Linear Prediction Based Semi-Blind Estimation of MIMO FIR Channels

Abdelkader Medles, Dirk T.M. Slock
Eurecom Institute[1]
2229 route des Crêtes, B.P. 193
06904 Sophia Antipolis Cedex
FRANCE
Email: {medles,slock}@eurecom.fr

Elisabeth De Carvalho
Stanford University
David Packard El. Eng. Bldg.
350 Serra Mall, Room 203
Stanford, CA 94305-9515, USA
Email: carvalho@dsl.stanford.edu

Abstract — **The multichannel aspect has led to the development of a wealth of blind channel estimation techniques over the last decade. However, most of these blind techniques are not very robust and only allow to estimate the channel up to a number of ambiguities, especially in the MIMO case. On the other hand, all current standardized communication systems employ some form of known inputs to allow channel estimation. The channel estimation performance in those cases can always be improved by a semiblind approach which exploits both training and blind information. The purpose of this paper is to introduce semiblind techniques of which the complexity is not immensely much higher than that of training based techniques. The semiblind criteria are quadratic and combine a training based least-squares criterion with a blind criterion based on linear prediction. A variety of convenient linear prediction approaches are considered.**

I. Introduction

Consider linear digital modulation over a linear channel with additive Gaussian noise. Assume that we have p transmitters and $m > p$ receiving channels (e.g. antennas in BLAST or SDMA). The received signals can be written in the baseband as

$$y_i(t) \;=\; \sum_{j=1}^{p}\sum_{k} a_j(k)h_{ij}(t-kT) \;+\; v_i(t) \qquad (1)$$

where the $a_j(k)$ are the transmitted symbols from source j, T is the common symbol period, $h_{ij}(t)$ is the (overall) channel impulse response from transmitter j to receiver antenna i. We assume the channels to be FIR. In particular, after sampling we assume the (vector) impulse response from source j to be of length N_j. W.l.o.g., we assume the first non-zero vector impulse response sample to occur at discrete time zero, and we can assume the sources to be ordered so that $N_1 \geq N_2 \geq \cdots \geq N_p$. Let $N = \sum_{j=1}^{p} N_j$. The discrete-time Rx signal can be represented in vector form as

$$\mathbf{y}(k) = \sum_{j=1}^{p}\sum_{i=0}^{N_j-1} \mathbf{h}_j(i)a_j(k-i) + \mathbf{v}(k) = \sum_{i=0}^{N_1-1} \mathbf{h}(i)\mathbf{a}(k-i) + \mathbf{v}(k)$$

$$= \sum_{j=1}^{p} \mathbf{H}_{j,N_j}A_{j,N_j}(k) + \mathbf{v}(k) \;=\; \mathbf{H}_N\mathbf{A}_N(k) + \mathbf{v}(k)\;,$$

$$\mathbf{y}(k) = \begin{bmatrix} y_1(k)\\ \vdots \\ y_m(k)\end{bmatrix}, \mathbf{v}(k) = \begin{bmatrix} v_1(k)\\ \vdots \\ v_m(k)\end{bmatrix}, \mathbf{h}_j(k) = \begin{bmatrix} h_{1j}(k)\\ \vdots \\ h_{mj}(k)\end{bmatrix}$$

$$\mathbf{H}_{j,N_j} = [\mathbf{h}_j(0)\cdots\mathbf{h}_j(N_j-1)]\,, \mathbf{H}_N = \left[\mathbf{H}_{1,N_1}\cdots\mathbf{H}_{p,N_p}\right],$$
$$\mathbf{h}(k) = [\mathbf{h}_1(k)\cdots\mathbf{h}_p(k)]\,,\; A_{j,N_j}(k) = [a_j(k)\cdots a_j(k-N_j+1)]^T,$$
$$\mathbf{a}(k) = [a_1(k)\cdots a_p(k)]^T\;,\; \mathbf{A}_N(k) = \left[A_{1,N_1}^T(k)\cdots A_{p,N_p}^T(k)\right]^T$$

where superscripts ^{T}, ^{H} denote transpose and Hermitian transpose respectively. The multichannel aspect leads to a signal subspace when $m > p$ since $\mathbf{y}(k) = \mathbf{H}(q)\,\mathbf{a}(k) + \mathbf{v}(k)$ with $\mathbf{H}(q) = \sum_{i=0}^{N_1-1}\mathbf{h}(i)\,q^{-i}$ and q^{-1} the unit delay operator ($q^{-1}\mathbf{a}(k) = \mathbf{a}(k-1)$) and hence we get for the power spectral density matrix $S_{\mathbf{yy}}(z) = \mathbf{H}(z)\,S_{\mathbf{aa}}(z)\,\mathbf{H}^\dagger(z) + S_{\mathbf{vv}}(z) = \sigma_a^2\,\mathbf{H}(z)\,\mathbf{H}^\dagger(z) + \sigma_v^2\,I_m$. The existence of this signal subspace has led to the development of a wealth of blind channel estimation techniques over the last decade. Some of these techniques are relatively simple due to the modeling of the unknown input symbols as either deterministic unknowns of uncorrelated random variables (as opposed to exploiting their finite alphabet nature). The latter (uncorrelated) case is also called the Gaussian case because (only) second-order statistics are exploited. However, most of these blind techniques are not very robust in the sense that they often require precise knowledge of the channel length(s) and if transmission zeros can be handled, they are required to be minimum-phase. Furthermore, the blind techniques leave channel ambiguities, which can range from a simple scalar ambiguity factor for Single-Input Multiple-Output (SIMO) channels or certain Multiple-Input Multiple-Output (MIMO) channels (for certain techniques), to instantaneous or even convolutive mixtures of the sources for other MIMO channels. On the other hand, all current standardized communication systems employ some form of known inputs to allow channel estimation. The channel estimation performance in those cases can always be improved by a semiblind approach which exploits both training and blind information. The training information allows to resolve the blind ambiguities and robustifies the channel estimates. The purpose of this paper is to introduce semiblind techniques of which the complexity is not immensely much higher than that of training based techniques.

In the case of SIMO channels, we previously [4] introduced a simple semi-blind technique, taking the SRM (Subchannel Response Matching) method (also known as the Cross-Relation (CR) method) for the blind criterion. In the SRM approach, we use a simple parameterization of the noise subspace that is linear in the channel parameters. A blind criterion is then obtained by expressing orthogonality between this parameterized noise subspace and the data (for use in a semiblind approach, the data covariance matric should be denoised). This leads to a simple quadratic semiblind criterion. However, a linear parameterization of the noise subspace in terms of the channel parameters only exists in the SIMO case.

II. MIMO Linear Prediction

In the MIMO case, we propose here as in [5] to use linear

[1]Eurécom's research is partially supported by its industrial partners: Ascom, Swisscom, France Télécom, La Fondation CEGETEL, Bouygues Télécom, Thales, ST Microelectronics, Motorola, Hitachi Europe and Texas Instruments.

prediction quantities for the blind information. Linear prediction is applicable equally well to both the SIMO and MIMO cases. Two flavors can be obtained, depending whether the transmitted symbols are modeled as deterministic unknowns or as uncorrelated random sequences (in the deterministic case, for the purpose of linear prediction, some considerations are more straightforward if the symbols are considered as stationary sequences with unknown correlation).

Consider the problem of predicting $\mathbf{y}(k)$ from $\mathbf{Y}_L(k-1) = [\mathbf{y}^T(k-1)\cdots\mathbf{y}^T(k-L)]^T$, for noiseless received signal. The prediction error can be written as

$$\widetilde{\mathbf{y}}(k)|_{\mathbf{Y}_L(k-1)} = \mathbf{y}(k) - \widehat{\mathbf{y}}(k)|_{\mathbf{Y}_L(k-1)} = \mathbf{P}_L \mathbf{Y}_{L+1}(k) \quad (2)$$

with $\mathbf{P}_L = [\boldsymbol{P}_{L,0}\ \boldsymbol{P}_{L,1}\cdots\boldsymbol{P}_{L,L}]$, $\boldsymbol{P}_{L,0} = I_m$. Minimizing the prediction error variance leads to the following optimisation problem

$$\min_{\mathbf{P}_L} \mathbf{P}_L R_{YY} \mathbf{P}_L^H = \sigma^2_{\widetilde{\mathbf{y}},L} \quad (3)$$

hence

$$\mathbf{P}_L R_{YY} = \left[\sigma^2_{\widetilde{\mathbf{y}},L}\ 0\cdots 0\right] . \quad (4)$$

Let $\underline{L} = \left\lceil \frac{N-p}{m-p} \right\rceil$. The rank profile of $\sigma^2_{\tilde{y},L}$ behaves as a function of L generically (for an irreducible and column reduced MIMO channel) like

$$\text{rank}\left(\sigma^2_{\widetilde{\mathbf{y}},L}\right) \begin{cases} = p & , L \geq \underline{L} \\ = m - \underline{m} \in \{p+1,..,m\} & , L = \underline{L}-1 \\ = m & , L < \underline{L}-1 \end{cases} \quad (5)$$

where $\underline{m} = \underline{L}(m-p) - N + p \in \{0,1,\ldots,m-1-p\}$ represents the degree of singularity of $R_{YY,\underline{L}}$. For $L \geq \underline{L}$, $\widetilde{\mathbf{y}}(k)|_{\mathbf{Y}_L(k-1)} = \mathbf{h}(0)\mathbf{a}(k)$. For such L, let V_i be the eigenvectors of $\sigma^2_{\widetilde{\mathbf{y}},L}$ in order of decreasing eigenvalue, then $V_{1:p} = [V_1\cdots V_p]$ has the same column space as $\mathbf{h}(0)$ and $\overline{\mathbf{P}}(z) = V^H_{p+1:m}\mathbf{P}(z)$ satisfies $\overline{\mathbf{P}}(z)\mathbf{H}(z) = 0$ ($\overline{\mathbf{P}}(z)$ represents a parameterization of the noise subspace). Note that $\mathbf{P}(z)$ changes if the symbols are correlated (hence $\mathbf{P}(z)$ contains information about the symbol correlation) whereas $\overline{\mathbf{P}}(z)$ is insensitive to such correlation. To obtain the noisefree prediction quantities, we need to denoise an estimated covariance matrix via $\widehat{R}^d_{YY} = \widehat{R}_{YY} - \widehat{\sigma^2_v}\, I$ (partial denoising) or $\widehat{R}^d_{YY} = \lfloor \widehat{R}_{YY} - \widehat{\sigma^2_v}\, I\rfloor_+$ (full denoising). In the case of partial denoising, we used a generalized version (to covariance windowing) of the MIMO Levinson algorithm, which applies in the nonsingular indefinite case. Singular components appear then as negative semidefinite. In the case of full denoising, we determined the prediction quantities directly from the normal equations, with a generalized inverse $R^{\#} = U^{-H} D^{\#} U^{-1}$ where $R = UDU^H$ is the UDL triangular factorization of R and $D^{\#}$ is the Moore-Penrose inverse of the singular diagonal matrix D. As in [1], the columns in U corresponding to zeros in D are taken to be all zero, except for a unit diagonal element. In both approaches, the overestimation of L leads to consistent in SNR $\overline{\mathbf{P}}(z)$, whereas for $\mathbf{P}(z)$ we only have consistency in amount M_B of (blind) data samples $\mathbf{y}(k)$ (the noiseless uncorrelated symbols case with finite amount of data is similar to a colored symbols case). Note that the partial and full denoising approaches correspond to resp. the first and second subspace estimates in [6]. Let $\mathbf{h}_i = [\mathbf{h}_i^T(0)\cdots\mathbf{h}_i^T(N_i-1)]^T = \mathbf{H}^{tT}_{i,N_i}$ where ^{t} denotes transposition of the block entries, and $\mathbf{h} = \mathbf{H}^{tT}_N$. Then a stretch of Rx signal $\mathbf{Y}$ can be written as

$$\mathbf{Y}_M = \mathcal{T}(\mathbf{h})\,\mathbf{A} + \mathbf{V}_M = \mathcal{A}\,\mathbf{h} + \mathbf{V}_M$$

where $\mathcal{T}(\mathbf{h}) = [\mathcal{T}_M(\mathbf{H}_{1,N_1})\cdots\mathcal{T}_M(\mathbf{H}_{p,N_p})]$ and $\mathcal{T}_M(\mathbf{H})$ denotes a block Toeplitz convolution matrix with M block rows and $[\mathbf{H}\ 0\cdots 0]$ as first block row. $\mathcal{A}$ is a structured matrix containing the multi-source symbols. Let TS denote the number of training sequence (TS) symbols per source (considered equal for all sources for most of what follows). The TSs for the differents users are considered to be simultaneous initially.

III. Deterministic Semi-Blind (DSB) Approach

In the semiblind approaches, we shall seek a channel estimate $\widehat{\mathbf{h}}$ with possibly overestimated channel lengths $\widehat{N}_i \geq N_i$ and we shall assume that $\widehat{N}_1$ remains the largest $\widehat{N}_i$. In the deterministic symbols setting, we shall work with $\overline{\mathbf{P}}$. $\overline{\mathbf{P}}(z)\,\widehat{\mathbf{H}}_i(z) = 0$ can be written in the time domain as $\mathcal{T}^T_{\widehat{N}_i}(\overline{\mathbf{P}}^t)\,\widehat{\mathbf{h}}_i = 0$. Let $\overline{\mathcal{B}} = \bigoplus_{i=1}^{p} \mathcal{T}^T_{\widehat{N}_i}(\overline{\mathbf{P}}^t)$ where $\bigoplus_{i=1}^{p} A_i = \text{blockdiag}\{A_1,\ldots,A_p\}$. We can now formulate a semiblind criterion as

$$\min_{\widehat{\mathbf{h}}} \left\{ \left\| \mathbf{Y}_{TS} - \mathcal{A}_{TS}\,\widehat{\mathbf{h}} \right\|^2 + \alpha \left\| \overline{\mathcal{B}}\,\widehat{\mathbf{h}} \right\|^2 \right\} \quad (6)$$

where α is a weighting factor, and Y_{TS} is the portion of Rx signal containing only training symbols. A more optimal approach introduces weighting involving the covariance matrix $\overline{C}$ of $\overline{\mathcal{B}}\,\mathbf{h}$ due to the estimation errors in $\overline{\mathbf{P}}$ and leads to

$$\min_{\widehat{\mathbf{h}}} \left\{ \left\| \mathbf{Y}_{TS} - \mathcal{A}_{TS}\,\widehat{\mathbf{h}} \right\|^2 + \sigma^2_v\, \widehat{\mathbf{h}}^H \overline{\mathcal{B}}^H\, \overline{C}^{\#}\, \overline{\mathcal{B}}\,\widehat{\mathbf{h}} \right\} \quad (7)$$

where a possible pseudo-inverse can be avoided by using an infinitesimal amount of regularization. Inspired by an approximate expression for $\overline{C}$ given in [2], we have taken $\sigma^2_v\,\overline{C}^{\#} = M_B\, I$ so that (7) reduces to (6) with $\alpha = M_B$.

With overestimated channel lengths, deterministic blind identification leads to an estimate $\widehat{\widehat{\mathbf{H}}}(z) = \mathbf{H}(z)\,\mathbf{S}(z)$ where $p\times p$ $\mathbf{S}(z)$ is also causal and polynomial and the length of $\mathbf{S}_{ij}(z)$ can be shown to be $(\widehat{N}_i - N_j + 1)^+$ where $(x)^+ = \max\{x,0\}$ (this is a generalization of a result in [3] for the case $\widehat{N}_i = N_i$).

As a result, the DSB approach has the following identifiability requirements

$$\sum_{i,j=1}^{p} (\widehat{N}_i - N_j + 1)^+ \leq m(TS - \widehat{N}_1 + 1)\,,\ \sum_{i=1}^{p} (\widehat{N}_i - N_j + 1)^+ \leq \min\{(\min(N_j, m))(TS - \widehat{N}_1 + 1), TS + N_j - \widehat{N}_1\},\ \forall j \quad (8)$$

IV. Gaussian Semi-Blind (GSB) Approach

In the Gaussian case, the blind estimation ambiguity gets reduced to an instantaneous unitary mixture of the sources (which gets even limited to mixtures of subsets of sources with identical channel length N_i). Since $\mathbf{h}(0)$ can only be determined up to an instantaneous mixture, we reduce the exploitation of $\mathbf{P}(z)\,\mathbf{H}(z) = \mathbf{h}(0)$ or $\mathbf{P}(q)\,\mathbf{h}(k) = \mathbf{h}(0)\,\delta_{k0}$ to $\overline{\mathbf{P}}_0\mathbf{h}(0) = 0$ and $\mathbf{P}(q)\,\mathbf{h}(k) = 0,\ k > 0$. We shall call this the reduced Gaussian case, in which all decorrelation is exploited except between symbols at the same time instant. This can be expressed by $\mathcal{B}\,\mathbf{h} = 0$ where $\mathcal{B} = \bigoplus_{i=1}^{p} \begin{bmatrix} \overline{\mathbf{P}}_0\ 0 \\ \overline{\mathcal{T}^T_{\widehat{N}_i}}(\mathbf{P}^t) \end{bmatrix}$ where $\overline{\mathcal{T}^T_{\widehat{N}_i}}(\mathbf{P}^t)$ is $\mathcal{T}^T_{\widehat{N}_i}(\mathbf{P}^t)$ with the first block row removed. The problem of recovering $\mathbf{h}$ from $\overline{\mathcal{T}^T_{\widehat{N}_i}}(\mathbf{P}^t)\,\mathbf{h}_i = 0$ in the SIMO case, with an optimal weighting between the nuller $\overline{\mathbf{P}}(z)$ and the equalizer portions of $\mathbf{P}(z)$ has been addressed in [2] and involves the covariance matrix of $\overline{\mathcal{T}^T_{\widehat{N}_i}}(\mathbf{P}^t)\,\mathbf{h}_i$ (a simple approximation is given also). This allows us to introduce a semi-blind criterion of the form

$$\min_{\mathbf{h}} \left\{ \|\mathbf{Y}_{TS} - \mathcal{A}_{TS}\,\mathbf{h}\|^2 + \sigma^2_v\, \mathbf{h}^H \mathcal{B}^H C^{\#} \mathcal{B}\,\mathbf{h} \right\} . \quad (9)$$

We took $C = M_B \bigoplus_{i=1}^{p}(I_m \oplus ((I_m + \sigma_v^{-2}\sigma^2_{\mathbf{y}}) \otimes I_{N_i-1}))$, inspired by [2]. The (restricted) GSB semiblind approach has the following identifiability requirements

$$\begin{array}{l} p^2 \leq m(TS - \widehat{N}_1 + 1) \\ p \leq \min\{(\min(N_j, m))(TS - \widehat{N}_1 + 1), TS + N_j - \widehat{N}_1\},\ \forall j \end{array} \tag{10}$$

For both semiblind methods, if the amount of blind data becomes very large, then the particular structure of the weighting matrix for the blind part becomes unimportant and the soft-constrained criterion approaches the hard constrained criterion, in which the TS criterion $\|Y_{TS} - \mathcal{A}_{TS}\,\mathbf{h}\|^2$ gets minimized subject to the blind constraints $\overline{\mathcal{B}}\,\widehat{\mathbf{h}} = 0$ or $\mathcal{B}\,\widehat{\mathbf{h}} = 0$.

In practice, σ_v^2 should be overestimated to obtain good denoising. If σ_v^2 gets that much overestimated that its subtraction cuts away a portion of the signal subspace, then this would lead to loss of the blindly identifiable (in a deterministic setting) part of $\mathbf{H}(z)$. However, in a semiblind approach, identifiability gets recovered and if a blindly identifiable portion got excised in this way, this means that it would have resulted in bad blind estimation quality. So even if the denoising gets done in an overzealous fashion and the order of $\mathbf{P}(z)/\overline{\mathbf{P}}(z)$ gets reduced w.r.t. its theoretical order, the resulting $\overline{\mathbf{P}}(z)$ still lies in the noise subspace and satisfies $\overline{\mathbf{P}}(z)\,\mathbf{H}(z) = 0/\mathbf{P}(z)\,\mathbf{H}(z) \approx \mathbf{h}(0)$ (though in that case this would not allow identification of the blindly identifiable part of $\mathbf{H}(z)$). So in this way, the badly blindly identifiable parameters also get estimated through the TS.

V. Augmented Training-Sequence Part

So far (classical TS approach) Y_{TS} denoted the Rx samples in which only TS symbols appear. In an augmented TS approach, Y_{TS} shall collect all Rx samples in which at least one TS symbol appears. In that case we can write $\mathbf{Y}_{TS} - \mathbf{V} = \mathcal{T}(\mathbf{h})\mathbf{A} = \mathcal{T}_K\mathbf{A}_K + \mathcal{T}_U\mathbf{A}_U$ in which $\mathbf{A}_{K/U}$ collect the known/unknown symbols and $\mathcal{T}_{K/U}$ the corresponding columns of $\mathcal{T}$. The TS part of the semiblind criteria becomes

$$(\mathbf{Y}_{TS} - \mathcal{A}_K\,\mathbf{h})^H(I + \frac{\sigma_a^2}{\sigma_v^2}\mathcal{T}_U\mathcal{T}_U^H)^{-1}(\mathbf{Y}_{TS} - \mathcal{A}_K\,\mathbf{h})\,. \tag{11}$$

Due to the parameter-dependent weighting, the semiblind criteria now require at least one iteration. In the Gaussian approach, the weighting can be determined blindly (and hence consistently). Identifiability conditions for the augmented approaches:

$$\begin{array}{l} \sum_{i,j=1}^{p}(\widehat{N}_i - N_j + 1)^+ \leq m(TS + \widehat{N}_1 - 1) \\ \sum_{i=1}^{p}(\widehat{N}_i - N_j + 1)^+ \leq TS - N_j + \widehat{N}_1\,,\ \forall j \end{array} \tag{12}$$

for DSBA, weheras for GSBA

$$\begin{array}{l} p^2 \leq m(TS + \widehat{N}_1 - 1) \\ p \leq TS - N_j + \widehat{N}_1\,,\ \forall j \end{array} \tag{13}$$

The augmented approach also allows us to handle the user-wise grouped TS approach (Y_{TS} contains TS symbols from only one user at a time) and the distributed TS approach (Y_{TS} contains only one TS symbol from any user at a time). The identifiability conditions in these cases reduce to having at least one TS symbol for every user.

VI. Flat Channel Case

For this case ($N_i = 1$), we propose a Gaussian semiblind approach which exploits more completely the uncorrelated symbols assumption. In this case, $\mathbf{h} = \mathbf{h}(0)$ is identifiable up to an instantaneous mixture (p^2 complex parameters) using the methods above. By exploiting the decorrelation between users, the mixture matrix can be reduced to a unitary matrix ($p^2/2$ parameters). We shall consider the limiting case in which $M_B \to \infty$. In this approach, the blind part is used to determine the blindly identifiable parameters while the TS is used only to determine the remaining parameters. We get

$$\widehat{R}_{\mathbf{YY}} = V\Lambda V^H = V\Lambda^{1/2}\Lambda^{1/2}V^H = \widehat{\widehat{\mathbf{h}}}\,\widehat{\widehat{\mathbf{h}}}^H = \sigma_a^2\widehat{\mathbf{h}}\,\widehat{\mathbf{h}}^H \tag{14}$$

where the eigendecomposition is limited to rank p, and the semiblind estimate $\widehat{\mathbf{h}} = \sigma_a^{-1}\widehat{\widehat{\mathbf{h}}}\,Q$ where $\widehat{\widehat{\mathbf{h}}} = V\Lambda^{1/2}$ is the blind estimate and $QQ^H = I_p$. The TS part gives (time index now denoted as subscript)

$$\begin{array}{l} \sum_k \|\mathbf{y}_k - \widehat{\mathbf{h}}\,\mathbf{a}_k\|^2 \ \rightarrow\ \sum_k \|\mathbf{y}_k^{'} - \sigma_a^{-1}\Lambda^{1/2}Q\,\mathbf{a}_k\|^2 \ = \\ \sum_k \|\mathbf{y}_k^{'}\|^2 - \frac{2}{\sigma_a}\Re\left\{\sum_k \mathbf{y}_k^{'H}\Lambda^{1/2}Q\mathbf{a}_k\right\} + \sigma_a^{-2}\sum_k \mathbf{a}_k^H Q^H\Lambda Q\mathbf{a}_k \end{array} \tag{15}$$

where $\mathbf{y}_k^{'} = V^H\mathbf{y}_k$. Assuming $\sum_k \mathbf{a}_k\mathbf{a}_k^H \sim I_p$, either exactly by TS design or approximately for long enough TS, the TS problem becomes $\max_Q \Re\mathrm{tr}\{Q\,A\}$ where $A = \left(\sum_k \mathbf{a}_k\mathbf{y}_k^H\right)V\Lambda^{1/2}$ has SVD $A = U\Sigma W^H$. Then with $Q^{'} = W^H QU$ which is also unitary, we get $\max_{Q^{'}} \Re\mathrm{tr}\left\{Q^{'}\Sigma\right\} = \max_{Q^{'}}\sum_{i=1}^{p}\Sigma_{ii}\,\Re\left\{Q_{ii}^{'}\right\}$ which leads to $Q_{ii}^{'} = 1$ and hence $Q^{'} = I_p$. The channel estimate finally becomes $\widehat{\mathbf{h}} = \sigma_a^{-1}V\,\Lambda^{1/2}W\,U^H$. Instead of $V\Lambda^{1/2}$, any $\widehat{R}^{1/2}_{\mathbf{YY}}$ could be used.

VII. Simulations

We consider the following scenarii:

scen.	(N_1, N_2)	$(\widehat{N}_1, \widehat{N}_2)$	TS	M_B
1	(3,1)	(3,1)	12	300
2	(3,1)	(3,3)	12	300
3	(3,1)	(3,3)	11	300
4	(4,1)	(4,1)	11	300
5	(3,1)	(3,3)	5	300

For the first two scenarii we used partial denoising, whereas for the last three we used full denoising. We compare the classical TS approach with the DSB/GSB/DSBA/GSBA approaches and with an “exact” version (e.g. DSBe) in which the blind quantities ($\mathbf{P}(z)$) are determined from an exact R_{YY} ($M_B = \infty$). We see that the semiblind approaches offer significant improvements over TS, especially using the augmented TS part. The performance of the deterministic approaches gets close to that of their exact versions, but not for the Gaussian approaches, which should yield better performance. For the curves in Fig. 5 that stay flat, the identifiability conditions are not satisfied.

References

[1] D.T.M. Slock. “Blind Joint Equalization of Multiple Synchronous Mobile Users Using Oversampling and/or Multiple Antennas”. In *Proc. 28th Asilomar Conf. on Signals, Systems and Computers*, pages 1154–1158, Pacific Grove, CA, Oct. 31 - Nov. 2 1994.

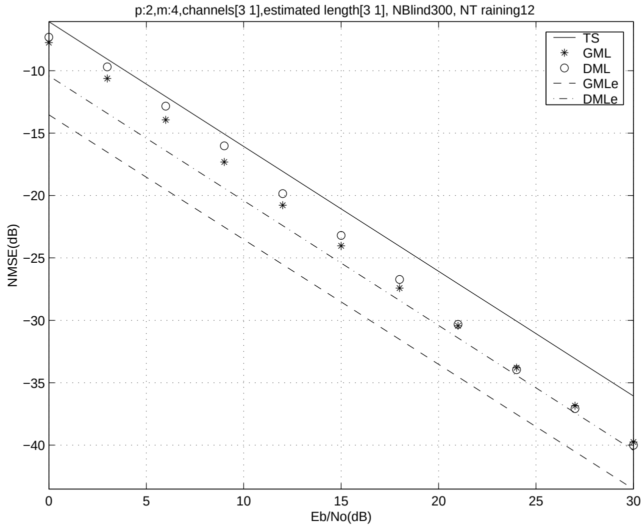

Figure 1: Normalized semiblind channel estimation MSE, scenario 1.

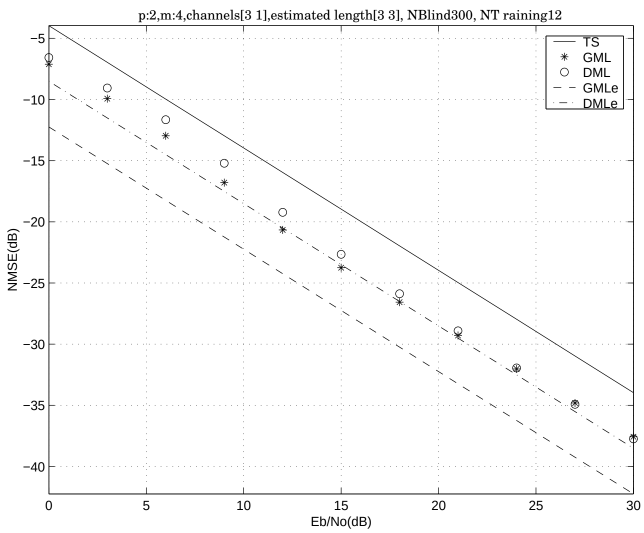

Figure 2: Normalized semiblind channel estimation MSE, scenario 2.

[2] A. Gorokhov and Ph. Loubaton. "Blind Identification of MIMO-FIR Systems: a Generalized Linear Prediction Approach". *EURASIP Signal Processing*, 73:105–124, 1999.

[3] A. Gorokhov and Ph. Loubaton. "Subspace based techniques for second-order blind separation of convolutive mixtures with temporally correlated sources". *IEEE Trans. Circuits and Systems*, 44:813–820, Sept. 1997.

[4] E. de Carvalho and D.T.M. Slock. "Semi-Blind Methods for FIR Multichannel Estimation". In Y. Hua G.B. Giannakis, P. Stoica and L. Tong, editors, *Signal Processing Advances in Communications*, volume 1: Trends in Channel Estimation and Equalization. Prentice Hall, 2000.

[5] E. de Carvalho, L. Deneire, and D.T.M. Slock. "Blind and Semi-Blind Maximum-Likelihood Techniques for Multi-user Multichannel Identification". In *European Signal Processing Conf. (EUSIPCO)*, Rhodes, Greece, Sept. 1998.

[6] J. Götze and A. van der Veen. "On-Line Subspace Estimation Using a Schur-Type Method". *IEEE Trans. Signal Processing*, SP-44(6):1585–1589, June 1996.

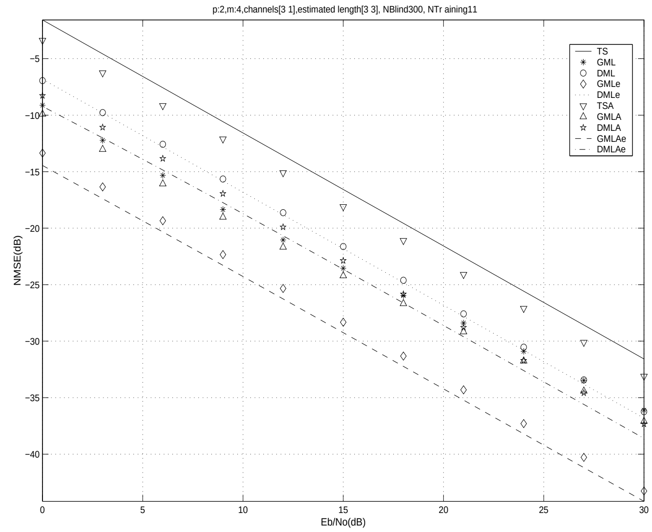

Figure 3: Normalized semiblind channel estimation MSE, scenario 3.

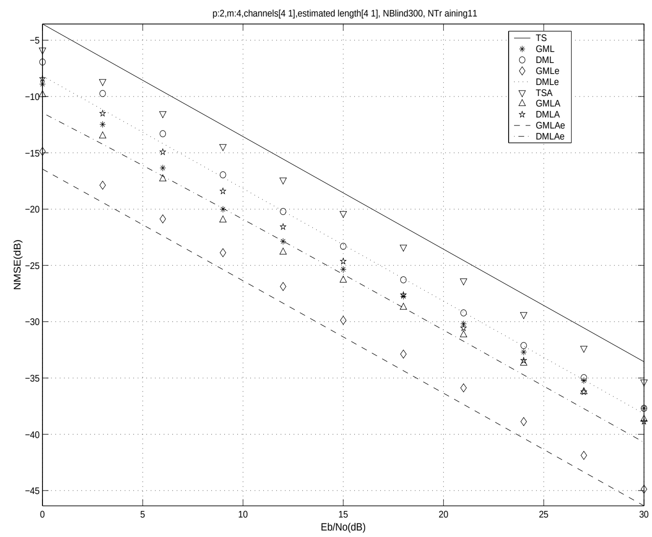

Figure 4: Normalized semiblind channel estimation MSE, scenario 4.

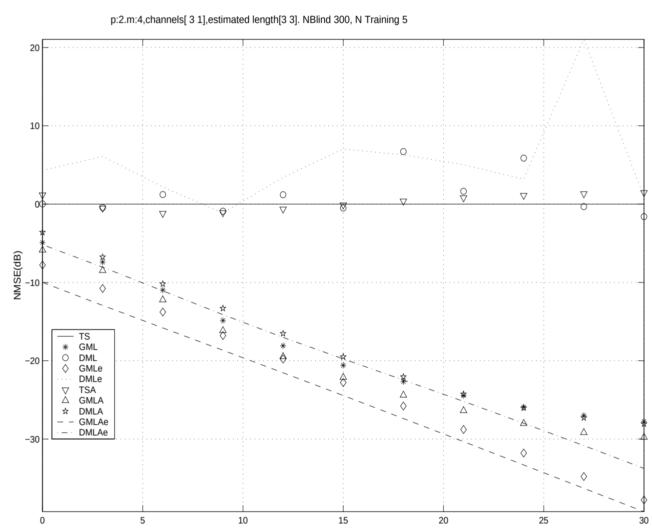

Figure 5: Normalized semiblind channel estimation MSE, scenario 5.