

# On the Step-Size of Leakage QN and Leakage LS Algorithms

Fabiano T. Castoldi and Marcello L. R. de Campos

**Abstract**—Recursive Least Squares (RLS) algorithms are known to have fast convergence in the presence of correlated input signals that have persistent excitation. However, when non-persistently exciting signals or quantization errors are present in the system, the RLS algorithms may diverge. This fact happens due to the loss of positive definiteness of the input signal correlation matrix. There is a class of quasi-Newton algorithms that use a correlation matrix that once initialized positive definite, it remains in this state independently of the input signal. In this paper some algorithms are developed using different approaches to the calculation of the step-size, such that the condition of positive definiteness of the correlation matrix is guaranteed.

**Keywords**—Leakage Least-Squares Algorithm, Variable Step-Size, Robustness, Optimization Problems, Logarithmic Barrier.

## I. INTRODUCTION

Many adaptation algorithms have been proposed in the past forty years, offering trade-offs in convergence speed, robustness, and computational complexity. Least-Squares (LS) algorithms usually have better convergence speed in comparison to Least Mean Squares (LMS) algorithms. However, many Recursive Least-Squares (RLS) algorithms lack robustness, due to accumulation of quantization errors, or due to loss of positive definiteness of the Hessian matrix caused by non-persistently exciting input signals [1].

The conventional RLS algorithm presents good convergence speed, but its robustness is not guaranteed unless we opt for a QR-decomposition implementation, or some other regularization scheme. Robust RLS-algorithm implementations with reduced computational complexity are usually based on QR-decompositions, which are complex to implement and maintain [2]. There are other algorithms that have been developed based on known convex optimization methods, like the quasi-Newton (QN) and IPLS (Interior Point Least Square) algorithms [1], [3]. These algorithms offer increased robustness at a cost of extra computational complexity, for they do not admit  $O(N)$  implementations.

The QN algorithm in [1] is a robust algorithm that presents good convergence capabilities without user-defined variables, which gives a further level of robustness. However, it does not allow fast implementation and it also has a constant misadjustment of 3 dB due its normalization, which cannot be reduced. A different approach to derive adaptation algorithms was developed in [4], and as a by product a new leakage QN

algorithm was also presented. This leakage QN algorithm also does not need user-defined quantities and has a slightly lower misadjustment in comparison to the QN algorithm.

The method to derive algorithms presented in [4] allows minimization of the cost function, in particular in LS problems, using convex optimization techniques and a minimum disturbance approach. The derivation of new variable step-sizes for LS algorithms is presented in this work, using different constraints to the input-signal correlation matrix.

This paper is organized as follows: In the following subsection some common notations will be presented for better understanding the text. Section II presents the Leakage QN algorithm developed in [4]. An alternative version, using inequality constraints, is derived in Section III. The Leakage LS algorithm and its modified version are presented in Sections IV and V, respectively. Simulations are carried out to show the convergence performance of the algorithms in Section VI and conclusions are given in Section VII.

### A. Common Notations

In this work vectors are represented by bold lower case letters and matrices are represented by bold upper case letters. The representation of variables in the text is similar to that of the RLS algorithm, which is presented here and will be used in comparisons in Section VI. The conventional version of the RLS algorithm can be implemented as:

$$e_n = d_n - \mathbf{w}_{n-1}^T \mathbf{x}_n \quad (1)$$

$$\mathbf{t}_n = \mathbf{R}_{n-1}^{-1} \mathbf{x}_n \quad (2)$$

$$\tau_n = \mathbf{x}_n^T \mathbf{t}_n \quad (3)$$

$$\mathbf{R}_n^{-1} = \frac{1}{\lambda} \left[ \mathbf{R}_{n-1}^{-1} - \frac{\mathbf{t}_n \mathbf{t}_n^T}{\lambda + \tau_n} \right] \quad (4)$$

$$\mathbf{w}_n = \mathbf{w}_{n-1} + e_n \mathbf{R}_n^{-1} \mathbf{x}_n \quad (5)$$

where  $e_n$  is the *a priori* error,  $d_n$  is the reference signal,  $\mathbf{x}_n$  is the input-signal vector,  $\mathbf{w}_n$  is the filter coefficient vector, all at iteration  $n$ , and the vectors have length  $N$ . The constant  $\lambda \in (0, 1]$  is purposely called forgetting factor because it controls how past information is weighted to form the estimates of the input-signal correlation matrix,  $\mathbf{R} = E[\mathbf{x}_n \mathbf{x}_n^T]$ . The operator  $(\cdot)^T$  is the transposed of a vector or a matrix.  $\mathbf{R}_n^{-1}$  is the inverse of the correlation matrix and is usually initialized as  $\mathbf{R}_0^{-1} = \gamma \mathbf{I}$ , where  $\mathbf{I}$  is the identity matrix with dimension  $N \times N$  and  $\gamma$  is a user-defined constant, usually  $\gamma = 1$ .

## II. LEAKAGE QN ALGORITHM

The leakage QN algorithm derived in [4] was developed by solving the following minimization problem with respect to the filter coefficients at instant  $n$ :

$$\begin{aligned} \min_{\mathbf{w}_n} \quad & \mu_n \|\mathbf{w}_n - \mathbf{w}_{n-1}\|_{\mathbf{R}_{n-1}}^2 + [d_n - \mathbf{x}_n^T \mathbf{w}_n]^2 \\ \text{s.t.} \quad & \mathbf{x}_n^T \mathbf{R}_n^{-1} \mathbf{x}_n = 1/2 \end{aligned} \quad (6)$$

For this algorithm the correlation matrix is  $\mathbf{R}_n = \mathbf{R}_{n-1} + \mu_n \mathbf{x}_n \mathbf{x}_n^T$  and its inverse, using the matrix inversion lemma, is

$$\mathbf{R}_n^{-1} = \frac{1}{\mu_n} \left[ \mathbf{R}_{n-1}^{-1} - \frac{\mathbf{t}_n \mathbf{t}_n^T}{\mu_n + \tau_n} \right] \quad (7)$$

with  $\mathbf{t}_n$  and  $\tau_n$  as in equations (2) and (3), respectively, and  $\mu_n$  is the step-size limited as  $0 \leq \mu_n \leq 1$ .

It can be seen from the minimization problem of equation (6) that the constraint imposed in the cost function does not include the coefficient vector,  $\mathbf{w}_n$ , having no influence in the coefficient update values. However, the constraint imposes to the correlation matrix a condition. Therefore, one can choose the step-size such that the constraint is satisfied, e.g., by directly applying the correlation matrix update formula [4] to the constraint:

$$\mathbf{x}_n^T \left\{ \frac{1}{\mu_n} \left[ \mathbf{R}_{n-1}^{-1} - \frac{\mathbf{t}_n \mathbf{t}_n^T}{\mu_n + \tau_n} \right] \right\} \mathbf{x}_n = 1/2 \quad (8)$$

The resulting algorithm is presented in Table I.

TABLE I  
LEAKAGE QN (LQN) ALGORITHM

|  |
|--|
| $\begin{aligned} e_n &= d_n - \mathbf{x}_n^T \mathbf{w}_{n-1} \\ \mathbf{t}_n &= \mathbf{R}_{n-1}^{-1} \mathbf{x}_n \\ \tau_n &= \mathbf{x}_n^T \mathbf{t}_n \\ \mu_n &= \tau_n \\ \mathbf{R}_n^{-1} &= \frac{1}{\mu_n} \left[ \mathbf{R}_{n-1}^{-1} - \frac{\mathbf{t}_n \mathbf{t}_n^T}{\mu_n + \tau_n} \right] \\ \mathbf{w}_n &= \mathbf{w}_{n-1} + \frac{e_n}{\mu_n + \tau_n} \mathbf{t}_n \end{aligned}$ |
|--|

Different constraints can be imposed to the algorithm, allowing the correlation matrix's eigenvalues to be lower and upper bounded within the  $(\alpha, \beta)$  interval. However, the use of inequalities cannot be easily solved using Lagrange multipliers, so the logarithmic barrier method can be used to help the solution.

## III. MODIFIED LEAKAGE QN ALGORITHM

### A. Logarithmic Barrier

The barrier method is an interior-point method that simplifies the inequality constraints in the cost function with a sum of indication functions  $I_-(f_i(x))$  in the cost function. The indication function for a constraint  $f_i(x) < 0$ , for  $i = 1, \dots, m$  (where  $m$  is the number of inequality constraints), has value zero for  $f_i(x) \leq 0$  and tends to  $\infty$  for  $f_i(x) \geq 0$ . The logarithmic barrier makes an approximation of this function as

$$\hat{I}_-(f_i(x)) = -(1/t) \log(-f_i(x)) \quad (9)$$

where the parameter  $t$  adjusts the approximation precision. As  $t$  increases, so does the precision.

The gradient of the logarithmic barrier is given as

$$\nabla \hat{I}_-(f_i(x)) = \frac{1}{-f_i(x)} \nabla f_i(x) \quad (10)$$

For more information about interior-point methods and the logarithmic barrier see [5].

### B. Algorithm Derivation

Using the same cost function as in equation (6) but with different constraints to limit both the lower and upper bound of the correlation matrix, the new minimization problem can be presented as follows:

$$\begin{aligned} \min_{\mathbf{w}_n} \quad & \mu_n \|\mathbf{w}_n - \mathbf{w}_{n-1}\|_{\mathbf{R}_{n-1}}^2 + [d_n - \mathbf{x}_n^T \mathbf{w}_n]^2 \\ \text{s.t.} \quad & \begin{cases} \mathbf{x}_n^T \mathbf{R}_n^{-1} \mathbf{x}_n > \alpha \\ \mathbf{x}_n^T \mathbf{R}_n^{-1} \mathbf{x}_n < \beta \end{cases} \end{aligned} \quad (11)$$

These constraints allow better control over the correlation matrix behavior, making explicit the positive definiteness property and limiting how much the matrix can grow to avoid numeric overflow problems. The update equations for the coefficient vector and for the inverse of correlation matrix are the same. The step-size now can be calculated by minimizing a convex function constructed as the combination of the two constraints, with the logarithmic barrier function used to approximate the inequality:

$$\begin{aligned} \min_{\mu_n} \quad & -(1/t_1) \log(\mathbf{x}_n^T \mathbf{R}_n^{-1} \mathbf{x}_n - \alpha) \\ & - (1/t_2) \log(-\mathbf{x}_n^T \mathbf{R}_n^{-1} \mathbf{x}_n + \beta) \end{aligned} \quad (12)$$

Solving the equation (12) by taking its gradient and setting it equal to zero, and using the update equation of the inverse of the correlation matrix, equation (7), then

$$\begin{aligned} -\frac{1}{t_1} \left[ \frac{-\tau_n}{\mu_n + \tau_n} + \alpha \right]^{-1} \left[ \frac{\tau_n}{(\mu_n + \tau_n)^2} \right] - \\ \frac{1}{t_2} \left[ \frac{\tau_n}{\mu_n + \tau_n} - \beta \right]^{-1} \left[ \frac{-\tau_n}{(\mu_n + \tau_n)^2} \right] = 0 \end{aligned} \quad (13)$$

After some manipulations, the step-size is given by

$$\mu_n = \frac{\tau_n [(t_1 + t_2) - (t_1 \alpha + t_2 \beta)]}{t_1 \alpha + t_2 \beta} \quad (14)$$

As this is also a modified version of the conventional RLS algorithm, the step-size needs to be kept inside the stability region,  $0 < \mu_n \leq 1$ . This can be accomplished in many ways, but probably the most simple one is to use a hard decision after the computation of the step-size: If it is off the limits then set it to the closest boundary value.

The resulting Leakage QN algorithm is summarized in Table II.

TABLE II  
MODIFIED LEAKAGE QN (MLQN) ALGORITHM

$$\begin{aligned}
 e_n &= d_n - \mathbf{x}_n^T \mathbf{w}_{n-1} \\
 \mathbf{t}_n &= \mathbf{R}_{n-1}^{-1} \mathbf{x}_n \\
 \tau_n &= \mathbf{x}_n^T \mathbf{t}_n \\
 \hat{\mu}_n &= \frac{\tau_n [(t_1+t_2) - (t_1\alpha + t_2\beta)]}{t_1\alpha + t_2\beta} \\
 \text{if } 0 \leq \hat{\mu}_n \leq 1 & \\
 \quad \mu_n &= \hat{\mu}_n \\
 \text{else} & \\
 \quad \mu_n &= \begin{cases} 0 & \text{for } \hat{\mu}_n < 0 \\ 1 & \text{for } \hat{\mu}_n > 1 \end{cases} \\
 \text{endif} & \\
 \mathbf{R}_n^{-1} &= \frac{1}{\mu_n} \left[ \mathbf{R}_{n-1}^{-1} - \frac{\mathbf{t}_n \mathbf{t}_n^T}{\mu_n + \tau_n} \right] \\
 \mathbf{w}_n &= \mathbf{w}_{n-1} + \frac{e_n}{\mu_n + \tau_n} \mathbf{t}_n
 \end{aligned}$$

#### IV. LEAKAGE LS ALGORITHM

An algorithm with good convergence performance and low misadjustment was presented in [4]. The cost function that is minimized is given by

$$\xi_n = \|\mathbf{w}_n - \mathbf{w}_{n-1}\|_{\mathbf{R}_{n-1}}^2 + \mu_n [d_n - \mathbf{w}_n^T \mathbf{x}_n]^2 \quad (15)$$

The solution of this minimization problem results in the following recursive update equations of the coefficient vector and the inverse correlation matrix:

$$\mathbf{w}_n = \mathbf{w}_{n-1} + \mu_n e_n \mathbf{R}_{n-1}^{-1} \mathbf{x}_n \quad (16)$$

and

$$\mathbf{R}_n^{-1} = \mathbf{R}_{n-1}^{-1} - \frac{\mu_n \mathbf{t}_n \mathbf{t}_n^T}{1 + \mu_n \tau_n} \quad (17)$$

respectively.

Using the QN constraint in this algorithm, one can find a new leakage quasi-Newton algorithm, that will be called herein leakage LS algorithm, where the step-size reinforces the equality  $\mathbf{x}_n^T \mathbf{R}_n^{-1} \mathbf{x}_n = 1/2$ . Solving this equality we have

$$\mathbf{x}_n^T \left\{ \mathbf{R}_{n-1}^{-1} - \frac{\mu_n \mathbf{t}_n \mathbf{t}_n^T}{1 + \mu_n \tau_n} \right\} \mathbf{x}_n = 1/2 \quad (18)$$

or, equivalently,

$$\left[ \frac{\tau_n (1 + \mu_n \tau_n) - \mu_n \tau_n^2}{1 + \mu_n \tau_n} \right] = 1/2 \quad (19)$$

which results in

$$\mu_n = \frac{2\tau_n - 1}{\tau_n} \quad (20)$$

The algorithm is presented in Table III.

#### V. MODIFIED LEAKAGE LS ALGORITHM

The same constraints used in the leakage QN algorithm can be employed in the leakage LS algorithm, allowing better

TABLE III  
LEAKAGE LS (LLS) ALGORITHM

$$\begin{aligned}
 e_n &= d_n - \mathbf{x}_n^T \mathbf{w}_{n-1} \\
 \mathbf{t}_n &= \mathbf{R}_{n-1}^{-1} \mathbf{x}_n \\
 \tau_n &= \mathbf{x}_n^T \mathbf{t}_n \\
 \mu_n &= \frac{2\tau_n - 1}{\tau_n} \\
 \mathbf{R}_n^{-1} &= \mathbf{R}_{n-1}^{-1} - \frac{\mu_n \mathbf{t}_n \mathbf{t}_n^T}{1 + \mu_n \tau_n} \\
 \mathbf{w}_n &= \mathbf{w}_{n-1} + \mu_n e_n \mathbf{R}_{n-1}^{-1} \mathbf{x}_n
 \end{aligned}$$

control over the inverse correlation matrix behavior. The problem to be solved now is

$$\begin{aligned}
 \min_{\mathbf{w}_n} \quad & \|\mathbf{w}_n - \mathbf{w}_{n-1}\|_{\mathbf{R}_{n-1}}^2 + \mu_n [d_n - \mathbf{x}_n^T \mathbf{w}_n]^2 \\
 \text{s.t.} \quad & \begin{cases} \mathbf{x}_n^T \mathbf{R}_{n-1}^{-1} \mathbf{x}_n > \alpha \\ \mathbf{x}_n^T \mathbf{R}_{n-1}^{-1} \mathbf{x}_n < \beta \end{cases} \quad (21)
 \end{aligned}$$

The coefficient vector and the inverse of correlation matrix update equations are the same, equations (16) and (17), respectively. However, the step-size is computed by minimizing a similar constraint problem as in equation (12). The solution becomes

$$\begin{aligned}
 -\frac{1}{t_1} \left[ \frac{-\tau_n}{1 + \mu_n \tau_n} + \alpha \right]^{-1} \left[ \frac{(\tau_n) \tau_n}{(1 + \mu_n \tau_n)^2} \right] - \\
 \frac{1}{t_2} \left[ \frac{\tau_n}{1 + \mu_n \tau_n} - \beta \right]^{-1} \left[ \frac{(-\tau_n) \tau_n}{(1 + \mu_n \tau_n)^2} \right] = 0 \quad (22)
 \end{aligned}$$

Solving equation (22) for  $\mu_n$ , we have

$$\mu_n = \frac{\tau_n (t_1 + t_2) - (\alpha t_1 + \beta t_2)}{\tau_n (\alpha t_1 + \beta t_2)} \quad (23)$$

The modified leakage LS algorithm can control the bounds of the correlation matrix, and its inverse, through the variable step-size. The resulting algorithm was first presented in [6] and is shown in Table IV.

TABLE IV  
MODIFIED LEAKAGE LS (MLLS) ALGORITHM

$$\begin{aligned}
 e_n &= d_n - \mathbf{x}_n^T \mathbf{w}_{n-1} \\
 \mathbf{t}_n &= \mathbf{R}_{n-1}^{-1} \mathbf{x}_n \\
 \tau_n &= \mathbf{x}_n^T \mathbf{t}_n \\
 \hat{\mu}_n &= \frac{\tau_n (t_1 + t_2) - (\alpha t_1 + \beta t_2)}{\tau_n (\alpha t_1 + \beta t_2)} \\
 \text{if } 0 \leq \hat{\mu}_n \leq 1 & \\
 \quad \mu_n &= \hat{\mu}_n \\
 \text{else} & \\
 \quad \mu_n &= \begin{cases} 0 & \text{for } \hat{\mu}_n < 0 \\ 1 & \text{for } \hat{\mu}_n > 1 \end{cases} \\
 \text{endif} & \\
 \mathbf{R}_n^{-1} &= \mathbf{R}_{n-1}^{-1} - \frac{\mu_n \mathbf{t}_n \mathbf{t}_n^T}{1 + \mu_n \tau_n} \\
 \mathbf{w}_n &= \mathbf{w}_{n-1} + \mu_n e_n \mathbf{R}_{n-1}^{-1} \mathbf{x}_n
 \end{aligned}$$

## VI. SIMULATIONS

Simulations were carried out to test and compare convergence speed and Mean Square Error (MSE) of the algorithms. Another objective is to verify the improved robustness of the algorithms using the variable step-size over the conventional RLS algorithm, which can diverge in situations of non-persistently exciting signals or quantization errors.

In our simulations, an unknown system was to be identified by the adaptive filter, and zero-mean white noise was added to the output signal of the system to form the reference signal. The signal-to-noise ratio (SNR) was constant and equal to 40dB for all simulations.

The unknown system was an FIR filter with transfer function given by

$$H(z) = 1 + z^{-1} + z^{-2} + z^{-3} + z^{-4} \quad (24)$$

normalized to unitary gain.

For each different algorithm, a sufficient order adaptive filter was used and the simulations were carried out until convergence was reached. Two different input signals were used to the unknown system and to the adaptive filter: white Gaussian and sinusoid signals. The MSE curves shown in the figures were obtained after averaging over an ensemble of 200 simulations. The user-defined parameters of algorithms in Tables II and IV were chosen so that the algorithms meet the misadjustment of the LQN algorithm, and the initialization parameter  $\gamma = 1$  for convenience.

### A. White Gaussian Signal

For this case, the input signal was generated using a White Gaussian Process (WGN) with unitary power. The user-defined parameters for the modified algorithms are:  $t_1 = t_2 = 50$ ,  $\alpha = 0.15$  and  $\beta = 0.75$  for the MLQN algorithm; and,  $t_1 = t_2 = 50$ ,  $\alpha = 0.01$  and  $\beta = 0.025$ , for the MLSS algorithm. The forgetting factor of the RLS algorithm for this experiment was set to  $\lambda = 0.85$ . Figure 1 shows the MSE learning curve for the 5 algorithms. As can be seen in the figure, the MLLS algorithm has the best convergence speed, whereas the LQN and MLQN algorithms converge at similar speed as the RLS algorithm. This is a good achievement for the LQN algorithm that does not need or use user parameters defined. The LLS algorithm has the worst behavior, and being an algorithm that does not have user-defined parameters it can't be better adjusted.

### B. Non-Persistently Exciting Signal

A sinusoid signal with sample frequency of  $1/N$  was used. Using the same forgetting factor for the RLS algorithm,  $\lambda = 0.85$ , the parameters for the MLLS algorithm are still the same as for the WGN input signal case to yield similar misadjustment. In this case, the parameters  $\alpha$  and  $\beta$  for the MLQN algorithm need to be set closer to each other to avoid divergence:  $t_1 = t_2 = 50$ ,  $\alpha = 0.0001$  and  $\beta = 0.001$ . Figure 2 presents the MSE convergence of the algorithms, where it can be seen that the modified algorithms converged and the RLS and the LQN algorithms diverged. Also, the MLLS algorithm has the best convergence performance followed by

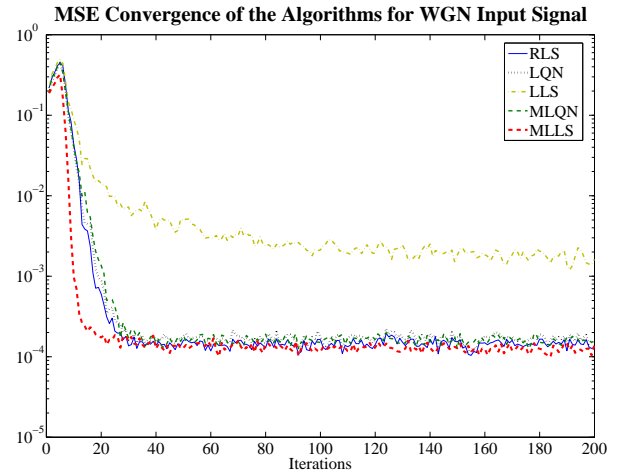


Fig. 1. Learning curves of the algorithms for WGN signal.

the MLQN algorithm. In this figure, the LLS algorithm was not present due to its poor performance, similar to the one obtained in the WGN case.

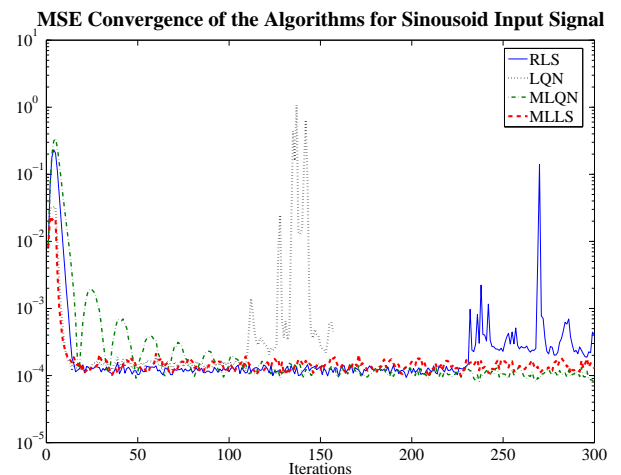


Fig. 2. Learning curves of the algorithms for non-persistently exciting signal.

## VII. CONCLUSIONS

This paper presented the development of variable step-size algorithms for both quasi-Newton and Least Squares algorithms with a leakage approach. The four algorithms obtained imposed constraints to the correlation matrix, so that once it is initialized positive definite, it remains in this state independently of the input signal statistics or presence of quantization errors.

The variable step-size equations were derived from equality and inequality constraints, where the logarithmic barrier was used to help solving the inequality optimization problems. In the algorithms using equality constraints, a variable step-size that does not need user-defined parameters was obtained, which gives a different level of robustness to the algorithm, specially in non-stationary environments. However, simulation results showed that the LLS algorithm presents bad performance at high SNR levels.

The modified algorithms that were derived from inequality-constrained optimization problems presented the best results in the simulations carried out. Even in situations of non-persistently exciting signals the algorithms still converged fast. However, these algorithms need user-defined parameters that may influence the algorithm's convergence performance.

In general, the most appropriate algorithm depends on the application. If the system does not present many variations, the MLQN and MLSS algorithms can be used. If one knows that the signal in the system has persistence of excitation and wants to avoid tuning the algorithm for the best performance, then the LQN algorithm can be employed. An article with the theoretical convergence analysis of the algorithms is in development.

#### REFERÊNCIAS

- [1] M. L. R. de Campos and A. Antoniou, "A new quasi-newton adaptive filtering algorithm," in *IEEE Transactions on Circuits and Systems*, November 1997, vol. 44, pp. 924–934.
- [2] P. S. R. Diniz, *Adaptive Filtering: Algorithms and Practical Implementation*, Kluwer Academic Publishers, Norwell, Massachusetts, 2002.
- [3] K. H. Afkhamie, Z.-Q. Luo, and K. M. Wong, "Adaptive linear filtering using interior point optimization techniques," in *IEEE Transactions on Signal Processing*, June 2000, vol. 48, pp. 1637–1647.
- [4] Fabiano T. Castoldi and M. L. R. de Campos, "A leakage quasi-newton adaptation algorithm," in *XXV Simpósio Brasileiro de Telecomunicações*, September 2007.
- [5] S. Boyd and L. Vandenberghe, *Convex Optimization*, Cambridge University Press, New York, New York, 2004.
- [6] Fabiano T. Castoldi, "Desenvolvimento de algoritmos adaptativos usando a descrição de distúrbio mínimo com restrições," M.S. thesis, Universidade Federal do Rio de Janeiro, 2007.