Reduced-Rank Adaptive Filtering Based on Joint Iterative Optimization of Adaptive Filters

Rodrigo C. de Lamare, *Member, IEEE*, and Raimundo Sampaio-Neto

***Abstract*—This letter proposes a novel adaptive reduced-rank filtering scheme based on joint iterative optimization of adaptive filters. The novel scheme consists of a joint iterative optimization of a bank of full-rank adaptive filters that forms the projection matrix and an adaptive reduced-rank filter that operates at the output of the bank of filters. We describe minimum mean-squared error (MMSE) expressions for the design of the projection matrix and the reduced-rank filter and low-complexity normalized least-mean squares (NLMS) adaptive algorithms for its efficient implementation. Simulations for an interference suppression application show that the proposed scheme outperforms in convergence and tracking the state-of-the-art reduced-rank schemes at significantly lower complexity.**

***Index Terms*—Adaptive filters, iterative methods.**

I. Introduction

In adaptive filtering [1], one can find a huge number of algorithms with different trade-offs between performance and complexity. They range from the simple and low-complexity least-mean squares (LMS) to the fast converging though complex recursive least squares (RLS) techniques. Several attempts to provide cost-effective adaptive filters with fast convergence performance have been made with variable step-size algorithms, data-reusing, sub-band and frequency-domain schemes, and RLS algorithms with linear complexity. A challenging problem which remains unsolved by conventional techniques is that when the number of elements in the filter is large, the algorithm requires a large number of samples to reach its steady-state behavior. In these situations, even RLS algorithms require an amount of data proportional to $2M$ [1] in stationary environments to reach steady state, where M is the filter length, and this may lead to unacceptable convergence performance. In dynamic scenarios, large filters usually fail or provide poor performance in tracking signals embedded in interference.

Reduced-rank filtering [2]–[9] is a powerful and effective technique in low sample support situations and in problems with large filters. The advantages of reduced-rank adaptive filters are their faster convergence speed and better tracking performance than full-rank techniques when dealing with large number of weights. Several reduced-rank methods and systems have been proposed in the last several years, namely, eigen-decomposition techniques [3], [4], the multistage Wiener filter (MWF) [6], [7], and the auxiliary vector filtering (AVF) algorithm [8]. The main problem with the above techniques is their high complexity and the existence of numerical problems for implementation.

In this letter, we propose an adaptive reduced-rank filtering scheme based on combinations of adaptive filters. Unlike related work on combinations of full-rank filters [10], the novel scheme consists of a joint iterative optimization of a bank of full-rank adaptive filters which constitutes the projection matrix and an adaptive reduced-rank filter that operates at the output of the bank of full-rank filters. Differently from [11], the proposed scheme estimates a scalar, allows filter updates for each successive observation, is adaptive, and has low complexity. The essence of the proposed approach is to change the role of adaptive filters. The bank of adaptive filters is responsible for performing dimensionality reduction, whereas the reduced-rank filter effectively estimates the desired signal. Despite the large dimensionality of the projection matrix and its associated slow learning behavior, the proposed and existing [7], [8] reduced-rank techniques enjoy in practice a very fast convergence. The reason is that even an inaccurate or rough estimation of the projection matrix is able to provide an appropriate dimensionality reduction for the reduced-rank filter, whose behavior will govern most of the performance of the overall scheme. We describe MMSE expressions for the design of the projection matrix and the reduced-rank filter along with simple NLMS adaptive algorithms for its computationally efficient implementation.

II. Reduced-Rank MMSE Parameter Estimation and Problem Statement

The MMSE filter is the vector $\mathbf{w} = [w_1\, w_2\, \dots\, w_M]^T$, which is designed to minimize the MSE cost function

$$J = E\left[|d(i) - \mathbf{w}^H \mathbf{r}(i)|^2\right] \quad (1)$$

where $d(i)$ is the desired signal; $\mathbf{r}(i) = [r_0^{(i)}\, \dots\, r_{M-1}^{(i)}]^T$ is the received data; $(\cdot)^T$ and $(\cdot)^H$ denote transpose and Hermitian transpose, respectively; and $E[\cdot]$ stands for expectation. The set of parameters $\mathbf{w}$ can be estimated via standard stochastic gradient or least-squares estimation techniques [1]. However, the laws that govern the convergence behavior of these estimation techniques imply that the convergence speed of these algorithms is proportional to M, the number of elements in the estimator. Thus, large M implies slow convergence. A reduced-rank algorithm attempts to circumvent this limitation in terms of speed of convergence by reducing the number of adaptive coefficients and extracting the most important features of the processed data. This dimensionality reduction is accomplished by projecting the

Manuscript received April 23, 2007; revised July 13, 2007. The associate editor coordinating the review of this manuscript and approving it for publication was Prof. Richard J. Kozick.

R. C. de Lamare is with the Department of Electronics, University of York, York Y010 5DD, U.K. (e-mail: rcdl500@ohm.york.ac.uk).

R. Sampaio-Neto is with CETUC/PUC-RIO, 22453-900 Rio de Janeiro, Brazil (e-mail: raimundo@cetuc.puc-rio.br).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/LSP.2007.907995

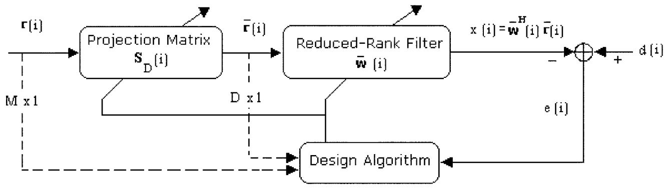

Fig. 1. Proposed reduced-rank scheme.

received vectors onto a lower dimensional subspace. Specifically, consider an $M \times D$ projection matrix $\mathbf{S}_D$ which carries out a dimensionality reduction on the received data as given by

$$\bar{\mathbf{r}}(i) = \mathbf{S}_D^H \mathbf{r}(i) \tag{2}$$

where, in what follows, all D-dimensional quantities are denoted with a "bar." The resulting projected received vector $\bar{\mathbf{r}}(i)$ is the input to a tapped-delay line filter represented by the D vector $\bar{\mathbf{w}} = [\bar{w}_1 \ \bar{w}_2 \ \ldots \bar{w}_D]^T$ for time interval i. The filter output corresponding to the ith time instant is

$$x(i) = \bar{\mathbf{w}}^H \bar{\mathbf{r}}(i). \tag{3}$$

If we consider the MMSE design in (1) with the reduced-rank parameters, we obtain

$$\bar{\mathbf{w}} = \bar{\mathbf{R}}^{-1} \bar{\mathbf{p}} \tag{4}$$

where $\bar{\mathbf{R}} = E[\bar{\mathbf{r}}(i)\bar{\mathbf{r}}^H(i)] = \mathbf{S}_D^H \mathbf{R} \mathbf{S}_D$ is the reduced-rank covariance matrix, $\mathbf{R} = E[\mathbf{r}(i)\mathbf{r}^H(i)]$ is the full-rank covariance matrix, $\bar{\mathbf{p}} = E[d^*(i)\bar{\mathbf{r}}(i)] = \mathbf{S}_D^H \mathbf{p}$, and $\mathbf{p} = E[d^*(i)\mathbf{r}(i)]$. The associated MMSE for a rank D estimator is expressed by

$$\text{MMSE} = \sigma_d^2 - \bar{\mathbf{p}}^H \bar{\mathbf{R}}^{-1} \bar{\mathbf{p}} = \sigma_d^2 - \mathbf{p}^H \mathbf{S}_D (\mathbf{S}_D^H \mathbf{R} \mathbf{S}_D)^{-1} \mathbf{S}_D^H \mathbf{p} \tag{5}$$

where σ_d^2 is the variance of $d(i)$. In the Appendix , we provide a necessary and sufficient condition for a projection $\mathbf{S}_D$ with dimensions $M \times D$ to not modify the MMSE and discuss the existence of multiple solutions. Based upon the problem statement above, the rationale for reduced-rank schemes can be simply put as follows. How to efficiently (or optimally) design a transformation matrix $\mathbf{S}_D$ with dimensions $M \times D$ that projects the observed data vector $\mathbf{r}(i)$ with dimensions $M\times 1$ onto a reduced-rank data vector $\bar{\mathbf{r}}(i)$ with dimensions $D\times 1$? In the next section, we present the proposed reduced-rank approach.

III. Proposed Reduced-Rank Scheme

Here we detail the principles of the proposed reduced-rank scheme using a projection operator based on adaptive filters. The novel scheme, depicted in Fig. 1, employs a projection matrix $\mathbf{S}_D(i)$ with dimensions $M \times D$ to process a data vector with dimensions $M\times 1$, that is responsible for the dimensionality reduction. The reduced-rank filter $\bar{\mathbf{w}}(i)$ with dimensions $D\times 1$ processes the reduced-rank data vector $\bar{\mathbf{r}}(i)$ in order to yield a scalar estimate $x(i)$. The projection matrix $\mathbf{S}_D(i)$ and the reduced-rank filter $\bar{\mathbf{w}}(i)$ are jointly optimized in the proposed scheme according to the MMSE criterion.

Specifically, the projection matrix is structured as a bank of D full-rank filters $\mathbf{s}_d(i) = [s_{1,d}(i) \ s_{2,d}(i) \ \ldots \ s_{M,d}(i)]^T$ ($d = 1, \ldots, D$) with dimensions $M\times 1$ as given by $\mathbf{S}_D(i) = [\ \mathbf{s}_1(i) \mid \mathbf{s}_2(i) \mid \ \ldots \ |\mathbf{s}_D(i)\ ]$. Let us now mathematically express the output estimate $x(i)$ of the reduced-rank scheme as a function of the received data $\mathbf{r}(i)$, the projection matrix $\mathbf{S}_D(i)$, and the reduced-rank filter $\bar{\mathbf{w}}(i)$

$$x(i) = \bar{\mathbf{w}}^H(i)\mathbf{S}_D^H(i)\mathbf{r}(i) = \bar{\mathbf{w}}^H(i)\bar{\mathbf{r}}(i). \tag{6}$$

Note that for $D = 1$, the novel scheme becomes a conventional full-rank filtering scheme with an addition weight parameter w_D that provides a gain. For $D > 1$, the signal processing tasks are changed and the full-rank filters compute a subspace projection and the reduced-rank filter estimates the desired signal.

The MMSE expressions for the filters $\mathbf{S}_D(i)$ and $\bar{\mathbf{w}}(i)$ can be computed through the following cost function:

$$\begin{aligned} J &= E\left[|d(i) - \bar{\mathbf{w}}^H(i)\mathbf{S}_D^H(i)\mathbf{r}(i)|^2\right] \\ &= E\left[|d(i) - \bar{\mathbf{w}}^H(i)\bar{\mathbf{r}}(i)|^2\right]. \end{aligned} \tag{7}$$

By fixing the projection $\mathbf{S}_D(i)$ and minimizing (7) with respect to $\bar{\mathbf{w}}(i)$, the reduced-rank filter weight vector becomes

$$\bar{\mathbf{w}}(i) = \bar{\mathbf{R}}^{-1}(i)\bar{\mathbf{p}}(i) \tag{8}$$

where $\bar{\mathbf{R}}(i) = E[\mathbf{S}_D^H(i)\mathbf{r}(i)\mathbf{r}^H(i)\mathbf{S}_D(i)] = E[\bar{\mathbf{r}}(i)\bar{\mathbf{r}}^H(i)]$, $\bar{\mathbf{p}}(i) = E[d^*(i)\mathbf{S}_D^H(i)\mathbf{r}(i)] = E[d^*(i)\bar{\mathbf{r}}(i)]$. We proceed with the proposed joint optimization by fixing $\bar{\mathbf{w}}(i)$ and minimizing (7) with respect to $\mathbf{S}_D(i)$. We then arrive at the following expression for the projection operator:

$$\mathbf{S}_D(i) = \mathbf{R}^{-1}(i)\mathbf{P}_D(i)\mathbf{R}_w(i) \tag{9}$$

where $\mathbf{R}(i) = E[\mathbf{r}(i)\mathbf{r}^H(i)]$, $\mathbf{P}_D(i) = E[d^*(i)\mathbf{r}(i)\mathbf{w}^H(i)]$, and $\mathbf{R}_w(i) = E[\mathbf{w}(i)\mathbf{w}^H(i)]$. The associated MMSE is

$$\text{MMSE} = \sigma_d^2 - \bar{\mathbf{p}}^H(i)\bar{\mathbf{R}}^{-1}(i)\bar{\mathbf{p}}(i) \tag{10}$$

where $\sigma_d^2 = E[|d(i)|^2]$. The filter expressions in (8) and (9) are not closed-form solutions for $\bar{\mathbf{w}}(i)$ and $\mathbf{S}_D(i)$ since (8) is a function of $\mathbf{S}_D(i)$ and (9) depends on $\bar{\mathbf{w}}(i)$, and thus, it is necessary to iterate (8) and (9) with an initial guess to obtain a solution. The MWF [6] employs the operator $\mathbf{S}_D = \left[\mathbf{p} \ \mathbf{R}\mathbf{p} \ \ldots \ \mathbf{R}^{D-1}\mathbf{p}\right]$ that projects the data onto the Krylov subspace. Unlike the MWF approach, the new scheme provides an iterative exchange of information between the reduced-rank filter and the projection matrix and leads to a much simpler adaptive implementation than the MWF. The projection matrix reduces the dimension of the input data, whereas the reduced-rank filter attempts to estimate the desired signal. The key strategy lies in the joint optimization of the filters. The rank D must be set by the designer to ensure appropriate performance, and the reader is referred to [12] for rank selection methods. In the next section, we seek iterative solutions via adaptive algorithms.

IV. Adaptive Algorithms

Here we describe an adaptive NLMS implementation, convergence conditions, and detail the computational complexity in arithmetic operations of the proposed reduced-rank scheme.

A. Adaptive Algorithms

Let us consider the following Lagrangian cost function:

$$\begin{aligned} \mathcal{L} &= \|\mathbf{w}(i+1) - \mathbf{w}(i)\|^2 + \|\mathbf{S}_D(i+1) - \mathbf{S}_D(i)\|^2 \\ &\quad + \Re[\lambda_1^*(d(i) - \mathbf{w}^H(i+1)\mathbf{S}_D^H(i)\mathbf{r}(i)] \\ &\quad + \Re[\lambda_2^*(d(i) - \mathbf{w}^H(i)\mathbf{S}_D^H(i+1)\mathbf{r}(i)] \end{aligned} \tag{11}$$

where λ_1, λ_2 are scalar Lagrange multipliers, $||\cdot||^2$ denotes the Frobenius norm, and the operator $\Re[\cdot]$ retains the real part of the argument. By computing the gradient terms of (11) with respect to $\bar{\mathbf{w}}(i+1)$, $\mathbf{S}_D(i+1)$, λ_1 and λ_2, setting them to 0 and solving the resulting equations, we obtain

$$\nabla_{\bar{\mathbf{w}}(i+1)}\mathcal{L}=2(\bar{\mathbf{w}}(i+1)-\bar{\mathbf{w}}(i))+\mathbf{S}_D^H(i)\mathbf{r}(i)\lambda_1=\mathbf{0} \quad (12)$$

$$\nabla_{\mathbf{S}_D(i+1)}\mathcal{L}=2(\mathbf{S}_D(i+1)-\mathbf{S}_D(i))+\mathbf{r}(i)\bar{\mathbf{w}}^H(i)\lambda_2=\mathbf{0} \quad (13)$$

$$\nabla_{\lambda_1}\mathcal{L}=d(i)-\bar{\mathbf{w}}^H(i+1)\mathbf{S}_D^H(i)\mathbf{r}(i)=0 \quad (14)$$

$$\nabla_{\lambda_2}\mathcal{L}=d(i)-\bar{\mathbf{w}}^H(i)\mathbf{S}_D^H(i+1)\mathbf{r}(i)=0. \quad (15)$$

By solving the above equations and introducing the convergence factors μ_0 and η_0, the proposed jointly optimized and iterative NLMS algorithms for parameter estimation become

$$\bar{\mathbf{w}}(i+1)=\bar{\mathbf{w}}(i)+\mu(i)e^*(i)\bar{\mathbf{r}}(i) \quad (16)$$

$$\mathbf{S}_D(i+1)=\mathbf{S}_D(i)+\eta(i)e^*(i)\mathbf{r}(i)\bar{\mathbf{w}}^H(i) \quad (17)$$

where $e(i)=d(i)-\bar{\mathbf{w}}^H(i)\mathbf{S}_D^H(i)\mathbf{r}(i)$, $\mu(i)=\mu_0/(\mathbf{r}^H(i)\mathbf{r}(i))$, and $\eta(i)=\eta_0/(\bar{\mathbf{w}}^H(i)\bar{\mathbf{w}}(i)\mathbf{r}^H(i)\mathbf{r}(i))$ are the time-varying step sizes. The algorithms described in (16)-(17) have a complexity $O(DM)$. The proposed scheme trades-off a full-rank filter against D full-rank adaptive filters as the projection matrix $\mathbf{S}_D(i)$ and one reduced-rank adaptive filter $\bar{\mathbf{w}}(i)$ operating simultaneously and exchanging information. The iteration and convergence occurs over several observations, and here we consider only one iteration per symbol (i).

B. Convergence Conditions

Define the error matrices at time index i as $\mathbf{e}_{\bar{\mathbf{w}}}(i)=\bar{\mathbf{w}}(i)-\bar{\mathbf{w}}_{\rm opt}$ and $\mathbf{e}_{\mathbf{S}_D}(i)=\mathbf{S}_D(i)-\mathbf{S}_{D,\rm opt}$, where $\bar{\mathbf{w}}_{\rm opt}$ and $\mathbf{S}_{D,\rm opt}$ are the optimal parameter estimators. Because of the joint optimization procedure, both filters have to be considered jointly. By substituting the expressions of $\mathbf{e}_{\bar{\mathbf{w}}}(i)$ and $\mathbf{e}_{\mathbf{S}_D}(i)$ in (16) and (17), taking expectations and simplifying the terms, we obtain

$$\begin{bmatrix} E[\mathbf{e}_{\bar{\mathbf{w}}}(i+1)] \\ E[\mathbf{e}_{\mathbf{S}_D}(i+1)] \end{bmatrix} = \mathbf{A}\begin{bmatrix} E[\mathbf{e}_{\bar{\mathbf{w}}}(i)] \\ E[\mathbf{e}_{\mathbf{S}_D}(i)] \end{bmatrix} + \mathbf{B} \quad (18)$$

where

$$\mathbf{A}=\begin{bmatrix} (\mathbf{I}-E[\mu(i)]\bar{\mathbf{R}}) & \mathbf{0} \\ E[\nu(i)]\sigma_w^2\mathbf{R}\mathbf{S}_{D,\rm opt} & (\mathbf{I}-E[\nu(i)]\sigma_w^2\mathbf{R}) \end{bmatrix}$$

$$\mathbf{B}=\begin{bmatrix} E[\mu(i)](\mathbf{R}\mathbf{S}_D(i)\bar{\mathbf{w}}_{\rm opt}-\bar{\mathbf{p}}) \\ E[\nu(i)]\sigma_w^2(\mathbf{R}\mathbf{S}_{D,\rm opt}\bar{\mathbf{w}}_{\rm opt}-\mathbf{p}) \end{bmatrix}$$

and $\sigma_w^2=E[||\bar{\mathbf{w}}(i)||^2]$. The above equation implies that the stability of the algorithms depends on the spectral radius of $\mathbf{A}$. For convergence, the step sizes should be chosen such that the eigenvalues of $\mathbf{A}^H\mathbf{A}$ are less than one.

C. Computational Complexity

Here we detail the computational complexity in terms of additions and multiplications of the proposed schemes with NLMS and other existing algorithms, namely, the Full-rank with NLMS and RLS, the MWF [7] with NLMS and RLS, and the AVF [8], as shown in Table I. The MWF [7] has a complexity $O(D\bar{M}^2)$, where the variable dimension of the vectors $\bar{M}=M-d$ varies according to the rank $d=1,\ldots,D$. The proposed scheme is much simpler than the Full-rank with RLS, the MWF, and the AVF and slightly more complex than the Full-rank with NLMS (for $D \ll M$, as will be explained later).

TABLE I
COMPUTATIONAL COMPLEXITY OF ALGORITHMS

	Number of operations per symbol	
Algorithm	Additions	Multiplications
Full-rank-NLMS	$3M-1$	$3M+2$
Full-rank-RLS	$3(M-1)^2+M^2+2M$	$6M^2+2M+2$
Proposed-NLMS	$2DM+M+4D-2$	$3DM+M+3D+6$
MWF-NLMS	$D(2\bar{M}^2-3\bar{M}+1)$	$D(2\bar{M}^2+5\bar{M}+7)$
MWF-RLS	$D(4(\bar{M}-1)^2+2\bar{M})$	$D(4\bar{M}^2+2\bar{M}+3)$
AVF	$D(M^2+3(M-1)^2)-1$ $+D(5(M-1)+1)+2M$	$D(4M^2+4M+1)$ $+4M+2$

V. SIMULATIONS

In this section, we assess the proposed reduced-rank scheme and algorithms in a CDMA interference suppression application. We consider the uplink of a symbol synchronous BPSK DS-CDMA system with K users, N chips per symbol, and L propagation paths. Assuming that the channel is constant during each symbol interval and the randomly generated spreading codes are repeated from symbol to symbol, the received signal after filtering by a chip-pulse matched filter and sampled at chip rate yields the M-dimensional received vector

$$\mathbf{r}(i)=\sum_{k=1}^{K}\mathbf{H}_k(i)A_k\mathbf{C}_k\mathbf{b}_k(i)+\mathbf{n}(i) \quad (19)$$

where $M=N+L-1$, $\mathbf{n}(i)=[n_1(i)\ \ldots\ n_M(i)]^T$ is the complex Gaussian noise vector with $E[\mathbf{n}(i)\mathbf{n}^H(i)]=\sigma^2\mathbf{I}$, the symbol vector is $\mathbf{b}_k(i)=[b_k(i+L_s-1)\ \ldots\ b_k(i)\ \ldots\ b_k(i-L_s+1)]^T$, the amplitude of user k is A_k, L_s is the intersymbol interference span, the $((2L_s-1)\cdot N)\times(2L_s-1)$ block diagonal matrix $\mathbf{C}_k$ is formed with N-chips shifted versions of the signature $\mathbf{s}_k=[a_k(1)\ldots a_k(N)]^T$ of user k, and the $M\times(2\cdot L_s-1)\cdot N$ convolution matrix $\mathbf{H}_k(i)$ is constructed with shifted versions of the $L\times 1$ channel vector $\mathbf{h}_k(i)=[h_{k,0}(i)\ \ldots\ h_{k,L_p-1}(i)]^T$ on each column and zeros elsewhere. For all simulations, we use $\bar{\mathbf{w}}(0)=\mathbf{0}_{D,1}$, $\mathbf{S}_D(0)=[\mathbf{I}_D\ \mathbf{0}_{D,M-D}]^T$, assume $L=9$ as an upper bound, use 3-path channels with relative powers given by 0, -3, and -6 dB, where in each run, the spacing between paths is obtained from a discrete uniform random variable between 1 and 2 chips and average the experiments over 200 runs. The system has a power distribution amongst the users for each run that follows a log-normal distribution with associated standard deviation equal to 1.5 dB.

We compare the proposed scheme with the Full-rank [1], the MWF [7], and the AVF [8] techniques for the design of linear receivers, where the reduced-rank filter $\bar{\mathbf{w}}(i)$ with D coefficients provides an estimate of the desired symbol for the desired user (user 1 in all experiments) using the signal-to-interference-plus-noise ratio (SINR) [7]. We consider the SINR performance versus the rank D with optimized parameters (μ_0, ν_0, and forgetting factors λ) for all schemes. The results in Fig. 2 indicate that the best rank for the proposed scheme is $D=4$ (which will be used in the remaining experiments), and it is very close to the optimal full-rank MMSE. Studies with systems with different processing gains show that D is invariant to the system size, which brings considerable computational savings. In practice, the rank D can be adapted in order to obtain fast convergence and ensure good steady-state performance and tracking after convergence.

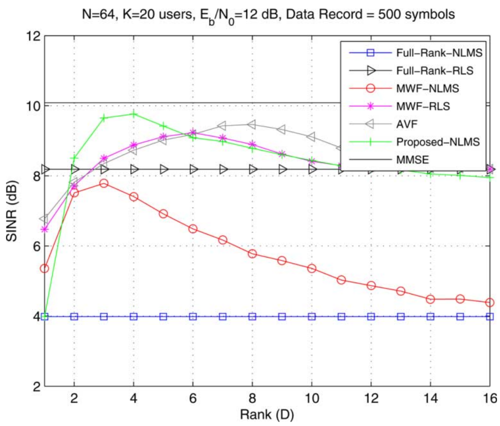

Fig. 2. SINR performance versus rank (D).

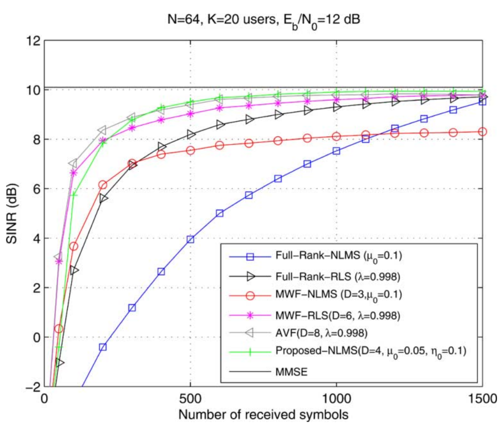

Fig. 3. SINR performance versus number of received symbols.

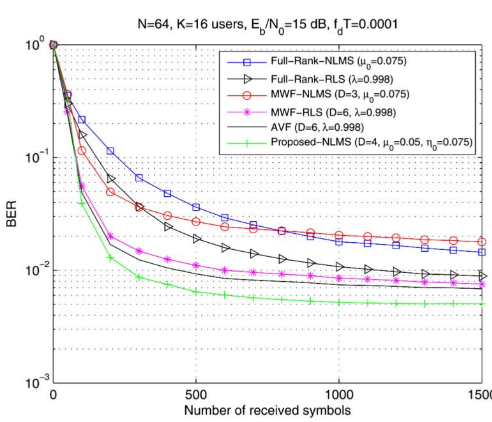

Fig. 4. BER performance versus number of received symbols.

We show an experiment in Fig. 3 where the adaptive filters are set to converge to the same SINR. The NLMS version of the MWF is known to have problems in these situations since it does not tridiagonalize its covariance matrix [7] and thus is unable to approach the MMSE. The curves show an excellent performance for the proposed scheme and algorithms, which converge much faster than the full-rank filter, are comparable to the more complex MWF-RLS and AVF schemes, at much lower complexity.

The BER convergence performance in a mobile communications situation is shown in Fig. 4. The channel coefficients are obtained with Clarke's model [13], and the adaptive filters of all methods are trained with 250 symbols and then switch to decision-directed mode. The results show that the proposed scheme has a much better performance than the existing approaches and is able to adequately track the desired signal.

VI. Conclusions

We proposed a novel reduced-rank scheme based on joint iterative optimization of adaptive filters with a low-complexity implementation using NLMS algorithms. In the proposed scheme, the full-rank adaptive filters are responsible for estimating the subspace projection rather than the desired signal, which is estimated by a small reduced-rank filter. The results for CDMA interference suppression show a performance significantly better than existing schemes and close to the optimal full-rank MMSE.

Appendix

Given a $M \times D$ projection matrix $\mathbf{S}_D$, where $D \leq M$, the MMSE is achieved if and only if $\mathbf{w}$ which minimizes (1) belongs to the $\mathrm{Range}\{\mathbf{S}_D\}$, i.e., $\mathbf{w}$ lies in the subspace generated by $\mathbf{S}_D$. In this case, we have $\mathrm{MMSE}(\bar{\mathbf{w}}) = \mathrm{MMSE}(\mathbf{w}) = \sigma_d^2 - \mathbf{p}^H\mathbf{R}^{-1}\mathbf{p}$. For a general $\mathbf{S}_D$, we have $\mathrm{MMSE}(\bar{\mathbf{w}}) \geq \sigma_d^2 - \mathbf{p}^H\mathbf{R}^{-1}\mathbf{p}$. From the above analysis, we can conclude that there exists multiple solutions to the proposed optimization problem. However, our studies indicate that there are no local minima and the performance is insensitive to initialization, provided we select the initial values $\bar{\mathbf{w}}(0)$ and $\mathbf{S}_D(0)$ which do not instabilize the algorithm and annihilate the signal, respectively.

References

[1] P. S. R. Diniz, *Adaptive Filtering: Algorithms and Practical Implementations*, 2nd ed. Boston, MA: Kluwer, 2002.

[2] L. L. Scharf and D. W. Tufts, "Rank reduction for modeling stationary signals," *IEEE Trans. Acoust., Speech, Signal Process.*, vol. ASSP-35, pp. 350–355, Mar. 1987.

[3] A. M. Haimovich and Y. Bar-Ness, "An eigenanalysis interference canceler," *IEEE Trans. Signal Process.*, vol. 39, no. 1, pp. 76–84, Jan. 1991.

[4] Y. Song and S. Roy, "Blind adaptive reduced-rank detection for DS-CDMA signals in multipath channels," *IEEE J. Sel. Areas Commun.*, vol. 17, no. 9, pp. 1960–1970, Nov. 1999.

[5] R. Singh and L. B. Milstein, "Adaptive interference suppression for DS/CDMA," *IEEE Trans. Commun.*, vol. 50, no. 12, pp. 1902–1905, Nov. 2002.

[6] J. S. Goldstein, I. S. Reed, and L. L. Scharf, "A multistage representation of the Wiener filter based on orthogonal projections," *IEEE Trans. Inf. Theory*, vol. 44, no. 11, pp. 2943–2959, Nov. 1998.

[7] M. L. Honig and J. S. Goldstein, "Adaptive reduced-rank interference suppression based on the multistage Wiener filter," *IEEE Trans. Commun.*, vol. 50, no. 6, pp. 986–994, Jun. 2002.

[8] D. A. Pados and G. N. Karystinos, "An iterative algorithm for the computation of the MVDR filter," *IEEE Trans. Signal Process.*, vol. 49, no. 2, pp. 290–300, Feb. 2001.

[9] R. C. de Lamare and R. Sampaio-Neto, "Adaptive reduced-rank MMSE filtering with interpolated FIR filters and adaptive interpolators," *IEEE Signal Process. Lett.*, vol. 12, no. 3, pp. 177–180, Mar. 2005.

[10] J. Arenas-Garcia, A. R. Figueiras-Vidal, and A. H. Sayed, "Mean-square performance of a convex combination of two adaptive filters," *IEEE Trans. Signal Process.*, vol. 54, no. 3, pp. 1078–1090, Mar. 2006.

[11] Y. Hua and M. Nikpour, "Computing the reduced rank Wiener filter by IQMD," *IEEE Signal Process. Lett.*, vol. 6, no. 9, pp. 240–242, Sep. 1999.

[12] Q. Haoli and S. N. Batalama, "Data record-based criteria for the selection of an auxiliary vector estimator of the MMSE/MVDR filter," *IEEE Trans. Commun.*, vol. 51, no. 10, pp. 1700–1708, Oct. 2003.

[13] T. S. Rappaport, *Wireless Communications*. Englewood Cliffs, NJ: Prentice-Hall, 1996.