Error Bounds in Parameter Estimation Under Mismatch

Demetrios Kazakos, Kami S. Makki
College of Science and Technology, Texas Southern University
3100 Cleburne Str. Houston, Texas 77004, U.S.A
Department of Electrical Engineering & Computer Science, University of Toledo
2801 W. Bancroft, Toledo, Ohio 43606, U.S.A

Abstract— In this paper we develop a new upper bound for the mean square estimation error of a parameter that takes values on a bounded interval. The bound is based on the discretization of the region into a finite number of points, and the determination of the estimate by a maximum likelihood procedure. It is assumed that inaccurate versions of the true spectra are utilized in the implementation of the maximum likelihood estimator. As a special case, we develop new upper bounds to the performance of time delay estimation schemes.

I. Introduction

The development of performance measures for parameter estimation schemes is an important step for evaluating the performance of various estimators. Due to the analytical difficulties involved in the evaluation of the performance measures of interest, several authors have developed bounds to the performance measures that are relevant to the operational quality of a system. Within the class of parameter estimators from signals in the presence of noise, several authors in the engineering literature have applied lower bounds available from the statistical literature. Others have developed various improved lower bounds [1, 2, 3, 4]. All of these have been derived under the assumption of accurate knowledge of the statistical models of the problem.

In this paper we develop new upper bounds to the mean square estimation error. By discretizing the parameter space, we relate the mean square estimation error to the probability of error in a two-class hypothesis testing problem. We subsequently develop upper bounds to the probability of error when there is a mismatch between the actual and assumed statistical models. To be specific, we assume that inaccurate versions of the signal and noise spectra are employed in the discrete maximum likelihood estimation scheme.

II. The New Estimation Bounds

Suppose that a sequence of n observations $Y_n = \{y_1 ... y_n\}$ is available for the estimation of a parameter Δ, which is continuous and lies in the interval [0, D]. Let $p(Y_n \mid \theta, \Delta)$ be the probability density function of Yn conditioned on a specific value of Δ and a particular model parametrized by θ. In other words, all of the parameters of the problem for which there may be a mismatch, are summarized in θ. The only parameter we wish to estimate is Δ, and the algorithm to be utilized is the Discretized Maximum Likelihood. Instead of allowing Δ to take values in the continuous interval [0, D], we assume that it takes only one of the m uniformly spaced discrete values: $\{\Delta_k = D^{-1} \cdot k; k = 1,..., m\}$

Thus, the problem of estimating a continuous parameter Δ has been transformed into the problem of choosing one of m hypotheses. Let $\hat{\Delta}$ denote the estimate of Δ, derived by the Discrete Maximum Likelihood (DML) procedure. Then:

$$\hat{\Delta} = \Delta_k; \text{ if } p(Y_n \mid \theta, \Delta_k) = \max_{l \le j \le m} p(Y_n \mid \theta, \Delta_j) \tag{1}$$

We will assume that θ is the parameter value utilized in the DML estimator, while the true parameter θ_0 may be distinct from θ. We will also assume that a uniform prior distribution is valid for the true parameter Δ, i.e.: $p(\Delta = \Delta_j) = m^{-1}$

The mean square estimation error for our problem is:

$$E[\Delta - \hat{\Delta}]^2 = \sum_{k=1}^{m} \sum_{j=1}^{m} (\Delta_k - \Delta_j)^2 P_m(k, j) \tag{2}$$

where $P_m(k, j) = \Pr[\hat{\Delta} = \Delta_k, \Delta = \Delta_j]$ (3)

is the joint probability of the events $\hat{\Delta} = \Delta_k$ and $\Delta = \Delta_j$, and depends on the decision rule chosen. Here we assume that the DML has been utilized. We observe that:

$$P_m(k, j) = \Pr[p(Y_n \mid \theta, \Delta_k) = \max_{l \le i \le m} p(Y_n \mid \theta, \Delta_i) \mid \Delta = \Delta_j] \cdot \Pr[\Delta = \Delta_j] \le m^{-1} \Pr[p(Y_n \mid \theta, \Delta_k) \ge p(Y_n \mid \theta, \Delta_j) \mid \Delta = \Delta_j] = m^{-1} P_2(k \mid j) \tag{4}$$

$$P_m(j, k) = \Pr[p(Y_n \mid \theta, \Delta_j) = \max_{l \le i \le m} p(Y_n \mid \theta, \Delta_i) \mid \Delta = \Delta_k] \cdot \Pr[\Delta = \Delta_k] \le m^{-1} \Pr[p(Y_n \mid \theta, \Delta_j) \ge p(Y_n \mid \theta, \Delta_k) \mid \Delta = \Delta_k] = m^{-1} P_2(j \mid k) \tag{5}$$

where $P_2(k \mid j)$ is the probability of deciding $\hat{\Delta} = \Delta_k$ when $\hat{\Delta} = \Delta_j$ is true, for a two-hypotheses case. In other words, the error probability $P_m(k, j)$ for m hypotheses has been bounded from above by the error probability $P_2(k \mid j)$ for the two-hypotheses problem. The reason is that the latter can be bounded from above, using bounds that will be developed in the sequel. The mean square estimation error is bounded as follows:

$$E[\Delta - \hat{\Delta}]^2 \le \sum_{k=1}^{m} \sum_{j=1}^{m} (\Delta_k - \Delta_j)^2 (2m)^{-1} [P_2(k \mid j) + P_2(j \mid k)] = \tag{6}$$

$$= m^{-1} \sum_{k=1}^{m} \sum_{j=1}^{m} (\Delta_k - \Delta_j)^2 P_2(k, j)$$

Note that:

$$P_2(k, j) = 2^{-1} [P_2(k \mid j) + P_2(j \mid k)] \tag{7}$$

is the average probability of error in a two-class hypothesis testing problem, with the assumption that a Maximum Likelihood decision rule is used for testing between a pair of values $\hat{\Delta} = \Delta_k$ and $\hat{\Delta} = \Delta_j$.

III. DETECTION BOUNDS UNDER MISMATCH

In this section we will develop bounds to the error probabilities for a two-hypotheses detection problem, when a suboptimal threshold type detection scheme is used. The objective is to find an upper bound to the probabilities $P_2(k \mid j)$ of (7). Our development in the present section generalizes this lower bound, for the situation where there is a condition of mismatch or suboptimal detector structure.

Suppose that there are two hypotheses, H1 and H0. Let L be the test statistic and A the threshold value. The decision rule is:

Decide H1 if L > A, otherwise Decide H0

We denote by P1(L), P0(L) the probability density functions of L under H1, H0 respectively. The two error probabilities are:

$$P_2(1 \mid 0) = \int_A^{+\infty} p_0(L)dL;$$

$$P_2(0 \mid 1) = \int_{-\infty}^{A} p_1(L)dL = \int_B^{+\infty} q_1(J)dJ$$

where: $q_1(J) = p_1(-L),\ B = -A,\ J = -L$

We define the logarithms of two moment generating functions:

$$M_0(s) \stackrel{\Delta}{=} \ln E[e^{sL} \mid H_0] = \ln \int_{-\infty}^{+\infty} e^{sL} p_0(L)dL \tag{8}$$

$$M_1(s) \stackrel{\Delta}{=} \ln E[e^{sJ} \mid H_1] = \ln \int_{-\infty}^{+\infty} e^{sJ} q_1(J)dJ = \ln \int_{-\infty}^{+\infty} e^{-sL} p_1(L)dL \tag{9}$$

and two new probability density functions:

$$p_{0s}(L) = p_0(L)e^{sL - M_0(s)} \tag{10}$$

$$q_{1s}(J) = q_1(J)e^{sJ - M_1(s)} \tag{11}$$

By direct calculation we can verify the relationships:

$$\dot{M}_0(s) = M_0^{-1}(s)\int_{-\infty}^{\infty} Le^{sL} p_0(L)dL$$

$$\dot{M}_1(s) = M_1^{-1}(s)\int_{-\infty}^{\infty} Je^{sJ} q_1(J)dJ$$

$$\dot{M}_0(0) = E[L \mid H_0];\quad \dot{M}_1(0) = E[J \mid H_1]$$

$$\ddot{M}_0(s) = \int_{-\infty}^{\infty} (L - \dot{M}_0(s))^2 p_{0s}(L)dL > 0$$

$$\dot{M}_1(s) = \int_{-\infty}^{\infty} (J - \dot{M}_1(s))^2 q_{1s}(J)dJ > 0$$

Furthermore, we have the error probability expressions:

$$P_2(1 \mid 0) = \int_A^{+\infty} p_0(L)dL = \int_A^{+\infty} p_{0s}(L)e^{M_0(s) - sL}dL \tag{12}$$

$$P_2(0 \mid 1) = \int_B^{+\infty} q_1(J)dJ = \int_B^{+\infty} q_{1s}(J)e^{M_1(s) - sJ}dJ \tag{13}$$

The functions M1(s), M0(s) are convex and take the value 0 at s = 0.

We will now seek a bound for $P_2(1 \mid 0)$. Suppose that there is a positive s, for which: $\dot{M}_0(s)$=A

Substituting (10) in (12), we find:

$$P_2(1 \mid 0) = \exp[M_0(s) - s\dot{M}_0(s)] \int_{\dot{M}_0(s)}^{+\infty} P_{0s}(L)\exp\{s[\dot{M}_0(s) - L]\}dL \tag{14}$$

If we define the new random variable

$$z = [L - \dot{M}_0(s)][\ddot{M}_0(s)]^{-1/2} \tag{14a}$$

it can be easily shown that $E[z] = 0, Ez^2 = 1$

Hence, (14) takes the form:

$$P_2(1 \mid 0) = \exp[M_0(s) - s\dot{M}_0(s)] \int_0^{+\infty} f_z(z)\exp[-sz(\ddot{M}_0(s))^{1/2}]dz \tag{15}$$

where z is a random variable with zero mean and unit variance.

The integral in (15) is less than 1, so we find the following upper bound:

$$P_2(1 \mid 0) \le \exp[M_0(s) - s\dot{M}_0(s)] \tag{16}$$

where s is a positive number, satisfying the equation

$$\dot{M}_0(s) = A. \tag{17}$$

We must investigate under what conditions (17) has a positive solution. Note that M0(s) is convex, and M0(0) = 0. Also:

$\dot{M}_0(0)$=E[L|H0].

Thus, a necessary and sufficient condition for eq. (17) to have a positive solution s, is that:

$$A > E[L|H0]. \tag{18}$$

If (18) is satisfied, the convexity of M0(s) guarantees that M0(s) − $s\dot{M}_0(s) < 0$, and that the bound (16) is less than 1.

In a completely analogous manner, dealing with the random variable J rather than L, we find the expression:

$$P_2(0 \mid 1) = \exp[M_1(s) - s\dot{M}_1(s)] \int_0^{+\infty} f_w(w)\exp[-sw[\ddot{M}_1(s))^{1/2}]dw \tag{19}$$

where

$$w = [J - \dot{M}_1(s)][\ddot{M}_1(s)]^{-1/2} \tag{19a}$$

Ew=0, Ew2=1.

The value of s must be the positive solution of the equation

$$\dot{M}_1(s) = B = -A \tag{20}$$

assuming such a positive solution exists. To investigate under what conditions eq. (20) has a positive s solution, we observe that M1(0) = 0, and that M1(s) is convex in s. Also:

$$\dot{M}_1(0) = E[J \mid H_1] = -E[L \mid H_1].$$

Thus, the condition for having a positive solution to (20) is

M1(0) < B, or: $E[L \mid H_1] > A$. (21)

Since the integral in (19) is less than 1, we have the bound:

$$P_2(0 \mid 1) \le \exp[M_1(s) - s\dot{M}_1(s)] \tag{22}$$

as where s is the unique positive solution of:

$\dot{M}$ 1(s) = -A

If (21) holds, the exponent in (22) is negative, and we have a useful bound.

For the case of optimal detection, we have L = log[p1(x)/p0(x)] and

$$M_0(s) = \ln \int_{-\infty}^{+\infty} [p_1(x)]^s [p_0(x)]^{1-s} dx$$

$$M_1(s) = M_0(1 - s).$$

If we define M(s) = M0(s), then the two bounds (16), (22) for the optimal detection case, take the form:

$$P_2(1 \mid 0) \le \exp[M\ (s) - s\dot{M}(s)2] \tag{23}$$

$$P_2(0 \mid 1) \le \exp[M\ (s) - (1 - s)\dot{M}(s)] \tag{24}$$

where s > 0, $\dot{M}$ (s) =A;

The bounds (23), (24) are a special case of (16), (22) for the situation of optimal detection. They are available in the literature [6, [7]. The generalized bounds (16), (19), for the case of mismatch, are new.

Suppose now that the test statistic L is constructed from a set of observations, $(y_1 \ldots y_n) \stackrel{\Delta}{=} Y_n$, distributed according to one of two true probability density functions, {pn(Yn|Hk); k=0, 1}. The test statistic L(Yn) has the form:

$$L(Y_n) = \ln[q_n(Y_n \mid H_1)/q_n(Y_n \mid H_0)] \quad (25)$$

where $\{q_n(Y_n H_k)$; k=0, 1}, are inaccurate versions of $\{p_n(Y_n \mid H_k)$; k=0, 1}, respectively.

The binary decision rule for testing between hypotheses H1, H0 based on Yn and using the test statistic L(Yn), as the form:

Decide H1 if L(Yn) > nA, otherwise Decide H0

Note that the threshold is nA, which is equivalent to comparing $n^{-1}L(Y_n)$ against a fixed threshold A.

The logarithms of the moment generating functions have the forms:

$$\mathrm{M}_{0,n}(s) \stackrel{\Delta}{=} \ln \mathrm{E}[\exp(s\mathrm{L}(\mathrm{Y}_n)) \mid \mathrm{H}_0] \quad (26)$$

$$\mathrm{M}_{1,n}(s) \stackrel{\Delta}{=} \ln \mathrm{E}[\exp(-s\mathrm{L}(\mathrm{Y}_n)) \mid \mathrm{H}_1] \quad (27)$$

$$\mathrm{M}_{0,n}(s) = \ln \int [q_n(\mathrm{Y}_n \mid \mathrm{H}_1)/q_n(\mathrm{Y}_n \mid \mathrm{H}_0)]^s p_n(\mathrm{Y}_n \mid \mathrm{H}_0) d\mathrm{Y}_n \quad (28)$$

$$\mathrm{M}_{1,n}(s) = \ln \int [q_n(\mathrm{Y}_n \mid \mathrm{H}_0)/q_n(\mathrm{Y}_n \mid \mathrm{H}_1)]^s p_n(\mathrm{Y}_n \mid \mathrm{H}_1) d\mathrm{Y}_n \quad (29)$$

Suppose now that for the stochastic models under consideration, the limiting expressions

$$\mathrm{M}_0(s) \stackrel{\Delta}{=} \lim_{n\to\infty} n^{-1}\mathrm{M}_{0,n}(s); \mathrm{M}_1(s) \stackrel{\Delta}{=} \lim_{n\to\infty} n^{-1}\mathrm{M}_{1,n}(s) \quad (30)$$

exist. Then, as n becomes large, we have the approximations:

$$\mathrm{M}_{0,n}(s) \cong n\mathrm{M}_0(s), \mathrm{M}_{1,n}(s) \cong n\mathrm{M}_1(s). \quad (31)$$

From previous arguments, we conclude that in order for the upper bounds (16), (22) to be less than 1, the threshold nA must be in the interval:

$$\dot{\mathrm{M}}_{0,n}(s) < n\mathrm{A} < -\dot{\mathrm{M}}_{1,n}(0) \quad (32)$$

(32) implies that the right hand side of the inequality must be larger than the left hand side. We now define the informational divergence measures:

$$I_n(p_n(Y_n \mid H_k), q_n(Y_n \mid H_j)) = \int p_n(Y_n \mid H_k) \ln[p_n(Y_n \mid H_k)/q_n(Y_n \mid H_j)] dY_n \quad (33)$$

In is nonnegative and takes the value of 0 if and only if the density functions pn and qn are equal almost everywhere. Thus, it is a measure of distance between the two densities.

Combining (32), (33), (28), (29), we find that (32) takes the form:

$$\begin{aligned} &n^{-1}\mathrm{I}_n[p_n(\mathrm{Y}_n \mid \mathrm{H}_0), q_n(\mathrm{Y}_n \mid \mathrm{H}_0)] - n^{-1}\mathrm{I}_n[p_n(\mathrm{Y}_n \mid \mathrm{H}_0), q_n(\mathrm{Y}_n \mid \mathrm{H}_1)] < \mathrm{A} < \\ &< n^{-1}\mathrm{I}_n[p_n(\mathrm{Y}_n \mid \mathrm{H}_1), q_n(\mathrm{Y}_n \mid \mathrm{H}_0)] - n^{-1}\mathrm{I}_n[p_n(\mathrm{Y}_n \mid \mathrm{H}_1), q_n(\mathrm{Y}_n \mid \mathrm{H}_1)]. \end{aligned} \quad (34)$$

If the limiting expressions (30) exist, then the limiting expressions

$$\mathrm{I}(p^k, q^j) \stackrel{\Delta}{=} \lim_{n\to\infty} n^{-1}\mathrm{I}_n[\mathrm{P}_n(\mathrm{Y}_n \mid \mathrm{H}_k), q_n(\mathrm{Y}_n \mid \mathrm{H}_j)] \quad (35)$$

exist as well [4]. In (35) we use pk, qj to denote the generic statistical models. As $n \to \infty$, (34) becomes:

$$\dot{M}_0(0) = I(p^0, q^0) - I(p^0, q^1) < A < I(p^1, q^0) - I(p^1, q^1) = -\dot{M}_1(0) \quad (36)$$

Finally, if we assume (36) to be active, the bounds (16), (22) take the form:

$$\mathrm{P}_2(1 \mid 0) \le \exp n[\mathrm{M}_0(s) - s\dot{\mathrm{M}}_0(s)] \quad (37)$$

$$\mathrm{P}_2(0 \mid 1) \le \exp n[\mathrm{M}_1(v) - v\dot{\mathrm{M}}_1(v)] \quad (38)$$

where the positive numbers s, v satisfy the equations:

$$\dot{\mathrm{M}}_0(s) = \mathrm{A} > \dot{\mathrm{M}}_0(0),\ \dot{\mathrm{M}}_1(v) = -\mathrm{A} > \dot{\mathrm{M}}_1(0) \quad (39)$$

and the function Mk(s) is the limiting value of n-1Mk,n(s), as defined by eq. (30). The bounds (37), (38) are valid for two hypotheses, and are less than 1 and hence meaningful, if inequalities (36) are satisfied. The threshold A can be set equal to zero, if the value A=0 satisfies inequalities (36). For A=0, the conditions for the bounds (37), (38) to be less than 1, are:

$$\dot{M}_0(0) = I(p^0, q^0) - I(p^0, q^1) < 0 \quad (40)$$

$$\dot{\mathrm{M}}_1(0) = \mathrm{I}(p^1, q^1) - \mathrm{I}(p^1, q^0) < 0 \quad (41)$$

For this case, the bounds (37), (38) take the form:

$$\mathrm{P}_2(1 \mid 0) \le \exp\cdot n\mathrm{M}_0(s); \dot{\mathrm{M}}_0(s) = 0 \quad (42)$$

$$\mathrm{P}_2(0 \mid 1) \le \exp\cdot n\mathrm{M}_1(v); \dot{\mathrm{M}}_1(v) = 0 \quad (43)$$

Note that the log-moment generating functions M0(s), M1(v) are determined for a pair of statistical hypotheses. In order to utilize (42), (43) in the bound (7), we must change notation. Let:

$$\mathrm{M}_{kj,n}(s) \stackrel{\Delta}{=} \ln \int [q_n(\mathrm{Y}_n \mid \mathrm{H}_k)/q_n(\mathrm{Y}_n \mid \mathrm{H}_j)]^s p_n(\mathrm{Y}_n \mid \mathrm{H}_j) d\mathrm{Y}_n \quad (44)$$

$$\mathrm{M}_{kj}(s) \stackrel{\Delta}{=} \lim_{n\to\infty} n^{-1}\mathrm{M}_{kj,n}(s) \quad (45)$$

Under the stated notation, the bound (7) becomes:

$$\mathrm{E}[\Delta - \hat{\Delta}]^2 \le (2m)^{-1} \sum_{k=1}^{m} \sum_{j=1}^{m} (\Delta_k - \Delta_j)^2 [\exp\cdot n\mathrm{M}_{kj}(s_{kj}) + \exp\cdot n\mathrm{M}_{jk}(s_{jk})] \quad (46)$$

where:

$$\dot{\mathrm{M}}_{kj}(s_{kj}) = 0; k, j = 1, \ldots, m;\ k \ne j$$

In order for values $\{M_{kj}(s_{kj})\}$ in the exponents of the bound (46) to be negative, the set of necessary and sufficient conditions is:

$$\dot{\mathrm{M}}_{kj}(0) = \mathrm{I}(p^j, q^j) - \mathrm{I}(p^j, q^k) < 0 \quad (47)$$

for all $j \ne k$; j, k=1, …, m

Conditions (47) require that the “approximation” qj to the real model pj must be closer to it than any other approximating model qk, in the sense of the information divergence I.

IV. GAUSSIAN PROCESS MODELS IN DISCRETE TIME

In this section we will consider a class of stationary Gaussian processes. For the case of discrete time data, discussed in the previous section, {yi; i=1, …, n} is a sequence of d-dimensional observations of a Gaussian, stationary, d-dimensional process. Let $F_k(\lambda), G_k(\lambda)$, $\lambda \varepsilon [-\pi, \pi]$ be the true and assumed spectral density matrix of the process {yi} under hypothesis Hk. In [5], we demonstrated that in this case the limits (45) exist, and are expressed in terms of the spectra $\{F_k(\lambda), G_k(\lambda)$; k=1, …, m}. The expressions are of the form: [5]

$$\begin{aligned} \mathrm{M}_{kj}(s) = -(4\pi)^{-1} \int_{-\pi}^{\pi} \{&\ln | \mathrm{F}_j^{-1}(\lambda) + s \cdot \mathrm{G}_k^{-1}(\lambda) - s \cdot \mathrm{G}_j^{-1}(\lambda) | - \\ &- s \cdot \ln | \mathrm{G}_j(\lambda) | + s \cdot \ln | \mathrm{G}_k(\lambda) | + \ln | \mathrm{F}_j(\lambda) | \} d\lambda \end{aligned} \quad (48)$$

Equation (48) is valid for $0 \le s \le s_0$, where s0 is the maximum positive value of s for which the first determinant is positive for all $\lambda \varepsilon [-\pi, \pi]$.

The limiting expression for the informational divergence (35), for the Gaussian case, has the form: [5]

$$I(p^j, q^k) = (2\pi)^{-1} \int_{-\pi}^{\pi} \{\mathrm{trace} F_j(\lambda) G_k^{-1}(\lambda) - \ln | F_j(\lambda) G_k^{-1}(\lambda) | - d\} d\lambda =$$

$$\stackrel{\Delta}{=} I(F_j, G_k) \tag{49}$$

The condition for all of the exponents in the bound (46) to be negative, is:

$$I(F_j, G_j) < I(F_j, G_k); \text{for all } k \neq j \tag{50}$$

It should be noted that condition (50) is also necessary and sufficient for the mean square estimation error to converge to zero as $n \to \infty$. This conclusion can be inferred from the theory in[5], where it was shown that (50) is a necessary and sufficient condition for the probability of error to converge to zero, in a multihypothesis testing situation, when inaccurate statistical models are used in a maximum likelihood decision rule. If (50) is not satisfied, and hence:

$$I(F_j, G_j) > I(F_j, G_k); \text{for some } k \neq j$$

then $P_2(k \mid j) \to 1$ as $n \to \infty$ [5]. As a consequence, the mean square error is then bounded from below by $(\Delta_k - \Delta_j)^2$.

V. Approximations to the Error Probability

In this section we consider approximations to the error probabilities $P_2(1|0)$, $P_2(0|1)$, for a two-class hypothesis testing problem. Our starting point is the exact expressions (14) and (19). We utilize the additional assumption that the random variables z, w, as defined by (14a), (19a) are approximately Gaussian. This situation occurs when a large number of samples are involved in developing the test statistics L, J, and thus the central limit theorem can be invoked, ensuring the approximate Gaussianity of z, w.

For Gaussian z, w, we have, from (15), (19):

$$(2\pi)^{-1/2} \int_0^{+\infty} \exp[-z^2/2 - s(\ddot{M}_k(s))^{1/2} \cdot z] dz = \tag{51}$$

$$= \exp[s^2 \ddot{M}_k(s)/2] \cdot \phi[s(\ddot{M}_k(s))^{1/2}]$$

for k = 0, 1, where: $\phi(x) = (2\pi)^{-1/2} \int_x^{-\infty} \exp(-z^2/2) dz$

For A = B = 0, utilizing (6) in (15), (19), we find:

$$P_2(1 \mid 0) = \exp[M_0(s_0) + \frac{s_0^2}{2} \ddot{M}_0(s_0)] \cdot \phi[s_0(\ddot{M}_0(s_0))^{1/2}] \tag{52}$$

where: $\dot{M}_0(s_0) = 0$

and $P_2(0 \mid 1) = \exp[M_1(s_1) + \frac{s_1^2}{2} \ddot{M}_1(s_1)] \cdot \phi[s_1[\ddot{M}_1(s_1))^{1/2}]$ (53)

where: $\dot{M}_1(s_1) = 0$. Equations (52), (53) are generalizations for the case of mismatch, of the original expressions found in Van Trees [8]. If we utilize the exact expressions (52), (53) instead of the bounds (37), (38) in the bounding expression (7) for the mean square error, we develop tighter upper bounds to the mean square error, under the condition of Gaussian test statistic.

For the case of accurately known spectra, Weiss and Weinstein [9] developed lower bounds to the mean square estimation error in estimating a time delay between two sensors. They utilize approximations to the error probability in a two class decision problem. Their approximations are special cases of (52), (53) for the case of known statistics.

VI. Conclusions

In this paper we have developed a new upper bound to the mean square error in estimating a bounded parameter. We assumed that a discretized maximum likelihood estimation rule was utilized, and we related the mean square estimation error to the probability of detection error in a two-class detection problem.

The new upper bound is valid when inaccurate versions of the signal spectra are utilized in the estimation scheme. We also establish the degree of inaccuracy that can be tolerated, and still attain improved performance for increasing amounts of observations.

Our bounds are applicable for Gaussian statistical models, for both discrete and continuous time observations. The work of other investigators has basically concentrated on lower bounds to the mean square estimation error, and under the assumption of exactly known statistical models. In the specific case of time delay estimation [7], most of the research has been directed to the development of lower bounds, and the case of high signal to noise ratio has been mainly emphasized. The results of the present paper are applicable for arbitrary signal to noise ratio.

Extensions and generalizations of the present results will be pursued in the near future. One direction is the departure from purely Gaussian statistics, and the inclusion of impulse noise modeled as a point process. A second direction is the design of robust estimation schemes for time delay estimation problems.

VII. References

[1] S. Basu and Y. Bresler. "A Global Lower Bound on Parameter Estimation Error with Periodic Distortion Functions". IEEE *Transactions on Information Theory*, 46(3): pp. 1145-50, 2000.

[2] K. L. Bell, Y. Ephraim, and H. L. Van Trees, "Explicit Ziv-Zakai lower bound for bearing estimation," IEEE Trans. Signal Processing, vol. 44, pp. 2810–2824, Nov. 1996.

[3] J. A. Gubner, "Distributed Estimation and Quantization", *IEEE Transactions on Information Theory*, IT39 (4):1456--1459, July 1993.

[4] R. J. McAulay and E. M. Hofstetter, "Barankin Bounds on Parameter Estimation", *IEEE Transactions on Information Theory*, vol. 1T-17, no. 6, Nov. 1971, pp. 669-676.

[5] D. Kazakos, "Statistical Discrimination Using Inaccurate Models", IEEE *Transactions on Information Theory*, vol. IT-28, No. 5, Sept. 1982, pp. 720-728.

[6] D. Kazakos, "Spectral Distance Measures Between Continuous-Time Vector Gaussian Processes", IEEE *Transactions on Information Theory*, vol. IT-28, no. 4, July 1982, pp. 679-681.

[7] Special Issue of IEEE Transactions on Acoustics, Speech and Signal Processing, on Time Delay Estimation. Part II of the June 1981 issue.

[8] H. L. Van Trees. Detection, estimation and modulation theory. Part 1, New York, Wiley, 1968, Ch. 2.

[9] A. Weiss and E. Weinstein, "Composite Bound on the Attainable Mean-Square Error in Passive Time-Delay Estimation from Ambiguity Prone Signals", IEEE Trans. on IT, vol. IT-28, no. 6, Nov. 1982, pp. 977-979.