HAMMERSLEY-CHAPMAN-ROBBINS LOWER BOUNDS ON POLE AND RESIDUE ESTIMATES FROM IMPULSE RESPONSE DATA

Abdullah Al Maruf and Sandip Roy

School of Electrical Engineering and Computer Science at Washington State University

ABSTRACT

The estimation of nonrandom pole and residue parameters from impulse-response data is studied. Specifically, the Hammersley-Chapman-Robbins lower bound (HCRB) on the estimation error variance is analyzed for single-input single-output systems with multiple but distinct poles. The HCRB is compared with the widely used Cramer-Rao lower bound (CRB) in examples. The HCRB is found to be significantly tighter than the CRB when noise levels are high compared to the impulse response signal, while the bounds become close for small noise levels (equivalently, large residues).

Index Terms— Estimation, Identification, Stochastic systems, Hammersley-Chapman-Robbins bound.

1. INTRODUCTION

There is a very wide literature on estimation of linear-system model parameters (e.g., poles, residues, zeros) from noisy impulse response data [1], [2], [3], [4], [5], [6], [7]. One focus has been on developing Cramer-Rao lower bounds on the error variances of parameter estimates, for the case that the model parameters are assumed to be nonrandom [1], [2], [3]. These bounds recently have found application in infrastructure-monitoring contexts (e.g., monitoring of power-grid swings and flexible-structure dynamics), as new sensing technologies have made possible the inference of modal dynamics from test responses or ambient data [8], [9], [10]. The bounds are useful for developing confidence intervals on modal estimates, and also for sensor placement under cost constraints.

The problem of estimating linear-system parameters from impulse response data typically does not satisfy the regularity conditions which guarantee that the Cramer-Rao bound is tight, even in the limit of a long data horizon. In a preliminary study, the Cramer-Rao bound has been compared with the Hammersley-Chapman-Robbins lower bound for a single pole system [11]. The gap between the bounds shows that the Cramer-Rao bound may be far from tight even in this simple case. This is concerning for infrastructure-monitoring applications, since it may lead to overly-optimistic confidence intervals and suboptimal sensor placement.

The purpose of this work is to pursue computation of the Hammersley-Chapman-Robbins bound (HCRB) on the error variances of pole and residue estimates [12], [13], for multi-dimensional linear-system models. Specifically, computation of the HCRB for single-input single-output linear systems with multiple non-repeated poles is studied. The main contribution of the work is to phrase the HCRB as the minimization of a cost function that has an explicit form. This explicit formulation permits application of numerical optimization techniques and also serves as a starting point for developing formal characterizations and bounds, as briefly discussed here.

2. PRIOR WORK AND BRIEF REVIEW

This study contributes to a literature on the Hammersley-Chapman-Robbins bound (HCRB), which is relatively sparse due to the analytical and computational difficulty inherent to applying the bound. The theory of the HCRB, among other Barankin-type bounds, has been developed over a number of years [12], [13]. Regarding applications, the HCRB has recently been used for estimating sparse non-random vectors in the presence of Gaussian white noise [15], [16]. The HCRB has also been used in estimating multiple change points in time series, since the widely-used CRB is not applicable when the change-point location parameters are discrete [17]. Similarly, the HCRB has been applied to threshold prediction in direction-of-arrival (DOA) estimation and source localization, since CRB is not satisfactorily tight in the case of low SNR and limited observation points [18], [19]. In our previous work [11], the HCRB was derived for the pole and residue estimation problem, for a single-pole system; this study extends the analysis to multi-dimensional systems.

The multi-parameter HCRB is a Barankin-type lower bound on the error covariance for a nonrandom parameter vector's estimate. A formulaic description of the HCRB, which is used to develop the results for pole and residue estimation here, is briefly reviewed [20]. Formally, consider an unbiased estimator $\mathbf{T}(\mathbf{x})$ for an unconstrained nonrandom parameter vector $\boldsymbol{\theta} = [\theta_1\ \theta_2 \cdots \theta_k]^T \in \mathbb{R}^k$ based on a set of observations $\mathbf{x} = \left[\mathbf{x_1}, \mathbf{x_2}, \cdots, \mathbf{x_n}\right]$. The observations are modeled as random variables generated according

This work is supported by the United States National Science Foundation under grants CNS-1545104 and CMMI-1635184, and of Lawrence Berkeley National Laboratory.

to the joint probability density function $f(\mathbf{x_1} = x_1, \mathbf{x_2} = x_2, \cdots, \mathbf{x_n} = x_n; \boldsymbol{\theta})$ or succinctly $f(\mathbf{x}; \boldsymbol{\theta})$. The HCRB $HCRB_{\boldsymbol{\theta}}$ provides a lower bound on the covariance matrix $COV\{\mathbf{T}(\mathbf{x})\} = E\{[\mathbf{T}(\mathbf{x}) - \boldsymbol{\theta}][\mathbf{T}(\mathbf{x}) - \boldsymbol{\theta}]^T\}$, in the sense that $HCRB_{\boldsymbol{\theta}} - COV\{\mathbf{T}(\mathbf{x})\}$ is negative definite. The HCRB is given by

$$HCRB_{\boldsymbol{\theta}} = \sup\left(\mathbf{G}_{HCRB}{}^{\dagger}\right) \tag{1}$$

where

$$\mathbf{G}_{HCRB} = [\boldsymbol{V}]\, E_{\boldsymbol{\theta}}\left\{[\frac{\delta f_{\boldsymbol{\theta}}}{f_{\boldsymbol{\theta}}}]^T\, [\frac{\delta f_{\boldsymbol{\theta}}}{f_{\boldsymbol{\theta}}}]\right\} [\boldsymbol{V}]^T. \tag{2}$$

Here $[\boldsymbol{V}] = [\mathbf{v}_1 \;\; \mathbf{v}_2 \;\; \cdots \;\; \mathbf{v}_k]$ where $\mathbf{v}_1, \mathbf{v}_2, \cdots, \mathbf{v}_k \in \mathbb{R}^k$ are mutually independent direction vectors, $\delta f_{\boldsymbol{\theta}} = [\delta_1 f_{\boldsymbol{\theta}} \;\; \delta_2 f_{\boldsymbol{\theta}} \;\; \cdots \;\; \delta_k f_{\boldsymbol{\theta}}]$ where each $\delta_i f_{\boldsymbol{\theta}}$ is a finite difference of the density function due to a change in the parameter vector $\boldsymbol{\theta}$ in the direction $\mathbf{v}_i$: $\delta_i f_{\boldsymbol{\theta}} = \frac{f_{\boldsymbol{\theta}+h_i\mathbf{v}_i} - f_{\boldsymbol{\theta}}}{h_i}$, $\quad i = 1, 2, \cdots, k$, and $h_1, h_2, \cdots, h_k$ are scalars. Note that, in (1), $\{\cdot\}^{\dagger}$ denotes the Moore-Penrose pseudo inverse, and the supremum is taken over all possible direction vectors and respective magnitude scalars. Also, the matrix $\left(\mathbf{G}_{HCRB}{}^{\dagger}\right)$ is supremized in the sense that any particular scalar quadratic form $\mathbf{z}^T\left(\mathbf{G}_{HCRB}{}^{\dagger}\right)\mathbf{z}$ is supremized. We notice that each chosen scalar quadratic form will give a different bound; it is natural to choose the quadratic form to develop an optimally-tight bound on a particular scalar statistic of interest.

The HCRB does not require any of the regularity assumptions of the CRB, but does require the weak condition that the support of $f(\mathbf{x}; \boldsymbol{\theta} + h_i\mathbf{v}_i)$ is subset of the support of $f(\mathbf{x}; \boldsymbol{\theta})$ for $i = 1, 2, \cdots, k$. Furthermore, the HCRB is at least as tight as the CRB, with the two bounds coinciding when the supremum is achieved at $h_i \to 0$ for all i with the direction vectors taken as the unit vectors in $\mathbb{R}^k$.

3. PROBLEM FORMULATION

Parameter estimation for a discrete-time system with distinct poles is studied. Specifically, a system with the transfer function $H(z) = \sum_{l=1}^{r} \frac{A_l}{(1-a_l z^{-1})}$ is considered, where the nonrandom parameters a_l and A_l are the poles and residues of the system, respectively. Noisy measurements are made of the system's response at the times $k = 0, \ldots, n$, upon impulsive stimulation at time $k = 0$. Specifically, the measured impulse response is given by

$$y(k) = \sum_{l=1}^{r} A_l a_l^k + w(k) \tag{3}$$

for $k = 0, \ldots, n$. Here $h(k) = \sum_{l=0}^{r} A_l a_l^k$ is the true impulse response of the system, and $w(k)$ is a zero-mean Gaussian white noise with variance σ^2.

Our aim here is to derive the HCRB on the estimation error covariance for the nonrandom parameter vector $\boldsymbol{\theta} = [a_1\; a_2\; \cdots a_r\; A_1\; A_2\; \cdots\; A_r]^T$.

4. ANALYSIS OF THE HCRB

The HCRB for the pole and residue estimation problem is developed using the formulaic expression reviewed in Section 2. Here is the main result:

Theorem 1 *The HCRB for the pole and residue estimation problem is*

$$HCRB_{\boldsymbol{\theta}} = \sup\left(\mathbf{G}_{HCRB}{}^{\dagger}\right) = \sup\left(\left[\begin{array}{cc} \mathbf{G}^{aa} & \mathbf{G}^{aA} \\ \mathbf{G}^{Aa} & \mathbf{G}^{AA} \end{array}\right]^{\dagger}\right)$$

Here, $\mathbf{G}^{aa}, \mathbf{G}^{aA}, \mathbf{G}^{Aa}$ and $\mathbf{G}^{AA}$ are $r \times r$ matrices whose (i, j) th entries are given by

$$\begin{aligned} G_{ij}^{aa} &= \frac{1}{h_{a_i} h_{a_j}}[\,\exp(\,\frac{A_i A_j}{\sigma^2} S_{a_i a_j}) - 1] \\ G_{ij}^{AA} &= \frac{1}{h_{A_i} h_{A_j}}[\,\exp(\,\frac{h_{A_i} h_{A_j}}{\sigma^2} S_{A_i A_j}) - 1] \\ G_{ij}^{aA} &= G_{ji}^{Aa} = \frac{1}{h_{A_i} h_{A_j}}[\,\exp(\,\frac{A_i h_{A_j}}{\sigma^2} S_{a_i A_j}) - 1] \end{aligned}$$

where

$$\begin{aligned} S_{a_i a_j} &= \sum_{k=0}^{n} \big[(a_i + h_{a_i})^k (a_j + h_{a_j})^k + (a_i a_j)^k \\ &\qquad - a_i^k (a_j + h_{a_j})^k - (a_i + h_{a_i})^k a_j^k\big] \\ S_{A_i A_j} &= \sum_{k=0}^{n} (a_i a_j)^k \\ S_{a_i A_j} &= \sum_{k=0}^{n} \left[(a_i + h_{a_i})^k a_j^k - (a_i a_j)^k\right]. \end{aligned}$$

Here, the supremum is found with respect to $h_{a_1}, \ldots, h_{a_r}$ and $h_{A_1}, \ldots, h_{A_r}$, and any scalar quadratic form $\mathbf{z}^T\left(\mathbf{G}_{HCRB}{}^{\dagger}\right)\mathbf{z}$ may be supremized to obtain a bound.

Proof: The HCRB is developed by substitution into Equations 1 and 2, from [20]. To simplify, first notice that the bound does not depend on the choice of the direction vectors, since the parameter vector is unconstrained. Hence without loss of generality, we choose $[\boldsymbol{V}] = I_{2r}$. Substituting, we thus obtain that $HCRB_{\boldsymbol{\theta}} = \sup(\mathbf{G}_{HCRB}{}^{\dagger})$, where the Barankin information matrix $\mathbf{G}_{HCRB}$ has the form: $\mathbf{G}_{HCRB} = \left[\begin{array}{cc} \mathbf{G}^{aa} & \mathbf{G}^{aA} \\ \mathbf{G}^{Aa} & \mathbf{G}^{AA} \end{array}\right]$, and $\mathbf{G}^{aa}, \mathbf{G}^{aA}, \mathbf{G}^{Aa}$ and $\mathbf{G}^{AA}$ are $r \times r$ matrices whose (i, j)th entries are: $G_{ij}^{aa} = \frac{1}{h_{a_i} h_{a_j}}[E_{\boldsymbol{\theta}}\{\frac{f(\mathbf{y};\, \boldsymbol{\theta}+h_{a_i}\mathbf{v}_i)}{f(\mathbf{y};\boldsymbol{\theta})} \frac{f(\mathbf{y};\, \boldsymbol{\theta}+h_{a_j}\mathbf{v}_j)}{f(\mathbf{y};\boldsymbol{\theta})}\} - 1]$, $G_{ij}^{aA} = G_{ji}^{Aa} = \frac{1}{h_{a_i} h_{A_j}}[E_{\boldsymbol{\theta}}\{\frac{f(\mathbf{y};\, \boldsymbol{\theta}+h_{a_i}\mathbf{v}_i)}{f(\mathbf{y};\boldsymbol{\theta})} \frac{f(\mathbf{y};\, \boldsymbol{\theta}+h_{A_j}\mathbf{v}_{r+j})}{f(\mathbf{y};\boldsymbol{\theta})}\} - 1]$, and $G_{ij}^{AA} = \frac{1}{h_{A_i} h_{A_j}}[E_{\boldsymbol{\theta}}\{\frac{f(\mathbf{y};\, \boldsymbol{\theta}+h_{A_i}\mathbf{v}_{r+i})}{f(\mathbf{y};\boldsymbol{\theta})} \frac{f(\mathbf{y};\, \boldsymbol{\theta}+h_{A_j}\mathbf{v}_{r+j})}{f(\mathbf{y};\boldsymbol{\theta})}\} - 1]$.

Since the noise is zero mean, white, and Gaussian with variance σ^2, the joint probability density function (pdf) for $y(k)$, $k = 0, 1, 2, \cdots, n$, can be written as $f(\mathbf{y};\boldsymbol{\theta}) = (2\pi\sigma^2)^{-\frac{n+1}{2}} \exp\left(\frac{-1}{2\sigma^2}\sum_{k=0}^{n}[y(k) - \sum_{l=0}^{r} A_l a_l^k]^2\right)$. Substitution of the joint pdf yields $\frac{f(\mathbf{y};\,\boldsymbol{\theta}+h_{a_i}\mathbf{v}_i)}{f(\mathbf{y};\boldsymbol{\theta})} = \exp\Bigg(\frac{1}{2\sigma^2}\sum_{k=0}^{n}\Big[2A_i y(k)[(a_i+h_{a_i})^k - a_i^k] + A_i^2[a_i^{2k} - (a_i+h_{a_i})^{2k}] + 2A_i[a_i^k - (a_i+h_{a_i})^k](\sum_{\substack{l=0\\l\neq i}}^{r} A_l a_l^k)\Big]\Bigg)$. Hence, G_{ij}^{aa} can be written as: $G_{ij}^{aa} = \frac{1}{h_{a_i}h_{a_j}}[E_{\boldsymbol{\theta}}\{\exp\left(\sum_{k=0}^{n}\left({}_{a_ia_j}C_k\; y(k) + {}_{a_ia_j}\; D_k\right)\right)\} - 1]$ where ${}_{a_ia_j}C_k = \frac{1}{\sigma^2}\left(A_i[(a_i+h_{a_i})^k - a_i^k] + A_j[(a_j+h_{a_j})^k - a_j^k]\right)$ and ${}_{a_ia_j}D_k = \frac{1}{2\sigma^2}\Big(A_i^2[a_i^{2k} - (a_i+h_{a_i})^{2k}] + 2A_i[a_i^k - (a_i + h_{a_i})^k](\sum_{\substack{l=0\\l\neq i}}^{r} A_l a_l^k)] + A_j^2[a_j^{2k} - (a_j+h_{a_j})^{2k}] + 2A_j[a_j^k - (a_j + h_{a_j})^k](\sum_{\substack{l=0\\l\neq j}}^{r} A_l a_l^k)]\Big)$.

Since each observation is Gaussian, the exponent $u_{a_ia_j} = \sum_{k=0}^{n}({}_{a_ia_j}C_k\; y(k) + {}_{a_ia_j}D_k)$ in the expression for G_{ij}^{aa} is also Gaussian. The mean of $u_{a_ia_j}$, denoted as $\mu_{u_{a_ia_j}}$ can be evaluated as $\mu_{u_{a_ia_j}} = \sum_{k=0}^{n}({}_{a_ia_j}C_k\; \sum_{l=0}^{r} A_l a_l^k + {}_{a_ia_j}D_k)$. Similarly, the variance of $u_{a_ia_j}$, denoted as $\sigma^2_{u_{a_ia_j}}$ can be evaluated as $\sigma^2_{u_{a_ia_j}} = \frac{1}{\sigma^2}\sum_{k=0}^{n}{}_{a_ia_j}C_k^2$. Thus, using the moment generating function for the Gaussian distribution, we can simplify $E_{\boldsymbol{\theta}}\{\exp(u_{a_ia_j})\}$ as $E_{\boldsymbol{\theta}}\{\exp(u_{a_ia_j})\} = \exp(\mu_{u_{a_ia_j}} + \frac{\sigma^2_{u_{a_ia_j}}}{2}) = \exp(\frac{A_iA_j}{\sigma^2}S_{a_ia_j})$, where $S_{a_ia_j} = \sum_{k=0}^{n}\left[(a_i+h_{a_i})^k(a_j+h_{a_j})^k + (a_ia_j)^k - a_i^k(a_j+h_{a_j})^k - a_j^k(a_i+h_{a_i})^k\right]$. Substituting, we recover the expression for G_{ij}^{aa} in the theorem statement.

Similarly, to find G_{ij}^{AA}, we can substitute the joint pdf of the observations to simplify $\frac{f(\mathbf{y};\,\boldsymbol{\theta}+h_{A_i}\mathbf{v}_{r+i})}{f(\mathbf{y};\boldsymbol{\theta})}$. Doing so, we get $\frac{f(\mathbf{y};\,\boldsymbol{\theta}+h_{A_i}\mathbf{v}_{r+i})}{f(\mathbf{y};\boldsymbol{\theta})} = \exp\Bigg(\frac{1}{2\sigma^2}\sum_{k=0}^{n}\Big[2h_{A_i}y(k)a_i^k - 2h_{A_i}a_i^k(\sum_{\substack{l=0\\l\neq i}}^{r} A_l a_l^k) - a_i^{2k}(2A_ih_{A_i} + h_{A_i}^2)\Big]\Bigg)$. Thus, G_{ij}^{AA} can be written as $G_{ij}^{AA} = \frac{1}{h_{A_i}h_{A_j}}[E_{\boldsymbol{\theta}}\{\exp\left(\sum_{k=0}^{n}\left({}_{A_iA_j}C_k\; y(k) + {}_{A_iA_j}D_k\right)\right)\} - 1]$ where ${}_{A_iA_j}C_k = \frac{1}{\sigma^2}\left(h_{A_i}a_i^k + h_{A_j}a_j^k\right)$ and ${}_{A_iA_j}D_k = \frac{-1}{2\sigma^2}\Big(2h_{A_i}a_i^k(\sum_{\substack{l=0\\l\neq i}}^{r} A_l a_l^k) + a_i^{2k}(2A_ih_{A_i} + h_{A_i}^2) + 2h_{A_j}a_j^k(\sum_{\substack{l=0\\l\neq j}}^{r} A_l a_l^k) + a_j^{2k}(2A_jh_{A_j} + h_{A_j}^2)\Big)$.

Since each observation is Gaussian, the exponent $u_{A_iA_j} = \sum_{k=0}^{n}({}_{A_iA_j}C_k\; y(k) + {}_{A_iA_j}D_k)$ in the expression for G_{ij}^{AA} is also Gaussian. The mean of $u_{A_iA_j}$, denoted as $\mu_{u_{A_iA_j}}$ can be evaluated as $\mu_{u_{A_iA_j}} = \sum_{k=0}^{n}({}_{A_iA_j}C_k\; \sum_{l=0}^{r} A_l a_l^k + {}_{A_iA_j}D_k)$. Similarly, the variance of $u_{A_iA_j}$, denoted as $\sigma^2_{u_{A_iA_j}}$ can be evaluated as $\sigma^2_{u_{A_iA_j}} = \frac{1}{\sigma^2}\sum_{k=0}^{n}{}_{A_iA_j}C_k^2$. Using the moment-generating function for the Gaussian distribution, $E_{\boldsymbol{\theta}}\{\exp(u_{A_iA_j})\}$ can be simplified as $E_{\boldsymbol{\theta}}\{\exp(u_{A_iA_j})\} = \exp(\mu_{u_{A_iA_j}} + \frac{\sigma^2_{u_{A_iA_j}}}{2}) = \exp(\frac{h_{A_i}h_{A_j}}{\sigma^2}S_{A_iA_j})$, where $S_{A_iA_j} = \sum_{k=0}^{n}(a_ia_j)^k$. Substituting, we immediately recover the expression for G_{ij}^{AA} in the theorem statement.

Similarly to find G_{ij}^{aA} we use previously obtained simplified expressions of $\frac{f(\mathbf{y};\,\boldsymbol{\theta}+h_{A_i}\mathbf{v}_{r+i})}{f(\mathbf{y};\boldsymbol{\theta})}$ and $\frac{f(\mathbf{y};\,\boldsymbol{\theta}+h_{a_i}\mathbf{v}_i)}{f(\mathbf{y};\boldsymbol{\theta})}$. Doing so, we get $G_{ij}^{aA} = \frac{1}{h_{a_i}h_{A_j}}[E_{\boldsymbol{\theta}}\{\exp\left(\sum_{k=0}^{n}\left({}_{a_iA_j}C_k\; y(k) + {}_{a_iA_j}D_k\right)\right)\} - 1]$ where ${}_{a_iA_j}C_k = \frac{1}{\sigma^2}\left(A_i[(a_i+h_{a_i})^k - a_i^k] + h_{A_j}a_j^k\right)$ and, ${}_{a_iA_j}D_k = \frac{1}{2\sigma^2}\Big(A_i^2[a_i^{2k} - (a_i+h_{a_i})^{2k}] + 2A_i[a_i^k - (a_i+h_{a_i})^k](\sum_{\substack{l=0\\l\neq i}}^{r} A_l a_l^k)] - 2h_{A_j}a_j^k\;(\sum_{\substack{l=0\\l\neq j}}^{r} A_l a_l^k) + a_j^{2k}(2A_jh_{A_j} + h_{A_j}^2)\Big)$

Since each observation is Gaussian, the exponent $u_{a_iA_j} = \sum_{k=0}^{n}({}_{a_iA_j}C_k\; y(k) + {}_{a_iA_j}D_k)$ in the expression for G_{ij}^{aA} is also Gaussian. The mean of $u_{a_iA_j}$, denoted as $\mu_{u_{a_iA_j}}$ can be evaluated as $\mu_{u_{a_iA_j}} = \sum_{k=0}^{n}({}_{a_iA_j}C_k\; \sum_{l=0}^{r} A_l a_l^k + {}_{a_iA_j}D_k)$. Similarly, the variance of $u_{a_iA_j}$, denoted as $\sigma^2_{u_{a_iA_j}}$ can be evaluated as $\sigma^2_{u_{a_iA_j}} = \frac{1}{\sigma^2}\sum_{k=0}^{n}{}_{a_iA_j}C_k^2$. Thus, using the moment generating function of the Gaussian distribution, we can simplify $E_{\boldsymbol{\theta}}\{\exp(u_{a_iA_j})\}$ as $E_{\boldsymbol{\theta}}\{\exp(u_{a_iA_j})\} = \exp(\mu_{u_{a_iA_j}} + \frac{\sigma^2_{u_{a_iA_j}}}{2}) = \exp(\frac{A_ih_{A_j}}{\sigma^2}S_{a_iA_j})$ where $S_{a_iA_j} = \sum_{k=0}^{n}\left[(a_i+h_{a_i})^k a_j^k - (a_ia_j)^k\right]$. Substituting, we immediately recover the expressions for G_{ij}^{aA} and G_{ji}^{aA} in the theorem statement. □

We remark that the terms $S_{a_ia_j}$, $S_{A_iA_j}$ and $S_{a_iA_j}$ can readily be written in closed form using geometric-sum formulas. For instance, $S_{a_ia_j} = \frac{1-p_{ij}^{n+1}}{1-p_{ij}} + \frac{1-q_{ij}^{n+1}}{1-q_{ij}} - \frac{1-m_{ij}^{n+1}}{1-m_{ij}} - \frac{1-m_{ji}^{n+1}}{1-m_{ji}}$, and where $p_{ij} = (a_i+h_{a_i})(a_j+h_{a_j})$, $q_{ij} = a_ia_j$, and $m_{ij} = a_i(a_j+h_{a_j})$. Further details are omitted. We also stress that different bounds can be obtained through selection of the vector $\mathbf{z}$, which specifies the scalar metric to be optimized. Typically, there is interest in tightly bounding estimator error variances for individual mode and residue estimates. To develop tight such bounds, the optimization problem given in Theorem 1 can be solved upon selecting $\mathbf{z}$ to be each possible unit basis vector (indicator vector).

The expression for the HCRB in Theorem 1 is useful as a starting point for numerical computation. Specifically, the expression gives an explicit closed-form expression for the Barankin information matrix, and hence the HCRB can be computed by optimizing a quadratic form of the inverse of a matrix which has closed-form expressions for the entries. Noting that derivatives of matrix inverses have been extensively studied, standard unconstrained optimization algorithms can potentially be applied with low computational cost. While the cost function is not convex, it is smooth and optima lie within a closed set; in practice, standard algorithms have been effective in finding the optimum (see following Example Section).

The HCRB expression in Theorem 1 is also appealing because it provides a starting point for formal analysis of the bound. For instance, the expression allows comparison of the CRB with the HCRB (see [1] , [2] for expressions for the CRBs on pole and residue estimates). It is found the

CRB is close to the HCRB if the noise level is sufficiently small (equivalently, the residues are sufficiently large), i.e. the signal-to-noise ratio is high. Meanwhile, the ratio between the HCRB and CRB increases but reaches an asymptote as the signal-to-noise ratio is decreased. The expression also gives insight into how the bound scales with parameters of the model, and allows development of simpler analytical bounds. These formal results will be developed in detail in future work.

5. NUMERICAL COMPUTATIONS

Numerical computations of the CRB and HCRB are undertaken for the two-pole case ($r = 2$) to gain further insight into the gaps between the bounds and their dependencies on the pole and residue locations. For the Figures 1 and 2, we fix the second pole a_2 at 0.9 and vary the first pole a_1 from 0.1 to 0.8 for two sets of residue values. For the Figure 3 we fix the poles a_1 and a_2 at 0.2 and 0.9 respectively and vary the residues A_1 and A_2 as $0.1k$ and $0.3k$ respectively for $k = 1, 1.5, 2, \cdots, 5$. A long time horizon ($N = 10000$) is assumed, and the noise level is assumed to be $\sigma = 0.5$ for all examples.

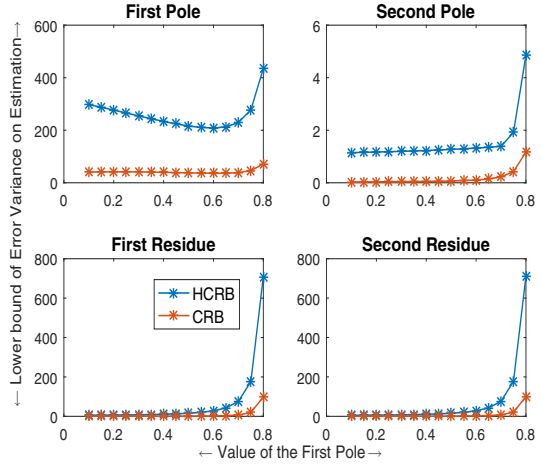

Fig. 1. Lower bounds on the pole and residue estimation error covariance as a function of the location of pole 1 (a_1), with $a_2 = 0.9$, $A_1 = 0.1$ and $A_2 = 0.3$.

Figures 1 and 2 indicate that the HCRB is larger than the CRB, as expected. Comparing Figures 1 and 2 and also from Figure 3, we see that the gap between the HCRB and CRBs is significant for smaller values of the residues (small signal components compared to noise levels). This matches with related works on the HCRB and CRB in other application areas, which have shown that the HCRB significantly improves on the CRB in low signal-to-noise ratio (SNR) settings. As the first pole gets closer to the second pole, both the HCRBs and CRBs increase in Figures 1 and 2, as expected.

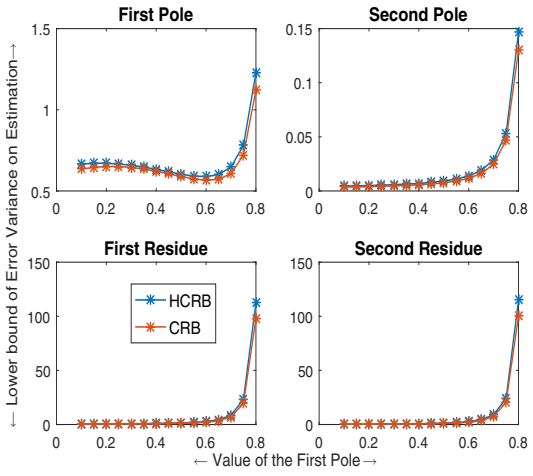

Fig. 2. Lower bounds on the pole and residue estimation error covariance as a function of the location of pole 1 (a_1), with $a_2 = 0.9$, $A_1 = 0.8$ and $A_2 = 0.9$.

Figure 3 further illustrates how the gap between the HCRB and CRB depends on the signal-to-noise ratio. As the residues are up-scaled (i.e., the signal-to-noise increases), the ratio of the HCRB to the CRB rapidly transitions from being large to being close to 1.

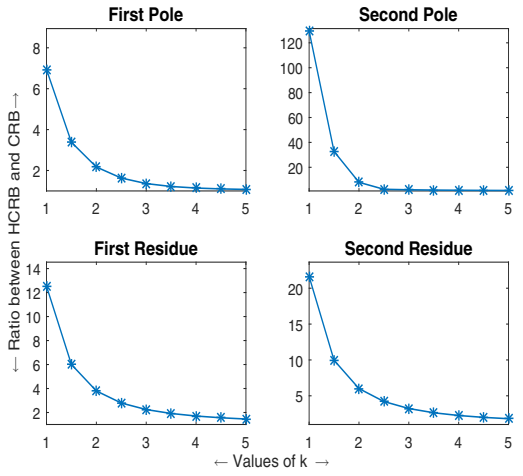

Fig. 3. Ratio between the HCRB and CRB as the residues are scaled, for the pole locations $a_1 = 0.2$ and $a_2 = 0.9$

6. REFERENCES

[1] L.T. McWhorter and L.L. Scharf, "Cramer-Rao bounds for deterministic modal analysis", *IEEE Trans. Signal Proc.*, SP-41, pp. 1847-1862, May 1993.

[2] Y. Wan and S. Roy, "On inference of network time constants from impulse response data: graph-theoretic Cramer-Rao bounds", *Proc. IEEE Conf. Decision and Control* pp. 4111-4116 Dec. 2009.

[3] M. P. Clark, "Cramer-Rao bounds for two-dimensional deterministic modal analysis", *Signals, Systems and Computers Conference Record of The Twenty-Seventh Asilomar*, 1993.

[4] T. K. Sarkar and O. Pereira, "Using the matrix pencil method to estimate the parameters of a sum of complex exponentials", *IEEE Antennas and Propagation Magazine*, vol. 37, no. 1, pp. 48-55, 1955.

[5] R. Kumaresan and D. Tufts, "Estimating the parameters of exponentially damped sinusoids and pole-zero modeling in noise", *IEEE Transactions on Acoustics, Speech, and Signal Processing*, vol. 30, no. 6, pp. 833-840, 1982.

[6] R. Kumaresan, L. Scharf, and A. Shaw, "An algorithm for pole-zero modeling and spectral analysis." *IEEE transactions on acoustics, speech, and signal processing* vol. 34, no. 3, pp. 637-640, 1986.

[7] R. S. Tsay and George C. Tiao, "Consistent estimates of autoregressive parameters and extended sample autocorrelation function for stationary and nonstationary ARMA models", *Journal of the American Statistical Association*, vol. 79, no. 385, pp. 84-96, 1984.

[8] Y. Wan, S. Roy, M. Xue and V. Katragadda, "Estimating modes of a complex dynamical network from impulse response data: Structural and graph-theoretic characterizations", *International Journal of Robust and Nonlinear Control* vol. 25, no. 10, pp. 1438-1453, 2015.

[9] A. Chakrabortty and C. F. Martin, "Optimal sensor placement for parametric model identification of electrical networks, part I: Open loop estimation." *IEEE Conference on Decision and Control*, pp. 5804-5809, 2010.

[10] P. H. Kirkegaard and R. Brincker, "On the optimal location of sensors for parametric identification of linear structural systems", Dept. of Building Technology and Structural Engineering, Aalborg University, 1992.

[11] A. A. Maruf and S. Roy, "Tighter lower bounds on the error variance of pole and residue estimates from impulse response data: an expository example", to appear in *Proc. IEEE Conf. on Decision and Control*, 2017.

[12] D. G. Chapman and H. Robbins, "Minimum variance estimation without regularity assumptions", *Ann. Math. Statist.*, vol. 22, no. 4, pp. 581586, Dec. 1951.

[13] J. M. Hammersley, "On estimating restricted parameters", J. Roy. Statist. Soc. B, vol. 12, no. 2, pp. 192-240, 1950.

[14] J. Abad Torres and S. Roy, "Cramer-Rao bounds on eigenvalue estimates from impulse response data: The multi-observation case", *Proc. IEEE Conference on Decision and Control*, December 10-13, 2012.

[15] A. Jung, Z. Ben-Haim, F. Hlawatsch and Y. C. Eldar, "Unbiased estimation of a sparse vector in white Gaussian noise", *IEEE Trans. Inf. Theory* vol. 57 no. 12 pp. 7856-7876 Dec. 2011.

[16] Y. Tang, L. Chen and Y. Gu "On the performance bound of sparse estimation with sensing matrix perturbation", *IEEE Trans. Signal Processing* vol. 61, no. 17 pp. 4372-4386, Sep. 2013.

[17] P. S. La Rosa, A. Renaux, C. H. Muravchik, and A. Nehorai, "Barankin-type lower bound on multiple change-point estimation", *IEEE Trans. on Signal Processing*, vol. 58, no. 11, Nov 2010.

[18] M. N. El Korso, A. Renaux, R. Boyer and S. Marcos, "Deterministic performance bounds on the mean square error for near field source localization", *IEEE Trans. on Signal Processing*, vol. 61, no. 4, pp. 871-877, Feb. 2013.

[19] T. Li J. Tabrikian and A. Nehorai, "A Barankin-type bound on direction estimation using acoustic sensor arrays", *IEEE Trans. on Signal Processing*, vol. 59, no. 1, pp. 431-435, Jan. 2011.

[20] J. D. Gorman and A. O. Hero, "Lower bounds for parametric estimation with constraints," *IEEE Trans. Information Theory*, vol. 36, no. 6, pp. 1285-1301, Nov. 1990.