ANNALES
DE L'INSTITUT
HENRI
POINCARÉ
PROBABILITÉS
ET STATISTIQUES

Evaluating default priors with a generalization of Eaton's Markov chain

Brian P. Shea and Galin L. Jones

School of Statistics, University of Minnesota, Minneapolis, MN, 55455, USA. E-mail: bshea@stat.umn.edu; galin@umn.edu

Abstract. We consider evaluating improper priors in a formal Bayes setting according to the consequences of their use. Let Φ be a class of functions on the parameter space and consider estimating elements of Φ under quadratic loss. If the formal Bayes estimator of every function in Φ is admissible, then the prior is strongly admissible with respect to Φ. Eaton's method for establishing strong admissibility is based on studying the stability properties of a particular Markov chain associated with the inferential setting. In previous work, this was handled differently depending upon whether $\varphi \in \Phi$ was bounded or unbounded. We consider a new Markov chain which allows us to unify and generalize existing approaches while simultaneously broadening the scope of their potential applicability. We use our general theory to investigate strong admissibility conditions for location models when the prior is Lebesgue measure and for the p-dimensional multivariate Normal distribution with unknown mean vector θ and a prior of the form $\nu(\|\theta\|^2)\,d\theta$.

Résumé. Nous considérons l'évaluation de lois a priori impropres dans un cadre Bayésien formel en fonction des conséquences de leur utilisation. Soit Φ une classe de fonctions sur l'espace des paramètres, que l'on cherche à estimer sous une fonction de perte quadratique. Si l'estimateur Bayésien de toute fonction dans Φ est admissible, alors la loi a priori est fortement admissible par rapport à Φ. La méthode d'Eaton pour établir l'admissibilité forte est basée sur l'étude des propriétés de stabilité d'une certaine chaîne de Markov associé au cadre inférentiel. Dans des travaux précédents, nous considérions une nouvelle chaîne de Markov qui permet d'unifier et de généraliser les approches existantes tout en élargissant simultanément son champ d'application. Nous utilisons cette théorie générale pour étudier des conditions d'admissibilité forte pour des modèles à paramètre de position, une loi a priori donnée par la mesure de Lebesgue et la loi normale multivariée de dimension p et moyenne θ, et une loi a priori de la forme $\nu(\|\theta\|^2)\,d\theta$.

1. Introduction

Suppose we are in a parametric setting, and we are considering use of an improper prior measure that yields a proper posterior distribution. Such priors arise in the absence of honest prior belief about parameter values and are typically derived from structural arguments based on the likelihood or the parameter space [18]. Thus, an improper prior, rather than being a statement of beliefs specific to a situation, is a default. Such priors, proposed from likelihood or invariance arguments, require evaluation, just as estimators proposed from likelihood or invariance arguments require evaluation, and an attractive avenue is to evaluate the prior according to the consequences of its use. That is, we can evaluate the prior by examining properties of the resulting posterior inferences. The criterion we use to judge posterior inferences is known as *strong admissibility*. This concept was introduced by Eaton [7] and has given rise to a substantial theory [7–15,19]. Our goal in the rest of this section is to convey the basic idea behind strong admissibility and the way

it is studied here. We also summarize our main results without delving too far into the details, which are dealt with carefully later.

Suppose the sample space $\mathcal{X}$ is a Polish space with Borel σ-algebra $\mathcal{B}$ and the parameter space $\boldsymbol{\vartheta}$ is a Polish space with Borel σ-algebra $\mathcal{C}$. Let $\{P(\cdot|\theta), \theta \in \boldsymbol{\vartheta}\}$ be a family of sampling distributions where we assume that for each $B \in \mathcal{B}$, $P(B|\cdot)$ is $\mathcal{C}$-measurable and for each $\theta \in \boldsymbol{\vartheta}$, $P(\cdot|\theta)$ is a probability measure on $(\mathcal{X}, \mathcal{B})$. Let ν be a σ-finite measure on the parameter space $\boldsymbol{\vartheta}$ with $\nu(\boldsymbol{\vartheta}) = \infty$. Throughout the marginal on $\mathcal{X}$

$$M(\mathrm{d}x) := \int_{\boldsymbol{\vartheta}} P(\mathrm{d}x|\theta)\nu(\mathrm{d}\theta) \tag{1}$$

is assumed to be σ-finite. In this case, the disintegration

$$Q(\mathrm{d}\theta|x)M(\mathrm{d}x) = P(\mathrm{d}x|\theta)\nu(\mathrm{d}\theta) \tag{2}$$

generalizes Bayes theorem and implicitly defines formal posterior distributions $Q(\cdot|x)$ on the parameter space. Note that for each $x \in \mathcal{X}$, $Q(\cdot|x)$ is a probability measure on $(\boldsymbol{\vartheta}, \mathcal{C})$ and for each $C \in \mathcal{C}$, $Q(C|\cdot)$ is $\mathcal{B}$-measurable. Taraldsen and Lindqvist [23] provide a recent, accessible introduction to the existence of formal posterior distributions while one can consult Eaton [6,7] and Johnson [17] for more details and references.

Suppose $\varphi : \boldsymbol{\vartheta} \to \mathbb{R}^p$ for $p \geq 1$, and consider estimating $\varphi(\theta)$. The formal Bayes estimator of φ under squared error loss is the posterior mean

$$\hat{\varphi}(x) := \int_{\boldsymbol{\vartheta}} \varphi(\theta) Q(\mathrm{d}\theta|x). \tag{3}$$

Let $\| \cdot \|$ denote the usual Euclidean norm. If $\delta(x)$ is any estimator of $\varphi(\theta)$, the risk function of δ is

$$\mathsf{R}(\delta; \theta) := \int_{\mathcal{X}} \|\varphi(\theta) - \delta(x)\|^2 P(\mathrm{d}x|\theta). \tag{4}$$

The estimator δ is *almost-ν admissible* if for any other estimator δ_1 such that $\mathsf{R}(\theta; \delta_1) \leq \mathsf{R}(\theta; \delta)$ for all $\theta \in \boldsymbol{\vartheta}$, then the set $\{\theta : \mathsf{R}(\theta; \delta_1) < \mathsf{R}(\theta; \delta)\}$ has ν-measure 0.

Since we will use admissibility to judge the prior, our interests are more ambitious than establishing admissibility of a single estimator. Let Φ be a class of functions defined on the parameter space. If the formal Bayes estimator of every $\varphi \in \Phi$ is almost-ν admissible, then we say the prior (equivalently the posterior) is *strongly admissible with respect to Φ*. A prior is strongly admissible if it is robust against risk dominance within the class Φ. Since a default prior will undergo repeated use, it is important for the range of appropriate uses to be clearly defined and desirable that the range be as large as possible. We can then endorse the improper prior insofar as it avoids unreasonable actions in a variety of such problems.

Previous work on strong admissibility focused on the case where Φ consisted of a single unbounded function [1,2, 9], or all bounded functions [7,8,10–14,19]. In either case, strong admissibility was established by verifying sufficient conditions for the admissibility of an estimator established by Brown [4] or via Markov chain arguments using an approach introduced by Eaton [7,9]. We study the latter method.

Eaton's method for establishing almost-ν admissibility of formal Bayes estimators is based on the recurrence properties of a Markov chain associated with the inferential setting; the relevant notion of recurrence is defined in the next section. However, different Markov chains were required depending upon whether φ was bounded [7] or unbounded [9]. We introduce and study a new Markov chain which allows us to unify and generalize these existing approaches while simultaneously broadening the scope of their potential applicability. The expected posterior

$$R(\mathrm{d}\theta|\eta) = \int_{\mathcal{X}} Q(\mathrm{d}\theta|x) P(\mathrm{d}x|\eta) \tag{5}$$

is a Markov kernel on $\boldsymbol{\vartheta}$. We study transformations of R, which are now described. Let $f : \boldsymbol{\vartheta} \times \boldsymbol{\vartheta} \to \mathbb{R}^+$ satisfy $f(\theta, \eta) = f(\eta, \theta)$ and set

$$\mathcal{T}(\eta) = \int_{\boldsymbol{\vartheta}} f(\theta, \eta) R(\mathrm{d}\theta|\eta).$$

If $0 < \mathcal{T}(\eta) < \infty$ for all $\eta \in \vartheta$, then

$$T(\mathrm{d}\theta|\eta) = \frac{f(\theta, \eta)}{\mathcal{T}(\eta)} R(\mathrm{d}\theta|\eta)$$

is a Markov kernel on ϑ. Recurrence of the Markov chain associated with T implies the almost-ν admissibility of formal Bayes estimators with respect to a large class of functions. Define

$$\Phi_f := \big\{\varphi \colon \big\|\varphi(\theta) - \varphi(\eta)\big\|^2 \le M_\varphi f(\theta, \eta) \text{ some } 0 < M_\varphi < \infty\big\}. \tag{6}$$

We prove that if the Markov chain defined by T is recurrent, then the formal Bayes estimator of every function in Φ_f is almost-ν admissible, and we say the prior ν is *strongly admissible with respect to* Φ_f. The following example illustrates this technique.

Example 1. *Let X be a p-dimensional Normal random variable with identity covariance I_p and unknown mean θ. Let p-dimensional Lebesgue measure $\nu(\mathrm{d}\theta) = \mathrm{d}\theta$ be our improper prior. The proper posterior for θ is a Normal with mean x – the observed value of the random variable X – and covariance I_p. The kernel*

$$R(\mathrm{d}\theta|\eta) = \int_{\mathcal{X}} Q(\mathrm{d}\theta|x) P(\mathrm{d}x|\eta) = (4\pi)^{-p/2} \exp\left(-\frac{1}{4}\|\theta - \eta\|^2\right) \mathrm{d}\theta$$

describes a $N(\eta, 2I_p)$ random variable. Let d be an arbitrary positive constant and note that $f(\theta, \eta) = \|\theta - \eta\|^2 + d$ is symmetric and uniformly bounded away from 0. Further

$$\mathcal{T}(\eta) = \int f(\theta, \eta) R(\mathrm{d}\theta|\eta) = \int \big(\|\theta - \eta\|^2 + d\big) R(\mathrm{d}\theta|\eta) = 2p + d.$$

Thus,

$$T(\mathrm{d}\theta|\eta) = \frac{\|\theta - \eta\|^2 + d}{2p + d} R(\mathrm{d}\theta|\eta)$$

and the chain with kernel T is a random walk on $\mathbb{R}^p$. Note that

$$\int \|\theta - \eta\|^p T(\mathrm{d}\theta|\eta) < \infty$$

since a Normal distribution has moments of all orders and hence for $p = 1$ or $p = 2$, the chain is recurrent [5,21]. We conclude that for $p = 1$ or $p = 2$ Lebesgue measure is strongly admissible. That is, the formal Bayes estimators of all functions φ satisfying

$$\big\|\varphi(\theta) - \varphi(\eta)\big\| \le M_\varphi\big(\|\theta - \eta\|^2 + d\big) \quad \text{for some } 0 < M_\varphi < \infty$$

are almost-ν admissible. Since d is arbitrary this includes all bounded functions as well as many unbounded functions.

Lebesgue measure is not strongly admissible with respect to Φ_f when $p \ge 3$ since the James–Stein estimator dominates the formal Bayes estimator of $\varphi(\theta) = \theta$. In Sections 4.1 and 4.2 we extend this example in two ways; in Section 4.1 we consider Lebesgue measure as a prior for general location models when $p = 1$ or $p = 2$ while in Section 4.2 we consider the Normal means problem with an alternative prior when $p \ge 3$. We will return to this example below.

Recurrence of the Markov chain defined by T implies more than we have so far claimed. Suppose the function $u : \vartheta \to \mathbb{R}^+$ is bounded away from zero and infinity so that $1/c < u(\theta) < c$ for some constant $c > 1$ and every $\theta \in \vartheta$. The measure defined by

$$\nu_u(\mathrm{d}\theta) := u(\theta)\nu(\mathrm{d}\theta) \tag{7}$$

is a *bounded perturbation* of ν. Let $\mathcal{F}_\nu$ be the family of all bounded perturbations of ν. Observe that $\nu \in \mathcal{F}_\nu$ and that the other elements of $\mathcal{F}_\nu$ are measures with tail behavior similar to ν. Eaton [7] showed that recurrence of the chain with kernel R implied the formal Bayes estimators of any bounded function is almost-ν admissible for every prior in $\mathcal{F}_\nu$. We extend this result and show that recurrence under T is sufficient for the strong admissibility with respect to Φ_f of every element of $\mathcal{F}_\nu$. In this case, we say the family $\mathcal{F}_\nu$ is *strongly admissible with respect to Φ_f*.

Example 2. *Recall the setting of Example* 1. *Let $\mathcal{F}_\nu$ be the bounded perturbations of p-dimensional Lebesgue measure $\nu(\mathrm{d}\theta) = \mathrm{d}\theta$. That is, elements of $\mathcal{F}_\nu$ are measures of the form $\nu_u(\mathrm{d}\theta) = u(\theta)\,\mathrm{d}\theta$ where $1/c < u(\theta) < c$ for some constant $c > 1$. Recall that Φ_f is the class of all bounded functions and all functions satisfying*

$$\left\| \varphi(\theta) - \varphi(\eta) \right\| \leq M_\varphi \left(\| \theta - \eta \|^2 + d \right) \quad \text{for some } 0 < M_\varphi < \infty.$$

The recurrence of the Markov chain governed by T when $p = 1$ or $p = 2$ implies the formal Bayes estimators of every $\varphi \in \Phi_f$ are almost-ν_u admissible for every prior $\nu_u \in \mathcal{F}_\nu$. That is, $\mathcal{F}_\nu$ is strongly admissible with respect to Φ_f.

Our main results generalize existing work [7,9] and in fact unify the analysis for bounded and unbounded functions. We show that the Markov kernel R can be transformed to define many Markov chains, any one of which might be used to demonstrate strong admissibility, thus greatly broadening the scope of potential applications. We are never concerned narrowly with a single admissibility problem but broadly with species of problems. Moreover, solving a single representative problem, which representative we are free to elect, solves any problem within a bounded rate of change – whether of the function to be estimated or the prior used to estimate it.

The remainder is organized as follows. Section 2 gives some background on recurrence for general state space Markov chains. Section 3 presents new results for establishing strong admissibility via Markov chain arguments. In Section 4 we investigate strong admissibility conditions for location models when the prior is Lebesgue measure and for the p-dimensional multivariate Normal distribution with unknown mean vector θ and a prior of the form $\nu(\|\theta\|^2)\,\mathrm{d}\theta$. Section 5 contains a few concluding remarks. Finally, our proofs and many technical details are deferred to the appendices.

2. Recurrence of Markov chains

The goal of this section is to introduce a general notion of recurrence for Markov chains. Let $\mathcal{W}$ be a Polish space and denote the Borel σ-algebra by $\mathcal{D}$. Let $K : \mathcal{D} \times \mathcal{W} \to [0, 1]$. Then K is a *Markov transition kernel* on the measurable space $(\mathcal{W}, \mathcal{D})$ if $K(D|\cdot)$ is a nonnegative measurable function for every $D \in \mathcal{D}$ and $K(\cdot|w)$ is a probability measure for every $w \in \mathcal{W}$.

The kernel K determines a time-homogeneous Markov chain $W = \{W_0, W_1, W_2, \ldots\}$ on the product space $\mathcal{W}^\infty$ which is equipped with the product σ-algebra $\mathcal{D}^\infty$. Note that conditional on $W_n = w$, the law of W_{n+1} is $K(\cdot|w)$. Given $W_0 = w$, let $\Pr(\cdot|w)$ be the law of W on $\mathcal{W}^\infty$.

Suppose $D \in \mathcal{D}$. The random variable

$$\tau_D = \begin{cases} \infty & \text{if } W_n \notin D \text{ for all } n \geq 1, \text{ or} \\ \text{the smallest } n \geq 1 \text{ such that } W_n \in D & \text{otherwise} \end{cases}$$

is a stopping time for D, and $E_D = \{\tau_D < \infty\}$ is the set of paths that encounter D after initialization. Let ξ be a non-trivial, σ-finite measure on $(\mathcal{W}, \mathcal{D})$ and recall that a set is ξ-*proper* if its measure under ξ is positive and finite.

Definition 2.1. *A ξ-proper set D is locally ξ-recurrent if $\Pr(E_D|w) = 1$ for all but a ξ-null set of initial values in D. Call the Markov chain W locally ξ-recurrent if every ξ-proper set is locally ξ-recurrent.*

This notion of recurrence is more general than that often encountered in general state space Markov chain theory [20], but is appropriate since the chains we will consider in the next section may not be irreducible [11,15]. We consider a method for establishing local recurrence in Section 3.2.

The connection between the above general Markov chain theory and the notion of strong admissibility relies heavily on the special structure of symmetric Markov chains. Let ξ be a non-trivial, σ-finite measure. Then the kernel K is ξ-symmetric if

$$\omega(\mathrm{d}w_1, \mathrm{d}w_2) := K(\mathrm{d}w_1|w_2)\xi(\mathrm{d}w_2)$$

is a symmetric measure on $(\mathcal{W} \times \mathcal{W}, \mathcal{D} \times \mathcal{D})$.

Throughout the remainder we restrict attention to symmetric Markov kernels. Eaton [8,10] provides some background on the theory of symmetric Markov chains underlying strong admissibility.

3. Strong admissibility via Markov chains

In a ground-breaking paper Eaton [7] connected the local recurrence of a Markov chain and the almost-ν admissibility of the formal Bayes estimator for bounded φ. Later, Eaton [9] showed that the local recurrence of a different Markov chain was required to establish almost-ν admissibility when φ is unbounded. Since our work builds on them, these foundational definitions and results are stated carefully here. The expected posterior distribution at the parameter value η, $R(\mathrm{d}\theta|\eta)$ defined in (5), is a ν-symmetric Markov kernel on ϑ since it satisfies the detailed balance condition

$$R(\mathrm{d}\theta|\eta)\nu(\mathrm{d}\eta) = R(\mathrm{d}\eta|\theta)\nu(\mathrm{d}\theta), \tag{8}$$

which follows from (2) and (5). Call R an *Eaton kernel*–analogous kernels exist for the sample space [13] and the product of the sample and parameter spaces [11]. Eaton [7] established the following basic result.

Theorem 3.1. *If the Markov chain with kernel R is locally-ν recurrent, then the formal Bayes estimator of every bounded function is almost-ν admissible.*

Eaton [7] also showed that the result holds for bounded perturbations of ν. Theorem 3.1 has found substantial application [7,10–14,19].

Now suppose we want to use Markov chains to study formal Bayes estimators of unbounded functions on ϑ. We need a basic assumption on the risk function R, defined at (4), to ensure existence of the integrated risk difference which is studied in the appendices.

Assumption 3.1. *Suppose there exist sets $K_1 \subseteq K_2 \subseteq \cdots$ with $\bigcup K_i = \vartheta$. Also assume that for all i, $0 < \nu(K_i) < \infty$ and*

$$\int_{K_i} \mathsf{R}(\hat{\varphi}; \theta)\nu(\mathrm{d}\theta) < \infty.$$

Let $\varphi: \vartheta \to \mathbb{R}^p$ be measurable and unbounded and set

$$\mathcal{S}(\eta) = \int_{\vartheta} \|\varphi(\theta) - \varphi(\eta)\|^2 R(\mathrm{d}\theta|\eta)$$

so that if $0 < \mathcal{S}(\eta) < \infty$ for all η, then

$$S(\mathrm{d}\theta|\eta) = \frac{\|\varphi(\theta) - \varphi(\eta)\|^2}{\mathcal{S}(\eta)} R(\mathrm{d}\theta|\eta) \tag{9}$$

is a Markov kernel defining a Markov chain on ϑ. Moreover S satisfies detailed balance with respect to $\mathcal{S}(\eta)\nu(\mathrm{d}\eta)$:

$$S(\mathrm{d}\theta|\eta)\mathcal{S}(\eta)\nu(\mathrm{d}\eta) = S(\mathrm{d}\eta|\theta)\mathcal{S}(\theta)\nu(\mathrm{d}\theta).$$

Eaton [9] connected the local recurrence of S with the almost-ν admissibility of the formal Bayes estimator of $\varphi(\theta)$.

Theorem 3.2. *Suppose $0 < \mathcal{S}(\eta) < \infty$ for each $\eta \in \boldsymbol{\vartheta}$ and there exist ν-proper sets $K_1 \subseteq K_2 \subseteq \cdots$ such that $\bigcup K_i = \boldsymbol{\vartheta}$, and for each i*

$$\int_{K_i} \mathcal{S}(\eta)\nu(\mathrm{d}\eta) < \infty.$$

If the Markov chain with kernel S is locally-ν recurrent, then the formal Bayes estimator of φ is almost-ν admissible.

In the next section, we unify and generalize these results. We show that by analyzing an appropriate Markov chain we can recover the conclusions of both theorems and, in fact, achieve something stronger. Moreover, we broaden the class of Markov chains that can be studied to obtain strong admissibility results.

3.1. *A new Markov chain connection*

For any $\eta \in \boldsymbol{\vartheta}$, let $\psi(\cdot|\eta)$ be a nontrivial σ-finite measure on $(\boldsymbol{\vartheta}, \mathcal{C})$ such that $\psi(\cdot|\eta)$ is absolutely continuous with respect to $R(\cdot|\eta)$. For any element C of the Borel sets $\mathcal{C}$, let $\psi(C|\cdot)$ be a nonnegative measurable function. Let $f(\cdot, \eta)$ be a Radon–Nikodym derivative of $\psi(\cdot|\eta)$ with respect to $R(\cdot|\eta)$:

$$\psi(C|\eta) = \int_C f(\theta, \eta)R(\mathrm{d}\theta|\eta) \quad \text{for all } (\eta, C) \in (\boldsymbol{\vartheta}, \mathcal{C}). \tag{10}$$

Define

$$T(C|\eta) = \frac{\psi(C|\eta)}{\psi(\boldsymbol{\vartheta}|\eta)} \quad \text{for all } (\eta, C) \in (\boldsymbol{\vartheta}, \mathcal{C}). \tag{11}$$

Let $\mathcal{T}(\eta) = \psi(\boldsymbol{\vartheta}|\eta)$ for all $\eta \in \boldsymbol{\vartheta}$. We will make the following basic assumptions on $\mathcal{T}$.

Assumption 3.2. *For all $\eta \in \boldsymbol{\vartheta}$ we have $0 < \mathcal{T}(\eta) < \infty$ and there exist ν-proper sets $K_1 \subseteq K_2 \subseteq \cdots$ such that $\bigcup K_i = \boldsymbol{\vartheta}$, and for each i*

$$\int_{K_i} \mathcal{T}(\eta)\nu(\mathrm{d}\eta) < \infty.$$

The first part of the assumption ensures that the kernel T is well-defined. If $\mu(\mathrm{d}\eta) = \mathcal{T}(\eta)\nu(\mathrm{d}\eta)$, then the second part of the assumption implies that μ is σ-finite. In Proposition B.1 we establish that

$$T(\mathrm{d}\theta|\eta)\mu(\mathrm{d}\eta) = f(\theta, \eta)R(\mathrm{d}\theta|\eta)\nu(\mathrm{d}\eta).$$

If $f(\theta, \eta) = f(\eta, \theta)$ for all η and θ in $\boldsymbol{\vartheta}$, then by using (8) it is easy to see that T is symmetric with respect to μ.

Note that $\mathcal{T}(\eta)$ is almost-ν uniformly bounded away from 0 if there exists $\varepsilon > 0$ such that $\mathcal{T}(\eta) \geq \varepsilon$ except possibly on a set of ν-measure 0. We are now in a position to state the main result of this section. The proof is given in Appendix B. Recall the definition of Φ_f from (6).

Theorem 3.3. *Let $\mathcal{T}(\eta)$ be almost-ν uniformly bounded away from zero and suppose that for all η and θ in $\boldsymbol{\vartheta}$*

$$f(\theta, \eta) = f(\eta, \theta). \tag{12}$$

If the Markov chain having kernel $T(\mathrm{d}\theta|\eta)$ is locally ν-recurrent, then ν is strongly admissible with respect to Φ_f.

The following extension of Theorem 3.3 is also proved in Appendix B.

Theorem 3.4. *Assume the conditions of Theorem 3.3. Then every bounded perturbation ν_u is strongly admissible, that is, the family $\mathcal{F}_\nu$ is strongly admissible with respect to Φ_f.*

If we take $f(\theta, \eta) = d$ for some $d > 0$, then we completely recover the results of Theorem 3.1 while if $f(\theta, \eta) = \|\varphi(\theta) - \varphi(\eta)\|^2$, then we extend the results of Theorem 3.2. If, as in Examples 1 and 2, we set $f(\theta, \eta) = \|\varphi(\theta) - \varphi(\eta)\|^2 + d$ with $d > 0$, then we obtain results stronger than if we had established local recurrence of each of the chains associated with the kernels R and S and relied on Theorems 3.1 and 3.2. Moreover, since the analyst has the freedom to choose an appropriate f, this result extends the range of potential applicability of Eaton's method.

3.2. *Reducing dimension*

The Markov kernel T naturally takes the same dimension as the parameter space. This dimension may be quite large, making the required analysis difficult. In this section we prove that the conclusions of Theorem 3.4 (hence Theorem 3.3) hold if we can establish the local recurrence of a particular Markov chain which lives on $[0, \infty)$.

Denote the Borel subsets of $[0, \infty)$ by $\mathcal{A}$. A measurable mapping t from $(\boldsymbol{\vartheta}, \mathcal{C})$ to $([0, \infty), \mathcal{A})$ induces a measure $\tilde{\nu}$ given by

$$\tilde{\nu}(A) = \nu\big(t^{-1}(A)\big), \quad A \in \mathcal{A}.$$

We will need the next assumption throughout the remainder of this section.

Assumption 3.3. *There exists a partition $\{A_i\}$ of $[0, \infty)$ such that each A_i is measurable and each $C_i = t^{-1}(A_i)$ is ν-proper.*

Eaton et al. [12] showed that under Assumption 3.3 there exists a Markov transition function $\pi(\mathrm{d}\theta | \beta)$ on $\mathcal{C} \times [0, \infty)$ such that

$$\nu(\mathrm{d}\theta) = \pi(\mathrm{d}\theta | \beta)\tilde{\nu}(\mathrm{d}\beta) \tag{13}$$

which means for all measurable nonnegative functions f_1 on $\boldsymbol{\vartheta}$ and f_2 on $[0, \infty)$

$$\int_{\boldsymbol{\vartheta}} f_2\big(t(\theta)\big) f_1(\theta)\nu(\mathrm{d}\theta) = \int_0^\infty f_2(\beta)\left(\int_{\boldsymbol{\vartheta}} f_1(\theta)\pi(\mathrm{d}\theta | \beta)\right)\tilde{\nu}(\mathrm{d}\beta).$$

Define

$$\tilde{P}(\mathrm{d}x | \beta) = \int_{\boldsymbol{\vartheta}} P(\mathrm{d}x | \theta)\pi(\mathrm{d}\theta | \beta), \quad \beta \in [0, \infty). \tag{14}$$

The conditional probabilities $\tilde{P}(\cdot | \beta)$ form a parametric family indexed by β, and $\tilde{\nu}(\mathrm{d}\beta)$ is a σ-finite prior. Eaton et al. [12] also showed that the marginal measure on $\mathcal{X}$ is the same as at (1). That is,

$$\int_0^\infty \tilde{P}(\mathrm{d}x | \beta)\tilde{\nu}(\mathrm{d}\beta) = M(\mathrm{d}x),$$

which is assumed σ-finite. Thus, there is a Markov kernel $\tilde{Q}(\mathrm{d}\beta | x)$ satisfying

$$\tilde{Q}(\mathrm{d}\beta | x)M(\mathrm{d}x) = \tilde{P}(\mathrm{d}x | \beta)\tilde{\nu}(\mathrm{d}\beta). \tag{15}$$

In fact, a version of the posterior [12] is

$$\tilde{Q}(A | x) := Q\big(t^{-1}(A) | x\big) \quad \text{for all } A \in \mathcal{A}. \tag{16}$$

The expected posterior

$$\tilde{R}(\mathrm{d}\beta | \alpha) = \int \tilde{Q}(\mathrm{d}\beta | x)\tilde{P}(\mathrm{d}x | \alpha) \tag{17}$$

is a $\tilde{v}$-symmetric Eaton kernel.

If η and θ are elements of $\boldsymbol{\vartheta}$, let $\alpha = t(\eta)$ and $\beta = t(\theta)$. For any $\alpha \in [0, \infty)$, let $\tilde{\psi}(\cdot|\alpha)$ be a nontrivial σ-finite measure on $([0, \infty), \mathcal{A})$ such that $\tilde{\psi}(\cdot|\alpha)$ is absolutely continuous with respect to $\tilde{R}(\cdot|\alpha)$. Let $\tilde{f}(\cdot, \alpha)$ be a Radon–Nikodym derivative of $\tilde{\psi}(\cdot|\alpha)$ with respect to $\tilde{R}(\cdot|\alpha)$ – i.e., for all $(\alpha, A) \in ([0, \infty), \mathcal{A})$

$$\tilde{\psi}(A|\alpha) = \int_A \tilde{f}(\beta, \alpha) \tilde{R}(\mathrm{d}\beta|\alpha).$$

Define $\tilde{\mathcal{T}}(\alpha) = \tilde{\psi}([0, \infty)|\alpha)$ and set

$$\tilde{T}(A|\alpha) = \frac{\tilde{\psi}(A|\alpha)}{\tilde{\mathcal{T}}(\alpha)}, \quad (\alpha, A) \in ([0, \infty), \mathcal{A})$$

and let $\tilde{\mu}(\mathrm{d}\alpha) = \tilde{\mathcal{T}}(\alpha)\tilde{v}(\mathrm{d}\alpha)$. The following assumption ensures that $\tilde{T}$ is a well-defined kernel and that $\tilde{\mu}$ is σ-finite.

Assumption 3.4. *For all α we have $0 < \tilde{\mathcal{T}}(\alpha) < \infty$ and there exist $\tilde{v}$-proper sets $K_1 \subseteq K_2 \subseteq \cdots$ such that $\bigcup K_i = [0, \infty)$, and for each i*

$$\int_{K_i} \tilde{\mathcal{T}}(\alpha)\tilde{v}(\mathrm{d}\alpha) < \infty.$$

In Appendix C we show that $\tilde{T}$ is $\tilde{\mu}$-symmetric. The following theorem shows that we can analyze the recurrence properties of the chain defined by $\tilde{T}$ to achieve the conclusions of Theorem 3.4 (hence Theorem 3.3). The proof is given in Appendix C. Recall that $\mathcal{F}_v$ is the family of bounded perturbations of v.

Theorem 3.5. *Let $\tilde{\mathcal{T}}(\cdot)$ be almost-$\tilde{v}$ uniformly bounded away from 0 and suppose that for all α and β in $[0, \infty)$*

$$\tilde{f}(\beta, \alpha) = \tilde{f}(\alpha, \beta).$$

If the Markov chain with kernel $\tilde{T}$ is locally $\tilde{v}$-recurrent, then $\mathcal{F}_v$ is strongly admissible with respect to Φ_f.

Since $\tilde{T}$ lives on $[0, \infty)$ it would be convenient to have conditions which guarantee the local recurrence of a Markov chain on $[0, \infty)$. This is discussed in the following section.

3.2.1. *Recurrence of Markov chains on $[0, \infty)$*

To this point we have said little about establishing local recurrence. The following theorem presents one method for doing so and is a distillation of several existing results [10,12]. It applies generally to Markov chains on $[0, \infty)$ and hence the notation in this section is consistent with that of Section 2.

Let $\mathcal{W} = [0, \infty)$ and $\mathcal{D}$ be the Borel σ-algebra. Let $K : \mathcal{D} \times \mathcal{W} \to [0, 1]$ be a Markov kernel which defines a time-homogeneous Markov chain $W = \{W_0, W_1, W_2, \ldots\}$ on $\mathcal{W}^\infty$. Define the kth moment of K about its current state as

$$m_k(v) = \int_0^\infty (y - v)^k K(\mathrm{d}y|v). \tag{18}$$

Theorem 3.6. *Assume for each positive integer n there exists $\delta(n) < 1$ such that*

$$\sup_{v \in [0,n]} K([0, n]|v) \le \delta(n). \tag{19}$$

Suppose

$$\lim_{v \to \infty} \frac{\log v}{v} \frac{m_3(v)}{m_2(v)} = 0 \tag{20}$$

and there exists a function ϕ and an integer n_0 such that for $v \in [n_0, \infty)$

$$m_1(v) \leq \frac{m_2(v)}{2v}\big[1 + \phi(v)\big] \quad and \quad \lim_{v \to \infty} \phi(v) \log v = 0. \tag{21}$$

If ξ is a non-trivial, σ-finite measure, $0 < \xi([0, n_0)) < \infty$ and K is ξ-symmetric, then the Markov chain W is locally ξ-recurrent.

We will use Theorem 3.6 in conjunction with Theorem 3.5 in one of our applications in Section 4.2.

4. Applications

We turn our attention to two applications of the general theory developed above. Both applications can be viewed as extensions of the setting in Example 1. In Section 4.1 we establish strong admissibility of the bounded perturbations of Lebesgue measure in 1 or 2 dimensions for general location models. Then in Section 4.2 we establish strong admissibility in 3 or more dimensions of the bounded perturbations of a family of priors for the multivariate normal mean.

4.1. *Admissible priors for location models*

Let dx be Lebesgue measure on $\mathbb{R}^p$, $p \geq 1$ and λ be a σ-finite measure on a Polish space $\mathcal{Y}$. Suppose g is a density with respect to $dx\lambda(dy)$ on $\mathbb{R}^p \times \mathcal{Y}$ and let the model be

$$P(dx, dy|\theta) = g(x - \theta, y) \, dx\lambda(dy).$$

We will take the prior on $\boldsymbol{\vartheta} = \mathbb{R}^p$ to be Lebesgue measure, that is, $\nu(d\theta) = d\theta$. Let

$$m(y) = \int g(x, y) \, dx$$

and q_0 be a density on $\mathbb{R}^p$ such that $E_{q_0} Y^2 < \infty$. A version of the formal posterior is then given by $Q(d\theta|x, y) = q(\theta|x, y) \, d\theta$ with

$$q(\theta|x, y) = \begin{cases} \frac{g(x-\theta, y)}{m(y)} & 0 < m(y) < \infty, \\ q_0(x - \theta) & \text{otherwise.} \end{cases}$$

Eaton [9] uses a Markov chain argument based on Theorem 3.2 to prove that the Bayes estimator of $\varphi(\theta) = \theta$ is admissible when $p = 1$ or $p = 2$, a result originally proved using Blyth's method [16,22]. We use Theorem 3.4 to add to these results. Let $d > 0$ be arbitrary and define

$$f(\theta, \eta) = \|\theta - \eta\|^2 + d, \quad \eta \in \mathbb{R}^p$$

so that

$$\Phi_f = \big\{\varphi \colon \big\|\varphi(\theta) - \varphi(\eta)\big\| \leq M_\varphi f(\theta, \eta) \text{ some } 0 < M_\varphi < \infty\big\}.$$

Theorem 4.1. *Assume that for $p = 1$ or $p = 2$*

$$\int \|x\|^{2+p} g(x, y) \, dx\lambda(dy) < \infty.$$

Then $\mathcal{F}_\nu$, the family of bounded perturbations of Lebesgue measure, is strongly admissible with respect to Φ_f.

The proof, given in Appendix D, is based on the properties of random walks – recall Example 1 – so it is not surprising that it fails in higher dimensions since nontrivial random walks on $\mathbb{R}^p$, $p \geq 3$ are not recurrent.

4.2. Admissible priors for the multivariate normal mean

Let $X \sim N_p(\theta, I_p)$ with $p \geq 1$. Consider the family of σ-finite measures on $\vartheta = \mathbb{R}^p$ described by

$$\left(\frac{1}{a + \|\theta\|^2}\right)^b d\theta, \quad a \geq 0, b \geq 0. \tag{22}$$

For $a = 0$, the prior is improper for all $b > 0$, but the induced marginal distributions on $\mathcal{X}$ are only σ-finite for $b < p/2$. For $a > 0$, the family yields improper prior distributions when $b \leq p/2$ and proper prior distributions when $b > p/2$. In fact, if $b = (a + p)/2$, the prior is the kernel of a multivariate t distribution with a degrees of freedom. When $a > 0$ and $b = 0$, this is p-dimensional Lebesgue measure which was considered in Example 1.

Now suppose $p \geq 3$. Berger, Strawderman and Tan [2] established that the formal Bayes estimator of θ is admissible when $a = 1$ and $b = (p - 1)/2$ while Eaton et al. [12] use Theorem 3.1 to prove that if $a > 0$ and $b \in [p/2 - 1, p/2]$, then the formal Bayes estimator of every bounded function is almost admissible. We use Theorems 3.5 and 3.6 to add to these results. Let $d > 0$ be arbitrary and define

$$f(\theta, \eta) = \|\theta - \eta\|^2 + d, \quad \eta \in \mathbb{R}^p$$

so that

$$\Phi_f = \left\{\varphi \colon \|\varphi(\theta) - \varphi(\eta)\| \leq M_\varphi f(\theta, \eta) \text{ some } 0 < M_\varphi < \infty\right\}.$$

The proof of the following theorem is given in Appendix E.

Theorem 4.2. *For $p \geq 3$ let $X \sim N_p(\theta, I_p)$ and with $a > 0$ set*

$$\nu(d\theta) = \left(\frac{1}{a + \|\theta\|^2}\right)^{p/2} d\theta.$$

Then $\mathcal{F}_\nu$, the family of bounded perturbations of ν, is strongly admissible with respect to Φ_f.

5. Concluding remarks

Given the results in Section 4 and some of the developments in Stein estimation (see e.g. [3]) it would be natural to seek to extend our work and study strong admissibility of priors for multivariate exponential families and spherically symmetric distributions. Indeed, Wen-Lin Lai [19] used Eaton's method to study strong admissibility with respect to bounded functions of priors for a multivariate Poisson model. However, to our knowledge, there have been no applications of Eaton's method to the situation where the model is spherically symmetric.

Appendix A: Preliminaries

We begin by stating some existing results concerning local recurrence and introduce Blyth's method. This material plays a fundamental role in our proofs of Theorems 3.3, 3.4 and 3.5.

A.1. Local recurrence

The purpose here is to give two characterizations of local recurrence for general symmetric Markov chains, hence the notation is consistent with that of Sections 2 and 3.2.1.

Let $\mathcal{W}$ be a Polish space and denote the Borel σ-algebra by $\mathcal{D}$. Let $K \colon \mathcal{D} \times \mathcal{W} \to [0, 1]$ be a Markov kernel on $(\mathcal{W}, \mathcal{D})$. Let ξ be a non-trivial σ-finite measure and recall that a set is ξ-proper if its measure under ξ is positive and finite. Throughout this section K is assumed to be ξ-symmetric.

Theorem A.1 (Eaton [10]). *The Markov chain W is locally ξ-recurrent if and only if there exists a sequence of ξ-proper sets increasing to the state space such that each is locally ξ-recurrent.*

Let $L^2(\xi)$ be the space of ξ-square integrable functions. Then the quantity

$$\Delta(h; K, \xi) = \frac{1}{2} \iint \big(h(\theta) - h(\eta)\big)^2 K(\mathrm{d}\theta|\eta)\xi(\mathrm{d}\eta), \quad h \in L^2(\xi) \tag{23}$$

is called a *Dirichlet form*. Also, if D is a ξ-proper set, define

$$\mathcal{H}_\xi(D) := \big\{h \in L^2(\xi): h \geq I_D\big\},$$

where I_D is the indicator function of the set D. A characterization of local ξ-recurrence in terms of Δ is given by the following result.

Theorem A.2 (Eaton [9]). *A set $D \in \mathcal{D}$ is locally ξ-recurrent if and only if*

$$\inf_{h \in \mathcal{H}_\xi(D)} \Delta(h; K, \xi) = 0.$$

A.2. Blyth's method

Consider the posterior distributions obtained from perturbations of the prior measure ν. Let $g : \vartheta \to \mathbb{R}^+$ be such that the perturbation

$$\nu_g(\mathrm{d}\theta) = g(\theta)\nu(\mathrm{d}\theta) \tag{24}$$

is σ-finite. Assume

$$\hat{g}(x) := \int g(\theta) Q(\mathrm{d}\theta|x) < \infty. \tag{25}$$

Letting

$$M_g(\mathrm{d}x) := \int P(\mathrm{d}x|\theta)g(\theta)\nu(\mathrm{d}\theta)$$

it is easy to see that $M_g(\mathrm{d}x) = \hat{g}(x)M(\mathrm{d}x)$ is σ-finite and that the posterior obtained from the perturbed prior is

$$Q_g(\mathrm{d}\theta|x) = \frac{g(\theta)}{\hat{g}(x)} Q(\mathrm{d}\theta|x). \tag{26}$$

Let $L^1(\nu)$ be the set of all ν-integrable functions and define

$$\mathcal{G}_\nu = \big\{g \in L^1(\nu): g \geq 0, g\text{bounded and } \nu_g(\vartheta) > 0\big\}. \tag{27}$$

Let I_C be the indicator function of the set C. Of particular interest are the subfamilies

$$\mathcal{G}_\nu(C) = \{g \in \mathcal{G}_\nu: g \geq I_C\}.$$

Let $\hat{\varphi}$ be the Bayes estimator of φ under the prior ν, and let $\hat{\varphi}_g$ be the Bayes estimator of φ under the prior ν_g. Also recall the definition of the risk function R at (4). A key quantity in connecting Markov chains to admissibility is the *integrated risk difference* at $g \in \mathcal{G}_\nu$ with respect to ν against $\hat{\varphi}$:

$$\mathrm{IRD}_\nu(g; \hat{\varphi}) := \int \big[\mathsf{R}(\theta; \hat{\varphi}) - \mathsf{R}(\theta; \hat{\varphi}_g)\big]\nu_g(\mathrm{d}\theta). \tag{28}$$

Notice that IRD_ν is nonnegative and the integrability assumptions on R in Assumption 3.1 ensure that it is well-defined. The connection of IRD_ν with almost-ν admissibility is given by Blyth's method.

Theorem A.3 (Blyth's method). *Let $\hat{\varphi}$ be an estimator and ν a σ-finite measure on the parameter space. If $\inf_{g \in \mathcal{G}_\nu(C)} \mathrm{IRD}_\nu(g; \hat{\varphi}) = 0$ for every $C \subseteq \boldsymbol{\vartheta}$ such that $0 < \nu(C) < \infty$, then $\hat{\varphi}$ is almost-ν admissible.*

Appendix B: Proof of Theorem 3.3

We will develop a connection between the kernel T defined in (11) and the IRD_ν which will be key to proving Theorem 3.3 via Theorems A.2 and A.3. We begin with some preliminary results before we prove Theorem 3.3.

B.1. Preliminary results

Recall $\mathcal{T}(\eta) = \psi(\boldsymbol{\vartheta}|\eta)$ where ψ is defined at (10). Define

$$\mu(C) = \int_C \mathcal{T}(\eta)\nu(\mathrm{d}\eta) \quad \text{for all } C \in \mathcal{C}. \tag{29}$$

Proposition B.1. *Suppose Assumption 3.2 and condition (12) hold. Then*

(a) *the measure μ is σ-finite and equivalent to ν,*
(b) *$\mathcal{T}$ is a Radon–Nikodym derivative of μ with respect to ν,*
(c) *T is a μ-symmetric Markov transition kernel,*
(d) *$T(\mathrm{d}\theta|\eta)\mu(\mathrm{d}\eta) = f(\theta, \eta)R(\mathrm{d}\theta|\eta)\nu(\mathrm{d}\eta)$,*
(e) *if μ' is proportional to μ, T is a μ'-symmetric Markov transition kernel.*

Proof. That μ is σ-finite follows easily from Assumption 3.2. Since $\mathcal{T}$ is a nonnegative measurable function, for any ν-null set C,

$$\mu(C) = \int_C \mathcal{T}(\eta)\nu(\mathrm{d}\eta) = 0.$$

Furthermore, since $\mathcal{T}(\eta) > 0$ for all η in the parameter space, every ν-positive set is a μ-positive set. Thus, if C is μ-null, it is also ν-null. Therefore, μ and ν are equivalent measures. Furthermore, $\mathcal{T}$ is a Radon–Nikodym derivative of μ with respect to ν since μ is absolutely continuous with respect to ν, and by assumption $\mathcal{T}$ is a nonnegative measurable function such that for any measurable set C, $\mu(C)$ is given by Eq. (29).

At any point η in the parameter space, since $\psi(\cdot|\eta)$ is a nontrivial finite measure, normalizing by $\mathcal{T}(\eta) = \psi(\boldsymbol{\vartheta}|\eta)$ produces a probability measure. Recall that $\psi(C|\cdot)$ is a nonnegative measurable function where C is an element of the Borel sets $\mathcal{C}$. Since $\boldsymbol{\vartheta}$ is a Borel measurable set, $\mathcal{T}$ is a Borel measurable function. Also by hypothesis, $\mathcal{T}(\eta)$ is positive and finite for all η. The reciprocal function is continuous, hence Borel measurable, on the positive real numbers, and the composition $1/\mathcal{T}$ of Borel measurable functions is a Borel measurable function. Thus, the product of Borel measurable functions $1/\mathcal{T}$ and $\psi(C|\cdot)$ is a Borel measurable function for any $C \in \mathcal{C}$.

Let A and B be measurable sets. Since

$$
\begin{aligned}
S(A, B) &= \iint I_A(\eta) I_B(\theta) T(\mathrm{d}\theta|\eta)\mu(\mathrm{d}\eta) \\
&= \iint I_A(\eta) I_B(\theta) f(\theta, \eta) R(\mathrm{d}\theta|\eta)\nu(\mathrm{d}\eta) \quad \text{by (10), (11), and (29)} \\
&= \iint I_B(\theta) I_A(\eta) f(\eta, \theta) R(\mathrm{d}\eta|\theta)\nu(\mathrm{d}\theta) \quad \text{by (8) and (12)} \\
&= \iint I_B(\theta) I_A(\eta) T(\mathrm{d}\eta|\theta)\mu(\mathrm{d}\theta) \quad \text{by (10), (11), and (29)} \\
&= S(B, A),
\end{aligned}
$$

the kernel T is μ-symmetric. Note that the substitution

$$T(d\theta|\eta)\mu(d\eta) = f(\theta, \eta)R(d\theta|\eta)\nu(d\eta) \tag{30}$$

holds as a consequence.

Let $\mu' = c\mu$ for some $c > 0$. Then by (8) and (30)

$$\begin{aligned}
T(d\theta|\eta)\mu'(d\eta) &= cT(d\theta|\eta)\mu(d\eta) \\
&= cf(\theta, \eta)R(d\theta|\eta)\nu(d\eta) \\
&= cf(\eta, \theta)R(d\eta|\theta)\nu(d\theta) \\
&= cT(d\eta|\theta)\mu(d\theta) \\
&= T(d\eta|\theta)\mu'(d\theta)
\end{aligned}$$

which implies T is μ'-symmetric. $\qquad\square$

We can now develop a connection between the Markov kernel T and the integrated risk difference IRD_ν. Our argument will require the following known result; recall the definition of $\mathcal{G}_\nu$ from (27).

Proposition B.2 (Eaton [9]). *If $g \in \mathcal{G}_\nu$ and R is Eaton's kernel, then*

$$\mathrm{IRD}_\nu(g; \hat{\varphi}) \le \iint \|\varphi(\theta) - \varphi(\eta)\|^2 \big(\sqrt{g}(\theta) - \sqrt{g}(\eta)\big)^2 R(d\theta|\eta)\nu(d\eta).$$

By Proposition B.1 we see that T is μ-symmetric so that the relevant Dirichlet form, recall (23), is

$$\Delta(h; T, \mu) = \frac{1}{2} \iint \big(h(\eta) - h(\theta)\big)^2 T(d\theta|\eta)\mu(d\eta), \quad h \in L^2(\mu).$$

Proposition B.3. *Suppose Assumption* 3.2 *and condition* (12) *hold. If $\varphi \in \Phi_f$ and $g \in \mathcal{G}_\nu$, then there is a measure μ_φ which is proportional to μ such that*

$$\mathrm{IRD}_\nu(g; \hat{\varphi}) \le 2\Delta(\sqrt{g}; T, \mu_\varphi).$$

Proof. Since $\varphi \in \Phi_f$ there exists $0 < M_\varphi < \infty$ such that

$$\|\varphi(\theta) - \varphi(\eta)\|^2 \le M_\varphi f(\theta, \eta) \quad \text{for all } \theta, \eta.$$

For $C \in \mathcal{C}$ define $\mu_\varphi(C) = M_\varphi \mu(C)$ and suppose $g \in \mathcal{G}_\nu$. In the following the first inequality is from Proposition B.2 and the second is obtained from Proposition B.1(d) and that $\varphi \in \Phi_f$

$$\begin{aligned}
\mathrm{IRD}_\nu(g; \hat{\varphi}) &\le \iint \|\varphi(\theta) - \varphi(\eta)\|^2 \big(\sqrt{g}(\theta) - \sqrt{g}(\eta)\big)^2 R(d\theta|\eta)\nu(d\eta) \\
&\le \iint \big(\sqrt{g}(\theta) - \sqrt{g}(\eta)\big)^2 T(d\theta|\eta) M_\varphi \mu(d\eta) \\
&= 2\Delta(\sqrt{g}; T, \mu_\varphi). \qquad\square
\end{aligned}$$

B.2. Proof of Theorem 3.3

Let the measures $\psi(\cdot|\eta)$ and μ and the transition kernel T be as defined at equations (10), (29), and (11), respectively. Also, suppose $\varphi \in \Phi_f$. By Proposition B.3 if $g \in \mathcal{G}_\nu$, then

$$\mathrm{IRD}_\nu(g; \hat{\varphi}) \le 2\Delta(\sqrt{g}; T, \mu_\varphi),$$

where $\mu_\varphi = M_\varphi \mu$. By Proposition B.1 we have that μ and ν are equivalent measures and T is a μ_φ-symmetric Markov kernel. Since the measures μ_φ and ν are equivalent and the chain with kernel T is locally ν-recurrent, it is also locally μ_φ-recurrent – that is, every μ_φ-proper set C is locally μ_φ-recurrent. Thus, by Theorem A.2,

$$\inf_{h \in \mathcal{H}_{\mu_\varphi}(C)} \Delta(h; T, \mu_\varphi) = 0,$$

where $\mathcal{H}_{\mu_\varphi}(C)$ collects the square-integrable dominators of I_C.

Now let $h \in \mathcal{H}_{\mu_\varphi}(C)$ and recall that by assumption $\mathcal{T}$ is a measurable function uniformly bounded away from zero ν-almost everywhere. Hence there exists some $\varepsilon > 0$ such that

$$\varepsilon \int h^2(\eta) \nu(d\eta) \le \int h^2(\eta) \mathcal{T}(\eta) \nu(d\eta)$$

$$= \int h^2(\eta) \mu(d\eta)$$

$$= \frac{1}{M_\varphi} \int h^2(\eta) \mu_\varphi(d\eta)$$

$$< \infty.$$

Hence $h \in L^2(\nu)$ and we conclude that $\mathcal{H}_{\mu_\varphi}(C) \subseteq \mathcal{G}_\nu(C)$. Moreover, if $\sqrt{h} \in \mathcal{H}_{\mu_\varphi}(C)$, then $h \in \mathcal{G}_\nu(C)$. Thus we obtain

$$\inf_{g \in \mathcal{G}_\nu(C)} \text{IRD}_\nu(g; \hat\varphi) = 0.$$

Therefore, $\hat\varphi$ is an almost-ν admissible estimator by Theorem A.3.

B.3. *Proof of Theorem 3.4*

Recall that $\nu_u \in \mathcal{F}_\nu$ is a bounded perturbation of ν. The sampling and posterior distributions define the ν_u-symmetric Eaton kernel

$$R_u(d\theta | \eta) = \int_{\mathcal{X}} Q_u(d\theta | x) P(dx | \eta),$$

where the perturbed posterior was defined at (26). The mean of φ with respect to $Q_u(\cdot | x)$ is the formal Bayes estimator of $\varphi(\theta)$ under squared error loss. We denote it $\hat\varphi_u$ to emphasize its dependence on the perturbed prior.

Since $\nu_u \in \mathcal{F}_\nu$, there exists $0 < c < \infty$ such that $\nu_u(d\theta) = u(\theta) \nu(d\theta)$ and $1/c < u(\theta) < c$. Recall the definition of μ_φ from the proof of Proposition B.3. In the following the first inequality is from Proposition B.2, the second follows by noting that u and $1/\hat{u}$ are both bounded above by c while the third is obtained from Proposition B.1(d) and that $\varphi \in \Phi_f$

$$\text{IRD}_{\nu_u}(g; \hat\varphi_u) \le \iint \left\| \varphi(\theta) - \varphi(\eta) \right\|^2 \left(\sqrt{g}(\theta) - \sqrt{g}(\eta) \right)^2 R_u(d\theta | \eta) \nu_u(d\eta)$$

$$\le c^3 \iint \left\| \varphi(\theta) - \varphi(\eta) \right\|^2 \left(\sqrt{g}(\theta) - \sqrt{g}(\eta) \right)^2 R(d\theta | \eta) \nu(d\eta)$$

$$\le c^3 \iint \left(\sqrt{g}(\theta) - \sqrt{g}(\eta) \right)^2 T(d\theta | \eta) M_\varphi \mu(d\eta)$$

$$= 2c^3 \Delta(\sqrt{g}; T, \mu_\varphi).$$

The remainder of the proof follows the proof of Theorem 3.3 exactly.

Appendix C: Proof of Theorem 3.5

Before proving Theorem 3.5 we require an analogue of Proposition B.1. Given a nonnegative function $\tilde{h}$ on $[0, \infty)$, define

$$h(\theta) = \tilde{h}\big(t(\theta)\big). \tag{31}$$

Proposition C.1. *Suppose Assumptions 3.3 and 3.4 hold. If $\tilde{f}(\beta, \alpha) = \tilde{f}(\alpha, \beta)$, then*

(a) $\tilde{\mu}$ *is σ-finite and equivalent to $\tilde{v}$,*
(b) $\tilde{\mathcal{T}}$ *is a Radon–Nikodym derivative of $\tilde{\mu}$ with respect to $\tilde{v}$,*
(c) $\tilde{T}$ *is a $\tilde{\mu}$-symmetric Markov transition kernel,*
(d) $\tilde{T}(d\beta|\alpha)\tilde{\mu}(d\alpha) = \tilde{f}(\beta, \alpha)\tilde{R}(d\beta|\alpha)\tilde{v}(d\alpha)$,
(e) $\tilde{h} \in L^2(\tilde{\mu})$ *implies h is in $L^2(\mu)$,*
(f) $\Delta(\tilde{h}; \tilde{T}, \tilde{\mu}) = \Delta(h; T, \mu)$, *and*
(g) $\mathcal{T}(\eta)$ *is almost-v bounded away from zero.*

Proof. The proof of the first 3 assertions follows exactly the proof of the first 3 assertions in Proposition B.1 with $[0, \infty)$, $\mathcal{A}$, α, β, $\tilde{v}$, and $\tilde{R}$ substituted for ϑ, $\mathcal{C}$, η, θ, v, and R, respectively. The substitution

$$\tilde{T}(d\beta|\alpha)\tilde{\mu}(d\alpha) = \tilde{f}(\beta, \alpha)\tilde{R}(d\beta|\alpha)\tilde{v}(d\alpha) \tag{32}$$

follows as a consequence. If $\tilde{h} \in L^2(\tilde{\mu})$, then $h \in L^2(\mu)$ since

$$
\begin{aligned}
\int \tilde{h}^2(\alpha)\tilde{\mu}(d\alpha) &= \int \tilde{h}^2(\alpha)\tilde{\mathcal{T}}(\alpha)\tilde{v}(d\alpha) \\
&= \int \tilde{h}^2(\alpha) \int \tilde{f}(\beta, \alpha) \int \tilde{Q}(d\beta|x)\tilde{P}(dx|\alpha)\tilde{v}(d\alpha) \\
&= \int h^2(\eta) \int f\big(t(\theta), t(\eta)\big) \int Q(d\theta|x)P(dx|\eta)v(d\eta) \\
&= \int h^2(\eta)\mathcal{T}(\eta)v(d\eta) \\
&= \int h^2(\eta)\mu(d\eta).
\end{aligned}
$$

Since

$$
\begin{aligned}
\iint \big(\tilde{h}(\beta) &- \tilde{h}(\alpha)\big)^2 \tilde{T}(d\beta|\alpha)\tilde{\mu}(d\alpha) \\
&= \iint \big(\tilde{h}(\beta) - \tilde{h}(\alpha)\big)^2 \tilde{f}(\beta, \alpha)\tilde{R}(d\beta|\alpha)\tilde{v}(d\alpha) \quad \text{by (32)} \\
&= \iint \big(\tilde{h}(\beta) - \tilde{h}(\alpha)\big)^2 \tilde{f}(\beta, \alpha) \int \tilde{Q}(d\beta|x)\tilde{P}(dx|\alpha)\tilde{v}(d\alpha) \quad \text{by (17)} \\
&= \iint \big(h(\theta) - h(\eta)\big)^2 f\big(t(\theta), t(\eta)\big) \int Q(d\theta|x)P(dx|\eta)v(d\eta) \quad \text{by (13), (14), (16), and (31)} \\
&= \iint \big(h(\theta) - h(\eta)\big)^2 f\big(t(\theta), t(\eta)\big)R(d\theta|\eta)v(d\eta) \quad \text{by (5)} \\
&= \iint \big(h(\theta) - h(\eta)\big)^2 T(d\theta|\eta)\mu(d\eta) \quad \text{by (30),}
\end{aligned}
$$

it follows that $\Delta(\tilde{h}; \tilde{T}, \tilde{\mu}) = \Delta(h; T, \mu)$. Finally, recall that $\tilde{T}$ is almost-$\tilde{\nu}$ bounded away from zero by some positive constant c. Note that for any $\tilde{\nu}$-proper set A,

$$\int_{t^{-1}(A)} T(\eta)\nu(\mathrm{d}\eta) = \int_A \tilde{T}(\alpha)\tilde{\nu}(\mathrm{d}\alpha) \geq c\tilde{\nu}(A) = c\nu(t^{-1}(A)).$$

It follows that T is almost-ν bounded away from zero. □

We are now ready to prove Theorem 3.5.

C.1. *Proof of Theorem* 3.5

By Proposition C.1, the measures $\tilde{\mu}$ and $\tilde{\nu}$ are equivalent, and the kernel $\tilde{T}$ is $\tilde{\mu}$-symmetric. Thus, the chain with kernel $\tilde{T}$ is locally $\tilde{\mu}$-recurrent, and by Theorem A.2

$$\inf_{\tilde{h} \in \mathcal{H}_{\tilde{\mu}}(A)} \Delta(\tilde{h}; \tilde{T}, \tilde{\mu}) = 0$$

for any $\tilde{\mu}$-proper set A.

Let (A_i) be a sequence of $\tilde{\mu}$-proper sets increasing to $[0, \infty)$. Letting $C_i = t^{-1}(A_i)$ defines a sequence of μ-proper sets increasing to ϑ. Since $h(\theta) = \tilde{h}(t(\theta))$, $\tilde{h} \geq I_{A_i}$ implies that $h \geq I_{C_i}$. By Proposition C.1, $\tilde{h} \in L^2(\tilde{\mu})$ implies $h \in L^2(\mu)$ and the corresponding Dirichlet forms are equal. Thus,

$$\inf_{\tilde{h} \in \mathcal{H}_{\mu}(C_i)} \Delta(h; T, \mu) = \inf_{\tilde{h} \in \mathcal{H}_{\tilde{\mu}}(A_i)} \Delta(\tilde{h}; \tilde{T}, \tilde{\mu}) = 0$$

for every C_i in the sequence. By Theorem A.1, the chain with kernel T is locally μ-recurrent. Since μ and ν are equivalent measures, the chain is locally ν-recurrent as well.

By Proposition C.1, T is almost-ν bounded away from zero, and thus all of the conditions for Theorem 3.4 are satisfied. Therefore, under squared error loss, the family $\mathcal{F}_{\nu}$ is Φ_f-admissible.

Appendix D: Location models

Proof of Theorem 4.1. We will verify the conditions of Theorem 3.3. Clearly f is symmetric, positive and dominates $\|\theta - \eta\|^2$. It is straightforward to see that Eaton's kernel is given by $R(\mathrm{d}\theta|\eta) = r(\theta - \eta)\,\mathrm{d}\theta$ where

$$r(u) = r(-u) = \int \int q(u|x, y)g(x, y)\,\mathrm{d}x\lambda(\mathrm{d}y).$$

Notice that

$$T(\eta) = \int f(\theta, \eta)R(\mathrm{d}\theta|\eta) = \int \|u\|^2 r(u)\,\mathrm{d}u + d$$

is a finite constant. Setting $T(\eta) = c$ we see that

$$T(\mathrm{d}\theta|\eta) = \frac{\|\theta - \eta\|^2 + d}{c} r(\theta - \eta)\,\mathrm{d}\theta$$

and hence has the form $T(\mathrm{d}\theta|\eta) = t(\theta - \eta)\,\mathrm{d}\theta$ where t is symmetric. The remainder of the proof closely follows that of Theorem 7.1 in [9] and hence is omitted. □

Appendix E: Multivariate Normal

Proof of Theorem 4.2. Let $\gamma_a(z) = (a + z)^{-p/2}$, so that our family of priors (22) can be expressed as $\gamma_a(\|\theta\|^2)\,d\theta$. The function $t(\theta) = \|\theta\|^2$ fulfills the requirements of Assumption 3.3 as can be seen by letting the sets $A_i = [i - 1, i)$ partition the nonnegative real numbers. Letting $\pi(d\theta|\beta)$ denote the uniform distribution on the hypersphere of radius $\sqrt{\beta}$, it can be shown that $\nu(d\theta) = \pi(d\theta|\beta)\tilde{\nu}(d\beta)$ with

$$\tilde{\nu}(d\beta) = \frac{[\Gamma(1/2)]^p}{\Gamma(p/2)} \gamma_a(\beta)\beta^{p/2-1}\,d\beta$$

on $[0, \infty)$. Let π_1 be the uniform distribution on the unit hypersphere Ξ. Then the reduced sampling distribution has density

$$\tilde{p}(x|\beta) = \int_\Xi (2\pi)^{-p/2} \exp\left(-\frac{1}{2}\|x - \xi\sqrt{\beta}\|\right)\pi_1(d\xi)$$

with respect to Lebesgue measure on $\mathcal{X}$. If

$$m(x) = \int_0^\infty \tilde{p}(x|\beta)\tilde{\nu}(d\beta),$$

then the formal posterior has density

$$\tilde{q}(\beta|x) = \frac{[\Gamma(1/2)]^p}{\Gamma(p/2)} \frac{\tilde{p}(x|\beta)\gamma_a(\beta)\beta^{p/2-1}}{m(x)}$$

with respect to Lebesgue measure on $[0, \infty)$. The expected posterior

$$\tilde{R}(d\beta|\alpha) = \int_\mathcal{X} \tilde{q}(\beta|x)\tilde{p}(x|\alpha)\,dx\,d\beta \tag{33}$$

is a $\tilde{\nu}$-symmetric Markov transition kernel on $[0, \infty)$.

Let $\tilde{f}(\beta, \alpha) = 2(\beta + \alpha + c)$ with c a positive constant. It is clear that $\tilde{f}$ is symmetric in α and β and bounded away from zero. For any nonnegative real number α,

$$\tilde{T}(\alpha) = \int \tilde{f}(\beta, \alpha)\tilde{R}(d\beta|\alpha) = 2\int \beta\tilde{R}(d\beta|\alpha) + 2\alpha + 2c$$

is greater than or equal to c. Thus the kernel

$$\tilde{T}(d\beta|\alpha) = \frac{(\beta + \alpha + c)\tilde{R}(d\beta|\alpha)}{\int(v + \alpha + c)\tilde{R}(dv|\alpha)} \tag{34}$$

defines a $\tilde{\mu}$-symmetric Markov chain on $[0, \infty)$ where

$$\tilde{\mu}(A) = \int_A \tilde{T}(\alpha)\tilde{\nu}(d\alpha), \quad A \in \mathcal{A}.$$

Clearly, for any integer n, $\tilde{\mu}([0, n)) < \infty$.

The next step is to verify the conditions of Theorem 3.6 which will imply the Markov chain associated with $\tilde{T}$ is locally $\tilde{\mu}$-recurrent. Since $\tilde{\mu}$ and $\tilde{\nu}$ are equivalent measures we will also conclude that the chain is locally $\tilde{\nu}$-recurrent. In Appendix E.2 it is shown that $\tilde{T}([0, m]|\alpha)$ is continuous as a function of α implying condition (19) of Theorem 3.6. For k a nonnegative integer set

$$m_k(\alpha) = \int (\beta - \alpha)^k \tilde{T}(d\beta|\alpha).$$

Additional calculations given in Appendix E.3 show that

$$\frac{m_3(\alpha)}{m_2(\alpha)} = \mathrm{O}(1) \quad \text{as } \alpha \to \infty \tag{35}$$

and

$$m_1(\alpha) = \frac{8\alpha + \psi_1(\alpha)}{\int (\beta + \alpha + c)\tilde{R}(\mathrm{d}\beta|\alpha)}, \tag{36}$$

where $\psi_1(\alpha) = \mathrm{O}(1)$ as $\alpha \to \infty$. Letting $\phi(\alpha) = 1/\sqrt{\alpha}$, note that $\lim_{\alpha \to \infty} \phi(\alpha)\log(\alpha) = 0$ and, by calculations in Appendix E.3,

$$\frac{m_2(\alpha)}{2\alpha}\big[1 + \phi(\alpha)\big] = \frac{8\alpha + 8\sqrt{\alpha} + \psi_2(\alpha)}{\int (\beta + \alpha + c)\tilde{R}(\mathrm{d}\beta|\alpha)}, \tag{37}$$

where $\psi_2(\alpha) = \mathrm{O}(1)$ as $\alpha \to \infty$. It is clear by inspection of (35), (36), and (37) that the conditions (20) and (21) of Theorem 3.6 are satisfied for α large enough. Hence the chain is locally $\tilde{\nu}$-recurrent.

By Theorem 3.5, the family of priors $\mathcal{F}_\nu$ is strongly admissible with respect to Φ_f. $\qquad\square$

Here we give the supplemental arguments required for the proof of Theorem 4.2.

E.1. *Existence of the integrated risk difference*

In order to appeal to Blyth's method, we need to know that the integrated risk differences are defined. Otherwise the bounding inequality (B.2) for the integrated risk differences is meaningless. The proper Bayes estimators necessarily have finite integrated risks, so it is sufficient to show that

$$\int \mathsf{R}(\theta; \hat{\theta}) g_n(\theta)\nu(\mathrm{d}\theta) < \infty,$$

where g_n is the indicator of C_n a closed ball around zero with radius n. Since ν is σ-finite, $\nu(C_n)$ is finite. Since the risk function is real-valued and continuous, it attains a finite maximum on C_n; call it M_n. Therefore,

$$\int \mathsf{R}(\theta; \hat{\theta}) g_n(\theta)\nu(\mathrm{d}\theta) \leq M_n \nu(C_n) < \infty$$

and the integrated risk differences are defined.

E.2. *Continuity of the transition kernel*

Let m be a nonnegative integer. We wish to show that $T([0, m]|\alpha)$ is continuous as a function of α. Since by definition

$$T\big([0, m]|\alpha\big) = \frac{\int_0^m \beta\tilde{R}(\mathrm{d}\beta|\alpha) + \alpha\tilde{R}([0, m]|\alpha) + c\tilde{R}([0, m]|\alpha)}{\int \beta\tilde{R}(\mathrm{d}\beta|\alpha) + \alpha + c}$$

it is sufficient to show that $\tilde{R}([0, m]|\alpha)$ is continuous and that continuity of

$$\int_0^m \beta\tilde{R}(\mathrm{d}\beta|\alpha) \quad \text{and} \quad \int \beta\tilde{R}(\mathrm{d}\beta|\alpha)$$

follows from there.

Fix $\alpha^* \in [0, \infty)$ and δ greater than zero. Let $S_\delta(\alpha^*) = [0, (\sqrt{\alpha^*} + \delta)^2)$. Let (α_n) be a sequence with limit α^* whose elements are in $S_\delta(\alpha^*)$. Define

$$A_\delta\big(\alpha^*\big) = \big\{x: \|x\| < 2\big(\sqrt{\alpha^*} + \delta\big)\big\},$$

an open ball in $\mathcal{X}$ with radius greater than twice $\sqrt{\alpha^*}$. Let Ξ denote the surface of the unit hypersphere in $\mathbb{R}^p$. For any $x \in A_\delta^c(\alpha^*)$ – that is, any point such that $\|x\|$ is greater than twice $\sqrt{\alpha^*} + \delta$ – and ξ on the unit hypersphere

$$\|x - \xi\sqrt{\alpha_n}\| \geq \left\|x\left(1 - \sqrt{\alpha_n}/\|x\|\right)\right\| \geq \|x/2\|$$

since x cannot be closer to any point with radius $\sqrt{\alpha_n}$ than it is to $x\sqrt{\alpha_n}/\|x\|$, the point with radius $\sqrt{\alpha_n}$ on the common ray and since $\sqrt{\alpha_n}$ is less than $\sqrt{\alpha^*} + \delta$ by construction. Recall that

$$\tilde{p}(x|\alpha) = \int_\Xi (2\pi)^{-p/2} e^{-(1/2)\|x - \xi\sqrt{\alpha}\|} \pi(d\xi),$$

where π is the uniform distribution on Ξ. Let

$$g_1(x) = (2\pi)^{-p/2}\left[I_{A_\delta(\alpha^*)}(x) + I_{A_\delta^c(\alpha^*)}(x)e^{-(1/8)\|x\|^2}\right].$$

Since $\int g_1(x)\pi(d\xi) < \infty$ and $g_1(x)$ dominates the integrand of $\tilde{p}(x|\alpha_n)$, we can say that

$$\lim_{n\to\infty} \tilde{p}(x|\alpha_n) = \tilde{p}(x|\alpha^*).$$

Furthermore, $\int g_1(x)\,dx < \infty$ and $g_1(x)$ dominates $\tilde{p}(x|\alpha_n)$ by monotonicity of the integral. Therefore,

$$\lim_{n\to\infty} \int \tilde{p}(x|\alpha_n)\,dx = \int \lim_{n\to\infty} \tilde{p}(x|\alpha_n)\,dx = \int \tilde{p}(x|\alpha^*)\,dx.$$

Denote the Borel σ-algebra on $[0, \infty)$ by $\mathcal{B}$, and choose $B \in \mathcal{B}$. Since the densities $\tilde{p}$ and $\tilde{q}$ are necessarily nonnegative and measurable, Fubini says

$$\tilde{R}(B|\alpha) = \int_B \int_\mathcal{X} \tilde{Q}(d\beta|x)\tilde{P}(dx|\alpha) = \int_\mathcal{X} \tilde{p}(x|\alpha)\tilde{Q}(B|x)\,dx.$$

Similarly,

$$\int_B \beta\tilde{R}(d\beta|\alpha) = \int_B \beta \int_\mathcal{X} \tilde{Q}(d\beta|x)\tilde{P}(dx|\alpha) = \int_\mathcal{X} \tilde{p}(x|\alpha) \int_B \beta\tilde{q}(\beta|x)\,d\beta\,dx.$$

Let

$$f(x) = \tilde{p}(x|\alpha^*)\tilde{Q}(B|x) \quad \text{and} \quad f_n(x) = \tilde{p}(x|\alpha_n)\tilde{Q}(B|x)$$

so that

$$\lim_{n\to\infty} f_n(x) = f(x).$$

Since $\tilde{Q}(B|x)$ is a probability,

$$f_n(x) \leq \tilde{p}(x|\alpha_n) \leq g_1(x).$$

Hence, by the dominated convergence theorem,

$$\lim_{n\to\infty} \int f_n(x)\,dx = \int \lim_{n\to\infty} f_n(x)\,dx = \int f(x)\,dx.$$

That is,

$$\lim_{n\to\infty} \tilde{R}(B|\alpha_n) = \tilde{R}(B|\alpha^*).$$

Therefore, $\tilde{R}(B|\alpha)$ is a continuous function of α.

Now note that

$$\int_B \beta \tilde{q}(\beta|x)\,d\beta \le \int \beta \tilde{q}(\beta|x)\,d\beta.$$

By Proposition A.7 of Eaton et al. [12], for some bounded $\psi(\|x\|^2)$,

$$\int \beta \tilde{q}(\beta|x)\,d\beta = -p + \|x\|^2 + \psi(\|x\|^2)$$

so that there exists a constant $k > 0$ such that

$$\int \beta \tilde{q}(\beta|x)\,d\beta \le k + \|x\|^2.$$

Let $g_2(x) = k + \|x\|^2$, and let $g(x) = g_1(x)g_2(x)$. Let

$$f_n(x) = \tilde{p}(x|\alpha_n)\int_B \beta \tilde{q}(\beta|x)\,d\beta \quad \text{and} \quad f(x) = \tilde{p}(x|\alpha^*)\int_B \beta \tilde{q}(\beta|x)\,d\beta.$$

Note that

$$\lim_{n\to\infty} f_n(x) = f(x)$$

and $g(x)$ dominates $f_n(x)$. One can show that

$$\int g(x)\,dx < \infty$$

by expanding the product $g_1(x)g_2(x)$ and integrating the components. Now, by the Dominated Convergence Theorem,

$$\lim_{n\to\infty}\int f_n(x)\,dx = \int \lim_{n\to\infty} f_n(x)\,dx = \int f(x)\,dx.$$

That is,

$$\lim_{n\to\infty}\int_B \beta R(d\beta|\alpha_n) = \int_B \beta R(d\beta|\alpha^*).$$

Therefore, $\int_B \beta R(d\beta|\alpha)$ is continuous as a function of α. Since B is any Borel set, the continuity holds for both $[0, m]$ and $[0, \infty)$. Finally, since $T([0, m]|\alpha)$ is an algebraic combination of continuous functions, it is itself a continuous function.

E.3. *Moment conditions*

We are interested in moments of the transition kernel about the current state:

$$m_k(\alpha) = \int (\beta - \alpha)^k \tilde{T}(d\beta|\alpha).$$

We can express these moments in terms of $\tilde{R}$ as

$$m_k(\alpha) = \frac{\int (\beta - \alpha)^k (\beta + \alpha)\tilde{R}(d\beta|\alpha) + c\int (\beta - \alpha)^k \tilde{R}(d\beta|\alpha)}{\int (\beta + \alpha + c)\tilde{R}(d\beta|\alpha)}.$$

We know from Appendix A of Eaton et al. [12] that

$$\int \beta \tilde{R}(\mathrm{d}\beta|\alpha) = \alpha + \phi_1(\alpha), \tag{38}$$

$$\int \beta^2 \tilde{R}(\mathrm{d}\beta|\alpha) = \alpha^2 + 8\alpha + \phi_2(\alpha), \quad \text{and} \tag{39}$$

$$\int \beta^3 \tilde{R}(\mathrm{d}\beta|\alpha) = \alpha^3 + 24\alpha^2 + \phi_3(\alpha), \tag{40}$$

where, as $\alpha \to \infty$, $\phi_1(\alpha) = \mathrm{O}(\alpha^{-1})$, $\phi_2(\alpha) = \mathrm{O}(1)$, and $\phi_3(\alpha) = \mathrm{O}(\alpha)$. A similar argument shows that

$$\int \beta^4 \tilde{R}(\mathrm{d}\beta|\alpha) = \alpha^4 + 48\alpha^3 + \phi_4(\alpha), \tag{41}$$

where $\phi_4(\alpha) = \mathrm{O}(\alpha^2)$ as $\alpha \to \infty$.

Let $g_0(z) = (a + z)^{-p/2}$, and let $t_k(y) = \mathrm{E}[g_0(U)U^k|y]$ with $U \sim \chi_p^2(y)$. Proposition A.2 of Eaton et al. [12] establishes that if $Y \sim \chi_p^2(\alpha)$, then

$$\int \beta^k \tilde{R}(\mathrm{d}\beta|\alpha) = \mathrm{E}\left[\frac{t_k(Y)}{t_0(Y)}\Big|\alpha\right] < \infty.$$

Let

$$w_k(n) = \int_0^\infty \frac{g_0(z/2)^{n+p/2+k-1}\mathrm{e}^{-z/2}}{2\Gamma(n + p/2)}\,\mathrm{d}z$$

and note that $\mathrm{E}[g_0(U)U^k|y] = 2^k\mathrm{E}[w_k(N)|y]$ where $N|y \sim \mathrm{Poisson}(y/2)$. This last equality follows from expressing U as a Poisson mixture of χ_p^2 random variables. From the definition of t_k and w_k, we have that $r_k(y) = 2^k\mathrm{E}[w_k(N)|y]$. Therefore,

$$\int \beta^k \tilde{R}(\mathrm{d}\beta|\alpha) = 2^k\mathrm{E}\left[\frac{\mathrm{E}(w_k(N)|Y)}{\mathrm{E}(w_0(N)|Y)}\Big|\alpha\right].$$

Proposition A.7 of Eaton et al. [12] establishes that

$$\frac{\mathrm{E}[w_k(N)|y]}{\mathrm{E}[w_{k-1}(N)|y]} = \frac{y}{2} + 2(k-1) - \frac{p}{2} + \psi_k(y),$$

where $|\psi_k(y)| \leq d_k/y$ for some finite positive constant d_k. This allows us to evaluate the right hand side of as

$$\int \beta^k \tilde{R}(\mathrm{d}\beta|\alpha) = 2^k\mathrm{E}\left[\frac{\mathrm{E}(w_k(N)|Y)}{\mathrm{E}(w_{k-1}(N)|Y)} \cdots \frac{\mathrm{E}(w_1(N)|Y)}{\mathrm{E}(w_0(N)|Y)}\Big|\alpha\right].$$

Furthermore, $\mathrm{E}[\psi_k(Y)|\alpha] = \mathrm{O}(\alpha^{-1})$ by Proposition A.8 of Eaton et al. [12]. Note that the third moment of a non-central $\chi_p^2(\alpha)$ is $\alpha^3 + \mathrm{O}(\alpha^2)$ and the fourth moment is

$$\alpha^4 + 24\alpha^3 + 4p\alpha^3 + \mathrm{O}(\alpha^2).$$

Combining these results leads to (41).

We now begin to find expressions for the transitional moments m_k in terms of equations (38), (39), (40), and (41) as α becomes large. First, note that if $0 < c < \infty$, then

$$c\int(\beta - \alpha)^2\tilde{R}(\mathrm{d}\beta|\alpha) = \mathrm{O}(\alpha) \quad \text{and} \quad c\int(\beta - \alpha)^3\tilde{R}(\mathrm{d}\beta|\alpha) = \mathrm{O}(\alpha^2).$$

Express the first transitional moment as

$$m_1(\alpha) = \frac{\int (\beta - \alpha)(\beta + \alpha)\tilde{R}(\mathrm{d}\beta|\alpha) + c \int (\beta - \alpha)\tilde{R}(\mathrm{d}\beta|\alpha)}{\int (\beta + \alpha + c)\tilde{R}(\mathrm{d}\beta|\alpha)}$$

$$= \frac{\int \beta^2 \tilde{R}(\mathrm{d}\beta|\alpha) - \alpha^2 + c \int \beta \tilde{R}(\mathrm{d}\beta|\alpha) - c\alpha}{\int (\beta + \alpha + c)\tilde{R}(\mathrm{d}\beta|\alpha)}$$

$$= \frac{8\alpha + \mathrm{O}(1)}{\int (\beta + \alpha + c)\tilde{R}(\mathrm{d}\beta|\alpha)}.$$

Since $(\beta - \alpha)^2(\beta + \alpha) = \beta^3 - \beta^2\alpha - \beta\alpha^2 + \alpha^3$, we have that

$$\int (\beta - \alpha)^2(\beta + \alpha)\tilde{R}(\mathrm{d}\beta|\alpha) = \alpha^3 + 24\alpha^2 + \mathrm{O}(\alpha) - \alpha^3 - 8\alpha^2 - \mathrm{O}(\alpha) - \alpha^3 - \mathrm{O}(\alpha) + \alpha^3$$

$$= 16\alpha^2 + \mathrm{O}(\alpha).$$

Express the second transitional moment as

$$m_2(\alpha) = \frac{\int (\beta - \alpha)^2(\beta + \alpha)\tilde{R}(\mathrm{d}\beta|\alpha) + c \int (\beta - \alpha)^2 \tilde{R}(\mathrm{d}\beta|\alpha)}{\int (\beta + \alpha + c)\tilde{R}(\mathrm{d}\beta|\alpha)} = \frac{16\alpha^2 + \mathrm{O}(\alpha)}{\int (\beta + \alpha + c)\tilde{R}(\mathrm{d}\beta|\alpha)}.$$

Since $(\beta - \alpha)^3(\beta + \alpha) = \beta^4 - 2\beta^3\alpha + 2\beta\alpha^3 - \alpha^4$, we have that

$$\int (\beta - \alpha)^3(\beta + \alpha)\tilde{R}(\mathrm{d}\beta|\alpha) = \alpha^4 + 48\alpha^3 + \mathrm{O}(\alpha^2) - 2\alpha^4 - 48\alpha^3 - \mathrm{O}(\alpha^2)$$

$$+ 2\alpha^4 + \mathrm{O}(\alpha) - \alpha^4$$

$$= \mathrm{O}(\alpha^2).$$

Since our expressions for the transitional moments all share a common denominator, the ratio $m_3(\alpha)/m_2(\alpha)$ may be evaluated as

$$\frac{m_3(\alpha)}{m_2(\alpha)} = \frac{\int (\beta - \alpha)^3(\beta + \alpha)\tilde{R}(\mathrm{d}\beta|\alpha) + c \int (\beta - \alpha)^3 \tilde{R}(\mathrm{d}\beta|\alpha)}{\int (\beta - \alpha)^2(\beta + \alpha)\tilde{R}(\mathrm{d}\beta|\alpha) + c \int (\beta - \alpha)^2 \tilde{R}(\mathrm{d}\beta|\alpha)}$$

$$= \frac{\mathrm{O}(\alpha^2)}{16\alpha^2 + \mathrm{O}(\alpha)}$$

$$= \mathrm{O}(1).$$

Acknowledgments

References

[1] J. Berger and W. E. Strawderman. Choice of hierarchical priors: Admissibility of Normal means. *Ann. Statist.* **24** (1996) 931–951. MR1401831

[2] J. Berger, W. E. Strawderman and D. Tan. Posterior propriety and admissibility of hyperpriors in Normal hierarchical models. *Ann. Statist.* **33** (2005) 606–646. MR2163154

[3] A. C. Brandwein and W. E. Strawderman. Stein estimation for spherically symmetric distributions: Recent developments. *Stat. Sci.* **27** (2012) 11–23. MR2953492

[4] L. D. Brown. Admissible estimators, recurrent diffusions, and insoluble boundary value problems. *Ann. Math. Statist.* **42** (1971) 855–903. MR0286209

[5] K. L. Chung and W. H. Fuchs. On the distribution of values of sums of random variables. *Mem. Amer. Math. Soc.* **6** (1951) 1–12. MR0040610

[6] M. L. Eaton. A method for evaluating improper prior distributions. In *Statistical Decision Theory and Related Topics III*. S. S. Gupta and J. O. Berger (Eds). Academic Press, Inc., New York, 1982. MR0705296

[7] M. L. Eaton. A statistical diptych: Admissible inferences – Recurrence of symmetric Markov chains. *Ann. Statist.* **20** (1992) 1147–1179. MR1186245

[8] M. L. Eaton. Admissibility in quadratically regular problems and recurrence of symmetric Markov chains: Why the connection? *J. Statist. Plan. Inference* **64** (1997) 231–247. MR1621615

[9] M. L. Eaton. Markov chain conditions for admissibility in estimation problems with quadratic loss. In *State of the Art in Probability and Statistics: Festschrift for Willem R. van Zwet* 223–243. M. de Gunst, C. Klaasen and A. van der Vaart (Eds). *IMS Lecture Notes Ser.* **36**. IMS, Beechwood, OH, 2001. MR1836563

[10] M. L. Eaton. Evaluating improper priors and recurrence of symmetric Markov chains: An overview. In *A Festschrift for Herman Rubin* 5–20. A. DasGupta (Ed.). *IMS Lecture Notes Ser.* **45**. IMS, Beechwood, OH, 2004. MR2126883

[11] M. L. Eaton, J. P. Hobert and G. L. Jones. On perturbations of strongly admissible prior distributions. *Ann. Inst. Henri Poincaré Probab. Stat.* **43** (2007) 633–653. MR2347100

[12] M. L. Eaton, J. P. Hobert, G. L. Jones and W.-L. Lai. Evaluation of formal posterior distributions via Markov chain arguments. *Ann. Statist.* **36** (2008) 2423–2452. MR2458193

[13] J. P. Hobert and C. P. Robert. Eaton's Markov chain, its conjugate partner, and $\mathcal{P}$-admissibility. *Ann. Statist.* **27** (1999) 361–373. MR1701115

[14] J. P. Hobert and J. Schweinsberg. Conditions for recurrence and transience of a Markov chain on $\mathbb{Z}^+$ and estimation of a geometric success probability. *Ann. Statist.* **30** (2002) 1214–1223. MR1926175

[15] J. P. Hobert, A. Tan and R. Liu. When is Eaton's Markov chain irreducible? *Bernoulli* **13** (2007) 641–652. MR2348744

[16] W. James and C. Stein (1961). Estimation with quadratic loss. In *Proc. Fourth Berkeley Symp. Math. Statist. Probab., Vol. 1* 361–380. Univ. California Press, Berkeley. MR0133191

[17] B. W. Johnson. On the admissibility of improper Bayes inferences in fair bayes decision problems. Ph.D. thesis, Univ. Minnesota, 1991. MR2686274

[18] R. E. Kass and L. Wasserman. The selection of prior distributions by formal rules. *J. Amer. Statist. Assoc.* **91** (1996) 1343–1370. MR1478684

[19] W.-L. Lai. Admissibility and recurrence of Markov chains with applications. Ph.D. thesis, Univ. Minnesota, 1996.

[20] S. P. Meyn and R. L. Tweedie. *Markov Chains and Stochastic Stability*. Springer, London, 1993. MR1287609

[21] D. Revuz. *Markov Chains*, 2nd edition. North-Holland, Amsterdam, 1984. MR0758799

[22] C. Stein. The admissibility of Pitman's estimator of a single location parameter. *Ann. Math. Statist.* **30** (1959) 970–979. MR0109392

[23] G. Taraldsen and B. H. Lindqvist. Improper priors are not improper. *Amer. Statist.* **64** (2010) 154–158. MR2757006