

Robustness of Least-Squares and Subspace Methods for Blind Channel Identification/Equalization with Respect to Effective Channel Undermodeling/Overmodeling

Athanasios P. Liavas, Phillip A. Regalia, *Senior Member, IEEE*, and Jean-Pierre Delmas

Abstract—The least-squares and the subspace methods are two well-known approaches for blind channel identification/equalization. When the order of the channel is known, the algorithms are able to identify the channel, under the so-called *length* and *zero* conditions. Furthermore, in the noiseless case, the channel can be perfectly equalized. Less is known about the performance of these algorithms in the practically inevitable cases in which the channel possesses long tails of “small” impulse response terms. We study the performance of the m th-order least-squares and subspace methods using a perturbation analysis approach. We partition the true impulse response into the m th-order significant part and the tails. We show that the m th-order least-squares or subspace methods estimate an impulse response that is “close” to the m th-order significant part. The closeness depends on the diversity of the m th-order significant part and the size of the tails. Furthermore, we show that if we try to model not only the “large” terms but also some “small” ones, then the quality of our estimate may degrade dramatically; thus, we should avoid modeling “small” terms. Finally, we present simulations using measured microwave radio channels, highlighting potential advantages and shortcomings of the least-squares and subspace methods.

Index Terms—Communications, equalization, multichannel system identification.

I. INTRODUCTION

INTERSYMBOL interference (ISI) is the distortion introduced by practically all channels during signal transmission. Adaptive channel equalization has been a successful technique toward the elimination of ISI. Traditional implementations of adaptive equalizers are based on the periodic transmission of a known training sequence, which permits the identification and/or equalization of the channel. However, there are important applications, such as digital TV broadcasting, in which the use of a training sequence is very costly. In these cases, blind equalization techniques have proved viable alternatives.

It is well established that when the receiver’s matched filter output is sampled at the symbol rate, the resulting sequence

is *stationary*, and blind channel identification/equalization techniques must use, implicitly or explicitly, higher (than second) order statistics (HOS) in order to identify mixed phase channels. On the other hand, if the matched filter is sampled faster than the symbol rate (fractionally spaced case), then the resulting sequence is *cyclostationary*, and second-order statistics (SOS) are sufficient for blind identification of most channels. Working with SOS-based instead of HOS-based techniques is advantageous, especially in a time-varying environment, because SOS can be estimated accurately with far fewer data samples than their higher order counterparts.

The recent development of SOS-based blind identification/equalization methods under a single-input/multi-output (SIMO) channel setting [1], derived either from fractional sampling (FS) of the receiver or from the use of an array of sensors at the receiver, has been considered a major breakthrough and has spawned intensive research in the area. As a result, many novel schemes have been developed that can claim exact channel identification/equalization, in the noiseless case, under the so-called *zero forcing* conditions. The most well-known approaches are the least-squares (LS) [2], the subspace (SS) [3], and the linear prediction (LP) [4] methods.

Furthermore, SIMO implementations of implicitly HOS-based blind equalization techniques have resulted in very interesting schemes, such as the FS constant modulus algorithm (CMA), which avoid, in the noiseless case, drawbacks related to their traditional single-input/single-output (SISO) counterparts, such as potentially large equalizer length and potential convergence to local minima [5].

While all the aforementioned methods claim exact channel identification/equalization in the noiseless case, under the zero forcing conditions, their behavior may change dramatically under practically inevitable “less ideal” conditions such as

- the presence of non-negligible additive channel noise;
- the presence of long tails of “small” leading and/or trailing impulse response terms.

The robustness of blind identification/equalization methods to additive channel noise and long channel “tails” is a very critical issue that is directly related to their applicability under “real-world” conditions.

The efforts toward a deeper understanding of the robustness properties of blind identification/equalization techniques, with

Manuscript received December 3, 1997; revised June 9, 1998. This work was supported by the Training and Mobility of Researchers (T.M.R.) Program of the European Commission under Contract ERBFMBICT960659. The associate editor coordinating the review of this paper and approving it for publication was Dr. Xiang-Gen Xia.

The authors are with the Département Signal et Image, Institut National des Télécommunications, Evry, France (e-mail: liavas@sim.int-evry.fr; regalia@sim.int-evry.fr; delmas@sim.int-evry.fr).

Publisher Item Identifier S 1053-587X(99)03681-8.

respect to channel noise and/or channel “tails,” are in the early stages [6]–[10] and admit the following shortcomings.

- It is not always clear to what extent the results apply in practice, due to various approximations and/or simplifications during the development of the analysis.
- The approaches often involve complicated techniques, lacking the engineering insight necessary for a lucid understanding of the various phenomena and for suggesting improved solutions.

These shortcomings have impeded a clear understanding of the behavior of blind identification/equalization methods under realistic operating conditions.

In this paper, we consider the behavior of the m th-order LS and SS algorithms in the two-channel, noiseless, exact statistics case. In Section II, we present the SIMO channel setting, and we review the LS and SS blind channel identification methods for the known channel order case. In Section III, in order to study the behavior of the m th-order LS and SS methods, we decompose the true impulse response into the m th-order significant part and the tails; we show that the m th-order LS and SS methods estimate an impulse response that is close to the m th-order significant part; the closeness depends on the diversity of the m th-order significant part and the size of the tails. Furthermore, we show that when we try to model not only the “large” terms but also some “small” ones, then the quality of our estimate may degrade dramatically. Thus, we should avoid modeling “small” terms. We then assess the performance of the $(m-1)$ st-order “zero forcing” equalizers. In Section IV, we check our theoretical results by simulations; furthermore, we present simulations using measured microwave radio channels, and we highlight potential advantages and shortcomings of the methods in realistic cases.

II. LS AND SS METHODS FOR BLIND CHANNEL IDENTIFICATION/EQUALIZATION

A. Two-Channel Model

In this section, we describe the basic steps of the LS and the SS methods for blind channel identification for the single-input/two-output channel setting, which is presented in Fig. 1; this setting can be obtained by channel oversampling by a factor of 2, which is quite common in telecommunications, or by using two sensors at the receiver. If the true channel order is M , the output of the j th channel $x_i^{(j)}$, for $j = 1, 2$, is given by

$$x_i^{(j)} = s_i \otimes h_i^{(j)} + n_i^{(j)} = \sum_{k=0}^M h_k^{(j)} s_{i-k} + n_i^{(j)}$$

where

- \otimes convolution operator;
- s_i input sequence, which is assumed to be zero-mean unit-variance i.i.d. sequence;
- $h_k^{(j)}$ impulse response of the j th channel;
- $n_i^{(j)}$ additive white channel noise.

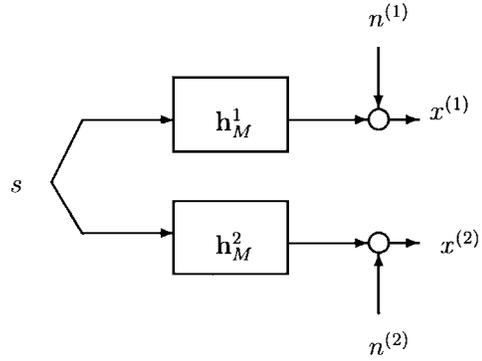


Fig. 1. Single-input/two-output channel setting.

We denote the impulse response of the j th channel, $j = 1, 2$, by $\mathbf{h}_M^j \triangleq [h_0^{(j)} \dots h_M^{(j)}]^T$ and the entire channel parameter vector by $\mathbf{h}_M \triangleq [\mathbf{h}_M^1; \mathbf{h}_M^2]$.

By stacking the $(L+1)$ most recent samples of each channel, we construct the data vector

$$\mathbf{x}_L(i) \triangleq [x_i^{(1)}, \dots, x_{i-L}^{(1)}, x_i^{(2)}, \dots, x_{i-L}^{(2)}]^T$$

which can be expressed as

$$\mathbf{x}_L(i) = \mathcal{H}_L(\mathbf{h}_M) \mathbf{s}_{L+M}(i) + \mathbf{n}_L(i)$$

using input and noise vectors

$$\begin{aligned} \mathbf{s}_{L+M}(i) &\triangleq [s_i, \dots, s_{i-L-M}]^T, \\ \mathbf{n}_L(i) &\triangleq [n_i^{(1)}, \dots, n_{i-L}^{(1)}, n_i^{(2)}, \dots, n_{i-L}^{(2)}]^T. \end{aligned}$$

The convolution matrix $\mathcal{H}_L(\mathbf{h}_M)$ is defined as

$$\begin{aligned} \mathcal{H}_L(\mathbf{h}_M) &\triangleq \begin{bmatrix} \mathcal{F}_L(\mathbf{h}_M^1) \\ \mathcal{F}_L(\mathbf{h}_M^2) \end{bmatrix} \\ \mathcal{F}_L(\mathbf{h}_M^i) &\triangleq \begin{bmatrix} h_0^{(i)} & \dots & \dots & h_M^{(i)} \\ & \ddots & & \vdots \\ & & h_0^{(i)} & \dots & \dots & h_M^{(i)} \end{bmatrix}_{(L+1) \times (M+L+1)}. \end{aligned}$$

In order to review the LS and the SS methods for the identification of \mathbf{h}_M , we consider first the case $L = M$. We assume that the subchannels do not share common zeros, guaranteeing their identifiability.

B. LS Methods

The SIMO channel structure implies that in the noiseless case [2]

$$x_i^{(1)} = h_i^{(1)} \otimes s_i, \quad x_i^{(2)} = h_i^{(2)} \otimes s_i$$

yielding

$$h_i^{(2)} \otimes x_i^{(1)} = h_i^{(2)} \otimes h_i^{(1)} \otimes s_i = h_i^{(1)} \otimes x_i^{(2)}. \quad (1)$$

If we define

$$\mathbf{T}_M \triangleq \begin{bmatrix} \mathbf{0} & \mathbf{I}_{M+1} \\ -\mathbf{I}_{M+1} & \mathbf{0} \end{bmatrix}$$

where \mathbf{I}_{M+1} is the $(M+1)$ -dimensional identity matrix, then (1) can be expressed, for all time indices i , as

$$\mathbf{h}_M^T \mathbf{T}_M \mathbf{x}_M(i) = 0.$$

The LS estimate is given by

$$\begin{aligned} \mathbf{h}_M^{LS} &\triangleq \arg \min_{\|\mathbf{h}\|_2=1} E\{\|\mathbf{h}^T \mathbf{T}_M \mathbf{x}_M(i)\|_2^2\} \\ &= \arg \min_{\|\mathbf{h}\|_2=1} \mathbf{h}^T \mathbf{T}_M \mathbf{R}_M \mathbf{T}_M^T \mathbf{h} \end{aligned}$$

where

$$\mathbf{R}_M \triangleq E[\mathbf{x}_M(i) \mathbf{x}_M^T(i)]$$

is the data autocorrelation matrix. As shown in [11], in this case, the LS estimate is given by

$$\mathbf{h}_M^{LS} = \mathbf{T}_M \mathbf{n}_M$$

where \mathbf{n}_M is the eigenvector associated with the smallest eigenvalue of \mathbf{R}_M . In the noiseless or the temporally and spatially white additive noise cases, the two-channel LS method identifies the unknown channel, that is,

$$\mathbf{h}_M^{LS} = \mathbf{h}_M.$$

C. SS Methods

Subspace methods are based on the orthogonality of the so-called noise and signal subspaces [3]. In the two-channel case, the noise subspace is spanned by \mathbf{n}_M , whereas the signal subspace is spanned by the columns of $\mathcal{H}_M(\mathbf{h}_M)$. Thus

$$\mathbf{n}_M^T \mathcal{H}_M(\mathbf{h}_M) \equiv \mathbf{h}_M^T \mathcal{H}_M(\mathbf{n}_M) = \mathbf{0}^T$$

where we have used the commutativity property of the convolution. The subspace estimate \mathbf{h}_M^{SS} is computed as

$$\mathbf{h}_M^{SS} \triangleq \arg \min_{\|\mathbf{h}\|_2=1} \mathbf{h}^T \mathcal{H}_M(\mathbf{n}_M) \mathcal{H}_M^T(\mathbf{n}_M) \mathbf{h}.$$

As shown in [11], in the two-channel case the LS and SS estimates coincide with probability 1.

D. Zero Forcing Equalization

Having identified the channel \mathbf{h}_M , we can equalize it perfectly, in the noiseless case, by using the zero forcing equalizers of order $(M-1)$ for delays $i = 0, \dots, 2M-1$ as

$$\mathbf{g}_{M-1,i}^{ZF} = \mathcal{H}_{M-1}^{-T}(\mathbf{h}_M) \mathbf{e}_i \quad (2)$$

where \mathbf{e}_i denotes the vector with 1 at the $(i+1)$ st position and zeros elsewhere.

III. UNDERMODELING/OVERMODELING

A. The Framework

In the previous section, we reviewed ways for the identification/equalization of an unknown channel under the SIMO framework, assuming that the channel order is given. We must appreciate, nonetheless, that physical microwave radio channel lengths are not unambiguously determined, due to possibly long “heads” and “tails” of small impulse response terms [7].

In order to study the behavior of the m th-order LS/SS method, we partition the channel impulse response into the following parts:

- 1) the m th-order significant part, which is usually found near the middle of the impulse response and contains $(m+1)$ “consecutive large” terms; if the channel possesses less than $(m+1)$ “large” terms, then the m th-order significant part contains some “small” terms as well;
- 2) the tails, which is the complementary part to the m th-order significant part; this part is usually composed of “small” leading and/or trailing impulse response terms; if the channel possesses more than $(m+1)$ “large” terms, then the tails contain some “large” terms as well.

This partition can be expressed notationally, for $0 \leq m_1 < m_2 \triangleq m_1 + m \leq M$, as [10], [12]

$$\mathbf{h}_M = \mathbf{h}_{m_1, m_2}^z + \mathbf{d}_{m_1, m_2}^z \quad (3)$$

where

$$\mathbf{h}_{m_1, m_2}^z \triangleq \begin{bmatrix} \mathbf{h}_{m_1, m_2}^{z1} \\ \mathbf{h}_{m_1, m_2}^{z2} \end{bmatrix}, \quad \mathbf{d}_{m_1, m_2}^z \triangleq \begin{bmatrix} \mathbf{d}_{m_1, m_2}^{z1} \\ \mathbf{d}_{m_1, m_2}^{z2} \end{bmatrix} \quad (4)$$

with

$$\mathbf{h}_{m_1, m_2}^{zj} \triangleq \underbrace{[0 \dots 0]_{m_1}}_{m_1} \underbrace{[h_{m_1}^{(j)} \dots h_{m_2}^{(j)}]_{m+1}}_{m+1} \underbrace{[0 \dots 0]_{M-m_2}}_{M-m_2}^T, \quad j = 1, 2 \quad (5)$$

$$\mathbf{d}_{m_1, m_2}^{zj} \triangleq \underbrace{[h_0^{(j)} \dots h_{m_1-1}^{(j)}]_{m_1}}_{m_1} \underbrace{[0 \dots 0]_{m+1}}_{m+1} \underbrace{[h_{m_2+1}^{(j)} \dots h_M^{(j)}]_{M-m_2}}_{M-m_2}^T, \quad j = 1, 2. \quad (6)$$

With \mathbf{h}_{m_1, m_2} , we denote the corresponding nonzero-padded vectors, i.e.,

$$\mathbf{h}_{m_1, m_2} \triangleq \begin{bmatrix} \mathbf{h}_{m_1, m_2}^1 \\ \mathbf{h}_{m_1, m_2}^2 \end{bmatrix}, \quad \mathbf{h}_{m_1, m_2}^j \triangleq [h_{m_1}^{(j)} \dots h_{m_2}^{(j)}]^T, \quad j = 1, 2. \quad (7)$$

In the sequel, we study the m th-order LS and SS methods, and we explore the relationship between the “identified” m th-order impulse response and the true M th-order channel \mathbf{h}_M .

Our principal concern is deducing the mean asymptotic performance obtainable using the m th-order LS and SS methods. All results, therefore, are expressed in terms of true second-order statistics. An important practical issue, of course, is to gauge the variances of different estimators versus the data length and their influence on the equalization quality. By definition of variance, however, such calculations require knowledge of mean values of estimators. In order to keep a manageable presentation, we shall not pursue a variance analysis for the m th-order case in this work. We believe, however, that our results, particularly those related to problem conditioning, will prove very useful in subsequent variance analyses. We remark finally that spatially and temporally white channel noise does not bias the mean asymptotic solution obtained using the LS and SS methods (although such noise will, of course, increase estimation variance) and, hence, will not alter expressions concerning mean asymptotic performance. For simplicity, we thus remove channel noise from our analysis.

B. m th-Order LS/SS Blind Channel Identification

If the true channel order is m and its impulse response is \mathbf{h}_{m_1, m_2} , then the autocorrelation matrix of \mathbf{x}_m , which is denoted by \mathbf{R}_m , provides sufficient information for the identification of the \mathbf{h}_{m_1, m_2} via the sequence of computations

$$\mathbf{R}_m = \mathcal{H}_m(\mathbf{h}_{m_1, m_2}) \mathcal{H}_m^T(\mathbf{h}_{m_1, m_2}) \rightarrow \mathbf{n}_m \rightarrow \mathbf{h}_{m_1, m_2} = \mathbf{T}_m \mathbf{n}_m$$

where \mathbf{n}_m denotes the minimum eigenvector of \mathbf{R}_m . If the true channel impulse response is \mathbf{h}_{m_1, m_2}^z , then it is easy to show that the autocorrelation matrix of \mathbf{x}_m remains \mathbf{R}_m because

$$\mathcal{H}_m(\mathbf{h}_{m_1, m_2}^z) \mathcal{H}_m^T(\mathbf{h}_{m_1, m_2}^z) = \mathcal{H}_m(\mathbf{h}_{m_1, m_2}) \mathcal{H}_m^T(\mathbf{h}_{m_1, m_2})$$

meaning that, in this particular case in which the subchannels possess common zeros at infinity, the m th-order LS or SS methods “identify” the nonzero part of \mathbf{h}_{m_1, m_2}^z , namely \mathbf{h}_{m_1, m_2} . This result is directly related to the blind nature of the algorithm, that is, the exploitation of solely channel output statistics, and will prove very useful in the sequel.

Now, let us consider what happens when the true impulse response is \mathbf{h}_M , with $\|\mathbf{h}_M\|_2 = 1$, under the assumption that \mathbf{d}_{m_1, m_2}^z is “small,” i.e.,

$$\|\mathbf{d}_{m_1, m_2}^z\|_2 = \epsilon_m, \quad \text{with } \epsilon_m \ll 1. \quad (8)$$

In this case

$$\|\mathbf{h}_{m_1, m_2}^z\|_2 = \|\mathbf{h}_{m_1, m_2}\|_2 = \sqrt{1 - \epsilon_m^2} \equiv \gamma_m. \quad (9)$$

The autocorrelation matrix of \mathbf{x}_m is

$$\begin{aligned} \tilde{\mathbf{R}}_m &= \mathcal{H}_m(\mathbf{h}_M) \mathcal{H}_m^T(\mathbf{h}_M) \\ &= \mathcal{H}_m(\mathbf{h}_{m_1, m_2}^z + \mathbf{d}_{m_1, m_2}^z) \mathcal{H}_m^T(\mathbf{h}_{m_1, m_2}^z + \mathbf{d}_{m_1, m_2}^z) \\ &= \{\mathcal{H}_m(\mathbf{h}_{m_1, m_2}^z) + \mathcal{H}_m(\mathbf{d}_{m_1, m_2}^z)\} \\ &\quad \times \{\mathcal{H}_m^T(\mathbf{h}_{m_1, m_2}^z) + \mathcal{H}_m^T(\mathbf{d}_{m_1, m_2}^z)\} \\ &= \mathbf{R}_m + \mathbf{E}_m \end{aligned}$$

where \mathbf{E}_m denotes the resulting perturbation. The m th-order LS/SS blind channel identification method “identifies” $\tilde{\mathbf{h}}_{m_1, m_2}$ through the sequence of computations

$$\tilde{\mathbf{R}}_m \rightarrow \tilde{\mathbf{n}}_m \rightarrow \tilde{\mathbf{h}}_{m_1, m_2} = \mathbf{T}_m \tilde{\mathbf{n}}_m$$

where the variables with tildes denote perturbed quantities. Here, the vector $\tilde{\mathbf{n}}_m$ is the minimum eigenvector of $\tilde{\mathbf{R}}_m$.

At first, we address how close $\tilde{\mathbf{n}}_m$ is to \mathbf{n}_m . For this purpose, we may consider $\tilde{\mathbf{R}}_m$ as a perturbation of \mathbf{R}_m and apply eigenvector perturbation results. However, since \mathbf{n}_m and $\tilde{\mathbf{n}}_m$ are the minimum right singular vectors of $\mathcal{H}_m^T(\mathbf{h}_{m_1, m_2}^z)$ and $\mathcal{H}_m^T(\mathbf{h}_M)$, respectively, it is preferable to use singular vector perturbation results instead. We thus consider $\mathcal{H}_m^T(\mathbf{h}_M)$ as a perturbation of $\mathcal{H}_m^T(\mathbf{h}_{m_1, m_2}^z)$, and we give an upper bound for $\|\mathbf{n}_m - \tilde{\mathbf{n}}_m\|_2$.

We recall that under the *no common zero* assumption, $\text{rank}(\mathcal{H}_m^T(\mathbf{h}_{m_1, m_2}^z)) = 2m + 1$, yielding $\sigma_{2(m+1)}(\mathcal{H}_m^T(\mathbf{h}_{m_1, m_2}^z)) = 0$, with associated right singular vector \mathbf{n}_m ; in this case, \mathbf{n}_m defines the null space of

$\mathcal{H}_m^T(\mathbf{h}_{m_1, m_2}^z)$. We denote by δ_m the smallest nonzero singular value of $\mathcal{H}_m^T(\mathbf{h}_{m_1, m_2}^z)$, i.e.,

$$\delta_m \triangleq \sigma_{2m+1}(\mathcal{H}_m^T(\mathbf{h}_{m_1, m_2}^z)). \quad (10)$$

Since δ_m measures the distance in the matrix 2-norm of $\mathcal{H}_m^T(\mathbf{h}_{m_1, m_2}^z)$ from the matrices of rank $2m$ [13, p. 73], thus violating our assumption concerning its rank, it may be interpreted as a measure of *diversity* of the channel \mathbf{h}_{m_1, m_2} .

Using (3), we identify the perturbation on $\mathcal{H}_m^T(\mathbf{h}_{m_1, m_2}^z)$ as

$$\Delta_m \triangleq \mathcal{H}_m^T(\mathbf{h}_M) - \mathcal{H}_m^T(\mathbf{h}_{m_1, m_2}^z) = \mathcal{H}_m^T(\mathbf{d}_{m_1, m_2}^z)$$

which, using the matrix 2-norm/ F -norm inequality [13, p. 57], the structure of $\mathcal{H}_m^T(\mathbf{d}_{m_1, m_2}^z)$, and (8), yields

$$\begin{aligned} \|\Delta_m\|_2 &= \|\mathcal{H}_m^T(\mathbf{d}_{m_1, m_2}^z)\|_2 \leq \|\mathcal{H}_m^T(\mathbf{d}_{m_1, m_2}^z)\|_F \\ &= \sqrt{m+1} \|\mathbf{d}_{m_1, m_2}^z\|_2 = \epsilon_m \sqrt{m+1} \equiv \mathcal{E}_m. \end{aligned} \quad (11)$$

In the sequel, we shall use the concept of the angle between the unit 2-norm singular vectors \mathbf{n}_m and $\tilde{\mathbf{n}}_m$, which is defined as [14, p. 15]

$$\begin{aligned} \angle(\mathbf{n}_m, \tilde{\mathbf{n}}_m) &\triangleq \arccos |\mathbf{n}_m^T \tilde{\mathbf{n}}_m| \quad \text{with} \\ 0 &\leq \angle(\mathbf{n}_m, \tilde{\mathbf{n}}_m) \leq \frac{\pi}{2}. \end{aligned} \quad (12)$$

By definition, $\angle(\mathbf{n}_m, \tilde{\mathbf{n}}_m)$ is a measure of the distance between the subspaces spanned by \mathbf{n}_m and $\tilde{\mathbf{n}}_m$, and the cosine between \mathbf{n}_m and $\tilde{\mathbf{n}}_m$ is non-negative.

We can now proceed to the following theorem, which provides an upper bound for $\|\mathbf{n}_m - \tilde{\mathbf{n}}_m\|_2$.

Theorem 1: Assume that $\text{rank}(\mathcal{H}_m^T(\mathbf{h}_{m_1, m_2}^z)) = 2m + 1$. Denote by \mathbf{n}_m the minimum right singular vector of $\mathcal{H}_m^T(\mathbf{h}_{m_1, m_2}^z)$, by δ_m the minimum nonzero singular value of $\mathcal{H}_m^T(\mathbf{h}_{m_1, m_2}^z)$, and by $\tilde{\mathbf{n}}_m$ the minimum right singular vector of $\mathcal{H}_m^T(\mathbf{h}_M)$. If $\mathcal{E}_m \leq (\delta_m/2)$, then

$$\|\mathbf{n}_m - \tilde{\mathbf{n}}_m\|_2 \leq 2\sqrt{2} \frac{\mathcal{E}_m}{\delta_m} \equiv \mathcal{D}_m. \quad (13)$$

Proof: Under the assumptions of the theorem, we obtain [15, p. 267, ex. 2], [16]

$$\sin \angle(\mathbf{n}_m, \tilde{\mathbf{n}}_m) \leq \frac{\mathcal{E}_m}{\delta_m - \mathcal{E}_m} \leq 2 \frac{\mathcal{E}_m}{\delta_m} \equiv \alpha_m. \quad (14)$$

Now, for the unit 2-norm vectors \mathbf{n}_m and $\tilde{\mathbf{n}}_m$, we obtain

$$\|\mathbf{n}_m - \tilde{\mathbf{n}}_m\|_2^2 = 2(1 - \cos \angle(\mathbf{n}_m, \tilde{\mathbf{n}}_m))$$

which, using (14), gives

$$\|\mathbf{n}_m - \tilde{\mathbf{n}}_m\|_2^2 \leq 2(1 - \sqrt{1 - \alpha_m^2}) \leq 2\alpha_m^2.$$

This verifies relation (13), to prove the theorem. \blacksquare

Corollary 1: Let \mathbf{h}_{m_1, m_2} be the m th-order significant part of the true channel and $\hat{\mathbf{h}}_{m_1, m_2}$ the estimate of the m th-order LS/SS method. If $\mathcal{E}_m \leq (\delta_m/2)$, then

$$\begin{aligned} \|\mathbf{h}_{m_1, m_2} - \hat{\mathbf{h}}_{m_1, m_2}\|_2 &\leq (1 - \gamma_m) + \mathcal{D}_m \leq \epsilon_m^2 + \mathcal{D}_m \leq \epsilon_m + \mathcal{D}_m \equiv \mathcal{K}_m. \end{aligned} \quad (15)$$

Proof: If $\mathcal{E}_m \leq (\delta_m/2)$, then Theorem 1 holds. Then, since $\mathbf{h}_{m_1, m_2} = \gamma_m \mathbf{T}_m \mathbf{n}_m$, $\tilde{\mathbf{h}}_{m_1, m_2} = \mathbf{T}_m \tilde{\mathbf{n}}_m$, and \mathbf{T}_m is orthogonal, (9) and (13) altogether give

$$\begin{aligned} & \|\mathbf{h}_{m_1, m_2} - \tilde{\mathbf{h}}_{m_1, m_2}\|_2 \\ &= \|(\gamma_m \mathbf{T}_m \mathbf{n}_m - \mathbf{T}_m \tilde{\mathbf{n}}_m) + \mathbf{T}_m (\mathbf{n}_m - \tilde{\mathbf{n}}_m)\|_2 \\ &\leq (1 - \gamma_m) + \mathcal{D}_m \leq \epsilon_m^2 + \mathcal{D}_m \leq \epsilon_m + \mathcal{D}_m \end{aligned}$$

to prove the corollary. ■

Recapitulating, we may say that in order to study the performance of the m th-order LS/SS methods, we partition the impulse response into the m th-order significant part and the tails. Results (13) and (15) are *generic*; that is, they are valid for *every* m , as long as $m \leq M$ and $\mathcal{E}_m \leq (\delta_m/2)$. They imply simply that if the diversity of the m th-order significant part of the channel δ_m is sufficiently large and, at the same time, the size of the tails ϵ_m is sufficiently small, then the m th-order LS and SS methods compute an impulse response that is close to the m th-order significant part of the channel.

Two important questions that cannot be answered by (15) are the following:

- “Is it always possible to find an m such that the m th-order LS/SS method provides a good channel approximation?”
- “When this is possible, how can we find such an m ?”

To answer the first question, extensive experimentation with measured microwave radio channels is required. The second question is the topic of [20].

In the sequel, we show that we should favor, in general, the “smallest possible” m . More specifically, we show that when we try to model not only the “large” impulse response terms but also some “small” ones, then the quality of our estimate may degrade dramatically.

Thus, let us imagine the hypothetical case in which we know *a priori* that the “large” channel terms occupy the m th-order part of the impulse response between indices m_1 and m_2 , implying that \mathbf{d}_{m_1, m_2}^z fulfills (8), whereas we apply the m' th-order LS/SS method, with $m' > m$; hence, $m' = m'_2 - m'_1$, with $m'_1 < m_1$ and/or $m'_2 > m_2$.

In this case, the autocorrelation matrix of $\mathbf{x}_{m'}$ can be expressed as

$$\begin{aligned} \tilde{\mathbf{R}}_{m'} &= \mathcal{H}_{m'}(\mathbf{h}_M) \mathcal{H}_{m'}^T(\mathbf{h}_M) \\ &= \mathcal{H}_{m'}(\mathbf{h}_{m'_1, m'_2}^z + \mathbf{d}_{m'_1, m'_2}^z) \mathcal{H}_{m'}^T(\mathbf{h}_{m'_1, m'_2}^z + \mathbf{d}_{m'_1, m'_2}^z) \\ &= \{\mathcal{H}_{m'}(\mathbf{h}_{m'_1, m'_2}^z) + \mathcal{H}_{m'}(\mathbf{d}_{m'_1, m'_2}^z)\} \\ &\quad \times \{\mathcal{H}_{m'}^T(\mathbf{h}_{m'_1, m'_2}^z) + \mathcal{H}_{m'}^T(\mathbf{d}_{m'_1, m'_2}^z)\}. \end{aligned}$$

From (13) and (15), it becomes clear that the factor that determines the accuracy of the estimation of $\mathbf{h}_{m'_1, m'_2}$ is

$$\delta_{m'} \triangleq \sigma_{2m'+1}(\mathcal{H}_{m'}^T(\mathbf{h}_{m'_1, m'_2}^z)) \quad (16)$$

that is, the minimum nonzero singular value of $\mathcal{H}_{m'}^T(\mathbf{h}_{m'_1, m'_2}^z)$. The next theorem gives a relationship between $\delta_{m'}$ and $h_{m'_1}^{(j)}$, $h_{m'_2}^{(j)}$ for $j = 1, 2$, which provides significant insight into the behavior of the algorithm, when $h_{m'_1}^{(j)}$ and/or $h_{m'_2}^{(j)}$, for $j = 1, 2$ are “small.”

Theorem 2: If $\delta_{m'}$ denotes the minimum nonzero singular value of $\mathcal{H}_{m'}^T(\mathbf{h}_{m'_1, m'_2}^z)$, then

$$\delta_{m'} \leq \min \left(\sqrt{(h_{m'_1}^{(1)})^2 + (h_{m'_1}^{(2)})^2}, \sqrt{(h_{m'_2}^{(1)})^2 + (h_{m'_2}^{(2)})^2} \right). \quad (17)$$

Proof: It is well known that $\delta_{m'}$ is the distance in the matrix 2-norm between the $(2m' + 1) \times 2(m' + 1)$ matrix $\mathcal{H}_{m'}^T(\mathbf{h}_{m'_1, m'_2}^z)$ and the matrices with rank $2m'$. A very simple way to decrease the rank of $\mathcal{H}_{m'}^T(\mathbf{h}_{m'_1, m'_2}^z)$ is to null its first row by adding a perturbation matrix E , with only two nonzero elements at the appropriate positions of the first row, with values $-h_{m'_1}^{(1)}$ and $-h_{m'_1}^{(2)}$. Using the matrix 2-norm/F-norm inequality, we obtain

$$\|E\|_2 \leq \|E\|_F = \sqrt{(h_{m'_1}^{(1)})^2 + (h_{m'_1}^{(2)})^2}.$$

An analogous statement holds for the case in which we null the last row of $\mathcal{H}_{m'}^T(\mathbf{h}_{m'_1, m'_2}^z)$. Since $\delta_{m'}$ equals the minimum 2-norm of a perturbation matrix, which decreases the rank of $\mathcal{H}_{m'}^T(\mathbf{h}_{m'_1, m'_2}^z)$, (17) follows, which proves the theorem. ■

Now, let us consider the implications of Theorem 2 to our study. We recall that

$$\|\mathbf{d}_{m_1, m_2}^z\|_2 = \epsilon_m, \quad \text{with } \epsilon_m \ll 1.$$

Thus, since $h_{m'_1}^{(j)}$ and/or $h_{m'_2}^{(j)}$ for $j = 1, 2$ belong to \mathbf{d}_{m_1, m_2}^z , Theorem 2 implies that

$$\delta_{m'} = \sigma_{2m'+1}(\mathcal{H}_{m'}^T(\mathbf{h}_{m'_1, m'_2}^z)) = O(\epsilon_m).$$

We recall that

$$\sigma_{2(m'+1)}(\mathcal{H}_{m'}^T(\mathbf{h}_{m'_1, m'_2}^z)) = 0$$

and that the perturbation $\Delta_{m'} \triangleq \mathcal{H}_{m'}^T(\mathbf{d}_{m'_1, m'_2}^z)$ is $O(\epsilon_{m'})$. Physical considerations imply that for $m' \geq m$

$$\|\mathbf{d}_{m'_1, m'_2}^z\|_2 \approx \|\mathbf{d}_{m_1, m_2}^z\|_2 \quad (18)$$

giving

$$\epsilon_{m'} \approx \epsilon_m. \quad (19)$$

To justify this, recall that now, m is the correct order of the actual significant part of the channel and *not* a hypothesized order; thus, the assumed tails are the actual channel tails; measured microwave radio channels possess long tails of small terms and decrease slowly [7]; this means that neighboring “small” terms are of the same order of magnitude, validating (18).

Hence, $\mathbf{n}_{m'}$, which is the minimum right singular vector of $\mathcal{H}_{m'}^T(\mathbf{h}_{m'_1, m'_2}^z)$, is a typical example of an *ill-conditioned* or *unstable* singular vector [13, p. 430] because the *separation* between its corresponding singular value, i.e., zero, and the remaining singular values, i.e., $\delta_{m'}$, is $O(\epsilon_m)$, which, due to (19), is of the order of the perturbation $\mathcal{H}_{m'}^T(\mathbf{d}_{m'_1, m'_2}^z)$. The fact that minimum right singular vectors $\mathbf{n}_{m'}$ and $\tilde{\mathbf{n}}_{m'}$ and

channels $\mathbf{h}_{m'_1, m'_2}$ and $\tilde{\mathbf{h}}_{m'_1, m'_2}$ are related through the orthogonal transformation $\mathbf{T}_{m'}$ implies that our estimate $\tilde{\mathbf{h}}_{m'_1, m'_2}$ is unstable as well.

In practice, in the majority of the cases, the condition $\mathcal{E}_{m'} \leq (\delta_{m'}/2)$ of Theorem 1 is not satisfied, and the only upper bound we can give for $\sin \angle(\mathbf{n}_{m'}, \tilde{\mathbf{n}}_{m'})$ is 1. This means that we do not have any *a priori* knowledge for the “distance” between $\mathbf{n}_{m'}$ and $\tilde{\mathbf{n}}_{m'}$ and, consequently, for the distance between $\mathbf{h}_{m'_1, m'_2}$ and $\tilde{\mathbf{h}}_{m'_1, m'_2}$. This is analogous to what has been considered as *channel overmodeling* [17], [18]. The role of the “zero impulse response terms” in these studies is played by the “small” terms in our study. In a numerical analysis parlance, we may say that efforts toward modeling “small” impulse response terms lead generically to an *ill-conditioned* or *unstable* problem and, thus, should be avoided.

From these considerations, we may define as *effective channel* that part of the channel containing all the “large” terms; let us denote its order by m^* . A case is called *undermodeled* (resp., *overmodeled*) if the assumed channel order m is smaller (resp. larger) than m^* . In both the undermodeled and overmodeled cases, good effective channel approximation seems difficult; in the former cases, the approximation may be poor due to large undermodeling error, whereas in the latter, it may be poor due to lack of diversity. Even assuming that we know m^* , good effective channel approximation is not guaranteed; the quality of the approximation depends on δ_{m^*} and ϵ_{m^*} . The fact that we *cannot* approximate a channel arbitrarily well by increasing the complexity of our model is probably a significant obstacle toward general applicability of these methods.

C. Zero Forcing Equalization

Having “identified” the m th-order channel $\tilde{\mathbf{h}}_{m_1, m_2}$, we can equalize it perfectly in the noiseless case, for delays $i = 0, \dots, 2m - 1$, by using the “zero forcing” equalizers of order $(m - 1)$

$$\mathbf{g}_{m-1, i} = \mathcal{H}_{m-1}^{-T}(\tilde{\mathbf{h}}_{m_1, m_2}) \mathbf{e}_i. \quad (20)$$

Of course, $\mathbf{g}_{m-1, i}$ does not, in general, equalize perfectly the true channel \mathbf{h}_M , even in the noiseless case, due to the influence of the tails. Since

$$\mathcal{H}_{m-1}^T(\tilde{\mathbf{h}}_{m_1, m_2}^z) \mathbf{g}_{m-1, i} = \mathbf{e}_{m_1+i} \quad (21)$$

and, under the assumptions stated in the previous subsection, $\tilde{\mathbf{h}}_{m_1, m_2}^z \approx \mathbf{h}_{m_1, m_2}^z$, we expect that

$$\mathcal{H}_{m-1}^T(\mathbf{h}_M) \mathbf{g}_{m-1, i} \approx \mathbf{e}_{m_1+i}$$

and we denote the corresponding residual as

$$\mathbf{r}_{m-1, i} \triangleq \mathbf{e}_{m_1+i} - \mathcal{H}_{m-1}^T(\mathbf{h}_M) \mathbf{g}_{m-1, i}. \quad (22)$$

The next theorem provides an upper bound for this quantity.

Theorem 3: If $\mathbf{r}_{m-1, i}$ denotes the residual of the “zero forcing” equalizer of order $(m - 1)$, $\mathbf{g}_{m-1, i}$, then for $i = 0, \dots, 2m - 1$

$$\|\mathbf{r}_{m-1, i}\|_2 \leq \frac{\sqrt{m}(\mathcal{K}_m + \epsilon_m)}{\sigma_{2m}(\mathcal{H}_{m-1}^T(\tilde{\mathbf{h}}_{m_1, m_2}))}. \quad (23)$$

Proof: The theorem can be proved as follows.

$$\begin{aligned} \|\mathbf{r}_{m-1, i}\|_2 &= \|\mathbf{e}_{m_1+i} - \mathcal{H}_{m-1}^T(\mathbf{h}_M) \mathbf{g}_{m-1, i}\|_2 \\ &= \|(\mathcal{H}_{m-1}^T(\tilde{\mathbf{h}}_{m_1, m_2}^z) - \mathcal{H}_{m-1}^T(\mathbf{h}_M)) \mathbf{g}_{m-1, i}\|_2 \\ &\leq \|\mathcal{H}_{m-1}^T(\tilde{\mathbf{h}}_{m_1, m_2}^z) - \mathcal{H}_{m-1}^T(\mathbf{h}_M)\|_2 \|\mathbf{g}_{m-1, i}\|_2 \end{aligned}$$

where we have used (21). From the definition of the convolution matrix $\mathcal{H}_{m-1}(\cdot)$, we have

$$\begin{aligned} \mathcal{H}_{m-1}^T(\tilde{\mathbf{h}}_{m_1, m_2}^z) - \mathcal{H}_{m-1}^T(\mathbf{h}_M) \\ = \mathcal{H}_{m-1}^T(\tilde{\mathbf{h}}_{m_1, m_2}^z - \mathbf{h}_{m_1, m_2}^z - \mathbf{d}_{m_1, m_2}^z) \end{aligned}$$

yielding

$$\begin{aligned} \|\mathcal{H}_{m-1}^T(\tilde{\mathbf{h}}_{m_1, m_2}^z - \mathbf{h}_{m_1, m_2}^z - \mathbf{d}_{m_1, m_2}^z)\|_2 \\ \leq \sqrt{m} \|\tilde{\mathbf{h}}_{m_1, m_2}^z - \mathbf{h}_{m_1, m_2}^z - \mathbf{d}_{m_1, m_2}^z\|_2 \\ \leq \sqrt{m} (\|\tilde{\mathbf{h}}_{m_1, m_2}^z - \mathbf{h}_{m_1, m_2}^z\|_2 + \|\mathbf{d}_{m_1, m_2}^z\|_2) \\ \leq \sqrt{m} (\mathcal{K}_m + \epsilon_m). \end{aligned}$$

Finally, from (20), we obtain

$$\|\mathbf{g}_{m-1, i}\|_2 \leq \|\mathcal{H}_{m-1}^{-T}(\tilde{\mathbf{h}}_{m_1, m_2})\|_2 = \frac{1}{\sigma_{2m}(\mathcal{H}_{m-1}^T(\tilde{\mathbf{h}}_{m_1, m_2}))}$$

to prove the theorem. \blacksquare

In the sequel, we modify slightly bound (23), and we provide a bound for $\|\mathbf{r}_{m-1, i}\|_2$ in terms of quantities related to the true channel and not to the computed estimates. Using the singular value perturbation bound [13, p. 428]

$$\begin{aligned} |\sigma_{2m}(\mathcal{H}_{m-1}^T(\tilde{\mathbf{h}}_{m_1, m_2})) - \sigma_{2m}(\mathcal{H}_{m-1}^T(\mathbf{h}_{m_1, m_2}))| \\ \leq \|\mathcal{H}_{m-1}^T(\tilde{\mathbf{h}}_{m_1, m_2}) - \mathcal{H}_{m-1}^T(\mathbf{h}_{m_1, m_2})\|_2 \leq \sqrt{m} \mathcal{K}_m \end{aligned}$$

we deduce that if $\sqrt{m} \mathcal{K}_m < \sigma_{2m}(\mathcal{H}_{m-1}^T(\mathbf{h}_{m_1, m_2}))$, then

$$\|\mathbf{r}_{m-1, i}\|_2 \leq \frac{\sqrt{m}(\mathcal{K}_m + \epsilon_m)}{\sigma_{2m}(\mathcal{H}_{m-1}^T(\mathbf{h}_{m_1, m_2})) - \sqrt{m} \mathcal{K}_m}.$$

If $\sqrt{m}(\mathcal{K}_m + \epsilon_m) \leq (\sigma_{2m}(\mathcal{H}_{m-1}^T(\mathbf{h}_{m_1, m_2}))/2)$, then

$$\begin{aligned} \|\mathbf{r}_{m-1, i}\|_2 \leq \frac{4\sqrt{2m(m+1)}\epsilon_m}{\sigma_{2m}(\mathcal{H}_{m-1}^T(\mathbf{h}_{m_1, m_2}))\sigma_{2m+1}(\mathcal{H}_m^T(\mathbf{h}_{m_1, m_2}))} \\ + \frac{4\sqrt{m}\epsilon_m}{\sigma_{2m}(\mathcal{H}_{m-1}^T(\mathbf{h}_{m_1, m_2}))}. \quad (24) \end{aligned}$$

Term $\sigma_{2m}(\mathcal{H}_{m-1}^T(\mathbf{h}_{m_1, m_2}))$ may be interpreted as a measure of diversity of \mathbf{h}_{m_1, m_2} , just like $\sigma_{2m+1}(\mathcal{H}_m^T(\mathbf{h}_{m_1, m_2}))$. These terms are not orderable, that is, one is not always larger than the other; extensive simulations have shown that they are very close each other.

Thus, the diversity of the m th-order significant part of the channel and the size of the tails are the factors that determine the performance of blind channel approximation/equalization.

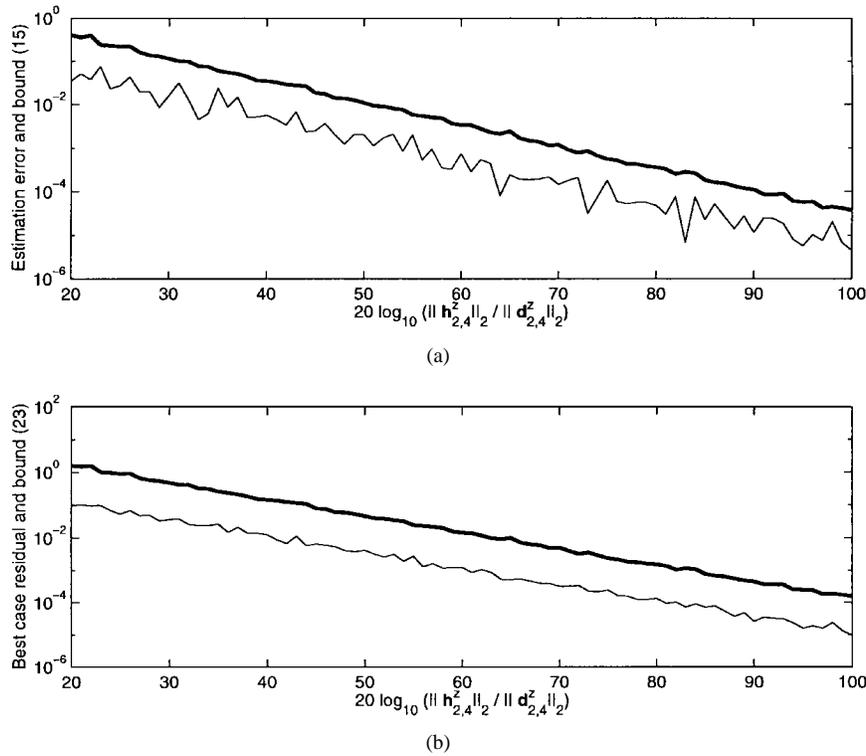


Fig. 2. (a) Estimation error $\|\tilde{\mathbf{h}}_{2,4} - \mathbf{h}_{2,4}\|_2$ (solid line) and bound (15) (thick line). (b) Best-case residual $\|\mathbf{r}_{1,i}\|_2$ (solid line) and bound (23) (thick line) for varying the size of the tails.

IV. SIMULATIONS

In the previous section, we studied the behavior of the m th-order LS and SS methods, and we derived bounds (15) and (23), which provide a measure of their performance. The bounds are not tight, in general. However, they are given by reasonably simple expressions, which identify the cases in which the algorithms perform well or may perform poorly. In this section, we validate our theoretical results using simulations.

In the first simulation, we provide results concerning blind channel identification/equalization for varying the size of the tails. The significant part of each subchannel starts at $m_1 = 2$ and has length 3, i.e., order 2; it is given by $\mathbf{h}_{2,4} = [-0.6804, 0.1777, -0.0902; 0.4281, -0.2446, -0.5043]^T$. In this case, $\delta_2 = \sigma_5(\mathcal{H}_2^T(\mathbf{h}_{2,4})) = 0.4158 = O(1)$; this implies that the significant part of the channel offers great diversity. In order to study the influence of the tails on the estimation of the significant part, relation (11) and Theorem 1 apply whenever the 2-norm of the tails does not exceed $\epsilon_2 = 0.12$, which gives

$$20 \log_{10} \frac{\|\mathbf{h}_{2,4}^z\|_2}{\|\mathbf{d}_{2,4}^z\|_2} \geq 18.3536 \text{ dB.}$$

We construct \mathbf{h}_{12} by extending the impulse response of each subchannel by random tails, which are composed of 10 nonzero terms; we add two terms before and eight terms after the significant part; using the tails, we construct $\mathbf{d}_{2,4}^z$, as in (6). In order to get the desired $(\|\mathbf{h}_{2,4}^z\|_2 / \|\mathbf{d}_{2,4}^z\|_2)$ ratio, we adjust the size of $\mathbf{h}_{2,4}^z$ and $\mathbf{d}_{2,4}^z$. Then, we apply the second-order LS/SS method on \mathbf{h}_{12} , and we compute the estimation error $\mathbf{h}_{2,4} - \tilde{\mathbf{h}}_{2,4}$. In Fig. 2(a), we plot $\|\mathbf{h}_{2,4} - \tilde{\mathbf{h}}_{2,4}\|_2$, as

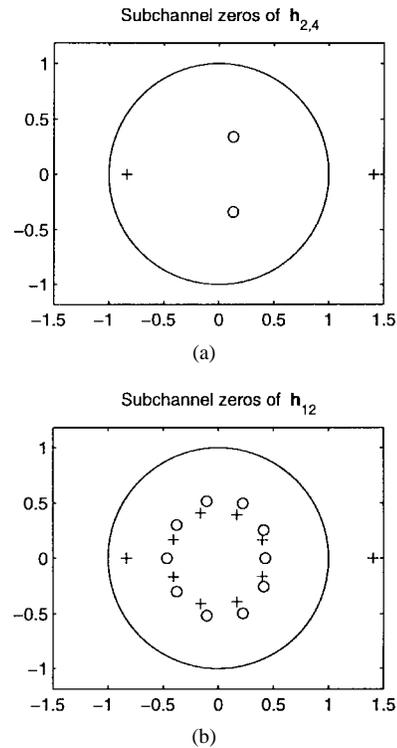


Fig. 3. (a) Subchannel zeros of $\mathbf{h}_{2,4}$. (b) Subchannel zeros of \mathbf{h}_{12} with $20 \log_{10}(\|\mathbf{h}_{2,4}^z\|_2 / \|\mathbf{d}_{2,4}^z\|_2) = 50$.

well as bound (15). We observe that the bound provides a good estimation of the "identification" error. In Fig. 2(b), we plot the smallest residual, over the different delays, of the first-order "zero forcing" equalizer, i.e., the equalizer that equalizes

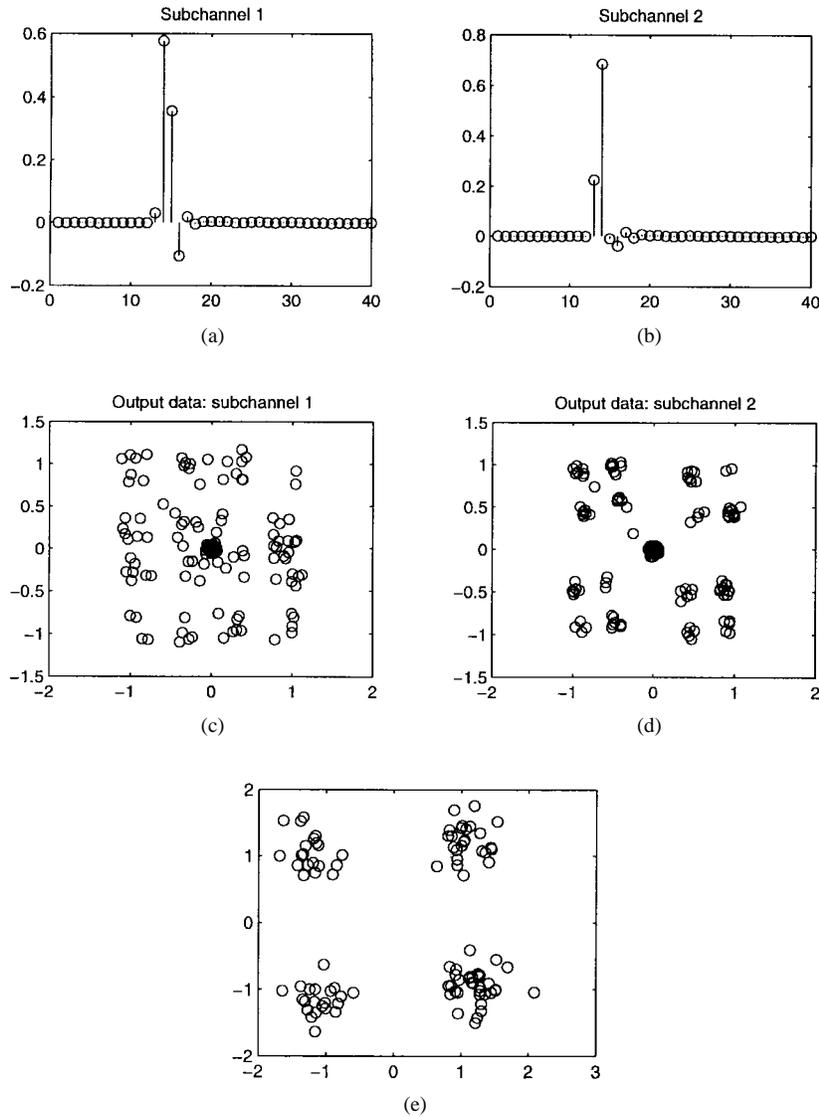


Fig. 4. (a), (b) Portion of the real part of subchannels. (c), (d) Subchannel outputs. (e) Best-case output of first-order zero forcing equalizer computed by impulse response estimated by second-order SS method.

perfectly $\tilde{\mathbf{h}}_{2,4}$, in the noiseless case, and bound (23). We observe that, at least for large $(\|\mathbf{h}_{2,4}^z\|_2/\|\mathbf{d}_{2,4}^z\|_2)$, bound (23) indicates that the channel can be equalized sufficiently well by a first-order equalizer.

In Fig. 3(a), we plot the zeros of $\mathbf{h}_{2,4}^{(1)}(z)$ and $\mathbf{h}_{2,4}^{(2)}(z)$, where

$$\mathbf{h}_{m_1, m_2}^{(j)}(z) \triangleq \sum_{i=m_1}^{m_2} h_i^{(j)} z^{-(i-m_1)}.$$

We observe that the subchannel zeros of $\mathbf{h}_{2,4}$ are not close, implying large diversity; we recall that the same fact is implied by the relatively large value of δ_2 . In Fig. 3(b), we plot most of the subchannel zeros of \mathbf{h}_{12} , which is the one used in Fig. 2, for $20 \log_{10}(\|\mathbf{h}_{2,4}^z\|_2/\|\mathbf{d}_{2,4}^z\|_2) = 50$ (two zeros of each subchannel are far away from the unit circle due to the small leading terms). We observe that the subchannel zeros of \mathbf{h}_{12} are very close, implying that \mathbf{h}_{12} offers very small diversity; however, since, as we see in Fig. 2, the corresponding residual is smaller than 10^{-2} , it is clear that

\mathbf{h}_{12} can be equalized sufficiently well by the first-order LS/SS “zero forcing” equalizer. The reason for this is that if the size of the tails is sufficiently small, as it is in our case, the performance of the first-order equalizer is determined by the diversity of $\mathbf{h}_{2,4}$, which is sufficiently large, and the size of $\mathbf{d}_{2,4}^z$, which is sufficiently small, and not by the diversity of \mathbf{h}_{12} , which is very small.

In the final simulation, we process data obtained by using the oversampled (by a factor of 2) microwave radio channel *chan4.mat*, which is found at <http://spib.rice.edu/spib/microwave.html>. The channel possesses long tails of small leading and trailing terms. In Fig. 4(a) and (b), we plot a portion of the real part of the two subchannels; each subchannel possesses 150 nonzero terms; the “small” terms are about two orders of magnitude smaller than the significant terms. In Fig. 4(c) and (d), we plot the output of each subchannel, in the noiseless case, with input 100 samples of an i.i.d. 4-QAM signal. In order to estimate the effective channel

length, we compute the “overmodeled” $\tilde{\mathbf{R}}_{30}$, and we apply the AIC and MDL criteria [19], which, unfortunately, lead to overmodeling. We found very useful the criterion

$$\text{rank}(\mathbf{R}_L) = \arg \min_i \frac{\lambda_{i+1}(\tilde{\mathbf{R}}_L)}{\lambda_i(\tilde{\mathbf{R}}_L)}, i = 1, \dots, 2L + 1. \quad (25)$$

This criterion provides a *stable* decomposition of the range space of the data autocorrelation matrix into signal and noise subspaces; an extensive study of information theoretic criteria for rank detection, as well as the development and study of (25), can be found in [20]. Using (25), we estimate the effective channel order as 2, i.e., three taps. This estimate is not only intuitively satisfying, taking into account Fig. 4(a) and (b), but it is also very useful since it will lead, as we shall see shortly, to sufficiently good equalization of *chan4.mat*. We apply the second-order LS/SS to the outputs of the two subchannels; then, we compute the corresponding first-order zero forcing equalizers. The best-case output of the zero forcing equalizers is plotted in Fig. 4(e). We see that we can open the eye by simply using first-order zero forcing equalizers.

Unfortunately, we did not manage to always reliably process data obtained by some channels available at this site (for example, *chan3.mat*). In addition, we found it difficult to process reliably more complicated input constellations, such as 16-QAM; this results from the fact that we *cannot* approximate the various channels arbitrarily well.

In the cases in which we can open the eye, 100 data samples seem to be enough; this is a clear advantage in rapidly time-varying environments. However, only channels whose significant part provides enough diversity and, at the same time, whose unmodeled tails are sufficiently small can be approximated and, subsequently, equalized sufficiently well; this is a clear shortcoming of the methods.

A more detailed simulation study, including consideration of the noisy case, can be found in [20].

V. CONCLUSIONS

In order to study the behavior of the m th-order LS and SS methods for blind channel identification, we partitioned the true channel into the m th-order significant part and the tails. We showed that the m th-order LS/SS method estimates a channel that is close to the m th-order significant part. The closeness depends on the diversity of the m th-order significant part δ_m and the size of the “unmodeled” part ϵ_m . Furthermore, we showed that if we try to model not only the “large” terms but also some “small” ones, then the blind channel estimation problem becomes generically ill-conditioned. Thus, we should avoid modeling “small” terms.

Consequently, we defined as effective channel that part of the channel containing all the “large” terms. If its order is m^* , then we called a case undermodeled (resp. overmodeled) if the assumed channel order m is smaller (resp. bigger) than m^* .

The development and the study of an efficient procedure for the determination of the effective channel order m^* is the topic of [20].

Concerning the equalization part, we considered the performance of “zero forcing” equalizers; the diversity of \mathbf{h}_{m_1, m_2} and the “identification error” determine their performance.

Finally, we performed simulations that were in general agreement with our theoretical results; more specifically, they showed that sufficiently good equalization of unknown channels, using “zero forcing” equalizers of order $(m - 1)$, is possible, if the diversity of the m th-order significant part of the true impulse response is sufficiently large and, at the same time, the size of the unmodeled part is sufficiently small. Results with a similar flavor concerning the LP method have been derived in [21].

The fact that we cannot approximate a channel arbitrarily well by increasing the complexity of our model is perhaps the most significant obstacle against general applicability of the LS and SS methods.

REFERENCES

- [1] L. Tong, G. Xu, and T. Kailath, “Blind identification based on second-order statistics: A time domain approach,” *IEEE Trans. Inform. Theory*, vol. 40, pp. 340–349, Apr. 1994.
- [2] G. Xu, H. Liu, L. Tong, and T. Kailath, “A least-squares approach to blind channel identification,” *IEEE Trans. Signal Processing*, vol. 43, pp. 2982–2993, Dec. 1995.
- [3] E. Moulines, P. Duhamel, J. F. Cardoso, and S. Mayrargue, “Subspace methods for the blind identification of multichannel FIR filters,” *IEEE Trans. Signal Processing*, vol. 43, pp. 516–525, Feb. 1995.
- [4] D. Slock, “Blind fractionally-spaced equalization, perfect reconstruction filterbanks, and multilinear prediction,” in *Proc. ICASSP*, Adelaide, Australia, Apr. 1994.
- [5] Y. Li and Z. Ding, “Global convergence of fractionally spaced Godard (CMA) adaptive equalizers,” *IEEE Trans. Signal Processing*, vol. 44, pp. 818–826, Apr. 1996.
- [6] T. J. Endres, S. D. Halford, C. R. Johnson, and G. B. Giannakis, “Blind adaptive channel equalization using fractionally-spaced receivers: A comparison study,” in *Proc. Conf. Inform. Sci. Syst.*, Princeton, NJ, Mar. 1996.
- [7] J. R. Treichler, I. Fijalkow, and C. R. Johnson, Jr., “Fractionally spaced equalizers. How long should they really be?,” *IEEE Signal Processing Mag.*, pp. 65–81, May 1996.
- [8] I. Fijalkow, A. Touzni, and J. R. Treichler, “Fractionally spaced equalization using CMA: Robustness to channel noise and lack of disparity,” *IEEE Trans. Signal Processing*, vol. 45, pp. 56–66, Jan. 1997.
- [9] H. Zeng and L. Tong, “Mean-squared error performance of constant modulus receiver for singular channels,” in *Proc. ICASSP*, Munich, Germany, Apr. 1997.
- [10] T. J. Endres, B. D. O. Anderson, C. R. Johnson, and M. Green, “On the robustness of the fractionally-spaced constant modulus criterion to channel order undermodeling: Part I,” in *Proc. IEEE Workshop Signal Process. Adv. Wireless Commun.*, Paris, France, Apr. 1997.
- [11] H. Zeng and L. Tong, “Connections between the least-squares and subspace approaches to blind channel estimation,” *IEEE Trans. Signal Processing*, vol. 44, pp. 1593–1596, June 1996.
- [12] A. P. Liavas, P. A. Regalia, and J.-P. Delmas, “Robustness of least-squares and subspace methods for blind channel identification/equalization with respect to channel undermodeling,” in *Proc. EUSIPCO*, Rhodes, Greece, Sept. 1998.
- [13] G. Golub and C. Van Loan, *Matrix Computations*. Baltimore, MD: Johns Hopkins Univ. Press, 1989.
- [14] R. A. Horn and C. R. Johnson, *Matrix Analysis*. Cambridge, U.K.: Cambridge Univ. Press, 1985.
- [15] G. Stewart and J. Sun, *Matrix Perturbation Theory*. New York: Academic, 1990.
- [16] P.-A. Wedin, “Perturbation bounds in connection with singular value decomposition,” *BIT*, vol. 12, pp. 99–111, 1972.
- [17] K. Abed-Meraim, E. Moulines, and P. Loubaton, “Prediction error method for second-order blind identification,” *IEEE Trans. Signal Processing*, vol. 45, pp. 694–705, Mar. 1997.
- [18] A. Touzni and I. Fijalkow, “Robustness of blind fractionally-spaced identification/equalization to loss of channel disparity,” in *Proc. ICASSP*, Munich, Germany, Apr. 1997.

- [19] M. Wax and T. Kailath, "Detection of signals by information theoretic criteria," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. 33, pp. 387–392, Apr. 1985.
- [20] A. P. Liavas, P. A. Regalia, and J.-P. Delmas, "Blind channel approximation: Effective channel length determination," *IEEE Trans. Signal Processing*, to be published.
- [21] ———, "Robustness of the linear prediction method for blind channel identification with respect to effective channel undermodeling/overmodeling," *IEEE Trans. Signal Processing*, submitted for publication.

Athanasios P. Liavas was born in Pyrgos, Greece, in 1966. He received the diploma and the Ph.D. degrees in computer engineering from the University of Patras, Patras, Greece, in 1989 and 1993, respectively.

He is currently a Research Fellow at the Institut National des Télécommunications, Evry, France, under the framework of the Training and Mobility of Researchers (T.M.R.) program of the European Commission. His research interests include adaptive signal processing algorithms, blind system identification, and biomedical signal processing.

Dr. Liavas is a member of the Technical Chamber of Greece.

Phillip A. Regalia (SM'96) was born in Walnut Creek, CA, in 1962. He received the B.Sc. (with highest honors), M.Sc., and Ph.D. degrees in electrical and computer engineering from the University of California, Santa Barbara, in 1985, 1987, and 1988, respectively, and the Habilitation à Diriger des Recherches degree from the University of Paris, Orsay, France, in 1994.

He is presently a Professor at the Institut National des Télécommunications, Evry, France, with research interests focused in adaptive signal processing.

Dr. Regalia serves as an Associate Editor for the IEEE TRANSACTIONS ON SIGNAL PROCESSING and as an Editor for the *International Journal of Adaptive Control and Signal Processing*.



Jean-Pierre Delmas was born in France in 1950. He received the engineering degree from the Ecole Centrale de Lyon, Lyon, France, in 1973 and the Certificat d' études supérieures from the Ecole Nationale Supérieure des Télécommunications, Paris, France, in 1982.

Since 1980, he has been with the Institut National des Télécommunications, Evry, France, where he is currently Maître de Conférences. His research interests are in statistical signal processing.