Fast Least Mean *M*-Estimate Algorithms For Robust Adaptive Filtering In Impulse Noise

Yuexian Zou, Shing-Chow Chan, and Tung-Sang Ng
Department of Electrical and Electronic Engineering
The University of Hong Kong, Pokfulam Road, Hong Kong
Email: {yxzou, scchan, tsng}@eee.hku.hk

ABSTRACT

Adaptive filters with suitable nonlinear devices are very effective in suppressing the adverse effect due to impulse noise. In a previous work, the authors have proposed a new class of nonlinear adaptive filters using the concept of robust statistics [1, 2]. The robust *M*-estimator is used as the objective function, instead of the mean square errors, to suppress the impulse noise. The optimal coefficient vector for such nonlinear filter is governed by a normal equation which can be solved by a recursive least squares like algorithm with $O(N^2)$ arithmetic complexity, where N is the length of the adaptive filter. In this paper, we generalize the robust statistic concept to least mean square (*LMS*) and transform domain *LMS* algorithms. The new fast nonlinear adaptive filtering algorithms called the least mean *M*-estimate (*LMM*) and transform domain *LMM* (*TLMM*) algorithms are derived. Simulation results show that they are robust to impulsive noise in the desired and input signals with an arithmetic complexity of order $O(N)$.

I. 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 [3, 4], 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 [5], combats the impulsive noise in the desired signal by minimizing 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 [1], a *RLS*-liked algorithm, called the *M-RLS* algorithm, was proposed for impulse noise suppression, by minimizing 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 the 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 complexities of the *M-RLS* algorithm and its improvement version, called the *RLM* algorithm [2], however, are still rather high ($O(N^2)$) compared with the *LMS* type algorithms which is $O(N)$. It is the purpose of this paper to develop fast transversal filtering algorithms using the concept of robust statistics as proposed in [1, 2]. In particular, a new robust least mean *M*-estimate (*LMM*) algorithm and its transform domain counterpart, called the transform domain least mean *M*-estimate algorithm (*TLMM*), are developed. They can be viewed, respectively, as the generalization of the conventional *LMS* and transform domain *LMS* (*TLMS*) algorithms with an arithmetic complexity of order $O(N)$. Simulation results show that they are more robust than the *ATNA*, *N-RLS*, and *RMN* algorithms in impulsive noise environment.

II. OVERVIEW OF THE ROBUST ALGORITHMS

Let's consider the system identification problem in Fig. 1 with an adaptive linear transversal filter. The signals $x(n)$, $y(n)$ and $d(n)$ are the input, output and desired signals of this filter, respectively. The estimation error $e(n)$ of the adaptive filter at time instant n is given by

$$e(n) = d(n) - \boldsymbol{w}^T(n)\boldsymbol{X}(n), \tag{1}$$

where $\boldsymbol{w}(n) = [w_0(n), w_1(n), ..., w_{N-1}(n)]^T$ is the filter coefficient vector, $\boldsymbol{X}(n) = [x(n), x(n-1), ..., x(n-N+1)]^T$ is its input vector. In practice, $d(n)$ and $x(n)$ may be corrupted by additive interference signal $\eta(n) = \eta_g(n) + \eta_{im}(n)$, where $\eta_g(n)$ and $\eta_{im}(n)$ are additive Gaussian and impulsive noise, respectively. In what follows, the principle of the *ATNA*, *RMN*, *N-RLS* and *RLM* algorithms will be briefly described.

A. Adaptive Threshold Nonlinear Algorithm

The nonlinear least mean squares (*ATNA*) algorithm is a stochastic gradient based algorithm. The cost function $J_{MSE} = E[e^2(n)]$ is minimized by updating the filter coefficient vector as [3]

$$\hat{\boldsymbol{w}}(n) = \hat{\boldsymbol{w}}(n-1) + \mu f_c(e(n))\boldsymbol{X}(n), \quad f_c(e(n)) = \begin{cases} e(n) & |e(n)| \le h \\ 0 & |e(n)| > h \end{cases},$$

$$h = \sqrt{2}\hat{\sigma}_e(n), \quad \hat{\sigma}_e^2(n) = \lambda_\sigma \hat{\sigma}_e^2(n-1) + (1-\lambda_\sigma)e^2(n). \tag{2}$$

where μ is the stepsize, h is the threshold parameter, $0 < \lambda_\sigma \le 1$ is the forgetting factor and $f_c(\cdot)$ is a truncated linear function applied to the error signal $e(n)$ to reduce its influence when it is abnormally large. The limitations of *ATNA* are its slow convergence speed and slow tracking ability.

B. Robust Mixed-Norm Adaptive Filter Algorithm

In the mixed-norm adaptive filter (*RMN*) algorithm [5], the coefficient vector is updated to minimize the cost function $J_{MIX} = \lambda_m E[e^2(n)] + (1-\lambda_m)E[|e(n)|]$. The resulting algorithm is similar to the combination of the well-known *LMS* algorithm and the least absolute difference (*LAD*) algorithm:

$$\hat{\boldsymbol{w}}(n) = \hat{\boldsymbol{w}}(n-1) + \mu\{2\lambda_m e(n) + (1-\lambda_m)sign(e(n))\}\boldsymbol{X}(n),$$
$$\lambda_m = 2erfc(|e(n)|/\hat{\sigma}_e(n)), \tag{3}$$

where λ_m is a mixing parameter, $erfc(.)$ is the complementary error function and $\hat{\sigma}_e(n)$ is estimated by the trimming window method (details can be found in [5]). The *RMN* algorithm also suffers from slow convergence speed and the increased steady-error due to the use of the mixed-norm.

C. Nonlinear *RLS* Algorithm

The nonlinear recursive least squares (*N-RLS*) algorithm is derived from the conventional *RLS* algorithm based on the cost function $J_{LS}(n) = \sum_{i=1}^{n} \lambda^{n-i} e^2(i)$. The coefficient vector is updated as [4],

$$\hat{\boldsymbol{w}}(n) = \hat{\boldsymbol{w}}(n-1) + f_c(e(n))\boldsymbol{K}(n), \quad f_c(e(n)) = \begin{cases} e(n) & |e(n)| \le h \\ h & |e(n)| > h \end{cases},$$

$$h = 2.24\hat{\sigma}_e(n), \quad \hat{\sigma}_e^2(n) = \lambda_\sigma \hat{\sigma}_e^2(n-1) + (1-\lambda_\sigma)e^2(n). \qquad (4)$$

Where, $\boldsymbol{K}(n)$ is the Kalman gain. $f_c(\cdot)$ is the clipper function. The advantage of the *N-RLS* algorithm is its fast convergence speed and low steady-state error. However, the complexity of the *N-RLS* algorithm is relatively high with an order of $O(N^2)$ per iteration.

All of these algorithms, unfortunately, are not robust to impulses in the input signal and consecutive impulses in the desired signal.

D. Robust *M*-Estimate Adaptive Filter And Recursive Least *M*-Estimate (*RLM*) Algorithm

In [1, 2], the authors have proposed a new approach to robust adaptive filtering, which can effectively suppress the adverse effects of the impulse in the desired signal $d(n)$ and/or the input signal $x(n)$. The proposed cost function is based on the robust *M*-estimate function, $J_{L\rho}(n) = \sum_{i=1}^{n} \lambda^{n-i}\rho(e(i))$. Here, $\rho(\cdot)$ is a robust *M*-estimate function for suppressing impulsive noise. In [1], $\rho(\cdot)$ was chosen to be the modified Huber *M*-estimate function. Later in [2], the following more general Hampel's three parts redescending *M*-estimate function was studied [8].

$$\rho(e) = \begin{cases} e^2/2, & 0 < |e| < \xi \\ \xi|e| - \xi^2/2, & \xi \le |e| < \Delta_1 \\ \frac{\xi}{2}(\Delta_2 + \Delta_1) - \frac{\xi^2}{2} - \frac{\xi}{2}\frac{(|e| - \Delta_2)^2}{\Delta_2 - \Delta_1}, & \Delta_1 \le |e| < \Delta_2 \\ \frac{\xi}{2}(\Delta_2 + \Delta_1) - \frac{\xi^2}{2}, & \Delta_2 \le |e| \end{cases}. \qquad 5)$$

The advantage of this *M*-estimate is that its first order derivative is continuous. As shown in Fig. 2, $\rho(\cdot)$ is an even real-valued function and it is quadratic when e is smaller than ξ. For $e \in [\xi, \Delta_1]$, the function is linear. For $e > \Delta_2$, the function is equal to a constant. The threshold parameters ξ, Δ_1, and Δ_2 are used to control the degree of suppression of the outliers. Smaller values of ξ, Δ_1, and Δ_2 imply greater suppression of the outliers. Therefore, $J_{L\rho}(n)$ is capable of smoothing out momentary fluctuation caused by the impulsive interference. The optimal coefficient vector can be obtained by setting the first order partial derivatives of $J_{L\rho}(n)$, w.r.t. $\boldsymbol{w}(n)$, to zero. This yields,

$$\sum_{i=1}^{n} \lambda^{n-i}\psi(e(i))\boldsymbol{X}(i) = \sum_{i=1}^{n} \lambda^{n-i} q(e(i))e(i)\boldsymbol{X}(i) = 0, \qquad (6)$$

where $\psi(r) = \partial\rho(r)/\partial r$ and $q(r) = \psi(r)/r$. Substituting (1) into (6), one gets

$$\boldsymbol{R}(n)\boldsymbol{w}(n) = \boldsymbol{P}(n), \qquad (7)$$

where

$$\boldsymbol{R}(n) = \sum_{i=1}^{n} \lambda^{n-i} q(e(i))\boldsymbol{X}(i)\boldsymbol{X}^T(i) = \lambda\boldsymbol{R}(n-1) + q(e(n))\boldsymbol{X}(n)\boldsymbol{X}^T(n),$$

$$\boldsymbol{P}(n) = \sum_{i=1}^{n} \lambda^{n-i} q(e(i))d(i)\boldsymbol{X}(i) = \lambda\boldsymbol{P}(n-1) + q(e(n))d(n)\boldsymbol{X}(n), \qquad (8)$$

are the correlation matrix of $\boldsymbol{X}(n)$ and the cross-correlation vector between $d(n)$ and $\boldsymbol{X}(n)$, respectively. We shall call (7) the *M*-estimate normal equation, which is a system of nonlinear equations. An effective recursive algorithm, called the recursive least *M*-estimate algorithm (*RLM*) was derived in [2] for solving this normal equation. Simulation results showed that the *RLM* algorithm could effectively suppress the adverse effects of individual and consecutive impulses occurring in the desired and the input signals. Its convergence speed and steady-state error is relatively unaffected by the impulses with a performance similar to the *RLS* algorithm in the noise-free case. However, its arithmetic complexity, like the *RLS* algorithm, is approximately $O(N^2)$ per iteration.

III. ROBUST LEAST MEAN *M*-ESTIMATE ADAPTIVE FILTER ALGORITHMS

In what follows, we shall generalize the concept of robust statistics to the *LMS* and *TLMS* algorithms to derive robust adaptive algorithms with $O(N)$ arithmetic complexity.

Least Mean *M*-Estimate (*LMM*) Algorithm

For any coefficient vector $\boldsymbol{w}$, the gradient vector of the mean square error function $J_{MSE} = E[e^2(n)]$ is

$$\nabla_w(J_{MSE}) = \frac{\partial}{\partial \boldsymbol{w}} E[e^2(n)] = 2E[e(n)\boldsymbol{X}(n)] = -2\boldsymbol{b} + 2\boldsymbol{R}_X\boldsymbol{w}. \qquad (9)$$

where, $\boldsymbol{R}_X = E[\boldsymbol{X}(n)\boldsymbol{X}^T(n)]$ and $\boldsymbol{b} = E[d(n)\boldsymbol{X}(n)]$ are the correlation matrix of $\boldsymbol{X}(n)$ and the cross-correlation vector between $d(n)$ and $\boldsymbol{X}(n)$, respectively. In the *LMS* algorithm, J_{MSE} is minimized by updating the coefficient vector $\hat{\boldsymbol{w}}(n)$ in the negative direction of the instantaneous gradient vector, $\hat{\nabla}_w$,

$$\nabla(J_{MSE}) \approx \hat{\nabla}_w = \frac{\partial}{\partial \boldsymbol{w}}(e^2(n)) = -2e(n)\boldsymbol{X}(n), \qquad (10)$$

$$\hat{\boldsymbol{w}}(n) = \hat{\boldsymbol{w}}(n-1) - \mu\hat{\nabla}_w = \hat{\boldsymbol{w}}(n-1) + 2\mu e(n)\boldsymbol{X}(n). \qquad (11)$$

In the proposed *LMM* algorithm, the robust statistic based objective function $J_{M\rho} = E[\rho(e(n))]$ is minimized instead of J_{MSE}. The coefficient vector $\hat{\boldsymbol{w}}(n)$ is also updated in the negative direction of the instantaneous gradient vector $\tilde{\nabla}_w$, which can be written as,

$$\nabla_w(J_{M\rho}) = \frac{\partial J_{M\rho}}{\partial \boldsymbol{w}} \approx \tilde{\nabla}_w = \frac{\partial}{\partial \boldsymbol{w}}(\rho(e(n))) = -q(e(n))e(n)\boldsymbol{X}(n), \qquad (12)$$

$$\hat{\boldsymbol{w}}(n) = \hat{\boldsymbol{w}}(n-1) + \mu q(e(n))\boldsymbol{X}(n)e(n). \qquad (13)$$

where $q(\cdot)$ is the weight function defined previously in (6). Eq.(13) can be viewed as the extension of the *LMS* algorithm in (11).

Transform Domain Least Mean *M*-Estimate (*TLMM*) Algorithm

The limitation of the *LMS*-type algorithms is their slow convergence speed, especially when the distribution of the input signal is colored. The transform domain *LMS* (*TLMS*) proposed in [9] greatly improved the convergence speed of the *LMS* algorithm. Various aspects of the *TLMS* and their performance analysis can be found in [10-12]. The input signal, $x(n)$, is first transformed by an $(N \times N)$ orthogonal matrix $\boldsymbol{Q}$ to produce transform outputs, $x_t(n,i)$, $i = 1,..., N$. Each of the N outputs is then normalized by its power estimate p_i^2, which are usually estimated as, $p_i^2(n) = \lambda_p p_i^2(n-1) + (1-\lambda_p)x_t^2(n,i)$, where λ_p is a forgetting factor and is usually chosen as 0.95. Let $\boldsymbol{X}_t(n) = [x_t(n,1), \cdots, x_t(n,N)]^T$ be the input vector of the *TLMS* adaptive filter, then the normalized input vector can be expressed as

$$\boldsymbol{X}_N(n) = (\Lambda^2)^{-\frac{1}{2}}\boldsymbol{X}_t(n) = \Lambda^{-1}\boldsymbol{Q}\boldsymbol{X}(n), \qquad (14)$$

where Λ^2 is an $(N \times N)$ diagonal matrix whose $(i,i)-th$ element is equal to $p_i^2(n)$. This normalization helps to reduce the eigenvalue spread of the autocorrelation matrix, $\boldsymbol{R}_X$, of $\boldsymbol{X}(n)$. The autocorrelation matrix of $\boldsymbol{X}_N(n)$ is given by,

$$\boldsymbol{R}_{X_N} = E[\boldsymbol{X}_N(n)\boldsymbol{X}_N^T(n)] = \Lambda^{-1}\boldsymbol{QR}_X\boldsymbol{Q}^T\Lambda^{-1} = \Lambda^{-1}\boldsymbol{QR}_X\boldsymbol{Q}^{-1}\Lambda^{-1}, \qquad (15)$$

If $\boldsymbol{Q}$ is properly chosen, say as the *KLT* of $\boldsymbol{R}_X$, $\boldsymbol{R}_X$ will be diagonalized with $\boldsymbol{R}_{X_N} = \boldsymbol{I}$. Unfortunately, it is very computational expensive to compute the *KLT* of $\boldsymbol{R}_X$ and sub-optimal transforms such as the discrete cosine transform (*DCT*), the discrete Fourier transform (*DFT*), or the discrete Hartley transform (*DHT*) are used. We now consider the derivation of the *TLMS* and *TLMM* algorithms. Pre-multiplying both side of (9) by $\boldsymbol{QR}_X^{-1}$ results in $\boldsymbol{QR}_X^{-1}\nabla_w(J_{MSE}) = -2\boldsymbol{QR}_X^{-1}\boldsymbol{b} + 2\boldsymbol{Qw} = -2\boldsymbol{Qw}_{opt} + 2\boldsymbol{Qw}$, where $\boldsymbol{w}_{opt} = \boldsymbol{R}_X^{-1}\boldsymbol{b}$, which is the optimal solution obtained by letting $\nabla_w(J_{MSE})$ in (9) to zero. Therefore, we obtain

$$\boldsymbol{Qw}_{opt} = \boldsymbol{Qw} - \frac{1}{2}\Lambda^{-1}\left(\Lambda\boldsymbol{QR}_X^{-1}\boldsymbol{Q}^{-1}\Lambda\right)\left(\Lambda^{-1}\boldsymbol{Q}\nabla_w(J_{MSE})\right). \qquad (16)$$

Assuming that the autocorrelation matrix $\boldsymbol{R}_X$ is approximately diagonalized by $\boldsymbol{Q}$, we have $\left(\Lambda\boldsymbol{QR}_X^{-1}\boldsymbol{Q}^{-1}\Lambda\right) \approx \boldsymbol{I}$ and by letting $\boldsymbol{w}_t = \boldsymbol{Qw}$, one gets

$$\boldsymbol{w}_{t,opt} \approx \boldsymbol{w}_t - \tfrac{1}{2}(\Lambda^{-2}\boldsymbol{Q}\nabla_w(J_{MSE})). \qquad (17)$$

The gradient vector can be estimated from the instantaneous *MSE* error as in (10). Therefore one gets the *TLMS* algorithm as, $\boldsymbol{w}_t(n) = \boldsymbol{w}_t(n-1) + \mu e(n)\Lambda^{-2}\boldsymbol{X}_t(n)$, where $\boldsymbol{X}_t(n) = \boldsymbol{QX}(n)$, and $e(n) = d(n) - \boldsymbol{w}_t^T(n)\boldsymbol{X}_t(n)$. Alternatively, the gradient vector can be estimated from the instantaneous robust distortion $\rho(e(n))$ as in eqn. (12), the resulting transform domain least mean *M*-estimate (*TLMM*) algorithm is obtained as follows

$$\boldsymbol{w}_t(n) = \boldsymbol{w}_t(n-1) + \mu\Lambda^{-2}q(e(n))\boldsymbol{X}_t(n)e(n). \qquad (18)$$

Since $J_{M\rho}$ is used instead of J_{MSE}, (13) and (18) are called the least mean *M*-estimate (*LMM*) and transform domain *LMM* algorithms, respectively. It can be seen that when $e(n)$ is smaller than ξ, the weight function $q(e(n))$ is equal to one and (13) and (18) become identical to the corresponding equations in the *LMS* and *TLMS* algorithms, respectively. When $e(n)$ is larger than ξ, $q(e(n))$ becomes smaller and smaller and is zero when $e(n)$ is greater than Δ_2. Thus the *LMM* and *TLMM* algorithms effectively de-emphasize the effect of large signal error during the updating of the filter coefficients. The effectiveness in suppressing the influence of the impulses depends very much on how the threshold parameters ξ, Δ_1 and Δ_2 are estimated. This is addressed in the following section.

IV. PARAMETER ESTIMATION

As mentioned earlier, the choice of the threshold parameters ξ, Δ_1 and Δ_2 can significantly affect the performance of the *LMM* and *TLMM* algorithms. For simplicity, the error signal $e(n)$ is assumed to be Gaussian distributed and corrupted by additive impulsive noise. By estimating the variance of $e(n)$ without the impulses, it is possible to detect and reject the impulse noise in $e(n)$. Specifically, the probability of $|e(n)|$ greater than a given threshold T is given by $\theta_T(n) = P_r\{|e(n)| > T\} = 1 - erf(T/\sqrt{2}\hat{\sigma}_e(n))$ [1], where $erf(r)$ is the error function and $\hat{\sigma}_e(n)$ is the estimated standard deviation of $e(n)$. Using different threshold parameters *T*, we can detect the impulse 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 this 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 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 $|e(n)| > \Delta_2$. The corresponding threshold parameters are determined to be $\xi = 1.96\hat{\sigma}_e(n)$, $\Delta_1 = 2.24\hat{\sigma}_e(n)$, and $\Delta_2 = 2.576\hat{\sigma}_e(n)$. A commonly used estimate of $\hat{\sigma}_e^2(n)$ is $\hat{\sigma}_e^2(n) = \lambda_\sigma\hat{\sigma}_e^2(n-1) + (1-\lambda_\sigma)e^2(n)$ [4]. It is, however, not robust to impulse noise. In fact, a single impulse with large amplitude can substantially increase the value of $\hat{\sigma}_e^2(n)$, and hence the values of ξ, Δ_1 and Δ_2. A more robust but complex estimator is the median absolute deviation from the median (*MAD*) ([8], pp.105). In the paper, the following robust recursive estimator with much lower complexity than the *MAD* for $\hat{\sigma}_e(n)$ is proposed

$$\hat{\sigma}_e^2(n) = \lambda_\sigma\hat{\sigma}_e^2(n-1) + 1.483\left(1 + \tfrac{5}{N_w - 1}\right)(1-\lambda_\sigma)\,\mathrm{med}(A_e(n)), \qquad (19)$$

where $A_e(n) = \{e^2(n), \cdots, e^2(n - N_w + 1)\}$, and N_w is the length of the estimation window. From the above discussion, it can be seen that the arithmetic complexity of the proposed *LMM* and *TLMM* algorithms is close to that of the conventional *LMS* and *TLMS* algorithm, except for the order $O(N_w \log N_w)$ operations in (19).

V. SIMULATION

The performance of the proposed *LMM* and *TLMM* algorithms are evaluated and compared to the conventional *LMS* [13], *TLMS* [9], *RLS* [13], *RLM* [2], *N-RLS* [4], *RMN* [5], and *ATNA* [3] algorithms through simulations of the system identification problem shown in Fig.1. The *DCT* is used as the orthogonal transformation in the *TLMS* and *TLMM* algorithms due to its good performance and low implementation complexity. The unknown system is a *FIR* filter with coefficients $\boldsymbol{w} = [0.2, -0.4, 0.6, -0.8, 1, -0.8, 0.6, -0.4, 0.2]^T$, which is suddenly changed to $-\boldsymbol{w}$ at $n = 1536$. The desired signal is corrupted by the zero mean additive Gaussian noise, $\eta_g(n)$, with variance σ_g^2 and the additive impulsive noise, $\eta_{im}(n)$, which is generated by the Gaussian-Bernoulli process with $P_{ar} \approx 5*10^{-2}$ and $\gamma^2 = 100$ [5] in $n = 1 \sim 900$. The locations of the impulses are fixed in all independent runs, while their amplitudes are governed by a zero-mean Gaussian process [5]. Specifically, these impulses are at time instants $n = 245, 246, 247, 461, 589$. The signal-to-noise ratio at the system output is given by $SNR = 20\log_{10}(\sigma_{d_0}^2/\sigma_g^2)$, where $\sigma_{d_0}^2$ is the variance of the unknown system output, $d_0(n)$, and it is set to $40dB$. The input signal $x(n)$ is a colored signal and it is generated by passing a zero-mean, unit variance white Gaussian process through a linear time-invariant filter with coefficients $[0.3887, 1, 0.3887]$ [13]. There is just one impulse appearing at $n = 944$ in $x(n)$. The length of the adaptive filter is $N = 9$ and that of the estimation window is $N_w = 14$. The forgetting factor λ_σ is chosen as 0.99 except for the *ATNA* algorithm, where it is chosen as $\lambda_\sigma = .995$ according to [3]. All step sizes are set to 0.02 except for the *RMN* algorithm, where it is $\mu = 0.014$. For the *RLS*, *N-RLS* and *RLM* algorithms, $\lambda = 0.99$ with initial conditions $\boldsymbol{w}(0) = \boldsymbol{0}$, $\boldsymbol{R}^{-1}(0) = 20 \cdot \boldsymbol{I}$, $\boldsymbol{P}(0) = \boldsymbol{0}$, and $\hat{\sigma}_e^2(0) = |d(0)|$. The mean squared error (*MSE*) obtained by averaging over 200 independent runs is plotted in Fig. 3. It can be seen from Fig. 3(a) that the performances of the *RLS*, *LMS*, *TLMS* are significantly degraded by the impulses. The proposed *LMM* and *TLMM* algorithms are robust to the impulses appearing either in the desired or input signals. The *LMM* algorithm, however, converges slowly due to its *LMS*-nature. From Fig. 3(b), it is clear that the *TLMM* algorithm provides faster tracking ability compared to the *RLM*, *N-RLS*, *RMN* and *ATNA* algorithms. Its steady-error, however, is higher than that of the *RLM* algorithm. The *RLM*, *N-RLS* and *RLS* algorithms have almost the same initial convergence speed, tracking ability and low steady-state error. It is also noted that the performances of the *N-RLS*, *ATNA*, and *RMN* algorithms were degraded significantly by the impulse in the input signal at $n = 944$. Therefore, it can be concluded that, under the experimental condition specified, the *TLMM* algorithm is more effective and robust than *ATNA*, *N-RLS*, *RMN* in mitigating the adverse effects due to the impulses either in the desired signal or in the input signal. It is therefore a very attractive sub-optimal alternative to the *RLM* algorithm with a much lower computation complexity of $O(N)$.

VI. CONCLUSION

Two new adaptive filtering algorithms called the least mean *M*-estimate (*LMM*) and the transform domain least mean *M*-estimate (*TLMM*) algorithms are proposed for robust adaptive filtering in impulse noise environment. They can be viewed, respectively, as the generalization of the conventional *LMS* and transform domain *LMS* (*TLMS*) algorithms using the robust statistic concept. The arithmetic complexity of the algorithms is of order $O(N)$, which is much lower than that of the *RLM* algorithm proposed previously by the authors. Simulation results show that they are in general more robust than the *ATNA*, the *N-RLS* and the *RMN* algorithms when the desired or input signals are corrupted by individual and consecutive impulses. Simulation results show that the *TLMM* algorithm is a very attractive sub-optimal alternative to the *RLM* algorithm with a much lower computational complexity and faster tracking speed.

REFERENCES

[1] Yuexian 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.

[2] Yuexian Zou, S. C. Chan, and T. S. Ng, "A Recursive *M*-Estimate Algorithm for robust Adaptive Filtering in Impulsive Noise," being prepared for submission, April 2000.

[3] 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.

[4] 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.

[5] 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.

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

[9] S. S. Narayan, A. M. Peterson, and M. J. Narasimha, "Transform domain *LMS* algorithm," *IEEE Transactions On Acoustics, Speech, and Signal Processing*, vol. ASSP-31, No.3, pp. 609-615, 1983.

[10] J. C. Lee and C. K. Un, "Performance of Transform Domain LMS Adaptive Digital Filters," *IEEE Transactions On Acoustics, Speech, and Signal Processing*, vol. ASSP-34, No. 3, pp. 499-510, 1986.

[11] B. F. Boroujeny and S. Gazor, "Selection of orthonormal transforms for improving the performance of the transform domain normalised LMS algorithm," *IEE Proceedings-F*, vol. 139, No. 5, pp. 327-335, 1992.

[12] D. F. Marshall and W. K. Jenkins, "The Use of Orthogonal Transforms for Improving Performance of Adaptive Filters," *IEEE Transactions On Circuits and Systems*, vol. 36, No.4, pp. 474-484, 1989.

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

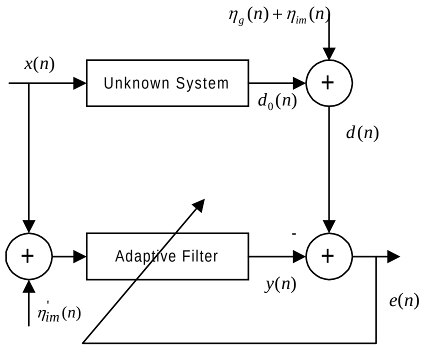

Fig.1 System Identification Structure

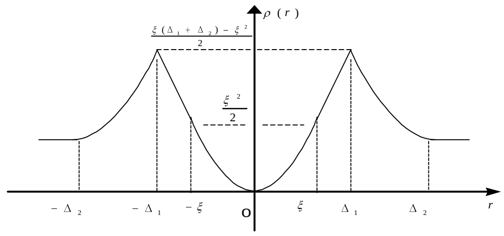

Fig.2 Hampel's three parts redescending *M*-estimate

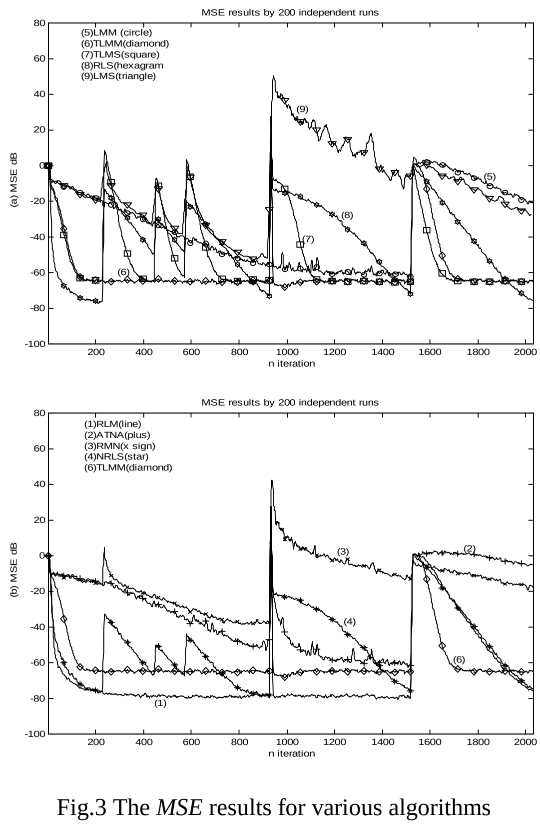

Fig.3 The *MSE* results for various algorithms