Probability Surveys
Vol. 11 (2014) 393–440
ISSN: 1549-5787
DOI: 10.1214/13-PS217

The trace problem for Toeplitz matrices and operators and its impact in probability

Mamikon S. Ginovyan[*]

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

Artur A. Sahakyan

Department of Mathematics and Mechanics
Yerevan State University, Yerevan, Armenia
e-mail: sart@ysu.am

and

Murad S. Taqqu

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

Abstract: The trace approximation problem for Toeplitz matrices and its applications to stationary processes dates back to the classic book by Grenander and Szegö, *Toeplitz forms and their applications* (University of California Press, Berkeley, 1958). It has then been extensively studied in the literature.

In this paper we provide a survey and unified treatment of the trace approximation problem both for Toeplitz matrices and for operators and describe applications to discrete- and continuous-time stationary processes.

The trace approximation problem serves indeed as a tool to study many probabilistic and statistical topics for stationary models. These include central and non-central limit theorems and large deviations of Toeplitz type random quadratic functionals, parametric and nonparametric estimation, prediction of the future value based on the observed past of the process, hypotheses testing about the spectrum, etc.

We review and summarize the known results concerning the trace approximation problem, prove some new results, and provide a number of applications to discrete- and continuous-time stationary time series models with various types of memory structures, such as long memory, anti-persistent and short memory.

AMS 2000 subject classifications: Primary 60G10, 62G20; secondary 47B35, 15B05.
Keywords and phrases: Stationary process, spectral density, long-memory, central limit theorem, Toeplitz operator, trace approximation, singularity.

Received April 2013.

[*]Corresponding author.

Contents

1. Introduction

Toeplitz matrices and operators, which have great independent interest and a wide range of applications in different fields of science (economics, engineering, finance, hydrology, physics, signal processing, etc.), arise naturally in the context of stationary processes. This is because the covariance matrix of a discrete-time stationary process is a truncated Toeplitz matrix generated by the spectral density of that process. Conversely, any non-negative summable function generates a Toeplitz matrix, which can be considered as a spectral density of some discrete-time stationary process, and therefore the corresponding truncated Toeplitz matrix will be the covariance matrix of that process. In the continuous-time case, Toeplitz matrix is replaced by a Toeplitz operator.

Truncated Toeplitz matrices and operators are of particular importance, and serve as tools, to study many topics in the spectral and statistical analysis of discrete- and continuous-time stationary processes, such as central and non-central limit theorems and large deviations of Toeplitz type random quadratic forms and functionals, estimation of the spectral parameters and functionals, asymptotic expansions of the estimators, hypotheses testing about the spectrum, prediction of the future value based on the observed past of the process, etc. (see, e.g., [1]–[7], [10, 12, 13, 15, 17], [20]–[70], and references therein).

The present work is devoted to the problem of approximation of the traces of products of truncated Toeplitz matrices and operators generated by integrable real symmetric functions defined on the unit circle (resp. on the real line). We discuss estimation of the corresponding errors, and describe applications to discrete- and continuous-time stationary time series models with various types of memory structures (long-memory, anti-persistent and short-memory).

The paper is organized as follows. In the remainder of this section we state the trace approximation problem and describe the statistical model. In Section 2 we discuss the trace problem for Toeplitz matrices. Section 3 considers the same problem for Toeplitz operators. Section 4 is devoted to some applications of the trace problem to discrete- and continuous-time stationary processes and also contains some new results. These are proved within the section. Also, the paper contains a number of new theorems both for Toeplitz matrices and operators, which are stated in Sections 2 and 3, respectively. Those in Section 3 involving Toeplitz operators (Theorems 3.1–3.4) are proved in Section 5. The corresponding theorems of Section 2 involving Toeplitz matrices can be proved in a similar way and hence their proofs are omitted. An Appendix contains the proofs of technical lemmas.

1.1. The trace approximation problem

We first define the main objects to be studied in this work, namely the truncated Toeplitz matrices and operators, generated by integrable real symmetric functions.

Let $h(\lambda)$ be an integrable real symmetric function defined on $\mathbb{T} := (-\pi, \pi]$. For $T = 1, 2, \ldots$, the $(T \times T)$-*truncated Toeplitz matrix* generated by $h(\lambda)$, denoted by $B_T(h)$, is defined by the following equation (see, e.g., [33]):

$$B_T(f) := \|\widehat{h}(s-t)\|_{s,t=1,2,\ldots,T}, \tag{1.1}$$

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

Given a number $T > 0$ and an integrable real symmetric function $h(\lambda)$ defined on $\mathbb{R} := (-\infty, \infty)$, the T-*truncated Toeplitz operator* generated by $h(\lambda)$, denoted by $W_T(h)$, is defined by the following equation (see, e.g., [33, 40, 45]):

$$[W_T(h)u](t) = \int_0^T \hat{h}(t-s)u(s)ds, \quad u(s) \in L^2[0,T], \tag{1.2}$$

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

The problem of approximating traces of products of truncated Toeplitz matrices and operators can be stated as follows.

Let $\mathcal{H} = \{h_1, h_2, \ldots, h_m\}$ be a collection of integrable real symmetric functions defined on the domain Λ, where $\Lambda = \mathbb{R} := (-\infty, \infty)$ or $\Lambda = \mathbb{T} := (-\pi.\pi]$.

For a given number $T > 0$, let $A_T(h_k)$ denote either the $(T \times T)$-truncated Toeplitz matrix $(B_T(h_k))$, or the T-truncated Toeplitz operator $(W_T(h_k))$ generated by function h_k, defined respectively by (1.1) and (1.2) with h_k instead of h. Further, let

$$\tau := \{\tau_k : \tau_k \in \{-1, 1\},\ k = 1, 2, \ldots, m\} \tag{1.3}$$

be a given sequence of ± 1's. Define

$$S_{A,\mathcal{H},\tau}(T) := \frac{1}{T} \operatorname{tr} \left[\prod_{k=1}^{m} \{A_T(h_k)\}^{\tau_k} \right], \tag{1.4}$$

where tr$[A]$ stands for the trace of A,

$$M_{\Lambda,\mathcal{H},\tau} := (2\pi)^{m-1} \int_\Lambda \prod_{k=1}^{m} [h_k(\lambda)]^{\tau_k} d\lambda, \tag{1.5}$$

and

$$\Delta_{A,\Lambda,\mathcal{H},\tau}(T) := |S_{A,\mathcal{H},\tau}(T) - M_{\Lambda,\mathcal{H},\tau}|. \tag{1.6}$$

The problem is to approximate $S_{A,\mathcal{H},\tau}(T)$ by $M_{\Lambda,\mathcal{H},\tau}$ and estimate the error rate for $\Delta_{A,\Lambda,\mathcal{H},\tau}(T)$ as $T \to \infty$. More precisely, for a given sequence $\tau = \{\tau_k \in \{-1, 1\}, k = 1, 2, \ldots, m\}$ find conditions on functions $\{h_k(\lambda), k = 1, 2, \ldots, m\}$ such that:

$$\text{Problem (A)}: \quad \Delta_{A,\Lambda,\mathcal{H},\tau}(T) = o(1) \quad \text{as} \quad T \to \infty, \quad \text{or}$$
$$\text{Problem (B)}: \quad \Delta_{A,\Lambda,\mathcal{H},\tau}(T) = O(T^{-\gamma}), \quad \gamma > 0, \quad \text{as} \quad T \to \infty.$$

The trace approximation problem goes back to the classical monograph by Grenander and Szegö [40], and has been extensively studied in the literature (see, e.g., Kac [49], Rosenblatt [57], [58], Ibragimov [45], Taniguchi [62], Avram [4], Fox and Taqqu [20], Taqqu [66], Dahlhaus [17], Giraitis and Surgailis [37], Ginovyan [24], Taniguchi and Kakizawa [64], Lieberman and Phillips [51], Giraitis et al. [36], Ginovyan and Sahakyan [32]–[35], and references therein).

In this paper we review and summarize the known results concerning Problems (A) and (B), prove some new results, as well as provide a number of applications to discrete- and continuous-time stationary time series models that have various types of memory structures (short-, intermediate-, and long-memory).

We focus on the following special case which is important from an application viewpoint, and is commonly discussed in the literature: $m = 2\nu$, $\tau_k = 1$, $k = 1, 2, \ldots, m$ (or $\tau_k = (-1)^k$, $k = 1, 2, \ldots, m$), and

$$\begin{aligned} h_1(\lambda) = h_3(\lambda) = \cdots = h_{2\nu-1}(\lambda) := f(\lambda) \\ h_2(\lambda) = h_4(\lambda) = \cdots = h_{2\nu}(\lambda) := g(\lambda). \end{aligned} \tag{1.7}$$

This case is of importance from an application viewpoint because it appears in many problems involving statistical analysis of stationary processes: asymptotic distributions and large deviations of Toeplitz-type quadratic functionals, estimation of the spectral parameters and functionals, asymptotic expansions of the estimators, hypotheses testing about the spectrum, etc. (see, e.g., [4], [17], [20], [24], [32]–[37], [45], [51], [62], [64], and references therein).

Throughout the paper the letters C and c, with or without index, are used to denote positive constants, the values of which can vary from line to line. Also, all functions defined on $\mathbb{T}$ are assumed to be 2π-periodic and periodically extended to $\mathbb{R}$.

1.2. The model: Short, intermediate and long memory processes

Let $\mathbb{U}$ denote either the real line $\mathbb{R} := (-\infty, \infty)$, or the set of integers $\mathbb{Z} := \{0, \pm 1, \pm 2, \ldots\}$. Let $\{X(u),\ u \in \mathbb{U}\}$ be a centered, real-valued, continuous-time or discrete-time second-order stationary process with covariance function $r(u)$, possessing a spectral density function $f(\lambda)$, $\lambda \in \Lambda$, where $\Lambda = \mathbb{R}$ or $\Lambda = \mathbb{T}$, that is,

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

and $r(u)$ and $f(\lambda)$ are connected by the Fourier integral (see, e.g., Brockwell and Davis [12], Section 4.3):

$$r(u) = \int_{\Lambda} \exp\{i\lambda u\} f(\lambda) d\lambda, \quad u \in \mathbb{U}. \tag{1.8}$$

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

The set $\mathbb{U}$, called the time domain of $X(u)$, is the real line $\mathbb{R}$ in the continuous-time case, and the set of integers $\mathbb{Z}$ in the discrete-time case. The set Λ, called the frequency domain of $X(u)$, is $\mathbb{R}$ in the continuous-time case, and $\Lambda = \mathbb{T} = (-\pi, \pi]$ in the discrete-time case. 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 $t \to s$. This assumption is equivalent to that of the covariance function $r(u)$ be continuous at $u = 0$ (see, e.g., Cramér and Leadbetter [16], Section 5.2).

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 $f(\lambda)$.

Dependence (memory) structure of the model. 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 long-lag autocorrelations, or in the frequency domain in terms of rates of explosion of low frequency spectra (see, e.g., Beran [7], Guegan [41], Robinson [55], Giraitis et al. [36], Beran et al. [8], and references therein).

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

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

$$0 < C_1 \le f(\lambda) \le C_2 < \infty.$$

A typical short memory model example is the stationary Autoregressive Moving Average (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 discrete-time white noise, that is, a sequence of zero-mean, uncorrelated random variables with variance σ^2. Notice that the covariance $r(k)$ of (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 $h(\lambda)$ is a rational function (see, e.g., Brockwell and Davis [12], Section 3.1):

$$h(\lambda) = \frac{\sigma^2}{2\pi} \cdot \frac{|\theta_q(e^{-i\lambda})|^2}{|\psi_p(e^{-i\lambda})|^2}. \tag{1.9}$$

Discrete-time long-memory and anti-persistent models. Data in many fields of science (economics, finance, hydrology, etc.), however, is well modeled by a stationary process with *unbounded* or *vanishing* (at some fixed points) spectral density (see, e.g., Beran [7], Guegan [41], Palma [53], Taqqu [65] 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 [12], 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 discrete-time white noise defined above. The spectral density $f(\lambda)$ of $X(t)$ is given by

$$f(\lambda) = |1 - e^{-i\lambda}|^{-2d} = (2\sin(\lambda/2))^{-2d}, \quad 0 < \lambda \leq \pi, \quad 0 < d < 1/2. \tag{1.10}$$

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 discrete-time white noise, and ψ_p and θ_q are polynomials of degrees p and q, respectively. The spectral density $f(\lambda)$ of $X(t)$ is given by

$$f(\lambda) = |1 - e^{-i\lambda}|^{-2d} h(\lambda), \quad d < 1/2, \tag{1.11}$$

where $h(\lambda)$ is the spectral density of an ARMA(p, q) process, given by (1.9). Observe that for $0 < d < 1/2$ the model $X(t)$ specified by (1.11) displays long-memory, for $d < 0$ – intermediate-memory, and for $d = 0$ – short-memory. For $d \geq 1/2$ the function $f(\lambda)$ in (1.11) is not integrable, and thus it cannot represent a spectral density of a stationary process. Also, if $d \leq -1$, 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., [10, 12]).

The ARFIMA(p, d, q) processes, first introduced by Granger and Joyeux [39], and Hosking [44], 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., Henry and Zaffaroni [43] for a survey on this topic).

Another important long-memory model is the fractional Gaussian noise (fGn). To define fGn 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) $\{B_H(at), t \in \mathbb{R}\} \stackrel{d}{=} \{a^H B_H(t), t \in \mathbb{R}\}$ for any $a > 0$;

(c) $\{B_H(t + u) - B_H(u), t \in \mathbb{R}\} \stackrel{d}{=} \{B_H(t), t \in \mathbb{R}\}$ 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 $\stackrel{d}{=}$ 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 discrete-time centered Gaussian stationary process with covariance function

$$r(k) = \frac{\sigma_0^2}{2}\left[|k+1|^{2H} - |k|^{2H} - |k-1|^{2H}\right], \quad k \in \mathbb{Z} \tag{1.12}$$

and spectral density function

$$f(\lambda) = c\,|1 - e^{-i\lambda}|^2 \sum_{k=-\infty}^{\infty} |\lambda + 2\pi k|^{-(2H+1)}, \quad -\pi \le \lambda \le \pi, \tag{1.13}$$

where c is a positive constant.

It follows from (1.13) 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 (1.10) and (1.13), 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 [59] and Taqqu [65].

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 fractional Brownian motion (fBm) B_H with Hurst index H, defined above. It can be regarded as Gaussian process having a spectral density:

$$f(\lambda) = c|\lambda|^{-(2H+1)}, \quad c > 0, \;\; 0 < H < 1, \;\; \lambda \in \mathbb{R}. \tag{1.14}$$

The form (1.14) can be understood in a generalized sense: either in the sense of time-scale analysis (Flandrin [19]), or in a limiting sense (Solo [61]), since the fBm B_H is a nonstationary process (see, also, Anh et al. [3] and Gao et al. [23]).

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 continuous-time Gaussian process $X(t)$ with spectral density

$$f(\lambda) = c\,|\lambda|^{-2\alpha}(1+\lambda^2)^{-\beta}, \quad \lambda \in \mathbb{R}, 0 < c < \infty, 0 < \alpha < 1, \beta > 0. \tag{1.15}$$

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. [23]).

Notice that the process $X(t)$, specified by (1.15), 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 (1.15) 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 (1.15) 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 (1.14) and (1.15), we observe that the spectral density of fBm is the limiting case as $\beta \to 0$ that of fRBm with Hurst index $H = \alpha - 1/2$.

Remark 1.1. Recall (see Yaglom [72], Section 4.23) that a centered mean-square continuous process $\{X(t),\ t \in \mathbb{R}\}$ is said to have second-order stationary increments if its structure function:

$$D(\tau_1, \tau_2) := E\left\{[(X(s+\tau_1) - X(s)][(X(s+\tau_2) - X(s)]\right\}$$

is independent of s for all $s, \tau_1, \tau_2 \in \mathbb{R}$. Then the function $D(\tau_1, \tau_2)$ has the spectral representation

$$D(\tau_1, \tau_2) = \int_{-\infty}^{\infty} \left(e^{i\tau_1\lambda} - 1\right)\left(e^{i\tau_2\lambda} - 1\right) dF(\lambda), \tag{1.16}$$

where $F(\lambda)$ is a left-continuous real-valued non-decreasing function on $\mathbb{R}$, called *spectral distribution function* of $X(t)$, such that for any $\varepsilon > 0$

$$\int_0^{\varepsilon} \lambda^2 dF(\lambda) + \int_{\varepsilon}^{\infty} dF(\lambda) < \infty.$$

If the spectral distribution function $F(\lambda)$ possesses a derivative $f(\lambda)$ (with respect to Lebesgue measure), then f is called spectral density function of $X(t)$. In particular, if $f(\lambda)$ has the form (1.15), then the model $X(t)$ is a fractional Riesz-Bessel motion.

1.3. A link between stationary processes and the trace problem

As was 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 discrete- and continuous-time stationary processes.

To understand the relevance of the trace approximation problem to stationary processes, consider a question concerning the asymptotic distribution (as $T \to \infty$) of the following Toeplitz type quadratic functionals of a Gaussian stationary process $\{X(u),\ u \in \mathbb{U}\}$ with spectral density $f(\lambda)$, $\lambda \in \Lambda$ and covariance function $r(t) := \widehat{f}(t)$, $t \in \mathbb{U}$ (here $\mathbb{U}$ and Λ are as in Section 1.2):

$$Q_T := \begin{cases} \int_0^T \int_0^T \widehat{g}(t-s) X(t) X(s)\, dt ds & \text{in the continuous-time case} \\ \sum_{k=1}^T \sum_{j=1}^T \widehat{g}(k-j) X(k) X(j) & \text{in the discrete-time case,} \end{cases} \tag{1.17}$$

where $\widehat{g}(t) = \int_\Lambda e^{i\lambda t} g(\lambda)\, d\lambda,\ t \in \mathbb{U}$ is the Fourier transform of some real, even, integrable function $g(\lambda)$, $\lambda \in \Lambda$. We will refer $g(\lambda)$ as a *generating function* for the functional Q_T.

The form (1.17) of Q_T comes from statistical applications, where such a Q_T provides a good approximation, called Whittle approximation, of the likelihood function of the observed sample: $\{X(t), t = 1, 2, \dots, T\}$ in the discrete-time case, and $\{X(t), 0 \le t \le T\}$ in the continuous-time case (see, e.g., [17], [20], [24], [36], [37], [64], and references therein).

The limit distributions of the functionals (1.17) are completely determined by the spectral density $f(\lambda)$ and the generating function $g(\lambda)$, and depending on their properties the limit distributions can be either Gaussian (that is, Q_T with an appropriate normalization obeys central limit theorem), or non-Gaussian.

The following two questions arise naturally:

(a) Under what conditions on $f(\lambda)$ and $g(\lambda)$ will the limits be Gaussian?
(b) Describe the limit distributions, if they are non-Gaussian.

These questions will be discussed in detail in Section 4.2.

Let $A_T(f)$ be the covariance matrix (or operator) of the process $\{X(u), u \in \mathbb{U}\}$, that is, $A_T(f)$ denote either the $T \times T$ Toeplitz matrix, or the T-truncated Toeplitz operator 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 (see Section 1.1).

Our study of the asymptotic distribution of the quadratic functionals (1.17) is based on the following well–known results (see, e.g., [40, 45]):

1. The quadratic functional Q_T in (1.17) has the same distribution as the sum $\sum_{k=1}^{\infty} \lambda_k^2 \xi_k^2$ ($\sum_{k=1}^{T} \lambda_k^2 \xi_k^2$ in the discrete-time case), where $\{\xi_k, k \ge 1\}$ are independent $N(0,1)$ Gaussian random variables and $\{\lambda_k, k \ge 1\}$ are the eigenvalues of the operator $A_T(f)A_T(g)$. (Observe that the sets of non–zero eigenvalues of the operators $A_T(f)A_T(g)$, $A_T(g)A_T(f)$ and $A_T^{1/2}(f)A_T(g)A_T^{1/2}(f)$ coincide, where $A_T^{1/2}(f)$ denotes the positive definite square root of $A_T(f)$.
2. The characteristic function $\varphi(t)$ of Q_T is given by
$$\varphi(t) = \prod_{k=1}^{\infty} |1 - 2it\lambda_k|^{-1/2}. \tag{1.18}$$
3. The k–th order cumulant $\chi_k(\cdot)$ of Q_T is given by
$$\chi_k(Q_T) = 2^{k-1}(k-1)! \sum_{j=1}^{\infty} \lambda_j^k = 2^{k-1}(k-1)! \, \mathrm{tr}\, [A_T(f)A_T(g)]^k. \tag{1.19}$$

Thus, to describe the asymptotic distributions of the quadratic functionals (1.17), we have to control the corresponding traces of the products of Toeplitz matrices (or operators), yielding the trace approximation problem with generating functions specified by (1.7).

2. The trace problem for Toeplitz matrices

Let $f(\lambda)$ be an integrable real symmetric function defined on $\mathbb{T} = (-\pi, \pi]$. For $T = 1, 2, \dots$ denote by $B_T(f)$ the $(T \times T)$ Toeplitz matrix generated by the function f, that is,

$$B_T(f) := \|\widehat{f}(s-t)\|_{s,t=1,2,\ldots,T} = \begin{pmatrix} \widehat{f}(0) & \widehat{f}(-1) & \cdots & \widehat{f}(1-T) \\ \widehat{f}(1) & \widehat{f}(0) & \cdots & \widehat{f}(2-T) \\ \cdots & \cdots & \cdots & \cdots \\ \widehat{f}(T-1) & \widehat{f}(T-2) & \cdots & \widehat{f}(0) \end{pmatrix}, \tag{2.1}$$

where

$$\widehat{f}(t) = \int_{\mathbb{T}} e^{i\lambda t} f(\lambda)\, d\lambda, \quad t \in \mathbb{Z}, \tag{2.2}$$

are the Fourier coefficients of f.

Observe that

$$\frac{1}{T}\text{tr}\,[B_T(f)] = \frac{1}{T}\cdot T\widehat{f}(0) = \int_{-\pi}^{\pi} f(\lambda) d\lambda. \tag{2.3}$$

What happens to the relation (2.3) when the matrix $B_T(f)$ is replaced by a product of Toeplitz matrices? Observe that the product of Toeplitz matrices is not a Toeplitz matrix.

The idea is to approximate the trace of the product of Toeplitz matrices by the trace of a Toeplitz matrix generated by the product of the generating functions. More precisely, let $\mathcal{H} = \{h_1, h_2, \ldots, h_m\}$ be a collection of integrable real symmetric functions defined on $\mathbb{T}$. Define

$$S_{B,\mathcal{H}}(T) := \frac{1}{T}\text{tr}\left[\prod_{i=1}^{m} B_T(h_i)\right], \quad M_{\mathbb{T},\mathcal{H}} := (2\pi)^{m-1}\int_{-\pi}^{\pi}\left[\prod_{i=1}^{m} h_i(\lambda)\right] d\lambda, \tag{2.4}$$

and let

$$\Delta(T) := \Delta_{B,\mathbb{T},\mathcal{H}}(T) = |S_{B,\mathcal{H}}(T) - M_{\mathbb{T},\mathcal{H}}|. \tag{2.5}$$

Observe that by (2.3)

$$M_{\mathbb{T},\mathcal{H}} = (2\pi)^{m-1}\int_{-\pi}^{\pi}\left[\prod_{i=1}^{m} h_i(\lambda)\right] d\lambda = \frac{1}{T}\text{tr}\left[B_T\left(\prod_{i=1}^{m} h_i(\lambda)\right)\right]. \tag{2.6}$$

How well is $S_{B,\mathcal{H}}(T)$ approximated by $M_{\mathbb{T},\mathcal{H}}$? What is the rate of convergence to zero of approximation error $\Delta_{B,\mathbb{T},\mathcal{H}}(T)$ as $T \to \infty$? These are Problems (A) and (B), stated in Section 1.1.

Note that the Problems (A) and (B) are important not only for their theoretical interest (see, e.g., [11, 40, 49]), but also because of their applications. For instance, in view of formulas (1.18) and (1.19), to describe the asymptotic distributions of the quadratic functionals (1.17), and then to apply to statistical estimation problems, we have to control the traces of the products of Toeplitz matrices $B_T(f)$ and $B_T(g)$ (see, also, [17, 20, 36, 37, 64]).

2.1. Problem (A) for Toeplitz matrices

Recall that Problem (A) involves finding conditions on the functions $h_1(\lambda)$, $h_2(\lambda), \dots, h_m(\lambda)$ in (2.4) such that $\Delta_{B,\mathbb{T},\mathcal{H}}(T) = o(1)$ as $T \to \infty$.

In Theorem 2.1 and Remark 2.2 we summarize the results concerning Problems (A) for Toeplitz matrices in the case where the exponents τ_k, $k = 1, 2, \dots, m$ (see (1.4)) are all equal to 1 as in (2.4).

Theorem 2.1. *Let $\Delta_{B,\mathbb{T},\mathcal{H}}(T)$ be as in (2.5). Each of the following conditions is sufficient for*

$$\Delta_{B,\mathbb{T},\mathcal{H}}(T) = o(1) \quad \text{as} \quad T \to \infty. \tag{2.7}$$

(A1) *$h_i \in L^{p_i}(\mathbb{T})$, $1 \le p_i \le \infty$, $i = 1, 2, \dots, m$, with $1/p_1 + \cdots + 1/p_m \le 1$.*
(A2) *The function $\varphi(\mathbf{u})$ given by*

$$\varphi(\mathbf{u}) := \int_{-\pi}^{\pi} h_1(\lambda) h_2(\lambda - u_1) h_3(\lambda - u_2) \cdots h_m(\lambda - u_{m-1})\, d\lambda, \tag{2.8}$$

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

Remark 2.1. Assertion (A1) was proved by Avram [4]. For the special case $p_i = \infty$, $i = 1, 2, \dots, m$, that is, when all h_i are bounded functions, it was first established by Grenander and Szegö ([40], Sec. 7.4). For $m = 4$; $p_1 = p_3 = 2$; $p_2 = p_4 = \infty$, (A1) was proved by Ibragimov [45] and Rosenblatt [57].

Assertion (A2), for $m = 4$, $h_1 = h_3 := f$ and $h_2 = h_4 := g$ was proved in Ginovyan and Sahakyan [32].

Remark 2.2. For the special case $m = 4$, $h_1 = h_3 := f$ and $h_2 = h_4 := g$, in Giraitis and Surgailis [37] (see also Giraitis et al. [36]), and in Ginovyan and Sahakyan [32], it was proved that the following conditions are also sufficient for (2.7):

(A3) (Giraitis and Surgailis [37]). $f \in L^2(\mathbb{T})$, $g \in L^2(\mathbb{T})$, $fg \in L^2(\mathbb{T})$ and

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

(A4) (Ginovyan and Sahakyan [32]). The functions f and g satisfy

$$f(\lambda) \le |\lambda|^{-\alpha} L_1(\lambda) \quad \text{and} \quad |g(\lambda)| \le |\lambda|^{-\beta} L_2(\lambda) \quad \text{for} \quad \lambda \in [-\pi, \pi],$$

for some $\alpha < 1$, $\beta < 1$ with $\alpha + \beta \le 1/2$, and $L_i \in SV(\mathbb{R})$, $\lambda^{-(\alpha+\beta)} L_i(\lambda) \in L^2(\mathbb{T})$, $i = 1, 2$, where $SV(\mathbb{R})$ is the class of slowly varying at zero functions $u(\lambda)$, $\lambda \in \mathbb{R}$, namely $\lim_{a \to 0} \frac{u(a\lambda)}{u(\lambda)} = 1$ for all $a > 0$, satisfying also $u(\lambda) \in L^\infty(\mathbb{R})$, $\lim_{\lambda \to 0} u(\lambda) = 0$, $u(\lambda) = u(-\lambda)$ and $0 < u(\lambda) < u(\mu)$ for $0 < \lambda < \mu$.

Remark 2.3. Case (A4), with $\alpha + \beta < 1/2$, was first obtained by Fox and Taqqu [20].

Remark 2.4. It would be of interest to extend the results of (A3) and (A4) to arbitrary $m > 4$.

We now consider the case when the product in (1.4) involves also *inverse matrices*, that is, $\tau_k = (-1)^k$, $k = 1, 2, \dots, m$. We assume that $m = 2\nu$, and the functions from the collection $\mathcal{H} = \{h_1, h_2, \dots, h_m\}$ that involve Toeplitz matrices we denote by g_i, $i = 1, 2, \dots, \nu$, while those involving inverse Toeplitz matrices we denote by f_i, $i = 1, 2, \dots, \nu$. We set

$$SI_{B,\mathcal{H}}(T) := \frac{1}{T}\mathrm{tr}\left[\prod_{i=1}^{\nu}[B_T(f_i)]^{-1}B_T(g_i)\right], \tag{2.9}$$

$$MI_{\mathbb{T},\mathcal{H}} := \frac{1}{2\pi}\int_{-\pi}^{\pi}\left[\prod_{i=1}^{\nu}\frac{g_i(\lambda)}{f_i(\lambda)}\right]d\lambda, \tag{2.10}$$

$$\Delta I_{B,\mathbb{T},\mathcal{H}}(T) := |SI_{B,\mathcal{H}}(T) - MI_{\mathbb{T},\mathcal{H}}|. \tag{2.11}$$

The following theorem was proved by Dahlhaus (see [17], Theorem 5.1).

Theorem 2.2. *Let $\nu \in \mathbb{N}$, and $\alpha, \beta \in \mathbb{R}$ with $0 < \alpha, \beta < 1$ and $\nu(\beta - \alpha) < 1/2$. Suppose $f_i(\lambda)$ and $g_i(\lambda)$, $i = 1, 2, \dots, \nu$, are symmetric real valued functions satisfying the conditions:*

(C1) *$f_i(\lambda)$, $i = 1, 2, \dots, \nu$, are nonnegative and continuous at all $\lambda \in \mathbb{T} \setminus \{0\}$, $f_i^{-1}(\lambda)$ are continuous at all $\lambda \in \mathbb{T}$, and $f_i(\lambda) = O\left(|\lambda|^{-\alpha-\delta}\right)$ as $\lambda \to 0$ for all $\delta > 0$ and $i = 1, 2, \dots, \nu$;*
(C2) *$\partial/(\partial\lambda) f_i^{-1}(\lambda)$ and $\partial^2/(\partial\lambda)^2 f_i^{-1}(\lambda)$ are continuous at all $\lambda \in \mathbb{T} \setminus \{0\}$, and for all $\delta > 0$ and $i = 1, 2, \dots, \nu$*

$$\partial^k/(\partial\lambda)^k f_i^{-1}(\lambda) = O\left(|\lambda|^{-\alpha-k-\delta}\right) \quad \text{for} \quad k = 0, 1.$$

(C3) *$g_i(\lambda)$ are continuous at all $\lambda \in \mathbb{T} \setminus \{0\}$ and $g_i(\lambda) = O\left(|\lambda|^{-\beta-\delta}\right)$ as $\lambda \to 0$ for all $\delta > 0$ and $i = 1, 2, \dots, \nu$.*

Then

$$\Delta I_{B,\mathbb{T},\mathcal{H}}(T) = o(1) \quad \text{as} \quad T \to \infty.$$

2.2. Problem (B) for Toeplitz matrices

Recall that Problem (B) involves finding conditions on the functions $h_1(\lambda)$, $h_2(\lambda), \dots, h_m(\lambda)$ in (2.4) such that $\Delta_{B,\mathbb{T},\mathcal{H}}(T) = O(T^{-\gamma})$ as $T \to \infty$ for some $\gamma > 0$.

In Theorem 2.3 below we summarize the results concerning Problem (B) for Toeplitz matrices in the case where $\tau_k = 1$, $k = 1, 2, \dots, m$. First we introduce some classes of functions (see, e.g., [14, 62, 64]). Recall that $\mathbb{T} = (-\pi, \pi]$ and denote

$$\mathcal{F}_1(\mathbb{T}) := \left\{f \in L^1(\mathbb{T}) : \sum_{k=-\infty}^{\infty}|k||\widehat{f}(k)| < \infty;\right\}, \tag{2.12}$$

where $\widehat{f}(k)=\int_{\mathbb{T}} e^{i\lambda k} f(\lambda)\,d\lambda$, $\widehat{f}(-k)=\widehat{f}(k)$, and

$$\mathcal{F}_2(\mathbb{T})=\mathcal{F}_{ARMA}(\mathbb{T}):=\left\{f:\ f(\lambda)=\frac{\sigma^2}{2\pi}\frac{|\theta_q(e^{-i\lambda})|^2}{|\psi_p(e^{-i\lambda})|^2}\right\},\tag{2.13}$$

where $0<\sigma^2<\infty$, $\theta_q(z):=\sum_{k=0}^{q} a_k z^k$ $(q\in\mathbb{N})$ and $\psi_p(z):=\sum_{k=0}^{p} b_k z^k$ $(p\in\mathbb{N})$ are both bounded away from zero for $|z|\le 1$.

Remark 2.5. The following implications were established in [62]:

(a) If $f_1, f_2\in\mathcal{F}_1(\mathbb{T})$, then $f_1f_2\in\mathcal{F}_1(\mathbb{T})$.
(b) If $f\in\mathcal{F}_2(\mathbb{T})$, then $f\in\mathcal{F}_1(\mathbb{T})$ and $f^{-1}\in\mathcal{F}_2(\mathbb{T})$.

For $\psi\in L^p(\mathbb{T})$, $1\le p\le\infty$, let $\omega_p(\psi,\delta)$ denote the L^p–modulus of continuity of ψ:

$$\omega_p(\psi,\delta):=\sup_{0<h\le\delta}\|\psi(\cdot+h)-\psi(\cdot)\|_p,\quad \delta>0.$$

Definition 2.1. Given numbers $0<\gamma\le 1$ and $1\le p\le\infty$, we denote by $\mathrm{Lip}(\mathbb{T};p,\gamma)$ the L^p-Lipschitz class of functions defined on $\mathbb{T}$ (see, e.g., [14]):

$$\mathrm{Lip}(\mathbb{T};p,\gamma)=\{\psi(\lambda)\in L^p(\mathbb{T});\quad \omega_p(\psi;\delta)=O(\delta^\gamma),\quad \delta\to 0\}.$$

Observe that if $\psi\in\mathrm{Lip}(p,\gamma)$, then there exists a constant C such that $\omega_p(\psi;\delta)\le C\,\delta^\gamma$ for all $\delta>0$.

Theorem 2.3. *Assume that $\tau_k=1$, $k=1,2,\dots,m$, $\mathcal{H}=\{h_1,h_2,\dots,h_m\}$, and let $\Delta_{B,\mathbb{T},\mathcal{H}}(T)$ be as in (2.5). The following assertions hold:*

(B1) *If $h_i\in\mathcal{F}_1(\mathbb{T})$, $i=1,2,\dots,m$, then*

$$\Delta_{B,\mathbb{T},\mathcal{H}}(T)=O(T^{-1})\quad\text{as } T\to\infty.$$

(B2) *If the functions $h_i(\lambda)$, $i=1,2,\dots,m$, have uniformly bounded derivatives on $\mathbb{T}:=(-\pi,\pi]$, then for any $\epsilon>0$*

$$\Delta_{B,\mathbb{T},\mathcal{H}}(T)=O(T^{-1+\epsilon})\quad \textit{as } T\to\infty.$$

(B3) *Assume that the function $\varphi(\mathbf{u})$ given by (2.8) with some constants $C>0$ and $\gamma\in(0,1]$ satisfies*

$$|\varphi(\mathbf{u})-\varphi(\mathbf{0})|\le C|\mathbf{u}|^\gamma,\quad \mathbf{u}=(u_1,u_2,\dots,u_{m-1})\in\mathbb{T}^{m-1},$$

where $\mathbf{0}=(0,0,\dots,0)$ and $|\mathbf{u}|=|u_1|+|u_2|+\cdots+|u_{m-1}|$. Then for any $\varepsilon>0$

$$\Delta_{B,\mathbb{T},\mathcal{H}}(T)=O\left(T^{-\gamma+\varepsilon}\right)\quad \textit{as}\quad T\to\infty.\tag{2.14}$$

(B4) *Let $h_i(\lambda)\in\mathrm{Lip}(\mathbb{T};p_i,\gamma)$, $p_i>1$, $i=1,2,\dots,m$, $1/p_1+\cdots+1/p_m\le 1$ and $\gamma\in(0,1]$. Then (2.14) holds for any $\varepsilon>0$.*

(B5) *Let $h_i(\lambda)$, $i = 1, 2, \ldots, m$, be differentiable functions defined on $\mathbb{T} \setminus \{0\}$, such that for some constants $C_{1i} > 0$, $C_{2i} > 0$, and α_i, $i = 1, 2, \ldots, m$, satisfying $0 < \alpha_i < 1$, $\alpha := \sum_{i=1}^m \alpha_i < 1$*

$$|h_i(\lambda)| \le C_{1i}|\lambda|^{-\alpha_i}, \ |h_i'(\lambda)| \le C_{2i}|\lambda|^{-(\alpha_i+1)}, \quad \lambda \in \mathbb{T}\setminus\{0\}, \ i = 1, 2, \ldots, m.$$

Then (2.14) holds for any $\varepsilon > 0$ with

$$\gamma = \frac{1}{m}(1-\alpha). \tag{2.15}$$

Remark 2.6. Assertion (B1) was proved in Taniguchi [62] (see, also, [64]). Assertion (B2), which is weaker than (B1), but holding under weaker conditions than those in (B1), was proved in Lieberman and Phillips [51]. Assertions (B3)–(B5) for $m = 4$ were proved in Ginovyan and Sahakyan [35].

Remark 2.7. It is easy to see that under the conditions of **(B2)** we have $h_i \in \mathrm{Lip}(\mathbb{T}; p, 1)$ for any $i = 1, 2, \ldots, m$ and $p \ge 1$. Hence **(B4)** implies **(B2)**.

Example 2.1. Let $h_i(\lambda) = |\lambda|^{-\alpha_i}$, $\lambda \in [-\pi, \pi]$, $i = 1, 2, \ldots, m$, with $0 < \alpha_i < 1$ and $\alpha := \sum_{i=1}^m \alpha_i < 1$. It is easy to see that the conditions of **(B5)** are satisfied, and hence we have (2.14) with γ as in (2.15).

The next results (cf. Ginovyan [28]) show that for special case $m = 2$ the rates in Theorem 2.3 (B4) and (B5) can be substantially improved.

Theorem 2.4. *Let $h_i(\lambda) \in \mathrm{Lip}(\mathbb{T}; p_i, \gamma_i)$ with $p_i > 1$, $1/p_1 + 1/p_2 = 1$ and $\gamma_i \in (0, 1]$, $i = 1, 2$, and let*

$$\Delta_{2,B}(T) := \left| \frac{1}{T} \mathrm{tr}[B_T(h_1)B_T(h_2)] - 2\pi \int_{\mathbb{T}} h_1(\lambda)h_2(\lambda)\, d\lambda \right|.$$

Then

$$\Delta_{2,B}(T) = \begin{cases} O(T^{-(\gamma_1+\gamma_2)}), & \text{if} \quad \gamma_1 + \gamma_2 < 1 \\ O(T^{-1}\ln T), & \text{if} \quad \gamma_1 + \gamma_2 = 1 \\ O(T^{-1}), & \text{if} \quad \gamma_1 + \gamma_2 > 1. \end{cases}$$

Theorem 2.5. *Assume that the functions $h_i(\lambda)$, $i = 1, 2$, satisfy the conditions of Theorem 2.3 (B5) with $m = 2$. Then*

$$\Delta_{2,B}(T) = O\left(T^{-1+(\alpha_1+\alpha_2)}\right) \quad \text{as} \quad T \to \infty. \tag{2.16}$$

The next result, due to Taniguchi [62] (see, also, [64]), concerns the case when the product in (1.4) involves also inverse matrices, that is, $\tau_k = (-1)^k$, $k = 1, 2, \ldots, m$.

Theorem 2.6. *Let $SI_{B,\mathcal{H}}(T)$, $MI_{\mathbb{T},\mathcal{H}}$ and $\Delta I_{B,\mathbb{T},\mathcal{H}}(T)$ be as in (2.9)–(2.11), and let $\mathcal{F}_1(\mathbb{T})$ and $\mathcal{F}_2(\mathbb{T})$ be as in (2.12) and (2.13), respectively. If $f_i \in \mathcal{F}_2$ and $g_i \in \mathcal{F}_1(\mathbb{T})$, $i = 1, 2, \ldots, \nu$, then*

$$\Delta I_{B,\mathbb{T},\mathcal{H}}(T) = O(T^{-1}) \quad \text{as } T \to \infty. \tag{2.17}$$

3. The trace problem for Toeplitz operators

In this section we consider Problems (A) and (B) for Toeplitz operators, that is, in the case where the generating functions are defined on the real line. Again, Problem (A) involves $o(1)$ approximation and Problem (B) involves $O(T^{-\gamma})$ approximation with $\gamma > 0$. The theorems in this section are proved in Section 5.

Let $f(\lambda)$ be an integrable real symmetric function defined on $\mathbb{R}$. The analogue of the Fourier coefficients $\widehat{f}(k)$ in (2.2) is the Fourier transform $\widehat{f}(t)$ of $f(\lambda)$:

$$\widehat{f}(t) = \int_{-\infty}^{+\infty} e^{i\lambda t} f(\lambda)\, d\lambda, \quad t \in \mathbb{R}. \tag{3.1}$$

The $\widehat{f}$ in (3.1) will play the role of kernel in an integral operator.

Given $T > 0$ and an integrable real symmetric function $f(\lambda)$ defined on $\mathbb{R}$, the *T-truncated Toeplitz operator* generated by $f(\lambda)$, denoted by $W_T(f)$, is defined by the following equation (see, e.g., [33, 40, 45]):

$$[W_T(f)u](t) = \int_0^T \hat{f}(t-s)u(s)ds, \quad u(s) \in L^2[0,T], \tag{3.2}$$

where $\widehat{f}$ is as in (3.1).

It follows from (3.1), (3.2) and the formula for traces of integral operators (see, e.g., Gohberg and Krein [38], p. 114) that

$$\mathrm{tr}\,[W_T(f)] = \int_0^T \hat{f}(t-t)dt = T\hat{f}(0) = T\int_{-\infty}^{+\infty} f(\lambda)d\lambda. \tag{3.3}$$

We pose the same question as in the case of Toeplitz matrices: what happens to the relation (3.3) when the single operator $W_T(f)$ is replaced by a product of such operators? Observe that the product of Toeplitz operators again is not a Toeplitz operator.

The approach is similar to that of Toeplitz matrices – to approximate the trace of the product of Toeplitz operators by the trace of a Toeplitz operator generated by the product of generating functions. More precisely, let $\mathcal{H} = \{h_1, h_2, \ldots, h_m\}$ be a collection of integrable real symmetric functions defined on $\mathbb{R}$. Define

$$S_{W,\mathcal{H}}(T) := \frac{1}{T}\mathrm{tr}\left[\prod_{i=1}^m W_T(h_i)\right], \quad M_{\mathbb{R},\mathcal{H}} := (2\pi)^{m-1}\int_{-\infty}^{\infty}\left[\prod_{i=1}^m h_i(\lambda)\right] d\lambda, \tag{3.4}$$

and let

$$\Delta(T) := \Delta_{W,\mathbb{R},\mathcal{H}}(T) = |S_{W,\mathcal{H}}(T) - M_{\mathbb{R},\mathcal{H}}|. \tag{3.5}$$

Observe that by (3.3),

$$M_{\mathbb{R},\mathcal{H}} = (2\pi)^{m-1}\int_{-\infty}^{\infty}\left[\prod_{i=1}^m h_i(\lambda)\right] d\lambda = \frac{1}{T}\mathrm{tr}\left[W_T\left(\prod_{i=1}^m h_i(\lambda)\right)\right]. \tag{3.6}$$

How well is $S_{W,\mathcal{H}}(T)$ approximated by $M_{\mathbb{R},\mathcal{H}}$? What is the rate of convergence to zero of approximation error $\Delta_{W,\mathbb{R},\mathcal{H}}(T)$ as $T \to \infty$? These are Problems (A) and (B) in this case.

3.1. Problem (A) for Toeplitz operators

In Theorem 3.1 and Remark 3.1 we summarize the results concerning Problem (A) for Toeplitz operators in the case where $\tau_k = 1$, $k = 1, 2, \ldots, m$.

Theorem 3.1. *Let $\Delta(T) := \Delta_{W,\mathbb{R},\mathcal{H}}(T)$ be as in (3.5). Each of the following conditions is sufficient for*

$$\Delta(T) = o(1) \quad \text{as} \quad T \to \infty. \tag{3.7}$$

(A1) $h_i \in L^1(\mathbb{R}) \cap L^{p_i}(\mathbb{R})$, $p_i > 1$, $i = 1, 2, \ldots, m$, *with* $1/p_1 + \cdots + 1/p_m \leq 1$.
(A2) *The function* $\varphi(\mathbf{u})$ *defined by*

$$\varphi(\mathbf{u}) := \int_{-\infty}^{+\infty} h_1(\lambda) h_2(\lambda - u_1) h_3(\lambda - u_2) \cdots h_m(\lambda - u_{m-1})\, d\lambda, \tag{3.8}$$

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

Remark 3.1. For the special case $m = 4$, $h_1 = h_3 := f$ and $h_2 = h_4 := g$, Ginovyan and Sahakyan [33] proved that the following conditions are also sufficient for (3.7):

(A3) $f \in L^1(\mathbb{R}) \cap L^2(\mathbb{R}),\ g \in L^1(\mathbb{R}) \cap L^2(\mathbb{R}),\ fg \in L^2(\mathbb{R})$ and

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

(A4) The functions f and g are integrable on $\mathbb{R}$, bounded on $\mathbb{R} \setminus (-\pi, \pi)$, and satisfy

$$f(\lambda) \leq |\lambda|^{-\alpha} L_1(\lambda) \quad \text{and} \quad |g(\lambda)| \leq |\lambda|^{-\beta} L_2(\lambda) \quad \text{for} \quad \lambda \in [-\pi, \pi],$$

for some $\alpha < 1,\ \beta < 1$ with $\alpha + \beta \leq 1/2$, and $L_i \in SV(\mathbb{R}),\ \lambda^{-(\alpha+\beta)} L_i(\lambda) \in L^2(\mathbb{T}),\ i = 1, 2$, where $SV(\mathbb{R})$ is the class of slowly varying at zero functions $u(\lambda)$, $\lambda \in \mathbb{R}$, satisfying $u(\lambda) \in L^\infty(\mathbb{R})$, $\lim_{\lambda \to 0} u(\lambda) = 0$, $u(\lambda) = u(-\lambda)$ and $0 < u(\lambda) < u(\mu)$ for $0 < \lambda < \mu$.

Remark 3.2. It would be of interest to extend the results of (A3) and (A4) to arbitrary $m > 4$.

3.2. Problem (B) for Toeplitz operators

In Theorem 3.2 below we summarize the results concerning Problem (B) for Toeplitz operators in the case where $\tau_k = 1$, $k = 1, 2, \ldots, m$. Let

$$\mathcal{F}_1(\mathbb{R}) := \left\{ f \in L^1(\mathbb{R}) : \int_{-\infty}^{\infty} |t| |\widehat{f}(t)| dt < \infty \right\}, \tag{3.9}$$

where $\widehat{f}(t)=\int_{\mathbb{R}} e^{i\lambda t} f(\lambda)\,d\lambda$, $\widehat{f}(-t)=\widehat{f}(t)$.

For $\psi \in L^p(\mathbb{R})$, $1\le p\le\infty$ let $\omega_p(\psi,\delta)$ denote the L^p–modulus of continuity of ψ:

$$\omega_p(\psi,\delta) := \sup_{0<h\le\delta} \|\psi(\cdot+h)-\psi(\cdot)\|_p, \quad \delta>0.$$

Given numbers $0<\gamma\le 1$ and $1\le p\le\infty$, we denote by $\mathrm{Lip}(\mathbb{R};p,\gamma)$ the L^p-Lipschitz class of functions defined on $\mathbb{R}$ (see, e.g., [14]):

$$\mathrm{Lip}(\mathbb{R};p,\gamma)=\{\psi(\lambda)\in L^p(\mathbb{R}) \ : \ \omega_p(\psi;\delta)=O(\delta^\gamma) \text{ as } \delta\to 0\}.$$

Theorem 3.2 is the continuous version of Theorem 2.3.

Theorem 3.2. *Let $\mathcal{H}=\{h_1,h_2,\ldots,h_m\}$, and $\Delta_{W,\mathbb{R},\mathcal{H}}(T)$ and $\varphi(\mathbf{u})$ be as in (3.5) and (3.8), respectively. The following assertions hold:*

(B1) *If $h_i\in\mathcal{F}_1(\mathbb{R})$, $i=1,2,\ldots,m$, then*

$$\Delta_{W,\mathbb{R},\mathcal{H}}(T)=O(T^{-1}) \quad \text{as} \quad T\to\infty. \tag{3.10}$$

(B2) *Assume that $\varphi(\mathbf{u})\in L^\infty(\mathbb{R}^{m-1})$ and with some constants $C>0$ and $\gamma\in(0,1]$*

$$|\varphi(\mathbf{u})-\varphi(\mathbf{0})|\le C|\mathbf{u}|^\gamma, \quad \mathbf{u}=(u_1,u_2,\ldots,u_{m-1})\in\mathbb{R}^{m-1}, \tag{3.11}$$

where $\mathbf{0}=(0,0,\ldots,0)$ and $|\mathbf{u}|=|u_1|+|u_2|+\cdots+|u_{m-1}|$. Then for any $\varepsilon>0$

$$\Delta_{W,\mathbb{R},\mathcal{H}}(T)=O\left(T^{-\gamma+\varepsilon}\right) \quad \text{as} \quad T\to\infty. \tag{3.12}$$

(B3) *Let $h_i(\lambda)\in\mathrm{Lip}(\mathbb{R};p_i,\gamma)$, $i=1,2,\ldots,m$, $1/p_1+\cdots+1/p_m\le 1$ and $\gamma\in(0,1]$. Then (3.12) holds for any $\varepsilon>0$.*

(B4) *Let $h_i(\lambda)$, $i=1,2,\ldots,m$, be differentiable functions defined on $\mathbb{R}\setminus\{0\}$, such that for some constants $C_i>0$, $\sigma_i>0$ and $\delta_i>1$, $i=1,2,\ldots,m$ with $\sigma:=\sum_{i=1}^m\sigma_i<1$*

$$|h_i(\lambda)|\le\begin{cases} C_i|\lambda|^{-\sigma_i} & \text{if} \quad |\lambda|\le 1\\ C_i|\lambda|^{-\delta_i} & \text{if} \quad |\lambda|>1\end{cases}, \quad |h_i'(\lambda)|\le\begin{cases} C_i|\lambda|^{-\sigma_i-1} & \text{if} \quad |\lambda|\le 1\\ C_i|\lambda|^{-\delta_i-1} & \text{if} \quad |\lambda|>1\end{cases} \tag{3.13}$$

for all $i=1,2,\ldots,m$. Then for any $\varepsilon>0$

$$\Delta_{W,\mathbb{R},\mathcal{H}}(T)=O\left(T^{-\gamma+\varepsilon}\right) \quad \text{as} \quad T\to\infty \tag{3.14}$$

with

$$\gamma=\frac{1}{m}(1-\sigma). \tag{3.15}$$

The next results, which are continuous versions of Theorems 2.4 and 2.5, respectively, show that for the special case $m=2$, the rates in Theorem 3.2 (B3) and (B4) can be substantially improved.

Theorem 3.3. *Let* $h_i(\lambda) \in \mathrm{Lip}(\mathbb{R}; p_i, \gamma_i)$ *with* $p_i > 1$, $1/p_1 + 1/p_2 = 1$ *and* $\gamma_i \in (0, 1]$, $i = 1, 2$, *and let*

$$\Delta_{2,W}(T) := \left| \frac{1}{T} \mathrm{tr}[W_T(h_1) W_T(h_2)] - 2\pi \int_{\mathbb{R}} h_1(\lambda) h_2(\lambda)\, d\lambda \right|.$$

Then

$$\Delta_{2,W}(T) = \begin{cases} O(T^{-(\gamma_1+\gamma_2)}), & \text{if} \quad \gamma_1 + \gamma_2 < 1 \\ O(T^{-1} \ln T), & \text{if} \quad \gamma_1 + \gamma_2 = 1 \\ O(T^{-1}), & \text{if} \quad \gamma_1 + \gamma_2 > 1. \end{cases}$$

Theorem 3.4. *Assume that the functions* $h_i(\lambda)$, $i = 1, 2$, *satisfy the conditions of Theorem 3.2 (B4) with* $m = 2$. *Then*

$$\Delta_{2,W}(T) = O\left(T^{-1+(\sigma_1+\sigma_2)}\right) \quad as \quad T \to \infty. \tag{3.16}$$

Remark 3.3. It would be of interest to prove the continuous analogs of Dahlhaus theorem (Theorem 2.2) and Taniguchi theorem (Theorem 2.6) for Toeplitz operators.

4. Applications to stationary processes

In this section we provide some applications of the trace problem to discrete- and continuous-time stationary processes: ARFIMA and Fractional Riesz-Bessel motions; central and non-central limit theorems, Berry-Esséen bounds, and large deviations for Toeplitz quadratic forms and functionals.

4.1. Applications to ARFIMA time series and fractional Riesz-Bessel motions

In this subsection we apply the results of Sections 2 and 3 to the important special cases where the generating functions are spectral densities of a discrete-time ARFIMA$(0, d, 0)$ stationary processes or continuous-time stationary fractional Riesz-Bessel motions.

We use the following notation: $m = 2\nu$;

$$h_1(\lambda) = h_3(\lambda) = \cdots = h_{2\nu-1}(\lambda) := f_1(\lambda)$$
$$h_2(\lambda) = h_4(\lambda) = \cdots = h_{2\nu}(\lambda) := f_2(\lambda);$$

and

$$S_{\nu,A}(T) = \frac{1}{T} \mathrm{tr}[A_T(f_1) A_T(f_2)]^\nu, \tag{4.1}$$

$$\Delta_{\nu,A}(T) := \left| S_{\nu,A}(T) - (2\pi)^{2\nu-1} \int_\Lambda [f_1(\lambda) f_2(\lambda)]^\nu\, d\lambda \right|, \tag{4.2}$$

where either $A_T(f_i) = B_T(f_i)$ and $\Lambda = \mathbb{T}$ or $A_T(f_i) = W_T(f_i)$ and $\Lambda = \mathbb{R}$, $i = 1, 2$.

4.1.1. Applications to ARFIMA time series

The next theorem gives an error bound for $\Delta_{2,B}(T)$ in the case where the corresponding Toeplitz matrices are generated by spectral densities of two discrete-time ARFIMA$(0,d,0)$ stationary processes.

Theorem 4.1. *Let $f_i(\lambda)$, $i=1,2$, be the spectral density functions of two ARFIMA$(0,d,0)$ stationary processes defined as*

$$f_i(\lambda) = \frac{\sigma_i^2}{2\pi}\left|1-e^{-i\lambda}\right|^{-2d_i}, \quad i=1,2 \tag{4.3}$$

with $0<\sigma_i^2<\infty$ and $0<d_i<1/2$. Then under $d:=d_1+d_2<1/(2\nu)$, $\nu\in\mathbb{N}$, for any $\varepsilon>0$,

$$\Delta_{\nu,B}(T) = O\left(T^{-\gamma+\varepsilon}\right) \quad as \quad T\to\infty \tag{4.4}$$

with

$$\gamma = \frac{1}{2\nu}-(d_1+d_2). \tag{4.5}$$

Proof. Assuming that $\lambda\in(0,\pi]$ (the case $\lambda\in[-\pi,0)$ is treated similarly), and taking into account $|1-e^{-i\lambda}| = 2\sin(\lambda/2)$, we have for $i=1,2$

$$\begin{aligned} f_i(\lambda) &= \frac{\sigma_i^2}{2\pi}\cdot 2^{-2d_i}\left[\sin\frac{\lambda}{2}\right]^{-2d_i}, \\ f_i'(\lambda) &= \frac{\sigma_i^2}{2\pi}\cdot\left[-2d_i 2^{-2d_i-1}\left(\sin\frac{\lambda}{2}\right)^{-2d_i-1}\cos\frac{\lambda}{2}\right]. \end{aligned} \tag{4.6}$$

It is clear that the conditions of Theorem 2.3 **(B5)** are satisfied with $\alpha_i=2d_i$ and $C_{1i}=C_{2i}=\sigma_i^2$, $i=1,2$, and the result follows. ☐

The next theorem, which was proved in Lieberman and Phillips [51], gives an explicit second-order asymptotic expansion for $S_{1,B}(T)$ in the case where the Toeplitz matrices are generated by the spectral densities given by (4.3), and shows that in this special case a second-order asymptotic expansion successfully removes the singularity and delivers a substantially improved approximation.

Theorem 4.2. *Let $f_i(\lambda)$, $i=1,2$, be the spectral density functions of two ARFIMA$(0,d,0)$ stationary processes defined by (4.3) with $0<\sigma_i^2<\infty$ and $0<d_i<1/2$, $i=1,2$. Then under $d:=d_1+d_2<1/2$*

$$\begin{aligned} S_{1,B}(T) &:= \frac{1}{T}\mathrm{tr}[B_T(f_1)B_T(f_2)] \\ &= 2\pi\int_{-\pi}^{\pi} f_1(\lambda)f_2(\lambda)\,d\lambda - \frac{C(d_1,d_2)}{T^{1-2d}} + o\left(\frac{1}{T^{1-2d}}\right) \end{aligned} \tag{4.7}$$

as $T\to\infty$, where

$$C(d_1,d_2) = \frac{2\sigma_1^2\sigma_2^2\pi^2}{\cos(\pi d_1)\cos(\pi d_2)\Gamma(2d_1)\Gamma(2d_2)}\cdot\frac{1}{2d(1-2d)}. \tag{4.8}$$

Remark 4.1. The asymptotic relation (4.7) in Lieberman and Phillips [51] was established by direct calculations using the explicit forms of functions f_i given by (4.3). On the other hand, as it follows from (4.6), the functions f_i $(i = 1, 2)$ satisfy conditions of Theorem 2.5 with $\alpha_i = 2d_i$ $(i = 1, 2)$, and hence (4.7) is a special case of Theorem 2.5.

4.1.2. Applications to fractional Riesz-Bessel motions

Now we assume that the underlying model is a fractional Riesz-Bessel motion with spectral density given by (1.15). The following result is an immediate consequence of Theorem 3.1 **(A1)**.

Theorem 4.3. *Let $f_1(\lambda) = f(\lambda)$ be the spectral density of a fractional Riesz-Bessel motion defined by (1.15), and let $f_2(\lambda) = g(\lambda)$ be an integrable real symmetric function on $\mathbb{R}$. If for some $p, q \geq 1$ with $1/p + 1/q \leq 1/\nu$ $(\nu \in \mathbb{N})$ we have $g(\lambda) \in L^q(\mathbb{R})$ and $0 < \alpha < 1/(2p)$, $\alpha + \beta > 1/2$, then*

$$\Delta_{\nu,W}(T) = o(1) \quad \text{as} \quad T \to \infty.$$

Theorem 4.4. *Let $f_1(\lambda) = f(\lambda)$ be as in (1.15) with $0 < \alpha < 1/(2p)$ and $\alpha + \beta > 1/2$, and let $f_2(\lambda) = g(\lambda)$ be an integrable real symmetric function from the class $\mathrm{Lip}(q, 1/p - 2\alpha)$ with $1/p + 1/q \leq 1/\nu$, $\nu \in \mathbb{N}$. Then for any $\varepsilon > 0$*

$$\Delta_{\nu,W}(T) = O\left(T^{-\gamma+\varepsilon}\right) \quad \textit{as} \quad T \to \infty. \tag{4.9}$$

with $\gamma = 1/p - 2\alpha$.

Proof. The result follows from Theorem 3.2 **(B3)** and the following lemma, which is proved in the Appendix.

Lemma 4.1. *Let $p > 1$, $0 < \sigma < 1/p$ and let $f(\lambda)$ be differentiable function defined on $\mathbb{R} \setminus \{0\}$, such that for some constant $C > 0$*

$$|f(\lambda)| \leq \begin{cases} C|\lambda|^{-\sigma} & \textit{if} \quad |\lambda| \leq 1 \\ C|\lambda|^{-\delta} & \textit{if} \quad |\lambda| > 1 \end{cases}, \quad |f'(\lambda)| \leq \begin{cases} C|\lambda|^{-\sigma-1} & \textit{if} \quad |\lambda| \leq 1 \\ C|\lambda|^{-\delta-1} & \textit{if} \quad |\lambda| > 1 \end{cases}. \tag{4.10}$$

Then $f \in \mathrm{Lip}(p, 1/p - \sigma)$.

Now to prove (4.9) observe that by (1.15),

$$f'(\lambda) = -\frac{2\alpha + 2(\alpha+\beta)\lambda^2}{\lambda^{2\alpha+1}(1+\lambda)^{\beta+1}}. \tag{4.11}$$

It follows from (1.15) and (4.11) that the functions f and f' satisfy conditions (4.10) with $\sigma = 2\alpha$ and $\delta = 2\alpha + 2\beta$. Hence by Lemma 4.1, $f(\lambda) \in \mathrm{Lip}(p, 1/p - 2\alpha)$. Therefore the result follows from Theorem 3.2 **(B3)**. ☐

Theorem 4.5. *Let $f_i(\lambda)$, $i = 1, 2$, be the spectral density functions of two fractional Riesz-Bessel motions defined as*

$$f_i(\lambda) = \frac{C_i}{|\lambda|^{2\alpha_i}(1+\lambda^2)^{\beta_i}}, \quad 0 < \alpha_i < 1/2, \quad \alpha_i + \beta_i > 1/2, \quad i = 1, 2. \quad (4.12)$$

If $\alpha_1 + \alpha_2 < 1/(2\nu)$, *then (4.9) holds for any* $\varepsilon > 0$ *with*

$$\gamma = \frac{1}{2\nu} - (\alpha_1 + \alpha_2). \quad (4.13)$$

Proof. It follows from (1.15) and (4.11) that the functions f_1 and f_2 satisfy conditions of Theorem 3.2 **(B4)** with $\sigma_i = 2\alpha_i$ and $\delta_i = 2\alpha_i + 2\beta_i$, $i = 1, 2$, and the result follows. □

The next theorem, which is a continuous version of Theorem 4.2, contains an explicit second-order asymptotic expansion for $S_{1,W}(T)$ in the case where the Toeplitz operators are generated by the spectral densities given by (4.12), and shows that in this special case a second-order asymptotic expansion successfully removes the singularity and delivers a substantially improved approximation.

Theorem 4.6. *Let $f_i(\lambda)$, $i = 1, 2$, be as in (4.12). Then under $\alpha := \alpha_1 + \alpha_2 < 1/2$*

$$\begin{aligned} S_{1,W}(T) &:= \frac{1}{T}\mathrm{tr}[W_T(f_1)W_T(f_2)] \\ &= 2\pi \int_{-\infty}^{+\infty} f_1(\lambda)f_2(\lambda)\,d\lambda - \frac{C(\alpha_1,\alpha_2)}{T^{1-2\alpha}} + o\left(\frac{1}{T^{1-2\alpha}}\right) \end{aligned} \quad (4.14)$$

as $T \to \infty$, where

$$C(\alpha_1, \alpha_2) = \frac{2C_1C_2\pi^2}{\cos(\pi\alpha_1)\cos(\pi\alpha_2)\Gamma(2\alpha_1)\Gamma(2\alpha_2)} \cdot \frac{1}{2\alpha(1-2\alpha)}. \quad (4.15)$$

The proof is based on the following lemma, which contains an asymptotic formula for the covariance function of a fRBm process. It is proved in the Appendix.

Lemma 4.2. *Let $f(\lambda)$ be as in (1.15) with $0 < \alpha < 1/2$ and $\beta > 1/2$, and let $r(t) := \hat{f}(t)$ be the Fourier transform of $f(\lambda)$. Then*

$$r(t) = t^{2\alpha-1} \cdot \frac{\pi C}{\cos(\pi\alpha)\Gamma(2\alpha)} \cdot (1 + o(1)) \quad \text{as} \quad t \to \infty. \quad (4.16)$$

Remark 4.2. Taking into account the reflection formula $\Gamma(2\alpha)\Gamma(1-2\alpha) = \pi/\sin(2\pi\alpha)$, the asymptotic relation (4.16) can be written in the following form

$$r(t) = Ct^{2\alpha-1}\sin(\pi\alpha)\Gamma(1-2\alpha) \cdot (1 + o(1)) \quad \text{as} \quad t \to \infty. \quad (4.17)$$

Proof of Theorem 4.6. By Lemma 5.2 1) and Parseval-Plancherel theorem we have

$$S_{T,1} = \int_{-T}^{T}\left(1 - \frac{|t|}{T}\right) r_1(t)r_2(t)\,dt = 2\pi \int_{-\infty}^{+\infty} f_1(\lambda)f_2(\lambda)\,d\lambda - I_1 - I_2, \quad (4.18)$$

where

$$I_1 = \int_{|t|>T} r_1(t) r_2(t)\, dt \tag{4.19}$$

and

$$I_2 = \frac{1}{T}\int_{-T}^{T} |t| r_1(t) r_2(t)\, dt. \tag{4.20}$$

Hence, for $\alpha := \alpha_1 + \alpha_2 < 1/2$ from Lemma 4.2 and (4.19), we have as $T \to \infty$

$$\begin{aligned} I_1 &= \int_{|t|>T} r_1(t) r_2(t)\, dt = 2\int_{t>T} r_1(t) r_2(t)\, dt \\ &= \frac{2C_1C_2\pi^2}{\cos(\pi\alpha_1)\cos(\pi\alpha_2)\Gamma(2\alpha_1)\Gamma(2\alpha_2)} \cdot \int_{t>T} t^{2(\alpha-1)}\,(1+o(1)) \\ &= \frac{2C_1C_2\pi^2}{\cos(\pi\alpha_1)\cos(\pi\alpha_2)\Gamma(2\alpha_1)\Gamma(2\alpha_2)} \cdot \frac{1}{1-2\alpha} \cdot T^{2\alpha-1}\,(1+o(1)). \end{aligned} \tag{4.21}$$

For I_2, from Lemma 4.2 and (4.20), we have as $T \to \infty$ and $\alpha := \alpha_1 + \alpha_2 < 1/2$

$$\begin{aligned} I_2 &= \frac{1}{T}\int_{-T}^{T} |t| r_1(t) r_2(t)\, dt = \frac{2}{T}\int_0^T t\, r_1(t) r_2(t)\, dt \\ &= \frac{2}{T} \cdot \frac{C_1C_2\pi^2}{\cos(\pi\alpha_1)\cos(\pi\alpha_2)\Gamma(2\alpha_1)\Gamma(2\alpha_2)} \int_0^T t^{2\alpha-1}\, dt\,(1+o(1)) \\ &= \frac{2C_1C_2\pi^2}{\cos(\pi\alpha_1)\cos(\pi\alpha_2)\Gamma(2\alpha_1)\Gamma(2\alpha_2)} \cdot \frac{1}{2\alpha} \cdot T^{2\alpha-1}\,(1+o(1)). \end{aligned} \tag{4.22}$$

From (4.18), (4.21) and (4.22) the result follows. Theorem 4.6 is proved. □

Remark 4.3. (a) As in Remark 4.1, the rate in (4.14) can be obtained from Theorem 3.4.

(b) We analyze the behavior of the approximations as $\alpha_1, \alpha_2 \to 1/4$. First observe that the first-order asymptotic formula has a pole when $\alpha := \alpha_1 + \alpha_2 = 1/2$. In particular, denoting $\beta := \beta_1 + \beta_2$, and using the change of variable $\lambda^2 = u$, we have

$$\begin{aligned} 2\pi \int_{-\infty}^{+\infty} f_1(\lambda) f_2(\lambda)\, d\lambda &= 2\pi C_1 C_2 \int_{-\infty}^{+\infty} \frac{1}{|\lambda|^{2\alpha}(1+\lambda^2)^{\beta}} d\lambda \\ &= 2\pi C_1 C_2 \int_0^{\infty} \frac{u^{-1/2-\alpha}}{(1+u)^{\beta}} du. \end{aligned} \tag{4.23}$$

Applying the formula (see, e.g., [18])

$$\int_0^{\infty} \frac{u^{m-1}}{(1+u)^{m+n}} du = \frac{\Gamma(m)\Gamma(n)}{\Gamma(m+n)}, \quad m > 0,\ n > 0, \tag{4.24}$$

with $m = 1/2 - \alpha$ and $n = \beta - m = \alpha + \beta - 1/2$, from (4.23) we find

$$2\pi C_1 C_2 \int_{-\infty}^{+\infty} f_1(\lambda) f_2(\lambda)\, d\lambda = 2\pi C_1 C_2 \frac{\Gamma(1/2-\alpha)\Gamma(\alpha+\beta-1/2)}{\Gamma(\beta)}. \tag{4.25}$$

Then, using the Laurent expansion of the gamma function $\Gamma(1/2-\alpha)$ around the pole $\alpha=1/2$, from (4.25) we obtain: as $\alpha=\alpha_1+\alpha_2\to 1/2$

$$2\pi\int_{-\infty}^{+\infty} f_1(\lambda)f_2(\lambda)\,d\lambda=\frac{4\pi C_1C_2}{1-2\alpha}+O(1). \tag{4.26}$$

However, the asymptotic behavior of the second-order term $C(\alpha_1,\alpha_2)$ (see (4.15)), as $\alpha_1,\alpha_2\to 1/4$ is readily seen to be

$$\begin{aligned} C(\alpha_1,\alpha_2)&=\frac{2\pi^2C_1C_2}{\cos(\pi\alpha_1)\cos(\pi\alpha_2)\Gamma(2\alpha_1)\Gamma(2\alpha_2)}\cdot\frac{1}{2\alpha(1-2\alpha)}\\ &=\frac{2\pi^2C_1C_2}{\cos^2(\pi/4)[\Gamma(1/2)]^2}\cdot\frac{1}{1-2\alpha}+O(1)\\ &=\frac{4\pi C_1C_2}{1-2\alpha}+O(1). \end{aligned} \tag{4.27}$$

Thus, the pole in the first-order approximation is removed by the second-order approximation, so that the approximation

$$2\pi\int_{-\infty}^{+\infty} f_1(\lambda)f_2(\lambda)\,d\lambda-\frac{C(\alpha_1,\alpha_2)}{T^{1-2(\alpha_1+\alpha_2)}} \tag{4.28}$$

is bounded as $\alpha_1,\alpha_2\to 1/4$.

This good behavior explains why the second-order approximation produces a good approximation that does uniformly well over $\alpha_1,\alpha_2\in[0,1/4]$, including the limits of the domain.

(c) The second-order equivalence holds along an arbitrary ray for which $\alpha=\alpha_1+\alpha_2\to 1/2$. Indeed, let $\alpha_1^0\in[0,1/2]$ be any fixed number such that $\alpha_1\to\alpha_1^0$ and $\alpha=\alpha_1+\alpha_2\to 1/2$, then the representation (4.26) for the first-order asymptotic term continues to apply.

On the other hand, using the reflection formula $\Gamma(z)\Gamma(1-z)=\pi/\sin(\pi z)$ with $z=2\alpha_1^0$, we have from (4.15) as $\alpha_1\to\alpha_1^0$ and $\alpha=\alpha_1+\alpha_2\to 1/2$

$$\begin{aligned} C(\alpha_1,\alpha_2)&=\frac{2\pi^2C_1C_2}{\cos(\alpha_1\pi)\cos(\alpha_2\pi)\Gamma(2\alpha_1)\Gamma(2\alpha_2)}\cdot\frac{1}{2\alpha(1-2\alpha)}\\ &=\frac{2\pi^2C_1C_2}{\cos(\alpha_1^0\pi)\cos[(1-2\alpha_1^0)\pi/2]\Gamma(2\alpha_1^0)\Gamma(1-2\alpha_1^0)}\cdot\frac{1}{1-2\alpha}+O(1)\\ &=\frac{2\pi^2C_1C_2}{\cos(\alpha_1^0\pi)\sin(\alpha_1^0\pi)}\cdot\frac{\sin(2\alpha_1^0\pi)}{\pi}\cdot\frac{1}{1-2\alpha}+O(1)\\ &=\frac{4\pi C_1C_2}{1-2\alpha}+O(1), \end{aligned}$$

and again the second-order equivalence holds.

Thus, in this special case, the second-order asymptotic expansion removes the singularity in the first-order approximation, and provides a substantially improved approximation to the original functional.

4.2. Limit theorems for Toeplitz quadratic functionals

In this section we examine the limit behavior of quadratic forms and functionals of discrete- and continuous-time stationary Gaussian processes with possibly long-range dependence. The matrix and the operator that characterize the quadratic form and functional are Toeplitz.

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 1.2. We are interested in the asymptotic distribution (as $T\to\infty$) of the following Toeplitz type quadratic functionals of the process $X(u)$:

$$Q_T := \begin{cases} \int_0^T\int_0^T \widehat{g}(t-s)X(t)X(s)\,dt\,ds & \text{in the continuous-time case}\\ \sum_{k=1}^T\sum_{j=1}^T \widehat{g}(k-j)X(k)X(j) & \text{in the discrete-time case,}\end{cases} \tag{4.29}$$

where

$$\widehat{g}(t)=\int_\Lambda e^{i\lambda t}\,g(\lambda)\,d\lambda,\ \ t\in\mathbb{U} \tag{4.30}$$

is the Fourier transform of some real, even, integrable function $g(\lambda)$, $\lambda\in\Lambda$. We will refer $g(\lambda)$ as a generating function for the functional Q_T. In the discrete-time case the functions $f(\lambda)$ and $g(\lambda)$ are assumed to be 2π-periodic and periodically extended to $\mathbb{R}$.

The limit distributions of the functionals (4.29) are completely determined by the spectral density $f(\lambda)$ and the generating function $g(\lambda)$, and depending on their properties the limit distributions can be either Gaussian (i.e., Q_T with an appropriate normalization obeys central limit theorem), or non-Gaussian. The following two questions arise naturally:

(a) Under what conditions on $f(\lambda)$ and $g(\lambda)$ will the limits be Gaussian?
(b) Describe the limit distributions, if they are non-Gaussian.

4.2.1. Central limit theorems for Toeplitz quadratic functionals

We first discuss the question (a), that is, finding conditions on the spectral density $f(\lambda)$ and the generating function $g(\lambda)$ under which the functional Q_T, defined by (4.29), obeys central limit theorem.

This question goes back to the classical monograph by Grenander and Szegö [40], where the problem was considered for discrete time processes, as an application of the authors' theory of the asymptotic behavior of the trace of products of truncated Toeplitz matrices (see [40], p. 217–219).

Later the problem (a) was studied by Ibragimov [45] and Rosenblatt [57], in connection to the statistical estimation of the spectral ($F(\lambda)$) and covariance ($r(t)$) functions, respectively. Since 1986, there has been a renewed interest in both questions (a) and (b), related to the statistical inferences for long-memory processes (see, e.g., Avram [4], Fox and Taqqu [20], Giraitis and Surgailis [37], Giraitis et al. [36], Terrin and Taqqu [70], Taniguchi [63], Taniguchi and Kakizawa [64], Ginovyan and Sahakyan [32], and references therein). In particular, Avram [4], Fox and Taqqu [20], Giraitis and Surgailis [37], Ginovyan and Sahakyan [32] have obtained sufficient conditions for the quadratic form Q_T to obey the central limit theorem (CLT), when the model $X(t)$ is a discrete-time process.

For continuous time processes the question (a) was studied in Ibragimov [45], Ginovyan [25, 27], and Ginovyan and Sahakyan [33].

Let Q_T be as in (4.29). We will use the following notation: By $\widetilde{Q}_T$ we denote the standard normalized quadratic functional:

$$\widetilde{Q}_T = \frac{1}{\sqrt{T}}\left(Q_T - EQ_T\right). \tag{4.31}$$

The notation

$$\widetilde{Q}_T \Longrightarrow N(0,\sigma^2) \quad \text{as} \quad T\to\infty \tag{4.32}$$

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.

Our study of the asymptotic distribution of the quadratic functionals (4.29) is based on the following representation of the k–th order cumulant $\chi_k(\cdot)$ of $\widetilde{Q}_T$, which follows from (1.19) (see, also, [40, 45]):

$$\chi_k(\widetilde{Q}_T) = \begin{cases} 0, & \text{for } k=1 \\ T^{-k/2}2^{k-1}(k-1)!\,\mathrm{tr}\,[A_T(f)A_T(g)]^k, & \text{for } k\geq 2, \end{cases} \tag{4.33}$$

where $A_T(f)$ and $A_T(g)$ denote either the T-truncated Toeplitz operators (for continuous-time case), or the $T\times T$ Toeplitz matrices (for discrete-time case) generated by the functions f and g respectively, and $\mathrm{tr}[A]$ stands for the trace of an operator A.

The next result contains sufficient conditions in terms of $f(\lambda)$ and $g(\lambda)$ ensuring central limit theorems for standard normalized quadratic functionals $\widetilde{Q}_T$ both for discrete-time and continuous-time processes.

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

$$\sigma_0^2 := 16\pi^3\int_\Lambda f^2(\lambda)g^2(\lambda)\,d\lambda. \tag{4.34}$$

As usual $\Lambda=\mathbb{T}=(-\pi,\pi]$ in the discrete-time case and $\Lambda=\mathbb{R}=(-\infty,\infty)$ in the continuous-time case.

Theorem 4.7. *Each of the following conditions is sufficient for*

$$\widetilde{Q}_T \Longrightarrow N(0,\sigma_0^2) \quad \text{as} \quad T\to\infty, \tag{4.35}$$

with σ_0^2 given by (4.34).

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

$$\chi_2(\widetilde{Q}_T) := \frac{2}{T}tr\big[B_T(f)B_T(g)\big]^2 \longrightarrow \sigma_0^2 < \infty. \tag{4.36}$$

(B) *The function*

$$\varphi(\mathbf{u}) := \varphi(u_1, u_2, u_3) = \int_\Lambda f(\lambda)g(\lambda - u_1)f(\lambda - u_2)g(\lambda - u_3)\, d\lambda \quad (4.37)$$

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*

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

(E) *The spectral density* $f(\lambda)$ *and the generating function* $g(\lambda)$ *satisfy*

$$f(\lambda) \leq |\lambda|^{-\alpha} L_1(\lambda), \quad |g(\lambda)| \leq |\lambda|^{-\beta} L_2(\lambda), \quad \lambda \in \Lambda,$$

for some $\alpha < 1,\ \beta < 1$ *with*

$$\alpha + \beta \leq 1/2 \quad \text{and} \quad L_i \in SV(\mathbb{R}),\ \lambda^{-(\alpha+\beta)} L_i(\lambda) \in L^2(\Lambda), \quad i = 1, 2,$$

where $SV(\mathbb{R})$ *is the class of slowly varying at zero functions* $u(\lambda)$, $\lambda \in \mathbb{R}$, *satisfying* $u(\lambda) \in L^\infty(\mathbb{R})$, $\lim_{\lambda\to 0} u(\lambda) = 0$, $u(\lambda) = u(-\lambda)$ *and* $0 < u(\lambda) < u(\mu)$ *for* $0 < \lambda < \mu$.

In the continuous-time case, we also assume that the functions $f(\lambda)$ *and* $g(\lambda)$ *are bounded on* $\mathbb{R} \setminus (-\pi, \pi)$.

Remark 4.4. For discrete-time case: assertions (A) and (D) were proved in Giraitis and Surgailis [37] (see also Giraitis et al. [36]); assertions (B) and (E) were proved in Ginovyan and Sahakyan [32]; assertion (E) with $\alpha + \beta < 1/2$ was first obtained by Fox and Taqqu [20]; assertion (C) for $p = q = \infty$ was first established by Grenander and Szegö ([40], Section 11.7), while the case $p = 2$, $q = \infty$ was proved by Ibragimov [45] and Rosenblatt [57], in the general discrete-time case assertion (D) was proved by Avram [4].

For continuous-time case assertions (A)–(E) were proved in Ginovyan [27] and Ginovyan and Sahakyan [33].

Remark 4.5. Assertion (A) implies assertions (B)–(E). Assertion (B) implies assertions (C) and (D). On the other hand, for functions $f(\lambda) = \lambda^{-3/4}$ and $g(\lambda) = \lambda^{3/4}$ satisfying the conditions of (E), the function $\varphi(t_1, t_2, t_3)$ is not defined for $t_2 = 0$, $t_1 \neq 0$, $t_3 \neq 0$, showing that assertion (B) generally does not imply assertion (E) (see Ginovyan and Sahakyan [32]).

Remark 4.6. Examples of spectral density $f(\lambda)$ and generating function $g(\lambda)$ satisfying the conditions of Theorem 4.7 (E) are provided by the functions

$$f(\lambda) = |\lambda|^{-\alpha} |\ln|\lambda||^{-\gamma} \quad \text{and} \quad g(\lambda) = |\lambda|^{-\beta} |\ln|\lambda||^{-\gamma}, \quad (4.38)$$

where $\alpha < 1$, $\beta < 1$, $\alpha + \beta \leq 1/2$ and $\gamma > 1/2$ (see Ginovyan and Sahakyan [32]).

Remark 4.7. The functions $f(\lambda)$ and $g(\lambda)$ in Theorem 4.7 (E) 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 $\lambda_0 \in [-\pi, \pi]$. Using the properties of the products of Toeplitz matrices and operators it can be shown that Theorem 4.7 (E) remains valid when $f(\lambda)$ and $g(\lambda)$ have singularities of the form (4.38) at the same finite number of points of the segment $[-\pi, \pi]$ (see Ginovyan and Sahakyan [32]).

Remark 4.8. The next proposition shows that the condition of positiveness and finiteness of asymptotic variance of quadratic form Q_T is not sufficient for Q_T to obey CLT as was conjectured in Giraitis and Surgailis [37], and Ginovyan [26].

Proposition 4.1. *There exist a spectral density $f(\lambda)$ and a generating function $g(\lambda)$ such that*

$$0 < \int_{-\pi}^{\pi} f^2(\lambda)\, g^2(\lambda)\, d\lambda < \infty \tag{4.39}$$

and

$$\lim_{T\to\infty} \sup \chi_2(\widetilde{Q}_T) = \lim_{n\to\infty} \sup \frac{2}{T} tr\left(B_T(f)B_T(g)\right)^2 = \infty, \tag{4.40}$$

that is, the condition (4.39) does not guarantee convergence in (4.36).

To construct functions $f(\lambda)$ and $g(\lambda)$ satisfying (4.39) and (4.40), 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

$$f_0(\lambda) = \begin{cases} \left(\frac{2^s}{s^2}\right)^{1/p}, & \text{if } 2^{-s-1} \leq \lambda \leq 2^{-s},\ s = 2m \\ 0, & \text{if } 2^{-s-1} \leq \lambda \leq 2^{-s},\ s = 2m+1 \end{cases} \tag{4.41}$$

$$g_0(\lambda) = \begin{cases} \left(\frac{2^s}{s^2}\right)^{1/q}, & \text{if } 2^{-s-1} \leq \lambda \leq 2^{-s},\ s = 2m+1 \\ 0, & \text{if } 2^{-s-1} \leq \lambda \leq 2^{-s},\ s = 2m, \end{cases} \tag{4.42}$$

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 (4.39) and (4.40) (for details we refer to Ginovyan and Sahakyan [32]. Consequently, for these functions the standard normalized quadratic form Q_T does not obey CLT, and it is of interest to describe the limiting non-Gaussian distribution of Q_T in this special case.

4.2.2. Non-central limit theorems

The problem (b) for discrete-time processes, that is, the description of the limit distribution of the quadratic form

$$Q_T := \sum_{k=1}^{T} \sum_{j=1}^{T} \widehat{g}(k-j) X(k) X(j), \quad T \in \mathbb{N} \tag{4.43}$$

if it is non-Gaussian, goes back to the papers by Rosenblatt [56]–[58].

Later this problem was studied in a series of papers by Taqqu, and Terrin and Taqqu (see, e.g., [65], [68], [69], [70], and references therein). Specifically, suppose that the spectral density $f(\lambda)$ and the generating function $g(\lambda)$ are regularly varying functions at the origin:

$$f(\lambda) = |\lambda|^{-\alpha} L_1(\lambda) \quad \text{and} \quad g(\lambda) = |\lambda|^{-\beta} L_2(\lambda), \quad \alpha < 1, \beta < 1, \tag{4.44}$$

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 4.7(E) the standard normalized quadratic form

$$T^{-1/2}\,(Q_T - EQ_T)$$

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

Consider the embedding of the discrete sequence $\{Q_T, T \in \mathbb{N}\}$ into a continuous-time process $\{Q_T(t), T \in \mathbb{N}, t \in \mathbb{R}\}$ defined by

$$Q_T(t) := \sum_{k=1}^{[Tt]} \sum_{j=1}^{[Tt]} \widehat{g}(k-j) X(k) X(j), \tag{4.45}$$

where $[\,\cdot\,]$ stands for the greatest integer. Denote by $Z(\cdot)$ the complex-valued Gaussian random measure defined on the Borel σ-algebra $\mathcal{B}(\mathbb{R})$, and satisfying $EZ(B) = 0$, $E|Z(B)|^2 = |B|$, and $\overline{Z(-B)} = Z(B)$ for any $B \in \mathcal{B}(\mathbb{R})$.

The next result, proved in Terrin and Taqqu [69], describes the non-Gaussian limit distribution of the suitable normalized process $Q_T(t)$.

Theorem 4.8. *Let $f(\lambda)$ and $g(\lambda)$ be as in (4.44) 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 $Q_T(t)$ be as in (4.45). Then for $\alpha + \beta > 1/2$*

$$\widehat{Q}_T(t) := \frac{1}{T^{\alpha+\beta} L_1(1/T) L_2(1/T)}\,(Q_T(t) - E[Q_T(t)]) \tag{4.46}$$

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

$$Q(t) := \int_{\mathbb{R}^2}'' K_t(x,y)\, dZ(x) dZ(y), \tag{4.47}$$

where

$$K_t(x,y) = |xy|^{-\alpha/2} \int_{\mathbb{R}} \frac{e^{it(x+u)} - 1}{i(x+u)} \cdot \frac{e^{it(y-u)} - 1}{i(y-u)} |u|^{-\beta} du, \tag{4.48}$$

The double prime in the integral (4.47) indicates that the integration excludes the diagonals $x = \pm y$.

Remark 4.9. The limiting process in (4.47) is real-valued, non-Gaussian, and satisfies $EQ(t) = 0$ and $EQ^2(t) = \int_{\mathbb{R}^2} |K_t(x,y)|^2 dxdy$. It is self-similar with parameter $H = \alpha + \beta \in (1/2, 2)$, that is, the processes $\{Q(at), t \geq 0\}$ and $\{a^H Q(t), t \geq 0\}$ have the same finite dimensional distributions for all $a > 0$.

Remark 4.10. In [56] (see also [58]) Rosenblatt showed that if a discrete-time 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

$$Q_T := T^{-\alpha} \sum_{k=1}^{T} \left[X^2(k) - 1\right]$$

has a non-Gaussian limiting distribution, and described this distribution in terms of characteristic functions. This is a special case of Theorem 4.8 with $t = 1$, $1/2 < \alpha < 1$ and $\beta = 0$. In [65] (see also [68]) Taqqu extended Rosenblatt's result by showing that the stochastic process

$$Q_T(t) := T^{-\alpha} \sum_{k=1}^{[Tt]} \left[X^2(k) - 1\right]$$

converges (as $T \to \infty$) weakly to a process (called the Rosenblatt process) which has the double Wiener-Itô integral representation

$$Q(t) := C_\alpha \int_{\mathbb{R}^2}^{''} \frac{e^{it(x+y)} - 1}{i(x+y)} |x|^{-\alpha/2} |y|^{-\alpha/2} dZ(x) dZ(y). \tag{4.49}$$

The distribution of the random variable $Q(t)$ in (4.49) for $t = 1$ is described in Veillette and Taqqu [71].

Remark 4.11. The slowly varying functions L_1 and L_2 in (4.44) are of importance because they provide a great flexibility in the choice of spectral density f and generating function g. Observe that in Theorem 4.8 the functions L_1 and L_2 influence only the normalization (see (4.46)), but not the limit $Q(t)$. Theorem 4.7(E) shows that in the critical case $\alpha + \beta = 1/2$ the limit distribution of the standard normalized quadratic form $\widetilde{Q}_T$ given by (4.31) is Gaussian and depends on the slowly varying factors L_1 and L_2 through f and g.

Note also that the critical case $\alpha + \beta = 1/2$ was partially investigated by Terrin and Taqqu in [70]. Starting from the limiting random variable $Q(1) = Q(1; \alpha, \beta)$, which exists only when $\alpha + \beta > 1/2$, they showed that the random variable

$$(\alpha + \beta - 1/2) Q(1; \alpha, \beta)$$

converges in distribution to a Gaussian random variable as $\alpha + \beta$ approaches to $1/2$.

Remark 4.12. For continuous-time processes the problem (b) has not been investigated, and it would be of interest to describe the limiting non-Gaussian distribution of the quadratic functional Q_T.

4.3. Berry-Esséen bounds and large deviations for Toeplitz quadratic functionals

In this section, we briefly discuss Berry-Esséen bounds in the CLT and large deviations principle for quadratic functionals both for continuous-time and discrete-time Gaussian stationary processes (for more about these topics we refer to [6, 13, 15, 22, 47, 52, 63], and reference therein).

Berry-Esséen Bounds Let Q_T and $\widetilde{Q}_T$ be as in (4.29) and (4.31), respectively. Denote $\widehat{Q}_T := \widetilde{Q}_T/\sqrt{\mathrm{Var}(\widetilde{Q}_T)}$, and let Z be the standard normal random variable: $Z \sim N(0,1)$. The CLT for Q_T (Theorem 4.7) tells us that $\widehat{Q}_T \longrightarrow Z$ in distribution as $T \to \infty$. The natural next step concerns the closeness between the distribution of $\widehat{Q}_T$ and standard normal distribution, which means asking for the rate of convergence in the CLT. Results of this type are known as Berry-Esséen bounds (or asymptotics).

In the discrete-time case, for special quadratic functionals, Berry-Esséen bounds were established in Tanoguchi [63], while for the continuous-time case, Berry-Esséen-type bounds were obtained in Nourdin and Peccati [52]. The next theorem captures both cases.

Theorem 4.9. *Let $\widetilde{Q}_T$ be as in (4.31), $\widehat{Q}_T := \widetilde{Q}_T/\sqrt{\mathrm{Var}(\widetilde{Q}_T)}$, and $\Phi(z) = P(Z \le z)$, where $Z \sim N(0,1)$. Assume that $f(\lambda) \in L^1(\Lambda) \cap L^p(\Lambda)$ $(p > 1)$ and $g(\lambda) \in L^1(\Lambda) \cap L^q(\Lambda)$ $(q > 1)$. The following assertions hold.*

1. *If $1/p + 1/q \le 1/4$, then there exists a constant $C = C(f,g) > 0$ such that for all $T > 0$ we have*

$$\sup_{z\in\mathbb{R}} |P(\widehat{Q}_T \le z) - \Phi(z)| \le \frac{C}{\sqrt{T}}. \tag{4.50}$$

2. *If $1/p + 1/q \le 1/8$ and $\int_\Lambda f^3(\lambda) g^3(\lambda)\, d\lambda \ne 0$, then there exist a constant $c = c(f,g) > 0$ and a number $T_0 = T_0(f,g) > 0$ such that $T > T_0$ implies*

$$\sup_{z\in\mathbb{R}} |P(\widehat{Q}_T \le z) - \Phi(z)| \ge \frac{c}{\sqrt{T}}. \tag{4.51}$$

More precisely, for any $z \in \mathbb{R}$, we have as $T \to \infty$

$$\sqrt{T} |P(\widehat{Q}_T \le z) - \Phi(z)| \longrightarrow \sqrt{\frac{2}{3}} \frac{\int_\Lambda f^3(\lambda) g^3(\lambda) d\lambda}{\left(\int_\Lambda f^2(\lambda) g^2(\lambda) d\lambda\right)^{3/2}} (1 - z^2) e^{-z^2/2}. \tag{4.52}$$

Remark 4.13. In the continuous-time case, Theorem 4.9 was proved in Nourdin and Peccati [52], by appealing to a general CLT of Section 4.2 (Theorem 4.7), and Stein's method. The proof, in the discrete-time case, is similar to that of the continuous-time case.

Large Deviations We now present sufficient conditions that ensure large deviations principle (LDP) for Toeplitz type quadratic functionals of stationary Gaussian processes. For more about LDP we refer to Bryc and Dembo [13], Bercu et al. [6], Sato et al. [60], Taniguchi and Kakizawa [64], and references therein.

First observe that large deviation theory can be viewed as an extension of the law of large numbers (LLN). The LLN states that certain probabilities converge to zero, while the large deviation theory focuses on the rate of convergence. Specifically, consider a sequence of random variables $\{\xi_n,\, n \geq 1\}$ converging in probability to a real constant m. Note that ξ_n could represent, for instance, the n-th partial sum of another sequence of random variables: $\xi_n = \frac{1}{n}\sum_{k=1}^{n} \eta_k$, where the sequence $\{\eta_k\}$ may be independent identically distributed, or dependent as in an observed stretch of a stochastic process. By the LLN, we have for $\varepsilon > 0$

$$\mathbb{P}\{|\xi_n - m| > \varepsilon\} \to 0 \quad \text{as} \quad n \to \infty. \tag{4.53}$$

It is often the case that the convergence in (4.53) is exponentially fast, that is,

$$\mathbb{P}\{|\xi_n - m| > \varepsilon\} \approx R(\cdot)\exp[-nI(\varepsilon, m)] \quad \text{as} \quad n \to \infty, \tag{4.54}$$

where $R(\cdot) = R(\varepsilon, m, n)$ is a slowly varying (relative to an exponential) function of n and $I(\varepsilon, m)$ is a positive quantity. Loosely, if (4.54) holds, we say that the sequence $\{\xi_n\}$ satisfies a large deviations principle. One of the basic problems of the large deviation theory is to determine $I(\varepsilon, m)$ and $R(\varepsilon, m, n)$. To be more precise, we recall the definition of Large Deviation Principle (LDP) (see, e.g., [13, 64]).

Definition 4.1. Let $\{\xi_n,\, n \in \mathbb{Z}\}$ be a sequence of real–valued random variables defined on the probability space $(\Omega, \mathcal{F}, \mathbb{P})$. We say that $\{\xi_n\}$ satisfies a Large Deviation Principle (LDP) with speed $a_n \to 0$ and rate function $I : \mathbb{R} \to [0, \infty]$, if $I(x)$ is lower semicontinuous, that is, if $x_n \to x$ then $\liminf_{n\to\infty} I(x_n) \geq I(x)$, and

$$\liminf_{n\to\infty} a_n \log \mathbb{P}\{\xi_n \in A\} \geq -\inf_{x\in A} I(x)$$

for all open subsets $A \subset \mathbb{R}$, while

$$\limsup_{n\to\infty} a_n \log \mathbb{P}\{\xi_n \in B\} \leq -\inf_{x\in B} I(x)$$

for all closed subsets $B \subset \mathbb{R}$. The function $I(x)$ is called a good rate function if its level sets are compact, that is, the set $\{x \in \mathbb{R} : I(x) \leq b\}$ is compact for each $b \in \mathbb{R}$.

Now let Q_T be the Toeplitz type quadratic functionals of a process $X(u)$ defined by (4.29) with spectral density $f(\lambda)$ and generating function $g(\lambda)$.

The next result states sufficient conditions in terms of $f(\lambda)$ and $g(\lambda)$ to ensure that the LDP for normalized quadratic functionals $\{\frac{1}{T}Q_T\}$ holds both for discrete-time and continuous-time processes.

Theorem 4.10. *Assume that $f(\lambda)g(\lambda) \in L^\infty(\Lambda)$. Then the random variable $\{\frac{1}{T}Q_T\}$ satisfies a LDP with the speed $a_T = \frac{1}{T}$ and a good rate function $I(x)$:*

$$I(x) = \sup_{-\infty<y<1/(2C)} \{xy - V(f,g;y)\},$$

where $C = ess\ sup f(\lambda)|g(\lambda)|$ and for $y < 1/(2C)$

$$V(f,g;y) = -\frac{1}{4\pi}\int_\Lambda \log(1-2yf(\lambda)g(\lambda))d\lambda.$$

Remark 4.14. In the special case of $g(\lambda) = 1$, Theorem 4.10 was proved by Bryc and Dembo [13]. For the general case we refer to Bercu et al. [6].

5. Proof of Theorems 3.1–3.4

We only prove the results concerning Toeplitz operators (Theorems 3.1–3.4). The proofs of the corresponding results for Toeplitz matrices are similar. First we state a number of technical lemmas, which are proved in the Appendix.

The following result is known (see, e.g., [33], or [36], p. 8).

Lemma 5.1. *Let $D_T(u)$ be the Dirichlet kernel*

$$D_T(u) = \frac{\sin(Tu/2)}{u/2}. \tag{5.1}$$

Then, for any $\delta \in (0,1)$

$$|D_T(u)| \le 2\,T^\delta |u|^{\delta-1}, \quad u \in \mathbb{R}. \tag{5.2}$$

Denote

$$G_T(u) := \int_0^T e^{iTu}dt = e^{iTu/2}D_T(u), \quad u \in \mathbb{R}, \tag{5.3}$$

$$\Phi_T(\mathbf{u}) := \frac{1}{(2\pi)^{m-1}T} \cdot D_T(u_1)\cdots D_T(u_{m-1})D_T(u_1+\cdots+u_{m-1}), \tag{5.4}$$

$$\Psi(\mathbf{u}) := \varphi(u_1, u_1+u_2, \ldots, u_1+\cdots+u_{m-1}), \tag{5.5}$$

where $\mathbf{u} = (u_1,\ldots,u_{m-1}) \in \mathbb{R}^{m-1}$ and the function $\varphi(\mathbf{u})$, corresponding to the collection $\mathcal{H} = \{h_1, h_2, \ldots, h_m\}$, is defined by (3.8).

The next lemma follows from (3.4) and (5.3)–(5.5) (cf. [33], Lemma 1).

Lemma 5.2. *Let $\mathcal{H} = \{h_1, h_2, \ldots, h_m\}$ be a collection of integrable real symmetric functions on $\mathbb{R}$, $\hat{h}_k$ be the Fourier transform of function h_k $(k = 1,\ldots,m)$, and let $S(T) := S_{W,\mathcal{H}}(T)$ be as in (3.4). The following equalities hold.*

$$1)\ S(T) = \frac{1}{T}\int_0^T \ldots \int_0^T \hat{h}_1(u_1-u_2)\hat{h}_2(u_2-u_3)\ldots\hat{h}_m(u_m-u_1)$$

$$2)\ S(T) = \frac{1}{T}\int_{\mathbb{R}^m} h_1(u_1)\dots h_m(u_m)G_T(u_1-u_2)G_T(u_2-u_3)\cdots \times G_T(u_m-u_1)du_1\dots du_m.$$

$$3)\ S(T) = (2\pi)^{m-1}\int_{\mathbb{R}^{m-1}} \Psi(\mathbf{u})\Phi_T(\mathbf{u})d\mathbf{u}.$$

For $m = 3, 4, \dots$ and $\delta > 0$ we denote

$$\mathbb{E}_\delta = \{(u_1,\dots,u_{m-1}) \in \mathbb{R}^{m-1} : |u_i| \le \delta,\, i = 1,\dots,m-1\}, \quad \mathbb{E}_\delta^c = \mathbb{R}^{m-1}\setminus \mathbb{E}_\delta$$

and

$$p_3 = 2, \quad p(m) = \frac{m-2}{m-3} \quad (m > 3).$$

Lemma 5.3. *The kernel $\Phi_T(\mathbf{u})$, $\mathbf{u} \in \mathbb{R}^{m-1}$, $m \ge 3$ possesses the following properties:*

a) $\sup_T \int_{\mathbb{R}^{m-1}} |\Phi_T(\mathbf{u})|\, d\mathbf{u} = C_1 < \infty;$

b) $\int_{\mathbb{R}^{m-1}} \Phi_T(\mathbf{u})\, d\mathbf{u} = 1;$

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

d) for any $\delta > 0$ there exists a constant $C_\delta > 0$ such that

$$\int_{\mathbb{E}_\delta^c} |\Phi_T(\mathbf{u})|^{p(m)}\, d\mathbf{u} \le C_\delta \quad \text{for} \quad T > 0, \tag{5.6}$$

The proof of the next lemma can be found in [36], p. 161.

Lemma 5.4. *Let $0 < \beta < 1$, $0 < \alpha < 1$, and $\alpha + \beta > 1$. Then for any $y \in \mathbb{R}$, $y \ne 0$,*

$$\int_{\mathbb{R}} \frac{1}{|x|^\alpha |x+y|^\beta} dx = \frac{C}{|y|^{\alpha+\beta-1}}, \tag{5.7}$$

where C is a constant depending on α and β.

Denote $E_1 = \{(u_1, u_2, \dots, u_n) \in \mathbb{R}^n : |u_i| \le 1,\ i = 1, 2, \dots, n\}$ and let $E_1^c = \mathbb{R}^n \setminus E_1$.

Lemma 5.5. *Let $0 < \alpha \le 1$ and $\frac{n}{n+1} < \beta < \frac{n+\alpha}{n+1}$. Then*

$$B_i := \int_{E_1} \frac{|u_i|^\alpha}{|u_1\cdots u_n(u_1+\dots+u_n)|^\beta}\, du_1\cdots du_n < \infty, \quad i = 1,\dots,n. \tag{5.8}$$

Lemma 5.6. *Let $\frac{n}{n+1} < \beta < 1$. Then*

$$I := \int_{E_1^c} \frac{1}{|u_1\cdots u_n(u_1+\dots+u_n)|^\beta}\, du_1\cdots du_n < \infty. \tag{5.9}$$

Proof of Theorem 3.1. We start with **(A2)**. By Lemma 5.2, 3) and 5.3, b), we have $\Delta(T) = (2\pi)^{m-1}|R(T)|$, where

$$R(T) = \int_{\mathbb{R}^3} [\Psi(\mathbf{u}) - \Psi(\mathbf{0})]\Phi_T(\mathbf{u}) d\mathbf{u}.$$

It follows from (5.5) that the function $\Psi(\mathbf{u})$ belongs to $L^{m-2}(\mathbb{R}^{m-1})$ and is continuous at $\mathbf{0} = (0, \ldots, 0) \in \mathbb{R}^{m-1}$. Hence for any $\varepsilon > 0$ we can find $\delta > 0$ to satisfy

$$|\Psi(\mathbf{u}) - \Psi(\mathbf{0})| < \frac{\varepsilon}{C_1}, \quad \mathbf{u} \in \mathbb{E}_\delta, \tag{5.10}$$

where C_1 is the constant from Lemma 5.3, a). Consider the decomposition $\Psi = \Psi_1 + \Psi_2$ such that

$$\|\Psi_1\|_{(m-2)} \le \frac{\varepsilon}{\sqrt{C_\delta}} \quad \text{and} \quad \|\Psi_2\|_\infty < \infty, \tag{5.11}$$

where C_δ is as in Lemma 5.3, d).

Observe that $\frac{1}{m-2} + \frac{1}{p(m)} = 1$, where $p(m) = \frac{m-2}{m-3}$. Hence, applying Lemma 5.3 and (5.10), (5.11) for sufficiently large T we obtain

$$\begin{aligned} |R(T)| \quad &\le \int_{\mathbb{E}_\delta} |\Psi(\mathbf{u}) - \Psi(\mathbf{0})||\Phi_T(\mathbf{u})| d\mathbf{u} + C_m \int_{\mathbb{E}_\delta^c} |\Psi_1(\mathbf{u})||\Phi_T(\mathbf{u})| d\mathbf{u} \\ &+ \int_{\mathbb{E}_\delta^c} |\Psi_2(\mathbf{u}) - \Psi(\mathbf{0})||\Phi_T(\mathbf{u})| d\mathbf{u} \le \frac{\varepsilon}{C_1} \int_{\mathbb{E}_\delta} |\Phi_T(\mathbf{u})| d\mathbf{u} \\ &+ \|\Psi_1\|_{(m-2)} \left[\int_{\mathbb{E}_\delta^c} \Phi_T^{p(m)}(\mathbf{u}) d\mathbf{u}\right]^{1/p(m)} + C_2 \int_{\mathbb{E}_\delta^c} |\Phi_T(\mathbf{u})| d\mathbf{u} \le 3\,\varepsilon, \end{aligned}$$

and the result follows. ☐

Proof of **(A1)**. According to Theorem (A2) it is enough to prove that the function

$$\varphi(\mathbf{u}) := \int_{-\infty}^{+\infty} h_1(\lambda) h_2(\lambda - u_1) h_3(\lambda - u_2) \cdots h_m(\lambda - u_{m-1})\, d\lambda, \tag{5.12}$$

where $\mathbf{u} = (u_1, \ldots, u_{m-1}) \in \mathbb{R}^{m-1}$, belongs to $L^{m-2}(\mathbb{R}^{m-1})$ and is continuous at $\mathbf{0} = (0, \ldots, 0) \in \mathbb{R}^{m-1}$, provided that

$$h_i \in L^1(\mathbb{R}) \bigcap L^{p_i}(\mathbb{R}), \quad 1 \le p_i \le \infty, \quad i = 1, \ldots, m, \quad \sum_{i=1}^m \frac{1}{p_i} \le 1. \tag{5.13}$$

It follows from Hölder's inequality, (5.12) and (5.13) that

$$|\varphi(\mathbf{u})| \le \prod_{i=1}^m ||f_i||_{L^{p_i}(\mathbb{R})} < \infty, \quad \mathbf{u} \in \mathbb{R}^{m-1}.$$

Hence, $\varphi \in L^\infty(\mathbb{R}^{m-1})$. On the other hand, the condition $h_i \in L^1(\mathbb{R})$ and (5.12) imply $\varphi \in L^1(\mathbb{R}^{m-1})$. Therefore $\varphi \in L^{m-2}(\mathbb{R}^{m-1})$.

To prove the continuity of $\varphi(\mathbf{u})$ at the point $\mathbf{0}$ we consider three cases.

Case 1. $p_i < \infty, \quad i = 1, \ldots, m.$
For an arbitrary $\varepsilon > 0$ we can find $\delta > 0$ satisfying (see (5.13))

$$\|h_i(\lambda - u) - h_i(\lambda)\|_{L^{p_i}(\mathbb{R})} \leq \varepsilon, \quad i = 2, \ldots, m, \quad \text{if} \quad |u| \leq \delta. \tag{5.14}$$

We fix $\mathbf{u} = (u_1, \ldots, u_{m-1})$ with $|\mathbf{u}| < \delta$ and denote

$$\overline{h}_i(\lambda) = h_i(\lambda - u_{i-1}) - h_i(\lambda), \quad i = 2, \ldots, m.$$

Then in view of (5.12) we have

$$\varphi(\mathbf{u}) = \int_{\mathbb{R}} h_1(\lambda) \prod_{i=2}^{m-1} \left(\overline{h}_i(\lambda) + h_i(\lambda)\right) d\lambda = \varphi(\mathbf{0}) + W.$$

It follows from (5.14) that $\|\overline{h}_i\|_{p_i} \leq \varepsilon$, $i = 2, \ldots, m$. Observe that each of the integrals comprising W contains at least one function $\overline{h}_i$ and can be estimated as follows:

$$\left| \int_{\mathbb{R}} h_1(u)\overline{h}_2(\lambda)h_3(\lambda) \ldots h_{m-1}(\lambda) d\lambda \right|$$
$$\leq \|h_1\|_{L^{p_1}} \|\overline{h}_2\|_{L^{p_2}} \|h_3\|_{L^{p_3}} \ldots \|h_m\|_{L^{p_m}} \leq C\varepsilon.$$

Case 2. $p_i \leq \infty, \ i = 1, \ldots, m, \ \ \sum_{i=m}^{m} \frac{1}{p_i} < 1.$
There exist finite numbers $p_i' < p_i, \ i = 1, \ldots, m$, such that $\sum_{i=1}^{m} 1/p_i' \leq 1$. Hence according to (5.13) we have $h_i \in L^{p_i'}$ $i = 1, \ldots, m$ and φ is continuous at $\mathbf{0}$ as in the case 1.

Case 3. $p_i \leq \infty, \ i = 1, \ldots, m, \ \sum_{i=1}^{m} \frac{1}{p_i} = 1.$
First observe that at least one of the numbers p_i is finite. Suppose, without loss of generality, that $p_1 < \infty$. For any $\varepsilon > 0$ we can find functions h_1', h_1'' such that

$$h_1 = h_1' + h_1'', \quad h_1' \in L^{\infty}(\mathbb{R}), \quad \|h_1''\|_{L^{p_1}} < \varepsilon. \tag{5.15}$$

Therefore

$$\varphi(\mathbf{u}) = \varphi'(\mathbf{u}) + \varphi''(\mathbf{u}),$$

where the functions φ' and φ'' are defined as φ in (5.12) with h_1 replaced by h_1' and h_1'', respectively. It follows from (5.15) that φ' is continuous at $\mathbf{0}$ (see Case 2), while by Hölder's inequality $|\varphi''(\mathbf{u})| \leq C \cdot \varepsilon$. Hence, for sufficiently small $|\mathbf{u}|$

$$|\varphi(\mathbf{u}) - \varphi(\mathbf{0})| \leq |\varphi'(\mathbf{u}) - \varphi'(\mathbf{0})| + |\varphi''(\mathbf{u}) - \varphi''(\mathbf{0})| \leq (C + 1)\varepsilon,$$

and the result follows. Theorem 3.1 is proved. □

Proof of Theorem 3.2. We start with **(B1)**. First observe that the condition $h_i \in \mathcal{F}_1$ implies that

$$\hat{h}_i \in L^1(\mathbb{R}) \quad \text{and} \quad |\hat{h}_i(t)| \leq A, \quad t \in \mathbb{R}, \quad i = 1, 2, \ldots, m \tag{5.16}$$

for some constant $A>0$. By Lemma 5.2 we have

$$T\cdot S(T)=\int_0^T\dots\int_0^T\hat h_1(u_1-u_2)\hat h_2(u_2-u_3)\dots\hat h_m(u_m-u_1)du_1\dots du_m.$$

Making the change of variables

$$u_1-u_2=t_1,\quad u_2-u_3=t_2,\dots,u_{m-1}-u_m=t_{m-1},$$

and observing that $t_1+\cdots+t_{m-1}=u_1-u_m$, we get (below we use the notation: $\mathbf{t}_{m-1}=(t_1,\dots,t_{m-1})$ and $d\mathbf{t}_{m-1}=dt_1\cdots dt_{m-1}$),

$$T\cdot S(T)=\int_0^T\int_{u_m-t_1-\cdots-t_{m-1}-T}^{u_m-t_1-\cdots-t_{m-1}}\int_{u_m-t_1-\cdots-t_{m-2}-T}^{u_m-t_1-\cdots-t_{m-2}}\cdots \tag{5.17}$$

$$\cdots\int_{u_m-t_1-T}^{u_m-t_1}\hat h_1(t_1)\cdots\hat h_{m-1}(t_{m-1})\hat h_m(-t_1-\cdots-t_{m-1})d\mathbf{t}_{m-1}du_m$$

$$=\int_{-T}^T\cdots\int_{-T}^T\hat h_1(t_1)\cdots\hat h_{m-1}(t_{m-1})\hat h_m(-t_1-\cdots-t_{m-1})\left[T-l(\mathbf{t}_{m-1})\right]d\mathbf{t}_{m-1},$$

where

$$|l(\mathbf{t}_{m-1})|=|l(t_1,\dots,t_{m-1})|\le 2\left(|t_1|+\cdots+|t_{m-1}|\right). \tag{5.18}$$

On the other hand, by (3.4) and Parseval's equality we have

$$\begin{aligned} M \quad &:= \quad M_{\mathbb{R},\mathcal{H}}=(2\pi)^{m-1}\int_{-\infty}^{\infty}\left[\prod_{i=1}^m h_i(\lambda)\right]d\lambda \\ &= \quad \int_{-\infty}^{\infty}\cdots\int_{-\infty}^{\infty}\hat h_1(t_1)\cdots\hat h_{m-1}(t_{m-1})\hat h_m(-t_1-\cdots-t_{m-1})d\mathbf{t}_{m-1}. \end{aligned} \tag{5.19}$$

It follows from (3.4), (5.17) and (5.19) that

$$S(T)-M:=S_{W,\mathcal{H}}(T)-M_{\mathbb{R},\mathcal{H}}$$

$$\begin{aligned} &=-\frac{1}{T}\int_{[-T,\,T]^{m-1}}\hat h_1(t_1)\cdots\hat h_{m-1}(t_{m-1})\hat h_m(-t_1-\cdots-t_{m-1})l(\mathbf{t}_{m-1})d\mathbf{t}_{m-1} \\ &\quad+\int_{\mathbb{R}^{m-1}\setminus[-T,\,T]^{m-1}}\hat h_1(t_1)\cdots\hat h_{m-1}(t_{m-1})\hat h_m(-t_1-\cdots-t_{m-1})d\mathbf{t}_{m-1} \\ &\quad=:\Delta_T^1+\Delta_T^2. \end{aligned} \tag{5.20}$$

By (5.16), (5.18) and (5.20) we have

$$|T\cdot\Delta_T^1|\le 2A\sum_{i=1}^{m-1}\int_{\mathbb{R}^{m-1}}\left|\hat h_1(t_1)\cdots\hat h_{m-1}(t_{m-1})t_i\right|d\mathbf{t}_{m-1}=:A_1<\infty, \tag{5.21}$$

since $h_i\in\mathcal{F}_1$, $i=1,2,\dots,m$.

Further, observe that

$$\mathbb{R}^{m-1} \setminus [-T,\, T]^{m-1} \subset \bigcup_{i=1}^{m} \big\{(t_1,\ldots,t_{m-1}) \in \mathbb{R}^{m-1} : |t_i| > T\big\} =: \bigcup_{i=1}^{m} E_i.$$

Hence by (5.16) and (5.20) we have

$$|T \cdot \Delta_T^2| \le 2A \sum_{i=1}^{m-1} \int_{E_i} \big|\hat{h}_1(t_1)\ldots\hat{h}_{m-1}(t_{m-1})t_i\big| d\mathbf{t}_{m-1} =: A_2 < \infty. \tag{5.22}$$

From (5.20)–(5.22) we get **(B1)**. □

Proof of **(B2)**. By Lemma 5.2, 3) and Lemma 5.3, b), and (3.5) we have

$$\Delta(T) = (2\pi)^{m-1} \left| \int_{\mathbb{R}^{m-1}} [\Psi(\mathbf{u}) - \Psi(\mathbf{0})]\Phi_T(\mathbf{u})d\mathbf{u} \right|, \quad \mathbf{0} = (0,\ldots,0). \tag{5.23}$$

It follows from (3.11) and (5.5) that for $\mathbf{u} = (u_1,\ldots,u_{m-1}) \in \mathbb{R}^{m-1}$

$$|\Psi(\mathbf{u}) - \Psi(\mathbf{0})| \le (m-1)C\left(|u_1|^\gamma + \cdots + |u_{2\nu-1}|^\gamma\right). \tag{5.24}$$

Let $\varepsilon \in (0,\gamma)$. Then, applying Lemma 5.1 with $\delta = \frac{1+\varepsilon-\gamma}{m}$, and using (5.23) and (5.24), we can write

$$\begin{aligned} \Delta(T) &\le C_m \int_E |\Psi(\mathbf{u}) - \Psi(\mathbf{0})||\Phi_T(\mathbf{u})|d\mathbf{u} \\ &+ C_m \int_{E^c} |\Psi(\mathbf{u}) - \Psi(\mathbf{0})||\Phi_T(\mathbf{u})|d\mathbf{u} \\ &\le \frac{C_m}{T^{1-m\delta}} \sum_{i=1}^{m-1} \int_E \frac{|u_i|^\gamma}{|u_1\cdots u_{m-1}(u_1+\cdots+u_{m-1})|^{1-\delta}}\, du_1\cdots du_{m-1} \\ &+ 2\|\varphi\|_\infty \frac{C_m}{T^{1-m\delta}} \int_{E^c} \frac{1}{|u_1\cdots u_{m-1}(u_1+\cdots+u_{m-1})|^{1-\delta}}\, du_1\cdots du_{m-1}, \end{aligned} \tag{5.25}$$

where $E = \{(u_1,u_2,\ldots,u_{m-1}) \in \mathbb{R}^{m-1} : |u_i| \le 1,\ i = 1,2,\ldots,m-1\}$ and $E^c = \mathbb{R}^{m-1} \setminus E$. Since

$$\frac{m-1}{m} < 1-\delta < \frac{m-1+\gamma}{m},$$

we can apply Lemmas 5.5 and 5.6 with $\alpha = \gamma$, $n = m-1$ and $\beta = 1-\delta$ to conclude that all the integrals in (5.25) are finite. Since $1 - m\delta = \gamma - \varepsilon$, from (5.25) follows the statement **(B2)**. □

Proof of **(B3)**. According to **(B2)** it is enough to prove that the function

$$\varphi(\mathbf{u}) := \int_{\mathbb{R}} h_1(\lambda)h_2(\lambda-u_1)\cdots h_m(\lambda-u_{m-1})d\mathbf{u}, \tag{5.26}$$

where $\mathbf{u} = (u_1, \dots, u_{m-1}) \in \mathbb{R}^{m-1}$, belongs to $L^\infty(\mathbb{R}^{m-1})$, and with some positive constant C

$$|\varphi(\mathbf{u}) - \varphi(\mathbf{0})| \le C|\mathbf{u}|^\gamma, \quad \mathbf{u} \in \mathbb{R}^{m-1}, \tag{5.27}$$

provided that

$$h_i \in \mathrm{Lip}(\mathbb{R}; p_i, \gamma), \quad 1 \le p_i \le \infty, \quad i = 1, 2, \dots, m, \quad \text{and} \quad \sum_{i=1}^m \frac{1}{p_i} \le 1. \tag{5.28}$$

It follows from Hölder's inequality and (5.28) that

$$|\varphi(\mathbf{u})| \le \prod_{i=1}^m ||h_i||_{p_i} < \infty, \quad \mathbf{u} \in \mathbb{R}^{m-1}.$$

Hence $\varphi \in L^\infty(\mathbb{R}^{m-1})$.

To prove (5.27) we fix $\mathbf{u} = (u_1, \dots, u_{m-1}) \in \mathbb{R}^{m-1}$ and denote

$$\overline{h}_i(\lambda) = h_i(\lambda - u_{i-1}) - h_i(\lambda), \quad \lambda \in \mathbb{R}, \quad i = 2, \dots, m. \tag{5.29}$$

Since $h_i \in \mathrm{Lip}(\mathbb{R}; p_i, \gamma)$ we have

$$\|\overline{h}_i\|_{p_i} \le C_i |\mathbf{u}|^\gamma, \quad i = 2, \dots, m. \tag{5.30}$$

By (5.26) and (5.29),

$$\varphi(\mathbf{u}) = \int_{\mathbb{R}} h_1(\lambda) \prod_{i=2}^m \left(\overline{h}_i(\lambda) + h_i(\lambda)\right) d\lambda = \varphi(\mathbf{0}) + W.$$

Each of the $(2^{m-1} - 1)$ integrals comprising W contains at least one function $\overline{f}_i$, and in view of (5.30), can be estimated as follows:

$$\left| \int_{\mathbb{R}} h_1(\lambda) \overline{h}_2(\lambda) h_3(\lambda) \cdots h_m(\lambda) d\lambda \right|$$

$$\le \|h_1\|_{p_1} \|\overline{h}_2\|_{p_2} \|h_3(u)\|_{p_2} \cdots \|h_m\|_{p_m} \le C|\mathbf{u}|^\gamma.$$

This completes the proof of **(B3)**. □

Proof of **(B4)**. We set

$$\frac{1}{p_i} := \sigma_i + \frac{1}{m} \left[1 - (\sigma_1 + \cdots + \sigma_m)\right], \quad i = 1, 2, \dots, m.$$

Then

$$\sum_{i=1}^m \frac{1}{p_i} = 1 \quad \text{and} \quad \frac{1}{p_i} - \sigma_i + \frac{1}{m} \left[1 - (\sigma_1 + \cdots + \sigma_m)\right] = \gamma > 0, \quad i = 1, 2, \dots, m,$$

and

$$0 < \sigma_i < \frac{1}{p_i} < 1 < \delta_i, \qquad i = 1, 2, \dots, m.$$

Hence, according to Lemma 4.1, $h_i \in \mathrm{Lip}(p_i, \gamma)$, $i = 1, 2, \dots, m$. Applying (B3), we get (3.14) with γ as in (3.15). □

Proof of Theorem 3.3. First observe that by Lemma 7 from [34]

$$S_{2,W}(T) := \frac{1}{T}\mathrm{tr}[W_T(h_1)W_T(h_2)] = 2\pi \int_{\mathbb{R}}\int_{\mathbb{R}} F_T(s-t)h_1(s)h_2(t)\,dt\,ds, \quad (5.31)$$

where $F_T(u)$ is the Fejér kernel:

$$F_T(u) = \frac{1}{2\pi T}\left(\frac{\sin Tu/2}{u/2}\right)^2, \quad t \in \mathbb{R}.$$

Below we use the following properties of $F_T(u)$ (see, e.g., [14]):

$$\int_{\mathbb{R}} F_T(u)\,du = 1, \quad (5.32)$$

$$\int_{u\geq 1} F_T(u)du \leq C\,T^{-1}, \quad (5.33)$$

$$\int_0^1 F_T(u)u^\alpha du \leq \begin{cases} CT^{-\alpha}, & \text{if } \alpha \leq 1 \\ CT^{-1}\ln T, & \text{if } \alpha = 1 \\ CT^{-1}, & \text{if } \alpha > 1. \end{cases} \quad (5.34)$$

Since the function $F_T(u)$ is even, in view of (5.31) we can write

$$S_{2,W}(T) = \pi \int_{\mathbb{R}}\int_{\mathbb{R}} F_T(u)\left[h_1(u+t)h_2(t) + h_1(t)h_2(u+t)\right] dudt. \quad (5.35)$$

Consequently, taking into account (5.32) and the equality

$$\int_{\mathbb{R}} h_1(t)h_2(t)dt = \int_{\mathbb{R}} h_1(u+t)h_2(u+t)dt,$$

by (5.35) we get

$$\begin{aligned} \Delta_{2,W}(T) &:= \left|\frac{1}{T}\mathrm{tr}[W_T(h_1)W_T(h_2)] - 2\pi\int_{\mathbb{R}} h_1(t)h_2(t)\,dt\right| \quad (5.36) \\ &= \left|\pi\int_{\mathbb{R}} F_T(u)\int_{\mathbb{R}}(h_1(t)-h_1(u+t))(h_2(u+t)-h_2(t))dtdu\right|. \end{aligned}$$

Using Hölder's inequality, we find from (5.36)

$$\Delta_{2,W}(T) \leq \pi\int_{\mathbb{R}} F_T(u)||h_1(u+\cdot)-h_1(\cdot)||_p||h_2(u+\cdot)-h_2(\cdot)||_q du. \quad (5.37)$$

In view of (5.37) we have

$$\Delta_{2,W}(T) \leq C_1\int_0^1 F_T(u)|u|^{\gamma_1+\gamma_2}du + C_2||h_1||_p||h_2||_q\int_{u>1} F_T(u)du. \quad (5.38)$$

Therefore, the result follows from (5.33), (5.34) and (5.38). □

Proof of Theorem 3.4. It follows from Lemma 4.1 that under the assumptions of theorem $h_i \in \mathrm{Lip}(p_i, 1/p_i - \sigma_i)$, $i = 1, 2$. Hence, applying Theorem 3.3 with $\gamma_i = 1/p_i - \sigma_i$, we obtain (3.16). □

Appendix: Proof of technical lemmas

In this section we give proofs of technical lemmas stated and used in Sections 4 and 5.

Proof of Lemma 4.1. Let $h \in (0, 1/2)$ be fixed. Then

$$\int_{|\lambda|\leq 2h} |f(\lambda+h) - f(\lambda)|^p d\lambda \leq (2C)^p \int_0^{3h} |\lambda|^{-p\sigma} d\lambda \leq C_1 h^{1-p\alpha}. \tag{A.1}$$

Next, for $|\lambda| > 2h$ we have $f(\lambda+h) - f(\lambda) = f'(\xi) \cdot h$ with some $\xi \in (\lambda, \lambda+h)$. Hence

$$\int_{2h<|\lambda|<1/2} |f(\lambda+h) - f(\lambda)|^p d\lambda \leq C^p h^p \int_h^{1/2} |\lambda|^{-p(\sigma+1)} d\lambda \leq C h^{1-p\sigma} \tag{A.2}$$

and

$$\begin{aligned} \int_{1/2<|\lambda|<\infty} |f(\lambda+h) - f(\lambda)|^p d\lambda \quad &\leq \quad C^p h^p \int_{1/2}^1 |\lambda|^{-p(\sigma+1)} d\lambda \\ &\qquad + C^p h^p \int_1^\infty |\lambda|^{-p(\delta+1)} d\lambda \\ &\leq \quad C h^p \leq C h^{1-p\sigma}. \end{aligned} \tag{A.3}$$

From (A.1), (A.2) and (A.3) we get

$$\|f(\lambda+h) - f(\lambda)\|_p \leq C h^{1/p-\sigma},$$

implying $f \in \mathrm{Lip}(p, 1/p - \sigma)$. □

Proof of Lemma 4.2. We use the technique of [48], where (4.16) was proved for $\beta = 1$. Since the underlying process $X(t)$ is real-valued, we have for $t > 0$

$$r(t) := \int_{-\infty}^{+\infty} e^{it\lambda} f(\lambda)\, d\lambda = 2 \int_0^\infty \frac{C}{\lambda^{2\alpha}(1+\lambda^2)^\beta} \cos(t\lambda)\, d\lambda. \tag{A.4}$$

Using the change of variable $\lambda = 1/(tu)$, we obtain

$$\begin{aligned} r(t) \quad &= \quad 2C \cdot t^{2\alpha-1} \int_0^\infty \left(\frac{(tu)^2}{1+(tu)^2} \right)^\beta u^{2\alpha-2} \cos(1/u)\, du \\ &= \quad 2C \cdot t^{2\alpha-1} \int_0^\infty L(tu) k(u)\, du, \end{aligned} \tag{A.5}$$

where

$$L(u) = \frac{u^{2\beta}}{(1+u^2)^\beta} \quad \text{and} \quad k(u) = u^{2\alpha-2} \cos(1/u).$$

Choose $\delta > 0$ such that $\delta < \min(1-2\alpha, 2\alpha)$. Then the improper integrals

$$\int_{0+}^{1} u^{-\delta} k(u)\, du \quad \text{and} \quad \int_{1}^{\infty-} u^{\delta} k(u)\, du$$

exist. Therefore, by the Bojanic-Karamata theorem (see [9], Th. 4.1.5) we have

$$\int_{0}^{\infty} k(u) L(tu)\, du \longrightarrow \int_{0+}^{\infty-} k(u)\, du \quad \text{as} \quad t \to \infty. \tag{A.6}$$

Next, using the change of variable $1/u = v$ and the formula (see, e.g., [18])

$$\int_{0}^{\infty-} x^{-p} \cos(mx)\, dx = \frac{\pi m^{p-1}}{2\cos(p\pi/2)\Gamma(p)}, \quad 0 < p < 1, \quad m > 0,$$

with $p = 2\alpha$ and $m = 1$, we obtain

$$\int_{0+}^{\infty-} k(u)\, du \;=\; \int_{0}^{\infty-} u^{2\alpha-2} \cos(1/u)\, du = \frac{\pi}{2\cos(\pi\alpha)\Gamma(2\alpha)}. \tag{A.7}$$

From (A.4)–(A.7) we get (4.16). □

Proof of Lemma 5.3. The proof of properties a)–c) can be found in [5], Lemma 3.2 (see also [33], Lemma 2). To prove d) first observe that for $T > 0$

$$\int_{\mathbb{R}} |D_T(u)|^{p(m)}\, du \le C \cdot T^{p(m)-1} \quad \text{and} \quad |D_T(u)| \le C_\delta \quad \text{for} \quad |u| > \delta. \tag{A.8}$$

For $\mathbf{u} = (u_1, \ldots, u_{m-1}) \in \mathbb{R}^{m-1}$ we have

$$\begin{aligned} \int_{\mathbb{E}_\delta^c} \Phi_T^{p(m)}(\mathbf{u}) d\mathbf{u} \quad &\le \quad \int\limits_{|u_1|>\delta} \Phi_T^{p(m)}(\mathbf{u}) d\mathbf{u} + \int\limits_{|u_2|>\delta} \Phi_T^{p(m)}(\mathbf{u}) d\mathbf{u} \\ &\quad + \ldots + \int\limits_{|u_{m-1}|>\delta} \Phi_T^{p(m)}(\mathbf{u}) d\mathbf{u} \\ &=: \quad I_1 + I_2 + \cdots + I_{m-1}. \end{aligned} \tag{A.9}$$

It is enough to estimate I_1 ($I_2, \ldots, I_n$ can be estimated in the same way). We have

$$\begin{aligned} I_1 \quad &\le \quad \int\limits_{|u_1|>\delta,\ |u_2|>\delta/m} \Phi_T^{p(m)}(\mathbf{u}) d\mathbf{u} + \cdots + \int\limits_{|u_1|>\delta,\ |u_{m-1}|>\delta/m} \Phi_T^{p(m)}(\mathbf{u}) d\mathbf{u} \\ &+ \int\limits_{|u_1|>\delta,\ |u_2|\le\delta/m,\ |u_{m-1}|\le\delta/m} \Phi_T^{p(m)}(\mathbf{u}) d\mathbf{u} \\ &=: \quad I_1^{(2)} + \cdots + I_1^{(m-1)} + I_1^{(m)}. \end{aligned} \tag{A.10}$$

According to (A.8)

$$I_1^{(2)} \leq C_\delta \cdot \frac{1}{T^{p(m)}} \cdot \int\limits_{|u_2|>\delta/m} |D_T(u_2)|^{p(m)} \cdots |D_T(u_{m-1})|^{p(m)}$$
$$\times |D_T(u_1 + \cdots + u_{m-1})|^{p(m)} \, du_1 du_{m-1} \ldots du_2$$
$$\leq C_\delta \cdot \frac{1}{T^{p(m)}} \cdot T^{(p(m)-1)(m-2)} \int\limits_{|u_2|>\delta/m} \frac{1}{|u_2|^{p(m)}} \, du_2 \leq C_\delta. \tag{A.11}$$

Likewise,

$$I_1^{(j)} \leq C_\delta, \qquad j = 3, \ldots, m-1. \tag{A.12}$$

Next, observe that in the integral $I_1^{(m)}$, we have $|u_1 + \cdots + u_{m-1}| > \delta/m,$. Hence by (A.8)

$$I_1^{(m)} \leq C_\delta \cdot \frac{1}{T^{p(m)}} \int\limits_{|u_1|>\delta} D_T^{p(m)}(u_1) \cdots D_n^{p(m)}(u_{m-1}) du_2 \ldots du_{m-1} du_1$$
$$\leq C_\delta \int\limits_{|u_1|>\delta} \frac{1}{|u_1|^{p(m)}} \, du_1 \leq C_\delta. \tag{A.13}$$

From (A.9)–(A.13) we obtain (5.6). Lemma 5.3 is proved. □

Proof of Lemma 5.5. Using Lemma 5.4 and the notation $d\mathbf{u} = du_n du_{n-1} \cdots du_1$, we can write

$$B_1 \leq \int\limits_{\{|u_1|\leq 1\}} \frac{1}{|u_1|^{\beta-\alpha}} \int\limits_{\mathbb{R}^{n-2}} \frac{1}{|u_2 \cdots u_{n-1}|^\beta} \int\limits_{\mathbb{R}} \frac{1}{|u_n(u_1 + \cdots + u_n)|^\beta} \, d\mathbf{u}$$

$$\leq C \int\limits_{\{|u_1|\leq 1\}} \frac{1}{|u_1|^{\beta-\alpha}} \int\limits_{\mathbb{R}^{n-2}} \frac{1}{|u_2 \cdots u_{n-1}|^\beta |(u_1 + \cdots + u_{n-1})|^{2\beta-1}} du_{n-1} \cdots du_1$$

$$\leq C^2 \int\limits_{\{|u_1|\leq 1\}} \frac{1}{|u_1|^{\beta-\alpha}} \int\limits_{\mathbb{R}^{n-3}} \frac{1}{|u_2 \cdots u_{n-2}|^\beta |(u_1 + \cdots + u_{n-2})|^{3\beta-2}} du_{n-2} \cdots du_1$$

$$\leq \quad \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots$$
$$\leq \quad C^{n-2} \int\limits_{\{|u_1|\leq 1\}} \frac{1}{|u_1|^{\beta-\alpha}} \int\limits_{\mathbb{R}} \frac{1}{|u_2|^\beta |(u_1+u_2)|^{(n-1)\beta-n+2}} du_2 du_1$$
$$\leq \quad C^{n-1} \int\limits_{\{|u_1|\leq 1\}} \frac{1}{|u_1|^{(n+1)\beta-\alpha-n+1}} du_1 < \infty,$$

yielding (5.8) for $i = 1$. The quantities $B_2, \ldots, B_n$ can be estimated in the same way. □

Proof of Lemma 5.6. We have

$$I \leq \int\limits_{|u_1|>1} + \cdots + \int\limits_{|u_n|>1} \frac{1}{|u_1 \cdots u_n(u_1 + \cdots + u_n)|^\beta} \, du_1 \cdots du_n =: I_1 + \cdots + I_n. \tag{A.14}$$

Using Lemma 5.4 we get

$$\begin{aligned} I_1 &\leq \int\limits_{|u_1|>1} \frac{1}{|u_1|^\beta} \int\limits_{\mathbb{R}^{n-2}} \frac{1}{|u_2 \cdots u_{n-1}|^\beta} \int\limits_{\mathbb{R}} \frac{1}{|u_n(u_1 + \cdots + u_n)|^\beta} \, du_n \cdots du_1 \\ &\leq C \int\limits_{|u_1|>1} \frac{1}{|u_1|^\beta} \int\limits_{\mathbb{R}^{n-2}} \frac{1}{|u_2 \cdots u_{n-1}|^\beta |u_1 + \cdots + u_{n-1}|^{2\beta-1}} \, du_{n-1} \cdots du_1 \\ &\leq \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ &\leq C^{n-1} \int\limits_{|u_1|>1} \frac{1}{|u_1|^{(n+1)\beta-n+1}} \, du_1 < \infty. \end{aligned}$$

The quantities $I_2, \ldots, I_n$ can be estimated in the same way, and by (A.14) the result follows. □

Acknowledgements

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

This work was supported in part by NSF Grants DMS-0706786 and DMS-1309009 at Boston University.

References

[1] Adenstedt, R.K. (1974). On large-sample estimation for the mean of a stationary random sequence. *Ann. Math. Statist* **2** 1095–1107. MR0368354

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

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

[6] Bercu, B., Gamboa, F. and Rouault, A. (1997). Large deviations for quadratic forms of Gaussian stationary processes. *Stochastic Process. Appl.* **71** 75–90. MR1480640

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

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

[9] Bingham, N.H., Goldie, C.M. and Teugels, J.L. (1989). *Regular Variation*. Cambridge University Press, New York. MR1015093

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

[11] Böttcher, A. and Silbermann, B. (2005). *Analysis of Toeplitz Operators*, 2nd ed. Springer-Verlag, Berlin. MR2223704

[12] Brockwell, P.J and Davis, R.A. (1991). *Time Series: Theory and Methods*, 2nd ed. Springer-Verlag, New York. MR1093459

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

[14] Butzer, P.L. and Nessel, R.J. (1971). *Fourier Analysis and Approximation I*. Academic Press, New York. MR0510857

[15] Coursol, J. and Dacunha-Castelle, D. (1979). Sur la formule de Chernoff pour deux processus gaussiens stationnaires. *C. R. Acad. Sci. Paris Ser.* **A-B 288** A769–A770. MR0535808

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

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

[18] Dwight, H.B. (1961). *Tables of Integrals and Other Mathematical Data*. The Macmillan Company, New York. MR0129577

[19] Flandrin, P. (1989). On the spectrum of fractional Brownian motions. *IEEE Trans. Inform. Theory* **35** 197–199. MR0995341

[20] 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

[21] 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

[22] Gamboa, F., Rouault, A. and Zani, M. (1999). A functional large deviations principle for quadratic forms of Gaussian stationary processes. *Statist. Probab. Lett.* **43** 299–308. MR1708097

[23] 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

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

[25] 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

[26] 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
[27] Ginovyan, M.S. (1994). On Toeplitz type quadratic functionals in Gaussian stationary process. *Probab. Theory Related Fields* **100** 395–406. MR1305588
[28] Ginovyan, M.S. (1995). Asymptotic properties of spectrum estimate of stationary Gaussian process. *J. Cont. Math. Anal.* **30** 1–17. MR1643528
[29] Ginovyan, M.S. (2003). Asymptotically efficient nonparametric estimation of nonlinear spectral functionals. *Acta Appl. Math.* **78** 145–154. MR2024019
[30] Ginovyan, M.S. (2011). Efficient estimation of spectral functionals for Gaussian stationary models. *Communications on Stochastic Analysis* **5** 211–232. MR2808543
[31] Ginovyan, M.S. (2011). Efficient estimation of spectral functionals for continuous-time stationary models. *Acta Appl. Math.* **115** 233–254. MR2818916
[32] Ginovyan, M.S. and Sahakyan, A.A. (2005). On the central limit theorem for Toeplitz quadratic forms of stationary sequences. *Theory Probab. Appl.* **49** 612–628. MR2142560
[33] Ginovyan, M.S. and Sahakyan, A.A. (2007). Limit theorems for Toeplitz quadratic functionals of continuous-time stationary process. *Probab. Theory Related Fields* **138** 551–579. MR2299719
[34] Ginovyan, M.S. and Sahakyan, A.A. (2012). Trace approximations of products of truncated Toeplitz operators. *Theory Probab. Appl.* **56** 57–71. MR2848419
[35] Ginovyan, M.S. and Sahakyan, A.A. (2013). On the trace approximations of products of Toeplitz matrices. *Statist. Probab. Lett.* **83** 753–760. MR3040300
[36] Giraitis, L., Koul, H. and Surgailis, D. (2012). *Large Sample Inference for Long Memory Processes*. Imperial College Press, London. MR2977317
[37] 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
[38] Gohberg, I.C. and Krein, M.G. (1969). *Introduction to the Theory of Linear Nonselfadjoint Operators in Hilbert Space*. American Mathematical Society, Providence, Rhode Island. MR0246142
[39] 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
[40] Grenander, U. and Szegö, G. (1958). *Toeplitz Forms and Their Applications*. University of California Press, Berkeley and Los Angeles. MR0094840

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

[42] 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

[43] 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., Birkhäuser, Boston, 417–438. MR1957502

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

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

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

[47] Ihara, S. (2000). Large deviation theorems for Gaussian processes and their applications in information theory. *Acta Appl. Math.* **63** 165–174. MR1831254

[48] Inoue, A. and Kasahara, Y. (1999). On the asymptotic behavior of the prediction error of a stationary process. In *Trends in Probability and Related Analysis*, Kono, N., Shieh, N.R., Eds., World Scientific, River Edghe, NJ, 207–218. MR1819207

[49] Kac, M. (1954). Toeplitz matrices, translation kernels and a related problem in probability theory. *Duke Math. J.* **21** 501–509. MR0062867

[50] Lieberman, O., Rousseau, J. and Zucker, D.M. (2001). Valid Edgeworth expansion for the sample autocorrelation function under long dependence. *Econometric Theory.* **17** 257–275. MR1863573

[51] 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

[52] Nourdin, I. and Peccati, G. (2009). Stein's method and exact Berry-Esséen asymptotics for functionals of Gaussian fields. *Ann. Probab.* **37** 2231–2261. MR2573557

[53] Palma, W. (2007). *Long-Memory Time Series.* Wiley, New York. MR2297359

[54] Robinson, P.M. (1995). Gaussian semiparametric estimation of long-range dependence. *Ann. Statist.* **23** 1630–1661. MR1370301

[55] 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

[56] Rosenblatt, M. (1961). Independence and dependence. In *Proc. 4th Berkeley Symp. Math. Stat. Probab.* **2** 431–443. MR0133863

[57] Rosenblatt, M. (1962). Asymptotic behavior of eigenvalues of Toeplitz forms. *J. Math. and Mech.* **11** 941–950. MR0150841

[58] Rosenblatt, M. (1979). Some limit theorems for partial sums of quadratic forms in stationary Gaussian variables. *Z. Wahr. verw. Gebiete* **49** 125–132. MR0543988

[59] Samorodnisky, G. and Taqqu, M.S. (1994). *Stable Non-Gaussian Random Peocesses.* Chapman & Hall/CRC, New York. MR1280932

[60] Sato, T., Kakizawa, Y. and Tahiguchi, M. (1998). Large deviations results for statistics of short– and long–memory Gaussian processes. *Austral. & New Zeland J. Statist.* **40** 17–29. MR1628204

[61] Solo, V. (1992). Intrinsic random functions and the paradox of 1/f noise. *SIAM J. Appl. Math.* **52** 270–291. MR1148328

[62] Taniguchi, M. (1983). On the second order asymptotic efficiency of estimators of Gaussian ARMA processes. *Ann. Statist.* **11** 157–169. MR0684873

[63] Taniguchi, M. (1986). Berry-Esséen theorems for quadratic forms of Gaussian stationary processes. *Probab. Theory Related Fields* **72** 185–194. MR0836274

[64] Taniguchi, M. and Kakizawa, Y. (2000). *Asymptotic Theory of Statistical Inference for Time Series.* Academic Press, New York. MR1785484

[65] Taqqu, M.S. (1975). Weak convergence to fractional Brownian motion and the Rosenblatt process. *Z. Wahr. verw. Gebiete* **31** 287–302. MR0400329

[66] 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

[67] 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., Birkhäuser, Boston, 5–38. MR1956042

[68] 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.

[69] 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

[70] 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

[71] Veillette, M. and Taqqu, M.S. (2013). Properties and numerical evaluation of the Rosenblatt distribution. *Bernoulli* **19** 982–1005. MR3079303

[72] Yaglom, A.M. (1986). *The Correlation Theory of Stationary and Related Random Processes, Vol. 1.* Springer, New York.