Received: 21 November 2019 | Revised: 8 February 2021 | Accepted: 20 February 2021

DOI: 10.1111/sjos.12521

ORIGINAL ARTICLE

Scandinavian Journal of Statistics

Asymptotic theory for statistics based on cumulant vectors with applications

Sreenivasa Rao Jammalamadaka[1] | **Emanuele Taufer**[2] | **György H. Terdik**[3]

[1]Department of Statistics and Applied Probability, University of California, Santa Barbara, California

[2]Department of Economics and Management, University of Trento, Trento, Italy

[3]Department of Information Technology, University of Debrecen, Debrecen, Hungary

Correspondence
Emanuele Taufer, Department of Economics and Management, University of Trento, Via Inama 5, 38122 Trento, Italy.
Email: emanuele.taufer@unitn.it

Funding information
European Union and European Social Fund, Grant/Award Number: EFOP3.6.2-16-2017-00015

Abstract

For any given multivariate distribution, explicit formulae for the asymptotic covariances of cumulant vectors of the third and the fourth order are provided here. General expressions for cumulants of elliptically symmetric multivariate distributions are also provided. Utilizing these formulae one can extend several results currently available in the literature, as well as obtain practically useful expressions in terms of population cumulants, and computational formulae in terms of commutator matrices. Results are provided for both symmetric and asymmetric distributions, when the required moments exist. New measures of skewness and kurtosis based on distinct elements are discussed, and other applications to independent component analysis and testing are considered.

KEYWORDS

elliptically symmetric distributions, Gram–Charlier expansion, multivariate cumulants, multivariate skewness and kurtosis

MOS SUBJECT CLASSIFICATION

MSC (2020) 62H10, 60H12, 60H15

1 | INTRODUCTION AND MOTIVATION

Cumulant-based skewness and kurtosis measures for random vectors play a central role in multivariate statistics going back to the early work of Mardia (1970). Beyond more traditional applications to estimation and testing, these measures play an important role in such areas as signal detection, clustering, invariant coordinate selection, as well as in pricing and portfolio analysis. Some related literature, without any presumption of completeness, includes Malkovich and Afifi (1973), Srivastava (1984), Koziol (1989), Móri et al. (1994), Oja et al. (2006), Balakrishnan et al. (2007), Kollo (2008), Tyler et al. (2009), Ilmonen et al. (2010), Peña et al. (2010), Tanaka et al. (2010), Huang et al. (2014), Lin et al. (2015), León and Moreno (2017), Nordhausen et al. (2017), Jammalamadaka et al. (2020).

Several asymptotic results for multivariate skewness and kurtosis statistics are available; see for instance Koziol (1987), Baringhaus (1991), Baringhaus and Henze (1991, 1992, 1994a, 1994b, 1997a, 1997b), Klar (2002), Henze (2002), Ilmonen et al. (2010), Nordhausen et al. (2017).

In this paper we develop results on the asymptotic theory of vector cumulants of the third and the fourth order in a completely general setting. Our main results concern explicit expressions for the asymptotic covariance matrices based on population parameters, and their computational formulae. Typically, in the literature, asymptotic covariances are expressed in terms of expectations of sample statistics while here the expectation step is solved in total generality, thus providing an explicit connection to the cumulants of the underlying model.

As an example, consider theorem 2.1 of Klar (2002) which provides a very nice unified treatment of the asymptotic properties of vectors of third-order cumulants: the asymptotic covariance matrix is expressed there in implicit terms as an expectation of a quadratic function of sample cumulants and their derivatives. This fact makes it quite difficult to implement the results in practice beyond the Gaussian case.

The results developed here are useful in estimation and testing, for efficiency and power comparisons, as well as in deriving simple estimators of the asymptotic covariances. For example, for the important subclass of elliptically symmetric distributions it turns out that estimation of the asymptotic covariance matrix can be carried out by estimating only a few univariate parameters, for which we provide a complete solution here. For asymmetric cases like multivariate skew normal and skew t-distributions, the reader may consult Jammalamadaka et al. (2020).

To demonstrate the usefulness of the methodology discussed, we consider applications to new weighted measures of skewness and kurtosis and tests-based thereon, for which the asymptotic distributions under the null and under alternatives are given. Connections to existing indexes of multivariate skewness and kurtosis as well as to independent component analysis (ICA) is also discussed.

The Gram–Charlier (GC) expansion of a density function is the fundamental tool that will lead us to the main results. The GC expansion is coupled with a vectorial approach to cumulants; as we shall see, this fact considerably simplifies the derivation of asymptotic results for the higher order cumulants of multivariate distributions.

To formally introduce the problem, consider a random d-vector $\underline{X}$ with mean vector $\underline{\mu}$ and covariance matrix $\mathbf{\Sigma}$; if $\phi_{\underline{X}}(\underline{\lambda})$ and $\psi_{\underline{X}}(\underline{\lambda}) = \log \phi_{\underline{X}}(\underline{\lambda})$ are, respectively, the characteristic function and the cumulant function of $\underline{X}$, then the kth order cumulant of $\underline{X}$ is given by

$$\kappa_{\underline{X},k} = \underline{\mathrm{Cum}}_k(\underline{X}) = (-\mathrm{i})^k D_{\underline{\lambda}}^{\otimes k} \boldsymbol{\psi}_{\underline{X}}(\underline{\lambda})\Big|_{\underline{\lambda}=0}. \tag{1}$$

Note that $\underline{\text{Cum}}_k(\underline{X})$ is a vector of dimension d^k that contains all possible cumulants of order k formed by $X_1, \ldots, X_d$. For instance, in Equation (1), one can see $\kappa_{\underline{X},2} = \text{Vec}\, \mathbf{\Sigma}$. The operator $D^{\otimes}_{\underline{\lambda}}$ in (1), which we refer to as the T-derivative, for any function $\boldsymbol{\phi}(\underline{\lambda})$, is defined as

$$D^{\otimes}_{\underline{\lambda}} \boldsymbol{\phi}(\underline{\lambda}) = \text{Vec}\left(\frac{\partial \boldsymbol{\phi}(\underline{\lambda})}{\partial \underline{\lambda}^{\top}} \right)^{\top} = \boldsymbol{\phi}(\underline{\lambda}) \otimes \frac{\partial}{\partial \underline{\lambda}}.$$

Here and in what follows, the symbol $\otimes$ denotes the Kronecker product. Assuming $\boldsymbol{\phi}$ is k times differentiable, the kth T-derivative is given by

$$D^{\otimes k}_{\underline{\lambda}} \boldsymbol{\phi}(\underline{\lambda}) = D^{\otimes}_{\underline{\lambda}} \left(D^{\otimes k-1}_{\underline{\lambda}} \boldsymbol{\phi}(\underline{\lambda}) \right).$$

Jammalamadaka et al. (2020) show that linear and nonlinear functions of $\kappa_{\underline{X},3}$ and $\kappa_{\underline{X},4}$ cover all existing cumulant-based indexes of skewness and kurtosis, revealing several connections and equivalences among them, that have not been noticed before. Let $\mathbf{\Sigma}^{-1/2}$ denote the symmetric positive definite square root of $\mathbf{\Sigma}^{-1}$; in the paper, $\underline{Y}$ will always denote the normalized version of $\underline{X}$, that is,

$$\underline{Y} = \mathbf{\Sigma}^{-1/2} \left(\underline{X} - \underline{\mu} \right). \tag{2}$$

Also, given a random sample $\underline{X}_1, \ldots \underline{X}_n$ of identical copies of $\underline{X}$ we define $\underline{Z}_j = \widehat{\mathbf{\Sigma}}^{-1/2} \left(\underline{X}_j - \overline{\underline{X}} \right)$, $j = 1, \ldots, n$, to be the sample version of $\underline{Y}$; $\overline{\underline{X}}$ and $\widehat{\mathbf{\Sigma}}$ are the usual sample estimates of $\underline{\mu}$ and $\mathbf{\Sigma}$; recall that we need $n > d + 1$ and absolute continuity of the distribution of $\underline{X}$ in order to have a nonsingular $\widehat{\mathbf{\Sigma}}$ almost surely (see Eaton & Perlman, 1973).

The paper is organized as follows. Section 2 provides the GC expansion for the density of a multivariate distribution and the main results of the paper. Section 3 discusses applications of the results by presenting new measures of skewness and kurtosis and providing new insights into ICA analysis based on scatter matrices. Section 4 provides some simulations which support the asymptotic results of the previous sections. Two final sections present a real data example and conclusions. An appendix contains the proofs, some technical details, and a general theorem on the cumulant vectors of multivariate elliptical distributions.

2 | MAIN RESULTS

2.1 | Density expansion and vector Hermite polynomials

The GC expansion expresses the density function of a random variable in terms of its cumulants (see e.g., Brenn & Anfinsen, 2017). Exploiting Equation (B1) in Section B1, which relates moments to cumulants via multivariate Bell polynomials B_k, the characteristic function ϕ of a d-variate random vector $\underline{X}$ can be written as

$$\phi_{\underline{X}}\left(\underline{\lambda}\right) = \sum_{k=0}^{\infty} \frac{\mathrm{i}^k}{k!} \mathrm{E}\underline{X}^{\top \otimes k} \underline{\lambda}^{\otimes k} = \sum_{k=0}^{\infty} \frac{\mathrm{i}^k}{k!} B_k \left(\underline{\kappa}_{\underline{X},1}, \underline{\kappa}_{\underline{X},2}, \ldots \underline{\kappa}_{\underline{X},k} \right)^{\top} \underline{\lambda}^{\otimes k}.$$

Let $\underline{\xi}$ be a Gaussian random d-vector with expected value $\mathrm{E}\underline{\xi} = \mathrm{E}\underline{X} = \underline{\mu}$, and variance vector $\underline{\kappa}_{\underline{\xi},2} = \underline{\mathrm{Cum}}_2\left(\underline{\xi}\right) = \underline{\mathrm{Cum}}_2\left(\underline{X}\right) = \underline{\kappa}_{\underline{X},2} = \mathrm{Vec}\,\boldsymbol{\Sigma}$, so that we have

$$\begin{aligned}\phi_{\underline{X}}\left(\underline{\lambda}\right) &= \exp\left(\sum_{k=0}^{\infty}\frac{\mathrm{i}^k}{k!}\left(\underline{\kappa}_{\underline{X},k}^{\top} - \underline{\kappa}_{\underline{\xi},k}^{\top}\right)\underline{\lambda}^{\otimes k}\right)\phi_{\underline{\xi}}\left(\underline{\lambda}\right)\\ &= \left(1 + \sum_{k=3}^{\infty}\frac{\mathrm{i}^k}{k!}\underline{B}_k\left(0,0,\underline{\kappa}_{\underline{X},3},\dots\underline{\kappa}_{\underline{X},k}\right)^{\top}\underline{\lambda}^{\otimes k}\right)\phi_{\underline{\xi}}\left(\underline{\lambda}\right).\end{aligned}$$

Finally, defining $\underline{Y} = \boldsymbol{\Sigma}^{-1/2}\left(\underline{X} - \underline{\mu}\right)$ and using the inverse Fourier transform, one can write the density of $\underline{Y}$ in the form of a GC series:

$$f_{\underline{Y}}\left(\underline{y}\right) = \left(1 + \sum_{k=3}^{5}\frac{1}{k!}\underline{\kappa}_{\underline{Y},k}^{\top}\underline{H}_k\left(\underline{y}\right) + \sum_{k=6}^{8}\frac{1}{k!}B_k\left(0,0,\underline{\kappa}_{\underline{Y},3},\dots,\underline{\kappa}_{\underline{Y},k}\right)^{\top}\underline{H}_k\left(\underline{y}\right)\right)\varphi\left(\underline{y}\right) + \mathcal{O}. \tag{3}$$

Note that in (3) we use an approximation exploiting only terms up to order 8 which are enough for our purposes, while $\mathcal{O}$ includes the remainder terms. In Equation (3), $\varphi\left(\underline{x}\right)$ denotes the density of a multivariate standard normal distribution, while $\underline{\kappa}_{\underline{Y},k}$ denotes the kth cumulant vector of $\underline{Y}$; recall that $\underline{\kappa}_{\underline{Y},1} = \underline{\mu} = 0$, and $\underline{\kappa}_{\underline{Y},2} = \mathrm{Vec}\,\mathbf{I}_d$.

$\underline{H}_k$ is a vector Hermite polynomial of order k, with variance vector $\underline{\kappa}_{\underline{Y}}$. In order to compute vector-multivariate Hermite polynomials, we use the definition in Holmquist (1996) who uses the symmetrizer matrix $\mathbf{S}_{d\mathbf{1}_k}$ for symmetrization of a T-product of k vectors with the same dimension d (see section B0.2). In particular $\mathbf{S}_{d\mathbf{1}_4}\left(\underline{a}_1\otimes\underline{a}_2\otimes\underline{a}_3\otimes\underline{a}_4\right)$ is a vector of dimension d^4, which is symmetric in $\underline{a}_j$. For example we have:

$$\begin{aligned}\underline{H}_3\left(\underline{y}\right) &= \mathbf{S}_{d\mathbf{1}_3}\left(\underline{y}^{\otimes 3} - 3\underline{\kappa}_{\underline{Y},2}\otimes\underline{y}\right),\\ \underline{H}_4\left(\underline{y}\right) &= \mathbf{S}_{d\mathbf{1}_4}\left(\underline{y}^{\otimes 4} - 6\underline{\kappa}_{\underline{Y},2}\otimes\underline{y}^{\otimes 2} + 3\underline{\kappa}_{\underline{Y},2}^{\otimes 2}\right).\end{aligned}$$

It can be verified that for $k = 3, 4, 5$, the Bell polynomials have the simple forms:

$$\begin{aligned}B_3\left(0,0,\underline{\kappa}_{\underline{Y},3}\right) &= \underline{\kappa}_{\underline{Y},3},\\ B_4\left(0,0,\underline{\kappa}_{\underline{Y},3},\underline{\kappa}_{\underline{Y},4}\right) &= \underline{\kappa}_{\underline{Y},4},\\ B_5\left(0,0,\underline{\kappa}_{\underline{Y},3},\underline{\kappa}_{\underline{Y},4},\underline{\kappa}_{\underline{Y},5}\right) &= \underline{\kappa}_{\underline{Y},5}.\end{aligned}$$

For higher-order terms we have a bit more complicated expressions, namely,

$$\begin{aligned}B_6\left(0,0,\underline{\kappa}_{\underline{Y},3},\underline{\kappa}_{\underline{Y},4},\underline{\kappa}_{\underline{Y},5},\underline{\kappa}_{\underline{Y},6}\right) &= \mathbf{S}_{d\mathbf{1}_6}\left(\underline{\kappa}_{\underline{Y},6} + 10\underline{\kappa}_{\underline{Y},3}^{\otimes 2}\right),\\ B_7\left(0,0,\underline{\kappa}_{\underline{Y},3},\underline{\kappa}_{\underline{Y},4},\underline{\kappa}_{\underline{Y},5},\underline{\kappa}_{\underline{Y},6},\underline{\kappa}_{\underline{Y},7}\right) &= \mathbf{S}_{d\mathbf{1}_7}\left(\underline{\kappa}_{\underline{Y},7} + 35\underline{\kappa}_{\underline{Y},3}\otimes\underline{\kappa}_{\underline{Y},4}\right),\\ B_8\left(0,0,\underline{\kappa}_{\underline{Y},3},\underline{\kappa}_{\underline{Y},4},\underline{\kappa}_{\underline{Y},5},\underline{\kappa}_{\underline{Y},6},\underline{\kappa}_{\underline{Y},7},\underline{\kappa}_{\underline{Y},8}\right) &= \mathbf{S}_{d\mathbf{1}_8}\left(\underline{\kappa}_{\underline{Y},8} + 56\underline{\kappa}_{\underline{Y},4}\otimes\underline{\kappa}_{\underline{Y},5} + 35\underline{\kappa}_{\underline{Y},4}^{\otimes 2}\right).\end{aligned}$$

By orthogonality of Hermite polynomials, one has $\mathrm{E}\underline{H}_k\left(\underline{Y}\right) = B_k\left(0,0,\underline{\kappa}_{\underline{Y},3},\dots,\underline{\kappa}_{\underline{Y},k}\right)$.

2.2 | Covariances

Using the GC expansion (3) and the results discussed above, we now provide general formulae for the covariances of $\underline{H}_3(\underline{Y})$ and $\underline{H}_4(\underline{Y})$. The results are presented in a vector form, that is, in the form of $\underline{\mathrm{Cum}}_2(\underline{H}_3(\underline{Y}))$ and $\underline{\mathrm{Cum}}_2(\underline{H}_4(\underline{Y}))$, and the proof is given in Appendix C.

Theorem 1. *Let $\underline{Y}$ be as defined in (2). Then, assuming the required moments exist, we have*

$$\underline{\mathrm{Cum}}_2\left(\underline{H}_3\left(\underline{Y}\right)\right) = \underline{\kappa}_{\underline{Y},6} + 10\,\mathbf{S}_{d\mathbf{1}_6}\underline{\kappa}_{\underline{Y},3}^{\otimes 2} + \mathbf{K}_{H4,2}^{-1}\left(\underline{\kappa}_{\underline{Y},4}\otimes\underline{\kappa}_{\underline{Y},2}\right) + \mathbf{K}_{3!}^{-1}\underline{\kappa}_{\underline{Y},2}^{\otimes 3} - \underline{\kappa}_{\underline{Y},3}^{\otimes 2}. \tag{4}$$

$$\begin{aligned}\underline{\mathrm{Cum}}_2\left(\underline{H}_4\left(\underline{Y}\right)\right) = {} & \underline{\kappa}_{\underline{Y},8} + \mathbf{S}_{d\mathbf{1}_8}\left(56\underline{\kappa}_{\underline{Y},4}\otimes\underline{\kappa}_{\underline{Y},5} + 35\underline{\kappa}_{\underline{Y},4}^{\otimes 2}\right)\\ & + \mathbf{K}_{H6,2}^{-1}\left(\left(\underline{\kappa}_{\underline{Y},6} + 10\mathbf{S}_{d\mathbf{1}_6}\underline{\kappa}_{\underline{Y},3}^{\otimes 2}\right)\otimes\underline{\kappa}_{\underline{Y},2}^{\otimes 2}\right)\\ & + \mathbf{K}_{H4,2,2}^{-1}\left(\boldsymbol{\kappa}_{\underline{Y},4}\otimes\underline{\kappa}_{\underline{Y},2}^{\otimes 2}\right) + \mathbf{K}_{4!}^{-1}\underline{\kappa}_{\underline{Y},2}^{\otimes 4} - \underline{\kappa}_{\underline{Y},4}^{\otimes 2}.\end{aligned} \tag{5}$$

The matrices $\mathbf{K}_{H\,4,2}$, $\mathbf{K}_{H\,2,2,2}$, $\mathbf{K}_{3!}$ $\mathbf{K}_{H\,6,2}$, $\mathbf{K}_{H\,4,2,2}$, and $\mathbf{K}_{4!}$ are commutator matrices whose computational formulae are given in Appendix B2. Note that in particular, $\mathbf{K}_{\mathfrak{p}}^{-1} = \mathbf{K}_{\mathfrak{p}}^{\top}$.

2.2.1 | The case of elliptically symmetric distributions

A d-vector $\underline{W}$ has a *spherically* symmetric distribution if its distribution is invariant under the group of rotations in $\mathbb{R}^d$. This is equivalent to saying that $\underline{W}$ has the stochastic representation $\underline{W} = R\underline{U}$, where R is a nonnegative random variable, $\underline{U}$ is uniform on sphere $\mathbb{S}_{d-1}$, and R and $\underline{U}$ are *independent* (see e.g. ,Fang et al., 2017, theorem 2.5). A d-vector $\underline{X}$ has an *elliptically* symmetric distribution if it has the representation

$$\underline{X} = \underline{\mu} + \boldsymbol{\Sigma}^{1/2}\underline{W},$$

where $\underline{\mu}\in\mathbb{R}^d$, $\boldsymbol{\Sigma}$ is a variance–covariance matrix, and $\underline{W}$ has a spherically symmetric distribution. Hence the cumulants of $\underline{X}$ are just constant times the cumulants of $\underline{W}$ except for the mean, that is, $\underline{\kappa}_{\underline{X},1} = \underline{\mu}$ and

$$\underline{\kappa}_{\underline{X},k} = \left(\boldsymbol{\Sigma}^{1/2}\right)^{\otimes k}\underline{\kappa}_{\underline{W},k}, \qquad k \geq 2. \tag{6}$$

Theorem 3 in Appendix A1 provides detailed formulae for the cumulant vectors of d-variate elliptically symmetric distributions. Note that from Theorem 1, under the assumption of elliptical symmetry where odd cumulants vanish, one has (note also that $\underline{\kappa}_{\underline{Y},2} = \mathrm{Vec}\,\mathbf{I}_d$)

$$\underline{\mathrm{Cum}}_2\left(\underline{H}_3\left(\underline{Y}\right)\right) = \underline{\kappa}_{\underline{Y},6} + \mathbf{K}_{H4,2}^{-1}\left(\underline{\kappa}_{\underline{Y},4}\otimes\mathrm{Vec}\,\mathbf{I}_d\right) + \mathbf{K}_{3!}^{-1}(\mathrm{Vec}\,\mathbf{I}_d)^{\otimes 3}. \tag{7}$$

$$\begin{aligned}\underline{\mathrm{Cum}}_2\left(\underline{H}_4\left(\underline{Y}\right)\right) = {} & \underline{\kappa}_{\underline{Y},8} + 35\mathbf{S}_{d\mathbf{1}_8}\underline{\kappa}_{\underline{Y},4}^{\otimes 2} + \mathbf{K}_{H6,2}^{-1}\left(\underline{\kappa}_{\underline{Y},6}\otimes(\mathrm{Vec}\,\mathbf{I}_d)^{\otimes 2}\right)\\ & + \mathbf{K}_{H4,2,2}^{-1}\left(\underline{\kappa}_{\underline{Y},4}\otimes(\mathrm{Vec}\,\mathbf{I}_d)^{\otimes 2}\right) + \mathbf{K}_{4!}^{-1}(\mathrm{Vec}\,\mathbf{I}_d)^{\otimes 4} - \underline{\kappa}_{\underline{Y},4}^{\otimes 2}.\end{aligned} \tag{8}$$

If in addition, we assume $\underline{\kappa}_{Y,4} = 0$, these expressions further simplify as

$$\underline{\mathrm{Cum}}_2\left(\underline{H}_3\left(\underline{Y}\right)\right) = \underline{\kappa}_{Y,6} + \mathbf{K}_{3!}^{-1}(\mathrm{Vec}\,\mathbf{I}_d)^{\otimes 3}.$$
$$\underline{\mathrm{Cum}}_2\left(\underline{H}_4\left(\underline{Y}\right)\right) = \underline{\kappa}_{Y,8} + \mathbf{K}_{H6,2}^{-1}\left(\underline{\kappa}_{Y,6} \otimes (\mathrm{Vec}\,\mathbf{I}_d)^{\otimes 2}\right) + \mathbf{K}_{4!}^{-1}(\mathrm{Vec}\,\mathbf{I}_d)^{\otimes 4}.$$

2.2.2 | The Gaussian case

Under a further assumption of Gaussianity where all higher-order cumulants are null, we have

$$\underline{\mathrm{Cum}}_2\left(\underline{H}_3\left(\underline{Y}\right)\right) = \mathbf{K}_{3!}^{-1}(\mathrm{Vec}\,\mathbf{I}_d)^{\otimes 3},$$
$$\underline{\mathrm{Cum}}_2\left(\underline{H}_4\left(\underline{Y}\right)\right) = \mathbf{K}_{4!}^{-1}(\mathrm{Vec}\,\mathbf{I}_d)^{\otimes 4}.$$

2.2.3 | Computational aspects

Note that Theorem 1 provides an explicit form for the asymptotic covariances using two basic elements: the cumulant vectors $\underline{\kappa}_{Y,k}$ and commutator matrices (the symmetrizer is indeed a sum of commutators, see Appendix B2).

Further simplifications of the above formulae for the case of elliptically symmetric distributions are possible by exploiting the results of Theorem 3 in Appendix A1. These results connect the cumulants $\underline{\kappa}_{W,k}$, of a spherical random vector $\underline{W}$ to the corresponding cumulant of any element of $\underline{W}$ for any even k (odds cumulants are null). Using the result we have

$$\underline{\kappa}_{Y,k} = \mathrm{Cum}_k(W_1)\mathbf{S}_{d\mathbf{1}_k}(\mathrm{Vec}\,\mathbf{I}_d)^{\otimes k/2}, \quad k = 2, 4, \ldots \tag{9}$$

that is, we can use univariate marginal cumulants to compute the vector cumulants. Formulae for the commutators $\mathbf{K}_{H\,4,2}$, $\mathbf{K}_{H\,2,2,2}$, $\mathbf{K}_{3!}$ $\mathbf{K}_{H\,6,2}$, $\mathbf{K}_{H\,4,2,2}$, and $\mathbf{K}_{4!}$, together with some more technical details are provided in Appendix B2. We point out that the use of symmetrizers, although quite useful in theoretical development, may quickly get intractable from the computational point of view since computing $\mathbf{S}_{d\mathbf{1}_k}$ requires $k!$ operations.

The use of the symmetrizers can be avoided by using sums of commutators which turn out to be much faster to compute with a software code. Details on these can be obtained from the authors.

Overall, Theorem 1 provides a feasible way, from both estimation and computational points of view, to determine the covariance matrices $\mathbf{C}_{\underline{H}_3}$ and $\mathbf{C}_{\underline{H}_4}$.

2.3 | A general CLT

This section provides general theorems concerning the asymptotic normality of the estimated third and fourth cumulant vectors. Consider a random sample $\underline{X}_1, \ldots, \underline{X}_n$ and the corresponding standardized version $\underline{Z}_j, j = 1, \ldots, n$. For any function g, let

$$\overline{g(\underline{Z})} = \frac{1}{n}\sum_{j=1}^{n} g(\underline{Z}_j).$$

The cumulant vector $\underline{\kappa}_{\underline{Y},3}$ (which is actually the third-order central moment) and $\underline{\kappa}_{\underline{Y},4}$, can be simply estimated by the method of moments; noting that $\overline{\underline{Z}} = 0$, we have

$$\hat{\underline{\kappa}}_{\underline{Z},3} = \overline{\underline{Z}^{\otimes 3}} = \overline{\underline{H}_3\left(\underline{Z}\right)} \quad \text{and} \quad \hat{\underline{\kappa}}_{\underline{Z},4} = \overline{\underline{H}_4\left(\underline{Z}\right)} = \overline{\underline{Z}^{\otimes 4}} - \mathbf{K}_{2,2}[\text{Vec}\,\mathbf{I}_d]^{\otimes 2}. \tag{10}$$

The formula for $\mathbf{K}_{2,2}$ is given in Appendix B2. The following theorem provides a result on the asymptotic normality of the estimated cumulant vectors, and the proof is given in Appendix C.

Theorem 2. *Let $\underline{X}_1, \dots, \underline{X}_n$ denote a random sample from a d-variate distribution with mean vector $\underline{\mu}$ and covariance matrix $\mathbf{\Sigma}$. If the required moments exist, then for $q = 3, 4$,*

$$\sqrt{n}\left(\hat{\underline{\kappa}}_{\underline{Z},q} - \kappa_{\underline{Y},q}\right),$$

is asymptotically normal with mean 0 and covariance matrix $\mathbf{C}_{\underline{H}_q}$, where $\text{Vec}\,\mathbf{C}_{\underline{H}_3} = \underline{\text{Cum}}_2\left(\underline{H}_3\left(\underline{Y}\right)\right)$ *and* $\text{Vec}\,\mathbf{C}_{\underline{H}_4} = \underline{\text{Cum}}_2\left(\underline{H}_4\left(\underline{Y}\right)\right) + 16\underline{\kappa}_{\underline{Y},3}^{\otimes 2} \otimes \text{Vec}\,\mathbf{I}_d$. *Formulae for* $\underline{\text{Cum}}_2\left(\underline{H}_q\left(\underline{Y}\right)\right)$, $q = 3, 4$ *are given in Theorem 1.*

Remark 1. Comparing Theorem 2 with theorem 2.1 of Klar (2002), we note that in the latter covariances are expressed in terms of sample statistics $\underline{Z}$ while in our Theorem 2, only the population model $\underline{Y}$ is considered. Moreover in Theorem 2 formulae for $C_{\underline{H}_q}$ are explicit, which makes it more practical for real applications.

This difference can be better noticed by observing that, although Theorem 2 states that, for symmetric multivariate distributions, the asymptotic distributions of $\sqrt{n}\overline{H_q\left(\underline{Z}\right)}$ and $\sqrt{n}\overline{H_q\left(\underline{Y}\right)}$, $q = 3, 4$ are the same, the Hermite polynomials of $\underline{Z}$ and $\underline{Y}$ are not the same since $\overline{\underline{Z}} = 0$ but $\overline{\underline{Y}}$ may not be necessarily null. Indeed note that, unlike in Equation (10), $\underline{H}_3\left(\underline{Y}\right) = \mathbf{S}_{d\mathbf{1}_3}\left(\underline{Y}^{\otimes 3} - 3\underline{Y} \otimes \underline{\kappa}_{\underline{Y},2}\right) \neq \underline{Y}^{\otimes 3}$ and $\underline{H}_4\left(\underline{Y}\right) = \mathbf{S}_{d\mathbf{1}_4}\left(\underline{Y}^{\otimes 4} - 6\underline{\kappa}_{\underline{Y},2} \otimes \underline{Y}^{\otimes 2} + 3\underline{\kappa}_{\underline{Y},2}^{\otimes 2}\right)$.

3 | APPLICATIONS

3.1 | New measures of skewness and kurtosis based on distinct cumulants

The vectors $\underline{H}_q(\underline{Y})$, $q = 3, 4$, contain d^q elements, which are not all distinct. Just as the covariance matrix of a d-dimensional vector contains only $d(d+1)/2$ distinct elements, a simple computation shows that $\underline{H}_q(\underline{Y})$ contains $\binom{d+q-1}{q}$ distinct elements.

It makes good sense in some cases to work only with these *distinct* elements of the cumulant vector. The selection of distinct elements can be accomplished via linear transformations and through the so-called elimination matrix which we denote here as $\mathbf{E}_{d,q}^{+}$ (see Meijer, 2005 and Jammalamadaka et al., 2020 for some details). The linear transformation

$$\underline{H}_{q,D}\left(\underline{Y}\right) = \mathbf{E}_{d,q}^{+}\underline{H}_q\left(\underline{Y}\right), \tag{11}$$

contains only the distinct values of $\underline{H}_q(\underline{Y})$ and has variance $\mathbf{C}_{\underline{H}_{q,D}} = \mathbf{E}_{d,q}^{+}\mathbf{C}_{\underline{H}_q}\mathbf{E}_{d,q}^{+\top}$. Given the availability of explicit formulae for the covariance matrices, it is natural to consider the weighted skewness and kurtosis measures based on $\underline{\kappa}_{\underline{Y},q,D} = \mathbf{E}^{+}_{d,q}\underline{\kappa}_{\underline{Y},q}$, $q = 3, 4$, the third- and fourth-order cumulant vectors with distinct cumulants, respectively; denote these as $\beta_{1,T,d}$ (skewness) and $\beta_{2,T,d}$ (kurtosis). The corresponding sample version is obtained by replacing the cumulant vectors with their estimators, that is,

$$b_{q-2,T,d} = \left\| \mathbf{C}^{-1/2}_{\underline{H}_{q,D}} \; \underline{\hat{\kappa}}_{\underline{Y},q,D} \right\|^2, \quad q = 3, 4. \tag{12}$$

As far as the asymptotic distributions of $b_{i,T,d}$ $i = 1, 2$ is concerned, exploiting Theorem 2 (see also example A, p. 130 in Serfling (2009)), one has the following

Proposition 1. *Let $\underline{X}_1, \ldots, \underline{X}_n$ be a random sample from a d-variate distribution with mean vector $\underline{\mu}$ and covariance matrix $\mathbf{\Sigma}$ and let $q = 3, 4$. If the required moments exist, the asymptotic distribution of $nb_{q-2,T,d}$ is a noncentral χ^2 distribution with $\binom{d+q-1}{q}$ degrees of freedom and noncentrality parameter $\beta_{q-2,T,d}$.*

Remark 2. A Slutsky-type argument shows that the above result also holds asymptotically when $\mathbf{C}_{\underline{H}_{q,D}}$ is estimated from the data.

Remark 3. In the Gaussian case, the computational formula for $b_{1,T,d}$, for $1 \leq j, k, l \leq d$, is

$$\hat{b}_{1,T,d} = \frac{1}{6}\sum_{j}\left(\frac{1}{n}\sum_{i=1}^{n} Z_{ji}^3\right)^2 + \frac{1}{2}\sum_{j\neq k}\left(\frac{1}{n}\sum_{i=1}^{n} Z_{ji}^2 Z_{ki}\right)^2 + \sum_{j<k<l}\left(\frac{1}{n}\sum_{i=1}^{n} Z_{ji}Z_{ki}Z_{li}\right)^2. \tag{13}$$

This expression coincides with the first nonzero component of Neyman's smooth test. Observing (13) one may also note that the popular Mardia's index of skewness (Mardia, 1970) $b_{1,d}$ satisfies the equation $b_{1,d} = 6 \cdot b_{1,T,d}$. Thus Mardia's index can be interpreted as a weighted skewness measure in the Gaussian case.

Since the distribution of $b_{1,T,d}$ with estimated covariance matrix is asymptotically a χ^2 with $\binom{d+2}{3}$ degrees of freedom either in the Gaussian or in the more general case of elliptically symmetric multivariate distributions, a consistent test for elliptical symmetry can be based on $b_{1,T,d}$ with the estimated covariance $\mathbf{C}_{\underline{H}_{3,D}}$. Given the results in Sections A1 and 2.2.1, under the null hypothesis, estimation of the covariance matrix only requires estimation of the marginal cumulants.

Under the alternative hypothesis of asymmetry, the asymptotic distribution will involve a noncentrality parameter $\beta_{1,T,d}$.

3.2 | Skewness and kurtosis measures of Móri et al. (1994)

As an illustration, in this section, we consider the indexes discussed in Móri, Székely and Rohatgi (Móri et al. (1994) and connect them to the skewness and kurtosis vectors; for further examples along these lines, see Jammalamadaka et al. (2020). Móri et al. (1994) define the "*skewness vector*" $b\left(\underline{Y}\right)$ of $\underline{Y}$ as the quantity

$$b\left(\underline{Y}\right) = \mathrm{E}\left[\underline{Y}^\top \underline{Y}\right]\underline{Y} = \left((\mathrm{Vec}\,\mathbf{I}_d)^\top \otimes \mathbf{I}_d\right)\underline{\kappa}_{\underline{Y},3}. \tag{14}$$

Let $\mathbf{C}_b$ denote the covariance matrix of $b(\underline{Y})$. Then $\operatorname{Vec}\mathbf{C}_b = \left((\operatorname{Vec}\mathbf{I}_d)^\top \otimes \mathbf{I}_d\right)^{\otimes 2}\underline{\operatorname{Cum}}_2\left(\underline{H}_3\left(\underline{Y}\right)\right)$ which, under the Gaussian assumption, $\mathbf{C}_b$ takes the form

$$\operatorname{Vec}\mathbf{C}_b = \left((\operatorname{Vec}\mathbf{I}_d)^\top \otimes \mathbf{I}_d\right)^{\otimes 2}\mathbf{K}_{3!}^{-1}(\operatorname{Vec}\mathbf{I}_d)^{\otimes 3}.$$

The index of skewness of Móri et al. (1994) is defined as $\tilde{b}_{1,d} = \left\|\left((\operatorname{Vec}\mathbf{I}_d)^\top \otimes \mathbf{I}_d\right)\overline{\underline{Z}^{\otimes 3}}\right\|^2$. From Theorem 2 and the computations above, one has (see also Klar, 2002 and Henze (2002), under the assumption of Gaussianity,

$$\frac{n\tilde{b}_{1,d}}{2(d+2)} \xrightarrow{D} \chi_d^2. \tag{15}$$

In the special case that the distribution is elliptically symmetric, evaluation of $\mathbf{C}_b$ (details omitted), shows that

$$\frac{n\tilde{b}_{1,d}}{(15+(d-1)(d+7))\,\kappa_6/15+\left((d+1)^2+23\right)\kappa_4/3+2(d+2)} \xrightarrow{D} \chi_d^2. \tag{16}$$

For measuring kurtosis, Móri et al. (1994) suggest the *kurtosis matrix* $B\left(\underline{Y}\right)$ of $\underline{Y}$ which is defined by the quantity

$$B\left(\underline{Y}\right) = \operatorname{E}\left(\underline{Y}^\top\underline{Y}\right)\underline{YY}^\top-(d+2)\,\mathbf{I}_d = \operatorname{Vec}B\left(\underline{Y}\right) = \left(\mathbf{I}_{d^2}\otimes(\operatorname{Vec}\mathbf{I}_d)^\top\right)\underline{\kappa}_{\underline{Y},4}. \tag{17}$$

One can compute the variance of $\left(\mathbf{I}_{d^2}\otimes(\operatorname{Vec}\mathbf{I}_d)^\top\right)\underline{H}_4\left(\underline{Y}\right)$, and derive its asymptotic distributions, under different hypotheses, using the results of Theorems 1 and 2.

3.2.1 | A note on ICA

There are some interesting connections between ICA based on scatter matrices and Móri et al. (1994) indexes defined above which allow one to exploit Theorems 1 and 2 for deriving asymptotic results and tests. Consider the model

$$\underline{X} = \mathbf{A}\underline{V} + \underline{b},$$

where the matrices of constants $\mathbf{A}$ and $\underline{b}$ are $d\times d$ and $d\times 1$, respectively, and $\underline{V}$ is a d-vector of independent random variables with null mean vector and unit covariance matrix; multivariate normality in $\underline{V}$ is excluded. Oja et al. (2006) then define *Location vector* and *Scatter matrix* functionals. The location vector $\underline{T}\left(\underline{X}\right)$ is a d-vector, which is 'location-affine equivariant' in the sense that $\underline{T}\left(\mathbf{A}\underline{X}+\underline{b}\right) = \mathbf{A}\underline{T}\left(\underline{X}\right)+\underline{b}$. In practice the location vector of order 1 is the mean, that is, $\underline{T}_1\left(\underline{X}\right) = E\underline{X}$, and the location vector of order 2 based on third moments is a vector, which can be connected to Móri et al. (1994) skewness vector $b(\underline{Y})$ defined in (14), since

$$\begin{aligned}\underline{T}_2\left(\underline{X}\right) &= \frac{1}{d}E\left(\underline{X}-\underline{\mu}\right)^\top\boldsymbol{\Sigma}^{-1}\left(\underline{X}-\underline{\mu}\right)\left(\underline{X}-\underline{\mu}\right) = \frac{1}{d}\boldsymbol{\Sigma}^{1/2}E\left[\underline{Y}^\top\underline{Y}\right]\underline{Y} = \frac{1}{d}\boldsymbol{\Sigma}^{1/2}E\underline{YY}^\top\underline{Y}\\ &= \frac{1}{d}\boldsymbol{\Sigma}^{1/2}b\left(\underline{Y}\right) = \frac{1}{d}\boldsymbol{\Sigma}^{1/2}\left((\operatorname{Vec}\mathbf{I}_d)^\top\otimes\mathbf{I}_d\right)\underline{\kappa}_{\underline{Y},3}^{\otimes}.\end{aligned}$$

The Scatter matrix $\mathbf{S}\left(\underline{X}\right)$ is a $d\times d$-matrix, which is positive definite and "scatter-affine equivariant" in the sense that $\mathbf{S}\left(\mathbf{A}\underline{X}+\underline{b}\right)=\mathbf{A}\mathbf{S}\left(\underline{X}\right)\mathbf{A}^{\top}$. The Scatter matrix of order 1 is the covariance matrix, that is, $\mathbf{S}_1\left(\underline{X}\right)=\mathrm{Var}\underline{X}$, and the Scatter matrix of order 2 is connected to the "kurtosis" matrix (17), since:

$$\begin{aligned}\mathbf{S}_2\left(\underline{X}\right)&=\frac{1}{d+2}E\left(\left(\underline{X}-\underline{\mu}\right)^{\top}\boldsymbol{\Sigma}^{-1}\left(\underline{X}-\underline{\mu}\right)\right)\left(\underline{X}-\underline{\mu}\right)\left(\underline{X}-\underline{\mu}\right)^{\top}\\&=\frac{1}{d+2}\boldsymbol{\Sigma}^{1/2}E\left(\underline{Y}^{\top}\underline{Y}\right)\underline{Y}\underline{Y}^{\top}\boldsymbol{\Sigma}^{1/2}=\frac{1}{d+2}\boldsymbol{\Sigma}^{1/2}B\left(\underline{Y}\right)\boldsymbol{\Sigma}^{1/2}+\boldsymbol{\Sigma}.\end{aligned}$$

The scatter matrix $\mathbf{S}_2\left(\underline{X}\right)$ can be written in terms of the kurtosis $\underline{\kappa}_{\underline{Y},4}$ of $\underline{Y}$ as:

$$\mathrm{Vec}\,\mathbf{S}_2\left(\underline{X}\right)=\frac{1}{d+2}\left(\boldsymbol{\Sigma}^{1/2}\right)^{\otimes 2}\left(\mathbf{I}_{d^2}\otimes\left(\mathrm{Vec}\,\mathbf{I}_d\right)^{\top}\right)\underline{\kappa}^{\otimes}_{\underline{Y},4}+\underline{\kappa}^{\otimes}_{\underline{X},2}. \tag{18}$$

If $\underline{V}$ satisfies the *independence property*, then $\boldsymbol{\Sigma}=\mathbf{I}_d$ and

$$\begin{aligned}B\left(\underline{Y}\right)&=\left(\mathbf{I}_{d^2}\otimes\left(\mathrm{Vec}\,\mathbf{I}_d\right)^{\top}\right)\underline{\kappa}_{\underline{V},4}=\sum_{i=1}^{d}\kappa_{V_i,4}\left(\mathbf{I}_{d^2}\otimes\left(\mathrm{Vec}\,\mathbf{I}_d\right)^{\top}\right)\underline{\mathbf{e}}_i^{\otimes 4}=\sum_{i=1}^{d}\kappa_{V_i,4}\,\underline{\mathbf{e}}_i^{\otimes 2},\\\mathbf{S}_2\left(\underline{V}\right)&=\frac{1}{d+2}\mathrm{diag}\left[\kappa_{V_i,4}\right]+\mathbf{I}_d.\end{aligned}$$

To test the hypothesis of independence of the components of $\underline{V}$, a natural test statistic is the fourth order weighted statistics $b_{2,T,d}$ (12). In order to implement the test in practice, let $\underline{\kappa}_{\underline{V}\setminus d,4}$ denote the kurtosis vector $\underline{\kappa}_{\underline{V},4}$ without $\kappa_{V_1,4},\dots,\kappa_{V_d,4}$ and let $b^*_{2,T,d}$ denote the measure given in (12) computed using $\underline{\kappa}_{\underline{V}\setminus d,4}$ with the corresponding covariance terms. Then, from Proposition 1, the test statistic $nb^*_{2,T,d}$, under the null, has a central χ^2 distribution with $\binom{d+3}{4}-d$ degrees of freedom. Under the alternative, this test statistic will have a noncentral chi square distribution with $\binom{d+3}{4}-d$ degrees of freedom and noncentrality parameter $\beta^*_{2,T,d}$.

4 | SIMULATIONS

In this section, some numerical results will be provided in support of, and utilizing the different theoretical results obtained here.

4.1 | Testing hypotheses on skewness

In this subsection, we will consider an application for testing the hypothesis

$$H_0:\underline{\kappa}_{Y,3}=0, \tag{19}$$

which holds under symmetry. Here a Monte Carlo experiment is performed in order to compare: (a) Mardia's index $b_{1,d}$ (see Remark 3); (b) Mori et al. index $\tilde{b}_{1,d}$ (see Section 3.2), and (c) the new index based on distinct elements $b_{1,T,d}$ (see Section 3.1), as competing criteria for testing the null hypothesis (19). The frequencies of rejection of the hypothesis in the tables are based on $M = 1000$ replications of the same experiment for each sample size $n = 250, 500, 1000$, and 2000. The following trivariate distributions are considered: (a) standard Normal; (b) t with 10 degrees of freedom; (c) Skew-normal (see Azzalini & Dalla Valle, 1996) with skew vector $\alpha = (-1, 1, 1)^{\top}$ and covariance matrix

$$\boldsymbol{\Sigma} = \begin{pmatrix} 1 & 0.5 & 0 \\ 0.5 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}. \tag{20}$$

For the asymptotic distribution of $b_{1,d}$ we may refer to Klar (2002) (see also Baringhaus & Henze, 1992), who shows that $nb_{1,d}$ has an asymptotic distribution which is a weighted sum of independent χ^2 distributions, namely

$$nb_{1,d} \xrightarrow{D} \alpha_1 \chi^2_d + \alpha_2 \chi^2_{d(d+1)(d+4)/6}, \tag{21}$$

where $\alpha_1 = \frac{3}{d}\left(\frac{E\|\underline{Y}\|^6}{d+2} - 2E\|\underline{Y}\|^4 + d(d+2)\right)$ and $\alpha_2 = \frac{6E\|\underline{Y}\|^6}{d(d+2)(d+4)}$. In particular, under the assumption of Gaussianity, $\alpha_1 = \alpha_2 = 6$ and the limiting distribution in (21) reduces to a $\chi^2_{\binom{d+2}{3}}$ distribution.

Remark 4. Observe that nowhere in this section, the assumption of normality is imposed on the data. The availability of consistent estimators of moments (and cumulants) assures that result (21) actually provides an asymptotic result for $b_{1,d}$ in the general case of elliptical symmetry. The same reasoning applies to (16) in the case of $\tilde{b}_{1,d}$. Also, estimation of the covariance matrix of $\hat{\kappa}_{\underline{Z},3}$ will be carried out under the symmetry assumption exploiting the formulae given in Theorem 3.

The covariance matrix to compute $b_{T,d}$ has been estimated exploiting (7) and (9); also $\hat{\kappa}_6 = \hat{\mu}_6 - 15\hat{\mu}_4 + 30$, $\hat{\kappa}_4 = \hat{\mu}_4 - 3$ with

$$\hat{\mu}_k = \frac{1}{dn}\sum_{j=1}^{d}\sum_{i=1}^{n} Z_{ij}^k.$$

The values $\hat{\mu}_k$ have also been used to estimate the parameters for determining the asymptotic distributions of $b_{1,d}$ and $\tilde{b}_{1,d}$.

Tables 1,2, and 3 report, for the three indexes discussed, the empirical frequencies of samples beyond the quantiles (0.9, 0.95, 0.99) of the corresponding asymptotic distribution. Results of these tables can be used to verify the correctness of the theoretical results and the relative performance of the three tests in testing H_0. Recall that in all the tables, the asymptotic distribution of $b_{1,d}$ and $\tilde{b}_{1,d}$ is always determined, respectively, from (21) and (16) with estimated parameters.

Note that in Tables 1 and 2 all indexes have an actual rejection rate very close to the nominal level for all sample sizes. On the other hand, in the case of Table 3, $\tilde{b}_{1,d}$ shows slightly lower power with respect to the other two indexes, both of which have comparable performance.

TABLE 1 Relative frequencies of samples declared significant by the symmetry indexes for tests at size 0.10, 0.05, 0.01. Results are relative to 1000 samples of varying size from a $N(0, \mathbf{I}_3)$-distribution

		$b_{1,d}$			$\tilde{b}_{1,d}$			$b_{1,T,d}$		
	Sig →	**0.10**	**0.05**	**0.01**	**0.10**	**0.05**	**0.01**	**0.10**	**0.05**	**0.01**
$n\rightarrow$	250	0.114	0.056	0.006	0.088	0.046	0.010	0.111	0.052	0.005
	500	0.106	0.057	0.012	0.103	0.059	0.014	0.110	0.057	0.013
	1000	0.117	0.057	0.015	0.083	0.049	0.007	0.116	0.059	0.013
	2000	0.105	0.060	0.010	0.099	0.049	0.009	0.102	0.061	0.010

TABLE 2 Relative frequencies of samples declared significant by the symmetry indexes for tests at size 0.10, 0.05, 0.01. Results are relative to 1000 samples of varying size from a t_{10}-distribution with covariance matrix $\frac{8}{10}\mathbf{I}_3$

		$b_{1,d}$			$\tilde{b}_{1,d}$			$b_{1,T,d}$		
	Sig →	**0.10**	**0.05**	**0.01**	**0.10**	**0.05**	**0.01**	**0.10**	**0.05**	**0.01**
$n\rightarrow$	250	0.158	0.086	0.025	0.099	0.036	0.006	0.154	0.084	0.026
	500	0.129	0.073	0.024	0.088	0.004	0.008	0.124	0.065	0.021
	1000	0.114	0.060	0.018	0.096	0.051	0.007	0.104	0.063	0.011
	2000	0.120	0.072	0.018	0.093	0.048	0.013	0.118	0.072	0.018

TABLE 3 Relative frequencies of samples declared significant by the symmetry indexes for tests at size 0.10, 0.05, 0.01. Results are relative to 1000 samples of varying size from a Skew-Normal-distribution with $\alpha = (-1, 1, 1)^\top$ and $\mathbf{\Sigma}$ as in (20)

		$b_{1,d}$			$\tilde{b}_{1,d}$			$b_{1,T,d}$		
	Sig →	**0.10**	**0.05**	**0.01**	**0.10**	**0.05**	**0.01**	**0.10**	**0.05**	**0.01**
$n\rightarrow$	250	0.298	0.184	0.068	0.292	0.165	0.048	0.309	0.188	0.065
	500	0.480	0.361	0.167	0.443	0.310	0.131	0.486	0.366	0.171
	1000	0.734	0.614	0.388	0.702	0.576	0.360	0.736	0.612	0.383
	2000	0.959	0.935	0.847	0.946	0.902	0.766	0.957	0.936	0.847

4.2 | Testing hypotheses on kurtosis

This subsection provides results of a Monte Carlo experiment that is performed in order to compare the kurtosis indexes $\tilde{b}_{2,d}$ and $b_{2,T,d}$ defined in the previous section as criteria for testing the hypothesis:

$$H_0 : \underline{\kappa}_{Y,4} = 0. \tag{22}$$

Recall also that the test for independence of Section 3.2.1 can be seen as a subcase of (22). For varying sample sizes, using $M = 1000$ replications, respectively from the following trivariate distributions: (a) standard Normal; (b) t with 20 degrees of freedom;

Note that in Table 4 both indexes have an actual rejection rate very close to the nominal level for all sample sizes. Table 5 shows that the two indexes have comparable performance, with a

TABLE 4 Relative frequencies of samples declared significant by the kurtosis indexes for tests at size 0.10, 0.05, 0.01. Results are relative to 1000 samples of varying size from a $N(0, \mathbf{I}_3)$-distribution

		$\tilde{b}_{2,d}$			$b_{2,T,d}$		
	Sig.→	**0.10**	**0.05**	**0.01**	**0.10**	**0.05**	**0.01**
n→	250	0.098	0.045	0.005	0.096	0.066	0.033
	500	0.115	0.052	0.012	0.093	0.059	0.014
	1000	0.131	0.064	0.011	0.085	0.052	0.016
	2000	0.096	0.045	0.011	0.095	0.053	0.012

TABLE 5 Relative frequencies of samples declared significant by the kurtosis indexes for tests at size 0.10, 0.05, 0.01. Results are relative to 1000 samples of varying size from a 3-variate t-distribution with 20 degrees of freedom.

		$\tilde{b}_{2,d}$			$b_{2,T,d}$		
	Sig.→	**0.10**	**0.05**	**0.01**	**0.10**	**0.05**	**0.01**
n→	250	0.573	0.480	0.325	0.568	0.501	0.379
	500	0.844	0.791	0.647	0.771	0.712	0.586
	1000	0.989	0.978	0.933	0.948	0.922	0.847
	2000	1.000	0.999	0.996	0.998	0.997	0.987

slight advantage for $b_{2,T,d}$ over $\tilde{b}_{2,d}$ for small sample sizes, and the other way around for larger sample sizes.

5 | A PRACTICAL APPLICATION

The Australian Institute of Sport (AIS) data (Weisberg, 2002) contains various biomedical measurements on a group of 202 athletes (102 males and 100 females). Here, we use a subset of the variables in the data set (viz., *Percentage of Body Fat*, *Body Mass Index*, *Lean Body Mass* and *Sum of Skin Folds*). The same variables have been used by Azzalini and Dalla Valle (1996) and modeled using a multivariate ($d = 4$) skew normal distribution. Here we use these data to compute and compare measures of skewness and kurtosis. Note from Figure 1, that a closer look at the data shows that distributions by gender are quite different and this may partly explain the phenomenon. We adapt indexes of skewness and kurtosis to the whole group of data, and separately for the subgroups of males and females.

Table 6 reports the estimated values of the skewness indexes $b_{1,d}$, $\tilde{b}_{1,d}$, and $b_{T,d}$ and those of the kurtosis indexes $\tilde{b}_{2,d}$ and $b_{2,T,d}$ for the three groups of athletes. The p-value in the (estimated) asymptotic distribution are used to interpret correctly the position of the estimated value within its distribution. Note that, although the three indexes of skewness exhibit slightly different tendencies (e.g., $b_{T,d}$ increases in both the subgroups of males and females), the p-values for all three skewness indexes indicate a marked asymmetry in all the groups.

For the case of kurtosis, there is a discrepancy between $\tilde{b}_{2,d}$ and $b_{2,T,d}$ in the female subgroup, where the latter does not detect a marked kurtosis. This may be due to a higher sensitivity of the index $b_{2,T,d}$ which uses all the mixed cumulants of the fourth order, while $\tilde{b}_{2,d}$ does not.

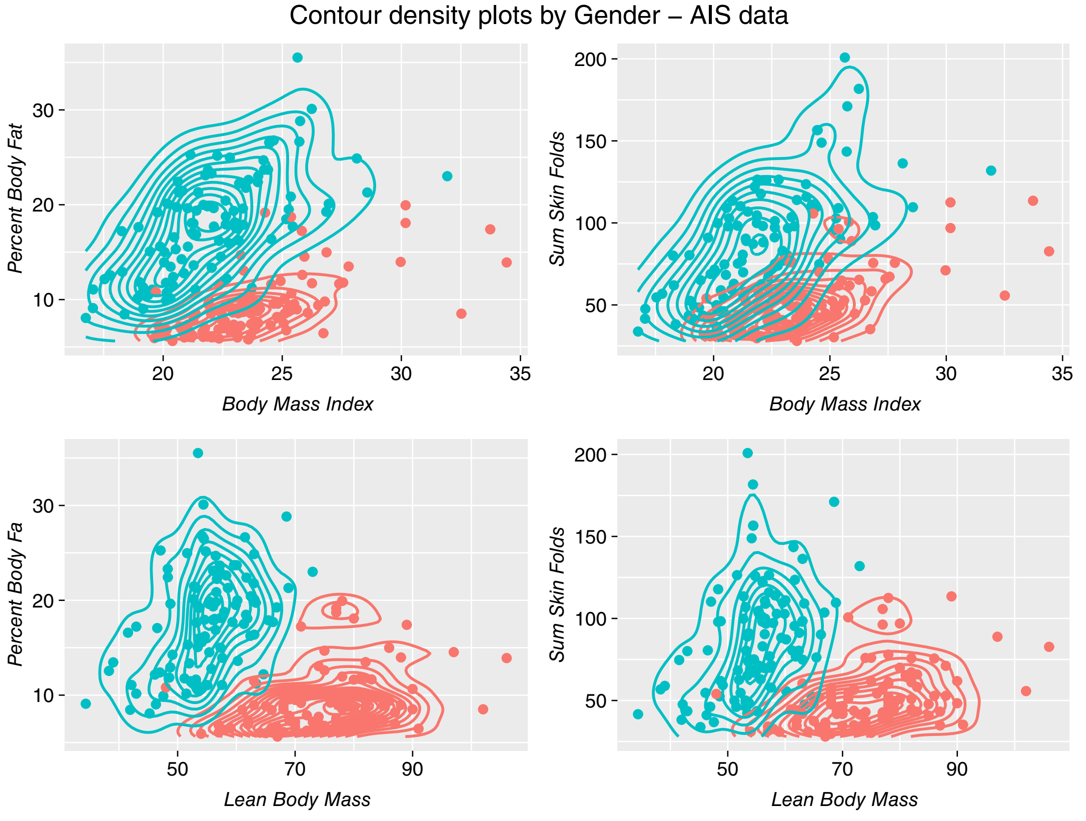

FIGURE 1 Australian Institute of Sport data: contour density plots. Blue points: Males; Red points: Females [Colour figure can be viewed at wileyonlinelibrary.com]

TABLE 6 Australian Institute of Sport data: skewness and kurtosis indexes and their *p*-values for all athletes and the subgroups of males and females

	$b_{1,d}$		$\tilde{b}_{1,d}$		$b_{T,d}$		$\tilde{b}_{2,d}$		$b_{2,T,d}$	
	Est.	**Sig.**	**Est.**	**Sig.**	**Est.**	**Sig.**	**Est.**	**Sig.**	**Est.**	**Sig.**
Males	7.84	0.000	9.82	0.000	0.64	0.000	11.0	0.000	2.77	0.000
Females	4.71	0.001	2.40	0.000	0.46	0.000	2.13	0.124	0.91	0.000
All	6.10	0.000	4.61	0.000	0.39	0.000	4.17	0.000	1.56	0.000

6 | SUMMARY AND CONCLUSIONS

A general framework for analyzing the asymptotic distributions of cumulant vectors of multivariate distributions is presented here, focusing in particular on the third- and fourth-order cumulant vectors and various statistics based on them. Formulae for asymptotic covariances, which can be obtained through computational algorithms, are provided.

The availability of such general formulae for the covariances helps, for instance, in obtaining simple new measures of skewness and kurtosis or in developing asymptotic distributions under different situations that prove useful in estimation as well as in efficiency and power comparisons.

Moreover, the approach through cumulant vectors allows reinterpretation of existing statistics and to readily obtain their asymptotic covariances, as demonstrated here for the ICA and Móri et al. (1994) indexes of skewness and kurtosis.

ACKNOWLEDGEMENTS

The authors would like to thank the reviewers for very helpful suggestions that improved the presentation. The work of Dr. Gy. H. Terdik was partially supported by the Project EFOP3.6.2-16-2017-00015 by the European Union, and co-financed by the European Social Fund.

ORCID

Emanuele Taufer https://orcid.org/0000-0001-5140-9912

REFERENCES

Azzalini, A., & Dalla Valle, A. (1996). The multivariate skew-normal distribution. *Biometrika*, *83*(4), 715–726.

Balakrishnan, N., Brito, M. R., & Quiroz, A. J. (2007). A vectorial notion of skewness and its use in testing for multivariate symmetry. *Communications in Statistics, Theory and Methods*, *36*(9), 1757–1767.

Baringhaus, L. (1991). Testing for spherical symmetry of a multivariate distribution. *The Annals of Statistics*, *19*(2), 899–917.

Baringhaus, L., & Henze, N. (1991). Limit distributions for measures of multivariate skewness and kurtosis based on projections. *Journal of Multivariate Analysis*, *38*(1), 51–69.

Baringhaus, L., & Henze, N. (1992). Limit distributions for Mardia's measure of multivariate skewness. *The Annals of Statistics*, *20*(4), 1889–1902.

Berkane, M., & Bentler, P. M. (1986). Moments of elliptically distributed random variates. *Statistics & Probability Letters*, *4*(6), 333–335.

Brenn, T., & Anfinsen, S. N. (2017). *A revisit of the Gram-Charlier and Edgeworth series expansions*. Artikler Rapporter Og Annet (fysikk og teknologi), Department of Physics and Technology, University of Troms, The Arctic University of Norway.

Eaton, M. L., & Perlman, M. D. (1973). The non-singularity of generalized sample covariance matrices. *The Annals of Statistics*, *1*(4), 710–717.

Fang, K. W., Kotz, S., & Ng, K. W. (2017). *Symmetric multivariate and related distributions*. Chapman & Hall/CRC Press.

Graham, A. (2018). *Kronecker products and matrix calculus with applications*. Courier Dover Publications.

Henze, N. (1994a). On Mardia's kurtosis test for multivariate normality. *Communications in Statistics-Theory and Methods*, *23*(4), 1031–1045.

Henze, N. (1994b). The asymptotic behavior of a variant of multivariate kurtosis. *Communications in Statistics-Theory and Methods*, *23*(4), 1047–1061.

Henze, N. (1997a). Extreme smoothing and testing for multivariate normality. *Statistics & Probability Letters*, *35*(3), 203–213.

Henze, N. (1997b). Limit laws for multivariate skewness in the sense of Móri, Rohatgi and Székely. *Statistics & Probability Letters*, *33*(3), 299–307.

Henze, N. (2002). Invariant tests for multivariate normality: A critical review. *Statistical Papers*, *43*(4), 467–506.

Holmquist, B. (1996). The d-variate vector Hermite polynomial of order. *Linear Algebra and Its Applications*, *237/238*, 155–190.

Huang, H. H., Lin, S. H., Wang, C. P., & Chiu, C. Y. (2014). Adjusting mv-efficient portfolio frontier bias for skewed and non-mesokurtic returns. *The North American Journal of Economics and Finance*, *29*, 59–83.

Ilmonen, P., Nevalainen, J., & Oja, H. (2010). Characteristics of multivariate distributions and the invariant coordinate system. *Statistics & Probability Letters*, *80*(23-24), 1844–1853.

Jammalamadaka, S. R., Taufer, E., & Terdik, G. H. (2020). On multivariate skewness and kurtosis. *Sankhya A*, 1–38. https://dx.doi.org/10.1007/s13171-020-00211-6.

Jammalamadaka, S., Taufer, E., & Terdik, G. H. (2021). On higher order cumulants of multivariate symmetric and skew symmetric distributions. *Technical Report, ResearchGate*(pp. 1–20).

Klar, B. (2002). A treatment of multivariate skewness, kurtosis, and related statistics. *Journal of Multivariate Analysis*, *83*(1), 141–165.

Kollo, T. (2008). Multivariate skewness and kurtosis measures with an application in ICA. *Journal of Multivariate Analysis*, *99*(10), 2328–2338.

Koziol, J. A. (1987). An alternative formulation of Neyman's smooth goodness of fit tests under composite alternatives. *Metrika*, *34*(1), 17–24.

Koziol, J. A. (1989). A note on measures of multivariate kurtosis. *Biometrical Journal*, *31*(5), 619–624.

León, A., & Moreno, M. (2017). One-sided performance measures under gram-Charlier distributions. *Journal of Banking & Finance*, *74*, 38–50.

Lin, S. H., Huang, H. H., & Li, S. H. (2015). Option pricing under truncated gram–Charlier expansion. *The North American Journal of Economics and Finance*, *32*, 77–97.

Magnus, J. R., & Neudecker, H. (1999). *Matrix differential calculus with applications in statistics and econometrics*. John Wiley & Sons Ltd Revised reprint of the 1988 original.

Malkovich, J. F., & Afifi, A. A. (1973). On tests for multivariate normality. *Journal of the American Statistical Association*, *68*(341), 176–179.

Mardia, K. V. (1970). Measures of multivariate skewness and kurtosis with applications. *Biometrika*, *57*, 519–530.

Meijer, E. (2005). Matrix algebra for higher order moments. *Linear Algebra and Its Applications*, *410*, 112–134.

Móri, T. F., Rohatgi, V. K., & Székely, G. J. (1994). On multivariate skewness and kurtosis. *Theory of Probability & Its Applications*, *38*(3), 547–551.

Nordhausen, K., Oja, H., Tyler, D. E., & Virta, J. (2017). Asymptotic and bootstrap tests for the dimension of the non-Gaussian subspace. *IEEE Signal Processing Letters*, *24*(6), 887–891.

Oja, H., Sirkiä, S., & Eriksson, J. (2006). Scatter matrices and independent component analysis. *Austrian Journal of Statistics*, *35*(2&3), 175–189.

Peña, D., Prieto, F. J., & Viladomat, J. (2010). Eigenvectors of a kurtosis matrix as interesting directions to reveal cluster structure. *Journal of Multivariate Analysis*, *101*(9), 1995–2007.

Serfling, R. J. (2009). *Approximation theorems of mathematical statistics* (Vol. *162*). John Wiley & Sons.

Srivastava, M. S. (1984). A measure of skewness and kurtosis and a graphical method for assessing multivariate normality. *Statistics & Probability Letters*, *2*(5), 263–267.

Tanaka, K., Yamada, T., & Watanabe, T. (2010). Applications of Gram–Charlier expansion and bond moments for pricing of interest rates and credit risk. *Quantitative Finance*, *10*(6), 645–662.

Tyler, D. E., Critchley, F., Dümbgen, L., & Oja, H. (2009). Invariant co-ordinate selection. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, *71*(3), 549–592.

Weisberg, S. (2002). Dimension reduction regression in R. *Journal of Statistical Software*, *7*(1), 1–22.

How to cite this article: Rao Jammalamadaka S, Taufer E, Terdik GH. Asymptotic theory for statistics based on cumulant vectors with applications. *Scand J Statist*. 2021;48:708–728. https://doi.org/10.1111/sjos.12521

APPENDIX A. MOMENTS AND CUMULANTS FOR SYMMETRIC MULTIVARIATE DISTRIBUTIONS

Recall the definition of a spherical random vector $\underline{W} = (W_1, \dots, W_d)^\top$ given in Section 2.2.1. We now provide general results for the cumulants of such a $\underline{W}$. In what follows, we use the multifactorial !! notation, which stands for the product of integers in steps of two. Recall that the moments of the components of $\underline{W}$, when they exist, can be expressed in terms of a one-dimensional integral (Fang et al., 2017, theorem 2.8, p. 34), and the characteristic function has the form

$$\phi_{\underline{W}}\left(\underline{\lambda}\right) = g\left(\underline{\lambda}^{\intercal}\underline{\lambda}\right) = \sum_{j=1}^{\infty} \underline{\mu}_j^{\intercal}\frac{i^j}{j!}\underline{\lambda}^{\otimes j}, \tag{A1}$$

where g is called the *characteristic generator*, and R is the generating variate with a generating distribution F (say). The coefficients in the expansion are the moments $\underline{\mu}_j = (-i)^j D_{\underline{\lambda}}^{\otimes j}\phi_{\underline{W}}\left(\underline{\lambda}\right)\Big|_{\underline{\lambda}=0}$.

Marginal moments

Moments of univariate elliptically symmetric distributions have been discussed by Berkane and Bentler (1986). Let us denote $\log(g) = f$, and define $\nu_k = g^{(k)}(0)$; in general,

$$EW_j^n = \begin{cases} 0 & \text{if} \quad n \quad \text{odd}, \\ (-1)^{\ell} 2^{\ell}(2\ell-1)!!\nu_{\ell} & \text{if} \quad n = 2\ell \quad \text{even}. \end{cases}$$

Notice that the right hand side above does not depend on j, and then all marginals are identically distributed. The odd cumulants are also zero. Note that ν_k **is not** the moment of R, the relationship between the distribution F of R and g is given through the characteristic function of the uniform distribution on the sphere (see Fang et al., 2017 , p. 30).

Multivariate moments and cumulants

Now, the characteristic generator g is a function of one variable with the series expansion

$$g(u) = \sum_{j=1}^{\infty} g_j \frac{(-1)^j}{j!}u^j,$$

such that $g_j = (-1)^j g^{(j)}(0) = (-1)^j \nu_j$. Rewrite $\phi_{\underline{W}}\left(\underline{\lambda}\right) = g\left(\underline{\lambda}^{\intercal}\underline{\lambda}\right) = \sum_{j=1}^{\infty} g_j \frac{(-1)^j}{j!}\left(\underline{\lambda}^{\intercal}\underline{\lambda}\right)^j$, and calculate the moments by

$$\underline{\mu}_k = (-i)^k D_{\underline{\lambda}}^{\otimes k}\phi_{\underline{W}}\left(\underline{\lambda}\right)\Big|_{\underline{\lambda}=0} = (-i)^k \sum_{j=1}^{\infty} g_j \frac{(-1)^j}{j!} D_{\underline{\lambda}}^{\otimes k}\left(\underline{\lambda}^{\intercal}\underline{\lambda}\right)^j\Bigg|_{\underline{\lambda}=0} = \begin{cases} 0 & \text{if} \quad k \neq 2j, \\ \frac{1}{j!}g_j\underline{c}_j & \text{if} \quad 2j = k, \end{cases}$$

and the vector $\underline{c}_j$ does not depend on g. Use $g^{(j)}(0) = i^j g_j$, hence $\underline{c}_j = (-1)^j D_{\underline{\lambda}}^{\otimes 2j}\left(\underline{\lambda}^{\intercal}\underline{\lambda}\right)^j$, and conclude that

$$E\underline{W}^{\otimes 2\ell} = \frac{(-1)^{\ell}}{\ell!} g^{(\ell)}(0) D_{\underline{\lambda}}^{\otimes 2\ell}\left(\underline{\lambda}^{\intercal}\underline{\lambda}\right)^{\ell} = \frac{EW_j^{2\ell}}{\ell! 2^{\ell}(2\ell-1)!!} D_{\underline{\lambda}}^{\otimes 2\ell}\left(\underline{\lambda}^{\intercal}\underline{\lambda}\right)^{\ell}.$$

We apply the same argument for the cumulant generator $\psi_{\underline{W}}\left(\underline{\lambda}\right) = \log\phi_{\underline{W}}\left(\underline{\lambda}\right)$ and get

$$\psi_{\underline{W}}\left(\underline{\lambda}\right) = \sum_{j=1}^{\infty} \underline{\kappa}_j^{\intercal}\frac{i^j}{j!}\underline{\lambda}^{\otimes j},$$

with $\underline{\operatorname{Cum}}_n\left(\underline{W}\right) = (-i)^n D_{\underline{\lambda}}^{\otimes n}\psi_{\underline{W}}\left(\underline{\lambda}\right)\Big|_{\underline{\lambda}=0}$. We can conclude directly from the series expansions of cumulant functions, that odd orders are all zero and

$$\underline{\mathrm{Cum}}_{2\ell}\left(\underline{W}\right) = \frac{(-1)^{\ell}}{\ell!}\rho_\ell D_{\underline{\lambda}}^{\otimes 2\ell}\left(\underline{\lambda}^{\intercal}\underline{\lambda}\right)^{\ell}.$$

Since $\rho_\ell = \left(\log(g)\right)^{(\ell)}(0)$ is connected to the ℓ^{th} cumulant of a component of $\underline{W}$, we write

$$\underline{\mathrm{Cum}}_{2\ell}\left(\underline{W}\right) = \frac{\mathrm{Cum}_{2\ell}(W_1)}{2^{\ell}\ell!(2\ell-1)!!}D_{\underline{\lambda}}^{\otimes 2\ell}\left(\underline{\lambda}^{\intercal}\underline{\lambda}\right)^{\ell}.$$

The above cumulant can be calculated if either the generator function is known, or by using the stochastic representation $\underline{W} = R\underline{U}$, of $\underline{W}$ where R is a non-negative random variable, and $\underline{U}$ is uniform on sphere $\mathbb{S}_{d-1}$. Furthermore, R and $\underline{U}$ are *independent*. Let U_j be a component of $\underline{U}$, then

$$\mathrm{Cum}_4(W_1) = \mathrm{E}R^4\mathrm{E}U_j^4 - 3\left(\mathrm{E}R^2\mathrm{E}U_j^2\right)^2,$$
$$\mathrm{Cum}_6(W_1) = \mathrm{E}R^6\mathrm{E}U_j^6 - 15\mathrm{E}R^4\mathrm{E}R^2\mathrm{E}U_j^4\mathrm{E}U_j^2 + 30\left(\mathrm{E}R^2\mathrm{E}U_j^2\right)^3.$$

These formulae are very useful in applications and simulations where, given a univariate random variable R, we can arrive at a complete specification of the cumulants of $\underline{W}$.

Theorem 3. *Let $\underline{W}$ have a spherically symmetric distribution. We then have*

$$\begin{aligned}
\mathrm{E}W_1^{2\ell} &= (-1)^{\ell}2^{\ell}(2\ell-1)!!\nu_\ell,\\
\mathrm{E}\underline{W}^{\otimes 2\ell} &= \mathrm{E}W_1^{2\ell}\mathbf{S}_{d\mathbf{1}_{2\ell}}(\mathrm{Vec}\,\mathbf{I}_d)^{\otimes \ell},\\
\mathrm{Cum}_{2\ell}\left(W_j\right) &= (-1)^{\ell}2^{\ell}(2\ell-1)!!\rho_\ell,\\
\underline{\mathrm{Cum}}_{2\ell}\left(\underline{W}\right) &= \mathrm{Cum}_{2\ell}(W_1)\mathbf{S}_{d\mathbf{1}_{2\ell}}(\mathrm{Vec}\,\mathbf{I}_d)^{\otimes \ell}.
\end{aligned}$$

Proof. For a full proof of the above theorem, see Jammalamadaka et al. (2021). ■

APPENDIX B. SOME TECHNICAL BACKGROUND

In what follows, we use the notation $1:d$ to denote $1, 2, \ldots, d$.

From cumulants to moments

Moments can be expressed in terms of cumulants via the formula

$$\mathrm{E}\underline{X}^{\otimes k} = \sum_{\mathcal{L}\in\mathcal{P}_{(1:k)}}\mathbf{K}_{\mathfrak{p}(\mathcal{L})}^{-1}\prod_{\mathbf{b}_j\in\mathcal{L}}^{\otimes}\underline{\mathrm{Cum}}_{|\mathbf{b}_j|}(\underline{X}) = \underline{B}_k\left(\underline{\kappa}_{\underline{X},1}, \underline{\kappa}_{\underline{X},2}, \ldots \underline{\kappa}_{\underline{X},k}\right), \tag{B1}$$

where the summation is over all partitions $\mathcal{L} = \{\mathbf{b}_1, \mathbf{b}_2, \ldots, \mathbf{b}_k\}$ of $1:n$, and $\underline{B}_k$ are multivariate Bell polynomials.

Commutators and symmetrizers

Some preliminary discussion of commutation matrices is helpful. Let $\mathbf{E}_{i,j}$ denote the $(d\times d)$ elementary matrix, i.e. one for which all entries are zero except for the (i,j)th element, which is 1. Set $\mathbf{K}_{d\bullet d} = \left[\mathrm{Vec}\mathrm{E}_{i,j}^{\intercal}\right]$, so that $\mathbf{K}_{d\bullet d}$ has dimension $d^2 \times d^2$. The vector $\mathrm{Vec}\mathrm{E}_{i,j}^{\intercal} = \mathrm{Vec}\,\mathrm{E}_{j,i}$, is the $((i-1)\,d+j)$th, $i=1:d$, $j=1:d$, unit vector of the unit matrix $\mathbf{I}_{d^2}$, and $\mathbf{K}_{d\bullet d}\left(\underline{a}_1 \otimes \underline{a}_2\right) = \underline{a}_2 \otimes \underline{a}_1$. The matrix $\mathbf{K}_{d\bullet d}$ is called a commutation matrix (see e.g., Graham, 2018; Magnus & Neudecker, 1999). By changing the neighboring elements of a Kronecker product, we can obtain any permutation of them. For instance, if $\mathfrak{p} = (i_1, i_2, i_3, i_4)$ is a permutation of the numbers 1,2,3,4, we can introduce the commutator matrix $\mathbf{K}_{\mathfrak{p}}$ for changing the order of a Kronecker product, namely $\mathbf{K}_{\mathfrak{p}}\left(\underline{a}_1 \otimes \underline{a}_2 \otimes \underline{a}_3 \otimes \underline{a}_4\right) = \underline{a}_{i_1} \otimes \underline{a}_{i_2} \otimes \underline{a}_{i_3} \otimes \underline{a}_{i_4}$. In particular, if we set $\mathfrak{p}_1 = (1,3,2,4)$, then

$$\mathbf{K}_{\mathfrak{p}_1}\left(\underline{a}_1 \otimes \underline{a}_2 \otimes \underline{a}_3 \otimes \underline{a}_4\right) = \left(\mathbf{I}_d \otimes \mathbf{K}_{d\bullet d} \otimes \mathbf{I}_d\right)\left(\underline{a}_1 \otimes \underline{a}_2 \otimes \underline{a}_3 \otimes \underline{a}_4\right) = \underline{a}_1 \otimes \underline{a}_3 \otimes \underline{a}_2 \otimes \underline{a}_4.$$

It is worth noting that $\mathbf{K}_{d\bullet d}$, and in general each commutator matrix, depends on the dimensions of the vectors under consideration. In our example, the dimension of $\mathbf{K}_{d\bullet d}$ is $d^2 \times d^2$, while the dimension of $\mathbf{K}_{\mathfrak{p}_1}$ is $d^4 \times d^4$. As far as the commutators used in Section 2 are concerned, the formulae are:

$$\mathbf{K}_{2,2} = \mathbf{K}_{(3,4,1,2)}^{-1} + \mathbf{K}_{(2,4,1,3)}^{-1} + \mathbf{K}_{(2,3,1,4)}^{-1},$$
$$\mathbf{K}_{H2,2,2}^{-1} = \sum_{j=1:3,k=4:6} \mathbf{K}_{(j,k,p_{2!}((1:6)\setminus(j,k)))}^{-1}, \qquad \mathbf{K}_{H4,2}^{-1} = \sum_{j=1:3,k=4:6} \mathbf{K}_{((1:6)\setminus(j,k),j,k)}^{-1}.$$

For instance, $p_{2!}\left((1:6)\setminus(j,k)\right)|_{j=1,k=5} = (2,4,3,6) + (2,6,3,4)$,

$$\mathbf{K}_{(j,k,p_{2!}((1:6)\setminus(j,k)))}^{-1}|_{j=1,k=5} = \mathbf{K}_{(1,5,2,3,4,6)}^{-1} + \mathbf{K}_{(1,5,2,6,3,4)}^{-1},$$
$$\mathbf{K}_{3!}^{-1} = \mathbf{K}_{(1,4,2,5,3,6)}^{-1} + \mathbf{K}_{(1,4,2,6,3,5)}^{-1} + \mathbf{K}_{(1,5,2,4,3,6)}^{-1} + \mathbf{K}_{(1,5,2,6,3,4)}^{-1} + \mathbf{K}_{(1,6,2,4,3,5)}^{-1} + \mathbf{K}_{(1,6,2,5,3,4)}^{-1}.$$

$\mathbf{K}_{3!}$ provides a specific permutation needed since $\mathrm{E}_{\varphi}\underline{H}_3\left(\underline{y}\right)^{\otimes 2}$ is not symmetric. We remark that the results in Holmquist (1996) are valid for symmetric vectors. next

$$\mathbf{K}_{H6,2}^{-1} = \sum_{j=1:4,\ k=5:8} \mathbf{K}_{([(1:4)\setminus j],[(5:8)\setminus k],j,k)}^{-1},$$
$$\mathbf{K}_{H4,2,2}^{-1} = \sum_{\substack{j_1,j_2 = 1:4,\ k_1,k_2 = 5:8 \\ j_2>j_1 \quad k_2>k_1}} \mathbf{K}_{(j_1,j_2,k_1,k_2,p_{2!}((1:8)\setminus(j_1,j_2,k_1,k_2)))}^{-1}.$$

Also, $\mathbf{K}_{4!}^{-1} = \sum_{4!}\mathbf{K}_{(1,k_1,2,k_2,3,k_3,4,k_4)}^{-1}$, with the sum taken over all permutations (k_1, k_2, k_3, k_4) of the numbers $(5:8)$. Algorithms for computing $\mathbf{K}_{2,2}$ $\mathbf{K}_{3!}$, $\mathbf{K}_{H\,2,2,2}$, $\mathbf{K}_{H\,4,2}$, $\mathbf{K}_{4!}$, $\mathbf{K}_{H\,4,2,2}$, and $\mathbf{K}_{H\,6,2}$ are available.

Holmquist (1996) uses the symmetrizer matrix $\mathbf{S}_{d\mathbf{1}_q}$ for symmetrization of a T-product of q vectors with the same dimension d. That is, $\mathbf{S}_{d\mathbf{1}_4}\left(\underline{a}_1 \otimes \underline{a}_2 \otimes \underline{a}_3 \otimes \underline{a}_4\right)$ is a vector of dimension d^4, which is symmetric in $\underline{a}_j$. It can be computed as $\mathbf{S}_{d\mathbf{1}_q} = \frac{1}{q!}\sum_{p\in P_q}\mathbf{K}_p$, where P_q denotes the set of all permutations of the numbers $1:q$; the sum includes $q!$ terms. The symmetrizer $\mathbf{S}_{d\mathbf{1}_q}$ provides an orthogonal projection to the subspace of $\mathbb{R}^{d^q}$ which is invariant under the transformation $\mathbf{S}_{d\mathbf{1}_q}$. A vector will be called symmetrical if it belongs to that subspace.

APPENDIX C. PROOFS

Proof of Theorem 1. Consider first $\underline{H}_3(\underline{Y})$ and note that $\underline{\mathrm{Cum}}_2\left(\underline{H}_3\left(\underline{Y}\right)\right) = \mathrm{E}\underline{H}_3\left(\underline{Y}\right)^{\otimes 2} - \underline{\kappa}_{Y,3}^{\otimes 2}$, where

$$\underline{H}_3(\underline{Y})^{\otimes 2} = \underline{H}_6\left(\underline{Y}\right) + \mathbf{K}^{-1}_{H4,2}\left(\underline{H}_4\left(\underline{Y}\right) \otimes \underline{\kappa}_{Y,2}\right) + \mathbf{K}^{-1}_{H2,2,2}\left(\underline{H}_2\left(\underline{Y}\right) \otimes \underline{\kappa}_{Y,2}^{\otimes 2}\right) + \mathbf{K}^{-1}_{3!}\underline{\kappa}_{Y,2}^{\otimes 3}.$$

Noting that $\mathrm{E}\underline{H}_2\left(\underline{Y}\right) = 0$, it follows that

$$\mathrm{E}\underline{H}_3\left(\underline{Y}\right)^{\otimes 2} = \mathrm{E}\underline{H}_6\left(\underline{Y}\right) + \mathbf{K}^{-1}_{H4,2}\left(\mathrm{E}\underline{H}_4\left(\underline{Y}\right) \otimes \underline{\kappa}_{Y,2}\right) + \mathbf{K}^{-1}_{3!}\underline{\kappa}_{Y,2}^{\otimes 3},$$

from which result (4) follows.

In the case of $\underline{H}_4(\underline{Y})$ we have $\underline{\mathrm{Cum}}_2\left(\underline{H}_4\left(\underline{Y}\right)\right) = \mathrm{E}\underline{H}_4\left(\underline{Y}\right)^{\otimes 2} - \underline{\kappa}_{Y,4}^{\otimes 2}$, where

$$\underline{H}_4(\underline{Y})^{\otimes 2} = \underline{H}_8\left(\underline{Y}\right) + \mathbf{K}^{-1}_{H6,2}\left(\underline{H}_6\left(\underline{Y}\right) \otimes \underline{\kappa}_{Y,2}\right) + \mathbf{K}^{-1}_{H4,2,2}\left(\underline{H}_4\left(\underline{Y}\right) \otimes \underline{\kappa}_{Y,2}^{\otimes 2}\right)$$
$$+ \mathbf{K}^{-1}_{H2,2,2}\left(\underline{H}_2\left(\underline{Y}\right) \otimes \underline{\kappa}_{Y,2}^{\otimes 3}\right) + \mathbf{K}^{-1}_{4!}\underline{\kappa}_{Y,2}^{\otimes 4}.$$

Taking expectations and recalling that $\mathrm{E}\underline{H}_2\left(\underline{Y}\right) = 0$, we have

$$\mathrm{E}\underline{H}_4\left(\underline{Y}\right)^{\otimes 2} = \mathrm{E}\underline{H}_8\left(\underline{Y}\right) + \mathbf{K}^{-1}_{H6,2}\left(\mathrm{E}\underline{H}_6\left(\underline{Y}\right) \otimes \underline{\kappa}_{Y,2}\right) + \mathbf{K}^{-1}_{H4,2,2}\left(\mathrm{E}\underline{H}_4\left(\underline{Y}\right) \otimes \underline{\kappa}_{Y,2}^{\otimes 2}\right) + \mathbf{K}^{-1}_{4!}\underline{\kappa}_{Y,2}^{\otimes 4}.$$

By the GC expansion, where the expected values are expressed in terms of Bell polynomials, we finally get expression (C). ∎

Proof of Theorem 2. Details are provided for the case $q = 4$ which requires some more computations; the proof for the case $q = 3$ is similar. In the proof, when the use of the symmetrizer $\mathbf{S}_{\mathbf{d1}_4}$ is needed, we denote this fact as $\overset{\circledS}{=}$. Note that

$$\underline{H}_4\left(\underline{Z}\right) = \underline{H}_4\left(\widehat{\boldsymbol{\Sigma}}^{-1/2}\left(\underline{X} - \overline{\underline{X}}\right)\right) = \left(\widehat{\boldsymbol{\Sigma}}^{-1/2}\right)^{\otimes 4}\underline{H}_4\left(\underline{X} - \overline{\underline{X}}\right).$$

Since $\widehat{\boldsymbol{\Sigma}} \to \boldsymbol{\Sigma}$ in probability, by Slutsky theorem, we only need to consider the distribution of $\underline{H}_4\left(\underline{X} - \overline{\underline{X}}\right)$. Since the kurtosis is affine invariant, we assume $\underline{\mu} = 0$ and $\boldsymbol{\Sigma} = \mathbf{I}_d$. It follows that $\underline{X}$ corresponds to its standardized version $\underline{Y}$ (see formula (2)). Then

$$\sqrt{n}\overline{\underline{H}_4\left(\underline{X} - \overline{\underline{X}}\right)} \overset{\circledS}{=} \sqrt{n}\overline{\left(\underline{X} - \overline{\underline{X}}\right)^{\otimes 4}} - 6\left(\mathrm{Vec}\,\mathbf{I}_d\right) \otimes \overline{\left(\underline{X} - \overline{\underline{X}}\right)^{\otimes 2}} + 3\left(\mathrm{Vec}^{\otimes 2}\,\mathbf{I}_d\right)$$
$$\overset{\circledS}{=} \sqrt{n}\,\overline{\underline{H}_4\left(\underline{X}\right)} - 4\overline{\underline{H}_3\left(\underline{X}\right)} \otimes \sqrt{n}\,\overline{\underline{X}} + 6\sqrt{n}\,\overline{\underline{X}^{\otimes 2}} \otimes \overline{\underline{X}}^{\otimes 2} - 4\sqrt{n}\,\overline{\underline{X}_i} \otimes \overline{\underline{X}}^{\otimes 3}$$
$$+ \sqrt{n}\,\overline{\underline{X}}^{\otimes 4} - 6\mathrm{Vec}\,\mathbf{I}_d \otimes \overline{\underline{X}}^{\otimes 2} + 3\left(\mathrm{Vec}^{\otimes 2}\,\mathbf{I}_d\right)$$
$$\overset{D}{\simeq} \sqrt{n}\,\overline{\underline{H}_4\left(\underline{X}\right)} - 4\underline{\kappa}_{X,3} \otimes \sqrt{n}\,\overline{\underline{X}},$$

since, using Slutsky's argument, when $n \to \infty$, we obtain that $\sqrt{n}\underline{X}$ is asymptotically multivariate standard normal, $\overline{\underline{X}}^{\otimes k} = o_p(1)$ for $k \geq 1$ and $\overline{\underline{H}_3(\underline{X})} \to \mathrm{E}\underline{H}_3(\underline{X}) = \underline{\kappa}_{\underline{X},3}$. The variance

$$\underline{\mathrm{Cum}}_2\left(\underline{H}_4(\underline{X}) - 4\underline{\kappa}_{\underline{X},3} \otimes \underline{H}_1(\underline{X})\right) = \underline{\mathrm{Cum}}_2\left(\underline{H}_4(\underline{X})\right) + 16\underline{\kappa}_{\underline{X},3}^{\otimes 2} \otimes \mathrm{Vec}\,\mathbf{I}_d$$

where $\underline{\mathrm{Cum}}_2\left(\underline{H}_4(\underline{X})\right)$ is given in (5). ∎