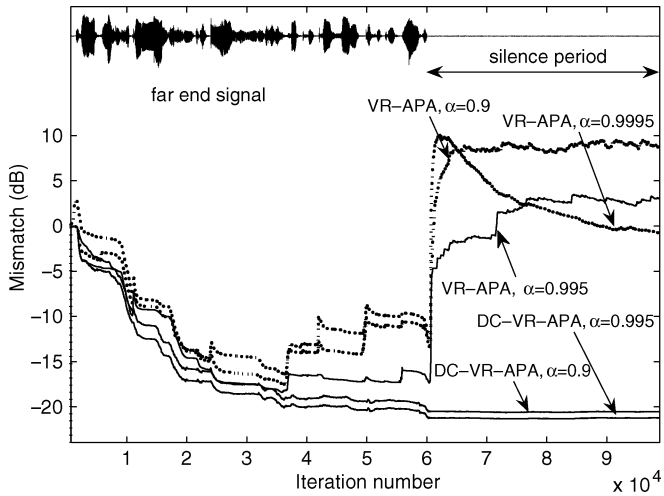

Fig. 1. Performance comparison of DC-VR-APA and VR-APA for various α. NESA is absent and input signal is speech signal with silence periods. Adaptive filter length is 512. Results were obtained in the context of a subband affine projection algorithm [3] with two subbands and projection order equal to 4.

and 6×10^4 during pauses between words. The DC-VR-APA does not diverge in these circumstances. When the far-end signal power tends to zero, we have $E\{\|\mathbf{e}_i\|^2\} - K\sigma_v^2 \rightarrow 0$. If $\delta_i > E\{\|\mathbf{e}_i\|^2\} - K\sigma_v^2$, it follows that $\hat{\beta}_i < 0.5\beta_i$ and the VR-APA can diverge. This happens at time instant 6×10^4 in Fig. 1. To a lesser extent, it also happens between time instants 4×10^4 and 6×10^4 where there are instances when the far-end signal energy is low. Once the VR-APA diverges, it quickly reconverges for a higher value of α. This is because for a higher value of α, there are lesser fluctuations[5] of $\sigma_{\mathbf{e}_i}^2$ around its true value $E\{\|\mathbf{e}_i\|^2\}$ and there are lesser instances when $\hat{\beta}_i$ is negative. Consequently, there are fewer instances when the adaptive filter is frozen and reconvergence happens quickly.

The divergence of the VR-APA during the silence periods can be addressed by using a lower bound, $\beta_{\min}$, on the computed value of $\hat{\beta}_i$. The value of $\beta_{\min}$ should be much larger than the low values of far-end signal energy that is encountered during silence periods. The use of $\beta_{\min}$ prevents filter divergence due to near-end noise amplification [5]. A high value (low value) of $\beta_{\min}$ ensures more (less) stability against divergence but results in slower (faster) convergence. The optimal value of $\beta_{\min}$ needs further investigation. However, note that the use of $\beta_{\min}$ does not solve the problem of divergence of the VR-APA when NESA is present. It is also to be noted that it is not necessary to use the lower bound $\beta_{\min}$ in the computation of $\hat{\beta}_i^{(\mathrm{dc})}$ in DC-VR-APA.

ACKNOWLEDGMENT

The authors would like to thank STMicroelectronics Asia Pacific Pte. Ltd. for supporting this research activity.

REFERENCES

[1] H. Rey, L. R. Vega, S. Tressens, and J. Benesty, "Variable explicit regularization in affine projection algorithm: Robustness issues and optimal choice," *IEEE Trans. Signal Process.*, vol. 55, no. 5, pp. 2096–2109, May 2007.

[2] V. Myllyla and G. Schmidt, "Pseudo-optimal regularization for affine projection algorithms," in *Proc. IEEE Int. Conf. Acoustics, Speech, Signal Processing (ICASSP)*, May 2002, pp. 1917–1920.

[3] M. Karthik, K. H. Li, and S. George, "Delay coefficients based variable regularization subband affine projection algorithms in acoustic echo cancellation applications," *IEICE Trans. Fundam. Electron., Commun. Comput. Sci.*, vol. E-92 A, no. 7, pp. 1699–1703, Jul. 2009.

[4] A. Mader, H. Puder, and G. Schmidt, "Step-size control for echo cancellation filters-An overview," *Signal Process.*, vol. 80, no. 9, pp. 1697–1719, Sep. 2000.

[5] S. Haykin, *Adaptive Filter Theory*, Low Price Edition ed. Delhi, India: Pearson Education, Indian Branch, 2002.

[6] M. A. Iqbal and S. L. Grant, "Novel variable step size NLMS algorithm for echo cancellation," in *Proc. IEEE Int. Conf. Acoust., Speech, Signal Process. (ICASSP)*, Mar. 2008, pp. 241–244.

[5]These reasons were reported in [1, p. 2102], although in a different context.

Author's Reply to "Comments on 'Variable Explicit Regularization in Affine Projection Algorithm: Robustness Issues and Optimal Choice' "

Hernán Rey, Leonardo Rey Vega, *Student Member, IEEE*, Jacob Benesty, and Sara Tressens

In [1], Muralidhar *et al.* point out that when the VR-APA [2] is applied in echo cancellation it will diverge during the presence of a near-end signal or a silence period. The first thing we would like to clarify is that what they probably meant is that in these situations the mismatch will show an increasing transient period. This should not be confused with the divergent behavior of adaptive algorithms, e.g., an NLMS with step-size larger than 2.

Regarding the behavior of the VR-APA during silence periods, Muralidhar *et al.* argue that fluctuations in $\hat{\sigma}_{\mathbf{e}_i}^2$, the estimate of $E[\|\mathbf{e}_i\|^2]$, may take the estimate of β_i, i.e., $\hat{\beta}_i$, below half its optimal value, causing instabilities. Besides the approximations made in their analysis, it is impossible to know how small the optimal β_i would be at a certain time given that $E[\|\mathbf{e}_i\|^2]$ is unknown. However, one might think that if $\hat{\beta}_i$ is small, it is possibly small enough to be below the optimal β_i and, according to them, cause instability. One thing that must be checked in this case is the condition number of the matrix $\mathbf{S}_i = (\mathbf{X}_i^H\mathbf{X}_i + \hat{\beta}_i\mathbf{I}_K)$, to ensure that instabilities of the algorithm are not caused by numerical errors due to poor regularization.

Consider now the effect of a long silence period in the dynamics of the algorithm. It is well known that when dealing with stationary inputs, the lower the output signal-to-noise ratio (SNR), the higher the steady-state mismatch. If we say we are in an echo cancellation scenario with SNR = 20 dB, the idea is that the power of the background additive noise at the output of the system is 20 dB below the power of the output signal, computed over a certain time frame. During a silence

Manuscript received November 24, 2009; accepted March 24, 2010. Date of publication April 8, 2010; date of current version June 16, 2010. The associate editor coordinating the review of this manuscript and approving it for publication was Dr. Dominic Ho.

H. Rey and L. Rey Vega are with the University of Buenos Aires and CONICET, Paseo Colon 850 (1063), Buenos Aires, Argentina (e-mail: hrey@fi.uba.ar; lrey@fi.uba.ar).

J. Benesty is with the INRS-EMT, Université du Québec, Montréal, QC H5A 1K6, Canada (e-mail: benesty@emt.inrs.ca).

S. Tressens is with the University of Buenos Aires, Buenos Aires, Argentina (e-mail: stres@fi.uba.ar).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TSP.2010.2047722

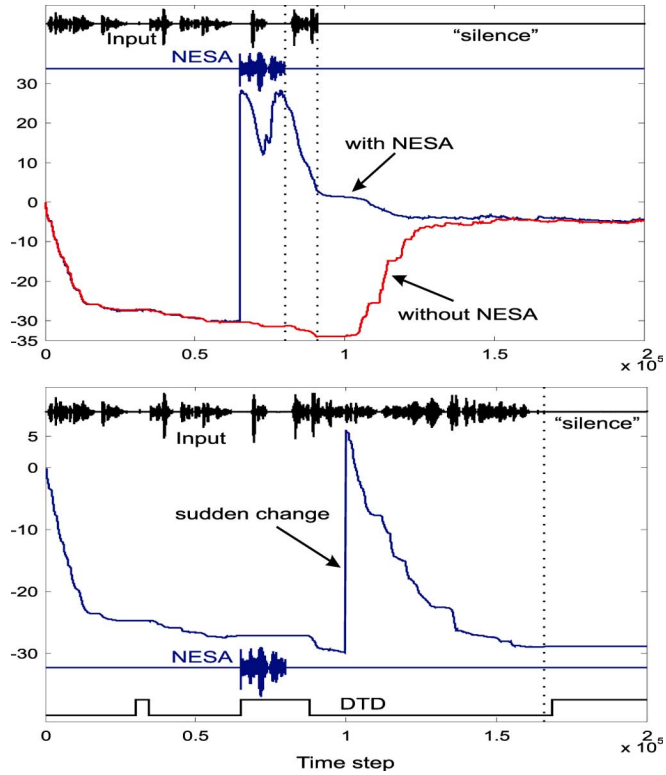

Fig. 1. Mismatch (in dB). Speech is sampled at 8 kHz. SNR = 20 dB. $M = 1024$. $K = 4$. $\eta = \eta_{\mathrm{DTD}} = 1$. $\eta_{\mathrm{input}} = 3$. $\delta = 0.01$. The ratio between the powers of the output signal and the NESA is -4.5 dB. $T = 0.15$. Top: No DTD is used. Bottom: DTD is used.

period, if the input results in $\mathbf{x}_i = \mathbf{0}$, the filter estimate will not be updated. But in practice, during a "silence" period the system is actually excited with some background noise at the input (at best 40 dB below the input power). This lead to a drop in the output signal power so that the output SNR can be seen as degraded. This analysis can be applied in general to any adaptive filter. The only way to prevent the filter to evolve towards the resulting new state of the mismatch would be to freeze the adaptation soon after the beginning of the silence period. In the VR-APA, this means that $\hat{\beta}_i$ should be very large. For computing $\hat{\beta}_i$, $E[\|\mathbf{e}_i\|^2]$ and σ_x^2 in [2, eq. (23)] are replaced by $\hat{\sigma}^2_{\mathbf{e}_i}$ and $\hat{\sigma}^2_{x_i}$, which are computed according to [2, eq. (36)] with parameters η and η_{input} respectively. Therefore, during the silence period, numerator and denominator of [2, eq. (36)] will be small. Fluctuations in $\hat{\sigma}^2_{\mathbf{e}_i}$ might turn the denominator negative, leading to a large value $\hat{\beta}_i = \beta_{\max}$ [2, eq. (26)], but they might also be large enough to lead to a relatively small $\hat{\beta}_i$. We argue that when $\hat{\beta}_i = \beta_{\max}$, and no near-end speech activity (NESA) is present, the mismatch will be almost unchanged. Nevertheless, a small $\hat{\beta}_i$ will allow large mismatch updates. Whether this update would lead to an increase on the mismatch depends on the relation between the current mismatch and the new mismatch state defined by the degraded SNR.

In the top part of Fig. 1 we simulated two different scenarios. The "silence " period is simulated with a zero mean white noise with power 30 dB below σ_x^2, the input power computed over the nonsilent period. When NESA is not present, the mismatch at the beginning of the silence period is close to -35 dB. After a transient period $\hat{\sigma}^2_{x_i}$ becomes small enough so that $\hat{\beta}_i$ starts to fluctuate between small values and $\beta_{\max}$ (not shown). When $\hat{\beta}_i$ is small, the mismatch is increased, until it reaches a steady-state around -5 dB. In the other scenario, NESA is present a few seconds before the beginning of the silence period, causing the mismatch at the beginning of the silence period to be larger than -5 dB. Yet again, $\hat{\beta}_i$ starts to fluctuate between small values and $\beta_{\max}$ in a similar way to the previous scenario. However, in this case, the updates lead to a decrease in the mismatch until it reaches the same steady-state value of the previous scenario. In both scenarios, the small values of $\hat{\beta}_i$ are large enough to provide regularization to the algorithm as the condition number of $\mathbf{S}_i$ was close to 1 (not shown).

The other issue pointed out in [1] is related to the NESA. When NESA is present it will have a strong impact on the filter update through $\mathbf{e}_i$. To prevent this, $\hat{\beta}_i$ must take a very large value in order to freeze the adaptation. But how large? In the setup of Fig. 1, $\hat{\beta}_i$ should be $\sigma_x^2 \times 10^7$ during the period when NESA is present. If this value is reduced by one order of magnitude, an increase in the mismatch will be noticeable (not shown). This high sensitivity poses a problem when implementing a variable regularization algorithm that should perform well in double-talk situations. In any case, the regularization control of the VR-APA would lead to a low value of $\hat{\beta}_i$ since the denominator of [2, eq. (36)] will be large. The robustness of VR-APA discussed in [2] means that small perturbations lead to small estimation errors. If an algorithm showing slight sensitivity to large perturbations is preferred, the approach used in [3] will be more appropriate.

We have seen that whether NESA is present or a silence period occurs, adaptation should be frozen. Fortunately, this can be accomplished with the use of a double-talk detector (DTD). We implement a simple DTD based on cross correlation [4]. The cross-correlation between the input and observed output, the observed output power and the system output power are estimated by $\hat{r}_{\mathbf{x}_i,d_i}$, $\hat{\sigma}^2_{d_i}$ and $\hat{\sigma}^2_{\hat{y}_i}$, respectively (computed as in [2, eq. (36)] with parameter η_{DTD}). They are used to evaluate the quantity

$$\xi_i = \frac{\left|\hat{\mathbf{w}}_i^H \hat{r}_{\mathbf{x}_i,d_i}\right|}{\left|\hat{\mathbf{w}}_i^H \hat{r}_{\mathbf{x}_i,d_i}\right| + \left|\hat{\sigma}^2_{d_i} - \hat{\sigma}^2_{\hat{y}_i}\right|}. \tag{1}$$

If ξ_i is below a threshold T the adaptation is frozen for 500 iterations. The bottom part of Fig. 1 presents a period with NESA, a sudden change in the system and a silence period in the input signal. It can be seen that the DTD freezes the adaptation during NESA and silence periods, resulting in an overall good performance of the VR-APA in an echo cancellation scenario.

REFERENCES

[1] K. Muralidhar, K. H. Li, and S. George, "Comments on 'Variable explicit regularization in affine projection algorithm: Robustness issues and optimal choice'," *IEEE Trans. Signal Process.*, vol. 58, no. 7, Jul. 2010.

[2] H. Rey, L. R. Vega, S. Tressens, and J. Benesty, "Variable explicit regularization in affine projection algorithm: Robustness issues and optimal choice," *IEEE Trans. Signal Process.*, vol. 55, no. 5, pp. 2096–2109, May 2007.

[3] L. R. Vega, H. Rey, J. Benesty, and S. Tressens, "A new robust variable step-size NLMS algorithm," *IEEE Trans. Signal Process.*, vol. 56, no. 5, pp. 1878–1893, May 2008.

[4] J. Benesty, D. Morgan, and J. H. Cho, "A new class of doubletalk detectors based on cross-correlation," *IEEE Trans. Speech Audio Process.*, vol. 8, no. 2, pp. 168–172, Mar. 2000.