ADAPTIVE COMBINATION OF IPNLMS FILTERS FOR ROBUST SPARSE ECHO CANCELLATION

Jerónimo Arenas-García and Aníbal R. Figueiras-Vidal*

Dep. of Signal Theory and Communications
Universidad Carlos III de Madrid
28911 Leganés-Madrid, Spain
{jarenas,arfv}@tsc.uc3m.es

ABSTRACT

Proportionate adaptive filters, such as the improved proportionate normalized least-mean-square (IPNLMS) algorithm, have been proposed for echo cancellation as an interesting alternative to the normalized least-mean-square (NLMS) filter. Proportionate schemes offer improved performance when the echo path is sparse, but are still subject to some compromises. In this paper, we study how combination schemes, where the output of two independent adaptive filters are adaptively mixed together, can be used to increase IPNLMS robustness to channels with different degrees of sparsity, as well as to alleviate the rate of convergence vs steady-state misadjustment tradeoff imposed by the selection of the step size. The advantages of these combined filters are illustrated in several echo cancellation scenarios.

Index Terms— Combination filters, proportionate filters, echo cancellation, sparse channel identification

1. INTRODUCTION

Adaptive echo cancellation, both acoustic and electrical, is a key component of modern communication networks. The overall echo cancellation process is illustrated in Fig. 1. The echo is produced when the far-end signal activates a (possibly time-varying) echo path, $\mathbf{w}_o(n)$. This echo signal is superimposed upon the near-end signal, $s(n)$, which is possibly contaminated by additive noise $e_0(n)$. The goal of an echo canceler is to produce a replica $y(n)$ of the echo signal which can be used to remove the echo before the signal is delivered to the far-end.

In this paper, we are interested in modelling sparse or quasi sparse echo channels, in which only a small fraction of the weights of the impulse response are significantly different from zero (the so-called active coefficients). Such echo paths are typically encountered in acoustic and network echo cancellation [1, 2], including also internet telephony [3, 4], where the echo-path impulse responses are of short duration, but present an unknown or even time-varying delay. Therefore, it becomes necessary to use echo cancelers with a long memory, which can be implemented with adaptive filters with hundreds or even thousands of weights, of which only a few will significantly differ from zero after convergence. Following [4], we define the degree of sparsity of a channel as a qualitative measure ranging from strongly dispersive (when most of the coefficients of $\mathbf{w}_o(n)$ are active) to strongly sparse.

*This work was partly supported by by the Spanish Ministry of Education and Science under grant CICYT TEC-2005-00992 and by Madrid Community grant S-505/TIC/0223.

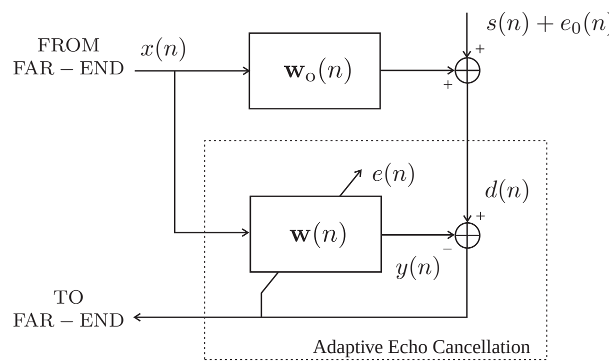


Fig. 1. Block diagram for an adaptive echo cancellation configuration.

It is a well-known fact that adaptive schemes which distribute the adaptation energy equally among all filter coefficients, such as least-mean-square (LMS) and normalized LMS (NLMS), exhibit a very slow convergence for filters with many taps [5, 6], making the application of such schemes unpractical for the cancellation of sparse echo channels. To alleviate this problem, the proportionate NLMS algorithm (PNLMS) [7] makes the adaptation step for each tap proportional to the current absolute value of the estimated weight, i.e,

$$w_m(n+1) = w_m(n) + \mu_m(n)\frac{e(n)}{\mathbf{x}^T(n)\mathbf{x}(n)}x_m(n), \qquad (1)$$

with $\mu_m(n) \propto |w_m(n)|$, for $m = 1, \ldots, M$, where M is the filter length. In the above expression, $w_m(n)$ and $x_m(n)$ are the mth components of the filter weights and the input regressor at time n, $\mathbf{w}(n)$ and $\mathbf{x}(n)$, respectively, and $e(n) = d(n) - y(n)$ is the error incurred by the filter, $d(n)$ being the desired response and $y(n) = \mathbf{w}^T(n)\mathbf{x}(n)$ the filter output.

By making the step size associated to each filter tap proportional to the current estimation of the corresponding weight component, the weights corresponding to the active region of the sparse echo path are adapted faster, and PNLMS achieves faster convergence than NLMS. A major drawback of the PNLMS filter, however, is that its behavior degrades significantly when identifying not-so-sparse echo channels.

Another proportionante scheme which tries to improve the robustness of PNLMS to dispersive channels is the so-called improved PNLMS algorithm (IPNLMS) [8], to which we will devote our attention in this paper. The different coefficients of an IPNLMS filter are adapted according to

$$w_m(n+1) = w_m(n) + \mu_m(n)e(n)x_m(n), \quad m = 1, \ldots, M \quad (2)$$

$$\mu_m(n) = \frac{\mu\, g_m(n)}{\delta + \sum_{k=1}^{M} g_k(n)x_k^2(n)}, \quad (3)$$

where μ is the step size of the filter and δ is a small constant to avoid division by 0. The gain for each weight, $g_m(n)$, is calculated using

$$g_m(n) = (1-\kappa)\frac{1}{2M} + (1+\kappa)\frac{|w_m(n)|}{\epsilon + 2\|\mathbf{w}(n)\|_1}, \quad (4)$$

where ϵ is a small positive constant, $\|\mathbf{w}(n)\|_1 = \sum_k |w_k(n)|$, and κ is a constant between -1 and 1, which allows to establish a tradeoff between the standard NLMS filter ($\kappa = -1$) and basic PNLMS ($\kappa = 1$).

From (3) and (4), it is evident that the step size associated to each coefficient increases with the absolute value of that coefficient. Consequently, and like in the PNLMS case, IPNLMS spends more energy adapting the active coefficients, thus converging faster than NLMS. It can also be shown that both IPNLMS and NLMS incur in approximately the same steady-state misalignment for a given step-size value.

In spite of the fact that IPNLMS generally achieves better performance than NLMS, the selection of its parameters μ and κ must consider the following compromises:

1) As any other gradient-based adaptive filter, IPNLMS is subject to a speed vs precision tradeoff, i.e., a large step size results in faster convergence, while the residual misadjustment is reduced for small μ.
2) Parameter κ imposes a behavior tradeoff for channels with different degrees of sparsity [8]. For strongly sparse channels the fastest convergence is obtained by the PNLMS filter ($\kappa = 1$), but such a selection leads to suboptimal performance in not-so-sparse channels. Therefore, the best value of κ for a given scenario depends on the actual degree of sparsity of the echo channel.

The adaptive combination of adaptive filters has proved to be an effective approach to improve the performance of adaptive schemes and to simplify their use. The basic idea of the algorithms in [9, 10] is to combine adaptive filters with different settings, so that the combination behaves, at each iteration, as the best component filter (or even better than any of them [11]). In previous works, this approach was used to improve the speed vs precision tradeoff of adaptive filters. In this paper, we will explain how the combination approach can also be useful to increase the robustness of the IPNLMS filter to channels with different degrees of sparsity. The efectiveness of the new schemes for echo cancellation will also be illutrated for echo paths with different degrees of sparsity.

2. ADAPTIVE COMBINATION OF IPNLMS FILTERS

The configuration of the adaptive combination scheme of two adaptive filters presented in [9] is illustrated in Fig. 2. To get

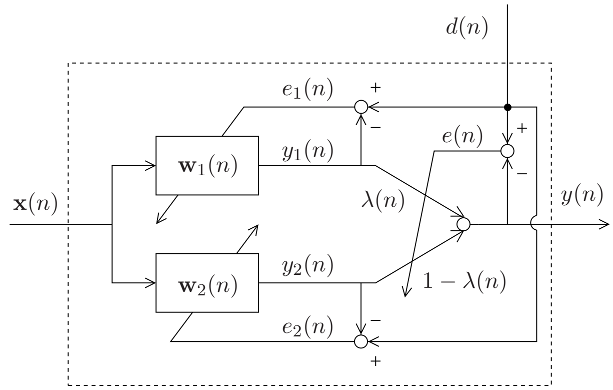


Fig. 2. Adaptive combination of two adaptive filters. Each component is adapted according to its own rules and error, while the mixing parameter, $\lambda(n)$, is updated to minimize the overall error.

a good performance from the combination, both component filters, $\mathbf{w}_1$ and $\mathbf{w}_2$, are independently adapted, using their own rules and settings. The overall filter output, $y(n)$, is calculated as a convex combination of the outputs of the two component filters, $y_i(n),\ i = 1, 2$:

$$y(n) = \lambda(n)y_1(n) + [1 - \lambda(n)]y_2(n), \quad (5)$$

where $\lambda(n)$ is a mixing parameter which is kept in the range $[0, 1]$ by defining it as the output of a sigmoid activation function,

$$\lambda(n) = \text{sgm}[a(n)] = \frac{1}{1 + e^{-a(n)}}. \quad (6)$$

At each iteration step, $a(n)$ will be adapted as described below, and then $\lambda(n)$ will be obtained from (6). Since the relation between $a(n)$ and $\lambda(n)$ is one to one, we will indistinctly refer to them as mixing parameters.

If the mixing parameter is appropriately adapted, the combination will keep the best characteristics of each component. Following [9], $a(n)$ can be adapted to minimize the square error of the overall filter according to the following stochastic gradient descent rule[1]:

$$\begin{aligned} a(n+1) &= a(n) - \mu_a \frac{\partial e^2(n)}{\partial a(n)} \\ &= a(n) + \mu_a e(n)[y_1(n) - y_2(n)]\lambda(n)[1 - \lambda(n)], \end{aligned} \quad (7)$$

μ_a being a step-size parameter for the adaptation of $a(n)$. For practical reasons, $a(n)$ is kept within $[-4, 4]$, so that its adaptation does not stop because of factor $\lambda(n)[1 - \lambda(n)]$ in (7) being too close to 0.

The steady-state performance of the combination scheme we have just reviewed has been theoretically analyzed in [11], showing that it enjoys *universal capabilities* with respect to the component filters. In other words, the combination performs at least as well as the best component filter and, under certain conditions, better than any of them (and, in this case, we say that the scheme performance is *better-than-universal*).

[1]Other adaptation schemes for the mixing parameter are also possible (see, e.g., [12, 13]).

Note that this result holds independently of the kind of adaptive filters used for the combination. In Section 3 we will show that this *better-than-universal* behavior is observed for the particular case of a combination of IPNLMS adaptive filters.

The adaptive combination scheme can be used to improve the convergence rate vs misadjustment tradeoff of IPNLMS filters, as well as their robustness to echo paths with different degrees of sparsity. To do so, we propose to combine two IPNLMS filters, selecting their parameters, $\{\mu_1, \kappa_1\}$ and $\{\mu_2, \kappa_2\}$, in one of the following ways:

a) $\mu_1 > \mu_2$ and $\kappa_1 = \kappa_2$: This configuration is intended to keep the faster convergence of the filter with step size μ_1, while achieving the lower steady-state error of the filter with small step size; thus the overall convergence rate vs steady-state performance tradeoff is improved.

b) $\mu_1 = \mu_2$, $\kappa_1 < 0$, $\kappa_2 \approx 1$: With this selection, the combined filter shows improved convergence and robustness against channels with different degrees of sparsity, i.e., the combination retains the faster initial convergence of PNLMS for very sparse channels[2], while behaving as an IPNLMS with κ_1 in a later stage of the adaptation, when the PNLMS filter convergence is known to slowdown. The first filter provides also robustness against channels which are less sparse than expected.

Switching between the PNLMS and the NLMS algorithms, which can be seen close in spirit to the second configuration proposed above, has already been explored in [14, 15]. The present combination of IPNLMS filters can be seen as a more flexible approach in the sense that, rather than using a hard commutation between both adaptation schemes, it uses a soft combination of two component filters which are adapted in parallel, and which can be different in any form. We will see in the experiments section that, by doing so, not only a faster convergence is achieved, but in many situations it is also possible to get a lower identification error than from any of the components alone.

The computational complexity of the proposed combination is roughly twice that of the basic IPNLMS (for the two IPNLMS filters running in parallel). Application of the combination scheme requres six additional multiplications, which is usually much less than the number of products required by the adaptation of the components. If necessary, the computational cost could be signficantly reduced in echo cancellation applications introducing selective-tap updates for the component filters [4], [16]-[18].

3. EXPERIMENTS

In this section, we evaluate the ability of combination schemes to improve the performance of IPNLMS using echo cancellation scenarios similar to those encountered in, e.g., [3, 4, 15, 8]. Different echo paths will be used, all with length $M = 512$ and an attenuation of 10 dB. Unless otherwise stated, the far-end signal, $x(n)$, from which input regressors are taken as $\mathbf{x}(n) = [x(n), x(n-1), \ldots, x(n-M+1)]^T$, is a white Gaussian noise with zero mean and variance 1. The output added noise, $e_0(n)$, is also a white Gaussian noise, whose power is set to get an SNR of 20 dB in the reference signal. We assume that no near-end signal is present (i.e., $s(n) = 0$), since the adaptation of echo cancelers is usually halted when double-talk situations are detected [19].

[2]Note that, for $\kappa_2 \approx 1$, the second filter behaves approximately as a PNLMS filter [8].

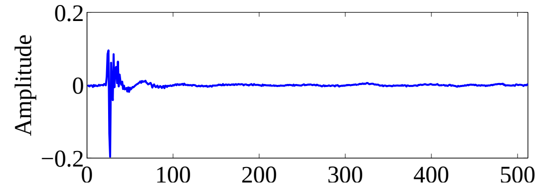


Fig. 3. Impulse response of the echo path identified by both adaptive schemes.

Parameter settings for the IPNLMS component filters are $\epsilon = 10^{-6}$ and $\delta = 0$, while different μ and κ values are used to illustrate the different benefits of the combination. The step size for the mixing parameter is set to $\mu_a = 100$ in all cases.

For comparison purposes, we will consider the estimation of the excess mean-square error (EMSE), $\text{EMSE}(n) = E\{[e(n) - e_0(n)]^2\}$, using an average of 1000 runs of the algorithms.

3.1. Improving the convergence speed vs steady-state misadjustment tradeoff

As we did in [11] for the particular case of LMS filters, we illustrate here how the convex combination scheme can be exploited, in a very easy and efficient way, to improve the speed of convergence vs steady-state performance compromise of IPNLMS filters.

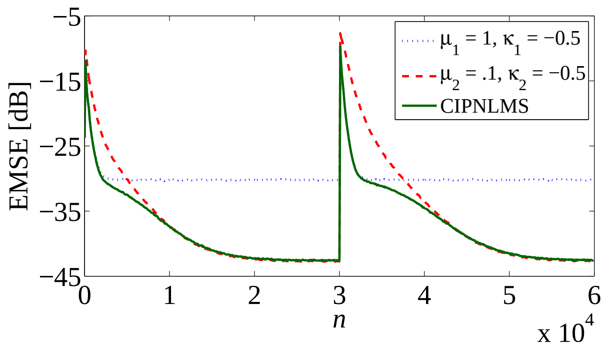
For this first set of experiments we have combined two IPNLMS filters with $\kappa_1 = \kappa_2 = -0.5$, as recommended in [8], using a different step size for each component: $\mu_1 = 1$ and $\mu_2 = 0.1$. The real echo path used in this subsection[3] corresponds to the impulse response of a hybrid sampled at a rate of 8 kHz, and has been depicted in Fig. 3. To study the ability of the algorithm to reconverge, a change in the echo path is introduced after a number of steps (see Fig. 4) by circularly shifting all coefficients 50 positions to the right.

Fig. 4(a) displays EMSE evolution for both component filters, as well as for their combination using the algorithm in Section 2, to which we will refer hereafter as combined IPNLMS (CIPNLMS). We can see that the μ_1 IPNLMS filter shows a faster convergence, both initially and after the change in the echo path. The filter with μ_2 presents a slower convergence, but it achieves a steady-state EMSE more than 10 dB smaller. As expected, the CIPNLMS filter is able to put together fast convergence and reduced steady-state misadjustment. In addition to this, we can check that CIPNLMS is able to simultaneously outperform both components at some iterations during the convergence.

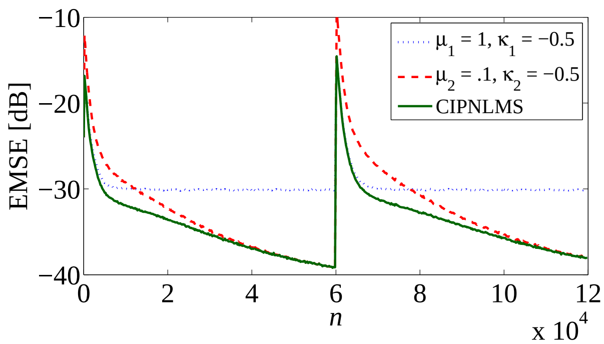
In Fig. 4(b) we illustrate the performance achieved for $x(n)$ being USASI noise with a speech-like spectrum [20][4]. The presence of a colored input causes a slower convergence

[3]This echo channel has been previously used in [8, 19].

[4]USASI noise was generated using the VOICEBOX toolbox for MATLAB (`http://www.ee.ic.ac.uk/hp/staff/dmb/voicebox/voicebox.html`).



(a)



(b)

Fig. 4. Mean-square performance of two IPNLMS filters with different step sizes, and of their adaptive combination. A change in the real echo path is introduced after 30000 and 60000 steps (for the upper and lower subfigures) by circularly shifting all coefficients 50 positions to the right. (a) $x(n)$ is white Gaussian noise. (b) $x(n)$ is USASI noise with a speech-like spectrum.

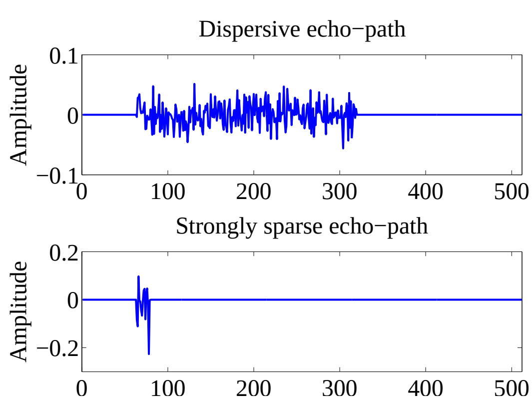
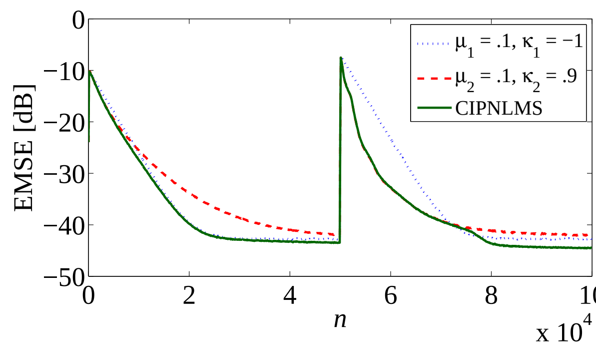
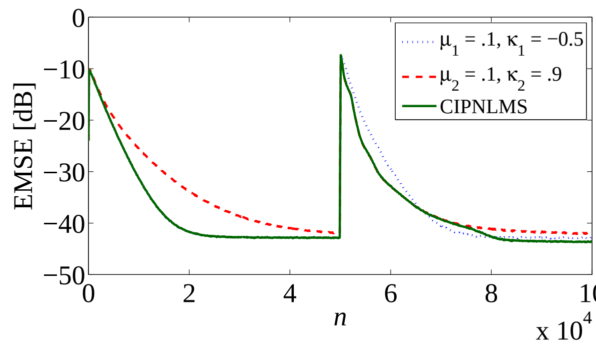


Fig. 5. Impulse response of two artificially generated echo channels, the first being a dispersive channel with 256 active coefficients, and the second being strongly sparse, with only 16 active coefficients.



(a)



(b)

Fig. 6. Performance of the combination scheme when combining IPNLMS filters with the same step size and different κ. The echo path has initially 256 active weights, while only 16 non-zero coefficients are present after the change at $n = 50000$.

of the component IPNLMS filters and, therefore, of their combination. Apart from this, the results shown in Fig. 4(b) can be discussed in very similar terms to the case of white Gaussian input.

3.2. Improving the robustness to channels with different degrees of sparsity

Next, we would like to illustrate how the CIPNLMS algorithm can be exploited to improve the robustness of IPNLMS to channels with different degrees of sparsity, even when the number of active coefficients changes during the simulation. To do so, we have generated a synthetic channel with 256 active coefficients taken from a random Gaussian distribution, and scaled to introduce 10 dB channel attenuation. In the second part of the simulation the echo path commutes to a different channel with only 16 active coefficients. These two artificial echo paths, which have been represented in Fig. 5, will be referred as the dispersive and the strongly sparse channels, respectively.

As it was explained at the end of Section 2, the adaptive combination approach can be used to improve IPNLMS robustness to channels with different degrees of sparsity by selecting $\mu_1 = \mu_2$, $\kappa_1 < 0$, and $\kappa_2 \approx 1$. Consequently, in this subsection, step sizes are adjusted to $\mu_1 = \mu_2 = 0.1$.

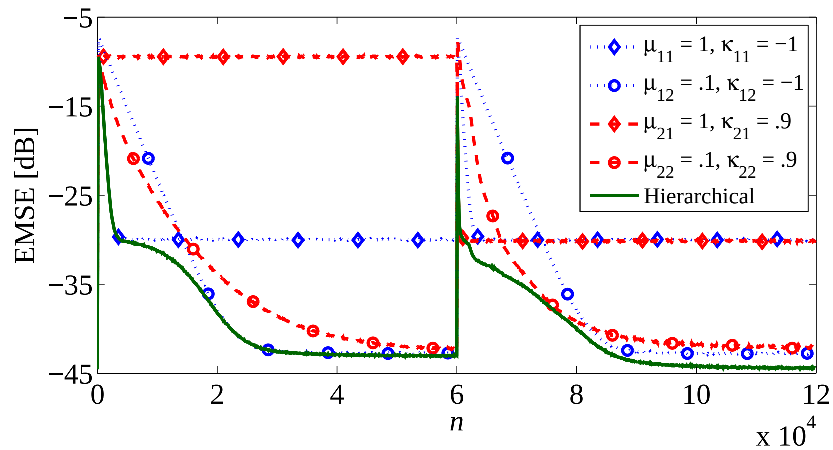


Fig. 7. EMSE evolution of the hierarchical combination scheme with four IPNLMS filters with different settings. The echo path has initially 256 active weights, while only 16 non-zero coefficients are present after the change.

As for parameter κ, different settings have been explored: $\kappa_1 = -1$ (for which the filter behaves as an NLMS filter) and $\kappa_1 = -0.5$ are used for the first component, while the second uses $\kappa_2 = 0.9$ in both cases, so that it behaves similarly to a PNLMS filter.

Fig. 6 shows that the combination scheme keeps again the best characteristics of each component. During the first part of the simulation, when the channel is not very sparse, the second (PNLMS) component has a slow convergence, but CIPNLMS remains robust to this situation by following the first filter. After the change at $n = 50000$, the echo path becomes very sparse, and the second component exhibits a fast convergence, which is also inherited by the combination. Then, we can conclude that the combination filter is robust to channels with different degrees of sparsity, in the sense that it achieves the very fast convergence of PNLMS if possible (i.e., for very sparse echo paths), while remaining robust to channels with wider active regions.

Finally, it is also interesting to notice that the steady-state EMSE of CIPNLMS can be simultaneusly smaller than those of both components, something which is specially clear during the last iterations in Fig. 6(a). This *better-than-universal* steady-state behavior, which has already been analyzed in [11], is a side advantage of the combination approach, and can be theoretically justified by a low cross-correlation between the *a priori* errors of the component filters.

3.3. Hierarchical combination of IPNLMS filters

In the two previous subsections we have illustrated that the adaptive combination of two adaptive filters can be used to improve two different aspects of the IPNLMS algorithm; now, we explain how the combination approach can be exploited to build an iterated architecture with the goal of simultaneously improving the convergence vs residual error tradeoff and the robustness to channels with different degrees of sparseness. The basic idea of the proposed hierarchical two-layer combination is to combine four IPNLMS filters. In the first level, we combine filters with the same κ and different step sizes, whereas in the second level we combine the outputs of the filters resulting from the first-level combinations.

EMSE evolution for such a hierarchical scheme has been depicted in Fig. 7. As before, the dispersive echo channel is used initially, and an abrupt change is simulated by commuting to the strongly sparse channel at step 60000. Parameter settings for the four component filters are shown in the legend of the figure. During the first part of the experiment, IPNLMS components with $\kappa = 0.9$ (thus, behaving like PNLMS filters) perform poorly, and the hierarchical scheme behaves like a convex combination of two NLMS filters with large and small step sizes. In the second half of the simulation, when the echo path is strongly sparse, the situation reverses: PNLMS-like components are the best performing among the four component filters, and the hierarchical scheme inherits their superior convergence.

Obviously, the cost we pay for using four component filters in parallel is an increased computational complexity, which is roughly twice that of the CIPNLMS scheme.

4. CONCLUSIONS

Proportionate schemes offer better behavior than standard adaptive filters for the cancellation of sparse echo-path impulse responses, but the selection of their parameters is subject to different compromises. In this paper we have shown how combination schemes can help to improve the performance of proportionate filters, by alleviating the speed vs precision tradeoff, as well as by increasing robustness to channels with different degrees of sparsity. The performance of such combination schemes has been illustrated when combining IPNLMS filters with different settings, showing in all cases significant advantages over the use of a single filter.

Current work includes the theoretical steady-state analysis of the presented schemes. Furthermore, we are also working in the design of new combination schemes which are specifically design to get the best out of this approach in echo cancellation applications.

Acknowledgements

The authors thank Dr. J. Benesty for kindly supplying the real echo-channel data.

5. REFERENCES

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

[2] C. Breining et. al., "Acoustic Echo Control. An application of very-high-order adaptive filters," in *IEEE Signal Process. Mag.*, vol. 16, pp. 42–69, 1999.

[3] J. Cui, P. A. Naylor, and D. T. Brown, "An improved IPNLMS algorithm for echo cancellation in packet-switched networks," in *Proc. ICASSP'04*, Montreal, Canada, vol. IV, 2004, pp. 141–144.

[4] P. A. Naylor, J. Cui, and M. Brookes, "Adaptive algorithms for sparse echo cancellation," *Signal Processing*, vol. 86, pp. 1182–1192, 2004.

[5] S. Haykin, *Adaptive Filter Theory*, Upper Saddle River, NJ: Prentice-Hall, 2002.

[6] A. H. Sayed, *Fundamentals of Adaptive Filtering*, NY: Wiley, 2003.

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

[8] J. Benesty and S. L. Gay, "An improved PNLMS algorithm," in *Proc. ICASSP'02*, Orlando, FL, vol. II, 2002, pp. 1881–1883.

[9] J. Arenas-García, M. Martínez-Ramón, A. Navia-Vázquez, and A. R. Figueiras-Vidal, "Plant identification via adaptive combination of transversal filters," *Signal Processing*, vol. 86, pp. 2430–2438, 2006.

[10] J. Arenas-García, V. Gómez-Verdejo, and A. R. Figueiras-Vidal, "New algorithms for improved adaptive convex combination of LMS transversal filers," *IEEE. Trans. Instr. Meas.*, vol. 54, pp. 2239–2249, 2005.

[11] J. Arenas-García, A. R. Figueiras-Vidal, and A. H. Sayed, "Mean-square performance of a convex combination of two adaptive filters," *IEEE Trans. Signal Process.*, vol. 54, 1078–1090, 2006.

[12] L. A. Azpicueta Ruiz, A. R. Figueiras-Vidal, and J. Arenas-García, "A normalized adaptation scheme for the convex combination of two adaptive filters," in *Proc. ICASSP'08*, Las Vegas, NV, pp. 3301-3304, 2008.

[13] N. J. Bershad, J. C. M. Bermudez, and J.-Y. Tourneret, "An affine combination of two LMS Adaptive Filters–Transient Mean-Square Analysis," *IEEE Trans. Signal Process.* vol. 56, pp. 1853-1864, 2008.

[14] S. L. Gay, "An efficient, fast converging adaptive filter for network echo cancellation," in *Proc. Asilomar Conf. on Signals, Systems and Computers*, Pacific Grove, CA, vol. 1, 1998, pp. 394–398.

[15] M. Nekuii and M. Atarodi, "A fast converging algorithm for network echo cancellation," *IEEE Signal Process. Lett.*, vol. 11, pp. 427–430, 2004.

[16] A. W. H. Khong and P. A. Naylor, "Selective-tap adaptive filtering with performance analysis for the identification of time-varying systems," *IEEE Trans. Audio, Speech and Language Process.*, vol. 15, 1681–1695, 2007.

[17] A. W. H. Khong and P. A. Naylor, "Stereophonic acoustic echo cancellation employing selective-tap adaptive algorithms," *IEEE Trans. Audio, Speech and Language Process.*, vol. 14, 785–796, 2007.

[18] T. Aboulnasr and K. Mayyas, "Complexity reduction of the NLMS algorithm via selective partial updates," *IEEE Trans. Signal Process.*, vol. 47, 1421–1424, 1999.

[19] T. Gänsler, S. L. Gay, M. Sondhi, and J. Benesty, "Double-talk robust fast converging algorithms for network echo cancellation," *IEEE Trans. Speech Audio Process.*, vol. 8, pp. 656–663, 2000.

[20] Standard ANSI/EIA-549-88, "NRSC AM Preemphasis/Deemphasis and Broadcast Audio Transmission Bandwidth Specifications (ANSI/EIA-549-88)," 1988.