Asymptotic Performance of ML Methods for Semi-Blind Channel Estimation

Elisabeth de Carvalho and Dirk T.M. Slock†
EURECOM Institute
2229 route des Crêtes, B.P. 193, 06904 Sophia Antipolis Cedex, FRANCE
{carvalho, slock}@eurecom.fr

Abstract

Two channel estimation methods are often opposed: training sequence methods which use the information coming from known symbols and blind methods which use the information coming from the received signal without integrating the possible knowledge of symbols. Semi-blind methods combine both informations and appear more powerful than both methods separately. Two Maximum-Likelihood approaches to semi-blind SIMO channel estimation are presented, one based on a deterministic model and another on a Gaussian model. Their asymptotic performance are studied and compared to the Cramer-Rao Bounds. The superiority of semi-blind over blind and training sequence methods, and of the Gaussian approach is demonstrated.

1. Introduction

Most of the actual mobile communication standards, like GSM, include a training sequence to estimate the channel, or simply some known symbols used for synchronization or as guard intervals. Training sequence methods base the parameter estimation on the received signal containing known symbols only and all the other observations, containing (some) unknown symbols, are ignored. Blind methods are based on the whole received signal, containing known and unknown symbols, but no use is made of the knowledge of some input symbols. The purpose of semi-blind methods is to combine both training sequence and blind informations.

Semi-blind techniques can then avoid the possible ill-conditioning of blind techniques, like the case of channel with closely-spaced zeros. With few known symbols, any channel becomes semi-blindly identifiable. Coupling blind and training sequence informations improves performance w.r.t. purely blind estimation but also w.r.t. to training sequence based estimation, possibly allowing good estimation even when the training sequence is short. All those aspects were already mentioned in [1] and [2]. We propose two approaches to semi-blind channel estimation based on ML, which we believe are among the less complex and more powerful methods. We study their performance asymptotically in the number of known and unknown symbols and compare it to the Cramer-Rao Bounds (CRB).

We consider a single-user multichannel model: this model results from the oversampling of the received signal and/or from reception by multiple antennas. Consider a sequence of symbols $a(k)$ received through m channels of length N and coefficients $\boldsymbol{h}(i)$:

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

$\boldsymbol{v}(k)$ is an additive independent white Gaussian noise with $r_{\boldsymbol{vv}}(k \Leftrightarrow i) = \mathrm{E}\, \boldsymbol{v}(k)\boldsymbol{v}(i)^H = \sigma_v^2 I_m\, \delta_{ki}$.

We consider the symbol constellation as known. When the input symbols are real, it will be advantageous to consider separately the real and imaginary parts of the channel and received signal as:

$$\begin{bmatrix} \mathrm{Re}(\boldsymbol{y}(k)) \\ \mathrm{Im}(\boldsymbol{y}(k)) \end{bmatrix} = \sum_{i=0}^{N-1} \begin{bmatrix} \mathrm{Re}(\boldsymbol{h}(i)) \\ \mathrm{Im}(\boldsymbol{h}(i)) \end{bmatrix} a(k \Leftrightarrow i) + \begin{bmatrix} \mathrm{Re}(\boldsymbol{v}(k)) \\ \mathrm{Im}(\boldsymbol{v}(k)) \end{bmatrix} \qquad (2)$$

Let's rename $\boldsymbol{y}(k) = [\mathrm{Re}^H(\boldsymbol{y}(k))\ \mathrm{Im}^H(\boldsymbol{y}(k))]^H$, and idem for $\boldsymbol{h}(i)$ and $\boldsymbol{v}(k)$, we get again (1), but this time, all the quantities are real. The number of channels gets doubled, which has for advantage to increase diversity. Note that the monochannel case does not exist in the real case.

Assume we receive M samples, concatenated in the vector $\boldsymbol{Y}_M(k)$:

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

$\boldsymbol{Y}_M(k) = [\boldsymbol{y}^H(k \Leftrightarrow M+1) \cdots \boldsymbol{y}^H(k)]^H$, similarly for $\boldsymbol{V}_M(k)$, and $A_M(k) = \left[a^H(k \Leftrightarrow M \Leftrightarrow N+2) \cdots a^H(k)\right]^H$, where $(.)^H$ denotes Hermitian transpose. $\mathcal{T}_M(h)$ is a

† 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 by Laboratoires d'Electronique PHILIPS under contract LEP 95FAR008 and by the EURECOM Institute

block Toeplitz matrix filled out with the channel coefficients grouped in the vector h. We shall simplify the notation in (2) with $k = M-1$ to:

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

We assume that the known symbols are grouped and for notational simplicity at the beginning of the burst: $A = [A_k^H \;\; A_u^H]^H$, A_k contains the M_k known symbols and A_u, the M_u unknown symbols.

2. ML Methods

2.1. Deterministic ML (DML)

In the deterministic model both input symbols and channel coefficients are considered as deterministic. We are interested in the joint estimation of h and the unknown symbols, which is decoupled from the estimation of σ_v^2. This estimation is based on the following DML criterion:

$$\max_{A_u,h} f(\boldsymbol{Y}|h) \quad \Leftrightarrow \quad \min_{A_u,h} \|\boldsymbol{Y} - \mathcal{T}(h)A\|^2 \qquad (5)$$

$f(\boldsymbol{Y}|h)$ is the complex probability density function when A is complex (which exists as $\boldsymbol{V}$ is circular) and the real one when A is real. $\boldsymbol{Y} = \mathcal{T}_k(h)A_k + \mathcal{T}_u(h)A_u + \boldsymbol{V}$, and optimizing w.r.t. the unknown symbols, we get:

$$A_u = \left(\mathcal{T}_u^H(h)\mathcal{T}_u(h)\right)^{-1}\mathcal{T}_u^H(h)\left(\boldsymbol{Y} - \mathcal{T}_k(h)A_k\right) \qquad (6)$$

which is the output of the non-causal minimum mean squared error zero-forcing decision feedback equalizer with feedback of the known symbols. Substituting (6) in (5) we get the following minimization criterion for h:

$$\min_h \left(\boldsymbol{Y} - \mathcal{T}_k(h)A_k\right)^H P^{\perp}_{\mathcal{T}_u(h)}\left(\boldsymbol{Y} - \mathcal{T}_k(h)A_k\right) \qquad (7)$$

where $P^{\perp}_{\mathcal{T}_u(h)} = I - \mathcal{T}_u(h)\left(\mathcal{T}_u^H(h)\mathcal{T}_u(h)\right)^{-1}\mathcal{T}_u^H(h)$. We will denote $\mathcal{C}(h)$ the cost function. For commodity reasons, when A is complex, it is taken equal to $\frac{1}{\sigma_v^2}$ times the expression in (7), when A is real it is $\frac{2}{\sigma_v^2}$ times this expression.

2.2. Gaussian ML (GML)

In the Gaussian Model [1],[2], the channel coefficients are still considered as deterministic but the input symbols as Gaussian random variables. This hypothesis, although false, allows to robustify the estimation problem and improves performance w.r.t. DML, as will be seen.

In the Gaussian model for (4), $\boldsymbol{V} \sim \mathcal{N}(0, C_{VV})$ is independent of $A \sim \mathcal{N}(A^o, C_{AA})$. A^o is the prior mean for the symbols. In the Gaussian case, the estimation of the channel can be done without the estimation of the unknown input symbols: GML considers the joint estimation of h and the coefficients of C_{VV}. $\boldsymbol{Y} \sim \mathcal{N}(\mathcal{T}(h)A^o, C_{YY})$, $C_{YY} = \mathcal{T}(h)C_{AA}\mathcal{T}^H(h) + C_{VV}$ and the GML criterion is $\max\limits_{h,\sigma_v^2} f(Y|h)$, or:

$$\min_{h,\sigma_v^2}\left\{\ln\det C_{YY} + \left(\boldsymbol{Y} - \mathcal{T}(h)A^o\right)^H C_{YY}^{-1}\left(\boldsymbol{Y} - \mathcal{T}(h)A^o\right)\right\} \qquad (8)$$

We will specialize this general model to the semi-blind case as follows: $A^o = \begin{bmatrix} A_k \\ 0 \end{bmatrix}$ and $C_{AA} = \begin{bmatrix} \epsilon I & 0 \\ 0 & \sigma_a^2 I \end{bmatrix}$, where ϵ is arbitrarily small. Furthermore, as already mentioned, we take $C_{VV} = \sigma_v^2 I$. We will denote $\mathcal{C}(h)$ the cost function. When A is complex, it is taken equal to the expression in (7), when A is real, it is 2 times this expression.

2.3. Identifiability Conditions

Let θ be the parameter vector to be estimated. θ is said to be identifiable if $\forall \boldsymbol{Y}$, $f(\boldsymbol{Y}|\theta) = f(\boldsymbol{Y}|\theta') \Rightarrow \theta = \theta'$. For Gaussian distributions, which is our case, identifiability is based on the first and second-order moments.

We give here identifiability conditions on the channel characteristics only. We assume that the burst length is sufficiently long and, for DML, that the entry contains at least $2N-1$ modes, which is a sufficient identifiablity condition for DML. For GML, the uncorrelated symbols are maximally excited.

2.3.1 Blind

Blind DML cannot estimate monochannels as well as the zeros and a scale factor of a multichannel. A multichannel can be estimated up to a complex scale factor if and only if it is irreducible [3], [4].

Blind GML requires less demanding conditions. Monochannels can be identified up to a phase factor if and only if they are minimum-phase. Multichannels can be estimated, up to a phase factor, if and only if the zeros, if any, are minimum-phase.

2.3.2 Semi-Blind

Semi-blind allows the estimation of any channel. Monochannels as well as the zeros or the ambiguous blind scale factor are estimated thanks to the training sequence. One needs $2N_z-3$ known symbols containing at least N_z-1 modes, where N_z is the number of zeros. 1 known symbol is sufficient when the channel is irreducible, and $2N-1$ are required for a monochannel.

For GML, identification is possible from the mean alone: 1 known symbol (not located at the edges of the burst) is sufficient to allow the identification of any channel. Note then that the blind part of the GML brings information on monochannels and zeros of multichannels.

Continuous ambiguities for identifiability corresponds to singularities in the (Fisher-like) information matrices (IM) J and $\mathcal{J}$ below, as will be seen. Binary ambiguities, like a sign for example, will not lead to singularity for the IM.

3. Asymptotic Performance

We explain here the general procedure to compute the asymptotic ML performance. θ is the complex parameter vector, $\theta_R = [\mathrm{Re}^H(\theta)\ \mathrm{Im}^H(\theta)]^H$, the real associated parameter vector, θ_o and θ_{o_R} the true values, $\hat{\theta}$ and $\hat{\theta}_R$ the ML estimates and $\Delta\theta = \hat{\theta} - \theta_o$, $\Delta\theta_R = \hat{\theta}_R - \theta_{o_R}$, the errors. Assuming consistency, we can proceed to the following Taylor development of $\mathcal{C}(\theta)$, the cost function, around θ_o.

$$\left.\frac{\partial \mathcal{C}(\theta)}{\partial \theta_R}\right|_{\theta=\hat{\theta}} = \left.\frac{\partial \mathcal{C}(\theta)}{\partial \theta_R}\right|_{\theta=\theta_o} + \left.\frac{\partial}{\partial \theta_R}\left(\frac{\partial \mathcal{C}(\theta)}{\partial \theta_R}\right)^H\right|_{\theta=\theta_o} \Delta\theta_R + o(\Delta\theta_R) \tag{9}$$

with $\left.\frac{\partial \mathcal{C}(\theta)}{\partial \theta_R}\right|_{\theta=\hat{\theta}} = 0$, then asymptotically:

$$\Delta\theta_R = \left(\frac{\partial}{\partial \theta_R}\left(\frac{\partial \mathcal{C}(\theta)}{\partial \theta_R}\right)^H\right)^{-1} \frac{\partial \mathcal{C}(\theta)}{\partial \theta_R}. \tag{10}$$

Let's denote:

$$\mathcal{J}^{(1)}_{\theta_R} = E\left(\frac{\partial \mathcal{C}(\theta)}{\partial \theta_R}\right)\left(\frac{\partial \mathcal{C}(\theta)}{\partial \theta_R}\right)^H,\ \mathcal{J}^{(2)}_{\theta_R} = -E\frac{\partial}{\partial \theta_R}\left(\frac{\partial \mathcal{C}(\theta)}{\partial \theta_R}\right)^H \tag{11}$$

In our specific cases, asymptotically, by the law of large numbers:

$$\frac{\partial}{\partial \theta_R}\left(\frac{\partial \mathcal{C}(\theta)}{\partial \theta_R}\right)^H \sim \mathcal{J}^{(2)}_{\theta_R} \tag{12}$$

Then, $\Delta\theta_R \sim \mathcal{N}(0, C_{\Delta\theta_R})$, with error covariance matrix:

$$C_{\Delta\theta_R} = \left(\mathcal{J}^{(2)}_{\theta_R}\right)^{-1} \mathcal{J}^{(1)}_{\theta_R} \left(\mathcal{J}^{(2)}_{\theta_R}\right)^{-1} \tag{13}$$

As we are working with complex quantities, we found it easier to manipulate derivation w.r.t. complex vectors, defined as: $\frac{\partial}{\partial \theta} = \frac{1}{2}\left(\frac{\partial}{\partial \alpha} - j\frac{\partial}{\partial \beta}\right)$, with $\theta = \alpha + j\beta$. Let:

$$J^{(1)}_{\varphi\psi} = E\left(\frac{\partial \mathcal{C}(\theta)}{\partial \varphi^*}\right)\left(\frac{\partial \mathcal{C}(\theta)}{\partial \psi^*}\right)^H,\ J^{(2)}_{\varphi\psi} = -E\frac{\partial}{\partial \varphi^*}\left(\frac{\partial \mathcal{C}(\theta)}{\partial \psi^*}\right)^H \tag{14}$$

$\mathcal{J}^{(1)}_{\theta_R}$ and $\mathcal{J}^{(2)}_{\theta_R}$ can be expressed in terms of $J_{\theta\theta}$ and $J_{\theta\theta^*}$, in particular, $\mathcal{J}^{(2)}_{\theta_R}$ equals:

$$2\begin{bmatrix} \mathrm{Re}(J^{(2)}_{\theta\theta}) & -\mathrm{Im}(J^{(2)}_{\theta\theta}) \\ \mathrm{Im}(J^{(2)}_{\theta\theta}) & \mathrm{Re}(J^{(2)}_{\theta\theta}) \end{bmatrix} + 2\begin{bmatrix} \mathrm{Re}(J^{(2)}_{\theta\theta^*}) & \mathrm{Im}(J^{(2)}_{\theta\theta^*}) \\ \mathrm{Im}(J^{(2)}_{\theta\theta^*}) & -\mathrm{Re}(J^{(2)}_{\theta\theta^*}) \end{bmatrix} \tag{15}$$

In the DML case, $J_{\theta\theta^*} = 0$, so when the input symbols are complex, we can work directly with complex quantities, and (13) can be compactly written as:

$$C_{\Delta\theta} = \left(J^{(2)}_{\theta\theta}\right)^{-1} J^{(1)}_{\theta\theta} \left(J^{(2)}_{\theta\theta}\right)^{-1} \tag{16}$$

We will treat the general case of a reducible channel: $\mathbf{H}(z) = \mathbf{H}_I(z)H_c(z)$, $\mathbf{H}_I(z)$ of length N_I is irreducible, $H_c(z)$ is a monochannel of length N_c and admits as zeros the $N_c - 1$ zeros of $\mathbf{H}(z)$. We assume that the first coefficient of $H_c(z)$ is equal to 1. For an irreducible channel, $H_c(z) = 1$ and is known. For a monochannel $\boldsymbol{H}(z) = H_c(z)$. $\mathcal{T}_M(h) = \mathcal{T}_M(h_I)\mathcal{T}_{M+N_I-1}(h_c) = \mathcal{T}(h_I)\mathcal{T}(h_c)$.

Furthermore, the asymptotic conditions will be:

(i) $M_k \to \infty$ and $M_u \to \infty$

(ii) $\frac{\sqrt{M_u}}{M_k} \to 0$.

We will not treat the case where M_k is finite and M_u infinite: the performance in that case is that of the blind method up to some ambiguities, that get estimated by the known symbols part.

One of the main challenges (for algorithm development) and interests of semi-blind methods is to give good performance when both blind and training sequence methods fail, and especially when the training sequence is too short to allow good channel estimation. Our asymptotic study does not allow to elucidate this phenomenon which happens because semi-blind methods also takes the observations into account that contain both known and unknown symbols.

3.1. Semi-Blind DML

As already mentioned the blind part does not contain any information on $\boldsymbol{H}_c(z)$, and $P^{\perp}_{\mathcal{T}_u(h)} = P^{\perp}_{\mathcal{T}_{u'}(h_I)}$, where $\mathcal{T}_{u'}(h_I)$ is $\mathcal{T}(h_I)$ with the first $M_k - N_c + 1$ columns removed. Under condition (i), the observations containing both known and unknown symbols can be neglected and the training sequence and blind contributions can be separated in the criterion as:

$$\|\boldsymbol{Y}_{TS} - \mathcal{T}_{TS}(h)A_k\|^2 + \boldsymbol{Y}_B^H P^{\perp}_{\mathcal{T}_{B(h_I)}} \boldsymbol{Y}_B \tag{17}$$

with $\mathcal{T}_{TS}(h) = \mathcal{T}_{M_k-N+1}(h)$ and $\mathcal{T}_B(h) = \mathcal{T}_{M-M_k}(h)$, $\boldsymbol{Y}_{TS}$ and $\boldsymbol{Y}_B$ designate resp. the observations with known and unknown symbols only.

The estimation of $\theta = [h_I^H\ \bar{h}_c^H]^H$ (where $\bar{h}_c$ is deduced from h_c by eliminating its first element equal to 1) can be proven to be consistent under conditions (i) and (ii). Condition (ii) allows the training sequence part not to be neglected in (17) and in (12). The different quantities of interest are:

$$\left\{\begin{array}{l} J^{(1)}_{h_I h_I} = J^{(2)}_{h_I h_I} + J'_{h_I h_I} \\ J^{(2)}_{h_I h_I} = \frac{1}{\sigma_v^2}\mathcal{A}^H_{I_{TS}} P^{\perp}_{\bar{\mathcal{A}}_{TS}} \mathcal{A}_{I_{TS}} + \frac{1}{\sigma_v^2}\mathcal{A}^H_{I_B} P^{\perp}_{\mathcal{T}_{B(h_I)}} \mathcal{A}_{I_B} \\ J'_{h_I h_I}(i,j) = \\ \quad \mathrm{tr}\left\{\mathcal{T}_B^H(\frac{\partial h_I}{\partial h_{I_i}}) P^{\perp}_{\mathcal{T}_{B(h_I)}} \mathcal{T}_B(\frac{\partial h_I}{\partial h_{I_j}}) (\mathcal{T}_B^H(h_I)\mathcal{T}_B(h_I))^{-1}\right\} \\ J^{(2)}_{h_c h_c} = \frac{1}{\sigma_v^2}\mathcal{A}^H_{c_{TS}} \mathcal{T}^H_{TS}(h_I) \\ \left[I - \mathcal{A}^H_{I_{TS}}\left(\mathcal{A}^H_{I_B} P^{\perp}_{\mathcal{T}_{B(h_I)}} \mathcal{A}_{I_B}\right)^{-1} \mathcal{A}_{I_{TS}}\right] \mathcal{T}_{TS}(h_I)\mathcal{A}_{c_{TS}} \end{array}\right. \tag{18}$$

with the notations defined by: $\bar{\mathcal{A}}_{TS} = \mathcal{T}_{TS}(h_I)\bar{\mathcal{A}}_{c_{TS}}$, $\mathcal{T}_{TS}(\bar{h}_c)A_k = \bar{\mathcal{A}}_{c_{TS}}\bar{h}_c$, $\mathcal{T}_{TS}(h_I)\mathcal{T}_{TS}(h_c)A_k = \mathcal{A}_{I_{TS}}h_I$, $\mathcal{T}_B(h_I)\mathcal{T}_B(h_c)A_u = \mathcal{A}_{I_B}h_I$.

Let CRB_{h_I} be the CRB for h_I and CRB_{h_c} the CRB for h_c, then asymptotically:

$$CRB_{h_I} = \left(J^{(2)}_{h_I h_I}\right)^{-1} \text{ and } CRB_{h_c} = \left(J^{(1)}_{h_c h_c}\right)^{-1}. \quad (19)$$

$\Delta h_I \sim \mathcal{N}(0, C_{\Delta h_I})$. Using equation (16):

$$C_{\Delta h_I} = CRB_{h_I} + \left(J^{(2)}_{h_I h_I}\right)^{-1} J'_{h_I h_I} \left(J^{(2)}_{h_I h_I}\right)^{-1}. \quad (20)$$

$\Delta h_c \sim \mathcal{N}(0, C_{\Delta h_c})$, hence $\Delta h_c = O_p(\frac{1}{\sqrt{M_k}})$ and

$$\begin{aligned} C_{\Delta h_c} = CRB_{h_c} + \left(\bar{\mathcal{A}}^H_{TS}\bar{\mathcal{A}}_{TS}\right)^{-1} \bar{\mathcal{A}}^H_{TS}\mathcal{A}_{I_{TS}} \left(J^{(2)}_{h_I h_I}\right)^{-1} \\ J'_{h_I h_I} \left(J^{(2)}_{h_I h_I}\right)^{-1} \mathcal{A}^H_{I_{TS}}\bar{\mathcal{A}}_{TS} \left(\bar{\mathcal{A}}^H_{TS}\bar{\mathcal{A}}_{TS}\right)^{-1} \end{aligned} \quad (21)$$

The DML ambiguous scale factor, not identifiable by blind estimation, is estimated thanks to the training sequence and its error evolves as $\frac{1}{\sqrt{M_k}}$. Note then that one component in Δh_I does not evolve as $\frac{1}{\sqrt{M_u}}$ whereas the remaining components evolve as $\frac{1}{\sqrt{M_k}}$.

The second terms in (20) and (21) are positive: DML for h_I and h_c does not reach the CRB. Their estimation is indeed coupled with the estimation of the unknown symbols which cannot be estimated consistently. The coupling prevents the channel estimates from being efficient. At high SNR however, these second terms being of order σ_v^2 and the CRBs of order σ_v^2 become negligible and the CRB is attained.

3.2. Blind DML

Here $M_u = M{+}N{\Leftrightarrow}1 \to \infty$. $H_c(z)$ cannot be estimated by blind DML: we assume $\boldsymbol{H}(z) = \boldsymbol{H}_I(z)$, which is blindly identifiable up to a complex scale factor: $J^{(2)}_{hh}$ will have one singularity spanned by h_o corresponding to this ambiguity. For regularization purpose, blind estimation performance will be computed under the constraint $h^H h_o = h_o^H h_o$ where h_o is the true channel value, or:

$$h = V_o\xi + h_o \quad (22)$$

where the columns of V_o form an orthonormal basis of the orthogonal complement of h_o. It can be shown that the estimation of ξ is consistent and (16) can be applied to ξ. $C_{\Delta h} = V_o C_{\Delta\xi} V_o^H$ implies:

$$C_{\Delta h} = J^{(2)}_{hh}{}^{+} J^{(1)}_{hh} J^{(2)}_{hh}{}^{+} \quad (23)$$

where $^+$ denotes the Moore-Penrose pseudo-inverse. $J^{(1)}_{hh}$ and $J^{(2)}_{hh}$ are given by (18) with $A_k = 0$. The asymptotic CRB with constraint (22) is the pseudo-inverse of the Fisher information matrix [1], which equals the IM $J^{(2)}_{hh}$:

$$CRB_h = J^{(2)}_{hh}{}^{+} \quad (24)$$

Asymptotically, $\Delta h \sim \mathcal{N}(0, C_{\Delta h})$ with:

$$C_{\Delta h} = CRB + J^{(2)}_{hh}{}^{+} J'_{hh} J^{(2)}_{hh}{}^{+} \quad (25)$$

The second term of (25) is positive semi-definite: blind DML does not attain the CRB asymptotically in M, but it does at high SNR (as mentioned in [4] also).

3.3. Semi-Blind GML

In the Gaussian, the estimation of σ_v^2 is not decoupled from the estimation of h. We have $\theta = [h_I^H \; \bar{h}_c^H \; \sigma_v^2]^H$. As DML, the GML criterion can be decomposed into training sequence and blind contributions as:

$$\begin{aligned} \tfrac{1}{\sigma_v^2}\left\|\boldsymbol{Y}_{TS} \Leftrightarrow \mathcal{T}_{TS}(h)A_k\right\|^2 + M_k \ln\sigma_v^2 \\ + \ln\det C_{Y_BY_B} + \boldsymbol{Y}_B^H C^{-1}_{Y_BY_B}\boldsymbol{Y}_B \end{aligned} \quad (26)$$

When the input symbols are complex, let $h_R = [\mathrm{Re}^H(h_I) \;\; \mathrm{Im}^H(h_I) \;\; \mathrm{Re}^H(\bar{h}_c) \;\; \mathrm{Im}^H(\bar{h}_c)]^H$ and $\theta'_R = [h_R^H \;\; \sigma_v^2]^H$. $\mathcal{J}^{(1)}_{\theta'_R} = \mathcal{J}^{(2)}_{\theta'_R} = \mathcal{J}_{\theta'_R}$, which we get by (15) thanks to the quantities:

$$\begin{aligned} J_{\theta\theta}(i,j) = \tfrac{1}{\sigma_v^2}\left([\mathcal{A}_o\;\bar{\mathcal{A}}_c]^H[\mathcal{A}_o\;\bar{\mathcal{A}}_c]\right)_{i,j}\delta_{\sigma_v^2} \Leftrightarrow \tfrac{M_k}{4\sigma_v^4}\bar{\delta}_{\sigma_v^2} \\ +\mathrm{tr}\left\{C^{-1}_{Y_BY_B}\left(\tfrac{\partial C_{Y_BY_B}}{\partial\theta_i^*}\right)C^{-1}_{Y_BY_B}\left(\tfrac{\partial C_{Y_BY_B}}{\partial\theta_j^*}\right)^H\right\} \end{aligned} \quad (27)$$

$$\begin{aligned} J_{\theta\theta^*}(i,j) = \mathrm{tr}\left\{C^{-1}_{Y_BY_B}\left(\tfrac{\partial C_{Y_BY_B}}{\partial\theta_i^*}\right)C^{-1}_{Y_BY_B}\left(\tfrac{\partial C_{Y_BY_B}}{\partial\theta_j^*}\right)\right\} \\ \Leftrightarrow \tfrac{M_k}{4\sigma_v^4}\bar{\delta}_{\sigma_v^2}, \text{ where:} \end{aligned} \quad (28)$$

$$\begin{cases} \frac{\partial C_{Y_BY_B}}{\partial h^*_{I_i}} = \sigma_a^2\mathcal{T}(h)\mathcal{T}^H(h_c)\mathcal{T}^H\left(\frac{\partial h_I^*}{\partial h^*_{I_i}}\right) \\ \frac{\partial C_{Y_BY_B}}{\partial h^*_{c_i}} = \sigma_a^2\mathcal{T}(h)\mathcal{T}^H\left(\frac{\partial h_c^*}{\partial h^*_{c_i}}\right)\mathcal{T}^H(h_I) \\ \frac{\partial C_{Y_BY_B}}{\partial\sigma_v^2} = \frac{1}{2}I \\ \delta_{\sigma_v^2} = 0 \quad \text{for } i \text{ or } j = N_o m + N_c \text{ and 1 elsewhere} \\ \bar{\delta}_{\sigma_v^2} = 1 \quad \text{for } i = j = N_o m + N_c \text{ and 0 elsewhere.} \end{cases} \quad (29)$$

When comparing asymptotically with the CRBs:

$$CRB_{h_I} = \left(\mathcal{J}^{-1}_{\theta'_R}\right)_{h_{I_R}} \text{ and } CRB_{h_c} = \left(\mathcal{J}^{-1}_{\theta'_R}\right)_{h_{c_R}}. \quad (30)$$

$\Delta h_{I_R} \sim \mathcal{N}(0, CRB_{h_I})$ and evolves as $\frac{1}{\sqrt{M_u}}$; $\Delta h_{c_R} \sim \mathcal{N}(0, CRB_{h_c})$ and evolves as $\frac{1}{\sqrt{M_u}}$. However, the phase component of Δh_{I_R} evolves as $\frac{1}{\sqrt{M_k}}$.

When the input symbols are real, again $\mathcal{J}^{(1)}_\theta = \mathcal{J}^{(2)}_\theta = J_{\theta\theta}$ of (27). $CRB_{h_I} = \left(\mathcal{J}^{-1}_\theta\right)_{h_oh_o}$ and $CRB_{h_c} =$

$\left(\mathcal{J}_{\theta}^{-1}\right)_{h_c h_c}$ $\Delta h_o \sim \mathcal{N}(0, CRB_{h_I})$ and evolves as $\frac{1}{\sqrt{M_u}}$; $\Delta h_c \sim \mathcal{N}(0, CRB_{h_c})$ and evolves as $\frac{1}{\sqrt{M_u}}$.

The CRB is asymptotically attained. Note that the CRB when the number of known symbols is finite (derived in [1]) is not attained. At high SNR, the influence of the estimation of σ_v^2 on the estimation of the channel becomes negligible, and performance for the estimation of h_I is the same as in the deterministic case.

3.4. Blind GML

Blind GML cannot estimate the channel phase factor. When the input symbols are real, this ambiguity is only about a sign and does not lead to singularity of $J_{\theta\theta}$. Also the ambiguities of the common zeros being minimum or maximum phase are binary. Discrete ambiguities do not lead to singularity of the FIM unlike continuous ambiguities. The error covariance matrices for h_I and h_c are the appropriate submatrices of $J_{\theta\theta}^{-1}$.

When the input symbols are complex, $\mathcal{J}_{\theta'_R}$ has one singularity spanned by $h'_s = [h_s^H \; 0 \; \cdots \; 0]^H$, where $h_s = [\Leftrightarrow\text{Im}^H(h_I) \;\; \text{Re}^H(h_I)]^H$ corresponding to the continuous ambiguity in the phase; again the non-minimum-phase channel zeros ambiguity does not appear in $\mathcal{J}_{\theta'_R}$. In this case, blind GML performance will be computed under the regularization constraint $h_{o_R}^H h_s = 0$, or:

$$h_{o_R} = V_{o_R}\xi + h_{o_s} \tag{31}$$

where the columns of V_{o_R} form an orthonormal basis of the orthogonal complement of h_{o_s}. Then:

$$C_{\Delta h_{I_R}} = \left(\mathcal{J}_{h_{I_R} h_{I_R}} \Leftrightarrow J_{h_{I_R} h_{\bar{I}_R}} (J_{h_{\bar{I}_R} h_{\bar{I}_R}})^{-1} J_{h_{\bar{I}_R} h_{\bar{I}_R}}\right)^{+} \tag{32}$$

$$C_{\Delta h_{c_R}} = \left(\mathcal{J}_{h_{c_R} h_{c_R}} \Leftrightarrow J_{h_{c_R} h_{\bar{c}_R}} (J_{h_{\bar{c}_R} h_{\bar{c}_R}})^{+} J_{h_{\bar{c}_R} h_{\bar{c}_R}}\right)^{-1} \tag{33}$$

$h_{\bar{I}_R} = [h_{o_R}^H \; \sigma_v^2]^H$ and $h_{\bar{c}_R} = [h_{I_R}^H \; \sigma_v^2]^H$. These quantities correspond to the CRBs under constraint (31): GML for h_I and h_c attains asymptotically the blind CRB.

4. Numerical Evaluations

For the simulations, we use an irreducible channel ($N = 5$, $m = 2$) and a reducible channel ($N_o = 2$, $N_c = 2$ $m = 4$), both randomly chosen, under SNR=10dB. The input symbols belong to a QPSK constellation and are i.i.d.. We plot the quantity: $\sqrt{\text{tr}(C_{\Delta h})}/||h||$.

In fig.1 (left), the irreducible semi-blind curves for DML and GML and the deterministic CRB is plotted w.r.t. the number of known symbols for a burst of length 150. The training sequence estimation mode (based on the known symbols of the semi-blind mode) is also shown as well as the case where all the input symbols are known, for reference. We essentially see the gain of semi-blind techniques

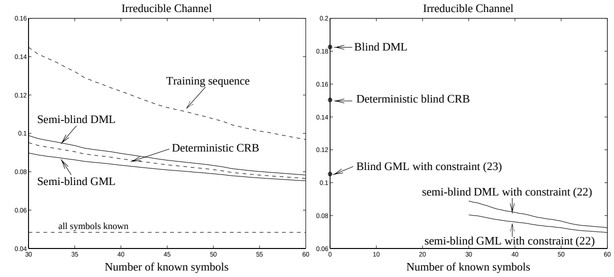

Figure 1. Semi-blind DML and GML: irreducible case

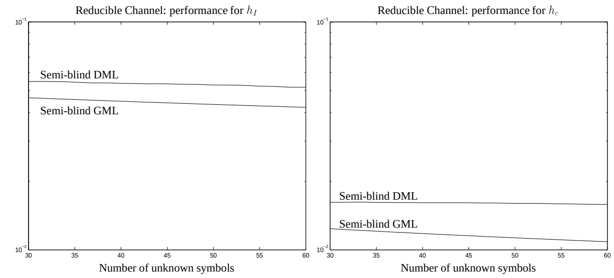

Figure 2. Semi-blind DML and GML: reducible case

w.r.t. the training sequence technique. In fig.1 (right), the blind and semi-blind performance with constraint (22) are shown: semi-blind appears better than blind. In both figures, GML appears better than DML. Other comments on such curves can be found in [1].

In fig.2, the reducible case is shown. For a fixed number of known symbols we plot the error variance w.r.t. the number of unknown symbols. The performance for the estimation of $H_c(z)$ by DML will tend to be constant as the number of unknown symbols grows. GML profits from the blind information, and the slope of the curve will eventually evolve in $\frac{1}{M_u}$.

References

[1] E. de Carvalho and D. T. M. Slock. “Cramer-Rao Bounds for Semi-blind, Blind and Training Sequence based Channel Estimation”. In *Proc. SPAWC 97 Conf.*, Paris, France, April 1997.

[2] 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.

[3] G. Xu, H. Liu, L. Tong, and T. Kailath. “A Least Squares Approach to Blind Channel Identification”. *IEEE Transactions on Signal Processing*, 43(12):2982–2993, Dec. 1995.

[4] Y. Hua. “Fast Maximum Likelihood for Blind Identification of Multiple FIR Channels”. *IEEE Transactions on Signal Processing*, 44(3):661–672, March 1996.