Title	A robust statistics based adaptive lattice-ladder filter in impulsive noise
Author(s)	Zou, YX; Chan, SC; Ng, TS
Citation	The 2000 IEEE International Symposium on Circuits and Systems, Geneva, Switzerland, 28-31 May 2000. In Historic Title IEEE International Symposium on Circuits and Systems Proceedings, 2000, v. 3, p. 539-542
Issued Date	2000
URL	http://hdl.handle.net/10722/46172
Rights	©1999 IEEE. Personal use of this material is permitted. However, permission to reprint/republish this material for advertising or promotional purposes or for creating new collective works for resale or redistribution to servers or lists, or to reuse any copyrighted component of this work in other works must be obtained from the IEEE.

A Robust Statistics Based Adaptive Lattice-Ladder Filter In Impulsive Noise

Yue-Xian Zou, Shing-Chow Chan and Tung-Sang Ng

Department of Electrical and Electronic Engineering
The University of Hong Kong, Pokfulam Road, Hong Kong
{yxzou,scchan,tsng}@eee.hku.hk

ABSTRACT

In this paper, a new robust adaptive lattice-ladder filter for impulsive noise suppression is proposed. The filter is obtained by applying the non-linear filtering technique in [1] and the robust statistic approach to the gradient adaptive lattice filter. A systematic method is also developed to determine the corresponding threshold parameters for impulse suppression. Simulation results showed that the performance of the proposed algorithm is better than the conventional *RLS*, *N-RLS*, the gradient adaptive lattice normalised-*LMS* (*GAL-NLMS*), *RMN* and *ATNA* algorithms when the input and desired signals are corrupted by individual and consecutive impulses. The initial convergence, steady-state error, computational complexity and tracking capability of the proposed algorithm are also comparable to the conventional *GAL-NLMS* algorithm.

1. INTRODUCTION

Recently, there has been considerable interest in studying adaptive filtering algorithms that are robust to impulsive interference. Under such adverse condition, the performance of the conventional linear adaptive filters can deteriorate significantly. Nonlinear techniques are often employed to reduce the hostile effects of the impulsive noise. In the nonlinear *LMS* (*ATNA*) and nonlinear *RLS* (*N-RLS*) algorithms [2, 3], nonlinear clipping functions are used to limit the transient fluctuation of the estimation error in conventional adaptive filters caused by the impulses. The mixed-norm *LMS* (*RMN*) algorithm, proposed in [4], combats the impulsive noise in the desired signal by minimising a combination of L_1 and L_2 norms using the stochastic gradient method. All of these methods are not robust to impulses that appear at the input signal. In [5], a *RLS*-liked algorithm, called the *M*-estimate recursive least (*M-RLS*) algorithm, was proposed for impulsive noise suppression, by minimising an *M*-estimate cost function instead of the conventional mean square error. The *M-RLS* algorithm is more robust to the conventional *RLS*, *N-RLS*, *RMN* and *ATNA* algorithms when the input and desired signals are corrupted by individual and consecutive impulses. It is also more suitable to real-time processing than the Huber adaptive filter [6], which treats the filtering problem as a block fitting problem using general *M*-estimator (*GM*-estimator) ([7], pp.12). The Huber adaptive filter is not recursive and a system of nonlinear equation has to be solved in each iteration. The arithmetic complexity of the *M-RLS* algorithm, however, is still rather high ($O(N^2)$) compared with the *LMS* and gradient lattice algorithms ($O(N)$). In this paper, a robust lattice-ladder adaptive filter with $O(N)$ complexity, called the *RGAL-RNLMS* algorithm, is proposed. The *M*-estimate distortion measure in [5] is used to combat the impulses in the estimation error while the non-linear filtering technique in [1] is used to remove the impulses in the input signal.

2. OVERVIEW OF THE GRADIENT ADAPTIVE LATTICE ALGORITHM

Without loss of generality, the gradient adaptive lattice (*GAL*) filter is configured as an adaptive channel equalizer as shown in Fig.1. The signals $x(n)$, $\hat{d}(n)$ and $d(n)$ are the input, output and desired signals of the lattice-ladder filter, respectively. The *GAL* filter consists of an optimal minimum mean squares (*MMSE*) lattice predictor of order M to whiten the input signal $x(n)$ and a linear estimator of $d(n)$ from the prediction errors, $b_m(n)$. In the m-th stage of the lattice predictor, the forward and backward prediction errors, $f_m(n)$ and $b_m(n)$, are computed as follows

$$f_m(n) = f_{m-1}(n) - k_m^b(n)b_{m-1}(n-1),$$
$$b_m(n) = b_{m-1}(n-1) - k_m^f(n)f_{m-1}(n), \quad 1 \le m \le M,$$
$$f_0(n) = b_0(n) = x(n), \quad P_0^b(n-1) = x^2(n), \tag{1}$$

where $k_m^f(n)$ and $k_m^b(n)$ are the forward and backward reflection coefficients, respectively. When $x(n)$ is a stationary process, the forward and backward reflection coefficients are equal, i.e., $k_m^f(n) = k_m^b(n) = k_m(n)$. Minimising the cost function $J_m = E[|f_m(n)|^2 + |b_m(n)|^2]$ [8] with respect to $k_m(n)$ and using the *LMS* concept, the reflection coefficients in the gradient adaptive lattice (*GAL*) algorithm can be updated as

$$k_m(n+1) = k_m(n) + \mu_{k,m}(n)\{f_m(n)b_{m-1}(n-1) + b_m(n)f_{m-1}(n)\} \tag{2}$$

where $\mu_{k,m}(n)$ is the normalised step-size given by

$$\mu_{k,m}(n) = \frac{2\mu_{k,0}}{P_{m-1}^b(n) + \varepsilon},$$

and $P_m^b(n+1) = \beta P_m^b(n) + \frac{1}{2}(1-\beta)\left(f_m^2(n) + b_{m-1}^2(n-1)\right).$ \tag{3}

$\mu_{k,0}$ is a constant stepsize, β is the forgetting factor, $P_{m-1}^b(n)$ is the signal power of the input at the $m-th$ stage of the predictor, and ε is a small positive constant to prevent the instability of the algorithms when $P_{m-1}^b(n)$ is close to zero. The desired signal is estimated as

$$\hat{d}(n) = \sum_{m=0}^{M} w_m(n)b_m(n), \tag{4}$$

where $w_m(n)$ is the ladder gain or regression coefficient [9]. The lattice predictor serves as an adaptive orthogonal transform to decorrelate the input signal $x(n)$. The *LMS* algorithm can be used to update the weights $w_m(n)$ as follows

$$w_m(n+1) = w_m(n) + 2\mu_{w,m}(n)e(n)b_m(n), \quad e(n) = d(n) - \hat{d}(n). \tag{5}$$

where $m = 0, \cdots, M$, $\mu_{w,m}(n)$ is the normalised stepsize parameter, which can be calculated by (3) with different stepsize constant, $\mu_{w,0}$. Equations (1~5) constitute the gradient adaptive lattice normalised-*LMS* (*GAL-NLMS*) algorithm [8].

3. ROBUST ADAPTIVE LATTICE–LADDER FILTERING ALGORITHM

From (1) and (2), we can see that only one impulse appearing in the $x(n)$ will greatly fluctuate the $b_m(n)$, $f_m(n)$ and $k_m(n)$ at each lattice stage due to the order recursion property. An effective method is to suppress the impulses before they enter the lattice filter. In our work, the approach in [1] is adopted to filter the impulses in the input signal (section 3.3). Basically, the input signal is predicted from its past samples using a linear predictor. If the prediction error is greater than a given threshold, an impulse is said to be detected and the input sample will be replaced by its estimated value. The adverse effect of the impulses is therefore reduced. Furthermore, if $d(n)$ is corrupted by impulsive noise, it can be seen from (5) that $e(n)$ and $w_m(n)$ will fluctuate significantly, which requires many iterations to recover. In Section 3.1, we shall extend the robust statistic approach in [5] to stabilise the weight vector $w_m(n)$ in the *GAL* algorithm.

3.1. Robust Normalised *LMS* (*RNLMS*) Algorithm

Following the same concept in [5], the following M-estimate based cost function is used instead of the mean squares error:

$$\overline{\gamma}_M = E[\rho(e(n))] = E[\rho(d(n) - \hat{d}(n))] . \qquad (6)$$

where, $\rho(\cdot)$ is a robust M-estimate. Without loss of generality, we shall consider the Hampel's three parts redescending M-estimate function ([10], pp. 150) in this paper, which is defined as

$$\rho(e) = \begin{cases} e^2 / 2, & 0 < |e| < \xi \\ \xi|e| - \xi^2/2, & \xi \le |e| < \Delta_1 \\ \dfrac{\xi}{2}(\Delta_2 + \Delta_1) - \dfrac{\xi^2}{2} + \dfrac{\xi}{2}\dfrac{(|e| - \Delta_2)^2}{\Delta_2 - \Delta_1}, & \Delta_1 \le |e| < \Delta_2 \\ \dfrac{\xi}{2}(\Delta_2 + \Delta_1) - \dfrac{\xi^2}{2}, & \Delta_2 \le |e| \end{cases} \qquad (7)$$

The advantage of this M-estimate is that its first order derivative is continuous and it is a piecewise approximation of the maximum likelihood estimator when the input and additive noises are modelled as a mixture of Gaussian processes. As shown in Fig. 2, $\rho(\cdot)$ is an even real-valued function and it is quadratic when e is smaller than ξ. For larger values of e in the interval $[\xi, \Delta_1]$, the function is linear. For values greater than Δ_2, the function is equal to a constant. It becomes apparent that the M-estimator is capable of suppressing outliers with large amplitude. The threshold parameters ξ, Δ_1, and Δ_2 are used to control the degree of suppression of the outliers. The smaller the values of ξ, Δ_1, and Δ_2, the greater the suppression will be of the outliers. The threshold parameters are usually chosen according to the applications or estimated continuously. $\overline{\gamma}_M(n)$, as defined as (6), is therefore capable of smoothing out momentary fluctuation caused by the impulsive interference. Differentiating (6) with respect to w, one gets the following gradient vector of $\overline{\gamma}_M$

$$\nabla_w = \frac{\partial \overline{\gamma}_M}{\partial w} = \frac{\partial [E(\rho(e(n)))]}{\partial w} = E\left[\frac{\partial(\rho(e(n)))}{\partial(e(n))}\frac{\partial(e(n))}{\partial w}\right] = E\left[q(e(n))e(n)\frac{\partial(e(n))}{\partial w}\right],$$

$$q(e) = \frac{\partial \rho(e)}{\partial e} / e = \begin{cases} 1, & 0 < |e| < \xi \\ \xi \, \mathrm{sgn}(e) / e, & \xi \le |e| < \Delta_1 \\ \mathrm{sgn}(e)\left[(|e| - \Delta_2)\dfrac{\xi}{\Delta_1 - \Delta_2}\right] / e, & \Delta_1 \le |e| < \Delta_2 \\ 0, & \Delta_2 \le |e| \end{cases} \qquad (8)$$

is the weighting function. Using the instantaneous gradient vector, $\hat{\nabla}_w = q(e(n))e(n)\dfrac{\partial(e(n))}{\partial w}$, as the estimate of the gradient vector, ∇_w, one gets the weight update equation of the proposed robust normalised *LMS* (*RNLMS*) algorithm:

$$w_m(n+1) = w_m(n) + 2\mu_{w,m}(n)q(e(n))e(n)b_m(n) . \qquad (9)$$

The thresholds ξ, Δ_1 and Δ_2 will be estimated continuously so that the corrupted estimation error can be detected. The values of $q(e(n))$, as given by (8), will be set to zero. Hence the weight vector is unaffected by the impulses.

3.2. Parameter estimation

Though the distribution of the error signal $e(n)$ is in general unknown, it is assumed as Gaussian distributed with additive impulsive noise. By estimating the variance of $e(n)$ without the impulses, it is possible to detect and reject the impulsive noise in $e(n)$. More specifically, the probability of $|e(n)|$ greater than a given threshold T is given by [5]

$$\theta_T(n) = P_r\{|e(n)| > T\} = 1 - erf(\frac{T}{\sqrt{2}\hat{\sigma}_e(n)}) . \qquad (10)$$

where $erf(r) = \frac{2}{\sqrt{\pi}}\int_0^r e^{-x^2}dx$ is the error function and $\hat{\sigma}_e(n)$ is the estimated standard deviation of the estimation error. Using different threshold parameter T, we can detect the impulsive noise with different degrees of confidence. Let $\theta_\xi = P_r\{|e(n)| > \xi\}$, $\theta_{\Delta_1} = P_r\{|e(n)| > \Delta_1\}$, and $\theta_{\Delta_2} = P_r\{|e(n)| > \Delta_2\}$ be the probabilities that $|e(n)|$ is greater than ξ, Δ_1 and Δ_2, respectively. In our work, θ_ξ, θ_{Δ_1} and θ_{Δ_2} are chosen to be 0.05, 0.025 and 0.01, respectively, so that we have 95% confidence to down weight or reject the impact of the $e(n)$ when $|e(n)| > \xi$. Using these values for θ_ξ, θ_{Δ_1} and θ_{Δ_2}, the threshold parameters are determined to be

$$\xi = 1.96\hat{\sigma}_e(n), \Delta_1 = 2.24\hat{\sigma}_e(n), \Delta_2 = 2.576\hat{\sigma}_e(n) . \qquad (11)$$

A common estimate of $\hat{\sigma}_e(n)$ is $\hat{\sigma}_e^2(n) = \lambda_e\hat{\sigma}_e^2(n-1) + (1-\lambda_e)e^2(n)$ [3]. It is, however, not robust to impulses. In fact, one single large impulse can substantially increase the value of $\hat{\sigma}_e(n)$, and hence the values of ξ, Δ_1 and Δ_2. A more robust and complex estimate is the median absolute deviation from the median (*MAD*) ([10], pp.105). In this paper, a new recursive estimate for $\hat{\sigma}_e(n)$ is proposed:

$$\hat{\sigma}_e^2(n) = \lambda_e\hat{\sigma}_e^2(n-1) + 1.483\left(1 + \frac{5}{N_w - 1}\right)(1 - \lambda_e)\mathrm{med}(A_e(n)) . \qquad (12)$$

where $A_e(n) = \{e^2(n), \cdots, e^2(n - N_w + 1)\}$, N_w is the length of the estimation window, and λ_e is the forgetting factor. Due to the recursive nature of the estimation, the estimation window is of infinite length, giving rise to a more stable estimation against impulsive noise. The tracking ability of this estimation is also very good, as will be shown in the simulation section.

3.3. Robust Gradient Adaptive Lattice (RGAL) Algorithm

Based on the lattice predictor, the one-step-ahead prediction of the input signal $\hat{x}(n+1)$ and the prediction error $e_p(n)$ can be written as [11]:

$$\hat{x}(n+1) = -\sum_{i=0}^{M-1} k_{i+1}(n)b_i(n),$$

and $e_p(n) = x(n+1) + \sum_{i=0}^{M-1} k_{i+1}(n)b_i(n)$. (13)

To detect the impulses in $e_p(n)$, the parameter estimation technique described in section 3.2 is employed to estimate the variance of $e_p(n)$, with different forgetting factor λ_p and window length N_p. Consequently, the threshold parameters and the variance of $e_p(n)$, $\hat{\sigma}_{e_p}^2(n)$, can be computed by (11) and (12) with $e(n)$, $\hat{\sigma}_e^2(n)$, N_w and λ_e replaced, respectively, by $e_p(n)$, $\hat{\sigma}_{e_p}^2(n)$, N_p and λ_p. If $e_p(n)$ is greater than the estimated thresholds, an impulse is said to occur in the input signal, the input sample will be replaced by a corresponding weighted combination of the input sample and its predicted value as follows

$$x(n+1) = q(e_p(n))x(n+1) + (1 - q(e_p(n)))\hat{x}(n+1). \quad (14)$$

where $q(\cdot)$ is the weighting function defined in (8), and $\hat{x}(n+1)$ is the predicted value of $x(n+1)$ in (13). The differences between our approach and that in [1] are: i) the different weighting function is used. ii) The suppressing or filtering of the impulses in $x(n)$ is derived from the lattice-ladder filtering process. Therefore an additive pre-processor is not needed. The arithmetic complexity can be greatly reduced. In summary, the proposed adaptive algorithm, called robust gradient lattice and robust step-normalised LMS (RGAL-RNLMS) algorithm consists of equations (1-4, and 11-14). Compared to the conventional GAL-NLMS algorithm, the increase of the arithmetic complexity is mainly due to the M more multiplications in (13), and the median operations in (12), which require $O(N_w \log N_w)$ and $O(N_p \log N_p)$ operations, respectively.

4. SIMULATION RESULTS

In this section, the performances of the proposed RGAL-RNLMS algorithm, the conventional RLS [9], GAL-NLMS [8], N-RLS [3], RMN [4], and ATNA [2] algorithms in impulsive noise environments will be evaluated, using the channel equalization problem shown in Fig.1. $s(n)$ is a random binary sequence with zero mean and variance one. $c_1^* = [.2194,1,.2194]$ is the impulse response of the communication channel [9] with eigenvalue spread 6.0782. To evaluate the tracking performance of the algorithms, the coefficients of the channel are suddenly changed to

$c_2^* = [.35,1,.35]$ [8] at $n = 1024$. The mean square error (MSE) is used as the performance measure. $\eta_g(n)$ and $\eta_i(n)$ are respectively the additive Gaussian and impulsive noises. $\eta_i(n)$ is generated by the Gaussian-Bernoulli process [1, 4] with an arrival probability of $P_{ar} \approx 5 * 10^{-3}$ and variance ratio $\gamma^2 = 100$. In the simulation, the impulses in the desired signal occur at $n = 302,609,610,1301,1511 - 1515,1807$. On the other hand, the channel output signal is assumed to be corrupted by $\eta_g(n)$ and the impulses appearing at $n = 301,302,1309 - 1314,1807$. The positions of the impulses are fixed in all independent runs but their amplitudes are governed by zero mean Gaussian process. The signal to noise ratio, $SNR = 20\log_{10}(\sigma_x^2/\sigma_g^2)$, at the equaliser input is 30 dB, where σ_x^2 is the variance of the channel output signal. The parameters for all algorithms are summarised in Table 1. The MSE results in the presence of the impulses in the desired and the input signal are plotted in Fig.3 and 4, respectively. From Fig.3 (a) and Fig. 4 (a), it can be seen that the performance of the GAL-NLMS (line 2 with hexagrams) and RLS (line 3 with plus sign) algorithms is significantly affected by the impulses, especially for those impulses in the input signal. The initial convergence speed, tracking ability and steady error are identical for the GAL-NLMS and the RGALR-NLMS (line 1 with stars) algorithms. From Fig.3 (b) and Fig. 4 (b), the following are observed: i) The RGAL-RNLMS algorithm is more robust to impulses and has the best tracking performance than all the other algorithms considered. Its initial convergence and steady error are inferior to that of the N-RLS algorithm, because of the sub-optimal nature of the gradient lattice algorithm. ii) The RMN (line 3 with squares) algorithm is not robust to the impulses in the input signal. iii) The performance of the N-RLS (line 2 with circles) and ATNA (line 4 with diamonds) algorithms is significantly degraded in the presence of the consecutive impulses in the desired signal or impulses in the input signal.

5. CONCLUSION

In this paper, a new adaptive lattice-ladder filter for impulsive noise suppression is presented. The filter is obtained by applying the non-linear filtering technique in [1] and the robust statistic approach to the gradient adaptive lattice filter. A systematic procedure is also developed to determine the corresponding threshold parameters for impulse suppression. Simulation results showed that the performance of the proposed algorithm is better than the conventional RLS, N-RLS, GAL-NLMS, RMN and ATNA algorithms when the input and desired signals are corrupted by individual and consecutive impulses. The initial convergence, steady-state error, computational complexity and tracking capability of the RGAL-RNLMS algorithm are also comparable to the conventional GAL-NLMS algorithm.

6. REFERENCES

[1]   S. R. Kim and A. Efron, "Adaptive Robust Impulsive Noise Filtering," IEEE Transactions on Signal Processing, vol. 43, No. 8, pp. 1855-1866, 1995.

[2]   S. Koike, "Adaptive Threshold Nonlinear Algorithm for Adaptive Filters with Robustness Against Impulsive Noise," IEEE Transactions on Signal Processing, vol. 3, pp. 2391-2395, 1997.

[3]   J. F. Weng and S. H. Leung, "Adaptive Nonlinear RLS Algorithm for Robust Filtering in Impulse Noise," IEEE

International Symposium on Circuits and Systems (ISCAS'97), Hong Kong, vol. 4, pp. 2337 -2340, 1997.

[4] J. A. Chambers and A. Avlonitis, "A Robust Mixed-Norm (RMN) Adaptive Filter Algorithm," IEEE Signal Processing Letters, vol. 4, No. 2, pp. 46-48, 1997.

[5] Y. Zou, S. C. Chan, and T. S. Ng, "A Robust M-Estimate Adaptive Filter For Impulse Noise Suppression," IEEE International Conference on Acoustic Speech Signal Processing (ICASSP'99), Phoenix, Arizona, USA, vol. 4, pp. 1765-1768, 1999.

[6] P. Petrus, "Robust Huber Adaptive Filter," IEEE Transactions on Signal Processing, vol. 47, No. 4, pp. 1129-1133, 1999.

[7] P. J. Rousseeuw and A. M. Leroy, Robust Regression and Outlier Detection: John Wiley & Sons Inc., New York, 1987.

[8] F. B. Boroujeny, Adaptive Filters : Theory and Applications: Chichester : John Wiley & Sons, 1999.

[9] S. Haykin, Adaptive Filter Theory: Prentice - Hall, 2nd edition, 1991.

[10] R. E. Frank and M. Hampel, Robust Statistics: The Approach Based On Influence Functions: John Wiley & Sons Inc., 1986.

[11] G. C. Goodwin and K. S. Sin, Adaptive Filtering Prediction and Control: Prentice-Hall, Inc. Englewood Cliffs, New Jersey, 1984.
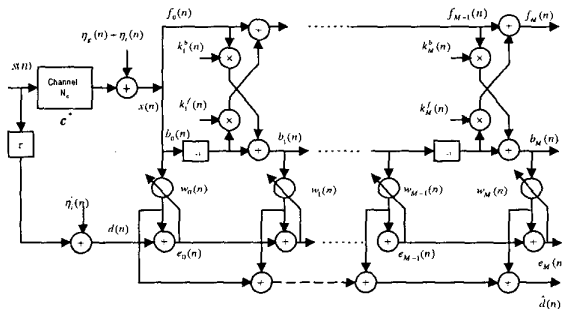
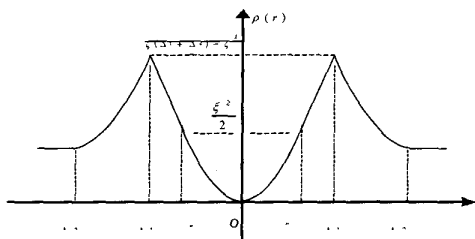Fig.1 Adaptive Channel Equalization Scheme

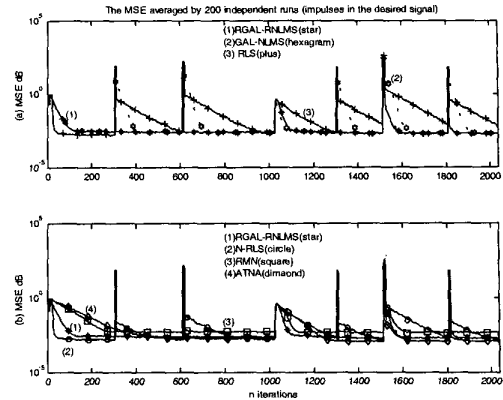Fig.2 Three parts redescending M-estimate function

Fig. 3 The performance comparisons in the presence of the individual and consecutive impulses in the desired signal
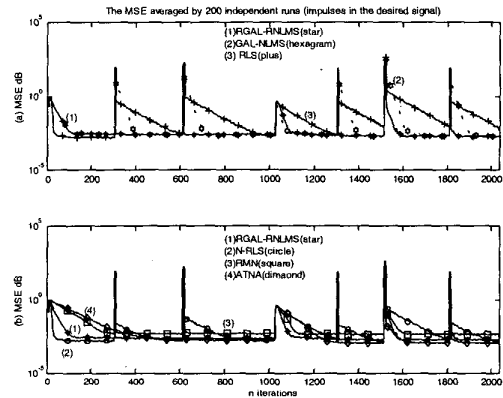

Fig. 4 The performance comparisons in the presence of the individual and consecutive impulses in the input signal

Table 1. The parameters for computer simulations

	N	λ / λ_e / λ_p / β	$\mu_{w,0}$ / $\mu_{k,0}$	N_w / N_p
RGAL_NLMS	7	-/.99 /.99 /.99	.03/.001	5/5
GAL-NLMS	7	-/ - / -/.99	.03/.001	-/-
RMN	11	-/ - / - / -	.01/-	7/-
ATAN	11	-/.995 / - /-	.01/-	$M = 1.4142$ [2]
N-RLS	11	.99 /.99/ - / -	$h_0 = 2.24$, $R^{-1}(0) = 20 \times I$, $P(0) = 0$ [3]	
RLS	11	.99 / - / - / -	$R^{-1}(0) = 20 \times I$, $P(0) = 0$	