Bounds on the Bayes and Minimax Risk for Signal Parameter Estimation

Lawrence D. Brown and Richard C. Liu

***Abstract*—In estimating the parameter θ from a parametrized signal problem (with $0 \le \theta \le L$) observed through Gaussian white noise, four useful and computable lower bounds for the Bayes risk were developed. For problems with different L and different signal to noise ratios, some bounds are superior to the others. The lower bound obtained from taking the maximum of the four, serves not only as a good lower bound for the Bayes risk but also as a good lower bound for the minimax risks. Threshold behavior of the Bayes risk is also evident as shown in our lower bound.**

***Index Terms*—Bayes risk, minimax risk, threshold effect.**

I. Introduction

CONSIDER a system involving a transmitted signal of the form $s_\theta(t)$, $0 \le t \le T$, $0 \le \theta \le L$, and a received signal

$$r(t) = s_\theta(t) + \sigma\, db(t), \qquad 0 \le t \le T,$$

where $b(t)$ denotes Brownian motion, so that $\sigma db(t)$ is white noise with intensity σ. The form of s is known but θ is an unknown parameter to be estimated. If $\tilde{\theta} = \tilde{\theta}(r(\cdot))$ denotes an estimator of θ, the squared-error is $R(\theta, \tilde{\theta}) = E_\theta(\tilde{\theta} - \theta)^2$. Of interest here will be bounds—especially lower bounds—for the Bayes risk under a uniform prior,

$$B(L) = \inf_{\tilde{\theta}} \frac{1}{L} \int_0^L R(\theta, \tilde{\theta})\, d\theta.$$

Also of some interest is the minimax risk, $M(L) = \inf_{\tilde{\theta}} \sup_{0 \le \theta \le L} R(\theta, \tilde{\theta})$. Note that any bound for $B(L)$ is automatically one for $M(L)$ since

$$B(L) \le M(L). \tag{1.1}$$

Many of the results in this paper deal only with the location model, where

$$s_\theta(t) = s(t - \theta), \tag{1.2}$$

and several require further that s have compact support—i.e., $s_\theta(t) = s(t - \theta) = 0$ if $t < \theta$ or $t > \theta + W$. In this case, we always assume that $T \ge L + W$. Of particular interest as a specific example is the case where s is a rectangular signal, i.e.,

$$s_\theta(t) = s(t - \theta) = \begin{cases} S, & \theta \le t \le \theta + W \\ 0, & \text{otherwise.} \end{cases} \tag{1.3}$$

Manuscript received April 13, 1992; revised October 29, 1992. L. D. Brown was supported in part by NSF Grant DMS 8809016. R. C. Liu was supported by the NSF Postdoctoral Research Fellowships and in part by the Army Research Grant through Cornell MSI.

The authors are with the Department of Mathematics, White Hall, Cornell University, Ithaca, NY 14853.

IEEE Log Number 9209454.

It is also of interest to study a *discretized version* of this compact support problem. Here, θ is assumed to be restricted to the values $\theta = 0, W, 2W, \cdots, \tau W$; $\tau = [L/W]$. (Note that $\tilde{\theta}$ is not restricted to take only the values $0, W, \cdots, \tau W$.) Let $B_D(\tau)$ and $M_D(\tau)$ denote the uniform prior Bayes risk and the minimax risk in this problem. Then,

$$\begin{cases} B_D(\tau) \le M_D(\tau) \le M(L), \\ B(L) \le M(L). \end{cases} \tag{1.4}$$

We also suspect that $B_D(\tau) \le B(L)$ but we can only prove that $B_D(\tau - 1) \le B(L)$. (See Lemma 4.1.) Finally, if $B_D(\tau, W)$ denotes this Bayes risk as a function of τ and W then $B_D(\tau, W)/W^2$ is independent of W, i.e.,

$$B_D(\tau, W)/W^2 = B_D(\tau, \omega)/\omega^2. \tag{1.5}$$

Section II of this paper studies in detail the situation when θ is known to take one of only two different values. This includes the discretized problem having $\tau = 1$. It is not difficult in this case to find the exact value of $B_D(1)$ by numerical integration. However, it is also useful for other purposes (such as in Section III) to have a good analytic bound. This bound is derived in Section II, and is compared with the exact values in Table I. It is also proved that the efficiency of the maximum likelihood estimator relative to the Bayes (and minimax) estimator in this setting varies between 50% and 75%.

Section III builds on the results of Section II to produce bounds which are useful for the case where L/W or τ are moderate (between roughly 1 and 10 to 100, depending on s). Section IV presents a different style of bound that is appropriate when L/W (or τ) is large. The asymptotic situation as $L/W \to \infty$ (or $\tau \to \infty$) is also studied. It is shown that the asymptotic efficiency in this sense of the maximum likelihood estimator is 3/4.

Bounds of a similar style were given by Van Trees [13], Ziv and Zakai [14], Bellini and Tartara [1], Chazan, Zakai, and Ziv [5], etc. Brown and Liu [2] studied the situation in which t, L, W are fixed and $S \to \infty$. This situation was also studied earlier in Terentyev [12] and Ibragimov and Hasminskii [8].

Section V presents some tables and graphs which combine the results of the preceding sections. All proofs are deferred to Section VI.

The Threshold Phenomenon: Van Trees [13], Ziv and Zakai [14], and others have noticed a threshold behavior of the Bayes risk as the signal to noise ratio increases. This behavior is evident from our bound, and is seen graphically in Fig. 1 for larger values of L. There is also another weaker threshold effect in this section. It appears as L increases with signal

TABLE I
VALUES OF THE TWO-POINT BAYES RISK (B^*), THE BOUND (b_2) OF THEOREM 2.1, AND THE RISK ($\hat{R}$) OF THE MLE

λ	$B^*(\lambda)$	$\rho(\lambda)$	$b_2(\lambda)$	$b_2(\lambda)/B^*(\lambda)$	$\hat{R}(\lambda)$	$B^*(\lambda)/\hat{R}(\lambda)$
0.1	0.2475	0.5364	0.2468	0.9973	0.4602	0.5379
0.5	0.1990	0.6295	0.1942	0.9761	0.3085	0.6449
1.0	0.1124	0.6843	0.1083	0.9638	0.1587	0.7084
1.5	0.04931	0.7082	0.04731	0.9594	0.06681	0.7381
2.0	0.01715	0.7218	0.01642	0.9576	0.02275	0.7538
2.5	0.00474	0.7299	0.00453	0.9567	0.00621	0.7630
3.0	0.00104	0.7351	0.00099	0.9561	0.00134	0.7688
3.5	0.00018	0.7386	0.00017	0.9558	0.00023	0.7726
∞		0.75		0.9549		0.7854

and signal to noise ratio held constant. (It thus appears in the case of having to locate the center of moderately informative limited bandwidth signal transmitted somewhere in a large duration of time.) This behavior is seen graphically in Fig. 2 for larger values of the signal to noise ratio (Q). This weaker threshold effect is predicted by common sense, and mentioned in Ziv and Zakai [14], but our results provide some theoretical and numerical validation.

II. THE TWO POINT PARAMETER SPACE

Suppose θ is restricted to take one of the two values $\theta = 0$ or $\theta > 0$. Then a sufficient statistic for inference about θ is

$$X = (2\lambda\sigma^2)^{-1}\left\{\int [s_{\theta_1}(t) - s_0(t)]r(t)\,dt - \overline{a}\right\}:$$

$$\lambda^2 = \lambda^2(\theta_1) = \frac{1}{4\sigma^2}\int (s_{\theta_1}(t) - s_0(t))^2\,dt$$

$$\overline{a} = \frac{1}{2}\int (s_{\theta_1}^2(t) - s_0^2(t))\,dt. \tag{2.1}$$

Under θ_i, X has a normal distribution with mean $(-1)^{i+1}\lambda$, and variance 1 for $i = 0, 1$.

For a location model with compact signal support of length W,

$$\lambda^2(\theta) = \frac{\int s^2(t)\,dt}{2\sigma^2} \equiv \frac{Q}{2} \quad \text{for } \theta \geq W, \tag{2.2}$$

where Q is the total-signal-to-noise-rate ratio. In particular, for s rectangular as in (1.3),

$$\lambda^2(\theta) = \begin{cases} \frac{Q\theta}{2W}, & \text{if } 0 \leq \theta \leq W, \\ \frac{Q}{2} = \frac{S^2W}{2\sigma^2}, & \text{if } \theta \geq W. \end{cases} \tag{2.3}$$

Let $B^*(\lambda)$ denote the Bayes risk for this problem as a function of $\lambda = \lambda(\theta_1)$ when $\theta_1 = 1$. Thus, $B^*(\lambda)$ is the Bayes risk for the symmetric prior when X is observed with either $X \sim N(-\lambda, 1)$ or $X \sim N(\lambda, 1)$ and the loss from an estimate d is $(0-d)^2$ or $(1-d)^2$, respectively. The Bayes risk for general θ_1 and $\lambda = \lambda(\theta_1)$ is then $\theta_1^2 B^*(\lambda)$. This is also the minimax risk for this problem.

Let Φ and ϕ denote the standard normal cumulative distribution function (c.d.f.) and density function respectively. Here is our basic two point bound. All proofs appear in Section VI.

Theorem 2.1:

$$B^*(\lambda) \geq \rho(\lambda)\Phi(-\lambda) \equiv b_2(\lambda), \tag{2.4}$$

where

$$\rho(\lambda) = \left[1 + e^{4\lambda^2}\frac{\Phi(-3\lambda)}{\Phi(-\lambda)}\right]^{-1}.$$

This lower bound is balanced by the following slightly cruder upper bound.

Theorem 2.2:

$$B^*(\lambda) \leq \left[1 - e^{4\lambda^2}\frac{\Phi(-3\lambda)}{2\Phi(-\lambda)}\right]\Phi(-\lambda) \equiv \tilde{b}_2(\lambda). \tag{2.5}$$

The first four columns of Table I show selected values of $B^*(\lambda)$, $\rho(\lambda)$, $b_2(\lambda)$, and the ratio $b_2(\lambda)/B^*(\lambda)$. It is clear that $b_2(\lambda)$ is a good lower bound. The numerical evidence is quite convincing that $\rho(\lambda)$ is increasing in λ and that $b_2(\lambda)/B^*(\lambda)$ is decreasing in λ. (Table I is selected from a more complete table having λ varying from 0 to 5.8 with interval length 0.1.) However, we have only been able to analytically prove the following.

Theorem 2.3: ρ is increasing in λ with

$$\begin{cases} \frac{1}{2} \leq \rho(\lambda) \leq \frac{3}{4}, \\ \rho(0) = \frac{1}{2}, \\ \lim_{\lambda\to\infty}\rho(\lambda) = \frac{3}{4}. \end{cases}$$

The ratio $b_2(\lambda)/B^*(\lambda)$ satisfies

$$\begin{cases} \lim_{\lambda\to 0}(b_2(\lambda)/B^*(\lambda)) = 1, \\ \lim_{\lambda\to\infty}(b_2(\lambda)/B^*(\lambda)) = \frac{3}{\pi} = 0.9549\cdots, \\ \lim_{\lambda\to\infty}B^*(\lambda)/\Phi(-\lambda) = \frac{\pi}{4}. \end{cases}$$

The numerical evidence, plus Theorem 2.3 convincingly shows that

$$1 \geq \frac{b_2(\lambda)}{B^*(\lambda)} \geq \frac{3}{\pi} = 0.9549\cdots. \tag{2.6}$$

Since we have not been able to prove this, we note that Theorems 2.1–2.3 do establish that

$$1 \geq \frac{b_2(\lambda)}{B^*(\lambda)} \geq \inf_\lambda\left(\frac{b_2(\lambda)}{\tilde{b}_2(\lambda)}\right) = 0.9. \tag{2.6'}$$

The maximum likelihood estimator (mle) $\hat{\theta}$, is a common multipurpose suggestion. In the present context it is given by

$$\hat{\theta} = \begin{cases} \theta, & \text{if } X > 0, \\ \theta_0 (= 0), & \text{if } X < 0. \end{cases}$$

Hence, it has risk $R(\theta, \hat{\theta}) = \theta_1^2\Phi(-\lambda)$ at both $\theta = \theta_0$ and $\theta = \theta_1$. Table I also contains values of $\hat{R}(\lambda) = R(\theta, \hat{\theta})/\theta^2 = \Phi(-\lambda)$, and of $B^*(\lambda)/\hat{R}(\lambda)$, which is the Bayes risk efficiency of the maximum likelihood estimator. The pattern apparent in the table is verified in the following theorem.

Theorem 2.4: The Bayes risk efficiency of the mle, $B^*(\lambda)/\hat{R}(\lambda)$ is increasing in λ. It satisfies

$$\frac{1}{2} \le \frac{B^*(\lambda)}{\hat{R}(\lambda)} \le \frac{\pi}{4} = 0.7854\cdots \tag{2.7}$$

with

$$\begin{cases} \lim_{\lambda\to 0}\frac{B^*(\lambda)}{\hat{R}(\lambda)} = \frac{1}{2}, \\ \lim_{\lambda\to\infty}\frac{B^*(\lambda)}{\hat{R}(\lambda)} = \frac{\pi}{4}. \end{cases}$$

Since the Bayes and minimax risks are equal, Theorem 2.4 also provides a statement about the minimax risk efficiency of the mle. The minimax risk efficiency of the mle in a somewhat different problem also involving the normal location family was studied in Donoho, Liu, and MacGibbon [6]. For their problem they found an efficiency varying between 4/5 (approximately) and 1.

The bound $\Phi(-Q)/4$ for $B_2(\lambda)$ is given in Ziv and Zakai [14]. Our bound b_2 is better by a factor between 2 and 3. Values of $B_2(\lambda)$ have been computed elsewhere. For example, Casella and Strawderman [4], motivated by a somewhat different problem, computed values for $\lambda \le 1.7$ (approximately). See also Kuks and Olman [10].

III. Bounds for Moderate L or τ

A simple technical argument in Ziv and Zakai [14, following their (9a)] converts two point bounds into bounds for arbitrary L or τ. This argument can be somewhat improved, and the two point bound of Section II is also better than what they used. The result is summarized in the following theorem both for the general model and for the discretized model.

Theorem 3.1: In the local model (1.2),

$$B(L) \ge L^{-1}\int_0^L \theta^2 B^*(\lambda(\theta))\,d\theta.$$

In the discretized model,

$$B_D(\tau) \ge \frac{\tau^2 + 2\tau}{3} W^2 B^*((Q/2)^{1/2}).$$

In both cases, replacing B^* by the bound of Section II results in a further lower bound, e.g.,

$$\begin{aligned} B(L) &\ge L^{-1}\int_0^L \theta^2 b_2(\lambda(\theta))\,d\theta \\ &= L^{-1}\int_0^L \theta^2 \rho(\lambda(\theta))\Phi(-\lambda(\theta))\,d\theta \\ &= L^{-1}\overline{\rho}\int_0^L \theta^2\Phi(-\lambda(\theta))\,d\theta, \end{aligned} \tag{3.1}$$

where $1/2 < \overline{\rho} < 3/4$.

For a rectangular signal (1.3), substitution of (2.3) and (2.4) into (3.1) and integration by parts yields for $L \ge W$,

$$\begin{aligned} B(L) \ge\; & L^{-1}\left\{W^3\overline{\rho}_W\left[\frac{1}{3}\Phi\left(-\left[\frac{Q}{2}\right]^{1/2}\right) + \left[\frac{2}{Q}\right]^3\frac{5}{2}K_7\left[\frac{Q}{2}\right]\right]\right. \\ & \left. +\frac{1}{3}(L^3 - W^3)\rho\left(\left[\frac{Q}{2}\right]^{1/2}\right)\Phi\left(-\left[\frac{Q}{2}\right]^{1/2}\right)\right\}, \end{aligned} \tag{3.2}$$

where $\overline{\rho}_W$ is defined by (3.1) with $L = W$ and λ as in (2.3), and

$$K_k(X) = \frac{1}{2^{k/2}\Gamma(k/2)}\int_0^X t^{(k/2)-1}e^{-t/2}\,dt$$

is the c.d.f. of a χ_k^2 variable. (This formula can also be made correct for $L \le W$ if in that case one substitutes L for W and $QL/2W$ for $Q/2$. The term in (3.2) involving $(L^3 - W^3)$ thus becomes zero.) Note that since $\overline{\rho} < 1/2$, the value 1/2 may be substituted for $\overline{\rho}$ in (3.2) to get a more convenient lower bound. Also note that $\overline{\rho}_W < \rho\ ([Q/2]^{1/2}) < 3/4$.

Chazan, Ziv, and Zakai [5] give a different bound in place of (3.1). Their bound has a different derivation. Unlike (3.1) it does not seem to readily generalize to situations not involving a uniform prior or a location family. In the case of the rectangular signal (and uniform prior) their bound is

$$\begin{aligned} B(L) \ge L^{-1}\Bigg\{ & W^3\left[\frac{1}{6}\Phi\left(-\left[\frac{Q}{2}\right]^{1/2}\right) + \frac{3}{4}\left(\frac{2}{Q}\right)^2\right. \\ & \left.\cdot K_5\left(\frac{Q}{2}\right) - \frac{5}{2}\left(\frac{2}{Q}\right)^3 K_7\left(\frac{Q}{2}\right)\right] \\ & + \frac{1}{6}(L^3 - W^3)\Phi\left(-\left[\frac{Q}{2}\right]^{1/2}\right)\Bigg\}. \end{aligned} \tag{3.3}$$

Note that as $L \to \infty$, the two bounds satisfy

$$\frac{(3.2)}{(3.3)} \sim 2\rho(Q/2)(\text{as } L \to \infty) \sim (3/2)(\text{as } Q \to \infty). \tag{3.4}$$

but for L fixed and $Q \to \infty$, (3.3) $\sim (W^3/L)(3/Q^2)$ whereas (3.2) $= O(1/Q^3)$. (Actually, as $Q \to \infty$, $B(L) \sim (W^3/L)(4.80/Q^2)$. See Brown and Liu [2].) Clearly, sometimes one bound is preferable and sometimes the other. A numerical comparison we performed indicates that (3.3) is always better when $L = W$ (or $L < W$), and (3.2) only dominates for L moderately larger than W, and Q not too small or too large. As suggested by the above and (3.4), the ratio of the two bounds apparently always satisfies $2/3 \le (3.3)/(3.2) < \infty$. For these reasons we conclude that bounds in the style of Theorem 3.1 seem most suitable when L/W is moderate (certainly > 1) but not too large. (See Section IV.) Q should also be moderate.

It is relatively straightforward to generalize Theorem 3.1 to treat other prior distributions. Thus, suppose θ has an *a priori* distribution G, so that one is interested in the corresponding

Bayes risk

$$B(G) = \inf_{\tilde\theta} \int R(\theta, \tilde\theta) G(d\theta).$$

Two different bounds are possible.

Corollary 3.2: In the location model (1.2) with prior distribution G,

$$B(G) \geq \int\int (\theta - \xi)^2 B^*(\lambda(|\theta - \xi|)) G(d\theta) G(d\xi). \quad (3.5)$$

If G is symmetric about some value $(\bar\theta)$, then

$$B(G) \geq \int \theta^2 B^*(\lambda(\theta)) G(d\theta). \quad (3.6)$$

Analogous bounds are valid for the discretized model. When G is the uniform distribution, (3.6) is of course the same as Theorem 3.1. In the case of the rectangular signal, (3.6) is better than (3.5) (but $(3.6)/(3.5) \leq 2$) unless Q is large and L is small. Since also, $(3.5)/(3.6) \leq 2$, neither inequality is very good in this latter range, and neither is nearly as goo there as (3.3).

IV. Bounds for Large L or τ

Consider, for simplicity, the rectangular model. As $L \to \infty$ (with W and Q fixed) the bound of Theorem 3.1 yields only $B(L) \geq (L^2/3)B^*((\phi/2)^{1/2})$ as $L \to \infty$. Since $B^*(\alpha) < 1/4$ for $\alpha > 0$, this bound is less than $L^2/12$. However, it is shown below that $B(L) \sim L^2/12$ as $L \to \infty$ for any fixed W and Q. (This result is implicit in Van Trees [13] and elsewhere.) Hence, Theorem 3.1 is not sharp, and a different style of inequality is needed to accurately describe the difficulty of estimating the location of a single signal over a wide range of θ.

The bounds in this section begin with results for the discretized problem. These bounds for the discretized problem are then converted to bounds for the continuous problem by applying Lemma 4.1.

Lemma 4.1: Consider a location model with compact signal support, $[\theta, W]$. Let $\tau = [L/W] \geq 2$. Then,

$$B(L) \geq \frac{L^2 B_D(\tau - 1)}{\tau^2 W^2}. \quad (4.1)$$

Our first main result is in Theorem 4.2. Its proof (in Section VI) begins like the proof of Theorem 2.1, but concludes rather differently.

Theorem 4.2: For the discretized model,

$$B_D(\tau) \geq \frac{\tau^2 + 2\tau}{12} W^2 \frac{\tau + 1}{\tau + 1 + 2(e^Q - 1)}, \quad (4.2)$$

where Q is defined in (2.2). Consequently, by (4.1) and (4.2), for a location model with compact support

$$B(L) \geq \frac{L^2(\tau^2 - 1)}{12\tau(\tau + 2e^Q - 2)}. \quad (4.3)$$

The bounds (4.2) and (4.3) take the asymptotic form

$$\begin{cases} B_D(\tau) \geq \frac{\tau^2+2\tau}{12} W^2 \left[1 - \frac{2(e^Q-1)}{\tau} + O(\tau^{-2})\right] \\ B(L) \geq \frac{L^2}{12}\left[1 - \frac{2(e^Q-1)}{\tau} + O(\tau^{-2})\right]. \end{cases} \quad (4.4)$$

Therefore, $B(L) \sim L^2/12$, as previously claimed. (Trivially, $B(L) \leq L^2/12$ for all L.) However, the $O(1/\tau)$ terms in (4.4) are not correct. The following stronger bound yields the correct term of $O(1/\tau)$ for $B_D(\tau)$. We conjecture that the corresponding term is also correct for $B(L)$, but we have not been able to prove this.

As it is shown in Fig. 1, the bounds in Theorem 4.2 can sometimes dominate the lower bounds in Theorem 4.3 and vice versa, so both bounds need to be checked. Some numerical examples appear in Section V.

Theorem 4.3: Let $M = (\tau + 1)$ and $\Delta = e^Q - 1$. Then,

$$\begin{aligned} &\frac{(M^2-1)}{12} W^2 \frac{M}{M + 2\Delta} \\ &\quad \cdot \left\{1 + \frac{\Delta(M + 2((\Delta+1)^2 - 1))}{(M + 2\Delta)^2} + O(M^{-2})\right\} \\ &\geq B_D(\tau) \geq \frac{M^2 - 1}{12} W^2 \frac{M}{M + 2\Delta} \\ &\quad \cdot \left\{1 + \frac{\Delta(M + 2((\Delta+1)^2 - 1))}{(M + 2\Delta)^2}\right\} \\ &\quad \cdot \left[1 + \frac{\Delta^2(\Delta+3)(M + 2((\Delta+1)^3 - 1))}{(M+2\Delta)^3}\right]^{-1}. \end{aligned} \quad (4.5)$$

Consequently,

$$\begin{aligned} B(L) \geq &\frac{L^2(\tau^2 - 1)}{12\tau^2} \frac{\tau}{\tau + 2\Delta} \\ &\cdot \left\{1 + \frac{\Delta(\tau + 2((\Delta+1)^2 - 1))}{(\tau + 2\Delta)^2}\right\} \\ &\cdot \left[1 + \frac{\Delta^2(\Delta+3)(\tau + 2((\Delta+1)^3 - 1))}{(\tau + 2\Delta)^3}\right]^{-1}. \end{aligned} \quad (4.6)$$

Lengthy (but routine) calculations following on (6.9) would yield an explicit upper bound for the $O(M^{-2})$ term in (4.5). However the main interest here is in the asymptotic expansion discussed previously. This is formally described in the following corollary.

Corollary 4.4: As $\tau \to \infty$ for fixed Q and W,

$$B_D(\tau) = \frac{\tau^2 + 2\tau}{12} W^2 \left[1 - \frac{e^Q - 1}{\tau} + O(\tau^{-2})\right].$$

(The corollary is proved by expanding both sides of (4.5) to terms of $O(\tau^{-2})$. (4.6) yields the bound

$$B(L) \geq \frac{L^2}{12}\left[1 - \frac{(e^Q - 1)}{\tau} + O(\tau^{-2})\right].$$

As mentioned, we conjecture that this bound is sharp to the terms of order $O(\tau^{-2})$.)

It is fairly evident from this, and can be shown explicitly, that as $L \to \infty$ the Bayes procedure. $\tilde\theta_{\text{Bayes}}$ converges to $L/2$ in probability, uniformly for $\theta \in (0, L)$. Consequently $\sup_{\theta \in (0, L)} R(\theta, \tilde\theta_{\text{Bayes}}) = L^2/4$. It can further be shown that $M(L) \sim L^2/4$ as $L \to \infty$; we do not know what is the next term in the asymptotic expansion. Similar calculations also show that $\sup_{\theta \in (0, L)} R(\theta, \hat\theta) \sim L^2/3$ as $L \to \infty$, where $\hat\theta$ denotes the mle. Consequently, the asymptotic minimax risk efficiency of the mle is 3/4.

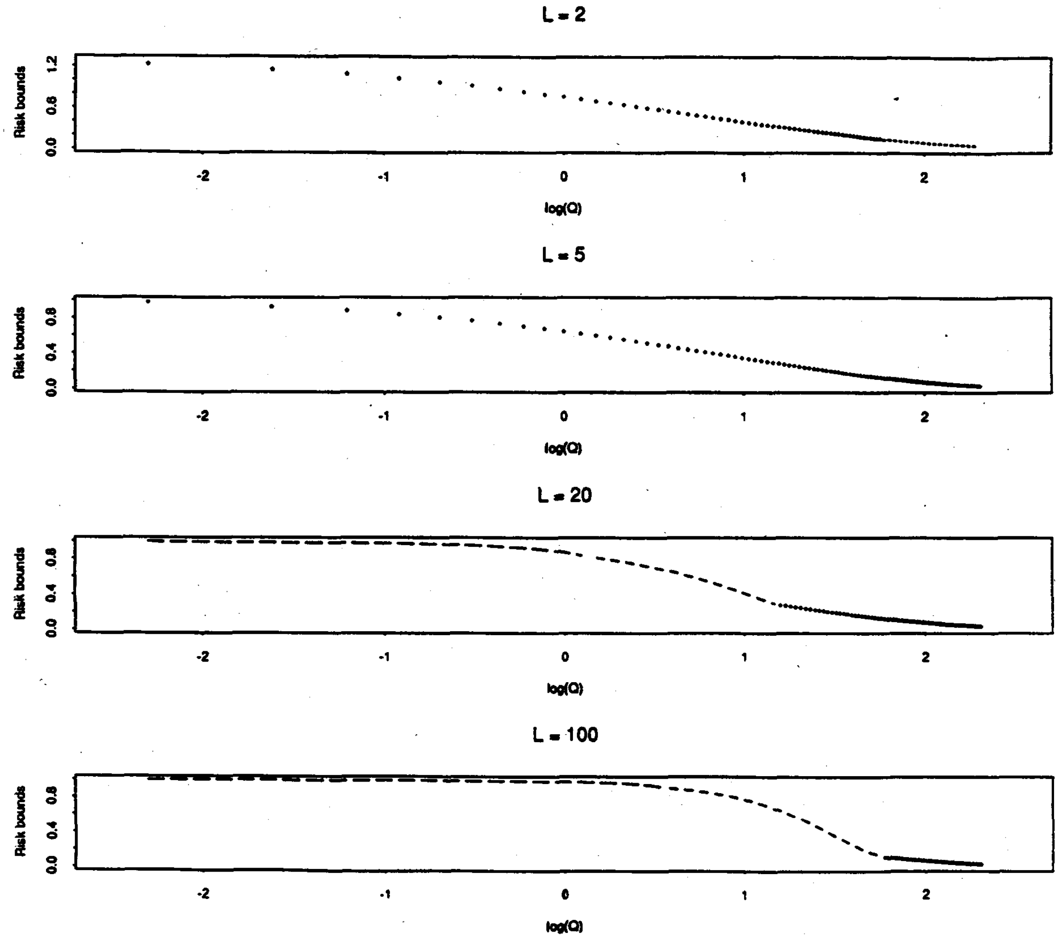

Fig. 1. Maximum of the normalized lower bounds as a function of $\ln(Q)$ where $W = 1$, $***$ represents (3.2), $\cdots$ represents (3.3), --- represents (4.3), ___ represents (4.6). All of them were normalized by the factor $L^2/12$.

V. Numerical Examples

A convenient way to compare lower bounds in Section III and IV for the discretized model is to multiply them by $12/W^2(\tau^2 + 2\tau) = 12/W^2(M^2 - 1)$. In this way, the normalized bound in Theorem 3.1 becomes $4B^*((Q/2)^{1/2}) < 1$, and the normalized bounds of Theorems 4.2 and 4.3 converge to 1 as $\tau \to \infty$. A comparison of bounds for the continuous model with $L = MW$ would be merely identical, except for the case $M = 2$ $(L = 2W)$ where the bound (3.2) after renormalization would be slightly different than $4B^*((Q/2)^{1/2})$. See also Fig. 1.

Van Trees [13], Ziv and Zakai [14], and others have identified a threshold effect as Q increases. This behavior is displayed in Fig. 1. This figure shows six plots of the various lower bounds for $B(L)$ in the rectangular case as functions of $\ln(Q)$. We take $W = 1$ with no real loss of generality. (This makes Q into the signal-rate-to-noise-rate ratio.) For easy comparison of the graphs with each other and with Table II, we normalize the bounds via multiplication by $12/(\tau^2 - 1)$ where $\tau = L/W(= L)$. (Thus, the Fig. 1 bounds for τ compare to the Table II bounds for $M = \tau + 1$ since the bounds in (4.3) and (4.6) derive from $B_D(\tau - 1)$.) The graphs compare four normalized bounds: (3.2) and (3.3) from Chazan, Ziv, and Zakai [5], (4.3) from Theorem 4.2, and (4.6) from Theorem 4.3.

The Bayes risk, $B(L)$, is of course larger than the maximum of the bounds in Fig. 1. However, it seems reasonable to presume that this maximum yields a moderately good idea as to the true behavior of $B(L)$. We see from the figures that, especially for moderate to large $\tau = L/W$, the threshold certainly occurs later and is presumably steeper than one might suppose on the basis of the previously published (3.3). Even though the computations for large L (e.g., $L = 1000$, 10000) are not graphed here, the new bound (4.6) is obviously the best for $\log(Q) < 2$ (for details, see Brown and Liu [3]). The existence of a threshold as L increases for fixed Q, W has also been mentioned. (See, e.g., Ziv and Zakai [14, section IV], but very little numerical information seems to exist. This threshold behavior can be seen in the bound of Section IV. It is displayed in Fig. 2. Again we take $W = 1$. For six values of Q, this figure shows the value of the largest of the bounds

TABLE II
COMPARISON OF NORMALIZED LOWER BOUNDS

$Q=$		0.5	1.0	2.0	3.0	5.0	10.0
$4B^*((Q/2)^{1/2})=$		0.7959	0.6509	0.4496	0.3202	0.1692	0.0387
M	LB						
2	A	0.6065	0.3668	0.1353	0.0498	0.0067	0.0000
2	B	0.5804	0.1526	0.0058	0.0003	0.0000	0.0000
5	A	0.79430	0.5926	0.2812	0.1158	0.0167	0.0001
5	B	0.8416	0.4232	0.0163	0.0007	0.0000	0.0000
20	A	0.9390	0.8534	0.6102	0.3438	0.0635	0.0004
20	B	0.9664	0.8851	0.1074	0.0293	0.0000	0.0000
100	A	0.9872	0.0668	0.8867	0.7237	0.2533	0.0023
100	B	0.9935	0.9826	0.7884	0.0272	0.0000	0.0000
1000	A	0.9987	0.9966	0.9874	0.9632	0.7723	0.0222
1000	B	0.9994	0.9983	0.9935	0.8846	0.0004	0.0000
10000	A	0.99987	0.99966	0.9987	0.9962	0.9714	0.1850
10000	B	0.99994	0.99983	0.9994	0.9980	0.0507	0.0000
100000	A	0.999987	0.999966	0.99987	0.99962	0.9971	0.6942
100000	B	0.999994	0.999983	0.99994	0.99981	0.9781	0.0000

(3.2), (3.3), (4.3), and (4.6), normalized via multiplication by $12/(L^2-1)$, as a function of $\ln(L)$. In order to obtain smooth interpolation in the figure, we graph (4.3) and (4.6) as though $\tau = L/W(=L)$ rather than the true expression $\tau = [L/W]$.

VI. PROOFS

Proof of Theorem 2.1: Recall that under θ_i, $X \sim N((-1)^{i+1}\lambda, 1)$, $i = 0, 1$. Let $\theta_1 = 1$. Le Cam [11, p. 49] yields

$$B^*(\lambda) = (1/2)\int \frac{\phi(x+\lambda)\phi(x-\lambda)}{\phi(x+\lambda)+\phi(x-\lambda)}\, dx. \tag{6.1}$$

Simplify to get

$$\begin{aligned} B^*(\lambda) &= (1/2)(2\pi)^{1/2}\int \frac{e^{-x^2/2-\lambda^2/2}}{e^{\lambda x}+e^{-\lambda x}}\, dx \\ &= \int_0^\infty \frac{\phi(x+\lambda)}{1+e^{-2\lambda x}}\, dx, \end{aligned} \tag{6.2}$$

by symmetry. Now note that for positive functions h, g,

$$\int (h(t)/g(t))\, d\mu(t) \geq \frac{(\int h(t)d\mu(t))^2}{\int h(t)g(t)d\mu(t)}. \tag{6.3}$$

Hence,

$$\begin{aligned} B^*(\lambda) &\geq \frac{\left(\int_0^\infty \phi(x+\lambda)\, dx\right)^2}{\int_0^\infty \phi(x+\lambda)(1+e^{-2\lambda x})\, dx} \\ &= \frac{(\Phi(-\lambda))^2}{\Phi(-\lambda)+e^{4Q^2}\Phi(-3\lambda)}, \end{aligned}$$

upon completing the square to evaluate the integral in the denominator. This is the desired inequality. □

Proof of Theorem 2.2: Note that for $0 \leq \alpha \leq 1$, $(1+\alpha)^{-1} \leq 1-\alpha/2$. Apply this inequality in the final integrand of (6.2) (with $\alpha = e^{-2\lambda x}$) and evaluate to get (2.5). □

Proof of Theorem 2.3: Note that

$$\begin{aligned} \frac{e^{4\lambda^2}\Phi(-3\lambda)}{\Phi(-\lambda)} &= \frac{\int_0^\infty e^{-2\lambda x}\phi(x+\lambda)\, dx}{\int_0^\infty \phi(x+\lambda)\, dx} \\ &= \frac{\int_0^\infty e^{-2y}\phi\left(\frac{y+\lambda^2}{\lambda}\right) dy}{\int_0^\infty \phi\left(\frac{y+\lambda^2}{\lambda}\right) dy}. \end{aligned}$$

Now, $\partial^2/\partial\lambda\,\partial y \ln[\phi((y+\lambda^2)/\lambda)] > 0$ for $y > 0$, $\lambda > 0$, and e^{-2y} is decreasing. Hence, standard mle arguments yield that $e^{4\lambda^2}\Phi(-3\lambda)/\Phi(-\lambda)$ is decreasing. (See, e.g., Karlin [9].) It follows that $\rho(\lambda)$ is increasing in λ. Clearly $\rho(0) = 1/2$. It is well known that $\Phi(-\lambda) \sim \lambda^{-1}\phi(\lambda)$ as $\lambda \to \infty$. (See, e.g., Feller [7].) Hence, $e^{4\lambda^2}\Phi(-3\lambda)/\Phi(-\lambda) \sim 1/3$ and $\lim_{\lambda\to\infty}\rho(\lambda) = 3/4$. This verifies all claims of the proposition concerning ρ.

It is evident from (6.2), that $\lim_{Q\to 0} B^*(\lambda) = 1/4$. Hence, $\lim_{\lambda\to 0}(b_2(\lambda)/B^*(\lambda)) = 1$. For the other extreme, expand

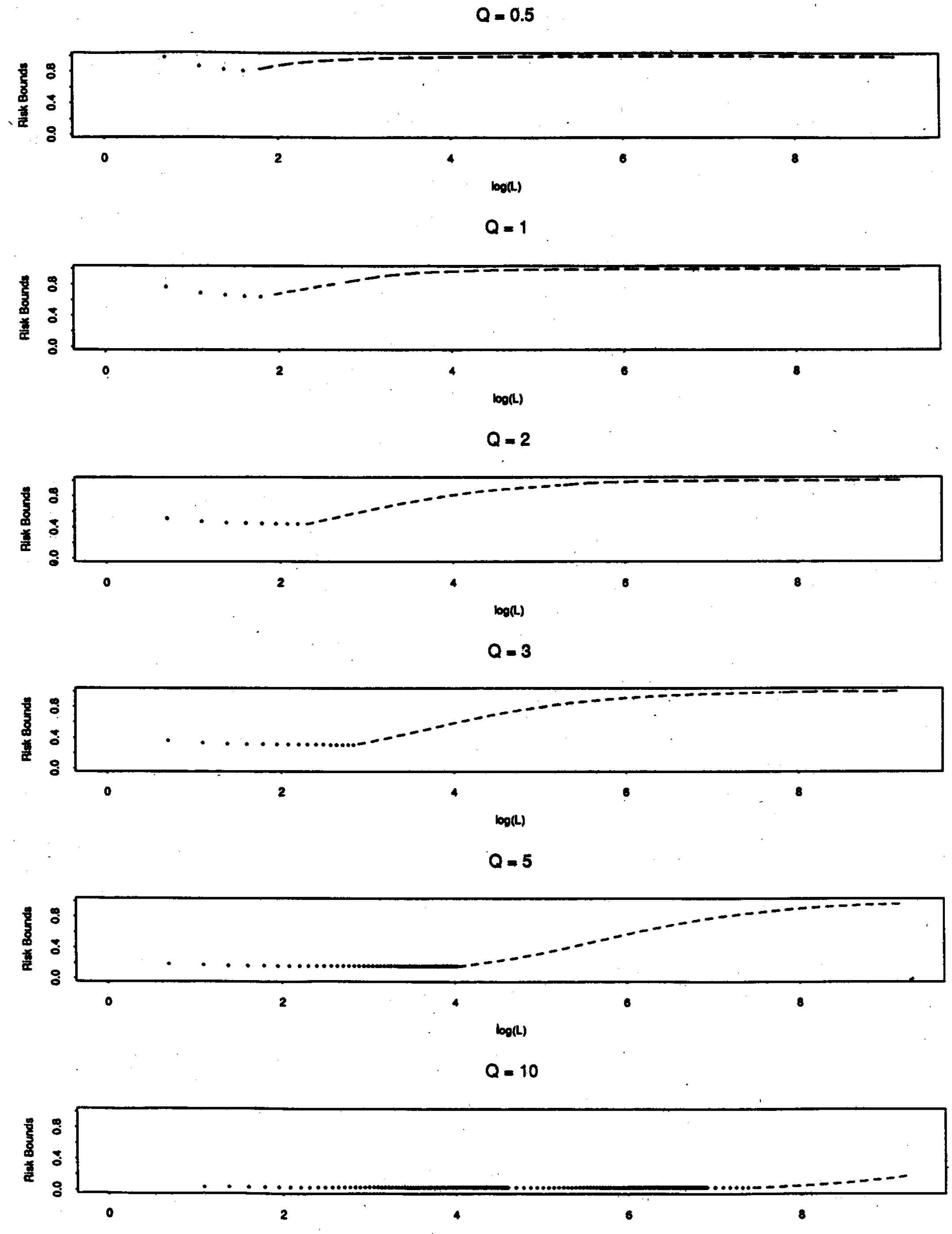

Fig. 2. Maximum of the normalized lower bounds as a function of $\ln(L)$ where the largest lower bound is (3.2), (3.3), (4.3), (4.6) is represented, respectively, by $***$, $\cdots$, ---, ___ and all were normalized by $L^2/12$.

$(1+e^{-2\lambda x})^{-1}$ in the integrand of (6.2) to get

$$B^*(\lambda) = \sum_{k=0}^{\infty} \int_0^{\infty} (-1)^k e^{-2k\lambda x} \phi(x+\lambda)\, dx$$

$$= \sum_{k=0}^{\infty} (-1)^k e^{[(2k+1)^2-1]\lambda^2/2} \Phi(-(2k+1)\lambda).$$

Again apply the expression $\Phi(-\lambda) = \lambda^{-1}\phi(\lambda)$ as $\lambda \to \infty$

to yield

$$\frac{b_2(\lambda)}{B^*(\lambda)} \sim \frac{[\lambda^{-1}\phi(\lambda)+(3\lambda)^{-1}\phi(\lambda)]^{-1}}{\sum_{k=0}^{\infty}(-1)^k[(2k+1)\lambda]^{-1}\phi(\lambda)} = \frac{3/4}{\sum_{k=0}^{\infty}(-1)^k/(2k+1)} = \frac{3}{\pi}.$$

Since $\sum_{k=0}^{\infty}(-1)^k/(2k+1) = \pi/4$. This also shows that $\lim_{\lambda\to\infty} B^*(\lambda)/\Phi(-\lambda) = \pi/4$. □

Proof of Theorem 2.4: From (6.2),

$$\frac{B^*(\lambda)}{\hat{R}(-\lambda)} = \frac{\int_0^\infty (1+e^{-2\lambda x})^{-1}\phi(x+\lambda)\,dx}{\int_0^\infty \phi(x+\lambda)\,dx} = \frac{\int_0^\infty (1+e^{-2y})^{-1}\phi\left(\frac{y+\lambda^2}{\lambda}\right)dy}{\int_0^\infty \phi\left(\frac{y+\lambda^2}{y}\right)dy}.$$

Now, $(1+e^{-2y})^{-1}$ is increasing. Hence, $B^*(\lambda)/\hat{R}(-\lambda)$ is also increasing, as in the proof of Theorem 2.3. The limiting values of $B^*(\lambda)/\hat{R}(\lambda)$ have already been established in that theorem. □

Proof of Theorem 3.1:

$$\begin{aligned} B(L) &= \inf_{\tilde\theta} \frac{1}{L}\int_0^L R(\theta,\tilde\theta)\,d\theta \\ &= \inf_{\tilde\theta}\left\{\frac{1}{2}\left\{\left[\frac{1}{L}\int_0^L R(\theta,\tilde\theta)\,d\theta\right] + \left[\frac{1}{L}\int_0^L R(L-\theta,\tilde\theta)\,d\theta\right]\right\}\right\} \\ &\ge \inf_{\tilde\theta}\frac{1}{L}\int_0^L\left[\frac{1}{2}R(\theta,\tilde\theta)+\frac{1}{2}R(L-\theta,\tilde\theta)\right]d\theta \\ &= \frac{1}{L}\int_0^\infty (L-2\theta)^2 B^*(\lambda(|L-2\theta|))\,d\theta \\ &= \frac{1}{L}\int_0^L t^2 B^*(\lambda(t))\,dt. \end{aligned}$$

In the discretized model $\lambda(t) = (Q/2)^{1/2}$ for $t = W,\cdots,\tau W$, and one similarly gets

$$\begin{aligned} B_D(\tau) &= \frac{1}{\tau+1}\left(\sum_{k=0}^{\tau}(2k-\tau)^2W^2\right)B^*((Q/2)^{1/2}) \\ &= \frac{\tau^2+2\tau}{3}W^2B^*((Q/2)^{1/2}). \end{aligned}$$ □

Proof of Corollary 3.2: To prove (3.5), proceed similarly to the previous proof to write

$$\begin{aligned} B(G) &= \int\int \inf_{\tilde\theta}\left[\frac{1}{2}R(\theta,\tilde\theta)+\frac{1}{2}R(\xi,\tilde\theta)\right]G(d\theta)G(d\xi) \\ &= \int\int(\theta-\xi)^2B^*(\lambda(|\theta-\xi|))G(d\theta)G(d\xi). \end{aligned}$$

(3.6) is proved similarly. □

Proof of Lemma 4.1: Let $\omega = L/\tau$. Then, $W \le \omega \le W + W/\tau$ and

$$\begin{aligned} B(L) &= \frac{1}{L}\sum_{k=0}^{\tau-1}\int_{k\omega}^{(k+1)\omega} R(\theta,\tilde\theta)\,d\theta \\ &= \int_0^\infty\left[\sum_{k=0}^{\tau-1}R(k\omega+t,\tilde\theta)\right]dt \\ &\ge \frac{1}{L}\int_0^\omega \tau B_D(\tau-1,\omega)\,dt \\ &= B_D(\tau-1,\omega) = B_D(\tau-1)\omega^2/W^2 \end{aligned}$$

by (1.5), where $B_D(\tau-1) = B_D(\tau-1, W)$. □

Proof of Theorem 4.2: LeCam [11, p. 49] can be extended to yield the following general results. Let $y \in Y$ have density $f_\theta(y)$ with respect to ν, $\theta \in \boldsymbol{R}$. Let G denote a prior distribution of θ and let B_G^* denote the Bayes risk for estimating θ under squared error loss. Then,

$$B_G^* \ge \left(\frac{1}{2}\right)\int\int\int(\theta-\xi)^2\frac{f_\theta(y)f_\xi(y)}{\int f_\zeta(y)G(d\zeta)} \cdot G(d\theta)G(d\xi)\nu(dy). \quad (6.4)$$

((6.4) is also valid for multidimensional θ when $\|\theta-\xi\|^2$ is substituted for $(\theta-\xi)^2$ in the formula.)

To apply (6.4) to the discretized problem let $\boldsymbol{Y} = \boldsymbol{R}^{\tau+1}$ and ν is a Lebesgue measure. The kth coordinate of Y is defined by

$$Y_k = (\sqrt{2}\lambda\sigma^2)^{-1}\int r(t)\dot{s}_{(k-1)W}(t)\,dt, \qquad k = 1,\cdots,\tau+1.$$

with $\lambda^2 = \int s_\theta^2(t)\,dt/2\sigma^2 = Q/2$, as in (2.2). It is easy to check that $Y = (Y_1,\cdots,Y_{\tau+1})$ is a sufficient statistic, and that when $\theta = (k-1)W$ then Y has a multivariate normal distribution with identity covariance matrix and with mean η having $\eta_k = \sqrt{2}\lambda$ and $\eta_i = 0$ for $i \ne k$. (To relate this to (2.1), note that $\tau = 1$ and $X = (Y_2 - Y_1)/\sqrt{2}$.) Substitution in (6.4) and simplification yields

$$B_D(\tau) \ge \left(\frac{W^2}{2}\right)\sum_{i=1}^{\tau+1}\sum_{j=1}^{\tau+1}\int_{\boldsymbol{R}^{\tau+1}}(i-j)^2 \cdot \frac{(2\pi)^{-(\tau+1)/2}\exp(-\|y\|^2/2+\sqrt{Q}(y_i+y_j)-Q)}{(\tau+1)\sum_{t=0}^{\tau+1}\exp(\sqrt{Q}y_k - Q/2)}. \quad (6.5)$$

Now apply (6.3) directly to (6.5) and evaluate to get the equation found at the top of the next page.

(In evaluating this equation, note that the integral evaluates to one if $k \ne i, j$ (but $i \ne j$) and to e^Q if $k = i$ or j (but $i \ne j$). The value when $i = j$ is unimportant, for then $(i-j)^2 = 0$.) This proves (4.2), (4.3) follows directly from (4.1) and (4.2). □

Proof of Theorem 4.3: In preparation for substitution in (6.5), fix i, j with $i \ne j$. Let

$$\nu(y) = (M+2\Delta)^{-1} \cdot \left\{\sum_{k\ne i,j}(e^{\sqrt{Q}y_k - Q/2}-1) + \sum_{k=i,j}(e^{\sqrt{Q}y_k-Q/2}-e^Q)\right\}$$

$$\begin{aligned} B_D &\geq \frac{W^2}{2(\tau+1)}\sum\sum(i-j)^2 \frac{1}{\sum_k \int (2\pi)^{-(\tau+1)/2} \exp\left(-\|y\|^2/2 + \sqrt{Q}(y_i+y_j+y_k) - 3Q/2\right) dy} \\ &= \frac{W^2}{2(\tau+1)}\sum\sum(i-j)^2 \frac{1}{\tau - 1 + 2e^Q} \\ &= \frac{W^2(\tau^2+2\tau)(\tau+1)}{12((\tau-1)+2e^Q)}. \end{aligned}$$

$$w(y) = (2\pi)^{-M/2} \exp\left(-\|y\|^2/2 + \sqrt{Q}(y_i+y_j) - Q\right).$$

Then, write

$$\int \frac{w(y)}{1+\nu(y)}\, dy = \int \left(1 - \nu(y) + \nu^2(y) - \frac{\nu^3(y)}{1+\nu(y)}\right) w(y)\, dy \quad (6.6)$$

and note

$$\begin{cases} \int \nu(y) w(y)\, dy = 0 \\ \int \nu^2(y) w(y)\, dy = \frac{\Delta(M+2(e^{2Q}-1))}{(M+2\Delta)^2}. \end{cases}$$

Note that

$$\int \frac{1+\nu^3(y)}{1+\nu(y)} w(y)\, dy \leq \left(\int \frac{w(y)}{1+\nu(y)}\, dy\right)\left(\int (1+\nu^3(y)) w(y)\, dy\right)$$

since $(1+\nu^3)$ is increasing in ν while $(1+\nu)^{-1}$ is decreasing in ν; and calculate that

$$\int (1+\nu^3(y)) w(y)\, dy = 1 + \frac{\Delta^2(\Delta+3)(M+2(e^{3Q}-1))}{(M+2\Delta)^3}. \quad (6.7)$$

Substitute (6.7) into (6.6), then into (6.5), and simplify to get the lower bound in (4.5).

For the upper bound, in place of (6.6) write

$$\int \frac{w(y)}{1+\nu(y)}\, dy = \int \left[1 - \nu(y) + \nu^2(y) - \nu^3(y) + \frac{\nu^4(y)}{1+\nu(y)}\right] w(y)\, dy,$$

and note, from (6.7), that $\int \nu^3(y) w(y)\, dy \geq 0$. Then,

$$\frac{1}{1+\nu(y)} \leq (M+2\Delta) e^{-\sqrt{Q}y_k + Q/2}, \qquad k = i, j. \quad (6.8)$$

The expression for $\nu^4(y)$ can be expanded into a sum of several terms. By choosing $k = i$ or j in (6.8) appropriate to the term involved and letting $Z \sim N(0,1)$ one gets after some simplification that

$$\begin{aligned} &\int \frac{\nu^4(y)}{1+\nu(y)} w(y)\, dy \\ &\leq (M+2\Delta)^{-3}\{(M-2)E(e^{\sqrt{Q}Z - Q/2} - 1)^4 \\ &\quad + 2e^Q E(e^{\sqrt{Q}Z-Q/2} - 1)^4 \\ &\quad + 6\binom{M-2}{2}(E(e^{\sqrt{Q}Z-Q/2} - 1)^2)^2 \\ &\quad + 12(M-2)E(e^{\sqrt{Q}Z-Q/2} - 1)^2 \\ &\quad \cdot E(e^{\sqrt{Q}Z-Q/2} - e^Q)^2 \\ &\quad + 6E(e^{\sqrt{Q}Z-Q/2} - e^Q)^2 \\ &\quad \cdot E((e^{\sqrt{Q}Z-Q/2} - e^Q)^2 e^{\sqrt{Q}Z-Q/2})\} \\ &= O(M^{-1}), \qquad \text{for fixed } Q. \end{aligned} \quad (6.9)$$

This verifies the upper bound in (4.5). Equation (4.6) follows directly from (4.5) and Lemma 4.1. □

Acknowledgment

The authors wish to thank J. Ziv and M. Zakai for several valuable discussions. The authors also wish to thank the helpful discussions from the associate editor, as well as the referees in the revision.

References

[1] S. Bellini and G. Tartara, "Bounds on error in signal parameter estimation," *IEEE Trans. Commun.*, vol. COM-22, pp. 340–342, Mar. 1974.
[2] L. D. Brown and R. C. Liu, "The asymptotic risk in a signal parameter estimation problem," to appear *IEEE Trans. Inform. Theory*, vol. 39, no. 1, pp. 254–257, Jan. 1993.
[3] ——, "Bounds on the Bayes and Minimax Risk for Signal Parameter Estimation," Cornell Tech. Rep., 1992.
[4] G. Casella and W. Strawderman, "Estimating a bounded normal mean," *Ann. Statist.*, vol. 9, pp. 870–878, 1981.
[5] D. Chazan, M. Zakai, and J. Ziv, "Improved lower bounds on signal parameter estimation," *IEEE Trans. Inform. Theory*, vol. 21, pp. 90–92, Jan. 1975.
[6] D. L. Donoho, R. C. Liu, and B. MacGibbon, "Minimax risks over hyperrectangles, and implications," *Ann. Statist.*, vol. 18, no. 3, pp. 1416–1437, 1990.
[7] W. Feller, *An Introduction to Probability Theory and Its Applications*, vol. I. New York: Wiley, 1970.
[8] I. A. Ibragimov and R. Z. Hasminiskii, *Statistical Estimation—Asymptotic Theory*. New York: Springer-Verlag, 1981.
[9] S. Karlin, *Total Positivity*. Stanford, CA: Stanford Univ. Press, 1968.
[10] J. A. Kuks and V. Olman, "A minimax estimator of regression coefficients," (in Russian), *Izv. Akad. Nauk. Eston. SSR*, vol. 21, pp. 66–72, 1972.
[11] L. M. Le Cam, *Asymptotic Methods in Statistical Decision Theory*. New York: Springer-Verlag, 1986.
[12] A. S. Terent'yev, "Probability distribution of a time location of an absolute maximum at the output of a synchronized filter," *Radioengineering and Electron.*, vol. 13, pp. 652–657, 1968.
[13] H. L. Van Trees, *Detection, Estimation, and Modulation Theory*. New York: Wiley, 1968.
[14] J. Ziv and M. Zakai, "Some lower bounds on signal parameter estimation," *IEEE Trans. Inform. Theory*, vol. IT-15, pp. 386–391, May 1969.