Joint Channel and Symbol Estimation by Oblique Projections

Xiang Yu and Lang Tong, *Senior Member, IEEE*

***Abstract*—The problem of simultaneous blind channel and symbol estimation of a single-input multiple-output (SIMO) communication channel is considered in this paper. It is shown that the outer product of the channel vector and the channel input sequence can be obtained by a linear estimator that has the finite sample convergence property. Furthermore, this estimator can be obtained by the use of oblique projections. An order detection algorithm that avoids the use of subjective thresholding is also proposed. Applications to multiuser detection are also considered.**

***Index Terms*—Blind channel identification, blind symbol estimation, oblique projection.**

I. Introduction

We consider the problem of joint estimation of a single-input multiple-output (SIMO) channel and its input sequence. The SIMO model is widely used in data transmission and diversity receptions, and the problem of joint channel and symbol estimation has important applications in packet transmissions where the use of training symbols may impose substantial overhead. The majority of existing techniques are iterative algorithms based on the maximization of the likelihood function [11], which require good initializations. To obtain accurate initializations, it is desirable to use closed-form algorithms (see [11] and references therein) for either channel or symbol estimation. Among these techniques, the class of deterministic algorithms have the finite sample convergence (FSC) property, which enables the estimator to obtain parameters perfectly with a finite number of noiseless observations. FSC is highly desirable in short data-length situations. Most existing closed-form blind identification techniques that have the FSC property are either channel or symbol estimators. The former include many (column) subspace techniques such as the subspace channel estimation algorithm [8], the cross relation algorithm [18], and the least squares smoothing (LSS) algorithm [12], [19]. The dual of the column space methods is the row space techniques for symbol estimations [14] or direct equalizer construction [3], [4].

The only closed-form algorithm capable of estimating channel and symbol simultaneously[1] and having the FSC property was proposed by Vandaele and Moonen [15]. Unlike those techniques using linear predictions [10] and the least squares smoothing [12], [19], where projections are orthogonal, Vandaele and Moonen used oblique projections that played a key role in obtaining the channel response and symbols jointly. This motivates us to explore systematically the idea of using oblique projection techniques for joint channel and symbol estimation.

The significance of using oblique projections is twofold. First, the use of oblique projections leads to rich geometrical interpretations of many applications, as demonstrated in [1]. Second, oblique projections can be casted within the linear least squares framework, which enables the application of existing adaptive techniques.

In this paper, we present a geometrical approach to joint channel and symbol identification. By decomposing the observation space into the past, current, and future subspaces and utilizing the isomorphic relation between the input and output subspaces, we formulate the problem of joint channel and symbol estimation as one of estimating the outer product of the channel vector and the symbol sequence using the *linear least squares estimator*. This estimation can be obtained in three different ways of using oblique projections. Unlike the Vandaele-Moonen (VM) algorithms [15] in which the oblique projection is used followed by solving the channel vector from a triangular system, our approach obtains the channel-source outer product directly from oblique projections. This implies an implementation of the joint channel source estimation based entirely on recursive least squares. We also present several extensions, including the use of total least squares projection and the application to code division multiple access (CDMA) systems.

The paper is organized as follows. Section II presents a list of key notations, data model, and a brief description of oblique projections. In Section III, the idea of intersymbol interference (ISI) removal is presented, and three projectors are defined to estimate the channel and the symbols simultaneously. Extensions to noisy environment, the problem of order detection, and the application in multiuser detection problem are also given. Simulation examples are presented in Section IV, followed by concluding remarks.

Manuscript received July 24, 2000; revised September 26, 2001. This work was supported in part by the National Science Foundation under Contract CCR-9804019 and the Multidisciplinary University Research Initiative (MURI) under the Office of Naval Research Contract N00014-00-1-0564. The associate editor coordinating the review of this paper and approving it for publication was Prof. Michail K. Tsatsanis.

X. Yu is with Aware, Inc., Bedford, MA 01730 USA (e-mail: xiangyu@aware.com).

L. Tong is with the School of Electrical and Computer Engineering, Cornell University, Ithaca, NY 14853 USA (e-mail: ltong@ece.cornell.edu).

Publisher Item Identifier S 1053-587X(01)10480-0.

[1] Here we exclude those methods that estimate channel first and then the symbols or vice versa.

II. Preliminary

A. Notations

Vectors and matrices are boldfaced letters. For most cases, we use uppercase and lowercase boldfaced letters for matrices and vectors, respectively, with $(\cdot)^T$ and $(\cdot)^H$ denoting the transpose and Hermitian operators. For a matrix with its singular value decomposition (SVD) $\mathbf{A}=\mathbf{U\Sigma V}^H$, $\mathbf{A}^\dagger$ denotes the pseudo-inverse [9] obtained by $\mathbf{A}^\dagger = \mathbf{V}\mathrm{diag}(1/\sigma_1,\ldots,1/\sigma_r,0,\ldots,0)\mathbf{U}^H$, where r is the rank of $\mathbf{A}$, and σ_i are the nonzero singular values. Calligraphic letters denote subspaces, and we use $\mathcal{C}^n$ for the n-dimensional complex Euclidean space. For a given matrix $\mathbf{A}, \mathcal{R}\{\mathbf{A}\}$, $(\mathcal{C}\{\mathbf{A}\})$ is the row (column) space of the matrix. For a given subspace $\mathcal{R}$,, $\mathbf{P}_{\mathcal{R}}$ denotes the corresponding orthogonal projection matrix on $\mathcal{R}$, and $\mathbf{x}_{\mathcal{R}}$ denotes the orthogonal projection of $\mathbf{x}$ onto $\mathcal{R}$. Similarly, $\mathbf{P}_{\mathcal{R}}^{\perp}$ denotes the corresponding orthogonal projection on $\mathcal{R}^\perp$—the orthogonal complement of $\mathcal{R}$. For two given subspaces $\mathcal{R}$ and $\mathcal{N}$, $\mathbf{E}_{\mathcal{R}\,|\,\mathcal{N}}$ denotes the corresponding oblique projection matrix with $\mathcal{R}$ as the range space and $\mathcal{N}$ as the null space, and $\mathbf{x}_{\mathcal{R}\,|\,\mathcal{N}}$ denotes the oblique projection of $\mathbf{x}$ onto $\mathcal{R}$ along $\mathcal{N}$. We use $\mathcal{R}\oplus\mathcal{N}$ to denote the direct sum of $\mathcal{R}$ and $\mathcal{N}$ and $\mathcal{R}\cap\mathcal{N}$ for the intersection of the two spaces. For a set of vectors $\mathbf{x}_1,\ldots,\mathbf{x}_n$, $sp\{\mathbf{x}_1,\ldots,\mathbf{x}_n\}$ denotes the linear subspace spanned by $\mathbf{x}_1,\ldots,\mathbf{x}_n$.

B. Oblique Projection

The idea of oblique projection and its application in signal processing are well known [1], [9]. For the sake of establishing notations, we briefly present a few necessary definitions and results.

The oblique projection of a vector obtains the component of a vector in a particular direction (the range space) while eliminating the component of the vector along a different direction (the null space). The basic idea is illustrated in Fig. 1. Consider two matrices $\mathbf{R}$ and $\mathbf{N}$ with row spaces $\mathcal{Z}_R$ and $\mathcal{Z}_N$, respectively. Assume that $\mathcal{Z}_R$ and $\mathcal{Z}_N$ are disjoint, and let $\mathcal{Z}=\mathcal{Z}_R\oplus\mathcal{Z}_N\subset\mathcal{C}^n$. For any vector $\mathbf{x}\in\mathcal{C}^n$, the oblique projection of $\mathbf{x}$ onto $\mathcal{Z}_R$ along $\mathcal{Z}_N$, which is denoted by $\mathbf{x}_{\mathcal{Z}_R\,|\,\mathcal{Z}_N}$, is obtained by the following two steps: i) the orthogonal projection $\mathbf{x}_{\mathcal{Z}}$ of $\mathbf{x}$ onto $\mathcal{Z}$; ii) finding the component of $\mathbf{x}_{\mathcal{Z}}$ in the direction of $\mathcal{Z}_R$ while eliminating that in the direction of $\mathcal{Z}_N$. Algebraically, the oblique projection can be computed as [1]

$$\mathbf{x}_{\mathcal{Z}_R\,|\,\mathcal{Z}_N}=\mathbf{x}\mathbf{E}_{\mathcal{Z}_R\,|\,\mathcal{Z}_N} \tag{1}$$

where $\mathbf{E}_{\mathcal{Z}_R\,|\,\mathcal{Z}_N}$ is the projection operator (projector) given by

$$\mathbf{E}_{\mathcal{Z}_R\,|\,\mathcal{Z}_N}=(\mathbf{R}^H,\mathbf{N}^H)\begin{pmatrix}\mathbf{RR}^H & \mathbf{RN}^H\\ \mathbf{NR}^H & \mathbf{NN}^H\end{pmatrix}^{\dagger}\begin{pmatrix}\mathbf{R}\\ \mathbf{0}\end{pmatrix} \tag{2}$$

$$=\mathbf{P}_{\mathcal{Z}_N}^{\perp}\mathbf{R}^H\left(\mathbf{R}^H\mathbf{P}_{\mathcal{Z}_N}^{\perp}\mathbf{R}\right)^{\dagger}\mathbf{R}. \tag{3}$$

The two subspaces $\mathcal{Z}_R$ and $\mathcal{Z}_N$ are called the *range* space and the *null* space of the projector, and they completely specify the oblique projection.

From the definition, we can see that the Euclidean space $\mathcal{C}^n$ is decomposed into a direct sum of three different directions, i.e., $\mathcal{C}^n=\mathcal{Z}^\perp\oplus\mathcal{Z}_\mathcal{R}\oplus\mathcal{Z}_\mathcal{N}$; for $\mathbf{x}\in\mathcal{Z}_R$, $\mathbf{x}\mathbf{E}_{\mathcal{Z}_R\,|\,\mathcal{Z}_N}=\mathbf{x}$, and $\mathbf{x}\mathbf{E}_{\mathcal{Z}_R\,|\,\mathcal{Z}_N}=\mathbf{0}$ for $\mathbf{x}\in\mathcal{Z}_N\oplus\mathcal{Z}^\perp$, i.e., the projector on $\mathcal{Z}_\mathcal{R}$ along $\mathcal{Z}_\mathcal{N}$ preserves any vector in $\mathcal{Z}_R$ unchanged and nullifies vectors in $\mathcal{Z}_N\oplus\mathcal{Z}^\perp$.

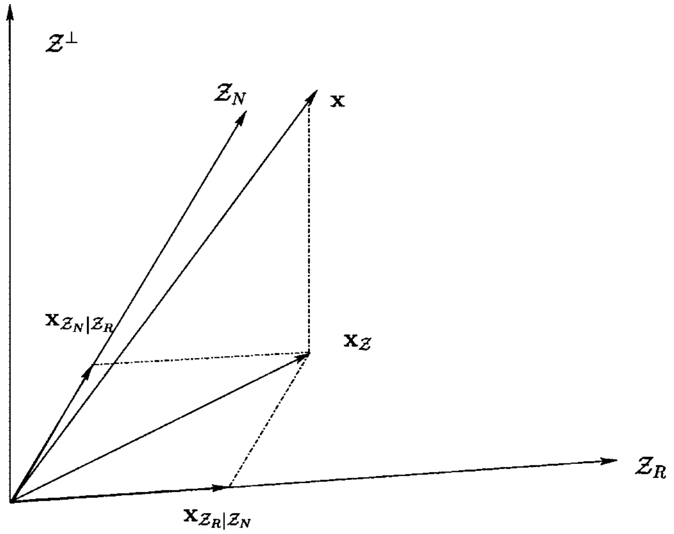
$\mathcal{Z}^\perp$
$\mathcal{Z}_N$
$\mathbf{x}$
$\mathbf{x}_{\mathcal{Z}_N|\mathcal{Z}_R}$
$\mathbf{x}_{\mathcal{Z}}$
$\mathcal{Z}_R$
$\mathbf{x}_{\mathcal{Z}_R|\mathcal{Z}_N}$

Fig. 1. Oblique projection illustration: The oblique projection of $\mathbf{x}$ onto $\mathcal{Z}_\mathcal{R}$ along $\mathcal{Z}_\mathcal{N}$ is $\mathbf{x}_{\mathcal{Z}_\mathcal{R}\,|\,\mathcal{Z}_\mathcal{N}}$.

The oblique projection can also be viewed as part of the linear least squares problem. Specifically, given two matrices $\mathbf{R}$ and $\mathbf{N}$, the oblique projection $\mathbf{x}_{\mathcal{Z}_\mathbf{R}\,|\,\mathcal{Z}_\mathbf{N}}$ of a vector $\mathbf{x}$ can be obtained from

$$\min_{\alpha,\beta}||\mathbf{x}-\alpha\mathbf{R}-\beta\mathbf{N}||^2\Rightarrow\mathbf{x}_{\mathcal{Z}_R\,|\,\mathcal{Z}_N}=\alpha_*\mathbf{R} \tag{4}$$

where α_* is the least squares solution. Thus, obtaining the oblique projection of vector $\mathbf{x}$ is equivalent to solving a linear least squares problem, which can be implemented by standard fast recursive techniques.

C. System Model

Consider the following single input P outputs FIR system:

$$\mathbf{x}(t)=\sum_{l=0}^{L}\mathbf{h}_l s(t-l),\quad \mathbf{y}(t)=\mathbf{x}(t)+\mathbf{n}(t) \tag{5}$$

where

$s(t)$	scaler complex input sequence;
$\mathbf{x}(t)\in\mathcal{C}^P$	noiseless vector channel output;
$\mathbf{h}_l$	vector channel impulse response.

With additive complex noise $\mathbf{n}(t)$, the received signal is $\mathbf{y}(t)$. Our goal is to estimate both the channel $\mathbf{h}\stackrel{\Delta}{=}[\mathbf{h}_L^T,\ldots,\mathbf{h}_0^T]^T$ and the symbols $s(t)$ from $\mathbf{y}(t)$. Define row vectors of input $s(t)$ and noiseless observation $\mathbf{x}(t)$ as

$$\mathbf{s}_t\stackrel{\Delta}{=}[s(t),s(t+1),\ldots],$$
$$\mathbf{x}_t\stackrel{\Delta}{=}[\mathbf{x}(t),\mathbf{x}(t+1),\ldots].$$

We now have

$$\mathbf{x}_t=\sum_{l=0}^{L}\mathbf{h}_l\mathbf{s}_{t-l},\quad \mathbf{y}_t=\mathbf{x}_t+\mathbf{n}_t. \tag{6}$$

We will use the block representation of the above SIMO model by stacking m blocks of $\mathbf{x}_t$ and denoting $\mathbf{X}_m(t) \triangleq [\mathbf{x}_t^T, \ldots, \mathbf{x}_{t-m+1}^T]^T$. We then have

$$\begin{aligned} \mathbf{X}_m(t) &= \mathcal{F}_m(\mathbf{h})\mathbf{S}_{L+m}(t) \\ \mathbf{Y}_m(t) &= \mathbf{X}_m(t) + \mathbf{N}_m(t) \end{aligned} \tag{7}$$

where $\mathbf{Y}_m(t), \mathbf{N}_m(t)$, and $\mathbf{S}_{L+m}(t)$ are similarly defined as $\mathbf{X}_m(t)$, and $\mathcal{F}_m(\mathbf{h})$ is the filtering matrix

$$\mathcal{F}_m(\mathbf{h}) \triangleq \begin{pmatrix} \mathbf{h}_0 & \cdots & \mathbf{h}_L & & \\ & \ddots & \cdots & \ddots & \\ & & \mathbf{h}_0 & \cdots & \mathbf{h}_L \end{pmatrix}_{mP\times(m+L)}. \tag{8}$$

We make the following two assumptions:

A1) There exists an m_0 such that the filtering matrix $\mathcal{F}_{m_0}$ has full column rank.

A2) The input sequence $s(t)$ has the linear complexity [2] greater than $L_* = 2m_0 + 2L$.

Assumptions A1 and A2 play a critical role in all subspace methods; they imply an isomorphic relation between the (noiseless) output and input subspaces. In addition, note that this relation is valid for all $m \geq m_0$. Under A1 and A2, for $m \geq m_0$, we have

$$\mathcal{R}\{\mathbf{X}_m(t)\} = \mathcal{R}\{\mathbf{S}_{m+L}(t)\}. \tag{9}$$

This property tells us that any input subspace spanned by no less than m_0+L consecutive row vectors can be directly constructed from the corresponding output subspace. By this property, we can use such subspace techniques such as projection, intersection, and union to obtain the estimates of the channel and input symbols.

III. Joint Channel and Symbol Estimation by Oblique Projection

Given $\{\mathbf{y}(t)\}$, our goal is to find a linear estimator $\mathcal{L}$ of the quantity $\mathbf{h}\mathbf{s}_t$, where $\mathbf{h}$ includes all parameters of the channel impulse response. The estimator is a smoother in the sense that for some $\bar{N}$, it takes $\{\mathbf{y}(t-\bar{N}), \ldots, \mathbf{y}(t+\bar{N})\}$ to generate an estimate of $\mathbf{h}\mathbf{s}_t$. To ensure the finite sample convergence property, we require $\mathcal{L}$ to have the following property:

$$\mathcal{L}[\mathbf{x}(t-\bar{N}), \ldots, \mathbf{x}(t+\bar{N})] = \mathbf{h}\mathbf{s}_t. \tag{10}$$

The existence and the construction of such an estimator, which is shown next, have not been established previously.

A. Subspace Decomposition

The following development is based on the noiseless observation, and we defer the discussion of the noisy case to Section III-D.

To obtain $\mathbf{h}\mathbf{s}_t$, we consider the observation matrix that contains the "current" input $\mathbf{s}_t$. As shown in Fig. 2, define the "current" data matrix

$$\mathbf{C}_t \triangleq \mathbf{X}_{L+1}(t+L) = \mathcal{F}_{L+1}(\mathbf{h})\mathbf{S}_{2L+1}(t+L).$$

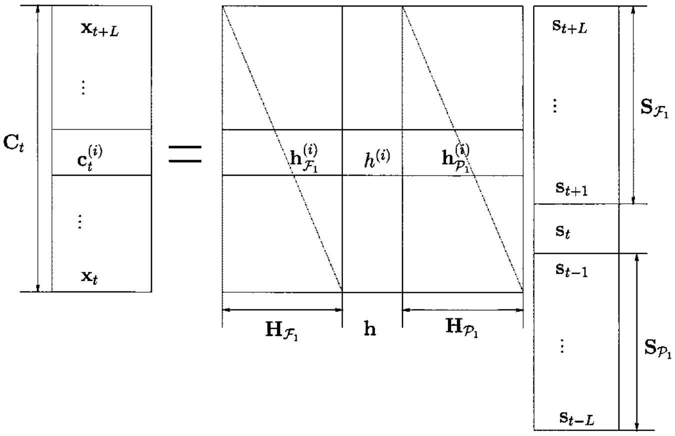

Fig. 2. Decomposition of the current data matrix $\mathbf{C}_t$.

By definition, an observation vector $\mathbf{x}_{t+i}$ can be written as a linear combination of input vectors involving $\mathbf{s}_t$ if and only if it is a row vector of $\mathbf{C}_t$. The ith row of $\mathbf{C}_t$ satisfies

$$\begin{aligned} \mathbf{c}_t^{(i)} &= h^{(i)}\mathbf{s}_t + \underbrace{\mathbf{h}_{\mathcal{F}_1}^{(i)}\mathbf{S}_{\mathcal{F}_1}(t) + \mathbf{h}_{\mathcal{P}_1}^{(i)}\mathbf{S}_{\mathcal{P}_1}(t)}_{\substack{\text{future and} \\ \text{past interference}}} \\ &\triangleq h^{(i)}\mathbf{s}_t + \mathbf{v}_t^{(i)} \end{aligned} \tag{11}$$

where $h^{(i)}$ is the ith component in the channel vector $\mathbf{h}$, and $\mathbf{h}_{\mathcal{F}_1}^{(i)}$, $\mathbf{h}_{\mathcal{P}_1}^{(i)}$ are row subvectors from $\mathcal{F}_{L+1}(\mathbf{h})$, as illustrated in Fig. 2. Past and future input matrices $\mathbf{S}_{\mathcal{P}_1}(t)$ and $\mathbf{S}_{\mathcal{F}_1}(t)$ are also shown in Fig. 2. In the observation $\mathbf{c}_t^{(i)}$

$$\mathbf{v}_t^{(i)} \triangleq \mathbf{h}_{\mathcal{F}_1}^{(i)}\mathbf{S}_{\mathcal{F}_1}(t) + \mathbf{h}_{\mathcal{P}_1}^{(i)}\mathbf{S}_{\mathcal{P}_1}(t)$$

is the interference to $\mathbf{s}_t$ from the past and future symbols. From here, our goal is to remove $\mathbf{v}_t^{(i)}$ from the noiseless observation.

Although the row spaces of $\mathbf{S}_{\mathcal{F}_1}(t)$ and $\mathbf{S}_{\mathcal{P}_1}(t)$ are not directly available from the observations, they are contained in the spaces spanned by future and past observations. Specifically, as illustrated in Fig. 3, for a large enough m

$$\begin{aligned} \mathcal{R}\{\mathbf{S}_{\mathcal{F}_1}(t)\} &\subset \mathcal{R}\{\mathbf{S}_{L+m}(t+L+m)\} \\ &= \mathcal{R}\{\mathbf{X}_m(t+L+m)\} \triangleq \mathcal{F}_1 \end{aligned} \tag{12}$$

$$\begin{aligned} \mathcal{R}\{\mathbf{S}_{\mathcal{P}_1}(t)\} &\subset \mathcal{R}\{\mathbf{S}_{L+m}(t-1)\} \\ &= \mathcal{R}\{\mathbf{X}_m(t-1)\} \triangleq \mathcal{P}_1 \end{aligned} \tag{13}$$

and $\mathcal{F}_1$ and $\mathcal{P}_1$ can be directly obtained from the observations according to (9). Thus, the interference is contained in $\mathcal{P}_1 \oplus \mathcal{F}_1$, i.e.,

$$\mathbf{v}_t^{(i)} \in \mathcal{P}_1 \oplus \mathcal{F}_1, \quad \forall i.$$

If $\mathbf{x}_{t-m}, \ldots, \mathbf{x}_{t+L+m}$ form the input of the linear estimator, the input space of the linear estimator has the following decomposition:

$$\mathcal{R}\{\mathbf{X}_{L+2m+1}(t+L+m)\} = \mathcal{P}_1 \oplus \mathcal{F}_1 \oplus sp\{\mathbf{s}_t\}. \tag{14}$$

The key idea of our approach is revealed in the above equation. To obtain the estimate of $\mathbf{h}\mathbf{s}_t$, we need to project the observation onto the signal subspace sp$\{\mathbf{s}_t\}$ along the interference subspace

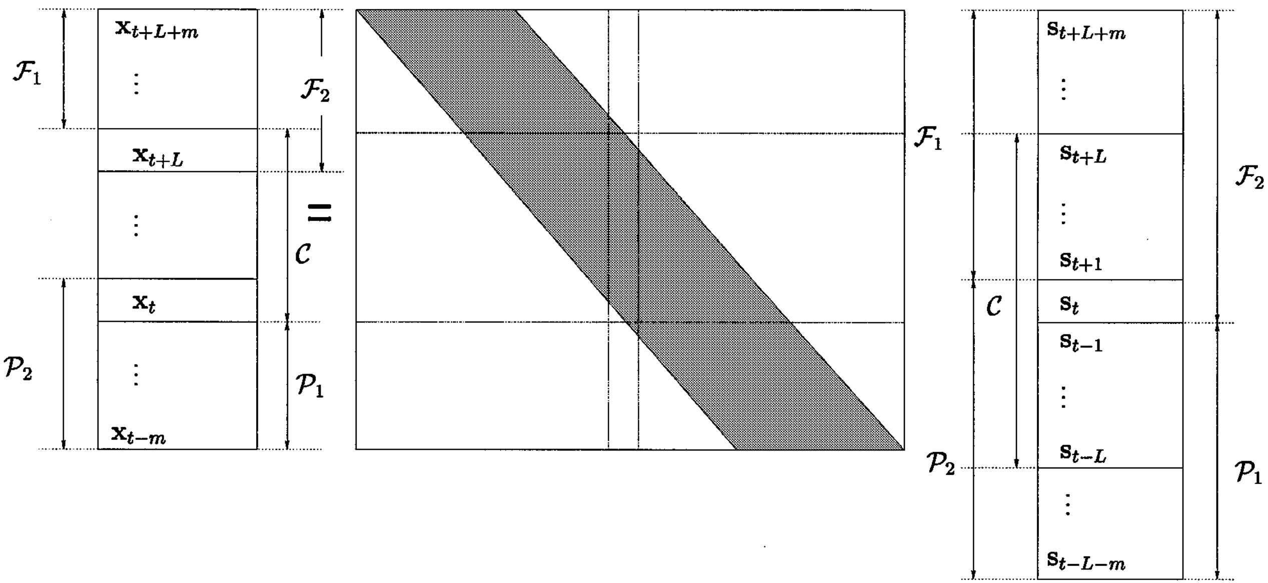

Fig. 3. Signal and Interference subspaces.

$\mathcal{P}_1 \oplus \mathcal{F}_1$. Since $\text{sp}\{\mathbf{s}_t\}$ is not available directly, we again need to construct this subspace from the observations. This can be done by extending $\mathcal{F}_1$ and $\mathcal{P}_1$ to include $\mathbf{s}_t$. Specifically, define

$$\mathcal{P}_2 \triangleq \mathcal{P}_1 \oplus \text{sp}\{\mathbf{s}_t\} = \mathcal{R}\{\mathbf{X}_{m+1}(t)\} \tag{15}$$

$$\mathcal{F}_2 \triangleq \mathcal{F}_1 \oplus \text{sp}\{\mathbf{s}_t\} = \mathcal{R}\{\mathbf{X}_{m+1}(t+L+m)\}. \tag{16}$$

This leads to $\text{sp}\{\mathbf{s}_t\} = \mathcal{P}_2 \bigcap \mathcal{F}_2$, and

$$\mathcal{R}\{\mathbf{X}_{L+2m+1}(t+L+m)\} = \mathcal{P}_1 \oplus \mathcal{F}_1 \oplus \left(\mathcal{P}_2 \bigcap \mathcal{F}_2\right). \tag{17}$$

The above equation serves as the basis of a number of techniques to be presented next. It should be pointed out that van der Veen *et al.* first proposed the use of intersection of $\mathcal{F}_2$ and $\mathcal{P}_2$ to obtain a direct input sequence estimation scheme [4], [14]. The computation of intersection, especially in the presence of noise, can be quite involved [5], [14]. In our approach, the direct intersection is avoided by the use of oblique projections.

B. Oblique Projection Algorithms

Recall $\mathbf{c}_t^{(i)} = h^{(i)}\mathbf{s}_t + \mathbf{v}_t^{(i)}$ from (11). From the above subspace definitions, we have

$$h^{(i)}\mathbf{s}_t \in \mathcal{P}_2 \cap \mathcal{F}_2, \quad \mathbf{v}_t^{(i)} \in \mathcal{P}_1 \oplus \mathcal{F}_1.$$

Therefore, $h^{(i)}\mathbf{s}_t$ can be obtained from the oblique projection of $\mathbf{c}_t^{(i)}$ onto $\mathcal{P}_2 \cap \mathcal{F}_2$ along $\mathcal{P}_1 \oplus \mathcal{F}_1$. To compute this projection, we call the following lemma that leads directly to three estimators.

Lemma: Let $\mathcal{F}_1, \mathcal{F}_2$ and $\mathcal{P}_1, \mathcal{P}_2$ be future and past subspaces defined in (12), (13), (15), (16), and $\mathbf{c}_t^{(i)}$, which is the ith row of the observation matrix $\mathbf{X}_{L+1}(t+L)$. Then

$$h^{(i)}\mathbf{s}_t = \mathbf{c}_t^{(i)}\mathbf{E} \Rightarrow \mathbf{h}\mathbf{s}_t = \mathbf{X}_{L+1}(t+L)\mathbf{E} \tag{18}$$

where $\mathbf{E}$ can be any one of the following projectors:

$$\mathbf{E}_1 \triangleq \mathbf{E}_{\mathcal{P}_2 \,|\, \mathcal{F}_1}\mathbf{E}_{\mathcal{F}_2 \,|\, \mathcal{P}_1} \tag{19}$$

$$\mathbf{E}_2 \triangleq \mathbf{E}_{\mathcal{P}_2 \,|\, \mathcal{F}_1} - \mathbf{E}_{\mathcal{P}_1 \,|\, \mathcal{F}_2} \tag{20}$$

$$\mathbf{E}_3 \triangleq \mathbf{I} - \left(\mathbf{E}_{\mathcal{F}_1 \,|\, \mathcal{P}_2} + \mathbf{E}_{\mathcal{P}_1 \,|\, \mathcal{F}_2}\right). \tag{21}$$

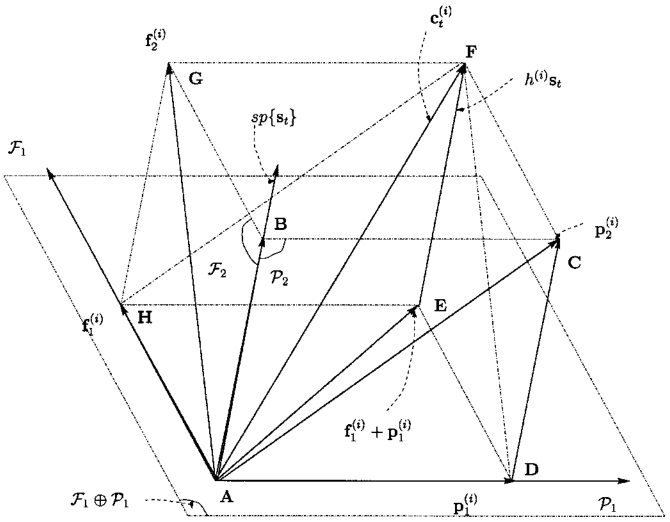

Fig. 4. Channel and symbol estimation by projection.

Although the proof of the above lemma can be found in the Appendix, we present here a geometrical interpretation of the three projectors described for (18). From (11), define the future and past interferences as

$$\mathbf{f}_1^{(i)} \triangleq \mathbf{h}_{\mathcal{F}_1}^{(i)}\mathbf{S}_{\mathcal{F}_1}(t), \quad \mathbf{p}_1^{(i)} \triangleq \mathbf{h}_{\mathcal{P}_1}^{(i)}\mathbf{S}_{\mathcal{P}_1}(t).$$

Here, we ignore the time index for ease of illustrations. Equation (11) can be rewritten as

$$\mathbf{c}_t^{(i)} = \underbrace{h^{(i)}\mathbf{s}_t + \mathbf{p}_1^{(i)}}_{\mathbf{p}_2^{(i)}} + \mathbf{f}_1^{(i)} = \mathbf{p}_2^{(i)} + \mathbf{f}_1^{(i)} \tag{22}$$

$$= \underbrace{h^{(i)}\mathbf{s}_t + \mathbf{f}_1^{(i)}}_{\mathbf{f}_2^{(i)}} + \mathbf{p}_1^{(i)} = \mathbf{f}_2^{(i)} + \mathbf{p}_1^{(i)}. \tag{23}$$

These quantities are illustrated in Fig. 4, where subspaces are represented as directed lines or planes, such as $\mathcal{P}_2, \mathcal{P}_1, \text{sp}\{\mathbf{s}_t\}$, whereas matrices or data are represented as vectors with specific lengths, such as $\mathbf{p}_2^{(i)}, \mathbf{p}_1^{(i)}, \mathbf{c}_t^{(i)}$.

From plane $\overline{ABCD}$, we have

$$h^{(i)}\mathbf{s}_t = \mathbf{p}_2^{(i)} - \mathbf{p}_1^{(i)} = \mathbf{p}_2^{(i)}\mathbf{E}_{\mathcal{F}_2 \mid \mathcal{P}_1}.$$

From plane $\overline{ACFH}$, we have

$$\mathbf{p}_2^{(i)} = \mathbf{c}_t^{(i)} - \mathbf{f}_1^{(i)} = \mathbf{c}_t^{(i)}\mathbf{E}_{\mathcal{P}_2 \mid \mathcal{F}_1}.$$

Hence, for every i

$$h^{(i)}\mathbf{s}_t = \mathbf{c}_t^{(i)}\mathbf{E}_{\mathcal{P}_2 \mid \mathcal{F}_1}\mathbf{E}_{\mathcal{F}_2 \mid \mathcal{P}_1}$$

which is the same as in (18).

The other two projectors can also be justified from the geometrical illustrations in Fig. 4. For $\mathbf{E}_2$, if we have $\mathbf{p}_2^{(i)}$ and $\mathbf{p}_1^{(i)}$ from (22)–(23), then $h^{(i)}\mathbf{s}_t$ is given by $\mathbf{p}_2^{(i)} - \mathbf{p}_1^{(i)}$. This is illustrated in the triangle $\widehat{\mathbf{ACD}} : \mathbf{DC} = \mathbf{AC} - \mathbf{AD}$. Note that this procedure can also be done with $\mathbf{f}_2^{(i)}$ and $\mathbf{f}_1^{(i)}$. For $\mathbf{E}_3$, if we have $\mathbf{f}_1^{(i)}, \mathbf{p}_1^{(i)}$ from (22) and (23), then $h^{(i)}\mathbf{s}_t$ is given by $\mathbf{c}_t^{(i)} - (\mathbf{f}_1^{(i)} + \mathbf{p}_1^{(i)})$. The relation is ready to show in the triangle $\widehat{\mathbf{AEF}} : \mathbf{EF} = \mathbf{AF} - \mathbf{AE}$.

Projectors $\mathbf{E}_2$ and $\mathbf{E}_3$ are, in fact, identical with noiseless observations. Because the row space of the data to be projected $\mathcal{R}\{\mathbf{X}_{L+1}(t+L)\} \subset \mathcal{F}_2 \oplus \mathcal{P}_1 = \mathcal{P}_2 \oplus \mathcal{F}_1$, the following equations hold:

$$\begin{aligned}\mathbf{X}_{L+1}(t+L)\mathbf{E}_{\mathcal{F}_2 \mid \mathcal{P}_1} &= \mathbf{X}_{L+1}(t+L)\left(\mathbf{I} - \mathbf{E}_{\mathcal{P}_1 \mid \mathcal{F}_2}\right)\\ \mathbf{X}_{L+1}(t+L)\mathbf{E}_{\mathcal{P}_2 \mid \mathcal{F}_1} &= \mathbf{X}_{L+1}(t+L)\left(\mathbf{I} - \mathbf{E}_{\mathcal{F}_1 \mid \mathcal{P}_2}\right).\end{aligned}$$

The use of oblique projection to estimate the input sequence in packet transmissions has a drawback at the two ends of the packet. In particular,the output of the proposed linear estimator is the outer product of $\mathbf{h}$ and $\mathbf{s}_t$, which is one row of the whole input matrix $\mathbf{S}_{2L+2m+1}(t+L+m)$. As a result, the first $m+L$ and the last $m+L$ symbols are not in the estimates. This can be explained in Fig. 3. It shows that an overall input subspace dimension as large as $2m+2L+1$ is required in order to construct the future and the past subspaces. This effectively reduces the number of estimated symbols to $N-2m-2L$. However, these symbols can be obtained once the channel is estimated using either direct inverse or, perhaps more efficiently, using a decision feedback approach, and the overall performance is not affected by the detection of these symbols.

C. Connections With Other Algorithms

The proposed identification scheme uses the row space of the output data. Thus, it is not surprising to see its relation in several aspects with other row-space based algorithms.

It is in [15] that the oblique projection is first applied into SIMO system for blind estimation of channel and symbols, but in [15], the decomposition of the subspaces and the estimation scheme are different from the proposed algorithms. The current data defined in [15] is $\mathbf{X}_L(t+L-1)$, which is contained in the future and past subspaces. After one oblique projection separating $\mathbf{X}_L(t+L-1)$ along future and past directions, the shifting-invariant Toeplitz structure of the input matrix enables the further cancellation of the remaining ISI in the two directions by doing direct subtraction.

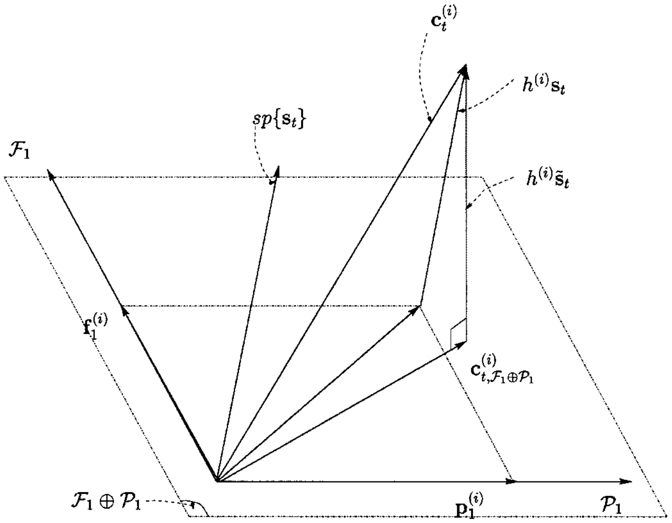

Fig. 5. Connection with LSS.

The major computational complexity of the proposed algorithms is the QR decomposition. In [15], one QR is used to decompose the current data into two parts. In $\mathbf{E}_1$, $\mathbf{E}_2$, and $\mathbf{E}_3$, the current data is decomposed into two parts but via two different ways. Therefore, two QRs are used. To be specific, for a $m \times n$ matrix, one QR decomposition needs $3n^2(m-n/3) \sim \mathcal{O}(n^2m)$ flops (Givens QR, [5]), and one SVD uses $14mn^2 - 2n^3 \sim \mathcal{O}(n^2m)$ [5] flops. Hence, for the algorithms in this paper and in [15], the computation complexity is $\mathcal{O}((m+L)^2N)$ and for SS-symbol in [14], the complexity is $\mathcal{O}((m+L)^4N)$.

From Fig. 5, the difference between the proposed oblique projection-based algorithm and the least square smoothing (LSS) [19] algorithm is illustrated. In LSS, the estimated channel is obtained from the orthogonal projection error of the current data $\mathbf{X}_{L+1}(t+L)$ onto the future and past spaces $\mathcal{F}_1$ and $\mathcal{P}_1$. The projection error is denoted as $\mathbf{h}\tilde{\mathbf{s}}_t$, where $\tilde{\mathbf{s}}_t$ is the error of projecting $\mathbf{s}_t$ onto $\mathcal{F}_1 \oplus \mathcal{P}_1$. Because the row space of the projection result is spanned by the projection error, we can only take $\tilde{\mathbf{s}}_t$ as the estimate of the symbol sequence. Hence, LSS has the finite sample convergence (FSC) property for channel estimation but not for symbol estimation. In the illustration, we can see the difference between the two vectors representing $\mathbf{h}\tilde{\mathbf{s}}_t$ and $\mathbf{h}\mathbf{s}_t$. The first one is orthogonal to the plane spanned by $\mathcal{F}_1$ and $\mathcal{P}_1$ (the plane is called interference plane), whereas the second is skew to it. When the input $\mathbf{s}_t$ is orthogonal to $\mathcal{F}_1 \oplus \mathcal{P}_1$, which is asymptotically true for an uncorrelated input sequence, the projection error $\tilde{\mathbf{s}}_t$ will converge to $\mathbf{s}_t$.

D. Effects of Noisy Data: Total Least Squares Oblique Projection

We have so far assumed the noiseless observations. While our simulation indicates that the algorithm performs reasonably well in the presence of noise, it is desirable to modify the formulation, taking explicitly account of the presence of the noise.

The idea is to "remove" noise from the observation samples that are used in constructing oblique projections. As shown in (4), the oblique projection can be obtained as the solution of a least squares problem. Let matrices $\mathbf{R}$ and $\mathbf{N}$ be the noisy range and null space matrices of the projector, and let $\mathbf{A} \triangleq \begin{pmatrix}\mathbf{R}\\ \mathbf{N}\end{pmatrix}$. The (least squares) oblique projection of a noisy vector $\mathbf{b}$ onto $\mathbf{R}$ along $\mathbf{N}$ can be viewed as removing the noise from $\mathbf{b}$ by

$$\min \|\Delta\mathbf{b}\|_2, \quad \text{subject to } \mathbf{b} - \Delta\mathbf{b} = [\alpha, \beta]\mathbf{A}. \tag{24}$$

The oblique projection is then given by

$$\mathbf{b}_{\mathbf{R}\,|\,\mathbf{N}} = \alpha_*\mathbf{R}, \quad \mathbf{b}_{\mathbf{N}\,|\,\mathbf{R}} = \beta_*\mathbf{N} \tag{25}$$

where α_* and β_* are the optimal vectors from (24).

Because $\mathbf{R}$ and $\mathbf{N}$ are themselves noisy, we can modify (24) to obtain the total least squares oblique projection [7] as follows:

$$\begin{aligned}&\min \|[\Delta\mathbf{b}^T, \Delta\mathbf{A}^T]^T\|_F\\ &\quad \text{subject to } (\mathbf{b} - \Delta\mathbf{b}) = [\alpha, \beta](\mathbf{A} - \Delta\mathbf{A}) && (26)\\ &[\alpha^*, \beta^*] = \mathbf{b}\mathbf{A}^H(\mathbf{A}\mathbf{A}^H - \sigma^2\mathbf{I})^\dagger && (27)\\ &\mathbf{b}_{\mathbf{R}\,|\,\mathbf{N}} = \alpha^*(\mathbf{R} - \Delta\mathbf{R}), \quad \mathbf{b}_{\mathbf{N}\,|\,\mathbf{R}} = \beta^*(\mathbf{N} - \Delta\mathbf{N}) && (28)\end{aligned}$$

where σ is the smallest singular value of $[\mathbf{A}^T, \mathbf{b}^T]^T$, and $\Delta\mathbf{R}$ and $\Delta\mathbf{N}$ are perturbations corresponding to $\mathbf{R}$ and $\mathbf{N}$, respectively.

E. Order Detection

Order detection is always difficult for channel equalization problem, especially for those multipath channels with small head and tail taps. Specifically, the intersection between $\mathcal{F}_2$ and $\mathcal{P}_2$ is ill defined when the channel has small head and/or tail taps. From Fig. 3, it can be seen that to compute $\mathcal{F}_2$ from the output matrix, $\mathbf{s}_t$ is scaled by the tail of the channel. Similarly, in $\mathcal{P}_2$, $\mathbf{s}_t$ is scaled by the head of the channel, and $\mathbf{s}_t$ is the only common part between $\mathcal{F}_2$ and $\mathcal{P}_2$. Therefore, for noisy data, the intersection of $\mathcal{F}_2$ and $\mathcal{P}_2$ will not give out desired results. In general, for this kind of channel, good performance can be obtained by underestimating the order in low SNR region. In this section, we present a technique based on the property of oblique projection to estimate the channel order. The order can be determined with the only assumption of an upper bound on the true order.

Let us consider the case in a SIMO system for which the channel order is over or underdetermined. Let the over/underdetermined order be L'. The *future* and *past* data are represented in Fig. 6. Given L', consider the sum of two projections

$$\begin{aligned}\mathbf{G_I} &\triangleq \mathbf{E}_{\mathcal{F}_2\,|\,\mathcal{P}_1}\mathbf{E}_{\mathcal{P}_2\,|\,\mathcal{F}_1} + \mathbf{E}_{\mathcal{P}_2\,|\,\mathcal{F}_1}\mathbf{E}_{\mathcal{F}_2\,|\,\mathcal{P}_1} && (29)\\ \mathbf{E}_t' &= \mathbf{X}_{L'+1}(t + L') \cdot \mathbf{G}_I. && (30)\end{aligned}$$

If $L = L'$, i.e., the order is correctly given, the two projectors in the summation are identical, and the resulting $\mathbf{E}_t'$ is rank one. When $L < L'$, the order is overdetermined. Both of the two terms in (29) are of rank one but have different row and column spaces. Hence, in general, the projection will have rank two. When the order is underdetermined, i.e., $L > L'$, since the preset order is smaller than the true one, the oblique projection is not well defined due to the overlap of the range and the null spaces, and the usage of pseudo-inverse in (3) makes the accurate analysis very complex. However, extensive simulations show that the projection has rank greater than one for randomly generated input sequences. Similar results also hold for the following projector:

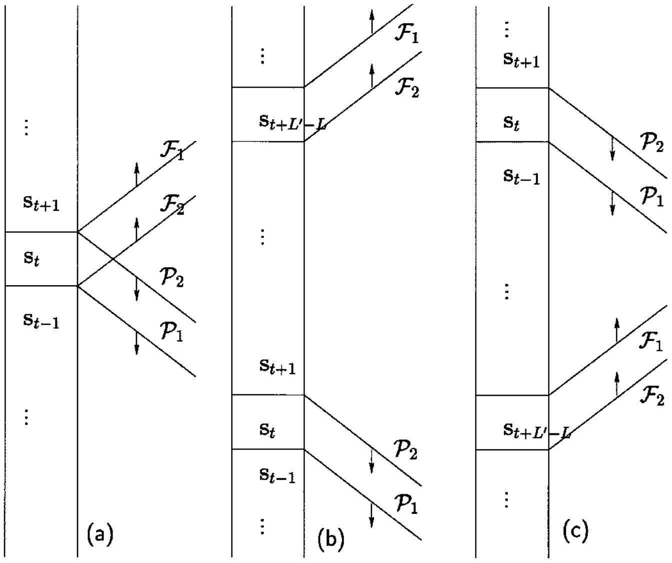

Fig. 6. (a) Correct order. (b) Overdetermined order. (c) Underdetermined order.

$$\mathbf{G_{II}} \triangleq \mathbf{E}_{\mathcal{F}_1\,|\,\mathcal{P}_2} + \mathbf{E}_{\mathcal{P}_1\,|\,\mathcal{P}_2} - \mathbf{E}_{\mathcal{F}_2\,|\,\mathcal{P}_1} - \mathbf{E}_{\mathcal{P}_2\,|\,\mathcal{F}_1}.$$

In summary, the order detection can be achieved by checking the rank of the projection result numerically. Under noisy condition, the numerical rank checking can be implemented by comparing the ratio of the second largest singular value to the largest one for orders from one to the upper bound.

F. Extension to CDMA MIMO Channels

In this subsection, the oblique projection techniques are applied to multiple-inputs multiple-outputs (MIMO) channels, in particular, to the multiuser detection (MUD) in code division multiple access (CDMA) systems. Here, we can see that by making reasonable assumptions, the ISI of each user can be first removed by using oblique projection, which makes the detection in the second step more flexible.

Consider a short-code CDMA system with processing gain of P. For each user-i, $i = 1, \ldots, K$, let $\mathbf{c}_i$ and $\mathbf{g}_i$ be, respectively, the spreading code and the propagation channel (including the effects of the relative time delays among the users). Thus, the received signal for user i can then be modeled as a single-input P-outputs linear channel with vector impulse response $\mathbf{h}_i$ obtained from the convolution of the code with the propagation channel $\mathbf{g}_i$. In particular, the vector $\mathbf{h}_i$ that contains all channel coefficients for user i can be written[2] as

$$\mathbf{h}_i = \text{Toeplitz}(\mathbf{c}_i, \mathbf{0})\mathbf{g}_i \triangleq \mathbf{C}_i\mathbf{g}_i.$$

[2] Here, Toeplitz$(\mathbf{x}, \mathbf{y})$ denotes a Toeplitz matrix with first row $\mathbf{y}$ and, except the first element, the first column $\mathbf{x}$.

Let the composite channel length of user i be $(L_i + 1)$. By stacking m blocks of data, we have the noiseless observation at the receiver

$$\mathbf{X}_m(t) = \sum_{i=1}^{K} \mathcal{F}_m(\mathbf{h}_i)\mathbf{S}_{i,L_i+m}(t).$$

Details of this model can be found in [16].

As in most subspace-based techniques in blind multiuser detection, we assume that

$$\mathcal{F}_m(\mathbf{h}) \triangleq [\mathcal{F}_m(\mathbf{h}_1), \ldots, \mathcal{F}_m(\mathbf{h}_K)] \tag{31}$$

has full column rank and that

$$\mathbf{S}(t) \triangleq \left[\mathbf{S}_{1,L_1+m}(t)^T, \ldots, \mathbf{S}_{K,L_K+m}(t)^T\right]^T \tag{32}$$

has full row rank. Further, we also assume that all the users have the same composite channel length,[3] i.e., $L_i = L, \forall i$. Using the same definition of future and past subspaces, we obtain

$$\mathbf{X}_{L+1}(t+L)\mathbf{E} = \sum_{i=1}^{K} \mathbf{h}_i \mathbf{s}_{i,t} \tag{33}$$

where $\mathbf{E}$ is the projector defined in Lemma 3.2, and $\mathbf{s}_{i,t}$ is the symbol sequence from user i. Note that after the oblique projection, all ISI has been completely removed, and the resulting model is one of the instantaneous mixture of input sequences.

While the channel estimation from (33) can be obtained the same way as in the subspace algorithm [13], [16], it is interesting to note that symbol estimation can be obtained directly using projections in a way similar to the zero-forcing [16] or minimum output energy (MOE) [6] detector but without explicit estimation of the channels. In fact, a projector predetermined from the knowledge of all codes can be used directly. Specifically, let $\mathbf{f}$ be a vector such that $\mathbf{f}_i \perp \mathcal{C}\{\mathbf{C}_j\}$ for $j \neq i$, and $\mathbf{f}_i \in \mathcal{C}\{\mathbf{C}_i\}$; then

$$\mathbf{f}_i^H \mathbf{X}_{L+1}(t+L)\mathbf{E} = \gamma \mathbf{s}_{i,t}$$

for some constant γ.

IV. Simulations

A. Simulation Conditions and Performance Measure

In this section, we present several simulation examples on channel and symbol estimation by the proposed algorithms and some existing deterministic algorithms.

Algorithms were compared by Monte Carlo simulations. We used normalized root mean square error (NRMSE) for channel estimation and root mean square error (RMSE) per symbol for symbol estimation:

$$\text{NRMSE} \triangleq \sqrt{\frac{1}{N_k \|\mathbf{h}\|^2} \sum_{k=1}^{N_k} \left\|\hat{\mathbf{h}}^{(k)} - \mathbf{h}\right\|^2} \tag{34}$$

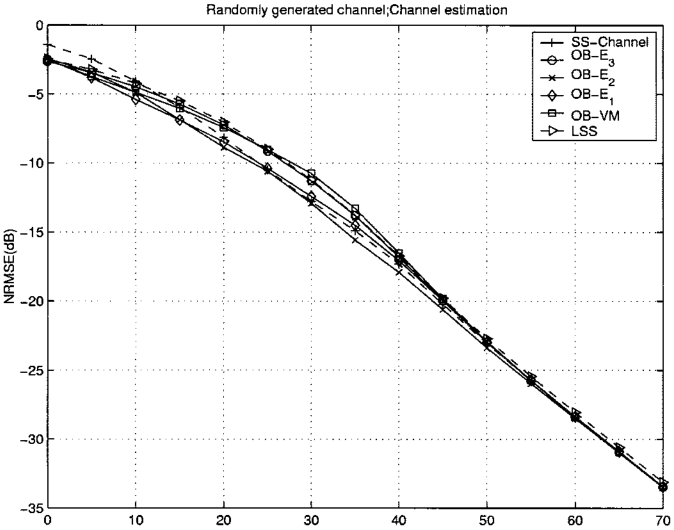

Fig. 7. Channel estimation comparison with randomly generated channels. Subspace channel estimation algorithm (SS-Channel), projector $\mathbf{E}_3, \mathbf{E}_2, \mathbf{E}_1$ proposed in this paper, Algorithm I proposed by Vandaele and Moonen (OB-VM), and the LSS algorithm (LSS).

$$\text{RMSE} \triangleq \sqrt{\frac{1}{N_k N_0} \sum_{k=1}^{N_k} \left\|\hat{\mathbf{s}}_t^{(k)} - \mathbf{s}_t\right\|^2} \tag{35}$$

where $\hat{\mathbf{h}}^{(k)}$ and $\hat{\mathbf{s}}_t^{(k)}$ were the estimated channel and symbol sequence from the kth run, and N_0 was the number of symbols in the estimated sequence. The additive noises were generated from i.i.d. zero mean complex Gaussian random sequence, and the signal-to-noise ratio (SNR) was defined as:

$$\text{SNR} \triangleq \frac{1}{P\sigma^2} E\left\{\sum_{j=1}^{P} \left|x_i^{(j)}\right|^2\right\} \tag{36}$$

where P was the output number, and σ^2 is the noise variance. The input sequence was an i.i.d. equally probable quadrature phase shift keying (QPSK) complex sequence.

B. Randomly Generated Channel

In the first example, we demonstrate the performances of the proposed algorithms for well-conditioned channels with known channel order. The channels were randomly generated for each run by the following three steps.

1) Generate $L+1$ tap coefficients according to i.i.d. complex Gaussian distribution.
2) Interpolate the channel according to the channel number P (over-sampling rate).
3) Normalize the channel to $||\mathbf{h}|| = 1$.

In this example, $L = 4$, and $P = 2$. The smoothing order m is chosen to be $m = L$. Note for these settings, the minimum necessary data length is $4m + 4L + 1 = 33$, which is the condition for $\mathbf{S}_{2m+2L+1}(t + m + L)$ to be a "fat" matrix. There are 500 channels used for each SNR and 100 symbols used for each Monte Carlo run. The performances curves for channel and symbol estimation are shown in Figs. 7 and 8. The performance comparisons with different data length are given in Figs. 9 and 10.

[3] This is valid for asynchronous system in which the summation of the time delay and the propagation channel order for each user is smaller than the code length.

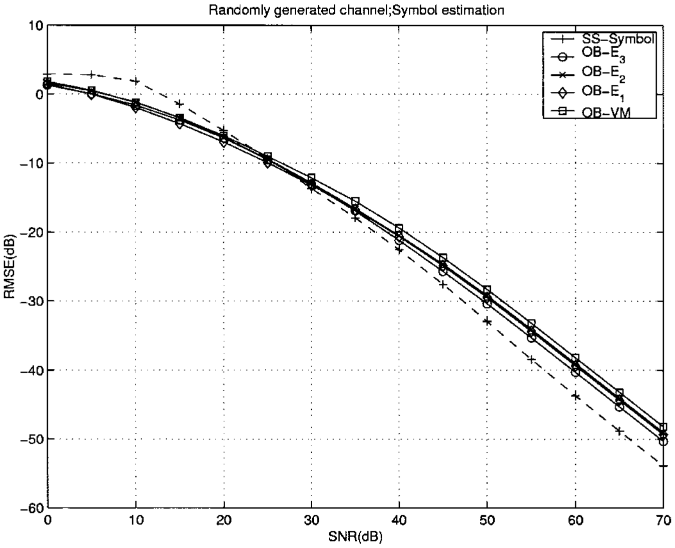

Fig. 8. Symbol estimation comparison with randomly generated channels. Subspace intersection symbol estimation algorithm (SS-Symbol), projector $\mathbf{E}_3, \mathbf{E}_2, \mathbf{E}_1$ proposed in this paper, and Algorithm I proposed by Vandaele and Moonen (OB-VM).

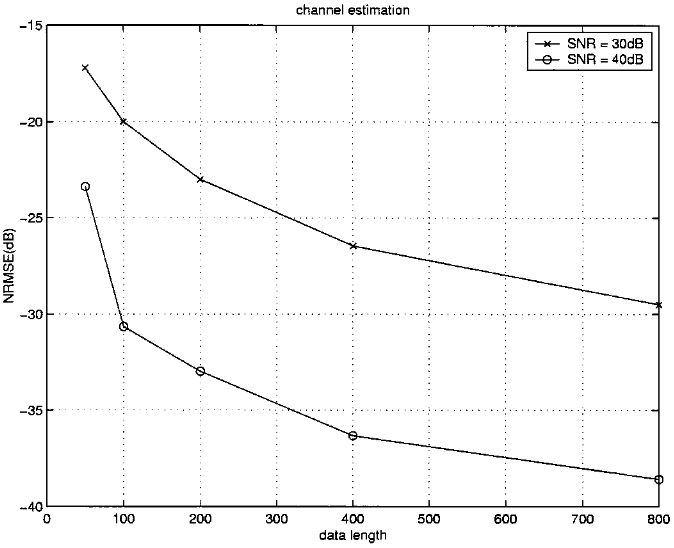

Fig. 9. Channel estimation with different data sample length ($\mathbf{E}_2$).

For channel estimation, we observe from Fig. 7 that the three oblique projection algorithms considered in this paper and the Algorithm I by Vandaele and Moonen (denoted as OB-VM) performed similarly as the subspace algorithm [8] and the LSS algorithm [12]. For symbol estimations, we considered the oblique projection techniques with the subspace intersection symbol estimation technique proposed by van der Veen [14]. From Fig. 8, we can see that these symbol estimation algorithms also performed comparably. It is interesting to point out that the "SS-symbol" outperformed the other algorithms by 5 dB in high SNR region. One possible reason is that in oblique projection-based algorithms, symbols were estimated by forming the intersection of two subspaces $\mathcal{F}_2$ and $\mathcal{P}_2$, whereas in the "SS-symbol," symbols were gotten by performing intersection of all subspaces containing $\mathbf{s}_t$.

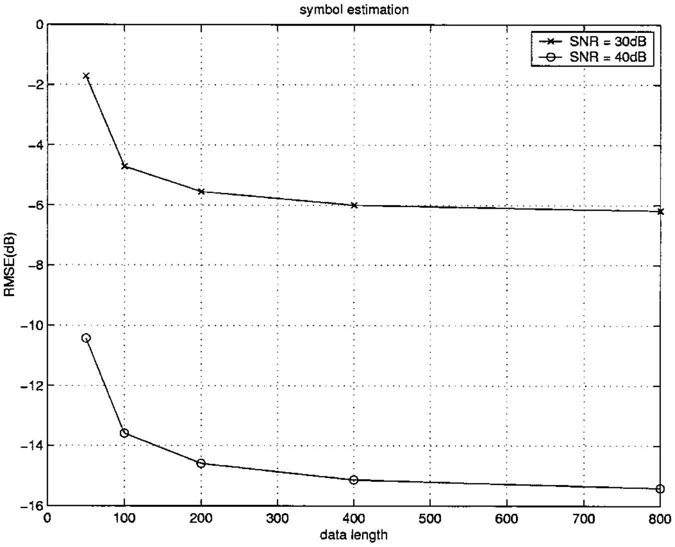

Fig. 10. Symbol estimation with different data sample length ($\mathbf{E}_2$).

TABLE I
CHANNEL COEFFICIENTS OF A MULTIPATH CHANNEL

Tap	1	2	3	4	5	6
Channel-1:real	−0.0031	−0.0109	+0.1522	+0.3789	−0.0301	−0.0032
Channel-1:imaginary	−0.0017	−0.0025	+0.0705	+0.5930	−0.0348	−0.0017
Channel-2:real	−0.0016	−0.0263	+0.4409	+0.0766	−0.0042	−0.0017
Channel-2:imaginary	−0.0047	−0.0433	+0.4736	+0.2168	−0.0154	−0.0044

The simulation comparison between TLS and LS projections are also made with projector $\mathbf{E}_2$. The results showed that TLS projections only provide negligible performance improvement and, hence, were omitted here. As a result, the LS solution is more preferable based on the consideration of the computation complexity and adaptivity of the implementation.

The simulation results with different length of samples are shown in Figs. 9 and 10. The proposed approaches depend on the estimates of the row spaces of the input matrix from the output matrix. Unfortunately, the row space estimation will not converge to the true basis as the data length goes to infinity, unless special considerations are given as in [14]. Therefore, the plots in Figs. 9 and 10 flatten, as expected when data lenth increases.

C. Multipath Channel

In the second example, we present the simulation results for multipath channels that have small head and tail taps. The channel has two outputs and channel order of 5, and the coefficients are given in Table I. In the simulation, 200 Monte Carlo runs are used for each SNR, and 100 symbols are used for each run. The performance curves are plotted in Figs. 11 and 12. As a comparison, we also include, in Figs. 13 and 14, the curves of several deterministic algorithms using the channel order given by the minimum description length (MDL) model selection criterion [17].

From Fig. 11, we can see that the proposed order-detection scheme together with the corresponding projector considerably outperformed those algorithms without order detection capability. It also performed better than J-LSS[12] for SNR less than 30 dB. In the high SNR region, the curve finally coincided

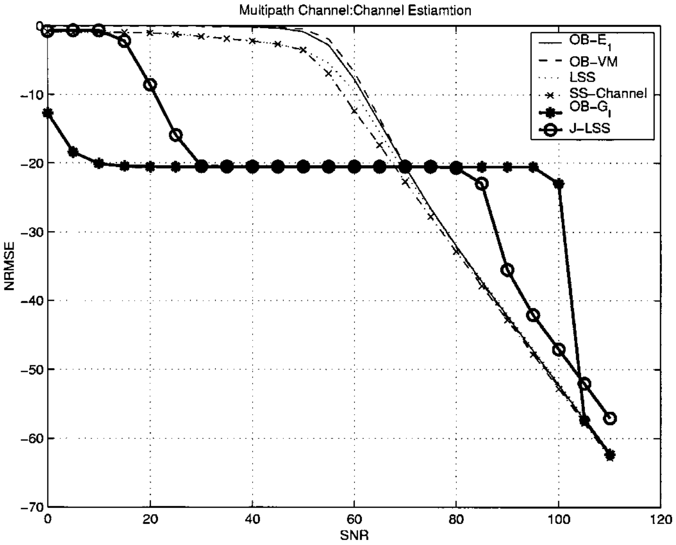

Fig. 11. Channel estimation comparison with multipath channels. Projector $\mathbf{E}_1$, Algorithm I proposed by Vandaele and Moonen (OB-VM), LSS algorithm (LSS), subspace channel estimation algorithm (SS-Channel), projector $\mathbf{G}_I$ proposed in this paper (OB-$\mathbf{G}_I$), and joint-LSS algorithm (J-LSS).

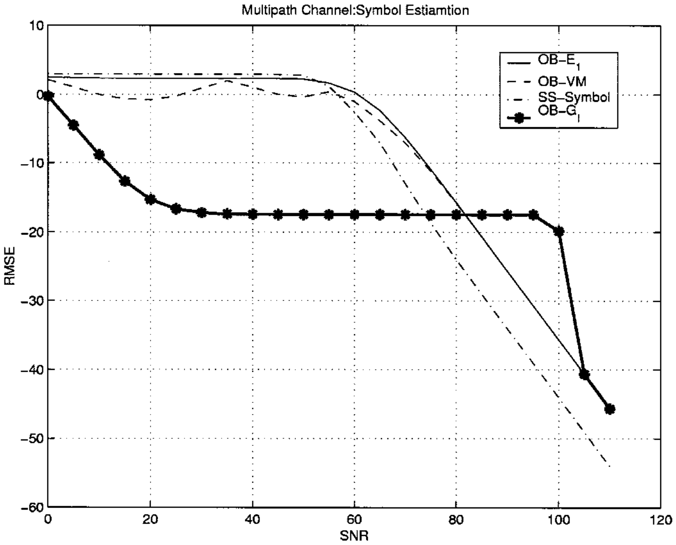

Fig. 12. Symbol estimation comparison with multipath channels. Projector $\mathbf{E}_1$, Algorithm I proposed by Vandaele and Moonen (OB-VM), subspace intersection symbol estimation algorithm (SS-Symbol), and projector $\mathbf{G}_I$ proposed in this paper (OB-$\mathbf{G}_I$).

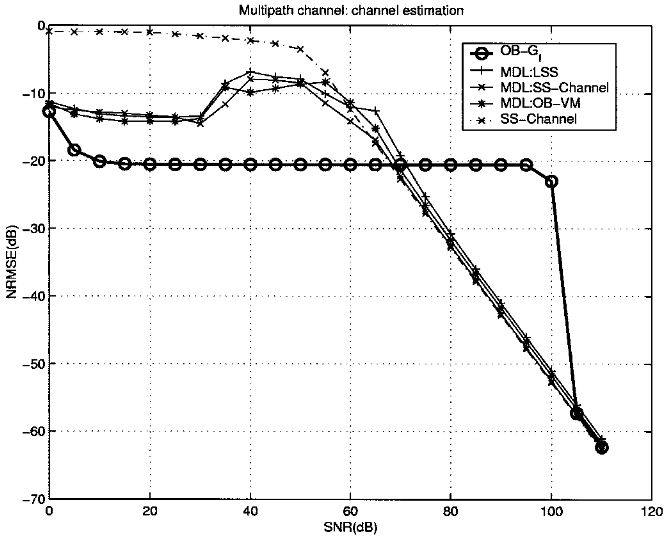

Fig. 13. Channel estimation comparison with multipath channels. Projector $\mathbf{G}_I$ (OB-$\mathbf{G}_I$), LSS algorithm with MDL order selection (MDL:LSS), subspace channel estimation with MDL order selection (MDL:SS-Channel), Algorithm I proposed by Vandaele and Moonen with MDL order selection (MDL:OB-VM), subspace channel estimation with true channel order (SS-Channel).

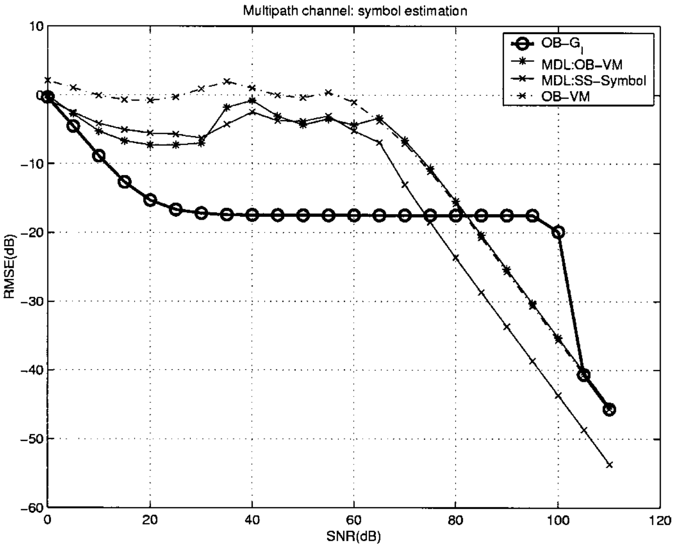

Fig. 14. Symbol estimation comparison with multipath channels. Projector $\mathbf{G}_I$ (OB-$\mathbf{G}_I$), SS-symbol with MDL order selection (MDL:SS-symbol), and Algorithm I proposed by Vandaele and Moonen with MDL order selection (MDL:OB-VM) and without order selection (OB-VM).

with other deterministic methods when the order was correctly chosen by the algorithm at high SNR. For symbol estimation, which is shown in Fig. 12, the proposed algorithm also worked better than the other algorithms without order selection capability before SNR $= 70$ dB.

The saturation behavior of the the estimator using $\mathbf{G}_I$ at high SNR was observed. This is due to the bias of the detection algorithm. Specifically, for the given channel, when the order is underestimated, only several dominant coefficients need to be estimated. In this case, $\mathbf{G}_I$ offers a good estimate, even in a low SNR region. Starting from 10 dB, the dominant part of NRMSE for $\mathbf{G}_I$ is the omitted coefficients due to the underestimated order. When the SNR is high, $\mathbf{G}_I$ is still using the underestimated channel order. Therefore, the performance of $\mathbf{G}_I$ (J-LSS) is worse than the other algorithms around 100 dB. When $\mathbf{G}_I$ gives out the correct channel order, the performace curves converge.

In Figs. 13 and 14, it can be seen that the MDL order selection resulted in about a 10–dB performance improvement for the corresponding channel and symbol estimation algorithms in the low SNR region. However, the performances were still worse than the proposed oblique projection algorithm with projector $\mathbf{G}_I$. The peculiar behavior of the MDL-based techniques that lead to an increase of NRMSE around 30 dB can be explained by the fact that as a good order detection algorithm, MDL began to provide correct order detection with high probability, which worsens the performance of the channel and symbol estimation (see the curve with true channel order for all SNRs).

V. Conclusion

Joint channel and symbol estimation algorithm using oblique projections is considered in this paper. The central idea of this approach is to formulate the problem of channel-symbol estimation as one of least squares smoothing and decompose the observation space into past, future, and current subspaces. The use of oblique projection leads to several new estimation algorithms, all of which can be implemented directly from oblique projections. One of the main advantages of the approaches presented in this paper, aside from the simple geometrical formulation of the problem, is that these algorithms can be easily implemented using recursive techniques while maintaining the performance similar to that of subspace techniques.

Appendix

Proof of Lemma 3.2:

Proof: The proof follows the steps in the general projector construction lemmas in [9]. We prove the details of (18) only for $\mathbf{E}_1$ since the proof for the other two cases is similar. First, we prove that projection matrices $\mathbf{A} \triangleq \mathbf{E}_{\mathcal{F}_2 \,|\, \mathcal{P}_1}$ and $\mathbf{B} \triangleq \mathbf{E}_{\mathcal{P}_2 \,|\, \mathcal{F}_1}$ commute with each other. To verify this, we only need to prove that for any vector $\mathbf{x} \in \mathcal{C}^n$, $\mathbf{xAB} = \mathbf{xBA}$. Since $\mathcal{F}_1, \mathcal{P}_1$ and sp$\{\mathbf{s}_t\}$ are disjoint, any vector $\mathbf{x} \in \mathcal{C}^n$ can be uniquely decomposed into four parts $\mathbf{x} = \mathbf{x}_{\mathcal{F}_1 \,|\, \mathcal{P}_1} + \mathbf{x}_{\mathcal{P}_1 \,|\, \mathcal{F}_1} + \mathbf{x}_{\mathrm{sp}\{\mathbf{s}_t\}} + \mathbf{x}_{(\,\cdot\,)^\perp}$, where $\mathbf{x}_{(\,\cdot\,)^\perp}$ denotes the part of $\mathbf{x}$ in the orthogonal complement subspace of $\mathcal{F}_1 \oplus \mathcal{P}_1 \oplus \mathrm{sp}\{\mathbf{s}_t\}$. Thus, by the definition of $\mathbf{E}_{\mathcal{F}_2 \,|\, \mathcal{P}_1}, \mathbf{E}_{\mathcal{P}_2 \,|\, \mathcal{F}_1}$ and the property of the oblique projection, we can show that

$$\mathbf{x} \cdot \mathbf{E}_{\mathcal{F}_2 \,|\, \mathcal{P}_1} \mathbf{E}_{\mathcal{P}_2 \,|\, \mathcal{F}_1} = \mathbf{x}_{\mathrm{sp}\{\mathbf{s}_t\}} = \mathbf{x} \cdot \mathbf{E}_{\mathcal{P}_2 \,|\, \mathcal{F}_1} \mathbf{E}_{\mathcal{F}_2 \,|\, \mathcal{P}_1}.$$

Next, we show that $\mathbf{E}_1 = \mathbf{AB}$ is a projector, i.e., it is idempotent [9]

$$\begin{aligned} (\mathbf{AB})^2 &= (\mathbf{AB})(\mathbf{BA}) \\ &= \mathbf{ABA} \\ &= \mathbf{AAB} \\ &= \mathbf{AB}. \end{aligned}$$

Finally, we show that the range space and null space of the projector are $\mathcal{F}_2 \cap \mathcal{P}_2$ and $\mathcal{F}_1 \oplus \mathcal{P}_1$, respectively. For all $\mathbf{x}$ in the range space of $\mathbf{AB}, \mathbf{xA} = \mathbf{x}(\mathbf{AB})\mathbf{A} = \mathbf{xBA} = \mathbf{xAB} = \mathbf{x}$. Similarly, $\mathbf{xB} = \mathbf{x}$. Therefore, the range space of $\mathbf{AB}$ is contained in $\mathcal{F}_2 \cap \mathcal{P}_2$. On the other hand, any $\mathbf{x} \in \mathcal{F}_2 \cap \mathcal{P}_2, \mathbf{x} = \mathbf{xAB}$. Therefore, the range space of $\mathbf{AB}$ is equal to $\mathcal{F}_2 \cap \mathcal{P}_2$. For vectors in the null space of $\mathbf{AB}$, $\mathbf{xAB} = \mathbf{0}$, which implies $\mathbf{xA} \in \mathcal{F}_1$ (note that $\mathcal{F}_1$ is the null space of projector $\mathbf{B}$), but $\mathbf{x} = \mathbf{xA} + \mathbf{x}(\mathbf{I} - \mathbf{A})$, and $\mathbf{x}(\mathbf{I} - \mathbf{A}) \in \mathcal{P}_1$ (the null space of projector $\mathbf{A}$); thus, $\mathbf{x} \in \mathcal{F}_1 \oplus \mathcal{P}_1$. Conversely, $\mathbf{x} \in \mathcal{F}_1 \oplus \mathcal{P}_1 \Rightarrow \mathbf{x} = \mathbf{x}_1 + \mathbf{x}_2$, where $\mathbf{x}_1 \in \mathcal{F}_1$ and $\mathbf{x}_2 \in \mathcal{P}_1$. Therefore, $\mathbf{xAB} = \mathbf{xBA} = \mathbf{0}$, implying that $\mathbf{x}$ is in the null space of $\mathbf{AB}$. Therefore, the null space of $\mathbf{AB}$ is equal to $\mathcal{F}_1 \oplus \mathcal{P}_1$.

We have, therefore, proved that $\mathbf{E}_{\mathcal{F}_2 \,|\, \mathcal{P}_1} \mathbf{E}_{\mathcal{P}_2 \,|\, \mathcal{F}_1}$ is a projector with $\mathcal{P}_2 \cap \mathcal{F}_2$ as the range space and $\mathcal{P}_1 \oplus \mathcal{F}_1$ as the null space. Hence, by the property of oblique projection, (18) is proved. □

References

[1] R. T. Behrens and L. L. Sharf, "Signal processing applications of oblique projection operators," *IEEE Trans. Signal Processing*, vol. 42, pp. 1413–1424, June 1994.

[2] R. E. Blahut, *Algebraic Methods for Signal Processing and Communications Coding*. New York: Springer-Verlag, 1992.

[3] D. Gesbert, P. Duhamel, and S. Mayrargue, "On-line blind multichannel equalization based on mutually referenced filters," *IEEE Trans. Signal Processing*, vol. 45, pp. 2307–2317, Sept. 1997.

[4] D. Gesbert, A. J. van der Veen, and A. Paulraj, "On the equivalence of blind equalizers based on MRE and subspace intersections," *IEEE Trans. Signal Processing*, vol. 47, pp. 856–859, Mar. 1999.

[5] G. H. Golub and C. F. Van Loan, *Matrix Computations*. Baltimore, MD: The Johns Hopkins Univ. Press, 1990.

[6] M. Honig, U. Madhow, and N. Verdú, "Blind adaptive multiuser detection," *IEEE Trans. Inform. Theory*, vol. 41, pp. 944–960, July 1995.

[7] S. Van Huffel and J. Vandewalle, *The Total Least Squares Problem:Computational Aspects and Analysis*. Philadelphia, PA: SIAM, 1991.

[8] 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.

[9] C. R. Rao and S. K. Mitra, *Generalized Inverse of Matrices and Its Applications*. New York: Wiley, 1971.

[10] D. Slock, "Blind fractionally-spaced equalization, perfect reconstruction filterbanks, and multilinear prediction," in *Proc. ICASSP Conf.*, Adelaide, Australia, Apr. 1994.

[11] L. Tong and S. Perreau, "Multichannel blind channel estimation: From subspace to maximum likelihood methods," in *Proc. IEEE*, vol. 86, Oct. 1998, pp. 1951–1968.

[12] L. Tong and Q. Zhao, "Joint order detection and blind channel estimation by least squares smoothing," *IEEE Trans. Signal Processing*, vol. 47, pp. 2345–2355, Sept. 1999.

[13] M. Tsatsanis and G. B. Giannakis, "Multirate filter banks for code-division multiple access systems," in *Proc. ICASSP*, vol. II, 1995, pp. 1484–1487.

[14] A. van der Veen, S. Talwar, and A. Paulraj, "A subspace approach to blind space-time signal processing for wireless communication systems," *IEEE Trans. Signal Processing*, vol. 45, pp. 173–190, Jan. 1997.

[15] P. Vandaele and M. Moonen, "Two deterministic blind channel estimation algorithms based on oblique projections," *Signal Process.*, vol. 80, no. 3, pp. 481–495, Mar. 2000.

[16] X. Wang and H. V. Poor, "Blind equalization and multiuser detection in dispersive CDMA channels," *IEEE Trans. Commun.*, vol. 46, pp. 91–103, Jan. 1998.

[17] M. Wax and T. Kailath, "Detection of signals by information theoretic criteria," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. ASSP-33, no. 2, pp. 387–392, Apr. 1985.

[18] 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.

[19] Q. Zhao and L. Tong, "Adaptive blind channel estimation by least squares smoothing," *IEEE Trans. Signal Processing*, vol. 47, pp. 3000–3012, Nov. 1999.

Xiang Yu received the B.E. and M.E. degrees from Tsinghua University, Beijing, China, in 1996 and 1999, respectively. In 2001, he received the M.S. degree in electrical engineering from Cornell University, Ithaca, NY.

Currently, he is a system engineer in Aware, Inc., Bedford, MA.

Lang Tong (S'87–M'91–SM'01) received the B.E. degree from Tsinghua University, Beijing, China, in 1985 and the M.S. and Ph.D. degrees in electrical engineering in 1987 and 1990, respectively, from the University of Notre Dame, Notre Dame, IN.

He was a Postdoctoral Research Affiliate with the Information Systems Laboratory, Stanford University, Stanford, CA, in 1991. Currently, he is an Associate Professor with the School of Electrical and Computer Engineering, Cornell University, Ithaca, NY. His areas of interest include statistical signal processing, adaptive receiver design for communication systems, signal processing for communication networks, and information theory.

Dr. Tong received the Young Investigator Award from the Office of Naval Research in 1996 and the Outstanding Young Author Award from the IEEE Circuits and Systems Society.