

ASYMPTOTICALLY OPTIMAL AND ADMISSIBLE DECISION RULES IN COMPOUND COMPACT GAUSSIAN SHIFT EXPERIMENTS

BY SUMAN MAJUMDAR

University of Connecticut

Asymptotically optimal and admissible compound decision rules are obtained in a Hilbert-parameterized Gaussian shift experiment. The component parameter set is restricted to compact. For the squared error loss, every compound Bayes estimator is admissible and every compound estimator Bayes versus full support hyperprior mixture of iid priors on the compound parameter is asymptotically optimal. For the latter class of rules induced by full support hyperpriors, asymptotic optimality and admissibility extend to equi- (in decisions) uniformly continuous and bounded risk functions. Normality of certain mixtures of the standard Gaussian process and qualitative robustness of the component Bayes estimator (results of independent interest used in the paper) are derived.

1. Introduction, notation and preliminaries. In this paper, we derive *asymptotically optimal* [in the sense of Robbins (1951)] and *admissible compound* decision rules when the *component* experiment is a *Gaussian shift* experiment [in the sense of Le Cam (1986)]. The next paragraph describes the notational conventions adopted. Then the idea of compounding a decision problem (called the component problem), first espoused by Robbins (1951), is discussed, followed (in that order) by a one-paragraph summary and some preliminaries relating the paper to the pertinent literature.

For probabilities (sets) P_1, \dots, P_n (A_1, \dots, A_n), $\times_{i=1}^n P_i$ ($\times_{i=1}^n A_i$) denotes their measure (set) theoretic product; when $P_i = P$ ($A_i = A$) $\forall i$, $\times_{i=1}^n P_i$ ($\times_{i=1}^n A_i$) is denoted by P^n (A^n). To denote the integral of a function f with respect to (wrt) a measure μ , we shall interchangeably use the standard integral notation $\int f d\mu$ [with the dummy variable of integration sometimes (partially) displayed] and the left operator notation $\mu(f)$ (or even μf). Sets (probabilities) are always identified with their indicator functions (induced expectations); \Re stands for the real line. If X is a random element on a probability space (\cdot, \cdot, P) , the PX^{-1} denotes the P -induced distribution of X on the range space. The notation $a := b$ will mean that a is defined to be b . For any n -tuple $\mathbf{x} := (x_1, \dots, x_n)$, \mathbf{x}_α will denote the α -delete vector $(x_1, \dots, x_{\alpha-1}, x_{\alpha+1}, \dots, x_n)$.

Received August 1993; revised October 1994.

AMS 1991 subject classifications. Primary 62C25; secondary 62C15, 62E15, 62F35.

Key words and phrases. Compound Bayes rules, component Bayes estimator, asymptotic optimality, admissibility, full support hyperprior, consistency, Gaussian shift experiment, qualitative robustness.

Roughly speaking, in standard iid statistics one models the data available as independent realizations of a random variable whose distribution is governed by the unknown state of nature θ ; in other words, the data are considered to be generated under identical states of nature. However, from a modeling perspective, it is often appropriate to consider the scenario when the data are generated under *similar*, but not identical, states of nature. In that case, it is only sensible to use the available data to infer about the entire collection of similar states of nature. In the compound formulation of a decision problem this idea is formally dealt with. Asymptotics, as in the standard iid case, is in terms of n , the number of data points, going to infinity. Thus one looks at the decision problem with parameter set Θ^n , family of probability measures $\{\mathbb{P}_\theta: \theta \in \Theta^n\}$ on the measurable space $(\mathcal{X}^n, \mathcal{F}^n)$, where $\mathbb{P}_\theta := \times_{\alpha=1}^n P_{\theta_\alpha}$, observations $\mathbf{X} \sim \mathbb{P}_\theta$ under θ , action space \mathcal{A}^n , decision rules $\mathbf{t}: \mathcal{X}^n \mapsto \mathcal{A}^n$ such that each $L(t_\alpha, \theta_\alpha)$ is measurable, loss and corresponding risk

$$L_n(\mathbf{t}, \theta) := n^{-1} \sum_{\alpha=1}^n L(t_\alpha, \theta_\alpha),$$

$$R_n(\mathbf{t}, \theta) := \mathbb{P}_\theta L_n(\mathbf{t}, \theta).$$

The decision problem $(\mathcal{X}, \mathcal{F}, \{P_\theta: \theta \in \Theta\}, \mathcal{A}, L)$ is traditionally referred to as the component problem.

For consideration of Bayes rules, we fix a σ -algebra of subsets of Θ such that each of the maps $(x, \theta) \mapsto L(t(x), \theta)$ is jointly measurable. Let $\Omega = \{\omega: \omega \text{ is a probability on } \Theta\}$. For $\omega \in \Omega$, let $r(\omega)$ and τ_ω , respectively, denote the minimum Bayes risk and a Bayes rule versus ω (we assume existence of τ_ω for every ω); that is,

$$r(\omega) = \inf \left\{ \int_{\Theta} R(t, \theta) d\omega(\theta) : t \right\} = \int_{\Theta} R(\tau_\omega, \theta) d\omega(\theta).$$

A compound rule \mathbf{t} , for which $t_\alpha(\mathbf{x}) = t(x_\alpha) \forall \alpha = 1, \dots, n$, where t is a component rule, is called *simple symmetric*. Let G_n denote the empirical distribution of θ . The compound risk at θ of a simple symmetric \mathbf{t} reduces to the component Bayes risk of t versus G_n , where $t_\alpha(\mathbf{x}) = t(x_\alpha) \forall \alpha = 1, \dots, n$; as such it is at least $r(G_n)$, which is referred to as the *simple envelope* at θ . For a compound rule \mathbf{t} , the difference $D_n(\mathbf{t}, \theta) = R_n(\mathbf{t}, \theta) - r(G_n)$ is called the *modified regret* of \mathbf{t} at θ and a sequence of compound rules $\{\mathbf{t}: n \geq 1\}$ is said to be *asymptotically optimal* (a.o.) if

$$\sup\{D_n(\mathbf{t}, \theta): \theta\} \rightarrow 0 \quad \text{as } n \rightarrow \infty.$$

Now, it has long been recognized [Hannan and Robbins (1955)] that the compound problem is invariant under the group of $n!$ permutations of coordinates. Thus, one might consider using the more stringent *equivariant envelope*, the minimum compound risk of equivariant rules, to judge the performance of a compound rule [see Gilliland and Hannan (1986)]. However, the uniform (in states) asymptotic equivalence of the two envelopes, estab-

lished by Mashayekhi (1990) under fairly general conditions, enables one to start with the simple envelope and extend (the asymptotic optimality result) to the equivariant envelope by verifying Mashayekhi's conditions.

A sequence of compound rules $\{\mathbf{t}: n \geq 1\}$ is said to be *admissible* if, for every n , \mathbf{t} is admissible in the usual sense.

The component experiment considered in this paper is a *Gaussian shift* experiment (indexed by a real separable Hilbert space H) in the sense of Le Cam (1986) (Section 2). When $\Theta(\subseteq \mathcal{A} \subseteq H)$ is strongly compact and $L(a, \theta) = \|a - \theta\|^2$, we derive (Section 3) that all compound Bayes rules are *admissible* and compound rules that are Bayes versus a hyperprior mixture of iid priors on the compound parameter are *asymptotically optimal* if the mixing hyperprior has full support. We extend (Section 4) the *asymptotic optimality* and *admissibility* of the Bayes rules induced by full support hyperpriors to loss functions L that yield an equi- (in t) uniformly continuous and bounded (in θ) risk function R .

Since the formulation of the compound problem by Robbins (1951), a lot of different component problems have been *compounded*, generating a huge literature. Of special pertinence are Gilliland, Hannan and Huang (1976), Datta (1988, 1991), Mashayekhi (1990, 1993), Zhu (1992), Majumdar, Gilliland and Hannan (1993) and Majumdar (1994). The common feature of these works (including this one) is the consideration of compound Bayes rules versus a hyperprior Λ (on Ω) mixture of iid priors on the compound parameter. The α -th component of such a rule is a component Bayes rule (evaluated at X_α) versus the posterior mean of ω (under Λ) given \mathbf{X}_α when $\omega \sim \Lambda$ and given ω , $\mathbf{X}_\alpha \sim P_\omega^{n-1}$. As shown below, this is a consequence of disintegrability of joint distributions under the Polish assumption [Le Cam (1986)], a fact obscured by unnecessary domination assumptions in the previous expositions.

Assuming that Θ is a Polish space, by Theorem II.6.2 and II.6.5 of Parthasarathy (1967), Ω with the topology of weak convergence is also a Polish space; let $\mathcal{B}(\Omega)$ denote its Borel σ -field. Let Λ be a hyperprior on $(\Omega, \mathcal{B}(\Omega))$. We take Λ -mixture of iid priors on Θ^n (for each n) and denote that prior by $\beta_{\Lambda, n}$. [The measure $\beta_{\Lambda, n}$ is defined on the class of measurable rectangles by

$$(1.1) \quad \beta_{\Lambda, n}(B_1 \times B_2 \times \cdots \times B_n) = \int_{\Omega} \prod_{i=1}^n \omega(B_i) d\Lambda(\omega),$$

and then extended to the product σ -field. Note that by Lemma A.1 of Majumdar (1993) the above integrand is measurable.]

Let $\mathbf{t} = (t_1, \dots, t_n)$, where $t_\alpha: \mathcal{X}^n \mapsto \mathcal{A}$ is a measurable function, be a decision rule in the compound problem. The α -th component Bayes risk of \mathbf{t} versus $\beta_{\Lambda, n}$ is

$$(1.2) \quad R(t_\alpha, \beta_{\Lambda, n}) = \int_{\Omega} \int_{\mathcal{X}^{n-1}} \left[\int_{\Theta} \int_{\mathcal{X}} L(t_\alpha, \theta_\alpha) dP_{\theta_\alpha} d\omega \right] dP_\omega^{n-1} d\Lambda.$$

Disintegrating the joint probability on $\mathcal{X}^{n-1} \times \Omega$ determined by $dP_\omega^{n-1} d\Lambda$ as $d\Lambda_{\alpha,n} d\mathbf{P}_{\beta_{\Lambda_{\alpha,n}}}$, where $\Lambda_{\alpha,n}$ is the posterior distribution of ω (under Λ) given \mathbf{X}_α when $\omega \sim \Lambda$ and, given ω , $\mathbf{X}_\alpha \sim P_\omega^{n-1}$ [since Ω is a Polish space, by Theorem 10.2.2 of Dudley (1989), such a disintegration exists], we get

$$(1.3) \quad [\text{l.h.s. (1.2)}] = \int_{\mathcal{X}^{n-1}} \left[\int_{\Theta} \int_{\mathcal{X}} L(t_\alpha, \theta_\alpha) dP_{\theta_\alpha} d\beta_{\Lambda_{\alpha,n}} \right] d\mathbf{P}_{\beta_{\Lambda_{\alpha,n}}},$$

where $\beta_{\Lambda_{\alpha,n}}$ denotes the $\Lambda_{\alpha,n}$ -mixture of ω 's. Clearly, the r.h.s. of (1.3) is minimized by $t_\alpha(\mathbf{x}) = \tau_{\beta_{\Lambda_{\alpha,n}}}(x_\alpha)$. Since the compound risk is the average of the component risks, a compound rule Bayes versus $\beta_{\Lambda,n}$ is given by \mathbf{t}_Λ , where

$$(1.4) \quad t_{\Lambda,\alpha}(\mathbf{x}) = \tau_{\beta_{\Lambda_{\alpha,n}}}(x_\alpha).$$

2. The Gaussian shift experiment. Let H be a real separable Hilbert space (with $\|f\|$ denoting the norm of an element f in H and $\langle \cdot, \cdot \rangle$ denoting the inner product). Let $\{P_\theta: \theta \in H\}$ be a family of probabilities on a measurable space $(\mathcal{X}, \mathcal{F})$ specified by (strictly positive) densities $\{p_\theta: \theta \in H\}$ wrt $\mu := P_0$ such that the following holds:

- (A) the map $\theta \mapsto l_\theta(\cdot) := \ln p_\theta(\cdot) + \|\theta\|^2/2$ is linear from H into the linear space of real-valued measurable functions on $(\mathcal{X}, \mathcal{F})$.

Each of the following conclusions can be deduced from a combination of (A) and the preceding ones [for details, see Majumdar (1993)]:

$$(2.1) \quad P_\xi l_\theta^{-1} = \mathcal{N}(\langle \theta, \xi \rangle, \|\theta\|^2) \quad \forall \theta, \xi \in H;$$

$$(2.2) \quad \{l_\theta: \theta \in H\} \text{ is a centered Gaussian process on } (\mathcal{X}, \mathcal{F}, \mu);$$

$$(2.3) \quad P_\xi([\!|l_\theta - \langle \theta, \xi \rangle|])[\!|l_\eta - \langle \eta, \xi \rangle|]) = \langle \theta, \eta \rangle \quad \forall \theta, \eta, \xi \in H.$$

[Actually, the linearity assumption in (A) can be *weakened* to a.s. equality of l_ξ and $(l_\theta + l_\eta)/2$, where $\xi = (\theta + \eta)/2$, $\forall \theta$ and η ; see Le Cam (1986), Lemma 9.3.2. Also, for the derivation of (2.1)–(2.3), see Millar (1983), V.3]. Thus, $\{l_\theta: \theta \in H\}$ is a *standard Gaussian process* in the sense of Definition 68.3 of Strasser (1985). Clearly, that identifies the experiment $\{P_\theta: \theta \in H\}$ to be a *Gaussian shift* one, not only in the broader sense of Definition 2 of Le Cam [(1986), Chapter 9], but in the somewhat narrower (compared to Le Cam) sense of Definition 69.2 of Strasser (1985) as well [via the characterization in Strasser (1985), Theorem 69.4]. Note that, the experiment $\{P_\theta: \theta \in H\}$ is what Le Cam (1986) calls a *standard Gaussian shift experiment* of H .

In the remainder of this section, we state and prove Theorem 2.1 (establishing normality of certain mixtures of the standard Gaussian process), which is used in the proof of Theorem 3.1, and give four examples of Gaussian shift experiments.

THEOREM 2.1. *Let ω be a Borel probability measure on H such that $\omega(\|\cdot\|)$ is finite. For $x \in \mathcal{H}$, let*

$$l_\omega(x) := \int_H l_\theta(x) d\omega(\theta).$$

Then, $\forall \xi \in H, P_\xi l_\omega^{-1} = \mathcal{N}(\omega(\langle \cdot, \xi \rangle), \omega^2(\langle \cdot, \cdot \rangle))$.

PROOF. Let G be the closure [in $L_2(P_\xi)$] of the linear subspace $\{l_\theta - \langle \theta, \xi \rangle : \theta \in H\}$. Note that $P_\xi f^{-1}$ is a centered Gaussian distribution on the line for every $f \in G$. The (crucial) normality assertion is proved by showing that $l_\omega - \omega(\langle \cdot, \xi \rangle)$ is an element of G . The three main steps of the proof are as follows:

- (i) $P_\xi l_\omega^{-1}$ is well defined;
- (ii) $l_\omega \in L_2(P_\xi)$;
- (iii) $P_\xi(l_\omega - \omega(\langle \cdot, \xi \rangle))^2 = \omega^2(\langle \cdot, \cdot \rangle) = P_\xi(\Pi_G)^2$, where Π_G is the projection of $l_\omega - \omega(\langle \cdot, \xi \rangle)$ on G .

Note that, since $l_\omega - \omega(\langle \cdot, \xi \rangle) = \Pi_G + \Pi_{G^\perp}$ and $P_\xi(\Pi_G)^2 + P_\xi(\Pi_{G^\perp})^2 = P_\xi(l_\omega - \omega(\langle \cdot, \xi \rangle))^2$ by the properties of projection [Rudin (1987), Theorem 4.11], where G^\perp is the ortho-complement of G , step (iii) proves the theorem, including the expression for the variance.

Now, the proofs of (i), (ii) and the first equality in (iii) are essentially careful applications of the Fubini theorem [Rudin (1987), Theorem 8.8]. Note that the map $(\theta, x) \mapsto l_\theta(x)$ is measurable [see Majumdar (1993), Lemma 3.0]. Use that, (2.1) and the Fubini theorem to conclude $l_\omega \in L_1(\mu)$, whence $x \mapsto l_\omega(x)$ is measurable and l_ω is finite P_ξ -a.s., establishing (i).

Use the measurability of $(\theta, \eta, x) \mapsto [l_\theta(x) - \langle \theta, \xi \rangle][l_\eta(x) - \langle \eta, \xi \rangle]$, the Cauchy-Schwarz inequality in $L_2(P_\xi)$, (2.1), the Fubini theorem and (2.3) (not necessarily in that order), and conclude that

$$P_\xi(l_\omega - \omega(\langle \cdot, \xi \rangle))^2 = \omega^2(\langle \cdot, \cdot \rangle) \leq (\omega(\|\cdot\|))^2 < \infty,$$

by the assumption of the theorem.

For the second equality of (iii), note that

$$(2.4) \quad P_\xi(h\Pi_G) = P_\xi\left(h\left[l_\omega - \int_H \langle \eta, \xi \rangle d\omega(\eta)\right]\right) \quad \forall h \in G,$$

which on interchange of the order of integration gives

$$(2.5) \quad [\text{l.h.s. (2.4)}] = \int_H P_\xi(h[l_\theta - \langle \theta, \xi \rangle]) d\omega(\theta).$$

Using (2.5) with $h = \Pi_G$,

$$(2.6) \quad P_\xi(\Pi_G)^2 = \int_H P_\xi(\Pi_G[l_\eta - \langle \eta, \xi \rangle]) d\omega(\eta);$$

by (2.5) again, with $h = l_\eta - \langle \eta, \xi \rangle$,

$$\begin{aligned}
 P_\xi([l_\eta - \langle \eta, \xi \rangle] \Pi_G) &= \int_H P_\xi([l_\eta - \langle \eta, \xi \rangle][l_\theta - \langle \theta, \xi \rangle]) d\omega(\theta) \\
 (2.7) \qquad \qquad \qquad &= \int_H \langle \eta, \theta \rangle d\omega(\theta) \quad \text{by (2.3)}.
 \end{aligned}$$

The second equality of (iii) is now obtained from (2.6) and (2.7). \square

2.1. Examples of Gaussian shift experiments. Multivariate normal distributions with a common covariance constitute a Gaussian shift experiment when the usual Euclidean inner product is replaced by an inverse covariance weighted version. The translation of the standard Brownian motion on $[0, 1]$ by the indefinite integrals of square integrable functions on $[0, 1]$ generates a Gaussian shift experiment [Millar (1983), Example V.3.13]. The log-likelihood ratio of a LAN experiment differs from that of a Gaussian shift experiment by an additive remainder term, which is asymptotically zero in probability [Ibragimov and Khas'minskii (1991), Section 2]. Gaussian shift experiments also occur as limits of products of nonparametric experiments (in the sense of weak convergence of experiments introduced by Le Cam, as the number of factors goes to infinity) indexed by square integrable mean-zero functions on an arbitrary probability space $(\mathcal{X}, \mathcal{F}, P)$, so that for *eligible* functions the corresponding distribution has a *smooth* density wrt P and otherwise the experiment is a contraction map in the Hellinger distance [Le Cam and Yang (1990), Lemma 6.4.1, and Strasser (1985), Example 80.4].

3. Compound Bayes estimation of the Gaussian shift parameter.

In this section, we consider the component problem with Θ (initially) a bounded subset of H (we shall further restrict Θ later), $\{P_\theta: \theta \in \Theta\}$ a sub-experiment of the H -indexed Gaussian shift experiment on $(\mathcal{X}, \mathcal{F})$, $\Theta \subseteq \mathcal{A} \subseteq H$ and $L(a, \theta) = \|a - \theta\|^2$. In Section 3.1, we specialize the compound rule t_Λ of (1.4) to this context and reduce (via Proposition 3.1 and Theorem 3.1) the question of asymptotic optimality of such rules to that of consistency of certain posteriors (Theorem 3.2), which is settled in Majumdar (1993) for hyperpriors with full topological support. Remark 3.1 interprets the result of Theorem 3.1 as a qualitative robustness property of the component Bayes estimator wrt the prior specification. Remark 3.3 verifies Mashayekhi's (1990) condition for asymptotic equivalence of the simple and the equivariant envelopes. To begin with, we show that *the* component Bayes estimator is the Petis integral of the identity wrt the posterior distribution (Lemma 3.1); the uniqueness is used to establish admissibility of every compound Bayes estimator (Section 3.2).

Before proceeding further, let us note that Ω (with the topology of weak convergence) is a separable metric space [see Parthasarathy (1967), Theorem II.6.2] and note that $(\theta, x) \mapsto p_\theta(x)$ and $(\omega, x) \mapsto p_\omega(x) := \int p_\theta(x) d\omega(\theta)$ are jointly measurable when Ω is endowed with the Borel σ -field [see Majumdar

(1993), Lemma 3.0]. That justifies the interchange of the order of integration in subsequent calculations without further comment. The component Bayes estimator versus ω is denoted by τ_ω . Throughout the remainder of this section, let

$$(3.1) \quad D := \sup\{\|\theta\|: \theta \in \Theta\},$$

and let $\|f\|_q$ denote the $L_q(\mu)$ norm of a function f .

LEMMA 3.1. *On the common support of $\{P_\nu: \nu \in \Omega\}$, τ_ω is the unique mapping into H satisfying $\langle \tau_\omega, h \rangle = \int_\Theta \langle \eta, h \rangle (p_n/p_\omega) d\omega(\eta) \forall h \in H$.*

PROOF. Since the map $h \mapsto \int_\Theta \langle \eta, h \rangle d\pi(\eta)$ is a linear functional on H whose norm is bounded by $D \forall \pi \in H$, by the Riesz–Fréchet theorem [see Dudley (1989), Theorem 5.5.1], \exists a unique element $v(\pi)$ in H satisfying

$$(3.2) \quad \langle v(\pi), h \rangle = \int_\Theta \langle \eta, h \rangle d\pi(\eta) \quad \forall h \in H.$$

Note that if $p_\omega(x)$ is positive, the map $\theta \mapsto p_\theta(x)/p_\omega(x)$ is a density (wrt ω) of the posterior measure $\tilde{\omega}_x$ on Θ . By (3.2), it is enough to show that $\tau_\omega = v(\tilde{\omega})$ on the common support of $\{P_\nu: \nu \in \Omega\}$. Now, by Fubini's theorem, the Bayes risk (versus ω) of an estimator t is equal to

$$(3.3) \quad \int_{\mathcal{X}} \int_\Theta \|t(x) - \theta\|^2 d\tilde{\omega}_x p_\omega(x) d\mu(x).$$

Triangulating around $v(\tilde{\omega}_x)$ and expanding the norm square of the sum, the inner integral in (3.3) is

$$\|t(x) - v(\tilde{\omega}_x)\|^2 + \int_\Theta \|v(\tilde{\omega}_x) - \theta\|^2 d\tilde{\omega}_x,$$

which is minimized iff $t(x) = v(\tilde{\omega}_x)$, completing the proof. \square

3.1. *Asymptotic optimality of \mathbf{t}_Λ .* In this subsection, we first obtain a useful upper bound on the absolute modified regret of \mathbf{t}_Λ in Proposition 3.1. In Theorem 3.1, we obtain an upper bound on the expected distance between two component Bayes estimators in terms of the total variation distance between the corresponding mixtures, which is used in conjunction with Proposition 3.1 to reduce the asymptotic optimality of \mathbf{t}_Λ to consistency of certain posteriors (Theorem 3.2).

PROPOSITION 3.1. *We have*

$$|D_n(\mathbf{t}_\Lambda, \boldsymbol{\theta})| \leq 4Dn^{-1} \sum_{\alpha=1}^n \mathbf{P}_\theta P_{\theta_\alpha} \|\tau_{\beta_{\Lambda, \alpha, n}} - \tau_{G_n}\|,$$

where G_n is the empirical distribution of $\boldsymbol{\theta}$ and $\beta_{\Lambda, \alpha, n}$ is as in (1.3).

PROOF. By definition of the quantities involved, for any compound estimator \mathbf{t} ,

$$D_n(\mathbf{t}, \boldsymbol{\theta}) = n^{-1} \sum_{\alpha=1}^n \mathbb{P}_{\boldsymbol{\theta}}(\|t_{\alpha} - \theta_{\alpha}\|^2 - \|\tilde{t}_{\alpha} - \theta_{\alpha}\|^2),$$

where $\tilde{t}_{\alpha}(\mathbf{x}) = \tau_{G_n}(x_{\alpha})$. Using the Cauchy–Schwarz inequality to bound the absolute difference between $\|d\|^2$ and $\|b\|^2$ by $\|d + b\|$ times $\|d - b\|$, triangle inequality in H and (3.1), we get

$$(3.4) \quad |D_n(\mathbf{t}, \boldsymbol{\theta})| \leq 4Dn^{-1} \sum_{\alpha=1}^n \mathbb{P}_{\boldsymbol{\theta}}\|t_{\alpha} - \tilde{t}_{\alpha}\|.$$

Since $t_{\Lambda, \alpha}(\mathbf{x}) = \tau_{\beta_{\Lambda, \alpha, n}}(x_{\alpha})$,

$$(3.5) \quad \mathbb{P}_{\boldsymbol{\theta}}\|t_{\Lambda, \alpha} - \tilde{t}_{\alpha}\| = \mathbb{P}_{\boldsymbol{\theta}}P_{\theta_{\alpha}}\|\tau_{\beta_{\Lambda, \alpha, n}} - \tau_{G_n}\|,$$

completing the proof. \square

In Theorem 3.1 we derive a uniform (in θ) bound on $P_{\theta}\|\tau_{\omega} - \tau_{\pi}\|$ in terms of the total variation distance between P_{ω} and P_{π} . Abusing notation, we shall let $\|\sigma\|$ denote the total variation norm of a signed measure σ on $(\mathcal{X}, \mathcal{F})$ as well.

The next three lemmas are used to prove Theorem 3.1.

LEMMA 3.2. For $(y, z, Y, Z, L) \in \mathfrak{R}^5$ such that $z \neq 0$ and $L \geq 0$,

$$|z| \left| \left| \frac{y}{z} - \frac{Y}{Z} \right| \wedge L \right| \leq |y - Y| + \left(\left| \frac{y}{z} \right| + L \right) |z - Z|.$$

PROOF. See Datta [(1988), Lemma A.1].

LEMMA 3.3. For every finite sequence $\{\theta_i: 1 \leq i \leq k\} \subset H$ and $\{a_i: 1 \leq i \leq k\} \subset \mathfrak{R}$,

$$2 \log \left(\int \prod_{i=1}^k p_{\theta_i}^{a_i} d\mu \right) = \left\| \sum_{i=1}^k a_i \theta_i \right\|^2 - \sum_{i=1}^k a_i \|\theta_i\|^2.$$

PROOF. See Majumdar [(1993), Lemma 3.1].

LEMMA 3.4. For every $\omega \in \Omega$ and every integer $q \geq 1$,

$$p_{\omega} \in L_q(\mu) \quad \text{and} \quad \|p_{\omega}\|_q \leq e^{(q-1)D^2/2}.$$

PROOF. See Majumdar [(1993), Lemma 3.2].

THEOREM 3.1 (The key step and the robustness result). Assume Θ is strongly totally bounded. Then, for every $\gamma > 0$, \exists a number \mathcal{N} such that

$$P_{\theta}\|\tau_{\omega} - \tau_{\pi}\| \leq 5\gamma + \mathcal{N}\|P_{\omega} - P_{\pi}\|.$$

PROOF. By Lemma 3.1,

$$(3.6) \quad \|\tau_\omega - \tau_\pi\| = \bigvee_{\mathscr{H}} \left\| \frac{\int_{\Theta} \langle \eta, h \rangle p_\eta d\omega(\eta)}{p_\omega} - \frac{\int_{\Theta} \langle \eta, h \rangle p_\eta d\pi(\eta)}{p_\pi} \right\|,$$

where $\mathscr{H} := \{h \in H: \|h\| \leq 1\}$.

Applying Lemma 3.2 (to get rid of the difference of ratios) with $z = p_\omega$, $y = \int_{\Theta} \langle \eta, h \rangle p_\eta d\omega(\eta)$, $Z = p_\pi$, $Y = \int_{\Theta} \langle \eta, h \rangle p_\eta d\pi(\eta)$ and $L = 2D$,

$$(3.7) \quad p_\omega \left| \frac{\int_{\Theta} \langle \eta, h \rangle p_\eta d\omega(\eta)}{p_\omega} - \frac{\int_{\Theta} \langle \eta, h \rangle p_\eta d\pi(\eta)}{p_\pi} \right| \\ \leq \left| \int_{\Theta} \langle \eta, h \rangle p_\eta d(\omega - \pi)(\eta) \right| + 3D|p_\omega - p_\pi|.$$

It turns out that we can bound p_θ uniformly (in θ) by a multiple of p_ω on a subset of \mathscr{X} with μ -measure arbitrarily close to 1. Since $\|\tau_\omega - \tau_\pi\| \leq 2D$ and $\|p_\theta\|_2 \leq \exp(D^2/2)$ (see Lemma 3.4), P_θ -expectation of $\|\tau_\omega - \tau_\pi\|$ on the complement can be made arbitrarily small, which allows us to concentrate on P_θ -expectation of $\|\tau_\omega - \tau_\pi\|$ on the set of large μ -measure.

More specifically, for all real numbers a and b ,

$$(3.8) \quad \exp(-D^2/2 - a + b)p_\theta[l_\theta \leq a][l_\omega > b] \leq p_\omega.$$

Because, $p_\omega = \int \exp(l_\theta - \|\theta\|^2/2) d\omega$ by definition, whence using (3.1) to bound $\exp(-\|\theta\|^2/2)$ below, applying Jensen's inequality and noting that $p_\theta[l_\theta \leq a]\exp(-a) \leq 1$ and $\exp(l_\omega) \geq \exp(b)[l_\omega > b]$, (3.8) follows. Note that the μ -measure of the complement of $[l_\theta \leq a][l_\omega > b]$ is bounded by

$$(3.9) \quad \mu[l_\theta > a] + \mu[l_\omega \leq b].$$

Since $\mu l_\theta^{-1} = \mathcal{N}(0, \|\theta\|^2)$ [(2.1)] and $\mu l_\omega^{-1} = \mathcal{N}(0, \omega^2(\langle \cdot, \cdot \rangle))$ (Theorem 2.1), and since both the variances are bounded by D^2 , we have, for $a > 0$ and $b < 0$,

$$(3.10) \quad (3.9) \leq \Pr[Z \geq a/D] + \Pr[Z \leq b/D],$$

where $Z \sim \mathcal{N}(0, 1)$. Therefore, by appropriate choice of a and b (to be made later), (3.9) can be made arbitrarily small uniformly in θ and ω .

For $a > 0$ and $b < 0$, partitioning \mathscr{X} into $[l_\theta \leq a][l_\omega > b]$ and its complement, bounding $\|\tau_\omega - \tau_\pi\|$ by $2D$ on the complement and P_θ -measure of the complement using the Cauchy-Schwarz inequality in $L_2(\mu)$ along with $\|p_\theta\|_2 \leq \exp(D^2/2)$ and (3.10), we get

$$(3.11) \quad P_\theta \|\tau_\omega - \tau_\pi\| \leq 2D \exp(D^2/2) \{\text{r.h.s. (3.10)}\}^{1/2} \\ + \mu(p_\theta \|\tau_\omega - \tau_\pi\| [l_\theta \leq a][l_\omega > b]).$$

In the next three paragraphs, we develop a bound for the second term above, culminating in the bound (3.17) for the l.h.s. of (3.11).

By (3.8),

$$(3.12) \quad [\text{second term in r.h.s. (3.11)}] \leq \exp(D^2/2 + a - b) \mu(p_\omega \|\tau_\omega - \tau_\pi\|).$$

By (3.6) and (3.7),

$$(3.13) \quad \mu(p_\omega \| \tau_\omega - \tau_\pi) \leq \mu \left(\bigvee_{\mathscr{H}} \left| \int_{\Theta} \langle \eta, h \rangle p_\eta d(\omega - \pi)(\eta) \right| \right) + 3D \|P_\omega - P_\pi\|.$$

Now, \mathscr{H} is weakly compact by a very special application of the Banach–Alaoglu theorem [Rudin (1973), Theorem 3.15]. Since H is separable, the weak topology of \mathscr{H} is metrizable [Rudin (1973), Theorem 3.16]. Since Θ is strongly totally bounded, given $\varepsilon > 0$, $\exists \delta > 0$ such that $d_w(h, h') < \delta$ implies

$$(3.14) \quad \bigvee_{\Theta} |\langle \theta, h \rangle - \langle \theta, h' \rangle| \leq \varepsilon,$$

where d_w is a metric for the weak topology of \mathscr{H} . By (3.14),

$$(3.15) \quad \mu \left| \int (\langle \theta, h \rangle - \langle \theta, h' \rangle) p_\theta d\omega \right| \leq \varepsilon \quad \text{if } d_w(h, h') < \delta.$$

If d_w -balls of radius δ around $\{h_1, \dots, h_I\}$ cover \mathscr{H} , then triangulating around appropriate h_i , using (3.15) and dominating the maximum of I nonnegative terms by their sum, we get

$$(3.16) \quad \mu \left(\bigvee_{\mathscr{H}} \left| \int \langle \theta, h \rangle p_\theta d(\omega - \pi) \right| \right) \leq 2\varepsilon + \sum_{i=1}^I \mu \left| \int \langle \theta, h_i \rangle p_\theta d(\omega - \pi) \right|.$$

Substituting the bound on the second term [obtained via (3.12), (3.13) and (3.16)] in the r.h.s. of (3.11), we get that, given $\varepsilon > 0$, $\exists \delta > 0$ such that

$$(3.17) \quad \begin{aligned} P_\theta \| \tau_\omega - \tau_\pi \| &\leq 2D \exp(D^2/2) \{\text{r.h.s. (3.10)}\}^{1/2} \\ &+ \exp(D^2/2 + a - b) \\ &\times \left[2\varepsilon + \sum_{i=1}^I \mu \left| \int \langle \theta, h_i \rangle p_\theta d(\omega - \pi) \right| + 3D \|P_\omega - P_\pi\| \right], \end{aligned}$$

where $a > 0$ and $b < 0$, and d_w -balls of radius δ around $\{h_1, \dots, h_I\}$ cover \mathscr{H} .

The culminating point of the next segment (spread over three paragraphs) is the bound (3.22) for $\mu \left| \int \langle \theta, h \rangle p_\theta d(\omega - \pi) \right|$, $h \in \mathscr{H}$.

Expanding the function $\lambda \mapsto \exp(\lambda \langle \theta, h \rangle)$ in a Taylor series around $\lambda = 0$ up to second order, collecting the terms in the l.h.s. of (3.18) on one side of the equality, and using the Cauchy-Schwarz inequality in H and (3.1) to bound the other side, we get, for $\lambda > 0$ and $h \in \mathscr{H}$,

$$(3.18) \quad \left| \langle \theta, h \rangle - \frac{1}{\lambda} (\exp(\lambda \langle \theta, h \rangle) - 1) \right| \leq \lambda D^2 \exp(\lambda D) / 2.$$

By (3.18) and the triangle inequality,

$$(3.19) \quad \begin{aligned} &\mu \left| \int \langle \theta, h \rangle p_\theta d(\omega - \pi) \right| \\ &\leq \lambda D^2 \exp(\lambda D) \\ &+ \frac{1}{\lambda} [\|P_\omega - P_\pi\| + \mu \left| \int \exp(\lambda \langle \theta, h \rangle) p_\theta d(\omega - \pi) \right|]. \end{aligned}$$

We now show

$$(3.20) \quad \mu \left| \int \exp(\lambda \langle \theta, h \rangle) p_\theta d(\omega - \pi) \right| = P_{\lambda h} |p_\omega - p_\pi|$$

as a consequence of

$$(3.21) \quad \mu \left(\int \exp(\lambda \langle \theta, h \rangle) p_\theta d\omega, \int \exp(\lambda \langle \theta, h \rangle) p_\theta d\pi \right)^{-1} = P_{\lambda h} (p_\omega, p_\pi)^{-1}.$$

By (2.1), linearity of inner product and the map in (A), we get, $\forall m \geq 1$ and $\forall (\theta_1, \dots, \theta_m) \in \Theta^m$,

$$\mu \left(\left\{ \lambda \langle \theta_i, h \rangle + l_{\theta_i} - \frac{\|\theta_i\|^2}{2} \right\}_{i=1}^{i=m} \right)^{-1} = P_{\lambda h} \left(\left\{ l_{\theta_i} - \frac{\|\theta_i\|^2}{2} \right\}_{i=1}^{i=m} \right)^{-1},$$

or, equivalently,

$$\mu \left(\left\{ \exp(\lambda \langle \theta_i, h \rangle) p_{\theta_i} \right\}_{i=1}^{i=m} \right)^{-1} = P_{\lambda h} \left(\left\{ p_{\theta_i} \right\}_{i=1}^{i=m} \right)^{-1}.$$

Hence, if ω and π are finitely supported, (3.21) holds. Since, by Theorem II.6.3 of Parthasarathy (1967), Ω has a dense subset consisting of finitely supported measures, to prove (3.21) for arbitrary ω and π it suffices to show that if $\nu_k \rightarrow \nu$ in Ω , $\int \exp(\lambda \langle \theta, h \rangle) p_\theta d\nu_k(\theta)[p_{\nu_k}]$ goes to $\int \exp(\lambda \langle \theta, h \rangle) p_\theta d\nu(\theta)[p_\nu]$ in $L_2(\mu)[L_2(P_{\lambda h})]$. Actually we show the continuity of the map taking (ν, ν') to the $L_2(\mu)[L_2(P_{\lambda h})]$ inner product of $\int \exp(\lambda \langle \theta, h \rangle) p_\theta d\nu(\theta)[p_\nu]$ and $\int \exp(\lambda \langle \eta, h \rangle) p_\eta d\nu'(\eta)[p_{\nu'}]$, by interchanging the order of integration on \mathcal{X} and Θ^2 , using Lemma 3.3 (with $k = 2, a_1 = a_2 = 1$) to evaluate the μ -integral which is continuous (by continuities of vector addition and inner product and the exponential function) and bounded [via (3.1)] on Θ^2 , and applying Lemma III.1.1 of Parthasarathy (1967). The second [bracket] assertion is proved by a similar argument.

Combining (3.19) and (3.20), we get

$$(3.22) \quad [l.h.s. (3.19)] \leq \lambda D^2 \exp(\lambda D) + \frac{1}{\lambda} \|P_\omega - P_\pi\| + \frac{1}{\lambda} \mu(|p_\omega - p_\pi| p_{\lambda h}).$$

By partitioning \mathcal{X} into $[p_{\lambda h} > c]$ and $[p_{\lambda h} \leq c]$, and applying the Cauchy-Schwarz inequality in $L_2(\mu)$,

$$(3.23) \quad \mu(|p_\omega - p_\pi| p_{\lambda h}) \leq c \|P_\omega - P_\pi\| + \|p_\omega - p_\pi\|_2 \{ \mu p_{\lambda h}^2 [p_{\lambda h} > c] \}^{1/2}.$$

Weakening the bound (3.22) by enlarging its r.h.s. using (3.23) and substituting the resulting bound in the r.h.s. of (3.17), we obtain (via some rear-

range of terms) that, given $\varepsilon > 0$, $\exists \delta > 0$ such that

$$\begin{aligned}
P_\theta \|\tau_\omega - \tau_\pi\| &\leq 2D \exp\left(\frac{D^2}{2}\right) \left\{ \Pr\left[Z \geq \frac{a}{D}\right] + \Pr\left[Z \leq \frac{b}{D}\right] \right\}^{1/2} \\
&\quad + \exp\left(\frac{D^2}{2} + a - b\right) [3D + I\lambda^{-1}(1+c)] \|P_\omega - P_\pi\| \\
&\quad + 2\varepsilon \exp\left(\frac{D^2}{2} + a - b\right) \\
(3.24) \quad &\quad + \exp\left(\frac{D^2}{2} + a - b\right) I\lambda D^2 \exp(\lambda D) \\
&\quad + \lambda^{-1} \exp\left(\frac{D^2}{2} + a - b\right) \|p_\omega - p_\pi\|_2 \\
&\quad \times \sum_{i=1}^I \left\{ \mu p_{\lambda h_i}^2 [p_{\lambda h_i} > c] \right\}^{1/2} \\
&:= T_1 + T_2 + T_3 + T_4 + T_5,
\end{aligned}$$

where $a > 0$, $b < 0$, $\lambda > 0$ and $c > 0$ are arbitrary, d_w -balls of radius δ around $\{h_1, \dots, h_I\}$ cover \mathscr{H} and $Z \sim \mathcal{N}(0, 1)$.

Now choose a and b so that $T_1 < \gamma$. Then choose ε so that $\exp(D^2/2 + a - b) < \gamma/\varepsilon$, implying $T_3 < 2\gamma$. Now choose λ so that $\lambda D^2 \exp(\lambda D) < \varepsilon/I$, implying $T_4 < \gamma$. Then choose c so that, uniformly in ω , π and h ($\in \mathscr{H}$), $(1/\lambda) \|p_\omega - p_\pi\|_2 \{ \mu p_{\lambda h}^2 [p_{\lambda h} > c] \}^{1/2} \leq \varepsilon/I$ [which is possible since, by Lemma 3.4 and the triangle inequality in $L_2(\mu)$, $\|p_\omega - p_\pi\|_2 \leq 2 \exp(D^2/2)$ and, by the existence of uniformly bounded higher moments, the family $\{p_{\lambda h}^2: h \in \mathscr{H}\}$ is uniformly μ -integrable]. That makes $T_5 < \gamma$, completing the proof with $\mathscr{H} = [3D + I\lambda^{-1}(1+c)] \exp(D^2/2 + a - b)$. \square

Remark 3.1 (Robustness of the Bayes estimator). Since $\|P_\omega - P_\pi\|$ is a metric distance between ω and π for the topology of weak convergence on Ω [Majumdar (1993), Remark 3.1], Theorem 3.1 establishes equi- (in $\theta \in \Theta$) uniform continuity of $\omega \mapsto \tau_\omega$ ($\in H$) in $L_1(P_\theta)$, which can be interpreted as a (frequentist) qualitative robustness property of the component Bayes estimator wrt the prior specification.

Combining Proposition 3.1 and Theorem 3.1,

$$|D_n(\mathbf{t}_\Lambda, \boldsymbol{\theta})| \leq 4D(5\gamma) + 4D\mathscr{N}n^{-1} \sum_{\alpha=1}^n \mathbb{P}_{\boldsymbol{\theta}} \|P_{\beta_{\Lambda, \alpha, n}} - P_{G_n}\|.$$

Since the bound in the display above holds for arbitrary $\gamma > 0$, to establish asymptotic optimality of \mathbf{t}_Λ it suffices to show

$$(3.25) \quad \bigvee_{\alpha=1}^n \mathbb{P}_\theta \|P_{\beta_{\Lambda, \alpha, n}} - P_{G_n}\| \rightarrow 0 \quad \text{uniformly in } \theta, \text{ as } n \rightarrow \infty.$$

Theorem 3.2 formulates sufficient conditions for (3.25) and, hence, for asymptotic optimality of \mathbf{t}_Λ .

For a finite measure m on the Borel σ -field of a separable metric space \mathcal{S} , by topological support of m we mean the smallest closed subset of \mathcal{S} carrying the total mass; let S_m denote the topological support of m .

THEOREM 3.2 (Asymptotic optimality). *If Θ is strongly compact and $S_\Lambda = \Omega$, then \mathbf{t}_Λ is asymptotically optimal.*

PROOF. As already observed, it suffices to establish (3.25) even if Θ is only strongly totally bounded. Fix $\theta \in \Theta^n$ and $1 \leq \alpha \leq n$. Let $G_{n\alpha}$ denote the empirical distribution based on θ_α . Then [see Datta (1991), Lemma 4.3],

$$\mathbb{P}_\theta \|P_{\beta_{\Lambda, \alpha, n}} - P_{G_{n\alpha}}\|^{-1} = \mathbb{P}_{\theta_\alpha} \|P_{\beta_{\Lambda, n, n}} - P_{G_{n-1}}\|^{-1}.$$

Since θ and α are arbitrary, from the above we get

$$(3.26) \quad \bigvee_{\theta \in \Theta^n} \bigvee_{\alpha=1}^n \mathbb{P}_\theta \|P_{\beta_{\Lambda, \alpha, n}} - P_{G_{n\alpha}}\| \leq \bigvee_{\theta \in \Theta^{n-1}} \mathbb{P}_\theta \|P_{\beta_{\Lambda, n, n}} - P_{G_{n-1}}\| \rightarrow 0 \quad \text{as } n \rightarrow \infty,$$

by Theorem 3.1 of Majumdar (1993).

Next, note that $G_n - G_{n\alpha} = n^{-1}(\delta_{\theta_\alpha} - G_{n\alpha})$. Therefore, $\|P_{G_n} - P_{G_{n\alpha}}\| = n^{-1} \|P_{\theta_\alpha} - P_{G_{n\alpha}}\| \leq 2n^{-1}$ and the proof now ends by the triangle inequality and subadditivity of supremum. \square

REMARK 3.2. The reduction of asymptotic optimality of \mathbf{t}_Λ to (3.25) uses neither compactness of Θ (total boundedness suffices) nor $S_\Lambda = \Omega$. The proof of Theorem 3.2, which involves reduction of (3.25) to Theorem 3.1 of Majumdar (1993) via (3.26), uses no assumption either. However, Theorem 3.1 of Majumdar (1993) assumes that Θ is strongly compact and $S_\Lambda = \Omega$.

REMARK 3.3 (Asymptotic equivalence of the equivariant envelope and the simple envelope). If the component problem involves a compact (in total variation norm) class of mutually absolutely continuous probability measures, then the excess of the simple envelope over the equivariant envelope goes to zero uniformly in the measures [Mashayekhi (1990), Remark 4]. By assumption the measures $\{P_\theta: \theta \in \Theta\}$ are mutually absolutely continuous. Le Cam [(1986), page 158] shows that $\|P_\theta - P_\eta\| < \sqrt{2} \|\theta - \eta\|$, implying continuity of $\theta \mapsto P_\theta$. Therefore, if Θ is compact, $\{P_\theta: \theta \in \Theta\}$ is compact and the asymptotic equivalence is established.

3.2. *Admissibility.* The argument we use to prove admissibility of compound Bayes estimators is fairly standard in decision theory: a unique (up to equivalence) Bayes rule in a mutually absolutely continuous family is admissible [Ferguson (1967), Theorem 1, Section 2.3].

Let ξ be a prior on the compound parameter θ . Let Q denote the joint distribution $\xi \circ P_\theta$ on (\mathbf{x}, θ) . Note that $n^{-1} \sum_{\alpha=1}^n Q \|t_\alpha - \theta_\alpha\|^2$, the Bayes (versus ξ) risk of a compound estimator \mathbf{t} , is minimal iff $Q \|t_\alpha - \theta_\alpha\|^2$ is minimal for every α . Now $Q \|t_\alpha - \theta_\alpha\|^2$ can be represented as $\int P_\theta (\int \|t_\alpha - \theta_\alpha\|^2 dP_{\theta_\alpha} d\xi_\alpha) d\xi$, where ξ_α is the conditional distribution (under ξ) of θ_α given the rest of the θ 's. Since $\int \|t_\alpha - \theta_\alpha\|^2 dP_{\theta_\alpha} d\xi_\alpha$ has, by Lemma 3.1, a unique minimizer, a compound estimator Bayes versus ξ is unique and, hence, admissible.

4. Compound Bayes decisions for the Gaussian shift parameter.

The asymptotic optimality and admissibility results for \mathbf{t}_Λ obtained in the previous section are extended to include other loss (risk) functions. We restrict Θ to be a strongly compact subset of H . Using Theorem 4.1 [which is Theorem 3 of Zhu (1992), based on Theorem 1 of Mashayekhi (1993)], Theorem 4.2 establishes asymptotic optimality of \mathbf{t}_Λ when the component risk $R(t, \theta)$ is an equi- (in t) uniformly continuous and bounded function of θ and $S_\Lambda = \Omega$.

THEOREM 4.1. *Let Θ be compact metric and $\{P_\theta: \theta \in \Theta\}$ be compact in the total variation norm. Suppose that $\omega \mapsto P_\omega$ is one-to-one and $\theta \mapsto R(t, \theta)$ is equi- (in t) uniformly continuous and bounded. Then $\hat{\mathbf{t}}$ defined by*

$$(4.1) \quad \hat{t}_\alpha(\mathbf{x}) := \tau_{\hat{\omega}(\mathbf{x}_\alpha)}(x_\alpha),$$

where $\hat{\omega}$ is a symmetric mapping from \mathcal{X}^{n-1} into Ω , is asymptotically optimal if

$$(4.2) \quad \bigvee_{\theta \in \Theta^{n-1}} P_\theta \|P_{\hat{\omega}} - P_{G_n}\| \rightarrow 0 \quad \text{as } n \rightarrow \infty.$$

Note that, with $\hat{\omega}(\mathbf{x}_\alpha) = \beta_{\Lambda_{\alpha,n}}$, $\hat{\mathbf{t}}$ of (4.1) becomes \mathbf{t}_Λ of (1.4). Hence, to establish asymptotic optimality of \mathbf{t}_Λ under Theorem 4.1 assumption on the risk function, it suffices to verify (the other assumptions) that $\{P_\theta: \theta \in \Theta\}$ is compact in the total variation norm (Remark 3.3), $\omega \mapsto P_\omega$ is one-to-one [Majumdar (1993), Lemma 3.4] and (4.2) holds with $\hat{\omega}(\mathbf{x}_\alpha) = \beta_{\Lambda_{\alpha,n}}$. With $\hat{\omega}(\mathbf{x}_\alpha) = \beta_{\Lambda_{\alpha,n}}$, (4.2) is equivalent to the convergence of the r.h.s. of (3.26) to zero, which is the assertion of Majumdar [(1993), Theorem 3.1] for full support Λ . Summarizing the discussion of this paragraph, we have the following theorem.

THEOREM 4.2 (Asymptotic optimality). *If Θ is strongly compact and $S_\Lambda = \Omega$, then \mathbf{t}_Λ is asymptotically optimal if the component risk function $\theta \mapsto R(t, \theta)$ is equi- (in t) uniformly continuous and bounded.*

REMARK 4.1 (Admissibility of \mathbf{t}_Λ). Under Theorem 4.2 assumption on the component risk function, by Theorem 1 of Zhu (1992), the compound risk is a continuous function of $\boldsymbol{\theta}$. Hence the admissibility of every compound rule Bayes versus a full support compound prior [Ferguson (1967), Theorem 3, Section 2.3]. By Theorem 2 of Zhu (1992), $\beta_{\Lambda, n}$ has full support if Λ has. Therefore, \mathbf{t}_Λ is admissible under the assumptions of Theorem 4.2.

Acknowledgments. This paper extends some of the results of my Ph.D. dissertation, written at Michigan State University under the (invaluable) supervision of Professor James Hannan. The preparation of this manuscript was completed during the summer of 1993 in the hospitable environment of the Theoretical Statistics and Mathematics Division of the Indian Statistical Institute at Calcutta. An anonymous reviewer provided insightful comments leading to a more succinct presentation.

REFERENCES

- DATTA, S. (1988). Asymptotically optimal Bayes compound and empirical Bayes estimators in exponential families with compact parameter space. Ph.D. dissertation, Dept. Statistics and Probability, Michigan State Univ.
- DATTA, S. (1991). Asymptotic optimality of Bayes compound estimators in compact exponential families. *Ann. Statist.* **19** 354–365.
- DUDLEY, R. M. (1989). *Real Analysis and Probability*. Wadsworth & Brooks/Cole, Belmont, CA.
- FERGUSON, T. S. (1967). *Mathematical Statistics, A Decision Theoretic Approach*. Academic Press, New York.
- GILLILAND, D. C. and HANNAN, J. (1986). The finite state compound decision problem, equivariance and restricted risk components. In *Adaptive Statistical Procedures and Related Topics* (J. Van Ryzin, ed.) 129–145. IMS, Hayward, CA.
- GILLILAND, D. C., HANNAN, J. and HUANG, J. S. (1976). Asymptotic solutions to the two state component compound decision problem, Bayes versus diffuse priors on proportions. *Ann. Statist.* **4** 1101–1112.
- HANNAN, J. F. and ROBBINS, H. (1955). Asymptotic solutions of the compound decision problem for two completely specified distributions. *Ann. Math. Statist.* **36** 1743–1752.
- IBRAGIMOV, I. A. and KHAS'MINSKII, R. Z. (1991). Asymptotically normal families of distributions and efficient estimation. *Ann. Statist.* **19** 1681–1724.
- LE CAM, L. M. (1986). *Asymptotic Methods in Statistical Decision Theory*. Springer, New York.
- LE CAM, L. M. and YANG, G. L. (1990). *Asymptotics in Statistics, Some Basic Concepts*. Springer, New York.
- MAJUMDAR, S. (1993). Posterior consistency in Gaussian shift experiments. Technical Report 247, Dept. Statistics and Applied Probability, Univ. California, Santa Barbara.
- MAJUMDAR, S. (1994). Bayes compound and empirical Bayes estimation of the mean of a Gaussian distribution on a Hilbert space. *J. Multivariate Anal.* **48** 87–106.
- MAJUMDAR, S., GILLILAND, D. and HANNAN, J. (1993). Bounds for robust ML and posterior consistency in compound mixture state experiments. Unpublished manuscript.
- MASHAYEKHI, M. (1990). Stability of symmetrized probabilities and compact compound equivariant decisions. Ph.D. dissertation, Dept. Statistics and Probability, Michigan State Univ.
- MASHAYEKHI, M. (1993). On equivariance and the compound decision problem. *Ann. Statist.* **21** 736–745.
- MILLAR, P. W. (1983). *The Minimax Principle in Asymptotic Statistical Theory. Lecture Notes in Math.* **976**. Springer, New York.

- PARTHASARATHY, K. R. (1967). *Probability Measures on Metric Spaces*. Academic Press, New York.
- ROBBINS, H. (1951). Asymptotically subminimax solutions of compound statistical decision problems. *Proc. Second Berkeley Symp. Math. Statist. Probab.* 131–148. Univ. California Press, Berkeley.
- RUDIN, W. (1973). *Functional Analysis*. McGraw-Hill, New York.
- RUDIN, W. (1987). *Real and Complex Analysis*. McGraw-Hill, Singapore.
- STRASSER, H. (1985). *Mathematical Theory of Statistics: Statistical Experiments and Asymptotic Decision Theory*. de Gruyter, Berlin.
- ZHU, J. (1992). Asymptotic behavior of compound rules in compact families. Unpublished manuscript.

DEPARTMENT OF STATISTICS
UNIVERSITY OF CONNECTICUT
U-120
STORRS, CONNECTICUT 06269