On Asymptotic Efficiency and Robustness of Multivariate Medians

by

Subhra Sankar Dhar[1] and Probal Chaudhuri[2]

Stat-Math Unit, Calcutta Centre,

Stat-Math Unit, Indian Statistical Institute, Calcutta

203, B. T. Road, Calcutta 700 108, India

e mails: subhra_r@isical.ac.in[1], probal@isical.ac.in[2]

Abstract

Univariate median is a well-known location estimator, which is $\sqrt{n}$-consistent, asymptotically Gaussian and affine equivariant. It is also a robust estimator of location with the highest asymptotic breakdown point (50%). While there are several versions of multivariate median proposed and extensively studied in the literature, many of those statistical properties of univariate median fail to hold for some of those multivariate medians. In this article, we make a comprehensive review of the known robustness and other asymptotic properties of different versions of multivariate median as well as some other well known multivariate location estimators with 50% breakdown points and derive some new results on asymptotic efficiency of those estimators. Our investigation shows that the affine equivariant version of spatial median obtained using the transformation-retransformation technique, which has 50% breakdown point, has optimal asymptotic properties among the popular versions of multivariate median considered in the literature for data following elliptic distributions with exponential and polynomial tails, which include multivariate normal and multivariate Cauchy distributions as special cases. Further, when compared with some other well-known robust estimates of location, the transformation-retransformation spatial median tends to have superior asymptotic performance in a number of cases.

Keywords and phrases: asymptotic normality, breakdown point, data depth, spatial median, transformation-retransformation technique.

1 Introduction

Some of the attractive statistical properties of sample median as an estimate of univariate location are its affine equivariance, high breakdown point (50%), $\sqrt{n}$-consistency and asymptotic normality under very general conditions. There are many proposals in the literature for generalizing the median in multi-dimension (see e.g., Small [34], DasGupta [17]). Among multivariate medians, the vector of co-ordinatewise median and the spatial median are perhaps the simplest ones (see e.g., Brown [5], Babu and Rao [4], Chaudhuri [14, 15]). By extending the property that the univariate median minimizes the sum of absolute deviations from the data points, Oja [27] proposed a multivariate version of median that minimizes the sum of the volumes of simplices formed by the data points and the median. Some other multivariate medians based on notions of data depth are half-space median (see Tukey [36]), simplicial volume median (see Oja [27]), simplicial median (see Liu [24]) and projection median (see Zuo [38]). Some of these versions of multivariate median are not affine equivariant (e.g., spatial median and co-ordinatewise median). However, all of them are $\sqrt{n}$-consistent estimate of multivariate location and some of them have asymptotic distributions that are non-Gaussian. Further, many of them have breakdown points much smaller than 50%.

Among the mentioned versions of multivariate median, except spatial median and co-ordinatewise median, all others are equivariant under arbitrary affine transformations. On the other hand, co-ordinatewise median and spatial median are asymptotically Gaussian, and they have 50% asymptotic breakdown point. It is well-known that spatial median and co-ordinatewise median do not perform well when the components of the data vector are correlated because of their lack of affine equivariance, and one may fix this problem by adopting the transformation-retransformation (TR) approach (see

Chakraborty [7]). Affine equivariant versions of co-ordinatewise and spatial medians (i.e., TR co-ordinatewise median and TR spatial median) have been extensively studied by Chakraborty and Chaudhuri [8, 9, 10], Chakraborty et al. [13].

In the following table, we give the asymptotic breakdown points and the asymptotic distributions of some of the affine equivariant multivariate versions of median.

Table 1: Asymptotic breakdown points and asymptotic distributions of multivariate medians

Estimator	Asymptotic breakdown point	Asymptotic distribution
TR spatial median	1/2 (Chakraborty and Chaudhuri [10])	Gaussian (Chakraborty et al. [13])
TR co-ordinatewise median	1/2 (Chakraborty and Chaudhuri [10])	Gaussian (Chakraborty and Chaudhuri [8])
Oja's median	0 (Oja et al. [29])	Gaussian (Arcones et al. [3])
Liu's median	$1/(d+2)$ (Chen [16])	Gaussian (Arcones et al. [3])
Tukey's median	1/3 (Donoho and Gasko [20])	Non-Gaussian (Masse [25])
Projection median	1/2 (Zuo [38])	Non-Gaussian (Zuo [38])

In this article, we make an attempt to identify that affine equivariant version of multivariate median, which has optimal asymptotic distribution as well as breakdown point. For this purpose, we consider multivariate data following elliptic distributions with exponential as well as polynomial tails and compare the asymptotic distributions of different estimates of the centre of elliptic symmetry, which possess 50% breakdown point and are $\sqrt{n}$-consistent with an asymptotically normal distribution. As we will see in subsequent sections, the asymptotic efficiency of TR spatial median is superior to most of the other robust estimates of multivariate location considered here for varying choices of the distribution of the data with different tail behaviors.

2 Comparison among robust estimators of multivariate location

We first compare the asymptotic dispersions of different robust estimators of multivariate location that include different versions of multivariate median mentioned in the preceding section as well as some other well known estimates based on different notions of multivariate trimming. We begin with the following theorem that established superiority of TR spatial median over TR co-ordinatewise median for data following elliptic distribution.

Theorem 1: *Let n i.i.d. observations be generated from an elliptically symmetric density function $(\det \Sigma)^{-1/2} g\{(\mathbf{X} - \theta)^T \Sigma^{-1} (\mathbf{X} - \theta)\}$, where Σ is a positive definite matrix, θ is a location parameter, and $g(.)$ is a univariate symmetric density function. If the asymptotic dispersions of TR spatial median and TR co-ordinatewise median are denoted by Σ_1/n and Σ_2/n respectively, we have $\Sigma_1 = \sigma_1^2 \Sigma$, $\Sigma_2 = \sigma_1^2 \Sigma$, where $\sigma_1^2 < \sigma_2^2$ for all $d \geq 2$.*

Among other multivariate medians, simplicial volume median (see Oja [27]) and simplicical median (see [24]) are also $\sqrt{n}$-consistent and have asymptotically Gaussian distribution. It follows from Arcones et al. [3] and Oja and Niinimaa [28] that the asymptotic efficiencies of simplicial volume median and simplicial median are exactly same as that of TR spatial median for multivariate normal distribution. However, both of these two versions of multivariate median have poor breakdown points. On the other hand, both of half space median (see Tukey [36]) and projection median (see Zuo [38]) have non-Gaussian asymptotic distributions.

2.1 Comparison among TR spatial median and some other robust estimators of multivariate location

We have already seen that TR spatial median has some optimal behavior among multivariate medians in terms of the asymptotic distribution as well as the breakdown point. Here we compare the asymptotic performance of TR spatial median with that of some other robust multivariate location estimators. Among different robust multivariate location estimators, MVE estimator is quite well-known, which is the center of the ellipsoid with smallest volume covering 50% of the data. Asymptotically a more efficient multivariate location estimator with 50% breakdown point is the MCD estimator. It is defined as the mean of the 50% data points for which the determinant of the empirical covariance matrix is minimal.

The multivariate trimmed mean based on different depth functions are also well-known in the literature. The population version of the α-trimmed mean based on projection depth is defined as
$PTM^{\alpha}(F) = \int_{\{x:PD(x)\geq\lambda^{\alpha}\}} xF(dx)/\int_{\{x:PD(x)\geq\lambda^{\alpha}\}} F(dx)$, where $PD(x)$ denotes the projection depth at x as defined in Zuo [39, p.2213], $F_{PD}(\lambda) = P(PD(x) \geq \lambda)$ and $\lambda^{\alpha} = \sup\{\lambda : F_{PD}(\lambda) \geq 1 - \alpha\}$ (see Zuo [37, 39]). For a sample of i.i.d. observations, F will be replaced by F_n, where F_n is the empirical cumulative distribution function. Another well known trimming based robust multivariate location estimator is the Stahel-Donoho (see Stahel [35], Donoho [19]) estimator, which is a weighted trimmed mean, where the weight is a function of the projection depth.

Table 2: Asymptotic breakdown points and asymptotic distributions of some robust multivariate location estimators

Estimator	Asymptotic breakdown point	Asymptotic distribution
Minimum volume ellipsoid (MVE)	1/2 (Rousseeuw [30])	Non-Gaussian (Davies [18])
Multivariate least trimmed squares	1/2 (Agullo et al. [1])	Has not been derived
Stahel-Donoho	1/2 (Zuo [40])	Gaussian (Zuo [40])
Minimum covariance determinant (MCD)	1/2 (Rousseeuw and Leroy [32])	Gaussian (Butler et al. [6])
α-trimmed mean based on projection depth	α (Zuo [37, p.318])	Gaussian (Zuo [39])
α-trimmed mean based on Tukey's depth	$\leq 1/3$ (Zuo [37, p.318])	Non-Gaussian (Masse [26])

Among the above mentioned estimators, the efficiency of Stahel-Donoho estimator depends on the weight function of the trimming proportion. Zuo et al. [40] showed that this estimator performs better than the TR spatial median for Gaussian model when a truncated exponential type weight function is used.

Here we compare asymptotic dispersions of MCD and 1/2-trimmed mean based on projection depth with that of TR spatial median. The asymptotic dispersion of an estimate of a multivariate location greatly depends on the behavior of the tail of the distribution of the data. The families of multivariate elliptic distributions with exponential and polynomial tails, which we consider here, consist of the densities of the form $f(\mathbf{x}) = \frac{d\Gamma(d/2)(\det\Sigma)^{-1/2}}{\pi^{d/2}\Gamma(1+d/2k)2^{1+d/2k}} e^{-\frac{\{(\mathbf{x}-\theta)^T\Sigma^{-1}(\mathbf{x}-\theta)\}^k}{2}}, \mathbf{x} \in R^d, k > 0$ (see Gomez et al. [23]) and $f(\mathbf{x}) = \frac{\Gamma(k)}{\pi^{d/2}\Gamma(k-d/2)} \frac{1}{\{1+(\mathbf{x}-\theta)^T\Sigma^{-1}(\mathbf{x}-\theta)\}^k}, \mathbf{x} \in R^d, k > d/2$ (see Fang et al. [21]), respectively. Multivariate normal and multivariate Cauchy distributions are included in these two families of distributions.

2.2 Results related to the multivariate exponential power family

We first investigate the efficiency of different location estimates under multivariate normal distributions. We state the following theorem which proves optimality of TR spatial median under Gaussian distribution.

Theorem 2: *For n i.i.d. observations from a multivariate normal density with location parameter θ and dispersion parameter Σ, if the asymptotic dispersions of TR spatial median, $1/2$-trimmed mean based on projection depth and MCD are denoted by Σ_1/n, Σ_3/n and Σ_4/n respectively, we have $\Sigma_1 = \sigma_1^2\Sigma$, $\Sigma_3 = \sigma_3^2\Sigma$, $\Sigma_4 = \sigma_4^2\Sigma$, where $\sigma_1^2 < \sigma_4^2 < \sigma_3^2$ for all $d \geq 2$.*

Next, we give a result related to the family of multivariate elliptic distributions with exponential tails, which consists of densities of the form $f(\mathbf{x}) = \frac{d\Gamma(d/2)(\det \Sigma)^{-1/2}}{\pi^{d/2}\Gamma(1+d/2k)2^{1+d/2k}} e^{-\frac{\{(\mathbf{x}-\theta)^T\Sigma^{-1}(\mathbf{x}-\theta)\}^k}{2}}, \mathbf{x} \in R^d, k > 0$ (see Gomez et al. [23]). Note that $k = 1$ corresponds to d-variate Gaussian distribution.

Theorem 3: *For n i.i.d. observations from any density in the multivariate exponential power family described above, if the asymptotic dispersions of TR spatial median, $1/2$-trimmed mean based on projection depth and MCD are denoted by Σ_1/n, Σ_3/n and Σ_4/n respectively, we have $\Sigma_1 = \sigma_1^2\Sigma$, $\Sigma_3 = \sigma_3^2\Sigma$ and $\Sigma_4 = \sigma_4^2\Sigma$. Here $\sigma_1^2 < \sigma_3^2$ for all $d \geq 2$ and for all k satisfying $\frac{2^{d/2k}\Gamma(d/2k)}{4k\exp(m^{2k}/2)} > \frac{\Gamma(d/2)}{5d\pi^{d/2}}$, where m is the $3/4$-th quantile of the marginal density function. Further, $\sigma_1^2 < \sigma_4^2$ for all $d \geq 2$ and for all k satisfying $\frac{2(d-1)^2}{d^2} > \frac{(2k/d)^{1/k}\{\Gamma(d/2k)\}^2}{\{\Gamma((d-1)/2k)\}^2}$.*

As far as we could check by direct numerical evaluation, we have seen that $\frac{2^{d/2k}\Gamma(d/2k)}{4k\exp(m^{2k}/2)} > \frac{\Gamma(d/2)}{5d\pi^{d/2}}$ for all $d \geq 2$ and when either $.05 \leq k \leq 3$ or $k \geq 6.5$. We have checked that the inequality fails to hold for some values of $d \geq 2$ and $3 < k < 6.5$. Also, $\frac{2(d-1)^2}{d^2} > \frac{(2k/d)^{1/k}\{\Gamma(d/2k)\}^2}{\{\Gamma((d-1)/2k)\}^2}$ for all $d \geq 2$ and $k \geq .05$ as far as we could numerically verify.

2.3 Results related to the multivariate polynomial tail family

We first state a result that compares the asymptotic dispersions of TR spatial median, 1/2-trimmed mean based on projection depth and MCD for multivariate Cauchy distribution, which is a special case in the multivariate polynomial tail family corresponding to $k = (d+1)/2$ (see Fang et al. [21]).

Theorem 4: *For n i.i.d. observations from multivariate Cauchy density with location parameter θ and dispersion parameter Σ, if the asymptotic dispersions of TR spatial median, 1/2-trimmed mean based on projection depth and MCD are denoted by Σ_1/n, Σ_3/n and Σ_4/n respectively, we have $\Sigma_1 = \sigma_1^2\Sigma$, $\Sigma_3 = \sigma_3^2\Sigma$ and $\Sigma_4 = \sigma_4^2\Sigma$. Here $\sigma_1^2 < \sigma_3^2$ for $2 \le d \le 11$, and $\sigma_1^2 > \sigma_3^2$ for all $d \ge 12$. Further, we have $\sigma_1^2 < \sigma_4^2$ for all $d \ge 2$.*

Next, we give a result related to the family of multivariate elliptic distributions with polynomial tails having densities of the form
$f(\mathbf{x}) = \frac{\Gamma(k)}{\pi^{d/2}\Gamma(k-d/2)} \frac{1}{\{1+(\mathbf{x}-\theta)^T\Sigma^{-1}(\mathbf{x}-\theta)\}^k}, \mathbf{x} \in R^d, k > d/2$ (see Fang et al. [21]).

Theorem 5: *For n i.i.d. observations from any density in the multivariate polynomial tail family described above, if the asymptotic dispersions of TR spatial median, 1/2-trimmed mean based on projection depth and MCD are denoted by Σ_1/n, Σ_3/n and Σ_4/n respectively, we have $\Sigma_1 = \sigma_1^2\Sigma$, $\Sigma_3 = \sigma_3^2\Sigma$ and $\Sigma_4 = \sigma_4^2\Sigma$. Here $\sigma_1^2 < \sigma_3^2$ for all $d \ge 2$ and for all k satisfying $\frac{(2k-d)\Gamma(k-d/2)}{2^k} \ge \frac{2(d+1)}{d}$. Further, $\sigma_1^2 < \sigma_4^2$ for all $d \ge 2$ and for all $k \ge \frac{(d+3)/2+\sqrt{(d+3)^2/4-4\{(d+1)/2\}\{(\pi-2)/\pi\}}}{2(1-2/\pi)} \approx (4/3)d$.*

As far as we could numerically verify using Starlings approximation, $\frac{(2k-d)\Gamma(k-d/2)}{2^k} \ge \frac{2(d+1)}{d}$ for all $d \ge 2$ and for all $k \ge (5/2)d$.

3 Some Concluding Remarks

In terms of the asymptotic efficiency, the TR spatial median performs best among multivariate medians having Gaussian asymptotic distribution with 50% breakdown point. Recall that Tukey's median (see Masse [25]) and projection median (see Zuo [38]) have non-Gaussian asymptotic distributions. Those asymptotic distributions can only be described in terms of the distributions of the infimum or the supremum of some Gaussian processes, and they are not tractable any further. Also, quantiles of those distributions cannot be worked out in a convenient form, which makes Tukey's median and projection median difficult to use in asymptotic inference. For these reasons, we excluded Tukey's median and projection median in our comparative study of multivariate medians. Among other multivariate location estimators, the 1/2-trimmed mean based on projection depth is asymptotically more efficient than the TR spatial median for large dimensions in the case of multivariate Cauchy distribution. The MCD estimator does not perform well in all the cases considered here from the point of view of the asymptotic efficiency.

One has to choose the transformation matrix appropriately in order to achieve the optimal asymptotic efficiency of the TR spatial median. The procedure for obtaining the optimal transformation matrix is discussed in Chakraborty et al. [13]. However, the optimization problem described in Chakraborty et al. [13, p.770] cannot be solved exactly in practice even for moderately large dimension and sample size. Chakraborty and Chaudhuri [11, 12] proposed probabilistic optimization algorithms to solve this optimization problem. These authors demonstrated satisfactory performance of these algorithms by means of extensive theoretical study and empirical investigation.

4 Appendix: Proofs

Proof of Theorem 1: For n i.i.d. observations coming from an elliptically symmetric density function $(\det\Sigma)^{-1/2}g\{(\mathbf{X}-\theta)^T\Sigma^{-1}(\mathbf{X}-\theta)\}$, where Σ is a positive definite matrix, θ is a location parameter, and $g(.)$ is a univariate symmetric density function, the $\sqrt{n}$-consistency and the asymptotic normality of the TR spatial median and the TR co-ordinatewise median have been derived by Chakraborty and Chaudhuri [8] and Chakraboty et al. [13] respectively. From the results in those papers, we have

$$\Sigma_1 = \sigma_1^2\Sigma = \frac{d\{\Gamma(\frac{d}{2})\}^2}{f_1^2(0)\pi(d-1)^2\{\Gamma(\frac{d-1}{2})\}^2}\Sigma$$

and

$$\Sigma_2 = \sigma_2^2\Sigma = \frac{1}{4f_1^2(0)}\Sigma,$$

where d is the dimension, and f_1 is the marginal density function. In order to prove $\sigma_1^2 < \sigma_2^2$, we have to show that

$$\frac{d(\Gamma(d/2))^2}{\pi(d-1)^2\{\Gamma((d-1)/2)\}^2} < \frac{1}{4} \Leftrightarrow \frac{d(\Gamma(d/2))^2}{\{\Gamma((d+1)/2)\}^2} < \pi.$$

Define $f(d) = \frac{d\{\Gamma(d/2)\}^2}{\{\Gamma((d+1)/2)\}^2}$. Now, we have

$$\begin{aligned} f(d+2) &= \frac{(1+2/d)^2(d/2)^2}{\{(d+1)/2\}^2}f(d) \\ &= \frac{d^2+2d}{d^2+2d+1}f(d) \\ &< f(d), \end{aligned}$$

and $f(1) = \pi$ and $f(2) = 8/\pi < \pi$. The theorem is now proved using the principle of mathematical induction. $\square$

Proof of Theorem 2: The $\sqrt{n}$-consistency and the asymptotic normality of the MCD estimator have been derived by Butler et al. [6], and similar results for the multivariate trimmed mean based on projection depth are given in Zuo [39], when data follows elliptically symmetric distribution as in Theorem 1. From the results in those papers, we have $\Sigma_3 = \sigma_3^2\Sigma$, where

$$\begin{aligned}\sigma_3^2 &= \frac{\{E(x_1 I(||\mathbf{x}|| \le m)) + \int_{S_{d-1}} mf(m)|J(u,m)|K(\mathbf{x}, mu)u_1 du\}^2}{\{P(||\mathbf{x}|| \le m)\}^2} \\ &= \frac{k_1 + k_2 + k_3}{\{P(||\mathbf{x}|| \le m)\}^2} \text{ (say).}\end{aligned}$$

Here m is the 3/4-th quantile of the univariate marginal density function, f_1 is the marginal density function, x_1 is the first component of $\mathbf{x}$, $R = \{\mathbf{x} : x_1 \ge 0, ||\mathbf{x}|| \le m\}$, $k_1 = \{E(x_1 I(||\mathbf{x}|| \le m))\}^2$, $k_2 = 2E(x_1 I(||\mathbf{x}|| \le m))\{\int_{S_{d-1}} mf(m)|J(u,m)|K(\mathbf{x}, mu)u_1 du\}$, $k_3 = \{\int_{S_{d-1}} mf(m)|J(u,m)|K(\mathbf{x}, mu)u_1 du\}^2$, S_{d-1}, $J(.)$ and $K(.)$ are as defined in Zuo [39, Corollary 5], and

$$\Sigma_4 = \sigma_4^2\Sigma = \frac{d\Gamma(d/2) \int_0^{r(1/2)} y^{k+1} g(r^2)dy}{8\pi^{d/2}\left\{\int_0^{r(1/2)} y^{k+1} g'(r^2)dr\right\}^2}\Sigma,$$

where $g'(.)$ is the first derivative of $g(.)$, and $r(1/2)$ is such that $\frac{2\pi^{d/2}}{\Gamma(d/2)} \int_0^{r(1/2)} y^{d-1} g(y^2)dy = 1/2$ (see Butler et al. [6, p.1387]). Note that it follows from the proof of Theorem 1 that $\Sigma_1 = \sigma_1^2\Sigma$. We split the proof into two parts. In the first part, we prove that $\sigma_1^2 < \sigma_4^2$.

In order to prove $\sigma_1^2 < \sigma_4^2$, we have to show that

$$\frac{\Gamma(d/2+1)}{\int_0^{m_1^2/2} \exp(-z)z^{d/2}dz} > \frac{2d\{\Gamma(d/2)\}^2}{(d-1)^2\{\Gamma((d-1)/2)\}^2}$$

$$\Leftarrow \quad \frac{\Gamma(d/2+1)}{\int\limits_0^{m_1^2/2} \exp(-z) z^{d/2} dz} > \frac{\pi}{2},$$

where $m_1^2/2$ is the median of the gamma distribution with scale parameter 1 and shape parameter $d/2$, which follows from Butler et al. [6, p.1387]. The last implication follows from the fact that $\frac{d\{\Gamma(d/2)\}^2}{(d-1)^2\{\Gamma((d-1)/2)\}^2} \leq \frac{\pi}{4}$ for all $d \geq 2$. As gamma distribution is a positively skewed distribution, $m_1^2/2 \leq d/2$. This implies that it is enough to prove that

$$\frac{\Gamma(d/2+1)}{\int\limits_0^{d/2} \exp(-z) z^{d/2} dz} > \frac{\pi}{2}. \tag{1}$$

Now, we assume that the dimension d is an even integer, i.e., $d = 2k$, where k is an integer. Using the relationship between the incomplete gamma function and the cumulative distribution function of the Poisson distribution (see Feller [22]), proving (1) is equivalent to showing that

$$\exp(-k)(1 + k + k^2/2! + \cdots + k^k/k!) > 1 - (2/\pi).$$

Let us define $g(k) = exp(-k)(1 + k + k^2/2! + \cdots + k^k/k!)$, which is an increasing function of k, and $g(1) = \exp(-1) > 1 - (2/\pi)$. This implies that $g(k) > 1 - (2/\pi)$ for all k. This completes the proof for the case $d = 2k$. The proof remains same if d is an odd integer.

Next, we prove that $\sigma_4^2 < \sigma_3^2$. As the first and the second terms of the numerator of σ_3^2 are positive (using Corollary 5 of Zuo [39]), it is enough to show that

$$\frac{\frac{4\pi^{d-1}\{g(m^2)\}^2 m^{2d}}{\{\Gamma((d+1)/2)\}^2}\left\{\frac{1}{16 f_1^2(m)} + \frac{1}{4 f_1^2(0)}\right\}}{\left[\frac{\pi^{d/2}\int_0^{m^2} y^{d/2-1}\exp(-y)dy}{\Gamma(d/2)}\right]^2} > \frac{\Gamma(d/2+1)}{\int\limits_0^{m_1^2/2} \exp(-z) z^{d/2} dz}$$

$$\Leftarrow \quad \frac{\frac{4\pi^{d-1}\{g(m^2)\}^2 m^{2d}}{\{\Gamma((d+1)/2)\}^2}\{\frac{1}{16f_1^2(0)}+\frac{1}{4f_1^2(0)}\}}{\left[\frac{\pi^{d/2}\int_0^{m^2} y^{d/2-1}\exp(-y)dy}{\Gamma(d/2)}\right]^2} > \frac{\Gamma(d/2+1)}{\int\limits_0^{m_1^2/2}\exp(-z)z^{d/2}dz}$$

(since $f_1^2(0) \geq f_1^2(m)$)

$$\Leftarrow \quad \frac{\{\exp(m^2)+4\}\{\Gamma(\frac{d}{2})\}^2 d^2\pi^{\frac{d}{2}}2^{\frac{d}{2}}}{8\{\Gamma(\frac{d+1}{2})\}^2} > \frac{\Gamma(\frac{d}{2}+1)}{\int\limits_0^{m_1^2/2}\exp(-z)z^{d/2}dz} \tag{2}$$

(since $\exp(-y) \leq 1$)

The RHS of inequality (2) is a decreasing function of d, and let the LHS of inequality (2) be denoted by $g(d)$, i.e.,

$$g(d) = \frac{\{\exp(m^2)+4\}\{\Gamma(\frac{d}{2})\}^2 d^2\pi^{\frac{d}{2}}2^{\frac{d}{2}}}{8\{\Gamma(\frac{d+1}{2})\}^2}.$$

Now, we have

$$\begin{aligned} g(d+2) &= c_d\frac{\{\exp(m^2)+4\}\{\Gamma(\frac{d}{2})\}^2 d^2\pi^{\frac{d}{2}}2^{\frac{d}{2}}}{8\{\Gamma(\frac{d+1}{2})\}^2} \\ &= c_d g(d), \end{aligned}$$

where $c_d = \frac{d^2(1+2/d)^2 d^2 2\pi}{(d+1)^2}$. It can be shown that $c_d > 1$ for all $d \geq 2$ by using the principle of mathematical induction. Also, $g(2)$ and $g(3)$ are larger than the RHS of inequality (2) for $d = 2$. This proves the theorem. □

Proof of Theorem 3: It follows from the proofs of Theorems 1 and 2 that $\Sigma_1 = \sigma_1^2\Sigma$, $\Sigma_3 = \sigma_3^2\Sigma$ and $\Sigma_4 = \sigma_4^2\Sigma$.

In order to prove $\sigma_1^2 < \sigma_3^2$, it is enough to show that

$$\frac{\frac{4\pi^{d-1}d\Gamma(\frac{d}{2})\exp(-m^{\frac{2k}{2}})m^{2d}}{\pi^{\frac{d}{2}}\Gamma(1+\frac{d}{2k})2^{1+\frac{d}{2k}}\{\Gamma(\frac{d+1}{2})\}^2}\{\frac{1}{16f_1^2(m)}+\frac{1}{4f_1^2(0)}\}}{\left[\{\Gamma(\frac{d}{2k})\}^{-1}\int\limits_0^{m^{\frac{2k}{2}}}\exp(-y)y^{\frac{d}{2k}-1}dy\right]^2} > \frac{d(\Gamma(\frac{d}{2}))^2}{f_1^2(0)\pi(d-1)^2(\Gamma(\frac{d-1}{2}))^2}$$

since the first and the second terms of the numerator of σ_3^2 are non-negative (using Corollary 5 of Zuo [39]).

Now,

$$\frac{\frac{4\pi^{d-1}d\Gamma(\frac{d}{2})\exp(-m^{\frac{2k}{2}})m^{2d}}{\pi^{\frac{d}{2}}\Gamma(1+\frac{d}{2k})2^{1+\frac{d}{2k}}\{\Gamma(\frac{d+1}{2})\}^2}\{\frac{1}{16f_1^2(m)}+\frac{1}{4f_1^2(0)}\}}{\left[\{\Gamma(\frac{d}{2k})\}^{-1}\int\limits_0^{m^{\frac{2k}{2}}}\exp(-y)y^{\frac{d}{2k}-1}dy\right]^2} > \frac{d\{\Gamma(\frac{d}{2})\}^2}{f_1^2(0)\pi(d-1)^2\{\Gamma(\frac{d-1}{2})\}^2}$$

$$\Leftarrow \frac{\frac{4\pi^{d-1}d\Gamma(\frac{d}{2})\exp(-m^{\frac{2k}{2}})m^{2d}}{\pi^{\frac{d}{2}}\Gamma(1+\frac{d}{2k})2^{1+\frac{d}{2k}}\{\Gamma(\frac{d+1}{2})\}^2}\{\frac{1}{16f_1^2(0)}+\frac{1}{4f_1^2(0)}\}}{\left[\{\Gamma(\frac{d}{2k})\}^{-1}\int\limits_0^{m^{\frac{2k}{2}}}\exp(-y)y^{\frac{d}{2k}-1}dy\right]^2} > \frac{d\{\Gamma(\frac{d}{2})\}^2}{f_1^2(0)\pi(d-1)^2\{\Gamma(\frac{d-1}{2})\}^2}$$

(since $f_1^2(0) \geq f_1^2(m)$)

$$\Leftrightarrow \frac{\frac{4\pi^{d-1}d\Gamma(\frac{d}{2})\exp(-m^{\frac{2k}{2}})m^{2d}}{\pi^{\frac{d}{2}}\Gamma(1+\frac{d}{2k})2^{1+\frac{d}{2k}}\{\Gamma(\frac{d+1}{2})\}^2}\frac{5}{16f_1^2(0)}}{\left[\{\Gamma(\frac{d}{2k})\}^{-1}\int\limits_0^{m^{\frac{2k}{2}}}\exp(-y)y^{\frac{d}{2k}-1}dy\right]^2} > \frac{d(\Gamma(\frac{d}{2}))^2}{4f_1^2(0)\pi(\Gamma(\frac{d+1}{2}))^2}$$

$$\Leftrightarrow \frac{5\exp(-m^{\frac{2k}{2}})m^{2d}\pi^{d/2}\{\Gamma(d/2k)\}^2}{2^{1+\frac{d}{2k}}\Gamma(1+\frac{d}{2k})\left\{\int\limits_0^{m^{\frac{2k}{2}}}\exp(-y)y^{\frac{d}{2k}-1}dy\right\}^2} > \Gamma(d/2)$$

$$\Leftrightarrow \frac{5\exp(-m^{\frac{2k}{2}})m^{2d}\pi^{d/2}\{\Gamma(d/2k)\}^2}{2\times 2^{\frac{d}{2k}}\frac{d}{2k}\Gamma(\frac{d}{2k})\left\{\int\limits_0^{m^{\frac{2k}{2}}}\exp(-y)y^{\frac{d}{2k}-1}dy\right\}^2} > \Gamma(d/2)$$

$$\Leftrightarrow \frac{5\exp(-m^{\frac{2k}{2}})m^{2d}\pi^{d/2}k\Gamma(d/2k)}{2^{\frac{d}{2k}}d\left\{\int\limits_0^{m^{\frac{2k}{2}}}\exp(-y)y^{\frac{d}{2k}-1}dy\right\}^2} > \Gamma(d/2)$$

$$\Leftarrow \quad \frac{5\exp(-m^{\frac{2k}{2}})m^{2d}\pi^{d/2}k\Gamma(d/2k)}{2^{\frac{d}{2k}}d\left(\int\limits_0^{m^{\frac{2k}{2}}} y^{\frac{d}{2k}-1}dy\right)^2} > \Gamma(d/2) \text{ (since } \exp(-y) \leq 1)$$

$$\Leftrightarrow \quad \frac{2^{d/2k}\Gamma(d/2k)}{4k\exp(m^{2k}/2)} > \frac{\Gamma(d/2)}{5d\pi^{d/2}}.$$

This completes the proof of $\sigma_1^2 < \sigma_3^2$.

Next, we prove the other part, i.e., $\sigma_1^2 < \sigma_4^2$. For any density in the multivariate exponential power family, it follows from Butler et al. [6] that

$$\sigma_4^2 = \frac{d\Gamma(d/2k) \int\limits_0^{m_1^{2k}/2} \exp(-y)y^{d/2k+1/k-1}dy}{k^2 2^{2-1/k}\left\{\int\limits_0^{m_1^{2k}/2} \exp(-y)y^{d/2k}dy\right\}^2},$$

where $m_1^{2k}/2$ is the median of gamma distribution with scale parameter 1 and shape parameter $d/2k$ (see Butler et al. [6, p.1387]).

Gamma distribution is a positively skewed distribution. This implies that the median of Gamma distribution is smaller than its mean. Using this fact, we have $m_1^{2k}/2 \leq (d/2k)$.

Using Theorem 2.10 of Fang et al. [21, p.36] and putting $x_i = 0$ there, we have

$$f_1(0) = \frac{d\Gamma(d/2)\Gamma((d-1)/2k)}{k\sqrt{\pi}\Gamma((d-1)/2)\Gamma(1+d/2k)2^{1+d/2k}}.$$

This implies that

$$\sigma_1^2 = \frac{d\{\Gamma(\frac{d}{2})\}^2}{f_1^2(0)\pi(d-1)^2\{\Gamma(\frac{d-1}{2})\}^2} = \frac{2^{1/k}d\{\Gamma(d/2k)\}^2}{(d-1)^2\{\Gamma((d-1)/2k)\}^2}.$$

Now, we have to show that

$$\frac{d\Gamma(d/2k)\int\limits_0^{m_1^{2k}/2}\exp(-y)y^{d/2k+1/k-1}dy}{k^2 2^{2-1/k}\left\{\int\limits_0^{m_1^{2k}/2}\exp(-y)y^{d/2k}dy\right\}^2} > \frac{2^{1/k}d\{\Gamma(d/2k)\}^2}{(d-1)^2\{\Gamma((d-1)/2k)\}^2}$$

$$\Leftrightarrow \quad \frac{\Gamma(d/2k)\int\limits_0^{m_1^{2k}/2}\exp(-y)y^{d/2k+1/k-1}dy}{4k^2\left\{\int\limits_0^{m_1^{2k}/2}\exp(-y)y^{d/2k}dy\right\}^2} > \frac{d\{\Gamma(d/2k)\}^2}{(d-1)^2\{\Gamma((d-1)/2k)\}^2}$$

$$\Leftarrow \quad \frac{\Gamma(d/2k)\int\limits_0^{m_1^{2k}/2}\exp(-y)y^{d/2k}dy}{4k^2(m_1^{2k}/2)^{1-1/k}\left\{\int\limits_0^{m_1^{2k}/2}\exp(-y)y^{d/2k}dy\right\}^2} > \frac{d\{\Gamma(d/2k)\}^2}{(d-1)^2\{\Gamma((d-1)/2k)\}^2}$$

(since $y \leq m_1^{2k}/2$)

$$\Leftrightarrow \quad \frac{\Gamma(d/2k)}{4k^2(m_1^{2k}/2)^{1-1/k}\int\limits_0^{m_1^{2k}/2}\exp(-y)y^{d/2k}dy} > \frac{d\{\Gamma(d/2k)\}^2}{(d-1)^2\{\Gamma((d-1)/2k)\}^2}$$

$$\Leftarrow \quad \frac{\Gamma(d/2k)}{4k^2(d/2k)^{1-1/k}\int\limits_0^{m_1^{2k}/2}\exp(-y)y^{d/2k}dy} > \frac{d\{\Gamma(d/2k)\}^2}{(d-1)^2\{\Gamma((d-1)/2k)\}^2}$$

(since $m_1^{2k}/2 \leq d/2k$)

$$\Leftarrow \quad \frac{\Gamma(d/2k)}{4k^2(\frac{d}{2k})^{1-\frac{1}{k}}\int\limits_0^{m_1^{2k}/2}\exp(-y)y^{d/2k-1}dy(\frac{m_1^{2k}}{2})} > \frac{d\{\Gamma(d/2k)\}^2}{(d-1)^2\{\Gamma((d-1)/2k)\}^2}$$

(since $y \leq m_1^{2k}/2$)

$$\Leftarrow \quad \frac{\Gamma(d/2k)}{4k^2(\frac{d}{2k})^{2-\frac{1}{k}}\int\limits_0^{m_1^{2k}/2}\exp(-y)y^{d/2k-1}dy} > \frac{d\{\Gamma(d/2k)\}^2}{(d-1)^2\{\Gamma((d-1)/2k)\}^2}$$

(since $m_1^{2k}/2 \leq d/2k$)

$$\Leftrightarrow \quad \frac{\Gamma(d/2k)}{4k^2(\frac{d}{2k})^{2-\frac{1}{k}}(\frac{1}{2})\Gamma(\frac{d}{2k})} > \frac{\{\Gamma(d/2k)\}^2}{(d-1)^2\{\Gamma(\frac{d-1}{2k})\}^2}$$

$$\text{(since } \int_0^{m_1^{2k}/2} \exp(-y)y^{d/2k-1}dy = \tfrac{1}{2}\Gamma(d/2k))$$

$$\Leftrightarrow \quad \frac{2(d-1)^2}{d^2} > \frac{(2k/d)^{1/k}\{\Gamma(d/2k)\}^2}{\{\Gamma((d-1)/2k)\}^2}.$$

This completes the proof. □

Proof of Theorem 4: Recall that $\Sigma = \sigma_1^2\Sigma$, $\Sigma_3 = \sigma_3^2\Sigma$ and $\Sigma_4 = \sigma_4^2\Sigma$, which follows from proofs of Theorem 1 and Theorem 2. In order to prove $\sigma_3^2 < \sigma_1^2$, we have to show that

$$\frac{\pi d\{\Gamma(\frac{d}{2})\}^2}{4\{\Gamma(\frac{d+1}{2})\}^2} > \frac{\frac{\pi^{d+1}}{d^{\frac{d+1}{2}}2^{\frac{d-1}{2}}\Gamma(\frac{d+1}{2})} + \frac{\pi^{d-1}\int_0^1 \frac{y^{\frac{d-1}{2}}}{(1+y)^{\frac{d+1}{2}}}dy}{d^{d+1}2^{\frac{d-3}{2}}} + \frac{\pi^{\frac{d}{2}}\Gamma(\frac{d+1}{2})\int_0^1 \frac{y^{\frac{d}{2}}}{(1+y)^{\frac{d+1}{2}}}dy}{2d^{\frac{d+1}{2}}\Gamma(\frac{d}{2}+1)}}{\left[\{\Gamma(\frac{1}{2})\Gamma(\frac{d}{2})\}^{-1}\Gamma(\frac{d+1}{2})\int_0^1 \frac{y^{\frac{d}{2}-1}}{(1+y)^{\frac{d+1}{2}}}dy\right]^2}$$

$$\Leftrightarrow \quad \frac{d}{4} > \frac{\frac{\pi^{d+1}}{d^{\frac{d+1}{2}}2^{\frac{d-1}{2}}\Gamma(\frac{d+1}{2})} + \frac{\pi^{d-1}\int_0^1 \frac{y^{\frac{d-1}{2}}}{(1+y)^{\frac{d+1}{2}}}dy}{d^{d+1}2^{\frac{d-3}{2}}} + \frac{\pi^{\frac{d}{2}}\Gamma(\frac{d+1}{2})\int_0^1 \frac{y^{\frac{d}{2}}}{(1+y)^{\frac{d+1}{2}}}dy}{2d^{\frac{d+1}{2}}\Gamma(\frac{d}{2}+1)}}{\left\{\int_0^1 \frac{y^{\frac{d}{2}-1}}{(1+y)^{\frac{d+1}{2}}}dy\right\}^2}$$

$$\Leftrightarrow \quad \frac{1}{4} > \frac{\frac{\pi^{d+1}}{d^{\frac{d-1}{2}}2^{\frac{d-1}{2}}\Gamma(\frac{d+1}{2})} + \frac{\pi^{d-1}\int_0^1 \frac{y^{\frac{d-1}{2}}}{(1+y)^{\frac{d+1}{2}}}dy}{d^{d}2^{\frac{d-3}{2}}} + \frac{\pi^{\frac{d}{2}}\Gamma(\frac{d+1}{2})\int_0^1 \frac{y^{\frac{d}{2}}}{(1+y)^{\frac{d+1}{2}}}dy}{2d^{\frac{d-1}{2}}\Gamma(\frac{d}{2}+1)}}{\left\{\int_0^1 \frac{y^{\frac{d}{2}-1}}{(1+y)^{\frac{d+1}{2}}}dy\right\}^2}.$$

Since both $\int_0^1 \frac{y^{\frac{d-1}{2}}}{(1+y)^{\frac{d+1}{2}}}dy$ and $\int_0^1 \frac{y^{\frac{d-1}{2}}}{(1+y)^{\frac{d+1}{2}}}dy$ are smaller than 1, it is enough to show that

$$\frac{1}{4} > \frac{\frac{\pi^{d+1}}{d^{\frac{d-1}{2}}2^{\frac{d-1}{2}}\Gamma(\frac{d+1}{2})} + \frac{\pi^{d-1}}{d^d 2^{\frac{d-3}{2}}} + \frac{\pi^{\frac{d}{2}}\Gamma(\frac{d+1}{2})}{2d^{\frac{d-1}{2}}\Gamma(\frac{d}{2}+1)}}{\left\{\int_0^1 \frac{y^{\frac{d}{2}-1}}{(1+y)^{\frac{d+1}{2}}}dy\right\}^2}$$

$$\Leftrightarrow \quad \frac{1}{4} > \frac{\frac{\pi^{d+1}}{d^{\frac{d-1}{2}}2^{\frac{d-1}{2}}\Gamma(\frac{d+1}{2})} + \frac{\pi^{d-1}}{d^d 2^{\frac{d-3}{2}}} + \frac{\pi^{\frac{d}{2}}\Gamma(\frac{d+1}{2})}{2d^{\frac{d-1}{2}}\Gamma(\frac{d}{2}+1)}}{\left\{\int_0^1 \frac{y^{\frac{d}{2}}}{y\sqrt{1+y}(1+y)^{\frac{d}{2}}}dy\right\}^2}$$

$$\Leftarrow \quad \frac{1}{4} > 2\frac{\frac{\pi^{d+1}}{d^{\frac{d-1}{2}}2^{\frac{d-1}{2}}\Gamma(\frac{d+1}{2})} + \frac{\pi^{d-1}}{d^d 2^{\frac{d-3}{2}}} + \frac{\pi^{\frac{d}{2}}\Gamma(\frac{d+1}{2})}{2d^{\frac{d-1}{2}}\Gamma(\frac{d}{2}+1)}}{\left\{\int_0^1 \frac{y^{\frac{d}{2}}}{(1+y)^{\frac{d}{2}}}dy\right\}^2} \quad \text{(since } y\sqrt{1+y} \leq 2.\text{)}$$

$$\Leftarrow \quad \frac{1}{4} > 2\frac{\frac{\pi^{d+1}}{d^{\frac{d-1}{2}}2^{\frac{d-1}{2}}\Gamma(\frac{d+1}{2})} + \frac{\pi^{d-1}}{d^d 2^{\frac{d-3}{2}}} + \frac{\pi^{\frac{d}{2}}\Gamma(\frac{d+1}{2})}{2d^{\frac{d-1}{2}}\Gamma(\frac{d}{2}+1)}}{\left(\int_0^1 \frac{y^{\frac{d}{2}}}{2^{\frac{d}{2}}}dy\right)^2} \quad \text{(since}(1+y) \leq 2)$$

$$\Leftarrow \quad \frac{1}{8} > \frac{\frac{\pi^{d+1}2^d}{d^{\frac{d-1}{2}}2^{\frac{d-1}{2}}\Gamma(\frac{d+1}{2})} + \frac{2^d\pi^{d-1}}{d^d 2^{\frac{d-3}{2}}} + \frac{2^d\pi^{\frac{d}{2}}\Gamma(\frac{d+1}{2})}{2d^{\frac{d-1}{2}}\Gamma(\frac{d}{2}+1)}}{\left(\int_0^1 y^{\frac{d}{2}}dy\right)^2}$$

$$\Leftrightarrow \quad \frac{1}{2} > \frac{(d+2)^2\pi^{d+1}2^d}{d^{\frac{d-1}{2}}2^{\frac{d-1}{2}}\Gamma(\frac{d+1}{2})} + \frac{(d+2)^2 2^d\pi^{d-1}}{d^d 2^{\frac{d-3}{2}}} + \frac{(d+2)^2 2^d\pi^{\frac{d}{2}}\Gamma(\frac{d+1}{2})}{2d^{\frac{d-1}{2}}\Gamma(\frac{d}{2}+1)}$$

$$\Leftrightarrow \quad \frac{1}{2} > g_1(d) + g_2(d) + g_3(d) \text{ (say).} \tag{3}$$

By the principle of mathematical induction, it can be shown that $g_1(d)$, $g_2(d)$ and $g_3(d)$ in inequality (3) are decreasing functions of d for all $d \geq 4$. This implies that $g(d) = g_1(d) + g_2(d) + g_3(d)$ is a decreasing function of d for all $d \geq 4$. Also, we have that $g(19)$ and $g(20)$ are smaller than $1/2$. This implies that $g(d) < 1/2$ for all $d \geq 19$. Also, direct verification shows that $\sigma_3^2 < \sigma_1^2$ when d lies between 12 and 18, and $\sigma_3^2 > \sigma_1^2$ when $2 \leq d \leq 11$.

Next, we show that $\sigma_1^2 < \sigma_4^2$. For multivariate Cauchy distribution (i.e., when $k = (d+1)/2$ in the multivariate polynomial tail family, see Fang et al. [21]), we have

$$\sigma_4^2 = \frac{d^{\frac{d+3}{2}}\Gamma(d/2)\int\limits_0^{m^2} \frac{y^{d/2}}{(1+y)^{\frac{d+1}{2}}}dy}{(d+1)^2\pi^{\frac{d}{2}}\Gamma(\frac{d+1}{2})\left\{\int\limits_0^{m^2} \frac{y^{d/2}}{(1+y)^{\frac{d+3}{2}}}dy\right\}^2},$$

where m^2 is such that

$$\frac{\int\limits_0^{m^2} \frac{y^{\frac{d}{2}-1}}{(1+y)^{\frac{d+1}{2}}}dy}{\beta(\frac{d}{2},\frac{1}{2})} = \frac{1}{2}, \tag{4}$$

which follows from Butler et al. [6, p.1387].

In order to prove $\sigma_1^2 < \sigma_4^2$, we have to show that

$$\frac{d^{\frac{d+3}{2}}\Gamma(d/2)\int\limits_0^{m^2} \frac{y^{d/2}}{(1+y)^{\frac{d+1}{2}}}dy}{(d+1)^2\pi^{\frac{d}{2}}\Gamma(\frac{d+1}{2})\left\{\int\limits_0^{m^2} \frac{y^{d/2}}{(1+y)^{\frac{d+3}{2}}}dy\right\}^2} > \frac{\pi d\{\Gamma(d/2)\}^2}{4\{\Gamma((d+1)/2)\}^2}$$

$$\Leftarrow \quad \frac{d^{\frac{d+3}{2}}\Gamma(d/2)\int\limits_0^{m^2} \frac{y^{d/2}}{(1+y)^{\frac{d+1}{2}}}dy}{(d+1)^2\pi^{\frac{d}{2}}\Gamma(\frac{d+1}{2})\left\{\int\limits_0^{m^2} \frac{y^{d/2}}{(1+y)^{\frac{d+3}{2}}}dy\right\}^2} > \frac{\pi^2}{4} \text{ (since } \tfrac{d\{\Gamma(d/2)\}^2}{\{\Gamma((d+1)/2)\}^2} \leq \pi \text{ for all } d \geq 2)$$

$$\Leftarrow \quad \frac{d^{\frac{d+3}{2}}\Gamma(d/2)\int\limits_0^{m^2} \frac{y^{d/2}}{(1+y)^{\frac{d+1}{2}}}dy}{(d+1)^2\pi^{\frac{d}{2}}\Gamma(\frac{d+1}{2})\left\{\int\limits_0^{m^2} \frac{y^{d/2}}{(1+y)^{\frac{d+1}{2}}}dy\right\}^2} > \frac{\pi^2}{4} \text{ (since } (1+y) \geq 1)$$

$$\Leftrightarrow \quad \frac{d^{\frac{d+3}{2}}\Gamma(d/2)}{(d+1)^2\pi^{\frac{d}{2}}\Gamma(\frac{d+1}{2})\int\limits_0^{m^2} \frac{y^{d/2}}{(1+y)^{\frac{d+1}{2}}}dy} > \frac{\pi^2}{4}$$

$$\Leftrightarrow \quad \frac{d^{\frac{d+3}{2}}\Gamma(d/2)}{(d+1)^2\pi^{\frac{d}{2}}\Gamma(\frac{d+1}{2})\int\limits_0^{m^2}\frac{y^{d/2-1}y}{(1+y)^{\frac{d+1}{2}}}dy} > \frac{\pi^2}{4}$$

$$\Leftarrow \quad \frac{d^{\frac{d+3}{2}}\Gamma(d/2)}{(m^2)(d+1)^2\pi^{\frac{d}{2}}\Gamma(\frac{d+1}{2})\int\limits_0^{m^2}\frac{y^{d/2-1}}{(1+y)^{\frac{d+1}{2}}}dy} > \frac{\pi^2}{4} \text{ (since } y \leq m^2\text{)}$$

$$\Leftrightarrow \quad \frac{d^{\frac{d+3}{2}}\Gamma(d/2)}{(m^2)(d+1)^2\pi^{\frac{d}{2}}\Gamma(\frac{d+1}{2})(1/2)\beta(1/2,d/2)} > \frac{\pi^2}{4} \text{ (Using (4))}$$

$$\Leftrightarrow \quad \frac{d^{\frac{d+3}{2}}}{(d+1)^2\pi^{\frac{d+1}{2}}m^2} > \frac{\pi^2}{8}$$

$$\Leftarrow \quad \frac{d^{\frac{d+3}{2}}}{(d+1)^2\pi^{\frac{d+1}{2}}(2.5d)^2} > \frac{\pi^2}{8} \text{ (since } m^2 \leq 2.5d\text{).} \tag{5}$$

To show inequality (5), we define

$$f(d) = \frac{d^{\frac{d+3}{2}}}{(d+1)^2\pi^{\frac{d+1}{2}}(2.5d)^2}.$$

Now, we have

$$\begin{aligned} f(d+1) &= \frac{(d+1)^{\frac{d+4}{2}}}{(d+2)^2\pi^{\frac{d+2}{2}}\{2.5(d+1)\}^2} \\ &= c_d f(d), \end{aligned}$$

where $c_d = \frac{\sqrt{d+1}(1+\frac{1}{d})^{\frac{d}{2}}\sqrt{1+\frac{1}{d}}}{\sqrt{\pi}(1+\frac{1}{d+1})^2}$. It can be shown that $c_d > 1$ for all $d \geq 4$. Also, $f(10) > \frac{\pi^2}{8}$. Also, we have $\sigma_1^2 < \sigma_4^2$ when $2 \leq d \leq 9$ by direct verification. This implies that $\sigma_1^2 < \sigma_4^2$ if $d \geq 2$. □

Proof of Theorem 5: It follows from the proofs of Theorems 1 and 2 that $\Sigma_1 = \sigma_1^2\Sigma$, $\Sigma_3 = \sigma_3^2\Sigma$ and $\Sigma_4 = \sigma_4^2\Sigma$.

From Fang et al. [21, p.83], we have

$$f_1(0) = \frac{\Gamma(k-(d-1)/2)}{\sqrt{\pi}\Gamma(k-d/2)}.$$

Using the expression of $f_1(0)$, we have

$$\sigma_1^2 = \frac{d\Gamma(d/2)^2}{4\{\Gamma(k-(d-1)/2)\}^2\{\Gamma((d+1)/2)\}^2}$$

for any density in the multivariate polynomial tail family.

As the first and the third terms of the numerator of σ_3^2 are positive (using Corollary 5 of Zuo [39]), in order to prove $\sigma_1^2 < \sigma_3^2$, it is enough to show that

$$\frac{\frac{m^d\{\Gamma(k)\}^2 \int_0^{m^2} \frac{y^{\frac{d-1}{2}}}{(1+y)^k} dy}{\sqrt{\pi}(1+m^2)^k \Gamma(\frac{d+1}{2})^2 \Gamma(k-\frac{d}{2})\Gamma(k-\frac{d-1}{2})}}{\left[\{\Gamma(\frac{d}{2})\Gamma(k-\frac{d}{2})\}^{-1}\Gamma(k)\int_0^{m^2} \frac{y^{\frac{d}{2}-1}}{(1+y)^k} dy\right]^2} > \frac{d\Gamma(d/2)^2}{4\{\Gamma(k-(d-1)/2)\}^2\{\Gamma((d+1)/2)\}^2},$$

where m is the 3/4-th quantile of the marginal density function.

Now, we have

$$\frac{\frac{m^d\{\Gamma(k)\}^2 \int_0^{m^2} \frac{y^{\frac{d-1}{2}}}{(1+y)^k} dy}{\sqrt{\pi}(1+m^2)^k \Gamma(\frac{d+1}{2})^2 \Gamma(k-\frac{d}{2})\Gamma(k-\frac{d-1}{2})}}{\left[\{\Gamma(\frac{d}{2})\Gamma(k-\frac{d}{2})\}^{-1}\Gamma(k)\int_0^{m^2} \frac{y^{\frac{d}{2}-1}}{(1+y)^k} dy\right]^2} > \frac{d\Gamma(d/2)^2}{4\{\Gamma(k-(d-1)/2)\}^2\{\Gamma((d+1)/2)\}^2}$$

$$\Leftrightarrow \quad \frac{m^d\Gamma(k-\frac{d}{2})\Gamma(k-\frac{d-1}{2})\int_0^{m^2} \frac{y^{\frac{d-1}{2}}}{(1+y)^k} dy}{\sqrt{\pi}(1+m^2)^k \left\{\int_0^{m^2} \frac{y^{\frac{d}{2}-1}}{(1+y)^k} dy\right\}^2} > d/4$$

$$\Leftarrow \quad \frac{m^d\Gamma(k-\frac{d}{2})\Gamma(k-\frac{d-1}{2})\int_0^{m^2} y^{\frac{d-1}{2}} dy}{\sqrt{\pi}(1+m^2)^{2k} \left\{\int_0^{m^2} \frac{y^{\frac{d}{2}-1}}{(1+y)^k} dy\right\}^2} > d/4 \text{ (since } y \le m^2)$$

$$\Leftarrow \quad \frac{m^d\Gamma(k-\frac{d}{2})\Gamma(k-\frac{d-1}{2})\int_0^{m^2} y^{\frac{d-1}{2}} dy}{\sqrt{\pi}(1+m^2)^{2k} \left(\int_0^{m^2} y^{\frac{d}{2}-1} dy\right)^2} > d/4 \text{ (since } y \ge 0)$$

$$\Leftrightarrow \quad \frac{2dm\Gamma(k-d/2)\Gamma\{k-(d-1)/2\}}{\sqrt{\pi}(d+1)(1+m^2)^k} > 1. \tag{6}$$

From the marginal densities of the multivariate polynomial tail family (see Fang et al. [21]), we can write

$$\int_0^m \frac{1}{(1+x^2/(2k-d))^{(2k-d+1)/2}} dx = (1/4)(2k-d)\sqrt{\pi})/\Gamma((2k-d+1)/2).$$

$$\Rightarrow \quad m \geq \{(1/4)(2k-d)\sqrt{\pi}\}/\Gamma(k-(d-1)/2).$$

The last implication follows from the fact that $x \geq 0$ in the integrand above. Using the inequality $m \geq ((1/4)(2k-d)\sqrt{\pi})/(\Gamma(k-(d-1)/2))$ in (6), we now have to show that

$$\frac{d(2k-d)\{\Gamma(k-d/2)\}^2}{(d+1)(1+m^2)^k} \geq 2 \Leftarrow \frac{(2k-d)\Gamma(k-d/2)}{2^k} \geq \frac{2(d+1)}{d}.$$

The last implication follows from the fact that m is a decreasing function of k (see Shaw [33]). This implies that $m \leq 1$ for all k (note that the 3/4-th quantile of standard Cauchy distribution is 1). Hence, $\sigma_1^2 < \sigma_3^2$ under the condition given in Theorem 5.

Next, we prove the other part of the theorem namely, $\sigma_1^2 < \sigma_4^2$. For the multivariate polynomial tail family, using Theorem 4 of Butler et al. [6], we have

$$\sigma_4^2 = \frac{d\Gamma(d/2)\Gamma(k-d/2) \int_0^{m^2} \frac{y^{d/2}}{(1+y)^k} dy}{4k^2\Gamma(k) \left\{ \int_0^{m^2} \frac{y^{d/2}}{(1+y)^{k+1}} dy \right\}^2},$$

where m is such that

$$\frac{2\Gamma(k)}{\Gamma(d/2)\Gamma(k-d/2)} \int_0^m \frac{r^{d-1}}{(1+r^2)^k} dr = 1/2$$

$$\Leftrightarrow \quad \frac{\Gamma(k)}{\Gamma(d/2)\Gamma(k-d/2)} \int_0^{m^2} \frac{y^{d/2-1}}{(1+y)^k} dy = 1/2, \tag{7}$$

which follows from Butler et al. [6, p.1387]. This implies that $1/(1+m^2)$ is the median of the beta distribution with parameters $d/2$ and $k-d/2$.

Beta distribution with parameters $d/2$ and $k-d/2$ is positively skewed if $k \geq d$. So, $1/(1+m^2) \geq$ mode of the beta distribution with parameters $d/2$ and $k-d/2$, i.e, we have

$$\begin{aligned} & 1/(1+m^2) \geq (k-d/2-1)/(k-1) \\ \Leftrightarrow \quad & m^2 \leq d/(2k-d-2). \end{aligned}$$

In order to prove $\sigma_1^2 < \sigma_4^2$, we now have to show that

$$\frac{d\Gamma(d/2)\Gamma(k-d/2) \int_0^{m^2} \frac{y^{d/2}}{(1+y)^k} dy}{4k^2\Gamma(k) \left\{ \int_0^{m^2} \frac{y^{d/2}}{(1+y)^{k+1}} dy \right\}^2} > \frac{d\{\Gamma(d/2)\}^2}{4\{\Gamma(k-(d-1)/2)\}^2\{\Gamma((d+1)/2)\}^2}$$

$$\Leftarrow \quad \frac{d\Gamma(d/2)\Gamma(k-d/2) \int_0^{m^2} \frac{y^{d/2}}{(1+y)^{k+1}} dy}{4k^2\Gamma(k) \left\{ \int_0^{m^2} \frac{y^{d/2}}{(1+y)^{k+1}} dy \right\}^2} > \frac{d\{\Gamma(d/2)\}^2}{4\{\Gamma(k-(d-1)/2)\}^2\{\Gamma((d+1)/2)\}^2}$$

(since $1+y \geq 1$)

$$\Leftrightarrow \quad \int_0^{m^2} \frac{y^{d/2}}{(1+y)^{k+1}} dy < \frac{\{\Gamma(k-(d-1)/2)\}^2\{\Gamma((d+1)/2)\}^2}{k^2\Gamma(k)\Gamma(d/2)\Gamma(k-d/2)}$$

$$\Leftrightarrow \quad \int_0^{m^2} \frac{y^{d/2-1}}{(1+y)^k} \left(\frac{y}{(1+y)}\right) dy < \frac{\{\Gamma(k-(d-1)/2)\}^2\{\Gamma((d+1)/2)\}^2}{k^2\Gamma(k)\Gamma(d/2)\Gamma(k-d/2)}$$

$$\Leftarrow \quad \frac{1}{2} \frac{\Gamma(d/2)\Gamma(k-d/2)}{\Gamma(k)} \frac{m^2}{(1+m^2)} < \frac{\{\Gamma(\frac{k-(d-1)}{2})\}^2\{\Gamma(\frac{d+1}{2}\}^2}{k^2\Gamma(k)\Gamma(d/2)\Gamma(k-d/2)}$$

$$\text{(using (7)and } \tfrac{y}{(1+y)} \uparrow y)$$

$$\Leftarrow \quad \frac{1}{2}\frac{\Gamma(d/2)\Gamma(k-d/2)}{\Gamma(k)}\frac{d}{(2k-2)} < \frac{\{\Gamma(k-(d-1)/2)\}^2\{\Gamma((d+1)/2)\}^2}{k^2\Gamma(k)\Gamma(d/2)\Gamma(k-d/2)}$$

$$\text{(since } m^2 \leq \tfrac{d}{2k-d-2})$$

$$\Leftrightarrow \quad \frac{1}{2}\frac{\Gamma(d/2)}{\Gamma(k)}\frac{d}{(2k-2)} < \frac{\{\Gamma(k-(d-1)/2)\}^2\{\Gamma((d+1)/2)\}^2}{k^2\Gamma(k)\Gamma(d/2)(\Gamma(k-d/2))^2}.$$

Alzer [2] derived an inequality for the ratio of two gamma functions. From that result, we have

$$\frac{\{\Gamma(k-(d-1)/2)\}^2}{\{\Gamma(k-d/2)\}^2} \leq (2k-d-1). \tag{8}$$

Using (8), it is now enough to prove that

$$\frac{d\{\Gamma(d/2)\}^2}{\{\Gamma((d+1)/2)\}^2} < (2k-2)(2k-d-1)$$

$$\Leftarrow \quad (8/\pi)k^2 < (2k-2)(2k-d-1) \text{ (since, } \tfrac{d\{\Gamma(d/2)\}^2}{\{\Gamma((d+1)/2)\}^2} \leq \tfrac{8}{\pi})$$

$$\Leftrightarrow \quad k^2(1-2/\pi) - k((d+3)/2) + ((d+1)/2) > 0.$$

Solving the last quadratic inequality in k, we get

$$k \geq \frac{(d+3)/2 + \sqrt{(d+3)^2/4 - 4\{(d+1)/2\}\{(\pi-2)/\pi\}}}{2(1-2/\pi)} \approx (4d/3).$$

□

Acknowledgments: The first author is supported by a grant from Council of Scientific and Industrial Research, Government of India.

References

[1] Agullo, J., Croux, C. and Aelst, S. V. (2008) The Multivariate Least Trimmed Squares Estimator, *Journal of Multivariate Analysis* **99**, no. 3, 311–338.

[2] Alzer, H. (1993) Some Gamma Function Inequalities, *Mathematics of Computation* **6**0, no. 201, 337–346.

[3] Arcones, M. A., Chen, Z. and Gine, E. (1994) Estimators Related to U-Processes with Applications to Multivariate Medians: Asymptotic Normality, *The Annals of Statistics* **2**2, no. 3, 1460–1477.

[4] Babu, G. J. and Rao, C. R. (1988) Joint asymptotic distribution of marginal quantiles and quantiles functions in samples from multivariate population, *Journal of Multivariate Analysis* **2**7, 15–23.

[5] Brown, B. M. (1983) Statistical Uses of the Spatial Median, *Journal of the Royal Statistical Society. Series B (Methodological)* **4**5, no. 1, 25–30.

[6] Butler, R. W., Davies, P. L. and Jhun, M (1993) Asymptotics for the Minimum Covariance Determinant Estimator, *The Annals of Statistics* **2**1, no. 2, 1385–1400.

[7] Chakraborty, B. (1998) Multivariate and Regression Analysis Based on the Geometry of Data Cloud. Ph.D. Thesis, Indian Statistical Institute, Calcutta.

[8] Chakraborty, B. and Chaudhuri, P. (1996) On a Transformation and Re-Transformation Technique for Constructing an Affine Equivariant Multivariate Median, *Proceedings of the American Mathematical*

Society **1**24, no. 8, 2539–2547.

[9] Chakraborty, B. and Chaudhuri, P. (1998) On an adaptive transformation retransformation estimate of multivariate location, *Journal of the Royal Statistical Society, Series B* **6**0, 145–157.

[10] Chakraborty, B. and Chaudhuri, P. (1999) A note on the robustness of multivariate medians, *Statistics and Probability Letters* **4**5, 269–276.

[11] Chakraborty, B. and Chaudhuri, P. (2003) On The Use of Genetic Algorithm with Elitism in Robust and Nonparametric Multivariate Analysis, *Austrian Journal of Statistics* **3**2, 13–27.

[12] Chakraborty, B. and Chaudhuri, P. (2008) On an Optimization problem in Robust Statistics, *Journal of Computational and Graphical Statistics* To appear.

[13] Chakraborty, B., Chaudhuri, P. and Oja, H. (1998) Operating Transformation Retransformation on Spatial Median and Angle test, *Statistica Sinica* **8**, 767–784.

[14] Chaudhuri, P. (1992) Multivariate Location estimation Using Extension of R-Estimates Through U-Statistics Type Approach, *The Annals of Statistics*, **2**0, no. 2, 897–916.

[15] Chaudhuri, P. (1996) On a Geometric Notion of Quantiles for Multivariate Data, *Journal of the American Statistical Association*, **9**1, no. 434, 863–872.

[16] Chen, J. (1995) Robustness of the half-space median, *Journal of Statistical Planning Inference*, **4**6, 175–181.

[17] DasGupta, A. (2008) *Asymptotic Theory of Statistics and Probability* Springer Verlag, New York, March 2008 (Forthcoming).

[18] Davies, L. (1992) The asymptotics of Rousseeuw's minimum volume ellipsoid estimator, *The Annals of Statistics*, **2**0, 1828–1843.

[19] Donoho, D. L. (1982) Breakdown properties of multivariate location estimators. Qualifying paper, Department of Statistics, Harvard University.

[20] Donoho, D. L. and Gasko, M. (1992) Breakdown Properties of Location Estimates Based on Half-space Depth and Projected Outlyingness, *The Annals of Statistics* **2**0, 1803–1827.

[21] Fang, K. T., Kotz, S. and Ng, K. W. (1989) *Symmetric Multivariate and Related Distributions*. Chapman and Hall, London.

[22] Feller, W. (1957) *An Introduction to Probability Theory and Its Applications, Volume 1*, John Wiley & Sons, New York.

[23] Gomez, E., Gomez-Villegas, M. A. and Marin, J. M. (1998) A Multivariate Generalization of the power Exponential Family of Distributions, *Communication of Statistics - Theory and Methods* **2**7, no. 3, 589–600.

[24] Liu, R. Y. (1990) On a notion of data depth based on random simplices, *The Annals of Statistics* **1**8, 405–414.

[25] Masse, J. C. (2002) Asymptotics for the Tukey Median, *Journal of Multivariate Analysis* **8**1, 286–300.

[26] Masse, J. C. (2004) Asymptotics for the Tukey depth process, with an application to a multivariate trimmed mean, *Bernoulli* **1**0, no. 3, 397–419.

[27] Oja, H. (1983) Descriptive statistics for multivariate distributions, *Statistics and Probability Letters* **1**, 327–332.

[28] Oja, H. and Niinimaa, A. O. (1985) Asymptotical properties of the generalized median in the case of multivariate normality, *Journal of the Royal Statistical Society B* **4**7, no. 2, 372–377.

[29] Oja, H., Niinimaa, A. O., and Tableman, M. (1990) The finite sample breakdown point of the Oja bivariate median and of the corresponding half-samples version, *Statistics and Probability Letters* **1**0, 325–328.

[30] Rousseeuw, P. J. (1983) Multivariate Estimation with High Breakdown Point, paper presented at Fourth Pannonian Symposium on Mathematical Statistics and Probability, Bad Tatzmannsdorf, Austria, September 4–9, 1983.

[31] Rousseeuw, P. J. (1984) Least median of squares regression, *Journal of the American Statistical Association* **7**9, 871–880.

[32] Rousseeuw, P. J. and Leroy, A. M. (1987) *Robust Regression and Outlier Detection*, John Wiley & Sons, New York.

[33] Shaw, W. T. (2006) New Methods for Managing "Student" T Distribution, *http://www.mth.kcl.ac.uk/shaww.*

[34] Small, C. G. (1990) A survey of multidimensional medians, *International Statistical Review* **5**8, 263–277.

[35] Stahel, W. A. (1981) Robust estimation: Infinitesimal optimality and covariance matrix estimators. Ph.D. Thesis, ETH, Zurich (in German).

[36] Tukey, J.W. (1975) Mathematics and picturing data. In R.D. James (ed.), Proceedings of the 1974 International Congress of Mathematicians, Vol. 2, pp. 523-531. Vancouver: Canadian Mathematical Congress.

[37] Zuo, Y. (2002) Multivariate Trimmed Means Based on Data Depth, *Statistical Data Analysis Based on the L_1-Norm and Related Methods*, 313–322.

[38] Zuo, Y. (2003) Projection - Based Depth Functions and Associated Medians, *The Annals of Statistics* **3**1, no. 5, 1460–1490.

[39] Zuo, Y. (2006) Multi-Dimensional Trimming Based on Projection Depth, *The Annals of Statistics* **3**4, no. 5, 2211–2251.

[40] Zuo, Y., Cui, H. and He, X. (2004) On the Stahel-Donoho Estimator and Depth-Weighted Means of Multivariate Data, *The Annals of Statistics* **2**2, no. 1, 167–188.