Applied Mathematics, 2026, 17(7), 446-475
https://www.scirp.org/journal/am
ISSN Online: 2152-7393
ISSN Print: 2152-7385

General Depth-Based Trimmed Means and Trimmed Scatter Matrices

Jin Wang

Northern Arizona University, Flagstaff, USA
Email: jin.wang@nau.edu

How to cite this paper: Wang, J. (2026) General Depth-Based Trimmed Means and Trimmed Scatter Matrices. *Applied Mathematics*, **17**, 446-475.
https://doi.org/10.4236/am.2026.177027

Received: May 16, 2026
Accepted: July 28, 2026
Published: July 31, 2026

Abstract

Multivariate descriptive measures for location and scatter are the foundation of multivariate statistics and underpin many methods in the field. In this paper, we propose and study some new general depth-based trimmed means and trimmed scatter matrices. In addition to the basic properties and algorithms, we establish the asymptotic distributions of their sample versions. Using the asymptotic distributions, we study the asymptotic efficiencies of the sample trimmed means and the sample trimmed scatter matrices. Robustness is explored through finite-sample breakdown point. The results show that the sample trimmed means and trimmed scatter matrices are not only highly efficient but also exceptionally robust, making them very favorable estimators for multivariate location and scatter.

Keywords

Multivariate Trimmed Mean, Trimmed Scatter Matrix, Asymptotic Distribution, Asymptotic Efficiency, Robustness

1. Introduction and Preliminaries

Multivariate descriptive measures for location and scatter are the foundation of multivariate statistics and underpin many methods in the field. The classical notions of multivariate location and scatter are moment-based and are estimated by the sample mean vector and the sample covariance matrix. However, these estimators are not robust and are extremely sensitive to outlier(s). Many efforts have been made to solve this problem, and various estimators for multivariate location and scatter have been proposed. The widely used ones include the *M*-estimators [1], the minimum volume ellipsoid (MVE) and the minimum covariance determinant (MCD) estimators [2], the *S*-estimators [3] [4], the τ-estimators [5], the constrained *M*-estimators [6], and so on.

In addition, various nonparametric estimators for multivariate location and scatter have been developed via depth functions. Generally, a depth function $D_F(\boldsymbol{x})$ is a nonnegative real-valued mapping which provides a distribution-based center-outward ordering of points $\boldsymbol{x}$ in $\mathbb{R}^d$. Desirable properties for a depth function are: 1) affine invariance (*i.e.*, $D_{F_{\mathbf{A}X+\boldsymbol{b}}}(\mathbf{A}\boldsymbol{X}+\boldsymbol{b}) = D_{F_X}(\boldsymbol{x})$ for any nonsingular $d \times d$ matrix $\mathbf{A}$ and d -vector $\boldsymbol{b}$); 2) maximality at "center"; 3) monotonicity relative to the deepest point; 4) vanishing at infinity. Widely used depth functions include the halfspace depth [7], the projection depth [8] [9], the simplicial depth [10], the L^p depth [11], and the Mahalanobis depth [8] [12]. The general discussions on depth functions and their applications can be found in [11] [13] [14].

Given a depth function $D_F(\boldsymbol{x})$, the deepest point plays the role of "center" of the distribution F and thus is considered as a multivariate median, denoted by $\boldsymbol{M}_F$. The α *depth inner region* is defined as $I_F(\alpha) = \{\boldsymbol{x} \in \mathbb{R}^d : D_F(\boldsymbol{x}) \ge \alpha\}$, $\alpha \ge 0$, and the *pth central region* as $C_F(p) = I_F(\alpha(p))$ with $\alpha(p) = \sup_\alpha \{\alpha : P(I_F(\alpha)) \ge p\}$, *i.e.*, the smallest α -depth inner region having a probability weight of at least p, $0 \le p \le 1$. The boundary $\partial I_F(\alpha)$ of $I_F(\alpha)$ is called the α *depth contour* and can be characterized by the *radius function*,

$$r_F(\alpha, \boldsymbol{u}) = \inf\{r \ge 0 : r\boldsymbol{u} \notin I_F(\alpha)\}, \boldsymbol{u} \in S^{d-1} = \{\boldsymbol{u} \in \mathbb{R}^d : \|\boldsymbol{u}\| = 1\},$$

when $\boldsymbol{M}_F$ is the origin. If the underlying distribution F is continuous, the random depth $D_F(\boldsymbol{X})$ is continuous for all the depths mentioned above, and then $p = P(C_F(p)) = P(D_F(\boldsymbol{X}) \ge \alpha(p)) = 1 - F_D(\alpha(p))$. Therefore, $\alpha(p) = F_D^{-1}(1-p)$ and $C_F(p) = I_F(F_D^{-1}(1-p))$.

Utilizing a depth function $D_F(\boldsymbol{x})$, one can define a depth weighted mean as

$$L(F) = \int \boldsymbol{x} w_1(D_F(\boldsymbol{x}))dF(\boldsymbol{x}) / \int w_1(D_F(\boldsymbol{x}))dF(\boldsymbol{x})$$

and a depth weighted scatter matrix as

$$S(F) = \int (\boldsymbol{x} - L(F))(\boldsymbol{x} - L(F))' w_2(D_F(\boldsymbol{x}))dF(\boldsymbol{x}) / \int w_2(D_F(\boldsymbol{x}))dF(\boldsymbol{x}),$$

where $w_1(\cdot)$ and $w_2(\cdot)$ are suitable weight functions and can be different. Conceptionally, the general depth weighted means and scatter matrices include depth trimmed means and trimmed scatter matrices as a special case. However, when studying the properties of the depth weighted means and scatter matrices, some conditions on the weight functions were invoked such that the usual trimmed means and scatter matrices were excluded. For example, to establish the asymptotic distributions of the sample version of $L(F)$, Dümbgen [15], Massé [16], and Zuo, Cui, and He [17] all made the assumption that the weight function $w_1(t)$ is continuously differentiable, which excluded the usual trimmed mean for which the weight function is an indicator function. For the finite sample breakdown point of the sample version of $S(F)$, Zuo and Cui [18] assumed that both weight functions are positive, which ruled out any trimmed scatter matrices.

It is well-known that the univariate trimmed means are not only robust but also highly efficient (much more efficient than the univariate median). It is conjectured that multivariate trimmed means and scatter matrices have the same properties, which motivates this work. In fact, the α depth trimmed means based on the projection depth have been studied by Zuo [19], and the ones based on the halfspace depth have been studied by Donoho and Gasko [20], Massé [21], and Wang [22], with the α depth trimmed means defined as

$$\mu_F(\alpha) = \int_{I_F(\alpha)} \boldsymbol{x} w_1(D_F(\boldsymbol{x}))dF(\boldsymbol{x}) / \int_{I_F(\alpha)} w_1(D_F(\boldsymbol{x}))dF(\boldsymbol{x}).$$

However, since the depth value of a point depends on both the depth function and the underlying distribution, it is very hard to choose an appropriate α value for $\mu_F(\alpha)$ in practice, and thus their applications are thwarted. **Table 1** lists the theoretical trimming proportions $1-p$ for the $\mu_F(\alpha)$ based on the halfspace depth with $\alpha = 0.01$ when the underlying distributions are the d-dimensional standard normal distribution $N_d(\mathbf{0}, \mathbf{I}_{d\times d})$, the d-dimensional t distribution with 5 degrees of freedom $Mt_d(5, \mathbf{0}, \mathbf{I}_{d\times d})$, and the d-dimensional spherically symmetric uniform distribution $MU_d(\mathbf{0}, \mathbf{I}_{d\times d})$, respectively. The formula for the calculation is: $1-p = P(\|\boldsymbol{Z}\| > -F_1^{-1}(\alpha))$, where random vector $\boldsymbol{Z}$ follows a spherically symmetric distribution centered at $\mathbf{0}$, $\|\cdot\|$ denotes the Euclidean norm, and F_1^{-1} is its univariate marginal quantile function. The details for the formula can be seen from the proof of Corollary 1 and the following remark.

Table 1. The theoretical trimming proportions $1-p$ for the $\mu_F(\alpha)$ based on the half-space depth with $\alpha = 0.01$ when the underlying distributions are $N_d(\mathbf{0}, \mathbf{I}_{d\times d})$, $Mt_d(5, \mathbf{0}, \mathbf{I}_{d\times d})$, and $MU_d(0, \mathbf{I}_{d\times d})$, respectively.

$1-p$ \ d	2	3	4	5	6	7	8
$N_d(0, \mathbf{I}_{d\times d})$	0.0668	0.1440	0.2476	0.3677	0.4922	0.6098	0.7128
$Mt_d(5, 0, \mathbf{I}_{d\times d})$	0.0519	0.0934	0.1420	0.1953	0.2513	0.3084	0.3651
$MU_d(0, \mathbf{I}_{d\times d})$	0.1270	0.3134	0.5188	0.6948	0.8223	0.9040	0.9515

From the table, we see that the theoretical trimming proportion $1-p$ varies enormously from distribution to distribution. It is even harder to choose the α value for many practical problems, in which the underlying distribution is unknown. Therefore, it is necessary to restudy multivariate trimmed means. As for multivariate trimmed scatter matrices, we have not seen any investigation in the literature.

In this paper, we propose and study some new general depth-based trimmed means and trimmed scatter matrices. Their definitions, basic properties, and algorithms are given in Section 2. The asymptotic distributions of their sample versions are established in Section 3. Using the asymptotic distributions, we study the asymptotic efficiencies of the sample trimmed means and the sample trimmed scatter matrices in Section 4. The results show that both the trimmed means and trimmed scatter matrices are highly efficient. Section 5 is devoted to the robustness of the sample trimmed means and the sample trimmed scatter matrices, which is explored through finite-sample breakdown point. The proofs of main results are reserved for the **Appendix**.

Throughout this paper, matrices are denoted by bold letters and vectors by bold italicized letters. We use uppercase letters to denote distribution functions and their lowercase counterparts to denote density functions if they exist. For example, we denote by $F_{\boldsymbol{X}}$ and $f_{\boldsymbol{X}}$ the *cdf* and density of a random vector $\boldsymbol{X}$ in $\mathbb{R}^d$. When X is a random variable, the quantile function of X is denoted by F_X^{-1}. Without confusion, we will omit the subscript. The indicator function of a set A is denoted by $I_A(\cdot)$ and the transpose of a matrix or a vector $\boldsymbol{v}$ is denoted by $\boldsymbol{v}'$. For an $m\times k$ matrix $\mathbf{M}$, $vec(\mathbf{M})$ denotes the mk-dimensional vector formed by stacking the columns of $\mathbf{M}$, and $\mathbf{B}\otimes\mathbf{C}$ denotes the Kronecker product of matrices $\mathbf{B}$ and $\mathbf{C}$.

2. Definitions, Basic Properties, and Algorithms

Let $w_1(t)$ and $w_2(t)$ be suitable weight functions on $[0,\alpha_F^*]$, where $\alpha_F^* = \sup_{\boldsymbol{x}\in\mathbb{R}^d}(D_F(\boldsymbol{x}))$. To reduce or eliminate the impacts of outliers, we assume that $w_1(t)$ and $w_2(t)$ are nondecreasing. Given p_1 and p_2 in $(0,1]$, we define the general depth-based trimmed mean of trimming proportion $1-p_1$ as

$$\boldsymbol{\mu}_F(p_1) = \int_{C_F(p_1)} \boldsymbol{x} w_1(D_F(\boldsymbol{x}))dF(\boldsymbol{x}) \Big/ \int_{C_F(p_1)} w_1(D_F(\boldsymbol{x}))dF(\boldsymbol{x})$$

and define the general depth-based trimmed scatter matrix of trimming proportion $1-p_2$ using the location measure $\boldsymbol{\mu}_F(p_1)$ as

$$\boldsymbol{\Sigma}_F(p_1,p_2) = \frac{\int_{C_F(p_2)} (\boldsymbol{x}-\boldsymbol{\mu}_F(p_1))(\boldsymbol{x}-\boldsymbol{\mu}_F(p_1))' w_2(D_F(\boldsymbol{x}))dF(\boldsymbol{x})}{\int_{C_F(p_2)} w_2(D_F(\boldsymbol{x}))dF(\boldsymbol{x})}.$$

As a special case, $\boldsymbol{\mu}_F(1) = L(F)$ and $\boldsymbol{\Sigma}_F(1,1) = S(F)$. Furthermore, $\boldsymbol{\mu}_F(1) = E(\boldsymbol{X})$ if $w_1(t)=1$ on $[0,\alpha_F^*]$, and $\boldsymbol{\Sigma}_F(1,1) = Cov(\boldsymbol{X})$ if $w_1(t)=w_2(t)=1$ on $[0,\alpha_F^*]$. In general, if $p_1\neq 1$ and $p_2\neq 1$, $\boldsymbol{\mu}_F(p_1)$ and $\boldsymbol{\Sigma}_F(p_1,p_2)$ are much more broadly defined than $E(\boldsymbol{X})$ and $Cov(\boldsymbol{X})$, respectively. They do not require existence of any moments.

Remark. Unlike $L(F)$ and $S(F)$, we separate weighting and trimming in $\boldsymbol{\mu}_F(p_1)$ and $\boldsymbol{\Sigma}_F(p_1,p_2)$ with trimming controlled by $C_F(p_1)$ and $C_F(p_2)$, and thus weight functions are not affected by trimming. For example, the weight function for the usual trimmed mean can be $w_1(t)=1$ on $[0,\alpha_F^*]$, which is continuously differentiable.

For simplicity, we will confine attention to the case $w_1(t)=w_2(t):=w(t)$ and $p_1=p_2:=p$, and call $\boldsymbol{\mu}_F(p)$ and $\boldsymbol{\Sigma}_F(p)$ the $1-p$ proportion or $100(1-p)\%$ general depth-based trimmed mean and trimmed scatter matrix, respectively. Given a sample $\boldsymbol{X}^n=\{\boldsymbol{X}_1,\cdots,\boldsymbol{X}_n\}$, the sample versions of $D_F(\boldsymbol{x})$, $I_F(\alpha)$, $C_F(p)$, $r_F(\alpha,\boldsymbol{u})$, $\boldsymbol{\mu}_F(p)$, and $\boldsymbol{\Sigma}_F(p)$ are obtained by replacing F with F_n, the empirical distribution function of $\boldsymbol{X}^n$, and are denoted by

$D_n(\boldsymbol{x})$, $I_n(\alpha)$, $C_n(p)$, $r_n(\alpha,\boldsymbol{u})$, $\boldsymbol{\mu}_n(p)$, and $\boldsymbol{\Sigma}_n(p)$, respectively. For convenience, we will suppress F in $D_F(\boldsymbol{x})$, $I_F(\alpha)$, $C_F(p)$, and $r_F(\alpha,\boldsymbol{u})$ in the following discussions unless it is necessary for clarity. $\boldsymbol{\mu}_F(p)$, $\boldsymbol{\Sigma}_F(p)$, and their sample versions have the following basic properties. The proofs are relatively straightforward and thus omitted.

Theorem 1. *Suppose that the depth function* $D_F(\boldsymbol{x})$ *is affine invariant. Then we have the following results:*

1) $\boldsymbol{\mu}_F(p)$ and $\boldsymbol{\Sigma}_F(p)$ are affine equivariant, that is, $\boldsymbol{\mu}_{F_{\mathbf{A}\boldsymbol{X}+\boldsymbol{b}}}(p) = \mathbf{A}\boldsymbol{\mu}_{F_{\boldsymbol{X}}}(p)+\boldsymbol{b}$ and $\boldsymbol{\Sigma}_{F_{\mathbf{A}\boldsymbol{X}+\boldsymbol{b}}}(p) = \mathbf{A}\boldsymbol{\Sigma}_{F_{\boldsymbol{X}}}(p)\mathbf{A}'$ for any nonsingular $d\times d$ matrix $\mathbf{A}$ and d-vector $\boldsymbol{b}$, and so are their sample versions.

2) If F is centrally symmetric about $\boldsymbol{\theta}\in\mathbb{R}^d$ and has the first moment, then for any $p\in(0,1]$, $\boldsymbol{\mu}_F(p)=\boldsymbol{\theta}$ (*i.e.*, $\boldsymbol{\mu}_F(p)$ is Fisher consistent for $\boldsymbol{\theta}$) and $E(\boldsymbol{\mu}_n(p))=\boldsymbol{\theta}$.

3) If $\boldsymbol{X}$ follows an elliptically symmetric distribution F and has the second moment, then $\boldsymbol{\Sigma}_F(p)=kCov(\boldsymbol{X})$ and $E(\boldsymbol{\Sigma}_n(p))=k_nCov(\boldsymbol{X})$, where k and k_n are some positive constants.

Computation of $\boldsymbol{\mu}_n(p)$ and $\boldsymbol{\Sigma}_n(p)$ is quite easy. The following are the algorithms by R.

1) Compute the depths of all points in the sample. Several R packages are available for depth-based statistical analysis. For the halfspace depth, the projection depth, the L^p depth, and the Mahalanobis depth, one can use the command "depth" in the R package "DepthProc" of Zawadzki, Kosiorowski, Slomczynski, Bocian, and Wegrzynkiewicz [23].

2) Order the data points according to their depths decreasingly, denoted by $\boldsymbol{X}_{(1)},\cdots,\boldsymbol{X}_{(n)}$, and get the points in $C_n(p)$, $\boldsymbol{X}_{(1)},\cdots,\boldsymbol{X}_{(\lceil np\rceil)}$.

3) Compute $\boldsymbol{\mu}_n(p)$ as $\boldsymbol{\mu}_n(p)=\sum_{i=1}^{\lceil np\rceil}\boldsymbol{X}_{(i)}w(D_n(\boldsymbol{X}_{(i)}))/\sum_{i=1}^{\lceil np\rceil}w(D_n(\boldsymbol{X}_{(i)}))$ and compute $\boldsymbol{\Sigma}_n(p)$ as

$$\boldsymbol{\Sigma}_n(p)=\sum_{i=1}^{\lceil np\rceil}(\boldsymbol{X}_{(i)}-\boldsymbol{\mu}_n(p))(\boldsymbol{X}_{(i)}-\boldsymbol{\mu}_n(p))'w(D_n(\boldsymbol{X}_{(i)}))/\sum_{i=1}^{\lceil np\rceil}w(D_n(\boldsymbol{X}_{(i)})).$$

3. Asymptotic Distributions

To use an estimator or a statistic for inferences, its distribution or asymptotic distribution is necessary. To establish the asymptotic distributions of $\boldsymbol{\mu}_n(p)$ and $\boldsymbol{\Sigma}_n(p)$, we invoke the following conditions.

(C_1) $D(\boldsymbol{x})$ is affine invariant, and $n^{1/2}(D_n(\boldsymbol{x})-D(\boldsymbol{x}))=\int\varphi(\boldsymbol{x},\boldsymbol{y})dv_n(\boldsymbol{y})+o_p(1)$ uniformly in $\boldsymbol{x}$, where $v_n=n^{1/2}(F_n-F)$;

Then, by Theorem 1 1), both $\boldsymbol{\mu}_n(p)-\boldsymbol{\mu}_F(p)$ and $\boldsymbol{\Sigma}_n(p)-\boldsymbol{\Sigma}_F(p)$ are invariant for any translation. Thus, we can assume without loss of generality that the origin is the deepest point.

(C_2) $r(\alpha,\boldsymbol{u})$ has a continuous partial derivative with respect to α, and $n^{1/2}(r_n(\alpha,\boldsymbol{u})-r(\alpha,\boldsymbol{u}))=\int\psi(\boldsymbol{x},\alpha,\boldsymbol{u})dv_n(\boldsymbol{x})+o_p(1)$ with $\psi(\boldsymbol{x},\alpha,\boldsymbol{u})$ being continuous in α;

(C_3) There exists a Donsker class $\mathcal{D}$ of sets such that $I(\alpha)\in\mathcal{D}$ and $I_n(\alpha)\in\mathcal{D}$ almost surely for any n and $\alpha\in(0,\alpha_F^*)$;

(C_4) F has a positive continuous density f in some neighborhood of $\partial C(p)$;

(C_5) $w^{(1)}(t)=\frac{d}{dt}w(t)$ is continuous and $w(t)>0$ for $t\ge F_D^{-1}(1-p)$ with $p\in(0,1)$.

Remark. Uniform weak convergence of $n^{1/2}(D_n(\boldsymbol{x})-D_F(\boldsymbol{x}))$ has been established for the simplicial depth by Dümbgen [15] and Arcones and Giné [24], for the halfspace depth by Massé [16], for the projection depth and some generalized halfspace depth by Arcones, Cui, and Zuo [25], and thus (C_1) is satisfied for these depth functions. For a fixed $\alpha\in(0,\alpha_F^*)$, the asymptotic representation of $n^{1/2}(r_n(\alpha,\boldsymbol{u})-r_F(\alpha,\boldsymbol{u}))$ was derived for the halfspace depth by Nolan [26] and for the projection depth by Zuo [19]. Considering α there as a variable, we see that (C_2) is satisfied by these depth functions. As for Condition (C_3), we know that $I_n(\alpha)$ is an ellipsoid for the Mahalanobis depth, a convex set formed by halfspaces for the halfspace depth and projection depth, and a union of some convex sets formed by halfspaces for the simplicial depth. Since {all ellipsoids in $\mathbb{R}^d$} is a Vapnik-Červonenkis (VC) class [27] and so is {all sets in $\mathbb{R}^d$ formed by halfspaces} [24] [28] [29]. Condition (C_3) is satisfied for these depths.

Remark. It can be shown that every Donsker class is a Glivenko-Cantelli class almost surely [27].

Define $\Delta_F(p)=\dfrac{\int_{C(p)}\boldsymbol{x}\boldsymbol{x}'w(D(\boldsymbol{x}))dF(\boldsymbol{x})}{\int_{C(p)}w(D(\boldsymbol{x}))dF(\boldsymbol{x})}$. Then we have the following results.

Theorem 2. Under the conditions (C_1) - (C_5), for any $p\in(0,1)$

1) $\boldsymbol{\mu}_n(p)-\boldsymbol{\mu}_F(p)=\frac{1}{n}\sum_{i=1}^{n}l(\boldsymbol{X}_i)-E(l(\boldsymbol{X}_i))+o_p(1/\sqrt{n})$, and thus

$$\sqrt{n}(\boldsymbol{\mu}_n(p)-\boldsymbol{\mu}_F(p))\xrightarrow{d}N_d(\boldsymbol{0},Cov(l(\boldsymbol{X})),$$

where $l(\boldsymbol{x})=(l_1(\boldsymbol{x})+l_2(\boldsymbol{x})+l_3(\boldsymbol{x})+l_4(\boldsymbol{x}))/l_0(p)$ with

$$l_0(p)=\int_{C(p)}w(D(\boldsymbol{x}))dF(\boldsymbol{x}),$$

$$l_1(\boldsymbol{x})=(\boldsymbol{x}-\boldsymbol{\mu}_F(p))w(D(\boldsymbol{x}))I_{C(p)}(\boldsymbol{x}),$$

$$l_2(\boldsymbol{x})=\int_{C(p)}(\boldsymbol{y}-\boldsymbol{\mu}_F(p))w^{(1)}(D(\boldsymbol{y}))\varphi(\boldsymbol{y},\boldsymbol{x})dF(\boldsymbol{y}),\text{ where } w^{(1)}(t)=\frac{d}{dt}w(t),$$

$$\begin{aligned}l_3(\boldsymbol{x})=\int_{S^{d-1}}&(r(F_D^{-1}(1-p),\boldsymbol{u})\boldsymbol{u}-\boldsymbol{\mu}_F(p))w(D(r(F_D^{-1}(1-p),\boldsymbol{u})\boldsymbol{u}))\\&f(r(F_D^{-1}(1-p),\boldsymbol{u})\boldsymbol{u})\,|J(r(F_D^{-1}(1-p),\boldsymbol{u}),\boldsymbol{u})|\,\psi(\boldsymbol{x},F_D^{-1}(1-p),\boldsymbol{u})d\boldsymbol{u},\end{aligned}$$

where $S^{d-1}=\{\boldsymbol{u}\in\mathbb{R}^d:\|\boldsymbol{u}\|=1\}$ and $J(r,\boldsymbol{u})$ is the Jacobian of the transformation $\boldsymbol{x}=r\boldsymbol{u}$ with $r=\|\boldsymbol{x}\|$ and $\boldsymbol{u}=\dfrac{\boldsymbol{x}}{\|\boldsymbol{x}\|}$, and

$$\begin{aligned}l_4(\boldsymbol{x})=-l_0(p)\boldsymbol{\mu}_F^{(1)}(p)[I_{C(p)}(\boldsymbol{x})+\int_{S^{d-1}}&f(r(F_D^{-1}(1-p),\boldsymbol{u})\boldsymbol{u})\\&|J(r(F_D^{-1}(1-p),\boldsymbol{u}),\boldsymbol{u})|\,\psi(\boldsymbol{x},F_D^{-1}(1-p),\boldsymbol{u})d\boldsymbol{u}],\end{aligned}$$

where $\boldsymbol{\mu}_F^{(1)}(p) = \frac{d}{dp}\boldsymbol{\mu}_F(p)$.

2) $\boldsymbol{\Sigma}_n(p) - \boldsymbol{\Sigma}_F(p) = \frac{1}{n}\sum_{i=1}^n k(\boldsymbol{X}_i) - E(k(\boldsymbol{X}_i)) + o_p(1/\sqrt{n})$, and thus

$$\sqrt{n}(vec(\boldsymbol{\Sigma}_n(p)) - vec(\boldsymbol{\Sigma}_F(p))) \xrightarrow{d} N_{d^2}(\boldsymbol{0}, Cov(vec(k(\boldsymbol{X})))),$$

where $k(\boldsymbol{x}) = (k_1(\boldsymbol{x}) + k_2(\boldsymbol{x}) + k_3(\boldsymbol{x}) + k_4(\boldsymbol{x}))/l_0(p) - \boldsymbol{\mu}_F(p)l'(\boldsymbol{x}) - l(\boldsymbol{x})\boldsymbol{\mu}_F'(p)$ with

$$k_1(\boldsymbol{x}) = (\boldsymbol{x}\boldsymbol{x}' - \Delta_F(p))w(D(\boldsymbol{x}))I_{C(p)}(\boldsymbol{x}),$$

$$k_2(\boldsymbol{x}) = \int_{C(p)}(\boldsymbol{y}\boldsymbol{y}' - \Delta_F(p))w^{(1)}(D(\boldsymbol{y}))\varphi(\boldsymbol{y}, \boldsymbol{x})dF(\boldsymbol{y}),$$

$$k_3(\boldsymbol{x}) = \int_{S^{d-1}}(r^2(F_D^{-1}(1-p), \boldsymbol{u})\boldsymbol{u}\boldsymbol{u}' - \Delta_F(p))w(D(r(F_D^{-1}(1-p), \boldsymbol{u})\boldsymbol{u}))$$
$$f(r(F_D^{-1}(1-p), \boldsymbol{u})\boldsymbol{u}) \,|\, J(r(F_D^{-1}(1-p), \boldsymbol{u}), \boldsymbol{u}) \,|\, \psi(\boldsymbol{x}, F_D^{-1}(1-p), \boldsymbol{u})d\boldsymbol{u},$$

$$k_4(\mathrm{x}) = -l_0(p)\Delta_F^{(1)}(p)[I_{C(p)}(\boldsymbol{x}) + \int_{S^{d-1}} f(r(F_D^{-1}(1-p), \boldsymbol{u})\boldsymbol{u})$$
$$|\, J(r(F_D^{-1}(1-p), \boldsymbol{u}), \boldsymbol{u}) \,|\, \psi(\boldsymbol{x}, F_D^{-1}(1-p), \boldsymbol{u})d\boldsymbol{u}],$$

where $\Delta_F^{(1)}(p) = \frac{d}{dp}\Delta_F(p)$.

Both $Cov(l(\boldsymbol{X}))$ and $Cov(vec(k(\boldsymbol{X}))$ are quite complicated in general. However, the results are relatively simple when $w(t) = 1$ on $[0, \alpha_F^*]$ and F is an elliptically symmetric distribution. As an example, we concretize $Cov(l(\boldsymbol{X}))$ and $Cov(vec(k(\boldsymbol{X}))$ for the halfspace depth-based trimmed means and scatter matrices for the very important special case. A continuous random vector $\boldsymbol{X}$ in $\mathbb{R}^d$ is said to have an *elliptically symmetric* distribution, denoted by $E_d(h; \boldsymbol{\theta}, \Lambda)$, if it has a density of the form

$$f(\boldsymbol{x}) = |\Lambda|^{-1/2} h((\boldsymbol{x} - \boldsymbol{\theta})'\Lambda^{-1}(\boldsymbol{x} - \boldsymbol{\theta})), \boldsymbol{x} \in \mathbb{R}^d$$

for a nonnegative function $h(\cdot)$ with $\int_0^\infty r^{d/2-1}h(r)dr < \infty$ and a positive definite matrix Λ. If the first moment of $\boldsymbol{X}$ exists, $E(\boldsymbol{X}) = \boldsymbol{\theta}$. If the second moment of $\boldsymbol{X}$ exists, $Cov(\boldsymbol{X}) = [E(\|\Lambda^{-1/2}(\boldsymbol{X} - \boldsymbol{\theta})\|^2)/d]\Lambda$, where $\|\cdot\|$ denotes the Euclidean norm. When $\Lambda = \mathbf{I}_{d\times d}$, the $d \times d$ identity matrix, $\boldsymbol{X}$ is said to have a *spherically symmetric* distribution with center $\boldsymbol{\theta}$. If $\boldsymbol{X} \sim E_d(h; \boldsymbol{\theta}, \Lambda)$, $\boldsymbol{Z} = \Lambda^{-1/2}(\boldsymbol{X} - \boldsymbol{\theta}) \sim E_d(h; \boldsymbol{0}, \mathbf{I}_{d\times d})$. Let $\boldsymbol{U} = \frac{\boldsymbol{Z}}{\|\boldsymbol{Z}\|}$. Then $\boldsymbol{U}$ has a uniform distribution on $S^{d-1} = \{\boldsymbol{u} \in \mathbb{R}^d : \|\boldsymbol{u}\| = 1\}$, and thus $E(\boldsymbol{U}) = \boldsymbol{0}$ and $Cov(\boldsymbol{U}) = E(\boldsymbol{U}\boldsymbol{U}') = \frac{1}{d}\mathbf{I}_{d\times d}$. In addition, $\boldsymbol{U}$ and $\|\boldsymbol{Z}\|$ are independent. See Fang, Kotz, and Ng [30] for broad discussion about elliptically symmetric distributions.

Because the trimming based on the halfspace depth is a natural extension of the usual univariate trimming, we choose the halfspace depth here. The halfspace depth is first introduced by Tukey [7] and is defined as

$$HD_F(\boldsymbol{x}) = \inf\{P(H) : \boldsymbol{x} \in H \in \mathcal{H}\}, \boldsymbol{x} \in \mathbb{R}^d,$$

where $\mathcal{H}$ = {all closed halfspaces}. As a leading depth function, the halfspace depth has been widely used. The finite-sample breakdown point of the halfspace depth was obtained by Donoho and Gasko [20]. Romanazzi [31] derived the influence function of $HD_F(\boldsymbol{x})$. The asymptotic distribution of its sample version $HD_n(\boldsymbol{x})$ was established by Massé [16]. Nolan [26] studied the asymptotics of the sample α halfspace depth inner regions. An important feature of the trimming based on the halfspace depth is that it trims off the same proportion of data points in a closed halfspace at each direction normal to the boundary of $C_n(p)$ and outward with respect to $C_n(p)$.

Corollary 1. Suppose that $\boldsymbol{X} \sim E_d(h;\boldsymbol{\theta},\Lambda)$ with a positive continuous density f in a small neighborhood of $\partial C(p)$ and $w(t)=1$ on $[0,\alpha_F^*]$. Let f_1 and F_1 be the univariate marginal density and the marginal distribution function of $\boldsymbol{Z}=\boldsymbol{\Lambda}^{-1/2}(\boldsymbol{X}-\boldsymbol{\theta})$. If the trimming is based on the halfspace depth, then for any $p\in(0,1)$ and $d\ge 2$.

1) $\sqrt{n}(\boldsymbol{\mu}_n(p)-\boldsymbol{\mu}_F(p)) \xrightarrow{d} N_d(\mathbf{0},\mathbf{V}_1)$,

where $\mathbf{V}_1 = E[c_1^2(\|\boldsymbol{Z}\|)+c_2^2(\|\boldsymbol{Z}\|)]\dfrac{\boldsymbol{\Lambda}}{d}$ with

$c_1(\|\boldsymbol{z}\|) = \dfrac{1}{p}\|\boldsymbol{z}\| I_{\{\|z\|\le F_{\|\boldsymbol{Z}\|}^{-1}(p)\}}(\boldsymbol{z})$, and

$$c_2(\|\boldsymbol{z}\|) = \frac{F_{\|\boldsymbol{Z}\|}^{-1}(p) f_{\|\boldsymbol{Z}\|}(F_{\|\boldsymbol{Z}\|}^{-1}(p))}{(d-1)pf_1(F_{\|\boldsymbol{Z}\|}^{-1}(p))\int_0^\pi \sin^{d-2}\theta_1 d\theta_1}[1-(\frac{F_{\|\boldsymbol{Z}\|}^{-1}(p)}{\|\boldsymbol{z}\|})^2]^{(d-1)/2} I_{\{\|z\|>F_{\|\boldsymbol{Z}\|}^{-1}(p)\}}(\boldsymbol{z}).$$

2) $\sqrt{n}(vec(\boldsymbol{\Sigma}_n(p))-vec(\boldsymbol{\Sigma}_F(p))) \xrightarrow{d} N_{d^2}(\mathbf{0},\mathbf{V}_2)$,

where $\mathbf{V}_2 = \sigma_1(\mathbf{I}_{d^2\times d^2}+\mathbf{K}_{d,d})(\boldsymbol{\Lambda}\otimes\boldsymbol{\Lambda})+\sigma_2 vec(\boldsymbol{\Lambda})vec(\boldsymbol{\Lambda})'$ with

$$\sigma_1 = E(c_3^2(\|\boldsymbol{Z}\|))/(d(d+2)),$$

$$\sigma_2 = \sigma_1 + 2E(c_3(\|\boldsymbol{Z}\|)c_4(\|\boldsymbol{Z}\|))/d + E(c_4^2(\|\boldsymbol{Z}\|)) - [E(c_3(\|\boldsymbol{Z}\|))/d + E(c_4(\|\boldsymbol{Z}\|))]^2,$$

$$c_3(\|\boldsymbol{z}\|) = \frac{1}{p}\|\boldsymbol{z}\|^2 I_{\{\|z\|\le F_{\|\boldsymbol{Z}\|}^{-1}(p)\}}(z) + \frac{[F_{\|\boldsymbol{Z}\|}^{-1}(p)]^2 f_{\|\boldsymbol{Z}\|}(F_{\|\boldsymbol{Z}\|}^{-1}(p))}{dpf_1(F_{\|\boldsymbol{Z}\|}^{-1}(p))\int_0^\pi \sin^d\theta_1 d\theta_1}$$
$$[1-(\frac{F_{\|\boldsymbol{Z}\|}^{-1}(p)}{\|\boldsymbol{z}\|})^2]^{(d-1)/2}(\frac{F_{\|\boldsymbol{Z}\|}^{-1}(p)}{\|\boldsymbol{z}\|}) I_{\{\|z\|>F_{\|\boldsymbol{Z}\|}^{-1}(p)\}}(z),$$

$$c_4(\|\boldsymbol{z}\|) = \frac{[F_{\|\boldsymbol{Z}\|}^{-1}(p)]^2 f_{\|\boldsymbol{Z}\|}(F_{\|\boldsymbol{Z}\|}^{-1}(p))}{dpf_1(F_{\|\boldsymbol{Z}\|}^{-1}(p))}[\frac{\int_0^{\arccos(F_{\|\boldsymbol{Z}\|}^{-1}(p)/\|z\|)}\sin^d\theta_1 d\theta_1}{\int_0^\pi \sin^d\theta_1 d\theta_1}$$
$$-\frac{\int_0^{\arccos(F_{\|\boldsymbol{Z}\|}^{-1}(p)/\|z\|)}\sin^{d-2}\theta_1 d\theta_1}{\int_0^\pi \sin^{d-2}\theta_1 d\theta_1}] I_{\{\|z\|>F_{\|\boldsymbol{Z}\|}^{-1}(p)\}}(\boldsymbol{z})$$
$$-\frac{1}{dp}[F_{\|\boldsymbol{Z}\|}^{-1}(p)]^2 I_{\{\|z\|\le F_{\|\boldsymbol{Z}\|}^{-1}(p)\}}(z),$$

and

$\mathbf{K}_{d,d}$ being a $(d^2\times d^2)$-block matrix with (i,j)-block being δ_{ji}, which is a

$d \times d$ matrix with 1 at entry (j,i) and 0 elsewhere.

4. Asymptotic Relative Efficiencies

When we evaluate the performance of an estimator, two important criteria are efficiency and robustness. Are $\boldsymbol{\mu}_n(p)$ and $\Sigma_n(p)$ efficient? We will answer this question in this section. It is well-known that the sample mean vector $\bar{\boldsymbol{X}}_n$ and the sample covariance matrix $\mathbf{S}_n$ are the most efficient unbiased estimators for the population mean vector $\boldsymbol{\mu}$ and covariance matrix $\boldsymbol{\Sigma}$, respectively, when the population distribution is $N_d(\boldsymbol{\mu},\Sigma)$. It is natural to compare $\boldsymbol{\mu}_n(p)$ with $\bar{\boldsymbol{X}}_n$ and $\Sigma_n(p)$ with $\mathbf{S}_n$ for efficiency.

4.1. Asymptotic Relative Efficiencies of the Sample Trimmed Means

Given the asymptotic distribution of $\boldsymbol{\mu}_n(p)$, it is straightforward to compute its asymptotic efficiency. For comparison, we compute the asymptotic relative efficiency (ARE) of the $\boldsymbol{\mu}_n(p)$ based on the halfspace depth with $w(t)=1$ on $[0,\alpha_F^*]$ with respect to $\bar{\boldsymbol{X}}_n$ for some elliptically symmetric distributions $E_d(h;\boldsymbol{\theta},\Lambda)$,

$$ARE_F(\boldsymbol{\mu}_n(p),\bar{\boldsymbol{X}}_n)=\left(\frac{\det(ACov(\bar{\boldsymbol{X}}_n))}{\det(ACov(\boldsymbol{\mu}_n(p)))}\right)^{1/d},$$

where $ACov(\bar{\boldsymbol{X}}_n)$ and $ACov(\boldsymbol{\mu}_n(p))$ are the asymptotic covariance matrices of $\sqrt{n}(\bar{\boldsymbol{X}}_n-\boldsymbol{\mu})$ and $\sqrt{n}(\boldsymbol{\mu}_n(p)-\boldsymbol{\mu}_F(p))$, respectively. Noting that $ACov(\bar{\boldsymbol{X}}_n)=Cov(\boldsymbol{X})=[E(\|\boldsymbol{Z}\|^2)/d]\Lambda$ and $ACov(\boldsymbol{\mu}_n(p))=E[c_1^2(\|\boldsymbol{Z}\|)+c_2^2(\|\boldsymbol{Z}\|)]\dfrac{\Lambda}{d}$ by Corollary 1, where $\boldsymbol{Z}=\Lambda^{-1/2}(\boldsymbol{X}-\boldsymbol{\theta})$, we have

$$ARE_F(\boldsymbol{\mu}_n(p),\bar{\boldsymbol{X}}_n)=\frac{E(\|\boldsymbol{Z}\|^2)}{E[c_1^2(\|\boldsymbol{Z}\|)+c_2^2(\|\boldsymbol{Z}\|)]}.$$

When the population distribution F is $N_d(\boldsymbol{\mu},\Lambda)$, the results are reported in **Table 2**. It shows that $ARE_F(\boldsymbol{\mu}_n(p),\bar{\boldsymbol{X}}_n)$ decreases as the dimension d or the trimming proportion $1-p$ increases. Overall, $\boldsymbol{\mu}_n(p)$ is efficient. Zuo [19] computed the asymptotic relative efficiencies of the halfspace depth-induced median (HM) and the projection depth-induced median (PM) with respect to $\bar{\boldsymbol{X}}_n$ for $N_2(\mathbf{0},\mathbf{I}_{2\times 2})$, which are 0.76 and 0.77, respectively. Even if the trimming proportion $1-p=0.4$, the trimmed mean is much more efficient than HM and PM. **Table 3** lists the results for the d-dimensional t distribution with 5 degrees of freedom $Mt_d(5,\boldsymbol{\mu},\Lambda)$. Again, $ARE_F(\boldsymbol{\mu}_n(p),\bar{\boldsymbol{X}}_n)$ decreases as the dimension d increases for each trimming proportion $1-p$. The striking finding is that $ARE_F(\boldsymbol{\mu}_n(p),\bar{\boldsymbol{X}}_n)$ increases first and then decreases as the trimming proportion $1-p$ increases for each $d\le 6$. In addition, $\boldsymbol{\mu}_n(p)$ is more efficient than $\bar{\boldsymbol{X}}_n$ when $d\le 5$.

Table 2. $ARE_F\left(\boldsymbol{\mu}_n(p), \bar{\boldsymbol{X}}_n\right)$, where $\boldsymbol{\mu}_n(p)$ is based on the halfspace depth with $w(t)=1$ on $\left[0, \alpha_F^*\right]$, when the underlying distribution F is $N_d(\boldsymbol{\mu}, \boldsymbol{\Lambda})$.

$1-p$ \ d	2	3	4	5	6	7	8
0.01	0.9936	0.9855	0.9771	0.9679	0.9574	0.9449	0.9299
0.05	0.9784	0.9568	0.9357	0.9144	0.8917	0.8669	0.8392
0.10	0.9628	0.9321	0.9032	0.8753	0.8471	0.8176	0.7861
0.15	0.9479	0.9111	0.8772	0.8455	0.8146	0.7832	0.7506
0.20	0.9331	0.8915	0.8538	0.8196	0.7871	0.7549	0.7223
0.25	0.9182	0.8725	0.8317	0.7955	0.7620	0.7296	0.6973
0.30	0.9029	0.8536	0.8100	0.7723	0.7380	0.7057	0.6740
0.35	0.8872	0.8344	0.7883	0.7492	0.7144	0.6822	0.6513
0.40	0.8710	0.8147	0.7662	0.7257	0.6905	0.6585	0.6284

Table 3. $ARE_F\left(\boldsymbol{\mu}_n(p), \bar{\boldsymbol{X}}_n\right)$, where $\boldsymbol{\mu}_n(p)$ is based on the halfspace depth with $w(t)=1$ on $\left[0, \alpha_F^*\right]$, when the underlying distribution F is $Mt_d(5, \boldsymbol{\mu}, \boldsymbol{\Lambda})$.

$1-p$ \ d	2	3	4	5	6	7	8
0.01	1.0525	1.0377	1.0244	1.0121	1.0002	0.9886	0.9771
0.05	1.1172	1.0845	1.0564	1.0313	1.0080	0.9858	0.9645
0.10	1.1569	1.1127	1.0756	1.0433	1.0139	0.9865	0.9606
0.15	1.1807	1.1288	1.0859	1.0490	1.0161	0.9859	0.9577
0.20	1.1955	1.1378	1.0905	1.0506	1.0154	0.9834	0.9539
0.25	1.2042	1.1418	1.0910	1.0486	1.0118	0.9788	0.9485
0.30	1.2081	1.1417	1.0879	1.0436	1.0056	0.9718	0.9411
0.35	1.2082	1.1381	1.0816	1.0356	0.9967	0.9625	0.9316
0.40	1.2049	1.1312	1.0721	1.0247	0.9850	0.9505	0.9196

4.2. Asymptotic Relative Efficiencies of the Sample Trimmed Scatter Matrices

For the asymptotic efficiency of $\Sigma_n(p)$, we will focus on how well it estimates some *shape component* of $\boldsymbol{\Sigma}_F(p)$. Generally speaking, any function $H(\mathbf{V})$ of a scatter matrix $\mathbf{V}$ can be considered as a shape component of $\mathbf{V}$ if $H(\mathbf{V})$ is invariant for a positive scalar multiple, *i.e.*, $H(\mathbf{V}) = H(\lambda\mathbf{V})$ for any $\lambda > 0$. See Tyler [32] and Kent and Tyler [6] for detailed discussions. For efficiency studies, the shape component of most interest is the nonsphericity of $\mathbf{V}$ and the measure most used for nonsphericity is the function φ_0 originating from the likelihood ratio test statistic for nonsphericity,

$$\varphi_0(\mathbf{V}) = \det(\mathbf{V}) / [\mathrm{trace}(\mathbf{V}) / d]^d,$$

which was first derived for multivariate normal populations with $\mathbf{V} = \mathbf{S}_n$ by Mauchly [33]. Muirhead [34] showed that if F is an elliptically symmetric distribution $E_d(h;\boldsymbol{\theta},\boldsymbol{\Lambda})$ with marginal kurtosis 3κ and the null hypothesis $H_0 : \boldsymbol{\Lambda} = \lambda \mathbf{I}_{d\times d}$ is true,

$$-n\log(\varphi_0(\mathbf{S}_n)) \xrightarrow{d} (1+\kappa)\chi^2_{(d-1)(d+2)/2}.$$

For $\boldsymbol{\Sigma}_n(p)$, we have the following result.

Theorem 3. Suppose that the conditions of Corollary 1 hold and the trimming is based on the halfspace depth. Then, if $\boldsymbol{\Lambda} = \lambda \mathbf{I}_{d\times d}$ and $d \ge 2$,

$$-n\log(\varphi_0(\boldsymbol{\Sigma}_n(p))) \xrightarrow{d} \frac{\sigma_1}{c^2}\chi^2_{(d-1)(d+2)/2},$$

where σ_1 is given in Corollary 1 and $c = \frac{1}{pd} E[\|\boldsymbol{Z}\|^2 I_{\{\|\boldsymbol{Z}\| \le F^{-1}_{\|\boldsymbol{Z}\|}(p)\}}(\boldsymbol{Z})]$.

Based on this result and the result of Muirhead [34], the asymptotic relative efficiency of the $\boldsymbol{\Sigma}_n(p)$ based on the halfspace depth with $w(t) = 1$ on $[0, \alpha_F^*]$ with respect to $\mathbf{S}_n$ is given as

$$ARE_F(\boldsymbol{\Sigma}_n(p), \mathbf{S}_n) = \frac{1+\kappa}{\sigma_1 / c^2} = \frac{(1+\kappa)c^2}{\sigma_1}.$$

We compute $ARE_F(\boldsymbol{\Sigma}_n(p), \mathbf{S}_n)$ for $N_d(\mathbf{0}, \mathbf{I}_{d\times d})$ and $Mt_d(5, \mathbf{0}, \mathbf{I}_{d\times d})$, which are listed in **Table 4** and **Table 5**, respectively. When the population distribution F is $N_d(\mathbf{0}, \mathbf{I}_{d\times d})$, $ARE_F(\boldsymbol{\Sigma}_n(p), \mathbf{S}_n)$ decreases as the trimming proportion $1-p$ increases for each dimension d. However, it increases first and then decreases as dimension d increases for each trimming proportion $1-p \ge 0.10$. Overall, $\boldsymbol{\Sigma}_n(p)$ is efficient.

Table 4. $ARE_F(\boldsymbol{\Sigma}_n(p), \mathbf{S}_n)$, where $\boldsymbol{\Sigma}_n(p)$ is based on the halfspace depth with $w(t) = 1$ on $[0, \alpha_F^*]$, when the underlying distribution F is $N_d(0, \mathbf{I}_{d\times d})$.

$1-p$ \ d	2	3	4	5	6	7	8
0.01	0.9779	0.9711	0.9618	0.9513	0.9396	0.9260	0.9101
0.05	0.9148	0.9121	0.8994	0.8836	0.8657	0.8456	0.8231
0.10	0.8473	0.8537	0.8434	0.8281	0.8102	0.7901	0.7678
0.15	0.7860	0.8009	0.7939	0.7805	0.7640	0.7454	0.7248
0.20	0.7289	0.7510	0.7474	0.7361	0.7215	0.7047	0.6861
0.25	0.6748	0.7031	0.7025	0.6933	0.6806	0.6657	0.6492
0.30	0.6233	0.6566	0.6586	0.6514	0.6404	0.6274	0.6128
0.35	0.5739	0.6110	0.6153	0.6098	0.6005	0.5892	0.5765
0.40	0.5263	0.5661	0.5722	0.5682	0.5604	0.5507	0.5397

Table 5. $ARE_F(\mathbf{\Sigma}_n(p),\mathbf{S}_n)$, where $\mathbf{\Sigma}_n(p)$ is based on the halfspace depth with $w(t)=1$ on $[0,\alpha_F^*]$, when the underlying distribution F is $Mt_d(5,0,\mathbf{I}_{d\times d})$.

$1-p$ \ d	2	3	4	5	6	7	8
0.01	1.9313	1.8812	1.8437	1.8155	1.7934	1.7755	1.7606
0.05	2.1720	2.1344	2.0944	2.0609	2.0332	2.0102	1.9907
0.10	2.1722	2.1621	2.1307	2.1001	2.0734	2.0505	2.0307
0.15	2.1016	2.1176	2.0961	2.0701	2.0458	2.0243	2.0054
0.20	2.0039	2.0431	2.0309	2.0097	1.9882	1.9685	1.9510
0.25	1.8933	1.9523	1.9487	1.9319	1.9132	1.8954	1.8792
0.30	1.7760	1.8516	1.8554	1.8427	1.8265	1.8105	1.7958
0.35	1.6551	1.7443	1.7543	1.7451	1.7312	1.7170	1.7035
0.40	1.5325	1.6321	1.6472	1.6410	1.6292	1.6164	1.6042

When the population distribution F is $Mt_d(5,\mathbf{0},\mathbf{I}_{d\times d})$, $ARE_F(\mathbf{\Sigma}_n(p),\mathbf{S}_n)$ increases first and then decreases as trimming proportion $1-p$ increases for each dimension d. Meanwhile, it increases first and then decreases as dimension d increases for each trimming proportion $1-p\geq 0.15$. $\mathbf{\Sigma}_n(p)$ is much more efficient than $\mathbf{S}_n$.

5. Finite-Sample Breakdown Points

Now we explore the robustness of $\boldsymbol{\mu}_n(p)$ and $\mathbf{\Sigma}_n(p)$ based on the projection depth by finite-sample breakdown point. Donoho and Huber [35] discussed two types of finite-sample breakdown points: replacement breakdown point (RBP) and addition breakdown point (ABP). We employ the finite-sample replacement breakdown point in this section. Generally speaking, the finite-sample replacement breakdown point of an estimator is the minimum fraction of contaminated points (outliers or inliers) in a data set such that the estimator breaks down. Specifically, suppose that any m points in a sample $\boldsymbol{X}^n=\{\boldsymbol{x}_1,\cdots,\boldsymbol{x}_n\}$ are replaced (contaminated) by m arbitrary points with the contaminated data set denoted by $\boldsymbol{X}_m^n$. Then, the finite-sample replacement breakdown point of a location estimator T at $\boldsymbol{X}^n$ is defined as

$$\mathrm{RBP}(T,\boldsymbol{X}^n)=\min\{\frac{m}{n}:\sup_{\boldsymbol{X}_m^n}\left\|T(\boldsymbol{X}_m^n)-T(\boldsymbol{X}^n)\right\|=\infty\},$$

and the finite-sample replacement breakdown point of a scatter estimator V at $\boldsymbol{X}^n$ is defined as

$$\mathrm{RBP}(V,\boldsymbol{X}^n)=\min\{\frac{m}{n}:\mathrm{trace}[V(\boldsymbol{X}^n)V(\boldsymbol{X}_m^n)^{-1}+V(\boldsymbol{X}^n)^{-1}V(\boldsymbol{X}_m^n)]=\infty\}.$$

Essentially, the breakdown of V means that the largest eigenvalue $\lambda_1(V(\boldsymbol{X}_m^n))$ of $V(\boldsymbol{X}_m^n)$ can be made arbitrarily large (*explosion breakdown*) or the smallest eigenvalue $\lambda_d(V(\boldsymbol{X}_m^n))$ of $V(\boldsymbol{X}_m^n)$ can be made arbitrarily close to 0 (*implosion* *breakdown*). Usually, explosion breakdown is caused by outliers and implosion breakdown by inliers. However, the breakdown of a location estimator is only caused by outliers unless it is related to some scatter estimator. Before we give our main result, we briefly recall the projection depth.

The projection depth was first introduced by Liu [8] and was thoroughly studied by Zuo [9]. Given a univariate location estimator μ and a univariate scale estimator σ, the outlyingness of a point $\boldsymbol{x}$ in $\mathbb{R}^d$ with respect to a sample $\boldsymbol{X}^n = \{\boldsymbol{x}_1, \cdots, \boldsymbol{x}_n\}$ is defined as

$$O_{\boldsymbol{X}^n}(\boldsymbol{x}) = \sup_{\boldsymbol{u} \in S^{d-1}} \left| \frac{\boldsymbol{u}'\boldsymbol{x} - \mu(\boldsymbol{u}'\boldsymbol{X}^n)}{\sigma(\boldsymbol{u}'\boldsymbol{X}^n)} \right|,$$

where $S^{d-1} = \{\boldsymbol{u} \in \mathbb{R}^d : \|\boldsymbol{u}\| = 1\}$ and $\boldsymbol{u}'\boldsymbol{X}^n = \{\boldsymbol{u}'\boldsymbol{x}_1, \cdots, \boldsymbol{u}'\boldsymbol{x}_n\}$. Then the projection depth of $\boldsymbol{x}$ with respect to $\boldsymbol{X}^n$ is defined as

$$PD_{\boldsymbol{X}^n}(\boldsymbol{x}) = 1/(1 + O_{\boldsymbol{X}^n}(\boldsymbol{x})).$$

It is clear that $PD_{\boldsymbol{X}^n}(\boldsymbol{x})$ depends on μ and σ, and so does any estimator based on $PD_{\boldsymbol{X}^n}(\boldsymbol{x})$. For the best robustness of an estimator based on $PD_{\boldsymbol{X}^n}(\boldsymbol{x})$, the pair $(\mu, \sigma) = (\mathrm{Med}, \mathrm{MAD}_k)$ is often chosen, where Med is the univariate median and MAD_k is a modified median absolute deviation (MAD):
$\mathrm{MAD}_k(\boldsymbol{X}^n) = \mathrm{Med}_k\{|x_i - \mathrm{Med}(\boldsymbol{X}^n)|, i = 1, \cdots, n\}$ with
$\mathrm{Med}_k(\boldsymbol{X}^n) = (x_{(\lceil (n+k)/2 \rceil)} + x_{(\lfloor (n+k)/2 \rfloor + 1)})/2$ for $k = 0, \cdots, n-1$ and $x_{(1)}, \cdots, x_{(n)}$ being the ordered values of $\boldsymbol{X}^n = \{x_1, \cdots, x_n\}$ in $\mathbb{R}^1$ from the smallest to the largest, where $\lceil \cdot \rceil$ is the ceiling function and $\lfloor \cdot \rfloor$ the floor function. When $k = 0$, $\mathrm{MAD}_k = \mathrm{MAD}$ the usual median absolute deviation. It is well-known that $\mathrm{RBP}(\mathrm{Med}, \boldsymbol{X}^n) = \lfloor (n+1)/2 \rfloor / n$. As for MAD_k, Gather and Hilker [36] showed that

$$\mathrm{RBP}(\mathrm{MAD}_k, \boldsymbol{X}^n) = \lfloor (n - c(\boldsymbol{X}^n) + 1)/2 \rfloor / n$$

when $k = c(\boldsymbol{X}^n)$, where $c(\boldsymbol{X}^n)$ denotes the maximal multiplicity of points in $\boldsymbol{X}^n$.

Unlike the trimming based on the halfspace depth, the trimming based on the projection depth is done based on only the depth values of data points regardless of their directions. This feature can significantly increase the finite-sample breakdown points of $\boldsymbol{\mu}_n(p)$ and $\Sigma_n(p)$. That is why we choose the projection depth here. To make things clear, $C_n(p)$, $\boldsymbol{\mu}_n(p)$, and $\Sigma_n(p)$ at $\boldsymbol{X}^n$ are denoted by $C_{\boldsymbol{X}^n}(p)$, $\boldsymbol{\mu}_{\boldsymbol{X}^n}(p)$, and $\boldsymbol{\Sigma}_{\boldsymbol{X}^n}(p)$, respectively, in the following discussion. If a data set $\boldsymbol{X}^n$ in $\mathbb{R}^d$ is in general position, that is, no more than d points in $\boldsymbol{X}^n$ are contained in any $(d-1)$-dimensional hyperplane, we have the following results.

Theorem 4. Let $\mu_n(p)$ and $\Sigma_n(p)$, $p \in (0,1)$, be the $1-p$ proportion trimmed mean and trimmed scatter matrix based on the projection depth with $(\mu, \sigma) = (\mathrm{Med}, \mathrm{MAD}_d)$. Suppose that $\boldsymbol{X}^n$ is in general position and $\|zw(PD_n(z))\| \to \infty$ as $\|z\| \to \infty$ for any z in $\mathbb{R}^d$. Then

1) $\mathrm{RBP}(\boldsymbol{\mu}_n(p), \boldsymbol{X}^n) = \min\{n - \lceil np \rceil + 1, \lfloor (n-d+1)/2 \rfloor\}/n$ for $n > d$, and

2) $\mathrm{RBP}(\boldsymbol{\Sigma}_n(p), \boldsymbol{X}^n) = \min\{n - \lceil np \rceil + 1, \lceil np \rceil - d\}/n$ for n such that $\lceil np \rceil > d$.

From this theorem, we see that RBP $(\boldsymbol{\mu}_n(p), \boldsymbol{X}^n) = \lfloor (n-d+1)/2 \rfloor / n$ for all the trimmed means with $n - \lceil np \rceil + 1 \ge \lfloor (n-d+1)/2 \rfloor$, *i.e.*, with at least $\lfloor (n-d+1)/2 \rfloor - 1$ points trimmed off. However, RBP $(\boldsymbol{\Sigma}_n(p), \boldsymbol{X}^n) = \lfloor (n-d+1)/2 \rfloor / n$ for the only trimmed scatter matrix with $n - \lceil np \rceil + 1 = \lfloor (n-d+1)/2 \rfloor$, *i.e.*, with $\lfloor (n-d+1)/2 \rfloor - 1$ points trimmed off. It is well-known that $\lfloor (n-d+1)/2 \rfloor / n$ is the upper bound of RBP of any affine equivariant scatter estimators [3]. Also, $\lfloor (n-d+1)/2 \rfloor / n$ is the highest RBP of almost all existing affine equivariant location estimators.

Conflicts of Interest

The author declares no conflicts of interest regarding the publication of this paper.

References

[1] Maronna, R.A. (1976) Robust *M*-Estimators of Multivariate Location and Scatter. *The Annals of Statistics*, **4**, 51-67. https://doi.org/10.1214/aos/1176343347

[2] Rousseeuw, P. (1985) Multivariate Estimation with High Breakdown Point. In: Grossmann, W., Pflug, G., Vincze, I. and Wertz, W., Eds., *Mathematical Statistics and Applications*, Springer, 283-297. https://doi.org/10.1007/978-94-009-5438-0_20

[3] Davies, P.L. (1987) Asymptotic Behaviour of *S*-Estimates of Multivariate Location Parameters and Dispersion Matrices. *The Annals of Statistics*, **15**, 1269-1292. https://doi.org/10.1214/aos/1176350505

[4] Lopuhaa, H.P. (1989) On the Relation between *S*-Estimators and *M*-Estimators of Multivariate Location and Covariance. *The Annals of Statistics*, **17**, 1662-1683. https://doi.org/10.1214/aos/1176347386

[5] Lopuhaä, H.P. and Lopuhaa, H.P. (1991) Multivariate *τ*-Estimators for Location and Scatter. *The Canadian Journal of Statistics/La Revue Canadienne de Statistique*, **19**, 307-321. https://doi.org/10.2307/3315396

[6] Kent, J.T. and Tyler, D.E. (1996) Constrained *M*-Estimation for Multivariate Location and Scatter. *The Annals of Statistics*, **24**, 1346-1370. https://doi.org/10.1214/aos/1032526973

[7] Tukey, J.W. (1975) Mathematics and Picturing of Data. *Proceedings of the International Congress of Mathematicians*, Zürich, 3-11 August 1994, 523-531.

[8] Liu, R.Y. (1992) Data Depth and Multivariate Rank Tests. In: Dodge Y., Ed., *L1-Statistical Analysis and Related Methods*, North-Holland, 279-294.

[9] Zuo, Y. (2003) Projection-Based Depth Functions and Associated Medians. *The Annals of Statistics*, **31**, 1460-1490. https://doi.org/10.1214/aos/1065705115

[10] Liu, R.Y. (1990) On a Notion of Data Depth Based on Random Simplices. *The Annals of Statistics*, **18**, 405-414. https://doi.org/10.1214/aos/1176347507

[11] Serfling, R. and Zuo, Y. (2000) General Notions of Statistical Depth Function. *The Annals of Statistics*, **28**, 461-482. https://doi.org/10.1214/aos/1016218226

[12] Mahalanobis, P.C. (1936) On the Generalized Distance in Statistics. *The journal of the Asiatic Society of Bengal*, **XXVI**, 541-588.

[13] Liu, R.Y., Parelius, J.M. and Singh, K. (1999) Multivariate Analysis by Data Depth: Descriptive Statistics, Graphics and Inference, (with Discussion and a Rejoinder by Liu and Singh). *The Annals of Statistics*, **27**, 783-858. https://doi.org/10.1214/aos/1018031260

[14] Serfling, R. (2006) Depth Functions in Nonparametric Multivariate Inference. In: Liu, R.Y., Serfling, R. and Souvaine, D.L., Eds., Data Depth: *Robust Multivariate Analysis, Computational Geometry and Applications*, American Mathematical Society, 1-16. https://doi.org/10.1090/dimacs/072/01

[15] Dümbgen, L. (1992) Limit Theorems for the Simplicial Depth. *Statistics & Probability Letters*, **14**, 119-128. https://doi.org/10.1016/0167-7152(92)90075-g

[16] Massé, J. (2004) Asymptotics for the Tukey Depth Process, with an Application to a Multivariate Trimmed Mean. *Bernoulli*, **10**, 397-419. https://doi.org/10.3150/bj/1089206404

[17] Zuo, Y., Cui, H. and He, X. (2004) On the Stahel-Donoho Estimator and Depth-Weighted Means of Multivariate Data. *The Annals of Statistics*, **32**, 167-188. https://doi.org/10.1214/aos/1079120132

[18] Zuo, Y. and Cui, H. (2005) Depth Weighted Scatter Estimators. *The Annals of Statistics*, **33**, 381-413. https://doi.org/10.1214/009053604000000922

[19] Zuo, Y. (2006) Multidimensional Trimming Based on Projection Depth. *The Annals of Statistics*, **34**, 2211-2251. https://doi.org/10.1214/009053606000000713

[20] Donoho, D.L. and Gasko, M. (1992) Breakdown Properties of Location Estimates Based on Halfspace Depth and Projected Outlyingness. *The Annals of Statistics*, **20**, 1803-1827. https://doi.org/10.1214/aos/1176348890

[21] Massé, J. (2009) Multivariate Trimmed Means Based on the Tukey Depth. *Journal of Statistical Planning and Inference*, **139**, 366-384. https://doi.org/10.1016/j.jspi.2008.03.038

[22] Wang, J. (2018) A Note on Weak Convergence of General Halfspace Depth Trimmed Means. *Statistics & Probability Letters*, **142**, 50-56. https://doi.org/10.1016/j.spl.2018.07.005

[23] Zawadzki, Z., Kosiorowski, D., Slomczynski, K., Bocian, M. and Wegrzynkiewicz, A. (2025) DepthProc: Statistical Depth Functions for Multivariate Analysis (Ver. 2.1.6). R Package. https://cran.r-project.org/package=DepthProc

[24] Arcones, M.A. and Gine, E. (1993) Limit Theorems for *U*-Processes. *The Annals of Probability*, **21**, 1494-1542. https://doi.org/10.1214/aop/1176989128

[25] Arcones, M.A., Cui, H. and Zuo, Y. (2006) Empirical Depth Processes. *TEST*, **15**, 151-177. https://doi.org/10.1007/bf02595423

[26] Nolan, D. (1992) Asymptotics for Multivariate Trimming. *Stochastic Processes and Their Applications*, **42**, 157-169. https://doi.org/10.1016/0304-4149(92)90032-l

[27] van der Vaart, A.W. and Wellner, J.A. (1996) Weak Convergence and Empirical Processes, with Applications to Statistics. Springer. https://doi.org/10.1007/978-1-4757-2545-2

[28] Vapnik, V.N. and Chervonenkis, A.Y. (2015) On the Uniform Convergence of Relative Frequencies of Events to Their Probabilities. In: Vovk, V., Papadopoulos, H. and Gammerman, A., Eds., *Measures of Complexity*, Springer, 11-30. https://doi.org/10.1007/978-3-319-21852-6_3

[29] He, X. and Wang, G. (1997) Convergence of Depth Contours for Multivariate Datasets. *The Annals of Statistics*, **25**, 495-504. https://doi.org/10.1214/aos/1031833661

[30] Fang, K.T., Kotz, S. and Ng, K.W. (1990) Symmetric Multivariate and Related Distributions. Chapman and Hall. https://doi.org/10.1007/978-1-4899-2937-2

[31] Romanazzi, M. (2001) Influence Function of Halfspace Depth. *Journal of Multivariate Analysis*, **77**, 138-161. https://doi.org/10.1006/jmva.2000.1929

[32] Tyler, D.E. (1983) Robustness and Efficiency Properties of Scatter Matrices. *Biometrika*, **70**, 411-420. https://doi.org/10.1093/biomet/70.2.411

[33] Mauchly, J.W. (1940) Significance Test for Sphericity of a Normal *N*-Variate Distribution. *The Annals of Mathematical Statistics*, **11**, 204-209. https://doi.org/10.1214/aoms/1177731915

[34] Muirhead, R.J. (1982) Aspects of Multivariate Statistical Theory. Wiley. https://doi.org/10.1002/9780470316559

[35] Donoho, D.L. and Huber, P.J. (1983) The Notion of Breakdown Point. In: Bickel, P.J., Doksum, K.A., and Hodges Jr., J.L., Eds., *A Festschrift for Erich L. Lehmann*, Wadsworth, 157-184.

[36] Gather, U. and Hilker, T. (1997) A Note on Tyler's Modification of the MAD for the Stahel-Donoho Estimator. *The Annals of Statistics*, **25**, 2024-2026. https://doi.org/10.1214/aos/1069362384

[37] Wang, J. (2019) Asymptotics of Generalized Depth-Based Spread Processes and Applications. *Journal of Multivariate Analysis*, **169**, 363-380. https://doi.org/10.1016/j.jmva.2018.09.012

[38] Pollard, D. (1984) Convergence of Stochastic Processes. Springer. https://doi.org/10.1007/978-1-4612-5254-2

[39] Tyler, D.E. (1982) Radial Estimates and the Test for Sphericity. *Biometrika*, **69**, 429-436. https://doi.org/10.1093/biomet/69.2.429

Appendix: Proofs of Main Results

To prove Theorem 2, we prove two lemmas first.

Lemma 1. For any sequence of random vectors, $\boldsymbol{\xi}_1, \boldsymbol{\xi}_2, \ldots$, and constant vector $\boldsymbol{c}$ in $\mathbb{R}^d$, if $\sqrt{n}(\boldsymbol{\xi}_n - \boldsymbol{c})$ converges to a random vector $\boldsymbol{\xi}$ in distribution, then $\boldsymbol{\xi}_n$ converges to $\boldsymbol{c}$ in probability.

Proof. Since $\sqrt{n}(\boldsymbol{\xi}_n - \boldsymbol{c})$ converges to $\boldsymbol{\xi}$ in distribution,

$$P(\|\boldsymbol{\xi}_n - \boldsymbol{c}\| \le \varepsilon) = P(\|\sqrt{n}(\boldsymbol{\xi}_n - c)\| \le \sqrt{n}\varepsilon) \to 1$$

as $n \to \infty$ for any $\varepsilon > 0$, that is, $\boldsymbol{\xi}_n$ converges to $\boldsymbol{c}$ in probability. □

Lemma 2. If $f_D(F_D^{-1}(1-p)) > 0$ and $\frac{\partial}{\partial \alpha} r(\alpha, \boldsymbol{u})$ exists at $\alpha = F_D^{-1}(1-p)$, then

1)
$$\begin{aligned} l_0(p)\boldsymbol{\mu}_F^{(1)}(p) = -\int_{S^{d-1}} & (r(F_D^{-1}(1-p), \boldsymbol{u})\boldsymbol{u} - \boldsymbol{\mu}_F(p)) \\ & w(D(r(F_D^{-1}(1-p), \boldsymbol{u})\boldsymbol{u})) f(r(F_D^{-1}(1-p), \boldsymbol{u})\boldsymbol{u}) \\ & |J(r(F_D^{-1}(1-p), \boldsymbol{u}), \boldsymbol{u})| \frac{\partial}{\partial \alpha} r(\alpha, \boldsymbol{u})|_{\alpha = F_D^{-1}(1-p)} \\ & \frac{1}{f_D(F_D^{-1}(1-p))} d\boldsymbol{u}, \end{aligned}$$

and

2)
$$\begin{aligned} l_0(p)\Delta_F^{(1)}(p) = -\int_{S^{d-1}} & (r^2(F_D^{-1}(1-p), \boldsymbol{u})\boldsymbol{u}\boldsymbol{u}' - \Delta_F(p)) \\ & w(D(r(F_D^{-1}(1-p), \boldsymbol{u})\boldsymbol{u})) f(r(F_D^{-1}(1-p), \boldsymbol{u})\boldsymbol{u}) \\ & |J(r(F_D^{-1}(1-p), \boldsymbol{u}), \boldsymbol{u})| \frac{\partial}{\partial \alpha} r(\alpha, \boldsymbol{u})|_{\alpha = F_D^{-1}(1-p)} \\ & \frac{1}{f_D(F_D^{-1}(1-p))} d\boldsymbol{u}. \end{aligned}$$

Proof. Taking the transformation, $\boldsymbol{x} = r\boldsymbol{u}$ with $r = \|\boldsymbol{x}\|$ and $\boldsymbol{u} = \frac{\boldsymbol{x}}{\|\boldsymbol{x}\|}$, we have

$$\begin{aligned} l_0(p) &= \int_{C(p)} w(D(\boldsymbol{x}))dF(\boldsymbol{x}) \\ &= \int_{S^{d-1}} \{\int_0^{r(F_D^{-1}(1-p), \boldsymbol{u})} w(D(r\boldsymbol{u})) f(r\boldsymbol{u}) |J(r, \boldsymbol{u})| dr\} d\boldsymbol{u}, \end{aligned}$$

and

$$\begin{aligned} & \int_{C(p)} (\boldsymbol{x} - \boldsymbol{\mu}_F(p)) w(D(\boldsymbol{x}))dF(\boldsymbol{x}) \\ &= \int_{S^{d-1}} \{\int_0^{r(F_D^{-1}(1-p), \boldsymbol{u})} (r\boldsymbol{u} - \boldsymbol{\mu}_F(p)) w(D(r\boldsymbol{u})) f(r\boldsymbol{u}) |J(r, \boldsymbol{u})| dr\} d\boldsymbol{u}. \end{aligned}$$

Then,

$$\begin{aligned} & \frac{d}{dp}[\int_{C(p)} (\boldsymbol{x} - \boldsymbol{\mu}_F(p)) w(D(\boldsymbol{x}))dF(\boldsymbol{x})] = \int_{S^{d-1}} (r(F_D^{-1}(1-p), \boldsymbol{u})\boldsymbol{u} - \boldsymbol{\mu}_F(p)) \\ & w(D(r(F_D^{-1}(1-p), \boldsymbol{u})\boldsymbol{u})) f(r(F_D^{-1}(1-p), \boldsymbol{u})\boldsymbol{u}) |J(r(F_D^{-1}(1-p), \boldsymbol{u}), \boldsymbol{u})| \\ & \frac{\partial}{\partial \alpha} r(\alpha, \boldsymbol{u})|_{\alpha = F_D^{-1}(1-p)} [-\frac{1}{f_D(F_D^{-1}(1-p))}] d\boldsymbol{u} + \int_{S^{d-1}} \{\int_0^{r(F_D^{-1}(1-p), \boldsymbol{u})} (0 - \boldsymbol{\mu}_F^{(1)}(p)) \\ & w(D(r\boldsymbol{u})) f(r\boldsymbol{u}) |J(r, \boldsymbol{u})| dr\} d\boldsymbol{u}. \end{aligned}$$

Since $\int_{C(p)}(\boldsymbol{x}-\boldsymbol{\mu}_F(p))w(D(\boldsymbol{x}))dF(\boldsymbol{x})=0$, the first result follows.

In the same way, we obtain the second result. □

Proof of Theorem 2. Since $D_F(\boldsymbol{x})$ is affine invariant, by Theorem 1 1), $\boldsymbol{\mu}_n(p)-\boldsymbol{\mu}_F(p)$ and $\boldsymbol{\Sigma}_n(p)-\boldsymbol{\Sigma}_F(p)$ is invariant for any translation. Thus, we can assume without loss of generality that the origin is the deepest point. First

$$\boldsymbol{\mu}_n(p)-\boldsymbol{\mu}_F(p)=\frac{\int_{C_n(p)}(\boldsymbol{x}-\boldsymbol{\mu}_F(p))w(D_n(\boldsymbol{x}))dF_n(\boldsymbol{x})}{\int_{C_n(p)}w(D_n(\boldsymbol{x}))dF_n(\boldsymbol{x})}=\frac{I_{1n}+I_{2n}+I_{3n}}{\int_{C_n(p)}w(D_n(\boldsymbol{x}))dF_n(\boldsymbol{x})},$$

where $I_{1n}=\int_{C(p)}(\boldsymbol{x}-\boldsymbol{\mu}_F(p))w(D(\boldsymbol{x}))dF_n(\boldsymbol{x})$,

$$I_{2n}=\int_{C_n(p)}(\boldsymbol{x}-\boldsymbol{\mu}_F(p))[w(D_n(\boldsymbol{x}))-w(D(\boldsymbol{x}))]dF_n(\boldsymbol{x}),$$

$$I_{3n}=\int_{C_n(p)}(\boldsymbol{x}-\boldsymbol{\mu}_F(p))w(D(\boldsymbol{x}))dF_n(\boldsymbol{x})-\int_{C(p)}(\boldsymbol{x}-\boldsymbol{\mu}_F(p))w(D(\boldsymbol{x}))dF_n(\boldsymbol{x}).$$

Let $v_n=n^{1/2}(F_n-F)$. Since $\int_{C(p)}(\boldsymbol{x}-\boldsymbol{\mu}_F(p))w(D(\boldsymbol{x}))dF(\boldsymbol{x})=0$,

$$n^{1/2}I_{1n}=\int_{C(p)}(\boldsymbol{x}-\boldsymbol{\mu}_F(p))w(D(\boldsymbol{x}))dv_n(\boldsymbol{x})=\int l_1(\boldsymbol{x})dv_n(\boldsymbol{x}).$$

For I_{2n}, by the Mean Value Theorem, there exists $\theta_n(\boldsymbol{x})$ between $D_n(\boldsymbol{x})$ and $D(\boldsymbol{x})$ such that

$$n^{1/2}I_{2n}=\int_{C_n(p)}(\boldsymbol{x}-\boldsymbol{\mu}_F(p))w^{(1)}(\theta_n(\boldsymbol{x}))\{n^{1/2}(D_n(\boldsymbol{x})-D(\boldsymbol{x}))\}dF_n(\boldsymbol{x}).$$

Then, the continuity of $w^{(1)}(\cdot)$, the condition (C_1), which implies $D_n(\boldsymbol{x})\to D(\boldsymbol{x})$ uniformly in $\boldsymbol{x}$ and thus $\left\|n^{1/2}(D_n(\boldsymbol{x})-D(\boldsymbol{x}))\right\|_\infty$ is bounded in probability by Lemma 1, and the uniform strong convergence of F_n to F over $\mathcal{D}$, which contains $C_n(p)$ and $C(p)$, lead to

$$\begin{aligned}n^{1/2}I_{2n}&=\int_{C_n(p)}(\boldsymbol{x}-\boldsymbol{\mu}_F(p))w^{(1)}(D(\boldsymbol{x}))\{n^{1/2}(D_n(\boldsymbol{x})-D(\boldsymbol{x}))\}dF_n(\boldsymbol{x})+o_p(1)\\&=\int_{C_n(p)}(\boldsymbol{x}-\boldsymbol{\mu}_F(p))w^{(1)}(D(\boldsymbol{x}))\{n^{1/2}(D_n(\boldsymbol{x})-D(\boldsymbol{x}))\}dF(\boldsymbol{x})+o_p(1).\end{aligned}$$

Since (C_2) implies that $I_n(\alpha)\to I(\alpha)$ in probability in the Hausdorff distance by Lemma 1 and $F_{D_n}^{-1}(1-p)\to F_D^{-1}(1-p)$ in probability by Theorem 4 of Wang [37] and Lemma 1, $C_n(p)\to C(p)$ in probability in the Hausdorff distance, which leads to

$$n^{1/2}I_{2n}=\int_{C(p)}(\boldsymbol{x}-\boldsymbol{\mu}_F(p))w^{(1)}(D(\boldsymbol{x}))\{n^{1/2}(D_n(\boldsymbol{x})-D(\boldsymbol{x}))\}dF(\boldsymbol{x})+o_p(1).$$

Then, (C_1) and Fubini's Theorem yields

$$\begin{aligned}n^{1/2}I_{2n}&=\int_{C(p)}(\boldsymbol{x}-\boldsymbol{\mu}_F(p))w^{(1)}(D(\boldsymbol{x}))\{\int\varphi(\boldsymbol{x},\boldsymbol{y})dv_n(\boldsymbol{y})\}dF(\boldsymbol{x})+o_p(1)\\&=\int\{\int_{C(p)}(\boldsymbol{x}-\boldsymbol{\mu}_F(p))w^{(1)}(D(\boldsymbol{x}))\varphi(\boldsymbol{x},\boldsymbol{y})dF(\boldsymbol{x})\}dv_n(\boldsymbol{y})+o_p(1)\\&=\int l_2(\boldsymbol{y})dv_n(\boldsymbol{y})+o_p(1).\end{aligned}$$

Now we work on I_{3n}. Since $C_n(p)\to C(p)$ in probability in the Hausdorff distance, and $C_n(p)$ and $C(p)$ are contained in a Donsker class $\mathcal{D}$, we obtain by the equicontinuity of the empirical processes $\{v_n\}$ (see, e.g., Lemma VII.15 of Pollard [38])

$$\int_{C_n(p)} (\boldsymbol{x}-\boldsymbol{\mu}_F(p)) w(D(\boldsymbol{x})) dv_n(\boldsymbol{x}) = \int_{C(p)} (\boldsymbol{x}-\boldsymbol{\mu}_F(p)) w(D(\boldsymbol{x})) dv_n(\boldsymbol{x})\} + o_p(1).$$

Therefore,

$$\begin{aligned} n^{1/2} I_{3n} &= \{n^{1/2} \int_{C_n(p)} (\boldsymbol{x}-\boldsymbol{\mu}_F(p)) w(D(\boldsymbol{x})) dF(\boldsymbol{x}) \\ &\quad + \int_{C_n(p)} (\boldsymbol{x}-\boldsymbol{\mu}_F(p)) w(D(\boldsymbol{x})) dv_n(\boldsymbol{x})\} \\ &\quad - \{n^{1/2} \int_{C(p)} (\boldsymbol{x}-\boldsymbol{\mu}_F(p)) w(D(\boldsymbol{x})) dF(\boldsymbol{x}) \\ &\quad + \int_{C(p)} (\boldsymbol{x}-\boldsymbol{\mu}_F(p)) w(D(\boldsymbol{x})) dv_n(\boldsymbol{x})\} \\ &= n^{1/2} \{\int_{C_n(p)} (\boldsymbol{x}-\boldsymbol{\mu}_F(p)) w(D(\boldsymbol{x})) dF(\boldsymbol{x}) \\ &\quad - \int_{C(p)} (\boldsymbol{x}-\boldsymbol{\mu}_F(p)) w(D(\boldsymbol{x})) dF(\boldsymbol{x})\} + o_p(1). \end{aligned}$$

Taking the transformation $\boldsymbol{x} = r\boldsymbol{u}$ with $r = \|\boldsymbol{x}\|$ and $\boldsymbol{u} = \dfrac{\boldsymbol{x}}{\|\boldsymbol{x}\|}$, we have

$$\begin{aligned} n^{1/2} I_{3n} &= n^{1/2} \int_{S^{d-1}} \{\int_{r(F_D^{-1}(1-p),\boldsymbol{u})}^{r_n(F_{D_n}^{-1}(1-p),\boldsymbol{u})} (r\boldsymbol{u}-\boldsymbol{\mu}_F(p)) w(D(r\boldsymbol{u})) f(r\boldsymbol{u}) \\ &\quad |J(r,\boldsymbol{u})| dr\} d\boldsymbol{u}, \end{aligned}$$

where $J(r,\boldsymbol{u})$ is the Jacobian of the transformation.

By the Mean Value Theorem, there exists $\delta_n(p,\boldsymbol{u})$ between $r_n(F_{D_n}^{-1}(1-p),\boldsymbol{u})$ and $r(F_D^{-1}(1-p),\boldsymbol{u})$ such that

$$\begin{aligned} n^{1/2} I_{3n} &= \int_{S^{d-1}} [\delta_n(p,\boldsymbol{u})\boldsymbol{u}-\boldsymbol{\mu}_F(p)] w(D(\delta_n(p,\boldsymbol{u})\boldsymbol{u})) f(\delta_n(p,\boldsymbol{u})\boldsymbol{u}) \\ &\quad |J(\delta_n(p,\boldsymbol{u}),\boldsymbol{u})| \{n^{1/2}[r_n(F_{D_n}^{-1}(1-p),\boldsymbol{u}) - r(F_D^{-1}(1-p),\boldsymbol{u})]\} d\boldsymbol{u}. \end{aligned}$$

Since (C_2) implies that $r_n(\alpha,\boldsymbol{u}) \to r(\alpha,\boldsymbol{u})$ in probability by Lemma 1, $F_{D_n}^{-1}(1-p) \to F_D^{-1}(1-p)$ in probability, $w(\cdot)$ is continuous, and f is continuous in some neighborhood of $\partial C(p)$, which implies that $D(\boldsymbol{x})$ is continuous in the neighborhood, we have that $\delta_n(p,\boldsymbol{u}) = r(F_D^{-1}(1-p),\boldsymbol{u}) + o_p(1)$, $w(D(\delta_n(p,\boldsymbol{u})\boldsymbol{u})) = w(D(r(F_D^{-1}(1-p),\boldsymbol{u})\boldsymbol{u})) + o_p(1)$, $f(\delta_n(p,\boldsymbol{u})\boldsymbol{u}) = f(r(F_D^{-1}(1-p),\boldsymbol{u})\boldsymbol{u}) + o_p(1)$, and $|J(\delta_n(p,\boldsymbol{u}),\boldsymbol{u})| = |J(r(F_D^{-1}(1-p),\boldsymbol{u}),\boldsymbol{u})| + o_p(1)$ for n sufficiently large. Thus

$$\begin{aligned} n^{1/2} I_{3n} &= \int_{S^{d-1}} (r(F_D^{-1}(1-p),\boldsymbol{u})\boldsymbol{u}-\boldsymbol{\mu}_F(p)) w(D(r(F_D^{-1}(1-p),\boldsymbol{u})\boldsymbol{u})) \\ &\quad f(r(F_D^{-1}(1-p),\boldsymbol{u})\boldsymbol{u}) |J(r(F_D^{-1}(1-p),\boldsymbol{u}),\boldsymbol{u})| \\ &\quad \{n^{1/2}[r_n(F_{D_n}^{-1}(1-p),\boldsymbol{u}) - r(F_D^{-1}(1-p),\boldsymbol{u})]\} d\boldsymbol{u} + o_p(1). \end{aligned}$$

Decomposing $r_n(F_{D_n}^{-1}(1-p),\boldsymbol{u}) - r(F_D^{-1}(1-p),\boldsymbol{u})$, we obtain

$$n^{1/2} I_{3n} = n^{1/2}(I_{3n1} + I_{3n2}) + o_p(1),$$

where

$$\begin{aligned} n^{1/2} I_{3n1} &= \int_{S^{d-1}} (r(F_D^{-1}(1-p),\boldsymbol{u})\boldsymbol{u}-\boldsymbol{\mu}_F(p)) w(D(r(F_D^{-1}(1-p),\boldsymbol{u})\boldsymbol{u})) \\ &\quad f(r(F_D^{-1}(1-p),\boldsymbol{u})\boldsymbol{u}) |J(r(F_D^{-1}(1-p),\boldsymbol{u}),\boldsymbol{u})| \\ &\quad \{n^{1/2}[r_n(F_{D_n}^{-1}(1-p),\boldsymbol{u}) - r(F_{D_n}^{-1}(1-p),\boldsymbol{u})]\} d\boldsymbol{u}, \end{aligned}$$

$$n^{1/2}I_{3n2} = \int_{S^{d-1}} (r(F_D^{-1}(1-p),\boldsymbol{u})\boldsymbol{u} - \boldsymbol{\mu}_F(p))w(D(r(F_D^{-1}(1-p),\boldsymbol{u})\boldsymbol{u}))$$
$$f(r(F_D^{-1}(1-p),\boldsymbol{u})\boldsymbol{u})\,|\,J(r(F_D^{-1}(1-p),\boldsymbol{u}),\boldsymbol{u})\,|$$
$$\{n^{1/2}[r(F_{D_n}^{-1}(1-p),\boldsymbol{u}) - r(F_D^{-1}(1-p),\boldsymbol{u})]\}d\boldsymbol{u}.$$

By (C_2),

$$n^{1/2}I_{3n1} = \int_{S^{d-1}} (r(F_D^{-1}(1-p),\boldsymbol{u})\boldsymbol{u} - \boldsymbol{\mu}_F(p))w(D(r(F_D^{-1}(1-p),\boldsymbol{u})\boldsymbol{u}))$$
$$f(r(F_D^{-1}(1-p),\boldsymbol{u})\boldsymbol{u})\,|\,J(r(F_D^{-1}(1-p),\boldsymbol{u}),\boldsymbol{u})\,|$$
$$\{\int \psi(\boldsymbol{x},F_{D_n}^{-1}(1-p),\boldsymbol{u})dv_n(\boldsymbol{x})\}d\boldsymbol{u} + o_p(1).$$

Since $F_{D_n}^{-1}(1-p) \to F_D^{-1}(1-p)$ in probability and $\psi(\boldsymbol{x},\alpha,\boldsymbol{u})$ is continuous in α, the equicontinuity of the empirical processes $\{v_n\}$, and Fubini's Theorem lead to

$$n^{1/2}I_{3n1} = \int_{S^{d-1}} (r(F_D^{-1}(1-p),\boldsymbol{u})\boldsymbol{u} - \boldsymbol{\mu}_F(p))w(D(r(F_D^{-1}(1-p),\boldsymbol{u})\boldsymbol{u}))$$
$$f(r(F_D^{-1}(1-p),\boldsymbol{u})\boldsymbol{u})\,|\,J(r(F_D^{-1}(1-p),\boldsymbol{u}),\boldsymbol{u})\,|$$
$$\{\int \psi(\boldsymbol{x},F_D^{-1}(1-p),\boldsymbol{u})dv_n(\boldsymbol{x})\}d\boldsymbol{u} + o_p(1)$$
$$= \int\{\int_{S^{d-1}} (r(F_D^{-1}(1-p),\boldsymbol{u})\boldsymbol{u} - \boldsymbol{\mu}_F(p))w(D(r(F_D^{-1}(1-p),\boldsymbol{u})\boldsymbol{u}))$$
$$f(r(F_D^{-1}(1-p),\boldsymbol{u})\boldsymbol{u})\,|\,J(r(F_D^{-1}(1-p),\boldsymbol{u}),\boldsymbol{u})\,|$$
$$\psi(\boldsymbol{x},F_D^{-1}(1-p),\boldsymbol{u})d\boldsymbol{u}\}dv_n(\boldsymbol{x}) + o_p(1)$$
$$= \int l_3(\boldsymbol{x})dv_n(\boldsymbol{x}) + o_p(1)$$

For $n^{1/2}I_{3n2}$, by the Mean Value Theorem, there exists $\theta_n(p)$ between $F_{D_n}^{-1}(1-p)$ and $F_D^{-1}(1-p)$ such that

$$r(F_{D_n}^{-1}(1-p),\boldsymbol{u}) - r(F_D^{-1}(1-p),\boldsymbol{u}) = \frac{\partial}{\partial\alpha}r(\alpha,\boldsymbol{u})|_{\alpha=\theta_n(p)}$$
$$\{F_{D_n}^{-1}(1-p) - F_D^{-1}(1-p)\}.$$

Since $F_{D_n}^{-1}(1-p) \to F_D^{-1}(1-p)$ in probability and $\frac{\partial}{\partial\alpha}r(\alpha,\boldsymbol{u})$ is continuous in α, $\theta_n(p) = F_D^{-1}(1-p) + o_p(1)$ and $\frac{\partial}{\partial\alpha}r(\alpha,\boldsymbol{u})|_{\alpha=\theta_n(p)} = \frac{\partial}{\partial\alpha}r(\alpha,\boldsymbol{u})|_{\alpha=F_D^{-1}(1-p)} + o_p(1)$ uniformly in $\boldsymbol{u}\in S^{d-1}$ for n sufficiently large. In addition, by Theorem 4 of Wang [37],

$$n^{1/2}\{F_{D_n}^{-1}(1-p) - F_D^{-1}(1-p)\} = \frac{1}{f_D(F_D^{-1}(1-p))}\int\{I_{C(p)}(\boldsymbol{x})$$
$$+\int_{S^{d-1}} f(r(F_D^{-1}(1-p),\boldsymbol{u})\boldsymbol{u})$$
$$|\,J(r(F_D^{-1}(1-p),\boldsymbol{u}),\boldsymbol{u})\,|$$
$$\psi(\boldsymbol{x},F_D^{-1}(1-p),\boldsymbol{u})d\boldsymbol{u}\}dv_n(\boldsymbol{x}) + o_p(1).$$

Combining all these results and using Lemma 2 1), we obtain

$$
\begin{aligned}
n^{1/2} I_{3n2} = & n^{1/2} \int_{S^{d-1}} (r(F_D^{-1}(1-p), \boldsymbol{u})\boldsymbol{u} - \boldsymbol{\mu}_F(p)) w(D(r(F_D^{-1}(1-p), \boldsymbol{u})\boldsymbol{u})) \\
& f(r(F_D^{-1}(1-p), \boldsymbol{u})\boldsymbol{u}) \,|\, J(r(F_D^{-1}(1-p), \boldsymbol{u}), \boldsymbol{u}) \,| \\
& [\frac{\partial}{\partial \alpha} r(\alpha, \boldsymbol{u})|_{\alpha=\theta_n(p)} \{F_{D_n}^{-1}(1-p) - F_D^{-1}(1-p)\}] d\boldsymbol{u} \\
= & \{\int_{S^{d-1}} (r(F_D^{-1}(1-p), \boldsymbol{u})\boldsymbol{u} - \boldsymbol{\mu}_F(p)) w(D(r(F_D^{-1}(1-p), \boldsymbol{u})\boldsymbol{u})) \\
& f(r(F_D^{-1}(1-p), \boldsymbol{u})\boldsymbol{u}) \,|\, J(r(F_D^{-1}(1-p), \boldsymbol{u}), \boldsymbol{u}) \,| \\
& \frac{\partial}{\partial \alpha} r(\alpha, \boldsymbol{u})|_{\alpha=F_D^{-1}(1-p)} d\boldsymbol{u}\} \\
& \frac{1}{f_D(F_D^{-1}(1-p))} \int \{I_{C(p)}(\boldsymbol{x}) + \int_{S^{d-1}} f(r(F_D^{-1}(1-p), \boldsymbol{u})\boldsymbol{u}) \\
& |\, J(r(F_D^{-1}(1-p), \boldsymbol{u}), \boldsymbol{u}) \,|\, \psi(\boldsymbol{x}, F_D^{-1}(1-p), \boldsymbol{u}) d\boldsymbol{u}\} dv_n(\boldsymbol{x}) + o_p(1) \\
= & -l_0(p) \boldsymbol{\mu}_F^{(1)}(p) \int \{I_{C(p)}(\boldsymbol{x}) + \int_{S^{d-1}} f(r(F_D^{-1}(1-p), \boldsymbol{u})\boldsymbol{u}) \\
& |\, J(r(F_D^{-1}(1-p), \boldsymbol{u}), \boldsymbol{u}) \,|\, \psi(\boldsymbol{x}, F_D^{-1}(1-p), \boldsymbol{u}) d\boldsymbol{u}\} dv_n(\boldsymbol{x}) + o_p(1) \\
= & \int l_4(\boldsymbol{x}) dv_n(\boldsymbol{x}) + o_p(1).
\end{aligned}
$$

In the same way, we can get the asymptotic representation for $\int_{C_n(p)} w(D_n(\boldsymbol{x})) dF_n(\boldsymbol{x}) - \int_{C(p)} w(D(\boldsymbol{x})) dF(\boldsymbol{x})$, which implies $\int_{C_n(p)} w(D_n(\boldsymbol{x})) dF_n(\boldsymbol{x}) \to \int_{C(p)} w(D(\boldsymbol{x})) dF(\boldsymbol{x})$ in probability by Lemma 1, and thus we have the asymptotic representation for $\boldsymbol{\mu}_n(p) - \boldsymbol{\mu}_F(p)$. Then, the multivariate central limit theorem leads to the asymptotic distribution of $\sqrt{n}(\boldsymbol{\mu}_n(p) - \boldsymbol{\mu}_F(p))$.

2) Let $\Delta_F(p) = \dfrac{\int_{C(p)} \boldsymbol{x}\boldsymbol{x}' w(D(\boldsymbol{x})) dF(\boldsymbol{x})}{\int_{C(p)} w(D(\boldsymbol{x})) dF(\boldsymbol{x})}$ and let $\Delta_n(p)$ be its sample version. Then,

$$
\begin{aligned}
\boldsymbol{\Sigma}_n(p) - \boldsymbol{\Sigma}_F(p) = & \Delta_n(p) - \Delta_F(p) - \boldsymbol{\mu}_n(p)[\boldsymbol{\mu}_n(p) - \boldsymbol{\mu}_F(p)]' \\
& -[\boldsymbol{\mu}_n(p) - \boldsymbol{\mu}_F(p)]\boldsymbol{\mu}_F'(p).
\end{aligned}
$$

In the same way as for $n^{1/2}(\boldsymbol{\mu}_n(p) - \boldsymbol{\mu}_F(p))$, we obtain that $n^{1/2}(\Delta_n(p) - \Delta_F(p)) = \int [k_1(\boldsymbol{x}) + k_2(\boldsymbol{x}) + k_3(\boldsymbol{x}) + k_4(\boldsymbol{x})) / l_0(p)] dv_n(\boldsymbol{x}) + o_p(1)$. Since the weak convergence of $n^{1/2}(\boldsymbol{\mu}_n(p) - \boldsymbol{\mu}_F(p))$ implies that $\boldsymbol{\mu}_n(p) \to \boldsymbol{\mu}_F(p)$ in probability by Lemma 1, the asymptotic representation for $\boldsymbol{\Sigma}_n(p) - \boldsymbol{\Sigma}_F(p)$ and the asymptotic distribution of $\sqrt{n}(\boldsymbol{\Sigma}_n(p) - \boldsymbol{\Sigma}_F(p))$ follow immediately. □

To prove Corollary 1, the following results are needed.

Lemma 3. If $\boldsymbol{X} \sim E_d(h; \boldsymbol{0}, \mathbf{I}_{d\times d})$, then the density of $\|X\|^\beta$ for any $\beta > 0$ is

$$
f_{\|x\|^\beta}(r) = \frac{2\pi^{d/2}}{\beta\Gamma(d/2)} r^{d/\beta - 1} h(r^{2/\beta}), r \geq 0.
$$

Proof. It is well-known that the density of $\|X\|^2$ is

$$
f_{\|x\|^2}(r) = \frac{\pi^{d/2}}{\Gamma(d/2)} r^{d/2 - 1} h(r), r \geq 0.
$$

Then a transformation leads to the result. □

Lemma 4. *For* $d \ge 3$,

$$\int_0^{\pi}\cdots\int_0^{\pi}\int_0^{2\pi}\sin^{d-3}\theta_2\cdots\sin\theta_{d-2}d\theta_{n-1}\cdots d\theta_2=\frac{2\pi^{d/2}}{\Gamma(d/2)\int_0^{\pi}\sin^{d-2}\theta_1 d\theta_1}.$$

Proof. On one hand, the volume of the d-dimensional unit ball is

$$\frac{\pi^{d/2}}{\Gamma(d/2+1)}=\frac{2\pi^{d/2}}{d\Gamma(d/2)}.$$

On the other hand, the volume of the unit ball can be calculated as

$$\begin{aligned}\int_{\{y\in\mathbb{R}^d:\|y\|\le 1\}}d\boldsymbol{y}&=\int_0^1\int_0^{\pi}\cdots\int_0^{\pi}\int_0^{2\pi}r^{d-1}\sin^{d-2}\theta_1\sin^{d-3}\theta_2\cdots\sin\theta_{d-2}d\theta_{n-1}\cdots d\theta_1 dr\\&=\frac{1}{d}\left(\int_0^{\pi}\sin^{d-2}\theta_1 d\theta_1\right)\left(\int_0^{\pi}\cdots\int_0^{\pi}\int_0^{2\pi}\sin^{d-3}\theta_2\cdots\sin\theta_{d-2}d\theta_{n-1}\cdots d\theta_2\right).\end{aligned}$$

The result follows immediately. □

Lemma 5. *For* $d \ge 3$,

1) $$\int_0^{\pi}\cdots\int_0^{\pi}\int_0^{2\pi}\sin^{d-3}\theta_2\cdots\sin\theta_{d-2}d\theta_{n-1}\cdots d\theta_2=\frac{2\pi^{d/2}}{\Gamma(d/2)}\frac{d-1}{d\int_0^{\pi}\sin^{d}\theta_1 d\theta_1}.$$

2) $$\begin{aligned}&\int_0^{\pi}\cdots\int_0^{\pi}\int_0^{2\pi}(\sin\theta_2\cdots\sin\theta_{i-1}\cos\theta_i)^2\sin^{d-3}\theta_2\cdots\sin\theta_{d-2}d\theta_{n-1}\cdots d\theta_2\\&=\frac{2\pi^{d/2}}{\Gamma(d/2)}\frac{1}{d\int_0^{\pi}\sin^{d}\theta_1 d\theta_1}\text{ for } i=2,\cdots,d-1.\end{aligned}$$

3) $$\begin{aligned}&\int_0^{\pi}\cdots\int_0^{\pi}\int_0^{2\pi}(\sin\theta_2\cdots\sin\theta_{d-2}\sin\theta_{d-1})^2\sin^{d-3}\theta_2\cdots\sin\theta_{d-2}d\theta_{n-1}\cdots d\theta_2\\&=\frac{2\pi^{d/2}}{\Gamma(d/2)}\frac{1}{d\int_0^{\pi}\sin^{d}\theta_1 d\theta_1}.\end{aligned}$$

Proof. It is well-known that if $\boldsymbol{Z}\sim E_d(h;\boldsymbol{0},\mathbf{I}_{d\times d})$, then $\boldsymbol{U}=\frac{\boldsymbol{Z}}{\|\boldsymbol{Z}\|}$ is uniformly distributed on the unit sphere $S^{d-1}=\{\boldsymbol{u}\in\mathbb{R}^d:\|\boldsymbol{u}\|=1\}$, and $E(\boldsymbol{U}\boldsymbol{U}')=\int_{S^{d-1}}\frac{1}{area(S^{d-1})}\boldsymbol{u}\boldsymbol{u}'d\boldsymbol{u}=\frac{1}{d}\mathbf{I}_{d\times d}$. Using the polar coordinates representation for $\boldsymbol{u}$, *i.e.*, $\boldsymbol{u}=(\cos\theta_1,\sin\theta_1\cos\theta_2,\cdots,\sin\theta_1\cdots\sin\theta_{d-2}\cos\theta_{d-1},\sin\theta_1\cdots\sin\theta_{d-2}\sin\theta_{d-1})'$, and noting that $area(S^{d-1})=\frac{2\pi^{d/2}}{\Gamma(d/2)}$, we have

$$\begin{aligned}\frac{1}{d}\mathbf{I}_{d\times d}&=E(\boldsymbol{U}\boldsymbol{U}')\\&=\int_0^{\pi}\cdots\int_0^{\pi}\int_0^{2\pi}\frac{1}{area(S^{d-1})}\boldsymbol{u}\boldsymbol{u}'\sin^{d-2}\theta_1\sin^{d-3}\theta_2\cdots\sin\theta_{d-2}d\theta_{d-1}\cdots d\theta_1\\&=\int_0^{\pi}\cdots\int_0^{\pi}\int_0^{2\pi}\left(\frac{2\pi^{d/2}}{\Gamma(d/2)}\right)^{-1}\boldsymbol{u}\boldsymbol{u}'\sin^{d-2}\theta_1\sin^{d-3}\theta_2\cdots\sin\theta_{d-2}d\theta_{d-1}\cdots d\theta_1.\end{aligned}$$

1) On one hand, we have from the first diagonal elements on both sides

$$\frac{1}{d} = (\frac{2\pi^{d/2}}{\Gamma(d/2)})^{-1} \int_0^{\pi} \cdots \int_0^{\pi} \int_0^{2\pi} \cos^2\theta_1 \sin^{d-2}\theta_1 \sin^{d-3}\theta_2 \cdots \sin\theta_{d-2} d\theta_{n-1} \cdots d\theta_1,$$

and thus

$$\int_0^{\pi} \cdots \int_0^{\pi} \int_0^{2\pi} \sin^{d-3}\theta_2 \cdots \sin\theta_{d-2} d\theta_{n-1} \cdots d\theta_2 = \frac{2\pi^{d/2}}{\Gamma(d/2)} \frac{1}{d\int_0^{\pi} \cos^2\theta_1 \sin^{d-2}\theta_1 d\theta_1}.$$

On the other hand, by Lemma 4,

$$\frac{1}{d} = (\frac{2\pi^{d/2}}{\Gamma(d/2)})^{-1} \int_0^{\pi} \cdots \int_0^{\pi} \int_0^{2\pi} \cos^2\theta_1 \sin^{d-2}\theta_1 \sin^{d-3}\theta_2 \cdots \sin\theta_{d-2} d\theta_{n-1} \cdots d\theta_1$$
$$= (\frac{2\pi^{d/2}}{\Gamma(d/2)})^{-1} \int_0^{\pi} \cdots \int_0^{\pi} \int_0^{2\pi} (1-\sin^2\theta_1) \sin^{d-2}\theta_1 \sin^{d-3}\theta_2 \cdots \sin\theta_{d-2}$$
$$d\theta_{n-1} \cdots d\theta_1$$
$$= (\frac{2\pi^{d/2}}{\Gamma(d/2)})^{-1} \{\frac{2\pi^{d/2}}{\Gamma(d/2)} - (\int_0^{\pi} \sin^d\theta_1 d\theta_1)[\frac{2\pi^{d/2}}{\Gamma(d/2)} \frac{1}{\int_0^{\pi} \sin^{d-2}\theta_1 d\theta_1}]\}$$
$$= 1 - \frac{\int_0^{\pi} \sin^d\theta_1 d\theta_1}{\int_0^{\pi} \sin^{d-2}\theta_1 d\theta_1},$$

which leads to $\int_0^{\pi} \sin^{d-2}\theta_1 d\theta_1 = \frac{d}{d-1}\int_0^{\pi} \sin^d\theta_1 d\theta_1$. Therefore,

$$\int_0^{\pi} \cos^2\theta_1 \sin^{d-2}\theta_1 d\theta_1 = \frac{d}{d-1}\int_0^{\pi} \sin^d\theta_1 d\theta_1 - \int_0^{\pi} \sin^d\theta_1 d\theta_1 = \frac{1}{d-1}\int_0^{\pi} \sin^d\theta_1 d\theta_1,$$

and the desired result follows.

2) The assertion results from the i^{th} diagonal elements on both sides, $i = 2, \cdots, d-1$.

3) The result is obtained from the last diagonal elements on both sides. □

Proof of Corollary 1. 1) Since $w(t) = 1$ on $[0, \alpha_F^*]$, $l_2(z) = \mathbf{0}$. In addition, by Theorem 1 2), $\boldsymbol{\mu}_{F_{\boldsymbol{Z}}}(p) = \mathbf{0}$ and thus $l_4(z) = \mathbf{0}$. So, we only need to concretize $l_1(z)$ and $l_3(z)$. Since $\boldsymbol{Z} \sim E_d(h; \mathbf{0}, \mathbf{I}_{d\times d})$, $HD_{F_{\boldsymbol{Z}}}(z) = F_1(-\|z\|)$, $r_{F_{\boldsymbol{Z}}}(F_{HD}^{-1}(1-p), \boldsymbol{u}) = F_{\|\boldsymbol{Z}\|}^{-1}(p)$, and $C_{F_{\boldsymbol{Z}}}(p) = \{z \in \mathbb{R}^d : \|z\| \le F_{\|\boldsymbol{Z}\|}^{-1}(p)\}$. So,

$$l_1(z) = zI_{C_{F_{\boldsymbol{Z}}}(p)}(z) = \|z\|\frac{z}{\|z\|} I_{\{\|z\| \le F_{\|\boldsymbol{Z}\|}^{-1}(p)\}}(z) \quad \text{and}$$

$$l_3(z) = \int_{S^{d-1}} (F_{\|\boldsymbol{Z}\|}^{-1}(p)\boldsymbol{u}) f(F_{\|\boldsymbol{Z}\|}^{-1}(p)\boldsymbol{u}) \,|\, J(F_{\|\boldsymbol{Z}\|}^{-1}(p), \boldsymbol{u}) \,|\, \psi(z, F_{HD}^{-1}(1-p), \boldsymbol{u}) d\boldsymbol{u}.$$

By Theorem 1 of Nolan [26], for n sufficiently large

$$n^{1/2}(r_n(\alpha, \boldsymbol{u}) - r_{F_{\boldsymbol{Z}}}(\alpha, \boldsymbol{u})) = \int I_{H[\alpha,\boldsymbol{u}]}(z) d\nu_n(z) / g(\alpha, \boldsymbol{u}) + o_p(1)$$

uniformly in $\boldsymbol{u} \in S^{d-1}$ and thus,

$$\psi(z, F_D^{-1}(1-p), \boldsymbol{u}) = I_{H[\alpha,\boldsymbol{u}]}(z) / g(\alpha, \boldsymbol{u})|_{\alpha = F_{HD}^{-1}(1-p)},$$

where $g(\alpha, \boldsymbol{u}) = \boldsymbol{u}'\boldsymbol{v}(\alpha, \boldsymbol{u}) f_{\boldsymbol{v}(\alpha,\boldsymbol{u})'\boldsymbol{Z}}(r(\alpha, \boldsymbol{u})\boldsymbol{u}'\boldsymbol{v}(\alpha, \boldsymbol{u}))$, $H[\alpha, \boldsymbol{u}]$ is the closed half-space $H[z]$ at point $z = r(\alpha, \boldsymbol{u})\boldsymbol{u}$ such that $HD_{F_{\boldsymbol{Z}}}(z) = P(H[z])$, and $\boldsymbol{v}(\alpha, \boldsymbol{u})$ is the unit vector normal to $\partial H[\alpha, \boldsymbol{u}]$ and toward $H[\alpha, \boldsymbol{u}]$. It is clear that

$\boldsymbol{u} = \boldsymbol{v}(\alpha, \boldsymbol{u})$ and $f_{\boldsymbol{v}(\alpha,\boldsymbol{u})'\boldsymbol{Z}} = f_1$. Noting that $f_{\boldsymbol{Z}}(F_{\|\boldsymbol{Z}\|}^{-1}(p)\boldsymbol{u}) = h([F_{\|\boldsymbol{Z}\|}^{-1}(p)]^2)$ and using Lemma 3, we obtain

$$
\begin{aligned}
l_3(\boldsymbol{z}) &= \frac{F_{\|\boldsymbol{Z}\|}^{-1}(p)}{f_1(F_{\|\boldsymbol{Z}\|}^{-1}(p))} \int_{S^{d-1}} \boldsymbol{u} f(F_{\|\boldsymbol{Z}\|}^{-1}(p)\boldsymbol{u}) \,|\, J(F_{\|\boldsymbol{Z}\|}^{-1}(p), \boldsymbol{u}) \,|\, I_{H[F_{HD}^{-1}(1-p),\boldsymbol{u}]}(\boldsymbol{z}) d\boldsymbol{u} \\
&= \frac{\Gamma(d/2)}{2\pi^{d/2}} \frac{[F_{\|\boldsymbol{Z}\|}^{-1}(p)]^{-(d-2)} f_{\|\boldsymbol{Z}\|}(F_{\|\boldsymbol{Z}\|}^{-1}(p))}{f_1(F_{\|\boldsymbol{Z}\|}^{-1}(p))} \int_{S^{d-1}} \boldsymbol{u} \,|\, J(F_{\|\boldsymbol{Z}\|}^{-1}(p), \boldsymbol{u}) \,| \\
&\quad I_{H[F_{HD}^{-1}(1-p),\boldsymbol{u}]}(\boldsymbol{z}) d\boldsymbol{u}.
\end{aligned}
$$

i) If $HD_{F_{\boldsymbol{Z}}}(\boldsymbol{z}) > F_{HD}^{-1}(1-p)$, $\boldsymbol{z} \in C_{F_{\boldsymbol{Z}}}(p)$, $I_{H[F_{HD}^{-1}(1-p),\boldsymbol{u}]}(\boldsymbol{z}) = 0$ for any $\boldsymbol{u}$ and thus $l_3(\boldsymbol{z}) = \boldsymbol{0}$.

ii) If $HD_{F_{\boldsymbol{Z}}}(\boldsymbol{z}) = F_{HD}^{-1}(1-p)$, $I_{H[F_{HD}^{-1}(1-p),\boldsymbol{u}]}(\boldsymbol{z}) = 1$ for $\boldsymbol{u} = \frac{\boldsymbol{z}}{\|\boldsymbol{z}\|}$ only and thus $l_3(\boldsymbol{z}) = \boldsymbol{0}$ as well.

iii) if $HD_{F_{\boldsymbol{Z}}}(\boldsymbol{z}) < F_{HD}^{-1}(1-p)$, equivalently $\|\boldsymbol{z}\| > F_{\|\boldsymbol{Z}\|}^{-1}(p)$, $\boldsymbol{z} \neq \boldsymbol{0}$. Suppose that $\boldsymbol{z} = (z_1, 0, \cdots, 0)'$ with $z_1 > 0$. Using the polar coordinates representation for $\boldsymbol{u}$, *i.e.*,
$\boldsymbol{u} = (\cos\theta_1, \sin\theta_1\cos\theta_2, \cdots, \sin\theta_1\cdots\sin\theta_{d-2}\cos\theta_{d-1}, \sin\theta_1\cdots\sin\theta_{d-2}\sin\theta_{d-1})'$ and Lemma 4, we get for $d \geq 3$

$$
\begin{aligned}
l_3(\boldsymbol{z}) &= \frac{\Gamma(d/2)}{2\pi^{d/2}} \frac{[F_{\|\boldsymbol{Z}\|}^{-1}(p)]^{-(d-2)} f_{\|\boldsymbol{Z}\|}(F_{\|\boldsymbol{Z}\|}^{-1}(p))}{f_1(F_{\|\boldsymbol{Z}\|}^{-1}(p))} \int_0^{\arccos(F_{\|\boldsymbol{Z}\|}^{-1}(p)/\|\boldsymbol{z}\|)} \int_0^{\pi} \cdots \int_0^{\pi} \int_0^{2\pi} \\
&\quad (\cos\theta_1, \sin\theta_1\cos\theta_2, \cdots, \sin\theta_1\cdots\sin\theta_{d-2}\cos\theta_{d-1}, \\
&\quad \sin\theta_1\cdots\sin\theta_{d-2}\sin\theta_{d-1})' \\
&\quad [F_{\|\boldsymbol{Z}\|}^{-1}(p)]^{d-1} \sin^{d-2}\theta_1 \sin^{d-3}\theta_2 \cdots \sin\theta_{d-2} d\theta_{d-1} \cdots d\theta_2 d\theta_1 \\
&= \frac{F_{\|\boldsymbol{Z}\|}^{-1}(p) f_{\|\boldsymbol{Z}\|}(F_{\|\boldsymbol{Z}\|}^{-1}(p))[1 - (\frac{F_{\|\boldsymbol{Z}\|}^{-1}(p)}{\|\boldsymbol{z}\|})^2]^{(d-1)/2}}{(d-1) f_1(F_{\|\boldsymbol{Z}\|}^{-1}(p)) \int_0^{\pi} \sin^{d-2}\theta_1 d\theta_1} (1, 0, \cdots, 0)'.
\end{aligned}
$$

It is seen by a direct check that the expression also holds for $d = 2$.

If $\boldsymbol{z}$ is not on the positive part of the first coordinate axis, we take the rotation $\mathbf{O}_z$ such that $\mathbf{O}_z \boldsymbol{z} = (\|\boldsymbol{z}\|, 0, \cdots, 0)'$, equivalently $\boldsymbol{z} = \mathbf{O}_z^{-1}(\|\boldsymbol{z}\|, 0, \cdots, 0)'$. Since $\mathbf{O}_z \boldsymbol{Z} \sim E_d(h; \boldsymbol{0}, \mathbf{I}_{d\times d})$ as well,

$$
\begin{aligned}
l_3(\boldsymbol{z}) &= \mathbf{O}_z^{-1}[\frac{\Gamma(d/2)}{2\pi^{d/2}} \frac{[F_{\|\boldsymbol{Z}\|}^{-1}(p)]^{-(d-2)} f_{\|\boldsymbol{Z}\|}(F_{\|\boldsymbol{Z}\|}^{-1}(p))}{f_1(F_{\|\boldsymbol{Z}\|}^{-1}(p))} \int_{S^{d-1}} \boldsymbol{u} \,|\, J(F_{\|\boldsymbol{Z}\|}^{-1}(p), \boldsymbol{u}) \,| \\
&\quad I_{H[F_{HD}^{-1}(1-p),\boldsymbol{u}]}((\|\boldsymbol{z}\|, 0, \cdots, 0)') d\boldsymbol{u}] \\
&= \frac{F_{\|\boldsymbol{Z}\|}^{-1}(p) f_{\|\boldsymbol{Z}\|}(F_{\|\boldsymbol{Z}\|}^{-1}(p))[1 - (\frac{F_{\|\boldsymbol{Z}\|}^{-1}(p)}{\|\boldsymbol{z}\|})^2]^{(d-1)/2}}{(d-1) f_1(F_{\|\boldsymbol{Z}\|}^{-1}(p)) \int_0^{\pi} \sin^{d-2}\theta_1 d\theta_1} \frac{\boldsymbol{z}}{\|\boldsymbol{z}\|}.
\end{aligned}
$$

Combining i), ii), and iii), we obtain

$$l_3(z) = \frac{F_{\|z\|}^{-1}(p) f_{\|z\|}(F_{\|z\|}^{-1}(p))[1-(\frac{F_{\|z\|}^{-1}(p)}{\|z\|})^2]^{(d-1)/2} I_{\{\|z\|>F_{\|z\|}^{-1}(p)\}}(z)}{(d-1) f_1(F_{\|z\|}^{-1}(p)) \int_0^{\pi} \sin^{d-2}\theta_1 d\theta_1} \frac{z}{\|z\|}.$$

It is straightforward that $l_0(p) = \int_{C_{F_Z}(p)} dF_Z(z) = p$, and thus $l(z) = [l_1(z) + l_3(z)]/p = [c_1(\|z\|) + c_2(\|z\|)] \frac{z}{\|z\|}$. Since $\|Z\|$ and $U = \frac{Z}{\|Z\|}$ are independent, $E(U) = \mathbf{0}$, which implies $E(l(Z)) = \mathbf{0}$, $E(UU') = \frac{\mathbf{I}_{d\times d}}{d}$, and $c_1(\|Z\|)c_2(\|Z\|) = 0$,

$$\begin{aligned} Cov(l(Z)) &= Cov([c_1(\|Z\|) + c_2(\|Z\|)]U) \\ &= E([c_1(\|Z\|) + c_2(\|Z\|)]^2)\, E(UU') \\ &= E[c_1^2(\|Z\|) + c_2^2(\|Z\|)] \frac{\mathbf{I}_{d\times d}}{d}. \end{aligned}$$

Since $X = \Lambda^{1/2} Z + \theta$ and the halfspace depth is affine invariant, by Theorem 1 1), $l(X) = \Lambda^{1/2} l(Z)$. Therefore,

$$Cov(l(X)) = \Lambda^{1/2} Cov(l(Z)) (\Lambda^{1/2})' = E[c_1^2(\|Z\|) + c_2^2(\|Z\|)] \frac{\Lambda}{d}.$$

(2) First, when $w(t) = 1$ on $[0, \alpha_F^*]$, $k_2(z) = \mathbf{0}$ and $l_0(p) = p$. Secondly, $\mu_{F_Z}(p) = \mathbf{0}$ by Theorem 1 2). Thus, $k(z) = [k_1(z) + k_3(z) + k_4(z)]/p$ with

$$k_1(z) = (zz' - \Delta_{F_Z}(p)) I_{C_{F_Z}(p)}(z),$$

$$\begin{aligned} k_3(z) = \int_{S^{d-1}} &(r_{F_Z}^2(F_{HD}^{-1}(1-p), u)uu' - \Delta_{F_Z}(p)) f(r_{F_Z}(F_{HD}^{-1}(1-p), u)u) \\ &|J(r_{F_Z}(F_{HD}^{-1}(1-p), u), u)| \psi(z, F_{HD}^{-1}(1-p), u) du, \end{aligned}$$

and

$$\begin{aligned} k_4(z) = -p\Delta_{F_Z}^{(1)}(p)[&I_{C_{F_Z}(p)}(z) + \int_{S^{d-1}} f(r_{F_Z}(F_{HD}^{-1}(1-p), u)u) \\ &|J(r_{F_Z}(F_{HD}^{-1}(1-p), u), u)| \psi(z, F_{HD}^{-1}(1-p), u) du] \end{aligned}$$

In addition,

$$\begin{aligned} \Delta_{F_Z}(p) &= \frac{\int_{C_{F_Z}(p)} zz' dF(z)}{\int_{C_{F_Z}(p)} dF(z)} = \frac{1}{p} E(ZZ' I_{C_{F_Z}(p)}(Z)) \\ &= \frac{1}{p} E(\|Z\|^2 \frac{Z}{\|Z\|} \frac{Z'}{\|Z\|} I_{\{\|Z\| \le F_{\|Z\|}^{-1}(p)\}}(Z)) \\ &= \frac{1}{p} E[\|Z\|^2 I_{\{\|Z\| \le F_{\|Z\|}^{-1}(p)\}}(Z)] E(\frac{Z}{\|Z\|} \frac{Z'}{\|Z\|}) \\ &= [\frac{1}{p} \int_0^{F_{\|Z\|}^{-1}(p)} y^2 f_{\|Z\|}(y) dy] \frac{\mathbf{I}_{d\times d}}{d}, \end{aligned}$$

and then

$$\Delta^{(1)}_{F_{\boldsymbol{Z}}}(p)=\frac{[F^{-1}_{\|\boldsymbol{Z}\|}(p)]^2 p-\int_0^{F^{-1}_{\|\boldsymbol{Z}\|}(p)} y^2 f_{\|\boldsymbol{Z}\|}(y)dy}{p^2}\frac{\mathbf{I}_{d\times d}}{d}$$
$$=\frac{[F^{-1}_{\|\boldsymbol{Z}\|}(p)]^2}{dp}\mathbf{I}_{d\times d}-\frac{1}{p}\Delta_{F_{\boldsymbol{Z}}}(p).$$

Therefore,

$$k(\boldsymbol{z})=[k_1(\boldsymbol{z})+k_3(\boldsymbol{z})+k_4(\boldsymbol{z})]/p$$
$$=\{\boldsymbol{z}\boldsymbol{z}'I_{C_{F_{\boldsymbol{Z}}}(p)}(\boldsymbol{z})+\tau_1(\boldsymbol{z})-[\frac{(F^{-1}_{\|\boldsymbol{Z}\|}(p))^2}{d}\mathbf{I}_{d\times d}][I_{C_{F_{\boldsymbol{Z}}}(p)}(\boldsymbol{z})+\tau_2(\boldsymbol{z})]\}/p,$$

where $\tau_1(\boldsymbol{z})=\int_{S^{d-1}}(r^2_{F_{\boldsymbol{Z}}}(F^{-1}_{HD}(1-p),\boldsymbol{u})\boldsymbol{u}\boldsymbol{u}')f(r_{F_{\boldsymbol{Z}}}(F^{-1}_{HD}(1-p),\boldsymbol{u})\boldsymbol{u})$
$|J(r_{F_{\boldsymbol{Z}}}(F^{-1}_{HD}(1-p),\boldsymbol{u}),\boldsymbol{u})|\psi(\boldsymbol{z},F^{-1}_{HD}(1-p),\boldsymbol{u})d\boldsymbol{u}$

and

$$\tau_2(\boldsymbol{z})=\int_{S^{d-1}}f(r_{F_{\boldsymbol{Z}}}(F^{-1}_{HD}(1-p),\boldsymbol{u})\boldsymbol{u})|J(r_{F_{\boldsymbol{Z}}}(F^{-1}_{HD}(1-p),\boldsymbol{u}),\boldsymbol{u})|$$
$$\psi(\boldsymbol{z},F^{-1}_{HD}(1-p),\boldsymbol{u})d\boldsymbol{u}.$$

It is straightforward that $\boldsymbol{z}\boldsymbol{z}'I_{C_{F_{\boldsymbol{Z}}}(p)}(\boldsymbol{z})=\|\boldsymbol{z}\|^2\frac{\boldsymbol{z}}{\|\boldsymbol{z}\|}\frac{\boldsymbol{z}'}{\|\boldsymbol{z}\|}I_{\{\|\boldsymbol{z}\|\le F^{-1}_{\|\boldsymbol{Z}\|}(p)\}}(\boldsymbol{z})$ and $[\frac{(F^{-1}_{\|\boldsymbol{Z}\|}(p))^2}{d}\mathbf{I}_{d\times d}]I_{C_{F_{\boldsymbol{Z}}}(p)}(\boldsymbol{z})=\frac{[F^{-1}_{\|\boldsymbol{Z}\|}(p)]^2 I_{\{\|\boldsymbol{z}\|\le F^{-1}_{\|\boldsymbol{Z}\|}(p)\}}(\boldsymbol{z})}{d}\mathbf{I}_{d\times d}$. So, we only need to concretize $\tau_1(\boldsymbol{z})$ and $\tau_2(\boldsymbol{z})$. The same arguments as for $l_3(\boldsymbol{z})$ yield

$$\tau_1(\boldsymbol{z})=\int_{S^{d-1}}[F^{-1}_{\|\boldsymbol{Z}\|}(p)]^2\boldsymbol{u}\boldsymbol{u}')f(F^{-1}_{\|\boldsymbol{Z}\|}(p)\boldsymbol{u})|J(F^{-1}_{\|\boldsymbol{Z}\|}(p),\boldsymbol{u})|\psi(\boldsymbol{z},F^{-1}_{HD}(1-p),\boldsymbol{u})d\boldsymbol{u}$$
$$=\frac{\Gamma(d/2)}{2\pi^{d/2}}\frac{[F^{-1}_{\|\boldsymbol{Z}\|}(p)]^{-(d-1)}f_{\|\boldsymbol{Z}\|}(F^{-1}_{\|\boldsymbol{Z}\|}(p))}{f_1(F^{-1}_{\|\boldsymbol{Z}\|}(p))}\int_{S^{d-1}}([F^{-1}_{\|\boldsymbol{Z}\|}(p)]^2\boldsymbol{u}\boldsymbol{u}')|J(F^{-1}_{\|\boldsymbol{Z}\|}(p),\boldsymbol{u})|$$
$$I_{H[F^{-1}_{HD}(1-p),\boldsymbol{u}]}(\boldsymbol{z})d\boldsymbol{u},$$

which is $\mathbf{0}$ if $HD_{F_{\boldsymbol{Z}}}(\boldsymbol{z})\ge F^{-1}_{HD}(1-p)$.

If $HD_{F_{\boldsymbol{Z}}}(\boldsymbol{z})<F^{-1}_{D}(1-p)$, equivalently $\|\boldsymbol{z}\|>F^{-1}_{\|\boldsymbol{Z}\|}(p)$, $\boldsymbol{z}\ne\mathbf{0}$. Suppose that $\boldsymbol{z}=(z_1,0,\cdots,0)'$ with $z_1>0$. Using the polar coordinates representation for $\boldsymbol{u}$, *i.e.*,
$\boldsymbol{u}=(\cos\theta_1,\sin\theta_1\cos\theta_2,\cdots,\sin\theta_1\cdots\sin\theta_{d-2}\cos\theta_{d-1},\sin\theta_1\cdots\sin\theta_{d-2}\sin\theta_{d-1})'$, and Lemma 5, we get for $d\ge 3$

$$\tau_1(\boldsymbol{z})=\frac{\Gamma(d/2)}{2\pi^{d/2}}\frac{[F^{-1}_{\|\boldsymbol{Z}\|}(p)]^{-(d-1)}f_{\|\boldsymbol{Z}\|}(F^{-1}_{\|\boldsymbol{Z}\|}(p))}{f_1(F^{-1}_{\|\boldsymbol{Z}\|}(p))}\int_0^{\arccos(F^{-1}_{\|\boldsymbol{Z}\|}(p)/\|\boldsymbol{z}\|)}\int_0^{\pi}\cdots\int_0^{\pi}\int_0^{2\pi}$$
$$[F^{-1}_{\|\boldsymbol{Z}\|}(p)]^2\boldsymbol{u}\boldsymbol{u}'[F^{-1}_{\|\boldsymbol{Z}\|}(p)]^{d-1}\sin^{d-2}\theta_1\sin^{d-3}\theta_2\cdots\sin\theta_{d-2}d\theta_{d-1}\cdots d\theta_2 d\theta_1$$
$$=\frac{\Gamma(d/2)}{2\pi^{d/2}}\frac{[F^{-1}_{\|\boldsymbol{Z}\|}(p)]^2 f_{\|\boldsymbol{Z}\|}(F^{-1}_{\|\boldsymbol{Z}\|}(p))}{f_1(F^{-1}_{\|\boldsymbol{Z}\|}(p))}[\frac{2\pi^{d/2}}{\Gamma(d/2)}\frac{1}{d\int_0^{\pi}\sin^d\theta_1 d\theta_1}]$$
$$diag((d-1)\int_0^{\arccos(F^{-1}_{\|\boldsymbol{Z}\|}(p)/\|\boldsymbol{z}\|)}\cos^2\theta_1\sin^{d-2}\theta_1 d\theta_1,$$
$$\int_0^{\arccos(F^{-1}_{\|\boldsymbol{Z}\|}(p)/\|\boldsymbol{z}\|)}\sin^d\theta_1 d\theta_1,\cdots,\int_0^{\arccos(F^{-1}_{\|\boldsymbol{Z}\|}(p)/\|\boldsymbol{z}\|)}\sin^d\theta_1 d\theta_1)$$

$$
\begin{aligned}
&= \frac{[F_{\|z\|}^{-1}(p)]^2 f_{\|z\|}(F_{\|z\|}^{-1}(p))}{f_1(F_{\|z\|}^{-1}(p)) d\int_0^{\pi} \sin^d \theta_1 d\theta_1} \{[d\int_0^{\arccos(F_{\|z\|}^{-1}(p)/\|z\|)} \cos^2\theta_1 \sin^{d-2}\theta_1 d\theta_1 \\
&- \int_0^{\arccos(F_{\|z\|}^{-1}(p)/\|z\|)} \sin^{d-2}\theta_1 d\theta_1]\mathbf{1}\mathbf{1}' \\
&+ [\int_0^{\arccos(F_{\|z\|}^{-1}(p)/\|z\|)} \sin^d \theta_1 d\theta_1]\mathbf{I}_{d\times d}\},
\end{aligned}
$$

where $\mathbf{1} = (1,0,\cdots,0)'$. It is seen by a direct check that the expression also holds for $d = 2$.

If z is not on the positive part of the first coordinate axis, we take the rotation $\mathbf{O}_z$ such that $\mathbf{O}_z z = (\|z\|,0,\cdots,0)'$, equivalently $z = \mathbf{O}_z^{-1}(\|z\|,0,\cdots,0)'$. Since $\mathbf{\Sigma}_n(p) - \mathbf{\Sigma}_F(p)$ is affine equivariant, which implies $\tau_1(z)$ is affine equivariant, and $\mathbf{O}_z \boldsymbol{Z} \sim E_d(h;\mathbf{0},\mathbf{I}_{d\times d})$ as well,

$$
\begin{aligned}
\tau_1(z) &= \mathbf{O}_z^{-1}[\frac{\Gamma(d/2)}{2\pi^{d/2}} \frac{[F_{\|z\|}^{-1}(p)]^{-(d-1)} f_{\|z\|}(F_{\|z\|}^{-1}(p))}{f_1(F_{\|z\|}^{-1}(p))} \int_{S^{d-1}} ([F_{\|z\|}^{-1}(p)]^2 \boldsymbol{u}\boldsymbol{u}') \\
&\quad \left|J(F_{\|z\|}^{-1}(p),\boldsymbol{u})\right| I_{H[F_{HD}^{-1}(1-p),\boldsymbol{u}]}((\|z\|,0,\cdots,0)')d\boldsymbol{u}](\mathbf{O}_z^{-1})' \\
&= \frac{[F_{\|z\|}^{-1}(p)]^2 f_{\|z\|}(F_{\|z\|}^{-1}(p))}{f_1(F_{\|z\|}^{-1}(p)) d\int_0^{\pi} \sin^d \theta_1 d\theta_1} \{[d\int_0^{\arccos(F_{\|z\|}^{-1}(p)/\|z\|)} \cos^2\theta_1 \sin^{d-2}\theta_1 d\theta_1 \\
&\quad - \int_0^{\arccos(F_{\|z\|}^{-1}(p)/\|z\|)} \sin^{d-2}\theta_1 d\theta_1] \frac{z}{\|z\|}\frac{z'}{\|z\|} \\
&\quad + [\int_0^{\arccos(F_{\|z\|}^{-1}(p)/\|z\|)} \sin^d \theta_1 d\theta_1]\mathbf{I}_{d\times d}\}.
\end{aligned}
$$

Using integration by parts, we obtain

$$
\begin{aligned}
& d\int_0^{\arccos(F_{\|z\|}^{-1}(p)/\|z\|)} \cos^2\theta_1 \sin^{d-2}\theta_1 d\theta_1 - \int_0^{\arccos(F_{\|z\|}^{-1}(p)/\|z\|)} \sin^{d-2}\theta_1 d\theta_1 \\
&= \sin^{d-1}\theta_1 \cos\theta_1 \Big|_0^{\arccos(F_{\|z\|}^{-1}(p)/\|z\|)} \\
&= [1 - (\frac{F_{\|z\|}^{-1}(p)}{\|z\|})^2]^{(d-1)/2} (\frac{F_{\|z\|}^{-1}(p)}{\|z\|}).
\end{aligned}
$$

Thus,

$$
\begin{aligned}
\tau_1(z) &= \frac{[F_{\|z\|}^{-1}(p)]^2 f_{\|z\|}(F_{\|z\|}^{-1}(p)) I_{\{\|z\| > F_{\|z\|}^{-1}(p)\}}(z)}{d f_1(F_{\|z\|}^{-1}(p)) \int_0^{\pi} \sin^d \theta_1 d\theta_1} \\
&\quad \{[1 - (\frac{F_{\|z\|}^{-1}(p)}{\|z\|})^2]^{(d-1)/2} (\frac{F_{\|z\|}^{-1}(p)}{\|z\|}) \frac{z}{\|z\|}\frac{z'}{\|z\|} \\
&\quad + [\int_0^{\arccos(F_{\|z\|}^{-1}(p)/\|z\|)} \sin^d \theta_1 d\theta_1]\mathbf{I}_{d\times d}\}.
\end{aligned}
$$

In the same way as for $l_3(z)$, we have

$$
\begin{aligned}
\tau_2(z) &= \int_{S^{d-1}} f(r(F_D^{-1}(1-p),\boldsymbol{u})\boldsymbol{u}) \,|\, J(r(F_D^{-1}(1-p),\boldsymbol{u}),\boldsymbol{u}) \,|\, \psi(z, F_D^{-1}(1-p),\boldsymbol{u}) d\boldsymbol{u} \\
&= \frac{f_{\|z\|}(F_{\|z\|}^{-1}(p)) I_{\{\|z\| > F_{\|z\|}^{-1}(p)\}}(z)}{f_1(F_{\|z\|}^{-1}(p)) \int_0^{\pi} \sin^{d-2}\theta_1 d\theta_1} \int_0^{\arccos(F_{\|z\|}^{-1}(p)/\|z\|)} \sin^{d-2}\theta_1 d\theta_1.
\end{aligned}
$$

Plugging $\tau_1(z)$ and $\tau_2(z)$ into $k(z)$ yields

$$
\begin{aligned}
k(z) &= \frac{1}{p}\|z\|^2 \frac{z}{\|z\|}\frac{z'}{\|z\|} I_{\{\|z\|\le F_{\|Z\|}^{-1}(p)\}}(z) + \frac{[F_{\|Z\|}^{-1}(p)]^2 f_{\|Z\|}(F_{\|Z\|}^{-1}(p)) I_{\{\|z\|>F_{\|Z\|}^{-1}(p)\}}(z)}{pdf_1(F_{\|Z\|}^{-1}(p))\int_0^{\pi}\sin^d\theta_1 d\theta_1} \\
&\quad \times \{[1-(\frac{F_{\|Z\|}^{-1}(p)}{\|z\|})^2]^{(d-1)/2}(\frac{F_{\|Z\|}^{-1}(p)}{\|z\|})\frac{z}{\|z\|}\frac{z'}{\|z\|} \\
&\quad + [\int_0^{\arccos(F_{\|Z\|}^{-1}(p)/\|z\|)}\sin^d\theta_1 d\theta_1]\,\mathbf{I}_{d\times d}\} \\
&\quad - \frac{[F_{\|Z\|}^{-1}(p)]^2 I_{\{\|z\|\le F_{\|Z\|}^{-1}(p)\}}(z)}{pd}\mathbf{I}_{d\times d} \\
&\quad - \frac{[F_{\|Z\|}^{-1}(p)]^2 f_{\|Z\|}(F_{\|Z\|}^{-1}(p)) I_{\{\|z\|>F_{\|Z\|}^{-1}(p)\}}(z)}{pdf_1(F_{\|Z\|}^{-1}(p))\int_0^{\pi}\sin^{d-2}\theta_1 d\theta_1} \\
&\quad \times [\int_0^{\arccos(F_{\|Z\|}^{-1}(p)/\|z\|)}\sin^{d-2}\theta_1 d\theta_1]\,\mathbf{I}_{d\times d} \\
&= c_3(\|z\|)\frac{z}{\|z\|}\frac{z'}{\|z\|} + c_4(\|z\|)\,\mathbf{I}_{d\times d}.
\end{aligned}
$$

Let $U = \frac{Z}{\|Z\|}$. By Lemma 5.1 of Lopuhaä [4], $E(UU') = \frac{1}{d}\mathbf{I}_{d\times d}$ and $Cov(vec(UU')) = \frac{1}{d(d+2)}[\mathbf{I}_{d^2\times d^2} + \mathbf{K}_{d,d} + vec(\mathbf{I}_{d\times d})(vec(\mathbf{I}_{d\times d}))']$. Since $U = \frac{Z}{\|Z\|}$ and $\|Z\|$ are independent, we obtain

$$
\begin{aligned}
Cov(vec(k(Z))) &= E[vec(k(Z))(vec(k(Z)))'] - E(vec(k(Z))(E(vec(k(Z))))' \\
&= \frac{E(c_3^2(\|Z\|))}{d(d+2)}[I_{d^2\times d^2} + K_{d,d} + vec(I_{d\times d})(vec(I_{d\times d}))'] \\
&\quad + \frac{2}{d}E(c_3(\|Z\|)c_4(\|Z\|))vec(I_{d\times d})(vec(I_{d\times d}))' \\
&\quad + E(c_4^2(\|Z\|))vec(I_{d\times d})(vec(I_{d\times d}))' \\
&\quad - [\frac{1}{d}E(c_3(\|Z\|)) + E(c_4(\|Z\|))]^2 vec(I_{d\times d})(vec(I_{d\times d}))' \\
&= \sigma_1(I_{d^2\times d^2} + K_{d,d}) + \sigma_2 vec(I_{d\times d})(vec(I_{d\times d}))'.
\end{aligned}
$$

Since $X = \Lambda^{1/2}Z + \theta$ and the halfspace depth is affine invariant, by Theorem 1 1), $\sqrt{n}(\Sigma_{F_{X^n}}(p) - \Sigma_{F_X}(p)) = \Lambda^{1/2}[\sqrt{n}(\Sigma_{F_{Z^n}}(p) - \Sigma_{F_Z}(p)](\Lambda^{1/2})'$, where $Z^n = \Lambda^{-1/2}(X^n - \theta) = \{\Lambda^{-1/2}(X_1 - \theta), \cdots, \Lambda^{-1/2}(X_n - \theta)\}$. Then, the desired result follows by Lemma 5.2 of Lopuhaä [4]. □

Remark. From the proof of Corollary 1, we see that the theoretical trimming proportion for the α halfspace depth trimmed mean is $1 - p = P(HD_{F_Z}(Z) < \alpha) = P(F_1(-\|Z\|) < \alpha) = P(\|Z\| > -F_1^{-1}(\alpha)) = 1 - F_{\|Z\|}(-F_1^{-1}(\alpha))$.

Proof of Theorem 3. Since $\Sigma_n(p)$ is affine equivariant, $\Sigma_n(p)$ does not depend on θ. Thus, we assume without loss of generality that $\theta = \mathbf{0}$. Then, from the proof of Corollary 1, we have

$$\begin{aligned}\mathbf{\Sigma}_{F_X}(p) &= \mathbf{\Delta}_{F_X}(p) = \mathbf{\Lambda}^{1/2}\mathbf{\Delta}_{F_Z}(p)(\mathbf{\Lambda}^{1/2})' \\ &= \frac{1}{pd}E[\|\boldsymbol{Z}\|^2 I_{\{\|\boldsymbol{Z}\|\le F^{-1}_{\|\boldsymbol{Z}\|}(p)\}}(\boldsymbol{Z})]\mathbf{\Lambda} = c\mathbf{\Lambda},\end{aligned}$$

equivalently, $\mathbf{\Lambda} = \mathbf{\Sigma}_{F_X}(p)/c$.

By Corollary 1, $\sqrt{n}(vec(\mathbf{\Sigma}_n(p)) - vec(\mathbf{\Sigma}_F(p))) \xrightarrow{d} N_{d^2}(\mathbf{0}, \mathbf{V}_2)$ with

$$\begin{aligned}\mathbf{V}_2 &= \sigma_1(\mathbf{I}_{d^2\times d^2} + \mathbf{K}_{d,d})(\mathbf{\Lambda}\otimes\mathbf{\Lambda}) + \sigma_2 vec(\mathbf{\Lambda})vec(\mathbf{\Lambda})' \\ &= \frac{\sigma_1}{c^2}(\mathbf{I}_{d^2\times d^2} + \mathbf{K}_{d,d})(\mathbf{\Sigma}_{F_X}(p)\otimes\mathbf{\Sigma}_{F_X}(p)) + \frac{\sigma_2}{c^2}vec(\mathbf{\Sigma}_{F_X}(p))vec(\mathbf{\Sigma}_{F_X}(p))'.\end{aligned}$$

If $\mathbf{\Lambda} = \lambda\mathbf{I}_{d\times d}$, $\mathbf{\Sigma}_{F_X}(p) = c\lambda\mathbf{I}_{d\times d}$, and then the desired result follows by Theorem 2 of Tyler [39]. □

Proof of Theorem 4. $\boldsymbol{\mu}_{X^n}(p)$ and $\mathbf{\Sigma}_{X^n}(p)$ are obtained by trimming off $n - \lceil np\rceil$ data points in X^n and computing with the remaining $\lceil np\rceil$ points in $C_{X^n}(p)$. We first consider the case $n - \lceil np\rceil + 1 \le \lceil np\rceil - d$, which implies that $n - \lceil np\rceil + 1 \le (n-d+1)/2 \le \lceil np\rceil - d$, and show that $m = n - \lceil np\rceil + 1$ contaminating points $\{z_1,\cdots,z_m\}$ can break down $\boldsymbol{\mu}_{X^n}(p)$ and $\mathbf{\Sigma}_{X^n}(p)$. Replace m points of X^n with $\{z_1,\cdots,z_m\}$ such that each z_i has a lower depth than any remaining data points of X^n. Since we will only trim off $n - \lceil np\rceil$ points, at least one contaminating point (say z_1) must be in $C_{X^n_m}(p)$. Then, as $\|z_1\| \to \infty$, $\boldsymbol{\mu}_{X^n_m}(p)$ and $\mathbf{\Sigma}_{X^n_m}(p)$ explode since $\|z_1 w(PD_{F_n}(z_1))\| \to \infty$.

However, $m = n - \lceil np\rceil$ contaminating points $\{z_1,\cdots,z_m\}$ cannot break down $\boldsymbol{\mu}_{X^n}(p)$ and $\mathbf{\Sigma}_{X^n}(p)$. In fact, since $m = n - \lceil np\rceil < \lfloor (n-d+1)/2\rfloor$ and X^n is in general position, which implies that $c(\boldsymbol{u}'X^n) = d$ for any $\boldsymbol{u} \in S^{d-1} = \{\boldsymbol{u}\in\mathbb{R}^d : \|\boldsymbol{u}\| = 1\}$, by the result of Gather and Hilker [8], $\sigma = \mathrm{MAD}_d$ does not break down and thus $0 < \sigma(\boldsymbol{u}'X^n_m) < \infty$. Furthermore, since $m < \lfloor (n+1)/2\rfloor$, $\mu = \mathrm{Med}$ does not break down either and thus $\mu(\boldsymbol{u}'X^n_m) < \infty$ for any $\boldsymbol{u}\in S^{d-1}$. For any $z_i, i = 1,\cdots,m$, if $\|z_i\| \to \infty$,

$O_{X^n_m}(z_i) = \sup_{\boldsymbol{u}\in S^{d-1}} \left|\dfrac{\boldsymbol{u}'z_i - \mu(\boldsymbol{u}'X^n_m)}{\sigma(\boldsymbol{u}'X^n_m)}\right| \to \infty$, which implies $PD_{X^n_m}(z_i) \to 0$ and thus $z_i \notin C_{X^n_m}(p)$. Therefore, $\boldsymbol{\mu}_{X^n_m}(p)$ and $\mathbf{\Sigma}_{X^n_m}(p)$ do not explode. In addition, since $m = n - \lceil np\rceil \le \lceil np\rceil - d - 1$, $\lceil np\rceil - m \ge d+1$, which means that $C_{X^n_m}(p)$ contains at least $d+1$ points of X^n. Since X^n is in general position, $\mathbf{\Sigma}_{X^n_m}(p)$ does not implode either. Therefore,

$$\mathrm{RBP}(\boldsymbol{\mu}_n(p), X^n) = (n - \lceil np\rceil + 1)/n = \mathrm{RBP}(\mathbf{\Sigma}_n(p), X^n).$$

Now we consider the case $n - \lceil np\rceil + 1 > \lceil np\rceil - d$, which implies $n - \lceil np\rceil + 1 > (n-d+1)/2 > \lceil np\rceil - d$. For this case, $\boldsymbol{\mu}_n(p)$ breaks down if and only if μ = Med or σ = MAD$_d$ breaks down, which makes $\boldsymbol{\mu}_n(p)$ undefined. Thus, $\mathrm{RBP}(\boldsymbol{\mu}_n(p), X^n) = \min$ {the RBP of Med, the RBP of MAD$_d$}

$= \lfloor (n-d+1)/2 \rfloor / n$.

As for $\boldsymbol{\Sigma}_n(p)$, $m = \lceil np \rceil - d$ contaminating points can make it break down. In fact, we can replace the data points in $\boldsymbol{X}^n$ such that the points in $C_{\boldsymbol{X}_m^n}(p)$ consist of $m = \lceil np \rceil - d$ contaminating points and d points in $\boldsymbol{X}^n$, and all the contaminating points are in the $(d-1)$-dimensional affine subspace H determined by the d points in $\boldsymbol{X}^n$. Then, $\det(\boldsymbol{\Sigma}_{\boldsymbol{X}_m^n}(p)) = 0$ and thus $\boldsymbol{\Sigma}_{\boldsymbol{X}_m^n}(p)$ implodes.

However, $m = \lceil np \rceil - d - 1$ contaminating points cannot break down $\boldsymbol{\Sigma}_n(p)$. On one hand, since $C_{\boldsymbol{X}_m^n}(p)$ contains at least $\lceil np \rceil - m = d + 1$ points of $\boldsymbol{X}^n$ and $\boldsymbol{X}^n$ is in general position, $\boldsymbol{\Sigma}_{\boldsymbol{X}_m^n}(p)$ does not implode. On the other hand, since $n - \lceil np \rceil > m$, at least $n - \lceil np \rceil - m$ points of $\boldsymbol{X}^n$ must be outside of $C_{\boldsymbol{X}_m^n}(p)$. If any contaminating point $\boldsymbol{z}_i$ is in $C_{\boldsymbol{X}_m^n}(p)$, its depth must be greater than the depths of any points of $\boldsymbol{X}^n$ outside of $C_{\boldsymbol{X}_m^n}(p)$ and thus $\boldsymbol{z}_i$ must be bounded. Therefore, $\boldsymbol{\Sigma}_{\boldsymbol{X}_m^n}(p)$ does not explode either, and thus

$$\mathrm{RBP}(\boldsymbol{\Sigma}_n(p), \boldsymbol{X}^n) = (\lceil np \rceil - d)/n.$$

Combining the above arguments, we complete the proof. □