Recursive Least-Squares Algorithms for Echo Cancellation An Overview and Open Issues

Camelia Elisei-Iliescu and Constantin Paleologu

University Politehnica of Bucharest
Bucharest, Romania
Email: pale@comm.pub.ro

***Abstract*—The recursive least-squares (RLS) algorithm is very popular in many applications of adaptive filtering, especially due to its fast convergence rate. However, the computational complexity of this algorithm represents a major limitation in some applications that involve high length adaptive filters, like echo cancellation. Moreover, the specific features of this application require good tracking capabilities and double-talk robustness for the adaptive algorithm, which further implies an optimization process on its parameters. In case of most RLS-based algorithms, the performance can be controlled in terms of two main parameters, i.e., the forgetting factor and the regularization term. The goal of this paper is to outline the influence of these parameters on the overall performance of the RLS algorithms and to present several solutions to control their behavior, taking into account the specific requirements of echo cancellation application.**

Keywords–Adaptive filters; Echo cancellation; Recursive least-squares (RLS) algorithm.

I. INTRODUCTION

The recursive least-squares (RLS) algorithm [1][2] is one of the most popular adaptive filters. As compared to the normalized least-mean-square (NLMS) algorithm [1][2], the RLS offers a superior convergence rate especially for highly correlated input signals. Of course, there is a price to pay for this advantage, which is an increase in the computational complexity. For this reason, it is not very often involved in echo cancellation [3][4], where high length adaptive filters (e.g., hundreds of coefficients) are required.

The performance of the RLS algorithm is mainly controlled by two important parameters, i.e., the forgetting factor and the regularization term. Similar to the attributes of the step-size from the NLMS-based algorithms, the performance of RLS-type algorithms in terms of convergence rate, tracking, misadjustment, and stability depends on the forgetting factor [1][2]. The classical RLS algorithm uses a constant forgetting factor (between 0 and 1) and needs to compromise between the previous performance criteria. When the forgetting factor is very close to one, the algorithm achieves low misadjustment and good stability, but its tracking capabilities are reduced [5]. A small value of the forgetting factor improves the tracking but increases the misadjustment, and could affect the stability of the algorithm [6]. Motivated by these aspects, a number of variable forgetting factor RLS (VFF-RLS) algorithms have been developed, e.g., [7]–[10] (and references therein).

It should be mentioned that in the context of system identification (like in echo cancellation), where the output of the unknown system is corrupted by another signal (which is usually an additive noise), the goal of the adaptive filter is not to make the error signal goes to zero, because this will introduce noise in the adaptive filter. The objective instead is to recover the “corrupting signal” from the error signal of the adaptive filter after this one converges to the true solution. This was the approach behind the VFF-RLS algorithm proposed in [9], which is analyzed in Section II.

As compared to the forgetting factor, the regularization parameter has been less addressed in the literature. Apparently, it is required in matrix inversion when this matrix is ill conditioned, especially in the initialization stage of the algorithm. However, its role is of great importance in practice, since regularization is a must in all ill-posed problems (like in adaptive filtering), especially in the presence of additive noise [11]–[14]. Consequently, in Section III, we focus on the regularized RLS algorithm [2]. Following the development from [12], a method to select an optimal regularization parameter is presented, so that the algorithm could behave well in all noisy conditions. Since the value of this parameter is related to the echo-to-noise ratio (ENR), a simple and practical way to estimate the ENR in practice is also presented, which leads to a variable-regularized RLS (VR-RLS) algorithm.

The simulation results (presented in Section IV) are performed in the context of echo cancellation and support the theoretical findings. Finally, the conclusions are outlined in Section V, together with some open issues related to future works.

II. VARIABLE FORGETTING FACTOR RLS ALGORITHM

Let us consider a system identification problem (like in echo cancellation), where the desired signal at the discrete-time index n is obtained as

$$\begin{aligned} d(n) &= \mathbf{h}^T\mathbf{x}(n) + v(n) \\ &= y(n) + v(n), \end{aligned} \tag{1}$$

where $\mathbf{h} = [\ h_0 \quad h_1 \quad \cdots \quad h_{L-1}\ ]^T$ is the impulse response (of length L) of the system that we need to identify (i.e., the echo path), superscript ^{T} denotes transpose of a vector or a matrix,

$$\mathbf{x}(n) = [\ x(n) \quad x(n-1) \quad \cdots \quad x(n-L+1)\ ]^T \tag{2}$$

is a vector containing the most recent L samples of the zero-mean input signal $x(n)$ (i.e., the far-end signal), $v(n)$ is a zero-mean additive noise signal [which is independent of $x(n)$], and $y(n)$ represents the output of the unknown system (i.e., the echo signal). In the context of echo cancellation, the output of the echo path could be also corrupted by the

 ISBN: 978-1-61208-546-3

near-end speech (besides the background noise), which is usually known as the double-talk scenario [3][4]. The main objective is to estimate or identify $\mathbf{h}$ with an adaptive filter $\widehat{\mathbf{h}}(n) = \left[\ \widehat{h}_0(n) \quad \widehat{h}_1(n) \quad \cdots \quad \widehat{h}_{L-1}(n)\ \right]^T$.

Using the previous notation we may define the a priori error signal as

$$\begin{aligned} e(n) &= d(n) - \mathbf{x}^T(n)\widehat{\mathbf{h}}(n-1) \\ &= \mathbf{x}^T(n)\left[\mathbf{h} - \widehat{\mathbf{h}}(n-1)\right] + v(n). \end{aligned} \tag{3}$$

In this context, the relations that define the classical RLS algorithm are:

$$\mathbf{k}(n) = \frac{\mathbf{P}(n-1)\mathbf{x}(n)}{\lambda + \mathbf{x}^T(n)\mathbf{P}(n-1)\mathbf{x}(n)}, \tag{4}$$

$$\widehat{\mathbf{h}}(n) = \widehat{\mathbf{h}}(n-1) + \mathbf{k}(n)e(n), \tag{5}$$

$$\mathbf{P}(n) = \frac{1}{\lambda}\left[\mathbf{P}(n-1) - \mathbf{k}(n)\mathbf{x}^T(n)\mathbf{P}(n-1)\right], \tag{6}$$

where λ ($0 < \lambda \le 1$) is the exponential forgetting factor, $\mathbf{k}(n)$ is the Kalman gain vector, $\mathbf{P}(n)$ is the estimate of the inverse of the input correlation matrix, and $e(n)$ is the a priori error signal defined in (3). The a posteriori error signal can be defined using the adaptive filter coefficients at time n, i.e.,

$$\begin{aligned} \varepsilon(n) &= d(n) - \mathbf{x}^T(n)\widehat{\mathbf{h}}(n) \\ &= \mathbf{x}^T(n)\left[\mathbf{h} - \widehat{\mathbf{h}}(n)\right] + v(n), \end{aligned} \tag{7}$$

Using (3) and (5) in (7), it results in

$$\varepsilon(n) = e(n)\left[1 - \mathbf{x}^T(n)\mathbf{k}(n)\right]. \tag{8}$$

In the framework of system identification, it is desirable to recover the system noise from the error signal [5]. Consequently, we can impose the condition:

$$E\left[\varepsilon^2(n)\right] = \sigma_v^2, \tag{9}$$

where $E[\cdot]$ denotes mathematical expectation and $\sigma_v^2 = E\left[v^2(n)\right]$ is the power of the system noise. Furthermore, using (9) in (8) and taking (4) into account, it finally results in

$$E\left\{\left[1 - \frac{\theta(n)}{\lambda(n) + \theta(n)}\right]^2\right\} = \frac{\sigma_v^2}{\sigma_e^2(n)}, \tag{10}$$

where $\theta(n) = \mathbf{x}^T(n)\mathbf{P}(n-1)\mathbf{x}(n)$. In (10), we assumed that the input and error signals are uncorrelated, which is true when the adaptive filter has started to converge to the true solution. We also assumed that the forgetting factor is deterministic and time dependent. By solving the quadratic equation (10), it results a variable forgetting factor

$$\lambda(n) = \frac{\sigma_\theta(n)\sigma_v}{\sigma_e(n) - \sigma_v}, \tag{11}$$

where $E\left[\theta^2(n)\right] = \sigma_\theta^2(n)$. In practice, the variance of the error signal can be recursively estimated based on

$$\widehat{\sigma}_e^2(n) = \alpha\widehat{\sigma}_e^2(n-1) + (1-\alpha)e^2(n), \tag{12}$$

where $\alpha = 1 - 1/(KL)$, with $K \ge 1$. The variance of $\theta(n)$ is evaluated in a similar manner, i.e.,

$$\widehat{\sigma}_\theta^2(n) = \alpha\widehat{\sigma}_\theta^2(n-1) + (1-\alpha)\theta^2(n). \tag{13}$$

The estimate of the noise power, $\widehat{\sigma}_v^2(n)$ [which should be used in (11) from practical reasons], can be estimated in different ways, e.g., [9][15][16].

Theoretically, $\sigma_e(n) \ge \sigma_v$ in (11). Compared to the NLMS algorithm (where there is the gradient noise, so that $\sigma_e(n) > \sigma_v$), the RLS algorithm with $\lambda(n) \approx 1$ leads to $\sigma_e(n) \approx \sigma_v$. In practice (since power estimates are used), several situations have to be prevented in (11). Apparently, when $\widehat{\sigma}_e(n) \le \widehat{\sigma}_v$, it could be set $\lambda(n) = \lambda_{\max}$, where $\lambda_{\max}$ is very close or equal to 1. But this could be a limitation, because in the steady-state of the algorithm $\widehat{\sigma}_e(n)$ varies around $\widehat{\sigma}_v$. A more reasonable solution is to impose that $\lambda(n) = \lambda_{\max}$ when

$$\widehat{\sigma}_e(n) \le \rho\widehat{\sigma}_v, \tag{14}$$

with $1 < \rho \le 2$. Otherwise, the forgetting factor of the VFF-RLS algorithm [9] is evaluated as

$$\lambda(n) = \min\left[\frac{\widehat{\sigma}_\theta(n)\widehat{\sigma}_v(n)}{\zeta + |\widehat{\sigma}_e(n) - \widehat{\sigma}_v(n)|}, \lambda_{\max}\right], \tag{15}$$

where the small positive constant ζ prevents a division by zero. Before the algorithm converges or when there is an abrupt change of the system, $\widehat{\sigma}_e(n)$ is large as compared to $\widehat{\sigma}_v(n)$; thus, the parameter $\lambda(n)$ from (15) takes low values, providing fast convergence and good tracking. When the algorithm converges to the steady-state solution, $\widehat{\sigma}_e(n) \approx \widehat{\sigma}_v(n)$ [so that condition (14) is fulfilled] and $\lambda(n)$ is equal to $\lambda_{\max}$, providing low misadjustment. It can be noticed that the mechanism that controls the forgetting factor is very simple and not expensive in terms of multiplications and additions.

III. Variable Regularized RLS Algorithm

In this section, a different version of the RLS algorithm is presented, which allow us to outline the importance of the regularization parameter. Let us consider the regularized least-squares criterion:

$$J(n) = \sum_{i=0}^{n} \lambda^{n-i}\left[d(i) - \widehat{\mathbf{h}}^T(n)\mathbf{x}(i)\right]^2 + \delta\left\|\widehat{\mathbf{h}}(n)\right\|_2, \tag{16}$$

where λ is the same exponential forgetting factor, δ is the regularization parameter, and $\|\cdot\|_2$ is the ℓ_2 norm. From (16), the update of the regularized RLS algorithm [2] results in

$$\widehat{\mathbf{h}}(n) = \widehat{\mathbf{h}}(n-1) + \left[\widehat{\mathbf{R}}_{\mathbf{x}}(n) + \delta\mathbf{I}_L\right]^{-1}\mathbf{x}(n)e(n), \tag{17}$$

where

$$\begin{aligned} \widehat{\mathbf{R}}_{\mathbf{x}}(n) &= \sum_{i=0}^{n} \lambda^{n-i}\mathbf{x}(i)\mathbf{x}^T(i) \\ &= \lambda\widehat{\mathbf{R}}_{\mathbf{x}}(n-1) + \mathbf{x}(n)\mathbf{x}^T(n) \end{aligned} \tag{18}$$

is an estimate of the correlation matrix of $\mathbf{x}(n)$ at time n, $\mathbf{I}_L$ is the identity matrix of size $L \times L$, and

$$\begin{aligned} e(n) &= d(n) - \widehat{\mathbf{h}}^T(n-1)\mathbf{x}(n) \\ &= d(n) - \widehat{y}(n) \end{aligned} \tag{19}$$

is the a priori error signal as defined in (3); the signal $\widehat{y}(n)$ represents the output of the adaptive filter, which should be an estimate of the echo signal. We will assume that the matrix $\widehat{\mathbf{R}}_{\mathbf{x}}(n)$ has full rank, although it can be very ill conditioned. As a result, if there is no noise, regularization is not really required; however, the more the noise, the larger should be the value of δ.

Summarizing, the regularized RLS algorithm is defined by the relations (17)–(19). In the following, we present one reasonable way to find the regularization parameter δ. It can be noticed that the update equation of the regularized RLS can be rewritten as [12]

$$\widehat{\mathbf{h}}(n) = \mathbf{Q}(n)\widehat{\mathbf{h}}(n-1) + \widetilde{\mathbf{h}}(n), \tag{20}$$

where

$$\mathbf{Q}(n) = \mathbf{I}_L - \left[\widehat{\mathbf{R}}_{\mathbf{x}}(n) + \delta\mathbf{I}_L\right]^{-1} \mathbf{x}(n)\mathbf{x}^T(n) \tag{21}$$

and

$$\widetilde{\mathbf{h}}(n) = \left[\widehat{\mathbf{R}}_{\mathbf{x}}(n) + \delta\mathbf{I}_L\right]^{-1} \mathbf{x}(n)d(n) \tag{22}$$

is the correctiveness component of the algorithm, which depends on the new observation $d(n)$. In this context, we can notice that $\mathbf{Q}(n)$ does not depend on the noise signal and $\mathbf{Q}(n)\widehat{\mathbf{h}}(n-1)$ in (20) can be seen as a good initialization of the adaptive filter. In fact, (22) is the solution of the noisy linear system of L equations:

$$\left[\widehat{\mathbf{R}}_{\mathbf{x}}(n) + \delta\mathbf{I}_L\right] \widetilde{\mathbf{h}}(n) = \mathbf{x}(n)d(n). \tag{23}$$

Let us define

$$\widetilde{e}(n) = d(n) - \widetilde{\mathbf{h}}^T(n)\mathbf{x}(n), \tag{24}$$

the error signal between the desired signal and the estimated signal obtained from the filter optimized in (22). Consequently, we could find δ in such a way that the expected value of $\widetilde{e}^2(n)$ is equal to the variance of the noise, i.e.,

$$E\left[\widetilde{e}^2(n)\right] = \sigma_v^2. \tag{25}$$

This is reasonable if we want to attenuate the effects of the noise in the estimator $\widetilde{\mathbf{h}}(n)$.

For the sake of simplicity, let us assume that $x(n)$ is stationary and white. Apparently, this assumption is quite restrictive, even if it was widely used in many developments in the context of adaptive filtering [1][2]. However, the resulting VR-RLS algorithm will still use the full matrix $\widehat{\mathbf{R}}_{\mathbf{x}}(n)$ and, consequently, it will inherit the good performance feature of the RLS family in case of correlated inputs. In this case and for n large enough (also considering that the forgetting factor λ is on the order of $1 - 1/L$), we have

$$\begin{aligned}\left[\widehat{\mathbf{R}}_{\mathbf{x}}(n) + \delta\mathbf{I}_L\right] &\approx \left[\frac{\sigma_x^2}{1-\lambda} + \delta\right]\mathbf{I}_L \\ &\approx \left[L\sigma_x^2 + \delta\right]\mathbf{I}_L\end{aligned} \tag{26}$$

and $\mathbf{x}^T(n)\mathbf{x}(n) \approx L\sigma_x^2$, where $\sigma_x^2 = E\left[x^2(n)\right]$ is the variance of the input signal. Next, from (1), we can define the echo-to-noise ratio (ENR) as

$$\text{ENR} = \frac{\sigma_y^2}{\sigma_v^2}, \tag{27}$$

where $\sigma_y^2 = E\left[y^2(n)\right]$ is the variance of $y(n)$. Developing (25) and based on the previous approximations, we obtain the quadratic equation:

$$\delta^2 - 2\frac{L\sigma_x^2}{\text{ENR}}\delta - \frac{\left(L\sigma_x^2\right)^2}{\text{ENR}} = 0, \tag{28}$$

with the obvious solution:

$$\begin{aligned}\delta &= \frac{L\left(1+\sqrt{1+\text{ENR}}\right)}{\text{ENR}}\sigma_x^2 \\ &= \beta\sigma_x^2,\end{aligned} \tag{29}$$

where

$$\beta = \frac{L\left(1+\sqrt{1+\text{ENR}}\right)}{\text{ENR}} \tag{30}$$

is the normalized regularization parameter of the RLS algorithm.

As we can notice from (29), the regularization parameter δ depends on three elements, i.e., the length of the adaptive filter, the variance of the input signal, and the ENR. In most applications, the first two elements (L and σ_x^2) are known, while the ENR can be estimated. Using a proper evaluation of the ENR, the algorithm should own good robustness features against the additive noise.

Let us assume that the adaptive filter has converged to a certain degree, so that we can use the approximation

$$y(n) \approx \widehat{y}(n). \tag{31}$$

Hence,

$$\sigma_y^2 \approx \sigma_{\widehat{y}}^2, \tag{32}$$

where $\sigma_{\widehat{y}}^2 = E\left[\widehat{y}^2(n)\right]$. Since the output of the unknown system and the noise can be considered uncorrelated, (1) can be expressed in terms of power estimates as

$$\sigma_d^2 = \sigma_y^2 + \sigma_v^2, \tag{33}$$

where $\sigma_d^2 = E\left[d^2(n)\right]$. Using (32) in (33), we obtain

$$\sigma_v^2 \approx \sigma_d^2 - \sigma_{\widehat{y}}^2. \tag{34}$$

The power estimates can be evaluated in a recursive manner as

$$\widehat{\sigma}_d^2(n) = \alpha\widehat{\sigma}_d^2(n-1) + (1-\alpha)d^2(n), \tag{35}$$
$$\widehat{\sigma}_{\widehat{y}}^2(n) = \alpha\widehat{\sigma}_{\widehat{y}}^2(n-1) + (1-\alpha)\widehat{y}^2(n), \tag{36}$$

where $\alpha = 1 - 1/(KL)$, with $K \geq 1$ [similar to (12) and (13)]. Therefore, based on (32), (34), and (35), an estimation of the ENR is obtained as

$$\widehat{\text{ENR}}(n) = \frac{\widehat{\sigma}_{\widehat{y}}^2(n)}{\left|\widehat{\sigma}_d^2(n) - \widehat{\sigma}_{\widehat{y}}^2(n)\right|}, \tag{37}$$

so that the variable regularization parameter results in

$$\begin{aligned}\delta(n) &= \frac{L\left[1+\sqrt{1+\widehat{\text{ENR}}(n)}\right]}{\widehat{\text{ENR}}(n)}\sigma_x^2 \\ &= \beta(n)\sigma_x^2,\end{aligned} \tag{38}$$

where

$$\beta(n) = \frac{L\left[1+\sqrt{1+\widehat{\text{ENR}}(n)}\right]}{\widehat{\text{ENR}}(n)} \tag{39}$$

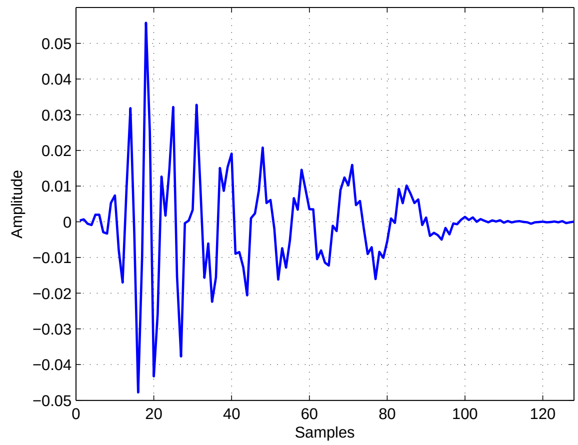

Figure 1. Impulse response used in simulations (the fourth echo path from G168 Recommendation [17]).

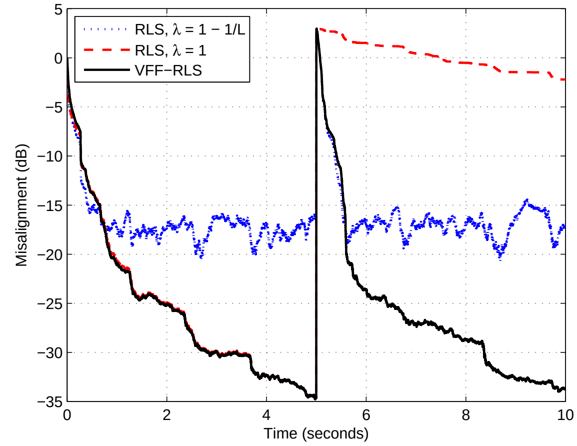

Figure 2. Misalignment of the RLS algorithm (using different constant values of the forgetting factor) and VFF-RLS algorithm in a single-talk scenario, including echo path change.

is the variable normalized regularization parameter. Consequently, based on (38), we obtain a variable-regularized RLS (VR-RLS) algorithm, with the update:

$$\widehat{\mathbf{h}}(n) = \widehat{\mathbf{h}}(n-1) + \left[\widehat{\mathbf{R}}_{\mathbf{x}}(n) + \delta(n)\mathbf{I}_L\right]^{-1}\mathbf{x}(n)e(n), \quad (40)$$

where $\widehat{\mathbf{R}}_{\mathbf{x}}(n)$ is recursively evaluated according to (18) and $\delta(n)$ is computed based on (35)–(38).

Finally, some practical issues should be outlined. The absolute values in (37) prevent any minor deviations (due to the use of power estimates) from the true values, which can make the denominator negative. It is a non-parametric algorithm, since all the parameters in (37) are available. Also, good robustness against the additive noise variations is expected. The main drawback is due to the approximation in (32). This assumption will be biased in the initial convergence phase or when there is a change of the unknown system. Concerning the initial convergence, we can use a constant regularization parameter δ in the first steps of the algorithm (e.g., in the first L iterations).

IV. Simulation Results

Let us consider a network echo cancellation scenario, in the framework of G168 Recommendation [17]. The echo path is depicted in Figure 1; it is the fourth impulse response (of length $L = 128$) from the above recommendation. The sampling rate is 8 kHz. All adaptive filters used in the experiments have the same length as the echo path. The far-end signal (i.e., the input signal) is a speech signal. The output of the echo path is corrupted by an independent white Gaussian noise with 20 dB ENR. An echo path change scenario is some experiments (in order to evaluate the tracking capabilities of the algorithms), by shifting the impulse response to the right by 8 samples in the middle of simulation. The performance measure is the normalized misalignment (in dB) evaluated as

$$\mathrm{Mis}(n) = 20\log_{10}\frac{\left\|\mathbf{h}(n) - \widehat{\mathbf{h}}(n)\right\|_2}{\left\|\mathbf{h}(n)\right\|_2}. \quad (41)$$

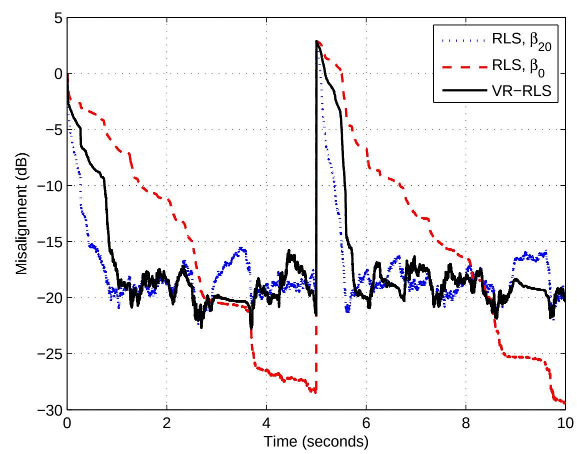

Figure 3. Misalignment of the regularized RLS algorithm (using different constant values of the regularization parameter) and VR-RLS algorithm in a single-talk scenario, including echo path change.

In the first the experiment, the performance of the VFF-RLS algorithm (presented in Section II) is evaluated, as compared to the classical RLS algorithm defined in (4)–(6), which uses different constant values of the forgetting factor. A single-talk case is considered and the echo path changes in the middle of simulation. It can be noticed in Figure 2 that the VFF-RLS algorithm achieves the same initial misalignment as the RLS with its maximum forgetting factor, but it tracks as fast as the RLS with the smaller forgetting factor. As expected, the classical RLS algorithm using constant forgetting factors has to compromise between these performance criteria, i.e., the larger the value of λ, the better the misalignment level but worse the tracking capability.

Next, the performance of the VR-RLS algorithm (from Section III) is investigated, as compared to the regularized RLS algorithm defined in (17)–(19), using different constant values of the regularization parameter. In real-world applications, the value of ENR is not available. However, based on

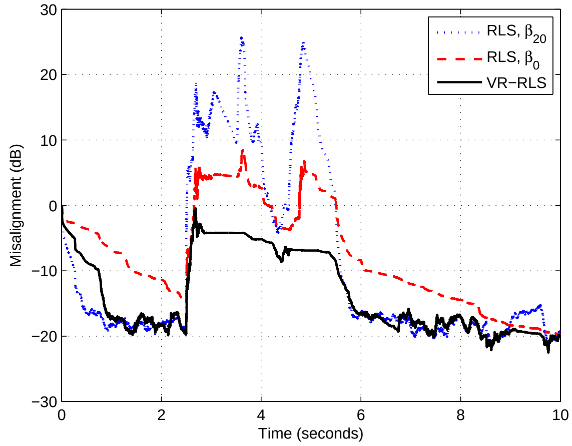

Figure 4. Misalignment of the regularized RLS algorithm (using different constant values of the regularization parameter) and VR-RLS algorithm in a double-talk scenario.

(30), we can determine the values of the optimal normalized regularization parameter of the RLS algorithm for different cases; for example, let us consider two values of the ENR, i.e., 20 dB (the true one) and 0 dB. Using appropriate notation, we obtain $\beta_{20} = 14.14$ and $\beta_0 = 309.01$, respectively. In the next set of experiments, we compare the regularized RLS algorithm using these constant regularization parameters with the VR-RLS algorithm. The constant forgetting factor is set to $\lambda = 1 - 1/(3L)$ for all the algorithms.

In Figure 3, a single-talk scenario is considered and an echo path change is introduced in the middle of the simulation. It can be noticed that the VR-RLS algorithm behaves similarly to the RLS algorithm using the constant parameter β_{20}, which is associated to the value of the true ENR. Also, it can be noticed that a larger value of the normalized regularization parameter (β_0) improves the misalignment but affects the convergence rate and tracking.

In Figure 4, a double-talk scenario [3][4] is considered. The near-end speech appears between time 2.5 and 5 seconds, so that the signal $v(n)$ is now non-stationary, since it contains both noise and speech. It is clear that the VR-RLS algorithm is more robust in this case as compared to the regularized RLS using constant values of β. It should be outlined that we do not use any double-talk detector (DTD) [3][4] with the VR-RLS algorithm, which is the regular approach in a double-talk situation. Therefore, the VR-RLS algorithm owns good robustness features against double-talk, which is an important gain in practice.

V. Conclusions and Perspectives

The RLS algorithms are very appealing due to their fast convergence rate. In this paper, we have focused on the main parameters that control the performance of these algorithms, i.e., the forgetting factor and the regularization term. In order to achieve a better compromise between the performance criteria (i.e., convergence and tracking versus misadjustment and robustness), these parameters could be controlled. In this context, the solutions presented in Sections II and III led to the VFF-RLS and VR-RLS algorithms, respectively.

The experiments were performed in the context of echo cancellation, which is a very challenging system identification problem. According to the simulation results, the VFF-RLS and VR-RLS algorithms perform very well as compared to their classical counterparts (which use constant values of the key parameters). On the other hand, the complexity of the RLS-based algorithms is $O(L^2)$, which represents a problematic issue for high values of L (like in echo cancellation). The alternative is to combine these VFF and VR methods with low-complexity versions of the RLS algorithm, e.g., [18]. Also, another interesting issue to address in future works could be a combination between the VFF and VR approaches, in order to inherent the advantages of both methods.

Acknowledgment

This work was supported by UEFISCDI Romania under Grant PN-II-RU-TE-2014-4-1880.

References

[1] S. Haykin, *Adaptive Filter Theory*. Fourth Edition, Upper Saddle River, NJ: Prentice-Hall, 2002.

[2] A. H. Sayed, *Adaptive Filters*. New York, NY: Wiley, 2008.

[3] J. Benesty, T. Gaensler, D. R. Morgan, M. M. Sondhi, and S. L. Gay, *Advances in Network and Acoustic Echo Cancellation*. Berlin, Germany: Springer-Verlag, 2001.

[4] C. Paleologu, J. Benesty, and S. Ciochină, *Sparse Adaptive Filters for Echo Cancellation*. Morgan & Claypool Publishers, 2010.

[5] S. Ciochină, C. Paleologu, J. Benesty, and A. A. Enescu, "On the influence of the forgetting factor of the RLS adaptive filter in system identification," in *Proc. IEEE ISSCS*, 2009, pp. 205–208.

[6] S. Ciochină, C. Paleologu, and A. A. Enescu, "On the behaviour of RLS adaptive algorithm in fixed-point implementation," in *Proc. IEEE ISSCS*, 2003, pp. 57–60.

[7] S. Song, J. S. Lim, S. J. Baek, and K. M. Sung, "Gauss Newton variable forgetting factor recursive least squares for time varying parameter tracking," *Electronics Lett.*, vol. 36, pp. 988–990, May 2000.

[8] S.-H. Leung and C. F. So, "Gradient-based variable forgetting factor RLS algorithm in time-varying environments," *IEEE Trans. Signal Processing*, vol. 53, pp. 3141–3150, Aug. 2005.

[9] C. Paleologu, J. Benesty, and S. Ciochină, "A robust variable forgetting factor recursive least-squares algorithm for system identification," *IEEE Signal Processing Lett.*, vol. 15, pp. 597–600, 2008.

[10] Y. J. Chu and S. C. Chan, "A new local polynomial modeling-based variable forgetting factor RLS algorithm and its acoustic applications," *IEEE/ACM Trans. Audio, Speech, Language Processing*, vol. 23, pp. 2059–2069, Nov. 2015.

[11] P. C. Hansen, *Rank-Deficient and Discrete Ill-Posed Problems: Numerical Aspects of Linear Inversion*. Philadelphia, PA: SIAM, 1998.

[12] J. Benesty, C. Paleologu, and S. Ciochină, "Regularization of the RLS algorithm," *IEICE Trans. Fundamentals*, vol. E94-A, pp. 1628–1629, Aug. 2011.

[13] Y. V. Zakharov and V. H. Nascimento, "Sparse sliding-window RLS adaptive filter with dynamic regularization," in *Proc. EUSIPCO*, 2016, pp. 145–149.

[14] C. Elisei-Iliescu, C. Stanciu, C. Paleologu, J. Benesty, C. Anghel, and S. Ciochină, "Robust variable regularized RLS algorithms," in *Proc. IEEE HSCMA*, 2017, pp. 171–175.

[15] C. Paleologu, S. Ciochină, and J. Benesty, "Variable step-size NLMS algorithm for under-modeling acoustic echo cancellation," *IEEE Signal Processing Lett.*, vol. 15, pp. 5–8, 2008.

[16] M. A. Iqbal and S. L. Grant, "Novel variable step size NLMS algorithms for echo cancellation," in *Proc. IEEE ICASSP*, 2008, pp. 241–244.

[17] *Digital Network Echo Cancellers*, ITU-T Rec. G.168, 2002.

[18] Y. V. Zakharov, G. P. White, and J. Liu, "Low-complexity RLS algorithms using dichotomous coordinate descent iterations," *IEEE Trans. Signal Processing*, vol. 56, pp. 3150–3161, July 2008.

 ISBN: 978-1-61208-546-3