and $6 \times 10^4$ 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 $E[||e||^2] - K \sigma^2_s \to 0$. If $\delta_i > E[||e||^2] - K \sigma^2_s$, it follows that $\beta_i < 0, \beta_i$, and the VR-APA can diverge. This happens at time instant $6 \times 10^4$ in Fig. 1. To a lesser extent, it also happens between time instants $4 \times 10^4$ and $6 \times 10^4$ where there are instances when the far-end signal energy is low. Once the VR-APA diverges, it quickly re- converges for a higher value of $\alpha$. This is because for a higher value of $\alpha$, there are lesser fluctuations of $\sigma^2_s$ around its true value $E[||e||^2]$ and there are lesser instances when $\beta_i$ 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, $\beta_{\text{min}}$, on the computed value of $\beta_i$. The value of $\beta_{\text{min}}$ should be much larger than the low values of far-end signal energy that is encountered during silence periods. The use of $\beta_{\text{min}}$ prevents filter divergence due to near-end noise amplification [5]. A high value (low value) of $\beta_{\text{min}}$ ensures more (less) stability against divergence but results in slower (faster) convergence. The optimal value of $\beta_{\text{min}}$ needs further investigation. However, note that the use of $\beta_{\text{min}}$ 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 $\beta_{\text{min}}$ in the computation of $\beta_i^{(d)}$ in DC-VR-APA.

ACKNOWLEDGMENT

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

REFERENCES


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 $\sigma^2_s$, the estimate of $E[||e||^2]$, may take the estimate of $\beta_i$, i.e., $\beta_i$, below half its optimal value, causing instabilities. Besides the approximations made in their analysis, it is impossible to know how small the optimal $\beta_i$ would be at a certain time given that $E[||e||^2]$ is unknown. However, one might think that if $\beta_i$ is small, it is possibly small enough to be below the optimal $\beta_i$ and, according to them, cause instability. One thing that must be checked in this case is the condition number of the matrix $S_i = (X_i X_i^H + \beta_i I_k)$, 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
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 $10^{-24}$. $K = 4$, $\eta = \eta_{\text{DTD}} = 1$, $\eta_{\text{NESA}} = 3$, $\delta = 0.01$. 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.

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 $e_i$. To prevent this, $\beta_i$ must take a very large value in order to freeze the adaptation. But how large? In the setup of Fig. 1, $\beta_i$ should be $\sigma_i^2 \times 10^7$ 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 $\beta_i$ 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 $\hat{r}_{x_i, e_i}$, $\hat{\sigma}_{x_i}^2$ and $\hat{\sigma}_{e_i}^2$, respectively (computed as in [2, eq. (36)] with parameter $\eta_{\text{DTD}}$). They are used to evaluate the quantity

$$\xi_i = \frac{|\hat{r}_{x_i, e_i}|^2}{\hat{\sigma}_{x_i}^2 \hat{\sigma}_{e_i}^2 / \hat{\sigma}_{x_i}^2}.$$  

If $\xi_i$ 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