Reduced-Rank Space-time Channel Estimation for DS-CDMA in MIMO Dispersive Channels

Fu-Hsuan Chiu, Sau-Hsuan Wu and C.-C. Jay Kuo
Department of Electrical Engineering and Integrated Media Systems Center
University of Southern California, Los Angeles, CA 90089-2564, USA
E-mails: fchiu@usc.edu, sauhsuaw@usc.edu, cckuo@sipi.usc.edu

***Abstract*— A reduced-rank space-time channel estimation scheme is presented for uplink multiuser DS-CDMA (direct-sequence code-division multiple access) systems in dispersive MIMO (multiple input and multiple output) channels. The proposed method is based on the alternative power (AP) method with the aid of pilot symbols. It is shown that the mean squared error (MSE) of channel estimation can be improved via rank order selection where the optimal performance can be achieved using the generalized likelihood ratio test (GLRT). Both analytical and simulated results are provided to demonstrate the performance of the proposed scheme with various measure criteria.**

I. Introduction

A dispersive channel has an impulse response with large delay spread and the energy is concentrated in several small regions within the overall delay spread. To mitigate the severe inter-symbol interference (ISI) in such a channel, the decision feedback equalizer (DFE) is often employed [1]. Even though the complexity of a DFE is lower than the trellis-based Viterbi algorithm, the performance is however degraded. Furthermore, the linear feedforward and feedback filter for a conventional DFE can still be long. In this case, the reduced-rank channel estimation method can lead to a reduced-state equalizer/detector [1] to obtain suboptimum performance with much lower complexity.

In view of the advantages of reduced-rank estimation for dispersive channels, the reduced-rank maximum likelihood (ML) estimation proposed in [2] for a linear regression model has been applied to multiuser CDMA case in [3]. However, the complexity for the singular-value-decomposition (SVD) and the square-root-inverse involved in this approach is high. On the other hand, the AP method developed in [4] is computationally more efficient. Based on the AP method, we have investigated herein the performance of reduced-rank estimation for MIMO DS-CDMA channels and evaluated analytically the Cramer-Rao bound for various reduced-rank schemes.

In MIMO channels, once the channel estimate is available, there are efficient ways to jointly estimate the angles of arrivals and path delays [5]. The accuracy of estimates can be improved by exploiting the stationary properties of a space-time channel. The rationale is that, for mobile channels, the angles of arrivals and path delays are relatively stationary while the fading gain is changing more rapidly. In [6], the slot-invariance properties of angles of arrivals and delay spreads were translated into stationary spatial and temporal channel subspaces to avoid complex joint angle and delay estimation. Nevertheless, this approach fails to separate the dynamic fading gain from the stationary angle and delay responses. Thus, it is difficult to be extended in tracking the time-varying channel gain.

In this work, we present a reduced-rank space-time channel estimation scheme for DS-CDMA in dispersive MIMO channels. The proposed scheme has lower complexity and can achieve a lower estimation error in comparison with the full-rank one. Moreover, this scheme can be extended to track time-varying block-fading channels using the EM-based approach developed in [7]. The reduced-rank scheme also outperforms the full-rank one when joint angle and delay estimation is considered. The research conducted here is particularly useful for systems that involve large channel delay spreads and a large number of receive antennas, where rank reduction is inherent in a multivariate system model.

The paper is organized as follow. The system model and the receiver structure are described in Section II. Section III presents a reduced-rank channel estimation scheme based on the AP method [4]. In Section IV, the performance of the proposed reduced-rank scheme is evaluated and contrasted with the full-rank scheme. Concluding remarks are given in Section V.

II. System Model

Consider an uplink asynchronous DS-CDMA system with K users transmitting sequences of binary phase-shift keying (BPSK) symbols through their respective multipath channels. During the training periods, the transmitted baseband signal due to the kth is given by

$$x_k(t) = A_k \sum c_k(j)\varphi(t - jT_c), \qquad (1)$$

where T_c is the chip interval, $\{c_k(j)\}$ is the spreading sequence of user k taking values on the set {-1,+1}, and $\varphi(t)$ is a normalized chip waveform of duration T_c. At the receiver, an antenna array of L_r elements is employed. Assume that each transmitter is equipped with a single antenna. Then, the baseband multipath channel between the transmitter of user k and the receiver of the base station can be modeled as a

This research has been funded in part by the Integrated Media Systems Center, a National Science Foundation Engineering Research Center, Cooperative Agreement No. EEC-9529152.

single-input multiple-output channel with the following vector impulse response.

$$\underline{h}(t) = \sum_{l=1}^{L_p} \underline{\phi}(\theta_{kl})\beta_{kl}\delta(t-\tau_{kl}), \qquad (2)$$

where L_p is the number of paths in each user's channel, β_{kl} and τ_{kl} are the complex fading gain and the delay of the *l*th path of the *k*th user, respectively, and $\underline{\phi}(\theta_{kl})$ is the corresponding array response vector determined by the array geometry and the angle of arrival θ_{kl}. At each receive antenna, the baseband received signal is passed through a chip matched filter. The discrete-time model for the received signal is obtained by sampling the output of the matched filter at the chip rate. The temporal support of the channel is assumed to be $[0, L_cT_c]$.

First, let us consider the single user case and the results will be generalized to multiuser case later. The discrete-time model can be written as

$$\mathbf{y}(n) = \mathbf{H}\mathbf{x}(n) + \mathbf{n}(n), \qquad (3)$$

where $\mathbf{H} = [\mathbf{h}(0), \mathbf{h}(T_c), \ldots, \mathbf{h}((L_c-1)T_c)]$ is the $L_r \times L_c$ space-time channel matrix and $\mathbf{n}(n)$ is the temporal white zero-mean Gaussian noise with covariance $\mathbf{R_n}$. The space-time channel matrix also satisfies the following factorization [5]

$$\begin{aligned}\mathbf{H} &= \begin{bmatrix}\underline{\phi}(\theta_1) & \cdots & \underline{\phi}(\theta_{L_p})\end{bmatrix} \begin{bmatrix}\beta_1 \cdots 0 \\ \vdots \ \ddots \ \vdots \\ 0 \cdots \beta_{L_p}\end{bmatrix} \begin{bmatrix}\mathbf{g}^T(\tau_1) \\ \vdots \\ \mathbf{g}^T(\tau_{L_p})\end{bmatrix} \\ &= \Phi(\underline{\theta})\mathrm{diag}(\underline{\beta})\mathbf{G}^T(\underline{\tau}), \qquad (4)\end{aligned}$$

where $\mathbf{g}(\tau_l)$ is the $L_c \times 1$ sampled delay waveform of the convolution of the transmitted pulse and the matched filter.

After collecting N samples, (3) leads to

$$\mathbf{Y} = \mathbf{H}\mathbf{X} + \mathbf{N}, \qquad (5)$$

where $\mathbf{X}$ is an $L_c \times N$ Toeplitz matrix of training sequences. The unconstrained maximum likelihood estimate (MLE) of the channel is found to be

$$\hat{\mathbf{H}}_{ML} = \mathbf{R_{yx}}\mathbf{R_{xx}}^{-1}, \qquad (6)$$

where $\mathbf{R_{yx}} = \frac{1}{N}\sum_{n=1}^{N}\mathbf{y}(n)\mathbf{x}^H(n) = \frac{1}{N}\mathbf{Y}\mathbf{X}^H$, $\mathbf{R_{xx}} = \frac{1}{N}\sum_{n=1}^{N}\mathbf{x}(n)\mathbf{x}^H(n) = \frac{1}{N}\mathbf{X}\mathbf{X}^H$. Let $\tilde{\mathbf{y}} = \mathrm{vec}(\mathbf{Y})$, $\tilde{\mathbf{h}} = \mathrm{vec}(\mathbf{H})$, and $\tilde{\mathbf{n}} = \mathrm{vec}(\mathbf{N})$, where $\mathrm{vec}(\cdot)$ denotes the operator of stacking columns. The MLE is unbiased and the mean-square-error (MSE) of the MLE can be obtained from the trace of the covariance of the estimation error vector

$$\Delta\tilde{\mathbf{h}}_{ML} = \mathrm{vec}(\hat{\mathbf{H}}_{ML} - \mathbf{H}) = \frac{1}{N}((\mathbf{R_{xx}}^{-1}\mathbf{X})^* \otimes \mathbf{I}_{L_r})\tilde{\mathbf{n}}.$$

where $\mathbf{X}^*$ denotes the conjugate of $\mathbf{X}$ and $\otimes$ denotes the Kronecker product. The covariance can be obtained as

$$\mathrm{Cov}\{\Delta\tilde{\mathbf{h}}_{ML}\} = \frac{1}{N}\mathbf{R_{xx}}^{*-1} \otimes \mathbf{R_n}.$$

Therefore, the MSE is given by

$$\mathrm{MSE}_{ML} = \frac{1}{N}\mathrm{tr}\{\mathbf{R_{xx}}^{-1}\}\mathrm{tr}\{\mathbf{R_n}\}. \qquad (7)$$

III. Reduced-Rank Channel Estimation

A. Reduced-rank Estimation

Although the full-rank ML scheme is unbiased, it does not take the space-time structure into account. We observe that the rank of the space-time channel matrix is governed by the rank of the spatial matrix $\Phi(\underline{\theta})$ and the temporal matrix $\mathbf{G}(\underline{\tau})$. Let $\hat{\mathbf{y}}(n) = \hat{\mathbf{H}}\mathbf{x}(n)$ be an estimate of $\mathbf{y}(n)$ from $\mathbf{x}(n)$. The correlation matrix of the error vector $\mathbf{z}(n) = \mathbf{y}(n) - \hat{\mathbf{y}}(n)$ can be represented as

$$\mathbf{R_{zz}} = E\{\mathbf{z}(n)\mathbf{z}(n)^H\} = \mathbf{R_{yy}} - \mathbf{R_{yx}}\hat{\mathbf{H}}^H - \hat{\mathbf{H}}\mathbf{R_{yx}}^H + \hat{\mathbf{H}}\mathbf{R_{xx}}\hat{\mathbf{H}}^H.$$

We assume that $\mathbf{R_{xx}}$ and $\mathbf{R_{yy}}$ are nonsingular. The optimum choice of the estimate $\hat{\mathbf{H}}$ depends on the measure applied to $\mathbf{R_{zz}}$. The reduced-rank minimum mean square error (RRMMSE) estimate of $\mathbf{H}$, denoted by $\hat{\mathbf{H}}_{RRMMSE}$, is given by minimizing the trace of $\mathbf{R_{zz}}$ subject to the rank constraint $r < \min(L_r, L_c)$. By writing $\hat{\mathbf{H}}_{RRMMSE} = \mathbf{A}\mathbf{B}^H$, where $\mathbf{A}$ and $\mathbf{B}$ have full column rank *r*, we observe that

$$\begin{aligned}\mathrm{tr}\{\mathbf{R_{zz}}\} &= \mathrm{tr}\{(\mathbf{A}\mathbf{B}^H\mathbf{R_{xx}}\mathbf{B} - \mathbf{R_{yx}}\mathbf{B})(\mathbf{B}^H\mathbf{R_{xx}}\mathbf{B})^{-1} \\ &\qquad (\mathbf{A}\mathbf{B}^H\mathbf{R_{xx}}\mathbf{B} - \mathbf{R_{yx}}\mathbf{B})^H\} + f(\mathbf{B}) \qquad (8) \\ &= \mathrm{tr}\{\mathbf{R_{xx}}^{-1}(\mathbf{R_{xx}}\mathbf{B}\mathbf{A}^H\mathbf{A} - \mathbf{R_{yx}}^H\mathbf{A})(\mathbf{A}^H\mathbf{A})^{-1} \\ &\qquad (\mathbf{R_{xx}}\mathbf{B}\mathbf{A}^H\mathbf{A} - \mathbf{R_{yx}}^H\mathbf{A})^H\} + g(\mathbf{A}), \qquad (9)\end{aligned}$$

where

$$\begin{aligned}f(\mathbf{B}) &= \mathrm{tr}\{\mathbf{R_{yy}} - \mathbf{R_{yx}}\mathbf{B}(\mathbf{B}^H\mathbf{R_{xx}}\mathbf{B})^{-1}\mathbf{B}^H\mathbf{R_{yx}}^H\}, \\ g(\mathbf{A}) &= \mathrm{tr}\{\mathbf{R_{yy}} - \mathbf{R_{xx}}^{-1}\mathbf{R_{yx}}^H\mathbf{A}(\mathbf{A}^H\mathbf{A})^{-1}\mathbf{A}^H\mathbf{R_{yx}}\}.\end{aligned}$$

According to (8) and (9), we can minimize $tr\{\mathbf{R_{zz}}\}$ with respect to $\mathbf{A}$ and $\mathbf{B}$ alternately, and obtain the following iterative equations:

$$\begin{aligned}\mathbf{A}(i+1) &= \mathbf{R_{yx}}\mathbf{B}(i)(\mathbf{B}(i)^H\mathbf{R_{xx}}\mathbf{B}(i))^{-1}, \\ \mathbf{B}(i+1) &= \mathbf{R_{xx}}^{-1}\mathbf{R_{yx}}^H\mathbf{A}(i+1)(\mathbf{A}(i+1)^H\mathbf{A}(i+1))^{-1} \qquad (10)\end{aligned}$$

This iterative approach (called the AP method in [4]) is a generalization of the power method [8] in computing the principal components of a given matrix. Let us define $\mathbf{R}_{tr} \triangleq \mathbf{R_{yx}}\mathbf{R_{xx}}^{-H/2}$, which has the singular value decomposition (SVD) as

$$\mathbf{R}_{tr} = \sum_{i=1}^{min(L_r,L_t)} \sigma_i\mathbf{u}_i\mathbf{v}_i^H = \mathbf{U}_{tr,1}\mathbf{\Sigma}_{tr,1}\mathbf{V}_{tr,1}^H + \mathbf{U}_{tr,2}\mathbf{\Sigma}_{tr,2}\mathbf{V}_{tr,2}^H,$$

where $\sigma_i > \sigma_j$, $i < j$, and $\mathbf{\Sigma}_{tr,1}$ contains the largest r singular values. It was shown in [4] that $\mathbf{A}(i)\mathbf{B}(i)^H$ converges globally and exponentially to the reduced-rank wiener filter (RRWF):

$$\begin{aligned}\mathbf{A}(i)\mathbf{B}(i)^H &\rightarrow \mathbf{R_{yx}}\mathbf{R_{xx}}^{-H/2}\mathbf{V}_{tr,1}\mathbf{V}_{tr,1}^H\mathbf{R_{xx}}^{-1/2} \qquad (11) \\ &= \mathbf{U}_{tr,1}\mathbf{U}_{tr,1}^H\mathbf{R_{yx}}\mathbf{R_{xx}}^{-1}. \qquad (12)\end{aligned}$$

Upon convergence, we have

$$\begin{aligned}\mathbf{A}(i) &= \mathbf{R_{yx}}\mathbf{R_{xx}}^{-H/2}\mathbf{V}_{tr,1}\mathbf{Q} \\ \mathbf{B}(i) &= \mathbf{R_{xx}}^{-H/2}\mathbf{V}_{tr,1}\mathbf{Q}^{-H}, \qquad (13)\end{aligned}$$

where $\mathbf{Q}$ is a nonsingular matrix depending on the set of the initial values for the first iteration. The AP method can be implemented using only $O(\max(L_r^2 r, L_c^2 r))$ operations per iteration.

It is further assumed that the color noise is temporally uncorrelated. Let us define $\mathbf{R}_{det} \triangleq \mathbf{R}_{\mathbf{yy}}^{-1/2}\mathbf{R}_{\mathbf{yx}}\mathbf{R}_{\mathbf{xx}}^{-H/2}$. By minimizing the negative log-likelihood function of $\mathbf{H}$, one can easily verify that the reduced-rank ML (RRML) estimate of $\mathbf{H}$ is given by

$$\begin{aligned}\hat{\mathbf{H}}_{RRML} &= \mathbf{R}_{\mathbf{yx}}\mathbf{R}_{\mathbf{xx}}^{-H/2}\mathbf{V}_{det,1}\mathbf{V}_{det,1}^H\mathbf{R}_{\mathbf{xx}}^{-1/2} &(14)\\ &= \mathbf{R}_{\mathbf{yy}}^{1/2}\mathbf{U}_{det,1}\mathbf{U}_{det,1}^H\mathbf{R}_{\mathbf{yy}}^{-1/2}\mathbf{R}_{\mathbf{yx}}\mathbf{R}_{\mathbf{xx}}^{-1}, &(15)\end{aligned}$$

where $\mathbf{U}_{det,1}$ and $\mathbf{V}_{det,1}$ are similarly defined as above. The RRMLE is used to minimize the determinant of $\mathbf{R}_{\mathbf{zz}}$.

The MSE of the RRMLE can be derived as (see Appendix for the proof)

$$\begin{aligned}\mathrm{MSE}_{RR} = &\frac{1}{N}(\mathrm{tr}\{\mathbf{R}_{\mathbf{xx}}^{*-1/2}\Pi_{\mathbf{B}^*}\mathbf{R}_{\mathbf{xx}}^{*-H/2}\}\mathrm{tr}\{\mathbf{R}_{\mathbf{n}}\} + \quad (16)\\ &\mathrm{tr}\{\mathbf{R}_{\mathbf{xx}}^{*-1/2}\Pi_{\mathbf{B}^*}^{\perp}\mathbf{R}_{\mathbf{xx}}^{*-H/2}\}\mathrm{tr}\{\mathbf{R}_{\mathbf{n}}^{*1/2}\Pi_{\mathbf{A}}\mathbf{R}_{\mathbf{n}}^{*H/2}\}),\end{aligned}$$

where $\Pi_{\mathbf{B}^*} = \mathbf{B}^*(\mathbf{B}^T\mathbf{B}^*)^{-1}\mathbf{B}^T$ is the projection matrix of $\mathbf{B}^*$, $\Pi_{\mathbf{B}^*}^{\perp} = \mathbf{I} - \Pi_{\mathbf{B}^*}$ and Π_A is similarly defined. Since $\mathrm{tr}\{\mathbf{R}_{\mathbf{n}}^{*1/2}\Pi_{\mathbf{A}}\mathbf{R}_{\mathbf{n}}^{*H/2}\} \leq \mathrm{tr}\{\mathbf{R}_{\mathbf{n}}\}$, we have

$$\begin{aligned}\mathrm{MSE}_{RR} &\leq \frac{1}{N}(\mathrm{tr}\{\mathbf{R}_{\mathbf{xx}}^{*-1/2}(\Pi_{\mathbf{B}^*} + \Pi_{\mathbf{B}^*}^{\perp})\mathbf{R}_{\mathbf{xx}}^{*-H/2}\}\mathrm{tr}\{\mathbf{R}_{\mathbf{n}}\})\\ &= \mathrm{MSE}_{ML}.\end{aligned}$$

B. Estimation Error and Rank Selection

Let the reduced-rank estimate of $\mathbf{H}$ of rank order r be $\hat{\mathbf{H}}_r = \hat{\mathbf{A}}_r\hat{\mathbf{B}}_r^H$. The normalized estimation error corresponding to the selected rank r can be written as

$$\Delta\mathbf{H}_r = \|\mathbf{H} - \hat{\mathbf{H}}_r\|/\|\mathbf{H}\|. \quad (17)$$

If the rank of the composite channel matrix $\mathbf{H}$ is unknown, the optimum rank order $\hat{r}$ is given by

$$\hat{r} = \arg\min_r \|\mathbf{H} - \hat{\mathbf{H}}_r\|^2, \quad (18)$$

which can be determined by the GLRT developed in [2]. Consider the test statistics

$$\xi_{\hat{r}} = -N\sum_{i=\hat{r}+1}^{\min(L_r,L_c)} \ln(1 - \hat{\sigma}_i^2), \quad (19)$$

where N is the length of the collected sample vectors. Let $\delta_\epsilon(\hat{r})$ be the threshold such that the following condition is met:

$$Pr\{\omega \geq \delta_\epsilon(\hat{r})\} = \epsilon, \quad (20)$$

where ϵ is a small number (ϵ is chosen to be 0.05 in the simulation), and ω is approximated as a chi-square distributed random variable with $(L_r - \hat{r})(L_c - \hat{r})$ degrees of freedom. Then, the rank should be the first $\hat{r}$ that satisfies

$$\xi_{\hat{r}} \leq \delta_\epsilon(\hat{r}). \quad (21)$$

The singular values, $\hat{\sigma}_i^2$, in (19) can be updated via

$$\hat{\sigma}_i^2 = \hat{\mathbf{a}}_i^H\mathbf{R}_{\mathbf{yx}}\hat{\mathbf{b}}_i, \quad (22)$$

where $\hat{\mathbf{a}}_i$ and $\hat{\mathbf{b}}_i$ are the ith columns of matrices $\hat{\mathbf{A}}$ and $\hat{\mathbf{B}}$, respectively. These vectors are also called canonical vectors [4].

Assume that the rank-r estimation error is $O(\varepsilon)$ where ε is a small number depending only on the data length and the noise level. We have

$$\hat{\mathbf{A}}_r\hat{\mathbf{B}}_r^H = \mathbf{A}\mathbf{B}^H + O(\varepsilon). \quad (23)$$

Then, it follows that $\hat{\mathbf{A}}_r = \mathbf{A}_r\mathbf{Q}_a + O(\varepsilon)$, and $\hat{\mathbf{B}}_r = \mathbf{B}_r\mathbf{Q}_b + O(\varepsilon)$, where $\mathbf{Q}_a\mathbf{Q}_b^H = \mathbf{I}$. According to (13) and (22), the estimated singular values are given by the diagonal elements of

$$\begin{aligned}\hat{\mathbf{\Lambda}}_{tr,1} &= \hat{\mathbf{A}}_r^H\mathbf{C}_{\mathbf{yx}}\hat{\mathbf{B}}_r\\ &= \mathbf{Q}_a^H\mathbf{Q}^H\mathbf{\Lambda}_{tr,1}\mathbf{Q}^{-H}\mathbf{Q}_b + O(\varepsilon). \quad (24)\end{aligned}$$

Since $\mathbf{Q}_a^H\mathbf{Q}^H\mathbf{Q}^{-H}\mathbf{Q}_b = \mathbf{I}$, we have $\hat{\sigma}_i^2 = \sigma_i^2 + O(\varepsilon)$ for $i = 1, \ldots, r$. Therefore, we conclude that when the estimated rank $\hat{r}$ is the actual rank of $\mathbf{H}$, the accuracy of the first r estimated singular values is at least as good as that of the first r pairs of estimated canonical components.

C. Space-time Channel Estimation

For time-varying channels, the angle of arrival and time delay of each path are more stationary than the fading channel gain. Thus, we assume that the spatial matrix $\mathbf{\Phi}(\underline{\theta})$ and the temporal matrix $\mathbf{G}(\underline{\tau})$ are time-invariant during M blocks, while the channel gain $\underline{\beta}(m)$ changes block by block. By the vectorization operator, (4) leads to

$$\tilde{\mathbf{h}}(m) = (\mathbf{G} \otimes \Phi)\mathrm{vec}(\mathrm{diag}(\underline{\beta}(m))) = (\mathbf{G} \diamond \Phi)\underline{\beta}(m), \quad (25)$$

where $\diamond$ denotes the Khatri-Rao product, which is a column-wise Kronecker product. Define $\mathbf{U} = \mathbf{G} \diamond \Phi$. By collecting the channel estimate $\hat{\mathbf{H}}(m)$ from different M blocks, we can form an estimation equation for both $\mathbf{U}$ and β.

The MLE of parameters $\{\mathbf{U}, \underline{\beta}\}$ can be reduced to the minimization of the loss function

$$\mathcal{L}(\mathbf{U}, \underline{\beta}) = \frac{1}{M}\sum_{m=1}^{M} \|\hat{\mathbf{h}}(m) - \mathbf{U}\underline{\beta}(m)\|^2, \quad (26)$$

subject to the constraint that $\mathrm{rank}(\mathbf{U}) = L_p$, where $\hat{\mathbf{h}}(m) = \mathrm{vec}(\hat{\mathbf{H}}(m))$. The loss function can be minimized by solving the two sub-optimization problems with respect to $\mathbf{U}$ and $\underline{\beta}$ alternatively and iteratively as

$$\begin{aligned}\hat{\underline{\beta}}(m) &= (\mathbf{U}^H\mathbf{U})^{-1}\mathbf{U}^H\hat{\mathbf{h}}(m), &(27)\\ \hat{\mathbf{U}} &= \arg\min_{\mathbf{U}} \mathrm{tr}((\mathbf{I}_{L_rL_c} - \Pi_{\mathbf{U}})\mathbf{R}_{\hat{\mathbf{h}}}), &(28)\end{aligned}$$

where $\Pi_{\mathbf{U}} = \mathbf{U}(\mathbf{U}^H\mathbf{U})^{-1}\mathbf{U}^H$ is the projection matrix onto the subspace spanned by the columns of $\mathbf{U}$, and $\mathbf{R}_{\hat{\mathbf{h}}} = \frac{1}{M}\sum_{m=1}^{M}\hat{\mathbf{h}}(m)\hat{\mathbf{h}}^H(m)$. Many of the well-known methods such as MUSIC, ESPRIT [9], and JADE-ESPRIT [5] can be used to solve (26). Please note that the minimizer in (28) should satisfy $\hat{\Pi}_{\mathbf{U}} = \mathbf{E}_{\mathbf{R}_{\hat{\mathbf{h}}}}\mathbf{E}_{\mathbf{R}_{\hat{\mathbf{h}}}}^H$ where $\mathbf{E}_{\mathbf{R}_{\hat{\mathbf{h}}}}$ contains the leading eigenvectors of correlation matrix $\mathbf{R}_{\hat{\mathbf{h}}}$.

After $\mathbf{U}$ is estimated, the block fading channel (5) can be rewritten as

$$\tilde{\mathbf{y}}(m) = \Gamma(\mathbf{X}(m), \mathbf{U})\underline{\beta}(m), \quad (29)$$

where $\Gamma(\mathbf{X}(m), \mathbf{U}) = (\mathbf{X}(m) \otimes \mathbf{I}_{Lr})\mathbf{U}$. Therefore, the recursive EM algorithm developed in [10] can be used to track the time-varying $\underline{\beta}(m)$.

D. Multiuser Case

For the multiuser case, the discrete-time model (5) can be written as

$$\mathbf{Y} = \tilde{\mathbf{H}}\tilde{\mathbf{X}} + \mathbf{N}, \tag{30}$$

where $\tilde{\mathbf{H}} = [\mathbf{H}_1, \ldots, \mathbf{H}_K]$ is the multiuser channel matrix and $\tilde{\mathbf{X}} = [\mathbf{X}_1^T, \ldots, \mathbf{X}_K^T]^T$ collects the multiuser training sequences. If we treat the interference signal plus noise as a color Gaussian vector, the reduced-rank channel estimation scheme developed above is still valid.

To improve the performance of channel estimates under multiple access interference (MAI), we adopt an interference cancellation scheme. The received signal for user k is replaced by

$$\mathbf{Y}_k = \mathbf{Y} - \sum_{i=1, i\neq k}^{K} \hat{\mathbf{H}}_i \mathbf{X}_i, \tag{31}$$

where $\hat{\mathbf{H}}_i$ is the estimated channel parameters of the previous stage of interference cancellation. A new channel estimate of $\mathbf{H}_k$ can be obtained for user k using the less noisy received signal $\mathbf{Y}_k$. The interference cancellation process is repeated until $\hat{\mathbf{H}}_k$ reaches it steady-state value. Note that the received signal for channel estimation differs from one user to the other.

IV. Simulation Results

Simulation results are presented to demonstrate the performance of the proposed reduced-rank maximum likelihood channel estimation scheme. It is compared with the full-rank scheme. A case of four users with an equal transmitted power is considered. User no. 1 is the desired user. A random spreading sequence is assigned to each user. An array of $L_r = 8$ half-wavelength spaced sensors are employed at the receiver, and the number of paths corresponding to each user's channel is set to $L_p = 10$. The spatial diversity of the desired user is fixed at $r = 4$ with the angles of arrivals equal to $[-5(2), 0(3), 5(2), 10(3)]$ and time delays $[0, 2, 5, 10, 15, 20, 25, 30, 35, 38]T_c$. The multipath fading coefficients are generated from a complex Gaussian distribution with a zero mean. The experimental MSEs of estimates are averaged over 500 Monte Carlo runs.

Fig. 1 shows that, after 6 iterations, the normalized estimation error is very small. A lower bound (10^{-3}) is reached when the SNR is fixed at 15dB. The straight line shows that the AP method converges exponentially. Each line represents the result from a different experimental environment with different fading coefficients.

Fig. 2 compares the analytical and simulated results for the full-rank as well as the reduced-rank estimates, where MSE is plotted as a function of the signal-to-noise-ratio (SNR). The reduced-rank MSE outperforms the full-rank one in all cases. For example, to achieve a $\mathrm{MSE} = 2 \times 10^{-2}$, a 3dB gain in SNR is provided if reduced-rank scheme is adopted.

Fig. 3 gives the relationship between the length of the training sequence, the input SNR and the accuracy of channel estimate. For example, to achieve $\mathrm{MSE} = 2 \times 10^{-2}$, a gain of 5dB in SNR can be obtained if the length of the training sequence is extended from 200 to 300.

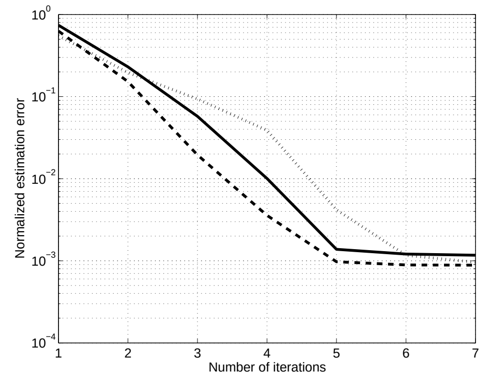

Fig. 1. The convergence of the normalized estimation error as a function of the iteration number with $L_r = 10$, $L_c = 40$, $N = 200$, $r = 4$ and SNR=15dB.

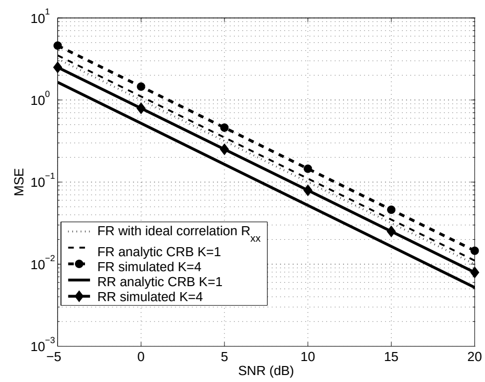

Fig. 2. The MSE-versus-SNR plot for the full-rank (FR) and the reduced-rank (RR) estimates of $\mathbf{H}$ with $L_r = 10$, $L_c = 40$, $N = 200$, and $r = 4$.

Fig. 4 compares results of using the full-rank and the reduce-rank schemes to estimate the space-time response matrix $\mathbf{U}$. The reduced-rank estimate still outperforms the full-rank one in the MSE performance. As shown in the figure, to achieve $\mathrm{MSE} = 2 \times 10^{-2}$, a gain of 5dB in SNR is achieved if the reduced-rank channel estimate $\hat{\mathbf{H}}_{RR}$ is used for the joint angle and delay estimation model in (26).

V. Conclusion

A reduced-rank space-time channel estimation scheme was proposed for multiuser DS-CDMA systems under dispersive frequency selective fading. The proposed method is based on the AP method with the aid of training sequences. Both analytical and simulation results demonstrated that the reduced-rank channel estimate outperforms the full-rank one in terms of MSE for all SNR values. The reduced-rank channel estimation scheme also provides better performance than the full-rank one when used as the input data for the joint angle and delay estimation scheme.

Appendix: Cramer-Rao Bound Derivation

Applying the vectorization operator to (5), we have

$$\tilde{\mathbf{y}} = \mathrm{vec}(\mathbf{Y}) = (\mathbf{X}^T \otimes \mathbf{I}_{L_r})\tilde{\mathbf{h}} + \tilde{\mathbf{n}}.$$

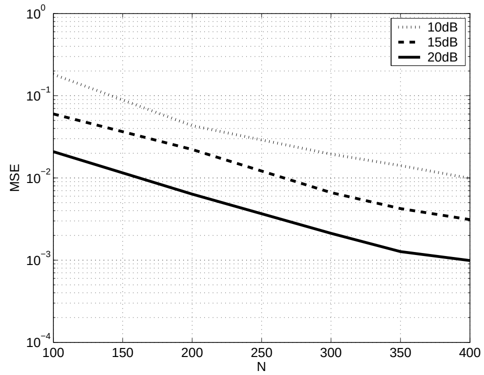

Fig. 3. The MSE versus the length N of the training sequence for the reduced-rank estimate $\hat{\mathbf{H}}_{RR}$.

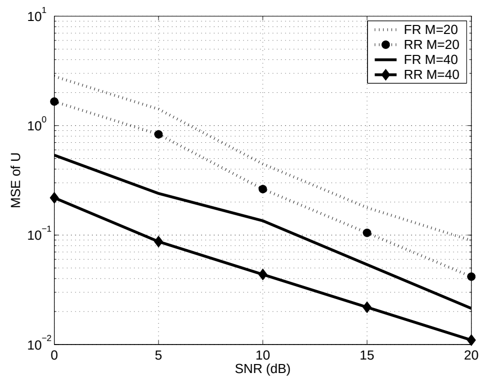

Fig. 4. The MSE-versus-SNR plot for the estimates of space-time response matrix $\mathbf{U}$ with the initial channel estimates from the full-rank (FR) and the reduced-rank (RR) schemes. The channel parameters are $L_r = 10$, $L_c = 40$, $N = 200$, $r = 4$, and $L_p = 10$.

Recall that $\mathbf{H} = \mathbf{A}\mathbf{B}^H$. Let $\tilde{\mathbf{a}} = \mathrm{vec}(\mathbf{A})$ and $\tilde{\mathbf{b}} = \mathrm{vec}(\mathbf{B}^H)$. Then, we have

$$\tilde{\mathbf{h}} = (\mathbf{B}^* \otimes \mathbf{I}_{L_r})\tilde{\mathbf{a}} = (\mathbf{I}_{L_c} \otimes \mathbf{A})\tilde{\mathbf{b}}.$$

Let $\mathbf{J}$ denote the so-called complex *Fisher information matrix* (FIM), which is associated with the estimation problem under discussion. Then, we have

$$\mathbf{J} = \mathrm{E}\{\Delta\Delta^H\}, \quad \Delta = \frac{\partial \ln p(y; \underline{\vartheta})}{\partial \underline{\vartheta}^*}. \tag{32}$$

It is well known that the covariance matrix of any unbiased estimator of $\alpha(\underline{\vartheta})$ is lower bounded by the Cramer-Rao bound (CRB) matrix, *i.e.*

$$\mathrm{E}\{(\hat{\alpha} - \mathrm{E}\{\hat{\alpha}\})(\hat{\alpha} - \mathrm{E}\{\hat{\alpha}\})^H\} \geq \mathbf{D}\mathbf{J}^{-1}\mathbf{D}^H,$$

where for any two Hermitian matrices $\mathbf{O}$ and $\mathbf{P}$, $\mathbf{O} \geq \mathbf{P}$ means that the difference $\mathbf{O} - \mathbf{P}$ is positive semi-definite, and

$$\mathbf{D} \triangleq \mathrm{E}\{(\hat{\alpha} - \mathrm{E}\{\hat{\alpha}\})\Delta^H\} = \frac{\partial b}{\partial \underline{\vartheta}^T} + \frac{\partial \alpha}{\partial \underline{\vartheta}^T}, \tag{33}$$

where $b = \mathrm{E}\{\alpha\} - \alpha$ is the bias of the estimate. The parameter vector is $\underline{\vartheta} = [\tilde{\mathbf{a}}^T \ \tilde{\mathbf{b}}^T]^T$. Assume that an unbiased estimate is obtained, *i.e.* $\hat{r} = r$. By substituting $\alpha = \tilde{\mathbf{h}}$ and $b = \mathbf{0}$ into (33), we have

$$\mathbf{D} = \left[\mathbf{B}^* \otimes \mathbf{I}_{L_r} \quad \mathbf{I}_{L_c} \otimes \mathbf{A}\right],$$

and (32) becomes

$$\mathbf{J} = N\mathbf{D}^H(\mathbf{R}_{\mathbf{xx}}^* \otimes \mathbf{R}_{\mathbf{n}}^{-1})\mathbf{D},$$

The rank of $\mathbf{J}$ is determined by the rank of $\mathbf{D}$ since matrix $\mathbf{R}_{\mathbf{xx}}^* \otimes \mathbf{R}_{\mathbf{n}}^{-1}$ is positive definite. Therefore, $\mathrm{rank}(\mathbf{J}) = \mathrm{rank}(\mathbf{D}) = r(L_r + L_c - r)$, and the CRB should be modified as [11]

$$\begin{aligned} \mathrm{Cov}\{\hat{\mathbf{h}}\} &\geq \mathrm{CRB}\{\hat{\mathbf{h}}\} = \mathbf{D}\mathbf{J}^\dagger\mathbf{D}^H \\ &= \frac{1}{N}(\mathbf{R}_{\mathbf{xx}}^{*-1/2} \otimes \mathbf{R}_{\mathbf{n}}^{1/2})\Pi_{\mathbf{W}}(\mathbf{R}_{\mathbf{xx}}^{*-H/2} \otimes \mathbf{R}_{\mathbf{n}}^{H/2}), \end{aligned} \tag{34}$$

where $\mathbf{J}^\dagger$ is the Moore-Penrose pseudoinverse, $\mathbf{W} = (\mathbf{R}_{\mathbf{xx}}^{*1/2} \otimes \mathbf{R}_{\mathbf{n}}^{-1/2})\mathbf{D}$ and $\Pi_{\mathbf{W}} = \mathbf{W}(\mathbf{W}^H\mathbf{W})^\dagger\mathbf{W}^H$ is the projection matrix onto the subspace spanned by the columns of $\mathbf{W}$. It can be shown that

$$\Pi_{\mathbf{W}} = \Pi_{\mathbf{B}^*}^{\perp} \otimes \Pi_{\mathbf{A}} + \Pi_{\mathbf{B}^*} \otimes \mathbf{I}_{L_r}.$$

After some simplification, $\mathrm{MSE}\{\hat{\mathbf{h}}\}$ can be evaluated as

$$\begin{aligned} \mathrm{MSE}\{\hat{\mathbf{h}}\} &= \frac{1}{N}(\mathrm{tr}\{\mathbf{R}_{\mathbf{xx}}^{*-1/2}\Pi_{\mathbf{B}^*}\mathbf{R}_{\mathbf{xx}}^{*-H/2}\}\mathrm{tr}\{\mathbf{R}_{\mathbf{n}}\} + \\ &\quad \mathrm{tr}\{\mathbf{R}_{\mathbf{xx}}^{*-1/2}\Pi_{\mathbf{B}^*}^{\perp}\mathbf{R}_{\mathbf{xx}}^{*-H/2}\}\mathrm{tr}\{\mathbf{R}_{\mathbf{n}}^{*1/2}\Pi_{\mathbf{A}}\mathbf{R}_{\mathbf{n}}^{*H/2}\}). \end{aligned} \tag{35}$$

If training sequences are designed with ideal correlation $\mathbf{R}_{\mathbf{xx}} = \sigma^2\mathbf{I}_{L_c}$ and the noise is white, the MSE above can be further simplified to

$$\mathrm{MSE}\{\hat{\mathbf{h}}\} = \frac{\sigma_n^2}{\sigma_x^2 N} r(L_r + L_c - r). \tag{36}$$

References

[1] K. M. Chugg, A. Anastasopoulos, and X. Chen, *Iterative Detection*, Kluwer Academic Publishers, 2001.

[2] P. Stoica and M. Viberg, “Maximum likelihood parameter and rank estimation in reduced-rank multivariate linear regressions,” *IEEE Trans. on Signal Processing*, vol. 44, pp. 3069–3078, Dec. 1996.

[3] M. Nicoli, M. Sternad, U. Spagnolini, and A. Ahlen, “Reduced-rank channel estimation and tracking in time-slotted CDMA systems,” in *Proc. IEEE ICC*, May 2002, vol. 1, pp. 533–537.

[4] Y. Hua, M. Nikpour, and P. Stoica, “Optimal reduced-rank estimation and filtering,” *IEEE Trans. on Signal Processing*, vol. 49, pp. 457–469, Mar. 2001.

[5] M. C. Vanderveen, A. van der Veen, and A. Paulraj, “Estiamtion of multipath parameters in wireless communications,” *IEEE Trans. on Signal Processing*, vol. 46, pp. 682–690, Mar. 1998.

[6] M. Nicoli, O. Simeone, and U. Spagnolini, “Multislot estimation of fast-varying space-time communication channels,” *IEEE Trans. on Signal Processing*, vol. 51, pp. 1184–1195, May 2003.

[7] S.-H. Wu, U Mitra, and C.-C. J. Kuo, “Joint channel estimation and multiuser dectection for multipath fading channels in DS-CDMA systems,” in *Proc. IEEE ICC*. Anchorage, Alaska, May 2003.

[8] G. H. Golub and C. F. Van Loan, *Matrix Computations*, Johns Hopkins Univ. Press, 1996.

[9] H. L. Van Trees, *Optimum Array Processing*, Wiley, 2002.

[10] H. Zamiri-Jafarian and S. Pasupathy, “EM-based recursive estimation of channel parameters,” *IEEE Trans. on Communications*, vol. 47, no. 9, pp. 1297–1302, Sept. 1999.

[11] P. Stoica and T. L. Marzetta, “Parameter estimation problems with singular information matrices,” *IEEE Trans. on Signal Processing*, vol. 49, pp. 87–90, Jan. 2001.