

A Robust Quasi-Newton Adaptive Filtering Algorithm For Impulse Noise Suppression

Yuexian Zou and Shing-Chow Chan
 Department of Electrical and Electronic Engineering
 The University of Hong Kong
yxzou@eee.hku.hk scchan@eee.hku.hk

ABSTRACT

This paper studies the problem of robust adaptive filtering in impulse noise environment using the Quasi-Newton (QN) adaptive filtering algorithm. An M -estimate based cost function is minimized instead of the commonly used mean square error (MSE) to suppress the adverse effect of the impulse noise on the filter coefficients. In particular, a new robust quasi-Newton (R-QN) algorithm using the self-scaling variable metric (SSV) method for unconstrained optimization is studied in details. Simulation results show that the R-QN algorithm is more robust to impulse noise in the desired signal than the RLS algorithm and other QN algorithm considered. Its initial convergence speed and tracking ability to sudden system change are also superior to those of the quasi-Newton algorithm proposed in [1].

1. INTRODUCTION

The performance of conventional linear adaptive filtering algorithms can deteriorate significantly when the input or desired signal is corrupted by impulse noise. Nonlinear techniques are usually employed to reduce the adverse effects of the impulse noise on the filter coefficient vector. For example, median filtering has been applied to the LMS and the RLS algorithms to protect the filter weights from the effects of impulsive interference, giving rise to the order statistic least mean square (OSLMS) [2] and the order statistic fast Kalman filtering (OSFKF) [3] algorithms. While the adaptive threshold nonlinear (ATNA) [4] and the nonlinear RLS (NRLS) [5] algorithms use nonlinear clipping functions to limit the transient fluctuation of the estimation error on filter coefficients in the LMS and the RLS adaptive algorithms. Recently, the authors have proposed a family of robust adaptive filtering algorithms based on the minimization of an M -estimate cost function. In particular, a recursive least M -estimate (RLM) algorithm [6], and a least mean M -estimate (LMM) [7] algorithms have been developed based on the Hampel's three part redescending M -estimate cost function [8] to suppress the impulses in the desired and the input signals. Simulation results showed that the robust statistic approach offer improved robustness to consecutive impulses in the desired and input signals than other approaches [6, 7, 15].

In this paper, we further generalize this approach to the conventional quasi-Newton algorithms to develop robust adaptive filtering algorithms in impulse noise environment. Particularly, a new robust quasi-Newton (R-QN) algorithm using the self-scaling variable metric (SSV) method for unconstrained optimization will be studied in details. Simulation results show that the R-QN algorithm is more robust to impulse noise in the desired signal than the RLS algorithm and other QN algorithm considered. Its initial convergence speed and tracking ability to sudden system change are also superior to those of the quasi-Newton algorithm proposed in [1].

The paper is organized as follows: A brief summary of the Quasi-Newton adaptive filtering algorithm is given in Section 2. The robust quasi-Newton algorithm is introduced in Section 3. Simulation result and comparison with other algorithms are described in Section 4. Finally, conclusions are drawn in Section 5.

2. QUASI-NEWTON ADAPTIVE FILTERING ALGORITHMS

Fig. 1 shows an adaptive linear transversal filter configured in a system identification application. Signals $x(n)$ and $y(n)$ are respectively the input and the output of the adaptive linear transversal filter. $d(n)$ is the desired signal, which consists of the output of the unknown system $d_s(n)$, and the additive interference $\eta_o(n)$, i.e. $d(n) = d_s(n) + \eta_o(n)$. The symbols n , N and the superscript T denote the time index, the filter length, and the transpose operator, respectively. The task of the adaptive filter is to minimize certain distortion measure between the output signal $y(n)$ and the desired signal $d(n)$. Therefore, the adaptive filtering problem can be formulated as a general unconstrained optimization problem. One frequently used distortion measure is the mean square error (MSE) as follows

$$\min J_{MSE}(\mathbf{w}) = \min E[e^2(n)], \quad (1)$$

$$\text{where } e(n) = d(n) - \mathbf{w}^T(n-1)\mathbf{X}(n), \quad (2)$$

$\mathbf{w}(n) = [w_1(n), \dots, w_N(n)]^T$, $\mathbf{X}(n) = [x(n), \dots, x(n-N+1)]^T$ and $J_{MSE}(\mathbf{w})$ are the estimation error, the filter coefficient vector, the input signal vector, and the MSE cost function, respectively. The Newton adaptive filtering algorithm is developed from the Wiener-Hopf equation and the filter coefficient vector is updated as [9]

$$\mathbf{w}(n) = \mathbf{w}(n-1) - \mu \mathbf{R}^{-1} \mathbf{g}(n-1), \quad (3)$$

where μ is the stepsize ($0 < \mu < 1$), \mathbf{R} is the $(N \times N)$ ensemble autocorrelation matrix of $x(n)$, and $\mathbf{g}(n)$ is the gradient vector of the cost function $J_{MSE}(\mathbf{w})$. In practice, \mathbf{R} and $\mathbf{g}(n)$ are not known in advance and they have to be estimated continuously. From an adaptive filtering point of view, different estimation method for \mathbf{R}^{-1} and $\mathbf{g}(n)$ constitute different Newton-type adaptive filtering algorithms, which have the following general form

$$\mathbf{w}(n) = \mathbf{w}(n-1) - \mu(n) \hat{\mathbf{R}}^{-1}(n-1) \hat{\mathbf{g}}(n-1), \quad (4)$$

where $\hat{\mathbf{R}}^{-1}(n)$ and $\hat{\mathbf{g}}(n)$ are the estimate of \mathbf{R}^{-1} and $\mathbf{g}(n)$, respectively. The stepsize $\mu(n)$ can either be time varying or constant. For example, one gets the LMS-Newton algorithm [10] if a constant stepsize ($\mu(n) = \mu$) and the following instantaneous gradient

$$\hat{\mathbf{g}}(n-1) = \hat{\mathbf{g}}_{MSE}(n-1) = -2e(n)\mathbf{X}(n), \quad (5)$$

of $J_{MSE}(w)$ are used. Substituting (5) into (4) yields the filter coefficient update equation of the LMS-Newton algorithms as follows [10, 11]

$$w(n) = w(n-1) + 2\mu e(n) \hat{R}^{-1}(n-1) X(n), \quad (6)$$

Various methods have been proposed to estimate $\hat{R}^{-1}(n)$ in (6) or solving the system $\hat{R}^{-1}(n-1)X(n)$, which constitutes the major complexity of the Newton-type algorithms. In the conventional RLS algorithm, $\hat{R}^{-1}(n)$ is recursively estimated from $X(n)$ using the inverse matrix lemma [12]

$$R^{-1}(n) = \frac{1}{\lambda} \left[R^{-1}(n-1) - \frac{R^{-1}(n-1)X(n)X^T(n)R^{-1}(n-1)}{\lambda + X^T(n)R^{-1}(n-1)X(n)} \right], \quad (7)$$

where λ is a forgetting factor. The RLS algorithm has a computational complexity of order $O(N^2)$ and provides the fastest convergence speed. Unfortunately, the RLS algorithm suffers from numerical problem when implemented in fixed-point or floating-point arithmetic [1, 12]. In [1], a new quasi-Newton (QN) algorithm with better numerical behavior was proposed, based on Fletcher optimization technique. More specifically, the coefficient vector of the QN algorithm is updated as follows:

$$t(n) = \hat{R}^{-1}(n-1)X(n), \quad (8)$$

$$v(n) = X^T(n)t(n), \quad \mu(n) = 1/(2v(n)), \quad (9)$$

$$\hat{R}^{-1}(n) = \hat{R}^{-1}(n-1) + (\mu(n)-1)t(n)t^T(n)/v(n), \quad (10)$$

$$w(n) = w(n-1) + e(n)t(n)/v(n), \quad (11)$$

where $e(n)$ is given by (2). Noted that, in the development of the QN algorithm, $g(n-1)$ is still estimated by (5), but in order to obtain the available regression for the adaptive filter, $g(n)$ is approximately estimated by

$$g(n) = \tilde{g}_{MSE}(n) = -2\varepsilon(n)X(n) \quad (\varepsilon(n) = d(n) - w^T(n)X(n)), \quad (12)$$

where $\varepsilon(n)$ is the posterior error. Finally, a linear search is performed along the direction $-\hat{R}^{-1}(n-1)\hat{g}(n-1)$ to minimize $\varepsilon^2(n)$ with respect to the time-varying stepsize $\mu(n)$. It can be shown that the update of $\hat{R}^{-1}(n)$ in (10) and (7) correspond to an update of $\hat{R}(n)$ respectively as follows [1, 12]

$$\hat{R}(n) = \hat{R}(n-1) + 2[1 - \mu(n)]X(n)X^T(n), \quad (\text{for QN}) \quad (13)$$

$$\hat{R}(n) = \lambda\hat{R}(n-1) + X(n)X^T(n), \quad (\text{for RLS}) \quad (14)$$

where $\mu(n) = 1/(2X^T(n)R^{-1}(n-1)X(n))$. Comparing (13) with (14), it can be seen that the $\hat{R}(n)$ in (13) will continuously accumulate its previous value and relies on the second term on the right hand side of (13) to give the correct estimate. In contrast, the $\hat{R}(n)$ in (14) will gradually forget the effects of its previous value. This makes the tracking ability of QN algorithm inferior to the RLS algorithm.

Both the LMS-Newton or the QN algorithms are not robust to impulse noise since they are all developed based on the minimization of the MSE cost function [2, 4, 6-8]. When there are impulses in the desired signal $d(n)$ or the input signal $x(n)$, $\hat{g}_{MSE}(n-1)$ in (5) will be corrupted by the impulse leading to a fault gradient with a large value. The corresponding $\mu(n)$ along this

'wrong' direction will further exaggerate the effect of the impulse on the filter coefficient vector $w(n)$ and the update of $\hat{R}^{-1}(n)$. Therefore, the performance of the QN algorithm will be significantly degraded.

3. ROBUST QUASI-NEWTON (R-QN) ADAPTIVE FILTERING ALGORITHM

In this section, we develop a new robust quasi-Newton (R-QN) algorithm based on the self-scaling variable metric (SSV) method [13, 14] and the robust statistics concept [6, 7, 15]. It has better robustness against impulse noise in the desired signal and better tracking ability to sudden system change than the QN algorithm in [1]. In what follows, the SSV method [13, 14] will be introduced. Then, the proposed R-QN algorithm will be described in details.

3.1 The SSV method [13, 14]

Starting with any positive definite matrix $\hat{R}^{-1}(0)$, $w(0) = \theta$. The n th iteration of the SSV method is given as follows:

Step One: set the searching direction $s(n)$ as

$$s(n) = -\hat{R}^{-1}(n-1)g(n-1), \quad (15)$$

Step Two: perform a linear search along $s(n)$ to determine the new coefficient vector

$$w(n) = w(n-1) + \mu(n)s(n), \quad (16)$$

Step Three: update $R^{-1}(n)$ from $R^{-1}(n-1)$

$$R^{-1}(n) = \kappa \left(R^{-1} - \frac{R^{-1}\Delta g \Delta g^T R^{-1}}{\Delta g^T R^{-1} \Delta g} \right) + \frac{\Delta w \Delta w^T}{\Delta w^T \Delta g} + \gamma v v^T \quad (17)$$

$$\text{and } v = (\Delta g^T R^{-1} \Delta g)^{1/2} \left(\frac{\Delta w}{\Delta w^T \Delta g} - \frac{R^{-1} \Delta g}{\Delta g^T R^{-1} \Delta g} \right), \quad (18)$$

where κ and γ are two parameters chosen to improve the approximation of the Hessian [13]. For clarity of presentation, in the right hand of equations (17) and (18), we have used the abbreviations $\Delta g \triangleq g(n) - g(n-1)$, $\Delta w \triangleq w(n) - w(n-1)$, and $R^{-1} = R^{-1}(n-1)$. As we shall see later in the simulation in section 4, the SSV provides a convenient tradeoff between tracking ability and steady state error for the QN algorithm by controlling the parameters κ and γ . In principle, QN algorithms based on the Fletcher or SSV formulas can be applied to nonlinear objective functions if the gradient in (4) can be estimated.

3.2 The Robust Quasi-Newton Algorithm

We can see that the various derivations of the QN algorithm are based on (3) and the use of the instantaneous gradient of the conventional MSE cost function as shown in (5). To provide more robustness to impulse noise, we estimate the gradient vector $g(n)$ by minimizing the mean M -estimate cost function $J_\rho(w) = E[\rho(e(n))]$ [7] instead of $J_{MSE}(w) = E[e^2(n)]$, where $\rho(e(n))$ is the Hampel's three-part redescending M -estimate function [7]. Then, the corresponding gradient vector $\hat{g}_\rho(n-1)$ can be computed as

$$\hat{g}_\rho(n-1) = \frac{\partial J_\rho(w)}{\partial w(n-1)} = E \left[\frac{\partial \rho(e(n))}{\partial e(n)} \frac{\partial e(n)}{\partial w(n-1)} \right] \approx -q(e(n))e(n)X(n), \quad (19)$$

and

$$q(e) = \frac{d\rho(e)}{de} / e = \begin{cases} 1, & 0 < |e| < \xi \\ \xi \operatorname{sgn}(e) / e, & \xi \leq |e| < \Delta_1 \\ \frac{\operatorname{sgn}(e)}{e} \left[(|e| - \Delta_2) \frac{\xi}{\Delta_1 - \Delta_2} \right], & \Delta_1 \leq |e| < \Delta_2 \\ 0, & \Delta_2 \leq |e| \end{cases} \quad (20)$$

where $q(e(n))$ is the weighting function of the Hampel's three part redescending M -estimate function $\rho(e(n))$ [7]. The reason for choosing the M -estimate function is to de-emphasize the contributions of abnormally large instantaneous error on the objective function. From (20), it can be seen that the weighting function decreases gradually from 1 to 0 when $e(n)$ is increased from ξ to Δ_2 . In fact, when $e(n)$ is smaller than ξ , the weight function $q(e(n))$ is equal to one and $\hat{g}_p(n-1)$ in (19) becomes identical to $\hat{g}_{MSE}(n-1)$ in (5). When $e(n)$ is larger than ξ , $q(e(n))$ starts to reduce and is equal to zero when $|e(n)| > \Delta_2$. As a result, the adverse effect of the impulse on $\hat{g}_p(n-1)$ will be suppressed, if these thresholds are properly chosen. In [6, 7], we proposed to estimate these thresholds continuously using the following formulas

$$\begin{aligned} \xi &= k_\xi \hat{\sigma}(n) = 1.96 \hat{\sigma}(n), \quad \Delta_1 = k_{\Delta_1} \hat{\sigma}(n) = 2.24 \hat{\sigma}(n), \\ \Delta_2 &= k_{\Delta_2} \hat{\sigma}(n) = 2.576 \hat{\sigma}(n), \end{aligned} \quad (21)$$

where $\hat{\sigma}(n)$ is the impulse-free variance of the estimation error which can be estimated as follows [7]

$$\hat{\sigma}^2(n) = \lambda_\sigma \hat{\sigma}^2(n-1) + 1.483(1 + 5/(N_w - 1))(1 - \lambda_\sigma) \operatorname{med}(A_e(n)), \quad (22)$$

where $A_e(n) = \{e^2(n), \dots, e^2(n - N_w + 1)\}$, N_w is the length of the estimation window, and λ_σ is the forgetting factor. The basic idea is to remove the impulsive components in the error signal through the median filtering operation so that an impulse-free variance of the estimation error can be obtained. Assuming that this impulse-free error signal to be Gaussian distributed, it is possible to determine the required thresholds to reject or de-emphasize the impulses with different degree of confidence. The robustness of the resulting algorithm is not sensitive to the choice of the constants k_ξ , k_{Δ_1} , and k_{Δ_2} provided that they are not too far away from those suggested here. The constant 1.483 in (22) is used to provide more accurate estimation of the variance if the input is Gaussian distributed. Interested readers are referred to [6, 7] for more details. Substituting (19) into (4) yields

$$w(n) = w(n-1) + q(e(n))\mu(n)e(n)R^{-1}(n-1)X(n). \quad (23)$$

Together with the method that we have proposed in [15] for suppressing input impulses, (23) serves as a framework for developing various Quasi-Newton algorithms for robust adaptive filtering in impulse noise. In what follows, we shall focus ourselves on the development of a new robust QN algorithm using the SSV method illustrated in Section 2.1

First of all, we consider the determination of the stepsize $\mu(n)$ in the linear search along the direction $s_p(n) = -\hat{R}^{-1}(n-1)\hat{g}_p(n-1)$ in Step Two of the SSV algorithm. From (19), it can be seen that when $e(n)$ is larger than Δ_2 , both $q(e(n))$ and $\hat{g}_p(n-1)$ will be zero. In this case, the weight vector is not updated, because of the possible risk of having an impulse in the desired or input signals. For other

situation, a linear search along the direction $s_p(n)$ is performed to minimize $\rho(e(n))$ with respect to $\mu(n)$ so that its optimal value can be determined as follows

$$\mu(n) = \frac{1}{q(e(n))X^T(n)R^{-1}(n-1)X(n)} \approx \frac{1}{q(e(n))\tau(n)}. \quad (24)$$

where $\tau(n) = X^T(n)R^{-1}(n-1)X(n)$. Here, we have approximated $q(e(n))$ by $q(e(n))$. To provide a more stable estimation against the impulse noise, the actual stepsize $\mu(n)$ is obtained by multiplying (24) with a stepsize parameter μ_c as follows

$$\mu(n) = \mu_c / (q(e(n))\tau(n)), \quad \text{with } 0 < \mu_c \leq 1. \quad (25)$$

Also, from (12) and (23), Δg and Δw in (17) and (18) can be estimated respectively as follows

$$\Delta \hat{g} = \tilde{g}_p(n) - \hat{g}_p(n-1) \approx q(e(n))(e(n) - \varepsilon(n))X(n), \quad (26)$$

$$\Delta \hat{w} = \mu(n)q(e(n))e(n)R^{-1}(n-1)X(n). \quad (27)$$

This completes the estimation of the matrix inverse $R^{-1}(n)$ in Step Three of the SSV algorithm. Inserting (26) and (27) into (17) and (18), we finally obtained the following recursion for the proposed R-QN algorithm

$$e(n) = d(n) - w^T(n-1)X(n), \quad (28)$$

$$A(n) = R^{-1}(n-1)X(n), \quad (29)$$

$$\tau(n) = X^T(n)A(n), \quad (30)$$

if $q(e(n)) = 0$, then

$$R^{-1}(n) = R^{-1}(n-1), \quad w(n) = w(n-1), \quad (31)$$

else if $q(e(n)) \neq 0$, then

$$R^{-1}(n) = \kappa R^{-1}(n-1) + \frac{1 - q(e(n))\kappa\tau(n)}{q(e(n))\tau^2(n)} A(n)A^T(n) \quad (32)$$

$$w(n) = w(n-1) + \mu_c e(n)A(n) / \tau(n), \quad (33)$$

where $q(e(n))$ is given by (20), and κ is chosen as $1/\lambda$, as suggested in [13], where λ is a forgetting factor. Comparing the R-QN algorithm to the QN algorithm given in Section 2, it can be seen that the arithmetic complexity of the R-QN algorithm is comparable to that of the QN algorithm, except for $O(N_w \log N_w)$ more operations required for computing $q(e(n))$ [6].

4. SIMULATIONS

The performance of the proposed R-QN algorithm is evaluated and compared with the RLS and the QN algorithms in impulse noise environment. The adaptive filter shown in Fig. 1 is used to identify the unknown system w^* , which is a 9th order lowpass FIR filter with coefficients $w^* = [2, -4, 6, -8, 1, -8, 6, -4, 2]^T$ and it is changed to $-w^*$ at $n=2500$ to evaluate its tracking ability to sudden system change. The input signal $x(n)$ is generated by passing a zero-mean, unit variance white Gaussian process through a linear time-invariant filter with coefficients [3887, 1, 3887] [12]. The output of the unknown system $d_s(n)$ is corrupted by the additive noise $\eta_s(n)$, which is modeled as the frequently used contaminated Gaussian (CG) noise, $\eta_s(n) = \eta_g(n) + b(n)\eta_w(n)$ [2, 6]. In fact, $\eta_g(n)$ and $\eta_w(n)$ are independent identically distributed (i.i.d.) zero mean Gaussian noises with variance σ_g^2 and σ_w^2 , respectively, $b(n)$ is an i.i.d

Bernoulli random variable with occurrence probability $P_r(b(n)=1)=p_r$. The ratio $\gamma_{in} = \sigma_{in}^2 / \sigma_s^2 = p_r \sigma_w^2 / \sigma_s^2$ determines the impulsive characteristic of $\eta_o(n)$. For fixed value of σ_s^2 , the larger the γ_{in} , the more impulsive $\eta_o(n)$ becomes. The signal-to-noise ratio at the system output is defined as $SNR = 10 \log_{10}(\sigma_d^2 / \sigma_s^2)$, where σ_d^2 is the variance of $d_s(n)$. The initial values of $R^{-1}(0)$ and $w(0)$ are set to identity matrix I and zero vector, respectively. Simulation parameters and the initial values for various algorithms are shown in Fig. 2. For illustration purpose, from $n=1$ to 999, $\eta_o(n)=\eta_s(n)$ is used. Whereas from $n=1000$ to 1950, $\eta_o(n)=\eta_s(n)+b(n)\eta_w(n)$ with $p_r=0.005$ and $r_{in}=300$ 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. The MSE results averaged over 200 independent runs are plotted in Fig. 2. It can be seen that the RLS and the QN algorithms are not robust to any impulses in the desired signal. The effect of a single impulse will last for more than 200 and 350 iterations for the RLS and the QN algorithms, respectively. In contrast, the performance of the R-QN algorithm is very robust to the added impulses. Additionally, the initial convergence and the tracking ability to the system sudden change of the R-QN algorithm are superior to those of the QN algorithm although they are slightly inferior to those of the RLS algorithm. This can be viewed as a reasonable tradeoff between the better numerical stability and the robustness to impulse noise and the convergence speed and tracking ability.

5. CONCLUSION

A new robust quasi-Newton (R-QN) adaptive filtering algorithm for impulse noise suppression is presented. It is developed on the self-scaling variable metric (SSV) method and the robust statistics concept. An M -estimate based cost function is minimized instead of the commonly used mean square error (MSE) to suppress the adverse effect of the impulse noise on the filter coefficients. Simulation results show that the proposed R-QN algorithm is more robust to impulse noise in the desired signal than the RLS algorithm and other QN algorithm considered. Its initial convergence speed and tracking ability to sudden system change are also superior to those of the quasi-Newton algorithm proposed in [1].

6. REFERENCES

- [1] M.L.R. De Campos, A. Antoniou, "A New Quasi-Newton Adaptive Filtering Algorithm," *IEEE Transactions on Circuits and Systems II: Analog and Digital Signal Processing*, vol. 44, No. 11, pp. 924-934, 1997.
- [2] T. I. Haweel and P. M. Clarkson, "A Class of Order Statistic LMS Algorithms," *IEEE Transactions on Signal Processing*, vol. 40, No. 1, pp. 44-53, 1992.
- [3] R. Settineri, M. Najim, and D. Ottaviani, "Order Statistic Fast Kalman Filter," *IEEE ISCAS-96*, vol. 2, pp. 116-119, 1996.
- [4] S. Koike, "Adaptive Threshold Nonlinear Algorithm for Adaptive Filters with Robustness against Impulsive Noise," *IEEE Transactions on Signal Processing*, vol. 45, No. 9, pp. 2391-2395, 1997.
- [5] J. F. Weng and S. H. Leung, "Adaptive Nonlinear RLS Algorithm for Robust Filtering in Impulsive Noise," *IEEE ISCAS-97, Hong Kong*, vol. 4, pp. 2337-2340, 1997.

- [6] Y. Zou, S. C. Chan, and T. S. Ng, "A Recursive Least M-Estimate (RLM) Adaptive Filter For Robust Filtering in Impulsive Noise," *IEEE Signal Processing Letters*, vol. 7, No. 11, pp. 5-8, 2000.
- [7] Y. Zou, S. C. Chan, and T. S. Ng, "Least Mean M-Estimate Algorithms for Robust Adaptive Filtering in Impulse Noise," *International European Signal Processing Conference (EUSIPCO-2000)*, Tampere, Finland, 2000.
- [8] R. E. Frank and M. Hampel, *Robust Statistics: The Approach Based on Influence Functions*, New York: John Wiley & Sons Inc., 1986.
- [9] B. Widrow and S. D. Stearns, *Adaptive Signal Processing*, Englewood Cliffs, N.J.: Prentice-Hall, 1985.
- [10] F. B. Boroujeny, *Adaptive Filters: Theory and Applications*, John Wiley & Sons, 1999.
- [11] D. F. Marshall and W. K. Jenkins, "A Fast Quasi Newton Adaptive Filtering Algorithm," *IEEE Transactions On Signal Processing*, vol. 40, No. 7, pp. 1652-1662, 1992.
- [12] S. Haykin, *Adaptive Filter Theory*, Prentice-Hall, 2nd edition, 1991.
- [13] R. Fletcher, *Practical Methods of Optimization*, Chichester [etc.]: Wiley, 1980.
- [14] S. S. Oren, "On the Selection of Parameters in Self Scaling Variable Metric Algorithms," *Mathematical Programming*, vol. 7, pp. 351-367, 1974.
- [15] Y. Zou, S. C. Chan, and T. S. Ng, "A Robust Statistics Based Adaptive Lattice-Ladder Filter in Impulsive Noise," *IEEE ISCAS-2000*, Geneva, Switzerland, 2000.

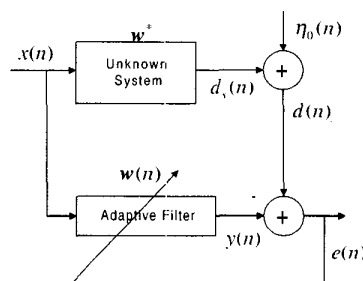


Fig. 1 System Identification Structure

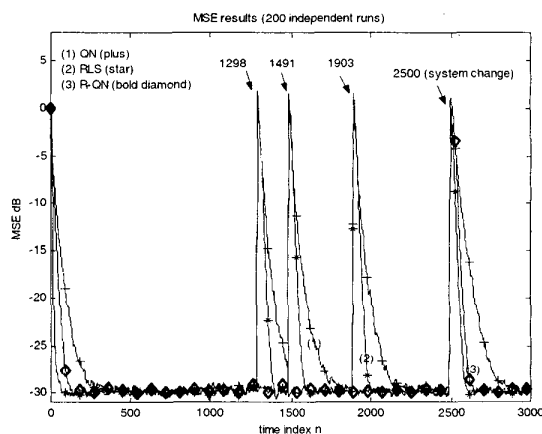


Fig. 2. The MSE performance of the various algorithms with impulse noise. (1) QN (with plus sign); (2) RLS (with stars); (3) R-QN (with bold diamonds). $N = 9$, $SNR = 30dB$, $\mu_r = 0.3$, $\lambda = \lambda_\sigma = 0.95$, $\sigma^2(0) = d^2(0)$ and $N_r = 5$ (the marked numbers (1298, 1491 and 1903) indicate the locations of the impulses in the desired).