CRAMER-RAO BOUNDS FOR SEMI-BLIND, BLIND AND TRAINING SEQUENCE BASED CHANNEL ESTIMATION†

Elisabeth De Carvalho *Dirk T.M. Slock*
Institut EURECOM
B.P. 193, 06904 Sophia Antipolis Cedex, FRANCE
{carvalho,slock}@eurecom.fr

ABSTRACT

Two channel estimation techniques are often opposed: training sequence based estimation in which a sequence of symbols known by the receiver is used and blind equalization in which the channel and/or the symbols are determined from the received signal only. The purpose of semi-blind techniques would be to adapt blind techniques in order to profit from the existence of a training sequence. A first approach to assess the performance of semi-blind methods is proposed. We study Cramer-Rao Bounds for blind, semi-blind and training-sequence based channel estimates for a deterministic as well as a Gaussian symbol model, and compare them theoretically as well as in numerical evaluations. Furthermore, an example of comparison between the corresponding Maximum Likelihood estimation methods is given.

1. INTRODUCTION

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

$$\mathrm{Y}_M(k)\;=\;\mathcal{T}_M(\mathrm{H}_N)\,A_{M+N-1}(k)+\mathrm{V}_M(k) \qquad (1)$$

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

$$\mathrm{Y}\;=\;\mathcal{T}A+\mathrm{V}\,. \qquad (2)$$

We assume that $mM>M+N-1$ in which case the channel convolution matrix $\mathcal{T}$ has more rows than columns. If the $\mathrm{H}_i(z)$, $i=1,\ldots,m$ have no zeros in common, then $\mathcal{T}$ has full column rank (which we will henceforth assume).

Most of the actual mobile communication standards include a training sequence used to estimate the channel, or simply some known symbols used for synchronization. Blind methods are based on the received signal only: the purpose of semi-blind methods is to incorporate this a priori knowlegde of a training sequence, which appears more relevant and robust than purely blind equalization. A performance comparison between blind, semi-blind and training sequence techniques is proposed in term of Cramer-Rao Bounds (CRB). Two models are discussed: the deterministic model in which both the channel and the symbols are considered as deterministic, and the stochastic model in which the symbols are assumed Gaussian, which will prove more robust than the deterministic version. We consider here σ_v^2 known to simplify the presentation.

† The work of Elisabeth de Carvalho was supported by Laboratoires d'Electronique PHILIPS under contract Cifre 297/95. The work of Dirk Slock was supported in part by Laboratoires d'Electronique PHILIPS under contract LEP 95FAR008.

2. REAL AND COMPLEX CRB

Let θ be a vector of complex parameters, and $\theta_R=\left[\mathrm{Re}(\theta)^H\ \ \mathrm{Im}(\theta)^H\right]^H$, the associated real parameters. The Fisher Information Matrix (FIM) associated with θ_R is:

$$J_R(\theta_R)=E_{Y/\theta_R}\left(\frac{\partial\ln f(Y/\theta_R)}{\partial\theta_R}\right)\left(\frac{\partial\ln f(Y/\theta_R)}{\partial\theta_R}\right)^T \qquad (3)$$

where $f(Y/\theta_R)$ is the conditional probability density function of the observations Y. Let $\hat{\theta}_R$ be an unbiased estimate, $\tilde{\theta}_R=\theta_R-\hat{\theta}_R$ the estimation error and $C_{\tilde{\theta}_R}=E\tilde{\theta}_R\tilde{\theta}_R^H$ the error covariance matrix. $J_R^{-1}(\theta_R)$ is the Cramer-Rao Bound:

$$C_{\tilde{\theta}_R}\geq CRB=J_R^{-1}(\theta_R) \qquad (4)$$

As we work with complex quantities, it may be easier to consider complex derivation defined as $\frac{\partial}{\partial\theta}=\frac{1}{2}\left(\frac{\partial}{\partial\alpha}-j\frac{\partial}{\partial\beta}\right)$ where $\theta=\alpha+j\beta$. Let $J_{\varphi\psi}$ be defined as:

$$J_{\varphi\psi}=E_{Y/\theta}\left(\frac{\partial\ln f(Y/\theta)}{\partial\varphi^*}\right)\left(\frac{\partial\ln f(Y/\theta)}{\partial\psi^*}\right)^H \qquad (5)$$

and we will consider $J_{\theta\theta}$ and $J_{\theta\theta^*}$. We get:

$$J_R(\theta_R)=2\begin{bmatrix}\mathrm{Re}(J_{\theta\theta}) & -\mathrm{Im}(J_{\theta\theta})\\ \mathrm{Im}(J_{\theta\theta}) & \mathrm{Re}(J_{\theta\theta})\end{bmatrix}+2\begin{bmatrix}\mathrm{Re}(J_{\theta\theta^*}) & -\mathrm{Im}(J_{\theta\theta^*})\\ \mathrm{Im}(J_{\theta\theta^*}) & \mathrm{Re}(J_{\theta\theta^*})\end{bmatrix} \qquad (6)$$

When $J_{\theta\theta^*}=0$, the real FIM is completely determined by $J_{\theta\theta}$. In that case, $J_{\theta\theta}$ can be considered as a complex FIM, and $C_{\tilde{\theta}}=E\tilde{\theta}\tilde{\theta}^H\geq J_{\theta\theta}^{-1}$, the complex CRB. If $J_{\theta\theta^*}\neq 0$, $J_{\theta\theta}^{-1}$ is also a bound on $C_{\tilde{\theta}}$, but not as tight as the CRB.

3. DETERMINISTIC MODEL

For simplicity reasons, as $J_{\theta\theta^*}=0$, we will consider the complex CRB. The (complex) parameter vector θ is:

0-7803-3944-4/97/$10.00

$\theta = \left[\begin{array}{cc} A_u^H & h^H \end{array}\right]^H$. A_u contains the unknown input symbols to be estimated (all the symbols (blind), part (semi-blind) or none (training sequence)), and h the channel coefficients. The complex probability density function is:

$$f(\mathrm{Y}/\theta) = \frac{1}{(\pi\sigma_v^2)^{Mm}} \exp\left(-\frac{1}{\sigma_v^2}\|\mathrm{Y} - \mathrm{Y}_S\|^2\right), \quad (7)$$

$\mathrm{Y}_S = \mathcal{T}A$ is the signal part of Y. The complex FIM is:

$$J(\theta) = \frac{1}{\sigma_v^2}\left(\frac{\partial \mathrm{Y}_S^H}{\partial \theta^*}\right)\left(\frac{\partial \mathrm{Y}_S^H}{\partial \theta^*}\right)^H = \frac{1}{\sigma_v^2}\left[\begin{array}{c} Q \\ R \end{array}\right]\left[\begin{array}{c} Q \\ R \end{array}\right]^H \quad (8)$$

where: $Q = \frac{\partial \mathrm{Y}_S^H}{\partial A_u^*}$ and $R = \frac{\partial \mathrm{Y}_S^H}{\partial h^*}$. $C_{\hat{h}}(\theta)$ is the error covariance matrix of the channel estimate corresponding to the joint estimation of the symbols and the channel:

$$C_{\hat{h}}(\theta) \geq \sigma_v^2\left(RP_{Q^H}^{\perp}R^H\right)^{-1} \quad (9)$$

$P_{Q^H}^{\perp} = I - P_{Q^H}$ and $P_{Q^H} = Q^H(QQ^H)^{-1}Q$. Next we elaborate this result for our specific cases.

3.1. Training Sequence based Channel Estimation

$$CRB_{TS} = \sigma_v^2\left[\mathcal{A}^H\mathcal{A}\right]^{-1} \quad (10)$$

$\mathcal{A}$ is a structured matrix filled with the input symbols: $\mathcal{T}A = \mathcal{A}h$. The CRB depends on the value of the symbols present in the training sequence. It is minimized for a given training sequence energy when $\mathcal{A}^H\mathcal{A}$ is a multiple of identity. The CRB is then equal to $\frac{\sigma_v^2}{M\sigma_a^2}I_{mN}$. This indicates a condition on the choice of a training sequence for good channel estimation. When the number of known uncorrelated symbols becomes large, it is interesting to notice, by the law of large numbers, that $\frac{1}{M}\mathcal{A}^H\mathcal{A}$ tends to $\sigma_a^2 I_{mN}$, which corresponds to the minimal value of the CRB already found. In this case, the previous optimization condition is then verified.

3.2. Blind Channel Estimation

$J(\theta)$ is singular: $\left[\begin{array}{cc} -A^H & h^H \end{array}\right] J(\theta) = 0$. This singularity is due to the fact that blind estimation is only possible up to a scale factor: the channel cannot be estimated, and the CRB is then $+\infty$. However, the Moore-Penrose pseudo-inverse of $J(\theta)$, $J^+(\theta)$, can be considered: it may be interpreted as the CRB knowing the component of θ along the null space of $J(\theta)$. Appendix 1 can be adapted to prove it.

$\mathcal{A}^H P_{\mathcal{T}}^{\perp}\mathcal{A}$ is also singular: h is its singular vector. Its Moore-Penrose pseudo-inverse (which is not the corresponding submatrix of $J^+(\theta)$) can again be interpreted as the CRB knowing the component of h along the null space of $\mathcal{A}^H P_{\mathcal{T}}^{\perp}\mathcal{A}$, which is equivalent to knowing the phase and module of the channel: see Appendix 1. More precisely, let $\hat{h}$ be an unbiased estimate (apart from an arbitrary module and phase) of h obtained by blind estimation. The previous pseudo-inverse is the CRB for $\hat{\hat{h}} = \alpha e^{j\varphi}\hat{h}$ where α and φ are adjusted so that $h^H\hat{\hat{h}} = h^H h$ ($= \mathcal{V}_2^{oH}h$ of Appendix 1).

We choose then as CRB for blind estimation:

$$CRB_B = \sigma_v^2\left[\mathcal{A}^H P_{\mathcal{T}}^{\perp}\mathcal{A}\right]^+ \quad (11)$$

The real CRB has two singular vectors:

$$h_R^{(1)} = h_R \text{ and } h_R^{(2)} \left[\begin{array}{cc} -\mathrm{Im}(h)^H & \mathrm{Re}(h)^H \end{array}\right]^H \quad (12)$$

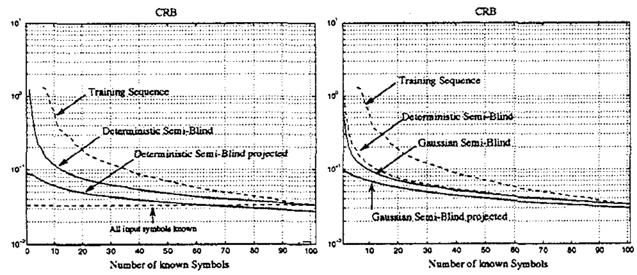

Figure 1. Comparison between semi-blind, training sequence and blind channel estimation

Appendix 1 holds also for real parameters. The first singular vector corresponds to the ambiguity in the module and the second one to the ambiguity in the phase. The pseudo-inverse corresponds to $\hat{\hat{h}}_R$ (defined as previously) where the module is determined using the equation: $h_R^T\hat{\hat{h}} = h_R^T h_R$, and the phase using $h_R^{(2)T}\hat{\hat{h}}_R = 0$. $h_R^T h_R$ and the 0 correspond to $\mathcal{V}_2^{oT}h_R$ of Appendix 1, $\mathcal{V}_2^o$ is now 2-dimensional.

3.3. Semi-Blind Channel Estimation

Part of the input symbols is known. The vector of input symbols can be written as: $A = \mathcal{P}\left[\begin{array}{c} A_k \\ A_u \end{array}\right]$ where A_k are the known symbols, A_u the unknown symbols and $\mathcal{P}$ is some permutation matrix. $\mathrm{Y}_S = \mathcal{T}A = \mathcal{T}_k A_k + \mathcal{T}_u A_u$. Then:

$$CRB_{SB} = \sigma_v^2\left[\mathcal{A}^H P_{\mathcal{T}_u}^{\perp}\mathcal{A}\right]^{-1} \quad (13)$$

3.4. Comparisons

We compare here the different estimation modes through their CRB. These comparisons are illustrated by curves showing the trace of the CRBs w.r.t. the number of known symbols M_k in the input burst. The SNR, defined as $\frac{\sigma_a^2\|h\|^2}{m\sigma_v^2}$ (average SNR per subchannel), is 10dB, M=100. The channel is the following:

$$\mathrm{H}_N = \left[\begin{array}{ccc} 1.1650+j0.2641 & 0.0751-j1.4462 & -0.6965+j1.2460 \\ 0.6268+j0.8717 & 0.3516-j0.7012 & 1.6961-j0.6390 \end{array}\right] \quad (14)$$

In fig. 1 the semi-blind CRB is shown. When very few symbols are known, performances are low, which is probably due to the difficulty of estimating the scale factor of the channel with few known symbols. However, we observe that after the introduction of very few more known symbols, performance increases dramatically. After this threshold of improvement, it is necessary to introduce a large number of known symbols to get a significant improvement. These numerical evaluations indicate that semi-blind techniques could improve performance drastically w.r.t. blind techniques with only a few known symbols.

3.4.1. Semi-Blind vs Training-Sequence

In fig. 1 (left), the training sequence used in the training sequence based channel estimation is the same (same symbols and same length) as the symbols known in the semi-blind mode. It can be noticed that semi-blind estimation represents an important gain w.r.t. training sequence mode, especially when few symbols are known. Besides, for the same performance, you do not need as many known symbols for semi-blind estimation compared to training sequence based estimation. The CRB when all the $M+N-1$ input symbols are known is given as reference.

3.4.2. Blind vs Semi-Blind

In this comparison, the input bursts are the same (same symbols and same length) but part of the symbols in the semi-blind mode is known. The CRB associated with blind ($M_k = 0$) channel estimates is theoretically $+\infty$, which allows us to trivially conclude that semi-blind ($M_k > 0$) methods have a better performance than blind methods.

The blind CRB in (11) corresponds to the estimation of the channel but with some a priori information. For the semi-blind CRB, this a priori knowledge is not used. This is why direct comparison of these two CRBs is not possible.

The projection of the CRB onto the space orthogonal to h can be interpreted as the CRB knowing the proper scaling factor for the channel (again the proof is similar to Appendix 1). $P_h^\perp CRB_B P_h^\perp = CRB_B$ which confirms our interpretation of CRB_B, and $P_h^\perp CRB_{SB} P_h^\perp \leq CRB_{SB}$. Fig. 1 (right) shows the different quantities ($M_k = 0$ corresponds to the blind case). The introduction of very few known symbols is sufficient to improve performance significantly w.r.t. to blind estimation (more evident on other channels we tested).

4. STOCHASTIC MODEL

The input symbols are no longer considered as deterministic, but Gaussian. $Y \sim \mathcal{N}(\mathcal{A}_k h, R_{YY})$, $R_{YY} = \sigma_a^2 \mathcal{T}_u \mathcal{T}_u^H + \sigma_v^2 I$:

$$f(Y/\theta) = \frac{1}{\pi^{Mm} \det R_{YY}} exp\left[-(Y - \mathcal{A}_k h)^H R_{YY}^{-1} (Y - \mathcal{A}_k h)\right] \tag{15}$$

where $\mathcal{A}_k h = \mathcal{T}_k A_k$. Unlike in the deterministic case, the input symbols are no longer nuisance parameters in the estimation of $\theta = h$. Since $J_{\theta\theta} \neq 0$ due to the factor $\det R_{YY}$, we consider the real CRB, which is determined via (6) thanks to the quantities:

$$J_{\theta\theta}(i,j) = \left(\mathcal{A}_k{}^H R_{YY}^{-1} \mathcal{A}_k\right)(i,j) + trace\left\{R_{YY}^{-1}\left(\frac{\partial R_{YY}}{\partial \theta_i^*}\right) R_{YY}^{-1} \left(\frac{\partial R_{YY}}{\partial \theta_j^*}\right)^H\right\}, \tag{16}$$

$$J_{\theta\theta^*}(i,j) = trace\left\{R_{YY}^{-1}\left(\frac{\partial R_{YY}}{\partial \theta_i^*}\right) R_{YY}^{-1}\left(\frac{\partial R_{YY}}{\partial \theta_j^*}\right)\right\}, \tag{17}$$

$$\frac{\partial R_{YY}}{\partial \theta_i^*} = \sigma_a^2 \mathcal{T}_u (\mathrm{H}_N) \mathcal{T}_u^H \left(\frac{\partial \mathrm{H}_N}{\partial \theta_i}\right) \tag{18}$$

The Gaussian model is more robust than the deterministic one: in the blind case, the module can indeed be estimated, but not the phase however, which leads to a singular FIM. The singular vector is $h_R^{(2)}$ (12). Again, the pseudo-inverse can be interpreted as the CRB knowing the phase of the channel. Let $\hat{h}_R$ be a blind estimate; this CRB corresponds to $\hat{\hat{h}}_R = e^{j\varphi} \hat{h}_R$, where φ is adjusted so that $h_R^{(1)T} \hat{\hat{h}}_R > 0$ and $h_R^{(2)T} \hat{\hat{h}}_R = 0$.

Fig. 1 (right) shows the Gaussian semi-blind curve for the channel in (14). Gaussian and deterministic performance seem equivalent when a significant number of symbols is known. When few symbols are known, Gaussian semi-blind performs better, which other simulations confirmed.

To compare Gaussian blind and semi-blind mode, we projected the different CRBs onto the space orthogonal to the singular vector of the Gaussian blind FIM, which can be interpreted as the CRBs knowing the phase of the channel (see fig. 1 (right)).

A Gaussian maximum likelihood method for semi-blind channel estimation is presented in [1].

5. MAXIMUM LIKELIHOOD METHODS

CRBs are considered as good performance indicators for Maximum Likelihood (ML) estimators. Indeed, under certain conditions, ML estimates are asymptotically efficient, i.e. unbiased and reaching the CRB. In the semi-blind case, although the estimation of the phase is not consistent (for a non-asymptotical number of known symbols), we conjecture that Gaussian ML (GML) estimates are efficient. For the blind case, we also conjecture that the CRB (the pseudo-inverse) is attained by $\hat{\hat{h}}_{G\bar{M}L}$.

However, deterministic ML (DML) estimates cannot reach the CRB: even asymptotically, the number of observations per symbol is finite, this is why the input symbols cannot be estimated consistently. Eliminating the symbols (as in the SRM method), one finds that there are essentially $m-1$ equations per symbol period for the m channel impulse responses. Hence these can be estimated consistently. But the inconsistency of the symbol estimates prevents efficiency of the channel estimate.

We derive the asymptotic DML performance (essentially for $N = 1$) in order to compare it to the corresponding CRB and GML. The log-likelihood function is:

$$\ln f(Y/\theta) = c' + \frac{1}{\sigma_v^2} \|(Y - \mathcal{T}_k A_k) - \mathcal{T}_u A_u\|^2 \tag{19}$$

Solving w.r.t. the unknown symbols A_u, we find:

$$A_u = (\mathcal{T}_u^H \mathcal{T}_u)^{-1} \mathcal{T}_u^H (Y - \mathcal{T}_k A_k) \tag{20}$$

which is the output of a MMSE-ZF equalizer, with feedback of A_k. Substituting:

$$\min_{A_u} \|(Y - \mathcal{T}_k A_k) - \mathcal{T}_u A_u\|^2 = (Y - \mathcal{T}_k A_k)^H P_{\mathcal{T}_u}^\perp (Y - \mathcal{T}_k A_k) \tag{21}$$

5.1. Purely Blind Case

The minimization criterion is:

$$\min_h Y^H P_{\mathcal{T}}^\perp Y. \tag{22}$$

This criterion is insensitive to the module and phase of the channel: those quantities cannot be determined. Consider putting $h = \alpha e^{j\varphi} \bar{h}$, where α is the module and φ the phase of the channel. For N=1, criterion (22) leads to:

$$\max_{\bar{h}} \bar{h}^H \widehat{r}_{yy} \bar{h} \tag{23}$$

$\widehat{r}_{yy} = \frac{1}{M} \sum_{i=1}^M y_i y_i^H$ is the sample covariance matrix, which converges to the true received signal covariance matrix: $r_{yy} = \sigma_a^2 h h^H + \sigma_v^2 I$. $\widehat{\bar{h}}$ is the maximal eigenvector of $\widehat{r}_{YY}$ which is asymptotically ($M = M_u \to \infty$) normally distributed with mean equal to the maximal eigenvector of r_{yy} i.e. $\bar{h}$, and covariance matrix:

$$C_{\tilde{\bar{h}}} = \frac{1}{\|h\|^2} \frac{1}{M} \frac{\sigma_v^2}{\sigma_a^2} \left[1 + \frac{\sigma_v^2}{\sigma_a^2 \|h\|^2}\right] P_h^\perp \tag{24}$$

The asymptotic blind CRB is ($\tilde{\hat{h}} = h - \hat{\hat{h}}$):

$$CRB_B = \frac{1}{M} \frac{\sigma_v^2}{\sigma_a^2} P_h^\perp \leq \|h\|^2 C_{\tilde{\bar{h}}} = C_{\tilde{\hat{h}}} \tag{25}$$

The blind DML estimate reaches asymptotically the CRB only when the SNR tends to ∞.

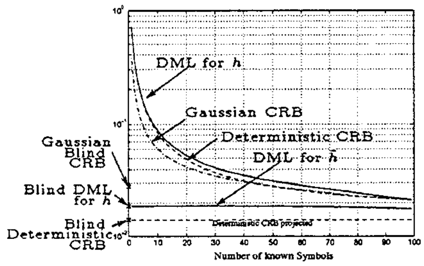

Figure 2. CRBs and DML and GML performance for N=1

5.2. Semi-blind Case

The minimization criterion is:

$$\min_{h}(\mathrm{Y}-\mathcal{T}_k A_k)^H P^{\perp}_{\mathcal{T}_u}(\mathrm{Y}-\mathcal{T}_k A_k) \tag{26}$$

Using the decomposition $h = \alpha e^{j\varphi}\bar{h}$ and, for $N=1$, $\mathcal{T}_u^H\mathcal{T}_k = 0$, the criterion is:

$$\min_{\bar{h},\alpha,\varphi}\left\{\mathrm{Y}^H\mathrm{Y} - 2\alpha\mathrm{Re}(\mathrm{Y}^H\mathcal{A}_k\bar{h}e^{j\varphi}) + \alpha^2\|A_k\|^2 - \bar{h}^H\widehat{r^u_{yy}}\bar{h}\right\} \tag{27}$$

We find:

$$\widehat{\varphi} = -arg\left\{\mathrm{Y}^H\mathcal{A}_k\bar{h}\right\},\ \widehat{\alpha} = \frac{|\mathrm{Y}^H\mathcal{A}_k h|}{\|A_k\|^2},\text{ resulting in:} \tag{28}$$

$$\max_{\bar{h}}\bar{h}^H\left\{M_u\widehat{r^u_{yy}} + vv^H\right\}\bar{h} \tag{29}$$

where $\widehat{r^u_{yy}} = \frac{1}{M_u}\sum_{(u)}\mathrm{y}_i\mathrm{y}_i^H$ (the sum is taken over the y_i containing unknown symbols only), $v = \sum_{(k)} a_i^*\mathrm{y}_i/\|A_k\|$ (the sum is taken over the y_i containing known symbols only). The problem decomposes into a blind part and a training sequence part (this is also true for $N>1$). It can be shown that asymptotically ($M_u \to \infty$ and $M_k \to \infty$) the error covariance matrix of $\hat{\bar{h}}$ is a linear combination of the error covariance matrices of both parts: We will not give the expression here for lack of space.

Fig. 2 shows an example with SNR=0dB, M=100 and the channel $\mathrm{H}_1 = [\ 0.4005\quad -1.3414\quad 0.3750\ ]^H$. The trace of the covariance matrices and the CRB is plotted. M_k varies from 1 to 99 and $M_u = M+N-1-M_k$: although the covariance expression is only asymptotically valid, it seems to make sense for small values of M_k or M_u also. The curve showing the results for GML is in fact the Gaussian CRB, which should be asymptotically attained. As predicted, DML does not reach its CRB asymptotically and its performance is below that of GML especially when few symbols are known. The performance of DML for $\bar{h}$ is also shown: in the case N=1, it is quite insensitive to variations in M_k. This number of known symbols is rather important though in the determination of the phase.

From this study, we can conclude the following asymptotical relationships:

$$CRB_{Gauss} = C_{GML} \leq C_{DML} \tag{30}$$

$$tr(CRB_{Gauss}) \leq tr(CRBdet) \leq tr(C_{DML}),\ (\text{for } M_k \text{ small}) \tag{31}$$

After the submission of this paper, we became aware of [2] in which a semiblind GML method and corresponding Cramer-Rao Bound have also been pursued; see [1] for some comments.

APPENDIX 1

Consider the vector of parameters $\theta = \left[\,A^H\ \ h^H\,\right]^H$ with true value $\theta^o = \left[\,A^{oH}\ \ h^{oH}\,\right]^H$ and $\hat{\theta}$ an unbiased estimate. The eigendecomposition of $J_{\tilde{h}\tilde{h}}(\theta) \triangleq \frac{1}{\sigma_v^2}\mathcal{A}^H P_{\mathcal{T}}\mathcal{A}$ is $V(\theta)\Lambda(\theta)V(\theta)^H = V_1(\theta)\Lambda_1 V_1(\theta)^H$ with $V(\theta) = [\ V_1(\theta)\quad V_2(\theta)\ ]$ where $V_2(\theta)$ spans the null space of $J_{\tilde{h}\tilde{h}}(\theta)$.

Let's consider the following change of variables: $\varphi = \begin{bmatrix} A \\ h' \end{bmatrix} = \begin{bmatrix} I & 0 \\ 0 & V(\theta^o)^H \end{bmatrix}\theta = \mathcal{V}^{oH}\theta$. $h' = V^{oH}h$. Suppose that we know $h_2'^{\,o} = V_2^{oH}h^o$, so that the channel parameter of interest becomes: $h_1' = V_1^{oH}h$. The overall parameter of interest is then: $\varphi_1 = \begin{bmatrix} I & 0 \\ 0 & V_1^{oH} \end{bmatrix}\theta = \mathcal{V}_1^{oH}\theta$. $\mathcal{V}^o = [\ \mathcal{V}_1^o\quad \mathcal{V}_2^o\ ]$. $\varphi_2 = \varphi_2^o = \mathcal{V}_2^{oH}\theta^o = h'^{\,o}_2$; $\varphi^H = [\ \varphi_1^H\quad \varphi_2^{oH}\ ]$.

Let $J_{\tilde{h}_1\tilde{h}_1}(\varphi) \triangleq \left((J^{-1}_{\varphi_1\varphi_1})_{h_1h_1}\right)^{-1}(\varphi)$, since $J_{\varphi_1\varphi_1}$ is non singular. From:

$$J_{\tilde{h}_1\tilde{h}_1}(\varphi) = \frac{1}{\sigma_v^2}\left(\frac{\partial \mathrm{Y}_S^H}{\partial h_1^*}\right)P_{\mathcal{T}}^{\perp}\left(\frac{\partial \mathrm{Y}_S^H}{\partial h_1^*}\right)^H \tag{32}$$

$$\text{and } \frac{\partial \mathrm{Y}_S^H}{\partial h^*} = V_1^o\frac{\partial \mathrm{Y}_S^H}{\partial h_1'^{\,*}} \tag{33}$$

$$\text{we get: } J_{\tilde{h}\tilde{h}}(\theta) = V_1^o J_{\tilde{h}_1'\tilde{h}_1'}(\varphi)V_1^{oH}. \tag{34}$$

Then:

$$\begin{aligned} &J_{\tilde{h}_1'\tilde{h}_1'}(\varphi^o) = V_1^{oH}J_{\tilde{h}\tilde{h}}(\theta^o)V_1^o = \Lambda_1^o \\ &\Rightarrow J^{+}_{\tilde{h}\tilde{h}}(\theta^o) = V_1^o J^{-1}_{\tilde{h}_1'\tilde{h}_1'}(\varphi^o)V_1^{oH}. \end{aligned} \tag{35}$$

Let's get back to the initial parameter $\theta = \mathcal{V}_1^o\varphi_1 + \mathcal{V}_2^o\varphi_2^o$. This is our initial parameter but knowing φ_2^o. The CRB for a transformation of parameters [3] gives:

$$C_{\tilde{\theta}} \geq \left(\frac{\partial\theta^H}{\partial\varphi_1^*}\right)^H J^{-1}_{\varphi_1\varphi_1}(\varphi^o)\left(\frac{\partial\theta^H}{\partial\varphi_1^*}\right). \tag{36}$$

Now $\theta = \mathcal{V}_1^o\varphi_1 + \mathcal{V}_2^o\varphi_2^o \Rightarrow \frac{\partial\theta^H}{\partial\varphi_1^*} = \mathcal{V}_1^{oH}$, hence:

$$\begin{aligned} C_{\tilde{\theta}} \geq\ & \mathcal{V}_1^o J^{-1}(\varphi_1)\mathcal{V}_1^{oH} \\ =\ & \begin{bmatrix} I & 0 \\ 0 & V_1^o \end{bmatrix}\begin{bmatrix} * & * \\ * & J^{-1}_{\tilde{h}_1'\tilde{h}_1'}(\varphi^o) \end{bmatrix}\begin{bmatrix} I & 0 \\ 0 & V_1^o \end{bmatrix}^H \\ =\ & \begin{bmatrix} * & * \\ * & V_1^o J^{-1}_{\tilde{h}_1'\tilde{h}_1'}(\varphi^o)V_1^{oH} \end{bmatrix} \end{aligned} \tag{37}$$

Using (35), we get: $C_{\tilde{h}\tilde{h}}(\theta^o) \geq J^{+}_{\tilde{h}\tilde{h}}(\theta^o)$.

Note that this is the CRB for any unbiased estimator that satisfies $\mathcal{V}_2^{oH}\hat{h} = \varphi_2^o$ and hence can be written as $\hat{h} = \mathcal{V}_1^o\hat{\varphi}_1 + \mathcal{V}_2^o\varphi_2^o$ with $\hat{\varphi}_1 = \mathcal{V}_1^{oH}\hat{h}$. Any arbitrary unbiased estimate $\hat{h} = \mathcal{V}_1^o\hat{\varphi}_1 + \mathcal{V}_2^o\hat{\varphi}_2$ can be transformed into an estimate $\hat{\hat{h}}$ by forcing $\hat{\hat{h}}$ such that $\mathcal{V}_2^{oH}\hat{\hat{h}} = \varphi_2^o$ in which case $\hat{\hat{h}} = \mathcal{V}_1^o\hat{\varphi}_1 + \mathcal{V}_2^o\varphi_2^o$.

REFERENCES

[1] E. de Carvalho and D. T. M. Slock. "Maximum-Likelihood FIR Multi-Channel Estimation with Gaussian Prior for the Symbols". In *Proc. ICASSP 97 Conf.*, Munich, Germany, April 1997.

[2] J.L. Bapat. *Partially Blind Identification of FIR Channels for QAM Signals.* PhD thesis, Pennsylvania State University, Aug. 1996.

[3] S. M. Kay. *Fundamentals of Statistical Signal Processing Estimation Theory.* Prentice-Hall, Englewood Cliffs, NJ, 1993.