On the generalized Brascamp-Lieb-Barthe inequality, a Szegö type limit theorem, and the asymptotic theory of random sums, integrals and quadratic forms *†

Florin Avram
Université de Pau

Murad S. Taqqu
Boston University

March 6, 2005

Abstract

This paper describes a purely analytic approach for establishing central limit theorems for sums or integrals $\int_{t\in I_T} P_m(X_t)dt$ and quadratic forms $\int_{t,s\in I_T} b(t-s)P_{m,n}(X_t,X_s)dsdt$ of Appell polynomials $P_m(X_t),\ P_{m,n}(X_t,X_s)$, where X_t is a linear process, that is, $X_t=\int_{u\in\mathbb{R}^{\mathbf{d}}} a(t-u)d\xi_u,\ t\in\mathbb{R}^{\mathbf{d}}$, where $\xi_t,\ t\in\mathbb{R}^{\mathbf{d}}$ has stationary independent increments. We review known results and outline some open problems related to this approach.

1 Introduction

1. Time series motivation. One-dimensional discrete time series with long memory have been extensively studied. The FARIMA family, for example, models the series $X_t,\ t\in\mathbb{Z}$ as the solution of an equation

$$\phi(B)(1-B)^d X_t=\theta(B)\epsilon_t \tag{1}$$

where B is the operator of backward translation in time, $\phi(B),\theta(B)$ are polynomials, d is a real number and ϵ_t is white noise. Using this family of models, it is usually possible via an extension of the classical Box-Jenkins methodology, to chose the parameter d and the coefficients of the polynomials $\phi(B),\theta(B)$ such that the residuals ϵ_t display white noise behaviour and hence may safely be discarded for prediction purposes.

In view of the prevalence of spatial statistics applications, it is important to develop models and methods which replace the assumption of a one-dimensional discrete time index with that of a multidimensional continuous one.

This review paper was motivated by the attempt to extend certain central limit theorems of Giraitis and Surgailis, Fox and Taqqu, Avram and Brown and Giraitis and Taqqu to the case of multidimensional continuous indices. We have not yet achieved this goal, but we offer some natural conjectures and

*This research was partially supported by the NSF Grant DMS-0102410 at Boston University.

†*AMS Subject classification.* 60F05, 62M10. *Keywords and phrases:* Quadratic forms, Appell polynomials, cumulants, diagram formula, asymptotic normality, sums with dependent indices, Hölder-Young inequality, Szegö type limit theorem.

the observation that the main part of the one-dimensional analytical approach does extend to the wider multidimensional setup.

The crucial tools are reviewed in two appendices. In Appendix A, section 2 we review several forms of the Brascamp-Lieb-Barthe inequality, a generalisation of the Hölder and Young inequalities which is instrumental for our calculations. In Appendix B, section 3 we review the well-known diagram formula for computing moments/cumulants of Wick products, and its application for studying discrete and continuous time series.

Throughour the paper, we will review the results in [6] and [8], which have been only established in the **classic discrete sum one-dimensional setup**, but we will adopt a unifying measure theoretic notation, in order to be able to discuss possible extensions.

2. The model. Let $\xi_A,\ A \subset \mathbb{R}^d$ denote a set indexed process with mean zero and finite second moments, independent values over disjoint sets and stationary distribution, and let $X_t,\ t \in \mathbb{R}^d$ denote a linear random field

$$X_t = \int_{u \in \mathbb{R}^d} \hat{a}(t-u)\xi(du), \quad t \in \mathbb{R}^d, \tag{2}$$

with a square-integrable kernel $\hat{a}(t),\ t \in \mathbb{R}^d$. For various other conditions which ensure that (2) is well-defined, see for example Anh, Heyde and Leonenko [2], pg. 733.

By choosing an appropriate "Green function" $\hat{a}(t)$, this wide class of processes includes, for example, the solutions of many differential equations with random noise $\xi(du)$.

The random field $X(t)$ is observed on a sequence I_T of increasing finite domains. In the discrete-time case, in keeping with tradition, the case $I_T = [0, T-1]^d$, $T \in \mathbb{Z}_+$ or $I_T = I_n = [1, n]^d$, will be assumed. In the continuous case, rectangles $I_T = \{t \in \mathbb{R}^d : -T_i/2 \leq t_i \leq T_i/2, i = 1, ..., d$ will be taken. For simplicity, we will assume always $T_1 = ... = T_d = T \to \infty$, but the extension to the case when all coordinates converge to ∞ at the same order of magnitude is immediate.

3. Asymptotic behavior of quadratic form estimators. The asymptotic behavior, in particular, the asymptotic normality of quadratic forms of stationary discrete time series, plays an important role in the choice of appropriate models. Let us recall, for example, the context of parametric estimation of Gaussian processes, where for an AR(1) process $X_{n+1} = \phi X_n + \epsilon_{n+1}$, the Yule-Walker estimator $\hat{\phi}$ of ϕ is given by

$$\hat{\phi} = \frac{\sum_{i=1}^T X_i X_{i-1}}{\sum_{i=1}^T X_i^2}$$

and thus the asymptotic behavior of these quadratic forms will be instrumental in providing confidence intervals for the estimation of $\hat{\phi}$.

Limit theory for quadratic forms is a subset of the more general one of providing limit theorems for sums/integrals and bilinear forms

$$S_T = \int_{t \in I_T} h(X_t)dt, \quad Q_T = \int_{t_1, t_2 \in I_T} \hat{b}(t_1 - t_2)\ h(X_{t_1}, X_{t_2})dt_1 dt_2 \tag{3}$$

where X_t is a stationary sequence. In the discrete time case, S_T and Q_T become respectively

$$S_T = \sum_{i=1}^T h(X_i), \quad Q_T = \sum_{i=1}^T \sum_{j=1}^T \hat{b}(i-j)\ h(X_i, X_j).$$

These topics, first studied by Dobrushin and Major [18] and Taqqu [37] in the Gaussian case, gave rise to very interesting non-Gaussian generalisations, and are still far from fully understood in the case of the continuous, "spatial" multidimensional indices arising in spatial statistics.

It is well-known in the context of discrete time series that the expansion in univariate/bivariate Appell polynomials determines the type of central limit theorem (CLT) or non-central limit theorem (NCLT) satisfied by the sums/quadratic forms (3). Hence, we consider below the problem (3) with h being an Appell polynomial.

4. The problem. In this paper we consider central limit theorems for quadratic forms

$$Q_T = Q_T^{(m,n)} = \int_{t,s\in I_T} \hat{b}(t-s) P_{m,n}(X_t, X_s) ds dt \tag{4}$$

involving the bivariate Appell polynomials

$$P_{m,n}(X_t, X_s) =: \underbrace{X_t, \ldots, X_t}_{m}, \underbrace{X_s, \ldots, X_s}_{n} : \;\; m, n \geq 0, m+n \geq 1,$$

which are defined via the Wick product $: X_1, \ldots, X_m :$ (see Appendix B).

For a warm-up, we consider also sums

$$S_T = S_T^{(m)} = \int_{t\in I_T} P_m(X_t) dt \tag{5}$$

involving the univariate Appell polynomials

$$P_m(X_t) =: \underbrace{X_t, \ldots, X_t}_{m} :$$

We will assume that $E|\xi_{I_1}|^{2(m+n)} < \infty$ in order to ensure that Q_T has a finite variance.

The variables X_t will be allowed to have short-range or long-range dependence (that is, with summable or non-summable sum of correlations), but the special short-range dependent case where the sum of correlations equals 0 will not be considered, since the tools described here are not sufficient in that case.

5. The method of cumulants. One of the classical approaches to establish asymptotic normality for processes having all moments, consists in computing all the scaled cumulants $\chi_{k,T}$ of the variables of interest, and in showing that they converge to those of a Gaussian distribution, that is, to 0 for $k \geq 3$.

For symmetric bilinear forms in stationary Gaussian discrete-time series X_t, with covariances $r_{i-j}, i, j \in \mathbb{Z}$, a direct computation yields the formula

$$\boxed{\chi_k = \chi(Q_T, ..., Q_T) = 2^{k-1}(k-1)! \text{ Tr}[(T_T(b)T_T(f))^k]} \tag{6}$$

where $T_T(b) = (\hat{b}_{i-j}, i, j = 1, ..., T)$, $T_T(f) = (r_{i-j}, i, j = 1, ..., T)$ denote Toeplitz matrices of dimension $T \times T$ and Tr denotes the trace. A limit theorem for products of Toeplitz matrices provided by Grenander and Szego(58) and strenghened by Avram(88) yields then the asymptotic normality, under the condition that $b(\lambda)f(\lambda) \in L_2$.

In the case of bilinear forms in Hermite/Appell polynomials $P_{m,n}(X,Y)$ of Gaussian/linear time series, or in continuous time, more complicated cumulants formulas arise from the so-called *diagram expansion*. Still these, can be expressed as sums/integrals involving a combinatorial structure similar to (6).

6. The spectral approach. We will assume throughout that all the existing cumulants of our stationary process X_t are expresses through Fourier transforms $c_k(t_1, t_2, ..., t_k)$ of "spectral densities" $f_k(\lambda_1, ..., \lambda_{k-1}) \in L_1$, i.e:

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

$$= \int_{\lambda_1,...,\lambda_k \in S} e^{i\sum_{j=1}^{k} \lambda_j t_j} f_k(\lambda_1, ..., \lambda_{k-1})\, \delta(\sum_{j=1}^{k} \lambda_j)\, \mu(d\lambda_1)...\mu(d\lambda_k)$$

where, throughout the paper, integrals involving delta functions will simply be used as a convenient notation for the corresponding integrals over lower dimensional subspaces. Throughout, S will denote the "spectral" space of discrete and continuous processes, i.e. $[-\pi, \pi]^d$ with Lebesgue measure normalized to unity, and $\mathbb{R}^d$ with Lebesgue measure, respectively.

Replacing now the time functions $c_k, \tilde{b}, \tilde{a}$ by their Fourier representations yields spectral integral representations for the cumulants $\chi_k(S_T), \chi_k(Q_T)$ corresponding to a particular graph (or matroid) structure.

7. Delta graph integrals. Let $G = (\mathcal{V}, \mathcal{E})$ denote a graph with V vertices, E edges and $co(G)$ components. Let M denote the $V \times E$ incidence matrix of the graph.

Definition 1 *The incidence matrix M of a graph has entries $M_{v,e} = \pm 1$ if the vertex v is the end/start point of the edge e, and 0 otherwise.*

Definition 2 (Delta graph integrals). *Suppose that associated to the edges of a graph $G = (\mathcal{V}, \mathcal{E})$ there is a set $f_e(\lambda),\ e = 1, \ldots, E$ of functions satisfying integrability conditions*

$$f_e \in L_{p_e}(\mu(d\lambda)),\ 1 \le p_e \le \infty,$$

where μ is normalized Lebesgue mesure on the torus $[-\pi, \pi]$ or Lebesgue mesure on $\mathbb{R}$.

A Delta graph integral is an integral of the form:

$$J_T = J_T(G, f_e, e = 1, ..., E) =$$

$$\boxed{\int_{\lambda_1,...,\lambda_E \in S} f_1(\lambda_1) f_2(\lambda_2)...f_E(\lambda_E) \prod_{v=1}^{V} \Delta_T(u_v) \prod_{e=1}^{E} \mu(d\lambda_e)} \tag{7}$$

where E, V and M denote respectively the number of edges, vertices, and the incidence matrix of the graph G, where

$$(u_1, ..., u_V)' = M(\lambda_1, ..., \lambda_E)'$$

and where a prime denotes a transpose. Finally,

$$\Delta_T(x) = \frac{\sin(Tx/2)}{x/2} \tag{8}$$

is the Fejer kernel.

This concept arose from the study of cumulants of sums/quadratic forms of stationary Gaussian processes. A simple computation based on the diagram formula [8] – see also Proposition 2– shows that the cumulants $\chi_k(S_T^{(m)})$ and $\chi_k(Q_T^{(m,n)})$ are sums of Delta graph integrals.

To obtain the central limit theorem by the method of cumulants, one wants then to show that $\lim_{T\to\infty} T^{-1}\chi_2(S_T^{(2)})$ is finite and that $\lim_{T\to\infty} T^{-k/2}\chi_k(S_T^{(2)}) = 0$ for $k \geq 3$.

To establish this one can use a "Szego-type" limit theorem for Delta graph integrals. Such a result (which extends the Grenander and Szego result for traces of Toeplitz matrices) was provided in Avram-Brown [6], in the discrete time case.

Quoting Tutte [38], it is probably true that "any theorem about graphs expressable in terms of edges and circuits exemplifies a more general result about matroids", 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 (thus, all matrices with the same rank function yield the same matroid). Indeed, Tutte's "conjecture" was true in this case; a matroid Szego-type limit theorem, in which the graph dependence structure is replaced with that of an arbitrary matroid, was given in Avram [8].

8. Delta matroid integrals. 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 $\mathcal{E}$. The most familiar matroids, associated to the set $\mathcal{E}$ of columns of a matrix and called *vector matroids*, may be specified uniquely 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). 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 excellent expositions on graphs and matroids, see [34], [35] and [39]. We ask the reader not familiar with this concept to consider only the particular case of **graphic matroids**, which are associated to the incidence matrix of an oriented graph. It turns out that the algebraic dependence structure translates in this case into graph-theoretic concepts, with circuits corresponding to cycles.

Here, we will only need to use the fact that to each matroid one may associate a dual matroid with rank function $r^*(A) = |A| - r(M) + r(M - A)$, and that in the case of graphic matroids, the dual coincides with the matroid associated with the $C \times E$ matrix M^* whose rows $c = 1, ..., C$ are obtained by assigning arbitrary orientations to the circuits (cycles) 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 [1].

Definition 3 *Let $f_e(\lambda)$, $e = 1, \ldots, E$ denote functions associated with the columns of M and satisfying integrability conditions*

$$f_e \in L_{p_e}(d\mu),\ 1 \leq p_e \leq \infty, \tag{9}$$

where μ is Lebesgue mesure on $\mathbb{R}$ or normalized Lebesgue mesure on the torus $[-\pi, \pi]$. Let M denote an arbitrary matrix in the first case, and with integer coefficients in the second case. Let $\hat{f}_e(k), k \in I$ denote the Fourier transform of $f_e(\lambda)$, i.e.

$$\hat{f}_e(k) = \int_S e^{ik\lambda} f_e(\lambda)\mu(d\lambda)$$

[1] It is enough to include in M^* a basis of cycles, thus excluding cycles which may be obtained via addition modulo 2 of other cycles, after ignoring the orientation.

A Delta matroid integral is defined by either one of the two equivalent expressions:

$$J_T = J_T(M, f_e, e=1,...,E) = \int_{j_1,...,j_m \in I_T} \hat{f}_1(i_1)\hat{f}_2(i_2)...\hat{f}_E(i_E) \prod_{v=1}^{V} dj_v \quad (10)$$

$$= \int_{\lambda_1,...,\lambda_E \in S} f_1(\lambda_1)f_2(\lambda_2)...f_E(\lambda_E) \prod_{v=1}^{V} \Delta_T(u_v) \prod_{e=1}^{E} \mu(d\lambda_e) \quad (11)$$

where $(i_1,...,i_E) = (j_1,...,j_V)M$ *and* $(u_1,...,u_V)' = M(\lambda_1,...,\lambda_E)'$ *and where, in the torus case, the linear combinations are computed modulo* $[-\pi, \pi]$.

Observe that (11) is the same expression as (7). A "Delta matroid integral" is also called a "(Delta) graph integral" when the matroid is associated to the incidence matrix M of a graph (graphic matroid).

9. The Szego-type limit theorem for Delta matroid integrals. Let

$$z_j = \frac{1}{p_j} \in [0,1], \quad j = 1, \cdots, E, \quad (12)$$

where p_j is defined in (9). Theorem 1 below, which is a summary of Theorems 1, 2 and Corollary 1 of [8], yields an upper bound and sometimes also the limit for Delta matroid integrals, in the case of discrete one-dimensional time series. The order of magnitude obtained is:

$$\boxed{\alpha_M(z) = V - r(M) + \max_{A \subset 1,...,E}[\sum_{j\in A} z_j - r^*(A)] = \max_{A \subset 1,...,E}[co(M-A) - \sum_{j \in A}(1-z_j)]} \quad (13)$$

where we define

$$co(M-A) = V - r(M-A). \quad (14)$$

Note: In the case of graph integrals, the function $co(M-A)$ represents the number of components, after the edges in A have been removed.

The function $\alpha_M(z)$ is thus found in the case of graph integrals by the following optimization problem:

<u>The "graph breaking" problem:</u> Find a set of edges whose removal maximizes the number of remaining components, with $\sum_{j\in A}(1-z_j) = \sum_{j \in A}(1-p_j^{-1})$ **as little as as possible.**

The function $\alpha_M(z)$ is then used in the following theorem which is a Szego-type limit theorem for Delta matroid integrals.

Theorem 1 *Let* $J_T = J_T(M, f_e, e=1,...,E)$ *denote a Delta 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:*

1.

$$\boxed{J_T(M, f_e, e=1,...,E) \le c_M T^{\alpha_M(z)}} \quad (15)$$

where c_M *is a constant and* $\alpha_M(z)$ *is given by (13).*

2. *If, moreover,* $\alpha_M(z) = V - r(M) = co(M)$ *(or, equivalently,* $\sum_{j\in A} z_j \le r^*(A), \forall A$*), then:*

$$\boxed{\lim_{T\to\infty} \frac{J_T(M)}{co(M)} = c_M \int_{S^C} f_1(\lambda_1) f_2(\lambda_2) ... f_E(\lambda_E) \prod_{c=1}^{C} \mu(dy_c) := J(M^*)} \tag{16}$$

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

3. *If a strict inequality* $\alpha_M(z) > co(M)$ *holds, then part a) may be strengthened to:*

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

Remark. The results of the theorem, that is, the expression of $\alpha_M(z)$ and the limit integral $J(M^*)$ depend on the matrix M only via the two equivalent rank functions $r(A), r^*(A)$, i.e. only via the matroid dependence structure between the columns.

10. Central limit theorems for variables whose cumulants are Delta matroid integrals. We draw now the attention to the convenient simplifications offered by these tools for establishing central limit theorems, cf. [6], [7], [8]. They arise from the fact that the cumulants are expressed as sums of integrals of the form (7) and their order of magnitude may be computed via the graph-optimization problem (13). Then a Szego-type limit theorem (Theorem 1) is used to conclude the proof. This shows that the central limit theorem can sometimes be reduced to a simple optimization problem.

We quote now Corollary 2 of [8].

Corollary 1 *Let* Z_T *be a sequence of zero mean random variables, whose cumulants of all orders are Delta matroid integrals:*

$$\chi_k(Z_T) = \sum_{G\in\mathcal{G}_k} J_T(G)$$

where $\mathcal{G}_k$ *are such that*

$$\alpha(G) \le k/2, \forall G \in \mathcal{G}_k$$

Then, a central limit theorem

$$\frac{Z_T}{\sigma T^{1/2}} \to N(0,1)$$

holds, with $\sigma^2 = \sum_{G\in\mathcal{G}_2} J(G^*)$*, where* $J(G^*)$ *is defined in the Theorem 1(b).*

This results reduce complicated central limit theorems for Gaussian processes to simple "graph breaking problems". For example, the result on bilinear forms of Avram(92), Theorem 4, and of Giraitis and Taqqu(97), Theorem 2.3, follow from the easily checked fact that the conditions of Corollary 1 hold in the extremal points of the polytope in Figure 1.

Remarks

1. The extremal points are solutions of equations $\alpha_G(z) = 1$, obtained for certain specific graphs $G \in \mathcal{G}_2$. The fact that at these points the inequalities $\alpha_G(z) \le k/2, \quad \forall G \in \mathcal{G}_k$ maybe checked via the graph-breaking problem described after equation (13). An example is presented at the end of the paper.

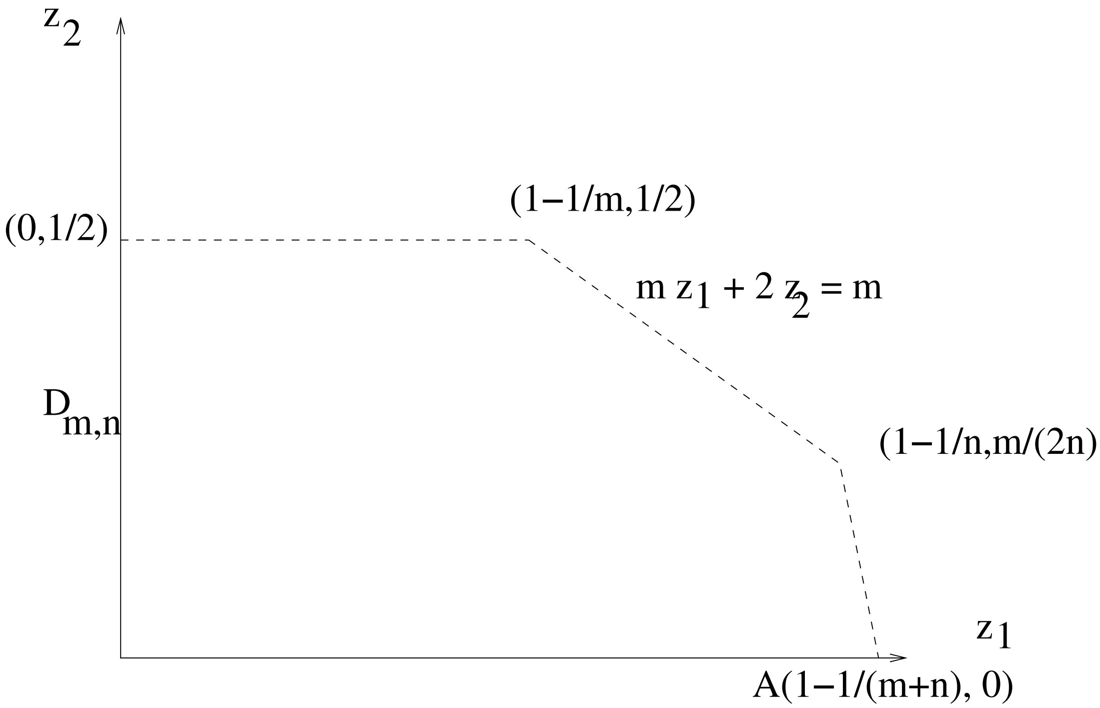

Figure 1: **The domain of the central limit theorem**

2. Giraitis and Taqqu expressed the upper boundary of this polytope in a convenient fashion:

$$d_m(z_1) + d_m(z_1) + 2z_2 = 1 \tag{17}$$

where $d_m(z) = 1 - m(1-z)^+$.

11. Sketch of the proof of Theorem 1. We sketch now the proof of the Szego-type result for Delta graph integrals given in [6], in the discrete one-dimensional setup, and for a connected graph. Note 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, the general formula (14) simplifies here to $r(M) = V-1$.

1. The first step for proving Theorem 1 is to establish that the measures

$$T^{-1}\Delta_T(-\sum_{v=1}^{V-1} u_v) \prod_{v=1}^{V-1} \Delta_T(u_v) \prod_{v=1}^{V-1} d\mu(u_v)$$

converge weakly to the measure $\delta_0(u_1, ..., u_{V-1})$ This convergence of measures holds since their Fourier coefficients converge –see [20], Lemma 7.1– and since the absolute variations of these measures are uniformly bounded, as may be seen by applying the corresponding Brascamp-Lieb-Barthe inequality (see Theorem 2 below) to the Delta graph integral, using estimates of the form

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

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

2. The main idea behind the proof of Theorem 1 is to integrate in (11) first over the complement of the space generated by the u_v's, $v = 1, \dots, V$. This is easier in the graph case, when, fixing some spanning tree $\mathcal{T}$ in the graph, we have a one to one correspondence between a set of independent cycles (with cardinality C) and the complementary set of edges $\mathcal{T}^c$. 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 $\tilde{M}$ is the first $V-1$ rows of M. Thus,

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

where the first rows are given by an identity matrix I_C completed by zeroes. Inverting this yields

$$(\lambda_1, ..., \lambda_E) = (y_1, ..., y_C, u_1, ..., u_{V-1}) \begin{pmatrix} M^* & | & N \end{pmatrix}$$

i.e. it turns out that the first columns of the inverse matrix are precisely the transpose of the dual matroid M^*. Note that $(\lambda_1, ..., \lambda_E)$ are linear functions of $(y_1, ..., y_C, u_1, ..., u_{V-1})$ such that when $u_1 = ... = u_{V-1} = 0$ the relation $(\lambda_1, ..., \lambda_E) = (y_1, ..., y_C)M^*$ is satisfied.

Definition 4 *The function*

$$J(u_1, ..., u_{r(M)}) = \int_{y_1, ..., y_C \in S} f_1(\lambda_1) f_2(\lambda_2) ... f_E(\lambda_E) \prod_{c=1}^{C} d\mu(y_c) \tag{18}$$

where λ_e are represented as linear combinations of $y_1, ..., y_C, u_1, ..., u_{V-1}$ via the linear transformation $(\lambda_1, ..., \lambda_E) = (y_1, ..., y_C, u_1, ..., u_{V-1}) \begin{pmatrix} M^* & | & N \end{pmatrix}$ defined above will be called a **graph convolution**.

Note: A key point in the discrete one-dimensional setup was then to show that under appropriate L_p conditions, the Brascamp-Lieb-Barthe inequality (GH) (see Theorem 2 below) ensures the continuity of the graph convolution functions $J(u_1, ..., u_{r(M)})$ in the variables $(u_1, ..., u_{r(M)})$. In the spatial statistics context [3], [4], this continuity was usually assumed and indeed checking if this assumption may be relaxed to L_p integrability conditions in the spectral domain is one of the outstanding difficulties for the spatial extension.

Recall now that $r(M) = V - 1$. The change to the variables $y_1, ..., y_C, u_1, ..., u_{V-1}$ and integration over $y_1, ..., y_C$ transforms the Delta graph integral into

$$\int_{u_1, ..., u_{V-1} \in S} J(u_1, ..., u_{V-1}) \prod_{v=1}^{V} \Delta_T(u_v) \prod_{v=1}^{V-1} d\mu(u_v)$$

Finally, the convergence of the Fejer kernel to a δ function implies the convergence of the scaled Delta graph integral $J_T(M, f_e, e = 1..., E)$ to $J(0, ..., 0) = J(M^*, f_e, e = 1..., E)$, establishing Part 2 of the theorem.

Remark. It is not difficult to extend this change of variables to the case of several components and then to the matroid setup. In the first case, one would need to choose 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

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

12. Conclusions. The problem of establishing the central limit theorem by the method of moments is related to some beautiful mathematics: the Brascamp-Lieb-Barthe inequality, the continuity of matroid convolutions and the matroid weak Szego theorem.

This leads to fascinating mathematical questions like strengthening the matroid weak Szego theorem to a strong one (i.e. providing correction terms).

The analytic methodology presented above suggests also the following conjecture:

Conjecture. A central limit theorem holds in the **continuous one-dimensional index case**, with the same normalization and limiting variance as in the discrete one-dimensional index case, if $f \in L_{z_1^{-1}}, \hat{b} \in L_{z_2^{-1}}$, where the exponents z_1, z_2 lie on the upper boundary of the polytope in the Figure 1.

These tools are also expected to be useful for studying processes with continuous multidimensional indices. Let us mention for example the versatile class of **isotropic** spatio-temporal models, of a form similar to (1) (with the Laplacian operator Δ replacing the operator B), recently introduced by Anh, Leonenko, Kelbert, McVinnish, Ruiz-Medina, Sakhno and coauthors [1], [2], [3],[4], [32]. These authors use the spectral approach as well and the tools described above hold the potential of simplifying their methods.

Finally, even for unidimensional discrete processes, these tools might be useful for strengthening the central limit theorem to sharp large deviations statements, as in the work on one-dimensional Gaussian quadratic forms of Bercu, Gamboa, Lavielle and Rouault [13], [14].

2 Appendix A: the Brascamp-Lieb-Barthe inequality

Set here $V = m$ and $E = n$. Let M be a $m \times n$ matrix, $\mathbf{x} = (x_1, \ldots, x_m)$ and let $l_1(\mathbf{x}), \ldots l_n(\mathbf{x})$ be n linear transformations such that

$$(l_1(\mathbf{x}), \ldots, l_n(\mathbf{x})) = (x_1, \ldots, x_m)M.$$

Let $f_j,\ j = 1, \ldots, n$ be functions belonging respectively to

$$L_{p_j}(d\mu),\ 1 \leq p_j \leq \infty,\ j = 1, \ldots, n.$$

We consider simultaneously three cases:

(C1) $\mu(dx)$ is the Lebesgue measure on the torus $[-\pi, \pi]^{n_j}$, and M has all its coefficients integers.

(C2) $\mu(dx)$ is the counting mesure on $\mathbb{Z}^{n_j}$, M has all its coefficients integers, and all its non-singular minors of dimension $m \times m$ have determinant ± 1.

(C3) $\mu(dx)$ is Lebesgue measure on $(-\infty, +\infty)^{n_j}$.

The following theorem, due when $n_j = 1, \forall j$ in the first case to [6], in the second to [7] and in the last to [10], with arbitrary n_j yields conditions on

$$z_j = \frac{1}{p_j},\ j = 1, \ldots, n,$$

so that a generalized Hölder inequality holds. The key idea of the proof in [6], [7], is that it is enough to find those points z= $(z_1,\ldots,z_n)$ with coordinates z_i equal to 0 or 1, for which the inequality (GH) below holds, then by the Riesz-Thorin interpolation theorem, (GH) will hold for the smallest convex set generated by these points. This yields:

Theorem 2 (Brascamp-Lieb-Barthe inequality). *Suppose, respectively, that the conditions (C1), (C2) and (C3) hold and let f_j, $j = 1,\ldots,n$ be functions $f_j \in L_{p_j}(\mu(dx))$, $1 \le p_j \le \infty$, where the integration space is either $[-\pi,\pi]^{n_j}$, $\mathbb{Z}^{n_j}$, or $\mathbb{R}^{n_j}$, and $\mu(dx)$ is respectively normalized Lebesgue measure, counting measure and Lebesgue measure. Let $z_j = (p_j)^{-1}$.*

For every subset A of columns of M (including the empty set $\emptyset$), denote by $r(A)$ the rank of the matrix formed by these columns, and suppose respectively:

(c1) $\sum_{j\in A} z_j \le r(A),\ \forall A$

(c2) $\sum_{j\in A} z_j + r(A^c) \ge r(M),\ \forall A$

(c3) $\sum_{j=1}^{T} z_j = m$, *and one of the conditions (c1) or (c2) is satisfied.*

Then, the following inequality holds:

$$(GH) \qquad \left| \int \prod_{j=1}^{T} f_j(l_j(\mathbf{x})) \prod_{i=1}^{m} d\mu(x_i) \right| \le K \prod_{j=1}^{T} \|f_j\|_{p_j}$$

where the constant K in (GH) is equal to 1 in the cases (C1) and (C2) and is finite in the case (C3) (and given by the supremum over centered Gaussian functions – see [10]).

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

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

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

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

Examples:

1) As an example of $(c1)$, consider the integral

$$J = \int_T \int_T f_1(x_1) f_2(x_2) f_3(x_1+x_2) f_4(x_1-x_2)\, dx_1 dx_2$$

where T denotes the torus $[0,1]$, so $f_j(x\pm 1) = f_j(x),\ j = 1,\ldots,4$. Here $m = 2,\ n = 4$ and the matrix

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

has rank $r(M) = 2$. The flats consist of $\emptyset$, the single columns and M. Only $\emptyset$ and M are flat, irreducible, and not singleton. Since $(a1')$ always holds for $\emptyset$, it is sufficient to apply it to M. The theorem yields

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

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

$$|J| \leq \left(\sup_{0 \leq x \leq 1} |f_1(x)| \right) \left(\int_0^1 |f_2(x)| dx \right) \left(\int_0^1 f_3^4(x) dx \right)^{1/4} \left(\int_0^1 f_4^2(x) dx \right)^{1/2}.$$

2. To illustrate $(c2)$, consider

$$S = \sum_{x_1=-\infty}^{\infty} \sum_{x_2=-\infty}^{\infty} f_1(x_1) f_2(x_2) f_3(x_1 + x_2) f_4(x_1 - x_2).$$

Since m, n and M are as in Example 1, we have $r(M) = m$ and the only flat and irreducible sets which are not singleton are $\emptyset$ and M. Since it is sufficient to apply $(c2)$ to $\emptyset$, the theorem yields $|S| \leq \|f_1\|_{1/z_1} \|f_2\|_{1/z_2} \|f_3\|_{1/z_3} \|f_4\|_{1/z_4}$ for any $z = (z_1, z_2, z_3, z_4) \in [0,1]^4$ satisfying $z_1 + z_2 + z_3 + z_4 \geq 2$, e.g. $|S| \leq \prod_{j=1}^4 \left(\sum_{x=-\infty}^{+\infty} f_j^2(x) \right)^{1/2}$.

3 Appendix B: the application of the diagram formula for computing moments/cumulants of Wick products of linear processes

1. Wick products. We start with some properties of the Wick products (cf. [24], [36]) and their application in our problem.

Definition 5 *The Wick products (also called Wick powers) are multivariate polynomials:*

$$: y_1, \ldots, y_n :^{(\nu)} = \frac{\partial^n}{\partial z_1 \ldots \partial z_n} \left[\exp(\sum_1^n z_j y_j) / \int_{\mathbb{R}^n} \exp(\sum_1^n z_j y_j) d\nu(y) \right] \Bigg|_{z_1 = \cdots = z_n = 0}$$

corresponding to a probability measure ν on $\mathbb{R}^n$. Interpret this as a formal expression if ν does not have a moment generating function, the Wick products being then obtained by formal differentiation.

A sufficient condition for the Wick products $: y_1, \ldots, y_n :^{(\nu)}$ to exist is $E|Y_i|^n < \infty$, $i = 1, \ldots, n$.

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 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; like them, they play an important role in the limit theory of quadratic forms of dependent variables (cf. [36], [24], [5]).

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 (4) is a weighted periodogram with its expectation removed.

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

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

be the ordinary product,

$$: Y^W :$$

the Wick product, and

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

be the (mixed) cumulant of the variables $Y_i, i \in W$, respectively, which is definedk as follows:

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

The following relations hold ([36], Prop. 1):

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

$$Y^W = \sum_{U \subset W} : Y^U : \sum_{\{V\}} \chi(Y^{V_1}) \cdots \chi(Y^{V_r}) = \sum_{U \subset W} : Y^U : E(Y^{W \backslash U})$$

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

2. 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 [33]. 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:

$$\begin{pmatrix} (1,1), \ldots, (1, n_1) \\ \ldots\ldots\ldots \\ (k,1), \ldots, (k, n_k) \end{pmatrix} = W. \tag{19}$$

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 $|V_1| = \cdots = |V_r| = 2$. 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

$$I_\gamma = \prod_{j=1}^{r} \chi(Y^{V_j}). \tag{20}$$

Proposition 1 (cf. [24], [36]) *Each of the numbers*

$$\begin{array}{ll} (i) & EY^W = E(Y^{W_1} \ldots Y^{W_k}), \\ (ii) & E(: Y^{W_1} : \cdots : Y^{W_k} :), \\ (iii) & \chi(Y^{W_1}, \ldots, Y^{W_k}), \\ (iv) & \chi(: Y^{W_1} :, \ldots, : Y^{W_k} :) \end{array}$$

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.

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

3. Multilinearity. An important property of Wick products and of cumulants is their multilinearity. This implies that for Q_T defined in (4) that

$$\chi_k(Q_T, ..., Q_T) =$$
$$\int_{t_i, s_i \in I_T} \chi(: X_{t_{1,1}}, \ldots, X_{t_{1,m}},\ X_{s_{1,1}}, \ldots, X_{s_{1,n}} :, \ldots, : X_{t_{k,1}}, \ldots, X_{t_{k,m}},\ X_{s_{k,1}}, \ldots, X_{s_{k,n}} :) \prod_{i=1}^{k} \hat{b}_{t_i - s_i} dt_i ds_i$$

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}.

A further application of part (iv) of Proposition 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.

Example 1: 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 (6) for the cumulants of discrete Gaussian bilinear forms.

Example 2: 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.

4. The cumulants of sums and quadratic forms of linear processes. Consider

$$S_T^m = \sum_{i=1}^{T} P_m(X_{t_i}), \quad Q_T^{m,n} = \sum_{i=1}^{T}\sum_{j=1}^{T} \hat{b}(i-j)\, P_{m,n}(X_{t_i}, X_{t_j}). \tag{21}$$

By part (iv) of proposition 1, applied to a table W of k rows $R_1 ...., R_k$, with $K = n_1 + ...n_k$ variables, and by the definition (20) and of $I\gamma$, we find the following formula for the cumulants of the Wick products of linear variables (2):

$$\chi(: X_{t_{1,1}}, \ldots, X_{t_{1,n_1}} :, \ldots, : X_{t_{k,1}}, \ldots, X_{t_{k,n_k}} :) = \sum_{\gamma \in \Gamma(n_1,\ldots,n_k)} \kappa_\gamma J_\gamma(\vec{t}) \tag{22}$$

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

$$J_\gamma(t_1, ..., t_K) = \prod_{j=1}^{r} J_{V_j}(t_{V_j}) \tag{23}$$

$$= \int_{s_1,\ldots,s_r \in I} \prod_{j=1}^{k} [\hat{a}(t_{j,1} - s_{j,1}) \ldots \hat{a}(t_{j,n_1} - s_{j,n_1}) \ldots \hat{a}(t_{k,1} - s_{k,1}) \ldots \hat{a}(t_{k,n_k} - s_{k,n_k})]\, ds_1, \ldots, ds_r$$

$$= \int_{\lambda_1,\ldots,\lambda_K} e^{i \sum_{j=1}^{K} t_j\, \lambda_j} \prod_{i=1}^{K} a(\lambda_i) \prod_{j=1}^{r} \delta(\sum_{i \in V_j} \lambda_i) \prod_{i=1}^{K} d\lambda_i \tag{24}$$

where $s_{i,j} \equiv s_l$ if $(i, j) \in V_l$, $l = 1, \ldots, r$.

We will apply now this formula to compute the cumulants of (21), in which case each row j contains just one, respectively two random variables. We will see below that this yields decompositions as sums of Delta graph integrals with a specific graph structure.

For example, 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 two possible diagrams of 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.

The general structure of the graphs that we get is as follows (see [8]):

1. In the case of cumulants of sums, we get graphs belonging to the set $\Gamma(m,k)$ of all connected graphs with no loops over k vertices, each of degree m.

2. In the case of cumulants of quadratic forms, we get – see Figure 2 – 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 : $X_{t_1},...,X_{t_m}$: in the diagram formula, and the "right" vertex of each pair arises out of the last n terms : $X_{s_1},...,X_{s_n}$: The edge set consists of:

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

   (b) 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.

   Thus, the k "left vertices" are of degree $m+1$, and the other k vertices are of degree $n+1$. (The "costs" mentionned in Figure 2 refer to (13)).

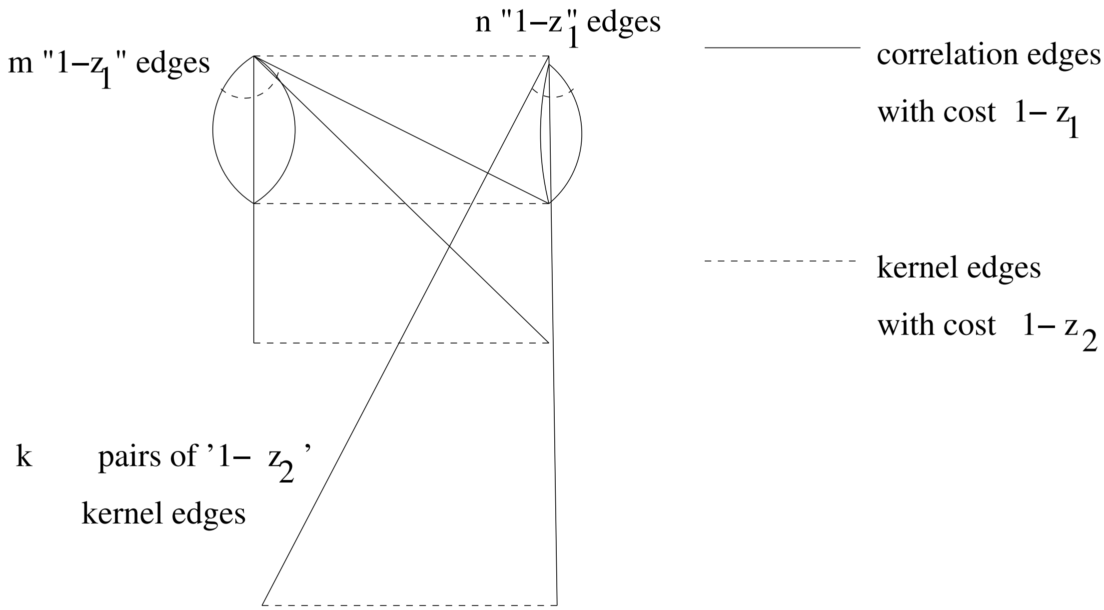

Figure 2: **The graphs appearing in the expansion of cumulants of quadratic forms. Here k=4, m=5, n=4. The figure displays only some of the k(m+n)/2=18 correlation edges.**

The following proposition is easy to check:

Proposition 2 *Let $X(t), t \in I_T$ denote a stationary Gaussian process. Then, the cumulants of the sums and quadratic forms defined in (21) are given respectively by:*

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

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 $\Delta_T(x) = \frac{\sin(Tx/2)}{x/2}$ *is the Fejer kernel,* $\Gamma(m,k)$*,* $\Gamma(m,n,k)$ *were defined above, and*

$$\sigma_\gamma(T) = \int_{\vec{t} \in I_T^k} J_\gamma(\vec{t}) dt = \int_{\lambda_1,...,\lambda_K} \prod_{j=1}^{k} \Delta_T(\sum_{i=m(j-1)+1}^{mj} \lambda_i) \prod_{i=1}^{K} a(\lambda_i) \prod_{j=1}^{r} \delta(\sum_{i \in V_j} \lambda_i) \prod_{i=1}^{K} d\lambda_i \quad (25)$$

$$\tau_\gamma(T) = \int_{\mu_1,...,\mu_k,\lambda_1,...,\lambda_K,\lambda'_1,...,\lambda'_{K'}} \prod_{j=1}^{k} \left[\Delta_T(\mu_j + \sum_{i=m(j-1)+1}^{mj} \lambda_i) \ \Delta_T(-\mu_j + \sum_{i=n(j-1)+1}^{nj} \lambda'_i) \ b(\mu_j) \right]$$
$$\times \prod_{i=1}^{K} a(\lambda_i) \prod_{i=1}^{K'} a(\lambda'_i) \prod_{j=1}^{r} \delta(\sum_{i \in V_j} \lambda_i + \sum_{i \in V_j} \lambda'_i) \prod_{i=1}^{K} d\lambda_i \prod_{i=1}^{K'} d\lambda'_i \prod_{i=1}^{k} d\mu_i. \quad (26)$$

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 [8].

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

$$\alpha_G(z_1) = \max_A p(A) = \max_A [co(G - A) - \sum_{e \in A} (1 - z_e)] = \max_A [co(G - A) - |A|(1 - z_1)] \leq k/2, \forall G \in \mathcal{G}_k \quad (27)$$

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 (27)) 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 $p(A)$.

References

[1] Anh V V, Leonenko N N and McVinish R (2001), Models for fractional Riesz-Bessel motion and related processes, Fractals, 9, N3, 329-346.

[2] 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), no. 4, 730-747.

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

[4] V. V. Anh, N. N. Leonenko, and L. M. Sakhno. Higher-order spectral densities of fractional random fields (2003). Journal of Statistical Physics, 111, nr. 3 - 4, pp 789 - 814.

[5] F. Avram and M. S. Taqqu. Noncentral limit theorems and Appell polynomials. *The Annals of Probability*, 15:767–775, 1987.

[6] F. Avram and Brown, L. (1989). A Generalized Holder Inequality and a Generalized Szego Theorem. Proceedings of Amer. Math. Soc., **107**, 687-695.

[7] F. Avram and Taqqu, M. (1989a). Holder's Inequality for Functions of Linearly Dependent Arguments. Soc. Ind. Appl. Math. J. of Math Anal., **20**, 1484-1489.

[8] F. Avram (1992). Generalized Szego Theorems and asymptotics of cumulants by graphical methods. Transactions of Amer. Math. Soc., **330**, 637-649.

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

[10] F. Barthe. On a reverse form of the Brascamp-Lieb inequality. Invent. Math. 134 (1998) 335-361.

[11] R. Bentkus. On the error of the estimate of the spectral function of a stationary process, Liet. Mat. Rink. 12, N 1, 1972a, p.55-71.(In Russian)

[12] J. Beran. *Statistics for Long-Memory Processes.* Chapman Hall, New York, 1994.

[13] B. Bercu, F. Gamboa and M. Lavielle, Sharp large deviations for Gaussian quadratic forms with applications. ESAIM PS, 4, pp. 1-24, 2000.

[14] B. Bercu, F. Gamboa and A. Rouault, Large deviations for quadratic forms of stationary Gaussian processes. Stochastic Processes and their Applications, 71, pp. 75-90, 1997.

[15] Chambers, M.J. (1996). The estimation of continuous parameter long- memory time series models. Economic Theory, 12, 373-390.

[16] Christakos, G. (2000). Modern Spatiotemporal Geostatistics. Oxford University Press, Oxford.

[17] R. B. Davies and D. S. Harte. Tests for Hurst effect. *Biometrika*, 74:95–101, 1987.

[18] Dobrushin, R. L. & Major, P. (1979), 'Non-central limit theorems for non-linear functions of Gaussian fields', *Zeitschrift für Wahrscheinlichkeitstheorie und verwandte Gebiete* **50**, 27–52.

[19] R. Fox and M. S. Taqqu. Large-sample properties of parameter estimates for strongly dependent stationary Gaussian time series. The Annals of Statistics, 14(2):517-532, 1986.

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

[21] Gay, R. and Heyde, C.C. (1990). On a class of random field models which allow long range dependence. Biometrika, 77, 401-403.

[22] Granger, C.W. and Joyeux, (1980). An introduction to long-memory time series models and fractional differencing. J. Time Series Anal., 10, 233- 257.

[23] L. Giraitis and M. S. Taqqu. Whittle estimator for finite variance non-Gaussian time series with long memory. Ann.Statist. 27(1):178-203, 1999.

[24] L. Giraitis and D. Surgailis. Multivariate Appell polynomials and the central limit theorem. In E. Eberlein and M. S. Taqqu, editors, *Dependence in Probability and Statistics*, pages 21–71. Birkhäuser, New York, 1986.

[25] X. Guyon. Random Fields on a Network: Modelling, Statistics and Applications. Springer-Verlag, New York, 1995.

[26] E. J. Hannan. Multiple Time Series. Springer-Verlag, New York, 1970.

[27] C. Heyde and R. Gay. On asymptotic quasi-likelihood. Stoch. Proc. Appl., 31:223-236, 1989.

[28] C. Heyde and R. Gay. Smoothed periodogram asymptotics and estimation for processes and fields with possible long-range dependence. Stoch. Proc. Appl., 45:169-182, 1993.

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

[30] Hosking, J.R.M. (1981). Fractional differencing. Biometrika, 68, 165-176.

[31] H. E. Hurst. Long-term storage capacity of reservoirs. *Transactions of the American Society of Civil Engineers*, 116:770–808, 1951.

[32] Kelbert M, Leonenko N N and Ruiz-Medina, M D (2005) Fractional random fields associated with stochastic fractional heat equation, Advanced of Applied Probability, 37, to appear.

[33] V. A. Malyshev. Cluster expansions in lattice models of statistical physics and the quantum theory of fields. *Russian Mathematical Surveys*, 35(2):1–62, 1980.

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

[35] J.G. Oxley, *What is a matroid?*. Preprint, www.math.lsu.edu/ oxley/survey4.pdf

[36] D. Surgailis. On Poisson multiple stochastic integral and associated equilibrium Markov process. In *Theory and Applications of Random Fields*, pages 233–238. Springer-Verlag, Berlin, 1983. In: *Lecture Notes in Control and Information Science*, Vol. **49**.

[37] Taqqu, M. S. (1979) 'Convergence of integrated processes of arbitrary Hermite rank', *Zeitschrift für Wahrscheinlichkeitstheorie und verwandte Gebiete* **50**, 53–83.

[38] W.T. Tutte , Matroids and graphs, Trans. A.M.S. 90 (1959), 527-552.

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

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

[41] P. Whittle. Estimation and information in stationary time series. Ark. Math. 2:423-434, 1953.

[42] P. Whittle (1963). Stochastic processes in several dimensions. Bull. Inst. Internat. Statist., 40, 974-994.