MSE OPTIMAL REGULARIZATION OF APA AND NLMS ALGORITHMS IN ROOM ACOUSTIC APPLICATIONS

[1,2] *Toon van Waterschoot,* [2] *Geert Rombouts, and* [2] *Marc Moonen*

[1] toon.vanwaterschoot@esat.kuleuven.be
[2] Katholieke Universiteit Leuven, ESAT-SCD, Kasteelpark Arenberg 10, B-3001 Leuven, Belgium

ABSTRACT

A common way of increasing the robustness of affine projection and normalized least mean squares adaptive filtering algorithms, is to add a scaled identity regularization matrix to the input signal correlation matrix before inversion. This ad-hoc method can also be interpreted as the result of minimizing a regularized underdetermined least squares criterion. Moreover, by relating this criterion to linear minimum mean square error estimation, we can derive MSE optimal APA and NLMS algorithms, which feature a regularization matrix that is not necessarily a scaled identity matrix. The proposed algorithms allow for incorporating prior knowledge on both the near-end signal and the true room impulse response, and are intimately linked to Levenberg-Marquardt regularization and proportionate adaptation. Simulation results of echo and feedback cancellation experiments confirm that the adaptive filter convergence speed and tracking properties may be considerably improved using the proposed algorithms.

1. INTRODUCTION

Several acoustic signal processing applications, such as acoustic echo cancellation (AEC) [1] or adaptive feedback cancellation (AFC) [2], require the identification of a room impulse response (RIR). A typical AEC scenario is depicted in Fig. 1. The true RIR coefficients of the echo path between the loudspeaker and the microphone are collected in the parameter vector

$$\mathbf{f} \triangleq \begin{bmatrix} f_0 & f_1 & \dots & f_{n_F} \end{bmatrix}^T, \tag{1}$$

of known length $n_F + 1$. The loudspeaker plays back the far-end signal $u(t)$, which generates an echo signal $x(t) = \mathbf{u}^T(t)\mathbf{f}$ at the microphone position, with the loudspeaker signal vector defined as

$$\mathbf{u}(t) \triangleq \begin{bmatrix} u(t) & u(t-1) & \dots & u(t-n_F) \end{bmatrix}^T. \tag{2}$$

The echo signal $x(t)$ is picked up by the microphone, in addition to a local signal $v(t)$ referred to as the "near-end" signal, hence the microphone signal can be written as $y(t) = x(t) + v(t)$.

[1] Toon van Waterschoot is a Research Assistant with the IWT (Flemish Institute for Scientific and Technological Research in Industry). This research work was carried out at the ESAT laboratory of the Katholieke Universiteit Leuven, in the frame of the Belgian Programme on Interuniversity Attraction Poles, initiated by the Belgian Federal Science Policy Office IUAP P5/22 ("Dynamical Systems and Control: Computation, Identification and Modelling"), the Concerted Research Action GOA-AMBioRICS, and IWT project 040803: "SMS4PA-II: Sound Management System for Public Address Systems". The scientific responsibility is assumed by its authors.

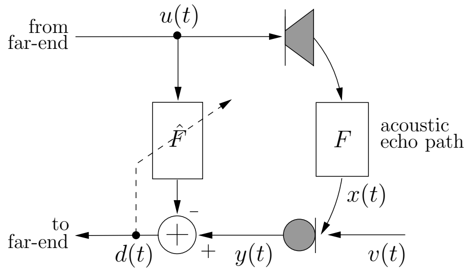

Figure 1: RIR identification in an AEC scenario.

It is important to note that the near-end signal is in most applications the signal of interest, however, from a system identification point of view, $v(t)$ is a disturbance to the estimation of the RIR. If at time t a RIR estimate $\hat{\mathbf{f}}(t)$ is available, then an echo-compensated signal can be calculated as $d(t) = y(t) - \mathbf{u}^T(t)\hat{\mathbf{f}}(t)$, which approximates the near-end signal $v(t)$.
A very popular recursive identification algorithm in room acoustic applications is the affine projection algorithm (APA) [1], due to its satisfactory convergence speed for colored input signals, and its $O(n_F)$ computational complexity per iteration. The APA with projection order M and step size μ is given by

$$\hat{\mathbf{f}}(t) = \hat{\mathbf{f}}(t-1) + \mu\mathbf{U}(t)[\mathbf{U}^T(t)\mathbf{U}(t) + \alpha\mathbf{I}]^{-1}\boldsymbol{\varepsilon}(t), \tag{3}$$

$$\boldsymbol{\varepsilon}(t) = \mathbf{y}(t) - \mathbf{U}^T(t)\hat{\mathbf{f}}(t-1), \tag{4}$$

with the data matrices defined as follows:

$$\mathbf{U}(t) \triangleq \begin{bmatrix} \mathbf{u}(t) & \dots & \mathbf{u}(t-M+1) \end{bmatrix}, \tag{5}$$

$$\mathbf{y}(t) \triangleq \begin{bmatrix} y(t) & \dots & y(t-M+1) \end{bmatrix}^T. \tag{6}$$

The well-known normalized least mean squares (NLMS) algorithm can then be obtained from the APA by setting $M = 1$:

$$\hat{\mathbf{f}}(t) = \hat{\mathbf{f}}(t-1) + \mu\frac{\mathbf{u}(t)\varepsilon(t)}{\mathbf{u}^T(t)\mathbf{u}(t) + \alpha}, \tag{7}$$

$$\varepsilon(t) = y(t) - \mathbf{u}^T(t)\hat{\mathbf{f}}(t-1). \tag{8}$$

In this paper, we focus on the regularization part of the above algorithms. A common problem in room acoustic applications is that the matrix $\mathbf{U}^T(t)\mathbf{U}(t)$, appearing in (3), is ill-conditioned or even singular, due to poor excitation. If no regularization is applied, the lack of excitation may lead to divergence of the adaptive filter, in particular during double-talk periods (i.e., when both the far-end and near-end signals are non-zero). The standard solution consists in adding an $M \times M$ scaled identity matrix $\alpha\mathbf{I}$ to $\mathbf{U}^T(t)\mathbf{U}(t)$ before inverting the matrix. It is known that the regularization parameter α should ideally be an a priori estimate of the near-end signal power [1], yet this parameter is often chosen to be an arbitrary, small number.

Whereas this traditional regularization approach is an ad-hoc method to avoid the inversion of an ill-conditioned matrix, the regularized APA and NLMS algorithms described above can also be obtained by minimization of a criterion which originates from minimizing the mean square error (MSE) between the estimated and true RIR. In this framework, it appears that the existing approach towards regularization can be optimized by taking into account any prior knowledge, not only on the near-end signal, but also on the true RIR.

In Section 2, we briefly present the MSE optimal approach towards regularization, referring to [3] for a more detailed explanation. We also indicate how a diagonal (and not necessarily scaled identity) regularization matrix can be constructed, that approaches the MSE optimal regularization matrix, using a previously proposed 3-parameter RIR model [3]. Then in Section 3, we show how minimization of the proposed criterion leads to MSE optimal APA and NLMS algorithms, which exhibit features known as leakage and proportionate adaptation. The resulting algorithms are moreover intimately linked to the widely used Levenberg-Marquardt regularization approach for recursive least squares (RLS) algorithms [3], [4], and provide new insight in the properties of the proportional NLMS (PNLMS) [5], [6] and proportionate APA (PAPA) [7] algorithms. Finally, in Section 4, simulation results are shown, which confirm the improved convergence behaviour of the proposed algorithms, both in echo and feedback cancellation problems.

2. MSE OPTIMAL REGULARIZATION

In [3], it is shown that there is a strong link between regularization and linear minimum mean square error estimation. In particular, minimizing the MSE criterion

$$\min_{\hat{\mathbf{f}}(t)} E\Big\{[\hat{\mathbf{f}}(t)-\mathbf{f}]^T[\hat{\mathbf{f}}(t)-\mathbf{f}]\Big\} \tag{9}$$

is shown to be equivalent to minimizing a weighted and regularized least squares criterion:

$$\min_{\hat{\mathbf{f}}(t)}\Big\{[\mathbf{y}(t)-\mathbf{U}^T(t)\hat{\mathbf{f}}(t)]^T\mathbf{R_v}^{-1}(t)[\mathbf{y}(t)-\mathbf{U}^T(t)\hat{\mathbf{f}}(t)] + [\hat{\mathbf{f}}(t)-\mathbf{f_0}]^T\mathbf{R_f}^{-1}[\hat{\mathbf{f}}(t)-\mathbf{f_0}]\Big\}. \tag{10}$$

The equivalence holds only in a Bayesian framework, in which not only the near-end signal vector $\mathbf{v}(t) \triangleq [v(t)\ \ldots\ v(t-M+1)]$, but also the true RIR is considered to be drawn from a stochastic vector process, on which some prior knowledge may be available through their means and covariance matrices, defined as

$$\begin{cases} E\{\mathbf{v}(t)\} = \mathbf{0}, & (11)\\ \mathrm{cov}\{\mathbf{v}(t)\} = E\big\{\mathbf{v}(t)\mathbf{v}^T(t)\big\} = \mathbf{R_v}(t), & (12)\end{cases}$$

$$\begin{cases} E\{\mathbf{f}\} = \mathbf{f_0}, & (13)\\ \mathrm{cov}\{\mathbf{f}\} = E\big\{(\mathbf{f}-\mathbf{f_0})(\mathbf{f}-\mathbf{f_0})^T\big\} = \mathbf{R_f}. & (14)\end{cases}$$

The criterion in (10) features both a least squares data term, and a regularization term penalizing the deviation of the RIR estimate from the true RIR expected value. The relative importance of these two terms is governed by the inverse covariance matrices of the near-end signal and true RIR distributions.

Before deriving the MSE optimal APA and NLMS algorithms, we comment on the choice of the covariance matrices $\mathbf{R_v}(t)$ and $\mathbf{R_f}$. In this paper, we assume that the near-end signal is drawn from a Gaussian white noise process with variance $\sigma_v^2(t)$, such that $\mathbf{R_v}(t)$ is a diagonal matrix:

$$\mathbf{R_v}(t) = \mathrm{diag}\{\sigma_v^2(t), \sigma_v^2(t-1), \ldots, \sigma_v^2(t-M+1)\}. \tag{15}$$

The general case of a non-white near-end signal is described in [3]. Also, in the sequel, the true RIR covariance matrix $\mathbf{R_f}$ is restricted to be diagonal. This leads to an $O(n_F)$ computational complexity when evaluating matrix-vector products involving $\mathbf{R_f}$, and still allows n_F+1 degrees of freedom in contrast to the traditional scaled identity matrix approach with only one degree of freedom. We may construct a diagonal estimate for $\mathbf{R_f}$ by collecting prior knowledge on the acoustic setup, e.g., the room volume, loudspeaker and microphone positions, wall absorption coefficients, etc., as described in [3].

3. REGULARIZED APA AND NLMS ALGORITHMS

Minimizing the criterion in (10), and subsequently applying the matrix inversion lemma, leads to the following underdetermined estimate:

$$\hat{\mathbf{f}}(t)=\mathbf{f_0}+\mathbf{R_f}\mathbf{U}(t)[\mathbf{U}^T(t)\mathbf{R_f}\mathbf{U}(t)+\mathbf{R_v}(t)]^{-1}[\mathbf{y}(t)-\mathbf{U}^T(t)\mathbf{f_0}].$$

A recursive algorithm can be obtained by explicitly bringing in the dependency on $\hat{\mathbf{f}}(t-1)$, which leads to:

$$\begin{aligned}\hat{\mathbf{f}}(t) &= \mathbf{f_0}\\ &+\{\mathbf{I}-[\mathbf{U}(t)\mathbf{R_v}^{-1}(t)\mathbf{U}^T(t)+\mathbf{R_f}^{-1}]^{-1}\mathbf{R_f}^{-1}\}[\hat{\mathbf{f}}(t-1)-\mathbf{f_0}]\\ &+\mathbf{R_f}\mathbf{U}(t)[\mathbf{U}^T(t)\mathbf{R_f}\mathbf{U}(t)+\mathbf{R_v}(t)]^{-1}[\mathbf{y}(t)-\mathbf{U}^T(t)\hat{\mathbf{f}}(t-1)].\end{aligned}$$

It can be seen that the minimizing estimate of the criterion in (10), consists of three terms: the mean value $\mathbf{f_0}$ of the true RIR distribution, a leakage term depending on the deviation $[\hat{\mathbf{f}}(t-1)-\mathbf{f_0}]$ of the previous estimate from the mean value, and a proportionate adaptation term. The leakage term disappears by choosing $\mathbf{f_0} = \hat{\mathbf{f}}(t-1)$, and hence we obtain a Levenberg-Marquardt type of regularization [3]. If we finally introduce a relaxation factor μ, we end up with the so-called Levenberg-Marquardt regularized affine projection algorithm (LMR-APA) and, when the projection order is set to $M=1$, the Levenberg-Marquardt regularized normalized least mean squares (LMR-NLMS) algorithm, which are shown in Table 1. Note that the traditionally regularized APA and NLMS algorithms, described in the Introduction, can be obtained as special cases of the proposed algorithms, by choosing $\mathbf{R_v}(t) = \sigma\mathbf{I}$ and $\mathbf{R_f} = \nu\mathbf{I}$ such that $\sigma\nu^{-1} = \alpha$.

The LMR-APA and LMR-NLMS algorithms are closely related to the proportionate APA (PAPA) [7] and proportionate NLMS (PNLMS) [5] algorithms. In the PNLMS RIR weight update,

$$\hat{\mathbf{f}}(t) = \hat{\mathbf{f}}(t-1) + \mu\frac{\mathbf{G}(t)\mathbf{u}(t)\varepsilon(t)}{\mathbf{u}^T(t)\mathbf{G}(t)\mathbf{u}(t)+\alpha}, \tag{16}$$

the diagonal matrix $\mathbf{G}(t)$ is constructed as follows:

$$\mathbf{G}(t) = \frac{1}{\bar{g}(t)}\ \mathrm{diag}\{g_0(t), \ldots, g_{n_F}(t)\}, \tag{17}$$

Table 1: MSE optimally regularized APA and NLMS algorithms

LMR-APA - Levenberg-Marquardt Regularized APA

$$\hat{\mathbf{f}}(t) = \hat{\mathbf{f}}(t-1) + \mu \mathbf{R_f} \mathbf{U}(t)[\mathbf{U}^T(t)\mathbf{R_f}\mathbf{U}(t) + \mathbf{R_v}(t)]^{-1}\boldsymbol{\varepsilon}(t),$$
$$\boldsymbol{\varepsilon}(t) = \mathbf{y}(t) - \mathbf{U}^T(t)\hat{\mathbf{f}}(t-1).$$

LMR-NLMS - Levenberg-Marquardt Regularized NLMS

$$\hat{\mathbf{f}}(t) = \hat{\mathbf{f}}(t-1) + \mu \frac{\mathbf{R_f}\mathbf{u}(t)\varepsilon(t)}{\mathbf{u}^T(t)\mathbf{R_f}\mathbf{u}(t) + \sigma_v^2(t)},$$
$$\varepsilon(t) = y(t) - \mathbf{u}^T(t)\hat{\mathbf{f}}(t-1).$$

with, for $k = 0, \ldots, n_F$,

$$g_k(t) \triangleq \max\Big\{\rho \cdot \max\big\{\delta, |\hat{f}_0(t-1)|, \ldots, |\hat{f}_{n_F}(t-1)|\big\}, |\hat{f}_k(t-1)|\Big\}, \quad (18)$$

$$\bar{g}(t) \triangleq \frac{1}{n_F+1}\sum_{k=0}^{n_F} g_k(t), \quad (19)$$

where ρ and δ are small positive parameters. This choice of $\mathbf{G}(t)$ was made somewhat intuitively by Duttweiler [5] with the aim of allocating a larger portion of the available adaptation energy to larger adaptive filter weights, to speed up the convergence. Later on, Chen et al. [6] provided an interpretation of the above choice of $\mathbf{G}(t)$ in terms of Bayesian priors.

The LMR-NLMS algorithm provides an alternative choice for the matrix $\mathbf{G}(t)$ in the PNLMS algorithm, which is optimal in the sense of minimizing the criterion in (10). The above comparison of LMR-NLMS with PNLMS can be done similarly for the LMR-APA and PAPA algorithms. An important difference with the existing proportionate adaptation algorithms is that in the proposed LMR-APA and LMR-NLMS algorithms, the regularization matrix $\mathbf{R_f}$ is fixed. As a consequence, the proposed algorithms may respond somewhat slower to a RIR change, but on the other hand they are much less computationally demanding than the PAPA and PNLMS algorithms, in which $\mathbf{G}(t)$ is recalculated with (17)-(19) in each iteration of the adaptive filter.

4. SIMULATION RESULTS

All simulations are done in Matlab at a sampling frequency $f_s =$ 8 kHz. The performance measure for comparing the different algorithms is the misadjustment, which is defined as the normalized Euclidian distance between the estimated and true RIR on a logarithmic scale:

$$\text{misadjustment (dB)} = 20\log_{10}\frac{\|\hat{\mathbf{f}}(t) - \mathbf{f}\|}{\|\mathbf{f}\|}. \quad (20)$$

4.1. Regularized NLMS Algorithms for Acoustic Echo Cancellation in a Non-Stationary Environment

In a first simulation, the performance of regularized NLMS algorithms is compared for acoustic echo cancellation in a non-stationary environment. We switch between three different room impulse responses of known length $n_F + 1 = 1000$, that were measured in our acoustic lab. The first RIR change occurs around $t/T_s = 1.2 \cdot 10^5$ samples, and consists in a change of 75 cm in the microphone position, in such a way that the distance between the loudspeaker and the microphone remains constant. After the second RIR change, around $t/T_s = 2.5 \cdot 10^5$ samples, the loudspeaker and microphone positions remain unchanged, but the room is made more reverberant by decreasing the absorption coefficients of the walls and ceiling. In this simulation, the far-end signal is a 46 s male speech signal (equivalent to $N = 368320$ samples), and the near-end signal is a stationary GWN signal with known variance $\sigma_v^2 = 3 \cdot 10^{-5}$, resulting in an echo-to-near-end ratio at the microphone $\text{ENR}_{\text{mic}} = 16$ dB.

The reference algorithms are the (unregularized) standard NLMS algorithm, the proportionate NLMS algorithm with $\rho = 5/(n_F+1)$ and $\delta = 0.01$ as suggested in [5], and the traditionally regularized NLMS algorithm as given in (7)-(8) with $\alpha = \sigma_v^2$, which is further called LMR-NLMS $\alpha\mathbf{I}$. These algorithms are compared to the proposed LMR-NLMS $\hat{\mathbf{R}}_{\mathbf{f},\mathbf{3}}$ and LMR-NLMS $\hat{\mathbf{R}}_{\mathbf{f},\mathbf{true}}$ algorithms in which the diagonal regularization matrix is based respectively on the 3-parameter model from [3] (with $d = 75$, $A = 0.1022$, $\tau = 70$ and $\beta = 10^{-6}$), and on exact knowledge of the first RIR (i.e., a "best-case" scenario). The regularization matrix is not altered after the two RIR changes, such that the robustness of the different regularized algorithms w.r.t. RIR changes can be evaluated. The step size μ is individually tuned for each of the algorithms such that the excess MSE in a stationary environment would approximately be the same for all algorithms.

The convergence curves and step sizes are shown in Fig. 2. First of all, we observe that the improvement in convergence speed of the existing PNLMS and LMR-NLMS $\alpha\mathbf{I}$ algorithms, as compared to the unregularized NLMS algorithm, is small. A significantly better convergence behaviour is obtained with the proposed LMR-NLMS $\hat{\mathbf{R}}_{\mathbf{f},\mathbf{3}}$ and LMR-NLMS $\hat{\mathbf{R}}_{\mathbf{f},\mathbf{true}}$ algorithms in the time interval $t/T_s = [0, 1.2 \cdot 10^5]$ samples, where the regularization matrix of the proposed algorithms is based on the "correct" RIR. After the first RIR change, the LMR-NLMS $\hat{\mathbf{R}}_{\mathbf{f},\mathbf{true}}$ algorithm's performance decreases dramatically, whereas the LMR-NLMS $\hat{\mathbf{R}}_{\mathbf{f},\mathbf{3}}$ algorithm converges as fast as initially. This is not much of a surprise, since the regularization matrix $\hat{\mathbf{R}}_{\mathbf{f},\mathbf{true}}$ based on the true first RIR, will be a bad model for the second RIR covariance matrix, whereas the 3-parameter model regularization matrix $\hat{\mathbf{R}}_{\mathbf{f},\mathbf{3}}$ of the first RIR will still be valid for the second RIR. Indeed, the microphone repositioning does not alter any of the three parameters on which $\hat{\mathbf{R}}_{\mathbf{f},\mathbf{3}}$ is based, since the distance between the loudspeaker and the microphone remains constant. However, after the second RIR change the parameter τ in $\hat{\mathbf{R}}_{\mathbf{f},\mathbf{3}}$ will have an inaccurate value, since the room reverberation has increased. This clearly affects the LMR-NLMS $\hat{\mathbf{R}}_{\mathbf{f},\mathbf{3}}$ convergence speed, but still this algorithm outperforms the other algorithms.

4.2. Regularized NLMS Algorithms with Prefiltering for Feedback Cancellation in a Stationary Environment

In a second simulation, the proposed LMR-NLMS algorithm is applied in a closed-loop scenario for performing adaptive feedback cancellation. In this case, the near-end signal is non-white such that prefiltering of the loudspeaker and microphone signals with the inverse near-end signal model is desirable [2]. The near-

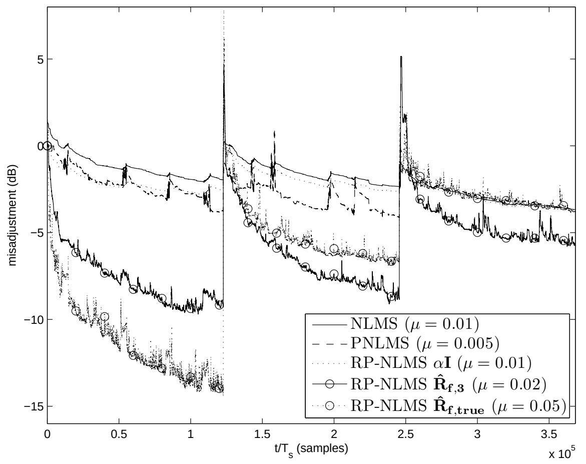

Figure 2: Convergence curves of regularized NLMS algorithms for an AEC application with two echo path changes.

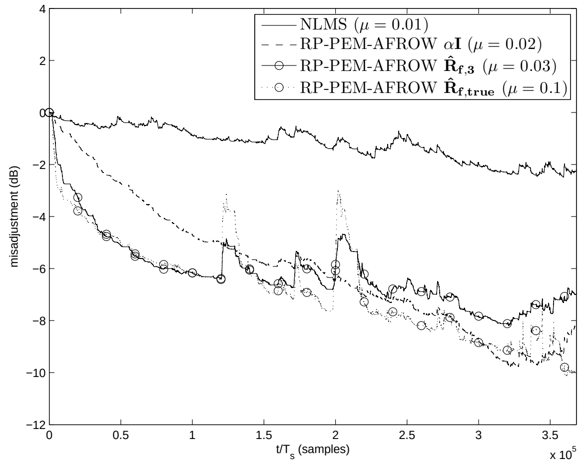

Figure 3: Convergence curves of regularized NLMS algorithms with prefiltering for an AFC application in a stationary environment.

end signal $v(t)$, equal to the same male speech signal as in the previous simulation, is added to the feedback signal before being amplified and delayed in the forward path. The resulting signal is sent to the loudspeaker, after which it is filtered in the feedback path to yield the feedback signal. An AR(12) model of the near-end signal as well as the near-end signal variance $\sigma_v^2(t)$ are identified by linear prediction, as in the PEM-AFROW algorithm described in [2]. In this simulation, the feedback path is a 1000-tap, measured RIR. The forward path delay is set equal to the linear prediction window length $L = 160$ samples, as suggested in [2], and the forward path gain is set to $K = -19$ dB, resulting in a closed-loop gain margin of 3 dB. These settings lead to $\text{ENR}_{\text{mic}} = -11$ dB.

The regularized algorithms are compared with the unregularized NLMS algorithm (without prefiltering) in Fig. 3. It is clear that the LMR-NLMS algorithms with prefiltering, are better suited to the closed-loop estimation problem than the NLMS algorithm. We further note that initially, the LMR-NLMS $\hat{\mathbf{R}}_{\mathbf{f},\mathbf{3}}$ and LMR-NLMS $\hat{\mathbf{R}}_{\mathbf{f},\mathbf{true}}$ algorithms converge considerably faster than the LMR-NLMS $\alpha\mathbf{I}$ algorithm, yet after some time the misadjustment of these three algorithms settles down to approximately the same level.

5. CONCLUSION

In this paper, we have shown how the use of more than one regularization parameter, may lead to an increased performance of regularized affine projection and normalized least mean squares adaptive filtering algorithms. Based on the equivalence between an MSE estimation criterion and a regularized least squares criterion, we have illustrated how optimal regularization can be achieved by taking into account any prior knowledge on the near-end signal and on the true room impulse response. The proposed Levenberg-Marquardt regularized APA and NLMS algorithms moreover provide a new interpretation of the existing proportionate adaptation algorithms. Simulations point out that the proposed algorithms exhibit improved convergence behaviour, and may have nice tracking properties if the regularization matrix is constructed in a proper way, e.g., using a previously proposed 3-parameter RIR model.

6. REFERENCES

[1] S. L. Gay, “Fast projection algorithms with application to voice echo cancellation,” Ph.D. dissertation, Rutgers, The State University of New Jersey, New Brunswick, New Jersey, USA, Oct. 1994.

[2] G. Rombouts, T. van Waterschoot, K. Struyve, and M. Moonen, “Acoustic feedback suppression for long acoustic paths using a nonstationary source model,” *IEEE Trans. Signal Process.*, to appear. [Online]. Available: ftp://ftp.esat.kuleuven.be/pub/sista/vanwaterschoot/abstracts/04-151.html

[3] T. van Waterschoot, G. Rombouts, and M. Moonen, “Optimally regularized adaptive filtering algorithms for room acoustic signal enhancement,” *Signal Processing*, submitted for publication. [Online]. Available: ftp://ftp.esat.kuleuven.be/pub/sista/vanwaterschoot/abstracts/05-184.html

[4] L. Ljung and T. Söderström, *Theory and practice of recursive identification*. Cambridge, Massachusetts: MIT Press, 1986.

[5] D. L. Duttweiler, “Proportionate normalized least-mean-squares adaptation in echo cancelers,” *IEEE Trans. Speech Audio Process.*, vol. 9, no. 5, pp. 508–518, Sept. 2000.

[6] Z. Chen, S. Haykin, and S. L. Gay, “Proportionate adaptation: New paradigms in adaptive filtering,” in *Advances in LMS filters*, S. Haykin and B. Widrow, Eds. New York: Wiley, 2003, ch. 8.

[7] T. Gänsler, J. Benesty, S. L. Gay, and M. M. Sondhi, “A robust proportionate affine projection algorithm for network echo cancellation,” in *Proc. 2000 IEEE Int. Conf. Acoust., Speech, Signal Process. (ICASSP ’00)*, vol. 2, Istanbul, Turkey, June 2000, pp. 793–796.