Next, the overall loop $T_3, T_{10}, T_1, T_2, T_{11}, T_4$ is analyzed

$$(I + I_o) \cdot [I_{\text{in}} + I_{\text{out}} - I_o - I_{d9} - I_{d13}] \cdot I_o = I_{\text{out}} \cdot [i_c + (I + I_o) - I_o - I_{d12}] \cdot I_o.$$

Substituting (21), reduces to the simple result

$$I \cdot I_{\text{in}} = i_c \cdot I_{\text{out}}. \tag{22}$$

The capacitor voltage V_c is the difference of the gate-source voltages of T_4 and T_{11} which have equal gate voltages. It follows from (5) that V_c is given by $U_T \ln(I_{\text{out}}/I_o)$. The capacitor current $i_c = C \cdot dV_c/dt$ therefore becomes

$$i_c = \frac{CU_T}{I_{\text{out}}} \cdot \frac{dI_{\text{out}}}{dt}. \tag{23}$$

Substituting (23) into (22) yields the same linear integrator function as before (20):

$$I_{\text{out}} = \frac{I}{CU_T} \int I_{\text{in}} \cdot dt. \tag{24}$$

It follows that the unity-gain bandwidth is

$$f_{\text{UG}} = \frac{I}{2\pi CU_T} \tag{25}$$

which is tunable by dc bias current I. This result is independent of the body-effect. The class AB companding integrator circuit proposed in [9] can be modified for minimum supply voltage in the same way as described here.

VI. Conclusion

The present drive toward low-voltage, low-power electronics will require translinear circuits capable of operating at minimum supply voltage. This is made possible by exploiting the symmetry of nonsaturated MOS transistors, together with the exponential MOS-characteristic in weak inversion. A technique of developing translinear circuits for low-supply voltage by allowing transistors to go below saturation was proposed. The forward and reverse modes then become equivalent to saturated transistors. This transforms all TL loops into alternating loops, providing two improvements. First, operation is extended to low-supply voltage. Second, accurate realization in a single substrate is enabled since the $I_{\square}(V_g)$-terms which represent the body-effect will cancel for the oppositely connected transistor-pairs in the loops.

The following three basic circuit topologies were modified for low-supply voltage operation: 1) the balanced TL loop; 2) the alternating TL loop; and 3) the instantaneous-companding integrator. The minimum value of supply voltage required for these circuit structures is given by the sum of the transistor threshold voltage and the drain-source saturation voltage. Since the circuits will operate in weak inversion, bandwidth will be limited and the circuits will be sensitive to threshold matching.

The techniques proposed in this paper are presently being developed further for low-voltage static and dynamic analog signal processing. Important properties with respect to bandwidth, noise, and errors due to transistor mismatching need to be studied. Experimental results of fabricated circuits employing these techniques will be published in a future paper.

References

[1] E. Vittoz and J. Fellrath, "CMOS analog integrated circuits based on weak inversion operation," *IEEE J. Solid-State Circuits*, vol. SC-12, pp. 224–231, June 1977.

[2] B. Gilbert, "Translinear circuits: A proposed classification," *Electron. Lett.*, vol. 11, pp. 14–16, Jan. 1975.

[3] E. Seevinck, "CMOS translinear circuits," in *Analog Circuit Design: MOST RF Circuits, Sigma–Delta Converters, and Translinear Circuits*, W. Sansen, R. J. v. d. Plassche, and J. H. Huijsing, Eds. Norwell, MA: Kluwer, 1996.

[4] A. G. Andreou and K. A. Boahen, "Translinear circuits in subthreshold MOS," *IEEE J. Analog Integr. Circuits Signal Processing*, vol. 9, pp. 141–166, 1996.

[5] T. Serrano-Gotarredona, B. Linares-Barranco, and A. G. Andreou, "A general translinear principle for subthreshold MOS transistors," *IEEE Trans. Circuits Syst. I*, vol. 46, pp. 607–616, May 1999.

[6] C. C. Enz, F. Krummenacher, and E. A. Vittoz, "An analytical MOS transistor model valid in all regions of operation and dedicated to low-voltage and low-current applications," *J. Analog Integr. Circuits Signal Processing*, vol. 8, pp. 83–114, 1995.

[7] M. Punzenberger and C. C. Enz, "A 1.2 V low-power BiCMOS class AB log-domain filter," *IEEE J. Solid-State Circuits*, vol. 32, pp. 1968–1978, Dec. 1997.

[8] B. Gilbert, "A new wide-band amplifier technique," *IEEE J. Solid-State Circuits*, vol. SC-3, pp. 353–365, Dec. 1968.

[9] E. Seevinck, "Companding current-mode integrator: A new circuit principle for continuous-time monolithic filters," *Electron. Lett.*, vol. 26, pp. 2046–2047, Nov. 1990.

Least Mean M-Estimate Algorithms for Robust Adaptive Filtering in Impulse Noise

Yuexian Zou, Shing-Chow Chan, and Tung-Sang Ng

***Abstract*—This paper proposes two gradient-based adaptive algorithms, called the least mean M-estimate and the transform domain least mean M-estimate (TLMM) algorithms, for robust adaptive filtering in impulse noise. A robust M-estimator is used as the objective function to suppress the adverse effects of impulse noise on the filter weights. They have a computational complexity of order $O(N)$ and can be viewed, respectively, as the generalization of the least mean square and the transform-domain least mean square algorithms. A robust method for estimating the required thresholds in the M-estimator is also given. Simulation results show that the TLMM algorithm, in particular, is more robust and effective than other commonly used algorithms in suppressing the adverse effects of the impulses.**

***Index Terms*—Adaptive filter, impulse noise suppression, least mean M-estimate algorithm (LMM), orthogonal transform, system identification, robust statistics.**

I. Introduction

The performance of conventional linear adaptive filtering algorithms can deteriorate significantly when the desired or the input signal is corrupted by impulse noise [1]–[8]. Nonlinear techniques are usually employed to reduce the adverse effects due to impulse noise. For example, in the order statistic least mean square (OSLMS) [1] and the order statistic recursive least square (OSRLS) [2] algorithms, the estimation of the instantaneous gradient weight vector is replaced by the

Manuscript received August 2000; revised October 2000. This work was supported by the Hong Kong Research Grants Council and by the CRCG of The University of Hong Kong. This paper was recommended by Associate Editor J. Chambers.

The authors are with the Department of Electrical and Electronic Engineering, The University of Hong Kong, Pokfulam Road, Hong Kong (e-mail: yxzou@eee.hku.hk; scchan@eee.hku.hk; tsng@eee.hku.hk).

Publisher Item Identifier S 1057-7130(00)11665-9.

output of the median filter. Another class of nonlinear techniques relies on nonlinear functions to limit the transient fluctuation in the estimation error. Typical examples in this class are the adaptive threshold nonlinear algorithm (ATNA) [3] and the nonlinear RLS (N-RLS) [4] algorithm. In the ATNA, a clipper function is applied to the error signal in the LMS algorithm to reduce its influence on the filter weights when the error signal is abnormally large, while in the N-RLS, a Huber function is applied to the error signal in the RLS algorithm. In [5], a robust mixed-norm (RMN) adaptive algorithm using a combination of L_1 and L_2 norms as the objective function and the stochastic gradient method was proposed.

Recently, a new class of adaptive filtering algorithms based on the concept of robust statistics has been proposed [6], [7]. Instead of minimizing the weighted least squares error criterion $J_{\rm LS}(n) = \sum_{i=1}^{n} \lambda^{n-i} e^2(i)$, where $0 < \lambda \leq 1$ is the forgetting factor and $e(i)$ is the estimation error, a weighted least M-estimate function criterion $J_\rho(n) = \sum_{i=1}^{n} \lambda^{n-i} \rho(e(i))$ was proposed. In [6], $\rho(\cdot)$ is chosen as a modified Huber function [9] and an RLS-liked algorithm, called the M-RLS algorithm, is derived. Later, in [7], using the more general Hampel's three-part redescending M-estimate function [10], the recursive least M-estimate (RLM) algorithm was derived. Simulation results and mean convergence analysis showed that these two algorithms are effective in suppressing the adverse effects of impulse noise. The convergence speed and steady-state error are relatively unaffected by impulses, and the performance is similar to that of the RLS algorithm in Gaussian noise alone. However, the main drawback of such RLS-like algorithms is the large computational complexity of $O(N^2)$ per iteration.

Motivated by the robustness of the M-RLS and RLM algorithms, we propose in this paper a new family of gradient-based adaptive algorithms for suppressing impulses with a lower computational complexity of order $O(N)$.

Instead of using the mean square objective function $J_{\rm MSE} = E[e^2(n)]$ to develop the LMS-type algorithms, the mean M-estimate error objective function $J_{M\rho} \triangleq E[\rho(e(n))]$ is used. Here, $E[\cdot]$ and $\rho(\cdot)$ are, respectively, the expectation operator and the M-estimate function. Two stochastic gradient based algorithms, called the least mean M-estimate (LMM) and the transform-domain least mean M-estimate (TLMM) algorithms, are derived. They can be viewed, respectively, as the generalization of the conventional LMS and the transform-domain LMS (TLMS) algorithms. The proposed algorithms have a computational complexity of order $O(N)$. Simulation results show that they are more robust than the ATNA, OSLMS, and RMN algorithms in impulse noise environment.

This paper is organized as follows. The LMM algorithm and the TLMM algorithm are derived in Sections II and III, respectively. Section IV briefly describes the threshold parameter estimation and the computational complexity of the LMM and TLMM algorithms. Their performances are evaluated and compared with other algorithms by simulations in Section V. Conclusions are drawn in Section VI.

II. Least Mean M-Estimate (LMM) Algorithm

Let us consider the system identification problem shown in Fig. 1. The signals $x(n)$ and $y(n)$ are, respectively, the input and output signals of the linear transversal adaptive filter. The estimation error is given by $e(n) = d(n) - \boldsymbol{w}^t(n-1)\boldsymbol{x}(n)$, where $\boldsymbol{w}(n) = [w_1(n), w_2(n), \ldots, w_N(n)]^t$ and $\boldsymbol{x}(n) = [x(n), x(n-1), \ldots, x(n-N+1)]^t$ are the weight vector and the input signal vector, respectively. Signal $d(n)$ is the reference or the desired signal, and the superscript t is the transpose operator. In practical applications, $x(n)$ and $d(n)$ may be corrupted by interference signals $\eta_s(n)$ and $\eta_o(n)$, respectively, which can be modeled

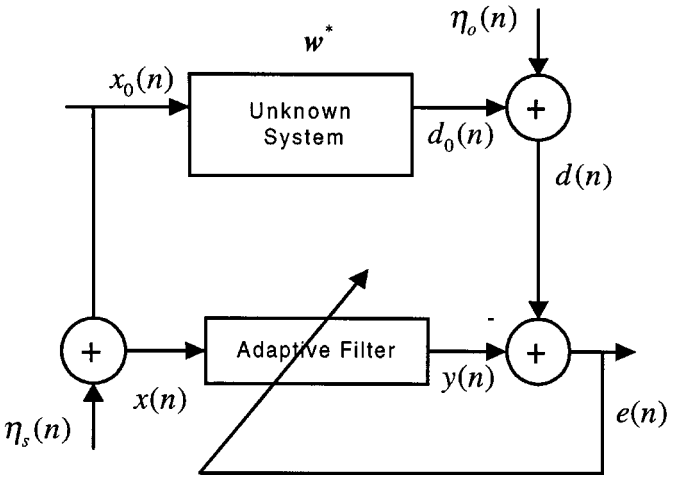


Fig. 1. System identification structure.

as contaminated Gaussian (CG) noise or alpha-stable noise. Based on robust statistical estimation, the following objective function for adaptive filter is proposed:

$$J_{M\rho} \triangleq E[\rho(e(n))] \tag{1}$$

where $\rho(\cdot)$ is a robust M-estimate function for suppressing impulse noise. In this paper, $\rho(\cdot)$ is chosen as the more general Hampel's three-part redescending M-estimate function, which is given as follows [10]:

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

where ξ, Δ_1 and Δ_2 are the threshold parameters. Function $\rho(\cdot)$ is a real-valued even function. The advantage of this M-estimate function is that it is a piecewise approximation of the maximum likelihood estimator when the input and additive noises are modeled as a contaminated Gaussian noise. Moreover, it reduces to the modified Huber function when Δ_1 is equal to Δ_2, which was studied in [6]. The optimal weight vector can be determined by setting the first-order partial derivatives of $J_{M\rho}$ in (1), with respect to $\boldsymbol{w}$, to zero. This yields

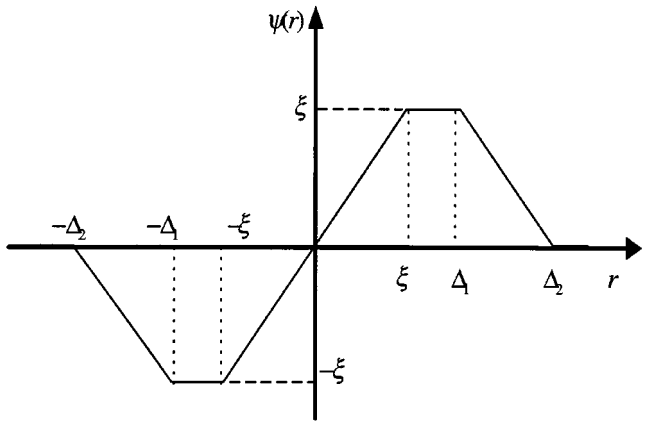
$$E[\psi(e(n))\boldsymbol{x}(n)] = 0 \tag{3}$$

where $\psi(e) \triangleq \partial\rho(e)/\partial e$ is called the score function, which is illustrated in Fig. 2. For notational convenience, we define the weight function $q(e)$ as $q(e) \triangleq \psi(e)/e$. Substituting $e(n) = d(n) - \boldsymbol{w}^t\boldsymbol{x}(n)$ into (3) and after some manipulations, the following M-estimate normal equation is obtained:

$$\boldsymbol{R}_{\boldsymbol{X}\rho}\boldsymbol{w} = \boldsymbol{P}_\rho \tag{4}$$

where

$$\boldsymbol{R}_{\boldsymbol{X}\rho} \triangleq E[q(e(n))x(n)\boldsymbol{x}^t(n)]$$



$$\psi(e)=\begin{cases} e, & 0\le |e|<\xi \\ \xi\,\mathrm{sgn}(e), & \xi\le |e|<\Delta_1 \\ \left[(|e|-\Delta_2)\dfrac{\xi}{\Delta_1-\Delta_2}\right]\mathrm{sgn}(e), & \Delta_1\le |e|<\Delta_2 \\ 0, & \Delta_2\le |e| \end{cases}.$$

Fig. 2. Score function of Hampel's three-part redescending M-estimate function.

and

$$\boldsymbol{P}_\rho \triangleq E[q(e(n))d(n)\boldsymbol{x}(n)] \tag{5}$$

are the M-estimate correlation matrix of $\boldsymbol{x}(n)$ and the M-estimate cross-correlation vector of $d(n)$ and $\boldsymbol{x}(n)$, respectively. They serve a similar purpose as the conventional correlation matrix of $\boldsymbol{x}(n)$ and the cross-correlation vector of $\boldsymbol{x}(n)$ and $d(n)$ in the Wiener filter. Since $q(e(n))$ depends on $\boldsymbol{w}$, (4) is a system of nonlinear equations. In what follows, two robust stochastic gradient-based adaptive algorithms with $O(N)$ computational complexity are derived to solve this normal equation.

In the proposed LMM algorithm, the mean M-estimate objective function $J_{M\rho}$ in (1) is minimized by updating the weight vector $\hat{\boldsymbol{w}}(n)$ in the negative direction of the gradient vector $\nabla(J_{M\rho})$ of $J_{M\rho}$, which can be approximated by

$$\begin{aligned}\nabla_w(J_{M\rho}) &= \frac{\partial J_{M\rho}}{\partial \boldsymbol{w}} \\ &\approx \hat{\nabla}_{w\rho} \\ &= \frac{\partial}{\partial \boldsymbol{w}}(\rho(e(n)) \\ &= -q(e(n))e(n)x(n).\end{aligned} \tag{6}$$

The weight vector is then updated as

$$\hat{\boldsymbol{w}}(n)=\hat{\boldsymbol{w}}(n-1)-\mu\hat{\nabla}\boldsymbol{w}_\rho=\hat{\boldsymbol{w}}(n-1)+\mu q(e(n))e(n)x(n). \tag{7}$$

where μ is the step size parameter. Equation (7) can be viewed as a generalization of the LMS algorithm and is called the least mean M-estimate algorithm. It can be seen that when $e(n)$ is smaller than ξ, the weight function $q(e(n))$ is equal to one and (7) becomes identical to the LMS algorithm. When $e(n)$ is larger than ξ, $q(e(n))$ starts to reduce and is equal to zero when $|e(n)|>\Delta_2$. Thus the LMM algorithm effectively reduces the effect of large signal error during the updating of the filter weights. The problem of how the threshold parameters ξ, Δ_1, and Δ_2 should be chosen will be addressed in Section IV.

III. Transform-Domain Least Mean M-Estimate (TLMM) Algorithm

The limitation of the LMS-type algorithms is their slow convergence speed, especially when the input signal is heavily colored. The TLMS algorithm proposed in [12] greatly improves the convergence speed of the LMS algorithm. Various aspects of the TLMS algorithm and the analysis of its performance can be found in [13]. The input signal $x(n)$ is first transformed by an $(N\times N)$ orthogonal matrix $\boldsymbol{Q}$ to produce N transform coefficients or outputs $x_T(n,\,i)$, $i=1,\ldots,N$. The subscript T indicates the variables in the transform domain. Each of the outputs is then normalized by its power estimate $p_i^2(n)$, which is 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. Let $\boldsymbol{X}_T(n)=[x_T(n,\,1),\ldots,x_T(n,\,N)]^t$ be the $(N\times 1)$ vector containing $x_T(n,\,i)$. Then, one gets

$$\boldsymbol{X}_{T_N}(n)=(\Lambda^2)^{-1/2}\boldsymbol{X}_T(n)=\Lambda^{-1}\boldsymbol{Q}\boldsymbol{X}(n) \tag{8}$$

where $\boldsymbol{X}_{T_N}(n)$ is the normalized input vector. Λ^2 is an $(N\times N)$ diagonal matrix whose $(i,\,i)$th element is equal to $p_i^2(n)$. This normalization is a simple and effective approach to reduce the eigenvalue spread of the autocorrelation matrix $\boldsymbol{R}_{E,\,\boldsymbol{X}}=E[\boldsymbol{x}(n)\boldsymbol{x}^t(n)]$ of $\boldsymbol{x}(n)$. From (8), the autocorrelation matrix of $\boldsymbol{X}_{T_N}(n)$ is

$$\begin{aligned}\boldsymbol{R}_{E,\,\boldsymbol{X}_{T_N}} &= E[\boldsymbol{X}_{T_N}(n)\boldsymbol{X}_{T_N}^t(n)] \\ &= \Lambda^{-1}\boldsymbol{Q}E[\boldsymbol{x}(n)\boldsymbol{x}^t(n)]\boldsymbol{Q}^t\Lambda^{-1} \\ &= \Lambda^{-1}\boldsymbol{Q}\boldsymbol{R}_{E,\,\boldsymbol{X}}\boldsymbol{Q}^t\Lambda^{-1}.\end{aligned} \tag{9}$$

If $\boldsymbol{Q}$ is chosen as the Karhunen–Loeve transform (KLT) of $\boldsymbol{x}(n)$ [13], we have $\boldsymbol{R}_{E,\,\boldsymbol{X}_{T_N}}=\boldsymbol{I}_{N\times N}$. It indicates that $\boldsymbol{R}_{E,\,\boldsymbol{X}}$ is diagonalized by $\boldsymbol{Q}$. Unfortunately, in practical applications, it is very computational expensive to compute the KLT. Some suboptimal transforms such as the discrete cosine transform (DCT), the discrete Fourier transform, and the discrete Hartley transform are generally employed.

We now consider the derivation of the transform domain least mean M-estimate algorithms. For an arbitrary weight vector $\boldsymbol{w}$, the gradient vector of $J_{\mathrm{MSE}}=E[e^2(n)]$ is

$$\nabla_{\boldsymbol{w}}(J_{\mathrm{MSE}})=-2E[e(n)\boldsymbol{x}(n)]=2\boldsymbol{R}_{E,\,\boldsymbol{X}}\boldsymbol{w}-2\boldsymbol{P}_E \tag{10}$$

where $\boldsymbol{R}_{E,\,\boldsymbol{X}}=E[\boldsymbol{x}(n)\boldsymbol{x}^t(n)]$ and $\boldsymbol{P}_E=E[d(n)\boldsymbol{x}(n)]$ are autocorrelation matrix of $\boldsymbol{x}(n)$ and cross-correlation vector between $d(n)$ and $\boldsymbol{x}(n)$, respectively. The optimal solution $\boldsymbol{w}_{\mathrm{opt}}$ of J_{MSE} satisfies

$$\boldsymbol{w}_{opt}=\boldsymbol{w}-\tfrac{1}{2}\boldsymbol{R}_{E,\,\boldsymbol{X}}^{-1}\nabla_w(J_{\mathrm{MSE}}). \tag{11}$$

Premultiplying both sides of (11) by $\boldsymbol{Q}$ and simplifying, one gets

$$\begin{aligned}\boldsymbol{Q}\boldsymbol{w}_{opt} &= \boldsymbol{Q}\boldsymbol{w}-\tfrac{1}{2}\boldsymbol{Q}\boldsymbol{R}_{E,\,\boldsymbol{X}}^{-1}\nabla_w(J_{\mathrm{MSE}}) \\ &= \boldsymbol{Q}\boldsymbol{w}-\tfrac{1}{2}\Lambda^{-1}(\Lambda^{-1}\boldsymbol{Q}\boldsymbol{R}_{E,\,\boldsymbol{X}}\boldsymbol{Q}^t\Lambda^{-1})^{-1} \\ &\quad\cdot(\Lambda^{-1}\boldsymbol{Q}\nabla_w(J_{\mathrm{MSE}})).\end{aligned} \tag{12}$$

Assuming that the autocorrelation matrix $\boldsymbol{R}_{E,\,\boldsymbol{X}}$ is approximately diagonalized by $\boldsymbol{Q}$ and normalized by the normalization process, we have $(\Lambda^{-1}\boldsymbol{Q}\boldsymbol{R}_{\boldsymbol{X}_\rho}\boldsymbol{Q}^t\Lambda^{-1})\approx\boldsymbol{I}_{N\times N}$ and $(\Lambda^{-1}\boldsymbol{Q}\boldsymbol{R}_{\boldsymbol{X}_\rho}\boldsymbol{Q}^t\Lambda^{-1})^{-1}\approx\boldsymbol{I}_{N\times N}$. Letting $\boldsymbol{w}_T=\boldsymbol{Q}\boldsymbol{w}$ and using (12), the following is obtained:

$$\begin{aligned}\boldsymbol{w}_{T,\,opt} &\approx \boldsymbol{w}_T-\tfrac{1}{2}\Lambda^{-2}\boldsymbol{Q}\nabla_w(J_{\mathrm{MSE}})\approx\boldsymbol{w}_T+\Lambda^{-2}\boldsymbol{Q}e(n)\boldsymbol{x}(n) \\ &= \boldsymbol{w}_T+\Lambda^{-2}e(n)\boldsymbol{X}_T(n)\end{aligned} \tag{13}$$

where the gradient vector $\nabla_w(J_{\rm MSE})$ is estimated from the instantaneous MSE error as $\nabla_w(J_{\rm MSE}) \approx -2e(n)\boldsymbol{x}(n)$. Letting $\boldsymbol{w}_T = \boldsymbol{w}_T(n)$ and denoting the stepsize for the gradient vector by μ_T, the following weight-updating equation for the conventional TLMS algorithm is obtained:

$$\boldsymbol{w}_T(n+1) = \boldsymbol{w}_T(n) + \mu_T \Lambda^{-2} e(n) \boldsymbol{X}_T(n). \tag{14}$$

In the proposed TLMM algorithm, the gradient vector is estimated from the instantaneous robust distortion $\rho(e(n))$ as in the LMM algorithm, $\nabla_w(J_{M\rho}) \approx -q(e(n))e(n)\boldsymbol{x}(n)$. Using the first equation in (13), we then arrive at the following TLMM algorithm:

$$\boldsymbol{X}_T(n) = \boldsymbol{Q}\boldsymbol{X}(n) \tag{15}$$

$$e(n) = d(n) - \boldsymbol{w}_T^t(n-1)\boldsymbol{X}_T(n) \tag{16}$$

$$\boldsymbol{w}_T(n) = \boldsymbol{w}_T(n-1) + \tfrac{1}{2}\mu_T \Lambda^{-2} q(e(n))e(n)\boldsymbol{X}_T(n). \tag{17}$$

Like the LMM algorithm, the weighting function $q(e(n))$ in (17) will reduce the effects of impulses in either the input or the desired signals on the filter weight update, resulting in better performance.

IV. Parameter Estimation and Computational Complexity

The choice of the threshold parameters ξ, Δ_1, and Δ_2 for the function $\rho(\cdot)$ is now addressed. Though the distribution of the estimation error $e(n)$ is in general unknown, it is assumed, for simplicity, to be Gaussian distributed but corrupted with additive impulse noise. By estimating the variance of $e(n)$ without impulses, it is possible to detect and reject the impulses in $e(n)$. More specifically, the probability of $e(n)$ greater than a given threshold T_h is given by [6], $\theta_T(n) = P_r\{|e(n)| > T_h\} = erfc(T_h/(\sqrt{2}\,\hat{\sigma}_e(n)))$, where $ercf(x) = (2/\sqrt{\pi})\int_x^\infty e^{-x^2}\,dx$ is the complementary error function and $\hat{\sigma}_e(n)$ is the estimated standard deviation of the "impulse free" error. Using different threshold parameters T_h, one can detect impulses with different degree of confidence. Let $\theta_\xi \triangleq P_r\{|e(n)| > \xi\}$, $\theta_{\Delta_1} \triangleq P_r\{|e(n)| > \Delta_1\}$, and $\theta_{\Delta_2} \triangleq P_r\{|e(n)| > \Delta_2\}$ be the probabilities that $e(n)$ is greater than ξ, Δ_1, and Δ_2, respectively. By appropriate choice of θ_ξ, θ_{Δ_1}, and θ_{Δ_2}, the values of ξ, Δ_1, and Δ_2 can be determined. The remaining problem is the robust estimation of $\hat{\sigma}_e(n)$. In [11], we examined a number of methods for estimating $\hat{\sigma}_e^2(n)$. Simulation shows that the following robust recursive estimator for $\hat{\sigma}_e^2(n)$ is both effective and computational inexpensive

$$\hat{\sigma}_e^2(n) = \lambda_\sigma \hat{\sigma}_e^2(n-1) + C_1(1-\lambda_\sigma)\,\text{med}(A_e(n)) \tag{18}$$

where $C_1 = 1.483(1 + 5/(N_w - 1))$ is a finite sample correction factor [9], $A_e(n) = \{e^2(n), \ldots, e^2(n - N_w + 1)\}$, and λ_σ is the forgetting factor. Interested readers are referred to [7] and [11] for a detailed comparison of the various estimation methods for $\hat{\sigma}_e^2(n)$.

The computational complexity of the proposed LMM and the TLMM algorithms is now briefly discussed. For the RLS algorithm, $O(N^2)$ arithmetic operations per iteration are required [14]. For the RLM algorithm [7], $O(N_w \log_2 N_w)$ more operations per iteration are needed to compute $\hat{\sigma}_e^2(n)$. The LMM algorithm, on the other hand, is an LMS-type algorithm that has $O(N)$ arithmetic complexity. Likewise, $O(N_w \log_2 N_w)$ more operations are needed to compute $\hat{\sigma}_e^2(n)$ per iteration. The $(N \times N)$ orthogonal transformation in the TLMM algorithm will require in general $O(N^2)$ operations. However, fast recursive algorithm for computing the running DCT requires only $O(N)$ operations per iteration [12]. Therefore, the computational complexity of the TLMM algorithm is of the order $O(N)$ together with $O(N_w \log_2 N_w)$ more operations for computing $\hat{\sigma}_e^2(n)$ when the running DCT is used.

V. Simulation Results

The performances of the proposed LMM and the TLMM algorithms are evaluated for the system identification problem shown in Fig. 1 with impulsive interferences. The unknown system is modeled as an FIR filter with impulse response $\boldsymbol{w}^* = [0.2, -0.4, 0.6, -0.8, 1, -0.8, 0.6, -0.4, 0.2]^t$. The input signal $x(n)$ is colored and 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] [14]. The length of the adaptive filter is set to nine ($N = 9$) and the initial values of the weights are set to zeros ($\boldsymbol{w}(0) = \boldsymbol{0}$ and $\boldsymbol{w}_T(0) = \boldsymbol{0}$). The DCT is used as the orthogonal transformation in the TLMS and the TLMM algorithms. The interference $\eta_o(n)$ at the desired signal is modeled as a CG noise, which is given by $\eta_o(n) = \eta_g(n) + \eta_{im}(n) = \eta_g(n) + b(n) * \eta_w(n)$, where $\eta_{im}(n)$ is the impulse noise, $\eta_g(n)$ and $\eta_w(n)$ are modeled as independently identically distributed (i.i.d.) zero-mean Gaussian noise with variance σ_g^2 and σ_w^2, respectively, and $b(n)$ is a switch sequence of ones and zeros, which is modeled as an i.i.d. Bernoulli random process with occurrence probability $P_r(b(n) = 1) = p_r$ and $P_r(b(n) = 0) = 1 - p_r$. The ratio $\gamma_{im} = \sigma_{im}^2/\sigma_g^2 = p_r\sigma_w^2/\sigma_g^2$ determines the impulsive characteristic of $\eta_o(n)$ [4], [15]. For a fixed value of σ_g^2, the larger the value of γ_{im}, the more impulsive $\eta_o(n)$ becomes. In the following simulations, 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, respectively, 95% and 97.5% confidence to down weight the error in the intervals $[\xi, \Delta_1]$ and $[\Delta_1, \Delta_2]$, and 99% confidence to reject it when $e(n) > \Delta_2$. The threshold parameters are obtained according to $\theta_T(n) = erfc(T_h/(\sqrt{2}\,\hat{\sigma}_e(n)))$ as follows: $\xi = k_\xi\hat{\sigma}_e(n) = 1.96\hat{\sigma}_e(n)$, $\Delta_1 = k_{\Delta_1}\hat{\sigma}_e(n) = 2.24\hat{\sigma}_e(n)$, and $\Delta_2 = k_{\Delta_2}\hat{\sigma}_e(n) = 2.576\hat{\sigma}_e(n)$, where $\hat{\sigma}_e(0) = 0$. The signal-to-noise ratio (SNR) at the system output is defined as SNR $= 10\log_{10}(\sigma_{d_0}^2/\sigma_g^2)$, where $\sigma_{d_0}^2$ is the variance of $d_0(n)$.

Example 1: Robustness and Convergence Performance: This experiment is carried out to evaluate the convergence speed and robustness of the proposed LMM and TLMM algorithms under contaminated Gaussian noise. Performances of these algorithms are compared with the ATNA [3], RMN [5], OSLMS [1], and TLMS [12] algorithms. In addition, the unknown system transfer function $\boldsymbol{w}^*$ is suddenly changed to $-\boldsymbol{w}^*$ at time instant $n = 7000$ to evaluate the behavior of the algorithms when the system parameters change suddenly. Step sizes for the various algorithms were chosen such that each algorithm produces the same average excess mean squared error at convergence. The resulting step sizes for all algorithms are illustrated in Fig. 3. For the LMM and TLMM algorithms, the forgetting factor λ_σ and the window length N_w are set to 0.99 and 14, respectively. The window length N_w for the OSLMS algorithm is set to seven. For the TLMS and the TLMM algorithms, λ_p is set to 0.98. For illustration purposes, from $n = 1$ to 1490 and 2801 to 9000, the interference consists of only Gaussian noise; whereas from $n = 1500$ to 2800, the contaminated Gaussian noise is used, which is generated by $\eta_o(n) = \eta_g(n) + b(n) * \eta_w(n)$ with $p_r = 0.005$ and $\gamma_{im} = 300$. To visualize more clearly the effect of impulses in $d(n)$, their locations generated by $b(n)$ are fixed and marked in Figs. 3 and 4 but their amplitudes are varied according to $\eta_w(n)$, which is generated independently in each run. Also, for simplicity in visualizing the effect of impulses in $x(n)$, only one impulse is added to $x(n)$ at $n = 3350$. The SNR is set to 35 dB. The mean squared errors (MSE) are obtained by averaging over 100 independent runs. The MSE results for the TLMM, LMM, TLMS, ATNA, RMN, and OSLMS algorithms are plotted in Figs. 3 and 4. It can be seen that the LMM and the TLMM algorithms are robust to impulses appearing in either the desired or input signals. The LMM algorithm, however, converges slower than that of the TLMS and TLMM algorithms due to

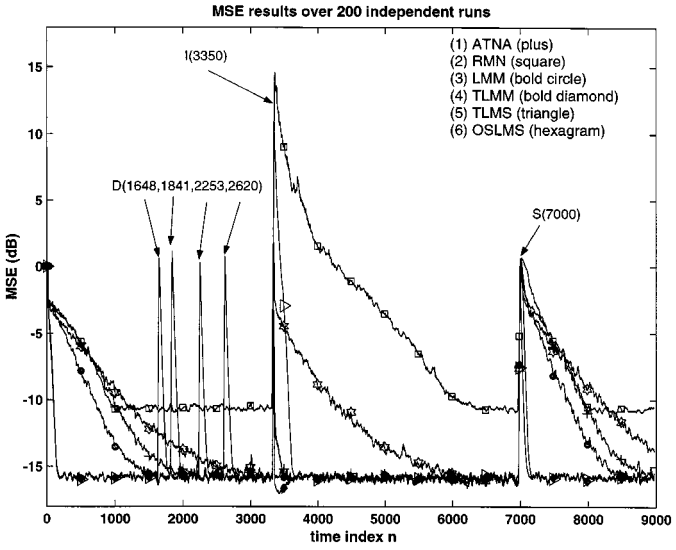


Fig. 3. MSE results versus time (n) (Example 1). $p_r = 0.005$, $N_w = 14$, SNR $= 35$ dB. (1) ATNA (plus); (2) RMN (square); (3) LMM (bold circle); (4) TLMM (bold diagonal); (5) TLMS (triangle) and (6) OSLMS (hexagram) (D (n), and I (n) indicate the locations of the impulses in the desired and the input signals at time instant n, respectively. S (n) indicates the time instant when the system changes suddenly). $\mu_{\mathrm{LMS}} = \mu_{\mathrm{LMM}} = \mu_{\mathrm{OSLMS}} = \mu_{\mathrm{ATNA}} = 0.016$, $\mu_{\mathrm{RMN}} = 0.01$, $\mu_{\mathrm{T,\,TLMS}} = 0.011$, and $\mu_{\mathrm{T,\,TLMM}} = 0.022$.

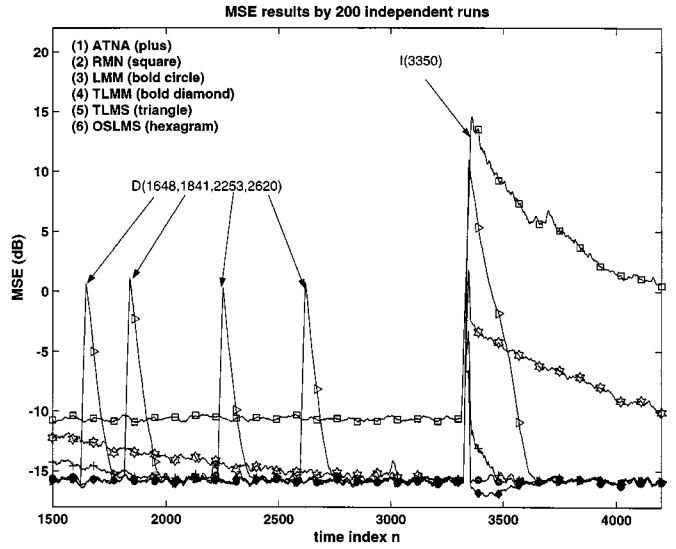


Fig. 4. MSE results ($n = 1500$ to 4200) (Example 1). $p_r = 0.005$, $N_w = 14$, SNR $= 35$ dB. (1) ATNA (plus); (2) RMN (square); (3) LMM (bold circle); (4) TLMM (bold diagonal); (5) TLMS (triangle), and (6) OSLMS (hexagram).

its LMS nature. The TLMM algorithm provides a much faster rate of convergence compared to the RMN, OSLMS, and ATNA algorithms. The TLMM algorithm is also more robust to sudden system parameter changes ($n = 7000$) than other algorithms. It is also found that the performance of the OSLMS algorithm is degraded significantly by the impulse in the input signal at $n = 3350$. Under the experiment conditions, the performance of the RMN algorithm is rather poor.

Example 2: MMSE Versus Probability of Occurrence of Impulses: This example evaluates the performance of various algorithms under different probability of occurrence of impulses. The parameter settings are identical to those in Example 1 except that the locations of impulses are not fixed, there is no impulse in the input signal, and N_w is set to nine for the LMM and TLMM algorithms. The mean MSE (MMSE) performance of various algorithms as a function of the probability of impulses p_r in the desired signal $d(n)$ is plotted in Fig. 5. The following are observed: the MMSE performance of the TLMM, LMM, ATNA, OSLMS, and RMN algorithms are only slightly impaired when the percentage of impulse noise is increased. The performance of these algorithms demonstrates the effectiveness of the robust statistics approach in suppressing the adverse influence of the impulse noise. The TLMM algorithm has the best MMSE performance compared with other algorithms considered under the experimental conditions, and there is about 10 dB improvement over other algorithms.

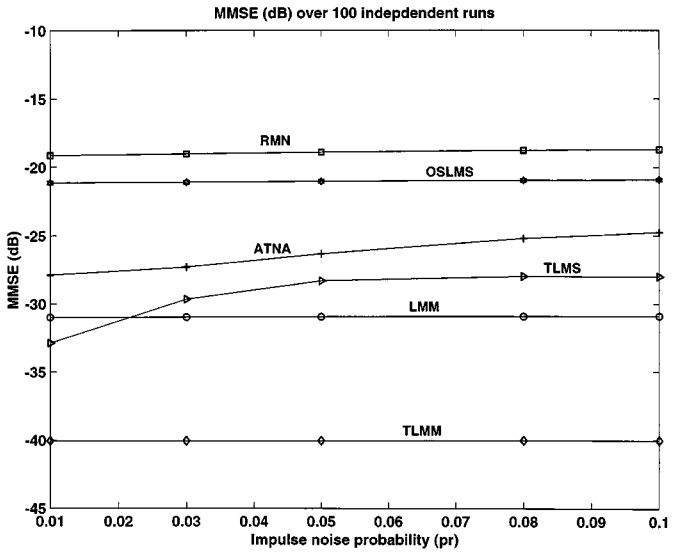


Fig. 5. MMMSE versus probability of the impulse of Example 2. $N_w = 9$ and SNR $= 35$ dB.

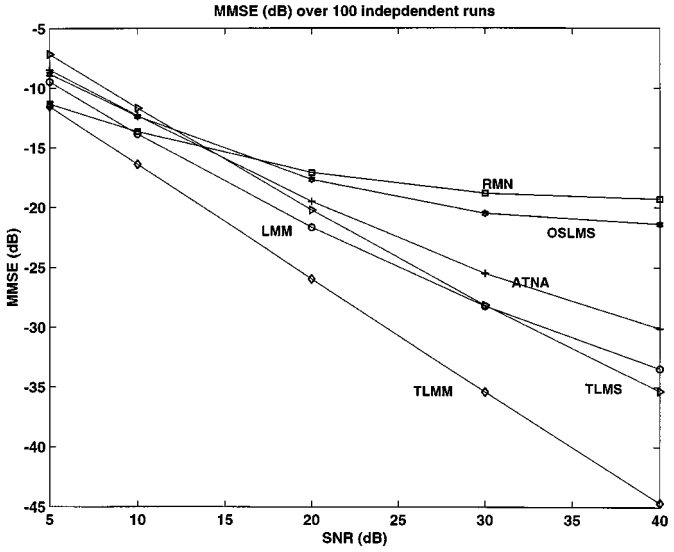


Fig. 6. MMSE versus SNR of Example 3. $N_w = 9$ and $P_r = 0.02$.

Example 3: MMSE Versus Signal-to-Noise Ratio: The experimental conditions are identical to those of Example 2 except that the parameters p_r and γ_{im} are set to 0.02 and 300, respectively. The MMSE performance of various algorithms as a function of the SNR is plotted in Fig. 6. It can be seen that the TLMM algorithm has the lowest MMSE. The improvement in the MMSE for the TLMM and the LMM algorithms decreases when the SNR is reduced, whereas for the OSLMS and RMN algorithms, the MMSE performance improves slightly with the increase of the SNR.

Due to page limitations, simulations regarding the sensitivity of the LMM and the TLMM algorithms to the choice of the threshold parameters k_ξ, k_{Δ_1}, and k_{Δ_2} for ξ, Δ_1 and Δ_2 are omitted in this paper. Interested readers are referred to [11], where extensive simulations were performed with the threshold parameters chosen from the ranges $1.287 \leq k_\xi \leq 2.576$, $1.43956 \leq k_{\Delta_1} \leq 3.091$, and $1.6449 \leq k_{\Delta_2} \leq 3.481$. This corresponds to (80–99)% and (85–99.98)% confidence to down weight the error in the intervals $[\xi, \Delta_1]$ and $[\Delta_1, \Delta_2]$, respectively, and (90 99.98)% confidence to reject it when $|e(n)| > \Delta_2$. It was found that the performance of the LMM and the TLMM algorithms together with the proposed threshold parameter estimation is robust to impulse disturbances within a wide range of threshold values for impulses appearing in the desired signal. Moreover, it is also not sensitive to the choices of k_ξ, k_{Δ_1}, and k_{Δ_2} when the input signal is corrupted by impulses provided that they are not at the tail part of the signal distributions [11].

From the simulation results of the above examples, it can be concluded that under the experimental conditions specified, the TLMM algorithm is more effective and robust than the ATNA, RMN, LMM, and OSLMS algorithms in mitigating the adverse effects due to impulses either in the desired signal or in the input signal. It is an attractive suboptimal algorithm with a much lower computational complexity of $O(N)$ when compared with other RLS-based algorithms.

VI. Conclusion

Two new adaptive filtering algorithms, called the least mean M-estimate and the transform-domain least mean M-estimate, have been proposed for robust adaptive filtering in impulse noise environment. They can be viewed, respectively, as the generalization of the LMS and the transform-domain LMS algorithms using the robust statistics concept. The arithmetic complexity of the algorithms is of order $O(N)$, which is much lower than that of the RLS-based algorithms. Simulation results have shown that the TLMM algorithm, in particular, is more robust and effective in suppressing the effects of impulsive disturbances when compared with the ATNA, OSLMS, and RMN algorithms.

References

[1] T. I. Haweel and P. M. Clarkson, "A class of order statistic LMS algorithms," *IEEE Trans. Signal Processing*, vol. 40, no. 1, pp. 44–53, 1992.

[2] R. Settineri, M. Najim, and D. Ottaviani, "Order statistic fast Kalman filter," in *Proc. IEEE Int. Symp. Circuits and Systems (ISCAS'96)*, vol. 2, 1996, pp. 116–119.

[3] S. Koike, "Adaptive threshold nonlinear algorithm for adaptive filters with robustness against impulsive noise," *IEEE Trans. Signal Processing*, vol. 45, no. 9, pp. 2391–2395, 1997.

[4] J. F. Weng and S. H. Leung, "Adaptive nonlinear RLS algorithm for robust filtering in impulse noise," in *Proc. IEEE Int. Symp. Circuits and Systems (ISCAS'97)*, vol. 4, Hong Kong, 1997, pp. 2337–2340.

[5] J. A. Chambers and A. Avlonitis, "A robust mixed-norm (RMN) adaptive filter algorithm," *IEEE Signal Processing Lett.*, vol. 4, no. 2, pp. 46–48, 1997.

[6] Y. Zou, S. C. Chan, and T. S. Ng, "A robust M-estimate adaptive filter for impulse noise suppression," in *Proc. Int. Conf. Acoustic Speech Signal Processing (ICASSP'99)*, vol. 4, 1999, pp. 1765–1768.

[7] 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 Lett., 2000, submitted for publication.

[8] P. Petrus, "Robust Huber adaptive filter," *IEEE Trans. Signal Processing*, vol. 47, no. 4, pp. 1129–1133, 1999.

[9] P. J. Rousseeuw and A. M. Leroy, *Robust Regression and Outlier Detection*. New York: Wiley, 1987.

[10] R. E. Frank and M. Hampel, *Robust Statistics: The Approach Based on Influence Functions*. New York: Wiley, 1986.

[11] Y. Zou, "Robust statistics based adaptive filtering algorithms for impulsive noise suppression," Ph.D. dissertation, Univ. of Hong Kong, May 2000.

[12] S. S. Narayan and A. M. Peterson, "Transform domain LMS algorithm," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. ASSP-31, no. 3, pp. 609–615, 1983.

[13] D. F. Marshall and W. K. Jenkins, "The use of orthogonal transforms for improving performance of adaptive filters," *IEEE Trans. Circuits Syst.*, vol. 36, no. 4, pp. 474–484, 1989.

[14] S. Haykin, *Adaptive Filter Theory*, 2nd ed. Englewood Cliffs, NJ: Prentice-Hall, 1991.

[15] S. R. Kim and A. Efron, "Adaptive robust impulsive noise filtering," *IEEE Trans. Signal Processing*, vol. 43, no. 8, pp. 1855–1866, 1995.

Dynamic Biasing for True Low-Voltage CMOS Class AB Current-Mode Circuits

G. Palmisano and S. Pennisi

***Abstract*—This paper presents a dynamic biasing approach for continuous-time CMOS class AB current-mode circuits. The method allows low-voltage circuits to be implemented whose supply requirements are restricted to one *threshold* voltage plus two *saturation* voltages. Fundamental limitations of the approach are analyzed and found to be compatible with a wide spectrum of analog applications, some of which are briefly discussed. A complementary current mirror, a current comparator, and a current amplifier were designed using the proposed technique. SPICE simulations using a 0.8-μm process are provided, which confirm the overall performance of these circuits especially in terms of low-voltage capability and speed without compromising linearity.**

***Index Terms*—Biasing, class AB, CMOS, current mirrors, current mode, low voltage, switched capacitor.**

I. Introduction

In recent years, current-mode (CM) signal processing has been widely investigated, and several works have demonstrated that this approach can solve many circuit and system problems. As can be expected from circuits exploiting CM techniques, performance in terms of low-voltage capability, slew rate, and bandwidth can in principle be maximized [1]. However, when class AB topologies have to be implemented, the need for a complementary structure has prevented, until now, the achievement of true low-voltage features. Compared to class A topologies, class AB versions provide better dynamic range [2] and reduced sensitivity to process tolerances [3]. In addition, they exhibit extremely high slew-rate values.

Although the well-known switched-current approach has been used in the past to achieve both class AB and low-voltage operations [4], no effective continuous-time approach providing the same performance exists at present.

For instance, a popular class AB CM input stage capable of managing a bipolar current is shown in Fig. 1. This has been used in a wide range of applications, such as in the input stages of current amplifiers [5], current conveyors [6], [7], current comparators [8]–[10],

Manuscript received December 1999. This paper was recommended by Associate Editor M. Ismail.

The authors are with the Dipartimento Elettrico Elettronico e Sistemistico, Universita' di Catania, Catania I-95125 Italy (e-mail: gpalmisano@dees.unict.it; spennisi@dees.unict.it).

Publisher Item Identifier S 1057-7130(00)11653-2.