New spatial sign tests for inference on signal dimension

Gaspard Bernard[1,2],

[1] *Department of Mathematics, University of Luxembourg, e-mail:* gaspard.bernard@uni.lu

[2] *Institute of Statistical Science, Academia Sinica*

Abstract: We consider triangular arrays of i.i.d. elliptical random vectors and study the Null hypothesis of equality of the last $(p-q)$ eigenvalues of the (normalized) scatter parameter. To attain robustness to heavy tails, and hence validity in the whole class of elliptical distributions, we consider procedures based on the spatial signs of the observations. We show that the existing spatial sign test exhibit very low asymptotic power when the qth and $(q+1)$th eigenvalues are too close to each other, which is highly problematic when trying to separate a signal from some spherically-distributed noise. We therefore consider two types of alternatives: (i) the qth and $(q+1)$th eigenvalues are well separated but the last $(p-q)$ ones are not equal, (ii) the qth and $(q+1)$th eigenvalues are too close to each other. We propose new spatial sign tests that are robust to heavy tails and display the same local asymptotic power as that of the classical spatial sign test in case (i) and arbitrarily large asymptotic power in case (ii), making these tests strictly better than the existing spatial sign test. We show how our new sign tests can be used to construct robust estimators of the dimension of heavy-tailed signals and also discuss how the proposed procedure can be used in a directional data framework.

MSC2020 subject classifications: Primary 62H15, 62F35; secondary 62F05.

Keywords and phrases: robust inference, shape matrix, hypothesis test, dimension reduction, elliptical model, spatial signs, directional data, latent roots, subsphericity, heavy tails.

1. Introduction

In a vast amount of classical multivariate models, it is common to consider collections of i.i.d. p-variate random vectors whose distribution is characterized by a *location* parameter $\boldsymbol{\theta}$ and a *scatter* parameter $\boldsymbol{\Sigma}$. Under certain symmetry assumptions, $\boldsymbol{\Sigma}$ is a symmetric positive-definite $p \times p$ matrix summarizing the dispersion of the population and performing *Principal Component Anlysis* (PCA) on $\boldsymbol{\Sigma}$ is therefore a very natural way to perform dimension reduction. The most classical case of such a model is that of multivariate Gaussian random vectors, where $\boldsymbol{\theta}$ is the expectation and $\boldsymbol{\Sigma}$ the covariance matrix. The *elliptical model*, where the characteristic function of the random vectors is assumed to be of the form $\mathbf{t} \mapsto e^{i\mathbf{t}'\boldsymbol{\theta}}\phi(\mathbf{t}'\boldsymbol{\Sigma}\mathbf{t})$ for some *characteristic generator* $\phi : \mathbb{R}^+ \to \mathbb{R}$, can be viewed as a more general model where strong symmetry–the probability contours are ellipses here–still makes it possible to characterize the dispersion through $\mathbf{\Sigma}$. Note that no moment assumption is required for $\mathbf{\Sigma}$ to be well-defined, allowing to perform PCA on $\mathbf{\Sigma}$ even under heavy-tailed assumptions where the covariance matrix does not exist. Given that many problems–such as the one studied in this contribution–are invariant with respect to scale transformations of the data, it is often more convenient to define a *shape* parameter $\mathbf{V} = \det(\mathbf{\Sigma})^{-1/p}\mathbf{\Sigma}$ and express the inference problem in terms of this matrix $\mathbf{V}$. Note that we know from [16] that this specific choice for the normalization of $\mathbf{V}$ is canonical in the elliptical framework. A large class of testing problems over $\mathbf{V}$ is particularly relevant when performing PCA-based dimension reduction techniques. In general, PCA makes use of the spectral decomposition $\mathbf{V} = \boldsymbol{\beta}\mathbf{\Lambda}_{\mathbf{V}}\boldsymbol{\beta}'$ of $\mathbf{V}$, where $\boldsymbol{\beta}$ belongs to the *special orthogonal group* $\mathcal{SO}_p$ and $\mathbf{\Lambda}_{\mathbf{V}}$ is a diagonal matrix with ordered diagonal elements $\lambda_{1,\mathbf{V}}, \ldots, \lambda_{p,\mathbf{V}}$. The objective of PCA-based dimension reduction is to project the data onto a space of dimension q, $0 < q < p$, spanned by the q eigenvectors in $\boldsymbol{\beta}$ that are associated with the largest eigenvalues. In this context, the *subsphericity* problem, i.e., the problem of testing for the equality of the last $(p-q)$ eigenvalues is very important. In the Gaussian case, a classical Likelihood Ratio Test (LRT) for $\mathcal{H}_{0q} : \lambda_{q,\mathbf{V}} > \lambda_{q+1,\mathbf{V}} = \ldots = \lambda_{p,\mathbf{V}}$ was introduced in [10]. In general, these tests are used successively for $\mathcal{H}_{0q}$ with $q = p-2, p-3, \ldots, 0$ with the convention $\lambda_{0,\mathbf{V}} = \lambda_{1,\mathbf{V}} + 1$, the aim being to identify the one k such that $\mathcal{H}_{0k}$ is not rejected. Indeed, if $\lambda_{k,\mathbf{V}} > \lambda_{k+1,\mathbf{V}} = \ldots = \lambda_{p,\mathbf{V}}$ holds, we can consider that the last $(p-k)$ principal components contain only spherical noise. In this situation, the dataset can be separated into a k-dimensional *signal* and some spherically-distributed noise that can be safely removed. The strategy of testing for $\mathcal{H}_{0q}$, $q = p-2, p-3, \ldots, 0$ has been initially proposed by [1] in a factor analysis context. In this context, the factors associated with the smallest equal eigenvalues are considered non-significant and no conclusion should be drawn from these since there is no consistent estimators of the last eigenvectors. This Bartlett procedure is still relevant today, and is used for example in [28] and [5] in very different domains as a preliminary step to factor analysis.

In recent years, the subsphericity problem and its ties to estimation of the dimension of a signal has been the subject of ongoing study; see [14], [13] and [15]. High-dimensional tests have been studied in [21], [7] and [24]. All these testing procedures have one common drawback: their asymptotic powers tend to be dramatically low under alternatives to $\mathcal{H}_{0q}$ of the type $\lambda_{q,\mathbf{V}} = \lambda_{q+1,\mathbf{V}} = \ldots = \lambda_{p,\mathbf{V}}$. Note that this is particularly problematic in practice when trying to estimate the dimension of the signal k. Indeed, assuming that $k < p-2$, successive testing for $\mathcal{H}_{0q}$, $q = p-2, p-3, \ldots, 0$ under $\mathcal{H}_{0k}$ will irremediably produce a situation where the data-generating process is under such an alternative. In [2], a *Gaussian* triangular array framework where the eigenvalues $\lambda_{i,\mathbf{V}}^{(n)}$, $i = 1, \ldots, p$ are depending on n has been considered, allowing to study the asymptotic behavior of the classical tests in scenarios where $\lambda_{q,\mathbf{V}}^{(n)} - \lambda_{q+1,\mathbf{V}}^{(n)} = o(1)$ as $n \to \infty$. In the aforementioned article, a new test based on the Gaussian LRT and the power-enhancement approach developed by [6], performing better in scenarios where $\lambda_{q,\mathbf{V}}^{(n)} - \lambda_{q+1,\mathbf{V}}^{(n)} = o(1)$, has been proposed. However, this strategy was proposed in a purely Gaussian framework, so that the procedures proposed in [2] are not robust to the presence of heavy tails. This lack of robustness can be partially solved by a pseudo-Gaussian correction in the spirit of [26], but this method requires finite moments of order 4. The aim of this contribution is to use the same power-enhancement strategy as in [2], but considering robust test statistics allowing for the new procedures to be asymptotically valid in the whole elliptical setting, irrespectively of the lack of finite fourth-order moments. A very natural way of obtaining robust tests is to use of the *spatial signs* of the observations when constructing the test statistic; that is using only the direction $\frac{\mathbf{X}_i-\boldsymbol{\theta}}{\|\mathbf{X}_i-\boldsymbol{\theta}\|}$ associated with each observation $\mathbf{X}_i$. In general, multivariate sign tests are elegantly simple, easy to compute, robust to heavy tails, remarkably robust in high-dimensional scenarios (see [19] and [27]) and can be applied to a certain class of directional data. In the context of PCA, multivariate sign tests have already been studied in [9], [18], [3]. More important, multivariate sign test for subsphericity have been proposed amongst other procedures in [4], even if we will show latter that this test lacks power against alternatives of the type $\lambda_{q,\mathbf{V}} = \lambda_{q+1,\mathbf{V}} = \ldots = \lambda_{p,\mathbf{V}}$. We will propose testing procedures based on the spatial sign that are asymptotically valid in the whole elliptical framework and show strong asymptotic powers with respect to the alternatives considered in [2] where $\lambda^{(n)}_{q,\mathbf{V}} - \lambda^{(n)}_{q+1,\mathbf{V}} = o(1)$ as $n \to \infty$. This point is discussed in more detail in Section 2. The new spatial sign tests will also be asymptotically valid in a directional data framework where we make the assumption that the data-generating process is *angular Gaussian* (see [22]). The aim is to combine power improvement in certain problematic situations where $\lambda^{(n)}_{q,\mathbf{V}} - \lambda^{(n)}_{q+1,\mathbf{V}} = o(1)$ as $n \to \infty$ and robustness to heavy tails in a single test statistic. Finally, we will use these new multivariate sign-based tests to construct robust estimators of the dimension of the signal, mimicking the construction of [15], and evaluate its performances through numerical comparison with the estimators based on the classical spatial sign test.

2. Elliptical model, null hypothesis and notations

We start the section by defining some notations that will be used throughout this contribution. First, if $\mathbf{e}_\ell$ is the ℓth vector of the canonical basis of $\mathbb{R}^p$, let $\mathbf{K}_p := \sum_{i,j=1}^p (\mathbf{e}_i\mathbf{e}_j')\otimes(\mathbf{e}_j\mathbf{e}_\mathbf{i}')$, the *commutation matrix*. For any matrix $\mathbf{A}$, let vec$(\mathbf{A})$ the vector obtained by stacking its columns on top of each others. We also define $\mathbf{J}_p := \mathrm{vec}\,(\mathbf{I}_p)\mathrm{vec}'\,(\mathbf{I}_p)$, with $\mathbf{I}_\ell$ the ℓ-dimensional identity matrix. Denoting by $\mathrm{dvec}\,(\mathbf{A}) =: (\mathbf{A}_{11}, \overset{\circ}{\mathrm{dvec}}\,'(\mathbf{A}))'$, the p-dimensional vector obtained by stacking the diagonal elements of $\mathbf{A}$, we let $\mathbf{H}_p$ be the $p\times p^2$ matrix such that $\mathbf{H}_p\mathrm{vec}\,(\mathbf{A}) = \mathrm{dvec}\,(\mathbf{A})$. Note that $\mathbf{H}_p\mathbf{H}_p' = \mathbf{I}_p$ and that if $\mathbf{A}$ is diagonal, then $\mathbf{H}_p'\mathrm{dvec}\,(\mathbf{A}) = \mathrm{vec}\,(\mathbf{A})$. Finally, we write diag$(\mathbf{B}_1, \ldots, \mathbf{B}_m)$ for the block-diagonal matrix with blocks $\mathbf{B}_1, \ldots, \mathbf{B}_m$ and $\mathbf{A}^{\otimes 2} := \mathbf{A}\otimes\mathbf{A}$, the classical Kronecker product between $\mathbf{A}$ and $\mathbf{A}$. For a symmetric and positive definite matrix $\mathbf{B}$, we will denote as $\mathbf{B}^{1/2}$ its symmetric and positive definite square root, as $\mathbf{B}^{-1/2}$ the inverse of its square root and as $\mathbf{B}^-$ its generalized Moore-Penrose inverse.

We will now examine the considered (elliptical) model in more details. Let us consider a triangular array $\mathbf{X}_{n1}, \ldots, \mathbf{X}_{nn}$ of p-variate i.i.d. random vectors with a common centred elliptical distribution with scatter parameter $\mathbf{\Sigma}^{(n)}$. Let

$$\boldsymbol{\beta}\mathbf{\Lambda}_{\mathbf{V}}^{(n)}\boldsymbol{\beta}' = \mathbf{V}^{(n)} := (\sigma^{(n)})^{-2}\,\mathbf{\Sigma}^{(n)}$$

with *scale* parameter $\sigma^{(n)} \in \mathbb{R}_0^+$ such that $\det(\mathbf{V}^{(n)}) = 1$, $\boldsymbol{\beta} \in \mathcal{SO}_p$ and with

$$\mathbf{\Lambda}_{\mathbf{V}}^{(n)} := \mathrm{diag}(\lambda_{1,\mathbf{V}}^{(n)}, \ldots, \lambda_{p,\mathbf{V}}^{(n)}),$$

a converging sequence of $p \times p$ diagonal matrices with ordered eigenvalues. If we assume that $\mathbf{X}_{n1}, \ldots, \mathbf{X}_{nn}$ are absolutely continuous with respect to the Lebesgue measure, the vectors $\mathbf{X}_{n1}, \ldots, \mathbf{X}_{nn}$ admit a common density function $f_{\mathbf{X}_{n1}}^{(n)} : \mathbb{R}^p \to \mathbb{R}^+$ of the form

$$f_{\mathbf{X}_{n1}}^{(n)}(\mathbf{x}) = c_{\sigma^{(n)},f_1}\; f_1\big(\frac{1}{\sigma^{(n)}}(\mathbf{x}'(\mathbf{V}^{(n)})^{-1}\mathbf{x})^{1/2}\big), \tag{1}$$

with f_1 belonging to the class of *standardized radial functions* $\mathcal{F}_1$ (see [8] for more details), and $c_{\sigma^{(n)},f_1} \in \mathbb{R}_0^+$ a normalization constant.

The case of unknown location is a straightforward extension of the centred case considered here, since it is well known that the location parameter can be estimated at no asymptotic cost under ellipticity, as shown by [8]. We will then restrict the model considered, without loss of generality, to densities of the form (1). In this model, the parameter of interest is the vector of the eigenvalues deprived of its first component, vector that we denote $\mathrm{d\overset{\circ}{v}ec}\,(\mathbf{\Lambda}_{\mathbf{V}}^{(n)})$. We will consider the sequence of testing problems characterized by the following null hypotheses, introduced in [2] in a Gaussian framework. Letting $q \geq 0$, the sequence of null hypotheses considered are

$$\mathcal{H}_{0q}^{(n)} : (\lambda_{q+1,\mathbf{V}}^{(n)} = \ldots = \lambda_{p,\mathbf{V}}^{(n)}) \cap (\sqrt{n}(\lambda_{q,\mathbf{V}}^{(n)} - \lambda_{q+1,\mathbf{V}}^{(n)}) \to \infty \text{ as } n \to \infty), \tag{2}$$

with $\lambda_{0,\mathbf{V}}^{(n)}$ defined such that $\sqrt{n}(\lambda_{0,\mathbf{V}}^{(n)} - \lambda_{1,\mathbf{V}}^{(n)}) \to \infty$ as $n \to \infty$. $\mathcal{H}_{0q}^{(n)}$ is both a standard subsphericity hypothesis concerning the last $(p-q)$ eigenvalues (the smallest $(p-q)$ latent roots are equal) and a separation assumption ($\lambda_{q,\mathbf{V}}^{(n)}$ and $\lambda_{q+1,\mathbf{V}}^{(n)}$ are separated in such a way that $\sqrt{n}(\lambda_{q,\mathbf{V}}^{(n)} - \lambda_{q+1,\mathbf{V}}^{(n)}) \to \infty$ as $n \to \infty$). It is logical to consider the separation part of the hypothesis $\mathcal{H}_{0q}^{(n)}$ as this specific root-n separation rate is necessary for consistent estimation of the q and $(q+1)$th eigenvectors, and then for any projection over the first q principal components to be carried out. For this reason, the Null (2) fully characterizes scenarios in which the data can be separated into a well-identified signal that can be consistently estimated and some spherical noise. For this reason, the alternatives to $\mathcal{H}_{0q}^{(n)}$ studied in [2] were classified by the authors into two distinct types:

(i) *alternatives of type* I, under which the last $(p-q)$ eigenvalues are not equal while $\sqrt{n}(\lambda_{q,\mathbf{V}}^{(n)} - \lambda_{q+1,\mathbf{V}}^{(n)}) \to \infty$ as $n \to \infty$;

(ii) *alternatives of type* II, under which $\lambda^{(n)}_{q+1,\mathbf{V}} = \ldots = \lambda^{(n)}_{p,\mathbf{V}}$ while $\sqrt{n}(\lambda^{(n)}_{q,\mathbf{V}} - \lambda^{(n)}_{q+1,\mathbf{V}}) \to c$, $c \in \mathbb{R}$, as $n \to \infty$.

Note that type II alternatives of course include the case where

$$\lambda^{(n)}_{q,\mathbf{V}} = \lambda^{(n)}_{q+1,\mathbf{V}} = \ldots = \lambda^{(n)}_{p,\mathbf{V}},$$

which was described in the Introduction as problematic when trying to estimate the dimension of the signal. We will study the asymptotic behavior under type I and type II alternatives of the classical multivariate sign test for $\mathcal{H}^{(n)}_{0q}$. First, we need to define $\forall i = 1, \ldots, n$, $n = 1, 2, \ldots$, the vectors of the spatial signs of the observations $\mathbf{U}_{ni} := \mathbf{X}_{ni}/\|\mathbf{X}_{ni}\|$. Under the elliptical assumption (1), the triangular array $\mathbf{U}_{n1}, \ldots, \mathbf{U}_{nn}$ is known to be angular Gaussian (see [22]), with shape parameter $\mathbf{V}^{(n)}$. In other words, $\mathcal{H}^{(n)}_{0q}$ is the exact same null hypothesis in the angular Gaussian model as in the elliptical model. Any asymptotic property of a test in the $\mathbf{U}_{n1}, \ldots, \mathbf{U}_{nn}$ model can then be transferred to a corresponding spatial sign test in the $\mathbf{X}_{n1}, \ldots, \mathbf{X}_{nn}$ model. In particular, all conclusions made in terms of asymptotic validity or local asymptotic powers in the angular Gaussian framework are still valid in the elliptical framework. In the following, we will therefore carry out the asymptotic analysis in the angular Gaussian model because of its simpler nature and the way in which it allows to treat directional data.

It should be noted that the interpretation of the null hypothesis of subsphericity in the presence of directional data is slightly different from that in the classical elliptical framework. As mentioned in sections 7 and 8 of [22], the projection on the first q principal components still captures the highest possible amount of variance, the only difference with the elliptical case being that the projected random vectors will be supported on a unit ball in dimension q. The second difference is that invariance with respect to the orthogonal transform for directional data implies uniformity. In particular, the Null (2) can be interpreted as the uniformity of the angular components with respect to the $(p-q)$ (well-identified) directions containing the least variance.

3. The angular Gaussian model and the classical spatial sign test statistic

As indicated in Section 2, we consider without loss of generality the limits of triangular arrays of angular Gaussian random vectors, entirely characterized by the shape parameter $\mathbf{V}^{(n)}$. Recall that a p-variate random vector $\mathbf{U} \in \mathcal{S}^{p-1} := \{\mathbf{u} \in \mathbb{R}^p, \mathbf{u}'\mathbf{u} = 1\}$ is said to follow an angular Gaussian distribution with shape parameter $\mathbf{V}^{(n)}$ if its density $f^{(n)}_{\mathbf{U}} : \mathcal{S}^{p-1} \to \mathbb{R}^+$ (with respect to the surface measure on $\mathcal{S}^{p-1}$) is of the form

$$f^{(n)}_{\mathbf{U}}(\mathbf{u}) = \frac{\Gamma(p/2)}{2\pi^{p/2}\det(\mathbf{V}^{(n)})^{1/2}}(\mathbf{u}'(\mathbf{V}^{(n)})^{-1}\mathbf{u})^{-p/2}, \tag{3}$$

with $\Gamma(k)$, $k \in \mathbb{C}$, representing the Euler gamma function.

Let $\mathbf{U}_{n1}, \dots \mathbf{U}_{nn}$, a triangular array of angular Gaussian vectors having densities (3) and denote this assumption by $\mathrm{P}^{(n)}_{\boldsymbol{\beta},\boldsymbol{\Lambda}^{(n)}_{\mathbf{V}}}$ (recall here that $\mathbf{V}^{(n)} = \boldsymbol{\beta}\boldsymbol{\Lambda}^{(n)}_{\mathbf{V}}\boldsymbol{\beta}'$). We will consider the natural multivariate sign-based test proposed in [4] (see that reference for in-depth discussions of the asymptotic properties of this test in a non-triangular array context where $\lambda_{q,\mathbf{V}} > \lambda_{q+1,\mathbf{V}}$), but first we need to introduce some new objects. Let $\boldsymbol{\beta}_{0q} := (\boldsymbol{\beta}_{q+1}, \dots, \boldsymbol{\beta}_p)$ be the $p\times(p-q)$ matrix of the last $(p-q)$ eigenvectors of $\mathbf{V}$. Let $\hat{\mathbf{V}}^{(n)}_{\mathrm{Tyler}}$ be the classical M-estimator for shape of [23] having determinant 1, eigenvalues $\hat{\boldsymbol{\Lambda}}_{\mathrm{Tyler}} = \mathrm{diag}(\hat{\lambda}_{1,\mathrm{Tyler}}, \dots, \hat{\lambda}_{p,\mathrm{Tyler}})$ and eigenvectors $\hat{\boldsymbol{\beta}}_{\mathrm{Tyler}} = (\hat{\boldsymbol{\beta}}_1, \dots, \hat{\boldsymbol{\beta}}_p)$. Let $\hat{\mathbf{V}}^{(n)} = \tilde{\mathbf{V}}^{(n)}/\det^{1/p}(\tilde{\mathbf{V}}^{(n)})$ a constrained (i.e., belonging the the Null) estimator of $\mathbf{V}^{(n)}$ with $\tilde{\mathbf{V}}^{(n)}$ defined as

$$\tilde{\mathbf{V}}^{(n)} := \sum_{j=1}^{q} \hat{\lambda}_{j,\mathrm{Tyler}} \hat{\boldsymbol{\beta}}_j \hat{\boldsymbol{\beta}}_j' + \underline{\hat{\lambda}}_{\mathbf{V}} \sum_{j=q+1}^{p} \hat{\boldsymbol{\beta}}_j \hat{\boldsymbol{\beta}}_j',$$

where $\underline{\hat{\lambda}}_{\mathbf{V}} := (p-q)^{-1} \sum_{j=q+1}^{p} \hat{\lambda}_{j,\mathrm{Tyler}}$. Note that even if the Tyler shape estimator was initially defined for random vectors belonging to $\mathbb{R}^p$, it has been shown in [22] that for angular Gaussian random vectors, the asymptotic distribution of this estimator is the same as in the elliptical case. We finally define

$$\mathbf{S}^{(n)}_{\mathrm{sign}} := \mathbf{S}^{(n)}_{\mathrm{sign}}(\mathbf{V}^{(n)}) = \frac{1}{n} \sum_{i=1}^{n} \frac{(\mathbf{V}^{(n)})^{-1/2} \mathbf{U}_{ni} \mathbf{U}'_{ni} (\mathbf{V}^{(n)})^{-1/2}}{\mathbf{U}'_{ni} (\mathbf{V}^{(n)})^{-1} \mathbf{U}_{ni}}$$

and

$$\hat{\mathbf{S}}^{(n)}_{\mathrm{sign}} := \mathbf{S}^{(n)}_{\mathrm{sign}}(\hat{\mathbf{V}}^{(n)}).$$

Letting $d(p,q) := (p-q+2)(p-q-1)/2$ and in accordance with the sign-based approach outlined earlier, we consider the sign test $\phi^{(n)}_{\mathrm{sign}}$ introduced in [4], which rejects the null hypothesis $\mathcal{H}^{(n)}_{0q}$ at the asymptotic level $\alpha \in (0,1)$ when,

$$\begin{aligned} S^{(n)}_{q,\mathrm{sign}} &= \frac{np(p+2)}{2} (\mathrm{tr}((\hat{\boldsymbol{\beta}}'_{0q} \hat{\mathbf{S}}^{(n)}_{\mathrm{sign}} \hat{\boldsymbol{\beta}}_{0q})^2) - (p-q)^{-1} \mathrm{tr}^2 (\hat{\boldsymbol{\beta}}'_{0q} \hat{\mathbf{S}}^{(n)}_{\mathrm{sign}} \hat{\boldsymbol{\beta}}_{0q})) \\ &> \chi^2_{d(p,q);1-\alpha}. \end{aligned} \tag{4}$$

Our objective is to study the asymptotic distribution of the test statistic (4) under the Null, type I and type II alternatives and to propose for $\phi^{(n)}_{\mathrm{sign}}$ a power improvement strategy of the type proposed for the Gaussian LRT in [2], allowing to achieve non-trivial asymptotic power against type II alternatives. These type II alternatives are the exact scenarios where we will show that $\phi^{(n)}_{\mathrm{sign}}$ lacks power in Section 5.

4. Asymptotic behavior under $\mathcal{H}_{0q}^{(n)}$ and against type I alternatives

In the classical framework where $\mathbf{\Lambda_V}$ does not depend on n and where $\lambda_q \neq \lambda_{q+1}$, the local asymptotic power of the test $\phi_{\text{sign}}^{(n)}$ has already been derived in [4]. However, in our present triangular array setting, no result of such type has been obtained. In this section, we then study the asymptotic properties of the test $\phi_{\text{sign}}^{(n)}$ under $\mathcal{H}_{0q}^{(n)}$ and (local) alternatives of type I, i.e., under local alternatives to $\mathcal{H}_{0q}^{(n)}$ such that $\sqrt{n}(\lambda_{q,\mathbf{V}}^{(n)} - \lambda_{q+1,\mathbf{V}}^{(n)}) \to \infty$ as $n \ \to \infty$. More precisely, we will consider local perturbations of $\overset{\circ}{\text{dvec}}\,(\mathbf{\Lambda_V}^{(n)}) \in \mathcal{H}_{0q}^{(n)}$ of the form

$$\overset{\circ}{\text{dvec}}\,(\mathbf{\Lambda_V}^{(n)} + n^{-1/2}\text{diag}(\boldsymbol{l}^{(n)})) =: \overset{\circ}{\text{dvec}}\,(\mathbf{\Lambda_V}^{(n)}) + n^{-1/2}\boldsymbol{\tau}^{(n)}, \tag{5}$$

where $\boldsymbol{l}^{(n)} =: (l_1^{(n)}, \ldots, l_p^{(n)})$ is a bounded sequence of $\mathbb{R}^p$ such that

$$\det(\mathbf{\Lambda_V}^{(n)} + n^{-1/2}\text{diag}(\boldsymbol{l}^{(n)})) = 1$$

and

$$\prod_{j=2}^{p}(\lambda_{j,\mathbf{V}}^{(n)} + n^{-1/2}l_j^{(n)})^{-1} = \lambda_{1,\mathbf{V}}^{(n)} + n^{-1/2}l_1^{(n)} \geq \ldots \geq \lambda_{p,\mathbf{V}}^{(n)} + n^{-1/2}l_p^{(n)}.$$

Note that since, as $n \to \infty$,

$$\begin{aligned} 0 &= \det(\mathbf{\Lambda_V}^{(n)} + n^{-1/2}\text{diag}(\boldsymbol{l}^{(n)})) - \det(\mathbf{\Lambda_V}^{(n)}) \\ &= n^{-1/2}\text{tr}((\mathbf{\Lambda_V}^{(n)})^{-1}\text{diag}(\boldsymbol{l}^{(n)})) + O(n^{-1}), \end{aligned}$$

we have that $\boldsymbol{l}^{(n)}$ must be such that $\text{tr}((\mathbf{\Lambda_V}^{(n)})^{-1}\text{diag}(\boldsymbol{l}^{(n)})) = O(n^{-1/2})$ as $n \to \infty$. To study the local asymptotic powers of $\phi_{\text{sign}}^{(n)}$, we consider $\overset{\circ}{\text{dvec}}\,(\mathbf{\Lambda_V}^{(n)}) \in \mathcal{H}_{0q}^{(n)}$ and sequences of local alternatives of the form

$$\text{P}^{(n)}_{\boldsymbol{\beta},\mathbf{\Lambda_V}^{(n)} + n^{-1/2}\text{diag}(\boldsymbol{l}^{(n)})}, \tag{6}$$

with $\boldsymbol{l}^{(n)} \in \mathbb{R}^p$ a bounded sequence satisfying (5). Note that under these local alternatives, the hypothesis of separation of the qth and $(q+1)$th eigenvalues $(n^{1/2}(\lambda_{q,\mathbf{V}}^{(n)} - \lambda_{q+1,\mathbf{V}}^{(n)}) \to \infty)$ is still valid. To study the asymptotic properties of $\phi_{\text{sign}}^{(n)}$, we will derive the asymptotic null distribution of the test statistic $S_{q,\text{sign}}^{(n)}$, as well as its asymptotic distribution under local alternatives of the type (6). The main tool to perform the asymptotic analysis of $S_{q,\text{sign}}^{(n)}$ under elliptical triangular array hypothesis will be a Local Asymptotic Normality (LAN) result for the eigenvalues $\overset{\circ}{\text{dvec}}\,(\mathbf{\Lambda_V}^{(n)})$. In the LAN result below, we will consider that $\mathbf{M}_p^{(n)} = \mathbf{M}_p(\overset{\circ}{\text{dvec}}\,(\mathbf{\Lambda_V}^{(n)}))$ is the $(p-1) \times p$ matrix such that

(i) $\mathbf{M}_p^{(n)}\text{dvec}\,((\mathbf{\Lambda_V}^{(n)})^{-1}) = \mathbf{0}$ and
(ii) $(\mathbf{M}_p^{(n)})'\overset{\circ}{\text{dvec}}\,(\mathbf{L}) = \text{dvec}(\mathbf{L})$ for any matrix $\mathbf{L}$ such that $\text{tr}((\mathbf{\Lambda_V}^{(n)})^{-1}\mathbf{L}) = 0$.

We can now state the Local Asymptotic Normality result.

Theorem 4.1. *Let* $\mathbf{\Lambda}_{\mathbf{V}}^{(n)}$ *be a sequence of diagonal matrices with* $\lim_{n\to\infty}\mathbf{\Lambda}_{\mathbf{V}}^{(n)} =: \mathbf{\Lambda}_{\mathbf{V}}$ *and let* $\boldsymbol{l}^{(n)} \in \mathbb{R}^p$ *a bounded sequence satisfying* (5) *with*

$$\boldsymbol{\tau}^{(n)} = \overset{\circ}{\mathrm{dvec}}\,(\mathrm{diag}(\boldsymbol{l}^{(n)})).$$

Then as $n \to \infty$ *under* $\mathrm{P}^{(n)}_{\boldsymbol{\beta},\mathbf{\Lambda}_{\mathbf{V}}^{(n)}}$,

$$\begin{aligned}\Lambda^{(n)} &= \log\Big(\frac{\mathrm{dP}^{(n)}_{\boldsymbol{\beta},\mathbf{\Lambda}_{\mathbf{V}}^{(n)}+n^{-1/2}\mathrm{diag}(\boldsymbol{l}^{(n)})}}{\mathrm{dP}^{(n)}_{\boldsymbol{\beta},\mathbf{\Lambda}_{\mathbf{V}}^{(n)}}}\Big)\\ &= (\boldsymbol{\tau}^{(n)})'\mathbf{\Delta}^{(n)}(\mathbf{\Lambda}_{\mathbf{V}}^{(n)},\boldsymbol{\beta}) - \frac{1}{2}(\boldsymbol{\tau}^{(n)})'\mathbf{\Gamma}(\mathbf{\Lambda}_{\mathbf{V}}^{(n)})\boldsymbol{\tau}^{(n)} + o_{\mathrm{P}}(1),\end{aligned}$$

with

$$\mathbf{\Delta}^{(n)}(\mathbf{\Lambda}_{\mathbf{V}}^{(n)},\boldsymbol{\beta}) = \frac{p\sqrt{n}}{2}\mathbf{M}_p^{(n)}\mathrm{dvec}((\mathbf{\Lambda}_{\mathbf{V}}^{(n)})^{-1/2}\boldsymbol{\beta}'\mathbf{S}^{(n)}_{\mathrm{sign}}\boldsymbol{\beta}(\mathbf{\Lambda}_{\mathbf{V}}^{(n)})^{-1/2})$$

and

$$\mathbf{\Gamma}(\mathbf{\Lambda}_{\mathbf{V}}^{(n)}) = \frac{p}{2(p+2)}\mathbf{M}_p^{(n)}(\mathbf{\Lambda}_{\mathbf{V}}^{(n)})^{-2}(\mathbf{M}_p^{(n)})'.$$

Moreover, $\mathbf{\Delta}^{(n)}(\mathbf{\Lambda}_{\mathbf{V}}^{(n)},\boldsymbol{\beta}) \xrightarrow{\mathcal{L}} \mathcal{N}_p(0,\mathbf{\Gamma}(\mathbf{\Lambda}_{\mathbf{V}}))$ *under* $\mathrm{P}^{(n)}_{\boldsymbol{\beta},\mathbf{\Lambda}_{\mathbf{V}}^{(n)}}$ *as* $n\to\infty$.

Following classical Le Cam Asymptotic Theory of experiment (see [11] and [12]), the distribution of the sign test statistics is now easy to derive. We first derive the asymptotic distribution of $S^{(n)}_{q,\mathrm{sign}}$ under $\mathcal{H}^{(n)}_{0q}$.

Theorem 4.2. *When* $\overset{\circ}{\mathrm{dvec}}\,(\mathbf{\Lambda}_{\mathbf{V}}^{(n)}) \in \mathcal{H}^{(n)}_{0q}$, *the test statistic* $S^{(n)}_{q,\mathrm{sign}}$ *converges weakly to the chi-square distribution with* $d(p,q)$ *degrees of freedom under* $\mathrm{P}^{(n)}_{\boldsymbol{\beta},\mathbf{\Lambda}_{\mathbf{V}}^{(n)}}$ *as* $n\to\infty$.

Theorem 4.2 means that asymptotically, the test $\phi^{(n)}_{\mathrm{sign}}$ behaves under $\mathcal{H}^{(n)}_{0q}$ exactly as in classical scenarios studied in [4] where $\mathbf{\Lambda}_{\mathbf{V}}^{(n)}$ does not depend on n and where the qth and $(q+1)$th eigenvalues are distinct from each other. The following theorem, obtained by a simple application of Le Cam's third Lemma alongside the use of Theorem 4.1, indicates that the same holds for the asymptotic behavior of $\phi^{(n)}_{\mathrm{sign}}$ under local alternatives of type I. We define $\underline{\lambda}_{\mathbf{V}}^{(n)}$, the common value of the last equal $(p-q)$ eigenvalues of $\mathbf{V}^{(n)}$ and $\lim_{n\to\infty}\underline{\lambda}_{\mathbf{V}}^{(n)} =: \underline{\lambda}_{\mathbf{V}}$.

Theorem 4.3. *Let* $\overset{\circ}{\mathrm{dvec}}\,(\mathbf{\Lambda}_{\mathbf{V}}^{(n)}) \in \mathcal{H}^{(n)}_{0q}$ *and let* $\boldsymbol{l}^{(n)} = (l_1^{(n)},\ldots,l_p^{(n)}) \in \mathbb{R}^p$ *be a bounded sequence satisfying* (5) *with* $\lim_{n\to\infty}\boldsymbol{l}^{(n)} =: \boldsymbol{l} = (l_1,\ldots,l_p)$. *Under* $\mathrm{P}^{(n)}_{\boldsymbol{\beta},\mathbf{\Lambda}_{\mathbf{V}}^{(n)}+n^{-1/2}\mathrm{diag}(\boldsymbol{l}^{(n)})}$, *the test statistic* $S^{(n)}_{q,\mathrm{sign}}$ *converges weakly as* $n\to\infty$ *to* *the noncentral chi-square distribution with $d(p,q)$ degrees of freedom and noncentrality parameter*

$$\frac{p}{2(p+2)(\underline{\lambda}_{\mathbf{V}})^2}\Big(\sum_{j=q+1}^{p}(l_j)^2-(p-q)^{-1}\big(\sum_{j=q+1}^{p}l_j\big)^2\Big).$$

The asymptotic distribution derived in Theorem 4.3 also coincides with the asymptotic distribution under local alternatives already derived in [4]. The conclusion of this section is therefore that the test $\phi^{(n)}_{\text{sign}}$ asymptotically behaves under local alternatives of type I exactly as it would in the classical non-triangular array setting where the qth and $(q+1)$th eigenvalues are distinct. In this sense, we can say that the natural sign-based test for this problem, introduced in [4], is still valid and retains all its desirable asymptotic properties under type I alternatives. As we will see in next section, $\phi^{(n)}_{\text{sign}}$ does not behave as well under type II alternatives.

5. Asymptotic behavior against type II alternatives

In a Gaussian triangular array framework, [2] showed that the LRT displays a very problematic lack of power under alternatives of type II. The aim of this section is to show that the same phenomenon is still present when studying considering $\phi^{(n)}_{\text{sign}}$ in an angular Gaussian triangular array framework. To achieve this objective, we will study the asymptotic distribution of the test statistic $S^{(n)}_{q,\text{sign}}$ when $n^{1/2}(\lambda^{(n)}_{q,\mathbf{V}}-\lambda^{(n)}_{q+1,\mathbf{V}})=O(1)$ and observe that the asymptotic power of $\phi^{(n)}_{\text{sign}}$ is below the nominal level α under type II alternatives. The main result of this section will be a proposition establishing the asymptotic distribution of $S_{q,\text{sign}}$ under any possible type II scenario.

We need to be slightly more specific and distinguish here a few cases, at this point. We consider type II alternatives characterized by

$$\mathbf{\Lambda}^{(n)}_{\mathbf{V}}=\text{diag}(\lambda^{(n)}_{1,\mathbf{V}},\dots,\lambda^{(n)}_{q,\mathbf{V}},\underline{\lambda}_{\mathbf{V}}\mathbf{1}'_{p-q})$$

with some blocks of the first q eigenvalues converging to $\underline{\lambda}_{\mathbf{V}}$ at various speeds. More precisely, we consider $0\leq s_1\leq s_2\leq s_3\leq q$ and we assume that

(i) $(\lambda^{(n)}_{j,\mathbf{V}}-\underline{\lambda}_{\mathbf{V}})=v_j>0$ for each $1\leq j\leq s_1$,
(ii) $(\lambda^{(n)}_{j,\mathbf{V}}-\underline{\lambda}_{\mathbf{V}})=o(1)$ with $n^{1/2}(\lambda^{(n)}_{j,\mathbf{V}}-\underline{\lambda}_{\mathbf{V}})\to\infty$ as $n\to\infty$, for each $s_1<j\leq s_2$,
(iii) $n^{1/2}(\lambda^{(n)}_{j,\mathbf{V}}-\underline{\lambda}_{\mathbf{V}})=v_j>0$ for each $s_2<j\leq s_3$ and
(iv) $(\lambda^{(n)}_{j,\mathbf{V}}-\underline{\lambda}_{\mathbf{V}})=o(n^{-1/2})$ as $n\to\infty$, for each $s_3<j\leq q$,

We also define the following quantity

$$\mathbf{\Theta}(v_1,\dots,v_{s_1}) \;=\; \text{diag}(v_1,\dots,v_{s_1},\underline{\lambda}_{\mathbf{V}}\mathbf{1}'_{p-s_1})^{\otimes 2}\,\frac{p+2}{p}(\mathbf{I}_{p^2}+\mathbf{K}_p+\frac{2}{p}\mathbf{J}_p), \tag{7}$$

where $\mathbf{K}_p$ and $\mathbf{J}_p$ have been defined at the beginning of Section 2. Theorem 5.1 below provides the asymptotic distribution of $S^{(n)}_{q,\mathrm{sign}}$ under any possible type II alternative.

Theorem 5.1. *Let $\mathbf{\Lambda}^{(n)}_{\mathbf{V}} = \mathrm{diag}(\lambda^{(n)}_{1,\mathbf{V}}, \ldots, \lambda^{(n)}_{q,\mathbf{V}}, \underline{\lambda}_{\mathbf{V}} \mathbf{1}'_{p-q})$ be as above and let*

$$\mathbf{Z}(v_1, \ldots, v_{s_1}) = \begin{pmatrix} \mathbf{Z}_{11} & \mathbf{Z}'_{21} \\ \mathbf{Z}_{21} & \mathbf{Z}_{22} \end{pmatrix}$$

be a $p \times p$ matrix where $\mathbf{Z}_{11}$ is the $s_2 \times s_2$ upper left block of $\mathbf{Z}(v_1, \ldots, v_{s_1})$, $\mathbf{Z}_{22}$ is the $(p-s_2) \times (p-s_2)$ lower right block of $\mathbf{Z}(v_1, \ldots, v_{s_1})$, etc, such that

$$\mathrm{vec}(\mathbf{Z}(v_1, \ldots, v_{s_1})) \sim \mathcal{N}_{p^2}(\mathbf{0}, \mathbf{\Theta}(v_1, \ldots, v_{s_1})).$$

Then as $n \to \infty$ and under $\mathrm{P}^{(n)}_{\boldsymbol{\beta}, \mathbf{\Lambda}^{(n)}_{\mathbf{V}}}$, $S^{(n)}_{q,\mathrm{sign}}$ converges weakly to

$$\frac{p}{2(p+2)\underline{\lambda}^2_{\mathbf{V}}} \Big(\sum_{j=q+1}^{p} \ell_j^2 - (p-q)^{-1} \big(\sum_{j=q+1}^{p} \ell_j \big)^2 \Big), \tag{8}$$

where $(\ell_{q+1}, \ldots, \ell_p)$ are the $p-q$ smallest roots of

$$\mathbf{Z}_{22} + \mathrm{diag}(v_{s_2+1}, \ldots, v_{s_3}, \mathbf{0}'_{q-s_3}, \mathbf{0}'_{p-q}).$$

Even if Theorem 5.1 can seem slightly hard to read, its implication is very clear. For instance, under the type II alternative such that $\mathbf{V} = \mathbf{I}_p$, we have that $s_1 = s_2 = s_3 = 0$. It is then straightforward to check that $(\ell_{q+1}, \ldots, \ell_p)$ are the smallest roots of a GOE matrix $\mathbf{Z}$ such that $\mathbf{\Theta}'\mathbf{Z}\mathbf{\Theta} \stackrel{\mathcal{D}}{=} \mathbf{Z}$, for all possible $\mathbf{\Theta} \in \mathcal{SO}_p$. It is not very surprising then that the test statistics $S_{q,\mathrm{sign}}$ tends to be smaller with high probability when p is large with respect to $(p-q)$.

To illustrate the power loss phenomenon under type II alternatives, we generated $M = 6000$ samples of sizes $n = 2000$ of $(p = 8)$-dimensional centred multivariate t-distributed random vectors with 1 degree of freedom and scatter parameter

$$\mathbf{\Sigma}^{(n)}(b, l) = \mathrm{diag}(2\, \mathbf{1}'_{p-l-2}, (1 + n^{-b})\, \mathbf{1}'_l, 1, 1).$$

The parameters b and l take values $b = 0, 1/4, 1/2, 1$ and $l = 1, \ldots, 6$. We performed the sign test $\phi^{(n)}_{\mathrm{sign}}$ based on $S^{(n)}_{q,\mathrm{sign}}$ for $\mathcal{H}^{(n)}_{06}$ $(q = 6)$, with l the number of eigenvalues converging to the last 2 ones and n^{-b} the speed of convergence. When the parameter b takes values 0 and 1/4, the data-generating processes belong to $\mathcal{H}^{(n)}_{06}$ while the values 1/2 and 1 yield data-generating processes belonging to type II alternatives. Inspection of figure 1 reveals that when l increases, the asymptotic power of the test essentially goes to 0 when $b = 1$. It is safe to say that $\phi^{(n)}_{\mathrm{sign}}$ shows essentially no power under type II alternatives. Our aim will then be, from now on, to correct this rather dramatic lack of power.

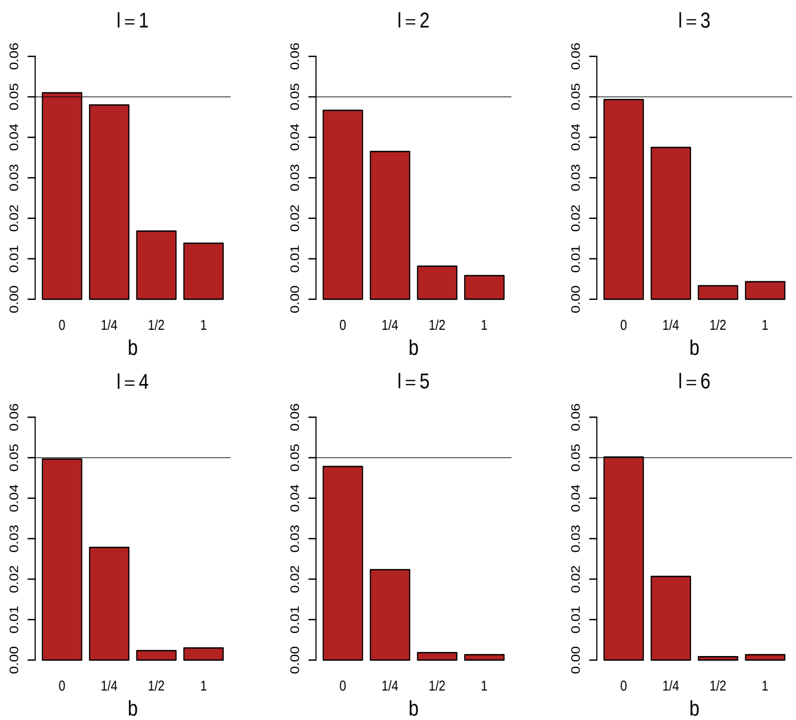


Fig 1: Empirical rejection frequencies of the test $\phi^{(n)}_{\text{sign}}$ under the null and type II alternatives, p=8 and $\mathbf{\Sigma}^{(n)}(b,l) = \text{diag}(2\ \mathbf{1}'_{p-l-2}, (1+n^{-b})\ \mathbf{1}'_{l}, 1, 1)$. All tests are performed at the asymptotic nominal level .05 with sample size $n = 2000$.

6. New spatial sign tests

The results obtained in the previous sections imply that the test $\phi_{\mathrm{sign}}^{(n)}$ achieves the same local asymptotic powers as the sign test of [4] against alternatives of type I but is (nearly) blind to alternatives of type II. We propose here a new multivariate sign test, based on the same approach as the one developed in [2] in the Gaussian framework. The new test is asymptotically equivalent to $\phi_{\mathrm{sign}}^{(n)}$ under $\mathcal{H}_{0q}^{(n)}$ and type I alternatives while showing non-trivial power against alternatives of type II. First, define $\hat{\mathbf{V}}_{q,q+1}^{(n)} = \tilde{\mathbf{V}}_{q,q+1}^{(n)}/\det^{1/p}(\tilde{\mathbf{V}}_{q,q+1}^{(n)})$, where

$$\tilde{\mathbf{V}}_{q,q+1}^{(n)} := \sum_{j=1}^{q-1} \hat{\lambda}_{j,\mathrm{Tyler}} \hat{\boldsymbol{\beta}}_j \hat{\boldsymbol{\beta}}_j' + \sum_{j=q+2}^{p} \hat{\lambda}_{j,\mathrm{Tyler}} \hat{\boldsymbol{\beta}}_j \hat{\boldsymbol{\beta}}_j' + \hat{\underline{\lambda}}_{\mathbf{V},q,q+1} \sum_{j=q}^{q+1} \hat{\boldsymbol{\beta}}_j \hat{\boldsymbol{\beta}}_j',$$

with $\hat{\underline{\lambda}}_{\mathbf{V},q,q+1} := \frac{1}{2}\sum_{j=q}^{q+1} \hat{\lambda}_{j,\mathrm{Tyler}}$. Now, letting $\hat{\mathbf{S}}_{q,q+1,\mathrm{sign}}^{(n)} := \mathbf{S}_{\mathrm{sign}}^{(n)}(\hat{\mathbf{V}}_{q,q+1}^{(n)})$ and $\hat{\boldsymbol{\beta}}_{q,q+1} := (\hat{\boldsymbol{\beta}}_q, \hat{\boldsymbol{\beta}}_{q+1})$, we will use the statistic

$$S_{q,q+1,\mathrm{sign}}^{(n)} := \frac{np(p+2)}{2}\Bigg(\mathrm{tr}\big((\hat{\boldsymbol{\beta}}_{q,q+1}' \hat{\mathbf{S}}_{q,q+1,\mathrm{sign}}^{(n)} \hat{\boldsymbol{\beta}}_{q,q+1})^2\big) - \frac{\mathrm{tr}^2\big(\hat{\boldsymbol{\beta}}_{q,q+1}' \hat{\mathbf{S}}_{q,q+1,\mathrm{sign}}^{(n)} \hat{\boldsymbol{\beta}}_{q,q+1}\big)}{2}\Bigg) \tag{9}$$

to take into account the potential deviation between $\lambda_{q,\mathbf{V}}^{(n)}$ and $\lambda_{q+1,\mathbf{V}}^{(n)}$. We can now define the new sign tests $\phi_{\mathrm{new,sign}}^{(n)}$. Letting $\alpha \in (0,1)$ and $\gamma \in (0,1)$, $\phi_{\mathrm{new,sign}}^{(n)}$ rejects $\mathcal{H}_{0q}^{(n)}$ at the asymptotic confidence level α when

$$\phi_{\mathrm{new,sign}}^{(n)} := \mathbb{I}[S_{q,\mathrm{sign}}^{(n)} > \chi^2_{d(p,q);1-\alpha}]\mathbb{I}[S_{q,q+1,\mathrm{sign}}^{(n)} > \chi^2_{2;1-\gamma}] + \mathbb{I}[S_{q,q+1,\mathrm{sign}}^{(n)} \leq \chi^2_{2;1-\gamma}] = 1. \tag{10}$$

In (10), we use the convention $S_{0,1,\mathrm{sign}}^{(n)} \equiv +\infty$. With this convention, $\phi_{\mathrm{new,sign}}^{(n)}$ and $\phi_{\mathrm{sign}}^{(n)}$ will coincide when performing the two tests for $\mathcal{H}_{00}^{(n)}$. For more details on the ideas behind this type of tests and their usage, we refer the reader to [2], [20] and [17], but the basic idea is to simultaneously test for the fact that the qth and $(q+1)$th eigenvalues are sufficiently separated and for the equality of the last $(p-q)$ eigenvalues. We have the following asymptotic result for the test $\phi_{\mathrm{new,sign}}^{(n)}$.

Theorem 6.1. *Let* $\mathrm{d}\overset{\circ}{\mathrm{vec}}\,(\boldsymbol{\Lambda}_{\mathbf{V}}^{(n)}) \in \mathcal{H}_{0q}^{(n)}$. *Under* $\mathrm{P}_{\boldsymbol{\beta},\boldsymbol{\Lambda}_{\mathbf{V}}^{(n)}}^{(n)}$, $\phi_{\mathrm{new,sign}}^{(n)} - \phi_{\mathrm{sign}}^{(n)} = o_{\mathrm{P}}(1)$ *as* $n \to \infty$.

Theorem 6.1 above shows that $\phi^{(n)}_{\text{new,sign}}$ is asymptotically valid and has same asymptotic powers as $\phi^{(n)}_{\text{sign}}$ under local alternatives of type I by contiguity. Moreover, it is very easy to show, using Theorem 5.1, that the asymptotic power of $\phi^{(n)}_{\text{new,sign}}$ under type II alternatives can be made arbitrarily large by taking γ arbitrarily small.

Then, $\phi^{(n)}_{\text{new,sign}}$ is a a strict improvement over $\phi^{(n)}_{\text{sign}}$, preserving the desirable properties of the latter test when it shows good performances while improving its power under type II scenarios. As noted in Section 2, these findings are true in the all elliptical model, and not only in the angular Gaussian setting.

To conclude this section, we present some finite n empirical powers obtained by simulation and illustrating the magnitude of the power enhancement granted by $\phi^{(n)}_{\text{new,sign}}$ in practical scenarios. We simulated $M = 1000$ independent samples $\mathbf{X}_1^{(b,\tau)}, \ldots, \mathbf{X}_n^{(b,\tau)}$ of $(p = 5)$-dimensional centred multivariate t-distributed random vectors with 1 degree of freedom. We used sample size $n = 2000$ and scatter parameter

$$\mathbf{\Sigma}^{(n)}(b,\tau) = \text{diag}(3, 1 + n^{-b}, 1 + n^{-b}, 1, 1 - n^{-1/2}\tau).$$

We used the values $\tau = 0, 1, 2, 4, 6, 8$ and $b = 0, 1/8, 1/4, 1/2, 1, 2$. The values $\tau = 0$ and $b < 1/2$ yield data-generating processes belonging to $\mathcal{H}^{(n)}_{03}$ $(q = 3)$ while all the other parameter values yield data-generating processes increasingly under the alternative. We performed the classical multivariate sign test $\phi^{(n)}_{\text{sign}}$ alongside three versions of $\phi^{(n)}_{\text{new,sign}}$ with parameter $\gamma = .9, .5, .05$. All tests are performed for $\mathcal{H}^{(n)}_{03}$ at the same asymptotic nominal level $\alpha = .05$. Inspection of figure 2 confirms our theoretical findings. The new test $\phi^{(n)}_{\text{new,sign}}$ asymptotically behaves as $\phi^{(n)}_{\text{sign}}$ when $b < 1/2$ (i.e. under the null and type I alternatives) while showing way larger power than $\phi^{(n)}_{\text{sign}}$ when $b \geq 1/2$ (i.e. under type II alternatives), even if some continuity phenomenons can of course be observed.

7. Robust sign-based estimator of the dimension of the signal

Rather naturally, we consider that the dimension of the signal k is the value $q \in \{0, \ldots, p-2\}$ for which $\mathcal{H}^{(n)}_{0q}$ holds. If $\mathcal{H}^{(n)}_{0q}$ does not hold for any $q \in \{0, \ldots, p-2\}$, we simply let $k = p - 1$. Note that the notion of signal and noise decomposition in a directional data framework readily follows from the analogy with the elliptical case.

The form of our new multivariate sign-based estimator of k is

$$\begin{aligned} \hat{k}_{\text{new,sign}} \quad = \quad & \min\{q \in \{0, \ldots, p-2\} \mid \mathbb{I}[S^{(n)}_{q,\text{sign}} > b^{(n)}_q]\mathbb{I}[S^{(n)}_{q,q+1,\text{sign}} > c^{(n)}] \\ & + \mathbb{I}[S^{(n)}_{q,q+1,\text{sign}} \leq c^{(n)}] = 0\}, \end{aligned} \tag{11}$$

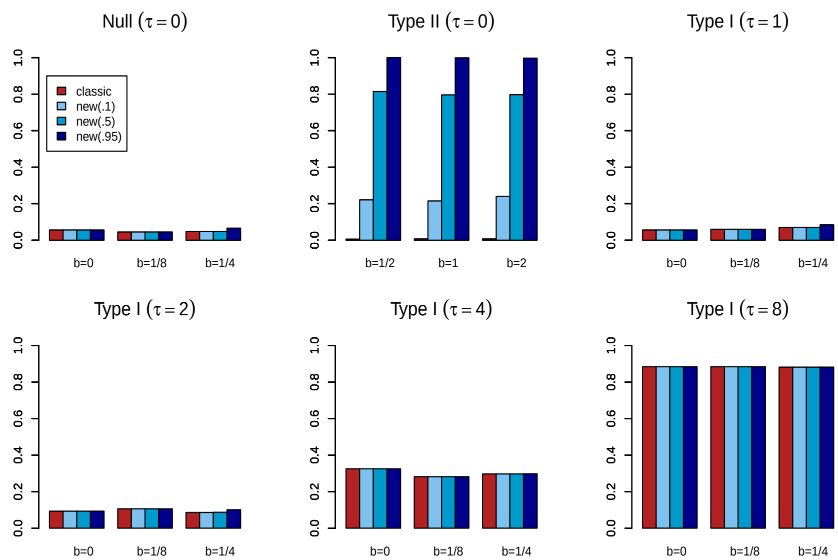


Fig 2: Empirical rejection frequencies of the tests $\phi^{(n)}_{\text{sign}}$ (in red, denoted classic) and $\phi^{(n)}_{\text{new,sign}}$ (denoted *new*) for $\gamma = .9, .5, .05$. All tests are performed at the asymptotic nominal level .05. The sample size is $n = 2000$.

for some positive sequences $b_q^{(n)}$, $q = 0, \ldots, p-2$ such that $b_q^{(n)} \to \infty$ and $b_q^{(n)} = o(n)$ as $n \to \infty$ for all q and for some other positive sequence $c^{(n)} \to \infty$ such that $c^{(n)} = o(n)$ as $n \to \infty$. These conditions are purely technical and serve only to guarantee the consistency of $\hat{k}_{\text{new,sign}}$. If the minimum is not achieved in (11), we use the definition $\hat{k}_{\text{new,sign}} = p - 1$. We compare our new estimator to its natural classical equivalent $\hat{k}_{\text{sign}}$, based on the classical test $\phi_{\text{sign}}^{(n)}$ and defined as follows

$$\hat{k}_{\text{sign}} = \min \{q \in \{0, \ldots, p-2\} \mid \mathbb{I}[S_{q,\text{sign}}^{(n)} > b_q^{(n)}] = 0\}. \tag{12}$$

Now, we assess by simulation whether the power gained by using $\phi_{\text{new,sign}}^{(n)}$ over $\phi_{\text{new}}^{(n)}$ under type II scenarios translates in better performances at finite n for the estimator $\hat{k}_{\text{new,sign}}$. For that purpose, we simulated $M = 2000$ independent samples of i.i.d. random vectors $\mathbf{X}_1^{(b,\tau^{(n)})}, \ldots, \mathbf{X}_n^{(b,\tau^{(n)})}$ with a common ($p = 3$)-dimensional centered multivariate t-distribution with 1 degree of freedom and scatter matrix

$$\mathbf{\Sigma}(b, \tau^{(n)}) = \text{diag}(1 + n^{-b}, 1, 1 - \frac{1}{2}\tau^{(n)}).$$

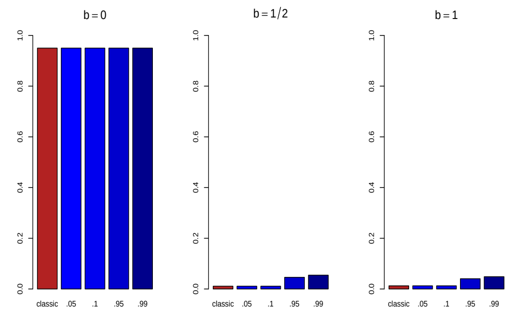
We used $\tau^{(n)} = 0, n^{-1}, n^{-1/2}, 1$ and $b = 0, \frac{1}{2}, 1$. The sample size is $n = 1000$. At each replication, we computed $\hat{k}_{\text{sign}}$ with $b_q^{(n)} = \chi^2_{d(p,q),.95}$, $q = 0, \ldots, p-2$. This specific choice for $b_q^{(n)}$ was recommended to be used in practice by [25]. We also computed 4 versions of the estimator $\hat{k}_{\text{new,sign}}$ with

$$c^{(n)} \in \{\chi^2_{2;.05}, \chi^2_{2;.1}, \chi^2_{2;.95}, \chi^2_{2;.99}\}.$$

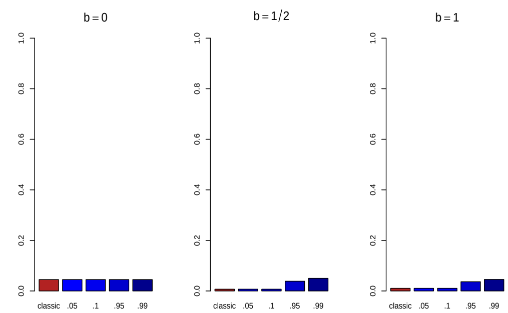
We compared the various estimators to the true value of k, i.e.

$$k = (p-1)\mathbb{I}[\tau^{(n)} > 0] + (\mathbb{I}[b < \frac{1}{2}] + (p-1)\mathbb{I}[b \geq \frac{1}{2}])\mathbb{I}[\tau^{(n)} = 0].$$

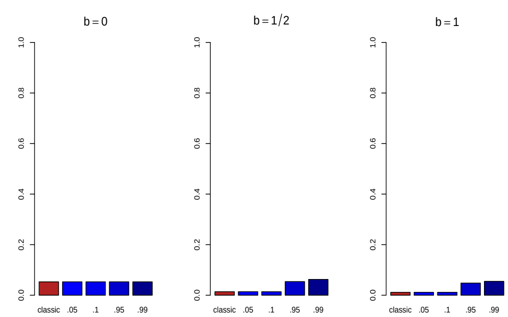
Figure 3 shows the proportion of good selections of k (i.e., $\hat{k}_{\text{sign}} = k$ or $\hat{k}_{\text{new,sign}} = k$) for each considered estimator. This shows that the new sign-based estimators perform as well as $\hat{k}_{\text{sign}}$ when $b = 0$ and outperform $\hat{k}_{\text{sign}}$ when $b > 0$. This is consistent with theory and with what was observed in the Gaussian framework in [2]. This makes of our new robust estimator an improvement over the existing ones, be it in terms of performance versus $\hat{k}_{\text{sign}}$ or robustness versus the pseudo-Gaussian procedures. It should also be noted that, as usual, if $\mathbf{X}_1^{(b,\tau^{(n)})}, \ldots, \mathbf{X}_n^{(b,\tau^{(n)})}$ had been angular Gaussian random vectors belonging to the unit sphere, every conclusions and simulation results would have been the exact same.



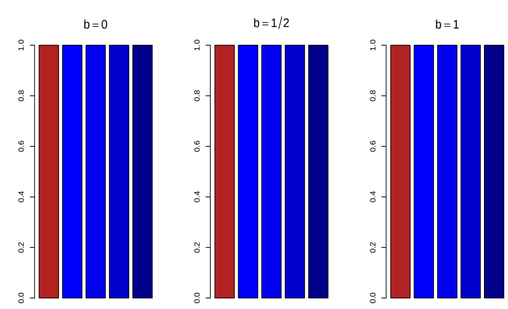
(a) $\tau^{(n)} = 0$



(b) $\tau^{(n)} = n^{-1}$



(c) $\tau^{(n)} = n^{-1/2}$



(d) $\tau^{(n)} = 1$

Fig 3: Proportion of good selection of k for the estimator $\hat{k}_{\text{sign}}$ (in red) and for four choices of estimators $\hat{k}_{\text{new,sign}}$ (in blue) with $n = 1000$, $b_q^{(n)} = \chi^2_{d(p,q),.95}$ and $c^{(n)} \in \{\chi^2_{2;.05}, \chi^2_{2;.1}, \chi^2_{2;.95}, \chi^2_{2;.99}\}$.

8. Conclusions

In this contribution, we considered the testing problem characterized by

$$\mathcal{H}_{0q}^{(n)} : (\lambda_{q+1,\mathbf{V}}^{(n)} = \ldots = \lambda_{p,\mathbf{V}}^{(n)}) \cap (n^{1/2}(\lambda_{q,\mathbf{V}}^{(n)} - \lambda_{q+1,\mathbf{V}}^{(n)}) \to \infty \text{ as } n \to \infty)$$

and studied the asymptotic behavior of the classical multivariate sign test $\phi_{\text{sign}}^{(n)}$ under elliptical triangular array assumptions. We have shown that when $n^{1/2}(\lambda_{q,\mathbf{V}}^{(n)} - \lambda_{q+1,\mathbf{V}}^{(n)}) \to \infty$ as $n \to \infty$, $\phi_{\text{sign}}^{(n)}$ behaves in our current framework exactly as in the classical non-triangular array framework studied in [4]. However, we also showed that $\phi_{\text{sign}}^{(n)}$ has dramatically low power when

$$n^{1/2}(\lambda_{q,\mathbf{V}}^{(n)} - \lambda_{q+1,\mathbf{V}}^{(n)}) = o(1).$$

We then proposed new tests $\phi_{\text{new,sign}}^{(n)}$, in the spirit of the new Gaussian tests proposed in [2], and showed that the tests $\phi_{\text{new,sign}}^{(n)}$ are retaining all the desirable asymptotic properties of $\phi_{\text{sign}}^{(n)}$ while showing arbitrarily large power when $n^{1/2}(\lambda_{q,\mathbf{V}}^{(n)} - \lambda_{q+1,\mathbf{V}}^{(n)}) = o(1)$. We also have shown through simulations that $\phi_{\text{new,sign}}^{(n)}$ can be used to estimate the dimension of a signal in such a way that the resulting estimators outperform the existing sign-based competitors at finite n.

The new spatial sign-based procedures proposed in this contribution are asymptotically valid in the elliptical distribution class and, in particular, make it possible to deal with cases where the covariance matrix does not exist. For this reason, our new tests and estimators seem particularly suitable for performing inference on the dimension of a signal when the data-generating process is suspected of being heavy-tailed.

Furthermore, it should be noted that as a by-product of the way we derived the asymptotic properties of our new tests, our procedure can be used in a directional data framework assuming an underlying angular Gaussian distribution. In this context, our procedure allows us to determine a set of $(p - q)$ directions that can be estimated consistently, whose linear span contains the smallest possible amount of variability and such that the angular components of the data with respect to these directions are independent and uniformly distributed.

For all the aforementioned reasons, the spatial sign procedures proposed in this contribution are indeed exceptionally robust and allow to perform inference on the dimension of a signal in an impressive number of scenarios. Its asymptotic distribution - both under the Null and local alternatives - is not affected by the assumptions made on the distribution of the radial parts of the elliptical random vectors considered or by the assumption that the data are restricted to the unit sphere. Thanks to our power enhancement strategy, we finally gained robustness with respect to scenarios where the qth and $(q + 1)$th eigenvalues are too close to each other for existing tests to show any significant asymptotic power. All these facts make our procedures very attractive for inference on the dimension of a signal when robustness properties are a priority.

Appendix A: Proofs of the various results

Proof of Theorem 4.1. To simplify the notations in this proof, we put

$$\mathbf{Z}_i := \boldsymbol{\beta}'\mathbf{U}_{ni}, \qquad \boldsymbol{\Lambda} := \boldsymbol{\Lambda}_{\mathbf{V}}^{(n)} \quad \text{and} \quad \boldsymbol{\Lambda}_n := \boldsymbol{\Lambda}_{\mathbf{V}}^{(n)} + n^{-1/2}\text{diag}(\boldsymbol{l}^{(n)}).$$

First,

$$\begin{aligned} \Lambda^{(n)} &= -\frac{n}{2}\big(\log(\det(\boldsymbol{\Lambda}_n)) - \log(\det(\boldsymbol{\Lambda}))\big) \\ &\quad -\frac{p}{2}\sum_{i=1}^{n}\big(\log(\mathbf{Z}_i'\boldsymbol{\Lambda}_n^{-1}\mathbf{Z}_i) - \log(\mathbf{Z}_i'\boldsymbol{\Lambda}^{-1}\mathbf{Z}_i)\big). \end{aligned} \tag{13}$$

Since $\det(\boldsymbol{\Lambda}) = 1$, we have that

$$-\frac{n}{2}\big(\log(\det(\boldsymbol{\Lambda}_n)) - \log(\det(\boldsymbol{\Lambda}))\big) = -\frac{n}{2}\log(\det(\mathbf{I}_p + n^{-1/2}\boldsymbol{\Lambda}^{-1}\text{diag}(\boldsymbol{l}^{(n)}))),$$

so that since for any $p \times p$ matrix $\mathbf{A}$,

$$\log(\det(\mathbf{I}_p + \mathbf{A})) = \text{tr}(\mathbf{A}) - \frac{1}{2}\text{tr}(\mathbf{A}^2) + o(\|\mathbf{A}\|)$$

when $\|\mathbf{A}\| \to \mathbf{0}$, we get from (13) that

$$\begin{aligned} \Lambda^{(n)} &= -\frac{\sqrt{n}}{2}\text{tr}(\boldsymbol{\Lambda}^{-1}\text{diag}(\boldsymbol{l}^{(n)})) + \frac{1}{4}\text{tr}(\boldsymbol{\Lambda}^{-2}\text{diag}^2(\boldsymbol{l}^{(n)})) \\ &\quad -\frac{p}{2}\sum_{i=1}^{n}\big(\log(\mathbf{Z}_i'\boldsymbol{\Lambda}_n^{-1}\mathbf{Z}_i) - \log(\mathbf{Z}_i'\boldsymbol{\Lambda}^{-1}\mathbf{Z}_i)\big) + o_{\mathrm{P}}(1) \end{aligned} \tag{14}$$

as $n \to \infty$ under $\mathrm{P}^{(n)}_{\boldsymbol{\beta},\boldsymbol{\Lambda}_n}$. Now, we have that

$$\begin{aligned} \log(\mathbf{Z}_i'\boldsymbol{\Lambda}_n^{-1}\mathbf{Z}_i) - \log(\mathbf{Z}_i'\boldsymbol{\Lambda}^{-1}\mathbf{Z}_i) &= \log(\mathbf{Z}_i'\boldsymbol{\Lambda}^{-1}\mathbf{Z}_i + \mathbf{Z}_i'(\boldsymbol{\Lambda}_n^{-1} - \boldsymbol{\Lambda}^{-1})\mathbf{Z}_i) \\ &\quad - \log(\mathbf{Z}_i'\boldsymbol{\Lambda}^{-1}\mathbf{Z}_i) \\ &= \log(1 + \frac{\mathbf{Z}_i'(\boldsymbol{\Lambda}_n^{-1} - \boldsymbol{\Lambda}^{-1})\mathbf{Z}_i}{\mathbf{Z}_i'\boldsymbol{\Lambda}^{-1}\mathbf{Z}_i}) \\ &= \frac{\mathbf{Z}_i'(\boldsymbol{\Lambda}_n^{-1} - \boldsymbol{\Lambda}^{-1})\mathbf{Z}_i}{\mathbf{Z}_i'\boldsymbol{\Lambda}^{-1}\mathbf{Z}_i} \\ &\quad -\frac{1}{2}\left(\frac{\mathbf{Z}_i'(\boldsymbol{\Lambda}_n^{-1} - \boldsymbol{\Lambda}^{-1})\mathbf{Z}_i}{\mathbf{Z}_i'\boldsymbol{\Lambda}^{-1}\mathbf{Z}_i}\right)^2 + R_{ni}, \end{aligned} \tag{15}$$

for some rest R_{ni}. Since

$$\mathbf{\Lambda}_n^{-1} - \mathbf{\Lambda}^{-1} = -n^{-1/2}\mathbf{\Lambda}^{-2}\mathrm{diag}(\boldsymbol{l}^{(n)}) + n^{-1}\mathbf{\Lambda}^{-3}\mathrm{diag}^2(\boldsymbol{l}^{(n)}) + O(n^{-3/2})$$

as $n \to \infty$, the boundedness of the $\mathbf{Z}_i$'s entails that

$$\begin{aligned}
\sum_{i=1}^n \frac{\mathbf{Z}_i'(\mathbf{\Lambda}_n^{-1} - \mathbf{\Lambda}^{-1})\mathbf{Z}_i}{\mathbf{Z}_i'\mathbf{\Lambda}^{-1}\mathbf{Z}_i} \quad &= \quad -n^{-1/2}\sum_{i=1}^n \frac{\mathbf{Z}_i'\mathbf{\Lambda}^{-2}\mathrm{diag}(\boldsymbol{l}^{(n)})\mathbf{Z}_i}{\mathbf{Z}_i'\mathbf{\Lambda}^{-1}\mathbf{Z}_i} \\
&\quad +n^{-1}\sum_{i=1}^n \frac{\mathbf{Z}_i'\mathbf{\Lambda}^{-3}\mathrm{diag}^2(\boldsymbol{l}^{(n)})\mathbf{Z}_i}{\mathbf{Z}_i'\mathbf{\Lambda}^{-1}\mathbf{Z}_i} + o_{\mathrm{P}}(1) \\
&= \quad -n^{-1/2}\sum_{i=1}^n \mathrm{tr}(\frac{\mathbf{Z}_i'\mathbf{\Lambda}^{-2}\mathrm{diag}(\boldsymbol{l}^{(n)})\mathbf{Z}_i}{\mathbf{Z}_i'\mathbf{\Lambda}^{-1}\mathbf{Z}_i}) \\
&\quad +n^{-1}\sum_{i=1}^n \mathrm{tr}(\frac{\mathbf{Z}_i'\mathbf{\Lambda}^{-3}\mathrm{diag}^2(\boldsymbol{l}^{(n)})\mathbf{Z}_i}{\mathbf{Z}_i'\mathbf{\Lambda}^{-1}\mathbf{Z}_i}) + o_{\mathrm{P}}(1) \\
&= \quad -n^{-1/2}\sum_{i=1}^n \mathrm{tr}(\mathrm{diag}(\boldsymbol{l}^{(n)})\frac{\mathbf{\Lambda}^{-1}\mathbf{Z}_i\mathbf{Z}_i'\mathbf{\Lambda}^{-1}}{\mathbf{Z}_i'\mathbf{\Lambda}^{-1}\mathbf{Z}_i}) \\
&\quad +n^{-1}\sum_{i=1}^n \mathrm{tr}(\mathrm{diag}(\boldsymbol{l}^{(n)})\frac{\mathbf{\Lambda}^{-3/2}\mathbf{Z}_i\mathbf{Z}_i'\mathbf{\Lambda}^{-3/2}}{\mathbf{Z}_i'\mathbf{\Lambda}^{-1}\mathbf{Z}_i}\mathrm{diag}(\boldsymbol{l}^{(n)})) \\
&\quad +o_{\mathrm{P}}(1) \\
&= \quad -n^{1/2}\mathrm{tr}(\mathrm{diag}(\boldsymbol{l}^{(n)})n^{-1}\sum_{i=1}^n \frac{\mathbf{\Lambda}^{-1}\mathbf{Z}_i\mathbf{Z}_i'\mathbf{\Lambda}^{-1}}{\mathbf{Z}_i'\mathbf{\Lambda}^{-1}\mathbf{Z}_i}) \\
&\quad +\mathrm{tr}(\mathrm{diag}(\boldsymbol{l}^{(n)})n^{-1}\sum_{i=1}^n \frac{\mathbf{\Lambda}^{-3/2}\mathbf{Z}_i\mathbf{Z}_i'\mathbf{\Lambda}^{-3/2}}{\mathbf{Z}_i'\mathbf{\Lambda}^{-1}\mathbf{Z}_i}\mathrm{diag}(\boldsymbol{l}^{(n)})) \\
&\quad +o_{\mathrm{P}}(1).
\end{aligned} \tag{16}$$

Using the definition of $\mathbf{S}_{\mathrm{sign}}^{(n)}$ alongside $\mathbf{S}_{\mathrm{sign}}^{(n)} = p^{-1}\mathbf{I}_p + o_{\mathrm{P}}(1)$, we get from (16) that

$$\begin{aligned}
\sum_{i=1}^n \frac{\mathbf{Z}_i'(\mathbf{\Lambda}_n^{-1} - \mathbf{\Lambda}^{-1})\mathbf{Z}_i}{\mathbf{Z}_i'\mathbf{\Lambda}^{-1}\mathbf{Z}_i} \quad &= \quad -n^{1/2}\mathrm{tr}(\mathrm{diag}(\boldsymbol{l}^{(n)})\mathbf{\Lambda}^{-1/2}\boldsymbol{\beta}'\mathbf{S}_{\mathrm{sign}}^{(n)}\boldsymbol{\beta}\mathbf{\Lambda}^{-1/2}) \\
&\quad +\mathrm{tr}(\mathrm{diag}(\boldsymbol{l}^{(n)})\mathbf{\Lambda}^{-1}\boldsymbol{\beta}'\mathbf{S}_{\mathrm{sign}}^{(n)}\boldsymbol{\beta}\mathbf{\Lambda}^{-1}\mathrm{diag}(\boldsymbol{l}^{(n)})) + o_{\mathrm{P}}(1) \\
&= \quad -n^{1/2}\mathrm{tr}(\mathrm{diag}(\boldsymbol{l}^{(n)})\mathbf{\Lambda}^{-1/2}\boldsymbol{\beta}'\mathbf{S}_{\mathrm{sign}}^{(n)}\boldsymbol{\beta}\mathbf{\Lambda}^{-1/2}) \\
&\quad +p^{-1}\mathrm{tr}(\mathrm{diag}(\boldsymbol{l}^{(n)})\mathbf{\Lambda}^{-2}\mathrm{diag}(\boldsymbol{l}^{(n)})) + o_{\mathrm{P}}(1).
\end{aligned} \tag{17}$$

Using the fact that for any $p \times p$ matrix $\mathbf{A}$ and diagonal $p \times p$ matrix $\mathbf{L}$, $\mathrm{tr}(\mathbf{LA}) = \mathrm{dvec}(\mathbf{L})'\mathrm{dvec}(\mathbf{A})$, we get from (17) that

$$\sum_{i=1}^{n} \frac{\mathbf{Z}_i'(\mathbf{\Lambda}_n^{-1} - \mathbf{\Lambda}^{-1})\mathbf{Z}_i}{\mathbf{Z}_i'\mathbf{\Lambda}^{-1}\mathbf{Z}_i} = -n^{1/2}\boldsymbol{l}^{(n)}\mathrm{dvec}(\mathbf{\Lambda}^{-1/2}\boldsymbol{\beta}'\mathbf{S}_{\mathrm{sign}}^{(n)}\boldsymbol{\beta}\mathbf{\Lambda}^{-1/2}) + p^{-1}\boldsymbol{l}^{(n)}\mathbf{\Lambda}^{-2}(\boldsymbol{l}^{(n)})' + o_{\mathrm{P}}(1). \quad (18)$$

Now, same Taylor expansion for the second term in (15) yields

$$\begin{aligned}
-\frac{1}{2}\sum_{i=1}^{n}\left(\frac{\mathbf{Z}_i'(\mathbf{\Lambda}_n^{-1} - \mathbf{\Lambda}^{-1})\mathbf{Z}_i}{\mathbf{Z}_i'\mathbf{\Lambda}^{-1}\mathbf{Z}_i}\right)^2 &= -\frac{1}{2n}\sum_{i=1}^{n}\left(\frac{\mathbf{Z}_i'\mathbf{\Lambda}^{-2}\mathrm{diag}(\boldsymbol{l}^{(n)})\mathbf{Z}_i}{\mathbf{Z}_i'\mathbf{\Lambda}^{-1}\mathbf{Z}_i}\right)^2 + o_{\mathrm{P}}(1)\\
&= -\frac{1}{2n}\sum_{i=1}^{n}\mathrm{tr}^2(\mathrm{diag}(\boldsymbol{l}^{(n)})\frac{\mathbf{\Lambda}^{-1}\mathbf{Z}_i\mathbf{Z}_i'\mathbf{\Lambda}^{-1}}{\mathbf{Z}_i'\mathbf{\Lambda}^{-1}\mathbf{Z}_i})\\
&\quad + o_{\mathrm{P}}(1)\\
&= -\frac{1}{2n}\sum_{i=1}^{n}\boldsymbol{l}^{(n)}\mathrm{dvec}(\frac{\mathbf{\Lambda}^{-1}\mathbf{Z}_i\mathbf{Z}_i'\mathbf{\Lambda}^{-1}}{\mathbf{Z}_i'\mathbf{\Lambda}^{-1}\mathbf{Z}_i})\\
&\quad \mathrm{dvec}'(\frac{\mathbf{\Lambda}^{-1}\mathbf{Z}_i\mathbf{Z}_i'\mathbf{\Lambda}^{-1}}{\mathbf{Z}_i'\mathbf{\Lambda}^{-1}\mathbf{Z}_i})(\boldsymbol{l}^{(n)})' + o_{\mathrm{P}}(1)\\
&= -\frac{1}{2}\mathrm{E}(\boldsymbol{l}^{(n)}\mathrm{dvec}(\frac{\mathbf{\Lambda}^{-1}\mathbf{Z}_i\mathbf{Z}_i'\mathbf{\Lambda}^{-1}}{\mathbf{Z}_i'\mathbf{\Lambda}^{-1}\mathbf{Z}_i})\\
&\quad \mathrm{dvec}'(\frac{\mathbf{\Lambda}^{-1}\mathbf{Z}_i\mathbf{Z}_i'\mathbf{\Lambda}^{-1}}{\mathbf{Z}_i'\mathbf{\Lambda}^{-1}\mathbf{Z}_i})(\boldsymbol{l}^{(n)})') + o_{\mathrm{P}}(1)\\
&= -\frac{1}{p(p+2)}\boldsymbol{l}^{(n)}\mathbf{\Lambda}^{-1}(\mathbf{I}_p + \frac{1}{2}\mathbf{1}_p\mathbf{1}_p')\mathbf{\Lambda}^{-1}(\boldsymbol{l}^{(n)})'\\
&\quad + o_{\mathrm{P}}(1). \qquad (19)
\end{aligned}$$

The last line is a well known fact (see, for example [9] and [8]). Finally note that

$$\sum_{i=1}^{n} R_{ni} = \frac{1}{3}\sum_{i=1}^{n}(1 + H_{ni})^{-3}\left(\frac{\mathbf{Z}_i'(\mathbf{\Lambda}_n^{-1} - \mathbf{\Lambda}^{-1})\mathbf{Z}_i}{\mathbf{Z}_i'\mathbf{\Lambda}^{-1}\mathbf{Z}_i}\right)^3,$$

for some H_{ni} between 0 and $\frac{\mathbf{Z}_i'(\mathbf{\Lambda}_n^{-1} - \mathbf{\Lambda}^{-1})\mathbf{Z}_i}{\mathbf{Z}_i'\mathbf{\Lambda}^{-1}\mathbf{Z}_i}$. The boundedness of the $\mathbf{Z}_i$'s implies that $|\sum_{i=1}^{n} R_{ni}| < C/n^{1/2}$ for some constant C so that $\sum_{i=1}^{n} R_{ni}$ is $o_{\mathrm{P}}(1)$ as $n \to \infty$ under $\mathrm{P}^{(n)}_{\boldsymbol{\beta},\mathbf{\Lambda}_n}$. Using this last fact and combining (15), (18) and (19), we get that, as $n \to \infty$ and under $\mathrm{P}^{(n)}_{\boldsymbol{\beta},\mathbf{\Lambda}_n}$,

$$\sum_{i=1}^{n} \log\left(\mathbf{Z}_i'\mathbf{\Lambda}_n^{-1}\mathbf{Z}_i\right) - \log\left(\mathbf{Z}_i'\mathbf{\Lambda}^{-1}\mathbf{Z}_i\right) = -n^{1/2}\boldsymbol{l}^{(n)}\mathrm{dvec}(\mathbf{\Lambda}^{-1/2}\boldsymbol{\beta}'\mathbf{S}^{(n)}_{\mathrm{sign}}\boldsymbol{\beta}\mathbf{\Lambda}^{-1/2}) + p^{-1}\boldsymbol{l}^{(n)}\mathbf{\Lambda}^{-2}(\boldsymbol{l}^{(n)})' - \frac{1}{p(p+2)}\boldsymbol{l}^{(n)}\mathbf{\Lambda}^{-1}(\mathbf{I}_p - \frac{1}{2}\mathbf{1}_p\mathbf{1}_p')\mathbf{\Lambda}^{-1}(\boldsymbol{l}^{(n)})' + o_{\mathrm{P}}(1). \tag{20}$$

Combining (14) and (20), we get that, as $n \to \infty$ and under $\mathrm{P}^{(n)}_{\boldsymbol{\beta},\,\mathbf{\Lambda}_n}$,

$$\begin{aligned}
\Lambda^{(n)} &= -\frac{n^{1/2}}{2}\boldsymbol{l}^{(n)}\mathrm{dvec}(\mathbf{\Lambda}^{-1}) + \frac{1}{4}\boldsymbol{l}^{(n)}\mathbf{\Lambda}^{-2}(\boldsymbol{l}^{(n)})' \\
&\quad + \frac{p}{2}n^{1/2}\boldsymbol{l}^{(n)}\mathrm{dvec}(\mathbf{\Lambda}^{-1/2}\boldsymbol{\beta}'\mathbf{S}^{(n)}_{\mathrm{sign}}\boldsymbol{\beta}\mathbf{\Lambda}^{-1/2}) - \frac{1}{2}\boldsymbol{l}^{(n)}\mathbf{\Lambda}^{-2}(\boldsymbol{l}^{(n)})' \\
&\quad + \frac{1}{2(p+2)}\boldsymbol{l}^{(n)}\mathbf{\Lambda}^{-1}(\mathbf{I}_p - \frac{1}{2}\mathbf{1}_p\mathbf{1}_p')\mathbf{\Lambda}^{-1}(\boldsymbol{l}^{(n)})' + o_{\mathrm{P}}(1) \\
&= \frac{p}{2}n^{1/2}\boldsymbol{l}^{(n)}\mathrm{dvec}(\mathbf{\Lambda}^{-1/2}\boldsymbol{\beta}'(\mathbf{S}^{(n)}_{\mathrm{sign}} - p^{-1}\mathbf{I}_p)\boldsymbol{\beta}\mathbf{\Lambda}^{-1/2}) \\
&\quad - \frac{p+2}{4(p+2)}\boldsymbol{l}^{(n)}\mathbf{\Lambda}^{-2}(\boldsymbol{l}^{(n)})' \\
&\quad + \frac{2}{4(p+2)}\boldsymbol{l}^{(n)}\mathbf{\Lambda}^{-1}(\mathbf{I}_p + \frac{1}{2}\mathbf{1}_p\mathbf{1}_p')\mathbf{\Lambda}^{-1}(\boldsymbol{l}^{(n)})' + o_{\mathrm{P}}(1) \\
&= \frac{p}{2}n^{1/2}\boldsymbol{l}^{(n)}\mathrm{dvec}(\mathbf{\Lambda}^{-1/2}\boldsymbol{\beta}'(\mathbf{S}^{(n)}_{\mathrm{sign}} - p^{-1}\mathbf{I}_p)\boldsymbol{\beta}\mathbf{\Lambda}^{-1/2}) \\
&\quad - \frac{p}{4(p+2)}\boldsymbol{l}^{(n)}\mathbf{\Lambda}^{-1}(\mathbf{I}_p - \frac{1}{p}\mathbf{1}_p\mathbf{1}_p')\mathbf{\Lambda}^{-1}(\boldsymbol{l}^{(n)})' + o_{\mathrm{P}}(1).
\end{aligned} \tag{21}$$

Recalling that the definition of $\boldsymbol{l}^{(n)}$ implies

$$\mathrm{tr}((\mathbf{\Lambda})^{-1}\mathrm{diag}(\boldsymbol{l}^{(n)})) = o(1)$$

as $n \to \infty$, it is easy to get that

$$(\mathbf{M}_p^{(n)})'\boldsymbol{\tau}^{(n)} = (\boldsymbol{l}^{(n)})' + o(1). \tag{22}$$

Now, using (21), (22) and properties of $\mathbf{M}_p^{(n)}$, we finally get that, as $n \to \infty$ and under $\mathrm{P}^{(n)}_{\boldsymbol{\beta},\,\mathbf{\Lambda}_n}$,

$$\begin{aligned}
\Lambda^{(n)} &= \frac{p\,n^{1/2}}{2}(\boldsymbol{\tau}^{(n)})'\mathbf{M}_p^{(n)}\mathrm{dvec}(\boldsymbol{\Lambda}^{-1/2}\boldsymbol{\beta}'(\mathbf{S}_{\mathrm{sign}}^{(n)} - p^{-1}\mathbf{I}_p)\boldsymbol{\beta}\boldsymbol{\Lambda}^{-1/2}) \\
&\quad -\frac{1}{2}(\boldsymbol{\tau}^{(n)})'\frac{p}{2(p+2)}\mathbf{M}_p^{(n)}\boldsymbol{\Lambda}^{-1}(\mathbf{I}_p - p^{-1}\mathbf{1}_p\mathbf{1}_p')\boldsymbol{\Lambda}^{-1}(\mathbf{M}_p^{(n)})'\boldsymbol{\tau}^{(n)} + o_{\mathrm{P}}(1) \\
&= \frac{p\,n^{1/2}}{2}(\boldsymbol{\tau}^{(n)})'\mathbf{M}_p\mathrm{dvec}(\boldsymbol{\Lambda}^{-1/2}\boldsymbol{\beta}'\mathbf{S}_{\mathrm{sign}}^{(n)}\boldsymbol{\beta}\boldsymbol{\Lambda}^{-1/2}) \\
&\quad -\frac{1}{2}(\boldsymbol{\tau}^{(n)})'\frac{p}{2(p+2)}\mathbf{M}_p^{(n)}\boldsymbol{\Lambda}^{-2}(\mathbf{M}_p^{(n)})'\boldsymbol{\tau}^{(n)} + o_{\mathrm{P}}(1).
\end{aligned} \tag{23}$$

Note that this is the desired result, which concludes the proof. □

The next Lemma provides the asymptotic behavior of

$$\mathbf{E}^{(n)} = \begin{pmatrix} \mathbf{E}_{11}^{(n)} & \mathbf{E}_{12}^{(n)} \\ \mathbf{E}_{21}^{(n)} & \mathbf{E}_{22}^{(n)} \end{pmatrix} := \hat{\boldsymbol{\beta}}_{\mathrm{Tyler}}'\boldsymbol{\beta},$$

where $\mathbf{E}_{11}^{(n)}$ is a $q \times q$ block, $\mathbf{E}_{22}^{(n)}$ a $(p-q)\times(p-q)$ block, etc. Its proof is similar as the one of Proposition 2 in [2] and is therefore omitted here.

Lemma A.1. *Using the above notations and letting* $\boldsymbol{\Lambda}_{1,\mathbf{V}}^{(n)} := \mathrm{diag}(\lambda_{1,\mathbf{V}}^{(n)}, \ldots, \lambda_{q,\mathbf{V}}^{(n)})$, *the following are true.*

(i) If $n^{1/2}(\lambda_{q,\mathbf{V}}^{(n)} - \lambda_{q+1,\mathbf{V}}^{(n)}) \to \infty$ *as* $n \to \infty$, *we have that*

$$n^{1/2}\mathbf{E}_{21}^{(n)}(\boldsymbol{\Lambda}_{1,\mathbf{V}}^{(n)} - \lambda_{q+1,\mathbf{V}}^{(n)}\mathbf{I}_q) = O_{\mathrm{P}}(1)$$

and

$$\mathbf{E}_{22}^{(n)}(\mathbf{E}_{22}^{(n)})' = \mathbf{I}_{p-q} + o_{\mathrm{P}}(n^{-1/2})$$

as $n \to \infty$ *under* $\mathrm{P}^{(n)}_{\boldsymbol{\beta},\boldsymbol{\Lambda}_{\mathbf{V}}^{(n)}}$;

(ii) if $n^{1/2}(\lambda_{q,\mathbf{V}}^{(n)} - \lambda_{q+1,\mathbf{V}}^{(n)}) \to c < \infty$ *as* $n \to \infty$, *we have that* $\mathbf{E}_{21}^{(n)}$ *is not* $o_{\mathrm{P}}(1)$ *as* $n \to \infty$ *under* $\mathrm{P}^{(n)}_{\boldsymbol{\beta},\boldsymbol{\Lambda}_{\mathbf{V}}^{(n)}}$.

Proof of Theorem 4.2. First,

$$\begin{aligned}
S_{q,\mathrm{sign}}^{(n)} &= \frac{np(p+2)}{2}\big(\mathrm{tr}((\hat{\boldsymbol{\beta}}_{0q}'\hat{\mathbf{S}}_{\mathrm{sign}}^{(n)}\hat{\boldsymbol{\beta}}_{0q})^2) - (p-q)^{-1}\mathrm{tr}^2(\hat{\boldsymbol{\beta}}_{0q}'\hat{\mathbf{S}}_{\mathrm{sign}}^{(n)}\hat{\boldsymbol{\beta}}_{0q})\big) \\
&= \frac{p(p+2)}{2}n^{1/2}\mathrm{vec}'(\hat{\boldsymbol{\beta}}_{0q}'\hat{\mathbf{S}}_{\mathrm{sign}}^{(n)}\hat{\boldsymbol{\beta}}_{0q})(\mathbf{I}_{(p-q)^2} - \frac{1}{p-q}\mathbf{J}_{p-q})n^{1/2}\mathrm{vec}(\hat{\boldsymbol{\beta}}_{0q}'\hat{\mathbf{S}}_{\mathrm{sign}}^{(n)}\hat{\boldsymbol{\beta}}_{0q}) \\
&= \frac{p(p+2)}{2}\mathbf{T}'(\hat{\mathbf{V}}^{(n)})\boldsymbol{\Sigma}_{p,q}\mathbf{T}(\hat{\mathbf{V}}^{(n)}),
\end{aligned} \tag{24}$$

where $\mathbf{\Sigma}_{p,q} := \mathbf{I}_{(p-q)^2} - \frac{1}{p-q}\mathbf{J}_{p-q}$ and

$$\mathbf{T}(\hat{\mathbf{V}}^{(n)}) := (\hat{\boldsymbol{\beta}}_{0q}' \otimes \hat{\boldsymbol{\beta}}_{0q}')n^{1/2}\text{vec}(\hat{\mathbf{S}}^{(n)}_{\text{sign}} - p^{-1}\mathbf{I}_p).$$

Using the same arguments as in the proof of Theorem 3.1 in [18], we have that

$$\begin{aligned} n^{1/2}\,\text{vec}\,(\hat{\mathbf{S}}^{(n)}_{\text{sign}} - \mathbf{S}^{(n)}_{\text{sign}}) &= -\frac{1}{p(p+2)}\Big[\mathbf{I}_{p^2} + \mathbf{K}_p + \mathbf{J}_p\Big]((\mathbf{V}^{(n)})^{1/2})^{\otimes 2} \\ &\quad n^{1/2}\text{vec}\,((\hat{\mathbf{V}}^{(n)})^{-1} - (\mathbf{V}^{(n)})^{-1}) \\ &\quad + \frac{n^{1/2}}{p}\Big[((\hat{\mathbf{V}}^{(n)})^{-1/2}(\mathbf{V}^{(n)})^{1/2})^{\otimes 2} - \mathbf{I}_{p^2}\Big]\text{vec}(\mathbf{I}_p) \\ &\quad + o_{\mathrm{P}}(1). \end{aligned} \tag{25}$$

as $n \to \infty$, under $\mathrm{P}^{(n)}_{\boldsymbol{\beta},\mathbf{\Lambda}^{(n)}_{\mathbf{V}}}$. Now, it follows from (25) that

$$\begin{aligned} \mathbf{\Sigma}_{p,q}\mathbf{T}^{(n)}(\hat{\mathbf{V}}^{(n)}) &= \mathbf{\Sigma}_{p,q}(\hat{\boldsymbol{\beta}}_{0q}' \otimes \hat{\boldsymbol{\beta}}_{0q}')n^{1/2}\text{vec}(\hat{\mathbf{S}}^{(n)}_{\text{sign}} - p^{-1}\mathbf{I}_p) \\ &= \mathbf{\Sigma}_{p,q}(\hat{\boldsymbol{\beta}}_{0q}' \otimes \hat{\boldsymbol{\beta}}_{0q}')n^{1/2}\text{vec}(\mathbf{S}^{(n)}_{\text{sign}} - p^{-1}\mathbf{I}_p) \\ &\quad - \mathbf{\Sigma}_{p,q}\mathbf{W}^{(n)}_1 + \mathbf{\Sigma}_{p,q}\mathbf{W}^{(n)}_2 + o_{\mathrm{P}}(1) \end{aligned} \tag{26}$$

as $n \to \infty$ under $\mathrm{P}^{(n)}_{\boldsymbol{\beta},\mathbf{\Lambda}^{(n)}_{\mathbf{V}}}$, where

$$\begin{aligned} \mathbf{W}^{(n)}_1 &:= (\hat{\boldsymbol{\beta}}_{0q}' \otimes \hat{\boldsymbol{\beta}}_{0q}')\frac{1}{p(p+2)}\Big[\mathbf{I}_{p^2} + \mathbf{K}_p + \mathbf{J}_p\Big]((\mathbf{V}^{(n)})^{1/2})^{\otimes 2} \\ &\quad n^{1/2}\text{vec}\,((\hat{\mathbf{V}}^{(n)})^{-1} - (\mathbf{V}^{(n)})^{-1}), \end{aligned}$$

and

$$\mathbf{W}^{(n)}_2 := (\hat{\boldsymbol{\beta}}_{0q}' \otimes \hat{\boldsymbol{\beta}}_{0q}')\frac{n^{1/2}}{p}\Big[((\hat{\mathbf{V}}^{(n)})^{-1/2}(\mathbf{V}^{(n)})^{1/2})^{\otimes 2} - \mathbf{I}_{p^2}\Big]\text{vec}(\mathbf{I}_p).$$

Since $\underline{\hat{\lambda}}_{\mathbf{V}}$ is the common value of the $(p-q)$ smallest eigenvalues of $\hat{\mathbf{V}}^{(n)}$, the Slutsky Lemma entails that

$$\begin{aligned} \mathbf{\Sigma}_{p,q}\mathbf{W}^{(n)}_1 &= \frac{2}{p(p+2)}\mathbf{\Sigma}_{p,q}(\hat{\boldsymbol{\beta}}_{0q}' \otimes \hat{\boldsymbol{\beta}}_{0q}')((\mathbf{V}^{(n)})^{1/2})^{\otimes 2}n^{1/2}\text{vec}\,((\hat{\mathbf{V}}^{(n)})^{-1} - (\mathbf{V}^{(n)})^{-1}) \\ &= \frac{2}{p(p+2)}\mathbf{\Sigma}_{p,q}(\hat{\boldsymbol{\beta}}_{0q}' \otimes \hat{\boldsymbol{\beta}}_{0q}')((\hat{\mathbf{V}}^{(n)})^{1/2})^{\otimes 2}n^{1/2}\text{vec}\,((\hat{\mathbf{V}}^{(n)})^{-1} - (\mathbf{V}^{(n)})^{-1}) \\ &\quad + o_{\mathrm{P}}(1) \\ &= \frac{2}{p(p+2)}\underline{\hat{\lambda}}_{\mathbf{V}}\mathbf{\Sigma}_{p,q}(\hat{\boldsymbol{\beta}}_{0q}' \otimes \hat{\boldsymbol{\beta}}_{0q}')n^{1/2}\text{vec}\,((\hat{\mathbf{V}}^{(n)})^{-1} - (\mathbf{V}^{(n)})^{-1}) + o_{\mathrm{P}}(1) \\ &= -\frac{2}{p(p+2)}\underline{\hat{\lambda}}_{\mathbf{V}}\mathbf{\Sigma}_{p,q}n^{1/2}\text{vec}\,(\hat{\boldsymbol{\beta}}_{0q}'(\mathbf{V}^{(n)})^{-1}\hat{\boldsymbol{\beta}}_{0q}) + o_{\mathrm{P}}(1). \end{aligned} \tag{27}$$

Using the notation $\mathbf{\Lambda}_{\mathbf{V}}^{(n)} = \mathrm{diag}(\lambda_{1,\mathbf{V}}^{(n)}, \ldots \lambda_{p,\mathbf{V}}^{(n)}) =: \mathrm{diag}(\mathbf{\Lambda}_{1,\mathbf{V}}^{(n)}, \underline{\lambda}_{\mathbf{V}}^{(n)}\mathbf{I}_{p-q})$, where $\mathbf{\Lambda}_{1,\mathbf{V}}^{(n)}$ is the $q \times q$ diagonal matrix with the q largest roots of $\mathbf{\Lambda}_{\mathbf{V}}^{(n)}$ as diagonal elements, simple computations and (27) yield

$$\begin{aligned}
\mathbf{\Sigma}_{p,q}\mathbf{W}_1^{(n)} &= -\frac{2}{p(p+2)}\hat{\underline{\lambda}}_{\mathbf{V}}\mathbf{\Sigma}_{p,q}n^{1/2}\mathrm{vec}\,(\hat{\boldsymbol{\beta}}_{0q}'(\mathbf{V}^{(n)} - \underline{\lambda}_{\mathbf{V}}^{(n)}\mathbf{I}_p)(-\underline{\lambda}_{\mathbf{V}}^{(n)}\mathbf{V}^{(n)})^{-1}\hat{\boldsymbol{\beta}}_{0q}) \\
&\quad + o_{\mathrm{P}}(1) \\
&= -\frac{2}{p(p+2)}\hat{\underline{\lambda}}_{\mathbf{V}}\mathbf{\Sigma}_{p,q}n^{1/2}\mathrm{vec}\,(\hat{\boldsymbol{\beta}}_{0q}'\boldsymbol{\beta}(\mathbf{\Lambda}_{\mathbf{V}}^{(n)} - \underline{\lambda}_{\mathbf{V}}^{(n)}\mathbf{I}_p)(-\underline{\lambda}_{\mathbf{V}}^{(n)}\mathbf{\Lambda}_{\mathbf{V}}^{(n)})^{-1}\boldsymbol{\beta}'\hat{\boldsymbol{\beta}}_{0q}) \\
&\quad + o_{\mathrm{P}}(1) \\
&= -\frac{2}{p(p+2)}\hat{\underline{\lambda}}_{\mathbf{V}}\mathbf{\Sigma}_{p,q}n^{1/2}\mathrm{vec}\,(\mathbf{E}_{21}^{(n)}(\mathbf{\Lambda}_{1,\mathbf{V}}^{(n)} - \underline{\lambda}_{\mathbf{V}}^{(n)}\mathbf{I}_q) \\
&\quad (-\underline{\lambda}_{\mathbf{V}}^{(n)}\mathbf{\Lambda}_{1,\mathbf{V}}^{(n)})^{-1}(\mathbf{E}_{21}^{(n)})') + o_{\mathrm{P}}(1) \\
&= o_{\mathrm{P}}(1), \qquad (28)
\end{aligned}$$

as $n \to \infty$ under $\mathrm{P}_{\boldsymbol{\beta},\,\mathbf{\Lambda}_{\mathbf{V}}^{(n)}}^{(n)}$, where we used Lemma A.1 in the last line. Now working along the same lines, we obtain that

$$\begin{aligned}
\mathbf{\Sigma}_{p,q}\mathbf{W}_2^{(n)} &= \mathbf{\Sigma}_{p,q}(\hat{\boldsymbol{\beta}}_{0q}' \otimes \hat{\boldsymbol{\beta}}_{0q}')\frac{n^{1/2}}{p}\Big[((\hat{\mathbf{V}}^{(n)})^{-1/2}(\mathbf{V}^{(n)})^{1/2})^{\otimes 2} - \mathbf{I}_{p^2}\Big]\mathrm{vec}(\mathbf{I}_p) \\
&= \mathbf{\Sigma}_{p,q}(\hat{\boldsymbol{\beta}}_{0q}' \otimes \hat{\boldsymbol{\beta}}_{0q}')\frac{n^{1/2}}{p}\Big[((\hat{\mathbf{V}}^{(n)})^{-1/2}(\mathbf{V}^{(n)})^{1/2})^{\otimes 2}\Big]\mathrm{vec}(\mathbf{I}_p) \\
&= \mathbf{\Sigma}_{p,q}\frac{n^{1/2}}{p}\hat{\underline{\lambda}}_{\mathbf{V}}^{-1}\ \mathrm{vec}((\mathbf{E}_{21}^{(n)}, \mathbf{E}_{22}^{(n)})\mathbf{\Lambda}_{\mathbf{V}}^{(n)}(\mathbf{E}_{21}^{(n)}, \mathbf{E}_{22}^{(n)})') \\
&= \mathbf{\Sigma}_{p,q}\frac{n^{1/2}}{p}\hat{\underline{\lambda}}_{\mathbf{V}}^{-1}\ \mathrm{vec}\,(\mathbf{E}_{21}^{(n)}(\mathbf{\Lambda}_{\mathbf{V},1}^{(n)} - \underline{\lambda}_{\mathbf{V}}^{(n)}\mathbf{I}_q)(\mathbf{E}_{21}^{(n)})') + o_{\mathrm{P}}(1) \\
&= o_{\mathrm{P}}(1) \qquad (29)
\end{aligned}$$

as $n \to \infty$ under $\mathrm{P}_{\boldsymbol{\beta},\,\mathbf{\Lambda}_{\mathbf{V}}^{(n)}}^{(n)}$. Combining (26), (28) and (29), we obtain that

$$\mathbf{\Sigma}_{p,q}\mathbf{T}^{(n)}(\hat{\mathbf{V}}^{(n)}) = \mathbf{\Sigma}_{p,q}(\hat{\boldsymbol{\beta}}_{0q}' \otimes \hat{\boldsymbol{\beta}}_{0q}')n^{1/2}\mathrm{vec}(\mathbf{S}_{\mathrm{sign}}^{(n)}) + o_{\mathrm{P}}(1) \qquad (30)$$

as $n \to \infty$ under $\mathrm{P}_{\boldsymbol{\beta},\,\mathbf{\Lambda}_{\mathbf{V}}^{(n)}}^{(n)}$. Using again Lemma A.1, the Slutsky Lemma and

(30), we obtain as $n \to \infty$ under $\mathrm{P}^{(n)}_{\boldsymbol{\beta},\, \boldsymbol{\Lambda}^{(n)}_{\mathbf{V}}}$ that

$$\begin{aligned}
S^{(n)}_{q,\mathrm{sign}} &= \frac{p(p+2)}{2}\Big(n^{1/2}\mathrm{vec}'((\mathbf{E}^{(n)}_{21}, \mathbf{E}^{(n)}_{22})\boldsymbol{\beta}'(\mathbf{S}^{(n)}_{\mathrm{sign}} - \frac{1}{p}\mathbf{I}_p)\boldsymbol{\beta}(\mathbf{E}^{(n)}_{21}, \mathbf{E}^{(n)}_{22})')\boldsymbol{\Sigma}_{p,q} \\
&\quad n^{1/2}\mathrm{vec}((\mathbf{E}^{(n)}_{21}, \mathbf{E}^{(n)}_{22})\boldsymbol{\beta}'(\mathbf{S}^{(n)}_{\mathrm{sign}} - \frac{1}{p}\mathbf{I}_p)\boldsymbol{\beta}(\mathbf{E}^{(n)}_{21}, \mathbf{E}^{(n)}_{22})')\Big) + o_{\mathrm{P}}(1) \\
&= \frac{p(p+2)}{2}\Big(n^{1/2}\mathrm{vec}'(\mathbf{E}^{(n)}_{22}\boldsymbol{\beta}'_{0q}\mathbf{S}^{(n)}_{\mathrm{sign}}\boldsymbol{\beta}_{0q}(\mathbf{E}^{(n)}_{22})')\boldsymbol{\Sigma}_{p,q} \\
&\quad n^{1/2}\mathrm{vec}(\mathbf{E}^{(n)}_{22}\boldsymbol{\beta}'_{0q}\mathbf{S}^{(n)}_{\mathrm{sign}}\boldsymbol{\beta}_{0q}(\mathbf{E}^{(n)}_{22})')\Big) + o_{\mathrm{P}}(1) \\
&= \frac{p(p+2)}{2} n\Big(\mathrm{tr}((\mathbf{E}^{(n)}_{22}\boldsymbol{\beta}'_{0q}\mathbf{S}^{(n)}_{\mathrm{sign}}\boldsymbol{\beta}_{0q}(\mathbf{E}^{(n)}_{22})')^2) \\
&\quad - \frac{1}{p-q}\mathrm{tr}^2(\mathbf{E}^{(n)}_{22}\boldsymbol{\beta}'_{0q}\mathbf{S}^{(n)}_{\mathrm{sign}}\boldsymbol{\beta}_{0q}(\mathbf{E}^{(n)}_{22})')\Big) + o_{\mathrm{P}}(1) \\
&= \frac{p(p+2)}{2} n\Big(\mathrm{tr}((\boldsymbol{\beta}'_{0q}\mathbf{S}^{(n)}_{\mathrm{sign}}\boldsymbol{\beta}_{0q})^2) - \frac{1}{p-q}\mathrm{tr}^2(\boldsymbol{\beta}'_{0q}\mathbf{S}^{(n)}_{\mathrm{sign}}\boldsymbol{\beta}_{0q})\Big) + o_{\mathrm{P}}(1).
\end{aligned} \tag{31}$$

The result follows exactly as in the proof of Proposition 5 in [2], using classical Rao-Mitra arguments. □

Proof of Theorem 4.3. The result follows from a direct application of the Le Cam third Lemma using the LAN property of Theorem 4.1 and (31). □

Now, the main tool to prove Theorem 5.1 will be the following Lemma, allowing to express the test statistic $S^{(n)}_{q,\mathrm{sign}}$ in terms of $\hat{\mathbf{V}}^{(n)}_{\mathrm{Tyler}}$ (see Section 3 for the definition of $\hat{\mathbf{V}}^{(n)}_{\mathrm{Tyler}}$) with probability 1 as $n \to \infty$. Note that Lemma A.2 requires that the last $(p-q)$ eigenvalues of $\mathbf{V}^{(n)}$ are equals, which does not allow us to use it to derive the asymptotic distribution of $S^{(n)}_{q,\mathrm{sign}}$ under local alternatives of the type (6) and justifies that the LAN property of Theorem 4.1 had to be derived.

Lemma A.2. *Let* $\boldsymbol{\Lambda}^{(n)}_{\mathbf{V}} = \mathrm{diag}(\lambda^{(n)}_{1,\mathbf{V}}, \ldots, \lambda^{(n)}_{q,\mathbf{V}}, \underline{\lambda}^{(n)}_{\mathbf{V}}\, \mathbf{1}'_{p-q})$ *a converging sequence of diagonal matrices with ordered eigenvalues. Under* $\mathrm{P}^{(n)}_{\boldsymbol{\beta},\boldsymbol{\Lambda}^{(n)}_{\mathbf{V}}}$ *and as* $n \to \infty$*, we have that*

$$S^{(n)}_{q,\mathrm{sign}} = S^{(n)}_{q,\mathrm{Tyler}} + o_{\mathrm{P}}(1),$$

with

$$S^{(n)}_{q,\mathrm{Tyler}} := \frac{np}{2(p+2)\hat{\underline{\lambda}}^2_{\mathbf{V}}}\Big(\sum_{i=q+1}^{p} \hat{\lambda}^2_{i,\mathrm{Tyler}} - (p-q)^{-1}\Big(\sum_{i=q+1}^{p} \hat{\lambda}_{i,\mathrm{Tyler}}\Big)^2\Big), \tag{32}$$

and with $\hat{\lambda}_{i,\mathrm{Tyler}}$*,* $i = 1, \ldots, p$ *and* $\hat{\underline{\lambda}}_{\mathbf{V}}$ *defined in Section 3.*

Proof of Lemma A.2. First, since $\hat{\mathbf{V}}^{(n)}_{\mathrm{Tyler}}$ satisfies

$$\frac{1}{n}\sum_{i=1}^{n}\frac{(\hat{\mathbf{V}}^{(n)}_{\mathrm{Tyler}})^{-1/2}\mathbf{U}_{ni}\mathbf{U}'_{ni}(\hat{\mathbf{V}}^{(n)}_{\mathrm{Tyler}})^{-1/2}}{\mathbf{U}'_{ni}(\hat{\mathbf{V}}^{(n)}_{\mathrm{Tyler}})^{-1}\mathbf{U}_{ni}} = p^{-1}\mathbf{I}_p,$$

we obtain, using the same arguments as in the proof of Theorem 3.1 in [18] that

$$\begin{aligned} n^{1/2}\,\mathrm{vec}\,(\hat{\mathbf{S}}^{(n)}_{\mathrm{sign}} - p^{-1}\mathbf{I}_p) &= -\frac{1}{p(p+2)}\Big[\mathbf{I}_{p^2}+\mathbf{K}_p+\mathbf{J}_p\Big]((\mathbf{V}^{(n)})^{1/2})^{\otimes 2} \\ &\quad n^{1/2}\mathrm{vec}\,((\hat{\mathbf{V}}^{(n)})^{-1} - (\hat{\mathbf{V}}^{(n)}_{\mathrm{Tyler}})^{-1}) + \frac{n^{1/2}}{p} \\ &\quad \Big[((\hat{\mathbf{V}}^{(n)})^{-1/2}(\mathbf{V}^{(n)})^{1/2})^{\otimes 2} \\ &\quad -((\hat{\mathbf{V}}^{(n)}_{\mathrm{Tyler}})^{-1/2}(\mathbf{V}^{(n)})^{-1/2})^{\otimes 2}\Big]\mathrm{vec}(\mathbf{I}_p) + o_{\mathrm{P}}(1) \\ &= -\frac{1}{p(p+2)}\Big[\mathbf{I}_{p^2}+\mathbf{K}_p+\mathbf{J}_p\Big]((\hat{\mathbf{V}}^{(n)}_{\mathrm{Tyler}})^{1/2})^{\otimes 2} \\ &\quad n^{1/2}\mathrm{vec}\,((\hat{\mathbf{V}}^{(n)})^{-1} - (\hat{\mathbf{V}}^{(n)}_{\mathrm{Tyler}})^{-1}) \\ &\quad +\frac{n^{1/2}}{p}\Big[((\hat{\mathbf{V}}^{(n)})^{-1/2}(\hat{\mathbf{V}}^{(n)}_{\mathrm{Tyler}})^{1/2})^{\otimes 2} - \mathbf{I}_{p^2}\Big]\mathrm{vec}(\mathbf{I}_p) \\ &\quad +o_{\mathrm{P}}(1) \end{aligned} \tag{33}$$

as $n\to\infty$ and under $\mathrm{P}^{(n)}_{\boldsymbol{\beta},\boldsymbol{\Lambda}^{(n)}_{\mathbf{V}}}$. Now using the notation $\boldsymbol{\Sigma}_{p,q} := \mathbf{I}_{(p-q)^2} - \frac{1}{p-q}\mathbf{J}_{p-q}$, (33) implies that

$$\begin{aligned} S^{(n)}_{q,\mathrm{sign}} &= \frac{p(p+2)}{2}\left(n^{1/2}\mathrm{vec}'(\hat{\boldsymbol{\beta}}'_{0q}\hat{\mathbf{S}}^{(n)}_{\mathrm{sign}}\hat{\boldsymbol{\beta}}_{0q})\boldsymbol{\Sigma}_{p,q}n^{1/2}\mathrm{vec}(\hat{\boldsymbol{\beta}}'_{0q}\hat{\mathbf{S}}^{(n)}_{\mathrm{sign}}\hat{\boldsymbol{\beta}}_{0q})\right) + o_{\mathrm{P}}(1) \\ &= \frac{p(p+2)}{2}(-\mathbf{W}^{(n)}_1+\mathbf{W}^{(n)}_2)'\boldsymbol{\Sigma}_{p,q}(-\mathbf{W}^{(n)}_1+\mathbf{W}^{(n)}_2) + o_{\mathrm{P}}(1) \end{aligned} \tag{34}$$

under $\mathrm{P}^{(n)}_{\boldsymbol{\beta},\boldsymbol{\Lambda}^{(n)}_{\mathbf{V}}}$ as $n\to\infty$, where

$$\begin{aligned} \mathbf{W}^{(n)}_1 &:= (\hat{\boldsymbol{\beta}}'_{0q}\otimes\hat{\boldsymbol{\beta}}'_{0q})\frac{1}{p(p+2)} \\ &\quad \Big[\mathbf{I}_{p^2}+\mathbf{K}_p+\mathbf{J}_p\Big]((\hat{\mathbf{V}}^{(n)})^{1/2})^{\otimes 2}n^{1/2}\mathrm{vec}\,((\hat{\mathbf{V}}^{(n)})^{-1} - (\hat{\mathbf{V}}^{(n)}_{\mathrm{Tyler}})^{-1}), \end{aligned}$$

and

$$\mathbf{W}^{(n)}_2 := (\hat{\boldsymbol{\beta}}'_{0q}\otimes\hat{\boldsymbol{\beta}}'_{0q})\frac{n^{1/2}}{p}\Big[((\hat{\mathbf{V}}^{(n)})^{-1/2}(\hat{\mathbf{V}}^{(n)}_{\mathrm{Tyler}})^{1/2})^{\otimes 2} - \mathbf{I}_{p^2}\Big]\mathrm{vec}(\mathbf{I}_p).$$

Now, denoting by $\underline{\hat{\lambda}}_{\mathbf{V}}$ the common value of the last $(p-q)$ eigenvalues of $\hat{\mathbf{V}}^{(n)}$ and using the fact that $\hat{\mathbf{V}}^{(n)}$ and $\hat{\mathbf{V}}^{(n)}_{\rm Tyler}$ share the same eigenvectors together with the fact that $\mathbf{\Sigma}_{p,q}(\hat{\boldsymbol{\beta}}_{0q}^{'} \otimes \hat{\boldsymbol{\beta}}_{0q}^{'})\mathrm{vec}\,(\mathbf{I}_p) = \mathbf{0}$, we obtain that

$$\begin{aligned}\mathbf{\Sigma}_{p,q}\mathbf{W}_1^{(n)} &= 2\mathbf{\Sigma}_{p,q}(\hat{\boldsymbol{\beta}}_{0q}^{'} \otimes \hat{\boldsymbol{\beta}}_{0q}^{'})\frac{1}{p(p+2)}((\hat{\mathbf{V}}^{(n)}_{\rm Tyler})^{1/2})^{\otimes 2} \\ &\quad n^{1/2}\mathrm{vec}\,((\hat{\mathbf{V}}^{(n)})^{-1} - (\hat{\mathbf{V}}^{(n)}_{\rm Tyler})^{-1}) + o_{\rm P}(1) \\ &= \frac{2}{\underline{\hat{\lambda}}_{\mathbf{V}}\, p(p+2)}\mathbf{\Sigma}_{p,q} n^{1/2}\mathrm{vec}\,(\mathrm{diag}(\hat{\lambda}_{q+1,\rm Tyler}, \ldots, \hat{\lambda}_{p,\rm Tyler})) + o_{\rm P}(1),\end{aligned} \tag{35}$$

under $\mathrm{P}^{(n)}_{\boldsymbol{\beta},\mathbf{\Lambda}_{\mathbf{V}}^{(n)}}$ as $n \to \infty$. Now still since $\mathbf{\Sigma}_{p,q}(\hat{\boldsymbol{\beta}}_{0q}^{'} \otimes \hat{\boldsymbol{\beta}}_{0q}^{'})\mathrm{vec}\,(\mathbf{I}_p) = \mathbf{0}$, we also have that

$$\begin{aligned}\mathbf{\Sigma}_{p,q}\mathbf{W}_2^{(n)} &= \mathbf{\Sigma}_{p,q}(\hat{\boldsymbol{\beta}}_{0q}^{'} \otimes \hat{\boldsymbol{\beta}}_{0q}^{'})\frac{n^{1/2}}{p}\Big[((\hat{\mathbf{V}}^{(n)})^{-1/2}(\hat{\mathbf{V}}^{(n)}_{\rm Tyler})^{1/2})^{\otimes 2} - \mathbf{I}_{p^2}\Big]\mathrm{vec}(\mathbf{I}_p) \\ &= \mathbf{\Sigma}_{p,q}(\hat{\boldsymbol{\beta}}_{0q}^{'} \otimes \hat{\boldsymbol{\beta}}_{0q}^{'})\frac{n^{1/2}}{p}\mathrm{vec}((\hat{\mathbf{V}}^{(n)})^{-1/2}(\hat{\mathbf{V}}^{(n)}_{\rm Tyler})(\hat{\mathbf{V}}^{(n)})^{-1/2}) \\ &= \frac{1}{\underline{\hat{\lambda}}_{\mathbf{V}}\, p}\mathbf{\Sigma}_{p,q} n^{1/2}\mathrm{vec}(\mathrm{diag}(\hat{\lambda}_{q+1,\rm Tyler}, \ldots, \hat{\lambda}_{p,\rm Tyler})).\end{aligned} \tag{36}$$

Combining (34), (35) and (36), the result follows easily from the fact that $\mathbf{\Sigma}_{p,q}$ is idempotent.

□

Proof of Theorem 5.1. From Lemma A.2, is it enough to show the weak convergence of $S^{(n)}_{q,\rm Tyler}$ in order to get the result. From [23], we have that

$$n^{1/2}\mathrm{vec}(\boldsymbol{\beta}'(\hat{\mathbf{V}}^{(n)}_{\rm Tyler} - \mathbf{V}^{(n)})\boldsymbol{\beta})$$

converges weakly to $\mathrm{vec}(\mathbf{Z}) \sim \mathcal{N}_{p^2}(0, \mathbf{\Theta}(v_1, \ldots, v_{s_1}))$ as $n \to \infty$ under $\mathrm{P}^{(n)}_{\boldsymbol{\beta},\mathbf{\Lambda}_{\mathbf{V}}^{(n)}}$. Now, it follows along the same lines as in the proof of Proposition 1 in [2] that $S^{(n)}_{q,\rm Tyler}$ converges weakly to

$$\frac{p}{2(p+2)\underline{\lambda}^2_{\mathbf{V}}}\Big(\sum_{j=q+1}^{p} \ell_j^2 - (p-q)^{-1}\big(\sum_{j=q+1}^{p} \ell_j\big)^2\Big). \tag{37}$$

This is the desired result.

□

Proof of Theorem 6.1. Using Theorem A.2, the result follows along the same lines as in the proof of Proposition 4 of [2]. □

Funding

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 811017.

References

[1] Bartlett, M. S. (1950). Tests of significance in factor analysis. *Br. J. Stat. Psychol.* **3** 77–85.
[2] Bernard, G. and Verdebout, T. (2024). Power enhancement for dimension detection of Gaussian signals. *Statist. Sinica* **34** 1–22.
[3] Bernard, G. and Verdebout, T. (2024). On some multivariate sign tests for scatter matrix eigenvalues. *Econom. Stat.* **29** 252–260.
[4] Bernard, G. and Verdebout, T. (2024). On testing the equality of latent roots of scatter matrices under ellipticity. *J. Multivariate Anal.* **199** 105232.
[5] Chakraborty, L., Rus, H., Thistlethwaite, J. and Scott, D. (2020). A place-based socioeconomic status index: Measuring social vulnerability to flood hazards in the context of environmental justice. *Int. J. Disaster Risk Reduct.* **43**.
[6] Fan, J., Liao, Y. and Yao, J. (2015). Power enhancement in high-dimensional cross-sectional tests. *Econometrica* **83** 1497–1541.
[7] Forzani, L., Gieco, A. and Carlos, T. (2017). Likelihood ratio test for partial sphericity in high and ultra-high dimensions. *J. Multivariate Anal.* **159** 18–38.
[8] Hallin, M. and Paindaveine, D. (2006). Semiparametrically efficient rank-based inference for shape. I. Optimal rank-based tests for sphericity. *Ann. Statist.* **34** 2707–2756.
[9] Hallin, M., Paindaveine, D. and Verdebout, T. (2010). Optimal rank-based testing for principal components. *Ann. Statist.* **38** 3245–3299.
[10] Lawley, D. N. (1956). Tests of significance for the latent roots of covariance and correlation matrices. *Biometrika* **43** 128–136.
[11] Le Cam, L. (1986). *Asymptotic methods in statistical decision theory*, 1st ed. Springer, New York.
[12] Le Cam, L. and Yang, G. L. (2000). *Asymptotics in statistics*, 2nd ed. Springer, New York.
[13] Luo, W. and Li, B. (2016). Combining eigenvalues and variation of eigenvectors for order determination. *Biometrika* **103** 875–887.
[14] Nadler, B. (2010). Nonparametric detection of signals by information theoretic criteria: performance analysis and an improved estimator. *IEEE Trans. Signal Process.* **58** 2746–2756.
[15] Nordhausen, K., Oja, H. and Tyler, D. E. (2022). Asymptotic and bootstrap tests for subspace dimension. *J. Multivariate Anal.* **188** 104830.
[16] Paindaveine, D. (2008). A canonical definition of shape. *Statist. Probab. Lett.* **78** 2240–2247.

[17] Paindaveine, D., Rasoafaraniaina, J. and Verdebout, T. (2021). Preliminary Multiple-Test Estimation, With Applications to k-Sample Covariance Estimation. *J. Amer. Statist. Assoc.* **117** 1904-1915.

[18] Paindaveine, D., Remy, J. and Verdebout, T. (2020). Sign tests for weak principal directions. *Bernoulli* **26** 2987–3016.

[19] Paindaveine, D. and Verdebout, T. (2016). On high-dimensional sign tests. *Bernoulli* **22** 1745–1769.

[20] Saleh, A. M. E. (2006). *Theory of preliminary test and Stein-type estimation with applications.* John Wiley & Sons, New Jersey.

[21] Schott, J. R. (2006). A high-dimensional test for the equality of the smallest eigenvalues of a covariance matrix. *J. Multivariate Anal.* **97** 827–843.

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

[23] Tyler, D. E. (1987). A distribution-free M-estimator of multivariate scatter. *Ann. Statist.* **15** 234–251.

[24] Virta, J. (2021). Testing for subsphericity when n and p are of different asymptotic order. *Statist. Probab. Lett.* **179** 109209.

[25] Virta, J. and Nordhausen, K. (2019). Estimating the number of signals using principal component analysis. *Stat* **8** e231.

[26] Waternaux, C. M. (1984). Principal components in the nonnormal case: the test of equality of q roots. *J. Multivariate Anal.* **14** 323–335.

[27] Zou, C., Peng, L., Feng, L. and Wang, Z. (2014). Multivariate-sign-based high-dimensional tests for sphericity. *Biometrika* **101** 229–236.

[28] Şahan, C., Baydur, H. and Demiral, Y. (2019). A novel version of Copenhagen Psychosocial Questionnaire-3: Turkish validation study. *Arch. Environ. Occup. Health* **74** 297-309.