The Annals of Statistics
2020, Vol. 48, No. 2, 1052–1071
https://doi.org/10.1214/19-AOS1837

ADMISSIBLE BAYES EQUIVARIANT ESTIMATION OF LOCATION VECTORS FOR SPHERICALLY SYMMETRIC DISTRIBUTIONS WITH UNKNOWN SCALE

BY YUZO MARUYAMA[1] AND WILLIAM E. STRAWDERMAN[2]

[1]*Mathematics and Informatics Center, University of Tokyo, maruyama@mi.u-tokyo.ac.jp*
[2]*Department of Statistics, Rutgers University, straw@stat.rutgers.edu*

This paper investigates estimation of the mean vector under invariant quadratic loss for a spherically symmetric location family with a residual vector with density of the form $f(x,u) = \eta^{(p+n)/2} f(\eta\{\|x-\theta\|^2 + \|u\|^2\})$, where η is unknown. We show that the natural estimator x is admissible for $p = 1, 2$. Also, for $p \geq 3$, we find classes of generalized Bayes estimators that are admissible within the class of equivariant estimators of the form $\{1 - \xi(x/\|u\|)\}x$. In the Gaussian case, a variant of the James–Stein estimator, $[1 - \{(p-2)/(n+2)\}/\{\|x\|^2/\|u\|^2 + (p-2)/(n+2) + 1\}]x$, which dominates the natural estimator x, is also admissible within this class. We also study the related regression model.

1. Introduction. This paper considers estimation of the p-dimensional mean vector, θ, of a spherically symmetric distribution in the presence of an unknown scale, $\sigma = \eta^{-1/2}$. The loss function is scale invariant quadratic loss. More specifically, we will study the question of admissibility within the class of equivariant procedures. Estimation of the mean vector has long been an important problem, but has become even more important since Stein's (1956) groundbreaking result that the usual unbiased estimator, which is also the generalized Bayes estimator with respect to the uniform prior, is inadmissible in 3 and higher dimensions.

The issue of admissibility of generalized Bayes estimators in the Gaussian case with known scale was largely settled by the monumental 1971 paper of Brown. Brown and Hwang (1982) studied the related issue of admissibility of the vector of expected values in a multiparameter exponential family. Maruyama and Takemura (2008) and Maruyama (2009) studied admissibility of generalized Bayes estimators of the mean vector of a spherically symmetric distributions with known scale. In the Gaussian case, Brown (1971) as well as Dasgupta and Strawderman (1997) gave a sufficient condition for generalized Bayes estimator to be inadmissible.

However, aside from various inadmissibility results (e.g., James–Stein-type estimators for $p \geq 3$), there has been little progress in the unknown scale case, even for Gaussian distributions. A notable exception is Strawderman (1973) which gives a class of proper Bayes, and hence admissible estimators dominating the usual unbiased estimator in the Gaussian setting for $p \geq 5$. Also, when $p = 1, 2$, the usual unbiased estimator is admissible among all estimators, as shown in Section 5. While not surprising, this result has not appeared in the literature as far as we know.

The most important subclass of improved estimators is arguably the class of scale equivariant estimators, particularly those that are generalized Bayes. The main contribution of this paper is to study admissibility of such estimators within the class of scale equivariant procedures. Our method of proof uses Blyth's (1951) method in a way closely related to that of

Received September 2018; revised March 2019.
MSC2010 subject classifications. Primary 62C15; secondary 62J07.
Key words and phrases. Admissibility, Stein's phenomenon, generalized Bayes, Bayes equivariance.

Brown and Hwang (1982). In Brown and Hwang (1982), the sequence h_i

$$h_i(\lambda) = \begin{cases} 1, & \lambda \le 1, \\ 1 - \log\lambda/\log i, & 1 \le \lambda \le i, \\ 0, & \lambda > i, \end{cases} \tag{1.1}$$

is the key for their proof. In this paper, we utilize a somewhat different sequence

$$h_i(\lambda) = 1 - \frac{\log\log(\lambda + e)}{\log\log(\lambda + e + i)}, \quad e = \exp(1), \tag{1.2}$$

that are smoother and adapt better to priors on the boundary of admissibility and inadmissibility within this class.

For example, in the known scale Gaussian case, Brown (1971) establishes admissibility of priors with tails behaving like $\|\theta\|^{2-p}\log\|\theta\|$, while the results of Brown and Hwang (1982) establish admissibility for priors with tail behavior $\|\theta\|^{2-p}$ but not $\|\theta\|^{2-p}\log\|\theta\|$. Roughly speaking, by our modification of Brown and Hwang's (1982) method with the new sequence $h_i(\lambda)$ given by (1.2), the generalized Bayes estimator with respect to

$$\eta^{-1} \times \eta^{p/2}\bar{\pi}(\eta\|\theta\|^2),$$

where $\bar{\pi}(\|\theta\|^2)$ satisfies Brown's (1971) sufficient condition for admissibility, is shown to be admissible within the class of scale equivariant procedures; see Remarks 4.1 and 4.2 for details.

The ultimate goal in this direction is to demonstrate admissibility among all estimators. However, considering general admissibility of equivariant estimators in the presence of nuisance parameter has been a longstanding unsolved problem as mentioned in James and Stein (1961) and Brewster and Zidek (1974). While our results do not resolve the general admissibility issue, they do advance substantially our understanding of admissibility within the class of scale equivariant estimators.

We consider the following model:

$$(X, U) \sim \eta^{(p+n)/2} f(\eta\{\|x - \theta\|^2 + \|u\|^2\}), \tag{1.3}$$

where $X \in \mathbb{R}^p$ and $U \in \mathbb{R}^n$ and where $\theta \in \mathbb{R}^p$ and $\eta \in \mathbb{R}_+$ are unknown. We mainly assume

$$p \ge 3 \quad \text{and} \quad n \ge 2, \tag{1.4}$$

and consider the problem of estimating θ under scaled quadratic loss

$$L(\delta, \theta, \eta) = \eta\|\delta - \theta\|^2. \tag{1.5}$$

In particular, we are interested in the admissibility among the class of equivariant estimators of the form

$$\delta_\xi(X, U) = \{1 - \xi(X/\|U\|)\}X, \quad \text{where } \xi : \mathbb{R}^p \to \mathbb{R}. \tag{1.6}$$

We assume that $f(\cdot) \ge 0$ is defined so that each coordinate has variance $1/\eta$. In particular, this implies that $f(\cdot)$ in (1.3), satisfies

$$\int_{\mathbb{R}^{p+n}} f(\|v\|^2)\,\mathrm{d}v = 1, \qquad \int_{\mathbb{R}^{p+n}} v_i^2 f(\|v\|^2)\,\mathrm{d}v = 1, \tag{1.7}$$

for $v = (v_1, \dots, v_{p+n})^{\mathrm{T}} \in \mathbb{R}^{p+n}$. Needless to say, this is a generalization of the Gaussian case where

$$f_G(t) = \frac{1}{(2\pi)^{(p+n)/2}}\exp(-t/2),$$

and hence $X \sim N_p(\theta, \eta^{-1}I)$ and $\|U\|^2 \sim \eta^{-1}\chi^2_n$ are mutually independent.

In Section 2, we show that if an estimator of the form

$$\delta_\psi(X,U) = \{1 - \psi(\|X\|^2/\|U\|^2)\}X, \quad \text{where } \psi : \mathbb{R}_+ \to \mathbb{R} \tag{1.8}$$

is admissible within the class of all such estimators, then it is also admissible within the larger class of estimators of the form $\delta_\xi(X,U)$ given by (1.6). Note that the risk of an equivariant estimator of the form (1.8) is a function of $\lambda = \eta\|\theta\|^2$.

Section 3 studies equivariant estimators of the form (1.8) which minimize the average risk with respect to a (proper) prior $\pi(\lambda)$ on the maximal invariant $\lambda = \eta\|\theta\|^2$. We give an expression for this average risk and for the equivariant estimator which effects the minimization. Additionally, we show that this proper Bayes equivariant estimator is equivalent to the generalized (but not proper) Bayes estimator corresponding to the prior on (θ,η),

$$\pi(\theta,\eta) = \eta^{-1}\{\eta\|\theta\|^2\}^{1-p/2}\pi(\eta\|\theta\|^2).$$

Further we demonstrate that such an estimator is admissible among the class of estimators of the form (1.8), and hence (1.6).

Section 4, using Blyth's (1951) method, extends the class of estimators which are admissible within the class of estimators of the form (1.8). One main result gives admissibility under $\pi(\lambda)$ including

$$\pi(\lambda) = \lambda^\alpha \quad \text{for } -1/2 < \alpha \le 0,$$

$$\text{equivalently} \quad \pi(\theta,\eta) = \eta^{-1}\eta^{p/2}\{\eta\|\theta\|^2\}^{\alpha+1-p/2},$$

for densities f including the normal distribution and many multivariate t distributions. For fixed η, this corresponds to a subclass of subharmonic priors including the fundamental harmonic suggested by Stein (1974) and sometimes referred to as Stein's prior. An interesting special case ($\alpha = 0$) gives admissibility (within the class of equivariant estimators) of the generalized Bayes equivariant estimator corresponding to $\pi(\lambda) \equiv 1$ or $\pi(\theta,\eta) = \eta^{-1}\eta^{p/2}\{\eta\|\theta\|^2\}^{1-p/2}$. Here, the form of the generalized Bayes estimator is independent of the underlying density $f(\cdot)$, as shown in Maruyama (2003). Further this estimator is minimax and dominates the James and Stein (1961) estimator

$$\left(1 - \frac{(p-2)/(n+2)}{\|X\|^2/\|U\|^2}\right)X$$

provided $f(\cdot)$ is nonincreasing. Another interesting result is on a variant of the James–Stein estimator of the simple form

$$\left(1 - \frac{(p-2)/(n+2)}{\|X\|^2/\|U\|^2 + (p-2)/(n+2) + 1}\right)X.$$

In the Gaussian case, this is the generalized Bayes equivariant estimator corresponding to

$$\pi(\lambda) = \lambda^{p/2-1}\int_0^\infty \frac{1}{(2\pi\xi)^{p/2}}\exp\left(-\frac{\lambda}{2\xi}\right)\left(\frac{\xi}{1+\xi}\right)^{n/2} d\xi.$$

It is admissible within the class of equivariant estimators, and is minimax. Section 4.2 numerically studies the risk functions for several of the estimators in the Gaussian and multivariate-t case.

In Section 5, we show that when $p = 1, 2$, the estimator X is admissible among all estimators. In the Gaussian case, Kubokawa (2001) in his unpublished lecture notes written in Japanese, showed the admissibility of X. Here, we give a generalization to the unknown scale case for underlying density f given by (1.3).

In Section 6, we demonstrate that our setting (1.3) may be regarded as a canonical form of a regression model with an intercept and a general spherically symmetric error distribution, where estimators of the form (1.8) corresponds to estimators of the vector of regression coefficients of the form $\{1-\psi_\star(R^2)\}\hat{\beta}$ where $\psi_\star:(0,1)\to\mathbb{R}$, $\hat{\beta}$ is the vector of least square estimators, and R^2 is the coefficient of determination. Hence, from the regression viewpoint, the class of equivariant estimators is quite natural.

Section 7 gives some concluding remarks. Most of the proofs are given in Appendix A through Appendix N in Supplementary Material (Maruyama and Strawderman (2020)).

2. Admissibility in a broader sense. We consider two groups of transformations. In the following, let $S=\|U\|^2$.

Group I.

$$X\to\gamma\Gamma X,\qquad \theta\to\gamma\Gamma\theta,\qquad S\to\gamma^2 S,\qquad \eta\to\eta/\gamma^2,$$

where $\Gamma\in\mathcal{O}(p)$, the group of $p\times p$ orthogonal matrices, and $\gamma\in\mathbb{R}_+$.

Group II.

$$X\to\gamma X,\qquad \theta\to\gamma\theta,\qquad S\to\gamma^2 S,\qquad \eta\to\eta/\gamma^2,$$

where $\gamma\in\mathbb{R}_+$.

Equivariant estimators for Group I should satisfy

$$\delta(\gamma\Gamma X,\gamma^2 S)=\gamma\Gamma\delta(X,S),$$

and reduce to estimators of the form

$$\delta_\psi=\{1-\psi(\|X\|^2/S)\}X, \tag{2.1}$$

where $\psi:\mathbb{R}_+\to\mathbb{R}$. The equivariant estimator for Group II should satisfy

$$\delta(\gamma X,\gamma^2 S)=\gamma\delta(X,S),$$

and reduce to estimator of the form

$$\delta_\xi=\{1-\xi(X/\sqrt{S})\}X, \tag{2.2}$$

where $\xi:\mathbb{R}^p\to\mathbb{R}$. It is useful to note the following.

LEMMA 2.1.

1. *The risk*, $R(\theta,\eta,\delta_\psi)=E[\eta\|\delta_\psi-\theta\|^2]$, *of an estimator* δ_ψ, *is a function of* $\eta\|\theta\|^2\in\mathbb{R}_+$.
2. *The risk*, $R(\theta,\eta,\delta_\xi)$, *of an estimator* δ_ξ, *is a function of* $\eta^{1/2}\theta\in\mathbb{R}^p$.

The standard proof is left to the reader.

Let two classes of estimators be

$$\mathcal{D}_\psi=\{\delta_\psi \text{ with } \psi:\mathbb{R}_+\to\mathbb{R} \text{ given by (2.1)}\},$$
$$\mathcal{D}_\xi=\{\delta_\xi \text{ with } \xi:\mathbb{R}^p\to\mathbb{R} \text{ given by (2.2)}\}.$$

Clearly, it follows that $\mathcal{D}_\psi\subset\mathcal{D}_\xi$. We shall show that if $\delta\in\mathcal{D}_\psi$ is admissible among the class $\mathcal{D}_\psi$, then it is admissible among the class $\mathcal{D}_\xi$. The proof is due to Section 3 of Stein (1956), based on the compactness of the orthogonal group $\mathcal{O}(p)$, and the continuity of the problem.

THEOREM 2.1. *If* $\delta\in\mathcal{D}_\psi$ *is admissible among the class* $\mathcal{D}_\psi$, *then it is admissible among the class* $\mathcal{D}_\xi$.

PROOF. See Appendix A. □

In this paper, we will investigate admissibility among the class $\mathcal{D}_\psi$. Admissibility among the class $\mathcal{D}_\xi$ then follows by Theorem 2.1.

3. Proper Bayes equivariant estimators. Recall that an equivariant estimator for Group I is given by

$$\delta_\psi = \{1 - \psi(\|X\|^2/S)\}X. \tag{3.1}$$

Since, as noted in Lemma 2.1, the risk function of the estimator $\delta_\psi \in \mathcal{D}_\psi$, $R(\theta, \eta, \delta_\psi)$, depends only on $\eta\|\theta\|^2 \in \mathbb{R}_+$, it may be expressed as

$$R(\theta, \eta, \delta_\psi) = \tilde{R}(\eta\|\theta\|^2, \delta_\psi). \tag{3.2}$$

Let $\lambda = \eta\|\theta\|^2 \in \mathbb{R}_+$. We assume the prior density on λ is $\pi(\lambda)$, and in this section, we assume the propriety of $\pi(\lambda)$, that is,

$$\int_0^\infty \pi(\lambda)\,\mathrm{d}\lambda = 1. \tag{3.3}$$

For an equivariant estimator δ_ψ, we define the Bayes equivariant risk as

$$B(\delta_\psi, \pi) = \int_0^\infty \tilde{R}(\lambda, \delta_\psi)\pi(\lambda)\,\mathrm{d}\lambda. \tag{3.4}$$

In this paper, the estimator δ_ψ which minimizes $B(\delta_\psi, \pi)$, is called the Bayes equivariant estimator and is denoted by δ_π. In the following, let

$$c_m = \pi^{m/2}/\Gamma(m/2) \quad \text{for } m \in \mathbb{N}_+ \tag{3.5}$$

and

$$\bar{\pi}(\lambda) = c_p^{-1}\lambda^{1-p/2}\pi(\lambda) \tag{3.6}$$

so that $\bar{\pi}(\|\mu\|^2)$ is a proper probability density on $\mathbb{R}^p$, that is,

$$\int_{\mathbb{R}^p} \bar{\pi}(\|\mu\|^2)\,\mathrm{d}\mu = 1. \tag{3.7}$$

THEOREM 3.1. *Assume $\int_0^\infty \pi(\lambda)\,\mathrm{d}\lambda = 1$ and that f satisfies* (1.7).

1. *The Bayes equivariant risk $B(\delta_\psi, \pi)$,* (3.4), *is given by*

$$\begin{aligned} B(\delta_\psi, \pi) = c_n \int_{\mathbb{R}^p} &\psi(\|z\|^2)\left\{\psi(\|z\|^2) - 2\left(1 - \frac{z^{\mathrm{T}}M_2(z,\pi)}{\|z\|^2 M_1(z,\pi)}\right)\right\} \\ &\times \|z\|^2 M_1(z,\pi)\,\mathrm{d}z + p, \end{aligned} \tag{3.8}$$

where c_n is given by (3.5) *and*

$$\begin{aligned} M_1(z,\pi) &= \iint \eta^{(2p+n)/2} f(\eta\{\|z-\theta\|^2+1\})\bar{\pi}(\eta\|\theta\|^2)\,\mathrm{d}\theta\,\mathrm{d}\eta, \\ M_2(z,\pi) &= \iint \theta\eta^{(2p+n)/2} f(\eta\{\|z-\theta\|^2+1\})\bar{\pi}(\eta\|\theta\|^2)\,\mathrm{d}\theta\,\mathrm{d}\eta. \end{aligned} \tag{3.9}$$

2. *Given $\pi(\lambda)$, the minimizer of $B(\delta_\psi, \pi)$ with respect to ψ is*

$$\psi_\pi(\|z\|^2) = \arg\min_\psi B(\delta_\psi, \pi) = 1 - \frac{z^{\mathrm{T}}M_2(z,\pi)}{\|z\|^2 M_1(z,\pi)}. \tag{3.10}$$

3. *The Bayes equivariant estimator*

$$\delta_\pi = \{1 - \psi_\pi(\|X\|^2/S)\}X \tag{3.11}$$

is equivalent to the generalized Bayes estimator of θ with respect to the joint prior density

$$\eta^{-1}\eta^{p/2}\bar{\pi}(\eta\|\theta\|^2),$$

where $\bar{\pi}(\lambda) = c_p^{-1}\lambda^{1-p/2}\pi(\lambda)$.

4. *The Bayes equivariant estimator δ_π is admissible among the class $\mathcal{D}_\psi$.*

PROOF. See Appendix B. □

REMARK 3.1. As shown in Appendix C, the generalized Bayes estimator of θ with respect to the joint prior density $\eta^\nu \eta^{p/2}\bar{\pi}(\eta\|\theta\|^2)$ for any ν is a member of the class $\mathcal{D}_\psi$. Part 3 of Theorem 3.1 applies only to the special case of $\nu = -1$. Additionally, the admissibility results of this section and of Section 4 apply only to this special case of $\nu = -1$ and imply neither admissibility or inadmissibility of generalized Bayes estimators if $\nu \neq -1$. Also note that while $\pi(\lambda)$ is assumed proper in this section, the prior on (θ, η), $\eta^{-1}\eta^{p/2}\bar{\pi}(\eta\|\theta\|^2)$, is never proper since

$$\begin{aligned}\int_0^\infty \int_{\mathbb{R}^p} \eta^{-1}\eta^{p/2}\bar{\pi}(\eta\|\theta\|^2)\,\mathrm{d}\theta\,\mathrm{d}\eta &= \int_0^\infty \int_{\mathbb{R}^p} \eta^{-1}\bar{\pi}(\|\mu\|^2)\,\mathrm{d}\mu\,\mathrm{d}\eta \\ &= 1 \times \int_0^\infty \frac{\mathrm{d}\eta}{\eta} = \infty.\end{aligned}$$

4. Admissible Bayes equivariant estimators through the Blyth method. Even if $\pi(\lambda)$ on $\mathbb{R}_+$ (and hence $\bar{\pi}(\|\mu\|^2)$ on $\mathbb{R}^p$) is improper, that is,

$$\int_{\mathbb{R}^p} \bar{\pi}(\|\mu\|^2)\,\mathrm{d}\mu = \int_0^\infty \pi(\lambda)\,\mathrm{d}\lambda = \infty, \tag{4.1}$$

the estimator δ_π given by (3.11) can be defined if $M_1(z,\pi)$ and $M_2(z,\pi)$ given by (3.9) are both finite, and the admissibility of δ_π within the class of equivariant estimators can be investigated through Blyth's (1951) method.

We consider the Bayes equivariant risk difference under $\pi_i(\lambda)$ which is proper, but not necessarily standardized; that is, $\int_0^\infty \pi_i(\lambda)\,\mathrm{d}\lambda < \infty$. Let δ_π and $\delta_{\pi i}$ be Bayes equivariant estimators with respect to $\pi(\lambda)$ and $\pi_i(\lambda)$, respectively. By Parts 1 and 2 of Theorem 3.1, the Bayes equivariant risk difference under $\pi_i(\lambda)$ is given as follows:

$$\begin{aligned}&\operatorname{diff} B(\delta_\pi, \delta_{\pi i}; \pi_i) \\ &\quad= \int_0^\infty \{R(\lambda,\delta_\pi) - R(\lambda,\delta_{\pi i})\}\pi_i(\lambda)\,\mathrm{d}\lambda \\ &\quad= c_n \int_{\mathbb{R}^p} (\{\psi_\pi^2(\|z\|^2) - 2\psi_\pi(\|z\|^2)\psi_{\pi i}(\|z\|^2)\} \\ &\qquad - \{\psi_{\pi i}^2(\|z\|^2) - 2\psi_{\pi i}(\|z\|^2)\psi_{\pi i}(\|z\|^2)\})\|z\|^2 M_1(z,\pi_i)\,\mathrm{d}z \\ &\quad= c_n \int_{\mathbb{R}^p} \overline{\operatorname{diff} B}(z; \delta_\pi, \delta_{\pi i}; \pi_i)\,\mathrm{d}z,\end{aligned} \tag{4.2}$$

where c_n is given by (3.5) and where

$$\overline{\operatorname{diff} B}(z; \delta_\pi, \delta_{\pi i}; \pi_i) = \{\psi_\pi(\|z\|^2) - \psi_{\pi i}(\|z\|^2)\}^2 \|z\|^2 M_1(z,\pi_i). \tag{4.3}$$

There are several versions of the Blyth method. For our purpose, the following version from Brown (1971) and Brown and Hwang (1982) is useful.

THEOREM 4.1. *Assume that the sequence $\pi_i(\lambda)$ for $i = 1, 2, \ldots$, satisfies*:

BL.1 $\pi_1(\lambda) \le \pi_2(\lambda) \le \cdots$ *for any* $\lambda \ge 0$ *and* $\lim_{i\to\infty} \pi_i(\lambda) = \pi(\lambda)$.
BL.2 $\int_0^\infty \pi_i(\lambda)\,\mathrm{d}\lambda < \infty$ *for any fixed* i.
BL.3 $\int_0^1 \pi_1(\lambda)\,\mathrm{d}\lambda > \gamma$ *for some positive* $\gamma > 0$.

BL.4 $\lim_{i\to\infty} \operatorname{diff} B(\delta_\pi, \delta_{\pi i}; \pi_i) = 0$.

Then δ_π *is admissible among the class* $\mathcal{D}_\psi$.

PROOF. See Appendix D. □

We consider the following assumptions on π in addition to (4.1).

ASSUMPTIONS ON π.

A.1 $\pi(\lambda)$ is differentiable.

A.2 (Behavior around the origin) For $\lambda \in [0, 1]$, there exist $\alpha > -1/2$ and $\nu(\lambda)$ such that

$$\pi(\lambda) = \lambda^\alpha \nu(\lambda),$$

where

$$0 < \nu(0) < \infty \quad \text{and} \quad \lim_{\lambda\to 0} \lambda\nu'(\lambda) = 0.$$

A.3 (Asymptotic behavior) Let $\kappa(\lambda) = \lambda\pi'(\lambda)/\pi(\lambda)$. Either A.3.1 or A.3.2 is assumed;

A.3.1 $-1 \le \lim_{\lambda\to\infty} \kappa(\lambda) < 0$;

A.3.2 $\lim_{\lambda\to\infty} \kappa(\lambda) = 0$. Further either A.3.2.1 or A.3.2.2 is assumed;

A.3.2.1 $\kappa(\lambda)$ is eventually monotone increasing and approaches 0 from below.

A.3.2.2 $\limsup_{\lambda\to\infty} \{\log\lambda\}|\kappa(\lambda)| < 1$.

Some preliminary results on π satisfying Assumptions A.1–A.3 are summarized in Appendix E.1. Assumption A.2 is a sufficient condition for propriety around the origin, $\int_0^1 \pi(\lambda)\,d\lambda < \infty$. The lower bound of α, $-1/2$ (not the necessary condition for propriety, -1), comes from the application of the Cauchy–Schwarz inequality in the proof. See also Remark 4.3. In Assumption A.3, $\lim_{\lambda\to\infty} \kappa(\lambda) < -1$ implies propriety at infinity, that is, $\int_1^\infty \pi(\lambda)\,d\lambda < \infty$, the case which has been considered in Section 3.

A typical prior $\pi(\lambda)$ satisfying Assumptions A.1–A.3, corresponding to a generalized Strawderman's (1971) prior, is given by

$$\pi(\lambda; \alpha, \beta, b) = c_p \lambda^{p/2-1} \int_b^\infty \frac{1}{(2\pi\xi)^{p/2}} \exp\left(-\frac{\lambda}{2\xi}\right)(\xi - b)^\alpha (1+\xi)^\beta\, d\xi, \tag{4.4}$$

which is clearly differentiable. Also, by a Tauberian theorem (see, e.g., Theorem 4 of Section 5 of Chapter 13 in Feller (1971)), we have

$$\lim_{\lambda\to\infty} \frac{\lambda\pi'(\lambda; \alpha, \beta, b)}{\pi(\lambda; \alpha, \beta, b)} = \alpha + \beta.$$

When $\alpha + \beta = 0$, Assumption A.3.2.2 is satisfied. For either $\{\alpha > -1, b > 0\}$ or $\{\alpha > -1/2, b = 0\}$, Assumption A.2 is satisfied. See Appendix F for the proof. In summary, Assumptions A.1–A.3 are satisfied when $\{-1 \le \alpha + \beta \le 0, \alpha > -1, b > 0\}$ or $\{-1 \le \alpha + \beta \le 0, \alpha > -1/2, b = 0\}$.

Note that the power prior

$$\pi(\lambda) = \lambda^\alpha \quad \text{for } -1/2 < \alpha \le 0,$$

which will be considered in Section 4.1, corresponds to the case $\beta = 0$ and $b = 0$ in (4.4).

For a generalized prior $\pi(\lambda)$ satisfying Assumptions A.1–A.3, consider the sequence given by

$$\pi_i(\lambda) = \pi(\lambda)h_i^2(\lambda),$$

where $h_i(\lambda)$, for $\lambda \geq 0$ and $i = 1, 2, \ldots$, is defined by

$$h_i(\lambda) = 1 - \frac{\log\log(\lambda + e)}{\log\log(\lambda + e + i)}, \tag{4.5}$$

and $e = \exp(1)$. Some preliminary results on π_i with π satisfying Assumptions A.1–A.3 are summarized in Appendix E.2. It is clear that π_i satisfies BL.1 of Theorem 4.1. In Lemma E.2 of Appendix E.2, we show that π_i also satisfies BL.2 and BL.3 of Theorem 4.1.

REMARK 4.1. The basic idea behind the sequence h_i given by (4.5) comes from the h_i of Brown and Hwang (1982),

$$h_i(\lambda) = \begin{cases} 1, & \lambda \leq 1, \\ 1 - \log\lambda/\log i, & 1 \leq \lambda \leq i, \\ 0, & \lambda > i. \end{cases} \tag{4.6}$$

A smoothed version of the above is

$$h_i(\lambda) = 1 - \frac{\log(\lambda + 1)}{\log(\lambda + 1 + i)}. \tag{4.7}$$

The sequence h_i given by (4.5) is more slowly changing in both λ and i, in order to handle priors with heavier tail than treated in Brown and Hwang (1982). As in Remark 4.2, the sequence (4.6) is optimized for the case $\pi(\lambda) = O(1)$ and does not work well for a prior with heavier tails such as $\pi(\lambda) \approx \{\log\lambda\}^{1-\epsilon}$ which satisfies the sufficient condition of Theorem 4.2 below as well as Brown's (1971) sufficient condition for admissibility. Using Brown and Hwang's (1982) idea, but with the smooth and heavier-tailed sequence h_i given by (4.5), we can approach the boundary of admissibility in Brown's (1971) paper.

For BL.4, note that diff $B(\delta_\pi, \delta_{\pi i}; \pi_i)$ given by (4.2) is a functional of f as well as π and π_i. Some additional assumptions on f (as well as (1.7)) are required as follows:

ASSUMPTIONS ON f.

F.1 $0 < f(t) < \infty$ for any $t \geq 0$.
F.2 f is differentiable.
F.3 Either F.3.1 or F.3.2 is assumed;

F.3.1 $\limsup_{t\to\infty} t\frac{f'(t)}{f(t)} < -\frac{p+n}{2} - 2$.
F.3.2 $\limsup_{t\to\infty} t\frac{f'(t)}{f(t)} < -\frac{p+n}{2} - 3$.

Some preliminary results on f satisfying Assumptions F.1–F.3 are summarized in Appendix E.3. We note that, in addition to the normal distribution,

$$f_G(t) = (2\pi)^{-(p+n)/2}\exp(-t/2),$$

an interesting heavier tailed class, also satisfying Assumptions F.1–F.3, is given by the multivariate Student t with

$$\begin{aligned} f(t; m, b) &= \int_0^\infty \frac{f_G(t/g)}{g^{(p+n)/2}} \frac{g^{-m/2-1}}{\Gamma(m/2)(2/b)^{m/2}} \exp\left(-\frac{b}{2g}\right) \mathrm{d}g \\ &= \frac{\Gamma((p+n+m)/2)}{(\pi b)^{(p+n)/2}\Gamma(m/2)} (1 + t/b)^{-(p+n+m)/2}. \end{aligned}$$

The usual multivariate-t random vector with m-degrees of freedom is with $b = m$. But the restriction (1.7) determines $b = m - 2$. The distribution with the density $f(t; m, m-2)$, which satisfy (1.7), may be called a scaled multivariate-t distribution. For Assumptions F.3.1 and F.3.2, $m > 4$ and $m > 6$ are needed, respectively.

The main result on admissibility of δ_π given by (3.11) among the class $\mathcal{D}_\psi$ through the Blyth method, Theorem 4.1, is as follows.

THEOREM 4.2.

Case I *Assume Assumptions* A.1, A.2 *and* A.3.1 *on* π *and Assumptions* F.1, F.2 *and* F.3.1 *on* f. *Then the estimator* δ_π *given by* (3.11) *is admissible among the class* $\mathcal{D}_\psi$.

Case II *Assume Assumptions* A.1, A.2 *and* A.3.2 *on* π *and Assumptions* F.1, F.2 *and* F.3.2 *on* f. *Then the estimator* δ_π *given by* (3.11) *is admissible among the class* $\mathcal{D}_\psi$.

SKETCH OF THE PROOF OF THEOREM 4.2. Assume Assumptions A.1, A.2 and A.3 on π. As in (4.5), set

$$\pi_i(\lambda) = \pi(\lambda)\left\{1 - \frac{\log\log(\lambda + e)}{\log\log(\lambda + e + i)}\right\}^2$$

with $\lambda \geq 0$ and $i = 1, 2, \dots$. Then the first three parts BL.1, BL.2 and BL.3 of Theorem 4.1 follow from Parts 1, 8 and 6 of Lemma E.2, respectively.

Considering BL.4, we first provide an alternative expression $\overline{\operatorname{diff} B}(z; \delta_\pi, \delta_{\pi i}; \pi_i)$ in (4.2) and (4.3). Recall

$$\begin{aligned}
&\operatorname{diff} B(\delta_\pi, \delta_{\pi i}; \pi_i) = c_n \int_{\mathbb{R}^p} \overline{\operatorname{diff} B}(z; \delta_\pi, \delta_{\pi i}; \pi_i)\, dz, \\
&\overline{\operatorname{diff} B}(z; \delta_\pi, \delta_{\pi i}; \pi_i) = \{\psi_\pi(\|z\|^2) - \psi_{\pi i}(\|z\|^2)\}^2 \|z\|^2 M_1(z, \pi_i),
\end{aligned} \tag{4.8}$$

with

$$\begin{aligned}
\psi_\pi(z) &= 1 - \frac{z^{\mathrm{T}} M_2(z, \pi)}{\|z\|^2 M_1(z, \pi)} = \frac{z^{\mathrm{T}} z M_1(z, \pi) - z^{\mathrm{T}} M_2(z, \pi)}{\|z\|^2 M_1(z, \pi)}, \\
M_1(z, \pi) &= \iint \eta^{(2p+n)/2} f(\eta\{\|z - \theta\|^2 + 1\}) \bar{\pi}(\eta\|\theta\|^2)\, d\theta\, d\eta, \\
M_2(z, \pi) &= \iint \theta \eta^{(2p+n)/2} f(\eta\{\|z - \theta\|^2 + 1\}) \bar{\pi}(\eta\|\theta\|^2)\, d\theta\, d\eta.
\end{aligned}$$

The numerator of $\psi_\pi(z)$ is rewritten as

$$\begin{aligned}
&z^{\mathrm{T}} z M_1(z, \pi) - z^{\mathrm{T}} M_2(z, \pi) \\
&\quad = z^{\mathrm{T}} \iint \eta(z - \theta) \eta^{(2p+n)/2 - 1} f(\eta\{\|z - \theta\|^2 + 1\}) \bar{\pi}(\eta\|\theta\|^2)\, d\theta\, d\eta \\
&\quad = z^{\mathrm{T}} \iint \eta^{(2p+n)/2 - 1} \nabla_\theta F(\eta\{\|z - \theta\|^2 + 1\}) \bar{\pi}(\eta\|\theta\|^2)\, d\theta\, d\eta \\
&\quad = -z^{\mathrm{T}} \iint \eta^{(2p+n)/2 - 1} F(\eta\{\|z - \theta\|^2 + 1\}) \nabla_\theta \bar{\pi}(\eta\|\theta\|^2)\, d\theta\, d\eta,
\end{aligned} \tag{4.9}$$

where $F(t) = (1/2)\int_t^\infty f(s)\, ds$ and the last equality follows from an integration by parts. To justify this integration by parts, note that, for fixed θ_i, the ith component of θ, we have

$$\lim_{\theta_i \to \pm\infty} F(\eta\{\|z - \theta\|^2 + 1\}) \bar{\pi}(\eta\|\theta\|^2) = 0$$

for any fixed η, z, $\theta_1, \dots, \theta_{i-1}, \theta_{i+1}, \dots, \theta_p$, since the asymptotic behavior of $\bar{\pi}$ and F are given by

$$\bar{\pi}(\lambda) = c_p^{-1}\lambda^{1-p/2}\pi(\lambda) = o(\lambda^{1-p/2}\log\lambda) \quad \text{and} \quad F(t) = o(t^{-(p+n)/2-1}),$$

as in Part 7 of Lemma E.1 and Part 1.A of Lemma E.3, respectively. Thus the last equality of (4.9) follows.

Therefore $\overline{\operatorname{diff} B}(z; \delta_\pi, \delta_{\pi i}; \pi_i)$ is reexpressed as

$$\begin{aligned} &\overline{\operatorname{diff} B}(z; \delta_\pi, \delta_{\pi i}; \pi_i) \\ &= \Bigg\| \frac{\iint \eta^{(2p+n)/2-1} F(\eta\{\|z-\theta\|^2+1\}) \nabla_\theta \bar{\pi}(\eta\|\theta\|^2)\, d\theta\, d\eta}{\iint \eta^{(2p+n)/2} f(\eta\{\|z-\theta\|^2+1\}) \bar{\pi}(\eta\|\theta\|^2)\, d\theta\, d\eta} \\ &\quad - \frac{\iint \eta^{(2p+n)/2-1} F(\eta\{\|z-\theta\|^2+1\}) \nabla_\theta \bar{\pi}_i(\eta\|\theta\|^2)\, d\theta\, d\eta}{\iint \eta^{(2p+n)/2} f(\eta\{\|z-\theta\|^2+1\}) \bar{\pi}_i(\eta\|\theta\|^2)\, d\theta\, d\eta} \Bigg\|^2 \\ &\quad \times \iint \eta^{(2p+n)/2} f(\eta\{\|z-\theta\|^2+1\}) \bar{\pi}_i(\eta\|\theta\|^2)\, d\theta\, d\eta. \end{aligned} \tag{4.10}$$

For two cases, Cases I and II, we will bound $\overline{\operatorname{diff} B}(z; \delta_\pi, \delta_{\pi i}; \pi_i)$ from above by some integrable functions independent of i. Then the theorem follows by the dominated convergence theorem because $\lim_{i\to\infty} \operatorname{diff} B(\delta_\pi, \delta_{\pi i}; \pi_i) = 0$ since $h_i^2 \to 1$ and $\delta_{\pi_i} \to \delta_\pi$ in the expression of (4.8).

Here is a reason why we have two cases. Recall $\operatorname{diff} B(\delta_\pi, \delta_{\pi i}; \pi_i)$ given by (4.2) is a functional of f as well as π and π_i. In Case II with Assumption A.3.2, $\pi(\lambda)$ satisfies $\lambda\pi'(\lambda)/\pi(\lambda) \to 0$ as $\lambda \to \infty$. In this case, we more carefully bound $\operatorname{diff} B(\delta_\pi, \delta_{\pi i}; \pi_i)$ from above, but need the more restrictive assumption on f as in Assumption F.3.2 as well as additional assumption on π, either Assumption A.3.2.1 or A.3.2.2.

More concretely, in Appendix G, we consider Case I, where $\overline{\operatorname{diff} B}(z; \delta_\pi, \delta_{\pi i}; \pi_i)$ is bounded as

$$\overline{\operatorname{diff} B}(z; \delta_\pi, \delta_{\pi i}; \pi_i) \le 2(\Delta_{1i} + \Delta_{2i}),$$

where

$$\Delta_{1i} = \frac{\| \iint \eta^{(2p+n)/2-1} F(\circ) \bar{\pi}(\bullet) \nabla_\theta h_i^2(\bullet)\, d\theta\, d\eta \|^2}{\iint \eta^{(2p+n)/2} f(\circ) \bar{\pi}_i(\bullet)\, d\theta\, d\eta}, \tag{4.11}$$

$$\begin{aligned} \Delta_{2i} &= \frac{\| \iint \eta^{(2p+n)/2-1} F(\circ) \nabla_\theta \bar{\pi}(\bullet)\, d\theta\, d\eta \|^2}{\iint \eta^{(2p+n)/2} f(\circ) \bar{\pi}(\bullet)\, d\theta\, d\eta} \\ &\quad + \frac{\| \iint \eta^{(2p+n)/2-1} F(\circ) \nabla_\theta \bar{\pi}(\bullet) h_i^2(\bullet)\, d\theta\, d\eta \|^2}{\iint \eta^{(2p+n)/2} f(\circ) \bar{\pi}_i(\bullet)\, d\theta\, d\eta}, \end{aligned} \tag{4.12}$$

where, for notational convenience and to control the size of expressions,

$$\bullet = \eta\|\theta\|^2, \quad \circ = \eta(\|z-\theta\|^2 + 1).$$

The integrability of $\sup_i \Delta_{1i}$ and $\sup_i \Delta_{2i}$ are shown in Appendices G.1 and G.2, respectively. More concretely, Appendix G.2 consists of two parts. We consider $\alpha > 0$ and $-1/2 < \alpha \le 0$ separately in Appendices G.2.1 and G.2.2, respectively.

In Appendix H, we consider Case II, where $\overline{\operatorname{diff} B}(z; \delta_\pi, \delta_{\pi i}; \pi_i)$ is bounded as

$$\overline{\operatorname{diff} B}(z; \delta_\pi, \delta_{\pi i}; \pi_i) \le 2\frac{(n+p)^2}{(n+2)^2}\{\Delta_{1i} + (p-2)^2\Delta_{3i} + 4\Delta_{4i}\},$$

where

$$\Delta_{3i} = \frac{1}{\|z\|^2} \frac{\{\iint (z^{\mathrm{T}}\theta/\|\theta\|^2 - 1)\eta^{(2p+n)/2-1} F(\circ)\bar{\pi}(\bullet)\,\mathrm{d}\theta\,\mathrm{d}\eta\}^2}{\iint \eta^{(2p+n)/2} f(\circ)\bar{\pi}(\bullet)\,\mathrm{d}\theta\,\mathrm{d}\eta} + \frac{1}{\|z\|^2} \frac{\{\iint (z^{\mathrm{T}}\theta/\|\theta\|^2 - 1)\eta^{(2p+n)/2-1} F(\circ)\bar{\pi}_i(\bullet)\,\mathrm{d}\theta\,\mathrm{d}\eta\}^2}{\iint \eta^{(2p+n)/2} f(\circ)\bar{\pi}_i(\bullet)\,\mathrm{d}\theta\,\mathrm{d}\eta},$$

$$\Delta_{4i} = \frac{\|\iint \theta\eta^{(2p+n)/2-1} F(\circ)\kappa(\bullet)\bar{\pi}(\bullet)\|\theta\|^{-2}\,\mathrm{d}\theta\,\mathrm{d}\eta\|^2}{\iint \eta^{(2p+n)/2} f(\circ)\bar{\pi}(\bullet)\,\mathrm{d}\theta\,\mathrm{d}\eta} + \frac{\|\iint \theta\eta^{(2p+n)/2-1} F(\circ)\kappa(\bullet)\bar{\pi}_i(\bullet)\|\theta\|^{-2}\,\mathrm{d}\theta\,\mathrm{d}\eta\|^2}{\iint \eta^{(2p+n)/2} f(\circ)\bar{\pi}_i(\bullet)\,\mathrm{d}\theta\,\mathrm{d}\eta}.$$

The integrability of $\sup_i \Delta_{3i}$ and $\sup_i \Delta_{4i}$ are shown in Appendices H.1 and H.2, respectively. □

Before giving some useful corollaries and examples, we give some remarks concerning the main result and method of proof, and indicate some of the differences between our proof and that of Brown and Hwang (1982).

REMARK 4.2. Assumption A.3 is a sufficient condition for

$$\int_1^\infty \frac{\mathrm{d}\lambda}{\lambda\pi(\lambda)} = \infty \quad \Leftrightarrow \quad \int_1^\infty \frac{\mathrm{d}\lambda}{\lambda^{p/2}\bar{\pi}(\lambda)} = \infty, \tag{4.13}$$

which is related to admissibility in the known variance case as follows. Maruyama (2009) showed that, in the problem of estimating μ of $X \sim N_p(\mu, I)$, regularly varying priors $\bar{\pi}(\|\mu\|^2)$ with

$$\int_1^\infty \frac{\mathrm{d}\lambda}{\lambda^{p/2}\bar{\pi}(\lambda)} = \infty \tag{4.14}$$

lead to admissibility, that is, the (generalized) Bayes estimator

$$X + \nabla \log m_{\bar{\pi}}(\|X\|^2),$$

where

$$m_{\bar{\pi}}(\|x\|^2) = \frac{1}{(2\pi)^{p/2}} \int \exp(-\|x-\mu\|^2/2)\bar{\pi}(\|\mu\|^2)\,\mathrm{d}\mu \tag{4.15}$$

is admissible. As Maruyama (2009) pointed out, the sufficient condition (4.14), which depends directly on the prior $\bar{\pi}(\|\mu\|^2)$, is closely related to Brown's (1971) sufficient condition for admissibility

$$\int_1^\infty \frac{\mathrm{d}r}{r^{p/2} m_{\bar{\pi}}(r)} = \infty,$$

which depends on the marginal distribution and only indirectly on the prior. Under Assumption A.3, as shown in Part 7 of Lemma E.1 in Appendix E, there exist $\epsilon \in (0, 1)$ and $\lambda_* > \exp(1)$ such that $\pi(\lambda)/(\log\lambda)^{1-\epsilon}$ for $\lambda \ge \lambda_*$ is bounded from above. This implies that Assumption A.3 is tight for the nonintegrability of (4.13) among the class

$$\bar{\pi}(\lambda) \approx \lambda^{1-p/2}(\log\lambda)^b \quad \text{with } b \in \mathbb{R},$$

or equivalently

$$\pi(\lambda) \approx (\log\lambda)^b \quad \text{with } b \in \mathbb{R}.$$

In particular, for $b = 1 - \epsilon$, $\int_1^\infty \mathrm{d}\lambda/\{\lambda(\log\lambda)^{1-\epsilon}\} = \infty$ and Assumption A.3.2 is satisfied, but for $b = 1 + \epsilon$, $\int_1^\infty \mathrm{d}\lambda/\{\lambda(\log\lambda)^{1+\epsilon}\} < \infty$ and Assumption A.3.2 is not satisfied. Actually, in the second case, the corresponding Bayes equivariant estimator is inadmissible as shown in Maruyama and Strawderman (2017). For the boundary case, $\epsilon = 0$ or $b = 1$, $\int_1^\infty \mathrm{d}\lambda/(\lambda\log\lambda) = \infty$ but Assumption A.3.2 is not satisfied. Our result does not settle the issue of admissibility within the class $\mathcal{D}_\psi$. Additionally, in Maruyama and Strawderman (2017), this case was also a boundary case and we were unable to settle the question of quasi-admissibility; see Remark 1 of Maruyama and Strawderman (2017).

Example 4.1 of Brown and Hwang (1982) gives two separate sufficient conditions for admissibility as follows:

BH.1 $\bar{\pi}(\lambda) \le \lambda^{1-p/2-\epsilon}$ for some $\epsilon > 0$ and

$$\lambda\bar{\pi}'(\lambda)/\bar{\pi}(\lambda) = O(1).$$

BH.2 $\bar{\pi}(\lambda) \le \lambda^{1-p/2}$ and

$$\lambda\bar{\pi}'(\lambda)/\bar{\pi}(\lambda) = O(1) \quad \text{and} \quad \lambda^2\bar{\pi}''(\lambda)/\bar{\pi}(\lambda) = O(1). \tag{4.16}$$

Theoretically, the sequence (4.6) of Brown and Hwang (1982) is optimized for the case $\bar{\pi}(\lambda) \approx \lambda^{1-p/2}$ and does not work well for a prior with heavier tails such as $\bar{\pi}(\lambda) \approx \lambda^{1-p/2}\{\log\lambda\}^{1-\epsilon}$ which satisfies Brown's (1971) sufficient condition for admissibility. This paper as well as Maruyama (2009) demonstrates that using Brown and Hwang's (1982) idea, but with the smooth and heavier-tailed sequence h_i given by (4.5), we can approach the boundary of admissibility in Brown's (1971) paper.

REMARK 4.3. In Remark 4.2, we paid attention to the tail behavior of the priors. Here, we consider the behavior of the priors around the origin. Brown and Hwang (1982) omitted a condition on the behavior around the origin in both BH.1 and BH.2. Suppose $\bar{\pi}(\lambda)$ satisfies the boundedness of $\lambda\bar{\pi}'(\lambda)/\bar{\pi}(\lambda)$, which is a necessary condition for A.1–A.3. Then, from their (3.3) on page 208,

$$\int_0^1 \lambda^{p/2-1}\frac{\bar{\pi}(\lambda)}{\lambda}\,\mathrm{d}\lambda < \infty$$

or equivalently

$$\int_0^1 \frac{\pi(\lambda)}{\lambda}\,\mathrm{d}\lambda < \infty \tag{4.17}$$

is needed for establishing that their $B_n \to 0$. This is clearly more restrictive than Assumption A.2. For example, $\pi(\lambda) = \lambda^\alpha$ for $\alpha \le 0$ does not satisfy the integrability (4.17) but does satisfy Assumption A.2 if $\alpha > -1/2$.

As in Brown and Hwang (1982), the key to the admissibility proof in this paper is to adequately apply the Cauchy–Schwarz inequality. When we treat the term corresponding to Brown and Hwang's (1982) B_n and apply the Cauchy–Schwarz inequality, we introduce

$$k(\lambda;\beta) = \lambda^\beta I_{[0,1]}(\lambda) + I_{(1,\infty)}(\lambda), \quad \text{for } \beta > 0,$$

which is effective for relaxing the condition around the origin. The reader can compare the inequality (G.6) without $k(\lambda;\beta)$ in Appendix G.2.1, with the inequality (G.13) with $k(\lambda;\beta)$ in Appendix G.2.2. When we apply the Cauchy–Schwarz inequality in (G.13), the first term involving $k(\lambda;\beta)$ gets smaller for larger β. On the other hand, the second term involving $1/k(\lambda;\beta)$ gets larger for larger β. Since the term (G.13) (the term corresponding to Brown and Hwang's (1982) B_n) is bounded when

$$\alpha > \max(-\beta, \beta - 1), \tag{4.18}$$

also since

$$\min_{\beta>0}\{\max(-\beta,\beta-1)\} = -\frac{1}{2}, \qquad \arg\min_{\beta>0}\{\max(-\beta,\beta-1)\} = \frac{1}{2},$$

the best choice is $\beta = 1/2$ and the corresponding lower bound of α is $-1/2$. This is why $\alpha > -1/2$ is assumed in Assumption A.2.

REMARK 4.4. Under Assumption A.3.2, $\pi(\lambda)$ satisfies $\lambda\pi'(\lambda)/\pi(\lambda) \to 0$ as $\lambda \to \infty$. In this case, to establish

$$\lim_{i\to\infty} \operatorname{diff} B(\delta_\pi, \delta_{\pi i}; \pi_i) = 0,$$

we carefully bound diff $B(\delta_\pi, \delta_{\pi i}; \pi_i)$ from above (see Appendix H), but need the more restrictive assumption on f as in Assumption F.3.2 as well as additional assumption on π, either Assumption A.3.2.1 or A.3.2.2. Thus, we have two cases in Theorem 4.2.

Notice that Brown and Hwang (1982) also have 2 sets of conditions BH.1 and BH.2 as noticed in Remark 4.2 even in the case of a normal distribution. Of course, in our case, the normal distribution satisfies both Assumptions F.3.1 and F.3.2, but for the scaled multivariate-t, Assumptions F.3.1 and F.3.2 are satisfied for $m > 4$ and $m > 6$, respectively.

REMARK 4.5. Note also (4.16) in BH.2 requires that $\bar{\pi}(\cdot)$ be twice differentiable while in our case II result, only differentiability is required; see Appendix F.3.1 for details.

4.1. *Some interesting cases.* Here, we present three interesting special cases of our main general theorem.

COROLLARY 4.1. *Assume Assumptions* F.1, F.2 *and* F.3.2 *on* f.

1. *Then* δ_π *with* $\pi \equiv 1$, *or equivalently the generalized Bayes estimator under the prior on* (θ, η) *given by*

$$\eta^{-1}\eta^{p/2}\{\eta\|\theta\|^2\}^{(2-p)/2},$$

is admissible among the class $\mathcal{D}_\psi$.

2. *The form of the generalized Bayes estimator does not depend on* f *and is given by* $\{1-\psi_0(W)\}X$ *where* $W = \|X\|^2/S$ *and*

$$\psi_0(w) = \frac{\int_0^\infty (1+\xi)^{n/2}(1+w+\xi)^{-(p+n)/2-1}\,\mathrm{d}\xi}{\int_0^\infty (1+\xi)^{n/2+1}(1+w+\xi)^{-(p+n)/2-1}\,\mathrm{d}\xi}.$$

3. *This estimator is minimax simultaneously for all such* f.

4. *This estimator dominates the James–Stein estimator*

$$\left(1 - \frac{p-2}{n+2}\frac{S}{\|X\|^2}\right)X$$

if f *is nonincreasing* (*i.e.*, *the distribution of* (X, U) *is unimodal*).

PROOF. For Part 1, Assumptions A.1, A.2 and A.3.2 are satisfied by $\pi(\lambda) \equiv 1$. Parts 2 and 4 are both shown by Maruyama (2003). Part 3 is shown by Cellier, Fourdrinier and Robert (1989). □

Note that the hierarchical structure in Corollary 4.1 is that $\mu = \eta^{1/2}\theta$ conditional on η has the fundamental harmonic prior $\|\mu\|^{2-p}$, sometimes referred as the Stein prior, while η^{-1} has the scale invariant prior. As shown in Maruyama and Strawderman (2017), this estimator is close to the boundary of quasi-admissibility and quasi-inadmissibility in that an estimator of the form $(1-aS/\|X\|^2)X$ is quasi-inadmissible (and inadmissible) if $a < (p-2)/(n+2)$ and quasi-admissible if $a \geq (p-2)/(n+2)$. As mentioned above, quasi-admissibility (inadmissibility) refers to the non-solvability (solvability) of the differential inequality $\Delta\text{SURE} \leq 0$ where ΔSURE is the Stein unbiased risk difference estimate.

COROLLARY 4.2. *Assume Assumptions F.1, F.2 and F.3.1 on f. Let $\alpha \in (-1/2, 0)$.*

1. *Then δ_π with $\pi(\lambda) = \lambda^\alpha$, or equivalently the generalized Bayes estimator under the prior on (θ, η) given by*

$$\eta^{-1}\eta^{p/2}\{\eta\|\theta\|^2\}^{\alpha+(2-p)/2},$$

is admissible among the class $\mathcal{D}_\psi$.

2. *The form of the estimator does not depend on f and is given by $\{1 - \psi_\alpha(W)\}X$ where $W = \|X\|^2/S$ and*

$$\psi_\alpha(w) = \frac{\int_0^\infty \xi^\alpha(1+\xi)^{n/2}(1+w+\xi)^{-(p+n)/2-1}\,d\xi}{\int_0^\infty \xi^\alpha(1+\xi)^{n/2+1}(1+w+\xi)^{-(p+n)/2-1}\,d\xi}. \tag{4.19}$$

3. *This estimator is minimax when*

$$-\left(5 + \frac{2}{p-2} + \frac{3p}{n+2}\right)^{-1} \leq \alpha < 0.$$

PROOF. For Part 1, Assumptions A.1, A.2 and A.3.1 are satisfied by $\pi(\lambda) = \lambda^\alpha$ for $\alpha \in (-1/2, 0)$. Part 2 is shown by Maruyama (2003). For Part 3, see Maruyama and Strawderman (2009) and Appendix J. □

The following corollary relates to the so-called "simple Bayes estimators" from Maruyama and Strawderman (2005).

COROLLARY 4.3. *Assume f is Gaussian. Then the simple Bayes estimator*

$$\left(1 - \frac{a}{(a+1)(b+1) + \|X\|^2/S}\right)X$$

with $a \geq (p-2)/(n+2)$ and $b \geq 0$ is admissible among the class $\mathcal{D}_\psi$. Within the region $\{(a,b) : a \geq (p-2)/(n+2)$ and $b \geq 0\}$, the subregion $\{(a,b) : a > (p-1)/(n+1)$ and $b \geq 0\}$ corresponds to proper priors, for which $\int_0^\infty \pi(\lambda)\,d\lambda < \infty$. Furthermore, the estimators with $(p-2)/(n+2) \leq a \leq 2(p-2)/(n+2)$ are minimax.

PROOF. The estimator is a (generalized) Bayes equivariant estimator with respect to $\pi(\lambda; \alpha, \beta, b)$ given by (4.4) with $\beta = -n/2$ and $\alpha = (p+n)/\{2(a+1)\} - 1$. The condition $\alpha + \beta < -1$ equivalent to $a > (p-1)/(n+1)$ corresponds to to a proper prior, for which $\int_0^\infty \pi(\lambda)\,d\lambda < \infty$; see Maruyama and Strawderman (2005) and Appendix K. □

4.2. *Numerical and asymptotic study of the risk.* The risk functions of several of the estimators in Corollaries 4.1 and 4.3 are presented in Figure 1. The first graph presents the risks in the Gaussian case, for $p = 10$ and $n = 10$, for four estimators,

$$\delta_{\text{JS}} = \left(1 - \frac{p-2}{n+2}\frac{S}{\|X\|^2}\right)X \quad \text{James–Stein estimator,}$$

$$\delta_{\text{JS}}^+ = \max\left(0, 1 - \frac{p-2}{n+2}\frac{S}{\|X\|^2}\right)X \quad \text{James–Stein positive part estimator,}$$

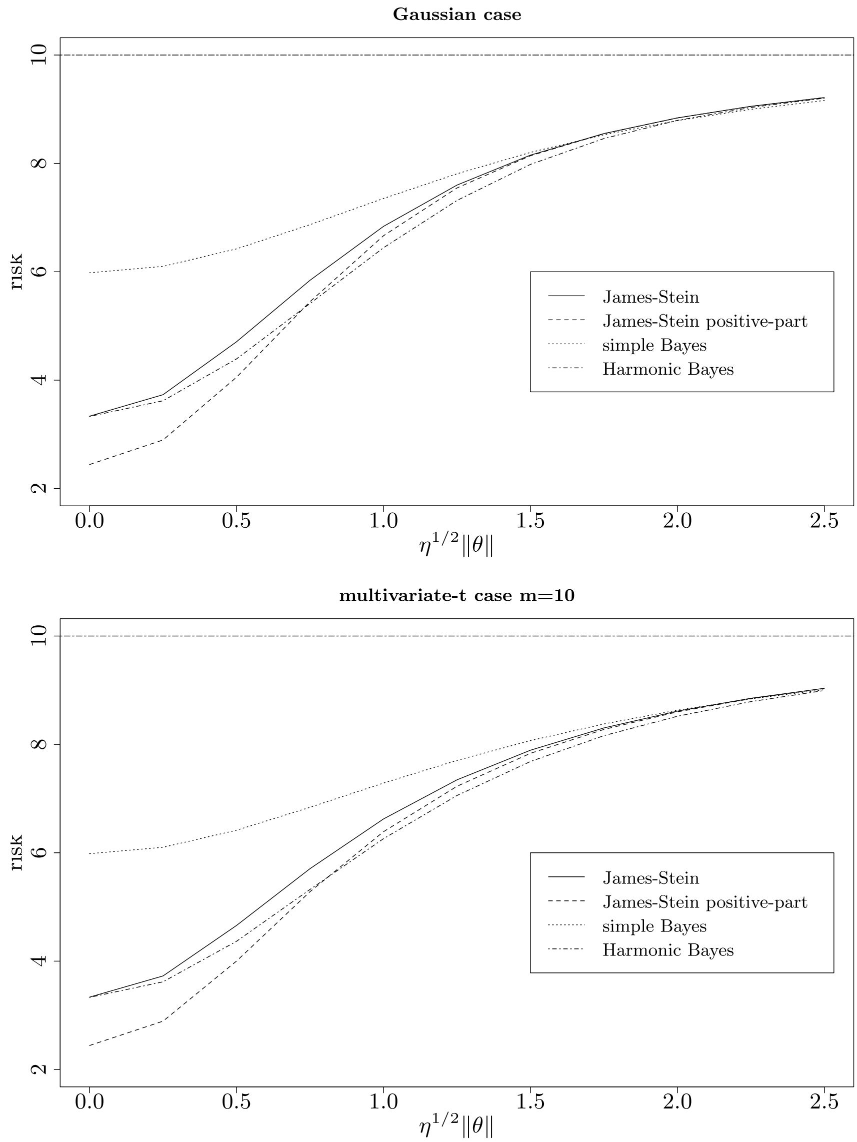

FIG. 1. *Risk* (*Gaussian and multivariate-t* $(m = 10)$) *with* $p = 10$ *and* $n = 10$.

$$\delta_{\mathrm{SB}} = \left(1 - \frac{(p-2)/(n+2)}{\|X\|^2/S + (p-2)/(n+2) + 1}\right)X \quad \text{by Corollary 4.3,}$$

$$\delta_0 = \{1 - \psi_0(\|X\|^2/S)\}X \quad \text{“Harmonic Bayes” by Corollary 4.1.}$$

As in mentioned in Part 4 of Corollary 4.1, the risk of δ_0 is uniformly smaller than that of James–Stein estimator. Also when $\eta\|\theta\|^2 = 0$, the two risks are equal as shown in Kubokawa (1994). Further, note δ_{JS}^+ and δ_{SB} are expressed as

$$\delta_{\mathrm{JS}}^+ = \{1 - \psi_{\mathrm{JS}}^+(\|X\|^2/S)\}X \quad \text{with } \psi_{\mathrm{JS}}^+(w) = \min\left(1, \frac{p-2}{n+2}\frac{1}{w}\right),$$

$$\delta_{\mathrm{SB}} = \{1 - \psi_{\mathrm{SB}}(\|X\|^2/S)\}X \quad \text{with } \psi_{\mathrm{SB}}^+(w) = \frac{(p-2)/(n+2)}{w + (p-2)/(n+2) + 1}.$$

Note the risk function of $\{1-\psi(\|X\|^2/S)\}X$ at $\eta\|\theta\|^2=0$ is

$$E\big[\{1-\psi(\|X\|^2/S)\}^2\eta\|X\|^2\big].$$

As in Lemma L.1 in Appendix L, $\psi_{\mathrm{JS}}^{+}(w)>\psi_0(w)>\psi_{\mathrm{SB}}(w)$ for all $w\ge 0$, and hence the risk functions of three estimators at $\eta\|\theta\|^2=0$ satisfy

$$R(\theta,\eta,\delta_{\mathrm{JS}}^{+})<R(\theta,\eta,\delta_0)<R(\theta,\eta,\delta_{\mathrm{SB}}),$$

which can be observed in Figure 1.

For larger $\lambda=\eta\|\theta\|^2$, by Lemma L.1, we have

$$\liminf_{\lambda\to\infty}\lambda^2\{R(\theta,\eta,\delta_0)-R(\theta,\eta,\delta_{\mathrm{SB}})\}\ge 4c_{p,n}(c_{p,n}+1)^2(n-2)^2,$$

$$\liminf_{\lambda\to\infty}\lambda^{n/2+2}\{R(\theta,\eta,\delta_{\mathrm{JS}}^{+})-R(\theta,\eta,\delta_0)\}\ge\frac{(n-2)^{n/2+2}}{B(p/2-1,n/2+2)},$$

which implies that risk plots of these three estimators surely cross each other and that the risk of δ_{SB} is asymptotically smallest among four estimators. This is natural since the simple Bayes estimator, δ_{SB}, is admissible among the class $\mathcal{D}_\psi$ and its risk for smaller $\eta\|\theta\|^2$ is relatively large.

The second graph of Figure 1 gives the corresponding risks for the case of a scaled multivariate-t distribution with 10 degrees of freedom. Graphs in the cases ($p=15$ and $n=5$) and ($p=5$ and $n=15$) are provided in Appendix M. The relative risk behaviors in these cases are largely similar.

When we numerically calculate the risk of the "Harmonic Bayes" estimator given by Corollary 4.1, a new form of ψ_0 is quite helpful. By an integration by parts and change of variables, $\psi_0(w)$ is rewritten as

$$\psi_0(w)=\frac{p-2}{n+2}\frac{1}{w}-\frac{2w^{p/2-2}(1+w)^{-(p+n)/2}}{(n+2)\mathrm{Ibeta}(w/(1+w),p/2-1,n/2+2)},$$

where $\mathrm{Ibeta}(x,\alpha,\beta)$ is the incomplete Beta function given by

$$\mathrm{Ibeta}(x,\alpha,\beta)=\int_0^x\lambda^{\alpha-1}(1-\lambda)^{\beta-1}\,\mathrm{d}\lambda.$$

Clearly $\mathrm{Ibeta}(x,\alpha,\beta)$ may be regarded as the product of the Beta function $B(\alpha,\beta)$ and the cumulative probability function of the Beta distribution $\mathcal{B}(\alpha,\beta)$, which can be easily coded in, for example, Python and R.

5. Proof of general admissibility of X for $p=1,2$. In the Gaussian case, $X\sim N_p(\theta,\eta^{-1}I)$ and $\eta\|U\|^2\sim\chi_n^2$, Kubokawa (2001) in his unpublished lecture notes written in Japanese, showed that when $p=1,2$, the estimator X is admissible among all estimators. Here, we generalize it for our general situation with the underlying density f given by (1.3). For a general prior $g(\theta,\eta)$, we have

$$\begin{aligned}\delta_g(x,u)&=\frac{\int_{\mathbb{R}^p}\int_0^\infty\theta\eta\eta^{(p+n)/2}f(\eta\{\|x-\theta\|^2+\|u\|^2\})g(\theta,\eta)\,\mathrm{d}\theta\,\mathrm{d}\eta}{\int_{\mathbb{R}^p}\int_0^\infty\eta\eta^{(p+n)/2}f(\eta\{\|x-\theta\|^2+\|u\|^2\})g(\theta,\eta)\,\mathrm{d}\theta\,\mathrm{d}\eta}\\&=x+\frac{\int_{\mathbb{R}^p}\int_0^\infty(\theta-x)\eta\eta^{(p+n)/2}f(\eta\{\|x-\theta\|^2+\|u\|^2\})g(\theta,\eta)\,\mathrm{d}\theta\,\mathrm{d}\eta}{\int_{\mathbb{R}^p}\int_0^\infty\eta^{(p+n)/2+1}f(\eta\{\|x-\theta\|^2+\|u\|^2\})g(\theta,\eta)\,\mathrm{d}\theta\,\mathrm{d}\eta}\\&=x-\frac{\int_{\mathbb{R}^p}\int_0^\infty\eta^{(p+n)/2}\nabla_\theta F(\eta\{\|x-\theta\|^2+\|u\|^2\})g(\theta,\eta)\,\mathrm{d}\theta\,\mathrm{d}\eta}{\int_{\mathbb{R}^p}\int_0^\infty\eta^{(p+n)/2+1}f(\eta\{\|x-\theta\|^2+\|u\|^2\})g(\theta,\eta)\,\mathrm{d}\theta\,\mathrm{d}\eta}\\&=x+\frac{\int_{\mathbb{R}^p}\int_0^\infty\eta^{(p+n)/2}F(\eta\{\|x-\theta\|^2+\|u\|^2\})\nabla_\theta g(\theta,\eta)\,\mathrm{d}\theta\,\mathrm{d}\eta}{\int_{\mathbb{R}^p}\int_0^\infty\eta^{(p+n)/2+1}f(\eta\{\|x-\theta\|^2+\|u\|^2\})g(\theta,\eta)\,\mathrm{d}\theta\,\mathrm{d}\eta},\end{aligned}$$

where $F(t) = (1/2)\int_t^\infty f(s)\,\mathrm{d}s$ and the last equality follows from an integration by parts. Hence the estimator X is the generalized Bayes estimator with respect to any improper prior for which $\nabla_\theta g(\theta, \eta) = 0$ and for which the integration by parts is valid, say $g(\theta, \eta) = \pi(\eta)$. Further, let

$$g_i(\theta, \eta) = h_i^2(\eta\|\theta\|^2)\pi(\eta),$$

where h_i is given by (4.5). Clearly $g_i(\theta, \eta)$ approaches $\pi(\eta)$ as $i \to \infty$. Also $g_i(\theta, \eta)$ for any fixed i is integrable under the condition

$$\int_0^\infty \eta^{-p/2}\pi(\eta)\,\mathrm{d}\eta < \infty \tag{5.1}$$

since

$$\int_{\mathbb{R}^p}\int_0^\infty g_i(\theta, \eta)\,\mathrm{d}\theta\,\mathrm{d}\eta = \int_{\mathbb{R}^p}\int_0^\infty \eta^{p/2}h_i^2(\eta\|\theta\|^2)\eta^{-p/2}\pi(\eta)\,\mathrm{d}\theta\,\mathrm{d}\eta$$
$$= c_p\int_0^\infty \lambda^{p/2-1}h_i^2(\lambda)\,\mathrm{d}\lambda\int_0^\infty \eta^{-p/2}\pi(\eta)\,\mathrm{d}\eta,$$

where, by Lemma E.2 of Appendix E, $\int_0^\infty \lambda^{p/2-1}h_i^2(\lambda)\,\mathrm{d}\lambda < \infty$ for $p = 1, 2$.

Let δ_{gi} be the proper Bayes estimator with respect to $g_i(\theta, \eta)$. Then the Bayes risk difference between X and δ_{gi} with respect to $g_i(\theta, \eta)$ is

$$\Delta_i = \int_{\mathbb{R}^p}\int_0^\infty \{R(\theta, \eta, X) - R(\theta, \eta, \delta_{gi})\}g_i(\theta, \eta)\,\mathrm{d}\theta\,\mathrm{d}\eta.$$

We show, in Appendix N, that under Assumptions F.1, F.2 and F.3.1 on f, $\Delta_i \to 0$ as $i \to \infty$, and hence, by Blyth's (1951) theorem (e.g., Theorem 5.6.1 of Brown (1971), not our version for admissibility among the class $\mathcal{D}_\psi$ given in Theorem 4.1), X is admissible among all estimators.

THEOREM 5.1. *Assume Assumptions* F.1, F.2 *and* F.3.1 *on* f. *Then the estimator* X *is admissible for* $p = 1, 2$.

6. Canonical form of the regression setup. Suppose a linear regression model is used to relate y to the p predictors $z_1, \dots, z_p$,

$$y = \alpha 1_m + Z\beta + \eta^{-1/2}\epsilon, \tag{6.1}$$

where α is an unknown intercept parameter, 1_m is an $m \times 1$ vector of ones, $Z = (z_1, \dots, z_p)$ is an $m \times p$ design matrix, and β is a $p \times 1$ vector of unknown regression coefficients. In the error term, η is an unknown scalar and $\epsilon = (\epsilon_1, \dots, \epsilon_m)^{\mathrm{T}}$ has a spherically symmetric distribution,

$$\epsilon \sim \tilde{f}(\|\epsilon\|^2), \tag{6.2}$$

where $\tilde{f}(\cdot)$ is the probability density, $E[\epsilon] = 0_m$, and $\mathrm{Var}[\epsilon] = I_m$. Hence the density of y is

$$y \sim \eta^{m/2}\tilde{f}(\eta\|y - \alpha 1_m - Z\beta\|^2), \tag{6.3}$$

where $\tilde{f}$ satisfies

$$\int_{\mathbb{R}^m}\tilde{f}(\|v\|^2)\,\mathrm{d}v = 1$$

for $v = (v_1, \dots, v_m)^{\mathrm{T}} \in \mathbb{R}^m$. We assume that the columns of Z have been centered so that $z_i^{\mathrm{T}}1_m = 0$ for $1 \le i \le p$. We also assume that $m > p + 1$ and $\{z_1, \dots, z_p\}$ are linearly independent, which implies that

$$\operatorname{rank} Z = p.$$

Let Q be an $m \times m$ orthogonal matrix of the form

$$Q = (1_m/\sqrt{m}, Z(Z^{\mathrm{T}}Z)^{-1/2}, W),$$

where W is $m \times (m-p-1)$ matrix which satisfies $W^{\mathrm{T}}1_m = 0$, $W^{\mathrm{T}}Z = 0$ and $W^{\mathrm{T}}W = I_{m-p-1}$. Also let $x = (Z^{\mathrm{T}}Z)^{-1/2}Z^{\mathrm{T}}y = (Z^{\mathrm{T}}Z)^{1/2}\hat{\beta}_{\mathrm{LSE}} \in \mathbb{R}^p$ where $\hat{\beta}_{\mathrm{LSE}} = (Z^{\mathrm{T}}Z)^{-1}Z^{\mathrm{T}}y$.

Let

$$Q^{\mathrm{T}}y = (\sqrt{m}\bar{y}, x^{\mathrm{T}}, u^{\mathrm{T}})^{\mathrm{T}},$$

where $u = W^{\mathrm{T}}y \in \mathbb{R}^{m-p-1}$. Then $(\sqrt{m}\bar{y}, x, u)$ are sufficient and the joint density of $(\sqrt{m}\bar{y}, x, u)$ is

$$\eta^{m/2}\tilde{f}(\eta\{m(\bar{y}-\alpha)^2 + \|x-\theta\|^2 + \|u\|^2\}),$$

where $\theta = (Z^{\mathrm{T}}Z)^{1/2}\beta$. Further the marginal density of (x, u) is

$$\eta^{(m-1)/2}f(\eta\{\|x-\theta\|^2 + \|u\|^2\}),$$

which we are considering in this paper, where $m-1 = p+n$ and

$$f(t) = \int_{-\infty}^{\infty}\tilde{f}(v^2+t)\,\mathrm{d}v.$$

Note that the loss function $\eta\|\delta-\theta\|^2$ corresponds to so-called "predictive loss" $\eta\|Z\hat{\beta}-Z\beta\|^2$ for estimation of the regression coefficient vector β.

In the equivariant estimator δ_ψ of θ,

$$\{1-\psi(\|x\|^2/s)\}x,$$

$\|x\|^2/s$ is $R^2/(1-R^2)$ in the regression context where R^2 is the coefficient of determination. It is natural to make use of R^2 for shrinkage since small R^2 corresponds to less reliability of the least squares estimator of β. We note that the corresponding "simple Bayes estimator" for regression coefficient β is rewritten as

$$\left(1 - \frac{a}{(a+1)(b+1) + R^2/(1-R^2)}\right)\hat{\beta}_{\mathrm{LSE}}$$

and has a shrinkage factor which is increasing in R^2.

In the equivariant estimator $\delta_\xi = \{1-\xi(x/\sqrt{s})\}x \in \mathcal{D}_\xi$,

$$\frac{x}{\sqrt{s}} = \frac{(Z^{\mathrm{T}}Z)^{1/2}\hat{\beta}_{\mathrm{LSE}}}{\sqrt{m-p-1}\hat{\sigma}} = \frac{(Z^{\mathrm{T}}Z)^{1/2}}{\sqrt{m-p-1}}\frac{\hat{\beta}_{\mathrm{LSE}}}{\hat{\sigma}}, \tag{6.4}$$

where $\hat{\sigma} = \sqrt{s/(m-p-1)}$. Under the Gaussian assumption, $\hat{\beta}_{\mathrm{LSE}}/\hat{\sigma}$ is a vector of the noncentral t-values.

Hence the restriction to $\mathcal{D}_\psi$ or $\mathcal{D}_\xi$ is quite natural in regression context. The minimaxity and admissibility results of Sections 3 and 4 provide some guidance as to reasonable shrinkage estimators in the regression context.

7. Concluding remarks. We have established admissibility of certain generalized Bayes estimators within the class of equivariant estimators, of the mean vector for a spherically symmetric distribution with unknown scale under invariant loss. In some cases, we establish simultaneous minimaxity and, equivariant admissibility for broader classes of sampling distributions. In the Gaussian case, we establish admissibility within the equivariant estimators of a class of generalized Bayes minimax estimators of a particularly simple form. We have also investigated similar issues in the setting of a general linear regression model with intercept and spherically symmetric error distribution. In this setting, the shrinkage factor of equivariant estimators of the regression coefficients depends on the coefficient of determination.

Acknowledgments. The authors thank the referees and Associate Editor for their constructive and insightful comments that helped improve this manuscript.

The first author was supported in part by KAKENHI #25330035, #16K00040.

The second author was supported in part by grants from the Simons Foundation (#209035 and #418098).

SUPPLEMENTARY MATERIAL

Supplement to "Admissible Bayes equivariant estimation of location vectors for spherically symmetric distributions with unknown scale" (DOI: 10.1214/19-AOS1837SUPP; .pdf). Supplementary information.

REFERENCES

BLYTH, C. R. (1951). On minimax statistical decision procedures and their admissibility. *Ann. Math. Stat.* **22** 22–42. MR0039966 https://doi.org/10.1214/aoms/1177729690

BREWSTER, J. F. and ZIDEK, J. V. (1974). Improving on equivariant estimators. *Ann. Statist.* **2** 21–38. MR0381098

BROWN, L. D. (1971). Admissible estimators, recurrent diffusions, and insoluble boundary value problems. *Ann. Math. Stat.* **42** 855–903. MR0286209 https://doi.org/10.1214/aoms/1177693318

BROWN, L. D. and HWANG, J. T. (1982). A unified admissibility proof. In *Statistical Decision Theory and Related Topics*, *III*, *Vol.* 1 (*West Lafayette*, *Ind.*, 1981) 205–230. Academic Press, New York. MR0705290

CELLIER, D., FOURDRINIER, D. and ROBERT, C. (1989). Robust shrinkage estimators of the location parameter for elliptically symmetric distributions. *J. Multivariate Anal.* **29** 39–52. MR0991055 https://doi.org/10.1016/0047-259X(89)90075-4

DASGUPTA, A. and STRAWDERMAN, W. E. (1997). All estimates with a given risk, Riccati differential equations and a new proof of a theorem of Brown. *Ann. Statist.* **25** 1208–1221. MR1447748 https://doi.org/10.1214/aos/1069362745

FELLER, W. (1971). *An Introduction to Probability Theory and Its Applications*. *Vol. II*. 2*nd ed*. Wiley, New York. MR0270403

JAMES, W. and STEIN, C. (1961). Estimation with quadratic loss. In *Proc*. 4*th Berkeley Sympos*. *Math*. *Statist*. *and Prob*., *Vol. I* 361–379. Univ. California Press, Berkeley, Calif. MR0133191

KUBOKAWA, T. (1994). A unified approach to improving equivariant estimators. *Ann. Statist.* **22** 290–299. MR1272084 https://doi.org/10.1214/aos/1176325369

KUBOKAWA, T. (2001). Shrinkage estimation and its applications. Unpublished lecture notes written in Japanese.

MARUYAMA, Y. (2003). A robust generalized Bayes estimator improving on the James–Stein estimator for spherically symmetric distributions. *Statist. Decisions* **21** 69–77. MR1985652 https://doi.org/10.1524/stnd.21.1.69.20318

MARUYAMA, Y. (2009). An admissibility proof using an adaptive sequence of smoother proper priors approaching the target improper prior. *J. Multivariate Anal.* **100** 1845–1853. MR2535391 https://doi.org/10.1016/j.jmva.2009.02.012

MARUYAMA, Y. and STRAWDERMAN, W. E. (2005). A new class of generalized Bayes minimax ridge regression estimators. *Ann. Statist.* **33** 1753–1770. MR2166561 https://doi.org/10.1214/009053605000000327

MARUYAMA, Y. and STRAWDERMAN, W. E. (2009). An extended class of minimax generalized Bayes estimators of regression coefficients. *J. Multivariate Anal.* **100** 2155–2166. MR2560360 https://doi.org/10.1016/j.jmva.2009.06.001

MARUYAMA, Y. and STRAWDERMAN, W. E. (2017). A sharp boundary for SURE-based admissibility for the normal means problem under unknown scale. *J. Multivariate Anal.* **162** 134–151. MR3719340 https://doi.org/10.1016/j.jmva.2017.09.002

MARUYAMA, Y. and STRAWDERMAN, W. E. (2020). Supplement to "Admissible Bayes equivariant estimation of location vectors for spherically symmetric distributions with unknown scale." https://doi.org/10.1214/19-AOS1837SUPP.

MARUYAMA, Y. and TAKEMURA, A. (2008). Admissibility and minimaxity of generalized Bayes estimators for spherically symmetric family. *J. Multivariate Anal.* **99** 50–73. MR2408443 https://doi.org/10.1016/j.jmva.2007.01.002

STEIN, C. (1956). Inadmissibility of the usual estimator for the mean of a multivariate normal distribution. In *Proceedings of the Third Berkeley Symposium on Mathematical Statistics and Probability*, 1954–1955, *Vol. I* 197–206. Univ. California Press, Berkeley and Los Angeles. MR0084922

STEIN, C. (1974). Estimation of the mean of a multivariate normal distribution. 345–381. MR0381062

STRAWDERMAN, W. E. (1971). Proper Bayes minimax estimators of the multivariate normal mean. *Ann. Math. Stat.* **42** 385–388. MR0397939 https://doi.org/10.1214/aoms/1177693528

STRAWDERMAN, W. E. (1973). Proper Bayes minimax estimators of the multivariate normal mean vector for the case of common unknown variances. *Ann. Statist.* **1** 1189–1194. MR0365806