Set-Membership Binormalized Data-Reusing LMS Algorithms

Paulo S. R. Diniz, *Fellow, IEEE,* and Stefan Werner, *Member, IEEE*

***Abstract*—This paper presents and analyzes novel data selective normalized adaptive filtering algorithms with two data reuses. The algorithms [the set-membership binormalized LMS (SM-BNDRLMS) algorithms] are derived using the concept of set-membership filtering (SMF). These algorithms can be regarded as generalizations of the recently proposed set-membership NLMS (SM-NLMS) algorithm. They include two constraint sets in order to construct a space of feasible solutions for the coefficient updates. The algorithms include data-dependent step sizes that provide fast convergence and low-excess mean-squared error (MSE). Convergence analyzes in the mean squared sense are presented, and closed-form expressions are given for both white and colored input signals. Simulation results show good performance of the algorithms in terms of convergence speed, final misadjustment, and reduced computational complexity.**

***Index Terms*—Adaptive filter, data-selective, normalized data-reusing algorithms, set-membership filtering.**

I. Introduction

The least mean square (LMS) algorithm has gained popularity due to its robustness and low computational complexity. The main drawback of the LMS algorithm is that the convergence speed depends strongly on the eigenvalue spread of the input-signal correlation matrix [1]. To overcome this problem, a more complex recursive least squares (RLS) type of algorithm can be used. However, the faster convergence of the RLS algorithm does not imply a better tracking capability in a time-varying environment [1]. An alternative to speed up the convergence at the expense of low additional complexity is to use the binormalized data-reusing LMS (BNDRLMS) algorithm [2], [3]. The BNDRLMS algorithm, which uses consecutive data pairs in each update, has shown fast convergence for correlated input signals. However, the fast convergence comes at the expense of higher misadjustment because the algorithm utilizes the data even if it does not imply innovation. In order to combat the conflicting requirements of fast convergence and low misadjustment, the objective function of the adaptive algorithm needs to be changed. Set-membership filtering (SMF) [4] specifies a bound on the magnitude of the estimation error. The SMF uses the framework of set-membership identification (SMI) [5]–[8] to include a general filtering problem. Consequently, many of the existing optimal bounding ellipsoid (OBE) algorithms [5], [10]–[13] can be applied to the SMF framework.

Most, if not all, of the SMF algorithms feature reduced computational complexity primarily due to (sparse) *data-selective* updates. Implementation of those algorithms essentially involves two steps: 1) information evaluation (innovation check) and 2) update of parameter estimate. If the update does not occur frequently and the information evaluation does not involve much computational complexity, the overall complexity is usually much less than that of their RLS counterparts. It was shown in [9] that the class of adaptive solutions, called *set-membership adaptive recursive techniques* (SMART), include a particularly attractive OBE algorithm, which is referred to as the quasi-OBE algorithm or the bounding ellipsoidal adaptive constrained least-squares (BEACON) algorithm [13], [14], with a complexity of $O(N)$ for the innovation check. In addition, in [9], an algorithm with recursions similar to those of the NLMS algorithm with an adaptive step size was derived. The algorithm known as the set-membership NLMS (SM-NLMS) algorithm, which is further studied in [4], was shown to achieve both fast convergence and low misadjustment. Applications of SMF include adaptive equalization, where it allows the sharing of hardware resources in multichannel communications systems [14], adaptive multiuser detection in CDMA systems [15], [16], and in filtering with deterministic constraints on the output-error sequence [17].

The SM-NLMS algorithm only uses the current input-desired signals in its update. Following the same pattern as the conventional NLMS algorithm, the convergence of SM-NLMS algorithm will slow down when the input signal is colored. In order to overcome this problem, this paper proposes two versions of an algorithm that uses data pairs from two successive time instants in order to construct a set of feasible solutions for the update. The new algorithms are also data-selective algorithms, leading to a low computational complexity per update. In addition, for correlated input signals, they retain the fast convergence of the BNDRLMS algorithms related to the smart reuse of input-desired data pairs. The low misadjustment is obtained due to the data-selective updating utilized by the new algorithms. The idea of data reuse was also exploited in the context of OBE algorithms in [12].

The organization of the paper is as follows. Section II reviews the concept of SMF and the SM-NLMS algorithm of [4]. The new algorithms are derived in Section III. Section IV contains analysis of the algorithms in the mean-squared sense, followed by simulations in Section V. Section VI contains the concluding remarks.

Manuscript received December 6, 2000; revised August 23, 2002. The associate editor coordinating the review of this paper and approving it for publication was Dr. Steven T. Smith.

P. S. R. Diniz is with the COPPE/Poli/Federal University of Rio de Janeiro, Rio de Janeiro, Brazil.

S. Werner is with the Signal Processing Laboratory, Helsinki University of Technology, Helsinki, Finland

Digital Object Identifier 10.1109/TSP.2002.806562

II. SMF

In SMF, the filter $\mathbf{w}$ is designed to achieve a specified bound on the magnitude of the output error. Assuming a sequence of input vectors $\{\mathbf{x}_k\}_{k=1}^{\infty}$ and a desired signal sequence $\{d_k\}_{k=1}^{\infty}$, we can write the sequence of estimation errors $\{e_k\}_{k=1}^{\infty}$ as

$$e_k = d_k - \mathbf{w}^T\mathbf{x}_k \tag{1}$$

where $\mathbf{x}_k$ and $\mathbf{w} \in \mathbb{R}^N$ with d_k and $e_k \in \mathbb{R}$. For a properly chosen bound γ on the estimation error, there are several valid estimates of $\mathbf{w}$.

Let $\mathcal{S}$ denote the set of all possible input-desired data pairs $(\mathbf{x}, d)$ of interest. Next, let Θ denote the set of all possible vectors $\mathbf{w}$ that result in an output error bounded by γ whenever $(\mathbf{x}, d) \in \mathcal{S}$. The set Θ, which is referred to as the *feasibility set*, is given by

$$\Theta = \bigcap_{(\mathbf{x},d)\in\mathcal{S}} \{\mathbf{w} \in \mathbb{R}^N : |d - \mathbf{w}^T\mathbf{x}| \le \gamma\}. \tag{2}$$

Assume that the adaptive filter is trained with k input-desired data pairs $\{\mathbf{x}_i, d_i\}_{i=1}^{k}$. Let $\mathcal{H}_k$ denote the set containing all vectors $\mathbf{w}$ for which the associated output error at time instant k is upper bounded in magnitude by γ. In other words

$$\mathcal{H}_k = \{\mathbf{w} \in \mathbb{R}^N : |d_k - \mathbf{w}^T\mathbf{x}_k| \le \gamma\}. \tag{3}$$

The set $\mathcal{H}_k$ is referred to as the *constraint set*, and its boundaries are hyperplanes. Finally, define the *exact membership set* ψ_k to be the intersection of the constraint sets over the time instants $i = 1, \ldots, k$, i.e.,

$$\psi_k = \bigcap_{i=1}^{k} \mathcal{H}_i. \tag{4}$$

It can be seen that the *feasibility set* Θ is a subset of the *exact membership set* ψ_k at any given time instant. The *feasibility set* is also the *limiting set* of the *exact membership set*, i.e., the two set will be equal if the training signal traverses all signal pairs belonging to $\mathcal{S}$.

The idea of SMART is to adaptively find an estimate that belongs to the feasibility set. One approach is to apply one of the many OBE algorithms, which tries to approximate the exact membership set ψ_k with ellipsoids. Another adaptive approach is to compute a point estimate through projections using, for example, the information provided by the constraint set $\mathcal{H}_k$, like in the set-membership NLMS (SM-NLMS) algorithm considered in the following subsection. It was also shown in [4] that the SM-NLMS algorithm can be associated with an optimal bounding spheroid (OBS).

A. Set-Membership Normalized LMS (SM-NLMS) Algorithm

The set-membership NLMS (SM-NLMS) algorithm derived in [4] is similar to the conventional NLMS algorithm in form. However, the philosophy behind the SM-NLMS algorithm derivation differs from that of the NLMS algorithm. The basic idea behind the algorithm is that if the previous estimate $\mathbf{w}_k$ lies outside the constraint set $\mathcal{H}_k$, i.e., $|d_k - \mathbf{w}_k^T\mathbf{x}_k| > \gamma$, the new

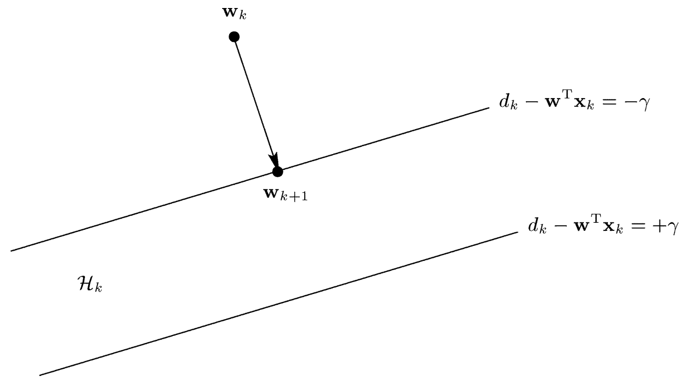

Fig. 1. SM-NLMS algorithm.

estimate $\mathbf{w}_{k+1}$ will lie on the closest boundary of $\mathcal{H}_k$ at a minimum distance, i.e., the SM-NLMS minimizes $\|\mathbf{w}_{k+1} - \mathbf{w}_k\|^2$ subject to $\mathbf{w}_{k+1} \in \mathcal{H}_k$. This is obtained by an orthogonal projection of the previous estimate onto the closest boundary of $\mathcal{H}_k$. A graphical visualization of the updating procedure of the SM-NLMS can be found in Fig. 1. Straightforward calculation leads to the following recursions for $\mathbf{w}_k$:

$$\mathbf{w}_{k+1} = \mathbf{w}_k + \alpha_k \frac{e_k\mathbf{x}_k}{\|\mathbf{x}_k\|^2} \tag{5}$$

with

$$\begin{aligned} e_k &= d_k - \mathbf{w}_k^T\mathbf{x}_k \\ \alpha_k &= \begin{cases} 1 - \frac{\gamma}{|e_k|}, & \text{if } |e_k| > \gamma \\ 0, & \text{otherwise} \end{cases} \end{aligned} \tag{6}$$

where e_k and α_k denote the *a priori* error and the time-dependent step-size, respectively. The update (5) and (6) resemble those of the conventional NLMS algorithm, except for the time-varying step-size α_k.

Note that since the conventional NLMS algorithm minimizes $\|\mathbf{w}_{k+1} - \mathbf{w}_k\|^2$ subject to the constraint that $\mathbf{w}_{k+1}^T\mathbf{x} = d_k$, it is a particular case of the above algorithm by choosing the bound $\gamma = 0$. Furthermore, using a step-size $\alpha_k = 1$ in the SM-NLMS whenever $\mathbf{w}_k \notin \mathcal{H}_k$ would result in a valid update because the hyperplane with zero *a posteriori* error lies in $\mathcal{H}_k$; however, the resulting algorithm does not minimize the Euclidean distance.

III. Set-Membership Binormalized Data-Reusing LMS Algorithms

The SM-NLMS algorithm in the previous subsection only considered the constraint set $\mathcal{H}_k$ in its update. The SM-NLMS algorithm has a low computational complexity per update, but its convergence speed appears to follow the trend of the normalized LMS algorithm, which depends on the eigenvalue spread of the input-signal correlation matrix. The exact membership set ψ_k defined in (4) suggests the use of more than one constraint set. In this subsection, two algorithms are derived, requiring that the solution belongs to the constraint sets at time instants k and $k-1$, i.e., $\mathbf{w}_{k+1} \in \mathcal{H}_k \cap \mathcal{H}_{k-1}$. The recursions of the algorithms are similar to the conventional BNDRLMS algorithm [2]. The set-membership binormalized data-reusing LMS (SM-BNDRLMS) algorithms can be seen as extensions of the SM-NLMS algorithm that use two consecutive constraint sets

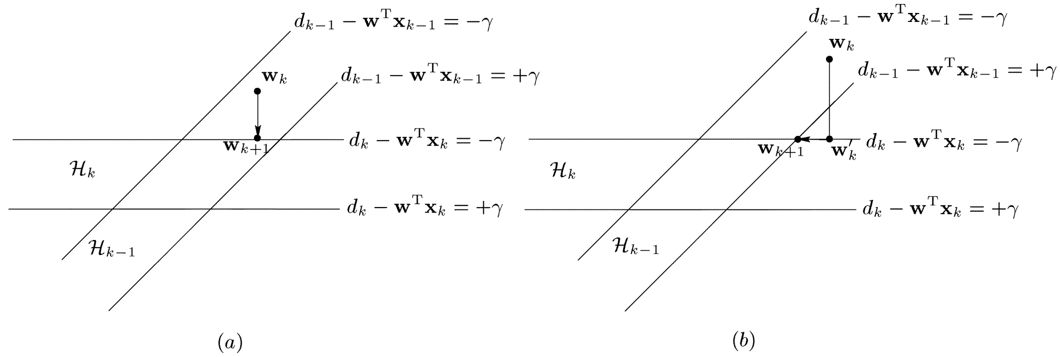

Fig. 2. SM-BNDRLMS-I algorithm. (a) Orthogonal projection onto the nearest boundary of $\mathcal{H}_k$ lies within $\mathcal{H}_{k-1}$, i.e., $\mathbf{w}'_k \in \mathcal{H}_{k-1}$. No further update. (b) Orthogonal projection onto the nearest boundary of $\mathcal{H}_k$, $\mathbf{w}'_k$ lies outside $\mathcal{H}_{k-1}$. Final solution at the nearest intersection of $\mathcal{H}_k$ and $\mathcal{H}_{k-1}$.

for each update. The first algorithm presented in Section III-A is a two-step approach minimizing the Euclidean distance between the old filter coefficients and the new update subjected to the constraints that the new update lies in both constraint sets $\mathcal{H}_k$ and $\mathcal{H}_{k-1}$. The second algorithm presented in Section III-B reduces the computational complexity per update, as compared with the first algorithm by choosing a different update strategy.

A. Algorithm I

The first set-membership binormalized data-reusing LMS algorithm (SM-BNDRLMS-I) performs an initial normalized step according to the SM-NLMS algorithm. If the solution to the first step belongs to both constraint sets $\mathcal{H}_k$ and $\mathcal{H}_{k-1}$, no further update is required. If the initial step moves the solution out of $\mathcal{H}_{k-1}$, a second step is taken such that the solution is at the intersection of $\mathcal{H}_k$ and $\mathcal{H}_{k-1}$ at a minimum distance from $\mathbf{w}_k$. Fig. 2 depicts the update procedure. The SM-BNDRLMS-I algorithm minimizes $\|\mathbf{w}_{k+1} - \mathbf{w}_k\|^2$ subject to the constraint that $\mathbf{w}_{k+1} \in \mathcal{H}_k \cap \mathcal{H}_{k-1}$.

The solution can be obtained by first performing an orthogonal projection of $\mathbf{w}_k$ onto the nearest boundary of $\mathcal{H}_k$, just like in the SM-NLMS algorithm

$$\mathbf{w}'_k = \mathbf{w}_k + \alpha_k \frac{e_k \mathbf{x}_k}{\|\mathbf{x}_k\|^2} \tag{7}$$

where α_k and e_k are defined in (6). If $\mathbf{w}'_k \in \mathcal{H}_{k-1}$, i.e., $|d_{k-1} - \mathbf{w}'^{T}_k \mathbf{x}_{k-1}| \le \gamma$, then $\mathbf{w}_{k+1} = \mathbf{w}'_k$. Otherwise, a second step is taken such that the solution lies at the intersection of $\mathcal{H}_k$ and $\mathcal{H}_{k-1}$ at a minimum distance. The second step in the algorithm will be in the direction of $\mathbf{x}_k^{\perp}$, which is orthogonal to the first step, i.e.,

$$\mathbf{w}_{k+1} = \mathbf{w}'_k + \beta_k \frac{e'_{k-1} \mathbf{x}_k^{\perp}}{\|\mathbf{x}_k^{\perp}\|^2} \tag{8}$$

where

$$\begin{aligned} \mathbf{x}_k^{\perp} &= \left(\mathbf{I} - \frac{\mathbf{x}_k \mathbf{x}_k^T}{\|\mathbf{x}_k\|^2}\right)\mathbf{x}_{k-1} \\ e'_{k-1} &= d_{k-1} - \mathbf{w}'^{T}_k \mathbf{x}_{k-1} \\ \beta_k &= 1 - \frac{\gamma}{|e'_{k-1}|}. \end{aligned} \tag{9}$$

In summary, the recursive algorithm for $\mathbf{w}_k$ is given by

$$\begin{aligned} \mathbf{w}'_k &= \mathbf{w}_k + \alpha_k \frac{e_k \mathbf{x}_k}{\|\mathbf{x}_k\|^2} \\ \mathbf{w}_{k+1} &= \mathbf{w}'_k + \lambda_1 \mathbf{x}_k + \lambda_2 \mathbf{x}_{k-1} \end{aligned} \tag{10}$$

where

$$\begin{aligned} e_k &= d_k - \mathbf{w}_k^T \mathbf{x}_k \\ e'_{k-1} &= d_{k-1} - \mathbf{w}'^{T}_k \mathbf{x}_{k-1} \\ \lambda_1 &= -\frac{\beta_k e'_{k-1} \mathbf{x}_{k-1}^T \mathbf{x}_k}{\|\mathbf{x}_k\|^2 \|\mathbf{x}_{k-1}\|^2 - [\mathbf{x}_{k-1}^T \mathbf{x}_k]^2} \\ \lambda_2 &= \frac{\beta_k e'_{k-1} \|\mathbf{x}_k\|^2}{\|\mathbf{x}_k\|^2 \|\mathbf{x}_{k-1}\|^2 - [\mathbf{x}_{k-1}^T \mathbf{x}_k]^2} \\ \alpha_k &= \begin{cases} 1 - \frac{\gamma}{|e_k|}, & \text{if } |e_k| > \gamma \\ 0, & \text{otherwise} \end{cases} \\ \beta_k &= \begin{cases} 1 - \frac{\gamma}{|e'_{k-1}|}, & \text{if } |e_k| > \gamma \text{ and } |e'_{k-1}| > \gamma \\ 0 & \text{otherwise.} \end{cases} \end{aligned} \tag{11}$$

Remark 1: If the constraint sets $\mathcal{H}_k$ and $\mathcal{H}_{k-1}$ are parallel, the denominator term of the λ_is in (11) will be zero. In this particular case, the second step of (10) is not performed to avoid division by zero.

It is easy to verify that if the bound of the estimation error is chosen to be zero, i.e., $\gamma = 0$, the update equations will be those of the conventional BNDRLMS algorithm with unity step-size [2].

B. Algorithm II

The SM-BNDRLMS-I algorithm in the previous subsection requires the intermediate check, that is, if $\mathbf{w}'_k \in \mathcal{H}_k$, to determine if a second step is needed. This check will add extra computational complexity. The algorithm proposed below (the SM-BNDRLMS-II) does not require this additional check to assure that $\mathbf{w}_{k+1} \in \mathcal{H}_k \cap \mathcal{H}_{k-1}$. Let $\mathcal{S}_{k-i+1}$ $(i = 1, 2)$ denote the hyperplanes that contain all vectors $\mathbf{w}$ such that $d_{k-i+1} - \mathbf{w}^T \mathbf{x}_{k-i+1} = g_{k-i+1}$, where g_{k-i+1} are extra variables chosen such that the bound constraints are valid. That is, if g_{k-i+1} are chosen such that $|g_{k-i+1}| \le \gamma$, then $\mathcal{S}_{k-i+1} \in \mathcal{H}_{k-i+1}$.

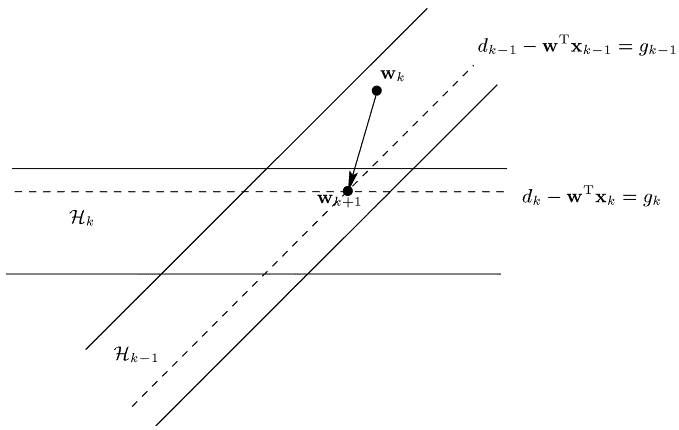

Fig. 3. General algorithm update.

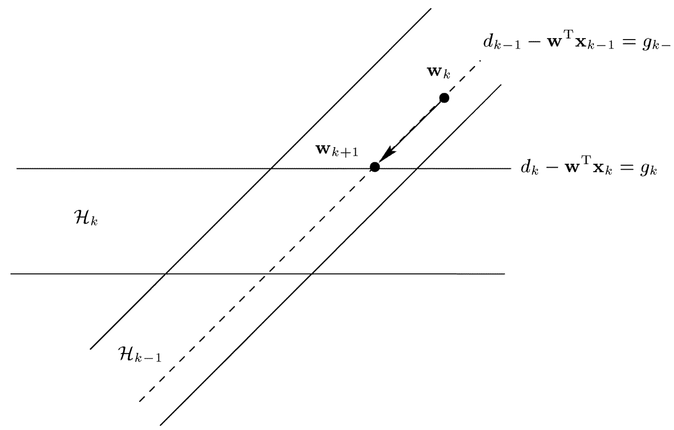

Fig. 4. SM-BNDRLMS-II algorithm.

Consider the following optimization criterion whenever $\mathbf{w}_k \notin \mathcal{H}_k \cap \mathcal{H}_{k-1}$:

$$\begin{aligned} \min \|\mathbf{w}_{k+1} - \mathbf{w}_k\|^2 \ \text{subject to} \\ d_k - \mathbf{x}_k^T \mathbf{w}_{k+1} &= g_k \\ d_{k-1} - \mathbf{x}_{k-1}^T \mathbf{w}_{k+1} &= g_{k-1}. \end{aligned} \tag{12}$$

The pair (g_k, g_{k-1}) specifies the point in $\mathcal{H}_k \cap \mathcal{H}_{k-1}$ where the final update will lie; see Fig. 3. In order to evaluate if an update according to (12) is required, we need to first check if $\mathbf{w}_k \in \mathcal{H}_k \cap \mathcal{H}_{k-1}$. Due to the concept of data reuse together with the constraint $|g_{k-i+1}| \leq \gamma$, this check reduces to $\mathbf{w}_k \in \mathcal{H}_k$. In what follows, we first solve for the general update and thereafter consider a specific choice of the pair (g_k, g_{k-1}), leading to a simplified form.

To solve the optimization problem in (12), we can apply the method of Lagrange multipliers leading to the following objective function:

$$\begin{aligned} f(\mathbf{w}_{k+1}) = \|\mathbf{w}_{k+1} - \mathbf{w}_k\|^2 + \lambda_1[d_k - \mathbf{x}_k^T \mathbf{w}_{k+1} - g_k] \\ + \lambda_2[d_{k-1} - \mathbf{x}_{k-1}^T \mathbf{w}_{k+1} - g_{k-1}]. \end{aligned} \tag{13}$$

After setting the gradient of (13) to zero and solving for the Lagrange multipliers, we get

$$\mathbf{w}_{k+1} = \begin{cases} \mathbf{w}_k + \frac{\lambda_1}{2}\mathbf{x}_k + \frac{\lambda_2}{2}\mathbf{x}_{k-1}, & \text{if } |e_k| > \gamma \\ \mathbf{w}_k, & \text{otherwise} \end{cases} \tag{14}$$

where

$$\frac{\lambda_1}{2} = \frac{[e_k - g_k]\,\|\mathbf{x}_{k-1}\|^2 - [\epsilon_{k-1} - g_{k-1}]\,\mathbf{x}_k^T \mathbf{x}_{k-1}}{\|\mathbf{x}_k\|^2 \|\mathbf{x}_{k-1}\|^2 - \left[\mathbf{x}_{k-1}^T \mathbf{x}_k\right]^2} \tag{15}$$

$$\frac{\lambda_2}{2} = \frac{[\epsilon_{k-1} - g_{k-1}]\,\|\mathbf{x}_k\|^2 - [e_k - g_k]\,\mathbf{x}_{k-1}^T \mathbf{x}_k}{\|\mathbf{x}_k\|^2 \|\mathbf{x}_{k-1}\|^2 - \left[\mathbf{x}_{k-1}^T \mathbf{x}_k\right]^2} \tag{16}$$

in which $e_k = d_k - \mathbf{w}_k^T \mathbf{x}_k$ and $\epsilon_{k-1} = d_{k-1} - \mathbf{w}_k^T \mathbf{x}_{k-1}$ are the *a priori* error at iteration k and the *a posteriori* error at iteration $k-1$, respectively.

Since $\mathbf{w}_k$ always belongs to $\mathcal{H}_{k-1}$ before a possible update, we have $\epsilon_{k-1} \leq \gamma$. Therefore, choosing $g_{k-1} = \epsilon_{k-1}$ satisfies $|g_{k-1}| \leq \gamma$. In the same way as in the SM-NLMS and SM-BNDRLMS-I algorithms, it is sufficient to choose g_k such that the update lies on the closest boundary of $\mathcal{H}_k$, i.e., $g_k = \gamma \operatorname{sign}(e_k)$.

TABLE I
COMPUTATIONAL COMPLEXITY PER UPDATE.

ALG.	MULT.	ADD.	DIV.
SM-NLMS	$3N+1$	$3N$	2
SM-BNDRLMS-I (1 step)	$4N+1$	$4N$	2
SM-BNDRLMS-I (2 steps)	$7N+8$	$7N+3$	4
SM-BNDRLMS-II	$5N+7$	$5N+3$	2

The above choices lead to the SM-BNDRLMS-II algorithm, where the new estimate $\mathbf{w}_{k+1}$ will lie at the nearest boundary of $\mathcal{H}_k$ such that the *a posteriori* error at iteration $k-1$, ϵ_{k-1} is kept constant. A graphical illustration of the update procedure is shown in Fig. 4. The update equations for the SM-BNDRLMS-II algorithm are given by

$$\mathbf{w}_{k+1} = \mathbf{w}_k + \frac{\lambda_1'}{2}\mathbf{x}_k + \frac{\lambda_2'}{2}\mathbf{x}_{k-1} \tag{17}$$

where

$$\begin{aligned} \frac{\lambda_1'}{2} &= \frac{\alpha_k e_k \|\mathbf{x}_{k-1}\|^2}{\|\mathbf{x}_k\|^2 \|\mathbf{x}_{k-1}\|^2 - [\mathbf{x}_{k-1}^T \mathbf{x}_k]^2} \\ \frac{\lambda_2'}{2} &= -\frac{\alpha_k e_k \mathbf{x}_{k-1}^T \mathbf{x}_k}{\|\mathbf{x}_k\|^2 \|\mathbf{x}_{k-1}\|^2 - [\mathbf{x}_{k-1}^T \mathbf{x}_k]^2} \\ \alpha_k &= \begin{cases} 1 - \frac{\gamma}{|e_k|}, & \text{if } |e_k| > \gamma \\ 0, & \text{otherwise.} \end{cases} \end{aligned} \tag{18}$$

As with the SM-BNDRLMS-I algorithm in the previous subsection, the problem with parallel constraint sets is avoided by using the SM-NLMS update of (5) whenever the denominator in the λ_i' is zero.

C. Computational Complexity

The computational complexity per update in terms of the number of additions, multiplications, and divisions for the three algorithms are shown in Table I. For the SM-BNDRLMS-I, the two possible update complexities are listed where the first corresponds to the total complexity when only the first step is necessary, i.e., when $\mathbf{w}_k' \in \mathcal{H}_{k-1}$, and the second corresponds to the total complexity when a full update is needed. Applying the SM-BNDRLMS algorithms slightly increases the computational complexity as compared with that of the SM-NLMS algorithm. However, the SM-BNDRLMS algorithms have a reduced number of updates and an increased convergence rate as compared to the SM-NLMS algorithm, as verified through simulations in Section V. Comparing the complexities of the SM-BNDRLMS-I and SM-BNDRLMS-II algorithms, we note that the difference in the overall complexity depends on the frequency the second step is required in Algorithm I. In the operation counts, the value of $\|\mathbf{x}_{k-1}\|^2$ at iteration k was assumed unknown. However, once $\|\mathbf{x}_k\|^2$ or $\|\mathbf{x}_{k-1}\|^2$ is known, one can compute the other using only two additional multiplications, e.g., $\|\mathbf{x}_{k-1}\|^2 = \|\mathbf{x}_k\|^2 - x_k^2 + x_{k-N}^2$. The relation between $\|\mathbf{x}_{k-1}\|^2$ and $\|\mathbf{x}_k\|^2$ has been used in the operation counts of the SM-BNDRLMS algorithms. If update occurs at two successive time instants, $\|\mathbf{x}_{k-1}\|^2$ and $\mathbf{x}_{k-1}^T\mathbf{x}_{k-2}$ are known from a previous update, and as a consequence, the number of multiplications and additions in such updates can be further reduced by approximately N for the SM-NLMS algorithm and $2N$ for the SM-BNDRLMS algorithms. Finally, note that if we continuously estimate $\|\mathbf{x}_k\|^2$ and $\mathbf{x}_k^T\mathbf{x}_{k-1}$, regardless of whether an update is required or not, the SM-BNDRLMS-II algorithm will always be more efficient than SM-BNDRLMS-I. These computational savings are crucial in applications where the filter order is high and computational resources are limited.

IV. Second-Order Statistical Analysis

This section addresses the steady-state analysis of the SM-BNDRLMS algorithms.

A. Coefficient-Error Vector

In this subsection, we investigate the convergence behavior of the coefficient vector $\mathbf{w}_k$. It is assumed that an unknown FIR $\mathbf{w}_o$ is identified with an adaptive filter $\mathbf{w}_k$ of the same order $N-1$ using the SM-BNDRLMS-II algorithm. The desired response is given by

$$d_k = \mathbf{x}_k^T\mathbf{w}_o + n_k \tag{19}$$

where n_k is measurement noise, which is assumed here to be Gaussian with zero mean and variance σ_n^2. We study the evolution of the coefficient error $\Delta\mathbf{w}_k = \mathbf{w}_k - \mathbf{w}_o$. The output error can now be written as

$$e_k = n_k - \mathbf{x}_k^T\Delta\mathbf{w}_k. \tag{20}$$

The update equations for the adaptive filter coefficients are given by

$$\mathbf{w}_{k+1} = \begin{cases} \mathbf{w}_k, & \text{if } |e_k| \le \gamma \\ \mathbf{w}_k + (e_k - \gamma)\,\mathbf{a}, & \text{if } e_k > +\gamma \\ \mathbf{w}_k + (e_k + \gamma)\,\mathbf{a}, & \text{if } e_k < -\gamma \end{cases} \tag{21}$$

where

$$\mathbf{a} = \frac{\|\mathbf{x}_{k-1}\|^2\mathbf{x}_k - \left(\mathbf{x}_{k-1}^T\mathbf{x}_k\right)\mathbf{x}_{k-1}}{\|\mathbf{x}_k\|^2\|\mathbf{x}_{k-1}\|^2 - \left(\mathbf{x}_k^T\mathbf{x}_{k-1}\right)^2}. \tag{22}$$

As a consequence, the coefficient error at time instant $k+1$ becomes

$$\Delta\mathbf{w}_{k+1} = \begin{cases} \Delta\mathbf{w}_k, & \text{if } |e_k| \le \gamma \\ [\mathbf{I}+\mathbf{A}]\,\Delta\mathbf{w}_k + \mathbf{b} - \mathbf{c}, & \text{if } e_k > +\gamma \\ [\mathbf{I}+\mathbf{A}]\,\Delta\mathbf{w}_k + \mathbf{b} + \mathbf{c}, & \text{if } e_k > -\gamma \end{cases} \tag{23}$$

where

$$\mathbf{A} = \frac{\mathbf{x}_{k-1}\mathbf{x}_{k-1}^T\mathbf{x}_k\mathbf{x}_k^T - \|\mathbf{x}_{k-1}\|^2\mathbf{x}_k\mathbf{x}_k^T}{\|\mathbf{x}_k\|^2\|\mathbf{x}_{k-1}\|^2 - \left(\mathbf{x}_k^T\mathbf{x}_{k-1}\right)^2} \tag{24}$$

and

$$\begin{aligned} \mathbf{b} &= n_k\mathbf{a} \\ \mathbf{c} &= \gamma\mathbf{a}. \end{aligned} \tag{25}$$

In the analysis, we utilize the following initial assumptions.

AS1) The filter is updated with the probability $P_{e,k} = P[|e_k| > \gamma]$, and $P[e_k > \gamma] = P[e_k < -\gamma]$.

Note that the probability $P_{e,k}$ will be time-varying because the variance of the output error e_k depends on the mean of the squared coefficient-error vector norm, and for Gaussian noise with zero mean and variance σ_n^2, we get $\sigma_e^2 = \sigma_n^2 + \mathrm{E}\left[\Delta\mathbf{w}_k^T\mathbf{R}\Delta\mathbf{w}_k\right]$. Since we are interested in the excess MSE and not the initial transient the following assumption is made.

AS2) The filter has reached the steady-state value.

From (23), we can now write the coefficient error as

$$\Delta\mathbf{w}_{k+1} = \Delta\mathbf{w}_k + P_{e,k}\left(\mathbf{A}\Delta\mathbf{w}_k + \mathbf{b}\right). \tag{26}$$

B. Input-Signal Model

In the evaluation of the excess MSE we use a simplified model for the input-signal vector $\mathbf{x}_k$. The model uses a simplified distribution for the input-signal vector by employing reduced and countable angular orientations for the excitation, which are consistent with the first- and second-order statistics of the actual input-signal vector. The model was used for analyzing the NLMS algorithm, [18] as well as the BNDRLMS algorithm [2], and was shown to yield good results.

The input signal vector for the model is

$$\mathbf{x}_k = s_k r_k \mathbf{v}_k \tag{27}$$

where

- s_k is ± 1 with probability 1/2
- r_k^2 has the same probability distribution as $\|\mathbf{x}_k\|^2$, and in the case of white Gaussian input signal, it is a sample of an independent process with χ-square distribution with N degrees of freedom, with $\mathrm{E}\left[r_k^2\right] = N\sigma_x^2$
- $\mathbf{v}_k$ is one of the N orthonormal eigenvectors $\mathbf{R} = \mathrm{E}\left[\mathbf{x}_k\mathbf{x}_k^T\right]$, say $\{\mathcal{V}_i,\, i = 1,\, \ldots\, N\}$.

For a white Gaussian input signal, it is assumed that $\mathbf{v}_k$ is uniformly distributed such that

$$P\left(\mathbf{v}_k = \mathcal{V}_i\right) = \frac{1}{N} \tag{28}$$

C. Excess MSE for White Input Signals

In this subsection, we investigate the excess MSE in the SM-BNDRLMS algorithms. In order to achieve this goal, we have to consider a simple model for the input signal vector that assumes a discrete set of angular orientations. The excess MSE is given by [1]

$$\xi_{exc} = \lim_{k\to\infty} \xi_k - \xi_{min} \tag{29}$$

where

$$\xi_k = \mathrm{E}\left[e_k^2\right] = \mathrm{E}\left[\left(n_k - \mathbf{x}_k^T \Delta\mathbf{w}_k\right)^2\right] \tag{30}$$

is the MSE at iteration k, and ξ_{min} is the minimum MSE. With these equations, we have that

$$\begin{aligned}\Delta\xi_k =& \mathrm{E}\left[\left(n_k - \mathbf{x}_k^T\Delta\mathbf{w}_k\right)^2\right] - \xi_{\min} = \mathrm{E}\left[\Delta\mathbf{w}_k^T \mathbf{R} \Delta\mathbf{w}_k\right] \\ =& \mathrm{tr}\left\{\mathbf{R}\mathrm{cov}\left[\Delta\mathbf{w}_k\right]\right\}.\end{aligned} \tag{31}$$

For the input-signal model presented in the previous subsection, $\Delta\xi_{k+1}$ can be written as

$$\begin{aligned}\Delta\xi_{k+1} = \Delta\xi_{k+1}|_{\mathbf{x}_k \| \mathbf{x}_{k-1}} \times P\left[\mathbf{x}_k \| \mathbf{x}_{k-1}\right] \\ + \Delta\xi_{k+1}|_{\mathbf{x}_k \perp \mathbf{x}_{k-1}} \times P\left[\mathbf{x}_k \perp \mathbf{x}_{k-1}\right].\end{aligned} \tag{32}$$

Conditions $\mathbf{x}_k \| \mathbf{x}_{k-1}$ and $\mathbf{x}_k \perp \mathbf{x}_{k-1}$ in the model are equivalent to $\mathbf{v}_k = \mathbf{v}_{k-1}$ and $\mathbf{v}_k \neq \mathbf{v}_{k-1}$, respectively, because $\mathbf{v}_k$ and $\mathbf{v}_{k-1}$ can only be parallel or orthogonal to each other. $P\left[\mathbf{x}_k \| \mathbf{x}_{k-1}\right]$ denotes the probability that $\mathbf{x}_k \| \mathbf{x}_{k-1}$, and $P\left[\mathbf{x}_k \perp \mathbf{x}_{k-1}\right]$ denotes the probability that $\mathbf{x}_k \perp \mathbf{x}_{k-1}$. For the case $\mathbf{x}_k \| \mathbf{x}_{k-1}$, the SM-BNDRLMS algorithm will behave like the SM-NLMS algorithm, which has the excess MSE (see Appendix A)

$$\begin{aligned}\Delta\xi_{k+1}|_{\mathbf{x}_k \| \mathbf{x}_{k-1}} = \left(1 - \frac{2P_{e,k} - P_{e,k}^2}{N}\right)\Delta\xi_k \\ + P_{e,k}^2 \frac{\sigma_n^2}{N+1-\nu_x}\end{aligned} \tag{33}$$

where $\nu_x = \mathrm{E}[x_k^4/\sigma_x^4]$ varies from 1 for binary distribution, to 3 for Gaussian distribution, to ∞ for a Cauchy distribution [3], [18]. For the case $\mathbf{x}_k \perp \mathbf{x}_{k-1}$, the expression for the coefficient error vector also reduces to the same as that of the SM-NLMS algorithm (see Appendix B), giving

$$\begin{aligned}\Delta\xi_{k+1}|_{\mathbf{x}_k \perp \mathbf{x}_{k-1}} = \left(1 - \frac{2P_{e,k} - P_{e,k}^2}{N}\right)\Delta\xi_k \\ + P_{e,k}^2 \frac{\sigma_n^2}{N+1-\nu_x}.\end{aligned} \tag{34}$$

Combining, we have

$$\begin{aligned}\Delta\xi_{k+1} =& \Delta\xi_{k+1}|_{\mathbf{x}_k \| \mathbf{x}_{k-1}} \times P\left[\mathbf{x}_k \| \mathbf{x}_{k-1}\right] \\ &+ \Delta\xi_{k+1}|_{\mathbf{x}_k \perp \mathbf{x}_{k-1}} \times P\left[\mathbf{x}_k \perp \mathbf{x}_{k-1}\right] \\ =& \left(P\left[\mathbf{x}_k \| \mathbf{x}_{k-1}\right] + P\left[\mathbf{x}_k \perp \mathbf{x}_{k-1}\right]\right)\Delta\xi_k^{\|} \\ =& \left(P\left[\mathbf{x}_k \| \mathbf{x}_{k-1}\right] + P\left[\mathbf{x}_k \perp \mathbf{x}_{k-1}\right]\right)\Delta\xi_k^{\perp} \\ =& \left(1 - \frac{2P_{e,k} - P_{e,k}^2}{N}\right)\Delta\xi_k + P_{e,k}^2 \frac{\sigma_n^2}{N+1-\nu_x}.\end{aligned} \tag{35}$$

Recall assumption *AS2)*, where the filter is in steady-state such that the probability $P_{e,k} \rightarrow P_e$ is constant. The stability and convergence of (35) holds since $0 \leq P_{e,k} \leq 1$. If we let $k \rightarrow \infty$, the excess MSE becomes

$$\xi_{exc} = \frac{N}{N+1-\nu_x} \cdot \frac{P_e \sigma_n^2}{2 - P_e}. \tag{36}$$

Assuming the filter has converged to its steady-state value, the probability of update for white Gaussian input signals is given by

$$P_e = 2Q\left(\frac{\gamma}{\sqrt{\sigma_n^2 + \sigma_x^2 \mathrm{E}\left[\|\Delta\mathbf{w}_\infty\|^2\right]}}\right) \tag{37}$$

where $Q(\cdot)$ is the complementary Gaussian cumulative distribution function given by

$$Q(x) = \int_x^\infty \frac{1}{\sqrt{2\pi}} e^{-t^2/2} dt \tag{38}$$

and $\mathrm{E}\left[\|\Delta\mathbf{w}_\infty\|^2\right]$ is the mean of the squared norm of the coefficient error after convergence. To be able to calculate the expression in (36), we need P_e, which in turn depends on $\sigma_x^2 \mathrm{E}\left[\|\Delta\mathbf{w}_\infty\|^2\right]$. Therefore, consider the following two cases of approximation.

AP1) The variance of the error is lower bounded by the noise variance, i.e., $\sigma_e^2 = \sigma_n^2 + \sigma_x^2 \mathrm{E}\left[\|\Delta\mathbf{w}_\infty\|^2\right] \geq \sigma_n^2$. Therefore, a simple lower bound is given by $\hat{P}_e \geq 2Q(\gamma/\sigma_n)$

AP2) We can rewrite the variance of the error as $\sigma_{e_k}^2 = \sigma_n^2 + \mathrm{E}\left[\tilde{e}_k^2\right]$, where $\tilde{e}_k = e_k - e_{opt}$ denotes the distance between the error at kth iteration and the optimal error. Assuming no update, we have $|e_k| \leq \gamma$, and with $\sigma_{opt}^2 = \sigma_n^2$, we get $\sigma_{e_k}^2 \leq 2\sigma_n^2 + \gamma^2$. Therefore, an upper bound of the probability of update is given by $\hat{P}_e = 2Q(\gamma/\sigma_e) \leq 2Q\left(\gamma/\sqrt{2\sigma_n^2 + \gamma^2}\right)$

The approximations of P_e together with (36) are used in the simulations to estimate the excess MSE for different thresholds γ.

D. Excess MSE for Colored Input Signals

When extending the analysis to colored input signals, we may still use the input-signal model in (27). The angular distributions of $\mathbf{x}_k$ will change, i.e., the probabilities $P\left[\mathbf{x}_k \| \mathbf{x}_{k-1}\right]$ and $P\left[\mathbf{x}_k \perp \mathbf{x}_{k-1}\right]$ will be different from those for white input signals. However, as with the case of white input signals, these probabilities will not have effect on the final results; see (35). In order to get an expression for the probability of update P_e for colored input signals, we assume that the input is correlated according to

$$x_k = r x_{k-1} + (1-r) v_k \tag{39}$$

where v_k is a white noise process with zero mean and variance σ_v^2. Straightforward calculations give the autocorrelation matrix

$$\mathbf{R} = \sigma_x^2 \begin{bmatrix} 1 & r & r^2 & \dots & r^{N-1} \\ r & 1 & r & \dots & r^{N-2} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ r^{N-1} & r^{N-2} & r^{N-3} & \dots & 1 \end{bmatrix} \tag{40}$$

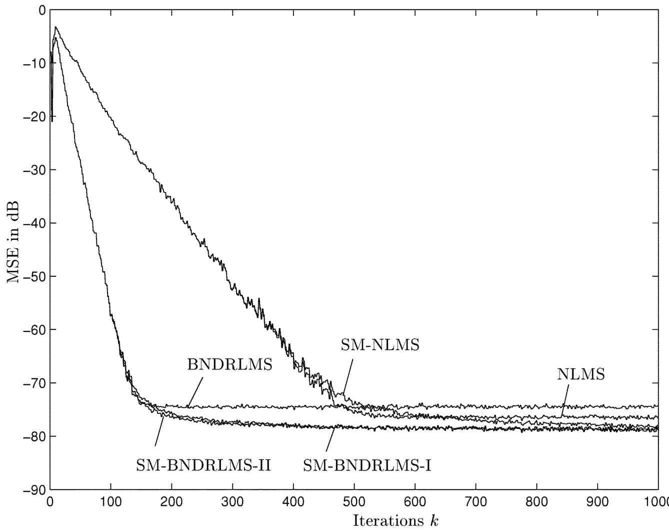

Fig. 5. Learning curves of the SM-BNDRLMS-I, the SM-BNDRLMS-II, the SM-NLMS, the BNDRLMS, and the NLMS algorithms. Condition number of the input-signal correlation matrix $= 100$, SNR$= 80$ dB, and $\gamma = \sqrt{5}\sigma_n$.

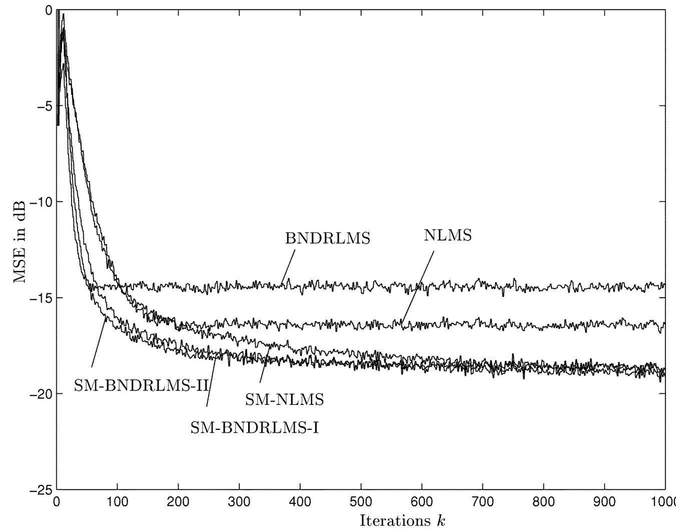

Fig. 6. Learning curves of the SM-BNDRLMS-I, the SM-BNDRLMS-II, the SM-NLMS, the BNDRLMS, and the NLMS algorithms. Condition number of the input-signal correlation matrix $= 100$, SNR$= 20$ dB, and $\gamma = \sqrt{5}\sigma_n$.

where

$$\sigma_x^2 = \frac{1-r}{1+r}\sigma_v^2 = b\sigma_v^2. \tag{41}$$

Assuming the filter has converged to its steady state, the variance of the output error can now be computed as

$$\begin{aligned}\sigma_e^2 =& \sigma_n^2 + \mathrm{E}[\Delta\mathbf{w}_\infty^T \mathbf{R} \Delta\mathbf{w}_\infty] \\ \leq& \sigma_n^2 + \frac{\sigma_x^2}{b}\mathrm{E}[\|\Delta\mathbf{w}_\infty\|^2]. \end{aligned}\tag{42}$$

The probability of update is now given by

$$P_e \leq Q\left(\frac{\gamma}{\sqrt{\sigma_n^2 + b^{-1}\sigma_x^2 \mathrm{E}[\|\Delta\mathbf{w}_\infty\|^2]}}\right). \tag{43}$$

To be able to evaluate the probability of update P_e, the same approximation is made as in AP2) for the case of white input signals, i.e., $\sigma_x^2\mathrm{E}[\|\Delta\mathbf{w}_\infty\|^2] \leq \sigma_n^2 + \gamma^2$. An upper bound for the case of colored input signals is now given by $\hat{P}_e \leq 2Q\left(\gamma/\sqrt{(1+b^{-1})\sigma_n^2 + b^{-1}\gamma^2}\right)$. The lower bound given in AP1) in the previous section is still valid.

V. Numerical Examples

In this section, the new algorithms are applied to a system identification problem. The order of the plant was $p = N - 1 = 10$ and the input signal was colored noise with condition number 100. The signal-to-noise ratio (SNR) was set to 80 and 20 dB in two different examples.

Fig. 5 shows the learning curves averaged over 500 simulations for the SM-BNDRLMS-I, the SM-BNDRLMS-II, the SM-NLMS, the BNDRLMS, and the NLMS algorithms for an $SNR = 80$ dB. The upper bound on the estimation error was set to $\gamma = \sqrt{5}\sigma_n$, and the step sizes used in the BNDRLMS and the NLMS algorithms were set to unity in order to obtain the fastest convergence.

Fig. 5 clearly shows how the SM-BNDRLMS-I and the SM-BNDRLMS-II algorithms combine the fast convergence of the BNDRLMS algorithm with the low misadjustment of the SM-NLMS algorithm. In an ensemble of 500 experiments of 1000 iterations, the average number of updates per experiment for the SM-BNDRLMS-I, SM-BNDRLMS-II, and the SM-NLMS algorithms were, 185, 180, and 436, respectively. For the SM-BNDRLMS-I, an average of 108 updates were full updates.

Fig. 6 shows the learning curve results for an SNR$= 20$ dB. The parameters used in the algorithms were the same as in the first example. As can be seen from the figure, the SM-BNDRLMS algorithms still have higher convergence speeds than the SM-NLMS algorithm.

In 1000 iterations, the average number of updates per experiment for the SM-BNDRLMS-I, SM-BNDRLMS-II, and the SM-NLMS algorithms were, 100, 95, and 129, respectively. For the SM-BNDRLMS-I, an average of 15 updates were full updates.

In the two examples above, the NLMS and the BNDRLMS algorithms were unable to reach the same low steady-state value as their set-membership versions, and a trade-off between convergence speed and final MSE was observed.

For the two examples above, we also plotted the overall complexity versus the total number of iterations for the SM-NLMS and the SM-BNDRLMS algorithms. The curves are normalized with respect to the number of filter coefficients N. To minimize the computational complexity for all the algorithms, we recursively estimated $\|\mathbf{x}_k\|^2$ and $\mathbf{x}_k^T\mathbf{x}_{k-1}$ at each iteration. Figs. 7 and 8 show the results based on the above simulations. For the case of high SNR, we see from Fig. 7 that the overall complexity of the SM-BNDRLMS algorithms are initially higher than the SM-NLMS algorithm. As time proceeds, the overall complexity of the SM-BNDRLMS-II algorithm becomes similar to that of the SM-NLMS algorithm. The SM-BNDRLMS-I, with its extra innovation check, tends to a slightly higher value. For a low SNR, the SM-NLMS algorithm will have a slightly lower overall complexity, as compared with the SM-BNDRLMS algorithms.

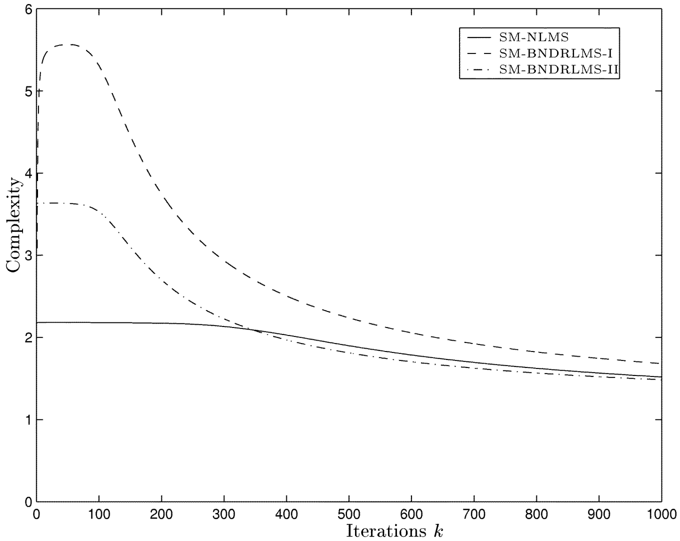

Fig. 7. Overall complexity normalized with N versus the number of data points in the simulation for SNR$= 80$ dB.

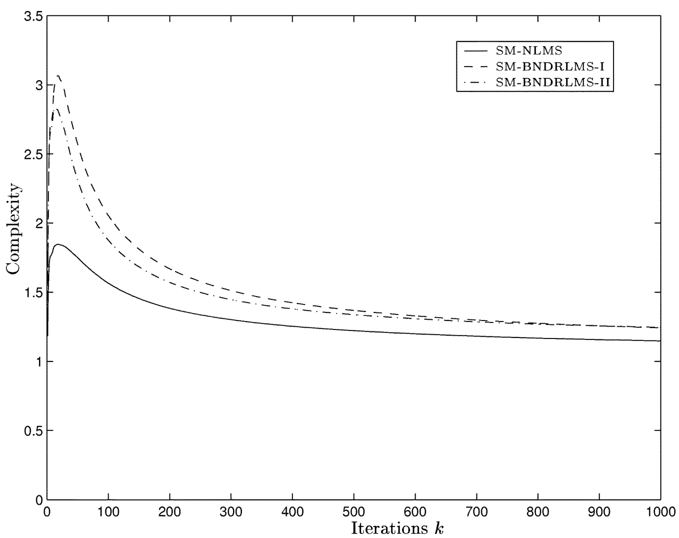

Fig. 8. Overall complexity normalized with N versus the number of data points in the simulation for SNR$= 20$ dB.

TABLE II
EXCESS MEAN-SQUARE ERROR IN NONSTATIONARY ENVIRONMENTS.

ALG.	ξ_{exc} (dB)
NLMS	-40.8
BNDRLMS	-43.5
SM-NLMS	-40.8
SM-BNDRLMS-I	-43.4
SM-BNDRLMS-II	-43.5

In order to test the algorithms in a time-varying environment, the system coefficients were changed according to the model $\mathbf{w}_{opt,k} = \mathbf{w}_{opt,k-1} + \mathbf{u}_k$, where $\mathbf{u}_k$ is a random vector with elements of zero mean and variance $\sigma_v^2 = 10^{-6}$. In the simulations, the additive noise was set to zero, and the bound on the estimation error was set to $\gamma = \sqrt{5}\sigma_v$. The results in terms of the excess MSE in decibels can be found in Table II. As can be noticed, the new proposed algorithms present tracking performance comparable with the BNDRLMS algorithm.

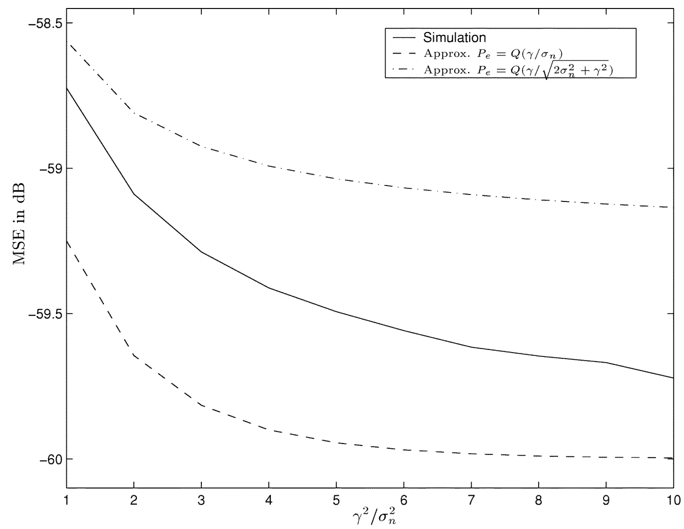

Fig. 9. MSE for $N = 10$ as function of γ^2/σ_n^2 for the input signals as modeled.

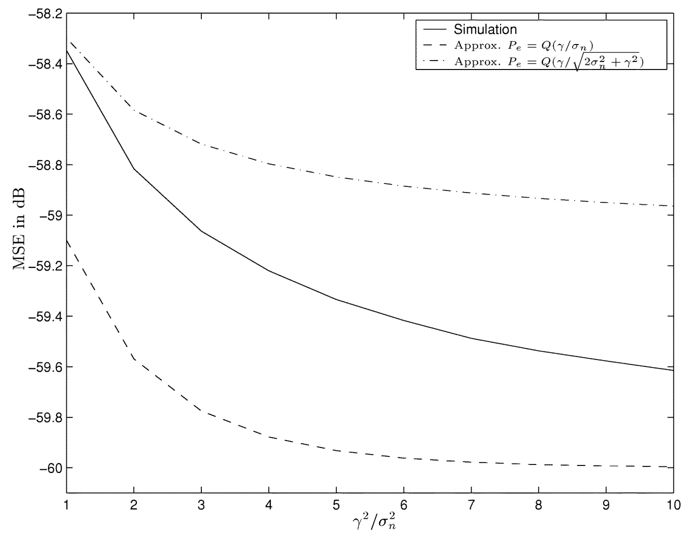

Fig. 10. MSE for $N = 10$ as function of γ^2/σ_n^2 for white input signals.

Finally, experiments were conducted to validate the theoretical results obtained in the MSE analysis. The MSE was measured for different values of γ (γ varied from σ_n to $\sqrt{10}\sigma_n$). The order of the plant was $N - 1 = 10$, and the SNR was chosen to 60 dB.

Fig. 9 shows the MSE versus γ^2/σ_n^2 for a modeled input signal, where the input vectors were chosen such that $\mathbf{v}_k$ and $\mathbf{v}_{k-1}$ were parallel or orthogonal with probabilities $1/N$ and $N - 1/N$, respectively. As can be seen from the figure, the theoretical curves can predict the behavior of the simulation for the assumed model. Figs. 10 and 11 show the results for white and colored input signals, respectively. In the case of colored input, the condition number of the input-signal correlation matrix was equal to 100. It was shown in [4] that the output error e_k is upper bounded by γ after convergence has taken place. Therefore, we can conclude that the MSE is upper bounded by γ^2. However, from the figures, it can be seen that the theoretical formulas for the MSE can provide a much tighter bound than simply considering $\sigma_e^2 = \gamma^2$. If we use this upper bound in AP1) together with

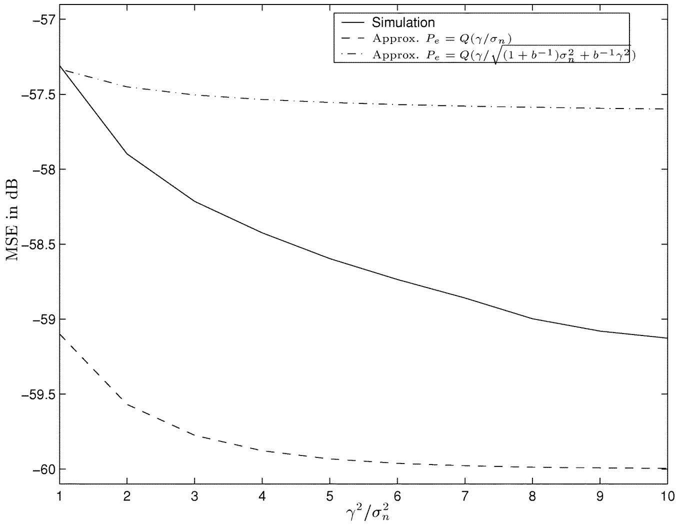

Fig. 11. MSE for $N = 10$ as function of γ^2/σ_n^2 for colored input signals.

(36), the difference for the white input case will be between 2.5 and 10 dB for γ^2/σ^2 in the range 2–10.

VI. CONCLUSIONS

This paper derived two novel adaptation algorithms based on the concept of set-membership filtering. The algorithms utilize consecutive data pairs in order to construct a space of feasible solutions for the updates. The new algorithms were applied to a system identification problem in order to verify the good performance of the algorithm when compared with the SM-NLMS algorithm in terms of high convergence speed, low misadjustment, and reduced number of updates. Analysis for the mean-squared error was carried out for both white and colored input signals, and closed-form expression for the excess MSE was provided. By no means did the algorithms presented form a complete family of SM algorithms with data reusing. A number of alternative algorithms can be derived, considering issues such as computational complexity, hardware implementation, as well as the usual performance criteria utilized for adaptive filtering algorithms, which include convergence speed and excess MSE in stationary and nonstationary environments. The results presented here indicate that the set-membership binormalized data-reusing algorithms represent a family of adaptive filtering algorithms that can provide favorable results in terms of the above-mentioned performance criteria, unlike the most widely used algorithms, such LMS and NLMS, where a tradeoff between convergence speed and excess MSE has to be made.

APPENDIX A

For the special case in which $\mathbf{x}_k \| \mathbf{x}_{k-1}$, the recursions of the SM-BNDRLMS algorithm will be equal to those of the SM-NLMS algorithm. In the derivations below, $\mathbf{x}_k$ is replaced by $s_k r_k \mathbf{v}_k$, and the second-order approximation $\mathrm{E}[1/r_k^2] \approx 1/\sigma_x^2(N+1-\nu_x)$ introduced in [18] is used. The coefficient error at time instant $k+1$ expressed in terms of the probability $P_{e,k}$ can be easily derived in the same manner as with the SM-BNDRLMS algorithms in Section IV and is given by

$$\Delta\mathbf{w}_{k+1} = \left[\mathbf{I} - P_{e,k}\frac{\mathbf{x}_k\mathbf{x}_k^T}{\|\mathbf{x}_k\|^2}\right]\Delta\mathbf{w}_k + P_{e,k}\frac{n_k\mathbf{x}_k}{\|\mathbf{x}_k\|^2}. \tag{44}$$

For the white input signal, we have $\mathbf{R} = \sigma_x^2\mathbf{I}$. The expression for $\Delta\xi_{k+1}$ is given by (45), shown at the bottom of the page, where we have (46), also shown at the bottom of the page, with

$$\rho_1 = \sigma_x^2 \mathrm{tr}\left(\mathrm{E}\left[\Delta\mathbf{w}_k\Delta\mathbf{w}_k^T\right]\right) = \Delta\xi_k \tag{47}$$

$$\begin{aligned}\rho_2 =& -\sigma_x^2 P_{e,k}\mathrm{tr}\left(\mathrm{E}\left[\frac{\Delta\mathbf{w}_k\Delta\mathbf{w}_k^T\mathbf{x}_k\mathbf{x}_k^T}{\|\mathbf{x}_k\|^2}\right]\right)\\ =& -\sigma_x^2 P_{e,k}\mathrm{tr}\left(\mathrm{E}\left[\Delta\mathbf{w}_k\Delta\mathbf{w}_k^T\mathbf{v}_k\mathbf{v}_k^T\right]\right)\\ =& -\frac{\sigma_x^2}{N}P_{e,k}\mathrm{tr}\left(\mathrm{E}\left[\Delta\mathbf{w}_k\Delta\mathbf{w}_k^T\right]\right) = -\frac{P_{e,k}}{N}\Delta\xi_k\\ =& \rho_3\end{aligned} \tag{48}$$

$$\begin{aligned}\Delta\xi_{k+1} =& \sigma_x^2\mathrm{tr}\left(\mathrm{cov}\left[\Delta\mathbf{w}_{k+1}\right]\right) = \sigma_x^2\mathrm{tr}\left(\mathrm{E}\left[\Delta\mathbf{w}_{k+1}\Delta\mathbf{w}_{k+1}^T\right]\right)\\ =& \sigma_x^2\mathrm{tr}\left(\mathrm{E}\left\{\left[\mathbf{I} - P_{e,k}\frac{\mathbf{x}_k\mathbf{x}_k^T}{\|\mathbf{x}_k\|^2}\right]\Delta\mathbf{w}_k\Delta\mathbf{w}_k^T\left[\mathbf{I} - P_{e,k}\frac{\mathbf{x}_k\mathbf{x}_k^T}{\|\mathbf{x}_k\|^2}\right]\right\}\right)\\ &+ \sigma_x^2 P_{e,k}^2\mathrm{tr}\left(\mathrm{E}\left\{\frac{n_k^2\mathbf{x}_k\mathbf{x}_k^T}{\left[\|\mathbf{x}_k\|^2\right]^2}\right\}\right) = \psi_1 + \psi_2\end{aligned} \tag{45}$$

$$\begin{aligned}\psi_1 =& \sigma_x^2\mathrm{tr}\left(\mathrm{E}\left\{\left[\mathbf{I} - P_{e,k}\frac{\mathbf{x}_k\mathbf{x}_k^T}{\|\mathbf{x}_k\|^2}\right]\Delta\mathbf{w}_k\Delta\mathbf{w}_k^T\left[\mathbf{I} - P_{e,k}\frac{\mathbf{x}_k\mathbf{x}_k^T}{\|\mathbf{x}_k\|^2}\right]\right\}\right)\\ =& \sigma_x^2\mathrm{tr}\left(\mathrm{E}\left\{\Delta\mathbf{w}_k\Delta\mathbf{w}_k^T\right\}\right) - \sigma_x^2 P_{e,k}\mathrm{tr}\left(\mathrm{E}\left[\frac{\Delta\mathbf{w}_k\Delta\mathbf{w}_k^T\mathbf{x}_k\mathbf{x}_k^T}{\|\mathbf{x}_k\|^2}\right]\right)\\ &- \sigma_x^2 P_{e,k}\mathrm{tr}\left(\mathrm{E}\left[\frac{\mathbf{x}_k\mathbf{x}_k^T\Delta\mathbf{w}_k\Delta\mathbf{w}_k^T}{\|\mathbf{x}_k\|^2}\right]\right) + \sigma_x^2 P_{e,k}^2\mathrm{tr}\left(\mathrm{E}\left[\frac{\mathbf{x}_k\mathbf{x}_k^T\Delta\mathbf{w}_k\Delta\mathbf{w}_k^T\mathbf{x}_k\mathbf{x}_k^T}{\left[\|\mathbf{x}_k\|^2\right]^2}\right]\right)\\ =& \rho_1 + \rho_2 + \rho_3 + \rho_4\end{aligned} \tag{46}$$

where the last equality is true since $\text{tr}(\mathbf{AB}) = \text{tr}(\mathbf{BA})$.

$$\begin{aligned}\rho_4 =& \sigma_x^2 P_{e,k}^2 \text{tr}\left(\text{E}\left[\mathbf{A}\Delta\mathbf{w}_k\Delta\mathbf{w}_k^T\mathbf{A}\right]\right)\\ =& \sigma_x^2 P_{e,k}^2 \text{tr}\left(\text{E}\left[\frac{\mathbf{x}_k\mathbf{x}_k^T\Delta\mathbf{w}_k\Delta\mathbf{w}_k^T\mathbf{x}_k\mathbf{x}_k^T}{[\|\mathbf{x}_k\|^2]^2}\right]\right)\\ =& \sigma_x^2 P_{e,k}^2 \text{tr}\left(\text{E}\left[\frac{\Delta\mathbf{w}_k^T\mathbf{x}_k\mathbf{x}_k^T\mathbf{x}_k\mathbf{x}_k^T\Delta\mathbf{w}_k}{[\|\mathbf{x}_k\|^2]^2}\right]\right)\\ =& \sigma_x^2 P_{e,k}^2 \text{tr}\left(\text{E}\left[\frac{\Delta\mathbf{w}_k^T\mathbf{x}_k\mathbf{x}_k^T\Delta\mathbf{w}_k}{\|\mathbf{x}_k\|^2}\right]\right)\\ =& \sigma_x^2 P_{e,k}^2 \text{tr}\left(\text{E}\left[\Delta\mathbf{w}_k^T\mathbf{v}_k\mathbf{v}_k^T\Delta\mathbf{w}_k\right]\right)\\ =& \frac{\sigma_x^2 P_{e,k}^2}{N} \text{tr}\left(\text{E}\left[\Delta\mathbf{w}_k\Delta\mathbf{w}_k^T\right]\right)\\ =& \frac{P_{e,k}^2}{N}\Delta\xi_k \end{aligned} \tag{49}$$

$$\begin{aligned}\psi_2 =& \sigma_x^2 P_{e,k}^2 \text{tr}\left(\text{E}\left\{\frac{n_k^2\mathbf{x}_k\mathbf{x}_k^T}{[\|\mathbf{x}_k\|^2]^2}\right\}\right)\\ =& \sigma_x^2 P_{e,k}^2 \text{tr}\left(\text{E}\left[n_k^2\frac{\mathbf{v}_k\mathbf{v}_k^T}{r^2}\right]\right)\\ =& \sigma_x^2 P_{e,k}^2\sigma_n^2\left(\text{E}\left[\frac{1}{r^2}\right]\right) = \frac{\sigma_n^2 P_{e,k}^2}{N+1-\nu_x}. \end{aligned} \tag{50}$$

Finally, we get

$$\Delta\xi_{k+1}|_{\mathbf{x}_k\|\mathbf{x}_{k-1}} = \left(1-\frac{2P_{e,k}-P_{e,k}^2}{N}\right)\Delta\xi_k + P_{e,k}^2\frac{\sigma_n^2}{N+1-\nu_x}. \tag{51}$$

Appendix B

For the case $\mathbf{x}_k\perp\mathbf{x}_{k-1}$, (24) and (25) reduce to

$$\begin{aligned}\mathbf{A}|_{\mathbf{x}_k\perp\mathbf{x}_{k-1}} =& \frac{\mathbf{x}_{k-1}\mathbf{x}_{k-1}^T\mathbf{x}_k\mathbf{x}_k^T - \|\mathbf{x}_{k-1}\|^2\mathbf{x}_k\mathbf{x}_k^T}{\|\mathbf{x}_k\|^2\|\mathbf{x}_{k-1}\|^2-\left(\mathbf{x}_k^T\mathbf{x}_{k-1}\right)^2}\\ =& \frac{-\|\mathbf{x}_{k-1}\|^2\mathbf{x}_k\mathbf{x}_k^T}{\|\mathbf{x}_k\|^2\|\mathbf{x}_{k-1}\|^2} = \frac{-\mathbf{x}_k\mathbf{x}_k^T}{\|\mathbf{x}_k\|^2}\end{aligned} \tag{52}$$

and

$$\begin{aligned}\mathbf{b}|_{\mathbf{x}_k\perp\mathbf{x}_{k-1}} =& n_k\frac{\|\mathbf{x}_{k-1}\|^2\mathbf{x}_k - \left(\mathbf{x}_{k-1}^T\mathbf{x}_k\right)\mathbf{x}_{k-1}}{\|\mathbf{x}_k\|^2\|\mathbf{x}_{k-1}\|^2-\left(\mathbf{x}_k^T\mathbf{x}_{k-1}\right)^2}\\ =& n_k\frac{\mathbf{x}_k}{\|\mathbf{x}_k\|^2}.\end{aligned} \tag{53}$$

The coefficient error vector in (26) now reduces to

$$\Delta\mathbf{w}_{k+1} = \left[\mathbf{I} - P_{e,k}\frac{\mathbf{x}_k\mathbf{x}_k^T}{\|\mathbf{x}_k\|^2}\right]\Delta\mathbf{w}_k + P_{e,k}\frac{n_k\mathbf{x}_k}{\|\mathbf{x}_k\|^2} \tag{54}$$

which is the same as in (44) for the case of the SM-NLMS algorithm. Consequently, we get

$$\Delta\xi_{k+1}|_{\mathbf{x}_k\perp\mathbf{x}_{k-1}} = \left(1-\frac{2P_{e,k}-P_{e,k}^2}{N}\right)\Delta\xi_k + P_{e,k}^2\frac{\sigma_n^2}{N+1-\nu_x}. \tag{55}$$

References

[1] P. S. R. Diniz, *Adaptive Filtering: Algorithms and Practical Implementations*, Second ed. Boston, MA: Kluwer, 2002.
[2] J. A. Apolinário, M. L. R. de Campos, and P. S. R. Diniz, "The binormalized data-reusing LMS algorithm," *IEEE Trans. Signal Processing*, vol. 48, pp. 3235–3242, Nov. 2000.
[3] J. A. Apolinário, M. L. R. de Campos, P. S. R. Diniz, and T. I. Laakso, "Binormalized data-reusing LMS algorithm with optimized step-size sequence," *Brazilian Telecommun. J. SBT*, vol. 13, pp. 49–55, 1998.
[4] S. Gollamudi, S. Nagaraj, S. Kapoor, and Y. F. Huang, "Set-membership filtering and a set-membership normalized LMS algorithm with an adaptive step size," *IEEE Signal Processing Lett.*, vol. 5, pp. 111–114, Apr. 1998.
[5] E. Fogel and Y. F. Huang, "On the value of information in system identification – bounded noise case," *Automatica*, vol. 18, pp. 229–238, 1982.
[6] J. R. Deller, "Set-membership identification in digital signal processing," *IEEE Acoust., Speech, Signal Processing Mag.*, vol. 6, pp. 4–22, Jan. 1989.
[7] J. R. Deller, M. Nayeri, and S. F. Odeh, "Least-squares identification with error bounds for real-time signal processing and control," *Proc. IEEE*, vol. 81, pp. 815–849, June 1993.
[8] E. Walter and H. Piet-Lahanier, "Estimation of parameter bounds from bounded-error: A survey," *Math. Comput. Simul.*, vol. 32, pp. 449–468, 1990.
[9] S. Gollamudi, S. Nagaraj, and Y. F. Huang, "SMART : A toolbox for set-membership filtering," in *Proc. IEEE Eur. Conf. Circuit Theory Des.*, 1997, pp. 879–884.
[10] S. Dasgupta and Y. F. Huang, "Asymptotically convergent modified recursive least-squares with data dependent updating and forgetting factor for systems with bounded noise," *IEEE Trans. Inform. Theory*, vol. IT-33, pp. 383–392, 1987.
[11] J. R. Deller, M. Nayeri, and M. Liu, "Unifying the landmark developments in optimal bounding ellipsoid identification," *Int. J. Adap. Contr. Signal Process.*, vol. 8, pp. 43–60, 1994.
[12] D. Joachim, J. R. Deller, and M. Nayeri, "Multiweight optimization in OBE algorithms for improved tracking and adaptive identification," in *Proc. IEEE Int. Acoust., Speech Signal Process.*, vol. 4, Seattle, WA, May 1998, pp. 2201–2204.
[13] J. R. Deller, S. Gollamudi, S. Nagaraj, and Y. F. Huang, "Convergence analysis of the QUASI-OBE algorithm and the performance implications," in *Proc. IFAC Int. Symp. Syst. Ident.*, vol. 3, Santa Barbara, CA, June 2000, pp. 875–880.
[14] S. Gollamudi, S. Nagaraj, S. Kapoor, and Y. F. Huang, "Set-membership adaptive equalization and updator-shared implementation for multiple channel communications systems," *IEEE Trans. Signal Processing*, vol. 46, pp. 2372–2384, Sept. 1998.
[15] S. Kapoor, S. Gollamudi, S. Nagaraj, and Y. F. Huang, "Adaptive multiuser detection and beamforming for interference suppression in CDMA mobile radio systems," *IEEE Trans. Veh. Technol.*, vol. 48, pp. 1341–1355, Sept. 1999.
[16] S. Nagaraj, S. Gollamudi, S. Kapoor, Y. F. Huang, and J. R. Deller, "Adaptive interference suppression for CDMA systems with a worst-case error criterion," *IEEE Trans. Signal Processing*, vol. 48, pp. 285–289, Jan. 2000.
[17] S. Nagaraj, S. Gollamudi, S. Kapoor, and Y. F. Huang, "BEACON : An adaptive set-membership filtering technique with sparse updates," *IEEE Trans. Signal Processing*, vol. 47, pp. 2928–2941, Nov. 1999.
[18] D. T. M. Slock, "On the convergence behavior of the LMS and the normalized LMS algorithms," *IEEE Trans. Signal Processing*, vol. 41, pp. 2811–2825, Sept. 1993.

Paulo S. R. Diniz (F'00) was born in Niterói, Brazil. He received the Electronics Eng. degree (Cum Laude) from the Federal University of Rio de Janeiro (UFRJ), Rio de Janeiro, Brazil, in 1978, the M.Sc. degree from COPPE/UFRJ in 1981, and the Ph.D. degree from Concordia University, Montreal, QC, Canada, in 1984, all in electrical engineering.

Since 1979, he has been with the Department of Electronic Engineering (undergraduate), UFRJ. He has also been with the Program of Electrical Engineering (graduate), COPPE/UFRJ, since 1984, where he is presently a Professor. He served as Undergraduate Course Coordinator and as Chairman of the Graduate Department. He is one of the three senior researchers and coordinators of the National Excellence Center in Signal Processing. From January 1991 to July 1992, he was a Docent Research Associate with the Department of Electrical and Computer Engineering, University of Victoria, Victoria, B.C., Canada. He also holds a visiting Professor position at Helsinki University of Technology, Helsinki, Finland. From January 2002 to June 2002, he was a Melchor Chair Professor with the Department of Electrical Engineering, University of Notre Dame, Notre Dame, IN. His teaching and research interests are in analog and digital signal processing, adaptive signal processing, digital communications, wireless communications, multirate systems, stochastic processes, and electronic circuits. He has published several refereed papers in some of these areas and wrote the books *Adaptive Filtering: Algorithms and Practical Implementation* (Boston, MA: Kluwer, Second Ed., 2002) and *Digital Signal Processing: System Analysis and Design* (Cambridge, U.K.: Cambridge Univ. Press, 2002: with E. A. B. da Silva and S. L. Netto). He was an associate editor of the *Circuits, Systems, and Signal Processing Journal* from 1998 to 2002.

Dr. Diniz was the Technical Program Chair of the 1995 MWSCAS, Rio de Janeiro, Brazil. He has been on the technical committee of several international conferences including ISCAS, ICECS, EUSIPCO, and MWSCAS. He has served Vice President for region 9 of the IEEE Circuits and Systems Society and as Chairman of the DSP technical committee of the same Society. He served as associate editor for the IEEE TRANSACTIONS ON CIRCUITS AND SYSTEMS II from 1996 to 1999 and the IEEE TRANSACTIONS ON SIGNAL PROCESSING from 1999 to 2002. He was a distinguished lecturer of the IEEE Circuits and Systems Society for the year 2000 to 2001. He received the Rio de Janeiro State Scientist award from the Governor of the state of Rio de Janeiro.

Stefan Werner (M'02) received the M.Sc. degree in electrical engineering from the Royal Institute of Technology, Stockholm, Sweden, in 1998, and the Dr.Tech. degree (with honors) from the Signal Processing Laboratory, Helsinki University of Technology, Helsinki, Finland, in 2002.

His research interests are in multiuser communications and adaptive filtering.