Tests and estimates of shape based on spatial signs and ranks

Seija Sirkiä[1] Sara Taskinen[1] Hannu Oja[2]
David E. Tyler[3]

June 11, 2008

Abstract

Nonparametric procedures for testing and estimation of the shape matrix in the case of multivariate elliptic distribution are considered. Testing for sphericity is an important special case. The tests and estimates are based on the spatial sign and rank covariance matrices. The estimates based on the spatial sign covariance matrix and symmetrized spatial sign covariance matrix are Tyler's (1987a) shape matrix and and Dümbgen's (1998) shape matrix, respectively. The test based on the spatial sign covariance matrix is the sign test statistic in the class of nonparametric tests proposed by Hallin and Paindaveine (2006). New tests and estimates based on the spatial rank covariance matrix are proposed. The shape estimates introduced in the paper play an important role in the inner standardization of the spatial sign and rank tests for multivariate location. Limiting distributions of the tests and estimates are reviewed and derived, and asymptotic efficiencies as well as finite sample efficiencies of the proposed tests are compared to those of the classical modified John's (1971,1972) test and the van der Waerden test (Hallin and Paindaveine, 2006). The symmetrized spatial sign and rank based estimates and tests seem to have a very high efficiency in the multivariate normal case and they are much better than classical estimate (shape matrix based on the regular covariance matrix) and test (John's test) for distributions with heavy tails. *Keywords:* Affine invariance, Pitman efficiency, Robustness, Spatial rank, Spatial sign, Spatial signed-rank

[1]Department of Mathematics and Statistics, FI-40014, University of Jyväskylä, Finland
[2]Tampere School of Public Health, FI-33014, University of Tampere, Finland
[3]Department of Statistics, Rutgers University, New Brunswick, New Jersey 08903

1 Introduction

Let $\boldsymbol{x}_1, ..., \boldsymbol{x}_n$ be a random sample from a p-variate elliptical distribution with a density function

$$det(\Sigma)^{-1/2} g\left(\|\Sigma^{-1/2}(\boldsymbol{x}-\boldsymbol{\theta})\|\right)$$

where $\boldsymbol{\theta}$ is the symmetry centre of the distribution and Σ is a positive definite symmetric $p \times p$ scatter matrix parameter. If symmetry centre $\boldsymbol{\theta} = \mathbf{0}$ and scatter $\Sigma \propto I_p$, then the distribution of $\boldsymbol{x}_i$ is spherically symmetric around the origin. Any random vector $\boldsymbol{x}$ can be decomposed into two parts, $\boldsymbol{x} = r\boldsymbol{u}$, where $r = \|\boldsymbol{x}\|$ is the modulus and $\boldsymbol{u} = r^{-1}\boldsymbol{x}$ is the direction vector. If $\boldsymbol{x}$ is spherically symmetric around the origin, its modulus r and direction $\boldsymbol{u}$ are independent, $\boldsymbol{u}$ is uniformly distributed on the unit sphere and the density of r is proportional to $r^{p-1}g(r)$.

The scatter parameter Σ describing the covariation of the p variables can be further divided into two components as $\Sigma = \sigma \cdot \Lambda$ where $\sigma = \sigma(\Sigma)$ is a scale parameter and $\Lambda = \sigma^{-1}\Sigma$ is a shape matrix. The scale parameter is assumed to satisfy $\sigma(I_p) = 1$ and $\sigma(a\Sigma) = a\sigma(\Sigma)$ for all $a > 0$. In the literature, the scale parameter $\sigma(\Sigma)$ has been defined as Σ_{11} or as an arithmetic or geometric mean of the eigenvalues of Σ,

$$\frac{1}{p}tr(\Sigma) \quad \text{and} \quad det(\Sigma)^{1/p},$$

for example. Paindaveine (2007) called the decomposition

$$\sigma = det(\Sigma)^{1/p} \quad \text{and} \quad \Lambda = det(\Sigma)^{-1/p}\Sigma$$

canonical, as the Fisher information matrix for scale and shape is block-diagonal for this decomposition only. For a thorough discussion of these concepts and references, see Paindaveine (2007).

In this paper, we wish to test the null hypothesis of sphericity $H_0 : \Lambda = I_p$. Note that the hypothesis does not depend on the decomposition $\Sigma = \sigma\Lambda$. If one wishes to test the hypothesis $H_0 : \Lambda = V$, one can use the standardized observations $V^{-1/2}\boldsymbol{x}_i$ instead of the original ones. Besides testing the shape, we also wish to find an estimate for an unknown Λ.

Several approaches for testing the shape are found in the literature. In the multivariate normal distribution case the likelihood ratio test for testing the null hypotheses $H_0 : \Lambda = I_p$ depends only on the sample covariance

matrix S and is given by

$$L = \left\{ \frac{det(S)^{1/p}}{p^{-1}tr(S)} \right\}^{n/(2p)},$$

a ratio of two scale functionals to the power $n/(2p)$ (Mauchly, 1940). Under the null hypotheses, the limiting distribution of $-2\log(L)$ is $\chi^2_{(p+2)(p-1)/2}$. Muirhead and Waternaux (1980) later showed that the adjusted test statistic $-2\log(L)/(1+\kappa)$, where κ is a kurtosis parameter, may be used to test for sphericity under any elliptical distribution with finite fourth order moments (with the same limiting null distribution as in the multivariate normal case). The regular covariance matrix can be further replaced by any other (robust) scatter matrix functional; under weak assumptions the limiting distribution of properly adjusted $-2\log(L)/b$ will again be $\chi^2_{(p+2)(p-1)/2}$. See Tyler (1982, 1983). Also John (1971,1972) considered the testing problem in the multivariate normal distribution case. He showed that the test

$$Q_J = \frac{n}{2} \left\| \frac{S}{tr(S)/p} - I_p \right\|^2, \tag{1}$$

where S is again the sample covariance matrix, is the locally most powerful invariant test for sphericity. Note that the test statistic now uses the shape matrix $pS/tr(S)$. The limiting distribution of Q_J is again $\chi^2_{(p+2)(p-1)/2}$. See also Sugiura (1972) and Nagao (1973). Hallin and Paindaveine (2006) showed that a modified version $Q_{J'} = Q_J/(1+\kappa)$ is asymptotically equivalent to the adjusted Mauchly test and therefore has the same limiting null distribution under any elliptic distribution with finite fourth order moments.

Hallin and Paindaveine (2006) proposed a family of signed rank test statistics which are based on the spatial sign vectors $\boldsymbol{u}_i$ and ranks R_i of the moduli r_i, $i = 1, ..., n$. In (1), the sample covariance matrix is replaced by a signed rank covariance matrix $S_K = \text{ave}\{K(R_i/(n+1))\boldsymbol{u}_i\boldsymbol{u}_i^T\}$ with a chosen score function K. The tests of Hallin and Paindaveine (2006) appear to be valid without any moment assumptions and can be made asymptotically optimal at given target densities g. Hallin et al. (2006) proposed optimal R-estimators for the shape of an elliptical population.

In this paper the tests and estimates based on the spatial signs and spatial ranks are considered. The spatial sign and rank vectors have been previously used for multivariate one-sample and several samples location problems in Möttönen and Oja (1995), Möttönen et al. (2003) and Oja and Randles

(2004) among others. Tyler (1987a) introduced an affine equivariant shape matrix estimator based on spatial sign vectors and showed that his estimator is distribution-free under elliptical model. Tyler (1987b) considered the problem of testing the uniformity of the directional distribution and introduces a sign test related to Tyler's scatter matrix estimate. In Visuri et al. (2000), the spatial sign covariance matrix and spatial rank covariance matrix were introduced and their usefulness in scatter matrix estimation was discussed.

Our plan is as follows. In Section 2 the concepts of spatial sign and ranks and their use is discussed, some useful matrix notations are also given. In sections 3, 4 and 5 the tests and estimates based on spatial signs, symmetrized spatial signs and and spatial ranks are derived. The limiting distributions are found for limiting power (tests) and efficiency (estimates) comparisons. In section 6 the limiting efficiencies of the derived tests and estimates with respect to classical ones are found. The finite sample efficiencies are compared in a small simulation study. Last, in section 7 some final remarks are discussed.

2 Definitions and notations

2.1 Spatial signs and ranks

The **spatial sign function** is defined as

$$\boldsymbol{U}(\boldsymbol{x}) = \begin{cases} \|\boldsymbol{x}\|^{-1}\boldsymbol{x}, & \boldsymbol{x} \neq \boldsymbol{0} \\ \boldsymbol{0}, & \boldsymbol{x} = \boldsymbol{0}, \end{cases}$$

where $\|\boldsymbol{x}\| = (\boldsymbol{x}^T\boldsymbol{x})^{1/2}$ is the Euclidean length of the vector $\boldsymbol{x}$. The **spatial rank function** of distribution (cdf) F is

$$\boldsymbol{R}_F(\boldsymbol{x}) = E_{\boldsymbol{y}}\left[\boldsymbol{U}(\boldsymbol{x}-\boldsymbol{y})\right],$$

where $\boldsymbol{y}$ is distributed according to F. (The notation $E_{\boldsymbol{y}}[\cdot]$ means that the expectation is taken with respect to the random vector $\boldsymbol{y}$.) If that distribution F is spherical it holds that for a vector $\boldsymbol{x} = r\boldsymbol{u}$ the spatial rank function $\boldsymbol{R}_F(\boldsymbol{x}) = q_F(r)\boldsymbol{u}$ for some bounded increasing function q_F depending on the distribution F (Möttönen et al. 1997). The distribution dependent constant

$$c_F^2 = E_F[q_F^2(r)] \tag{2}$$

is needed often in the calculations in this paper. The **spatial signed-rank function** is closely related to the spatial rank function and is defined as

$$\boldsymbol{Q}_F(\boldsymbol{x}) = (E_{\boldsymbol{y}}\left[\boldsymbol{U}(\boldsymbol{x}-\boldsymbol{y}) + \boldsymbol{U}(\boldsymbol{x}+\boldsymbol{y})\right])/2.$$

If F is symmetric with respect to the origin it holds that $\boldsymbol{Q}_F(\boldsymbol{x}) = \boldsymbol{R}_F(\boldsymbol{x})$.

Applying spatial functions to the empirical distribution given by data points $\boldsymbol{x}_1, \ldots, \boldsymbol{x}_n$ produces the so called **spatial sign vectors** $\boldsymbol{u}_i = \boldsymbol{U}(\boldsymbol{x}_i)$, the **spatial rank vectors** $\boldsymbol{R}_i = \text{ave}_j \boldsymbol{U}(\boldsymbol{x}_i - \boldsymbol{x}_j)$ and the **spatial signed-rank vectors** $\boldsymbol{Q}_i = (\text{ave}_j \boldsymbol{U}(\boldsymbol{x}_i - \boldsymbol{x}_j) + \text{ave}_j \boldsymbol{U}(\boldsymbol{x}_i + \boldsymbol{x}_j))/2$. We write also $\boldsymbol{x}_{ij} = \boldsymbol{x}_i - \boldsymbol{x}_j$ and $\boldsymbol{u}_{ij} = \boldsymbol{U}(\boldsymbol{x}_{ij})$, $i, j = 1, \ldots, n$.

The spatial signs, ranks and signed-ranks have been used for testing and estimation in the multivariate location problems. If $\boldsymbol{x}_1, \ldots, \boldsymbol{x}_n$ is a random sample from a distribution symmetric around $\boldsymbol{\theta}$, then the sign and signed-rank test statistics for testing $H_0 : \boldsymbol{\theta} = \boldsymbol{0}$ are

$$\text{ave}\left\{\boldsymbol{u}_i\right\} \quad \text{and} \quad \text{ave}\left\{\boldsymbol{Q}_i\right\},$$

respectively. In the two-sample location case with two independent random samples $\boldsymbol{x}_1, \ldots, \boldsymbol{x}_{n_1}$ and $\boldsymbol{x}_{n_1+1}, \ldots, \boldsymbol{x}_{n_1+n_2}$, the rank test statistic is

$$\frac{1}{n_2} \sum_{i=n_1+1}^{n_1+n_2} \boldsymbol{R}_i$$

where the spatial ranks $\boldsymbol{R}_i$ are based on the combined sample

$$\boldsymbol{x}_1, \ldots, \boldsymbol{x}_{n_1}, \boldsymbol{x}_{n_1+1}, \ldots, \boldsymbol{x}_{n_1+n_2}.$$

The test are then conditionally (and asymptotically) distribution-free with the covariance matrices proportional to sign, signed-rank and rank covariance matrices

$$\text{ave}\left\{\boldsymbol{u}_i\boldsymbol{u}_i^T\right\}, \quad \text{ave}\left\{\boldsymbol{Q}_i\boldsymbol{Q}_i^T\right\} \quad \text{and} \quad \text{ave}\left\{\boldsymbol{R}_i\boldsymbol{R}_i^T\right\},$$

respectively. The squared versions of the test statistics (with "outer standardization") are not affine invariant, however; they are only orthogonally invariant. The affine invariance is attained by an inner standardization: For each test statistic, one can find a linear transformation $\boldsymbol{x}_i \to A\boldsymbol{x}_i$, $i = 1, \ldots, n$ such that the corresponding covariance matrix is proportional to an identity matrix. For the location tests and estimates, see Möttönen and Oja (1995) and Oja and Randles (2004).

The **spatial sign and rank covariance matrices**

$$\underset{i}{\text{ave}}\left\{\boldsymbol{u}_i\boldsymbol{u}_i^T\right\}, \quad \underset{i<j}{\text{ave}}\left\{\boldsymbol{u}_{ij}\boldsymbol{u}_{ij}^T\right\}, \quad \underset{i}{\text{ave}}\left\{\boldsymbol{R}_i\boldsymbol{R}_i^T\right\} \quad \text{and} \quad \underset{i}{\text{ave}}\left\{\boldsymbol{Q}_i\boldsymbol{Q}_i^T\right\}$$

have also been used in robust principal component analysis (Locantore et al., 1999; Marden, 1999; Croux et al., 2002) and also applied in subspace DOA estimation for wireless communication (Visuri et al., 2001). For sign and rank covariances based on marginal signs and ranks and Oja signs and ranks, see Visuri et al. (2000). Note that the spatial sign and rank covariance matrices are only orthogonally equivariant.

2.2 Vector and matrix notations

Throughout the paper $\text{vec}(A)$ means the vectorization of a matrix A obtained by stacking the columns on top of each other. We will use notation $K_{p,p}$ for the commutation matrix, that is, a $p^2 \times p^2$ block matrix with (i,j)-block being equal to a $p \times p$ matrix that has 1 at entry (j,i) and zero elsewhere, and $J_{p,p}$ for $\text{vec}(I_p)\text{vec}(I_p)^T$. Note that matrices $K_{p,p}$ and $J_{p,p}$ satisfy

$$K_{p,p}\text{vec}(A) = \text{vec}(A^T) \quad \text{and} \quad J_{p,p}\text{vec}(A) = tr(A)\text{vec}\left(I_p\right).$$

Finally, the projection matrix $C_{p,p}$ is defined as

$$C_{p,p} = \frac{1}{2}(I_{p^2} + K_{p,p}) - \frac{1}{p}J_{p,p}.$$

Due to the properties of $K_{p,p}$ and $J_{p,p}$ this matrix projects a vectorized matrix A on symmetric and centered – in the sense that the trace is zero – matrices. We also write

$$C_{p,p}(V) = \frac{1}{2}(I_{p^2} + K_{p,p})(V \otimes V) - \frac{1}{p}vec(V)vec(V)^T. \tag{3}$$

Then $C_{p,p}(I_p) = C_{p,p}$.

Our tests and estimates in this paper are based on

$$Q(S) = \left\| \frac{S}{tr(S)/p} - I_p \right\|^2 = \frac{p^2}{(tr(S))^2}\|C_{p,p}\text{vec}(S)\|^2, \tag{4}$$

which is proportional to the variance of eigenvalues of shape matrix $pS/tr(S)$. Among symmetric matrices the only ones with equal eigenvalues are those

proportional to the identity matrix. On the other hand, among symmetric matrices those proportional to the identity matrix are the only ones that are projected to the origin in the $C_{p,p}$-projection. Thus it is apparent that small values of Q are connected to sphericity through a matrix that is nearly proportional to the identity matrix.

We denote the symmetric square root matrix of A by $A^{1/2}$ and its generalized inverse by A^{-}.

3 Inference based on sign covariance matrix

3.1 Testing for shape

Assume that $X = (\boldsymbol{x}_1 \ ... \ \boldsymbol{x}_n)$ is a random sample from an elliptically symmetric distribution with a symmetry centre $\boldsymbol{\theta} = \mathbf{0}$ and a shape parameter Λ. In the following we assume that the shape matrix is standardized so that $tr(\Lambda) = p$. We wish to test the null hypothesis

$$H_0 : \quad \Lambda = I_p.$$

In efficiency studies we consider the sequences of contiguous alternatives

$$H_n : \quad \Lambda \propto I_p + n^{-1/2} D,$$

where D is a symmetric matrix. The **spatial sign test statistic** is based on the spatial sign covariance matrix given in the following

Definition 1 *The spatial sign covariance matrix is defined as*

$$S_1 = \underset{i}{ave} \left\{ \boldsymbol{u}_i \boldsymbol{u}_i^T \right\}.$$

See e.g. Marden (1999) and Visuri et al. (2000). Note that $\frac{1}{n} S_1$ is the null covariance matrix of the sign test statistic

$$T_1 = \underset{i}{\text{ave}} \left\{ \boldsymbol{u}_i \right\}$$

for the location problem $H_0 : \ \boldsymbol{\theta} = \mathbf{0}$. The spatial sign covariance matrix S_1 is distribution-free under the null hypothesis with

$$E(vec(S_1)) = p^{-1} vec(I_p) \ \text{ and } \ Cov(vec(S_1)) = \frac{2}{np(p+2)} C_{p,p}.$$

See Tyler (1987a).

The test statistic based on the spatial sign covariance matrix is given in

Definition 2 *The spatial sign test statistic is defined as*

$$Q_1 = Q(S_1),$$

where $Q(\cdot)$ is as in (4).

The value of Q_1 is equal to zero if and only if S_1 is proportional to the identity matrix. As the trace of the sign covariance matrix S_1 is always 1, the latter is true if and only if S_1 is equal to $p^{-1}I_p$, its expected value under the null hypothesis $\Lambda = I_p$. The following theorem gives the asymptotic distribution of Q_1.

Theorem 1 *Under the alternative sequence H_n,*

$$\frac{n}{\gamma_1}Q_1 \rightarrow_d \chi^2_{(p+2)(p-1)/2}\left(\frac{1}{\gamma_1(p+2)^2}\|C_{p,p}vec(D)\|^2\right).$$

where $\gamma_1 = 2/\left(p(p+2)\right)$.

The limiting null distribution, a central chi-squared distribution, is naturally obtained by setting $D = 0$. If the true value of $\boldsymbol{\theta}$ is unknown it can be replaced by its $\sqrt{n}$ -consistent estimate $\hat{\theta}$ without changing the limiting behavior of the test statistic under the mild assumption that $E[\|\boldsymbol{x}_i\|^{-3/2}] < \infty$. This is true under the multivariate normal case, for example. See the proof of Theorem 4.2 in Tyler (1987a).

Finally note that, in the general (even non-elliptic) case,

$$\widehat{Cov}(vec(S_1)) = n^{-1}\left(\text{ave}\left\{\text{vec}(\boldsymbol{u}_i\boldsymbol{u}_i^T)\text{vec}(\boldsymbol{u}_i\boldsymbol{u}_i^T)^T\right\} - vec(S_1)vec(S_1)^T\right)$$

is a consistent estimate of $Cov\left(vec(S_1)\right)$ and, in the elliptic case, the statistic

$$\left(C_{p,p}vec(S_1)\right)^T \widehat{Cov}(vec(S_1))^- \left(C_{p,p}vec(S_1)\right)$$

is asymptotically equivalent to $(n/\gamma_1)Q_1$.

3.2 Estimation of shape

Let $S_1(V)$ and $Q_1(V)$ be the spatial sign covariance matrix and the test statistic computed on $V^{-1/2}\boldsymbol{x}_i$, $i = 1, \ldots, n$, the observations standardized by a matrix V. The null hypothesis $H_0 :\ \Lambda = V$ is then rejected for large values of $Q_1(V)$. Therefore, a natural shape estimate would be the one for which the test statistic $Q(V)$ is as small as possible:

Definition 3 *Tyler's shape estimate $\widehat{V}_1$ based on spatial signs is the matrix that minimizes $Q_1(V)$ and for which it holds that $det(\widehat{V}_1) = 1$.*

The estimate was given by Tyler (1987a). Here the absolute minimum $Q_1(V) = Q(S_1(V)) = 0$ can in fact be reached under weak assumptions. The solution to the estimating equation

$$\underset{i}{\text{ave}}\{\boldsymbol{U}(V^{-1/2}\boldsymbol{x}_i)\boldsymbol{U}(V^{-1/2}\boldsymbol{x}_i)^T\} = \frac{1}{p}I_p \tag{5}$$

was considered in Tyler (1987a); he only used the condition $tr(\widehat{V}_1) = p$ instead of $det(\widehat{V}_1) = 1$. This estimating equation suggests an iterative algorithm where the iteration step from V_k to V_{k+1}

$$V_{k+1} \leftarrow V_k^{1/2}\underset{i}{\text{ave}}\{\boldsymbol{U}(V_k^{-1/2}\boldsymbol{x}_i)\boldsymbol{U}(V_k^{-1/2}\boldsymbol{x}_i)^T\}V_k^{1/2}, \tag{6}$$

is repeated until the sequence converges and then the result is scaled correctly. The convergence of this algorithm is proven in Tyler (1987a). Note that the estimate is affine equivariant in the sense that

$$\widehat{V}_1(AX) = Det(A)^{-2/p}A\widehat{V}_1(X)A^T,$$

where $\widehat{V}_1(X)$ means the estimate computed on observations $X = (\boldsymbol{x}_1, \ldots, \boldsymbol{x}_n)$. Tyler (1987a) also proved

Theorem 2 *At elliptical distribution with shape parameter V, the limiting distribution of the shape estimate $\widehat{V}_1$ is given by*

$$\sqrt{n}\,vec(\widehat{V}_1 - V) \rightarrow_d N_{p^2}\left(0, (p+2)^2\gamma_1 C_{p,p}(V)\right).$$

It is notable that in the elliptic case this estimate is distribution-free.

If the location centre $\boldsymbol{\theta} \neq \boldsymbol{0}$ is known then Tyler's shape matrix is calculated using the centred observations $\boldsymbol{x}_1 - \boldsymbol{\theta}, ..., \boldsymbol{x}_n - \boldsymbol{\theta}$. The case of unknown location $\boldsymbol{\theta}$ was also considered in Tyler (1987a). If $E[\|\boldsymbol{x}_i\|^{-3/2}] < \infty$ then it is possible to replace $\boldsymbol{\theta}$ with a $\sqrt{n}$-consistent estimate $\widehat{\boldsymbol{\theta}}$ without affecting the asymptotic properties of $\widehat{V}_1$. Hettmansperger and Randles (2002) proposed a simultaneous estimation of the multivariate spatial median and Tyler's shape matrix. Unfortunately, there is no proof of the existence of the solution for their estimating equations, but the algorithm they propose works well in practice.

Finally note that, for known Λ, for the location sign test statistic

$$T_1(\Lambda) = \text{ave}\left\{\boldsymbol{U}(\Lambda^{-1/2}\boldsymbol{x}_i\right\}$$

it holds that $\sqrt{n}T_1 \rightarrow_d N_p(\mathbf{0}, \frac{1}{p}I_p)$. If Λ is replaced by Tyler's estimate $\widehat{V}_1$, then $\hat{T}_1 = T_1(\widehat{V}_1)$ (with inner standardization) is an affine invariant, distribution-free test statistic for $H_0: \boldsymbol{\theta} = \mathbf{0}$. See Randles (2000).

4 Inference based on symmetrised sign covariance matrix

4.1 Testing for shape

For the same null hypothesis as in the previous section, the **symmetrised spatial sign test statistic** is constructed in exactly the same way as the sign test statistic but for the data set of all pairwise differences. The pairwise differences and their signs are denoted by $\boldsymbol{x}_{ij} = \boldsymbol{x}_i - \boldsymbol{x}_j$ and $\boldsymbol{u}_{ij} = \boldsymbol{U}(\boldsymbol{x}_{ij})$.

Definition 4 *The symmetrised sign covariance matrix is defined as*

$$S_2 = \underset{i<j}{ave}\left\{\mathbf{u}_{ij}\mathbf{u}_{ij}^T\right\}.$$

This matrix is introduced and studied in Visuri et al. (2000). Now

$$vec(S_2) = \underset{i<j}{\text{ave}}\left\{\boldsymbol{h}(\boldsymbol{x}_i, \boldsymbol{x}_j)\right\},$$

is a U-statistic with kernel $h(\boldsymbol{x}_1, \boldsymbol{x}_2) = vec(\boldsymbol{u}_{12}\boldsymbol{u}_{12}^T)$. Clearly under the null hypothesis $H_0: \Lambda = I_p$,

$$E(vec(S_2)) = p^{-1}vec(I_p).$$

If we write

$$h_2(r) = E\left[[\boldsymbol{u}_{12}\boldsymbol{u}_{12}^T]_{11} - [\boldsymbol{u}_{12}\boldsymbol{u}_{12}^T]_{22} \mid \boldsymbol{x}_2 = r\boldsymbol{e}_1\right)$$

then (based on the general U-statistic theory)

$$\sqrt{n}\left(S_2 - \frac{1}{p}I_p\right) \;=\; \sqrt{n}\ \text{ave}\left\{2h_2(r_i)\left(\boldsymbol{u}_i\boldsymbol{u}_i^T - \frac{1}{p}I_p\right)\right\} + o_P(1)$$

where, as before, $r_i = ||\boldsymbol{x}_i||$ and $\boldsymbol{u}_i = ||\boldsymbol{x}_i||^{-1}\boldsymbol{x}_i$, $i = 1, \ldots, n$. So, under the null hypothesis

$$Cov(vec(S_2)) = \frac{\gamma_2}{n} C_{p,p} + o\left(\frac{1}{n}\right),$$

where $\gamma_2 = 4E[h_2^2(r)]$. See Appendix A and Appendix B for more details and alternative formulas. For the comparison, recall that

$$\sqrt{n}\left(S_1 - \frac{1}{p} I_p\right) \;=\; \sqrt{n}\ \text{ave}\left\{\boldsymbol{u}_i \boldsymbol{u}_i^T - \frac{1}{p} I_p\right\}.$$

The symmetrised sign test statistic is defined as follows.

Definition 5 *The symmetrised sign test statistic is defined as*

$$Q_2 = Q(S_2),$$

where $Q(\cdot)$ is as in (4).

Again, it holds that $tr(S_2) = 1$ and so $Q_2 = 0$ only when $S_2 = p^{-1} I_p$. Note that since the test statistics is based on pairwise differences the location needs not to be known. The asymptotic distribution is given by

Theorem 3 *Under the alternative sequence H_n,*

$$\frac{n}{\gamma_2} Q_2 \to_d \chi^2_{(p+2)(p-1)/2}\left(\frac{1}{(p+2)^2\gamma_2}\|C_{p,p} vec(D)\|^2\right).$$

In practice, the coefficient γ_2 is not known and has to be estimated. For details see Appendix B.

In the general nonelliptic case, the covariance matrix of $vec(S_2)$ may be estimated by

$$\widehat{Cov}(vec(S_2)) = \frac{4}{n}\left[\underset{i \neq j,k}{ave}\left\{\text{vec}(\boldsymbol{u}_{ij}\boldsymbol{u}_{ij}^T)\text{vec}(\boldsymbol{u}_{ik}\boldsymbol{u}_{ik}^T)^T\right\} - vec(S_2)vec(S_2)^T\right]. \tag{7}$$

As $n\,\widehat{Cov}(vec(S_2)) \to_p \gamma_2 C_{p,p}$ in the spherical case, one may alternatively use the statistic

$$(C_{p,p} vec(S_2))^T\, \widehat{Cov}(vec(S_2))^-\, (C_{p,p} vec(S_2)),$$

which is asymptotically equivalent to $(n/\gamma_2)Q_2$.

4.2 Estimation of shape

Let again $Q_2(V) = Q(S_2(V))$ be a test statistic for testing $H_0 : \Lambda = V$. A similar argument as in the previous section leads to the following estimate of shape.

Definition 6 *The Dümbgen's shape estimate $\widehat{V}_2$ based on the symmetrised spatial signs is the one that minimizes $Q_2(V) = Q(S_2(V))$ and for which it holds that $det(\widehat{V}_2) = 1$.*

The estimator was first introduced by Dümbgen (1998) and further studied by Sirkiä et al. (2006) as a member in a family of symmetrised M-estimators of scatter. Also here the absolute minimum $Q(S_2(V)) = 0$ can be reached. The solution is given by an estimating equation is as (5) but uses pairwise differences instead of the original observations. It can be found using a similar iterative algorithm as (6). The breakdown properties of $\widehat{V}_1$ and $\widehat{V}_2$ were considered in Dümbgen and Tyler (2005). The Dümbgen's estimate is affine equivariant in the sense that

$$\widehat{V}_2(AX + \boldsymbol{b}) = det(A)^{-2/p} A \widehat{V}_2(X) A^T,$$

The limiting distribution in the elliptic case is also analogous to the limiting distribution of the estimator based on the signs, as was the case with the test statistics.

Theorem 4 *At elliptical distribution with shape parameter V, the limiting distribution of the shape estimate $\widehat{V}_2$ is given by*

$$\sqrt{n}\, vec(\widehat{V}_2 - V) \rightarrow_d N_{p^2}\left(0, (p+2)^2 \gamma_2 C_{p,p}(V)\right).$$

5 Inference based on rank covariance matrix

5.1 Testing for shape

The spatial rank covariance matrix is the covariance matrix calculated using the spatial ranks:

Definition 7 *The spatial rank covariance matrix is*

$$S_3 = \underset{i}{ave}\left\{\boldsymbol{R}_i \boldsymbol{R}_i\right\} = \underset{i,j,k}{ave}\left\{\boldsymbol{u}_{ij} \boldsymbol{u}_{ik}^T\right\}.$$

This matrix is considered in Marden (1999) and Visuri et al. (2000). Now $vec(S_3)$ is (up to a constant) asymptotically equivalent to a U-statistic

$$\underset{i<j<k}{\text{ave}} \left\{ h(\boldsymbol{x}_i, \boldsymbol{x}_j, \boldsymbol{x}_k) \right\},$$

with a symmetric kernel

$$h(\boldsymbol{x}_1, \boldsymbol{x}_2, \boldsymbol{x}_3) = \frac{1}{6}\text{vec}\left(\boldsymbol{u}_{12}\boldsymbol{u}_{13}^T + \boldsymbol{u}_{13}\boldsymbol{u}_{12}^T + \boldsymbol{u}_{21}\boldsymbol{u}_{23}^T + \boldsymbol{u}_{23}\boldsymbol{u}_{21}^T + \boldsymbol{u}_{31}\boldsymbol{u}_{32}^T + \boldsymbol{u}_{32}\boldsymbol{u}_{31}^T \right).$$

Again, based on the U-statistics theory,

$$\sqrt{n}\left(S_3 - \frac{c_F^2}{p} I_p \right) \;=\; \sqrt{n}\ \text{ave}\left\{ 3h_3(r_i)\left(\boldsymbol{u}_i\boldsymbol{u}_i^T - \frac{1}{p} I_p \right) \right\} + o_P(1),$$

where

$$h_3(r) = E\left[[\boldsymbol{u}_{12}\boldsymbol{u}_{13}^T + 2\boldsymbol{u}_{21}\boldsymbol{u}_{23}^T]_{11} - [\boldsymbol{u}_{12}\boldsymbol{u}_{13}^T + 2\boldsymbol{u}_{21}\boldsymbol{u}_{23}^T]_{22} \;|\boldsymbol{x}_1 = r\boldsymbol{e}_1 \right].$$

Then under the null hypothesis,

$$E(C_{p,p}vec(S_3)) = \mathbf{0} \;\text{ and }\; Cov(C_{p,p}vec(S_3)) = \frac{\gamma_3}{n} C_{p,p} + o\left(\frac{1}{n}\right),$$

with $\gamma_3 = 9E[h_3^2(r)]$.

Definition 8 *The spatial rank test statistic is defined as*

$$Q_3 = Q(S_3),$$

where $Q(\cdot)$ is as in (4).

Note that, unlike S_1 and S_2, the trace of S_3 is not fixed but varies from sample to sample. Its expected value is $E[tr(S_3)] = pc_F^2$. Still, Q_3 is naturally equal to zero only when S_3 is proportional to the identity matrix. The asymptotic distribution of Q_3 is as follows.

Theorem 5 *Under the alternative sequence H_n,*

$$\frac{n}{\gamma_3/(c_F^2)^2} Q_3 \to_d \chi^2_{(p+2)(p-1)/2}\left(\frac{(c_F^2)^2}{(p+2)^2\gamma_3} \|C_{p,p}vec(D)\|^2 \right).$$

Again, the coefficients c_F^2 and γ_3 need to be estimated as the distribution is unknown. A natural estimate of c_F^2 is $tr(S_3)/p$. Alternative formulas for estimation of γ_3 can be found in Appendix A and Appendix B. In the general non-elliptic case the covariance matrix of $vec(S_3)$ may be estimated with

$$\widehat{Cov}(C_{p,p}vec(S_3)) = \frac{9}{n}\, ave\left\{h(\boldsymbol{x}_i, \boldsymbol{x}_j, \boldsymbol{x}_k)h(\boldsymbol{x}_i, \boldsymbol{x}_l, \boldsymbol{x}_m)^T\right\}.$$

and one may use the statistic

$$(C_{p,p}vec(S_3))^T\, \widehat{Cov}(C_{p,p}vec(S_3))^-\, (C_{p,p}vec(S_3))\,.$$

5.2 Estimation of shape

As in the previous sections, it is possible to define a shape estimate corresponding to the above test. Let $S_3(V)$ be the spatial rank covariance matrix calculated from $V^{-1/2}\boldsymbol{x}_i$, $i = 1, \ldots, n$. Then

Definition 9 *The shape estimate $\widehat{V}_3$ based on the spatial rank covariance matrix is the one that minimizes $Q_3(V) = Q(S_3(V))$ and for which it holds that $det(\widehat{V}_3) = 1$.*

So far, there is no proof that there is a unique solution to the estimating equation

$$C_{p,p}S_3(V) = 0.$$

However, the iterative algorithm with steps

$$V_{k+1} \leftarrow V_k^{1/2} \underset{i}{\text{ave}}\{\boldsymbol{R}(V_k^{-1/2}\boldsymbol{x}_i)\boldsymbol{R}(V_k^{-1/2}\boldsymbol{x}_i)^T\}V_k^{1/2}, \tag{8}$$

seems again to work in practice.

If we assume that there exists a consistent sequence of estimates $\hat{V}_3$ such that

$$\sqrt{n}\widehat{V}_3 = O_p(1) \;\text{ and }\; \sqrt{n}C_{p,p}S_3(\widehat{V}_3) = o_P(1).$$

then, as before,

$$\sqrt{n}\, vec(\widehat{V}_3 - V) \rightarrow_d N_{p^2}\left(0, (p+2)^2\gamma_3(c_F^2)^{-1}C_{p,p}(V)\right).$$

The small sample simulation studies considered in Chapter 6 seem to comply with this result.

Starting with Tyler's scatter matrix V_0, for example, and repeating steps in (8) k times ("k step rank estimator") gives a practical estimate with known asymptotical properties. With just a few steps one gets an estimate with an asymptotical behavior almost identical to that of $\hat{V}_3$. These estimates will be considered in a future paper.

Multivariate spatial rank test statistics for location, e.g.,

$$\frac{1}{n_2}\sum_{i=n_1+1}^{n_1+n_2} \boldsymbol{R}_i$$

with $n_1 + n_2 = n$ have a conditional and limiting covariance matrix which is proportional to $S_3(V)$. Therefore, $\widehat{V}_3$ is a natural shape matrix for the inner standardization of the spatial rank test statistic. Then the resulting test statistic is affine invariant and has a simple covariance structure.

5.3 Use of the spatial signed-rank covariance matrix

It is also possible to base the test and the estimate on the spatial signed-ranks instead of the ordinary spatial signs or the ordinary spatial ranks. We then again assume that the symmetry centre $\boldsymbol{\theta} = \mathbf{0}$ is known. Whether an unknown $\boldsymbol{\theta}$ can be replaced by its estimate, remains an open question.

Definition 10 *The signed-rank covariance matrix is defined as*

$$S_4 = \underset{i}{ave}\left\{\boldsymbol{Q}_i\boldsymbol{Q}_i^T\right\}.$$

The test statistic for the null hypothesis $\Lambda = I_p$ is as before

Definition 11 *The signed-rank test statistic is defined as*

$$Q_4 = Q(S_4),$$

where $Q(\cdot)$ is as in (4).

In the spherical case, the spatial signed-rank function and the rank function converge to the same theoretical function. The asymptotic null distributions of S_4 and S_3 are then the same as well. Further,

Theorem 6 *Under the alternative sequence H_n,*

$$\frac{n}{\gamma_3/(c_F^2)^2} Q_4 \to_d \chi^2_{(p+2)(p-1)/2} \left(\frac{(c_F^2)^2}{(p+2)^2\gamma_3} ||C_{p,p} vec(D)||^2 \right).$$

The corresponding estimate of shape may also be defined as before.

Definition 12 *The shape estimate $\widehat{V}_4$ based on the signed-rank covariance matrix is the one that minimizes $Q(S_4(V))$ and for which it holds that $det(\widehat{V}_4) = 1$.*

There is, however, the same problem on the existence and uniqueness of the estimate as with $\widehat{V}_3$. Again a k-step estimate with comparable limiting properties may be used in practice.

The main motivation for the use of the signed-rank covariance matrix is that the shape estimate $\widehat{V}_4$ is a natural choice for the inner standardization to obtain an affine invariant spatial signed-rank test for location. The spatial rank based $\widehat{V}_3$ on the other hand is genuinely location invariant.

6 Limiting and finite-sample efficiencies

In this section, we will compare the tests based on the spatial sign and rank covariance matrices, Q_1, Q_2 andQ_3, to the classical John's test (1971, 1972) through limiting efficiencies and a simple simulation study. The modified John's test statistic is

$$Q_{J'} = \frac{n}{2(1+\hat{\kappa})} \left\| \frac{S}{tr(S)/p} - I_p \right\|^2,$$

where S is the regular covariance matrix and $\hat{\kappa}$ is the estimated kurtosis. Also, for statistics Q_2 and Q_3, the unknown coefficients γ_2 and γ_3 have to be estimated from the sample. See also Muirhead and Waternaux (1980) and Tyler (1982, 1983). The limiting distribution of the modified John's test under the alternative sequence is as follows (Hallin and Paindaveine, 2006).

Theorem 7 *Under the alternative sequence H_n,*

$$Q_{J'} \to_d \chi^2_{(p+2)(p-1)/2} \left(\frac{1}{2(1+\kappa)} ||C_{p,p} vec(D)||^2 \right).$$

Since the limiting distributions of different test statistics are of the same type, the efficiency comparisons may be based on non-centrality parameters only. The asymptotic relative efficiencies of spatial sign and rank tests as compared to the modified John's test reduce to

$$ARE\,(Q_1, Q_{J'}) = \frac{p(1+\kappa)}{p+2} \quad \text{and} \quad ARE\,(Q_2, Q_{J'}) = \frac{2(1+\kappa)}{(p+2)^2\gamma_2},$$

and

$$ARE\,(Q_3, Q_{J'}) = \frac{2(c_F^2)^2(1+\kappa)}{(p+2)^2\gamma_3},$$

where γ_1, γ_2, γ_3 and c_F^2 are as given before. Note that these efficiencies are also relative efficiencies of the shape estimates (based on the spatial sign and rank tests) with respect to the shape estimate based on the regular covariance matrix.

In Table 1, the limiting efficiencies are given under t-distributions with selected dimensions p and degrees of freedom ν, with $\nu = \infty$ referring to the multinormal case. Note that for multinormal distributions $\kappa = 0$ and for t-distributions $\kappa = 2/(\nu-4)$. Formulas for calculating the c_F^2 coefficients can be found in Möttönen et al. (1997). A combination of numeric integration and Monte Carlo simulation was used to find the needed values of γ_2 and γ_3: in the considered cases the density function of the distribution of the (squared) length of the observation is known and thus the outer part of the two-part integral was computed with a simple application of Simpson's rule while the value of the inner integral for given values of r was found by a straightforward simulation.

From this table it can be seen that the Q_2 and Q_3 tests behave very similarly and are highly efficient even in the normal case. Q_1 is not as efficient but still outperforms the John's test in many considered cases. As the number of dimensions increase the efficiencies increase as well.

We also considered the finite-sample efficiencies of tests using simulation designs similar to those in Hallin and Paindaveine (2006). The tests included in the comparisons were modified John's test $Q_{J'}$, spatial sign test Q_1, symmetrized spatial sign test Q_2 (with estimated γ_2) and the van der Waerden test

$$Q_{vdW} = \frac{1}{2n}\sum_{i,j} \Psi_p^{-1}\left(\frac{R_i}{n+1}\right)\Psi_p^{-1}\left(\frac{R_j}{n+1}\right)\left(\boldsymbol{u}_i^T\boldsymbol{u}_j - \frac{1}{p}\right),$$

Table 1: *Asymptotic relative efficiencies of tests Q_1, Q_2 and Q_3 relative to the modified John's test at different t-distribution cases with selected values of dimension p and degrees of freedom ν.*

	$p=2$			$p=3$			$p=4$			$p=5$		
ν	Q_1	Q_2	Q_3	Q_1	Q_2	Q_3	Q_1	Q_2	Q_3	Q_1	Q_2	Q_3
5	1.50	2.43	2.42	1.80	2.53	2.49	2.00	2.62	2.56	2.14	2.71	2.63
6	1.00	1.63	1.62	1.20	1.69	1.65	1.33	1.76	1.68	1.43	1.80	1.69
8	0.75	1.26	1.25	0.90	1.29	1.28	1.00	1.32	1.30	1.07	1.35	1.31
15	0.59	1.04	1.04	0.71	1.05	1.05	0.79	1.07	1.06	0.84	1.08	1.07
∞	0.50	0.93	0.95	0.60	0.94	0.95	0.67	0.95	0.97	0.71	0.95	0.99

where Ψ_p stands for the cdf of chi-square distribution function with p degrees of freedom and R_i is the rank of r_i among the moduli $r_1, \ldots, r_n$ of the observations. This test was introduced in Hallin and Paindaveine (2006) and it was shown to be asymptotically optimal (locally asymptotically maximin) in the multivariate normal case. Its asymptotic relative efficiencies can also be found in that paper, being slightly lower than for Q_2 and Q_3 for the considered t-distribution cases with bigger difference in higher dimension and lower degrees of freedom. In the normal case the efficiency of the van der Waerden test is naturally equal to that of the regular shape estimate, outperforming Q_2 and Q_3.

In our simulation study, 1500 samples of sizes $n = 50, 200$ were generated from a spherical bivariate normal and spherical t_6 distributions and then multiplied by a matrix of the form

$$\left(I_2 + \frac{1}{\sqrt{n}} D \right)^{1/2},$$

where symmetric matrix $D = \delta[1_p 1_p^T - I_p]$ has zero as diagonal elements and δ as the off-diagonal element (1_p is a p-vector full of ones). We thus test the hypothesis of sphericity ($\delta = 0$) with a local alternative of compound symmetry ($\delta > 0$). (A similar setting in five dimensions was also considered but not reported here: The results were similar in the considered case of compound symmetry.) The spatial rank test statistic Q_3 was not included in comparisons as it is computationally rather heavy and its asymptotic efficiency is close to that of the Q_2 test.

Figure 1: Estimated power curves as a function of δ for modified John's test (solid line), spatial sign test (dashed), symmetrized spatial sign test (dotted) and van der Waerden test (dot-dashed) in the normal and t_6 distribution cases with sample sizes $n = 50$ and $n = 200$. Also the limiting power curve ($n = \infty$) is given. The horizontal dashed line shows the nominal 0.05 level.

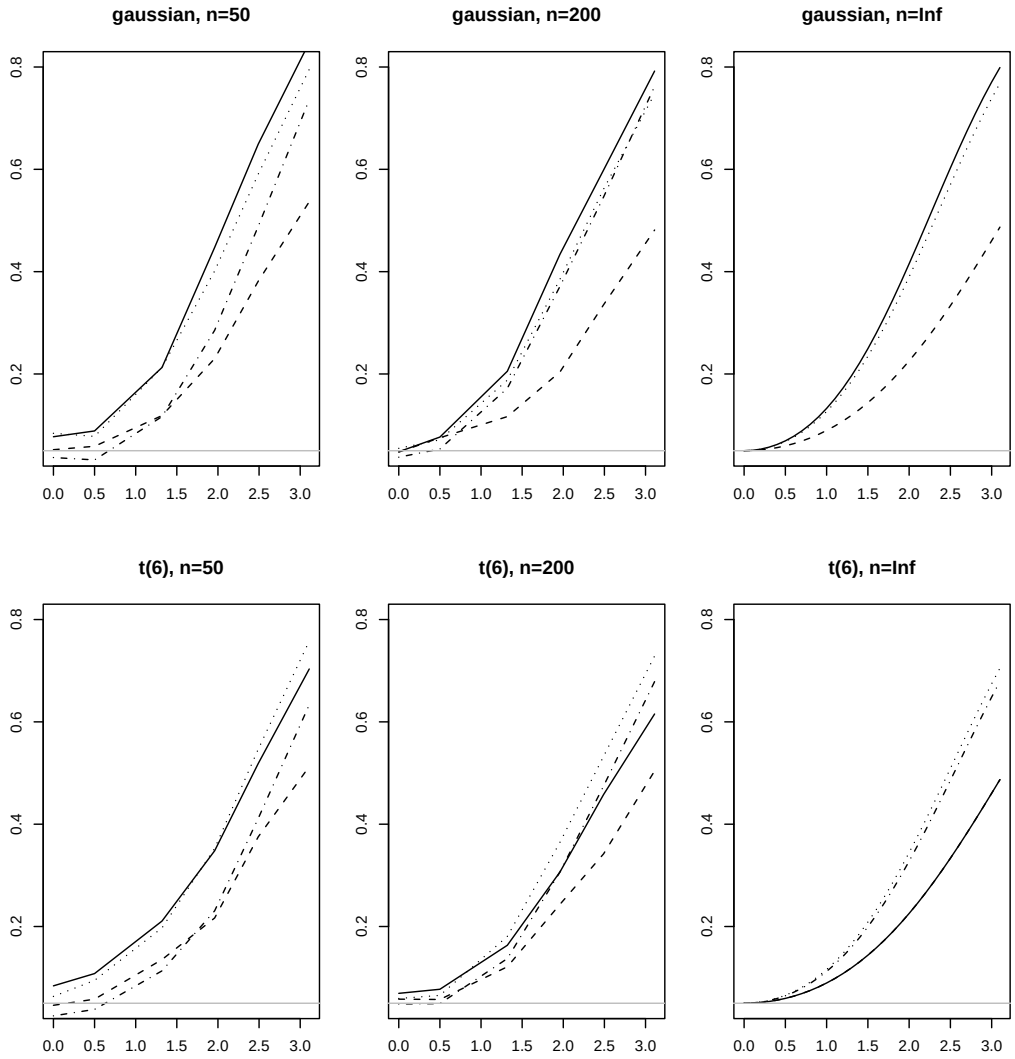

Figure 1 shows simulated power curves of the modified John's test, spatial sign test, symmetrized spatial sign test and van der Waerden test. Modified John's test is behaving surprisingly well also in the t_6 distribution case. As expected, the symmetrized spatial sign test is much more powerful than the spatial sign test.

The power estimates were based on the critical values obtained from the limiting chi square distribution. The observed sizes for the test statistics were then often quite far away from the nominal size 0.05 which makes the power comparison difficult. Table 2 lists the observed sizes for easier reference. For example the high power of the John's test reported above is at least partly due to the incorrect size of the test. It is worth noting that the size of the spatial sign test is practically correct even with sample size 50. The symmetrized spatial sign test also behaves adequately well in both cases with sample size 200. The observed sizes of John's test and the van der Waerden test are often far away from the nominal size.

Table 2: *Observed sizes of the modified John's test, the spatial sign test, the symmetrized spatial sign test and van der Waerden test in the bivariate normal and t_6 distribution cases and with sample sizes 50 and 200.*

		$Q_{J'}$	Q_1	Q_2	Q_W
gaussian	n=50	.077	.052	.084	.037
	n=200	.048	.047	.055	.037
t_6	n=50	.084	.046	.063	.025
	n=200	.069	.058	.059	.049

We have not included distributions with lighter than normal tails in the comparisons since signs and ranks are particularly suited for heavy tailed distributions. It is apparent that in such cases the proposed methods will be less efficient while modified John's test and the shape estimate based on regular covariance matrix will perform better.

7 Final remarks

Throughout the paper we have assumed that the data are coming from an elliptically symmetric distribution. This is motivated by the fact that the shape matrices considered are natural standardizing matrices for location tests in

the elliptic model. The spatial sign test which uses the inner standardization based on Tyler's shape matrix, for example, is strictly distribution-free in the elliptic model. The asymptotic and other theoretical properties are also fairly comfortably obtained, and the limiting behavior of the tests and estimates can be characterized by one constant (γ_1, γ_2 or γ_3) only.

However, shape matrices are meaningful also under more general models. Consider the elliptic model as the collection of distributions of all affine transformations of spherically symmetric (around the origin) random vectors, that is

$$\boldsymbol{x}_i = \Sigma^{-1/2}\boldsymbol{z}_i + \mu$$

where $Oz_i \sim \boldsymbol{z}_i$ for all orthogonal matrices O. $O\boldsymbol{z}_i \sim \boldsymbol{z}_i$ means that $O\boldsymbol{z}_i$ and $\boldsymbol{z}_i$ have the same distribution. The density functions of the generating random vectors $\boldsymbol{z}_i$ have contours that are concentric spheres, and the shape matrix is an identity matrix I_p. Now consider a wider model generated by exchangeable and marginally symmetric random vectors, that is, those for which it holds $PJ\boldsymbol{z}_i \sim \boldsymbol{z}_i$ for all permutation matrices P and sign-change matrices J. (A permutation matrix P is obtained from an identity matrix by permuting its rows and/or columns. A sign-change matrix J is a diagonal matrix with diagonal elements ± 1.) The distributions of $\boldsymbol{z}_i$ can be said to have unit shape in this case as well, as all shape matrices then are equal to identity matrix. The model includes distributions where $\boldsymbol{z}_i$ has independent and identically distributed components or a density of the general form

$$f(\boldsymbol{x}) = \exp\left(-\rho(||\boldsymbol{x}||)\right),$$

where $||\cdot||$ is some permutation and sign change invariant metric. In the family produced by affine transformations of such random vectors shape is still a meaningful and well defined concept, as for any such transformation all shape matrices will by definition be equal. Further, the spatial sign and rank based tests of sphericity will still be valid in this family as long as the test statistic is computed as given in 3.1 for the sign test statistic, and correspondingly for the other tests. Testing via direct estimation of coefficients γ_i, $i = 1, 2, 3$, as explained in the appendix, is valid only in the elliptic case.

The skew-elliptic model can be constructed as follows: Let $\boldsymbol{y}^* = (y_0, \boldsymbol{y})$ be a $(p+1)$-variate and spherical random vector around the origin. Write

$$\boldsymbol{z} = sign(y_0 - \alpha - \beta y_1)\boldsymbol{y}.$$

Then

$$\boldsymbol{x} = \Sigma^{-1/2}\boldsymbol{z} + \mu$$

has a skew-elliptic distribution. For skew-elliptical distributions, see e.g. Liu and Dey (2004). If $\mu = \mathbf{0}$ is known then any scatter matrix with respect to the origin satisfying $S(s\boldsymbol{x}) = S(\boldsymbol{x})$ for all random signs s, possibly depending on $\boldsymbol{x}$, is proportional to Σ. Under the skew elliptic model such shape matrices thus estimate again the same population quantity.

For shape and scatter estimation the so called independent components or IC-model is an interesting one. This is the family of affine transformations of distributions that have independent marginals (or components). Naturally, for a random vector $\boldsymbol{z}$ with independent components and for every scatter or shape matrix S it still holds

$$S(\Omega\boldsymbol{z} + \mu) = \Omega\Lambda\Omega^T.$$

with $\Lambda = S(\boldsymbol{z})$. We say that the shape or scatter matrix has independence property if it is diagonal when ever marginals are independent. This means that although in general shape and scatter estimators do not estimate the same population quantity even up to a constant the estimators with the independence property do have a similar structure, namely, matrix Λ in the above situation is diagonal. Because of this, two such estimators can be used to find an estimate of Ω or Ω^{-1}, a solution for the so called independent component analysis (ICA). For details see Oja et al. (2006). Of the estimates considered in this paper the one based on pair-wise differences, V_2, has the independence property.

Appendix A: Proofs of the results

To prove Theorem 1, we need the following result.

Lemma 1 *Under the alternative sequence*

$$\sqrt{n}vec\left(S_1 - \frac{1}{p}I_p\right) \to_d N_{p^2}\left(\frac{1}{p+2}vec(D), \gamma_1 C_{p,p}\right)$$

where $\gamma_1 = 2/(p(p+2))$.

Proof of Lemma 1. The statistic S_1 is in the family of signed rank statistics considered in Hallin and Paindaveine (2006), and the lemma follows from

their results. However, we give a simple heuristic proof. We assume without loss of generality that $tr(D) = 0$ for ease of notation.

Write $\Lambda = I_p + n^{-1/2}D$ and $\boldsymbol{x} = \sigma\Lambda^{\frac{1}{2}}\boldsymbol{z}$. Let $\boldsymbol{u} = ||\boldsymbol{z}||^{-1}\boldsymbol{z}$. Then

$$\|\boldsymbol{x}\|^{-2}\boldsymbol{x}\boldsymbol{x}^T = \boldsymbol{u}\boldsymbol{u}^T + \frac{1}{2\sqrt{n}}(\boldsymbol{u}\boldsymbol{u}^T D + D\boldsymbol{u}\boldsymbol{u}^T) - \frac{1}{\sqrt{n}}(\boldsymbol{u}^T D\boldsymbol{u})\boldsymbol{u}\boldsymbol{u}^T + o\left(\frac{1}{\sqrt{n}}\right)$$

uniformly in $\boldsymbol{z}$. Let next $\boldsymbol{z}_1, ...\boldsymbol{z}_n$ be a random sample from a spherically symmetric distribution, $\boldsymbol{u}_i = ||\boldsymbol{u}_i||^{-1}\boldsymbol{u}_i$ and $\boldsymbol{x}_i = \sigma\Lambda^{\frac{1}{2}}\boldsymbol{z}_i$, $i = 1, ..., n$. One easily sees that (Tyler, 1987a)

$$\sqrt{n}\text{vec}\left(\text{ave}\{\boldsymbol{u}_i\boldsymbol{u}_i^T\} - \frac{1}{p}I_p\right) \rightarrow_d N(\boldsymbol{0}, \gamma_1 C_{p,p}).$$

As

$$E\left[(\boldsymbol{u}_i^T D\boldsymbol{u}_i)\boldsymbol{u}_i\boldsymbol{u}_i^T\right] = \frac{2}{p(p+2)}\ D,$$

the law of large numbers gives

$$\text{vec}\left(\text{ave}\left\{\frac{1}{2}(\boldsymbol{u}_i\boldsymbol{u}_i^T D + D\boldsymbol{u}_i\boldsymbol{u}_i^T) - (\boldsymbol{u}_i^T D\boldsymbol{u}_i)\boldsymbol{u}_i\boldsymbol{u}_i^T\right\}\right)$$
$$\rightarrow_p \frac{1}{p+2}\text{vec}(D)$$

and the result follows.

Proof of Theorem 1. Using Lemma 1,

$$n(C_{p,p}\text{vec}(S_1))^T(\gamma_1^{-1}C_{p,p}^-)(C_{p,p}\text{vec}(S_1)) = \frac{n}{\gamma_1}||C_{p,p}S_1||^2$$

has a limiting noncentral chi square distribution with $(p+2)(p-1)/2$ degrees of freedom and noncentrality parameter $||C_{p,p}D||^2/(\gamma_1(p+2)^2)$, and the result follows. See also Rao (1965, section 3b.4).

Proof of Theorem 2. Tyler (1987) showed that in the spherical case, the limiting distribution of $\sqrt{n}\,vec(\widehat{V}_0 - I_p)$, where $\widehat{V}_0$ is standardized using the trace, is multivariate normal with mean zero and covariance matrix

$$(p+2)^2\gamma_1 C_{p,p} \tag{9}$$

It is straightforward to see that $\sqrt{n}\,vec(\widehat{V}_1 - I_p)$ (where the standardization is done with the determinant) has the same limiting distribution. Let $X = (\boldsymbol{x}_1 \; ... \; \boldsymbol{x}_n)$ be a random sample form a spherical distribution. Now, due to affine equivariance of $\hat{V}_1$, in the elliptical case $V^{1/2}X$ with $det(V) = 1$ and true shape V,

$$\begin{aligned}\sqrt{n}\,vec(\widehat{V}_1(V^{1/2}X) - V)) &= \sqrt{n}\,vec\,(V^{1/2}\widehat{V}_1(X)V^{1/2} - V) \\ &= (V^{1/2} \otimes V^{1/2})\sqrt{n}\,vec(\widehat{V}_1(X) - I_p)\end{aligned}.$$

Using (9) and the properties of vec-operator and Kronecker product given in Magnus and Neudecker (1979), the limiting variance of $\sqrt{n}\,vec(\widehat{V}_1(V^{1/2}X) - V))$ reduces to

$$\begin{aligned}&(p+2)^2\gamma_1(V^{1/2} \otimes V^{1/2})\left(\frac{1}{2}(I_{p^2} + K_{p,p}) - \frac{1}{p}vec(I_p)vec^T(I_p)\right)(V^{1/2} \otimes V^{1/2}) \\ &= (p+2)^2\gamma_1\left(\frac{1}{2}(I_{p^2} + K_{p,p})(V \otimes V) - \frac{1}{p}vec(V)vec(V)^T\right).\end{aligned}$$

To prove Theorem 3, we need the following result.

Lemma 2 *Under the null hypothesis*

$$\sqrt{n}\,vec\left(S_2 - \frac{1}{p}I_p\right) \to_d N_{p^2}\left(\mathbf{0}, \gamma_2 C_{p,p}\right).$$

Proof of Lemma 2. First note that $\mathrm{vec}(S_2)$ is a U-statistic

$$\mathrm{vec}(S_2) = \mathrm{vec}\left(\underset{i<j}{\mathrm{ave}}\{\boldsymbol{u}_{ij}\boldsymbol{u}_{ij}{}^T\}\right) = \binom{n}{2}^{-1}\sum_{1\le i<j\le n}\sum \boldsymbol{h}(\boldsymbol{x}_i, \boldsymbol{x}_j),$$

where

$$\boldsymbol{h}(\boldsymbol{x}_1, \boldsymbol{x}_2) = \mathrm{vec}\left(\frac{(\boldsymbol{x}_1 - \boldsymbol{x}_2)(\boldsymbol{x}_1 - \boldsymbol{x}_2)^T}{\|\boldsymbol{x}_1 - \boldsymbol{x}_2\|^2}\right)$$

is symmetric in its arguments. The limiting distribution of such a U-statistic is (see e.g. the Appendix in Lehmann (1998) for the univariate result)

$$\sqrt{n}(\boldsymbol{U} - \boldsymbol{\theta}) \to_d N(0, 4E[\boldsymbol{h}_2(\boldsymbol{x}_2)\boldsymbol{h}_2(\boldsymbol{x}_2)^T]),$$

where $\boldsymbol{\theta} = E[\boldsymbol{h}(\boldsymbol{x}_1, \boldsymbol{x}_2)]$ and $\boldsymbol{h}_2(\boldsymbol{x}) = E_{\boldsymbol{x}_1}[\boldsymbol{h}(\boldsymbol{x}_1, \boldsymbol{x})] - \boldsymbol{\theta}$.

Under the null hypothesis the expectation $\boldsymbol{\theta}$ is equal to $p^{-1}\text{vec}(I_p)$. We next find a formula for the limiting covariance matrix. For the covariance matrix under the assumption of sphericity (around origin, for ease of notation) write $\boldsymbol{z}_1$ and $\boldsymbol{z}_2$ instead of $\boldsymbol{x}_1$ and $\boldsymbol{x}_2$ for clarity, and consider first a case where $\boldsymbol{z}_2 = r\boldsymbol{e}_1 = (r, 0, \ldots, 0)^T$:

$$
\begin{aligned}
&E_{\boldsymbol{z}_1}\left[\frac{(\boldsymbol{z}_1 - r\boldsymbol{e}_1)(\boldsymbol{z}_1 - r\boldsymbol{e}_1)^T}{\|\boldsymbol{z}_1 - r\boldsymbol{e}_1\|^2}\right] \\
&= E_{\boldsymbol{z}_1}\left[\frac{((\boldsymbol{z}_1)_1 - r)^2}{\|\boldsymbol{z}_1 - r\boldsymbol{e}_1\|^2}\right]\boldsymbol{e}_1\boldsymbol{e}_1^T + E_{\boldsymbol{z}_1}\left[\frac{(\boldsymbol{z}_1)_2^2}{\|\boldsymbol{z}_1 - r\boldsymbol{e}_1\|^2}\right](I_p - \boldsymbol{e}_1\boldsymbol{e}_1^T) \\
&= \left(1 - pE_{\boldsymbol{z}_1}\left[\frac{(\boldsymbol{z}_1)_2^2}{\|\boldsymbol{z}_1 - r\boldsymbol{e}_1\|^2}\right]\right)\boldsymbol{e}_1\boldsymbol{e}_1^T + E_{\boldsymbol{z}_1}\left[\frac{(\boldsymbol{z}_1)_2^2}{\|\boldsymbol{z}_1 - r\boldsymbol{e}_1\|^2}\right]I_p.
\end{aligned}
$$

Here the first equation follows from the sphericity of $\boldsymbol{z}_1$, implying that all off-diagonal elements are equal to zero, and the second from the fact that the trace of the matrix in question is one. In the following an arbitrary $\boldsymbol{z}_2$ is decomposed as $r\boldsymbol{u}$ where $r = \|\boldsymbol{z}_2\|$ and $\boldsymbol{u} = r^{-1}\boldsymbol{z}_2$. Now, there exists an orthogonal matrix A such that $A\boldsymbol{e}_1 = \boldsymbol{u}$. Again because of the sphericity of $\boldsymbol{z}_1$ it holds that

$$
\begin{aligned}
&E_{\boldsymbol{z}_1}\left[\frac{(\boldsymbol{z}_1 - \boldsymbol{z}_2)(\boldsymbol{z}_1 - \boldsymbol{z}_2)^T}{\|\boldsymbol{z}_1 - \boldsymbol{z}_2\|^2}\right] = E_{\boldsymbol{z}_1}\left[\frac{A(\boldsymbol{z}_1 - r\boldsymbol{e}_1)(A(\boldsymbol{z}_1 - r\boldsymbol{e}_1))^T}{\|A(\boldsymbol{z}_1 - r\boldsymbol{e}_1)\|^2}\right] \\
&= \left(1 - pE_{\boldsymbol{z}_1}\left[\frac{(\boldsymbol{z}_1)_2^2}{\|\boldsymbol{z}_1 - r\boldsymbol{e}_1\|^2}\right]\right)\boldsymbol{u}\boldsymbol{u}^T + E_{\boldsymbol{z}_1}\left[\frac{(\boldsymbol{z}_1)_2^2}{\|\boldsymbol{z}_1 - r\boldsymbol{e}_1\|^2}\right]I_p.
\end{aligned}
$$

The covariance matrix can now be found using the matrix derived above and the covariance matrix of $\text{vec}(\boldsymbol{u}\boldsymbol{u}^T)$. We obtain

$$
Cov(vec(S_2)) = \frac{\gamma_2}{n}C_{p,p} + o(\frac{1}{n}),
$$

where $\gamma_2 = 4E[h_2^2(r)]$ with

$$
\gamma_2 = \frac{8}{p(p+2)}E_r\left[\left(1 - pE_{\boldsymbol{z}_1}\left[\frac{(\boldsymbol{z}_1)_2^2}{\|\boldsymbol{z}_1 - r\boldsymbol{e}_1\|^2}\right]\right)^2\right], \tag{10}
$$

where $r = \|\boldsymbol{z}\|$ and $\boldsymbol{z}$ and $\boldsymbol{z}_1$ are independent observations from a null distribution with symmetry centre $\boldsymbol{0}$.

Proof of Theorem 3. As in Lemma 1 one can first show that

$$\sqrt{n}\mathrm{vec}\left(\underset{i<j}{\mathrm{ave}}\{\boldsymbol{u}_{ij}\boldsymbol{u}_{ij}^T\} - \frac{1}{p}I_p\right) \to_d N\left(\frac{1}{p+2}vec(D), \gamma_2 C_{p,p}\right). \qquad (11)$$

The proof is then similar to that of Theorem 1.

Proof of Theorem 4. The results in Dümbgen (1998) imply that, if $\widehat{V}_2$ satisfies $tr(\widehat{V}_2) = p$ then, in the spherical case,

$$\sqrt{n}\,\mathrm{vec}(\widehat{V}_2 - I_p) \to_d N(\mathbf{0}, (p+2)^2\gamma_2 C_{p,p}),$$

with γ_2 is as in Theorem 3. In the general elliptic case with shape matrix V the limiting distribution of $\widehat{V}_2$ satisfying $det(\widehat{V})_2) = 1$ may be obtained as in the proof of Theorem 2.

To prove Theorem 5, we use the following two lemmas.

Lemma 3 *Under the null hypothesis*

$$\sqrt{n}C_{p,p}vec(S_3) \to_d N_{p^2}\left(\mathbf{0}, \gamma_3 C_{p,p}\right).$$

Lemma 4 *Under the alternative sequence*

$$\sqrt{n}C_{p,p}vec(S_3) \to_d N_{p^2}\left(\frac{c_F^2}{p+2}vec(D), \gamma_3 C_{p,p}\right).$$

Proof of Lemma 3. The proof is similar to that of Lemma 2. To see that $\mathrm{vec}(S_3)$ is asymptotically equivalent (up to a constant) to a U-statistic note first that

$$\begin{aligned}\mathrm{vec}(S_3) &= \mathrm{vec}\left(\frac{1}{n^3}\sum_i\sum_j\sum_k \boldsymbol{u}_{ij}\boldsymbol{u}_{ik}^T\right) \\ &= \mathrm{vec}\left(\frac{1}{n^3}\left(\sum_i\sum_{j\neq i}\sum_{k\neq i,j}\boldsymbol{u}_{ij}\boldsymbol{u}_{ik}^T\right)\right) + \mathrm{vec}\left(\frac{1}{n^3}\left(\sum_i\sum_{j\neq i}\boldsymbol{u}_{ij}\boldsymbol{u}_{ij}^T\right)\right).\end{aligned}$$

(Note that $2n(n-1)+n$ terms of the full sum are in fact equal to zero.) The second term is $o_P(n^{-1})$. The first term can in turn be written as

$$\frac{\binom{n}{3}}{n^3}\left(\binom{n}{3}^{-1}\sum_{i<j<k}\boldsymbol{h}(\boldsymbol{x}_i,\boldsymbol{x}_j,\boldsymbol{x}_k)\right)=\frac{\binom{n}{3}}{n^3}\boldsymbol{U}$$

where

$$\boldsymbol{h}(\boldsymbol{x}_1,\boldsymbol{x}_2,\boldsymbol{x}_3)=\text{vec}\left(\boldsymbol{u}_{12}\boldsymbol{u}_{13}^T+\boldsymbol{u}_{13}\boldsymbol{u}_{12}^T+\boldsymbol{u}_{21}\boldsymbol{u}_{23}^T+\ldots+\boldsymbol{u}_{32}\boldsymbol{u}_{31}^T\right)$$

A generalization of the result used in the proof of Lemma 2 states that

$$\sqrt{n}(\boldsymbol{U}-\boldsymbol{\theta})\rightarrow_d N(0,9E[\boldsymbol{h}_3(\boldsymbol{x})\boldsymbol{h}_3(\boldsymbol{x})^T]),$$

where $\boldsymbol{\theta}=E[\boldsymbol{h}(\boldsymbol{x}_1,\boldsymbol{x}_2,\boldsymbol{x}_3)]$ and $\boldsymbol{h}_3(\boldsymbol{x})=E_{\boldsymbol{x}_2,\boldsymbol{x}_3}[\boldsymbol{h}(\boldsymbol{x},\boldsymbol{x}_2,\boldsymbol{x}_3)]-\boldsymbol{\theta}$. To find $\boldsymbol{\theta}$ under the null hypothesis (again assuming that the symmetry center is the origin), note that due to the spherical symmetry $E[\boldsymbol{u}_{12}\boldsymbol{u}_{13}^T]$ is proportional to I_p and its trace is equal to $E\left[\boldsymbol{u}_{12}^T\boldsymbol{u}_{13}\right]=c_F^2$. The same holds for other five permutations as well and so $\boldsymbol{\theta}=p^{-1}6c_F^2vec(I_p)$.

We then find a formula for the covariance matrix under the null hypothesis. Write again $\boldsymbol{z}$, $\boldsymbol{z}_2$ and $\boldsymbol{z}_3$ instead of $\boldsymbol{x}$, $\boldsymbol{x}_2$ and $\boldsymbol{x}_3$ as in the proof of Theorem 2, and denote first

$$G_1(\boldsymbol{z})=E_{\boldsymbol{z}_2,\boldsymbol{z}_3}\left[\frac{(\boldsymbol{z}-\boldsymbol{z}_2)(\boldsymbol{z}-\boldsymbol{z}_3)^T}{\|\boldsymbol{z}-\boldsymbol{z}_2\|\|\boldsymbol{z}-\boldsymbol{z}_3\|}\right]$$

and

$$G_2(\boldsymbol{z})=E_{\boldsymbol{z}_2,\boldsymbol{z}_3}\left[\frac{(\boldsymbol{z}_2-\boldsymbol{z})(\boldsymbol{z}_2-\boldsymbol{z}_3)^T}{\|\boldsymbol{z}_2-\boldsymbol{z}\|\|\boldsymbol{z}_2-\boldsymbol{z}_3\|}\right].$$

With this notation $\boldsymbol{h}_3(\boldsymbol{z})=2\text{vec}(G_1(\boldsymbol{z})+2G_2(\boldsymbol{z}))-\boldsymbol{\theta}$. Like in the proof of Lemma 2 consider first the case of $\boldsymbol{z}_1=r\boldsymbol{e}_1$ and denote

$$g_1(r)=E_{\boldsymbol{z}_2,\boldsymbol{z}_3}\left[\frac{(r-(\boldsymbol{z}_2)_1)(r-(\boldsymbol{z}_3)_1)}{\|r\boldsymbol{e}_1-\boldsymbol{z}_2\|\|r\boldsymbol{e}_1-\boldsymbol{z}_3\|}\right]+2E_{\boldsymbol{z}_2,\boldsymbol{z}_3}\left[\frac{((\boldsymbol{z}_2)_1-r)(\boldsymbol{z}_2-\boldsymbol{z}_3)_1}{\|\boldsymbol{z}_2-r\boldsymbol{e}_1\|\|\boldsymbol{z}_2-\boldsymbol{z}_3\|}\right],$$

$$g_2(r)=E_{\boldsymbol{z}_2,\boldsymbol{z}_3}\left[\frac{(\boldsymbol{z}_2)_2(\boldsymbol{z}_3)_2}{\|\boldsymbol{z}_2-r\boldsymbol{e}_1\|\|r\boldsymbol{e}_1-\boldsymbol{z}_3\|}\right]+2E_{\boldsymbol{z}_2,\boldsymbol{z}_3}\left[\frac{(\boldsymbol{z}_2)_2(\boldsymbol{z}_2-\boldsymbol{z}_3)_2}{\|\boldsymbol{z}_2-r\boldsymbol{e}_1\|\|\boldsymbol{z}_2-\boldsymbol{z}_3\|}\right].$$

Now

$$G_1(r\boldsymbol{e}_1)+2G_2(r\boldsymbol{e}_1)=g_1(r)\boldsymbol{e}_1\boldsymbol{e}_1^T+g_2(r)(I_p-\boldsymbol{e}_1\boldsymbol{e}_1^T)$$

which reduces to $(g_1(r)-g_2(r))(\boldsymbol{e}_1\boldsymbol{e}_1^T+g_2(r)I_p)$. As in the proof of Lemma 2

$$G_1(\boldsymbol{z})+2G_2(\boldsymbol{z}))=(g_1(r)-g_2(r))\boldsymbol{u}\boldsymbol{u}^T+g_2(r)I_p$$

for arbitrary $\boldsymbol{z}=r\boldsymbol{u}$. This leads to

$$E\left[\boldsymbol{h}_3(\boldsymbol{z})\boldsymbol{h}_3(\boldsymbol{z})^T\right]=\frac{8}{p(p+2)}E[(g_1(r)-g_2(r))^2]C_{p,p}$$

and finally, since $\binom{n}{3}/n^3\to\frac{1}{6}$, as $n\to\infty$, it holds that

$$\sqrt{n}\left(\mathrm{vec}(S_3)-\boldsymbol{\theta}\right)\to_d N_{p^2}\left(\mathbf{0},\frac{2}{p(p+2)}E[(g_1(r)-g_2(r))^2]C_{p,p}\right)$$

and further

$$\sqrt{n}C_{p,p}\mathrm{vec}(S_3)\to_d N_{p^2}\left(\mathbf{0},\frac{2}{p(p+2)}E[(g_1(r)-g_2(r))^2]C_{p,p}\right)$$

Proof of Lemma 4. Proceeding as in the proof of Lemma 1 one has that

$$\begin{aligned}
&\sqrt{n}\,\mathrm{vec}\left(\operatorname*{ave}_{i,j,k}\left\{\frac{\boldsymbol{x}_{ij}\boldsymbol{x}_{ik}^T}{\|\boldsymbol{x}_{ij}\|\|\boldsymbol{x}_{ik}\|}-\frac{\boldsymbol{x}_{ij}^T\boldsymbol{x}_{ik}}{\|\boldsymbol{x}_{ij}\|\|\boldsymbol{x}_{ik}\|}\frac{1}{p}I_p\right\}\right)\\
&=\sqrt{n}\,\mathrm{vec}\left(\operatorname*{ave}_{i,j,k}\left\{\boldsymbol{u}_{ij}\boldsymbol{u}_{ik}^T-\boldsymbol{u}_{ij}^T\boldsymbol{u}_{ik}\frac{1}{p}I_p\right\}\right)+\frac{1}{2}\mathrm{vec}\left(\operatorname*{ave}_{i,j,k}\left\{\boldsymbol{u}_{ij}\boldsymbol{u}_{ik}^TD+D\boldsymbol{u}_{ij}\boldsymbol{u}_{ik}^T\right.\right.\\
&\quad-\boldsymbol{u}_{ij}\boldsymbol{u}_{ik}^T\cdot\boldsymbol{u}_{ik}^TD\boldsymbol{u}_{ik}-\boldsymbol{u}_{ij}\boldsymbol{u}_{ik}^T\cdot\boldsymbol{u}_{ij}^TD\boldsymbol{u}_{ij}-p^{-1}(2\boldsymbol{u}_{ij}^TD\boldsymbol{u}_{ik}+\boldsymbol{u}_{ij}^T\boldsymbol{u}_{ik}\cdot\boldsymbol{u}_{ik}^TD\boldsymbol{u}_{ik}\\
&\quad\left.\left.+\boldsymbol{u}_{ij}^T\boldsymbol{u}_{ik}\cdot\boldsymbol{u}_{ij}^TD\boldsymbol{u}_{ij})I_p\right\}\right)+o_p(1)
\end{aligned}$$

where $\boldsymbol{u}_{ij}=\boldsymbol{U}(\boldsymbol{z}_i-\boldsymbol{z}_j)$ and $\boldsymbol{x}_{ij}=\sigma\Lambda(\boldsymbol{z}_i-\boldsymbol{z}_j)$, with $\boldsymbol{z}_i$ being spherical. Now, for any symmetric D, it holds that

$$E\left[\boldsymbol{u}_{ij}\boldsymbol{u}_{ik}^T\cdot\boldsymbol{u}_{ik}^TD\boldsymbol{u}_{ik}\right]=\frac{c_F^2}{p(p+2)}(2D+tr(D)I_p)\tag{12}$$

and

$$E\left[\boldsymbol{u}_{ij}^T\boldsymbol{u}_{ik}\cdot\boldsymbol{u}_{ik}^TD\boldsymbol{u}_{ik}\right]=\frac{c_F^2}{p}tr(D).\tag{13}$$

These can be proved using

$$E\left[\frac{\boldsymbol{z}_1-\boldsymbol{z}_2}{\|\boldsymbol{z}_1-\boldsymbol{z}_2\|}\middle|\frac{\boldsymbol{z}_1-\boldsymbol{z}_3}{\|\boldsymbol{z}_1-\boldsymbol{z}_3\|}\right]=c_F^2\frac{\boldsymbol{z}_1-\boldsymbol{z}_3}{\|\boldsymbol{z}_1-\boldsymbol{z}_3\|},$$

which in turn can be proved as follows. The fact that

$$E\left[\boldsymbol{u}|\boldsymbol{v}\right]=c\boldsymbol{v}$$

where $\boldsymbol{u}=\boldsymbol{U}(\boldsymbol{z}_1-\boldsymbol{z}_2)$ and $\boldsymbol{v}=\boldsymbol{U}(\boldsymbol{z}_1-\boldsymbol{z}_3)$, for some c, follows from spherical symmetry by first considering the case $\boldsymbol{v}=\boldsymbol{e}_1$ and a matrix $D=diag(1,\pm1,\ldots,\pm1)$. It then holds that

$$E[\boldsymbol{u}|\boldsymbol{v}=\boldsymbol{e}_1]=E[D\boldsymbol{u}|D\boldsymbol{v}=\boldsymbol{e}_1]=E[D\boldsymbol{u}|\boldsymbol{v}=\boldsymbol{e}_1]=D\,E[\boldsymbol{u}|\boldsymbol{v}=\boldsymbol{e}_1]$$

and so $E[\boldsymbol{u}|\boldsymbol{v}=\boldsymbol{e}_1]=c\boldsymbol{e}_1$ for some c. The case of a general $\boldsymbol{v}$ is obtained by considering an orthogonal matrix A such that $A\boldsymbol{v}=\boldsymbol{e}_1$ and noting that $A\boldsymbol{u}\sim\boldsymbol{u}$. After this the fact that $c=c_F^2$ is obtained by

$$\frac{c_F^2}{p}I_p=E(\boldsymbol{u}\boldsymbol{v}^T)=E\left[E\left[\boldsymbol{u}\boldsymbol{v}^T|\boldsymbol{v}\right]\right]=cE\left[\boldsymbol{v}\boldsymbol{v}^T\right]=\frac{c}{p}I_p.$$

Proof of Theorem 5. This may be proved using Lemma 4 and proceeding as before.

Proof of Theorem 6. In the null hypothesis case, the limiting distribution of $C_{p,p}vec(S_4)$ is obtained as in the proof of Lemma 3. The kernel of the equivalent U-statistic is more complicated as it includes terms with signs of pairwise sums as well and also the ones including both sums and differences. However, due to symmetry, the conditional expectation of that kernel is the same as for the Spearman's rho covariance matrix, therefore the limiting distribution is the same. The limiting distribution under the alternative sequence as well as the the distribution of the test statistic therefore follows.

Appendix B: Estimation of γ_2 and γ_3

The coefficient γ_2 can be estimated from the data in a straightforward manner by taking the defining formula and replacing the expectations by averages:

$$\widehat{\gamma}_2=\frac{8}{p(p+2)}\frac{1}{n}\sum_i\left(1-\frac{p}{n-1}\sum_{j\neq i}\left(\frac{(\boldsymbol{x}_j)_2^2}{\|\boldsymbol{x}_j-\|\boldsymbol{x}_i\|\boldsymbol{e}_1\|^2}\right)\right)^2.$$

However, this requires the location of $\boldsymbol{x}_i$'s to be the origin while the test statistic itself does not. Also it is not affine invariant and is valid only in the elliptic case.

Other possibility is to use the covariance matrix estimator given by (7) and the approximation

$$n\,\widehat{Cov}(vec(S_2)) \approx \gamma_2 C_{p,p}.$$

This could be done for example by summing up the elements of $n\,\widehat{Cov}(vec(S_2))$ corresponding to non-zero elements of $C_{p,p}$ and dividing by $p(p-1)$, the sum of elements of $C_{p,p}$. This result then serves as $\widehat{\gamma}_2$.

For γ_3 the idea is the same. The estimate by averaging is

$$\widehat{\gamma}_3 = \frac{2}{p(p+2)}\frac{1}{n}\sum_i \left[\left(\frac{1}{n^2}\sum_{j,k}\left[\frac{(r_i-(\boldsymbol{x}_j)_1)(r_i-(\boldsymbol{x}_k)_1)-(\boldsymbol{x}_j)_2(\boldsymbol{x}_k)_2}{\|r_i\boldsymbol{e}_1-\boldsymbol{x}_j\|\|r_i\boldsymbol{e}_1-\boldsymbol{x}_k\|}\right.\right.\right.$$
$$\left.\left.\left. + 2\frac{((\boldsymbol{x}_j)_1-r_i)(\boldsymbol{x}_j-\boldsymbol{x}_k)_1-(\boldsymbol{x}_j)_2(\boldsymbol{x}_j-\boldsymbol{x}_k)_2}{\|\boldsymbol{x}_j-r_i\boldsymbol{e}_1\|\|\boldsymbol{x}_j-\boldsymbol{x}_k\|}\right]\right)^2\right],$$

where r_i stands for $\|\boldsymbol{x}_i\|$. The estimation by covariance matrix is as with γ_2.

References

Croux, C., Ollila, E. and Oja, H. (2002). Sign and rank covariance matrices: Statistical properties and application to principal component analysis. In *Statistical Data Analysis based on the L_1 Norm and Related Methods*, 257–269.

Dümbgen, L. (1998), On Tyler's M-functional of scatter in high dimension, *Annals of Institute Statistical Mathematics*, **50**, 471–491.

Dümbgen, L. and Tyler, D. (2005), On the Breakdown Properties of Some Multivariate M-Functionals, *Scandinavian Journal of Statistics*, **32**, 247–264.

Hallin, M. and Paindaveine, D. (2006), Semiparametrically efficient rank-based inference for shape. I. Optimal rank-based tests for sphericity, *Annals of Statistics*, **34**, 2707–2756.

Hallin, M. Oja, H., and Paindaveine, D. (2006), Semiparametrically efficient rank-based inference for shape. II. Optimal R-estimation of shape, *Annals of Statistics*, **34**, 2757–2789.

Hettmansperger, T.P. and Randles, R.H. (2002), A practical affine equivariant multivariate median, *Biometrika*, **89**, 851–860.

John, S. (1971), Some optimal multivariate tests, *Biometrika*, **58**, 123–127.

John, S. (1972), The distribution of a statistic used for testing sphericity of normal distributions, *Biometrika*, **59**, 169–173.

Lehmann, E.L. (1998), *Nonparametrics: Statistical methods based on ranks*, Prentice-Hall, New Jersey.

Liu, J. and Dey, K.D. (2004) Skew-elliptical distributions. In: *Skew-Elliptical Distributions and their Applications: A Journey beyond Normality*, 43–64, ed. by Genton, M., Chapman and Hall, New York.

Locantore, N., Marron, J.S., Simpson, D.G., Tripoli, N., Zhang, J.T., and Cohen, K.L. (1999), Robust principal components for functional data. *Test*, **8**, 1–73.

Magnus, J.R. and Neudecker, H. (1979), The commutation matrix: some properties and applications. *Annals of Statistics*, **7** 381–394.

Marden, J.I. (1999), Some robust estimates of principal components, *Statistics and Probability Letters*, **43**, 349–359.

Mauchly, J.W. (1940), Significance test for sphericity of a normal n-variate distribution, *Annals of Mathematical Statistics*, **11**, 204–209.

Muirhead, R.J. and Waternaux, C.M. (1980), Asymptotic distributions in canonical correlation analysis and other multivariate procedures for nonnormal populations, *Biometrika*, **67**, 31–43.

Möttönen, J. and Oja, H. (1995). Multivariate spatial sign and rank methods, *Journal of Nonparametric Statistics*, **5**, 201–213.

Möttönen J., Oja, H. and Tienari J. (1997), On the efficiency of multivariate spatial sign and rank tests, *Annals of Statistics*, **25**, 542–552.

Möttönen, J., Hüsler, J. and Oja, H. (2003), Multivariate nonparametric tests in randomized complete block design, *Journal of Multivariate Analysis*, **85**, 106–129.

Nagao, H. (1973), On some test criteria for covariance matrix, *Annals of Statistics*, **1**, 700–709.

Oja, H. and Randles, R. (2004), Multivariate nonparametric tests, *Statistical Science*, **19**, 598–605.

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

Paindaveine, D. (2007), A canonical definition of shape. *Statistics and Probability Letters*, in press, available online.

Randles, R.H. (2000), A simpler, affine equivariant multivariate, distribution-free sign test, *Journal of the American Statistical Association*, **95**, 1263–1268.

Rao, C.R. (1965), *Linear Statistical Inference and Its Applications*, Wiley, New York.

Sirkiä, S., Taskinen, S. and Oja H. (2007) Symmetrised M-estimators of multivariate scatter. *Journal of Multivariate Analysis*, **98**, 1611–1629

Sugiura, N. (1972), Locally best invariant test for sphericity and the limiting distributions, *Annals of Mathematical Statistics*, **43**, 1312–1316.

Tyler, D.E. (1982), Radial estimates and the test for the sphericity, *Biometrika*, **69**, 429–436.

Tyler, D.E. (1983), Robustness and efficiency properties of scatter matrices, *Biometrika*, **70**, 411–420.

Tyler, D.E. (1987a), A distribution-free M-estimator of multivariate scatter, *Annals of Statistics*, **15**, 234–251.

Tyler, D.E. (1987b), Statistical analysis for the angular central Gaussian distribution on the sphere, *Biometrika*, **74**, 579–589.

Visuri, S., Oja, H. and Koivunen, V. (2000), Sign and rank covariane matrices, *Journal of Statistical Planning and Inference*, **91**, 557–575.

Visuri, S., Oja; H. and Koivunen, V. (2001). Subspace-based direction of arrival estimation using nonparametric statistics. *IEEE Transactions on Signal Processing*, **49**, 2060-2073.