T E C H N I C A L

R E P O R T

0553

ESTIMATING THE SHAPE OF AN ELLIPTIC DISTRIBUTION PARAMETRIC VERSUS NONPARAMETRIC EFFICIENCIES

HALLIN M., and D. PANDAVEINE

I A P S T A T I S T I C S

N E T W O R K

INTERUNIVERSITY ATTRACTION POLE

http://www.stat.ucl.ac.be/IAP

Estimating the Shape of an Elliptic Distribution Parametric versus Nonparametric Efficiencies

Marc Hallin and Davy Paindaveine*
Université Libre de Bruxelles
Brussels, Belgium

Abstract

We consider the problem of estimating the shape of an elliptical distribution from a parametric (completely specified radial density) and a semiparametric point of view. For every radial density, we provide a closed form expression of the corresponding parametric and semiparametric efficiency bounds. For each case (specified and unspecified density), we exhibit an estimator of shape that achieves the corresponding efficiency bound in the multinormal case. We show that the efficiency loss due to the non-specification of the radial density is entirely caused by the non-specification of its scale. This loss however remains bounded, and reaches a maximum at arbitrarily heavy tails; we also show that under arbitrarily light tails, parametric and semiparametric efficiency bounds coincide, so that, under such densities, ignoring the scale (ignoring the exact density) asymptotically does not harm the estimation of shape.

AMS 1980 subject classification : 62M15, 62G35.

Key words and phrases : Elliptical densities, Shape matrices, Local asymptotic normality, Semiparametric efficiency.

1 Introduction.

Denote by $\mathrm{P}^n_{\boldsymbol{\theta},\mathbf{V},f}$ the distribution of the n-tuple of k-variate observations $\mathbf{X}^{(n)} = (\mathbf{X}_1, \ldots, \mathbf{X}_n)$, where the $\mathbf{X}_i$'s are i.i.d. with common elliptical density

$$\mathbf{x} \mapsto c_{k,f} \left(\det \mathbf{V}\right)^{-1/2} f\Big(\sqrt{(\mathbf{x}-\boldsymbol{\theta})' \mathbf{V}^{-1}(\mathbf{x}-\boldsymbol{\theta})}\Big). \tag{1.1}$$

The *center of symmetry* $\boldsymbol{\theta}$ is a k-dimensional real vector, the *shape matrix* $\mathbf{V}$ belongs to the collection $\mathcal{V}$ of all symmetric positive definite real $k \times k$ matrices with entry 1 in the upper-left corner, and the *radial density* $f : \mathbb{R}^+_0 \longrightarrow \mathbb{R}^+_0$ satisfies $\mu_{k-1,f} := \int_0^\infty r^{k-1} f(r)\, dr < \infty$; $c_{k,f}$ is a normalization factor. The class of all such radial densities will be denoted as $\mathcal{F}$.

The shape matrix $\mathbf{V}$, which coincides with the usual *scatter* matrix up to a positive scale factor, is a parameter of interest in a number of very standard problems in multivariate analysis: principal component analysis (PCA), canonical correlation analysis (CCA), and the problem of testing for sphericity, among others, only depend on shape—rather than on the scatter matrix. Inference on shape is thus an essential tool in the domain. However we mainly concentrate on

*Research supported by by a P.A.I. contract of the Belgian Federal Government and an Action de Recherche Concertée of the Communauté française de Belgique.

the two generic problems of (i) estimating $\mathbf{V}$, and of (ii) testing that $\mathbf{V}$ coincides with some given $\mathbf{V}_0 \in \mathcal{V}$.

These problems can be considered in parametric models of the form $\mathcal{P}^n_f := \{\mathrm{P}^n_{\boldsymbol{\theta},\mathbf{V},f}, \boldsymbol{\theta} \in \mathbb{R}^k, \mathbf{V} \in \mathcal{V}\}$, where the radial density f is completely specified. Such full specification of f however is too much of an assumption, and the classical approach consists in specifying f up to a scale parameter only. This leads to rewriting (1.1) as

$$\mathbf{x} \mapsto c_{k,f_1} \frac{1}{\sigma^k |\mathbf{V}|^{1/2}} f_1 \left(\frac{1}{\sigma} \left((\mathbf{x} - \boldsymbol{\theta})' \mathbf{V}^{-1} (\mathbf{x} - \boldsymbol{\theta}) \right)^{1/2} \right), \quad \mathbf{x} \in \mathbb{R}^k, \tag{1.2}$$

and $\mathrm{P}^n_{\boldsymbol{\theta},\mathbf{V},f}$ as $\mathrm{P}^n_{\boldsymbol{\theta},\sigma,\mathbf{V},f_1}$, where $\sigma > 0$ is a *scale parameter* and $f_1 : \mathbb{R}^+_0 \longrightarrow \mathbb{R}^+_0$ a *standardized radial density*. To ensure identifiability of σ and $c_{k,f_1} \times f_1$ without imposing any moment conditions, we require that

$$\mathrm{Med}\Big[\| \mathbf{V}^{-1/2} (\mathbf{X}^{(n)}_i - \boldsymbol{\theta}) \| \Big] = \sigma, \tag{1.3}$$

where $\mathrm{Med}[Y]$ denotes the median of Y under $\mathrm{P}^n_{\boldsymbol{\theta},\sigma,\mathbf{V},f_1}$. The resulting model, still of a parametric nature, is of the form $\mathcal{P}^n_{f_1} := \{\mathrm{P}^n_{\boldsymbol{\theta},\sigma,\mathbf{V},f_1}, \boldsymbol{\theta} \in \mathbb{R}^k, \sigma^2 > 0, \mathbf{V} \in \mathcal{V}\}$.

Special cases of (1.2) are the k-variate multinormal distribution, with radial density $f_1(r) = \phi_1(r) := \exp(-a_k r^2/2)$, the k-variate Student distributions, with radial densities (for ν degrees of freedom) $f_1(r) = f^t_{1,\nu}(r) := (1 + a_{k,\nu} r^2/\nu)^{-(k+\nu)/2}$, and the k-variate power-exponential distributions, with radial densities of the form $f_1(r) = f^e_{1,\eta}(r) := \exp(-b_{k,\eta} r^{2\eta})$, $\eta > 0$; the constants a_k, $a_{k,\nu} > 0$, and $b_{k,\eta}$ are such that (1.3) is satisfied.

Specifying the density up to a scale parameter still may be unrealistic, and one may prefer considering the same problems of inference on $\mathbf{V}$ in the semiparametric model $\mathcal{P}^n := \{\mathrm{P}^n_{\boldsymbol{\theta},\mathbf{V},f}, \boldsymbol{\theta} \in \mathbb{R}^k, \mathbf{V} \in \mathcal{V}, f \in \mathcal{F}\} = \bigcup_{f \in \mathcal{F}} \mathcal{P}^n_f$.

Optimal inference on $\mathbf{V}$ at some given $\mathrm{P}^n_{\boldsymbol{\theta},\mathbf{V},f}$, as well as the corresponding optimal performance (efficiency), depends on the model adopted. Clearly, the optimal performance achievable in $\mathcal{P}^n_f$, where f is completely specified, is highest (*parametric efficiency*), followed by the performance in $\mathcal{P}^n_{f_1}$, where f is only partially specified, and the performance in $\mathcal{P}^n$ (*semiparametric efficiency*) where f is completely unspecified. It can be shown however that, under mild regularity assumptions (ensuring LAN) on f and the family $\mathcal{F}$,

(i) the location parameter $\boldsymbol{\theta}$ has no influence on any of these efficiencies, which are the same whether $\boldsymbol{\theta}$ is known or not; in practice, any root-n consistent estimator $\hat{\boldsymbol{\theta}}^{(n)}$ thus can be substituted for $\boldsymbol{\theta}$, whereas in theoretical developments, we safely can assume that $\boldsymbol{\theta} = \mathbf{0}$;

(ii) parametric and semiparametric efficiencies at $\mathrm{P}^n_{\boldsymbol{\theta},\mathbf{V},f}$ do not depend on the actual values of $\boldsymbol{\theta}$ and σ: it makes sense, thus, to speak about f_1-*(semi)parametric efficiencies*;

(iii) at standardized radial density f_1, radial optimal performances in $\mathcal{P}^n_{f_1}$ and $\mathcal{P}^n$ coincide; the difference between parametric and semiparametric efficiencies at given $\mathrm{P}^n_{\boldsymbol{\theta},\mathbf{V},f}$ thus is entirely due to the non-specification of scale.

Natural questions in this context are: how do these f_1-parametric and f_1-semiparametric efficiencies compare to each other, that is, how large is, under given f_1, the cost of not knowing the scale (not knowing f) when performing inference on shape? Are there density types f_1 for which this cost would be minimal? Maximal? Zero? Arbitrarily large?

In the hypothesis testing context, efficiencies are measured in terms of local powers, which depend on quadratics in the difference between the parametric and semiparametric information matrices for $\mathbf{V}$. These information matrices, and the answers to the above questions, are explicitly provided in Hallin and Paindaveine (2005a). The corresponding problem for point estimation is much harder, as comparisons here involve asymptotic covariance matrices which are the inverses of information matrices; those information matrices are rather complex, and obtaining their inverse under closed form is far from trivial.

The paper is organized as follows. In Section 2, we introduce some further notation and state the ULAN result which provides the various efficiency bounds. In Section 3, we determine the asymptotic covariance matrix of the f_1-parametric estimators of shape. Section 4 states the corresponding result for the f_1-semiparametric estimators. We state the main results of the paper in Section 5, in which we compare both types of performance at f_1 and discuss our findings in terms of the dimension k and the tail weight of the underlying elliptical distribution. Technical results are proved in the appendix.

2 A ULAN result.

Writing $\mathrm{vech}\,\mathbf{M} := (M_{11}, (\mathrm{v\overset{\circ}{e}ch}\,\mathbf{M})')'$ for the $k(k+1)/2$-dimensional vector obtained by stacking the upper-triangular elements of a $k \times k$ symmetric matrix $\mathbf{M} = (M_{ij})$, we also write $\mathrm{P}^n_{\boldsymbol{\vartheta},f_1}$ for the distribution $\mathrm{P}^n_{\boldsymbol{\theta},\sigma,\mathbf{V},f_1}$ of $\mathbf{X}^{(n)}$ under given values of $\boldsymbol{\vartheta} := (\boldsymbol{\theta}', \sigma^2, (\mathrm{v\overset{\circ}{e}ch}\,\mathbf{V})')'$ and f_1. Below, we state the uniform local asymptotic normality (ULAN) result, with respect to $\boldsymbol{\vartheta} = (\boldsymbol{\theta}', \sigma^2, (\mathrm{v\overset{\circ}{e}ch}\,\mathbf{V})')'$, of the families of distributions $\mathcal{P}^n_{f_1} := \{\mathrm{P}^n_{\boldsymbol{\vartheta},f_1}, \boldsymbol{\vartheta} \in \boldsymbol{\Theta}\}$, where $\boldsymbol{\Theta} = \mathbb{R}^k \times \mathbb{R}^+_0 \times \mathrm{v\overset{\circ}{e}ch}\,\mathcal{V}$. Of course, ULAN requires some regularity condition on the radial density f_1. A minimal assumption is given in Hallin and Paindaveine (2005a); for the sake of simplicity, we rather provide the following sufficient one:

Assumption (A). The radial density f_1 is absolutely continuous, with a.e.-derivative $\dot{f}_1$, and, letting $\varphi_{f_1} := -\dot{f}_1/f_1$, the expectations (under $\mathrm{P}^n_{\mathbf{0},1,\mathbf{I}_k,f_1}$, where $\mathbf{I}_k$ denotes the k-dimensional identity matrix)

$$\mathcal{I}_k(f_1) := \mathrm{E}\Big[\varphi^2_{f_1}(\|\mathbf{X}_i\|)\Big] \quad \text{and} \quad \mathcal{J}_k(f_1) := \mathrm{E}\Big[\|\mathbf{X}_i\|^2\varphi^2_{f_1}(\|\mathbf{X}_i\|)\Big]$$

are finite.

Rather than introducing a new specific notation, we henceforth tacitly assume that $\mathcal{F}$ is the class of all radial densities satisfying Assumption (A).

This assumption is extremely mild, and does not imply any moment conditions; $\mathcal{I}_k(f_1)$ and $\mathcal{J}_k(f_1)$ can be interpreted as radial Fisher information for location and radial Fisher information for shape/scale, respectively. It can be checked that—provided that $k \geq 2$ (which is not a limitation, since the problem under consideration is void for $k = 1$)—Assumption (A) is satisfied at Gaussian densities, at all Student densities (including the Cauchy ones), as well as at all power-exponential densities. Using the notation of the previous section, the corresponding radial Fisher information values are given, for the Gaussian, the Student with ν degrees of freedom, and the power-exponential with parameter η, by

$$\mathcal{I}_k(\phi_1) = a_k\, k, \quad \mathcal{I}_k(f^t_{1,\nu}) = a_{k,\nu}\,\frac{k(k+\nu)}{k+\nu+2}, \quad \mathcal{I}_k(f^e_{1,\eta}) = 4\eta^2 b_{k,\eta}\,\frac{\Gamma\left((4\eta+k-2)/(2\eta)\right)}{\Gamma\left(k/(2\eta)\right)},$$

and

$$\mathcal{J}_k(\phi_1) = k(k+2), \quad \mathcal{J}_k(f^t_{1,\nu}) = \frac{k(k+2)(k+\nu)}{k+\nu+2}, \quad \mathcal{J}_k(f^e_{1,\eta}) = k(k+2\eta),$$

respectively, where Γ stands for Euler's Gamma function. Note that the lower bound k^2 of the radial information for shape/scale $\mathcal{J}_k(f_1)$ (see Hallin and Paindaveine 2005a) is achieved for arbitrarily heavy tails, that is, as $\nu \to 0$ and $\eta \to 0$ in the classes of k-variate Student and power-exponential distributions, respectively.

The following notation is needed in the statement of ULAN and will be used throughout the paper. Write $\mathbf{V}^{\otimes 2}$ for the Kronecker product $\mathbf{V} \otimes \mathbf{V}$. Denoting by $\mathbf{e}_\ell$ the ℓth vector of the canonical basis of $\mathbb{R}^k$, let $\mathbf{K}_k := \sum_{i,j=1}^k (\mathbf{e}_i \mathbf{e}_j') \otimes (\mathbf{e}_j \mathbf{e}_i')$ be the $k^2 \times k^2$ *commutation matrix*, and put $\mathbf{J}_k := \sum_{i,j=1}^k (\mathbf{e}_i \mathbf{e}_j') \otimes (\mathbf{e}_i \mathbf{e}_j') = (\mathrm{vec}\, \mathbf{I}_k)(\mathrm{vec}\, \mathbf{I}_k)'$ (as usual, $\mathrm{vec}\,(\mathbf{A})$ stands for the vector resulting from stacking the columns of $\mathbf{A}$ on top of each other). Also let $\mathbf{N}_k$ be such that $\mathbf{N}_k(\mathrm{vec}\, \mathbf{v}) = \mathrm{ve\overset{\circ}{c}h}(\mathbf{v})$ for any $k \times k$ matrix $\mathbf{v}$ and $\mathbf{M}_k$ be such that $\mathbf{M}_k'(\mathrm{ve\overset{\circ}{c}h}\mathbf{v}) = \mathrm{vec}\,(\mathbf{v})$ for any symmetric $k \times k$ matrix $\mathbf{v} = (v_{ij})$ such that $v_{11} = 0$. Finally, although any square root $\mathbf{V}^{1/2}$ of $\mathbf{V}$ (satisfying $\mathbf{V}^{1/2}\mathbf{V}^{1/2\prime} = \mathbf{V}$) can be used in the results below (provided, of course, it is used in a consistent way), we will use the symmetric root in order to save superfluous primes.

We then have the following ULAN result (See Hallin and Paindaveine 2005a for the explicit form of the *central sequence* $\boldsymbol{\Delta}^{(n)}_{\boldsymbol{\vartheta}^{(n)};f_1}$ and a proof).

Proposition 2.1 *Under Assumption (A), the family* $\mathcal{P}^n_{f_1}$ *is ULAN, that is, for any* $\boldsymbol{\vartheta}^{(n)} = (\boldsymbol{\theta}^{(n)\prime}, (\sigma^{(n)})^2, (\mathrm{ve\overset{\circ}{c}h}\mathbf{V}^{(n)})')' = \boldsymbol{\vartheta} + O(n^{-1/2})$ *and any bounded sequence* $\boldsymbol{\tau}^{(n)} := (\boldsymbol{\tau}_1^{(n)\prime}, \tau_2^{(n)}, \boldsymbol{\tau}_3^{(n)\prime})'$ $:= (\mathbf{t}^{(n)\prime}, s^{(n)}, (\mathrm{ve\overset{\circ}{c}h}\mathbf{v}^{(n)})')' \in \mathbb{R}^{k+k(k+1)/2}$, *we have, under* $\mathrm{P}^n_{\boldsymbol{\vartheta}^{(n)};f_1}$,

$$\log\left(d\mathrm{P}^n_{\boldsymbol{\vartheta}^{(n)}+n^{-1/2}\boldsymbol{\tau}^{(n)};f_1} / d\mathrm{P}^n_{\boldsymbol{\vartheta}^{(n)};f_1}\right) = (\boldsymbol{\tau}^{(n)})'\boldsymbol{\Delta}^{(n)}_{\boldsymbol{\vartheta}^{(n)};f_1} - \frac{1}{2}(\boldsymbol{\tau}^{(n)})'\boldsymbol{\Gamma}_{\boldsymbol{\vartheta};f_1}\boldsymbol{\tau}^{(n)} + o_{\mathrm{P}}(1),$$

for some sequence of random vectors $\boldsymbol{\Delta}^{(n)}_{\boldsymbol{\vartheta}^{(n)};f_1}$ *that are, still under* $\mathrm{P}^n_{\boldsymbol{\vartheta}^{(n)};f_1}$, *asymptotically normal with mean zero and covariance matrix*

$$\boldsymbol{\Gamma}_{\boldsymbol{\vartheta};f_1} := \begin{pmatrix} \boldsymbol{\Gamma}_{\boldsymbol{\vartheta};f_1;11} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \Gamma_{\boldsymbol{\vartheta};f_1;22} & \boldsymbol{\Gamma}'_{\boldsymbol{\vartheta};f_1;32} \\ \mathbf{0} & \boldsymbol{\Gamma}_{\boldsymbol{\vartheta};f_1;32} & \boldsymbol{\Gamma}_{\boldsymbol{\vartheta};f_1;33} \end{pmatrix}, \tag{2.1}$$

with

$$\boldsymbol{\Gamma}_{\boldsymbol{\vartheta};f_1;11} := \frac{\mathcal{I}_k(f_1)}{k\sigma^2}\mathbf{V}^{-1}, \quad \Gamma_{\boldsymbol{\vartheta};f_1;22} := \frac{\mathcal{J}_k(f_1) - k^2}{4\sigma^4}, \quad \boldsymbol{\Gamma}_{\boldsymbol{\vartheta};f_1;32} := \frac{\mathcal{J}_k(f_1) - k^2}{4k\sigma^2}\mathbf{M}_k\, \mathrm{vec}\,(\mathbf{V}^{-1}),$$

and

$$\boldsymbol{\Gamma}_{\boldsymbol{\vartheta};f_1;33} := \frac{1}{4}\mathbf{M}_k\left(\mathbf{V}^{\otimes 2}\right)^{-1/2}\left[\frac{\mathcal{J}_k(f_1)}{k(k+2)}(\mathbf{I}_{k^2} + \mathbf{K}_k + \mathbf{J}_k) - \mathbf{J}_k\right]\left(\mathbf{V}^{\otimes 2}\right)^{-1/2}\mathbf{M}_k'. \tag{2.2}$$

The block-diagonal structure of the information matrix (2.1) implies that the non-specification of the location centre $\boldsymbol{\theta}$ does not affect the optimal parametric performance when estimating $\mathbf{V}$; more precisely, the optimal asymptotic covariance matrix that can be achieved (at $\mathrm{P}^n_{\boldsymbol{\theta},\sigma,\mathbf{V},f_1}$) by an estimator $\mathbf{V}^{(n)}$ of $\mathbf{V}$ is the same in $\mathcal{P}^n_f = \{\mathrm{P}^n_{\boldsymbol{\theta},\sigma,\mathbf{V},f_1}, \boldsymbol{\theta} \in \mathbb{R}^k, \mathbf{V} \in \mathcal{V}\}$ as in $\mathcal{P}^n_{\boldsymbol{\theta},f} := \{\mathrm{P}^n_{\boldsymbol{\theta},\sigma,\mathbf{V},f_1}, \mathbf{V} \in \mathcal{V}\}$, where $\boldsymbol{\theta}$ is specified. Since moreover $\boldsymbol{\Gamma}_{\boldsymbol{\vartheta};f_1}$ itself does not depend on $\boldsymbol{\theta}$, this optimal performance does not depend on $\boldsymbol{\theta}$.

3 Parametric efficiency bounds.

The ULAN result of Proposition 2.1 is about the "unspecified scale" model $\mathcal{P}^n_{f_1}$, but automatically entails ULAN for the "specified scale" models $\mathcal{P}^n_f$, the information matrices of which are obtained by deleting the σ^2 rows and columns in (2.1).

Parametric efficiency at $\mathrm{P}^n_{\boldsymbol{\vartheta},f_1}$, thus is characterized by the *parametric information matrix for shape* $\boldsymbol{\Gamma}_{f_1}(\mathbf{V}) := \boldsymbol{\Gamma}_{\boldsymbol{\vartheta};f_1;33}$ in (2.2), which does not depend on σ nor on $\boldsymbol{\theta}$ (whence the notation). An estimator $\mathbf{V}^{(n)}$ of $\mathbf{V}$ thus is f_1*-parametrically efficient* (here again, the terminology is justified by the fact that the information matrix does not depend on the value of the scale parameter σ) iff, for all admissible $\boldsymbol{\vartheta} = (\boldsymbol{\theta}', \sigma^2, (\overset{\circ}{\mathrm{vech}}\mathbf{V})')'$, $n^{1/2}\overset{\circ}{\mathrm{vech}}\,(\mathbf{V}^{(n)} - \mathbf{V}) \xrightarrow{\mathcal{L}} \mathcal{N}(\mathbf{0}, (\boldsymbol{\Gamma}_{f_1}(\mathbf{V}))^{-1})$, under $\mathrm{P}^n_{\boldsymbol{\vartheta},f_1}$, as $n \to \infty$, or, in terms of vec $\mathbf{V}$, iff

$$n^{1/2}\mathrm{vec}\left(\mathbf{V}^{(n)} - \mathbf{V}\right) \xrightarrow{\mathcal{L}} \mathcal{N}\left(\mathbf{0}, \mathbf{M}_k'(\boldsymbol{\Gamma}_{f_1}(\mathbf{V}))^{-1}\mathbf{M}_k\right), \tag{3.1}$$

under $\mathrm{P}^n_{\boldsymbol{\vartheta},f_1}$, as $n \to \infty$. One of the main objectives of this paper is to provide an explicit ($\mathbf{M}_k$-free) expression for the asymptotic covariance in (3.1), allowing for a comparison with the corresponding semiparametric asymptotic covariance matrix achievable in the more realistic unspecified scale setup, that is, in $\mathcal{P}^n_{f_1} := \{\mathrm{P}^n_{\boldsymbol{\theta},\sigma,\mathbf{V},f_1}, \boldsymbol{\theta} \in \mathbb{R}^k, \sigma^2 > 0, \mathbf{V} \in \mathcal{V}\}$. This performance coincides with the f_1-semiparametric one, achievable in $\mathcal{P}^n$, which was determined by Hallin, Oja and Paindaveine (2005): see Section 4 below.

Denoting by $\mathbf{e}_{k^2,1}$ the first vector of the canonical basis of $\mathbb{R}^{k^2}$, let

$$\begin{aligned}\mathbf{Q}_k^{(r,s)}(\mathbf{V}) := r\left\{\left[\mathbf{I}_{k^2} + \mathbf{K}_k\right](\mathbf{V}^{\otimes 2})\right\} + (s-2r)\left\{(\mathbf{V}^{\otimes 2})\,\mathbf{e}_{k^2,1}\mathbf{e}_{k^2,1}'(\mathbf{V}^{\otimes 2})\right\} \\ + s\left\{(\mathrm{vec}\mathbf{V})\,(\mathrm{vec}\mathbf{V})' - (\mathrm{vec}\mathbf{V})\,\mathbf{e}_{k^2,1}'(\mathbf{V}^{\otimes 2}) - (\mathbf{V}^{\otimes 2})(\mathbf{e}_{k^2,1})\,(\mathrm{vec}\mathbf{V})'\right\}.\end{aligned} \tag{3.2}$$

The following lemma provides the key result in the derivation of an explicit expression of in (3.1) above (see the appendix for the proof).

Lemma 3.1 *Let* $\mathbf{V} = (V_{ij}) \in \mathcal{V}$*. Then,*
(i) for all $a \neq 0$ *and* $2a + (k-1)b \neq 0$,

$$\left\{\frac{1}{4}\mathbf{M}_k\left(\mathbf{V}^{\otimes 2}\right)^{-1/2}\left[a(\mathbf{I}_{k^2} + \mathbf{K}_k) + b\mathbf{J}_k\right]\left(\mathbf{V}^{\otimes 2}\right)^{-1/2}\mathbf{M}_k'\right\}^{-1} = \mathbf{N}_k\mathbf{Q}_k^{(A,B)}(\mathbf{V})\mathbf{N}_k', \tag{3.3}$$

with $A = a^{-1}$ *and* $B := -2a^{-1}b/(2a + (k-1)b)$*;*
(ii) for all $r, s \in \mathbb{R}$, $\mathbf{M}_k'\mathbf{N}_k\mathbf{Q}_k^{(r,s)}(\mathbf{V}) = \mathbf{Q}_k^{(r,s)}(\mathbf{V}) = \mathbf{Q}_k^{(r,s)}(\mathbf{V})\mathbf{N}_k'\mathbf{M}_k$*, so that*

(iii) $$\mathbf{M}_k'\left\{\frac{1}{4}\mathbf{M}_k\left(\mathbf{V}^{\otimes 2}\right)^{-1/2}\left[a(\mathbf{I}_{k^2} + \mathbf{K}_k) + b\mathbf{J}_k\right]\left(\mathbf{V}^{\otimes 2}\right)^{-1/2}\mathbf{M}_k'\right\}^{-1}\mathbf{M}_k = \mathbf{Q}_k^{(A,B)}(\mathbf{V}).$$

Using this lemma in (3.1), we directly obtain the following proposition.

Proposition 3.1 *The asymptotic (under* $\mathrm{P}^n_{\boldsymbol{\vartheta},f_1}$*, as* $n \to \infty$*) covariance matrix of* f_1*-parametrically efficient estimators of* $\mathbf{V}$ *is, for all admissible values* $\boldsymbol{\vartheta} = (\boldsymbol{\theta}', \sigma^2, (\overset{\circ}{\mathrm{vech}}\mathbf{V})')'$,

$$\mathbf{M}_k'(\boldsymbol{\Gamma}_{f_1}(\mathbf{V}))^{-1}\mathbf{M}_k = \frac{k(k+2)}{\mathcal{J}_k(f_1)}\,\mathbf{Q}_k^{(1,2\mathcal{M}_k(f_1))}(\mathbf{V}), \tag{3.4}$$

where

$$\mathcal{M}_k(f_1) := \frac{k(k+2) - \mathcal{J}_k(f_1)}{(k+1)(\mathcal{J}_k(f_1) - k^2) + 2k}. \tag{3.5}$$

Note that, since $\mathcal{J}_k(f_1) \geq k^2$, the quantity $\mathcal{M}_k(f_1)$ in (3.5) satisfies

$$-1/(k+1) \leq \mathcal{M}_k(f_1) \leq 1. \tag{3.6}$$

These lower and upper bounds for $\mathcal{M}_k(f_1)$ are achieved in the limiting cases $\eta \to \infty$ and $\eta \to 0$, respectively, within the class of k-variate power-exponential densities $f^e_{1,\eta}$.

As an illustration, we now provide an estimator $\mathbf{V}^{(n)}_{\mathcal{N}}$ that is ϕ_1-parametrically efficient, that is, parametrically efficient in the multinormal case. Under $\mathrm{P}^n_{\boldsymbol{\vartheta},\phi_1}$, the regular sample covariance matrix $\boldsymbol{\Sigma}^{(n)} := (n-1)^{-1}\sum_{i=1}^n(\mathbf{X}_i - \bar{\mathbf{X}})(\mathbf{X}_i - \bar{\mathbf{X}})'$ (with $\bar{\mathbf{X}} := n^{-1}\sum_{i=1}^n \mathbf{X}_i$) is consistent for $\boldsymbol{\Sigma}_k := a_k^{-1}\sigma^2\mathbf{V}$ (where a_k was defined in page 2). Actually, it is easy to show that

$$n^{1/2}\mathrm{vec}(\boldsymbol{\Sigma}^{(n)} - \boldsymbol{\Sigma}_k) \xrightarrow{\mathcal{L}} \mathcal{N}(\mathbf{0}, [\mathbf{I}_{k^2} + \mathbf{K}_k]\,(\boldsymbol{\Sigma}_k^{\otimes 2})),$$

under $\mathrm{P}^n_{\boldsymbol{\vartheta},\phi_1}$, as $n \to \infty$. If the scale parameter σ is known, one can define the estimator of shape

$$\mathbf{V}^{(n)}_{\mathcal{N}} := \frac{a_k\boldsymbol{\Sigma}^{(n)}}{\sigma^2} - \frac{1}{(\Sigma^{(n)}_{11})^2}\left(\frac{a_k\Sigma^{(n)}_{11}}{\sigma^2} - 1\right)\left(\boldsymbol{\Sigma}^{(n)}\mathbf{e}_1\right)\left(\boldsymbol{\Sigma}^{(n)}\mathbf{e}_1\right)' \tag{3.7}$$

($\mathbf{e}_\ell$ stands for the ℓth vector of the canonical basis of $\mathbb{R}^k$). Applying Slutzky's Lemma, we obtain, under $\bigcup_{\boldsymbol{\theta}}\{\mathrm{P}^n_{\boldsymbol{\theta},\sigma,\mathbf{V},\phi_1}\}$,

$$n^{1/2}\mathrm{vec}\left(\mathbf{V}^{(n)}_{\mathcal{N}} - \mathbf{V}\right) = \frac{a_k}{\sigma^2}\left[\mathbf{I}_{k^2} - (\mathbf{V}^{\otimes 2})\,\mathbf{e}_{k^2,1}\mathbf{e}'_{k^2,1}\right]\left[n^{1/2}\mathrm{vec}\left(\boldsymbol{\Sigma}^{(n)} - \boldsymbol{\Sigma}_k\right)\right] + o_{\mathrm{P}}(1) \xrightarrow{\mathcal{L}} \mathcal{N}(\mathbf{0}, \mathbf{A}),$$

as $n \to \infty$. The asymptotic covariance matrix $\mathbf{A}$, after some standard algebra, reduces to

$$\left[\mathbf{I}_{k^2} - (\mathbf{V}^{\otimes 2})\,\mathbf{e}_{k^2,1}\mathbf{e}'_{k^2,1}\right]\left[\mathbf{I}_{k^2} + \mathbf{K}_k\right](\mathbf{V}^{\otimes 2})\left[\mathbf{I}_{k^2} - (\mathbf{V}^{\otimes 2})\,\mathbf{e}_{k^2,1}\mathbf{e}'_{k^2,1}\right]' = \mathbf{Q}_k^{(1,0)}(\mathbf{V}),$$

which is the value of the asymptotic covariance matrix in (3.4) at $f_1 = \phi_1$. Consequently, the estimator $\mathbf{V}^{(n)}_{\mathcal{N}}$ in (3.7) is parametrically efficient in the multinormal case.

4 Semiparametric efficiency bounds.

In the more realistic setup where $\boldsymbol{\theta}$, σ^2, and f_1 remain unspecified and play the role of a nuisance, estimators are optimal if they reach semiparametric efficiency bounds, either at some prespecified radial density f_1, or at any density in $\mathcal{F}$. The corresponding semiparametric efficiency bound, at f_1, is the inverse of the *efficient Fisher information for shape*

$$\begin{aligned}\boldsymbol{\Gamma}^*_{f_1}(\mathbf{V}) &= \boldsymbol{\Gamma}_{\boldsymbol{\vartheta};f_1;33} - \boldsymbol{\Gamma}_{\boldsymbol{\vartheta};f_1;32}\Gamma^{-1}_{\boldsymbol{\vartheta};f_1;22}\boldsymbol{\Gamma}'_{\boldsymbol{\vartheta};f_1;32} \\ &= \frac{\mathcal{J}_k(f_1)}{4k(k+2)}\mathbf{M}_k\left(\mathbf{V}^{\otimes 2}\right)^{-1/2}\left[\mathbf{I}_{k^2} + \mathbf{K}_k - \frac{2}{k}\mathbf{J}_k\right]\left(\mathbf{V}^{\otimes 2}\right)^{-1/2}\mathbf{M}'_k; \end{aligned} \tag{4.1}$$

see Hallin and Paindaveine (2005a). More precisely, an estimator $\mathbf{V}_*^{(n)}$ of $\mathbf{V}$ is f_1-*semiparametrically efficient* iff, for all admissible $\boldsymbol{\vartheta} = (\boldsymbol{\theta}', \sigma^2, (\mathrm{v\mathring{e}ch}\mathbf{V})')'$, $n^{1/2}\mathrm{v\mathring{e}ch}\,(\mathbf{V}_*^{(n)} - \mathbf{V}) \xrightarrow{\mathcal{L}} \mathcal{N}(\mathbf{0}, (\boldsymbol{\Gamma}_{f_1}^*(\mathbf{V}))^{-1})$, under $\mathrm{P}^n_{\boldsymbol{\vartheta},f_1}$, as $n \to \infty$, or, in terms of $\mathrm{vec}\,\mathbf{V}$, iff

$$n^{1/2}\mathrm{vec}\left(\mathbf{V}_*^{(n)} - \mathbf{V}\right) \xrightarrow{\mathcal{L}} \mathcal{N}\left(\mathbf{0}, \mathbf{M}_k'(\boldsymbol{\Gamma}_{f_1}^*(\mathbf{V}))^{-1}\mathbf{M}_k\right), \tag{4.2}$$

under $\mathrm{P}^n_{\boldsymbol{\vartheta},f_1}$, as $n \to \infty$. Using Lemma 3.1 again, we obtain

Proposition 4.1 *The asymptotic (under $\mathrm{P}^n_{\boldsymbol{\vartheta},f_1}$, as $n \to \infty$) covariance matrix of f_1-semiparametrically efficient estimators of $\mathbf{V}$ is, for all admissible values $\boldsymbol{\vartheta} = (\boldsymbol{\theta}', \sigma^2, (\mathrm{v\mathring{e}ch}\mathbf{V})')'$,*

$$\mathbf{M}_k'(\boldsymbol{\Gamma}_{f_1}^*(\mathbf{V}))^{-1}\mathbf{M}_k = \frac{k(k+2)}{\mathcal{J}_k(f_1)}\,\mathbf{Q}_k^{(1,2)}(\mathbf{V}), \tag{4.3}$$

with $\mathbf{Q}_k^{(r,s)}$ is defined in (3.2).

We conclude this section by defining an estimator $\mathbf{V}_{\mathcal{N}*}^{(n)}$ that is semiparametrically efficient in the multinormal case (at $f_1 = \phi_1$). Let

$$\mathbf{V}_{\mathcal{N}*}^{(n)} := \frac{\boldsymbol{\Sigma}^{(n)}}{\Sigma_{11}^{(n)}}, \tag{4.4}$$

where $\boldsymbol{\Sigma}^{(n)}$ denotes the sample covariance matrix. Slutzky's Lemma yields

$$n^{1/2}\mathrm{vec}\left(\mathbf{V}_{\mathcal{N}*}^{(n)} - \mathbf{V}\right) = \frac{a_k}{\sigma^2}\left[\mathbf{I}_{k^2} - (\mathrm{vec}\mathbf{V})\,(\mathbf{e}_{k^2,1})'\right]\left[n^{1/2}\mathrm{vec}\left(\boldsymbol{\Sigma}^{(n)} - \boldsymbol{\Sigma}_k\right)\right] + o_{\mathrm{P}}(1) \xrightarrow{\mathcal{L}} \mathcal{N}\left(\mathbf{0}, \mathbf{B}\right),$$

under $\bigcup_{\boldsymbol{\theta}}\left\{\mathrm{P}^n_{\boldsymbol{\theta},\sigma,\mathbf{V},\phi_1}\right\}$, as $n \to \infty$, where the matrix $\mathbf{B}$, after standard algebra, reduces to

$$\left[\mathbf{I}_{k^2} - (\mathbf{V}^{\otimes 2})\,\mathbf{e}_{k^2,1}\mathbf{e}_{k^2,1}'\right]\left[\mathbf{I}_{k^2} + \mathbf{K}_k\right](\mathbf{V}^{\otimes 2})\left[\mathbf{I}_{k^2} - (\mathbf{V}^{\otimes 2})\,\mathbf{e}_{k^2,1}\mathbf{e}_{k^2,1}'\right]' = \mathbf{Q}_k^{(1,2)}(\mathbf{V}),$$

which is the value, at $f_1 = \phi_1$, of the asymptotic covariance matrix in (4.3). The semiparametric efficiency of $\mathbf{V}_{\mathcal{N}*}^{(n)}$ in the multinormal case follows.

5 The cost of unspecified scale/radial density.

We now are able to compare the asymptotic performances of f_1-parametrically and f_1-semiparametrically efficient estimators for shape, and to quantify the corresponding efficiency losses. The main result of this paper is stated in the next proposition; see the appendix for the proof. For the sake of simplicity, we restrict to the case where the shape matrix $\mathbf{V}$ is diagonal, and, in order to improve readability, write $\mathrm{AVar}[S^{(n)}]$ and $\mathrm{ACov}[S^{(n)}, T^{(n)}]$ for the asymptotic variance of S_n and asymptotic covariance of S^n and T^n, respectively.

Proposition 5.1 *Assume that $\mathbf{V} = (V_{ij}) \in \mathcal{V}$ is diagonal. Then, under $\bigcup_{\boldsymbol{\theta}}\bigcup_{\sigma}\left\{\mathrm{P}^n_{\boldsymbol{\theta},\sigma,\mathbf{V},f_1}\right\}$, for a f_1-parametrically efficient estimator $\mathbf{V}^{(n)} = (V_{ij}^{(n)})$ of $\mathbf{V}$, we have*

$$\mathrm{AVar}[n^{1/2}(V_{ii}^{(n)} - V_{ij})] = \frac{2k(k+2)}{\mathcal{J}_k(f_1)}\Big(\mathcal{M}_k(f_1) + 1\Big)V_{ii}^2, \quad i = 2, \ldots, k, \tag{5.1}$$

$$\mathrm{AVar}[n^{1/2}(V_{ij}^{(n)} - V_{ij})] = \frac{k(k+2)}{\mathcal{J}_k(f_1)}V_{ii}V_{jj}, \quad i, j = 1, \ldots, k,\ i \neq j, \quad \textit{and} \tag{5.2}$$

$$\mathrm{ACov}[n^{1/2}(V_{ii}^{(n)} - V_{ii}), n^{1/2}(V_{jj}^{(n)} - V_{jj})] = \frac{2k(k+2)}{\mathcal{J}_k(f_1)}\mathcal{M}_k(f_1)V_{ii}V_{jj},\ i, j = 2, \ldots, k,\ i \neq j, \tag{5.3}$$

whereas, for a f_1-semiparametrically efficient estimator $\mathbf{V}_*^{(n)} = (V_{*,ij}^{(n)})$ of $\mathbf{V}$,

$$\begin{aligned}
\mathrm{AVar}[n^{1/2}(V_{*,ii}^{(n)} - V_{ij})] &= \frac{4k(k+2)}{\mathcal{J}_k(f_1)} V_{ii}^2, \quad i = 2, \ldots, k, \\
\mathrm{AVar}[n^{1/2}(V_{*,ij}^{(n)} - V_{ij})] &= \frac{k(k+2)}{\mathcal{J}_k(f_1)} V_{ii}V_{jj}, \quad i,j = 1, \ldots, k,\ i \neq j, \quad \textit{and} \\
\mathrm{ACov}[n^{1/2}(V_{*,ii}^{(n)} - V_{ii}), n^{1/2}(V_{*,jj}^{(n)} - V_{jj})] &= \frac{2k(k+2)}{\mathcal{J}_k(f_1)} V_{ii}V_{jj}, \quad i,j = 2, \ldots, k,\ i \neq j.
\end{aligned}$$

In both cases, all other entries of asymptotic covariance matrices are zero.

Note that, both for $\mathbf{W}^{(n)} = \mathbf{V}^{(n)}$ and $\mathbf{W}^{(n)} = \mathbf{V}_*^{(n)}$,

$$\begin{aligned}
&\mathrm{ACov}[n^{1/2}(W_{ii}^{(n)} - V_{ii}), n^{1/2}(W_{jj}^{(n)} - V_{jj})] \\
&= \left(\mathrm{AVar}[n^{1/2}(W_{ii}^{(n)} - V_{ii})]\right)^{1/2} \left(\mathrm{AVar}[n^{1/2}(W_{jj}^{(n)} - V_{jj})]\right)^{1/2} - 2\,\mathrm{AVar}[n^{1/2}(W_{ij}^{(n)} - V_{ij})],
\end{aligned}$$

so that the asymptotic covariance matrices under consideration are completely determined by the quantities $\mathrm{AVar}[n^{1/2}(W_{ii}^{(n)} - V_{ii})]$ $(i = 2, \ldots, k)$ and $\mathrm{AVar}[n^{1/2}(W_{ij}^{(n)} - V_{ij})]$ $(i, j = 1, \ldots, k,\ i \neq j)$; see Ollila, Croux, and Oja (2004) and Ollila, Oja, and Croux (2003) for a similar phenomenon. The difference between the performances of f_1-parametrically efficient and f_1-semiparametrically efficient estimators of shape can therefore be fully characterized by the asymptotic relative efficiencies associated with the components $W_{ii}^{(n)}$ $(i = 2, \ldots, k)$ and $W_{ij}^{(n)}$ $(i, j = 1, \ldots, k,\ i \neq j)$ of such estimators. Now, it directly follows from Proposition 5.1 that

$$\mathrm{ARE}_{k,f_1}\left[V_{*,ii}^{(n)}/V_{ii}^{(n)}\right] = \frac{\mathcal{M}_k(f_1)+1}{2}, \quad i = 2, \ldots, k, \text{ and} \tag{5.4}$$

$$\mathrm{ARE}_{k,f_1}\left[V_{*,ij}^{(n)}/V_{ij}^{(n)}\right] = 1, \quad i, j = 1, \ldots, k,\ i \neq j.$$

As a first conclusion, this shows that, irrespective of the underlying radial density type f_1, there is no efficiency loss due to the non-specification of scale when estimating off-diagonal entries of shape matrices. However, there is some efficiency loss when estimating diagonal entries. In view of the lower and upper bounds on $\mathcal{M}_k(f_1)$ in (3.6), we have

$$\frac{k}{2(k+1)} \leq \mathrm{ARE}_{k,f_1}\left[V_{*,ii}^{(n)}/V_{ii}^{(n)}\right] \leq 1, \quad i = 2, \ldots, k. \tag{5.5}$$

As an illustration, at k-dimensional Student densities with ν degrees of freedom, we have (with obvious notation)

$$\mathrm{ARE}_{k,\nu}\left[V_{*,ii}^{(n)}/V_{ii}^{(n)}\right] = \frac{\nu+2}{2(\nu+1)}, \quad i = 2, \ldots, k, \tag{5.6}$$

which does not depend on the space dimension k. These AREs are decreasing with ν (equivalently, they are increasing with the tail weight of the underlying Student distribution); their value is minimal (the efficiency loss is maximal) at the multinormal case ($\nu \to \infty$), where the AREs in (5.6) take the pretty low value of .5. On the contrary, in the limit as $\nu \to 0$, no efficiency loss is incurred. This latter remark is compatible with the fact that Tyler's estimator of shape, which does not take any radial information into account—and, in particular, does not take advantage of any scale information—is optimal as $\nu \to 0$ (see Section 3.2 in Hallin, Oja, and Paindaveine 2005 for a precise statement).

At the k-dimensional power-exponential densities $f^e_{1,\eta}$, we obtain

$$\mathrm{ARE}_{k,\eta}\Big[V^{(n)}_{*ii}/V^{(n)}_{ii}\Big] = \frac{k\eta+2}{2((k+1)\eta+1)} \qquad i=2,\ldots,k.$$

Again, no efficiency loss is incurred under arbitrarily heavy tails (as $\eta \to 0$), whereas the maximal efficiency loss occurs at extremely light-tailed densities (as $\eta \to \infty$); in this case, note that the maximal efficiency loss within this class of densities coincides with the overall maximal efficiency loss in (5.5), namely, an ARE value of $k/[2(k+1)]$.

	degrees of freedom ν of the Student density						
k	0	1	3	8	15	20	∞
	1.000	0.750	0.625	0.556	0.531	0.524	0.500

	parameter η of the power-exponential density						
k	0	0.1	0.5	1	2	5	∞
2	1.000	0.846	0.600	0.500	0.429	0.375	0.333
3	1.000	0.821	0.583	0.500	0.444	0.405	0.375
4	1.000	0.800	0.571	0.500	0.455	0.423	0.400
6	1.000	0.765	0.556	0.500	0.467	0.444	0.429
10	1.000	0.714	0.538	0.500	0.478	0.464	0.455
∞	—	0.500	0.500	0.500	0.500	0.500	0.500

Table 1: Numerical values, for $k = 2, 3, 4, 6, 10$, and $k \to \infty$, of the AREs (5.4), under k-dimensional Student densities (with ν degrees of freedom, $\nu = 1, 3, 8, 15, 20$, along with the limiting values obtained for $\nu \to 0$ and $\nu \to \infty$), and under k-dimensional power-exponential densities (with $\eta = .1, .5, 1, 2, 5$, along with the limiting values obtained for $\eta \to 0$ and $\eta \to \infty$); for Student densities, the ARE values do not depend on the space dimension k.

Numerical values of the AREs given in (5.4) under various Student and power-exponential densities, and for various space dimensions k, are provided in Table 1. [Comments].

A Appendix.

In this final section, we prove Lemma 3.1 and Proposition 5.1.

Proof of Lemma 3.1. (i) Write $\mathbf{\Gamma}_{a,b}$ for the matrix in braces in the left-hand side of (3.3). Using part (ii) of the lemma (as we shall see, the proof of (ii) does not require (i)) and the identities $\mathbf{K}_k^2 = \mathbf{I}_{k^2}$, $\mathbf{K}_k\mathbf{J}_k = \mathbf{K}_k = \mathbf{J}_k\mathbf{K}_k$, $\mathbf{J}_k^2 = k\mathbf{J}_k$, $\mathrm{vec}\,(\mathbf{ABC}) = (\mathbf{C}' \otimes \mathbf{A})\,(\mathrm{vec}\,\mathbf{B})$, $(\mathrm{vec}\,\mathbf{A})'(\mathrm{vec}\,\mathbf{B}) = \mathrm{tr}\,(\mathbf{A}'\mathbf{B})$, $\mathbf{K}_k(\mathbf{A}\otimes\mathbf{B}) = (\mathbf{B}\otimes\mathbf{A})\mathbf{K}_k$, $\mathbf{K}_k(\mathrm{vec}\,\mathbf{A}) = \mathrm{vec}\,(\mathbf{A}')$ (holding for all $k\times k$ matrices $\mathbf{A}$, $\mathbf{B}$, and $\mathbf{C}$), lengthy but straightforward calculations yield

$$\mathbf{\Gamma}_{a,b}\Big(\mathbf{N}_k\mathbf{Q}_k^{(A,B)}(\mathbf{V})\mathbf{N}_k'\Big) = \frac{1}{2}\mathbf{M}_k\Big[\mathbf{I}_{k^2}+\mathbf{K}_k\Big]\mathbf{N}_k' + \mathbf{M}_k\mathbf{e}_{k^2,1}\Big[(2aB-4)\mathbf{e}_{k^2,1}'(\mathbf{V}^{\otimes 2}) - 2aB\,(\mathrm{vec}\mathbf{V})'\Big]\mathbf{N}_k'.$$

Now, since

(a) $\mathbf{M}_k\mathbf{e}_{k^2,1} = \mathbf{0}$ (it follows from the definition of $\mathbf{M}_k$ that its first column is the $(k(k+1)/2-1)$-dimensional zero-vector),

(b) $\mathbf{M}_k\mathbf{K}_k = \mathbf{M}_k$ (since $\mathbf{K}_k\mathbf{M}'_k(\overset{\circ}{\mathrm{vech}}\,\mathbf{w}) = \mathbf{K}_k(\mathrm{vec}\,\mathbf{w}) = \mathrm{vec}\,\mathbf{w} = \mathbf{M}'_k(\overset{\circ}{\mathrm{vech}}\,\mathbf{w})$ for all symmetric $k \times k$ matrix $\mathbf{w} = (w_{ij})$ such that $w_{11} = 0$), and

(c) $\mathbf{M}_k\mathbf{N}'_k = \mathbf{I}_{k(k+1)/2-1}$ (since $\mathbf{N}_k\mathbf{M}'_k(\overset{\circ}{\mathrm{vech}}\,\mathbf{w}) = \mathbf{N}_k(\mathrm{vec}\,\mathbf{w}) = \overset{\circ}{\mathrm{vech}}\,\mathbf{w}$ for all symmetric $k \times k$ matrix $\mathbf{w} = (w_{ij})$ such that $w_{11} = 0$),

we obtain that $\mathbf{N}_k\mathbf{Q}_k^{(A,B)}(\mathbf{V})\mathbf{N}'_k$ is a right-inverse of $\mathbf{\Gamma}_{a,b}$. Since both $\mathbf{N}_k\mathbf{Q}_k^{(A,B)}(\mathbf{V})\mathbf{N}'_k$ and $\mathbf{\Gamma}_{a,b}$ are symmetric, this right-inverse is also a left-inverse.

(ii) Since $\mathbf{Q}_k^{(r,s)}(\mathbf{V})$ is symmetric, it is clearly sufficient to prove that $\mathbf{M}'_k\mathbf{N}_k\mathbf{Q}_k^{(r,s)}(\mathbf{V}) = \mathbf{Q}_k^{(r,s)}(\mathbf{V})$. Now, it is easily seen that

$$\mathbf{Q}_k^{(r,s)}(\mathbf{V}) = \Big[\mathbf{I}_{k^2} - (\mathbf{V}^{\otimes 2})\,\mathbf{e}_{k^2,1}\mathbf{e}'_{k^2,1}\Big]\Big\{r\,\Big[\mathbf{I}_{k^2} + \mathbf{K}_k\Big](\mathbf{V}^{\otimes 2}) + s\,(\mathrm{vec}\mathbf{V})(\mathrm{vec}\mathbf{V})'\Big\}\Big[\mathbf{I}_{k^2} - (\mathbf{V}^{\otimes 2})\,\mathbf{e}_{k^2,1}\mathbf{e}'_{k^2,1}\Big]', \tag{A.1}$$

so that it is sufficient to show that

$$\mathbf{M}'_k\mathbf{N}_k\Big[\mathbf{I}_{k^2} - (\mathbf{V}^{\otimes 2})\,\mathbf{e}_{k^2,1}\mathbf{e}'_{k^2,1}\Big][\mathbf{I}_{k^2} + \mathbf{K}_k] = \Big[\mathbf{I}_{k^2} - (\mathbf{V}^{\otimes 2})\,\mathbf{e}_{k^2,1}\mathbf{e}'_{k^2,1}\Big][\mathbf{I}_{k^2} + \mathbf{K}_k] \tag{A.2}$$

and

$$\mathbf{M}'_k\mathbf{N}_k\Big[\mathbf{I}_{k^2} - (\mathbf{V}^{\otimes 2})\,\mathbf{e}_{k^2,1}\mathbf{e}'_{k^2,1}\Big](\mathrm{vec}\mathbf{V}) = \Big[\mathbf{I}_{k^2} - (\mathbf{V}^{\otimes 2})\,\mathbf{e}_{k^2,1}\mathbf{e}'_{k^2,1}\Big](\mathrm{vec}\mathbf{V}). \tag{A.3}$$

But, letting $\mathbf{E}_{ij} := \mathbf{e}_i\mathbf{e}'_j + \mathbf{e}_j\mathbf{e}'_i$, we have

$$\begin{aligned}\Big[\mathbf{I}_{k^2} - (\mathbf{V}^{\otimes 2})\,\mathbf{e}_{k^2,1}\mathbf{e}'_{k^2,1}\Big][\mathbf{I}_{k^2} + \mathbf{K}_k] &= \mathbf{I}_{k^2} + \mathbf{K}_k - 2\,(\mathbf{V}^{\otimes 2})\,\mathbf{e}_{k^2,1}\mathbf{e}'_{k^2,1}\\ &= \frac{1}{2}\sum_{i,j=1}^{k}(\mathrm{vec}\,\mathbf{E}_{ij})(\mathrm{vec}\,\mathbf{E}_{ij})' - 2\Big(\mathrm{vec}\Big((\mathbf{V}\mathbf{e}_1)(\mathbf{V}\mathbf{e}_1)'\Big)\Big)\mathbf{e}'_{k^2,1}\\ &= \frac{1}{2}\sum_{\substack{i,j=1\\(i,j)\neq(1,1)}}^{k}(\mathrm{vec}\,\mathbf{E}_{ij})(\mathrm{vec}\,\mathbf{E}_{ij})' + 2\Big(\mathrm{vec}\Big(\mathbf{e}_1\mathbf{e}'_1 - (\mathbf{V}\mathbf{e}_1)(\mathbf{V}\mathbf{e}_1)'\Big)\Big)\mathbf{e}'_{k^2,1}.\end{aligned}$$

This establishes (A.2) since $\mathbf{M}'_k\mathbf{N}_k\,(\mathrm{vec}\,\mathbf{w}) = (\mathrm{vec}\,\mathbf{w})$ for all symmetric $k \times k$ matrix $\mathbf{w} = (w_{ij})$ such that $w_{11} = 0$ (recall that $\mathbf{V} = (V_{ij}) \in \mathcal{V}$ is symmetric with $V_{11} = 1$). As for (A.3), it follows in a similar way from noting that

$$\begin{aligned}\Big[\mathbf{I}_{k^2} - (\mathbf{V}^{\otimes 2})\,\mathbf{e}_{k^2,1}\mathbf{e}'_{k^2,1}\Big](\mathrm{vec}\mathbf{V}) &= (\mathrm{vec}\mathbf{V}) - (\mathbf{V}^{\otimes 2})\,\mathbf{e}_{k^2,1}\\ &= \mathrm{vec}\Big(\mathbf{V} - (\mathbf{V}\mathbf{e}_1)(\mathbf{V}\mathbf{e}_1)'\Big).\end{aligned}$$

(iii) This part of the lemma trivially follows from (i) and (ii). □

Proof of Proposition 5.1. In view of Proposition 3.1 and (A.1), the asymptotic covariance matrix of $n^{1/2}\mathrm{vec}(\mathbf{V}^{(n)} - \mathbf{V})$ is

$$\frac{k(k+2)}{\mathcal{J}_k(f_1)}\,\mathbf{P}_k(\mathbf{V})\Big\{\Big[\mathbf{I}_{k^2} + \mathbf{K}_k\Big](\mathbf{V}^{\otimes 2}) + 2\mathcal{M}_k(f_1)\,(\mathrm{vec}\mathbf{V})(\mathrm{vec}\mathbf{V})'\Big\}\Big(\mathbf{P}_k(\mathbf{V})\Big)', \tag{A.4}$$

where $\mathbf{P}_k(\mathbf{V}) := \mathbf{I}_{k^2} - (\mathbf{V}^{\otimes 2})\,\mathbf{e}_{k^2,1}\mathbf{e}'_{k^2,1}$. As $\mathbf{V}$ is diagonal with $V_{11} = 1$, $\mathbf{P}_k(\mathbf{V}) = \mathbf{I}_{k^2} - \mathbf{e}_{k^2,1}\mathbf{e}'_{k^2,1}$ is the matrix of the projection on the orthogonal complement of $\{\lambda \mathbf{e}_{k^2,1},\, \lambda \in \mathbb{R}\}$ in $\mathbb{R}^{k^2}$. Consequently, the effect of premultiplying by $\mathbf{P}_k(\mathbf{V})$ and postmultiplying by $(\mathbf{P}_k(\mathbf{V}))'$ the matrix

$$\mathbf{R}_k(\mathbf{V}) := \frac{k(k+2)}{\mathcal{J}_k(f_1)} \left\{ \Big[\mathbf{I}_{k^2} + \mathbf{K}_k\Big](\mathbf{V}^{\otimes 2}) + 2\mathcal{M}_k(f_1)\,(\mathrm{vec}\mathbf{V})\,(\mathrm{vec}\mathbf{V})' \right\} \tag{A.5}$$

is to put to zero all components in the first row and first column of $\mathbf{R}_k(\mathbf{V})$—which reflects the fact that $\mathrm{Cov}\big[V_{11}^{(n)}, V_{ij}^{(n)}\big] = \mathrm{Cov}\big[1, V_{ij}^{(n)}\big] = 0$, for all i, j, n. Consequently, it is sufficient to study the structure of $\mathbf{R}_k(\mathbf{V})$. To this end, we plug $\mathbf{V} = \sum_{i=1} V_{ii}\,\mathbf{e}_i\mathbf{e}'_i$ in $\mathbf{R}_k(\mathbf{V})$. By using that $\mathbf{K}_k((\mathbf{e}_i\mathbf{e}'_i) \otimes (\mathbf{e}_j\mathbf{e}'_j)) = (\mathbf{e}_j\mathbf{e}'_i) \otimes (\mathbf{e}_i\mathbf{e}'_j)$ and that $(\mathrm{vec}\,\mathbf{e}_i\mathbf{e}'_i)(\mathrm{vec}\,\mathbf{e}_j\mathbf{e}'_j)' = (\mathbf{e}_i\mathbf{e}'_j) \otimes (\mathbf{e}_i\mathbf{e}'_j)$, we easily obtain

$$\begin{aligned}\frac{\mathcal{J}_k(f_1)}{k(k+2)}\mathbf{R}_k(\mathbf{V}) = {} & \sum_{i=1}^{k} (\mathbf{e}_i\mathbf{e}'_i) \otimes \left\{ \sum_{j=1}^{k} V_{ii}V_{jj}(2 + 2\mathcal{M}_k(f_1))^{\delta_{i,j}}\,\mathbf{e}_j\mathbf{e}'_j \right\} \\ & + \sum_{\substack{i,j=1 \\ i \neq j}}^{k} (\mathbf{e}_i\mathbf{e}'_j) \otimes \left\{ 2V_{ii}V_{jj}\mathcal{M}_k(f_1)\,\mathbf{e}_i\mathbf{e}'_j + V_{ii}V_{jj}\,\mathbf{e}_j\mathbf{e}'_i \right\},\end{aligned}$$

where $\delta_{i,j} = 1$ if $i = j$ and 0 otherwise. Denoting by $[\mathbf{A}]_{i,j}$ the $(k \times k)$-block in position (i, j) in the $(k^2 \times k^2)$ matrix $\mathbf{A}$, this means that

$$\left[\frac{\mathcal{J}_k(f_1)}{k(k+2)}\mathbf{R}_k(\mathbf{V})\right]_{i,i} = \begin{pmatrix} V_{ii}V_{11} & 0 & & \dots & & 0 \\ 0 & V_{ii}V_{22} & & & & \\ & & \ddots & & & \\ \vdots & & & 2V_{ii}^2(1 + \mathcal{M}_k(f_1)) & & \vdots \\ & & & & \ddots & \\ & & & & V_{ii}V_{k-1,k-1} & 0 \\ 0 & & \dots & & 0 & V_{ii}V_{kk} \end{pmatrix},$$

which proves (5.1) and (5.2), and

$$\left[\frac{\mathcal{J}_k(f_1)}{k(k+2)}\mathbf{R}_k(\mathbf{V})\right]_{i,j} = 2V_{ii}V_{jj}\mathcal{M}_k(f_1)\,\mathbf{e}_i\mathbf{e}'_j + V_{ii}V_{jj}\,\mathbf{e}_j\mathbf{e}'_i, \qquad i \neq j,$$

which proves (5.2) and (5.3). The entries in the asymptotic covariance matrix of the f_1-semiparametrically efficient estimator $\mathbf{V}_*^{(n)}$ are obtained directly by replacing $\mathcal{M}_k(f_1)$ by 1 (compare the asymptotic covariance matrices in Propositions 3.1 and 4.1). □

References

[1] Hallin, M. and D. Paindaveine (2005a). Optimal rank-based tests for sphericity. *Ann. Statist.*, to appear.

[2] Hallin, M. and D. Paindaveine (2005b). Hyperplane-based signed-rank tests for shape homogeneity. Submitted.

[3] Hallin, M., Oja, H., and Paindaveine, D. (2005). Optimal R-estimation of shape. Submitted.

[4] Hallin, M., and B. J. M. Werker (2003). Semiparametric efficiency, distribution-freeness, and invariance, *Bernoulli* **9**, 55-65.

[5] Ollila, E., Croux, C. and Oja, H. (2004). Influence function and asymptotic efficiency of the affine equivariant rank covariance matrix, *Statist. Sinica* **14**, 297-316

[6] Ollila, E., Oja, H. and Croux, C. (2003). The affine equivariant sign covariance matrix: Asymptotic behavior and efficiencies, *J. Multivariate Anal.* **87**, 328-355

Marc Hallin
Département de Mathématique,
I.S.R.O., and E.C.A.R.E.S.
Université Libre de Bruxelles
Campus de la Plaine CP 210
B-1050 Bruxelles Belgium
mhallin@ulb.ac.be

Davy Paindaveine
Département de Mathématique,
I.S.R.O., and E.C.A.R.E.S.
Université Libre de Bruxelles
Campus de la Plaine CP 210
B-1050 Bruxelles Belgium
dpaindav@ulb.ac.be