

A Normality Test for Multivariate Dependent Samples

Sara ElBouch and Olivier Michel and Pierre Comon

^a*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 and a mixing condition. The calculus is long and tedious but the final result is simple and is implemented with a low computational burden. The proposed expression of the test power exhibits good properties on various scenarios; this is illustrated by computer experiments by means of copulas.

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

1. Introduction

The interest in techniques involving higher order statistics has grown considerably during the past decades [1, 2, 3, 4]. Actually, first and second order statistics allow an exhaustive characterization of Gaussian processes and linear systems. Despite the practical importance of the Gaussian distribution, thanks to the central limit theorem, and the prevalence of linear dynamical systems in small fluctuations models, many situations do not resort to these assumptions. As a consequence, detecting departure from Gaussianity arose as a means to detect and characterize non linear behavior, detection of changes in dynamical regimes [5], etc. Higher-Order Statistics (HOS) were also shown to carry valuable information for blind identification problems, source separation and in measuring information theoretic quantities [4], to name a few applications.

The present growth of interest in sensor networks and our ability to simultaneously record time series representing the fluctuations of numerous

physical quantities, naturally leads to consider p -dimensional processes. Surprisingly enough, normality tests for such p -dimensional stochastic processes were not so much investigated. To be more precise, very few results concern both the multivariate nature of the time series and the fact that the p -dimensional time samples cannot in general be considered as being i.i.d. This will be referred to as the *non independent* identically distributed (n.i.d.) property. The difficulty in testing the Gaussian nature of p -dimensional stochastic processes arises from the necessity for the test to tackle both the joint Gaussianity of the p components and the time dependence of successive p -dimensional samples. To make this framework clear, the following notation is introduced: Let $\mathbf{x}(n) = [x_1(n), \dots, x_p(n)]^T$ be a real p -variate stochastic process, of which a sample of finite size, N , is observed. Hereafter, the stochastic processes under study will be assumed stationary and zero-mean with covariance matrices for delay τ defined by

$$\mathbf{S}(\tau) = \mathbb{E}\{\mathbf{x}(n)\mathbf{x}(n - \tau)^T\}. \quad (1)$$

Let $S_{ab}(\tau)$ denote the entries of matrix $\mathbf{S}(\tau)$, $(a, b) \in \{1, \dots, p\}$. The n.i.d. nature of the time samples corresponds to have $S_{ab}(\tau) \neq 0$ in general for $\tau \neq 0$. Thus, the process $X(n)$ enjoys both *spatial* and *temporal* dependencies (temporal refers to dependence w.r.t. τ and spatial w.r.t. to $(a, b) \in \{1, \dots, p\}$). The normality test *without alternative* can be formulated as follows:

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

$$H : \mathbf{X} \text{ is Gaussian} \quad \textit{versus} \quad \bar{H} \quad (2)$$

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

Solving this problem requires (i) to define a test variable, and (ii) to determine its asymptotic distribution (often itself normal) in order to assess the power of the test, that is, the probability to decide H whereas H is true.

For the scalar case ($p = 1$), since the so-called Chi-squared test proposed by Fisher and improved in [6], the most popular test is probably the omnibus test based on skewness and kurtosis [7]. The omnibus test combines estimated skewness and kurtosis weighted by the inverse of their respective

asymptotic variance, evaluated under the assumption that the samples are Gaussian i.i.d.; see also [8], [9]. The asymptotic distribution of the test is χ_2^2 when samples are i.i.d normal. However, as pointed out by [10], the Chi-square test is very sensitive to the dependence between samples; the process color yields a loss in apparent normality [11]. Actually, most of the tests proposed in the literature assume that observations are i.i.d., see [12] or [13]. This is also true for multivariate tests [14, 15]; see the survey of [16]. Only very few authors address the case of n.i.d samples, or so-called colored processes. One can mention Hinich’s bispectrum-based linearity test [17], or Brillinger’s trispectrum [18]. These multispectra (Fourier transform of 3rd and 4th order cumulant multicorrelations evaluated under stationarity assumption) induce in general an important computational load and have important estimation variance even for large values of N . An appealing alternative was proposed in [19] where non linear transforms of the samples allow to go beyond monomials of degree 3 or 4. For instance, some tests are based on the characteristic function [20, 19] and others on entropy [21]. These tests remain however complex to implement in practice and may hardly be executable in real time on a light processor when samples are colored (*i.e.* statistically time dependant).

Contribution. Taking an opposite direction, the purpose of this paper is to propose a normality test that is simple to implement, even for colored (time correlated) p -dimensional processes, eventually at the expense of quite complicated and lengthy calculus to derive the exact form of the test. For this reason, we shall focus on the multivariate kurtosis proposed by Mardia in [14], and derive its mean and variance when samples are assumed to be statistically n.i.d. Although the results presented apply easily in practice to a very wide class of applications, the process is assumed to satisfy statistical *mixing* properties. Within this framework, a general procedure is proposed to compute the asymptotic mean and variance of Mardia’s multivariate kurtosis when applied to *colored* processes, and is shown to converge in $O(1/\sqrt{N})$, thus fully characterizing the asymptotic normal distribution of the test statistic. The complete derivation is given for $p = 2$, which allows to test joint normality of arbitrary 2-dimensional projections, thus generalizing the tests proposed in [14, 22, 23] based on 1D projections and i.i.d samples. The benefits of using 2D is clear, as the resulting tests are subsequently shown to outperform 1D projection-based tests, via computer experiments. The importance of joint normality and the performance of our test is illus-

trated on *n.i.d.* copulas, *i.e.* with colored Gaussian marginals. Additionally, the particular case where p -dimensional observations are constructed by time embedding, thus mixing time and space dependencies, is developed.

This article is organized as follows. Section 2 contains the definition of the test statistic. The tools necessary to conduct the calculations are introduced in Section 3. The moments involved in the derivation of both the mean and variance of the test statistic are given in Sections 4-5, for arbitrary dimension p . Their expression in closed form for $p = 1$ and $p = 2$, and for the case where the multivariate process arises from a time embedding, are given in Sections 6-8. Section 9 reports some computer experiments. The expressions of moments and details of calculation are deferred to appendices in Section 11.

2. Mardia's Multivariate kurtosis

The test proposed by Mardia in [14] takes the form:

$$\beta_p = \mathbb{E}\{(\mathbf{x}^T \mathbf{S}^{-1} \mathbf{x})^2\}. \quad (3)$$

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:

$$B_p(N) = \frac{1}{N} \sum_{n=1}^N (\mathbf{x}(n)^T \mathbf{S}^{-1} \mathbf{x}(n))^2 \quad (4)$$

It is worth noticing that \mathbf{S} being the exact covariance matrix, all random realizations involved in the latter equation are standardized (remind that we assume zero-mean processes). Thus, the 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:

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

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 η 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 β_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 . 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 1. [14] *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 $\frac{8p(p+2)}{N} + o(\frac{1}{N})$.*

Note that the result above makes use of Landau notation $o(\frac{1}{N})$, to precise that the absolute approximation error is dominated by $\frac{1}{N}$. Landau notation $O(h(N))$ will also be often used, when the absolute approximation error will be of the order of $h(N)$. Both will be extensively used in the rest of this paper.

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

In this section, partial useful results are established. Each is associated with a lemma, and represents a step towards the derivation of the exact expression of the statistics of $\hat{B}_p(N)$ defined in (5) for multivariate *colored* processes:

-Lemma 3 proves that $\mathbf{\Delta} = \hat{\mathbf{S}} - \mathbf{S}$ varies as $O(1/\sqrt{N})$.

-Lemma 4 uses the preceding result in order to express the sample precision matrix $\hat{G} = (\mathbf{S} + \mathbf{\Delta})^{-1}$ as a function of the exact precision matrix \mathbf{G} and of the approximation matrix $\mathbf{\Delta}$, up to order $O(\|\mathbf{\Delta}\|^3)$.

-Finally, lemma 5 allows to derive the approximate expression of $\hat{B}_p(N)$ to order N^{-1} in $o(1/N)$.

From now on, we assume the following condition upon $\mathbf{x}(n)$, necessary to relax the i.i.d. property while maintaining convergence of various terms:

Assumption 2 (Mixing). $\sum_{\tau=0}^{\infty} |S_{ab}(\tau)|^2$ converges to a finite limit Ω_{ab} , $\forall (a, b) \in \{1, \dots, p\}^2$, where S_{ab} denote the entries of matrix \mathbf{S} .

3.1. Lemmas

The estimated multivariate kurtosis (5) 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. *The entries of matrix $\mathbf{\Delta}$ are of order $O(1/\sqrt{N})$.*

Proof. Under Hypothesis H , the covariance of entries Δ_{ab} take the form below :

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

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

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

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:

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

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

$$|Cov(\Delta_{ab}, \Delta_{cd})| \leq \frac{|\Omega_{abcd}|}{N} + \frac{1}{N} (\Omega_{ac} + \Omega_{bd} + \Omega_{ad} + \Omega_{bc}) \quad (6)$$

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

Lemma 4. *The inverse $\hat{\mathbf{G}}$ of $\hat{\mathbf{S}}$ can be approximated by*

$$\hat{\mathbf{G}} = \mathbf{G} - \mathbf{G}\mathbf{\Delta}\mathbf{G} + \mathbf{G}\mathbf{\Delta}\mathbf{G}\mathbf{\Delta}\mathbf{G} + o(1/N). \quad (7)$$

Proof. Notice that positive definite sample covariance matrix may be reexpressed as

$$\hat{\mathbf{S}} = \mathbf{S} + \mathbf{\Delta} = \mathbf{S}^{1/2} \mathbf{I} \mathbf{S}^{1/2} + \mathbf{S}^{1/2} \mathbf{S}^{-1/2} \mathbf{\Delta} \mathbf{S}^{-1/2} \mathbf{S}^{1/2}$$

Let \mathbf{E} be the symmetric matrix $\mathbf{E} = -\mathbf{S}^{-1/2} \mathbf{\Delta} \mathbf{S}^{-1/2}$. Then with this definition,

$$\hat{\mathbf{G}} = \mathbf{S}^{-1/2} (\mathbf{I} + \mathbf{E})^{-1} \mathbf{S}^{-1/2}$$

As 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 $\hat{\mathbf{G}}$, for N large enough to warrant that the spectral radius of \mathbf{E} is less than 1, 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 and taking $K = 3$ eventually yields (7). Note that the precise approximation order is $O(N^{-3/2})$, but only $o(1/N)$ will be useful in what follows. \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 (7), we obtain:

$$\hat{\mathbf{G}} = 3\mathbf{G} - 3\mathbf{G}\hat{\mathbf{S}}\mathbf{G} + \mathbf{G}\hat{\mathbf{S}}\mathbf{G}\hat{\mathbf{S}}\mathbf{G} + o(1/N). \quad (8)$$

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

$$\begin{aligned} \hat{B}_p(N) &= \frac{6}{N} \sum_{n=1}^N A_{nn}^2 - \frac{8}{N^2} \sum_{n=1}^N A_{nn} \sum_{i=1}^N A_{ni}^2 + \frac{1}{N^3} \sum_{n=1}^N \left(\sum_{i=1}^N A_{ni}^2 \right) \left(\sum_{j=1}^N A_{nj}^2 \right) \\ &\quad + \frac{2}{N^3} \sum_{n=1}^N \sum_{j=1}^N \sum_{k=1}^N A_{nn} A_{nj} A_{jk} A_{kn} + o(1/N) \end{aligned} \quad (9)$$

Proof. First inject (7) 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 $O(\|\mathbf{\Delta}\|^2)$; this yields:

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

Then replace Δ 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 (9) 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 calculus issues

When computing the mean and variance of $\hat{B}_p(N)$ given in (9), 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 [24], 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

$$\mu_{ab}^{tu} = \mathbb{E}\{x_a(t)x_b(u)\}, \quad \mu_{abc}^{tuv} = \mathbb{E}\{x_a(t)x_b(u)x_c(v)\} \quad (10)$$

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 (9) 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

$$\mathbb{E}\left\{\prod_{l=1}^L A_{\alpha^l \beta^l}\right\} = \sum_{r_1 \dots r_L, c_1 \dots c_L=1}^p \left(\prod_{i=1}^L G_{r_i, c_i} \right) \mu_{r_1 \dots r_L c_1 \dots c_L}^{\alpha_1 \dots \alpha_L \beta_1 \dots \beta_L} \quad (11)$$

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 $\hat{B}_p(N)$

According to Equation (9), we have four types of terms. The goal of this section is to provide the expectation of each of these terms. In the propositions below, all terms are developed as being sums and products of second order moments, as it is reminded that under H the process is Gaussian. Notice also that under the latter assumption, all higher-order moments of any order are finite. For sake of simplicity, Landau's approximation order $O(h(n))$ is omitted in most equations.

Lemma 6. *With the definition of A_{ij} given in Lemma 5, we have:*

$$\mathbb{E}\{A_{nn}^2\} = \sum_{a,b,c,d=1}^p G_{ab} G_{cd} \mu_{abcd}^{nnnn} \quad (12)$$

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

$$\mathbb{E}\{A_{ni}^2 A_{nj}^2\} = \sum_{a,b,c,d=1}^p \sum_{e,f,g,h=1}^p G_{ab} G_{cd} G_{ef} G_{gh} \mu_{acegbdfh}^{nnnnii jj} \quad (14)$$

$$\mathbb{E}\{A_{nn} A_{nj} A_{jk} A_{kn}\} = \sum_{a,b,c,d=1}^p \sum_{e,f,g,h=1}^p G_{ab} G_{cd} G_{ef} G_{gh} \mu_{abchdefg}^{nnnnjjkk} \quad (15)$$

Proposition 7. *Using expressions of moments given in Appendix 11.3, the expectations of the four terms defined in Lemma 6 take the form below*

$$\mathbb{E}\{A_{nn}^2\} = \sum_{klqr=1}^p G_{kl}G_{rq} \left\{ [3]\mu_{kl}^{nn}\mu_{qr}^{nn} \right\} \quad (16)$$

$$\mathbb{E}\{A_{nn}A_{ni}^2\} = \sum_{klqrst=1}^p G_{kl}G_{qr}G_{st} \left\{ [12]\mu_{kr}^{ni}\mu_{lt}^{ni}\mu_{qs}^{nn} + [3]\mu_{kl}^{nn}\mu_{qs}^{nn}\mu_{rt}^{ii} \right\} \quad (17)$$

$$\begin{aligned} \mathbb{E}\{A_{ni}^2A_{nj}^2\} = & \sum_{k,\ell,q,r} \sum_{s,t,u,v} G_{kl}G_{qr}G_{st}G_{uv} \left\{ [3]\mu_{kq}^{nn}\mu_{su}^{nn}\mu_{lr}^{ii}\mu_{tv}^{jj} \right. \\ & + [6]\mu_{kq}^{nn}\mu_{su}^{nn}\mu_{lt}^{ij}\mu_{rv}^{ij} + [12]\mu_{kq}^{nn}\mu_{sl}^{ni}\mu_{ur}^{ni}\mu_{tv}^{jj} + [24]\mu_{kt}^{nj}\mu_{qv}^{nj}\mu_{ls}^{in}\mu_{ur}^{ni} \\ & \left. + [48]\mu_{kl}^{ni}\mu_{rt}^{ij}\mu_{qv}^{nj}\mu_{su}^{nn} + [12]\mu_{kq}^{nn}\mu_{ts}^{jn}\mu_{uv}^{nj}\mu_{rl}^{ii} \right\} \quad (18) \end{aligned}$$

$$\begin{aligned} \mathbb{E}\{A_{nn}A_{nj}A_{jk}A_{kn}\} = & \sum_{m,\ell,q,r} \sum_{s,t,u,v} G_{m\ell}G_{qr}G_{st}G_{uv} \left\{ [3]\mu_{m\ell}^{nn}\mu_{qv}^{nn}\mu_{sr}^{jj}\mu_{tu}^{kk} \right. \\ & + [6]\mu_{m\ell}^{nn}\mu_{qv}^{nn}\mu_{rt}^{jk}\mu_{su}^{jk} + [12]\mu_{m\ell}^{nn}\mu_{qr}^{nj}\mu_{vs}^{nj}\mu_{tu}^{kk} + [24]\mu_{mv}^{nk}\mu_{lu}^{nk}\mu_{qr}^{nj}\mu_{vs}^{nj} \\ & \left. + [48]\mu_{mr}^{nj}\mu_{st}^{jk}\mu_{lu}^{nk}\mu_{qv}^{nn} + [12]\mu_{kl}^{nn}\mu_{qt}^{nk}\mu_{vu}^{nk}\mu_{rs}^{jj} \right\} \quad (19) \end{aligned}$$

The mean of $\hat{B}_p(N)$ then follows from (9).

5. Expression of the variance of $\hat{B}_p(N)$

From Lemma 5, we can also state what moments of A_{ij} will be required in the expression of the variance of $B_p(N)$.

Lemma 8. *By raising (9) to the second power and using the definition of*

A_{ij} given in Lemma 5, we can check that the following moments are required:

$$\mathbb{E}\{A_{nn}^2 A_{ii}^2\} = \sum_{a,b,c,d,e,f,g,h=1}^p G_{ab} G_{cd} G_{ef} G_{gh} \mu_{abcde fgh}^{nnnniiii} \quad (20)$$

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

$$\mathbb{E}\{A_{nn} A_{kk} A_{ni}^2 A_{kj}^2\} = \sum_{a,b,c,d=1}^p \sum_{e,f,g,h=1}^p \sum_{m,\ell,q,r=1}^p G_{ab} G_{cd} G_{ef} G_{gh} G_{m\ell} G_{qr} \mu_{abegcdmqf h\ell r}^{nnnnkkkkiiijj} \quad (22)$$

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

$$\mathbb{E}\{A_{nn}^2 A_{ii} A_{ij} A_{jk} A_{ki}\} = \sum_{a,b,c,d=1}^p \sum_{e,f,g,h=1}^p \sum_{m,\ell,q,r=1}^p G_{ab} G_{cd} G_{ef} G_{gh} G_{m\ell} G_{qr} \mu_{abcde fgrh m\ell q}^{nnnniiiijjkk} \quad (24)$$

$$\mathbb{E}\{A_{ni}^2 A_{nj}^2 A_{kt}^2 A_{kk}\} = \sum_{a,b,c,d=1}^p \sum_{e,f,g,h=1}^p \sum_{m,\ell,q,r=1}^p \sum_{s,u=1}^p G_{ab} G_{cd} G_{ef} G_{gh} G_{m\ell} G_{qr} G_{su} \mu_{acegmqs u b d f h\ell r}^{nnnnkkkkiiijjtt} \quad (25)$$

$$\mathbb{E}\{A_{ii}A_{it}^2A_{nn}A_{nj}A_{jk}A_{kn}\} = \sum_{a,b,c,d=1}^p \sum_{e,f,g,h=1}^p \sum_{m,\ell,q,r=1}^p \sum_{s,u=1}^p G_{ab}G_{cd}G_{ef}G_{gh}G_{ml} G_{qr}G_{su} \mu_{abceghmulqrsdf}^{iiiiinnnnjjkktt} \quad (26)$$

$$\mathbb{E}\{A_{ni}^2A_{kt}^2A_{nj}^2A_{ku}^2\} = \sum_{a,b,c,d=1}^p \sum_{e,f,g,h=1}^p \sum_{m,\ell,q,r=1}^p \sum_{s,v,w,z=1}^p G_{ab}G_{cd}G_{ef}G_{gh}G_{ml} G_{qr}G_{sv}G_{wz} \mu_{acmqegswbd\ell rf hvz}^{nnnnkkkkiiijttuu} \quad (27)$$

$$\mathbb{E}\{A_{nn}A_{nj}A_{jk}A_{kn}A_{ii}A_{it}A_{tu}A_{ui}\} = \sum_{a,b,c,d=1}^p \sum_{e,f,g,h=1}^p \sum_{m,\ell,q,r=1}^p \sum_{s,v,w,z=1}^p G_{ab}G_{cd}G_{ef}G_{gh}G_{ml} G_{qr}G_{sv}G_{wz} \mu_{abchmlqzdefgrsvw}^{nnnniiiijjkkttuu} \quad (28)$$

$$\mathbb{E}\{A_{nn}A_{nj}A_{jk}A_{kn}A_{it}^2A_{iu}^2\} = \sum_{a,b,c,d=1}^p \sum_{e,f,g,h=1}^p \sum_{m,\ell,q,r=1}^p \sum_{s,v,w,z=1}^p G_{ab}G_{cd}G_{ef}G_{gh}G_{ml} G_{qr}G_{sv}G_{wz} \mu_{abchmqswdefglrvz}^{nnnniiiijjkkttuu} \quad (29)$$

Then, as in Proposition 7, by using the results of Appendix 11.3, the moments μ_*^* could be in turn expressed as a function of second order moments. For readability, we do not substitute here these values.

Proposition 9.

$$\begin{aligned} Var\{\hat{B}_p\} &= \frac{36}{N^2} \sum_n \sum_i \mathbb{E}\{A_{nn}^2 A_{ii}^2\} - \frac{96}{N^3} \sum_j \sum_{n,i} \mathbb{E}\{A_{nn}^2 A_{ij}^2 A_{ii}\} \\ &+ \frac{64}{N^4} \sum_{n,i} \sum_{j,k} \mathbb{E}\{A_{nn} A_{kk} A_{ni}^2 A_{kj}^2\} + \frac{12}{N^4} \sum_{n,i,j,k} \mathbb{E}\{A_{kk}^2 A_{ni}^2 A_{nj}^2\} \\ &+ \frac{24}{N^4} \sum_{n,i,j,k} \mathbb{E}\{A_{nn}^2 A_{ii} A_{ij} A_{jk} A_{ki}\} - \frac{16}{N^5} \sum_{n,i} \sum_{j,k,t} \mathbb{E}\{A_{ni}^2 A_{nj}^2 A_{kt}^2 A_{kk}\} \\ &- \frac{32}{N^5} \sum_{i,t} \sum_{n,j,k} \mathbb{E}\{A_{ii} A_{it}^2 A_{nn} A_{nj} A_{jk} A_{kn}\} + \frac{1}{N^6} \sum_{n,i,j} \sum_{k,t,u} \mathbb{E}\{A_{ni}^2 A_{kt}^2 A_{nj}^2 A_{ku}^2\} \\ &+ \frac{4}{N^6} \sum_{n,j,k} \sum_{i,t,u} \mathbb{E}\{A_{nn} A_{nj} A_{jk} A_{kn} A_{ii} A_{it} A_{tu} A_{ui}\} \\ &+ \frac{4}{N^6} \sum_{n,j,k} \sum_{i,t,u} \mathbb{E}\{A_{nn} A_{nj} A_{jk} A_{kn} A_{it}^2 A_{iu}^2\} - (\mathbb{E}\{\hat{B}_p\})^2 \end{aligned} \quad (30)$$

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.

$$\mathbb{E}\{\hat{B}_1\} = 3 - \frac{6}{N} - \frac{12}{N^2} \sum_{\tau=1}^{N-1} (N - \tau) \frac{S(\tau)^2}{S^2} + o\left(\frac{1}{N}\right) \quad (31)$$

$$Var\{\hat{B}_1\} = \frac{24}{N} \left[1 + \frac{2}{N} \sum_{\tau=1}^{N-1} (N - \tau) \frac{S(\tau)^4}{S^4} \right] + o\left(\frac{1}{N}\right) \quad (32)$$

In particular in the i.i.d. case, $S(\tau) = 0$ for $\tau \neq 0$, and we get the well-known result [8] [25]:

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

The expressions of mean and variance above are identical to those given in Theorem 1, the difference being that here the ratio $\frac{N-1}{N+1}$ is replaced by its approximation of order N^{-1} , i.e. $\frac{N-1}{N+1} = 1 - \frac{2}{N} + o(1/N)$.

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 $\mu_{ab}^{ij} = S_{ab}(i - j)$.

$$\mathbb{E}\{\hat{B}_2\} = 8 - \frac{16}{N} - \frac{4}{N^2} \sum_{\tau=1}^{N-1} \frac{(N - \tau)Q_1(\tau)}{(S_{11}S_{22} - S_{12}^2)^2} + o\left(\frac{1}{N}\right) \quad (33)$$

with

$$\begin{aligned} Q_1(\tau) = & S_{11}S_{22} \left[(S_{12}(\tau) + S_{21}(\tau))^2 - 4S_{11}(\tau)S_{22}(\tau) \right] \\ & + S_{12}^2 \left[2(S_{12}(\tau) + S_{21}(\tau))^2 + 4S_{22}(\tau)S_{11}(\tau) \right] \\ & - 6S_{22}S_{12} \left(S_{11}(\tau)(S_{12}(\tau) + S_{21}(\tau)) \right) \\ & - 6S_{11}S_{12} \left(S_{22}(\tau)(S_{12}(\tau) + S_{21}(\tau)) \right) \\ & + 6S_{11}^2S_{22}^2(\tau) + 6S_{22}^2S_{11}^2(\tau). \end{aligned} \quad (34)$$

$$\boxed{Var\{\hat{B}_2\} = \frac{64}{N} + \frac{16}{N^2} \sum_{\tau=1}^{N-1} \frac{(N-\tau)Q_2(\tau)}{(S_{11}S_{22} - S_{12}^2)^4} + o(\frac{1}{N})} \quad (35)$$

with

$$\begin{aligned} Q_2(\tau) = & \left[2S_{11}^2(\tau)S_{22}^2(\tau) - 16S_{11}(\tau)S_{22}(\tau)S_{12}(\tau)S_{21}(\tau) \right. \\ & + 3(S_{21}^2(\tau) + S_{12}^2(\tau))^2 + 12S_{11}(\tau)S_{22}(\tau)(S_{12}(\tau) + S_{21}(\tau))^2 \\ & \left. - 4S_{12}^2(\tau)S_{21}^2(\tau) \right] S_{11}^2 S_{22}^2 \\ & + 2S_{11}^2 S_{12}^2 \left[8S_{11}(\tau)S_{22}(\tau) + 3(5S_{11}(\tau)S_{22}(\tau) + S_{21}(\tau)S_{12}(\tau)) \right. \\ & \left. (S_{21}(\tau) + S_{12}(\tau))^2 - 4S_{21}(\tau)S_{12}(\tau)(S_{22}^2(\tau) + S_{21}(\tau)S_{12}(\tau)) \right] \\ & + 2S_{22}^2 S_{12}^2 \left[8S_{22}(\tau)S_{11}(\tau) + 3(5S_{11}(\tau)S_{22}(\tau)) + S_{21}(\tau)S_{12}(\tau) \right. \\ & \left. (S_{21}(\tau) + S_{12}(\tau))^2 - 4S_{21}(\tau)S_{12}(\tau)(S_{11}^2(\tau) + S_{21}(\tau)S_{12}(\tau)) \right] \\ & + 3S_{11}^4 S_{22}^4(\tau) + 3S_{22}^4 S_{11}^4(\tau) \\ & + 8S_{12}^4 \left[S_{11}^2(\tau)S_{22}^2(\tau) + 4S_{11}(\tau)S_{22}(\tau)S_{12}(\tau)S_{12}(\tau) + S_{21}^2(\tau)S_{12}^2(\tau) \right] \\ & - 12S_{11}S_{12}S_{22}(\tau)(S_{12}(\tau) + S_{21}(\tau)) \left[(2S_{11}(\tau)S_{22}(\tau) + S_{21}^2(\tau) + S_{12}^2(\tau)) \right. \\ & \left. S_{11}S_{22} + 2(S_{11}(\tau)S_{22}(\tau) + S_{12}(\tau)S_{21}(\tau))S_{12}^2 \right] \\ & - 12S_{22}S_{12}S_{11}(\tau)(S_{12}(\tau) + S_{21}(\tau)) \left[(2S_{11}(\tau)S_{22}(\tau) + S_{21}^2(\tau) + S_{12}^2(\tau)) \right. \\ & \left. S_{11}S_{22} + 2(S_{11}(\tau)S_{22}(\tau) + S_{12}(\tau)S_{21}(\tau))S_{12}^2 \right]. \end{aligned} \quad (36)$$

Note that the latter expressions are complicated, but easy to implement as demonstrated in the remaining sections. Again for this case where $p = 2$, the approximation $\frac{N-1}{N+1} = 1 - \frac{2}{N} + o(1/N)$ was used in the expressions of the mean and variance of \hat{B}_2 .

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:

$$\mathbb{E}\{\hat{B}_2\} \approx 8 - \frac{16}{N} - \frac{4}{N^2} \sum_{\tau=1}^{N-1} \frac{(N-\tau)q_1(\tau)}{(C_0^2 - C_1^2)^2} \quad (37)$$

$$\text{Var}\{\hat{B}_2\} \approx \frac{64}{N} + \frac{16}{N^2} \sum_{\tau=1}^{N-1} \frac{(N-h)q_2(\tau)}{(C_0^2 - C_1^2)^4} \quad (38)$$

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

$$q_1(\tau) = \left[(\gamma_1 + \gamma_{-1})^2 + 8\gamma_0^2 \right] C_0^2 - 12C_0C_1\gamma_0(\gamma_1 + \gamma_{-1}) + \left[2(\gamma_1 + \gamma_{-1})^2 + 4\gamma_0^2 \right] C_1^2, \quad (39)$$

$$\begin{aligned} q_2(\tau) = & \left[8(\gamma_0^2 - \gamma_1\gamma_{-1})^2 + 3(\gamma_1^2 - \gamma_{-1}^2)^2 + 12\gamma_0^2(\gamma_1 + \gamma_{-1})^2 \right] C_0^4 \\ & + 4 \left[8\gamma_0^4 + 3(5\gamma_0^2 + \gamma_1\gamma_{-1})(\gamma_1 + \gamma_{-1})^2 - 4\gamma_1\gamma_{-1}(\gamma_0^2 + \gamma_1\gamma_{-1}) \right] C_0^2C_1^2 + \\ & 8 \left[\gamma_0^4 + 4\gamma_0^2\gamma_1\gamma_{-1} + \gamma_1^2\gamma_{-1}^2 \right] C_1^4 - 24C_0C_1\gamma_0(\gamma_1 + \gamma_{-1}) \left[(2\gamma_0^2 + \gamma_1^2 + \gamma_{-1}^2)C_0^2 \right. \\ & \left. + 2(\gamma_0^2 + \gamma_1\gamma_{-1})C_1^2 \right]. \end{aligned} \quad (40)$$

The exact computation for the trivariate embedding case have also been conducted; but because of their lengthy 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.

Remark 10. *Up to now, we have derived the mean and variance of a test variable \widehat{B}_p . In order to compute the power of the test, we need its distribution. First, \widehat{B}_p is shown in [14] to converge to B_p in probability. Next, B_p is a sum of n.i.d. random variables enjoying the mixing Property 2; for this reason B_p converges to a normal variable thanks to the Law of Large Numbers [26, ch.IV]. This guarantees that $(\widehat{B}_p - \mathbb{E}\{\widehat{B}_p\})/\sqrt{\text{VAR}\{\widehat{B}_p\}}$ is asymptotically $\mathcal{N}(0, 1)$.*

9.1. Gaussian Marginals under \bar{H}

Copulas are a classical framework, which is simple to implement 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 :

$$\mathcal{C}_R(u, v) = \int_{-\infty}^{\Phi^{-1}(u)} \int_{-\infty}^{\Phi^{-1}(v)} \frac{1}{2\pi(1 - R_{12}^2)^{1/2}} \exp\left\{-\frac{s^2 - 2R_{12}st + t^2}{2(1 - R_{12}^2)}\right\} ds dt \quad (41)$$

where Φ^{-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:

$$\text{Clayton: } \mathcal{C}_\theta(u, v) = \max\{u^{-\theta} + v^{-\theta} - 1; 0\}, \theta \in [-1, \infty) \setminus \{0\} \quad (42)$$

$$\text{Gumbel: } \mathcal{C}_\theta(u, v) = \exp\left\{-(-\log(u)^\theta + -\log(v)^\theta)^{\frac{1}{\theta}}\right\}, \theta \in [1, \infty) \quad (43)$$

Since Sklar's theorem does not impose *independence* of any variate u or v of $\mathcal{C}_\theta(u, v)$, we need to 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:

$$u = \Phi(y_1) \tag{44}$$

$$v = \Phi(y_2) \tag{45}$$

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}_θ . For more details about efficient sampling of copula see the (Marshall and Olkin 1988 algorithm) cited in [27].

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

$$\begin{aligned} x_1(n) &= \Phi^{-1}(u(n)) \\ x_2(n) &= \Phi^{-1}(v(n)) \end{aligned}$$

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 $p\text{-value} = 2(1 - \Phi(|t|))$ is compared to pre-specified significance levels α . For any p smaller than α , 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.

Test statistic	Gaussian $R_{12} = 0.8$		Clayton $\theta = 2$		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 : $\alpha = 5\%, 10\%$

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 [10] and [11] who showed that the correlation among samples is 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:

$$y(n) = x(n) + kb(n) \quad (46)$$

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 ϵ follows a double-exponential distribution with unit scale parameter. We perform 500 replications of $\{y(n)\}$ of total length $N_{tot} = n_{drop} + N$,

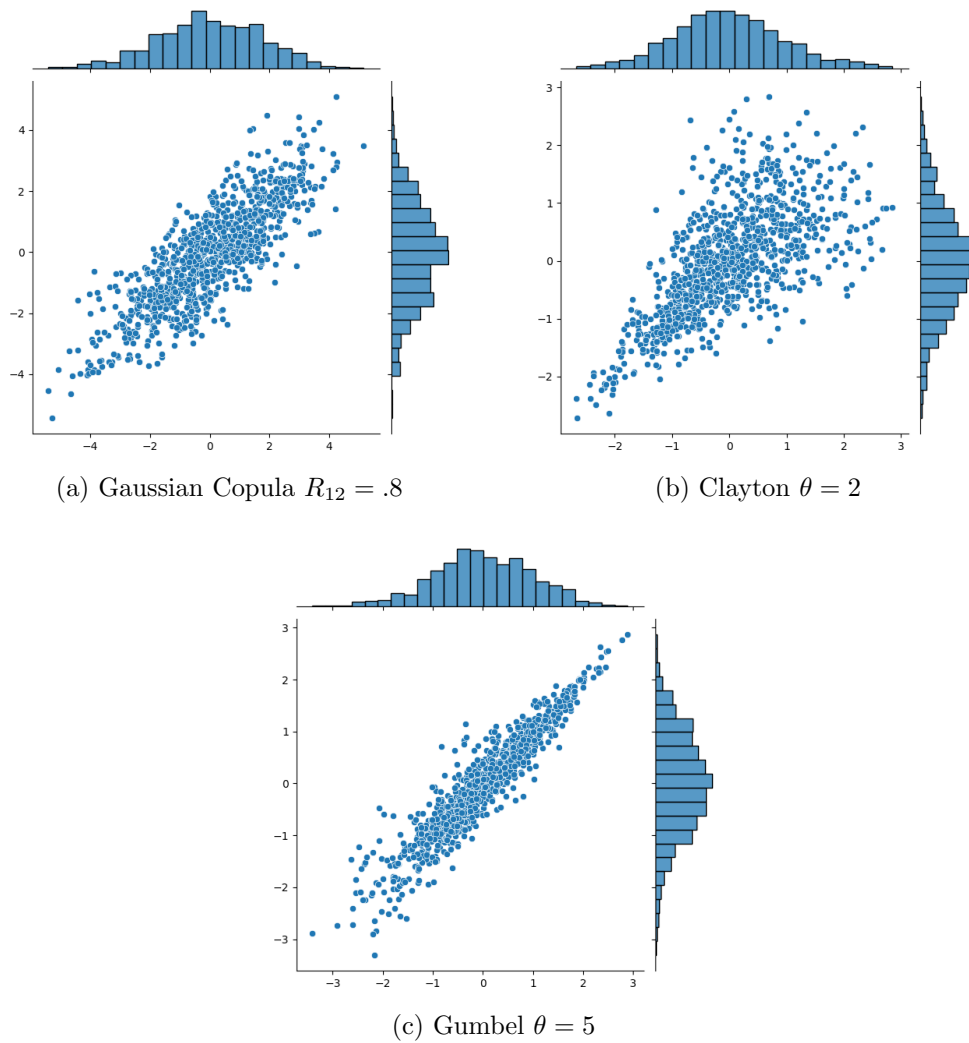


Figure 1: Examples of non-Gaussian process whose marginals are standard normal

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.

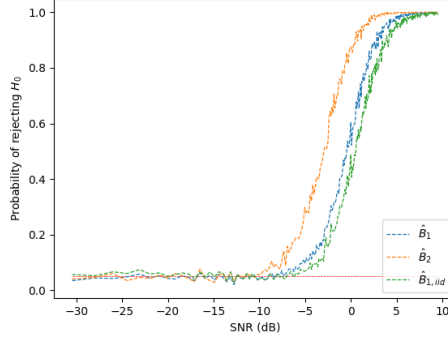


Figure 2: Empirical rejection rate at $\alpha = 5\%$ (in red dashed horizontal line) for 300 SNR values in logarithmic scale (dB)

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.

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.

Acknowledgment. This work has been partially supported by the MIAI chair “Environmental issues underground” of Institut MIAI@Grenoble Alpes (ANR-19-P3IA-0003).

References

- [1] C. L. Nikias, A. P. Petropulu, Higher-Order Spectra Analysis, Signal Processing Series, Prentice-Hall, Englewood Cliffs, 1993.
- [2] S. Haykin, Unsupervised Adaptive Filtering, Vol. 1 & 2, Wiley, 2000, series in Adaptive and Learning Systems for Communications, Signal Processing, and Control.
- [3] A. Cichocki, S.-I. Amari, Adaptive Blind Signal and Image Processing, Wiley, New York, 2002.
- [4] P. Comon, C. Jutten (Eds.), Handbook of Blind Source Separation, Independent Component Analysis and Applications, Academic Press, Oxford UK, Burlington USA, 2010.
- [5] M. Basseville, I. Nikiforov, Detection of Abrupt Changes, Theory and Application, Information and System Sciences Series, Prentice-Hall, Englewood Cliffs, 1993.
- [6] D. S. Moore, A chi-square statistic with random cell boundaries, The Annals of Statistics 42 (1) (1971) 147–156.

- [7] K. O. Bowman, L. R. Shenton, Omnibus contours for departures from normality based on b_1 and b_2 , *Biometrika* 62 (1975) 243–250.
- [8] K. V. Mardia, Applications of some measures of multivariate skewness and kurtosis for testing normality, *Sankhya B* 36 (1974) 115–128.
- [9] S. Kotz, N. L. Johnson, *Encyclopedia of Statistical Sciences*, Wiley, 1982.
- [10] D. S. Moore, The effect of dependence on chi squared tests of fit, *The Annals of Statistics* 10 (4) (1982) 1163–1171.
- [11] T. Gasser, Goodness-of-fit tests for correlated data, *Biometrika* 62 (3) (1975) 563–570.
- [12] S. S. Shapiro, M. B. Wilk, H. J. Chen, A comparative study of various tests for normality, *American Statistical Association Journal* 63 (1968) 1343–1372.
- [13] E. S. Pearson, R. B. D’agostino, K. O. Bowman, Tests for departure from normality: Comparison of powers, *Biometrika* 64 (2) (1977) 231–246.
- [14] K. V. Mardia, Measures of multivariate skewness and kurtosis with applications, *Biometrika* 57 (1970) 519–530.
- [15] D. F. Andrews, R. Gnanadesikan, J. L. Warner, Methods for assessing multivariate normality, in: P. R. Krishnaiah (Ed.), *Multivariate Analysis III*, Academic press, 1973, pp. 95–116.
- [16] R. Henze, Invariant tests for multivariate normality: a critical review, *Statistical papers* 43 (2002) 467–506.
- [17] M. Hinich, Testing for Gaussianity and linearity of a stationary time series, *Journal of Time Series Analysis* 3 (3) (1982) 169–176.
- [18] D. R. Brillinger, *Time Series, Data Analysis and Theory*, Holden-Day, 1981.
- [19] E. Moulines, K. Choukri, M. Charbit, Testing that a multivariate stationary time series is Gaussian, in: *Sixth SSAP Workshop on Stat. Signal and Array Proc.*, 1992, pp. 185–188.

- [20] T. W. Epps, Testing that a stationary time series is Gaussian, *The Annals of Statistics* 15 (4) (1987) 1683–1698.
- [21] Y. Steinberg, O. Zeitouni, On tests for normality, *IEEE Trans. on Inf. Theory* 38 (6) (1992) 1779–1787.
- [22] J. F. Malkovich, A. Afifi, On tests for multivariate normality, *Journal of the American statistical association* 68 (341) (1973) 176–179.
- [23] A. Nieto-Reyes, J. A. Cuesta-Albertos, F. Gamboa, A random-projection based test of gaussianity for stationary processes, *Computational Statistics & Data Analysis* 75 (2014) 124–141.
- [24] P. McCullagh, *Tensor Methods in Statistics*, Monographs on Statistics and Applied Probability, Chapman and Hall, 1987.
- [25] P. Comon, L. Deruaz, Normality tests for coloured samples, in: *IEEE-ATHOS Workshop on Higher-Order Statistics*, Begur, Spain, 1995, pp. 217–221.
- [26] E. J. Hannan, *Multiple time series*, Wiley, 1970.
- [27] M. Hofert, Sampling archimedean copulas, *Computational Statistics & Data Analysis* 52 (12) (2008) 5163–5174.
- [28] H. Cramér, A contribution to the theory of statistical estimation, *Scandinavian Actuarial Journal* 1946 (1) (1946) 85–94. doi:10.1080/03461238.1946.10419631.
- [29] S. ElBouch, Supplementary material, working paper or preprint (Sep. 2021).
URL <https://hal.archives-ouvertes.fr/hal-03343508>

11. Appendices

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

McCullagh’s bracket notation [24] 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} :

$$M_{abcd} = S_{ab}S_{cd} + S_{ac}S_{bd} + S_{ad}S_{bc} = [3]S_{ab}S_{cd} \quad (47)$$

Moments of higher order can be found easily:

$$\text{order 6:} \quad M_{abcdef} = [15]S_{ab}S_{cd}S_{ef} \quad (48)$$

$$\text{order 8:} \quad M_{abcdefgh} = [105]S_{ab}S_{cd}S_{ef}S_{gh} \quad (49)$$

$$\text{order 10:} \quad M_{abcdefghij} = [945]S_{ab}S_{cd}S_{ef}S_{gh}S_{ij} \quad (50)$$

$$\text{order 12:} \quad M_{abcdefghijk\ell} = [10395]S_{ab}S_{cd}S_{ef}S_{gh}S_{ij}S_{k\ell} \quad (51)$$

$$\text{order 14:} \quad M_{abcdefghijklmn} = [135135]S_{ab}S_{cd}S_{ef}S_{gh}S_{ij}S_{k\ell}S_{mn} \quad (52)$$

$$\text{order 16:} \quad M_{abcdefghijklmnpq} = [2027025]S_{ab}S_{cd}S_{ef}S_{gh}S_{ij}S_{k\ell}S_{mn}S_{pq} \quad (53)$$

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 5, $A_{\alpha^l \beta^l} = \mathbf{x}(\alpha_l)^T G \mathbf{x}(\beta_l)$, 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 (9), it appears that the indices (α_l, β_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

$$\mu_{r_1 \dots r_L c_1 \dots c_L}^{\alpha_1 \dots \alpha_L \beta_1 \dots \beta_L} = M_{i \eta_i j \eta_j k \eta_k \dots} \quad (54)$$

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 (10) 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 (11) it will be expanded as a sum of moments of order 8 (i.e. $L = 4$); using the

compact notation from equation (54), we get

$$\begin{aligned}\mathbb{E}\{A_{nn}A_{nj}A_{jk}A_{kn}\} &= \sum_{((r_i, c_i)_{i=1 \dots 4})=1}^p G_{r_1 c_1} G_{r_2 c_2} G_{r_3 c_3} G_{r_4 c_4} \mu_{r_1 c_1 r_2 c_2 r_3 c_3 r_4 c_4}^{nnnj j k k n} \\ &= \sum_{((r_i, c_i)_{i=1 \dots 4})=1}^p G_{r_1 c_1} G_{r_2 c_2} G_{r_3 c_3} G_{r_4 c_4} M_{n^4 j^2 k^2} \quad (55)\end{aligned}$$

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 $(n, r_1), (n, c_1), (n, r_2), (n, c_2), (j, r_3), (j, c_3), (k, r_4), (k, c_4)$ respectively, as it appears in equation (55). Plugging the above expansion in equation (55) leads to summing over 64×105 terms! However, in most cases of interest many terms may be grouped together and highlight the behavior of equation (11). 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 (55) 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 4×2 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 3×1 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 $p = 1$

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

which can be directly plugged into equation (55). 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

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

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 (54) 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 (56), 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.

$$M_{iiij} = [3]\mu_{ab}^{ii}\mu_{cd}^{ij} \quad (56)$$

$$M_{iijj} = [2]\mu_{ab}^{ij}\mu_{cd}^{ij} + \mu_{ab}^{ii}\mu_{cd}^{jj} \quad (57)$$

Order 4, D=3.

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

Order 6, D=2.

$$M_{i^5j} = [15]\mu_{ab}^{ii}\mu_{cd}^{ii}\mu_{ef}^{ij} \quad (59)$$

$$M_{i^4jj} = [12]\mu_{ae}^{ij}\mu_{bf}^{ij}\mu_{cd}^{ii} + [3]\mu_{ab}^{ii}\mu_{cd}^{ii}\mu_{ef}^{jj} \quad (60)$$

$$M_{iii jjj} = [6]\mu_{ad}^{ij}\mu_{be}^{ij}\mu_{df}^{ij} + [9]\mu_{ab}^{ii}\mu_{cd}^{ij}\mu_{ef}^{jj} \quad (61)$$

Order 6, D=3.

$$M_{i^4jk} = [3]\mu_{ab}^{ii}\mu_{cd}^{ii}\mu_{ef}^{jk} + [12]\mu_{ae}^{ij}\mu_{bf}^{ik}\mu_{cd}^{ii} \quad (62)$$

$$M_{iii jk} = [6]\mu_{ad}^{ij}\mu_{be}^{ij}\mu_{cf}^{ik} + [6]\mu_{ad}^{ij}\mu_{bc}^{ii}\mu_{ef}^{jk} + [3]\mu_{ab}^{ii}\mu_{de}^{jj}\mu_{cf}^{ik} \quad (63)$$

$$M_{iij jkk} = \mu_{ab}^{ii}\mu_{cd}^{jj}\mu_{ef}^{kk} + [2]\mu_{ab}^{ii}\mu_{ce}^{jk}\mu_{bf}^{jk} + [2]\mu_{cd}^{jj}\mu_{ae}^{ik}\mu_{bf}^{ik} + [2]\mu_{ef}^{kk}\mu_{ac}^{ij}\mu_{bd}^{ij} \\ + [8]\mu_{ac}^{ij}\mu_{de}^{jk}\mu_{bf}^{ik} \quad (64)$$

Order 8, D=2.

$$M_{i^7j} = [105]\mu_{ab}^{ii}\mu_{cd}^{ii}\mu_{ef}^{ii}\mu_{gh}^{ij} \quad (65)$$

$$M_{i^6jj} = [90]\mu_{ag}^{ij}\mu_{bh}^{ij}\mu_{cd}^{ii}\mu_{ef}^{ii} + [15]\mu_{ab}^{ii}\mu_{cd}^{ii}\mu_{ef}^{ii}\mu_{gh}^{jj} \quad (66)$$

$$M_{i^5 jjj} = [60]\mu_{af}^{ij}\mu_{bg}^{ij}\mu_{ch}^{ij}\mu_{de}^{ii} + [45]\mu_{ab}^{ii}\mu_{cd}^{ii}\mu_{ef}^{ij}\mu_{gh}^{jj} \quad (67)$$

$$M_{i^4 j^4} = [9]\mu_{ab}^{ii}\mu_{cd}^{ii}\mu_{ef}^{jj}\mu_{gh}^{jj} + [72]\mu_{ab}^{ii}\mu_{ce}^{ij}\mu_{df}^{ij}\mu_{gh}^{jj} + [24]\mu_{ae}^{ij}\mu_{bf}^{ij}\mu_{cg}^{ij}\mu_{dh}^{ij} \quad (68)$$

Order 8, D=3.

$$M_{i^6jk} = [15]\mu_{ab}^{ii}\mu_{cd}^{ii}\mu_{ef}^{ii}\mu_{gh}^{jk} + [90]\mu_{ab}^{ii}\mu_{cd}^{ii}\mu_{eg}^{ij}\mu_{fg}^{ik} \quad (69)$$

$$M_{i^5jjk} = [30]\mu_{ab}^{ii}\mu_{cd}^{ii}\mu_{ef}^{ij}\mu_{gh}^{jk} + [60]\mu_{af}^{ij}\mu_{bc}^{ii}\mu_{dh}^{ik} + [15]\mu_{ab}^{ii}\mu_{cd}^{ii}\mu_{fg}^{jj}\mu_{eh}^{ik} \quad (70)$$

$$M_{i^4jjjk} = [9]\mu_{ab}^{ii}\mu_{cd}^{ii}\mu_{ef}^{jj}\mu_{gh}^{jk} + [36]\mu_{ef}^{jj}\mu_{ag}^{ij}\mu_{bh}^{ik}\mu_{cd}^{ii} \\ + [24]\mu_{ae}^{ij}\mu_{bf}^{ij}\mu_{cg}^{ij}\mu_{dh}^{ik} + [36]\mu_{ab}^{ii}\mu_{ce}^{ij}\mu_{df}^{ij}\mu_{gh}^{jk} \quad (71)$$

$$M_{i^4jjkk} = [3]\mu_{ab}^{ii}\mu_{cd}^{ii}\mu_{ef}^{jj}\mu_{gh}^{kk} + [6]\mu_{ab}^{ii}\mu_{cd}^{ii}\mu_{eg}^{jk}\mu_{fh}^{jk} + [12]\mu_{ab}^{ii}\mu_{ce}^{ij}\mu_{df}^{ij}\mu_{gh}^{kk} \\ + [24]\mu_{ag}^{ik}\mu_{bh}^{ik}\mu_{ce}^{ij}\mu_{df}^{ij} + [48]\mu_{ae}^{ij}\mu_{fg}^{jk}\mu_{bh}^{ik}\mu_{cd}^{ii} \\ + [12]\mu_{ab}^{ii}\mu_{bg}^{ik}\mu_{ch}^{ik}\mu_{ef}^{jj} \quad (72)$$

$$M_{iii jjkk} = [9]\mu_{ab}^{ii}\mu_{cd}^{ij}\mu_{ef}^{jj}\mu_{gh}^{kk} + [18]\mu_{ab}^{ii}\mu_{cd}^{ij}\mu_{eg}^{jk}\mu_{fh}^{jk} + [6]\mu_{ad}^{ij}\mu_{be}^{ij}\mu_{cf}^{ij}\mu_{gh}^{kk} \\ + [18]\mu_{ag}^{ik}\mu_{bh}^{ik}\mu_{cd}^{ij}\mu_{ef}^{jj} + [36]\mu_{ad}^{ij}\mu_{be}^{ij}\mu_{cg}^{ik}\mu_{fh}^{jk} \\ + [18]\mu_{ag}^{ik}\mu_{dh}^{jk}\mu_{bc}^{ii}\mu_{ef}^{jj} \quad (73)$$

Order 10, D=2.

$$M_{i^9j} = [945]\mu_{ab}^{ii}\mu_{cd}^{ii}\mu_{ef}^{ii}\mu_{gh}^{ii}\mu_{ml}^{ij} \quad (74)$$

$$M_{i^8jj} = [105]\mu_{ab}^{ii}\mu_{cd}^{ii}\mu_{ef}^{ii}\mu_{gh}^{ii}\mu_{ml}^{jj} + [840]\mu_{ab}^{ii}\mu_{cd}^{ii}\mu_{ef}^{ii}\mu_{gm}^{ij}\mu_{hl}^{ij} \quad (75)$$

$$M_{i^7jjj} = [315]\mu_{ab}^{ii}\mu_{cd}^{ii}\mu_{ef}^{ii}\mu_{gh}^{ij}\mu_{ml}^{jj} + [630]\mu_{ah}^{ij}\mu_{bm}^{ij}\mu_{cl}^{ij}\mu_{de}^{ii}\mu_{fg}^{ii} \quad (76)$$

$$M_{i^6j^4} = [45]\mu_{ab}^{ii}\mu_{cd}^{ii}\mu_{ef}^{ii}\mu_{gh}^{jj}\mu_{ml}^{jj} + [360]\mu_{ag}^{ij}\mu_{bh}^{ij}\mu_{cm}^{ij}\mu_{dl}^{ij}\mu_{ef}^{ii} \\ + [540]\mu_{ab}^{ii}\mu_{cd}^{ii}\mu_{eg}^{ij}\mu_{fh}^{ij}\mu_{ml}^{jj} \quad (77)$$

$$M_{i^5j^5} = [120]\mu_{af}^{ij}\mu_{bg}^{ij}\mu_{ch}^{ij}\mu_{dm}^{ij}\mu_{el}^{ij} \\ + [225]\mu_{ab}^{ii}\mu_{cd}^{ii}\mu_{ef}^{ij}\mu_{gh}^{jj}\mu_{ml}^{jj} + [600]\mu_{fg}^{jj}\mu_{ah}^{ij}\mu_{bm}^{ij}\mu_{cl}^{ij}\mu_{de}^{ii} \quad (78)$$

Order 10, D=3.

$$M_{i^8jk} = [105]\mu_{ab}^{ii}\mu_{cd}^{ii}\mu_{ef}^{ii}\mu_{gh}^{ii}\mu_{ml}^{jk} + [840]\mu_{ab}^{ii}\mu_{cd}^{ii}\mu_{ef}^{ii}\mu_{gm}^{ij}\mu_{hl}^{ik} \quad (79)$$

$$M_{i^7jjk} = [210]\mu_{ab}^{ii}\mu_{cd}^{ii}\mu_{ef}^{ii}\mu_{gh}^{ij}\mu_{ml}^{jk} + [630]\mu_{ab}^{ii}\mu_{cd}^{ii}\mu_{eh}^{ij}\mu_{fm}^{ij}\mu_{hl}^{ik} \\ + [105]\mu_{ab}^{ii}\mu_{cd}^{ii}\mu_{ef}^{ii}\mu_{hm}^{jj}\mu_{gl}^{ik} \quad (80)$$

$$M_{i^6jjjk} = [45]\mu_{ab}^{ii}\mu_{cd}^{ii}\mu_{ef}^{ii}\mu_{gh}^{jj}\mu_{ml}^{jk} + [270]\mu_{ab}^{ii}\mu_{cd}^{ii}\mu_{eg}^{ij}\mu_{fh}^{ij}\mu_{ml}^{jk} \\ + [360]\mu_{ag}^{ij}\mu_{bh}^{ij}\mu_{cm}^{ij}\mu_{dl}^{ik}\mu_{ef}^{ii} + [270]\mu_{al}^{ik}\mu_{gh}^{jj}\mu_{bl}^{ij}\mu_{cd}^{ii}\mu_{ef}^{ii} \quad (81)$$

$$M_{i^6jjkk} = [15]\mu_{ab}^{ii}\mu_{cd}^{ii}\mu_{ef}^{ii}\mu_{gh}^{jj}\mu_{ml}^{kk} + [30]\mu_{ab}^{ii}\mu_{cd}^{ii}\mu_{ef}^{ii}\mu_{gm}^{jk}\mu_{hl}^{jk} \\ + [90]\mu_{ab}^{ii}\mu_{cd}^{ii}\mu_{eg}^{ij}\mu_{fh}^{ij}\mu_{ml}^{kk} + [90]\mu_{ab}^{ii}\mu_{cd}^{ii}\mu_{em}^{ik}\mu_{fl}^{ik}\mu_{gh}^{jj} \\ + [360]\mu_{ab}^{ii}\mu_{cg}^{ij}\mu_{dh}^{ij}\mu_{em}^{ik}\mu_{fl}^{ik} + [360]\mu_{ab}^{ii}\mu_{cd}^{ii}\mu_{eg}^{ij}\mu_{fm}^{ik}\mu_{hl}^{jk} \quad (82)$$

$$M_{i^5j^4k} = [45]\mu_{ab}^{ii}\mu_{cd}^{ii}\mu_{fg}^{jj}\mu_{hm}^{jj}\mu_{el}^{ik} + [360]\mu_{ab}^{ii}\mu_{cf}^{ij}\mu_{dg}^{jj}\mu_{hm}^{ij}\mu_{el}^{ik} \\ + [120]\mu_{ag}^{ij}\mu_{bf}^{ij}\mu_{cg}^{ij}\mu_{dh}^{ij}\mu_{el}^{ik} + [180]\mu_{ab}^{ii}\mu_{cd}^{ii}\mu_{ef}^{ij}\mu_{gh}^{jj}\mu_{ml}^{jk} \\ + [240]\mu_{ab}^{ii}\mu_{cf}^{ij}\mu_{dg}^{ij}\mu_{eh}^{ij}\mu_{ml}^{jk} \quad (83)$$

$$M_{i^5jjjkk} = [45]\mu_{ab}^{ii}\mu_{cd}^{ii}\mu_{ef}^{ij}\mu_{gh}^{jj}\mu_{ml}^{kk} + [60]\mu_{ab}^{ii}\mu_{cf}^{ij}\mu_{dg}^{ij}\mu_{eh}^{ij}\mu_{ml}^{kk} \\ + [90]\mu_{ab}^{ii}\mu_{cd}^{ii}\mu_{ef}^{ij}\mu_{gm}^{jk}\mu_{hl}^{jk} + 360\mu_{ab}^{ii}\mu_{cf}^{ij}\mu_{dg}^{ij}\mu_{em}^{ik}\mu_{hl}^{jk} \\ + [90]\mu_{ab}^{ii}\mu_{fg}^{jj}\mu_{cm}^{ik}\mu_{hl}^{jk} + [180]\mu_{ab}^{ii}\mu_{cf}^{ij}\mu_{gh}^{jj}\mu_{dm}^{ik}\mu_{el}^{ik} \\ + [120]\mu_{af}^{ij}\mu_{bg}^{ij}\mu_{ch}^{ij}\mu_{dm}^{ik}\mu_{el}^{ik} \quad (84)$$

$$M_{i^4jjjkkk} = [27]\mu_{ab}^{ii}\mu_{cd}^{ii}\mu_{ef}^{jj}\mu_{gh}^{jk}\mu_{ml}^{kk} + [18]\mu_{ab}^{ii}\mu_{cd}^{ii}\mu_{eh}^{jk}\mu_{fm}^{jk}\mu_{gl}^{ij} \\ + [108]\mu_{ab}^{ii}\mu_{ce}^{ij}\mu_{df}^{ij}\mu_{gh}^{jk}\mu_{ml}^{kk} + [108]\mu_{ab}^{ii}\mu_{ef}^{jj}\mu_{gh}^{jk}\mu_{cm}^{ik}\mu_{dl}^{ik} \\ + [108]\mu_{ab}^{ii}\mu_{ef}^{jj}\mu_{cg}^{ij}\mu_{dh}^{ik}\mu_{ml}^{kk} + [216]\mu_{ab}^{ii}\mu_{ce}^{ij}\mu_{dh}^{ik}\mu_{fm}^{jk}\mu_{gl}^{jk} \\ + [72]\mu_{ae}^{ij}\mu_{bf}^{ij}\mu_{cg}^{ij}\mu_{dh}^{ik}\mu_{ml}^{kk} + [216]\mu_{ah}^{ik}\mu_{bm}^{ik}\mu_{ce}^{ij}\mu_{df}^{ij}\mu_{gl}^{jk} \\ + [72]\mu_{ah}^{ik}\mu_{bm}^{ik}\mu_{cl}^{ik}\mu_{de}^{ij}\mu_{fg}^{jj} \quad (85)$$

$$M_{i^4j^4kk} = [9]\mu_{ab}^{ii}\mu_{cd}^{ii}\mu_{ef}^{jj}\mu_{gh}^{jj}\mu_{ml}^{kk} + [72]\mu_{ab}^{ii}\mu_{ce}^{ij}\mu_{df}^{ij}\mu_{gh}^{jj}\mu_{ml}^{kk} \\ + [24]\mu_{ae}^{ij}\mu_{bf}^{ij}\mu_{cg}^{ij}\mu_{dh}^{ij}\mu_{ml}^{kk} + [36]\mu_{ab}^{ii}\mu_{ef}^{jj}\mu_{gm}^{jk}\mu_{hl}^{jk} \\ + [144]\mu_{ab}^{ii}\mu_{ce}^{ij}\mu_{df}^{ij}\mu_{gm}^{jk}\mu_{hl}^{jk} + [36]\mu_{ab}^{ii}\mu_{ef}^{jj}\mu_{gh}^{jj}\mu_{cm}^{ik}\mu_{dl}^{ik} \\ + [144]\mu_{ae}^{ij}\mu_{bf}^{ij}\mu_{gh}^{jj}\mu_{cm}^{ik}\mu_{dl}^{ik} + [288]\mu_{ab}^{ii}\mu_{ef}^{jj}\mu_{cg}^{ij}\mu_{dm}^{ik}\mu_{hl}^{jk} \\ + [192]\mu_{ae}^{ij}\mu_{bf}^{ij}\mu_{cg}^{ij}\mu_{dm}^{ik}\mu_{hl}^{jk} \quad (86)$$

11.4. Particular results when $p = 1$

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

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

$$M_{i^{11}j} = 10395S_{ii}^5S_{ij} + 9450S_{ii}^4S_{ij}^2 \quad (87)$$

$$M_{i^9jjj} = 2835S_{ii}^4S_{ij}S_{jj} + 7560S_{ii}^3S_{ij}^3 \quad (88)$$

$$M_{i^8j4} = 5040S_{ij}^4S_{ii}^2 + 315S_{ii}^4S_{jj}^2 + 5040S_{ii}^3S_{ij}^2S_{jj} \quad (89)$$

$$M_{i^7j^5} = 1575S_{ii}^3S_{ij}S_{jj}^2 + 6300S_{ii}^2S_{ij}^3S_{jj} + 2520S_{ii}^5S_{ij} \quad (90)$$

$$M_{i^6j^6} = 720S_{ij}^6 + 225S_{ii}^3S_{jj}^3 + 5400S_{ii}S_{ij}^4S_{jj} + 4050S_{ii}^2S_{ij}^2S_{jj}^2 \quad (91)$$

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

$$M_{i^{10}jk} = 945S_{ii}^5S_{jk} + 9450S_{ik}S_{ij}S_{ii}^4 \quad (92)$$

$$M_{ijjk} = 945S_{ii}^4S_{jj}S_{ik} + 7560S_{ii}^3S_{ij}^2S_{ik} + 1890S_{ii}^4S_{ij}S_{jk} \quad (93)$$

$$M_{i^8jjjk} = 315S_{ii}^4S_{jj}S_{jk} + 2520S_{ii}^3S_{ij}S_{jj}S_{ik} + 2520S_{ii}^3S_{ij}^2S_{jk} \\ + 5040S_{ii}^2S_{ij}^3S_{ik} \quad (94)$$

$$M_{i^7j^4k} = 315S_{ii}^3S_{jj}^2S_{ik} + 3780S_{ii}^2S_{ij}^2S_{ik} + 1260S_{ii}^4S_{ij}S_{ik} + 1260S_{ii}^3S_{ij}S_{jj}S_{jk} \\ + 3780S_{ii}^2S_{ij}^3S_{jk} \quad (95)$$

$$M_{i^8jjkk} = 105S_{ii}^4S_{jj}S_{kk} + 210S_{ii}^4S_{jk}^2 + 840S_{ii}^3S_{ij}^2S_{kk} + 840S_{ii}^3S_{ik}^2S_{jj} \\ + 5040S_{ii}^2S_{ij}^2S_{ik}^2 + 3360S_{ii}^3S_{ik}S_{ij}S_{jk} \quad (96)$$

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

$$M_{i^4j^4kk\ell\ell} = 3S_{ii}^2[3S_{jj}^2S_{kk}S_{\ell\ell} + 6S_{jj}^2S_{k\ell}^2 + 12S_{jj}S_{jk}^2S_{\ell\ell} + 24S_{j\ell}^2S_{jk}^2 \\ + 48S_{jk}S_{k\ell}S_{j\ell}S_{jj} + 12S_{jj}S_{j\ell}^2S_{kk}] + 3S_{jj}^2[12S_{ii}S_{ik}^2S_{\ell\ell} \\ + 24S_{il}^2S_{ik}^2 + 48S_{ik}S_{k\ell}S_{il}S_{ii} + 12S_{ii}S_{il}^2S_{kk}] \\ + 24S_{ij}^4S_{kk}S_{\ell\ell} + 48S_{ij}^4S_{k\ell}^2 + 96S_{ij}^3[2S_{ik}S_{jk}S_{\ell\ell} + 2S_{il}S_{j\ell}S_{kk} \\ + 4S_{ik}S_{j\ell}S_{k\ell} + 4S_{il}S_{jk}S_{\ell k}] + 72S_{ij}^2[4S_{ik}^2S_{j\ell}^2 + 4S_{jk}^2S_{il}^2 \\ + 16S_{ik}S_{il}S_{jk}S_{j\ell} + S_{ii}S_{jj}S_{kk}S_{\ell\ell} + 2S_{ii}S_{jj}S_{k\ell}^2] + 12S_{ik}^2[12S_{ji}^2 \\ \times S_{jj}S_{\ell\ell} + 48S_{ij}S_{il}S_{j\ell}S_{jj} + 12S_{jj}S_{j\ell}^2S_{ii} + 12S_{il}^2[12S_{ji}^2S_{jj}S_{kk} \\ + 48S_{ij}S_{ik}S_{jk}S_{jj} + 12S_{jj}S_{jk}^2S_{ii}] + 12S_{j\ell}^2[12S_{ij}^2S_{ii}S_{kk} \\ + 48S_{ij}S_{jk}S_{ik}S_{ii}] + 12S_{jk}^2[12S_{ij}^2S_{ii}S_{\ell\ell} + 48S_{ij}S_{j\ell}S_{il}S_{ii}] \\ + 576S_{ii}[S_{ik}S_{il}S_{jj}S_{jk}S_{j\ell} + S_{ik}S_{ij}S_{jj}S_{j\ell}S_{k\ell} \\ + S_{ik}S_{ij}S_{jj}S_{jk}S_{\ell\ell} + S_{il}S_{ij}S_{jj}S_{jk}S_{\ell k} \\ + S_{il}S_{ij}S_{jj}S_{j\ell}S_{kk} + S_{ik}S_{ij}S_{jj}S_{\ell k}S_{j\ell}] \quad (97)$$

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:

$$\mathbb{E}\{A_{nn}^2\} = \sum_{a,b=1}^p \sum_{c,d=1}^p G_{ab}G_{cd}(S_{ab}S_{cd} + S_{ac}S_{bd} + S_{ad}S_{bc}) \quad (98)$$

For p=1.

$$\mathbb{E}\{A_{nn}^2\} = 3 \quad (99)$$

$$\mathbb{E}\{A_{nn}A_{ni}^2\} = 3 + 12\frac{S(n-i)^2}{S^2} \quad (100)$$

$$\begin{aligned} \mathbb{E}\{A_{ni}^2A_{nj}^2\} &= 3 + 6\frac{S(i-j)^2}{S^2} + 12\frac{S(n-i)^2}{S^2} + 12\frac{S(n-j)^2}{S^2} \\ &\quad + 24\frac{S(n-i)^2S_{11}(n-j)^2}{S^4} + 48\frac{S(n-i)S(n-j)S(i-j)}{S^3} \\ \mathbb{E}\{A_{nn}A_{nj}A_{jk}A_{kn}\} &= 3 + 6\frac{S(j-k)^2}{S^2} + 12\frac{S(n-k)^2}{S^2} + 12\frac{S(n-j)^2}{S^2} \\ &\quad + 24\frac{S(n-k)^2S(n-j)^2}{S^4} + 48\frac{S(n-k)S(j-k)S(n-j)}{S^3} \end{aligned} \quad (101)$$

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

$$\begin{aligned} \mathbb{E}\{\hat{B}_1\} &= 3 - \frac{6}{N^2} \sum_{n,i} \frac{S(n-i)^2}{S^2} + \frac{72}{N^3} \sum_{n,i,j} \frac{S(n-j)^2S(n-i)^2}{S^4} \\ &\quad + \frac{144}{N^3} \sum_{n,i,j} \frac{S(n-i)S(i-j)S(n-j)}{S^3} \end{aligned} \quad (102)$$

Based on the results in [28, 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.

$$\mathbb{E}\{A_{nn}^2\} = 8 \quad (103)$$

$$\begin{aligned} \mathbb{E}\{A_{nn}A_{ni}^2\} &= 8 + \frac{1}{(S_{11}S_{22} - S_{12}^2)^2} \left[S_{11}S_{22}[2S_{12}^2(n-i) + 2S_{12}^2(n-i)] \right. \\ &\quad + 12S_{21}(n-i)S_{12}(n-i) + 12S_{22}(n-i)S_{11}(n-i) \\ &\quad + S_{12}^2[12S_{12}^2(n-i) + 12S_{21}(n-i)S_{12}(n-i) \\ &\quad + 12S_{21}^2(n-i) + 16S_{11}(n-i)S_{22}(n-i)] \\ &\quad - 28S_{12}S_{11}[S_{22}(n-i)S_{12}(n-i) + S_{22}(n-i)S_{12}(n-i)] \\ &\quad - 28S_{12}S_{22}[S_{11}(n-i)S_{12}(n-i) + S_{11}(n-i)S_{12}(n-i)] \\ &\quad \left. + 14S_{11}^2S_{22}^2(n-i) + 14S_{22}^2S_{11}^2(n-i) \right] \quad (104) \end{aligned}$$

Bivariate embedding.

$$\begin{aligned} \mathbb{E}\{A_{nn}A_{ni}^2\} &= 8 + \frac{1}{(C_0^2 - C_1^2)^2} \times [C_0^2 \times [2(\gamma_1(n-i) + \gamma_{-1}(n-i))^2 \\ &\quad + 8\gamma_{-1}(n-i)\gamma_1(n-i) + 40\gamma_0(n-i)^2] + C_1^2 \\ &\quad \times [12(\gamma_1(n-i) + \gamma_{-1}(n-i))^2 - 8\gamma_{-1}(n-i)\gamma_1(n-i) + 16\gamma_0(n-i)^2] \\ &\quad - C_0C_1 \times [56\gamma_0(n-i)(\gamma_1(n-i) + \gamma_{-1}(n-i))]] \quad (105) \end{aligned}$$

$$\begin{aligned} \mathbb{E}\{A_{ni}^2A_{nj}^2\} &= 8 + \frac{1}{(C_0^2 - C_1^2)^2} \times [C_0^2 \times [2(\gamma_1(i-j) + \gamma_{-1}(i-j))^2 \\ &\quad + 2(\gamma_1(n-i) + \gamma_{-1}(n-i))^2 + 2(\gamma_1(n-j) + \gamma_{-1}(n-j))^2 \\ &\quad + 8\gamma_{-1}(n-i)\gamma_1(n-i) + 8\gamma_{-1}(n-j)\gamma_1(n-j) + 16\gamma_0(i-j)^2 \\ &\quad + 40\gamma_0(n-j)^2 + 40\gamma_0(n-i)^2] + C_1^2 \times [4(\gamma_1(i-j) + \gamma_{-1}(i-j))^2 \\ &\quad + 12(\gamma_1(n-j) + \gamma_{-1}(n-j))^2 + 12(\gamma_1(n-i) + \gamma_{-1}(n-i))^2 \\ &\quad - 8\gamma_{-1}(n-i)\gamma_1(n-i) - 8\gamma_{-1}(n-j)\gamma_1(n-j) + 8\gamma_0(i-j)^2 \\ &\quad + 16\gamma_0(n-j)^2 + 16\gamma_0(n-i)^2] - C_0C_1 \times [24\gamma_0(i-j) \\ &\quad \times (\gamma_1(i-j) + \gamma_{-1}(i-j)) + 56\gamma_0(n-i)(\gamma_1(n-i) + \gamma_{-1}(n-i)) \\ &\quad + 56\gamma_0(n-j)(\gamma_1(n-j) + \gamma_{-1}(n-j))] \quad (106) \end{aligned}$$

Following the same pattern as the mean, but with more moments involved, the computation of the variance may be conducted [[29]].

11.6. Sklar's theorem

Theorem 11. (Sklar's theorem 1959)

$$F_{X_1, X_2}(x_1, x_2) = Pr(X_1 \leq x_1, X_2 \leq x_2) = \mathcal{C}(F(x_1), G(x_2)) \quad (107)$$

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:

$$\mathcal{C}(u_1, u_2) = F_{X_1, X_2}(F^{-1}(u_1), G^{-1}(u_2)). \quad (108)$$