following properties:*

a) $\sup_{T>0} \int_{\mathbb{R}^{k-1}} \left| \Phi_{k,T}^h(\mathbf{u}) \right| d\mathbf{u} = C_1 < \infty$;

b) $\int_{\mathbb{R}^{k-1}} \Phi_{k,T}^h(\mathbf{u}) d\mathbf{u} = 1$;

c) $\lim_{T\to\infty} \int_{\mathbb{E}_\delta^c} \left| \Phi_{k,T}^h(\mathbf{u}) \right| d\mathbf{u} = 0$ *for any $\delta > 0$*;

d) *If $k > 2$ for any $\delta > 0$ there exists a constant $M_\delta > 0$ such that for $T > 0$*

$$\left\| \Phi_{k,T}^h \right\|_{L^{p_k}(\mathbb{E}_\delta^c)} \leq M_\delta, \qquad (8.16)$$

where $p_k = \frac{k-2}{k-3}$ for $k > 3$, $p_3 = \infty$, $\mathbb{E}_\delta^c = \mathbb{R}^{k-1} \setminus \mathbb{E}_\delta$, and

$$\mathbb{E}_\delta = \{\mathbf{u} = (u_1, \ldots, u_{k-1}) \in \mathbb{R}^{k-1} : |u_i| \leq \delta, \, i = 1, \ldots, k-1\}.$$

e) *If the function $\Psi \in L^1(\mathbb{R}^{k-1}) \bigcap L^{k-2}(\mathbb{R}^{k-1})$ is continuous at the point $\mathbf{v} = (v_1, \ldots, v_{k-1})$ (L^0 is the space of measurable functions), then*

$$\lim_{T\to\infty} \int_{\mathbb{R}^{k-1}} \Psi(\mathbf{u} + \mathbf{v})\Phi_{k,T}^h(\mathbf{u}) d\mathbf{u} = \Psi(\mathbf{v}). \qquad (8.17)$$

Denote

$$\Delta_{2,T}^h := \int_{\mathbb{R}^2} f(\lambda)g(\lambda + \mu)\Phi_{2,T}^h(\mu) d\lambda d\mu - \int_{\mathbb{R}} f(\lambda)g(\lambda) d\lambda, \qquad (8.18)$$

where $\Phi_{2,T}^h(\mu)$ is given by (8.14)–(8.15).

The next two propositions, which were used to prove Theorems 7.1 and 7.3, give information on the rate of convergence to zero of $\Delta_{2,T}^h$ as $T \to \infty$.

Proposition 8.6. *Assume that Assumptions 7.2 and 6.1 are satisfied. Then the following asymptotic relation holds:*

$$\Delta_{2,T}^h = o\left(T^{-1/2}\right) \quad \text{as} \quad T \to \infty. \qquad (8.19)$$

Proposition 8.7. *Assume that Assumptions 7.3 and 6.1are satisfied. Then the following inequality holds:*

$$|\Delta_{2,T}^h| \le C_h \begin{cases} T^{-(\beta_1+\beta_2)}, & \text{if } \beta_1+\beta_2 < 1 \\ T^{-1}\ln T, & \text{if } \beta_1+\beta_2 = 1 \qquad T > 0, \\ T^{-1}, & \text{if } \beta_1+\beta_2 > 1, \end{cases} \qquad (8.20)$$

where C_h is a constant depending on h.

Notice that for non-tapered case ($h(t) = \mathbb{I}_{[0,1]}(t)$), the above stated results were proved in Ginovyan and Sahakyan [53] (see also Ginovyan and Sahakyan [54, 55]).

8.5. Lévy-Itô-type and Stratonovich-type multiple stochastic integrals

To prove limit theorems for quadratic functionals of Lévy-driven c.t. linear models, was used the multiple *off-diagonal* (Itô-type) and *with-diagonal* (Stratonovich-type) stochastic integrals with respect to Lévy noise. In this subsection we introduce these integrals, and briefly discuss their properties (see, e.g., Bai et al. [9], Farré et al. [35], Peccati and Taqqu [84]).

Let f be a function in $L^2(\mathbb{R}^k)$, then the following off-diagonal multiple stochastic integral, called Itô-Lévy integral, is well-defined:

$$I_k^\xi(f) = \int_{\mathbb{R}^k}' f(x_1, \ldots, x_k)\xi(dx_1)\ldots\xi(dx_k), \qquad (8.21)$$

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 $x_i = x_j$, $i \ne j$. The multiple integral $I_k^\xi(\cdot)$ satisfies the following inequality:

$$\|I_k^\xi(f)\|_{L^2(\Omega)}^2 \le k!\sigma_\xi^{2k}\|f\|_{L^2(\mathbb{R}^k)}^2, \qquad (8.22)$$

and the inequality in (8.22) becomes equality if f is symmetric:

$$\|I_k^\xi(f)\|_{L^2(\Omega)}^2 = k!\sigma_\xi^{2k}\|f\|_{L^2(\mathbb{R}^k)}^2. \qquad (8.23)$$

Observe that if in (8.21), $\xi(\cdot) = B(\cdot)$, where $B(\cdot)$ is the real-valued Brownian motion, then the corresponding integral:

$$I_k^B(f) = \int_{\mathbb{R}^k}' f(x_1, \ldots, x_k)B(dx_1)\ldots B(dx_k) \qquad (8.24)$$

is called multiple Wiener-Itô integral (see Itô [78]).

The Wiener-Itô integral can also be defined with respect to the complex-valued Brownian motion:

$$I_k^W(g) = \int_{\mathbb{R}^k}'' g(u_1, \ldots, u_k)W(du_1)\ldots W(du_k), \qquad (8.25)$$

where $g \in L^2(\mathbb{R}^k)$ is a complex-valued function satisfying $g(-u_1, \ldots, -u_k) = \overline{g(u_1, \ldots, u_k)}$, 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 [34]), 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 [84] and Proposition 4.2 of Dobrushin [30] (see Bai et al. [9]), gives a relationship between the integrals $I_k^B(\cdot)$ and $I_k^W(\cdot)$, defined by (8.24) and (8.25), respectively.

Proposition 8.8. *Let $f_j(\cdot)$ be real-valued functions in $L^2(\mathbb{R}^{k_j})$, $j = 1, \ldots, J$, and let*

$$\widehat{f}_j(w_1, \ldots, w_{k_j}) = \int_{\mathbb{R}^{k_j}} f_j(x_1, \ldots, x_{k_j}) e^{i(x_1 w_1 + \ldots + x_{k_j} w_{k_j})} dx_1 \ldots dx_{k_j}$$

be the L^2-Fourier transform of $f_j(\cdot)$. Then

$$\left(I_{k_1}^B(f_1), \ldots, I_{k_J}^B(f_J) \right) \overset{d}{=}$$
$$\left((2\pi)^{-k_1/2} I_{k_1}^W \left(\widehat{f}_1 A^{\otimes k_1} \right), \ldots, (2\pi)^{-k_J/2} I_{k_J}^W \left(\widehat{f}_J A^{\otimes k_J} \right) \right),$$

for any function $A(u) : \mathbb{R} \to \mathbb{C}$ such that $|A(u)| = 1$ and $A(w) = \overline{A(-w)}$ almost everywhere, where $A^{\otimes k}(w_1, \ldots, w_k) := A(w_1) \cdots A(w_k)$.

In the next proposition we state a stochastic Fubini's theorem (see Bai et al. [9], Lemma 3.1, or Peccati and Taqqu [84], Theorem 5.12.1).

Proposition 8.9. *Let (S, μ) be a measure space with $\mu(S) < \infty$, and let $f(s, x_1, \ldots, x_k)$ be a function on $S \times \mathbb{R}^k$ such that*

$$\int_S \int_{\mathbb{R}^k} f^2(s, x_1, \ldots, x_k) dx_1 \ldots dx_k \mu(ds) < \infty,$$

then we can change the order of the multiple stochastic integration $I_k^\xi(\cdot)$ and the deterministic integration $\int_S f(s, \cdot) \mu(ds)$:

$$\int_S I_k^\xi \big(f(s, \cdot) \big) \mu(ds) = I_k^\xi \left(\int_S f(s, \cdot) \mu(ds) \right).$$

The *with-diagonal* counterpart of the Lévy-Itô integral $I_k^\xi(f)$, called a *Stratonovich-type* stochastic integral, is defined by

$$\mathring{I}_k^\xi(f) := \int_{\mathbb{R}^k} f(x_1, \ldots, x_k) \xi(dx_1) \ldots \xi(dx_k), \tag{8.26}$$

which includes all the diagonals. We refer to Farré et al. [35] for a comprehensive treatment of Stratonovich-type integrals $\mathring{I}_k^\xi(f)$. Observe that 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 $\mathbb{R}^k$ (see Bai et al. [9], Farré et al. [35]).

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. [35], Theorem 5.9). In the special case when $k = 2$, we have

$$\mathring{I}_2^\xi(f) = \int_{\mathbb{R}^2}' f(x_1, x_2)\xi(dx_1)\xi(dx_2) + \int_{\mathbb{R}} f(x, x)\xi_c^{(2)}(dx) + \int_{\mathbb{R}} f(x, x)dx, \quad (8.27)$$

where

$$\xi_c^{(2)}(t) = \xi^{(2)}(t) - \mathbb{E}\xi^{(2)}(t) = \xi^{(2)}(t) - |t|, \quad (8.28)$$

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. [35], equation (10)). The centered process $\xi_c^{(2)}(t)$ is called a second order *Teugels martingale*, 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. [35])

$$\mathbb{E}\left[\int_{\mathbb{R}} g(x)\xi_c^{(2)}(dx) \int_{\mathbb{R}} h(x)\xi_c^{(2)}(dx)\right] = \kappa_4 \int_{\mathbb{R}} f(x)g(x)dx. \quad (8.29)$$

The decomposition (8.27) implies that

$$\mathbb{E}\left[\mathring{I}_k^\xi(f)\right] = \int_{\mathbb{R}} f(x, x)dx.$$

Notice that for any $f \in L^2(\mathbb{R}^2)$ and $g \in L^2(\mathbb{R})$ 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 $\xi_c^{(2)}(t)$:

$$\int_{\mathbb{R}^2}' f(x_1, x_2)\xi(dx_1)\xi(dx_2) \quad \text{and} \quad \int_{\mathbb{R}} g(x)\xi_c^{(2)}(dx). \quad (8.30)$$

are uncorrelated (see Bai et al. [9]).

8.6. Power counting theorems

Power counting theorems provide convergence conditions for some classes of integrals on $\mathbb{R}^n$ whose integrands are products of functions bounded near zero by powers of linearly dependent affine functionals and near infinity by different powers of those functionals. These theorems are useful in studying asymptotic distributions of statistics of time series models with long-range dependence. The results stated below were used in Terrin and Taqqu [101] to establish non-central limit theorems for quadratic forms Q_T of d.t. stationary processes with long-range dependence (see Theorem 4.5).

First we introduce some notation. Let $\mathfrak{L} := \{L_1, \ldots, L_m\}$ be a collection of linear functionals $L_i(\mathbf{x})$ on $\mathbb{R}^n$, $i = 1, \ldots, m$. For numbers $0 < a_i \leq b_i$, $c_i > 0$, and real constants α_i and β_i we define

$$P_1(\mathbf{x}) := f_1(L_1(\mathbf{x})) f_2(L_2(\mathbf{x})) \cdots f_m(L_m(\mathbf{x})), \tag{8.31}$$

where the functions f_i, $(i = 1, \ldots, m)$ satisfy the condition:

$$|f_i(y)| \leq \left\{ \begin{array}{ll} c_i |y_i|^{\alpha_i} & \text{if} \quad |y| < a_i \\ c_i |y_i|^{\beta_i} & \text{if} \quad |y| > b_i, \end{array} \right. \tag{8.32}$$

and $|f_i(y)|$ is bounded above in the interval (a_i, b_i), $i = 1, ..., m$. The constants α_i and β_i are called the *exponents* of $|y|$ around 0 and ∞, respectively. For $W \subset T$ we set $s(W) := \text{span}(W) \cap T$, where $\text{span}(W)$ denotes the linear span of W. Define

$$d_0(W) := r(W) + \sum_{s(W)} \alpha_i \quad \text{and} \quad d_\infty(W) := r(T) - r(W) + \sum_{T \setminus s(W)} \beta_i. \tag{8.33}$$

A summation over a set E means summation over the set $\{i : L_i \in E\}$, $|E|$ denotes the cardinality of E and $r(E)$ is the rank of E, that is, the number of linearly independent functionals in E. We call a set W *padded* if for every L in W, L is also in $s(W \setminus \{L\})$, that is, L can be obtained as a linear combination of other elements in W. Observe that $W = \emptyset$ is padded, and if W is linearly independent, then $d_0(W) := |W| + \sum_{s(W)} \alpha_i$.

The proofs of the following results can be found in Terrin and Taqqu [101] (see also Fox and Taqqu [39], and Terrin and Taqqu [103]).

Proposition 8.10. *If $r(T) = n$ and (a) $d_0(W) > 0$ for every nonempty subset W of T with $s(W) = W$, while (b) $d_\infty(W) < 0$ for every proper subset W of T with $s(W) = W$, including the empty set. Then*

$$\int_{\mathbb{R}^n} |P_1(\mathbf{x})| d\mathbf{x} < \infty, \tag{8.34}$$

where $P_1(\mathbf{x})$ is as in (8.31).

Proposition 8.10 can be extended to the class of functionals of the form $L_i(\mathbf{x}) + \theta_{\mathbf{i}}$, where θ_i is a constant. For $i = 1, ..., m$ let α_i and θ_i be real constants, and let $M_i(\mathbf{x})$ be a linear functional on $\mathbb{R}^n$. Put $L_i(\mathbf{x}) := \mathbf{M_i}(\mathbf{x}) + \theta_{\mathbf{i}}$ and set

$$P_2(\mathbf{x}) := |L_1(\mathbf{x})|^{\alpha_1} \cdots |L_m(\mathbf{x})|^{\alpha_m}. \tag{8.35}$$

Proposition 8.11. *Let $s(W)$ and $d_0(W)$ be as above. Suppose that $d_0(W) > 0$ for every nonempty subset W of T with $s(W) = W$. Then*

$$\int_{[-t,t]^n} |P_2(\mathbf{x})| d\mathbf{x} < \infty \quad \text{for all} \quad t > 0, \tag{8.36}$$

where $P_2(\mathbf{x})$ is as in (8.35).

Remark 8.4. If $\alpha_i > -1$ and $\beta_i \geq -1$ in Proposition 8.10, then it suffices to verify the conditions (a) and (b) for sets W that are also padded. If $\alpha_i > -1$ in Proposition 8.11, then it suffices to consider subsets W that are also padded (see Terrin and Taqqu [101, 103]).

Acknowledgments

The authors would like to thank the anonymous referees for their careful review of the manuscript and valuable remarks and comments.

References

[1] ALOMARI, H. M., FRÍAS, M.P., LEONENKO, N. N., RUIZ-MEDINA, M.D. SAKHNO, L. and TORRES, A. (2017). Asymptotic properties of parameter estimates for random fields with tapered data, *Electronic Journal of Statistics* **11**, 3332–3367. MR3708540

[2] ANH, V.V., ANGULO, J.M. and RUIZ-MEDINA, M.D. (1999). Possible long-range dependence in fractional random fields. *J. Statist. Plann. Inference* **80**, 95–110. MR1713795

[3] ANH, V.V., LEONENKO, N.N. and MCVINISH, R. (2001). Models for fractional Riesz-Bessel motion and related processes. *Fractals* **9**, 329–346.

[4] ANH, V.V., LEONENKO, N.N. and SAKHNO, L.M. (2007). Minimum contrast estimation of random processes based on information of second and third orders. *J. Statist. Planning Inference.* **137**, 1302–1331. MR2301481

[5] ANH, V.V., LEONENKO, N.N. and SAKHNO, L.M. (2009). Evaluation of bias in high-order spectral estimation. *Theor. Probability and Math. Statist.* **80**, 1–14. MR2541947

[6] AVRAM, F. (1988). On bilinear forms in Gaussian random variables and Toeplitz matrices. *Probab. Theory Related Fields* **79**, 37–45. MR0952991

[7] AVRAM, F., LEONENKO, N. N. AND SAKHNO, L. (2010). On a Szegö 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. MR2741966

[8] BAI, S., GINOVYAN, M. S. and TAQQU, M. S. (2015). Functional limit theorems for Toeplitz quadratic functionals of continuous time Gaussian stationary processes. *Statistics and Probability Letters* **104**, 58-67. MR3360705

[9] BAI, S., GINOVYAN, M. S. and TAQQU, M. S. (2016). Limit theorems for quadratic forms of Lévy-driven continuous time linear processes. *Stochast. Process. Appl.* **126**, 1036–1065. MR3461190

[10] BENTKUS, R. (1972). On the error of the estimate of the spectral function of a stationary process. *Lit. Mat. Sb.* **12**, 55–71. MR0319332

[11] BERAN, J.(1994). *Statistics for Long-Memory Processes.* Chapman & Hall, New York. MR1304490

[12] BERAN, J., FENG, Y., GHOSH, S. and KULIK, R. (2013). *Long-Memory Processes: Probabilistic Properties and Statistical Methods.* Springer, Berlin. MR3075595

[13] BINGHAM, N.H., GOLDIE, C.M. and TEUGELS, J.L. (1989). *Regular Variation.* Cambridge University Press, New York. MR1015093

[14] BLOOMFIELD, P. (1976). *Fourier Annlysis of Time Series.* Wiley, New York. MR0654511

[15] BOGACHEV, V.I. (2007) *Measure Theory*, Vol. 1, Springer, New York. MR2267655

[16] BONDON, P. and PALMA, W. (2006). A class of antipersistent processes. *J. Time Ser. Anal.* **28**, 261–273. MR2345660

[17] BRILLINGER, D.R. (1981). *Time Series: Data Analysis and Theory*, Holden Day, San Francisco. MR0595684

[18] BROCKWELL, P.J (2014). Recent results in the theory and applications of CARMA processes. *Annals of the Institute of Statistical Mathematics* **66**(4), 647–685. MR3224604

[19] BROCKWELL, P.J and DAVIS, R.A. (1991) *Time Series: Theory and Methods*, 2nd ed. Springer, New York. MR2839251

[20] BRYC, W. and DEMBO, A. (1997) Large deviations for quadratic functionals of Gaussian processes. *J. Theoret. Probab.* **10**, 307–332. MR1455147

[21] CASAS, I. and GAO, J. (2008). Econometric estimation in long-range dependent volatility models: Theory and practice. *Journal of Econometrics.* **147**, 72–83. MR2472982

[22] COHEN, S. and LINDNER, A. (2013). A central limit theorem for the sample autocorrelations of a Lévy driven continuous time moving average process, *J. Statist. Plann. Inference* **143** (8), 1295–1306. MR3055748

[23] CRAMÉR, H. and LEADBETTER, M.R. (1967) *Stationary and Related Stochastic Processes: Sample Function Properties and Their Applications.* John Wiley & Sons, New York. MR0217860

[24] DAHLHAUS, R. (1983). Spectral analysis with tapered data, *J. Time Ser. Anal.* **4**, 163-174. MR0732895

[25] DAHLHAUS, R. (1984). Parameter estimation of stationary processes with spectra containing strong peaks. In "Robust and Nonlinear Time Series Analysis" (Eds. Franke, Hardle and Martin). *Lecture Notes in Statistics*, **26**, 50-67. MR0786303

[26] DAHLHAUS, R. (1985). A functional limit theorem for tapered empirical spectral functions, *Stoch. Process. Appl.* **19**, 135–149. MR0780726

[27] DAHLHAUS, R. (1989). Efficient parameter estimation for self-similar processes. *Ann. Statist.* **17** 1749–1766. MR1026311

[28] DAHLHAUS, R. and KÜNSCH, H (1987). Edge effects and efficient parameter estimation for stationary random fields, *Biometrika* **74**(4), 877–882. MR0919857

[29] DAHLHAUS, R. and WEFELMEYER, W. (1996) Asymptotically optimal estimation in misspecified time series models, *Ann. Statist.* **24**, 952–974. MR1401832

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

[31] DOOB, J.L. (1953). *Stochastic Processes*, Wiley, New York. MR0058896

[32] DOUKHAN, P., LEÓN, J. and SOULIER, P. (1996). Central and non central limit theorems for strongly dependent stationary Gaussian field. *Rebrape*, **10**, 205–223. MR1475443

[33] DZHAPARIDZE, K. (1986). *Parameter Estimation and Hypotesis Testing in Spectral Analysis of Stationary Time Series*, Springer, New York. MR0812272

[34] EMBRECHTS, P. and MAEJIMA, M. (2002). *Selfsimilar Processes*, Princeton University Press. MR1920153

[35] FARRÉ, M., JOLIS, M. and UTZET, F. (2010). Multiple Stratonovich integral and Hu-Meyer formula for Lévy processes, *Ann. Probab.* **38**(6), 2136–2169. MR2683627

[36] FLANDRIN, P. (1989). On the spectrum of fractional Brownian motions. *IEEE*

Trans. Inform. Theory **35**, 197–199. MR0995341

[37] Fox, R. and Taqqu, M.S. (1985). Non-central limit theorems for quadratic forms in random variables having long-range dependence. *Ann. Probab.* **13**(2), 428–446. MR0781415

[38] Fox, R. and Taqqu, M.S. (1986). Large-sample properties of parameter estimation for strongly dependent stationary Gaussian time series. *Ann. Statist.* **14**, 517–532. MR0840512

[39] Fox, R. and Taqqu, M.S. (1987). Central limit theorem for quadratic forms in random variables having long-range dependence. *Probab. Theory Related Fields* **74**, 213–240. MR0871252

[40] Gao, J. (2004). Modelling long-range dependent Gaussian processes with application in continuous time financial models. *J. Appl.Probab.* **41**, 467–482. MR2052585

[41] Gao J., Anh V., Heyde C. and Tieng Q. (2001). Parameter estimation of stochastic processes with long-range dependence and intermittency. *J. Time Ser. Anal.* **22**, 517–535. MR1859563

[42] Ginovyan, M.S. (1988). Asymptotically efficient nonparametric estimation of functionals on spectral density with zeros, *Theory Probab. Appl.* **33** 315–322. MR0954578

[43] Ginovyan, M.S. (1988). On estimate of the value of the linear functional in a spectral density of stationary Gaussian process. *Theory Probab. Appl.* **33**, 777–781. MR0979749

[44] Ginovyan, M.S. (1993). A note on central limit theorem for Toeplitz type quadratic forms in stationary Gaussian variables. *J. Cont. Math. Anal.* **28**, 78–81. MR1359889

[45] Ginovyan, M.S. (1994). On Toeplitz type quadratic functionals in Gaussian stationary process. *Probab. Theory Related Fields.* **100**, 395–406. MR1305588

[46] Ginovyan, M.S. (1995). Asymptotic properties of spectrum estimate of stationary Gaussian process. *J. Cont. Math. Anal.* **30** 1–17. MR1643528

[47] Ginovyan, M.S. (2003). Asymptotically efficient nonparametric estimation of nonlinear spectral functionals. *Acta Appl. Math.* **78**, 145–154. MR2024019

[48] Ginovyan, M.S. (2011). Efficient estimation of spectral functionals for Gaussian stationary models. *Communications on Stochastic Analysis* **5**, 211–232. MR2808543

[49] Ginovyan, M.S. (2011). Efficient estimation of spectral functionals for continuous time stationary models. *Acta Appl. Math.* **115**, 233–254. MR2818916

[50] Ginovyan, M.S. (2020). Parameter Estimation for Lévy-Driven continuous time Linear Models with Tapered Data. *Acta Appl. Math.* **169** (1), 79–97. MR4146891

[51] Ginovyan, M.S. (2021). Goodness-of-fit tests for stationary Gaussian processes with tapered data. *Acta Appl. Math.* **171**(1), 1–12. MR4187307

[52] Ginovyan, M.S. and Sahakyan, A. A. (2005). On the Central Limit Theorem for Toeplitz Quadratic Forms of Stationary Sequences. *Theory Probab. and Appl.* **49**, 612 - 628. MR2142560

[53] Ginovyan, M.S. and Sahakyan, A. A. (2007). Limit Theorems for Toeplitz quadratic functionals of continuous time stationary process, *Probab. Theory Relat. Fields* **138**, 551–579. MR2299719

[54] Ginovyan, M.S. and Sahakyan, A. A. (2012). Trace approximations of products of truncated Toeplitz operators. *Theory Probab. Appl.* **56**(1), 57–71. MR2848419

[55] Ginovyan, M.S. and Sahakyan, A. A. (2013). On the trace approximations of

products of Toeplitz matrices. *Statist. Probab. Lett.* **83**(3), 753-760. MR3040300

[56] GINOVYAN, M.S. and SAHAKYAN, A. A. (2019). Estimation of spectral functionals for Lévy-driven continuous time linear models with tapered data, *Electronic Journal of Statistics* **13**, 255–283. MR3905127

[57] GINOVYAN, M.S. and SAHAKYAN, A. A. (2019). Limit theorems for tapered Toeplitz quadratic functionals of continuous time Gaussian stationary processes. *J. Cont. Math. Anal.* **54**(4), 222–239. MR4019542

[58] GINOVYAN, M.S., SAHAKYAN, A. A. and TAQQU M. S. (2014). The trace problem for Toeplitz matrices and operators and its impact in probability. *Probability Surveys* **11**, 393–440. MR3290440

[59] GIRAITIS, L., KOUL, H. and SURGAILIS, D. (2012). *Large Sample Inference for Long Memory Processes*. Imperial College Press, London. MR2977317

[60] GIRAITIS, L. and SURGAILIS, D. (1990). A central limit theorem for quadratic forms in strongly dependent linear variables and its application to asymptotical normality of Whittle's estimate. *Probab. Theory Related Fields.* **86**, 87–104. MR1061950

[61] GIRAITIS, L. and TAQQU, M.S. (1997). Limit theorems for bivariate Appell polynomials: Part I. Central limit theorems. *Probab. Theory Related Fields*, **107**, 359–381. MR1440137

[62] GIRAITIS, L. and TAQQU, M.S. (1998). Central limit theorems for quadratic forms with time-domain conditions. *Annals of Probability.* **26**(1), 377–398. MR1617055

[63] GIRAITIS, L. and TAQQU, M.S. (1999). Convergence of normalized quadratic forms, *J. Statist. Plann. Inference* **80** (1), 15–35. MR1713799

[64] GIRAITIS, L. and TAQQU, M.S. (1999). Whittle estimator for finite-variance non-Gaussian long-memory time series. *Ann. Statist.* **27** (1), 178–203. MR1701107

[65] GIRAITIS, L. and TAQQU, M.S. (2001). Functional non-central and central limit theorems for bivariate Appell polynomials. *Journal of Theoretical Probability*, **14**(2), 393–426. MR1838735

[66] GRANGER, C.W.J. and JOYEUX, K. (1980). An introduction to long-memory time series and fractional differencing. *J. Time Ser. Anal.* **1**, 15–29. MR0605572

[67] GRENANDER, U. and SZEGŐ, G. (1958). *Toeplitz Forms and Their Applications*. University of California Press, Berkeley and Los Angeles. MR0094840

[68] GUÉGAN, D. (2005). How can we define the concept of long memory? An econometric survey. *Econometric Reviews* **24**, 113–149. MR2190313

[69] GUYON, X. (1995). *Random Fields on a Network: Modelling, Statistics and Applications*, Springer, New York. MR1344683

[70] HASMINSKII, R.Z. and IBRAGIMOV, I.A. (1986). Asymptotically efficient nonparametric estimation of functionals of a spectral density function. *Probab. Theory Related Fields* **73**, 447–461. MR0859842

[71] HENRY, M. and ZAFFARONI, P. (2003). The long range dependence paradigm for macroeconomics and finance. In *Long range dependence: Theory and applications*. Doukhan, P., Oppenheim, G., Taqqu, M.S. Eds. Birkhüser, Boston, 417–438. MR1957502

[72] HEYDE, C.C. and GAY, R. (1993). Smoothed periodogram asymptotics and estimation for processes and fields with possible long-range dependence. *Stochastic Processes and their Applications* **45**, 169–182. MR1204868

[73] HOSKING, J.R.M. (2002). Fractional differencing. *Biometrika* **68**, 165–176. MR0614953

[74] HUNT, R. A., MUCKENHOUPT, B. and WHEEDEN, R. L. (1973). Weighted Norm Inequalities for the Conjugate Function and Hilbert Transform. *Trans. of the*

AMS **176**, 227–251. MR0312139

[75] Ibragimov, I.A. (1963). On estimation of the spectral function of a stationary Gaussian process. *Theory Probab. Appl.* **8**, 366–401. MR0160274

[76] Ibragimov, I.A. (1967). On maximum likelihood estimation of parameters of the spectral of stationary time series. *Theory Probab. Appl.* **12**, 115–119. MR0228095

[77] Ibragimov, I.A. and Linnik, Yu. V. (1971). *Independent and Stationary Sequences of Random Variables.* Wolters-Noordhoff Publishing Groningen, The Netherlands. MR0322926

[78] Itô, K. (1951). Multiple Wiener Integral. *J. Math. Soc. Japan* **3**(1), 157-169. MR0044064

[79] Kac, M. (1954). Toeplitz matrices, translation kernels and a related problem in probability theory. *Duke Math. J.* **21**, 501–509. MR0062867

[80] Leonenko, N. N. and Sakhno, L. M. (2006). On the Whittle estimators for some classes of continuous-parameter random processes and fields. *Stat. Probab. Lett.* **76**, 781–795. MR2266092

[81] Lieberman, O. and Phillips, P.C.B. (2004). Error bounds and asymptotic expansions for Toeplitz product functionals of unbounded spectra. *J. Time Ser. Anal.* **25**, 733–753. MR2089192

[82] Ludeña, C. and Lavielle, M. (1999). The Whittle Estimator for Strongly Dependent Stationary Gaussian Fields. *Scand. J. Stat.* **26**, 433–450. MR1712039

[83] Palma, W. (2007). *Long-memory time series.* Wiley, New York. MR2297359

[84] Peccati, G. and Taqqu, M.S. (2011). *Wiener Chaos: Moments, Cumulants and Diagrams: as Survey With Computer Implementation*, Springer, New York. MR2791919

[85] Robinson, P.M. (2003) Long memory time series. In *Time series with long memory.* Robinson, P.M. Ed. Oxford University Press, Oxford, 4–32. MR2077520

[86] Rosenblatt M. (1961). Independence and dependence. In *Proc. 4th Berkeley Symp. Math. Stat. Probab.* **2**, 431–443. MR0133863

[87] Rosenblatt M. (1962). Asymptotic behavior of eigenvalues of Toeplitz forms. *J. Math. and Mech.* **11**, 941–950. MR0150841

[88] Rosenblatt M. (1979). Some limit theorems for partial sums of quadratic forms in stationary Gaussian variables. *Z. Wahr. verw. Gebiete* **49**, 125–132. MR0543988

[89] Samorodnisky, G. and Taqqu, M.S. (1994). *Stable Non-Gaussian Random Peocesses.* Chapman & Hall/CRC, New York. MR1280932

[90] Solev, V. N. and Gerville-Reache, L. (2006). Large Toeplitz operators and quadratic forms generated by a stationary Gaussian sequence. *Journal of Mathematical Sciences* **139**(3), 6625-6630. MR2214543

[91] Solo, V. (1992). Intrinsic random functions and the paradox of 1/f noise. *SIAM J. Appl. Math.* **52** 270–291. MR1148328

[92] Soulier, P. (2009). Some applications of regular variation in probability and statistics. Unpublished Notes.

[93] Spangenberg, F. (2015). A central limit theorem for the sample autocovariance of a continuous time moving average processwith long memory. *ArXiv Preprint* arXiv:1502.04851.

[94] Surgailis, D. (1982). Zones of attraction of self-similar multiple integrals. *Lithuanian Mathematical Journal* **22** (3), 327–340. MR0684472

[95] Taniguchi, M. (1987). Minimum contrast estimation for spectral densities of stationary processes. *J. R. Stat. Soc. Ser. B-Stat. Methodol.* **49**, 315–325. MR0928940

[96] Taniguchi, M. and Kakizawa, Y. (2000). *Asymptotic Theory of Statistical Inference for Time Series.* Academic Press, New York. MR1785484

[97] Taqqu, M.S. (1975). Weak convergence to fractional Brownian motion and the Rosenblatt process. *Z. Wahr. verw. Gebiete* **31**, 287–302. MR0400329

[98] Taqqu, M.S. (1987). Toeplitz matrices and estimation of time series with long-range dependence. In *Proc. of the First World Congress of the Bernoulli Society.* VNU Science Press, BV: Utrecht, The Netherlands **1**, 75–83. MR1092337

[99] Taqqu, M.S. (2003). Fractional Brownian motion and long-range dependence. In *Long range dependence: Theory and applications.* Doukhan, P., Oppenheim, G., Taqqu, M.S. Eds. Birkhüser, Boston, 5–38. MR1956042

[100] Taqqu, M.S. (2011). The Rosenblatt process. In *Selected works of Murray Rosenblatt, Selected works in Probab. and Statist.* Davis R.A. et al. Eds. Springer, Boston, 29–45. MR2742596

[101] Terrin, N. and Taqqu, M.S. (1990). A noncentral limit theorem for quadratic forms of Gaussian stationary sequences. *J. Theoret. Probab.* **3**, 449–475. MR1057525

[102] Terrin, N. and Taqqu, M.S. (1991). Convergence to a Gaussian limit as the normalization exponent tends to $1/2$. *Statist. Probab. Lett.* **11**, 419–427. MR1114532

[103] Terrin, N. and Taqqu, M.S. (1991). Power Counting Theorem in Euclidean Space. In: *Random Walks, Brownian Motion, and Interacting Particle Systems. Progress in Probability.* Durrett R., Kesten H. Eds. Vol. 28. Birkhäuser, Boston, MA, 425–440. MR1146462

[104] Tsai, H. and Chan, K. S. (2005). Quasi-maximum likelihood estimation for a class of continuous-time long memory processes. *J. Time Ser. Anal.* **26**(5), 691–713. MR2188305

[105] Veillette, M. and Taqqu, M.S. (2013). Properties and numerical evaluation of the Rosenblatt distribution. *Bernoulli* **19**, 982–1005. MR3079303

[106] Walker, A.M. (1964). Asymptotic properties of least-squares estimates of parameters of the spectrum of a stationary non-deterministic time-series. *J. Austr. Math. Soc.* **4**, 363–384. MR0171345

[107] Whittle, P. (1951). *Hypothesis Testing in Time Series.* Hafner, New York. MR0040634

[108] Yaglom, A.M. (1986). *The Correlation Theory of Stationary and Related Random Processes,* Vol. 1, Springer, New York. MR0915557