On the asymptotic behavior of the leading eigenvector of Tyler's shape estimator under weak identifiability

Davy Paindaveine and Thomas Verdebout

Abstract We consider point estimation in an elliptical Principal Component Analysis framework. More precisely, we focus on the problem of estimating the leading eigenvector $\boldsymbol{\theta}_1$ of the corresponding shape matrix. We consider this problem under asymptotic scenarios that allow the difference $r_n := \lambda_{n1} - \lambda_{n2}$ between both largest eigenvalues of the underlying shape matrix to converge to zero as the sample size n diverges to infinity. Such scenarios make the problem of estimating $\boldsymbol{\theta}_1$ challenging since this leading eigenvector is then not identifiable in the limit. In this framework, we study the asymptotic behavior of $\hat{\boldsymbol{\theta}}_1$, the leading eigenvector of Tyler's M-estimator of shape. We show that consistency and asymptotic normality survive scenarios where $\sqrt{n}r_n$ diverges to infinity as n does, although the faster the sequence (r_n) converges to zero, the poorer the corresponding consistency rate is. We also prove that consistency is lost if $r_n = O(1/\sqrt{n})$, but that $\hat{\boldsymbol{\theta}}_1$ still bears some information on $\boldsymbol{\theta}_1$ when $\sqrt{n}r_n$ converges to a positive constant. When $\sqrt{n}r_n$ diverges to infinity, we provide asymptotic confidence zones for $\boldsymbol{\theta}_1$ based on $\hat{\boldsymbol{\theta}}_1$. Our non-standard asymptotic results are supported by Monte-Carlo exercises.

1 Introduction

Many classical models in multivariate statistics include scatter parameters. The most common example is the elliptical model where observations are independent copies of a random p-vector $\mathbf{X}$ whose characteristic function is of the form

$$\mathbf{t} \mapsto e^{i\mathbf{t}'\boldsymbol{\mu}}\phi(\mathbf{t}'\boldsymbol{\Sigma}\mathbf{t}) \tag{1}$$

Davy Paindaveine
ECARES and Mathematics Department e-mail: Davy.Paindaveine@ulb.be

Thomas Verdebout
ECARES and Mathematics Department e-mail: Thomas.Verdebout@ulb.be

for some *characteristic generator* $\phi : \mathbb{R}^+ \to \mathbb{R}$. Here, the p-vector $\boldsymbol{\mu}$ is a location parameter and the $p \times p$ symmetric and positive definite matrix $\boldsymbol{\Sigma}$ is a scatter parameter. A very popular instance is of course the p-variate normal model that is obtained with $\phi(s) = \exp(-s^2/2)$. Inference on the scatter parameter in the elliptical model has been the subject of many contributions: to cite only a few, [6, 11, 20] studied the asymptotic properties of robust estimators of $\boldsymbol{\Sigma}$, [1] provided several properties of the Minimum Covariance Determinant estimator of $\boldsymbol{\Sigma}$, sphericity tests have been studied in [8, 13], [18] computed the influence functions of empirical canonical correlation coefficients, [9, 10, 17, 21, 22] considered Principal Component Analysis based on estimators of $\boldsymbol{\Sigma}$, whereas [4, 19] studied estimators of the eigenvalues of $\boldsymbol{\Sigma}$.

In the present paper, we consider estimation of the leading eigenvector of $\boldsymbol{\Sigma}$, that is, of the eigenvector, $\boldsymbol{\theta}_1$ say, associated with the largest eigenvalue of $\boldsymbol{\Sigma}$. This is of course the primary object of interest when conducting a Principal Component Analysis exercise. Since $\boldsymbol{\theta}_1$ does not change when $\boldsymbol{\Sigma}$ is replaced with $c\boldsymbol{\Sigma}$ for any $c > 0$, we actually want to estimate the leading eigenvector of the *shape matrix*

$$\mathbf{V} := \frac{\boldsymbol{\Sigma}}{(\det \boldsymbol{\Sigma})^{1/p}} \tag{2}$$

associated with $\boldsymbol{\Sigma}$, that is, of the version of $\boldsymbol{\Sigma}$ that is normalized to have determinant one (see [14]); note that this also takes care of the fact that, in (1), $\boldsymbol{\Sigma}$ was identified up to a positive scalar factor only. What makes our contribution original is that we will consider double asymptotic scenarios where, as the sample size n diverges to infinity, the underlying shape matrix $\mathbf{V} = \mathbf{V}_n$ has its two leading eigenvalues $\lambda_{n1} > \lambda_{n2}$ that satisfy $\lambda_{n1}/\lambda_{n2} \to 1$. This means that, while $\boldsymbol{\theta}_1$ is properly identifiable for any n (up to an unimportant sign change, as usual), it is no more identifiable in the limit as $n \to \infty$. Obviously, such *weak identifiability* scenarios make inference on $\boldsymbol{\theta}_1$ challenging for large n.

More precisely, we will consider throughout triangular arrays of observations $\mathbf{X}_{n1}, \ldots, \mathbf{X}_{nn}$, where, for each n, the $\mathbf{X}_{in}$'s form a random sample from the p-variate elliptical distribution with location $\boldsymbol{\mu}$, shape matrix

$$\mathbf{V}_n = \frac{\mathbf{I}_p + \delta_n \xi \boldsymbol{\theta}_1 \boldsymbol{\theta}_1'}{(\det(\mathbf{I}_p + \delta_n \xi \boldsymbol{\theta}_1 \boldsymbol{\theta}_1'))^{1/p}} = \frac{(1 + \delta_n \xi)}{(1 + \delta_n \xi)^{1/p}} \boldsymbol{\theta}_1 \boldsymbol{\theta}_1' + \frac{1}{(1 + \delta_n \xi)^{1/p}} (\mathbf{I}_p - \boldsymbol{\theta}_1 \boldsymbol{\theta}_1'), \tag{3}$$

and characteristic generator ϕ_n; in (3), $\boldsymbol{\theta}_1$ is a unit p-vector, ξ is a positive real number, and δ_n is a bounded positive sequence. We will denote the corresponding sequence of hypotheses as $\mathrm{P}_{\boldsymbol{\theta}_1, \delta_n, \xi, \phi_n}$. Throughout the paper, we tacitly assume that ϕ_n is such that $\mathbf{X}_{n1} \neq \mathbf{0}$ almost surely, which is needed to make Tyler's estimator of shape well-defined below. The second expression of $\mathbf{V}_n$ in (3) makes it clear that the leading eigenvalue of $\mathbf{V}_n$ is

$$\lambda_{n1} := (1 + \delta_n \xi)^{(p-1)/p}, \tag{4}$$

with corresponding eigenvector $\boldsymbol{\theta}_1$, and that its remaining eigenvalues are

$$\lambda_{n2} = \ldots = \lambda_{np} := (1 + \delta_n \xi)^{-1/p}, \tag{5}$$

with an eigenspace that is obviously the orthogonal complement to $\boldsymbol{\theta}_1$. If $\delta = 1$ for any n (which we will denote as $\delta \equiv 1$ in the sequel), then the classical setup in which λ_{n1} remains asymptotically well separated from the remaining eigenvalues is obtained. While we will cover this case as well, our main interest below will be on the weakly identifiable case where δ_n is $o(1)$, which provides $\lambda_{n1}/\lambda_{n2} \to 1$, hence makes $\boldsymbol{\theta}_1$ unidentifiable in the limit.

In the sequel, we will restrict to the case $\boldsymbol{\mu} = \mathbf{0}$, which is actually without any loss of generality in the distributional setup considered above. In elliptical models, the Fisher information matrix for location and shape parameters is indeed block-diagonal (see [8]), which entails that asymptotic inference for the shape parameter can be conducted in the same way under specified and unspecified location (block-diagonality of the Fisher information matrix guarantees in particular that parametric efficiency bounds for shape under known and unknown $\boldsymbol{\mu}$ do coincide). In the specified location case, the results of this paper actually trivially extend to the *generalized elliptical distributions* introduced in [5].

Quite naturally, $\boldsymbol{\theta}_1$ can be estimated by the leading eigenvector of a shape estimator $\hat{\mathbf{V}}_n$. For this purpose, we will focus in this paper on the shape estimator $\hat{\mathbf{V}}_n$ that was proposed by David Tyler in [20]. We will investigate the asymptotic behavior of the corresponding leading eigenvector $\hat{\boldsymbol{\theta}}_{n1}$ in the triangular distributional framework described above. In particular, we will show that $\hat{\boldsymbol{\theta}}_{n1}$ is consistent and asymptotically normal when $\sqrt{n}\delta_n \to \infty$, but that it is not consistent when $\delta_n = O(1/\sqrt{n})$. We will precisely derive the limiting distribution of $\hat{\boldsymbol{\theta}}_{n1}$ for any sequence (δ_n). Our results identify the same phase transitions as in the corresponding hypothesis testing framework, when testing $\mathcal{H}_0 : \boldsymbol{\theta}_1 = \boldsymbol{\theta}_1^0$ against $\mathcal{H}_1 : \boldsymbol{\theta}_1 \neq \boldsymbol{\theta}_1^0$ for some fixed unit p-vector $\boldsymbol{\theta}_1^0$; see [15, 16].

The rest of the paper is organized as follows: in Section 2, we recall the definition of Tyler's estimator of shape $\hat{\mathbf{V}}_n$ and provide its asymptotic distribution under weak identifiability. In Section 3, we derive the limiting behavior of $\hat{\boldsymbol{\theta}}_{n1}$ under weak identifiability and discuss the construction of confidence zones for $\boldsymbol{\theta}_1$ under sequences δ_n such that $\sqrt{n}\delta_n \to \infty$. In Section 4, we corroborate the results of Section 3 through Monte-Carlo exercises. A technical appendix collects the proofs.

For the sake of convenience, we collect here the notation that will be used in the paper. Throughout, $\mathbf{e}_\ell$ will denote the ℓth vector of the canonical basis of $\mathbb{R}^p$, so that $\mathbf{K}_p := \sum_{i,j=1}^p (\mathbf{e}_i\mathbf{e}_j') \otimes (\mathbf{e}_j\mathbf{e}_i')$ is the usual *commutation matrix*. Denoting as vec $\mathbf{A}$ the vector obtained by stacking the columns of the matrix $\mathbf{A}$ on top of each other, we let $\mathbf{J}_p := (\text{vec}\,\mathbf{I}_p)(\text{vec}\,\mathbf{I}_p)'$, where $\mathbf{I}_\ell$ is the ℓ-dimensional identity matrix. We will write $\text{diag}(a_1, \ldots, a_\ell)$ for the diagonal matrix collecting the real numbers $a_1, \ldots, a_\ell$ on its diagonal. For a symmetric and positive definite matrix $\mathbf{B}$, we will denote as $\mathbf{B}^{1/2}$ its symmetric and positive definite square root and as $\mathbf{B}^{-1/2}$ the inverse of this square root. Finally, $\to_{\mathcal{D}}$ will stand for convergence in distribution.

2 Tyler's estimator of shape under weak identifiability

As explained above, we consider the problem of estimating the eigenvector $\boldsymbol{\theta}_1$ associated with the largest eigenvalue of the underlying shape matrix $\mathbf{V}_n$. To estimate $\boldsymbol{\theta}_1$, we will use the leading eigenvector $\hat{\boldsymbol{\theta}}_{n1}$ of Tyler's estimator of shape $\hat{\mathbf{V}}_n$ from [20]. Under specified location $\boldsymbol{\mu} = \mathbf{0}$, this shape estimator $\hat{\mathbf{V}}_n$ is defined as the solution of

$$\frac{p}{n} \sum_{i=1}^{n} \frac{\mathbf{X}_{ni}\mathbf{X}'_{ni}}{\mathbf{X}'_{ni}\hat{\mathbf{V}}_n^{-1}\mathbf{X}_{ni}} = \hat{\mathbf{V}}_n, \tag{6}$$

normalized to have unit determinant. This can be seen as the estimator of shape for which the directions (or *spatial signs*) of the resulting sphericized observations

$$\frac{\hat{\mathbf{V}}_n^{-1/2}\mathbf{X}_{n1}}{\|\hat{\mathbf{V}}_n^{-1/2}\mathbf{X}_{n1}\|}, \ldots, \frac{\hat{\mathbf{V}}_n^{-1/2}\mathbf{X}_{nn}}{\|\hat{\mathbf{V}}_n^{-1/2}\mathbf{X}_{nn}\|}$$

have an empirical covariance matrix (with respect to specified location $\boldsymbol{\mu} = \mathbf{0}$) equal to $(1/p)\mathbf{I}_p$. Tyler's estimator of shape enjoys many nice properties. In particular, it is distribution-free in the (centered) elliptical model and it is consistent and asymptotically normal under a broad range of distributions without moment assumptions; see [20]. Distribution-freeness is an important property since it entails that the distribution of $\hat{\boldsymbol{\theta}}_{n1}$ does not depend on the underlying characteristic generator ϕ_n, that is, it does not depend on the type of elliptical distribution at hand (normal, t, etc.) nor on the scale of this elliptical distribution.

The following result provides the asymptotic distribution of Tyler's estimator of shape in a framework where $\boldsymbol{\theta}_1$ is possibly weakly identifiable.

Proposition 1 *Fix a unit vector $\boldsymbol{\theta}_1$, a positive real number ξ and a sequence (δ_n) that either is $\delta_n \equiv 1$ or is $o(1)$. Let $(\mathbf{V}_n)$ be the resulting sequence of shape matrices in (3). Let further (ϕ_n) be a sequence of characteristic generators. Then,*

$$\sqrt{n}\,\mathrm{vec}(\hat{\mathbf{V}}_n - \mathbf{V}_n) \to_{\mathcal{D}} \mathcal{N}\bigg(\mathbf{0}, \Big(1 + \frac{2}{p}\Big)\Big\{(\mathbf{I}_{p^2} + \mathbf{K}_p)(\mathbf{V} \otimes \mathbf{V}) - \frac{2}{p}\mathrm{vec}(\mathbf{V})\mathrm{vec}'(\mathbf{V})\Big\}\bigg)$$

under $\mathrm{P}_{\boldsymbol{\theta}_1,\delta_n,\xi,\phi_n}$ as $n \to \infty$, where $\mathbf{V}$ is the limit of $(\mathbf{V}_n)$.

This result shows that root-n consistency of Tyler's estimator of shape $\hat{\mathbf{V}}_n$ is robust to arbitrarily weakly identifiable scenarios, that is, to scenarios where (δ_n) converges to zero arbitrarily fast. As we will show in the next section, this is not the case for the leading eigenvector of $\hat{\mathbf{V}}_n$.

3 Asymptotic behavior of Tyler's leading eigenvector under weak identifiability

The main goal of this section is to derive the asymptotic behavior of the leading eigenvector $\hat{\boldsymbol{\theta}}_{n1}$ of $\hat{\mathbf{V}}_n$ under weak identifiability. Denoting as $\hat{\lambda}_{nj}$, $j = 1, \ldots, p$, the eigenvalues of $\hat{\mathbf{V}}_n$ in decreasing order (these sample eigenvalues are pairwise different almost surely), we first provide the following result, that shows that root-n consistency of these eigenvalues is robust to weak identifiability.

Proposition 2 *Fix a unit vector $\boldsymbol{\theta}_1$, a positive real number ξ and a sequence (δ_n) that either is $\delta_n \equiv 1$ or is $o(1)$. Let (ϕ_n) be a sequence of characteristic generators. Then, for any $j = 1, \ldots, p$, $\sqrt{n}(\hat{\lambda}_{nj} - \lambda_{nj})$ is $O_{\mathrm{P}}(1)$ as $n \to \infty$ under $\mathrm{P}_{\boldsymbol{\theta}_1, \delta_n, \xi, \phi_n}$.*

With $\boldsymbol{\theta}_1$ fixed, pick arbitrarily p-vectors $\boldsymbol{\theta}_2, \ldots, \boldsymbol{\theta}_p$ such that $\boldsymbol{\Gamma} := (\boldsymbol{\theta}_1, \boldsymbol{\theta}_2, \ldots, \boldsymbol{\theta}_p)$ is orthogonal. Let further $\hat{\boldsymbol{\Gamma}}_n := (\hat{\boldsymbol{\theta}}_{n1}, \ldots, \hat{\boldsymbol{\theta}}_{np})$ stand for the orthogonal matrix whose jth column vector is an eigenvector of $\hat{\mathbf{V}}_n$ associated with eigenvalue $\hat{\lambda}_{nj}$. To unambiguously fix the "signs" of $\hat{\boldsymbol{\theta}}_{nj}$, $j = 1, \ldots, p$, we impose that, with probability one, all entries in the first column of

$$\mathbf{E}_n := \hat{\boldsymbol{\Gamma}}_n' \boldsymbol{\Gamma} = \begin{pmatrix} E_{n,11} & \mathbf{E}_{n,12} \\ \mathbf{E}_{n,21} & \mathbf{E}_{n,22} \end{pmatrix} \tag{7}$$

are positive (note that all entries of $\mathbf{E}_n$ are almost surely non-zero). The following result then provides the asymptotic behavior of $\mathbf{E}_n$ in the present context.

Proposition 3 *Fix a unit vector $\boldsymbol{\theta}_1$, a positive real number ξ and a sequence (δ_n) that either is $\delta_n \equiv 1$ or is $o(1)$. Let (ϕ_n) be a sequence of characteristic generators. Let $\mathbf{Z}$ be a $p \times p$ random matrix such that*

$$\mathrm{vec}(\mathbf{Z}) \sim \mathcal{N}\bigg(\mathbf{0}, \Big(1 + \frac{2}{p}\Big)\Big\{(\mathbf{I}_{p^2} + \mathbf{K}_p) - \frac{2}{p}\mathbf{J}_p\Big\}\bigg),$$

and let $\mathbf{E}(\xi) := (\mathbf{w}_1(\xi), \ldots, \mathbf{w}_p(\xi))'$, where $\mathbf{w}_j(\xi) = (w_{j1}(\xi), \ldots, w_{jp}(\xi))'$ is the unit eigenvector associated with the jth largest eigenvalue of $\mathbf{Z} + \mathrm{diag}(\xi, 0, \ldots, 0)$ and such that $w_{j1}(\xi) > 0$ almost surely. Then, we have the following as $n \to \infty$ under $\mathrm{P}_{\boldsymbol{\theta}_1, \delta_n, \xi, \phi_n}$:

(i) if $\delta_n \equiv 1$, then $n(E_{n,11} - 1) = O_{\mathrm{P}}(1)$, $\mathbf{E}_{n,22}\mathbf{E}_{n,22}' = \mathbf{I}_{p-1} + o_{\mathrm{P}}(1)$, $\sqrt{n}\mathbf{E}_{n,21} = O_{\mathrm{P}}(1)$, and both $\sqrt{n}\mathbf{E}_{n,22}'\mathbf{E}_{n,21}$ and $\sqrt{n}\mathbf{E}_{n,12}'$ are asymptotically normal with mean zero and covariance matrix $\xi^{-2}(1 + \xi)(1 + \frac{2}{p})\mathbf{I}_{p-1}$;

(ii) if δ_n is $o(1)$ with $\sqrt{n}\delta_n \to \infty$, then $n\delta_n^2(E_{n,11} - 1) = O_{\mathrm{P}}(1)$, $\mathbf{E}_{n,22}\mathbf{E}_{n,22}' = \mathbf{I}_{p-1} + o_{\mathrm{P}}(1)$, $\sqrt{n}\delta_n\mathbf{E}_{n,21} = O_{\mathrm{P}}(1)$, and both $\sqrt{n}\delta_n\mathbf{E}_{n,22}'\mathbf{E}_{n,21}$ and $\sqrt{n}\delta_n\mathbf{E}_{n,12}'$ are asymptotically normal with mean zero and covariance matrix $\xi^{-2}(1 + \frac{2}{p})\mathbf{I}_{p-1}$;

(iii) if $\delta_n = 1/\sqrt{n}$, then $\mathbf{E}_n$ converges weakly to $\mathbf{E}(\xi)$;

(iv) if $\delta_n = o(1/\sqrt{n})$, then $\mathbf{E}_n$ converges weakly to $\mathbf{E} := \mathbf{E}(0)$.

The four regimes (i)–(iv) identified in this result will play a crucial role in the asymptotic behavior of $\hat{\boldsymbol{\theta}}_{n1}$ below. At this point, let us note that, in regimes (i)–(ii),

$$\|\sqrt{n}\delta_n(\hat{\boldsymbol{\theta}}_{n1} - \boldsymbol{\theta}_1)\|^2 = 2n\delta_n^2(1 - \hat{\boldsymbol{\theta}}'_{n1}\boldsymbol{\theta}_1) = 2n\delta_n^2(1 - E_{n,11}) = O_{\rm P}(1); \tag{8}$$

this is compatible with the well-known $\sqrt{n}$-consistency of $\hat{\boldsymbol{\theta}}_{n1}$ in the classical case obtained with $\delta_n \equiv 1$, and suggests that $\sqrt{n}$-consistency deteriorates into $(\sqrt{n}\delta_n)$-consistency in regime (ii), which in turn suggests that consistency is lost in regime (iii). The following result, which is the main result of the paper, shows that this is precisely what happens.

Theorem 1 *Fix a unit vector $\boldsymbol{\theta}_1$, a positive real number ξ and a sequence (δ_n) that either is $\delta_n \equiv 1$ or is $o(1)$. Let (ϕ_n) be a sequence of characteristic generators. Then, the leading eigenvector $\hat{\boldsymbol{\theta}}_{n1}$ of Tyler's estimator of shape satisfies the following as $n \to \infty$ under ${\rm P}_{\boldsymbol{\theta}_1,\delta_n,\xi,\phi_n}$:*

(i) if $\delta_n \equiv 1$, then $\sqrt{n}(\hat{\boldsymbol{\theta}}_{n1} - \boldsymbol{\theta}_1)$ is asymptotically normal with mean zero and covariance matrix

$$\frac{1}{\xi^2}(1+\xi)\left(1+\frac{2}{p}\right)(\mathbf{I}_p - \boldsymbol{\theta}_1\boldsymbol{\theta}'_1) = \left(1+\frac{2}{p}\right)\frac{\lambda_1\lambda_2}{(\lambda_1-\lambda_2)^2}(\mathbf{I}_p - \boldsymbol{\theta}_1\boldsymbol{\theta}'_1),$$

where λ_1 and λ_2 are the eigenvalues in (4)–(5) with $\delta_n \equiv 1$;

(ii) if δ_n is $o(1)$ with $\sqrt{n}\delta_n \to \infty$, then $\sqrt{n}\delta_n(\hat{\boldsymbol{\theta}}_{n1} - \boldsymbol{\theta}_1)$ is asymptotically normal with mean zero and covariance matrix

$$\frac{1}{\xi^2}\left(1+\frac{2}{p}\right)(\mathbf{I}_p - \boldsymbol{\theta}_1\boldsymbol{\theta}'_1); \tag{9}$$

(iii) if $\delta_n = 1/\sqrt{n}$, then $\hat{\boldsymbol{\theta}}_{n1}$ converges weakly to the unit eigenvector associated with the largest eigenvalue of $\mathbf{Z}+\xi\boldsymbol{\theta}_1\boldsymbol{\theta}'_1$, where $\mathbf{Z}$ is as in the statement of Proposition 3;

(iv) if $\delta_n = o(1/\sqrt{n})$, then $\hat{\boldsymbol{\theta}}_{n1}$ converges weakly to a random vector that is uniformly distributed over the unit sphere $\mathcal{S}^{p-1}$.

This result confirms that, while the consistency rate of $\hat{\boldsymbol{\theta}}_{n1}$ is of course $\sqrt{n}$ in the standard case $\delta_n \equiv 1$, this consistency rate goes down to $\sqrt{n}\delta_n$ when $\delta_n \to 0$ with $\sqrt{n}\delta_n \to \infty$. Asymptotic normality is obtained in both cases. In the threshold regime (iii) obtained with $\delta_n = 1/\sqrt{n}$, the estimator $\hat{\boldsymbol{\theta}}_{n1}$ is no more consistent for $\boldsymbol{\theta}_1$, yet it still bears some information on $\boldsymbol{\theta}_1$. Clearly, the larger ξ, the larger this information (in particular, the weak limit of $\hat{\boldsymbol{\theta}}_{n1}$ converges to the Dirac distribution at $\boldsymbol{\theta}_1$ as $\xi \to \infty$). Finally, if $\delta_n = o(1/\sqrt{n})$, then $\hat{\boldsymbol{\theta}}_{n1}$ behaves asymptotically as a random vector that is uniformly distributed on the unit sphere of $\mathbb{R}^p$, hence does not bare any information on $\boldsymbol{\theta}_1$. Incidentally, we stress that since $\boldsymbol{\theta}_{n1}$ (resp., $\hat{\boldsymbol{\theta}}_{n1}$) is a homogenous function of $\mathbf{V}_n$ (resp., $\hat{\mathbf{V}}_n$), Theorem 1 still holds true if, in (2), $\mathbf{V}_n$ is rather normalized so that it has trace p, or so that its upper-left entry is equal to one, etc.

The results in Theorem 1 allow one to build confidence zones for $\boldsymbol{\theta}_1$. Let us start with regime (i). Since the sample eigenvalues $\hat{\lambda}_{nj}$, $j = 1, 2$, are $\sqrt{n}$-consistent, confidence zones for $\boldsymbol{\theta}_1$ with asymptotic confidence level $1 - \alpha$ in this regime are given by

$$C_n^{1-\alpha} := \left\{ \boldsymbol{\theta}_1 \in \mathcal{S}^{p-1} : n\left(1 + \frac{2}{p}\right)^{-1} \frac{(\hat{\lambda}_1 - \hat{\lambda}_2)^2}{\hat{\lambda}_1 \hat{\lambda}_2} \hat{\boldsymbol{\theta}}'_{n1} (\mathbf{I}_p - \boldsymbol{\theta}_1 \boldsymbol{\theta}'_1) \hat{\boldsymbol{\theta}}_{n1} \leq \chi^2_{p-1,1-\alpha} \right\},$$

where $\chi^2_{p-1,1-\alpha}$ denotes the upper-α quantile of the chi-square distribution with $p-1$ degrees of freedom. Now, in regime (ii),

$$\begin{aligned}
&\frac{(\hat{\lambda}_{n1} - \hat{\lambda}_{n2})}{\sqrt{\hat{\lambda}_{n1} \hat{\lambda}_{n2}}} \sqrt{n} (\hat{\boldsymbol{\theta}}_{n1} - \boldsymbol{\theta}_1) \\
&\quad = \frac{\sqrt{n}(\hat{\lambda}_{n1} - \lambda_{n1}) - \sqrt{n}(\hat{\lambda}_{n2} - \lambda_{n2}) + \sqrt{n}(\lambda_{n1} - \lambda_{n2})}{\sqrt{\hat{\lambda}_{n1} \hat{\lambda}_{n2}}} (\hat{\boldsymbol{\theta}}_{n1} - \boldsymbol{\theta}_1) \\
&\quad = \frac{\sqrt{n}(\lambda_{n1} - \lambda_{n2})}{\sqrt{\lambda_{n1} \lambda_{n2}}} (\hat{\boldsymbol{\theta}}_{n1} - \boldsymbol{\theta}_1) + o_{\mathrm{P}}(1) = \delta_n \xi (1 + o(1)) \sqrt{n} (\hat{\boldsymbol{\theta}}_{n1} - \boldsymbol{\theta}_1) + o_{\mathrm{P}}(1) \\
&\quad \to_{\mathcal{D}} \mathcal{N}\left(\mathbf{0}, \left(1 + \frac{2}{p}\right)(\mathbf{I}_p - \boldsymbol{\theta}_1 \boldsymbol{\theta}'_1)\right),
\end{aligned}$$

where we used the fact that $\sqrt{n}(\hat{\lambda}_{nj} - \lambda_{nj})$, $j = 1, 2$, are still $O_{\mathrm{P}}(1)$ in this regime (Proposition 2). A direct consequence is that the asymptotic confidence zones $C_n^{1-\alpha}$ above are still valid in regime (ii).

To conclude this section, we turn to robustness issues by considering the influence function of $\hat{\boldsymbol{\theta}}_{n1}$ in regimes (i)–(ii). Using (27) (resp., (28)) in regime (i) (resp., regime (ii)), jointly with (20), (23) and the fact that $E_{n,11} = 1 + o_{\mathrm{P}}(1)$ in regimes (i)–(ii), we obtain

$$\begin{aligned}
\sqrt{n}\delta_n (\hat{\boldsymbol{\theta}}_{n1} - \boldsymbol{\theta}_1) &= (\boldsymbol{\theta}_2, \ldots, \boldsymbol{\theta}_p) \sqrt{n} \delta_n E_{n,11} \mathbf{E}'_{n,12} + o_{\mathrm{P}}(1) \\
&= -(\boldsymbol{\theta}_2, \ldots, \boldsymbol{\theta}_p) \sqrt{n} \delta_n \mathbf{E}'_{n,22} \mathbf{E}_{n,21} + o_{\mathrm{P}}(1) \\
&= (\boldsymbol{\theta}_2, \ldots, \boldsymbol{\theta}_p) \xi^{-1} (1 + \delta_n \xi)^{1/p} (\boldsymbol{\theta}_2, \ldots, \boldsymbol{\theta}_p)' \sqrt{n} (\hat{\mathbf{V}}_n - \mathbf{V}_n) \boldsymbol{\theta}_1 + o_{\mathrm{P}}(1) \\
&= \xi^{-1} (1 + \delta_n \xi)^{1/p} (\mathbf{I}_p - \boldsymbol{\theta}_1 \boldsymbol{\theta}'_1) \sqrt{n} (\hat{\mathbf{V}}_n - \mathbf{V}_n) \boldsymbol{\theta}_1 + o_{\mathrm{P}}(1). \qquad (10)
\end{aligned}$$

From (14) and (16) in the proof of Proposition 1, we have

$$
\begin{aligned}
&\sqrt{n}\,\mathrm{vec}\big(\mathbf{V}_n^{-1/2}\hat{\mathbf{V}}_n\mathbf{V}_n^{-1/2}-\mathbf{I}_p\big)\\
&\qquad\qquad = \left(\mathbf{I}_{p^2}-\frac{1}{p}\mathbf{J}_p\right)\sqrt{n}\,\mathrm{vec}\left(\frac{p\mathbf{V}_n^{-1/2}\hat{\mathbf{V}}_n\mathbf{V}_n^{-1/2}}{\mathrm{tr}[\mathbf{V}_n^{-1}\hat{\mathbf{V}}_n]}-\mathbf{I}_p\right)+o_{\mathrm{P}}(1)\\
&\qquad\qquad = (p+2)\left(\mathbf{I}_{p^2}-\frac{1}{p}\mathbf{J}_p\right)\sqrt{n}\,\mathrm{vec}\big(\mathbf{S}_n(\mathbf{V}_n)-\tfrac{1}{p}\mathbf{I}_p\big)+o_{\mathrm{P}}(1),
\end{aligned}
$$

which yields

$$
\sqrt{n}\,\mathrm{vec}(\hat{\mathbf{V}}_n-\mathbf{V}_n) = (p+2)\big(\mathbf{V}_n^{1/2}\otimes\mathbf{V}_n^{1/2}\big)\left(\mathbf{I}_{p^2}-\frac{1}{p}\mathbf{J}_p\right)\sqrt{n}\,\mathrm{vec}\big(\mathbf{S}_n(\mathbf{V}_n)-\tfrac{1}{p}\mathbf{I}_p\big)+o_{\mathrm{P}}(1).
$$

Since

$$
(\boldsymbol{\theta}_1'\otimes(\mathbf{I}_p-\boldsymbol{\theta}_1\boldsymbol{\theta}_1'))\big(\mathbf{V}_n^{1/2}\otimes\mathbf{V}_n^{1/2}\big) = \frac{(1+\delta_n\xi)^{1/2}}{(1+\delta_n\xi)^{1/p}}(\boldsymbol{\theta}_1'\otimes(\mathbf{I}_p-\boldsymbol{\theta}_1\boldsymbol{\theta}_1'))
$$

(which in particular entails that $(\boldsymbol{\theta}_1'\otimes(\mathbf{I}_p-\boldsymbol{\theta}_1\boldsymbol{\theta}_1'))\big(\mathbf{V}_n^{1/2}\otimes\mathbf{V}_n^{1/2}\big)(\mathrm{vec}\,\mathbf{I}_p) = \mathbf{0}$), plugging this in (10) then provides

$$
\begin{aligned}
\sqrt{n}\delta_n(\hat{\boldsymbol{\theta}}_{n1}-\boldsymbol{\theta}_1) &= \frac{(p+2)(1+\delta_n\xi)^{1/2}}{\xi}(\boldsymbol{\theta}_1'\otimes(\mathbf{I}_p-\boldsymbol{\theta}_1\boldsymbol{\theta}_1'))\sqrt{n}\,\mathrm{vec}(\mathbf{S}_n(\mathbf{V}_n))+o_{\mathrm{P}}(1)\\
&= \frac{(p+2)(1+\delta_n\xi)^{1/2}}{\xi\sqrt{n}}(\mathbf{I}_p-\boldsymbol{\theta}_1\boldsymbol{\theta}_1')\sum_{i=1}^n\frac{\mathbf{V}_n^{-1/2}\mathbf{X}_{ni}\mathbf{X}_{ni}'\mathbf{V}_n^{-1/2}}{\|\mathbf{V}_n^{-1/2}\mathbf{X}_{ni}\|^2}\boldsymbol{\theta}_1+o_{\mathrm{P}}(1).
\end{aligned}
$$

By applying the multivariate central limit theorem (and (15)), it is readily checked that this Bahadur representation result for $\sqrt{n}\delta_n(\hat{\boldsymbol{\theta}}_{n1}-\boldsymbol{\theta}_1)$ is compatible with the asymptotic normality statements in Theorem 1(i)–(ii). More importantly, this Bahadur representation result shows that the boundedness of the influence function of $\hat{\boldsymbol{\theta}}_{n1}$ does not only hold in the standard regime (i) but also in the weakly identifiable regime (ii).

4 Numerical illustration

In this section, we conduct Monte-Carlo simulation exercises to validate the various asymptotic results in Theorem 1. For any $\ell\in\{0,1,\ldots,7\}$, we generated $M = 10,000$ independent random samples of size $n = 100,000$ from the bivariate ($p = 2$) normal distribution with mean vector zero and covariance matrix

$$
\boldsymbol{\Sigma}_{n,\ell} = \mathbf{I}_2+\delta_{n,\ell}\xi\boldsymbol{\theta}_1\boldsymbol{\theta}_1', \tag{11}
$$

with $\delta_{n,\ell} = n^{-\ell/8}$, $\xi = 2$ and $\boldsymbol{\theta}_1 = \mathbf{e}_1 \in \mathbb{R}^2$. In each of these samples, we computed the leading eigenvector $\hat{\boldsymbol{\theta}}_{n1}$ of Tyler's estimator of scatter (still with respect to fixed location at the origin of $\mathbb{R}^p$); evaluation of Tyler's estimator of scatter was done by using the function `tyler.shape` from the R package *ICSNP* ([12]). We first focus on Theorem 1(i)–(ii), hence on the cases $\ell \in \{0, 1, 2, 3\}$. For each such ℓ, we provide in Figure 1 a histogram of the M corresponding values of

$$\sqrt{n}\delta_{n,\ell}\mathbf{e}_2'\hat{\boldsymbol{\theta}}_{n1} = \sqrt{n}\delta_{n,\ell}\mathbf{e}_2'(\hat{\boldsymbol{\theta}}_{n1} - \boldsymbol{\theta}_1). \tag{12}$$

Clearly, the results nicely agree with the corresponding asymptotic distribution of (12) in Theorem 1, namely $\mathcal{N}(0, \frac{3}{2})$ for $\ell = 0$ (regime (i)) and $\mathcal{N}(0, \frac{1}{2})$ for $\ell = 1, 2, 3$ (regime (ii)). We then turn to Theorem 1(iii)–(iv), hence to the cases $\ell \in \{4, 5, 6, 7\}$. For these values of ℓ, Figure 2 reports histograms of

$$\mathbf{e}_2'\hat{\boldsymbol{\theta}}_{n1}. \tag{13}$$

Here, the asymptotic distributions of (13) in Theorem 1(iii)–(iv) do not have a closed form density, and we are therefore plotting kernel density estimates obtained from a random sample of size 10^6 from the weak limit of (13) in Theorem 1(iii)–(iv). To avoid boundary effects (the support of this weak limit is of course $[-1, 1]$), we employed the function `kde.boundary` from the R package *ks* ([3]) with default parameters, which returns the kernel density estimate using the second form of the Beta boundary kernel in [2]. Irrespective of $\ell \in \{4, 5, 6, 7\}$, these empirical results fully support the corresponding asymptotic results in Theorem 1.

Acknowledgements This research is supported by the Program of Concerted Research Actions (ARC) of the Université libre de Bruxelles.

Appendix

The proof of Proposition 1 requires the following preliminary result, which follows from (3.7)–(3.8) in [20].

Lemma 1 *Fix a unit vector $\boldsymbol{\theta}_1$, a positive real number ξ and a sequence (δ_n) that either is $\delta_n \equiv 1$ or is $o(1)$. Let $(\mathbf{V}_n)$ be the resulting sequence of shape matrices in (3). Let further (ϕ_n) be a sequence of characteristic generators. Then, letting*

$$\mathbf{G}_p := \mathbf{I}_{p^2} - \tfrac{1}{p+2}(\mathbf{I}_{p^2} + \mathbf{K}_p - \mathbf{J}_p) \ \ \textit{and} \ \ \mathbf{S}_n(\mathbf{V}) := \frac{1}{n}\sum_{i=1}^{n} \frac{\mathbf{V}^{-1/2}\mathbf{X}_{ni}\mathbf{X}_{ni}'\mathbf{V}^{-1/2}}{\|\mathbf{V}^{-1/2}\mathbf{X}_{ni}\|^2},$$

we have that

$$\mathbf{G}_p\sqrt{n}\,\mathrm{vec}\left(\frac{p\mathbf{V}_n^{-1/2}\hat{\mathbf{V}}_n\mathbf{V}_n^{-1/2}}{\mathrm{tr}[\mathbf{V}_n^{-1}\hat{\mathbf{V}}_n]} - \mathbf{I}_p\right) = p\sqrt{n}\,\mathrm{vec}\big(\mathbf{S}_n(\mathbf{V}_n) - \tfrac{1}{p}\mathbf{I}_p\big) + o_{\mathrm{P}}(1),$$

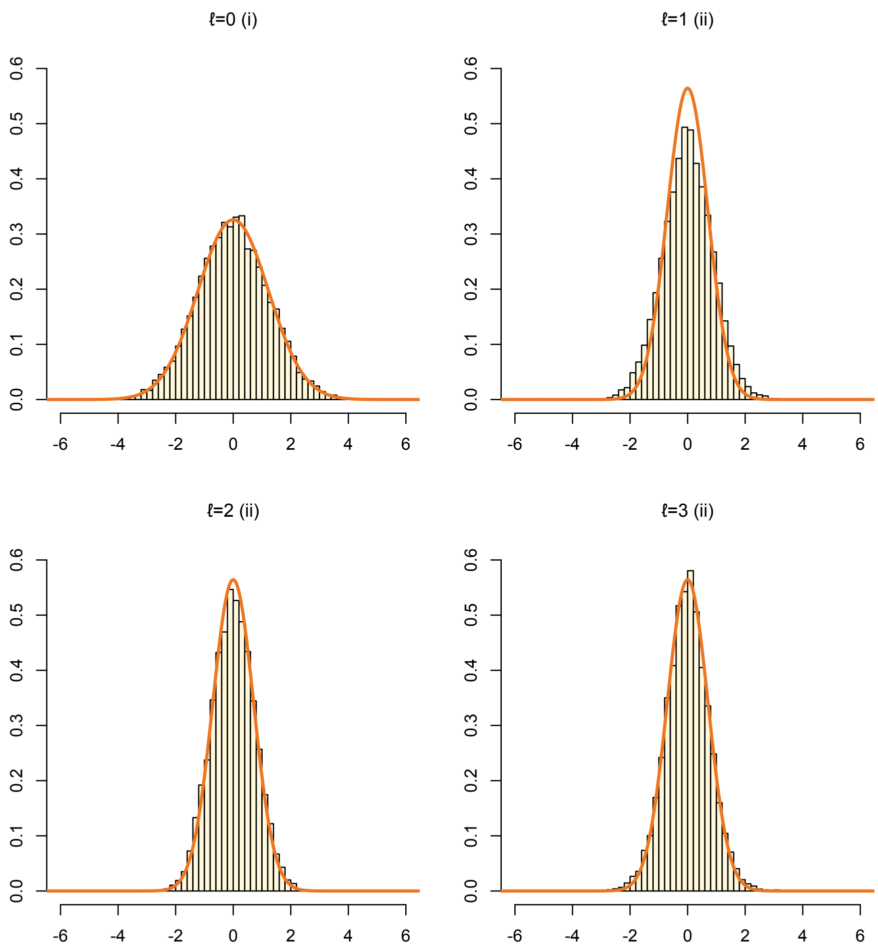

Fig. 1 For each $\ell \in \{0, 1, 2, 3\}$, histograms of the quantities $\sqrt{n}\delta_{n,\ell}\mathbf{e}_2'(\hat{\boldsymbol{\theta}}_{n1} - \boldsymbol{\theta}_1)$ computed from $M = 10,000$ independent random samples of size $n = 100,000$ from the bivariate normal distribution with mean vector zero and the covariance matrix $\boldsymbol{\Sigma}_{n,\ell}$ in (11), where $\hat{\boldsymbol{\theta}}_{n1}$ denotes the leading eigenvector of Tyler's estimator of scatter (with respect to fixed location at the origin of $\mathbb{R}^2$). In each panel, the solid curve is the density of the corresponding asymptotic distribution, namely $\mathcal{N}(0, \frac{3}{2})$ for $\ell = 0$ and $\mathcal{N}(0, \frac{1}{2})$ for $\ell = 1, 2, 3$.

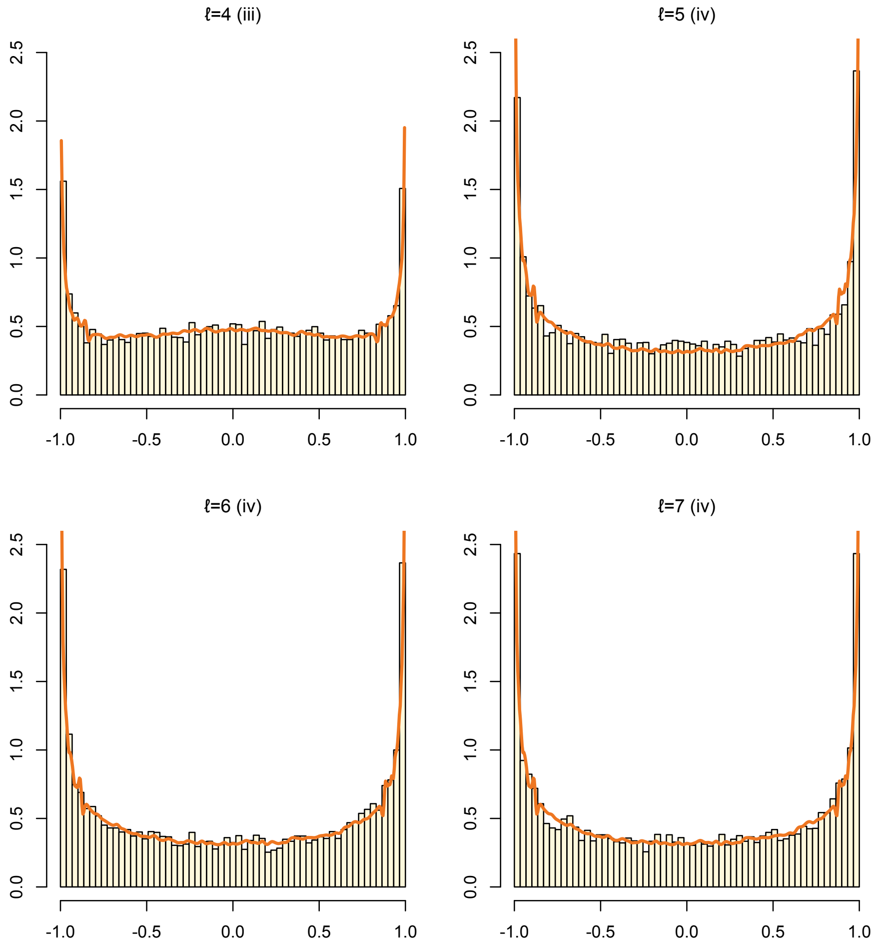

Fig. 2 For each $\ell \in \{4, 5, 6, 7\}$, histograms of the quantities $\mathbf{e}_2'\hat{\boldsymbol{\theta}}_{n1}$ computed from $M = 10,000$ independent random samples of size $n = 100,000$ from the bivariate normal distribution with mean vector zero and the covariance matrix $\boldsymbol{\Sigma}_{n,\ell}$ in (11), where $\hat{\boldsymbol{\theta}}_{n1}$ still denotes the leading eigenvector of Tyler's estimator of scatter (with respect to fixed location at the origin of $\mathbb{R}^2$). In each panel, the solid curve is a kernel estimate for the density of the corresponding weak limit obtained from Theorem 1(iii)–(iv); see Section 4 for details.

under $\mathrm{P}_{\boldsymbol{\theta}_1,\delta_n,\xi,\phi_n}$ *as* $n \to \infty$.

In all proofs below, stochastic convergences are as $n \to \infty$ under $\mathrm{P}_{\boldsymbol{\theta}_1,\delta_n,\xi,\phi_n}$.

Proof of Proposition 1. Letting $\mathbf{H}_p := \mathbf{I}_{p^2} + \mathbf{K}_p - \frac{2}{p}\mathbf{J}_p$, we have $\mathbf{H}_p\mathbf{J}_p = \mathbf{0}$ and $\mathbf{H}_p\mathbf{K}_p(\mathrm{vec}\,\mathbf{B}) = \mathbf{H}_p(\mathrm{vec}\,\mathbf{B})$ for any symmetric matrix $\mathbf{B}$, so that

$$\mathbf{H}_p\mathbf{G}_p(\mathrm{vec}\,\mathbf{B}) = (\mathbf{H}_p - \tfrac{2}{p+2}\mathbf{H}_p)(\mathrm{vec}\,\mathbf{B}) = \tfrac{p}{p+2}\mathbf{H}_p(\mathrm{vec}\,\mathbf{B})$$

for any symmetric matrix $\mathbf{B}$. Lemma 1 thus yields that

$$\tfrac{p}{p+2}\mathbf{H}_p\sqrt{n}\,\mathrm{vec}\bigg(\frac{p\mathbf{V}_n^{-1/2}\hat{\mathbf{V}}_n\mathbf{V}_n^{-1/2}}{\mathrm{tr}[\mathbf{V}_n^{-1}\hat{\mathbf{V}}_n]} - \mathbf{I}_p\bigg) = p\sqrt{n}\,\mathbf{H}_p\mathrm{vec}\big(\mathbf{S}_n(\mathbf{V}_n) - \tfrac{1}{p}\mathbf{I}_p\big) + o_{\mathrm{P}}(1).$$

Using the fact that $\mathbf{J}_p(\mathrm{vec}\,\mathbf{B}) = (\mathrm{tr}[\mathbf{B}])(\mathrm{vec}\,\mathbf{I}_k)$ and $\mathbf{K}_p(\mathrm{vec}\,\mathbf{B}) = \mathrm{vec}\,\mathbf{B}$ for any symmetric matrix $\mathbf{B}$, this rewrites

$$\sqrt{n}\,\mathrm{vec}\bigg(\frac{p\mathbf{V}_n^{-1/2}\hat{\mathbf{V}}_n\mathbf{V}_n^{-1/2}}{\mathrm{tr}[\mathbf{V}_n^{-1}\hat{\mathbf{V}}_n]} - \mathbf{I}_p\bigg) = (p+2)\sqrt{n}\,\mathrm{vec}\big(\mathbf{S}_n(\mathbf{V}_n) - \tfrac{1}{p}\mathbf{I}_p\big) + o_{\mathrm{P}}(1). \quad (14)$$

Now, Lemma A.3(ii) from [15] states that

$$\sqrt{n}\,\mathrm{vec}\big(\mathbf{S}_n(\mathbf{V}_n) - \tfrac{1}{p}\mathbf{I}_p\big) \to_{\mathcal{D}} \mathcal{N}\bigg(\mathbf{0}, \frac{1}{p(p+2)}\mathbf{H}_p\bigg), \quad (15)$$

so that

$$\sqrt{n}\,\mathrm{vec}\bigg(\frac{p\mathbf{V}_n^{-1/2}\hat{\mathbf{V}}_n\mathbf{V}_n^{-1/2}}{\mathrm{tr}[\mathbf{V}_n^{-1}\hat{\mathbf{V}}_n]} - \mathbf{I}_p\bigg) \to_{\mathcal{D}} \mathcal{N}\bigg(\mathbf{0}, \Big(1+\frac{2}{p}\Big)\mathbf{H}_p\bigg).$$

Hence, the argument in the bottom of page 341 in [7] yields that

$$\begin{aligned}
&\sqrt{n}\,\mathrm{vec}\bigg(\frac{\mathbf{V}_n^{-1/2}\hat{\mathbf{V}}_n\mathbf{V}_n^{-1/2}}{(\det(\mathbf{V}_n^{-1/2}\hat{\mathbf{V}}_n\mathbf{V}_n^{-1/2}))^{1/p}} - \mathbf{I}_p\bigg)\\
&\qquad = \bigg(\mathbf{I}_{p^2} - \frac{1}{p}\mathbf{J}_p\bigg)\sqrt{n}\,\mathrm{vec}\bigg(\frac{p\mathbf{V}_n^{-1/2}\hat{\mathbf{V}}_n\mathbf{V}_n^{-1/2}}{\mathrm{tr}[\mathbf{V}_n^{-1}\hat{\mathbf{V}}_n]} - \mathbf{I}_p\bigg) + o_{\mathrm{P}}(1) \qquad (16)\\
&\qquad \to_{\mathcal{D}} \mathcal{N}\bigg(\mathbf{0}, \frac{1}{p(p+2)}\mathbf{H}_p\bigg),
\end{aligned}$$

that is,

$$\sqrt{n}\,\mathrm{vec}\big(\mathbf{V}_n^{-1/2}\hat{\mathbf{V}}_n\mathbf{V}_n^{-1/2} - \mathbf{I}_p\big) \to_{\mathcal{D}} \mathcal{N}\bigg(\mathbf{0}, \frac{1}{p(p+2)}\mathbf{H}_p\bigg).$$

Since this rewrites

$$\big(\mathbf{V}_n^{-1/2}\otimes\mathbf{V}_n^{-1/2}\big)\sqrt{n}\,\mathrm{vec}\big(\hat{\mathbf{V}}_n - \mathbf{V}_n\big) \to_{\mathcal{D}} \mathcal{N}\bigg(\mathbf{0}, \frac{1}{p(p+2)}\mathbf{H}_p\bigg),$$

we finally obtain that

$$\sqrt{n}\,\mathrm{vec}\big(\hat{\mathbf{V}}_n - \mathbf{V}_n\big) \to_{\mathcal{D}} \mathcal{N}\bigg(\mathbf{0}, \Big(1+\frac{2}{p}\Big)\Big\{(\mathbf{I}_{p^2}+\mathbf{K}_p)(\mathbf{V}\otimes\mathbf{V}) - \frac{2}{p}\mathrm{vec}(\mathbf{V})\mathrm{vec}'(\mathbf{V})\Big\}\bigg),$$

with $\mathbf{V}$ the limiting value of $(\mathbf{V}_n)$. □

We do not prove Proposition 2 here since the proof follows along the exact same lines as the proof of Lemma 2.2 in [16]. We thus turn to the proof of Proposition 3, that requires the following linear algebra result.

Lemma 2 *Let* $\mathbf{A}$ *be a* $p \times p$ *matrix. Assume that* λ *is an eigenvalue of* $\mathbf{A}$ *and that the corresponding eigenspace* V_λ *has dimension one. Denoting as* $C = (C_{ij})$ *the cofactor matrix of* $\mathbf{A} - \lambda \mathbf{I}_p$*, assume that* $\mathbf{v} := (C_{11}, \ldots, C_{1p})' \neq \mathbf{0}$*. Then* $V_\lambda = \{t\mathbf{v} : t \in \mathbb{R}\}$*.*

Proof of Lemma 2. For any $j = 1, \ldots, p$, denote as $(\mathbf{A}-\lambda\mathbf{I}_p)_j$ the jth row of $\mathbf{A}-\lambda\mathbf{I}_p$. For $j = 2, \ldots, p$,

$$(\mathbf{A} - \lambda\mathbf{I}_p)_j v = \det\begin{pmatrix} (\mathbf{A} - \lambda\mathbf{I}_p)_j \\ (\mathbf{A} - \lambda\mathbf{I}_p)_2 \\ \vdots \\ (\mathbf{A} - \lambda\mathbf{I}_p)_p \end{pmatrix} = 0,$$

since this is the determinant of a matrix with (at least) twice the same row. Since λ is an eigenvalue of $\mathbf{A}$, this determinant is also zero for $j = 1$. Therefore, $(\mathbf{A}-\lambda\mathbf{I}_p)\mathbf{v} = 0$. The non-zero vector $\mathbf{v}$ thus belongs to V_λ. Since V_λ has dimension one by assumption, the result follows. □

Proof of Proposition 3. In this proof, we put

$$\mathbf{Z}_n := \sqrt{n}\boldsymbol{\Gamma}'(\hat{\mathbf{V}}_n - \mathbf{V}_n)\boldsymbol{\Gamma}. \tag{17}$$

and $\boldsymbol{\Lambda}_n := \boldsymbol{\Gamma}'\mathbf{V}_n\boldsymbol{\Gamma} = \mathrm{diag}(\lambda_{n1}, \ldots, \lambda_{np})$. First note that since

$$\mathbf{E}_n = \hat{\boldsymbol{\Gamma}}_n'\boldsymbol{\Gamma} = \begin{pmatrix} E_{n,11} & \mathbf{E}_{n,12} \\ \mathbf{E}_{n,21} & \mathbf{E}_{n,22} \end{pmatrix}$$

is an orthogonal matrix, we easily obtain that

$$\mathbf{E}_{n,21} = -\frac{1}{E_{n,11}}\mathbf{E}_{n,22}\mathbf{E}_{n,12}', \tag{18}$$

$$\mathbf{E}_{n,22}\mathbf{E}_{n,22}' = \mathbf{I}_{p-1} - \mathbf{E}_{n,21}\mathbf{E}_{n,21}' \tag{19}$$

and

$$E_{n,11}\mathbf{E}_{n,12}' = -\mathbf{E}_{n,22}'\mathbf{E}_{n,21}. \tag{20}$$

We start with the proof of (i)–(ii). The random matrix $\mathbf{Y}_n := \sqrt{n}\boldsymbol{\Gamma}'\hat{\mathbf{V}}_n\boldsymbol{\Gamma} - \sqrt{n}\lambda_{n1}\mathbf{I}_p$ admits the eigenvectors $\mathbf{w}_{nj} := \boldsymbol{\Gamma}'\hat{\boldsymbol{\theta}}_{nj}$, $j = 1, \ldots, p$, with corresponding eigenvalues $\zeta_{nj} := \sqrt{n}(\hat{\lambda}_{nj} - \lambda_{n1})$, $j = 1, \ldots, p$. Thus, with probability one, we have $\zeta_{n1} > \zeta_{n2} > \ldots > \zeta_{np}$, and the eigenspace of

$$\mathbf{Y}_n = \mathbf{Z}_n + \sqrt{n}(\mathbf{\Lambda}_n - \lambda_{n1}\mathbf{I}_p) = \mathbf{Z}_n - \mathrm{diag}\left(0, \frac{\sqrt{n}\delta_n\xi}{(1+\delta_n\xi)^{1/p}}, \ldots, \frac{\sqrt{n}\delta_n\xi}{(1+\delta_n\xi)^{1/p}}\right) \quad (21)$$

associated with eigenvalue ζ_{n1} is spanned by

$$\mathbf{w}_{n1} = \mathbf{\Gamma}'\hat{\boldsymbol{\theta}}_{n1} = \begin{pmatrix} E_{n,11} \\ \mathbf{E}'_{n,12} \end{pmatrix}.$$

Partitioning $\mathbf{Z}_n$ into

$$\mathbf{Z}_n = \begin{pmatrix} Z_{n,11} & \mathbf{Z}'_{n,21} \\ \mathbf{Z}_{n,21} & \mathbf{Z}_{n,22} \end{pmatrix},$$

where $Z_{n,11}$ is a scalar and $\mathbf{Z}_{n,22}$ is a $(p-1)\times(p-1)$ matrix, Lemma 2 then yields that $\mathbf{w}_{n1}$ is proportional to the vector of cofactors associated with the first row of

$$\mathbf{M}_{n,1} := \begin{pmatrix} Z_{n,11} - \zeta_{n1} & \mathbf{Z}'_{n,21} \\ \mathbf{Z}_{n,21} & \mathbf{Z}_{n,22} - \frac{\sqrt{n}\delta_n\xi}{(1+\delta_n\xi)^{1/p}}\mathbf{I}_{p-1} - \zeta_{n1}\mathbf{I}_{p-1} \end{pmatrix}, \quad (22)$$

or equivalently, that $\mathbf{w}_{n1}$ is proportional to the vector of cofactors associated with the first row of

$$\begin{pmatrix} Z_{n,11} - \zeta_{n1} & \mathbf{Z}'_{n,21} \\ \frac{(1+\delta_n\xi)^{1/p}}{\sqrt{n}\delta_n\xi}\mathbf{Z}_{n,21} & \frac{(1+\delta_n\xi)^{1/p}}{\sqrt{n}\delta_n\xi}\mathbf{Z}_{n,22} - \mathbf{I}_{p-1} - \frac{(1+\delta_n\xi)^{1/p}}{\sqrt{n}\delta_n\xi}\zeta_{n1}\mathbf{I}_{p-1} \end{pmatrix}.$$

Since $\mathbf{Z}_{n,21}$ and $\mathbf{Z}_{n,22}$ are $O_{\mathrm{P}}(1)$ (Proposition 1) and so is ζ_{n1} (Proposition 2), we obtain that

$$\begin{pmatrix} E_{n,11} \\ \mathbf{E}'_{n,12} \end{pmatrix} = \mathbf{e}_1 + o_{\mathrm{P}}(1)$$

(recall that $E_{n,11} > 0$ almost surely and that $\mathbf{e}_1$ is the first vector of the canonical basis of $\mathbb{R}^p$) and that

$$\sqrt{n}\delta_n\mathbf{E}'_{n,12} = O_{\mathrm{P}}(1).$$

Using the fact that $\mathbf{E}_n$ is orthogonal, it follows that

$$n\delta_n^2(1 - E_{n,11}) = \frac{\|\sqrt{n}\delta_n\mathbf{E}'_{n,12}\|^2}{1 + E_{n,11}} = \frac{1}{2}\|\sqrt{n}\delta_n\mathbf{E}'_{n,12}\|^2 + o_{\mathrm{P}}(1) = O_{\mathrm{P}}(1).$$

Since $\mathbf{E}_{n,22}$ is bounded, it also directly follows from (18) that $\sqrt{n}\delta_n\mathbf{E}_{n,21} = O_{\mathrm{P}}(1)$. In view of (19), we then obtain that $\mathbf{E}_{n,22}\mathbf{E}'_{n,22} - \mathbf{I}_{p-1}$ is $o_{\mathrm{P}}(1)$. Now, letting $\hat{\mathbf{\Lambda}}_n := \hat{\mathbf{\Gamma}}'_n\hat{\mathbf{V}}_n\hat{\mathbf{\Gamma}}_n = \mathrm{diag}(\hat{\lambda}_{n1}, \ldots, \hat{\lambda}_{np})$, we have

$$\begin{aligned}
\mathbf{Z}_{n,21} &= \sqrt{n}(\boldsymbol{\Gamma}'\hat{\mathbf{V}}_n\boldsymbol{\Gamma})_{21} = \sqrt{n}(\boldsymbol{\Gamma}'\hat{\boldsymbol{\Gamma}}_n\hat{\boldsymbol{\Lambda}}_n\hat{\boldsymbol{\Gamma}}_n'\boldsymbol{\Gamma})_{21} \\
&= \sqrt{n}(\mathbf{E}_n'\hat{\boldsymbol{\Lambda}}_n\mathbf{E}_n)_{21} = \sqrt{n}(\mathbf{E}_{n,12}'\,\mathbf{E}_{n,22}')\hat{\boldsymbol{\Lambda}}_n\begin{pmatrix} E_{n,11} \\ \mathbf{E}_{n,21}\end{pmatrix} \\
&= \sqrt{n}\hat{\lambda}_{n1}E_{n,11}\mathbf{E}_{n,12}' + \sqrt{n}\mathbf{E}_{n,22}'\mathrm{diag}(\hat{\lambda}_{n2},\ldots,\hat{\lambda}_{np})\mathbf{E}_{n,21}.
\end{aligned}$$

Writing $\ell_{nj} := \sqrt{n}(\hat{\lambda}_{nj}-\lambda_{nj})$ for $j = 1,\ldots,p$, using (4)–(5), then (20), thus provides

$$\begin{aligned}
\mathbf{Z}_{n,21} &= \ell_{n1}\mathbf{E}_{n,11}'\mathbf{E}_{n,12}' + \mathbf{E}_{n,22}'\mathrm{diag}(\ell_{n2},\ldots,\ell_{np})\mathbf{E}_{n,21} \\
&\qquad +\sqrt{n}(1+\delta_n\xi)^{(p-1)/p}E_{n,11}\mathbf{E}_{n,12}' + \sqrt{n}(1+\delta_n\xi)^{-1/p}\mathbf{E}_{n,22}'\mathbf{E}_{n,21} \\
&= \mathbf{E}_{n,22}'\mathrm{diag}(\ell_{n2}-\ell_{n1},\ldots,\ell_{np}-\ell_{n1})\mathbf{E}_{n,21} - \sqrt{n}\delta_n\xi(1+\delta_n\xi)^{-1/p}\mathbf{E}_{n,22}'\mathbf{E}_{n,21},
\end{aligned}$$

which, since the ℓ_{nj}'s are $O_{\mathrm{P}}(1)$ (Proposition 2), yields

$$\sqrt{n}\delta_n\mathbf{E}_{n,22}'\mathbf{E}_{n,21} = -\frac{(1+\delta_n\xi)^{1/p}}{\xi}\mathbf{Z}_{n,21} + o_{\mathrm{P}}(1). \tag{23}$$

Now, Proposition 1 directly entails that $\mathrm{vec}\,\mathbf{Z}_n = (\boldsymbol{\Gamma}'\otimes\boldsymbol{\Gamma}')\sqrt{n}\mathrm{vec}\,(\hat{\mathbf{V}}_n - \mathbf{V}_n)$ is asymptotically

$$\mathcal{N}\left(\mathbf{0}, \left(1+\frac{2}{p}\right)\left\{(\mathbf{I}_{p^2}+\mathbf{K}_p)(\boldsymbol{\Lambda}\otimes\boldsymbol{\Lambda}) - \frac{2}{p}(\mathrm{vec}\,\boldsymbol{\Lambda})(\mathrm{vec}\,\boldsymbol{\Lambda})'\right\}\right)$$

in case (i), where $\boldsymbol{\Lambda} := \mathrm{diag}((1+\xi)^{(p-1)/p}, (1+\xi)^{-1/p}\ldots,(1+\xi)^{-1/p})$ and

$$\mathcal{N}\left(\mathbf{0}, \left(1+\frac{2}{p}\right)\left\{(\mathbf{I}_{p^2}+\mathbf{K}_p) - \frac{2}{p}\mathbf{J}_p\right\}\right)$$

in case (ii). Therefore, straightforward computations yield

$$\mathbf{Z}_{n,21} = (\mathbf{e}_2,\ldots,\mathbf{e}_p)'\mathbf{Z}_n\mathbf{e}_1 = (\mathbf{e}_1'\otimes(\mathbf{e}_2,\ldots,\mathbf{e}_p)')\mathrm{vec}\,\mathbf{Z}_n \to_{\mathcal{D}} \mathcal{N}(\mathbf{0},\mathbf{B}),$$

where

$$\mathbf{B} := \left(1+\frac{2}{p}\right)(1+\xi)^{(p-2)/p}\mathbf{I}_{p-1} \quad \text{and} \quad \mathbf{B} := \left(1+\frac{2}{p}\right)\mathbf{I}_{p-1}$$

in case (i) and in case (ii), respectively. In view of (23), the desired asymptotic normality result for $\sqrt{n}\delta_n\mathbf{E}_{n,22}'\mathbf{E}_{n,21}$ follows. The one for $\sqrt{n}\delta_n\mathbf{E}_{n,12}'$ then follows from (20) and the fact that $E_{n,11} = 1 + o_{\mathrm{P}}(1)$.

We turn to the proof of (iii)–(iv). As above, $\mathbf{w}_{n1} = \boldsymbol{\Gamma}'\hat{\boldsymbol{\theta}}_{n1} = \mathbf{E}_n'\mathbf{e}_1$ is the unit eigenvector associated with the eigenvalue $\zeta_{n1} = \ell_{n1} = \sqrt{n}(\hat{\lambda}_{n1}-\lambda_{n1})$ of $\mathbf{Y}_n$ in (21), or equivalently, with the eigenvalue

$$\tilde{\ell}_{n1} = \ell_{n1} + \frac{\sqrt{n}\delta_n\xi}{(1+\delta_n\xi)^{1/p}} = \sqrt{n}(\hat{\lambda}_{n1}-\lambda_{n2})$$

of

$$\mathbf{Y}_n + \frac{\sqrt{n}\delta_n \xi}{(1+\delta_n\xi)^{1/p}} \mathbf{I}_p = \mathbf{Z}_n + \mathrm{diag}\bigg(\frac{\sqrt{n}\delta_n \xi}{(1+\delta_n\xi)^{1/p}}, 0, \ldots, 0\bigg). \tag{24}$$

Similarly, $\mathbf{w}_{nj} := \boldsymbol{\Gamma}'\hat{\boldsymbol{\theta}}_{nj} = \mathbf{E}_n'\mathbf{e}_j$, $j = 2, \ldots, p$, are the unit eigenvectors associated with the $p-1$ smallest eigenvalues $\ell_{n2} = \sqrt{n}(\hat{\lambda}_{n2} - \lambda_{n2}), \ldots, \ell_{np} = \sqrt{n}(\hat{\lambda}_{np} - \lambda_{np})$ of (24). Consequently, the joint distribution of $\mathbf{w}_{nj}$, $j = 1, \ldots, p$ — that is, the joint distribution of the columns of $\mathbf{E}_n'$ — converges weakly to the joint distribution of the unit eigenvectors (associated with eigenvalues in decreasing order, and with the signs fixed as in the statement of the theorem) of

$$\mathbf{Z} + \lim_{n\to\infty} \mathrm{diag}\bigg(\frac{\sqrt{n}\delta_n \xi}{(1+\delta_n\xi)^{1/p}}, 0, \ldots, 0\bigg)$$

(recall that, in cases (iii)–(iv), $\mathbf{Z}_n$ converges weakly to the random matrix $\mathbf{Z}$). This establishes the result. □

Proof of Theorem 1. (i) In this regime, the eigenvalues λ_{nj}, $j = 1, \ldots, p$, are fixed and given by

$$\lambda_1 := (1+\xi)^{(p-1)/p} \quad \text{and} \quad \lambda_j := (1+\xi)^{-1/p}, \ j = 2, \ldots, p,$$

respectively; see (4)–(5). Since

$$\frac{1}{\xi^2}(1+\xi) = \frac{\lambda_1\lambda_2}{(\lambda_1 - \lambda_2)^2},$$

Proposition 3(i) entails that

$$\sqrt{n}\mathbf{E}_{n,12}' \to_{\mathcal{D}} \mathcal{N}\bigg(\mathbf{0}, \Big(1 + \frac{2}{p}\Big) \frac{\lambda_1\lambda_2}{(\lambda_1 - \lambda_2)^2} \mathbf{I}_{p-1}\bigg). \tag{25}$$

Now, writing $\boldsymbol{\tau}_n := \sqrt{n}(\hat{\boldsymbol{\theta}}_{n1} - \boldsymbol{\theta}_1)$, we have

$$\frac{\|\boldsymbol{\tau}_n\|^2}{2\sqrt{n}} = \frac{\boldsymbol{\tau}_n'\boldsymbol{\tau}_n}{2\sqrt{n}} = \frac{\sqrt{n}}{2}(\hat{\boldsymbol{\theta}}_{n1} - \boldsymbol{\theta}_1)'(\hat{\boldsymbol{\theta}}_{n1} - \boldsymbol{\theta}_1) = \sqrt{n}(1 - \boldsymbol{\theta}_1'\hat{\boldsymbol{\theta}}_{n1}) = -\boldsymbol{\theta}_1'\boldsymbol{\tau}_n, \tag{26}$$

where we used the fact that $\hat{\boldsymbol{\theta}}_{n1}$ and $\boldsymbol{\theta}_1$ are unit vectors. Since $\boldsymbol{\tau}_n := \sqrt{n}(\hat{\boldsymbol{\theta}}_{n1} - \boldsymbol{\theta}_1)$ is $O_{\mathrm{P}}(1)$, it follows that

$$(\mathbf{I}_p - \boldsymbol{\theta}_1\boldsymbol{\theta}_1')\boldsymbol{\tau}_n = \boldsymbol{\tau}_n - (\boldsymbol{\theta}_1'\boldsymbol{\tau}_n)\boldsymbol{\theta}_1 = \boldsymbol{\tau}_n + \frac{\|\boldsymbol{\tau}_n\|^2}{2\sqrt{n}} = \boldsymbol{\tau}_n + o_{\mathrm{P}}(1)$$

as $n \to \infty$. Therefore,

$$\begin{aligned}\sqrt{n}(\hat{\boldsymbol{\theta}}_{n1}-\boldsymbol{\theta}_1) &= (\mathbf{I}_p-\boldsymbol{\theta}_1\boldsymbol{\theta}_1')\sqrt{n}(\hat{\boldsymbol{\theta}}_{n1}-\boldsymbol{\theta}_1)+o_{\rm P}(1)\\ &= \left(\sum_{j=2}^p \boldsymbol{\theta}_j\boldsymbol{\theta}_j'\right)\sqrt{n}(\hat{\boldsymbol{\theta}}_{n1}-\boldsymbol{\theta}_1)+o_{\rm P}(1)\\ &= \sqrt{n}\sum_{j=2}^p \boldsymbol{\theta}_j(\hat{\boldsymbol{\theta}}_{n1}'\boldsymbol{\theta}_j)+o_{\rm P}(1)\\ &= (\boldsymbol{\theta}_2,\ldots,\boldsymbol{\theta}_p)\sqrt{n}\mathbf{E}_{n,12}'+o_{\rm P}(1), \qquad (27)\end{aligned}$$

so that the asymptotic normality result in (25) entails that $\sqrt{n}(\hat{\boldsymbol{\theta}}_{n1}-\boldsymbol{\theta}_1)$ is asymptotically normal with mean zero and covariance matrix

$$\left(1+\frac{2}{p}\right)\frac{\lambda_1\lambda_2}{(\lambda_1-\lambda_2)^2}\sum_{j=2}^p \boldsymbol{\theta}_j\boldsymbol{\theta}_j' = \left(1+\frac{2}{p}\right)\frac{\lambda_1\lambda_2}{(\lambda_1-\lambda_2)^2}(\mathbf{I}_p-\boldsymbol{\theta}_1\boldsymbol{\theta}_1'),$$

as was to be shown. (ii) In this regime, $\boldsymbol{\tau}_n=\sqrt{n}(\hat{\boldsymbol{\theta}}_{n1}-\boldsymbol{\theta}_1)$ is $O_{\rm P}(1/\delta_n)$ (see (8)), so that (26) yields

$$(\mathbf{I}_p-\boldsymbol{\theta}_1\boldsymbol{\theta}_1')\delta_n\boldsymbol{\tau}_n=\delta_n\boldsymbol{\tau}_n+\frac{\delta_n\|\boldsymbol{\tau}_n\|^2}{2\sqrt{n}}=\delta_n\boldsymbol{\tau}_n+o_{\rm P}(1).$$

Therefore,

$$\begin{aligned}\sqrt{n}\delta_n(\hat{\boldsymbol{\theta}}_{n1}-\boldsymbol{\theta}_1) &= (\mathbf{I}_p-\boldsymbol{\theta}_1\boldsymbol{\theta}_1')\sqrt{n}\delta_n(\hat{\boldsymbol{\theta}}_{n1}-\boldsymbol{\theta}_1)+o_{\rm P}(1)\\ &= \left(\sum_{j=2}^p \boldsymbol{\theta}_j\boldsymbol{\theta}_j'\right)\sqrt{n}\delta_n(\hat{\boldsymbol{\theta}}_{n1}-\boldsymbol{\theta}_1)+o_{\rm P}(1)\\ &= (\boldsymbol{\theta}_2,\ldots,\boldsymbol{\theta}_p)\sqrt{n}\delta_n\mathbf{E}_{n,12}'+o_{\rm P}(1), \qquad (28)\end{aligned}$$

so that the result follows from the fact that

$$\sqrt{n}\delta_n\mathbf{E}_{n,12}' \to_{\mathcal{D}} \mathcal{N}\left(\mathbf{0},\frac{1}{\xi^2}\left(1+\frac{2}{p}\right)\mathbf{I}_{p-1}\right).$$

in this regime; see Proposition 3(ii).

(iii) Let $\mathbf{Z}$ be as in the statement of Proposition 3 and write again $\boldsymbol{\Gamma}=(\boldsymbol{\theta}_1,\ldots,\boldsymbol{\theta}_p)$. In the regimes (iii)–(iv),

$$\boldsymbol{\Gamma}'\hat{\boldsymbol{\theta}}_{n1}=\begin{pmatrix}E_{n,11}\\ \mathbf{E}_{n,12}'\end{pmatrix}$$

converges weakly to the unit eigenvector associated with the largest eigenvalue of $\mathbf{Z}+\mathrm{diag}(\xi,0,\ldots,0)$ with $\xi>0$ in regime (iii) and $\xi=0$ in regime (iv). This directly entails that

$$\hat{\boldsymbol{\theta}}_{n1} = \boldsymbol{\Gamma}\begin{pmatrix} E_{n,11} \\ \mathbf{E}'_{n,12} \end{pmatrix}$$

converges weakly to the unit eigenvector associated with the largest eigenvalue of

$$\boldsymbol{\Gamma}(\mathbf{Z} + \mathrm{diag}(\xi, 0, \ldots, 0))\boldsymbol{\Gamma}' = \boldsymbol{\Gamma}\mathbf{Z}\boldsymbol{\Gamma}' + \xi\boldsymbol{\theta}_1\boldsymbol{\theta}'_1. \qquad (29)$$

Part (iii) of the result then follows from the fact that the distribution of $\mathbf{Z}$ is invariant with respect to orthogonal transformations, in the sense that $\mathbf{OZO}'$ has the same distribution as $\mathbf{Z}$ for any $p \times p$ orthogonal matrix $\mathbf{O}$. (iv) The proof for $\xi > 0$ in (iii) above applies for ξ and shows that, in regime (iv), $\hat{\boldsymbol{\theta}}_{n1}$ converges weakly to the unit eigenvector associated with the largest eigenvalue of $\mathbf{Z} = \mathbf{Z}(0)$. Now, the orthogonal invariance of the distribution of $\mathbf{Z} = \mathbf{Z}(0)$ entails that the joint distribution of its eigenvectors is the invariant Haar distribution on the group of $p \times p$ orthogonal matrices, which implies in particular that each of these eigenvectors is uniformly distributed over $\mathcal{S}^{p-1}$. This establishes Part (iv) of the result. □

References

1. Cator, E.A., Lopuhaä, H.P.: Asymptotic expansion of the minimum covariance determinant estimators. J. Multivariate Anal. **101**, 2372–2388 (2010)
2. Chen, S.: Beta kernel estimators for density functions. Comput. Statist. Data Anal. **31**, 131–145 (1999)
3. Duong, T.: ks: Kernel smoothing, https://CRAN.R-project.org/package=ks, R package version 1.13.2 (2021)
4. Dürre, A., Tyler, D.E., Vogel, D.: On the eigenvalues of the spatial sign covariance matrix in more than two dimensions. Statist. Probab. Lett. **111**, 80–85 (2016)
5. Frahm, G.: Generalized elliptical distributions: Theory and applications. PhD thesis, Universität zu Köln (2004)
6. Frahm, G.: Asymptotic distributions of robust shape matrices and scales. J. Multivariate Anal. **100**, 1329–1337 (2009)
7. Hallin, M., Paindaveine, D.: Parametric and semiparametric inference for shape: the role of the scale functional. Statist. Decisions **24**, 327–350 (2006)
8. Hallin, M., Paindaveine, D.: Semiparametrically efficient rank-based inference for shape. i. optimal rank-based tests for sphericity. Ann. Statist. **34**, 2707–2756 (2006)
9. Hallin, M., Paindaveine, D., Verdebout, T.: Optimal rank-based testing for principal components. Ann. Statist. **38**, 3245–3299 (2010)
10. Hallin, M., Paindaveine, D., Verdebout, T.: Efficient R-estimation of principal and common principal components. J. Amer. Statist. Assoc. **109**, 1071–1083 (2014)
11. Lopuhaa, H.P.: Asymptotics of reweighted estimators of multivariate location and scatter. Ann. Statist. pp. 1638–1665 (1999)
12. Nordhausen, K., Sirkia, S., Oja, H., Tyler, D.E.: ICSNP: tools for multivariate nonparametrics, https://CRAN.R-project.org/package=ICSNP, R package version 1.1-1 (2018)
13. Onatski, A., Moreira, M., Hallin, M.: Signal detection in high dimension: the multispiked case. Ann. Statist. **42**, 225–254 (2014)
14. Paindaveine, D.: A canonical definition of shape. Statist. Probab. Lett. **78**, 2240–2247 (2008)
15. Paindaveine, D., Remy, J., Verdebout, T.: Sign tests for weak principal directions. Bernoulli **26**, 2987–3016 (2020)

16. Paindaveine, D., Remy, J., Verdebout, T.: Testing for principal component directions under weak identifiability. Ann. Statist. **48**, 324–345 (2020)
17. Salibián-Barrera, M., Van Aelst, S., Willems, G.: Principal components analysis based on multivariate MM estimators with fast and robust bootstrap. J. Amer. Statist. Assoc. **101**, 1198–1211 (2006)
18. Taskinen, S., Croux, C., Kankainen, A., Ollila, E., Oja, H.: Influence functions and efficiencies of the canonical correlation and vector estimates based on scatter and shape matrices. J. Multivariate Anal. **97**, 359–384 (2006)
19. Tyler, D.: The asymptotic distribution of principal component roots under local alternatives to multiple roots. Ann. Statist. **11**, 1232–1242 (1983)
20. Tyler, D.: A distribution-free M-estimator of multivariate scatter. Ann. Statist. **15**, 234–251 (1987)
21. Tyler, D.E.: Asymptotic inference for eigenvectors. Ann. Statist. **9**, 725–736 (1981)
22. Tyler, D.E.: A class of asymptotic tests for principal component vectors. Ann. Statist. **11**, 1243–1250 (1983)