A NOVEL PERSPECTIVE ON STEREOPHONIC ACOUSTIC ECHO CANCELLATION

Cristian Stanciu[†]*, Jacob Benesty*[⋆]*, Constantin Paleologu*[†]*, Tomas Gaensler*[‡]*, and Silviu Ciochină*[†]

[†] University Politehnica of Bucharest, Romania, e-mail: {cristian, pale, silviu}@comm.pub.ro
[⋆] INRS-EMT, University of Quebec, Montreal, Canada, e-mail: benesty@emt.inrs.ca
[‡] mh Acoustics, Summit, NJ, USA, e-mail: tfg@mhacoustics.com

ABSTRACT

The stereophonic acoustic echo is due to the coupling between two loudspeakers and two microphones. In the classical approach, this configuration is modelled by a two-input/two-output system with real random variables. In this paper, we propose to redesign this scheme as a single-input/single-output system with complex random variables. In this framework, we illustrate the behavior of some basic adaptive algorithms and present a distortion method which is more suitable for this model.

Index Terms— Stereophonic acoustic echo cancellation (SAEC), widely linear (WL) model, nonlinear distortion, adaptive filters.

1. INTRODUCTION

In hands-free teleconferencing systems, stereo transmission provides telepresence thanks to our binaural hearing system. These stereophonic systems give a realistic presence that actual single-channel systems cannot offer [1], [2]. In this context, stereophonic acoustic echo cancellation (SAEC) is necessary for full-duplex quality communication. For each microphone in the receiving (i.e., near-end) location, the SAEC consists of the identification of a two-input unknown system, consisting of the parallel combination of two acoustic echo paths (from the two loudspeakers to the microphone). Therefore, in the usual approach, an SAEC system consists of four adaptive filters aiming at identifying four echo paths from two loudspeakers to two microphones.

Despite the inherent similarities, SAEC is fundamentally different (and also more difficult) as compared to single-channel acoustic echo cancellation. The main challenge of SAEC is that the two channels may carry linearly related signals, which in turn may make the normal equation to be solved by the adaptive algorithm singular. This implies that there is no unique solution to the equation (as in the single-channel case) but an infinite number of solutions [3]. It was demonstrated that the only practical solution to the nonuniqueness problem is to reduce the coherence between the input (loudspeaker) signals [4]. Consequently, we need to distort these signals but without affecting too much the stereo perception and the sound quality.

In this paper, we propose a different approach for SAEC by recasting the classical two-input/two-output scheme with real random variables as a single-input/single-output system with complex random variables. In this framework, we present some basic adaptive algorithms and also a nonlinear distortion method which could be more suitable in this context.

This work was supported under the Grant POSDRU/107/1.5/S/76903 and Grant UEFISCDI PN-II-RU-TE no. 7/5.08.2010.

2. A NOVEL MODEL FOR SAEC

The stereophonic acoustic echo is usually modelled by a two-input/two-output system. In this classical setup, we have two input or loudspeaker signals denoted by $x_{\mathrm{L}}(n)$ and $x_{\mathrm{R}}(n)$ (i.e., "left" and "right"), and two output or microphone signals denoted by $d_{\mathrm{L}}(n)$ and $d_{\mathrm{R}}(n)$, where n is the time index. The microphone signals can be expressed as

$$\begin{aligned} d_{\mathrm{L}}(n) &= y_{\mathrm{L}}(n) + v_{\mathrm{L}}(n), \quad (1) \\ d_{\mathrm{R}}(n) &= y_{\mathrm{R}}(n) + v_{\mathrm{R}}(n), \quad (2) \end{aligned}$$

where $y_{\mathrm{L}}(n)$ and $y_{\mathrm{R}}(n)$ are the stereo echo signals, and $v_{\mathrm{L}}(n)$ and $v_{\mathrm{R}}(n)$ are the near-end signals. The echo signals can be obtained as [3], [4]

$$\begin{aligned} y_{\mathrm{L}}(n) &= \mathbf{h}_{\mathrm{t,LL}}^T \mathbf{x}_{\mathrm{L}}(n) + \mathbf{h}_{\mathrm{t,RL}}^T \mathbf{x}_{\mathrm{R}}(n), \quad (3) \\ y_{\mathrm{R}}(n) &= \mathbf{h}_{\mathrm{t,LR}}^T \mathbf{x}_{\mathrm{L}}(n) + \mathbf{h}_{\mathrm{t,RR}}^T \mathbf{x}_{\mathrm{R}}(n), \quad (4) \end{aligned}$$

where $\mathbf{h}_{\mathrm{t,LL}}, \mathbf{h}_{\mathrm{t,RL}}, \mathbf{h}_{\mathrm{t,LR}}, \mathbf{h}_{\mathrm{t,RR}}$ are L-dimensional vectors of the loudspeaker-to-microphone ("true") acoustic impulse responses, the superscript ^{T} denotes transpose of a vector or a matrix, and the vectors

$$\begin{aligned} \mathbf{x}_{\mathrm{L}}(n) &= \left[\begin{array}{cccc} x_{\mathrm{L}}(n) & x_{\mathrm{L}}(n-1) & \cdots & x_{\mathrm{L}}(n-L+1) \end{array} \right]^T \\ \mathbf{x}_{\mathrm{R}}(n) &= \left[\begin{array}{cccc} x_{\mathrm{R}}(n) & x_{\mathrm{R}}(n-1) & \cdots & x_{\mathrm{R}}(n-L+1) \end{array} \right]^T \end{aligned}$$

contain the most recent L samples of the loudspeaker signals. Consequently, the main goal of this application is to estimate the four acoustic impulse responses (i.e., $\mathbf{h}_{\mathrm{t,LL}}, \mathbf{h}_{\mathrm{t,RL}}, \mathbf{h}_{\mathrm{t,LR}}, \mathbf{h}_{\mathrm{t,RR}}$) from the microphone signals in order to cancel the echo due to the coupling between the loudspeakers and the microphones.

In the context of acoustic echo cancellation, the loudspeaker and microphone signals are all real random variables. In order to introduce the proposed model, let us use the complex notation

$$d(n) = d_{\mathrm{L}}(n) + j d_{\mathrm{R}}(n) = y(n) + v(n), \quad (5)$$

where $j = \sqrt{-1}$, $y(n) = y_{\mathrm{L}}(n) + j y_{\mathrm{R}}(n)$, and $v(n) = v_{\mathrm{L}}(n) + j v_{\mathrm{R}}(n)$. Furthermore, let us define the complex random vector

$$\begin{aligned} \mathbf{x}(n) &= \left[\begin{array}{cccc} x(n) & x(n-1) & \cdots & x(n-L+1) \end{array} \right]^T \\ &= \mathbf{x}_{\mathrm{L}}(n) + j \mathbf{x}_{\mathrm{R}}(n), \quad (6) \end{aligned}$$

where $x(n) = x_{\mathrm{L}}(n) + j x_{\mathrm{R}}(n)$, so that the complex echo signal can be expressed as

$$y(n) = \mathbf{h}_{\mathrm{t}}^H \mathbf{x}(n) + \mathbf{h}_{\mathrm{t}}'^H \mathbf{x}^*(n), \quad (7)$$

where the superscripts ^{H} and * denote transpose-conjugate and conjugate, respectively, and

$$\mathbf{h}_{\mathrm{t}} = \mathbf{h}_{\mathrm{t},1} + j\mathbf{h}_{\mathrm{t},2}, \quad (8)$$

$$\mathbf{h}'_{\mathrm{t}} = \mathbf{h}'_{\mathrm{t},1} + j\mathbf{h}'_{\mathrm{t},2}, \quad (9)$$

with

$$\mathbf{h}_{\mathrm{t},1} = \frac{\mathbf{h}_{\mathrm{t,LL}} + \mathbf{h}_{\mathrm{t,RR}}}{2}, \quad \mathbf{h}_{\mathrm{t},2} = \frac{\mathbf{h}_{\mathrm{t,RL}} - \mathbf{h}_{\mathrm{t,LR}}}{2},$$

$$\mathbf{h}'_{\mathrm{t},1} = \frac{\mathbf{h}_{\mathrm{t,LL}} - \mathbf{h}_{\mathrm{t,RR}}}{2}, \quad \mathbf{h}'_{\mathrm{t},2} = -\frac{\mathbf{h}_{\mathrm{t,RL}} + \mathbf{h}_{\mathrm{t,LR}}}{2}.$$

Consequently, (7) can be rewritten as

$$y(n) = \widetilde{\mathbf{h}}_{\mathrm{t}}^H \widetilde{\mathbf{x}}(n), \quad (10)$$

where

$$\widetilde{\mathbf{h}}_{\mathrm{t}} = \begin{bmatrix} \mathbf{h}_{\mathrm{t}} \\ \mathbf{h}'_{\mathrm{t}} \end{bmatrix}, \quad \widetilde{\mathbf{x}}(n) = \begin{bmatrix} \mathbf{x}(n) \\ \mathbf{x}^*(n) \end{bmatrix}.$$

Finally, the complex reference signal (7) becomes

$$d(n) = \widetilde{\mathbf{h}}_{\mathrm{t}}^H \widetilde{\mathbf{x}}(n) + v(n). \quad (11)$$

In this context, our new goal is to estimate the complex acoustic impulse response $\widetilde{\mathbf{h}}_{\mathrm{t}}$ (of length $2L$) from the complex microphone signal, $d(n)$, and the complex loudspeaker signal, $x(n)$. In fact, the classical two-input/two-output system with real random variables has been converted to a single-input/single-output system with complex random variables. Looking of (7) or (10), we can recognize the widely linear (WL) model for complex random variables proposed in [5]; also, this approach is in consistence with the duality principle explained in [6].

3. SOME BASIC ADAPTIVE ALGORITHMS

Let $\widetilde{\mathbf{h}}(n) = \begin{bmatrix} \widetilde{h}_0(n) & \widetilde{h}_1(n) & \cdots & \widetilde{h}_{2L-1}(n) \end{bmatrix}^T$, be an adaptive filter of length $2L$, which is an estimate of $\widetilde{\mathbf{h}}_{\mathrm{t}}$, and let

$$\widehat{y}(n) = \widetilde{\mathbf{h}}^H(n-1)\widetilde{\mathbf{x}}(n) \quad (12)$$

be the output of the adaptive filter at time n. Thus, the error signal is

$$e(n) = d(n) - \widehat{y}(n). \quad (13)$$

Based on (12) and (13), we can write the update of the normalized least-mean-square (NLMS) algorithm as

$$\widetilde{\mathbf{h}}(n) = \widetilde{\mathbf{h}}(n-1) + \frac{\alpha\widetilde{\mathbf{x}}(n)e^*(n)}{\delta + \widetilde{\mathbf{x}}^H(n)\widetilde{\mathbf{x}}(n)}, \quad (14)$$

where α is the normalized stepsize parameter ($0 < \alpha < 2$) and $\delta \geq 0$ is the regularization constant. The NLMS algorithm could be useful in practice mainly due to its simplicity. However, it converges slowly for long length adaptive filters and highly correlated inputs.

In order to improve the convergence rate, we can take advantage of the sparseness character of the echo paths, which inspired the idea to "proportionate" the algorithm behavior [7]. In other words, we can update each coefficient of the filter independently of the others, by adjusting the adaptation stepsize in proportion to the magnitude of the estimated filter coefficient. Hence, the adaptation gain is "proportionately" redistributed among all the coefficients to emphasize the large ones in order to speed up their convergence and, consequently, to increase the overall convergence rate. Among the many proportionate-type algorithms developed for echo cancellation (e.g., see [8] and the references therein), the improved proportionate NLMS (IPNLMS) algorithm [9] is one of the most attractive choice. The good features of this algorithm include its simplicity and the robustness to the sparseness degree of the echo path.

In the context of the proposed model for SAEC, the update of the IPNLMS algorithm can be expressed as

$$\widetilde{\mathbf{h}}(n) = \widetilde{\mathbf{h}}(n-1) + \frac{\alpha\mathbf{G}(n-1)\widetilde{\mathbf{x}}(n)e^*(n)}{\delta + \widetilde{\mathbf{x}}^H(n)\mathbf{G}(n-1)\widetilde{\mathbf{x}}(n)}, \quad (15)$$

where

$$\mathbf{G}(n-1) = \mathrm{diag}\left[g_0(n-1), g_1(n-1), \ldots, g_{2L-1}(n-1)\right], \quad (16)$$

is a diagonal matrix (of size $2L \times 2L$) containing the proportionate (or gain) factors, which are evaluated as

$$g_l(n-1) = \frac{1-\kappa}{4L} + (1+\kappa)\frac{\left|\widetilde{h}_l(n-1)\right|}{2\sum_{i=0}^{2L-1}\left|\widetilde{h}_i(n-1)\right|}, \; 0 \leq l \leq 2L-1, \quad (17)$$

where κ ($-1 \leq \kappa < 1$) is a parameter that controls the amount of proportionality in the IPNLMS algorithm [9].

Another very good candidate for echo cancellation is the affine projection algorithm (APA) [10], since it converges and tracks faster than the NLMS algorithm. Besides, it can be efficient from an arithmetic complexity viewpoint as compared to more complex algorithms from the recursive least-squares (RLS) family. In order to derive the APA in the context of the proposed model, let us write the $2L \times P$ input matrix

$$\widetilde{\mathbf{X}}(n) = \begin{bmatrix} \widetilde{\mathbf{x}}(n) & \widetilde{\mathbf{x}}(n-1) & \cdots & \widetilde{\mathbf{x}}(n-P+1) \end{bmatrix},$$

where P is the projection order. Also, we can define the $P \times 1$ a priori error vector as

$$\mathbf{e}(n) = \mathbf{d}(n) - \widetilde{\mathbf{X}}^T(n)\widetilde{\mathbf{h}}^*(n-1), \quad (18)$$

where $\mathbf{d}(n) = \begin{bmatrix} d(n) & d(n-1) & \cdots & d(n-P+1) \end{bmatrix}^T$. Using this notation, the update of the APA is

$$\widetilde{\mathbf{h}}(n) = \widetilde{\mathbf{h}}(n-1) + \alpha\widetilde{\mathbf{X}}(n)\left[\delta\mathbf{I}_P + \widetilde{\mathbf{X}}^H(n)\widetilde{\mathbf{X}}(n)\right]^{-1}\mathbf{e}^*(n), \quad (19)$$

where $\mathbf{I}_P$ is the $P \times P$ identity matrix. It can be noticed that, by taking $P = 1$, we obtain the update of the NLMS algorithm (14).

The "proportionate" idea can be also extended in the case of APA, in order to further increase its performance when identifying sparse impulse responses. For example, using the gain factors of the IPNLMS algorithm, we can derive the improved proportionate APA (IPAPA):

$$\widetilde{\mathbf{h}}(n) = \widetilde{\mathbf{h}}(n-1) + \alpha\mathbf{G}(n-1)\widetilde{\mathbf{X}}(n) \times \left[\delta\mathbf{I}_P + \widetilde{\mathbf{X}}^H(n)\mathbf{G}(n-1)\widetilde{\mathbf{X}}(n)\right]^{-1}\mathbf{e}^*(n), \quad (20)$$

where $\mathbf{G}(n-1)$ is defined in (16) and (17). Clearly, for $P = 1$ we find the IPNLMS algorithm [see (15)].

Of course, many other adaptive algorithms can be derived in the context of the proposed model for SAEC. However, due to the lack of space, we limit our presentation to these four basic algorithms, i.e., NLMS, IPNLMS, APA, and IPAPA.

4. SOLUTIONS TO THE NONUNIQUENESS PROBLEM

It is well known that in the SAEC problem, most of the time, the two input signals [i.e., $x_{\mathrm{L}}(n)$ and $x_{\mathrm{R}}(n)$] are obtained by filtering a common source, so that a problem of nonuniqueness is expected [3]. Also, it was found that preprocessing of these far-end loudspeaker signals that actually are transmitted to the near-end room is the only way to achieve a unique solution [4]. In other words, it may be required to distort the input signals $x_{\mathrm{L}}(n)$ and $x_{\mathrm{R}}(n)$, in order to reduce the coherence between these two signals, which can lead to the estimation of the true acoustic impulse responses. However, this distortion must be performed in such a way that the quality of the signals and the stereo effect are not degraded.

A simple but efficient method uses positive and negative half-wave rectifiers on each channel respectively [4], i.e.,

$$x'_{\mathrm{L}}(n) = x_{\mathrm{L}}(n) + \alpha_{\mathrm{r}} \frac{x_{\mathrm{L}}(n) + |x_{\mathrm{L}}(n)|}{2}, \quad (21)$$

$$x'_{\mathrm{R}}(n) = x_{\mathrm{R}}(n) + \alpha_{\mathrm{r}} \frac{x_{\mathrm{R}}(n) - |x_{\mathrm{R}}(n)|}{2}, \quad (22)$$

where α_{r} is a parameter used to control the amount of nonlinearity. Experiments show that stereo perception is not affected by this method even with α_{r} as large as 0.5.

In the context of the proposed model, the complex input signal can be expressed as

$$x(n) = x_{\mathrm{L}}(n) + j x_{\mathrm{R}}(n) = e^{j\theta_{\mathrm{r}}(n)} |x(n)|, \quad (23)$$

where $\theta_{\mathrm{r}}(n)$ [with $\tan\theta_{\mathrm{r}}(n) = x_{\mathrm{R}}(n)/x_{\mathrm{L}}(n)$] and $|x(n)| = \sqrt{x_{\mathrm{L}}^2(n) + x_{\mathrm{R}}^2(n)}$ are the phase and module of $x(n)$, respectively. In this formulation, we represent the stereo perception with $\theta_{\mathrm{r}}(n)$ and the quality of the stereo signals with $|x(n)|$. A modification of $\theta_{\mathrm{r}}(n)$ only, will mostly affect the stereo effect of $x(n)$; while a modification of $|x(n)|$ will mostly affect the quality of the stereo signals.

Similarly, using the complex notation, (21) and (22) can be expressed as

$$x'(n) = x'_{\mathrm{L}}(n) + j x'_{\mathrm{R}}(n) = e^{j\theta'_{\mathrm{r}}(n)} |x'(n)|, \quad (24)$$

where $\tan\theta'_{\mathrm{r}}(n) = x'_{\mathrm{R}}(n)/x'_{\mathrm{L}}(n)$ and $|x'(n)| = \sqrt{x'^2_{\mathrm{L}}(n) + x'^2_{\mathrm{R}}(n)}$. In order to preserve the quality of the stereo signals, we propose not to modify the module of the complex input signal $x(n)$, but only to change its phase. Therefore, we can use the new following transformations [2]:

$$x''_{\mathrm{L}}(n) = \cos\theta'_{\mathrm{r}}(n) |x(n)|, \quad (25)$$

$$x''_{\mathrm{R}}(n) = \sin\theta'_{\mathrm{r}}(n) |x(n)|, \quad (26)$$

where the phase $\theta'_{\mathrm{r}}(n)$ is computed from the half-wave rectifiers [see (24)] while the module corresponds to the module of the original signals.

5. SIMULATION RESULTS

Simulations are performed in the context of the proposed model for SAEC. The acoustic impulse responses in the far-end location have 2048 coefficients, while the length of the impulse responses in the near-end location [i.e., $h_{\mathrm{t,LL}}(n)$, $h_{\mathrm{t,RL}}(n)$, $h_{\mathrm{t,LR}}(n)$, and $h_{\mathrm{t,RR}}(n)$] is $L = 512$. The length of the adaptive filter $\widetilde{\mathbf{h}}(n)$ is $2L = 1024$ and the sampling rate is 8 kHz.

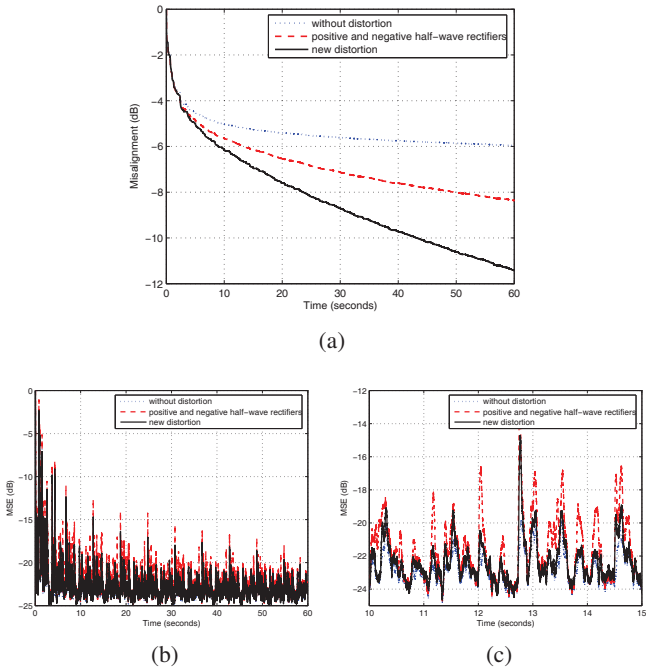

Fig. 1. Results of the NLMS algorithm for different types of distortion with $\alpha_{\mathrm{r}} = 0.3$. (a) Misalignment; (b) MSE; (c) MSE detail.

The source signal in the far-end location is a speech sequence. All simulations are performed in the single-talk scenario, i.e., absence of a near-end talker. In this case, the near-end signal $v(n)$ consists only of the background noise. We can define the stereo echo-to-noise ratio (SENR) [which is equivalent to the signal-to-noise ratio (SNR)] as SENR $= \sigma_y^2/\sigma_v^2$, where $\sigma_y^2 = E\left[|y(n)|^2\right]$ and $\sigma_v^2 = E\left[|v(n)|^2\right]$ are the variances of $y(n)$ and $v(n)$, respectively. In our simulations, the background noise in the near-end is an independent white Gaussian signal and its level is set such that SENR = 30 dB.

We choose for comparisons the four algorithms presented in Section 3, i.e., NLMS, IPNLMS, APA, and IPAPA. The stepsize for all the algorithms is $\alpha = 0.25$ and the regularization constants are $\delta_{\mathrm{NLMS}} = \delta_{\mathrm{APA}} = 20\sigma_x^2$ and $\delta_{\mathrm{IPNLMS}} = \delta_{\mathrm{IPAPA}} = 20\sigma_x^2/(2L)$ [11], where $\sigma_x^2 = E\left[|x(n)|^2\right]$ is the variance of $x(n)$. The proportionate-type algorithms (i.e., IPNLMS and IPAPA) use $\kappa = 0$. The performance of the algorithms is evaluated in terms of two measures, i.e., (a) the normalized misalignment (in dB), defined as $20\log_{10}\left\|\widetilde{\mathbf{h}}_{\mathrm{t}} - \widetilde{\mathbf{h}}(n)\right\|_2 / \left\|\widetilde{\mathbf{h}}_{\mathrm{t}}\right\|_2$ (with $\|\cdot\|_2$ denoting the ℓ_2 norm) and (b) the mean-square error (MSE) averaged over 256 points for the purpose of smoothing the results.

In all the experiments, we compare the performance of the algorithms using positive and negative half-wave rectifiers [see (21) and (22)] versus the new proposed distortion [see (25) and (26)]; the distortion parameter is set to $\alpha_{\mathrm{r}} = 0.3$. Also, the case without distortion is shown as a reference.

Figure 1 presents the performance of the NLMS algorithm. It can be noticed from Fig. 1(a) that the misalignment is greatly reduced by the new distortion. Also, as we can see in Fig. 1(b) and in the detail presented in Fig. 1(c), the new distortion leads to a better performance in terms of the MSE as compared to the positive and

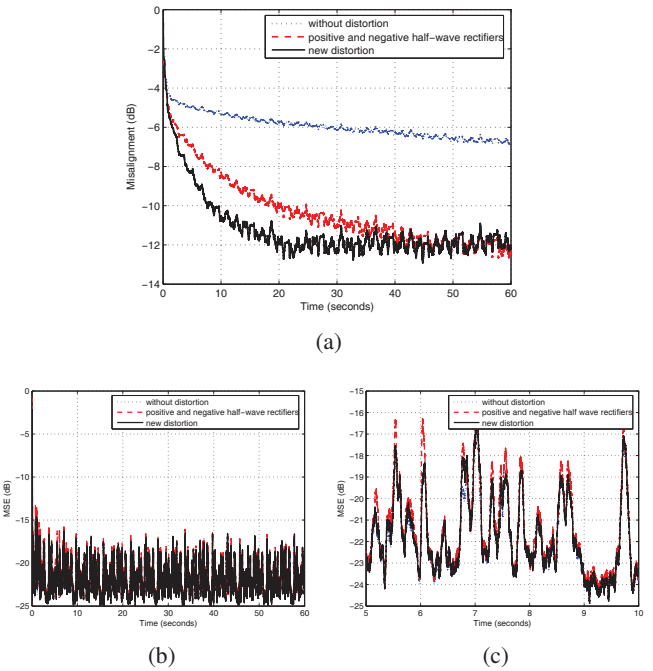

Fig. 2. Results of the APA (using $P = 8$) for different types of distortion with $\alpha_{\mathrm{r}} = 0.3$. (a) Misalignment; (b) MSE; (c) MSE detail.

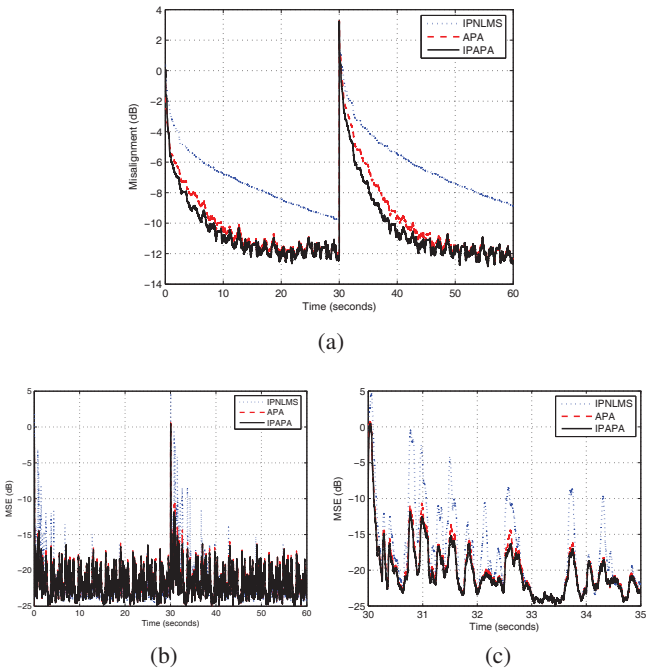

Fig. 3. Results of the IPNLMS, APA, and IPAPA (using $P = 8$) for different types of distortion with $\alpha_{\mathrm{r}} = 0.3$. Echo paths changes at time 30 seconds. (a) Misalignment; (b) MSE; (c) MSE detail.

negative half-wave rectifiers method.

Figure 2 shows the performance of the APA with $P = 8$; it was found that this value of the projection order offers a proper compromise between the performance and complexity. It can be noticed from Fig. 2(a) that the APA converges faster with the new distortion, outperforming by far the NLMS algorithm [see for comparison Fig. 1(a)]. Also, as we can see in Fig. 2(b) and in the detail presented in Fig. 2(c), the new distortion leads to a slightly better performance in terms of the MSE as compared to the positive and negative half-wave rectifiers.

Since the IPAPA has resulted as a combination between the IPNLMS algorithm and the APA, it is expected that the IPAPA should outperform both its predecessors. The last experiment outlines this aspect, by comparing these three algorithms in a tracking situation (the impulse responses in the near-end location are shifted to the right by 12 samples). The new distortion is used with $\alpha_{\mathrm{r}} = 0.3$. The projection order is $P = 8$ for the APA and IPAPA. The results are shown in Fig. 3. According to these plots, it is clear that IPAPA outperforms both the IPNLMS and APA.

6. CONCLUSIONS

In this paper, we proposed to recast the SAEC problem as a single-input/single-output system with complex random variables. As a consequence, the four real-valued acoustic impulse responses are converted to one complex-valued impulse response. The main advantage of this approach is that instead of handling two (real) output signals separately, we only handle one (complex) output signal. In this framework, we have presented some typical adaptive algorithms and a new distortion method suitable for this model.

7. REFERENCES

[1] 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.

[2] J. Benesty, C. Paleologu, T. Gänsler, and S. Ciochină, *A Perspective on Stereophonic Acoustic Echo Cancellation*. Springer-Verlag, Berlin, Germany, 2011.

[3] M. M. Sondhi, D. R. Morgan, and J. L. Hall, "Stereophonic acoustic echo cancellation–An overview of the fundamental problem," *IEEE Signal Process. Lett.*, vol. 2, pp. 148–151, Aug. 1995.

[4] J. Benesty, D. R. Morgan, and M. M. Sondhi, "A better understanding and an improved solution to the specific problems of stereophonic acoustic echo cancellation," *IEEE Trans. Speech, Audio Process.*, vol. 6, pp. 156–165, Mar. 1998.

[5] B. Picinbono and P. Chevalier, "Widely linear estimation with complex data," *IEEE Trans. Signal Process.*, vol. 43, pp. 2030–2033, Aug. 1995.

[6] D. P. Mandic, S. Still, and S. C. Douglas, "Duality between widely linear and dual channel adaptive filtering," in *Proc. IEEE ICASSP*, 2009, pp. 1729–1732.

[7] D. L. Duttweiler, "Proportionate normalized least-mean-squares adaptation in echo cancelers," *IEEE Trans. Speech, Audio Process.*, vol. 8, pp. 508–518, Sept. 2000.

[8] C. Paleologu, J. Benesty, and S. Ciochină, *Sparse Adaptive Filters for Echo Cancellation*. Morgan & Claypool Publishers, 2010.

[9] J. Benesty and S. L. Gay, "An improved PNLMS algorithm," in *Proc. IEEE ICASSP*, 2002, pp. 1881–1884.

[10] K. Ozeki and T. Umeda, "An adaptive filtering algorithm using an orthogonal projection to an affine subspace and its properties," *Electron. Commun. Jpn.*, vol. 67-A, pp. 19–27, May 1984.

[11] J. Benesty, C. Paleologu, and S. Ciochină, "On regularization in adaptive filtering," *IEEE Trans. Audio, Speech, Language Process.*, vol. 19, pp. 1734–1742, Aug. 2011.