Generalized Blind Subspace Channel Estimation

Wei Kang and Benoit Champagne
Department of Electrical and Computer Engineering, McGill University
3480 University, Montreal, Quebec, Canada H3A 2A7
wkang@tsp.ece.mcgill.ca, champagne@ece.mcgill.ca

***Abstract*— In this paper, we present a systematic study of the subspace-based blind channel estimation method. We first formulate a general signal model of multiple simultaneous signals transmitted through vector channels. Based on this model, we then propose a generalized subspace-based channel estimator by minimizing a novel cost function, which incorporates the set of kernel matrices of the signals sharing the target channel via a weighted sum of projection errors. We investigate the asymptotic performance of the proposed estimator, i.e. bias, covariance, mean square error (MSE) and Cramer-Rao bound, for large numbers of independent observations. We show that the performance of the estimator can be optimized by increasing the number of kernel matrices and by using a special set of weights in the cost function. Finally, we consider the application of the proposed estimator to a down-link CDMA system. The results of the computer simulations fully support our analytical developments.**

I. INTRODUCTION

Recently, blind channel estimation algorithms have received considerable attention due to their advantages in terms of bandwidth efficiency [1]. Of particular interest within the family of blind algorithms are the so-called *subspace-based* blind channel estimation algorithms, which derive their properties from the second-order statistics of the received signals.

During the past decade, subspace-based channel estimation algorithms have been developed for and applied to various vector channels, such as: SIMO channels [2]; frequency selective fading channel in DS-CDMA systems [3], [4]; multiple receiver antennae and/or multiple transmitter antennae channels in CDMA systems [5]. Although these algorithms were developed separately for certain specific transmission scenarios, the similarities among them indicate that there must exist some common features of the underlying system models, which provide for the feasibility of the subspace channel estimation. However, so far these common features have not been studied in the literature.

Besides, among the existing subspace-based channel estimation algorithms, a majority of them only utilize a single signal component to estimate the target channel, e.g. [3]. However, in many situations of interest, the target channel is shared by multiple signal components simultaneously, as in e.g. a typical downlink environment in cellular systems [4], time dispersive channels [4] or space-time block coded channels [5]. Then the problem of utilizing multiple signal components to estimate the target channel arises naturally. A pioneering work on this topic appeared in [2], which tackles the ISI channel estimation problem in SIMO systems. Extension to the ISI channel in CDMA can be found in [4]. So far, there has not been a study that quantifies the effects of using multiple signal components in subspace-based blind channel estimation.

Motivated by the above considerations, we present a systematic study of the subspace-based blind channel estimation method. We first formulate a general signal model of multiple simultaneous signals transmitted through vector channels, We then proposed a generalized subspace-based channel estimator by minimizing a novel cost function, which incorporates the set of kernel matrices of the signal components sharing the target channel via a weighted sum of projection errors. We investigate the asymptotic performance of the proposed estimator, i.e. bias, covariance, mean square error (MSE) and Cramer-Rao bound (CRB) for large numbers of independent observations. Finally, we consider the application of the proposed estimator to a down-link CDMA system operating in frequency selective fading channel with negligible ISI. The results of the computer simulations fully support our analysis.

II. PROBLEM FORMULATION

We consider the following model of an L-dimensional received signal vector in a communication system:

$$\mathbf{r} = \sum_{i=1}^{N} \gamma_i b_i \mathbf{C}_i \mathbf{h}_i + \mathbf{e} \tag{1}$$

where N is the number of individual symbols that comprise the received signal vector, γ_i is a real-valued received amplitude, b_i is the i-th information symbol, $\mathbf{C}_i$ is defined as a kernel matrix with size $L \times M$, $\mathbf{h}_i$ is an $M \times 1$ normalized channel vector (i.e. $\|\mathbf{h}_i\| = 1$), and $\mathbf{e}$ is an $L \times 1$ additive noise vector. We assume that the information symbols b_i, for $i = 1, \ldots, N$, are independent and identically distributed with zero mean and unit variance. The additive noise vector $\mathbf{e}$ is circularly complex Gaussian with covariance matrix $\sigma^2 \mathbf{I}_L$ and is independent from the information symbols b_i. We define

$$\mathbf{b} \triangleq [b_1, \ldots, b_N]^T \tag{2}$$

$$\mathbf{\Gamma} \triangleq \text{diag}[\gamma_1, \ldots \gamma_N] \tag{3}$$

$$\mathbf{W} \triangleq [\mathbf{w}_1, \ldots, \mathbf{w}_N] \tag{4}$$

where $\mathbf{w}_i \triangleq \mathbf{C}_i \mathbf{h}_i$ for $i = 1, \ldots, N$ is the effective signature waveform of the i-th information symbol, i.e. combined effect of channel and kernel matrix as seen by the receiver. Using the above matrix notations, the signal model (1) can be expressed more compactly as

$$\mathbf{r} = \mathbf{W}\mathbf{\Gamma}\mathbf{b} + \mathbf{e} \tag{5}$$

In the sequel, we refer to the individual products $\gamma_i b_i \mathbf{C}_i \mathbf{h}_i$ ($i = 1, ..., N$) in (1) as *signal components*. We assume that these N signal components experience J different channels, $1 \leq J \leq N$. Then we separate the N signal components into J groups, such that the signal components in each group share the same channel. We denote the number of signal components in the m-th group as K^m ($m = 1, \ldots, J$), so that $\sum_{m=1}^{J} K^m = N$. In the m-th group, we use the superscript m to denote group affiliation, as in the common channel parameter $\mathbf{h}^m$, and we use the superscript l to further distinguish among the K^m signal components, as in $\gamma^{m,l}$, $b^{m,l}$, $\mathbf{C}^{m,l}$ and $\mathbf{w}^{m,l}$.

This general model is applicable to a multitude of communication systems, e.g. [2]-[5]. For details, please check [6].

Within the above framework, the goal of blind channel estimation is to determine a target channel vectors $\mathbf{h}^m$, $m = 1, \ldots, J$, using T observations of the received signal vector in (1). In blind channel estimation, the transmitted information symbols, as represented by vector $\mathbf{b}$ (2), are unknown. To estimate the target channel vector $\mathbf{h}^m$, at least one kernel matrix in the m-th group needs to be known by the estimating algorithm. In practice, the specific available knowledge of the kernel matrices depends on the particular system under consideration. Below, we formulate a generalized cost function for subspace-based blind channel estimation, which incorporates the set of kernel matrices of the signal components sharing the target channel. We then investigate the asymptotic performance of the estimator when the number of independent observations T is large.

III. Generalized Blind Subspace Channel Estimation

A. Theoretical Foundation

Let $\mathbf{R}$ denote the covariance matrix of received signal vector $\mathbf{r}$ in (1):

$$\mathbf{R} = E[\mathbf{r}\mathbf{r}^H] = \mathbf{W}\mathbf{\Gamma}^2\mathbf{W}^H + \sigma^2 \mathbf{I}_L \tag{6}$$

Blind subspace methods exploit the special structure of $\mathbf{R}$ to estimate the channel parameters. Specifically, let us express the EigenValue Decomposition (EVD) of $\mathbf{R}$ in the form

$$\mathbf{R} = \mathbf{U}\mathbf{\Lambda}\mathbf{U}^H \tag{7}$$

where $\mathbf{\Lambda} = \text{diag}[\lambda_1, \ldots, \lambda_L]$ denotes the eigenvalue matrix, with the eigenvalues in a non-increasing order, and $\mathbf{U}$ is a unitary matrix that contains the corresponding eigenvectors. Since the rank of matrix $\mathbf{W}\mathbf{\Gamma^2}\mathbf{W}^H$ (6) is N, it follows that

$$\lambda_1 \geq \cdots \geq \lambda_N > \lambda_{N+1} = \cdots = \lambda_L = \sigma^2 \tag{8}$$

Thus, the eigenvalues can be separated into two distinct groups, the signal eigenvalues and the noise eigenvalues, respectively represented by matrices $\mathbf{\Lambda}_s \triangleq \text{diag}[\lambda_1, \ldots, \lambda_N]$ and $\mathbf{\Lambda}_n \triangleq \text{diag}[\lambda_{N+1}, \ldots, \lambda_L]$. Accordingly, the eigenvectors can be separated into the signal and noise eigenvectors, represented by matrices $\mathbf{U}_s$ and $\mathbf{U}_n$. With these notations, the EVD in (7) can be expressed as

$$\mathbf{R} = \begin{bmatrix} \mathbf{U}_s & \mathbf{U}_n \end{bmatrix} \begin{bmatrix} \mathbf{\Lambda}_s & \mathbf{0} \\ \mathbf{0} & \mathbf{\Lambda}_n \end{bmatrix} \begin{bmatrix} \mathbf{U}_s^H \\ \mathbf{U}_n^H \end{bmatrix} \tag{9}$$

The columns of $\mathbf{U}_s$ span the so-called signal subspace with dimension N, while those of $\mathbf{U}_n$ span its orthogonal complement, i.e. the noise subspace. The signal subspace is indeed equal to the space spanned by the columns of $\mathbf{W}$:

$$\text{Span}[\mathbf{W}] = \text{Span}[\mathbf{U}_s] \perp \text{Span}[\mathbf{U}_n] \tag{10}$$

To estimate the target channel vector $\mathbf{h}^m$, which is shared by the signal components in the m-th group, we select $1 \leq P \leq K^m$ effective signature waveforms from the m-th group, say $\mathbf{w}^{m,j}$ ($j = 1, \ldots, P$) without loss of generality, and construct a matrix $\bar{\mathbf{W}} \triangleq [\mathbf{w}^{m,1}, \ldots, \mathbf{w}^{m,P}]$. Then

$$\text{Span}[\bar{\mathbf{W}}] \subseteq \text{Span}[\mathbf{W}] \perp \text{Span}[\mathbf{U}_n] \tag{11}$$

Consequently, $\mathbf{U}_n^H \bar{\mathbf{W}} = \mathbf{0}$. Defining

$$\mathcal{U}_s \triangleq \mathbf{I}_P \otimes \mathbf{U}_s \tag{12}$$

$$\mathcal{U}_n \triangleq \mathbf{I}_P \otimes \mathbf{U}_n \tag{13}$$

$$\mathcal{C}^T \triangleq [(\mathbf{C}^{m,1})^T, \ldots, (\mathbf{C}^{m,P})^T] \tag{14}$$

where $\otimes$ represents the Kronecker product, and applying vectorization operation on $\mathbf{U}_n^H \bar{\mathbf{W}}$, we obtain

$$\text{vec}[\mathbf{U}_n^H \bar{\mathbf{W}}] = \mathcal{U}_n^H \text{vec}[\bar{\mathbf{W}}] = \mathcal{U}_n^H \mathcal{C} \mathbf{h}^m = \mathbf{0} \tag{15}$$

B. The Algorithm

In practice, the covariance matrix $\mathbf{R}$ is usually unknown and must be estimated from the observed data via time averaging. Assuming a locally stationary environment, one such estimate based on a rectangular window of T samples is given by $\hat{\mathbf{R}} = \frac{1}{T}\sum_{j=1}^{T} \mathbf{r}_j \mathbf{r}_j^H$, where $\mathbf{r}_j$ now denotes the received signal vector at the j-th time instant (with similar modifications for other quantities of interest $\mathbf{b} \rightarrow \mathbf{b}_j$, $\mathbf{e} \rightarrow \mathbf{e}_j$), for $j = 1, \ldots, T$. In practice, the EVD is applied to $\hat{\mathbf{R}}$, resulting in the noisy estimates of $\mathbf{U}_s$,$\mathbf{U}_n$, $\mathbf{\Lambda}_s$ and $\mathbf{\Lambda}_n$, respectively denoted as $\hat{\mathbf{U}}_s$,$\hat{\mathbf{U}}_n$, $\hat{\mathbf{\Lambda}}_s$ and $\hat{\mathbf{\Lambda}}_n$. Consequently, the noisy estimates of $\mathcal{U}_s$ (12) and $\mathcal{U}_n$ (13) are defined as $\hat{\mathcal{U}}_n \triangleq \mathbf{I}_P \otimes \hat{\mathbf{U}}_n$ and $\hat{\mathcal{U}}_s \triangleq \mathbf{I}_P \otimes \hat{\mathbf{U}}_s$.

In this work, we consider the following optimization criterion for the blind estimation of channel vector $\mathbf{h}^m$:

$$\hat{\mathbf{h}}^m = \arg \min_{||\mathbf{t}||=1} \mathbf{t}^H \mathbf{D} \mathbf{t} \tag{16}$$

where $\mathbf{D} \triangleq \mathcal{C}^H \hat{\mathcal{U}}_n \hat{\mathcal{U}}_n^H \mathcal{C} = \sum_{i=1}^{P} (\mathbf{C}^{m,i})^H \hat{\mathbf{U}}_n \hat{\mathbf{U}}_n^H \mathbf{C}^{m,i}$. Ideally, if $\hat{\mathcal{U}}_n = \mathcal{U}_n$ and the identifiability condition (see [3]) is satisfied, all the eigenvalues of $\mathcal{C}^H \hat{\mathcal{U}}_n \hat{\mathcal{U}}_n^H \mathcal{C}$ in the above criterion are positive except the smallest one, which is equal to 0. However, in practice, the estimation error in $\hat{\mathcal{U}}_n$ may result in a positive perturbation in the smallest eigenvalue so that the matrix $\mathcal{C}^H \hat{\mathcal{U}}_n \hat{\mathcal{U}}_n^H \mathcal{C}$ is positive definite. In this case, (15) does not have a (non-trivial) solution, but the target channel vector still can be estimated by minimizing the cost function

TABLE I
GENERALIZED BLIND SUBSPACE CHANNEL ESTIMATION ALGORITHM

$\hat{\mathbf{R}} = \frac{1}{T}\sum_{j=1}^{T}\mathbf{r}_j\mathbf{r}_j^H$
$\hat{\mathbf{R}} = [\,\hat{\mathbf{U}}_s\,\hat{\mathbf{U}}_n] \begin{bmatrix}\hat{\mathbf{\Lambda}}_s & \mathbf{0}\\ \mathbf{0} & \hat{\mathbf{\Lambda}}_n\end{bmatrix} \begin{bmatrix}\hat{\mathbf{U}}_s^H\\ \hat{\mathbf{U}}_n^H\end{bmatrix}$
P is user specified
$\hat{\mathcal{U}}_n \triangleq \mathbf{I}_P \otimes \hat{\mathbf{U}}_n$
$\mathcal{C}^T \triangleq [(\mathbf{C}^{m,1})^T, \ldots, (\mathbf{C}^{m,P})^T]$
α^i, $i = 1, \ldots, P$, are user specified
$\mathbf{A} \triangleq \text{diag}[\sqrt{\alpha^1}, \ldots, \sqrt{\alpha^P}]$
$\mathcal{A} \triangleq \mathbf{A} \otimes \mathbf{I}_N$
Construct the matrix $\mathcal{C}^H\hat{\mathcal{U}}_n\mathcal{A}\mathcal{A}\hat{\mathcal{U}}_n^H\mathcal{C}$
$\hat{\mathbf{h}}^m$ is the smallest eigenvector of $\mathcal{C}^H\hat{\mathcal{U}}_n\mathcal{A}\mathcal{A}\hat{\mathcal{U}}_n^H\mathcal{C}$

in (16). Thus, we conclude that the optimization criterion in (16) is more robust to the perturbation of $\mathcal{U}_n$ than (15).

The choice of kernel matrices $\mathbf{C}^{m,i}$ included in the proposed criterion (16) is specified by the user, allowing a generalization of previous works. For example, the single signal algorithm in [3] can be viewed as a special case of (16) with $P = K^m = 1$, while the multiple signals algorithms in [2] corresponds to $P = K^m$. Here we can use any value of $1 \le P \le K^m$.

A further modification to the above criterion is motivated by the consideration of performance (see Sections IV and V). Specifically, we shall allow the assignment of different weights to the different terms $(\mathbf{C}^{m,i})^H\hat{\mathbf{U}}_n\hat{\mathbf{U}}_n^H\mathbf{C}^{m,i}$, i.e.

$$\hat{\mathbf{h}}^m = \arg\min_{||\mathbf{t}||=1} \mathbf{t}^H[\sum_{i=1}^{P}\alpha^i(\mathbf{C}^{m,i})^H\hat{\mathbf{U}}_n\hat{\mathbf{U}}_n^H\mathbf{C}^{m,i}]\mathbf{t} \quad (17)$$

where $\alpha^i > 0$ are user-specified weight parameters. Defining $\mathcal{A} \triangleq \text{diag}[\sqrt{\alpha^1}, \ldots, \sqrt{\alpha^P}] \otimes \mathbf{I}_{L-N}$, criterion (17) can be expressed in matrix form as:

$$\hat{\mathbf{h}}^m = \arg\min_{||\mathbf{t}||=1} \mathbf{t}^H\mathcal{C}^H\hat{\mathcal{U}}_n\mathcal{A}\mathcal{A}^H\hat{\mathcal{U}}_n^H\mathcal{C}\mathbf{t} \quad (18)$$

From an algorithmic viewpoint, the solution $\hat{\mathbf{h}}^m$ of (18) can be calculated as the eigenvector corresponding to the smallest eigenvalue of $\mathcal{C}^H\hat{\mathcal{U}}_n\mathcal{A}\mathcal{A}^H\hat{\mathcal{U}}_n^H\mathcal{C}$. The resulting estimation algorithm is summarized in Table I; we call it the *generalized* blind subspace channel estimation algorithm.

IV. ASYMPTOTIC PERFORMANCE ANALYSIS

In this section, we investigate the asymptotic performance of the proposed generalized blind subspace channel estimator (see Table I). We define the estimation error as $\Delta\mathbf{h}^m \triangleq \hat{\mathbf{h}}^m - \mathbf{h}^m$, where $\hat{\mathbf{h}}^m$ and $\mathbf{h}^m$ respectively denote the estimated and true target channel vector for the m-th group. The performance criteria of interest are the bias, covariance and mean square error of the proposed estimator, respectively defined as

$$\text{Bias} \triangleq E[\Delta\mathbf{h}^m] \quad (19)$$

$$\text{Cov} \triangleq E[(\Delta\mathbf{h}^m - E[\Delta\mathbf{h}^m])(\Delta\mathbf{h}^m - E[\Delta\mathbf{h}^m])^H] \quad (20)$$

$$\text{MSE} \triangleq E[||\Delta\mathbf{h}^m||^2] \quad (21)$$

We assume that the number of time samples T is large, so that $\frac{1}{T}\sum_{j=1}^{T}\mathbf{b}_j\mathbf{b}_j^H \approx \mathbf{I}_N$. Thus, the algorithm performance shall not depend on the specific transmitted sequence $\{\mathbf{b}_j\}$.

Theorem 1 *The proposed generalized estimator $\hat{\mathbf{h}}^m$ is asymptotically unbiased (i.e. Bias $= \mathbf{0}$) with the covariance*

$$Cov = \frac{\sigma^2}{T}[(\mathcal{C}^H\mathcal{U}_n\mathcal{A})^\dagger]^H\mathcal{A}^H\mathbf{\Upsilon}^{-2}\mathcal{A}(\mathcal{C}^H\mathcal{U}_n\mathcal{A})^\dagger \quad (22)$$

where $\mathbf{\Upsilon} \triangleq diag[\gamma^{m,1}, \ldots, \gamma^{m,P}] \otimes \mathbf{I}_{L-N}$, and mean square error $MSE = Tr[Cov]$.

Due to the space limitation, the proofs of the Theorems is omitted in this paper, The detailed proofs can be found in [6].

Theorem 1 indicates that the performance of the proposed estimator depends on the user specified parameters, i.e. the weight matrix $\mathcal{A}$ and the compounded kernel matrix $\mathcal{C}$ (14), which is determined by the set of kernel matrices utilized in the estimator, i.e. $S \triangleq \{\mathbf{C}^{m,1}, \ldots, \mathbf{C}^{m,P}\}$.

We next investigate the optimal choice of parameters $\mathcal{A}$ and S that minimizes the mean square error and the covariance of the estimator. To this end, it is convenient to explicitly indicate the functional dependence of these measures on $\mathcal{A}$ and S, i.e. MSE$(\mathcal{A}, S)$ and Cov$(\mathcal{A}, S)$. We begin with the minimization of MSE$(\mathcal{A}, S)$, which proceeds in two steps. Firstly we minimize this measure by adjusting the weight matrix $\mathcal{A}$ in the case of a fixed set S; secondly we search for an optimal S to minimize MSE$(\mathcal{A}, S)$ when the optimal weight matrix determined in the first step is used. Then the resulting choice on the parameters $\mathcal{A}$ and S minimizes MSE$(\mathcal{A}, S)$.

Theorem 2 *$c\mathbf{\Upsilon}$ is the optimal weight matrix minimizing MSE$(\mathcal{A}, S)$ for a fixed set S:*

$$MSE^o(S) \triangleq \min_{\mathcal{A}} MSE(\mathcal{A}, S) = MSE(c\mathbf{\Upsilon}, S) = \frac{\sigma^2}{T}Tr[\mathcal{Q}^\dagger] \quad (23)$$

where c is an arbitrary constant and $\mathcal{Q} \triangleq \mathcal{C}^H\mathcal{U}_n\mathbf{\Upsilon}^2\mathcal{U}_n^H\mathcal{C}$.

Theorem 2 shows that the optimal weights α^i $(i = 1, \ldots, P)$ are proportional to the corresponding received powers $(\gamma^{m,i})^2$.

Next we consider the set of kernel matrices S utilized in the estimator with optimal weight matrix.

Theorem 3 *For any proper subset S_q of S, we have*

$$MSE^o(S) < MSE^o(S_q) \quad (24)$$

The above theorem implies qualitatively that enlarging the set of kernel matrices S in the estimator will decrease its mean square error. Consequently the minimum mean square error is achieved when the estimator utilizes the kernel matrices of all the signal components in this group:

$$\text{MSE}^o \triangleq \min_{S\subseteq U}\text{MSE}^o(S) = \text{MSE}^o(U) = \frac{\sigma^2}{T}Tr[(\mathcal{Q}^m)^\dagger] \quad (25)$$

where $U \triangleq \{\mathbf{C}^{m,1}, \ldots, \mathbf{C}^{m,K^m}\}$ and $\mathcal{Q}^m$ is defined as $\mathcal{Q}$ with $P = K^m$

Theorem 4 *Given an arbitrary partition of S into Q non-empty subsets S_q $(q = 1, \ldots, Q)$, we have*

$$MSE^o(S) \le \frac{\sum_{q=1}^{Q}c_q^2 MSE^o(S_q)}{(\sum_{q=1}^{Q}c_q)^2} \quad (26)$$

where c_q is an arbitrary positive integer for $q = 1, \ldots, Q$.

As a special case of Theorem 4, assume that for $q = 1, \ldots, Q$, $c_q = 1$, and subset S_q only has one element, i.e. $S_q = \{\mathbf{C}^{m,q}\}$, and consequently $Q = P$. Then $\text{MSE}^o(S_q)$ represents the mean square error of the single signal estimator (e.g. [3]) applied on $\mathbf{C}^{m,q}$. Thus according to Theorem 4

$$\text{MSE}^o(S) \leq \frac{1}{P^2} \sum_{q=1}^{P} \text{MSE}^o(S_q) = \frac{\overline{\text{MSE}}}{P} \tag{27}$$

where $\overline{\text{MSE}} \triangleq \frac{1}{P} \sum_{q=1}^{P} \text{MSE}^o(S_q)$ denotes the average mean square error of single signal estimators over the set S.

Based on the above considerations, we suggest the following principles for minimizing $\mathbf{MSE}(\mathcal{A}, S)$:

1) *Choose the weights proportional to the received powers;*
2) *Include the maximum number of kernel matrices.*

We now turn our attention to the optimization of the covariance of the proposed estimator, as defined in (20). So far, we have not been able to extend the results of Theorems 2 to 4 to the covariance matrix so that they remain valid in the form of matrix inequalities. Fortunately, we can use the Cramer-Rao bound (CRB) to judge the optimality of the parameter choice previously obtained in the case of mean square error. That is: if the covariance matrix with the parameters $\mathcal{A} = c\mathbf{\Upsilon}$ and $S = U$ achieves the CRB, this parameter setting is considered the optimal one to minimize the covariance of the estimator.

We notice that some constraints are usually imposed on the estimated channel vector, e.g. unit norm. In this case, the traditional CRB is no longer applicable. The CRB for parameter estimation under constraints was recently given in [7], where a so-called *constrained* CRB is derived which depends on the specific algebraic constraints imposed on the estimated parameters. In [8], the concept of *minimal constrained* CRB is further introduced, which corresponds to the CRB matrix with the smallest trace (i.e. MSE) among the various constrained CRB matrices within the constraint class.

Theorem 5 *The minimal constrained CRB for the channel vector of interest is given by*

$$CRB_{C,\mathbf{h}^m} = \frac{\sigma^2}{T} (\mathcal{Q}^m)^\dagger \tag{28}$$

From the previously derived expression (22) for the covariance matrix of the target channel, we find that the proposed generalized subspace estimator $\hat{\mathbf{h}}^m$ achieves the minimal constrained CRB when $\mathcal{A} = c\mathbf{\Upsilon}$ and $S = U$. Therefore, we conclude that the choice of parameters $\mathcal{A} = c\mathbf{\Upsilon}$ and $S = U$ not only minimizes the mean square error, but also minimizes the covariance of the estimator.

V. Computer Experiments

Consider a down-link DS-CDMA connection from a base station to N remote users. The information bit to the i-th user b_i is spread by a unique spreading code $\mathbf{c}^i \triangleq [c_1^i, \ldots, c_{L_c}^i]^H$, where L_c is the processing gain. The frequency-selective channel is modelled as an FIR filter. The normalized coefficient vector of the filter is represented by $\mathbf{h}$ with size $M \times 1$. The kernel matrix of the i-th user $\mathbf{C}^i$ is an $(L_c - M + 1) \times M$ Toeplitz matrix with the first column $[c_M^i, \ldots, c_{L_c}^i]^T$ and the first row $[c_M^i, \ldots, c_1^i]$ [3]. Assuming the received amplitude of the i-th user is γ^i and the signal of all the users are synchronized, the received signal can be represented as

$$\mathbf{r} = (\sum_{i=1}^{N} \gamma^i \mathbf{C}^i b^i)\mathbf{h} + \mathbf{e} \tag{29}$$

The algorithm in Table I was specialized to this situation, resulting in a novel blind channel estimator that utilizes multiple signal components. Computer experiments were then conducted to verify the theoretical performance results derived in the last Section.

In the simulations, the following parameter values are used: number of active users $N = 4$, processing gain $L_c = 12$ and length of the channel vector $M = 4$. The binary spreading codes were randomly generated and stored for later use. We assume that some power control technique is applied so that the received amplitudes $[\gamma^1, \gamma^2, \gamma^3, \gamma^4]$ are proportional to $[1, 2, 3, 4]$, respectively. The following sets of kernel matrices were considered in the evaluation: $S^1 = \{\mathbf{C}^1\} \subset S^2 = \{\mathbf{C}^1, \mathbf{C}^2\} \subset S^3 = \{\mathbf{C}^1, \mathbf{C}^2, \mathbf{C}^3\} \subset S^4 = \{\mathbf{C}^1, \mathbf{C}^2, \mathbf{C}^3, \mathbf{C}^4\}$. We use the average value of the square error in 10^4 independent experiments to approximate the mean square error.

According to the analysis in Section IV, we consider the asymptotic MSE performance of the proposed estimator. From (23), (24) and (27), it respectively follows that

$$\text{MSE}^o(S^4) = \text{MSE}(\mathbf{\Upsilon}, S^4) \leq \text{MSE}(\mathbf{I_{4(L-N)}}, S^4) \tag{30}$$

$$\text{MSE}^o(S^4) < \text{MSE}^o(S^3) < \text{MSE}^o(S^2) < \text{MSE}^o(S^1) \tag{31}$$

$$\text{MSE}^o(S^4) \leq \frac{1}{4^2} \sum_{i=1}^{4} \text{MSE}^o(\{\mathbf{C}^i\}) \tag{32}$$

where $\mathbf{\Upsilon} \triangleq \text{diag}[\gamma^1, \ldots, \gamma^4] \otimes \mathbf{I_{4(L-N)}}$ and $L = L_c - M + 1$.

Both the theoretical and experimental results simulation results are presented in Fig. 1 to 4. Fig. 1 to 3, respectively, show the MSEs in (30), (31) and (32) plotted as a function of SNR, with a number of observed samples $T = 10^4$. Fig. 4 shows the MSEs in (31) plotted as a function of the number of observed samples T, with the SNR set to 10dB. Clearly, the theoretical performance properties in (30), (31) an (32) are verified in our simulations. Generally, we find that all the experimental results match the theoretical results well, especially in the case of high SNR and large T. The former is because our theoretical results are derived from a first-order perturbation analysis, which is accurate for small perturbations (i.e. high SNR); the latter is because in the asymptotic analysis we assume a large number of samples T. Our results thus support the performance analysis of the general model presented in Section IV.

VI. Conclusion

We presented a systematic study of the subspace-based blind channel estimation method. We first introduced a general signal model of multiple simultaneous signals transmitted through vector channels, which can be applied to a multitude

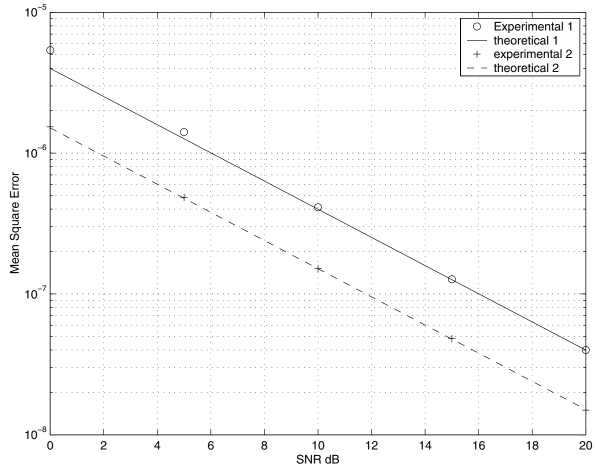

Fig. 1. Comparison in (30): 1: MSE($\mathbf{I_{4(L-N)}}, S^4$), 2: MSE$^o(S^4)$

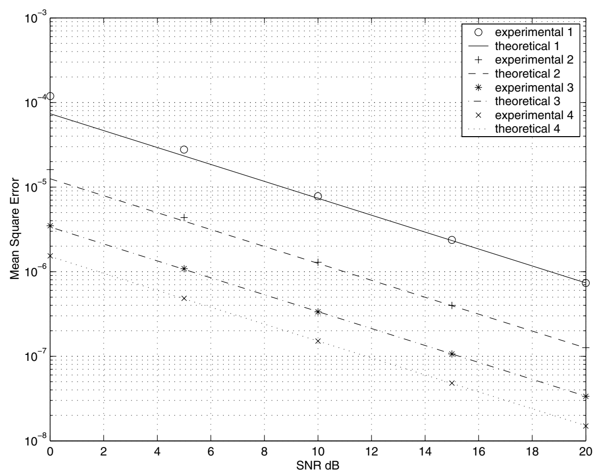

Fig. 2. Comparison in (31): 1: MSE$^o(S^1)$, 2: MSE$^o(S^2)$, 3: MSE$^o(S^3)$, 4: MSE$^o(S^4)$,

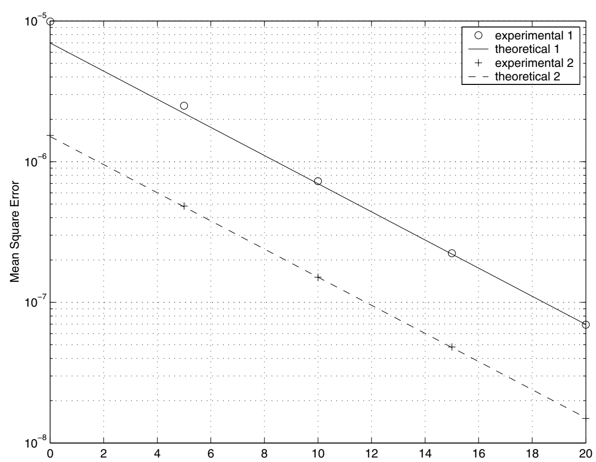

Fig. 3. Comparison in (32): 1: $\frac{1}{4^2}\sum_{i=1}^{4}$ MSE$^o(\{\mathbf{C}^i\})$, 2: MSE$^o(S^4)$

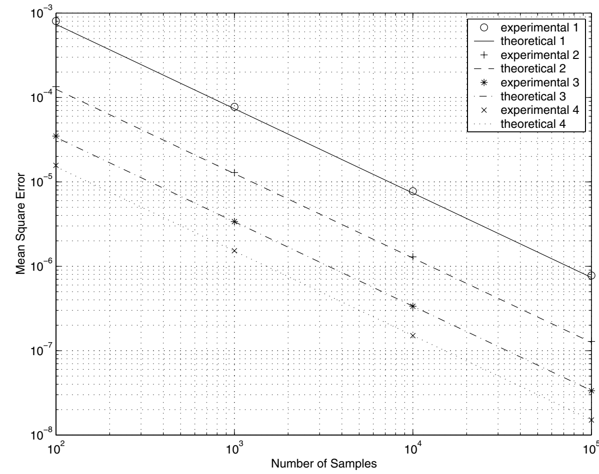

Fig. 4. Comparison in (31): 1: MSE$^o(S^1)$, 2: MSE$^o(S^2)$, 3: MSE$^o(S^3)$, 4: MSE$^o(S^4)$,

of modern digital communication systems. Based on this model, we formulated a generalized cost function for the purpose of subspace-based blind channel estimation, which incorporates the set of kernel matrices of the signals sharing the target channel via a weighted sum of projection errors. We investigated the asymptotic bias, covariance, MSE and CRB of the proposed estimator when the number of observations is large. We showed that the performance of the estimator can be optimized by using the maximum number of available kernel matrices and a special set of weights in the cost function. The results of the computer simulations fully support our analysis.

VII. Acknowledgement

Support for this work was provided in part by a grant from the Natural Sciences and Engineering Research Council of Canada and by a grant from McGill University, under Research & Development Fund.

References

[1] L. Tong and S. Perreau, "Multichannel blind identification: from subspace to maximum likelihood methods," *Proc. IEEE*, vol. 86, pp. 1951–1968, Oct. 1998.

[2] 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. 527–534, Feb. 1995.

[3] H. Liu and G. Xu, "A subspace method for signature waveform estimation in synchronous CDMA systems," *IEEE Trans. Commun.*, vol. 44, pp. 1346–1354, Oct. 1996.

[4] 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.

[5] D. Reynolds, X. Wang, and H. V. Poor, "Blind adaptive space-time multiuser detection with multiple transmitter and receiver antennas," *IEEE Trans. Signal Processing*, vol. 50, pp. 1261–1276, Jun. 2002.

[6] W. Kang, *Subspace-based Blind Channel Estimation: Generalization, Performance Analysis and Adaptive Implementation*. Master Thesis, McGill University, 2003.

[7] P. Stoica and B. C. Ng, "On the Cramer-Rao bound under parametric constraints," *IEEE Signal Processing Lett.*, vol. 5, pp. 177 –179, Jul. 1998.

[8] E. de Carvalho, J. Cioffi, and D. Slock, "Cramer-Rao bounds for blind multichannel estimation," in *Proc. GLOBECOM 2000*, vol. 2, 2000, pp. 1036 –1040.