SEMI-BLIND SECOND ORDER IDENTIFICATION OF CONVOLUTIVE CHANNELS

A. Gorokhov†

Télécom Paris, Département SIGNAL
46 rue Barrault 75634 Paris
Cedex 13 FRANCE
e-mail: gorokhov@sig.enst.fr

Ph. Loubaton

Université de Marne la Vallée, UF SPI
2 rue de la Butte Verte 93166 Noisy le Grand
Cedex FRANCE
e-mail: loubaton@univ-mlv.fr

ABSTRACT

Our contribution addresses the identification of multiple convolutive FIR channels. Many recently proposed blind techniques suffer from the imperfect knowledge of the channel order. Meanwhile the major part of the existing communication systems requires periodically transmitted reference sequences known at the reception site. This information can be used to ensure the robustness the existing blind approaches. We consider here joint utilization of the referenced snapshots with the non-referenced data and construct a combined estimator originating from the blind subspace based technique applied to the Single Input Multiple Output (SIMO) systems identification. The statistical efficiency analysis and numerical examples are presented.

1. INTRODUCTION

The problem of multiple filters deconvolution typically appears in digital communication where the transmitted signal is subject to the multipath propagation resulting in the severe inter-symbol interference (ISI). The classical approach is based on sending *reference sequences* periodically alterating the useful message and utilized for the channel acquisition and further signal extraction. A certain number of recently proposed blind methods [1]-[5] exploits the non-referenced observation. In particular several second order techniques [6]-[9] are based on multi-sensor reception or fractionally spaced observation both leading to the SIMO convolutive model identifiable from the second order statistics of the observed signals. Quite often the success of these methods is owing to the perfect knowledge of the channel order as well as good *inter-channel disparity*. However most of the existing communication systems still require some reference signals for various operational purposes (*e.g.* modem synchronization). These sequences may be usually done much shorter than the data necessary for the channel acquisition. Such an auxiliary information can be also used to enhance the blind identification techniques based on the non-referenced observation.

In this contribution we derive a combined estimator based on the blind subspace based criterion [8] recently analyzed in [10], [11]. Such choice of the blind technique is motivated by certain asymptotic properties of the subspace based approach. The related statistical result is discussed later in this paper.

The advantage of semi-blind identification techniques with respect to its purely blind counterparts within the high capacity communication systems depends on the possibility to operate with relatively short reference sequences. We focus on the case when the information data length essentially exceeds the duration of the reference sequence. Some analytic results presented in the article may be helpful in choosing these parameters. We finally offer the numerical validation of our conclusions.

†This study is supported by CNET (France Télécom), ENST and the SASPARC project of INTAS.

2. DATA MODEL

Let an M-variate times series $\{y(t)\}_{t\in\mathbb{Z}}$ represent a baseband noisy output of M antennas while the transmitted scalar series $\{s(t)\}_{t\in\mathbb{Z}}$ is subject to multipath propagation:

$$y(t) = x(t) + b(t), \quad x(t) = \sum_{\tau=0}^{L} \mathbf{h}(\tau)\, s(t-\tau), \quad t \in \mathbb{Z}. \quad (1)$$

Here the set of $M \times 1$ vectors $\mathbf{h}(\tau)$, $\tau = 0, \dots, L$ defines the overall channel impulse response. The associated multivariate transfer function is conventionally given by $h(z) = \sum_{\tau=0}^{L} \mathbf{h}(\tau)\, z^{-\tau}$, where $[h(z)]_p$ stands for the scalar transfer function to the p-th output. The objective is to construct the estimator of $h(z)$ based on the observation $\{y(t)\}_{t\in\mathbb{Z}}$ and partially known input signal. We specify the two separate observation sets: $Y_{T_r} \triangleq [y(t_0+T_r-1)^T, \dots, y(t_0)^T]^T$ and $Y_T \triangleq [y(t_1+T-1)^T, \dots, y(t_1)^T]^T$ associated with the input data sets $S_{T_r} \triangleq [s(t_0+T_r-1), \dots, s(t_0-L)]^T$ and $S_T \triangleq [s(t_1+T-1), \dots, s(t_1-L)]^T$ correspondingly. Here Y_T is non-referenced data with unknown inputs while Y_{T_r} corresponds to the referenced signal *i.e.* the entries of S_{T_r} are known. We also assume for simplicity that S_T and S_{T_r} do not overlap and rely upon the following hypotheses:

H1 The number of outputs M is strictly greater than one.

H2 The input signal $\{s(t)\}_{t\in\mathbb{Z}}$ is a zero-mean complex forth-order stationary process with the *unknown* power spectral density $\varrho(\omega)$.

H3 The additive noise $\{b(t)\}_{t\in\mathbb{Z}}$ is spatially and temporally white *complex circular Gaussian* zero-mean stationary process of power σ.

Note that the latter assumption may be relaxed to the case of spatially colored noise with the *known* covariance structure since preliminary spatial prewhitening leads back to the hypothesis **(H3)**. We now propose a brief review of the referenced and blind identification techniques.

3. REFERENCED AND BLIND TECHNIQUES

Let us first consider the referenced data Y_{T_r}. This vector can be written as $Y_{T_r} = X_{T_r} + B_{T_r}$, where $X_{T_r} = [x(t_0+T_r-1)^T, \dots, x(t_0)^T]^T$ and $B_{T_r} = [b(t_0+T_r-1)^T, \dots, b(t_0)^T]^T$. The convolution (1) may be rewritten in the algebraic form using the properties of *Hankel matrix*. Let us denote $\mathbf{h} = [\mathbf{h}(0)^T, \ldots, \mathbf{h}(L)^T, 0_{1\times M(L'-L)}]^T$ the vector of channel coefficients, here L' represents the *virtual* channel order *i.e.* the result of order determination or any sufficiently large value involved in estimation. Then $X_{T_r} = \mathcal{H}_{L'}(S_{T_r})\,\mathbf{h}$ with $\mathcal{H}_{L'}(S_{T_r})$ given by

$$\begin{bmatrix} s(t_0+T_r-1) & s(t_0+T_r-2) & \ldots & s(t_0+T_r-L'-1) \\ s(t_0+T_r-2) & s(t_0+T_r-3) & \ldots & s(t_0+T_r-L'-2) \\ \vdots & \iddots & \iddots & \vdots \\ s(t_0+L') & s(t_0+L'-1) & \ldots & s(t_0) \end{bmatrix} \otimes \mathrm{I}_M,$$

a block Hankel matrix of the input data with each block proportional to the identity matrix (here $(\otimes)$ denotes Kronecker product) and $L' \geq L$. According to the classical results [12]-[13], the assumption **(H3)** yields a simple minimum variance estimate of $\mathbf{h}$ from Y_{T_r} given by the quadratic minimizer $\hat{\mathbf{h}}_r = \arg\min_{\mathbf{g}} \mathcal{Q}_r(Y_{T_r}, S_{T_r}; \mathbf{g})$, where $\mathbf{g} \in \mathbb{C}^{M(L'+1)}$ and

$$\mathcal{Q}_r(Y_{T_r}, S_{T_r}; \mathbf{g}) = \| Y_{T_r} - \mathcal{H}_{L'}(S_{T_r})\,\mathbf{g}\|^2. \quad (2)$$

The problem of finding the statistically optimal solution is much more complicated in the blind context. Many recently proposed simplified approaches are based on the second-order analysis of the noise-free *space-time covariance matrix* of finite order N: $\hat{\mathbf{R}}_N = (T-N)^{-1}\sum_{t=N}^{T} Y_N(t)\,Y_N(t)^H$, where $Y_N(t) \triangleq [y(t)^T, \ldots, y(t-N)^T]^T$ is (stacked) space-time observation. In particular the noise subspace (NS) method [8] is based on unique correspondence between the cannel $h(z)$ and dominant (signal) subspace of $\mathbf{R}_N = \mathbb{E}\{\hat{\mathbf{R}}_N\}$. Indeed, model (1) may be written in the algebraic form $Y_N(t) = \mathcal{T}_N(h)\,S_N(t) + B_T$, where $S_N \triangleq [s(t)^T, \ldots, s(t-N-L)^T]^T$, $B_N \triangleq [b(t)^T, \ldots, b(t-N)^T]^T$ and *Sylvester resultant matrix*

$$\mathcal{T}_N(h) = \begin{bmatrix} \mathbf{h}(0) & \ldots & \mathbf{h}(L) & 0 & 0 \\ 0 & \ddots & & \ddots & 0 \\ 0 & 0 & \mathbf{h}(0) & \ldots & \mathbf{h}(L) \end{bmatrix}, \quad (3)$$

is a block Toeplitz matrix associated to the polynomial $h(z)$ of (virtual) degree L' and composed of $N+1$ vertical and $N+L'+1$ horizontal $M\times 1$ blocks. Clearly the signal subspace of $\mathbf{R}_N$ coincides with span$\{\,\mathcal{T}_N(h)\,\}$. As shown in [11],[14],[15], under the condition $N \geq L$ we have

$$\text{span}\,\{\,\mathcal{T}_N(g)\,\} = \text{span}\,\{\,\mathcal{T}_N(h)\,\} \Leftrightarrow g(z) = h(z)\,\alpha(z), \quad (4)$$

here $g(z)$ is $M\times 1$ polynomial vector and $\alpha(z)$ is a scalar polynomial. Denote $\mathbf{\Pi}_N$ the orthogonal projector onto the orthogonal (noise) subspace of $\mathbf{R}_N$. According to (4), $\mathbf{\Pi}_N\,\mathcal{T}_N(g) = 0$ implies $g(z) = h(z)\,\alpha(z)$. If the true order is known (*e.g.* it is possible to choose $L' = L$), one can check that $g(z)$ and $h(z)$ coincide up to a scalar constant factor *i.e.* the true channel is identifiable up to scale. In the original version [8] the empirical channel is found as $\hat{\mathbf{h}}_{NS} = \arg\min_{\|\mathbf{g}\|=1}\left\{\left\|\,\hat{\mathbf{\Pi}}_N\,\mathcal{T}_N(g)\,\right\|_F\right\}$, where $\hat{\mathbf{\Pi}}_N$ is the empirical noise subspace projector conventionally obtained from $\hat{\mathbf{R}}_N$. The authors of [10], [11] propose to consider more general class of *weighted* estimators obtained via replacing quadratic form $\|\,\hat{\mathbf{\Pi}}_N\,\mathcal{T}_N(g)\,\|_F$ by $\|\,\hat{\mathbf{\Pi}}_N\,\mathcal{T}_N(g)\,\|_W$, where W is the non-negative matrix providing the minimal asymptotic error variance. Numerical implementation of this estimator requires minor revisions of $\hat{\mathbf{h}}_{NS}$. First of all partition matrix $\mathbf{\Pi}_N$ in $M(N+1)\times M$ blocks: $\mathbf{\Pi}_N = [\mathbf{\Pi}_N(0), \ldots, \mathbf{\Pi}_N(N)]$ and construct an $M(N+1)\times M$ polynomial matrix $\Pi_N(z) \triangleq \sum_{\tau=0}^{N}\mathbf{\Pi}_N(\tau)\,z^{-\tau}$. We adopt the same notation for $\hat{\mathbf{\Pi}}_N$. Now associate to any finite order polynomial $P(z) = \sum_{\tau=0}^{n}\mathbf{P}(\tau)\,z^{-\tau}$ the block-wise transposed Sylvester matrix of order N: $\mathcal{B}_N(P) \triangleq \mathcal{T}_N(P^H)^H$, where $P^H(z) \triangleq \sum_{\tau=0}^{n}\mathbf{P}(\tau)^H\,z^{-\tau}$. One can check that $\text{vec}\left\{\hat{\mathbf{\Pi}}_N\,\mathcal{T}_N(g)\right\} = \mathcal{B}_L(\hat{\Pi}_N)\,\mathbf{g}$, where vec$\{.\}$ is a column vectorization operation: $\text{vec}\,\{\mathbf{P}\} \triangleq [\mathbf{P}_1^T, \ldots, \mathbf{P}_q^T]^T$ for any $p\times q$ matrix $\mathbf{P}$. The weighted quadratic criterion may be written as $\hat{\mathbf{h}}_W = \arg\min_{\|\mathbf{g}\|=1}\mathcal{Q}_W(\hat{\mathbf{R}}_N;\mathbf{g})$, where

$$\mathcal{Q}_W(\hat{\mathbf{R}}_N;\mathbf{g}) = \mathbf{g}^H\,\mathcal{B}_L(\hat{\Pi}_N)^H\,W\,\mathcal{B}_L(\hat{\Pi}_N)\,\mathbf{g}. \quad (5)$$

Here notation $\mathcal{Q}_W(\hat{\mathbf{R}}_N;\mathbf{g})$ underlines that $\hat{\mathbf{h}}_W$ is a function of the empirical second order statistics of $\{y(t)\}_{t\in\mathbf{Z}}$. As a matter of fact, minimization problem (5) yields somehow scaled true channel $\hat{\mathbf{h}}_W = \alpha\,\mathbf{h}$ in the noise-free case, this opportunity owing to the perfect subspace identifiability *i.e.* span$\{\,\hat{\mathbf{R}}_N\,\}$ = span$\{\,\mathcal{T}_N(h)\,\}$ if $\sigma = 0$. However the described estimator looses its consistency once the channel order is overestimated *e.g.* $L' > L$. Indeed according to (4) we have $\hat{h}_W(z) = \alpha(z)\,h(z)$ where $\alpha(z)$ is an arbitrary scalar polynomial of degree $L' - L$. In the following section we discuss how the problem can be treated in the semi-blind context.

4. SEMI-BLIND IDENTIFICATION

A straightforward fusion of the reference-based (2) and blind (5) cost functions leads to a combined criterion:

$$\hat{\mathbf{h}}_W = \arg\min_{\mathbf{g}}\left\{\,\mathcal{Q}_r(Y_{T_r}, S_{T_r};\mathbf{g}) + T\,\mathcal{Q}_W(\hat{\mathbf{R}}_N;\mathbf{g})\,\right\}, \quad (6)$$

here W is some non-negative definite matrix bounded uniformly over the possible values of T. The attenuation T is introduced to balance the information contributed by the the referenced and blind data. The quadratic form in (6) may be regarded as a pseudo-likelihood function obtained by modifying the likelihood function of the blind data. Notice that the column space of $\mathcal{B}_{L'}(\Pi_N)^H\,W$, should coincide with that of $\mathcal{B}_{L'}(\Pi_N)^H$, otherwise the blind contribution of (6) has spurious minima $\mathbf{g} \neq \mathbf{h}$, $W\,\mathcal{B}_{L'}(\Pi_N)\,\mathbf{g} = 0$. Hence we adopt only the *admissible* weightings: $\mathcal{W} \triangleq \{W \geq 0 : \text{span}\{\,\mathcal{B}_{L'}(\Pi_N)^H W\,\} = \text{span}\{\,\mathcal{B}_{L'}(\Pi_N)^H\,\}\}$, see [15]. The problem (6) yields a closed-form solution

$$\hat{\mathbf{h}}_W = \left(T_r\,\mathbb{I}_r + T\,\mathcal{B}_{L'}(\hat{\Pi}_N)^H W \mathcal{B}_{L'}(\hat{\Pi}_N)\right)^{-1}\mathcal{H}_{L'}(S_{T_r})^H Y_{T_r}.$$

here $\mathbb{I}_r = {T_r}^{-1}\,\mathcal{H}_{L'}(S_{T_r})^H\,\mathcal{H}_{L'}(S_{T_r})$ is the autocorrelation matrix of the reference sequence. The solution of purely referenced problem appears to be a particular case: $\hat{\mathbf{h}}_r = \mathcal{H}_{L'}(S_{T_r})^{\#}Y_{T_r}$ hence according to **(H3)**, $\mathbb{E}\left\{(\hat{\mathbf{h}}_r - \mathbf{h})\,(\hat{\mathbf{h}}_r - \mathbf{h})^H\right\} = {T_r}^{-1}\,\sigma\,\mathbb{I}_r^{-1}$. Maximization of this quantity clearly requires that $\mathbb{I}_r = \mathrm{I}_{L'+1}$. Note that such choice is often used in the actual systems, see [16]. We assume for simplicity $\mathbb{I}_r = \mathrm{I}_{L'+1}$. Unlike most of blind techniques, our method also treats the case of the *non-prime* channels *e.g.* when $h(z) = 0$ on some finite set. Let us specify $h(z) = \hbar(z)\,\varepsilon(z)$, where $\hbar(z) \neq 0$ has degree $L_o \leq L$ and $\varepsilon(z) = \sum_{\tau=0}^{L'-L_o}\epsilon(\tau)\,z^{-\tau}$, define also $\epsilon = [\,\epsilon(0), \ldots, \epsilon(L'-L_o)\,]$. The following asymptotic statement holds at $T_r \ll T$.

Theorem 1 *Assume that* **(H1)** - **(H3)** *hold. Then the estimate (6) is asymptotically* ($T_r, T \to \infty$) *Gaussian with* $\hat{h}_W(z) = \sum_{\tau=0}^{L'} \hat{\mathbf{h}}_W(\tau) z^{-\tau}$ *verifying* $\hat{h}_W(z) = \hbar(z)\hat{\varepsilon}(z) + \Delta\hat{h}_W(z)$, $\Delta\hat{h}_W(z) = \sum_{\tau=0}^{L'} \Delta\hat{\mathbf{h}}_W(\tau) z^{-\tau}$, *scalar* $\hat{\varepsilon}(z) = \sum_{\tau=0}^{L'-L_o} \hat{\epsilon}(\tau) z^{-\tau}$, *here* $\hat{\epsilon} = [\hat{\epsilon}(0), \ldots, \hat{\epsilon}(L' - L_o)]$ *and* $\Delta\hat{\mathbf{h}}_W = [\Delta\hat{\mathbf{h}}_W(0)^T, \ldots, \Delta\hat{\mathbf{h}}_W(L')^T]^T$ *are independent:*

$$\sqrt{T_r}(\hat{\epsilon} - \epsilon) \xrightarrow{d} \mathcal{N}_c(0, \mathbf{\Delta}_\epsilon), \ \lim_{T\to\infty} T\mathbb{E}\{\Delta\hat{\mathbf{h}}_W \Delta\hat{\mathbf{h}}_W{}^H\} = \mathbf{\Delta}_h^{(W)},$$

$$\mathbf{\Delta}_\epsilon = \sigma \left(\mathcal{B}_{L'-L_o}(\hbar)^H \mathcal{B}_{L'-L_o}(\hbar)\right)^{-1}, \quad \mathbf{\Delta}_h^{(W)} = \sigma \mathbf{\Delta}_N^{(W)}.$$

$\mathbf{\Sigma}_N \triangleq \lim_{T\to\infty} \frac{T}{\sigma}\mathbb{E}\{\mathcal{B}_{L'}(\hat{\Pi}_N - \Pi_N)\mathbf{h}\mathbf{h}^H \mathcal{B}_{L'}(\hat{\Pi}_N - \Pi_N)^H\}$, *define* $W_\gamma = (\mathbf{\Sigma}_N + \gamma \mathrm{I})^{-1}, \gamma \geq 0$. *Then there exists a lower bound* $\mathbf{\Delta}_N$ *:*

$$\mathbf{\Delta}_N^{(W)} \geq \mathbf{\Delta}_N \ \ \forall\, W \in \mathcal{W}, \qquad \lim_{\gamma\to 0} \mathbf{\Delta}_N^{(W_\gamma)} = \mathbf{\Delta}_N.$$

Analytic expressions for $\mathbf{\Delta}_N^{(W)}$ and $\mathbf{\Delta}_N$ are omitted because of space limitations. The essential conclusion of theorem 1 is that the estimation error consists of two contributions (*e.g.* $\hbar(z)(\hat{\varepsilon}(z) - \varepsilon(z))$ and $\Delta\hat{h}_W(z)$) converging at different rates. Moreover, the weighting choice affects only the "most convergent" part $\Delta\hat{h}_W(z)$. Actually the optimal solution requires the consistent estimate of $\mathbf{\Sigma}_N$ which is available only if the true order is known. In most of cases with $L' > L$, the *plain* weighting $W = \mathrm{I}$ yields rather good results even compared to the optimally weighted technique (*i.e.* calculated for the true order). Clearly the choice of W is not critical if $L' > L$. Indeed the total performance is dominated by the relatively slow convergence of $\hat{\varepsilon}(z)$. Let us focuse on the *regular* case characterized by the known order and inter-channel disparity (*e.g.* $L' = L_o = L$):

$$\hat{h}_W(z) = h(z)\,\hat{\epsilon} + \Delta\hat{h}_W(z), \qquad \hat{\epsilon} \xrightarrow{d} 1. \tag{7}$$

Note that imperfect knowledge of the scale factor $\hat{\epsilon}$ is fairly important for the signal extraction provided that $\hat{h}_W(z) \xrightarrow{P} h(z)$. Hence the actual performance is fully determined by $\mathbf{\Delta}_N^{(W)}$. The optimal weighting presented in theorem 1 is consistent in this case and shown equivalent to $W = (\mathbf{\Sigma}_N)^{\#}$, thus the lower bound $\mathbf{\Delta}_N$ is asymptotically achievable. This bound admits a close-form approximation for the large values of the analyzed order N.

Corollary 1.1 *Assume the conditions of theorem 1 and* $L' = L_o = L$. *Then*

$$\lim_{N\to\infty} (\mathbf{\Delta}_N)^{\#} = \frac{1}{2\pi} \int_{-\pi}^{\pi} \frac{\varrho(\omega)}{1 + \sigma[\varrho(\omega) h(e^{\imath\omega})^H h(e^{\imath\omega})]^{-1}} \times \left[E_L(e^{\imath\omega}) E_L(e^{\imath\omega})^H\right] \otimes \Pi(e^{\imath\omega})\, d\omega,$$

here $\Pi(e^{\imath\omega})$ *is the projector onto* null$\{\, h(e^{\imath\omega}) \,\}$.

Note that the performance of (6) in the regular case is determined by the unreferenced data length T *i.e.* the finite estimation accuracy of the second order statistics $\hat{\mathbf{R}}_N$. However it is also important to how "good" the second order information is exploited *i.e.* establish the relative efficiency of (6) within the class of second order techniques. Actually the second order asymptotic efficiency might be reached in the blind case using the *covariance matching* technique, [17]. However the algorithmic solution of this technique yields rather complex numerical implementation. We study here the asymptotic lower bound achievable in the regular case due to simultaneous use of the referenced data Y_{T_r} and the empirical second order statistics $\hat{\mathbf{R}}_N$.

Theorem 2 *Assume that the conditions of corollary 1.1 hold and* $\hat{\mathbf{h}}_{\mathcal{F}} = \mathcal{F}(Y_{T_r}, \hat{\mathbf{R}}_N)$ *a consistent estimator of* $\mathbf{h}$. *Then* $\hat{h}_{\mathcal{F}}(z) \triangleq \sum_{\tau=0}^{L} \hat{\mathbf{h}}_{\mathcal{F}}(\tau) z^{-\tau}$, $\hat{h}_{\mathcal{F}}(z) = h(z)\,\hat{\epsilon}_{\mathcal{F}} + \Delta\hat{h}_{\mathcal{F}}(z)$ *with* $\hat{\epsilon}_{\mathcal{F}} \xrightarrow{d} 1$, *here* $\Delta\hat{h}_{\mathcal{F}}(z) \triangleq \sum_{\tau=0}^{L} \Delta\hat{\mathbf{h}}_{\mathcal{F}}(\tau) z^{-\tau}$ *such that* $\Delta\hat{\mathbf{h}}_{\mathcal{F}} = [\Delta\hat{\mathbf{h}}_{\mathcal{F}}(0)^T, \ldots, \Delta\hat{\mathbf{h}}_{\mathcal{F}}(L)^T]^T$ *verifies*

$$\lim_{T,N\to\infty} T\,\mathbb{E}\left\{\Delta\hat{\mathbf{h}}_{\mathcal{F}}\, \Delta\hat{\mathbf{h}}_{\mathcal{F}}{}^H\right\} \geq \sigma\,\mathbf{\Delta}_N.$$

Consequently (6) is an asymptotic *minimum variance* semi-blind second order estimator. Actually the results of corollary 1.1 and theorem 2 also hold in the blind context *i.e.* for $T_r = 0$ ($\mathcal{Q}_r(.) \equiv 0$) and the family of consistent estimators $\hat{\mathbf{h}}_{\mathcal{G}} = \mathcal{G}(\hat{\mathbf{R}}_N)$. Note that the second order optimality *only* holds when $\varrho(\omega)$ is unknown *i.e.* the empirical knowledge of signal subspace span$\{\, h(e^{\imath\omega}) \,\}$ is the only statistical information about $h(z)$. More comprehensive analysis of this question is deferred to a regular publication.

One should notice that most of the statistical inference in this paper is asymptotic in T and T_r. Meanwhile possibility to operate with relatively short training sequences is of particular importance for the high capacity communication systems. The following statement provides a statistical description of the proposed technique at finite length reference sequences though (infinitely) long unreferenced observation.

Theorem 3 *Let* $\hat{\mathbf{h}}_W$ *be obtained via (6) at some* $W \in \mathcal{W}$. *The empirical transfer function* $\hat{h}_W(z) = \sum_{\tau=0}^{L'} \hat{\mathbf{h}}_W(\tau) z^{-\tau}$ *verifies* $\lim_{T\to\infty} \hat{h}_W(z) = \hbar(z)\,\hat{\varepsilon}(z)$ *with scalar* $\hat{\varepsilon}(z) = \sum_{\tau=0}^{L'-L_o} \hat{\epsilon}(\tau) z^{-\tau}$. *Denote* $\hat{\epsilon} = [\hat{\epsilon}(0), \ldots, \hat{\epsilon}(L' - L_o)]$, *then*

$$\hat{\epsilon} \sim \mathcal{N}_c(\epsilon, \mathbf{\Delta}_*), \quad \mathbf{\Delta}_* = \frac{\sigma}{T_r} [\mathcal{B}_{L'-L_o}(\hbar)^H \mathcal{B}_{L'-L_o}(\hbar)]^{-1}.$$

Apparently the distribution of $\hat{\mathbf{h}}_W$ at short reference sequences and $T \to \infty$ matches the asymptotic result. It is quite easy to show that the distribution of $\hat{\epsilon}$ yields $\lim_{T\to\infty} \mathbb{E}\{\Delta\hat{\mathbf{h}}_W\, \Delta\hat{\mathbf{h}}_W{}^H\} = T_r{}^{-1}\sigma P_h$ where P_h stands for the orthogonal projector onto span$\{\, \mathcal{B}_{L'-L_o}(\hbar) \,\}$. Hence at moderate over-modeling and reasonable inter-channel disparity (*e.g.* $L' - L_o$ is small), our approach offers certain improvement of the referenced one, remind that $\mathbb{E}\left\{(\hat{\mathbf{h}}_r - \mathbf{h})(\hat{\mathbf{h}}_r - \mathbf{h})^H\right\} = T_r{}^{-1}\sigma\,\mathrm{I}_{M(L'+1)}$. The result of theorem 3 also provides some general considerations about the choice of the reference sequence length.

5. NUMERICAL STUDY

According to the standard baseband representation $\{x(t)\}_{t\in\mathbf{Z}}$ is modeled as the output of the $M = 4$ identical half-wavelenght spaced antennas receiving several modes of the transmitted message delayed by the multiples of the Nyquist period. Each propagation path is defined by the slightly perturbed plane wave with the wavenumber ϑ_τ, $\tau = 0, \ldots, L$. and common spatial correlation factor $\nu = 0.99$. These propagation modes are independently driven by the multivariate circular Gaussian distribution with the covariance matrix $\mathbb{E}\{\mathbf{h}_p(\tau)\overline{\mathbf{h}_q(\tau)}\} = \nu^{|p-q|} e^{\imath(p-q)\vartheta_\tau}$. The observation process is contaminated by additive noise quantified here by the average signal-to-noise ratio per receiver: $SNR = \sum_{\tau=0}^{L} \|\mathbf{h}(\tau)\|^2/(\sigma M)$. The transmitted signal consists of BPSK preamble coded by a binary Gold sequence [18] of length $T_r = 31$ followed by the unknown data with QAM-4 modulation. We further apply the semi-blind techniques developed in this paper. The empirical channel

$\hat{h}(z)$ is used to construct a causal channel inverse (*zero-forcing equalizer*) $\hat{e}_h(z)$: $\hat{\delta}(z) = \hat{e}_h(z)\,h(z) \xrightarrow{P} 1$, here $\hat{\delta}(z) = \sum_{\tau=0}^{\infty} \hat{\delta}(\tau)\,z^{-\tau}$, therefore $\underline{\hat{\delta}}(0) \xrightarrow{P} 1$ and $\underline{\hat{\delta}}(\tau) \xrightarrow{P} 0$, $\tau > 0$. The residual deconvolution error (*e.g.* the residual ISI) $\sum_{\tau=1}^{\infty} |\underline{\hat{\delta}}(\tau)|^2$ is regarded here as a measure of the estimation accuracy. Let us consider the environment with five propagation modes ($L = 4$) having random direction parameters ϑ_τ uniformly spaced in $(-\pi;\ \pi)$. We fix the virtual channel order $L' = L+2$ and use the minimal analysis window $N = L'$, see [8] for more details. The unreferenced data length is chosen $T = 300$.

On Fig.1 the residual ISI (in dB) is plotted versus different values of SNR for various techniques. Here the solid line and "-x-" stand for theoretical and simulated ammounts of the residual ISI provided by the plain semi-blind estimator (*e.g.* $W = I$). All simulated quantities are averaged over 100 Monte-Carlo trails and plotted along with the 95% confidence intervals obtained via the bootstrap method. Similarly, dashed line and "-o-" depict theoretical and simulated performances of the optimally weighted version of (6). Note that the true order was exploited to calculate the optimal weighting matrix *i.e.* any realistic implementation with the imperfectly known order generally yields certain degradation. The dash-dotted line and "-*-" stand for theoretical and sample performances of the classical approach (2). One can observe the essential advantage of semi-blind techniques while the choice of weighting matrix is less important in practice as well as in theory. On Fig.2 we compare the plain and weighted estimators at different values of the virtual order and SNR equal to 20 dB. Apparently the relative degradation hardly exceeds 3dB even at rather severe over-modeling. Notice also that the plain estimator is often better than its asymptotically optimal adverse at finite sample sizes.

We finally study the accuracy bound stated in corollary 1.1 by tracing the residual ISI versus the factor N, see Fig.3. This results were obtained at worse spatial diversity ($\vartheta_\tau \in (-0.05\pi;\ 0.05\pi)$), the true channel order *i.e.* $L' = L$ and the same value of SNR. The dash-dotted line stands here for the residual error calculated from corollary 1.1. Apparently the theoretical performance of the optimal technique reaches the bound while the actual finite sample performance is noticeably worse. Meanwhile the plain estimator yields quite good accuracy.

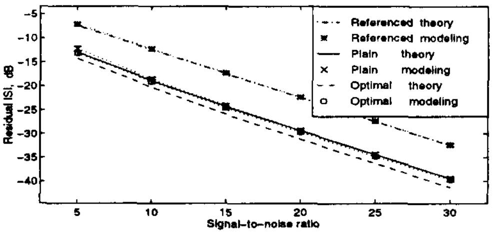

Fig.1. Residual error versus the signal-to-noise ratio.

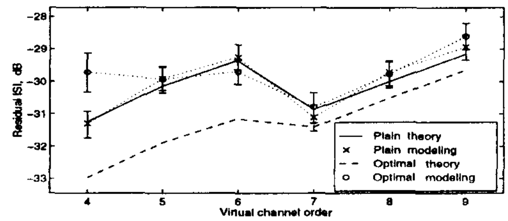

Fig.2. Residual error versus the virtual channel order L'.

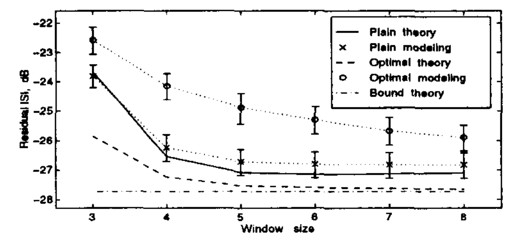

Fig.3. Residual error versus the analysis window size N.

REFERENCES

[1] W. Gardner, "A new method of channel identification", *IEEE Trans. on Comm.*, vol. 39, pp. 813–817, Aug. 1991.

[2] L. Tong, G.Xu, T. Kailath, "Fast blind equalization via antenna arrays", in *Proc. ICASSP'93*, vol. 4, pp. 272–275, 1993.

[3] Z. Ding, Y. Li, "On channel identification based on second order cyclic spectra", *IEEE Trans on Sig. Proc.*, vol. 42, pp. 1260–1264, May 1994.

[4] G. Giannakis, "A linear cyclic correlation approach for blind identification of FIR channels", in *Proc. Asilomar Conf. on Signals, Systems and Computers*, pp. 420–424, Oct. 1994.

[5] H. Liu and G. Xu, "A deterministic approach to blind symbol estimation", *IEEE Sig. Proc. Letters*, pp. 205–207, Dec. 1994.

[6] M. Guirelli, C. Nikias, "A new eigenvector based algorithm for multichannel blind deconvolution of input colored signals", in *Proc. of ICASSP'93*, vol. 4, pp. 448–451, Apr. 1993.

[7] D. Slock, "Blind fractionally-spaced equalization, perfect-reconstruction and filter-banks and multi-channel linear prediction", *in Proc. of ICASSP'94*, vol. 4, pp. 585–588, Adelaide, May 1994.

[8] E. Moulines, P. Duhamel, J.-F. Cardoso, S. Mayrargue, "Subspace methods for the blind identification of multichannel FIR filters", *IEEE Trans. on Sig. Proc.*, vol. 43, No. 2, pp. 516–525, Feb. 1995.

[9] Y. Hua, "Fast maximum likelihood for blind identification of multiple FIR channels", in *IEEE Trans. on Sig. Proc.*, vol. 44, No. 3, pp. 661–671, Feb. 1995.

[10] M. Kristensson, B. Ottersten, "Statistical analysis of a subspace methods for blind channel identification", in *Proc. of ICASSP'96*, vol. 5, pp. 2435–2438, Atlanta, May. 1996.

[11] E. Moulines, J.F. Cardoso, A. Gorokhov, Ph. Loubaton, "Subspace methods for blind identification of SIMO-FIR systems", in *Proc. ICASSP'96*, vol. 5, pp. 2449–2452, Atlanta, May 1996.

[12] T.C. Hsia, *Identification: Least squares methods.* Lexington Books, Lexington, Mass., 1977.

[13] T. Söderstorm, P. Stoica, *System Identification*, Prentice Hall International (UK) Ltd, 1989.

[14] K. Abed-Meraïm, P. Loubaton, E. Moulines, "A subspace algorithm for certain blind identification problems", *To appear in IEEE Trans. on Inf. Theory.*

[15] K. Abed-Meraïm, J.F. Cardoso, Gorokhov, Ph. Loubaton, E. Moulines, "On subspace methods for blind identification of single-input multiple-output FIR systems", to appear, *IEEE Special Issue on Comm.*, 1996.

[16] M. Mouly, M.B. Pautet, *The GSM system for Mobile Communications*, ISBN 2-9507190-0-7, 1994.

[17] H.H. Zeng, S. Zeng, L. Tong, "On the performance of blind equalization using the second-order statistics", *in Proc. of ICASSP'96*, vol. 5, pp. 2427–2430, Atlanta, May 1996.

[18] R. Gold, "Optimal binary sequences for for spread spectrum multiplexing", *IEEE Trans. on Inform. Theory*, vol. IT-13, pp. 619–621, Oct. 1967.