VARIABLE STEP SIZE DICHOTOMOUS COORDINATE DESCENT AFFINE PROJECTION ALGORITHM

Felix Albu, *Member IEEE*, Constantin Paleologu, *Member IEEE*, Jacob Benesty, *Senior Member IEEE* and Yuriy V. Zakharov, *Senior Member IEEE*

Abstract: **A new affine projection (AP) algorithm based on dichotomous coordinate descent (DCD) iterations has been recently proposed for acoustic echo cancellation (AEC). It uses a constant step size parameter and, therefore, has to compromise between fast convergence and tracking on the one hand, and low misadjustment and robustness to the presence and variations of a near-end signal on the other hand. In this paper we propose a variable step-size (VSS) version of the DCD-AP algorithm (VSS-DCD-AP) that does not require any a priori information about the acoustic environment. It is shown that the new algorithm is robust against near-end signal variations, including double-talk (DT).**

***Index Terms*: dichotomous coordinate descent algorithm, affine projection algorithm, acoustic echo cancellation, double-talk**

I. INTRODUCTION

In echo cancellation systems, an adaptive filter is used to reduce the echo. The echo path is usually modeled by a linear filter. The well known normalized least-mean-square (NLMS) algorithm has been widely used, but converges very slowly. The affine projection algorithm (APA) can be considered as a generalization of the NLMS algorithm that provides a much improved convergence speed compared to LMS-type algorithms, although it is sensitive to high level of noise [1]. It has a performance that rivals with the more complex recursive least-squares (RLS) algorithms in many situations. However, the fast affine projection (FAP) algorithm proposed in [2] suffers from numerical instability when implemented with an embedded fast RLS algorithm. A key element in other proposed FAP algorithms is the approach to solve the encountered linear system. The choice of the approach (i.e., direct or iterative) determines the stability and robustness of the FAP algorithm. Several proposed FAP algorithms use an approximation that leads to simpler algorithms if the step size is $\mu = 1$ (non-relaxed case) or close to 1 (e.g. $0.7 < \mu \le 1$) [3-6]. For such values, these algorithms have a fast convergence, but they exhibit a high sensitivity to noisy inputs. They employ some approximations that degrade the performance of the original APA. A low complexity implementation of the APA has been recently proposed in [7]. It uses a novel recursive filtering technique and filtering update that is incorporated in the dichotomous coordinate descent method (DCD), originally proposed in [8] and enhanced in [9]. This leads to an important reduction in the number of multiplications needed by the AP algorithm. However, the version presented in [7] has a fixed step size and a variable step size version could be a more reliable solution in case of near-end signal variations, including double-talk. Such a solution has been presented in [10] and the VSS-APA algorithm proved to be more robust in adverse conditions than the APA. The same step size computational method can be adapted to the less computationally demanding DCD-AP algorithm.

The outline of the paper is as follows. The VSS-DCD-AP algorithm is described in Section II. In the same section, the computational complexity of the proposed algorithm is derived and compared with that of VSS-APA. In Section III, the behavior of VSS-DCD-AP algorithm for echo cancellation in single-talk and double-talk scenarios is examined. A comparison of the proposed algorithm with VSS-APA is performed. Section IV concludes the paper.

II. THE VSS-DCD-AP ALGORITHM

Let us follow the notation used in deriving the DCD-AP (see [7]): L is the filter length, K is the projection order, γ a regularization parameter, λ a forgetting factor; x_n, d_n, $\hat{\mathbf{h}}_n$, $\boldsymbol{\mu}_n = \left[\mu_n^0, \ldots, \mu_n^{K-1}\right]$, are the excitation signal, desired signal, an $L \times 1$ vector of adaptive filter taps, and variable step size vector respectively at time instant n; $\mathbf{X}_n = \left[\mathbf{x}_n \; \mathbf{x}_{n-1} \ldots \mathbf{x}_{n-K+1}\right]^T$ where $\mathbf{x}_n = \left[x_n \; x_{n-1} \ldots x_{n-L+1}\right]^T$; $\mathbf{d}_n = \left[d_n \; d_{n-1} \ldots d_{n-K+1}\right]^T$; $\mathbf{I}_K$ is a $K \times K$ identity matrix and $(\cdot)^T$ denotes the matrix transpose. The DCD algorithm is used to solve the linear system $\mathbf{R}_n \boldsymbol{\varepsilon} = \mathrm{diag}\left\{\mu_n^0, \ldots, \mu_n^{K-1}\right\} \mathbf{e}_n$ (see Table 1). The original DCD algorithm updates a solution of a linear system of equations in directions of Euclidian coordinates in the cyclic order and with a step size α that takes one of M_b (number of bits) predefined values corresponding to a binary representation bounded by an interval $\left[-H, H\right]$ [8,9]. The algorithm starts the iterative search from the most significant bits of the solution and continues until the least significant bits were updated. The algorithm complexity is limited by N_u, the maximum number of "successful" iterations. More details about this original DCD version can be found in [8]. A more efficient DCD version was proposed in [9]. This new version finds a 'leading' (pth) element of the solution to be updated (step 11 of Table 1). It can be seen from Table 1 that the filtering update is incorporated in the DCD procedure (step 17). More details about the DCD-AP algorithm can be found in [7]. The step sizes are computed as in [10] (see steps 6-9 from Table 1).

Table 1. The VSS-DCD-AP algorithm.

Step	Equation
	Initialization $\hat{\boldsymbol{\varepsilon}}_{-1} = 0, \hat{\mathbf{h}}_{-1} = 0,\ \hat{\sigma}^2_{d,-1} = 0,\ \hat{\sigma}^2_{y,-1} = 0,$ $\hat{\sigma}^2_{e_{l+1},-1} = 0$, for $l = 0..K-1$ $\mathbf{x}_{-1} = 0$, and $\mathbf{y}_{-1} = 0$
	For $n = 0,1,\ldots$
1	$\mathbf{z}_n = \left[\mathbf{x}_n^T \hat{\mathbf{h}}_{n-2} \quad \mathbf{y}_{n-1}^T (0 : K-2)\right]^T$
2	Calculate $\mathbf{G}_n = \mathbf{X}_n \mathbf{X}_{n-1}^T$
3	$\mathbf{y}_n = \mathbf{z}_n + \mathbf{G}_n \hat{\boldsymbol{\varepsilon}}_{n-1}$
4	$\mathbf{e}_n = \mathbf{d}_n - \mathbf{y}_n$
5	Calculate $\mathbf{R}_n = \mathbf{X}_n \mathbf{X}_n^T + \gamma \mathbf{I}_K$
6	$\hat{\sigma}^2_{d,n} = \lambda \hat{\sigma}^2_{d,n-1} + (1-\lambda) d_n^2$
7	$\hat{\sigma}^2_{y,n} = \lambda \hat{\sigma}^2_{y,n-1} + (1-\lambda) y_n^2$
	For $l = 0$ to $K-1$
8	$\hat{\sigma}^2_{e_{l+1},n} = \lambda \hat{\sigma}^2_{e_{l+1},n-1} + (1-\lambda) e^2_{l+1,n}$
9	$\mu_n^l = \left\lvert 1 - \dfrac{\sqrt{\hat{\sigma}^2_{d,n-l} - \hat{\sigma}^2_{y,n-l}}}{\xi + \hat{\sigma}_{e_{l+1},n}} \right\rvert$
10	$\mathbf{r} = \mathrm{diag}\left\{\mu_n^0, \ldots, \mu_n^{K-1}\right\} \cdot \mathbf{e}_n,$ $\alpha = H/2, m = 1, \hat{\boldsymbol{\varepsilon}}_n = 0$
	For $k = 1, \ldots, N_u$
11	$p = \arg\max_{i=0,\ldots,K-1} \left\{ \lvert r^i \rvert \right\}$
12	while $\lvert r^p \rvert \le (\alpha/2)[\mathbf{R}_n]_{p,p}$ and $m \le M_b$
13	$m = m+1, \alpha = \alpha/2$
14	if $m > M_b$, go to step 1
15	$\hat{\varepsilon}_n^p = \hat{\varepsilon}_n^p + \mathrm{sign}(r^p)\alpha$
16	$\mathbf{r} = \mathbf{r} - \mathrm{sign}(r^p)\alpha \mathbf{R}_n^{(p)}$
17	$\hat{\mathbf{h}}_n = \hat{\mathbf{h}}_{n-1} + \mathrm{sign}(r^p)\alpha \mathbf{x}_{n-p}$
Total: $L + K^2 + 6K + 8$ Mults $+ K$ div $+ K$ sqrt $+$ $L(N_u+1) + K^2 + K(2N_u+7) + M_b + 3$ adds	

The step-size equations of the VSS-APA and proposed VSS-DCD-AP algorithm do not depend explicitly on the near-end signal, even though they were derived by taking into account its presence; consequently, a robust behaviour under near-end signal variations (e.g., background noise variations and double-talk) is expected. Moreover, since only the parameters available from the adaptive filter are required and there is no need for a priori information about the acoustic environment, the proposed algorithm is easy to control in practice [10].

The total computational complexity of the APA is $2KL + P_m + 3K$ multiplications and $2KL + P_m + 2K$ additions, where $P_m = O(K^3)$ and $P_a = O(K^3)$ [7]. The VSS part adds $6 + 3K$ multiplications, $2 + 4K$ additions, K square roots and K divisions. Table 2 shows the number of multiplications and additions of the investigated algorithms for different L, K and N_u. It can be seen that the number of multiplications of the VSS-DCD-AP algorithm is several times smaller than that of the VSS-AP algorithm.

Table 2. The number of multiplications and additions of the investigated algorithms for different L, K, and N_u.

Algorithm	L	K	+	×
VSS-APA	512	4	4186	4190
		8	8754	8758
	1024	4	8282	8286
		8	16946	16950
VSS-DCD-AP	512	4	$N_u = 1 \rightarrow 1095$	560
			$N_u = 4 \rightarrow 2655$	
			$N_u = 8 \rightarrow 4735$	
		8	$N_u = 1 \rightarrow 1179$	632
			$N_u = 4 \rightarrow 2763$	
			$N_u = 8 \rightarrow 4875$	
	1024	4	$N_u = 1 \rightarrow 2119$	1072
			$N_u = 4 \rightarrow 5215$	
			$N_u = 8 \rightarrow 9343$	
		8	$N_u = 1 \rightarrow 2203$	1144
			$N_u = 4 \rightarrow 5323$	
			$N_u = 8 \rightarrow 9483$	

Simulation results shown in the next section revealed that, in most cases, VSS-DCD-AP algorithm performance matches that of VSS-APA algorithm for $N_u \geq 4$. It can be seen in Table 2 that, for $N_u = 4$, the VSS-DCD-AP algorithm is less complex than the VSS-APA in terms of the total number of additions and multiplications. The same conclusion can be obtained if $N_u = 8$ and K=8. Therefore, longer filter lengths, higher projection orders can be used by VSS-DCD-AP algorithm for a similar complexity with the VSS-APA algorithm but with improved performance. The number of additions of the VSS-DCD-AP algorithm increases linearly with N_u and therefore, for smal projection orders (e.g. K=2 and high N_u (e.g. $N_u > 8$), it could be higher than that of the VSS-APA algorithm.

III. SIMULATION RESULTS

The simulations were performed in an AEC context and the VSS-APA and VSS-DCD-AP were compared. The length of the adaptive filter is set to 512 coefficients. The measured impulse response of the acoustic echo path is plotted in Fig. 1(a) (the sampling rate is 8 kHz); its entire length has 1024 coefficients. This length is truncated to the first 512 coefficients for a first set of experiments performed in an exact modeling case. Then, the entire length of the acoustic impulse response is used for a second set of experiments performed in the under-modeling case [10].

The simulations are performed in an exact modeling scenario ($N = L = 512$), and in an under-modeling scenario, using the entire acoustic impulse response from Fig. 1(a), while the length of the adaptive filter remains the same ($N = 1024, L = 512$). The performance for the first scenario is evaluated in terms of the normalized misalignment (in dB), defined as $20\log_{10}\left(\left\|\mathbf{h} - \hat{\mathbf{h}}_n\right\| / \left\|\mathbf{h}\right\|\right)$. In the second scenario, the expression of the normalized misalignment is evaluated by padding the vector of the adaptive filter coefficients with $N - L$ zeros, i.e., $20\log_{10}\left(\left\|\mathbf{h} - \left[\hat{\mathbf{h}}_n \; \mathbf{0}_{N\text{-}L}^T\right]\right\| / \left\|\mathbf{h}\right\|\right)$. An alternative to compare the performance of the algorithms is to use the Echo Return Loss Enhancement (ERLE). It is a standard parameter of adaptive filter quality evaluation in echo cancellation. However, our simulations have shown that the ERLE performances of the investigated algorithms are close. Four ERLE curves cannot be individually discerned in most plots and therefore, the ERLE plots were not included in this paper.

The regularization factor for both algorithms is $\gamma = 25K\sigma_x^2$ and λ is the forgetting factor chosen as $\lambda = 1 - \frac{1}{6 \cdot L}$ [10]. In all experiments the parameter M_b was set to 16 and $H = 2^{-10}$. In the following experiments, in order to approach the context of typical AEC applications, only the speech sequence from Fig. 1(b) will be used as the far-end signal.

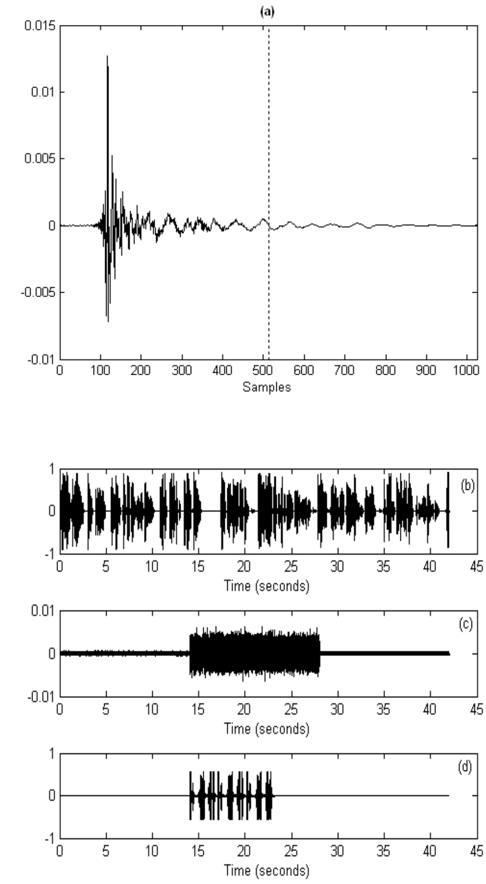

Fig. 1. (a) Measured room acoustic impulse response. Its length is truncated to the first 512 coefficients (before the dotted line) for the set of experiments performed in an exact modeling case (L = N = 512). The entire length is used for the set of experiments performed in the under-modeling case (L = 512, N = 1024); (b) Far-end continuous real speech signal used in the experiments; (c) The background noise with variable SNR; (d) Near-end speech signal used in the experiments performed in the double-talk case.

Single-talk and double-talk scenarios are considered.

The DCD procedure requires the execution of a number of iterations in order to obtain an acceptable accuracy in the solution of the linear system. All experiments confirmed that the performance of the VSS-DCD-AP algorithm is increasing when increasing the parameter N_u. In most cases, using more DCD iterations in the VSS-DCD-AP algorithm give closer performance with the more complex VSS-APA algorithm.

- *Single-talk scenario*

For the first set of simulations (Figs. 2 and 3) the value of the projection order was $K = 4$. Fig. 2 shows the misalignment curve in case of exact modeling while Fig. 3 considers the under-modelling case.

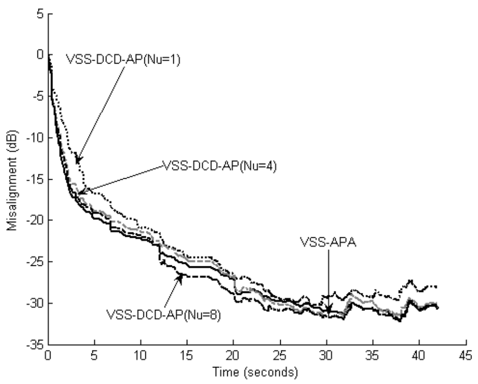

Fig. 2 Misalignment of VSS-APA and VSS-DCD-AP. Single-talk case, *L* = 512, *K* = 4, SNR = 30dB.

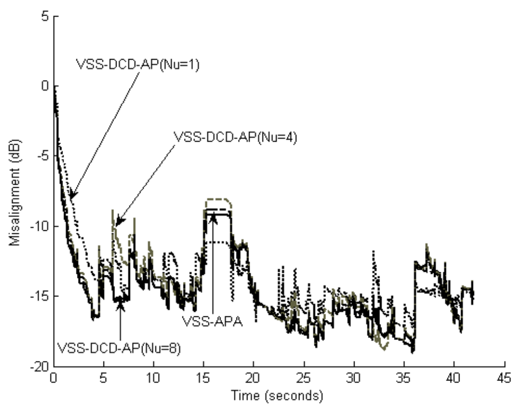

Fig. 3 Misalignment of VSS-APA and VSS-DCD-AP. Single-talk case, *L* = 1024, *K* = 4, SNR = 30dB (under-modeling case).

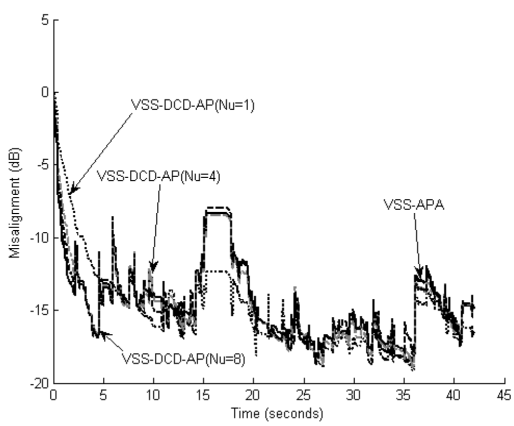

Fig. 4 Misalignment of VSS-APA and VSS-DCD-AP. Single-talk case, *L* = 1024, *K* = 8, SNR = 30dB (under-modeling case).

The robustness of the VSS-DCD-AP algorithm in the under-modelling case and its close performance to VSS-APA ($N_u \geq 4$) is verified. There are insignificant differences between the initial convergence rates in all three figures. Fig. 4 shows the convergence speed and misalignment that can be obtained with a higher projection order ($K = 8$).

A variation of the background noise shown in Fig. 1c is considered in Figs. 5 and 6. The SNR decreases from 30 dB to 10 dB at 14th second for a period of 14 seconds. The behavior of the algorithms is evaluated in the exact modeling case (Fig. 5) and under-modeling case (Fig. 6), for the projection order $K = 4$. It can be noticed that the proposed algorithm matches the VSS-APA algorithm in this situation if $N_u = 8$ and $N_u = 4$.

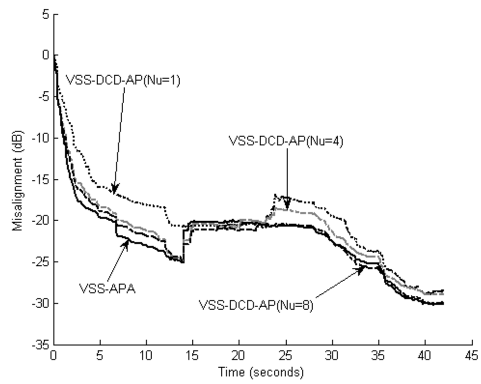

Fig. 5 Misalignment of VSS-APA and VSS-DCD-AP. Background noise variation at 14 s for a period of 14 seconds (SNR decreases from 30 dB to 10 dB). Other conditions are the same as in Fig. 2.

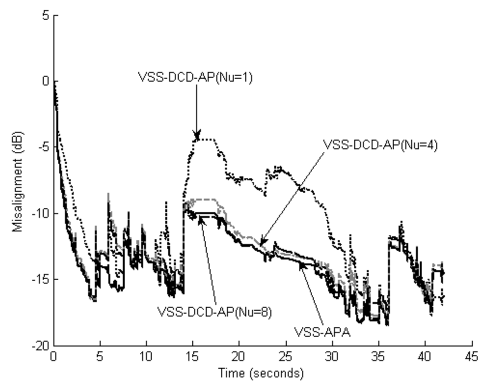

Fig. 6 Misalignment of VSS-APA and VSS-DCD-AP. Background noise variation at 14 s for a period of 14 seconds (SNR decreases from 30 dB to 10 dB). Other conditions are the same as in Fig. 3.

Another possible scenario in AEC is the change of the acoustic echo path. The results of such an experiment are depicted in Figs. 7 and 8, where the acoustic impulse response was shifted to the right by 12 samples after 21 seconds from the debut of the adaptive process.

Fig. 7 shows the results in the exact modeling case, while Fig. 8 shows the results in the under-modeling case. It can be noticed the loss in performance in the under-modeling case for both the algorithms. However, their tracking capabilities are good. The VSS-DCD-AP algorithm tracking ability increases with N_u. Other simulations (not shown here) have indicated that the tracking reaction of the VSS-DCD-AP algorithm improves as the projection order increases. This finding is consistent with the conclusion of previous works on VSS-APA [10].

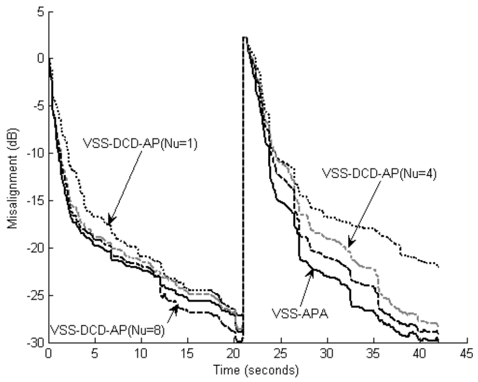

Fig. 7 Misalignment of VSS-APA and VSS-DCD-AP. Echo path change at 21 s. Other conditions are the same as in Fig. 2.

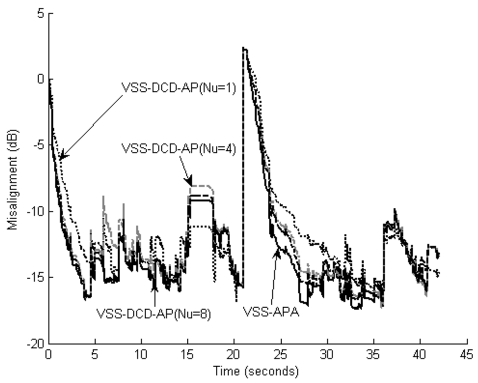

Fig. 8 Misalignment of VSS-APA and VSS-DCD-AP. Echo path change at 21 s. Other conditions are the same as in Fig. 3.

- *Double-talk scenario*

The most challenging problem in echo cancellation is considered to be the double-talk situation. Such a scenario is considered in the simulations using the speech signals from Fig. 1. In Figs. 9 and 10, the VSS-APA and VSS-DCD-AP algorithms are involved without the use of a DTD and $K = 4$. It is confirmed that a double talk detector is needed. The VSS-APA is the most affected. Also, the misalignment jump is

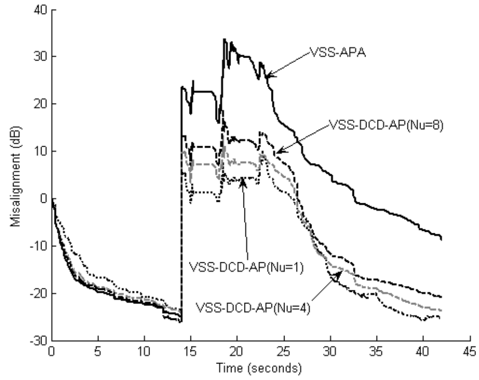

Fig. 9 Misalignment of VSS-APA and VSS-DCD-AP. Double-talk case and no DTD used. Other conditions are the same as in Fig. 2.

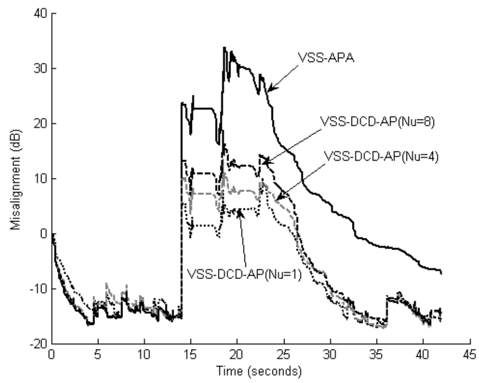

Fig. 10 Misalignment of VSS-APA, and VSS-DCD-AP. Double-talk case and no DTD used. Other conditions are the same as in Fig. 3.

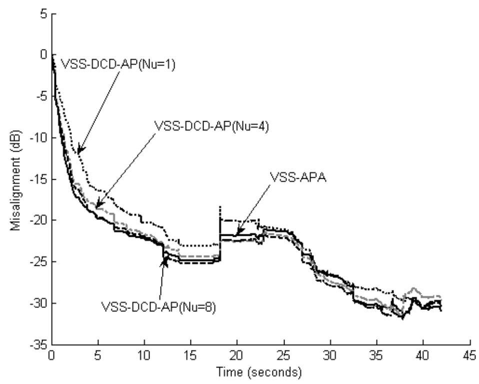

Fig. 11 Misalignment of VSS-APA and VSS-DCD-AP. Double-talk case and Geigel DTD used. Other conditions are the same as in Fig. 2.

higher as N_u increases. As shown in [10], a simple DTD (Geigel algorithm) is good enough for the VSS-

APA algorithm. Its settings are chosen assuming 6dB attenuation, i.e., the threshold is equal to 0.5 and the hangover time is set to 240 samples [12]. In Figs. 11 and 12, the VSS-APA and VSS-DCD-AP algorithms are used in conjunction with a Geigel DTD. It can be noticed from Figs. 11 and 12 that the robustness to double talk situation is greatly improved and the conclusions regarding the convergence performance are the same as in the other experiments in both exact modeling and under-modeling cases.

IV. CONCLUSION

A VSS-DCD-AP suitable for AEC applications has been derived in this paper. A variable step size was used in order to take into account the existence and the non-stationarity of the near-end signal as well as the under-modeling noise. Computation of the variable step-size requires no additional parameters from the acoustic environment. The simulation results performed in an AEC context showed its suitability in practice due to robustness to near-end signal variations like the increase of the background noise or double-talk. Concerning the last scenario, the VSS-DCD-AP can be combined with a simple Geigel DTD in order to enhance its performance. The proposed algorithm is much less computationally complex than the VSS-APA.

Acknowledgment: This work was supported by the UEFISCSU under Grant PN-II no. 331 / 01.10.2007
The authors want to thank reviewers for their comments that helped improve this paper.

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 "A fast converging, low complexity adaptive filtering algorithm", *Third International Workshop on Acoustic Echo Control*, Plestin les Greves, France, 1993, pp. 223-226.

[3] Q.G. Liu, B. Champagne, and K. C. Ho, "On the use of a modified FAP algorithm in subbands for acoustic echo cancellation," in *Proc. 7th IEEE DSP Workshop*, Loen, Norway, 1996, pp. 2570-2573

[4] H. Ding, "A stable fast affine projection adaptation algorithm suitable for low-cost processors", *ICASSP* 2000, Turkey, pp. 360-363

[5] F. Albu, J. Kadlec, N. Coleman, and A. Fagan, "The Gauss-Seidel fast affine projection algorithm," in *Proc. IEEE SIPS 2002*, pp. 109 - 114, San Diego, U.S.A, October 2002.

[6] Y. Zakharov, F. Albu, Coordinate descent iterations in fast affine projection algorithm, *IEEE Signal Processing Letters*, Vol. No. 5, May 2005, pp. 353-356.

[7] Y. Zakharov, "Low complexity implementation of the affine projection algorithm", *IEEE Signal Processing Letters*, vol. 15, pp. 557-560

[8] Y. Zakharov, and T. Tozer, "Multiplication-free iterative algorithm for LS problem", *Electron. Lett.*, 2004, 40, (9), pp. 567-569

[9] Y. Zakharov, G. White, and J. Liu, "Low complexity RLS algorithms using dichotomous coordinate descent iterations," *IEEE Trans. Signal Processing*, vol.56. No.7, pp.3150—3161, July 2008

[10] C. Paleologu, J. Benesty, and S. Ciochina, "Robust variable step-size affine projection algorithm suitable for acoustic echo cancellation", *Proc. Eusipco* 2008, Laussane, Switzerland

[11] J. Benesty, H. Rey, L. Rey Vega, and S. Tressens, "A nonparametric VSS NLMS algorithm," *IEEE Signal Process. Lett.*, vol. 13, no. 10, pp. 581–584, Oct. 2006.

[12] J. Benesty, T. Gaensler, D. R. Morgan, M. M. Sondhi, and S. L. Gay, *Advances in Network and Acoustic Echo Cancellation*. Berlin, Germany: Springer-Verlag, 2001.

How to cite this paper if used in research publications:

F. Albu, C. Paleologu, J. Benesty, and Y. V. Zakharov, "Variable Step Size Dichotomous Coordinate Descent Affine Projection Algorithm", EUROCON 2009, Saint Petersburg, Russia, May 2009, pp. 1366-1371

Websites with some documentation on related area:
http://falbu.50webs.com and
http://falbu.50webs.com/dcd.htm