EFFICIENT COMPUTATION OF THE BAYESIAN CRAMER-RAO BOUND ON
ESTIMATING PARAMETERS OF MARKOV MODELS

Joseph Tabrikian and Jeffrey L. Krollik

Department of Electrical and Computer Engineering,
Duke University, Durham, NC 27708-0291, USA

ABSTRACT
This paper presents a novel method for calculating the Hybrid Cramer-Rao lower bound (HCRBL) when the statistical model for the data has a Markovian nature. The method applies to both the non-linear/non-Gaussian as well as linear/Gaussian model. The approach solves the required expectation over unknown random parameters by several one-dimensional integrals computed recursively, thus simplifying a computationally-intensive multi-dimensional integration. The method is applied to the problem of refractivity estimation using radar clutter from the sea surface, where the backscatter cross section is assumed to be a Markov process in range. The HCRBL is evaluated and compared to the performance of the corresponding maximum a-posteriori estimator. Simulation results indicate that the HCRBL provides a tighter lower bound in this application.

1. INTRODUCTION
Bayesian bounds are used for evaluating the performance limitations in cases where the unknown parameters are random with known prior distribution. A major difficulty in calculating such bounds is computing the required expectation over the Fisher Information Matrix (FIM) with respect to the unknown random parameters. Such an expectation involves a multi-dimensional integration over these parameters. In cases where the expectation cannot be performed analytically, numerical or Monte-Carlo integration methods may be applied. However, in the presence of a large number of unknown parameters, such a solution is not viable because of the large number of computations required. In the presence of prior statistical information on some but not all the parameters, the hybrid form of Cramer-Rao lower bound (HCRBL) has previously been proposed [5]. The HCRBL is a variation of the Bayesian CRLB [10], but also requires a similar expectation over the random parameters. In [5] this problem is handled by assuming that the FIM corresponding to the measurements only (i.e. \( J_1 \)) is constant with respect to the random parameters involved in the multi-dimensional integral. In fact, under this assumption, no expectation over the random parameters is required, and derivation of the bound becomes straightforward. However, in most cases such an approximation cannot be justified.

Cramer-Rao bounds for Markov processes have been proposed assuming linear and Gaussian processes [4] where derivation of the bound can be performed analytically. In this paper, a novel method for calculating the Hybrid CRLB for non-Gaussian and/or non-linear cases is proposed. The method exploits the Markov property of the unknown random variables, and involves replacing the multi-dimensional integration with several one-dimensional ones, thereby significantly reducing computation required to calculate the bound.

2. THE HYBRID CRLB
Let \( s \) be a non-random vector and \( y, s \) be vector valued random variables with conditional probability density function (pdf)
\[
\begin{align*}
&f_{y,s|g}(y,s|g) = f_{y,s|g}(y,s|g)
\prod_{k=1}^{N} f_{g_k|s_k}(g_k|s_k) f_{s_k|s_{k-1}}(s_k|s_{k-1}, g_k).
\end{align*}
\]

The goal here is to derive the hybrid forms of CRLB on the estimation error of the random and non-random parameters \( \Theta \triangleq [s^T, g^T]^T \) using the measurements \( y \) and the a-priori statistical information on s.

Let \( \hat{\Theta} \) denote any unbiased estimator of \( \Theta \). Then, the HCRBL on the estimation errors of the vector \( \Theta \) is [10]
\[
\text{cov}(\hat{\Theta}) \geq (E_y [J_1(\Theta)] + J_2)^{-1}
\]
where \( J_1(\Theta) \) is the FIM of \( \Theta \) using the measurements \( y \), and it is given by
\[
J_1(\Theta) = E_{y|\Theta} \left\{ \frac{\partial \log f_y(g|\Theta)}{\partial \Theta} \right\}^T \frac{\partial \log f_y(g|\Theta)}{\partial \Theta},
\]
where the derivative of a scalar with respect to a vector \( \Theta \triangleq [\theta_1, \ldots, \theta_L]^T \), is defined as \( \frac{\partial \log f_y(g|\Theta)}{\partial \Theta} \). The matrix \( J_2 \) stands for the FIM of the vector s based on its prior statistical information.

In many papers concerning applications of the Bayesian or Hybrid CRLB, the prior distribution of the unknown random vector parameters has been assumed to be Gaussian. In this case, the
elements of the matrix $J_2$ are simply the covariance matrix of the corresponding random vector. It can be shown that if the prior distribution of the random parameters is uniform, the elements of the matrix $J_2$ are equal to zero.

For calculation of the bound, the matrix $J_1$ is required, and it is derived as follows. The matrix $J_1(\Theta)$ can be partitioned as

$$J_1(\Theta) = \begin{bmatrix} J_{ss} & J_{sg} \\ J_{gs}^T & J_{gg} \end{bmatrix}$$

(4)

By substituting the expression for $f_{yi|s,g}(y|s,g)$, and using the conditional independence between the elements of $y$, the submatrices in (4) can be expressed as

$$[J_{ss}]_{ij} = U_{si}U_{sj}(1 - \delta_{ij}) + U_{s=\partial ij}$$

(5)

$$[J_{sg}] = \sum_{i=1}^{N} U_{si}u_{gi}(1 - \delta_{ii}) + u_{s=gi}$$

(6)

$$J_{gg} = \sum_{i=1}^{N} \sum_{j=1}^{N} U_{gi}u_{gj}(1 - \delta_{ii}) + \sum_{k=1}^{N} U_{g=ki}$$

(7)

where

$$U_{si} \triangleq E_{y_i|s,g} \left\{ \frac{\partial \log f_{y_i|s,g}(y_i|s_i,g)}{\partial s_i} \right\}$$

(8)

$$U_{s=gi} \triangleq E_{y_i|s,g} \left\{ \left( \frac{\partial \log f_{y_i|s,g}(y_i|s_i,g)}{\partial s_i} \right)^2 \right\}$$

(9)

$$u_{gi} \triangleq E_{y_i|s,g} \left\{ \frac{\partial \log f_{y_i|s,g}(y_i|s_i,g)}{\partial g} \right\}$$

(10)

$$u_{s=gi} \triangleq E_{y_i|s,g} \left\{ \frac{\partial \log f_{y_i|s,g}(y_i|s_i,g)}{\partial s_i} \frac{\partial \log f_{y_i|s,g}(y_i|s_i,g)}{\partial g} \right\}$$

(11)

$$U_{g=ki} \triangleq E_{y_i|s,g} \left\{ \left( \frac{\partial \log f_{y_i|s,g}(y_i|s_i,g)}{\partial g} \right)^T \frac{\partial \log f_{y_i|s,g}(y_i|s_i,g)}{\partial g} \right\}$$

(12)

In the next section, a recursive method for performing the expectations in expressions in (8)-(12) is presented.

3. THE RECURSIVE INTEGRATION METHOD

In this section, the Markov property of the random unknown parameters is used in order to obtain a simple procedure for computing the expectations in (8)-(12), thus solving the required N-dimensional integration by a series of N one-dimensional integrals.

Note that each one of these matrices $J_{ss}$, $J_{sg}$ and $J_{gg}$ can be expressed as linear combinations of terms of the form $\prod_{k=0}^{N} V_{ij}^{(k)}(s_k)$, where for example, the $ij$th element of the matrix $J_{ss}$ is given by

$$V_{ij}^{(k)}(s_k) = \begin{cases} U_{ik} & \text{for } k = i = j \\ U_{ik} & \text{for } k = i \neq j \\ 1 & \text{otherwise} \end{cases}$$

(13)

Similarly, the matrices $J_{sg}$ and $J_{gg}$ can also be expressed by a linear combination of the terms of the form $\prod_{k=0}^{N} V_k(s_k)$. Performing the expectation over $J_1(\Theta)$, thus involves computing of expectations over terms of the form $\prod_{k=0}^{N} V_k(s_k)$. In what follows, a fast method for calculating $E_x \{ \prod_{k=0}^{N} V_k(s_k) \}$ is derived.

Using the Markov property of $s$, the expected value of $\prod_{k=0}^{N} V_k(s_k)$ after having integrated over $s_0$, is

$$E_x \{ \prod_{k=0}^{N} V_k(s_k) \} = \int_{s_0} f_{s_0}(s_0) V_0(s_0) f_{s_1|s_0}(s_1|s_0) f_{s_2|s_1}(s_2|s_1) \cdots f_{s_N|s_{N-1}}(s_N|s_{N-1}) V_N(s_N) h_1(s_1) ds_1 \cdots ds_N$$

(14)

where $h_1(s_1)$ is defined as

$$h_1(s_1) \triangleq V_1(s_1) \int_{s_0} \int_{s_0} f_{s_0}(s_0) f_{s_1|s_0}(s_1|s_0) f_{s_2|s_1}(s_2|s_1) \cdots f_{s_N|s_{N-1}}(s_N|s_{N-1}) ds_0 \cdots ds_N$$

(15)

Similarly, by performing the integrals over $s_1, \ldots, s_{N-1}$ in (14), one obtains a recursive method for computing of $E_x \{ \prod_{k=0}^{N} V_k(s_k) \}$ by writing:

$$E_x \{ \prod_{k=0}^{N} V_k(s_k) \} = \int_{s_N} h_N(s_N) ds_N$$

(16)

where

$$h_m(s_m) \triangleq V_m(s_m) \int_{s_{m-1}} h_{m-1}(s_{m-1}) f_{s_m|s_{m-1}}(s_m|s_{m-1}) ds_{m-1}, \quad m = 2, \ldots, N$$

(17)

and $h_1(s_1)$ is given by (15).

The above recursive method can further be simplified by noting that

$$V_k(s_k) = 1 \quad \text{for } k = 0, \ldots, q,$$

(18)

then

$$h_k(s_k) = f_{s_k|s_k}(s_k|g), \quad \text{for } k = 0, \ldots, q,$$

$$h_{q+1}(s_{q+1}) = V_{q+1}(s_{q+1}) f_{s_{q+1}|s_{q+1}}(s_{q+1}|g).$$

(19)

Therefore, no integration is required for computing $h_1(s_1), \ldots, h_{q+1}(s_{q+1})$. For a stationary sequence, $f_{s_k|s_k}(s_k|g)$ is independent of $k$, and it may be known or computed. Therefore under the condition (18), if the density functions $f_{s_k|s_k}(s_k|g)$ are known, then only $N - q$ integrations are required for calculating $h_{q+2}, \ldots, h_N$ and the integration in (16). For instance, for the $ij$th element of the matrix $J_{ss}$, (13) implies that the condition (18) is satisfied by setting $q = \min(i,j) - 1$. Thus, for calculating the $ij$th element of $J_{ss}$, $(N - \min(i,j) + 1)$ integrals need to be computed.

Let $Q$ denote the number of operations required for calculating one integral, and $K$ be the size of the unknown deterministic vector, $g$. Then, the number of operations for calculating each element in $E_x(J_1)$ is $Q^{N+1}$. By comparison, since $J_1$ is a symmetric matrix of size $N + K + 1$, the total number of required operations using a brute-force method would be $(N + K + 1)(N + K + 2)Q^{N+1}$.

The number of operations required for calculating the expectation over $J_1$ with respect to $s$ is derived in [8]. For $N \gg K \geq 1$, the number of required operations using the conventional versus the recursive integration is approximately $N^2/2Q^{N+1}$ versus $N^2/6(K^2 + 5K + 2)Q$, respectively. This comparison demonstrates that the proposed algorithm significantly reduces the number of required operations for calculating the bound.

The matrix $J_2$ can also be computed using the proposed recursive method by a similar approach.
4. APPLICATION: THE HCRLB ON REFRACTIVITY FROM CLUTTER ESTIMATION

In this section, the proposed method is used for calculating the HCRLB for the problem of refractivity estimation using clutter from sea surface. The refractivity profile in coastal regions to a large extent determines the performance of shipboard radar and communications systems. If the atmospheric conditions, particularly the water vapor spatial distribution, were known, numerical propagation models could be used for such purposes as predicting detection ranges, correcting altitude estimates, and estimating surface backscatter strength. The effect of the variability of atmospheric refractivity on propagation estimates using real data profiles has been investigated in [6]. In [3], a MAP approach for estimating refractivity from clutter (RFC) is proposed when both the refractivity profile and range-dependent backscatter cross-section (BSCS) are unknown. The BSCS as a function of range is considered in [3] to be a Markov process.

The radar clutter measurement, sampled at range bins \( \{x_k\}_{k=1}^N \), in radar “slant” range, is given by [3]:

\[
y_k = L(x_k; g)(a(x_k) + \eta(x_k)) + \eta_k, \quad k = 1, \cdots, N,
\]

where \( L(x_k; g) \) is the two-way propagation loss which depends on an unknown refractivity profile parameter vector, \( g \). The clutter return, \( a = [a(x_1), \cdots, a(x_N)]^T \), is a complex random vector whose elements have variances \( \sigma_a^2(x_k) \equiv E(|a(x_k)|^2) \) for \( k = 1, \cdots, N \). The additive noise components \( \eta(x_1), \cdots, \eta(x_N) \), are assumed to be i.i.d., complex Gaussian random variables with variance \( \sigma_\eta^2 \) and zero-mean. The elements of the clutter return vector \( a \), given the BSCS’s \( \sigma_a^2(x_1), \cdots, \sigma_a^2(x_N) \), are i.i.d., with approximately zero-mean complex Gaussian distribution. The unknown variance of \( a(x_k) \) is the BSCS of the sea surface at the \( k \)th range bin and it depends on the sea surface roughness, incident grazing angle and frequency of the illuminating radar. Let \( s_k \) denote the BSCS at the \( k \)th range bin \( s_k \equiv \sigma_a^2(x_k) \). The unknown BSCS is modeled here as a first order Markov random sequence with conditional transition pdf’s \( f_{s_k|s_{k-1}}(s_k|s_{k-1}) \) and prior \( f_{s_0}(s_0) \). The vector of BSCS’s, \( s \), is independent of the unknown deterministic parameter vector, \( g \). The pdf of \( s_k|s_{k-1} \) and \( s_0 \) are assumed to be uniformly distributed:

\[
s_k|s_{k-1} \sim {\mathcal U}[s_{k-1} - d, s_{k-1} + d], \quad s_0 \sim {\mathcal U}(-D, D),
\]

and \( s_0 \sim {\mathcal U}(-D, D) \), where \( d \) and \( D \) are positive. The objective in this application is to estimate the refractivity profile parameter, \( g \), and the nuisance parameters, \( s = [s_0, \cdots, s_N]^T \), given \( N \) samples of the clutter intensity, \( y \equiv [y_1, \cdots, y_N]^T \).

The above assumptions satisfy the conditions of the problem stated in section 2. The transition density functions

\[
\{f_{s_k|s_{k-1}}(s_k|s_{k-1})\}_{k=1}^N
\]

and the pdf \( f_{s_0}(s_0) \) are assumed to be known. According to the model stated above, the measurements, \( (y_1, \cdots, y_N) \), given the unknown parameters are independent, zero-mean Gaussian random variables with variance:

\[
\sigma^2 = L^2(x_k; g)s_k + \sigma_\eta^2, \quad k = 1, \cdots, N.
\]

For the HCRLB from (2), the FIM’s \( J_1 \) and \( J_2 \) are required. It can be shown that for a uniformly distributed random variable, \( z \), the FIM corresponding to the priors is zero. Therefore the diagonal of the matrix \( J_2 \) is zero, and further \( J_2 = 0 \) for this problem. This does not necessarily mean that the prior information is non-informative however. In particular, although \( J_2 = 0 \), the prior information reduces the bound by performing the integration of \( J_1(s) \) over the random parameters, \( s \), in (2). The matrix \( J_1 \) can be calculated using Eqs. (5)-(7). In [8] the blocks of the matrix \( J_1 \) are calculated. The matrix \( J_{ss} \) is found to be diagonal with:

\[
J_{ss}[i] = \frac{|L(x_i; g)|^4}{(s_i^2|L(x_i; g)|^2 + \sigma_\eta^2)^2}, \quad i = 1, \cdots, N,
\]

the rows of \( J_{ss} \) are given by:

\[
J_{ss}[i] = \frac{s_i^2|L(x_i; g)|^2|L(x_i; g)|^2}{(s_i^2|L(x_i; g)|^2 + \sigma_\eta^2)^2}, \quad i = 1, \cdots, N,
\]

and the matrix \( J_{gg} \) is:

\[
J_{gg} = \sum_{i=1}^N \frac{s_i^2|L(x_i; g)|^2|L(x_i; g)|^2}{(s_i^2|L(x_i; g)|^2 + \sigma_\eta^2)^2}.
\]

where \( |L(x_i; g)|^2 \equiv \theta^2(x_i; g) \). The first row and column of \( J_1 \), corresponding to \( s_0 \), are zero, and since also \( J_2 = 0 \), the hybrid FIM \( E_1(J_1) + J_2 \) is not invertible. This implies that \( s_0 \) cannot be jointly estimated with the other parameters \( (s_1, \cdots, s_N) \) and \( g \). Therefore the corresponding row and column of \( J \) are excluded from the FIM. The corresponding estimator estimates \( (s_1, \cdots, s_N) \) using both the prior statistics and the data, while the estimate of \( s_0 \) is achieved according to its prior statistics only.

As an illustration of the HCRLB, a tri-linear profile was used for a numerical experiment and is depicted in Fig. 1, where its base height is assumed to be the unknown non-random. The two-way propagation loss, assuming a base height of \( g_0 = 30 \) m, is illustrated in Fig. 2. The propagation loss is calculated using the Radio Optical Physics (RPO) code [2]. The BSCS is assumed to be a Markov sequence where 20 log \( s_1 \) is uniformly distributed on -10 and +10 dB, and 20 log \( \frac{s_{(k+1)}}{s_k} \) is uniformly distributed between -3 dB and 3 dB for \( k = 1, \cdots, N \). The BSCS is assumed to be constant over 2 km intervals and the clutter intensity, \( |y|^2 \), is available every 50 m. The clutter-to-noise ratio (CNR) is defined as:

\[
CNR = \min_{s_0} \frac{|L(x_0; g)|^2}{\sigma_\eta^2}
\]

where the minimum is taken over the range extent of the dwell.

The HCRLB and maximum a-posteriori (MAP) estimator for RFC have been computed for the scenario described above. The performance of the MAP estimator is evaluated as a function of CNR using Monte-Carlo simulations and the HCRLB. Fig. 3 illustrates the estimator performance, based on 100 independent realizations for the case where \( s \) is an unknown Markov sequence as described above versus the HCRLB. The HCRLB is calculated using the recursive method presented in this paper. This figure shows that the estimator results are in agreement with the bound. Note that above a threshold CNR, the refractivity profile can be estimated with a root mean-square error of less than 0.5 m.

5. CONCLUSIONS

In this paper a novel method for calculating the hybrid CRLB for non-linear/non-Gaussian Markov models has been presented. The
method replaces the required expectation over the unknown random parameters, which involves a multi-dimensional integration, with recursively computed one-dimensional integral. The method is applied to the problem of refractivity estimation using radar clutter from the sea surface, where the BSCS is assumed to be a Markov process in range. The HCRLB results show the bound is tight with respect to the corresponding MAP estimator.

6. REFERENCES


