Journal of Multivariate Analysis 101 (2010) 2372–2388

Contents lists available at ScienceDirect

Journal of Multivariate Analysis

journal homepage: www.elsevier.com/locate/jmva

Asymptotic expansion of the minimum covariance determinant estimators

Eric A. Cator, Hendrik P. Lopuhaä *

Delft University of Technology, The Netherlands

ARTICLE INFO

Article history:
Received 26 August 2009
Available online 23 June 2010

AMS 2008 subject classifications:
62G35
62E20
62H12

Keywords:
Minimum covariance determinant
Asymptotic normality
Influence function

ABSTRACT

In Cator and Lopuhaä (arXiv:math.ST/0907.0079) [3], an asymptotic expansion for the minimum covariance determinant (MCD) estimators is established in a very general framework. This expansion requires the existence and non-singularity of the derivative in a first-order Taylor expansion. In this paper, we prove the existence of this derivative for general multivariate distributions that have a density and provide an explicit expression, which can be used in practice to estimate limiting variances. Moreover, under suitable symmetry conditions on the density, we show that this derivative is non-singular. These symmetry conditions include the elliptically contoured multivariate location-scatter model, in which case we show that the MCD estimators of multivariate location and covariance are asymptotically equivalent to a sum of independent identically distributed vector and matrix valued random elements, respectively. This provides a proof of asymptotic normality and a precise description of the limiting covariance structure for the MCD estimators.

1. Introduction

The minimum covariance determinant (MCD) estimator [14] is one of the most popular robust methods to estimate multivariate location and scatter parameters. These estimators, in particular the covariance estimator, also serve as robust plug-ins in other multivariate statistical techniques, such as principal component analysis [5,16], multivariate linear regression [1,15], discriminant analysis [7], factor analysis [13], canonical correlations [17,18] and errors-in-variables models [6], among others (see also [8] for a more extensive overview). For this reason, the distributional and the robustness properties of the MCD estimators are essential for conducting inference and performing robust estimation in several statistical models.

The MCD estimators have the same high breakdown point as the minimum volume ellipsoid estimators (e.g., see [1,11]). The asymptotic properties were first studied by Butler, Davies and Jhun [2] in the framework of unimodal elliptically contoured densities; they showed that the MCD location estimator converges at $\sqrt{n}$-rate towards a normal distribution with mean equal to the MCD location functional. In the same framework, Croux and Haesbroeck [4] give the expression for the influence function of the MCD covariance functional and use this to compute limiting variances of the MCD covariance estimator. The asymptotic theory was extended and generalized by Cator and Lopuhaä [3], who studied the MCD estimators and the corresponding functional in a very general framework. They establish an asymptotic expansion of the type

$$\widehat{\theta}_n - \theta_0 = -\Lambda'(\theta_0)^{-1}\frac{1}{n}\sum_{i=1}^{n}(\Psi(X_i,\theta_0) - \mathbb{E}\Psi(X_i,\theta_0)) + o_{\mathbb{P}}(n^{-1/2}), \tag{1.1}$$

* Corresponding address: Delft Institute of Applied Mathematics, TU Delft, Mekelweg 4, 2628 CD Delft, The Netherlands.
E-mail address: h.p.lopuhaa@tudelft.nl (H.P. Lopuhaä).
URL: http://ssor.twi.tudelft.nl/~lopuhaa/ (H.P. Lopuhaä).

0047-259X/$ – see front matter
doi:10.1016/j.jmva.2010.06.009

where $\widehat{\theta}_n$ and θ_0 denote vectors consisting of the MCD estimators and the MCD functional at the underlying distribution, respectively, and $\Psi(\cdot, \theta_0)$ is a function that we will specify later on. In principle, from this expansion a central limit theorem for the MCD estimator can be derived. However, the expansion requires the existence and non-singularity of $\Lambda'(\theta_0)$. Moreover, a more explicit expression of its inverse is desirable from a practical point of view, since it determines the limiting variances.

In this paper we show that $\Lambda'(\theta_0)$ exists as long as the underlying distribution P has a density f. Moreover, we provide an explicit expression for $\Lambda'(\theta_0)$ in Theorem 3.1. The expression offers the possibility to estimate the limiting variances of the MCD estimators in *any* model where P has a density. This extends the applicability of the MCD estimator far beyond elliptically contoured models. We will also provide sufficient symmetry conditions on f for $\Lambda'(\theta_0)$ to be non-singular. This includes the special case of elliptically contoured densities

$$f(x) = \det(\Sigma)^{-1/2} h((x-\mu)\Sigma^{-1}(x-\mu)),$$

for which we show that the MCD location and the MCD covariance estimator are asymptotically equivalent to a sum of independent vector and matrix valued random elements, respectively. This exact expansion shows that at elliptically contoured densities the MCD location and MCD covariance estimator are asymptotically independent and yields an explicit central limit theorem for both MCD estimators separately, in such a way that the limiting covariances between elements of the location and covariance estimators can be obtained directly from the covariances between elements of the summands. Furthermore, the expansion for the MCD estimators is needed to obtain the limiting distribution of robustly reweighted least squares estimators for (μ, Σ), if one uses the MCD estimators to assign the weights (see [12]).

The paper is organized as follows. In Section 2, we define the MCD estimators and MCD functionals and discuss some results from [3] that are relevant for our setup. In Section 3, we establish the expression for $\Lambda'(\theta_0)$ in terms of a linear mapping and show that this mapping is non-singular under suitable symmetry conditions. The special case of elliptically contoured densities is considered in Section 4, where we obtain an explicit expression of $\Lambda'(\theta_0)^{-1}$. From this we derive an asymptotic expansion for the estimators, prove asymptotic normality, and derive the influence function of the MCD functionals. As special cases we recover results from [2,4] under weaker conditions.

All proofs have been postponed to an Appendix at the end of the paper.

2. Definition and preliminaries

For a sample $X_1, X_2, \ldots, X_n$ from a distribution P on $\mathbb{R}^k$, the MCD estimator is defined as follows. Fix a fraction $0 < \gamma \leq 1$ and consider subsamples $S \subset \{X_1, \ldots, X_n\}$ that contain $h_n \geq \lceil n\gamma \rceil$ points. Define a corresponding trimmed sample mean and sample covariance matrix by

$$\begin{aligned} \widehat{T}_n(S) &= \frac{1}{h_n}\sum_{X_i\in S} X_i, \\ \widehat{C}_n(S) &= \frac{1}{h_n}\sum_{X_i\in S}(X_i - \widehat{T}_n(S))(X_i - \widehat{T}_n(S))'. \end{aligned} \tag{2.1}$$

Note that each subsample S determines an ellipsoid $E(\widehat{T}_n(S), \widehat{C}_n(S), \widehat{r}_n(S))$, where, for each $\mu \in \mathbb{R}^k$, Σ symmetric positive definite, and $\rho > 0$,

$$E(\mu, \Sigma, \rho) = \left\{x \in \mathbb{R}^k : (x-\mu)'\Sigma^{-1}(x-\mu) \leq \rho^2\right\}, \tag{2.2}$$

and

$$\widehat{r}_n(S) = \inf\left\{s > 0 : P_n\left(E(\widehat{T}_n(S), \widehat{C}_n(S), s)\right) \geq \gamma\right\}, \tag{2.3}$$

where P_n denotes the empirical measure corresponding to the sample. Let S_n be a subsample that minimizes $\det(\widehat{C}_n(S))$ over all subsamples of size $h_n \geq \lceil n\gamma \rceil$; then the pair $(\widehat{T}_n(S_n), \widehat{C}_n(S_n))$ is an MCD estimator. Note that a minimizing subsample always exists, but it need not be unique. In [3], it is shown that a minimizing subsample S_n always has exactly $\lceil n\gamma \rceil$ points and is contained in the ellipsoid $E(\widehat{T}_n(S_n), \widehat{C}_n(S_n), \widehat{r}_n(S_n))$, which separates S_n from all other points in the sample. Note that in [2] (among others) one minimizes over subsamples of size $\lfloor n\gamma \rfloor$. This is somewhat unnatural, since it may lead to subsamples S for which $P_n(S) < \gamma$. Moreover, it may lead to situations where the trimmed subsample does not contain the majority of the points; for example, if $\gamma = 1/2$ and n is odd, then $\lfloor n\gamma \rfloor = (n-1)/2$. By considering subsamples S of size $h_n \geq \lceil n\gamma \rceil$ in definition (2.1), we always have $P_n(S) \geq \gamma$, and for any $1/2 \leq \gamma \leq 1$, the subsample contains the majority of points.

We define the MCD functionals in a similar fashion. Define a trimmed mean and covariance as follows:

$$\begin{aligned} T_P(\phi) &= \frac{1}{\int \phi \, \mathrm{d}P}\int x\phi(x)P(\mathrm{d}x), \\ C_P(\phi) &= \frac{1}{\int \phi \, \mathrm{d}P}\int (x - T_P(\phi))(x - T_P(\phi))'\phi(x)P(\mathrm{d}x) \end{aligned} \tag{2.4}$$

and define

$$r_P(\phi) = \inf\{s > 0 : P\left(E(T_P(\phi), C_P(\phi), s)\right) \geq \gamma\}$$

for measurable $\phi : \mathbb{R}^k \to [0, 1]$, such that $\int \phi\,\mathrm{d}P \geq \gamma$ and $\int \|x\|^2\phi(x)P(\mathrm{d}x) < \infty$. Note that, for $P = P_n$ and $\phi = \mathbb{1}_S$, we recover (2.1) and (2.3). If ϕ_P minimizes $\det(C_P(\phi))$ over all ϕ considered above, then the pair $(T_P(\phi_P), C_P(\phi_P))$ is called an MCD functional. In [3], it is shown that such a ϕ_P always exists, and a characterization of a minimizing ϕ is provided. From this characterization (Theorem 3.2 in [3]) it follows that, if P has a density, then

$$\phi_P = \mathbb{1}_{E_P} \quad \text{and} \quad P(E_P) = \gamma, \tag{2.5}$$

where $E_P = E(T_P(\phi_P), C_P(\phi_P), r_P(\phi_P))$. This means that the MCD functional defined by (2.4) coincides with the definition through minimization over bounded Borel sets given in [2].

Throughout the paper we will assume that the MCD functional at P is uniquely defined, and we write $(\mu_0, \Sigma_0) = (T_P(\phi_P), C_P(\phi_P))$ and $\rho_0 = r_P(\phi_P)$. This holds, for instance, if P has a unimodal elliptically contoured density (see Theorem 1 in [2]). We will also assume that P has a density f that satisfies the following condition:

(B) f is continuous and strictly positive on a open neighborhood of the boundary of $E(\mu_0, \Sigma_0, \rho_0)$.

In that case, it follows from Theorem 4.2 in [3] that $\widehat{\theta}_n \to \theta_0$ with probability 1, where

$$\begin{aligned}
\widehat{\theta}_n &= \left(\widehat{T}_n(S_n), \widehat{B}_n(S_n), \widehat{r}_n(S_n)\right), \quad \text{with } \widehat{B}_n(S_n)^2 = \widehat{C}_n(S_n),\\
\theta_0 &= \left(\mu_0, \Gamma_0, \rho_0\right), \quad \text{with } \Gamma_0^2 = \Sigma_0.
\end{aligned} \tag{2.6}$$

Moreover, Theorem 5.1 in [3] implies that expansion (1.1) holds, where $\Psi = (\Psi_1, \Psi_2, \Psi_3)$ is defined as

$$\begin{aligned}
\Psi_1(y, \theta) &= \mathbb{1}_{\{\|G^{-1}(y-m)\| \leq r\}} G^{-1}(y-m)\\
\Psi_2(y, \theta) &= \mathbb{1}_{\{\|G^{-1}(y-m)\| \leq r\}} \left(G^{-1}(y-m)(y-m)'G^{-1} - I_k\right)\\
\Psi_3(y, \theta) &= \mathbb{1}_{\{\|G^{-1}(y-m)\| \leq r\}} - \gamma,
\end{aligned} \tag{2.7}$$

and $\Lambda = (\Lambda_1, \Lambda_2, \Lambda_3)$, with

$$\Lambda_j(\theta) = \int \Psi_j(y, \theta)P(\mathrm{d}y), \quad \text{for } j = 1, 2, 3, \tag{2.8}$$

for $\theta = (m, G, r)$, with $y, t \in \mathbb{R}^k$, $r > 0$ and $G \in \mathrm{PDS}(k)$. Here, $\mathrm{PDS}(k)$ denotes the space of all positive definite symmetric $k \times k$ matrices.

3. Existence and non-singularity of $\Lambda'(\theta_0)$

Let $\theta_0 = (\mu_0, \Gamma_0, \rho_0)$ be the MCD functional at P. Due to the indicator function in the expression of $\Psi(x, \theta_0)$, it can be seen that the existence of a derivative of $\Lambda(\theta)$ at θ_0 cannot be expected in general if P does not satisfy condition (B). If P does satisfy (B), then the derivative will depend on the behavior of f on the boundary of $E(\mu_0, \Sigma_0, \rho_0)$. For $\rho > 0$ and $\mu \in \mathbb{R}^k$, define

$$B(\mu, \rho) = \left\{x \in \mathbb{R}^k : \|x - \mu\| \leq \rho\right\}.$$

The derivative of $\Lambda(\theta)$ at θ_0 turns out to be an integral over the boundary of $B(0, \rho_0)$. In order to keep things tidy in describing the derivative, we denote σ_0 as the Lebesgue surface measure on $\partial B(0, \rho_0)$ and introduce the measure

$$\nu(\mathrm{d}\omega) = \det(\Gamma_0)f(\Gamma_0\omega + \mu_0)\sigma_0(\mathrm{d}\omega) \quad \text{for } \omega \in \partial B(0, \rho_0). \tag{3.1}$$

This can be interpreted as the image measure of P restricted to $\partial E(\mu_0, \Sigma_0, \rho_0)$, after the affine map that transforms the ellipsoid to the ball around 0 with radius ρ_0. Note that our parameter θ_0 is an element of $\Theta = \mathbb{R}^k \times \mathrm{PDS}(k) \times \mathbb{R}$. This means that the derivative of Λ at θ_0, if it exists, can be described as a linear mapping on the tangent space of Θ in θ_0, which is given by $\mathbb{R}^k \times S(k) \times \mathbb{R}$. Here, $S(k)$ denotes the space of all symmetric $k \times k$ matrices. The derivatives of Λ_1, Λ_2 and Λ_3 are given as linear mappings by the following theorem.

Theorem 3.1. *Suppose that P satisfies* (B) *and let the MCD functional $\theta_0 = (\mu_0, \Gamma_0, \rho_0)$ be uniquely defined at P. For $j = 1, 2, 3$, the derivatives of Λ_j are given by the following linear mappings, with $(h, A, s) \in \mathbb{R}^k \times S(k) \times \mathbb{R}$:*

$$\begin{aligned}
\Lambda_1'(\theta_0)(h, A, s) &= -\gamma\Gamma_0^{-1}h + \int_{\partial B_0}\left(\frac{\omega'\Gamma_0^{-1}h}{\rho_0} + \frac{\omega'(\Gamma_0^{-1}A + A\Gamma_0^{-1})\omega}{2\rho_0} + s\right)\omega\,\nu(\mathrm{d}\omega)\\
\Lambda_2'(\theta_0)(h, A, s) &= -\gamma(\Gamma_0^{-1}A + A\Gamma_0^{-1}) + \int_{\partial B_0}\left(\frac{\omega'\Gamma_0^{-1}h}{\rho_0} + \frac{\omega'(\Gamma_0^{-1}A + A\Gamma_0^{-1})\omega}{2\rho_0} + s\right)\left(\omega\omega' - I\right)\nu(\mathrm{d}\omega)\\
\Lambda_3'(\theta_0)(h, A, s) &= \int_{\partial B_0}\left(\frac{\omega'\Gamma_0^{-1}h}{\rho_0} + \frac{\omega'(\Gamma_0^{-1}A + A\Gamma_0^{-1})\omega}{2\rho_0} + s\right)\nu(\mathrm{d}\omega),
\end{aligned}$$

where $B_0 = B(0, \rho_0)$ and $\nu(\mathrm{d}\omega)$ is defined in (3.1).

Note that Theorem 3.1 also has practical implications. According to Theorem 5.1 in [3], the MCD estimator $\widehat{\theta}_n = \big(\widehat{T}_n(S_n), \widehat{B}_n(S_n), \widehat{r}_n(S_n)\big)$, represented as a vector, is asymptotically normal, with mean zero and limiting variance given by the covariance matrix of $Z = \Lambda'(\theta_0)^{-1}\Psi(X_1, \theta_0)$. This means that the expression for $\Lambda'(\theta_0)$ enables one to estimate the limiting variance of the MCD estimators in any model where P has a density, which goes far beyond the traditional elliptically contoured densities. In the expressions of Theorem 3.1, the parameters μ_0, Γ_0 and ρ_0 can be estimated by the MCD estimators $\widehat{T}_n(S_n)$, $\widehat{B}_n(S_n)$ and $\widehat{r}_n(S_n)$ from (2.6). To estimate the density f on the boundary of $B(0, \widehat{\rho}_n)$, one can use a nonparametric estimate $\widehat{f}$, e.g., a histogram or kernel type estimate. Finally, the surface measure σ_0 on $\partial B(0, \rho_0)$ can be estimated by the surface measure $\widehat{\sigma}_n$ on $\partial B(0, \widehat{r}_n(S_n))$ and the measure $\nu(\mathrm{d}\omega)$ by

$$\widehat{\nu}(\mathrm{d}\omega) = \det(\widehat{B}_n(S_n))\widehat{f}\left(\widehat{B}_n(S_n)\omega + \widehat{T}_n(S_n)\right)\widehat{\sigma}_n(\mathrm{d}\omega), \quad \omega \in \partial B(0, \widehat{r}_n(S_n)).$$

The integrals over $\partial B(0, \widehat{r}_n(S_n))$ with respect to $\widehat{\nu}(\mathrm{d}\omega)$ can be approximated numerically by means of Riemann sums. It follows that the expressions in Theorem 3.1, with the parameters replaced by their estimates as just described, provide an estimate $\widehat{\Lambda}'(\widehat{\theta}_n)$ for the derivative as a linear mapping of (h, A, s). Being a linear mapping, $\widehat{\Lambda}'(\widehat{\theta}_n)$ can be represented by a matrix. The columns of this matrix can be determined by inserting first of all $(h, A, s) = (e_i, 0, 0)$, where the vector e_i has all elements equal to zero except for a 1 at position i, then inserting $(h, A, s) = (0, E_{ij}, 0)$, where the symmetric matrix E_{ij} has all entries equal to zero except for a 1 at positions (i, j) and (j, i), and finally inserting $(h, A, s) = (0, 0, 1)$. These vectors form a canonical orthogonal basis for $\mathbb{R}^k \times S(k) \times \mathbb{R}$. As long as $\widehat{\Lambda}'(\widehat{\theta}_n)$ turns out to be non-singular, the limiting covariance matrix of $\sqrt{n}(\widehat{\theta}_n - \theta_0)$ can be estimated by the sample covariance of the $Z_i = \widehat{\Lambda}'(\widehat{\theta}_n)^{-1}\Psi(X_i, \widehat{\theta}_n)$.

We proceed by finding sufficient conditions for $\Lambda'(\theta_0)$ to be non-singular. We would have non-singularity if, for all $(h, A, s) \in \mathbb{R}^k \times S(k) \times \mathbb{R}$, $\Lambda'(\theta_0)(h, A, s) = 0$ implies that $(h, A, s) = (0, 0, 0)$. Although it does imply that $\mathrm{Tr}(\Gamma_0^{-1}A) = 0$ (see Lemma A.2), from the expressions in Theorem 3.1 it can be seen that it cannot be expected that $\Lambda'(\theta_0)(h, A, s) = 0$ implies that $(h, A, s) = (0, 0, 0)$ without further assumptions on f. Suitable symmetry assumptions on f will simplify the expressions for the derivative, in which case non-singularity can be established. Point symmetry with respect to the center of $E(\mu_0, \Sigma_0, \rho_0)$, i.e.,

$$f(-\Gamma_0\omega + \mu_0) = f(\Gamma_0\omega + \mu_0), \quad \text{for } \omega \in \partial B(0, \rho_0), \tag{3.2}$$

allows us to express s in terms of A and if, for all $i = 1, 2 \dots, k$,

$$\int_{\partial B_0} \omega_i^2\, \nu(\mathrm{d}\omega) \neq \gamma\rho_0, \tag{3.3}$$

then $\Lambda'(\theta_0)(h, A, s) = 0$ implies that $h = 0$ (see Lemma A.3), but this will not be sufficient to conclude that $A = 0$ from $\Lambda'(\theta_0)(h, A, s) = 0$. The slightly stronger condition of half-space symmetry will suffice, i.e.,

$$f(\Gamma_0\omega_{(-i)} + \mu_0) = f(\Gamma_0\omega + \mu_0), \quad \text{where } \omega_{(-i)} = (\omega_1, \dots, \omega_{i-1}, -\omega_i, \omega_{i+1}, \dots, \omega_k), \tag{3.4}$$

for all $i = 1, 2, \dots, k$ and $\omega \in \partial B(0, \rho_0)$. To describe sufficient conditions for non-singularity, we define the matrix M with elements

$$M_{ij} = \int_{\partial B(0,\rho_0)} \omega_i^2\omega_j^2\, \nu(\mathrm{d}\omega) - \frac{1}{\nu_0}\int_{\partial B(0,\rho_0)} \omega_i^2\, \nu(\mathrm{d}\omega)\int_{\partial B(0,\rho_0)} \omega_j^2\, \nu(\mathrm{d}\omega) - 2\gamma\rho_0\mathbb{1}_{\{i=j\}}, \tag{3.5}$$

for $i, j = 1, 2, \dots, k$, where $\nu_0 = \nu(\partial B(0, \rho_0))$ and $\nu(\mathrm{d}\omega)$ is defined by (3.1). We then have the following theorem.

Theorem 3.2. *Suppose that P satisfies* (B) *and* (3.4)*, and let the MCD functional $\theta_0 = (\mu_0, \Gamma_0, \rho_0)$ be uniquely defined at P. Suppose that* (3.3) *holds, that for all $i, j = 1, 2 \dots, k$ with $i \neq j$,*

$$\int_{\partial B(0,\rho_0)} \omega_i^2\omega_j^2\, \nu(\mathrm{d}\omega) \neq \gamma\rho_0, \tag{3.6}$$

where ν is defined in (3.1)*, and that the matrix M defined in* (3.5) *is such that, for any $x \in \mathbb{R}^k$,*

$$Mx = 0 \text{ and } x_1 + \dots + x_k = 0 \Rightarrow x = 0. \tag{3.7}$$

Then, for $\theta_0 = (\mu_0, \Gamma_0, \rho_0)$, the derivative $\Lambda'(\theta_0)$ is non-singular as a linear map on $\mathbb{R}^k \times S(k) \times \mathbb{R}$.

Example of densities that satisfy (3.4) are elliptically contoured densities. However, also affine transformations of densities that have independent marginal densities that are symmetric around zero, i.e.,

$$f(x) = g(\Gamma^{-1}(x - \mu)), \quad \text{where } g(x_1, \dots, x_k) = g_1(x_1)\cdots g_k(x_k) \text{ and } g_i(x_i) = g_i(-x_i),$$

satisfy (3.4).

4. Elliptically contoured densities

Suppose that P has an elliptically contoured density, i.e.,

$$f(x)=\det(\Sigma)^{-1/2}h((x-\mu)\Sigma^{-1}(x-\mu)) \quad \text{where } \mu\in\mathbb{R}^k,\ \Sigma\in \mathrm{PDS}(k), \tag{4.1}$$

and $h:[0,\infty)\to[0,\infty)$ is decreasing so that P is unimodal. In this case, it follows from the characterization for the ϕ function that minimizes $\det(C_P(\phi))$ (see Theorem 3.2 in [3]) that our definition of the MCD functional coincides with the one used by Butler et al. [2], who show that the MCD functionals are unique:

$$\mu_0=\mu,\qquad \Sigma_0=\alpha(\gamma)^2\Sigma,\quad \text{and}\quad \rho_0^2=\frac{r(\gamma)^2}{\alpha(\gamma)^2}, \tag{4.2}$$

where

$$\alpha(\gamma)^2=\frac{2\pi^{k/2}}{\gamma k\Gamma(k/2)}\int_0^{r(\gamma)}h(r^2)r^{k+1}\,\mathrm{d}r, \tag{4.3}$$

and where $r(\gamma)$ is determined by

$$\frac{2\pi^{k/2}}{\Gamma(k/2)}\int_0^{r(\gamma)}h(r^2)r^{k-1}\,\mathrm{d}r=\gamma. \tag{4.4}$$

The next proposition shows that for elliptically contoured densities the derivative $\Lambda'(\theta_0)$ exists and is non-singular.

Proposition 4.1. *Let P have an elliptically contoured density as defined in (4.1) with h non-increasing such that P is unimodal. Then all conditions of Theorem 3.2 are satisfied.*

We proceed by obtaining asymptotic expansions for the MCD estimators in the case of elliptically contoured densities. Because the estimators are affine equivariant, it suffices to consider the spherically symmetric case $(\mu,\Sigma)=(0,I)$. The next theorem provides the expressions for $\Lambda'(\theta_0)$ and its inverse at spherically symmetric densities.

Theorem 4.1. *Let P have an spherically symmetric density $f(x)=h(\|x\|^2)$ with h decreasing such that P is unimodal. Let $r=r(\gamma)$ and $\alpha=\alpha(\gamma)$ be defined in (4.4) and (4.3), respectively, and let $D=\Lambda'(\theta_0)$, for $\theta_0=(\mu_0,\Gamma_0,\rho_0)=(0,\alpha I,r/\alpha)$. Then the linear mapping D is given by*

$$D_1:(h,A,s)\mapsto\beta_1 h$$
$$D_2:(h,A,s)\mapsto\beta_2 A+\beta_3\,\mathrm{Tr}(A)\cdot I+\beta_4 s\cdot I$$
$$D_3:(h,A,s)\mapsto\beta_5\,\mathrm{Tr}(A)+\beta_6 s,$$

and the inverse linear mapping D^{inv} is given by

$$\left[D^{\mathrm{inv}}\right]_1:(g,B,t)\mapsto\beta_1^{-1}g$$
$$\left[D^{\mathrm{inv}}\right]_2:(g,B,t)\mapsto\beta_2^{-1}B+\frac{\alpha(\beta_3\beta_6-\beta_4\beta_5)}{2\gamma\beta_2\beta_6}\mathrm{Tr}(B)\cdot I+\frac{\alpha\beta_4}{2\gamma\beta_6}t\cdot I$$
$$\left[D^{\mathrm{inv}}\right]_3:(g,B,t)\mapsto\frac{\alpha\beta_5}{2\gamma\beta_6}\mathrm{Tr}(B)-\frac{\alpha(\beta_2+k\beta_3)}{2\gamma\beta_6}t,$$

where

$$\beta_1=\frac{1}{\alpha}\left(\frac{\rho_0}{k}\nu_0-\gamma\right)<0,\qquad \beta_4=\frac{\rho_0^2}{k}\nu_0-\nu_0,$$
$$\beta_2=\frac{2\rho_0^3\nu_0}{\alpha k(k+2)}-\frac{2\gamma}{\alpha}<0,\qquad \beta_5=\frac{\rho_0\nu_0}{k\alpha},$$
$$\beta_3=\frac{\rho_0^3\nu_0}{\alpha k(k+2)}-\frac{\rho_0\nu_0}{k\alpha},\qquad \beta_6=\nu_0>0,$$

with $B_0=B(0,\rho_0)$ and

$$\nu_0=\nu(\partial B_0)=\frac{2\pi^{k/2}}{\Gamma(k/2)}h(r^2)r^{k-1}\alpha.$$

An immediate consequence of Theorem 4.1 is the next corollary, which shows that the MCD estimators of location and covariance are asymptotically equivalent to a sum of independent identically distributed vector and matrix valued random elements, respectively.

Corollary 4.1. *Suppose that P has a spherically symmetric density $f(x) = h(\|x\|^2)$ with h decreasing such that P is unimodal. Let $r = r(\gamma)$ and $\alpha = \alpha(\gamma)$ be as defined in (4.4) and (4.3), respectively. Then, for $n \to \infty$,*

$$\sqrt{n}\widehat{\mu}_n = \frac{1}{\sqrt{n}}\sum_{i=1}^{n} \tau \mathbb{1}_{\{\|X_i\|\le r\}} X_i + o_{\mathbb{P}}(1);$$

$$\sqrt{n}(\widehat{\Sigma}_n - \alpha^2 I) = \frac{1}{\sqrt{n}}\sum_{i=1}^{n}\left[\mathbb{1}_{\{\|X_i\|\le r\}}\left(\kappa_1 \cdot I + \kappa_2\|X_i\|^2 \cdot I + \kappa_3 X_i X_i'\right) + \kappa_4 \cdot I\right] + o_{\mathbb{P}}(1);$$

$$\sqrt{n}\left(\widehat{\rho}_n - \frac{r}{\alpha}\right) = \frac{1}{\sqrt{n}}\sum_{i=1}^{n}\left[\lambda_1 \mathbb{1}_{\{\|X_i\|\le r\}}\|X_i\|^2 + \lambda_2 \mathbb{1}_{\{\|X_i\|\le r\}} + \lambda_3\right] + o_{\mathbb{P}}(1),$$

where $\tau = -(\alpha\beta_1)^{-1}$ and

$$\kappa_1 = -\frac{r^2}{k\gamma}, \qquad \kappa_2 = \frac{\alpha\beta_2 + 2\gamma}{k\gamma\alpha\beta_2}, \qquad \kappa_3 = -\frac{2}{\alpha\beta_2}, \qquad \kappa_4 = \frac{r^2 - k\alpha^2}{k}$$

$$\lambda_1 = -\frac{r}{2k\gamma\alpha^3}, \qquad \lambda_2 = \frac{r^3}{2k\gamma\alpha^3} - \frac{1}{\beta_6}, \qquad \lambda_3 = \frac{\gamma}{\beta_6} + \frac{r}{2k\alpha^3}\left(k\alpha^2 - r^2\right),$$

with β_1, β_2 and β_6 defined in Theorem 4.1.

We proceed by obtaining the limit distribution of the MCD estimators. To describe the limiting distribution of a random matrix, we use the operator vec(·), which stacks the columns of a matrix M on top of each other, i.e.,

$$\mathrm{vec}(M) = (M_{11}, \ldots, M_{k1}, \ldots, M_{1k}, \ldots, M_{kk})'.$$

We will also need the commutation matrix $C_{k,k}$, which is a $k^2 \times k^2$ matrix consisting of $k \times k$ blocks: $C_{k,k} = (\Delta_{ij})_{i,j=1}^{k}$, where each (i, j)-th block is equal to a $k \times k$-matrix Δ_{ji}, which is 1 at entry (j, i) and zero everywhere else. Finally, for matrices M and N, the Kronecker product $M \otimes N$ is a $k^2 \times k^2$ matrix consisting of $k \times k$ blocks, with the (i, j)-th block equal to $m_{ij}N$.

Theorem 4.2. *Suppose that P has a spherically symmetric density $f(x) = h(\|x\|^2)$ with h decreasing such that P is unimodal. Let $r = r(\gamma)$ and $\alpha = \alpha(\gamma)$ be as defined in (4.4) and (4.3), respectively. Let $\widehat{\mu}_n$, $\widehat{\Sigma}_n$ and $\widehat{\rho}_n$ be the MCD estimators. Then*

(i) *$\widehat{\mu}_n$ and $(\widehat{\Sigma}_n, \widehat{\rho}_n)$ are asymptotically independent, the diagonal elements of $\widehat{\Sigma}_n$ are asymptotically independent from the off-diagonal elements and $\widehat{\rho}_n$, and the off-diagonal elements of $\widehat{\Sigma}_n$ are asymptotically mutually independent;*

(ii) *$\sqrt{n}\widehat{\mu}_n$ is asymptotically normal with mean zero and covariance matrix ξI, where*

$$\xi = \frac{k^2\gamma\alpha^4}{(k\gamma\alpha - r\nu_0)^2},$$

where ν_0 is defined in Theorem 4.1;

(iii) *$\sqrt{n}(\mathrm{vec}(\widehat{\Sigma}_n) - \alpha^2\,\mathrm{vec}(I))$ is asymptotically normal with mean zero and covariance matrix*

$$\sigma_1(I + C_{k,k})(I \otimes I) + \sigma_2\,\mathrm{vec}(I)\mathrm{vec}(I)',$$

where

$$\sigma_1 = \frac{\kappa_3^2}{k(k+2)}\mathbb{E}\mathbb{1}_{\{\|X_1\|\le r\}}\|X_1\|^4$$

$$\sigma_2 = -\frac{2}{k}\sigma_1 + \frac{1}{k^2\gamma^2}\mathbb{E}\mathbb{1}_{\{\|X_1\|\le r\}}\|X_1\|^4 - \frac{\gamma r^4 - 2k\gamma r^2\alpha^2 + k^2\gamma\alpha^4 + 2kr^2\alpha^2 - r^4}{\gamma k^2}$$

where κ_3 is defined in Corollary 4.3;

(iv) *$\sqrt{n}(\widehat{\rho}_n - r/\alpha)$ is asymptotically normal with mean zero and variance*

$$\sigma_\rho^2 = \lambda_1^2\mathbb{E}\mathbb{1}_{\{\|X_1\|\le r\}}\|X_1\|^4 + \frac{2k^2\nu_0 r\alpha^5\gamma - \nu_0^2 kr^4\alpha^2 + 4k^2\gamma^2\alpha^6 - 4kr^3\nu_0\alpha^3\gamma + r^6\nu_0^2}{4k^2\alpha^6\nu_0^2\gamma},$$

where λ_1 is defined in Corollary 4.1 and ν_0 is defined in Theorem 4.1.

Note that, at the multivariate normal, the asymptotic relative efficiency of the location MCD estimator equals $\tau = 1/(\gamma\alpha^2)$, which can be seen to be greater than 1 and tending to 1 as γ tends to 1. For values of asymptotic efficiencies at specific distributions, we refer to [4], who provide an extensive account of asymptotic and finite sample relative efficiencies for the MCD covariance estimator at the multivariate standard normal, a contaminated multivariate normal and at several multivariate Student distributions, for a variety of dimensions $k = 2, 3, 5, 10, 30$ and $\gamma = 0.5, 0.75$, as well as a comparison with S-estimators and reweighted versions. One should note however that they compute efficiencies for a Fisher consistent version of the MCD covariance estimator, i.e., $\widehat{\Sigma}_n/\alpha(\gamma)^2$. Taking this into account, our expressions coincide with the ones in [4]. This follows from the fact that the expressions in Theorem 4.2 are derived from the expansion given in Corollary 4.1, of which the right-hand side coincides with the expressions for the influence function given in Corollary 4.3. Our expression for the influence function of the covariance functional is identical to the one in [4] (see the comments right after Corollary 4.3).

With Theorem 4.2(i) we recover Theorem 4 in [2]. Note however, that the assumption of h being differentiable (see [2]) is not required in our approach. Furthermore, it can be seen from the expression of the limiting variance of $\widehat{\Sigma}_n$ that, in the spherically symmetric case,

$$\sqrt{n}(\widehat{\Sigma}_{n,ii}-\alpha^2)\to N(0,2\sigma_1+\sigma_2)$$
$$\sqrt{n}\widehat{\Sigma}_{n,ij}\to N(0,\sigma_1)$$
$$\sqrt{n}\begin{pmatrix}\widehat{\Sigma}_{n,ii}-\alpha^2\\ \widehat{\Sigma}_{n,jj}-\alpha^2\end{pmatrix}\to N\left(\begin{pmatrix}0\\0\end{pmatrix},\begin{pmatrix}2\sigma_1+\sigma_2 & \sigma_2\\ \sigma_2 & 2\sigma_1+\sigma_2\end{pmatrix}\right),\quad i\neq j,$$

for $i,j=1,2,\ldots,k$.

Because $\widehat{\mu}_n$ and $\widehat{\Sigma}_n$ are affine equivariant, the limiting distributions for the MCD estimators in the case of general $\mu\in\mathbb{R}^k$ and $\Sigma\in \mathrm{PDS}(k)$ can be obtained easily. When $X_1,\ldots,X_n$ are independent with density (4.1), then, because of affine equivariance, it follows immediately that $\sqrt{n}(\widehat{\mu}_n-\mu)$ is asymptotically normal with zero mean and covariance matrix $\Gamma(\tau I)\Gamma=\tau\Sigma$, where $\Gamma^2=\Sigma$. Similarly, $\sqrt{n}(\mathrm{vec}(\widehat{\Sigma}_n)-\alpha^2\,\mathrm{vec}(\Sigma))$ is asymptotically normal with mean zero and covariance matrix $\mathbb{E}\,\mathrm{vec}(\Gamma M\Gamma)\mathrm{vec}(\Gamma M\Gamma)'$, where the covariance matrix of $\mathrm{vec}(M)$ is given in Theorem 4.2(iii). The expression of $\mathbb{E}\,\mathrm{vec}(\Gamma M\Gamma)\mathrm{vec}(\Gamma M\Gamma)'$ follows from Lemma 5.2 in [9]. This means that we have the following general corollary of Theorem 4.2.

Corollary 4.2. *Suppose that $X_1,\ldots,X_n$ are independent with an elliptical contoured density*

$$f(x)=\det(\Sigma)^{-1/2}h((x-\mu)'\Sigma^{-1/2}(x-\mu)),\quad \mu\in\mathbb{R}^k,\ \Sigma\in\mathrm{PDS}(k),$$

where $h:[0,\infty)\to[0,\infty)$ is non-increasing such that f is unimodal. Let $(\widehat{\mu}_n,\widehat{\Sigma}_n)$ be the MCD estimators. Then $\widehat{\mu}_n$ and $\widehat{\Sigma}_n$ are asymptotically independent, $\sqrt{n}(\widehat{\mu}_n-\mu)$ has a limiting normal distribution with zero mean and covariance matrix $\tau\Sigma$ and $\sqrt{n}(\mathrm{vec}(\widehat{\Sigma}_n)-\alpha^2\,\mathrm{vec}(\Sigma))$ has a limiting normal distribution with zero mean and covariance matrix $\sigma_1(I+C_{k,k})(\Sigma\otimes\Sigma)+\sigma_2\,\mathrm{vec}(\Sigma)\mathrm{vec}(\Sigma)'$, where τ, σ_1 and σ_2 are given in Theorem 4.2.

Another corollary of Theorem 4.1 is the expression for the influence function of the MCD functional. The influence function of a functional $\Theta(\cdot)$ at P is defined as

$$\mathrm{IF}(x,\Theta,P)=\lim_{\varepsilon\downarrow 0}\frac{\Theta((1-\varepsilon)P+\varepsilon\delta_x)-\Theta(P)}{\varepsilon},\tag{4.5}$$

if this limit exists, where δ_x is the Dirac measure at $x\in\mathbb{R}^k$. Denote by $\mu(P)=T_P(\phi_P)$, $\Sigma(P)=C_P(\phi_P)$, and $\rho(P)=r_P(\phi_P)$ the MCD functionals at distribution P. We then have the following corollary.

Corollary 4.3. *Suppose that P has a spherically symmetric density $f(x)=h(\|x\|^2)$ with h decreasing such that P is unimodal. Let $r=r(\gamma)$ and $\alpha=\alpha(\gamma)$ be as defined in (4.4) and (4.3), respectively. Then, for $x\in\mathbb{R}^k$ such that $\|x\|\neq r$, the influence functions of the functionals $\mu(P)$, $\Sigma(P)$ and $\rho(P)$ are given by*

$$\mathrm{IF}(x;\mu,P)=\tau\mathbb{1}_{\{\|x\|\le r\}}x$$
$$\mathrm{IF}(x;\Sigma,P)=\mathbb{1}_{\{\|x\|\le r\}}\left(\kappa_1\cdot I+\kappa_2\|x\|^2\cdot I+\kappa_3 xx'\right)+\kappa_4\cdot I$$
$$\mathrm{IF}(x;\rho,P)=\lambda_1\mathbb{1}_{\{\|x\|\le r\}}\|x\|^2+\lambda_2\mathbb{1}_{\{\|x\|\le r\}}+\lambda_3,$$

where $\tau,\kappa_1,\kappa_2,\kappa_3,\kappa_4$ and $\lambda_1,\lambda_2,\lambda_3$ are defined in Corollary 4.1.

Clearly, all the expressions in Corollary 4.3 are bounded uniformly for $\|x\|\neq r(\gamma)$. For $x\in\mathbb{R}^k$ with $\|x\|=r(\gamma)$, it is not clear whether the limit in (4.5) exists, not even in the case of a unimodal spherically symmetric density. As a special case of Corollary 4.3, we recover Theorem 1 in [4]. However, we do not need the assumption that h is differentiable (see [4]). In order to see that our expressions coincide with the ones in [4], note that their quantities g,α,q_α and c_α correspond to our h, $1-\gamma$, $r(\gamma)^2$ and $1/\alpha(\gamma)^2$, respectively, and that they consider the Fisher consistent version of the covariance functional, i.e., $c_\alpha\times\Sigma(P)$. Moreover, their expression b_1-kb_2 is simply equal to 1. For further discussion on $\mathrm{IF}(x;\Sigma,P)$ at spherically symmetric densities and corresponding graphs, we refer to [4].

Appendix

A.1. Proofs for Section 3

The proof of Theorem 3.1 requires the following lemma, which helps to describe the derivative of Λ when $(\mu_0,\Gamma_0,s)=(0,I,r)$, in terms of a linear mapping. Let $M(k)$ be the space of all $k\times k$ matrices.

Lemma A.1. *Let $r > 0$ and $\phi : \mathbb{R}^k \to \mathbb{R}^m$, which is continuous on $\partial B(0, r)$. Define the mapping $L : \mathbb{R}^k \times M(k) \times \mathbb{R} \to \mathbb{R}^m$ by*

$$L(h, A, s) = \int_{E(h,(I+A)(I+A)',r+s)} \phi(y)\, \mathrm{d}y.$$

Then, the derivative of L at $(h_0, A_0, s_0) = (0, 0, 0)$ is given by the continuous linear mapping

$$L'(0, 0, 0)(h, A, s) = \int_{\partial B(0,r)} \left(\frac{\omega' h}{r} + \frac{\omega'(A + A')\omega}{2r} + s \right) \phi(\omega)\, \sigma_0(\mathrm{d}\omega),$$

with $(h, A, s) \in \mathbb{R}^k \times M(k) \times \mathbb{R}$.

Proof. The derivative can be found as the sum of the derivatives of

$$\int_{B(h,r)} \phi(y)\, \mathrm{d}y, \quad \int_{E(0,(I+A)(I+A)',r)} \phi(y)\, \mathrm{d}y, \quad \text{and} \quad \int_{B(0,r+s)} \phi(y)\, \mathrm{d}y. \tag{A.1}$$

For the first integral, consider

$$\int_{B(h,r)} \phi(y)\, \mathrm{d}y - \int_{B(0,r)} \phi(y)\, \mathrm{d}y = \int \left(\mathbb{1}_{B(h,r)} - \mathbb{1}_{B(0,r)} \right) \phi(y)\, \mathrm{d}y,$$

for $\|h\| \to 0$. In first order this reduces to integration over $\partial B(0, r)$. Let $\omega \in \partial B(0, r)$, let $v = (1 + \delta)\omega \in \partial B(h, r)$, and let α denote the angle between ω and h. Then the law of cosines yields that

$$r^2 = \|v\|^2 + \|h\|^2 - 2\|v\| \cdot \|h\| \cos\alpha = \left(\|v\| - \frac{\omega' h}{\|\omega\|} \right)^2 + \|h\|^2 (\sin\alpha)^2.$$

Since $\|\omega\| = r$, in first order we find that $r^2 = (1 + \delta)r^2 - \omega' h$, or $\delta = (\omega' h)/r^2$. This means that, for each $\omega \in \partial B(0, r)$, the length over which we integrate $\phi(\omega)$ is $\|v\| - \|\omega\| = \delta\|\omega\| = (\omega' h)/r$. Since ϕ is continuous at $\omega \in \partial B(0, r)$, we get, for $\|h\| \to 0$,

$$\int_{B(h,r)} \phi(y)\, \mathrm{d}y - \int_{B(0,r)} \phi(y)\, \mathrm{d}y = \int_{\partial B(0,r)} \frac{\omega' h}{r} \phi(\omega)\, \sigma_0(\mathrm{d}\omega) + o(\|h\|).$$

For the second integral in (A.1), we consider

$$\int \left(\mathbb{1}_{E(0,(I+A)(I+A)',r)} - \mathbb{1}_{B(0,r)} \right) \phi(y)\, \mathrm{d}y,$$

for $\|A\| \to 0$, which reduces to integration over $\omega \in \partial B(0, r)$. Let $v = (1 + \delta)\omega$ be such that $\|(I + A)^{-1} v\| = r$. Then

$$(1 + \delta)^2 = \frac{r^2}{\omega'(I + A')^{-1}(I + A)^{-1}\omega}.$$

Since, for $\|A\| \to 0$, we have $(I + A')^{-1}(I + A)^{-1} = I - A - A' + O(\|A\|^2)$, it follows that

$$\delta = \frac{\omega'(A + A')\omega}{2r^2} + O(\|A\|^2).$$

This means that, for each $\omega \in \partial B(0, r)$, the length over which we integrate $\phi(\omega)$ is

$$\|v\| - \|\omega\| = \delta\|\omega\| = \frac{\omega'(A + A')\omega}{2r} + O(\|A\|^2).$$

This implies that

$$\int \left(\mathbb{1}_{E(0,(I+A)(I+A)',r)} - \mathbb{1}_{B(0,r)} \right) \phi(y)\, \mathrm{d}y = \int_{\partial B(0,r)} \frac{\omega'(A + A')\omega}{2r} \phi(\omega)\sigma_0(\mathrm{d}\omega) + o(\|A\|).$$

Finally, for the third integral in (A.1) we obtain

$$\int \left(\mathbb{1}_{B(0,r+s)} - \mathbb{1}_{B(0,r)} \right) \phi(y)\mathrm{d}y = s \int_{\partial B(0,r)} \phi(\omega)\sigma_0(\mathrm{d}\omega) + o(s).$$

Summing the three linear mappings yields the desired result. □

Proof of Theorem 3.1. First note that everything can be rescaled to the situation with $\mu_0 = 0$ and $\Gamma_0 = I$; i.e., for any function $g(y)$, we have

$$\int \mathbb{1}_{E(\mu_0+h,(\Gamma_0+A)(\Gamma_0+A)',\rho_0+s)}(y)g(y)P(\mathrm{d}y) = \det(\Gamma_0)\int \mathbb{1}_{E(\widetilde{h},(I+\widetilde{A})^2,\rho_0+s)}(z)g(\Gamma_0 z+\mu_0)f(\Gamma_0 z+\mu_0)\,\mathrm{d}z, \tag{A.2}$$

where $\widetilde{h} = \Gamma_0^{-1}h$ and $\widetilde{A} = \Gamma_0^{-1}A$. To compute $\Lambda_3'(\theta_0)$, take $g(y) = 1$ in (A.2), and for $\eta = (h, A, s) \to (0, 0, 0)$, consider

$$\begin{aligned}\Lambda_3(\theta_0+\eta) - \Lambda_3(\theta_0) &= \int \mathbb{1}_{E(\mu_0+h,(\Gamma_0+A)(\Gamma_0+A)',\rho_0+s)}(y)P(\mathrm{d}y) - \int \mathbb{1}_{E(\mu_0,\Sigma_0,\rho_0)}(y)P(\mathrm{d}y)\\ &= \det(\Gamma_0)\int \left(\mathbb{1}_{E(\widetilde{h},(I+\widetilde{A})(I+\widetilde{A})',\rho_0+s)}(z) - \mathbb{1}_{E(0,I,\rho_0)}(z)\right)f(\Gamma_0 z+\mu_0)\,\mathrm{d}z\\ &= L'(0,0,0)(\widetilde{h},\widetilde{A},s) + o(\|(h,A,s)\|),\end{aligned}$$

by taking $\phi(z) = \det(\Gamma_0)f(\Gamma_0 z+\mu_0)$ in Lemma A.1. We conclude that

$$\Lambda_3'(\theta_0) = \det(\Gamma_0)\int_{\partial B_0}\left(\frac{\omega'\Gamma_0^{-1}h}{\rho_0} + \frac{\omega'(\Gamma_0^{-1}A + A'\Gamma_0^{-1})\omega}{2\rho_0} + s\right)f(\Gamma_0\omega+\mu_0)\sigma_0(\mathrm{d}\omega).$$

For the location functional, with $\theta = (m, G, r)$, we have

$$\Lambda_1'(\theta_0) = \frac{\partial}{\partial\theta}\left(\int_{E(\mu_0,\Sigma_0,\rho_0)} G^{-1}(y-m)P(\mathrm{d}y)\right)\bigg|_{\theta=\theta_0} + \frac{\partial}{\partial\theta}\left(\int_{E(m,G^2,r)} \Gamma_0^{-1}(y-\mu_0)P(\mathrm{d}y)\right)\bigg|_{\theta=\theta_0}. \tag{A.3}$$

The first term on the right-hand side of (A.3) can be decomposed as

$$\frac{\partial}{\partial\theta}\left(\int_{E_0} G^{-1}(y-\mu_0)P(\mathrm{d}y)\right)\bigg|_{\theta=\theta_0} + \frac{\partial}{\partial\theta}\left(\int_{E_0} \Gamma_0^{-1}(y-m)P(\mathrm{d}y)\right)\bigg|_{\theta=\theta_0}, \tag{A.4}$$

where $E_0 = E(\mu_0, \Sigma_0, \rho_0)$. Because of (2.4) and (2.5), it follows that

$$P(E(\mu_0,\Sigma_0,\rho_0)) = \gamma \quad\text{and}\quad \int_{E(\mu_0,\Sigma_0,\rho_0)} (y-\mu_0)P(\mathrm{d}y) = 0, \tag{A.5}$$

so the first derivative in (A.4) is equal to zero. To determine the second derivative in (A.4), write

$$\int_{E(\mu_0,\Sigma_0,\rho_0)} \Gamma_0^{-1}(y-\mu_0-h)P(\mathrm{d}y) - \int_{E(\mu_0,\Sigma_0,\rho_0)} \Gamma_0^{-1}(y-\mu_0)P(\mathrm{d}y) = -\gamma\Gamma_0^{-1}h,$$

which yields

$$\frac{\partial}{\partial\theta}\left(\int_{E(\mu_0,\Sigma_0,\rho_0)} \Gamma_0^{-1}(y-m)P(\mathrm{d}y)\right)\bigg|_{\theta=\theta_0} = -\gamma\Gamma_0^{-1}h.$$

For the second term on the right-hand side of (A.3), for $(h, A, s) \to (0, 0, 0)$, consider

$$\begin{aligned}&\int \mathbb{1}_{E(\mu_0+h,(\Gamma_0+A)(\Gamma_0+A)',\rho_0+s)}(y)\Gamma_0^{-1}(y-\mu_0)P(\mathrm{d}y) - \int \mathbb{1}_{E(\mu_0,\Sigma_0,\rho_0)}(y)\Gamma_0^{-1}(y-\mu_0)P(\mathrm{d}y)\\ &\quad= \det(\Gamma_0)\int \left(\mathbb{1}_{E(\widetilde{h},(I+\widetilde{A})(I+\widetilde{A})',\rho_0+s)}(z) - \mathbb{1}_{E(0,I,\rho_0)}(z)\right)zf(\Gamma_0 z+\mu_0)\,\mathrm{d}z\\ &\quad= L'(0,0,0)(\widetilde{h},\widetilde{A},s) + o(\|(h,A,s)\|),\end{aligned}$$

by taking $\phi(z) = \det(\Gamma_0)zf(\Gamma_0 z+\mu_0)$ in Lemma A.1. It follows that

$$\begin{aligned}\frac{\partial}{\partial\theta}\left(\int_{E(m,G^2,r)} \Gamma_0^{-1}(y-\mu_0)P(\mathrm{d}y)\right)\bigg|_{\theta=\theta_0} &= \det(\Gamma_0)\int_{\partial B_0}\left(\frac{\omega'\widetilde{h}}{\rho_0} + \frac{\omega'(\widetilde{A}+\widetilde{A}')\omega}{2\rho_0} + s\right)\omega f(\Gamma_0\omega+\mu_0)\sigma_0(\mathrm{d}\omega)\\ &= \det(\Gamma_0)\int_{\partial B_0}\left(\frac{\omega'\Gamma_0^{-1}h}{\rho_0} + \frac{\omega'(\Gamma_0^{-1}A + A'\Gamma_0^{-1})\omega}{2\rho_0} + s\right)\omega f(\Gamma_0\omega+\mu_0)\sigma_0(\mathrm{d}\omega).\end{aligned}$$

We conclude that

$$\Lambda_1'(\theta_0) = -\gamma\Gamma_0^{-1}h + \det(\Gamma_0)\int_{\partial B_0}\left(\frac{\omega'\Gamma_0^{-1}h}{\rho_0} + \frac{\omega'(\Gamma_0^{-1}A + A'\Gamma_0^{-1})\omega}{2\rho_0} + s\right)\omega f(\Gamma_0\omega+\mu_0)\sigma_0(\mathrm{d}\omega).$$

Then, similar to (A.3), we have

$$\Lambda_2'(\theta_0) = \frac{\partial}{\partial\theta}\left(\int_{E(\mu_0,\Sigma_0,\rho_0)}\Big[G^{-1}(y-m)(y-m)'G^{-1}-I\Big]P(\mathrm{d}y)\right)\Big|_{\theta=\theta_0} + \frac{\partial}{\partial\theta}\left(\int_{E(m,G^2,r)}\Big[\Gamma_0^{-1}(y-\mu_0)(y-\mu_0)'\Gamma_0^{-1}-I\Big]P(\mathrm{d}y)\right)\Big|_{\theta=\theta_0}. \tag{A.6}$$

The first term in (A.6) can be decomposed as

$$\begin{aligned}&\frac{\partial}{\partial\theta}\left(\int_{E(\mu_0,\Sigma_0,\rho_0)}\Big[G^{-1}(y-\mu_0)(y-\mu_0)'\Gamma_0^{-1}-I\Big]P(\mathrm{d}y)\right)\Big|_{\theta=\theta_0}\\
&+\frac{\partial}{\partial\theta}\left(\int_{E(\mu_0,\Sigma_0,\rho_0)}\Big[\Gamma_0^{-1}(y-m)(y-\mu_0)'\Gamma_0^{-1}-I\Big]P(\mathrm{d}y)\right)\Big|_{\theta=\theta_0}\\
&+\frac{\partial}{\partial\theta}\left(\int_{E(\mu_0,\Sigma_0,\rho_0)}\Big[\Gamma_0^{-1}(y-\mu_0)(y-m)'\Gamma_0^{-1}-I\Big]P(\mathrm{d}y)\right)\Big|_{\theta=\theta_0}\\
&+\frac{\partial}{\partial\theta}\left(\int_{E(\mu_0,\Sigma_0,\rho_0)}\Big[\Gamma_0^{-1}(y-\mu_0)(y-\mu_0)'G^{-1}-I\Big]P(\mathrm{d}y)\right)\Big|_{\theta=\theta_0}.\end{aligned} \tag{A.7}$$

For the first term in (A.7), for $\|A\|\to 0$, consider

$$\begin{aligned}&\int_{E(\mu_0,\Sigma_0,\rho_0)}\Big[(\Gamma_0+A)^{-1}(y-\mu_0)(y-\mu_0)'\Gamma_0^{-1}-I\Big]P(\mathrm{d}y)-\int_{E(\mu_0,\Sigma_0,\rho_0)}\Big[\Gamma_0^{-1}(y-\mu_0)(y-\mu_0)'\Gamma_0^{-1}-I\Big]P(\mathrm{d}y)\\
&=\int_{E(\mu_0,\Sigma_0,\rho_0)}\Big[(I-\Gamma_0^{-1}A)\Gamma_0^{-1}(y-\mu_0)(y-\mu_0)'\Gamma_0^{-1}-I\Big]P(\mathrm{d}y)+o(\|A\|)\\
&=-\Gamma_0^{-1}A\int_{E(\mu_0,\Sigma_0,\rho_0)}\Gamma_0^{-1}(y-\mu_0)(y-\mu_0)'\Gamma_0^{-1}P(\mathrm{d}y)+o(\|A\|)=-\gamma\Gamma_0^{-1}A+o(\|A\|),\end{aligned}$$

where in the last two steps we use

$$\int_{E(\mu_0,\Sigma_0,\rho_0)}\Gamma_0^{-1}(y-\mu_0)(y-\mu_0)'\Gamma_0^{-1}P(\mathrm{d}y)=\gamma I, \tag{A.8}$$

which follows from (2.4). For the second term in (A.7), consider

$$\begin{aligned}&\int_{E(\mu_0,\Sigma_0,\rho_0)}\Big[\Gamma_0^{-1}(y-\mu_0-h)(y-\mu_0)'\Gamma_0^{-1}-I\Big]P(\mathrm{d}y)-\int_{E(\mu_0,\Sigma_0,\rho_0)}\Big[\Gamma_0^{-1}(y-\mu_0)(y-\mu_0)'\Gamma_0^{-1}-I\Big]P(\mathrm{d}y)\\
&=-\Gamma_0^{-1}h\int_{E(\mu_0,\Sigma_0,\rho_0)}(y-\mu_0)'\Gamma_0^{-1}P(\mathrm{d}y)=0,\end{aligned}$$

where we use (A.5) and (A.8). Because G and Γ_0 are symmetric, the last two terms in (A.7) are the transpose of the first two terms in (A.7). This leads to the following derivative for the first term in (A.6):

$$\frac{\partial}{\partial\theta}\left(\int_{E(\mu_0,\Sigma_0,\rho_0)}\Big[G^{-1}(y-m)(y-m)'G^{-1}-I\Big]P(\mathrm{d}y)\right)\Big|_{\theta=\theta_0}=-\gamma(\Gamma_0^{-1}A+A'\Gamma_0^{-1}).$$

For the second term on the right-hand side of (A.6), for $(h,A,s)\to(0,0,0)$, consider

$$\begin{aligned}&\int\mathbb{1}_{E(\mu_0+h,(\Gamma_0+A)(\Gamma_0+A)',\rho_0+s)}(y)\left[\Gamma_0^{-1}(y-\mu_0)(y-\mu_0)'\Gamma_0^{-1}-I\right]P(\mathrm{d}y)\\
&\quad-\int\mathbb{1}_{E(\mu_0,\Sigma_0,\rho_0)}(y)\left[\Gamma_0^{-1}(y-\mu_0)(y-\mu_0)'\Gamma_0^{-1}-I\right]P(\mathrm{d}y)\\
&=\det(\Gamma_0)\int\left(\mathbb{1}_{E(\widetilde{h},(I+\widetilde{A})(I+\widetilde{A})',\rho_0+s)}(z)-\mathbb{1}_{E(0,I,\rho_0)}(z)\right)\left[zz'-I\right]f(\Gamma_0z+\mu_0)\,\mathrm{d}z\\
&=L'(0,0,0)(\widetilde{h},\widetilde{A},s)+o(\|(h,A,s)\|),\end{aligned}$$

by taking $\phi(z) = \det(\Gamma_0)[zz' - I]f(\Gamma_0 z + \mu_0)$ in Lemma A.1. It follows that

$$\frac{\partial}{\partial\theta}\left(\int_{E(m,G^2,r)}\left[\Gamma_0^{-1}(y-\mu_0)(y-\mu_0)'\Gamma_0^{-1} - I\right]P(\mathrm{d}y)\right)\Bigg|_{\theta=\theta_0}$$

$$= \det(\Gamma_0)\int_{\partial B_0}\left(\frac{\omega'\widetilde{h}}{\rho_0} + \frac{\omega'(\widetilde{A}+\widetilde{A}')\omega}{2\rho_0} + s\right)\left(\omega\omega' - I\right)f(\Gamma_0\omega+\mu_0)\sigma_0(\mathrm{d}\omega)$$

$$= \det(\Gamma_0)\int_{\partial B_0}\left(\frac{\omega'\Gamma_0^{-1}h}{\rho_0} + \frac{\omega'(\Gamma_0^{-1}A + A'\Gamma_0^{-1})\omega}{2\rho_0} + s\right)\left(\omega\omega' - I\right)f(\Gamma_0\omega+\mu_0)\sigma_0(\mathrm{d}\omega).$$

We conclude that

$$\Lambda_2'(\theta_0)(h,A,s) = -\gamma(\Gamma_0^{-1}A + A'\Gamma_0^{-1}) + \int_{\partial B_0}\left(\frac{\omega'\Gamma_0^{-1}h}{\rho_0} + \frac{\omega'(\Gamma_0^{-1}A + A'\Gamma_0^{-1})\omega}{2\rho_0} + s\right)\left(\omega\omega' - I\right)\nu(\mathrm{d}\omega).$$

Noting that we take A symmetric, this finishes the proof. □

Lemma A.2. *Suppose that P satisfies* (B) *and let the MCD functional $\theta_0 = (\mu_0, \Gamma_0, \rho_0)$ be uniquely defined at P. Let Λ be defined by* (2.8) *and suppose that $\Lambda'(\theta_0)(h, A, s) = 0$. Then $\mathrm{Tr}(\Gamma_0^{-1}A) = 0$.*

Proof. From Theorem 3.1,

$$0 = \Lambda_2'(\theta_0)(h,A,s)$$
$$= -\gamma(\Gamma_0^{-1}A + A\Gamma_0^{-1}) + \int_{\partial B_0}\left(\frac{\omega'\Gamma_0^{-1}h}{\rho_0} + \frac{\omega'(\Gamma_0^{-1}A + A\Gamma_0^{-1})\omega}{2\rho_0} + s\right)\left(\omega\omega' - I\right)\nu(\mathrm{d}\omega),$$

where $B_0 = B(0, \rho_0)$. Taking traces yields

$$0 = -2\gamma\,\mathrm{Tr}(\Gamma_0^{-1}A) + (\rho_0^2 - k)\int_{\partial B_0}\left(\frac{\omega'\Gamma_0^{-1}h}{\rho_0} + \frac{\omega'(\Gamma_0^{-1}A + A\Gamma_0^{-1})\omega}{2\rho_0} + s\right)\nu(\mathrm{d}\omega).$$

Because $\Lambda_3'(\theta_0) = 0$, it follows from Theorem 3.1 that the second term on the right-hand side is zero, which proves the lemma. □

Lemma A.3. *Suppose that P satisfies* (B) *and* (3.2)*, and let the MCD functional $\theta_0 = (\mu_0, \Gamma_0, \rho_0)$ be uniquely defined at P. Let Λ be defined by* (2.8) *and suppose that $\Lambda'(\theta_0)(h, A, s) = 0$. Then*

$$s = -\frac{1}{2\rho_0\nu_0}\int_{\partial B_0}\omega'(\Gamma_0^{-1}A + A\Gamma_0^{-1})\omega\,\nu(\mathrm{d}\omega), \tag{A.9}$$

where $B_0 = B(0, \rho_0)$ and $\nu_0 = \nu(\partial B_0)$, with ν defined in (3.1)*. If, in addition,* (3.3) *holds, then $h = 0$.*

Proof. If f satisfies (3.2), then

$$\int_{\partial B_0}\omega_i\omega_j\omega_m\,\nu(\mathrm{d}\omega) = \int_{\partial B_0}\omega_i\,\nu(\mathrm{d}\omega) = 0, \quad \text{for all } i,j,m = 1,2,\ldots,k, \tag{A.10}$$

where $\nu(\mathrm{d}\omega)$ is defined by (3.1). Hence, from Theorem 3.1, we get

$$0 = \Lambda_3'(\theta_0) = \frac{1}{2\rho_0}\int_{\partial B_0}\omega'(\Gamma_0^{-1}A + A\Gamma_0^{-1})\omega\,\nu(\mathrm{d}\omega) + s\nu(\partial B_0),$$

which yields the first statement. Moreover, (A.10) and Theorem 3.1 also yield that

$$0 = \Lambda_1'(\theta_0) = -\gamma\Gamma_0^{-1}h + \int_{\partial B_0}\frac{\omega'\Gamma_0^{-1}h}{\rho_0}\omega\nu(\mathrm{d}\omega) = \frac{1}{\rho_0}\left(\int_{\partial B_0}\omega\omega'\,\nu(\mathrm{d}\omega) - \gamma\rho_0 I\right)\Gamma_0^{-1}h.$$

Because f satisfies (3.2), the matrix on the right-hand side is a diagonal matrix with elements

$$\int_{\partial B_0}\omega_i^2\,\nu(\mathrm{d}\omega) - \gamma\rho_0.$$

Therefore, from (3.3), it follows that $h = 0$. □

The proof of Theorem 3.2 requires the following lemma.

Lemma A.4. *Let A be a $k \times k$ matrix and suppose that $\Gamma A + A\Gamma = 0$, for some $k \times k$ positive definite symmetric matrix Γ. Then $A = 0$.*

Proof. Since Γ is positive definite symmetric, there exists a basis of eigenvectors of Γ. Choose v to be an eigenvector with eigenvalue $\lambda > 0$. Then $\Gamma(Av) + \lambda(Av) = 0$. This means that either Av is an eigenvector of Γ with eigenvalue $-\lambda < 0$, which is impossible since Γ is positive definite, or $Av = 0$. This holds for all eigenvectors of Γ, and therefore $A = 0$. □

Proof of Theorem 3.2. Suppose that $\Lambda_j'(\theta_0)(h, A, s) = 0$ for $j = 1, 2, 3$; then it suffices to show that $(h, A, s) = (0, 0, 0)$. Because (3.4) implies (3.2), it follows from Lemma A.3 that $h = 0$. Furthermore, $\Lambda_2'(\theta_0) = 0$ and $\Lambda_3'(\theta_0) = 0$ imply that

$$0 = \Lambda_2'(\theta_0) = -\gamma S + \int_{\partial B_0} \left(\frac{\omega' S \omega}{2\rho_0} + s \right) \omega\omega' \, \nu(\mathrm{d}\omega), \tag{A.11}$$

where $B_0 = B(0, \rho_0)$ and $S = \Gamma_0^{-1} A + A \Gamma_0^{-1}$ is symmetric. Condition (3.4) implies that

$$\int_{\partial B_0} \omega_i \omega_j \omega_m \omega_n \, \nu(\mathrm{d}\omega) = \begin{cases} \int_{\partial B_0} \omega_i^2 \omega_m^2 \, \nu(\mathrm{d}\omega), & \text{for } m = n;\ i = j, \\ \int_{\partial B_0} \omega_m^2 \omega_n^2 \, \nu(\mathrm{d}\omega), & \text{for } m \neq n;\ \{i, j\} = \{m, n\}, \\ 0, & \text{otherwise}, \end{cases} \tag{A.12}$$

where $\nu(\mathrm{d}\omega)$ is defined by (3.1). Consider the (m, n)-th element of equation (A.11) for $m \neq n$. Then it follows from (A.12) that

$$0 = -2\gamma\rho_0 S_{mn} + 2S_{mn} \int_{\partial B_0} \omega_m^2 \omega_n^2 \, \nu(\mathrm{d}\omega) = 2 \left(\int_{\partial B_0} \omega_m^2 \omega_n^2 \, \nu(\mathrm{d}\omega) - \gamma\rho_0 \right) S_{mn}.$$

The factor in front of S_{mn} is non-zero by assumption (3.6), so $S_{mn} = 0$ for all $m \neq n$. Finally, consider the (m, m)-th element of (A.11) and insert (A.9), which is obtained from Lemma A.3. Then we get

$$0 = -2\gamma\rho_0 S_{mm} + \sum_{i=1}^{k} \left(\int_{\partial B_0} \omega_i^2 \omega_m^2 \, \nu(\mathrm{d}\omega) - \frac{1}{\nu_0} \int_{\partial B_0} \omega_i^2 \, \nu(\mathrm{d}\omega) \int_{\partial B_0} \omega_m^2 \, \nu(\mathrm{d}\omega) \right) S_{ii}.$$

The right-hand side is of the form Mx, where $x = \mathrm{diag}(S)$ and M is defined in (3.5). However, since $\mathrm{Tr}(S) = 0$ according to Lemma A.2, from (3.7) we conclude that $S_{mm} = 0$ for all $m = 1, 2, \ldots, k$. It follows that $S = \Gamma_0^{-1} A + A\Gamma_0^{-1} = 0$, and consequently, by (A.9), we have $s = 0$. Furthermore, from Lemma A.4, we conclude that $A = 0$. □

A.2. Proofs for Section 4

Proof of Proposition 4.1. Conditions (B) and (3.4) are immediate. From (4.2), we find that $f(\Gamma_0\omega + \mu_0)$ is constant on ∂B_0:

$$f(\Gamma_0\omega + \mu_0) = \det(\Sigma)^{-1/2} h(\alpha(\gamma)^2 \|\omega\|^2) = \alpha(\gamma)^k \det(\Gamma_0)^{-1} h(r(\gamma)^2),$$

and

$$\nu(\mathrm{d}\omega) = \alpha(\gamma)^k h(r(\gamma)^2) \sigma_0(\mathrm{d}\omega), \tag{A.13}$$

for $\omega \in \partial B_0$. One can easily check that, for all $i, j = 1, 2, \ldots, k$ (e.g., see Lemma 1 in [10]),

$$\begin{aligned} \int_{\partial B_0} \omega_i^2 \, \nu(\mathrm{d}\omega) &= \frac{1}{k} \int_{\partial B_0} \|\omega\|^2 \, \nu(\mathrm{d}\omega) = \frac{2\pi^{k/2}}{k\Gamma(k/2)} h(r(\gamma)^2) r(\gamma)^k \rho_0, \\ \int_{\partial B_0} \omega_i^2 \omega_j^2 \, \nu(\mathrm{d}\omega) &= \frac{1 + 2\delta_{ij}}{k(k+2)} \int_{\partial B_0} \|\omega\|^4 \, \nu(\mathrm{d}\omega) \\ &= (1 + 2\delta_{ij}) \frac{2\pi^{k/2}}{k(k+2)\Gamma(k/2)} h(r(\gamma)^2) r(\gamma)^k \rho_0^3, \\ \nu(\partial B_0) &= \frac{2\pi^{k/2}}{\Gamma(k/2)} h(r(\gamma)^2) r(\gamma)^{k-1} \alpha(\gamma) = \frac{k}{\rho_0^2} \int_{\partial B_0} \omega_i^2 \, \nu(\mathrm{d}\omega). \end{aligned} \tag{A.14}$$

Because h is decreasing and non-constant on $[0, r(\gamma))$, conditions (3.3) and (3.6) are fulfilled:

$$\begin{aligned}\gamma &= \frac{2\pi^{k/2}}{\Gamma(k/2)}\int_0^{r(\gamma)} h(r^2)r^{k-1}\,\mathrm{d}r > \frac{2\pi^{k/2}}{k\Gamma(k/2)}h(r(\gamma)^2)r(\gamma)^k,\\ \gamma\alpha(\gamma)^2 &= \frac{2\pi^{k/2}}{k\Gamma(k/2)}\int_0^{r(\gamma)} h(r^2)r^{k+1}\,\mathrm{d}r > \frac{2\pi^{k/2}}{k(k+2)\Gamma(k/2)}h(r(\gamma)^2)r(\gamma)^{k+2}.\end{aligned} \tag{A.15}$$

Finally, from the equations above, it follows that the matrix M defined in (3.5) can be decomposed as $M = c_1I + c_2\mathbf{1}\mathbf{1}'$, where $\mathbf{1} = (1, 1, \ldots, 1)'$, and

$$\begin{aligned}c_1 &= \frac{4\pi^{k/2}}{k(k+2)\Gamma(k/2)}h(r(\gamma)^2)r(\gamma)^k\rho_0^3 - 2\gamma\rho_0,\\ c_2 &= -\frac{4\pi^{k/2}}{k^2(k+2)\Gamma(k/2)}h(r(\gamma)^2)r(\gamma)^k\rho_0^3.\end{aligned}$$

Because $Mx = c_1x + c_2(x_1 + \cdots + x_k)\mathbf{1}$, it follows that, if $x_1 + \cdots + x_k = 0$, $Mx = 0$ implies that $x = 0$ as long as $c_1 \neq 0$, i.e.,

$$\frac{2\pi^{k/2}}{k(k+2)\Gamma(k/2)}h(r(\gamma)^2)r(\gamma)^{k+2} \neq \gamma\alpha(\gamma)^2,$$

which follows from (A.15). □

Proof of Theorem 4.1. From (A.13), it follows that

$$\begin{aligned}&\int_{\partial B_0} \omega_i\omega_j\,\nu(\mathrm{d}\omega) = 0, \quad \text{for } i \neq j\\ &\int_{\partial B_0} \omega_i\,\nu(\mathrm{d}\omega) = 0 \quad\text{and}\quad \int_{\partial B_0}\omega_i\omega_j\omega_k\,\nu(\mathrm{d}\omega) = 0, \quad \text{for all } i, j, k,\end{aligned} \tag{A.16}$$

Hence, from Theorem 3.1, we find that

$$\Lambda_1'(\theta_0)(h, A, s) = -\gamma\Gamma_0^{-1}h + \frac{1}{\rho_0}\int_{\partial B_0}\omega\omega'\Gamma_0^{-1}h\,\nu(\mathrm{d}\omega) = \beta_1 h,$$

where, according to (A.14) and (A.15),

$$\beta_1 = \frac{1}{k\alpha\rho_0}\int_{\partial B_0}\|\omega\|^2\,\nu(\mathrm{d}\omega) - \frac{\gamma}{\alpha} = \frac{1}{\alpha}\left(\frac{\rho_0}{k}\nu_0 - \gamma\right) < 0.$$

Next, consider $\Lambda_3'(\theta_0)$. From (A.16), we find that

$$\Lambda_3'(\theta_0)(h, A, s) = \frac{1}{\rho_0}\int_{\partial B_0}\omega'\Gamma_0^{-1}A\omega\,\nu(\mathrm{d}\omega) + s\nu_0.$$

From (A.14), the first term on the right-hand side is

$$\frac{1}{\alpha\rho_0}\int_{\partial B_0}\omega'A\omega\,\nu(\mathrm{d}\omega) = \frac{1}{k\alpha\rho_0}\int_{\partial B_0}\|\omega\|^2\,\mathrm{Tr}(A)\nu(\mathrm{d}\omega) = \frac{\rho_0\nu_0}{k\alpha}\,\mathrm{Tr}(A).$$

This means that $\Lambda_3'(\theta_0)(h, A, s) = \beta_5\,\mathrm{Tr}(A) + \beta_6 s$. Finally, from (4.2) and (A.16),

$$\begin{aligned}\Lambda_2'(\theta_0)(h, A, s) &= -\gamma(\Gamma_0^{-1}A + A\Gamma_0^{-1}) + \int_{\partial B_0}\left(\frac{\omega'(\Gamma_0^{-1}A + A\Gamma_0^{-1})\omega}{2\rho_0} + s\right)\left(\omega\omega' - I\right)\nu(\mathrm{d}\omega)\\ &= -\frac{2\gamma}{\alpha}A - \Lambda_3'(\theta_0)\cdot I + \frac{1}{\alpha\rho_0}\int_{\partial B_0}(\omega'A\omega)\omega\omega'\,\nu(\mathrm{d}\omega) + s\int_{\partial B_0}\omega\omega'\,\nu(\mathrm{d}\omega)\\ &= -\frac{2\gamma}{\alpha}A - (\beta_5\,\mathrm{Tr}(A) + \beta_6 s)\cdot I + \frac{1}{\alpha\rho_0}\int_{\partial B_0}(\omega'A\omega)\omega\omega'\,\nu(\mathrm{d}\omega) + \frac{\rho_0^2}{k}\nu_0\cdot sI.\end{aligned}$$

Consider the (m, n)-th element of the third integral on the right-hand side. From (A.12) and (A.14), it follows that this integral is equal to

$$\frac{1}{\alpha\rho_0}\sum_{i=1}^k\sum_{j=1}^k A_{ij}\int_{\partial B_0}\omega_i\omega_j\omega_m\omega_n\,\nu(\mathrm{d}\omega) = \frac{\rho_0^3\nu_0}{\alpha k(k+2)}\left(\mathrm{Tr}(A)\mathbb{1}_{\{m=n\}} + 2A_{mn}\right),$$

which means that

$$\frac{1}{\alpha\rho_0}\int_{\partial B_0}(\omega' A\omega)\omega\omega'\,\nu(\mathrm{d}\omega)=\frac{\rho_0^3\nu_0}{\alpha k(k+2)}\left(\mathrm{Tr}(A)\cdot I+2A\right).$$

Summarizing, in the expression of $\Lambda_2'(\theta_0)$, the coefficient of A is

$$\beta_2=\frac{2\rho_0^3\nu_0}{\alpha k(k+2)}-\frac{2\gamma}{\alpha},$$

the coefficient of $\mathrm{Tr}(A)\cdot I$ is

$$\beta_3=\frac{\rho_0^3\nu_0}{\alpha k(k+2)}-\frac{\rho_0\nu_0}{k\alpha},$$

and the coefficient of sI is

$$\beta_4=\frac{\rho_0^2}{k}\nu_0-\nu_0.$$

From (A.14) and (A.15), it can be seen that $\beta_2<0$.

To determine the expression of the inverse mapping, put $D(h,A,s)=(g,B,t)$ and solve for (h,A,s). For the vector valued component of D, we have $g=D_1(h,A,s)=\beta_1 h$. Since $\beta_1<0$, this immediately gives $h=\beta_1^{-1}g$. For the remaining mappings, put

$$\begin{aligned}B&=D_2(h,A,s)=\beta_2 A+\beta_3\,\mathrm{Tr}(A)\cdot I+\beta_4 s\cdot I\\ t&=D_3(h,A,s)=\beta_5\,\mathrm{Tr}(A)+\beta_6 s.\end{aligned}\tag{A.17}$$

By taking traces in the first equation, we can solve for $\mathrm{Tr}(A)$ and s:

$$\begin{aligned}c\,\mathrm{Tr}(A)&=\beta_6\,\mathrm{Tr}(B)-k\beta_4 t\\ cs&=(\beta_2+k\beta_3)t-\beta_5\,\mathrm{Tr}(B),\end{aligned}\tag{A.18}$$

where $c=\beta_2\beta_6+k\beta_3\beta_6-k\beta_4\beta_5=-2\gamma\beta_6/\alpha$. Since $\beta_2<0$ and $\beta_6>0$, from (A.17) and (A.18), it follows that

$$\begin{aligned}A&=\beta_2^{-1}\left(B-\beta_3\,\mathrm{Tr}(A)\cdot I-\beta_4 s\cdot I\right)\\ &=\beta_2^{-1}B-\frac{\beta_3}{c\beta_2}\left(\beta_6\,\mathrm{Tr}(B)-k\beta_4 t\right)\cdot I-\frac{\beta_4}{c\beta_2}\left(-\beta_5\,\mathrm{Tr}(B)+(\beta_2+k\beta_3)t\right)\cdot I\\ &=\beta_2^{-1}B+\frac{\alpha(\beta_3\beta_6-\beta_4\beta_5)}{2\gamma\beta_2\beta_6}\,\mathrm{Tr}(B)\cdot I-\frac{\alpha\beta_2\beta_3}{2\gamma\beta_2\beta_6}t\cdot I\end{aligned}$$

and

$$s=-\frac{\beta_5}{c}\,\mathrm{Tr}(B)+\frac{\beta_2+k\beta_3}{c}t=\frac{\alpha\beta_5}{2\gamma\beta_6}\,\mathrm{Tr}(B)-\frac{\alpha(\beta_2+k\beta_3)}{2\gamma\beta_6}t.\quad\square$$

Proof of Corollary 4.1. Since $\Lambda(\theta_0)^{-1}$ is a linear mapping and $\mathbb{E}\Psi(X_i,\theta_0)=0$, we obtain from (1.1)

$$\widehat{\theta}_n-\theta_0=-\frac{1}{n}\sum_{i=1}^n\Lambda'(\theta_0)^{-1}\Psi(X_i,\theta_0)+o_{\mathbb{P}}(n^{-1/2}),\tag{A.19}$$

where Ψ is defined in (2.7). In particular, we have

$$\begin{aligned}\sqrt{n}\widehat{\mu}_n&=-\frac{1}{\sqrt{n}}\sum_{i=1}^n\left[D^{\mathrm{inv}}\right]_1\Psi(X_i,\theta_0)+o_{\mathbb{P}}(1),\\ \sqrt{n}\left(\widehat{\rho}_n-\frac{r}{\alpha}\right)&=-\frac{1}{\sqrt{n}}\sum_{i=1}^n\left[D^{\mathrm{inv}}\right]_3\Psi(X_i,\theta_0)+o_{\mathbb{P}}(1).\end{aligned}\tag{A.20}$$

According to (4.2), $\theta_0=(\mu_0,\Gamma_0,\rho_0)=(0,\alpha I,r/\alpha)$, so

$$\begin{aligned}\Psi_1(x,\theta_0)&=\mathbb{1}_{\{\|x\|\le r\}}\alpha^{-1}x,\\ \Psi_2(x,\theta_0)&=\mathbb{1}_{\{\|x\|\le r\}}\left(\alpha^{-2}xx'-I\right),\\ \Psi_3(x,\theta_0)&=\mathbb{1}_{\{\|x\|\le r\}}-\gamma.\end{aligned}\tag{A.21}$$

Insert $g_0 = \Psi_1(x, \theta_0)$, $B_0 = \Psi_2(x, \theta_0)$ and $t_0 = \Psi_3(x, \theta_0)$ in the expressions for $D^{\text{inv}}(g, B, t)$ given in Theorem 4.1. Then we find that

$$\begin{aligned}\left[D^{\text{inv}}\right]_1 \Psi(x, \theta_0) &= (\alpha\beta_1)^{-1} \mathbb{1}_{\{\|x\| \leq r\}} x \\ \left[D^{\text{inv}}\right]_3 \Psi(x, \theta_0) &= \frac{\alpha\beta_5}{2\gamma\beta_6} \mathbb{1}_{\{\|x\| \leq r\}} \left(\frac{\|x\|^2}{\alpha^2} - k\right) - \frac{\alpha(\beta_2 + k\beta_3)}{2\gamma\beta_6} \left(\mathbb{1}_{\{\|x\| \leq r\}} - \gamma\right).\end{aligned} \tag{A.22}$$

Together with (A.20), this immediately yields the expansion for $\sqrt{n}\widehat{\mu}_n$ and the expansion for $\sqrt{n}\,(\widehat{\rho}_n - r/\alpha)$ with

$$\begin{aligned}\lambda_1 &= -\frac{\beta_5}{2\alpha\gamma\beta_6} = -\frac{r}{2k\gamma\alpha^3}, \\ \lambda_2 &= \frac{\alpha(\beta_2 + k\beta_3 + k\beta_5)}{2\gamma\beta_6} = \frac{r^3}{2k\gamma\alpha^3} - \frac{1}{\beta_6}, \\ \lambda_3 &= -\frac{\alpha(\beta_2 + k\beta_3)}{2\beta_6} = \frac{\gamma}{\beta_6} + \frac{r}{2k\alpha^3}\left(k\alpha^2 - r^2\right).\end{aligned}$$

To obtain the expansion for the covariance estimator, note that P satisfies the conditions of Theorem 4.2 in [3]. This means that $\widehat{\Gamma}_n \to \alpha I$ with probability 1, so

$$\widehat{\Sigma}_n - \alpha^2 I = (\widehat{\Gamma}_n + \alpha I)(\widehat{\Gamma}_n - \alpha I) = 2\alpha(\widehat{\Gamma}_n - \alpha I) + o(1),$$

with probability 1. Hence, from (A.19), we obtain

$$\sqrt{n}\left(\widehat{\Sigma}_n - \alpha^2 I\right) = -\frac{2\alpha}{\sqrt{n}} \sum_{i=1}^n \left[D^{\text{inv}}\right]_2 \Psi(X_i, \theta_0) + o_{\mathbb{P}}(1), \tag{A.23}$$

where

$$\begin{aligned}\left[D^{\text{inv}}\right]_2 \Psi(x, \theta_0) \;=\; &\beta_2^{-1} \mathbb{1}_{\{\|x\| \leq r\}} \left(\frac{xx'}{\alpha^2} - I\right) + \frac{\alpha(\beta_3\beta_6 - \beta_4\beta_5)}{2\gamma\beta_2\beta_6} \mathbb{1}_{\{\|x\| \leq r\}} \left(\frac{\|x\|^2}{\alpha^2} - k\right) \cdot I \\ &+ \frac{\alpha\beta_4}{2\gamma\beta_6} \left(\mathbb{1}_{\{\|x\| \leq r\}} - \gamma\right) \cdot I.\end{aligned} \tag{A.24}$$

This yields the expansion for $\sqrt{n}\left(\widehat{\Sigma}_n - \alpha^2 I\right)$ with

$$\begin{aligned}\kappa_1 &= \frac{2\alpha}{\beta_2} + \frac{k\alpha^2(\beta_3\beta_6 - \beta_4\beta_5)}{\gamma\beta_2\beta_6} - \frac{\alpha^2\beta_4}{\gamma\beta_6} = -\frac{r^2}{k\gamma}, \\ \kappa_2 &= \frac{\beta_4\beta_5 - \beta_3\beta_6}{\gamma\beta_2\beta_6} = \frac{\alpha\beta_2 + 2\gamma}{k\gamma\alpha\beta_2}, \\ \kappa_3 &= -\frac{2}{\alpha\beta_2}, \\ \kappa_4 &= \frac{\alpha^2\beta_4}{\beta_6} = \frac{r^2 - k\alpha^2}{k}. \quad \square\end{aligned}$$

Proof of Theorem 4.2. The expansion for $\sqrt{n}\widehat{\mu}_n$ given in Corollary 4.1, together with the fact that $\mathbb{E}\{\|X_1\| \leq r\}X_1 = 0$, yields that $\sqrt{n}\widehat{\mu}_n$ is asymptotically normal with mean zero and covariance matrix

$$\tau^2 \mathbb{E}\{\|X_1\| \leq r\} X_1 X_1' = \frac{\tau^2}{k} \mathbb{E}\{\|X_1\| \leq r\} \|X_1\|^2 \cdot I.$$

Since $\tau = -(\alpha\beta_1)^{-1}$, together with (4.3), we find that

$$\xi = \frac{\tau^2}{k} \mathbb{E}\{\|X_1\| \leq r\} \|X_1\|^2 = \frac{k^2\gamma\alpha^4}{(k\gamma\alpha - r\nu_0)^2},$$

which proves part (ii). To prove (iii), first note that, from Corollary 4.1, it follows that

$$\sqrt{n}(\widehat{\Sigma}_n - \alpha^2 I) = \frac{1}{\sqrt{n}} \sum_{i=1}^n \left(\ell(\|X_i\|) \frac{X_i X_i'}{\|X_i\|^2} + m(\|X_i\|) \cdot I\right) + o_{\mathbb{P}}(1), \tag{A.25}$$

where $\ell(y) = \kappa_3 \mathbb{1}_{\{\|y\|\le r\}} y^2$ and $m(y) = \mathbb{1}_{\{\|y\|\le r\}}(\kappa_1 + \kappa_2 y^2) + \kappa_4$. Note that, according to (4.3),

$$\mathbb{E}\ell(\|X_1\|) = -\frac{2}{\alpha\beta_2}\mathbb{E}\{\|X_1\| \le r\}\|X_1\|^2 = -\frac{2\alpha k\gamma}{\beta_2},$$
$$\mathbb{E}m(\|X_1\|) = -\frac{r^2}{k\gamma}\mathbb{E}\{\|X_1\| \le r\} + \frac{\alpha\beta_2 + 2\gamma}{k\gamma\alpha\beta_2}\mathbb{E}\{\|X_1\| \le r\}\|X_1\|^2 + \frac{r^2 - k\alpha^2}{k} = \frac{2\alpha\gamma}{\beta_2},$$

so $\mathbb{E}\left[\ell(\|X_1\|) + km(\|X_1\|)\right] = 0$. Since also $\mathbb{E}\ell^2(\|X_1\|) < \infty$ and $\mathbb{E}m^2(\|X_1\|) < \infty$, it follows from Lemma 5 in [10] that the sum on the right-hand side of (A.25) is asymptotically normal with mean zero and covariance matrix $\sigma_1(I + C_{k,k}) + \sigma_2 \,\mathrm{vec}(I)\mathrm{vec}(I)'$, where

$$\sigma_1 = \frac{\mathbb{E}\,\ell^2(\|X_1\|)}{k(k+2)},$$
$$\sigma_2 = \frac{\mathbb{E}\,\ell^2(\|X_1\|)}{k(k+2)} + \mathbb{E}\,m^2(\|X_1\|) + \frac{2}{k}\mathbb{E}\,\ell(\|X_1\|)m(\|X_1\|).$$

If we fill in the expressions for $\ell(\|X_1\|)$ and $m(\|X_1\|)$, we get

$$\sigma_1 = \frac{\kappa_3^2}{k(k+2)}\mathbb{E}\mathbb{1}_{\{\|X_1\|\le r\}}\|X_1\|^4$$
$$\sigma_2 = \left(\frac{\kappa_3^2}{k(k+2)} + \kappa_2^2 + \frac{2\kappa_2\kappa_3}{k}\right)\mathbb{E}\mathbb{1}_{\{\|X_1\|\le r\}}\|X_1\|^4 + \left(2(\kappa_1+\kappa_4)\kappa_2 + \frac{2}{k}\kappa_3(\kappa_1+\kappa_4)\right)\mathbb{E}\mathbb{1}_{\{\|X_1\|\le r\}}\|X_1\|^2 + \kappa_1(\kappa_1+2\kappa_4)\mathbb{E}\mathbb{1}_{\{\|X_1\|\le r\}} + \kappa_4^2.$$

Substituting the expressions for $\kappa_1, \kappa_2, \kappa_4$ given in Corollary 4.3 together with (4.3) and (4.4) proves (iii). For part (iv), note that

$$\mathbb{E}\left[\lambda_1\mathbb{1}_{\{\|X_i\|\le r\}}\|X_i\|^2 + \lambda_2\mathbb{1}_{\{\|X_i\|\le r\}} + \lambda_3\right] = \lambda_1 k\gamma\alpha^2 + \lambda_2\gamma + \lambda_3 = 0.$$

Therefore, from the expansion given in Corollary 4.1, it follows that $\sqrt{n}(\widehat{\rho}_n - r/\alpha)$ is asymptotically normal with variance

$$\sigma_\rho^2 = \mathbb{E}\left(\lambda_1\mathbb{1}_{\{\|X_i\|\le r\}}\|X_i\|^2 + \lambda_2\mathbb{1}_{\{\|X_i\|\le r\}} + \lambda_3\right)^2$$
$$= \lambda_1^2\mathbb{E}\mathbb{1}_{\{\|X_i\|\le r\}}\|X_i\|^4 + \lambda_1(\lambda_2+\lambda_3)\mathbb{E}\mathbb{1}_{\{\|X_i\|\le r\}}\|X_i\|^2 + \lambda_2(\lambda_2+\lambda_3)\mathbb{E}\mathbb{1}_{\{\|X_i\|\le r\}} + \lambda_3^2.$$

Substituting the expressions for λ_2, λ_3 given in Corollary 4.3 together with (4.3) and (4.4) proves (iv). Finally, for part (i), first note that, according to Theorem 5.1 in [3], $\widehat{\mu}_n$, $\widehat{\Sigma}_n$ and $\widehat{\rho}_n$ are mutually asymptotically normal. Hence, it suffices to prove that the quantities considered in part (i) are asymptotically uncorrelated. However, this follows directly from the expansions given in Corollary 4.1 together with the symmetry properties of spherically symmetric densities. □

Proof of Corollary 4.3. According to Theorem 1 in [2], the MCD functional $\theta_0 = (\mu_0, \Gamma_0, \rho_0)$ as defined in (2.6) is unique, and since P has a density, all conditions of Theorem 5.2 in [3] are satisfied. It follows from this theorem that the influence function for the functional $\Theta(P) = \big(\mu(P), \Gamma(P), \rho(P)\big)$, where $\Gamma(P)^2 = \Sigma(P)$, is given by

$$\mathrm{IF}(x; \Theta, P) = -\Lambda'(\theta_0)^{-1}\Psi(x, \theta_0), \tag{A.26}$$

where Ψ is defined in (2.7). The expressions for $\mathrm{IF}(x; \mu, P)$ and $\mathrm{IF}(x; \rho, P)$ follow directly from (A.22). To obtain the influence function for the covariance functional, first note that, according to the continuity of the MCD functional, $\Gamma(P_{\varepsilon,x}) \to \Gamma(P) = \alpha I$, as $\varepsilon \downarrow 0$, where $P_{\varepsilon,x} = (1-\varepsilon)P + \varepsilon\delta_x$. This means that

$$\Sigma(P_{\varepsilon,x}) - \Sigma(P) = \big(\Gamma(P_{\varepsilon,x}) + \Gamma(P)\big)\big(\Gamma(P_{\varepsilon,x}) - \Gamma(P)\big) = 2\alpha\big(\Gamma(P_{\varepsilon,x}) - \Gamma(P)\big) + o(\varepsilon).$$

It follows that

$$\mathrm{IF}(x; \Sigma, P) = 2\alpha \cdot \mathrm{IF}(x; \Gamma, P) = -2\alpha\left[D^{\mathrm{inv}}\right]_2 \Psi(x, \theta_0).$$

The expression then follows from (A.24). □

References

[1] J. Agulló, C. Croux, S. Van Aelst, The multivariate least-trimmed squares estimator, Journal of Multivariate Analysis 99 (2008) 311–338.

[2] R.W. Butler, P.L. Davies, M. Jhun, Asymptotics for the minimum covariance determinant estimator, Annals of Statistics 3 (1993) 1385–1400.

[3] E.A. Cator, H.P. Lopuhaä, Central limit theorem and influence function for the MCD estimators at general multivariate distributions (submitted for publication). Available at: arXiv:math.ST/0907.0079.

[4] C. Croux, G. Haesbroeck, Influence function and efficiency of the minimum covariance determinant scatter matrix estimator, Journal of Multivariate Analysis 71 (1999) 161–190.
[5] C. Croux, G. Haesbroeck, Principal component analysis based on robust estimators of the covariance or correlation matrix: influence functions and efficiencies, Biometrika 87 (2000) 603–618.
[6] M. Fekri, A. Ruiz-Gazen, Robust weighted orthogonal regression in the errors-in-variables model, Journal of Multivariate Analysis 88 (2004) 89–108.
[7] D.M. Hawkins, G.J. McLachlan, High-breakdown linear discriminant analysis, Journal of the American Statistical Association 92 (1997) 136–143.
[8] M. Hubert, P.J. Rousseeuw, S. Van Aelst, High-breakdown robust multivariate methods, Statistical Science 23 (2008) 92–119.
[9] H.P. Lopuhaä, On the relation between S-estimators and M-estimators of multivariate location and covariance, Annals of Statistics 17 (1989) 1662–1683.
[10] H.P. Lopuhaä, Asymptotic expansion of S-estimators of location and covariance, Statistica Neerlandica 51 (1997) 220–237.
[11] H.P. Lopuhaä, P.J. Rousseeuw, Breakdown points of affine equivariant estimators of multivariate location and covariance matrices, Annals of Statistics 19 (1991) 229–248.
[12] H.P. Lopuhaä, Asymptotics of reweighted estimators of multivariate location and scatter, Annals of Statistics 27 (1999) 1638–1665.
[13] G. Pison, P.J. Rousseeuw, P. Filzmoser, C. Croux, Robust factor analysis, Journal of Multivariate Analysis 84 (2003) 145–172.
[14] P.J. Rousseeuw, Multivariate estimation with high breakdown point, in: Mathematical Statistics and Applications, vol. B, Reidel, Dordrecht, 1985, pp. 283–297 (Bad Tatzmannsdorf, 1983).
[15] P.J. Rousseeuw, K. Van Driessen, S. Van Aelst, J. Agulló, Robust multivariate regression, Technometrics 46 (2004) 293–305.
[16] S. Serneels, T. Verdonck, Principal component analysis for data containing outliers and missing elements, Computational Statistics and Data Analysis 52 (2008) 1712–1727.
[17] S. Taskinen, C. Croux, A. Kankainen, E. Ollila, H. Oja, Influence functions and efficiencies of the canonical correlation and vector estimates based on scatter and shape matrices, Journal of Multivariate Analysis 97 (2006) 359–384.
[18] J. Zhou, Robust dimension reduction based on canonical correlation, Journal of Multivariate Analysis 100 (2009) 195–209.