BOUNDS ON MIMO CHANNEL ESTIMATION AND EQUALIZATION WITH SIDE INFORMATION

Brian M. Sadler[1], *Richard J. Kozick*[2], *and Terrence Moore*[1]

[1]Army Research Laboratory, Adelphi, MD 20783, bsadler@arl.army.mil
[2]Bucknell University, Lewisburg, PA 17837, kozick@bucknell.edu

ABSTRACT

We present constrained Cramér-Rao bounds for multi-input multi-output (MIMO) channel and source estimation. We find the MIMO Fisher information matrix (FIM) and consider its properties, including the maximum rank of the unconstrained FIM, and develop necessary conditions for the FIM to achieve full rank. Equality constraints provide a means to study the potential value of side information, such as training (semi-blind case), constant modulus (CM) sources, or source non-Gaussianity. Nonredundant constraints may be combined in an arbitrary fashion, so that side information may be different for different sources. The bounds are useful for evaluating various MIMO source and channel estimation algorithms. We present an example using the constant modulus blind equalization algorithm.

1. INTRODUCTION

Side information, such as the constant modulus property, non-Gaussianity, or training (the semi-blind case), is highly informative and often exploited in signal processing. This is especially true for multi-channel wireless communications, e.g., see [1]. In this paper we develop Cramér-Rao bounds (CRBs) for channel and source estimation in convolutive multi-input multi-output (MIMO) scenarios, when side information is available. We employ the constrained CRB formulation of Gorman / Hero [2] and Stoica / Ng [3, 4]. This approach provides a general framework that yields CRBs in a large variety of cases, and allows combination of different side information for different sources. We use a deterministic model for the sources and channels; a random Gaussian source model does not model the source attributes, such as CM and non-Gaussianity, that are commonly exploited in channel and source estimation algorithms. The constrained CRB approach may also be applied for narrowband instantaneous mixing (uncalibrated arrays [5] and calibrated arrays [6, 7]), and space-time coding [5].

2. SOURCE AND CHANNEL MODEL

We write the SIMO model as follows; e.g., see [1, 8]. The complex baseband representation of the M-channel FIR system is given by $y_i(k) = \sum_{l=0}^{L} s(k-l)h_i(l) + w_i(k)$, for $1 \leq i \leq M$ channels, and $0 \leq k \leq N-1$ output samples. The maximum channel order is denoted L, and is assumed known; there are $N+L$ input samples and N output samples. The noise $w_i(k)$ is assumed to be circular Gaussian and white with variance σ^2.

The model may be expressed as an $MN \times 1$ vector $\mathbf{y}$, given by

$$\mathbf{y} = H_M \mathbf{s} + \mathbf{w}, \tag{1}$$

with input $\mathbf{s} = [s(-L), \cdots, s(0), \cdots, s(N-1)]^T$, output $\mathbf{y} = [\mathbf{y}_1^T, \cdots, \mathbf{y}_M^T]^T$, $\mathbf{y}_i = [y_i(0), \cdots, y_i(N-1)]^T$, and channel matrix

$$H_M = \begin{bmatrix} H_{(1)} \\ \vdots \\ H_{(M)} \end{bmatrix}_{MN \times (N+L)}, \tag{2}$$

where $H_{(i)}$ is the ith channel (convolution) matrix given by

$$H_{(i)} = \begin{bmatrix} h_i(L) & \cdots & h_i(0) & & \\ & \ddots & & \ddots & \\ & & h_i(L) & \cdots & h_i(0) \end{bmatrix}_{N \times (N+L)}. \tag{3}$$

Based on the noise assumptions it follows that $E[\mathbf{w}\mathbf{w}^H] = \sigma^2 \mathbf{I}$.

The K-user MIMO model may be written

$$y_i(k) = s^{(1)}(k) * h_i^{(1)}(k) + \cdots + s^{(K)}(k) * h_i^{(K)}(k) + w_i(k), \tag{4}$$

where, for example, $s^{(K)}(k)$ denotes the Kth user and $h_i^{(K)}(k)$ denotes the ith sub-channel of the Kth user. We extend (1) to obtain

$$\mathbf{y} = \sum_{k=1}^{K} H_M^{(k)} \mathbf{s}^{(k)} + \mathbf{w}. \tag{5}$$

Here, for example, $H_M^{(k)} \mathbf{s}^{(k)}$ denotes the contribution of the kth user where

$$H_M^{(k)} = \begin{bmatrix} H_1^{(k)} \\ \vdots \\ H_M^{(k)} \end{bmatrix}, \tag{6}$$

and $\mathbf{s}^{(k)}$ is the kth user input. (The remaining notational extensions are straightforward.)

3. THE FISHER INFORMATION MATRIX

Next we derive the Fisher information matrix (FIM) for the model in (1). This has been derived by Hua [8]; we give a brief derivation using the complex form of the FIM and then extend these results to the multi-user case. Define the complex vector of unknown parameters as

$$\theta = \left[\mathbf{h}^T, \mathbf{s}^T\right]^T, \tag{7}$$

$$\mathbf{h} = [\mathbf{h}_1^T, \cdots, \mathbf{h}_M^T]^T \quad \text{and } \mathbf{h}_i = [h_i(0), \cdots, h_i(L)]^T. \tag{8}$$

Writing θ in terms of its real and imaginary parts as $\theta = \overline{\theta} + j\widetilde{\theta}$, define the real parameter vector as

$$\xi = \left[\overline{\theta}^T, \widetilde{\theta}^T\right]^T. \tag{9}$$

For (1), the FIM has the general block form

$$J(\xi) = 2\begin{bmatrix} E & -F \\ F & E \end{bmatrix}. \tag{10}$$

The complex FIM may be defined as $J_c(\theta) = E + jF$. Note that the vector $\mathbf{y}$ in (1) has complex normal distribution with mean vector $\boldsymbol{\mu}(\theta) = H_M \mathbf{s}$ and covariance matrix $\sigma^2 \mathbf{I}$. We can show that the complex FIM for (1) may be written

$$[J_c(\theta)]_{ij} = \frac{2}{\sigma^2}\frac{\partial \boldsymbol{\mu}^H(\theta)}{\partial \theta_i^*}\frac{\partial \boldsymbol{\mu}(\theta)}{\partial \theta_j}, \tag{11}$$

where θ_i^* denotes the complex conjugate of θ_i.

Now, it is straightforward to show that

$$\frac{\partial \boldsymbol{\mu}(\theta)}{\partial \theta^T} = [\mathbf{I}_M \otimes S, \;\; H_M] \triangleq Q, \tag{12}$$

where $\mathbf{I}_M$ is the $M \times M$ identity matrix, $\otimes$ denotes Kronecker product, and

$$S = \begin{bmatrix} s(0) & s(-1) & \cdots & s(-L) \\ s(1) & s(0) & \cdots & s(-L+1) \\ \vdots & \vdots & \ddots & \vdots \\ s(N-1) & s(N-2) & \cdots & s(N-L-1) \end{bmatrix}. \tag{13}$$

Using (11), the complex FIM for (1) is given by

$$J_c = \frac{2}{\sigma^2} Q^H Q. \tag{14}$$

The real-valued FIM is obtained by letting $E = \mathrm{Re}(J_c)$, and $F = \mathrm{Im}(J_c)$ in (10).

Next we develop the FIM for the multi-user model in (5). Now, with K users, the complex parameter vector is

$$\theta_K = \left[(\mathbf{h}^{(1)})^T, (\mathbf{s}^{(1)})^T, \cdots, (\mathbf{h}^{(K)})^T, (\mathbf{s}^{(K)})^T\right]^T, \tag{15}$$

and the mean of $\mathbf{y}$ is now

$$\boldsymbol{\mu}_K(\theta_K) = \sum_{k=1}^{K} H_M^{(k)} \mathbf{s}^{(k)}. \tag{16}$$

Using (12),

$$\frac{\partial \boldsymbol{\mu}_K(\theta_K)}{\partial \theta_K^T} = [Q_1, \cdots, Q_K] \triangleq \mathcal{Q}, \tag{17}$$

where

$$Q_k = \left[\mathbf{I}_M \otimes S^{(k)}, \;\; H_M^{(k)}\right], \;\; 1 \le k \le K. \tag{18}$$

Now using (14), the complex multi-user FIM is

$$J_c^K = \frac{2}{\sigma^2}\mathcal{Q}^H \mathcal{Q} = \frac{2}{\sigma^2}\begin{bmatrix} Q_1^H Q_1 & Q_1^H Q_2 & \cdots & Q_1^H Q_K \\ \vdots & \vdots & \ldots & \vdots \\ Q_K^H Q_1 & Q_K^H Q_2 & \cdots & Q_K^H Q_K \end{bmatrix}. \tag{19}$$

The real-valued FIM is obtained by letting $E = \mathrm{Re}(J_c^K)$, and $F = \mathrm{Im}(J_c^K)$ in (10), with corresponding real parameter vector

$$\xi_K = \left[\overline{\theta}_K^T, \; \widetilde{\theta}_K^T\right]^T. \tag{20}$$

Next we explore the properties of the MIMO FIM. Theorem proofs may be found in [9].

Theorem 1: $\mathrm{nullity}(J_c^K) \ge K^2$

Theorem 1 provides the maximum rank of J_c^K; this maximum will not be surpassed by increasing N, M, or L. As noted, the real parameter case FIM (denoted J^K) may be obtained from (10), and then $\mathrm{nullity}(J^K) = 2\cdot\mathrm{nullity}(J_c^K)$. In the single user case $\mathrm{nullity}(J_c) \ge 1$, corresponding to the multiplicative ambiguity in blind SIMO problems (see Theorem 2 of [8]). Intuitively, under certain conditions the K-user convolutional MIMO problem may be equalized, yielding a memoryless MIMO problem that has a $K \times K$ matrix ambiguity remaining (e.g., see the discussion of MIMO MA model identifiability in section III.A. of [10]).

The results of Theorem 1 may be related to the number of variables; notice that there are MN equations and $K(N + L + ML + M)$ unknowns in the MIMO model. We have the following.

Corollary 1: $\mathrm{nullity}(J_c^K) \ge \max\{K^2, K(N + L + ML + M) - MN\}$

Note that the Corollary implies that we must have $M > K$ in order for the FIM to achieve its maximum rank. More precisely, $K^2 < K(N + L + ML + M) - MN$ if $M = K$ or if

$$M < K \quad \text{and} \quad N > K - \frac{KL(M+1)}{K-M}, \tag{21}$$

where the constraint on N in (21) is nontrivial (> 1) when $L < \frac{K-1}{K}\frac{K-M}{M+1}$. Corollary 1 leads to a necessary condition for $\mathrm{nullity}(J_c) = K^2$, for in this case we must have

$$K^2 \ge K(N + L + ML + M) - MN. \tag{22}$$

Rearranging (22) we have that the number of equations must be greater than or equal to the number of unknowns minus K^2. Solving (22) for N we obtain a necessary condition on the data length,

$$N \ge K(L+1) + \frac{KL(K+1)}{M-K} = K + \frac{KL(M+1)}{M-K}. \tag{23}$$

Condition (23) is not always sufficient. However, numerical testing reveals that when (23) fails to be sufficient, it still provides a good approximation for the minimum value of N.

In the SIMO case ($K = 1$), equation (23) reduces to $N \geq L+1+\frac{2L}{M-1}$. Notice that for all $M \geq 2$, and for all L, we have $\frac{2L}{M-1} > 0$. Therefore, in the SIMO case, $N \geq L+2$ is always necessary.

It is also of interest to specify when nullity$(F_c^K) = K^2$. Generally, identifiability and regularity require the sources to be persistently exciting of sufficient order. Alternatively, for finite deterministic sequences, this idea can be expressed in terms of the number of modes necessary to be present in the source. The modes are independent basis functions that may be used to describe any finite length sequence. A sufficient condition for $K = 1$ is that the number of modes be greater than or equal to $2L + 1$; additional conditions are that the SIMO sub-channels do not have common zeros and that $N \geq 3L + 1$ [8]. However, the situation is more complicated with $K > 1$, for now the required value of N depends on M, K, and L. Equation (23) provides a good approximation when the input has sufficient modes; the resulting necessary values of N are not large, being on the order of KL.

In the SISO case ($K = 1, M = 1$), there are N equations in $L + 1$ channel plus $N + L$ source unknowns. Thus we always have $2L+1$ more unknowns than equations, and we find that nullity$(J_c) \geq 2L+1$ in this case, with equality achieved when the input has sufficient modes.

When nullity$(J_c^K) = K^2$ then it is of interest to specify K^2 complex parameters in θ so that the resulting row and column-reduced FIM will have full rank. This can be seen from the following Theorem for two users. The proof of Theorem 2 utilizes the null space basis vectors found in the proof of Theorem 1.

Theorem 2: Let $K = 2$ and assume nullity$(J_c^K) = 4$. Let $\mathcal{J}$ denote J_c^K with four row-column pairs removed, i.e., by specifying four complex parameters $\theta_i \in \theta$, $i = 1, 2, 3, 4$. Then, nullity$(\mathcal{J}) = 0$ if the θ_i are chosen in any of the following ways.

a) $\theta_1, \theta_2 \in \mathbf{h}^1, \theta_3, \theta_4 \in \mathbf{h}^2$, with at least one of $\theta_1, \theta_2 \neq 0$, and one of $\theta_3, \theta_4 \neq 0$.

b) $\theta_1, \theta_2 \in \mathbf{s}^1, \theta_3, \theta_4 \in \mathbf{s}^2$, with at least one of $\theta_1, \theta_2 \neq 0$, and one of $\theta_3, \theta_4 \neq 0$;

c) θ_i, $i = 1, 2, 3$, from each of any unique three in the set $\{\mathbf{h}^1, \mathbf{s}^1, \mathbf{h}^2, \mathbf{s}^2\}$, and θ_4 in any element of the set.

Unlike the $K = 1$ case, we cannot arbitrarily specify four parameters, e.g., we cannot specify four parameters in the set $\{\mathbf{h}^1, \mathbf{s}^1\}$ and achieve a full rank FIM. Rather, some parameters must be specified for both users. This idea generalizes for $K > 2$, although the proof is cumbersome. The lack of FIM regularity in blind single-user SIMO channel estimation problems is often circumvented by assuming that one of the complex channel coefficients is known, resulting in a full rank FIM. Theorem 2 states that, for $K = 2$, at least two parameters must be specified for one of the users to obtain a full rank FIM.

4. CONSTRAINED CRBS

We work with the real-valued FIM J, and the corresponding real-valued parameter vector ξ in (9). Consider N_c equality constraints on elements of θ, where $N_c < D_\theta = \dim(\theta)$. The constraints have the form $f_i(\theta) = 0$ for $i = 1, \ldots, N_c$. The constraints form a $N_c \times 1$ vector $f(\theta)$, and we define a $N_c \times D_\theta$ gradient matrix $F(\theta) = \frac{\partial f(\theta)}{\partial \theta}$ with elements $[F(\theta)]_{i,m} = \partial f_i(\theta)/\partial[\theta]_m$. $F(\theta)$ is assumed to have full row rank N_c for any θ satisfying the constraints $f_1(\theta), \ldots, f_{N_c}(\theta)$. Let U be a $D_\theta \times (D_\theta - N_c)$ matrix whose columns are an orthonormal basis for the null space of F, so that $FU = \mathbf{0}$, and $U^T U = \mathbf{I}$ Then, Stoica and Ng have shown that as long as $U^T J U$ is invertible, the constrained CRB is given by (Thrm. 1 of [3])

$$E\left[(\hat{\theta} - \theta)^T(\hat{\theta} - \theta)\right] \geq U(U^T J U)^{-1} U^T. \quad (24)$$

Notice that J in (24) is the unconstrained CRB, while U is solely a function of the constraints. Properties of inverses of partitioned matrices may be exploited to find closed form expressions for the resulting constrained CRBs [7]. This requires specification of U for a particular constraint set, which is typically not too difficult. Alternatively, given $F(\theta)$ and J, both U and (24) may be evaluated numerically. For examples, see [5, 6, 7, 9].

5. EXAMPLE AND DISCUSSION

To illustrate some of the ideas, we present a single-input single-output ($M = 1$) example. We compare two blind linear equalizers with a constrained CRB. The source is QPSK with iid symbols and unit modulus, and we use symbol synchronous sampling. The block length is N, and SNR $= \mathbf{h}^H \mathbf{h}/\sigma^2$. The equalizer length is $L_e + 1 = 31$. The channel is given by $\mathbf{h} = [0.53+j0.07, -0.24-j0.23, -0.54-j0.32; 0.11+j0.44, -0.036-j0.099]^T$. The equalizer yields source estimates

$$\hat{s}(k-d) = \sum_{\ell=-L_e/2}^{L_e/2} w_\ell \, y(k-\ell), \quad (25)$$

for $k = \frac{L_e}{2}, \ldots, N-1-\frac{L_e}{2}$, $d \in \{-\frac{L_e}{2}, \ldots, \frac{L_e}{2}\}$, where w_ℓ and d are the equalizer weights and delay. Only $N - L_e$ of the $N + L$ signal values are estimated, eliminating block edge effects.

We compare performance of the constant modulus algorithm (CMA) and the alphabet-matched algorithm (AMA) [11, 12]. Both CMA and AMA employ a block-averaged gradient method for smoother convergence, as in [11, 12]. The CRB incorporates the CM signal constraint with unit modulus. The CM constraint alone does not provide a full rank FIM, so we we additionally constrain one signal sample to be known that is block centered at $s(N/2 - 1)$. The CMA exploits the CM signal property directly, while AMA exploits the discrete alphabet constraint in a soft manner [12]. We note that applying the discrete alphabet constraint does not yield a useful bound on variance [7].

Because the FIR SISO channel cannot be perfectly equalized with an FIR filter, residual inter-symbol interference (ISI) will remain. This residual ISI can be bounded

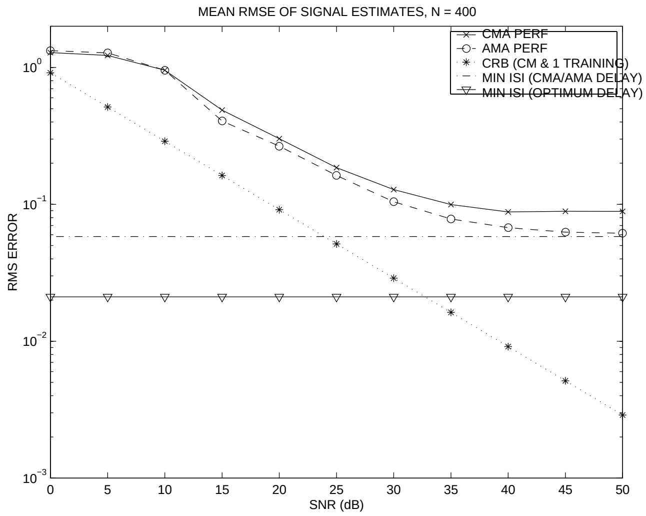

Figure 1: Simulation results and bounds on symbol estimation error versus SNR, with block length $N = 400$.

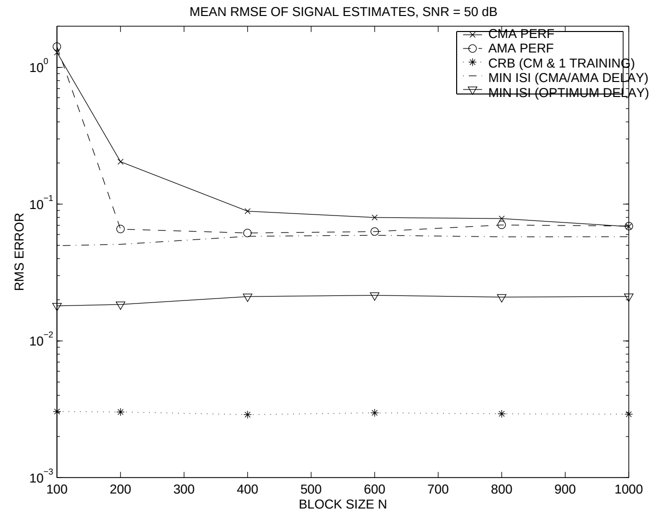

Figure 2: Simulation results and bounds on symbol estimation error versus block length N, with SNR = 50 dB.

by designing a noise-free known-signal equalizer, and measuring the ISI in this case. We refer to this value as the minimum ISI (MIN ISI). Note that this value is a function of the equalizer delay d.

Figure 1 shows AMA and CMA results for $N = 400$, along with the constrained CRB. The AMA and CMA converge to weight vectors with (non-optimal) delay $d = 2$, primarily because the gradient descent was center tap initialized with weight vector $w_0 = 1$ and $w_\ell = 0$ for $\ell \neq 0$. The line labeled "MIN ISI (CMA/AMA DELAY)" is the residual ISI of the noise-free, known-signal equalizer using the same value of delay $d = 2$ as the CMA and AMA. The line labeled "MIN ISI (OPTIMUM DELAY)" corresponds to the *minimum* ISI over all delays.

Figure 2 shows the effect of increasing the block length for a high SNR case. Both CMA and AMA asymptotically (in N and SNR) achieve the minimum ISI bound for the appropriate delay. They did not achieve the minimum possible ISI limit as they did not converge to the optimal delay. The AMA algorithm displays some advantage over CMA at lower block sizes, in the range of $200 \leq N \leq 400$. AMA also generally converges faster than CMA.

The constrained CRB provides a fundamental bound on the potential improvement if a nonlinear equalizer is employed (such as one employing decision feedback), over that attainable with the linear equalizer in the SISO case. Together, the constrained CRB and ISI bounds delineate SNR regimes in which the linear equalizer is "noise limited" (when the CRB is larger than ISI) and "residual ISI limited" (when the ISI is larger than the CRB). The particular bound in this example incorporates the CM source property. It is interesting to consider other constraints that may lead to potentially lower bounds, and to determine if appropriate algorithms might reach such bounds.

6. REFERENCES

[1] L. Tong, S. Perreau, "Multichannel blind identification: from subspace to maximum likelihood methods," *IEEE Proc.*, Vol. 86, No. 10, pp. 1951–1968, October 1998.

[2] J. D. Gorman, A. O. Hero "Lower bounds for parametric estimation with constraints," *IEEE Trans. Info. Theory*, Vol. 26, no. 6, pp. 1285–1301, November 1990.

[3] P. Stoica, B. C. Ng, "On the Cramer-Rao bound under parametric constraints," *IEEE Sig. Proc. Letters,* vol. 5, no. 7, pp. 177–179, July 1998.

[4] P. Stoica, B. C. Ng, "Performance bounds for blind channel estimation," to appear in *Sig. Proc. Adv. in Wireless Comm., Vol. 1*, G. B. Giannakis, Y. Hua, P. Stoica, and L. Tong, eds. (Prentice-Hall, 2001).

[5] B. M. Sadler, R. J. Kozick, "Bounds on uncalibrated array signal processing," *Proc. SSAP'2000,* pp. 73–77, Poconos, PA, August 2000.

[6] B. M. Sadler, R. J. Kozick, T. Moore, "Semi-blind array processing with constant modulus signals," *Proc. Multiconf. on Syst., Cyb. and Inform. (SCI 2000)*, Vol. VI, pp. 241–246, Orlando, FL, July 23-26, 2000.

[7] B. M. Sadler, R. J. Kozick. T. Moore, "Bounds on bearing and symbol estimation with side information," *IEEE Trans. Signal Processing*, to appear.

[8] Y. Hua, "Fast maximum likelihood for blind identification of multiple FIR channels," *IEEE Trans. Signal Processing*, Vol. 44, No. 3, pp. 661–672, March 1996.

[9] B. M. Sadler, R. J. Kozick. T. Moore, "Bounds On SIMO and MIMO channel estimation and equalization with side information ," *Jrnl. VLSI Sig. Proc.* (invited special issue), 2001.

[10] A. Swami, G. Giannakis, S. Shamsunder, "Multichannel ARMA processes," *IEEE Trans. Signal Processing*, Vol. 42, No. 4, pp. 898–913, April 1994.

[11] S. Barbarossa and A. Scaglione, "Blind equalization using cost functions matched to the signal constellation," *Proc. 31st Asilomar Conf. Sig. Sys. Comp.*, Pacific Grove, CA, vol. 1, pp. 550–554, Nov. 1997.

[12] A. Swami, S. Barbarossa, B. M. Sadler, "Blind source separation and signal classification," *Proc. 34th Asilomar Conf. Sig. Sys. Comp.*, Pacific Grove, CA, Nov. 2000.