DEPTH-BASED SIGNED-RANK TESTS FOR BIVARIATE CENTRAL SYMMETRY

Authors: Sakineh Dehghan
– Department of Statistics,
Shahid Beheshti University, Tehran, Iran
(sa_ dehghan@sbu.ac.ir)

Mohammad Reza Faridrohani
– Department of Statistics, Faculty of Mathematical Sciences,
Shahid Beheshti University, Tehran, Iran
(m_ faridrohani@sbu.ac.ir)

Abstract:

- In this paper, distribution-free, affine invariant, signed-rank test statistics are proposed for the hypothesis that a bivariate distribution is centrally symmetric about an arbitrary specified point. The proposed tests are based on the concept of data depth. However, our tests are inherently orthogonal invariant, an affine invariant version of them is provided by using Tyler's estimator of scatter. The limiting null distribution of proposed tests is derived and the performance of the proposed tests is evaluated through a Monte Carlo study. This study demonstrates that the tests always detect asymmetry and they are convenient to determine small departures from the null hypothesis with high power. Also it shows that the tests perform well comparing other procedures in the literature.

Key-Words:

- *Affine invariance; Central symmetry; Depth function; Distribution-free; Tyler's estimator of scatter.*

AMS Subject Classification:

- 49A05, 78B26.

1. INTRODUCTION

Let $\mathbf{X}_1, ..., \mathbf{X}_n$ denote independent copies of the bivariate random vector $\mathbf{X} = (X_1, X_2)^T$ from a continuous bivariate population. One problem which has been considered in the literature is to test whether the distribution is symmetric about an unknown center against the alternative that the symmetry is lost (Heathcote et al. [16], Koltchinskii and Li [22], Neuhaus and Zhu [33], Manzotti et al. [30] and Henze et al. [17]). Moreover, in the univariate case, we can mention to Cassart et al. [6]. Unlike the univariate case, there are several concepts of multivariate symmetry including spherical, elliptical, central and angular symmetry. It is worth noting that the mentioned arrangement of the multivariate symmetry concepts are ordered in increasing generality. To read more about different types of multivariate symmetry see Serfling [38].

A different problem is the testing of the hypothesis that the bivariate distribution is symmetric about a known center $\boldsymbol{\mu}_0$ against the alternative that the distribution is symmetric about $\boldsymbol{\mu} \neq \boldsymbol{\mu}_0$. There is a substantial literature for this problem. Under the multivariate normality assumption, it is common to use Hotelling's T^2 test [20]. A multivariate affine-invariant sign test based on counts called interdirections has been presented by Randles [35]. In the sequence, Peter and Randles [41] based on the notion of interdirection, provided affine invariant signed rank test and signed sum test, respectively. Optimal affine invariant tests based on interdirections and pseudo-Mahalanobis ranks have been developed by Hallin and Paindaveine [12]. Hallin and Paindaveine [11] also presented an alternative version of these procedures in which interdirections are replaced by angles between the observations standardized via Tyler's estimator of scatter [40]. Mottonen and Oja [32] developed the tests based on spatial signs and ranks. Hettmansperger et al. [18] and Hettmansperger et al. [19] extended the bivariate tests of Brown and Hettmansperger [5] to the multivariate case. An affine invariant sign test by applying the Tyler's transformation on data points has been presented by Randles [36]. The affine invariant signed rank test, modified from sign test of Randles [36], was suggested by Mahfoud and Randles [29]. The tests described in preceding paragraph can serve as important preliminaries before applying these corresponding location tests. Moreover, there are several tests for testing of the hypothesis that the bivariate distribution is symmetric against not only location parameter but also regression and serial dependence alternatives e.g. Hallin and Paindaveine [13], [14] and [15].

Another problem that has received attention is to test whether the distribution is symmetric about known center $\boldsymbol{\mu}_0$ against the alternative that either the symmetry is lost or the location parameter is changed. Our paper deals with the latter problem. Indeed, the purpose of this paper is to develop affine invariant tests for testing the central symmetry of the bivariate distribution about a known center $\boldsymbol{\mu}_0$. Baringhaus [4] introduced the rotation invariant tests, for testing the spherical symmetry of the multivariate distribution about known center. For central symmetry that it is a weaker assumption than spherical and elliptical symmetry, the tests have been developed employing the empirical characteristic functions by Ghosh and Ruymgaart [10]. Aki [1] proposed a rotation invariant test based on the empirical distribution function. An extension of McWilliams' univariate run test (Mcwilliams [31]) into a test of bivariate central symmetry based on the depth function have been presented by Dyckerhoff et al. [8]. Although, this test is affine invariant, it suffers from low power in distinguishing most of the alternative hypotheses to central symmetry. Recently Einmahl and Gan [9] proposed two versions of a rotation invariant test based on empirical measures of opposite regions.

In this paper, we aim to propose test statistics for central symmetry in such a way that they would be affine invariant, distribution-free and have good power against alternatives to the null hypothesis. The test statistics are created based on sum of the signed-ranks where the sign and rank functions are determined through the depth function. Based on a given depth function, this procedure results in an orthogonal invariant test statistic. An affine invariant version of this test is provided by applying Tyler's transformation (Tyler [40]) on data points. The affine invariance property ensures that the performance of the test does not depend on the underlying coordinate system.

The word of depth has been used for the first time by Tukey [39] to introduce the halfspace depth function. In the sequence, different depth functions have been introduced and the multivariate data have been ordered as center-outward based on them. This center-outward ranking has been widely applied in multivariate nonparametric inference. Liu and Singh [27] presented a quality index and provided some multivariate rank tests for difference between two independent distributions based on it. In the following, a distribution-free test was presented based on both the depth function and the principal components by Rousson [37] for the multivariate two-sample location-scale model. Based on DD plots (depth vs. depth plots) introduced by Liu et al. [26], two tests have been provided by Li and Liu [24] for location difference between two multivariate distributions. In addition, Liu and Singh [28] introduced some rank tests for multivariate scale difference between two or more independent populations. Depth-based run tests for bivariate central symmetry is introduced by Dyckerhoff et al. [8].

The remainder of this paper is organized as follows. In section 2, we review briefly the concept of depth function and ranking based on it. The proposed test statistics will be described in section 3 and the asymptotic properties of those are also investigated. Finally, in section 4, a Monte Carlo study evaluates the finite sample performance of the proposed test statistics in accordance with other tests. All technical proofs are deferred to the Appendix.

2. DEPTH FUNCTION

Let $\mathbf{X}$ be a p-dimensional random vector defined on a probability space $(\Omega, \mathcal{F}, P)$. We denote F as a distribution function corresponding to P. A depth function associated with a distribution function F on $\mathbb{R}^p$ is defined to provide a center-outward ordering of points of $\mathbb{R}^p$ relative to F. Based on depth function, a corresponding notion of center or multidimensional median could be defined. The higher depth values refer to the points near to the center, whereas the lower values refer to the outer points of the center. A formal definition of "statistical depth function" is presented by Zuo and Serfling [42] as a function $D(., F) : \mathbb{R}^p \to \mathbb{R}$ satisfying the following properties:

P1. affine invariance: for any nonsingular $p \times p$ matrix $\mathbf{A}$ and $p-$vector $\mathbf{b}$, $D(\mathbf{Ax} + \mathbf{b}, F_{\mathbf{Ax+b}}) = D(\mathbf{x}, F)$.

P2. Maximality at center: If F is symmetric about $\boldsymbol{\theta}$ in some sense, then $D(\boldsymbol{\theta}, F) = \sup_{\mathbf{x}\in\mathbb{R}^p} D(\mathbf{x}, F)$.

P3. Monotonicity relative to deepest point: if $D(\boldsymbol{\theta}, F) \geq D(\mathbf{x}, F)$ for any $\mathbf{x} \in \mathbb{R}^p$ then $D(\boldsymbol{\theta} + \alpha(\mathbf{x} - \boldsymbol{\theta}), F) \geq D(\mathbf{x}, F)$ for each $\alpha \in [0, 1]$ and $\mathbf{x} \in \mathbb{R}^p$.

P4. Vanishing at infinity: as $||\mathbf{x}|| \to \infty$, $D(\mathbf{x}, F) \to 0$.

Let $\mathbf{X}_1, ..., \mathbf{X}_n$ be a random sample from p-dimensional distribution function F. The sample version of the depth function $D(., F)$ will be obtained by replacing F with the sample distribution F_n.

Remark 2.1. If the sample depth function $D(., F_n)$ satisfies property P1, then it will also be invariant under data-dependent nonsingular transformations.

Different depth functions have been proposed by some authors, which the definition of some of them that we deal with in this paper are given as follows.

Definition 2.1. (Tukey [39]) The halfspace depth of $\mathbf{x} \in \mathbb{R}^p$ with respect to F is defined as

$$HD(\mathbf{x}, F) = \inf_H \{P(H) : \ H \ is \ a \ closed \ halfspace \ in \ \mathbb{R}^p \ and \ \mathbf{x} \in H\}$$

and the sample halfspace depth function is

$$HD(\mathbf{x}, F_n) = \frac{\min_{||\mathbf{u}||=1} \#\left\{i : \quad \mathbf{u}^T\mathbf{X}_i \leq \mathbf{u}^T\mathbf{x}, \quad i = 1, ..., n\right\}}{n}.$$

Definition 2.2. (Liu [25]) The simplicial depth of $\mathbf{x}$ with respect to F is defined as

$$SD(\mathbf{x}, F) = P_F(\mathbf{x} \in S[\mathbf{X}_1, ..., \mathbf{X}_{p+1}])$$

where $S[\ \mathbf{X}_1, ..., \mathbf{X}_{p+1}]$ is a closed simplex with $\mathbf{X}_1, ..., \mathbf{X}_{p+1}$ vertics. The sample version of $SD(\mathbf{x}, F)$ is given by the fraction of the sample random simplices containing the point $\mathbf{x}$.

Definition 2.3. (Liu [27]) The Mahalanobis depth of $\mathbf{x}$ with respect to F is given by

$$MD(\mathbf{x}, F) = \frac{1}{1 + (\mathbf{x} - \boldsymbol{\mu})^T \boldsymbol{\Sigma}^{-1} (\mathbf{x} - \boldsymbol{\mu})}$$

where $\boldsymbol{\mu}$ and $\boldsymbol{\Sigma}$ are the mean vector and dispersion matrix of F distribution, respectively. The sample version of Mahalanobis depth is provided by replacing $\boldsymbol{\mu}$ and $\boldsymbol{\Sigma}$ with their sample estimates.

Additionally, some other depth functions have been introduced such as Oja depth (Oja [34]) and zonoid depth (Koshevoy and Mosler [23]). A more recent proposal for data depth is the Monge-Kantorovich depth (Chernozhukov et al. [7]) based on the Monge-Kantorovich theory of measure transportation.

Now, we present the definition of center-outward ranking of data points.

Definition 2.4. Assume that $\mathbf{X}_1, ..., \mathbf{X}_n$ is a random sample from distribution function F in $\mathbb{R}^p$. The center-outward rank $\mathbf{X}_i$ within the sample $\mathbf{X}_1, ..., \mathbf{X}_n$ is

$$\#\{\mathbf{X}_j \in \{\mathbf{X}_1, ..., \mathbf{X}_n\} : \quad D(\mathbf{X}_j, F_n) \geq D(\mathbf{X}_i, F_n)\}$$

where F_n is the sample distribution function.

Thus, the center-outward ranking is defined in such a way that a larger rank is assigned to a more outlying point w.r.t. $\mathbf{X}_1, ..., \mathbf{X}_n$. If there are no ties, rank 1 and rank n are assigned to the deepest point and the most outlying point, respectively.

3. THE PROPOSED TESTS

Let $\mathbf{X}_1, ..., \mathbf{X}_n$ be independently and identically distributed as $\mathbf{X} = (X_1, X_2)^T$, where $\mathbf{X}$ has an arbitrary bivariate continuous distribution F. The null hypothesis of interest is that, the random vector $\mathbf{X}$ has a distribution centrally symmetric about the known point $\boldsymbol{\mu}_0$. The random vector $\mathbf{X}$ is centrally symmetric around $\boldsymbol{\mu}_0$ provided $\mathbf{X} - \boldsymbol{\mu}_0$ and $\boldsymbol{\mu}_0 - \mathbf{X}$ have the same distribution. Since it is assumed that the symmetry point is known, it is possible to take $\boldsymbol{\mu}_0 = 0$, without loss of generality. So, the hypothesis that the probability distribution is centrally symmetric about $\boldsymbol{\mu}_0$, reduces to the hypothesis $H_0 : \mathbf{X} \stackrel{d}{=} -\mathbf{X}$, where $\stackrel{d}{=}$ denotes "equal in distribution". We now describe the procedure for defining affine invariant tests. Let us look at tests that they are only invariant with respect to orthogonal transformations of the data in subsection 3.1, and then proceed to provide our main affine-invariant tests in subsection 3.2.

3.1. The orthogonal invariant tests

Let $D(., F)$ be a depth function on $\mathbb{R}^2$ associated with a distribution function F. Now, under the given depth function $D(., F)$, we derive a test statistic using depth-based ranks and signs of $\mathbf{X}_1, ..., \mathbf{X}_n$. To define the proposed test statistic, we need to order the points $\mathbf{X}_1, ..., \mathbf{X}_n$ in terms of the evidence they provide against the null hypothesis. To this end, we order the points $\mathbf{X}_1, ..., \mathbf{X}_n$ as center-outward, such that the larger ranks correspond to the closer points to the null symmetry center and the smaller ranks correspond to the outer ones. Let F_n and F_n^s denote the sample distribution function of random sample $\mathbf{X}_1, ..., \mathbf{X}_n$ and the symmetrized sample $(\pm\mathbf{X}_1, ..., \pm\ \mathbf{X}_n)$, respectively. Employing property P2 of the depth function, to obtain center-outward rank of points relative to the null symmetry center instead of the median of $\mathbf{X}_1, ..., \mathbf{X}_n$, the points are ordered based on $D(., F_n^s)$ rather than $D(., F_n)$. More precisely, define

$$R_i = \#\left\{\mathbf{X}_j \in \{\mathbf{X}_1, ..., \mathbf{X}_n\} : \quad D\left(\mathbf{X}_j, F_n^s\right) \geq D\left(\mathbf{X}_i, F_n^s\right)\right\}, \quad i = 1, ..., n. \tag{3.1}$$

If ties occur in this ranking, the ranks within each ties-class have been assigned based on increasing values at the corresponding index set of that. This assignment is allocated to induce invariance property on proposed test statistic.

The test statistic is sum of the signed-ranks of points. The sign of each bivariate point can be determined as the sign of the first or second component of it. Specifically, the sign of a bivariate point is equal to 1 if its first (or second) component is nonnegative and otherwise is equal to -1. This definition of sign, leads to a test statistic which is not only noninvariant, but also it is not able to detect all different types of departures from the null hypothesis. Moreover, The sign of $\mathbf{X}_i$, $i = 1, ..., n$, could be defined as the spatial sign vector $\mathbf{X}_i/||\mathbf{X}_i||$ where $||.||$ denoting the Euclidean norm in $\mathbb{R}^2$. By this definition of sign, the resulted test statistic is not strictly distribution-free. To overcome these limitations, we will determine sign of points based on a data-dependent line passing through the origin instead of the horizontal or vertical axis of the coordinate plane. In what follows, we will describe how to obtain this line.

Let sample median $\mathbf{M}_n$ is a point among $\mathbf{X}_1, ..., \mathbf{X}_n$ with maximum sample depth $D(., F_n^s)$. If there is more than one sample point with the highest depth value $D(., F_n^s)$, $\mathbf{M}_n$ will be the point with minimum index among those data points. Let

$$\theta_{\mathbf{M}_n} = -\arctan\left(\frac{M_{n1}}{M_{n2}}\right) + \frac{\pi}{2}$$

be the angle between the bivariate vector $\mathbf{M}_n = (M_{n1}, M_{n2})^T$ and the horizontal-axis. Note that $\theta_{\mathbf{M}_n} \in [0, \pi)$. Related point $\mathbf{Z}_{ni} = (Z_{ni1}, Z_{ni2})^T$ is given by rotating $\mathbf{X}_i$ counter-clockwise by angle $\frac{\pi}{2} - \theta_{\mathbf{M}_n}$, for all $i = 1, ..., n$. Based on the sample depth function $D(., F_n)$, the proposed test statistic is defined as

$$T_{n,D} = \frac{6}{n(n+1)(2n+1)} \left(\sum_{i=1}^{n} \delta_{ni} R_i\right)^2, \tag{3.2}$$

where R_i is expressed in (3.1) and the random variable δ_{ni} is defined as

$$\delta_{ni} = \begin{cases} 1 & Z_{ni2} \geq 0 \\ -1 & Z_{ni2} < 0 \end{cases} \tag{3.3}$$

for all $i = 1, ..., n$. The large values of the test statistic $T_{n,D}$ reject H_0 in favor of alternative hypothesis.

Note that the sign of bivariate points is determined based on a data-dependent line passing through the origin that is perpendicular to depth based median. Indeed, the reason for restricting to dimension two is that this procedure is employed to divide plane R^2 into two unique halfspaces based on two points (the origin and the depth based median), whereas by this procedure dividing hyperplane R^p, $(p > 2)$ into two unique halfspaces would not be possible.

In what follows, we present the desirable property of orthogonal invariance of $T_{n,D}$ and asymptotic distribution of $T_{n,D}$ under the null hypothesis is developed. The proofs are provided in the Appendix.

Theorem 3.1. *If the sample depth function $D(., F_n)$ satisfies property P1, then the test statistic $T_{n,D}$ will be invariant under orthogonal transformations; that is*

$$T_{n,D}(\mathbf{X}_1, ..., \mathbf{X}_n) = T_{n,D}(\boldsymbol{A}\mathbf{X}_1, ..., \boldsymbol{A}\mathbf{X}_n),$$

for any 2×2 orthogonal matrix $\boldsymbol{A}$.

Theorem 3.2. *If the sample depth function satisfies property P1, then under the null hypothesis of centrally symmetric about 0, $T_{n,D}$ converges in distribution to a chi-square random variable with 1 degree of freedom.*

By applying this theorem, the null hypothesis will be rejected at level α when

$$T_{n,D} \geq \chi^2_{1,1-\alpha}$$

where $\chi^2_{1,1-\alpha}$ denotes the $1 - \alpha$ quantile of the chi-square distribution with 1 degree of freedom.

As mentioned in Theorem 3.2, the asymptotic null distribution of the test statistics presented here is chi-square with one degree of freedom. One would expect, for location alternatives, a chi-square with two degrees of freedom (the dimension of the information matrix for location). It should be remembered that the main object of this paper is proposing several test statistics for testing that the distribution is symmetric about a specified value against the alternative that either the symmetry is lost or the location parameter is changed. Indeed, this alternative is different from location alternatives.

In our proof of Theorem 3.2 we show that R_i's, $i = 1, ..., n$, are identically and uniformly distributed on the set $\{1, 2, ..., n\}$ and δ_{ni}'s, $i = 1, ..., n$ are i.i.d. random variables as distributed independently of R_i and taking the values 1 and -1 each with probability 1/2. These traits immediately imply that under the null hypothesis and the conditions of Theorem 3.2, our test statistic $T_{n,D}$ is strictly distribution-free.

3.2. The affine invariant tests

As shown, Theorem 3.1 indicates that $T_{n,D}$ is orthogonal invariant. In this subsection, we would extend $T_{n,D}$ to be affine invariant, preserving the asymptotic behavior of $T_{n,D}$. To achieve the affine invariant version of the proposed test statistics, we can apply the Tyler's auxiliary transformation (Tyler [40]) on data points. Tyler [40] proposed the data-dependent $p \times p$ scatter matrix $\mathbf{V}_n$, that is a positive definite and symmetric matrix, satisfying $trace(\mathbf{V}_n) = p$ and

$$\frac{1}{n}\sum_{i=1}^{n}\left(\frac{\mathbf{\Gamma}_n\mathbf{X}_i}{||\mathbf{\Gamma}_n\mathbf{X}_i||}\right)\left(\frac{\mathbf{\Gamma}_n\mathbf{X}_i}{||\mathbf{\Gamma}_n\ \mathbf{X}_i||}\right)^T = \frac{1}{p}I_p \tag{3.4}$$

where $\mathbf{X}_i$, $i = 1, ..., n$ is a random vector in $\mathbb{R}^p$, $\mathbf{\Gamma}_n^T\mathbf{\Gamma}_n = \mathbf{V}_n^{-1}$ such that $\mathbf{\Gamma}_n$ is an upper triangular nonsingular matrix with 1 on the first element on the diagonal and $\mathbf{I}_p$ is the p-dimensional identity matrix. This scatter matrix is unique up to multiplication by a positive constant if the sample comes from a continuous p-dimensional distribution and $n > p(p-1)$ (Tyler [40]). An iterative computation scheme has been developed to compute this matrix by Randles [36].

We define $\mathbf{W}_{ni} = \mathbf{\Gamma}_n\mathbf{X}_i$, $i = 1, ..., n$ and $F^s_{\mathbf{w}_n}$ as the sample distribution of the symmetrized sample $(\pm\mathbf{W}_{n1}, ..., \pm\mathbf{W}_{nn})$. Let

$$R_{ni} = \#\left\{\mathbf{W}_{nj} \in \{\mathbf{W}_{n1}, ..., \mathbf{W}_{nn}\} : \quad D\left(\mathbf{W}_{nj}, F^s_{\mathbf{w}_n}\right) \geq D\left(\mathbf{W}_{ni}, F^s_{\mathbf{w}_n}\right)\right\}, \quad i = 1, ..., n, \tag{3.5}$$

and

$$\theta_{\mathbf{M}_{\mathbf{w}_n}} = -\arctan\left(\frac{M_{\mathbf{w}_n 1}}{M_{\mathbf{w}_n 2}}\right) + \frac{\pi}{2}$$

where $\mathbf{M}_{\mathbf{w}_n} = (M_{\mathbf{w}_n 1}, M_{\mathbf{w}_n 2})^T$ refers to the sample median among $\mathbf{W}_{n1}, ..., \mathbf{W}_{nn}$ based on $D(., F^s_{\mathbf{w}_n})$. In the following, points $\mathbf{W}_{n1}, ..., \mathbf{W}_{nn}$ are rotated counter-clockwise by angle $\frac{\pi}{2} - \theta_{\mathbf{M}_{\mathbf{w}_n}}$, which we call them as $\mathbf{V}_{n1}, ..., \mathbf{V}_{nn}$.

Now, based on $D(., F_n)$, the affine invariant test statistic is defined as

$$T^*_{n,D} = \frac{6}{n(n+1)(2n+1)} \left(\sum_{i=1}^{n} \gamma_{ni} R_{ni} \right)^2, \tag{3.6}$$

where γ_{ni} is specified in the same way as δ_{ni}, through $\mathbf{V}_{ni}$ instead of $\mathbf{Z}_{ni}$, for all $i = 1, ..., n$.

It is worth to note that, the test statistic $T^*_{n,D}$ is also distribution-free. The affine invariance property and asymptotic null distribution of $T^*_{n,D}$ are presented in the following Theorems.

Theorem 3.3. *If the sample depth function $D(., F_n)$ satisfies property P1 and $n > 2$, the test statistic $T^*_{n,D}$ will be affine invariant; that is,*

$$T^*_{n,D}\,(\;\boldsymbol{X}_1, ..., \boldsymbol{X}_n) = T^*_{n,D}\,(\boldsymbol{AX}_1, ..., \boldsymbol{AX}_n)$$

for any 2×2 nonsingular matrix $\boldsymbol{A}$.

Theorem 3.4. *If the sample depth function satisfies property P1, then under the null hypothesis of centrally symmetric about 0, $T^*_{n,D}$ converges in distribution to a chi-square random variable with 1 degree of freedom.*

4. Simulation study

In this section, an extensive simulation study is conducted to evaluate the finite sample behavior of the proposed test procedure. Two characteristics of interest are the empirical level and power of the proposed testing procedure. To assess the effects of different depth rankings on the performance of our test statistic, we determined three versions of $T_{n,D}$, derived from the simplicial, halfspace, and Mahalanobis depth functions as $T_{n,SD}$, $T_{n,HD}$ and $T_{n,MD}$, respectively. In the same way, $T^*_{n,SD}$, $T^*_{n,HD}$, $T^*_{n,MD}$ will be defined corresponding to $T^*_{n,D}$. The performance of our test statistics is compared with the affine invariant run test based on the simplicial depth function that we refer to $R_{n,SD}$ hereafter (Dyckerhoff et al. [8]) and the two rotation invariant tests Q^1_n and Q^2_n proposed by Einmahl and Gan [9]. Q^1_n refers to their main test, and Q^2_n is given by Q^1_n adding a weight function to it (we avoid presenting the details of these test statistics).

To illustrate the effect of the sample size on the finite sample behavior of our proposed test statistics, we set the sample sizes as $n = 100$ and 200. Moreover, the nominal level was set at 0.05 throughout. In each setting, 2000 independent random samples were generated to calculate the proportion of replications for which the null hypothesis is rejected. To examine the finite sample behavior of test statistics under the null and alternative hypotheses, we have simulated samples from several bivariate distribution families, including Azzalini's skew-normal distribution (Azzalini and Dalla Valle [3]), Azzalini's skew-t distribution (Azzalini and Capitanio [2]), perturbed symmetric beta distribution (Azzalini and Capitanio [2]) and sinh-arcsinh distribution (Jones and Pewsey [21]). Indeed, we consider different types of skewness over very light-tailed distributions to very heavy-tailed ones. In what follows, we provide an overview of these families.

- *Bivariate skew-normal distribution*: Let $\mathbf{X}$ be defined as

$$\mathbf{X} = \begin{cases} \mathbf{Y} & if \quad Z > \Delta^T \mathbf{Y}, \\ -\mathbf{Y} & if \quad Z \leq \Delta^T \mathbf{Y}, \end{cases}$$

  where $\mathbf{Y} \sim N_2(0, \mathbf{\Sigma})$, $\mathbf{\Delta} = (\Delta_1, \Delta_2)^T$ is the shape parameter, and Z is distributed independently of $\mathbf{Y}$ according to $N(0,1)$. The random vector $\mathbf{X}$ is known as bivariate skew-normal random vector and it may be written as $\mathbf{X} \sim SN_2(0, \Sigma, \Delta)$.

- *Bivariate skew-t distribution*: Let $\mathbf{T} = V^{-\frac{1}{2}}\mathbf{X}$, where the random vector $\mathbf{X}$ follows the distribution $SN_2(0, \mathbf{\Sigma}, \mathbf{\Delta})$ and νV is distributed independently of $\mathbf{X}$ according to a chi-squared distribution with ν degrees of freedom. We will say that $\mathbf{T}$ has a bivariate skew-t distribution and write $\mathbf{T} \sim ST_2(0, \mathbf{\Sigma}, \mathbf{\Delta}, \nu)$

- *Bivariate perturbed symmetric beta distribution:* Let $\mathbf{Y} = (2B_1 - 1, 2B_2 - 1)^T$, where B_1 and B_2 have beta distributions $B(a,a)$ and $B(b,b)$, respectively. The random vector $\mathbf{Y}$ can be treated as a central and non-elliptical symmetric random vector. Define the random vector $\mathbf{X}$ as

$$\mathbf{X} = \begin{cases} \mathbf{Y} & if \quad Z < w(\mathbf{Y}), \\ -\mathbf{Y} & if \quad Z > w(\mathbf{Y}), \end{cases}$$

  where Z (independently of $\mathbf{Y}$) has distribution function $G(.)$. The distribution function $G(.)$ and function $w(.)$ are given as

$$G(z) = \frac{e^z}{1+e^z} \qquad and \qquad w(\mathbf{y}) = \frac{sin(p_1 y_1 + p_2 y_2)}{1 + cos(q_1 y_1 + q_2 y_2)},$$

  where p_1, p_2, q_1 and q_2 are additional parameters. Then, we will say that $\mathbf{X}$ has a perturbed symmetric beta distribution.

- *Bivariate sinh-arcsinh distribution:* This family is generated by sinh-arcsinh transformation on a primary symmetric distribution. We consider the bivariate normal distribution as the primary distribution. The desirable property of this transformation is to induce skewness on the primary distribution and distributions with heavier/lighter tails than the primary one. Suppose random vector $\mathbf{Z} = (Z_1, Z_2)^T$ follows $N_2(0, \mathbf{\Sigma})$. Define the bivariate vector $\mathbf{X} = (X_1, X_2)^T$ as

$$X_j = \sinh\left[\tfrac{1}{\delta_j}(\sinh^{-1}(Z_j) + \Delta_j)\right], \qquad j = 1, 2, \tag{4.1}$$

where Δ_j and δ_j denote the measure of skewness and tail weight in direction of jth component of $\mathbf{Z}$, respectively. Amount of skewness increases with increasing positive Δ_j or decreasing negative Δ_j. Additionally, distributions with heavier and lighter tails than the bivariate normal distribution are generated by taking $0 < \delta_j < 1$ and $\delta_j > 1$, respectively.

In this study, we generate samples from the aforementioned distribution families with $\mathbf{\Sigma} = (1-\rho)\mathbf{I}_2 + \rho\mathbf{J}_2$ with $\rho = -0.5, 0$ and 0.5, and $\mathbf{J}_2$ denoting the 2×2 matrix with all entries equal 1 and $\Delta_i = k\eta$, $i = 1, 2$ with $\boldsymbol{\eta} = (0.15, 0.15)^T$ and $k = 0, 1, 2$ and 3. We consider $\nu = 1, 3, 6, 10$ and 20 for bivariate skew-t distribution and $\delta_i = 0.5, 0.75, 1, 2$ and 5, $i = 1, 2$ for bivariate sinh-arcsinh distribution.

Table 1 and Figures 1 and 2 provide the empirical rejection probabilities for sample size $n = 100$ and for bivariate skew-normal, skew-t and sinh-arcsinh distribution, respectively. Inspection of the table and figures confirms that the performance of our test statistics is not affected by different depth ranking. In all of them, the empirical rejection probabilities corresponding to $k = 0$ represents the proportion of rejection under the null hypothesis. These results demonstrate that all the tests would be accurate in estimating the nominal level, except $R_{n,SD}$ which it has been underestimated in some cases. Since the performance of test statistics, even affine invariant test statistics are affected by correlation structure of primary distribution, we provide three possibilities for ρ as -0.5, 0 and 0.5. From the represented results in Table 1 and Figure 1, it is obvious that all empirical powers will be increased by increasing the value of ρ for bivariate skew-normal and skew-t distributions. This situation is reversed for bivariate sinh-arcsinh distribution in Figure 2 except for $T_{n,D}$.

Table 1 shows that $T^*_{n,D}$ outperforms $R_{n,SD}$ and Q^2_n in all cases, performs virtually as well as Q^1_n for $k = 1$ and 2 and has slightly lower power than Q^1_n for $k = 3$. Moreover, $T_{n,D}$ outperforms $R_{n,SD}$ in all cases, Q^2_n when $\rho = 0$ and 0.5 and Q^1_n when $\rho = 0.5$. Figure 1 indicates that $T^*_{n,D}$ and $T_{n,D}$ outperform $R_{n,SD}$ for all values of ν and ρ. In addition $T^*_{n,D}$ has higher power than Q^2_n except when $\nu = 1$, and $T_{n,D}$ overcomes Q^2_n except when $\nu = 1$ and $\rho = -0.5$. In comparison on Q^1_n, $T^*_{n,D}$ performs better when $k = 1, 2$, $v = 6, 20$ and all values of ρ, and $T_{n,D}$ performs better when $k = 1$ and 2, $v = 6, 20$ and $\rho = 0$ and 0.5. Indeed, the empirical power of our tests increases as degrees of freedom increases. In Figure 2, superiority of our affine invariant tests is clear in most cases especially for $\rho = 0.5$ and $k = 1$ and 2.

Tables 2 provide the empirical rejection probabilities for sample size $n = 200$ and for bivariate skew-normal distribution. In Figures 3 and 4, we plot the empirical rejection probabilities against k corresponding to some values of parameters of the same populations and tests with Figures 1 and 2 respectively, for sample size $n = 200$. Note that, as expected, the empirical powers increase with the sample size. These simulations lead to almost the same conclusions as in $n = 100$.

These simulations demonstrate that our tests are more powerful for small and moderate departures from the null hypothesis and for light-tailed distributions. As expected, the performance of affine invariant tests $T^*_{n,D}$ is less affected by changing the value of ρ rather than the orthogonal invariant tests $T_{n,D}$. The results show that, compared to $T^*_{n,D}$ test, $T_{n,D}$ performs better when $\rho = 0.5$, is comparable when $\rho = 0$ and performs worse when $\rho = -0.5$.

Finally, to complete our simulations, for sample sizes $n = 100$ and 200, we generate samples from bivariate perturbed symmetric beta distribution with several choices of the parameters such that different situations of asymmetry can be considered. A thorough investigation of Table 3 and 4 indicated that our tests overcome $R_{n,SD}$ and Q^2_n in all cases and Q^1_n in some cases.

Table 1: Empirical rejection probabilities (out of 2000 replications) for bivariate skew-normal distribution with $n = 100$, $\rho = -0.5, 0$ and 0.5, and $\Delta_i = k\eta$, $i = 1,2$ with $\eta = (0.15, 0.15)^T$ and $k = 0, 1, 2, 3$.

	$\rho = -0.5$				$\rho = 0$				$\rho = 0.5$			
test	k=0	k=1	k=2	k=3	k=0	k=1	k=2	k=3	k=0	k=1	k=2	k=3
$T^*_{n,SD}$	0.046	0.105	0.294	0.520	0.046	0.175	0.488	0.694	0.046	0.245	0.607	0.763
$T^*_{n,HD}$	0.047	0.107	0.295	0.518	0.047	0.169	0.484	0.692	0.047	0.244	0.602	0.766
$T^*_{n,MD}$	0.048	0.103	0.291	0.511	0.048	0.177	0.486	0.688	0.048	0.241	0.604	0.760
$T_{n,SD}$	0.047	0.072	0.175	0.282	0.043	0.171	0.469	0.658	0.047	0.336	0.779	0.877
$T_{n,HD}$	0.044	0.074	0.169	0.280	0.044	0.171	0.462	0.644	0.044	0.340	0.779	0.881
$T_{n,MD}$	0.049	0.076	0.172	0.282	0.048	0.172	0.466	0.648	0.049	0.337	0.774	0.876
$R_{n,SD}$	0.043	0.048	0.081	0.159	0.043	0.057	0.135	0.320	0.043	0.070	0.190	0.460
Q^1_n	0.049	0.098	0.269	0.544	0.050	0.155	0.439	0.783	0.049	0.182	0.572	0.883
Q^2_n	0.054	0.080	0.173	0.355	0.048	0.098	0.250	0.507	0.053	0.127	0.342	0.657

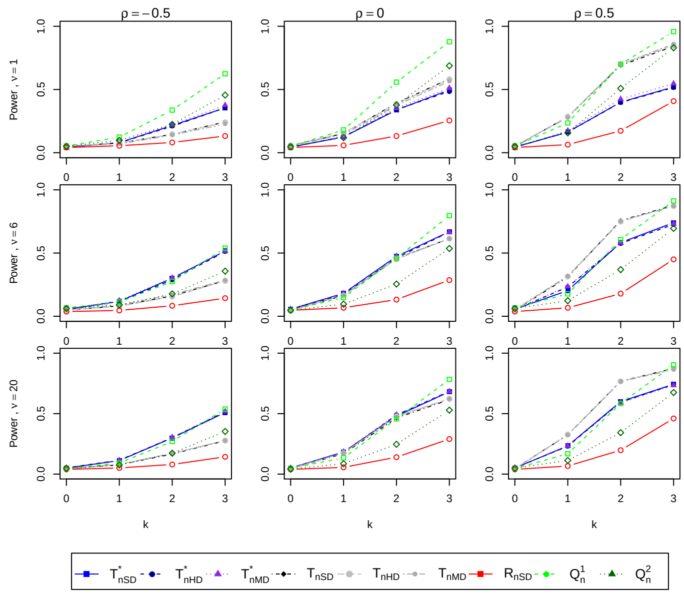

Figure 1: Empirical rejection probabilities (out of 2000 replications) for bivariate skew-t distribution with $n = 100$, $\rho = -0.5, 0$ and 0.5, $\nu = 1, 6, 20$ and $\Delta_i = k\eta$, $i = 1, 2$ with $\eta = (0.15, 0.15)^T$ and $k = 0, 1, 2, 3$.

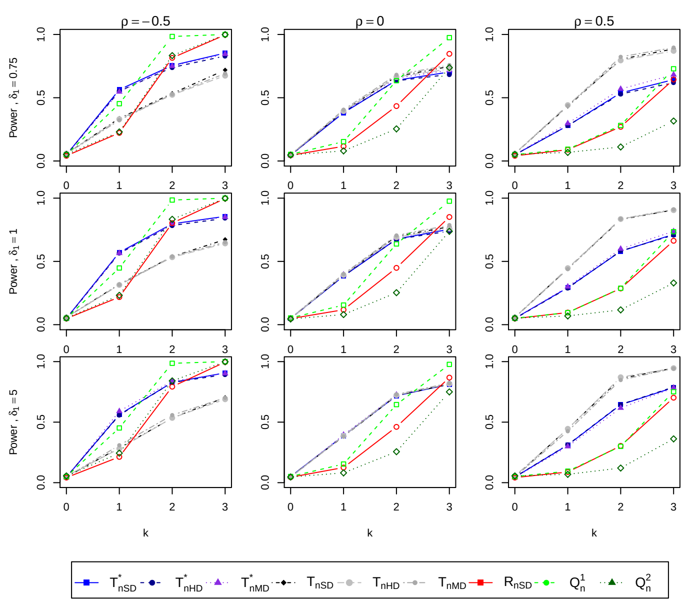

Figure 2: Empirical rejection probabilities (out of 2000 replications) for bivariate sinh-arcsinh distribution with $n = 100$, $\rho = -0.5, 0$ and 0.5, $\delta_i = 0.75, 1, 5$ and $\Delta_i = k\eta$, $i = 1, 2$ with $\eta = (0.15, 0.15)^T$ and $k = 0, 1, 2, 3$.

Table 2: Empirical rejection probabilities (out of 2000 replications) for bivariate skew-normal distribution with $n = 200$, $\rho = -0.5, 0$ and 0.5, and $\Delta_i = k\eta$, $i = 1, 2$ with $\eta = (0.15, 0.15)^T$ and $k = 0, 1, 2, 3$.

	$\rho = -0.5$				$\rho = 0$				$\rho = 0.5$			
test	k=0	k=1	k=2	k=3	k=0	k=1	k=2	k=3	k=0	k=1	k=2	k=3
$T^*_{n,SD}$	0.052	0.204	0.486	0.690	0.052	0.317	0.681	0.813	0.052	0.414	0.747	0.834
$T^*_{n,HD}$	0.047	0.200	0.486	0.692	0.047	0.313	0.678	0.808	0.047	0.419	0.740	0.828
$T^*_{n,MD}$	0.052	0.206	0.488	0.696	0.052	0.319	0.680	0.804	0.052	0.420	0.748	0.832
$T_{n,SD}$	0.050	0.122	0.276	0.387	0.053	0.311	0.652	0.776	0.050	0.554	0.888	0.921
$T_{n,HD}$	0.048	0.121	0.283	0.391	0.050	0.305	0.645	0.770	0.048	0.560	0.879	0.918
$T_{n,MD}$	0.052	0.125	0.279	0.391	0.051	0.313	0.645	0.769	0.052	0.559	0.886	0.920
$R_{n,SD}$	0.049	0.061	0.117	0.250	0.049	0.076	0.218	0.519	0.049	0.093	0.347	0.754
Q^1_n	0.054	0.162	0.506	0.850	0.053	0.257	0.766	0.984	0.054	0.318	0.875	0.998
Q^2_n	0.056	0.130	0.342	0.665	0.053	0.156	0.500	0.856	0.056	0.212	0.647	0.942

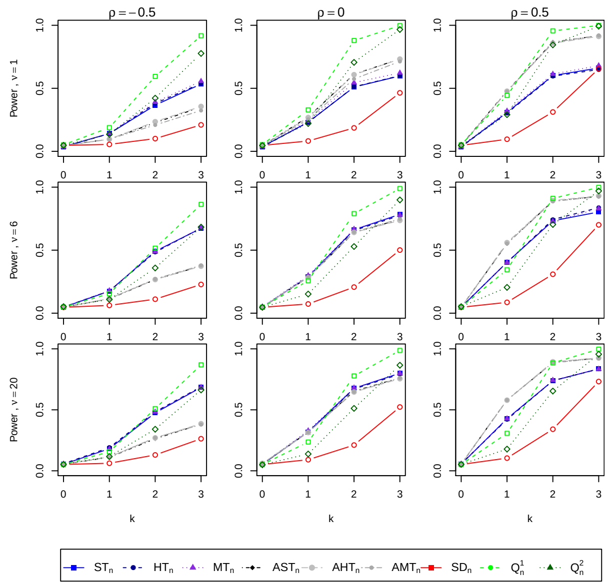

Figure 3: Empirical rejection probabilities (out of 2000 replications) for bivariate skew-t distribution with $n = 200$, $\rho = -0.5, 0$ and 0.5, $\nu = 1, 6, 20$ and $\Delta_i = k\eta$, $i = 1, 2$ with $\eta = (0.15, 0.15)^T$ and $k = 0, 1, 2, 3$.

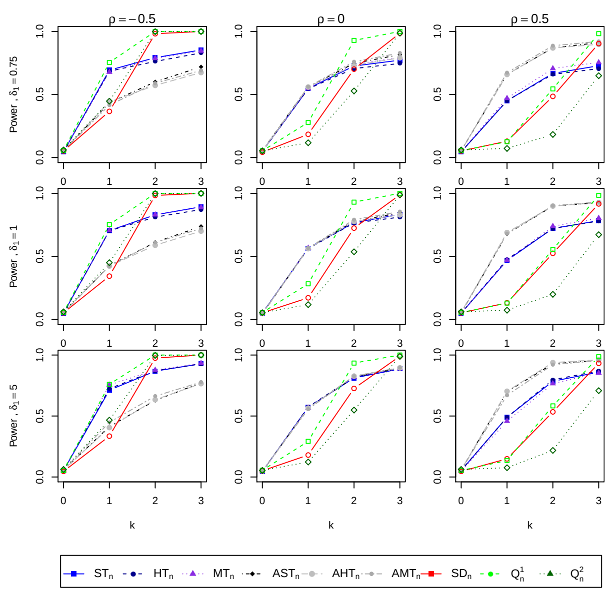

Figure 4: Empirical rejection probabilities (out of 2000 replications) for bivariate sinh-arcsinh distribution with $n = 200$, $\rho = -0.5, 0$ and 0.5, $\delta_i = 0.75, 1, 5$ and $\Delta_i = k\eta$, $i = 1, 2$ with $\eta = (0.15, 0.15)^T$ and $k = 0, 1, 2, 3$.

Table 3: Empirical rejection probabilities (out of 2000 replications) for bivariate perturbed symmetric beta distribution with $n = 100$, a and $b = 0.5, 1$ and 3, $p_1 = q_1 = 1$ and p_2 and $q_2 = 0.5, 1$ and 2

			test								
a, b	p_2	q_2	$T^*_{n,SD}$	$T^*_{n,HD}$	$T^*_{n,MD}$	$T_{n,SD}$	$T_{n,HD}$	$T_{n,MD}$	$R_{n,SD}$	Q^1_n	Q^2_n
	2	0.5	0.194	0.194	0.194	0.175	0.184	0.175	0.060	0.170	0.112
3,3	1	1	0.147	0.149	0.147	0.134	0.145	0.137	0.048	0.110	0.078
	0.5	2	0.165	0.168	0.164	0.147	0.158	0.154	0.051	0.098	0.073
	2	0.5	0.238	0.237	0.205	0.355	0.349	0.332	0.069	0.319	0.232
3,0.5	1	1	0.284	0.290	0.270	0.364	0.351	0.340	0.077	0.265	0.177
	0.5	2	0.598	0.604	0.619	0.636	0.632	0.618	0.163	0.415	0.216
	2	0.5	0.241	0.244	0.244	0.221	0.215	0.222	0.076	0.320	0.223
1,1	1	1	0.305	0.315	0.323	0.300	0.297	0.306	0.077	0.277	0.165
	0.5	2	0.581	0.593	0.606	0.558	0.573	0.585	0.165	0.397	0.201
	2	0.5	0.222	0.214	0.218	0.200	0.192	0.191	0.069	0.381	0.301
0.5,0.5	1	1	0.470	0.474	0.486	0.445	0.453	0.464	0.113	0.505	0.324
	0.5	2	0.815	0.838	0.869	0.779	0.807	0.833	0.353	0.835	0.563

Table 4: Empirical rejection probabilities (out of 2000 replications) for bivariate perturbed symmetric beta distribution with $n = 200$, a and $b = 0.5, 1$ and 3, $p_1 = q_1 = 1$ and p_2 and $q_2 = 0.5, 1$ and 2

			test								
a, b	p_2	q_2	$T^*_{n,SD}$	$T^*_{n,HD}$	$T^*_{n,MD}$	$T_{n,SD}$	$T_{n,HD}$	$T_{n,MD}$	$R_{n,SD}$	Q^1_n	Q^2_n
	2	0.5	0.310	0.303	0.310	0.300	0.299	0.300	0.087	0.329	0.204
3,3	1	1	0.252	0.250	0.250	0.235	0.237	0.233	0.083	0.194	0.120
	0.5	2	0.279	0.278	0.279	0.268	0.265	0.266	0.089	0.173	0.098
	2	0.5	0.399	0.399	0.338	0.624	0.619	0.595	0.111	0.611	0.471
3,0.5	1	1	0.469	0.472	0.444	0.631	0.627	0.608	0.112	0.491	0.310
	0.5	2	0.791	0.810	0.808	0.896	0.893	0.875	0.283	0.729	0.400
	2	0.5	0.370	0.369	0.380	0.364	0.363	0.365	0.105	0.578	0.411
1,1	1	1	0.509	0.520	0.531	0.499	0.499	0.505	0.117	0.512	0.315
	0.5	2	0.793	0.816	0.819	0.788	0.801	0.812	0.262	0.694	0.378
	2	0.5	0.352	0.350	0.352	325	0.336	0.339	0.101	0.656	0.559
0.5,0.5	1	1	0.700	0.738	0.747	0.686	0.725	0.733	0.175	0.804	0.592
	0.5	2	0.916	0.935	0.931	0.913	0.934	0.943	0.594	0.993	0.891

A second Monte Carlo study is provided in order to evaluate the performance of our tests for pure location alternatives. In this study the performance of tests considered in first Monte Carlo study compared with the Hotelling's T^2 and the tests due to Hallin and Paindaveine [12] computed with the sign score function, van der Waerden score function and Wilcoxson score function and denoted by HS_n, HN_n and HR_n, respectively.

We set the sample size as $n = 50$. In each setting, 2000 independent random samples were generated to calculate the proportion of replications for which the null hypothesis is rejected. For each replication, the all tests were performed at the significance level $\alpha = 0.05$. To examine the finite sample behavior of test statistics under the null and alternative hypotheses, we have simulated samples from the t family of distributions and the exponential power family of distributions. In what follows, we provide an overview of these families.

A p-dimensional random vector $\mathbf{X}$ has a multivariate t distribution with ν degree of freedom if its density function has the form

$$f_{\boldsymbol{\mu},\boldsymbol{\Sigma}}(\mathbf{x}) = \frac{\Gamma\left((p+\upsilon)/2\right)}{\Gamma\left(\upsilon/2\right)\left(\pi\upsilon\right)^{p/2}}|\boldsymbol{\Sigma}|^{-1/2}\left[1+\frac{1}{\upsilon}\left(\mathbf{x}-\boldsymbol{\mu}\right)^T\boldsymbol{\Sigma}^{-1}\left(\mathbf{x}-\boldsymbol{\mu}\right)\right]^{-(p+\upsilon)/2}$$

where $\boldsymbol{\mu} = (\mu_1, ..., \mu_p)^T \in \mathbb{R}^p$ and $\boldsymbol{\Sigma}$ is a symmetric $p \times p$ positive definite matrix.

The density function of a p-dimensional random vector $\mathbf{X}$ from the exponential power family of distributions is

$$f_{\boldsymbol{\mu},\boldsymbol{\Sigma}}(\mathbf{x}) = \frac{\upsilon\Gamma\left(p/2\right)}{\Gamma\left(p+2\upsilon\right)\left(\pi c_0\right)^{p/2}}|\boldsymbol{\Sigma}|^{-1/2}\exp\left\{-\frac{\left(\mathbf{x}-\boldsymbol{\mu}\right)^T\boldsymbol{\Sigma}^{-1}\left(\mathbf{x}-\boldsymbol{\mu}\right)}{c_0}\right\}^{\upsilon}$$

where

$$c_0 = \frac{p\Gamma\left(p/2\upsilon\right)}{\Gamma\left((p+2)/2\upsilon\right)},$$

and $\boldsymbol{\mu}$ and $\boldsymbol{\Sigma}$ are defined as above.

We generate samples from the aforementioned distribution families with $\boldsymbol{\Sigma} = \mathbf{I}$ and $\boldsymbol{\mu} = k\boldsymbol{\Delta}$ with $\boldsymbol{\Delta} = (0.2, 0.2)^T$ and $\boldsymbol{\Delta} = (0.1, 0.1)^T$ for the t family of distributions and the exponential power family of distributions, respectively and $k = 0, 1, 2, 3$. We consider $\nu = 1, 6$ and 10 for t-distribution family and $\nu = 0.5, 1$ and 2 for the exponential power family of distributions.

Inspection of Tables 5 and 6 demonstrated that the performance of our tests is comparable to the other tests. The proposed tests overcome $R_{n,SD}$ and Q_n^2 in most cases and Q_n^1 in some cases. It worth to note that all tests which are defined in the similar way of our proposed test e.g. $R_{n,SD}$, Q_n^1 and Q_n^2 are not expected to perform as well as T^2, HS_n, HN_n and HR_n. In other hand, the results confirm that the performance of our test statistics is not affected by different depth ranking.

Table 5: Empirical rejection probabilities (out of 2000 replications) for bivariate t distribution with $n = 50$, $\mathbf{\Sigma} = \mathbf{I}$, $\nu = 1, 6, 20$ and $\boldsymbol{\mu} = k\mathbf{\Delta}$ with $\mathbf{\Delta} = (0.2, 0.2)^T$ and $k = 0, 1, 2, 3$.

	$\nu = 1$				$\nu = 6$				$\nu = 20$			
test	k=0	k=1	k=2	k=3	k=0	k=1	k=2	k=3	k=0	k=1	k=2	k=3
T^2	0.015	0.033	0.084	0.164	0.058	0.287	0.809	0.987	0.056	0.374	0.911	1
HS_n	0.050	0.203	0.634	0.902	0.053	0.285	0.805	0.993	0.054	0.322	0.851	1
HN_n	0.040	0.122	0.351	0.622	0.048	0.270	0.801	0.988	0.049	0.345	0.894	1
HR_n	0.044	0.120	0.343	0.599	0.049	0.279	0.804	0.988	0.057	0.354	0.901	1
$T^*_{n,SD}$	0.042	0.085	0.150	0.431	0.050	0.212	0.542	0.757	0.056	0.236	0.601	0.7715
$T^*_{n,HD}$	0.040	0.083	0.146	0.425	0.053	0.204	0.541	0.750	0.056	0.223	0.592	0.768
$T^*_{n,MD}$	0.039	0.093	0.156	0.432	0.051	0.205	0.526	0.724	0.060	0.234	0.593	0.762
$T_{n,SD}$	0.052	0.100	0.171	0.429	0.046	0.188	0.523	0.749	0.047	0.219	0.570	0.762
$T_{n,HD}$	0.050	0.104	0.173	0.435	0.047	0.182	0.520	0.746	0.048	0.216	0.575	0.761
$T_{n,MD}$	0.051	0.090	0.155	0.389	0.045	0.183	0.510	0.722	0.048	0.223	0.567	0.749
$R_{n,SD}$	0.038	0.073	0.105	0.342	0.040	0.085	0.229	0.569	0.040	0.090	0.287	0.650
Q^1_n	0.057	0.223	0.450	0.937	0.066	0.203	0.671	0.945	0.058	0.197	0.661	0.946
Q^2_n	0.039	0.160	0.338	0.862	0.049	0.121	0.412	0.791	0.038	0.106	0.386	0.787

Table 6: Empirical rejection probabilities (out of 2000 replications) for bivariate power family of distributions with $n = 50$, $\mathbf{\Sigma} = \mathbf{I}$, $\nu = 0.5, 1, 2$ and $\boldsymbol{\mu} = k\mathbf{\Delta}$, with $\mathbf{\Delta} = (0.1, 0.1)^T$ and $k = 0, 1, 2, 3$.

	$\nu = 0.5$				$\nu = 1$				$\nu = 2$			
test	k=0	k=1	k=2	k=3	k=0	k=1	k=2	k=3	k=0	k=1	k=2	k=3
T^2	0.044	0.161	0.515	0.867	0.048	0.392	0.940	1	0.043	0.468	0.981	1
HS_n	0.043	0.203	0.633	0.930	0.044	0.311	0.873	1	0.047	0.317	0.868	0.999
HN_n	0.036	0.158	0.528	0.864	0.038	0.348	0.917	1	0.036	0.447	0.978	1
HR_n	0.040	0.155	0.519	0.849	0.044	0.361	0.922	1	0.038	0.467	0.982	1
$T^*_{n,SD}$	0.053	0.127	0.355	0.601	0.054	0.241	0.619	0.782	0.048	0.282	0.669	0.793
$T^*_{n,HD}$	0.056	0.133	0.363	0.605	0.057	0.240	0.612	0.775	0.051	0.282	0.671	0.779
$T^*_{n,MD}$	0.054	0.130	0.349	0.570	0.055	0.248	0.618	0.761	0.044	0.289	0.685	0.793
$T_{n,SD}$	0.049	0.118	0.344	0.583	0.047	0.214	0.605	0.777	0.036	0.273	0.646	0.784
$T_{n,HD}$	0.046	0.120	0.343	0.596	0.046	0.210	0.604	0.772	0.041	0.268	0.648	0.772
$T_{n,MD}$	0.048	0.123	0.335	0.565	0.048	0.222	0.600	0.758	0.035	0.273	0.660	0.786
$R_{n,SD}$	0.044	0.057	0.133	0.322	0.038	0.095	0.288	0.676	0.042	0.090	0.362	0.763
Q^1_n	0.049	0.179	0.524	0.854	0.045	0.185	0.640	0.949	0.059	0.141	0.533	0.935
Q^2_n	0.035	0.118	0.366	0.670	0.037	0.099	0.350	0.763	0.041	0.070	0.233	0.683

5. Conclusion

This paper concerns with the problem of detecting central symmetry of a bivariate distribution. To this end, based on depth function, we introduced a family of signed-rank test which is orthogonal invariant and distribution-free. Affine invariant tests were obtained by applying our proposed test to the standardized data with Tyler's matrix. The proposed orthogonal and affine invariant tests have the same asymptotic properties. In simulation study, the finite sample behavior of the proposed test procedure was evaluated over distributions family from very light to very heavy-tailed distributions with different kinds of skewness. The simulations confirmed that our affine invariant tests successfully can distinguish different asymmetries and shifting the location parameter. Moreover, we observed that they performed as good as their competitors and actually in many cases they even outperform them.

6. Appendix

Proof of Theorem 3.1: According to the construction of $\mathbf{Z}_{ni}$, it is clear that $\mathbf{Z}_{ni} = \mathbf{B}_{\mathbf{X}_n}\mathbf{X}_i$, $i = 1, ..., n$ where

$$\mathbf{B}_{\mathbf{X}_n} = \begin{bmatrix} \cos\left(\frac{\pi}{2} - \theta_{\mathbf{M}_n}\right) & -\sin\left(\frac{\pi}{2} - \theta_{\mathbf{M}_n}\right) \\ \sin\left(\frac{\pi}{2} - \theta_{\mathbf{M}_n}\right) & \cos\left(\frac{\pi}{2} - \theta_{\mathbf{M}_n}\right) \end{bmatrix}. \tag{6.1}$$

Let $\mathbf{A}$ be an arbitrary 2×2 orthogonal matrix. Define $\tilde{\mathbf{Z}}_{ni} = \mathbf{B}_{\mathbf{AX}_n}\mathbf{AX}_i$ for all $i = 1, ..., n$, where

$$\mathbf{B}_{\mathbf{AX}_n} = \begin{bmatrix} \cos\left(\frac{\pi}{2} - \theta_{\tilde{\mathbf{M}}_n}\right) & -\sin\left(\frac{\pi}{2} - \theta_{\tilde{\mathbf{M}}_n}\right) \\ \sin\left(\frac{\pi}{2} - \theta_{\tilde{\mathbf{M}}_n}\right) & \cos\left(\frac{\pi}{2} - \theta_{\tilde{\mathbf{M}}_n}\right) \end{bmatrix},$$

with $\theta_{\tilde{\mathbf{M}}_n} \in [0, \pi)$, as the angle between horizontal-axis and the sample median $\tilde{\mathbf{M}}_n$ that is obtained in the same way as $\mathbf{M}_n$, through $\mathbf{AX}_i$'s instead of $\mathbf{X}_i$'s, $i = 1, ..., n$. The orthogonality of matrix $\mathbf{A}$ implies that there exists an angle $\alpha \in [0, 2\pi)$ such that

$$\mathbf{A} = \begin{bmatrix} \cos(\alpha) & -\sin(\alpha) \\ \sin(\alpha) & \cos(\alpha) \end{bmatrix} \quad \text{or} \quad \mathbf{A} = \begin{bmatrix} \cos(\alpha) & \sin(\alpha) \\ \sin(\alpha) & -\cos(\alpha) \end{bmatrix}. \tag{6.2}$$

Property P1 of the sample depth function shows that

$$\tilde{\mathbf{M}}_n = \mathbf{AM}_n. \tag{6.3}$$

Let matrix $\mathbf{A}$ be defined as the left side of (6.2), then (6.3) results in

$$\theta_{\tilde{\mathbf{M}}_n} = \begin{cases} \alpha + \theta_{\mathbf{M}_n} & 0 \le \alpha + \theta_{\mathbf{M}_n} < \pi, \\ \alpha + \theta_{\mathbf{M}_n} - \pi & \pi \le \alpha + \theta_{\mathbf{M}_n} < 2\pi, \\ \alpha + \theta_{\mathbf{M}_n} - 2\pi & 2\pi \le \alpha + \theta_{\mathbf{M}_n} < 3\pi. \end{cases}$$

Using the trigonometric relationships, it is straightforward to verify that $\mathbf{B}_{\mathbf{AX}_n}\mathbf{A} = \mathbf{B}_{\mathbf{X}_n}$, or $\mathbf{B}_{\mathbf{AX}_n}\mathbf{A} = -\mathbf{B}_{\mathbf{X}_n}$. Thus

$$\tilde{\mathbf{Z}}_{ni} = \mathbf{Z}_{ni} \qquad or \qquad \tilde{\mathbf{Z}}_{ni} = -\mathbf{Z}_{ni}, \quad i = 1, ..., n. \tag{6.4}$$

Now, let matrix $\mathbf{A}$ be according to the right side of (6.2), similarly we have

$$\theta_{\tilde{\mathbf{M}}_n} = \begin{cases} \alpha - \theta_{\mathbf{M}_n} + \pi & -\pi < \alpha - \theta_{\mathbf{M}_n} < 0, \\ \alpha - \theta_{\mathbf{M}_n} & 0 \le \alpha - \theta_{\mathbf{M}_n} < \pi, \\ \alpha - \theta_{\mathbf{M}_n} - \pi & \pi \le \alpha - \theta_{\mathbf{M}_n} < 2\pi, \end{cases}$$

and

$$\mathbf{B}_{\mathbf{AX}_n}\mathbf{A} = \begin{bmatrix} -1 & 0 \\ 0 & 1 \end{bmatrix} \mathbf{B}_{\mathbf{X}_n} \qquad or \qquad \mathbf{B}_{\mathbf{AX}_n}\mathbf{A} = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} \mathbf{B}_{\mathbf{X}_n}.$$

Hence

$$\tilde{\mathbf{Z}}_{ni} = (-\mathbf{Z}_{ni1}, \mathbf{Z}_{ni2})^T \qquad or \qquad \tilde{\mathbf{Z}}_{ni} = (\mathbf{Z}_{ni1}, -\mathbf{Z}_{ni2})^T, \qquad i = 1, ..., n. \tag{6.5}$$

The proof of affine invariance of $T_{n,D}$ will be completed by using (6.4), (6.5) and property P1 of the sample depth function.

Proof of Theorem 3.2: Under the null hypothesis, $\mathbf{X}_1, ..., \mathbf{X}_n$ are i.i.d. from F, where $F(.)$ is centrally symmetric distribution about the origin. Hence, we have

$$(\mathbf{X}_1, ..., \mathbf{X}_n) \stackrel{d}{=} (\eta_1 \mathbf{X}_1, ..., \eta_n \mathbf{X}_n). \tag{6.6}$$

where η_i's, $i = 1, ..., n$ are i.i.d. random variables taking the values 1 and -1 each with probability 1/2. It is clear that

$$(\pm\mathbf{X}_1, ..., \pm\mathbf{X}_n) = (\pm\eta_1\mathbf{X}_1, ..., \pm\eta_n\mathbf{X}_n). \tag{6.7}$$

Additionally, $\mathbf{M}_n \equiv \mathbf{M}(\mathbf{X}_1, ..., \mathbf{X}_n)$ is considered as a point with maximum sample depth with respect to the symmetrized sample $(\pm\mathbf{X}_1, ..., \pm\mathbf{X}_n)$, (if there is more than one sample point with the highest depth value, $\mathbf{M}_n$ will be defined as the point with minimum index among those data points). By this definition of $\mathbf{M}_n$, there exists $i \in \{1, ..., n\}$ that

$$\mathbf{M}(\mathbf{X}_1, ..., \mathbf{X}_n) = \mathbf{X}_i. \tag{6.8}$$

From property $P1$ and equation (6.7),

$$\mathbf{M}(\eta_1\mathbf{X}_1, ..., \eta_n\mathbf{X}_n) = \eta_i\mathbf{X}_i \tag{6.9}$$

Hence from (6.8) and (6.9), we have

$$\mathbf{M}(\mathbf{X}_1, ..., \mathbf{X}_n) = \eta\mathbf{M}(\eta_1\mathbf{X}_1, ..., \eta_n\mathbf{X}_n) \tag{6.10}$$

where $\eta = 1$ or -1. Thus $\mathbf{B}_{\mathbf{X}_n}$ where is defined as (6.1) will be same whether it is obtained from either $\mathbf{X}_1, ..., \mathbf{X}_n$ or $\eta_1\mathbf{X}_1, ..., \eta_n\mathbf{X}_n$. Hence (6.6) implies that

$$(\mathbf{B}_{\mathbf{X}_n}\mathbf{X}_1, ..., \mathbf{B}_{\mathbf{X}_n}\mathbf{X}_n) \stackrel{d}{=} (\eta_1\mathbf{B}_{\mathbf{X}_n}\mathbf{X}_1, ..., \eta_n\mathbf{B}_{\mathbf{X}_n}\mathbf{X}_n). \tag{6.11}$$

This yields that δ_{ni}'s, $i = 1, ..., n$ are independent and identically distributed random variables that take the values 1 and -1 with probability 1/2. Let $\mathbf{Z}_{ni} = \delta_{ni}\mathbf{Y}_{ni}$ where $\mathbf{Y}_{ni} = (Y_{ni1}, Y_{ni2})^T$ for all $i = 1, ..., n$. (6.11) denotes that $\mathbf{Z}_{n1}, ..., \mathbf{Z}_{nn}$ distributed as centrally symmetric random vectors about origin. Thus for $\mathbf{y} = (y_1, y_2)^T \in \mathbb{R}^2$ and $i = 1, ..., n$

$$\begin{aligned}
P_{H_0}(Y_{ni1} \le y_1, Y_{ni2} \le y_2, \delta_{ni} = 1) &= P_{H_0}(\delta_{ni}Y_{ni1} \le y_1, \delta_{ni}Y_{ni2} \le y_2, \delta_{ni} = 1) \\
&= P_{H_0}(Z_{ni1} \le y_1, Z_{ni2} \le y_2, \delta_{ni} = 1) \\
&= P_{H_0}(Z_{ni1} \le y_1, Z_{ni2} \le y_2, Z_{ni2} > 0) \\
&= P_{H_0}(Z_{ni1} \le y_1, 0 < Z_{ni2} \le y_2) \\
&= P_{H_0}(-Z_{ni1} \le y_1, 0 < -Z_{ni2} \le y_2) \\
&= P_{H_0}(Z_{ni1} \ge -y_1, -y_2 \le Z_{ni2} < 0) \\
&= P_{H_0}(Z_{ni1} \ge -y_1, Z_{ni2} \ge -y_2, Z_{ni2} < 0) \\
&= P_{H_0}(-Z_{ni1} \le y_1, -Z_{ni2} \le y_2, \delta_{ni} = -1) \\
&= P_{H_0}(\delta_{ni}Z_{ni1} \le y_1, \delta_{ni}Z_{ni2} \le y_2, \delta_{ni} = -1) \\
&= P_{H_0}(Y_{ni1} \le y_1, Y_{ni2} \le y_2, \delta_{ni} = -1)
\end{aligned}$$

and for $j \neq i$

$$\begin{aligned}
P_{H_0}(Y_{ni1} \le y_1, Y_{ni2} \le y_2, \delta_{nj} = 1) &= P_{H_0}(Y_{ni1} \le y_1, Y_{ni2} \le y_2, \delta_{ni} = 1, \delta_{nj} = 1) + \\
&\quad P_{H_0}(Y_{ni1} \le y_1, Y_{ni2} \le y_2, \delta_{ni} = -1, \delta_{nj} = 1) \\
&= P_{H_0}(Z_{ni1} \le y_1, Z_{ni2} \le y_2, Z_{ni2} > 0, Z_{nj2} > 0) + \\
&\quad P_{H_0}(Z_{ni1} \ge -y_1, Z_{ni2} \ge -y_2, Z_{ni2} < 0, Z_{nj2} > 0) \\
&= P_{H_0}(Z_{ni1} \le y_1, Z_{ni2} \le y_2, Z_{ni2} > 0, Z_{nj2} < 0) + \\
&\quad P_{H_0}(Z_{ni1} \ge -y_1, Z_{ni2} \ge -y_2, Z_{ni2} < 0, Z_{nj2} < 0) \\
&= P_{H_0}(Y_{ni1} \le y_1, Y_{ni2} \le y_2, \delta_{ni} = 1, \delta_{nj} = -1) + \\
&\quad P_{H_0}(Y_{ni1} \le y_1, Y_{ni2} \le y_2, \delta_{ni} = -1, \delta_{nj} = -1) \\
&= P_{H_0}(Y_{ni1} \le y_1, Y_{ni2} \le y_2, \delta_{nj} = -1).
\end{aligned}$$

Hence these imply that δ_{ni} for $i = 1, ..., n$, is independent of $\mathbf{Y}_{n1}, ..., \mathbf{Y}_{nn}$. Now, suppose that $F^s_{\mathbf{Z}_n}$ and $F^s_{\mathbf{Y}_n}$ be the sample distribution functions of $\{\pm\mathbf{Z}_{n1}, ..., \pm\mathbf{Z}_{nn}\}$ and $\{\pm\mathbf{Y}_{n1}, ..., \pm\mathbf{Y}_{nn}\}$, respectively. Since $\{\pm\mathbf{Z}_{n1}, ..., \pm\mathbf{Z}_{nn}\} = \{\pm\mathbf{Y}_{n1}, ..., \pm\mathbf{Y}_{nn}\}$, it is clear that $F^s_{\mathbf{Z}_n} = F^s_{\mathbf{Y}_n}$. This equality, along with $D\left(\mathbf{Z}_{ni}, F^s_{\mathbf{Z}_n}\right) = D\left(-\mathbf{Z}_{ni}, F^s_{\mathbf{Z}_n}\right)$ (resulted from property P1 by considering $\mathbf{A} = -I_2$ and $\mathbf{b} = \mathbf{0}$) conclude that $D\left(\mathbf{Z}_{ni}, F^s_{\mathbf{Z}_n}\right) = D\left(\mathbf{Y}_{ni}, F^s_{\mathbf{Y}_n}\right)$, for all $i = 1, ..., n$. Additionally, from property $P1$ of the sample depth function and Remark 2.1, we see that $D\left(\mathbf{X}_i, F^s_n\right) = D\left(\mathbf{Z}_{ni}, F^s_{\mathbf{Z}_n}\right)$. Hence $D\left(\mathbf{X}_i, F^s_n\right) = D\left(\mathbf{Y}_{ni}, F^s_{\mathbf{Y}_n}\right)$. This shows that R_i is a function of $\mathbf{Y}_{n1}, ..., \mathbf{Y}_{nn}$ and thus is independent of δ_{ni}, $i = 1, 2, ..., n$. Under null hypothesis, $R_1, ..., R_n$ have the discrete uniform distribution on $\{1, ..., n\}$. Then the expectation and variance of $T^{1/2}_{n,D}$ are given as

$$E\left(T^{1/2}_{n,D}\right) = \sqrt{\frac{6}{n(n+1)(2n+1)}}E\left(\sum_{i=1}^{n} \delta_{in}R_i\right) = 0$$

and

$$Var\left(T_{n,D}^{1/2}\right) = \frac{6}{n(n+1)(2n+1)}\sum_{i=1}^{n} E\left({R_i}^2\right) + \frac{6}{n(n+1)(2n+1)}\sum_{i=1}^{n}\sum_{\substack{j=1\\ i\neq j}}^{n} E\left(\delta_{ni}\delta_{nj}R_iR_j\right)$$
$$= 1,$$

respectively. Because of the dependency between summands in $T_{n,D}^{1/2}$, the central limit theory is not applied. In the other hand, $T_{n,D}^{1/2}$ is equal in distribution to

$$K_n = \sqrt{\frac{6}{n(n+1)(2n+1)}}\sum_{i=1}^{n}\delta_i i$$

where δ_i's, $i = 1, ..., n$ are independent random variables with probability 1/2 of being 1 or -1. Since K_n is sum of independent random variables and the Lyapunov's condition

$$\lim_{n\to\infty}\left(\sqrt{\frac{6}{n(n+1)(2n+1)}}\right)^{(2+\delta)}\sum_{i=1}^{n}\left(E|\delta_i i|\right)^{(2+\delta)} = \lim_{n\to\infty}\left(\frac{n^3}{3}\right)^{-(2+\delta)/2}\sum_{i=1}^{n} i^{2+\delta} = 0$$

is satisfied for $\delta = 1$, then the asymptotic null distribution is obtained by Lyapunov's central limit theorem.

Proof of Theorem 3.3: It is clear that $\mathbf{V}_{ni} = \mathbf{B}_{\mathbf{W}_n}\mathbf{W}_{ni}$, $i = 1, ..., n$ where

$$\mathbf{B}_{\mathbf{W}_n} = \begin{bmatrix} \cos\left(\frac{\pi}{2}-\theta_{\mathbf{M}_{\mathbf{W}_n}}\right) & -\sin\left(\frac{\pi}{2}-\theta_{\mathbf{M}_{\mathbf{W}_n}}\right) \\ \sin\left(\frac{\pi}{2}-\theta_{\mathbf{M}_{\mathbf{W}_n}}\right) & \cos\left(\frac{\pi}{2}-\theta_{\mathbf{M}_{\mathbf{W}_n}}\right) \end{bmatrix}. \tag{6.12}$$

Let $\mathbf{A}$ be an arbitrary 2×2 nonsingular matrix and define $\tilde{\mathbf{V}}_{ni} = \mathbf{B}_{\mathbf{AW}_n}\tilde{\mathbf{W}}_{ni}$ where

$$\mathbf{B}_{\mathbf{AW}_n} = \begin{bmatrix} \cos\left(\frac{\pi}{2}-\theta_{\tilde{\mathbf{M}}_{\mathbf{W}_n}}\right) & -\sin\left(\frac{\pi}{2}-\theta_{\tilde{\mathbf{M}}_{\mathbf{W}_n}}\right) \\ \sin\left(\frac{\pi}{2}-\theta_{\tilde{\mathbf{M}}_{\mathbf{W}_n}}\right) & \cos\left(\frac{\pi}{2}-\theta_{\tilde{\mathbf{M}}_{\mathbf{W}_n}}\right) \end{bmatrix}.$$

with $\theta_{\tilde{\mathbf{M}}_{\mathbf{W}_n}} \in [0, \pi)$, as the angle between horizontal-axis and the sample median $\tilde{\mathbf{M}}_{\mathbf{W}_n}$ that is obtained in the same way as $\mathbf{M}_{\mathbf{W}_n}$, through $\tilde{\mathbf{W}}_{ni}$'s instead of $\mathbf{W}_{ni}$'s, $i = 1, ..., n$. Moreover, $\tilde{\mathbf{W}}_{ni} = \mathbf{\Gamma}_{\mathbf{AX}_n}\mathbf{AX}_i$, where $\mathbf{\Gamma}_{\mathbf{AX}_n}$ is Tyler's matrix defined in terms of the transformed data points $\mathbf{AX}_i$, for all $i = 1, ..., n$.

If $n > 2$, Randles [36] indicated that $\mathbf{\Gamma}_n$ satisfies the condition

$$\mathbf{A}^T\mathbf{\Gamma}_{\mathbf{AX}_n}^T\mathbf{\Gamma}_{\mathbf{AX}_n}\mathbf{A} = k\mathbf{\Gamma}_n^T\mathbf{\Gamma}_n \tag{6.13}$$

where k is a positive scalar that may depends on $\mathbf{A}$ and the data. This equation clearly shows that there exists an orthogonal matrix $\mathbf{H} = k^{-1/2}\mathbf{\Gamma}_{\mathbf{AX}_n}\mathbf{A}\mathbf{\Gamma}_n^{-1}$ such that

$$\sqrt{k}\mathbf{H}\mathbf{\Gamma}_n = \mathbf{\Gamma}_{\mathbf{AX}_n}\mathbf{A}. \tag{6.14}$$

It follows easily that

$$\tilde{\mathbf{W}}_{ni} = \mathbf{\Gamma}_{\mathbf{A}\mathbf{X}_n}\mathbf{A}\mathbf{X}_i = \sqrt{k}\mathbf{H}\mathbf{\Gamma}_n\mathbf{X}_i = \sqrt{k}\mathbf{H}\mathbf{W}_{ni}. \tag{6.15}$$

Additionally, property P1 of the sample depth function along with Remark 2.1 and equation (6.15) show that $\tilde{\mathbf{M}}_{\mathbf{W}_n} = \sqrt{k}\mathbf{H}\mathbf{M}_{\mathbf{W}_n}$. Thus, the result follows from Theorem 3.1.

Proof of Theorem 3.4: The Tyler's matrix $\mathbf{\Gamma}_n \equiv \mathbf{\Gamma}(\mathbf{X}_1, ..., \mathbf{X}_n)$ is invariant under sign changes among the $\mathbf{X}_i$'s (Randles [36]), that is

$$\mathbf{\Gamma}(\mathbf{X}_1, ..., \mathbf{X}_n) = \mathbf{\Gamma}(\eta_1\mathbf{X}_1, ..., \eta_n\mathbf{X}_n). \tag{6.16}$$

Hence, by (6.6) we have

$$(\mathbf{B}_{\mathbf{W}_n}\mathbf{\Gamma}_n\mathbf{X}_1, ..., \mathbf{B}_{\mathbf{W}_n}\mathbf{\Gamma}_n\mathbf{X}_n) \overset{d}{=} (\eta_1\mathbf{B}_{\mathbf{W}_n}\mathbf{\Gamma}_n\mathbf{X}_1, ..., \eta_n\mathbf{B}_{\mathbf{W}_n}\mathbf{\Gamma}_n\mathbf{X}_n). \tag{6.17}$$

where $\mathbf{B}_{\mathbf{W}_n}$ is defined as (6.12). Additionally, from property P1 of the sample depth function and Remark 2.1, it is straightforward to verify that $R_{ni} = R_i$ for all $i = 1, ..., n$. The rest of the proof proceeds as in Theorem 3.2.

REFERENCES

[1] Aki, S. (1993). On nonparametric tests for symmetry in R^m. *Annals of the Institute of Statistical Mathematics* **45**, 4, 787-800.

[2] Azzalini, A. and Capitanio, A. (2003). Distributions generated by perturbation of symmetry with emphasis on a multivariate skew t distribution. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, **65**, 2, 367-389.

[3] Azzalini, A. and Dalla Valle, A. (1996). The multivariate skew-normal distribution. *Biometrika*, **83**, 4, 715-726.

[4] Baringhaus, L. (1991). Testing for spherical symmetry of a multivariate distribution. *Annals of Statistics*, **19**, 2, 899-917.

[5] Brown, B. M. and Hettmansperger, T. P. (1987). Affine invariant rank methods in the bivariate location model. *Journal of the Royal Statistical Society. Series B (Methodological)*. 301-310.

[6] Cassart, D., Hallin, M. and Paindaveine, D. (2011). A class of optimal tests for symmetry based on local Edgeworth approximations. *Bernoulli*, **17**, 3, 1063-1094.

[7] Chernozhukov, V., Galichon, A., Hallin, M. and Henry, M. (2017). MongeKantorovich depth, quantiles, ranks and signs. *The Annals of Statistics*, **45**, 1, 223-256.

[8] Dyckerhoff, R., Ley, C. and Paindaveine, D. 2015). Depth-based runs tests for bivariate central symmetry. *Annals of the Institute of Statistical Mathematics*, **67**, 5, 917-941.

[9] Einmahl, J. H., and Gan, Z. (2016). Testing for central symmetry. *Journal of Statistical Planning and Inference*, **169**, 27-33.

[10] Ghosh, S., and Ruymgaart, F. H. (1992). Applications of empirical characteristic functions in some multivariate problems. *Canadian Journal of Statistics*, **20**, 4, 429-440.

[11] Hallin M., and Paindaveine D. (2002a) Multivariate Signed Ranks: Randles' Interdirections or Tyler's Angles. *Statistical Data Analysis Based on the L1-Norm and Related Methods*, 271-282.

[12] Hallin M., and Paindaveine D. (2002b). Optimal tests for multivariate location based on interdirections and pesudo-mahalanobis ranks, *The Annals of Statistics*, **30**, 4, 1103-1133.

[13] Hallin M., and Paindaveine D. (2002), Optimal procedures based on interdirections and pseudo-Mahalanobis ranks for testing multivariate elliptic white noise against ARMA dependence. *Bernoulli*, **8**, 6, 787-815.

[14] Hallin M., and Paindaveine D. (2004). Rank-based optimal tests of the adequacy of an elliptic VARMA model, *Annals of Statistics*, **32**, 2642-2678.

[15] Hallin M., and Paindaveine D. (2005), Affine-invariant aligned rank tests for the multivariate general linear model with VARMA errors. *Journal of Multivariate Analysis*, **93**, 1, 122-163.

[16] Heathcote, C. R., Rachev, S. T. and Cheng, B. (1995). Testing multivariate symmetry. *Journal of Multivariate Analysis*, **54**, 1, 91-112.

[17] Henze, N., Klar, B. and Meintanis, S. G. (2003). Invariant tests for symmetry about an unspecified point based on the empirical characteristic function. *Journal of Multivariate Analysis*, **87**, 2, 275-297.

[18] Hettmansperger, T. P., Nyblom,J. and Oja, H. (1994). Affine invariant multivariate one-sample sign tests. *Journal of the Royal Statistical Society. Series B (Methodological)*. 221-234.

[19] Hettmansperger, T. P., Mottonen,J, and Oja, H. (1997). Affine-invariant multivariate one-sample signed-rank tests. *Journal of the American Statistical Association*, **92**, 440, 1591-1600.

[20] Hotelling, H. (1931). The generalization of Student's ratio. *Annals of Mathematical Statistics*, **2**, 3, 360-378.

[21] Jones, M. C., and Pewsey, A. (2009). Sinh-arcsinh distributions. *Biometrika*, **96**, 4, 761-780.

[22] Koltchinskii, V. I., and Li, L. (1998). Testing for spherical symmetry of a multivariate distribution. *Journal of Multivariate Analysis*, **65**, 2, 228-244.

[23] Koshevey, G. and Mosler, K. (1997). Zonoid trimming for multivariate distributions. *The Annals of Statistics*, 1998-2017.

[24] Li, J., and Liu, R. V. (2004). New nonparametric tests of multivariate locations and scales using data depth. *Statistical Science*, **19**, 4, 686-696.

[25] Liu, R. Y. (1988). On a notion of simplicial depth. *Proceedings of the National Academy of Sciences*, **85**, 6, 1732-1734.

[26] Liu, R. Y., Parelius, J. M. and Singh, K. (1999). Multivariate analysis by data depth: descriptive statistics, graphics and inference,(with discussion and a rejoinder by Liu and Singh). *Annals of Statistics*, **27**, 3, 783-858.

[27] Liu, R. Y., and Singh, K. (1993). A quality index based on data depth and multivariate rank tests. *Journal of the American Statistical Association*, **88**, 421, 252-260.

[28] Liu, R. Y., and Singh, K. (2006). Rank tests for multivariate scale difference based on data depth. *DIAMCS Series in Discrete Mathematics and Theoretical Computer Science*, **72**, 17-35.

[29] Mahfoud, Z. R., and Randles, R. H. (2005). On multivariate signed-rank tests. *J. Nonparametric Statistics*, **17**, 201-216.

[30] Manzotti, A., Perez, F. J. and Quiroz, A. J. (2002). A statistic for testing the null hypothesis of elliptical symmetry. *Journal of Multivariate Analysis*, **81**, 2, 274-285.

[31] McWilliams, T. P. (1990). A distribution-free test for symmetry based on a runs statistic. *Journal of the American Statistical Association*, **85**, 412, 1130-1133.

[32] Mottonen, J. and Oja, H. (1995). Multivariate Spatial Sign and Rank Methods. *Journal of Nonparametric Statistics*, **5**, 201-213.

[33] Neuhaus, G., and L. X. Zhu (1998). Permutation tests for reflected symmetry. *Journal of Multivariate Analysis*, **67**, 2, 129-153.

[34] Oja, H. (1983). Descriptive statistics for multivariate distributions. *Statistics & Probability Letters*, **1**, 6, 327-332.

[35] Randles, R. H. (1989). A distribution-free multivariate sign test based on interdirections. *J. Amer. Statist. Assoc.*, 84, 1045-1050.

[36] Randles, R. H. (2000). A simpler, affine-invariant, multivariate, distribution-free sign test. *Journal of the American Statistical Association*, 95, 452, 1263-1268.

[37] Rousson, V. (2002). On distribution-free tests for the multivariate two-sample location-scale model. *Journal of Multivariate Analysis*, **80**, 1, 43-57.

[38] Serfling, R. J. (2006). Multivariate symmetry and asymmetry. *Encyclopedia of Statistical Sciences.*

[39] Tukey, J. W. (1975). Mathematics and the picturing of data. *In Proceedings of the International Congress of Mathematicians*, **2**, 523-531.

[40] Tyler, D. E. (1987). A distribution-free M-estimator of multivariate scatter. *Annals of Statistics*, **15**, 1, 234-251.

[41] Peters D. and Randles R. H. (1990). A multivariate signed-rank test for the one-sample location problem, *J. Amer. Statist. Assoc.*, **85**, 552-557.

[42] Zuo, Y., and Serfling, R. (2000). General notions of statistical depth function. *Annals of Statistics*, **28**, 2, 461-482.