

BLIND CHANNEL EQUALIZATION USING WEIGHTED SUBSPACE METHODS

Rafael Ruiz and Margarita Cabrera

Dept. of Signal Theory and Communications, E.T.S.I. Telecomunicación, UPC.
C/ Sor Eulàlia de Anzizu, s/n. 08034 Barcelona, SPAIN.
e-mail: rafael@gps.tsc.upc.es

ABSTRACT

This paper addresses the problems of blind channel estimation and symbol detection with second order statistics methods from the received data. It can be shown that this problem is similar to Direction Of Arrival (DOA) estimation, where many solutions like the *MUSIC* algorithm or "weighted" techniques (as *Deterministic Maximum Likelihood* or *Weighted Subspace Fitting* method) have been developed. In this proposal we extend these techniques to blind channel identification problem in an unified framework known as *Subspace Fitting*. In this framework the estimated and the received data are "fitting" through the subspaces in a least square sense. Then, in order to solve this problem and estimate the channel, a modified Gauss-Newton type algorithm is suggested. Simulations are carried out comparing the proposed solutions with a classical signal subspace-based blind channel identification scheme.

1. INTRODUCTION

Digital communications through channels with multipath phenomena are subjected to intersymbol interference. This problem can be so severe that correct reception of the transmitted symbols is not feasible anymore. It is necessary therefore to equalize the channel, which implies to estimate the channel. Since Tong, Xu and Kailath showed in [1] that it is possible to obtain an estimation of the channel from a second order statistic of the received signal, second order statistics have substituted to high order statistics in channel estimation.

Due to its high-resolution capability, many subspace-based methods have been developed in order to recover signal parameters. These methods are based on the singular value decomposition (SVD) of a matrix constructed from the observed signal, which provides a robust discrimination between desired and disturbing signals in terms of signal and noise subspaces. The general objective is to find a low-rank subspace with a shift structure that has minimal distance to the true signal space, or equivalently, that is as orthogonal to the noise subspace as possible. Viberg and Ottersten formulate in [3] different methods in a common subspace fitting based framework, providing an overview of the DOA estimation problem and clarifying the algebraic relations between the algorithms.

This work is partially supported by the National Research Plan of Spain, CICYT, TIC96-0500-C10-01 & TIC98-0412 and by the Generalitat of Catalonia, CIRIT, 1996SGR-00096.

In this contribution we focus on *MUSIC* [4], *WSF* [3] and *MDL* [5] algorithms. These methods are summarized and adapted to blind channel identification in a subspace fitting approach. The paper is organized as follows. In Section 2 data model and its subspace approach are formulated. In Section 3 we extend the subspace-fitting framework proposed by Viberg [3] to blind channel identification and formulate the cost functions whose minimization/maximization will allow estimating the channel. After that, the Gauss-Newton algorithm that solves this problem is developed. In section 4 some simulation results are presented and discussed. Finally, we present some conclusions and outline our future work in section 5.

2. PROBLEM FORMULATION

2.1 Data Model

For simplicity, we assume only one emitter. Let $s(n)$ denote the transmitted symbol at time nT where T is the symbol duration. This discrete time signal is modulated, filtered, transmitted over a Gaussian channel, filtered and down band converted. The resulting baseband signal has the expression:

$$y(t) = x(t) + n(t) = \sum_{k=-\infty}^{\infty} s(k)h(t - kT) + n(t) \quad (1)$$

Sampling this signal at baud rate, the resulting discrete signal will be a stationary process in wide sense. In that case, only minimum-phase channels could be recovered from second order statistics. As showed Tong, Xu and Kailath [1], when the received signal is oversampled (for example at M times the baud rate), the resulting discrete signal becomes cyclostationary, which allows to recover the phase of the channel from second order statistics. Moreover, this oversampling strategy allows modeling the channel as a single-input multiple-output FIR system. M -oversampling the received signal, we have:

$$y_i(n) = y(t_0 + (i-1)\Delta + nT) = \sum_{l=0}^{L-1} s(n-l)h_i(l) + n_i(n); \quad i=1..M \quad (2)$$

Where Δ is the new sample rate ($\Delta=T/M$) and the highest order of the resulting discrete M -multichannel system is $L+1$.

Let $\mathbf{h}_i = [h_i(0), h_i(1), \dots, h_i(L)]^T$ and defining $\mathbf{y}_i(n) = [y_i(n+W-1), y_i(n+W-2), \dots, y_i(n)]^T$ as a window of W consecutive samples:

$$\mathbf{y}_i(n) = \mathbf{H}_i(W) \cdot \mathbf{s}(n) + \mathbf{n}_i(n) \quad (3)$$

Where $\mathbf{s}(n) = [s(n+W-1), s(n+W-2), \dots, s(n-L)]^T$ and the $W \times (W+L)$ matrix $\mathbf{H}_i(W)$ is defined as:

$$\mathbf{H}_i(W) = \begin{bmatrix} h_i(0) & h_i(1) & \dots & h_i(L) & 0 & \dots & 0 \\ 0 & h_i(0) & h_i(1) & \dots & h_i(L) & \vdots & \\ \vdots & & & & & & 0 \\ 0 & \dots & 0 & h_i(0) & h_i(1) & \dots & h_i(L) \end{bmatrix} \quad (4)$$

Stacking the M vectors $\mathbf{y}_i(n)$ yields the MW -dimensional vector

$$\mathbf{y}(n) = [\mathbf{y}_1^T(n) \quad \mathbf{y}_2^T(n) \quad \dots \quad \mathbf{y}_M^T(n)]^T \quad (5)$$

And the received signal gives:

$$\mathbf{y}(n) = \mathbf{H}(W) \cdot \mathbf{s}(n) + \mathbf{n}(n) \quad (6)$$

With $\mathbf{H}(W)$ the $MW \times (W+L)$ matrix resulting of stacking the M $\mathbf{H}_i(W)$ matrices, $\mathbf{y}(n)$ a wide-sense cyclostationary process and $\mathbf{n}(n)$ a MW vector. On the other hand, to assure identifiability, \mathbf{H} should have full column rank. Conditions under this occur can be found in [2] and references therein.

2.2 Subspace Approach

The additive noise $\mathbf{n}(n)$ is modeled as a Gaussian, stationary, white and zero mean random process. For simplicity we take a spatially white noise, i. e., its $MW \times MW$ covariance matrix yields $\mathbf{R}_n = E[\mathbf{n}(n)\mathbf{n}^H(n)] = \sigma^2 \cdot \mathbf{I}$, where the function $E[\cdot]$ denotes mathematical expectation and superscript H is the hermitic operator. The covariance matrix of the received signal is:

$$\mathbf{R}_{yy} = E[\mathbf{y}(n)\mathbf{y}^H(n)] = \lim_{N_e \rightarrow \infty} \frac{1}{N_e} \sum_{n=1}^{N_e} \mathbf{y}(n)\mathbf{y}^H(n) = \mathbf{H}\mathbf{R}_s\mathbf{H}^H + \mathbf{R}_n \quad (7)$$

Where \mathbf{R}_s is the covariance matrix of \mathbf{s} and full column rank. The signal subspace is defined as the space spanned by columns of \mathbf{H} , and the noise subspace is its orthogonal complement. Both subspaces can be obtained from the SVD of covariance matrix:

$$\mathbf{R}_{yy} = \sum_{i=1}^{MW} \lambda_i \cdot \mathbf{e}_i \cdot \mathbf{e}_i^H = \mathbf{E}_s \cdot \Lambda_s \cdot \mathbf{E}_s^H + \mathbf{E}_o \cdot \Lambda_o \cdot \mathbf{E}_o^H \quad (8)$$

With $\lambda_i > \sigma^2$ for $i=1, \dots, r$ (signal eigenvalues, $r=W+L$), $\lambda_i = \sigma^2$ for $i=r+1, \dots, MW$ (noise eigenvalues), $\Lambda_s = \text{diag}(\lambda_1 \dots \lambda_r)$ and $\Lambda_o = \text{diag}(\lambda_{r+1} \dots \lambda_{MW})$. Defining the $MW \times r$ matrix \mathbf{E}_s and the $MW \times MW-r$ matrix \mathbf{E}_o :

$$\mathbf{E}_s = [\mathbf{e}_{s,1} \quad \mathbf{e}_{s,2} \quad \dots \quad \mathbf{e}_{s,r}]; \quad \mathbf{E}_o = [\mathbf{e}_{o,1} \quad \mathbf{e}_{o,2} \quad \dots \quad \mathbf{e}_{o,MW-r}] \quad (9)$$

Since columns of \mathbf{E}_s span the signal subspace and columns of \mathbf{E}_o the noise subspace, orthogonality between subspaces provides:

$$\mathbf{e}_{o,i}^H \cdot \mathbf{H} = 0 \quad \text{for } i=1, \dots, MW-(W+L) \quad (10)$$

3. WEIGTED SUBSPACE FITTING

3.1 Cost Functions

Since covariance matrix \mathbf{R}_{yy} is in practice estimated from a limited amount of received data, only an approximation of \mathbf{R}_{yy} is available. Then, (10) has to be solved in a least square sense. Adapting to blind channel identification the basic subspace

fitting problem in DOA estimation proposed by Viberg & Ottersen [3], we have that, given some representation of the data \mathbf{M} , we should to find an estimation of \mathbf{h} and \mathbf{T} such that

$$\mathbf{h}, \mathbf{T} = \arg \min_{\mathbf{h}, \mathbf{T}} \|\mathbf{M} - \mathbf{H}(\mathbf{h}) \cdot \mathbf{T}\|_F^2 \quad (11)$$

Where $\mathbf{h} = [\mathbf{h}_1^T \quad \mathbf{h}_2^T \quad \dots \quad \mathbf{h}_M^T]^T$ and $\|\cdot\|_F^2$ is the quadratic Froebius norm. This is a separable problem, and substituting the solution of the first part, $\mathbf{T} = \mathbf{H}^+ \cdot \mathbf{M}$ into (11) (the superscript $+$ denotes the pseudoinverse operator), this gives the generic subspace fitting cost function:

$$\hat{\mathbf{h}} = \arg \min_{\mathbf{h}} V(\mathbf{h}) \quad \text{with } V(\mathbf{h}) = \|\mathbf{I} - \mathbf{P}_{\mathbf{H}}\|_F^2 = \text{Tr}\{\mathbf{P}_{\mathbf{H}}^\perp \cdot \mathbf{M} \cdot \mathbf{M}^H\} \quad (12)$$

Where $\mathbf{P}_{\mathbf{H}} = \mathbf{H} \cdot \mathbf{H}^+$ is the LS projector onto the column space of \mathbf{H} , $\mathbf{P}_{\mathbf{H}}^\perp = \mathbf{I} - \mathbf{P}_{\mathbf{H}}$ is the orthogonal LS projector and $\text{Tr}\{\cdot\}$ denotes the trace operator. As \mathbf{M} is a representation of the data, different choices of \mathbf{M} will provide different cost functions (criteria).

3.1.1 MUSIC criterion.

The MUSIC criterion is based on orthogonality between signal and noise subspaces. The cost function derived from (10) can be written as:

$$V_{MUSIC}(\mathbf{h}) = \sum_{i=1}^{MW-r} \|\mathbf{e}_{o,i}^H \cdot \mathbf{H}\|^2 = \|\mathbf{E}_o^H \cdot \mathbf{H}\|_F^2 = \text{Tr}\{\mathbf{H}^H \cdot \mathbf{E}_o \cdot \mathbf{E}_o^H \cdot \mathbf{H}\} \quad (13)$$

This is the MUSIC cost function. Since in DOA problem (13) does not give accurate results when the signal are highly correlated, Schmidt [4] introduced a normalization matrix $(\mathbf{H}^H \mathbf{H})^{-1}$ into (13), resulting the *Multi-Dimensional (MD) MUSIC* algorithm:

$$\begin{aligned} \hat{\mathbf{h}} &= \arg \min_{\mathbf{h}} V_1(\mathbf{h}) \\ V_1(\mathbf{h}) &= \text{Tr}\{(\mathbf{H}^H \cdot \mathbf{H})^{-1} \cdot \mathbf{H}^H \cdot \mathbf{E}_o \cdot \mathbf{E}_o^H \cdot \mathbf{H}\} = \\ &= \text{Tr}\{\mathbf{P}_{\mathbf{H}} \cdot \mathbf{E}_o \cdot \mathbf{E}_o^H\} = \text{Tr}\{\mathbf{P}_{\mathbf{H}} \cdot (\mathbf{I} - \mathbf{E}_s \cdot \mathbf{E}_s^H)\} \end{aligned} \quad (14)$$

As the trace of a projection operator is equal to the dimension of the subspace on which it projects, minimizing $V_1(\mathbf{h})$ gives the same result that maximizing $V_2(\mathbf{h}) = \text{Tr}\{\mathbf{P}_{\mathbf{H}} \cdot \mathbf{E}_s \cdot \mathbf{E}_s^H\} = \text{Tr}\{(\mathbf{I} - \mathbf{P}_{\mathbf{H}}^\perp) \cdot \mathbf{E}_s \cdot \mathbf{E}_s^H\}$, and the cost function of *MD-MUSIC* algorithm can be expressed as:

$$\begin{aligned} \hat{\mathbf{h}} &= \arg \min_{\mathbf{h}} V_{MD-MUSIC}(\mathbf{h}) \\ V_{MD-MUSIC}(\mathbf{h}) &= \text{Tr}\{\mathbf{P}_{\mathbf{H}}^\perp \cdot \mathbf{E}_s \cdot \mathbf{E}_s^H\} \end{aligned} \quad (15)$$

Thus, the *MD-MUSIC* algorithm is a subspace fitting method where the representation of data is directly given by the signal subspace, $\mathbf{M} = \mathbf{E}_s$.

3.1.2 Deterministic ML criterion.

Described by Böhme [5], this method try to maximize the log likelihood of the received data \mathbf{Y} with respect to \mathbf{H} and \mathbf{S} (the columns of the $MW \times N_e$ matrix \mathbf{Y} are the N_e snapshots $\mathbf{y}(n)$ with

$n=1..N_e$, and the columns of the $(W+L) \times N_e$ matrix \mathbf{S} are the N_e transmitted symbol vectors $\mathbf{s}(n)$). This is equivalent to minimize with respect to \mathbf{h} the cost function $V_{MDL}(\mathbf{h}) = \|\mathbf{Y} - \mathbf{H} \cdot \mathbf{S}\|_F^2$. The solution of the first part gives $\hat{\mathbf{S}} = \mathbf{H}^+ \cdot \mathbf{Y}$ and the cost function becomes:

$$\hat{\mathbf{h}} = \arg \min_{\mathbf{h}} V_{MDL}(\mathbf{h}) \quad \text{with} \quad V_{MDL}(\mathbf{h}) = Tr\{\mathbf{P}_H^\perp \cdot \mathbf{R}_{yy}\} \quad (16)$$

Connection with subspace fitting can be made using asymptotic arguments (see [3]). For large N_e , we have that $\Lambda_o \rightarrow \sigma^2 \mathbf{I}$, $\mathbf{R}_{yy} \rightarrow \hat{\mathbf{E}}_s \cdot \bar{\Lambda} \cdot \hat{\mathbf{E}}_s^H + \sigma^2 \cdot \mathbf{I}$ and $\bar{\Lambda} = \Lambda_s - \sigma^2 \cdot \mathbf{I}$. As the trace of $\sigma^2 \mathbf{P}_H^\perp$ is a constant, the cost function is asymptotically (for large N_e) equivalent to:

$$V_{MDL}(\mathbf{h}) = Tr\{\mathbf{P}_H^\perp \cdot \hat{\mathbf{E}}_s \cdot \bar{\Lambda} \cdot \hat{\mathbf{E}}_s^H\} \quad (17)$$

Therefore, the deterministic ML method is a subspace fitting technique where the data are represented by the weighted signal subspace $\mathbf{M} = \hat{\mathbf{E}}_s \cdot \mathbf{W}_{MDL}^{1/2}$ (with $\mathbf{W}_{MDL} = \bar{\Lambda}$).

3.1.3 WSF criterion.

The *deterministic ML* method allows to us to introduce the weighted subspace-fitting concept. This is:

$$\hat{\mathbf{h}} = \arg \min_{\mathbf{h}} V_{WSF}(\mathbf{h}) \quad \text{with} \quad V_{WSF}(\mathbf{h}) = Tr\{\mathbf{P}_H^\perp \hat{\mathbf{E}}_s \mathbf{W}_{WSF} \hat{\mathbf{E}}_s^H\} \quad (18)$$

Where \mathbf{W}_{WSF} is a positive definite weighting matrix. The question is to find a weighting matrix \mathbf{W}_{WSF} that makes the estimation statistically (for large N_e) efficient, i. e.; that makes the WSF estimates asymptotically achieve the Cramer-Rao lower bound on the variance of the estimator error. Viberg and Ottersten [3] have shown that the optimal choice for \mathbf{W}_{WSF} is $\mathbf{W}_{WSF} = \bar{\Lambda}^2 \Lambda_s^{-1}$. Thus, the WSF criterion can be expressed in (13) when the representation of the data is $\mathbf{M} = \hat{\mathbf{E}}_s \cdot \mathbf{W}_{WSF}^{1/2}$.

3.2 Modified Variable Projection Algorithm

In order to estimate the channel, the criterion function in (18) must be minimized over the $M(L+1)$ -dimensional vector \mathbf{h} . The technique that we propose is the *Modified Variable Projection (MVP)* method (see [6] and references therein). This scheme is a Kaufmann's modification of the Gauss-Newton method: Consider the nonlinear least square problem given by (13), one of the more efficient, globally convergent, optimizing method for unconstrained smooth criteria in the damped Newton scheme. The estimate is iteratively calculated as:

$$\mathbf{h}_{k+1} = \mathbf{h}_k - \mu_k \cdot \mathbf{G}^{-1} \cdot \mathbf{v} \quad (19)$$

With μ_k the step length, \mathbf{G} the Hessian matrix of the cost function and \mathbf{v} the gradient. Every iteration the Hessian and the gradient are evaluated in \mathbf{h}_k . If \mathbf{h} is well initialized, the Newton method guarantees an ultimate quadratic convergence to $\hat{\mathbf{h}}$.

Consider first the gradient of the cost function. Introducing the \mathbf{r} vector by stacking the columns of $\mathbf{P}_H^\perp \cdot \mathbf{M}$, the cost function can be written as $V(\mathbf{h}) = |\mathbf{r}|^2$. The first derivative of V with respect to h_i^* (complex conjugate of the i^{th} element of \mathbf{h}) is:

$$\frac{\partial V}{\partial h_i^*} = 2 \left(\frac{\partial \mathbf{r}}{\partial h_i} \right)^H \mathbf{r} = 2 \mathbf{r}_i^H \cdot \mathbf{r} = 2 Tr \left\{ \left[\frac{\partial \mathbf{P}_H^\perp}{\partial h_i} \cdot \mathbf{M} \right]^H \cdot \mathbf{P}_H^\perp \cdot \mathbf{M} \right\} \quad (20)$$

The derivative of the projection matrix is:

$$\frac{\partial \mathbf{P}_H^\perp}{\partial h_i} = - \frac{\partial \mathbf{P}_H}{\partial h_i} = - \mathbf{P}_H^\perp \cdot \mathbf{H}_i \cdot \mathbf{H}^+ - \left(\mathbf{P}_H^\perp \cdot \mathbf{H}_i \cdot \mathbf{H}^+ \right)^H \quad (21)$$

Where now $\mathbf{H}_i = \frac{\partial \mathbf{H}}{\partial h_i}$ (do not confuse with $\mathbf{H}_i(W)$ defined in

(4)). Since $\mathbf{H}^+ \cdot \mathbf{P}_H^\perp = \mathbf{P}_H^\perp \cdot (\mathbf{H}^+)^H = \mathbf{0}$, the derivative yields:

$$\frac{\partial V}{\partial h_i^*} = -2 Tr \left\{ \mathbf{H}_i^H \cdot \mathbf{P}_H^\perp \cdot \mathbf{M} \cdot \mathbf{M}^H \cdot (\mathbf{H}^+)^H \right\} \quad (22)$$

Consider now the Hessian matrix. Differentiate (20) with respect to h_j^* we obtain:

$$\frac{\partial^2 V}{\partial h_i^* \partial h_j^*} = 2 \left[\mathbf{r}_i^H \cdot \mathbf{r}_j + \mathbf{r}_{ij}^H \cdot \mathbf{r} \right] \quad (23)$$

The Gauss modification of the Newton method assumes that the residual (i. e., the 2nd term in (23)) is small compared to the first term (see [6]). For large N_e , \mathbf{h}_k will be close to the optimum and the noise variance will be small. Then, we can approximate the Hessian matrix by $\mathbf{G}_{ij} = 2 \cdot \mathbf{r}_i^H \cdot \mathbf{r}_j$, and we have that $\mathbf{r}_i^H \mathbf{r}_j$ is:

$$Tr \left\{ \left[\mathbf{H}^+ \mathbf{H}_i^H \mathbf{P}_H^\perp \mathbf{H}_j^H \mathbf{H}^+ + \mathbf{P}_H^\perp \mathbf{H}_i \mathbf{H}^+ \mathbf{H}^+ \mathbf{H}_j^H \mathbf{P}_H^\perp \right] \mathbf{M} \mathbf{M}^H \right\} \quad (24)$$

The Kaufmann's modification of the algorithm consists on delete the second term in (24). This approximation can be hold because this term is small if the residual is small. Moreover the second term in (24) cancels part of the second term in (23). With these suppositions, the ij^{th} component of the approximate Hessian matrix is:

$$\mathbf{G}_{ij} = 2 Tr \left\{ \left[(\mathbf{H}^+)^H \cdot \mathbf{H}_i^H \cdot \mathbf{P}_H^\perp \cdot \mathbf{H}_j^H \cdot \mathbf{H}^+ \right] \cdot \mathbf{M} \cdot \mathbf{M}^H \right\} \quad (25)$$

Then, (19), (22) and (25) give the MVP algorithm. In order to obtain a good initialization of the estimates, we can use the MUSIC algorithm proposed by Moulines *et al.* [2]. This approach will provide a good estimation of the channel, and the algorithm will achieve a global minimum. The step length factor should be chosen in order to guarantee global convergence. It is known that quadratic convergence of Newton-type algorithms is only achieved if the step length factor converges to unity. Derivation of some convergence factors can be found in [7].

4. SIMULATION RESULTS

A Monte Carlo simulation is carried out in order to evaluate the performance of the algorithms in a digital communication system. We emitted a burst of $N=128$ BPSK symbols. The oversampling rate (i.e. the number of virtual channels) is $M=4$. The higher order of this virtual channels (ISI) is set to $L=2$ and the width of the temporal window is $W=4$.

First, we computed the *Mean Square estimation Error (MSE)* along the time (iterations) with a SNR of 10 dBs. We take $\mu_k=1-(0.1)^k$. The 0th iteration corresponds to the MSE provided by *MUSIC* algorithm proposed in [2]. Results of 100 Monte-Carlo runs are presented in Fig. 1.

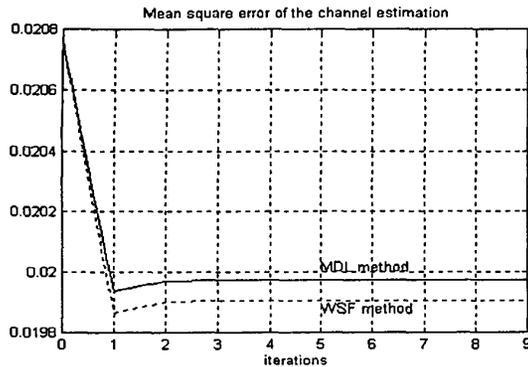


Figure 1. MSE along the iterations.

As can be seen, in less than four iterations the permanent state is achieved, and the minimum of the error is given at first iteration. As is expected, the *WSF* criterion gives the most accurate results. Unfortunately, the *MD-MUSIC* algorithm (i.e. weighting matrix set to identity) does not improve the MSE given by *MUSIC*.

The second experiment consists on computing the *Bit Error Rate (BER)* as function of the SNR with 40 Monte-Carlo simulations and 4 iterations of the *MVP* algorithm. In Fig. 2 results are presented. On the other hand, simulations of BER have shown that channel estimations provided by *MVP* algorithm with 1 and 4 iterations virtually give the same BER.

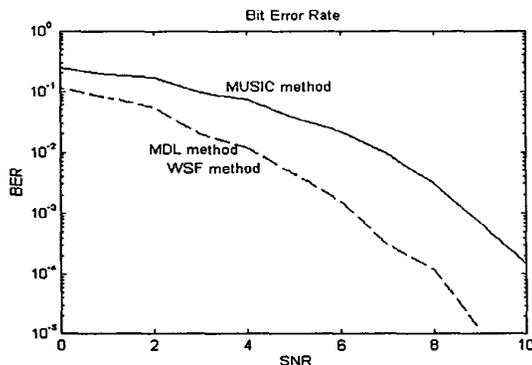


Figure 2. Bit error rate of the estimation.

As can be seen, both *MVP* techniques improve in 2 dBs the *MUSIC* algorithm proposed in [2]. On the other hand, *MDL* method has slightly better results in BER. Note that the minimum MSE does not guarantee the best BER.

5. CONCLUSIONS

The *Modified Variable Projection* algorithm has been developed and applied in a blind channel identification context. As it is shown in the figures, weighted subspace methods provide more accurate results than the *MUSIC* method developed by Moulines *et al.* in [2]. However, the *MVP* algorithm it is not a good choice in the unweighted method presented here (the *MD-MUSIC* method). On the other hand, the weighting matrices in *MDL* and *WSF* methods were defined for DOA problem; as channel estimation is not the same problem another weighting matrix can be derived for this specific situation as was done in [8]. Another simulations have been carried out in order to test the behavior of weighted algorithms with larger bursts. Results have shown that at larger sample size, less difference in MSE exists between *MUSIC* and the proposed algorithms. Therefore we can conclude that the weighting matrices defined in *MDL* and *WSF* criteria are useful in low sample sizes.

6. REFERENCES

- [1] Tong, L., Xu G. and Kailath T. "A new approach to blind identification and equalization of multipath channels". *Proc. 25th Asilomar Conf. On Signals, Systems and Computers*, pp. 856-860, Pacific Grove, CA., Nov. 1991.
- [2] E. Moulines, P. Dohamel, J. Cardoso & S. Mayrargue, "Subspace Methods for the blind identification of multichannel FIR filters", *IEEE Trans. On Signal Processing*, Vol. SP-43, pp. 516-525, Feb. 1995.
- [3] M. Viberg and B. Ottersten, "Sensor Array Processing based on subspace fitting", *IEEE Trans. On Signal Processing*, Vol. 39, pp. 1110-1121, May 1991.
- [4] R. O. Schmith, "A signal subspace approach to multiple emitter location and spectral estimation", *Ph. D. dissertation*, Stanford University, Stanford, CA, 1981.
- [5] J. F. Böhme, "Estimation of Spectral parameters of correlated signals in wavefields", *Signal Processing*, Vol. 11, pp. 329-337. Dec. 1986.
- [6] M. Viberg, B. Ottersten, "Detection and Estimation in Sensor Arrays Using Weighted Subspace Fitting", *IEEE Trans. On Signal Processing*, Vol. 39, pp. 2436-2449. Nov. 1991.
- [7] P. E. Gill, W. Murray and M. H. Wright, *Practical optimization*, London: Academic, 1981.
- [8] K. Abed-Meraim, J. F. Cardoso, A. Y. Gorokhov, P. Loubaton, E. Moulines, "On subspace methods for blind identification of single-input multiple-output FIR systems", *IEEE Trans. On Signal Processing*, Vol. 45, pp. 42-55. Jan. 1997.