A unified and generalized set of shrinkage bounds on minimax Stein estimates

Dominique Fourdrinier\textsuperscript{a,*}, William E. Strawderman\textsuperscript{b}

\textsuperscript{a} Université de Rouen, LITIS, BP 12, 76801 Saint-Étienne-du-Rouvray, France
\textsuperscript{b} Rutgers University, Hill Center, Department of Statistics, 10 Frelinghuysen Road, Piscataway, NJ 08854, USA

Received 18 July 2007
Available online 13 February 2008

Abstract

Consider the problem of estimating the mean vector $\theta$ of a random variable $X$ in $\mathbb{R}^p$, with a spherically symmetric density $f(\|x - \theta\|^2)$, under loss $\|\delta - \theta\|^2$. We give an increasing sequence of bounds on the shrinkage constant of Stein-type estimators depending on properties of $f(t)$ that unify and extend several classical bounds from the literature. The basic way to view the conditions on $f(t)$ is that the distribution of $X$ arises as the projection of a spherically symmetric vector $(X, U)$ in $\mathbb{R}^{p+k}$. A second way is that $f(t)$ satisfies $(-1)^j f^{(j)}(t) \geq 0$ for $0 \leq j \leq \ell$ and that $(-1)^\ell f^{(\ell)}(t)$ is non-increasing where $k = 2(\ell + 1)$. The case $\ell = 0$ ($k = 2$) corresponds to unimodality, while the case $\ell = k = \infty$ corresponds to complete monotonicity of $f(t)$ (or equivalently that $f(\|x - \theta\|^2)$ is a scale mixture of normals). The bounds on the minimax shrinkage constant in this paper agree with the classical bounds in the literature for the case of spherical symmetry, spherical symmetry and unimodality, and scale mixtures of normals. However, they extend these bounds to an increasing sequence (in $k$ or $\ell$) of minimax bounds.

© 2008 Elsevier Inc. All rights reserved.

AMS 2000 subject classifications: primary 62C10; 62C20

Keywords: Minimax estimators; Quadratic loss; Spherically symmetric distributions; Location parameters; Completely monotone functions; Unimodality

\* Corresponding author.
E-mail address: Dominique.Fourdrinier@univ-rouen.fr (D. Fourdrinier).
1. Introduction

Let \((X, U)\) be a random vector in \(\mathbb{R}^{p+k}\) spherically symmetric around \((\theta, 0)\) \(\in \mathbb{R}^{p+k}\) such that \(\dim X = \dim \theta = p \geq 3\) and \(\dim U = \dim 0 = k \geq 1\); we use the notation 
\[
(X, U) \sim \text{ss}_{p+k}(\theta, 0). \tag{1}
\]
We will study estimators \(\delta(X)\) of \(\theta\), that is, estimators which only depend on \(X\), under the usual quadratic loss function
\[
L(\theta, \delta) = \|\delta - \theta\|^2. \tag{2}
\]
To assure finiteness of the risk function of \(X\), we assume throughout that \(E_\theta [\|X - \theta\|^2] = E_0 [\|X\|^2] < \infty\) which is equivalent to \(E[R^2] < \infty\) where \(R^2 = \|X - \theta\|^2 + \|U\|^2\). We also assume that \(E_0 \left[ \frac{1}{\|X\|^2} \right] < \infty\) which implies \(E \left[ \frac{1}{R^2} \right] < \infty\).

An equivalent version of the problem is to observe
\[
X \sim f(\|x - \theta\|^2) \tag{3}
\]
where \(f(\cdot)\) is the density of the projection of \((X, U) \in \mathbb{R}^{p+k}\) onto \(\mathbb{R}^p\). Such a density exists for all \(k \geq 1\) even if the distribution of \((X, U)\) in \(\mathbb{R}^{p+k}\) has no density. In particular, the density \(f(\cdot)\) is unimodal for all \(k \geq 2\).

The main class of estimators that we consider are shrinkage estimators of the James–Stein or Baranchik form
\[
\delta_B(X) = \left(1 - b \frac{r(\|X\|^2)}{\|X\|^2}\right) X \tag{4}
\]
although our principal technical result applies to the more general class
\[
\delta(X) = X + ag(X) \tag{5}
\]
under conditions on \(g(\cdot)\) comparable to those in [5].

Our main result gives an increasing sequence of bounds on the shrinkage constant, \(a\), in (5) depending on \(k\), which ensures minimaxity of the estimator. When applied to the James–Stein estimator with \(p \geq 4\),
\[
\delta_{JS}(X) = \left(1 - \frac{b}{\|X\|^2}\right) X, \tag{6}
\]
the result implies minimaxity for
\[
0 \leq b \leq 2 \frac{p - 2 + k}{p + k} \frac{1}{E_0 \left[ \frac{1}{\|X\|^2} \right]} . \tag{7}
\]
This bound unifies and extends several bounds from the literature. The case \(k = \infty\) corresponds to the original bound of [10] in the normal case and to the bound of [12] in the scale mixture of the normal case. In each of these cases the \(p\)-dimensional distribution arises as a projection from a distribution in \(p + k\) dimensions for all \(k \geq 0\). The bound for \(k = 2\) corresponds to the bound given in [4] for spherically symmetric unimodal distributions (which arise as projections from spherically symmetric distributions in \(p + 2\) dimensions). Interestingly, the bound for \(k = 0\)
corresponds to the bound in [3,5] for general spherically symmetric distributions, although our theorem technically requires $k \geq 1$. Section 2 is devoted to a proof of these domination results.

In Section 3, we study properties of densities $f(\|x - \theta\|^2)$ in (3), which arise as projections from $p + k$ dimensions. In particular we show that if $(-1)^j f^{(j)}(t) \geq 0$ on $(0, \infty)$ for $0 \leq j \leq \ell$, then $f(\|x - \theta\|^2)$ is the density of a projection from $p + 2(\ell + 1)$ dimensions, provided that $f^{(\ell)}(t)$ is non-increasing and $f^{(\ell - 1)}(t)$ is a primitive of $f^{(\ell)}(t)$. Thus if $f(t) = f^{(0)}(t)$ is non-increasing, the density $f(\|x - \theta\|^2)$ is unimodal and arises as a projection from $p + 2$ dimensions. If $\ell = \infty$, then $f(t)$ is completely monotone and the corresponding distribution is a scale mixture of normals as noted in [2]. After completing this paper, we became aware of [7] which gives conditions for a distribution to arise as the marginal distribution of a spherically symmetric distribution in higher dimension. Eaton’s results imply Theorem 3.2 for which we give a self-contained proof.

It follows that the results of Section 2 may be reinterpreted as an increasing sequence of bounds on the shrinkage constant, $a$, which ensure minimaxity of Stein-type shrinkage estimators depending on the degree to which $f(t)$ is “completely monotone”. That is, if $(-1)^j f^{(j)}(t) \geq 0$ for $0 \leq j \leq \ell$, then the minimax bound for $b$ becomes

$$0 \leq a \leq 2 \frac{p + 2\ell}{p + 2(\ell + 1)} \frac{1}{E_0 \left[ \frac{1}{\|X\|^2} \right]},$$

i.e., the bound is that given by (7) with $k = 2(\ell + 1)$.

In Section 4, two examples illustrate the theory. Section 5 has some concluding remarks, and the Appendix gives several technical lemmas needed in Sections 2 and 3.

2. Dominating minimax results

In this section, we give our main theoretical result. In Section 3, we will use it to unify and extend the minimax results mentioned in the introduction. Here is the main result.

**Theorem 2.1.** Let $(X, U)$ be a random vector as in (1) and let $\delta(X) = X + ag(X)$ be an estimator of $\theta$ where $a$ is a real valued constant and $g(\cdot)$ is a weakly differentiable function such that $E_\theta \left[ \|g(X)\|^2 \right] < \infty$. Assume that $g$ satisfies the following conditions:

(a) there exists a function $h(\cdot)$ such that, for all $x \in \mathbb{R}^p$,

$$\frac{1}{2} \|g(x)\|^2 \leq h(x) \leq -\text{div} \, g(x);$$

(b) $h(\cdot)$ is superharmonic;

(c) $E_\theta \left[ R^2 h(X) \|X - \theta\|^2 + \|U\|^2 = R^2 \right]$ is non-decreasing in $R^2$.

Then, under the loss function (2), $\delta(X)$ dominates $X$ provided

$$0 < a < \frac{p - 2 + k}{(p + k)(p - 2)} \frac{1}{E_0 \left[ \frac{1}{\|X\|^2} \right]}.$$

**Proof.** We work throughout conditionally on $\|X - \theta\|^2 + \|U\|^2 = R^2$, so that the distribution of $(X, U)$ is the uniform distribution on the sphere of radius $R$ centered at $(\theta, 0)$, for which the expectation is denoted by $E^{(X, U)}_{\theta, 0}$. Note that Condition (c) can be expressed as $R^2 E^{(X, U)}_{\theta, 0} [h(X)]$ is non-decreasing in $R^2$ and hence is a condition on $g(\cdot)$ and does not depend on the underlying distribution of $(X, U)$. We will use the fact that an additional conditioning with respect to $U$
results in the uniform distribution (of $X$) on the sphere of radius $(R^2 - \|U\|^2)^{1/2}$ centered at $\theta$.

We denote by $E_{R,\theta}^{X|U}$ the corresponding expectation and by $E_U^{X}$ the expectation with respect to $U$ conditionally on $R$.

As in [9], we use the identity

$$E_R^{(X,U)}[(X - \theta)'g(X)] = E_R^{(X,U)}\left[\frac{\|U\|^2}{k} \text{div} g(X)\right].$$

Then the risk difference, conditional on $R$, between $\delta(X) = X + ag(X)$ and $X$ equals

$$\Delta_R(\theta) = E_R^{(X,U)}[a^2g(X) + 2a(X - \theta)'g(X)]$$

$$= E_R^{(X,U)}\left[a^2g(X)^2 + 2a \frac{\|U\|^2}{k} \text{div} g(X)\right]$$

$$\leq 2aE_R^{(X,U)}[h(X)\left(a - \frac{\|U\|^2}{k}\right)] \quad (8)$$

by Condition (a). Also conditioning on $U$, the function

$$H_\theta \left(R^2 - \|U\|^2\right) = E_{R,\theta}^{X|U}[h(X)] \quad (9)$$

is, by Condition (b), non-increasing in $R^2 - \|U\|^2$ (since the distribution of $X|U$, $R$ is uniform on a sphere of radius $(R^2 - \|U\|^2)^{1/2}$), and hence non-decreasing in $\|U\|^2$ for fixed $R^2$. Therefore, according to (8) and (9), we have

$$\Delta_R(\theta) \leq 2aE_U^{X}[E_{R,\theta}^{X|U}[h(X)]\left(a - \frac{\|U\|^2}{k}\right)]$$

$$= 2aE_U^{(X,U)}\left[H_\theta(R^2 - \|U\|^2)\left(a - \frac{\|U\|^2}{k}\right)\right]$$

$$\leq 2aE_U^{(X,U)}[H_\theta(R^2 - \|U\|^2)]E_U^{X}\left(a - \frac{\|U\|^2}{k}\right) \quad (10)$$

by the covariance inequality. Also

$$E_U^{\|U\|^2} = E_R^{(X,U)}\left[\frac{X - \theta\|U\|^2 + \|U\|^2}{p + k}\right] = \frac{R^2}{p + k} \quad (11)$$

by the exchangeability property of a spherically symmetric distribution. Then (10) and (11) lead to

$$\Delta_R(\theta) \leq 2aE_U^{(X,U)}\left[R^2H_\theta(R^2 - \|U\|^2)\left(a - \frac{1}{p + k}\right)\right]$$

$$= 2aE_U^{(X,U)}\left[R^2h(X)\left(a - \frac{1}{p + k}\right)\right]$$

according to (9).

By Condition (c) and the covariance inequality, the unconditional risk difference $\Delta(\theta)$ satisfies
\[ \Delta(\theta) = E R[\Delta_R(\theta)] \]
\[ \leq 2 a E R \left[ E_{R,\theta}^{(X,U)}[R^2 h(X)] \right] \left( a E R \left[ \frac{1}{R^2} \right] - \frac{1}{p + k} \right). \]

Finally we have \( \Delta(\theta) \leq 0 \) since \( h(\cdot) \geq 0 \) and the condition on \( a \) reduces to \( 0 < a < \frac{1}{p+k} \frac{1}{E R \left[ \frac{1}{R^2} \right]} \).

Indeed we have
\[
E_0 \left[ \frac{1}{\|X\|^2} \right] = E_0 \left[ \frac{1}{R^2} \frac{R^2}{\|X\|^2} \right] = E \left[ \frac{1}{R^2} \right] E_0 \left[ \frac{R^2}{\|X\|^2} \right] = E \left[ \frac{1}{R^2} \right] \left( \frac{p - 2 + k}{p - 2} \right)
\]

by independence of \( R^2 \) and \( \frac{R^2}{\|X\|^2} \) and the fact that, if \( \theta = 0, \frac{\|X\|^2}{R^2} \sim \text{Beta} \left( \frac{p}{2}, \frac{k}{2} \right). \)

We derive from Theorem 2.1 two corollaries valid for \( p \geq 4 \). This restriction, rather than \( p \geq 3 \), results from the fact that \( \frac{1}{\|X\|^2} \) is not superharmonic when \( p = 3 \).

**Corollary 2.1.** The James–Stein estimator \( \delta_{JS}(X) = X - \frac{b}{\|X\|^2} X \) dominates \( X \) provided \( p \geq 4 \) and
\[
0 < b < 2 \frac{p - 2 + k}{p + k} \frac{1}{E_0 \left[ \frac{1}{\|X\|^2} \right]}.
\]

**Proof.** Let \( g(x) = -\frac{2(p-2)}{\|x\|^2} x \) and note that \( \text{div} \ g(x) = -\frac{2(p-2)^2}{\|x\|^2} \) and \( \|g(x)\|^2 = \frac{4(p-2)^2}{\|x\|^2} \) so that a natural choice for the function \( h(\cdot) \) in Condition (a) is \( h(x) = \frac{2(p-2)^2}{\|x\|^2} \). Since \( \frac{4}{\|x\|^2} \) is superharmonic for \( p \geq 4 \), Condition (b) is satisfied. Finally Condition (c) is satisfied according to Lemma A.1 and then the corollary follows from Theorem 2.1 with \( a = \frac{b}{p-2} \).

The context of (1) necessitates \( k \geq 1 \). However, it is worth noting that \( k = 0 \) in (12) corresponds to the bound in [3] and also [5].

**Corollary 2.2.** Let \( r(\cdot) \) be a concave function bounded between 0 and 1. The Baranchik estimator \( \delta_B(X) = X - b \frac{r(\|X\|^2)}{\|X\|^2} X \) dominates \( X \) provided \( p \geq 4, k \geq 1 \) and
\[
0 < b < 2 \frac{p - 2 + k}{p + k} \frac{1}{E_0 \left[ \frac{1}{\|X\|^2} \right]}.
\]

**Proof.** Let \( g(x) = -2(p-2) \frac{r(\|x\|^2)}{\|x\|^2} x \) and note that \( \text{div} \ g(x) = -2(p-2) \left[ \frac{p-2}{\|x\|^2} r(\|x\|^2) + 2r'(\|x\|^2) \right] \) and \( \|g(x)\|^2 = 4(p-2)^2 \frac{r(\|x\|^2)}{\|x\|^2} \). Hence \( h(x) = 2(p-2)^2 \frac{r(\|x\|^2)}{\|x\|^2} \) satisfies Condition (a). Since \( r(\cdot) \) is concave, \( \frac{r(\|x\|^2)}{\|x\|^2} \) is superharmonic for \( p \geq 4 \) (see, for example, [8]). Therefore Condition (b) is satisfied. Finally, by Lemma A.2 and the fact that the concavity of \( r(\cdot) \) also implies that \( \frac{r(t)}{t} \) is non-increasing (see also [8]), Condition (c) is satisfied and the corollary follows from Theorem 2.1 with \( a = \frac{b}{p-2} \).

3. Unification and extension of minimax results

In this section, we concentrate on minimax domination results when the observed vector \( X \) has a spherically symmetric distribution in \( \mathbb{R}^p \) around a location parameter \( \theta \in \mathbb{R}^p \). The main feature of the distribution of \( X \) that we consider is that it can be viewed as the distribution of a projection from \( \mathbb{R}^{p+k} \) onto \( \mathbb{R}^p \). Two particularly interesting classes of distributions are the class of unimodal distributions studied in [4] and the class of scale mixtures of normal distributions studied in [12]. We will show that the first class corresponds to \( k = 2 \) and the second class to \( k = \infty \). Furthermore we will show that the bounds of the shrinkage constant given in Theorem 2.1 and its corollaries correspond to the respective bounds in the above mentioned papers.

Our first result makes explicit the projection phenomenon for these two classes of distributions.

**Theorem 3.1.**

1. Any unimodal spherically symmetric distribution in \( \mathbb{R}^p \) is the distribution of the projection, onto \( \mathbb{R}^p \), of a random vector in \( \mathbb{R}^{p+2} \) having a spherically symmetric distribution.

2. Any scale mixture of spherical normal distributions in \( \mathbb{R}^p \) is, for any \( k \geq 0 \), the distribution of the projection, onto \( \mathbb{R}^p \), of a random vector in \( \mathbb{R}^{p+k} \) having a spherically symmetric distribution (in particular, a scale mixture of normals).

**Proof.**
1. Recall that any unimodal spherically symmetric distribution \( P \) in \( \mathbb{R}^p \) is a radial mixture of uniform distributions on balls in \( \mathbb{R}^p \) with the same center (cf. [4]). Also note that, by Lemma A.3, the uniform distribution on a ball in \( \mathbb{R}^p \) is the distribution of the projection \( \Pi \) of a random vector having a uniform distribution on the corresponding sphere in \( \mathbb{R}^{p+2} \). By Lemma A.4, if \( G \) is the mixing distribution on the balls, then, for any Borel set \( A \) in \( \mathbb{R}^p \),

\[
P(A) = \int_{\mathbb{R}^{p+2}}_{U_{p+2}} (U_{R,\theta})^p(A) dG(R)
\]

\[
= \int_{\mathbb{R}^{p+2}} (U_{R,\theta})^p(A) dG(R)
\]

\[
= Q_{\Pi}(A)
\]

where, for any Borel set \( C \) in \( \mathbb{R}^{p+2} \),

\[
Q(C) = \int_{\mathbb{R}^{p+2}} (U_{R,\theta})^p(C) dG(R).
\]

Thus \( P = Q_{\Pi} \) where \( Q \) is the spherically symmetric distribution in \( \mathbb{R}^{p+2} \) defined through the same mixing distribution \( G \). Hence the first part follows.

2. The second part of the theorem follows similarly since, for any \( k \geq 0 \), the projection of a spherical normal distribution in \( \mathbb{R}^{p+k} \) is the spherical normal distribution in \( \mathbb{R}^p \) (with the same scale parameter). □

We now show that the classical bounds on the shrinkage constant for the case of spherically symmetric unimodal distributions and for the case of scale mixtures of normals follow from Theorem 2.1 and the results of Section 2.

**Corollary 3.1.** Let \( X \sim s.s.(\theta) \) in \( \mathbb{R}^p \). Let \( \delta_b(X) = X - b \frac{r(\|X\|^2)}{\|X\|^2} X \) be a Baranchik estimator with \( r \) a concave function bounded between 0 and 1. Then, under the loss function (2), \( \delta_b(X) \) dominates \( X \) provided \( p \geq 4 \) and
(1) \(0 < b < 2 \frac{p}{p+2} \frac{1}{E_0} \frac{1}{|x|^p}\) and the distribution of \(X\) is unimodal or

(2) \(0 < b < 2 \frac{1}{E_0} \frac{1}{|x|^p}\) and the distribution of \(X\) is a scale mixture of normal distributions.

The upper bound for the constant \(b\) in (1) agrees with that in [4] while the upper bound in (2) agrees with that in [12].

**Proof.** As noted in Theorem 3.1(1), if the spherically symmetric distribution of \(X\) is unimodal in \(\mathbb{R}^p\), it is the distribution of \(X\) in \((X, U)\) in (1) with \(k = 2\). Thus, by Corollary 2.2, part (1) follows.

Similarly, for part (2), Theorem 3.1(2) implies that, if the distribution of \(X\) in \(\mathbb{R}^p\) is a scale mixture of normals, for any \(k \geq 0\), it is the distribution of \(X\) in \((X, U)\) in (1). Hence, for any \(k > 0\), Corollary 2.2 implies that \(\delta_B(X)\) dominates \(X\) as soon as \(0 < b < 2 \frac{1}{E_0} \frac{1}{|x|^p}\) since

\[
\lim_{k \to \infty} \frac{p-2+k}{p+k} = 1. \quad \square
\]

We next extend the above discussion to classes of densities such that \((-1)^j f^{(j)}(t) \geq 0\) for \(j = 0, \ldots, \ell\) with \(f^{(\ell)}(t)\) monotone non-increasing. Recall that if \(\ell = 0\), \(f(\cdot)\) is non-increasing and the density is unimodal; and that if \(\ell = \infty\), the function \(f(\cdot)\) is completely monotone, and the corresponding density is representable as a scale mixture of normals (see [2]). Hence, if \(C_\ell\) represents the above class of distributions, \(C_0 = \{\text{unimodals}\}\), \(C_\infty = \{\text{scale mixture of normals}\}\) and \(C_j \supset C_{j+1}\) for all \(j \geq 0\).

The next result and particularly its corollary establish a connection between the classes \(C_\ell\) and projections from \(\mathbb{R}^{p+2(\ell+1)}\) into \(\mathbb{R}^p\).

**Theorem 3.2.** Suppose that \(f(\|x\|^2)\) is a density in \(\mathbb{R}^p\) such that \(f(\cdot)\) is \(\ell\)-times differentiable and such that, for any \(x \in [0, \infty]\),

(i) \(\int_0^x f^{(\ell)}(t) \, dt = f^{(\ell-1)}(x) - f^{(\ell-1)}(0)\).

Suppose also that

(ii) \((-1)^j f^{(j)}(t) \geq 0\) on \([0, \infty)\), for \(0 \leq j \leq \ell\)

and

(iii) \((-1)\ell f^{(\ell)}(t)\) is non-increasing on \((0, \infty)\).

Then \(f(\cdot) = f^{(0)}(\cdot)\) has the representation

\[
f(t) = \frac{C}{\ell!} \int_0^\infty \frac{(R^2 - t)^\ell I_{[0,R]}(t)}{R^p} \, dG(R),
\]

for some probability measure \(G\) on \((0, \infty)\) and some positive constant \(C\).

**Remark.** In Theorem 3.2, the derivatives of \(f\) at 0 are considered as right hand derivatives.

**Proof.** Since \(f(\|x\|^2)\) is a density in \(\mathbb{R}^p\) and \(f(\cdot)\) is non-increasing on \([0, +\infty)\), it follows that \(t^{p/2} f(t) \to 0\) as \(t \to \infty\), since

\[
\infty > \int_{t/2}^t u^{(p-2)/2} f(u) \, du > \frac{2}{p} f(t) \int_{t/2}^t u^{(p-2)/2} \, du = f(t) t^{p/2}(1 - 1/2^{p/2}).
\]
An integration by parts then gives that
\[
\int_0^\infty u^{p/2} (-1) f'(u) \, du = -u^{p/2} f(u) \bigg|_0^\infty + \frac{p}{2} \int_0^\infty u^{p/2-1} f(u) \, du
\]
\[\leq \frac{p}{2} \int_0^\infty u^{p/2-1} f(u) \, du < \infty .\]

In the same way, it follows by induction that
\[
\int_0^\infty u^{(p-2)/2} (-1)^j f^{(j)}(u) \, du < \infty \quad \text{for } 1 \leq j \leq \ell
\]
which implies, in conjunction with (i), that
\[
\int_0^\infty u^{(p-2)/2} (-1)^\ell f^{(\ell)}(u) \, du < \infty .\]

Therefore \((-1)^\ell f^{(\ell)}(\|x\|^2)\) is, up to a positive constant multiple \(C\), a spherically symmetric unimodal density on \(\mathbb{R}^p\), and hence (as in Theorem 3.1) has a representation as a mixture of uniform distributions on balls of radius \(R\), i.e.,
\[
(-1)^\ell f^{(\ell)}(\|x\|^2) = C \int_0^\infty \frac{I_{B(0,R^2)}(x)}{R^p} \, dG(R),
\]
or equivalently
\[
(-1)^\ell f^{(\ell)}(t) = C \int_0^\infty \frac{I_{[0,R^2]}(t)}{R^p} \, dG(R).
\]

By an argument similar to that used at the beginning of the proof, it follows from (i) and (ii) that, for \(1 \leq j \leq k - 1\), \(f^{(j)}(t)\) is bounded in a neighborhood of 0 and, for \(1 \leq j \leq k\),
\[
0 < \int_0^\infty (-1)^j f^{(j)}(t) \, dt = (-1)^{j-1} f^{(j-1)}(0) < \infty .
\]

Now the representation of \(f\) will be obtained through induction for \(j = 0, \ldots, \ell\) on \((-1)^{\ell-j} f^{(\ell-j)}(\|x\|^2)\). Using the finiteness of the one-dimensional integrals as shown above and Fubini’s theorem,
\[
f^{(\ell-1)}(t) = (-1) \int_t^\infty f^{(\ell)}(u) \, du
\]
\[= (-1)^{\ell-1} C \int_t^\infty \int_0^\infty \frac{I_{[0,R^2]}(u)}{R^p} dG(R) \, du
\]
\[= (-1)^{\ell-1} C \int_t^\infty \left\{ \int_0^\infty \frac{I_{[0,R^2]}(u) I_{[t,\infty]}(u)}{R^p} \, du \right\} dG(R)
\]
\[= (-1)^{\ell-1} D \int_t^\infty \left\{ \int_0^{R^2} \frac{I_{[0,R^2]}(t)}{R^p} \, du \right\} dG(R)
\]
\[= (-1)^{\ell-1} C \int_t^\infty (R^2 - t) \frac{I_{[0,R^2]}(t)}{R^p} \, dG(R).
\]
Now assume that for \( i = 1, \ldots, j \) (for \( j < \ell \)), we have that

\[
f^{(\ell-i)}(t) = \frac{(-1)^{\ell-i} C}{i!} \int_0^\infty \frac{(R^2 - t)^i I_{[0, R^2]}(t)}{R^p} dG(R).
\]

Then

\[
f^{(\ell-(j+1))}(t) = (-1)^j \int_0^\infty f^{(\ell-j)}(u) du
= \frac{(-1)^{\ell-(j+1)} C}{j!} \int_0^\infty \int_0^\infty \frac{(R^2 - u)^j I_{[0, R^2]}(u) I_{[t, \infty)}(u)}{R^p} dG(R) du
= \frac{(-1)^{\ell-(j+1)} C}{j!} \int_0^\infty \left\{ \int_0^R \frac{(R^2 - u)^j I_{[0, R^2]}(u)}{R^p} du \right\} dG(R)
= \frac{(-1)^{\ell-(j+1)} C}{(j+1)!} \int_0^\infty \frac{(R^2 - t)^{j+1} I_{[0, R^2]}(t)}{R^p} dG(R).
\]

The result then follows by taking \( j = \ell - 1 \) in the above (induction) step. \( \square \)

**Corollary 3.2.** Assume that \( f(\cdot) \) satisfies the assumptions of **Theorem 3.2**. Then there exists a random vector \((X, U)\) in \(\mathbb{R}^{p+2(\ell+1)}\) such that \((X, U) \sim ss_{p+2(\ell+1)}(\theta, 0)\) and the marginal density of \(X\) is \(f(\|x - \theta\|^2)\).

**Proof.** The exponent \(\ell\) in (13) corresponds to \(k = 2(\ell+1)\) in Lemma A.3 and hence the corollary follows by **Theorem 3.2**, **Lemmas A.3** and A.4. \( \square \)

**Corollary 3.3.** Let \(X \sim ss(\theta)\) in \(\mathbb{R}^p\) with density \(f(\|x\|^2)\) as in **Theorem 3.2**. Let \(\delta_B(X) = X - b \frac{r(X)^2}{\|X\|^2} X\) be a Baranchik estimator as in **Corollary 3.1**. Then, under the loss function (2), \(\delta_B(X)\) dominates \(X\) provided \(p \geq 4\) and

\[
0 \leq b \leq 2 \frac{p + 2\ell}{p + 2(\ell + 1)} \frac{1}{E_0 \left\{ \frac{1}{\|X\|^2} \right\}}.
\]

**Proof.** The result follows from **Theorem 3.2** and **Corollary 2.2**. \( \square \)

As noted in the introduction, Propositions 1 and 2 in [7] give necessary and sufficient conditions for the spherically symmetric distribution of \(X\) in dimension \(p\) to arise as the marginal distribution of a spherically symmetric distribution in dimension \(p + k\). Our **Theorem 3.2** follows from these propositions but we have elected to give a self-contained proof.

### 4. Examples

In this section, we give two examples which illustrate the results of Sections 2 and 3 for Baranchik-type estimators (4).
**Example 1.** Let \( X \sim f(\|x - \theta\|^2) \) in \( \mathbb{R}^p \) with \( f(t) \propto \left(1 - \frac{t}{R^2}\right)^A 1_{[0,R^2]}(t) \) where \( A > -1 \) and \( R > 0 \) are fixed. When \( A = \frac{k}{2} - 1 \) with \( k \) a positive integer, by Lemma A.3, \( f(\|x - \theta\|^2) \) is the marginal density of \( X \) if \( (X, U) \sim U_{R, \theta}^{p+k} \). Hence, by Corollary 2.2, the Baranchik estimator (4) dominates \( X \) provided \( p \geq 4 \) and

\[
0 \leq b \leq 2 \frac{p - 2 + k}{p + k} \frac{1}{E_0 \left[ \frac{1}{\|X\|^2} \right]}.
\] (14)

For general \( A \), the results of Section 3 are also applicable here. Note that, for any integer \( j < A \),

\[
(-1)^j f^{(j)}(t) = A(A - 1) \ldots (A - j + 1) \left(1 - \frac{t}{R^2}\right)^{A-j} \left(\frac{1}{R^2}\right)^j 1_{[0,R^2]}(t)
\]

Therefore the conditions of Theorem 3.2 are satisfied for \( \ell = [A] \) where \( [A] \) is the greatest integer less than or equal to \( A \). Then it follows from Corollary 3.3 that the Baranchik estimator (4) dominates \( X \) provided \( p \geq 4 \) and

\[
0 \leq b \leq 2 \frac{p + 2 [A]}{p + 2 ([A] + 1)} \frac{1}{E_0 \left[ \frac{1}{\|X\|^2} \right]}.
\] (15)

Note that, if \( A = \frac{k}{2} - 1 \) for an even integer, the two results in (14) and (15) agree. However, if \( k \) is odd, (15) leads to

\[
0 \leq b \leq 2 \frac{p + k - 3}{p + k - 1} \frac{1}{E_0 \left[ \frac{1}{\|X\|^2} \right]},
\]

so that the corresponding upper bound for \( b \) in (14) is larger.

**Example 2.** Let \( X \sim f(\|x - \theta\|^2) \) in \( \mathbb{R}^p \) with \( f(t) \sim (t + A) e^{-t/2} \) where \( A > 0 \) is fixed. When \( A = k \) with \( k \) a positive integer, \( f(\|x - \theta\|^2) \) in \( \mathbb{R}^p \) is the marginal density of \( X \) if

\[
(X, U) \sim \frac{1}{(p + k)(2\pi)^{(p+k)/2}} (\|x - \theta\|^2 + \|u\|^2) \exp \left(-\frac{1}{2}(\|x - \theta\|^2 + \|u\|^2)\right)
\]

where the dimension of \( U \) is \( k \), i.e. the Kotz’s distribution with parameter 1 (see [11]). Hence, by Corollary 2.2, the bound in (14) guarantees domination over \( X \) of the Baranchik estimator (4) when \( p \geq 4 \).

For general \( A \), note that, for any non-negative integer \( j \),

\[
(-1)^j f^{(j)}(t) = \left(-\frac{1}{2}\right)^j (A - 2j + t) e^{-t/2}.
\]

Hence, if \( j \leq \left[\frac{A}{2}\right] \), then \( (-1)^j f^{(j)}(t) \geq 0 \). Therefore Theorem 3.2 holds with \( \ell = \left[\frac{A}{2}\right] - 1 \) and the upper bound on \( b \) for domination of the Baranchik estimator over \( X \) is

\[
2 \frac{p + 2 \left[\frac{A}{2}\right] - 2}{p + 2 \left[\frac{A}{2}\right]} \frac{1}{E_0 \left[ \frac{1}{\|X\|^2} \right]}.
\]

Again, if \( A = k \) is an even integer, the upper bound agrees with the upper bound (14). However, if \( k \) is an odd integer, the bound in (14) is larger.
Remark. Propositions 1 and 2 of [7], when applied to these two examples in place of Theorem 3.2, give a slightly better bound in each. In particular, in Example 1, the bound becomes

\[ 0 < b \leq 2 \frac{p - 2 + [2(A + 1)]}{p + [2(A + 1)]} E_0 \left[ \frac{1}{\|X\|^2} \right] \]

and, in Example 2, it becomes

\[ 0 < b \leq 2 \frac{p - 2 + [A]}{p + [A]} E_0 \left[ \frac{1}{\|X\|^2} \right] \]

These bounds agree with (14) whenever the value of \( A \) in either example (\( A = \frac{k}{2} - 1 \), and \( A = k \) respectively) corresponds to any integer \( k \).

5. Concluding remarks

This paper has developed a sequence of increasing minimax shrinkage bounds for Stein-type estimators for a corresponding decreasing sequence of classes of spherically symmetric distributions. The classes of distributions in \( \mathbb{R}^p \) may be viewed as (for \( k \geq 1 \)) those spherically symmetric distributions with densities which arise as the projection from a spherically symmetric distribution on \( \mathbb{R}^{p+k} \). We have also shown that, if the generating function \( f \) satisfies \((-1)^j f^{(j)}(t) \geq 0 \) for \( 0 \leq j \leq \ell \) and \( f^{(\ell)}(t) \) is non-increasing, then the density \( f(\|x - \theta\|^2) \) arises as a projection from a spherically symmetric distribution on \( \mathbb{R}^{p+k} \) where \( \ell = 2(k + 1) \).

The sequence of bounds on \( b \), for domination of the James–Stein estimator (6) and the Baranchik-type estimator (4) over \( X \) is given by

\[ 0 < b \leq 2 \frac{p - 2 + k}{p + k} E_0 \left[ \frac{1}{\|X\|^2} \right] = 2 \frac{p + 2\ell}{p + 2(\ell + 1)} E_0 \left[ \frac{1}{\|X\|^2} \right] . \]

These bounds coincide with classical bounds for the case \( k = 0 \) (spherically symmetric distribution), \( k = 2 \) (spherically symmetric and unimodal distribution), and \( k = \infty \) (scale mixtures of normals). In this way the results of this paper unify several known classical bounds and extend them to adapt to different degrees of monotonicity of \( f(t) \).

Appendix

In this appendix, for \( \theta \in \mathbb{R}^p \) and \( R \geq 0 \), \( U_{R,\theta}^p \) denotes the uniform distribution on the sphere \( S_{R,\theta}^p \) in \( \mathbb{R}^p \) of radius \( R \) centered at \( \theta \). For notational convenience, we will occasionally denote \( U_{R,(\theta,0)}^{p+k} \) by \( U_{R,\theta}^{p+k} \) and \( S_{R,(\theta,0)}^{p+k} \) by \( S_{R,\theta}^{p+k} \). Similarly, \( B_{R,\theta}^p \) denotes the ball in \( \mathbb{R}^p \) of radius \( R \) centered at \( \theta \) and \( \mathcal{V}_{R,\theta}^p \) denotes the uniform distribution on this ball.

Lemma A.1. The function

\[ f_\theta : R \mapsto \mathbb{R}^2 \int_{S_{R,\theta}^{p+k}} \frac{1}{\|x\|^2} \; dU_{R,\theta}^{p+k}(x, u) \]

is non-decreasing for \( p \geq 4 \) and \( k \geq 0 \).
**Proof.** Note that, by invariance, $f_\theta$ depends on $\theta$ only through $\|\theta\|$. With the change of variable $(y, v) = \left( \frac{x - \theta}{R}, \frac{u}{R} \right)$, we have

$$f_\theta(R) = \int_{S^p_{1,0}} \frac{1}{\|y + \frac{\theta}{R}\|^2} \, dU^{p+k}_{1,0}(y, v).$$

Hence, for any $\theta \in \mathbb{R}^p$ such that $\|\theta\| = R_0$,

$$f_\theta(R) = \int_{S^p_{R_0,0}} \int_{S^{p+k}_{1,0}} \frac{1}{\|y + \frac{\theta}{R}\|^2} \, dU^{p+k}_{1,0}(y, v) \, dU^p_{R_0,0}(\theta)$$

$$= \int_{S^{p+k}_{1,0}} \int_{S^p_{R_0,0}} \frac{1}{\|y + \frac{\theta}{R}\|^2} \, dU^p_{R_0,0}(\theta) \, dU^{p+k}_{1,0}(y, v)$$

by Fubini’s theorem. In the inner integral, the change of variable $z = \frac{\theta}{R} + y$ leads to

$$f_\theta(R) = \int_{S^{p+k}_{1,0}} \int_{S^p_{R_0/R,y}} \frac{1}{\|z\|^2} \, dU^p_{R_0/R,y}(z) \, dU^{p+k}_{1,0}(y, v).$$

As the function $\frac{1}{\|z\|^2}$ is superharmonic for $p \geq 4$, the inner integral is non-increasing in $R_0/R$ for each $y$, and hence non-decreasing in $R$. □

**Lemma A.2.** Let $r(t)$ be a non-negative and non-decreasing function on $[0, \infty[$ such that $\frac{r(t)}{t}$ is non-increasing. Then, for any fixed $\theta \in \mathbb{R}^p$, the function

$$f_\theta : R \mapsto R^2 \int_{S^{p+k}_{R,0}} \frac{r(\|x\|^2)}{\|x\|^2} \, dU^{p+k}_{R,\theta}(x, u)$$

is non-decreasing for $p \geq 1$ and $k \geq 2$.

**Proof.** As noted in Lemma A.3, the marginal distribution of $X$ is absolutely continuous with unimodal density $\frac{1}{R^p} \psi \left( \frac{\|x - \theta\|^2}{R^2} \right)$ for all $k \geq 2$. Then $f_\theta$ can be written as

$$f_\theta(R) = \int_{B^p_{1,0}} \frac{r(R^2 \|z + \frac{\theta}{R}\|^2)}{\|z + \frac{\theta}{R}\|^2} \psi(\|z\|^2) \, dz.$$

For any $R_1 \leq R_2$, we have by monotonicity of $r(t)$

$$f_\theta(R_1) \leq \int_{B^p_{1,0}} \frac{r(R_2^2 \|z + \frac{\theta}{R_2}\|^2)}{\|z + \frac{\theta}{R_2}\|^2} \psi(\|z\|^2) \, dz.$$

Furthermore monotonicity of $\frac{r(t)}{t}$ implies that $\frac{r(R_2^2 \|z + \frac{\theta}{R_1}\|^2)}{\|z + \frac{\theta}{R_1}\|^2}$ is symmetric and unimodal in $z$ about $-\frac{\theta}{R_1}$. Hence, by Anderson’s theorem (see [1]),

$$\int_{B^p_{1,0}} \frac{r(R_2^2 \|z + \frac{\theta}{R_2}\|^2)}{\|z + \frac{\theta}{R_2}\|^2} \psi(\|z\|^2) \, dz \leq \int_{B^p_{1,0}} \frac{r(R_2^2 \|z + \frac{\theta}{R_2}\|^2)}{\|z + \frac{\theta}{R_2}\|^2} \psi(\|z\|^2) \, dz$$

$$= f_\theta(R_2). \quad \square$$
Lemma A.3 (See e.g., [6]). Let \((X, U) \sim U_{\mathbb{R}^p, \theta}^{p+k}\) with \(k \geq 1\). Then \(X\) is absolutely continuous in \(\mathbb{R}^p\) with density proportional to \(\frac{1}{|\mathbb{R}^p|} \psi(\frac{\|x-\theta\|^2}{R^2})\) where \(\psi(t) = (1-t)^{k/2-1}I_{[0,1]}(t)\). In particular, when \(k = 2\), the distribution of \(X\) is uniform on the ball \(B_{\mathbb{R}^p, \theta}^R\).

The next lemma deals with transformations of distributions which are mixtures of distributions. It is well known. We include it for completeness and to establish notation.

Lemma A.4. Let \(P\) be a distribution on \(X\) and \(\phi\) a transformation from \(X\) to a space \(Y\). Assume that \(P\) is defined as a mixture of a family of distributions \((P_\lambda)_{\lambda \in \Lambda}\) on \(X\) with mixing distribution \(G\),

\[
P(A) = \int_{\Lambda} P_\lambda(A) \, dG(\lambda).
\]

Then the distribution of \(\phi\) under \(P\) defined by

\[
P_\phi(B) = P\left(\phi^{-1}(B)\right)
\]

satisfies

\[
P_\phi(B) = \int_{\Lambda} (P_\lambda)_\phi(B) \, dG(\lambda).
\]

Proof. The result is immediate and follows from

\[
(P_\lambda)_\phi(B) = P_\lambda\left(\phi^{-1}(B)\right).
\]

References