Blind Channel Estimation via Subspace Approximation

Zhengyuan Xu
Dept. of Electrical Engineering
University of California
Riverside, CA 92521
dxu@ee.ucr.edu

***Abstract*— Subspace channel estimation technique has been well studied. It relies on subspace decomposition on the data covariance matrix to identify either signal subspace or noise subspace. However, the subspace decomposition process is computationally expensive. This paper applies a novel subspace approximation (SA) idea to bypass this process for channel estimation. It is based on a recently proposed "power of R" (POR) multiuser detection technique that raises power of estimated data covariance matrix to a positive integer to approximate rather than estimate the noise subspace component. Thus complexity is significantly reduced while satisfactory performance is still maintained. Channel estimation performance is studied in detail and compared with that of the well-known subspace method in the literature.** **Subspace channel estimation technique has been well studied. It relies on subspace decomposition on the data covariance matrix to identify either signal subspace or noise subspace. However, the subspace decomposition process is computationally expensive. This paper applies a novel subspace approximation (SA) idea to bypass this process for channel estimation. It is based on a recently proposed "power of R" (POR) multiuser detection technique that raises power of estimated data covariance matrix to a positive integer to approximate rather than estimate the noise subspace component. Thus complexity is significantly reduced while satisfactory performance is still maintained. Channel estimation performance is studied in detail and compared with that of the well-known subspace method in the literature.**

I. INTRODUCTION

A wireless communication channel usually introduces intersymbol interference (ISI) due to multipath propagation. At the receiver, channel equalization is necessary to mitigate ISI for successful symbol detection [1]. Due to unknown channel characteristics, design of equalizers can be facilitated by channel estimation. Blind methods rely solely on the channel outputs instead of requiring system to transmit training sequences. Among all blind methods, second-order statistics (SOS) based methods can yield fast convergence and acceptable performance [2]. They utilize auto-covariance of received data samples. Subspace methods are very efficient and widely studied. They apply subspace decomposition to the data covariance matrix to obtain either signal subspace or noise subspace first [3], [4]. Then, based on an observation that signature waveforms of transmitted symbols are orthogonal to the noise subspace, channel parameters can be identified under some conditions. The idea has been applied to channel estimation in a code-division multiple access (CDMA) multiuser system [5], [6] in order to build different multiuser receivers.

In applying the subspace technique, subspace decomposition such as eigenvalue decomposition (EVD) is computationally demanding due to possibly large dimension of the objective matrix. Therefore, low-complexity algorithms based on the SOS are desirable. Certainly, subspace tracking technique can be applied to adaptively obtain subspaces [7]. In either batch processing or adaptive implementation, channel order needs to be known or estimated. Performance of typical rank detection methods depends on the signal to noise ratio (SNR) and sample size in a practical condition [8]. It relies heavily on the distribution of small eigenvalues of estimated data covariance matrix. Some methods robust to order overestimation have been recently proposed such as linear prediction method [9], matrix outer product decomposition method [10], or a truly robust method based on shifted versions of the covariance matrix and the properties of the associated kernel [11]. However, their performance is inferior to the subspace method when the true channel order is known [11].

Recently, a power of R (POR) method closely related to the subspace method has been proposed for multiuser detection [12], [13]. It raises power of the inverse of the data covariance matrix $\boldsymbol{R}$ ($\boldsymbol{R}^{-1}$) to a positive integer m to approximate the noise subspace component rather than estimate it from subspace decomposition. Therefore, complexity can be significantly reduced. In this paper, we apply the subspace approximation (SA) idea to estimate multiple finite impulse response (FIR) channels. Though inaccurate, the approximation error decreases dramatically with increased exponent m irrespective of channel noise. Asymptotic equivalence of the proposed method to the subspace technique is demonstrated. Simulation results show that small m ($m = 2$ or 3) is sufficient for the proposed method to provide satisfactory channel estimate in practical conditions. Convergence to the subspace method is also experimentally observed while in particular, low complexity of the proposed algorithm can be easily achieved by simple and straightforward adaptive implementation. Needless to mention, our method requires channel order information, similar to the subspace method.

II. System Model

In digital communication, if duration of transmitted symbol $w(n)$ is T, then received baseband signal follows a model [4]

$$y(t) = \sum_{n=-\infty}^{\infty} w(n)h(t-nT) + v(t) \quad (1)$$

where $h(t)$ is the overall channel impulse response including the effects of transmitter filter, propagation channel and receiver filter, $v(t)$ is a white Gaussian stationary process.

For diversity reception and channel estimation purposes, we assume P sub-channels are available. They can be obtained from either oversampling a single sensor at a rate of multiples of symbol rate, or employing multiple sensors sampling at the symbol rate at the receiver [4]. In either case, all subchannels are assumed to have finite duration support with maximum order q. The discrete-time output of the pth subchannel can be written by

$$y_p(n) = \sum_{l=0}^{q} w(n-l)h_p(l) + v_p(n). \quad (2)$$

If we collect $y_p(n)$ for $p = 1, \cdots, P$ in a vector and then stack L such vectors corresponding to L current/past successive symbol intervals in a vector $\boldsymbol{y}_n$ of length $\nu = LP$, then we obtain a vector/matrix representation [4]

$$\boldsymbol{y}_n = \boldsymbol{H}\boldsymbol{w}_n + \boldsymbol{v}_n \quad (3)$$

where $\boldsymbol{H}$ is a block Toeplitz channel matrix of size $\nu \times (L+q)$, $\boldsymbol{w}_n = [w(n), \cdots, w(n-L-q+1)]^T$ is an input vector, $\boldsymbol{v}_n$ is a noise vector. Assume input and noise have zero mean and variance σ_w^2 and σ_v^2 respectively.

III. Review of Subspace Channel Estimation

All channel parameters in $\boldsymbol{H}$ can be estimated by the subspace technique. For convenience, collect all $P(q+1)$ unknown parameters in a vector $\boldsymbol{g}_d$

$$\boldsymbol{g}_d = [h_1(q), \cdots, h_P(q), \cdots, h_1(0), \cdots, h_P(0)]^T.$$

Also denote the $(j+q+1)$th column ($j = -q, \cdots, 0, \cdots, L-1$) of $\boldsymbol{H}$ by $\boldsymbol{h}_j$. Then $\boldsymbol{h}_0 = [\boldsymbol{g}_d^T, 0, \cdots, 0]^T$. If we define a $\nu \times P(q+1)$ matrix $\boldsymbol{S} = [\boldsymbol{I}_{P(q+1)}, \boldsymbol{0}]^T$ where $\boldsymbol{I}$ is an identity matrix, then $\boldsymbol{h}_0 = \boldsymbol{S}\boldsymbol{g}_d$. Other columns $\boldsymbol{h}_j$ also contain partial/full copies of vector $\boldsymbol{g}_d$ together with some possible leading/trailing zeros. To conveniently explore this relationship, we introduce a shift matrix $\boldsymbol{J}$ with all $1's$ in the first sub-diagonal. We also use the symbol $\boldsymbol{J}^{-1}$ to denote $\boldsymbol{J}^T$ although $\boldsymbol{J}$ is singular: $\boldsymbol{J}^{-1} \triangleq \boldsymbol{J}^T$ and define $\boldsymbol{J}^0$ as an identity matrix. Then $\boldsymbol{h}_j$ can be obtained by shifting all elements of $\boldsymbol{h}_0$ up or down by $|jP|$ positions

$$\boldsymbol{h}_j = \boldsymbol{J}^{jP}\boldsymbol{h}_0 = \boldsymbol{J}^{jP}\boldsymbol{S}\boldsymbol{g}_d. \quad (4)$$

If the data covariance matrix $\boldsymbol{R} = E\{\boldsymbol{y}_n\boldsymbol{y}_n^H\}$ is decomposed by EVD as

$$\boldsymbol{R} = \begin{bmatrix} \boldsymbol{U}_s & \boldsymbol{U}_n \end{bmatrix} \begin{bmatrix} \boldsymbol{\Lambda}_s + \sigma_v^2\boldsymbol{I} & \boldsymbol{0} \\ \boldsymbol{0} & \sigma_v^2\boldsymbol{I} \end{bmatrix} \begin{bmatrix} \boldsymbol{U}_s^H \\ \boldsymbol{U}_n^H \end{bmatrix} \quad (5)$$

where $\boldsymbol{\Lambda}_s = diag\{\lambda_1^2, \ \ldots \ , \lambda_{L+q}^2\}$, $\boldsymbol{U}_s$ and $\boldsymbol{U}_n$ represent the signal and noise subspaces respectively, then $\boldsymbol{U}_n^H\boldsymbol{h}_j = \boldsymbol{0}$ for all possible j. It is easily found that the rank of $\boldsymbol{U}_s$ is equal to the rank of $\boldsymbol{H}$ which is $L+q$ under some assumptions on all subchannels [4]. Denote it by $\xi_s = L+q$. Then the rank of $\boldsymbol{U}_n$ becomes $\xi_n = \nu - L - q$. Therefore considering (4), the subspace channel estimation criterion can be described as follows [4]

$$\widehat{\boldsymbol{g}} = \arg\min_{||\boldsymbol{x}||=1} \sum_{j=-q}^{L-1} \boldsymbol{x}^H\boldsymbol{S}^H\boldsymbol{J}^{-jP}\boldsymbol{U}_n\boldsymbol{U}_n^H\boldsymbol{J}^{jP}\boldsymbol{S}\boldsymbol{x}. \quad (6)$$

Under some identifiability conditions, (6) gives a unique channel vector up to a multiplicative scalar. In the subspace technique, $\boldsymbol{U}_n$ can be estimated from EVD of data covariance matrix estimated from N data vectors

$$\widehat{\boldsymbol{R}} = \frac{1}{N}\sum_{n=1}^{N}\boldsymbol{y}_n\boldsymbol{y}_n^H. \quad (7)$$

However, we will propose a subspace approximation (SA) technique to approximate the noise subspace component $\boldsymbol{U}_n\boldsymbol{U}_n^H$ in (6) from which $\boldsymbol{g}_d$ can be estimated.

IV. Proposed SA Channel Estimation

If (6) is examined once more, it is found that only $\boldsymbol{U}_n\boldsymbol{U}_n^H$ instead of $\boldsymbol{U}_n$ itself is required. This quantity is clearly related to $\boldsymbol{R}^{-1}$ according to (5)

$$\sigma_v^2\boldsymbol{R}^{-1} = \boldsymbol{U}_n\boldsymbol{U}_n^H + \boldsymbol{U}_s diag\{(\frac{\sigma_v^2}{\lambda_i^2 + \sigma_v^2})\}\boldsymbol{U}_s^H. \quad (8)$$

If we raise the power of $\sigma_v^2\boldsymbol{R}^{-1}$ to a positive integer m ($m = 1, 2, 3, \cdots$)

$$\sigma_v^{2m}\boldsymbol{R}^{-m} = \boldsymbol{U}_n\boldsymbol{U}_n^H + \boldsymbol{U}_s diag\{(\frac{\sigma_v^2}{\lambda_i^2 + \sigma_v^2})^m\}\boldsymbol{U}_s^H, \quad (9)$$

then the second term on the right hand side of (8) will reduce weight because $\frac{\sigma_v^2}{\lambda_i^2+\sigma_v^2} < 1$ for arbitrary noise power σ_v^2. Theoretically,

$$\lim_{m\to\infty}\sigma_v^{2m}\boldsymbol{R}^{-m} = \boldsymbol{U}_n\boldsymbol{U}_n^H \quad (10)$$

irrespective of the noise power. Therefore, estimating $\boldsymbol{U}_n\boldsymbol{U}_n^H$ from $\boldsymbol{R}^{-m}$ has an obvious advantage. It eliminates a need to estimate the subspace rank and thus mitigates the effects of noise and sample size by certain degree.

To conclude, $\boldsymbol{U}_n\boldsymbol{U}_n^H$ can be approximated by $\boldsymbol{R}^{-m}$ after ignoring a scalar σ_v^{2m} which is unimportant to channel estimation. Therefore, considering (6), our SA criterion becomes

$$\boldsymbol{g}_{sa} = \arg\min_{||\boldsymbol{x}||=1}\boldsymbol{x}^H\boldsymbol{A}\boldsymbol{x}, \quad \boldsymbol{A} = \sum_{j=-q}^{L-1}\boldsymbol{S}^H\boldsymbol{J}^{-jP}\boldsymbol{R}^{-m}\boldsymbol{J}^{jP}\boldsymbol{S}. \quad (11)$$

Channel estimate $\boldsymbol{g}_{sa}$ is the eigenvector of $\boldsymbol{A}$ corresponding to its minimum eigenvalue γ_{sa}. $\boldsymbol{R}^{-m}$ can be obtained from direct inversion on $\boldsymbol{R}^m$, or by recursive least-squares (RLS) estimation of $\boldsymbol{R}^{-1}$ from data vectors and raising its power to m. Parameter m can be properly chosen. Next, we will analyze effects of noise, parameter m and sample size N on the proposed SA estimator.

V. Performance of SA Channel Estimator

It is clear that the SA channel estimator is affected by additive noise (σ_v^2). It also depends on parameter m. In addition, since estimate $\boldsymbol{g}_{sa}$ depends on $\boldsymbol{R}$ which can only be estimated from finite received data samples, effect of sample size N will also be studied in this section. However, the analysis will be much involved with arbitrary σ_v^2 and N. For analytical convenience, we assume that σ_v^2 is small and N is large, leading to small perturbation to $\boldsymbol{R}$. Then the perturbation technique is readily applicable. A brief discussion will be carried out when σ_v^2 can not be assumed small.

A. Effect of Noise

We begin with examination of our channel estimator obtained from (11). $\boldsymbol{g}_{sa}$ is the eigenvector of $\boldsymbol{A}$ corresponding to its minimum eigenvalue. Under a small noise assumption, a perturbation by noise in $\boldsymbol{R}$ causes a perturbation in $\boldsymbol{A}$. Thus a perturbation in $\boldsymbol{g}_{sa}$ occurs. For convenience, assume $\|\boldsymbol{g}_d\| = 1$. The following two lemmas show dependence of the objective matrix in our criterion and its eigen-pairs to the noise. They can be easily proven using Taylor expansion.

Lemma 1: Under a small noise assumption, $\boldsymbol{A}$ can be expressed as a power series of σ_v^2 as follows

$$\sigma_v^{2m}\boldsymbol{A} \approx \boldsymbol{A}_0 + \sigma_v^{2m}\boldsymbol{A}_m - \sigma_v^{2m+2} m\boldsymbol{A}_{m+1} \tag{12}$$

where

$$\boldsymbol{A}_l = \sum_{j=-q}^{L-1} \boldsymbol{S}^H\boldsymbol{J}^{-jP}\boldsymbol{U}_s\boldsymbol{\Lambda}_s^{-l}\boldsymbol{U}_s^H\boldsymbol{J}^{jP}\boldsymbol{S}, \quad \boldsymbol{\Lambda}_s^0 = \boldsymbol{I}$$

for $l = 0, 1, \cdots, m, m+1, \cdots$. □

Lemma 2: Under channel identifiability conditions [4], the eigen-pair $(\gamma, \boldsymbol{v})$ of matrix $\sigma_v^{2m}\boldsymbol{A}$ has different power series expansions for different m. If $m = 1$, it is given by

$$\gamma_{sa} \approx \sigma_v^2\boldsymbol{g}_d^H\boldsymbol{A}_1\boldsymbol{g}_d - \sigma_v^4(\boldsymbol{g}_d^H\boldsymbol{A}_1\boldsymbol{A}_0^\dagger\boldsymbol{A}_1\boldsymbol{g}_d + \boldsymbol{g}_d^H\boldsymbol{A}_2\boldsymbol{g}_d) \tag{13}$$

$$\begin{aligned}\boldsymbol{g}_{sa} &\approx \boldsymbol{g}_d - \sigma_v^2\boldsymbol{A}_0^\dagger\boldsymbol{A}_1\boldsymbol{g}_d + \sigma_v^4\boldsymbol{A}_0^\dagger\boldsymbol{A}_1\boldsymbol{A}_0^\dagger\boldsymbol{A}_1\boldsymbol{g}_d \\ &+ \sigma_v^4(\boldsymbol{A}_0^\dagger\boldsymbol{A}_2\boldsymbol{g}_d - \boldsymbol{g}_d^H\boldsymbol{A}_1\boldsymbol{g}_d\boldsymbol{A}_0^{\dagger 2}\boldsymbol{A}_1\boldsymbol{g}_d)\end{aligned} \tag{14}$$

where † represents pseudo-inverse. If $m > 1$, it has a form

$$\gamma_{sa} \approx \sigma_v^{2m}\boldsymbol{g}_d^H\boldsymbol{A}_m\boldsymbol{g}_d - \sigma_v^{2m+2} m\boldsymbol{g}_d^H\boldsymbol{A}_{m+1}\boldsymbol{g}_d, \tag{15}$$

$$\boldsymbol{g}_{sa} \approx \boldsymbol{g}_d - \sigma_v^{2m}\boldsymbol{A}_0^\dagger\boldsymbol{A}_m\boldsymbol{g}_d + \sigma_v^{2m+2} m\boldsymbol{A}_0^\dagger\boldsymbol{A}_{m+1}\boldsymbol{g}_d. \tag{16}$$

□

In these lemmas, terms with order higher than σ_v^{2m+2} are omitted. Although the eigen-pair $(\gamma, \boldsymbol{v})$ of $\sigma_v^{2m}\boldsymbol{A}$ has different expansions for different m, it is observed that corresponding terms with lowest order (except order zero) of σ_v still obey the same form. Difference arises only in those terms with higher order. Based on *Lemma 2*, the channel estimation error can be easily obtained. For convenience, denote the error due to the effect of noise by preceding the corresponding quantity by Δ.

Theorem 1: For small σ_v^2, the channel estimation error is given by

$$\Delta\boldsymbol{g} = \boldsymbol{g}_{sa} - \boldsymbol{g}_d = -\sigma_v^{2m}\boldsymbol{A}_0^\dagger\boldsymbol{A}_m\boldsymbol{g}_d + \boldsymbol{o}(\sigma_v^{2m+2}), \tag{17}$$

and the convergence can be established

$$\boldsymbol{g}_{sa} \to \boldsymbol{g}_d \quad as \quad \sigma_v^2 \to 0. \tag{18}$$

□

The previous results are based on a small noise assumption, i.e. $\sigma_v^2 << \lambda_i^2$ for $i = 1, \ldots, \xi$, which may be violated for a system with large eigenvalue spread. In that situation, (17) does not predict the channel estimation error satisfactorily. A more accurate expression for $\Delta\boldsymbol{g}$ is thus desirable. According to (9), we obtain

$$\sigma_v^{2m}\boldsymbol{A} = \boldsymbol{A}_0 + \Delta\boldsymbol{A}_0, \tag{19}$$

where $\Delta\boldsymbol{A}_0 = \sum_j \boldsymbol{S}^H\boldsymbol{J}^{-jP}\boldsymbol{U}_s diag\{(\frac{\sigma_v^2}{\lambda_i^2+\sigma_v^2})^m\}\boldsymbol{U}_s^H\boldsymbol{J}^{jP}\boldsymbol{S}$. For sufficiently large m, $\Delta\boldsymbol{A}_0$ can be viewed as a perturbation. Notice that $\boldsymbol{g}_d$ is the minimum eigenvector of $\boldsymbol{A}_0$. When channel is estimated from the minimum eigenvector of $\boldsymbol{A}$, its first-order perturbation is then given by [14]

$$\Delta\boldsymbol{g} \approx -\boldsymbol{A}_0^\dagger\Delta\boldsymbol{A}_0\boldsymbol{g}_d. \tag{20}$$

According to (20) and $\Delta\boldsymbol{A}_0$, the channel estimation error can still be very small for large m even in the case of $\sigma_v^2 > \lambda_i^2$ for some i. On the other hand, for small σ_v^2, (20) can be easily shown to reduce to (17) after further invoking Taylor expansion.

B. Effect of Parameter m

The SA channel estimation method depends on m. After noticing (19) and $\boldsymbol{g}_d$ is the minimum eigenvector of $\boldsymbol{A}_0$, a theorem similar to *Theorem 1* can be proved.

Theorem 2: (18) holds for any fixed σ_v^2 as $m \to \infty$. □

This theorem shows that the proposed method converges to the subspace one for sufficiently large m.

C. Effect of Sample Size

As observed from (11), the estimated channel vector $\boldsymbol{g}_{sa}$ depends on $\boldsymbol{R}$ which is typically estimated from finite data samples by (7). Perturbation arises in the estimated data covariance matrix when it is estimated from N data vectors [15]. We will denote a perturbation due to this effect by preceding the corresponding quantity by δ, and perturbed quantity with $\widehat{\ }$

$$\delta\boldsymbol{g} = \widehat{\boldsymbol{g}}_{sa} - \boldsymbol{g}_{sa}, \quad \delta\boldsymbol{R} = \widehat{\boldsymbol{R}} - \boldsymbol{R}.$$

Although $\widehat{\boldsymbol{R}}$ converges to $\boldsymbol{R}$ as $N \to \infty$, a perturbation $\delta\boldsymbol{R}$ due to finite N will cause $\boldsymbol{A}$ perturbed and finally our SA solution. Here, we investigate how $\delta\boldsymbol{R}$ affects the performance of the SA channel estimator.

According to (11), $\boldsymbol{A}$ depends on $\boldsymbol{R}^{-1}$. Under small perturbation assumption (large N) and using Taylor's expansion up to the first order, we have

$$(\boldsymbol{R} + \delta\boldsymbol{R})^{-m} \approx \boldsymbol{R}^{-m} - \sum_{k=1}^{m} \boldsymbol{R}^{-k}\delta\boldsymbol{R}\boldsymbol{R}^{-(m+1-k)} \tag{21}$$

Then the first order perturbation of $\boldsymbol{A}$ is

$$\delta\boldsymbol{A} = -\sum_{j=-q}^{L-1}\sum_{k=1}^{m}\boldsymbol{S}^H\boldsymbol{J}^{-jP}\boldsymbol{R}^{-k}\delta\boldsymbol{R}\boldsymbol{R}^{-(m+1-k)}\boldsymbol{J}^{jP}\boldsymbol{S}. \quad (22)$$

Due to $\delta\boldsymbol{A}$, $\boldsymbol{g}_{sa}$ corresponding to eigenvalue γ_{sa} is perturbed with perturbation $\delta\boldsymbol{g}$. It can be found that [14], [15]

$$\delta\boldsymbol{g} \approx -(\boldsymbol{A}-\gamma_{sa}\boldsymbol{I})^{\dagger}\delta\boldsymbol{A}\boldsymbol{g}_{sa}. \quad (23)$$

After substituting (22) in (23), we obtain

$$\delta\boldsymbol{g} \approx \sum_{j=-q}^{L-1}\sum_{k=1}^{m}\boldsymbol{T}_{j,k}\delta\boldsymbol{R}\boldsymbol{t}_{j,k} \quad (24)$$

where $\boldsymbol{T}_{j,k}$ and $\boldsymbol{t}_{j,k}$ are deterministic quantities

$$\boldsymbol{T}_{j,k} = (\boldsymbol{A}-\gamma_{sa}\boldsymbol{I})^{\dagger}\boldsymbol{S}^H\boldsymbol{J}^{-jP}\boldsymbol{R}^{-k},$$

$$\boldsymbol{t}_{j,k} = \boldsymbol{R}^{-(m+1-k)}\boldsymbol{J}^{jP}\boldsymbol{S}\boldsymbol{g}_{sa}.$$

It can be proved that due to concurrence of $\boldsymbol{T}_{j,k}$ and $\boldsymbol{t}_{j,k}$, the channel estimation error will not be significantly amplified by noise although $\boldsymbol{R}^{-k}$ for $k=1,\ldots,m$ are involved. The covariance of $\delta\boldsymbol{g}$ can be found from (24)

$$\mathrm{Cov}_g \approx \sum_{j_1,j_2,k_1,k_2}\boldsymbol{T}_{j_1,k_1}E\{\delta\boldsymbol{R}\boldsymbol{t}_{j_1,k_1}\boldsymbol{t}_{j_2,k_2}^H\delta\boldsymbol{R}\}\boldsymbol{T}_{j_2,k_2}^H, \quad (25)$$

and the mean-square-error is equal to the trace of Cov_g. Both are dependent on the weighted covariance of $\delta\boldsymbol{R}$ in a form $\boldsymbol{\Psi} = E\{\delta\boldsymbol{R}\boldsymbol{D}\delta\boldsymbol{R}\}$ where $\boldsymbol{D}$ can be replaced by quantities $\boldsymbol{t}_{j_1,k_1}\boldsymbol{t}_{j_2,k_2}^H$. Its closed form has been derived in [16] if covariance matrix $\boldsymbol{R}$ is estimated from N independent data vectors by (7) for given data model (3).

Considering both noise and sample size, we further obtain the channel estimation error.

Theorem 3: For small σ_v^2 and large N, channel estimation mean-square-error (MSE) is given by $MSE \approx ||\Delta\boldsymbol{g}||^2 + tr(\mathrm{Cov}_g)$ where the first term is obtained from either (17) or (20) and the second term by (25). □

So far, we have studied channel estimation performance due to noise, design parameter m and sample size N. The proposed method requires $\boldsymbol{R}^{-m}$. If inversion of $\boldsymbol{R}$ is directly invoked, then it is computationally complex. We consider update of $\boldsymbol{R}^{-m}$ next to reduce complexity.

VI. Update of $\boldsymbol{R}^{-m}$

Update of matrix $\boldsymbol{R}^{-m}$ can be derived by applying the matrix inversion lemma. If the covariance matrix is updated from its previous value and received data vector as follows [17]

$$\boldsymbol{R}_n = \beta\boldsymbol{R}_{n-1} + \boldsymbol{y}_n\boldsymbol{y}_n^H \quad (26)$$

where β is a forgetting factor, then $\boldsymbol{R}_n^{-1}$ can be updated by

$$\boldsymbol{R}_n^{-1} = \beta^{-1}\boldsymbol{R}_{n-1}^{-1} - \frac{\beta^{-2}\boldsymbol{R}_{n-1}^{-1}\boldsymbol{y}_n\boldsymbol{y}_n^H\boldsymbol{R}_{n-1}^{-1}}{1+\beta^{-1}\boldsymbol{y}_n^H\boldsymbol{R}_{n-1}^{-1}\boldsymbol{y}_n} \quad (27)$$

from which $\boldsymbol{R}_n^{-m}$ can be found to be determined by $\boldsymbol{R}_{n-1}^{-l}$ for $l=m,m-1,\cdots,1$. Therefore, update of $\boldsymbol{R}_n^{-m}$ requires update of $\boldsymbol{R}_n^{-l}$ for $l=1,\cdots,m-1$. Starting from $\boldsymbol{R}_n^{-1}$ as in (27), we can successively update $\boldsymbol{R}_n^{-2}$, $\cdots$, up to $\boldsymbol{R}_n^{-m}$ based on either a power operation or multiplication by $\boldsymbol{R}_n^{-k}$ (with appropriate k) on both sides of (27). With an initial value of $\boldsymbol{R}_0^{-1}$, all other initial values $\boldsymbol{R}_0^{-l}$ follow. Each time, matrix $\boldsymbol{A}$ can then be constructed and the desired eigenvector can be obtained from its EVD. For a moderate channel order, this process is not demanding. However, one may still consider subspace tracking to reduce computations at this step but is beyond the scope of the paper.

VII. Simulation Results

We test channel estimation performance based on the channel parameters used in [4] which form a set of four subchannels with five coefficients each. Transmitted signals are taken from 4-QAM constellations. L is set to be 10. A normalized mean-square-error (MSE) for $\boldsymbol{g}_d$ is adopted as a performance measure.

The effect of design parameter m is studied for different SNRs with results presented in Fig. 1 when $N=\infty$. m varies from 1 to 6 corresponding to lines from top to bottom. It is found that there is significant improvement for small m, for example from $m=1$ to $m=2$, or $m=2$ to $m=3$, all monotonically with increased m. After certain point, slight performance improvement results. For a finite number of symbol intervals however, monotonic improvement with increased m does not always hold due to additional error from effect of sample size, as predicted by *Theorem 3* and shown in Fig. 2 with $N=1000$ and averaged from 100 independent realizations. There are totally 10 solid lines corresponding to $m=1$ to $m=10$ and a dashed line for the subspace method. It is found that all solid lines for m larger than 3 are close to each other except for small m. Three upper lines in the region of $10\sim15dB$ correspond to $m=1,2,3$. $m=1$ gives the worst performance while $m=3$ gives MSEs very close to the subspace method. It can thus be concluded that the proposed method with small m can provide performance as sufficiently good as the subspace method.

To verify *Theorem 3*, we use up to 10^5 independent data vectors (no overlapping in time) and consider $m=3$ and $SNR=20dB$ only. The solid line in Fig. 3 is obtained from simulation, and dashed line from analysis. Convergence of the solid line to the dashed line is observed for large N. We also test the proposed adaptive algorithm using $\beta=0.995$, $m=2$, and $SNR=20dB$. Data covariance matrix is initialized as $0.01\boldsymbol{I}$. One iteration corresponds to one data vector successively collected. It can be observed from Fig. 4 that after 1000 iterations, the MSE reaches about 10^{-2}.

References

[1] J. K. Tugnait, L. Tong and Z. Ding, "Single-user channel estimation and equalization," *IEEE Signal Processing Magazine*, vol. 17, no. 3, May 2000.

[2] L. Tong and S. Perreau, "Blind channel estimation: from subspace to maximum likelihood methods", *IEEE Proceedings*, vol. 86, no. 10, pp 1951-1968, 1998.

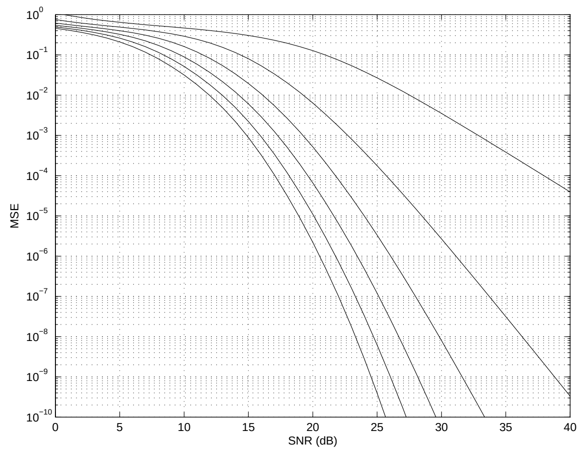

Fig. 1. Effect of m ($N = \infty$, different SNR).

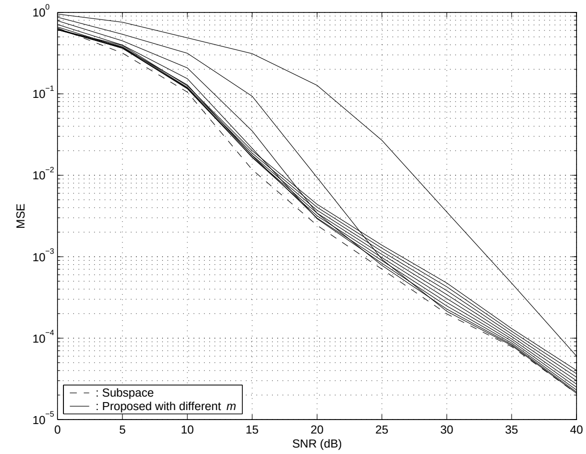

Fig. 2. Effect of m ($N = 1000$, different SNR).

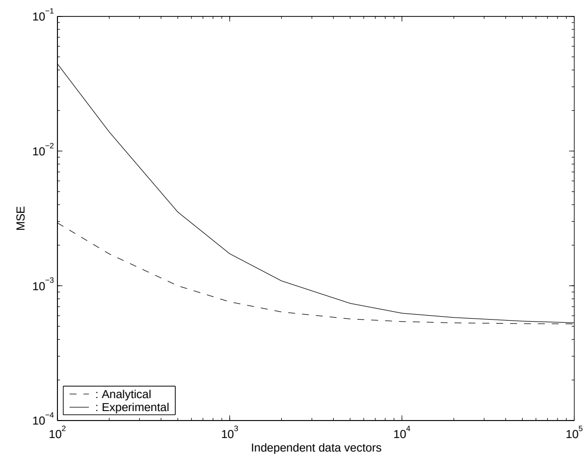

Fig. 3. Effect of sample size (independent data vectors, $m = 3$, $SNR = 20dB$).

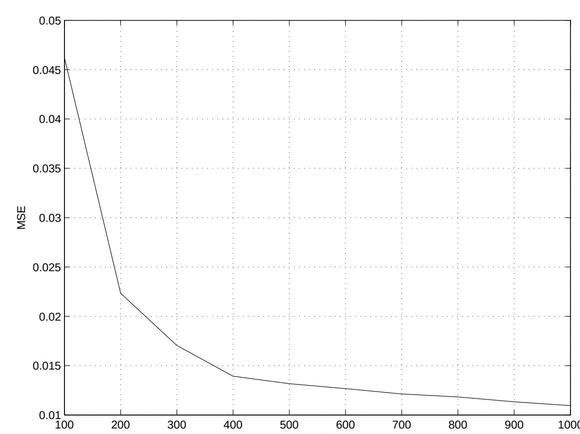

Fig. 4. Channel MSE from adaptive implementation ($m = 2$, $SNR = 20dB$).

[3] L. Tong, G. Xu and T. Kailath, “Blind channel identification and equalization using second-order statistics: a time-domain approach,” *IEEE Trans. Information Theory*, vol. 40, no. 2, pp. 340-349, March 1994.

[4] 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, no. 2, pp. 516-525, February 1995.

[5] X. Wang and H. V. Poor, “Blind multiuser detection: a subspace approach,” *IEEE Trans. Information Theory*, vol. 44, no. 2, pp. 677-690, March 1998.

[6] X. Wang and H. Poor, “Blind equalization and multiuser detection in dispersive CDMA channels,” *IEEE Trans. Commun.*, vol. 46, no. 1, pp. 91-103, January 1998.

[7] B. Yang, “Projection approximation subspace tracking,” *IEEE Trans. Signal Processing*, vol. 44, no. 1, pp. 95-107, Jan. 1995.

[8] A. Liavas and P. Regalia, “On the behavior of information theoretic criteria for model order selection,” *IEEE Trans. Signal Processing*, vol. 49, no. 8, pp. 1689-95, Aug. 2001.

[9] K. Abed-Meraim, E. Moulines and P. Loubaton, “Prediction error method for second-order blind identification,” *IEEE Trans. Signal Processing*, vol. 45, no. 3, pp. 694-705, March 1997.

[10] Z. Ding, “Matrix outer product decomposition method for blind multiple channel identification,” *IEEE Trans. Signal Processing*, vol. 45, no. 12, pp. 3053-3061, Dec. 1997.

[11] H. Gazzah, P. A. Regalia, J-P Delmas and K. Abed-Meraim, “A blind multichannel identification algorithm robust to order overestimation,” *IEEE Trans. Signal Processing*, vol. 50, no. 6, pp. 1449-1458, June 2002.

[12] Z. Xu, P. Liu and X. Wang, “Towards closing the gap between MOE and subspace methods,” *Proc. Asilomar Conf. Signals, Systems, and Computers*, pp. 689-693, Nov. 2002.

[13] Z. Xu, P. Liu and X. Wang, “Blind multiuser detection: from MOE to subspace methods,” *IEEE Trans. Signal Processing*, vol. 52, no. 2, February 2004.

[14] Z. Xu, “Perturbation analysis for subspace decomposition with applications in subspace-based algorithms,” *IEEE Trans. Signal Processing*, vol. 50, no. 11, pp. 2820-2830, November 2002.

[15] B. Friedlander and A. Weiss, “On the second-order statistics of the eigenvectors of sample covariance matrices,” *IEEE Trans. Signal Processing*, vol. 46, no. 11, pp. 3136-3139, Nov. 1998.

[16] Z. Xu, “On the second-order statistics of the weighted sample covariance matrix,” *IEEE Trans. Signal Processing*, vol. 51, no. 2, pp. 527-534, February 2003.

[17] S. Haykin, *Adaptive Filter Theory*, 3rd edition, Prentice Hall, Upper Saddle River, NJ, 1996.