Further Results on Bayesian and Deterministic CRBs in the Context of Blind SIMO Channel Estimation

Samir-Mohamad Omar, Dirk Slock
Mobile Communications Department, Eurecom
BP 193, 06904 Sophia Antipolis Cdx, France
omar@eurecom.fr, slock@eurecom.fr

Oussama Bazzi
Department of Physics and Electronics
Faculty of Sciences I, Lebanese University
Hadath ,Beirut, Lebanon
obazzi@ul.edu.lb

***Abstract*—The performance of channel estimation is often assessed by deriving the proper Cramér-Rao Bound (CRB). However, in the blind case a special treatment is required due to the singularity of the Fisher Information Matrix (FIM). Usually a constraint is introduced to overcome the blind ambiguity and ensuing singularity. Hence, a constrained CRB has been derived in the literature since a long time ago. Although this constrained CRB has been proven to be a valid lower bound in the medium and high SNR regimes, it fails completely in the low SNR regime because unlike the MSE it does not saturate. Motivated by the shortcoming of the constrained CRB, we derive in this paper a modified constrained CRB (MCCRB) for deterministic blind channel estimation. The MCCRB is valid over the whole SNR regime. In the second part of the paper we address Bayesian blind channel estimation and explore the apparent discrepancy between channel unidentifiability with a non-singular FIM. We highlight that in the less familiar Bayesian case this relationship needs to be interpreted differently. The analytical formulas for the introduced bounds are validated by some Monte-Carlo simulations.**

I. Introduction

It is well known that blind channel estimation yields a channel estimate with a scalar ambiguity in the context of SIMO transmission systems. Traditionally, this ambiguity is solved by forcing a constraint. There are many constraints available in the literature, e.g. linear constraints, least-squares constraints, fixing one tap constraints. However, all these constraints assume that the true channel or part of it is provided by a genie. Nevertheless, people have used extensively these constraints to evaluate and compare blind channel estimation algorithms. Practically, the true channel is not available at all, hence one should resort to other techniques to fix this scalar ambiguity, for eg. differential modulation. On the other hand, as we have indicated previously, people have derived a performance lower bound that constitutes a reference for comparing different blind channel estimation algorithms. In the training sequence case, this lower bound is usually computed from the inverse of the FIM. However, in the blind channel estimation case, the FIM is singular due to the scalar ambiguity and consequently, it can't be inverted. Hence, to fix this ambiguity people usually resort to impose constraints as we indicated. In [1], a constrained CRB has been derived for the singular FIM case. In the same spirit, in [2] the authors prove that taking the pseudo inverse of the FIM would yield the constrained CRB that has the lowest MSE. However, the main drawback of this CRB is its incapability to track the estimation error in the low to very low SNR regime. In other words, the existing constrained CRB is no longer a lower bound in this SNR regime. The main reason for this shortcoming is that the constraint usually imposed on the blind channel estimate renders the estimation error bounded at low SNR (reflecting a bias). Hence, we notice from various simulations that we have conducted that the Normalized Mean Square Error (NMSE) levels off at low SNR and doesn't grow unboundedly as the noise variance gets smaller and smaller. This behavior is not reflected in the constrained CRB that has been derived by [1] and [2]. Motivated by the need to have a valid lower bound also at the low and very low SNR regimes for the blind channel estimation, we investigate and derive in the sequel, a modified constrained CRB that takes into consideration that the error is bounded at low SNR.

On the other hand, in [3] the Bayesian CRB in the context of cooperative OFDM was derived where the authors claimed (section III-B) that the knowledge of the prior information of the channel eliminates the ambiguity of the blind channel estimation. We will show in this paper that the prior information of the channel doesn't provide any information about the phase while it provides some information about the amplitude of the channel response scale factor. Consequently, the ambiguity is not completely removed and the singularity persists.

II. SIMO FIR Tx System Model

In (semi-)blind channel identification, a multichannel framework can be obtained from oversampling a received signal and leads to a Single Input Multiple Output (SIMO) vector channel representation. The multiple FIR channels we obtain in this representation can also be obtained from multiple received signals from an array of antennas (in the context of mobile digital communications [4]) from a combination of both. To further develop the case of oversampling, consider a linear digital modulation over a linear channel with additive noise so that the received signal $y(t)$ has the following form

$$y(t) \ = \ \sum_k h(t-kT)a(k) \ + \ v(t). \qquad (1)$$

In (1) $a(k)$ are the transmitted symbols, T is the symbol period and $h(t)$ is the channel impulse response. The channel is assumed to be FIR with length NT. If the received signal is oversampled at the rate $\frac{m}{T}$ (or if m different samples of the received signal are captured by m sensors every T seconds, or a combination of both), the discrete input-output relationship can be written as:

$$\boldsymbol{y}(k) = \sum_{i=0}^{N-1} \boldsymbol{h}(i)a(k-i) + \boldsymbol{v}(k) = \boldsymbol{H}A_N(k) + \boldsymbol{v}(k) \qquad (2)$$

where $\boldsymbol{y}(k) = [y_1^H(k)\cdots y_m^H(k)]^H, \boldsymbol{h}(i) = \left[h_1^H(i)\cdots h_m^H(i)\right]^H$, $\boldsymbol{v}(k) = [v_1^H(k)\cdots v_m^H(k)]^H$, $\boldsymbol{H} = [\boldsymbol{h}(N-1)\cdots\boldsymbol{h}(0)]$, $A_N(k) = \left[a(k-N+1)^H\cdots a(k)^H\right]^H$ and superscript ^{H} denotes Hermitian transpose. Let $\mathbf{H}(z) = \sum_{i=0}^{N-1}\boldsymbol{h}(i)z^{-i} = [\mathrm{H}_1^H(z)\cdots\mathrm{H}_m^H(z)]^H$ be the SIMO channel transfer function, and $\boldsymbol{h} = \left[\boldsymbol{h}^H(N-1)\cdots\boldsymbol{h}^H(0)\right]^H$. Consider additive independent white Gaussian circular noise $\boldsymbol{v}(k)$ with $r_{\boldsymbol{vv}}(k-i) = \mathrm{E}\,\boldsymbol{v}(k)\boldsymbol{v}(i)^H = \sigma_v^2 I_m\,\delta_{ki}$. Assume we receive M samples:

$$\boldsymbol{Y}_M(k) \;=\; \mathcal{T}_M(\boldsymbol{h})\,A_{M+N-1}(k) + \boldsymbol{V}_M(k) \qquad (3)$$

where $\boldsymbol{Y}_M(k) = [\boldsymbol{y}^H(k-M+1)\cdots\boldsymbol{y}^H(k)]^H$ and similarly for $\boldsymbol{V}_M(k)$, and $\mathcal{T}_M(\boldsymbol{h})$ is a block Toepltiz matrix with M block rows and $[\boldsymbol{H}\;\;0_{m\times(M-1)}]$ as first block row. We shall simplify the notation in (3) with $k = M-1$ to

$$\boldsymbol{Y} \;=\; \mathcal{T}(\boldsymbol{h})\,A + \boldsymbol{V} \qquad (4)$$

III. Preliminary Formulas

In [5] we have presented a complete framework that permits the derivation of various deterministic and Bayesian CRBs. Those different CRBs were classified into two categories. The first category constitutes of the CRBs that correspond to the cases where we jointly estimate the channel and the symbols, whereas the second category corresponds to the cases where we estimate the channel and marginalize the symbols. In this paper we are are going to deal with CRBs that belong to the first category. Hence, we shall reintroduce the framework presented originally in [5] to elaborate our ideas. If we denote by θ the unknown parameters to be estimated then it is given by:

$$\theta = [A^H, \boldsymbol{h}^H]^H \qquad (5)$$

The likelihood function is given by:

$$f(Y,\theta) = f(Y/\theta)f(\theta) \qquad (6)$$

Where $f(\theta)$ stands for the probability density function (pdf) of θ, $f(Y,\theta)$ stands for the joint probability density function of Y and θ and $f(Y/\theta)$ stands for the pdf of Y conditioned on θ is given or known. Once we substitute θ in (6) by its elements we get:

$$f(Y,A,\boldsymbol{h}) = f(Y/A,\boldsymbol{h})f(A)f(\boldsymbol{h}) \qquad (7)$$

Since the symbols and the channel are independent of each other we can write $f(\theta) = f(A)f(\boldsymbol{h})$. Of course on the basis of how we treat the symbols and the channel both $f(A)$ and $f(\boldsymbol{h})$ differs from one estimator to another as we shall see in the sequel. Knowing that the CRB and consequently the Fisher Information Matrix (FIM) requires the log-likelihood function, hence we apply the log function on both sides of (7) to get:

$$\ln[f(Y,A,\boldsymbol{h})] = \ln[f(Y/A,\boldsymbol{h})] + \ln[f(A)] + \ln[f(\boldsymbol{h})] \qquad (8)$$

Now let J represents the Fisher Information matrix (FIM), it is given by:

$$\begin{aligned} J_{\theta\theta} &= \mathrm{E}\left(\frac{\partial\ln[f(Y,A,\boldsymbol{h})]}{\partial\theta^*}\right)\left(\frac{\partial\ln[f(Y,A,\boldsymbol{h})]}{\partial\theta^*}\right)^H \\ &= -\,\mathrm{E}\,\frac{\partial}{\partial\theta^*}\left(\frac{\partial\ln[f(Y,A,\boldsymbol{h})]}{\partial\theta^*}\right)^H \end{aligned} \qquad (9)$$

As we shall observe later, since we are treating complex parameters we need besides $J_{\theta\theta}$ also $J_{\theta\theta^*}$ which is defined by:

$$\begin{aligned} J_{\theta\theta^*} &= \mathrm{E}\left(\frac{\partial\ln[f(Y,A,\boldsymbol{h})]}{\partial\theta^*}\right)\left(\frac{\partial\ln[f(Y,A,\boldsymbol{h})]}{\partial\theta}\right)^H \\ &= -\,\mathrm{E}\,\frac{\partial}{\partial\theta}\left(\frac{\partial\ln[f(Y,A,\boldsymbol{h})]}{\partial\theta^*}\right)^H \end{aligned} \qquad (10)$$

When $J_{\theta\theta^*} \neq 0$ we shall resort to θ_R defined below:

$$\theta_R = \begin{bmatrix} Re(\theta) \\ Im(\theta) \end{bmatrix} = \mathcal{M}\begin{bmatrix} \theta \\ \theta^* \end{bmatrix}, \mathcal{M} = \frac{1}{2}\begin{bmatrix} I & I \\ -jI & jI \end{bmatrix} \qquad (11)$$

Knowing that $J_{\theta\theta} = J^*_{\theta^*\theta^*}$ and $J_{\theta\theta^*} = J^*_{\theta^*\theta}$ then (11) yields:

$$J_{\theta_R\theta_R} = \mathcal{M}\begin{bmatrix} J_{\theta\theta} & J_{\theta\theta^*} \\ J^*_{\theta\theta^*} & J^*_{\theta\theta} \end{bmatrix}\mathcal{M}^H \qquad (12)$$

On the other side, when $J_{\theta\theta^*} = 0$ then $J_{\theta_R\theta_R}$ is determined totally by $J_{\theta\theta}$. This holds true for all the cases where we jointly estimate the channel and the symbols, as we shall notice later. Under some assumptions and regularity conditions [6] the error covariance matrix of an unbiased channel estimator $\widehat{\boldsymbol{h}}(Y)$, which is defined as:

$$C(\widehat{h}) = \mathrm{E}\left\{\left[\widehat{\boldsymbol{h}}(Y) - \boldsymbol{h}\right]\left[\widehat{\boldsymbol{h}}(Y) - \boldsymbol{h}\right]^H\right\} \qquad (13)$$

and which satisfies the following inequality:

$$C(\widehat{\boldsymbol{h}}) \geq \{J_{\theta_R\theta_R}\}^{-1} \triangleq CRB \qquad (14)$$

We usually focus on comparing the Mean Square Error, MSE = $\mathrm{tr}\left\{C(\widehat{\boldsymbol{h}})\right\}$ to the minimum error variance which is defined by $\mathrm{tr}\,\{CRB\}$ where tr stands for the trace of a matrix.

A. CRB_{det}

This is the traditional lower bound [7] for the case where both the symbols and the channel are considered as deterministic unknowns to be estimated jointly. After a little bit treatment, (9) yields:

$$J_{\theta\theta} = \frac{1}{\sigma_v^2}\begin{bmatrix} \mathcal{T}^H(\boldsymbol{h})\mathcal{T}(\boldsymbol{h}) & \mathcal{T}^H(\boldsymbol{h})\mathcal{A} \\ \mathcal{A}^H\mathcal{T}(\boldsymbol{h}) & \mathcal{A}^H\mathcal{A} \end{bmatrix} \qquad (15)$$

Moreover, we can easily show that $J_{\theta\theta^*} = 0$. Hence, by applying the Schur's complement on (15) we can extract the FIM for the channel. Normally, the CRB can be obtained by inverting the FIM however, in the blind channel estimation case the FIM is singular and consequently can't be inverted. In [2] it was shown that taking the pseudo inverse of the FIM corresponds to the constrained CRB with the lowest MSE. Hence, we can write:

$$CRB_{det} = J^{+}_{\boldsymbol{hh}} = \sigma_v^2 \left(\mathcal{A}^H P^{\perp}_{\mathcal{T}(\boldsymbol{h})}\mathcal{A}\right)^{+} \tag{16}$$

Where $+$ denotes the Moore-Penrose pseudoinverse, $P^{\perp}_{\mathcal{T}(\boldsymbol{h})} = I - P_{\mathcal{T}(\boldsymbol{h})}$ and $P_{\mathcal{T}(\boldsymbol{h})} = \mathcal{T}(\boldsymbol{h})(\mathcal{T}^H(\boldsymbol{h})\mathcal{T}(\boldsymbol{h}))^{-1}\mathcal{T}^H(\boldsymbol{h})$ is a projection matrix on $\mathcal{T}(\boldsymbol{h})$. The main shortcoming of this CRB is its incapability to comply with the deterministic blind channel estimators that normally saturate at low SNR due to the imposed constraint. In other words, this CRB can't be considered a valid lower bound in this SNR regime. This encourages us to derive a modified CRB that tries to deal with the effect of the constraint directly. Although the methodology we shall implement is considered to be rigorous, nonetheless it provides a concrete formula for an upper lower bound which is valid at the whole SNR range.

B. Modified Constrained CRB_{det} ($MCCRB_{det}$)

To commence, we show below the relation between the blind channel estimate $\widehat{\boldsymbol{h}}$, the true channel and the channel estimation error $\tilde{\boldsymbol{h}}$:

$$\widehat{\boldsymbol{h}} = \beta(\boldsymbol{h} + \tilde{\boldsymbol{h}}) \tag{17}$$

Where β denotes the scalar ambiguity. As mentioned earlier, we can get rid of this scalar ambiguity by using one of the constraints available in the literature. In our case, we choose the least squares constraint. However, it should be noted that the formula for the modified constrained CRB that we are going to derive is affected well by the choice of the constraint.

$$\min_{\alpha} \ ||\boldsymbol{h} - \alpha\widehat{\boldsymbol{h}}||^2 \tag{18}$$

which yields $\alpha = \frac{\widehat{\boldsymbol{h}}^H \boldsymbol{h}}{||\widehat{\boldsymbol{h}}||^2}$. Hence, we have:

$$\begin{aligned} \hat{\hat{\boldsymbol{h}}} &= \alpha\widehat{\boldsymbol{h}} \\ &= \alpha\beta(\boldsymbol{h} + \tilde{\tilde{\boldsymbol{h}}}) \end{aligned} \tag{19}$$

On the other hand, we know that $\tilde{\tilde{\boldsymbol{h}}}$ is orthogonal to $\hat{\hat{\boldsymbol{h}}}$ (well known property of the least squares estimate). Therefore, we can write: $\tilde{\tilde{\boldsymbol{h}}} = P^{\perp}_{\hat{\hat{\boldsymbol{h}}}}\boldsymbol{h}$ and consequently,

$$C_{\tilde{\tilde{\boldsymbol{h}}}} = \mathrm{E}\ \tilde{\tilde{\boldsymbol{h}}}\tilde{\tilde{\boldsymbol{h}}}^H = \mathrm{E}\ P^{\perp}_{\hat{\hat{\boldsymbol{h}}}}\boldsymbol{h}\boldsymbol{h}^H P^{\perp}_{\hat{\hat{\boldsymbol{h}}}}. \tag{20}$$

Where E stands for the Expectation operator. Since we are usually interested in computing NMSE for different algorithms, then we are only interested in the diagonal elements of $C_{\tilde{\tilde{\boldsymbol{h}}}}$. Hence, we apply the trace operator as follows:

$$tr(C_{\tilde{\tilde{\boldsymbol{h}}}}) = ||\boldsymbol{h}||^2 - \mathrm{E}\ \frac{\boldsymbol{h}^H\hat{\hat{\boldsymbol{h}}}\hat{\hat{\boldsymbol{h}}}^H\boldsymbol{h}}{||\hat{\hat{\boldsymbol{h}}}||^2} \tag{21}$$

Since it is difficult to carry out the expectation operator in (21), we shall resort to Jensen's inequality which permits us to split the expectation operator between the numerator and the denominator as follows:

$$\mathrm{E}\ \frac{\hat{\hat{\boldsymbol{h}}}}{||\hat{\hat{\boldsymbol{h}}}||}\frac{\hat{\hat{\boldsymbol{h}}}^H}{||\hat{\hat{\boldsymbol{h}}}||} \geq \frac{\mathrm{E}\ \hat{\hat{\boldsymbol{h}}}\hat{\hat{\boldsymbol{h}}}^H}{\mathrm{E}\ ||\hat{\hat{\boldsymbol{h}}}||||\hat{\hat{\boldsymbol{h}}}||} \tag{22}$$

As we shall observe in the sequel, by using Jensen's inequality we are no longer computing the exact channel error covariance matrix but rather an upper bound for it. Now, substituting (19) in (22), we can write:

$$\mathrm{E}\ \frac{\hat{\hat{\boldsymbol{h}}}}{||\hat{\hat{\boldsymbol{h}}}||}\frac{\hat{\hat{\boldsymbol{h}}}^H}{||\hat{\hat{\boldsymbol{h}}}||} \geq \frac{\mathrm{E}\ (\boldsymbol{h}\boldsymbol{h}^H + \boldsymbol{h}\tilde{\tilde{\boldsymbol{h}}}^H + \tilde{\tilde{\boldsymbol{h}}}^H\boldsymbol{h} + \tilde{\tilde{\boldsymbol{h}}}\tilde{\tilde{\boldsymbol{h}}}^H)}{\mathrm{tr}\left\{\mathrm{E}\ (\boldsymbol{h}\boldsymbol{h}^H + \boldsymbol{h}\tilde{\tilde{\boldsymbol{h}}}^H + \tilde{\tilde{\boldsymbol{h}}}^H\boldsymbol{h} + \tilde{\tilde{\boldsymbol{h}}}\tilde{\tilde{\boldsymbol{h}}}^H)\right\}} \tag{23}$$

It should be noted that the $|\alpha\beta|^2$ terms from the numerator and the denominator cancel each other. Moreover, knowing that $\mathrm{E}\ \boldsymbol{h}\tilde{\tilde{\boldsymbol{h}}}^H = \mathrm{E}\ \tilde{\tilde{\boldsymbol{h}}}\boldsymbol{h}^H = 0$ since $\mathrm{E}\ \tilde{\tilde{\boldsymbol{h}}} = 0$ (error with zero mean) and substituting $\mathrm{E}\ \tilde{\tilde{\boldsymbol{h}}}\tilde{\tilde{\boldsymbol{h}}}$ by $C_{\tilde{\tilde{\boldsymbol{h}}}}$ in (24) we get:

$$\mathrm{E}\ \frac{\hat{\hat{\boldsymbol{h}}}}{||\hat{\hat{\boldsymbol{h}}}||}\frac{\hat{\hat{\boldsymbol{h}}}^H}{||\hat{\hat{\boldsymbol{h}}}||} \geq \frac{(\boldsymbol{h}\boldsymbol{h}^H + C_{\tilde{\tilde{\boldsymbol{h}}}})}{\mathrm{tr}\,((\boldsymbol{h}\boldsymbol{h}^H + C_{\tilde{\tilde{\boldsymbol{h}}}}))} \tag{24}$$

Thus, making use of (24) in (21) we get:

$$\begin{aligned} \mathrm{tr}\,(C_{\tilde{\tilde{\boldsymbol{h}}}}) &\leq ||\boldsymbol{h}||^2 - \frac{\boldsymbol{h}^H(\boldsymbol{h}\boldsymbol{h}^H + C_{\tilde{\tilde{\boldsymbol{h}}}})\boldsymbol{h}}{tr(\boldsymbol{h}\boldsymbol{h}^H + C_{\tilde{\tilde{\boldsymbol{h}}}})} \\ &\leq ||\boldsymbol{h}||^2 - \frac{||\boldsymbol{h}||^4}{||\boldsymbol{h}||^2 + tr(C_{\tilde{\tilde{\boldsymbol{h}}}})} \\ &\leq \frac{||\boldsymbol{h}||^2 tr(C_{\tilde{\tilde{\boldsymbol{h}}}})}{||\boldsymbol{h}||^2 + tr(C_{\tilde{\tilde{\boldsymbol{h}}}})} \end{aligned} \tag{25}$$

where we have used in the above derivation the fact that $\boldsymbol{h}^H C_{\tilde{\tilde{\boldsymbol{h}}}}\boldsymbol{h} = 0$. This is true since $C_{\tilde{\tilde{\boldsymbol{h}}}}$ admits $\boldsymbol{h}$ as a singular vector [8]. Moreover, we have also substituted $\mathrm{tr}\,(\boldsymbol{h}\boldsymbol{h}^H)$ by $||\boldsymbol{h}||^2$. Now we can define the trace of the modified error covariance matrix as follows:

$$\mathrm{tr}\,(C^{mod}_{\tilde{\tilde{\boldsymbol{h}}}}) = \frac{||\boldsymbol{h}||^2 tr(C_{\tilde{\tilde{\boldsymbol{h}}}})}{||\boldsymbol{h}||^2 + tr(C_{\tilde{\tilde{\boldsymbol{h}}}})} \tag{26}$$

A close look at (26) suggests that at very low SNR ($tr(C_{\tilde{\tilde{\boldsymbol{h}}}}) >> ||\boldsymbol{h}||^2$) we can neglect $||\boldsymbol{h}||^2$ in the denominator. This yields a simple formula for the modified Mean Square Error (MSE^{mod}) as follows:

$$MSE^{mod} = \mathrm{tr}\,(C^{mod}_{\tilde{\tilde{\boldsymbol{h}}}}) = ||\boldsymbol{h}||^2 \tag{27}$$

It is obvious from (27) that the channel error is bounded and can't exceed the norm of the channel even when the noise variance grows infinitely. However, at high SNR ($tr(C_{\tilde{\boldsymbol{h}}}) << ||\boldsymbol{h}||^2$) we can neglect $tr(C_{\tilde{\boldsymbol{h}}})$ in the denominator. By doing so, we get the same MSE giving by the traditional formula for the error covariance matrix:

$$MSE^{mod} = \text{tr}\,(C^{mod}_{\tilde{\boldsymbol{h}}}) = \text{tr}\,(C_{\tilde{\boldsymbol{h}}}) = MSE \qquad (28)$$

Consequently, we deduce that our proposed formula for the constrained channel estimation error has the potential to keep track of the actual channel estimation error at low SNR, by leveling off as expected, and at high SNR, by providing the same traditional analytical formula.

Well, substituting now $C_{\tilde{\boldsymbol{h}}}$ by CRB_{det} in (26), we will get the formula for the modified constrained CRB:

$$\text{tr}\,(MCCRB_{det}) = \frac{||\boldsymbol{h}||^2 tr(CRB_{det})}{||\boldsymbol{h}||^2 + tr(CRB_{det})} \qquad (29)$$

The same discussion illustrated in the case of the modified error covariance matrix is still applicable here in the case of the modified CRB. As a result, we can state that our modified CRB levels off at low SNR whereas it is congruent to the traditional CRB at high SNR. Furthermore, we know that $\text{tr}\,(C_{\tilde{\boldsymbol{h}}}) \geq \text{tr}\,(CRB_{det})$, hence substituting $C_{\tilde{\boldsymbol{h}}}$ by CRB_{det} in the right hand side of the second line in (25) yields:

$$\begin{aligned} ||\boldsymbol{h}||^2 - \frac{||\boldsymbol{h}||^4}{||\boldsymbol{h}||^2 + tr(C_{\tilde{\boldsymbol{h}}})} &\geq ||\boldsymbol{h}||^2 - \frac{||\boldsymbol{h}||^4}{||\boldsymbol{h}||^2 + \text{tr}\,(CRB_{det})} \\ \frac{||\boldsymbol{h}||^2\,\text{tr}\,(C_{\tilde{\boldsymbol{h}}})}{||\boldsymbol{h}||^2 + tr(C_{\tilde{\boldsymbol{h}}})} &\geq \frac{||\boldsymbol{h}||^2 tr(CRB_{det})}{||\boldsymbol{h}||^2 + \text{tr}\,(CRB_{det})} \\ \text{tr}\,(C^{mod}_{\tilde{\boldsymbol{h}}}) &\geq \text{tr}\,(MCCRB_{det}) \end{aligned} \qquad (30)$$

Therefore, this modified constrained CRB still constitutes a lower bound for the modified error covariance matrix.

IV. Bayesian Cramer Rao bound with Deterministic symbols (BCRB$_{det,joint}$)

In this section we treat the Bayesian blind case where the channel is considered as random with Gaussian distribution and the symbols are considered as deterministic unknowns to be jointly estimated with the channel. Lets denote by $\rho e^{j\phi}$ to the a scalar ambiguity, where ρ stands for the amplitude and ϕ stands for the phase. From the pdf of the channel ($\boldsymbol{h} \sim \mathcal{N}(0, C^o_h)$), we can easily notice that the prior Fisher information Matrix (FIM) is given by $C^{o^{-1}}_h$. Usually the total FIM is the sum of the prior FIM and the FIM of the data. The latter is singular while the the former has usually a full rank. Hence, the total FIM has a full rank. At the first glance this will lead to the same conclusion that was drawn in [3] namely, the prior information eliminates the blind channel ambiguity. However, a closer look at the problem will prove that this result is inaccurate at all.

First of all the notion of parameter identifiability needs to be considered with care in the case of Bayesian parameters. For deterministic parameters, local identifiability means that no continuously valued unknowns remain in the parameters and corresponds to the Fisher Information Matrix (FIM) being non-singular and hence its inverse, the Cramer-Rao bound (CRB) being finite. Local unidentifiability on the other hand means that the CRB and hence the MSE (for which the CRB is a lower bound) is infinite. Local identifiability with global unidentifiability means that there are some discrete valued unknowns remaining in the parameter estimates. In Bayesian blind channel estimation on the other hand, the FIM coming from the prior channel distribution (expressing the PDP) is non-singular and leads to non-singularity of the overall FIM. On the other hand, it is clear that the channel response remains only identifiable up to a global phase factor (unit magnitude scalar), which is a continuously varying unknown. On the other hand, the channel estimation MSE remains finite. Indeed, any wrong hypothesis on the phase factor leads to finite MSE since the channel itself has finite power (due to the prior distribution with finite PDP). So in the Bayesian case, the concepts of identifiability need to be considered with care.

Now, the question is how to show that the prior FIM is rank deficient? In order to answer this question and show that the prior FIM is singular we should reparametrize the problem between our hands. Moreover, we should also resort to splitting the complex channel parameters into their real and imaginary parts. When we accomplish the two previous steps and derive the FIM for the new reparametrized prior we will find it singular for sure. To commence with this task, lets take the first tap of the channel as a common factor we get $\boldsymbol{h} = \rho e^{j\phi}\boldsymbol{h}^{'}$ where $\boldsymbol{h}^{'} = [1 \;\bar{\boldsymbol{h}}^H]^H$. Denote by $\boldsymbol{h}_R = [\bar{\boldsymbol{h}}^{rT}, \bar{\boldsymbol{h}}^{sT}, \rho, \phi]$ the set of parameters to be estimated where $\bar{\boldsymbol{h}}^{rT}$ and $\bar{\boldsymbol{h}}^{sT}$ denotes respectively the real and the imaginary parts of $\bar{\boldsymbol{h}}$. Due to the lack of space we will not go into the detailed derivation nevertheless we will show below the resulting prior FIM (2mN x 2mN)which is given by:

$$J^{prior}_{\boldsymbol{h}_R\boldsymbol{h}_R} = \begin{bmatrix} C^o_h(1,1)\bar{C}^{o^{-1}}_h & 0 & 0 & 0 \\ 0 & C^o_h(1,1)\bar{C}^{o^{-1}}_h & 0 & 0 \\ 0 & 0 & C^{o^{-1}}_h(1,1) & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \qquad (31)$$

Where $C^o_h(1,1)$ denotes the element that lies in the first row and first column of C^o_h and $\bar{C}^{o^{-1}}_h$ can be obtained from $C^{o^{-1}}_h$ by omitting the first row and the first column. It is evident now that the prior FIM admits one singularity that corresponds to the phase and it provides only the variance of the ambiguous amplitude $C^o_h(1,1)$ and not the amplitude itself. Hence, this information is considered limited and incomplete. Now to pursue the derivation of the BCRB we should play the same game with the FIM of the data. Denote by $A^{'} = \rho e^{j\phi}A$ then (4) can be written as: $Y = \mathcal{T}(\boldsymbol{h}^{'})A^{'} + \boldsymbol{V}$ or in the following form: Y = $\mathcal{A}^{'}\boldsymbol{h}^{'} + \boldsymbol{V}$. We shall work out the FIM_{data} first in its complex form hence we take $\theta = [\bar{\boldsymbol{h}}^H, \rho e^{j\phi}, A^{'H}]$. Using (9) we get:

$$J^{data}_{\theta\theta} = \text{E}_{\boldsymbol{h}}\; \frac{1}{\sigma^2_v} \begin{bmatrix} \bar{\mathcal{A}}^{'H}\bar{\mathcal{A}}^{'} & 0 & \bar{\mathcal{A}}^{'H}\mathcal{T}(\boldsymbol{h}^{'}) \\ 0 & 0 & 0 \\ \mathcal{T}^H(\boldsymbol{h}^{'})\bar{\mathcal{A}}^{'} & 0 & \mathcal{T}^H(\boldsymbol{h}^{'})\mathcal{T}(\boldsymbol{h}^{'}) \end{bmatrix} \qquad (32)$$

Where $\bar{\mathcal{A}}'$ is obtained from $\mathcal{A}'$ by omitting the first column. On the other hand, we can prove after a little bit manipulation that $J_{\theta\theta^*} = 0$. Now in order to be able to compute the total FIM of the channel which is obtained by adding both FIM_{data} and FIM_{prior} of the channel. We should transform the former to make it corresponds to $\theta_R = [\bar{\boldsymbol{h}}^{rT}, \bar{\boldsymbol{h}}^{sT}, \rho, \phi, A'^{rT}, A'^{sT}]$. We shall accomplish this task in two steps. In the first step we use the Jacob matrix $\mathcal{M}$ in ([5], eqn. 12) to get $J_{\theta'_R\theta'_R}$ from (32) as in ([5], eqn. 13) where $\theta'_R = [\bar{\boldsymbol{h}}^{rT}, \rho cos\phi, A'^{rT}, \bar{\boldsymbol{h}}^{sT}, \rho sin\phi, A'^{sT}]$. However, in the second step we use another Jacob matrix ($\mathcal{M}'$) to compute $J_{\theta_R\theta_R}$ which is given by: $J_{\theta_R\theta_R} = \mathcal{M}' J_{\theta'_R\theta'_R} \mathcal{M}'^H$ where $\mathcal{M}'$ is given by:

$$\begin{bmatrix} I_{mN-1} & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & I_{mN-1} & 0 & 0 \\ 0 & cos\phi & 0 & 0 & -\rho sin\phi & 0 \\ 0 & sin\phi & 0 & 0 & \rho cos\phi & 0 \\ 0 & 0 & I_{M+N-1} & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & I_{M+N-1} \end{bmatrix} \quad (33)$$

Now using Schur's complement we can readily extract $J^{data}_{\boldsymbol{h}_R\boldsymbol{h}_R}$ from $J_{\theta_R\theta_R}$. Taking a close look at $J_{\boldsymbol{h}_R\boldsymbol{h}_R}$ we realize that it admits two singularities, one corresponds to the amplitude and the other corresponds to the phase. However, the total FIM which is the sum of the prior and the data FIMs is given by:

$$J_{\boldsymbol{h}_R\boldsymbol{h}_R} = \mathrm{E}_{\boldsymbol{h}}\; J^{data}_{\boldsymbol{h}_R\boldsymbol{h}_R} + J^{prior}_{\boldsymbol{h}_R\boldsymbol{h}_R} \quad (34)$$

Checking $J_{\boldsymbol{h}_R\boldsymbol{h}_R}$ closely, one can show that it admits one singularity that corresponds to the phase. This is due to the fact that the prior FIM ameliorates only the singularity that corresponds to the amplitude which results from the FIM of the data. Therefore, the prior FIM only contributes to fix one singularity while it has no means to deal with the other. As a consequence, the resulting BCRB which is defined as the inverse of the total FIM is still singular and needs an additional constraint to fix the phase ambiguity. If we consider for instance that the phase and/or the amplitude of the first channel tap is given then the BCRB is obtained by the Moore-Penrose pseudo inverse of $J_{\boldsymbol{h}_R\boldsymbol{h}_R}$.

V. Simulations

In this section we try to verify the analytical formulas we have derived throughout this paper by means of Monte Carlo simulations. Since we are dealing with a deterministic channel case, we used the channel in ([9], table II) which is composed of four taps and corresponds to having 4 antennas at the receiver.

However, in each Monte Carlo simulation we generate different realizations of symbols and white Gaussian noise. The symbols are drawn from a 8PSK constellation. The performance of the different channel estimators is evaluated by means of the Normalized MSE (NMSE) vs. SNR. The SNR is defined as: SNR $= \frac{||\mathcal{T}(h)A||^2}{mM\,\sigma_v^2}$ while the NMSE is defined as $\frac{||\boldsymbol{h}-\hat{\hat{\boldsymbol{h}}}||^2}{||\boldsymbol{h}||^2}$ where $\hat{\hat{\boldsymbol{h}}} = \frac{\hat{\boldsymbol{h}}^H\boldsymbol{h}}{||\hat{\boldsymbol{h}}||^2}\boldsymbol{h}$ is the channel estimate adjusted by the least squares constraint. We can notice from fig. 1 how the analytical performance for Deterministic Maximum Likelihood (DML) levels off in a consistent way like the simulated DML. Moreover, we can notice also that our modified constrained deterministic CRB (MCCRB$_{det}$) remains the lower bound for the modified analytical performance of DML after leveling off. It should be noted here that we use the analytical performance of DML that has been derived in [8]. It is worth noting that the simulated DML has been initialized by Subchannel Response Matching (SRM) [9].

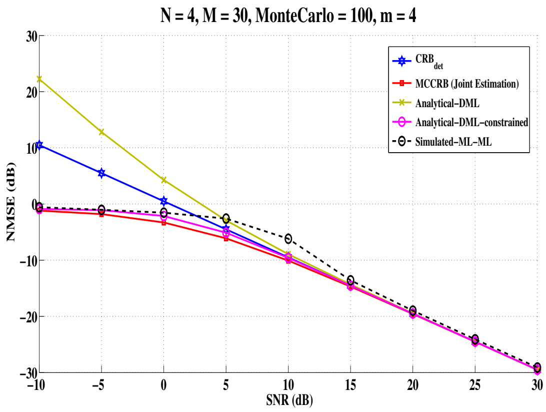

Fig. 1. Least squares constrained ML performance.

VI. Acknowledgment

EURECOM's research is partially supported by its industrial members: BMW Group Research Technology, Bouygues Telecom, Cisco, Hitachi, ORANGE, SFR, Sharp, STMicroelectronics, Swisscom, Thales. The research reported herein was also partially supported by the French ANR project SESAME and by the EU FET project CROWN and Strep WHERE2.

VII. Conclusion

We have derived in this paper an analytical formula for both the least squares constrained blind channel estimation error and its corresponding constrained CRB. In the low SNR regimes, our formulas are totally capable of tracking the error resulting in the simulations by providing an upper bound for this error, whereas the traditional analytical performance formulas and constrained CRB fail to accomplish this task. However, in the high SNR regimes our proposed formulas matches the traditional one, hence it provides the exact channel estimation error and its corresponding lower bound rather than an upper bound for both. On the other hand, we have also derived in this paper a reparametrized BCRB in the context of blind channel estimation, showing the incapability of the prior information to ameliorate the singularity due to the phase ambiguity. Moreover, we have shown that it offers only a limited information concerning the amplitude of the ambiguous scalar.

REFERENCES

[1] P. Stoica and B. Ng, "On the Cramér–Rao Bound Under Parametric Constraints," *IEEE Signal Processing Letters*, vol. 5, no. 7, pp. 177–179, July 1998.

[2] E. de Carvalho, J. Cioffi, and D. Slock, "Cramér–Rao Bounds for Blind Multichannel Estimation," in *Proc. GLOBECOM 2000 Conf.*, San Francisco, USA, December 2000.

[3] W. Wei, J. Xin, C. Yueming, and W. X. ;, "Bayesian Cramer-rao Bound for channel estimation in cooperative OFDM," in *IEEE WCSP*, Suzhou, China, Oct. 2010.

[4] D. Slock, "Blind Fractionally-Spaced Equalization, Perfect-Reconstruction Filter Banks and Multichannel Linear Prediction," in *Proc. ICASSP 94 Conf.*, Adelaide, Australia, April 1994.

[5] S.-M. Omar, D. Slock, and O. Bazzi, "Bayesian and Deterministic CRBs for Semi-Blind Channel Estimation in SIMO Single Carrier Cyclic Prefix Systems," in *Proc. IEEE PIMRC*, Toronto, Canada, Sep. 2011.

[6] E. Weinstein and A. Weiss, "A general class of lower bounds in parameter estimation," *IEEE Trans. Inform. Theory*, vol. 34, no. 34, p. 338342, Mar. 1988.

[7] E. de Carvalho and D. Slock, "Cramér-Rao Bounds for Semi-blind, Blind and Training Sequence based Channel Estimation," in *Proc. SPAWC 97 Conf.*, Paris, France, April 1997.

[8] ——, "Asymptotic Performance of ML Methods for Semi-Blind Channel Estimation," in *Proc. Asilomar Conference on Signals, Systems & Computers*, Pacific Grove, CA, Nov. 1997.

[9] G. Xu, H. Liu, L. Tong, and T. Kailath, "A Least Squares Approach to Blind Channel Identification," *IEEE Transactions on Signal Processing*, vol. 43, no. 12, pp. 2982–2993, Dec. 1995.