Ann Inst Stat Math (2019) 71:137–162
https://doi.org/10.1007/s10463-017-0630-0

A class of uniform tests for goodness-of-fit of the multivariate L_p-norm spherical distributions and the l_p-norm symmetric distributions

Jiajuan Liang[1] · **Kai Wang Ng**[2] · **Guoliang Tian**[3]

Received: 20 October 2012 / Revised: 14 July 2017 / Published online: 1 December 2017

Abstract In this paper we employ the conditional probability integral transformation (CPIT) method to transform a d-dimensional sample from two classes of generalized multivariate distributions into a uniform sample in the unit interval (0, 1) or in the unit hypercube $[0,\ 1]^{d-1}$ ($d \geq 2$). A class of existing uniform statistics are adopted to test the uniformity of the transformed sample. Monte Carlo studies are carried out to demonstrate the performance of the tests in controlling type I error rates and power against a selected group of alternative distributions. It is concluded that the proposed tests have satisfactory empirical performance and the CPIT method in this paper can serve as a general way to construct goodness-of-fit tests for many generalized multivariate distributions.

Keywords Goodness-of-fit · Monte Carlo study · L_p-norm spherical distribution · l_p-norm symmetric distribution · Uniformity

The research was supported in part by the University of Hong Kong Research Grant and University of New Haven Research Scholar Grant.

✉ Jiajuan Liang
jliang@newhaven.edu

1 College of Business, University of New Haven, 300 Boston Post Road, West Haven, CT 06516, USA

2 Department of Statistics and Actuarial Science, The University of Hong Kong, Pokfulam Road, Hong Kong, China

3 Department of Mathematics, Southern University of Science and Technology, 1088 Xueyuan Road, Nanshan District, Shenzhen 518055, China

1 Introduction

Methodologies for multivariate statistical inference under the normal distribution have been comprehensively studied, and various approaches to testing goodness-of-fit for the multivariate normal distribution are available in the literature. Because various violations of the normal assumption were noticed in modern high-dimensional data analysis, researchers have been making efforts in developing theory for statistical inference under a much wider class of multivariate distributions that possess similar properties to those of the multivariate normal distribution. The family of the elliptically contoured distributions (ECD for simplicity) is a natural extension to the family of multivariate normal distributions (MVN for simplicity), see, for example, Fang and Zhang (1990), and Gupta and Varga (1993). Some kind of symmetry is a special characteristic that all ECDs commonly possess. For example, if a random vector $\boldsymbol{x}$ has an ECD, then its linear transformation $\boldsymbol{y} = \boldsymbol{A}x + \boldsymbol{b}$ ($\boldsymbol{A}$ is a matrix of suitable dimension and $\boldsymbol{b}$ a constant vector of suitable dimension) also has an ECD; similarly, if a random vector $\boldsymbol{z}$ has a spherically symmetric distribution (SSD for simplicity), then $\boldsymbol{\Gamma z}$ also has an SSD for any orthogonal constant matrix $\boldsymbol{\Gamma}$. An SSD is a standardized ECD with zero mean and a covariance matrix of the form $\sigma^2\boldsymbol{I}$ ($\boldsymbol{I}$ stands for an identity matrix). The symmetry of ECD or MVN is called affine invariance, and the symmetry of SSD is called rotational invariance. Invariance is a common approach to developing a family of multivariate distributions. Fang et al. (1990) summarizes various approaches to generalizing the multivariate normal distribution to different families of symmetric multivariate distributions.

Among the various symmetric multivariate distributions, the family of ECD possesses the closest properties to those of MVN, and the family of SSD possesses the closest properties to those of the standard MVN. Statistical inference under SSD or ECD received earlier attention compared with other symmetric multivariate distributions. For example, Zellner (1976) and Lange et al. (1989) studied statistical models under SSD; Anderson et al. (1986), Fang and Zhang (1990), and Anderson (1993) developed some theory on statistical inference under ECD; and Osiewalski and Steel (1993) proposed robust Bayesian inference for the l_q-spherical models. Some methods for statistical inference under SSD with non-independent samples were proposed by Kariya and Eaton (1977) and Gupta and Kabe (1993). Some early methods for testing spherical and elliptical symmetry were summarized in Fang and Liang (1999). A lot of new methods for statistical inference under SSD or ECD have been developed since the past ten years, see, for example, Liang and Fang (2000), Manzottia et al. (2002), Schott (2002), Zhu (2003), Huffer and Park (2007), and Liang et al. (2008).

Compared to the relatively rich literature on ECD and SSD, there are few methodologies for statistical inference under a wide class of symmetric multivariate distributions. The class of L_p-norm spherical distributions and the class of l_p-norm symmetric distributions are two classes of symmetric multivariate distributions that have been receiving attention in the area of generalized multivariate analysis (see, e.g., Osiewalski and Steel 1993).

Definition 1 (Gupta and Song 1997). A random vector $\boldsymbol{u}_d = (U_1, \ldots, U_d)'$ is said to have an L_p-norm uniform distribution, denoted by $\mathcal{U}(d, p)$, if $\|\boldsymbol{u}_d\|_p =$

$(\sum_{i=1}^d |U_i|^p)^{1/p} = 1$ ($p > 0$) and the joint p.d.f. (probability density function) of $U_1, \ldots, U_{d-1}$ is given by

$$f(u_1, \ldots, u_{d-1}) = \frac{p^{d-1}\Gamma(d/p)}{2^{d-1}\Gamma^d(1/p)} \left(1 - \sum_{i=1}^{d-1} |u_i|^p\right)^{(1-p)/p},$$

$$-1 < u_i < 1,\ i = 1, \ldots, d-1,\ \sum_{i=1}^{d-1} |u_i|^p < 1. \quad (1)$$

Definition 2 (Gupta and Song 1997). A d-variate random vector $\boldsymbol{x}$ is said to have an L_p-norm spherical distribution if $\boldsymbol{x} = R\boldsymbol{u}_d$, denoted by $\boldsymbol{x} \sim \mathcal{S}(d, p)$, where $\boldsymbol{u}_d \sim \mathcal{U}(d, p)$ and R, which is independent of $\boldsymbol{u}_d$, is a nonnegative random variable.

Definition 3 (Yue and Ma 1995). Suppose that $W_1, \ldots, W_d$ are i.i.d. Weibull distributed with a fixed shape parameter p and density $g(w) = pw^{p-1}\exp(-w^p)$ ($w > 0$). Let $\boldsymbol{w} = (W_1, \ldots, W_d)'$ ($W_i > 0, i = 1, \ldots, d$) and $\boldsymbol{u} \stackrel{d}{=} (U_1, \ldots, U_d)' = \boldsymbol{w}/\|\boldsymbol{w}\|_p^+$, where $\|\boldsymbol{w}\|_p^+ = (\sum_{i=1}^d W_i^p)^{1/p}$ and "$\stackrel{d}{=}$" means that the two sides of the random variables (r.v.) have the same probability distribution, denote the distribution of $\boldsymbol{u}$ by $\boldsymbol{u} \sim UW_d(p)$. Define a class of distributions by

$$\mathcal{L}_d(p) = \{\boldsymbol{x} : \boldsymbol{x} \stackrel{d}{=} r\boldsymbol{u},\ r \geq 0 \text{ and is a r.v. independent of } \boldsymbol{u} \sim UW_d(p)\}, \quad (2)$$

$\boldsymbol{x}$ is said to have a multivariate l_p-norm symmetric distribution if $\boldsymbol{x} \in \mathcal{L}_d(p)$.

From the above definitions, it can be easily understood that the family of $\mathcal{S}(d, p)$ contains SSD as a special case ($p = 2$) and the family of $\mathcal{L}_d(p)$ contains as its special cases ($p = 1$) the family of the multivariate l_1-norm symmetric distributions (Fang et al. 1990), and the family of mixtures of Weibull distributions. Some theoretical properties of the distributions in $\mathcal{S}(d, p)$ and the distributions in $\mathcal{L}_d(p)$ can be referred to Gupta and Song (1997) and Yue and Ma (1995), respectively.

Similar to testing goodness-of-fit for ECD or SSD, testing goodness-of-fit for $\mathcal{S}(d, p)$ and $\mathcal{L}_d(p)$ is even more challenging because much fewer characterization properties for distributions in $\mathcal{S}(d, p)$ or $\mathcal{L}_d(p)$ are known. The purpose of this paper is to construct a class of uniform tests for uniformity of two classes of generalized multivariate uniform distributions and then apply the uniform tests to testing goodness-of-fit for distributions in either $\mathcal{S}(d, p)$ or $\mathcal{L}_d(p)$. This paper is organized as follows. Section 2 is devoted to the theoretical construction of the tests. Section 3 presents Monte Carlo studies on the performance of the tests. Some concluding remarks are summarized in the last section.

2 Theoretical construction of the uniform tests

In this section we will construct two types of goodness-of-fit tests:

Type I test Testing goodness-of-fit for $\mathcal{U}(d, p)$
Let $\{\boldsymbol{x}_1, \ldots, \boldsymbol{x}_n\}$ be an i.i.d. sample, $\boldsymbol{x}_i \in R^d$ (the d-dimensional Euclidean space) for $i = 1, \ldots, n$. We want to test the null hypothesis

$$H_0 : \{\boldsymbol{x}_1, \ldots, \boldsymbol{x}_n\} \text{ is a sample from } \mathcal{U}(d, p) \text{ for some given } p > 0 \tag{3}$$

versus the alternative hypothesis H_1: the sample $\{\boldsymbol{x}_1, \ldots, \boldsymbol{x}_n\}$ is not from any c.d.f. in $\mathcal{U}(d, p)$, where c.d.f. = cumulative distribution function.

Type II test Testing goodness-of-fit for $UW_d(p)$
Let $\{\boldsymbol{y}_1, \ldots, \boldsymbol{y}_n\}$ be an i.i.d. sample, $\boldsymbol{y}_i \in R_+^d = \{\boldsymbol{y} = (y_1, \ldots, y_d)' : \ y_l > 0, \ l = 1, \ldots, d\}$. We want to test the null hypothesis

$$H_0 : \ \{\boldsymbol{y}_1, \ldots, \boldsymbol{y}_n\} \text{ is a sample from } UW_d(p) \text{ for some given } p > 0 \tag{4}$$

versus the alternative hypothesis H_1: the sample $\{\boldsymbol{y}_1, \ldots, \boldsymbol{y}_n\}$ is not from any c.d.f. in $UW_d(p)$.

Testing hypotheses (3) and (4) will be transferred to testing uniformity in the unit interval $(0, 1)$ or to testing uniformity in the unit hypercube $[0, 1]^{d-1}$. Liang et al. (2008) gave a review on some uniform tests as follows.

2.1 Tests for uniformity in (0, 1)

1. Watson's U^2-statistic
Let $W^2 = 1/(12n) + \sum_{i=1}^n [(2i - 1)/2n - U_{(i)}]^2$, Watson (1962) proposed the statistic

$$WU^2 = W^2 - n(\bar{U} - 0.5)^2 \tag{5}$$

for testing uniformity in $(0, 1)$, where $\bar{U}$ is the sample mean from an i.i.d. sample $\{U_1, \ldots, U_n\}$ and $U_{(1)} \leq \cdots \leq U_{(n)}$ is ordered array. Tables of critical values for WU^2 are usually given for the modified form of WU^2:

$$MU^2 = \left(WU^2 - \frac{1}{10n} + \frac{1}{10n^2}\right)\left(1 + \frac{0.8}{n}\right). \tag{6}$$

The critical values of MU^2 are found to be only slightly dependent on the sample size n, and they are 0.267 ($\alpha = 0.01$), 0.187 ($\alpha = 0.05$) and 0.152 ($\alpha = 0.10$) from Stephens (1970). Large values of MU^2 indicate evidence of non-uniformity of the sample. For example, if $MU^2 > 0.187$, one rejects the null hypothesis of uniformity in $(0, 1)$ at the significance level $\alpha = 0.05$.

2. Neyman's smooth test

Let

$$\pi_0(y) = 1, \quad \pi_3(y) = \sqrt{7}[20(y-1/2)^3 - 3(y-1/2)],$$
$$\pi_1(y) = \sqrt{12}(y-1/2), \quad \pi_4(y) = 210(y-1/2)^4 - 45(y-1/2)^2 + 9/8,$$
$$\pi_2(y) = \sqrt{5}[6(y-1/2)^2 - 1/2],$$

which are Legendre polynomials, $y \in [0, 1]$. Denote by

$$t_r = \sum_{i=1}^{n} \pi_r(U_i), \quad r = 1, 2, 3, 4, \tag{7}$$

where $\{U_1, \ldots, U_n\}$ is an i.i.d. sample in $(0, 1)$. Neyman's smooth test (Neyman 1937) with the 4th-degree polynomials is defined by

$$P_4^2 = \frac{1}{n} \sum_{r=1}^{4} t_r^2. \tag{8}$$

Large values of P_4^2 indicate evidence of non-uniformity of the sample. Critical values for P_4^2 for some small sample size n and for large n ($n \to \infty$) were provided by Miller and Quesenberry (1979). For example, for $n > 50$, the critical values for P_4^2 were given as 13.28 ($\alpha = 0.01$), 9.49 ($\alpha = 0.05$) and 7.78 ($\alpha = 0.10$).

2.2 Tests for uniformity in $[0, 1]^{d-1}$

Testing uniformity in the unit hypercube $[0, 1]^{d-1}$ is to test whether an i.i.d. $(d-1)$ dimensional sample $\{z_1, \ldots, z_n\}$ can be considered from the uniform distribution in $[0, 1]^{d-1}$. The hypothesis can be set up as

$$H_0: \ z_1, \ldots, z_n \text{ are uniformly distributed in } [0, 1]^{d-1}. \tag{9}$$

The alternative hypothesis H_1 implies rejection for H_0 in (9). Liang et al. (2001) proposed the following types of uniform statistics for testing uniformity in $[0, 1]^{d-1}$.

Type 1 Approximate $N(0, 1)$-statistics

$$A_n = \sqrt{n}\left[\left(U_1 - M^{d-1}\right) + 2\left(U_2 - M^{d-1}\right)\right] \Big/ (5\sqrt{\zeta_1}) \xrightarrow{\mathcal{D}} N(0, 1), \quad (n \to \infty) \tag{10}$$

under H_0 in (9), where "$\xrightarrow{\mathcal{D}}$" means convergence in probability distribution. There are three choices for A_n according to the three measures of discrepancy: symmetric, centered, and star (Hickernell 1998).

Type 2 Approximate χ^2-statistics

$$T_n = n\left[\left(U_1 - M^{d-1}\right), \left(U_2 - M^{d-1}\right)\right] \mathbf{\Sigma}_n^{-1} \left[\left(U_1 - M^{d-1}\right), \left(U_2 - M^{d-1}\right)\right]' \xrightarrow{\mathcal{D}} \chi^2(2), \ (n \to \infty) \tag{11}$$

under H_0 in (9), where

$$\mathbf{\Sigma}_n = \begin{pmatrix} \zeta_1 & 2\zeta_1 \\ 2\zeta_1 & \frac{4(n-2)}{n-1}\zeta_1 + \frac{2}{n-1}\zeta_2 \end{pmatrix}, \tag{12}$$

and ζ_1 and ζ_2 are calculated differently according to the three measures of discrepancies given as follows. There are also three choices for T_n.

The calculation of A_n in (10) and that of T_n in (11) are obtained according to any of the following three measures of discrepancy. From an i.i.d. $(d-1)$-dimensional sample $\{z_1, \ldots, z_n\}$ in $[0, 1]^{d-1}$, let $z_k = (z_{k1}, \ldots, z_{k,d-1})'$ $(k = 1, \ldots, n)$.

1. The symmetric discrepancy gives

$$U_1 = \frac{1}{n}\sum_{k=1}^{n}\prod_{j=1}^{d-1}\left(1 + 2z_{kj} - 2z_{kj}^2\right),$$
$$U_2 = \frac{2^d}{n(n-1)}\sum_{k<l}^{n}\prod_{j=1}^{d-1}\left(1 - |z_{kj} - z_{lj}|\right), \tag{13}$$

with $M = 4/3$, $\zeta_1 = (9/5)^{d-1} - (6/9)^{d-1}$ and $\zeta_2 = 2^{d-1} - (16/9)^{d-1}$;

2. The centered discrepancy gives

$$U_1 = \frac{1}{n}\sum_{k=1}^{n}\prod_{j=1}^{d-1}\left(1 + \frac{1}{2}\left|z_{kj} - \frac{1}{2}\right| - \frac{1}{2}\left|z_{kj} - \frac{1}{2}\right|^2\right),$$
$$U_2 = \frac{2}{n(n-1)}\sum_{k<l}^{n}\prod_{j=1}^{d-1}\left(1 + \frac{1}{2}\left|z_{kj} - \frac{1}{2}\right| + \frac{1}{2}\left|z_{lj} - \frac{1}{2}\right| - \frac{1}{2}\left|z_{kj} - z_{lj}\right|\right), \tag{14}$$

with $M = 13/12$, $\zeta_1 = (47/40)^{d-1} - (13/12)^{2(d-1)}$ and $\zeta_2 = (57/48)^{d-1} - (13/12)^{2(d-1)}$;

3. The star discrepancy gives

$$U_1 = \frac{1}{n}\sum_{k=1}^{n}\prod_{j=1}^{d-1}\left(\frac{3 - z_{kj}}{2}\right),$$
$$U_2 = \frac{2}{n(n-1)}\sum_{k<l}^{n}\prod_{j=1}^{d-1}\left[2 - \max(z_{kj}, z_{lj})\right], \tag{15}$$

with $M = 4/3$, $\zeta_1 = (9/5)^{d-1} - (16/9)^{d-1}$ and $\zeta_2 = (11/6)^{d-1} - (16/9)^{d-1}$.

The empirical finite-sample percentiles of A_n and T_n under the above three discrepancies were provided in Liang et al. (2001) for some selected sample sizes ($n = 25$, $n = 50$, $n = 100$ and $n = 200$). A large value of $|A_n|$ or T_n indicates evidence of non-uniformity for the underlying distribution of a sample from $[0, 1]^{d-1}$.

2.3 Construction of the uniform tests for $\mathcal{U}(d, p)$

Theorem 1 *Let* $\{\boldsymbol{u}_1, \ldots, \boldsymbol{u}_n\}$ *be an i.i.d. sample from* $\mathcal{U}(d, p)$. *Denote by* $\boldsymbol{u}_i = (U_{i1}, \ldots, U_{id})'$ *for* $i = 1, \ldots, n$. *For each fixed* i $(1 \le i \le n)$, *define the following random variables:*

$$
\begin{aligned}
B_1(i) &= |U_{i1}|^p, \\
B_2(i) &= \{(1 - |U_{i1}|^p)^{-1}|U_{i2}|^p, \\
&\vdots \quad \vdots \quad \vdots \\
B_k(i) &= \{(1 - \sum_{j=1}^{k-1} |U_{ij}|^p)^{-1}|U_{ik}|^p,
\end{aligned}
\tag{16}
$$

for $k = 2, \ldots, d-1$. *Then the following assertions are true:*

(1) *For each fixed* i $(1 \le i \le n)$, *the random variables* $\{B_1(i), \ldots, B_{d-1}(i)\}$ *are mutually independent, and* $B_k(i)$ *has a beta distribution* $\beta(1/p, (d-k)/p)$ $(k = 1, \ldots, d-1)$.
(2) *The* $n(d-1)$ *random variables* $\{B_1(i), \ldots, B_{d-1}(i) : i = 1, \ldots, n\}$ *are mutually independent.*

Proof For each fixed $1 \le i \le n$, $\boldsymbol{u}_i = (U_{i1}, \ldots, U_{id})' \sim \mathcal{U}(d, p)$, Theorem 2.1 in Gupta and Song (1997) gives the marginal density function of $(U_{i1}, \ldots, U_{im})'$ $(1 \le m \le d-1)$ as

$$
f_m(u_1, \ldots, u_m) = \frac{p^m \Gamma(d/p)}{2^m \Gamma^m(1/p)\Gamma((d-m)/p)} \left(1 - \sum_{l=1}^{m} |u_l|^p\right)^{\frac{d-m}{p}-1},
$$
$$
-1 < u_l < 1, \ \sum_{l=1}^{m} |u_l|^p < 1. \tag{17}
$$

In particular, let $m = 1$, U_{i1} has a density function

$$
f_{u_{i1}}(u_1) = \frac{p\Gamma(d/p)}{2\Gamma(1/p)\Gamma((d-1)/p)} \left(1 - |u_1|^p\right)^{(d-1)/p-1}, \qquad -1 < u_1 < 1.
$$

It is easy to derive $B_1(i) = |U_{i1}|^p$ has a density function

$$
g_{B_1(i)}(b_1) = \frac{\Gamma(d/p)}{\Gamma(1/p)\Gamma((d-1)/p)} b_1^{\frac{1}{p}-1} (1 - b_1)^{\frac{d-1}{p}-1}, \qquad 0 < b_1 < 1.
$$

That is, $B_1(i) \sim \beta(1/p, (d-1)/p)$, the beta distribution with parameters $(1/p, (d-1)/p)$. From marginal density function (17), for each fixed $k = 2, \ldots, d-1$, the conditional density function of $U_{ik}|(U_{i1}, \ldots, U_{i,k-1})$ can be obtained as

$$\begin{aligned} &f_{u_{ik}|(u_{i1},\ldots,u_{i,k-1})}(u_k) \\ &= \frac{p\Gamma\left(\frac{d-k+1}{p}\right)}{2\Gamma\left(\frac{1}{p}\right)\Gamma\left(\frac{d-k}{p}\right)}\left(1-\frac{|u_k|^p}{1-\sum_{j=1}^{k-1}|u_{ij}|^p}\right)^{\frac{d-k}{p}-1}\left(1-\sum_{j=1}^{k-1}|u_{ij}|^p\right), \\ &\sum_{j=1}^{k-1}|u_{ij}|^p < 1. \end{aligned} \tag{18}$$

It is easy to derive the density function of the conditional distribution for $B_k(i) = (1-\sum_{j=1}^{k-1}|U_{ij}|^p)^{-1}|U_{ik}|^p\big|(U_{i1}, \ldots, U_{i,k-1})$ is

$$f_{B_k(i)}(b_k) = \frac{\Gamma\left(\frac{d-k+1}{p}\right)}{\Gamma\left(\frac{1}{p}\right)\Gamma\left(\frac{d-k}{p}\right)} b_k^{\frac{1}{p}-1}(1-b_k)^{\frac{d-k}{p}-1}, \quad 0 < b_k < 1.$$

That is, $B_k(i) = (1-\sum_{j=1}^{k-1}|U_{ij}|^p)^{-1}|U_{ik}|^p\big|(U_{i1}, \ldots, U_{i,k-1}) \sim \beta(1/p, (d-k)/p)$. Because this conditional distribution does not depend on $(U_{i1}, \ldots, U_{i,k-1})$, we can conclude $B_k(i) \sim \beta(1/p, (d-k)/p)$,and $B_k(i)$ is independent of $(U_{i1}, \ldots, U_{i,k-1})$ for each fixed $1 \le i \le n$ and for $k = 2, \ldots, d-1$. This leads to the mutual independence of $\{B_1(i), \ldots, B_{d-1}(i)\}$ for each fixed $1 \le i \le n$. Because $\{\boldsymbol{u}_1, \ldots, \boldsymbol{u}_n\}$ is an i.i.d. sample from $\mathcal{U}(d, p)$ by assumption, and $\{B_1(i), \ldots, B_{d-1}(i)\}$ depend on $\boldsymbol{u}_i$ only, we can conclude that the $n(d-1)$ random variables $\{B_1(i), \ldots, B_{d-1}(i) : i = 1, \ldots, n\}$ are mutually independent. This completes the proof. □

Denote by $F_{b_j}(\cdot; p)$ the c.d.f. of $\beta(1/p, (d-j)/p)$ $(j = 1, \ldots, d-1)$. Let random variables $\boldsymbol{v}_i = (V_{i1}, \ldots, V_{i,d-1})'$ be given by

$$V_{ij} = F_{b_j}(B_j(i); p), \quad i = 1, \ldots, n; \ j = 1, \ldots, d-1. \tag{19}$$

Based on Theorem 1, a class of goodness-of-fit tests for $\mathcal{U}(d, p)$ can be constructed. Let $\boldsymbol{u}_1, \ldots, \boldsymbol{u}_n$ be i.i.d. observations. Perform the transformation

$$\begin{aligned} &\text{Original i.i.d. sample in Theorem 1: } \{\boldsymbol{u}_1, \ldots, \boldsymbol{u}_n\} \\ &\Rightarrow \{B_1(i), \ldots, B_{d-1}(i) : i = 1, \ldots, n; \text{ in } R^{d-1} \text{ by (16)}\} \\ &\Rightarrow \{\boldsymbol{v}_1, \ldots, \boldsymbol{v}_n : \text{ in } [0,1]^{d-1} \text{ by (19)}\}. \end{aligned} \tag{20}$$

If the null hypothesis H_0 in (3) is true for the underlying distribution of the i.i.d. sample $\{\boldsymbol{u}_1, \ldots, \boldsymbol{u}_n\}$, the random variables V_{ij}'s given by (19) are mutually independent and the V_{ij} has a uniform distribution $U(0, 1)$. Considering the random vector

$\boldsymbol{v}_i = (V_{i1}, \ldots, V_{i,d-1})' \in [0, 1]^{d-1}$, we can transfer a test for (3) to a test for multivariate uniformity

$$H_0: \quad \text{the } \boldsymbol{v}_i\text{'s are uniformly distributed in } [0, 1]^{d-1}, \tag{21}$$

versus H_1 that implies that H_0 in (21) is not true. Similarly, considering the random variables $V_{ij} \in (0,\ 1)$, we can transfer a test for (3) to a test for univariate uniformity

$$H_0: \quad \text{the } V_{ij}\text{'s are uniformly distributed in } (0, 1), \tag{22}$$

versus H_1 that implies that H_0 in (22) is not true. The statistics A_n in (10) and T_n in (11) can be employed to test (21), and the statistics MU^2 in (6) and P_4^2 in (8) can be employed to test (22). As a result, each of these four statistics can be used as a test for hypothesis (3).

It should be pointed out that acceptance of H_0 in (21) can lead to acceptance of H_0 in (22), but the contrary is usually not true. This implies that univariate uniformity of all one-dimensional marginal distributions does not automatically lead to multivariate uniformity for the joint distribution, see the illustration by Fig. 1 in Liang et al. (2008).

2.4 Construction of the uniform tests for $UW_d(p)$

Theorem 2 *Let* $\{\boldsymbol{h}_i = (H_{i1}, \ldots, H_{id})' : i = 1, \ldots, n\}$ *be an i.i.d. sample from* $UW_d(p)$. *For each fixed i* $(1 \le i \le n)$, *define the following random variables:*

$$\begin{aligned}
D_1(i) &= H_{i1}^p, \\
D_2(i) &= (1 - H_{i1}^p)^{-1} H_{i2}^p, \\
&\vdots \quad \vdots \quad \vdots \\
D_k(i) &= \left(1 - \sum_{j=1}^{k-1} H_{ij}^p\right)^{-1} H_{ik}^p,
\end{aligned} \tag{23}$$

for $k = 2, \ldots, d-1$. *Then the following assertions are true:*

(1) *For each fixed i* $(1 \le i \le n)$, *the random variables* $\{D_1(i), \ldots, D_{d-1}(i)\}$ *are mutually independent, and* $D_k(i)$ *has a beta distribution* $\beta(1, d-k)$ $(k = 1, \ldots, d-1)$.
(2) *The* $n(d-1)$ *random variables* $\{D_1(i), \ldots, D_{d-1}(i) : i = 1, \ldots, n\}$ *are mutually independent.*

Proof For each fixed $i = 1, \ldots, n$, Corollary 3 in Yue and Ma (1995) gives the marginal density function of $(H_{i1}, \ldots, H_{im})'$ $(1 \le m \le d-1)$ as

$$g_m(h_1,\ldots,h_m)=\frac{p^m\Gamma(d)}{\Gamma(d-m)}\cdot\prod_{l=1}^{m}h_l^{p-1}\left(1-\sum_{l=1}^{m}h_l^p\right)^{d-m-1},$$
$$h_l>0,\ \sum_{l=1}^{m}h_l^p<1. \tag{24}$$

In particular, let $m=1$, H_{i1} has a density function

$$g_1(h_1)=\frac{p\Gamma(d)}{\Gamma(d-1)}\cdot h_1^{p-1}\left(1-h_1^p\right)^{d-2},\quad 0<h_1<1.$$

The density function of $D_1(i)=H_{i1}^p$ can be easily derived as

$$f_{D_1}(d_1)=\frac{\Gamma(d)}{\Gamma(d-1)\Gamma(1)}(1-d_1)^{d-2},\quad 0<d_1<1.$$

That is, $D_1(i)\sim\beta(1,d-1)$. From marginal density function (24), for each fixed $k=2,\ldots,d-1$, the conditional density function of $H_{ik}|(H_{i1},\ldots,H_{i,k-1})$ can be obtained as

$$f_{h_{ik}|(h_{i1},\ldots,h_{i,k-1})}(h_k)$$
$$=\frac{p\Gamma(d-k+1)}{\Gamma(d-k)}h_k^{p-1}\left(1-\frac{h_k^p}{1-\sum_{j=1}^{k-1}h_{ij}^p}\right)^{d-k-1}\left(1-\sum_{j=1}^{k-1}h_{ij}^p\right)^{-1},$$
$$0<h_{ij}<1,\quad \sum_{j=1}^{k-1}h_{ij}^p<1. \tag{25}$$

It is easy to derive the density function of the conditional distribution for $D_k(i)=(1-\sum_{j=1}^{k-1}H_{ij}^p)^{-1}H_{ik}^p\big|(H_{i1},\ldots,H_{i,k-1})$ is

$$f_{D_k(i)}(d_k)=\frac{\Gamma(d-k+1)}{\Gamma(d-k)}(1-d_k)^{d-k-1},\quad 0\quad<d_k<1.$$

That is, $D_k(i)=(1-\sum_{j=1}^{k-1}H_{ij}^p)^{-1}H_{ik}^p\big|(H_{i1},\ldots,H_{i,k-1})\sim\beta(1,d-k)$ $(k=2,\ldots,d-1)$. Because this conditional distribution does not depend on $(H_{i1},\ldots,H_{i,k-1})$, we can conclude $H_k(i)\sim\beta(1,d-k)$,and $H_k(i)$ is independent of $(H_{i1},\ldots,H_{i,k-1})$ for each fixed $1\le i\le n$ and for $k=2,\ldots,d-1$. This leads to the mutual independence of $\{D_1(i),\ldots,D_{d-1}(i)\}$ for each fixed $1\le i\le n$. Because $\{\boldsymbol{h}_1,\ldots,\boldsymbol{h}_n\}$ is an i.i.d. sample from $UW_d(p)$ by assumption, and $\{H_1(i),\ldots,H_{d-1}(i)\}$ depend on $\boldsymbol{h}_i$ only, we can conclude that the $n(d-1)$ random variables $\{D_1(i),\ldots,D_{d-1}(i):\ i=1,\ldots,n\}$ are mutually independent. This completes the proof. □

Denote by $G_{b_j}(\cdot)$ the c.d.f. of $\beta(1, d-j)$ $(j = 1, \ldots, d-1)$. Let $\boldsymbol{c}_i = (C_{i1}, \ldots, C_{i,d-1})'$ be given by

$$C_{ij} = G_{b_j}(D_j(i)), \quad i = 1, \ldots, n; \; j = 1, \ldots, d-1. \tag{26}$$

Based on Theorem 2, a class of tests for $UW_d(p)$ can be constructed. Let $\boldsymbol{h}_1, \ldots, \boldsymbol{h}_n$ be i.i.d. observations. Perform the transformation

$$\begin{aligned} &\text{Original i.i.d. sample in Theorem 2: } \{\boldsymbol{h}_1, \ldots, \boldsymbol{h}_n\} \\ &\Rightarrow \{D_1(i), \ldots, D_{d-1}(i) : i = 1, \ldots, n; \;\text{ in }\; R_+^{d-1} \;\text{ by (23)}\} \\ &\Rightarrow \{\boldsymbol{c}_1, \ldots, \boldsymbol{c}_n : \;\text{ in }\; [0, 1]^{d-1} \;\text{ by (26)}\}, \end{aligned} \tag{27}$$

where $R_+^d = \{\boldsymbol{y} = (y_1, \ldots, y_d)' \in R^d, \; y_i \geq 0, \; i = 1, \ldots, d\}$. If the null hypothesis H_0 in (4) is true for the underlying distribution of the i.i.d. sample $\{\boldsymbol{h}_1, \ldots, \boldsymbol{h}_n\}$, the random variables C_{ij}'s given by (26) are mutually independent and the C_{ij} has a uniform distribution $U(0, 1)$. Considering the random vector $\boldsymbol{c}_i = (C_{i1}, \ldots, C_{i,d-1})' \in [0, 1]^{d-1}$, we can transfer a test for (4) to a test for multivariate uniformity

$$H_0 : \;\text{ the } \boldsymbol{c}_i\text{'s are uniformly distributed in } [0, 1]^{d-1}, \tag{28}$$

versus H_1 that implies that H_0 in (28) is not true. Similarly, considering the random variables $C_{ij} \in (0, \; 1)$, we can transfer a test for (4) to a test for univariate uniformity

$$H_0 : \;\text{ the } C_{ij}\text{'s are uniformly distributed in } (0, 1), \tag{29}$$

versus H_1 that implies that H_0 in (29) is not true. Therefore, the statistics introduced in Sect. 2.1 can be adopted to test hypotheses (3) and (4).

2.5 Extension of the uniform tests

The uniform distribution on the unit sphere in the sense of traditional Euclidean norm, the L_p-norm, or the l_p-norm plays an important role in constructing generalized multivariate distributions through a stochastic representation. For example, the family of spherical distributions in Fang et al. (1990) is generated by $\boldsymbol{x} \stackrel{d}{=} R\mathcal{U}_d$ with $\mathcal{U}_d$ having a uniform distribution on the unit sphere of the Euclidean space R^d and $R > 0$ being independent of $\mathcal{U}_d$; the ECD is generated by $\boldsymbol{y} \stackrel{d}{=} \boldsymbol{u} + RA\mathcal{U}_d$ with $\boldsymbol{u} \in R^d$ and $\boldsymbol{A}$ being a constant matrix; the family of the L_p-norm symmetric distribution $\mathcal{S}(d, p)$ in Definition 2 is generated by $\boldsymbol{x} \stackrel{d}{=} R\boldsymbol{u}_d$ with $\boldsymbol{u}_d \sim \mathcal{U}(d, p)$ (the uniform distribution on the L_p-norm unit sphere); the family of the l_p-norm symmetric distribution $\mathcal{L}_d(p)$ in Definition 3 is generated by $\boldsymbol{x} \stackrel{d}{=} r\boldsymbol{u}$, with $r \geq 0$ being independent of $\boldsymbol{u} \sim UW_d(p)$ in Definition 3. As a result of the stochastic representation, testing goodness-of-fit of any distribution in these families can be easily transferred to testing uniformity of some uniform distribution in the unit sphere. If we allow the stochastic representation some kind of dependence, the L_p-norm symmetric distribution $\mathcal{S}(d, p)$ in Definition 2 can be generalized to a much wider family:

$$\mathcal{GS}(d, p) = \{\boldsymbol{x},\ \boldsymbol{x} \stackrel{d}{=} R\boldsymbol{u}_d,\ P(\boldsymbol{x} = \boldsymbol{0}) = 0,\ \boldsymbol{u}_d \sim \mathcal{U}(d, p),\ \text{and } R > 0 \text{ is any random variable.}\} \tag{30}$$

It is obvious that $\mathcal{S}(d, p)$ in Definition 2 is a subfamily of $\mathcal{GS}(d, p)$ with the restriction that $R > 0$ is independent of $\boldsymbol{u}_d \sim \mathcal{U}(d, p)$. Similarly, the l_p-norm symmetric distribution $\mathcal{L}_d(p)$ in Definition 3 can be generalized to a much wider family:

$$\mathcal{GL}_d(p) = \{\boldsymbol{y} :\ \boldsymbol{y} \stackrel{d}{=} r\boldsymbol{u},\ P(\boldsymbol{y} = \boldsymbol{0}) = 0,\ \boldsymbol{u} \sim UW_d(p),\ \text{and } r \geq 0 \text{ is any random variable.}\} \tag{31}$$

It is easy to see that if $\{\boldsymbol{x}_1, \ldots, \boldsymbol{x}_n\}$ is an i.i.d. sample from $\mathcal{GS}(d, p)$, then

$$\{\boldsymbol{u}_1 = \boldsymbol{x}_1/\|\boldsymbol{x}_1\|,\ \ldots,\ \boldsymbol{u}_n = \boldsymbol{x}_n/\|\boldsymbol{x}_n\|\} \tag{32}$$

is an i.i.d. sample from $\mathcal{U}(d, p)$. Similarly, if $\{\boldsymbol{y}_1, \ldots, \boldsymbol{y}_n\}$ is an i.i.d. sample from $\mathcal{GL}_d(p)$, then

$$\left\{\boldsymbol{h}_1 = \boldsymbol{y}_1/\|\boldsymbol{y}_1\|_p^+,\ \ldots,\ \boldsymbol{h}_n = \boldsymbol{y}_n/\|\boldsymbol{y}_n\|_p^+\right\} \tag{33}$$

is an i.i.d. sample from $UW_d(p)$. Based on (32) and (33), we can transfer the goodness-of-fit test for the family of $\mathcal{GS}(d, p)$ or $\mathcal{GL}_d(p)$ to a test for uniformity by using the uniform tests in Sect. 2.4. Suppose that $\{\boldsymbol{x}_1, \ldots, \boldsymbol{x}_n\}$ is an i.i.d. sample. The goodness-of-fit test for the family of $\mathcal{GS}(d, p)$ is to test the null hypothesis

$$H_0 : \{\boldsymbol{x}_1, \ldots, \boldsymbol{x}_n\} \text{ is from } \mathcal{GS}(d, p) \tag{34}$$

versus the alternative hypothesis that $\{\boldsymbol{x}_1, \ldots, \boldsymbol{x}_n\}$ is not from $\mathcal{GS}(d, p)$. The goodness-of-fit test for the family of $\mathcal{GL}_d(p)$ based on an i.i.d. sample $\{\boldsymbol{y}_1, \ldots, \boldsymbol{y}_n\}$ is to test the null hypothesis

$$H_0 :\ \{\boldsymbol{y}_1, \ldots, \boldsymbol{y}_n\} \text{ is from } \mathcal{GL}_d(p) \tag{35}$$

versus the alternative hypothesis that $\{\boldsymbol{y}_1, \ldots, \boldsymbol{y}_n\}$ is not from $\mathcal{GL}_d(p)$. Perform the following transformation for the sample $\{\boldsymbol{x}_1, \ldots, \boldsymbol{x}_n\}$:

$$\begin{aligned}
&\text{Original i.i.d. sample: } \{\boldsymbol{x}_1, \ldots, \boldsymbol{x}_n :\ \text{in } R^d\} \\
&\quad \Rightarrow \{\boldsymbol{u}_1, \ldots, \boldsymbol{u}_n :\ \text{in } R^d \text{ by (32)}\} \\
&\quad \Rightarrow \{B_1(i), \ldots, B_{d-1}(i) :\ i = 1, \ldots, n;\ \text{in } R^{d-1} \text{ by (16)}\} \\
&\quad \Rightarrow \{\boldsymbol{v}_1, \ldots, \boldsymbol{v}_n :\ \text{in } [0, 1]^{d-1} \text{ by (19)}\}.
\end{aligned} \tag{36}$$

If the null hypothesis H_0 in (3) is true for the underlying distribution of the i.i.d. sample $\{\boldsymbol{x}_1, \ldots, \boldsymbol{x}_n\}$, the random variables V_{ij}'s obtained through (36) are mutually independent and the V_{ij} has a uniform distribution $U(0, 1)$. Considering the random vector $\boldsymbol{v}_i = (V_{i1}, \ldots, V_{i,d-1})' \in [0, 1]^{d-1}$, we can transfer a test for (3) to a test for multivariate uniformity

$$H_0: \text{ the } \boldsymbol{v}_i\text{'s are uniformly distributed in } [0, 1]^{d-1}, \tag{37}$$

versus H_1 that implies that H_0 in (37) is not true. Similarly, considering the random variables $V_{ij} \in (0, 1)$, we can transfer a test for (3) to a test for univariate uniformity

$$H_0: \text{ the } V_{ij}\text{'s are uniformly distributed in } (0, 1), \tag{38}$$

versus H_1 that implies that H_0 in (38) is not true. The statistics A_n in (10) and T_n in (11) can be employed to test (37), and the statistics MU^2 in (6) and P_4^2 in (8) can be employed to test (38). As a result, each of these four statistics can be used as a test for hypothesis (3).

It should be pointed out that acceptance of H_0 in (37) can lead to acceptance of H_0 in (38), but the contrary is usually not true. This implies that univariate uniformity of all one-dimensional marginal distributions does not automatically lead to multivariate uniformity for the joint distribution, see the illustration by Fig. 1 in Liang et al. (2008).

Similar to the transformation in (36), we perform the following transformation for an i.i.d. sample $\{\boldsymbol{y}_1, \ldots, \boldsymbol{y}_n\}$:

$$\begin{aligned} &\text{Original i.i.d. sample: } \{\boldsymbol{y}_1, \ldots, \boldsymbol{y}_n: \text{ in } R_+^d\} \\ &\Rightarrow \{\boldsymbol{h}_1, \ldots, \boldsymbol{h}_n: \text{ in } R_+^d \text{ by (33)}\} \\ &\Rightarrow \{D_1(i), \ldots, D_{d-1}(i): i = 1, \ldots, n; \text{ in } R_+^{d-1} \text{ by (23)}\} \\ &\Rightarrow \{\boldsymbol{c}_1, \ldots, \boldsymbol{c}_n: \text{ in } [0, 1]^{d-1} \text{ by (26)}\}, \end{aligned} \tag{39}$$

where $R_+^d = \{\boldsymbol{y} = (y_1, \ldots, y_d)' \in R^d, \; y_i \geq 0, \; i = 1, \ldots, d\}$. If the null hypothesis H_0 in (4) is true for the underlying distribution of the i.i.d. sample $\{\boldsymbol{y}_1, \ldots, \boldsymbol{y}_n\}$, the random variables C_{ij}'s obtained by (26) are mutually independent and the C_{ij} has a uniform distribution $U(0, 1)$. Considering the random vector $\boldsymbol{c}_i = (C_{i1}, \ldots, C_{i,d-1})' \in [0, 1]^{d-1}$, we can transfer a test for (4) to a test for multivariate uniformity

$$H_0: \text{ the } \boldsymbol{c}_i\text{'s are uniformly distributed in } [0, 1]^{d-1}, \tag{40}$$

versus H_1 that implies that H_0 in (40) is not true. Similarly, considering the random variables $C_{ij} \in (0, 1)$, we can transfer a test for (4) to a test for univariate uniformity

$$H_0: \text{ the } C_{ij}\text{'s are uniformly distributed in } (0, 1), \tag{41}$$

versus H_1 that implies that H_0 in (41) is not true. Therefore, the statistics introduced in Sect. 2.1 can be adopted to test hypotheses (3) and (4).

3 Monte Carlo study

In this section we carry out Monte Carlo studies on the performance of the uniform tests A_n in (10), T_n in (11), MU^2 in (6) and P_4^2 in (8) for testing goodness-of-fit for the two families of generalized multivariate distributions $\mathcal{GS}(d, p)$ and $\mathcal{GL}_d(p)$ as stated by two hypotheses (34) and (35) in Sect. 2.5. For the convenience of generating i.i.d. samples, we employ the two subfamilies: (1) the L_p-norm spherical distribution $\mathcal{S}(d, p)$ (Gupta and Song 1997); and (2) the l_p-norm symmetric distribution $\mathcal{L}_d(p)$ (Yue and Ma 1995) for illustrating the power performance of testing hypothesis (34) and (35), respectively. The empirical type I error rates for MU^2 and P_4^2, and the type I error rates for A_n and T_n under the three discrepancies (symmetric, centered, and star) and the power of all uniform statistics will be studied. In calculating the type I error rates and power, the corresponding percentiles for the multivariate uniform statistics A_n and T_n are chosen as those of their limiting distributions, respectively. That is, we consider $A_n \sim N(0, 1)$ and $T_n \sim \chi^2(2)$ for all sample sizes n. The percentiles of MU^2 and P_4^2 are given in Sect. 2.1. That is, for MU^2, they are 0.267 $(1-\alpha = 99\%)$, 0.187 $(1-\alpha = 95\%)$ and 0.152 $(1-\alpha = 90\%)$; for P_4^2, they are 13.28 $(1-\alpha = 99\%)$, 9.49 $(1-\alpha = 95\%)$ and 7.78 $(1-\alpha = 90\%)$.

3.1 Type I error rates in testing the L_p-norm spherical distribution

In the study on type I error rates of testing the L_p-norm spherical distribution, the null distribution is chosen as $\boldsymbol{x} \sim R\boldsymbol{u}$, where $R \sim \chi^2(5)$ (the Chi-square distribution with 5 degrees of freedom) and $\boldsymbol{u} \sim \mathcal{U}(d, p)$ with different choices of (d, p). The sample size n is chosen as $n = 25, 50, 100,$ and200 for each choice of (d, p). The stochastic representation given by Theorem 2 of Liang and Ng (2008) is employed to generate empirical samples from $\mathcal{S}(d, p)$. The empirical type I error rates are computed by

$$\text{Type I error rate} = \frac{\text{Number of rejections}}{\text{Number of replications}}. \tag{42}$$

Table 1 presents the simulation results on the type I error rates of A_n and T_n, and those of MU^2 and P_4^2, when testing goodness-of-fit for the L_p-norm spherical distribution with the significance level $\alpha = 0.05$ and number of replications $= 2000$. Similar simulation results were also obtained for the significance levels $\alpha = 0.01$ and 0.10, but these are not presented to save space. The following empirical conclusions can be summarized.

(1) When using the limiting distributions instead of the finite-sample distributions, the two multivariate uniform statistics A_n and T_n, and the two univariate uniform statistics MU^2 and P_4^2 can maintain feasible control of the type I error rates for the sample size n as small as $n = 25$ and the population dimension d as high as $d = 10$;

Table 1 Type I error rates of the test statistics ($\alpha = 0.05$)

Statistics	Discrepancy	n	Dimension $d = 5$			Dimension $d = 10$		
			$p = 1/2$	$p = 1$	$p = 3$	$p = 1/2$	$p = 1$	$p = 3$
A_n	Symmetric	25	0.0695	0.0720	0.0675	0.0740	0.0670	0.0730
		50	0.0600	0.0605	0.0550	0.0675	0.0585	0.0600
		100	0.0385	0.0490	0.0460	0.0635	0.0525	0.0520
		200	0.0480	0.0430	0.0465	0.0490	0.0575	0.0430
	Centered	25	0.0570	0.0525	0.0620	0.0555	0.0540	0.0630
		50	0.0540	0.0565	0.0570	0.0615	0.0575	0.0535
		100	0.0570	0.0570	0.0600	0.0555	0.0510	0.0425
		200	0.0490	0.0485	0.0450	0.0575	0.0495	0.0505
	Star	25	0.0470	0.0505	0.0485	0.0500	0.0525	0.0645
		50	0.0400	0.0455	0.0530	0.0555	0.0585	0.0540
		100	0.0485	0.0500	0.0440	0.0540	0.0550	0.0425
		200	0.0490	0.0510	0.0510	0.0510	0.0495	0.0500
T_n	Symmetric	25	0.0540	0.0575	0.0525	0.0530	0.0435	0.0490
		50	0.0455	0.0570	0.0460	0.0510	0.0425	0.0515
		100	0.0520	0.0560	0.0585	0.0505	0.0460	0.0540
		200	0.0535	0.0490	0.0550	0.0460	0.0555	0.0465
	Centered	25	0.0585	0.0605	0.0540	0.0480	0.0465	0.0525
		50	0.0515	0.0580	0.0570	0.0475	0.0530	0.0500
		100	0.0515	0.0505	0.0555	0.0525	0.0525	0.0585
		200	0.0530	0.0495	0.0530	0.0585	0.0540	0.0445
	Star	25	0.0560	0.0605	0.0610	0.0600	0.0580	0.0670
		50	0.0540	0.0580	0.0605	0.0590	0.0640	0.0625
		100	0.0500	0.0535	0.0585	0.0630	0.0660	0.0515
		200	0.0605	0.0565	0.0635	0.0640	0.0570	0.0565
MU^2		25	0.0445	0.0415	0.0420	0.0420	0.0465	0.0515
		50	0.0430	0.0450	0.0515	0.0525	0.0465	0.0460
		100	0.0395	0.0485	0.0555	0.0505	0.0490	0.0460
		200	0.0500	0.0460	0.0435	0.0550	0.0500	0.0475
P_4^2		25	0.0535	0.0510	0.0425	0.0535	0.0435	0.0565
		50	0.0455	0.0470	0.0535	0.0555	0.0490	0.0505
		100	0.0395	0.0460	0.0500	0.0475	0.0550	0.0540
		200	0.0495	0.0490	0.0570	0.0515	0.0555	0.0480

(2) The population dimension seems to have little influence on the type I error rates of the statistics A_n, T_n, MU^2, and P_4^2. This is shown in Table 1 for the cases of the same sample sizes but doubled dimension. This is a good indication in testing high-dimensional goodness-of-fit in the sense of avoiding curse of dimensionality.

3.2 Power in testing the L_p-norm spherical distribution

The empirical power of the tests is computed by (42) when choosing the following six non-L_p-norm spherical alternative distributions.

(1) the population $\boldsymbol{x} = (X_1, \ldots, X_d)' \sim N_d(\boldsymbol{0}, \boldsymbol{I}_d)$, the d-dimensional standard normal distribution;
(2) the population $\boldsymbol{x} = (X_1, \ldots, X_d)' \sim$ multivariate t-distribution with zero mean and identity covariance matrix, and the parameter $m = 5$ in Fang, Kotz, and Ng (1990, Chapter 3);
(3) the population $\boldsymbol{x} = (X_1, \ldots, X_d)' \sim$ multivariate Kotz-type distribution with zero mean and identity covariance matrix, and the parameter $(N, r, s) = (1, 1, 1)$ in Fang et al. (1990, Chapter 3);
(4) the population $\boldsymbol{x} = (X_1, \ldots, X_d)' \sim N + L_p S$, which means that $\boldsymbol{x} = (\boldsymbol{x}_1', \boldsymbol{x}_2')'$ with $\boldsymbol{x}_1 \sim \mathcal{S}(f, p)$ and $\boldsymbol{x}_2 \sim N_{d-f}(\boldsymbol{0}, \boldsymbol{I}_{d-f})$, where $f = [d/2]$, and $[\cdot]$ means the integer part of a real number (e.g., $[5/2] = 2$, $[10/2] = 5$);
(5) the population $\boldsymbol{x} = (X_1, \ldots, X_d)' \sim t + L_p S$, which means that $\boldsymbol{x} = (\boldsymbol{x}_1', \boldsymbol{x}_2')'$ with $\boldsymbol{x}_1 \sim \mathcal{S}(f, p)$ and $\boldsymbol{x}_2 \sim (d - f)$-dimensional t-distribution with zero mean and identity covariance matrix, and the parameter $m = 5$;
(6) the population $\boldsymbol{x} = (X_1, \ldots, X_d)' \sim K + L_p S$, which means that $\boldsymbol{x} = (\boldsymbol{x}_1', \boldsymbol{x}_2')'$ with $\boldsymbol{x}_1 \sim \mathcal{S}(f, p)$ and $\boldsymbol{x}_2 \sim (d - f)$-dimensional Kotz-type distribution with zero mean and identity covariance matrix, and the parameter $(N, r, s) = (1, 1, 1)$.

The simulation for empirical power of the four statistics was carried out with 2000 replications by MATLAB code (available from the authors upon request). Table 2 ($p = 1/2$) and Table 3 ($p = 1$), respectively, present the power values (the rejection rates) in testing goodness-of-fit for the six alternative distributions, where Table 2 only presents the results for the sample size $n = 25$ because almost all power values for $n = 50$ are equal to 1 for $p = 1/2$.

Based on Tables 2 and 3, we can summarize our empirical conclusions on the power performance of the uniform statistics in testing goodness-of-fit for the L_p-norm spherical distribution as follows.

(1) Similar to the performance in their type I error rates, all four uniform statistics A_n, T_n, MU^2, and P_4^2 are not sensitive to the increase of the dimension. For example, when the dimension d increases from $d = 5$ to $d = 10$, the four statistics have similar power performance for each of the selected alternative distributions;
(2) For the two multivariate uniform statistics A_n and T_n, their power performance is different for different choices of discrepancy measures. The symmetric discrepancy tends to be the best, centered discrepancy the second, and the star discrepancy the worst in all cases;
(3) The two multivariate uniform statistics A_n and T_n cannot outperform the two univariate uniform statistics MU^2 and P_4^2 in all cases and vice versa. So in practical applications, all of these statistics can be used together to give more confidence in drawing conclusions.

Table 2 Empirical power of the test statistics ($\alpha = 0.05$)

Statistics	Discrepancy	n	Alternative distributions, dimension $d = 5$, $p = 1/2$					
			Normal	Multi. t	Kotz	$N + L_pS$	$t + L_pS$	$K + L_pS$
A_n	Symmetric	25	0.9995	0.9995	0.9985	0.9385	0.9150	0.9220
	Centered	25	0.9345	0.9410	0.9350	0.3165	0.4135	0.1160
	Star	25	0.0545	0.0510	0.0445	0.0530	0.0210	0.3175
T_n	Symmetric	25	0.9950	0.9975	0.9975	0.9140	0.8635	0.9775
	Centered	25	0.9740	0.9775	0.9655	0.8370	0.7675	0.9625
	Star	25	0.4405	0.4585	0.4475	0.6685	0.4370	0.9610
MU^2		25	0.9955	0.9970	0.9960	0.8180	0.7765	0.7780
P_4^2		25	0.9920	0.9935	0.9900	0.7240	0.6840	0.7545
Statistics	Discrepancy	n	Alternative distributions, dimension $d = 10$, $p = 1/2$					
			Normal	Multi. t	Kotz	$N + L_pS$	$t + L_pS$	$K + L_pS$
A_n	Symmetric	25	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
	Centered	25	0.9920	0.9870	0.9890	1.0000	1.0000	1.0000
	Star	25	0.0185	0.0215	0.0250	1.0000	1.0000	1.0000
T_n	Symmetric	25	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
	Centered	25	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
	Star	25	0.8445	0.8510	0.8450	1.0000	1.0000	1.0000
MU^2		25	1.0000	1.0000	1.0000	0.7950	0.8515	0.4010
P_4^2		25	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000

Table 3 Empirical power of the test statistics ($\alpha = 0.05$)

Statistics	Discrepancy	n	Alternative distributions, dimension $d = 5$, $p = 1$					
			Normal	Multi. t	Kotz	$N + L_p S$	$t + L_p S$	$K + L_p S$
A_n	Symmetric	25	0.6260	0.6410	0.6205	0.3830	0.3040	0.4735
	Centered	25	0.3615	0.3590	0.3650	0.8125	0.6710	0.9840
	Star	25	0.0285	0.0270	0.0295	0.9390	0.8775	0.9905
T_n	Symmetric	25	0.4790	0.4865	0.4725	0.9930	0.9630	1.0000
	Centered	25	0.3380	0.3335	0.3255	0.9930	0.9645	1.0000
	Star	25	0.0735	0.0680	0.0655	0.9980	0.9915	1.0000
MU^2		25	0.4840	0.4880	0.4725	0.5720	0.4905	0.8090
P_4^2		25	0.3795	0.3900	0.3615	0.9395	0.8550	0.9965
A_n	Symmetric	50	0.8960	0.8865	0.8970	0.5740	0.4880	0.7660
	Centered	50	0.6135	0.6250	0.6180	0.9750	0.8900	1.0000
	Star	50	0.0305	0.0290	0.0360	1.0000	0.9960	1.0000
T_n	Symmetric	50	0.8060	0.8060	0.8055	1.0000	1.0000	1.0000
	Centered	50	0.6360	0.6240	0.6345	1.0000	1.0000	1.0000
	Star	50	0.1460	0.1430	0.1515	1.0000	1.0000	1.0000
MU^2		50	0.8145	0.8030	0.8100	0.9245	0.8260	0.9960
P_4^2		50	0.7415	0.7395	0.7390	1.0000	0.9940	1.0000

Table 3 continued

Statistics	Discrepancy	n	Alternative distributions, dimension $d = 10,\ p = 1$					
			Normal	Multi. t	Kotz	$N + L_pS$	$t + L_pS$	$K + L_pS$
A_n	Symmetric	25	0.9105	0.8960	0.9070	0.9730	0.9655	0.9560
	Centered	25	0.5005	0.5115	0.5220	0.9780	0.9755	0.9940
	Star	25	0.0190	0.0245	0.0235	0.3630	0.5420	0.0350
T_n	Symmetric	25	0.8490	0.8315	0.8340	1.0000	1.0000	1.0000
	Centered	25	0.5755	0.5660	0.5755	1.0000	1.0000	1.0000
	Star	25	0.1220	0.1150	0.1190	1.0000	1.0000	1.0000
MU^2		25	0.8845	0.8810	0.8885	0.0695	0.1070	0.0945
P_4^2		25	0.8690	0.8625	0.8655	0.1365	0.2505	0.3050
A_n	Symmetric	50	0.9955	0.9980	0.9955	0.9995	1.0000	0.9990
	Centered	50	0.7920	0.8085	0.7690	1.0000	1.0000	1.0000
	Star	50	0.0290	0.0255	0.0250	0.5765	0.8165	0.0435
T_n	Symmetric	50	0.9850	0.9925	0.9850	1.0000	1.0000	1.0000
	Centered	50	0.8945	0.9085	0.8935	1.0000	1.0000	1.0000
	Star	50	0.3215	0.3335	0.3285	1.0000	1.0000	1.0000
MU^2		50	0.9975	0.9980	0.9965	0.1365	0.2195	0.1820
P_4^2		50	0.9985	0.9985	0.9970	0.2600	0.4515	0.5465

Table 4 Type I error rates of the test statistics ($\alpha = 0.05$)

Statistics	Discrepancy	n	Dimension $d = 5$			Dimension $d = 10$		
			$p = 1/2$	$p = 1$	$p = 3$	$p = 1/2$	$p = 1$	$p = 3$
A_n	Symmetric	25	0.0710	0.0605	0.0460	0.0470	0.0515	0.0570
		50	0.0635	0.0430	0.0575	0.0560	0.0530	0.0600
		100	0.0495	0.0445	0.0590	0.0515	0.0535	0.0560
		200	0.0555	0.0475	0.0540	0.0440	0.0445	0.0480
	Centered	25	0.0590	0.0575	0.0655	0.0530	0.0500	0.0575
		50	0.0740	0.0445	0.0600	0.0720	0.0590	0.0575
		100	0.0465	0.0555	0.0650	0.0440	0.0490	0.0515
		200	0.0560	0.0485	0.0535	0.0530	0.0515	0.0450
	Star	25	0.0650	0.0665	0.0625	0.0505	0.0470	0.0505
		50	0.0710	0.0480	0.0570	0.0675	0.0600	0.0610
		100	0.0565	0.0535	0.0655	0.0650	0.0555	0.0530
		200	0.0585	0.0505	0.0480	0.0405	0.0395	0.0515
T_n	Symmetric	25	0.0750	0.0700	0.0690	0.0615	0.0525	0.0670
		50	0.0540	0.0410	0.0560	0.0600	0.0665	0.0715
		100	0.0580	0.0470	0.0515	0.0445	0.0495	0.0565
		200	0.0525	0.0530	0.0580	0.0620	0.0570	0.0555
	Centered	25	0.0680	0.0690	0.0650	0.0660	0.0740	0.0745
		50	0.0645	0.0495	0.0480	0.0785	0.0655	0.0630
		100	0.0590	0.0545	0.0620	0.0470	0.0590	0.0625
		200	0.0550	0.0515	0.0505	0.0650	0.0635	0.0520
	Star	25	0.0685	0.0670	0.0460	0.0465	0.0400	0.0425
		50	0.0535	0.0420	0.0540	0.0545	0.0515	0.0490
		100	0.0525	0.0530	0.0625	0.0565	0.0465	0.0470
		200	0.0455	0.0395	0.0345	0.0320	0.0365	0.0375
MU^2		25	0.0685	0.0640	0.0590	0.0460	0.0440	0.0520
		50	0.0535	0.0345	0.0520	0.0585	0.0565	0.0515
		100	0.0440	0.0475	0.0555	0.0450	0.0485	0.0500
		200	0.0505	0.0485	0.0445	0.0455	0.0465	0.0505
P_4^2		25	0.0450	0.0490	0.0595	0.0460	0.0415	0.0445
		50	0.0580	0.0335	0.0540	0.0605	0.0465	0.0410
		100	0.0345	0.0450	0.0545	0.0435	0.0455	0.0475
		200	0.0485	0.0505	0.0450	0.0530	0.0500	0.0465

3.3 Type I error rates in testing the l_p-norm symmetric distribution

The null distribution is chosen as $\boldsymbol{x} \sim r\boldsymbol{u}$, where $r \sim \chi^2(2)$ and $\boldsymbol{u} \sim UW_d(p)$ with different choices of (d, p). The sample size n is chosen as $n = 25, 50, 100$, and 200 for each choice of (d, p). Stochastic representation (2) is employed to generate empirical

Table 5 Empirical power of the test statistics ($\alpha = 0.05$)

Statistics	Discrepancy	n	Alternative distributions, dimension $d = 5$, $p = 1/2$					
			$\chi^2(2)$	$F(2, 5)$	Rayleigh	$\chi^2 + l_p n$	$F + l_p n$	$R + l_p n$
A_n	Symmetric	25	1.0000	1.0000	1.0000	1.0000	0.9985	1.0000
	Centered	25	0.9970	0.9770	1.0000	0.4845	0.5550	0.9955
	Star	25	0.0330	0.0375	0.2500	0.9915	0.8245	0.9785
T_n	Symmetric	25	1.0000	1.0000	1.0000	1.0000	0.9940	1.0000
	Centered	25	1.0000	0.9940	1.0000	1.0000	0.9720	1.0000
	Star	25	0.9605	0.4985	1.0000	1.0000	0.9225	1.0000
MU^2		25	1.0000	1.0000	1.0000	0.9940	0.9420	1.0000
P_4^2		25	1.0000	1.0000	1.0000	0.9990	0.9400	1.0000
Statistics	Discrepancy	n	Alternative distributions, dimension $d = 10$, $p = 1/2$					
			$\chi^2(2)$	$F(2, 5)$	Rayleigh	$\chi^2 + l_p n$	$F + l_p n$	$R + l_p n$
A_n	Symmetric	25	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
	Centered	25	1.0000	0.9995	1.0000	1.0000	0.9990	0.8905
	Star	25	0.0065	0.0135	0.0540	1.0000	1.0000	1.0000
T_n	Symmetric	25	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
	Centered	25	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
	Star	25	1.0000	0.8310	1.0000	1.0000	1.0000	1.0000
MU^2		25	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
P_4^2		25	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000

Table 5 continued

Statistics	Discrepancy	n	Alternative distributions, dimension $d = 5$, $p = 3$					
			$\chi^2(2)$	$F(2, 5)$	Rayleigh	$\chi^2 + l_p n$	$F + l_p n$	$R + l_p n$
A_n	Symmetric	25	1.0000	1.0000	0.9645	0.9795	0.9995	0.9965
	Centered	25	1.0000	1.0000	0.9065	1.0000	1.0000	0.9980
	Star	25	0.2685	0.4030	0.1250	0.7260	0.3865	0.2730
T_n	Symmetric	25	1.0000	1.0000	0.9865	1.0000	1.0000	1.0000
	Centered	25	1.0000	1.0000	0.8500	1.0000	1.0000	0.9955
	Star	25	1.0000	1.0000	0.3370	1.0000	1.0000	0.8990
MU^2		25	1.0000	1.0000	0.9635	1.0000	1.0000	1.0000
P_4^2		25	1.0000	1.0000	0.9970	1.0000	1.0000	1.0000
Statistics	Discrepancy	n	Alternative distributions, dimension $d = 10$, $p = 3$					
			$\chi^2(2)$	$F(2, 5)$	Rayleigh	$\chi^2 + l_p n$	$F + l_p n$	$R + l_p n$
A_n	Symmetric	25	0.9990	0.9950	0.9965	0.4600	0.4725	0.5765
	Centered	25	1.0000	1.0000	0.9980	1.0000	1.0000	1.0000
	Star	25	0.9435	0.9960	0.2405	1.0000	1.0000	0.8815
T_n	Symmetric	25	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
	Centered	25	1.0000	1.0000	0.9970	1.0000	1.0000	1.0000
	Star	25	1.0000	1.0000	0.5860	1.0000	1.0000	0.9985
MU^2		25	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
P_4^2		25	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000

samples from $UW_d(p)$. Table 4 presents the simulation results when testing goodness-of-fit for the l_p-norm symmetric distribution with the significance level $\alpha = 0.05$ and number of replications = 2000. Similar simulation results were also obtained for the significance levels $\alpha = 0.01$ and 0.10, but these are not presented to save space. The same conclusions as those from testing goodness-of-fit of the L_p-norm spherical distribution can be obtained: all statistics can maintain feasible control of the type I error rates for the sample size as small as $n = 25$.

3.4 Power in testing the l_p-norm symmetric distribution

The following six non-l_p-norm symmetric alternative distributions are chosen for testing goodness-of-fit of the l_p-norm symmetric distribution.

(1) the population $\boldsymbol{x} = (X_1, \ldots, X_d)' \sim \chi^2(2)$, which means that $X_1, \ldots, X_d$ are i.i.d. and $X_1 \sim \chi^2(2)$;
(2) the population $\boldsymbol{x} = (X_1, \ldots, X_d)' \sim F(2, 5)$, which means that $X_1, \ldots, X_d$ are i.i.d. and $X_1 \sim F(2, 5)$ (the F-distribution with degrees of freedom (2, 5));
(3) the population $\boldsymbol{x} = (X_1, \ldots, X_d)' \sim$Rayleigh, which means that $X_1, \ldots, X_d$ are i.i.d. and X_1 has a Rayleigh distribution with p.d.f. $2x\exp(-x^2)$ $(x > 0)$;
(4) the population $\boldsymbol{x} = (X_1, \ldots, X_d)' \sim \chi^2 + l_p n$, which means that $\boldsymbol{x} = (\boldsymbol{x}_1', \boldsymbol{x}_2')'$ with $\boldsymbol{x}_1 \sim \mathcal{L}_f(p)$ $(f = [d/2])$ and $\boldsymbol{x}_2 \sim \chi^2(2)$, (the components of $\boldsymbol{x}_2$ are i.i.d. $\chi^2(2)$);
(5) the population $\boldsymbol{x} = (X_1, \ldots, X_d)' \sim F + l_p n$, which means that $\boldsymbol{x} = (\boldsymbol{x}_1', \boldsymbol{x}_2')'$ with $\boldsymbol{x}_1 \sim \mathcal{L}_f(p)$ and $\boldsymbol{x}_2 \sim F(2, 5)$ (the components of $\boldsymbol{x}_2$ are i.i.d. $F(2, 5)$);
(6) the population $\boldsymbol{x} = (X_1, \ldots, X_d)' \sim R + L_p S$, which means that $\boldsymbol{x} = (\boldsymbol{x}_1', \boldsymbol{x}_2')'$ with $\boldsymbol{x}_1 \sim \mathcal{L}_f(p)$ and $\boldsymbol{x}_2 \sim$Rayleigh distribution (the components of $\boldsymbol{x}_2$ are i.i.d., and each of them has a Rayleigh distribution with p.d.f. $2x\exp(-x^2)$ $(x > 0)$.

The simulation was carried out with 2000 replications by MATLAB code (available from the authors upon request). Table 5 presents the power values (the rejection rates) in testing goodness-of-fit for the six alternative distributions, where only the results for the sample size $n = 25$ are presented because most of the power values for $n \geq 50$ are equal to 1. The results in Table 5 show that the statistics are very powerful in testing goodness-of-fit of these non-l_p-norm symmetric distributions in most cases.

4 Concluding remarks

Testing goodness-of-fit for the two families of generalized multivariate distributions $\mathcal{GS}(d, p)$ and $\mathcal{GL}_d(p)$ is a challenging topic. $\mathcal{GS}(d, p)$ contains the L_p-norm spherical distributions as its special case, and $\mathcal{GL}_d(p)$ contains the l_p-norm symmetric distributions as its special case. The main difficulty in constructing the goodness-of-fit tests is the challenge of developing the distributional characterization properties for these types of distributions. The uniform tests in this paper are developed from the principle of conditional probability integral transformation (CPIT), which was dated back to Rosenblatt (1952) and O'Reilly and Quesenberry (1973). One of the benefits from CPIT is the easiness in test construction and numerical computation of the test statistics due to the fact that most popular computer programs such as MATLAB and SAS provide the internal functions for computing many cumulative distribution functions and their inverse functions. Another benefit from CPIT is the fast convergence of the null distributions of the test statistics to their limiting null distributions because CPIT transforms a sample into a uniform one in the unit interval (0, 1) or into one in the hypercube $[0, 1]^{d-1}$ ($d \geq 2$), which is always bounded with probability 1. This is verified by the simulation results in Sect. 3, where it shows that even for the population dimension d as high as $d = 10$ and the sample size n as low as $n = 25$, the type I error rates of the test statistics are regularly controlled within a feasible range of the preassigned significance level. Compared to many traditional approaches to constructing goodness-of-fit tests that are based on the large sample theory, although we use their large sample percentiles for the Monte Carlo studies on type I error rates and power, the uniform tests in this paper turn out to control their type I error rates very well. Liang et al. (2001) carried out a comprehensive Monte Carlo study on the convergence of the multivariate uniform statistics A_n given by (10) and T_n given by (11). The fast convergence of A_n and T_n in testing hypotheses (3) and (4) can be also verified by the simulation results on the type I error rates in Tables 1 and 4.

In the choice of statistics for testing uniformity, it is noted that the multivariate uniform tests A_n or T_n cannot always outperform the univariate ones MU^2 and P_4^2. It is no wonder that different measures for uniformity were developed under different points of view. The multivariate measures of uniformity A_n and T_n were both developed through approximating a figure of merit for measuring non-uniformity for a given sample in a high-dimensional unit cube, while the univariate indices MU^2 and P_4^2 only measure the linear uniformity of a sample in the unit interval (0, 1). Non-uniformity of a sample in a high-dimensional unit cube usually implies non-uniformity of any one-dimensional marginal distribution of the sample, whereas the uniformity in (0, 1) of a one-dimensional marginal distribution of a high-dimensional sample usually does not reflect uniformity of the sample in the hypercube $[0, 1]^d$ ($d \geq 2$). Because of the different starting points for measuring uniformity in the hypercube $[0, 1]^d$, A_n and T_n have different performance for the three types of discrepancies (symmetric, centered, and star). This is reflected by the results in Tables 2, 3, and 5. The effect from choosing different discrepancies to construct the two measures A_n and T_n was studied in Liang et al. (2001). The empirical conclusion is that A_n and T_n under symmetric discrepancy usually perform the best in identifying non-uniformity in the hypercube $[0, 1]^d$; A_n and T_n under centered discrepancy perform the second; and A_n and T_n under the star discrepancy perform the worst. The results in Tables 2, 3, and 5 basically verify the empirical conclusion in Liang et al. (2001) about the choice of discrepancy in measuring uniformity in $[0, 1]^d$. Although it is an open theoretical problem to prove that the symmetric discrepancy is always the best, the centered discrepancy the second, and the star discrepancy the worst, it helps to get better understanding by thinking about the difference of their starting points. The star discrepancy is defined by measuring the generalized distance between the empirical distribution function of a set of projected sample on $[0, 1]^d$ and the distribution function of a set of d i.i.d. uniform distributions $U(0, 1)$. The centered discrepancy is anchored to the center point $(1/2, \ldots, 1/2)$ of the hypercube $[0, 1]^d$ compared to the star discrepancy that is anchored to the origin $(0, \ldots, 0)$ of $[0, 1]^d$. The center discrepancy is invariant under some kind of reflections. The symmetric discrepancy is anchored to an average of some function values over all possible vertices of the hypercube. This kind of average may reduce the drawback from choosing a fixed reference point like the star and centered discrepancies. Averaging over possible vertices may act as a kind of compensation for choosing a fixed vertex or the center of the cube. This could shed some light on the fact that the two measures of uniformity A_n and T_n under the symmetric discrepancy perform better than those cases under the star and the centered discrepancies in most cases as demonstrated by the empirical results in Tables 2, 3, and 5.

It is pointed out that CPIT only results in necessary tests for goodness-of-fit. This implies that when the test rejects null hypothesis (3) or (4), one can be confident that the null hypothesis is not true at the given significance level, but the non-rejection of the test, in general, does not lead to the truth of the null hypothesis. In other words, some distributions that are not in the family of $\mathcal{GS}(d, p)$ or $\mathcal{GL}_d(p)$ could pass the uniform tests for hypothesis (3) or (4) at a given significance level. The necessity of goodness-of-fit tests without sufficiency is a common drawback that most goodness-of-fit tests in the literature possess. Construction of necessary and sufficient test statistics is a challenging topic in goodness-of-fit area. This is true even for the normal distribution. The uniform tests based on the CPIT method in this paper provide a general way to construct tests for goodness-of-fit for many generalized multivariate distributions for which the CPIT is easy to obtain. Furthermore, the uniform tests in this paper seem to control their type I error rates very well and have satisfactory power performance. Because no existing goodness-of-fit tests for the same purpose are available in the literature, we are not able to carry out a simple Monte Carlo comparison between the uniform tests and others. We hope to see some comparable tests for the same purpose will be available in the future.

References

Anderson, T. W. (1993). Nonnormal multivariate distributions: Inference based on elliptically contoured distributions. In C. R. Rao (Ed.), *Multivariate analysis: Future directions* (pp. 1–24). Amsterdam, London: Elsevier (North Holland Publishing Corporation).

Anderson, T. W., Fang, K. T., Hsu, H. (1986). Maximum likelihood estimates and likelihood ratio criteria for multivariate elliptically contoured distributions. *Canadian Journal of Statistics*, *14*, 55–59.

Fang, K. T., Liang, J. (1999). Testing spherical and elliptical symmetry. In S. Kotz, C. B. Read, & D. L. Banks (Eds.), *Encyclopedia of statistical sciences (update)* (Vol. 3, pp. 686–691). New York: Wiley.

Fang, K. T., Zhang, Y. (1990). *Generalized multivariate analysis*. Berlin: Springer.

Fang, K. T., Kotz, S., Ng, K. W. (1990). *Symmetric multivariate and related distributions*. London and New York: Chapman and Hall.

Gupta, A. K., Kabe, D. G. (1993). Multivariate robust tests for spherical symmetry with applications to multivariate least squares regression. *Journal of Applied Statistical Science*, *1*(2), 159–168.

Gupta, A. K., Song, D. (1997). L_p-norm spherical distributions. *Journal of Statistical Planning & Inference*, *60*, 241–260.

Gupta, A. K., Varga, T. (1993). *Elliptically contoured models in statistics*. Dordrecht, The Netherlands: Kluwer Academic Publishers.

Hickernell, F. J. (1998). A generalized discrepancy and quadrature error bound. *Mathematics of Computation*, *67*, 299–322.

Huffer, F. W., Park, C. (2007). A test for elliptical symmetry. *Journal of Multivariate Analysis*, *98*, 256–281.

Kariya, T., Eaton, M. L. (1977). Robust tests for spherical symmetry. *The Annals of Statistics*, *5*, 206–215.

Lange, K. L., Little, R. J. A., Taylor, J. M. G. (1989). Robust statistical modeling using the *t*-distribution. *Journal of the American Statistical Association*, *84*, 881–896.

Liang, J., Fang, K. T. (2000). Some applications of Läuter's technique in tests for spherical symmetry. *Biometrical Journal*, *42*, 923–936.

Liang, J., Ng, K. W. (2008). A method for generating uniformly scattered points on the L_p-norm unit sphere and its applications. *Metrika*, *68*, 83–98.

Liang, J., Fang, K. T., Hickernell, F. J., Li, R. (2001). Testing multivariate uniformity and its applications. *Mathematics of Computation*, *70*, 337–355.

Liang, J., Fang, K. T., Hickernell, F. J. (2008). Some necessary uniform tests for spherical symmetry. *Annals of the Institute of Statistical Mathematics*, *60*, 679–696.

Manzottia, A., Pérez, F. J., Quiroz, A. J. (2002). A Statistic for testing the null hypothesis of elliptical symmetry. *Journal of Multivariate Analysis*, *81*, 274–285.

Miller, F. L, Jr., Quesenberry, C. P. (1979). Power studies of tests for uniformity, II. *Communications of Statistics-Simulation and Computation B*, *8*(3), 271–290.

Neyman, J. (1937). "Smooth" test for goodness of fit. *Journal of the American Statistical Association*, *20*, 149–199.

O'Reilly, F. J., Quesenberry, C. P. (1973). The conditional probability integral transformation and applications to obtain composite chi-square goodness-of-fit tests. *The Annals of Statistics*, *1*, 74–83.

Osiewalski, J., Steel, M. F. J. (1993). Robust Bayesian inference in l_q-spherical models. *Biometrika*, *80*, 456–460.

Rosenblatt, M. (1952). Remarks on a multivariate transformation. *The Annals of Mathematical Statistics*, *23*, 470–472.

Schott, J. R. (2002). Testing for elliptical symmetry in covariance-matrix-based analysis. *Statistics & Probability Letters*, *60*, 395–404.

Stephens, M. A. (1970). Use of the Kolmogorov Smirnov, Cramér-von Mises and related statistics without extensive tables. *Journal of the Royal Statistical Society (Series B)*, *32*, 115–122.

Watson, G. S. (1962). Goodness-of-fit tests on a circle. *II. Biometrika*, *49*, 57–63.

Yue, X., Ma, C. (1995). Multivariate l_p-norm symmetric distributions. *Statistics & Probability Letters*, *24*, 281–288.

Zellner, A. (1976). Bayesian and non-Bayesian analysis of the regression model with multivariate Student-t error terms. *Journal of the American Statistical Association*, *71*, 400–405.

Zhu, L. X. (2003). Conditional tests for elliptical symmetry. *Journal of Multivariate Analysis*, *84*, 284–298.