ON A SZEGÖ TYPE LIMIT THEOREM, THE HÖLDER-YOUNG-BRASCAMP-LIEB INEQUALITY, AND THE ASYMPTOTIC THEORY OF INTEGRALS AND QUADRATIC FORMS OF STATIONARY FIELDS *

Florin Avram[1], Nikolai Leonenko[2] and Ludmila Sakhno[3]

Abstract. Many statistical applications require establishing central limit theorems for sums/integrals $S_T(h) = \int_{t \in I_T} h(X_t)dt$ or for quadratic forms $Q_T(h) = \int_{t,s \in I_T} \hat{b}(t-s)h(X_t, X_s)dsdt$, where X_t is a stationary process. A particularly important case is that of Appell polynomials $h(X_t) = P_m(X_t)$, $h(X_t, X_s) = P_{m,n}(X_t, X_s)$, since the "Appell expansion rank" determines typically the type of central limit theorem satisfied by the functionals $S_T(h), Q_T(h)$. We review and extend here to multidimensional indices, along lines conjectured in [F. Avram and M.S. Taqqu, *Lect. Notes Statist.* **187** (2006) 259–286], a functional analysis approach to this problem proposed by [Avram and Brown, *Proc. Amer. Math. Soc.* **107** (1989) 687–695] based on the method of cumulants and on integrability assumptions in the spectral domain; several applications are presented as well.

1. Introduction

Motivation. Following the work of Szegö, Schur, Wiener, Kolmogorov, and others, the theory of stationary one-dimensional discrete time series, obtained by choosing $I = \mathbb{Z}$, became a very well developed topic, tightly interwoven with several branches of mathematics. It brought forth jewels like for example the Wiener-Kolmogorov formula identifying the variance of the prediction error with respect to the past with the integral of the log of the spectral density. However, transposing these jewels to the multidimensional and continuous parameter case has not always been straightforward. Indeed, modifying the nature of the index set may involve subtle challenges, as may be seen for example from the extension of the Wiener-Kolmogorov formula provided by Krein for the continuous time case $I = \mathbb{R}$.

[1] Dépt de Mathématiques, Université de Pau, France; `Florin.Avram@univ-Pau.fr`

[2] Cardiff School of Mathematics, Cardiff University, Senghennydd Road, Cardiff, CF24 4AG, UK; `LeonenkoN@Cardiff.ac.uk`

[3] Dept. of Probability Theory and Mathematical Statistics, Kyiv National Taras Shevchenko University, Ukraine, and Cardiff School of Mathematics, Cardiff University, Senghennydd Road, Cardiff, CF24 4AG, UK; `lms@univ.kiev.ua`

The convenience of time series comes largely from the FARIMA family of parametric models, defined as solutions of equations

$$\phi(B)(1 - B)^d X_t = \theta(B)\xi_t,$$

where B is the operator of backward translation in time, $\phi(B), \theta(B)$ are polynomials (with the usual assumption that these are the polynomials without common zeros and without zeros in the unit circle in the complex plane), d is a real number and ξ_t is white noise [22,40,45,46,73]. Using the FARIMA family of models, one attempts, *via* an extension of the Box-Jenkins methodology, to estimate the parameter d and the coefficients of the polynomials $\phi(B), \theta(B)$ such that the residuals ξ_t display white noise behavior (and hence may be safely discarded for prediction purposes).

An extension of this approach to continuous time and to multi-parameter processes (spatial statistics) has been long time missing. Only recently, the FICARMA [5,25] and the Riesz-Bessel families of processes (which essentially replace the operator B by the Laplacian – see Appendix E), have allowed a similar approach for processes with continuous and multidimensional indices (at least in the rotationally invariant case).

These examples illustrate the point that extension of time series results to continuous and multidimensional indices, preferably in a unified framework, is an important challenge.

Such a framework can be provided by separable locally compact abelian groups [60]. Let I denote such a group, let S denote the dual locally compact abelian group of its characters, let $\left(\int_{t \in I} \nu(\mathrm{d}t), \int_{l \in S} \mu(\mathrm{d}s) \right)$ denote respectively the integrals with respect to the normalized Haar measures, let $\hat{f}$ denote the Fourier transform, for $f \in L_1(\mathrm{d}\nu) \cap L_2(\mathrm{d}\mu)$ (which ensures the validity of the Plancherel formula), and assume that the Haar measures are associated, *i.e.* that the Parseval identity holds for every continuous complex-valued function of compact support on I. Under this setup, it is possible to provide a unified treatment of discrete and continuous (multi-dimensional) index sets I.

Model. Consider a real, centered, stationary in the strict sense random field $X_t, t \in I$, where I denotes a separable locally compact abelian group endowed with its Haar measure, and $\int_{t \in I} \nu(\mathrm{d}t)$ is integral with respect to the Haar measure. We will focus on the cases when I is $\mathbb{R}^d$ endowed with Lebesgue measure, or $\mathbb{Z}^d$ with the counting measure, to be referred to respectively as the continuous and discrete setups. Note that in these cases, a vector structure is available as well. For the continuous case, we assume that $X_t, t \in \mathbb{R}^d$, is a measurable random field.

Note. Discrete sums might be written below either as integrals (in the statement of theorems), or in traditional sum notation (in the expository part).

We will assume throughout the existence of all order cumulants $c_k(t_1, t_2, ..., t_k)$ for our stationary random field X_t, and also that they are representable as Fourier transforms of "cumulant spectral densities" $f_k(\lambda_1, ..., \lambda_{k-1}) \in L_1$, *i.e.*:

$$c_k(t_1, t_2, ..., t_k) = c_k(t_1 - t_k, .., t_{k-1} - t_k, 0)$$
$$= \int_{\lambda_1, ..., \lambda_{k-1} \in S} \mathrm{e}^{\mathrm{i} \sum_{j=1}^{k-1} \lambda_j (t_j - t_k)} f_k(\lambda_1, ..., \lambda_{k-1}) \, \mu(\mathrm{d}\lambda_1)...\mu(\mathrm{d}\lambda_{k-1}).$$

Throughout, S will denote the "spectral" space of discrete/continuous processes and $\mu(\mathrm{d}\lambda)$ will denote the associated Haar measure, *i.e.* Lebesgue measure normalized to unity on $[-\pi, \pi]^d$ and Lebesgue measure on $\mathbb{R}^d$, respectively. The functions $f_k(\lambda_1, ..., \lambda_{k-1})$ are symmetric and complex valued in general, for the real field X_t.

The random field X_t is observed on a sequence I_T of increasing dilations of a bounded convex domain I_1, *i.e.*

$$I_T = T \, I_1, \quad T \to \infty$$

(for $I = \mathbb{Z}^d$, by convex domains we mean those which can be represented as an intersection $D \bigcap \mathbb{Z}^d$, where D is a convex domain in $\mathbb{R}^d$).

Correspondingly, we will consider linear and bilinear forms S_T and Q_T (see (1.2), (1.3) below), where summation/integration is performed over domains I_T.

In the discrete-time case, the cases $I_T = [1, T]^d$, $T \in \mathbb{Z}_+$ (in keeping with tradition) or $I_T = [-T/2, T/2]^d$, $T \in 2\mathbb{Z}_+$ will be assumed. In the continuous case, we focus on rectangles $I_T = \{t \in \mathbb{R}^d : -T/2 \le t_i \le T/2, i = 1, ..., d\}$.

Later on we will see that the choice of a specific set I_1 leads (when working in the spectral domain) to using an associated Dirichlet type kernel

$$\Delta_T(\lambda) = \int_{t \in I_T} e^{it\lambda} \nu(dt) \tag{1.1}$$

and a multivariate Fejér kernel (1.4). Explicit well-known formulas are available when I_1 is a rectangle or a ball, both for the discrete and continuous case (see Appendix D). We will work in the paper with rectangular domains, however, extensions are possible for balls by adapting the corresponding estimates for the norms of the kernels $\Delta_T(\lambda)$ (see Appendix D).

Central limit theorems for quadratic forms. Asymptotic statistical theory, and in particular, estimation of the parameters of FARIMA and Riesz-Bessel processes, requires often establishing central limit theorems concerning

$$\text{sums/integrals} \qquad S_T(h) = \int_{t \in I_T} h(X_t) \nu(dt) \tag{1.2}$$

$$\text{and bilinear forms} \qquad Q_T(h) = \int_{t,s \in I_T} \hat{b}(t-s) h(X_t, X_s) \nu(ds) \nu(dt) \tag{1.3}$$

of functions of stationary processes X_t with discrete or continuous time and their generalizations to multidimensional setup (or continuous multidimensional parameter).

Note. The kernel of the quadratic form will be denoted by $\hat{b}(t)$, so that we may reserve $b(\lambda)$ for its Fourier transform.

A particularly important case is that of Appell polynomials[1] $h(X_t) = P_m(X_t)$, $h(X_t, X_s) = P_{m,n}(X_t, X_s)$ associated to the distribution of X_t, which are the building block of the so called "chaos/Fock expansions". Two main cases were distinguished, depending on whether the limit is Gaussian or not (the latter case being referred to as a non-central limit theorem).

This line of research, initiated by Dobrushin and Major [27] and Taqqu [70] in the Gaussian case (see also Ivanov and Leonenko [49] for Gaussian continuous case), by Giraitis and Surgailis [35,36], and by Avram and Taqqu [14] in the linear case, continues to be of interest today [33,38,39,56,58,64] to name only a few papers.

Our interest here is in obtaining extensions to continuous and multi-parameter processes of the central limit theorem for sums and quadratic form, obtained in the case of discrete time series by Breuer and Major [24], by the method of moments. Note, that applying this method, that is, showing that all the moments converge to those of Gaussian distribution (a distribution which is uniquely identifiable by its moments), we will get convergence in distribution, but convergence of all moments give us even stronger statement. In the results below, where we state convergence in distribution to Gaussian law, in fact, convergence of all moments holds (see, Thm. 1.3, Cor. 3.3, and Thms. 3.5, 4.1, 4.3, 4.7, E.2).

Analytic tools. A key unifying role in our story is played by the multivariate Fejér kernels

$$\Phi_T^*(u_1, ..., u_{n-1}) = \frac{1}{(2\pi)^{(n-1)d} \mu(I_1) T^d} \Delta_T \left(-\sum_{e=1}^{n-1} u_e \right) \prod_{e=1}^{n-1} \Delta_T(u_e), \tag{1.4}$$

and by their well-known kernel property, *i.e.* the fact that when $T \to \infty$, the multivariate Fejér kernels Φ_T^* converge weakly to a δ measure. More precisely

---

[1] For the definition of Appell polynomials see, for instance, Avram and Taqqu [14], Giraitis and Surgailis [36], or Appendix A.

Lemma 1.1. The kernel property. *For any bounded function $C(u_1, ..., u_{n-1})$, which is continuous at a point $(\alpha_1, ..., \alpha_{n-1})$, the multivariate Fejér kernel with I_1 a rectangle satisfies:*

$$\lim_{T \to \infty} \int_{S^{n-1}} C(u_1 - \alpha_1, ..., u_{n-1} - \alpha_{n-1}) \Phi_T^*(u_1, ..., u_{n-1}) \prod_{i=1}^{n-1} \mu(\mathrm{d}u_i) = C(\alpha_1, ..., \alpha_{n-1}).$$

For the proof of this result, we refer when $I = \mathbb{Z}$, for example, to [12], which is easily extendable to $I = \mathbb{Z}^d$ [2]. For the continuous (as well as discrete) multidimensional case, with I_1 a rectangle, we refer to Bentkus [20,21] (see also [9], Prop. 1), where it is shown that the kernels $\Phi_T^*(u_1, ..., u_{n-1})$ constitute an approximate identity for convolution (as defined, *e.g.*, in Rudin [62]).

Developing limit theory for multivariate Fejér kernels was the key point in several papers [11–13,15] which generalized the Breuer and Major central limit theorem [24]. The papers above introduced a new mathematical object to be called "Fejér graph integrals" – see (3.4) in Section 3, which captures the common structure of several cumulant computations. Replacing the cyclic graph encountered in the specific case of quadratic forms in Gaussian random variables by an arbitrary graph, these papers reduce the central limit theorem for a large class of problems involving Appell polynomials in Gaussian or moving average summands to an application of three analytical tools:

(1) Identifying the graphs involved by applying the well-known diagram formula for computing moments/cumulants of Wick products – see Appendix B.
(2) Applying a generalization of a Grenander-Szegö theorem on the trace of products of Toeplitz matrices to the Fejér graph integrals – see Section 3, to obtain the asymptotic variance. This theorem is valid under some general integrability assumptions furnished by the Hölder-Young-Brascamp-Lieb inequality.
(3) The resolution of certain combinatorial graph optimization problems, specific to each application, which clarify the geometric structure of the polytope of valid integrability exponents on the functions involved (spectral density, kernel of the quadratic form, etc.) and allows therefore to evaluate the integrals representing higher order (≥ 3) cumulants.

Here, we observe that a similar approach works in the multidimensional and continuous indices case. More precisely, the only changes are (a) the normalizations, which change from T to T^d; and (b) the condition for the validity of the Hölder-Young-Brascamp-Lieb inequality (see Appendix C) in the continuous case. Therefore, the previously obtained central limit theorems continue to hold in the multidimensional case, including continuous indices, after simply adjusting the normalizations and integrability conditions.

The two cases most easy to study are that of Gaussian and that of linear processes, defined in the discrete case by

$$X_t = \sum_{u \in \mathbb{Z}^d} \hat{a}(t - u)\xi(u), \ \sum_{u \in \mathbb{Z}^d} \hat{a}^2(u) < \infty, \ t \in \mathbb{Z}^d,$$

where $\xi(u)$, $u \in \mathbb{Z}^d$, are independent random variables (see (5.1) and (5.3) in the applications Section 5 for the continuous case and further details). This assumption has the advantage of implying a product representation of the cumulant spectral densities – see for example Theorem 2.1 of [5]. Namely, for the cumulants we get the explicit formula

$$c_k(t_1, ..., t_k) = d_k \int_{s \in I} \prod_{j=1}^{k} \hat{a}(t_j - s) \ \nu(\mathrm{d}s),$$

---

[2]The convergence of measures there was derived as a consequence of the Hölder-Young-Brascamp-Lieb inequality (see Thm. C.1), using estimates of the form

$$||\Delta_T||_{s_v^{-1}} \leq k(s_v) T^{d(1-s_v)}$$

with optimally chosen s_v, $v = 1, ..., V$.

where d_k is the k'th cumulant of $\xi(0)$. This yields in the spectral domain

$$f_k(\lambda_1, ..., \lambda_{k-1}) = d_k \, a\left(-\sum_{i=1}^{k-1}\lambda_i\right)\prod_{i=1}^{k-1}a(\lambda_i) = \mathrm{d}_k\prod_{i=1}^{k}a(\lambda_i)\delta\left(\sum_{j=1}^{k}\lambda_j\right). \tag{1.5}$$

For $k = 2$, the spectral density $f_2(\lambda) = d_2 a(\lambda)a(-\lambda)$ will be denoted simply by $f(\lambda)$.

We introduce now a class of spaces needed below:

Definition 1.2. Let

$$\mathbf{L}_p(\mathrm{d}\mu) = \begin{cases} L_p(d\mu) & \text{if } p < \infty, \\ C_0 & \text{if } p = \infty \end{cases} \tag{1.6}$$

that is, the only difference in comparison with the usual definition of $L_p(\mathrm{d}\mu)$ spaces is that for $p = \infty$ it will be supposed that $\mathbf{L}_\infty$ is the space C_0 of all continuous functions which "vanish at infinity" (that is, functions f such that for every $\varepsilon > 0$ there exists a compact set K such that $|f(x)| < \varepsilon$ for all x not in K). With this definition our spaces $\mathbf{L}_p$, $1 \le p \le \infty$, will be the closure of the continuous functions of bounded support under the L_p norm, with L_∞ norm being, as usual, uniform supremum norm (see, *e.g.*, Rudin [61], Chap. 3). Note, that in the discrete case our $\mathbf{L}_p$ spaces just coincide with L_p for all $1 \le p \le \infty$.

Note. In the continuous case (when "spectral" space S is $\mathbb{R}^d$) we will need these $\mathbf{L}_p$ spaces for the parts b) and c) of Theorem 4.1: in such a way we can easily adopt the approach used in the torus case ($S = [-\pi, \pi]^d$). Recall that in the torus case [11] the approach was used, when the statements were established first for complex exponentials, and extended then to the Banach space of functions which may be approximated arbitrarily close in $L_p(\mathrm{d}\mu)$ sense by linear combinations of complex exponentials, endowed with the L_p norm). Analogously, here we will show that the statements hold for continuous functions with bounded support and, therefore, can be extended to spaces $\mathbf{L}_p$.

Considering bilinear forms Q_T we will work under the following integrability assumption on the spectral density f and the Fourier transform of the kernel of the quadratic form b:

Assumption A. For some $1 \le p_i \le \infty$, $i = 1, 2$,

$$f \in \mathbf{L}_{p_1}(\mathrm{d}\mu), b \in \mathbf{L}_{p_2}(\mathrm{d}\mu).$$

Note. 1) While a general stationary model is parameterized by a sequence of functions $f_k(\lambda_1, ..., \lambda_{k-1})$, $k \in \mathbb{N}$, the linear model (1.5) is considerably simpler, being parameterized by a single function $a(\lambda)$.

2) We note that all our results may be formulated directly in terms of characteristics of the field X_t, which suggests that the moving average assumption may be unnecessary; indeed, more general results which make direct assumptions that functions $f_k(\lambda_1, ..., \lambda_{k-1})$ belong to some special L_p-type spaces, have been obtained in certain cases – see, for example, [13].

Contents. We present a warmup example involving quadratic Gaussian forms in Section 2. The results here are closely connected to those of the paper by Ginovian and Sahakyan (2007). We define the concept of Fejér graph and matroid integrals in Section 3. We will consider here only the first case (*i.e.* graphic matroids associated to the incidence matrix of a graph).

Our main result is Theorem 3.1, a limit theorem (of Grenander-Szegö type) for Fejér graph integrals in the discrete and continuous multiparameter cases. Its importance is amply demonstrated by several new limit theorems for the Gaussian and linear model, presented in Sections 2, 3, 5. In particular, see Theorem 4.5, analogous to a result by Ivanov and Leonenko [49], however stated here under spectral domain conditions, and Theorem 1.3, generalizing to random fields results by Ginovian [33] and Ginovian and Sahakyan [34]. For the class of linear random fields with continuous parameter based on the stochastic integration with respect

to independently scattered random measures with infinite divisible distributions, we present a central limit theorem for quadratic forms (Thm. 4.1). Also, we present a CLT for linear and quadratic forms of the linear Riesz-Bessel motion (Thm. E.2). We consider possible applications of these results to the minimum contrast estimation of unknown parameter of linear stochastic processes and state the results on asymptotic normality of the estimates under weaker conditions than in [8,52].

To make the paper self-contained we supply in Appendices the material we refer to in the main part of the paper. Various estimates concerning kernels are collected in Appendix B.4, and a particular case of the Hölder-Young-Brascamp-Lieb inequality required here is presented in Appendix C.

2. An example: the central limit theorem for Gaussian bilinear forms

We present first our method in the simplest case of symmetric bilinear forms

$$Q_T = Q_T^{(1,1)} = Q_T(P_{1,1}) = \int_{t,s \in I_T} [X_t X_s - \mathbb{E} X_t X_s] \hat{b}(t-s) \nu(\mathrm{d}s) \nu(\mathrm{d}t), \tag{2.1}$$

in stationary zero-mean Gaussian fields X_t, with covariances $r(t-s)$, $t, s \in I$ and spectral density $f(\lambda)$ (note that $S_T^{(1)} = S_T(P_1) = \int_{t \in I_T} X_t \nu(\mathrm{d}t)$ is "too simple" for our purpose, since it is already Gaussian and its kth order cumulants $\chi_k(S_T^{(1)}) = 0$, $\forall k \neq 2$). The presentation follows [11] for the discrete case and [33] for the continuous case, except that we clarify the point that the previous results hold in any dimension d.

To obtain the central limit theorem for $T^{-d/2} Q_T^{(1,1)}$ by the method of cumulants it is enough to show that:

$$\lim_{T \to \infty} \chi_2 \left(\frac{Q_T^{(1,1)}}{T^{d/2}} \right) \text{ is finite, and} \quad \lim_{T \to \infty} \chi_k \left(\frac{Q_T^{(1,1)}}{T^{d/2}} \right) = 0, \ \forall k \geq 3.$$

A direct computation based on multilinearity (see, *e.g.* [41,47], or apply Appendix B.2) yields the cumulants of $Q_T^{(1,1)}$:

$$\chi_k = \chi(Q_T, ..., Q_T) = 2^{k-1}(k-1)! \operatorname{Tr}[(T_T(b)T_T(f))^k]. \tag{2.2}$$

Here, Tr denotes the trace and

$$T_T(b) = (\hat{b}(t-s), \ t, s \in I_T), \qquad T_T(f) = (r(t-s), \ t, s \in I_T)$$

denote Toeplitz matrices (with multidimensional indices) of dimension $T^d \times T^d$ in the discrete case and truncated Toeplitz-type operators in the continuous case[3].

While the cumulants in (2.2) may be expressed using powers of two Toeplitz matrices, it turns out more convenient in fact to consider more general products with all terms potentially different (taking advantage thus of multilinearity).

Suppose therefore given a set $f_e(\lambda) : S \to \mathbb{R}$, $e = 1, \ldots, n$ of "symbols" associated to the set of Toeplitz operators, where $(S, \mathrm{d}\mu)$ denotes either $\mathbb{R}^d$ with Lebesgue measure, or the torus $[-\pi, \pi]^d$ with normalized Lebesgue measure. Assume the symbols satisfy integrability conditions

$$f_e \in \mathbf{L}_{p_e}(S, \mathrm{d}\mu), \ 1 \leq p_e \leq \infty. \tag{2.3}$$

---

[3]Recall that in continuous case, the truncated Toeplitz-type operator generated by a function $f \in L_1$ ($\hat{f} \in L_\infty$) is defined for $u \in L_2$ as follows

$$T_T(f)u(t) = \int_{I_T} \hat{f}(t-s)u(s)\nu(\mathrm{d}s).$$

Let $\hat{f}_e(t), t \in I$, be the Fourier transform of $f_e(\lambda)$:

$$\hat{f}_e(t) = \int_S e^{it\lambda} f_e(\lambda)\mu(d\lambda), \quad t \in I,$$

where $I = \mathbb{Z}^d$ in the torus case and $I = \mathbb{R}^d$ in the case $S = \mathbb{R}^d$, respectively. In this last case, we would also need to assume that $f_e \in L_1(\mathbb{R}^d, d\mu)$, for the Fourier transform to be well defined. Consider the extension of our cumulants:

$$\tilde{J}_T = \text{Tr}\left[\prod_{i=1}^n T_T(f_i)\right] = \int_{t_1,...,t_n \in I_T} \hat{f}_1(t_2 - t_1)\hat{f}_2(t_3 - t_2)...\hat{f}_n(t_1 - t_n) \prod_{v=1}^n \nu(dt_v). \tag{2.4}$$

Replacing the sequences $\hat{f}_e(t)$ by their Fourier representations $\hat{f}_e(t) = \int_S f_e(\lambda)e^{it\lambda}\mu(d\lambda)$ in (2.4) yields the following alternative spectral integral representation for traces of products of Toeplitz matrices or truncated Toeplitz operators[4]:

$$\begin{aligned}
J_T &= \int_{\lambda_1,...,\lambda_n \in S} f_1(\lambda_1)f_2(\lambda_2)...f_n(\lambda_n) \prod_{e=1}^n \Delta_T(\lambda_{e+1} - \lambda_e) \prod_{e=1}^n \mu(d\lambda_e) \tag{2.5} \\
&= \int_{u_1,...,u_{n-1} \in S} \left(\int_{\lambda \in S} f_1(\lambda)f_2(\lambda + u_1)...f_n\left(\lambda + \sum_1^{n-1} u_e\right)\mu(d\lambda)\right) \Phi_T(u_1,...,u_{n-1}) \prod_{e=1}^{n-1} \mu(du_e) \\
&= \int_{u_1,...,u_{n-1} \in S} C(u_1,...,u_n)\Phi_T(u_1,...,u_{n-1}) \prod_{e=1}^{n-1} \mu(du_e),
\end{aligned}$$

where the index $n + 1$ is defined to be equal 1, where

$$\Phi_T = \Delta_T\left(-\sum_1^{n-1} u_e\right) \prod_{e=1}^{n-1} \Delta_T(u_e),$$

is, after normalization with the factor $\frac{1}{(2\pi)^{(n-1)d}T^d}$, the "multivariate Fejér kernel" (1.4), and where the inner integral

$$C(u) = C_{(f_1,...,f_n)}(u_1,...,u_{n-1}) := \int_{\lambda \in S} f_1(\lambda)f_2(\lambda + u_1)...f_n(\lambda + \sum_1^{n-1} u_e)\mu(d\lambda), \tag{2.6}$$

will be called a graph convolution.

The first expression in (2.5) is our first example of a "Fejér graph integral" to be introduced in general in the next section. These are integrals involving products of Dirichlet kernels Δ_T and functions, applied to linear combinations, which are related to the vertex-edge incidence structure of a certain directed graph (in the occurrence, the cyclic graph on the vertices $\{1,...,n\}$).

The second expression in the RHS of (2.5) reveals the asymptotic behavior of Fejér graph integrals, since, as noted, when $T \to \infty$, the multivariate Fejér kernel Φ_T^* convergences weakly to a δ measure.

Note now that the inner integral/graph convolution $C_{(f_1,...,f_n)}(u_1,...,u_{n-1})$ is well defined precisely under the classical Hölder conditions, *i.e.* when the integrability indices in (2.3) satisfy:

$$\begin{cases} \sum_e p_e^{-1} & \leq 1 \text{ when } S = \mathbb{Z}^d, \\ \sum_e p_e^{-1} & = 1 \text{ when } S = \mathbb{R}^d, \end{cases} \tag{2.7}$$

---

[4] Of course, the two expressions $\tilde{J}_T, J_T$ are equal if $f_e \in L_1, e = 1,...,n$. Note however that the "spectral representation" (2.5) is well defined even without the last condition.

(and it is only under these conditions that the second expression in (2.5) is valid). It turns out that even more is true under these conditions:

Lemma 2.1. *The "graph convolution" function $C(u_1, ..., u_{n-1})$ defined in (2.6) is bounded and continuous if (2.3) holds with integrability indices satisfying the condition (2.7).*

Proof. Note that the function $C : S^{n-1} \to \mathbb{R}$ is a composition

$$C(u_1, ..., u_{n-1}) = \mathcal{C}_{(T_1(u_1, ..., u_{n-1}), ..., T_n(u_1, ..., u_{n-1}))}$$

of the functional

$$\mathcal{C}_{(f_1, ..., f_n)} : \prod_{e=1}^{n} \mathbf{L}_{p_e} \to \mathbb{R}.$$

defined by[5]

$$\mathcal{C}_{(f_1, ..., f_n)} := \int_{\lambda \in S} \prod_{e=1}^{n} f_e(\lambda) \mu(\mathrm{d}\lambda)$$

with the continuous functionals

$$T_e(u_1, ..., u_{n-1}) : S^{n-1} \to \mathbf{L}_{p_e}$$

defined by $T_e(u_1, ..., u_{n-1}) = f_e(\cdot + \sum_{v=1}^{e-1} u_v), e = 1, ..., n.$

Indeed, the continuity of the functionals T_e is clear when f_e is a function which is continuous and of bounded support, and this continues to be true for functions $f_e \in \mathbf{L}_{p_e}$, by our definition of these spaces. In conclusion, under the Hölder assumptions, the continuity of the functional $C(u_1, ..., u_{n-1})$ will follow automatically from that of $\mathcal{C}_{(f_1, ..., f_n)}$.

Finally, under the "Hölder conditions" (2.7), the continuity as well as boundedness of the function $\mathcal{C}_{(f_1, ..., f_E)}$ follow from its multilinearity and from the Hölder inequality:

$$|\mathcal{C}_{(f_1, ..., f_n)}| \leq \prod_{e=1}^{n} \|f_e\|_{p_e}. \qquad \square$$

Finally, since the function $C(u_1, ..., u_{n-1})$ is bounded and continuous at the point $(0, ..., 0)$, we may apply the kernel Lemma 1.1 to the last expression in (2.5), to conclude:

$$T^{-d} J_T \to (2\pi)^{(n-1)d} C(0, ..., 0) = (2\pi)^{(n-1)d} \int_{\lambda \in S} f_1(\lambda) f_2(\lambda) ... f_n(\lambda) \mu(\mathrm{d}\lambda).$$

We summarize this in the next result:

Theorem 2.2. *Let $f_e \in \mathbf{L}_{p_e}, e = 1, ..., n$, where $1 \leq p_e \leq \infty, e = 1, ..., n$, and let $J_T, \tilde{J}_T$ be defined by (2.5), (2.4) respectively. Then, it follows that:*

(1) *If the integrability indices satisfy the Hölder conditions*

$$\begin{cases} \sum_e p_e^{-1} & \leq 1 \text{ when } S = \mathbb{Z}^d \\ \sum_e p_e^{-1} & = 1 \text{ when } S = \mathbb{R}^d, \end{cases}$$

then

$$\lim_{T \to \infty} T^{-d} J_T = (2\pi)^{(n-1)d} \int_{\lambda \in S} \prod_{e=1}^{n} f_e(\lambda) \mu(\mathrm{d}\lambda). \qquad (2.8)$$

---

[5]Note again this is well defined precisely under the classical Hölder conditions.

(2) *If* $\alpha := \sum_e p_e^{-1} > 1$, *then it holds that:*

$$J_T = o(T^{\alpha d}).$$

(3) *If in the continuous case it holds in addition that if* $f_e \in \mathbf{L}_{p_e} \cap L_1, e = 1, ..., n$, *then Fourier coefficients* $\hat{f}_e(k)$ *may be defined, and the previous results hold also for* $\tilde{J}_T = \mathrm{Tr}[\prod_{e=1}^n T_T(f_e)]$.

Proof. Following the proof of [12], we see that part 1 of Theorem 2.2 follows from Lemmas 1.1, 2.1 above, for part 2 we refer to the proof of Theorem 3.1(c) bellow, where more general case is considered.   □

Notes. 1) In the case (2), the exact exponent of magnitude is unknown, except for the upper bound αd.

2) The result above is a refinement of a limit theorem of Grenander and Szegö [41] concerning traces of products of truncated Toeplitz operators. Under the current strengthened integrability conditions, it was obtained when $d = 1$ in the discrete case in [11]; as we show below, the result holds in fact in any dimension d, in discrete and continuous case (after modifying the necessary integrability condition in accordance to the Hölder-Young-Brascamp-Lieb inequality).

We may apply now the previous result to establish the CLT for $Q_T^{(1,1)}$. The kth cumulant is of the form (2.4) with spectral representation (2.5), where $n = 2k, f_{2i-1} = f, f_{2i} = b, i = 1, ..., k$. In particular, part 1 of the theorem above yields the convergence of the normalized variance:

$$\lim_{T\to\infty} \frac{\chi_{2,T}(Q_T^{(1,1)})}{T^d} = \lim_{T\to\infty} \frac{2\mathrm{Tr}[(T_T(b)T_T(f))^2]}{T^d}$$
$$= 2(2\pi)^{3d}C_{f,f,b,b}(0,...,0) = 2(2\pi)^{3d} \int_{\lambda\in S} f^2(\lambda)b^2(\lambda)\mu(\mathrm{d}\lambda)$$

provided the conditions of part (1) are satisfied for $n = 4, f_1 = f_2 = f, f_3 = f_4 = b$, *i.e.* that

$$2\left(\frac{1}{p_f} + \frac{1}{p_b}\right) \le 1 \quad \text{in the discrete case,} \tag{2.9}$$
$$2\left(\frac{1}{p_f} + \frac{1}{p_b}\right) = 1 \quad \text{in the continuous case.}$$

We note now that the conditions (2.9) imply also the fact that the cumulants of order $k \ge 3$ converge to 0

$$\lim_{T\to\infty} \frac{\chi_{k,T}}{T^{\mathrm{d}k/2}} = 0.$$

Indeed, let $z = z_f + z_b$. If $kz \le 1$, convergence to 0 is implied by part 1 and by the normalization; if $1 < \alpha_k := kz \le k/2$, it is implied by part 2 (finally, $kz > k/2$ is impossible by (2.9)).

We arrive thus at the following asymptotic normality result, generalizing Avram [10], Ginovian [33] and Ginovian and Sahakyan [34].

Theorem 2.3. *Consider the quadratic functional* $Q_T^{(1,1)}$ *defined by (2.1), where* $X_t, \quad t \in I$, *is a zero-mean Gaussian random field with spectral density* $f(\lambda) \in \mathbf{L}_{p_f}$. *Assume the generating function* b *of the quadratic functional is such that* $b(\lambda) \in \mathbf{L}_{p_b}$, *and in the discrete and continuous cases it holds respectively that:*

$$\frac{1}{p_f} + \frac{1}{p_b} \le \frac{1}{2}, \quad \frac{1}{p_f} + \frac{1}{p_b} = \frac{1}{2}.$$

In the continuous case, assume also that $b(\lambda) \in L_2, f(\lambda) \in L_1$.

Then, the central limit theorem holds:

$$T^{-d/2}Q_T \to N(0, \sigma^2), \quad T \to \infty,$$

where

$$\sigma^2 := 2(2\pi)^{3d} \int_S b^2(\lambda)f^2(\lambda)\mu(\mathrm{d}\lambda). \tag{2.10}$$

Let us summarize the key steps leading to this CLT, with a view to generalizing:

(1) The cumulants of interest have an explicit formula (2.2) which gives rise in the spectral domain to the "Fejér graph integral" (2.5): the integral of a product involving spectral densities and Fejér kernels.

(2) Consider the Fejér graph integral corresponding to the variance. Once integrability conditions are assumed ensuring that its inner part (involving the spectral densities) called "graph convolution", is well-defined, the graph convolution is automatically bounded and continuous, and therefore a limit theorem for the variance holds as well (Thm. 2.2).

(3) Finally, we need to check whether the previous integrability conditions ensuring that the limiting variance is well-defined are also sufficient to imply that the normalized higher order cumulants converge to 0. We will see later that when checking this, the specific structure of the graphs involved may efficiently be exploited (which is the main advantage of this approach).

3. Fejér graph/matroid integrals and graph/matroid convolutions

In this section, we introduce a unifying graph-theoretical framework for problems similar to those of the previous section.

Definition 3.1. Let $G = (\mathcal{V}, \mathcal{E})$ denote a directed graph with V vertices, E edges, a basis of C independent cycles[6] and $co(G)$ components. The incidence matrix $M = \{M_{v,e}\}_{v \in \mathcal{V}, e \in \mathcal{E}}$ of the graph is the $V \times E$ matrix with entries

$$M_{v,e} = \begin{cases} 1 & \text{if the vertex } v \text{ is the end point of the edge } e, \\ -1 & \text{if the vertex } v \text{ is the start point of the edge } e, \\ 0 & \text{if } v \notin e. \end{cases}$$

A circuit matrix M^* is a $C \times E$ matrix whose rows are obtained by assigning arbitrary orientations to a basis of circuits (cycles) $c = 1, ..., C$ of the graph, and by writing each edge as a sum of $\pm$ the circuits it is included in, with the $\pm$ sign indicating a coincidence or opposition to the orientation of the cycle .

Besides the graph framework, we will hint also to possible matroid generalizations (not surprising, since, quoting Tutte [71]: "it is probably true that any theorem about graphs expressible in terms of edges and circuits exemplifies a more general result about vector matroids").

However, while we use occasionally matroid terminology in the exposition, we will only state results for the graph case, since the general case is not yet fully understood.

Let us recall briefly that matroids are a concept which formalizes the properties of the "rank function" $r(A)$ obtained by considering the rank of an arbitrary set of columns A in a given arbitrary matrix M. More precisely, a matroid is a pair $\mathcal{E}, r : 2^{\mathcal{E}} \to \mathbb{N}$ of a set $\mathcal{E}$ and a "rank like function" $r(A)$ defined on the subsets of $\mathcal{E}$. Matroids may also be defined in equivalent ways *via* their independent sets, *via* their bases (maximal independent sets), *via* their circuits (minimal dependent sets), *via* their spanning sets (sets containing a basis), or *via* their flats (sets which may not be augmented without increasing the rank). For precise definitions and for excellent expositions on graphs and matroids, see [57**?**] or [72].

The most familiar matroids, called *vectorial matroids*, are defined by the set $\mathcal{E}$ of columns of a matrix and by the rank function $r(A)$ which gives the rank of any set of columns A (matrices with the same rank function yield the same matroid).

Some useful facts from matroid theory are the fact that to each matroid M one may associate a dual matroid, with rank function

$$r^*(A) = |A| - r(M) + r(\mathcal{E} - A).$$

---

[6]A basis of cycles is a set of cycles, none of which may be obtained *via* addition modulo 2 of other cycles, after ignoring the orientation.

For vectorial matroids, the dual matroid is also vectorial, associated to any matrix whose rows span the space orthogonal to the rows of M. Furthermore, in the case of graphic matroids, the dual matroid is associated to the circuit matrix.

Tutte's "conjecture" holds true in our case: a matroid Szegö-type limit theorem was already given in [11]. However, for simplicity, we will restrict ourselves here to the particular case of graphic matroids associated to the incidence matrix M of an oriented graph. In this case, the proofs are more intuitive, due to the fact that the algebraic dependence structures translate into graph-theoretic concepts, like circuits corresponding to cycles, etc.

Let $(S, \mathrm{d}\mu)$ denote either $\mathbb{R}^d$ with Lebesgue measure, or the torus $[-\pi, \pi]^d$ with normalized Lebesgue measure, and let $f_e(\lambda) : S \to \mathbb{R}, \ e = 1, \ldots, E$ denote a set of functions ("base functions") associated to the columns of M, which satisfy integrability conditions

$$f_e \in \mathbf{L}_{p_e}(S, \mathrm{d}\mu), \text{ for some } p_e : \ 1 \le p_e \le \infty. \tag{3.1}$$

Let $\hat{f}_e(t), t \in I$, be the Fourier transform of $f_e(\lambda)$:

$$\hat{f}_e(t) = \int_S \mathrm{e}^{\mathrm{i}t\lambda} f_e(\lambda)\mu(\mathrm{d}\lambda), \quad t \in I,$$

where $I = \mathbb{Z}^d$ in the torus case and $I = \mathbb{R}^d$ in the case $S = \mathbb{R}^d$, respectively. In this last case, we would also need to assume that $f_e \in L_1(\mathbb{R}^d, \mathrm{d}\mu)$, for the Fourier transform to be well defined. However, all our analytic results concern the spectral domain, and hence this assumption will not be necessary.

Our object of interest, in its "time domain representation", is:

$$\tilde{J}_T = \tilde{J}_T(M, f_e, e = 1, ..., E) = \int_{s_1, ..., s_V \in I_T} \hat{f}_1(t_1)\hat{f}_2(t_2)...\hat{f}_E(t_E) \prod_{v=1}^V \nu(\mathrm{d}s_v), \tag{3.2}$$

where $t = (t_1, ..., t_E) = (s_1, ..., s_V)M = sM$, where $\nu(\mathrm{d}s_v)$ stands for Lebesgue measure and counting measure, respectively.

Note. To keep the transparent analogy with the case $d = 1$, we make the following convention concerning notations. Here and in what follows we will treat a product of a vector, whose components are d-dimensional, and a matrix (or another vector) with scalar components in a specific sense: we will still perform multiplication component-wise according to the usual rule, and as a result we obtain a vector, whose components are d-dimensional again (or, correspondingly, just a d-dimensional vector). Thus, writing above $t = (t_1, ..., t_E) = (s_1, ..., s_V)M = sM$ we mean that each d-dimensional $t_k, \ k = 1, ..., E$, is representable as a linear combination of d-dimensional vectors $s_l, \ l = 1, ..., V$, with coefficients $M_{k,l}$: $t_k = \sum_{l=1}^V s_l M_{l,k}$ (that is, as multiplication of the row $(s_1, ..., s_v)$ by kth column of M).

A Fejér graph integral is the expression obtained by replacing the sequences $\hat{f}_e(t)$ in (3.2) by their Fourier representations $\hat{f}_e(t) = \int_S f_e(\lambda)\mathrm{e}^{\mathrm{i}t\lambda}\mu(\mathrm{d}\lambda)$: under the assumption $f_e \in L_1(S, \mathrm{d}\mu)$, an easy computation (see [12], Lem. 1) shows that (3.2) may be written also as the integral (3.4) below. We introduce however a more general concept.

Definition 3.2. Let $(S, \mathrm{d}\mu)$ denote either $\mathbb{R}^d$ with Lebesgue measure, or the torus $[-\pi, \pi]^d$ with normalized Lebesgue measure. Let M be a matrix of dimensions $V \times E$, with arbitrary coefficients in the first case and with integer coefficients in the second case. Let $f_e(\lambda) : S \to \mathbb{R}, \ e = 1, \ldots, E$ denote a set of functions associated to the columns of M. Suppose these functions satisfy integrability conditions

$$f_e \in \mathbf{L}_{p_e}(S, \mathrm{d}\mu), \ 1 \le p_e \le \infty. \tag{3.3}$$

A Fejér matroid integral (FMI) is defined by the following "spectral representation":

$$
\begin{aligned}
J_T &= J_T(M, f_e, e = 1, ..., E) && (3.4) \\
&= \int_{\lambda_1,...,\lambda_E \in S} f_1(\lambda_1)f_2(\lambda_2)...f_E(\lambda_E) \prod_{v=1}^{V} \Delta_T(u_v) \prod_{e=1}^{E} \mu(\mathrm{d}\lambda_e), \\
&= \int_{\lambda_1,...,\lambda_E \in S} \int_{s_1,...,s_V \in I_T} f_1(\lambda_1)f_2(\lambda_2)...f_E(\lambda_E) \mathrm{e}^{\langle sM,\lambda \rangle} \prod_{v=1}^{V} \nu(\mathrm{d}s_v) \prod_{e=1}^{E} \mu(\mathrm{d}\lambda_e), && (3.5)
\end{aligned}
$$

where $\Delta_T(u)$ is the kernel defined by (1.1), where $(u_1, ..., u_V)' = M(\lambda_1, ..., \lambda_E)'$ (in the torus case the linear combinations being computed modulo $[-\pi, \ \pi]^d$), and where $s = (s, ..., s_V)$.

A Fejér matroid integral will be called a Fejér graph integral for graphic matroids associated to the incidence matrix M of a directed graph G. In this case, the functions and kernels in (3.4) are associated respectively to the edges and vertices of the graph.

The cycle graph/Toeplitz example. Consider the particular case of a cyclic graph with n edges. In this case, the matrix M with n columns and rows, is:

$$
M = \begin{pmatrix}
-1 & 0 & 0 & . & ... & 0 & 1 \\
1 & -1 & 0 & 0 & ... & 0 & 0 \\
0 & 1 & -1 & \ddots & ... & \vdots & \vdots \\
\vdots & & \ddots & \ddots & & & \\
0 & 0 & & \ddots & & -1 & 0 \\
0 & 0 & . & . & . & 1 & -1
\end{pmatrix}
$$

and its Fejér graph integral is given by:

$$
J_T = \int_{\lambda_1,...,\lambda_n \in S} f_1(\lambda_1)f_2(\lambda_2)...f_n(\lambda_n) \prod_{v=1}^{n} \Delta_T(\lambda_v - \lambda_{v+1}) \prod_{e=1}^{n} \mu(\mathrm{d}\lambda_e),
$$

the integral, which we have already encountered in Section 1.

Note. For analytical results concerning only Fejér matroid integrals as defined by (3.4), the condition $f_e \in L_1(S, \mathrm{d}\mu)$ is unnecessary.

4. LIMIT THEORY FOR FEJÉR GRAPH INTEGRALS

The main points of the limit theory for Fejér graph integrals, to be presented now, are that:

(1) Upper bounds for Fejér graph integrals may always be obtained. They involve a specific function $\alpha_M(z)$, related to the graph structure and to the integrability indices p_e, which is defined in (4.6) below (as the value of an optimization problem). Its computation being a key point, we expound on it in Section 4.1.

(2) Under certain Hölder-Young-Brascamp-Lieb conditions necessary to ensure the existence of the limiting integral, the following convergence holds as $T \to \infty$:

$$
T^{-d \, co(M)} J_T(M, f_1, ..., f_E) \to \int_{S^C} f_1(\lambda_1)f_2(\lambda_2)...f_E(\lambda_E) \prod_{c=1}^{C} \mu(\mathrm{d}y_c), \tag{4.1}
$$

where $(\lambda_1, ... \lambda_E) = (y_1, ..., y_C)M^*$ (with every λ_e reduced modulo $[-\pi, \pi]^d$ in the torus case), M^* being any matrix whose rows span the space orthogonal to the rows of M, and C being the rank of M^*. Informally, the kernels disappear in the limit, giving rise to the "dual matroid" M^*.

(3) When the Hölder-Young-Brascamp-Lieb conditions do not hold, then, *cf.* part c) of the theorem below, the normalizing by the factor $T^{d\,\alpha_M(z)}$ appearing in part a) will lead to a zero limit.

Theorem 4.1. *Suppose that $f_e \in L_{p_e}(d\mu)$ for part a), and $f_e \in \mathbf{L}_{p_e}(d\mu)$ for parts b),c) and set $z = (p_1^{-1}, ..., p_E^{-1})$.*

Let $J_T = J_T(M, f_1, ..., f_E)$ denote a Fejér matroid integral and let $r(A), r^*(A)$ denote respectively the ranks of a set of columns in M and in the dual matroid M^*.

Suppose that for every row l of the matrix M, one has $r(M) = r(M_l)$, where M_l is the matrix with the row l removed. Then:

a)

$$J_T(M, f_1, ..., f_E) \leq c_M T^{d\,\alpha_M(z)} \prod_{e=1}^{E} \|f_e\|_{p_e} \tag{4.2}$$

where c_M *is some constant independent of z (see Rem. 3) below) and*

$$\begin{cases} \text{in the discrete case} & \alpha_M(z) \text{ is given by } (4.6) \\ \text{in the continuous case} & \alpha_M(z) \text{ is given by } (4.10). \end{cases}$$

b) *If $\alpha_M(z) = V - r(M) = co(M)$ or, equivalently, if*

$$\text{in the discrete case} \qquad \sum_{j \in A} z_j \leq r^*(A), \ \forall A \tag{4.3}$$

$$\text{in the continuous case, in addition} \quad \sum_{e \in \mathcal{G}_k} z_e = C_k, k = 1, ..., co(M)$$

where $\mathcal{G}_k, C_k$ *denote respectively the k'th component of the graph and its number of cycles, then*

$$\lim_{T \to \infty} \frac{J_T(M)}{T^{d\,co(M)}} = \mathcal{J}(M^*, f_1, ..., f_E), \tag{4.4}$$

where

$$\mathcal{J}(M^*, f_1, ..., f_E) = \int_{S^C} f_1(\lambda_1) f_2(\lambda_2) ... f_E(\lambda_E) \prod_{c=1}^{C} \mu(dy_c) \tag{4.5}$$

where $(\lambda_1, ... \lambda_E) = (y_1, ..., y_C)M^*$ (with every λ_e reduced modulo $[-\pi, \pi]^d$ in the discrete case), where C denotes the rank of the dual matroid M^*.

c) *If a strict inequality $\alpha_M(z) > co(M)$ holds, then the inequality (4.2) of Theorem 4.1 a) may be strengthened to:*

$$J_T(M) = o(T^{d\,\alpha_M(z)}).$$

Remarks.

1) Note that for connected graphs the exponent in part b) reduces to d, and that for cycle graphs the power counting conditions $\sum_{j \in A} z_j \leq r^*(A)$ reduce to $\sum_{j=1}^{n} z_j \leq 1$, as in Theorem 2.2.
2) The exponent bounds $\alpha_M(z)$ can never be tight when $\alpha_M(z) > co(M)$. Even the first-order asymptotics of the Fejér graph integrals is unknown in this case.
3) The constant c_M is the product of the constant appearing in the Hölder-Young inequality (C.1) with the constant $C_{\frac{V}{V-1}}^{V}$ defined in the kernel estimate (D.2) in Appendix D, due to the monotonicity of C_p in p (see, Appendix D) and to the inequality $s_v \leq 1 - 1/V$, by Lemma 2 of [11].

4) The result follows trivially from the case $co(M) = 1$. This formulation is however worthwhile as an intermediate step towards the general nonunimodular matroid case (where an extra constant will appear in the limit (4.4) of part b)).

5) Note that the results of this theorem: the expression of $\alpha_M(z)$, the limit integral $\mathcal{J}(M^*) = \mathcal{J}(M^*, f_1, ..., f_E)$, and the convergence conditions of the integrals depend on M, M^* only *via* the two equivalent rank functions $r(A), r^*(A)$, *i.e.* only *via* the matroid dependence structure between the columns, and not on the chosen representing matrices.

4.1. **The upper bound for the order of magnitude of Fejér matroid integrals**

We turn now to the "upper bound exponent" $\alpha_M(z)$ for the order of magnitude of Fejér matroid integrals (useful when it is not precisely d). In the discrete case [11], the exponent $\alpha_M(z)$ of this upper bound turns out to be d times the solution of a graph optimization problem:

$$\alpha_M(z) = co(M) + \max_{A \subset 1, ..., E} \left[\sum_{j \in A} z_j - r^*(A) \right] \tag{4.6}$$

or equivalently,

$$\alpha_M(z) = \max_{A \subset 1, ..., E} f(A), \quad f(A) := \left[co(M - A) - \sum_{j \in A} (1 - z_j) \right] \tag{4.7}$$

where $co(M - A)$ represents in the graph case the number of remaining components, after the edges in A have been removed, and, for a general Fejér matroid integral, we define

$$co(M - A) := V - r(M - A). \tag{4.8}$$

We will call the problem (4.7) a graph breaking problem: find a set of edges whose removal maximizes the difference between the number of remaining components and $\sum_{j \in A} (1 - z_j)$, or, equivalently, the difference between $\sum_{j \in A} z_j$ and the dual rank $r^*(A)$.

Remark 4.2. In examples involving connected graphs, the solution of the optimal breaking problem is often obtained at the "total breaking" (TB) set $A = M$ with $f(M) = V + \sum_e (z_e - 1)$ or at the "no breaking" (NB) set $A = \emptyset$ with $f(\emptyset) = co(M)$ (geometrically, these are the most important facets of the power counting polytope (PCP)). It is useful to introduce therefore the lower bound:

$$\begin{aligned}
\alpha_M^c(z) &= \max_{A \in \{M, \emptyset\}} \left[co(M - A) - \sum_{j \in A} (1 - z_j) \right] = \max \left\{ co(M), \sum_e (z_e - 1) + V \right\} \\
&= \max \left\{ co(M), co(M) + \sum_e z_e - C \right\} = 1 + \left(\sum_e z_e - C \right)_+ ,
\end{aligned} \tag{4.9}$$

where the equality before the last holds by Euler's relation $C = E - V + co(M)$ and where C denotes the number of cycles in a graph, and, more generally, the rank of the dual matroid M^*.

Note that in the case of a cycle graph of size m, this reduces to

$$\alpha_m^c(z) = \max \left\{ 1, \sum_{e=1}^m z_e \right\} = 1 + \left(\sum_{e=1}^m z_e - 1 \right)_+$$

as stated in Theorem 2.2.

In the case of graphs with several components, this gives rise to the additive extension

$$\alpha_M^c(z) = co(M) + \sum_{j=1}^{co(M)} \left(\sum_{e \in \mathcal{G}_j} z_e - C_j \right)_+ , \tag{4.10}$$

where $\mathcal{G}_j$, C_j denote the set of columns and the number of cycles belonging to the j'th component.

4.2. Computation of $\alpha_M(z)$ and applications

As observed in Corollary 2 of [11], and as in the particular case of Theorem 2.2, Theorem 4.1 may be used for establishing asymptotic normality, whenever the cumulants may be written as sums of Fejér graph integrals.

Corollary 4.3. *Let $X_T, T \to \infty$, be a sequence of centered random variables with cumulants of all orders $\chi_{k,T}$, for which*

$$\chi_{k,T} = \sum_{G \in \mathcal{G}_k} F_T(G),$$

where $F_T(G)$ are d-dimensional Fejér graph integrals, and the summation runs over some families of connected graphs $\mathcal{G}_k$. Then, if the "exponent inequalities"

$$\alpha_G \leq k/2, \quad \forall G \in \mathcal{G}_k$$

are satisfied for all Fejér graph integrals intervening in the expansion of the cumulant of order k, then the central limit theorem

$$\frac{X_T}{T^{d/2}} \to N(0, \sigma^2), \quad \text{where } \sigma^2 = \sum_{G \in \mathcal{G}_k} \mathcal{J}(G),$$

holds.

Proof. Since the CLT normalization here is $T^{d/2}$, it is enough to notice that for all the Fejér graph integrals appearing in the expansion of the kth cumulant it holds that

$$F_T(G) = O(T^{d\alpha_G(z)}) = O(T^{kd/2}),$$

and

$$F_T(G) = o(T^{dk/2}),$$

since if the second "O" is not an "o", then $\alpha_G > 1$ and the first "O" must be an "o", by Theorem 4.1 c).   □

Remark 4.4. Note since the functions $\alpha_G(z)$ are nondecreasing in z, it is enough to establish the exponent inequalities on the nondominated part of the boundary of the PCP polytope.

In conclusion, establishing normality is reduced to computing the functions $\alpha_G(z)$ *i.e.*, to solving a sequence of graph breaking problems.

We present now one more result for Gaussian fields, which is related in the discrete case to the classical result of Breuer and Major [24], and in the continuous case to the result of Ivanov and Leonenko [49]. We note that these authors worked under time-domain assumptions, however, reasonings in the spectral domain with the methodology of [11] and the present paper immediately lead to the following result.

Theorem 4.5. *Let X_t be a d-parameter zero-mean Gaussian random field with spectral density $f(\lambda) \in \mathbf{L}_p$. Let $S_T = \int_{t \in I_T} P_m(X_t) \nu(\mathrm{d}t)$, where $P_m(X_t)$ are univariate Appell (orthonormal Hermite) polynomials and $m \geq 2$. Assume that:*

$$z := p^{-1} \leq 1 - \frac{1}{m} \tag{4.11}$$

with equality holding in the continuous case.

Then,

$$
\begin{aligned}
\sigma^2 \quad &:= \quad (2\pi)^d f^{(*,m)}(0) \\
&= \quad (2\pi)^d \int_{y_1,\ldots,y_{m-1}\in S} f(y_1)f(y_2-y_1)...f(y_{m-1}-y_{m-2})f(-y_{m-1}) \prod_{i=1}^{m-1} \mu(\mathrm{d}y_i) < \infty. \quad (4.12)
\end{aligned}
$$

If, moreover, $\sigma \neq 0$, then the central limit theorem holds

$$
T^{-d/2}S_T \to N(0,\sigma^2), \quad T \to \infty.
$$

Remark 4.6. The discrete and continuous case may be unified here by asking for integrability at $z = 1 - \frac{1}{m}$, since in the discrete case the extension to smaller values of z is immediate. In fact, it is always enough in the discrete case to prove the result for the undominated boundary of the PCP polytope, which reduces here to one point.

Proof. For an integer $m \geq 1$, let $r^m(\cdot)$ denote the mth power of $r(\cdot)$ and let

$$
\begin{aligned}
f^{(*,m)}(\lambda) \quad &= \quad \int_{\lambda_1,\ldots,\lambda_{m-1}\in S} f(\lambda_1)f(\lambda_2)...f(\lambda_{m-1})f(\lambda-\lambda_1-...-\lambda_{m-1}) \prod_{i=1}^{m-1} \mu(\mathrm{d}\lambda_i) \\
&= \quad \int_{y_1,\ldots,y_{m-1}\in S} f(y_1)f(y_2-y_1)...f(y_{m-1}-y_{m-2})f(\lambda-y_{m-1}) \prod_{i=1}^{l-1} \mu(\mathrm{d}y_i)
\end{aligned}
$$

be the m-tuple convolution of f with itself, which is well defined if $f(\lambda) \in L_p$, $z = \frac{1}{p} \leq 1 - \frac{1}{m}$, (with equality in the continuous case), by the Hölder-Young inequality, since (see Appendix C). Indeed, the dependence matrix M with m columns has rank $r(M) = m-1$, and hence the necessary condition $z_1 + \ldots + z_m \leq m-1$, with $z_1 = \ldots = z_m = z$ requires $zm \leq m-1$, or $z \leq \frac{m-1}{m}$.

The idea of the proof (appearing already in [27] and [49], p. 54), consists in showing that a) for the orthonormal Hermite polynomials $H^{(m)}(x) = P^{(m)}(x)$, it holds that

$$
\begin{aligned}
T^{-d}\mathrm{var}\int_{I_T} H^{(m)}(X_t)\nu(\mathrm{d}t) \quad &= \quad T^{-d}\int_{I_T}\int_{I_T} r^m(t-s)\nu(\mathrm{d}t)\nu(\mathrm{d}s) \quad (4.13) \\
&= \quad T^{-d}\int_{I_T}\int_{I_T}\int_S \mathrm{e}^{\mathrm{i}\lambda(t-s)} f^{(*,m)}(\lambda) \, \mathrm{d}\lambda \, \nu(\mathrm{d}t)\nu(\mathrm{d}s) \\
&= \quad \int_S T^{-d}\Delta_T(\lambda)\Delta_T(-\lambda)f^{(*,m)}(\lambda) \, \mu(\mathrm{d}\lambda) \to (2\pi)^d f^{(*,m)}(0), \quad T \to \infty,
\end{aligned}
$$

where the last limit follows from the kernel property, and that b) under the same hypothesis $z \leq \frac{m-1}{m}$, similar computations imply that the normalized higher order cumulants converge to 0.

We will review now these computations in more detail. By the multilinearity of cumulants (see Appendix B.2)

$$
\begin{aligned}
\chi_k(T^{-d/2}S_T) \quad &= \quad T^{-kd/2}\mathrm{cum}_k\left(\int_{I_T} H^{(m)}(X_{t_1})\nu(\mathrm{d}t_1),\ldots,\int_{I_T} H^{(m)}(X_{t_k})\nu(\mathrm{d}t_k)\right) \quad (4.14) \\
&= \quad T^{-kd/2}\int_{I_T}\cdots\int_{I_T} \mathrm{cum}_k\left(H^{(m)}(X_{t_1}),\ldots,H^{(m)}(X_{t_k})\right) \nu(\mathrm{d}t_1)...\nu(\mathrm{d}t_k).
\end{aligned}
$$

By the diagram formula (see Appendix B)

$$
\mathrm{cum}_k\left(H^{(m)}(X_{t_1}),\ldots,H^{(m)}(X_{t_k})\right) = \sum_{P\in\mathcal{P}} r(t_{u_1}-t_{v_1})\ldots r(t_{u_{mk/2}}-t_{v_{mk/2}}),
$$

where $P = (t_{u_1}, t_{v_1}), \ldots, (t_{u_{mk/2}}, t_{v_{mk/2}})$ and $\mathcal{P}$ is the set of partitions in pairs of the table

$$
\begin{array}{cccc}
t_1 & \ldots & t_1 & (m \text{ times}) \\
\vdots & \ddots & \vdots & \\
t_k & \ldots & t_k & (m \text{ times})
\end{array}
$$

such that each pair connects two distinct rows, and no subset of pairs has as its union a subset of rows strictly included in $\{1, \ldots, k\}$.

We note now that the set of partitions $\mathcal{P}$ intervening in the k'th cumulant of sums S_T (4.14) (of the m'th Appell polynomial of a Gaussian sequence) can be put in one-to-one correspondence with the set of connected unoriented graphs with no loops $\Gamma(m, k)$ with k vertices of degree m (one for each index t_s, $s = 1, \ldots, k$), by representing each partition pair (t_u, t_v) as an edge between the vertices associated to t_u, t_v.

In conclusion, the cumulants (4.14) may be represented as sums of Fejér graph integrals:

$$
\chi_k(T^{-d/2} S_T) = T^{-kd/2} \sum_{\Gamma \in \Gamma(m,k)} J_T(\Gamma).
$$

To conclude that the central limit theorem $S_T / T^{d/2} \to N(0, \sigma^2)$, $T \to \infty$, holds, it remains only to show that the associated "base functions" (see (3.3)) intervening in the Fejér integrals satisfy integrability conditions which imply the exponent inequalities $\alpha_\Gamma(z) \leq \frac{k}{2}$, for all $\Gamma \in \Gamma(m, k)$, for $z = z_c := 1 - 1/m$ (*cf.* Rem. 4.4).

Note first that at z_c, the lower bound implied by the total breaking of the k'th cumulant graph is $k/2$:

$$
\alpha_k(z_c) = V - \sum_e (1 - z_c) = k - \frac{km}{2}(1 - z_c) = \frac{k}{2}.
$$

Let now

$$
p(A) = co(\mathcal{E} - A) - \sum_{e \in A}(1 - z_e)
$$

denote the "profit" of a set of edges A.

We will show now that the unique "maximal with respect to inclusion optimal breaking" (MOB) A is the total breaking.

Indeed, suppose that after applying a MOB, one vertex is still connected to another. Cutting now all edges around this vertex, we obtain a new breaking, whose profit has increased by 1, at a cost of at most m/m, contradicting thus that we had started with a MOB.

In conclusion, the cumulant exponents $\alpha_k(z_c) = \frac{k}{2}$, ensuring convergence of the variance for $k = 2$ and negligibility of the cumulants for $k \geq 3$ (since then $\alpha_k > 1$). $\qquad \square$

Remark 4.7. A similar analysis holds in the case of cumulants of quadratic forms in Appell polynomials $P_{m,n}(X_t, X_s)$. Note that while the number of graphs intervening increases considerably, the number of extremal points of the PCP is just 4 – see Figure 2.

The graphs intervening are – see Figure 1 – all the graphs belonging to the set $\Gamma(m, n, k)$ of all connected bipartite graphs with no loops whose vertex set consists of k pairs of vertices. The "left" vertex of each pair arises out of the first m terms : $X_{t_1}, \ldots, X_{t_m}$: in the diagram formula, and the "right" vertex of each pair arises out of the last n terms : $X_{s_1}, \ldots, X_{s_n}$: the edge set consists of:

(1) k "kernel edges" pairing each left vertex with a right vertex. The kernel edges will contribute below terms involving the function $b(\lambda)$.

(2) A set of "correlation edges", always connecting vertices in different rows, and contributing below terms involving the function $f(\lambda)$). They are arranged such that each left vertex connects to m and each right vertex connects to n such edges, yielding a total of $k(m + n)/2$ correlation edges.
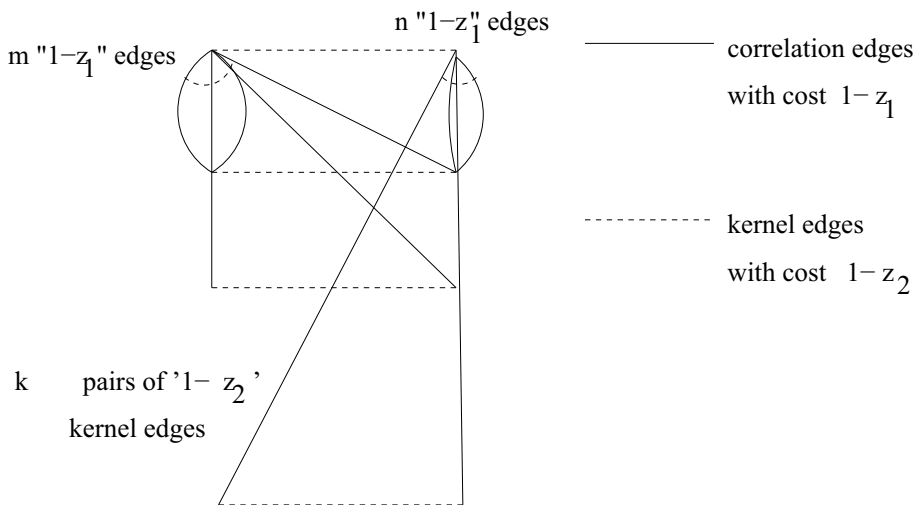
FIGURE 1. The graphs appearing in the expansion of cumulants of quadratic forms. Here $k = 4$, $m = 5$, $n = 4$. The figure displays only some of the $k(m + n)/2 = 18$ correlation edges.
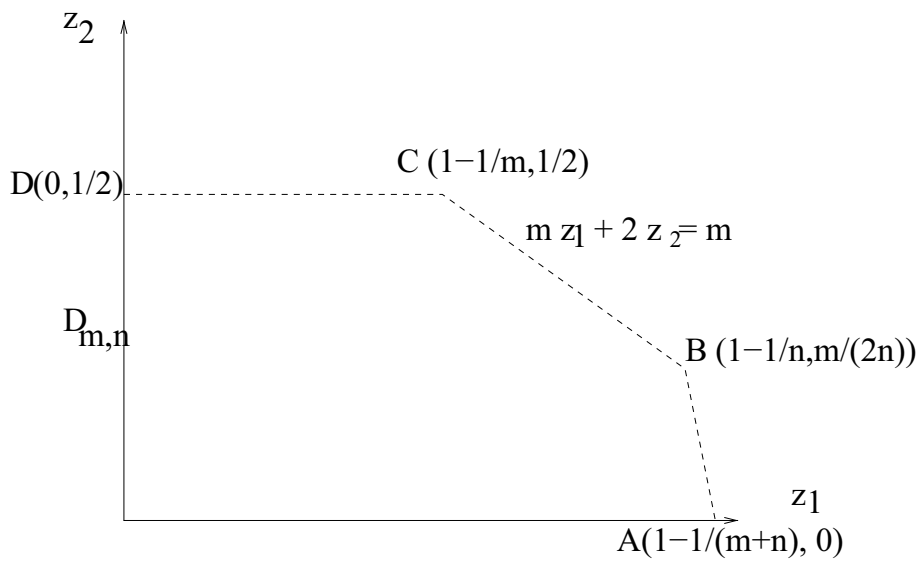

FIGURE 2. The domain of the central limit theorem, discrete case.

Thus, the k "left vertices" are of degree $m + 1$, and the other k vertices are of degree $n + 1$. (The "costs $1 - z_1, 1 - z_2$" mentioned in Fig. 1 refer to (4.6)).

The PCP domain in the discrete case (which is precisely the convergence domain of the integrals defining the limiting variance), is indicated below, when $m < n$, in terms of the integrability indices $z = (z_1, z_2)$ of f and b (i.e. $f \in L_{p_1}, b \in L_{p_2}$).

When $m < n$, there are only three segments on the undominated boundary of the PCP, connecting respectively the extremal points (A, B), (B, C) and (C, D)(with coordinates $A(1 - 1/(m+n), 0)$, $B(1 - 1/n, m/(2n))$,

$C(1 - 1/m, 1/2)$, $D(0, 1/2)$), and correspond respectively to the breakings indicated below:

$$
\begin{cases}
\text{total breaking} & 2k - \frac{k(m+n)}{2}(1-z_1) - k(1-z_2) \leq \frac{k}{2} \\
& \Leftrightarrow \frac{3}{2} \leq \frac{(m+n)}{2}(1-z_1) + (1-z_2) \\
\text{breaking all } z_2 \text{ and the left } z_1 \text{ edges} & \max_k \frac{k+1}{k} - m(1-z_1) - (1-z_2) \leq \frac{1}{2} \\
& \Leftrightarrow 1 \leq m(1-z_1) + (1-z_2) \\
\text{breaking all } z_2 \text{ edges} & \max_k \frac{2}{k} - (1-z_2) = \leq \frac{1}{2} \\
& \Leftrightarrow \frac{1}{2} \leq z_2.
\end{cases}
$$

In the continuous case, the domain is just the lower segment between the points A and B in Figure 2.

4.3. **Proof of Theorem 4.1**

We turn now to the proof of Theorem 4.1, following the proofs of Theorem 2 and Corollary 1 of [11]. The first step is to assume w.l.o.g. (by multiplicativity of the R.H.S) that $co(M) = 1$.

For part a), let us apply the Hölder-Young-Brascamp-Lieb inequality with optimally chosen integrability parameters s_v^{-1}:

$$|J_T(M, f_1, ..., f_E)| \leq K \prod_{v=1}^{V} \|\Delta_T\|_{s_v^{-1}} \prod_{e=1}^{E} \|f_e\|_{z_e^{-1}} \leq c_M T^{d \sum_{v=1}^{V}(1-s_v^{-1})} \prod_{e=1}^{E} \|f_e\|_{z_e^{-1}} \tag{4.15}$$

under the constraint that $(s_1, ..., s_V, z_1, ..., z_E)$ satisfy the power counting conditions, and where we used the kernel estimate

$$\|\Delta_T(\lambda)\|_{s^{-1}} \leq C_s T^{d(1-s)}, \quad \forall s \in [0, 1)$$

(see Appendix D), the monotonicity of C_s in s and the fact that the optimal s_v may be all chosen so that $s_v \leq 1 - 1/V$, by Lemma 2 of [11].

The optimization problem for s_v in the discrete case:

$$\min_{s_1, ..., s_V} d \sum_{v=1}^{V}(1 - s_v^{-1}), \text{ where } (s_1, ..., s_V, z_1, ..., z_E) \in \text{ PCP},$$

has the same constraints as Lemma 2 in [11], except that the objective is multiplied by d. Hence, in the torus case, the exponent is simply d times the one dimensional exponent of Theorem 1, [11].

In the continuous case $S = \mathbb{R}^d$, we note first that when $\sum_e z_e \leq C$ the result holds by applying the Hölder-Young inequality with $s_v = 1 - 1/V$, while in the other case $\sum_e z_e \geq C$, the extra constraint $\sum_v s_v = E - \sum_e z_e$ forces the one-dimensional exponent to be $\alpha_M(z) = co(M) + \sum_e z_e - C$.

The proof of part b) of Theorem 4.1 is essentially identical with that given in [12] (see also [11] for the matroid case), up to the modification of the integrability conditions. For completeness, we sketch now this proof, for a connected graph.

Note first that in a connected graph there are only $V - 1$ independent rows of the incidence matrix M (or independent variables u_j), since the sum of all the rows is 0 (equivalently, $u_V = -\sum_{v=1}^{V-1} u_v$). Thus, $r(M) = V - 1$, $co(M) = 1$, and the order of magnitude appearing in the normalization is just T^d.

The main ideas behind the proof of Theorem 4.1 b) are a change of variables and applying the continuity of graph convolutions:

(1) **Change of variables.** Fix a basis $y_1, ..., y_C$ in the complement of the space generated by the u_v's, $v = 1, ..., V$, switch to the variables $u_1, ..., u_{V-1}, y_1, ..., y_C$ and integrate in (3.4) first over the variables y_c's, $c = 1, ..., C$. This is more convenient in the graphic case, since, after fixing an arbitrary spanning tree $\mathcal{T}$ in the graph, the complementary set of edges $\mathcal{T}^c$ furnishes a maximal set of independent cycles

(with cardinality C). Assume w.l.o.g. that in the list $(\lambda_1, ..., \lambda_E)$, the edges in $\mathcal{T}^c$ are listed first, namely $(\lambda_e, e \in \mathcal{T}^c) = (\lambda_1, ..., \lambda_C)$. We make the change of variables $y_1 = \lambda_1, ..., y_C = \lambda_C$, and $(u_1, ..., u_{V-1})' = \tilde{M}(\lambda_1, ..., \lambda_E)'$, where $\tilde{M}$ denotes the first $V-1$ rows of the incidence matrix M. Thus,

$$(y_1, ..., y_C, u_1, ..., u_{V-1})' = \begin{pmatrix} I_C & 0 \\ \tilde{M}_C & \tilde{M}_V \end{pmatrix} (\lambda_1, ..., \lambda_E)',$$

where the first rows are given by an identity matrix I_C completed by zeroes and where $\tilde{M}_C, \tilde{M}_V$ denote the first C columns/ next $V-1-C$ columns of the matrix $\tilde{M}$, and where the unimodularity of M ensures that $\det(\tilde{M}_V) = 1$ and therefore the Jacobian of the transformation is 1.

Inverting the transformation above yields:

$$(\lambda_1, ..., \lambda_E) = (y_1, ..., y_C, u_1, ..., u_{V-1}) \begin{pmatrix} I_C & -\tilde{M}_C' \tilde{M}_V^{-1} \\ 0 & \tilde{M}_V^{-1} \end{pmatrix} = (y_1, ..., y_C, u_1, ..., u_{V-1}) \begin{pmatrix} M^* \\ N \end{pmatrix}, \qquad (4.16)$$

that is, it turns out that the first rows of the inverse matrix are precisely the dual matroid M^*.

Definition 4.8. The function

$$h_{M^*,N}(u_1, ..., u_{r(M)}) = \int_{y_1, ..., y_C \in S} f_1(\lambda_1) f_2(\lambda_2) ... f_E(\lambda_E) \prod_{c=1}^{C} d\mu(y_c), \qquad (4.17)$$

where λ_e are represented as linear combinations of $y_1, ..., y_C, u_1, ..., u_{V-1}$ *via* the linear transformation (4.16) will be called a matroid/graph convolution depending on whether the matroid is graphic or not.

Note that the function $h_{M^*,N}$ is well defined as soon as we suppose integrability conditions (4.3) to hold due to the Hölder-Brascamp-Lieb-Barthe inequality (see Appendix C and Lem. 4.9 below).

Thus, the change to the variables $y_1, ..., y_C, u_1, ..., u_{V-1}$ and integration over $y_1, ..., y_C$ transforms the Fejér graph integral into the integral of the product of a "graph convolution" and a Fejér kernel:

$$\begin{aligned} J_T(M) &= \int_{u_1, ..., u_{V-1} \in S} h_{M^*,N}(u_1, ..., u_{V-1}) \prod_{v=1}^{V} \Delta_T(u_v) \prod_{v=1}^{V-1} d\mu(u_v) \\ &= (2\pi)^{(n-1)d} \int_{u_1, ..., u_{V-1} \in S} h_{M^*,N}(u_1, ..., u_{V-1}) \Phi_T^*(u_1, ..., u_{V-1}) \prod_{v=1}^{V-1} d\mu(u_v). \end{aligned}$$

(2) The continuity of the graph convolutions $h(u_1, ..., u_{r(M)})$ in the variables $(u_1, ..., u_{r(M)})$ continues to hold here, with a proof essentially identical with that of Lemma 2.1 from the cycle case.

Lemma 4.9. The continuity of graph convolutions. *The "graph convolution" function defined in (4.17) is bounded and continuous if (3.3) holds with integrability indices satisfying the power counting condition (4.3).*

The proof of the above lemma is essentially the same as in the cycle case (see Lem. 2.1).

In conclusion, the convergence of the Fejér kernels to Lebesgue measure on the set $u_1 = ... = u_{V-1} = 0$ (Lem. 1.1) implies, just as in the cycle case, the convergence of the scaled Fejér graph integral

$$J_T(M, f_e, e = 1..., E) \quad \text{to} \quad \mathcal{J}(M^*, f_e, e = 1..., E),$$

establishing Part b) of the theorem.

For the proof of part c), we approximate our functions by continuous functions of bounded support, for which conditions (3.3) of part b) hold. For these approximants, it follows from the convergence $\frac{J_T(M)}{T^{d\,co(M)}} \to k_M \, \mathcal{J}(M^*, f_1, ..., f_E)$ that $J_T(M) = o(T^{d\,co(M)+a})$, $\forall a > 0$. The result follows then for the functions from $\mathbf{L}_p$ spaces, by the definition of these spaces. Note that in the discrete case, the same argument was applied based on trigonometric polynomials – see [11], proof of Corollary 1.

Remark 4.10. In the spatial statistics papers [7,8], the continuity of the graph convolutions $h_{M^*,N}(u_1, ..., u_{V-1})$ was assumed to hold, and indeed checking whether this assumption may be relaxed was one of the outstanding difficulties for the spatial extension.

Remark 4.11. It is immediate to extend this approach to the case of graphs with several components. One would choose in each component independent cycle and vertex variables $y_1, ..., y_{r(M^*)}$ and $u_1, ..., u_{r(M)}$, note the block structure of the matrices, with each block corresponding to a graph component, use the fact that for graphs with several components, the rank of the graphic matroid is $r(M) = V - co(G)$ and finally Euler's relation $E - V = C - co(G)$, which ensures that

$$E = \big(V - co(G)\big) + C = r(M) + r(M^*).$$

For further extension to the matroid setup, one would need to add to part b) the Jacobian k_M of the linear transformation, as sketched in Remark 4.15 below.

4.4. **An open problem: refined asymptotics for Fejér matroid integrals**

With the purpose of suggesting how to refine the asymptotics in the future, let us consider now the simplest discrete case, with $d = 1, I_1 = [0, 1), co(M) = 1$ and with f_e being complex exponentials.

In this case, we have also the bonus of an immediate proof of the limit relation (4.4) in Lemma 4.12 below, which implies immediately Theorem 4.1(b), by the multilinearity of $\mathcal{J}(M^*, f_1, ..., f_E)$ and $T^{-1}J(M, f_1, ..., f_E)$.

Lemma 4.12. *For any matrix with $r(M) = V - 1$, any set of integers $b = (b_e, e = 1, ..., E)$, and any functions $f_e(\lambda_e) = \mathrm{e}^{-\mathrm{i}\lambda_e b_e}, e = 1, ..., E$, Theorem 4.1(b) holds, i.e.:*

$$\lim_{T \to \infty} \frac{\int_{S^E} \mathrm{e}^{-\mathrm{i}\langle b, \lambda\rangle} \prod_{v=1}^V \Delta_T(u_v) \prod_{e=1}^E \mu(\mathrm{d}\lambda_e)}{T} = \int_{S^C} \mathrm{e}^{-\mathrm{i}\langle \lambda, b\rangle} \prod_{c=1}^C \mu(\mathrm{d}y_c),$$

where $(\lambda_1, ...\lambda_E) = (y_1, ..., y_C)M^*$.

Remark 4.13. This lemma is related to the kernel Lemma 1.1; indeed, since converge of the Fourier coefficients implies weak convergence, its statement is equivalent to the weak converge as $T \to \infty$ of the measures on $(S)^E$ given by the multiple Fejér kernels

$$T^{-1}\Delta_T \left(-\sum_{v=1}^{V-1} \lambda_v \right) \prod_{v=1}^{V-1} (\Delta_T(\lambda_v)\nu(\mathrm{d}\lambda_\nu))$$

to uniform measure on the subspace $\lambda = yM^*$.

Proof. It turns out that for complex exponential integrands, the stated integral appearing in the RHS is an indicator:

$$\int_{S^C} e^{-i\langle\lambda,b\rangle} \prod_{c=1}^{C} \mu(dy_c) = \int_{S^C} e^{-i\langle y, M^*b\rangle} \prod_{c=1}^{C} \mu(dy_c) = 1_{\{M^*b=0\}} = 1_{\{b\in R(M)\}},$$

where $R(M)$ denotes the subspace generated by the rows of M.

Now the numerator of the LHS in (4.4), using the symmetric definition (3.5) for FMI's, is:

$$
\begin{aligned}
\int_{S^E} e^{-i\langle b,\lambda\rangle} \int_{I_T^V} e^{i\langle sM,\lambda\rangle} \prod_{v=1}^{V} \nu(ds_v) \prod_{e=1}^{E} \mu(d\lambda_e) &= \int_{I_T^V} \prod_{v=1}^{V} \nu(ds_v) \left(\int_{S^E} e^{i\langle sM-b,\lambda\rangle} \prod_{e=1}^{E} \mu(d\lambda_e) \right) \\
&= \int_{I_T^V} \prod_{v=1}^{V} \nu(ds_v) 1_{\{sM-b=0\}} \\
&= \nu(s : sM = b, s \in I_T) = \# \left(s : sM = b, s \in I_T, s \in \mathbb{N}^V \right).
\end{aligned}
$$

Example 4.14. a) If $b = 0$, and M is the incidence matrix of a graph, an elementary computation yields the function of interest $\# \left(s : sM = b, s \in I_T, s \in \mathbb{N}^V \right) = \#(s : s = x(1,...,1), x \in [0,T), x \in \mathbb{N}) = T$.

b) If M is the incidence matrix of a cycle graph, let $s_k = \sum_{i=1}^{k} b_i, k = 1, ..., E-1, s_E = 0$. An elementary computation yields the function of interest $\# \left(s : sM = b, s \in I_T, s \in \mathbb{N}^V \right) = \#(s : s = x(1,...,1), x \in \mathbb{N}, x + \max[s_k] < T, x + \min[s_k] \geq 0) = T - \max[s_k] + \min[s_k]$.

Returning to our lemma, we see that as $T \to \infty$, the effect of b is negligible:

$$\# \left(s : sM = b, s \in TI_1, s \in \mathbb{N}^V \right) = \# \left(x : xM = b/T, x \in I_1, x \in T^{-1}\mathbb{N}^V \right) \sim \# \left(x : xM = 0, x \in I_1, x \in T^{-1}\mathbb{N}^V \right).$$

Thus, we may approximate our function by its particular value obtained when all the components of b are zero:

$$E(T) = \# \left(x : xM = 0, x \in I_1, x \in T^{-1}\mathbb{N}^V \right) = \# \left(s : sM = 0, s \in TI_1, s \in \mathbb{N}^V \right). \qquad \square$$

Remark 4.15. Note that a lot is known about the function

$$\phi_M(b,T) := \# \left(s : sM = b, s \in TI_1, s \in \mathbb{N}^V \right)$$

for general integral matrices M and after including the translation by b. This is an "Ehrhart polynomial" in T [65,67], which suggests the possibility of developing correction terms to Theorem 4.1 b).

The first modification is the appearance of an extra constant k_M in the RHS of the limit result (4.4) of Theorem 4.1, as indicated in [11]. Indeed, letting $K, \mathcal{K}$ denote the affine lattice defined by the set of all integral points of the subspace $\ker(M) := \{\sigma : sM = 0\}$ (of dimension $co(M)$) and its fundamental parallelepiped, it follows that:

$$\# \left(s : sM = b, s \in TI_1, s \in \mathbb{N}^V \right) \begin{cases} = 0 & \text{if } b \notin R(M), \\ \sim T^{co(M)} \tilde{\mu}(I_1 \cap \ker(M)) & \text{if } b \in R(M), \end{cases}$$

where $\tilde{\mu}$ denotes "normalized Lebesgue measure" on the affine lattice K (normalized so that the fundamental parallelepiped of this lattice has unit being the volume). Thus, the limit result will involve the extra constant

$$k_M = \tilde{\mu}(I_1 \cap \ker(M)) = \frac{\mu(I_1 \cap \ker(M))}{\mu(\mathcal{K})}. \tag{4.18}$$

Note that $\mu(\mathcal{K})$ is the determinant of an arbitrary basis of $\mathcal{K}$, which explains why it does not depend on the change of variables of the previous section.

5. APPLICATIONS

5.1. **Central limit theorems for bilinear forms of moving averages**

We assume below that our stationary random field X_t, $t \in I$, admits a representation as a linear/moving averages random field.

For discrete parameter it means that

$$X_t = \sum_{u \in \mathbb{Z}^d} \hat{a}(t - u)\xi(u), \ \sum_{u \in \mathbb{Z}^d} \hat{a}^2(u) < \infty, \ t \in \mathbb{Z}^d, \tag{5.1}$$

where $\xi(u)$, $u \in \mathbb{Z}^d$, are independent random variables indexed by $\mathbb{Z}^d$ with $\mathbb{E}\xi(0) = 0$ and such that $E\,|\xi(0)|^k \leq c_k < \infty$, $k = 1, 2, \ldots$. In this case

$$c_k(t_1, ..., t_k) = \mathrm{cum}_k\{X_{t_1}, ..., X_{t_k}\} = d_k \sum_{s \in \mathbb{Z}^d} \prod_{j=1}^{k} \hat{a}(t_j - s), \tag{5.2}$$

where d_k is the k'th cumulant of $\xi(0)$.

For continuous parameter we assume that

$$X_t = \int_{u \in \mathbb{R}^d} \hat{a}(t - u)\xi(\mathrm{d}u), \quad t \in \mathbb{R}^d, \tag{5.3}$$

with a square-integrable kernel $\hat{a}(t)$, $t \in I$, with respect to a independently scattered random measure with finite second moment, that is a homogeneous random measure $\xi(A)$, $A \subset \mathbb{R}^d$, with finite second moments and independent values over disjoint sets (see, for instance, Rajput and Rosinski [59] or Kwapien and Woyczynski [51]). That is, for each Borel A, $\xi(A)$ is an infinitely divisible random variable whose cumulant function can be written as

$$\kappa(z) = \log E \mathrm{e}^{\mathrm{i}z\xi(A)} = \mathrm{i}z m_0(A) - \frac{1}{2} z^2 m_1(A) + \int_{\mathbb{R}} \left(\mathrm{e}^{\mathrm{i}zx} - 1 - \mathrm{i}z\tau(x) \right) Q(A, \mathrm{d}x), \tag{5.4}$$

where m_0 is a signed measure, m_1 is a positive measure, $Q(A, \mathrm{d}x)$ (for fixed A) is a measure on $\mathbb{R}^1$ without atoms at 0, such that $\int_{\mathbb{R}} \min\left\{1, |x|^2\right\} Q(A, \mathrm{d}x) < \infty$, and where $\tau(x) = x$ if $|x| \leq 1$, and $\tau(x) = x/|x|$, if $|x| > 1$.

For example, if $I = \mathbb{R}$, then $\xi(A)$ is a set indexed Lévy process with finite second moments and stationary intensity proportional to the Lebesgue measure.

We also assume that Q factorizes as $Q(A, \mathrm{d}x) = M(A)W(\mathrm{d}x)$, where $M(A)$ is a σ-finite measure, and W is some Lévy measure on $\mathbb{R}^1$, such that for some $\varepsilon > 0$ and $\lambda > 0$

$$\int_{(-\varepsilon, \varepsilon)} \mathrm{e}^{\lambda u} W(\mathrm{d}u) < \infty.$$

This implies that

$$\int_{\mathbb{R}} |u|^k \, W(\mathrm{d}u) < \infty, \ k \geq 2,$$

and that the cumulant function $\kappa(z)$ is analytical in a neighborhood of 0.

Necessary and sufficient conditions of existence of the integral (as limit in probability of integrals of simple functions)

$$\int_A f(s)\mathrm{d}\xi(s)$$

can be found in [59]. Note that, for $d = 1$, an integrals become integrals with respect to Lévy process $L(t)$, $t \in \mathbb{R}^1$, and $\kappa(z) = \log E\mathrm{e}^{\mathrm{i}zL(1)}$.

For Lebesgue measures m_0, m_1 and Q, one can prove (by using product integration) that

$$\log E \exp\{\mathrm{i}(z_1 X_{t_1} + \ldots + z_k X_{t_k}\} = \int_{\mathbb{R}^d} \kappa\left(\sum_{j=1}^{k} z_j \hat{a}(t_j - s)\right) \mathrm{d}s \qquad (5.5)$$

if $\hat{a} \in L_1 \cap L_2$. From (5.5) if can be seen that random field (5.3) is homogeneous in a strict sense.

We assume that $m_1 = 0$, that is, $\mathbb{E}\xi(I_1) = 0$, then the last formula holds for $\hat{a} \in L_2$.

We obtain that

$$c_k(t_1, ..., t_k) = \mathrm{cum}_k\{X_{t_1}, ..., X_{t_k}\} = d_k \int_{\mathbb{R}^d} \prod_{j=1}^{k} \hat{a}(t_j - s)\mathrm{d}s, \qquad (5.6)$$

where d_k is the kth cumulant of $\xi(I_1)$ with I_1 being the unit rectangle, that is $d_k = \kappa^k(0)/i^k$, $k \geq 2$.

We assume from now on that $\mathbb{E}\xi(I_1) = 0$, and use the same notation for both discrete and continuous cases

$$X_t = \int_{u \in I} \hat{a}(t - u)\xi(\mathrm{d}u), \quad t \in I, \qquad (5.7)$$

where $I = \mathbb{Z}^d$ in the discrete case and $I = \mathbb{R}^d$ in the continuous case.

For various conditions which ensure that (5.7) is well-defined, see, for example, Anh *et al.* [5], p. 733, and references therein.

Note. By choosing an appropriate "Green function" $\hat{a}(t)$, this very general class of processes includes the solutions of many interesting differential equations with random noise $\xi(\mathrm{d}u)$, like, for example, generalized Ornstein-Uhlenbeck processes in $\mathbb{R}$ [5].

We will assume here that all moments for our stationary field X_t exist.

The advantage of the linear representation assumption (5.7) and (5.4), (5.5) is the explicit representation of cumulants – see for example Theorem 2.1 of [5]:

$$c_k(t_1, ..., t_k) = d_k \int_{s \in I} \prod_{j=1}^{k} \hat{a}(t_j - s)\,\nu(\mathrm{d}s), \qquad (5.8)$$

where d_k is the kth cumulant of $\xi(I_1)$ with I_1 being the unit rectangle in the continuous case and d_k is the kth cumulant of $\xi(0)$ in the discrete case, that is, in particular,

$$d_2 = \mathbb{E}\xi(I_1)^2, \quad d_4 = \mathbb{E}(\xi(I_1)^4) - 2[\mathbb{E}(\xi(I_1)^2)]^2 \qquad (5.9)$$

or

$$d_2 = \mathbb{E}\xi(0)^2, \quad d_4 = \mathbb{E}(\xi(0)^4) - 2[\mathbb{E}(\xi(0)^2)]^2 \qquad (5.10)$$

for the continuous and discrete cases respectively.

In the spectral domain, we get

$$f_k(\lambda_1, ..., \lambda_{k-1}) = d_k \, a\left(-\sum_{i=1}^{k-1} \lambda_i\right) \prod_{i=1}^{k-1} a(\lambda_i) = d_k \, \prod_{i=1}^{k} a(\lambda_i)\delta\left(\sum_{j=1}^{k} \lambda_j\right). \tag{5.11}$$

For $k = 2$, we will denote the spectral density by $f(\lambda) = f_2(\lambda) = d_2 a(\lambda)a(-\lambda)$.

We can formulate now a central limit theorems for quadratic functional of a linear field, which is a generalization of the results of Giraitis and Surgailis [36] and Giraitis and Taqqu [38] (see also references therein) for discrete time processes. This next theorem follows from the results of Section 3 (the proof is almost identical to the proof of Theorem 4 of Avram [11], the expression for the variance can be obtained by direct computations).

Theorem 5.1. *Let* $X_t = \int_{u \in \, I} \hat{a}(t - u)\xi(\mathrm{d}u), \quad t \in \, I$, *be a linear random field with a square integrable kernel* $\hat{a}(t), t \in I$, *and a random measure* $\xi(\mathrm{d}u)$ *admitting all moments and let* $Q_T = Q_T^{(1,1)}$ *is defined by (2.1).*

We assume that $f(\lambda) = (2\pi)^d |a(\lambda)|^2 \in \mathbf{L}_p$ *and* $b(\lambda) \in \mathbf{L}_q$, *and in the continuous case we assume also that* $b(\lambda) \in \mathbf{L}_q \cap L_1$.

Assume that:

$$\frac{1}{p} + \frac{1}{q} \le \frac{1}{2}$$

with equality holding in continuous case.

Then, the central limit theorem holds:

$$T^{-d/2}Q_T \to N(0, \sigma^2), \quad T \to \infty,$$

where

$$\sigma^2 := 2(2\pi)^d d_2^2 \int_S b^2(\lambda)f^2(\lambda)\mathrm{d}\lambda + (2\pi)^d d_4 \left(\int_S b(\lambda)f(\lambda)\mathrm{d}\lambda\right)^2$$

and d_2 *and* d_4 *are defined by (5.9)–(5.10).*

5.2. **Minimum contrast estimation based on the Whittle contrast function**

In what follows we will consider continuous time linear processes ($d = 1$) whose spectral densities of all orders exist and admit the representation of the form (5.11).

The class of Whittle estimators is the most popular in applications (see Whittle [74], [75], Giraitis and Surgailis [37], Fox and Taqqu [28], Heyde and Gay [43,44], Heyde [42], Gao *et al.* [31], Leonenko and Sakhno [52], see also the references therein).

We begin with the following assumption.

A.I.. Let X_t, $t \in I_T = \left[-\frac{T}{2}, \frac{T}{2}\right]$, be an observation of a real-valued measurable stationary linear process X_t, $t \in \mathbb{R}^1$, with zero mean and the family of spectral densities (5.11). Let $a(\lambda) = a\left(\lambda; \theta^{(1)}\right), d_k = d_k\left(\theta^{(2)}\right)$, that is, $f_2(\lambda) = f(\lambda, \theta)$, $\lambda \in \mathbb{R}^1$, $\theta = \left(\theta^{(1)}, \theta^{(2)}\right)$, $\theta \in \Theta \subset \mathbb{R}^m$, where Θ is a compact set, and the true value of the parameter $\theta_0 \in int\Theta$, the interior of Θ. Suppose further that $f(\lambda; \theta_1) \not\equiv f(\lambda; \theta_2)$ for $\theta_1 \ne \theta_2$, almost everywhere in $\mathbb{R}^1$ with respect to the Lebesgue measure.

Consider the Whittle contrast process (or objective function)

$$U_T(\theta) = \frac{1}{4\pi} \int_{\mathbb{R}^1} \left(\log f(\lambda; \theta) + \frac{I_T(\lambda)}{f(\lambda; \theta)}\right) w(\lambda) \, \mathrm{d}\lambda, \tag{5.12}$$

where $I_T(\lambda)$ is the periodogram of the second order

$$I_T(\lambda) = \frac{1}{2\pi T} \left|\int_{I_T} X_t \mathrm{e}^{-\mathrm{i}t\lambda} dt\right|^2, \quad \lambda \in \mathbb{R}^1, \tag{5.13}$$

and $w(\lambda)$ is a symmetric about $\lambda = 0$ function such that all considered integrals are well defined and which will satisfy some conditions given below; in some cases we can choose $w(\lambda) = \frac{1}{1+\lambda^2}$.

Introduce the Whittle contrast function

$$K(\theta_0; \theta) = \frac{1}{4\pi} \int_{\mathbb{R}^1} \left(\frac{f(\lambda; \theta_0)}{f(\lambda; \theta)} - 1 - \log \frac{f(\lambda; \theta_0)}{f(\lambda; \theta)} \right) w(\lambda) \, d\lambda. \tag{5.14}$$

To state the result on consistency of the minimum contrast estimator based on the contrast process (5.12) we will need the following conditions on the spectral density $f(\lambda; \theta)$ and the weight function $w(\lambda)$.

A.II.. $f(\lambda; \theta_0) w(\lambda) \frac{1}{f(\lambda; \theta)} \in L_1(\mathbb{R}^1) \cap L_2(\mathbb{R}^1), \quad \forall \theta \in \Theta.$

A.III.. There exists a function $v(\lambda), \lambda \in \mathbb{R}^1$, such that

(i) the function $h(\lambda; \theta) = v(\lambda) \frac{1}{f(\lambda; \theta)}$ is uniformly continuous in $\mathbb{R}^1 \times \Theta$;

(ii) $f(\lambda; \theta_0) \frac{w(\lambda)}{v(\lambda)} \in L_1(\mathbb{R}^1) \cap L_2(\mathbb{R}^1).$

Theorem 5.2. *Let the assumptions A.I to A.III be satisfied. Then the function $K(\theta_0; \theta)$ defined by (5.14) is the contrast function for the contrast process $U_T(\theta)$ defined by (5.12). The minimum contrast estimator $\widehat{\theta}_T$ defined as*

$$\widehat{\theta}_T = \underset{\theta \in \Theta}{\arg \min} U_T(\theta) \tag{5.15}$$

is a consistent estimator of the parameter θ, that is, $\widehat{\theta}_T \to \theta_0$ in P_0-probability as $T \to \infty$.

The above theorem can be obtained as a consequence of a more general result by Leonenko and Sakhno [52] (Thm. 3), one needs just to rewrite for the case of linear processes the corresponding conditions on spectral densities, which become of much simpler form.

Next set of assumptions (in addition to the above ones) is needed to state the result on asymptotic normality of the estimator (5.15).

A.IV. The function $\frac{1}{f(\lambda; \theta)}$ is twice differentiable in a neighborhood of the point θ_0 and

(i) $f(\lambda; \theta_0) w(\lambda) \frac{\partial^2}{\partial \theta_i \partial \theta_j} \frac{1}{f(\lambda; \theta)} \in L_1(\mathbb{R}^1) \cap L_2(\mathbb{R}^1), i, j = 1, ..., m, \theta \in \Theta$;

(ii) $f(\lambda; \theta_0) \in \mathbf{L}_p(\mathbb{R}^1), w(\lambda) \frac{\partial}{\partial \theta_i} \frac{1}{f(\lambda; \theta)} \in \mathbf{L}_q(\mathbb{R}^1) \cap L_1(\mathbb{R}^1), i = 1, ..., m, \theta \in \Theta$,
for some p, q such that $\frac{1}{p} + \frac{1}{q} \le \frac{1}{2}$, with equality holding in continuous case;

(iii) $T^{1/2} \int_{\mathbb{R}^1} (EI_T(\lambda) - f(\lambda; \theta_0)) w(\lambda) \frac{\partial}{\partial \theta_i} \frac{1}{f(\lambda; \theta)} \, d\lambda \to 0$ as $T \to \infty$,
for all $i = 1, ..., m, \theta \in \Theta$;

(iv) the second order derivatives $\frac{\partial^2}{\partial \theta_i \partial \theta_j} \frac{1}{f(\lambda; \theta)}, i = 1, ..., m$, are continuous in θ.

A.V. The matrices $W_1(\theta) = \left(w_{ij}^{(1)}(\theta) \right)_{i,j=1,...,m}, W_2(\theta) = \left(w_{ij}^{(2)}(\theta) \right)_{i,j=1,...,m}, V(\theta) = (v_{ij}(\theta))_{i,j=1,...,m}$ are positive definite, where

$$w_{ij}^{(1)}(\theta) = \frac{1}{4\pi} \int_{\mathbb{R}^1} w(\lambda) \frac{\partial}{\partial \theta_i} \log f(\lambda; \theta) \frac{\partial}{\partial \theta_j} \log f(\lambda; \theta) \, d\lambda, \tag{5.16}$$

$$w_{ij}^{(2)}(\theta) = \frac{1}{4\pi} \int_{\mathbb{R}^1} w^2(\lambda) \frac{\partial}{\partial \theta_i} \log f(\lambda; \theta) \frac{\partial}{\partial \theta_j} \log f(\lambda; \theta) \, d\lambda, \tag{5.17}$$

$$v_{ij}(\theta) = \frac{1}{8\pi} \frac{d_4}{d_2^2} \int_{\mathbb{R}^1} w(\lambda) \frac{\partial}{\partial \theta_i} \log f(\lambda; \theta) \, d\lambda \int_{\mathbb{R}^1} w(\lambda) \frac{\partial}{\partial \theta_j} \log f(\lambda; \theta) \, d\lambda. \tag{5.18}$$

Theorem 5.3. *Let the assumptions A.I to A.V be satisfied. Then as $T \to \infty$*

$$T^{1/2} \left(\widehat{\theta_T} - \theta_0 \right) \to \mathcal{N}_m \left(0, W_1^{-1} (\theta_0) (W_2 (\theta_0) + V (\theta_0)) W_1^{-1} (\theta_0) \right),$$

where $N_m (\cdot, \cdot)$ denotes the m-dimensional Gaussian law.

Reasonings for the proof of Theorem 5.3 will be given in the next section.

Comparing the above theorem with a more general result stated in [52], one can see that the set of conditions for the case of linear processes becomes of much simpler form, but the most important improvement is in condition A.IV(ii), which was achieved due to the application of the Theorem 5.1 (see the proof). Note that corresponding condition for the case of general processes, formulated in [52], unfortunately, is difficult to check in general situation.

Remark 5.4. Condition A.IV(iii) will hold, *e.g.*, if $f (\lambda; \theta)$ is differentiable with respect to λ and

$$\int_{\mathbb{R}^1} f'_\lambda (\lambda; \theta_0)) w (\lambda) \frac{\partial}{\partial \theta_i} \frac{1}{f (\lambda; \theta)} \mathrm{d}\lambda < \infty,$$

or under any conditions which assure

$$\int_{\mathbb{R}^1} |f (\lambda + h; \theta_0) - f (\lambda; \theta_0)| w (\lambda) \frac{\partial}{\partial \theta_i} \frac{1}{f (\lambda; \theta)} \mathrm{d}\lambda \le C|h|^a,$$

for $a > \frac{1}{2}$ and C being a constant.

Example. Estimation of fractional Riesz-Bessel motion (FRBM) (see Appendix E for details and definition of FRBM in non-Gaussian case). Let X_t, $t \in \mathbb{R}^1$, be a non-Gaussian Riesz-Bessel stationary motion, that is a stationary linear process with the spectral density of the form

$$f (\lambda) = f (\lambda, \theta) = \frac{c}{|\lambda|^{2\alpha} (1 + \lambda^2)^\gamma}, \quad \lambda \in \mathbb{R}^1, \tag{5.19}$$

where the unknown vector parameter $\theta = (\gamma, \alpha, c)' \in \Theta$, Θ being a compact subset of $\left[\frac{1}{2}, \infty \right) \times \left(0, \frac{1}{2} \right) \times (0, \infty)$. Note that the index α determines the long-range dependence of FRBM, and the parameter γ is another fractal index connected to Hausdorff dimension of paths of the stochastic process. Note that procedure of discretization leads to the loss of information of one parameter γ, which is important for applications in both turbulence and finance theory. That is why a direct method of estimation of both parameters from continuous data looks appropriate.

For this model we can choose the weight function $w (\lambda) = \frac{1}{1+\lambda^2}$, $\lambda \in \mathbb{R}^1$, to satisfy the conditions needed for consistency of the estimator (5.15), that is, for Theorem 5.2 to hold. However, to satisfy all the conditions needed for Theorem 5.3 we choose the weight function $w (\lambda) = \frac{\lambda^{2b}}{(1+\lambda^2)^a}$, $\lambda \in \mathbb{R}^1$, where a and b satisfy the restrictions: $\{b > 1\} \wedge \{a > b + 2\} \wedge \{a > A + 2\}$, where we have denoted by A the length of the finite interval carrying the admissible values of the parameter γ. With such a choice of the weight function we have the convergence

$$T^{1/2} \left(\widehat{\theta_T} - \theta_0 \right) \to N_3 \left(0, W_1^{-1} (\theta_0) (W_2 (\theta_0) + V (\theta_0)) W_1^{-1} (\theta_0) \right) \quad \text{as} \ \ T \to \infty,$$

where the elements of the matrices W_1 and W_2 are of the following form:

$$w_{11}^{(1\vee 2)} = \frac{1}{4\pi}\int_{\mathbb{R}^1} w^{1\vee 2}(\lambda)\left(\ln\left(1+\lambda^2\right)\right)^2 d\lambda;$$

$$w_{22}^{(1\vee 2)} = \frac{1}{4\pi}\int_{\mathbb{R}^1} w^{1\vee 2}(\lambda)\left(\ln\left(\lambda^2\right)\right)^2 d\lambda;$$

$$w_{33}^{(1\vee 2)} = \frac{1}{4\pi}c_0^{-2}\int_{\mathbb{R}^1} w^{1\vee 2}(\lambda) d\lambda;$$

$$w_{12}^{(1\vee 2)} = w_{21}^{(1\vee 2)} = \frac{1}{4\pi}\int_{\mathbb{R}^1} w^{1\vee 2}(\lambda)\ln\left(1+\lambda^2\right)\ln\left(\lambda^2\right) d\lambda;$$

$$w_{13}^{(1\vee 2)} = w_{31}^{(1\vee 2)} = -\frac{1}{4\pi}c_0^{-1}\int_{\mathbb{R}^1} w^{1\vee 2}(\lambda)\ln\left(1+\lambda^2\right) d\lambda;$$

$$w_{23}^{(1\vee 2)} = w_{32}^{(1\vee 2)} = -\frac{1}{4\pi}c_0^{-1}\int_{\mathbb{R}^1} w^{1\vee 2}(\lambda)\ln\left(\lambda^2\right) d\lambda;$$

$$v_{11} = \frac{1}{8\pi}\frac{d_4}{d_2^2}\left(\int_{\mathbb{R}^1} w(\lambda)\ln\left(1+\lambda^2\right) d\lambda\right)^2;$$

$$v_{22} = \frac{1}{8\pi}\frac{d_4}{d_2^2}\left(\int_{\mathbb{R}^1} w(\lambda)\ln\left(\lambda^2\right) d\lambda\right)^2;$$

$$v_{33} = \frac{1}{8\pi}\frac{d_4}{d_2^2}c_0^{-2}\left(\int_{\mathbb{R}^1} w(\lambda) d\lambda\right)^2;$$

$$v_{12} = v_{21} = \frac{1}{8\pi}\frac{d_4}{d_2^2}\int_{\mathbb{R}^1} w(\lambda)\ln\left(1+\lambda^2\right) d\lambda \int_{\mathbb{R}^1} w(\lambda)\ln\left(\lambda^2\right) d\lambda;$$

$$v_{13} = v_{31} = -\frac{1}{8\pi}\frac{d_4}{d_2^2}c_0^{-1}\int_{\mathbb{R}^1} w(\lambda)\ln\left(1+\lambda^2\right) d\lambda \int_{\mathbb{R}^1} w(\lambda) d\lambda;$$

$$v_{23} = v_{32} = -\frac{1}{8\pi}\frac{d_4}{d_2^2}c_0^{-1}\int_{\mathbb{R}^1} w(\lambda)\ln\left(\lambda^2\right) d\lambda \int_{\mathbb{R}^1} w(\lambda) d\lambda.$$

In the above formulae we mean that the weight function $w(\lambda)$ is involved to the expressions for $w_{ij}^{(1)}$ in the 1st power and to the expressions for $w_{ij}^{(2)}$ in the 2nd power. From the above formulae we see that the covariance matrix of the limiting normal law has the charming feature that it appears not depending on the values α_0 and γ_0.

Remark 5.5. Continuous version of Gauss-Whittle objective function with the weight function $w(\lambda) = \frac{1}{1+\lambda^2}$ had been used in [31] for the estimation of the Gaussian processes in stationary and nonstationary cases respectively.

5.3. Minimum contrast estimation based on the Ibragimov contrast function

We consider now the minimum contrast functional motivated by the paper of Ibragimov [48], see also Anh et al. [8].

We assume condition **A.I** and introduce the following condition
B.I. There exists a nonnegative function $w(\lambda), \lambda \in \mathbb{R}$, such that

(i) $w(\lambda)$ is symmetric about $\lambda = 0: w(\lambda) = w(-\lambda)$;
(ii) $w(\lambda) f(\lambda;\theta)$ is in $L_1(\mathbb{R})$ for $\forall\theta \in \Theta$.

Under the condition B.I, we set

$$\sigma^2(\theta) = \int_{\mathbb{R}} f(\lambda;\theta) w(\lambda) d\lambda$$

and consider the factorization of the spectral density

$$f(\lambda; \theta) = \sigma^2(\theta) \psi(\lambda; \theta), \quad \lambda \in \mathbb{R}, \theta \in \Theta.$$

For the function $\psi(\lambda, \theta), \lambda \in \mathbb{R}, \theta \in \Theta$, we have

$$\int_{\mathbb{R}} \psi(\lambda; \theta) w(\lambda) \, d\lambda = 1$$

and we additionally suppose
B.II. The derivatives $\nabla_\theta \psi(\lambda; \theta)$ exist and

$$\nabla_\theta \int_{\mathbb{R}} \psi(\lambda; \theta) w(\lambda) \, d\lambda = \int_{\mathbb{R}} \nabla_\theta \psi(\lambda; \theta) w(\lambda) \, d\lambda = 0,$$

that is we can differentiate under the integral sign in the above integral.

Consider the following contrast process (or objective function):

$$U_T(\theta) = -\int_{\mathbb{R}} I_T(\lambda) w(\lambda) \log \psi(\lambda; \theta) \, d\lambda, \quad \theta \in \Theta. \tag{5.20}$$

Define also the function

$$K(\theta_0; \theta) = \int_{\mathbb{R}} f(\lambda; \theta_0) w(\lambda) \log \frac{\psi(\lambda; \theta_0)}{\psi(\lambda; \theta)} d\lambda, \quad \theta_0, \theta \in \Theta. \tag{5.21}$$

B.III. $f(\lambda; \theta_0) w(\lambda) \log \psi(\lambda; \theta) \in L_1(\mathbb{R}^1) \cap L_2(\mathbb{R}^1), \quad \forall \theta \in \Theta.$
B.IV. There exists a function $v(\lambda), \lambda \in \mathbb{R}^1$, such that

   (i) the function $h(\lambda; \theta) = v(\lambda) \log \psi(\lambda; \theta)$ is uniformly continuous in $\mathbb{R}^1 \times \Theta$;
   (ii) $f(\lambda; \theta_0) \frac{w(\lambda)}{v(\lambda)} \in L_1(\mathbb{R}^1) \cap L_2(\mathbb{R}^1).$

Theorem 5.6. *Let conditions A.I, B.I–B.IV be satisfied. Then the function $K(\theta_0; \theta)$ defined by (5.21) is the contrast function for the contrast process $U_T(\theta)$ defined by (5.20). Moreover the minimum contrast estimator $\widehat{\theta}_T$ defined as*

$$\widehat{\theta}_T = \arg\min_{\theta \in \Theta} U_T(\theta), \tag{5.22}$$

is a consistent estimator of the parameter θ, that is, $\widehat{\theta}_T \to \theta_0$ in P_0-probability as $T \to \infty$, and the estimator

$$\widehat{\sigma}_T^2 = \int_{\mathbb{R}^n} I_T(\lambda) w(\lambda) \, d\lambda$$

is a consistent estimator of the parameter $\sigma^2(\theta)$, that is, $\widehat{\sigma}_T^2 \to \sigma^2(\theta_0)$ in P_0-probability as $T \to \infty$.

To formulate the result on the asymptotic distribution of the minimum contrast estimator (5.22) we need some further conditions.
B.V. The function $\psi(\lambda; \theta)$ is twice differentiable in a neighborhood of the point θ_0 and

   (i) $f(\lambda; \theta) w(\lambda) \frac{\partial^2}{\partial\theta_i \partial\theta_j} \log \psi(\lambda, \theta) \in L_1(\mathbb{R}) \cap L_2(\mathbb{R}), i, j = 1, ..., m, \theta \in \Theta$;
   (ii) $f(\lambda; \theta_0) \in \mathbf{L}_p(\mathbb{R}^1), w(\lambda) \frac{\partial}{\partial\theta_i} \log \psi(\lambda, \theta) \in \mathbf{L}_q(\mathbb{R}^1) \cap L_1(\mathbb{R}^1), i = 1, ..., m, \theta \in \Theta$,
      for some p, q such that $\frac{1}{p} + \frac{1}{q} \le \frac{1}{2}$, with equality holding in continuous case;
   (iii) $T^{1/2} \int_{\mathbb{R}^1} (EI_T(\lambda) - f(\lambda; \theta_0)) w(\lambda) \frac{\partial}{\partial\theta_i} \log \psi(\lambda; \theta) \, d\lambda \to 0$ as $T \to \infty$,
      for all $i = 1, ..., m, \theta \in \Theta$;
   (iv) the second order derivatives $\frac{\partial^2}{\partial\theta_i \partial\theta_j} \log \psi(\lambda; \theta), i = 1, ..., m$, are continuous in θ.

B.VI. The matrices $S(\theta) = (s_{ij}(\theta))_{i,j=1,\ldots,m}$ and $A(\theta) = (a_{ij}(\theta))_{i,j=1,\ldots,m}$ are positive definite where

$$
\begin{aligned}
s_{ij}(\theta) &= \int_{\mathbb{R}} f(\lambda;\theta)\, w(\lambda)\, \frac{\partial^2}{\partial\theta_i\partial\theta_j} \log\psi(\lambda;\theta)\, \mathrm{d}\lambda \\
&= \sigma^2(\theta) \int_{\mathbb{R}} w(\lambda) \left[\frac{\partial^2}{\partial\theta_i\partial\theta_j}\psi(\lambda,\theta) - \frac{1}{\psi(\lambda,\theta)} \frac{\partial}{\partial\theta_i}\psi(\lambda,\theta) \frac{\partial}{\partial\theta_j}\psi(\lambda,\theta) \right] \mathrm{d}\lambda, \\
a_{ij}(\theta) &= 4\pi \int_{\mathbb{R}^1} f^2(\lambda;\theta)\, w^2(\lambda)\, \frac{\partial}{\partial\theta_i} \log\psi(\lambda;\theta) \frac{\partial}{\partial\theta_j} \log\psi(\lambda;\theta)\, \mathrm{d}\lambda \\
&\quad + 2\pi\frac{\mathrm{d}_4}{\mathrm{d}_2^2} \int_{\mathbb{R}} \frac{w(\lambda)\, f(\lambda;\theta)}{\psi(\lambda;\theta)} \frac{\partial}{\partial\theta_i}\psi(\lambda;\theta)\, \mathrm{d}\lambda \int_{\mathbb{R}} \frac{w(\lambda)\, f(\lambda;\theta)}{\psi(\lambda;\theta)} \frac{\partial}{\partial\theta_j}\psi(\lambda;\theta)\, \mathrm{d}\lambda \\
&= 4\pi\left(\sigma^2(\theta)\right)^2 \int_{\mathbb{R}} w^2(\lambda)\, \frac{\partial}{\partial\theta_i}\psi(\lambda;\theta) \frac{\partial}{\partial\theta_j}\psi(\lambda;\theta)\, \mathrm{d}\lambda \\
&\quad + 2\pi\frac{\mathrm{d}_4}{\mathrm{d}_2^2}\left(\sigma^2(\theta)\right)^2 \int_{\mathbb{R}} w(\lambda)\, \frac{\partial}{\partial\theta_i}\psi(\lambda;\theta)\, \mathrm{d}\lambda \int_{\mathbb{R}} w(\lambda)\, \frac{\partial}{\partial\theta_j}\psi(\lambda;\theta)\, \mathrm{d}\lambda.
\end{aligned}
$$

Theorem 5.7. *Let the conditions A.I, B.I–B.VI be satisfied. Then as $T \longrightarrow \infty$*

$$
T^{1/2}\left(\widehat{\theta}_T - \theta_0\right) \to N_m\left(0, S^{-1}(\theta_0)\, A(\theta_0)\, S^{-1}(\theta_0)\right),
$$

where $N_m(\cdot,\cdot)$ denotes the m-dimensional Gaussian law.

Proofs of Theorems 5.2, 5.3, 5.6, 5.7. The results on consistency of estimators (Thms. 5.2 and 5.6) are consequences of corresponding theorems stated for the general case in [52] for the Whittle functional and in [8,9] for the case of Ibragimov functional. We present here reasonings for the proofs of Theorems 5.3 and 5.7, which make use of CLT for bilinear forms (Thm. 5.1 above). For the proofs the standard arguments based on Taylor's formula for $\nabla_\theta U_T\left(\widehat{\theta}_T\right)$ are used. Namely, we can write the relation

$$
\nabla_\theta U_T\left(\widehat{\theta}_T\right) = \nabla_\theta U_T(\theta_0) + \nabla_\theta\nabla'_\theta U_T(\theta_T^*)\left(\widehat{\theta}_T - \theta_0\right),
$$

where $|\theta_T^* - \theta_0| < \left|\widehat{\theta}_T - \theta_0\right|$.

It follows from the definition of minimum contrast estimators that for sufficiently large T

$$
\nabla_\theta U_T(\theta_0) = -\nabla_\theta\nabla'_\theta U_T(\theta_T^*)\left(\widehat{\theta}_T - \theta_0\right),
$$

therefore, to state the asymptotic normality for the estimator $\widehat{\theta}_T$, by Slutsky's arguments, one needs to deduce: (1) limit in probability for $\nabla_\theta\nabla'_\theta U_T(\theta_T^*)$ and (2) limiting normal law for $T^{1/2}\nabla_\theta U_T(\theta_0)$.

For the 1st task we can use the same arguments as in the mentioned above papers, and to rewrite (simplify) corresponding conditions for the case of linear processes.

However, for the step (2) we can appeal now to Theorem 5.1. We provide the details below.

Consider firstly the case of Whittle functional. Limit in P_0-probability for $\nabla_\theta\nabla'_\theta U_T(\theta_T^*)$ is given by the matrix $W_1(\theta_0)$.

Next, consider

$$
\nabla_\theta U_T(\theta_0) = \frac{1}{4\pi} \int_{\mathbb{R}} \left(\nabla_\theta \log f(\lambda;\theta)|_{\theta=\theta_0} + \nabla_\theta\left(\frac{1}{f(\lambda;\theta)}\right)\Big|_{\theta=\theta_0} I_T(\lambda) \right) w(\lambda)\, \mathrm{d}\lambda,
$$

which can be written in the form

$$
\begin{aligned}
\nabla_\theta U_T\left(\theta_0\right) &= \left(J_T(\varphi_i) - J(\varphi_i)\right)_{i=1,\ldots,m} \\
&= \left(\int_{\mathbb{R}} \varphi_i\left(\lambda;\theta_0\right) I_T\left(\lambda\right) \mathrm{d}\lambda - \int_{\mathbb{R}} \varphi_i\left(\lambda;\theta_0\right) f\left(\lambda;\theta_0\right) \mathrm{d}\lambda\right)_{i=1,\ldots,m}
\end{aligned}
$$

where

$$
\varphi_i = \varphi_i\left(\lambda;\theta_0\right) = -\frac{1}{4\pi} \frac{1}{f^2\left(\lambda;\theta_0\right)} w\left(\lambda\right) \left.\left(\frac{\partial}{\partial\theta_i} f\left(\lambda;\theta\right)\right)\right|_{\theta=\theta_0}, \quad i=1,\ldots,m.
$$

Under the assumptions of Theorem 5.3 (see A.IV(ii)) in view of Theorem 5.1 we have the convergence

$$
T^{1/2}\left(J_T(\varphi_i) - EJ_T(\varphi_i)\right)_{i=1,\ldots,m} \to N_m\left(0, W_2\left(\theta_0\right) + V\left(\theta_0\right)\right), \tag{5.23}
$$

where the the matrices $W_2\left(\theta_0\right)$ and $V\left(\theta_0\right)$ are defined in the assumption A.V.

Further, in view of the assumption A.IV(iii)

$$
T^{1/2}\left(EJ_T(\varphi_i) - J(\varphi_i)\right) \to 0, \text{ as } T \to \infty
$$

which, combined with (5.23), implies

$$
T^{1/2}\left(J_T(\varphi_i) - J(\varphi_i)\right)_{i=1,\ldots,m} = T^{1/2}\nabla_\theta U_T\left(\theta_0\right) \to N_m\left(0, W_2\left(\theta_0\right) + V\left(\theta_0\right)\right).
$$

The case of Ibragimov functional is treated analogously. We have that $\nabla_\theta \nabla'_\theta U_T\left(\theta_T^*\right)$ converges in P_0-probability to the matrix $S\left(\theta_0\right)$. Further,

$$
\nabla_\theta U_T\left(\theta_0\right) = -\int_{\mathbb{R}} I_T\left(\lambda\right) \left.\nabla_\theta \log \psi\left(\lambda;\theta\right)\right|_{\theta=\theta_0} w\left(\lambda\right) \mathrm{d}\lambda.
$$

In view of B.II

$$
\int_{\mathbb{R}} f\left(\lambda;\theta_0\right) \left.\nabla_\theta \log \psi\left(\lambda;\theta\right)\right|_{\theta=\theta_0} w\left(\lambda\right) \mathrm{d}\lambda = 0,
$$

and we can write

$$
\begin{aligned}
\nabla_\theta U_T\left(\theta_0\right) &= \left(J_T(\varphi_i) - J(\varphi_i)\right)_{i=1,\ldots,m} \\
&= \left(\int_{\mathbb{R}} \varphi_i\left(\lambda;\theta_0\right) I_T\left(\lambda\right) \mathrm{d}\lambda - \int_{\mathbb{R}} \varphi_i\left(\lambda;\theta_0\right) f\left(\lambda;\theta_0\right) \mathrm{d}\lambda\right)_{i=1,\ldots,m},
\end{aligned}
$$

where now

$$
\varphi_i = \varphi_i\left(\lambda;\theta_0\right) = w\left(\lambda\right) \left.\frac{\partial}{\partial\theta_i} \log \psi\left(\lambda;\theta\right)\right|_{\theta=\theta_0}, \quad i=1,\ldots,m.
$$

Again in view of Theorem 5.1, under the assumption B.V(ii), we obtain the convergence

$$
T^{1/2}\left(J(\varphi_i) - EJ_T(\varphi_i)\right) \to N_m\left(0, A\left(\theta_0\right)\right), \text{ as } T \to \infty,
$$

where the matrix $A\left(\theta_0\right)$ is defined in B.VI. By assumption B.V(iii) the convergence $T^{1/2}\nabla_\theta U_T(\theta_0) \to N_m(0, A(\theta_0))$ follows. $\qquad\square$

APPENDIX A. WICK PRODUCTS AND APPELL POLYNOMIALS

Let W be a finite set and Y_i, $i \in W$ be a system of random variables. Let

$$Y^W = \prod_{i \in W} Y_i$$

denote the ordinary product, with $Y^\emptyset = 1$, let $m^W = E \prod_{i \in W} Y_i$ be the (mixed) moment, and let

$$\chi(Y^W) = \chi(Y_i, i \in W)$$

denote the (mixed) cumulant of the variables $Y_i, i \in W$, defined recursively as the solutions of the equations:

$$m^W = \sum_{\{V\}|-W} \chi(Y^{V_1}) \ldots \chi(Y^{V_r}), \tag{A.1}$$

where the sum $\sum_{\{V\}|-W}$ is over all partitions $\{V\} = (V_1, \ldots, V_r)$, $r \geq 1$ of the set W, and where $\chi(Y^\emptyset) = 1$.

Notes. 1) Equation (A.1) is the formal power series expression of the "exponential relation $m = e^\chi$" between moments and cumulants, viewed as functions on the lattice of subsets [55].

2) The inverse of equation (A.1), the formal power series expression of the "logarithmic relation $\chi = \log(m)$" may also be computed by:

$$\chi(Y_1, \ldots, Y_n) = \frac{\partial^T}{\partial z_1 \ldots \partial z_n} \log E \exp \left(\sum_{i=1}^T z_j Y_j \right) \Big|_{z_1 = \ldots = z_n = 0},$$

where the differentiation is interpreted formally if the moment generating function does not exist.

Definition A.1. The *Wick products* $: Y^W :$ are defined as the solutions of the recursion:

$$Y^W = \sum_{U \subset W} : Y^U : E(Y^{W \setminus U}) = \sum_{U \subset W} : Y^U : \sum_{\{V\}|-W \setminus U} \chi(Y^{V_1}) \ldots \chi(Y^{V_r}),$$

where the sum $\sum_{U \subset W}$ is taken over all subsets $U \subset W$, including $U = \emptyset$, the sum $\sum_{\{V\}|-W \setminus U}$ is over all partitions $\{V\} = (V_1, \ldots, V_r)$, $r \geq 1$ of the set $W \setminus U$, and the starting value is $: Y^\emptyset := 1$.

Notes. 1) Inverting the recursion yields [68] (Prop. 1):

$$: Y^W := \sum_{U \subset W} Y^U \sum_{\{V\}|-W \setminus U} (-1)^r \chi(Y^{V_1}) \ldots \chi(Y^{V_r}),$$

as may be formally seen by replacing m^{-1} by $e^{-\chi}$.

2) When some variables appear repeatedly, it is convenient to use the notation

$$: \underbrace{Y_{t_1}, \ldots, Y_{t_1}}_{n_1}, \ldots, \underbrace{Y_{t_k}, \ldots, Y_{t_k}}_{n_k} := P_{n_1, \ldots, n_k}(Y_{t_1}, \ldots, Y_{t_k})$$

(the indices in P correspond to the number of times that the variables in ":   :" are repeated). The resulting multivariate polynomials $P_{n_1, \ldots, n_k}$ are known as Appell polynomials. These polynomials are a generalization of the Hermite polynomials, which are obtained if Y_t are Gaussian, and like them they play an important role in the limit theory of quadratic forms of dependent variables (*cf.* [14,36,68]).

4) The Appell polynomials may also be directly defined by "power-type" recursions like:

$$\frac{\partial}{\partial x_j} P_{n_1,\ldots,n_k}(x_1,\ldots,x_k) = n_j P_{n_1,\ldots,n_j-1,\ldots,n_k}(x_1,\ldots,x_k), \quad E P_{n_1,\ldots,n_k}(X_1,\ldots,X_k) = 0$$

$$\forall n_j \geq 0, j = 1,\ldots,k, \ \sum_j n_j \geq 1,$$

$$P_{0,\ldots,0}(x_1,\ldots,x_k) = 1.$$

For example, when $m = n = 1$, $P_{1,1}(X_t, X_s) = X_t X_s - \mathbb{E} X_t X_s$, and the bilinear form $Q_T(P_{1,1})$ is a weighted periodogram with its expectation removed.

Note that the multivariate Appell polynomials can be defined by using characteristic functions as well (see, *e.g.*, [69]).

APPENDIX B. THE DIAGRAM FORMULA AND THE MOMENTS/CUMULANTS OF SUMS/BILINEAR FORMS OF WICK PRODUCTS

B.1. **The cumulants diagram representation**

An important property of the Wick products is the existence of simple combinatorial rules for calculation of the (mixed) cumulants, analogous to the familiar diagrammatic formalism for the mixed cumulants of the Hermite polynomials with respect to a Gaussian measure [54]. Let us assume that W is a union of (disjoint) subsets $W_1,\ldots,W_k$. If $(i,1),(i,2),\ldots,(i,n_i)$ represent the elements of the subset W_i, $i = 1,\ldots,k$, then we can represent W as a table consisting of rows $W_1,\ldots,W_k$, as follows:

$$\begin{pmatrix} (1,1),\ldots,(1,n_1) \\ \ldots\ldots\ldots \\ (k,1),\ldots,(k,n_k) \end{pmatrix} = W. \tag{B.1}$$

By a *diagram* γ we mean a partition $\gamma = (V_1,\ldots,V_r)$, $r = 1,2,\ldots$ of the table W into nonempty sets V_i (the "edges" of the diagram) such that $|V_i| \geq 1$. We shall call the edge V_i of the diagram γ *flat*, if it is contained in one row of the table W; and *free*, if it consists of one element, *i.e.* $|V_i| = 1$. We shall call the diagram *connected*, if it does not split the rows of the table W into two or more disjoint subsets. We shall call the diagram $\gamma = (V_1,\ldots,V_r)$ *Gaussian*, if $|V_1| = \ldots = |V_r| = 2$. Suppose given a system of random variables $Y_{i,j}$ indexed by $(i,j) \in W$. Set for $V \subset W$,

$$Y^V = \prod_{(i,j)\in V} Y_{i,j}, \quad \text{and} \quad :Y^V := :(Y_{i,j},(i,j) \in V):.$$

For each diagram $\gamma = (V_1,\ldots,V_r)$ we define the number

$$I_\gamma = \prod_{j=1}^r \chi(Y^{V_j}). \tag{B.2}$$

Proposition B.1 (*cf.* [36,68]))**. *Each of the numbers*

$$(i) \quad EY^W = E(Y^{W_1}\ldots Y^{W_k}),$$
$$(ii) \quad E(:Y^{W_1}:\ldots:Y^{W_k}:),$$
$$(iii) \quad \chi(Y^{W_1},\ldots,Y^{W_k}),$$
$$(iv) \quad \chi(:Y^{W_1}:,\ldots,:Y^{W_k}:)$$

is equal to

$$\sum I_\gamma,$$

where the sum is taken, respectively, over

    (i) *all diagrams;*
    (ii) *all diagrams without flat edges;*
    (iii) *all connected diagrams;*
    (iv) *all connected diagrams without flat edges.*

If $EY_{i,j} = 0$ for all $(i,j) \in W$, then the diagrams in (i)–(iv) have no singletons.

Notes. 1) Part (i) is just the exponential relation between moments and cumulants.

2) From part (ii) follows, for example, that $E : Y^W := 0$ (take $W = W_1$, then W has only 1 row and all diagrams have flat edges).

B.2. **Multilinearity**

An important property of Wick products and of cumulants is their multilinearity. For sums and bilinear forms

$$S_T = S_T^m = \int_{I_T} P_m(X_t)\nu(dt), \quad Q_T = Q_T^{m,n} = \int_{I_T} \int_{I_T} \hat{b}(t-s)P_{m,n}(X_t, X_s)\nu(dt)\nu(ds)$$

this implies that:

$$\chi_k(S_T, ..., S_T) \;=\; \int_{t_i \in I_T} \chi(: X_{t_1,1}, \ldots, X_{t_1,m} :, \ldots, : X_{t_k,1}, \ldots, X_{t_k,m} :) \prod_{i=1}^{k} \nu(dt_i),$$

where the cumulant in the integral is taken for a table W of k rows $R_1...., R_k$, each containing the Wick product of l variables identically equal to X_{t_k};

$$\chi_k(Q_T, ..., Q_T) \;=\; \int_{t_i, s_i \in I_T} \chi(: X_{t_1,1}, \ldots, X_{t_1,m}, \; X_{s_1,1}, \ldots, X_{s_1,n} :,$$

$$\ldots, : X_{t_k,1}, \ldots, X_{t_k,m}, \; X_{s_k,1}, \ldots, X_{s_k,n} :) \prod_{i=1}^{k} \hat{b}(t_i - s_i)\nu(dt_i)\nu(ds_i),$$

where the cumulant in the integral needs to be taken for a table W of k rows $R_1...., R_k$, each containing the Wick product of m variables identically equal to X_{t_k} and of n variables identically equal to X_{s_k}.

B.3. **The cumulants of sums and quadratic forms of moving average tables**

By part (iv) of Proposition B.1, applied to a table W of k rows $R_1...., R_k$, with $K = n_1 + ...n_k$ variables, and by Definition (B.2) and of $I\gamma$, we find the following formula for the cumulants of the Wick products of linear variables (5.7):

$$\chi(: X_{t_1,1}, \ldots, X_{t_1,n_1} :, \ldots, : X_{t_k,1}, \ldots, X_{t_k,n_k} :) = \sum_{\gamma \in \Gamma(n_1,...,n_k)} \kappa_\gamma J_\gamma(\vec{t}), \tag{B.3}$$

where $\Gamma(n_1, \ldots, n_k)$ denotes the set of all connected diagrams $\gamma = (V_1, \ldots, V_r)$ without flat edges of the table W, $\kappa_\gamma = \chi_{|V_1|}(\xi_{I_1}) \ldots \chi_{|V_r|}(\xi_{I_1})$ and

$$
\begin{aligned}
J_\gamma(t_1, ..., t_K) &= \prod_{j=1}^{r} J_{V_j}(t_{V_j}) \quad\quad\quad\quad\quad\quad\quad\quad\quad\quad\quad\quad\quad\quad\quad\quad\quad\text{(B.4)}\\
&= \int_{s_1, \ldots, s_r \in \ I} \prod_{j=1}^{k} \Big[\hat{a}(t_{j,1} - s_{j,1}) \hat{a}(t_{j,n_1} - s_{j,n_1}) \ldots \\
&\quad\quad \ldots \hat{a}(t_{k,1} - s_{k,1}) \ldots \hat{a}(t_{k,n_k} - s_{k,n_k}) \Big] \nu(ds_1), \ldots, \nu(ds_r) \\
&= \int_{\lambda_1, \ldots, \lambda_K} e^{i \sum_{j=1}^{K} t_j \lambda_j} \prod_{i=1}^{K} a(\lambda_i) \prod_{j=1}^{r} \delta\left(\sum_{i \in V_j} \lambda_i \right) \prod_{i=1}^{K} \mu(d\lambda_i) \\
&= \int_{\lambda_1, \ldots, \lambda_K} e^{i \sum_{j=1}^{K} t_j \lambda_j} \prod_{j=1}^{r} \left(f_{|V_j|}(\lambda_{j,1}, ..., \lambda_{j,|V_j|-1}) \delta\left(\sum_{i \in V_j} \lambda_i \right) \right) \prod_{i=1}^{K} \mu(d\lambda_i),
\end{aligned}
$$

where $s_{k,i} \equiv s_j$ if $(k,i) \in V_j$, $j = 1, \ldots, r$ and $\lambda_{j,i} \equiv \lambda_{i+\sum_{l=1}^{j-1} |V_l|}$ if $(j,i) \in V_j$, $j = 1, \ldots, r$.

B.4. **The cumulants of sums and quadratic forms of moving average processes**

We will apply now the formula (B.4) to compute the cumulants of $S_T^{(m)}, Q_T^{(m,n)}$. In this case, each row j contains just one, respectively two random variables.

It is easy to check that the variance of $S_T^{(2)}$ is:

$$
\chi_2(S_T^{(2)}) = 2 \int_{\lambda_1, \lambda_2 \in S} f(\lambda_1) f(\lambda_2) \Delta_T(\lambda_1 - \lambda_2) \Delta_T(\lambda_2 - \lambda_1) \prod_{e=1}^{2} \mu(d\lambda_e).
$$

Note that there are only two possible diagrams on a table with two rows of size 2, and that they yield both a graph on two vertices (corresponding to the rows), connected one to the other *via* two edges.

For another example, the third cumulant $\chi_3(S_T^{(2)})$ is a sum of terms similar to:

$$
2^2 \int_{\lambda_1, \lambda_2, \lambda_3 \in S} f(\lambda_1) f(\lambda_2) f(\lambda_3) \Delta_T(\lambda_1 - \lambda_2) \Delta_T(\lambda_2 - \lambda_3) \Delta_T(\lambda_3 - \lambda_1) \prod_{e=1}^{3} \mu(d\lambda_e).
$$

This term comes from the 2^2 diagrams in which the row 1 is connected to row 2, 2 to 3 and 3 to 1.

For quadratic forms, a further application of part (iv) of Proposition B.1 will decompose this as a sum of the form

$$
\sum_{\gamma \in \Gamma(n_1, \ldots, n_k)} \int_{t_i, s_i \in I_T} R_\gamma(t_i, s_i) \prod_{i=1}^{k} \hat{b}_{t_i - s_i} dt_i ds_i,
$$

where $\Gamma(n_1, \ldots, n_k)$ denotes the set of all connected diagrams $\gamma = (V_1, \ldots, V_r)$ without flat edges of the table W and $R_\gamma(t_i, s_i)$ denotes the product of the cumulants corresponding to the partition sets of γ. This easy to check formula is also an illustration of the diagram formula.

When $m = n = 1$ and $k = 2$, besides the Gaussian diagrams we have also one diagram including all the four terms, which makes intervene the fourth order cumulant of X_t.

When $m = n = 1$, the Gaussian diagrams are all products of correlations and the symmetry of $\hat{b}$ implies that all these $2^{k-1}(k-1)!$ terms are equal. We get thus the well-known formula for the cumulants of discrete Gaussian bilinear forms.

In general, we find decompositions as sums of certain "Fejér graph integrals", associated to specific graph structures.

The general structure of the intervening graphs for the cumulants of sums S_T and quadratic forms Q_T have been discussed in Section 3.2.

The following proposition is easy to check.

Proposition B.2. *Let $X_t, t \in I_T$ denote a stationary linear process given by (5.7) with $d = 1$. Then, the cumulants of the sums and quadratic forms defined in (B.4) are given respectively by:*

$$\chi_{k,l} = \chi_k(S_T^{(m)}, ..., S_T^{(m)}) = \sum_{\gamma \in \Gamma(m,k)} \kappa_\gamma \, \sigma_\gamma(T)$$

and

$$\chi_{k,m,n} = \chi_k(Q_T^{(m,n)}, ..., Q_T^{(m,n)}) = \sum_{\gamma \in \Gamma(m,n,k)} \kappa_\gamma \, \tau_\gamma(T),$$

where $\Delta_T(x)$ is the Fejér kernel, $\Gamma(l,k)$, $\Gamma(m,n,k)$ were defined above, and

$$
\begin{aligned}
\sigma_\gamma(T) &= \int_{\vec{t} \in I_T^k} J_\gamma(\vec{t}) \mathrm{d}t \qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad (B.5)\\
&= \int_{\lambda_1,...,\lambda_K} \prod_{j=1}^k \Delta_T \left(\sum_{i=m(j-1)+1}^{mj} \lambda_i \right) \prod_{i=1}^K a(\lambda_i) \prod_{j=1}^r \delta \left(\sum_{i \in V_j} \lambda_i \right) \prod_{i=1}^K \mathrm{d}\lambda_i, \\
\tau_\gamma(T) &= \int_{\mu_1,...,\mu_k,\lambda_1,...,\lambda_K,\lambda'_1,...,\lambda'_{K'}} \prod_{j=1}^k \left[\Delta_T \left(\mu_j + \sum_{i=m(j-1)+1}^{mj} \lambda_i \right) \Delta_T \left(-\mu_j + \sum_{i=n(j-1)+1}^{nj} \lambda'_i \right) b(\mu_j) \right] \\
&\quad \times \prod_{i=1}^K a(\lambda_i) \prod_{i=1}^{K'} a(\lambda'_i) \prod_{j=1}^r \delta \left(\sum_{i \in V_j} \lambda_i + \sum_{i \in V_j} \lambda'_i \right) \prod_{i=1}^K d\lambda_i \prod_{i=1}^{K'} \mathrm{d}\lambda'_i \prod_{i=1}^k \mathrm{d}\mu_i.
\end{aligned}
$$

These graph structures are simple enough to allow a quick evaluation of the orders of magnitude $\alpha_M(z)$, *via* the corresponding graph-breaking problems; for the case of bilinear forms we refer to Lemma 1 in [11].

For the case of sums, the domain of applicability of the CLT is $1 - z_1 \geq 1/m$. We check now that at the extremal point $1 - z_1 = 1/m$ we have

$$
\begin{aligned}
\alpha_G(z_1) &= \max_A p(A) \\
&= \max_A \left[co(G - A) - \sum_{e \in A} (1 - z_e) \right] \\
&= \max_A [co(G - A) - |A|(1 - z_1)] \\
&\leq k/2, \; \forall G \in \mathcal{G}_k,
\end{aligned}
$$

where we interpret $p(A)$ as a "profit", equal to the "gain" $co(G - A)$ minus the "cost" $\sum_{e \in A}(1 - z_e)$. We thus need to show that at the extremal point $1 - z_1 = 1/m$,

$$co(G - A)] \leq |A|/m + k/2, \; \forall G \in \mathcal{G}_k.$$

Indeed, this inequality holds with equality for the "total breaking" $A = \mathcal{E}$ (which contains $(km)/2$ edges). It is also clear that no other set of edges A can achieve a bigger "profit" $p(A)$ (defined in (B.6)) than the total breaking, since for any other set A which leaves some vertex still attached to the others, the vertex could be

detached from the others with an increase of the number of components by 1 and a cost no more than $m\frac{1}{m}$; thus the profit is nondecreasing with respect to the number of vertices left unattached and thus the total breaking achieves the maximum of $p(A)$.

APPENDIX C. THE HOMOGENEOUS HÖLDER-YOUNG-BRASCAMP-LIEB INEQUALITY

Subtle modifications of the conditions of the Hölder inequality must be made when the arguments of the functions involved are restricted to some subspaces [30]. Starting with Brascamp and Lieb [23] and Lieb [53] (who considered only the case $S = \mathbb{R}$), and following with Ball [17], Barthe [18] and Carlen *et al.* [26], this generalization of the classical inequalities of Hölder and Young seems to have attained now its definite form in the work of Bennett *et al.* [19**?**].

We review now a particular case of this result.

Let
$$x = (x_1, \ldots, x_m) \in S^m,$$
where S may be either the multidimensional torus, integers or reals
$$S = \begin{cases} [-\pi, \pi]^d \\ \mathbb{Z}^d \\ \mathbb{R}^d \end{cases}$$
endowed with the respective normalized Haar measure $\mu(\mathrm{d}x)$.

When $d = 1$, the convergence of integrals of the form:
$$\int_{x \in S^m} \frac{\mu(\mathrm{d}x_1)...\mu(\mathrm{d}x_m)}{l_1(x)^{z_1}...l_k(x)^{z_k}} \tag{C.1}$$
where $(l_1, ..., l_k)$ are linear transformations
$$l_j : S^m \rightarrow S, \quad l_j(x) = \langle \alpha_j, x \rangle, \quad j = 1, ..., k$$
and where in the first two cases α_j are supposed to have integer coefficients, is a fundamental question arising in many applications.

Let M denote the matrix with columns $\alpha_j, j = 1, ..., k$. It was for long known to physicists that, when M is fixed, convergence holds for $z = (z_1, ..., z_k)$ belonging to a certain "power counting polytope" PCP (these are relatively similar in all the three cases – see Theorem C.1 below).

It was first noticed in [12] and [15], in the easier case of unimodal matrices M, that under the same "power counting conditions" on $z_j = p_j^{-1}$, a Hölder-type inequality

$$(GH) \qquad \left| \int_{S^m} \prod_{j=1}^k f_j(l_j(x)) \mu(\mathrm{d}x) \right| \leq K \prod_{j=1}^k \|f_j\|_{p_j}$$

holds, with the powers in (C.1) being replaced by arbitrary functions satisfying integrability conditions $f_j \in L_{p_j}, j = 1, ..., k$, and with $K = 1$. Note that Brascamp and Lieb [23] had already studied the analog harder inequality for general matrices (in the case $S = \mathbb{R}^d$), but without pinpointing exactly the polytope; this was done later by Barthe [18].

For an example, consider the integral
$$J = \int_S \int_S f_1(x_1) f_2(x_2) f_3(x_1 + x_2) f_4(x_1 - x_2) \mathrm{d}x_1 \mathrm{d}x_2$$

where $S = \mathbb{R}$. Here $m = 2$, $k = 4$ and the matrix

$$M = \begin{pmatrix} 1 & 0 & 1 & 1 \\ 0 & 1 & 1 & -1 \end{pmatrix}$$

has rank $r(M) = 2$. The theorem below will ensure that

$$|J| \leq \|f_1\|_{1/z_1} \, \|f_2\|_{1/z_2} \, \|f_3\|_{1/z_3} \|f_4\|_{1/z_4}$$

for any $z = (z_1, z_2, z_3, z_4) \in [0,1]^4$ satisfying $z_1 + z_2 + z_3 + z_4 = 2$, *e.g.* if $z = (0, 1, 1/2, 1/2)$, then

$$|J| \leq \left(\sup_{0 \leq x \leq 1} |f_1(x)| \right) \left(\int_0^1 |f_2(x)| \mathrm{d}x \right) \left(\int_0^1 f_3^2(x) \mathrm{d}x \right)^{1/2} \left(\int_0^1 f_4^2(x) \mathrm{d}x \right)^{1/2}.$$

It is easy to check (and true in general) that the extremal points of the PCP have only 0 and 1 coordinates (which may be exploited for establishing the result). Note also that the matrix M in this example is not unimodal; as a consequence, the optimal constant $K = K(z)$ is not 1 at all the extremal points, the exception being $(0, 0, 1, 1)$, where it is 2^{-1}; also, the functions achieving equality must be Gaussian (which holds in general, *cf.* Brascamp-Lieb [23]).

We will formulate now simultaneously the Hölder-Young-Brascamp-Lieb inequality in the three cases:

- (C1) $\mu(\mathrm{d}x_j)$ is normalized Lebesgue measure on the torus $[-\pi, \pi]^d$, and M has all its coefficients integers.
- (C2) $\mu(\mathrm{d}x_j)$ is counting measure on $\mathbb{Z}^d$, M has all its coefficients integers, and is unimodular, *i.e.* all its non-singular minors of dimension $m \times m$ have determinant ± 1.
- (C3) $\mu(\mathrm{d}x_j)$ is Lebesgue measure on $(-\infty, +\infty)^d$.

The result below specifies the domain of validity of Hölder's inequality, called power counting polytope, in terms of linear inequalities involving the rank $r(A)$ of arbitrary subsets A of columns of the matrix M (including the empty set $\emptyset$). It is also possible to express the inequalities in terms of the "dual rank" $r^*(A)$ defined by a dual matrix M^* whose lines are orthogonal to those of M, by using the duality relation

$$r^*(A) = |A| - r(M) + r(A^c), \ \forall A$$

Theorem C.1 (Homogeneous Hölder-Young-Brascamp-Lieb inequality). *Let $l_j(x) = x^t \alpha_j, j = 1, ..., k$ be linear functionals $l_j : S^m \to S$ where the space S is either the torus $[-\pi, \pi]^d$, $\mathbb{Z}^d$, or $\mathbb{R}^d$. Let M denote the matrix with columns $\alpha_j, j = 1, ..., k$, and let $r(A), r^*(A)$, denote the rank and dual rank of any set A of columns of M.*

Let $f_j, \ j = 1, \ldots, k$ be functions $f_j \in L_{p_j}(\mu(\mathrm{d}x)), \ 1 \leq p_j \leq \infty$, defined on S, where $\mu(\mathrm{d}x)$ is respectively normalized Lebesgue measure, counting measure and Lebesgue measure.

Let $z_j = \frac{1}{p_j}, \ j = 1, \ldots, k$, and $z = (z_1, ..., z_k)$. The Hölder-Young-Brascamp-Lieb inequality (GH) will hold (with $K = K(z) < \infty$) throughout the "power counting polytopes" PCP defined respectively by:

- (c1) $\sum_{j \in A} z_j \leq r(A), \ \forall A$;
- (c2) $\sum_{j \in A} z_j \geq r(M) - r(A^c), \quad \Leftrightarrow \quad \sum_{j \in A}(1 - z_j) \leq r^*(A) \ \forall A$;
- (c3) $\sum_{j=1}^k z_j = m$, *and one of the conditions (c1) or (c2) is satisfied.*

Alternatively, the conditions (c1–c3) in the theorem are respectively equivalent to:

- (1) $z = (z_1, \ldots, z_k)$ *lies in the convex hull of the indicators of the sets of independent columns of M, including the void set.*

(2) $z = (z_1, \ldots, z_k)$ *lies in the convex hull of the indicators of the sets of columns of* M *which span its range.*

(3) $z = (z_1, \ldots, z_k)$ *lies in the convex hull of the indicators of the sets of columns of* M *which form a basis.*

If, moreover, the matrix M *is unimodal, then the Hölder-Young-Brascamp-Lieb inequality (GH) holds with constant* $K = 1$.

Notes. 1) Polytopes defined by the type of rank constraints appearing in cases (c1–c2) are called "polymatroids" (associated to M and M^*) – see Welsh [72], 18.3, Theorem 1. The third polytope is the intersection of the first two.

2) The first two cases of Theorem C.1 were obtained for unimodal matrices in [12] and [15], respectively.

3) Some further important issues, like the precise formula for $K(z)$, and the nonhomogeneous extension where l_j may be linear operators with possibly different images, were resolved only recently – see Lieb [53], Bennett *et al.* [19].

Proof sketch. By Edmonds theorem (see Welsh [72], 18.4, Thm. 1) the extremal points of the above polymatroids have only 0 and 1 coordinates for any matrix M. This fact leads to an easy proof, since at the extremal points the result is immediate. For example, in the first and third cases, the extremal points are in one to one correspondence with the indicators of independent sets and bases A, respectively, and the constant at such a point, by a change of variables, is seen to be:

$$K_A = \sqrt{\det(AA^t)}.$$

Since $K(z)$ is finite at the extremal points, and Riesz-Thorin interpolation ensures the convexity of $\log(K(z))$, it follows that $K(z)$ will remain finite throughout the polytope generated by the indicators 1_A. $\square$

APPENDIX D. KERNEL ESTIMATES

Consider the Dirichlet type kernel

$$\Delta_T(\lambda) = \int_{t \in I_T} e^{it\lambda} \nu(dt).$$

When $d = 1$ and $I_1 = [-1/2, 1/2]$ $(I_T = I_1 T)$, one gets the classical discrete/continuous time Dirichlet kernels:

$$\Delta_T(\lambda) = \sum_{-T/2}^{T/2} e^{it\lambda} = \frac{\sin((T+1)\lambda/2)}{\sin(\lambda/2)}, \quad \Delta_T(\lambda) = \int_{-T/2}^{T/2} e^{it\lambda} dt = \frac{\sin(T\lambda/2)}{\lambda/2},$$

respectively. For general d and $I_T = [-T/2, T/2]^d$, putting $\lambda = (\lambda_1, .., \lambda_d)$, it follows that $\Delta_T(\lambda) = \prod_{j=1}^{d} \Delta_T(\lambda_j)$.

Note that in the continuous case, by scaling, one finds

$$\|\Delta_T(\lambda)\|_p = T^{1-1/p} C_p, \quad \lambda \in \mathbb{R}, \quad 1 < p < \infty, \tag{D.1}$$

$$\|\Delta_T(\lambda)\|_p = T^{d(1-1/p)} C_p^d, \quad \lambda \in \mathbb{R}^d, \quad 1 < p < \infty \tag{D.2}$$

with $C_p = (2 \int_R |\frac{\sin(z)}{z}|^p dz)^{\frac{1}{p}}$. Note that $C_{p_1} > C_{p_2}$ for $p_1 < p_2$.

In the discrete case, similar estimates may be obtained by using the inequality

$$\left| \frac{\sin((T+1)\lambda/2)}{\sin(\lambda/2)} \right| \leq \tilde{C} \frac{T}{1 + T\|\lambda\|}, \lambda \in [-\pi, \pi).$$

We find then:

$$\|\Delta_T(\lambda)\|_p \leq T_p^{1-1/p} \tilde{C}^{\frac{1}{p}} \left(\int_{\mathbb{R}} \frac{dz}{(1 + |z|^p)} \right)^{\frac{1}{p}}, \quad \lambda \in [-\pi, \pi), \ 1 < p < \infty,$$

$$\|\Delta_T(\lambda)\|_p \leq T_p^{d(1-1/p)}\left[\tilde{C}^{\frac{1}{p}}\left(\int_{\mathbb{R}}\frac{\mathrm{d}z}{(1+|z|^p)}\right)^{\frac{1}{p}}\right]^d, \quad \lambda = (\lambda_1,..,\lambda_d) \in [-\pi,\pi)^d,\ 1 < p < \infty.$$

In the case of the Euclidean ball $I_T = B_T = \{t \in \mathbb{R}^d : \|t\| \leq T/2\}$, we find again by scaling in the continuous case

$$\Delta_T(\lambda) = \int_{B_T}\mathrm{e}^{\mathrm{i}t\lambda}\mathrm{d}t = (2\pi)^{\frac{d}{2}}J_{d/2}\left(\|\lambda\|\frac{T}{2}\right)/\|\lambda\|^{d/2}, \quad \lambda \in \mathbb{R}^d,$$

where $J_\nu(z)$ is the Bessel function of the first kind and order ν. It is known that $J_\nu(z) \leq \mathrm{const.}/\sqrt{z}$ for a large z, thus for the ball

$$\|\Delta_T(\lambda)\|_p = \begin{cases} T^{(1-\frac{1}{p})}C_p, & d = 1, \quad p > 1, \\ T^{d(\frac{1}{2}-\frac{1}{p})}\bar{C}_p, & d \geq 2, \quad p > \frac{2d}{d+1}, \end{cases}$$

$$\bar{C}_p = 2^{d(\frac{1}{2}-\frac{1}{p})}(2\pi)^{\frac{d}{2}}|s(1)|\left(\int_0^\infty \rho^{d-1}\left|\frac{J_{\frac{d}{2}}(\rho)}{\rho^{d/2}}\right|^p\right)^{1/p},$$

where $|s(1)|$ is the surface area of the unit ball in $\mathbb{R}^d$, $d \geq 2$.

Similar estimates may be obtained for the L_p $(1 < p \leq \infty)$ norms of the discrete Dirichlet kernel.

Note. These results are particular cases of the so-called Hardy-Littlewood theorem (see, for instance, Zygmund [76], V. II, XII, Sect. 6), which can be formulated as follows:

Theorem D.1. *Let $a_n \geq a_{n+1} \geq \ldots$ and $a_n \to 0$. Consider the series*

$$\sum_{n=1}^\infty a_n \cos n\lambda \tag{D.3}$$

and

$$\sum_{n=1}^\infty a_n \sin n\lambda \tag{D.4}$$

and define by $f(\lambda)$ and $g(\lambda)$ the sums of the series (D.3) and (D.4) respectively at the points where the series converge. A necessary and sufficient condition that the function f (or g) belongs to L_p, $1 < p < \infty$, is the following

$$\sum_{n=1}^\infty a_n^p n^{p-2} < \infty.$$

Moreover,

$$\|f\|_p^p \asymp \sum_{n=1}^\infty a_n^p n^{p-2}.$$

Clearly, $1 = 1 = \ldots = 1 > 0 = \ldots$ is nonincreasing, and we arrive thus to the following estimate for Dirichlet kernels:

$$\left\|\sum_{t=1}^T \mathrm{e}^{\mathrm{i}t\lambda}\right\|_p \leq C\left(\sum_{t=1}^T 1^p t^{p-2}\right)^{\frac{1}{p}} \leq CT^{\frac{p-1}{p}},\ 1 < p < \infty,$$

and

$$\int_0^1\left|\sum_{t=0}^{T-1}\mathrm{e}^{2\pi\mathrm{i}t\lambda}\right|^p\mathrm{d}\lambda = T^{p-1}\frac{2}{\pi}\int_0^\infty\left|\frac{\sin u}{u}\right|^p\mathrm{d}u + R_p(T),$$

where the error term

$$R_p(T) = \begin{cases} O_p(T^{p-3}), & p > 3 \\ O(\log T), & p = 3 \\ O_p(1), & 1 < p < 3, \end{cases}$$

where O_p means that constants depend on p. (See, *e.g.*, [1].)

Note that for $p = 1$ for the Lebesgue constant L_T we have (see, Zygmund [76], p. 67):

$$L_T = \int_0^1 \left| \sum_{t=0}^{T-1} e^{2\pi i t \lambda} \right| d\lambda \asymp \frac{4}{\pi^2} \log T.$$

APPENDIX E. FRACTIONAL RIESZ-BESSEL MOTION

In this Appendix we mainly review a number results discussed in Gay and Heyde [32], Anh *et al.* [4], Anh *et al.* [6], Anh and Leonenko [2], [3], Kelbert *et al.* [50] (see also references therein). Also we introduce a not necessarily Gaussian Riesz-Bessel stationary process and formulate the central limit theorem for such a processes as well as for quadratic forms of such a processes.

The fractional operators are natural mathematical objects to describe the singular phenomena of random fields such as long range dependence or/and intermittency.

In particular Gay and Heyde [32] introduced a class of random fields as solutions of fractional Helmholtz equation driven by white noise, contained the fractional operator $(cI - \Delta)^{\alpha/2}$, $c \geq 0$ (and its limit as $c \to 0$ $(-\Delta)^{\alpha/2}$), where Δ is the d-dimensional Laplacian and I is the identity operator (see also [50] for properties of such fields and possible generalization). Anh *et al.* [4] (see also [2,3,6] and references therein) generalized the fractional stochastic equation of Gay and Heyde in which the fractional Helmholtz operator $(cI - \Delta)^{\alpha/2}$, $c \geq 0$ or the d-dimensional Laplacian $(c \to 0)$ is replaced by a fractional Laplace-type operator of the form $-(I - \Delta)^{\gamma/2} (-\Delta)^{\alpha/2}, \alpha > 0$, $\gamma \geq 0$, where the operators $-(I - \Delta)^{\gamma/2}, \gamma \geq 0$, and $(-\Delta)^{\alpha/2}$, $\alpha > 0$, are interpreted as inverses to the Bessel and Riesz potentials (see [66], pp. 134–138), that is integral operators, whose kernels have a Fourier transforms $(2\pi)^{-d/2} \left(1 + \|\lambda\|^2 \right)^{-\gamma/2}$, $\lambda \in \mathbb{R}^d$, and $(2\pi)^{-d/2} \|\lambda\|^{-\alpha}$, $\lambda \in \mathbb{R}^d$, respectively. Then there exists a generalized random field $\zeta(x)$, $x \in \mathbb{R}^d$, on fractional Sobolev space, which is defined by the equation

$$(I - \Delta)^{\gamma/2} (-\Delta)^{\alpha/2} \zeta(x) = e(x), \ x \in \mathbb{R}^d, \tag{E.1}$$

where $\{e(x), \ x \in \mathbb{R}^d\}$ is a Gaussian white noise or equivalently (in the sense of second-order moments) there exists a random field with the spectral density

$$f(\lambda) = \frac{c}{\|\lambda\|^{2\alpha} \left(1 + \|\lambda\|^2 \right)^\gamma}, \ \lambda \in \mathbb{R}^d, \ c > 0. \tag{E.2}$$

These random fields were named the fractional Riesz-Bessel motion.

For the random fields with stationary increments we assume $\alpha \in \left(\frac{d}{2}, \frac{d}{2} + 1 \right), \gamma \geq 0$. In particular, for $d = 1$, there exists a Gaussian stochastic process with stationary increments and the spectral density (E.2), where $\alpha \in \left(\frac{1}{2}, \frac{3}{2} \right), \ \gamma \geq 0$. This fractional Riesz-Bessel motion (FRBM) is a generalization of the fractional Brownian motion (FBM) (see, for instance, Samorodnitsky and Taqqu [63]). FBM is a limiting case of the Riesz-Bessel (non-stationary) motion with $\gamma = 0$ (in terms of the Hurst parameter $H \in (0,1)$, the spectral density of the FBM with long-range dependence ($H \in (\frac{1}{2}, 0)$) is equal to $\frac{1}{|\lambda|^{2H+1}}$). The FRBM is not self-similar (unless when $\gamma = 0$), but it is locally self-similar.

For $d \geq 1$ the presence of the Bessel operator is essential for a study of homogeneous (and isotropic) solutions of (E.1) with spectral density (E.2), which requires $0 \leq \alpha < \frac{d}{2}$, $\alpha + \gamma > \frac{d}{2}$; that is the condition $\gamma > 0$ is necessary for $f(\lambda) \in L_1(\mathbb{R}^d)$. Thus the homogeneous isotropic FRBM can be defined as a Gaussian random field

with zero mean and covariance function of the form

$$B_{\alpha,\gamma}(x) = \int_{\mathbb{R}^d} e^{i\langle\lambda,x\rangle} \frac{c}{\|\lambda\|^{2\alpha}\left(1+\|\lambda\|^2\right)^{\gamma}} d\lambda, \ x \in \mathbb{R}^d, \tag{E.3}$$

where $0 \le \alpha < \frac{d}{2}$, $\alpha + \gamma > \frac{d}{2}$. Note that for $\alpha = 0$, the covariance structure (E.3) belongs the Matérn class, that is with

$$c = \frac{\Gamma(\gamma)}{\pi^{d/2}2^{d-1}\Gamma\left(\frac{2\gamma-d}{2}\right)};$$

$$B_{\alpha,\gamma}(x) = \frac{1}{2^{\frac{2-d}{2}}\Gamma\left(\frac{2\gamma-d}{2}\right)} \frac{K_{\frac{2\gamma-d}{2}}}{\|x\|^{\frac{d-2\gamma}{2}}}, \ x \in \mathbb{R}^d, \ \gamma > \frac{d}{2}, \tag{E.4}$$

where

$$K_{\nu}(z) = \frac{1}{2}\int_0^{\infty} s^{\nu-1}\exp\left\{-\frac{1}{2}\left(s+\frac{1}{s}\right)z\right\}ds, \ z \ge 0, \ \nu \in R,$$

is the modified Bessel function of the third kind of order ν or Mc Donald's function. Note that

$$K_{\nu}(z) = K_{-\nu}(z), \quad K_{\nu}(z) \sim \Gamma(\nu)2^{\nu-1}z^{-\nu},$$

for $\nu > 0$ as $z \to 0$,

$$K_{\frac{1}{2}}(z) = \sqrt{\frac{\pi}{2}}\frac{e^{-z}}{\sqrt{z}}.$$

Thus, we have $B_{\alpha,\gamma}(0) = 1$.

Note that for $d = 1$, $\alpha = 0$, $\gamma = 1$, the covariance structure (E.4) becomes $B_{0,1}(x) = e^{-x}$, $x \ge 0$, that is stationary Gaussian Riesz-Bessel motion is identical to the Gaussian Ornstein-Uhlenbeck process.

Remark E.1. These results can be generalized to the case when the above fractional operator is replaced by the operator

$$H = \frac{\partial^{\beta}}{\partial t^{\beta}} + \mu\left(I-\Delta\right)^{\gamma/2}\left(-\Delta\right)^{\alpha/2}, \ 0 \le \beta \le 2, \ \alpha > 0, \ \gamma \ge 0,$$

where $\frac{\partial^{\beta}}{\partial t^{\beta}}$ is the regularized fractional derivative. In particular, the Green function of the fractional heat equation: $Hu(t,x) = 0$, $t > 0$, $x \in \mathbb{R}^d$, can be given as inverse Fourier transform of the function

$$E_{\beta}\left(-\mu t^{\beta}\|\lambda\|^{\alpha}\left(1+\|\lambda\|^2\right)^{\gamma/2}\right), \ t > 0, \ \lambda \in \mathbb{R}^d,$$

where

$$E_{\beta}(z) = \sum_{j=1}^{\infty}\frac{z^j}{\Gamma(\beta_j+1)}, \ z \in \mathbb{C}, \ \beta > 0$$

is the Mittag-Leffler function (see [2,3] for details and references).

In order to introduce a Riesz-Bessel motion driven by Lévy noise, we restrict our attention to the stationary case and $d = 1$ (replacing the space parameter x into t). For the function

$$a(\lambda) = \frac{\sqrt{c}}{(i\lambda)^{\alpha}(1+i\lambda)^{\gamma}}, \ \lambda \in \mathbb{R}, \ a+\gamma > 1, \ \alpha \ge 0,$$

we introduce the function

$$\hat{a}(t) = \int_{\mathbb{R}} e^{it\lambda} a(\lambda) d\lambda = \begin{cases} \frac{2\pi}{\Gamma(\alpha+\gamma)} t^{\alpha+\gamma-1} e^{-1} {}_1F_1(\gamma, \alpha+\gamma; t), & t \geq 0, \\ 0, & t < 0, \end{cases} \tag{E.5}$$

where the confluent hypergeometric function

$${}_1F_1(a, b; z) = \sum_{n=0}^{\infty} \frac{(a)_n}{(b)_n} \frac{z^n}{n!}, \ (c)_n = c(c+1)\ldots(c+n-1), \ (c)_0 = 1.$$

The Riesz-Bessel motion driven by Lévy noise can be defined as the linear process

$$X_t = \int_{\mathbb{R}} \hat{a}(t - s) d\xi(s), \tag{E.6}$$

where $\xi(t)$, $t \in R$ is a Lévy process with cumulant function

$$\kappa(z) = \log E \exp\{iz\xi(1)\},$$

such that $\kappa^{(k)}(0) \neq 0$, $k \geq 2$, and $\hat{a}(.)$ is defined by (E.5). The kth order spectral densities of the Riesz-Bessel motion driven by Lévy noise (E.6) take the form:

$$f_k(\lambda_1, \ldots, \lambda_{k-1}) = (2\pi)^{-k+1} i^{-k} \kappa^{(k)}(0) a(\lambda_1) \ldots a(\lambda_{k-1}) \overline{a(\lambda_1 + \ldots + \lambda_{k-1})}, \tag{E.7}$$

which reduces to the second-order spectral density

$$\begin{aligned} f_2(\lambda) &= \frac{c}{|\lambda|^{2\alpha} (1 + \lambda^2)^{\gamma}}, \\ c &= \frac{\kappa^{(2)}(0)}{2\pi}, \ 0 \leq \alpha < \frac{1}{2}, \ \alpha + \gamma > \frac{1}{2}, \ \lambda \in \mathbb{R}^1. \end{aligned} \tag{E.8}$$

For the Gaussian case, of course, $\kappa^{(k)}(0) = 0$, $k \geq 3$.

Note that for $\alpha = 0$, $\gamma = 1$ we arrive to the Ornstein-Uhlenbeck process driven by Lévy noise ([5]).

As a consequence of the theorems of Sections 3 and 4 we obtain the following result for the linear process (E.6). (*Cf.* also with Thm. 4.1.)

Theorem E.2. *Consider the Riesz-Bessel stationary motion (E.6) and assume that all cumulants of Lévy process are finite. Let*

$$S_T = \int_{-T/2}^{T/2} X_s ds, \quad Q_T = \int_{-T/2}^{T/2} \int_{-T/2}^{T/2} \hat{b}(t - s) \left[X_t X_s - E X_t X_s \right] dt ds,$$

then:

    i) *if*

$$\alpha + \gamma > \frac{1}{2}, \quad \alpha \leq 0,$$

then the central limit theorem holds:

$$T^{-1/2} S_T \rightarrow N(0, \sigma^2), T \rightarrow \infty,$$

where

$$\sigma^2 = \kappa^{(2)}(0);$$

ii) *if for some $p > 1, q > 1$, such that $\frac{1}{p} + \frac{1}{q} \leq \frac{1}{2}$, with equality holding in continuous case, we have*

$$b(\lambda) \in \mathbf{L_q} \cap L_1, \quad \alpha + \gamma > \frac{1}{2p}, \quad \alpha < \frac{1}{2p},$$

then the central limit theorem holds:

$$T^{-1/2}Q_T \to N(0, \sigma^2), \, T \to \infty,$$

where

$$\sigma^2 = 2\kappa^{(2)}(0) \int_{\mathbb{R}} b^2(\lambda) \frac{1}{|\lambda|^{4\alpha}(1+\lambda^2)^{2\gamma}} \mathrm{d}\lambda + \kappa^{(4)}(0) \frac{1}{(2\pi)^2} \int_{\mathbb{R}} b(\lambda) \frac{1}{|\lambda|^{2\alpha}(1+\lambda^2)^{\gamma}} \mathrm{d}\lambda.$$

Remark E.3. The central limit theorem for a linear processes with continuous time of the form (E.6) can be proved also by methods different from those presented in our paper, for instance, *via* truncation and using the central limit theorem for m-dependent random fields, or by the method of characteristic functions.

REFERENCES

[1] B. Anderson, J.M. Ash, R. Jones, D.G. Rider and B. Saffari, Exponential sums with coefficients 0 or 1 and concentrated L^p norms. *Ann. Inst. Fourier* **57** (2007) 1377–1404.

[2] V.V. Anh and N.N. Leonenko, Spectral analysis of fractional kinetic equations with random data. *J. Statist. Phys.* **104** (2001) 1349–1387.

[3] V.V. Anh and N.N. Leonenko, Renormalization and homogenization of fractional diffusion equations with random data. *Probab. Theory Relat. Fields* **124** (2002) 381–408.

[4] V.V. Anh, J.M. Angulo and M.D. Ruiz-Medina, Possible long-range dependence in fractional random fields. *J. Statist. Plann. Infer.* **80** (1999) 95–110.

[5] V.V. Anh, C.C. Heyde and N.N. Leonenko, Dynamic models of long-memory processes driven by Lévy noise. *J. Appl. Probab.* **39** (2002) 730–747.

[6] V.V. Anh, N.N. Leonenko and R. McVinish, Models for fractional Riesz-Bessel motion and related processes. *Fractals* **9** (2001) 329–346.

[7] V.V. Anh, N.N. Leonenko and L.M. Sakhno, Higher-order spectral densities of fractional random fields. *J. Stat. Phys.* **111** (2003) 789–814.

[8] V.V. Anh, N.N. Leonenko and L.M. Sakhno, Quasi-likelihood-based higher-order spectral estimation of random fields with possible long-range dependence. Stochastic Methods and their Applications. *J. Appl. Probab. A* **41** (2004) 35–53.

[9] V.V. Anh, N.N. Leonenko and L.M. Sakhno, On a class of minimum contrast estimators. *J. Statist. Plann. Infer.* **123** (2004) 161–185.

[10] F. Avram, On Bilinear Forms in Gaussian Random Variables and Toeplitz Matrices. *Probab. Theory Relat. Fields* **79** (1988) 37–45.

[11] F. Avram, Generalized Szegö Theorems and asymptotics of cumulants by graphical methods. *Trans. Amer. Math. Soc.* **330** (1992) 637–649.

[12] F. Avram and L. Brown, A Generalized Hölder Inequality and a Generalized Szegö Theorem. *Proc. Amer. Math. Soc.* **107** (1989) 687–695.

[13] F. Avram and R. Fox, Central limit theorems for sums of Wick products of stationary sequences. *Trans. Amer. Math. Soc.* **330** (1992) 651–663.

[14] F. Avram and M.S. Taqqu, Noncentral limit theorems and Appell polynomials. *Ann. Probab.* **15** (1987) 767–775.

[15] F. Avram and M.S. Taqqu, Hölder's Inequality for Functions of Linearly Dependent Arguments. *SIAM J. Math. Anal.* **20** (1989) 1484–1489.

[16] F. Avram and M.S. Taqqu, On a Szegö type limit theorem and the asymptotic theory of random sums, integrals and quadratic forms. *Dependence in probability and statistics*, Lect. Notes Statist. **187**. Springer, New York (2006) 259–286.

[17] K. Ball, Volume ratios and a reverse isoperimetric inequality. *J. London Math. Soc.* **44** (1991) 351–359.

[18] F. Barthe, On a reverse form of the Brascamp-Lieb inequality. *Inventiones Mathematicae* **134** (2005) 335–361.

[19] J. Bennett, A. Carbery, M. Christ and T. Tao, The Brascamp-Lieb inequalities: finiteness, structure and extremals, *Geom. Funct. Anal.* **17** (2008) 1343–1415.

[20] R. Bentkus, On the error of the estimate of the spectral function of a stationary process. *Lietuvos Matematikos Rinkinys* **12** (1972) 55–71 (In Russian).

[21] R. Bentkus, and R. Rutkauskas, On the asymptotics of the first two moments of second order spectral estimators. *Liet. Mat. Rink.* **13** (1973) 29–45.

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

[23] H.J. Brascamp and E. Lieb, Best constants in Young's inequality, its converse, and its generalization to more than three functions. *Adv. Math.* **20** (1976) 151–173.

[24] P. Breuer and P. Major, Central limit theorems for nonlinear functionals of Gaussian fields. *J. Multiv. Anal.* **13** (1983) 425–441.

[25] P.J. Brockwell, Representations of continuous-time ARMA processes. Stochastic Methods and their Applications. *J. Appl. Probab. A* **41** (2004) 375–382.

[26] E.A. Carlen, E.H. Lieb and M. Loss, A sharp analog of Young's inequality on S^n and related entropy inequalities. *J. Geom. Anal.* **14** (2004) 487–520.

[27] R.L. Dobrushin and P. Major, Non-central limit theorems for non-linear functions of Gaussian fields. *Z. Wahrscheinlichkeitstheorie Verw. Geb.* **50** (1979) 27–52.

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

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

[30] E. Friedgut, Hypergraphs, entropy, and inequalities. *Amer. Math. Monthly* **111** (2004) 749–760.

[31] J. Gao, V.V. Anh and C.C. Heyde, Statistical estimation of nonstationary Gaussian process with long-range dependence and intermittency. *Stoch. Process. Appl.* **99** (2002) 295–321.

[32] R. Gay and C.C. Heyde, On a class of random field models which allows long range dependence. *Biometrika* **77** (1990) 401–403.

[33] M.S. Ginovian, On Toeplitz type quadratic functionals of stationary Gaussian processes. *Probab. Theory Relat. Fields* **100** (1994) 395–406.

[34] M.S. Ginovian and A.A. Sahakyan, Limit theorems for Toeplitz quadratic functionals of continuous-time stationary processes. *Probab. Theory Relat. Fields* **138** (2007) 551–579.

[35] L. Giraitis, Central limit theorem for functionals of linear processes. *Lithuanian Math. J.* **25** (1985) 43–57.

[36] L. Giraitis and D. Surgailis, Multivariate Appell polynomials and the central limit theorem, in *Dependence in Probability and Statistics*. Edited by E. Eberlein and M.S. Taqqu. Birkhäuser, New York (1986) 21–71.

[37] L. Giraitis and D. Surgailis, A central limit theorem for quadratic forms in strongly dependent linear variables and its application to asymptotic normality of Whittle estimate. *Probab. Theory Relat. Fields* **86** (1990) 87-104.

[38] L. Giraitis and M.S. Taqqu, Limit theorems for bivariate Appell polynomials, Part 1: Central limit theorems. *Probab. Theory Relat. Fields* **107** (1997) 359–381.

[39] L. Giraitis and M.S. Taqqu, Whittle estimator for finite variance non-Gaussian time series with long memory. *Ann. Statist.* **27** (1999) 178–203.

[40] C.W. Granger and R. Joyeux, An introduction to long-memory time series models and fractional differencing. *J. Time Ser. Anal.* **10** (1990) 233– 257.

[41] V. Grenander and G. Szegö, *Toeplitz forms and their applications*. University of California Press, Berkeley (1958).

[42] C.C. Heyde, *Quasi-Likelihood And Its Applications: A General Approach to Optimal Parameter Estimation*. Springer-Verlag, New York (1997).

[43] C. Heyde and R. Gay, On asymptotic quasi-likelihood. *Stoch. Process. Appl.* **31** (1989) 223–236.

[44] C. Heyde and R. Gay, Smoothed periodogram asymptotics and estimation for processes and fields with possible long-range dependence. *Stoch.c Process. Appl.* **45** (1993) 169–182.

[45] J.R.M. Hosking, Fractional differencing. *Biometrika* **68** (1981) 165–176.

[46] H.E. Hurst, Long-term storage capacity of reservoirs. *Trans. Amer. Soc. Civil Eng.* **116** (1951) 770–808.

[47] I.A. Ibragimov, On estimation of the spectral function of a stationary Gaussian process. *Theory Probab. Appl.* **8** (1963) 391-430.

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

[49] A.V. Ivanov and N.N. Leonenko, *Statistical Analysis of Random Processes*. Kluwer Academic Publisher, Dordrecht (1989).

[50] M. Kelbert, N.N. Leonenko and M.D. Ruiz-Medina, Fractional random fields associated with stochastic fractional heat equation. *Adv. Appl. Probab.* **37** (2005) 108–133.

[51] S. Kwapien and W.A. Woyczynski, *Random Series and Stochastic Integrals: Single and Multiple*. Birkhaäser, Boston (1992).

[52] N.N. Leonenko and L.M. Sakhno, On the Whittle estimators for some classes of continuous parameter random processes and fields. *Stat. Probab. Lett.* **76** (2006) 781–795.

[53] E.H. Lieb, Gaussian kernels have only Gaussian maximizers. *Invent. Math.* **102** (1990) 179–208.

[54] V.A. Malyshev, Cluster expansions in lattice models of statistical physics and the quantum theory of fields. *Russ. Math. Surveys* **35** (1980) 1–62.

[55] I. Niven, Formal power series. *Amer. Math. Monthly* **76** (1969) 871–889.

[56] D. Nualart and G. Peccati, Central limit theorems for sequences of multiple stochastic integrals. *Ann. Probab.* **33** (2005) 177–193.

[57] J.G. Oxley, *Matroid Theory*. Oxford University Press, New York (1992).

[58] G. Peccati and C.A. Tudor, Gaussian limits for vector-valued multiple stochastic integrals. *Séminaire de Probabilités XXXVIII*. Lect. Notes Math. **1857** 247–262. Springer-Verlag, Berlin (2004).

[59] B. Rajput and J. Rosinski, Spectral representations of infinity divisible processes. *Probab. Theory Relat. Fields* **82** (1989) 451–487.

[60] H. Reiter and J.D. Stegeman, *Classical Harmonic Analysis and Locally Compact Groups*. Oxford University Press, USA (2000).

[61] W. Rudin, *Real and Comlex Analysis*. McGraw-Hill, London, New York (1970).

[62] W. Rudin, *Functional Analysis*. McGraw-Hill, London, New York (1991).

[63] G. Samorodnitsky and M.S. Taqqu, *Stable Non-Gaussian Random Processes*. Chapman and Hall, New York (1994).

[64] V. Solev and L. Gerville-Reache, A sufficient condition for asymptotic normality of the normalized quadratic form $\Psi_n(f, g)$. *C. R. Acad. Sci. Paris, Ser. I* **342** (2006) 971–975.

[65] R. Stanley, *Enumerative combinatorics*. Cambridge University Press (1997).

[66] E.M. Stein, *Singular Integrals and Differential Properties of Functions*. Princeton University Press (1970).

[67] B. Sturmfels, *Grobner bases and convex polytopes*. Volume 8 of University lecture Series. AMS, Providence, RI (1996).

[68] D. Surgailis, On Poisson multiple stochastic integral and associated Markov semigroups. *Probab. Math. Statist.* **3** (1984) 217–239.

[69] D. Surgailis, Long-range dependence and Appel rank, *Ann. Probab.* **28** (2000) 478–497.

[70] M.S. Taqqu, Convergence of integrated processes of arbitrary Hermite rank. *Z. Wahrscheinlichkeitstheorie Verw. Geb.* **50** (1979) 53–83.

[71] W.T. Tutte, Matroids and graphs. *Trans. Amer. Math. Soc.* **90** (1959) 527–552.

[72] D. Welsh, *Matroid Theory*. Academic Press, London (1976).

[73] W. Willinger, M.S. Taqqu and V. Teverovsky, Stock market prices and long-range dependence. *Finance and Stochastics* **3** (1999) 1–13.

[74] P. Whittle, *Hypothesis Testing in Time Series*. Hafner, New York (1951).

[75] P. Whittle, Estimation and information in stationary time series. *Ark. Mat.* **2** (1953) 423–434.

[76] A. Zygmund, *Trigonometric Series*. Volumes I and II. Third edition. Cambridge University Press (2002).