LMS ALGORITHM WITH REUSE OF COEFFICIENTS AND ROBUSTNESS AGAINST IMPULSIVE NOISE

Leonardo C. Resende*, Newton N. Siqueira†, Rodrigo M.S. Pimenta‡, Diego B. Haddad§, Mariane R. Petraglia¶

*Instituto Federal do Rio de Janeiro
Paracambi, RJ, Brasil

†Coordenação de Telecomunicações
CEFET-RJ Nova Iguaçu, RJ, Brasil.

‡Coordenação de Telecomunicações
CEFET-RJ Maracanã
Rio de Janeiro-RJ, Brasil.

§Coordenação de Engenharia de Computação
CEFET-RJ Petrópolis, RJ, Brasil

¶Programa de Engenharia Elétrica (PEE)
COPPE/UFRJ, Rio de Janeiro-RJ, Brasil.

Emails: leonardo.resende@ifrj.edu.br, newton.siqueira@cefet-rj.br, rodrigo.pimenta@cefet-rj.br, diego@pads.ufrj.br, mariane@pads.ufrj.br.

Abstract— This article contributes with two major improvements related to the learning performance of the least mean squares (LMS) adaptive filtering algorithm under challenging scenarios, low signal-to-noise ratio and in the presence of impulsive noise. The first problem is solved through the insertion of the coefficient reuse technique in the update equation, while the second one uses the maximum correntropy criterion to make the algorithm less sensitive to impulsive noise, which is very often present in real applications.

Keywords— Adaptive filtering, LMS, Reuse of Coefficients, Maximum Correntropy.

1 Introduction

Adaptive filtering (AF) techniques can be employed to electronically emulate the acoustic coupling between a loudspeaker and a microphone. Acoustic echo cancellers commonly use this system identification technique to establish a duplex communication (Kumar et al., 2017). Other applications can still be contemplated by the AF strategies, such as control, channel equalization and time series prediction (Diniz, 1997). For the most part, the AF algorithms perform a nonlinear estimation of the coefficients in a transversal structure, which are collected in a vector[1] $\boldsymbol{w}(k) \in \mathbb{R}^N$, where N is an adjustable parameter.

Among the various AF algorithms, the LMS (*Least Mean Squares*) is the most popular. Usually, it is argued that such widespread usage is due to its low computational complexity (Venkatesan et al., 2018) and its ability to approximate the supposedly optimal Wiener solution. More recently, several articles have highlighted some LMS optimization properties that demonstrate their superiority over the Wiener filter under certain contexts. These properties further motivate the adoption of the LMS algorithm. The delay in realizing that the LMS can overcome the Wiener filter seems to be due to the fact that classical analysis techniques, such as the independence hypotheses (Quirk et al., 2000), relax the dependence of the LMS estimator on the previous input vectors $\boldsymbol{x}(k)$, defined by

$$\boldsymbol{x}(k) \triangleq \begin{bmatrix} x(k) & x(k-1) & \dots & x(k-N+1) \end{bmatrix}^T, \tag{1}$$

where $x(k)$ is k-th input signal sample (Quirk et al., 1998).

According to (Ikuma et al., 2008), the authors argue that the mean coefficients of an equalizer based on the steady-state LMS solution differs from the Wiener coefficients in the presence of narrow band additive noise $\nu(k)$, and that the mean square error of the LMS filter may be significantly better than the Wiener filter solution. Interestingly, this result is not restricted to small values of the learning factor β (Ikuma et al., 2007; Ikuma et al., 2008). The approach of (Reuter and Zeidler, 1999), approximates analytic expressions for the mean square error (MSE) of the LMS algorithm that is developed through the construction of transfer functions. Depending on the signal-to-noise ratio (SNR), the signal-to-interference ratio (SIR), the equalizer length N and the learning factor, the LMS presents advantages when compared to the Wiener filter.

The results of (Reuter and Zeidler, 1999) were refined by (Ikuma and Beex, 2008), with similar

[1] All vectors in this article are column type.

conclusions. These properties (including the fact that the LMS presents optimality $\mathcal{H}^\infty$ (Hassibi et al., 1996)) demonstrate that this algorithm can overcome, under certain situations, its normalized versions, such as the normalized LMS (NLMS) algorithm. Instead of minimizing the MSE, the NLMS minimizes the following cost function by means of the stochastic gradient technique:

$$\mathcal{F}[\boldsymbol{w}(k)] = \mathbb{E}\left[\frac{e^2(k)}{\|\boldsymbol{x}(k)\|^2}\right], \tag{2}$$

whose minimum may differ[2] from the MSE $\xi(k) \triangleq \mathbb{E}\left[e^2(k)\right]$, where the error $e(k)$ consists of the difference between the reference signal $d(k)$ and the filter output $y(k)$:

$$e(k) \triangleq d(k) - \overbrace{\boldsymbol{w}^T(k)\boldsymbol{x}(k)}^{\triangleq y(k)}. \tag{3}$$

It is known that the LMS shows a steady-state performance deterioration when the SNR is low (Sayed, 2011). This article presents an optimization problem whose solution is an LMS-type algorithm with coefficient reuse (RC), a technique originally proposed for normalized algorithms (Kim et al., 2011). The resulting algorithm (RC-LMS), however, exhibits great sensitivity to impulsive noise, a phenomenon that can be caused, for example, by double-talk in acoustic echo cancellation systems (Petraglia et al., 2016) or by atmospheric phenomena in telecommunication systems (Das and Narwaria, 2017). In order to address these ubiquitous phenomena, this paper also uses the maximum correntropy criterion to make the performance of the proposed RC-LMS algorithm relatively insensitive to impulsive noise.

This article is structured as follows. In Section 2, we describe the coefficient reuse strategy (in qualitative terms), which motivates the derivation of the first algorithm proposed in Section 3, whose unbiased property of the estimation is demonstrated by means of a theorem. The adoption of the maximum correntropy criterion, described in Section 4, allows insertion of the robustness to impulsive interferences in the previously derived algorithm, which is described in Section 5. Section 6 verifies the advantages in the performances of the proposed algorithms, while Section 7 contains the final conclusions of the article. The three theorems enunciated in Sections 3 and 5 are finally demonstrated in the Appendix.

2 Reuse of Coefficients

Adaptive filtering algorithms that employ the coefficient reuse technique present better convergence rates with good steady-state responses (Cho et al., 2009). Such reusing takes explicitly into account the adaptive coefficients obtained in past iterations, which mitigates the magnitude of the oscillations of the adaptive estimation. Several algorithms, such as the APA (Affine-Projection Algorithm) increase the convergence rate using the input data reuse criterion (Simon, 2002). These algorithms usually present a loss of performance in steady-state condition, being considered dual in relation to the algorithms that use the coefficient reuse technique, like the algorithm proposed in (Kim et al., 2011). The RC family algorithms show a good steady-state performance and a (sometimes imperceptible) loss in the convergence rate (Kim et al., 2011), which can be explained by the fact that the RC strategy uses the last L vectors of adaptive coefficients ($\boldsymbol{w}(k-l)$, $l \in \{0, \cdots, L-1\}$), whose effect is to smooth the oscillations of the parameters to be estimated.

3 RC-LMS

The first contribution of this article is to propose a constrained optimization problem whose solution (obtained by the Lagrange multiplier technique) gives rise to an LMS algorithm with coefficient reuse, capable of improving the steady-state performance of the LMS algorithm. The resulting algorithm is obtained by Theorem 1 below.

Theorem 1. Consider the following optimization problem:

$$\min_{\boldsymbol{w}(k+1)} \mathcal{F}_{\mathrm{RC}}[\boldsymbol{w}(k+1)] \tag{4}$$

$$\text{s.t. } e_p(k) = (1 - \beta\|\boldsymbol{x}(k)\|^2)\overline{e}(k),$$

where

$$\mathcal{F}_{\mathrm{RC}}[\boldsymbol{w}(k+1)] \triangleq \sum_{l=0}^{L-1} \rho^l \|\boldsymbol{w}(k+1) - \boldsymbol{w}(k-l)\|^2 \tag{5}$$

and

$$\overline{e}(k) \triangleq d(k) - \left(\frac{\rho - 1}{\rho^L - 1}\right) \sum_{l=0}^{L-1} \rho^l \boldsymbol{w}^T(k-l)\boldsymbol{x}(k), \tag{6}$$

where $\rho \in (0, 1]$ and $L \in \mathbb{N}$ are factors at the discretion of the designer and the error *a posteriori* is defined by

$$e_p(k) \triangleq d(k) - \boldsymbol{w}^T(k+1)\boldsymbol{x}(k). \tag{7}$$

It can be proved that the update equation that solves (4) is given by

$$\boldsymbol{w}(k+1) = \left(\frac{\rho - 1}{\rho^L - 1}\right) \sum_{l=0}^{L-1} \rho^l \boldsymbol{w}(k-l) + \beta\overline{e}(k)\boldsymbol{x}(k). \tag{8}$$

Demonstration: see Appendix.
Remarks: Considering that $\mathcal{F}_{\mathrm{RC}}[\boldsymbol{w}(k+1)]$ is a cost function that penalizes solutions $\boldsymbol{w}(k+1)$ that are

[2] Such difference tends to be emphasized when the equalizer size is small, in which case the correlation between the random variables $\|\boldsymbol{x}(k)\|^2$ and $e^2(k)$ tends to be greater.

very distant from the previous vectors $\boldsymbol{w}(k-l)$, $l \in \{0, 1, \ldots, L-1\}$, it gives rise to the update equation (8) of the proposed RC-LMS algorithm. When $L = 1$, $\mathcal{F}_{\mathrm{RC}}[\boldsymbol{w}(k+1)]$ degenerates in the classic *Minimum Distortion Principle* cost function (Simon, 2002), giving rise to the LMS (particular case of the proposed algorithm).

The relative weight given to the previous vectors can be controlled by the factor ρ, with higher values of ρ tending to give similar weights to the L previous vectors. The L parameter controls the magnitude of the coefficient reuse, and can be changed dynamically to maximize the transient performance (Kim et al., 2011). An important feature of the LMS is that it is an unbiased estimator of $\boldsymbol{w}^\star \in \mathbb{R}^N$ (the optimal filter). The following theorem demonstrates that RC-LMS inherits this LMS property.

Theorem 2. Let the input signal $x(k)$ be stationary with full rank autocorrelation matrix $\boldsymbol{R} \triangleq \mathbb{E}\left[\boldsymbol{x}(k)\boldsymbol{x}^T(k)\right]$, and $\boldsymbol{w}^\star$ be the ideal (and unknown) plant that the algorithm intends to identify, with the reference signal described by:

$$d(k) = \left(\boldsymbol{w}^\star\right)^T \boldsymbol{x}(k) + \nu(k). \tag{9}$$

The following hypotheses are considered:
$H1$. The filter converges at the steady state;
$H2$. The additive noise $\nu(k)$ presents zero mean and is independent of the other random variables involved;
$H3$. The filter coefficients $\boldsymbol{w}(k)$ are independent of the input $\boldsymbol{x}(k)$.
Considering $H1 - H3$, we can assert that the RC-LMS is an unbiased estimator. It should be noted that $H3$ is the strongest hypothesis of all (particularly in a transverse structure), being known in the literature as *independence hypothesis* (Haykin and Widrow, 2003).
Demonstration: see Appendix.

4 Robustness to Impulsive Noise

The use of the maximum correntropy criterion (MCC) may result in robust AF algorithms for impulsive noise (Singh and Principe, 2009). The derivation of such algorithms, however, employs a stochastic gradient, not a constrained formulation like (4). The insertion of the MCC into algorithms that were derived through Lagrange multipliers was first proposed in (Haddad et al., 2016). The resulting algorithms are able to cope with non-Gaussian noise scenarios, which are common in real-world applications (Liu et al., 2017). Correntropy is a measure of local similarity between random variables X and Y given by (Liu et al., 2007)

$$V(X, Y) = \iint_{x,y} \kappa_\sigma(x-y) f_{XY}(x, y)\,dxdy, \tag{10}$$

where $f_{XY}(x, y)$ is the joint probability density function of random variables X and Y, and the Gaussian kernel[3] $\kappa_\sigma(x-y)$ is

$$\kappa_\sigma(x-y) \triangleq \frac{1}{\sqrt{2\pi}\sigma} \exp\left[-\frac{(x-y)^2}{2\sigma^2}\right] \tag{11}$$

and σ determines the width of the Gaussian kernel, which influences the trade-off between steady-state performance and convergence rate (Liu et al., 2011). The Gaussian kernel function in (10) transforms the data into an infinite dimension Hilbert space $\mathbb{F}$, so that a nonlinear mapping $\Phi(\cdot)$ yields

$$\kappa_\sigma(x-y) = \langle \Phi(x), \Phi(y) \rangle_{\mathbb{F}}\,, \tag{12}$$

where $\langle \cdot, \cdot \rangle_{\mathbb{F}}$ denotes the inner product in $\mathbb{F}$.

The MCC-LMS algorithm is obtained from the stochastic gradient maximization of the cost function (Chen et al., 2014)

$$\mathcal{F}_{\mathrm{MCC}}[\boldsymbol{w}(k)] \triangleq \mathbb{E}\left\{\exp\left[-\frac{e^2(k)}{2\sigma^2}\right]\right\}, \tag{13}$$

resulting in the following update equation (Singh and Principe, 2009)

$$\boldsymbol{w}(k+1) = \boldsymbol{w}(k) + \beta \exp\left[-\frac{e^2(k)}{2\sigma^2}\right] e(k)\boldsymbol{x}(k), \tag{14}$$

which degenerates into the LMS algorithm when $\sigma \to \infty$. The normalized version (MCC-NLMS) of (14) is given by (Liu et al., 2011)

$$\boldsymbol{w}(k+1) = \boldsymbol{w}(k) + \beta \exp\left[-\frac{e^2(k)}{2\sigma^2}\right] \frac{e(k)\boldsymbol{x}(k)}{\|\boldsymbol{x}(k)\|^2}. \tag{15}$$

The update equations (14) and (15) can be derived by solving the following deterministic optimization problem (Haddad et al., 2016):

$$\min_{\boldsymbol{w}(k+1)} \mathcal{F}_{\mathrm{MDP}}\left[\boldsymbol{w}(k+1)\right] \triangleq \|\boldsymbol{w}(k+1) - \boldsymbol{w}(k)\|^2 \tag{16}$$
$$\text{s.t. } e_p(k) = \left\{1 - \gamma \exp\left[-\frac{e^2(k)}{2\sigma^2}\right]\right\} e(k),$$

where $\gamma = \beta \|\boldsymbol{x}(k)\|^2$ and $\gamma = \beta$ for the MCC-LMS and MCC-NLMS, respectively. In this work, we focus on the non-normalized MCC-LMS algorithm.

5 MCC-RC-LMS Algorithm

The adoption of the MCC technique allows to strengthen the RC-LMS algorithm in the presence of impulsive noise. The resulting algorithm, named MCC-RC-LMS, combines the advantages of data reuse and robustness to impulsive interference. To derive a new algorithm that reuses the previous L adaptive coefficient vectors

[3]There are other options for the kernel, but the Gaussian kernel is preferred because of the resulting computational simplification (Erdogmus and Principe, 2002).

(Cho, 2009), we formulate the following optimization problem:

$$\min_{\boldsymbol{w}(k+1)} \mathcal{F}_{\mathrm{RC}}[\boldsymbol{w}(k+1)] \tag{17}$$

$$\text{s.t. } e_p(k) = \left[1 - \gamma e^{-\frac{e^2(k)}{2\sigma^2}}\right] \overline{e}(k),$$

whose solution is given in the theorem below.

Theorem 3. The solution of the constraint optimization problem (17) is given by:

$$\boldsymbol{w}(k+1) = \left(\frac{\rho-1}{\rho^L-1}\right) \sum_{l=0}^{L-1} \rho^l \boldsymbol{w}(k-l)$$

$$+\beta \exp^{-\frac{e^2(k)}{2\sigma^2}} \overline{e}(k) \boldsymbol{x}(k), \tag{18}$$

which is the update equation of the proposed MCC-RC-NLMS algorithm.
Demonstration: See Appendix.

6 Simulations

The algorithms used for comparison with the proposed MCC-RC-LMS algorithm were LMS, MCC-LMS and RC-LMS, with the following parameters: $L = 7$, $\rho = 0,9$ and $\sigma^2_{\mathrm{MCC}} = 2$, chosen for performance optimization. The metric used to evaluate the performance of the algorithms is the MSD, defined by

$$\mathrm{MSD}(k) \triangleq \|\boldsymbol{w}(k) - \boldsymbol{w}^\star\|^2. \tag{19}$$

The additive noise $\nu(k)$ was generated by

$$\nu(k) = (1 - \omega(k))\varphi(k) + \omega(k)\phi(k), \tag{20}$$

where $\omega(k)$ is a Bernoulli process with $\Pr[\omega(k) = 1] = 0.99$, $\varphi(k)$ and $\phi(k)$ consist of white Gaussian noise sequences with zero-mean and variances $\sigma^2_\phi = 10^{-1}$ and $\sigma^2_\varphi = 1$, respectively. Note that $\varphi(k)$ simulates a possible occurrence of impulsive noise. The mean results are from 500 independent Monte Carlo trials. The transfer function of the unknown system was the Model 1 of (TSG, 2004). Figures 1 and 2 show the MSD evolution as a function of β for white noise and impulsive noise, respectively, with $\sigma^2_\nu = 10^{-1}$. From these figures, it can be seen that the proposed RC-LMS and MCC-RC-LMS algorithms present better performance at steady state than the LMS and MCC-LMS algorithms. The MCC-RC-LMS presents better performance than the RC-LMS mainly in the presence of impulsive noise (Fig. 2).

Figure 3 shows the MSD evolutions of the algorithms in the presence of white noise, using the same parameters mentioned above, except for the probability of occurrence of impulsive noise which was arbitrated at 2%. The learning factors were selected so that all algorithms presented similar convergence rates. The resulting values were:

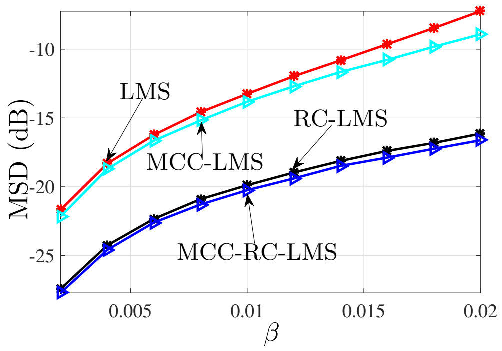

Figure 1: Steady-state MSD as a function of β in the presence of white noise.

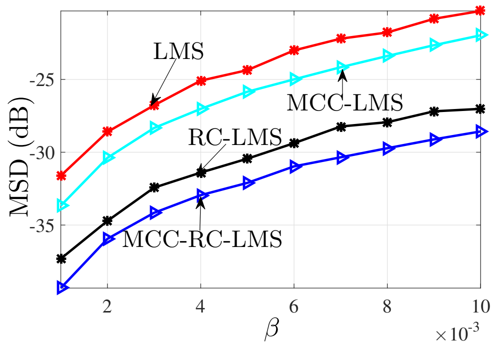

Figure 2: Steady-state MSD as a function of β in the presence of impulsive noise.

$\beta_{\mathrm{LMS}} = 10^{-2}$, $\beta_{\mathrm{MCC\text{-}LMS}} = 10^{-2}$, $\beta_{\mathrm{RC\text{-}LMS}} = 1,5 \cdot 10^{-2}$ and $\beta_{\mathrm{MCC\text{-}RC\text{-}LMS}} = 2 \cdot 10^{-2}$. It can be noticed that the proposed RC-LMS and MCC-RC-LMS algorithms present better steady-state behavior than the LMS and MCC-LMS. Again, the MCC-RC-LMS presented the best performance.

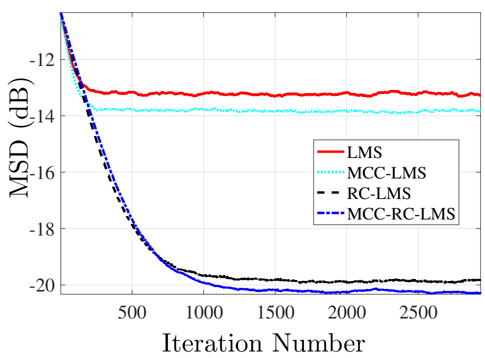

Figure 3: MSD evolution (in dB) on the iterations in the presence of white noise.

Figure 4 presents the MSD evolution of the algorithms in the presence of impulsive noise, with the same parameters used in the experiment with white noise. It may be noted again that the adoption of the maximum correntropy strategy associated to the reuse of coefficients resulted in better convergence properties.

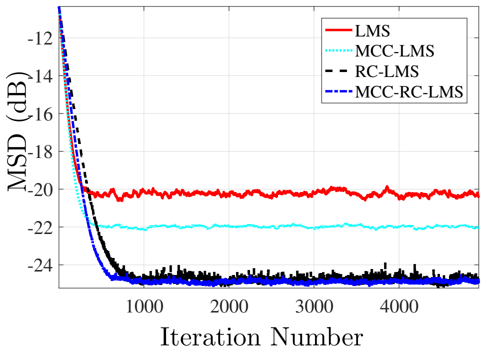

Figure 4: MSD evolution (in dB) on the iterations in the presence of impulsive noise.

7 Conclusions

In this article, new adaptive filters that generalize the LMS algorithm were proposed, being derived by the technique of Lagrange multipliers, able to solve problems of optimization with constraints. It has been demonstrated, through hypotheses commonly used in the literature, that the proposed RC-LMS and MCC-RC-LMS algorithms correspond to unbiased estimators of the optimal filter coefficients. The presented results confirm the superiority of the proposed algorithms, which can obtain MSD gains of more than 6 dB.

Acknowledgments

This work was partially supported by CNPq, FAPERJ, and CAPES.

Appendix

Theorem 1. The constrained optimization problem (4) can be converted into an equivalent problem without restriction by the Lagrange multipliers technique, giving rise to:

$$\mathcal{F}[\boldsymbol{w}(k+1)] = \sum_{l=0}^{L-1} \rho^l \|\boldsymbol{w}(k+1) - \boldsymbol{w}(k-l)\|^2$$

$$+\lambda(e_p(k) - (1-\beta||\boldsymbol{x}(k)||^2)\overline{e}(k)). \qquad (21)$$

Equating to zero the gradient of (21) for minimization effects, we obtain

$$\sum_{l=0}^{L-1} \rho^l \boldsymbol{w}(k+1) = \sum_{l=0}^{L-1} \rho^l \boldsymbol{w}(k-l) + \frac{\lambda}{2}\boldsymbol{x}(k)$$

or, equivalently,

$$\boldsymbol{w}(k+1) = \left(\frac{\rho-1}{\rho^L-1}\right) \sum_{l=0}^{L-1} \rho^l \boldsymbol{w}(k-l)$$

$$+\left(\frac{\rho-1}{\rho^L-1}\right)\frac{\lambda}{2}\boldsymbol{x}(k) \qquad (22)$$

Replacing (22) in the constraint of (4) results in

$$\left(\frac{\rho^L-1}{\rho-1}\right)\frac{\lambda}{2} = \beta\overline{e}(k), \qquad (23)$$

which, when inserted in (22), provides (8). □

Theorem 2. Substituting (6) in (8), we get

$$\boldsymbol{w}(k+1) = \left(\frac{\rho-1}{\rho^L-1}\right) \sum_{l=0}^{L-1} \rho^l \boldsymbol{w}(k-L) + \beta\boldsymbol{x}(k)\nu(k)$$

$$+\beta\left(\frac{\rho-1}{\rho^L-1}\right) \sum_{l=0}^{L-1} \rho^l \boldsymbol{x}(k)\boldsymbol{x}^T(k)\tilde{\boldsymbol{w}}(k-l) \qquad (24)$$

where $\tilde{\boldsymbol{w}}(k) \triangleq \boldsymbol{w}^\star - \boldsymbol{w}(k)$ is the deviation vector, which reflects the discrepancy between the vector of adaptive coefficients at instant k and their respective ideal values. Expressing (24) only in terms of the deviation vector, we find

$$\tilde{\boldsymbol{w}}(k+1) = \left(\frac{\rho-1}{\rho^L-1}\right) \sum_{l=0}^{L-1} \rho^l \left[\boldsymbol{I} - \boldsymbol{x}(k)\boldsymbol{x}^T(k)\right]\tilde{\boldsymbol{w}}(k-l)$$

$$-\beta\boldsymbol{x}(k)\nu(k), \qquad (25)$$

where $\boldsymbol{I}$ is the identity matrix of dimensions $N \times N$. Applying the statistical mean operator to Eq. (25) and using hypotheses $H2$ and $H3$, we obtain

$$\mathbb{E}\left[\tilde{\boldsymbol{w}}(k+1)\right] = \left(\frac{\rho-1}{\rho^L-1}\right) \sum_{l=0}^{L-1} \rho^l \left[\boldsymbol{I} - \boldsymbol{R}\right]\mathbb{E}\left[\tilde{\boldsymbol{w}}(k-l)\right]. \qquad (26)$$

Hypothesis $H1$ allows us to write

$$\lim_{k\to\infty} \mathbb{E}\left[\tilde{\boldsymbol{w}}(k)\right] = \mathbb{E}\left[\tilde{\boldsymbol{w}}(k-l)\right] \triangleq \mathbb{E}\left[\tilde{\boldsymbol{w}}_\infty\right], \qquad (27)$$

for $l \in \{0, 1, \ldots, L-1\}$. Substituting (27) in (26), we get

$$\left\{\left(\frac{\rho-1}{\rho^L-1}\right) \sum_{l=0}^{L-1} \rho^l \boldsymbol{R}\right\} \mathbb{E}\left[\tilde{\boldsymbol{w}}_\infty\right] = \boldsymbol{0}, \qquad (28)$$

which implies $\mathbb{E}\left[\tilde{\boldsymbol{w}}_\infty\right] = \boldsymbol{0}$. □

Theorem 3. Applying the Lagrange multipliers technique to problem (17) with $\gamma = \beta\|\boldsymbol{x}(k)\|^2$, we obtain

$$\mathcal{F}[\boldsymbol{w}(k+1)] = \sum_{l=0}^{L-1} \rho^l \|\boldsymbol{w}(k+1) - \boldsymbol{w}(k-l)\|^2$$

$$+\lambda(e_p(k) - (1-\beta\exp\left[-\frac{e^2(k)}{2\sigma^2}\right]||\boldsymbol{x}(k)||^2)\overline{e}(k)) \qquad (29)$$

By equating the derivative of (29) with respect to $\boldsymbol{w}(k+1)$ to zero, we obtain

$$\boldsymbol{w}(k+1) = \left(\frac{\rho-1}{\rho^L-1}\right) \sum_{l=0}^{L-1} \rho^l \boldsymbol{w}(k-l)$$

$$+\left(\frac{\rho-1}{\rho^L-1}\right)\frac{\lambda}{2}\boldsymbol{x}(k), \qquad (30)$$

which, after substituted in the constraint of (17), generates

$$\frac{\lambda}{2}\left(\frac{\rho^L - 1}{\rho - 1}\right) = \beta \exp\left[-\frac{e^2(k)}{2\sigma^2}\right]\bar{e}(k). \tag{31}$$

Replacement of (31) in (30) gives rise to (18). □

References

Chen, B., Xing, L., Liang, J., Zheng, N. and Principe, J. C. (2014). Steady-state mean-square error analysis for adaptive filtering under the maximum correntropy criterion, *IEEE Signal Processing Letters* **21**(7): 880–884.

Cho, H. e. a. (2009). Affine projection algorithm with coefficient vector reusing, *ECTI-CON 2009. 6th International Conference on. IEEE, 2009* pp. 1148–1150.

Cho, H., Lee, C. W. and Kim, S. W. (2009). Derivation of a new normalized least mean squares algorithm with modified minimization criterion, *Signal Processing* **89**(4): 692 – 695.

Das, R. L. and Narwaria, M. (2017). Lorentzian based adaptive filters for impulsive noise environments, *IEEE Transactions on Circuits and Systems I: Regular Papers* **64**(6): 1529–1539.

Diniz, P. S. (1997). *Adaptive filtering*, Springer.

Erdogmus, D. and Principe, J. C. (2002). An error-entropy minimization algorithm for supervised training of nonlinear adaptive systems, *IEEE Transactions on Signal Processing* **50**(7): 1780–1786.

Haddad, D. B., Petraglia, M. R. and Petraglia, A. (2016). A unified approach for sparsity-aware and maximum correntropy adaptive filters, *2016 24th European Signal Processing Conference (EUSIPCO)*, pp. 170–174.

Hassibi, B., Sayed, A. H. and Kailath, T. (1996). H-∞ optimality of the LMS algorithm, *IEEE Transactions on Signal Processing* **44**(2): 267–280.

Haykin, S. and Widrow, B. (2003). *Least-mean-square adaptive filters*, Vol. 31, John Wiley & Sons.

Ikuma, T. and Beex, A. A. (2008). Improved mean-square error estimate for the LMS transversal equalizer with narrowband interference, *IEEE Transactions on Signal Processing* **56**(10): 5273–5277.

Ikuma, T., Beex, A. A. and Zeidler, J. R. (2007). Non-Wiener weight behavior of LMS transversal equalizers, *2007 IEEE International Conference on Acoustics, Speech and Signal Processing - ICASSP '07*, Vol. 3, pp. III–1297–III–1300.

Ikuma, T., Beex, A. A. and Zeidler, J. R. (2008). Non-Wiener mean weight behavior of LMS transversal equalizers with sinusoidal interference, *IEEE Transactions on Signal Processing* **56**(9): 4521–4525.

Kim, S.-E., Lee, J.-W. and Song, W.-J. (2011). Steady-state analysis of the NLMS algorithm with reusing coefficient vector and a method for improving its performance, *Acoustics, Speech and Signal Processing (ICASSP), 2011 IEEE International Conference on*, IEEE, pp. 4120–4123.

Kumar, M. N., PrasannaVani, V. and Saravanan, S. (2017). Evaluation of LMS and Volterra adaptive filters for AEC, *2017 IEEE International Conference on Signal Processing, Informatics, Communication and Energy Systems (SPICES)*, pp. 1–6.

Liu, W., Pokharel, P. P. and Principe, J. C. (2007). Correntropy: Properties and applications in non-gaussian signal processing, *IEEE Transactions on Signal Processing* **55**(11): 5286–5298.

Liu, W., Principe, J. C. and Haykin, S. (2011). *Kernel adaptive filtering: a comprehensive introduction*, Vol. 57, John Wiley & Sons.

Liu, X., Chen, B., Zhao, H., Qin, J. and Cao, J. (2017). Maximum correntropy Kalman filter with state constraints, *IEEE Access* **5**: 25846–25853.

Petraglia, M. R., Marques, E. L. and Haddad, D. B. (2016). Low-complexity affine projection subband algorithm for robust adaptive filtering in impulsive noise, *2016 IEEE Sensor Array and Multichannel Signal Processing Workshop (SAM)*, pp. 1–5.

Quirk, K. J., Milstein, L. B. and Zeidler, J. R. (2000). A performance bound for the LMS estimator, *IEEE Transactions on Information Theory* **46**(3): 1150–1158.

Quirk, K. J., Zeidler, J. R. and Milstein, L. B. (1998). Bounding the performance of the LMS estimator for cases where performance exceeds that of the finite Wiener filter, *Acoustics, Speech and Signal Processing, 1998. Proceedings of the 1998 IEEE International Conference on*, Vol. 3, pp. 1417–1420 vol.3.

Reuter, M. and Zeidler, J. R. (1999). Nonlinear effects in LMS adaptive equalizers, *IEEE Transactions on Signal Processing* **47**(6): 1570–1579.

Sayed, A. H. (2011). *Adaptive filters*, John Wiley & Sons.

Simon, H. (2002). Adaptive filter theory, *Prentice Hall* **2**.

Singh, A. and Principe, J. C. (2009). Using correntropy as a cost function in linear adaptive filters, *2009 International Joint Conference on Neural Networks*, pp. 2950–2955.

TSG, I. (2004). *5, Digital Network Echo Cancellers (recommendation*, tech. rep., Tech. Rep. G. 168, ITU-T.

Venkatesan, C., Karthigaikumar, P., Paul, A., Satheeskumaran, S. and Kumar, R. (2018). ECG signal preprocessing and SVM classifier-based abnormality detection in remote healthcare applications, *IEEE Access* **6**: 9767–9773.