

# Cramer-Rao Bounds on Eigenvalue Estimates from Impulse Response Data: The Multi-Observation Case

Jackeline Abad Torres and Sandip Roy

**Abstract**— We examine the effect of having multiple observations in the estimation of non-random modes of linear dynamical systems from noisy impulse response data. Specifically, for this estimation problem, we develop an explicit algebraic characterization of the Fisher information matrix and hence Cramer-Rao bound in terms of the eigenvalues and residues of the transfer function, and so develop some simple bounds on the minimum possible error variance for eigenvalue estimates in terms of the eigenvalues' locations. We focus especially on developing a relationship between the Cramer-Rao bound on pole estimates for the multi-observation case, and those when each single observation is used separately for estimation.

## I. INTRODUCTION

ESTIMATION of model parameters from measurement data has been extensively studied. The particular problem of mode (eigenvalue) estimation for dynamical systems, which is the focus of this article, was first studied in 1795 in the classical work of Prony [1]. Surprisingly, however, Cramer-Rao (CR) bounds on the optimal performance of non-random mode estimators were only developed relatively recently. Specifically motivated by ARMA (auto-regressive moving average) modeling applications, the article [2] developed Cramer-Rao bounds on pole and residue estimates (as well as pole and zero estimates) from impulse-response data, for single-input-single-output (SISO) discrete time linear systems. The subsequent work [3] continued analyzing the CR lower bound on optimal mode estimates and found the dependence of estimation error on the mode locations. As an extension, Cramer-Rao bounds (CRB) have also been developed for mode estimates from response data in two dimensional systems, i.e. ones in which modal responses are functions of two independent variables rather than a single temporal variable [4, 5].

However, because efforts to develop the CR bounds on mode estimates have largely been motivated by ARMA modeling and signal processing applications, there has been little focus on mode estimation in dynamical models with multiple observations. In fact, a couple of efforts have considered parameter-estimation performance in dynamical models with multiple observations, however these have been focused on estimating parameters of the noise rather than

modes of the system [6, 7]. To the best of our knowledge, an explicit Cramer-Rao bound has not been given for estimates of the modes of a linear dynamical system from multiple concurrent observations of the system state.

It is worthwhile to connect and distinguish our work to the extensive and insightful work on inference of system parameters from measurement data when stochastic stimulations are applied (e.g., [8, 9]) or through arbitrary temporal stimulation [2]. While these approaches achieve verifiable model identification, they depend on deliberate application of persistent stimulations: unfortunately, in the network applications of interest to us (see below for further discussion), we do not have the freedom to apply such persistent stimulation, and instead are limited to identifying model parameters from impulse-response data, which further is corrupted by noise. Because of the sparsity of such impulse response data, full model identification is often infeasible, but estimates for the modes can still sometimes be obtained (and are informative to our applications). Therefore, our focus is on maximum likelihood (ML) mode estimation based on the impulse response. Particularly, we are concerned about the impact of the system structure of these models and the benefit of having multiple observations in the estimation performance.

In this paper, we consider the problem of estimating the non-random modes of a linear dynamical system from noisy impulse response data, in the case that the system has two or more response variables (or observations or outputs). Specifically, for this estimation problem, we develop several characterizations of the Cramer-Rao bound on the mode estimate. Our particular interest in this mode estimation problem for SIMO (single-input multiple-output) systems stems from our ongoing effort to estimate network models and model statistics from noisy response data [10], including estimating virus-spread models from infection counts and estimating biological (e.g., neuronal) network models from cell-culture data. We note that the multi-observation case studied here also has direct application in such domains power system dynamic analysis, see e.g. [11]. More broadly, our development of Cramer-Rao bounds for mode estimates can facilitate sensor selection in dynamical systems, and hence inform the nascent focus on *dynamical network security* in the control engineering community.

The results presented in this article, on mode estimation from multi-observation impulse response data, are of three types. First, we provide algebraic characterizations of the Fisher information matrix and so the Cramer-Rao bound, focusing especially on the asymptotic case that a long

Manuscript received March 6<sup>th</sup>, 2011. This work was partially supported by National Science Foundation grants ECS-0725589, ECS-0901137, ECS-1035369, and ECS-1058124. Both authors are with the School of Electrical Engineering and Computer Science at Washington State University, Pullman, WA, USA. Correspondence can be sent to {jabadtors, sroy}@eeecs.wsu.edu.

sequence of data is observed. We also relate the Cramer-Rao bound on pole estimates for the multi-observation case to those when each single observation is used separately for estimation. Further, we present some simple bounds on pole inference error in terms of the pole location, by simplifying the Cramer-Rao bound.

The remainder article is organized as follows. In section II, we formulate the problem of estimating residues and poles for a single-input two-output system, and pose the task of finding the Cramer-Rao bound for this estimation problem. In section III, we present our main results on the Cramer-Rao lower bound for multiple observations. In section IV, we discuss a generalization of our results to multiple observation systems.

## II. PROBLEM FORMULATION

We are concerned with inference of the eigenvalues of an unknown linear discrete-time system from multiple noisy impulse response measurements. Specifically, our primary interest is in determining the Cramer-Rao lower bound on the estimation error covariance, and in relating this bound to the structure of the system and the multiple observation locations. We limit ourselves to a system whose eigenvalues are unknown but observable and consider specifically the case that two diverse observations of the systems state are made.

Formally, we consider a discrete-time model, whose state evolves according to the following difference equation in response to an impulse stimulus into one state variable:

$$\mathbf{x}[k+1] = L\mathbf{x}[k] + \mathbf{e}_h u[k], \quad (1)$$

where  $\mathbf{x}[k] \in \mathbb{R}^n$  is the full state of the system,  $\mathbf{e}_h$  is  $0-1$  indicator vector with  $h^{th}$  entry equal to 1,  $u[k]$  is a scalar deterministic impulse signal (stimulus) and the state matrix  $L$  is symmetric and stable (i.e. has eigenvalues strictly inside the unit circle). We consider identification of the eigenvalues of  $L$  from noisy observations of two state variables, say the  $i^{th}$  and  $j^{th}$  ones. In particular, we assume that the following observation signal is available:

$$\mathbf{y}[k] = C\mathbf{x}[k] + \boldsymbol{\omega}[k], \quad (2)$$

where  $\boldsymbol{\omega}[k]$  is a discrete-time stationary zero mean Gaussian white noise process with covariance matrix  $\Sigma = \sigma^2 I$  with  $\sigma^2 = 1$ , and the output matrix  $C = [\mathbf{e}_i \ \mathbf{e}_j]^T$  captures that the state variables  $i$  and  $j$  are observed. We are interested in estimating the eigenvalues of  $L$  from a long (infinite) data sequence of the observations

$$\mathbf{y}^T = \begin{bmatrix} y_1[0] & y_1[1] & y_1[2] & y_1[3] & \dots \\ y_2[0] & y_2[1] & y_2[2] & y_2[3] & \dots \end{bmatrix}.$$

We assume that the eigenvalues of the system are both observable and controllable, i.e. the pair  $(C, L)$  is observable and the pair  $(L, \mathbf{e}_h)$  is controllable.

Our primary goal is to characterize the minimum possible error in the poles (or eigenvalues) estimate as specified in the Cramer-Rao bound. To characterize the CR bound, we find it convenient to use a pole-residue representation of the transfer functions from the stimulus to the two observations.

This form simplifies relating the CR bound with the magnitudes of the system parameters, including the poles (eigenvalues) and the residues specified of each observation.

As per the system definition, all the poles are real with magnitude less than 1, and each pole is observable. We thus see that the transfer function can be written in the following form:

$$H(z) = \sum_{l=0}^{n-1} \left[ \frac{A_l}{z-p_l} \ \frac{B_l}{z-p_l} \right]^T \quad (3)$$

Where  $p_l, l = 0, 1, 2, \dots, n-1$ , are the poles of the system and  $A_l$  and  $B_l, l = 0, 1, 2, \dots, n-1$ , are the residues associated with each pole and observation.

In this formulation, we consider the estimation of residues and poles from the noisy impulse response. Specifically, we study the CR lower bound on these pole and residue estimates as a representation of the estimator performance. Specifically, we can establish several characterizations, including explicit algebraic descriptions of the bound, comparison of the two observation case to the single observation one, and simple bounds in the estimator performance.

## III. RESULTS

In this section, we provide several analyses of the Cramer-Rao bound on pole and residue estimates, for the multi-observation model developed in section II.

We begin with an explicit algebraic characterization of the Fisher information matrix and hence Cramer-Rao bound in terms of the eigenvalues and residues of the transfer function (section III.A). Then, we relate the Cramer-Rao bound in the two-observation case to the Cramer-Rao bounds that would be obtained if each observation was used on its own for estimation (section III.B). Finally, simpler bounds on pole estimates in terms of the pole locations are found.

### A. Algebraic expressions for the Cramer-Rao bound

Let us begin with an algebraic characterization of the Fisher information matrix and hence Cramer-Rao bound for the pole and residues estimates.

We define  $\boldsymbol{\theta} = [\mathbf{A} \ \mathbf{B} \ \mathbf{p}]$  as the parameter vector, where  $\mathbf{A} = [A_0 \ A_1 \ \dots \ A_{n-1}]$ ,  $\mathbf{B} = [B_0 \ B_1 \ \dots \ B_{n-1}]$  and  $\mathbf{p} = [p_0 \ p_1 \ \dots \ p_{n-1}]$ . We also use the notation  $x_i[k]$  and  $x_j[k]$  to represent the impulse responses of state variables  $i$  and  $j$ . Note that the observation sequences  $y_1$  and  $y_2$  are the states  $i$  and  $j$  corrupted by noise (2).

Noting that the observations together constitute a set of multivariate normal random variables [2], it is straightforward to derive the Fisher information matrix for the parameter vector. Following on the analysis given in [2], we immediately find that the Fisher information matrix can be calculated as:

$$F(\boldsymbol{\theta}) = \frac{\partial x_i(\boldsymbol{\theta})^T}{\partial \boldsymbol{\theta}} \frac{\partial x_i(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}} + \frac{\partial x_j(\boldsymbol{\theta})^T}{\partial \boldsymbol{\theta}} \frac{\partial x_j(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}}. \quad (4)$$

Expressing the derivatives in terms of the residue and pole parameters, we obtain

$$F(\theta) = \begin{bmatrix} \frac{\partial x_i^T}{\partial A} \\ 0 \\ \frac{\partial x_i^T}{\partial p} \end{bmatrix} \begin{bmatrix} \frac{\partial x_i}{\partial A} & 0 & \frac{\partial x_i}{\partial p} \end{bmatrix} + \begin{bmatrix} 0 \\ \frac{\partial x_j^T}{\partial B} \\ \frac{\partial x_j^T}{\partial p} \end{bmatrix} \begin{bmatrix} 0 & \frac{\partial x_j}{\partial B} & \frac{\partial x_j}{\partial p} \end{bmatrix}. \quad (5)$$

From the transfer function (3) we obtain  $x_i[k] = \sum_{l=0}^{n-1} A_l p_l^k$  and  $x_j[k] = \sum_{l=0}^{n-1} B_l p_l^k$ . This result leads to:

$$\frac{dx_i^T}{dA} = \frac{dx_j^T}{dB} = \begin{bmatrix} 1 & p_0 & p_0^2 & p_0^3 & \dots \\ 1 & p_1 & p_1^2 & p_1^3 & \dots \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 1 & p_{n-1} & p_{n-1}^2 & p_{n-1}^3 & \dots \end{bmatrix}, \quad (6)$$

and  $\frac{dx_i^T}{dp} = A_d P^T$ ,  $\frac{dx_j^T}{dp} = B_d P^T$ , where

$$A_d = \begin{bmatrix} A_0 & & & \\ & \ddots & & \\ & & A_{n-1} & \\ & & & \ddots \end{bmatrix}, B_d = \begin{bmatrix} B_0 & & & \\ & \ddots & & \\ & & B_{n-1} & \\ & & & \ddots \end{bmatrix}$$

$$\text{And } P^T = \begin{bmatrix} 0 & 1 & 2p_0 & 3p_0^2 & \dots \\ 0 & 1 & 2p_1 & 3p_1^2 & \dots \\ \vdots & \vdots & \vdots & \vdots & \dots \\ 0 & 1 & 2p_{n-1} & 3p_{n-1}^2 & \dots \end{bmatrix} \quad (7)$$

$F(\theta)$  can be written in the following block form:

$$F(\theta) = \begin{bmatrix} F_R & F_{RP} \\ F_{RP}^T & F_P \end{bmatrix} \quad (8)$$

Where  $F_R, F_{RP}, F_P$  represent the blocks of the information matrix related to the residues, residue-pole correlations, and poles, respectively. With some further algebra, the blocks in the Fisher information matrix can be written as:

$$F_R = \begin{bmatrix} F_{R_1} & 0 \\ 0 & F_{R_2} \end{bmatrix}, F_{RP} = \begin{bmatrix} F_{RP_1} \\ F_{RP_2} \end{bmatrix}, F_P = [F_{P_1} + F_{P_2}], \quad (9)$$

where we have the entry wise expressions:

$$\{F_{R_1}\}_{i,j} = \{F_{R_2}\}_{i,j} = \frac{1}{1-p_i p_j} \text{ for } i, j = 0, 1, \dots, n-1 \quad (10)$$

$$\{F_{RP_1}\}_{i,j} = \frac{A_j p_i}{(1-p_i p_j)^2} \text{ for } i \neq j \text{ and } i, j = 0, 1, \dots, n-1$$

$$\{F_{RP_1}\}_{i,i} = \frac{2A_j p_i}{(1-p_i p_j)^2} \text{ for } i = 0, 1, \dots, n-1 \quad (11)$$

$$\{F_{RP_2}\}_{i,j} = \frac{B_j p_i}{(1-p_i p_j)^2} \text{ for } i \neq j \text{ and } i, j = 0, 1, \dots, n-1$$

$$\{F_{RP_2}\}_{i,i} = \frac{2B_j p_i}{(1-p_i p_j)^2} \text{ for } i = 0, 1, \dots, n-1 \quad (12)$$

$$\{F_{P_1}\}_{i,j} = \frac{A_i A_j (1+p_i p_j)}{(1-p_i p_j)^3} \text{ for } i, j = 0, 1, \dots, n-1 \quad (13)$$

and

$$\{F_{P_2}\}_{i,j} = \frac{B_i B_j (1+p_i p_j)}{(1-p_i p_j)^3} \text{ for } i, j = 0, 1, \dots, n-1 \quad (14)$$

It is worth noting here that the matrices  $F_{R_i}, F_{RP_i}$ , and  $F_{P_i}$  represent the blocks Fisher information matrix, related to the residues, residue-pole correlations, and poles, assuming that only the observation  $y_i$ ,  $i = 1, 2$ , is available at the time of the estimation. In other words, let call  $F_1(\theta)$  the Fisher information matrix on the parameter estimation from the noisy impulse response  $y_1$ .  $F_1(\theta)$  can be expressed as

follows:

$$F_1(\theta) = \begin{bmatrix} F_{R_1} & F_{RP_1} \\ F_{RP_1}^T & F_{P_1} \end{bmatrix} \quad (15)$$

Once the Fisher information matrix has been computed, the Cramer-Rao bound can be obtained as its inverse. Specifically, the CR bound on the parameter estimate [2] is given by:

$C_\theta = (F(\theta))^{-1}$ , and the bound on the error variance of a particular parameter  $\theta_i$  in the parameter vector is the  $(i, i)$  entry of the matrix  $C_\theta$ .

### B. Relationship between the CRB and the system structure

We emphasize that the Cramer-Rao bound gives a lower bound on the variance of any unbiased estimator of the parameter [3].

Beyond obtaining algebraic expressions for the CRB, we would like to relate the CRB to the structure of the system under investigation. With this goal in mind, let us first examine the effect of the additional sequence of data in the inference of the poles.

**Theorem 1:** Let  $C_p$  be the minimum achievable error covariance (or CRB) for the poles estimates, for the pole-residue estimation problem introduced in section II. Let  $C_{p_1}$  and  $C_{p_2}$  be the Cramer-Rao bounds on the pole estimates, assuming that only the first and second observation, respectively, are available at each time for the estimation. Then  $C_p^{-1} = C_{p_1}^{-1} + C_{p_2}^{-1}$ .

**Proof:**

The Cramer-Rao bound can be expressed as follows, based on the block representation of  $F(\theta)$  and the formula for block-matrix inverses:

$$C_\theta = \begin{bmatrix} (F_R - F_{RP} F_P^{-1} F_{RP}^T)^{-1} & (F_R - F_{RP} F_P^{-1} F_{RP}^T)^{-1} F_{RP} F_P^{-1} \\ -(F_P - F_{RP}^T F_R^{-1} F_{RP})^{-1} F_{RP}^T F_R^{-1} & (F_P - F_{RP}^T F_R^{-1} F_{RP})^{-1} \end{bmatrix} \quad (16)$$

Noting that the south east block of  $C_\theta$  captures the error variance in pole estimates, we obtain the following result:

$$C_p = (F_P - F_{RP}^T F_R^{-1} F_{RP})^{-1} \quad (17)$$

$F_R, F_{RP}, F_P$  can be also written as block matrices relating the parameters to be estimated with the observations (9). Doing so, we obtain:

$$C_p = \left( F_{P_1} + F_{P_2} - [F_{RP_1}^T F_{RP_2}^T] \begin{bmatrix} F_{R_1}^{-1} & 0 \\ 0 & F_{R_2}^{-1} \end{bmatrix} \begin{bmatrix} F_{RP_1} \\ F_{RP_2} \end{bmatrix} \right)^{-1} \quad (18)$$

We obtain the following expression for  $C_p^{-1}$  by completing the matrix multiplications and inverting both sides of the equation,

$$C_p^{-1} = F_{P_1} + F_{P_2} - F_{RP_1}^T F_{R_1}^{-1} F_{RP_1} - F_{RP_2}^T F_{R_2}^{-1} F_{RP_2} \quad (19)$$

Finally, rearranging  $C_p^{-1}$  to separate the matrices related to each observation and also considering the inverse of the Cramer-Rao bounds on pole estimates when each observation sequence is used individually, we can conclude that:

$$C_p^{-1} = C_{p_1}^{-1} + C_{p_2}^{-1} \blacksquare \quad (20)$$

**Corollary:** Let  $C_p$  be the minimum achievable error covariance (or CRB) for the poles estimates, for the pole-residue estimation problem introduced in section II. Let  $C_{p_1}$  and  $C_{p_2}$  be the Cramer-Rao bounds on the pole estimates, assuming that only the first and second observation, respectively, are available at each time for the estimation. The minimum achievable error variance satisfies the following condition:

$$C_p \leq C_{p_1} \text{ and } C_p \leq C_{p_2}.$$

We have thus far related the minimum achievable error covariance (CRB) for poles estimates from two observations with the CRBs in the corresponding single-observation cases. We observe that the error variance is reduced, but have not quantified the amount of reduction yet. To do so, let us consider some particular measures defined from the error covariance matrix.

In general, the maximum (or minimum) eigenvalue of a covariance matrix is the maximum (or minimum) variance among unitary linear combinations of the uncertain variables, where the maximum (or minimum) is reached when the weights in the linear combination are the corresponding eigenvector's components. As such, these extremal eigenvalues are of interest as scalar performance measures defined from the error covariance. The determinant of an error covariance matrix is also used as a performance measure in nonrandom estimation, because it indicates the volume of the error ellipsoid around the true parameter value [12], and also admits interpretation as a measure of mutual information between the unknown parameter and the observations [13, 14]. Since the CRB lower-bounds estimator covariances for any estimate, we are motivated to use the extremal eigenvalues and also the determinant of the CRB (which also serve as lower bounds for the respective measures of the error covariance when any unbiased estimator is used) as performance bounds.

Let us give bounds on the maximum (and minimum) eigenvalue of the covariance matrix and an upper bound on the determinant of CRB in the multi-observation case:

**Theorem 2:** For the multi-observation estimation problem, consider  $\lambda_1(C_p)$  and  $\lambda_n(C_p)$  be the maximum and minimum eigenvalues of the error covariance matrix (CRB). Also, Let  $\lambda_1(C_{p_i})$  and  $\lambda_n(C_{p_i})$ ,  $i = 1, 2$  be the maximum and minimum eigenvalues of the CRB assuming that only the  $i^{\text{th}}$  observation is available. These eigenvalues satisfy the following four conditions:

- 1)  $\min \left( \left( \lambda_1^{-1}(C_{p_1}) + \lambda_n^{-1}(C_{p_2}) \right)^{-1}, \left( \lambda_n^{-1}(C_{p_1}) + \lambda_1^{-1}(C_{p_2}) \right)^{-1} \right) \leq \lambda_1(C_p)$
- 2)  $\lambda_1(C_p) \leq \left( \lambda_1^{-1}(C_{p_1}) + \lambda_1^{-1}(C_{p_2}) \right)^{-1}$
- 3)  $\left( \lambda_n^{-1}(C_{p_1}) + \lambda_n^{-1}(C_{p_2}) \right)^{-1} \leq \lambda_n(C_p)$
- 4)  $\lambda_n(C_p) \leq \max \left( \left( \lambda_1^{-1}(C_{p_1}) + \lambda_n^{-1}(C_{p_2}) \right)^{-1}, \left( \lambda_n^{-1}(C_{p_1}) + \lambda_1^{-1}(C_{p_2}) \right)^{-1} \right)$

**Proof:**

The proof of Theorem 2 is omitted to save space. The inequalities claimed in the theorem follow easily from Equation (20), using Weyl's inequalities and the relationship between the eigenvalues of a matrix and its inverse. Please see the extended document [15] for the details.

**Theorem 3:** For the multi-observation estimation problem, consider the determinant of the error covariance matrix  $\det(C_p)$  associated with the CRB of the pole estimate. Let  $\det(C_{p_i})$ ,  $i = 1, 2$  be the determinant of the CRB in the case that only the  $i^{\text{th}}$  observation is used. The determinant of the covariance matrix  $\det(C_p)$  satisfies the following condition:

$$\det(C_p) \leq \frac{\det(C_{p_1}) \det(C_{p_2})}{\prod_{i=1}^n (\lambda_i(C_{p_1}) + \lambda_i(C_{p_2}))}$$

**Proof:**

The proof of Theorem 3 is omitted to save space. The result follows from the classical spectral bounds on the determinant [16] together with Weyl's inequalities. Please see extended document [15] for details.

Now, let us present a simple bound for the estimation of a specific pole, and so give insight into the role played by the pole location and the multiple-observation structure in the error variance. This result is a direct generalization of a bound given in [10], for the single-observation case.

**Theorem 4:** For the multi-observation estimation problem, consider the Cramer-Rao lower bound  $C_{\theta_i}$  on the error variance for the estimate of a particular pole  $p_i$  (where  $p_i$  is inside the unit circle by assumption). The error variance satisfies  $C_{\theta_i} \geq \frac{(1-p_i^2)^3}{(A_i^2+B_i^2)(1+p_i^2)}$ . This bound is approximately  $\frac{1}{(A_i^2+B_i^2)}$  for a pole  $p_i$  near 0 and  $\frac{4(1-p_i)^3}{(A_i^2+B_i^2)}$  for a pole  $p_i$  near 1.

**Proof:** With a little algebra, we find that the diagonal entry of the Fisher information matrix corresponding to the pole  $p_i$  is  $\frac{(A_i^2+B_i^2)(1+p_i^2)}{(1-p_i^2)^3}$ . Therefore, the error variance is bounded by the inverse of this entry.  $\blacksquare$

Let us make some observations about this lower bound and also compare it to the analogous lower bound for the single observation case. As we expect, there is a reduction in

the lower bound from that calculated for one observation [10], specifically  $\frac{(1-p_i^2)^3}{A_i^2(1+p_i^2)}$ .

We have characterized the relationship between the pole location and the minimum achievable error variance on the estimates of a particular pole. Next, we study the consequence of having two poles near each other on estimation error variances, for the multiple observations problem. The result closely follows an analysis in [10].

**Theorem 5:** Consider a system that has two poles  $p_0$  and  $p_1$  close to each other, specifically  $p_1 = p_0 - \delta$  for some small  $\delta$  (where  $p_0$  and  $p_1$  are inside the unit circle). There is a lower bound on the error variance of the efficient estimate that is on the order of  $\frac{1}{\delta^2}$ .

**Proof:** Analogously to [10], we know that the Cramer-Rao bound on pole estimate satisfies:

$$C_p \geq \left( \frac{\partial x_i^T}{\partial p} \frac{\partial x_i}{\partial p} + \frac{\partial x_j^T}{\partial p} \frac{\partial x_j}{\partial p} \right)^{-1}, \quad (21)$$

or equivalently

$$C_p \geq (F_{px_i} + F_{px_j})^{-1} = (A_d P^T P A_d + B_d P^T P B_d)^{-1} \quad (22)$$

Following [10], let us use the decomposition  $P^T = U^T \hat{P}^T$  where  $U^T = \begin{bmatrix} 1 & -1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & I_{n-2} \end{bmatrix}$ . In this case,  $\hat{P}^T$  is

identical to  $P^T$  except that the first row of  $\hat{P}^T$  is the difference between the first and second rows of  $P^T$ . Then the Cramer-Rao bound for the pole estimate is given by:

$$C_p \geq (A_d U^T \hat{P}^T \hat{P} U^{-1} A_d + B_d U^T \hat{P}^T \hat{P} U^{-1} B_d)^{-1} \quad (23)$$

We then note that

$$\hat{P}^T \hat{P} = \begin{bmatrix} 4\delta^2 & 4\delta(p_0 - \delta) & 4\delta p_2 & \dots \\ 4\delta(p_0 - \delta) & 1 + 4(p_0 - \delta)^2 & 1 + 4(p_0 - \delta)p_2 & \dots \\ 4\delta p_2 & 1 + 4(p_0 - \delta)p_2 & 1 + 4p_2^2 & \dots \\ \vdots & \vdots & \vdots & \ddots \end{bmatrix} \quad (24)$$

Now consider  $\hat{C}_p = (A_d \hat{P}^T \hat{P} A_d + B_d \hat{P}^T \hat{P} B_d)^{-1}$ . Through matrix multiplication, we obtain:

$$\hat{C}_p = \left( \begin{bmatrix} 4A_0^2 \delta^2 & 4A_0 A_1 \delta(p_0 - \delta) & \dots \\ 4\delta A_0 A_1 (p_0 - \delta) & A_1^2 (1 + 4(p_0 - \delta)^2) & \dots \\ 4A_0 A_2 \delta p_2 & A_1 A_2 (1 + 4(p_0 - \delta)p_2) & \dots \\ \vdots & \vdots & \ddots \end{bmatrix} + \begin{bmatrix} 4B_0^2 \delta^2 & 4B_0 B_1 \delta(p_0 - \delta) & \dots \\ 4\delta B_0 B_1 (p_0 - \delta) & B_1^2 (1 + 4(p_0 - \delta)^2) & \dots \\ 4B_0 B_2 \delta p_2 & B_1 B_2 (1 + 4(p_0 - \delta)p_2) & \dots \\ \vdots & \vdots & \ddots \end{bmatrix} \right)^{-1} \quad (25)$$

Adding both matrices inside the parentheses, we obtain

$$\hat{C}_p = \left( \begin{bmatrix} 4\delta^2 (A_0^2 + B_0^2) & 4(A_0 A_1 + B_0 B_1) \delta(p_0 - \delta) & \dots \\ 4\delta(A_0 A_1 + B_0 B_1)(p_0 - \delta) & (A_1^2 + B_1^2)(1 + 4(p_0 - \delta)^2) & \dots \\ 4(A_0 A_2 + B_0 B_2) \delta p_2 & (A_1 A_2 + B_1 B_2)(1 + 4(p_0 - \delta)p_2) & \dots \\ \vdots & \vdots & \ddots \end{bmatrix} \right)^{-1} \quad (26)$$

With a little effort, we can prove that  $\hat{C}_{p_0} \geq \frac{1}{4\delta^2(A_0^2 + B_0^2)}$ , and hence  $C_{p_0} \geq \frac{1}{4\delta^2(A_0^2 + B_0^2)}$  ■

#### IV. GENERALIZATION

Thus far, we have characterized the Cramer-Rao bound on eigenvalue estimation error, for the problem of eigenvalue inference from impulse response data in the two-observation case. We find it instructive to generalize the results to the case that  $m$  state variables are observed (where  $m$  is less than the number of state variables  $n$ ). Let us here briefly present the results for the  $m$  observation case; we omit the proofs, since they are so similar to those for the two-observation case.

First, it is straightforward to show that the block of the Fisher information matrix associated with the pole estimates is  $F_p = \sum_{k=1}^m F_{p_k}$ , where  $F_{p_k}$  is the corresponding block of the Fisher information matrix for the estimation problem where only the  $k^{th}$  observation is available. The Cramer-Rao Bound then satisfies  $C_p^{-1} = \sum_{k=1}^m C_{p_k}^{-1}$ , where  $C_{p_k}$  is the Cramer-Rao bound on pole estimates assuming only that the  $k^{th}$  observation was available. Moreover, we find that the lower bound of an optimal estimator for the multi-observation problem satisfies the following inequality:

$$C_p \leq C_{p_k} \text{ for all } k = 1, 2, \dots, m \quad (27)$$

Thus, as in the two-observation case, pole estimation using  $m$  observations yields a lower error covariance than estimation using any single observation.

Let us present results on the maximum and minimum eigenvalue and determinant of the CRB, when  $m$  observations are available.

First of all, let us introduce the variable  $\tilde{C}_{p_i}$  which represents the minimum achievable error covariance (CRB) when the observations  $i, i+1, i+2, \dots, m$  are available, for each  $i = 1, 2, 3, \dots, m$ . The CRB  $\tilde{C}_{p_i}$  satisfies:

$$\tilde{C}_{p_i}^{-1} = C_{p_i}^{-1} + \tilde{C}_{p_{i+1}}^{-1} \text{ for } i = 1, 2, 3, \dots, m-1 \quad (28)$$

while

$$\tilde{C}_{p_m} = C_{p_m} \quad (29)$$

Analogously to the proof of theorem 2, the following relation between the maximum eigenvalue of the CRB holds:

$$\lambda_1^{-1}(C_p) \geq (\lambda_1^{-1}(C_{p_1}) + \lambda_n(\tilde{C}_{p_2}^{-1})). \quad (30)$$

Since  $\lambda_n(\tilde{C}_{p_2}^{-1})$  can be written as the sum of the eigenvalues of  $C_{p_2}^{-1}$  and  $\tilde{C}_{p_3}^{-1}$ , we have that:

$$\lambda_1^{-1}(C_p) \geq (\lambda_1^{-1}(C_{p_1}) + \lambda_1^{-1}(C_{p_2}) + \lambda_n(\tilde{C}_{p_3}^{-1})). \quad (31)$$

If we keep replacing the minimum eigenvalue of  $\tilde{C}_{p_i}^{-1}$ ,  $i = 3, 4, \dots, m$ , for the sum of eigenvalues of  $C_{p_i}^{-1}$  and  $\tilde{C}_{p_{i+1}}^{-1}$ , we find that:

$$\lambda_1^{-1}(C_p) \geq \sum_{k=1}^m \lambda_1^{-1}(C_{p_k}) \quad (32)$$

Therefore, the maximum eigenvalue of the CRB is bounded by:

$$\lambda_1(C_p) \leq (\sum_{k=1}^m \lambda_1^{-1}(C_{p_k}))^{-1} \quad (33)$$

We use a similar argument to find the bound for the minimum eigenvalue of the CRB:

$$\lambda_n(C_p) \geq (\sum_{k=1}^m \lambda_n^{-1}(C_{p_k}))^{-1} \quad (34)$$

It is not difficult to argue that when more information (observations) is available for the pole estimation, the performance of the estimator cannot be worse than having just one observation. Moreover, because of (33), we know that:

$$\lambda_1(C_p) \leq \min(\lambda_1(C_{p_i})) \quad (35)$$

On the other hand, from the proof of theorem 3 we know that:

$$\det(C_p^{-1}) \geq \prod_{i=1}^n (\lambda_i(C_{p_1}^{-1}) + \lambda_i(\tilde{C}_{p_2}^{-1})) \quad (36)$$

By writing the eigenvalue of  $\tilde{C}_{p_2}^{-1}$  as the sum of eigenvalues of  $C_{p_2}^{-1}$  and  $\tilde{C}_{p_3}^{-1}$ , we have that:

$$\det(C_p^{-1}) \geq \prod_{i=1}^n (\lambda_i(C_{p_1}^{-1}) + \lambda_i(C_{p_2}^{-1}) + \lambda_n(\tilde{C}_{p_3}^{-1})). \quad (37)$$

We write the minimum eigenvalue of  $\tilde{C}_{p_3}^{-1}$  as the sum of eigenvalues of  $C_{p_3}^{-1}$  and  $\tilde{C}_{p_4}^{-1}$ , to find that:

$$\det(C_p^{-1}) \geq \prod_{i=1}^n (\lambda_i(C_{p_1}^{-1}) + \lambda_i(C_{p_2}^{-1}) + \lambda_n(C_{p_3}^{-1}) + \lambda_n(\tilde{C}_{p_4}^{-1})) \quad (38)$$

We keep writing  $\tilde{C}_{p_i}^{-1}$ ,  $i = 3, 4, \dots, m$ , as the sum of  $C_{p_i}^{-1}$  and  $\tilde{C}_{p_{i+1}}^{-1}$ , thus finding that:

$$\det(C_p) \leq \prod_{i=1}^n (\lambda_i^{-1}(C_{p_1}) + \lambda_i^{-1}(C_{p_2}) + \sum_{k=3}^m \lambda_i^{-1}(C_{p_k}))^{-1} \quad (39)$$

Or equivalently:

$$\det(C_p) \leq \prod_{i=1}^n (\lambda_i^{-1}(C_{p_1}) + \lambda_i^{-1}(C_{p_2}) + \sum_{k=3}^m \lambda_i^{-1}(C_{p_k}))^{-1} \quad (40)$$

Hence, the determinant of the CRB, when  $m$  observations are available, is bounded by:

$$\det(C_p) \leq \prod_{i=1}^n (\sum_{k=1}^m \lambda_i^{-1}(C_{p_k}))^{-1} \leq \min(\lambda_1(C_{p_i}))^n \quad (41)$$

Let us now consider the case of inference of a specific pole  $p_i$  from  $m$  observations. In this case, the error variance satisfies:

$$C_{p_i} \geq \frac{(1-p_i^2)^3}{\sum_{k=1}^m Z_{k,i}^2 (1+p_i^2)^2}, \quad (42)$$

where  $Z_{k,i}$  is the residue corresponding to the  $k^{th}$  observation and the pole  $p_i$ . This bound is approximately  $\frac{1}{\sum_{k=1}^m Z_{k,i}^2}$  for the pole  $p_i$  near 0 and  $\frac{4(1-p_i^2)^3}{\sum_{k=1}^m Z_{k,i}^2}$  for the pole  $p_i$  near 1.

The lower bound on error variance for estimation of two poles close to each other, say  $p_1 = p_0 - \delta$  in the multi-observation problem, is still on the order of  $\frac{1}{\delta^2}$  independently of the number of observations.

## V. CONCLUSION

We have characterized the CRB on pole estimates from impulse-response data for a single-input multiple-output system, in terms of the poles and residues of its transfer function. Specifically, we have developed a relationship on the CRB for multiple- and single- observation case, which shows that CRB for the multiple observation case is a lower bound for the CRB for any single observation case. Moreover, the maximum error variance when multiple observations are made is bounded by the minimum among the error variances for each single-observation case. We have emphasized the role played for both pole location and residue in the CRB. Of interest, we have found that

estimating poles that are close to each other is difficult, independent of the number of observations made. Another interesting conclusion is that the lower bound on the pole estimate depends on the sum of the residues, in contrast with the single observation case.

## REFERENCES

- [1] G. R. deProny and R. Baron, Sur Essai experimental et analytique: les los de la dilatabilité de fluidselastiques et sur celle de la forcée expansive de la vapeur de l'eau et de la vapeur de l'alcool a diferentes temperaturas, vol. 1, Paris: J. l'EcolePolytechnique, 1795, pp. 24-76.
- [2] L. L. Scharf, Statistical Signal Processing: Detection, Estimation, and Time-Series Analysis, Addison-Wesley: Reading MA, 1991.
- [3] L. T. McWhorter and L. L. Scharf, "Cramer-Rao Bounds for Deterministic Modal Analysis," in *IEEE Trans. Signal Proc.*, 1993.
- [4] M. P. Clark, "Cramer-Rao bounds for two-dimensional deterministic modal analysis," in *Signals, Systems and Computers Conference Record of The Twenty-Seventh Asilomar*, 1993.
- [5] M. P. Clark and L. L. Scharf, "Two-dimensional modal analysis based on maximum likelihood," in *EEE Trans. Signal Proc.*, 1994.
- [6] A. O. Hero, "A Cramer Rao Bound Type Lower Bound for Essentially Unbiased Parameter Estimation," Lincoln Laboratory, Massachusetts Institute for Technology, 1992.
- [7] R. C. Dahyia and R. M. Korwar, "Maximun Likelihood Estimates for a Bivariate Normal Distribution with Missing Data," in *The Annals of Statiscis*, 1980.
- [8] L. Guo and D. Huang, "Least-squares identification for ARMAX models without the positive real condition," in *IEEE Transactions on Automatic Control*, 1989.
- [9] L. Guo and D. Huang, "Estimation of Nonstationary Armax Models Based on the Hannan-Rissanen Method," in *The Annals of Statistics*, 1990.
- [10] Y. Wan and S. Roy, "On inference of network time constants from impulse-response data: graph-theoretic Cramer-Rao bounds," in *Proceedings of the 48th IEEE Conference on Decision and Control held jointly with the 2009 28th Chinese Control Conference*, 2009.
- [11] D. J. Trudnowski, J. M. Johnson and J. F. Hauer, "Making Prony analysis more accurate using multiple signals," in *EEE Transactions on Power Systems*, 1999.
- [12] O. P. Bagai, "The Distribution of the Generalized Variance," in *The Annals of Mathematical Statistics*, 1965.
- [13] V. V. Fedorov, Theory of Optimal Experiments, New York: Academic Press, 1972.
- [14] M. Xue, S. Roy, Y. Wan and S. K. Das, "Security and Vulnerability of Cyber-Physical Infrastructure Networks: A Control-Theoretic Approach," in *Handbook on Securing Cyber-Physical Critical Infrastructures: Foundations and Challenges*, Morgan-Kauffman, 2012.
- [15] J. Abad Torres and S. Roy, "Cramer-Rao Bounds on Eigenvalue Estimates from Impulse Response Data: The Multi-Observation Case (extended version)," *PDF available at www.eecs.edu/~jabad*.
- [16] F. Miroslav, "Bounds for the Determinant of the Sum of Hermitian Matrices," in *Proceedings of the American Mathematical Society*, 1971.