

# Improved Variable Forgetting Factor Recursive Least Square Algorithm

Felix Albu

Department of Electronics & Telecommunications  
Valahia University of Targoviste  
Targoviste, Romania  
felix.albu@valahia.ro

**Abstract**—In this paper an improved variable forgetting factor recursive least square (IVFF-RLS) algorithm is proposed. The forgetting factor is adjusted according to the square of a time-averaging estimate of the autocorrelation of a priori and a posteriori errors. The proposed algorithm has fast convergence, and robustness against variable background noise, near-end signal variations and echo path change. The simulation results indicate the superior performances of IVFF-RLS when compared to the RLS and VFF-RLS algorithms.

**Keywords**—recursive least squares, variable forgetting factor, system identification, adaptive control, echo cancellation.

## I. INTRODUCTION

The adaptive algorithms are widely used in applications such as adaptive control, system identification etc. In system identification problem, an adaptive filter is used to identify an unknown system, (e.g. the echo path in case of echo cancellation systems [1]). Typically, the least mean square (LMS) or normalized LMS (NLMS) are used. Unfortunately, their convergence is very slow and better algorithms are needed. A variable step size (VSS) version (NPVSS-NLMS) proposed in [2] proved to be a more reliable solution in case of near-end signal variations, including double-talk. The affine projection algorithm (APA) [3] and its variable step size or fast versions (e.g. [4-7] and the references therein) offer superior convergence rate as compared to NLMS, but they are still insufficient for many applications [1]. It is well known that the recursive least square algorithm has the fastest convergence, especially for highly correlated input signals [1], [8-9]. Its computational complexity is also high ( $O(L^2)$ ,  $L$  being the filter length). The forgetting factor  $\lambda$  plays an important role in the behavior of the RLS algorithm in terms of convergence, misalignment and stability. It is fixed in the classical RLS algorithm with values between 0 and 1. It is known that, if  $\lambda$  is closer to one, the algorithm has not only good stability and fast convergence, but also reduced tracking ability. The tracking abilities can be improved by lowering the value of  $\lambda$ , but the speed convergence and stability of the algorithm suffers too. In order to meet these conflicting requirements several variable forgetting factor (VFF) recursive least squares algorithms have been developed (e.g. [10-12] and

the references therein).

In this paper, we propose an improved VFF-RLS algorithm (IVFF-RLS) suitable for system identification applications. The proposed approach takes into account the square of a time-averaging estimate of the autocorrelation of a priori and a posteriori errors. A similar approach has been used in [13] to derive the non-parametric variable step size NLMS algorithm for transversal filters (NVS-NLMS). It was also used in [14] in order to derive the non-parametric variable step size AP algorithm (NVSS-APA).

The paper is organized as follows. Section 2 introduces the classical RLS, followed by the derivation of the proposed IVFF-RLS. The simulation results are presented in Section 3. Finally, Section 4 concludes this work.

## II. THE PROPOSED ALGORITHM

The unknown system is identified using an adaptive filter. The system model is shown in Fig.1. Both the unknown system and the adaptive filter have finite impulse responses, defined by the real-valued vectors  $\mathbf{h} = [h_0 \ h_1 \ \dots \ h_{L-1}]^T$  and  $\hat{\mathbf{h}}(n) = [\hat{h}_0(n) \ \hat{h}_1(n) \ \dots \ \hat{h}_{L-1}(n)]^T$ , respectively, where superscript  $T$  denotes transposition and  $n$  is the time index;  $L$  is the length of the echo path.

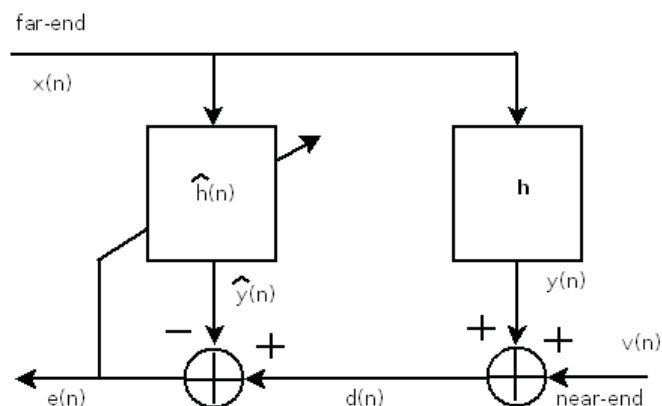


Fig. 1 System model

The signal  $x(n)$  is the far-end speech which goes through the acoustic impulse response  $\mathbf{h}$ , resulting the echo signal,  $y(n)$ .

This signal is picked up by the microphone together with the near-end signal  $v(n)$ , resulting the microphone signal  $d(n)$ . The near-end signal can contain both the background noise,  $w(n)$ , and the near-end speech,  $u(n)$ . The output of the adaptive filter,  $\hat{y}(n)$ , provides a replica of the echo, which will be subtracted from the microphone signal. The input signal vector is  $\mathbf{x}(n) = [x(n), x(n-1), \dots, x(n-L+1)]^T$ . The RLS algorithm [1] is defined by the following relations:

$$e(n) = d(n) - \hat{\mathbf{h}}^T(n-1)\mathbf{x}(n), \quad (1)$$

$$\mathbf{k}(n) = \frac{\mathbf{P}(n-1)\mathbf{x}(n)}{\lambda + \mathbf{x}^T(n)\mathbf{P}(n-1)\mathbf{x}(n)}, \quad (2)$$

$$\hat{\mathbf{h}}(n) = \hat{\mathbf{h}}(n-1) + \mathbf{k}(n)e(n), \quad (3)$$

$$\mathbf{P}(n) = \frac{1}{\lambda} \left[ \mathbf{P}(n-1) - \mathbf{k}(n)\mathbf{x}^T(n)\mathbf{P}(n-1) \right], \quad (4)$$

where  $\lambda$  is the forgetting factor,  $\mathbf{k}(n)$  is the Kalman gain vector and  $\mathbf{P}(n)$  is the inverse of the input correlation matrix  $\Phi(n)$  defined by

$$\Phi(n) = \sum_{i=1}^n \lambda^{n-i} \mathbf{x}(i)\mathbf{x}^T(i). \quad (5)$$

The a posteriori error [12] is defined as

$$\varepsilon(n) = d(n) - \hat{\mathbf{h}}^T(n)\mathbf{x}(n). \quad (6)$$

It results from (1), (3) and (6)

$$\varepsilon(n) = e(n) \left[ 1 - \mathbf{x}^T(n)\mathbf{k}(n) \right]. \quad (7)$$

Next, the adjustment of the forgetting factor according to the square of the square of the time-averaged estimation of the autocorrelation of  $\varepsilon(n)$  and  $e(n)$  is proposed. A similar approach was used to derive a non-parametric NLMS algorithm in [13]. Therefore, the goal is to find an expression for the forgetting factor  $\lambda(n)$  such that

$$E\{\varepsilon(n)e(n)\} = E\{v^2(n)\}, \quad (8)$$

where  $E\{\cdot\}$  denotes the mathematical expectation. The right side of (8) is the power of the system noise (i.e.  $E\{v^2(n)\} = \sigma_v^2$ ).

$$\begin{aligned} E\{\varepsilon(n)e(n)\} &= E\left\{e^2(n) \left[ 1 - \mathbf{x}^T(n)\mathbf{k}(n) \right] \right\} = \\ &= E\left\{e^2(n) \left[ 1 - \frac{\mathbf{x}^T(n)\mathbf{P}(n-1)\mathbf{x}(n)}{\lambda(n) + \mathbf{x}^T(n)\mathbf{P}(n-1)\mathbf{x}(n)} \right] \right\}. \end{aligned} \quad (9)$$

Using equations (8) and (9) after several mathematical manipulations it results:

$$\lambda(n) = \frac{\bar{q}(n)\sigma_v^2}{\sigma_e^2(n) - \sigma_v^2}, \quad (10)$$

where  $q(n) = \mathbf{x}^T(n)\mathbf{P}(n-1)\mathbf{x}(n)$ , and  $E\{e^2(n)\} = \sigma_e^2(n)$  is the power of the a priori error signal. Like in [12] it was assumed that the input and error signals are uncorrelated and the forgetting factor is deterministic and time dependent. The estimated powers or averages are computed as follows

$$\hat{\sigma}_e^2(n) = \alpha \hat{\sigma}_e^2(n-1) + (1-\alpha)e^2(n), \quad (11)$$

$$\hat{\sigma}_v^2(n) = \beta \hat{\sigma}_v^2(n-1) + (1-\beta)e^2(n), \quad (12)$$

$$\bar{q}(n) = \alpha_0 \bar{q}(n-1) + (1-\alpha_0)q(n), \quad (13)$$

with  $\alpha_0$ ,  $\alpha = 1 - 1/(K_a L)$  and  $\beta = 1 - 1/(K_b L)$  being the weighting factors,  $K_a \geq 2$  and  $K_b > K_a$  respectively [11-12]. A superior limit has to be imposed on  $\lambda$  value, for cases when the denominator of (10) is close to zero. Like in [12] this limit is set to  $\lambda_{\max}$  (i.e.  $\lambda(n) = \lambda_{\max}$ ) when

$$\hat{\sigma}_e(n) \leq \gamma \sigma_v(n), \quad (14)$$

with  $1 < \gamma \leq 2$ . If condition (14) is not fulfilled, the forgetting factor of the proposed IVFF-RLS is computed as follows:

$$\lambda(n) = \min \left\{ \frac{\bar{q}(n)\sigma_v^2(n)}{\xi + |\sigma_e^2(n) - \sigma_v^2(n)|}, \lambda_{\max} \right\}, \quad (15)$$

where  $\xi$  is a small constant. It can be noticed that if  $\hat{\sigma}_e^2(n) \cong \sigma_v^2(n)$  (it happens around the steady-state solution),  $\lambda(n)$  goes to  $\lambda_{\max}$ . Also, lower  $\lambda(n)$  values are obtained if  $\hat{\sigma}_e^2(n) \gg \sigma_v^2(n)$ . Therefore, a faster convergence is expected at the initial convergence phase.

The VFF-RLS algorithm [12] uses a different formula for the forgetting factor:

$$\lambda(n) = \min \left\{ \frac{\sigma_q(n)\sigma_v(n)}{\xi + |\sigma_e(n) - \sigma_v(n)|}, \lambda_{\max} \right\}, \quad (16)$$

where

$$\hat{\sigma}_q^2(n) = \alpha \hat{\sigma}_q^2(n-1) + (1-\alpha)q^2(n). \quad (17)$$

The equations that define the proposed IVFF-RLS algorithm are summarized in Table 1.

TABLE I. THE IVFF-RLS ALGORITHM

Inputs:	$\alpha_0, \gamma, K_a, K_b, \xi, \lambda_{\max}, \delta$
Initialization:	$\hat{\mathbf{h}}(0) = \mathbf{0}, \bar{q}(0) = 10, \sigma_e^2(0) = 0, \sigma_v^2(0) = 0,$ $\mathbf{P}(0) = \delta^{-1} \mathbf{I}_L$
Loop:	$n=1, 2, \dots$
	$q(n) = \mathbf{x}^T(n) \mathbf{P}(n-1) \mathbf{x}(n)$
	$\bar{q}(n) = \alpha_0 \bar{q}(n-1) + (1-\alpha_0)q(n)$
	$e(n) = d(n) - \hat{\mathbf{h}}^T(n-1) \mathbf{x}(n)$
	$\hat{\sigma}_e^2(n) = \alpha \hat{\sigma}_e^2(n-1) + (1-\alpha)e^2(n)$
	$\hat{\sigma}_v^2(n) = \beta \hat{\sigma}_v^2(n-1) + (1-\beta)e^2(n)$
	$\lambda(n) = \begin{cases} \lambda_{\max} & \text{if } \hat{\sigma}_e(n) \leq \gamma \hat{\sigma}_v(n), \text{ else} \\ \min \left\{ \frac{\bar{q}(n) \hat{\sigma}_v^2(n)}{\xi +  \hat{\sigma}_e^2(n) - \hat{\sigma}_v^2(n) }, \lambda_{\max} \right\} \end{cases}$
	$\mathbf{k}(n) = \frac{\mathbf{P}(n-1) \mathbf{x}(n)}{\lambda(n) + \mathbf{x}^T(n) \mathbf{P}(n-1) \mathbf{x}(n)}$
	$\hat{\mathbf{h}}(n) = \hat{\mathbf{h}}(n-1) + \mathbf{k}(n) e(n)$
	$\mathbf{P}(n) = \frac{1}{\lambda(n)} \left[ \mathbf{P}(n-1) - \mathbf{k}(n) \mathbf{x}^T(n) \mathbf{P}(n-1) \right]$

Both VFF-RLS and IVFF-RLS algorithms are only slightly less complex than RLS. The additional complexity of VFF-RLS over RLS is given by the variable forgetting factor formula (16), the comparison of (14) and the estimated powers (equations (11)-(12) and (17)). Therefore the VFF-RLS algorithm has two comparisons, 12 multiplications, 5 additions, one division, and three square roots more than the RLS algorithm.

The additional complexity of IVFF-RLS over RLS is given by the variable forgetting factor formula (15), the comparison of (14) and the equations (11)-(13). Therefore the VFF-RLS algorithm has two comparisons, 10 multiplications, 5 additions, and one division more than the RLS algorithm.

It can be noticed that IVFF-RLS has three square roots and two multiplications less than VFF-RLS. Therefore, IVFF-RLS is slightly less complex than VFF-RLS. Other advantages of the new IVFF-RLS algorithm over VFF-RLS algorithm will be revealed in the Simulation sections.

### III. SIMULATIONS

The simulations were performed in a system identification context. The length of the adaptive filter is set to 64 coefficients and the echo path is the first impulse response from ITU G.168 [15]. The impulse response of the acoustic echo path is plotted in Fig. 2(a) (the sampling rate is 8 kHz). The far-end signal,  $x(n)$ , is either a white Gaussian sequence, an AR(1) process generated by filtering a white Gaussian noise through a first-order system  $1/(1-0.9z^{-1})$ , or a speech sequence (Fig. 2(b)). For the double-talk scenario, the near-end speech  $u(n)$  is plotted in Fig. 2(c). An independent white Gaussian noise signal  $w(n)$  is added to the echo signal  $y(n)$ , with 30 dB signal-to-noise ratio (SNR) for most of the experiments. The value of the  $\xi$  in the denominator of (15) and (16) is  $\xi = 10^{-8}$ ,  $\lambda_{\max} = 0.999999$  and  $\alpha_0 = 7/8$ .

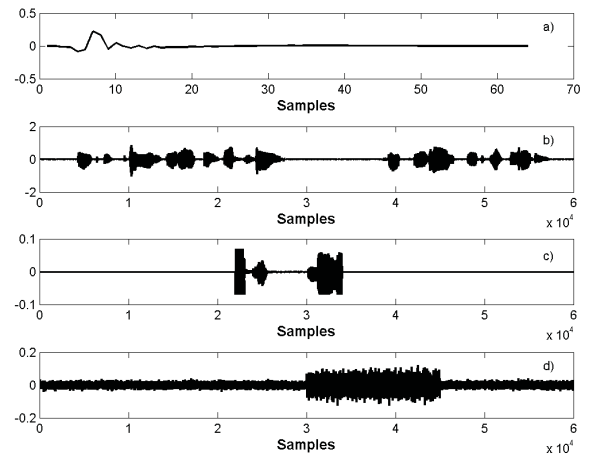


Fig. 2 a) Measured room acoustic impulse response; b) Far-end speech signal; c) Near-end speech signal; d) the variable background noise, SNR decreases from 30 dB to 20dB between iterations 30000 and 45000.

In the first case, the input signal is a white Gaussian noise (Fig. 3). The impulse response was delayed by 12 samples after 20000 iterations and by 24 samples after 40000 iterations, in order to verify the tracking abilities of the investigated algorithms. It can be noticed from Fig. 3 that IVFF-RLS has superior tracking ability to both VFF-RLS and RLS. Both IVFF-RLS and VFF-RLS achieve significantly lower misalignment than RLS. The misalignment difference between IVFF-RLS and VFF-RLS increases more after the second larger echo path disturbance than following the first echo path change. Figure 4 shows the  $\lambda(n)$  values of VFF-RLS and IVFF-RLS algorithms. It can be noticed that for most of iterations  $\lambda(n)$  values of both algorithms are very similar.

In the steady state regions  $\lambda(n)$  is very close or equal with  $\lambda_{\max}$ .

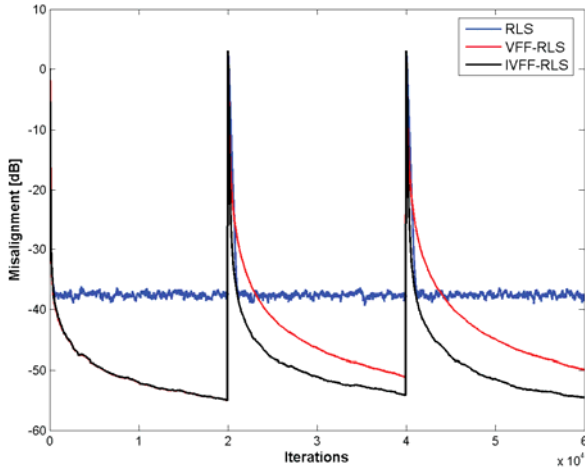


Fig. 3 Misalignment of the RLS algorithm ( $L = 64$ ,  $\lambda = 1 - 1/(3L)$ ), VFF-RLS and IVFF-RLS algorithms. The impulse response changes at iterations 20000 and 40000. The input signal is a white Gaussian noise,  $K_a = 2$ ,  $K_b = 10$ ,  $\lambda_{\max} = 0.999999$ ,  $\gamma = 1.5$ , and SNR = 30dB.

However, slight differences appear when the system is disturbed. Figure 4c shows a zoomed portion around the second echo path change. In this region, the  $\lambda(n)$  values of the IVFF-RLS algorithm have more values smaller than  $\lambda_{\max}$  than VFF-RLS algorithm. This fact explains the superior tracking abilities and the lower misalignment of IVFF-RLS. The same conclusions drawn from Fig. 4 regarding the  $\lambda(n)$  values of the VFF-RLS and IVFF-RLS algorithms were observed in the subsequent simulations.

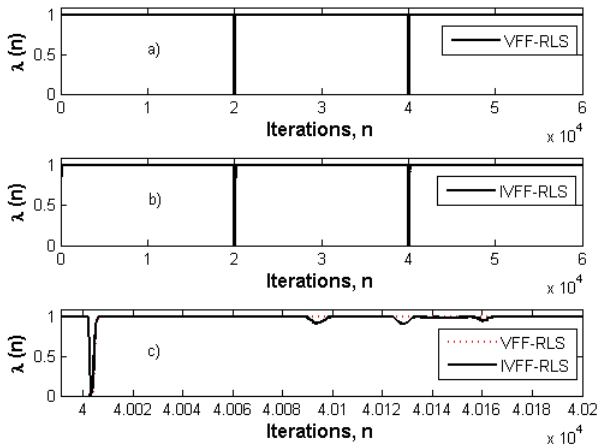


Fig.4 The forgetting factors  $\lambda(n)$  for: a) the VFF-RLS algorithm; b) the IVFF-RLS algorithm; c) zoom around the second echo path change.

In Fig. 5, the input signal is the AR(1) process (pole at 0.9), while the other conditions are the same as those of Fig. 3. The same conclusions as above are derived regarding the

convergence and tracking abilities. IVFF-RLS has the fastest convergence and the best tracking ability among the considered algorithms. However, the misalignment difference between IVFF-RLS and VFF-RLS is reduced in comparison with that obtained while using white excitation input signal.

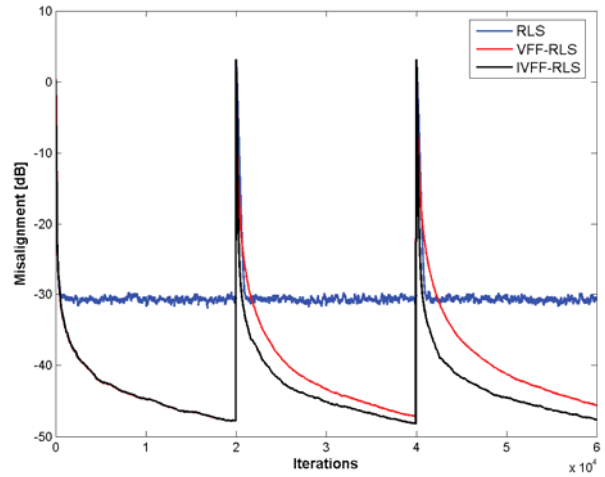


Fig. 5. Misalignment of the RLS ( $L = 64$ ,  $\lambda = 1 - 1/(3L)$ ), VFF-RLS and IVFF-RLS algorithms. The impulse response changes at iterations 20000 and 40000. The input signal is a colored Gaussian noise,  $K_a = 2$ ,  $K_b = 10$ ,  $\lambda_{\max} = 0.999999$ ,  $\gamma = 1.5$ , and SNR = 30dB.

In Fig. 6 the background noise SNR decreases from 30 dB to 20 dB between iterations 30000 and 45000 (Fig. 2d). The other conditions are the same from Fig. 3. It can be noticed that RLS is the most affected by the increased system noise. Also, it can be seen from Figs. 3 and 6 that the IVFF-RLS algorithm behavior is better than that of the VFF-RLS algorithm in case of echo path change, while it is similar when there is a variable background noise situation.

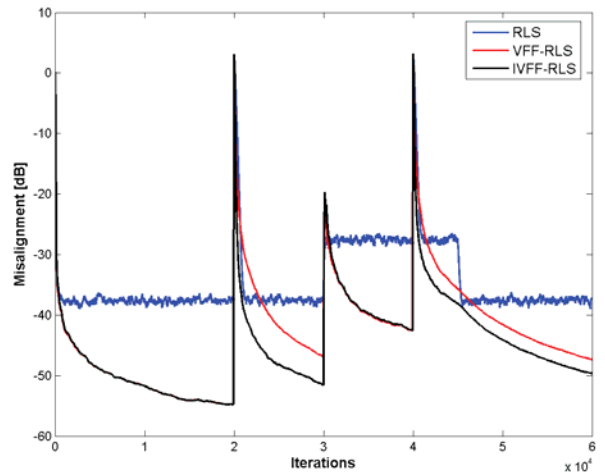


Fig. 6. Misalignment of the RLS, VFF-RLS and IVFF-RLS algorithms. SNR decreases from 30 dB to 20dB between iterations 30000 and 45000. Other conditions are the same as in Fig. 3.

The result using the speech signal (Fig. 2b) as an input signal is presented in Fig. 7. A speech signal corrupts the output of the unknown system between iterations 22000 and 34000 (Fig. 2c), such that the input signal-to-corrupting speech ratio is 12dB. In terms of performance, both VFF based RLS algorithms outperform the RLS algorithm. In terms of robustness, the proposed IVFF-RLS algorithm outperforms RLS and VFF-RLS algorithms. In case of using a speech signal, as for the variable background noise, the echo path change disturbs more the VFF based algorithms than the near end signal.

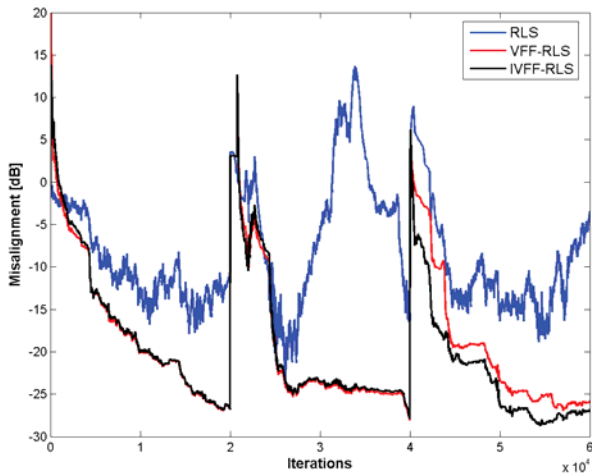


Fig. 7. Misalignment of the RLS ( $L = 64$ ,  $\lambda = 1 - 1/(10L)$ ), VFF-RLS and IVFF-RLS algorithms. The input signal is speech,  $K_a = 6$ ,  $K_b = 18$ . A speech corrupts the output of the unknown system between iterations 22000 and 34000. Other conditions are the same as in Fig. 3.

In the above simulations, when identical parameters are used, the convergence behavior of VFF-RLS and IVFF-RLS is similar for the first 20000 samples (before the first echo path change). However, the IVFF-RLS is more robust than RLS or VFF-RLS algorithms when the system is disturbed by a variable noise background, near-end signal or echo path change. The proposed IVFF-RLS algorithm has good tracking ability and low misalignment for both stationary and non-stationary input signals. It is also robust against different system noise variations. Its behavior can be adapted to various applications and input signals by varying  $\alpha_0$ ,  $\gamma$ ,  $K_a$ , and  $K_b$  parameters. The numerical complexity of all RLS-type algorithms (including IVFF-RLS) limits their use in acoustic echo cancellation systems, but IVFF-RLS could be useful if implemented in subbands. An implementation using the logarithmic number system [16] is also envisaged as a future option.

#### IV. CONCLUSIONS

In this paper the IVFF-RLS algorithm is proposed for system identification. The simulation results indicate the superior performances of the IVFF-RLS algorithm for both stationary and non-stationary input signals in comparison with RLS and VFF-RLS algorithms. It is shown that IVFF-RLS is slightly less complex than VFF-RLS and has better tracking ability, especially in case of echo path changes. Therefore, the IVFF-RLS algorithm could be a suitable choice in system identification applications.

#### ACKNOWLEDGMENT

This work was supported by a grant of the Romanian National Authority for Scientific Research, CNCS-UEFISCDI, project number PN-II-ID-PCE-2011-3-0097.

#### REFERENCES

- [1] A. Sayed, Adaptive filters, John Wiley & Sons, 2008
- [2] J. Benesty, H. Rey, L. Rey Vega, and S. Tressens, A nonparametric VSS NLMS algorithm, *IEEE Signal Process. Lett.*, 13, (10), pp. 581–584, 2006.
- [3] K. Ozeki and T. Umeda, An adaptive filtering algorithm using an orthogonal projection to an affine subspace and its properties, *Electron. Commun. Jpn.* 67-A., (5), pp. 19–27, 1984.
- [4] S. L. Gay and S. Tavathia, The fast affine projection algorithm, in *Proc. IEEE ICASSP*, vol. 5, pp. 3023–3026, 1995
- [5] F. Albu, H.K. Kwan, Fast block exact Gauss-Seidel pseudo affine projection algorithm, *Electronics Letters*, pp. 1451-1453, Vol. 40, (22), 2004
- [6] F. Albu, C. Kotropoulos, Modified Gauss-Seidel affine projection algorithm for acoustic echo cancellation, in *Proc. of ICASSP 2005*, pp. 121-124, 2005.
- [7] F. Albu, C. Paleologu, J. Benesty, “A Variable Step Size Evolutionary Affine Projection Algorithm”, in *Proc. of ICASSP 2011*, pp. 429-432, 2011.
- [8] J. Benesty and T. Gänslér, “New insights into the RLS algorithm,” *EURASIP J. Applied Signal Processing*, vol. 3, pp. 331 – 339, 2004.
- [9] G. O. Glentis, K. Berberidis, S. Theodoridis, “Efficient least squares adaptive algorithms for FIR transversal filtering: a unified view”, *IEEE Signal Processing Magazine*, pp. 13-41, July 1999
- [10] 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 Letters*, 36, (11), pp. 988–990, 2000.
- [11] S.-H. Leung and C. F. So, Gradient-based variable forgetting factor RLS algorithm in time-varying environments, *IEEE Trans. Signal Process.*, 53, (8), pp. 3141–3150, 2005.
- [12] C. Paleologu, J. Benesty, S. Ciochina, A robust variable forgetting factor recursive least-squares algorithm for system identification, *IEEE Signal Process. Lett.*, 15, pp. 597-600, 2008.
- [13] J.-C. Liu and H.-R. Li, A nonparametric variable step-size NLMS algorithm, *Appl. Math. Comput.*, 217, (14), pp. 7365-7371, 2011.
- [14] F. Albu, C. Paleologu, S. Ciochina “New variable step size affine projection algorithms”, in *Proc. of IEEE COMM 2012*, Bucharest, Romania, pp.63-66, 2012.
- [15] Digital Network Echo Cancellers, ITU-T Rec. G.168, 2002.
- [16] F. Albu, C. Paleologu, A Recursive Least Square algorithm for Active Noise Control based on the Gauss-Seidel Method, *Proc. of ICECS*, pp.830-833,2008.