IEEE Access
Multidisciplinary ⋮ Rapid Review ⋮ Open Access Journal

Received December 12, 2018, accepted January 2, 2019, date of publication January 9, 2019, date of current version January 29, 2019.

Digital Object Identifier 10.1109/ACCESS.2019.2891549

Robust Normalized Least Mean Absolute Third Algorithms

KUI XIONG[1,2], SHIYUAN WANG[1,2], (Senior Member, IEEE), AND BADONG CHEN[3], (Senior Member, IEEE)

[1]College of Electronic and Information Engineering, Southwest University, Chongqing 400715, China
[2]Chongqing Key Laboratory of Nonlinear Circuits and Intelligent Information Processing, Chongqing 400715, China
[3]Institute of Artificial Intelligence and Robotics, Xi'an Jiaotong University, Xi'an 710049, China

Corresponding author: Shiyuan Wang (wsy@swu.edu.cn)

This work was supported in part by the National Natural Science Foundation of China under Grant 61671389, and in part by the Fundamental Research Funds for the Central Universities under Grant XDJK2017D177.

ABSTRACT This paper addresses the stability issues of the least mean absolute third (LMAT) algorithm using the normalization based on the third order in the estimation error. A novel *robust normalized least mean absolute third* (RNLMAT) algorithm is therefore proposed to be stable for all statistics of the input, noise, and initial weights. For further improving the filtering performance of RNLMAT in different noises and initial conditions, the variable step-size RNLMAT (VSSRNLMAT) and the switching RNLMAT (SWRNLMAT) algorithms are proposed using the statistics of the estimation error and a switching method, respectively. The filtering performance of RNLMAT is improved by VSSRNLMAT and SWRNLMAT at the expense of affordable computational cost. RNLMAT with less computational complexity than other normalized adaptive filtering algorithms, can provide better filtering accuracy and robustness against impulsive noises. The steady-state performance of RNLMAT and SWRNLMAT in terms of the *excess mean-square error* is performed for theoretical analysis. Simulations conducted in system identification under different noise environments confirm the theoretical results and the superiorities of the proposed algorithms from the aspects of filtering accuracy and robustness against large outliers.

INDEX TERMS Least mean absolute third algorithm, normalization, robustness, variable step-size, switching, performance analysis, impulsive noise.

I. INTRODUCTION

In adaptive filters, the least mean square (LMS) algorithm using the minimum mean squared error criterion is the optimum under the Gaussian assumption [1]. However, in non-Gaussian situations, LMS may suffer from performance degradation and instability issues. To address these issues, the adaptive filters using higher order moments of error have therefore been proposed. For example, the least mean fourth (LMF) algorithm using the mean fourth order of error outperforms LMS in sub-Gaussian (light-tailed) noises [1], [2]. The least mean $2L$ (LM2L) algorithm [3] minimizes the $2L$th order of error, where L is an integer greater than 1. Therefore, LMF can be regarded as a special case of LM2L with $L = 2$. The least mean absolute third (LMAT) algorithm using the mean absolute third power of error [4]–[6] is superior to LMS for most of noise probability densities [6]. Even for Gaussian noise, LMAT can provide faster convergence rate than LMS and LMF [6]. However, the convergence conditions of LMF and LMAT strongly depend on input variance, noise variance, and weight initialization [2], [7]–[11]. To address these convergence dependencies, the normalization regarding the weight update form is generally used to improve the stability of the aforementioned filters.

For LMF, different normalized terms in the weight update form generate different versions of normalized least mean fourth (NLMF) algorithms [12]–[17]. In LMF, the numerator of the weight update term is fourth order in the input. For improving the stability, the normalized terms are given by a squared norm of the input [12] and a weighted sum of the squared norm of the input and the error [13], [16], which are called NLMF1 and NLMF2 in this paper respectively. Since in NLMF1 and NLMF2, the denominator is second order in the input, NLMF1 and NLMF2 cannot converge when the input variance exceeds a threshold value which is related to their step-sizes. To overcome this issue, the normalized term is modified by the fourth power of the norm of the input [17], which is called NLMF3. However, the stability of NLMF3 still depends on both the noise variance and the weight initialization [15]. A global stability of NLMF (namely NLMF4) is therefore proposed using fourth order in the input and second order in the error as the normalized term [18]. The mean square deviation (MSD) of NLMF4 has been shown to be bounded for all finite values of the input variance, noise variance, initial MSD, and mean-square plant parameter increment [19], [20]. To further address these limitations of NLMF4 efficiently, a switchable normalization is constructed by the product of the squared norm of the input and the maximum of the squared error and the scaled squared norm of the input, leading to a stable NLMF5 [21].

A nonparametric variable step-size least mean absolute third (NVSLMAT) algorithm is proposed to improve the capability of LMAT against non-Gaussian noises [22]. Further, to combat Gaussian and non-Gaussian noises in the time-varying unknown system under low signal-to-noise ratio, an optimized least mean absolute third (OPLMAT) algorithm is proposed in [23]. However, there exist the stability issues in NVSLMAT and OPLMAT. The convergence performance of the aforementioned LMAT and its extensions still strongly depend on the input variance and the initialization of filter weights, which is similar to LMF [6]. The normalization strategy is also used to improve the stability of LMAT. Unlike the aforementioned normalization methods used in LMF, the upper bound of the squared error is combined with the squared norm of the input as the normalization, leading to a normalized least mean absolute third (NLMAT) algorithm [24]. NLMAT can provide the robustness against impulsive noise thanks to its limitation of the squared error. However, the stability issues based on input variance, noise variance, and weight initialization cannot be solved absolutely since the normalization term in NLMAT is that the numerator is second order in the estimation error while the denominator is second order in the input. In addition, other robust adaptive filters based on different error criteria, e.g., the correntropy-based algorithms [25]–[28], logarithm-based algorithms [29], [30], mixed norm algorithms [31]–[34], and others [35], [36], are proposed for impulsive noises. However, all these algorithms cannot provides performance improvement in both Gaussian and non-Gaussian noises, simultaneously.

In this paper, to address the stability issues of LMAT, a novel *robust normalized least mean absolute third* (RNLMAT) algorithm is proposed by using the third order in the estimation error as the normalization. The purpose of the normalization in RNLMAT is to generate stability with the increase of both the input variance and the noise variance. Therefore, RNLMAT can provide the stability features regarding the input, the noise, and the initialization of filter weights. RNLMAT is developed by the variable step-size RNLMAT (VSSRNLMAT) and the switching RNLMAT (SWRNLMAT) algorithms for improvement of filtering performance. Four contributions of this paper are summarized as follows. (1) The stability of the proposed RNLMAT is not influenced by the input variance, the noise variance, and initialization of filter weights. And RNLMAT provides filtering performance improvement and robustness against impulsive noises including large outliers, simultaneously. (2) To dramatically improve the filtering performance of RNLMAT in impulsive noise environments, VSSRNLMAT is proposed using the statistics of the estimation error. (3) Using the switching method for LMAT, SWRNLMAT is proposed for the case of large weight initialization in the absence of impulsive noises. And SWRNLMAT can provide excellent convergence performance when the initial weight vector is very large. (4) The steady-state performance of RNLMAT and SWRNLMAT in terms of the excess mean-square error (EMSE) is performed under different noise distributions.

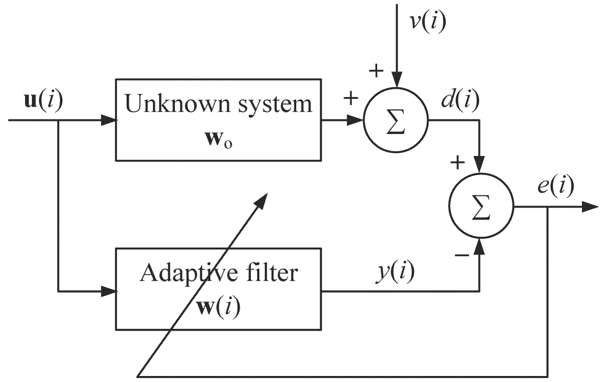


FIGURE 1. System identification structure.

The rest of the paper is organized as follows. Section II describes the formulation of the system identification and its algorithms. A novel robust NLMAT is presented in Section III. The extensions of RNLMAT, i.e., VSSRNLMAT and SWRNLMAT, are also proposed in this section. In Section IV, the steady-state convergence performance of RNLMAT and SWRNLMAT are obtained for theoretical analysis. Monte Carlo simulations are conducted in the non-impulsive and impulsive noises to validate the theoretical results and the MSD performance of the proposed algorithms in Section V. Section VI concludes this paper.

II. PROBLEM FORMULATION

The system identification model based on the finite impulse response (FIR) is shown in Fig. 1. The desired response $d(i)$ in Fig. 1 is given by

$$d(i) = \mathbf{w}_o^{\mathrm{T}}\mathbf{u}(i) + v(i), \tag{1}$$

where $\mathbf{w}_o = [w_{o,1}, w_{o,2}, \ldots, w_{o,N}]^{\mathrm{T}}$ is the unknown optimal weight vector of the FIR system with length N; $\mathbf{u}(i) = [u_i, u_{i-1}, \ldots, u_{i-N+1}]^{\mathrm{T}}$ is the input vector at discrete time i with $(\cdot)^{\mathrm{T}}$ being the transpose operator; $v(i)$ is the system background noise. The system identification is also modeled by an FIR filter with the same structure as that of the unknown system $\mathbf{w}_o$, and the estimation error is therefore defined by

$$e(i) = d(i) - \mathbf{w}^{\mathrm{T}}(i)\mathbf{u}(i), \tag{2}$$

where $\mathbf{w}(i) = [w_{i,1}, w_{i,2}, \ldots, w_{i,N}]^{\mathrm{T}}$ is the weight vector of the adaptive filter.

The LMAT algorithm minimizes the mean absolute third error, and its update form regarding weight vector $\mathbf{w}(i)$ is therefore shown by [4]

$$\mathbf{w}(i+1) = \mathbf{w}(i) + \mu e^2(i)\mathrm{sign}[e(i)]\mathbf{u}(i), \tag{3}$$

where $\mu > 0$ is the step-size; sign$[e(i)]$ denotes the signum function of $e(i)$, i.e., if $e(i) > 0$, then sign$[e(i)] = 1$, and if $e(i) = 0$, then sign$[e(i)] = 0$, otherwise sign$[e(i)] = -1$.

The convergence performance of LMAT depends on the power of the input [4], [6]. To overcome this dependency, the normalized LMAT (NLMAT) using the squared norm of the input is introduced in [24], and its weight update form is described by

$$\mathbf{w}(i+1) = \mathbf{w}(i) + \mu \frac{e^2(i)\mathrm{sign}[e(i)]}{\delta + \|\mathbf{u}(i)\|^2}\mathbf{u}(i), \tag{4}$$

where $\|\mathbf{u}(i)\|$ is the Euclidean norm of $\mathbf{u}(i)$, i.e., $\|\mathbf{u}(i)\|^2 = \mathbf{u}^{\mathrm{T}}(i)\mathbf{u}(i)$, and δ is a small positive constant.

To improve the stability of NLMAT, (4) is modified by [24]

$$\mathbf{w}(i+1) = \mathbf{w}(i) + \mu \frac{\mathrm{sign}[e(i)]\mathbf{u}(i)}{\delta + \|\mathbf{u}(i)\|^2} \min\left\{e^2(i), e_{up}\right\}, \tag{5}$$

where e_{up} is the upper-bound of $e^2(i)$ and given by

$$e_{up} < \frac{\sqrt{2\pi}\sigma_e(i)}{\mu}, \tag{6}$$

where $\sigma_e(i)$ is the standard deviation of $e(i)$, i.e., $\sigma_e(i) = \sqrt{E[e^2(i)]}$ with E being the mathematical expectation and $\sigma_e(i)$ being estimated by [24] and [31]

$$\sigma_{\mathrm{e}}(i) = \sqrt{\frac{1}{N_w - K}\mathbf{O}^{\mathrm{T}}(i)\mathbf{T}_w\mathbf{O}(i)}, \tag{7}$$

where $\mathbf{O}(i) = O([|e(i)|, |e(i-1)|, \ldots, |e(i-N_w+1)|]^{\mathrm{T}})$ contains the N_w most recent absolute values of the error ($N_w = N$ is set in [24]) with $O(\cdot)$ sorting the elements in $(\cdot)$ from the smallest to the largest value; $\mathbf{T}_w = \mathrm{diag}[1, \ldots, 1, 0, \ldots, 0]$ denotes a diagonal matrix whose last K elements are set to zero. The largest K elements in $\mathbf{O}(i)$ are set to zero by $\mathbf{T}_w$ and the remainder is used to achieve an unbiased estimate of $\sigma_e(i)$. In addition, in the absence of impulsive noise, $\sigma_e^2(i)$ is estimated by [24]

$$\sigma_e^2(i) = \beta\sigma_e^2(i-1) + (1-\beta)e^2(i), \tag{8}$$

where $\beta = 1 - 1/\kappa N$ with $\kappa > 2$ being an exponential weighting factor.

In [24], to guarantee the mean convergence of NLMAT, the range of its step-size is given by

$$0 < \mu < \sqrt{\frac{\pi}{2}}\frac{1}{\sigma_e(i)}. \tag{9}$$

From (5) and (6), we see that the normalization term of $\|\mathbf{u}(i)\|^2$ in NLMAT is used to solve the stability issue based on the input power, and the restricted error term of $\min\{e^2(i), e_{up}\}$ is used to alleviate the stability issue based on the noise power. However, it can be seen from (9) that the upper bound of μ is influenced by the standard deviation $\sigma_e(i)$. And $\sigma_e(i)$ is related to the variances of the input and the noise. At the beginning of filtering process, $\sigma_e(i)$ is large generally. The range of the step-size of NLMAT for convergence is therefore very small. Therefore, given a step-size, the convergence of NLMAT strongly depends on the initialization of weights and the variances of the input and the noise. This dependence is caused by the normalization based on a second order polynomial of the input, which is shown in (4). Although there exists the limitation of the upper bound of squared error in NLMAT (5), it cannot guarantee filtering accuracy in the presence of large outliers. To this end, a novel *robust normalized least mean absolute third* (RNLMAT) algorithm and its extensions are proposed in the following.

III. PROPOSED ALGORITHMS

A. RNLMAT ALGORITHM

First, we consider the following adaptive filter with error nonlinearity [37]

$$\mathbf{w}(i+1) = \mathbf{w}(i) + \mu f(e(i))\mathbf{u}(i), \tag{10}$$

with $f(e(i))$ being the error nonlinear function.

The optimal error nonlinear function is therefore obtained by minimizing the steady-state mean square error, i.e., [37]

$$f^{opt}(e(i)) = \frac{p_e'(e(i))}{p_e(e(i))}, \tag{11}$$

where $p_e(e(i))$ denotes the probability density function (PDF) of estimation error $e(i)$, and $p_e'(e(i))$ is the first derivative of $p_e(e(i))$. Generally, owing to the unavailable $p_e(e(i))$, it is difficult to calculate the optimal nonlinearity (11).

Inspired by the optimal nonlinearity (11), for simplicity, we replace $p_e(e(i))$ in (11) with an even function $h(e(i))$, to obtain the following error nonlinear function:

$$f(e(i)) = \frac{h'(e(i))}{h(e(i))}, \tag{12}$$

where $h'(e(i))$ is the first derivative of $h(e(i))$. Then, according to the LMAT algorithm (3), by setting $h'(e(i)) = e^2(i)\mathrm{sign}[e(i)]$, we can obtain $h(e(i)) = \alpha + \frac{1}{3}|e(i)|^3$ with $\alpha > 0$.

Substituting $h(e(i))$ and $h'(e(i))$ into (12) and (10), generates a novel robust normalized least mean absolute third (RNLMAT) algorithm as follows:

$$\mathbf{w}(i+1) = \mathbf{w}(i) + \mu \frac{e^2(i)\mathrm{sign}[e(i)]}{1 + \beta|e(i)|^3}\mathbf{u}(i), \tag{13}$$

where $\beta > 0$ together with step-size μ can balance the transient and the steady-state performance. Since $e^2(i)$ is a function of $\mathbf{u}(i)$ from (2), the numerator $e^2(i)\mathbf{u}(i)$ in (13) is third order in $\mathbf{u}(i)$. Therefore, third order in $e(i)$ is used as the normalized term in (13), which is to combat the input variance increase and the unboundedness of the input distribution.

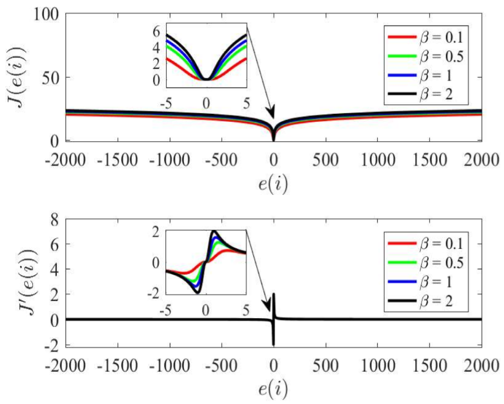


FIGURE 2. The cost function and its derivative of RNLMAT with different β.

In addition, (13) can also be regarded as a variable step-size LMS algorithm with step-size being $\mu \frac{|e(i)|}{1+\beta|e(i)|^3}$, which is close to zero whenever $e(i)$ contaminated by an impulsive interference appears. Thus, the normalized term of (13) can also combat the increases of the weight initialization and the noise variance, efficiently.

In order to further illustrate the robustness of RNLMAT, we need to consider its cost function, which can reflect the essence of algorithm generally. Therefore, the cost function of RNLMAT can be obtained from (13) as:

$$J(e(i)) = \ln\left(1 + \beta|e(i)|^3\right), \tag{14}$$

where factor 3β is assumed to be absorbed in step-size μ. The cost function $J(e(i))$ and its first derivative $J'(e(i))$ versus $e(i)$ are plotted in Fig. 2. As can be seen from Fig. 2, a sufficiently small or large error results in a fairly smooth cost function or a derivative that tends to zero. This illustrates that the proposed RNLMAT algorithm based on the cost function (14) can provide the robustness against large errors and the smoothness to small errors. We also see from Fig. 2 that the curves of $J(e(i))$ and $J'(e(i))$ are steep when the error is of the medium size. This illustrates that RNLMAT can provide fast convergence rate in the transient stage. In addition, we see from Fig. 2 that the value of β affects the steepness of the cost function, which can provide a trade-off between the convergence rate and steady-state accuracy of RNLMAT at different stages of filtering process.

It can be seen from (13) that, RNLMAT generally provides slower convergence rate than LMAT in the absence of impulsive noise thanks to $\frac{1}{1+\beta|e(i)|^3} \leq 1$ in (13). For the case of impulsive noises, $\frac{e^2(i)}{1+\beta|e(i)|^3} \ll 1$ in (13) also reduces the convergence rate of RNLMAT. Thus, we next propose the variable step-size version of RNLMAT to improve its convergence rate.

B. VSSRNLMAT ALGORITHM

Generally, the variable step-size (VSS) strategies are applied to improve the convergence rate of adaptive filters [38], [39]. In the VSS methods, a reasonably large step-size is needed to improve the convergence rate in the transient filtering phase, and a small value to achieve high filtering accuracy in the steady-state phase. Since $|e(i)| = e(i)\text{sign}[e(i)]$, (13) can be rewritten as:

$$\mathbf{w}(i+1) = \mathbf{w}(i) + \mu \frac{e(i)\,|e(i)|}{1+\beta|e(i)|^3}\mathbf{u}(i). \tag{15}$$

To improve robustness and filtering performance simultaneously in the presence of impulsive noises, we apply the following method to estimate $\bar{e}(i) = \sqrt{E[|e(i)|^2]}$, instead of $|e(i)|$ in (15) as follows:

$$\bar{e}(i) = \gamma\bar{e}(i-1) + (1-\gamma)\hat{\sigma}_{e_{\min}}(i), \tag{16}$$

where $0 \ll \gamma < 1$ is the smoothing factor; $\hat{\sigma}_{e_{\min}}(i)$ is estimated by [40]

$$\hat{\sigma}_{e_{\min}}(i) = \sqrt{\frac{\mathbf{O}^{\mathrm{T}}(i)\mathbf{T}_w\mathbf{O}(i)}{K}}, \tag{17}$$

where vector $\mathbf{O}(i)$ is the same as $\mathbf{O}(i)$ in (7); $\mathbf{T}_w = \text{diag}[1, \ldots, 1, 0, \ldots, 0]$ is a diagonal matrix whose last $N_w - K$ elements are set to zero; the largest $N_w - K$ elements in $\mathbf{O}(i)$ are nullified by $\mathbf{T}_w$ and the remaining K smallest elements in $\mathbf{O}(i)$ are used to obtain $\hat{\sigma}_{e_{\min}}(i)$. Note that $\hat{\sigma}_{e_{\min}}(i)$ (17) is different from $\sigma_e(i)$ (7) since only K (usually small) smallest elements of $\mathbf{O}(i)$ are utilized to obtain $\hat{\sigma}_{e_{\min}}(i)$ while the more elements of $\mathbf{O}(i)$ are utilized to estimate $\sigma_e(i)$, generally. In addition, when $K = 1$, i.e., only the smallest element of $\mathbf{O}(i)$ is used to estimate $\bar{e}(i)$, (17) is equivalent to the method used in [36]. Essentially, unlike the traditional VSS methods, the proposed VSS scheme in this paper uses the statistics of the error, which can significantly improve the convergence rate and steady-state filtering performance of RNLMAT in impulsive noises. In addition, (16) can be seen as the improvement of the estimate method proposed in [40] since smoothing factor γ in (16) emphasizes the influence of the recent values of $\hat{\sigma}_{e_{\min}}(i)$ and forgets the past ones, thus improving the traceability of data statistical variations.

Finally, we summarize the proposed variable step-size RNLMAT (VSSRNLMAT) as follows:

$$\mathbf{w}(i+1) = \mathbf{w}(i) + \mu(i)\frac{e(i)}{1+\beta|e(i)|^3}\mathbf{u}(i), \tag{18}$$

where $\mu(i) = \mu\bar{e}(i)$ with $\mu > 0$ being the scale parameter and $\bar{e}(i)$ being obtained from (16) and (17).

In the VSSRNLMAT algorithm, variable step-size $\mu(i)$ is used to improve both the convergence speed and steady-state performance, and the error nonlinearity $\frac{e(i)}{1+\beta|e(i)|^3}$ to guarantee its convergence. In the transient phase, $\mu(i)$ is generally large, which can speed up the convergence rate. And in the steady-state phase, $\mu(i)$ is very small, which can generate the low steady-state error. Note that only the first order statistic of error is used in $\mu(i)$, and thus the numerator in (18) is the second order of the error. Therefore, VSSRNLMAT can guarantee the convergence in adaption.

TABLE 1. Computational complexity of RNLMAT, SWRNLMAT, and VSSRNLMAT per iteration.

Algorithm	Operation	Multiplication	Division	Addition	Comparison
RNLMAT	$e(i)=d(i)-\mathbf{w}^{\mathrm{T}}(i)\mathbf{u}(i)$	N	0	N	0
	$\mathbf{w}(i+1)=\mathbf{w}(i)+\mu\frac{e^2(i)\text{sign}[e(i)]}{1+\beta\lvert e(i)\rvert^3}\mathbf{u}(i)$	$N+3$	1	$N+1$	1
	Total	$2N+3$	1	$2N+1$	1
SWRNLMAT	$e(i)=d(i)-\mathbf{w}^{\mathrm{T}}(i)\mathbf{u}(i)$	N	0	N	0
	$\lvert\lvert\mathbf{u}(i)\rvert\rvert$	N	0	$N-1$	0
	$\mathbf{w}(i+1)=\mathbf{w}(i)+\frac{e^2(i)\text{sign}[e(i)]}{\lVert\mathbf{u}(i)\rVert^2\max\{\zeta\lVert\mathbf{u}(i)\rVert,\lvert e(i)\rvert\}}\mathbf{u}(i)$	$N+3$	1	N	2
	Total	$3N+3$	1	$3N-1$	2
VSSRNLMAT	$e(i)=d(i)-\mathbf{w}^{\mathrm{T}}(i)\mathbf{u}(i)$	N	0	N	0
	$\mathbf{O}(i)$	0	0	0	$N_w\ln N_w$
	$\hat{\sigma}_{e_{\min}}(i)=\sqrt{\frac{\mathbf{O}^{\mathrm{T}}(i)\mathbf{T}_w\mathbf{O}(i)}{K}}$	N_w	1	N_w	0
	$\bar{e}(i)=\gamma\bar{e}(i-1)+(1-\gamma)\hat{\sigma}_{e_{\min}}(i)$	2	0	2	0
	$\mu(i)=\mu\bar{e}(i)$	1	0	0	0
	$\mathbf{w}(i+1)=\mathbf{w}(i)+\mu(i)\frac{e(i)}{1+\beta\lvert e(i)\rvert^3}\mathbf{u}(i)$	$N+2$	1	$N+1$	0
	Total	N_w+2N+5	2	N_w+2N+3	$N_w\ln N_w$

TABLE 2. Comparison of computational complexity between the proposed algorithms and the representative algorithms per iteration.

Algorithm	Multiplication	Division	Addition	Comparison
NLMS	$3N+1$	1	$3N-1$	0
NLMAT with (7)	N_w+3N+3	3	N_w+3N+1	$N_w\ln N_w+2$
NLMAT with (8)	$3N+5$	2	$3N+2$	2
NLMF4	$3N+2$	1	$3N+1$	0
NLMF5	$3N+1$	1	$3N-1$	1
RNLMAT	$2N+3$	1	$2N+1$	1
SWRNLMAT	$3N+3$	1	$3N-1$	2
VSSRNLMAT	N_w+2N+5	2	N_w+2N+3	$N_w\ln N_w$

C. SWRNLMAT ALGORITHM

In the absence of impulsive noise, and when the initial weight vector is assumed to be very large, both RNLMAT and VSSRNLMAT cannot provide fast convergence rate. The switching methods [3], [21] can solve this issue effectively, which is achieved by adaptively switching the original algorithm from one algorithm to the other in different filtering phases.

Inspired by NLMF5 [21], we propose a novel switching adaptive filtering algorithm, namely, the switching RNLMAT (SWRNLMAT) algorithm based on the LMAT algorithm (3) as follows:

$$\mathbf{w}(i+1)=\mathbf{w}(i)+\frac{e^2(i)\text{sign}[e(i)]}{\lVert\mathbf{u}(i)\rVert^2\max\{\zeta\lVert\mathbf{u}(i)\rVert,\lvert e(i)\rvert\}}\mathbf{u}(i), \quad (19)$$

where $\max\{\cdot\}$ denotes the maximum function and $\zeta>0$ is the scaling parameter. Note that there is no extra step-size parameter in (19), and the scaling parameter can achieve a trade-off between the transient and the stead-state performance.

Since SWRNLMAT behaves as a normalized LMS (NLMS) algorithm [41] with a unit step-size when the error is very large, it can be stable for large values of input variance, noise variance, and initial weights in non-impulsive noises [1], [3], [18], [21]. Since SWRNLMAT behaves as a LMAT algorithm with a step-size being $\frac{1}{\zeta\lVert\mathbf{u}(i)\rVert^3}$ when the error is small, it can provide performance improvement in most noise environments [6]. Especially, in the case of large initial weights, SWRNLMAT achieves performance improvement of RNLMAT and VSSRNLMAT in Gaussian and sub-Gaussian noises. However, SWRNLMAT cannot provide robustness against impulsive noises like NLMF5 [21].

D. COMPLEXITY ANALYSIS

In this section, we discuss the computational complexities of the proposed RNLMAT, SWRNLMAT, and VSSRNLMAT algorithms. The detailed computational complexity of RNLMAT, SWRNLMAT, and VSSRNLMAT is shown in Table 1. Here, for VSSRNLMAT, the standard sort algorithms are used to obtain $\mathbf{O}(i)$, and thus $N_w\ln N_w$ comparisons are required. In addition, using the dot product, N_w multiplications and N_w additions are required for the calculation of $\mathbf{O}^{\mathrm{T}}(i)\mathbf{T}_w\mathbf{O}(i)$. Further, Table 2 shows the comparison of the computational complexity of RNLMAT, SWRNLMAT, and VSSRNLMAT with NLMS [41], NLMAT [24], NLMF4 [18], and NLMF5 [21] in terms of the total number of multiplications, divisions, additions, and comparisons at each iteration. We see from Table 2 that RNLMAT has less computational complexity than other normalized algorithms, and the computational complexity of SWRNLMAT is comparable to those of NLMS and NLMF. In addition, the computational complexity of VSSRNLMAT depends on N_w. Thus, we choose N_w according to practical applications.

IV. PERFORMANCE ANALYSIS

In this section, we perform the steady-state performance analysis of the proposed algorithms. Since the probability density function (PDF) of impulsive noise is generally unknown, we only perform the theoretical analysis of RNLMAT and SWRNLMAT algorithms in the absence of impulsive noise.

Generally, the mean square deviation (MSD) [1] defined below is used to evaluate the filtering accuracy of adaptive filters

$$y(i) = E[\|\widetilde{\mathbf{w}}(i)\|^2], \tag{20}$$

where $\widetilde{\mathbf{w}}(i) = \mathbf{w}(i) - \mathbf{w}_o$ is the weight deviation vector. The excess mean-square error (EMSE) is therefore defined as:

$$\xi(i) = E[e_a^2(i)], \tag{21}$$

where $e_a(i) = \widetilde{\mathbf{w}}^{\mathrm{T}}(i)\mathbf{u}(i)$ is the *a priori* estimation error. The EMSE can also be used to evaluate the steady-state performance of filters, effectively [1]. Combining Eqs. (1) and (2), we have

$$e(i) = \widetilde{\mathbf{w}}^{\mathrm{T}}(i)\mathbf{u}(i) + v(i) = e_a(i) + v(i). \tag{22}$$

Then, the mean-square error (MSE) can be obtained by

$$E[e^2(i)] = \sigma_v^2 + \sigma_u^2 E[\|\widetilde{\mathbf{w}}(i)\|^2]. \tag{23}$$

It can be seen from (23) that the convergence of MSD or EMSE is equivalent to the convergence of MSE. Thus, we analyze the steady-state performance of RNLMAT and SWRNLMAT in terms of EMSE. To make the analysis mathematically tractable, the following assumptions are used.

A1: The input $\{\mathbf{u}(i)\}$ is a stationary zero-mean independently and identically distributed (i.i.d.) Gaussian sequence with a finite variance σ_u^2.

A2: The noise $\{v(i)\}$ is a stationary zero-mean i.i.d. sequence with a finite variance σ_v^2 and zero odd order moments.

A3: The input and the noise are mutually independent.

A4: The weight vector $\mathbf{w}(i)$ is independent of the input.

A5: The *a priori* estimation error $e_a(i)$ is zero-mean and independent of the noise.

A6: $\|\mathbf{u}(i)\|^2$ is asymptotically uncorrelated with $f^2(e(i))$, where $f(e(i))$ is a nonlinear function regarding error.

Assumptions $A1-A5$ are commonly used in the theoretical analysis of adaptive filters to enable mathematical tractability for theoretical analysis [1]. The validity of Assumption $A6$ can be guaranteed when the filter is long enough such that $\|\mathbf{u}(i)\|^2 \approx N\sigma_u^2$ [19], [21], [24], and can be found in [37] for more details.

A. PERFORMANCE OF RNLMAT

Subtracting $\mathbf{w}_o$ from both sides of (13) and combining the weight deviation vector $\widetilde{\mathbf{w}}(i)$ generate

$$\widetilde{\mathbf{w}}(i+1) = \widetilde{\mathbf{w}}(i) - \mu f_1(e(i))\mathbf{u}(i), \tag{24}$$

with $f_1(e(i)) = \frac{e^2(i)\mathrm{sign}[e(i)]}{1+\beta|e(i)|^3}$. Premultiplying both sides of (24) by their transposes, using (22), and taking the expected value, we obtain

$$E\left[\|\widetilde{\mathbf{w}}(i+1)\|^2\right] = E\left[\|\widetilde{\mathbf{w}}(i)\|^2\right] - 2\mu E\left[e_a(i)f_1(e(i))\right] + \mu^2 E\left[\|\mathbf{u}(i)\|^2 f_1^2(e(i))\right], \tag{25}$$

When RNLMAT approaches the steady-state, we have $E[\|\widetilde{\mathbf{w}}(i+1)\|^2] = E[\|\widetilde{\mathbf{w}}(i)\|^2]$ in (25). This implies that

$$2E\left[e_a(i)f_1(e(i))\right] = \mu E\left[\|\mathbf{u}(i)\|^2 f_1^2(e(i))\right]. \tag{26}$$

According to Assumptions $A1$ and $A6$, (26) becomes

$$2E\left[e_a(i)f_1(e(i))\right] = \mu \mathrm{Tr}(\mathbf{R_u}) E\left[f_1^2(e(i))\right], \tag{27}$$

where $\mathbf{R_u} = E[\mathbf{u}(i)\mathbf{u}^{\mathrm{T}}(i)]$ is the covariance matrix of the input and $\mathrm{Tr}(\mathbf{R_u})$ denotes the trace of $\mathbf{R_u}$. Since the distributions of $e_a(i)$ and $e(i)$ are independent of i in the steady-state, we omit the time index i and rewrite (27) as

$$2E\left[e_a f_1(e)\right] = \mu \mathrm{Tr}(\mathbf{R_u}) E\left[f_1^2(e)\right]. \tag{28}$$

Next, we use the Taylor expansion method [42] to derive the expectations in (28). Thus, taking the Taylor expansion of $f_1(e)$ regarding e_a around the noise v, we have

$$\begin{aligned} f_1(e) &= f_1(e_a + v) \\ &= f_1(v) + f_1'(v)e_a + \frac{1}{2}f_1''(v)e_a^2 + o(e_a^2), \end{aligned} \tag{29}$$

with $f_1'(v)$ and $f_1''(v)$ being the first and second derivatives, and $o(e_a^2)$ being the third and other higher-order terms. Define the steady-state EMSE of RNLMAT by $\xi_1 = \lim_{i\to\infty} E\left[e_a^2(i)\right] = E\left[e_a^2\right]$. Then, if $o(e_a^2)$ is small enough, we can obtain the following results according to Assumptions $A3-A5$

$$\begin{aligned} E\left[e_a f_1(e)\right] &= E\left[e_a f_1(v) + e_a^2 f_1'(v) + o(e_a^2)\right] \\ &\approx E\left[f_1'(v)\right]\xi_1, \end{aligned} \tag{30}$$

$$E\left[f_1^2(e)\right] \approx E\left[f_1^2(v)\right] + E\left[f_1(v)f_1''(v) + |f_1'(v)|^2\right]\xi_1. \tag{31}$$

Substituting (30) and (31) into (28) generates

$$2E\left[f'_1(v)\right]\xi_1 = \mu\mathrm{Tr}(\mathbf{R_u}) \times \left(E\left[f_1^2(v)\right] + E\left[f_1(v)f''_1(v) + |f'_1(v)|^2\right]\xi_1\right). \tag{32}$$

Finally, from (32), we can obtain the steady-state EMSE of RNLMAT as follows:

$$\xi_1 = \frac{\mu \mathrm{Tr}(\mathbf{R_u})E\left[f_1^2(v)\right]}{2E\left[f_1'(v)\right] - \mu\mathrm{Tr}(\mathbf{R_u})E\left[f_1(v)f_1''(v) + |f_1'(v)|^2\right]}, \tag{33}$$

where $f_1(v)$, $f_1'(v)$ and $f_1''(v)$ are given respectively by

$$f_1(v) = \frac{|v|v}{1+\beta|v|^3}, \tag{34}$$

$$f_1'(v) = \frac{2|v| - \beta|v|^4}{\left(1+\beta|v|^3\right)^2}, \tag{35}$$

$$f_1''(v) = \frac{2\mathrm{sign}[v] - 14v^3 + 2\beta^2|v|^5 v}{\left(1+\beta|v|^3\right)^3}. \tag{36}$$

Thus, given a PDF of noise $p_v(\cdot)$, one can obtain the expectations $E[f_1^2(v)]$, $E[f_1'(v)]$ and $E[f_1(v)f_1''(v) + |f_1'(v)|^2]$ in (33) by numerical integration. A simple way to calculate these expectations is shown as follows:

$$\begin{cases} E\left[f_1^2(v)\right] = \frac{1}{L}\sum_{i=1}^{L} f_1^2(v(i)), \\ E\left[f_1'(v)\right] = \frac{1}{L}\sum_{i=1}^{L} f_1'(v(i)), \\ E\left[f_1(v)f_1''(v) + |f_1'(v)|^2\right] \\ \quad = \frac{1}{L}\sum_{i=1}^{L}\left(f_1(v(i))f_1''(v(i)) + |f_1'(v(i))|^2\right), \end{cases} \tag{37}$$

where L denotes the number of noise samples. Therefore, the theoretical steady-state EMSEs of RNLMAT for different noises are obtained by combining (33) and (37).

B. PERFORMANCE OF SWRNLMAT

When SWRNLMAT approaches the steady-state, assuming that ζ is sufficient large, we obtain

$$\mathbf{w}(i+1) = \mathbf{w}(i) + \frac{e^2(i)\text{sign}[e(i)]}{\zeta\,\|\mathbf{u}(i)\|^3}\mathbf{u}(i). \tag{38}$$

Subtracting $\mathbf{w}_o$ from both sides of (38) and combining the weight deviation vector $\widetilde{\mathbf{w}}(i)$ generate

$$\widetilde{\mathbf{w}}(i+1) = \widetilde{\mathbf{w}}(i) - \frac{e^2(i)\text{sign}[e(i)]}{\zeta\,\|\mathbf{u}(i)\|^3}\mathbf{u}(i). \tag{39}$$

Under Assumption $A1$, the dimensionality N is assumed to be sufficiently large such that $\|\mathbf{u}(i)\|^2 \approx N\sigma_u^2$, which is usually used in [19], [21], and [24]. Actually this is also an ergodic assumption, which means that the time average over the taps is equal to its ensemble average. Then, according to Assumptions $A1 - A4$, we rewrite (39) as

$$\widetilde{\mathbf{w}}(i+1) = \widetilde{\mathbf{w}}(i) - \mu f_2(e(i))\mathbf{u}(i). \tag{40}$$

with $\mu = \frac{1}{\zeta N\sqrt{N}\sigma_u^3}$ and $f_2(e(i)) = e^2(i)\text{sign}[e(i)]$.

Define the steady-state EMSE of SWRNLMAT by ξ_2. Similar to the steady-state EMSE of RNLMAT, the steady-state EMSE of SWRNLMAT can be derived as follows:

$$\begin{aligned} \xi_2 &= \frac{\mu \text{Tr}(\mathbf{R_u})E\left[f_2^2(v)\right]}{2E\left[f_2'(v)\right] - \mu\text{Tr}(\mathbf{R_u})E\left[f_2(v)f_2''(v) + |f_2'(v)|^2\right]} \\ &= \frac{\mu\text{Tr}(\mathbf{R_u})E\left[v^4\right]}{4E\left[|v|\right] - 6\mu\text{Tr}(\mathbf{R_u})E\left[v^2\right]} \\ &= \frac{\text{Tr}(\mathbf{R_u})E\left[v^4\right]}{4\zeta N\sqrt{N}\sigma_u^3 E\left[|v|\right] - 6\text{Tr}(\mathbf{R_u})E\left[v^2\right]} \\ &\approx \frac{E\left[v^4\right]}{4\zeta\sqrt{N}\sigma_u E\left[|v|\right] - 6E\left[v^2\right]}, \end{aligned} \tag{41}$$

where we use the approximation $\text{Tr}(\mathbf{R_u}) \approx N\sigma_u^2$ for large N [19], [21], [24] in the last line.

The theoretical results of (33) and (41) can be suitable for any distribution of noise if its PDF is known. It is worth noting that the steady-state EMSE of (33) or (41) is valid only under the assumption that the steady-state *a priori* estimation error is small enough to make its third and higher-order terms negligible. In addition, the validity of (41) can only be guaranteed when the dimension N is sufficient large owing to the used approximation $\|\mathbf{u}(i)\|^2 \approx N\sigma_u^2$.

V. SIMULATION RESULTS

To validate the performance of the proposed RNLMAT, SWRNLMAT, and VSSRNLMAT algorithms, we perform Monte Carlo (MC) independent runs in the FIR system identification shown in Fig. 1. The input signal is a Gaussian sequence with zero-mean and variance σ_u^2. The optimal weight vector $\mathbf{w}_o$ with length $N = 20$ and equal parameters $w_{o,i} = \frac{1}{\sqrt{N}}, i = 1, 2, \ldots, N$, are used to model the unknown system. The initial adaptive weight vector of adaptive filters is a zero vector with length N, unless otherwise specified. And the noises with different distributions are used to model impulsive and non-impulsive noise environments. Here, the Gaussian and uniform noises are chosen to denote non-impulsive noises, and the α-stable noise and the mixed Gaussian and Laplace noises are chosen to denote impulsive noises. In each simulation, the simulated results are obtained from the ensemble averages of 100 independent runs.

A. THEORETICAL VERIFICATION

The simulations are conducted to validate the theoretical steady-state EMSEs of RNLMAT and SWRNLMAT. The initial weight is a zero vector, the input is a Gaussian sequence with zero-mean and unit variance, and the noise is a zero-mean uniform sequence. To verify the theoretical results of RNLMAT, the theoretical steady-state EMSE in (33) is compared with the simulated one under the cases of different values of step-size and noise variance. In each simulation, 20000 iterations are performed, and the steady-state EMSEs are obtained as averages over the last 1000 iterations. The theoretical and simulated steady-state EMSEs of RNLMAT are shown in Fig. 3. As can be seen from Fig. 3, the simulated steady-state EMSEs match well with the theoretical ones when the step-size is small ($\mu < 0.05$).

For SWRNLMAT, the theoretical steady-state EMSE is derived under the assumption that N is sufficient large. Thus, we choose $N = 64$ for the simulation, and other settings are the same as those in Fig. 3. To verify the theoretical results of SWRNLMAT, the theoretical steady-state EMSE in (41) is compared with the simulated one under the cases of different values of scaling parameter ζ and noise variance. The theoretical and simulated steady-state EMSEs of SWRNLMAT are shown in Fig. 4. From Fig. 4, we see that the simulated steady-state EMSEs match well with the theoretical ones in the considered uniform noises.

To further confirm the theoretical steady-state EMSEs of RNLMAT and SWRNLMAT in different noises, the zero-mean Gaussian noise with unit variance, the binary noise at -1 or 1 with the same probability, and the zero-mean Laplace noise with unit variance are considered for simulations. The simulation settings of RNLMAT and SWRNLMAT are the same as those in Fig. 3 and Fig. 4, respectively.

TABLE 3. Theoretical and simulated steady-state EMSEs for different noises.

Algorithm	Parameter	Noise	Theory	Simulation
RNLMAT	$\mu = 0.001,\ \beta = 1$	Gaussian	0.0038112	0.0038646 ± 0.0010584
	$\mu = 0.001,\ \beta = 1$	Binary	0.010037	0.0099795 ± 0.0027938
	$\mu = 0.001,\ \beta = 1$	Laplace	0.0028234	0.0028604 ± 0.00085707
SWRNLMAT	$\zeta = 5$	Gaussian	0.024123	0.025481 ± 0.003815
	$\zeta = 5$	Binary	0.0064109	0.006843 ± 0.0010152
	$\zeta = 5$	Laplace	0.054922	0.054968 ± 0.010884

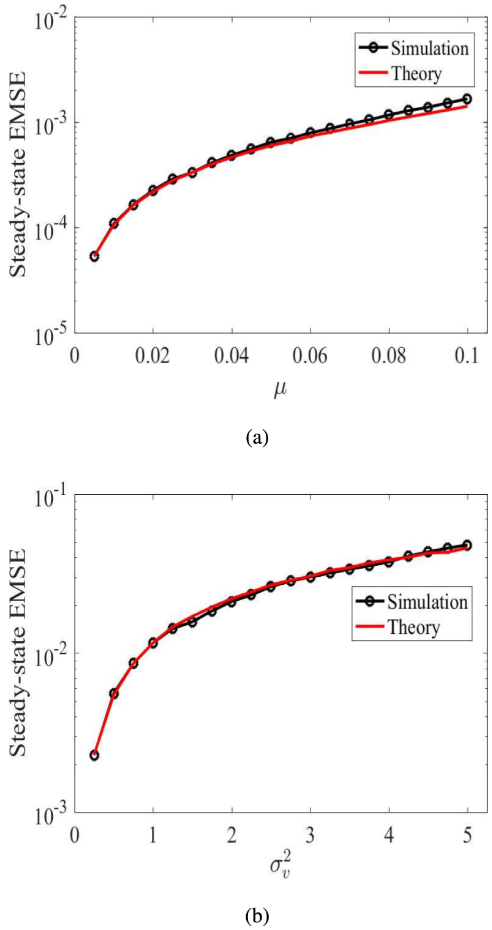


FIGURE 3. Theoretical and simulated steady-state EMSEs of RNLMAT in uniform noises: (a) with different step-sizes ($\beta = 1$, $\sigma_v^2 = 0.01$); (b) with different noise variances ($\beta = 1$, $\mu = 0.002$).

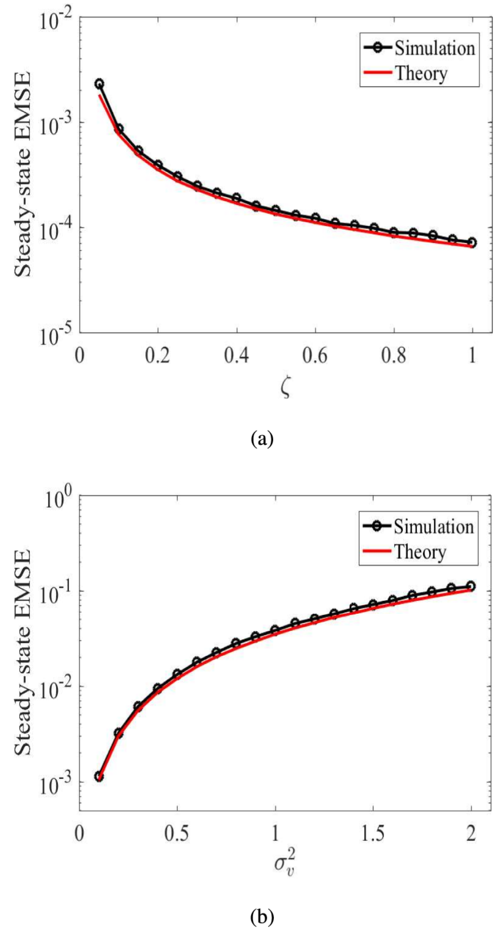


FIGURE 4. Theoretical and simulated steady-state EMSEs of SWRNLMAT in uniform noises: (a) with different scaling parameter ζ ($\sigma_v^2 = 0.01$); (b) with different noise variances ($\zeta = 2$).

The theoretical and simulated steady-state EMSEs of RNLMAT and SWRNLMAT are presented in Table 3. We see from Table 3 that the theoretical steady-state performance analyses of RNLMAT and SWRNLMAT are also valid for different noise distributions.

B. PERFORMANCE COMPARISON

1) NON-IMPULSIVE NOISES

We first compare the MSDs of RNLMAT and SWRNLMAT to those of NLMS [41], NLMAT [24], and NLMF5 [21] by the simulations conducted in the non-impulsive noise environments, i.e., the Gaussian and uniform noises, respectively.

2) GAUSSIAN NOISE:

In the Gaussian noise environment, both the input and noise are zero-mean white Gaussian sequences with variance $\sigma_u^2 = 1$ and variance $\sigma_v^2 = 0.01$ respectively. We first discuss the influence of β on the performance of RNLMAT. The learning curves based on the MSDs of RNLMAT with different β are shown in Fig. 5, where the step-size is chosen to achieve almost the same steady-state performance for each curve.

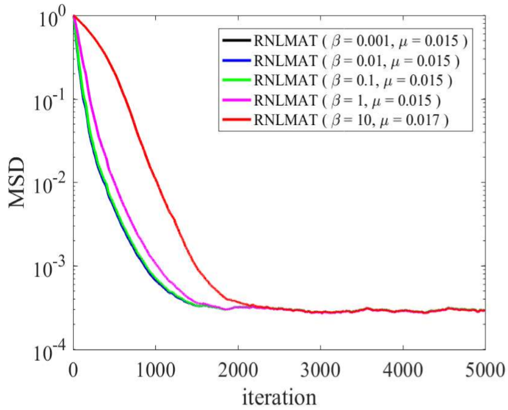


FIGURE 5. MSDs of RNLMAT with different β in the Gaussian noise environment.

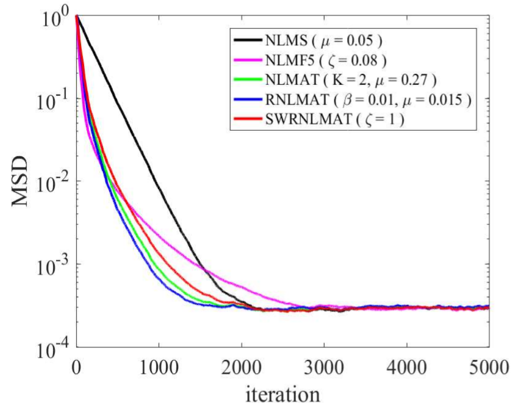


FIGURE 6. MSDs of RNLMAT, SWRNLMAT, NLMS, NLMAT, and NLMF5 with zero weight initialization in the Gaussian noise environment.

From Fig. 5, we see that the convergence rate of RNLMAT can only be influenced by large values of β. Thus, we can find an optimal β by trial and error methods in practice. In the following simulations, we choose β such that the algorithm achieves the desirable performance.

Next, the simulations are conducted to compare the filtering performance of RNLMAT and SWRNLMAT with the other algorithms. The parameters of filters are chosen such that the compared algorithms have almost the same steady-state performance. Hence, we compare the convergence rate of filters in Gaussian noises. First, the initial weight vector of filters is assumed to be a zero vector, and thus the initial MSD is 1. The compared results are shown in Fig. 6. As can be seen from Fig. 6, RNLMAT has faster convergence rate than SWRNLMAT, NLMS, NLMAT, and NLMF5, and SWRNLMAT is slower than NLMAT but faster than NLMS and NLMF5. We next consider the large weight initialization in the filters. The initial weight vector is set as $w_{1,i} = \frac{1+m}{\sqrt{N}}$ with $i = 1, 2, \ldots, N$ and $m = 100$, and thus the corresponding initial MSD is 10^4. The compared MSDs of above algorithms are shown in Fig. 7. We see from Fig. 7 that both RNLMAT and SWRNLMAT provide faster convergence rate than the other three algorithms, and SWRNLMAT has the fastest convergence rate in all the compared algorithms.

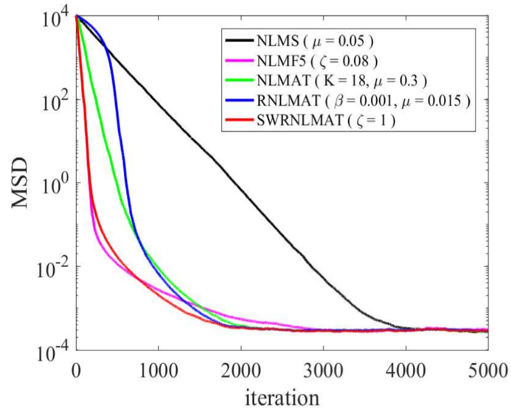


FIGURE 7. MSDs of RNLMAT, SWRNLMAT, NLMS, NLMAT, and NLMF5 with large weight initialization in the Gaussian noise environment.

3) UNIFORM NOISE:

In the uniform noise environment, the uniform noise with zero-mean and variance $\sigma_v^2 = 0.01$ is chosen for simulations. The input signal and the criterion for choosing parameters of filters are the same as those in Fig. 6. Similar to the settings in Figs. 6 and 7, we also consider the zero and large weight initializations of filters. The simulation results with zero and large weight initializations are shown in Figs. 8 and 9, respectively. From Fig. 8, we see that RNLMAT has slightly slower convergence rate than NLMF5 but faster than SWRNLMAT, NLMAT, and NLMS. It can be seen from Fig. 9 that, SWRNLMAT has slightly slower convergence rate than NLMF5 but dramatically faster than RNLMAT, NLMAT and NLMS. Therefore, in uniform noises, the proposed RNLMAT and SWRNLMAT approach the convergence rate of NLMF5 for the cases of small and large weight initializations, respectively. It is worth noting that, RNLMAT has the lowest computational complexity in all the compared algorithms. Thus, for the zero initial weight

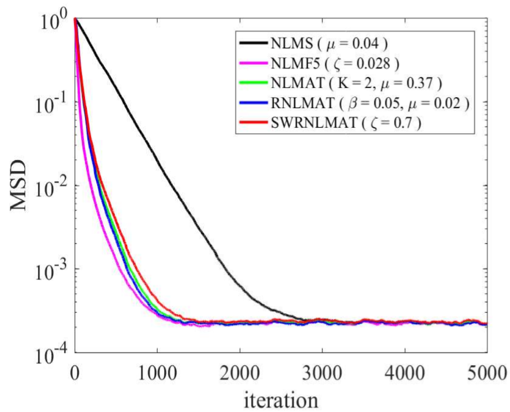


FIGURE 8. MSDs of RNLMAT, SWRNLMAT, NLMS, NLMAT, and NLMF5 with zero weight initialization in the uniform noise environment.

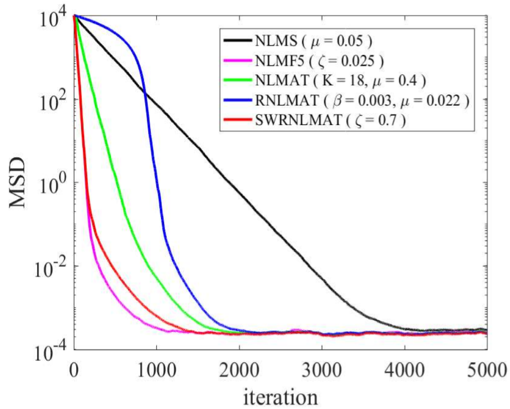


FIGURE 9. MSDs of RNLMAT, SWRNLMAT, NLMS, NLMAT, and NLMF5 with large weight initialization in the uniform noise environment.

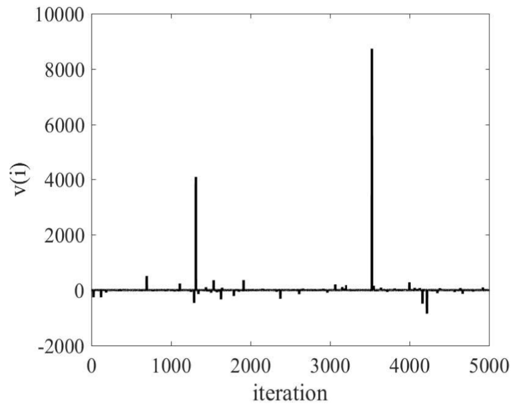


FIGURE 10. α-stable noises with $V_{\alpha-\text{stable}}(0.8, 0, 0.1, 0)$.

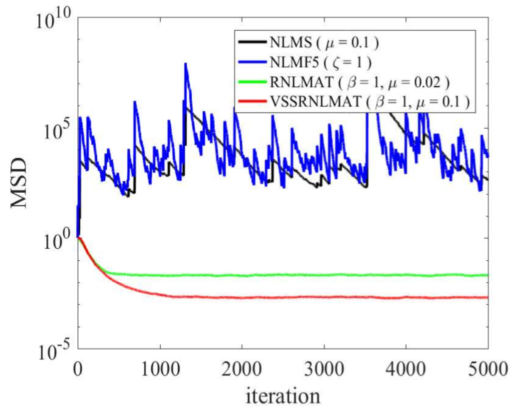


FIGURE 11. MSDs of RNLMAT, VSSRNLMAT, NLMS, and NLMF5 in the α-stable noise environment.

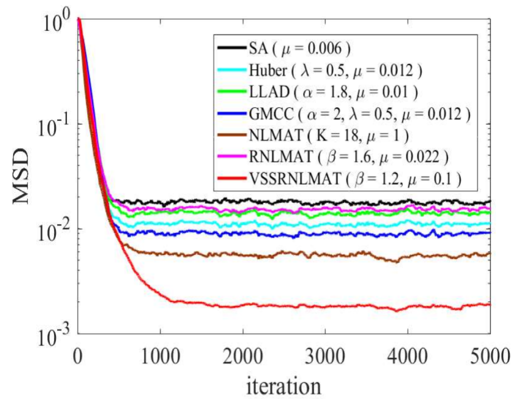


FIGURE 12. MSDs of RNLMAT, VSSRNLMAT, NLMAT, SA, LLAD, Huber, and GMCC in the α-stable noise environment.

vector case, RNLMAT is the best choice in the uniform noise environment.

4) IMPULSIVE NOISES

To validate the robustness of RNLMAT and VSSRNLMAT against impulsive noise, the α-stable noise and the mixed Gaussian and super-Gaussian noises are chosen for modeling impulsive noises in the following.

5) α-STABLE NOISE:

The α-stable noise with a heavy-tailed PDF is usually used to model the impulsive noise, and its characteristic function can be described as follows [43]:

$$f(t) = \exp\{j\delta t - \gamma |t|^{\alpha}[1 + j\beta \text{sign}(t)S(t, \alpha)]\},$$

where

$$S(t, \alpha) = \begin{cases} \tan \frac{\alpha\pi}{2}, & \text{if } \alpha \neq 1 \\ \frac{2}{\pi} \log |t|, & \text{if } \alpha = 1 \end{cases},$$

and $\alpha \in (0, 2]$ is the characteristic factor, $\beta \in [-1, 1]$ is the symmetry parameter, $\gamma > 0$ is the dispersion parameter, and $-\infty < \delta < \infty$ is the location parameter. For simplicity, the parameters of the characteristic function are set as a parameter vector $V_{\alpha-\text{stable}}(\alpha, \beta, \gamma, \delta)$.

In this simulation, the parameters of α-stable noise is set as $V_{\alpha-\text{stable}}(0.8, 0, 0.1, 0)$. The input signal is a Gaussian sequence with zero-mean and variance $\sigma_u^2 = 1$. First, we compare the MSD performance of RNLMAT with those of VSSRNLMAT, NLMS and, NLMF5. The noise sequence is shown in Fig. 10, and we see from Fig. 10 that there exist large impulses in the noise sequence. The corresponding compared MSDs are shown in Fig. 11. As we can see from Fig. 11, both RNLMAT and VSSRNLMAT can combat impulsive noises effectively and VSSRNLMAT has better filtering accuracy than RNLMAT. Fig. 11 also illustrates that NLMS and NLMF5 cannot converge in α-stable noises. Therefore, the typical robust adaptive filters, i.e., NLMAT [24], sign algorithm (SA) [44], least logarithmic absolute difference (LLAD) [29], robust Huber (Huber) [32], and generalized maximum correntropy criterion (GMCC) algorithms are chosen for comparisons in the same environment as in Fig. 11. The parameters are chosen such that each algorithm achieves the desirable performance. The compared MSDs of all algorithms are shown in Fig. 12.

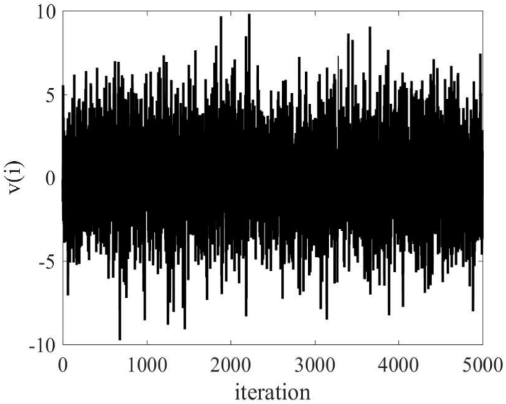


FIGURE 13. Mixed Gaussian and Laplace noises.

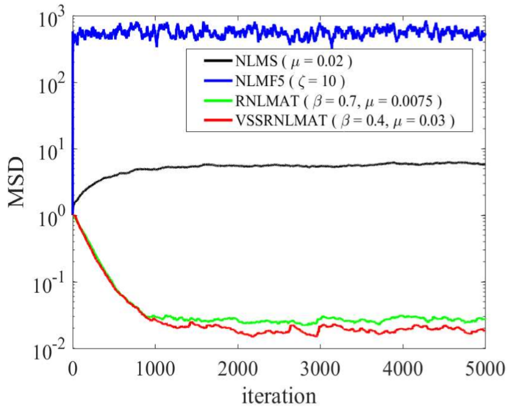


FIGURE 14. MSDs of RNLMAT, VSSRNLMAT, NLMS, and NLMF5 in the mixed Gaussian and Laplace noises environment.

Here, $K = 18$ is configured in NLMAT to guarantee its convergence. From Fig. 12, we obtain that VSSRNLMAT dramatically outperforms the other robust adaptive filters including NLMAT and GMCC in the considered impulsive noise.

6) MIXED GAUSSIAN AND LAPLACE NOISES:

Generally, the mixed Gaussian and other noises are also used to model the impulsive noises [26], [29]. Here, we use the mixed Gaussian and Laplace noises to model the impulsive noise. The noise model is described as $v(i) = v_1(i)+b(i)v_2(i)$, where $v_1(i)$ is the Laplace noise with zero-mean and unit variance, $v_2(i)$ is a white Gaussian sequence with zero-mean and variance $\sigma_{v_2}^2 = 10^4$ to represent large outliers, and $b(i)$ is generated using a Bernoulli random process with $Pr\{b(i) = 1\} = c$, $Pr\{b(i) = 0\} = 1 - c$ ($0 \leq c \leq 1$ is an occurrence probability). In the following, we select $c = 0.05$. The initial weight vector of adaptive filters is a zero vector and the input is a Gaussian sequence with zero-mean and unit variance.

Similar to Fig. 11 and Fig. 12, we first compare the MSD performance of RNLMAT with those of VSSRNLMAT, NLMS, and NLMF5. Then, the robust adaptive filters are chosen to validate the superiorities of RNLMAT and VSSRNLMAT. Fig. 13 shows the mixed Gaussian and Laplace noises. And the compared results are shown in Fig. 14 and Fig. 15. As can be seen from Fig. 14, RNLMAT and VSSRNLMAT can combat impulsive noises efficiently while NLMS and NLMF5 cannot provide robustness against such noises. It can be seen from Fig. 15 that VSSRNLMAT provides the best filtering accuracy in all the compared robust adaptive filters. And RNLMAT also has the comparable filtering accuracy to GMCC. In addition, NLMAT cannot combat the mixed Gaussian and Laplace noises, efficiently. Therefore, both RNLMAT and VSSRNLMAT can efficiently combat α-stable and mixed impulsive noises, and VSSRNLMAT provides better filtering performance than other robust adaptive filters.

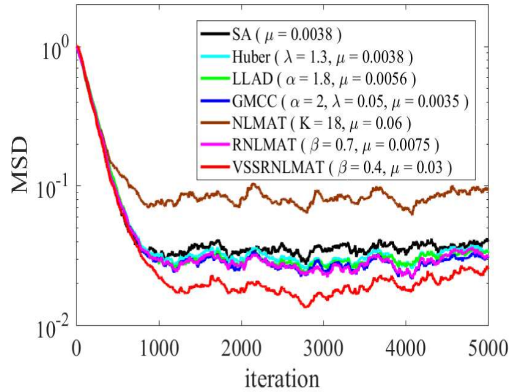


FIGURE 15. MSDs of RNLMAT, VSSRNLMAT, NLMAT, SA, LLAD, Huber, and GMCC in the mixed Gaussian and Laplace noises environment.

To sum up, we see from all the above simulations that when the weight initialization of filters is small, RNLMAT is preferred in non-impulsive noises. For the large weight initialization of filters in non-impulsive noises, SWRNLMAT provides the best filtering performance in the Gaussian noise and the comparable performance to NLMF5 in the uniform noise, respectively. In impulsive noises, VSSRNLMAT is preferred from the aspect of filtering accuracy, and RNLMAT also can provide the similar filtering accuracy to GMCC.

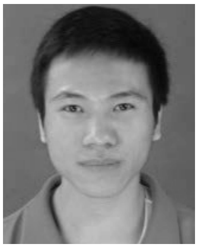
VI. CONCLUSION

A novel robust normalized least mean absolute third (RNLMAT) algorithm is proposed using the third order in the estimation error as the normalization in this paper. The influence of the input and the noise on the stability of RNLMAT can be removed by third order in the estimation error efficiently. RNLMAT is therefore stable for all statistics of the input, noises and the weight initialization, can improve filtering accuracy and robustness in impulsive noises, simultaneously. Using the variable step-size (VSS) and switching (SW) schemes in the RNLMAT and the least mean absolute third (LMAT) algorithms, respectively, the VSSRNLMAT and the SWRNLMAT algorithms are developed for accuracy improvement in impulsive noises and non-impulsive noises respectively. The steady-state performance of RNLMAT and SWRNLMAT in terms of the excess mean-square error (EMSE) is also performed using the

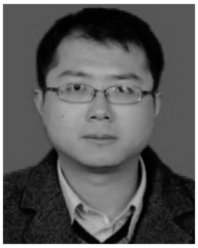
Taylor expansion method. Simulations on system identification validate the theoretical results and the superiorities of RNLMAT, VSSRNLMAT, and SWRNLMAT for different noises. In Gaussian and uniform noises, RNLMAT and SWRNLMAT improve the convergence rate for small and large weight initializations, respectively. In impulsive noises, VSSRNLMAT provides better robustness and filtering accuracy than the other robust adaptive filters including the generalized maximum correntropy criterion (GMCC) and the normalized least mean absolute third (NLMAT) algorithms.

REFERENCES

[1] A. H. Sayed, *Fundamentals of Adaptive Filtering*. Hoboken, NJ, USA: Wiley, 2003.
[2] E. Walach and B. Widrow, "The least mean fourth (LMF) adaptive algorithm and its family," *IEEE Trans. Inf. Theory*, vol. IT-30, no. 2, pp. 275–283, Mar. 1984.
[3] E. Eweda, "Stabilization of high-order stochastic gradient adaptive filtering algorithms," *IEEE Trans. Signal Process.*, vol. 65, no. 15, pp. 3948–3959, Aug. 2017.
[4] S. H. Cho and S. D. Kim, "Adaptive filters based on the high order error statistics," in *Proc. IEEE Asia Pacific Conf. Circuits Syst.*, Seoul, South Korea, Nov. 1996, pp. 109–112.
[5] S. H. Cho, S. D. Kim, H. P. Moom, and J. Y. Na, "Least mean absolute third (LMAT) adaptive algorithm: Mean and mean-squared convergence properties," in *Proc. 6th Western Pacific Reg. Aconst. Conf.*, vol. 22, no. 10, Hong Kong, Nov. 1997, pp. 2303–2309.
[6] Y. H. Lee, J. D. Mok, S. D. Kim, and S. H. Cho, "Performance of least mean absolute third (LMAT) adaptive algorithm in various noise environments," *Electron. Lett.*, vol. 34, no. 3, pp. 241–242, Jan. 1998.
[7] S. H. Cho, S. D. Kim, and K. Y. Jen, "Statistical convergence of the adaptive least mean fourth algorithm," in *Proc. Int. Conf. Signal Process. (ICSP)*, Oct. 1996, pp. 610–613.
[8] P. I. Hübscher and J. C. M. Bermudez, "An improved statistical analysis of the least mean fourth (LMF) adaptive algorithm," *IEEE Trans. Signal Process.*, vol. 51, no. 3, pp. 664–671, Mar. 2003.
[9] P. I. Hübscher, V. H. Nuscimento, and J. C. M. Bermudez, "New results on the stability analysis of the LMF (least mean fourth) adaptive algorithm," in *Proc. ICASSP*, Apr. 2003, pp. VI369–VI372.
[10] V. H. Nascimento and J. C. M. Bermudez, "Probability of divergence for the least-mean fourth algorithm," *IEEE Trans. Signal Process.*, vol. 54, no. 4, pp. 1376–1385, Apr. 2006.
[11] P. I. Hübscher, J. C. M. Bermudez, and V. H. Nascimento, "A mean-square stability analysis of the least mean fourth adaptive algorithm," *IEEE Trans. Signal Process.*, vol. 55, no. 8, pp. 4018–4028, Aug. 2007.
[12] A. Zerguine, "Convergence behavior of the normalized least mean fourth algorithm," in *Proc. 34th Annu. ASILOMAR Conf. Signals, Syst., Comput.*, Pacific Grove, CA, USA, Oct./Nov. 2000, pp. 275–278.
[13] M. K. Chan and C. F. N. Cowan, "Using a normalised LMF algorithm for channel equalisation with co-channel interference," in *Proc. EUSIPCO*, Toulouse, France, Sep. 2002, pp. 48–51.
[14] A. Zerguine, "Convergence and steady-state analysis of the normalized least mean fourth algorithm," *Digit. Signal Process.*, vol. 17, no. 1, pp. 17–31, Jan. 2007.
[15] E. Eweda and A. Zerguine, "New insights into the normalization of the least mean fourth algorithm," *Signal, Image, Video Process.*, vol. 7, no. 2, pp. 255–262, Mar. 2013.
[16] A. Zerguine, M. K. Chan, T. Y. Al-Naffouri, M. Moinuddin, and C. F. N. Cowan, "Convergence and tracking analysis of a variable normalised LMF (XE-NLMF) algorithm," *Signal Process.*, vol. 89, no. 5, pp. 778–790, May 2009.
[17] E. Eweda and A. Zerguine, "A normalized least mean fourth algorithm with improved stability," in *Proc. 44th Asilomar Conf. Signals, Syst. Comput.*, Pacific Grove, CA, USA, Nov. 2010, pp. 1002–1005.
[18] E. Eweda, "Global stabilization of the least mean fourth algorithm," *IEEE Trans. Signal Process.*, vol. 60, no. 3, pp. 1473–1477, Mar. 2012.
[19] E. Eweda and N. J. Bershad, "Stochastic analysis of a stable normalized least mean fourth algorithm for adaptive noise canceling with a white Gaussian reference," *IEEE Trans. Signal Process.*, vol. 60, no. 12, pp. 6235–6244, Dec. 2012.
[20] E. Eweda, "Mean-square stability analysis of a normalized least mean fourth algorithm for a Markov plant," *IEEE Trans. Signal Process.*, vol. 62, no. 24, pp. 6545–6553, Dec. 2014.
[21] E. Eweda, "A stable normalized least mean fourth algorithm with improved transient and tracking behaviors," *IEEE Trans. Signal Process.*, vol. 64, no. 18, pp. 4805–4816, Sep. 2016.
[22] S. Guan and Z. Li, "Nonparametric variable step-size LMAT algorithm," *Circuits, Syst., Signal Process.*, vol. 36, no. 3, pp. 1322–1339, 2017.
[23] S. H. Guan and Z. Li, "Optimal step size of least mean absolute third algorithm," *Signal, Image Video Process.*, vol. 11, no. 6, pp. 1105–1113, Feb. 2017.
[24] H. Zhao, Y. Yu, S. Gao, X. Zeng, and Z. He, "A new normalized LMAT algorithm and its performance analysis," *Signal Process.*, vol. 105, no. 12, pp. 399–409, Jun. 2014.
[25] W. Liu, P. P. Pokharel, and J. C. Príncipe, "Correntropy: Properties and applications in non-Gaussian signal processing," *IEEE Trans. Signal Process.*, vol. 55, no. 11, pp. 5286–5298, Nov. 2007.
[26] B. Chen, L. Xing, H. Zhao, N. Zheng, and J. C. Príncipe, "Generalized correntropy for robust adaptive filtering," *IEEE Trans. Signal Process.*, vol. 64, no. 13, pp. 3376–3387, Jul. 2016.
[27] G. Qian and S. Wang, "Generalized complex correntropy: Application to adaptive filtering of complex data," *IEEE Access*, vol. 6, pp. 19113–19120, 2018.
[28] X. Liu, B. Chen, H. Zhao, J. Qin, and J. Cao, "Maximum correntropy Kalman filter with state constraints," *IEEE Access*, vol. 5, pp. 25846–25853, 2017.
[29] M. O. Sayin, N. D. Vanli, and S. S. Kozat, "A novel family of adaptive filtering algorithms based on the logarithmic cost," *IEEE Trans. Signal Process.*, vol. 62, no. 17, pp. 4411–4424, Sep. 2014.
[30] L. Lu, H. Zhao, and B. Chen, "Improved-variable-forgetting-factor recursive algorithm based on the logarithmic cost for Volterra system identification," *IEEE Trans. Circuits Syst. II, Exp. Briefs.*, vol. 63, no. 6, pp. 588–592, Jun. 2016.
[31] J. Chambers and A. Avlonitis, "A robust mixed-norm adaptive filter algorithm," *IEEE Signal Process. Lett.*, vol. 4, no. 2, pp. 46–48, Feb. 1997.
[32] P. Petrus, "Robust Huber adaptive filter," *IEEE Trans. Signal Process.*, vol. 47, no. 4, pp. 1129–1133, Apr. 1999.
[33] E. V. Papoulis and T. Stathaki, "A normalized robust mixed-norm adaptive algorithm for system identification," *IEEE Signal Process. Lett.*, vol. 11, no. 1, pp. 56–59, Jan. 2004.
[34] W. Ma, J. Duan, W. Man, J. Liang, and B. Chen, "General mixed-norm-based diffusion adaptive filtering algorithm for distributed estimation over network," *IEEE Access*, vol. 5, pp. 1090–1102, 2017.
[35] S. Zhang, W. X. Zheng, J. Zhang, and H. Han, "A family of robust M-shaped error weighted least mean square algorithms: Performance analysis and echo cancellation application," *IEEE Access*, vol. 5, pp. 14716–14727, Aug. 2017.
[36] F. Huang, J. Zhang, and S. Zhang, "Maximum Versoria criterion-based robust adaptive filtering algorithm," *IEEE Trans. Circuits Syst. II, Exp. Briefs.*, vol. 64, no. 10, pp. 1252–1256, Oct. 2017.
[37] T. Y. Al-Naffouri and A. H. Sayed, "Adaptive filters with error nonlinearities: Mean-square analysis and optimum design," *EURASIP J. Adv. Signal Process.*, vol. 4, no. 1, pp. 192–205, Dec. 2001.
[38] H.-C. Huang and J. Lee, "A new variable step-size NLMS algorithm and its performance analysis," *IEEE Trans. Signal Process.*, vol. 60, no. 4, pp. 2055–2060, Apr. 2012.
[39] F. Zhang, Y. Wang, and B. Ai, "Variable step-size MLMS algorithm for digital predistortion in wideband OFDM systems," *IEEE Trans. Consum. Electron.*, vol. 61, no. 1, pp. 10–15, Feb. 2015.
[40] K. Xiong and S. Wang, "Robust least mean logarithmic square adaptive filtering algorithms," *J. Franklin Inst.*, vol. 356, no. 1, pp. 654–674, Jan. 2019.
[41] M. Tarrab and A. Feuer, "Convergence and performance analysis of the normalized LMS algorithm with uncorrelated Gaussian data," *IEEE Trans. Inf. Theory*, vol. IT-34, no. 4, pp. 680–691, Jul. 1988.
[42] B. Chen, L. Xing, J. Liang, N. Zheng, and J. C. Príncipe, "Steady-state mean-square error analysis for adaptive filtering under the maximum correntropy criterion," *IEEE Signal Process. Lett.*, vol. 21, no. 7, pp. 880–884, Jul. 2014.
[43] B. Weng and K. E. Barner, "Nonlinear system identification in impulsive environments," *IEEE Trans. Signal Process.*, vol. 53, no. 7, pp. 2588–2594, Jul. 2005.
[44] V. Mathews and S. Cho, "Improved convergence analysis of stochastic gradient adaptive filters using the sign algorithm," *IEEE Trans. Acoust., Speech, Signal Process.*, vol. ASSP-35, no. 4, pp. 450–454, Apr. 1987.



KUI XIONG received the B.Eng. degree in electronic information science and technology from Zhejiang Sci-Tech University, Hangzhou, China, in 2017. He is currently pursuing the M.Eng. degree with the College of Electronic and Information Engineering, Southwest University, Chongqing, China. His current research interests include kernel adaptive filtering, information theoretic learning, and optimization.



SHIYUAN WANG (M'13–SM'18) received the B.Eng. and M.Eng. degrees in electronic and information engineering from Southwest Normal University, Chongqing, China, in 2002 and 2005, respectively, and the Ph.D. degree in circuit and system from Chongqing University, Chongqing, China, in 2011. He was a Research Associate with The Hong Kong Polytechnic University, from 2012 to 2013. He is currently a Professor with the College of Electronic and Information Engineering, Southwest University, China. He has published one book and more than 50 papers. His research interests include adaptive signal processing, nonlinear dynamics and chaos, and bioinformatics. He is an Associate Editor of the IEEE TRANSACTIONS ON CIRCUITS AND SYSTEMS II, from 2018 to 2019.

BADONG CHEN (M'10–SM'13) received the B.S. and M.S. degrees in control theory and engineering from Chongqing University, in 1997 and 2003, respectively, and the Ph.D. degree in computer science and technology from Tsinghua University, in 2008, where he was a Postdoctoral Researcher from 2008 to 2010. He was a Postdoctoral Associate with the Computational NeuroEngineering Laboratory, University of Florida, from 2010 to 2012. He visited Nanyang Technological University as a Visiting Research Scientist in 2015. He is currently a Professor with the Institute of Artificial Intelligence and Robotics, Xi'an Jiaotong University. He has published two books, four chapters, and over 200 papers in various journals and conference proceedings. His research interests include signal processing, information theory, machine learning, and their applications in cognitive science and the engineering. He is an IEEE Senior Member, and a Technical Committee Member of the IEEE SPS Machine Learning for Signal Processing and of the IEEE CIS Cognitive and Developmental Systems. He is an Associate Editor of the IEEE TRANSACTIONS ON COGNITIVE AND DEVELOPMENTAL SYSTEMS, the IEEE TRANSACTIONS ON NEURAL NETWORKS AND LEARNING SYSTEMS, and the *Journal of The Franklin Institute*. He has been on the Editorial Board of *Entropy*.