BLIND MULTICHANNEL IDENTIFICATION IN THE STATIONARY COLORED NOISE CASE

Jaouhar Ayadi and *Dirk T.M. Slock*

Institut EURECOM,
B.P. 193, 06904 Sophia Antipolis Cedex, FRANCE
{ayadi, slock}@eurecom.fr

ABSTRACT

We investigate a new approach to identify a multichannel blindly when the additive noise is no longer considered as white but only stationary. We consider the case when the multiple FIR channels are obtained from an array of n antennas and from oversampling the received signal with a factor m. In this case the covariance matrix of the stationary noise is block Toeplitz with $n \times n$ blocks, whereas the covariance matrices of the signal part and the total received signal are block Toeplitz with $mn \times mn$ blocks. In this paper, we shall mainly concentrate on a subspace method based on an appropriate displacement of the covariance matrix in which the noise contribution disappears. The technique developed appears to give acceptable performance for the channel estimate, compared to the Cramer-Rao bound. Furthermore, since the proposed method is based on a parameterization in terms of the channel impulse reponse, prior knowledge of the transmission filter can easily be incorporated [4]. The method mentioned previously can be extended to exploit this knowledge and hence its performance can be improved. Simulations are presented to illustrate the performance of the method.

1. INTRODUCTION

Blind single-user multichannel identification techniques exploit a multichannel formulation corresponding to a Single Input Multiple Output (SIMO) vector channel. The channel is assumed to have a finite delay spread NT. The multiple FIR channels can be obtained by oversampling a single received signal, but can also be obtained as multiple received signals from an array of antennas (in the context of mobile digital communications [1],[2]) or from a combination of both. We consider in the sequel an oversampling factor m and n antennas. For these mn channels the discrete-time input-output relationship can be written as:

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

where $\boldsymbol{y}(k) = [y_1^H(k) \cdots y_{mn}^H(k)]^H$, $\boldsymbol{h}(i) = \left[h_1^H(i) \cdots h_{mn}^H(i)\right]^H$, $\boldsymbol{v}(k) = [v_1^H(k) \cdots v_{mn}^H(k)]^H$, $\boldsymbol{H} = [\boldsymbol{h}(N-1) \cdots \boldsymbol{h}(0)]$, $A_N(k) = \left[a(k-N+1)^H \cdots a(k)^H\right]^H$ and superscript ^{H} denotes Hermitian transpose. Let $\mathbf{H}(z) = \sum_{i=0}^{N-1} \boldsymbol{h}(i) z^{-i} = [\mathrm{H}_1^H(z) \cdots \mathrm{H}_{mn}^H(z)]^H$ be the SIMO channel transfer function. Consider the symbols i.i.d. if required and additive independant (Gaussian if required) noise $\boldsymbol{v}(k)$ with $r_{\boldsymbol{vv}}(k-i) = \mathrm{E}\, \boldsymbol{v}(k)\boldsymbol{v}(i)^H$. Note that due to stationarity the $r_{\boldsymbol{vv}}(k)$ are block Toeplitz with $n \times n$ blocks. Assume we receive M samples:

$$\boldsymbol{Y}_M(k) \;=\; \mathcal{T}_M(\boldsymbol{h})\, A_{M+N-1}(k) + \boldsymbol{V}_M(k) \qquad (2)$$

where $\boldsymbol{Y}_M(k) = [\boldsymbol{y}^H(k-M+1) \cdots \boldsymbol{y}^H(k)]^H$ and similarly for $\boldsymbol{V}_M(k)$. $\mathcal{T}_M(\boldsymbol{h})$ is a block Toeplitz matrix filled out with the channel coefficients grouped in $\boldsymbol{h} = [\boldsymbol{h}^H(N-1) \;\cdots\; \boldsymbol{h}^H(0)]^H$.

We shall simplify the notation in (2) with $k = M-1$ to

$$\boldsymbol{Y} \;=\; \mathcal{T}(\boldsymbol{h})\, A + \boldsymbol{V}\ . \qquad (3)$$

We assume that $mnM > M+N-1$ in which case the channel convolution matrix $\mathcal{T}(\boldsymbol{h})$ has more rows than columns. If the $\mathrm{H}_i(z)$, $i = 1, \ldots, mn$ have no zeros in common (the channel is said irreducible), then $\mathcal{T}(\boldsymbol{h})$ has full column rank (which we will henceforth assume). For obvious reasons, the column space of $\mathcal{T}(\boldsymbol{h})$ is called the signal subspace and its orthogonal complement the noise subspace. The signal subspace is parameterized linearly by $\boldsymbol{h}$.

The case of DOA estimation in colored noise was analyzed in [3], where some methods and performance bounds are presented to estimate the directions of signal sources. In the sequel we present a subspace based approach to estimate the channel in a mobile communications context where the additive noise is assumed to be stationary.

2. THE IDENTIFICATION METHOD

The main idea is to observe that the covariance matrix R_{VV} of the colored noise is block Toeplitz with blocks of size $n \times n$. On the other hand, $\mathcal{T}(\boldsymbol{h})R_{AA}\mathcal{T}^H(\boldsymbol{h})$ is block Toeplitz with blocks of size $mn \times mn$ (assuming the transmitted symbols a_k to be stationary so that R_{AA} is Toeplitz). So

$$R_{YY} \;=\; \mathcal{T}(\boldsymbol{h})R_{AA}\mathcal{T}(\boldsymbol{h})^H + R_{VV} \qquad (4)$$

is block Toeplitz with $mn \times mn$ blocks. However, we shall consider now the displacement of R_{YY} at the level of $n \times n$ blocks. This is done by extracting two $(mnM-n) \times (mnM-n)$ submatrices $\underline{R}_{YY}$ and $\overline{R}_{YY}$ defined as follows: $\underline{R}_{YY}$ and $\overline{R}_{YY}$ are submatrices of R_{YY} with the last, resp. first, n rows and columns removed (see Fig. 1). $\underline{R}_{VV}$ and $\overline{R}_{VV}$ are defined similarly w.r.t. R_{VV}. Note that $\underline{R}_{VV} = \overline{R}_{VV}$.

The first matrix $\underline{R}_{YY}$ can be written as

$$\underline{R}_{YY} = \underline{\mathcal{T}}(\boldsymbol{h})\, R_{AA}\, \underline{\mathcal{T}}(\boldsymbol{h})^H + \underline{R}_{VV}, \qquad (5)$$

where $\underline{\mathcal{T}}(\boldsymbol{h})$ corresponds to the matrix $\mathcal{T}(\boldsymbol{h})$ from which we have omitted the n last rows. Similarly, by considering $\overline{\mathcal{T}}(\boldsymbol{h})$ as the

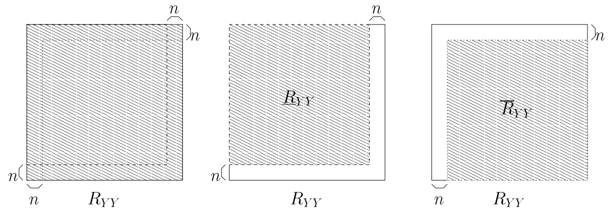

Figure 1: R_{YY}, $\underline{R}_{YY}$ and $\overline{R}_{YY}$

matrix $\mathcal{T}(\boldsymbol{h})$ from which we omit the first n rows, $\overline{R}_{YY}$ can be written as

$$\overline{R}_{YY} = \overline{\mathcal{T}}(\boldsymbol{h})\, R_{AA}\, \overline{\mathcal{T}}(\boldsymbol{h})^H + \overline{R}_{VV}\,. \tag{6}$$

Hence the displacement $\underline{R}_{YY} - \overline{R}_{YY}$ is given by

$$\underline{R}_{YY} - \overline{R}_{YY} = \underline{\mathcal{T}}(\boldsymbol{h})\, R_{AA}\, \underline{\mathcal{T}}(\boldsymbol{h})^H - \overline{\mathcal{T}}(\boldsymbol{h})\, R_{AA}\, \overline{\mathcal{T}}(\boldsymbol{h})^H\,, \tag{7}$$

which is parameterized by the channel $\boldsymbol{h}$. Consider now the eigendecomposition of the matrix $\underline{R}_{YY} - \overline{R}_{YY}$ in which the eigenvalues are ordered in descending order:

$$\underline{R}_{YY} - \overline{R}_{YY} = \sum_{i=1}^{mnM-n} \lambda_i V_i V_i^H = \mathcal{V}_+ \Lambda_+ \mathcal{V}_+^H \;+\; \mathcal{V}_- \Lambda_- \mathcal{V}_-^H, \tag{8}$$

where Λ_+, resp. Λ_-, denote the set of positive, resp. negative, eigenvalues of $\underline{R}_{YY} - \overline{R}_{YY}$, and $\mathcal{V}_+$, resp. $\mathcal{V}_-$, their corresponding eigenvectors. In what follows, we assume R_{AA} to have full rank and the number of channels to be at least three: $mn \geq 3$. In that case, $\underline{R}_{YY} - \overline{R}_{YY}$ has a noise subspace for M sufficiently large, the positive signal subspace dimension (size of $\mathcal{V}_+$) is equal to the full column rank of $\underline{\mathcal{T}}(\boldsymbol{h})$, while the negative signal subspace dimension (size of $\mathcal{V}_-$) is equal to the full column rank of $\overline{\mathcal{T}}(\boldsymbol{h})$, and both positive and negative signal subspaces have the same dimension. The column space of $\mathcal{V}_+$ can be interpreted as the components of the columnspace of $\underline{\mathcal{T}}(\boldsymbol{h})$ that are orthogonal to $\mathcal{V}_-$ (and similarly for $\mathcal{V}_-$ and $\overline{\mathcal{T}}(\boldsymbol{h})$). This leads to the following equivalences:

$$\begin{aligned} Range\,\{\mathcal{V}_+\} &= Range\,\left\{P^{\perp}_{V_-}\,\underline{\mathcal{T}}(\boldsymbol{h})\right\}, \\ Range\,\{\mathcal{V}_-\} &= Range\,\left\{P^{\perp}_{V_+}\,\overline{\mathcal{T}}(\boldsymbol{h})\right\}, \end{aligned} \tag{9}$$

where $P_X^{\perp} = I - X\left(X^H X\right)^+ X^H$ denotes the projection operator on the orthogonal complement of $Range\,\{X\}$. A natural signal subspace fitting criterion can now be formulated as follows:

$$\min_{\boldsymbol{h}, T_1, T_2} \|P^{\perp}_{V_-}\,\underline{\mathcal{T}}(\boldsymbol{h}) - \mathcal{V}_+ T_1\|_F^2 + \|P^{\perp}_{V_+}\,\overline{\mathcal{T}}(\boldsymbol{h}) - \mathcal{V}_- T_2\|_F^2, \tag{10}$$

where the Frobenius norm of a matrix Z can be defined in terms of the trace operator $\|Z\|_F^2 = \mathrm{tr}\left\{Z^H Z\right\}$. After minimization w.r.t. T_1,T_2, the problem in (10) boils down to

$$\begin{aligned} \min_{\|\boldsymbol{h}\|=1} &\ \mathrm{tr}\left\{\underline{\mathcal{T}}^H(\boldsymbol{h}) P^{\perp}_{V_-} P^{\perp}_{V_+} P^{\perp}_{V_-}\,\underline{\mathcal{T}}(\boldsymbol{h})\right\} \\ &+ \mathrm{tr}\left\{\overline{\mathcal{T}}^H(\boldsymbol{h}) P^{\perp}_{V_+} P^{\perp}_{V_-} P^{\perp}_{V_+}\,\overline{\mathcal{T}}(\boldsymbol{h})\right\} \end{aligned} \tag{11}$$

which can be simplified to

$$\begin{aligned} \min_{\|\boldsymbol{h}\|=1} &\ \mathrm{tr}\left\{\underline{\mathcal{T}}^H(\boldsymbol{h}) P_{V_\mathcal{N}}\underline{\mathcal{T}}(\boldsymbol{h})\right\} + \mathrm{tr}\left\{\overline{\mathcal{T}}^H(\boldsymbol{h}) P_{V_\mathcal{N}}\overline{\mathcal{T}}(\boldsymbol{h})\right\} \\ &= \min_{\|\boldsymbol{h}\|=1} \boldsymbol{h}^H\left(A_1 + A_2\right)\boldsymbol{h}, \end{aligned} \tag{12}$$

where $V_\mathcal{N}$ are the noise subspace eigenvectors ($P_{V_\mathcal{N}} = I - P_{V_+} - P_{V_-}$), and the matrices A_1, A_2 can be determined from $P_{V_\mathcal{N}}$ and $\underline{\mathcal{T}}(\boldsymbol{h})$, $\overline{\mathcal{T}}(\boldsymbol{h})$. The solution for $\boldsymbol{h}$ is the eigenvector correponding to the minimum eigenvalue of $A_1 + A_2$. In the following, we will discuss the Cramer-Rao bounds in both Gaussian and deterministic contexts.

3. CRAMER-RAO BOUNDS

Let θ be the complex parameter vector to be estimated, and $\theta_R = \left[\begin{array}{cc}\mathrm{Re}(\theta)^H & \mathrm{Im}(\theta)^H\end{array}\right]^H$ the associated real parameters. The real Fisher Information Matrix (FIM) associated to θ_R is:

$$J_{\theta_R\theta_R} = E_{\boldsymbol{Y}|\theta_R}\left(\frac{\partial \ln f(\boldsymbol{Y}|\theta_R)}{\partial\theta_R}\right)\left(\frac{\partial \ln f(\boldsymbol{Y}|\theta_R)}{\partial\theta_R}\right)^T, \tag{13}$$

$\boldsymbol{Y}$ are the observations and $f(\boldsymbol{Y}|\theta_R)$ is their probability density function. Let $\hat{\theta}_R$ be an unbiased parameter estimate, $\tilde{\theta}_R = \theta_R - \hat{\theta}_R$ the estimation error and $C_{\tilde{\theta}_R\tilde{\theta}_R} = E\tilde{\theta}_R\tilde{\theta}_R^H$ the error covariance matrix. $J^{-1}_{\theta_R\theta_R}$ is the Cramer-Rao Bound:

$$C_{\tilde{\theta}_R\tilde{\theta}_R} \geq CRB_{\hat{\theta}_R} = J^{-1}_{\theta_R\theta_R}. \tag{14}$$

As we work with complex quantities, it may be better 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 $\theta = \alpha + j\beta$. The FIM $J_{\varphi\psi}$ for complex parameters φ, ψ (parts of θ) is defined as:

$$J_{\varphi\psi} = E_{Y|\theta}\left(\frac{\partial \ln f(Y|\theta)}{\partial\varphi^*}\right)\left(\frac{\partial \ln f(Y|\theta)}{\partial\psi^*}\right)^H. \tag{15}$$

Let us introduce the extended complex parameter vector $\theta_C = [\theta^T\ \theta^H]^T$. Then $J_{\theta_C\theta_C}$ contains the same information as $J_{\theta_R\theta_R}$.

If $J_{\theta\theta^*} = 0$, the matrix $J_{\theta\theta}$ can be considered as a complex FIM, and the covariance matrix of the unbiased estimation error $\tilde{\theta} = \theta - \hat{\theta}$ is $C_{\tilde{\theta}\tilde{\theta}} \geq J^{-1}_{\theta\theta}$, the complex CRB. If $J_{\theta\theta^*} \neq 0$ (as in the cases to be considered below), $J^{-1}_{\theta\theta}$ is also a bound on $C_{\tilde{\theta}\tilde{\theta}}$, but not as tight as the actual CRB = $C_{\tilde{\theta}_C\tilde{\theta}_C} \geq J^{-1}_{\theta_C\theta_C}$. The quantity we are usually interested in is the MSE =

$$E\|\tilde{\theta}\|^2 = E\|\tilde{\theta}_R\|^2 = \frac{1}{2}E\|\tilde{\theta}_C\|^2 \geq \frac{1}{2}\,\mathrm{tr}\{J^{-1}_{\theta_C\theta_C}\} \geq \mathrm{tr}\{J^{-1}_{\theta\theta}\}. \tag{16}$$

In the blind channel estimation application considered here, an identifiability problem arises since the channel can only be identified up to a scalar multiple. This leads to singularity of the FIM. For the computation of CRBs, we replace the inverses of FIMs $J^{-1}_{\theta_C\theta_C}$ by Moore-Penrose pseudo-inverses $J^{+}_{\theta_C\theta_C}$. For the resulting inverse still to be a valid CRB, the unidentifiable channel factor should be adjusted in a particular fashion that will be explained in the simulations section (see also [5]).

3.1. The Gaussian case

In the circular Gaussian symbols case, $\boldsymbol{Y} \sim \mathcal{N}(0, R_{YY})$ with $R_{YY} = \sigma_a^2\mathcal{T}(\boldsymbol{h})\mathcal{T}^H(\boldsymbol{h}) + R_{VV}$ (so $R_{AA} = \sigma_a^2 I$). The negative log likelihood to be minimized is

$$\mathcal{L}(\boldsymbol{h}, R_{VV}) = c^t + \ln\det R_{YY} + \boldsymbol{Y}^H R_{YY}^{-1}\boldsymbol{Y}. \tag{17}$$

We have to estimate jointly the channel coefficients and the noise covariance coefficients. Since the covariance matrix R_{VV} of the colored noise is block Toeplitz with blocks of size $n \times n$, its parameters to be estimated are: the elements of the lower or upper triangular part of $r_{\boldsymbol{vv}}(0)$ and the elements of the $n \times n$ matrices $r_{\boldsymbol{vv}}(1)...r_{\boldsymbol{vv}}(L-1)$, where L denotes the length of the FIR filter used to generate the MA colored noise. One can notice that the diagonal elements of $r_{\boldsymbol{vv}}(0)$ are real and hence we have a mix of real and complex parameters. We consider the following parameter vector $\theta = \left[\theta_1^H \;\; \theta_2^H\right]^H$, in which θ_1 denotes the vector obtained by stacking the diagonal elements of $r_{\boldsymbol{vv}}(0)$ and θ_2 denotes the vector obtained by concatenating the channel coefficients $\boldsymbol{h}$ and the vector formed by stacking the colums of the matrix $\left[\; \underline{r}_{\boldsymbol{vv}}^H(0) \quad r_{\boldsymbol{vv}}^H(1) \quad \cdots \quad r_{\boldsymbol{vv}}^H(L-1)\;\right]^H$, where $\underline{r}_{\boldsymbol{vv}}(0)$ is the strict lower triangular part of $r_{\boldsymbol{vv}}(0)$. Let $\theta_C = \left[\theta_1^T \;\; \theta_2^T \;\; \theta_2^H\right]^T$ be the extended complex parameter vector. We get for the FIM $J_{\theta_C\theta_C}$

$$J_{\theta_C\theta_C} = \begin{bmatrix} J_{\theta_1\theta_1} & J_{\theta_1\theta_2} & J_{\theta_1\theta_2^*} \\ J_{\theta_2\theta_1} & J_{\theta_2\theta_2} & J_{\theta_2\theta_2^*} \\ J_{\theta_2^*\theta_1} & J_{\theta_2^*\theta_2} & J_{\theta_2^*\theta_2^*} \end{bmatrix} \tag{18}$$

where $J_{\varphi\psi}$ is given by (15) and $J_{\varphi^*\psi^*} = J_{\varphi\psi}^*$. Exploiting this information and the Hermitian symmetry of $J_{\theta_C\theta_C}$, one has to compute only $J_{\theta_1\theta_1}$, $J_{\theta_1\theta_2}$, $J_{\theta_2\theta_2}$ and $J_{\theta_2\theta_2^*}$. We are interested here in the MSE on the channel estimates $\widehat{\boldsymbol{h}}$. Let $\mathcal{P}$ be a permutation matrix such that $\mathcal{P}\theta_C = [\boldsymbol{h}_C^T \;\; \eta_C^T]^T$ where η_C represents the other (nuisance) parameters. Then

$$\mathcal{P} J_{\theta_C\theta_C} \mathcal{P}^H = \begin{bmatrix} J_{\boldsymbol{h}_C\boldsymbol{h}_C} & J_{\boldsymbol{h}_C\eta_C} \\ J_{\eta_C\boldsymbol{h}_C} & J_{\eta_C\eta_C} \end{bmatrix}. \tag{19}$$

So we get the following CRB

$$E\|\tilde{\boldsymbol{h}}\|^2 = \frac{1}{2}E\|\tilde{\boldsymbol{h}}_C\|^2 \geq \frac{1}{2}\,\mathrm{tr}\{(J_{\boldsymbol{h}_C\boldsymbol{h}_C} - J_{\boldsymbol{h}_C\eta_C} J_{\eta_C\eta_C}^{-1} J_{\eta_C\boldsymbol{h}_C})^{+}\}. \tag{20}$$

3.2. The deterministic case

In the deterministic model, both the channel $\boldsymbol{h}$ and the symbols A are considered as deterministic quantities. The complex parameter vector θ is: $\theta = \left[\; \boldsymbol{r}^H \quad \boldsymbol{h}^H \quad A^H \;\right]^H$. A contains the input symbols, $\boldsymbol{h}$ the channel coefficients and $\boldsymbol{r}$ the colored noise covariance coefficients to be estimated. The complex probability density function is:

$$f(\boldsymbol{Y}|\theta) = \frac{1}{\pi^{mM}\det R_{VV}} e^{-\left[\boldsymbol{Y}-\boldsymbol{Y}^{(S)}\right]^H R_{VV}^{-1}\left[\boldsymbol{Y}-\boldsymbol{Y}^{(S)}\right]} \tag{21}$$

where $\boldsymbol{Y}^{(S)} = \mathcal{T}(\boldsymbol{h})A$ is the signal part of $\boldsymbol{Y}$. The negative log likelihood to be minimized is

$$\mathcal{L}(\theta) = c^t + \ln\det R_{VV} + (\boldsymbol{Y} - \mathcal{T}(\boldsymbol{h})A)^H R_{VV}^{-1} (\boldsymbol{Y} - \mathcal{T}(\boldsymbol{h})A)\,. \tag{22}$$

Following the same reasoning as for the Gaussian case, we can rewrite θ as $\theta = \left[\; \theta_1^H \quad \theta_2^H \;\right]^H$, in which θ_1 denotes the vector obtained by stacking the diagonal elements of $r_{\boldsymbol{vv}}(0)$ and θ_2 denotes the vector obtained by concatenating the symbols A, the channel coefficients $\boldsymbol{h}$ and the complex correlations of the colored noise. We consider again the FIM associated to the extended complex parameter vector θ_C. $J_{\theta_C\theta_C}$ is again given by (18), where the different matrices $J_{\varphi\psi}$ are computed by deriving $\mathcal{L}(\theta)$ defined in (22). And we are again interested in the MSE on the channel estimates.

4. TRANSMISSION (TX) FILTER KNOWLEDGE

Since the elimination of the stationary noise in the subspace fitting approach outlined above is based on oversampling and hence on the exploitation of excess bandwidth, the use of prior knowledge of the transmission filter (which shapes the excess bandwidth) should be useful. In [4], we adressed the exploitation of the Transmission/Reception (TX/RX) filters knowledge and we presented a set of blind channel estimation methods that we extended to exploit the prior knowledge of the TX/RX filters. We shall review the basic approach. Consider a certain oversampling factor m, and let the oversampled transfer function $\mathrm{H}(z) = \mathrm{C}(z)\mathrm{G}(z)$ of the overall channel be the cascade of the actual oversampled anti-aliasing filtered channel $\mathrm{C}(z)$ and the oversampled combined TX/RX filter $\mathrm{G}(z)$ (the oversampling factor should satisfy the Nyquist criterion for the TX/RX filter). Each of these transfer functions can be decomposed into its polyphase components at the symbol rate, e.g. $\mathrm{H}(z) = \sum_{i=0}^{m-1} z^{-i}\mathrm{H}_i(z^m)$. These components can also be represented in the SIMO form, $\mathbf{G}(z) = [\mathrm{G}_1^H(z)\cdots\mathrm{G}_m^H(z)]^H = \sum_{k=0}^{K-1}\boldsymbol{g}(k)z^{-k}$ and $\mathbf{C}(z) = [\mathrm{C}_1^H(z)\cdots\mathrm{C}_m^H(z)]^H = \sum_{k=0}^{L-1}\boldsymbol{c}(k)z^{-k}$ with $K+L-1 = N$. The relations between the polyphase components can be obtained from

$$\sum_{i=0}^{m-1} z^{-i}\mathrm{H}_i(z^m) = \left(\sum_{k=0}^{m-1} z^{-k}\mathrm{G}_k(z^m)\right)\left(\sum_{l=0}^{m-1} z^{-l}\mathrm{C}_l(z^m)\right) \tag{23}$$

In particular for $m = 2$ we get

$$\begin{aligned} \begin{bmatrix} \mathrm{H}_0(z) \\ \mathrm{H}_1(z) \end{bmatrix} &= \begin{bmatrix} \mathrm{G}_0(z) & z^{-1}\mathrm{G}_1(z) \\ \mathrm{G}_1(z) & \mathrm{G}_0(z) \end{bmatrix} \begin{bmatrix} \mathrm{C}_0(z) \\ \mathrm{C}_1(z) \end{bmatrix} \\ &= \begin{bmatrix} \mathrm{C}_0(z) & z^{-1}\mathrm{C}_1(z) \\ \mathrm{C}_1(z) & \mathrm{C}_0(z) \end{bmatrix} \begin{bmatrix} \mathrm{G}_0(z) \\ \mathrm{G}_1(z) \end{bmatrix} \end{aligned} \tag{24}$$

or $\mathbf{H}(z) = \underline{\mathbf{G}}(z)\mathbf{C}(z) = \underline{\mathbf{C}}(z)\mathbf{G}(z)$. In the time domain, we get

$$\mathcal{T}_M(\boldsymbol{H}) \;=\; \mathcal{T}_M(\underline{\boldsymbol{G}})\mathcal{T}_{M+K-1}(\boldsymbol{C}) \tag{25}$$

where $\mathcal{T}_M(\boldsymbol{X})$ is a block Toeplitz matrix with M block rows and $[\boldsymbol{X} \;\; 0_{p\times(M-1)q}]$ as first block row, $\boldsymbol{X}$ being considered as a block row vector with $p \times q$ blocks, $\boldsymbol{C}$ is similar to $\boldsymbol{H}$ and

$$\underline{\boldsymbol{G}} = \left[\underline{\boldsymbol{g}}(K-1)\cdots\underline{\boldsymbol{g}}(0)\right],\; \underline{\boldsymbol{g}}(k) = \begin{bmatrix} g_0(k) & g_1(k-1) \\ g_1(k) & g_0(k) \end{bmatrix} \tag{26}$$

and we assume $g_1(K-1) = 0$. The relation between $\boldsymbol{h}$ and $\boldsymbol{c}$ is $\boldsymbol{h} = \mathcal{T}_L^T(\underline{\boldsymbol{G}}^t)\boldsymbol{c}$ where ^{t} denotes transposition of the blocks: $\underline{\boldsymbol{G}}^t = \left[\underline{\boldsymbol{g}}^T(K-1)\cdots\underline{\boldsymbol{g}}^T(0)\right]$.

In the case of an array of n antennas, $\mathbf{H}_i(z) = \mathbf{G}(z)\mathbf{C}_i(z)$ for every antenna signal $i = 1\ldots n$, $\mathbf{H}(z) = [\mathbf{H}_1^H(z)\cdots\mathbf{H}_n^H(z)]^H =$ blockdiag$\{\mathbf{G}(z)\cdots\mathbf{G}(z)\}\mathbf{C}(z)$ where now $\mathbf{H}(z)$ and $\mathbf{C}(z)$ regroup mn channels and can be expressed as folllows

$$\mathbf{H}(z) = (I_n \otimes \mathbf{G}(z))\,\mathbf{C}(z), \tag{27}$$

where $\otimes$ denotes the Kronecker product. (25) becomes

$$\mathcal{T}_M(\boldsymbol{H}) = \mathcal{T}_M([I_n \otimes \underline{\boldsymbol{g}}(K-1)\cdots I_n \otimes \underline{\boldsymbol{g}}(0)])\mathcal{T}_{M+K-1}(\boldsymbol{C}). \tag{28}$$

Prior TX/RX filter knowledge gets exploited by expressing $\boldsymbol{h} = \boldsymbol{G}\boldsymbol{c}$ and searching for $\boldsymbol{c}$, where $\boldsymbol{G} = \mathcal{T}_L^T([I_n \otimes \underline{\boldsymbol{g}}^T(K-1)\cdots I_n \otimes$

$\underline{g}^T(0)]$). Since the subspace identification method discussed above is of the form $\min_{\|\boldsymbol{h}\|=1} \boldsymbol{h}^H(A_1+A_2)\boldsymbol{h}$, we get

$$\min_{\boldsymbol{c}} \boldsymbol{c}^H \boldsymbol{G}^H (A_1+A_2)\boldsymbol{G}\boldsymbol{c}, \tag{29}$$

which can be solved under the non-triviality constraint $\|\boldsymbol{c}\| = 1$. The method thus obtained is a channel identification method With TX Filter Knowledge (WTXFK).

5. CRAMER-RAO BOUNDS WTXFK

So WTXFK, we obtain $\widehat{\boldsymbol{h}} = \boldsymbol{G}\widehat{\boldsymbol{c}}$ from $\widehat{\boldsymbol{c}}$. The FIM for $\boldsymbol{c}$ can be obtained from the FIM for the unstructured $\boldsymbol{h}$ which we found earlier. From $\boldsymbol{h} = \boldsymbol{G}\boldsymbol{c}$, we get $\boldsymbol{h}_C = \boldsymbol{G}_C\boldsymbol{c}_C$ where $\boldsymbol{G}_C =$ blockdiag$\{\boldsymbol{G},\boldsymbol{G}^*\}$. Hence $\frac{\partial \ln f(\boldsymbol{Y}|\theta)}{\partial \boldsymbol{c}_C^*} = \frac{\partial \boldsymbol{h}_C^H}{\partial \boldsymbol{c}_C^*}\frac{\partial \ln f(\boldsymbol{Y}|\theta)}{\partial \boldsymbol{h}_C^*} = \boldsymbol{G}_C^H \frac{\partial \ln f(\boldsymbol{Y}|\theta)}{\partial \boldsymbol{h}_C^*}$ which implies e.g. $J_{\boldsymbol{c}_C\varphi} = \boldsymbol{G}_C^H J_{\boldsymbol{h}_C\varphi}$. Then $J_{\boldsymbol{c}_C\boldsymbol{c}_C}$ for $\boldsymbol{c}_C$ with elimination of the nuisance parameters η_C becomes $J_{\boldsymbol{c}_C\boldsymbol{c}_C} = \boldsymbol{G}_C^H(J_{\boldsymbol{h}_C\boldsymbol{h}_C} - J_{\boldsymbol{h}_C\eta_C}J^{-1}_{\eta_C\eta_C}J_{\eta_C\boldsymbol{h}_C})\boldsymbol{G}_C$. where the matrix in the middle is the one appearing in (20). Now, $\boldsymbol{c}$ is only an intermediate quantity in the estimation of $\boldsymbol{h}$. To find the FIM for $\boldsymbol{h}_C$ from the FIM for $\boldsymbol{c}_C$, we get from $\boldsymbol{h}_C = \boldsymbol{G}_C\boldsymbol{c}_C$ that $\boldsymbol{c}_C = \boldsymbol{G}_C^{+}\boldsymbol{h}_C$. As before, we can find $J_{\boldsymbol{h}_C\varphi} = \boldsymbol{G}_C^{+H}J_{\boldsymbol{c}_C\varphi}$. Since $\boldsymbol{G}_C\boldsymbol{G}_C^{+} = P_{\boldsymbol{G}_C}$, we finally get the CRB

$$E\|\tilde{\boldsymbol{h}}\|^2 \geq \frac{1}{2}\,\mathrm{tr}\{[P_{\boldsymbol{G}_C}(J_{\boldsymbol{h}_C\boldsymbol{h}_C} - J_{\boldsymbol{h}_C\eta_C}J^{-1}_{\eta_C\eta_C}J_{\eta_C\boldsymbol{h}_C})P_{\boldsymbol{G}_C}]^{+}\} \tag{30}$$

for the case WTXFK, which should be compared to (20) for the unstructured $\boldsymbol{h}$ case. The pseudo-inverse in (30) indicates that the ambiguity factor gets fixed at the level of $\widehat{\boldsymbol{h}}$.

6. SIMULATION RESULTS

We consider a burst length of 100 symbol periods, a complex channel $\boldsymbol{H}$ randomly generated, of length $N = 4$. The number of antennas is $n = 2$ and the oversampling factor is $m = 2$. The input symbols are drawn from an i.i.d. BPSK sequence. The colored noise is MA generated by filtering a complex white Gaussian $n \times 1$ process with an $n \times n$ FIR $\boldsymbol{f}$ filter of length equal to 3. The SNR is defined as SNR $= \frac{\left(\|\boldsymbol{h}\|^2/mn\right)\sigma_a^2}{\left(\|\boldsymbol{f}\|^2/n\right)\sigma_v^2} = \frac{\|\boldsymbol{h}\|^2\sigma_a^2}{m\|\boldsymbol{f}\|^2\sigma_v^2}$. We use a sample covariance matrix $\widehat{R}_{YY}$ of size $M = 20$. Blind estimation gives a channel estimate $\hat{\hat{\boldsymbol{h}}}$ with $\|\hat{\hat{\boldsymbol{h}}}\| = 1$, we adjust the right scale factor α so that $\boldsymbol{h}_o^H(\alpha\hat{\hat{\boldsymbol{h}}}) = \boldsymbol{h}_o^H\boldsymbol{h}_o$ where $\boldsymbol{h}_o$ is the true channel vector (see [5]): the final estimate is $\widehat{\boldsymbol{h}} = \alpha\hat{\hat{\boldsymbol{h}}}$. The performance measure is the Normalized MSE: NMSE, averaged over 100 Monte-Carlo runs and defined as NMSE $=$ E$\|\boldsymbol{h} - \widehat{\boldsymbol{h}}\|^2/\|\boldsymbol{h}\|^2$. We simulated the previously described subspace fitting (SSF) method and we evaluated its performance. In Fig. 2, we plot the NMSE versus the SNR: it is clear that the method works. On the same figure we plot the normalized Gaussain CRB computed as tr$\{CRB_{\boldsymbol{h}}\}/\|\boldsymbol{h}\|^2$. Whereas the method is a deteministic one, and since the Gaussian CRB is lower than the deterministic one, it can be seen than the NMSE curve is not close to the Gaussian CRB unless the SNR is high. This means that potentially a large gain in performance can be obtained by exploiting the information on the second-order statistics of the symbols. In Fig. 3, we exploit the prior knowledge of the transmission filter and we measure the NMSE obtained by the SSF WTXFK method. We consider a burst length of 200 symbol periods, $n = 1$ antenna, an oversampling factor of $m = 3$. The propagation channel is of length 3 and the transmission filter is a linarized GMSK filter truncated to 4 symbol periods. Our simulation results show that the purely blind SSF identification method suffers from channel zeros that are almost in common (due to the limited excess bandwidth), whereas the SSF identification approach WTXFK performs well.

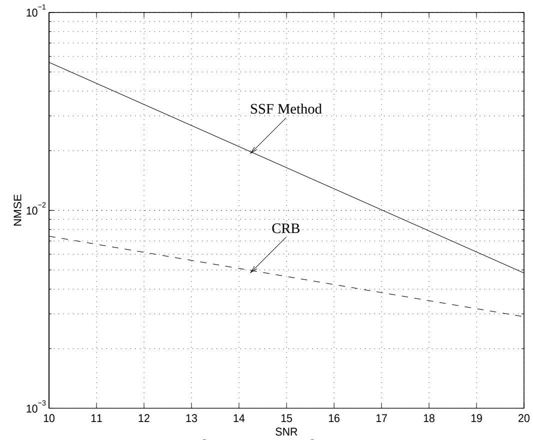

Figure 2: Performance of the SSF method

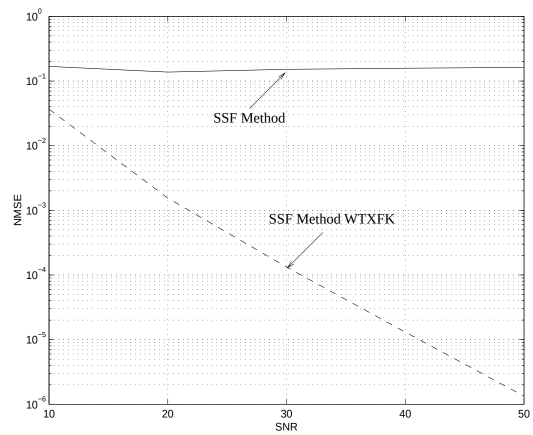

Figure 3: Performance of the SSF WTXFK method

7. REFERENCES

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

[2] D.T.M. Slock and C.B. Papadias. "Blind Fractionally-Spaced Equalization Based on Cyclostationarity". In *Proc. Vehicular Technology Conf.*, Stockholm, Sweden, June 1994.

[3] B. Göransson. *On Parametric Methods for Source Localization*. PhD thesis, KTH, Stockholm, Sweden, 1997.

[4] J. Ayadi and D.T.M. Slock. "Cramer-Rao Bounds and Methods for Knowledge Based Estimation of Multiple FIR Channels". In *Proc. SPAWC Workshop*, Paris, France, Apr. 1997.

[5] E. de Carvalho and D.T.M. Slock. "Cramer-Rao Bounds for Semi-blind, Blind and Training Sequence Based Channel Estimation". In *Proc. SPAWC Conf.*, Paris, France, Apr. 1997.