A New Proportionate Adaptive Filtering Algorithm with Coefficient Reuse and Robustness Against Impulsive Noise

Rodrigo M. S. Pimenta*, Leonardo C. Resende†, Newton N. Siqueira‡, Diego B. Haddad§, Mariane R. Petraglia¶
*Telecom. Eng., CEFET-RJ - Maracanã
†Engineering Coordination, Instituto Federal do Rio de Janeiro
‡Telecom. Eng., CEFET-RJ - Nova Iguaçu
§Computer Eng., CEFET-RJ - Petrópolis
¶Electrical Eng. Program, COPPE, UFRJ
Rio de Janeiro, Brazil
Email: {rodrigo.pimenta,newton.siqueira}@cefet-rj.br, leonardo.resende@ifrj.edu.br, {diego,mariane}@pads.ufrj.br

***Abstract*—An adaptive algorithm should ideally present high convergence rate, good steady-state performance, and robustness against impulsive noise. Few algorithms can simultaneously meet these requirements. This paper proposes a local and deterministic optimization problem whose solution gives rise to an adaptive algorithm that presents a higher convergence rate in the identification of sparse systems due to the use of the proportionate adaptation technique. In addition, a correntropy-based cost function is employed in order to enhance its robustness against non-Gaussian noise. Finally, the adoption of coefficient reuse approach results in a good system identification performance in steady-state conditions, especially in low SNR scenarios.**

***Index Terms*—Adaptive Filtering, Sparse systems, Proportionate Adaptation, Coefficients Reuse, Maximum Correntropy Criterion.**

I. Introduction

Digital signal processing techniques allowed the advent of many implementations of adaptive filtering (AF) algorithms capable of addressing complex applications with challenging requirements, such as echo acoustic cancellation (AEC), equalization, prediction, interference reduction, antenna arrays spatial-temporal processing, spectrum analysis and system identification [1].

In the case of (but not restricted to) AEC tasks, AF techniques may be employed to electrically emulate the acoustic echo coupling between the loudspeaker and the microphone signals [2]. The performance or the convergence rate of such identification procedure, crucial in telecommunications systems [3], can be impaired when *i)* the acoustic impulse response is long [4], *ii)* in presence of impulsive noise [5] or *iii)* the signal-to-noise ratio (SNR) is low [6]. Before describing our integrated solution that addresses all these issues, common approaches to mitigate their impacts will be presented succinctly.

Fortunately, the problem *i)* can be mitigated by the use of sparsity-aware identification schemes (such as the proportionate approach [7]), which take into account the fact that frequently most elements of the transfer function to be identified are close to zero [8]. The insertion of such prior knowledge can increase the convergence rate of identification tasks, as compared to the more naive sparsity-agnostic approaches [7].

Regarding to the problem *ii)*, it is noteworthy that near-end speech in AEC with double-talk is a common source of impulsive noise [9], as well as sudden atmospheric phenomena in telecommunication systems [5]. Although the mean square error (MSE) minimization has been widely used as a statistical measure for the development of adaptive filtering algorithms [10], this cost function is prone to instability in the presence of impulsive noise [11]. This fact claims robust cost functions, such as the maximum correntropy criterion (MCC) [12].

The problem *iii)* can be addressed by minimizing the weighted summation of squared Euclidean norms of the difference between the updated coefficient vector and previous ones [13]. This reusing coefficient (RC) strategy presents reduced steady-state mean-square deviation with convergence rate similar to the normalized least mean squares (NLMS) algorithm in the case of high energy measurement noise [6].

Since the above solutions are derived by different approaches[1], a formal derivation of an algorithm that incorporates their capabilities is not straightforward. This paper achieves this goal through a unified derivation framework.

Section II introduces the fundamentals of classic adaptive filtering techniques, whose performance can be improved by sparsity-aware, MCC and RC schemes.The proposed algorithm is presented in Section III, and its energy conservation relation is derived in Section IV. The performance of the proposed solution is evaluated in Section V. Finally, Section VI presents the concluding remarks.

II. Adaptive Filtering Fundamentals

Despite their simplicity (from both conceptual and computational viewpoints), AF techniques consist of powerful methods

[1]For example, the proportionate approach uses the Lagrange multiplier method, while the MCC based derivations usually employ the stochastic gradient method.

ISBN 978-90-827970-1-5 © EURASIP 2018

for addressing crucial tasks in actual digital signal processing systems. In their most popular form[2], such algorithms should recursively obtain at the kth iteration a vector $\boldsymbol{w}(k) \in \mathbb{R}^N$ which emulates the ideal one, $\boldsymbol{w}^\star$, which is unknown to the algorithm designer. In supervised settings, one has access to a reference (or desired) signal $d(k)$ given by

$$d(k) \triangleq \boldsymbol{x}^T(k)\boldsymbol{w}^\star + \nu(k), \tag{1}$$

where $\nu(k)$ is a measurement noise (which can also incorporate modeling errors) and

$$\boldsymbol{x}(k) \triangleq \begin{bmatrix} x(k) & x(k-1) & \dots & x(k-N+1) \end{bmatrix}^T. \tag{2}$$

The least mean squares (LMS) updating equation [14] - the most popular AF algorithm - may be derived by means of the stochastic gradient technique, that is,

$$\boldsymbol{w}(k+1) = \boldsymbol{w}(k) - \beta \nabla_{\boldsymbol{w}(k)} \mathcal{F}_{\text{LMS}}[\boldsymbol{w}(k)], \tag{3}$$

where β is a step-size parameter (whose value should be chosen in order to address the convergence rate versus steady-state performance trade-off) and $\mathcal{F}_{\text{LMS}}[\boldsymbol{w}(k)]$ is the cost function, namely a stochastic approximation of the MSE,

$$\mathcal{F}_{\text{LMS}}[\boldsymbol{w}(k)] \triangleq \frac{1}{2} e^2(k), \tag{4}$$

where

$$e(k) \triangleq d(k) - \overbrace{\boldsymbol{w}^T(k)\boldsymbol{x}(k)}^{\triangleq y(k)}. \tag{5}$$

After such definitions, it is straightforward to derive the LMS update equation:

$$\boldsymbol{w}(k+1) = \boldsymbol{w}(k) + \beta \boldsymbol{x}(k) e(k), \tag{6}$$

which presents some undesirable features, such as an upper bound for β that avoids divergence (a catastrophic phenomenon) which is highly dependent of statistical properties of the input signal [14]. Normalized schemes (e.g., the NLMS algorithm) do not share several of such critical issues.

The δ-NLMS update equation can be understood as the solver of the following local and deterministic optimization problem [15]:

$$\min_{\boldsymbol{w}(k+1)} \mathcal{F}_{\text{MDP}}[\boldsymbol{w}(k+1)] \text{ s.t. } e_p(k) = \left(1 - \beta \frac{\|\boldsymbol{x}(k)\|^2}{\|\boldsymbol{x}(k)\|^2 + \delta}\right) e(k), \tag{7}$$

where δ is a regularization parameter (hereinafter supposed to be zero), $e_p(k)$ is the posterior error

$$e_p(k) \triangleq d(k) - \boldsymbol{w}^T(k+1)\boldsymbol{x}(k) \tag{8}$$

and $\mathcal{F}_{\text{MDP}}[\boldsymbol{w}(k+1)]$ is a cost function based on the conservative minimum disturbance principle (MDP):

$$\mathcal{F}_{\text{MDP}}[\boldsymbol{w}(k+1)] \triangleq \|\boldsymbol{w}(k+1) - \boldsymbol{w}(k)\|^2. \tag{9}$$

Note that $\mathcal{F}_{\text{MDP}}[\boldsymbol{w}(k+1)]$ penalizes solutions distant from the previous coefficient vector $\boldsymbol{w}(k)$ (which is coherent to the MDP principle), while presenting a controlled posterior error (which can be zeroed under the choice $\beta = 1$). The resulting NLMS algorithm can be written as

$$\boldsymbol{w}(k+1) = \boldsymbol{w}(k) + \beta \frac{\boldsymbol{x}(k) e(k)}{\|\boldsymbol{x}(k)\|^2}. \tag{10}$$

A. Enhancing the Convergence Rate

Taking advantage of a prior knowledge of high sparsity in systems may enhance the convergence rate of adaptation schemes. One of such strategies is the family of proportionate algorithms, in which the Proportionate NLMS (PNLMS) [7] pioneered. Proportionate algorithms distribute the updating energy proportionally to the magnitude of the adaptive coefficients. Proportionate algorithms perform a natural gradient procedure in a warped coefficient space defined by a specific coefficient metric [16]. The proportional steps (specific for each adaptive coefficient) are implemented by a diagonal matrix $\boldsymbol{\Lambda}(k)$ and the update equation of the algorithm is given by

$$\boldsymbol{w}(k+1) = \boldsymbol{w}(k) + \beta \frac{\boldsymbol{\Lambda}(k)\boldsymbol{x}(k) e(k)}{\|\boldsymbol{x}(k)\|^2_{\boldsymbol{\Lambda}(k)}}, \tag{11}$$

where $\|\boldsymbol{x}\|^2_{\boldsymbol{A}} \triangleq \boldsymbol{x}^T \boldsymbol{A} \boldsymbol{x}$. The above update equation can be understood as the solver of the following constraint optimization problem:

$$\min_{\boldsymbol{w}(k+1)} \|\boldsymbol{w}(k+1) - \boldsymbol{w}(k)\|^2_{\boldsymbol{\Lambda}^{-1}(k)} \tag{12}$$
$$\text{s.t. } e_p(k) = (1-\beta)\, e(k).$$

Different choices of diagonal elements of $\boldsymbol{\Lambda}(k)$ give rise to different proportional AFs, such as the MPNLMS [17], the IPNLMS [18] and the IMPNLMS [19].

B. Robustness against Impulsive Noise

It is an established fact that the MSE criterion-based adaptive filters may not perform well under non-Gaussian noise. A more robust alternative is the correntropy, which is a local similarity measure between random variables X and Y given by [20]

$$V(X,Y) = \iint_{x,y} \kappa_\sigma(x-y) f_{XY}(x,y)\, dx dy, \tag{13}$$

where $\kappa_\sigma(x-y)$ is the Gaussian kernel[3]

$$\kappa_\sigma(x-y) \triangleq \frac{1}{\sqrt{2\pi}\sigma} \exp\left[-\frac{(x-y)^2}{2\sigma^2}\right] \tag{14}$$

and σ is the kernel size that induces a trade-off between steady-state performance and convergence rate [22]. The kernel function (13) transforms data to an infinite dimensional reproducing kernel Hilbert space $\mathbb{F}$, so that the following nonlinear mapping Φ holds:

$$\kappa_\sigma(x-y) = \langle \Phi(x), \Phi(y) \rangle_{\mathbb{F}}, \tag{15}$$

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

[2]This paper focus on system identification tasks.

[3]There are other possible choices for the kernel function, but the Gaussian kernel is the preferred one due to the resulting computational simplification in the algorithm design [21].

Under the MCC criterion, a stochastic gradient ascent method with cost function [23]

$$\mathcal{F}_{\text{MCC}}[\boldsymbol{w}(k)] \triangleq \mathbb{E}\left\{\exp\left[-\frac{e^2(k)}{2\sigma^2}\right]\right\} \tag{16}$$

can be employed to derive the MCC-LMS AF algorithm [12]:

$$\boldsymbol{w}(k+1) = \boldsymbol{w}(k) + \beta\exp\left[-\frac{e^2(k)}{2\sigma^2}\right] e(k)\boldsymbol{x}(k), \tag{17}$$

which reduces to the LMS algorithm as $\sigma \to \infty$. The normalized version (MCC-NLMS) of (17) is given by [22]

$$\boldsymbol{w}(k+1) = \boldsymbol{w}(k) + \beta\exp\left[-\frac{e^2(k)}{2\sigma^2}\right] \frac{e(k)\boldsymbol{x}(k)}{\|\boldsymbol{x}(k)\|^2}. \tag{18}$$

Both update equations (17) and (18) can be derived in an unified way from the following deterministic optimization problem [24]:

$$\min_{\boldsymbol{w}(k+1)} \mathcal{F}_{\text{MDP}}[\boldsymbol{w}(k+1)] \text{ s.t. } e_p(k) = \left\{1 - \gamma\exp\left[-\frac{e^2(k)}{2\sigma^2}\right]\right\} e(k), \tag{19}$$

where $\gamma = \beta\|\boldsymbol{x}(k)\|^2$ for obtaining (17) and $\gamma = \beta$ for obtaining (18). This paper focus on normalized algorithms, so that hereinafter the choice $\gamma = \beta$ is assumed.

C. Enhancing the Steady-State Performance in Low SNR Regimes

The popular Affine Projection Algorithm (APA) [25] increases the convergence rate through the reuse of input data. Such algorithm tends to present poor performance in steady state as a disadvantage. These characteristics can be considered dual to those presented by the RC algorithm [13], which implements a reuse of the last L vectors of adaptive coefficients. The RC family of algorithms combine better performance at steady-state with a loss in the convergence rate [26]. In order to take advantage of such complementary features, [27] proposes and analyzes the joint use of both reuse strategies.

Defining the weighted error $e'(k)$ as

$$e'(k) \triangleq d(k) - \frac{\rho - 1}{\rho^L - 1}\sum_{l=0}^{L-1} \rho^l \boldsymbol{w}^T(k-l)\boldsymbol{x}(k), \tag{20}$$

the RC-NLMS algorithm can be described as a solver of the following optimization problem:

$$\min_{\boldsymbol{w}(k+1)} \mathcal{F}_{\text{RC}}[\boldsymbol{w}(k+1)] \text{ s.t. } e_p(k) = (1-\beta)e'(k), \tag{21}$$

where

$$\mathcal{F}_{\text{RC}}[\boldsymbol{w}(k+1)] \triangleq \sum_{l=0}^{L-1} \rho^l \|\boldsymbol{w}(k+1) - \boldsymbol{w}(k-l)\|^2 \tag{22}$$

and $\rho \in (0, 1]$ is a parameter at the discretion of the designer that controls the influence of older adaptive coefficients in the update mechanism. The solution of (21) gives rise to the following update equation:

$$\boldsymbol{w}(k+1) = \frac{\rho - 1}{\rho^L - 1}\sum_{l=0}^{L-1} \rho^l \boldsymbol{w}(k-l) + \beta\frac{\boldsymbol{x}(k)e'(k)}{\|\boldsymbol{x}(k)\|^2}, \tag{23}$$

so that $\boldsymbol{w}(k+1)$ depends on a weighted sum of the L vectors $\boldsymbol{w}(k-l)$ (with $l \in \{0, 1, \ldots, L-1\}$), which softens the filter oscillations. This feature is responsible for the performance improvement in steady-state regime, especially in configurations with low SNR [13].

III. Proposed RC-MCC-PNLMS Algorithm

The previous discussion highlighted the connection between features of specific deterministic optimization problems and the overcoming advantages of algorithms obtained as solutions of such problems. This relationship motivates a problem-building procedure, whose solution (adaptive algorithm) presents the desirable properties. In order to address the issues described in Section II, we propose the following local optimization problem:

$$\min_{\boldsymbol{w}(k+1)} \sum_{l=0}^{L-1} \rho^l \|\boldsymbol{w}(k+1) - \boldsymbol{w}(k)\|^2_{\boldsymbol{\Lambda}^{-1}(k)} \tag{24}$$

$$\text{s.t. } e_p(k) = \left\{1 - \beta\exp\left[-\frac{e^2(k)}{2\sigma^2}\right]\right\} e'(k).$$

Theorem 1. The vector $\boldsymbol{w}(k+1)$ that exactly solves (24) can be expressed as

$$\boldsymbol{w}(k+1) = \theta(\rho)\sum_{l=0}^{L-1} \rho^l \boldsymbol{w}(k-l) + \frac{\beta\exp\left[-\frac{e^2(k)}{2\sigma^2}\right] e'(k)\boldsymbol{\Lambda}(k)\boldsymbol{x}(k)}{\|\boldsymbol{x}(k)\|^2_{\boldsymbol{\Lambda}(k)}}, \tag{25}$$

where $\theta(\rho) \triangleq \frac{1-\rho}{1-\rho^L}$.

Proof: Using the Lagrange multiplier technique, the optimization problem (24) can be converted into the minimization of the following unconstrained cost function:

$$\begin{aligned}\mathcal{F}_{\text{U}}[\boldsymbol{w}(k+1)] = &\sum_{l=0}^{L-1} \|\boldsymbol{w}(k+1) - \boldsymbol{w}(k)\|^2_{\boldsymbol{\Lambda}^{-1}(k)} \\ &+ \lambda\left\{e_p(k) - \left[1 - \beta\exp\left(-\frac{e^2(k)}{2\sigma^2}\right)\right] e'(k)\right\}.\end{aligned} \tag{26}$$

Differentiating (26) w.r.t. $\boldsymbol{w}(k+1)$ and zeroing the resulting expression, one finds

$$\boldsymbol{w}(k+1) = \theta(\rho)\sum_{l=0}^{L-1} \rho^l \boldsymbol{w}(k-l) + \frac{\lambda\theta(\rho)}{2}\boldsymbol{\Lambda}(k)\boldsymbol{x}(k), \tag{27}$$

which gives the format of the update equation. By replacing (27) in the linear constraint of (24), one concludes that

$$\frac{\lambda}{2}\theta(\rho) = \frac{\beta\exp\left[-\frac{e^2(k)}{\sigma^2}\right]}{\|\boldsymbol{x}(k)\|^2_{\boldsymbol{\Lambda}(k)}}. \tag{28}$$

At last, the application of (28) in (27) establishes the identity (25). □

The update equation (25) consists of the proposed solution, which simultaneously incorporates features of (12), (19), and (21), so that the resulting algorithm presents high convergence rate in the identification of sparse responses, robustness against impulsive noise, and good steady-state performance in low SNR contexts. We denote the algorithm of Eq. (25) as RC-MCC-PNLMS (Reusing Coefficient Maximum Correntropy Proportionate NLMS Algorithm).

ISBN 978-90-827970-1-5 © EURASIP 2018

IV. Energy Conservation Relationship

A popular approach for the performance prediction of adaptive filtering algorithms relies on an energy conservation identity [28]. Let us define the following error-related quantities:

$$e_{\mathrm{p}}^{\boldsymbol{\Sigma}}(k) \triangleq \boldsymbol{x}^T(k)\boldsymbol{\Sigma}\tilde{\boldsymbol{w}}(k+1), \tag{29}$$

$$e_{\mathrm{a},l}^{\boldsymbol{\Sigma}}(k) \triangleq \boldsymbol{x}^T(k)\boldsymbol{\Sigma}\tilde{\boldsymbol{w}}(k-l), \;\text{ for } l \in \{0,1,\ldots,L-1\}, \tag{30}$$

where $\tilde{\boldsymbol{w}}(k) \triangleq \tilde{\boldsymbol{w}}^\star - \boldsymbol{w}(k)$ is the deviation vector and $\boldsymbol{\Sigma} \in \mathbb{R}^{N\times N}$ is an arbitrary symmetric matrix. Combining the approaches of [6], [24], [29], [30] one obtains from Eq. (25) the following weighted variance relation of the RC-MCC-PNLMS algorithm:

$$\|\tilde{\boldsymbol{w}}(k+1)\|^2 + \frac{2\theta(\rho)}{\|\boldsymbol{x}(k)\|^2_{\boldsymbol{\Sigma}\boldsymbol{\Lambda}_k}} \sum_{l=0}^{L-1} \rho^l e_{\mathrm{a},l}^{\boldsymbol{\Sigma}}(k) e_{\mathrm{p}}^{\boldsymbol{\Lambda}_k}(k)$$
$$+\frac{\theta^2(\rho)\|\boldsymbol{x}(k)\|^2_{\boldsymbol{\Lambda}_k^2}}{\|\boldsymbol{x}(k)\|^4_{\boldsymbol{\Sigma}\boldsymbol{\Lambda}_k}} \sum_{l_1=0}^{L-1}\sum_{l_2=0}^{L-1} \rho^{l_1}\rho^{l_2} e_{\mathrm{a},l_1}^{\boldsymbol{\Sigma}}(k) e_{\mathrm{a},l_2}^{\boldsymbol{\Sigma}}(k) \tag{31}$$
$$= \theta^2(\rho) \sum_{l_1=0}^{L-1}\sum_{l_2=0}^{L-1} \rho^{l_1}\rho^{l_2} \tilde{\boldsymbol{w}}^T(k-l_1)\tilde{\boldsymbol{w}}(k-l_2)$$
$$+2\frac{\theta(\rho)}{\|\boldsymbol{x}(k)\|^2_{\boldsymbol{\Sigma}\boldsymbol{\Lambda}_k}} \sum_{l=0}^{L-1} \rho^l e_{\mathrm{a},l}^{\boldsymbol{\Lambda}_k}(k) e_{\mathrm{p}}^{\boldsymbol{\Sigma}}(k) + \frac{\left[e_{\mathrm{p}}^{\boldsymbol{\Sigma}}(k)\right]^2 \|\boldsymbol{x}(k)\|^2_{\boldsymbol{\Lambda}_k}}{\|\boldsymbol{x}(k)\|^2_{\boldsymbol{\Sigma}\boldsymbol{\Lambda}_k}},$$

which is an exact identity that can be used to perform a stochastic transient analysis of the proposed algorithm [29]. Note that in the case of steady-state analysis, the choice $\boldsymbol{\Sigma} = \boldsymbol{\Lambda}_k$ simplifies Eq. (31). Due to the lack of space, a complete theoretical analysis of the RC-MCC-PNLMS algorithm is not presented in this paper.

V. Simulations

The algorithms used for comparison with the proposed RC-MCC-PNLMS[4] algorithm were the NLMS, MCC-NLMS and MCC-RC-NLMS with parameters $L = 1$ and $L = 7$, $\rho = 0.9$, $\epsilon = 10^{-3}$, $\lambda = 0.96$ (see [19]), and $\sigma^2_{\mathrm{MCC}} = 2$. The measurement noise $\nu(k)$ is given by

$$\nu(k) = (1-\omega(k))\varphi(k) + \omega(k)\phi(k), \tag{32}$$

where $\omega(k)$ is a Bernoulli process with $\Pr[\omega(k) = 1] = 0.99$, and $\varphi(k)$ and $\phi(k)$ are white Gaussian noises with zero means and variances $\sigma^2_\varphi = 1$ and $\sigma^2_\phi = 10^{-1}$, respectively. Note that $\phi(k)$ emulates an eventual occurrence of impulsive noise.

All results were obtained by averaging 1000 independent Monte Carlo runs. The ideal transfer function contained 64 coefficients, of which the first three consisted of ones and the others of zeros. Figure 1 displays the steady-state mean-square deviation (MSD $\triangleq \|\boldsymbol{w}^\star - \boldsymbol{w}(k)\|^2$) as a function of β. These results demonstrate that the proposed algorithm presents a steady-state performance superior to most of the competing algorithms.

Figure 2 shows the evolution of the MSD (equalized for all algorithms in steady state) with the same parameters of Fig. 1 and with the transfer function of Model 1 of [31]. It can be observed that the proposed algorithm presents faster convergence than the other algorithms.

[4]The chosen proportionate algorithm is the IMPNLMS [19].

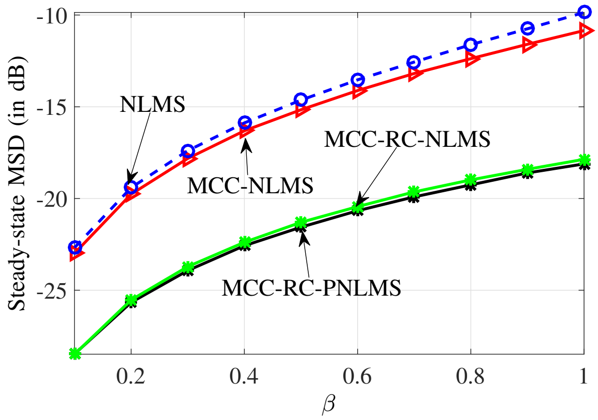

Fig. 1. Steady-state MSD for different β values.

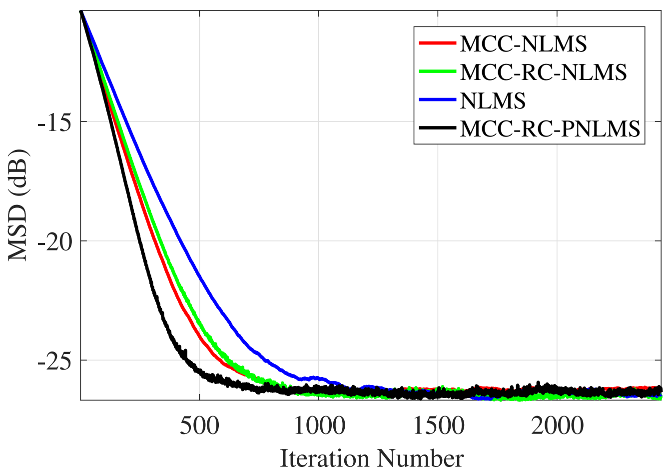

Fig. 2. MSD evolution for NLMS, MCC-NLMS, MCC-RC-NLMS and RC-MCC-PNLMS, at equal steady-state performance.

To evaluate the tracking ability of the algorithms, an experiment with an abrupt change of the ideal transfer function was performed, so that its coefficients are $\boldsymbol{w}_1^\star(k)$ in the first half of the iterations and $\boldsymbol{w}_2^\star(k)$ in the second half, given by

$$w_1^\star(k) = \begin{cases} 1, & \text{for } k=0 \\ -0.8, & \text{for } k=1 \\ 0.3, & \text{for } k=2 \\ 0, & \text{otherwise} \end{cases}, \quad w_2^\star(k) = \begin{cases} 1, & \text{for } k=0 \\ 1, & \text{for } k=1 \\ 1, & \text{for } k=2 \\ 0, & \text{otherwise} \end{cases}.$$

Figure 3 shows the MSD evolutions, from which it can be observed that the proposed algorithm outperforms the others in relation to the tracking performance.

VI. Conclusions

Owing to their widespread use in digital signal processing tasks, the design of state-of-the-art AF algorithms still attracts the attention of the scientific community. This paper advances a derivation methodology that smoothly incorporates to the NLMS algorithm properties that enhance its learning abilities

ISBN 978-90-827970-1-5 © EURASIP 2018

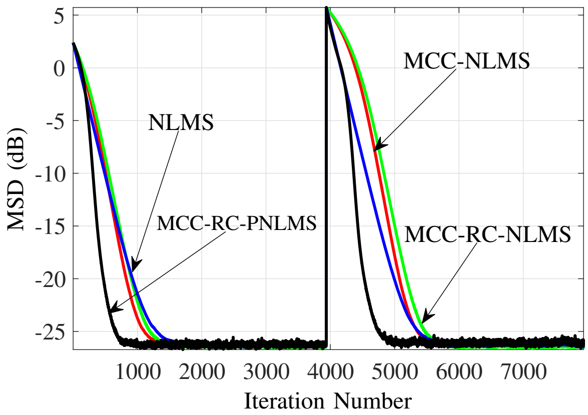

Fig. 3. MSD evolution for NLMS, MCC-NLMS, MCC-RC-NLMS and RC-MCC-PNLMS, at equal steady-state performance for two transfer functions.

in adversarial environments or in the case of sparse transfer functions. More specifically, in such scenarios the proposed solution attained faster convergence rates than the traditional algorithms, when the steady-state performances are matched.

Acknowledgments

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

References

[1] H. Simon, "Adaptive filter theory," *Prentice Hall*, vol. 2, 2002.

[2] Z. Zheng, Z. Liu, H. Zhao, Y. Yu, and L. Lu, "Robust set-membership normalized subband adaptive filtering algorithms and their application to acoustic echo cancellation," *IEEE Transactions on Circuits and Systems I: Regular Papers*, vol. 64, pp. 2098–2111, Aug 2017.

[3] I. Albu, C. Anghel, and C. Paleologu, "Adaptive filtering in acoustic echo cancellation systems - a practical overview," in *2017 9th International Conference on Electronics, Computers and Artificial Intelligence (ECAI)*, pp. 1–6, June 2017.

[4] A. Tedjani and A. Benallal, "Performance study of three different sparse adaptive filtering algorithms for echo cancellation in long acoustic impulse responses," in *2017 5th International Conference on Electrical Engineering - Boumerdes (ICEE-B)*, pp. 1–7, Oct 2017.

[5] R. L. Das and M. Narwaria, "Lorentzian based adaptive filters for impulsive noise environments," *IEEE Transactions on Circuits and Systems I: Regular Papers*, vol. 64, pp. 1529–1539, June 2017.

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

[7] D. L. Duttweiler, "Proportionate normalized least-mean-squares adaptation in echo cancelers," *IEEE Transactions on Speech and Audio Processing*, vol. 8, pp. 508–518, Sep 2000.

[8] M. Yamagishi, M. Yukawa, and I. Yamada, "Automatic shrinkage tuning based on a system-mismatch estimate for sparsity-aware adaptive filtering," in *2017 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pp. 4800–4804, March 2017.

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

[10] M. Xiang, S. C. Douglas, and D. P. Mandic, "The quaternion least mean magnitude phase adaptive filtering algorithm," in *2017 22nd International Conference on Digital Signal Processing (DSP)*, pp. 1–5, Aug 2017.

[11] J. Hur, I. Song, and P. Park, "A variable step-size normalized subband adaptive filter with a step-size scaler against impulsive measurement noise," *IEEE Transactions on Circuits and Systems II: Express Briefs*, vol. 64, pp. 842–846, July 2017.

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

[13] H. Cho, C. W. Lee, and S. W. Kim, "Derivation of a new normalized least mean squares algorithm with modified minimization criterion," *Signal Processing*, vol. 89, no. 4, pp. 692 – 695, 2009.

[14] S. Haykin and B. Widrow, *Least-mean-square adaptive filters*, vol. 31. John Wiley & Sons, 2003.

[15] A. H. Sayed, *Adaptive filters*. John Wiley & Sons, 1 ed., 2011.

[16] S. L. Gay and S. C. Douglas, "Normalized natural gradient adaptive filtering for sparse and non-sparse systems," in *2002 IEEE International Conference on Acoustics, Speech, and Signal Processing*, vol. 2, pp. II–1405–II–1408, May 2002.

[17] H. Deng and M. Doroslovacki, "Improving convergence of the pnlms algorithm for sparse impulse response identification," *IEEE Signal Processing Letters*, vol. 12, no. 3, pp. 181–184, 2005.

[18] J. Benesty and S. L. Gay, "An improved pnlms algorithm," in *Acoustics, Speech, and Signal Processing (ICASSP), 2002 IEEE International Conference on*, vol. 2, pp. II–1881, IEEE, 2002.

[19] M. Fukumoto, S. Saiki, *et al.*, "An improved mu-law proportionate nlms algorithm," in *Acoustics, Speech and Signal Processing, 2008. ICASSP 2008. IEEE International Conference on*, pp. 3797–3800, IEEE, 2008.

[20] W. Liu, P. P. Pokharel, and J. C. Principe, "Correntropy: Properties and applications in non-gaussian signal processing," *IEEE Transactions on Signal Processing*, vol. 55, pp. 5286–5298, Nov 2007.

[21] D. Erdogmus and J. C. Principe, "An error-entropy minimization algorithm for supervised training of nonlinear adaptive systems," *IEEE Transactions on Signal Processing*, vol. 50, pp. 1780–1786, Jul 2002.

[22] W. Liu, J. C. Principe, and S. Haykin, *Kernel adaptive filtering: a comprehensive introduction*, vol. 57. John Wiley & Sons, 2011.

[23] B. Chen, L. Xing, J. Liang, N. Zheng, and J. C. Principe, "Steady-state mean-square error analysis for adaptive filtering under the maximum correntropy criterion," *IEEE Signal Processing Letters*, vol. 21, pp. 880–884, July 2014.

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

[25] K. Ozeki and T. Umeda, "An adaptive filtering algorithm using an orthogonal projection to an affine subspace and its properties," *Electronics and Communications in Japan (Part I: Communications)*, vol. 67, no. 5, pp. 19–27, 1984.

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

[27] S.-E. Kim, J.-W. Lee, and W.-J. Song, "A noise-resilient affine projection algorithm and its convergence analysis," *Signal Processing*, vol. 121, pp. 94–101, 2016.

[28] T. Y. Al-Naffouri and A. H. Sayed, "Transient analysis of data-normalized adaptive filters," *IEEE Transactions on Signal Processing*, vol. 51, pp. 639–652, March 2003.

[29] H.-C. Shin and A. H. Sayed, "Mean-square performance of a family of affine projection algorithms," *IEEE Transactions on Signal Processing*, vol. 52, pp. 90–102, Jan 2004.

[30] D. B. Haddad and M. R. Petraglia, "Transient and steady-state mse analysis of the impnlms algorithm," *Digital Signal Processing*, vol. 33, pp. 50 – 59, 2014.

[31] I. TSG, *5, digital network echo cancellers (recommendation*. tech. rep., Tech. Rep. G. 168, ITU-T, 2004.

ISBN 978-90-827970-1-5 © EURASIP 2018