

A Recursive Least M-Estimate (RLM) Adaptive Filter for Robust Filtering in Impulse Noise

Y. Zou, S. C. Chan, and T. S. Ng

Abstract—This paper proposes a recursive least M-estimate (RLM) algorithm for robust adaptive filtering in impulse noise. It employs an M-estimate cost function, which is able to suppress the effect of impulses on the filter weights. Simulation results showed that the RLM algorithm performs better than the conventional RLS, NRLS, and the OSFKF algorithms when the desired and input signals are corrupted by impulses. Its initial convergence, steady-state error, computational complexity, and robustness to sudden system change are comparable to the conventional RLS algorithm in the presence of Gaussian noise alone.

Index Terms—Adaptive filter, impulse noise suppression, recursive least M-estimate algorithm, robust statistics, system identification.

I. INTRODUCTION

THE performance of conventional linear adaptive filters can deteriorate significantly when the desired or the input signal is corrupted by impulse noise. Several nonlinear algorithms, such as the order statistic least mean square (OSLMS) [1], the order statistic fast Kalman filtering (OSFKF) [2], the adaptive threshold nonlinear (ATNA) [3], the robust mixed-norm (RMN) [4], and the nonlinear recursive least square (NRLS) algorithms [5] have been developed to combat the adverse effects due to impulses. In this paper, an RLS-like algorithm called the recursive least M-estimate (RLM) algorithm and a systematic method for estimating its threshold parameters are proposed. It employs an M-estimate cost function, which is able to suppress the hostile effect of large estimation error, due to impulses, on the filter weights. Simulation results showed that the performance of the RLM algorithm is better than the conventional RLS, NRLS, and OSFKF algorithms, when the desired and input signals are corrupted by impulses. The initial convergence, steady-state error, computational complexity, and the robustness to sudden system change of the RLM algorithm are also found to be comparable to the conventional RLS algorithm with Gaussian noise alone. The RLM algorithm differs from the Huber adaptive filter recently reported in [6] in the following ways. The Huber adaptive filter suppresses the adverse effect of the impulses in the input signal on the general M-estimator ([7, p. 12]) by down-weighting the distortion measure when the input signal amplitude is large. Also, it is not recursive in nature and a system of nonlinear equations must be solved for each data block. The proposed RLM algorithm on the other hand, is a recursive algorithm with a much lower complexity

Manuscript received June 6, 2000. The associate editor coordinating the review of this manuscript and approving it for publication was Prof. G. Ramponi. The authors are with the Department of Electrical and Electronic Engineering, The University of Hong Kong, Hong Kong (e-mail: scchan@eee.hku.hk).

Publisher Item Identifier S 1070-9908(00)09401-3.

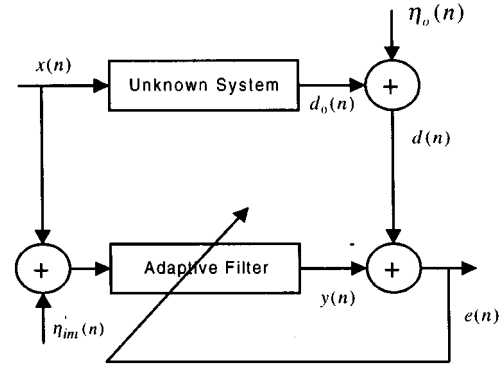


Fig. 1. System identification structure.

of $O(N^2)$ per iteration. Furthermore, LMS and transform domain LMS-like generalizations of the proposed algorithm with complexity $O(N)$ per iteration have also been proposed recently by the authors [8], [9].

II. ROBUST M-ESTIMATE ADAPTIVE FILTER

Let us consider the system identification problem in Fig. 1. The signals $x(n)$ and $y(n)$ are, respectively, the input and output of the adaptive linear transversal filter. The estimation error at time instant n is given by $e(n) = d(n) - \mathbf{w}^T(n)\mathbf{X}(n)$, where $\mathbf{w}(n) = [w_1(n), \dots, w_N(n)]^T$ and $\mathbf{X}(n) = [x(n), \dots, x(n - N + 1)]^T$ are the weight vector and the input signal vector, respectively. $d(n)$ is the desired signal, which consists of the output of the unknown system $d_o(n)$ and the additive interference $\eta_o(n)$, i.e., $d(n) = d_o(n) + \eta_o(n)$. Instead of the commonly used least square (LS) cost function $J_{LS}(n) \triangleq \sum_{i=1}^n \lambda^{n-i} e^2(i)$, the following M-estimate cost function is proposed:

$$J_\rho(n) \triangleq \sum_{i=1}^n \lambda^{n-i} \rho(e(i)) \quad (1)$$

where λ is a forgetting factor and $\rho(\cdot)$ is an M-estimate function. In this paper, the Hampel's three-part redescending M-estimate function ([7, p. 148]) is considered due to its computational simplicity and more flexibility in choosing the interval parameters for impulse noise suppression

$$\rho(e) \triangleq \begin{cases} e^2/2, & 0 < |e| < \xi \\ \xi|e| - \xi^2/2, & \xi \leq |e| < \Delta_1 \\ \frac{\xi}{2}(\Delta_2 + \Delta_1) - \frac{\xi^2}{2} + \frac{\xi}{2} \frac{(|e| - \Delta_2)^2}{\Delta_1 - \Delta_2}, & \Delta_1 \leq |e| < \Delta_2, \\ \frac{\xi}{2}(\Delta_2 + \Delta_1) - \frac{\xi^2}{2}, & \Delta_2 \leq |e| \end{cases} \quad (2)$$

where $\rho(\cdot)$ is a real-valued even function and is quadratic when e is smaller than ξ . For values of e in $[\xi, \Delta_1]$, the function is linear. For values of e greater than Δ_2 , the function is equal to a constant. The M-estimator is capable of suppressing outliers with large amplitude, and 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 estimated continuously. The cost function $J_\rho(n)$ is therefore capable of smoothing out momentary fluctuation caused by impulses. The optimal weight vector can be obtained by setting the first order partial derivatives of $J_\rho(n)$ with respect to $\mathbf{w}(n)$ to zero. This yields

$$\mathbf{R}_{X\rho}(n)\mathbf{w}(n) = \mathbf{P}_{X\rho}(n)$$

where

$$\begin{aligned} \mathbf{R}_{X\rho}(n) &= \sum_{i=1}^n \lambda^{n-i} q(e(i)) \mathbf{X}(i)\mathbf{X}^T(i) \\ &= \lambda \mathbf{R}_{X\rho}(n-1) + q(e(n)) \mathbf{X}(n)\mathbf{X}^T(n), \\ \mathbf{P}_{X\rho}(n) &= \sum_{i=1}^n \lambda^{n-i} q(e(i)) d(i)\mathbf{X}(i) \\ &= \lambda \mathbf{P}_{X\rho}(n-1) + q(e(n)) d(n)\mathbf{X}(n) \end{aligned} \quad (3)$$

$q(e) \triangleq \psi(e)/e$ and $\psi(e) \triangleq \partial \rho(e)/\partial e$. The first equation in (3) is referred to as the M-estimate normal equation, $\mathbf{R}_{X\rho}(n)$ and $\mathbf{P}_{X\rho}(n)$ are called the M-estimate correlation matrix of $\mathbf{X}(n)$, and the M-estimate cross-correlation vector of $\mathbf{X}(n)$ and $d(n)$, respectively. They serve similar purposes as the conventional correlation matrix of $\mathbf{X}(n)$ and the cross-correlation vector of $\mathbf{X}(n)$ and $d(n)$ in the RLS algorithm. In the following section, a recursive algorithm, called the RLM algorithm, will be derived for solving (3).

III. RECURSIVE LEAST M-ESTIMATE (RLM) ALGORITHM

Apply the matrix inversion lemma $(\mathbf{A} + \mu\mathbf{xy}^T)^{-1} = \mathbf{A}^{-1}\{\mathbf{I} - (\mu\mathbf{xy}^T\mathbf{A}^{-1})/(1 + \mu\mathbf{y}^T\mathbf{A}^{-1}\mathbf{x})\}$ to (3) and letting $\mathbf{A} = \lambda\mathbf{R}_{X\rho}(n-1)$, $\mathbf{x} = \mathbf{y} = \mathbf{X}(n)$, and $\mu = q(e(n))$, $\mathbf{V}(n) = \mathbf{R}_{X\rho}^{-1}(n)$ can be computed recursively as

$$\begin{aligned} \mathbf{V}(n) &= \lambda^{-1} (\mathbf{I} - \mathbf{K}(n)\mathbf{X}^T(n)) \mathbf{V}(n-1), \\ \mathbf{K}(n) &= \frac{q(e(n)) \mathbf{V}(n-1)\mathbf{X}(n)}{\lambda + q(e(n)) \mathbf{X}^T(n)\mathbf{V}(n-1)\mathbf{X}(n)} \end{aligned} \quad (4)$$

where $\mathbf{K}(n)$ is the M-estimate gain vector. Using (3) and (4), the filter weights can be updated by $\mathbf{w}(n) = \mathbf{w}(n-1) + (d(n) - \mathbf{X}^T(n)\mathbf{w}(n-1))\mathbf{K}(n)$. Therefore, when $e(n)$ is smaller than ξ , $q(e(n))$ is equal to one, and (4) is identical to the gain vector in the conventional RLS algorithm. When $e(n)$ is larger than ξ , $q(e(n))$ starts to reduce and is equal to zero when $e(n) > \Delta_2$. The latter property makes the proposed algorithm more robust to consecutive impulses than the NRLS algorithm [5]. In fact, the corrupted $e(n)$ in the NRLS algorithm, though limited by the clipping devices, will also be added to $\mathbf{w}(n)$, which makes them sensitive to consecutive impulses. In contrast, for the RLM algorithm, if an impulse is detected ($|e(n)| > \Delta_2$), $\mathbf{K}(n)$ is set to zero, and $\mathbf{w}(n)$ is not updated. If $\xi < |e(n)| < \Delta_2$, the contribution of $\mathbf{K}(n)$ to $\mathbf{w}(n)$ will be reduced as well.

IV. PARAMETER ESTIMATION

To provide robust estimation under impulse noise environment, the threshold parameters ξ , Δ_1 , and Δ_2 need to be estimated continuously. Although the distribution of the error signal $e(n)$ is in general unknown due to the presence of impulses, it can be assumed, for simplicity, to be Gaussian distributed but corrupted by additive impulses. By estimating the variance of ‘‘impulse-free’’ estimation error $\hat{\sigma}^2(n)$, it is possible to detect and reject the impulses in $e(n)$. Specifically, the probability of $|e(n)|$ greater than a given threshold T is $\theta_T(n) = Pr\{|e(n)| > T\} = \text{erfc}(T/\sqrt{2}\hat{\sigma}(n))$ [10], where $\text{erfc}(r) = 2 \int_0^r e^{-x^2} dx/\sqrt{\pi}$ is the complementary error function, and $\hat{\sigma}(n)$ is the standard deviation of the ‘‘impulse-free’’ estimation error. Using a different threshold parameter T , one can detect the impulse noise with different degrees of confidence. Let $\theta_\xi = Pr\{|e(n)| > \xi\}$, $\theta_{\Delta_1} = Pr\{|e(n)| > \Delta_1\}$, and $\theta_{\Delta_2} = Pr\{|e(n)| > \Delta_2\}$ be the probabilities that $|e(n)|$ is greater than ξ , Δ_1 , and Δ_2 , respectively. If θ_ξ , θ_{Δ_1} , and θ_{Δ_2} are chosen to be 0.05, 0.025, and 0.01, respectively, we have 95% confidence to down weight the error in the interval $[\xi, \Delta_1]$, 97.5% confidence to down weight the error signal in the interval $[\Delta_1, \Delta_2]$ and 99% confidence to reject it when $|e(n)| > \Delta_2$. Then, the thresholds are determined to be $\xi = k_\xi \hat{\sigma}(n) = 1.96\hat{\sigma}(n)$, $\Delta_1 = k_{\Delta_1} \hat{\sigma}(n) = 2.24\hat{\sigma}(n)$, $\Delta_2 = k_{\Delta_2} \hat{\sigma}(n) = 2.576\hat{\sigma}(n)$. The commonly used estimator $\hat{\sigma}^2(n) = \lambda_\sigma \hat{\sigma}^2(n-1) + (1 - \lambda_\sigma)e^2(n)$ [5] is not used here due to its sensitivity to impulses with large amplitude. A more robust estimator using the median of the absolute deviation is given in [7, p. 45]. Its complexity, however, is rather high because of the considerable amount of median operations required. In this paper, a new recursive estimator for $\hat{\sigma}^2(n)$ is proposed as follows:

$$\hat{\sigma}^2(n) = \lambda_\sigma \hat{\sigma}^2(n-1) + C_1(1 - \lambda_\sigma)\text{med}(A_e(n)) \quad (5)$$

where

$$\begin{aligned} A_e(n) &= \{e^2(n), \dots, e^2(n - N_w + 1)\}; \\ N_w &= \text{length of the estimation window}; \\ \lambda_\sigma &= \text{forgetting factor}; \\ C_1 = 1.483(1 + 5/(N_w - 1)) &= \text{finite sample correction factor [7, p. 44]}. \end{aligned}$$

Due to the recursive nature and the median operation in (5), this estimator provides more stable estimation of the variance $\hat{\sigma}^2(n)$. It can be seen from the above derivation that the arithmetic complexity of the proposed RLM algorithm is comparable to that of the conventional RLS algorithm, except for the N more multiplications in (4) and $O(N_w \log N_w)$ operations required in (5).

V. SIMULATION RESULTS

The performance of the proposed RLM algorithm is evaluated and compared with the RLS, NRLS [5], and the OSFKF [2] algorithms for the system identification problem shown in Fig. 1. The impulse response of the unknown system is $\mathbf{w}^* = [0.2, -0.4, 0.6, -0.8, 1, -0.8, 0.6, -0.4, 0.2]^T$, which is changed to $-\mathbf{w}^*$ at $n = 5500$ to evaluate its robustness to sudden system change. $x(n)$ is colored and is generated by passing a zero-mean, unit variance white Gaussian

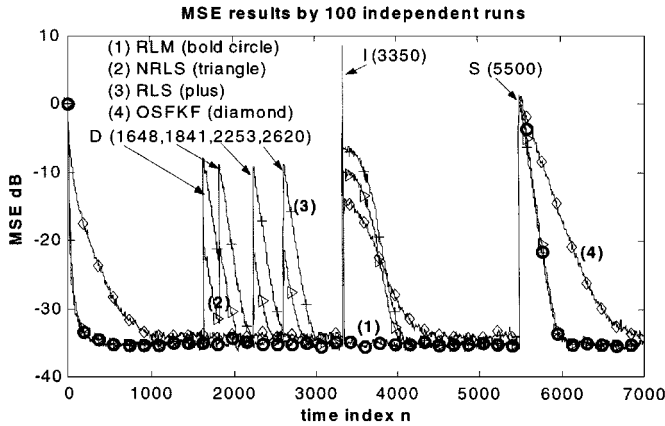


Fig. 2. MSE performance of the various algorithms under different impulsive interference (Example 1). (1) RLM (bold circle); (2) NRLS (triangle); (3) RLS (plus); (4) OSFKF (diamond). $N = 9$, $\text{SNR} = 35$ dB, $\lambda = \lambda_\sigma = 0.99$, $\hat{\sigma}^2(0) = d^2(0)$, $\mathbf{w}(0) = \mathbf{P}(0) = \mathbf{O}$, $\mathbf{V}(0) = 20\mathbf{I}$, and $N_w = 14$ in (5), $\{D(n)\}$, and $I(n)$ indicate the locations of the impulses in the desired and the input signals at time instant n , respectively. $S(n)$ indicates the time instant when the system changes suddenly.

process through a linear time-invariant filter with coefficients $[\cdot, 3887, 1, \cdot, 3887] \cdot \eta_0(n)$ is modeled as the frequently used contaminated Gaussian (CG) noise $\eta_0(n) = \eta_g(n) + b(n)\eta_w(n)$ [1], [2], where $\eta_g(n)$ and $\eta_w(n)$ are independently identically distributed (i.i.d.) zero mean Gaussian noises with variance σ_g^2 and σ_w^2 , respectively, and $b(n)$ is an i.i.d Bernoulli random variable with occurrence probability $P_r(b(n) = 1) = p_r$. The ratio $\gamma_{im} = \sigma_{im}^2/\sigma_g^2 = p_r\sigma_w^2/\sigma_g^2$ determines the impulsive characteristic of $\eta_0(n)$. For fixed value of σ_g^2 , the larger the γ_{im} , the more impulsive $\eta_0(n)$ becomes. The SNR at the system output is defined as $\text{SNR} = 10 \log_{10}(\sigma_{d_0}^2/\sigma_g^2)$, where $\sigma_{d_0}^2$ is the variance of $d_0(n)$. Simulation parameters and the initial values for various algorithms are shown in Fig. 2. The constant 20 in $\mathbf{V}(0)$ is chosen to ensure that $\mathbf{R}_{X\rho}(0)$ consists of reasonable values. For illustration purposes, from $n = 1-1549$, and $2801-7000$, $\eta_0(n) = \eta_g(n)$ is used, whereas from $n = 1550-2800$, $\eta_0(n) = \eta_g(n) + b(n)\eta_w(n)$ with $p_r = 0.005$ and $r_{im} = 300$ is used. To visualize clearly the effect of impulses in $d(n)$, their locations generated by $b(n)$ are fixed and marked in Fig 2, but their amplitudes are varied according to $\eta_w(n)$, which is generated statistically independent in each run. Also, for simplicity in visualizing the effect of impulses in $x(n)$, only one impulse is added to $x(n)$ at $n = 3350$. $N_w = 14$ is used for the OSFKF algorithm. The MSE results averaged over 100 independent runs are plotted in Fig. 2. The RLM, NRLS, and the RLS algorithms have almost identical initial convergence speed, lower steady-state error, and robustness to sudden system change. The RLS algorithm, however, is not robust to any impulse, and the NRLS algorithm is sensitive to impulses in $d(n)$ and $x(n)$. The performances of the RLS and the OSFKF algorithms are also deteriorated significantly by the impulse in $x(n)$ at $n = 3350$, which lasts for several hundred iterations. On the other hand, the RLM algorithm is able to recover very quickly in about $3N$ iterations. We also observed that the OSFKF algorithm has a slower convergence speed, slower response to sudden system change, and higher steady-state error than other algorithms. Other experiments have been performed to evaluate the influence of SNR, p_r and k_ξ , k_{Δ_1} , and k_{Δ_2} on the performance of the RLM algorithm.

The details are omitted here due to space limitation. We only summarize the observations here. First, compared with other robust RLS-like algorithms, the RLM algorithm has better performance at different SNR and impulse density p_r . Second, the performance of the RLM algorithm is not sensitive to the choices of k_ξ , k_{Δ_1} , and k_{Δ_2} , provided that they are not much larger than that suggested in Section IV. Comparisons with other robust LMS-like algorithms, such as ATNA, RMN, and the OSLMS algorithms, have also been performed. The RLM algorithm is found to have faster initial convergence, lower steady-state error, and better robustness to sudden system change than those algorithms due to its RLS-like nature. In addition, the authors have evaluated the mean and mean square convergence of the RLM algorithm using the CG noise model. Results show that the theoretical and simulation results agree with each other. Interested readers are referred to [8], [9], and [11] for details.

VI. CONCLUSION

An RLM algorithm for robust filtering in impulse noise is presented. It employs an M-estimate cost function, which is able to suppress the adverse effect of impulses on the filter weights. Simulation results showed that the performance of the RLM algorithm is better than the conventional RLS, NRLS, and the OSFKF algorithms in impulse noise environment. Its initial convergence, steady-state error, computational complexity, and robustness to sudden system change are also found to be comparable to the conventional RLS algorithm in Gaussian noise alone.

ACKNOWLEDGMENT

The authors would like to thank the reviewers for carefully reading this manuscript and for their constructive comments, which greatly improved the presentation of this paper.

REFERENCES

- [1] T. I. Haweel and P. M. Clarkson, "A class of order statistic LMS algorithms," *IEEE Trans. Signal Processing*, vol. 40, pp. 44-53, Jan. 1992.
- [2] R. Settineri, M. Najim, and D. Ottaviani, "Order statistic fast Kalman filter," *IEEE ISCAS'96*, vol. 2, pp. 116-119, 1996.
- [3] S. Koike, "Adaptive threshold nonlinear algorithm for adaptive filters with robustness against impulsive noise," *IEEE Trans. Signal Processing*, vol. 45, pp. 2391-2395, Sept. 1997.
- [4] J. A. Chambers and A. Avlonitis, "A robust mixed-norm (RMN) adaptive filter algorithm," *IEEE Signal Processing Lett.*, vol. 4, pp. 46-48, Feb. 1997.
- [5] J. F. Weng and S. H. Leung, "Adaptive nonlinear RLS algorithm for robust filtering in impulse noise," *IEEE ISCAS'97*, vol. 4, pp. 2337-2340, 1997.
- [6] P. Petrus, "Robust Huber adaptive filter," *IEEE Trans. Signal Processing*, vol. 47, pp. 1129-1133, Apr. 1999.
- [7] P. J. Rousseeuw and A. M. Leroy, *Robust Regression and Outlier Detection*. New York: Wiley, 1987.
- [8] Y. Zou, S. C. Chan, and T. S. Ng, "Least mean M-estimate algorithms for robust adaptive filtering in impulse noise," in *ENSIPCO-2000*, Tampere, Finland.
- [9] —, "A robust statistics based adaptive lattice-ladder filter in impulsive noise," in *IEEE ISCAS 2000*, Geneva, Switzerland.
- [10] —, "A robust M-estimate adaptive filter for impulse noise suppression," *IEEE ICASSP'99*, vol. 4, pp. 1765-1768.
- [11] S. C. Chan, Y. Zou, and T. S. Ng, "A recursive least M-estimate algorithm for robust adaptive filtering in impulsive noise and its performance analysis," Int. Rep., Dept. Elect. Electron. Eng., Univ. Hong Kong, Hong Kong, May 2000.