



**HAL**  
open science

# A Normality Test for Multivariate Dependent Samples

Sara Elbouch, Olivier Michel, Pierre Comon

► **To cite this version:**

Sara Elbouch, Olivier Michel, Pierre Comon. A Normality Test for Multivariate Dependent Samples. 2021. hal-03344745v1

**HAL Id: hal-03344745**

**<https://hal.science/hal-03344745v1>**

Preprint submitted on 17 Sep 2021 (v1), last revised 1 Mar 2022 (v2)

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# A Normality Test for Multivariate Dependent Samples

Sara ElBouch and Olivier Michel and Pierre Comon

*<sup>a</sup>Univ. Grenoble Alpes, CNRS, Grenoble INP, GIPSA-Lab, Grenoble Campus,  
BP.46, Grenoble, 38000, France*

---

## Abstract

Most normality tests in the literature are performed for scalar and independent samples. Thus, they become unreliable when applied to colored processes, hampering their use in realistic scenarios. We focus on Mardia's multivariate kurtosis, derive closed-form expressions of its asymptotic distribution for statistically dependent samples, under the null hypothesis of normality. Included experiments illustrate, by means of copulas, that it does not suffice to test a one-dimensional marginal to conclude normality. The proposed test also exhibits good properties on other typical scenarios, such as the detection of a non-Gaussian process in the presence of an additive Gaussian noise.

*Keywords:* Multivariate normality test, kurtosis, colored process, copula

---

## 1. Introduction

In recent decades, techniques using high-order statistics (HOS) (Nikias and Petropulu, 1993; Haykin, 2000; Cichocki and Amari, 2002; Comon and Jutten, 2010) have grown considerably. The reason is that some problems were not solved by using simple techniques based on first and second order statistics only. Therefore, it is important to know whether or not HOS are providing information in a given data set. On the other hand, normality tests are also of interest in their own right; they can, for example, be used to detect abrupt changes in dynamical systems (Basseville and Nikiforov, 1993).

Let  $\mathbf{x}(n) = [x_1(n), \dots, x_p(n)]^T$  be a real  $p$ -variate stochastic process. We assume that a sample of finite size  $N$  is observed,  $1 \leq n \leq N$ . Our goal is to implement the following test *without alternative*, in such a way that it can be executed in real-time, e.g. over a sliding window.

**Problem P1:** Given a finite sample of size  $N$ ,  $\mathbf{X} \stackrel{def}{=} \{\mathbf{x}(1), \dots, \mathbf{x}(N)\}$ :

$$H : \mathbf{X} \text{ is Gaussian } \textit{versus} \bar{H} \quad (1)$$

where variables  $\mathbf{x}(n) \in \mathbb{R}^p$  are identically distributed, but not statistically independent.

This well-known problem is twofold: (i) define a test variable, and (ii) determine its asymptotic distribution (often itself normal) in order to assess the power of the test, that is, the probability to decide  $\bar{H}$  whereas  $H$  is true.

One should distinguish between *scalar* and *multivariate* tests, the latter addressing joint normality of several variables. Since the so-called Chi-squared test proposed by Fisher and improved in (Moore, 1971), the most popular scalar test is probably the omnibus test based on skewness and kurtosis by (Bowman and Shenton, 1975). The omnibus test proposed by (Bowman and Shenton, 1975) combines estimated skewness  $\hat{b}_1$  and kurtosis  $\hat{b}_2$  weighted by the inverse of their respective asymptotic variance:

$$SK = \frac{N}{6} \hat{b}_1^2 + \frac{N}{24} (\hat{b}_2 - 3)^2 \quad (2)$$

where  $\hat{b}_1 = \hat{m}_3 / \hat{m}_2^{3/2}$ ,  $\hat{b}_2 = \hat{m}_4 / \hat{m}_2^2$ ,  $\hat{\mu}$  is the sample mean and  $\hat{m}_i = \frac{1}{N} \sum_{n=1}^N (x(n) - \hat{\mu})^i$ , for  $i > 1$ . The asymptotic variance of  $\hat{b}_1$  and  $\hat{b}_2$  is indeed  $6/N$  and  $24/N$  under the assumption that samples  $x(n)$  are independently and identically distributed (i.i.d.) and normal; see (Mardia, 1974), (Kotz and Johnson, 1982). The asymptotic distribution of the test is  $\chi_2^2$  when samples are i.i.d normal. However, as pointed out by (Moore, 1982), the Chi-square test is very sensitive to the dependence between samples; the process color yields a loss in apparent normality (Gasser, 1975).

Actually, most of the tests proposed in the literature assume that observations  $\mathbf{x}(n)$  are i.i.d., see (Shapiro et al., 1968) or (Pearson et al., 1977). This is also true for multivariate tests (Mardia, 1970; Andrews et al., 1973); see the survey of (Henze, 2002).

Even if samples are often correlated in practice, few tests are dedicated to colored processes. For instance, the linearity test of (Hinich, 1982) can serve as a normality test; it is indeed based on the bispectrum, is constant if the process is linear, and that constant is null in the Gaussian case. One could

build a similar test based on the trispectrum, since estimated multispectra of higher order are also asymptotically normal (Brillinger, 1981).

In practice, nonlinear functions applied to  $\mathbf{x}(n)$  can go beyond monomials of degree 3 or 4 (Moulines et al., 1992). For instance, some tests are based on the characteristic function (Epps, 1987; Moulines et al., 1992) and others on entropy (Steinberg and Zeitouni, 1992). These tests are complex to implement in practice.

Except tests based on arbitrary 1D projections (Mardia, 1970; Malkovich and Afifi, 1973; Nieto-Reyes et al., 2014), which we shall discuss later, all the tests we have reviewed above are hardly executable in real-time on a light processor, as soon as they are valid for statistically dependent samples. For this reason, we shall focus on the multivariate kurtosis proposed in (Mardia, 1970), and derive its mean and variance when samples are assumed to be statistically dependent.

When deriving theoretical properties in the remainder, it is supposed that  $\mathbf{x}(n)$  is a zero-mean stationary process, with finite moments up to order 16. Its covariance matrix function is denoted by

$$\mathbf{S}(\tau) = \mathbb{E}\{\mathbf{x}(n)\mathbf{x}(n - \tau)^T\}. \quad (3)$$

For the sake of conciseness,  $\mathbf{S}(0)$  will be merely denoted by  $\mathbf{S}$ . In addition, we assume the following mixing condition upon  $\mathbf{x}(n)$ :  $\sum_{\tau=0}^{\infty} |S_{ab}(\tau)|^2$  converges to a finite limit  $\Omega_{ab}$ ,  $\forall(a, b) \in \{1, \dots, p\}^2$ , where  $S_{ab}$  denote the entries of matrix  $\mathbf{S}$ .

*Contribution.* Our main contributions are the following. We provide a *Multivariate* test for Gaussianity, which can be implemented in real-time, as most of the conventional ones are univariate. We could use univariate tests on each of the  $p$  components, but this would not test for *joint normality* and can lead to misdetections; this fact is subsequently illustrated with copula. A general procedure is provided to compute the asymptotic mean and variance of Mardia’s Multivariate Kurtosis when samples are *statistically dependent*. Then we provide the complete expressions of mean and variance of the test variable in the general case when  $\mathbf{x}(n)$  is of dimension  $p = 2$ , which allows to test joint normality if two arbitrary projections are performed in a first stage, in the same spirit as done in (Malkovich and Afifi, 1973) in the i.i.d. case. These results are summarized in Section 7. Additionally, the particular case when  $\mathbf{x}(n)$  has the form  $\mathbf{x}(n)^T = [y(n\delta + 1), y(n\delta + 2), \dots, y(n\delta + p)]$  is addressed, where  $y(n)$  is a scalar colored process.

This article is organized as follows. Section 2 contains the definition of the test statistic, followed by Section 3 where necessary tools are introduced to conduct the calculations. The moments involved in deriving both the mean and variance of the test statistic are given in sections 4-5, then their exact expressions for various cases are given in sections 6-8. Section 9 reports some computer experiments. We defer the expressions of the moments and details about the computation to appendices in Section 11.

## 2. Multivariate kurtosis

The test proposed in (Mardia, 1970) takes the form:

$$\beta_p = \mathbb{E}\{(\mathbf{x}^T \mathbf{S}^{-1} \mathbf{x})^2\}. \quad (4)$$

For  $\mathbf{x} \sim \mathcal{N}_p(0, \mathbf{S})$ , one can show that  $\beta_p = p(p+2)$ . Its sample counterpart for a sample of size  $N$  is:

$$B_p(N) = \frac{1}{N} \sum_{n=1}^N (\mathbf{x}(n)^T \mathbf{S}^{-1} \mathbf{x}(n))^2 \quad (5)$$

One advantage of this test variable is that it is invariant with respect to linear transformations, i.e.,  $\mathbf{y} = \mathbf{A}\mathbf{x}$ . In practice, the covariance matrix  $\mathbf{S}$  is unknown and is replaced by its sample estimate,  $\hat{\mathbf{S}}$ , so that we end up with the following test variable:

$$\hat{B}_p(N) = \frac{1}{N} \sum_{n=1}^N (\mathbf{x}(n)^T \hat{\mathbf{S}}^{-1} \mathbf{x}(n))^2 \quad (6)$$

with

$$\hat{\mathbf{S}} = \frac{1}{N} \sum_{k=1}^N \mathbf{x}(k) \mathbf{x}(k)^T.$$

The multivariate normality test can be formulated in terms of the multivariate Kurtosis: the variable  $\mathbf{x}$  is said to be normal if  $|\hat{B}_p(N) - \mathbb{E}\{\hat{B}_p(N)|H_0\}| \leq \eta$ , where  $\eta$  is a threshold to be determined as a function of the power of the test. The fact that  $\hat{B}_p(N)$  is a good estimate of  $\beta_p$  or not is relevant; what is important is to have a sufficiently accurate estimation of the power of the test. In order to do that, we need to assess the mean and variance of  $\hat{B}_p(N)$  under  $H_0$ . Under the assumption that  $\mathbf{x}(n)$  are i.i.d. realizations of variable  $\mathbf{x}$ , the mean and variance of  $\hat{B}_p(N)$  have been calculated:

**Theorem 2.1.** (Mardia, 1970) Let  $\{\mathbf{x}(n)\}_{1 \leq n \leq N}$  be i.i.d. of dimension  $p$ . Then under the null hypothesis  $H_0$ ,  $\hat{B}_p(N)$  is asymptotically normal, with mean  $p(p+2)\frac{N-1}{N+1}$  and variance  $\frac{8p(p+2)}{N} + o(\frac{1}{N})$ .

Our purpose is now to state a similar theorem when  $\{\mathbf{x}(n)\}$  are not independent. Since this involves heavy calculations, we need to introduce some tools to make them possible.

### 3. Statistical and combinatorial tools

#### 3.1. Lemmas

The estimated multivariate kurtosis (6) is a rational function of degree 4. Since we wish to calculate its asymptotic first and second order moments, when  $N$  tends to infinity, we may expand this rational function about its mean. The first step is to expand the estimated covariance  $\hat{\mathbf{S}}$ . Let  $\hat{\mathbf{S}} = \mathbf{S} + \mathbf{\Delta}$ , where  $\mathbf{\Delta}$  is small compared to  $\mathbf{S}$ ; in fact :

**Lemma 3.1.** *The entries of matrix  $\mathbf{\Delta}$  are of order  $O(1/\sqrt{N})$ .*

*Proof.* Under Hypothesis  $H$ , the covariance of entries  $\Delta_{ab}$  take the form below :

$$Cov(\mathbf{\Delta}_{ab}, \mathbf{\Delta}_{cd}) = \frac{1}{N^2} \sum_{n=1}^N \sum_{m=1}^N \mathbb{E}\{x_a(n)x_b(n)x_c(m)x_d(m)\} - S_{ab}S_{cd}$$

and letting  $\tau = n - m$ , and  $\Omega_{abcd} = S_{ac}S_{bd} + S_{ad}S_{bc}$  we have after some manipulation:

$$\begin{aligned} Cov(\mathbf{\Delta}_{ab}, \mathbf{\Delta}_{cd}) &= \frac{1}{N} \Omega_{abcd} + \frac{2}{N} \sum_{\tau=1}^{N-1} \left(1 - \frac{\tau}{N}\right) \{S_{ac}(\tau)S_{bd}(\tau) + S_{ad}(\tau)S_{bc}(\tau)\} \\ &\leq \frac{1}{N} \Omega_{abcd} + \frac{2}{N} \sum_{\tau} \{|S_{ac}(\tau)||S_{bd}(\tau)| + |S_{ad}(\tau)||S_{bc}(\tau)|\}. \end{aligned}$$

Next, using the inequalities  $|\sum_i u_i v_i| \leq \sum_i |u_i||v_i| \leq \frac{1}{2} \sum_i (u_i^2 + v_i^2)$ , we have:

$$|Cov(\mathbf{\Delta}_{ab}, \mathbf{\Delta}_{cd})| \leq \frac{|\Omega_{abcd}|}{N} + \frac{1}{N} \sum_{\tau} |S_{ac}(\tau)|^2 + |S_{bd}(\tau)|^2 + |S_{ad}(\tau)|^2 + |S_{bc}(\tau)|^2.$$

Now using the mixing condition,  $\sum_{\tau=0}^{\infty} |S_{ij}(\tau)|^2 \leq \Omega_{ij}$ , we eventually obtain:

$$|Cov(\mathbf{\Delta}_{ab}, \mathbf{\Delta}_{cd})| \leq \frac{|\Omega_{abcd}|}{N} + \frac{1}{N}(\Omega_{ac} + \Omega_{bd} + \Omega_{ad} + \Omega_{bc}) \quad (7)$$

which shows that  $Cov(\mathbf{\Delta}_{ab}, \mathbf{\Delta}_{cd}) = O(1/N)$ .  $\square$

If we denote by  $\mathbf{G}$  and  $\hat{\mathbf{G}}$  the inverse of  $\mathbf{S}$  and  $\hat{\mathbf{S}}$ , respectively, we have the lemma below.

**Lemma 3.2.** *The inverse  $\hat{\mathbf{G}}$  of  $\hat{\mathbf{S}}$  can be approximated by*

$$\hat{\mathbf{G}} = \mathbf{G} - \mathbf{G}\mathbf{\Delta}\mathbf{G} + \mathbf{G}\mathbf{\Delta}\mathbf{G}\mathbf{\Delta}\mathbf{G} + o(1/N). \quad (8)$$

*Proof.* Let  $\mathbf{E}$  be the symmetric matrix  $\mathbf{E} = -\mathbf{S}^{-1/2}\mathbf{\Delta}\mathbf{S}^{-1/2}$ . Then with this definition,  $\hat{\mathbf{G}} = \mathbf{S}^{-1/2}(\mathbf{I} + \mathbf{E})^{-1}\mathbf{S}^{-1/2}$ . Now we know that for any matrix  $\mathbf{E}$  with spectral radius smaller than 1 the series  $\sum_{k=0}^{\infty} \mathbf{E}^k$  converges to  $(\mathbf{I} - \mathbf{E})^{-1}$ . If we plug this series in the expression of  $\hat{\mathbf{G}}$  we get  $\hat{\mathbf{G}} = \mathbf{S}^{-1/2} \sum_{k=0}^K \mathbf{E}^k \mathbf{S}^{-1/2} + o(\|\mathbf{E}\|^K)$ . Replacing  $\mathbf{E}$  by its definition eventually yields (8).  $\square$

Now it is desirable to express  $\hat{\mathbf{G}}$  as a function of  $\hat{\mathbf{S}}$ . If we replace  $\mathbf{\Delta}$  by  $\hat{\mathbf{S}} - \mathbf{S}$  in (8), we obtain:

$$\hat{\mathbf{G}} = 3\mathbf{G} - 3\mathbf{G}\hat{\mathbf{S}}\mathbf{G} + \mathbf{G}\hat{\mathbf{S}}\mathbf{G}\hat{\mathbf{S}}\mathbf{G} + o(1/N). \quad (9)$$

With this approximation,  $\hat{\mathbf{G}}$  is now a polynomial function of  $\hat{\mathbf{S}}$  of degree 2, and hence of degree 4 in  $\mathbf{x}$ . We shall show that the mean of  $\hat{B}_p(N)$  involves moments of  $\mathbf{x}$  up to order 8, whereas its variance involves moments up to order 16.

**Lemma 3.3.** *Denote  $A_{ij} = \mathbf{x}(i)^T \mathbf{S}^{-1} \mathbf{x}(j)$ . Then:*

$$\begin{aligned} \hat{B}_p(N) &= \frac{6}{N} \sum_{n=1}^N A_{nn}^2 - \frac{8}{N^2} \sum_{n=1}^N A_{nn} \sum_{i=1}^N A_{ni}^2 + \frac{1}{N^3} \sum_{n=1}^N \left( \sum_{i=1}^N A_{ni}^2 \right) \left( \sum_{j=1}^N A_{nj}^2 \right) \\ &+ \frac{2}{N^3} \sum_{n=1}^N \sum_{j=1}^N \sum_{k=1}^N A_{nn} A_{nj} A_{jk} A_{kn} + o(1/N) \end{aligned} \quad (10)$$

*Proof.* First inject (8) in the expression  $\hat{B}_p(N) = \frac{1}{N} \sum_{n=1}^N \left( \mathbf{x}(n)^T \hat{\mathbf{G}} \mathbf{x}(n) \right)^2$ , and keep terms up to order  $O(\|\Delta\|^2)$ ; this yields:

$$\begin{aligned} \hat{B}_p(N) = \frac{1}{N} \sum_n \left[ A_{nn}^2 - 2A_{nn} \mathbf{x}(n)^T \mathbf{G} \Delta \mathbf{G} \mathbf{x}(n) + (\mathbf{x}(n)^T \mathbf{G} \Delta \mathbf{G} \mathbf{x}(n))^2 \right. \\ \left. + 2A_{nn} \mathbf{x}(n)^T \mathbf{G} \Delta \mathbf{G} \Delta \mathbf{G} \mathbf{x}(n) \right] + o(\|\Delta\|^2). \end{aligned}$$

Then replace  $\Delta$  by  $\hat{\mathbf{S}} - \mathbf{S}$ . This leads to

$$\begin{aligned} \hat{B}_p(N) = \frac{1}{N} \sum_n \left[ 6A_{nn}^2 - 8A_{nn} (\mathbf{x}(n)^T \mathbf{G} \hat{\mathbf{S}} \mathbf{G} \mathbf{x}(n)) \right. \\ \left. + (\mathbf{x}(n)^T \mathbf{G} \hat{\mathbf{S}} \mathbf{G} \mathbf{x}(n))^2 + 2A_{nn} (\mathbf{x}(n)^T \mathbf{G} \hat{\mathbf{S}} \mathbf{G} \hat{\mathbf{S}} \mathbf{G} \mathbf{x}(n)) \right] + o(\|\Delta\|^2). \end{aligned}$$

Equation (10) is eventually obtained after replacing  $\hat{\mathbf{S}}$  by  $\frac{1}{N} \sum_k \mathbf{x}(k) \mathbf{x}(k)^T$  and all terms of the form  $\mathbf{x}(q)^T \mathbf{G} \mathbf{x}(r)$  by  $A_{qr}$ .  $\square$

### 3.2. Additional notations and computational issues

When computing the mean and variance of  $\hat{B}_p(N)$  given in (10), higher order moments of the multivariate random variable  $\mathbf{x}$  will arise. Under the normal (null) hypothesis, these moments are expressed as functions of second order moments only. To keep notations reasonably concise, it is proposed to use McCullagh's bracket notation (McCullagh, 1987), briefly reminded in Appendix 11.1. Furthermore, for all moments of order higher than  $p$ , some components appear multiple times; counting the number of identical terms in the expansion of the higher moments is a tedious task. All the moment expansions that are necessary for the derivations presented in this paper are developed in Appendix 11.3.

In order to keep notations as explicit and concise as possible, while keeping explicit the role of both coordinate (or space) indices and time indices, let the moments of  $\mathbf{x}(t)$ , whose  $p$  components are  $x_a(t)$ ,  $1 \leq a \leq p$  be noted

$$\mu_{ab}^{tu} = \mathbb{E}\{x_a(t)x_b(u)\}, \quad \mu_{abc}^{tuv} = \mathbb{E}\{x_a(t)x_b(u)x_c(v)\} \quad (11)$$

and so forth for higher orders. It shall be emphasized that different time and coordinate indices appear here as the components are assumed to be colored (time correlated) and dependent to each others (spatially correlated).



Computation of the mean and variance of  $\hat{B}_p$  defined by equation (10) involves the computation of moments of order noted  $2L$  whose generic expression is

$$\mathbb{E}\left\{\prod_{l=1}^L A_{\alpha_l \beta_l}\right\} = \sum_{r_1 \dots r_L, c_1 \dots c_L=1}^p \left( \prod_{i=1}^L G_{r_i, c_i} \right) \mathbb{E}\{x_{r_1}(\alpha_1) x_{c_1}(\beta_1) \dots x_{r_L}(\alpha_L) x_{c_L}(\beta_L)\}$$

or equivalently

$$\mathbb{E}\left\{\prod_{l=1}^L A_{\alpha^l \beta^l}\right\} = \sum_{r_1 \dots r_L, c_1 \dots c_L=1}^p \left( \prod_{i=1}^L G_{r_i, c_i} \right) \mu_{r_1 \dots r_L c_1 \dots c_L}^{\alpha_1 \dots \alpha_L \beta_1 \dots \beta_L} \quad (12)$$

In the above equation, the  $2L$ -order moment  $\mu_{r_1 \dots r_L c_1 \dots c_L}^{\alpha_1 \dots \alpha_L \beta_1 \dots \beta_L}$  has superscripts indicating the time indices involved, whereas the subscripts indicate the coordinate (or space) indices.

While being general, the above formulation may take simpler, or more explicit forms in practice. The detailed methodology for computing the expressions of the mean and variance of  $\hat{B}_p$  as functions of second order moments is deferred to Appendix 11.2. The resulting expressions of Mardia's statistics are given and discussed in the sections to come.

#### 4. Expression of the mean of $B_p(N)$

According to Equation (10), we have four types of terms. The goal of this section is to provide the expectation of each of these terms.

**Lemma 4.1.** *With the definition of  $A_{ij}$  given in Lemma 3.3, we have:*

$$\mathbb{E}\{A_{nn}^2\} = \sum_{a,b,c,d=1}^p G_{ab} G_{cd} \mu_{abcd}^{nnnn} \quad (13)$$

$$\mathbb{E}\{A_{nn} A_{ni}^2\} = \sum_{a,b,c,d=1}^p \sum_{e,f=1}^p G_{ab} G_{cd} G_{ef} \mu_{abcdf}^{nnnnii} \quad (14)$$

$$\mathbb{E}\{A_{ni}^2 A_{nj}^2\} = \sum_{a,b,c,d=1}^p \sum_{e,f,g,h=1}^p G_{ab} G_{cd} G_{ef} G_{gh} \mu_{acegbdhf}^{nnnnijjj} \quad (15)$$

$$\mathbb{E}\{A_{nn} A_{nj} A_{jk} A_{kn}\} = \sum_{a,b,c,d=1}^p \sum_{e,f,g,h=1}^p G_{ab} G_{cd} G_{ef} G_{gh} \mu_{abchdefg}^{nnnnjjkk} \quad (16)$$

**Proposition 4.1.** *Using expressions of moments given in Appendix 11.3, the expectations of the four terms defined in Lemma 4.1 take the form below*

$$\mathbb{E}\{A_{nn}^2\} = \sum_{klqr=1}^p G_{kl}G_{rq} \left\{ [3]\mu_{kl}^{nn}\mu_{qr}^{nn} \right\} \quad (17)$$

$$\mathbb{E}\{A_{nn}A_{ni}^2\} = \sum_{klqrst=1}^p G_{kl}G_{qr}G_{st} \left\{ [12]\mu_{kr}^{ni}\mu_{lt}^{ni}\mu_{qs}^{nn} + [3]\mu_{kl}^{nn}\mu_{qs}^{nn}\mu_{rt}^{ii} \right\} \quad (18)$$

$$\begin{aligned} \mathbb{E}\{A_{ni}^2A_{nj}^2\} = & \sum_{k,\ell,q,r} \sum_{s,t,u,v} G_{kl}G_{qr}G_{st}G_{uv} \left\{ [3]\mu_{kq}^{nn}\mu_{su}^{nn}\mu_{lr}^{ii}\mu_{tv}^{jj} \right. \\ & + [6]\mu_{kq}^{nn}\mu_{su}^{nn}\mu_{lt}^{ij}\mu_{rv}^{ij} + [12]\mu_{kq}^{nn}\mu_{sl}^{ni}\mu_{ur}^{ni}\mu_{tv}^{jj} + [24]\mu_{kt}^{nj}\mu_{qv}^{nj}\mu_{ls}^{in}\mu_{ur}^{ni} \\ & \left. + [48]\mu_{kl}^{ni}\mu_{rt}^{ij}\mu_{qv}^{nj}\mu_{su}^{nn} + [12]\mu_{kq}^{nn}\mu_{ts}^{jn}\mu_{uv}^{nj}\mu_{rl}^{ii} \right\} \quad (19) \end{aligned}$$

$$\begin{aligned} \mathbb{E}\{A_{nn}A_{nj}A_{jk}A_{kn}\} = & \sum_{m,\ell,q,r} \sum_{s,t,u,v} G_{ml}G_{qr}G_{st}G_{uv} \left\{ [3]\mu_{ml}^{nn}\mu_{qv}^{nn}\mu_{sr}^{jj}\mu_{tu}^{kk} \right. \\ & + [6]\mu_{ml}^{nn}\mu_{qv}^{nn}\mu_{rt}^{jk}\mu_{su}^{jk} + [12]\mu_{ml}^{nn}\mu_{qr}^{nj}\mu_{vs}^{nj}\mu_{tu}^{kk} + [24]\mu_{mv}^{nk}\mu_{lu}^{nk}\mu_{qr}^{nj}\mu_{vs}^{nj} \\ & \left. + [48]\mu_{mr}^{nj}\mu_{st}^{jk}\mu_{lu}^{nk}\mu_{qv}^{nn} + [12]\mu_{kl}^{nn}\mu_{qt}^{nk}\mu_{vu}^{nk}\mu_{rs}^{jj} \right\} \quad (20) \end{aligned}$$

The mean of  $\hat{B}_p(N)$  then follows from (10).

## 5. Expression of the variance of $B_p(N)$

From Lemma 3.3, we can also state what moments of  $A_{ij}$  will be required in the expression of the variance of  $B_p(N)$ .

**Lemma 5.1.** *By raising (10) to the second power and using the definition of  $A_{ij}$  given in Lemma 3.3, we can check that the following moments are*

required:

$$\mathbb{E}\{A_{nn}^2 A_{ii}^2\} = \sum_{a,b,c,d,e,f,g,h=1}^p G_{ab} G_{cd} G_{ef} G_{gh} \mu_{abcdefgh}^{nnniii} \quad (21)$$

$$\mathbb{E}\{A_{nn}^2 A_{ij}^2 A_{ii}\} = \sum_{a,b,c,d=1}^p \sum_{e,f,g,h=1}^p \sum_{m,\ell=1}^p G_{ab} G_{cd} G_{ef} G_{gh} G_{m\ell} \mu_{abcdegm\ell fh}^{nnniiiij} \quad (22)$$

$$\mathbb{E}\{A_{nn} A_{kk} A_{ni}^2 A_{kj}^2\} = \sum_{a,b,c,d=1}^p \sum_{e,f,g,h=1}^p \sum_{m,\ell,q,r=1}^p G_{ab} G_{cd} G_{ef} G_{gh} G_{m\ell} G_{qr} \mu_{abegcdmqfhlr}^{nnnnkkkiiij} \quad (23)$$

$$\mathbb{E}\{A_{kk}^2 A_{ni}^2 A_{nj}^2\} = \sum_{a,b,c,d=1}^p \sum_{e,f,g,h=1}^p \sum_{m,\ell,q,r=1}^p G_{ab} G_{cd} G_{ef} G_{gh} G_{m\ell} G_{qr} \mu_{abcdegmqfhlr}^{kkkknnniiij} \quad (24)$$

$$\mathbb{E}\{A_{nn}^2 A_{ii} A_{ij} A_{jk} A_{ki}\} = \sum_{a,b,c,d=1}^p \sum_{e,f,g,h=1}^p \sum_{m,\ell,q,r=1}^p G_{ab} G_{cd} G_{ef} G_{gh} G_{m\ell} G_{qr} \mu_{abcdefgrhmlq}^{nnniiiijkk} \quad (25)$$

$$\mathbb{E}\{A_{ni}^2 A_{nj}^2 A_{kt}^2 A_{kk}\} = \sum_{a,b,c,d=1}^p \sum_{e,f,g,h=1}^p \sum_{m,\ell,q,r=1}^p \sum_{s,u=1}^p G_{ab} G_{cd} G_{ef} G_{gh} G_{m\ell} G_{qr} G_{su} \mu_{acegmqsudfhlr}^{nnnnkkkiiijtt} \quad (26)$$

$$\mathbb{E}\{A_{ii}A_{it}^2A_{nn}A_{nj}A_{jk}A_{kn}\} = \sum_{a,b,c,d=1}^p \sum_{e,f,g,h=1}^p \sum_{m,\ell,q,r=1}^p \sum_{s,u=1}^p G_{ab}G_{cd}G_{ef}G_{gh}G_{ml} G_{qr}G_{su} \mu_{abceghmulqrsdf}^{iiiiinnnjjkktt} \quad (27)$$

$$\mathbb{E}\{A_{ni}^2A_{kt}^2A_{nj}^2A_{ku}^2\} = \sum_{a,b,c,d=1}^p \sum_{e,f,g,h=1}^p \sum_{m,\ell,q,r=1}^p \sum_{s,v,w,z=1}^p G_{ab}G_{cd}G_{ef}G_{gh}G_{ml} G_{qr}G_{sv}G_{wz} \mu_{acmqegswbdlr}^{nnnnkkkkiiijttuu} \quad (28)$$

$$\mathbb{E}\{A_{nn}A_{nj}A_{jk}A_{kn}A_{ii}A_{it}A_{tu}A_{ui}\} = \sum_{a,b,c,d=1}^p \sum_{e,f,g,h=1}^p \sum_{m,\ell,q,r=1}^p \sum_{s,v,w,z=1}^p G_{ab}G_{cd}G_{ef}G_{gh}G_{ml} G_{qr}G_{sv}G_{wz} \mu_{abchmlqzdefgrsvw}^{nnnniiiijjkkttuu} \quad (29)$$

$$\mathbb{E}\{A_{nn}A_{nj}A_{jk}A_{kn}A_{it}^2A_{iu}^2\} = \sum_{a,b,c,d=1}^p \sum_{e,f,g,h=1}^p \sum_{m,\ell,q,r=1}^p \sum_{s,v,w,z=1}^p G_{ab}G_{cd}G_{ef}G_{gh}G_{ml} G_{qr}G_{sv}G_{wz} \mu_{abchmqswdefglrvz}^{nnnniiiijjkkttuu} \quad (30)$$

Then, as in Proposition 4.1, by using the results of Appendix 11.3, the moments  $\mu_*^*$  could be in turn expressed as a function of second order moments. For readability, we do not substitute here these values.

**Proposition 5.1.**

$$\begin{aligned} \text{Var}\{\hat{B}_p\} &= \frac{36}{N^2} \sum_n \sum_i \mathbb{E}\{A_{nn}^2A_{ii}^2\} - \frac{96}{N^3} \sum_j \sum_{n,i} \mathbb{E}\{A_{nn}^2A_{ij}^2A_{ii}\} \\ &+ \frac{64}{N^4} \sum_{n,i} \sum_{j,k} \mathbb{E}\{A_{nn}A_{kk}A_{ni}^2A_{kj}^2\} + \frac{12}{N^4} \sum_{n,i,j,k} \mathbb{E}\{A_{kk}^2A_{ni}^2A_{nj}^2\} \\ &+ \frac{24}{N^4} \sum_{n,i,j,k} \mathbb{E}\{A_{nn}^2A_{ii}A_{ij}A_{jk}A_{ki}\} - \frac{16}{N^5} \sum_{n,i} \sum_{j,k,t} \mathbb{E}\{A_{ni}^2A_{nj}^2A_{kt}^2A_{kk}\} \\ &- \frac{32}{N^5} \sum_{i,t} \sum_{n,j,k} \mathbb{E}\{A_{ii}A_{it}^2A_{nn}A_{nj}A_{jk}A_{kn}\} + \frac{1}{N^6} \sum_{n,i,j} \sum_{k,t,u} \mathbb{E}\{A_{ni}^2A_{kt}^2A_{nj}^2A_{ku}^2\} \\ &+ \frac{4}{N^6} \sum_{n,j,k} \sum_{i,t,u} \mathbb{E}\{A_{nn}A_{nj}A_{jk}A_{kn}A_{ii}A_{it}A_{tu}A_{ui}\} \\ &+ \frac{4}{N^6} \sum_{n,j,k} \sum_{i,t,u} \mathbb{E}\{A_{nn}A_{nj}A_{jk}A_{kn}A_{it}^2A_{iu}^2\} - (\mathbb{E}\{\hat{B}_p\})^2 \end{aligned} \quad (31)$$

## 6. Mean and variance of $\hat{B}_1(N)$ in the scalar case ( $p = 1$ )

The complicated expressions obtained in the previous sections simplify drastically in the scalar case, and we get the results below.

$$\mathbb{E}\{\hat{B}_1\} = 3 - \frac{6}{N} - \frac{12}{N^2} \sum_{\tau=1}^{N-1} (N - \tau) \frac{S(\tau)^2}{S^2} + o\left(\frac{1}{N}\right) \quad (32)$$

$$Var\{\hat{B}_1\} = \frac{24}{N} \left[ 1 + \frac{2}{N} \sum_{\tau=1}^{N-1} (N - \tau) \frac{S(\tau)^4}{S^4} \right] + o\left(\frac{1}{N}\right) \quad (33)$$

In particular in the i.i.d. case,  $S(\tau) = 0$  for  $\tau \neq 0$ , and we get the well-known result (Mardia, 1974) (Comon and Deruaz, 1995):

$$\mathbb{E}\{\hat{B}_1\} \approx 3 - \frac{6}{N}, \quad \text{and} \quad Var\{\hat{B}_1\} \approx \frac{24}{N}.$$

## 7. Mean and variance of $\hat{B}_2(N)$ in the bivariate case ( $p = 2$ )

In the bivariate case, expressions become immediately more complicated, but we can still write them explicitly, as reported below. We remind that  $\mu_{ab}^{ij} = S_{ab}(i - j)$ .

$$\mathbb{E}\{\hat{B}_2\} = 8 - \frac{16}{N} - \frac{4}{N^2} \sum_{\tau=1}^{N-1} \frac{(N - \tau)Q_1(\tau)}{(S_{11}S_{22} - S_{12}^2)^2} + o\left(\frac{1}{N}\right) \quad (34)$$

with

$$\begin{aligned} Q_1(\tau) = & S_{11}S_{22} \left[ (S_{12}(\tau) + S_{21}(\tau))^2 - 4S_{11}(\tau)S_{22}(\tau) \right] \\ & + S_{12}^2 \left[ 2(S_{12}(\tau) + S_{21}(\tau))^2 + 4S_{22}(\tau)S_{11}(\tau) \right] \\ & - 6S_{22}S_{12} \left( S_{11}(\tau)(S_{12}(\tau) + S_{21}(\tau)) \right) \\ & - 6S_{11}S_{12} \left( S_{22}(\tau)(S_{12}(\tau) + S_{21}(\tau)) \right) \\ & + 6S_{11}^2S_{22}^2(\tau) + 6S_{22}^2S_{11}^2(\tau). \end{aligned} \quad (35)$$

$$Var\{\hat{B}_2\} = \frac{64}{N} + \frac{16}{N^2} \sum_{\tau=1}^{N-1} \frac{(N - \tau)Q_2(\tau)}{(S_{11}S_{22} - S_{12}^2)^4} + o\left(\frac{1}{N}\right) \quad (36)$$

with

$$\begin{aligned}
Q_2(\tau) = & \left[ -4S_{11}^2(\tau)S_{22}^2(\tau) + 16S_{11}(\tau)S_{22}(\tau)S_{12}(\tau)S_{21}(\tau) \right. \\
& + 2S_{12}^2(\tau)S_{21}^2(\tau) + 3(S_{21}^4(\tau) + S_{12}^4(\tau)) \\
& \left. + 12S_{11}(\tau)S_{22}(\tau)(S_{12}(\tau) + S_{21}(\tau)) \right] S_{11}^2 S_{22}^2 \\
& + 2S_{11}^2 S_{12}^2 \left[ 8S_{11}(\tau)S_{22}(\tau) + 3(5S_{11}(\tau)S_{22}(\tau) + S_{21}(\tau)S_{12}(\tau) \right. \\
& \left. (S_{21}(\tau) + S_{12}(\tau))^2 - 4S_{21}(\tau)S_{12}(\tau)(S_{22}^2(\tau) + S_{21}(\tau)S_{12}(\tau)) \right] \\
& + 2S_{22}^2 S_{12}^2 \left[ 8S_{22}(\tau)S_{11}(\tau) + 3(5S_{11}(\tau)S_{22}(\tau) + S_{21}(\tau)S_{12}(\tau) \right. \\
& \left. (S_{21}(\tau) + S_{12}(\tau))^2 - 4S_{21}(\tau)S_{12}(\tau)(S_{11}^2(\tau) + S_{21}(\tau)S_{12}(\tau)) \right] \\
& + 6S_{11}^4 S_{22}^4(\tau) + 6S_{22}^4 S_{11}^4(\tau) \\
& + 8S_{12}^4 \left[ S_{11}^2(\tau)S_{22}^2(\tau) + 4S_{11}(\tau)S_{22}(\tau)S_{12}(\tau)S_{12}(\tau) + S_{21}^2(\tau)S_{12}^2(\tau) \right] \\
& - 12S_{11}S_{12}S_{22}(\tau)(S_{12}(\tau) + S_{21}(\tau)) \left[ (2S_{11}(\tau)S_{22}(\tau) + S_{21}^2(\tau) + S_{12}^2(\tau)) \right. \\
& \left. S_{11}S_{22} + 2(S_{11}(\tau)S_{22}(\tau) + S_{12}(\tau)S_{21}(\tau))S_{12}^2 \right] \\
& - 12S_{22}S_{12}S_{11}(\tau)(S_{12}(\tau) + S_{21}(\tau)) \left[ (2S_{11}(\tau)S_{22}(\tau) + S_{21}^2(\tau) + S_{12}^2(\tau)) \right. \\
& \left. S_{11}S_{22} + 2(S_{11}(\tau)S_{22}(\tau) + S_{12}(\tau)S_{21}(\tau))S_{12}^2 \right]. \tag{37}
\end{aligned}$$

## 8. Particular case: multidimensional embedding of a scalar process

In this section, we consider the particular case where the multivariate process consists of the embedding of a scalar process. More precisely, we assume that

$$\mathbf{x}(n) = \begin{pmatrix} x_1(n) \\ \dots \\ x_p(n) \end{pmatrix} = \begin{pmatrix} y(n\delta + 1) \\ \dots \\ y(n\delta + p) \end{pmatrix}.$$

where  $y(k)$  is a scalar wide-sense stationary process of correlation function  $C(\tau) = \mathbb{E}\{y(k)y(k-\tau)\} = S_{11}(\tau/\delta)$ . Note that now, because of the particular form of  $\mathbf{x}(n)$ , we can exploit the translation invariance by remarking that  $S_{ab}(\tau) = \mathbb{E}\{x_a(n\delta)x_b(n\delta-\tau\delta)\}$  implies  $S_{ab}(\tau) = C(\tau\delta+a-b)$ , for  $1 \leq a, b \leq p$ .

To keep results as concise as possible, we assume the notation  $\gamma_i(\tau) = C(\tau\delta + i)$ , and the shortcut  $C_j = C(j)$ . The main goal targeted by defining these multiple notations is to obtain more compact expressions.

### 8.1. Bivariate embedding

The bivariate case is more difficult but the expressions still have a simple form:

$$\mathbb{E}\{\hat{B}_2\} \approx 8 - \frac{16}{N} - \frac{4}{N^2} \sum_{\tau=1}^{N-1} \frac{(N-\tau)q_1(\tau)}{(C_0^2 - C_1^2)^2} \quad (38)$$

$$\text{Var}\{\hat{B}_2\} \approx \frac{64}{N} + \frac{16}{N^2} \sum_{\tau=1}^{N-1} \frac{(N-h)q_2(\tau)}{(C_0^2 - C_1^2)^4} \quad (39)$$

with  $q_1(\tau)$  and  $q_2(\tau)$  defined below, where  $\gamma_i$  stands for  $\gamma_i(\tau)$ :

$$q_1(\tau) = \left[ (\gamma_1 + \gamma_{-1})^2 + 8\gamma_0^2 \right] C_0^2 - 12C_0C_1 \gamma_0(\gamma_1 + \gamma_{-1}) + \left[ 2(\gamma_1 + \gamma_{-1})^2 + 4\gamma_0^2 \right] C_1^2, \quad (40)$$

$$q_2(\tau) = \left[ 8(\gamma_0^2 - \gamma_1\gamma_{-1})^2 + 3(\gamma_1^2 - \gamma_{-1}^2)^2 + 12\gamma_0^2(\gamma_1 + \gamma_{-1})^2 \right] C_0^4 + 4 \left[ 8\gamma_0^4 + 3(5\gamma_0^2 + \gamma_1\gamma_{-1})(\gamma_1 + \gamma_{-1})^2 - 4\gamma_1\gamma_{-1}(\gamma_0^2 + \gamma_1\gamma_{-1}) \right] C_0^2 C_1^2 + 8 \left[ \gamma_0^4 + 4\gamma_0^2\gamma_1\gamma_{-1} + \gamma_1^2\gamma_{-1}^2 \right] C_1^4 - 24C_0C_1 \gamma_0(\gamma_1 + \gamma_{-1}) \left[ (2\gamma_0^2 + \gamma_1^2 + \gamma_{-1}^2) C_0^2 + 2(\gamma_0^2 + \gamma_1\gamma_{-1}) C_1^2 \right]. \quad (41)$$

The exact computation for the trivariate embedding case have also been conducted; but because of their lengthily expressions (especially that of the variance), they are not detailed here and can be given as supplementary material upon request.

## 9. Computer experiments

In this section, the preceding results are illustrated on dedicated computer experiments. To emphasize the importance of the univariate and the bivariate normality tests on colored random process, we simulate correlated

bivariate random processes with Gaussian marginals. The generation procedure is briefly described in the next section. Then tests are performed to detect non Gaussian nature of the joint distribution while the marginals remain Gaussian.

### 9.1. Gaussian Marginals under $\bar{H}$

Copulas are a simple to implement, classical framework for defining multivariate distributions with controlled joint distribution function. It is known that there is a *unique* copula -called the Gaussian copula  $\mathcal{C}_R$  - that produces the bivariate Gaussian distribution, fully specified by the correlation matrix  $R$ :

$$\mathcal{C}_R(u, v) = \int_{-\infty}^{\Phi^{-1}(u)} \int_{-\infty}^{\Phi^{-1}(v)} \frac{1}{2\pi(1 - R_{12}^2)^{1/2}} \exp\left\{-\frac{s^2 - 2R_{12}st + t^2}{2(1 - R_{12}^2)}\right\} ds dt \quad (42)$$

where  $\Phi^{-1}$  is the inverse of the cumulative distribution function of the standard normal distribution. As Sklar's theorem (cf. Appendix 11.6) guarantees the uniqueness of the copula generating a given bivariate distribution, non Gaussian distributions can easily be obtained by using other types of copulas. Namely here, Clayton and Gumbel bivariate copulas are used as examples:

$$\text{Clayton: } \mathcal{C}_\theta(u, v) = \max\{u^{-\theta} + v^{-\theta} - 1; 0\}, \theta \in [-1, \infty) \setminus \{0\} \quad (43)$$

$$\text{Gumbel: } \mathcal{C}_\theta(u, v) = \exp\left\{-(-\log(u)^\theta + -\log(v)^\theta)^{\frac{1}{\theta}}\right\}, \theta \in [1, \infty) \quad (44)$$

Since Sklar's theorem does not impose *independence* of any variate  $u$  or  $v$  of  $\mathcal{C}_\theta(u, v)$ , we propose the following algorithm to generate a bivariate copula with *colored* Gaussian marginals.

- Generate two *i.i.d* centered normalized Gaussian variables:  $\eta_1, \eta_2 \underset{i.i.d}{\sim} \mathcal{N}(0, 1)$
- Make the previous variables correlated in time by a first-order autoregressive filter:

$$\begin{aligned} y_1(n) &= 0.8y_1(n-1) + \eta_1(n) \\ y_2(n) &= 0.8y_2(n-1) + \eta_2(n) \end{aligned}$$



Thus  $\mathbb{E}\{y_1(n)y_1(n-k)\} = 0.8^{|k|}$ , for all  $k \in \mathbb{Z}$ .

- Transform  $y_1$  and  $y_2$  as:

$$u = \Phi(y_1) \tag{45}$$

$$v = \Phi(y_2) \tag{46}$$

Note that  $u$  and  $v$  are uniformly distributed on  $[0, 1]$ . Thus, we can generate new samples  $u'$ ,  $v'$  coupled by a given copula  $\mathcal{C}_\theta$ . For more details about efficient sampling of copula see the (Marshall and Olkin 1988 algorithm) cited in (Hofert, 2008).

- Transform  $u'$  and  $v'$  to obtain Gaussian standard marginals:  $\mathbf{x} = (x_1(n), x_2(n))^T$ :

$$x_1(n) = \Phi^{-1}(u(n))$$

$$x_2(n) = \Phi^{-1}(v(n))$$

### *Simulation study*

For a given copula  $\mathcal{C}$ , we perform  $M = 2000$  realizations of  $\mathbf{x}(n) = (x_1(n), x_2(n))^T$  of total length  $N = 1000$ . First, the  $p$ -values of the two-sided tests are computed based on:

$$t = \frac{\hat{B}_{(\cdot)} - \mathbb{E}\{\hat{B}_{(\cdot)}\}}{\sqrt{\text{Var}\{\hat{B}_{(\cdot)}\}}}$$

Recall that this statistic is standard normal. Then  $p$ -value =  $2(1 - \Phi(|t|))$  is compared to pre-specified significance levels  $\alpha$ . For any  $p$  smaller than  $\alpha$ , it is considered heuristically that the test rejected  $H$ . The empirical rejection rates, defined by  $\frac{\text{Number of rejections}}{M}$  for each statistic  $\hat{B}_{1,i.i.d}$ ,  $\hat{B}_1$  and  $\hat{B}_2$  are reported in Table 1.

### *Mardia's test*

- $\hat{B}_{1,i.i.d}$ : Under the null hypothesis  $H$ , the rejection rate surpasses the nominal level. That  $\hat{B}_{1,i.i.d}$  over-rejects  $H$  is due to the one-dimensional marginal being time-correlated. Such observation was already formulated by (Moore, 1982) and (Gasser, 1975) who showed that the correlation among samples is

Test statistic	Gaussian $R_{12} = 0.8$		Clayton $\theta = 1.5$		Gumbel $\theta = 5$	
	$\alpha = 5\%$	$\alpha = 10\%$	$\alpha = 5\%$	$\alpha = 10\%$	$\alpha = 5\%$	$\alpha = 10\%$
$\hat{B}_{1,i.i.d}$	0.1660	0.2460	0.1011	0.1651	0.1189	0.1930
$\hat{B}_1$	0.0450	0.0730	0.1060	0.1701	0.0390	0.0860
$\hat{B}_2$	0.0480	0.0801	0.9890	0.9920	0.9920	0.9960

Table 1: Empirical Rejection rate at two significance levels :  $\alpha = 5\%, 10\%$

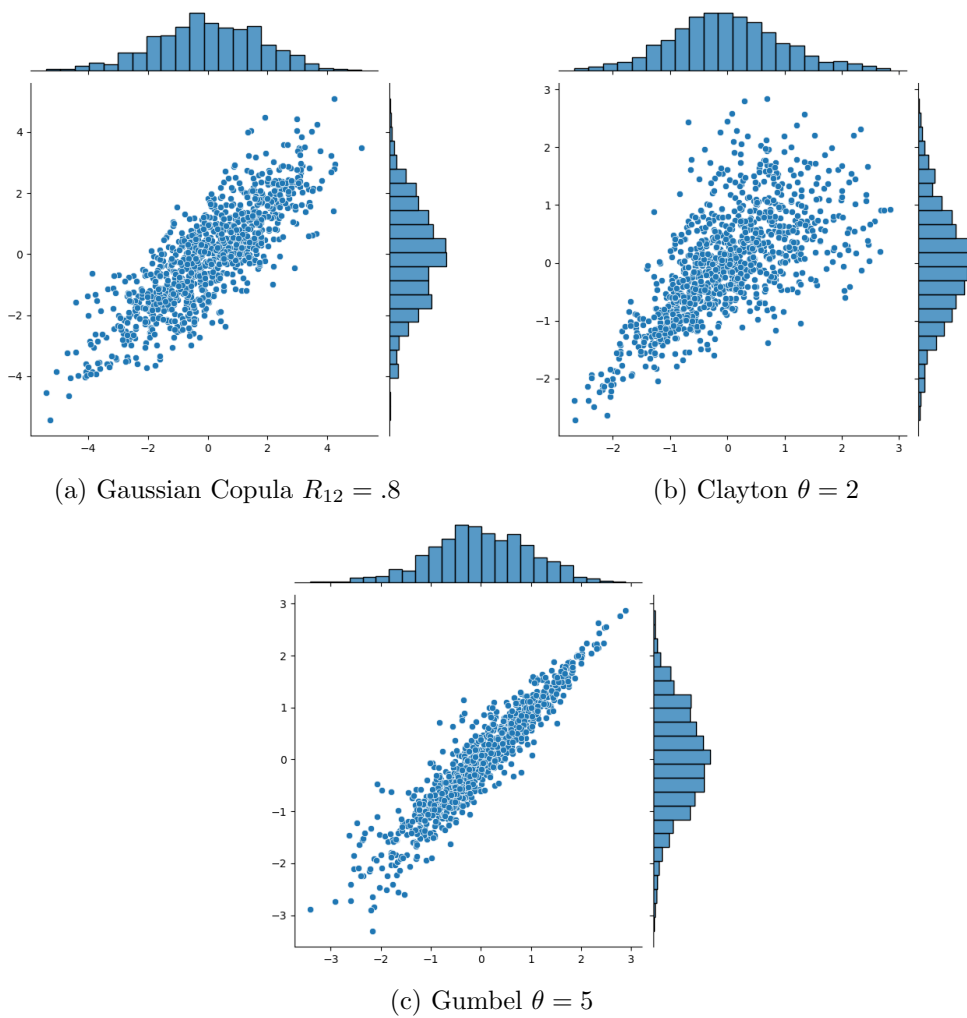


Figure 1: Examples of non-Gaussian process whose marginals are standard normal

confounded with lack of Normality.

- $\hat{B}_{1,i.i.d}$  and  $\hat{B}_1$  test one-dimensional marginals only, therefore they are always conservative.

- $\hat{B}_2$ : The rejection rates do not differ substantially from the nominal level when data is distributed according to bivariate Gaussian. Under  $\bar{H}$ , this test has very high rejection rates, which confirms the necessity of taking into account the full dimension to design a powerful test.

### 9.2. Detection of a time-series embedded in Gaussian noise

In this simulation, the detection of an additive corruption in a Gaussian process is considered:

$$y(n) = x(n) + kb(n) \quad (47)$$

where  $x(n)$  is a first order auto-regressive process AR(1):  $x(n) = 0.8x(n-1) + \eta(n)$  and where  $\eta \underset{iid}{\sim} \mathcal{N}(0, S)$ ;  $b(n) = 0.8b(n-1) - 0.5b(n-2) + \epsilon(n)$ , where  $\epsilon$  follows a double-exponential distribution with unit scale parameter. We perform 500 replications of  $\{y(n)\}$  of total length  $N_{tot} = n_{drop} + N$ , the first  $n_{drop} = 1000$  observations at the beginning of the sample are discarded to alleviate side effects and reduce the dependence on initial values:  $x(1) = \eta(1)$  and  $b(1) = \epsilon(1)$ . For each data record, the covariance function  $\gamma_{a,b}(i)$  is estimated once for a fixed dimension  $p$  for all the test statistics. Testing the normality of the process  $y(n)$  can be accomplished by standard scalar tests. By exploiting the results in Section 8, we propose to test the joint normality of its successive values:  $\mathbf{x}(n) = (y(2n+1), y(2n+2))^T$ ; Note that here  $\delta = 2$ .

The normality test can be reformulated in terms of the detection of an *unknown* non-Gaussian signal embedded in Gaussian noise. The ability of the test to detect the presence of  $b(n)$  for different  $SNR = k^2 \frac{\mathbb{E}\{b(n)^2\}}{\mathbb{E}\{x(n)^2\}}$  is reported in Figure 2.

As SNR increases, statistic  $\hat{B}_2$  is the first to detect the presence of an additive non-Gaussian process, followed by  $\hat{B}_1$  and  $\hat{B}_{1,i.i.d}$  whose behaviors do not differ substantially.

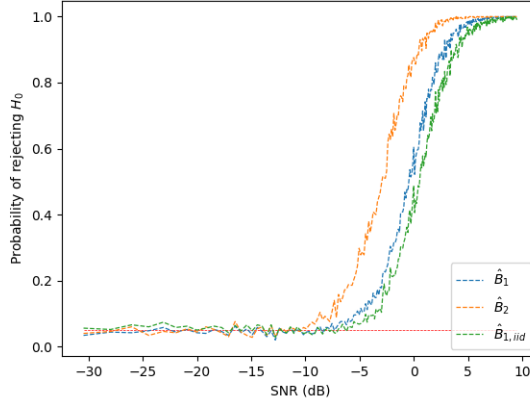


Figure 2: Empirical rejection rate at  $\alpha = 5\%$  (in red dashed horizontal line) for 300 SNR values in logarithmic scale (dB)

## 10. Concluding remarks

Mardia’s multivariate kurtosis,  $\hat{B}_p$ , is intended to test the *joint normality* when statistically independent realizations are available. Without assuming the latter independence, we derive in this paper the asymptotic distribution of the multivariate kurtosis under the null hypothesis. Limited by the length of the expressions for  $p > 3$ , the exact expressions are reported only in the bivariate case.

There are many ways to construct non-Gaussian processes with Gaussian marginals, as illustrated by copulas, and scalar tests often lead to misdetections, whereas our test continues to be powerful. Our test also proves to be useful for scalar processes, for example by testing the joint normality of successive values of a time-series.

## References

- Andrews, D.F., Gnanadesikan, R., Warner, J.L., 1973. Methods for assessing multivariate normality, in: Krishnaiah, P.R. (Ed.), *Multivariate Analysis III*. Academic press, pp. 95–116.
- Basseville, M., Nikiforov, I., 1993. *Detection of Abrupt Changes, Theory and Application*. Information and System Sciences Series, Prentice-Hall, Englewood Cliffs.

- Bowman, K.O., Shenton, L.R., 1975. Omnibus contours for departures from normality based on  $b_1$  and  $b_2$ . *Biometrika* 62, 243–250.
- Brillinger, D.R., 1981. *Time Series, Data Analysis and Theory*. Holden-Day.
- Cichocki, A., Amari, S.I., 2002. *Adaptive Blind Signal and Image Processing*. Wiley, New York.
- Comon, P., Deruaz, L., 1995. Normality tests for coloured samples, in: *IEEE-ATHOS Workshop on Higher-Order Statistics*, Begur, Spain. pp. 217–221.
- Comon, P., Jutten, C. (Eds.), 2010. *Handbook of Blind Source Separation, Independent Component Analysis and Applications*. Academic Press, Oxford UK, Burlington USA.
- Cramér, H., 1946. A contribution to the theory of statistical estimation. *Scandinavian Actuarial Journal* 1946, 85–94. doi:10.1080/03461238.1946.10419631.
- ElBouch, S., 2021. Supplementary material. URL: <https://hal.archives-ouvertes.fr/hal-03343508>. working paper or preprint.
- Epps, T.W., 1987. Testing that a stationary time series is Gaussian. *The Annals of Statistics* 15, 1683–1698.
- Gasser, T., 1975. Goodness-of-fit tests for correlated data. *Biometrika* 62, 563–570.
- Haykin, S., 2000. *Unsupervised Adaptive Filtering*. volume 1 & 2. Wiley. Series in Adaptive and Learning Systems for Communications, Signal Processing, and Control.
- Henze, R., 2002. Invariant tests for multivariate normality: a critical review. *Statistical papers* 43, 467–506.
- Hinich, M., 1982. Testing for Gaussianity and linearity of a stationary time series. *Journal of Time Series Analysis* 3, 169–176.
- Hofert, M., 2008. Sampling archimedean copulas. *Computational Statistics & Data Analysis* 52, 5163–5174.

- Kotz, S., Johnson, N.L., 1982. Encyclopedia of Statistical Sciences. Wiley.
- Malkovich, J.F., Afifi, A., 1973. On tests for multivariate normality. *Journal of the American statistical association* 68, 176–179.
- Mardia, K.V., 1970. Measures of multivariate skewness and kurtosis with applications. *Biometrika* 57, 519–530.
- Mardia, K.V., 1974. Applications of some measures of multivariate skewness and kurtosis for testing normality. *Sankhya B* 36, 115–128.
- Mccullagh, P., 1987. *Tensor Methods in Statistics. Monographs on Statistics and Applied Probability*, Chapman and Hall.
- Moore, D.S., 1971. A chi-square statistic with random cell boundaries. *The Annals of Statistics* 42, 147–156.
- Moore, D.S., 1982. The effect of dependence on chi squared tests of fit. *The Annals of Statistics* 10, 1163–1171.
- Moulines, E., Choukri, K., Charbit, M., 1992. Testing that a multivariate stationary time series is Gaussian, in: *Sixth SSAP Workshop on Stat. Signal and Array Proc.*, pp. 185–188.
- Nieto-Reyes, A., Cuesta-Albertos, J.A., Gamboa, F., 2014. A random-projection based test of gaussianity for stationary processes. *Computational Statistics & Data Analysis* 75, 124–141.
- Nikias, C.L., Petropulu, A.P., 1993. *Higher-Order Spectra Analysis. Signal Processing Series*, Prentice-Hall, Englewood Cliffs.
- Pearson, E.S., D’agostino, R.B., Bowman, K.O., 1977. Tests for departure from normality: Comparison of powers. *Biometrika* 64, 231–246.
- Shapiro, S.S., Wilk, M.B., Chen, H.J., 1968. A comparative study of various tests for normality. *American Statistical Association Journal* 63, 1343–1372.
- Steinberg, Y., Zeitouni, O., 1992. On tests for normality. *IEEE Trans. on Inf. Theory* 38, 1779–1787.

## 11. Appendices

### 11.1. McCullagh's bracket notation and expression of the higher moments under the null hypothesis

McCullagh's bracket notation (McCullagh, 1987) allows to write into a compact form a sum of terms that can be deduced from each other by generating all possible partitions of the same type. For instance, we have the following expression for fourth order moments  $M_{abcd}$  of a zero-mean multivariate normal variable with covariance  $\mathbf{S}$ :

$$M_{abcd} = S_{ab}S_{cd} + S_{ac}S_{bd} + S_{ad}S_{bc} = [3]S_{ab}S_{cd} \quad (48)$$

Moments of higher order can be found easily:

$$\text{order 6:} \quad M_{abcdef} = [15]S_{ab}S_{cd}S_{ef} \quad (49)$$

$$\text{order 8:} \quad M_{abcdefgh} = [105]S_{ab}S_{cd}S_{ef}S_{gh} \quad (50)$$

$$\text{order 10:} \quad M_{abcdefghij} = [945]S_{ab}S_{cd}S_{ef}S_{gh}S_{ij} \quad (51)$$

$$\text{order 12:} \quad M_{abcdefghijkl} = [10395]S_{ab}S_{cd}S_{ef}S_{gh}S_{ij}S_{kl} \quad (52)$$

$$\text{order 14:} \quad M_{abcdefghijklmn} = [135135]S_{ab}S_{cd}S_{ef}S_{gh}S_{ij}S_{kl}S_{mn} \quad (53)$$

$$\text{order 16:} \quad M_{abcdefghijklmnpq} = [2027025]S_{ab}S_{cd}S_{ef}S_{gh}S_{ij}S_{kl}S_{mn}S_{pq} \quad (54)$$

since it is well known that there are  $\lfloor \frac{2r!}{2^r r!} \rfloor$  terms in the moment of order  $2r$ .

### 11.2. Calculation methodology

Remind that, as introduced in Lemma 3.3,  $A_{\alpha_l \beta_l} = \mathbf{x}(\alpha_l)^T G \mathbf{x}(\beta_l)$ , where  $G$  stands for the true precision matrix of the process whose terms are  $G_{r,c}$ , and where  $(r, c) \in \{1, \dots, p\}^2$ .

Referring to the expression of  $\hat{B}_p$  or  $\hat{B}_p^2$  as derived from equation (10), it appears that the indices  $(\alpha_l, \beta_l)$  take values on a restricted set  $\mathcal{S} = \{i, j, k, \dots\}$ , and  $|\mathcal{S}| \ll N$ . The following compact notation is therefore introduced

$$\mu_{r_1 \dots r_L c_1 \dots c_L}^{\alpha_1 \dots \alpha_L \beta_1 \dots \beta_L} = M_{i \eta_i j \eta_j k \eta_k \dots} \quad (55)$$

where

$$\eta_i = \sum_{l=1}^L (\mathbb{I}_{[\alpha_l=i]} + \mathbb{I}_{[\beta_l=i]}), \forall i \in \mathcal{S}$$

Note that the subscripts  $r_1, \dots, c_1, \dots$  are skipped here for sake of readability, though any permutation of the superscripts in equation (11) requests the corresponding permutation of the subscripts. It is easier to describe the general methodology by the typical example below.

*Example*

Consider the moment  $\mathbb{E}\{A_{nn}A_{nj}A_{jk}A_{kn}\}$ . According to equation (12) it will be expanded as a sum of moments of order 8 (i.e.  $L = 4$ ); using the compact notation from equation (55), we get

$$\begin{aligned} \mathbb{E}\{A_{nn}A_{nj}A_{jk}A_{kn}\} &= \sum_{((r_i, c_i)_{i=1 \dots 4})=1}^p G_{r_1 c_1} G_{r_2 c_2} G_{r_3 c_3} G_{r_4 c_4} \mu_{r_1 c_1 r_2 c_2 r_3 c_3 r_4 c_4}^{nnnj j k k n} \\ &= \sum_{((r_i, c_i)_{i=1 \dots 4})=1}^p G_{r_1 c_1} G_{r_2 c_2} G_{r_3 c_3} G_{r_4 c_4} M_{n^4 j^2 k^2} \end{aligned} \quad (56)$$

The sum involves  $2^{2L} = 64$  terms. It is reminded that the coefficients  $r_i$  or  $c_i$  indicate the coordinate of the vector process (or space coordinate, thus taking values on  $\{1, \dots, p\}$ ), whereas time indices  $n, j, k$  take values on  $\{1, \dots, N\}$ . Following McCullagh's notations, under the assumption ( $H_0$ ) that the  $p$ -dimensional process is centered and jointly Gaussian, for this particular 8-th order moment

$$M_{abcdefgh} = [105] S_{ab} S_{cd} S_{ef} S_{gh}$$

which expresses that under  $H_0$ , higher even order moments (odd-order moments are zero) may be expanded as sums of products of second order moments. It must be reminded that here,  $a, b, c, d, e, f, g, h$  stand for 'meta-indices' defined in the present example by  $(n, r_1), (n, c_1), (n, r_2), (n, c_2), (j, c_2), (j, r_3), (k, c_3), (k, r_4)$  respectively, as it appears in equation (56). Plugging the above expansion in equation (56) leads to summing over  $64 \times 105$  terms! However, in most cases of interest many terms may be grouped together and highlight the behavior of equation (12). The case  $p = 1$  is briefly sketched below as an illustration.

The case  $p = 1$  implies that  $r_i = c_i = 1 \forall i \in \{1, \dots, (L = 4)\}$ ; the particular 8-th order moment in equation (56) may be simply written as  $M_{n^4 j^2 k^2}$ , whose expansion into sum of products of second order moments will involve the following products : (as there is no ambiguity in this case, we set



$$M_{ij} \stackrel{\text{nota.}}{=} S_{ij}),$$

$S_{nn}S_{nn}S_{jj}S_{kk}$	appearing 3 times
$S_{nn}S_{nn}S_{jk}S_{jk}$	appearing 6 times
$S_{nn}S_{nj}S_{nj}S_{kk}$	appearing 12 times
$S_{nk}S_{nk}S_{nj}S_{nj}$	appearing 24 times
$S_{nj}S_{jk}S_{nk}S_{nn}$	appearing 48 times
$S_{nn}S_{nk}S_{nk}S_{jj}$	appearing 12 times

For example the number of occurrences of the term of type  $S_{nk}S_{nk}S_{nj}S_{nj}$  is given by

$$(4 \times 2 \times 3 \times 1)/2 \times (2 \times 2 \times 1 \times 1)/2 = 24$$

where  $4 \times 2$  stand for the number of possible choices for index  $i$  (one out of 4) times the number of possible choices for index  $k$  (one out of 2); then  $3 \times 1$  stand for the number of remaining possibilities to select index  $i$  times the remaining choices for  $k$ ; Division by 2 accounts for the fact that permutations of terms  $S_{ik}$  were counted twice. All other occurrence calculations follow the same guidelines. Finally, one gets for the case  $p = 1$

$$M_{n^4 j j k k} = 3S_{nn}^2 S_{jj} S_{kk} + 6S_{nn}^2 S_{jk} S_{jk} + 12S_{nn} S_{ij}^2 S_{kk} + 24S_{nk}^2 S_{nj}^2 + \dots \\ 48S_{nj} S_{jk} S_{nk} S_{nn} + 12S_{nn} S_{nk}^2 S_{jj}$$

which can be directly plugged into equation (56). Note that the sum of all coefficient is actually 105, as expected for an 8-th order moment.

The cases  $p \geq 2$  turns out to be a bit more complicated, as one has to deal with the 'meta-indices' directly. However counting the number of configurations involving the same time indices follows the same lines as in the case  $p = 1$ . Going back to the example introduced above for  $p = 2$ , one gets

$$\mathbb{E}\{A_{nn}A_{nj}A_{jk}A_{kn}\} = \sum_{((r_i, c_i)_{i=1 \dots 4})=1}^p G_{r_1 c_1} G_{r_2 c_2} G_{r_3 c_3} G_{r_4 c_4} \{ [3] \mu_{r_1 c_1}^{nn} \mu_{r_2 c_2}^{nn} \mu_{c_2 r_3}^{jj} \mu_{c_3 r_4}^{kk} + \\ [6] \mu_{r_1 c_1}^{nn} \mu_{r_2 c_2}^{nn} \mu_{c_2 c_3}^{jk} \mu_{r_3 r_4}^{jk} + [12] \mu_{r_1 c_1}^{nn} \mu_{r_2 c_2}^{nj} \mu_{c_4 r_3}^{nj} \mu_{c_3 r_4}^{kk} + [24] \mu_{r_1 c_3}^{nk} \mu_{c_1 r_4}^{nk} \mu_{r_2 c_2}^{nj} \mu_{c_4 r_3}^{nj} + \\ [48] \mu_{r_1 c_2}^{nj} \mu_{r_3 c_3}^{jk} \mu_{c_1 r_4}^{nk} \mu_{r_2 c_4}^{nn} + [12] \mu_{r_1 c_1}^{nn} \mu_{r_2 c_3}^{nk} \mu_{c_4 r_4}^{nk} \mu_{c_2 r_3}^{jj} \}$$

where we have used notations  $\mu_{rc}^{\alpha\beta}$  to emphasize that the permutations (whose number is indicated using McCullagh's brackets) are applied on the

'meta-indices' and grouped such that they share the same 'time structure'; This allow to get the same values as in the case  $p = 1$ , though replacing the scalar coefficients by McCullagh's brackets.

### 11.3. Multivariate moments up to order 12

In this section, we give all moments of a zero-mean multivariate normal variable of even order. Most of these expressions have not been reported in the literature. In addition, for the sake of readability, when an index is repeated more than three times, we assume an alternative notation, for instance at order 10:

$$M_{iiiijjjjk} = M_{i^5j^4k}$$

Furthermore, we use notation introduced in (55) involving meta-indices; more precisely, since each subscript is always associated with a superscript, we may omit the subscript. In order to lighten notation, especially when terms need to be raised to a power, we put the latter superscript in subscript. For instance in (57),  $M_{abcd}^{iiij}$  is replaced by  $M_{iiij}$ . In the list below, moments are sorted by increasing  $D$ , where  $D$  denotes the number of distinct indices.

#### Order 4, D=2.

$$M_{iiij} = [3]\mu_{ab}^{ii}\mu_{cd}^{ij} \quad (57)$$

$$M_{iijj} = [2]\mu_{ab}^{ij}\mu_{cd}^{ij} + \mu_{ab}^{ii}\mu_{cd}^{jj} \quad (58)$$

#### Order 4, D=3.

$$M_{iijk} = \mu_{ab}^{ii}\mu_{cd}^{jk} + [2]\mu_{ab}^{ij}\mu_{cd}^{ik} \quad (59)$$

#### Order 6, D=2.

$$M_{i^5j} = [15]\mu_{ab}^{ii}\mu_{cd}^{ii}\mu_{ef}^{ij} \quad (60)$$

$$M_{i^4jj} = [12]\mu_{ae}^{ij}\mu_{bf}^{ij}\mu_{cd}^{ii} + [3]\mu_{ab}^{ii}\mu_{cd}^{ii}\mu_{ef}^{jj} \quad (61)$$

$$M_{iiiijj} = [6]\mu_{ad}^{ij}\mu_{be}^{ij}\mu_{df}^{ij} + [9]\mu_{ab}^{ii}\mu_{cd}^{ij}\mu_{ef}^{jj} \quad (62)$$

#### Order 6, D=3.

$$M_{i^4jk} = [3]\mu_{ab}^{ii}\mu_{cd}^{ii}\mu_{ef}^{jk} + [12]\mu_{ae}^{ij}\mu_{bf}^{ik}\mu_{cd}^{ii} \quad (63)$$

$$M_{iiiijk} = [6]\mu_{ad}^{ij}\mu_{be}^{ij}\mu_{cf}^{ik} + [6]\mu_{ad}^{ij}\mu_{bc}^{ii}\mu_{ef}^{jk} + [3]\mu_{ab}^{ii}\mu_{de}^{jj}\mu_{cf}^{ik} \quad (64)$$

$$M_{iijjkk} = \mu_{ab}^{ii}\mu_{cd}^{jj}\mu_{ef}^{kk} + [2]\mu_{ab}^{ii}\mu_{ce}^{jk}\mu_{bf}^{jk} + [2]\mu_{cd}^{jj}\mu_{ae}^{ik}\mu_{bf}^{ik} + [2]\mu_{ef}^{kk}\mu_{ac}^{ij}\mu_{bd}^{ij} \\ + [8]\mu_{ac}^{ij}\mu_{de}^{jk}\mu_{bf}^{ik} \quad (65)$$

**Order 8, D=2.**

$$M_{i^7j} = [105]\mu_{ab}^{ii}\mu_{cd}^{ii}\mu_{ef}^{ii}\mu_{gh}^{ij} \quad (66)$$

$$M_{i^6jj} = [90]\mu_{ag}^{ij}\mu_{bh}^{ij}\mu_{cd}^{ii}\mu_{ef}^{ii} + [15]\mu_{ab}^{ii}\mu_{cd}^{ii}\mu_{ef}^{ii}\mu_{gh}^{jj} \quad (67)$$

$$M_{i^5jjj} = [60]\mu_{af}^{ij}\mu_{bg}^{ij}\mu_{ch}^{ij}\mu_{de}^{ii} + [45]\mu_{ab}^{ii}\mu_{cd}^{ii}\mu_{ef}^{ij}\mu_{gh}^{jj} \quad (68)$$

$$M_{i^4j^4} = [9]\mu_{ab}^{ii}\mu_{cd}^{ii}\mu_{ef}^{jj}\mu_{gh}^{jj} + [72]\mu_{ab}^{ii}\mu_{ce}^{ij}\mu_{df}^{ij}\mu_{gh}^{jj} + [24]\mu_{ae}^{ij}\mu_{bf}^{ij}\mu_{cg}^{ij}\mu_{dh}^{ij} \quad (69)$$

**Order 8, D=3.**

$$M_{i^6jk} = [15]\mu_{ab}^{ii}\mu_{cd}^{ii}\mu_{ef}^{ii}\mu_{gh}^{jk} + [90]\mu_{ab}^{ii}\mu_{cd}^{ii}\mu_{eg}^{ij}\mu_{fg}^{ik} \quad (70)$$

$$M_{i^5jjk} = [30]\mu_{ab}^{ii}\mu_{cd}^{ii}\mu_{ef}^{ij}\mu_{gh}^{jk} + [60]\mu_{af}^{ij}\mu_{bc}^{ii}\mu_{dh}^{ik} + [15]\mu_{ab}^{ii}\mu_{cd}^{ii}\mu_{fg}^{jj}\mu_{eh}^{ik} \quad (71)$$

$$M_{i^4jjjk} = [9]\mu_{ab}^{ii}\mu_{cd}^{ii}\mu_{ef}^{jj}\mu_{gh}^{jk} + [36]\mu_{ef}^{jj}\mu_{ag}^{ij}\mu_{bh}^{ik}\mu_{cd}^{ii} \\ + [24]\mu_{ae}^{ij}\mu_{bf}^{ij}\mu_{cg}^{ij}\mu_{dh}^{ik} + [36]\mu_{ab}^{ii}\mu_{ce}^{ij}\mu_{df}^{ij}\mu_{gh}^{jk} \quad (72)$$

$$M_{i^4jjkk} = [3]\mu_{ab}^{ii}\mu_{cd}^{ii}\mu_{ef}^{jj}\mu_{gh}^{kk} + [6]\mu_{ab}^{ii}\mu_{cd}^{ii}\mu_{eg}^{jk}\mu_{fh}^{jk} + [12]\mu_{ab}^{ii}\mu_{ce}^{ij}\mu_{df}^{ij}\mu_{gh}^{kk} \\ + [24]\mu_{ag}^{ik}\mu_{bh}^{ik}\mu_{ce}^{ij}\mu_{df}^{ij} + [48]\mu_{ae}^{ij}\mu_{fg}^{jk}\mu_{bh}^{ik}\mu_{cd}^{ii} \\ + [12]\mu_{ab}^{ii}\mu_{bg}^{ik}\mu_{ch}^{ik}\mu_{ef}^{jj} \quad (73)$$

$$M_{iii jjkk} = [9]\mu_{ab}^{ii}\mu_{cd}^{ij}\mu_{ef}^{jj}\mu_{gh}^{kk} + [18]\mu_{ab}^{ii}\mu_{cd}^{ij}\mu_{eg}^{jk}\mu_{fh}^{jk} + [6]\mu_{ad}^{ij}\mu_{be}^{ij}\mu_{cf}^{ij}\mu_{gh}^{kk} \\ + [18]\mu_{ag}^{ik}\mu_{bh}^{ik}\mu_{cd}^{ij}\mu_{ef}^{jj} + [36]\mu_{ad}^{ij}\mu_{be}^{ij}\mu_{cg}^{ik}\mu_{fh}^{jk} \\ + [18]\mu_{ag}^{ik}\mu_{dh}^{jk}\mu_{bc}^{ii}\mu_{ef}^{jj} \quad (74)$$

**Order 10, D=2.**

$$M_{i^9j} = [945]\mu_{ab}^{ii}\mu_{cd}^{ii}\mu_{ef}^{ii}\mu_{gh}^{ii}\mu_{ml}^{ij} \quad (75)$$

$$M_{i^8jj} = [105]\mu_{ab}^{ii}\mu_{cd}^{ii}\mu_{ef}^{ii}\mu_{gh}^{ii}\mu_{ml}^{jj} + [840]\mu_{ab}^{ii}\mu_{cd}^{ii}\mu_{ef}^{ij}\mu_{gm}^{ij}\mu_{hl}^{ij} \quad (76)$$

$$M_{i^7jjj} = [315]\mu_{ab}^{ii}\mu_{cd}^{ii}\mu_{ef}^{ij}\mu_{gh}^{ij}\mu_{ml}^{jj} + [630]\mu_{ah}^{ij}\mu_{bm}^{ij}\mu_{cl}^{ij}\mu_{de}^{ii}\mu_{fg}^{ii} \quad (77)$$

$$M_{i^6j^4} = [45]\mu_{ab}^{ii}\mu_{cd}^{ii}\mu_{ef}^{jj}\mu_{gh}^{jj}\mu_{ml}^{jj} + [360]\mu_{ag}^{ij}\mu_{bh}^{ij}\mu_{cm}^{ij}\mu_{dl}^{ij}\mu_{ef}^{ii} \\ + [540]\mu_{ab}^{ii}\mu_{cd}^{ii}\mu_{eg}^{ij}\mu_{fh}^{ij}\mu_{ml}^{jj} \quad (78)$$

$$M_{i^5j^5} = [120]\mu_{af}^{ij}\mu_{bg}^{ij}\mu_{ch}^{ij}\mu_{dm}^{ij}\mu_{el}^{ij} \\ + [225]\mu_{ab}^{ii}\mu_{cd}^{ii}\mu_{ef}^{ij}\mu_{gh}^{jj}\mu_{ml}^{jj} + [600]\mu_{fg}^{jj}\mu_{ah}^{ij}\mu_{bm}^{ij}\mu_{cl}^{ij}\mu_{de}^{ii} \quad (79)$$

Order 10, D=3.

$$M_{i^8jk} = [105]\mu_{ab}^{ii}\mu_{cd}^{ii}\mu_{ef}^{ii}\mu_{gh}^{ii}\mu_{ml}^{jk} + [840]\mu_{ab}^{ii}\mu_{cd}^{ii}\mu_{ef}^{ii}\mu_{gm}^{ij}\mu_{hl}^{ik} \quad (80)$$

$$M_{i^7jjk} = [210]\mu_{ab}^{ii}\mu_{cd}^{ii}\mu_{ef}^{ii}\mu_{gh}^{ij}\mu_{ml}^{jk} + [630]\mu_{ab}^{ii}\mu_{cd}^{ii}\mu_{eh}^{ij}\mu_{fm}^{ij}\mu_{hl}^{ik} \\ + [105]\mu_{ab}^{ii}\mu_{cd}^{ii}\mu_{ef}^{ii}\mu_{hm}^{jj}\mu_{gl}^{ik} \quad (81)$$

$$M_{i^6jjjk} = [45]\mu_{ab}^{ii}\mu_{cd}^{ii}\mu_{ef}^{ii}\mu_{gh}^{jj}\mu_{ml}^{jk} + [270]\mu_{ab}^{ii}\mu_{cd}^{ii}\mu_{eg}^{ij}\mu_{fh}^{ij}\mu_{ml}^{jk} \\ + [360]\mu_{ag}^{ij}\mu_{bh}^{ij}\mu_{cm}^{ij}\mu_{dl}^{ik}\mu_{ef}^{ii} + [270]\mu_{al}^{ik}\mu_{gh}^{jj}\mu_{bl}^{ij}\mu_{cd}^{ii}\mu_{ef}^{ii} \quad (82)$$

$$M_{i^5jjkk} = [15]\mu_{ab}^{ii}\mu_{cd}^{ii}\mu_{ef}^{ii}\mu_{gh}^{jj}\mu_{ml}^{kk} + [30]\mu_{ab}^{ii}\mu_{cd}^{ii}\mu_{ef}^{ii}\mu_{gm}^{jk}\mu_{hl}^{jk} \\ + [90]\mu_{ab}^{ii}\mu_{cd}^{ii}\mu_{eg}^{ij}\mu_{fh}^{ij}\mu_{ml}^{kk} + [90]\mu_{ab}^{ii}\mu_{cd}^{ii}\mu_{em}^{ik}\mu_{fl}^{ik}\mu_{gh}^{jj} \\ + [360]\mu_{ab}^{ii}\mu_{cg}^{ij}\mu_{dh}^{ij}\mu_{em}^{ik}\mu_{fl}^{ik} + [360]\mu_{ab}^{ii}\mu_{cd}^{ii}\mu_{eg}^{ij}\mu_{fm}^{ik}\mu_{hl}^{jk} \quad (83)$$

$$M_{i^5j^4k} = [45]\mu_{ab}^{ii}\mu_{cd}^{ii}\mu_{fg}^{jj}\mu_{hm}^{jj}\mu_{el}^{ik} + [360]\mu_{ab}^{ii}\mu_{cf}^{ij}\mu_{dg}^{jj}\mu_{hm}^{ij}\mu_{el}^{ik} \\ + [120]\mu_{ag}^{ij}\mu_{bf}^{ij}\mu_{cg}^{ij}\mu_{dh}^{ij}\mu_{el}^{ik} + [180]\mu_{ab}^{ii}\mu_{cd}^{ii}\mu_{ef}^{ij}\mu_{gh}^{jj}\mu_{ml}^{jk} \\ + [240]\mu_{ab}^{ii}\mu_{cf}^{ij}\mu_{dg}^{ij}\mu_{eh}^{ij}\mu_{ml}^{jk} \quad (84)$$

$$M_{i^5jjjkk} = [45]\mu_{ab}^{ii}\mu_{cd}^{ii}\mu_{ef}^{ij}\mu_{gh}^{jj}\mu_{ml}^{kk} + [60]\mu_{ab}^{ii}\mu_{cf}^{ij}\mu_{dg}^{ij}\mu_{eh}^{ij}\mu_{ml}^{kk} \\ + [90]\mu_{ab}^{ii}\mu_{cd}^{ii}\mu_{ef}^{ij}\mu_{gm}^{jk}\mu_{hl}^{jk} + 360\mu_{ab}^{ii}\mu_{cf}^{ij}\mu_{dg}^{ij}\mu_{em}^{ik}\mu_{hl}^{jk} \\ + [90]\mu_{ab}^{ii}\mu_{fg}^{jj}\mu_{cm}^{ik}\mu_{hl}^{jk} + [180]\mu_{ab}^{ii}\mu_{cf}^{ij}\mu_{gh}^{jj}\mu_{dm}^{ik}\mu_{el}^{ik} \\ + [120]\mu_{af}^{ij}\mu_{bg}^{ij}\mu_{ch}^{ij}\mu_{dm}^{ik}\mu_{el}^{ik} \quad (85)$$

$$M_{i^4jjjkkk} = [27]\mu_{ab}^{ii}\mu_{cd}^{ii}\mu_{ef}^{jj}\mu_{gh}^{jk}\mu_{ml}^{kk} + [18]\mu_{ab}^{ii}\mu_{cd}^{ii}\mu_{eh}^{jk}\mu_{fm}^{jk}\mu_{gl}^{ij} \\ + [108]\mu_{ab}^{ii}\mu_{ce}^{ij}\mu_{df}^{ij}\mu_{gh}^{jk}\mu_{ml}^{kk} + [108]\mu_{ab}^{ii}\mu_{ef}^{jj}\mu_{gh}^{jk}\mu_{cm}^{ik}\mu_{dl}^{ik} \\ + [108]\mu_{ab}^{ii}\mu_{ef}^{jj}\mu_{cg}^{ij}\mu_{dh}^{ik}\mu_{ml}^{kk} + [216]\mu_{ab}^{ii}\mu_{ce}^{ij}\mu_{dh}^{ik}\mu_{fm}^{jk}\mu_{gl}^{jk} \\ + [72]\mu_{ae}^{ij}\mu_{bf}^{ij}\mu_{cg}^{ij}\mu_{dh}^{ik}\mu_{ml}^{kk} + [216]\mu_{ah}^{ik}\mu_{bm}^{ik}\mu_{ce}^{ij}\mu_{df}^{ij}\mu_{gl}^{jk} \\ + [72]\mu_{ah}^{ik}\mu_{bm}^{ik}\mu_{cl}^{ik}\mu_{de}^{ij}\mu_{fg}^{jj} \quad (86)$$

$$M_{i^4j^4kk} = [9]\mu_{ab}^{ii}\mu_{cd}^{ii}\mu_{ef}^{jj}\mu_{gh}^{jj}\mu_{ml}^{kk} + [72]\mu_{ab}^{ii}\mu_{ce}^{ij}\mu_{df}^{ij}\mu_{gh}^{jj}\mu_{ml}^{kk} \\ + [24]\mu_{ae}^{ij}\mu_{bf}^{ij}\mu_{cg}^{ij}\mu_{dh}^{ik}\mu_{ml}^{kk} + [36]\mu_{ab}^{ii}\mu_{ef}^{jj}\mu_{gm}^{jk}\mu_{hl}^{jk} \\ + [144]\mu_{ab}^{ii}\mu_{ce}^{ij}\mu_{df}^{ij}\mu_{gm}^{jk}\mu_{hl}^{jk} + [36]\mu_{ab}^{ii}\mu_{ef}^{jj}\mu_{gh}^{ij}\mu_{cm}^{ik}\mu_{dl}^{ik} \\ + [144]\mu_{ae}^{ij}\mu_{bf}^{ij}\mu_{gh}^{jj}\mu_{cm}^{ik}\mu_{dl}^{ik} + [288]\mu_{ab}^{ii}\mu_{ef}^{jj}\mu_{cg}^{ij}\mu_{dm}^{ik}\mu_{hl}^{jk} \\ + [192]\mu_{ae}^{ij}\mu_{bf}^{ij}\mu_{cg}^{ij}\mu_{dm}^{ik}\mu_{hl}^{jk} \quad (87)$$

11.4. Particular results when  $p = 1$

Here we remind that  $\mu_{11}^{ij} = S_{ij}$ .

**Order 12, p=1, D=2.**

$$M_{i^{11}j} = 10395S_{ii}^5S_{ij} + 9450S_{ii}^4S_{ij}^2 \quad (88)$$

$$M_{i^9jjj} = 2835S_{ii}^4S_{ij}S_{jj} + 7560S_{ii}^3S_{ij}^3 \quad (89)$$

$$M_{i^8j^4} = 5040S_{ij}^4S_{ii}^2 + 315S_{ii}^4S_{jj}^2 + 5040S_{ii}^3S_{ij}^2S_{jj} \quad (90)$$

$$M_{i^7j^5} = 1575S_{ii}^3S_{ij}S_{jj}^2 + 6300S_{ii}^2S_{ij}^3S_{jj} + 2520S_{ii}S_{ij}^5 \quad (91)$$

$$M_{i^6j^6} = 720S_{ij}^6 + 225S_{ii}^3S_{jj}^3 + 5400S_{ii}S_{ij}^4S_{jj} + 4050S_{ii}^2S_{ij}^2S_{jj}^2 \quad (92)$$

**Order 12, p=1, D=3.**

$$M_{i^{10}jk} = 945S_{ii}^5S_{jk} + 9450S_{ik}S_{ij}S_{ii}^4 \quad (93)$$

$$M_{ijjk} = 945S_{ii}^4S_{jj}S_{ik} + 7560S_{ii}^3S_{ij}^2S_{ik} + 1890S_{ii}^4S_{ij}S_{jk} \quad (94)$$

$$M_{i^8jjjk} = 315S_{ii}^4S_{jj}S_{jk} + 2520S_{ii}^3S_{ij}S_{jj}S_{ik} + 2520S_{ii}^3S_{ij}^2S_{jk} \\ + 5040S_{ii}^2S_{ij}^3S_{ik} \quad (95)$$

$$M_{i^7j^4k} = 315S_{ii}^3S_{jj}^2S_{ik} + 3780S_{ii}^2S_{ij}^2S_{ik} + 1260S_{ii}S_{ij}^4S_{ik} + 1260S_{ii}^3S_{ij}S_{jj}S_{jk} \\ + 3780S_{ii}^2S_{ij}^3S_{jk} \quad (96)$$

$$M_{i^8jjkk} = 105S_{ii}^4S_{jj}S_{kk} + 210S_{ii}^4S_{jk}^2 + 840S_{ii}^3S_{ij}^2S_{kk} + 840S_{ii}^3S_{ik}^2S_{jj} \\ + 5040S_{ii}^2S_{ij}^2S_{ik}^2 + 3360S_{ii}^3S_{ik}S_{ij}S_{jk} \quad (97)$$

**Order 12, p=1, D=4.**

$$M_{i^4j^4kkll} = 3S_{ii}^2[3S_{jj}^2S_{kk}S_{ll} + 6S_{jj}^2S_{kl}^2 + 12S_{jj}S_{jk}^2S_{ll} + 24S_{jl}^2S_{jk}^2 \\ + 48S_{jk}S_{kl}S_{jl}S_{jj} + 12S_{jj}S_{jl}^2S_{kk}] + 3S_{jj}^2[12S_{ii}S_{ik}^2S_{ll} \\ + 24S_{il}^2S_{ik}^2 + 48S_{ik}S_{kl}S_{il}S_{ii} + 12S_{ii}S_{il}^2S_{kk}] \\ + 24S_{ij}^4S_{kk}S_{ll} + 48S_{ij}^4S_{kl}^2 + 96S_{ij}^3[2S_{ik}S_{jk}S_{ll} + 2S_{il}S_{jl}S_{kk} \\ + 4S_{ik}S_{jl}S_{kl} + 4S_{il}S_{jk}S_{lk}] + 72S_{ij}^2[4S_{ik}^2S_{jl}^2 + 4S_{jk}^2S_{il}^2 \\ + 16S_{ik}S_{il}S_{jk}S_{jl} + S_{ii}S_{jj}S_{kk}S_{ll} + 2S_{ii}S_{jj}S_{kl}^2] + 12S_{ik}^2[12S_{ji}^2 \\ \times S_{jj}S_{ll} + 48S_{ij}S_{il}S_{jl}S_{jj} + 12S_{jj}S_{jl}^2S_{ii} + 12S_{il}^2[12S_{ji}^2S_{jj}S_{kk} \\ + 48S_{ij}S_{ik}S_{jk}S_{jj} + 12S_{jj}S_{jk}^2S_{ii}] + 12S_{jl}^2[12S_{ij}^2S_{ii}S_{kk} \\ + 48S_{ij}S_{jk}S_{ik}S_{ii}] + 12S_{jk}^2[12S_{ij}^2S_{ii}S_{ll} + 48S_{ij}S_{jl}S_{il}S_{ii}] \\ + 576S_{ii}[S_{ik}S_{il}S_{jj}S_{jk}S_{jl} + S_{ik}S_{ij}S_{jj}S_{jl}S_{kl} \\ + S_{ik}S_{ij}S_{jj}S_{jk}S_{ll} + S_{il}S_{ij}S_{jj}S_{jk}S_{lk} \\ + S_{il}S_{ij}S_{jj}S_{jl}S_{kk} + S_{ik}S_{ij}S_{jj}S_{lk}S_{jl}] \quad (98)$$

11.5. Computation of the mean of  $\hat{B}_p(N)$

The first step is to unfold McCullagh's bracket notation to have the explicit summation terms. For instance:

$$\mathbb{E}\{A_{nn}^2\} = \sum_{a,b=1}^p \sum_{c,d=1}^p G_{ab}G_{cd}(S_{ab}S_{cd} + S_{ac}S_{bd} + S_{ad}S_{bc}) \quad (99)$$

**For p=1.**

$$\mathbb{E}\{A_{nn}^2\} = 3 \quad (100)$$

$$\mathbb{E}\{A_{nn}A_{ni}^2\} = 3 + 12\frac{S(n-i)^2}{S^2} \quad (101)$$

$$\begin{aligned} \mathbb{E}\{A_{ni}^2A_{nj}^2\} &= 3 + 6\frac{S(i-j)^2}{S^2} + 12\frac{S(n-i)^2}{S^2} + 12\frac{S(n-j)^2}{S^2} \\ &\quad + 24\frac{S(n-i)^2S_{11}(n-j)^2}{S^4} + 48\frac{S(n-i)S(n-j)S(i-j)}{S^3} \end{aligned}$$

$$\begin{aligned} \mathbb{E}\{A_{nn}A_{nj}A_{jk}A_{kn}\} &= 3 + 6\frac{S(j-k)^2}{S^2} + 12\frac{S(n-k)^2}{S^2} + 12\frac{S(n-j)^2}{S^2} \\ &\quad + 24\frac{S(n-k)^2S(n-j)^2}{S^4} + 48\frac{S(n-k)S(j-k)S(n-j)}{S^3} \end{aligned} \quad (102)$$

The exact computation of  $\mathbb{E}\{\hat{B}_1\}$  yields the following result:

$$\begin{aligned} \mathbb{E}\{\hat{B}_1\} &= 3 - \frac{6}{N^2} \sum_{n,i} \frac{S(n-i)^2}{S^2} + \frac{72}{N^3} \sum_{n,i,j} \frac{S(n-j)^2S(n-i)^2}{S^4} \\ &\quad + \frac{144}{N^3} \sum_{n,i,j} \frac{S(n-i)S(i-j)S(n-j)}{S^3} \end{aligned} \quad (103)$$

Based on the results in (Cramér, 1946, p. 346-347), it can be shown that  $\frac{1}{N^3} \sum_{n,i,j} \frac{S(n-j)^2S(n-i)^2}{S^4}$  and  $\frac{1}{N^3} \sum_{n,i,j} \frac{S(n-i)S(i-j)S(n-j)}{S^3}$  will contribute quantities of order lower than  $N^{-1}$ .

**For  $p=2$ .**

$$\mathbb{E}\{A_{nn}^2\} = 8 \quad (104)$$

$$\begin{aligned} \mathbb{E}\{A_{nn}A_{ni}^2\} &= 8 + \frac{1}{(S_{11}S_{22} - S_{12}^2)^2} \left[ S_{11}S_{22} [2S_{12}^2(n-i) + 2S_{12}^2(n-i)] \right. \\ &\quad + 12S_{21}(n-i)S_{12}(n-i) + 12S_{22}(n-i)S_{11}(n-i) \\ &\quad + S_{12}^2 [12S_{12}^2(n-i) + 12S_{21}(n-i)S_{12}(n-i) \\ &\quad + 12S_{21}^2(n-i) + 16S_{11}(n-i)S_{22}(n-i)] \\ &\quad - 28S_{12}S_{11} [S_{22}(n-i)S_{12}(n-i) + S_{22}(n-i)S_{12}(n-i)] \\ &\quad - 28S_{12}S_{22} [S_{11}(n-i)S_{12}(n-i) + S_{11}(n-i)S_{12}(n-i)] \\ &\quad \left. + 14S_{11}^2S_{22}^2(n-i) + 14S_{22}^2S_{11}^2(n-i) \right] \quad (105) \end{aligned}$$

*Bivariate embedding.*

$$\begin{aligned} \mathbb{E}\{A_{nn}A_{ni}^2\} &= 8 + \frac{1}{(C_0^2 - C_1^2)^2} \times [C_0^2 \times [2(\gamma_1(n-i) + \gamma_{-1}(n-i))^2 \\ &\quad + 8\gamma_{-1}(n-i)\gamma_1(n-i) + 40\gamma_0(n-i)^2] + C_1^2 \\ &\quad \times [12(\gamma_1(n-i) + \gamma_{-1}(n-i))^2 - 8\gamma_{-1}(n-i)\gamma_1(n-i) + 16\gamma_0(n-i)^2] \\ &\quad - C_0C_1 \times [56\gamma_0(n-i)(\gamma_1(n-i) + \gamma_{-1}(n-i))] \quad (106) \end{aligned}$$

$$\begin{aligned} \mathbb{E}\{A_{ni}^2A_{nj}^2\} &= 8 + \frac{1}{(C_0^2 - C_1^2)^2} \times [C_0^2 \times [2(\gamma_1(i-j) + \gamma_{-1}(i-j))^2 \\ &\quad + 2(\gamma_1(n-i) + \gamma_{-1}(n-i))^2 + 2(\gamma_1(n-j) + \gamma_{-1}(n-j))^2 \\ &\quad + 8\gamma_{-1}(n-i)\gamma_1(n-i) + 8\gamma_{-1}(n-j)\gamma_1(n-j) + 16\gamma_0(i-j)^2 \\ &\quad + 40\gamma_0(n-j)^2 + 40\gamma_0(n-i)^2] + C_1^2 \times [4(\gamma_1(i-j) + \gamma_{-1}(i-j))^2 \\ &\quad + 12(\gamma_1(n-j) + \gamma_{-1}(n-j))^2 + 12(\gamma_1(n-i) + \gamma_{-1}(n-i))^2 \\ &\quad - 8\gamma_{-1}(n-i)\gamma_1(n-i) - 8\gamma_{-1}(n-j)\gamma_1(n-j) + 8\gamma_0(i-j)^2 \\ &\quad + 16\gamma_0(n-j)^2 + 16\gamma_0(n-i)^2] - C_0C_1 \times [24\gamma_0(i-j) \\ &\quad \times (\gamma_1(i-j) + \gamma_{-1}(i-j)) + 56\gamma_0(n-i)(\gamma_1(n-i) + \gamma_{-1}(n-i)) \\ &\quad + 56\gamma_0(n-j)(\gamma_1(n-j) + \gamma_{-1}(n-j))] \quad (107) \end{aligned}$$

Following the same pattern as the mean, but with more moments involved, the computation of the variance may be conducted [ElBouch (2021)].

### 11.6. Sklar's theorem

**Theorem 11.1.** (Sklar's theorem 1959)

$$F_{X_1, X_2}(x_1, x_2) = Pr(X_1 \leq x_1, X_2 \leq x_2) = \mathcal{C}(F(x_1), G(x_2)) \quad (108)$$

where  $F_{X_1, X_2}$  is the joint cumulative distribution function (cdf) of  $(X_1, X_2)$ , and  $F$  (resp.  $G$ ) is the cdf of  $X_1$  (resp.  $X_2$ ). If  $F, G$  are continuous, then  $\mathcal{C}$  is unique, and is defined by:

$$\mathcal{C}(u_1, u_2) = F_{X_1, X_2}(F^{-1}(u_1), G^{-1}(u_2)). \quad (109)$$