Analysis of the Normalized LMS Optimum Solution in the Context of Channel Equalization

Levy Boccato, Rafael Ferrari, Romis Attux

***Abstract*—Albeit being presented as an alternative to the classical least-mean-square (LMS) algorithm, the normalized LMS (NLMS) actually deals with a modified mean squared error (MSE) cost function, so that the expected optimum solution may differ from the Wiener solution. In this work, we perform an investigation concerning the question as to whether such difference may arise in the context of the channel equalization problem by considering a representative set of transmitted signal modulations, channel models and signal-to-noise ratio (SNR) conditions. Additionally, we analyze the influence of the potential deviation from the optimal solution on the performance of the equalizer.**

***Index Terms*—Adaptive filtering, Channel equalization, Normalized LMS**

I. Introduction

THE problem of adaptive filtering comprises three main choices: a) that of an adequate filter structure; b) that of a statistical criterion that expresses in mathematical terms what is expected from the filtering process; c) that of an optimization method that estimate the system parameters according to the chosen criterion. It is possible to state that the most classical setup is that of a linear finite impulse response (FIR) filter whose adaptation is based on the minimum mean squared error (MSE) criterion and carried out with the aid of gradient-based algorithms [1].

Undoubtedly, the problem of channel equalization represents an emblematic application of the adaptive filtering framework, in which a carefully designed filter (termed equalizer) is employed to cancel out the noxious effects of a certain communication channel, thereby, in a certain sense, inverting it. The problem is supervised when the equalizer parameters are modified online with the aid of reference samples taken from the transmitted signal [1].

L. Boccato (Corresponding Author), R. Ferrari and R. Attux are with the School of Electrical and Computer Engineering (FEEC), University of Campinas, SP, Brazil (email: {lboccato,rferrari,attux}@dca.fee.unicamp.br).

Digital Object Identifier: 10.14209/jcis.2018.3.

In this context, given that the MSE criterion calls for knowledge of statistical expectations of certain terms related to information signals, the optimization process is carried out according to two main possibilities: 1) the use of instantaneous estimates instead of statistical expectations and 2) the approximation of the MSE by means of time averages. The first of these alternatives engenders the least-mean-square (LMS) algorithm, whereas the second is the basis of the recursive least squares (RLS) algorithm [1].

Despite the importance of the RLS algorithm, it is possible to consider the LMS as the canonical adaptive algorithm within the MSE-based framework. The algorithm has a very simple and elegant expression:

$$\mathbf{w}(n+1) = \mathbf{w}(n) + \mu e^*(n)\mathbf{x}(n), \qquad (1)$$

where $\mathbf{x}(n) = [x(n)x(n-1)\dots x(n-K+1)]^{\mathrm{T}}$ is the input vector of a FIR filter of order K, $e^*(n)$ is the complex conjugate of $e(n)$, the difference between a reference signal $d(n)$ and the filter output $y(n) = \mathbf{w}^{\mathrm{H}}(n)\mathbf{x}(n)$, μ is a step-size parameter and $\mathbf{w}(n)$ is the filter coefficient vector at instant n.

The practical use of the LMS algorithm raises a crucial tradeoff regarding the choice of the step-size parameter μ in terms of convergence rate and misadjustment. The normalized LMS algorithm [1], [2] (NLMS) is an attempt to find a compromise by means of the adoption of a variable step-size parameter. This is done by introducing the term $\mathbf{x}^{\mathrm{H}}(n)\mathbf{x}(n)$ as a regulator, as it is possible to show that this term is crucial to determine the variation of the instantaneous squared error [1], [2]:

$$\mathbf{w}(n+1) = \mathbf{w}(n) + \frac{\mu}{\mathbf{x}^{\mathrm{H}}(n)\mathbf{x}(n)} e^*(n)\mathbf{x}(n). \qquad (2)$$

The introduction of the normalization term in (2) ends up modifying the underlying cost function, so that the NLMS algorithm is not, in fact, minimizing the instantaneous squared error, as the LMS. Hence,

it is possible that the optimum solution for this modified cost function be different from the Wiener solution. In other words, the NLMS may converge to a coefficient vector that is not the Wiener solution.

This aspect is of paramount importance since the NLMS algorithm usually is presented as an alternative to the conventional LMS for adapting the filter parameters. However, if these algorithms should converge to different solutions, there actually would be two distinct approaches for the adaptive filtering problem. In [3], it was demonstrated that the mean behavior of the NLMS converges to the Wiener solution considering white Gaussian input signal, which means that the modified optimal solution is almost equivalent to the classical Wiener solution.

Nevertheless, this result cannot be directly applied to the channel equalization problem, since the hypothesis that the input signal is a white process is not correct here due to the intersymbol interference (ISI). Additionally, depending on the characteristics of the channel, the received signal is not adequately represented as a Gaussian random variable. Other aspects regarding the NLMS algorithm, such as the convergence rate (speed) and the steady-state performance were analyzed in several works [4]–[10]. However, these studies generally focused on the system identification task and, in most cases, with Gaussian inputs.

Therefore, to the best of our knowledge, an analysis of the behavior of NLMS – in terms of the expected optimum filter it may obtain – specifically in the context of digital channel equalization has not been performed yet. In this work, we aim at contributing to bridge this gap, giving particular attention to the question as to whether the NLMS algorithm can converge to a solution different from the Wiener filter and, in such case, analyze the adequacy of the obtained filter in performing the desired task. Hence, the applicability of the NLMS in this kind of task will be further clarified.

This paper is organized as follows: Section II describes the main concepts of the channel equalization problem, as well as highlights the expected solutions obtained by the LMS and the NLMS. Then, Section III presents the experimental results, considering a representative set of input modulations, channels (both FIR and infinite impulse response (IIR) filters) and noise, and analyzes the behavior of the NLMS optimum solution when compared with the Wiener solution. Finally, Section IV brings the conclusions and perspectives for future works.

II. Problem Statement

The main elements involved in the supervised channel equalization problem are depicted in Figure 1.

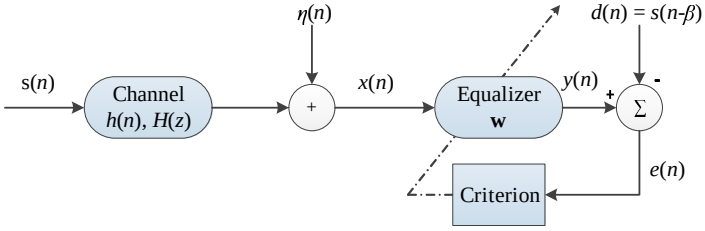

Fig. 1. Basic diagram of the supervised channel equalization problem.

The transmitted signal $s(n)$ is composed of independent and identically distributed (i.i.d.) samples belonging to the discrete alphabet $\mathcal{S}$ associated with the chosen modulation (e.g., BPSK or 4-QAM). The channel transfer function $H(z)$ models the effect known as intersymbol interference (ISI) [1] and $\eta(n)$ represents an additive white Gaussian noise with zero mean and variance σ_η^2. The objective in channel equalization is to remove as much as possible the ISI in an attempt to recover $s(n)$ or a delayed version thereof ($s(n-\beta)$), reaching the so-called zero-forcing (ZF) condition [1].

In this work, the equalizer consists of a FIR filter with K coefficients, defined as $\mathbf{w} = [w_0 \dots w_{K-1}]^T$, which shall be adapted with the aid of the LMS and the NLMS algorithms. When the channel and the involved signals do not change their properties during the communication process, by properly selecting the value of the step-size in (1), it is expected that the LMS converge to the Wiener solution. On the other hand, the NLMS algorithm actually deals with a modified cost function, so that the optimal solution it may find can, in theory, be different.

This aspect can be more clearly understood if we interpret the NLMS algorithm as an instance of the standard LMS applied to modified data. By rearranging the terms in (2), we can write:

$$\begin{aligned}\mathbf{w}(n+1) &= \mathbf{w}(n) + \frac{\mu}{\|\mathbf{x}(n)\|}(d(n)-y(n))^*\frac{\mathbf{x}(n)}{\|\mathbf{x}(n)\|} \\ &= \mathbf{w}(n) + \mu\tilde{e}^*(n)\tilde{\mathbf{x}}(n), \end{aligned} \quad (3)$$

where $\tilde{e}(n) = \tilde{d}(n) - \tilde{y}(n)$, $\tilde{d}(n) = \frac{d(n)}{\|\mathbf{x}(n)\|}$ is the modified reference signal, $\tilde{y}(n) = \mathbf{w}^{\mathrm{H}}(n)\frac{\mathbf{x}(n)}{\|\mathbf{x}(n)\|}$ and $\tilde{\mathbf{x}}(n) = \frac{\mathbf{x}(n)}{\|\mathbf{x}(n)\|}$ is the modified input signal.

So, while the LMS seeks the solution that minimizes $E\{e(n)e^*(n)\}$, the underlying cost function of NLMS corresponds to $E\{\tilde{e}(n)\tilde{e}^*(n)\}$. Therefore, there are two potentially distinct optimal solutions [6]: the conventional Wiener filter, denoted as $\mathbf{w}_{\text{wiener}}$, which uses the statistical information of the input autocorrelation matrix $\mathbf{R}_x = E\{\mathbf{x}(n)\mathbf{x}^{\mathrm{H}}(n)\}$ and of the cross-correlation vector $\mathbf{p}_{xd} = E\{\mathbf{x}(n)d^*(n)\}$, and a modified solution, which makes use of normalized data and is given by:

$$\mathbf{w}_{\text{mod}} = \tilde{\mathbf{R}}_x^{-1}\tilde{\mathbf{p}}_{xd}, \tag{4}$$

where $\tilde{\mathbf{R}}_x = E\{\tilde{\mathbf{x}}(n)\tilde{\mathbf{x}}^{\mathrm{H}}(n)\}$ and $\tilde{\mathbf{p}}_{xd} = E\{\tilde{\mathbf{x}}(n)\tilde{d}^*(n)\}$.

In this scenario, one may wonder whether the aforementioned solutions are equivalent or, at least, approximately equal. This is the main aspect to be investigated in this work, particularly in the context of of supervised equalization of digital signals. Even though it is not feasible to compute in analytical terms the statistical entities involved in the modified (normalized) solution, we can resort to simple estimates using a sufficiently large number of samples, since the signals considered in this work are stationary and ergodic. Hence, the analysis carried out in this work shall be based on a set of experimental results in different scenarios of the channel equalization problem. Notwithstanding, by considering representative conditions with respect to the input signal, the channel and the noise, we believe that the obtained results are capable of providing a broader view concerning the adequacy of NLMS for the equalizer design.

III. Experimental Results

In this section, we are interested in comparing the characteristics and attainable performances associated with the Wiener solution and with the modified solution, defined in (4), which, as discussed in Section II, establishes the actual optimum filter that the NLMS algorithm pursues. Hence, we shall analyze the theoretical MSE value associated with each solution, defined as [1]:

$$\begin{aligned}\text{MSE}(\mathbf{w}) &= E\{|s(n-\beta)|^2\} - \mathbf{w}^{\mathrm{H}}\mathbf{p}_{xd} \\ &\quad -\mathbf{p}_{xd}^{\mathrm{H}}\mathbf{w} + \mathbf{w}^{\mathrm{H}}\mathbf{R}_x\mathbf{w}\end{aligned} \tag{5}$$

where $\mathbf{w}$ can be $\mathbf{w}_{\text{wiener}}$ or $\mathbf{w}_{\text{mod}}$. The term $E\{|s(n-\beta)|^2\}$ is the mean energy of the transmitted signal.

In the experiments, we use the Normalized MSE Difference (NMD), given by

$$\text{NMD} = \frac{\text{MSE}(\mathbf{w}_{\text{mod}}) - \text{MSE}(\mathbf{w}_{\text{wiener}})}{\text{MSE}(\mathbf{w}_{\text{wiener}})}, \tag{6}$$

as a performance metric to compare the solutions, which represents the percentage of deviation between the MSE values associated with $\mathbf{w}_{\text{wiener}}$ and $\mathbf{w}_{\text{mod}}$.

Additionally, we also assess the normalized Euclidean distance (NED) between the corresponding filter coefficient vectors:

$$\text{NED} = \frac{\|\mathbf{w}_{\text{wiener}} - \mathbf{w}_{\text{mod}}\|}{\|\mathbf{w}_{\text{wiener}}\|}. \tag{7}$$

It is important to mention that the modified solution ($\mathbf{w}_{\text{mod}}$) is computed as an average of $N_r = 50$ independent estimates, where each estimate is obtained via (4) using sample mean approximations of $\tilde{\mathbf{R}}_x$ and $\tilde{\mathbf{p}}_{xd}$, considering a set of $T = 10000$ transmitted symbols. On the other hand, the Wiener solution ($\mathbf{w}_{\text{wiener}}$) is analytically calculated for the considered scenarios.

We shall consider two types of channel: (*i*) a minimum-phase FIR system, and (*ii*) an IIR system. The equalization delay was $\beta = 0$, which represents an adequate choice for these channels having in view the attainable performance of the Wiener solution.

The input signals considered in this work are related to traditional digital modulation schemes, *viz.*, BPSK, 4-QAM, 8-PSK and 16-QAM [11], and present unitary mean energy (i.e., $E\{s^2(n)\} = 1$) in order to allow a direct comparison of the results under the same SNR conditions.

A. Minimum-Phase FIR Channel

In this scenario, the transfer function of the channel is given by $H(z) = 1 + \alpha z^{-1}$, with $0 < \alpha < 1$. Hence, the closer is α to unity, the more difficult is the inversion of the channel when using a FIR equalizer, or, in other words, more coefficients are necessary to reach an adequate cancellation of the channel.

For the case in which the equalizer has $K = 2$ coefficients, we derived the exact expression of

the modified solution by computing the statistical expectation considering all the possible received signal vectors ($\mathbf{x}(n)$) in the absence of noise. In this case, the modified solution is given by

$$\mathbf{w}_{\text{mod}} = \begin{bmatrix} \frac{1}{1-\alpha^4} \\ \frac{-\alpha}{1-\alpha^4} \end{bmatrix}. \quad (8)$$

On the other hand, the Wiener solution for the same scenario corresponds to:

$$\mathbf{w}_{\text{wiener}} = \begin{bmatrix} \frac{1+\alpha^2}{1+\alpha^2+\alpha^4} \\ \frac{-\alpha}{1+\alpha^2+\alpha^4} \end{bmatrix}. \quad (9)$$

These expressions shall be useful for the analysis of the results in the sequence of the text.

Figure 2 shows the NMD values as a function of α in the absence of noise for a BPSK transmitted signal and for several equalizer lengths. The NED between the solutions is depicted in Figure 3.

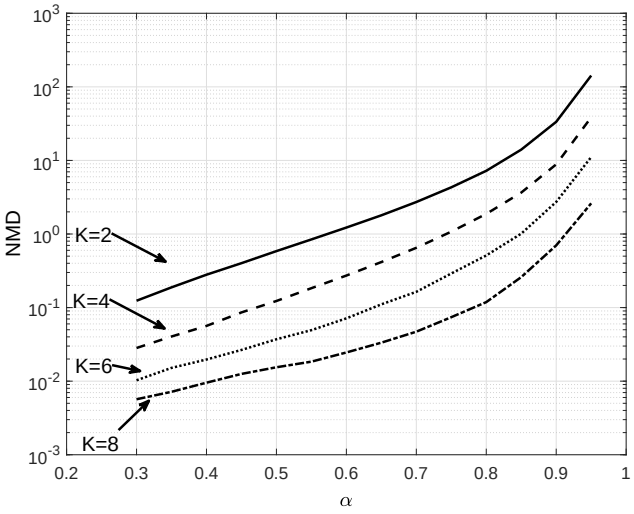

Fig. 2. Normalized MSE difference between $\mathbf{w}_{\text{wiener}}$ and $\mathbf{w}_{\text{mod}}$ for several equalizer lengths considering the FIR channel, no noise and BPSK modulation.

It is possible to notice that there can be a significant deviation of MSE and of distance between $\mathbf{w}_{\text{wiener}}$ and $\mathbf{w}_{\text{mod}}$. In particular, we observe that both NMD and NED values increase as the channel coefficient α is increased.

For $K = 2$, as α approaches the unity, the coefficients of $\mathbf{w}_{\text{mod}}$, given by (8), tend to increase in magnitude and, in the limit, diverge. On the other hand, the Wiener solution always preserve a limited magnitude for each coefficient, according to (9), which explains the behavior of the NED curve in this case. Additionally, based on the expressions of

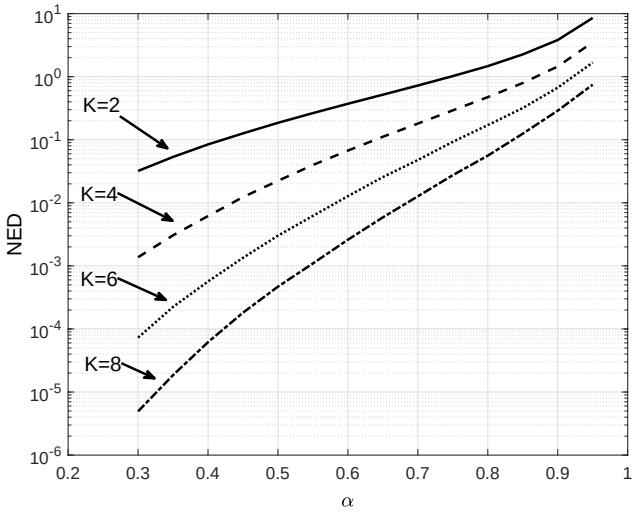

Fig. 3. Normalized Euclidean distance between $\mathbf{w}_{\text{wiener}}$ and $\mathbf{w}_{\text{mod}}$ for several equalizer lengths considering the FIR channel, no noise and BPSK modulation.

$\mathbf{w}_{\text{mod}}$ and $\mathbf{w}_{\text{wiener}}$, it also becomes evident why the solutions are more similar when $\alpha \ll 1$, reaching equality for $\alpha = 0$. Hence, we notice a potential connection between the NED values and the difficulty for inverting the channel.

This connection is also corroborated by analyzing the impact of the length of the equalizer: for any value of α, the more coefficients the equalizer has, the smaller is the difference between $\mathbf{w}_{\text{mod}}$ and $\mathbf{w}_{\text{wiener}}$, and the better is the approximation of the channel inverse by means of the equalizer.

The influence of the noise in the NMD and in the NED is shown on Figures 4 and 5, respectively. The results were obtained using $K = 2$ and BPSK modulation, which is the configuration that attained the largest deviation in the previous scenario.

As we can observe, the NMD and the NED are no longer monotonically increasing with α as in the noiseless case. As α increases, both performance metrics increase until they reach a maximum value and, then, start to decrease as α approaches unity. The peaks of NMD and NME, as well the value of α at which they occur, decrease as the SNR is reduced. Interestingly, the noise significantly reduces the difference between $\mathbf{w}_{\text{wiener}}$ and $\mathbf{w}_{\text{mod}}$ when the ZF condition is harder to attain.

Having in view the potential connection between the difficulty for inverting the channel and the expected differences between $\mathbf{w}_{\text{wiener}}$ and $\mathbf{w}_{\text{mod}}$, raised during the analysis of the noiseless case, we believe

that the presence of noise ends up introducing, in a certain sense, a regularization factor in the computation of $\mathbf{w}_{\text{mod}}$, which avoids the divergence of the solution for α close to the unity and, ultimately, makes $\mathbf{w}_{\text{mod}}$ more similar to the Wiener solution.

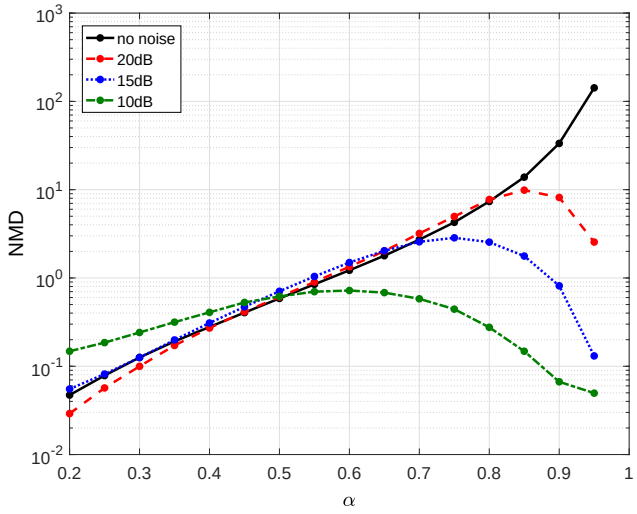

Fig. 4. Normalized MSE difference between $\mathbf{w}_{\text{wiener}}$ and $\mathbf{w}_{\text{mod}}$ for several SNR values considering the FIR channel, $K = 2$ and BPSK modulation.

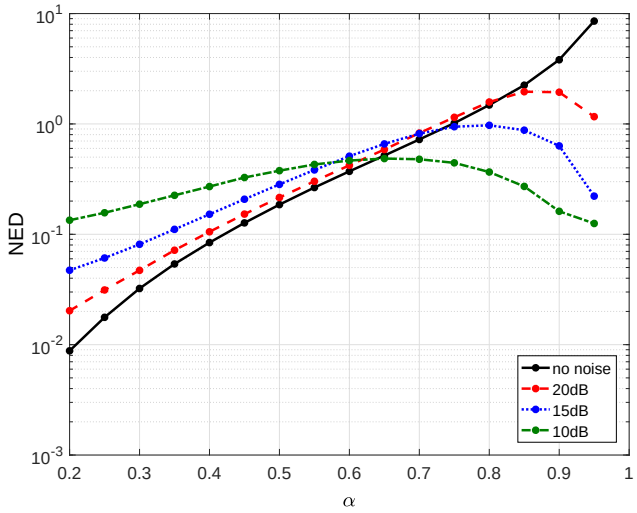

Fig. 5. Normalized Euclidean distance between $\mathbf{w}_{\text{wiener}}$ and $\mathbf{w}_{\text{mod}}$ for several SNR values considering the FIR channel, $K = 2$ and BPSK modulation.

Next, we assess the impact of the modulation of the transmitted signal. We keep the equalizer length constant and equal to $K = 2$, which is the worst case observed in the first experiment. Figures 6 and 7 show the NMD and NED respectively as functions of α for several modulations in the absence of noise.

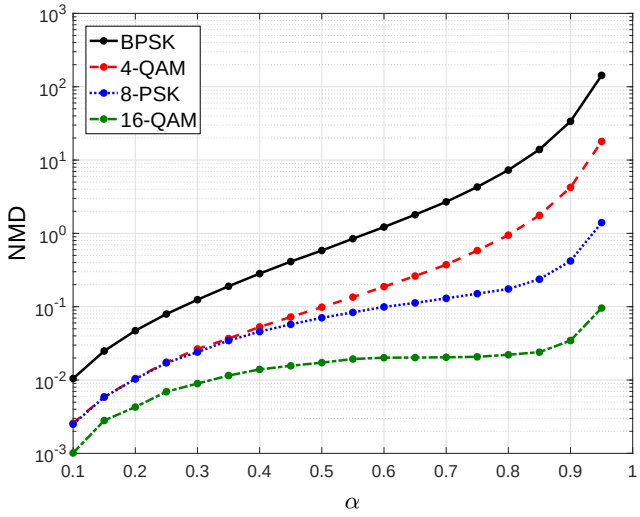

Fig. 6. Normalized MSE difference between $\mathbf{w}_{\text{wiener}}$ and $\mathbf{w}_{\text{mod}}$ for several modulations considering the FIR channel, no noise and $K = 2$.

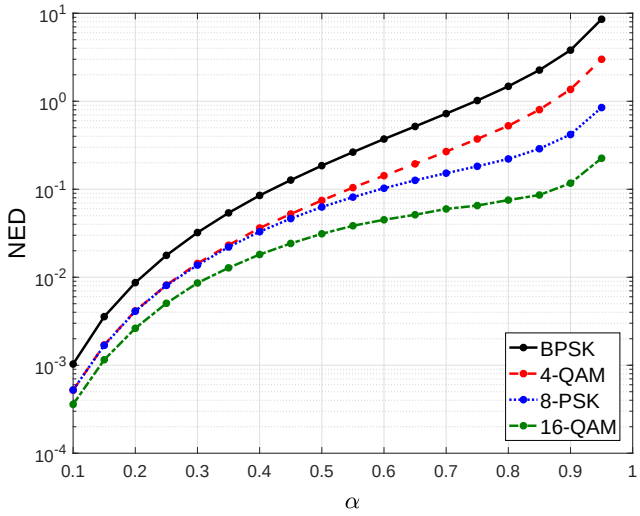

Fig. 7. Normalized Euclidean distance between $\mathbf{w}_{\text{wiener}}$ and $\mathbf{w}_{\text{mod}}$ for several modulations considering the FIR channel, no noise and $K = 2$.

The results show that the larger the modulation cardinality, the smaller the difference between the modified and the Wiener solution. In fact, we verified through experimental simulations that if $s(n)$ has a uniform distribution, which is equivalent to a modulation with infinite cardinality, the difference between the solutions is negligible, independently of α. The same behavior was observed when the

distribution of $s(n)$ is Gaussian. Therefore, the distribution of the transmitted signal has an important impact on the filter attained by the NLMS.

The difference between $\mathbf{w}_{\text{wiener}}$ and $\mathbf{w}_{\text{mod}}$ also impacts on the probability of error of the system. Figure 8 shows the theoretical probability of error of the BPSK [12] as a function of α for equalizers with $K = 2$, $K = 4$ and $K = 8$ coefficients, when the SNR is set to 20 dB. The curves shown in Figure 9 are obtained when the SNR is 15 dB.

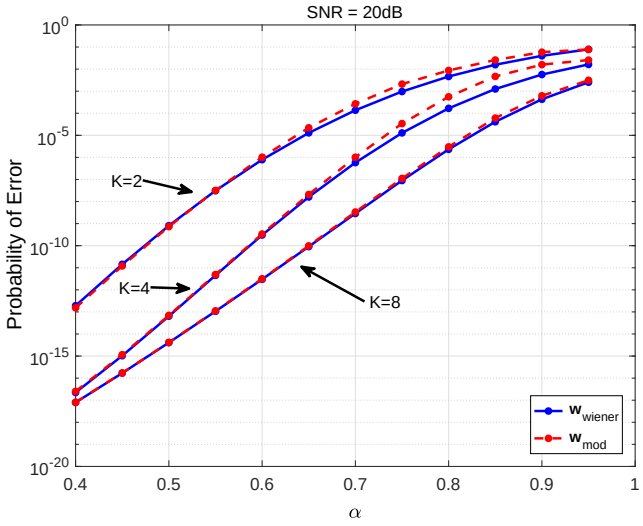

Fig. 8. Theoretical probability of error for several equalizer lengths considering the FIR channel, BPSK modulation and SNR= 20dB.

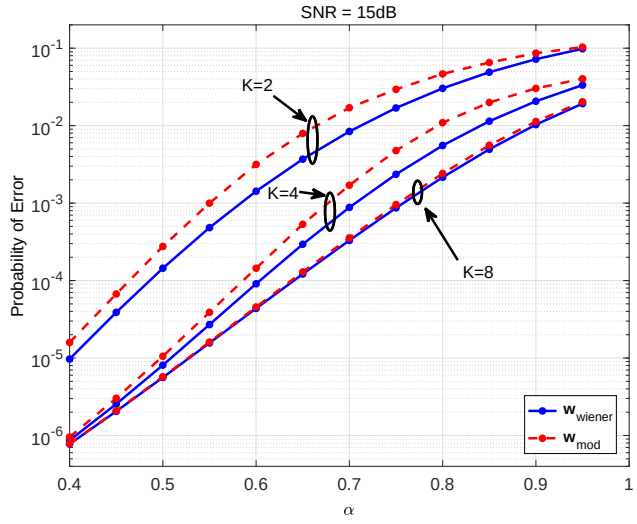

Fig. 9. Theoretical probability of error for several equalizer lengths considering the FIR channel, BPSK modulation and SNR= 15dB.

Note that $\mathbf{w}_{\text{mod}}$ provides a worse performance than $\mathbf{w}_{\text{wiener}}$ for intermediate values of α, which is in accordance with the results shown in Figures 4 and 5, since the largest deviation between the solutions also occurs for intermediate values of α. However, the difference between the curves is smaller for larger equalizer lengths, which is expected because the deviation between the solutions is smaller in those conditions, as previously discussed. Additionally, we can observe that the performance gap increases as the SNR is reduced.

Finally, we show in Figure 10 the contours of the standard and modified MSE surfaces along with the trajectories associated with the LMS and the NLMS considering BPSK modulation, SNR of 20 dB and $\alpha = 0.8$, a scenario where a large difference between $\mathbf{w}_{\text{mod}}$ and $\mathbf{w}_{\text{wiener}}$ is expected based on the results obtained so far. The initial condition and the adopted step size were equal to $\mathbf{w} = [-2; -2]^T$ and $\mu = 0.001$, respectively, for both the algorithms.

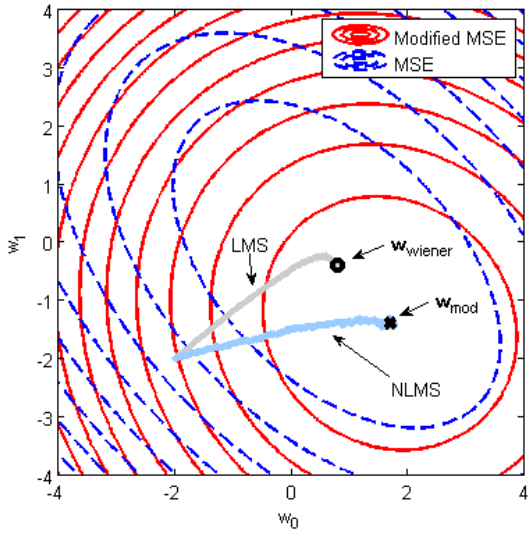

Fig. 10. Contours of the MSE and modified MSE surfaces considering the FIR channel with $\alpha = 0.8$ and the SNR of 20 dB. The X mark refers to $\mathbf{w}_{\text{mod}}$, whereas the circle represents the Wiener solution.

As expected, the normalization of data in the NLMS improves the conditioning of the filter input autocorrelation matrix, since the contours of the modified MSE surface are more similar to circles, which may accelerate the convergence of a gradient-based algorithm towards the optimum solution. However, the normalization has an undesirable effect: the NLMS may not lead to the Wiener filter, as can bee seen by the difference between $\mathbf{w}_{\text{mod}}$ and $\mathbf{w}_{\text{wiener}}$ in Figure 10.

B. IIR Channel

Now, the transfer function of the channel is given by $H(z) = \frac{1}{1+\alpha z^{-1}}$, which means that the ZF condition can be attained through the use of a FIR equalizer with only two coefficients ($K = 2$). Therefore, having in mind the observations raised in the previous scenario, it is expected that the difference between the Wiener and the modified solutions be less pronounced here. Additionally, since the potential difference between them is reduced as the cardinality of the alphabet associated with the input modulation increases, we shall concentrate the analysis on the BPSK modulation.

Figure 11 exhibits the NMD values as a function of α considering the SNRs of 10 dB and 30 dB, whereas Figure 12 displays the NED between these solutions.

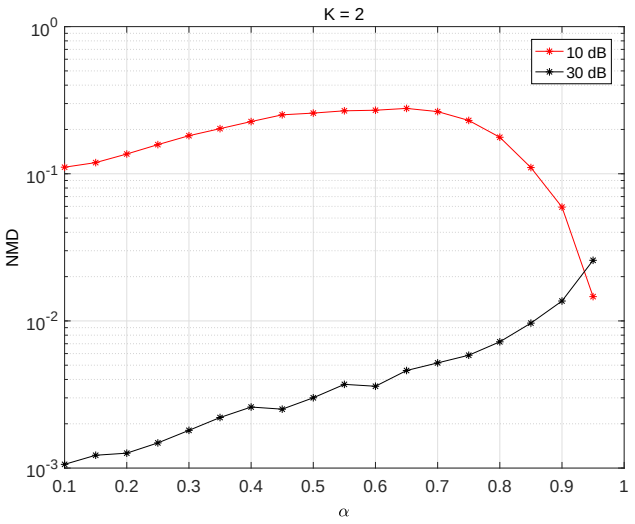

Fig. 11. NMD values associated with classical and modified Wiener solutions for the IIR channel and the SNRs of 10 dB and 30 dB.

It is possible to notice that when the noise power is small, $\mathbf{w}_{\text{wiener}}$ and $\mathbf{w}_{\text{mod}}$ are almost identical and, consequently, the NMD values are significantly small. On the other hand, when the SNR is 10 dB, the solutions differ in a relatively higher degree, and the maximum NMD value is close to 30%. So, differently from the case with the FIR channel, as more noise is present in the received signal, the optimum solution that the NLMS algorithm should converge to becomes more distinct from the Wiener solution. Notwithstanding, the distances between the solutions are quite smaller when compared with those observed for the FIR channel, which means

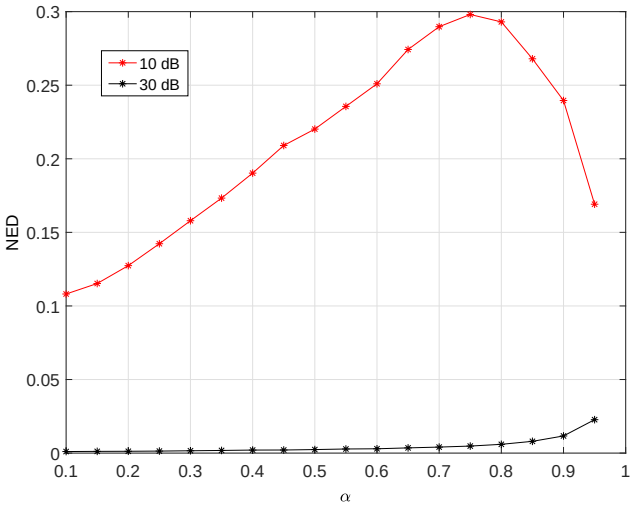

Fig. 12. Normalized Euclidean distance between $\mathbf{w}_{\text{wiener}}$ and $\mathbf{w}_{\text{mod}}$ considering the IIR channel and the SNRs of 10 dB and 30 dB.

that the obtained results have confirmed our expectation with respect to the potential difference between $\mathbf{w}_{\text{mod}}$ and $\mathbf{w}_{\text{wiener}}$ when the equalizer approaches the ZF condition.

In order to complete our analysis, we shall verify the behavior of the NLMS algorithm in this scenario. Thus, we show in Figure 13 the contours of the standard and modified MSE surfaces along with the trajectories associated with the LMS and the NLMS considering the SNR of 10 dB and $\alpha = 0.8$, which was a case with a large difference between $\mathbf{w}_{\text{mod}}$ and $\mathbf{w}_{\text{wiener}}$. The initial condition and the adopted step size were equal to $\mathbf{w} = [1.5; -2.2]^T$ and $\mu = 0.005$, respectively, for both the algorithms.

Similarly to the previous scenario, we can infer that the normalization of data contributes to a faster convergence of a gradient-based algorithm, but, due to the presence of noise, the optimum solution is slightly different than the Wiener filter.

IV. Conclusion

In this work, we investigated the behavior of the NLMS algorithm in terms of the expected optimum filter it may obtain in the context of the channel equalization problem. The analysis considered different situations regarding the transmitted signal modulation, the channel and the noise power, aiming at verifying the circumstances that may lead to a deviation from the Wiener solution.

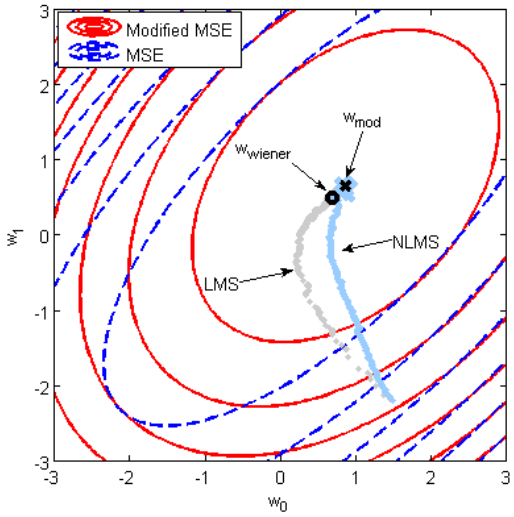

Fig. 13. Contours of the MSE and modified MSE surfaces considering the IIR channel and the SNR of 10 dB. The X mark refers to $\mathbf{w}_{\text{mod}}$, whereas the circle represents the Wiener solution.

The obtained results indicate that the optimum solution found by NLMS can be considerably different when compared with the Wiener solution in certain cases. In particular, the difference between these solutions tends to be more pronounced when the equalizer does not have as many coefficients as needed for a proper inversion of the channel. In such condition, the optimum equalizer found by NLMS yielded a worse performance in terms of the bit error probability when compared with the Wiener filter.

On the other hand, as we increase the size of the input modulation, the difference between the LMS and NLMS solutions is reduced, reaching negligible values when the input signal has a continuous distribution, even when the equalizer is not capable of perfectly inverting the channel. With respect to the effect of noise, two different aspects have been observed: (*i*) for the IIR channel, the presence of noise may increase the difference between $\mathbf{w}_{\text{mod}}$ and $\mathbf{w}_{\text{wiener}}$; (*ii*) for the FIR channel, the addition of noise makes the modified solution more similar to the Wiener solution. As perspective for future works, a theoretical analysis of the NLMS optimum solution is certainly pertinent.

ACKNOWLEDGMENTS

This work was supported by CNPq.

REFERENCES

[1] S. Haykin, *Adaptive filter theory*, 5th ed. Prentice Hall, 2013.

[2] P. S. R. Diniz, *Adaptive Filtering: Algorithms and Practical Implementation*, 4th ed. Springer, 2013.

[3] N. J. Bershad, "Analysis of the normalized LMS algorithm with gaussian inputs," *IEEE Trans. on Acoust., Speech, Signal Process.*, vol. 34, no. 4, pp. 793–806, 1986. doi: 10.1109/TASSP.1986.1164914

[4] M. Tarrab and A. Feuer, "Convergence and performance analysis of the normalized LMS algorithm with uncorrelated gaussian data," *IEEE Trans. Inf. Theory*, vol. 34, no. 4, pp. 680–691, 1988. doi: 10.1109/18.9768

[5] D. T. M. Slock, "On the convergence behavior of the LMS and the normalized LMS algorithms," *IEEE Trans. on Signal Process.*, vol. 41, pp. 2811–2825, 1993. doi: 10.1109/78.236504

[6] P. E. An, M. Brown, and C. J. Harris, "On the convergence rate performance of the normalized least-mean-square adaptation," *IEEE Trans. Neural Netw.*, vol. 8, no. 5, pp. 1211–1214, 1997. doi: 10.1109/72.623223

[7] E. Eweda, "A new approach for analyzing the limiting behavior of the normalized lms algorithm under weak assumptions," *Signal Processing*, vol. 89, no. 11, pp. 2143 – 2151, 2009. doi: https://doi.org/10.1016/j.sigpro.2009.04.040

[8] J. E. Kolodziej, "Algoritmos adaptativos com sinal de entrada normalizado: Modelagem estatística e aprimoramentos," Ph.D. dissertation, Federal University of Santa Catarina, Santa Catarina, Brazil, 2010.

[9] E. V. Kuhn, F. C. de Souza, R. Seara, and D. R. Morgan, "On the steady-state analysis of pnlms-type algorithms for correlated Gaussian input data," *IEEE Signal Process. Lett.*, vol. 21, no. 11, pp. 1433–1437, 2014. doi: 10.1109/LSP.2014.2332751

[10] M. V. Matsuo and R. Seara, "On the stochastic analysis of the NLMS algorithm for white and correlated Gaussian inputs in time-varying environments," *Signal Processing*, vol. 128, pp. 291–302, 2016. doi: http://dx.doi.org/10.1016/j.sigpro.2016.03.018

[11] J. R. Barry, E. A. Lee, and D. G. Messerschmitt, *Digital Communication*, 3rd ed. Springer, 2003.

[12] L. Boccato, D. G. Fantinato, D. G. Silva, R. Ferrari, A. Neves, and R. Attux, "Analysis of ITL criteria in the context of FIR channel equalization," *Journal of Communications and Information Systems*, vol. 31, no. 1, pp. 1–29, 2016. doi: http://dx.doi.org/10.14209/jcis.2016.1

Levy Boccato was born in Piracicaba, São Paulo, Brazil, in 1986. He received the titles of Computer Engineer (2008), Master in Electrical Engineering (2010) and Doctor in Electrical Engineering (2013), all from the University of Campinas (UNICAMP), São Paulo, Brazil. Currently, he is an Assistant Professor at the same university. His main research interests include computational intelligence, adaptive filtering, machine learning and signal processing.

Rafael Ferrari received his B.S. (2001), M.Sc.(2005), and Ph.D. (2011) degrees in Electrical Engineering from the University of Campinas (UNICAMP). From 2011 to 2015, he has been a researcher at the Center for Petroleum Studies (CEPETRO-UNICAMP). Currently, he is an Assistant Professor at the School of Electrical and Computer Engineering (FEEC) of UNICAMP. His main research interests include computational intelligence and digital signal processing, especially their application to communication systems and to seismic data analysis.

Romis Attux was born in Goiânia, Brazil, in 1978. He obtained the degrees of Electrical Engineer (1999), Master in Electrical Engineering (2001) and Doctor in Electrical Engineering from the University of Campinas (UNICAMP), Brazil. He is currently an Associate Professor at the same university. His main research areas are unsupervised signal processing, computational intelligence, dynamical systems / chaos and brain-computer interfaces.