CONE CONSTRAINED ADAPTIVE ALGORITHMS AND MULTIPLE ACCESS INTERFERENCE CANCELLATION

Miloš I. Doroslovački and Branimir R. Vojčić

The George Washington University
Department of Electrical Engineering and Computer Science
Washington, DC 20052, USA

ABSTRACT

A parameterization of the cone around a given vector in N-dimensional vector space is derived. Vectors obtained by changing the parameter values can not escape the cone. The constraint is useful when it is known that an optimal vector of filter coefficients is close in direction with the given vector. Adaptive filtering algorithms that use the steepest descent method and stochastic gradient approximation are developed for the case of filter coefficients constrained in the cone. There is no need for monitoring the constraint since it is always satisfied by the construction. Two optimization criteria are explicitly considered: the least mean square and constant modulus. The cone constrained constant modulus algorithm (CMA) is applied to the problem of user detection in a synchronous direct sequence code division multiple access system. Its convergence is compared with the plain CMA and back projection CMA. Under severe conditions the cone constrained CMA is the only one who locks to the desired user.

1. PROBLEM STATEMENT

Let $\mathbf{x}(t)$ be an N-dimensional vector of observations obtained in the tth time interval. In order to obtain a scalar estimate $y(t)$, the observables $\mathbf{x}(t)$ are linearly processed with a vector of weights $\mathbf{w}(t)$, so that

$$y(t) = \mathbf{w}^H(t)\mathbf{x}(t), \tag{1}$$

where H denotes the Hermitian transpose. The above equation describes a linear time-varying system. Furthermore, it is known that the direction of the vector $\mathbf{w}(t)$ is close to the direction of a given unit-norm vector $\mathbf{u}_1$. That is, the vector $\mathbf{w}(t)$ must be within a cone around the vector $\mathbf{u}_1$ in a M-dimensional space ($M \leq N$). The cone aperture is defined as $s = \tan(\theta)$, where θ is defined in Figure 1. The length of the vector $\mathbf{w}(t)$ is not restricted. One way of guaranteeing that the vector $\mathbf{w}(t)$ is within the cone, is to check for each t whether this is a case. If the vector is out of the cone, it should be projected back to the cone. In order to avoid testing of the condition for each t, it is possible to parameterize the coordinates of the vector $\mathbf{w}(t)$, in a way that the condition is always satisfied.

2. CONE PARAMETERIZATION

Let $\mathbf{U}$ be a unitary matrix of dimension $N \times N$. Also, let the first column of $\mathbf{U}$ be the vector $\mathbf{u}_1$ which defines the center of the cone. Moreover, assume that the first M columns of $\mathbf{U}$ define the M-dimensional space of the cone. The vector of projections of $\mathbf{w}(t)$ to the columns of $\mathbf{U}$ is given as

$$\mathbf{v}(t) = \mathbf{U}^H\mathbf{w}(t). \tag{2}$$

It can be parameterized in the following way:

$$\mathbf{v}(t) = \begin{bmatrix} p_1(t) \\ sp_1(t)\cos p_2(t)\sin p_3(t) \\ sp_1(t)\cos p_2(t)\cos p_3(t)\sin p_4(t) \\ \vdots \\ sp_1(t)\cos p_2(t)\cdots\cos p_i(t)\sin p_{i+1}(t) \\ \vdots \\ sp_1(t)\cos p_2(t)\cos p_3(t)\cdots\cos p_M(t) \\ p_{M+1}(t) \\ \vdots \\ p_N(t) \end{bmatrix}. \tag{3}$$

The first M components of the vector $\mathbf{v}(t)$ satisfy

$$\frac{\sum_{i=2}^{M} |v_i(t)|^2}{p_1^2(t)} = s^2\cos^2 p_2(t). \tag{4}$$

That is, the projection of $\mathbf{w}(t)$ to the M-dimensional space is within the cone specified by the vector $\mathbf{u}_1$ and the aperture s. The vector $\mathbf{v}(t)$ can be rewritten as

$$\mathbf{v}(t) = [\mathbf{v}_M^T(t)\mathbf{v}_{N-M}^T(t)]^T, \tag{5}$$

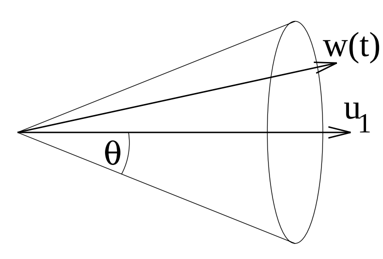

Figure 1: Cone around a given vector

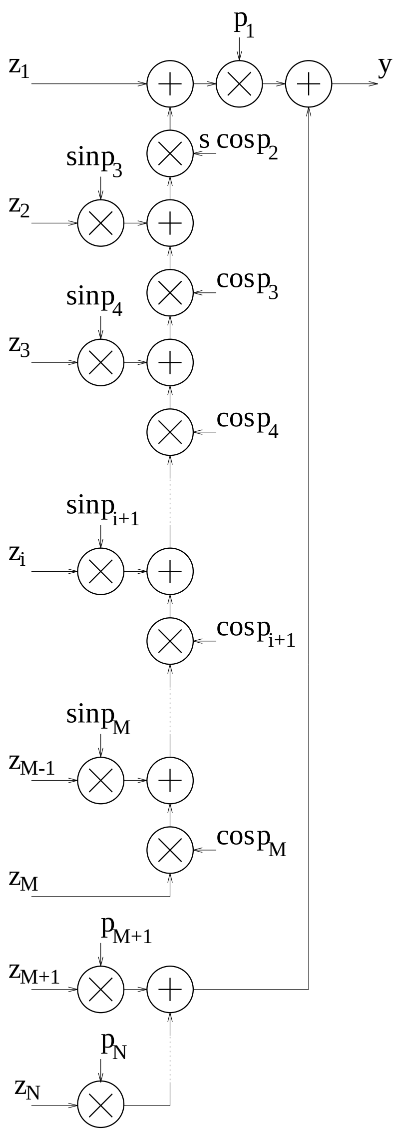

Figure 2: Cone-constraint processing scheme

where

$$\mathbf{v}_M(t) = p_1(t) \begin{bmatrix} 1 & \mathbf{0} \\ \mathbf{0} & s\cos p_2(t)\mathbf{I}_{M-1} \end{bmatrix}$$
$$\times \prod_{k=3}^{M} \begin{bmatrix} \mathbf{I}_{k-2} & 0 & \mathbf{0} \\ 0 & \sin p_k(t) & 0 \\ \mathbf{0} & 0 & \cos p_k(t)\mathbf{I}_{M-k+1} \end{bmatrix} \mathbf{1}_M, \quad (6)$$

and

$$\mathbf{v}_{N-M}(t) = [p_{M+1}(t) \cdots p_N(t)]^T. \quad (7)$$

Here $\mathbf{I}_k$ is the $k \times k$ identity matrix, and $\mathbf{1}_M$ is the M-dimensional vector having all entries equal to one. The estimate $y(t)$ can be written as

$$y(t) = \mathbf{v}^H(t)\mathbf{z}(t), \quad (8)$$

where $\mathbf{z}(t) = \mathbf{U}^H\mathbf{x}(t)$. The processing scheme corresponding to (8) is given in Figure 2.

3. ADAPTIVE PROCESSING

Typically, the linear processing (1) should be such that it is optimal in some sense. The weights $\mathbf{w}(t)$, i.e. the parameter vector $\mathbf{p}(t) = [p_1(t) \cdots p_N(t)]^T$, can be such that minimize the average value of an error function of the estimate $y(t)$, $E\{f(y(t))\}$. Using the steepest descent method and stochastic gradient approximation [1], the optimal parameter vector can be approached through the sequence of parameter vectors defined by:

$$\mathbf{p}(t+1) = \mathbf{p}(t) - \mu \frac{\partial f(y(t))}{\partial y(t)} (\nabla_{\mathbf{p}(t)} \mathbf{v}^H(t))\mathbf{U}^H\mathbf{x}(t), \quad (9)$$

where

$$\nabla_{\mathbf{p}(t)} \mathbf{v}^H(t) = \begin{bmatrix} \frac{\partial \mathbf{v}^H(t)}{\partial p_1(t)} \\ \frac{\partial \mathbf{v}^H(t)}{\partial p_2(t)} \\ \vdots \\ \frac{\partial \mathbf{v}^H(t)}{\partial p_N(t)} \end{bmatrix} \quad (10)$$

is the $N \times N$ matrix of derivatives. The matrix can be rewritten as

$$\nabla_{\mathbf{p}(t)} \mathbf{v}^H(t) = \begin{bmatrix} \nabla_{\mathbf{p}_M(t)} \mathbf{v}_M^H(t) & \mathbf{0} \\ \mathbf{0} & \mathbf{I}_{N-M} \end{bmatrix}. \quad (11)$$

The rows of the upper left submatrix in (11) are:

$$\frac{\partial \mathbf{v}_M(t)}{\partial p_1(t)} = \mathbf{1}^H \begin{bmatrix} 1 & \mathbf{0} \\ \mathbf{0} & s\cos p_2(t)\mathbf{I}_{M-1} \end{bmatrix}$$
$$\times \prod_{k=3}^{M} \begin{bmatrix} \mathbf{I}_{k-2} & 0 & \mathbf{0} \\ 0 & \sin p_k(t) & 0 \\ \mathbf{0} & 0 & \cos p_k(t)\mathbf{I}_{M-k+1} \end{bmatrix} \quad (12)$$

$$\frac{\partial \mathbf{v}_M(t)}{\partial p_2(t)} = \mathbf{1}^H p_1(t) \begin{bmatrix} 0 & \mathbf{0} \\ \mathbf{0} & -s\sin p_2(t)\mathbf{I}_{M-1} \end{bmatrix}$$
$$\times \prod_{k=3}^{M} \begin{bmatrix} \mathbf{I}_{k-2} & 0 & \mathbf{0} \\ 0 & \sin p_k(t) & 0 \\ \mathbf{0} & 0 & \cos p_k(t)\mathbf{I}_{M-k+1} \end{bmatrix} \quad (13)$$

$$\frac{\partial \mathbf{v}_M(t)}{\partial p_i(t)} = \mathbf{1}^H p_1(t) \begin{bmatrix} 1 & \mathbf{0} \\ \mathbf{0} & s\cos p_2(t)\mathbf{I}_{M-1} \end{bmatrix}$$
$$\times \prod_{\substack{k=3 \\ k \neq i}}^{M} \begin{bmatrix} \mathbf{I}_{k-2} & 0 & \mathbf{0} \\ 0 & \sin p_k(t) & 0 \\ \mathbf{0} & 0 & \cos p_k(t)\mathbf{I}_{M-k+1} \end{bmatrix}$$
$$\times \begin{bmatrix} \mathbf{0}_{i-2} & 0 & \mathbf{0} \\ 0 & \cos p_i(t) & 0 \\ \mathbf{0} & 0 & -\sin p_i(t)\mathbf{I}_{M-i+1} \end{bmatrix}. \quad (14)$$

One more motivation for using the cone constraint, is to help the adaptive algorithm to avoid the convergence to points corresponding to unwanted minima of $E\{f(y(t))\}$. In the case when the estimated quantity is known, e.g. when a training sequence $d(t)$ is used to help the adaptation, the error function whose average value is minimized, can take the form of

$$f(y(t)) = \frac{1}{2}|y(t) - d(t)|^2. \quad (15)$$

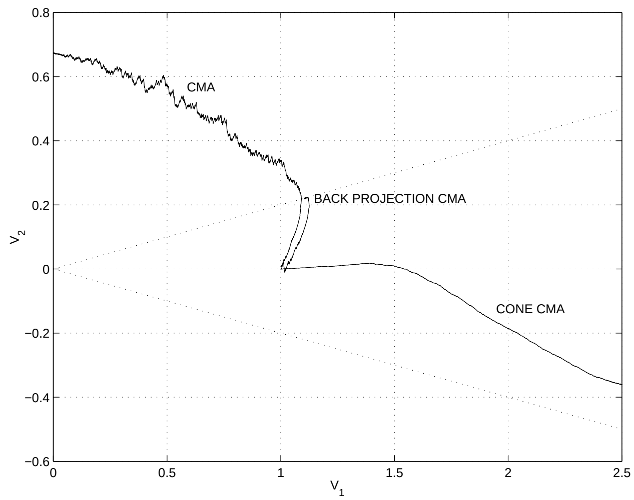

Figure 3: Weight trajectories (noiseless case)

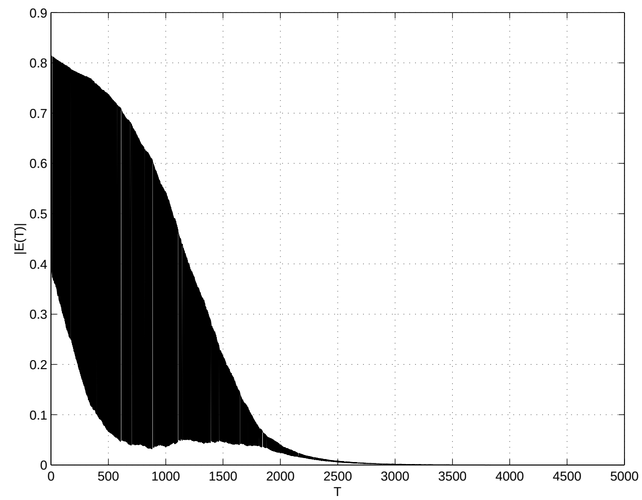

Figure 4: Bit estimation error for User-1 obtained by cone constrained CMA (noiseless case)

Then, in Equation (9)

$$\frac{\partial f(y(t))}{\partial y(t)} = (y(t) - d(t))^*, \tag{16}$$

where * denotes the complex conjugate operator. On the other hand, when it is known that the modulus of estimated quantity is constant, the error function can take the form of

$$f(y(t)) = \frac{1}{4}(|y(t)|^2 - \xi)^2. \tag{17}$$

Then, the derivative needed in Equation (9) is

$$\frac{\partial f(y(t))}{\partial y(t)} = (|y(t)|^2 - \xi)y^*(t). \tag{18}$$

4. INTERFERENCE SUPPRESSION IN DS-CDMA SYSTEMS

A received signal in the synchronous direct sequence code division multiple access (DS-CDMA) system can be represented in the following equivalent discrete-time form [2, 3]:

$$\mathbf{x}(t) = \mathbf{CAb}(t) + n(t), \tag{19}$$

where $\mathbf{C}$ is a $N \times M$ matrix whose columns are user code sequences. The $M \times M$ matrix $\mathbf{A}$ is diagonal, having as entries the attenuations for user signals. The vector $\mathbf{b}(t)$ is M-dimensional and has as entries user symbols that are transmitted in the tth time interval. The N-dimensional noise vector is denoted by $\mathbf{n}(t)$. In general, the user code sequences in $\mathbf{C}$ are not orthogonal. Even if they are orthogonal at the transmitter side, the corresponding user code sequences at the receiver side may not be, because of the intersymbol interference and difference in time lags for user signals. Also, the attenuations of the signals can be quite different and unknown.

Using the processing scheme (1), the $y(t)$ should provide an estimate of the transmitted symbol which corresponds to a specific user. The adaptive estimation tries to decrease the estimation error. That can be done by using a known training sequence and the error criterion (15). On the other hand, it can be done blindly by minimizing the criterion (17). The latter approach can be implemented by using the constant modulus algorithm (CMA) [4]. A problem with the CMA is that the receiver (i.e. the estimate $y(t)$) may lock to the signal of a wrong user, if the signal is much stronger than the signal of a desired user [5]. To alleviate the problem constraints can be applied to the weight adaptation using the CMA. In [6], the projection of the weight vector $\mathbf{w}(t)$ to the code sequence of a desired user $\mathbf{c}_1$, should be equal to one, i.e. $\mathbf{w}^H(t)\mathbf{c}_1 = 1$. Here, the application of the cone constraint is advocated. The cone is around a vector which is assumed appropriate for detection of symbols of a desired user. Typically that vector can be $\mathbf{c}_1$. Note that $\mathbf{w}(t) = \mathbf{c}_1$ corresponds to the correlation receiver. The cone aperture should be chosen narrow enough to prohibit locking to other users, but also wide enough to permit the adaptation to the signal of a desired user by minimizing the constant modulus criterion (17).

Let us consider a DS-CDMA system with only two users. User-1 and User-2 use for transmission the code sequences

$$\mathbf{c}_1 = [+1 + 1 + 1 - 1 - 1 + 1 - 1]^T/\sqrt{7}$$
$$\mathbf{c}_2 = [-1 - 1 + 1 - 1 + 1 + 1 + 1]^T/\sqrt{7},$$

respectively. User-1 is the desired user. The users transmit symbols ± 1. The attenuation of User-1 is 0.4 and of User-2 is 1.5, that is User-2 is 11.5dB stronger than User-1. The cone aperture is $s = 0.2$ and the cone dimension is $M = 2$. The vector $\mathbf{w}(k)$ which corresponds to the de-correlation receiver is within the cone. In general, the choice of the aperture should be based on the knowledge and/or predictions of the user code cross-correlations and the user power distribution. On a sequence of 5000 bits, we investigated the convergence of the three constant modulus adaptive algorithms: the plain CMA, back projection CMA, and cone constraint CMA. The back projection CMA checks whether $\mathbf{w}(k)$ is out of the cone and if yes, it orthogonally projects

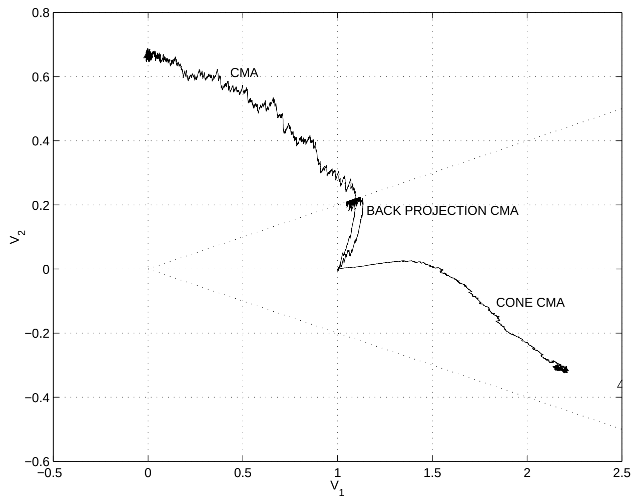

Figure 5: Weight trajectories (noisy case)

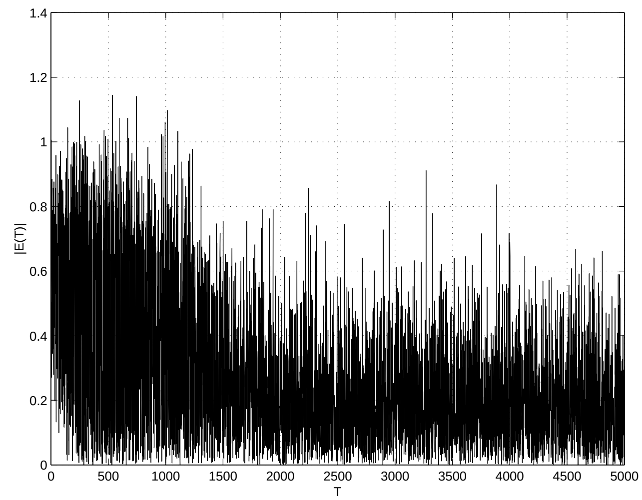

Figure 6: Bit estimation error for User-1 obtained by cone constrained CMA (noisy case)

the weight vector to the closes point in the cone. The matrix $\mathbf{U}$ is obtained by the QR decomposition of $[\mathbf{c}_1\mathbf{c}_2]$. The adaptation step in (9) is $\mu = 0.01$. Weight trajectories obtained by the three algorithms in the noiseless case are plotted in Figure 3. The horizontal coordinate of a trajectory point gives the projection of $\mathbf{w}(k)$ to $\mathbf{c}_1$ and the vertical coordinate the projection of $\mathbf{w}(k)$ to the second column of $\mathbf{U}$. Of course, the column is orthogonal to $\mathbf{c}_1$ and it is in the plane defined by $\mathbf{c}_1$ and $\mathbf{c}_2$. All algorithms start with $\mathbf{w}(1) = \mathbf{c}_1$. The plain CMA converged to the de-correlation receiver weights for User-2, i.e. it is locked to the unwanted user. The back projection CMA was stuck on the cone border. Only, the cone constrained CMA converged to the weights which provide the User-1 symbol detection. The weights correspond to the de-correlation receiver weights, what is an expected result in the no-noise case. The corresponding estimation error $|y(t) - b_1(t)|$ is given in Figure 4. Note that only errors larger than one can produce wrong bit decisions. Figure 4 shows that the cone constrained CMA provided an errorless User-1 symbol detection for all 5000 bits.

In the presence of a stationary white zero-mean Gaussian noise $\mathbf{n}(t)$, the weight trajectories for the three algorithms are given in Figure 5. The average noise energy is 0.07 and this corresponds to the User-1 signal-to-noise ratio of 3.6dB. All other conditions are same as in the noiseless simulation experiment. Again, the plain CMA was useless since it was locked to User-2. The back projection CMA was stuck on the cone border. The cone constrained CMA continued to serve User-1. The weights did not converged to the de-correlation receiver weights (denoted by a triangle in Figure 5). The estimation error obtained by the cone constrained CMA is plotted in Figure 6. As can be seen from the figure, at the beginning of adaptation the error is sometimes larger than one. That is, some bit decisions are wrong. Also, after the convergence is achieved, the decision errors are possible when very high noisy spikes occur. But that happens with lower probability. Of course, the bit error rate can be estimated through long-term simulations.

5. CONCLUSION

The derived parameterization of the cone around a given vector enabled the development of adaptive algorithms of training or blind type that automatically satisfy the cone constraint. Among them are the least mean square algorithm and constant modulus algorithm (CMA). The cone constraint helps the algorithm to avoid unwanted convergence points. Experiments show that under severe working conditions the cone constrained CMA provides the detection of a desired user in a DS-CDMA system, while the plain CMA and back projection CMA fail to do that.

6. REFERENCES

[1] S. Haykin. *Adaptive Filter Theory.* Prentice Hall, Englewood Cliffs, third edition, 1996.

[2] R. Lupas and S. Verdu. Linear multiuser detectors for synchronous code-division multiple access channels. *IEEE Trans. Inform. Theory*, 35(1):123–136, Jan. 1989.

[3] M. Honig, U. Madhow, and S. Verdu. Blind adaptive multiuser detection. *IEEE Trans. Inform. Theory*, 41(4):944–960, July 1995.

[4] J. R. Treichler and M. G. Larimore. The tone capture properties of cma-based interference suppressors. *IEEE Trans. Acoust., Speech, Signal Processing*, 33(4):946–958, Aug. 1985.

[5] W. Lee, B. R. Vojčić, and R. L. Pickholtz. Constant modulus algorithm for blind multiuser detection. In *Proc. IEEE Int. Symp. Spread Spectrum Techniques and Applications, vol. 3*, pages 1262–1266, Mainz, Germany, Sep. 1996.

[6] J. Míguez and Lois Castedo. A linearly constrained constant modulus approach to blind adaptive multiuser interference suppression. *IEEE Communications Letters*, 2, 1998. preprint.