BILINEAR METHODS FOR BLIND CHANNEL EQUALIZATION: (NO) LOCAL MINIMUM ISSUE

Eric Pité and Pierre Duhamel

ENST/SIG, 46, rue Barrault, 75013 Paris, France
fax: 33 (1) 45 88 79 35 e-mail: duhamel@sig.enst.fr

ABSTRACT

Bilinear methods for jointly estimating the channel coefficients and the symbols emitted through these channels are very appealing. However, they can be trapped by local minima. This paper provides **(i)** a full characterization of the local minima, **(ii)** a simple criterion for checking whether the procedure has converged to the global minimum, **(iii)** a simple algorithm for obtaining this solution, with a proof of convergence.

1. INTRODUCTION

Blind equalization and channel identification is a field that has received increased interest recently, mainly due to the availability of second-order based methods that have the potential for providing fast and efficient algorithms. These second-order methods rely on channel diversity when the received signal can be seen as the output of a single-input /multiple-output (SIMO) channel. Such a setup may be applicable either in the case of reception through an antenna array or in the case of fractionnaly spaced receivers.

Many methods are based on statistical considerations [7, 6, 9], tending to assume that the input sequence is long enough, so that its statistics can be estimated precisely. Some methods solve the problem by treating the input as a *deterministic* signal, resulting in the so-called *deterministic* approach to blind identification [5, 4, 1]. While some results are easily derived in the statistical framework, their practical use can be increased by translating them into the deterministic one. The results below are an example of such a situation.

Among the many methods that have been recently proposed, a very promising one is the so-called "bilinear" one, which has been proposed by various authors under various names. This method is a blind least-squares approach for joint data/channel identification. The goal is to minimize a distance between the received signal and the estimated one, which results in a bilinear cost function when the channel coefficients are unknown. The main interest of such bilinear approaches is that one does not *invert* the channel, which always has drawbacks (noise amplification, resulting colored noise at the output of the equalizer, $\cdots$). Most of these methods exploit the finite alphabet property of the data to provide an iterative algorithm based on Maximum likelihood, [8, 10]. As a consequence, the algorithm can be trapped in the numerous local minima of the cost function, but works very well when it converges to the actual optimum. Some methods do not rely on this finite alphabet property, [3] but the local minima problems are not fully solved, while experimentally being recognized as being of smaller importance. This class of methods is usually denoted as Iterative Quadratic Maximum Likelihood (IQML) [3], or Deterministic Maximum Likelihood [8].

This paper provides an explicit characterization of the local minima of such methods. It is shown that the possible local minima do not meet the required assumption that the channels have "maximum diversity", or that the input sequence is "persistently exciting". Hence, we provide a tractable way of checking whether the algorithm has been trapped in a local minimum or not. More practically, we derive a simple iterative algorithm (similar to IQML), with a convergence proof, that can be used in a context of short data records. We then draw a parallel with the deterministic blind maximum likelihood approach.

Note that, for mathematical convenience and simplicity of the derivation, all results are given in the noiseless case.

2. PROBLEM STATEMENT

This paper considers a multichannel model to represent the SIMO equivalent of a digital communication system.

Let $x_i(\cdot)$ denote the output from the i^{th} channel with the FIR channel impulse response $\{h_i(\cdot)\}$, all channels being driven by the same input $s(\cdot)$. For linearly modulated communication signals, we have

$$x_i(n) = \sum_{k=0}^{M} h_i(k)s(n-k) + b_i(n) \quad i = 1, \ldots, L \qquad (1)$$

where L is the number of channels, M is the maximum order of the channels. $b_i(\cdot)$ $(i = 1, \ldots, L)$ are supposed to be i.i.d, mutually uncorrelated processes.

Consider the vectorized processes

$$\begin{aligned}
\mathbf{x}(n) &= (x_1(n), \ldots x_L(n))^T && (2)\\
\mathbf{b}(n) &= (b_1(n), \ldots, b_L(n))^T && (3)\\
\mathbf{h}(k) &= (h_1(k) \ldots, h_L(k))^T && (4)\\
\mathbf{h}_j &= (h_j(0), \ldots, h_j(M))^T && (5)\\
\mathbf{h} &= (\mathbf{h}_1, \mathbf{h}_L)^T && (6)
\end{aligned}$$

where the superscript ^{T} denotes transposition. The data model is written as the following convolution equation:

$$\mathbf{x}(n)=\sum_{k=0}^{M}\mathbf{h}(k)s(n-k)+\mathbf{b}(n) \tag{7}$$

Denote by $X_N(n) = \left(\mathbf{x}(n)^T,\ldots,\mathbf{x}(n-N+1)^T\right)^T$ a space time data record of size LN, $\mathbf{s}_N(n) = (s(n),s(n-1)\ldots,s(n-M-N+1))^T$ the $M+N$ vector involving a total of $M+N$ interfering symbols and $B_N(n)=\left(\mathbf{b}(n)^T,\ldots,\mathbf{b}(n-N+1)^T\right)^T$. The following linear model holds:

$$X_N(n)=\mathcal{T}_N(\mathbf{h})\mathbf{s}_N(n)+B_N(n) \tag{8}$$

where $\mathcal{T}_N(\mathbf{h})$ is the $LN\times(M+N)$ generalized Sylvester matrix

$$\mathcal{T}_N(\mathbf{h})=\begin{pmatrix}\mathbf{h}(0) & \cdots & \mathbf{h}(M) & \mathbf{0} & \cdots & \mathbf{0}\\ \mathbf{0} & \ddots & & \ddots & \ddots & \vdots\\ \vdots & \ddots & \ddots & & \ddots & \mathbf{0}\\ \mathbf{0} & \cdots & \mathbf{0} & \mathbf{h}(0) & \cdots & \mathbf{h}(M)\end{pmatrix}$$

Permuting the role of $s(\cdot)$ and $\mathbf{h}$ we also have:

$$X_N(n)=\mathcal{U}\left(\mathbf{s}_N(n)\right)\mathbf{h}+B_N(n) \tag{9}$$

where $\mathcal{U}\left(\mathbf{s}_N(n)\right)$ is the $LN\times L(M+1)$ data matrix:

$$\mathcal{U}\left(\mathbf{s}_N(n)\right)=\begin{pmatrix}I_L\otimes\mathbf{s}_1(n)^T\\ I_L\otimes\mathbf{s}_1(n-1)^T\\ \vdots\\ I_L\otimes\mathbf{s}_1(n-N+1)^T\end{pmatrix} \tag{10}$$

$\otimes$ denotes the Kronecker product and I_L is the identity matrix of size L. The specificity of the multichannel framework compared to the classical single channel situation is that $BOTH$ equations (10) and (8) form an overdetermined set of equations under appropriate conditions.

Assume that M is known or correctly estimated and consider the problem of identifying both $\mathbf{h}$ and $\mathbf{s}_N(n)$ from $X_N(n)$ only. An obvious way to solve this, is to consider the minimization of the criterion

$$J\left(\hat{\mathbf{h}},\hat{\mathbf{s}}_N(n)\right)=||X_N(n)-\mathcal{T}_N(\hat{\mathbf{h}})\hat{\mathbf{s}}_N(n)||^2 \tag{11}$$

where $||\cdot||$ is the L_2-norm. Our results rely on the assumption that $\mathcal{T}_N(\mathbf{h})$ has full rank. It is now well known (see [7]) that the full rank conditions are given by

Lemma 1 *$\mathcal{T}_N(\mathbf{h})$ has full rank, i.e.* $\operatorname{rank}(\mathcal{T}_\mathrm{N}(\mathbf{h}))=\mathrm{M}+\mathrm{N}$, *if i) the polynomials $H_j(z)=\sum_{i=0}^{M}h_j(i)z^i$ have no common zero, ii) $N\geqslant M$ and iii) at least one polynomial $H_j(z)$ has degree M.*

3. BLIND IDENTIFICATION

This section provides a sufficient condition for the criterion in (11) to have only the true cofficients channel and symbols (up to a scalar factor) as a global minimum. This is another proof of the same result given in [2].

Definition 1 *The linear complexity of sequence $\mathbf{s}_N(n)$ is defined as the smallest value of c for which there exists $\lambda_1,\ldots\lambda_c\in\mathbb{C}$ such as*

$$s(k)=\sum_{i=1}^{c}\lambda_i s(k-i)\ ,\quad k=n-N-M-c+1,\ldots,n$$

Linear complexity measures the predictability of a finite-length deterministic sequence.

Consider the $2M+1\times(N-M)$ matrix $\mathcal{V}_M\left(\mathbf{s}_N(n)\right)$

$$\begin{pmatrix}s(n) & s(n-1) & \cdots & s(n-N+M+1)\\ s(n-1) & s(n-2) & \cdots & s(n-N+M)\\ \vdots & \vdots & & \vdots\\ s(n-2M) & s(n-2M-1) & \cdots & s(n-N-M+1)\end{pmatrix}$$

If $\mathbf{s}_N(n)$ has linear complexity greater than $2M$ and $N\geqslant 3M+1$, then $\operatorname{rank}(\mathcal{V}_\mathrm{M}(\mathbf{s}_\mathrm{N}(\mathrm{n})))=2\mathrm{M}+1$.

Theorem 1 *In the noiseless case, if $\mathcal{T}_N(\mathbf{h})$ has full rank and if the linear complexity of $\mathbf{s}_N(n)$ is greater than $2M$, $J\left(\hat{\mathbf{h}},\hat{\mathbf{s}}_N(n)\right)=0$ iff $\exists\alpha\in\mathbb{C}^*$ such as $\hat{\mathbf{h}}=\alpha\mathbf{h}$ and $\hat{\mathbf{s}}_N(n)=\mathbf{s}_N(n)/\alpha$.*

Proof is outlined in appendix A

Remarks: **(i)** In a statistical context, Moulines *et al.*, [7], provide a mechanism for estimating the channel coefficients (*i.e.* $\mathbf{h}$) in the cases where the autocorrelation matrix of the transmitted symbols is unknown, provided it has full rank. Our result only extend this mechanism in providing explicit conditions for the empirical autocorrelation estimate to be full rank. **(ii)** Even if Theorem 1 says that the bilinear criterion has the right global minimum, it does not provide information on the convergence towards the minimum of (11) because J has possible local minima.

These local minima are now characterized.

4. LOCAL MINIMA

Considering equations (10) and (8), it is easily seen that $\left(\hat{\mathbf{h}},\hat{\mathbf{s}}_N(n)\right)$ is a stationnary point of J iff one can find "wrong" estimates $\hat{\mathbf{s}}_N(n)$ and $\hat{\mathbf{h}}$ of s and $\mathbf{h}$ characterized by:

$$\begin{aligned}\hat{\mathbf{s}}_N(n)^*\mathcal{T}_N(\hat{\mathbf{h}})^*\mathcal{T}_N(\hat{\mathbf{h}}) &= \mathbf{s}_N(n)^*\mathcal{T}_N(\mathbf{h})^*\mathcal{T}_N(\hat{\mathbf{h}}) && (12)\\ \hat{\mathbf{h}}^*\mathcal{U}(\hat{\mathbf{s}}_N(n))^*\mathcal{U}\left(\hat{\mathbf{s}}_N(n)\right) &= \mathbf{h}^*\mathcal{U}(\mathbf{s}_N(n))^*\mathcal{U}(\hat{\mathbf{s}}_N(n)) && (13)\end{aligned}$$

where * denotes trans-conjugaison.

Unfortunatly, it seems very difficult to find the explicit solution of (12) and (13) in the general case. And for small values of L, M, N, when (12) is solvable, calculation of the Hessian matrix of J (required for deciding if the solutions of (12) and (13) are local minima or not) is even more difficult.

However, the following theorem shows that there is no local minimum in J if the estimated channels do not have common zeroes and if the estimated input sequence is persistently exciting :

Theorem 2 *In the noiseless case, suppose that* $\left(\hat{\mathbf{h}}, \hat{\mathbf{s}}_N(n)\right)$ *is a local minimum of* J*. If* $\mathcal{T}_N(\mathbf{h})$*,* $\mathcal{T}_N(\hat{\mathbf{h}})$ *have full rank and if the linear complexity of* $\mathbf{s}_N(n)$ *and* $\hat{\mathbf{s}}_N(n)$ *is greater than* $2M$*; then* $\hat{\mathbf{h}} = \alpha\mathbf{h}$ *and* $\hat{\mathbf{s}}_N(n) = \mathbf{s}_N(n)/\alpha$*.*

proof is outlined in appendix B.

5. BLOCK ALGORITHM

If $\mathcal{T}_N(\hat{\mathbf{h}})$ has full rank and if $\hat{\mathbf{s}}_N(n)$ has linear complexity greater than $2M$ (hence $\mathcal{U}(\hat{\mathbf{s}}_N(n))$ has full rank) equations (12) and (13) become:

$$\hat{\mathbf{s}}_N(n) = \left(\mathcal{T}_N(\hat{\mathbf{h}})^* \mathcal{T}_N(\hat{\mathbf{h}})\right)^{-1} \mathcal{T}_N(\hat{\mathbf{h}})^* X_N(n) \quad (14)$$

$$\hat{\mathbf{h}} = (\mathcal{U}(\hat{\mathbf{s}}_N(n))^* \mathcal{U}(\hat{\mathbf{s}}_N(n)))^{-1} \mathcal{U}(\hat{\mathbf{s}}_N(n))^* X_N(n) \quad (15)$$

We derive from (14) an Iterative Quadratic Least Square algorithm (IQML), similar to Hua's TSML algorithm [3]. Given a value $\mathbf{h}^{(k)}$ of $\hat{\mathbf{h}}$, we deduce $\mathbf{s}_N^{(k)}(n)$ by solving in the least mean-square sense (11), $\mathbf{s}_N^{(k)}(n) = \mathcal{T}_N(\mathbf{h}^{(k)})^{\#} X_N(n)$ (the superscript $^{\#}$ denotes Moore-Penrose pseudo inverse). And from $\mathbf{s}_N^{(k)}(n)$ and the observation $X_N(n)$, we deduce $\mathbf{h}^{(k+1)} = \mathcal{U}\left(\mathbf{s}_N^{(k)}(n)\right)^{\#} X_N(n)$. The process is then iterated until convergence. Each computation being a least squares problem, both steps have no local minimum problem. The question now is about the convergence of the global procedure.

Another way to express this algorithm with projections is the following:

Take an initialization of $\mathbf{s}_N^{(0)}(n)$ and $\mathbf{h}^{(0)}$, then

$$\mathbf{s}_N^{(k+1)}(n) = \Pi_{\mathbf{h}^{(k)},N} \mathbf{s}_N^{(k)}(n) \quad (16)$$

$$\mathbf{h}^{(k+1)} = \Pi_{\mathbf{s}^{(k+1)},N} \mathbf{h}^{(k)} \quad (17)$$

where $\Pi_{\mathbf{h}^{(k)},N}$ (resp. $\Pi_{\mathbf{s}^{(k+1)},N}$) is the orthogonal projection onto range $\left(\mathcal{T}_{\mathrm{N}}(\hat{\mathbf{h}})\right)$ (resp. range $(\mathcal{U}(\hat{\mathbf{s}}_{\mathrm{N}}(\mathrm{n})))$.

$$\Pi_{\hat{\mathbf{h}},N} = \mathcal{T}_N(\hat{\mathbf{h}})^{\#} \mathcal{T}_N(\hat{\mathbf{h}})^*, \ \Pi_{\hat{\mathbf{s}},N} = \mathcal{U}(\hat{\mathbf{s}}_N(n))^{\#} \mathcal{U}(\hat{\mathbf{s}}_N(n))^*$$

Each step of the algorithm decreases $||X_N(n) - \mathcal{T}_N(\mathbf{h}^{(k)}) \mathbf{s}_N^{(k)}(n)||^2$ so it converges to a point $(\mathbf{h}^\infty, \mathbf{s}_N^\infty(n))$. One easily verifies that $(\mathbf{h}^\infty, \mathbf{s}_N^\infty(n))$ is a local minimum of J. So we have the

Theorem 3 *If* $(\mathbf{h}^\infty, \mathbf{s}_N^\infty(n))$ *statisfies asumptions of theorem 2, then* $(\mathbf{h}^\infty, \mathbf{s}_N^\infty(n))$ *is equal to* $(\mathbf{h}, \mathbf{s}_N(n))$ *up to a scalar factor.*

Figure 1 shows the convergence of the algorithm to the global minimum while figure 2 shows a convergence to a local minimum which is detected because the estimated channels share a common zero.

Practically, simple and reliable ways for checking the globality of the minimum, and restart the algorithm at a right location still remains to be found.

6. LINK WITH THE DETERMINISTIC MAXIMUM LIKELIHOOD APPROACH

As its name indicates, DML considers some of the stochastic parameters $(s(n))$ as deterministic quantities.

The additive noise samples $\{\mathbf{b}(n)\}$ are assumed to be i.i.d complex random variables that are gaussian with zero mean and unknown variance $\sigma_{\mathbf{b}}^2$, the real and the imaginary parts being independent. The likelihood function is the pdf of the noise

$$p\left(X_N(n); \Theta(n)\right) = \frac{1}{(2\pi\sigma_{\mathbf{b}}^2)^{LN/2}} e^{-\frac{1}{2\sigma_{\mathbf{b}}^2}||X_N(n) - \mathcal{T}_N(\hat{\mathbf{h}})\hat{\mathbf{s}}_N(n)||^2} \quad (18)$$

where $\Theta(n) = (\hat{\mathbf{s}}_N(n)^T, \hat{\mathbf{h}}^T)^T$.

The goal to determine the parameter $\Theta(n)$ that maximizes $p\left(X_N(n); \Theta(n)\right)$. The maximization or the likelihood fonction therefore boils to the following least-square problem: $\min_{\hat{\mathbf{h}}, \hat{\mathbf{s}}_N(n)} J\left(\hat{\mathbf{h}}, \hat{\mathbf{s}}_N(n)\right)$.

This optimization is separable, so we have to minimize $\left|\left|\left(I_{LN} - \mathcal{T}_N(\hat{\mathbf{h}}) \left(\mathcal{T}_N(\hat{\mathbf{h}})^* \mathcal{T}_N(\hat{\mathbf{h}})\right)^{-1} \mathcal{T}_N(\hat{\mathbf{h}})^*\right) X_N(n)\right|\right|^2$. So

$$\mathbf{h} = \alpha \arg\max_{\hat{\mathbf{h}}} \left|\left|\Pi_{\hat{\mathbf{h}},N} X_N(n)\right|\right|^2 \quad (19)$$

where α is a scale factor. However, as described in the work by Slock and Papadias, the solution is obtained through a highly non linear minimization procedure, as seen above. Since the equivalence between the "bilinear methods" and the Deterministic ML has been recognized, we thus have shown that the solution of DML can be obtained through a sequence of least squares problems.

7. CONCLUSION

The characterization of the local minima provided in this paper opens several new possibilities for these so-called bilinear or deterministic ML methods.

- this provides the base for methods that will converge to the global minimum.
- As shown in [2], recursive solutions are easily derived form the block methods derived in the paper.
- These methods have the potential for providing the equivalent in the multichannel framework of the "Decision Feedback Equalizer" (with soft decisions), with the additional property of a guaranteed convergence.

Further work will be reported.

A. PROOF OF THEOREM 1

The idea of the proof is due to D. Gesbert [2]. A different proof can be found in [11].

Rewrite $X_N(n) = \mathcal{T}_N(\mathbf{h})\mathbf{s}_N(n) = \mathcal{T}_N(\hat{\mathbf{h}})\hat{\mathbf{s}}_N(n)$ as $\mathcal{T}_{M+1}(\mathbf{h})\mathcal{V}_M(\mathbf{s}_N(n)) = \mathcal{T}_{M+1}(\hat{\mathbf{h}})\mathcal{V}_M(\hat{\mathbf{s}}_N(n))$.
range $(\mathcal{T}_{\mathrm{M+1}}(\mathbf{h})\mathcal{V}_{\mathrm{M}}(\mathbf{s}_{\mathrm{N}}(\mathrm{n})))$ = range $(\mathcal{T}_{\mathrm{M+1}}(\mathbf{h}))$ (because rank $(\mathcal{V}_{\mathrm{M}}(\mathbf{s}_{\mathrm{N}}(\mathrm{n})))$ = $2\mathrm{M}+1$). Then range $(\mathcal{T}_{\mathrm{M+1}}(\mathbf{h}))$ = range $\left(\mathcal{T}_{\mathrm{M+1}}(\hat{\mathbf{h}})\right)$.

Using [7, theorem 2], there exists $\alpha \in \mathbb{C}$ such as $\hat{\mathbf{h}} = \alpha \mathbf{h}$, but $\hat{\mathbf{h}} \neq \mathbf{0}$ implies $\alpha \neq 0$. Plus, $\mathcal{T}_N(\mathbf{h})$ has full column rank, so it is left invertible, then $\hat{\mathbf{s}}_N(n) = \mathbf{s}_N(n)/\alpha$.

B. PROOF OF THEOREM 2

For simplicity assume that $N = 3M + 1$.

Sketch of the proof:

1 – If $\left(\hat{\mathbf{h}}, \hat{\mathbf{s}}_N(n)\right)$ is a local minimum of $||X_N(n) - \mathcal{T}_N(\hat{\mathbf{h}})\hat{\mathbf{s}}_N(n)||^2$ then $\left(\hat{\mathbf{h}}, \hat{\mathbf{s}}_{M+1}(n-k)\right)$ is a local minimum of $||X_{M+1}(n-k) - \mathcal{T}_{M+1}(\hat{\mathbf{h}})\hat{\mathbf{s}}_{M+1}(n-k)||^2$ ($\forall k \in \{0, \ldots, 2M\}$).
Let $\hat{X}_{M+1}(n-k) = \mathcal{T}_{M+1}(\hat{\mathbf{h}})\hat{\mathbf{s}}_{M+1}(n-k)$.

2 – Denote $\widetilde{R}_{X_{M+1}}(n)$, $\widetilde{R}_{\hat{X}_{M+1}}(n)$ and $\widetilde{R}_{\hat{\mathbf{s}}_{M+1}}(n)$ the sample covariance of the vector $X_{M+1}(n)$, $\hat{X}_{M+1}(n)$ and $\hat{\mathbf{s}}_{M+1}(n)$. using (12), we have

$$\hat{X}_{M+1} = \Pi_{\hat{\mathbf{h}},M+1} X_{M+1} \quad (20)$$

$$\widetilde{R}_{\hat{X}_{M+1}}(n) = \Pi_{\hat{\mathbf{h}},M+1} \widetilde{R}_{X_{M+1}}(n) \Pi_{\hat{\mathbf{h}},M+1} \quad (21)$$

3 – Write the EVD of $\widetilde{R}_{X_{M+1}}(n)$, $\widetilde{R}_{\hat{X}_{M+1}}(n)$ and $\Pi_{\hat{\mathbf{h}},M+1}$, then show that $\mathrm{range}(\mathrm{R_{X_N}}) = \mathrm{range}\,(\mathcal{T}_\mathrm{N}(\mathbf{h}))$ (because
$\widetilde{R}_{X_{M+1}}(n) = \mathcal{T}_{M+1}(\mathbf{h})\widetilde{R}_{\hat{\mathbf{s}}_{M+1}}(n)\mathcal{T}_{M+1}(\hat{\mathbf{h}})^*$ and $\widetilde{R}_{\hat{\mathbf{s}}_{M+1}}(n)$ has full rank.).

4 – Deduce that $\mathrm{range}\,(\mathcal{T}_\mathrm{N}(\mathbf{h})) = \mathrm{range}\left(\mathcal{T}_\mathrm{N}(\hat{\mathbf{h}})\right)$, then $\hat{\mathbf{h}} = \alpha \mathbf{h}$ ([7, theorem 2]).

5 – Finally, $\left|\left|\mathcal{T}_N(\mathbf{h})\left[\mathbf{s}_N(n) - \alpha\hat{\mathbf{s}}_N(n)\right]\right|\right|^2$ has a local minimum, but is quadratic in $\hat{\mathbf{s}}_N(n)$, hence the minimum is global. Finally, due to theorem 1, $\hat{\mathbf{s}}_N(n) = \alpha \mathbf{s}_N(n)$.

C. REFERENCES

[1] A. Van der Veen, S. Talwar, and A. Paulraj. Blind estimation of multiple digital signals transmitted over fir channels. *IEEE S. P. Letters*, 2(5):99–102, May 1995.

[2] D. Gesbert, P. Duhamel, and S. Mayrargue. Blind least-square approaches for joint data/channel estimation. In *IEEE D. S. P. Workshop, Loen, Norway*, pages 450–453, September 1996.

[3] Y. Hua. Fast maximum likelihood for blind identification of multiple fir channels. *IEEE Trans. on S. P.*, 44(3):661–672, March 1996.

[4] H. Liu and G. Xu. A deterministic approach to blind symbol estimation. *IEEE S. P. Letters*, 1(12):205–207, December 1994.

[5] H. Liu, G. Xu, and L. Tong. A deterministic approach to blind identification of multi-channel fir systems. In *Pro. ICASSP*, pages 581–584, 1994.

[6] K. Abed-Meraim P. Loubaton and E. Moulines. A subspace algorithm for certain blind identification problems. *IEEE Trans. on I. T.*, 43(2):499–511, March 1997.

[7] E. Moulines, P. Duhamel, J.-F. Cardoso, and S. Mayrargue. Subspace methods for blind identification of multichannel fir filters. *IEEE Trans. on S. P.*, 43(2):516–525, February 1995.

[8] D. Slock and C. Papadias. Furthers results on blind identification and equalization of multiple fir channels. In *Poceedings ICASSP*, 1995.

[9] L. Tong, G. Xu, and T. Kailath. Blind identification and equalization based on second-order statistics: a time-domain approach. *IEEE Trans. on I. T.*, 40(2):340–349, March 1994.

[10] L. White and V. Krishnamurthy. Adaptative blind equalization of fir channels using hidden markov models, 1993.

[11] G. Xu, H. Liu, L. Tong, and T. Kailath. A least-square approach to blind channel identification. *IEEE Trans. on S. P.*, 43(12):2982–2993, December 1995.

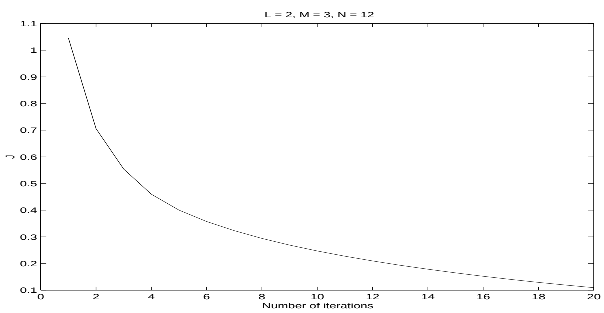

Figure 1: $||X_N(n) - \mathcal{T}_N(\mathbf{h}^{(k)})\ \mathbf{s}_N^{(k)}(n)||$ versus k

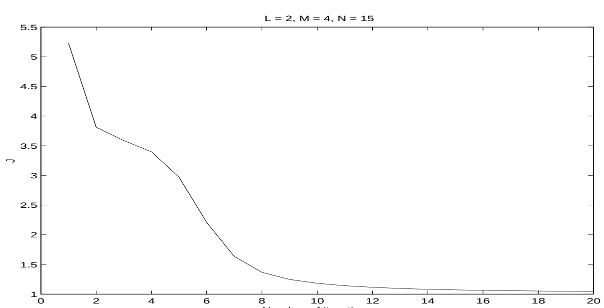

Figure 2: $||X_N(n) - \mathcal{T}_N(\mathbf{h}^{(k)})\ \mathbf{s}_N^{(k)}(n)||$ versus k