# On the estimation of a restricted location parameter for symmetric distributions

Éric Marchand[a, 1, 2] , Idir Ouassou[b], Amir T. Payandeh[c], François Perron[d,2]

*a Département de mathématiques, Université de Sherbrooke, Sherbrooke, Qc, CANADA, J1K 2R1*

*b ÉNSA, Avenue Abdelkrim El Khattabi BP 575, Marrakech, Maroc*

*c Departement of Mathematics and Statistics, University of New Brunswick, Fredericton, N.B. CANADA E3B 5A3*

*d Département de mathématiques et de statistique, Université de Montréal, C.P. 6128, Succursale Centre-Ville, Montréal, Qc, CANADA, H3C 3J7*

## Summary

For estimating the median $\theta$ of a spherically symmetric univariate distribution under squared error loss, when $\theta$ is known to be restricted to an interval $[-m,m]$; $m$ known; we derive sufficient conditions for estimators $\delta$ to dominate the maximum likelihood estimator $\delta_{\text{mle}}$. Namely: (i) we identify a large class of models where for sufficiently small $m$, **all** Bayesian estimators with respect to symmetric about 0 priors supported on $[-m,m]$ dominate $\delta_{\text{mle}}$, and (ii) we provide sufficient dominance conditions of the form $m \leq c_\pi$ which are applicable to various models and priors $\pi$. In terms of the models, applications include Cauchy and Student distributions, as well as logconcave densities with logconvex derivatives such as normal, logistic, double-exponential among others. In terms of priors $\pi$, applications include the uniform prior as well as priors which are symmetric about 0, continuous, and nondecreasing and logconcave on $(0,m)$.

*AMS 2000 subject classifications:* 62F10, 62F15, 62F30 *Keywords and phrases*: Maximum likelihood estimator, restricted parameter space, Bayes estimator, squared error loss, dominance, symmetric location families.

---

1. research supported by NSERC of Canada

2. Corresponding author; e-mail address: eric.marchand@usherbrooke.ca

# 1. Introduction

Consider the problem of estimating under squared-error loss, based on an observable $X$, the median $\theta$ of a spherically symmetric univariate model where $\theta$ is known to be restricted to an interval $[a,b]$. Without loss of generality, we set $a = -m$, $b = m$, with $m > 0$. Interesting questions pertain to the frequentist performance of Bayesian estimators, such as the determination or description of Bayesian estimators that improve upon the benchmark (but inadmissible; e.g., Charras and van Eeden, 1993) maximum likelihood estimator (mle). With respect to the last question, findings have been obtained, for normal models and among others, by Casella and Strawderman (1981); Gatsonis, MacGibbon and Strawderman (1987); and Marchand and Perron (2001); and for more general models by Moors (1981, 1985) and Marchand and Perron (2007).

The results of Marchand and Perron (2001), which are developed for multivariate versions with $X \sim N_p(\theta,\sigma^2 I_p)$, $\|\theta\| \leq m$, and known $\sigma^2$, will be of particular relevance to us. Namely, their results for $p = 1$ imply that:

(i) all Bayes estimators with respect to a symmetric about 0 prior dominates the mle as soon as $m \leq c_0\sigma$ with $c_0 \approx 0{,}4837$;

(ii) the Bayes estimator $\delta_{BU}$ with respect to a boundary uniform prior on $\{-m,m\}$ dominates the mle as soon as $m \leq \sigma$;

(iii) the Bayes estimator $\delta_U$ with respect to a fully uniform on $[-m,m]$ prior dominates the mle as soon as $m \leq c_1\sigma$ with $c_1 \approx 0{,}5230$.

The technical arguments relied upon to arrive at these findings are conditional risk decompositions of the type $|X| \leq m$, $|X| > m$, and $|X| = r$, $r > 0$. Perron (2003) exploited these types of decompositions to obtain analogous dominance results for a estimating Binomial proportion $p$ when $|p-\frac{1}{2}|$ is constrained above. The main objectives of this paper are to make it apparent that these conditional risk techniques are not specific to a normal model, and to obtain explicit dominance results (e.g., Corollary 3) for symmetric location models analogous to (i) and (iii). With respect to (ii), Marchand and Perron (2007) have already obtained extensions for a large class of symmetric models.

Sections 3 and 4 contain various dominance results and illustrations. More specifically, Section 3 gives for various models, conditions on $(m,\delta)$ for which symmetric about 0 estimators $\delta$ dominate the mle. Section 4 makes connections between priors $\pi$ and associated $\delta_\pi$'s arriving at dominance conditions on $\pi$ for Bayesian estimators $\delta_\pi$. For several applications, such as Corollary 3, the results and the technical arguments leading to these beginning in Section 2, apply to logconcave densities with log convex derivatives (examples include Normal, Laplace, Logistic, Hyperbolic secant). However, not all the dominance results require such assumptions, as illustrated by the univariate Cauchy and Student distributions (Example 2).

## 2. Definitions and technical lemmas

Our work here relates to the estimation of the parameter $\theta$, under squared error loss $(d-\theta)^2$, for symmetric location families with densities $f(\cdot|\theta)$ on $\mathbb{R}$ (with respect to Lebesgue measure) of the form:

$$f(x|\,\theta) = e^{-h(x-\theta)}, \theta \in \Theta(m) = [-m,m], \,(m\,\text{fixed}). \tag{1}$$

The corresponding cdf's will be denoted $F_{\theta,h}$. In (1), $h$ is a given continuous and even function. As our model and loss being invariant under sign changes, we will consider equivariant estimators only (i.e., odd functions); and corresponding risk functions will be symmetric functions on $\Theta(m)$. We will be concerned with the specification of such estimators that dominate $\delta_0$, with $\delta_0(x) = (x \wedge m)$ for $x > 0$. For unimodal models in (1) or cases where $h$ is increasing, $\delta_0$ coincides with the maximum likelihood estimator of $\theta$ (in such cases, we will refer to this target estimator as $\delta_{\text{mle}}$; otherwise, we will write $\delta_0$).

Our dominance results are based on the conditional risk decompositions: **(i)** $|X| > m$, **(ii)** $|X| \leq m$, and **(iii)** $|X| = r$; and make use of the following preliminary results and definitions concerning various conditional expectations and how their behaviour depends on $h$.

**Definition 1.** *(a) For $m > 0$, set*

$$\begin{aligned} H^{**}(m) &= \{h : h \textit{ increasing and convex on } (0,2m)\}; \\ H^{*}(m) &= \{h : h \in H^{**}(m) \textit{ and } h' \textit{ is concave on } (0,2m)\}; \\ H^{***}(m) &= \{h :\ h \textit{ increasing and } h(\theta+x) - h(\theta-x) \textit{ increasing in } \theta \textit{ on } (x,m), \textit{ for all } x \in (0,m)\}. \end{aligned}$$

*(b) For fixed increasing $h$, set $b_h^*$, $b_h^{**}$, $b_h^{***}$ as*

$$\begin{aligned} b_h^* &= \sup\{m \geq 0 :\ H^*(m) \ni h\}; \\ b_h^{**} &= \sup\{m \geq 0 :\ H^{**}(m) \ni h\}; \\ b_h^{***} &= \sup\{m \geq 0 :\ H^{***}(m) \ni h\}. \end{aligned}$$

(note: if $h$ is increasing and convex on $(0,\infty)$, we will say that $h \in H^{**}(\infty)$ and $b_h^{**} = \infty$. We adopt similar conventions for $H^*$, $H^{***}$, $b_h^*$, and $b_h^{***}$.)

**Remark 1.** *It is immediate with the above definitions that:* **(i)** *for all $m > 0$, $H^*(m) \subseteq H^{**}(m) \subseteq H^{***}(m)$; and* **(ii)** *for all fixed increasing $h$, $b_h^* \leq b_h^{**} \leq b_h^{***}$. From Marchand and Perron (2007), we also have that $H^{***}(m) \subseteq H^{**}(\frac{m}{2})$ for all $m > 0$.*

We also require the following quantities, defined for any $h$.

$$\begin{aligned}
\rho_h(\theta,x) &= \tanh(\frac{h(\theta+x)-h(\theta-x)}{2}),\ \theta,x \in \mathbb{R};\\
\bar{\rho_h}(m,x) &= \sup\{\rho_h(\theta,x)\ :\ \theta \in [0,m]\}\ ,m,x > 0;\\
\alpha_h(m,\theta) &= E_\theta[\rho_h(\theta,|X|)||X| > m];\ 0 \leq \theta \leq m;\\
\bar{\alpha_h}(m) &= \sup\{\alpha_h(m,\theta)\,;\ \theta \in [0,m]\},\ m > 0;\\
\beta_h(m,\theta) &= E_\theta[\frac{\theta\rho_h(\theta,X)}{X}||X| \leq m],\ 0 \leq \theta \leq m;\\
\bar{\beta_h}(m) &= \sup\{\beta_h(m,\theta)\,;\ \theta \in [0,m]\},\ m > 0;\\
G_h(m) &= mh'(m),\ m > 0;\ \text{for } h \text{ differentiable at } m.
\end{aligned}$$

Observe that for fixed $h$ such that $b_h^{**} > 0$, $G_h$ necessarily increases on $(0,b_h^{**})$. In such cases, the quantities $G_h^{-1}(a) = \inf\{m : mh'(m) > a\}$ will play a key role (e.g., Corollary 2, Remark 3, Table 1).

Here are some useful properties concerning the quantities above.

**Lemma 1.** *For a fixed nondecreasing $h$, we have*

*(a) $\rho_h(\cdot\,,x)$ is nondecreasing on $[0,m]$ for all $x > 0$ if and only if $h \in H^{***}(m)$;*

*(b) $\rho_h(\theta\,,\cdot)$ is nondecreasing on $(0,\infty)$ for all $\theta \in (0,m)$ if and only if $h \in H^{**}(\infty)$;*

*(c) $E_\theta[sgn(X)||X| = r] = \rho_h(\theta\,,r);\ r \geq 0\,;$*

*(d) for any fixed $\theta \geq 0$, $\lim_{x\to 0^+} \frac{\rho_h(\theta\,,x)}{x} = h'(\theta)$; ($h$ differentiable at $\theta$);*

*(e) for any fixed $\theta \in [0,m]$, $\frac{\rho_h(\theta\,,x)}{x}$ is decreasing in $x$ on $(0,m)$ whenever $h \in H^*(m)$;*

*(f) whenever $h \in H^*(m)$, we have $\bar{\rho}_h(m,x) \le x h'(m)$ for all $x \in (0,m)$.*

**Proof.** Parts (a,b,c) are given by Marchand and Perron (2007, lemmas 1 and 2), while part (d) follows by a direct evaluation. Part (f) is a direct consequence of (d,e). In order to establish the result in (e), it will suffice to show that, for any fixed $\theta \in [0,m]$, the function $\rho_h(\theta\,, x)$ is concave in $x$ on $(0,m)$, given that $\rho_h(\theta,0) = 0$ for all $\theta$. Since the hyperbolic tangent function is concave on $[0,\infty)$, and that a composition $\tanh \circ \varphi$ is necessarily concave for nondecreasing and concave $\varphi$, it will suffice to show that $\varphi(x) = h(\theta+x) - h(\theta-x)$ is nondecreasing and concave in $x$ on $(0,m)$ and takes its values on $[0,\infty)$. Since $h$ is nondecreasing on $(0,m)$, even and convex on $(-2m,2m)$, we have $\varphi(0) = 0$ and $\varphi$ nondecreasing on $(0,m)$. Also, we have for $(x_0,x_1)$ such that $0 < x_0 < x_1 < m$:

$$(i)\ \varphi'(x_1) - \varphi'(x_0) = [\{h'(\theta + x_1) - h'(\theta + x_0)\} - \{h'(\theta - x_0) - h'(\theta - x_1)\}] \le 0 \text{ if } x_1 \le \theta$$

and

$$(ii)\ \varphi'(x_1) - \varphi'(x_0) = [\{h'(x_1 + \theta) - h'(x_1 - \theta)\} - \{h'(x_0 + \theta) - h'(x_0 - \theta)\}] \le 0 \text{ if } \theta \le x_0,$$

because $h'$ is concave on $(0,2m)$. Therefore, $\varphi$ is concave on both the intervals $(0,\theta)$ and $(\theta,m)$. Finally, $\varphi$ is continuous at $\theta$ and $\varphi'(\theta-) = h'(2\theta) - h'(0-) \ge h'(2\theta) - h'(0+) = \varphi'(\theta+)$, implying that $\varphi$ is concave on $(0,m)$.

**Lemma 2.** *For any fixed $h$,*

*(a) the function $\alpha_h(m,\cdot)$ increases on $[0,m]$ with $\bar{\alpha_h}(m) = \alpha_h(m,m) = \frac{1-2F_{0,h}(-2m)}{1+2F_{0,h}(-2m)}$, $\lim_{m\to 0^+} \bar{\alpha_h}(m) = 0$, $\lim_{m\to\infty} \bar{\alpha_h}(m) = 1$, and $\bar{\alpha_h}^{-1}(1/2) = \frac{1}{2}F_{0,h}^{-1}(\frac{5}{6})$;*

*(b) $\bar{\beta}_h(m) \le G_h(m)$ for any fixed $(m,h)$ such that $h \in H^*(m)$.*

**Proof.** For part (a), a direct evaluation gives us for $\theta \in [0,m]$:

$$\begin{aligned}\alpha_h(m,\theta) &= E_\theta[\tanh(\frac{h(\theta+|X|)-h(\theta-|X|)}{2})||X|>m] \\ &= \frac{\int_m^\infty (e^{-h(\theta-x)} - e^{-h(\theta+x)})dx}{\int_m^\infty (e^{-h(\theta-x)} + e^{-h(\theta+x)})dx} \\ &= \frac{F_{0,h}(m+\theta) - F_{0,h}(m-\theta)}{2 - F_{0,h}(m+\theta) - F_{0,h}(m-\theta)} \\ &= 1 - 2\frac{1}{1+\frac{1-F_{0,h}(m-\theta)}{1-F_{0,h}(m+\theta)}}.\end{aligned}$$

From the above, the increasing property of $\alpha_h(m,\cdot)$ follows at once, with $\bar{\alpha_h}(m) = \alpha_h(m,m)$ as given. The other given properties of (a) follow in a straightforward manner.

For (b), we have indeed from parts (d,e) of Lemma 1, and for all $\theta \in [0,m]$: $\beta_h(m,\theta) \leq \theta h'(\theta) \leq mh'(m) = G_h(m)$, since $h \in H^*(m)$.

**Lemma 3.** *For $h \in H^*(\infty)$, the density of $R = |X|$ has increasing monotone likelihood ratio in $R$, with $\lambda = |\theta|$ taken as the parameter.*

**Proof.** It is sufficient (and necessary) to show that the mixed derivative $\frac{\partial^2}{\partial\lambda\partial r}\log g_h(\lambda,r) \geq 0$ for all $\lambda,r > 0$; $g_h(\lambda,\cdot)$ being the density of $R$ under (1). Now set $z_1 = \frac{1}{2}h(\lambda+r)$, $z_2 = \frac{1}{2}h(\lambda-r)$. Under (1), we have directly $g_h(\lambda,r) = \frac{1}{2}(e^{-2z_1} + e^{-2z_2})$, which we can also write as

$$g_h(\lambda,r) = \cosh(z_1 - z_2)e^{-(z_1+z_2)}.$$

Working with this last representation, it is easy to establish that:

$$\frac{\partial^2}{\partial\lambda\partial r}\log g_h(\lambda,r) = (1-\tanh^2(z_1-z_2))\frac{\partial}{\partial\lambda}(z_1-z_2)\frac{\partial}{\partial r}(z_1-z_2)+\tanh(z_1-z_2)\frac{\partial^2}{\partial\lambda\partial r}(z_1-z_2)-\frac{\partial^2}{\partial\lambda\partial r}(z_1+z_2).$$

The result follows as

(i) $|\tanh| \leq 1$ and $\frac{\partial}{\partial\lambda}(z_1-z_2)\frac{\partial}{\partial r}(z_1-z_2) = \frac{1}{2}\{h'(\lambda+r)^2 - h'(\lambda-r)^2\} \geq 0$, given that $h'$ is increasing and positive;

(ii) $\tanh(z_1 - z_2) \geq 0$ with $\tanh(\cdot)$ increasing, $z_1 \geq z_2$ because $h$ is increasing on $(0,\infty)$ and even on $(-\infty,\infty)$, and $\frac{\partial^2}{\partial\lambda\partial r}(z_1 - z_2) = \frac{1}{2}\{h''(\lambda + r) + h''(\lambda - r)\} \geq 0$ given that $h$ is convex;

(iii) and, finally, $\frac{\partial^2}{\partial\lambda\partial r}(z_1 + z_2) = \frac{1}{2}\{h''(\lambda + r) - h''(\lambda - r)\} \leq 0$ given that $h$ is even and $h'$ is concave on $(0,\infty)$.

# 3. General dominance results

Here, for an equivariant estimator $\delta$, we provide conditions on $(h,m,\delta)$ for which $\delta$ dominates $\delta_0$ on $\Theta(m)$, with risk analyzes for $\theta \in [0,m]$ sufficing given the symmetry of the risk functions involved. As in Marchand and Perron (2001), we proceed by studying improvements with respect to the conditional risks for: **(i)** $|X| > m$, **(ii)** $|X| \leq m$, and **(iii)** $|X| = r$, $r > 0$. Notwithstanding the fact that a large portion of the development below does indeed represent an extension of the univariate normal case analyzed by Marchand and Perron (2001) to general models in (1), the technical arguments used match those given by Marchand and Perron (2001), with results of Section 2 called upon for various bounds on monotonicity properties. Hence, we have relegated some proofs to an Appendix.

**Theorem 1.** *For a model in (1) with $h \in H^{**}(\infty)$, and $\delta$ a nondecreasing function on $(m,\infty)$, the condition*

$$m(2\bar{\alpha}_h(m) - 1) \leq \delta(x) \leq m, \text{ for all } x > m, \tag{2}$$

*with strict inequalities for some $x \in (m,\infty)$, is sufficient to have*

$$E_\theta\left[|\theta - \delta_{mle}(X)|^2 - |\theta - \delta(X)|^2 \big| |X| > m\right] > 0\,, \qquad for \quad all \quad \theta \in \Theta(m).$$

**Proof.** See Appendix.

**Remark 2.** *Lemma 2 tells us that the left-hand side of (2) can also be expressed explicitly (or written as we will do frequently hereafter) as* $m\,(\frac{1-6F_{0,h}(-2m)}{1+2F_{0,h}(-2m)})$.

**Theorem 2.** *For a model in (1) with* $(m,h)$ *such that* $h \in H^*(m)$ *(i.e.,* $m \leq b_h^*$*) and* $\bar{\beta}_h(m) \leq 1$*, and* $g$ *a nondecreasing function on* $(0,m)$ *such that* $x(x-\delta(x))$ *increases in* $x$ *for* $x \in (0,m)$*, the condition*

$$(2\bar{\beta}_h(m)-1)x \leq \delta(x) \leq x, \textit{ for all } x \in (0,m),$$

*with strict inequalities for some* $x \in (0,m)$*, is sufficient to have*

$$E_\theta\left[|\theta - \delta_{mle}(X)|^2 - |\theta - \delta(X)|^2 \big| |X| \leq m\right] > 0 \qquad for \quad all \quad \theta \in \Theta(m).$$

**Proof.** See Appendix.

Combining the dominance conditions of Theorem 1 and Theorem 2 leads directly to the following.

**Corollary 1.** *For a model in (1) with* $(m,h)$ *such that* $h \in H^*(m) \cap H^{**}(\infty)$ *and* $\bar{\beta}_h(m) \leq 1$*, and* $\delta$ *a nondecreasing function on* $(0,\infty)$ *such that* **(i)** $\delta(x) < x \wedge m$ *for* $x > 0$*,* **(ii)** $x(x-\delta(x))$ *increases in* $x$ *on* $(0,m)$*, the conditions*

$$\frac{\delta(m)}{m} > \frac{1-6F_{0,h}(-2m)}{1+2F_{0,h}(-2m)}, \textit{ and } \frac{\delta(x)}{x} > 2\bar{\beta}_h(m) - 1 \textit{ for all } x \in (0,m) \tag{3}$$

*are jointly sufficient for* $\delta$ *to dominate* $\delta_{mle}$ *on* $\Theta(m)$*.*

**Remark 3.** *For densities in (1) that are logconcave with logconvex derivatives (i.e.,* $h \in H^*(\infty)$*), the conditions on* $\delta$ *of Corollary 1 yield many dominating estimators as long as* $\bar{\beta}_h(m) \leq 1$*, which in turn holds whenever* $m \leq G_h^{-1}(1)$ *(see Lemma 2). Furthermore, a wide class of solutions consist in selecting* $\delta$*'s such that* $\frac{\delta(x)}{x}$ *decreases for* $x \in (0,m)$ *(which together with (i) implies (ii)), with* $\frac{\delta(m)}{m} > (2\bar{\beta}_h(m)-1) \vee (\frac{1-6F_{0,h}(-2m)}{1+2F_{0,h}(-2m)})$.

**Example 1.** *(Truncated Linear and truncated linear minimax estimators) Interesting and simple alternatives to* $\delta_{mle}$ *are truncated linear estimators of the form* $\delta_a(x) = (ax \wedge m)sgn(x)$*, for* $a \in (0,1)$*. A particulary appealing choice, available for cases where the variance in (1) exists (say equal to* $V$*), is given by the choice* $a_0 = \frac{m^2}{m^2+V}$ *corresponding to the truncation onto* $[-m,m]$ *of the linear minimax estimator for* $\theta \in \Theta(m)$*. Although* $a_0 X$ *dominates* $X$ *on* $\Theta(m)$*, there is no guarantee that dominance will carry over to the truncated versions.*

*For the estimator* $\delta_a$ *with* $a \in (0,1)$*, it is easy to see that conditions (i) and (ii) of Corollary 1 are satisfied. Hence for models in (1), Corollary 1 (or Remark 3) tells us, for* $(m,h)$ *satisfying its conditions, that* $\delta_a$ *dominates* $\delta_{mle}$ *whenever:*

$$(2\bar{\beta}_h(m) - 1) \vee (\frac{1 - 6F_{0,h}(-2m)}{1 + 2F_{0,h}(-2m)}) \leq a < 1. \tag{4}$$

*For the truncated linear minimax estimator* $\delta_{a_0}$*, the evaluation above is more delicate since* $a_0$ *depends on* $m$*. The following is specific to the normal case, but similar analyzes can be arrived at for other logconcave densities.*

**Theorem 3.** *For a normal* $N(\theta,\sigma^2)$ *model with* $|\theta| \leq m$*,* $\delta_{a_0}$ *dominates* $\delta_{mle}$ *whenever* $m \leq c_1 \approx 0.60936\sigma$*, with* $c_1$ *being the unique positive solution in* $c$ *to the equation* $\psi(c) = 0$*, with* $\psi(c) = \frac{c^2}{c^2+1} - (\frac{1-6\Phi(-2c)}{1+2\Phi(-2c)})$*;* $\Phi$ *being the standard normal cdf.*

**Proof.** *Setting* $\sigma = 1$ *without loss of generality and applying (4) with* $a = \frac{m^2}{m^2+1}$*, we seek to determine the values of* $m$ *such that: (1)* $\frac{m^2}{m^2+1} \geq 2\bar{\beta}_h(m) - 1$*, and (2)* $\psi(m) \geq 0$*. For (1), make use of part (b) of Lemma 2 with* $h(y) = (y^2)/2$ *to infer that (1) holds as long as* $\frac{m^2}{m^2+1} \geq 2m^2 - 1$*; i.e.,* $m \leq (\frac{1}{2})^{1/4}$*. With regards to (2), observe that* $\psi(0) > 0$*, and that*

$$\psi(m) \geq 0 \Longleftrightarrow T(m) \geq \frac{1}{2}, \text{ with } T(m) = (4m^2 + 3)\Phi(-2m).$$

*A direct evaluation yields* $T'(m) = 2\Phi(-2m)(4m - (4m^2+3)\frac{\phi(2m)}{\Phi(-2m)})$; $\phi$ *being the standard normal pdf. From this and by virtue of the inequality* $\frac{\phi(x)}{\Phi(-x)} \geq x$ $(x \geq 0)$, *we infer that* $T(m)$ *decreases in* $m$ *on* $[0,\infty)$. *Since* $\lim_{m\to\infty} T(m) = 0$, *we conclude that* $\psi(m)$ *changes signs once from* $+$ *to* $-$ *as* $m$ *ranges on* $(0,\infty)$. *Finally, a direct evaluation tells us that* $\psi((\frac{1}{2})^{1/4}) \leq 0$, *yielding the result.*

**Theorem 4.** *(Marchand & Perron, 2005) For a model in (1), the (symmetric) estimator* $\delta$ *dominates* $\delta_0$ *on* $\Theta(m)$ *whenever, for all* $x > 0$,

$$2\sup\{\theta\rho_h(\theta,x) : \theta \in [0,m]\} - (m \wedge x) < \delta(x) < (m \wedge x), \tag{5}$$

*for all* $x \in A_{h,m} = \{x > 0 : \sup\{\theta\rho_h(\theta,x) : \theta \in [0,m]\} < x\}$, *and* $\delta(x) = x$ *otherwise.*

The above condition is also used in Marchand and Perron (2007) to determine conditions for which the Bayes estimator with respect to a uniform 2-point prior on $\{-m,m\}$, given by $\delta_{BU}(x) = m\rho_h(m,x)$ for $x > 0$, dominates $\delta_0$ on $\Theta(m)$. Two key observations lead to their Theorem 2 result. First, if $(m,h)$ are such that $h \in H^{***}(m)$, then we have $\bar{\rho}_h(m,x) = \rho_h(m,x)$ for all $x > 0$ (part (a) of Lemma 1). Secondly, if $(m,h)$ are such that $h \in H^*(m)$ and $m \leq G_h^{-1}(1)$ (see Table 1), then (5) is satisfied for $\delta = \delta_{BU}$ as seen as a consequence of part (f) of Lemma 1 (they actually give somewhat weaker conditions along with illustrations). We will not use Theorem 4 in the same way. Instead, we will, as in the next result and the examples that follow, exploit conditions for which the left-hand side of (5) is bounded above, namely by 0, which leaves a large window in (5) for $\delta$ to vary and for $\delta$ to dominate $\delta_0$.

**Corollary 2.** *For a model in (1), an equivariant estimator* $\delta$ *of* $\theta$, *such that* $\delta \neq \delta_0$ *and* $0 \leq \delta(x) \leq (m \wedge x)$ *for all* $x > 0$, *dominates* $\delta_0$ *when either:*

*(i)* $2m\bar{\rho}_h(m,x) \leq (m \wedge x)$ *for all* $x > 0$,

*(ii) or $\delta$ is nondecreasing on $(m,\infty)$, and $(m,h)$ are such that $h \in H^*(m)$ with $m \leq G_h^{-1}(\frac{1}{2}) \wedge \frac{1}{2}F_{0,h}^{-1}(\frac{5}{6})$.*

**Proof.** The sufficiency of (i) is a direct consequence of (5). For (ii), we verify (5) for all $x \in (0,m)$ and (2) for all $x > m$. For $x \in (0,m)$, apply part (f) of Lemma 1 to infer that $2m\bar{\rho}_h(m,x) \leq 2mh'(m)x \leq x$ as soon as $m \leq G_h^{-1}(1/2)$. For $x > m$, it is straightforward to see that the mere nonnegativity of $g$ implies that (2) is satisfied whenever $\bar{\alpha}_h(m) \leq \frac{1}{2}$, that is $m \leq \frac{1}{2}F_{0,h}^{-1}(\frac{5}{6})$ as seen in Lemma 2.

**Remark 4.** *For densities in (1) that are logconcave with logconvex derivatives (i.e., $h \in H^*(\infty)$), condition (ii) of Corollary 2 gives, for shrinkage estimators $\delta$ which increase on $[m,\infty)$, universal dominance with no further conditions on $g$, but requiring only small enough $m$. Furthermore, it will be established below (i.e., Lemma 4, Corollary 3) that the conditions on $g$ are shared by all Bayes estimators with respect to symmetric about 0 priors on $\Theta(m)$. This universal dominance result, which illustrates vividly the inadequacy of $\delta_{mle}$ for small parameter spaces (under squared error loss) extends the univariate normal result of Marchand and Perron (2001) (they also give a multivariate normal version). Observe that the estimator $\delta(X) = 0$, which has been further studied by Dou and van Eeden (2006) in the normal case, belongs to the class of dominating estimators. A selection of further examples of models for which $h \in H^*(\infty)$; which include members of the Exponential power family such as Laplace (or Double-Exponential) distributions, Hyperbolic Secant, and Logistic distributions (see Marchand and Perron, 2007, for more details and examples); along with the corresponding upper bounds on $m$, are given in Table 1.*

**Example 2.** *(Student and Cauchy distributions) Consider families of distributions in (1) with* *densities*

$$f(x|\,\theta) = \frac{1}{B(\nu/2,1/2)\sigma\sqrt{\nu}}\{1 + \frac{1}{\nu}(\frac{x-\theta}{\sigma})^2\}^{-\frac{(\nu+1)}{2}}, \tag{6}$$

*corresponding to Student distributions with positive scale and shape parameters* $\sigma$ *and* $\nu$. *Even though the associated* $h$ *functions belong to* $H^{**}(\frac{\sigma\sqrt{\nu}}{2})$, *they do not belong to* $H^{**}(\infty)$ *(see Marchand and Perron, 2007), and part (ii) of Corollary 2 cannot be applied. However, turning to the potential applicability of part (ii) of Corollary 2,* $\bar{\rho}_h$ *admits a useful lower bound as established by Marchand and Perron (2005, Lemma 2 and equation A.4) which tells us that*

$$m\bar{\rho}_h(m,x) \leq \frac{m^2 x}{m^2 + x^2 + \sigma^2\nu}(1+\nu), \text{for all } x > 0,$$

*with equality for all* $x > 0$ *if and only if* $\nu = 1$ *(i.e., Cauchy case). Simple analysis with the upper bound tells us that:* **(a)** $2m\bar{\rho}_h(m,x) \leq x$ *for* $x \in (0,m)$ *whenever* $m \leq \sigma\sqrt{\frac{\nu}{1+2\nu}}$, *and* **(b)** $2m\bar{\rho}_h(m,x) \leq m$ *for* $x > m$ *whenever* $m \leq \sigma\sqrt{\frac{1}{\nu+2}}$. *Therefore, unifying both conditions for* $\nu \geq 1$, *we infer that any shrinkage estimator* $\delta$ *(towards* $0$*) with respect to* $\delta_0$ *is a dominator of* $\delta_{mle}$ *under (6) as soon as* $m \leq \sigma\frac{1}{\sqrt{\nu+2}}$. *In the Cauchy case, the bound becomes* $m \leq \sigma\sqrt{\frac{1}{3}}$. *We conclude by mentioning that, for the specific case of the Bayes estimator* $\delta_{BU}$ *with respect to the boundary uniform prior on* $\{-m,m\}$, *Marchand and Perron (2005) establish the stronger condition* $m \leq \sigma$ *for dominance. But the new condition above is universal in the sense that it applies to* **all** *shrinkage estimators.*

**Remark 5.** *Although this paper does not address multivariate extensions with a constraint of the form* $\|\theta\| \leq m$, *as in Marchand and Perron (2001,2005), one wonders about such extensions. Several difficulties persist, namely with regards to Bayesian estimators, but we point out that the above Student case admits the following relatively straightforward extension, which we establish in the Appendix.*

**Theorem 5.** *For $X - \theta$ distributed p-variate student with $d$ degrees of freedom and $||\theta|| \leq m$, any equivariant shrinkage estimator $\delta$ with respect to $\delta_{mle}$ (i.e., $\delta(x) = g(\|x\|)\frac{x}{\|x\|}$, with $0 \leq g(r) \leq m \wedge r$ for all $r > 0$) dominates $\delta_{mle}$ as soon as $m \leq \sqrt{\frac{p^2}{d+2p} \wedge \frac{d}{3}}$.*

# 4. Bayesian dominance results

We now turn our attention to Bayesian estimators $\delta_\pi$, associated with symmetric about 0 priors $\pi$, and work towards conditions on $\pi$ which will justify the applicability of the dominance results of Section 3. The next lemma summarizes the key relationships between $\pi$ and $\delta_\pi$ which we will make use of.

**Lemma 4.** *Let $\pi$ be a symmetric about 0 prior density on $\Theta(m)$ (with respect to a symmetric $\sigma$-finite measure $\nu$), and $\delta_\pi$ the corresponding Bayes estimator for squared error loss under model (1). Then,*

- *(a) if $h \in H^{**}(\infty)$, then the family of posterior densities $\pi(\cdot|x)$ possess increasing monotone likelihood ratio;*
- *(b) $\delta_\pi$ is nondecreasing whenever $h \in H^{**}(\infty)$;*
- *(c) $\delta_\pi$ admits the representation: $\delta_\pi(x) = E_x[U\rho_h(U,x)]$ $x \geq 0$, where $U$ has density (with respect to $\nu$) proportional to $\{(e^{-h(x-u)} + e^{-h(x+u)})\, I_{(0\,,\,m]}(u) + e^{-h(x)}\, I_{\{0\}}(u)\ \}\pi(u)$;*
- *(d) if $(m,h)$ are such that $h \in H^{*}(m)$, then $\delta_\pi(x) \leq x \wedge m$, for all $x > 0$, whenever* **(i)** *$E_x[Uh'(U)] \leq 1$, or* **(ii)** *$m \leq G_h^{-1}(1)$;*
- *(e) if $h$ is an increasing function on $(0,\infty)$, and if $\pi$ is a continuous, logconcave, and nondecreasing density on $(0,m)$, then both $x - \delta_\pi(x)$ and $x(x - \delta_\pi(x))$ increase in $x$, $x \geq 0$;*

*(f) If $\pi^*$ is another prior (also with respect to $\nu$) such that $\frac{\pi^*(\theta)}{\pi(\theta)}$ increases in $\theta$ on $[0,m]$, and if $(m,h)$ are such that $h \in H^{***}(m) \cap H^{**}(\infty)$, then $\delta_{\pi^*}(x) \geq \delta_\pi(x)$ for all $x \geq 0$;*

*(g) For $h \in H^*(\infty)$ any any prior $\pi$, the family of posterior densities of $|\theta|$ given $x$, $x \geq 0$, have increasing monotone likelihood ratio in $|\theta|$, with $x$ taken as the parameter.*

**Proof. (a,b)** Part (b) follows immediately from (a) since $\delta_\pi(x)$ is a posterior expectation. For (a), observe that, for any $(x_1,x_2)$ such that $x_1 < x_2$,

$$\frac{\pi(\theta|x_2)}{\pi(\theta|x_1)} \propto e^{h(\theta-x_1)-h(\theta-x_2)}\,,$$

which is indeed increasing in $\theta$, $\theta \in \Theta(m)$, whenever $h$ is convex; i.e, $h \in H^{**}(\infty)$.

**(c)** Start with the expression

$$\delta_\pi(x) = \frac{\int_{[-m,m]} te^{-h(x-t)}\,\pi(t)\,d\nu(t)}{\int_{[-m,m]} e^{-h(x-t)}\,\pi(t)\,d\nu(t)}\,,$$

set $u = -t$ for $t \in [-m,0\,]$, and exploit the symmetry of both $h$ and $\pi$ to obtain

$$\delta_\pi(x) = \frac{\int_{(0,m]} t\{e^{-h(x-t)} - e^{-h(x+t)}\}\pi(t)\,d\nu(t)}{\int_{(0,m]} \{e^{-h(x-t)} + e^{-h(x+t)}\}\,\pi(t)\,d\nu(t) + e^{-h(x)}\,\pi(0)\nu(\{0\})}\,.$$

Finally, the representation holds with the definition of $\rho_h$, and since $\frac{e^{-A}-e^{-B}}{e^{-A}+e^{-B}} = \tanh(\frac{A-B}{2})$, for all $A$,$B$.

**(d)** Clearly $\delta_\pi(x) \leq m$ for all $x \geq m$. Now, to have the property $\delta_\pi(x) \leq x$ for $x \in [0,m]$, condition (i) suffices as seen by making use of part (f) of Lemma 1. Condition (ii), which says that $mh'(m) \leq 1$, implies (i) given that $h$ is convex on $(0,m)$ by assumption that $h \in H^*(m)$.

**(e)** The increasing property of $x - \delta_\pi(x)$ is established within Proposition 3.1 of Kubokawa (2005), and implies as well that $x(x - \delta_\pi(x))$ increases for $x \geq 0$ (since $\delta_\pi(0) = 0$ given the symmetry of $\pi$ and $h$).

**(f)** In part (c) above, set $g_{\pi,x}$ and $g_{\pi^*,x}$ as the densities of $U$ associated with observation $x$ and with priors $\pi$ and $\pi^*$, respectively. Since the increasing property of the ratio $\frac{\pi^*(u)}{\pi(u)}$ transfers over to the ratio $\frac{g_{\pi^*,x}(u)}{g_{\pi,x}(u)}$; $u \in [0,m]$; we infer that $\delta_{\pi^*}(x) \geq \delta_\pi(x)$; $x \geq 0$; given that $u\rho_h(u,x)$ is, by part (a) of Lemma 1 and since $h \in H^{***}(m)$, increasing in $u$; $u \in [0,m]$. From the above, we now infer the following as a consequence of Corollary 2.

**(g)** The result and proof are perfectly analogous to those of Lemma 3.

**Corollary 3.** *(a) For models in (1), with* $(m,h)$ *such that* $h \in H^*(m) \cap H^{**}(\infty)$*,* **any** *Bayes estimator* $\delta_\pi$ *with respect to a symmetric about* 0 *prior dominates* $\delta_{mle}$ *on* $\Theta(m)$ *as soon as* $m \leq G_h^{-1}(\frac{1}{2}) \wedge \frac{1}{2}F_{0,h}^{-1}(\frac{5}{6})$.

*(b) For models in (1), and an arbitrary Bayes estimator* $\delta_\pi$ *with respect to a symmetric about* 0 *prior, condition (i) of Corollary 2 suffices for* $\delta_\pi$ *to dominate* $\delta_0$*.*

**Proof.** By virtue of Corollary 2, we only need to show that $\delta_\pi(x)$ is a nondecreasing shrinkage estimator in (a); and a shrinkage estimator in (b). With the given assumption on $(m,h)$, the nondecreasing and shrinkage properties required in (a) follow from parts (a) and (d, ii) (respectively) of Lemma 4. Now, suppose condition (i) of Corollary 2 holds. Then, the shrinkage property required in (b) follows from part (c) of Lemma 4 as $U\rho_h(U,x) \leq m\bar{\rho}_h(m,x)$ (with probability one, for all $x > 0$), which implies $\delta_\pi(x) \leq m\bar{\rho}_h(m,x) \leq 2m\bar{\rho}_h(m,x) \leq m \wedge x$, for all $x > 0$.

**Example 2** (continued)

For Student densities in (6), $\nu \geq 1$, it is immediate from the above that dominance holds universally for **all** symmetric Bayes estimators as soon as $m \leq \sigma\frac{1}{\sqrt{\nu+2}}$.

Now, concerning the applicability of Corollary 1 to Bayesian estimators, the results of Lemma 4 lead to the following.

**Corollary 4.** *For models in (1), with $(m,h)$ such that $h \in H^*(m) \cap H^{**}(\infty)$ and $m \leq G_h^{-1}(1)$, and for prior densities $\pi$ which are symmetric, continuous, logconcave, and nondecreasing on $(0,m)$, condition (3) is sufficient for $\delta_\pi$ to dominate $\delta_{mle}$.*

**Proof.** It suffices to check that conditions (i) and (ii) of Corollary 1 are satisfied. On one hand, (i) holds given part (d, ii) of Lemma 4 and part (b) of Lemma 2. On the other hand, part (e) of Lemma 4 tells us that (ii) is satisfied with the given assumptions on the prior $\pi$.

**Example 3.** *(Uniform and other symmetric, logconcave priors)*

*Corollary 4 applies to the important and benchmark choice of a fully uniform prior on the parameter space $\Theta(m)$ and the corresponding Bayes estimator $\delta_U$. The universal dominance result of Corollary 3 already applies, but the conditions here of Corollary 1 (or 4) give us weaker conditions (on $m$) for $\delta_U$ to dominate $\delta_{mle}$. Hence, we proceeded to evaluate, for a selection of admissible models with $h \in H^*(\infty)$, the set:*

$$C_U = \{m : m \leq G_h^{-1}(1) \text{ and } (3) \text{ is satisfied for } \delta_U\}.$$

*In Table 1, these sets $C_U$ are reported on and contrasted with the universal cutoff points, as well as the dominance cutoffs points (i.e., $m < G_h^{-1}(1)$) for the boundary uniform prior.*

*For other continuous, symmetric, logconcave priors $\pi$, which are also nondecreasing on $(0,m)$, it follows from part (f) of Lemma 4 and Corollary 4 that the dominance set*

$$C_\pi = \{m : m \leq G_h^{-1}(1) \text{ and } (3) \text{ is satisfied for } \delta_\pi\}$$

*will contain the set $C_U$. Furthermore, again by virtue of part (f) of Corollary 4), the steeper $\pi$ is on $(0,m)$, the larger $C_\pi$ is, with the steepest $\pi$ being boundary uniform in the limit with corresponding maximal $C_\pi$ given by the interval $[0,G_h^{-1}(1)]$. Finally, the above elements are partially illustrated in*

*Table 2 for various models and the logconcave priors* $\pi(\theta) \propto |\theta|^k\, I_{(-m,m)}(\theta)\,;\, k = 0,1,2,4.$

*TABLE 1*

*Sufficient conditions for dominance of* $\delta_{mle}$ *by different Bayesian estimators*

| ***Density*** $(f(x\|\theta) \propto)$ | ***Type of Prior*** | | |
|---|---|---|---|
| | ***Symmetric*** $(m \leq G_h^{-1}(\frac{1}{2}) \wedge \frac{1}{2}F_h^{-1}(\frac{5}{6}))$ | ***Boundary Uniform*** $(m \leq G_h^{-1}(1))$ | ***Uniform*** $(\{m : m \leq G_h^{-1}(1)$ and (3) is satisfied for $\delta_U\})$ |
| Normal $(e^{-\frac{1}{2}(\frac{\|x-\theta\|}{\sigma})^2})$ | $m \leq 0.4837\sigma$ | $m \leq \sigma$ | $m \leq 0.5230\sigma$ |
| Laplace $(e^{-\|\frac{x-\theta}{\sigma}\|})$ | $m \leq 0.5\sigma$ | $m \leq \sigma$ | $m \leq 0.59720\sigma$ |
| Logistic $(\frac{e^{(\frac{\|x-\theta\|}{\sigma})}}{(1+e^{(\|\frac{x-\theta}{\sigma}\|)})^2})$ | $m \leq 0.8047\sigma$ | $m \leq 1.5434\sigma$ | $m \leq 0.9004\sigma$ |
| Hyp. Sech. $(sech(\|\frac{x-\theta}{\sigma}\|))$ | $m \leq 0.6585\sigma$ | $m \leq 1.1997\sigma$ | $m \leq 0.7640\sigma$ |

*TABLE 2*

*Dominance set* $C_\pi$ *for different models and priors*

| | ***Type of Prior*** $\pi(\theta) \propto$ | | | |
|---|---|---|---|---|
| ***Density*** $(f(x\|\theta) \propto)$ | $I_{(-m,m)}(\theta)$ | $\|\theta\| I_{(-m,m)}(\theta)$ | $\theta^2 I_{(-m,m)}(\theta)$ | $\theta^4 I_{(-m,m)}(\theta)$ |
| Normal $(e^{\frac{1}{2}(\frac{\|x-\theta\|}{\sigma})^2})$ | (0,0.5230] | (0,0.6213] | (0,0.6624] | (0,0.7189] |
| Laplace $(e^{-\|\frac{x-\theta}{\sigma}\|})$ | (0,0.5972] | (0,0.7270] | (0,0.7805] | (0,0.8441] |
| Logistic $(\frac{e^{(\frac{\|x-\theta\|}{\sigma})}}{(1+e^{(\|\frac{x-\theta}{\sigma}\|)})^2})$ | (0,0.9004] | (0,0.9783] | (0,1.0222] | (0,1.0760] |
| Hyperbolic secant $(sech(\|\frac{x-\theta}{\sigma}\|))$ | (0,0.7640] | (0,0.8938] | (0,0.9534] | (0,1.0237] |

# Acknowledgements

The research support of NSERC of Canada for both Marchand and Perron is gratefully acknowledged by both authors.

## References


Casella, G. & Strawderman, W.E. (1981). Estimating a bounded normal mean. *Annals of Statistics*, **9**, 870-878.

Charras, A. & van Eeden, C. (1991). Bayes and admissibility properties of estimators in truncated parameter spaces, *Canadian Journal of Statistics*, **19**, 121-134.

Dou, Y. & van Eeden, C. (2006). Comparisons of the performances of estimators of a bounded normal mean under squared-error loss. *Technical Report 223, Department of Statistics, The University of British Columbia.*

Gatsonis, C, MacGibbon, B. & Strawderman, W.E. (1987). On the estimation of a restricted normal mean, *Statistics and Probability Letters*, **6**, 21-30.

Kubokawa, T. (2005). Estimation of bounded location and scale parameters. *Journal of the Japanese Statistical Society*, **35**, 221-249.

Marchand, É. & Strawderman, W. E. (2004). Estimation in restricted parameter spaces: A review. *Festschrift for Herman Rubin*, IMS Lecture Notes-Monograph Series, **45**, pp. 21-44.

Marchand, É. & Perron, F. (2001). Improving on the MLE of a bounded normal mean. *Annals of Statistics*, **29**, 1078-1093.

Marchand, É. & Perron, F. (2005). Improving on the MLE of a bounded mean for spherical distributions. *Journal of Multivariate Analysis*, **92**, 227-238.

Marchand, É. & Perron, F. (2007). Estimating a bounded parameter for symmetric distributions. *Annals of the Institute of Mathematical Statistics*, to appear.

Moors, J.J.A. (1981). Inadmissibility of linearly invariant estimators in the truncated parameter spaces, *Journal of the American Statistical Association*, **76**, 910-915.

Moors, J.J.A. (1985). Estimation in truncated parameter spaces. Ph.D. thesis, Tilburg University.

Perron, F. (2003). Improving on the mle of $p$ for a Binomial($n$,$p$) when $p$ is around 1/2. *Festschrift for Constance van Eeden*, IMS Lecture Notes-Monograph Series, **42**, pp. 45-64.

# 5. Appendix

The property that the covariance $Cov(f_1(X),f_2(X))$ is nonnegative for increasing $f_1$ and $f_2$ is used below.

**Proof of Theorem 1** Decomposing the difference in conditional risks for $\theta \in [0,m]$, we obtain

$$\begin{aligned}
& E_\theta\left[|\theta - \delta_{mle}(X)|^2 - |\theta - \delta(X)|^2 \,||X| > m\right] \\
= \;& E_\theta\left[\{\,(m - \delta(|X|))\mathrm{sgn}(X)\,\}\{\,(\,(m + \delta(|X|)\,\mathrm{sgn}(X)\,) - 2\theta\,\}||X| > m\right] \\
= \;& E_\theta\left[(m - \delta(|X|))(\delta(|X|) - \{2\theta\rho_h(\theta,|X|) - m\})||X| > m\right] \\
\geq \;& E_\theta\left[(m - \delta(|X|))(\delta(|X|) - \{2\theta\alpha_h(m,\theta) - m\})||X| > m\right] \\
\geq \;& E_\theta\left[(m - \delta(|X|))(\delta(|X|) - \{2\theta\bar{\alpha}_h(m) - m\})||X| > m\right] \\
> \;& 0,
\end{aligned}$$

where **(i)** both the second equality and inequality exploit property (c) of Lemma 1 as well as the definitions of $\rho_h$ and $\bar{\alpha}_h$, **(ii)** the first inequality holds with the increasing properties of $\rho_h(\theta,\cdot)$ (Lemma 1) and $\delta(\cdot)$, and **(iii)** the strict inequality follows with the given bounds on $\delta$.

**Proof of Theorem 2** Decomposing the difference in conditional risks for $\theta \in [0,m]$, we obtain

$$
\begin{aligned}
& E_\theta\left[|\theta-\delta_{mle}(X)|^2 - |\theta-\delta(X)|^2 \,|\, |X|\le m\right] \\
= \;& E_\theta\left[\{\,(|X|-\delta(|X|))\mathrm{sgn}(X)\,\}\{\,(\,(|X|+\delta(|X|)\,\mathrm{sgn}(X)\,)-2\theta\,\} \,|\, |X|\le m\right] \\
= \;& E_\theta\left[(|X|-\delta(|X|))\left(\delta(|X|)-\left\{2\theta\frac{\rho_h(\theta,|X|)}{|X|}-1\right\}|X|\right) \,|\, |X|\le m\right] \\
\ge \;& E_\theta\left[(|X|-\delta(|X|))\,(\delta(|X|)-\{2\beta_h(m,\theta)-1\}|X|) \,|\, |X|\le m\right] \\
\ge \;& E_\theta\left[(|X|-\delta(|X|))\,(\delta(|X|)-\{2\bar{\beta}_h(m)-1\}|X|) \,|\, |X|\le m\right] \\
> \;& 0
\end{aligned}
$$

where **(i)** both the second equality and inequality exploit property (c) of Lemma 1 as well as the definitions of $\rho_h$ and $\bar{\beta}_h$, **(ii)** the first inequality holds given that $\delta(x)$, $x(x-\delta(x))$ increase in $x$; $x \in (0,m)$; and $\frac{\rho_h(\theta,x)}{x}$ decreases in $x$; $x > 0$; (part e of Lemma 1), and **(iii)** the strict inequality follows with the given assumptions on $\delta$.

**Proof of Theorem 5.** It follows from Marchand and Perron (2005) (Lemma 2, Theorem 2) that a sufficient condition for a shrinkage $\delta$ to dominate $\delta_{mle}$ is:

$$\frac{2m^2 r\gamma}{m^2+r^2+d} \le m \wedge r, \text{ for all } r>0, \text{ with } \gamma = 1+(1\vee\frac{d}{p}).$$

For $r \le m$, the above requires: $\frac{2m^2\gamma}{m^2+d} \le 1 \Leftrightarrow m^2 \le \frac{d}{2\gamma-1}$.

For $r > m$, we require:

$$
\begin{aligned}
\frac{2m^2 r\gamma}{m^2+r^2+d} \le m \text{ for all } r>m \quad &\Leftrightarrow \quad \frac{2m^2\,\gamma\,\sqrt{m^2+d}}{m^2+(m^2+d)+d} \le m \\
&\Leftrightarrow \quad m^2 \le \frac{d}{\gamma^2-1}.
\end{aligned}
$$

Finally, the result follows since $\gamma^2-1 \ge 2\gamma-1$ given that $\gamma \ge 2$, and by verifying that $\frac{d}{\gamma^2-1} = \frac{p^2}{d+2p}\wedge\frac{d}{3}$.