LINEARLY CONstrained ADAPTIVE FILTERING ALGORITHMS DESIGNED USING CONTROL LIAPUNOV FUNCTIONS

OUMAR DIENE*, AMIT BHAYA*

*Universidade Federal do Rio de Janeiro
COPPE - Programa de Engenharia Elétrica
Núcleo de Atendimento em Computação de Alto Desempenho
Cidade Universitária - Ilha do Fundão
21.945-970 - Rio de Janeiro - R. J.

Emails: oumar@coep.ufrj.br, amit@nacad.ufrj.br

Abstract— The standard conjugate gradient (CG) method uses orthogonality of the residues to simplify the formulas for the parameters necessary for convergence. In adaptive filtering, the sample-by-sample update of the correlation matrix and the cross-correlation vector causes a loss of the residue orthogonality in a modified online algorithm, which, in turn, results in loss of convergence and an increase of the filter quadratic mean error. This paper extends a recently proposed optimality and convergence proof of the degenerated CG method to the case of linearly constrained adaptive filtering, and proposes a constrained Steepest Descent (CSD) method.

Keywords— Linearly constrained adaptive filtering algorithms, single-user DS-CDMA, control Liapunov function, conjugate gradient method, steepest descent method.

1 Introduction

The Conjugate Gradient (CG) is considered to be one of the best iterative methods for linear systems of equations with symmetric positive definite (spd) coefficient matrices (Saad, 1996). The standard CG method uses orthogonality of the residues to simplify the formulas for the parameters necessary for convergence (Hestenes and Stiefel, 1952). In applications such as adaptive filtering, the sample-by-sample update of the correlation matrix and the cross-correlation vector causes a loss of the residue orthogonality in a modified online algorithm. This loss of orthogonality results in loss of convergence and an increase of the filter quadratic mean error (Diene and Bhaya, 2006). Boray and Srinath (1992) proposed a minimization of the output-error energy of an adaptive finite impulse response (FIR) filter by using the CG method with a sliding data window. In this case, several CG iterations are made at each sample. Chang and Willson (2000) proposed an alternative degenerated CG method in which an exponentially decaying window performs only one iteration of the CG method at each sample and leads to an algorithm with performance comparable to the RLS algorithm. Using the recently proposed control Liapunov (CLF) analysis of the CG method viewed as a dynamic system in the standard feedback configuration (Bhya and Kaszkurewicz, 2006), the optimality of the parameters and the convergence of the degenerated CG method were proved, and a Steepest Descent (SD) method was proposed in Diene and Bhaya (2006). The constrained version of the degenerated CG algorithm of Chang and Willson (2000) was proposed, in Apolinário et al. (2000), for the minimization of the output-error energy of an adaptive FIR filter subject to a set of linear constraints, i.e., \(\min_{\mathbf{w}} E[\mathbf{e}^2] \) subject to \( \mathbf{C}^T \mathbf{w} = \mathbf{f} \), where \( \mathbf{w} \) is the length \( M \) coefficient vector, \( e \) is the filter output error, \( \mathbf{C} \) is the \( M \times p \) constraint matrix, and \( \mathbf{f} \) is the length \( p \) gain vector. This paper extends the optimality and convergence proof of the degenerated CG method (Diene and Bhaya, 2006) to the case of linearly constrained adaptive filtering, and proposes a constrained Steepest Descent (CSD) method. The main technique consists of an appropriate choice of a CLF, and has been used in Bhaya and Kaszkurewicz (2003) and Diene and Bhaya (2004) to analyze and design Krylov and new iterative methods.

2 Preliminaries

Consider the linear system

\[
\mathbf{R}_w = \mathbf{b}.
\]

Iterative methods for solving the system (1) can be described by the following equation

\[
\mathbf{w}_{k+1} = \mathbf{w}_k + \mathbf{u}_k
\]

where \( \mathbf{w}_k \) is an approximation of the solution and \( \mathbf{u}_k \) is a correction calculated in order to increase the accuracy of the approximated solution \( \mathbf{w}_{k+1} \) at each iteration (Saad, 1996), and then drive the residue, defined as \( \mathbf{g}_k = \mathbf{b} - \mathbf{R}_w \mathbf{w}_k \), to zero in a finite number of iterations. The problem of solving (1) by iterative methods can be represented by the signal regulation problem:

\[
\begin{align*}
\mathbf{w}_{k+1} &= \mathbf{w}_k + \mathbf{u}_k, \\
\mathbf{y}_k &= \mathbf{R}_w \mathbf{w}_k, \\
\mathbf{g}_k &= \mathbf{b} - \mathbf{y}_k, \\
\mathbf{u}_k &= f(\mathbf{g}_k),
\end{align*}
\]
Figure 1: Block diagram representing the regulation problem equivalent to resolution (1) of by iterative methods. The problem is to design a controller such that \( \lim_{k \to \infty} g_k = 0 \Leftrightarrow \lim_{k \to \infty} Rw_k = b \).

i.e., designing control signal \( u_k \) in order to zero the steady state error by choice of the control law \( f(\cdot) \) i.e., \( \lim g_k = 0 \), or, equivalently, \( \lim Rw_k = b \).

System (3), represented in block diagram form in figure 1, models an iterative method as a feedback control system (plant \( \{ I, I, R, 0 \} \) with controller in unitary feedback configuration). In control theory, both classical and modern, there are various techniques of solving the signal regulation problem (3) representing the solution of (1), which correspond to different choices of the controller \( f(\cdot) \) in (3). However, as shown in Diene (2004), many of these techniques (pole allocation, state feedback) result in controllers that explicitly use the inverse matrix \( R^{-1} \). In Bhaya and Kaszkurewicz (2003) and Bhaya and Kaszkurewicz (2006), the authors show that, by using appropriate quadratic techniques, choosing a proportional derivative controller for \( f(\cdot) \) in (3) leads to the CG method, shown in algorithm 1. The CLF method proposed in Bhaya and Kaszkurewicz (2003) and Bhaya and Kaszkurewicz (2006) gives an idea of the robustness of the CG method, since it does not use the assumption \( p_0 = g_0 \) (initialization of conjugate directions), used to obtain residue orthogonality (see, e.g. Hestenes and Stiefel (1952)). This indicates that, in situations where the assumption does not hold, the determination of CG parameters by CLF still ensures the decrease of some norm of the residual vector. When the assumption \( p_0 = g_0 \) holds, it can be shown that the parameters determined by CLFs are the same as those commonly found in the literature (Hestenes and Stiefel, 1952, p.411, eqs 3.1b,3.1e).

**Algorithm 1** Hestenes-Stiefel CG algorithm for constant Linear Systems

Calculate \( g_0 = b - Rw_0, p_0 = g_0 \)

For \( k = 0, 1, \ldots, \) until convergence

\[
\begin{align*}
\alpha_k &= \frac{(g_k, p_k)}{(p_k, R p_k)} \\
\omega_{k+1} &= \omega_k + \alpha_k p_k \\
g_{k+1} &= g_k - \alpha_k R p_k \\
\beta_k &= \frac{(g_{k+1}, p_k)}{(p_k, R p_k)} \\
p_{k+1} &= g_{k+1} + \beta_k p_k
\end{align*}
\]

**3 Designing new Linearly Constrained Adaptive Filtering Algorithms**

The problem of linearly constrained adaptive filters can be formulated as follows:

\[
\begin{align*}
\text{Solve} & \quad R_k w_k = b_k, \\
\text{subject to} & \quad C^T w_k = f
\end{align*}
\]

where \( b_k = E[d_k x_k] \) is the length \( M \) cross-correlation vector between the scalar desired response \( d_k \) and the length \( M \) input signal \( x_k \), \( R_k = E[x_k x_k^T] \) is the \( M \times M \) correlation matrix of the input signal, \( w_k \) is the length \( M \) coefficient vector, \( C \) is the \( M \times p \) constraint matrix, and \( f \) is the length \( p \) gain vector. To solve this problem, a structure was presented in Griffiths and Jim (1982) in which only a small set of coefficients are updated, and this set is confined to the subspace orthogonal to the space spanned by the constraint matrix \( C \). This strategy, referred to as the general- ized sidelobe cancelation (GSC), is able to transform the linearly constrained minimization problem into an unconstrained minimization problem, and subsequently any adaptation algorithm can be used (Apolinário et al., 2000). Use of GSC consists of solving the linear unconstrained minimization problem:

\[
R_k^\prime w_k^\prime = b_k^\prime, \quad (5)
\]

where \( b_k^\prime = E[d_k x_k^\prime] \), \( R_k^\prime = E[x_k^\prime x_k'^T] \), \( x_k^\prime = B^T x_k \), \( B \) is the blocking matrix such that \( B^T C = 0 \). The last step is to determine the solution of the constrained problem (4) using equation (6) below

\[
w_k = F - B w_k^\prime, \quad (6)
\]

where \( F = C (C^T C)^{-1} f \). Notice that equation (6) guarantees that \( C^T w_k = f \). Using the exponentially decaying data window, proposed in Chang and Willson (2000), the recurrences of the auto-correlation matrix \( R_k^\prime \) and the cross-correlation vector \( b_k^\prime \) are given by:

\[
\begin{align*}
R_{k+1}^\prime &= \lambda_f R_k^\prime + x_{k+1}'x_{k+1}'^T \\
b_{k+1}' &= \lambda_f b_k' + d_{k+1}x_{k+1}'
\end{align*}
\]

where \( \lambda_f < 1 \) is the forgetting factor. For sample-by-sample processing, a recursive formulation for the residual vector can be found by using (2), (7) and (8), resulting in

\[
g_{k+1} = \lambda_f g_k - \alpha_k R_{k+1}^\prime p_k' + x_{k+1}'(d_{k+1} - x_{k+1}'^T w_k^\prime). \quad (9)
\]

Then, for the unconstrained minimization problem (5), the system (3) can be rewritten as

\[
\begin{align*}
w_{k+1}' &= w_k' + \alpha_k p_k' \\
R_{k+1}^\prime &= \lambda_f R_k^\prime + x_{k+1}'x_{k+1}'^T \\
g_{k+1} &= \lambda_f g_k - \alpha_k R_{k+1}^\prime p_k + x_{k+1}'(d_{k+1} - x_{k+1}'^T w_k') \\
p_{k+1}' &= g_{k+1}' + \beta_k p_k'
\end{align*}
\]

(10)
Choosing a blocking matrix $B$ such that $B^T B = I$, it can be seen that

$$ R_k' = B^T R_k B \tag{11} $$

Then, multiplying both sides of (9) by $B$ and choosing

$$ g_k = Bg_k' \tag{12} $$
$$ p_k = Bp_k' \tag{13} $$

it can be shown that (Apolinário et al., 2000)

$$ g_{k+1} = \lambda_f g_k - \alpha_k R_{k+1} p_k - x_{k+1} e_{k+1} \tag{14} $$

where

$$ R_{k+1} = PR_{k+1} P = BR_k B^T \tag{15} $$
$$ x_{k+1} = Ps_{k+1} \tag{16} $$
$$ P = BB^T \tag{17} $$
$$ e_{k+1} = d_{k+1} - x_{k+1}^T w_k \tag{18} $$

In order to design new constrained algorithms, it is necessary to determine the system of equations representing the equivalent regulation problem (3). This can be done noting that, from (7), (15), (10), (12) and (13) it is straightforward to show that

$$ R_{k+1} = \lambda_f R_k + x_{k+1}^T x_{k+1} \tag{19} $$
$$ p_{k+1} = g_{k+1} + \beta_k p_k \tag{20} $$

The weight updating equation (6) can be rewritten as

$$ w_{k+1} = F - B w_{k+1} = P w_k + F - \alpha_k p_k \tag{21} $$

Then, for the constrained minimization problem (4), the system (3) can be rewritten as

$$ \begin{cases} w_{k+1} = P w_k + F - \alpha_k p_k \\ e_{k+1} = d_{k+1} - x_{k+1}^T w_k \\ R_{k+1} = \lambda_f R_k + x_{k+1}^T x_{k+1} \\ g_{k+1} = \lambda_f g_k - \alpha_k R_{k+1} p_k - x_{k+1} e_{k+1} \\ p_{k+1} = g_{k+1} + \beta_k p_k. \end{cases} \tag{22} $$

For the system (22), the main result of this paper is stated as follows:

**Theorem 1** The following assumption is made:

$$ \inf_k \lambda_{\min}(R_k) = \gamma_1 > 0, \quad \inf_k \lambda_{\min}(R_k^{-1}) = \gamma_2 > 0. \tag{23} $$

With this assumption, the choices

$$ \alpha_k = \frac{\langle \lambda_f g_k, p_k \rangle}{\langle p_k, R_{k+1} p_k \rangle} \tag{24} $$
$$ \beta_k = -\frac{\langle g_{k+1}, R_{k+1} p_k \rangle}{\langle p_k, R_{k+1} p_k \rangle} \tag{25} $$

are optimal and ensure that the equilibrium of the system (10) is asymptotically stable (i.e. $g_k \to 0$ and $p_k \to 0$).

**Proof:** It is desired that $e_{k+1} \to 0$. Thus (14) can be rewritten as

$$ g_{k+1} = \lambda_f g_k - \alpha_k R_{k+1} p_k. \tag{26} $$

Choose the control Lyapunov function $V_{g_k}(k) = \langle g_k, R_k^{-1} g_k \rangle = \|g_k\|^2_{R_k^{-1}}$, then

$$ \Delta V_g = \langle g_{k+1}, R_{k+1}^{-1} g_{k+1} \rangle - \langle g_k, R_k^{-1} g_k \rangle. \tag{27} $$

Substituting (26) into (27) results in

$$ \Delta V_g = \lambda_f^2 \langle g_k, R_k^{-1} R_{k+1} g_k \rangle - 2\alpha_k \langle \lambda_f g_k, p_k \rangle + \alpha^2 \langle p_k, R_{k+1} p_k \rangle - \langle g_k, R_k^{-1} g_k \rangle. \tag{28} $$

The optimal value of $\alpha_k$ (which makes $\Delta V_g$ as negative as possible), given by (24), is determined calculating $\frac{\partial \Delta V_g}{\partial \alpha_k}$ and setting it to zero. Substituting (24) into (28) yields

$$ \Delta V_g = -\frac{\langle \lambda_f g_k, p_k \rangle^2}{\langle p_k, R_{k+1} p_k \rangle} + \lambda_f^2 \langle g_k, R_k^{-1} R_{k+1} g_k \rangle - \langle g_k, R_k^{-1} g_k \rangle \tag{29} $$

Using the matrix inversion lemma (Haykin, 1991), one can write

$$ R_{k+1}^{-1} = \lambda_f R_k^{-1} - c \lambda_f^2 R_k^{-1} x_{k+1} x_{k+1}^T R_k^{-1}, \tag{30} $$

where $c = (1/(1 + \lambda_f^2 x_{k+1}^T R_k^{-1} x_{k+1})) > 0$. Inserting this into (29) results in

$$ \Delta V_g = -\frac{\langle \lambda_f g_k, p_k \rangle^2}{\langle p_k, R_{k+1} p_k \rangle} + \lambda_f^2 \langle g_k, R_k^{-1} R_{k+1} g_k \rangle - c \langle g_k, R_k^{-1} x_{k+1} x_{k+1}^T R_k^{-1} g_k \rangle. \tag{31} $$

Then, from hypothesis (23), it follows that $\lambda_f < 1$ implies that

$$ \Delta V_g < -c \|g_k\|^2_{R_k^{-1}} < 0, \tag{32} $$

which implies that $\|g_k\|_{R_k^{-1}}$ is a decreasing sequence.

For the choice of $\beta_k$, the choice of the CLF $V_{p_k}(k) = \langle p_k, R_k p_k \rangle = \|p_k\|^2_{R_k}$, then

$$ \Delta V_p = \langle g_{k+1}, R_{k+1} p_k \rangle + 2\beta_k \langle g_{k+1}, R_{k+1} p_k \rangle \tag{33} $$

In this case, the optimal value of $\beta_k$, given by (25), is obtained calculating $\frac{\partial \Delta V_p}{\partial \beta_k}$ and setting it to zero, yielding

$$ V_{p_{k+1}} (k+1) = \|g_{k+1}\|^2_{R_{k+1}} - \frac{\langle g_{k+1}, R_{k+1} p_k \rangle^2}{\langle p_k, R_{k+1} p_k \rangle} < \|g_{k+1}\|^2_{R_{k+1}}. \tag{33} $$

From (32), the equivalence of norms and hypothesis (23), it can be concluded that $\|g_k\|_{R_k^{-1}}$ is a decreasing sequence. Thus (33) implies that $p_{k+1}$ decreases in $R_{k+1}$-norm, not necessarily monotonically, concluding the proof. \qed
Observation 1 Notice that equation (31) can be used to determine the value of \( \lambda_f \) which makes \( \Delta V_g \) as negative as possible: using the expression (31) and solving the unidimensional minimization problem \( \min_{\lambda_f} \Delta V_g \), allows, in theory, the determination of an optimal value of \( \lambda_f \).

Observation 2 Also notice that for the constrained algorithm, the projection matrix \( P = I - C(C^T C)^{-1} C^T \), thus it is not necessary to determine explicitly the blocking matrix \( B \).

The pseudocode of the extension of the Hestenes-Stiefel CG method for linearly constrained adaptive filtering algorithms, based on theorem 1 and referenced in this paper as CCG-CLF, is presented in algorithm 2.

Algorithm 2 CCG-CLF method for linearly constrained adaptive filters

Calculate \( w_0 = F = C(C^T C)^{-1} f \)

\[ R_0 = P = I - C(C^T C)^{-1} C^T \]

\[ g_0 = p_0 = 0 \]

For \( k = 0, 1, \ldots, N - 1 \)

\[ \alpha_k = \frac{\langle \lambda_f g_k, p_k \rangle}{\langle p_k, R_{k+1} p_k \rangle} \]

\[ w_{k+1} = P w_k + F - \alpha_k p_k \]

\[ x_{k+1} = P x_{k+1} \]

\[ R_{k+1} = \lambda_f R_k + x_{k+1} x_{k+1}^T \]

\[ e_{k+1} = d_k - x_{k+1}^T w_k \]

\[ g_{k+1} = \lambda_f g_k - \alpha_k R_{k+1} p_k - x_{k+1} e_{k+1} \]

\[ \beta_k = \langle g_k, p_k \rangle \]

\[ p_{k+1} = g_{k+1} + \beta_k p_k \]

End

Note that the algorithm in algorithm 2 does not introduce any additional computation and requires 3M floating point operations (FLOP) less than the algorithm in Apolinário et al. (2000). Choosing the ad hoc parameter \( \eta \) in Apolinário et al. (2000) equal to \( \lambda_f \) gives our formula for \( \alpha_k \) which is above shown to be an optimal choice in the sense of making \( \Delta V_g \) as negative as possible, although, in non-stationary environments, a smaller value (\( \eta < \lambda_f \) which keeps \( \Delta V_g < 0 \)) may be advisable.

In Diene and Bhaya (2006), a steepest descent (SD) method for adaptive filters was proposed. The results in Diene and Bhaya (2006) show that it is possible for the SD method for adaptive filters to have a performance comparable with the CG and RLS methods. The same extension, made above for the CG method, can be done for the SD method in order to design, by CLFs, a SD method for linearly constrained adaptive filters (CSD-CLF). The system of equations representing the regulation problem, equivalent to solving the linearly constrained adaptive filtering problem (4) by the CSD-CLF method, is given by system (34) below

\[
\begin{align*}
 w_{k+1} & = P w_k + F - \alpha_k g_k \\
 e_{k+1} & = d_k - x_{k+1}^T w_k \\
 R_{k+1} & = \lambda_f R_k + x_{k+1} x_{k+1}^T \\
 g_{k+1} & = \lambda_f g_k - \alpha_k R_{k+1} g_k - x_{k+1} e_{k+1}. \\
\end{align*}
\]

(34)

For the system (34), another result of this paper is stated as follows:

**Theorem 2** With assumption (23) valid, the choice

\[ \alpha_k = \frac{\langle \lambda_f g_k, g_k \rangle}{\langle g_k, R_{k+1} g_k \rangle} \]

(35)

is optimal and ensures that the equilibrium of the system (34) is asymptotically stable (i.e. \( g_k \rightarrow 0 \)).

**Proof:** Once again, it is desired that \( e_{k+1} \rightarrow 0 \). Thus substituting in (34) gives

\[ g_{k+1} = \lambda_f g_k - \alpha_k R_{k+1} g_k. \]

(36)

Again choose the control Liapunov function \( V_{g_k}(k) = \langle g_k, R_k^{-1} g_k \rangle = \|g_k\|^2/R_k \)

\[ \Delta V_g = \langle g_{k+1}, R_k^{-1} g_{k+1} \rangle - \langle g_k, R_k^{-1} g_k \rangle. \]

(37)

Substituting (36) into (37) results in

\[ \Delta V_g = \lambda_f^2 \langle g_k, R_k^{-1} g_k \rangle - 2 \alpha_k \langle \lambda_f g_k, g_k \rangle + \alpha_k^2 \langle g_k, R_k^{-1} g_k \rangle \]

(38)

The optimal value of \( \alpha_k \) (which makes \( \Delta V_g \) as negative as possible), given by (35), is determined calculating \( \partial \Delta V_g / \partial \alpha_k \) and setting it to zero. Substituting (35) into (38) yields

\[ \Delta V_g = -\frac{\langle \lambda_f g_k, g_k \rangle^2}{\langle g_k, R_k^{-1} g_k \rangle} + \lambda_f^2 \langle g_k, R_k^{-1} g_k \rangle \]

(39)

Inserting equation (30) into (39) results in

\[ \Delta V_g = -\frac{\langle \lambda_f g_k, g_k \rangle^2}{\langle g_k, R_k^{-1} g_k \rangle} - (1 - \lambda_f) \langle g_k, R_k^{-1} g_k \rangle \]

(40)

Then, from hypothesis (23), it follows that \( \lambda_f < 1 \) implies that

\[ \Delta V_g < -\gamma \|g_k\|^2/R_k \]

(41)

which implies that \( \|g_k\|^2/R_k \) is a decreasing sequence.

**The pseudocode of the extension of the SD method for linearly constrained adaptive filtering algorithms, based on theorem 2 and referenced in this paper as CSD-CLF, is presented in algorithm 3.**
Figure 2: Simple model for a downlink synchronous transmission of $K$ users. The structure of each array processor is shown in figure 3 below.

Figure 3: Adaptive array processor. Note that $d_i$ represents an estimate or replica of the desired signal for the $i^{th}$ user at the array output.

Algorithm 3 CSD-CLF method for linearly constrained adaptive filters

\begin{align*}
\text{Calculate } w_0 &= F = C(C^T C)^{-1} f \\
\mathbf{R}_0 &= P = I - C(C^T C)^{-1} C^T \\
g_0 &= 0 \\
\text{For } k = 0, 1, \ldots, N - 1 \\
\alpha_k &= \left( \frac{1}{\lambda} \text{sgn}(g_k) g_k \right) \\
w_{k+1} &= Pw_k + F - \alpha_k g_k \\
x_{k+1} &= P x_{k+1} \\
\mathbf{R}_{k+1} &= \lambda_f \mathbf{R}_k + x_{k+1} x_{k+1}^T \\
c_{k+1} &= d_{k+1} - x_{k+1}^T w_k \\
g_{k+1} &= \lambda_f g_k - \alpha_k c_{k+1} g_k - x_{k+1} c_{k+1} \\
\text{End}
\end{align*}

Also note that algorithm 3 has 2 steps and 7M FLOP less than the algorithm in Apolinário et al. (2000).

4 Simulation Results

In order to test the proposed algorithms, we compare the performance of the proposed CLF algorithms with the algorithms in Apolinário et al. (2000), applied to a single-user detection in a DS-CDMA mobile communication system. For this experiment, we assumed a simple model for a downlink synchronous transmission of $K$ users, as shown in figure 2 (Liberti and Rappaport, 1999). The received continuous-time signal is passed through a chip-matched filter and is sampled at a chip rate such that the received discrete time input-signal vector may be expressed as $x_k = S A c_k + n_k$ where $S = \begin{bmatrix} s_1 & s_2 & \cdots & s_K \end{bmatrix}$ is the spreading matrix containing the sampled spreading sequences of users 1 to $K$, $A = \text{diag} \left[ A_1, A_2, \ldots, A_K \right]$ contains the amplitudes of signals for each user, $c_k = \begin{bmatrix} c_{1k} & c_{2k} & \cdots & c_{Kk} \end{bmatrix}^T$ contains the information bits, and $n_k$ is the sampled noise sequence. For this example, the constraint is such that $C = s_1$, and $f = 1$. The number of users was set to $K = 5$, with Gold codes of length 7 used for spreading. The SNR for user one was made equal to 8 dB, and the relative power of interfering users was set to 20 dB, i.e., $10 \log(P_i/P_1) = 20$.

In order to compare the CLF algorithms with the algorithms in Apolinário et al. (2000), the same parameters are used. Figures 4 and 5 show the learning curves and the coefficient errors of the linearly constrained adaptive filters applied to a single-user detection in a DS-CDMA mobile communication system. It can be observed that the CLF methods proposed in this paper have convergence properties comparable with the degenerated CG algorithm used in Apolinário et al. (2000) and with the linearly constrained RLS algorithm. It is important to notice that, although it has a rate of convergence lower than that of the RLS algorithm in a noisy environment, the CSD-CLF method can be used in cases where the LMS algorithm (which has lower rate of convergence, higher steady state mean squared error and comparable computational cost) is applicable.

In order to compare the computational cost of the proposed algorithms with the CCG of Apolinário et al. (2000) and the CLMS, the number of basic floating point operations FLOP (addition, subtraction, multiplication and division) of the principal loop is determined. The results, presented in table 4, show that the CSD-CLF has less 4M FLOP than the CCG-CLF, which has less 3M FLOP than the CCG. The CLMS presents
a lower computational cost than the other algorithms because it does not compute the correlation matrix $R_{m+1}$ and the matrix/vector multiplication $R_{k+1}g_k$ or $R_{k+1}p_k$.

### 5 Conclusion

The control viewpoint (Bhya and Kaszkurewicz, 2003; Bhaya and Kaszkurewicz, 2006) reveals that the formulas for the CG parameters $\alpha_k$ and $\beta_k$ in Hestenes and Stiefel (1952, p.411,eqs 3:2a and 3:2b), actually enhance the robustness of the CG algorithm with respect to the standard equivalent choices $(\alpha_k = \langle g_k, g_k \rangle / \langle p_k, R_k p_k \rangle)$ and $(\beta_k = \langle g_{k+1}, g_{k+1} \rangle / \langle h_k, h_k \rangle)$, since the latter make use of the assumption $p_0 = g_0$, to obtain orthogonality of residues. Hestenes and Stiefel (1952, p. 411) recognize that “although these formulas are slightly more complicated, they have advantage that scale factors (introduced to increase accuracy) are more easily changed during the course of the computation”. Our results, both theoretical and simulated, show that since the “more complicated formulas” correspond to optimal choices in the CLF method, they endow the system with robustness, since there is some leeway corresponding to neighboring non-optimal choices of the parameters $\alpha_k$ and $\beta_k$. The authors in Apolinário et al. (2000) have presented a constrained CG method for linearly constrained adaptive filtering problems, however their convergence analysis was carried out by numerical simulations. This paper proposes an analysis using the linear structure of the time varying system (3) and the bilinear structure of the dynamical system solvers (22) and (34) considered, and the asymptotic stability of the algorithms is proved in theorems 1 and 2.

### Acknowledgment

This research was partially financed by Project Nos 141566/2004-3, 551863/2002-1 of CNPq and also by the agencies CAPES and FAPERJ. The authors would like to thank Profs. Campos and Apolinário for making their CCG software available.

### References


