

Coordinate Descent Iterations in Pseudo Affine Projection Algorithm

Felix Albu, Constantine Kotropoulos

Department of Informatics
Aristotle University of Thessaloniki
Thessaloniki 54124, Greece

felix_albu@ieee.org, costas@zeus.csd.auth.gr

Abstract— This paper introduces a new affine projection algorithm based on the dichotomous coordinate descent algorithm. A generalization for simplified Volterra filters is also made and its stability is improved by using a voice activity detector. The performance of the proposed algorithms for acoustic echo cancellation is assessed.

I. INTRODUCTION

In echo cancellation systems, an adaptive filter is used to reduce echo. The echo path is usually modeled by a linear filter. The well known normalized Least Mean Squares (NLMS) algorithm has been widely used, but it has a slow asymptotic convergence. The affine projection algorithm (APA) can be considered as a generalization of the NLMS algorithm that provides a much improved convergence speed compared to stochastic gradient descent algorithms [1]. It is known that its convergence performance lies between those of NLMS and the Recursive Least Squares (RLS) algorithm. However, the RLS algorithm and its fast versions are too complex to be used in acoustic echo cancellers [2]. For usual filter lengths used in AEC (250-4000 coefficients) RLS is at least 200 times more computationally expensive than any FAP algorithm.

Several fast affine projection (FAP) alternatives have been proposed (see [2-4] and the references therein). They are suitable for values of the step size close to 1 (e.g. $0.7 < \mu \leq 1$). A promising algorithm was proposed in [4]. The Gauss-Seidel Pseudo Affine Projection (GSPAP) algorithm is numerically less complex than other similar FAPs for typical values of the projection order. As in [5], we propose the replacement of the Gauss-Seidel method with a recently introduced dichotomous coordinate descent (DCD) algorithm [6]. It is a multiplier-free and division-free iterative technique for solving linear systems. The new algorithm will be abbreviated as DCDPAP algorithm.

The outline of the paper is as follows. In Section II, the DCDPAP algorithm is described and a generalization for simplified Volterra filters (SVF), abbreviated as DCDPAP-

SVF, is proposed in a similar way with that in [7]. Moreover, a voice activity detector (VAD) is employed in order to tackle the sensitivity of SVF structure. In Section III, the algorithm performance for echo cancellation and their complexity are examined. Section IV concludes the paper.

II. THE ALGORITHMS

Henceforth, most of the notations in [2] and [4] are used. Let $x(n)$ be the input signal, $y(n)$ be the desired output signal and $e(n)$ be the output error. We denote by L the filter length, and by N the affine projection order. In the following, $\mathbf{R}(n)$ is the auto-correlation matrix of the input signal, \mathbf{I} is the identity matrix, $\delta(n)$ is a regularization factor that prevents the input auto-correlation matrix from becoming ill-conditioned, and μ is the step size. Let $\mathbf{h}(n) = [h_1(n), \dots, h_L(n)]$ be the $L \times 1$ filter coefficient vector, $\mathbf{x}(n) = [x(n), \dots, x(n-L+1)]$, $\mathbf{a}(n) = [x(n), \dots, x(n-N+1)]$, $\mathbf{u}(n) = [u(n), \dots, u(n-L+1)]$ and $\mathbf{b} = [1, 0, \dots, 0]^T$. $\mathbf{R}(n)$ is updated by replacing its first row and first column by $\mathbf{r}(n)$ whose elements are $r_p(n)$, $p = 0, \dots, N-1$, with $r_p(n)$ denoting the p -th element of $\mathbf{r}(n)$ and moving down on the diagonal the left top $(N-1) \times (N-1)$ sub-matrix of $\mathbf{R}(n-1)$. $\lambda(n)$ is a variable and $\hat{\mathbf{u}}(n)$ is a tapped-delay line storing every sample of $u(n)$. Here are the equations of the DCDPAP algorithm obtained from the original GSPAP algorithm [4].

Initialization

$$\begin{aligned} \mathbf{h}(-1) &= \mathbf{0}, \mathbf{R}(-1) = \delta(-1)\mathbf{I}, \mathbf{u}(-1) = \mathbf{0} \\ \mathbf{a}(-1) &= \mathbf{0}, \mathbf{r}(-1) = [\delta(-1), \mathbf{0}]^T, \mathbf{P}(-1) = \mathbf{b} / \delta(-1) \end{aligned} \quad (1)$$

For each $n \geq 0$ and $p = 0, \dots, N-1$

$$r_p(n) = r_p(n-1) + x(n)x(n-p) - x(n-N+1)x(n-p-N+1) \quad (2)$$

$$e(n) = y(n) - \mathbf{x}^T(n)\mathbf{h}(n-1) \quad (3)$$

$$\text{update } \mathbf{R}(n) \text{ using } \mathbf{r}(n) \quad (4)$$

$$\bar{\mathbf{R}}(n) = \mathbf{R}(n) + \delta(n)\mathbf{I} \quad (5)$$

$$\text{Solve the linear system } \bar{\mathbf{R}}(n) \cdot \mathbf{P}(n) = \mathbf{b} \text{ with DCD [6]} \quad (6)$$

$$\mathbf{u}(n) = \frac{1}{\mathbf{P}_0(n)} \sum_{i=0}^{N-1} \mathbf{P}_i(n) \mathbf{x}(n-i) \quad (7)$$

$$\lambda(n) = \frac{\mu}{\hat{\mathbf{u}}^T(n)\mathbf{x}(n) + \delta(n)} \quad (8)$$

$$\mathbf{h}(n) = \mathbf{h}(n-1) + \lambda(n)\hat{\mathbf{u}}(n)e(n) \quad (9)$$

The DCD algorithm is based on a binary representation of elements of the solution vector with M_b bits. The complexity is limited by imposing a limited number of iterations, N_{it} . The elements of the solution vector are restricted in the amplitude range $[-H, H]$. It starts an iterative approximation of the solution vector $\mathbf{P}(n)$ from the most significant bit [6]. After the most significant bit has been found for all vector elements, the algorithm starts updating the next less significant bits. If a bit update happens the vector \mathbf{b} is changed [6]. If H is a power of 2, the multiplications are replaced by bit shifts. The algorithm has only shift and accumulate operations (SAC) and no divisions. It will be shown in the next section that the DCD method approximates very well the exact solution of the linear system given by (8) if enough iterations are executed. More details about the DCD algorithm can be obtained from [5-6]. The maximum complexity of DCD algorithms is

$$N(2N_{it} + M_b) \text{ SACs} \quad (10)$$

The regularization is a necessary part of the algorithm, especially in noisy conditions and double-talk scenarios. Often when there is near-end speech activity, SNR fall sharply well below 0 dB. We used a simple regularization method based on the approach described in [8-9]. Let $\rho_x(n)$ and $\rho_y(n)$ be the time-averaged powers of $x(n)$ and $y(n)$ respectively. The regularization parameter is $\delta(n) = \rho_x(n)$ if $\rho_x(n) > \gamma\rho_y(n)$ and $\delta(n) = 20L\rho_y(n)$ otherwise, with $\gamma = 8$.

However, it is known that low cost loudspeakers or microphones introduce nonlinear distortions, especially at high volumes. In such cases, the performance of a linear acoustic canceller degrades. A common approach is to consider nonlinear models [10-11]. Sometimes, in the case of the nonlinear distortions, SVFs can achieve a better system identification than a linear one can provide, but at the price of a much higher complexity. We consider memory-less non-linearities that have been proposed as a

reliable model for the nonlinearities that occur in amplifiers and loudspeakers [10]. The algorithm can be generalized for SVFs in the same way described in [7]. The computation of the input vector, approximated tapped-line vector and the generalized correlations are affected. We denote by O the order of the SVF. Let $k = 1, \dots, O$,

$$\mathbf{x}_k(n) = [x^k(n), \dots, x^k(n-L+1)], \quad (11)$$

$$\mathbf{h}_k(n) = [h_{1+(k-1)L}(n), \dots, h_{kL}(n)], \quad (12)$$

$$\mathbf{h}(n) = [\mathbf{h}_1(n), \dots, \mathbf{h}_k(n), \dots, \mathbf{h}_O(n)], \quad (13)$$

$$\mathbf{x}(n) = [\mathbf{x}_1(n), \dots, \mathbf{x}_k(n), \dots, \mathbf{x}_O(n)], \quad (14)$$

$$\hat{\mathbf{u}}(n) = [\hat{\mathbf{u}}_1(n), \dots, \hat{\mathbf{u}}_k(n), \dots, \hat{\mathbf{u}}_O(n)], \quad (15)$$

$$\mathbf{h}_k(n) = \mathbf{h}_k(n-1) + \frac{\mu}{\hat{\mathbf{u}}_k^T(n)\mathbf{x}_k(n) + \delta(n)} \hat{\mathbf{u}}_k(n)e(n) \quad (16)$$

The Eq. 18 replaces Eq. 2 as follows for $p = 0, \dots, N-1$

$$r_p(n) = r_p(n-1) + \sum_{o=1}^O x(n)^o x(n-p)^o - \sum_{o=1}^O x(n-N+1)^o x(n-p-N+1)^o \quad (17)$$

As shown in [7], the nonlinear parameters were adapted only during voiced sections. Therefore, a voice activity detection algorithm (VAD) is needed. Its performance requirements are not high. It should determine relatively well the voiced sections, since the stability of the algorithm is greatly improved by the regularization method. Many approaches for VAD were proposed. For example, there are algorithms based on the short-term energy, zero crossing rate, cepstral features, periodicity measures, statistical models, etc. (see [12] and the references therein). The energy-based approaches to speech detection suffer from a lack of robustness and fail to work well in a wide range of SNR, unless some care is taken. We used ITU-T G.729B used in speech coding standards [13]. The algorithm achieves a good performance, and it is relatively simple to implement. In integrated systems, some included speech coding functions can be re-used in this case.

III. SIMULATIONS

In our acoustic echo cancellation simulations the excitation signal is a speech signal, sampled at 8 kHz. The convergence of the algorithm was compared by using the squared norm of the difference between the car cabin impulse response (taken from [14]) and the adaptive filter coefficients in dB. From Figs. 1 and 2 it can be seen that the performance of DCDPAP is improved with the number of DCD iterations. The results for $N_{it} = 32$ were not plotted because they gave almost identical results to the exact solution. A performance very close to the ideal one is obtained for $N_{it} \geq 8$. The influence of the step size is also

visible. A lower step size leads to a slower convergence rate. However, the simulations have shown that the use of lower step sizes increases the robustness of the algorithm to noisy conditions.

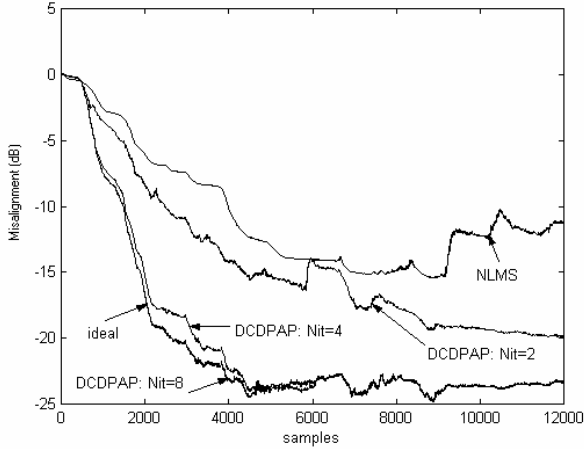


Figure 1. The misalignment curves for NLMS, DCDPAP ($N_{it} = 2$, $N_{it} = 4$, and $N_{it} = 8$) and ideal PAP algorithms ($SNR = 30dB$, $\mu = 1$, $L = 256$, $N = 5$).

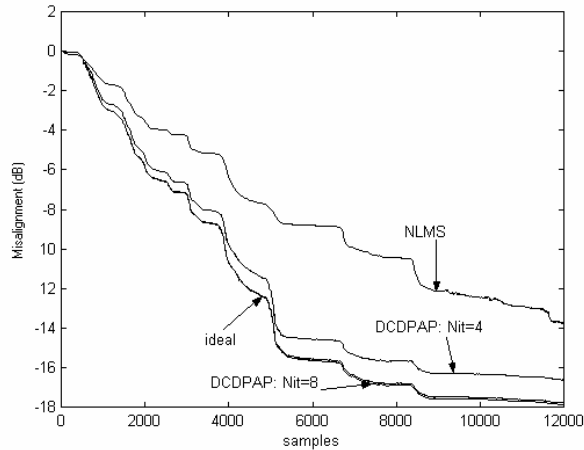


Figure 2. The misalignment curves for NLMS, DCDPAP ($N_{it} = 4$ and $N_{it} = 8$) and ideal PAP algorithms ($SNR = 30dB$, $\mu = 1/4$, $L = 256$, $N = 5$).

It can also be seen that more DCD iterations are needed for a low step size (cf. the misalignment curves for $N_{it} = 4$ in both Figs. 1 and 2). Moreover, the DCDPAP performance is better than that of NLMS. In all our simulations the performance of the GSPAP with one iteration was virtually identical to that of the ideal algorithm or DCDPAP with at least 8 iterations. This is confirmed by Fig. 3 where the error norm between the exact solution and the iterated GS or DCD solution is shown.

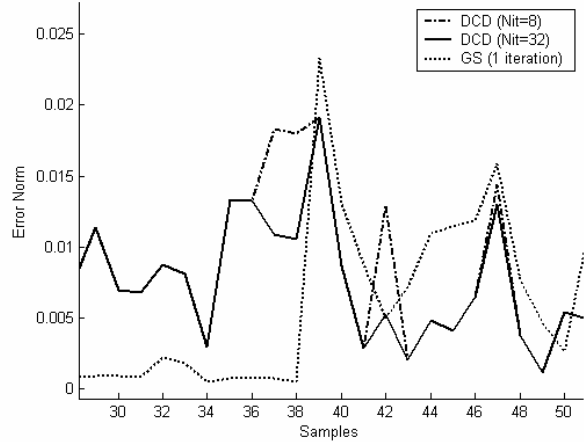


Figure 3. The error norm between the exact solution and the iterated solutions ($\mu = 1$).

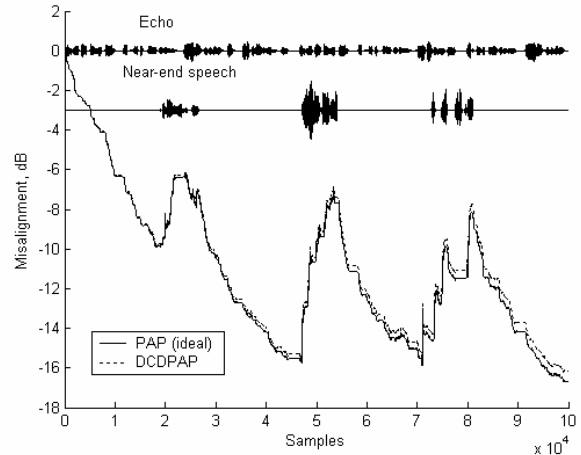


Figure 4. Comparison of the misalignment (in dB) provided by the ideal PAP and DCDPAP with time-varying regularization in a double-talk scenario for $SNR = 30 dB$, $\mu = 1/4$, $N_{it} = 32$, $L = 256$, $N = 5$.

According to (9) the peak DCD complexity for $N_{it} = 2$, $N_{it} = 4$, $N_{it} = 8$ and $N_{it} = 32$ are 100, 120, 160 and 400 SACs, respectively. However, the average complexity is about 30-40% smaller. All the other alternatives to DCD in solving the linear system require many multiplications and accumulate (MAC) operations and divisions (complicate to implement and a source of numerical instability). The DCDPAP algorithm complexity reported in multiplications and divisions per recursion (MADPR) is

$$C_{DCDPAP} = 2L + 3N + 5 \quad (18)$$

The behavior of the algorithm using the proposed regularization procedure is investigated in the double-talk scenario. It can be seen from Fig. 4 that DCDPAP with 8

iterations closely matches the performance of the ideal PAP algorithm.

Next, we examined the performance of the DCDPAP-SVF algorithm. We simulated the algorithm performance using test signals recorded in a car cabin. The echo return loss enhancement (ERLE) is computed in windows of 100 ms as follows

$$ERLE = 10 \log_{10} \frac{E[y^2(n)]}{E[e^2(n)]} \quad (19)$$

The results are similar with those obtained in [7]. Fig. 5a shows the VAD decision where 1 indicates speech frames, while 0 indicates silence frames. The simulations showed that the convergence is slightly slowed down when VAD is activated, but the stability of the algorithm is improved. The ERLE performance of the DCDPAP for SVF is lower in high volume level than in the normal volume level (Fig. 5c). Moreover, it can be seen from Fig. 5b that the difference between the NLMS and DCDPAP-SVF is rather small for a high volume level. These small benefits were obtained at the cost of much increased complexity than that of the linear algorithms. Also, these improvements depends on how well the non-linear and multi-path echo is modelled by the approach proposed in [7] and [10].

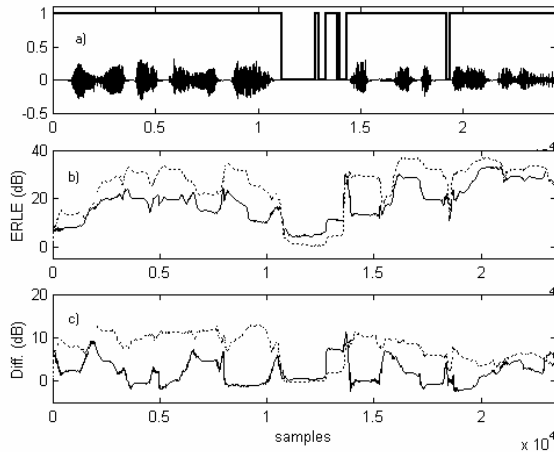


Figure 5. a) The speech signal, and the VAD decision; b) ERLE of the DCDPAP-SVF algorithm ($N = 2$, $N_{it} = 32$) on high (solid line) and normal (dotted line) volume levels; c) Improvement achieved over NLMS using a 3rd order SVF filter, in the same situations.

The complexity of the DCDPAP-SVF algorithm is much higher than that of DCDPAP or NLMS.

$$C_{DCDPAP-SVF} = 2 \cdot O \cdot L + (2 \cdot O + 1)N + O + 4 \quad (20)$$

For example, in the reported experiment it needed approximately 1578 MADPR, while NLMS needed only 516 MADPR, and DCDPAP needed 532 MADPR.

IV. CONCLUSIONS

We have proposed the dichotomous coordinate descent pseudo affine projection algorithm. We have demonstrated that the new algorithm has a fast convergence, low complexity and behaves well in a double-talk scenario. Also, we extended it for simplified Volterra filters. Therefore, the proposed algorithms represent interesting options for acoustic echo cancellation systems.

REFERENCES

- [1] K. Ozeki and T. Umeda, "An adaptive filtering algorithm using an orthogonal projection to an affine subspace and its properties," *Electronics and Communications in Japan*, vol. 67-A, no. 5, 1984.
- [2] S. L. Gay "Fast Projection Algorithms with Application to Voice Echo Cancellation", Ph.D. Dissertation, Rutgers The State University of New Jersey, New Brunswick, New Jersey, 1994.
- [3] F. Bouteille, P. Scalart, and M. Corazza, "Pseudo affine projection algorithm new solution for adaptive identification", in *Proc. Eurospeech 1999*, vol. 1, pp. 427-430.
- [4] F. Albu and A. Fagan, "The Gauss-Seidel pseudo affine projection algorithm and its application for echo cancellation," in *Proc. of Asilomar Conference*, Asilomar Grove, CA, U.S.A, November 2003.
- [5] Y. V. Zakharov and F. Albu, "Coordinate descent iterations in fast affine projection algorithm", *IEEE Signal Processing Letters*, accepted
- [6] Y. V. Zakharov and T. C. Tozer, "Multiplication-free iterative algorithm for LS problem," *Electronics Letters*, vol. 40, no. 9, pp. 567-569, April 2004.
- [7] R. Niemisto, T. Makela, and V. Mllyla, "Robust fast affine projection algorithm for nonlinear acoustic echo cancellation," in *Proc. XI European Signal Processing Conf.* pp. 523-526, Toulouse, France, September 2002.
- [8] V. Myllyla and G. Schmidt, "Pseudo-optimal regularization for affine projection algorithms," in *Proc. 2002 IEEE Int. Conf. Acoustics Speech, and Signal Processing*, vol. II, pp. 1917-1920, 2002.
- [9] H. Sheikhzadeh, R. L. Brennan, and K. R. L. Whyte, "Near-end distortion in over-sampled subband adaptive implementation of affine projection algorithm", in *Proc. XII European Signal Processing Conf.*, pp. 413-416, September 2004, Vienna, Austria.
- [10] A. Stenger and W. Kellerman, "Adaptation of a memory-less preprocessor for nonlinear acoustic echo cancelling", *Signal Processing*, vol. 80, pp. 1747-1760, 2000.
- [11] A. Fermo, A. Carini, and G. L. Sicuranza, 'Simplified volterra filters for acoustic echo cancellation in GSM receivers', in *Proc. X European Signal Processing Conf.*, vol. 4, pp. 2413-2416, September 2000.
- [12] S. G. Tanyer and H. Özer, "Voice activity detection in nonstationary noise," *IEEE Trans. Speech Audio Processing*, vol. 8, pp. 478-482, July 2000.
- [13] ITU-T Rec. G.729, Annex B, A silence compression scheme for G.729 optimized for terminals conforming to ITU-T V.70.
- [14] C. Breining, P. Dreitseitel, E. Hansler, A. Mader, B. Nitsch, H. Pudeer, T. Scheitler, G. Schmidt, and J. Tilp, "Acoustic echo control-An application of very high order adaptive filters", *IEEE Signal Processing Magazine*, pp. 42-69, July 1999