Bidirectional MMSE Algorithms for Interference Suppression in DS-CDMA Systems over Fast Fading Channels

Patrick Clarke and Rodrigo C. de Lamare
Department of Electronics, University of York, York, UK, YO10 5DD
Email: pjc500@ohm.york.ac.uk, rcdl500@ohm.york.ac.uk

Abstract— **This paper presents adaptive bidirectional minimum mean-square error (MMSE) parameter estimation algorithms for fast-fading channels. The time correlation between successive channel gains is exploited to improve the estimation and tracking capabilities of adaptive algorithms and provide robustness against time-varying channels. Bidirectional normalized least mean-square (NLMS) and conjugate gradient (CG) algorithms are devised along with adaptive mixing parameters that adjust to the time-varying channel correlation properties. An analysis of the proposed algorithms is provided along with a discussion of their performance advantages. Simulations for an application to interference suppression in DS-CDMA systems show the advantages of the proposed algorithms.**

Keywords— **Bidirectional signal processing, adaptive interference suppression, fast-fading channels, adaptive receivers.**

I. Introduction

Low-complexity reception and interference suppression is essential in multiuser mobile systems if battery power is to be conserved, data-rates improved, and quality of service enhanced. Conventional adaptive schemes fulfil many of these requirements and have been a significant focus of the research literature [1], [2]. However, in time-varying fading channels commonly associated with highly mobile systems, adaptive techniques encounter problems when estimating and tracking parameters such as the receive filter and the channel coefficients. The development of cost-effective parameter estimation and tracking techniques for highly dynamic channels remains a very challenging problem.

Existing strategies to enhance the performance of estimation techniques include the use of optimized convergence parameters in conventional adaptive algorithms to extend their ability to deal with fading and improve their convergence and tracking performances [3], [4], [5], [6]. However, the stability of adaptive step-sizes and forgetting factors is a concern unless they are constrained to lie within a predefined region [4]. Furthermore, the fundamental problem of demodulating the transmitted data symbols whilst suppressing multiuser interference (MUI) remains. Approaches to avoid and/or improve the tracking and estimation of the fading coefficients have been reported in [7], [8], [9], [10], [11], [12]. Although a channel might be highly time-variant, adjacent fading coefficients can be approximately equal and have a significant level of correlation. These properties can be exploited to obtain a sequence of faded symbols where the primary purpose of the receive filter is to suppress interference and track the ratio between successive fading coefficients, dispensing with the estimation of the fading coefficients themselves. However, this scheme has a number of limitations due to the fact that only one correlation time instant is employed, which results in instability and a difficulty to track highly time-variant signals.

In this paper, a bidirectional minimum mean-square error (MMSE) based interference suppression scheme for highly-dynamic fading channels is presented. The channel correlation between adjacent time instants is exploited to improve the robustness, tracking and convergence performances of existing adaptive schemes. Bidirectional normalized least mean squares (NLMS) and conjugate gradient (CG) algorithms are devised along with a mixing strategy that adaptively weights the contribution of the considered time instants. An analysis of the proposed schemes is given and establishes the factors behind their behaviour and improved performance. The proposed schemes are applied to DS-CDMA systems and simulations show that they significantly outperform existing schemes.

The remainder of the paper is structured as follows, Section II describes the DS-CDMA signal model, states the problem and explains the motivation of the work. Section III presents the proposed bidirectional processing scheme. The algorithmic implementations of the proposed bidirectional methods are given in Section IV, and analysis of the proposed algorithms in Section V. Simulations and performance evaluation are given in Section VI and conclusions in Section VII.

II. DS-CDMA Signal Model and Problem Statement

Consider the uplink of a multiuser DS-CDMA system with K users, processing gain N and multipath channels with L paths. The $M \times 1$ received signal after chip-pulse matched filtering and sampling at the chip rate is given by

$$\mathbf{r}[i] = A_1 b_1[i]\mathbf{H}_1[i]\mathbf{c}_1[i] + \underbrace{\sum_{k=2}^{K} A_k b_k[i]\mathbf{H}_k[i]\mathbf{c}_k[i]}_{\text{MUI}} + \boldsymbol{\eta}[i] + \mathbf{n}[i], \quad (1)$$

where $M = N + L - 1$, and $\mathbf{c}_k[i]$ and A_k are the spreading sequence and signal amplitude of the k^{th} user, respectively. The $M \times N$ matrix $\mathbf{H}_k[i]$ models the channel propagation effects for the k^{th} user, $b_k[i]$ corresponds to the transmitted symbol of the k^{th} user, $\boldsymbol{\eta}[i]$ is the intersymbol interference (ISI) vector and $\mathbf{n}[i]$ is the noise vector.

The design of linear receivers consists of processing the received vector $\boldsymbol{r}[i]$ with the receive filter $\boldsymbol{w}_k[i]$ with M coefficients that provide an estimate of the desired symbol as follows

$$z_k[i] = \boldsymbol{w}_k^H[i]\boldsymbol{r}[i], \quad (2)$$

where the detected symbol is given by $\hat{b}_k[i] = Q(\boldsymbol{w}_k^H[i]\boldsymbol{r}[i])$, where $\mathrm{Q}(\cdot)$ is a function that performs the detection according to the constellation employed. It is also possible to use non-linear receiver techniques. The problem we are interested in solving in this work is how to estimate the parameter vector $\boldsymbol{w}_k[i]$ of the receive filter in fast time-varying channels.

III. Proposed Bidirectional Processing Scheme

Adaptive parameter estimation techniques have two primary objectives: estimation and tracking of the desired parameters. However, in fast fading channels the combination of these two objectives places unrealistic demands on conventional estimation schemes. Differential techniques reduce these demands by relieving the adaptive filter of the task of tracking fading coefficients. This is achieved by posing an optimization problem where the ratio between two successive received samples is the quantity to be tracked. Such an approach is enabled by the presumption that, although the fading is fast, there is a significant level of correlation between the adjacent channel samples

$$f_1[i] = E\left[h_1[i]h_1^*[i+1]\right] \geq 0, \quad (3)$$

where $h_1[i]$ is the channel coefficient of the desired user. The interference suppression of the resulting filter is improved in fast fading environments compared to conventional adaptive filters but only the ratio of adjacent fading samples is obtained. Consequently, differential modulation, where the ratio between adjacent symbols is the data carrying mechanism, are suited to differential MMSE schemes.

However, limiting the optimization process to two adjacent samples exposes the differential MMSE process to the negative effects of uncorrelated samples

$$E\left[h_1[i]h_1^*[i+1]\right] \approx 0, \quad (4)$$

and does not exploit the correlation that may be present between two or more adjacent samples, i.e.

$$\begin{aligned} f_2[i] &= E\left[h_1[i]h_1^*[i-1]\right] > 0 \text{ and} \\ f_3[i] &= E\left[h_1[i+1]h_1^*[i-1]\right] > 0. \end{aligned} \quad (5)$$

To address these weaknesses, we propose a bidirectional MSE cost function based on adjacent received data vectors so that the number of channel fading scenarios under which an algorithm performs reliable estimation and tracking is increased and the performance improved. Termed the bidirectional MMSE, due to the use of multiple and adjacent time instants, the proposed scheme can exploit the correlation between successive received signals and reuse data [2]. The optimization problem of the proposed scheme is given by

$$\begin{aligned} \boldsymbol{w}_{\text{MMSE}} = \arg\min_{\boldsymbol{w}} E\Bigg[\sum_{d=0}^{D-2} \sum_{l=d+1}^{D-1} \rho_n[i] |b[i-d]\boldsymbol{w}^H \boldsymbol{r}[i-l] \\ - b[i-l]\boldsymbol{w}^H \boldsymbol{r}[i-d]|^2 \Bigg], \end{aligned} \quad (6)$$

where $\boldsymbol{w}_{\text{MMSE}}$ is the expected value of the filter, $\rho_n[i]$ is a weighting factor used in the cost function to address problems with uncorrelated fading coefficients and $n = d(D-3)+l+1$. Note that the time instants of interest have been altered to avoid the use of future samples. In addition to (6), an output power constraint is required to avoid the trivial zero correlator solution

$$E\left[\left|\boldsymbol{w}^H \boldsymbol{r}[i]\right|^2\right] = 1. \quad (7)$$

In fast-fading channels, the correlation between the considered time instant is unlikely to be equal. Therefore variable weighting or mixing of the cost function will be required to obtain improved performance. However, the setting of the weights is problematic if they are to be fixed. An adaptive scheme is preferable which can take account of the time-varying channels. The errors extracted from the cost function (6) are chosen as the metric for this approach. This provides an input to the weighting factor calculation process that is directly related to the optimization in (6). The time-varying mixing factors are given by

$$\rho_n[i] = \lambda_e \rho_n[i-1] + (1-\lambda_e)\frac{e_T[i] - |e_n[i]|}{e_T[i]} \quad (8)$$

where $e_T[i] = |e_1[i]| + |e_2[i]| + \ldots + |e_D[i]|$ and the individual errors terms are calculated for $d = 0, \ldots, D-2$ and $l = d+1, \ldots, D-1$

$$e_n[i] = b[i-d]\boldsymbol{w}^H[i]\boldsymbol{r}[i-l] - b[i-l]\boldsymbol{w}^H[i]\boldsymbol{r}[i-d]. \quad (9)$$

The forgetting factor, $0 \leq \lambda_e \leq 1$, is user defined and, along with the normalization by the total error, $e_T[i]$, and $\sum_{n=1}^{D} \rho_n[0] = 1$, ensures $\sum_{n=1}^{D} \rho_n[i] = 1$ and a convex combination at each time instant.

IV. Proposed Bidirectional Algorithms

In this section, the proposed bidirectional MMSE-based algorithms based on (6) are derived. In particular, we concentrate on the case where $D = 3$ as it captures most of the gains of the proposed scheme, and develop bidirectional NLMS and CG adaptive algorithms. Let us consider the following cost function

$$\begin{aligned} C(\boldsymbol{w}[i], \rho_n[i]) = E\big[&\rho_1[i]|b[i]\boldsymbol{w}^H[i]\boldsymbol{r}[i-1] \\ &- b[i-1]\boldsymbol{w}^H[i]\boldsymbol{r}[i]|^2 \\ &+ \rho_2[i]|b[i]\boldsymbol{w}^H[i]\boldsymbol{r}[i-2] \\ &- b[i-2]\boldsymbol{w}^H[i]\boldsymbol{r}[i]|^2 \\ &+ \rho_3[i]|b[i-1]\boldsymbol{w}^H[i]\boldsymbol{r}[i-2] \\ &- b[i-2]\boldsymbol{w}^H[i]\boldsymbol{r}[i-1]|^2\big] \end{aligned} \quad (10)$$

The time-varying mixing factors are adjusted by

$$\rho_n[i] = \lambda_e \rho_n[i-1] + (1-\lambda_e)\frac{e_T[i] - |e_n[i]|}{e_T[i]} \quad (11)$$

where $e_T[i] = |e_1[i]| + |e_2[i]| + |e_3[i]|$ and the individual error terms are given by

$$\begin{aligned} e_1[i] &= b[i]\boldsymbol{w}^H[i-1]\boldsymbol{r}[i-1] - b[i-1]\boldsymbol{w}^H[i-1]\boldsymbol{r}[i] \\ e_2[i] &= b[i]\boldsymbol{w}^H[i-1]\boldsymbol{r}[i-2] - b[i-2]\boldsymbol{w}^H[i-1]\boldsymbol{r}[i] \\ e_3[i] &= b[i-1]\boldsymbol{w}^H[i-1]\boldsymbol{r}[i-2] - b[i-2]\boldsymbol{w}^H[i-1]\boldsymbol{r}[i-1]. \end{aligned} \quad (12)$$

The forgetting factor, $0 \leq \lambda_e \leq 1$, is user defined and, along with the normalization by the total error, $e_T[i]$, and $\sum_{n=1}^{3} \rho_n[0] = 1$, ensures $\sum_{n=1}^{3} \rho_n[i] = 1$ and a convex combination at each time instant.

A. Bidirectional NLMS Algorithm

We first devise a low-complexity bidirectional NLMS algorithm that iteratively computes the solution of (10). The instantaneous gradient of (10) is taken with respect to $\boldsymbol{w}^*[i]$, and the errors terms of (12) are incorporated to yield the update equation

$$\begin{aligned}\boldsymbol{w}[i] = \quad & \boldsymbol{w}[i-1] + \tfrac{\mu}{M[i]}\left[\rho_1[i]b[i-1]\boldsymbol{r}[i]e_1[i]\cdots\right. \\ & \left.+\rho_2[i]b[i-2]\boldsymbol{r}[i]e_2[i] + \rho_3[i]b[i-2]\boldsymbol{r}[i-1]e_3[i]\right],\end{aligned} \tag{13}$$

where μ is the step-size and the adaptive mixing parameters have been included. The normalization factor, $M[i]$, is given by

$$M[i] = \lambda_M M[i-1] + (1-\lambda_M)\boldsymbol{r}^H[i]\boldsymbol{r}[i] \tag{14}$$

where λ_M is an exponential forgetting factor [9]. The enforcement of the constraint is performed by the denominator of (13) and ensures that the filter $\boldsymbol{w}[i]$ does not tend towards a zero correlator. The complexity of this algorithm is $O(DM)$, which corresponds to roughly $D = 3$ times that of the NLMS.

B. Bidirectional Conjugate Gradient Algorithm

Due to the incongruous form of the bidirectional formulation and the conventional matrix inversion lemma based recursive least-squares (RLS) algorithm, an alternative bidirectional CG algorithm is now derived. We begin with the time-averaged autocorrelation and crosscorrelation structures $\bar{\mathbf{R}}$ and $\bar{\mathbf{t}}$ from (10)

$$\begin{aligned}\bar{\boldsymbol{R}}_1[i] &= \lambda\bar{\boldsymbol{R}}_1[i-1] + b[i-1]\boldsymbol{r}[i]\boldsymbol{r}^H[i]b^*[i-1] \\ \bar{\boldsymbol{R}}_2[i] &= \lambda\bar{\boldsymbol{R}}_2[i-1] + b[i-2]\boldsymbol{r}[i]\boldsymbol{r}^H[i]b^*[i-2] \\ \bar{\boldsymbol{R}}_3[i] &= \lambda\bar{\boldsymbol{R}}_3[i-1] + b[i-2]\boldsymbol{r}[i-1]\boldsymbol{r}^H[i-1]b^*[i-2]\end{aligned} \tag{15}$$

and

$$\begin{aligned}\bar{\boldsymbol{t}}_1[i] &= \lambda\bar{\boldsymbol{t}}_3[i-1] + b[i-1]\boldsymbol{r}[i]\boldsymbol{r}^H[i-1]\boldsymbol{w}[i-1]b^*[i] \\ \bar{\boldsymbol{t}}_2[i] &= \lambda\bar{\boldsymbol{t}}_2[i-1] + b[i-2]\boldsymbol{r}[i]\boldsymbol{r}^H[i-2]\boldsymbol{w}[i-1]b^*[i] \\ \bar{\boldsymbol{t}}_3[i] &= \lambda\bar{\boldsymbol{t}}_3[i-1] + b[i-2]\boldsymbol{r}[i-1]\boldsymbol{r}^H[i-2]\boldsymbol{w}[i-1]b^*[i-1]\end{aligned}, \tag{16}$$

respectively. After some algebraic manipulations with the terms, the final correlation structures are given by

$$\bar{\boldsymbol{R}}[i] = \rho_1[i]\bar{\boldsymbol{R}}_1[i] + \rho_2[i]\bar{\boldsymbol{R}}_2[i] + \rho_3[i]\bar{\boldsymbol{R}}_3[i] \tag{17}$$

$$\bar{\boldsymbol{t}}[i] = \rho_1[i]\bar{\boldsymbol{t}}_1[i] + \rho_2[i]\bar{\boldsymbol{t}}_2[i] + \rho_3[i]\bar{\boldsymbol{t}}_3[i] \tag{18}$$

where the adaptive mixing factors have been included. Inserting these structures into the standard CG quadratic form yields

$$J(\boldsymbol{w}[i], \rho_n[i]) = \boldsymbol{w}^H[i]\boldsymbol{R}[i]\boldsymbol{w}[i] - \boldsymbol{t}^H[i]\boldsymbol{w}[i]. \tag{19}$$

From [13], the unique minimizer of (19) is also the minimizer of

$$\boldsymbol{R}[i]\boldsymbol{w}[i] = \boldsymbol{t}[i]. \tag{20}$$

At each time instant, a number of iterations of the following method are required to reach an accurate solution, where the iterations are indexed with the variable j. At the i^{th} time instant the gradient and direction vectors are initialized as

$$\boldsymbol{g}_0[i] = \nabla_{\boldsymbol{w}^*[i]}J(\boldsymbol{w}[i], \rho_n[i]) = \boldsymbol{R}[i]\boldsymbol{w}_0[i] - \boldsymbol{t}[i] \tag{21}$$

and

$$\boldsymbol{d}_0[i] = -\boldsymbol{g}_0[i], \tag{22}$$

respectively, where the gradient expression is equivalent to those used in the derivation of the NLMS algorithm. The vectors $\boldsymbol{d}_j[i]$ and $\boldsymbol{d}_{j+1}[i]$ are $\boldsymbol{R}[i]$ orthogonal with respect to $\boldsymbol{R}[i]$ such that $\boldsymbol{d}_j[i]\boldsymbol{R}[i]\boldsymbol{d}_l[i] = 0$ for $j \neq l$. At each iteration, the filter is updated as

$$\boldsymbol{w}_{j+1}[i] = \boldsymbol{w}_j[i] + \alpha_j[i]\boldsymbol{d}_j[i] \tag{23}$$

where $\alpha_j[i]$ is the minimizer of $J_{LS}(\boldsymbol{w}_{j+1}[i])$ such that

$$\alpha_j = \frac{-\boldsymbol{d}_j^H\boldsymbol{g}_j[i]}{\boldsymbol{d}_j^H[i]\boldsymbol{R}[i]\boldsymbol{d}_j[i]}. \tag{24}$$

The gradient vector is then updated according to

$$\boldsymbol{g}_{j+1}[i] = \boldsymbol{R}[i]\boldsymbol{w}_j[i] - \boldsymbol{t}[i] \tag{25}$$

and a new CG direction vector found

$$\boldsymbol{d}_{j+1}[i] = -\boldsymbol{g}_{j+1}[i] + \beta_j[i]\boldsymbol{d}_j[i] \tag{26}$$

where

$$\beta_j[i] = \frac{\boldsymbol{g}_{j+1}^H[i]\boldsymbol{R}[i]\boldsymbol{d}_j[i]}{\boldsymbol{d}_j^H[i]\boldsymbol{R}[i]\boldsymbol{d}_j[i]} \tag{27}$$

ensures the $\boldsymbol{R}[i]$ orthogonality between $\boldsymbol{d}_j[i]$ and $\boldsymbol{d}_l[i]$ where $j \neq l$. The iterations (23) - (27) are repeated until $j = j_{max}$.

V. Analysis of the Proposed Algorithms

The form of the bidirectional MSE cost function precludes the application of standard MSE analysis. Consequently, we concentrate on the signal to interference plus noise ratio (SINR) of the proposed NLMS algorithm to analyze its performance.

A. SINR Analysis

To begin, we convert the SINR expression given by

$$\text{SINR} = \frac{\boldsymbol{w}^{\text{H}}[\text{i}]\boldsymbol{R}_{\text{S}}\boldsymbol{w}[\text{i}]}{\boldsymbol{w}^{\text{H}}[\text{i}]\boldsymbol{R}_{\text{I}}\boldsymbol{w}[\text{i}]}, \tag{28}$$

where $\boldsymbol{R}_S$ and $\boldsymbol{R}_I$ are the signal and interference and noise correlation matrices, into a form amenable to analysis. Substituting in the filter error weight vector, $\boldsymbol{\varepsilon}[i] = \boldsymbol{w}[i] - \boldsymbol{w}_o[i]$, where $\boldsymbol{w}_o$ is the instantaneous standard optimal linear MMSE receiver, and taking the trace of the expectation yields

$$\text{SINR} = \frac{\boldsymbol{K}[\text{i}]\boldsymbol{R}_{\text{S}} + \boldsymbol{G}[\text{i}]\boldsymbol{R}_{\text{S}} + \text{P}_{\text{S,opt}}[\text{i}] + \boldsymbol{G}^{\text{H}}[\text{i}]\boldsymbol{R}_{\text{S}}}{\boldsymbol{K}[\text{i}]\boldsymbol{R}_{\text{I}} + \boldsymbol{G}[\text{i}]\boldsymbol{R}_{\text{I}} + \text{P}_{\text{I,opt}}[\text{i}] + \boldsymbol{G}^{\text{H}}[\text{i}]\boldsymbol{R}_{\text{I}}}, \tag{29}$$

where $\boldsymbol{K}[i] = E[\boldsymbol{\varepsilon}[i]\boldsymbol{\varepsilon}^H[i]]$, $\boldsymbol{G}[i] = E[\boldsymbol{w}_o[i]\boldsymbol{\varepsilon}^H[i]]$, $P_{\text{S,opt}}[i] = [\boldsymbol{w}_o^H[i]\boldsymbol{R}_{\text{S}}\boldsymbol{w}_o[i]]$ and $P_{\text{I,opt}}[i] = E[\boldsymbol{w}_o^H[i]\boldsymbol{R}_{\text{I}}\boldsymbol{w}_o[i]]$. From (29) it is clear that we need to pursue expressions for $\boldsymbol{K}[i]$ and $\boldsymbol{G}[i]$ to reach an analytical interpretation of the bidirectional scheme.

Substituting the filter error weight vector into the filter update expression of (13) yields a recursive expression for $\boldsymbol{\varepsilon}[i]$

$$\begin{aligned}\boldsymbol{\varepsilon}[i] = \ & \boldsymbol{\varepsilon}[i-1] \\ & + \big[\boldsymbol{I} + \mu\boldsymbol{r}[i]b[i-1]\boldsymbol{r}^H[i-1]b^*[i] - \mu\boldsymbol{r}[i]b[i-1]\boldsymbol{r}^H[n]b^*[i-1] \\ & +\mu\boldsymbol{r}[i]b[i-2]\boldsymbol{r}^H[i-2]b^*[i] - \mu\boldsymbol{r}[i]b[i-2]\boldsymbol{r}^H[n]b^*[i-2] \\ & +\mu\boldsymbol{r}[i-1]b[i-2]\boldsymbol{r}^H[i-2]b^*[i-1] \\ & - \mu\boldsymbol{r}[i-1]b[i-2]\boldsymbol{r}^H[i-1]b^*[i-2]\big]\,\boldsymbol{\varepsilon}[i-1] \\ & +\mu\boldsymbol{r}[i]b[i-1]e_{o,1}^*[i] + \mu\boldsymbol{r}[i]b[i-2]e_{o,2}^*[i] \\ & +\mu\boldsymbol{r}[i-1]b[i-2]e_{o,3}^*[i]\end{aligned}, \tag{30}$$

where the terms $e_{o,1}$, $e_{o,2}$ and $e_{o,3}$ are the error terms of (12) when the optimum filter $\boldsymbol{w}_o$ is used. Using the direct averaging approach of Kushner [14], the solution to the stochastic difference equation of (30) can be approximated by the solution to a second equation [2], such that

$$\begin{aligned} E\,\big[&\boldsymbol{I} + \mu\boldsymbol{r}[i]b[i-1]\boldsymbol{r}^H[i-1]b^*[i] - \mu\boldsymbol{r}[i]b[i-1]\boldsymbol{r}^H[n]b^*[i-1] \\ &+\mu\boldsymbol{r}[i]b[i-2]\boldsymbol{r}^H[i-2]b^*[i] - \mu\boldsymbol{r}[i]b[i-2]\boldsymbol{r}^H[n]b^*[i-2] \\ &+\mu\boldsymbol{r}[i-1]b[i-2]\boldsymbol{r}^H[i-2]b^*[i-1] \\ &-\mu\boldsymbol{r}[i-1]b[i-2]\boldsymbol{r}^H[i-1]b^*[i-2]\big] \\ &= \boldsymbol{I} + \mu\boldsymbol{F}_1 - \mu\boldsymbol{R}_1 + \mu\boldsymbol{F}_2 - \mu\boldsymbol{R}_2 + \mu\boldsymbol{F}_3 - \mu\boldsymbol{R}_3 \end{aligned} \tag{31}$$

where $\boldsymbol{F}$ and $\boldsymbol{R}$ are correlations matrices. Specifically, $\boldsymbol{R}_1$, $\boldsymbol{R}_2$ and $\boldsymbol{R}_3$ are autocorrelation matrices given by

$$\begin{aligned} \boldsymbol{R}_1 &= E\left[\mu\boldsymbol{r}[i]b^*[i-1]\boldsymbol{r}^H[i]b^*[i-1]\right] \\ \boldsymbol{R}_2 &= E\left[\mu\boldsymbol{r}[i]b^*[i-2]\boldsymbol{r}^H[i]b^*[i-2]\right] \\ \boldsymbol{R}_3 &= E\left[\mu\boldsymbol{r}[i-1]b^*[i-2]\boldsymbol{r}^H[i-1]b^*[i-1]\right] \end{aligned} \tag{32}$$

and $\boldsymbol{F}_1$, $\boldsymbol{F}_2$ and $\boldsymbol{F}_3$ cross-time-instant correlation matrices, given by

$$\begin{aligned} \boldsymbol{F}_1 &= E\left[\mu\boldsymbol{r}[i]b^*[i-1]\boldsymbol{r}^H[i-1]b^*[i]\right] \\ \boldsymbol{F}_2 &= E\left[\mu\boldsymbol{r}[i]b^*[i-2]\boldsymbol{r}^H[i-2]b^*[i]\right] \\ \boldsymbol{F}_3 &= E\left[\mu\boldsymbol{r}[i-1]b^*[i-2]\boldsymbol{r}^H[i-2]b^*[i-1]\right]. \end{aligned} \tag{33}$$

Using (31) and the independence assumptions of $E\left[e_{o,n}[i]\boldsymbol{\varepsilon}[i]\right] = 0$ for $n = 1, 2, 3$, $E\left[\boldsymbol{r}^H[i]\boldsymbol{r}[i-1]\right] = 0$ and $E\left[b_k[i]b_k[i-1]\right] = 0$, we arrive at an expression for $\boldsymbol{K}[i]$

$$\begin{aligned} \boldsymbol{K}[i] = \;& \left[\boldsymbol{I} + \mu\boldsymbol{F}_1 - \mu\boldsymbol{R}_1 + \mu\boldsymbol{F}_2 - \mu\boldsymbol{R}_2 + \mu\boldsymbol{F}_3 - \mu\boldsymbol{R}_3\right]\boldsymbol{K}[i-1] \\ & \left[\boldsymbol{I} + \mu\boldsymbol{F}_1 - \mu\boldsymbol{R}_1 + \mu\boldsymbol{F}_2 - \mu\boldsymbol{R}_2 + \mu\boldsymbol{F}_3 - \mu\boldsymbol{R}_3\right] \\ & +\mu^2\boldsymbol{R}_1 J_{min,1}[i] + \mu^2\boldsymbol{R}_2 J_{min,2}[i] + \mu^2\boldsymbol{R}_1 J_{min,3}[i] \end{aligned} \tag{34}$$

where $J_{min,j}[i] = |e_{o,j}|^2$. Following a similar method, an expression for $\boldsymbol{G}[i]$ can also be reached

$$\boldsymbol{G}[i] = \boldsymbol{G}[i-1]\left[\mu\boldsymbol{F}_1 - \mu\boldsymbol{R}_1 + \mu\boldsymbol{F}_2 - \mu\boldsymbol{R}_2 + \mu\boldsymbol{F}_3 - \mu\boldsymbol{R}_3\right]. \tag{35}$$

At this point we study the derived expression to gain an insight into the operation of the bidirectional algorithm and the origins of its advantages over the conventional differential scheme. Equivalent expressions for the existing differential NLMS scheme are given by

$$\begin{aligned} \boldsymbol{K}[i] = \;& \left[\boldsymbol{I} + \mu\boldsymbol{F}_1 - \mu\boldsymbol{R}_1\right]\boldsymbol{K}[i-1]\left[\boldsymbol{I} + \mu\boldsymbol{F}_1 - \mu\boldsymbol{R}_1\right] \\ & +\mu^2\boldsymbol{R}_1 J_{min,1}[i] \\ \boldsymbol{G}[i] = \;& \boldsymbol{G}[i-1]\left[\mu\boldsymbol{F}_1 - \mu\boldsymbol{R}_1\right]. \end{aligned} \tag{36}$$

The bidirectional scheme has a number of additional correlation terms compared to the existing scheme. Evaluating the cross-time-instant matrices with regards to the independence assumptions yields

$$\begin{aligned} &\boldsymbol{F}_1 = |a_1|^2\boldsymbol{c}_1\boldsymbol{c}_1^H \underbrace{E\left[h[i]h^*[i-1]\right]}_{f_1[i]},\; \boldsymbol{F}_2 = |a_1|^2\boldsymbol{c}_1\boldsymbol{c}_1^H \underbrace{E\left[h[i]h^*[i-2]\right]}_{f_2[i]}, \\ &\text{and } \boldsymbol{F}_3 = |a_1|^2\boldsymbol{c}_1\boldsymbol{c}_1^H \underbrace{E\left[h[i-1]h^*[i-2]\right]}_{f_3[i]}. \end{aligned} \tag{37}$$

From the expression above it is clear that the underlying factor that governs the SINR performance of the algorithms is the correlation between the considered time instants, f_1, f_2, and f_3, and similarity between data-reuse and the use of f_1 and f_2. Accordingly, it is the additional correlation factors of the bidirectional algorithm that enhance its performance compared to the conventional techniques, confirming the initial motivation behind the proposition of the bidirectional approach. Lastly, the f_1, f_2, and f_3 expressions of (37) can be seen as the factors that influence the number of considered time instants.

Central to the performance of the bidirectional schemes are the correlation factors f_{1-3} and the related assumption of $h_1[i] \approx h_1[i-1]$. Examining the effect of the fading rate on the value of f_{1-3} shows that $f_1 \approx f_2 \approx f_3$ at fading rates of up to $T_s f_d = 0.01$, where $T_s f_d$ is the normalized fading parameter. Consequently, after a large number of received symbols with high total receive power

$$3\left[\boldsymbol{I} + \mu\boldsymbol{F}_1 - \mu\boldsymbol{R}_1\right] \approx \left[\boldsymbol{I} + \mu\boldsymbol{F}_1 - \mu\boldsymbol{R}_1 + \mu\boldsymbol{F}_2 - \mu\boldsymbol{R}_2 + \mu\boldsymbol{F}_3 - \mu\boldsymbol{R}_3\right], \tag{38}$$

due to the decreasing significance of the identity matrix. This indicates that the expected value of the SINR of the bidirectional scheme, once $f_1 \approx f_2 \approx f_3$ have stabilized, should be similar to the differential scheme. A second implication is that the bidirectional scheme should converge towards the MMSE level due to the equivalence between the bidirectional scheme and the MMSE solution. Fig. 1 illustrates the analytical performance using the above expressions.

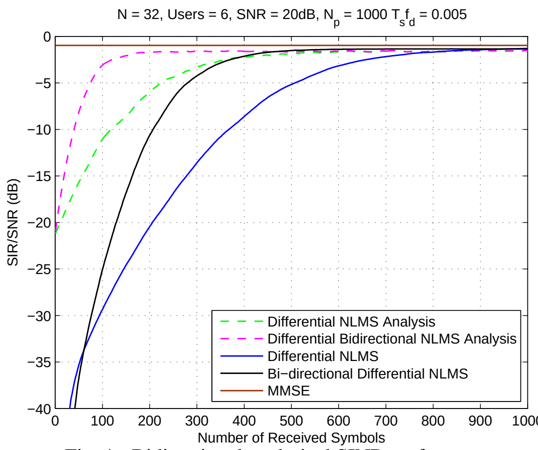

Fig. 1. Bidirectional analytical SINR performance.

The correlation matrices are calculated via ensemble averages prior to commencement of the algorithm and $\boldsymbol{G}[0] = \boldsymbol{K}[0] = \boldsymbol{I}$. In Fig. 1 one can see the convergence of the simulated schemes to the analytical and MMSE plots, validating the analysis. Due to the highly dynamic nature of the channel, using the expected values of the correlation matrix alone cannot capture the true transient performance of the algorithms. However, the convergence period of the analytical plots within the first 200 iterations can be considered to be within the coherence time and therefore give an indication of the transient performance relative to other analytical plots. Using this justification and the aforementioned analysis, it is clear that the advantages brought by the bidirectional scheme are predominantly in the transient phase due to the additional correlation information supplied by $\boldsymbol{F}_2$ and $\boldsymbol{F}_3$ and their analogy with data reuse algorithms. This observation is supported by the similar forms of the analytical and simulated results and their subsequent convergence. Note that these advantages are also verified for least-squares-based algorithms.

VI. SIMULATIONS

We apply the proposed bidirectional adaptive algorithms to interference suppression in the uplink of the DS-CDMA system described in Section II. The simulations employ multipath fading channels with $L = 3$ paths with relative powers equal to 0, -3 and -6 dB, Clarke's model [6] and are averaged over N_p packets and the parameters are specified in each plot. Conventional schemes use BPSK modulation and the differential schemes employ differential phase shift keying where the sequence of data symbols to be transmitted by user k are given by $b_k[i] = a_k[i]b_k[i-1]$ where $a_k[i]$ is the unmodulated data.

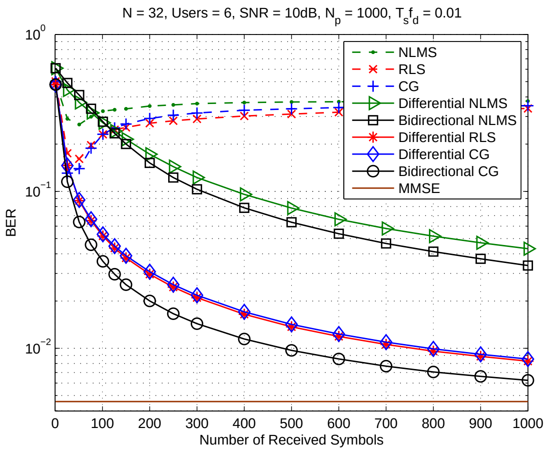

Fig. 2. BER performance comparison of proposed schemes during training.

The BER performance of existing and bidirectional schemes is illustrated in Fig. 2. The existing RLS and proposed CG algorithms converge to near the MMSE level with the bidirectional scheme providing a clear performance advantage. However, the NLMS schemes have a slower convergence performance due to their reduced adaptation rate compared to the CG algorithms.

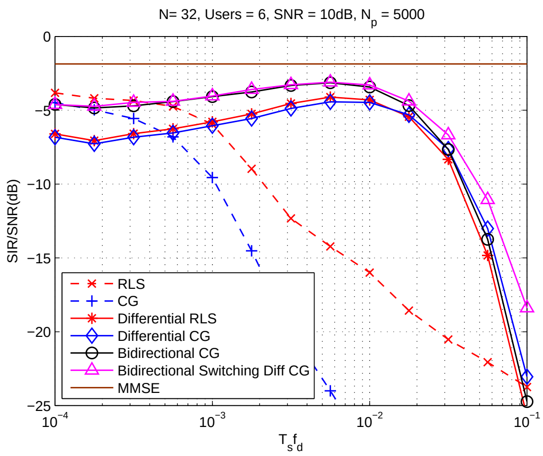

Fig. 3. SINR performance versus fading rate of the proposed CG schemes after 200 training symbols.

Fig. 3 illustrates the performance of the proposed CG and existing RLS algorithms as the fading rate is increased, where the SINR is normalized by the instantaneous SNR. The conventional schemes are unable to cope with fading rates in excess of $T_s f_d = 0.005$ and begin to diverge at the completion of the training sequence. The bidirectional scheme outperforms the differential schemes but the performance begins to decline once fading rates above $f_d T_s = 0.01$ are reached. The increase in performance of the bidirectional scheme can be accounted for by the increased correlation information supplied by the matrices $\boldsymbol{F}_2$ and $\boldsymbol{F}_3$ and effective data reuse. A second benefit of the bidirectional scheme is the improved performance at low fading rate. The introduction of the mixing factors into the bidirectional algorithm improves performance further, especially at higher fading rates. An explanation for this can be established by referring back to the observations on the correlation factors f_1, f_2 and f_3. Although fading rates of 0.01 may be fast, the assumption $h[i-2] \approx h[i-1] \approx h[i]$ is still valid. Consequently, $f_1 \approx f_2 \approx f_3$ and equal weighting is optimum. However, as the fading rate increases beyond $T_s f_d = 0.01$ this assumption breaks down and the correlation information requires unequal weighting for optimum performance, a task fulfilled by the adaptive mixing factors.

VII. CONCLUSIONS

We have presented bidirectional MMSE-based parameter estimation algorithms that exploit the time correlation of rapidly varying fading channels. The ratio between successive received vectors is tracked using correlation information gathered at adjacent time instants to avoid tracking of the faded or unfaded symbols. The results show that the proposed algorithms applied to interference suppression in DS-CDMA systems significantly outperform existing algorithms.

REFERENCES

[1] S. Verdu, *Multiuser Detection*. Cambridge University Press, NY, 1998.
[2] S. Haykin, *Adaptive Filter Theory*, 4th ed. NJ: Prentice Hall, 2002.
[3] B. Toplis and S. Pasupathy, "Tracking improvements in fast RLS algorithm using a variable forgetting factor," *IEEE Trans. Acoust., Speech, Signal Process.*, vol. 36, no. 2, pp. 206–227, February 1988.
[4] S. Gelfand, Y. Wei, and J. Krogmeier, "The stability of variable step-size LMS algorithms," *IEEE Trans. Sig. Proc.*, vol. 47, no. 12, pp. 3277–3288, December 1999.
[5] R. C. de Lamare and P. S. R. Diniz, "Set-Membership Adaptive Algorithms based on Time-Varying Error Bounds for CDMA Interference Suppression", *IEEE Trans. on Vehicular Technology*, vol. 58, no. 2, February 2009 , pp. 644 - 654.
[6] P. Clarke and R. C. de Lamare, "Low-Complexity Reduced-Rank Linear Interference Suppression based on Set-Membership Joint Iterative Optimization for DS-CDMA Systems", IEEE Trans. on Vehicular Technology, vol. 60, no. 9, 2011, pp. 4324-4337.
[7] M. L. Honig, S. L. Miller, M. J. Shensa, and L. B. Milstein, "Performance of adaptive linear interference suppression in the presence of dynamic fading," *IEEE Trans. Commun.*, vol. 49, no. 4, Apr. 2001.
[8] W. Chen and U. Mitra, "An Improved Blind Adaptive MMSE Receiver for Fast Fading DS-CDMA Channels", *IEEE Journal on Selected Areas of Communication*, vol. 19, no. 8, August 2001, pp. 1531–1543.
[9] U. Madhow, K. Bruvold, and L. J. Zhu, "Differential MMSE: A framework for robust adaptive interference suppression for DS-CDMA over fading channels," *IEEE Trans. Commun.*, vol. 53, no. 8, pp. 1377–1390, 2005.
[10] R. C. de Lamare and R. Sampaio-Neto, "Reduced-Rank Adaptive Filtering Based on Joint Iterative Optimization of Adaptive Filters", *IEEE Signal Processing Letters*, Vol. 14, no. 12, December 2007.
[11] R. C. de Lamare and R. Sampaio-Neto, "Reduced-Rank Space-Time Adaptive Interference Suppression With Joint Iterative Least Squares Algorithms for Spread-Spectrum Systems," *IEEE Trans. Veh. Technology*, vol.59, no.3, March 2010, pp.1217-1228.
[12] R. C. de Lamare and R. Sampaio-Neto, "Adaptive Reduced-Rank Processing Based on Joint and Iterative Interpolation, Decimation, and Filtering," *IEEE Trans. Sig. Proc.*, vol. 57, no. 7, July 2009, pp. 2503 - 2514.
[13] D. G. Luenburger and Y. Ye, *Linear and Nonliner Programming*, 3rd ed. Springer, 2008.
[14] H. Kushner, *Approximation and Weak Convergence Methods for Random Processes with Applications to Stochastic System Theory*, MIT Press, April 1984.