

Received November 14, 2020, accepted November 26, 2020, date of publication December 3, 2020,
date of current version December 16, 2020.

Digital Object Identifier 10.1109/ACCESS.2020.3042284

$\mathcal{L}_1/\mathcal{L}_2$ -Mode Switching Adaptive Filter Algorithm Based on Novel Mean Square Deviation Analysis

MINHO LEE^{ID}, (Graduate Student Member, IEEE), TAESU PARK^{ID}, (Student Member, IEEE),
JUNWOONG HUR^{ID}, AND POOGYEON PARK^{ID}, (Senior Member, IEEE)

Department of Electrical Engineering, Pohang University of Science and Technology, Pohang 37673, South Korea

Corresponding author: Poogyeon Park (ppg@postech.ac.kr)

This work was supported in part by the Ministry of Science and ICT (MSIT), South Korea, through the ICT under Grant IITP-2019-2011-1-00783, and in part by the Korea Institute of Energy Technology Evaluation and Planning (KETEP) grant funded by the Korea Government (MOTIE) under Grant 20182020109350.

ABSTRACT This paper proposes an $\mathcal{L}_1/\mathcal{L}_2$ -mode switching adaptive filter algorithm by comparing the performances in \mathcal{L}_1 and \mathcal{L}_2 modes. In the \mathcal{L}_1 or \mathcal{L}_2 mode, the proposed algorithm adopts, as the update equation, that of the normalized sign (NS) algorithm or that of the normalized least mean square (NLMS) algorithm, respectively. By analyzing the mean square deviations (MSDs) of the NS algorithm as well as of the NLMS algorithm, the algorithm selects the better mode in the sense that the next MSD of the algorithm in the selected mode becomes smaller than that in the other mode. The algorithm mainly operates in the \mathcal{L}_1 mode when the impulsive noises occur but in the \mathcal{L}_2 mode otherwise, which leads to robustness against the impulsive noises like the NS algorithm and also leads to a low steady-state misalignment like the NLMS algorithm. Furthermore, the proposed algorithm is faster than the NS and the NLMS algorithms in terms of the convergence rate. A modified reset algorithm is also applied to maintain performance when the unknown system is abruptly changed. Simulations conducted in various system identification scenarios show that the proposed algorithm outperforms the conventional algorithms in terms of the convergence rate and the steady-state misalignment, whenever impulsive noises exist or not.

INDEX TERMS Impulsive noise, $\mathcal{L}_1/\mathcal{L}_2$ -mode switching adaptive filter, mean square deviation (MSD) analysis, normalized least mean square (NLMS) algorithm, normalized sign (NS) algorithm.

I. INTRODUCTION

Adaptive filters have been widely used in many applications such as channel estimation, echo cancellation, system identification, active noise control, and biological signal processing [1]–[4]. Least mean square (LMS) algorithm and normalized LMS (NLMS) algorithm are consistently studied owing to their simplicity and low computational complexity [5]–[9]. However, these algorithms suffer from performance deterioration in the impulsive noise environments because they are derived from an \mathcal{L}_2 -norm optimization.

To surmount this defect of the algorithms based on the \mathcal{L}_2 -norm optimization, several approaches are proposed in the literature. Adaptive filters based on Huber function are

The associate editor coordinating the review of this manuscript and approving it for publication was Jingen Ni^{ID}.

proposed, where the cost functions are designed by using the Huber function or an modified Huber functions to retain robustness against to the impulsive noises [10]–[12]. Algorithms using step size scaler (SSS) are also proposed in [13], [14], where the SSS is derived from \mathcal{L}_2 -norm based cost functions that is modified by utilizing the tangent hyperbolic function and log function.

The \mathcal{L}_1 -norm based algorithms that use the sign of the error signal in the weight update equation are also presented in [15]–[17]. These \mathcal{L}_1 -norm based algorithms have intuitively simple structure and are not affected by the impulsive noises. However, the convergence rates of them are slower than those of the \mathcal{L}_2 -norm-based algorithms in the non-impulsive noise environments as they just use the sign of the error signal which has partial information [18], [19].

To improve the performance, the algorithms using the \mathcal{L}_1 and \mathcal{L}_2 norms together have been proposed. Among

them, the family of mixed-norm adaptive filters has been presented by optimizing the weighted sum of both \mathcal{L}_1 and \mathcal{L}_2 norms [20]–[23]. Firstly, the robust mixed-norm (RMN) algorithm was proposed, which linearly combines the LMS and least absolute deviation (LAD) algorithms [20]. Then, the normalized version of RMN (NRMN) was presented, which shows better performance in a non-stationary environments [21]. The continuous mixed p -norm (CMPN) adaptive filtering algorithm was also put forward, where the continuous p -norms for $1 \leq p \leq 2$ are combined using a probability density-like function and takes advantage of their strengths [22]. Subsequently, the generalized version of the VSS-CMPN (GVSS-CMPN) was proposed, which introduces a rotated linear function of the probability density-like function [23]. In [22], the probability density-like function which mixes the continuous norms is assumed to be uniform, but GVSS-CMPN generalizes it to the rotated linear function rotated around $(\frac{3}{2}, 1)$ to improve the performance. Although these algorithms based on the mixed-norm show the improved performance, there is a limitation that cannot use the merits of \mathcal{L}_2 -norm-based algorithms in the environments with strong impulsive noises.

To overcome this limitation, the proposed $\mathcal{L}_1/\mathcal{L}_2$ -mode switching adaptive filter algorithm adopts a switching mechanism. In [24] and [25], a switching mechanism is applied to solve the trade-off problem between a convergence rate and a steady-state error, where the switching is activated in variable step size scheme when the adaptive filter is in the steady-state. However, the proposed algorithm switches between \mathcal{L}_1 and \mathcal{L}_2 modes, which have different advantages. The proposed algorithm compares the MSDs in \mathcal{L}_1 and \mathcal{L}_2 modes and selects the better mode at each iteration so that the forthcoming MSD in the selected mode becomes lower than that in the other mode. By analyzing the MSDs of the two modes in a new way, the MSDs can be estimated more accurately and utilized in the switching process in the impulsive noise environments. The algorithm primarily operates in the \mathcal{L}_1 mode when the impulsive noise occurs, otherwise in the \mathcal{L}_2 mode. Therefore, the algorithm can take advantage of both NS and NLMS algorithms. In other words, the algorithm is robust to the impulsive noises like the NS algorithm and has a low steady-state misalignment like the NLMS algorithm. To maintain the performance in the system sudden change, a modified reset algorithm is also applied. The simulation is conducted in various system identification scenarios and the algorithm is compared with other conventional algorithms.

This paper is composed as follows. In section II, the NS and the NLMS algorithms adopted as the update equations in the \mathcal{L}_1 and \mathcal{L}_2 modes are summarized. In section III, the MSDs of the NS and the NLMS algorithms are analyzed in the novel methods and the mode switching scheme is proposed. The simulations are conducted to verify the proposed algorithm and to evaluate the theoretical MSD analyses in section IV. Finally, the conclusion is presented in V.

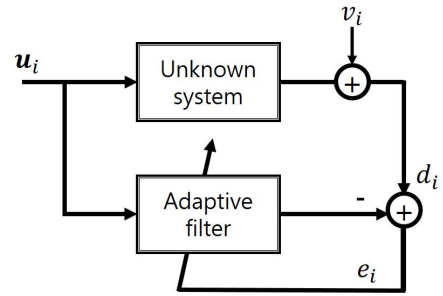


FIGURE 1. System identification structure.

II. REVIEW NS AND NLMS ALGORITHM

In the system identification model (Fig. 1), the desired output of the adaptive filter d_i is obtained as

$$d_i = \mathbf{u}_i^T \mathbf{w}_o + v_i, \quad (1)$$

where \mathbf{w}_o is an M -dimensional optimal weight vector which has to be estimated, v_i is the measurement noise, $\mathbf{u}_i = [u_i, u_{i-1}, \dots, u_{i-M+1}]^T \in \mathbb{R}^M$ is the input vector with the variance σ_u^2 and superscript T stands for the transpose operator of a vector. The measurement noise v_i is modeled as

$$v_i = \theta_i + \eta_i, \quad (2)$$

where θ_i is the white Gaussian sequences with zero-mean and variance σ_θ^2 and η_i is the impulsive noise. *A priori* error e_i is represented as

$$e_i = d_i - \mathbf{u}_i^T \hat{\mathbf{w}}_i, \quad (3)$$

where $\hat{\mathbf{w}}_i = [w_{0,i}, w_{1,i}, \dots, w_{M-1,i}]^T \in \mathbb{R}^M$ is the weight vector of the adaptive filter that is the estimate of the unknown system. It is assumed that the tap length of the filter is same with that of the unknown system.

A. \mathcal{L}_1 -NORM OPTIMIZATION ALGORITHM: NS ALGORITHM

The NS Algorithm can be derived by minimizing the \mathcal{L}_1 -norm of *a priori* error. The cost function is described as follows:

$$J_1(\hat{\mathbf{w}}_i) = \frac{\|e_i\|_1}{\sqrt{\mathbf{u}_i^T \mathbf{u}_i}} \quad (4)$$

where $\|\cdot\|_1$ stands for \mathcal{L}_1 -norm. By applying the adaptive gradient method to the cost function [15], [16], the update equation of the NS algorithm can be derived as

$$\begin{aligned} \hat{\mathbf{w}}_{i+1} &= \hat{\mathbf{w}}_i - \mu_1 \nabla_{\hat{\mathbf{w}}_i} J_1(\hat{\mathbf{w}}_i) \\ &= \hat{\mathbf{w}}_i + \mu_1 \frac{\mathbf{u}_i \text{sgn}(e_i)}{\sqrt{\mathbf{u}_i^T \mathbf{u}_i}}. \end{aligned} \quad (5)$$

where ∇ is the gradient operator and μ_1 is the step size of the NS algorithm. Alternatively, (5) can be derived by applying the Lagrange multipliers method to the \mathcal{L}_1 norm of *a priori* error with the constraint about minimal disturbances [17].

B. \mathcal{L}_2 -NORM OPTIMIZATION ALGORITHM: NLMS ALGORITHM

The NLMS can be derived by minimizing the \mathcal{L}_2 norm of a priori error. The cost function is described as follows:

$$J_2(\hat{\mathbf{w}}_i) = \frac{1}{2} \frac{\|e_i\|_2^2}{\mathbf{u}_i^T \mathbf{u}_i}, \quad (6)$$

where $\|\cdot\|_2$ stands for \mathcal{L}_2 -norm. By applying the adaptive gradient method to the cost function (6), the update equation of the NLMS algorithm can be derived as follows [26], [27]:

$$\begin{aligned} \hat{\mathbf{w}}_{i+1} &= \hat{\mathbf{w}}_i - \mu_2 \nabla_{\hat{\mathbf{w}}_i} J_2(\hat{\mathbf{w}}_i) \\ &= \hat{\mathbf{w}}_i + \mu_2 \frac{\mathbf{u}_i e_i}{\mathbf{u}_i^T \mathbf{u}_i}. \end{aligned} \quad (7)$$

where μ_2 is the step size of the NLMS algorithm. Alternatively, (7) can be derived by applying the Lagrange multipliers method and the principle of minimal disturbance with the constraint about the updated filter's output [1].

III. PROPOSED ALGORITHM

The NS and the NLMS algorithm mentioned in Section II have advantages and disadvantages, respectively. The NS algorithm is robust to the impulsive noises as it uses the sign of the error in the update equation, but its convergence rate is slower than that of the NLMS. On the other hand, the convergence rate of the NLMS is faster than that of the NS algorithm as it is derived from an \mathcal{L}_2 -norm, where the weight update equation uses the direction and size of the gradient unlike the NS algorithm, but its performance is degraded severely by the impulsive noises. Therefore, a mode switching scheme is proposed to compensate for the shortcomings of each algorithm. The \mathcal{L}_1 mode and \mathcal{L}_2 mode use the update equations of the NS and the NLMS algorithms, respectively.

A. NOVEL MSD ANALYSIS OF NS ALGORITHM

From [13] and [28], MSD is defined as

$$\begin{aligned} \text{MSD}_i &\triangleq E(\|\mathbf{w}_o - \hat{\mathbf{w}}_i\|^2 | \mathcal{U}_i \cup \mathcal{D}_i) \\ &\triangleq \bar{E}(\|\mathbf{w}_o - \hat{\mathbf{w}}_i\|^2) \\ &= \bar{E}(\|\tilde{\mathbf{w}}_i\|^2) \triangleq p_i, \end{aligned} \quad (8)$$

where $\tilde{\mathbf{w}}_i = \mathbf{w}_o - \hat{\mathbf{w}}_i$ is the weight error vector, $\mathcal{D}_i = \{d_k | 0 \leq k < i\}$ and $\mathcal{U}_i = \{\mathbf{u}_k | 0 \leq k < i\}$. The weight error vector $\tilde{\mathbf{w}}_{i+1}$ is derived as

$$\tilde{\mathbf{w}}_{i+1} = \mathbf{w}_o - \hat{\mathbf{w}}_{i+1} = \tilde{\mathbf{w}}_i - \frac{\mu_1 \mathbf{u}_i \text{sgn}(e_i)}{\sqrt{\mathbf{u}_i^T \mathbf{u}_i}}. \quad (9)$$

From (9), MSD estimation of NS algorithm is represented as

$$\begin{aligned} p_{i+1} &= \bar{E}(\tilde{\mathbf{w}}_{i+1}^T \tilde{\mathbf{w}}_{i+1}) \\ &= p_i - 2\mu_1 \bar{E} \left(\frac{\text{sgn}(e_i) \mathbf{u}_i^T \tilde{\mathbf{w}}_i}{\sqrt{\mathbf{u}_i^T \mathbf{u}_i}} \right) + \mu_1^2 \end{aligned} \quad (10)$$

As $\text{sgn}(e_i)$ can be represented by $\text{sgn}(e_i) = e_i/|e_i|$, equation (10) is derived as

$$\begin{aligned} p_{i+1} &= p_i - 2\mu_1 \bar{E} \left(\frac{e_i \mathbf{u}_i^T \tilde{\mathbf{w}}_i}{|e_i| \sqrt{\mathbf{u}_i^T \mathbf{u}_i}} \right) + \mu_1^2 \\ &= p_i - 2\mu_1 \bar{E} \left(\frac{(\tilde{\mathbf{w}}_i^T \mathbf{u}_i + v_i) \mathbf{u}_i^T \tilde{\mathbf{w}}_i}{|e_i| \sqrt{\mathbf{u}_i^T \mathbf{u}_i}} \right) + \mu_1^2 \\ &= p_i - 2\mu_1 \bar{E} \left(\frac{\tilde{\mathbf{w}}_i^T \mathbf{u}_i \mathbf{u}_i^T \tilde{\mathbf{w}}_i + \mathbf{u}_i^T \tilde{\mathbf{w}}_i v_i}{|e_i| \sqrt{\mathbf{u}_i^T \mathbf{u}_i}} \right) + \mu_1^2. \end{aligned} \quad (11)$$

The measurement noise v_i and the weight error vector $\tilde{\mathbf{w}}_i$ are assumed to be uncorrelated [29]–[31]. From this assumption, (11) is represented as

$$\begin{aligned} p_{i+1} &= p_i - 2\mu_1 \bar{E} \left(\frac{\tilde{\mathbf{w}}_i^T \mathbf{u}_i \mathbf{u}_i^T \tilde{\mathbf{w}}_i}{|e_i| \sqrt{\mathbf{u}_i^T \mathbf{u}_i}} \right) + \mu_1^2 \\ &= p_i - 2\mu_1 \sqrt{\mathbf{u}_i^T \mathbf{u}_i} \bar{E} \left(\frac{\tilde{\mathbf{w}}_i^T \mathbf{u}_i \mathbf{u}_i^T \tilde{\mathbf{w}}_i}{\mathbf{u}_i^T \mathbf{u}_i |e_i|} \right) + \mu_1^2. \end{aligned} \quad (12)$$

By ensemble averaging, MSD estimation (12) can be approximated as

$$\begin{aligned} p_{i+1} &= p_i - 2\mu_1 \sqrt{\mathbf{u}_i^T \mathbf{u}_i} \bar{E} \left(\frac{\tilde{\mathbf{w}}_i^T \tilde{\mathbf{w}}_i}{q_1 M |e_i|} \right) + \mu_1^2 \\ &= p_i - \frac{2\mu_1}{q_1 M} \bar{E} \left(\frac{\tilde{\mathbf{w}}_i^T \tilde{\mathbf{w}}_i}{|e_i| \sqrt{\mathbf{u}_i^T \mathbf{u}_i}} \right) + \mu_1^2, \end{aligned} \quad (13)$$

where q_1 is the scaling factor which reflects the correlatedness of the input vectors. To get the recursion form of MSD estimation, the approximation used in [32]–[35] is applied to (13). Then, the MSD estimation p_{i+1} is approximated as

$$\begin{aligned} p_{i+1} &\approx p_i - \frac{2\mu_1}{q_1 M} \frac{\bar{E}(\tilde{\mathbf{w}}_i^T \tilde{\mathbf{w}}_i)}{\bar{E}(|e_i|/\sqrt{\mathbf{u}_i^T \mathbf{u}_i})} + \mu_1^2 \\ &= p_i - \frac{2\mu_1}{q_1 M} \frac{p_i}{\bar{E}(|e_i|/\sqrt{\mathbf{u}_i^T \mathbf{u}_i})} + \mu_1^2 \\ &= \left(1 - \frac{2}{q_1 M} \frac{\mu_1}{\bar{E}(|e_i|/\sqrt{\mathbf{u}_i^T \mathbf{u}_i})} \right) p_i + \mu_1^2 \end{aligned} \quad (14)$$

Owing to the expectation, the estimation (14) cannot be used in the practical situation. Accordingly, the following moving average is applied to get the expectation term.

$$\beta_{i+1} = \lambda_1 \beta_i + (1 - \lambda_1) \frac{|e_i|}{\sqrt{\mathbf{u}_i^T \mathbf{u}_i}}, \quad (15)$$

where $0 \leq \lambda_1 < 1$ is a smoothing factor. Therefore, the update equation of MSD estimation of NS algorithm is finally derived as

$$p_{i+1} = \left(1 - \frac{2\mu_1}{q_1 M \beta_{i+1}} \right) p_i + \mu_1^2. \quad (16)$$

B. NOVEL MSD ANALYSIS OF NLMS ALGORITHM

The weight error vector of the NLMS algorithm at the next iteration is derived as

$$\tilde{\mathbf{w}}_{i+1} = \tilde{\mathbf{w}}_i - \mu_2 \frac{\mathbf{u}_i e_i}{\mathbf{u}_i^T \mathbf{u}_i} \quad (17)$$

From [13], the upper bound of MSD of the NLMS algorithm is analyzed as

$$p_{i+1} \leq \left(1 - \frac{2\mu_2 - \mu_2^2}{\rho M}\right) p_i + \frac{\mu_2^2 \sigma_\theta^2}{\mathbf{u}_i^T \mathbf{u}_i} \quad (18)$$

where $\rho \leq 1$ and it is an user-defined value which is determined by using the characteristics of the input signals. ρ is set to one when the inputs are white Gaussian process, but ρ is set to a value greater than one when the inputs are correlated signal. However, this estimation does not work well in the impulsive noise environments. To reflect abrupt changes of the error in the impulsive noise environments, the MSD estimation analyzed in the diffusion NLMS algorithm [36] is modified to be used in the NLMS algorithm as follows:

$$\begin{aligned} p_{i+1} &= \bar{E}(\tilde{\mathbf{w}}_{i+1}^T \tilde{\mathbf{w}}_{i+1}) \\ &= p_i - 2\mu_2 \bar{E} \left(\frac{\tilde{\mathbf{w}}_i^T \mathbf{u}_i e_i}{\mathbf{u}_i^T \mathbf{u}_i} \right) + \mu_2^2 \bar{E} \left(\frac{e_i^2}{\mathbf{u}_i^T \mathbf{u}_i} \right) \\ &= p_i - 2\mu_2 \bar{E} \left(\frac{\tilde{\mathbf{w}}_i^T \mathbf{u}_i (\mathbf{u}_i^T \tilde{\mathbf{w}}_i + v_i)}{\mathbf{u}_i^T \mathbf{u}_i} \right) + \mu_2^2 \bar{E} \left(\frac{e_i^2}{\mathbf{u}_i^T \mathbf{u}_i} \right) \\ &\approx p_i - 2\mu_2 \bar{E} \left(\frac{\tilde{\mathbf{w}}_i^T \mathbf{u}_i \mathbf{u}_i^T \tilde{\mathbf{w}}_i}{\mathbf{u}_i^T \mathbf{u}_i} \right) + \mu_2^2 \bar{E} \left(\frac{e_i^2}{\mathbf{u}_i^T \mathbf{u}_i} \right), \end{aligned} \quad (19)$$

where the measurement noise v_i and the weight error vector $\tilde{\mathbf{w}}_i$ are assumed to be uncorrelated [29]–[31]. Using the approximation used in (13), (19) is derived as

$$\begin{aligned} p_{i+1} &= p_i - 2\mu_2 \frac{\bar{E}(\tilde{\mathbf{w}}_i^T \tilde{\mathbf{w}}_i)}{q_2 M} + \mu_2^2 \bar{E} \left(\frac{e_i^2}{\mathbf{u}_i^T \mathbf{u}_i} \right) \\ &= p_i - \frac{2\mu_2}{q_2 M} p_i + \mu_2^2 \bar{E} \left(\frac{e_i^2}{\mathbf{u}_i^T \mathbf{u}_i} \right) \\ &= (1 - \frac{2\mu_2}{q_2 M}) p_i + \mu_2^2 \bar{E} \left(\frac{e_i^2}{\mathbf{u}_i^T \mathbf{u}_i} \right). \end{aligned} \quad (20)$$

In the practical situation, the estimation (20) cannot be used directly owing to the expectation. Unlike the NS algorithm, the instantaneous values are used to reflect a sudden surge of the error caused by the impulsive noise in the NLMS algorithm. Therefore, the update equation of MSD estimation of the NLMS algorithm is derived as

$$p_{i+1} = (1 - \frac{2\mu_2}{q_2 M}) p_i + \mu_2^2 \frac{e_i^2}{\mathbf{u}_i^T \mathbf{u}_i}. \quad (21)$$

C. MODE SWITCHING SCHEME

In the mode switching scheme, the mode with lower MSD estimation value is selected. The MSD estimations of two modes analyzed in (16), (21) are used to compare the MSDs

TABLE 1. The proposed mixed adaptive filter algorithm.

Initialization: $p_{1,0} = p_{2,0} = p_0 = 1, \hat{\mathbf{w}}_0 = \mathbf{0}$
 $\beta_0 = |d_0|$

Parameters setting: $q_1, q_2, \mu_1, \mu_2, \lambda_1 = \lambda_2 = 0.99$

For each iterations i ,

$e_i = d_i - \mathbf{u}_i^T \hat{\mathbf{w}}_i$

$\beta_{i+1} = \lambda_1 \beta_i + (1 - \lambda_1) \frac{|e_i|}{\|\mathbf{u}_i\|}$

$p_{1,i+1} = (1 - \frac{2\mu_1}{q_1 M \beta_{i+1}}) p_i + \mu_1^2$,

$p_{2,i+1} = (1 - \frac{2\mu_2}{q_2 M}) p_i + \mu_2^2 \frac{e_i^2}{\mathbf{u}_i^T \mathbf{u}_i}$.

$\hat{\mathbf{w}}_{i+1} = \begin{cases} \hat{\mathbf{w}}_i + \frac{\mu_1 \mathbf{u}_i \text{sgn}(e_i)}{\sqrt{\mathbf{u}_i^T \mathbf{u}_i}}, & \text{if } p_{1,i+1} \leq p_{2,i+1} \\ \hat{\mathbf{w}}_i + \mu_2 \frac{\mathbf{u}_i e_i}{\mathbf{u}_i^T \mathbf{u}_i}, & \text{if } p_{1,i+1} > p_{2,i+1} \end{cases}$

$p_{i+1} = \begin{cases} p_{1,i+1}, & \text{if } p_{1,i+1} \leq p_{2,i+1} \\ p_{2,i+1}, & \text{if } p_{1,i+1} > p_{2,i+1} \end{cases}$

end

of each mode. The proposed switching scheme is described as

$$\hat{\mathbf{w}}_{i+1} = \begin{cases} \hat{\mathbf{w}}_i + \frac{\mu_1 \mathbf{u}_i \text{sgn}(e_i)}{\sqrt{\mathbf{u}_i^T \mathbf{u}_i}}, & \text{if } p_{1,i+1} \leq p_{2,i+1} \\ \hat{\mathbf{w}}_i + \mu_2 \frac{\mathbf{u}_i e_i}{\mathbf{u}_i^T \mathbf{u}_i}, & \text{if } p_{1,i+1} > p_{2,i+1}, \end{cases} \quad (22)$$

where $p_{1,i+1}$ and $p_{2,i+1}$ are the MSD estimation values of the NS and the NLMS algorithm. Here, the MSD estimation values $p_{1,i+1}$ and $p_{2,i+1}$ are calculated at each iteration as follows

$$p_{1,i+1} = (1 - \frac{2\mu_1}{q_1 M \beta_{i+1}}) p_i + \mu_1^2, \quad (23)$$

$$p_{2,i+1} = (1 - \frac{2\mu_2}{q_2 M}) p_i + \mu_2^2 \frac{e_i^2}{\mathbf{u}_i^T \mathbf{u}_i}. \quad (24)$$

From the switching scheme (22), the mode with lower MSD estimation value is selected. The MSD estimation value at the next iteration is decided as

$$p_{i+1} = \begin{cases} p_{1,i+1}, & \text{if } p_{1,i+1} \leq p_{2,i+1} \\ p_{2,i+1}, & \text{if } p_{1,i+1} > p_{2,i+1} \end{cases} \quad (25)$$

The proposed $\mathcal{L}_1/\mathcal{L}_2$ -mode switching adaptive filter is summarized in Table 1.

D. RESET ALGORITHM FOR SYSTEM ABRUPT CHANGE

From the mode switching scheme (22), the proposed algorithm can minimize the bad effects caused by the

TABLE 2. Computational Complexity of Various Algorithms.

Algorithms	Multiplication	Algorithms	Multiplication
NLMS	$3M + 1$	VSS-CMPN	$2M + 4$
NS	$3M + 1$	GVSS-CMPN	$2M + 10$
LAD	$2M + 1$	Proposed	$3M + 10$
RMN	$2M + 2N_w + 4$	-	-

impulsive noises and converge fast to the steady-state. However, the algorithm loses its tracking capacity if the unknown system to be estimated has changed. To overcome this drawback in a non-stationary environments, a reset algorithm has to be used. The reset algorithms used in [19] cannot be directly applied to the proposed algorithm. Therefore, the modified reset algorithm is proposed as follows:

$$\begin{aligned}
 & \text{if } \text{mod}(i, V_T) = 0 \\
 & \quad C_{\text{new}} = \frac{\mathbf{S}^T \mathbf{D} \mathbf{S}}{V_T - V_D} \\
 & \quad \text{end} \\
 & \quad \text{if } (C_{\text{new}} - C_{\text{old}})/\mu_m > \xi \\
 & \quad \quad p_{i+1} = 1 \\
 & \quad \quad \beta_{i+1} = |d_i| \\
 & \quad \quad \text{end} \\
 & \quad C_{\text{old}} = C_{\text{new}}
 \end{aligned}$$

where V_T, V_D are the positive integers ($V_D < V_T$), $\text{mod}(a, b)$ stands for the remainder of the division between a and b , and $\mathbf{S} = \text{sort}(|e(n)|/(||\mathbf{u}(n)||_2 + \epsilon), \dots, |e(n - V_T + 1)|/(||\mathbf{u}(n - V_T + 1)||_2 + \epsilon))^T$ where $\text{sort}(\cdot)$ is the ascending order operator, $\mathbf{D} = \text{diag}(1, \dots, 1, 0, \dots, 0)$ is a diagonal matrix with its first $V_T - V_D$ elements set to one, $\mu_m = (\mu_1 + \mu_2)/2$, and ξ is a threshold value.

E. COMPUTATIONAL COMPLEXITY

Unlike the conventional algorithms, the noise variance estimation process is not necessary because the proposed mixed $\mathcal{L}_1/\mathcal{L}_2$ -norm adaptive filter does not use the noise variance σ_θ^2 . Considering these points, the proposed algorithm has merit in terms of computational complexity. Table 2 shows the computational complexity of various algorithms. The computational complexity of the proposed algorithm is a little higher than those of the NLMS and NS algorithm. However, the performance of the proposed algorithm is much better than that of the other algorithms.

IV. SIMULATION RESULTS

To verify the performance of the proposed algorithm, the simulations are conducted in the system identification models. The channel of the unknown system is generated by the white Gaussian random process with zero mean and unit variance, and it has 64 taps ($M = 64$). It is assumed that the tap

length of the adaptive filter is the same as that of the unknown system. The signal-to-noise ratio (SNR) between filter output $y_i = u_i^T w_o$ and the measurement noise is defined as

$$SNR = 10 \log_{10} \left(\frac{E[y_i^2]}{E[\theta_i^2]} \right). \quad (26)$$

In the following simulations, the SNR is set to 30dB. The white input signal is obtained by a white Gaussian random process with zero mean and unit variance. The correlated input signals are obtained by filtering through these systems:

$$G_1(z) = \frac{1}{1 - 0.7z^{-1}} \quad (27)$$

$$G_2(z) = \frac{1 + 0.6z^{-1}}{1 + 1.0z^{-1} + 0.21z^{-2}} \quad (28)$$

The impulsive noise η_i is modeled by a Bernoulli-Gaussian process as follows:

$$\eta_i = s_i b_i \quad (29)$$

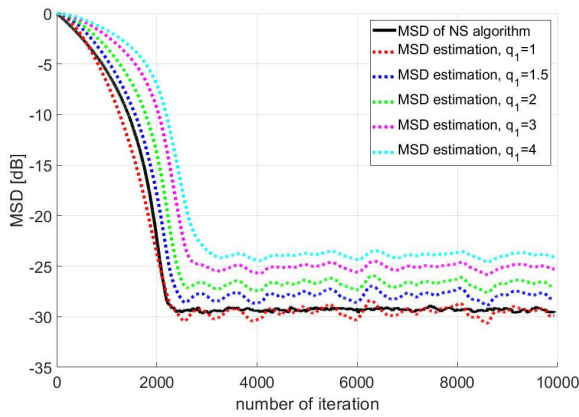
where b_i is a white Gaussian process and s_i is a Bernoulli process with the probabilities $P[s_i = 1] = pr$ and $P[s_i = 0] = 1 - pr$. Here, pr is the impulsive noise occurrence probability that is set to 0.001 and 0.01 in our simulations. The signal-to-impulsive noise ratio (SIR) between the filter output and the impulsive noise is defined as

$$SIR = 10 \log_{10} \left(\frac{E[y_i^2]}{E[b_i^2]} \right). \quad (30)$$

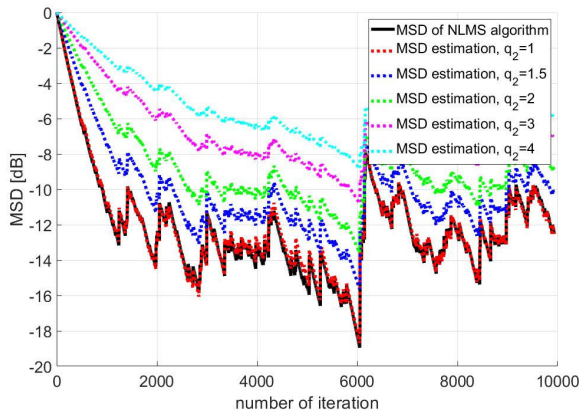
In the following simulations, the SIR is set to -30dB and -20dB .

A. MSD ANALYSES COMPARISON

Fig. 2 shows the curves of MSD estimations of the NS and the NLMS algorithms for different q_1 and q_2 . The smoothing factors are set to a value close to one, where $\lambda_1 = \lambda_2 = 0.99$ which are used in the following simulations. The step sizes are chosen so that the real MSD converges to an appropriate steady-state misalignment value, where $\mu_1 = 0.005$ and $\mu_2 = 0.1$. We can see that the larger q_1 and q_2 , the larger the MSD estimation values. In the proposed algorithm, the MSD estimation values should be slightly larger than the real MSD values because the worst cases have to be considered in the comparison step of the mode-switching scheme. However, the MSD estimations are regarded as faulty analyses if they are largely different from the real MSDs. Therefore, the optimal values that satisfy these conditions have to be set as q_1 and q_2 . Based on the results of Fig. 2, q_1 and q_2 are set as slightly a larger value than 1 that satisfies the above conditions when the input is white Gaussian process, which $q_1 = 1.1$ and $q_2 = 1.3$. On the other hand, q_1 and q_2 are set as a moderately large value that satisfies the above conditions when the input is correlated, where $q_1 = 3$ and $q_2 = 3.5$. The proposed algorithm has similar performance as long as the q_1 and q_2 are set as an appropriately larger value than the previously chosen values.



(a)



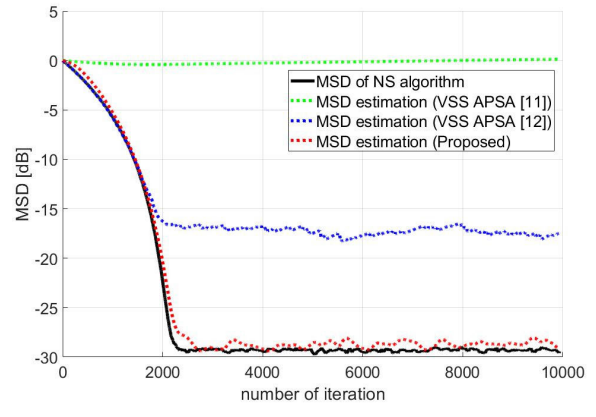
(b)

FIGURE 2. Learning curves of real MSD and MSD estimations in impulsive noise environments (a) MSD estimations of NS algorithm for different q_1 values (b) MSD estimations of NLMS algorithm for different q_2 values.

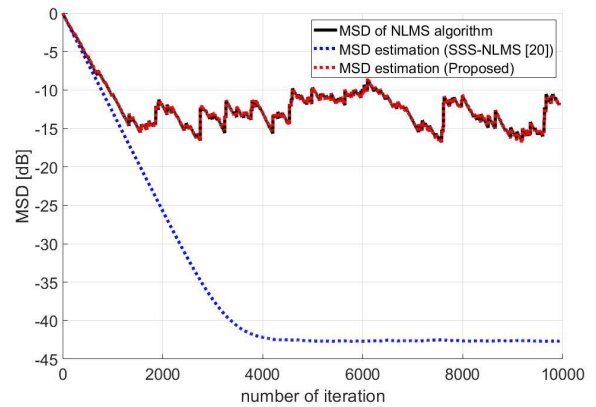
Fig. 3 shows the curves of MSD estimations of the NS and the NLMS algorithms when the input is white Gaussian process. The step sizes are set to same values used in Fig. 2. The MSD estimations of the NS algorithm in the existing studies [18], [19] do not track the curve of the real MSD value but the proposed estimation properly tracks the curve of the real one (Fig. 3(a)). The MSD estimation of the NLMS algorithm in the existing study (18) also does not track the curve of the real MSD but the proposed estimation properly tracks the curve of the real one in the impulsive noise environments (Fig. 3(b)).

B. SYSTEM IDENTIFICATION WITH STEADY ENVIRONMENTS

In Figs 4-6, the simulations are conducted in the steady environments where the system parameters and impulsive noise ratio are not changed. The step sizes are chosen so that their MSDs at the steady-state have similar values. In Fig. 4, the step sizes are set to $\mu_1 = 0.005$ and $\mu_2 = 0.1$. Firstly, Fig. 4(a) shows the MSD curves for the NS, the NLMS and the proposed algorithm in non-impulsive noise environments. Since the proposed algorithm compares the updated MSD estimations and adopts the better mode



(a)

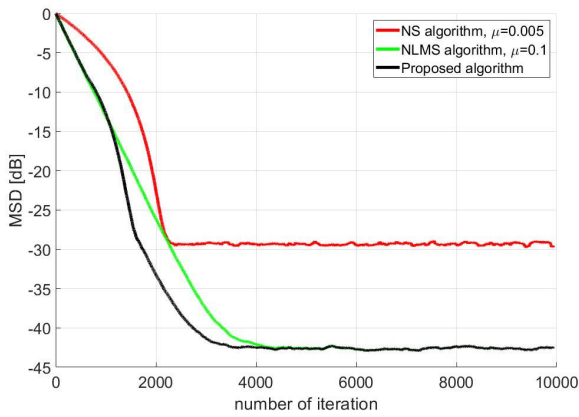


(b)

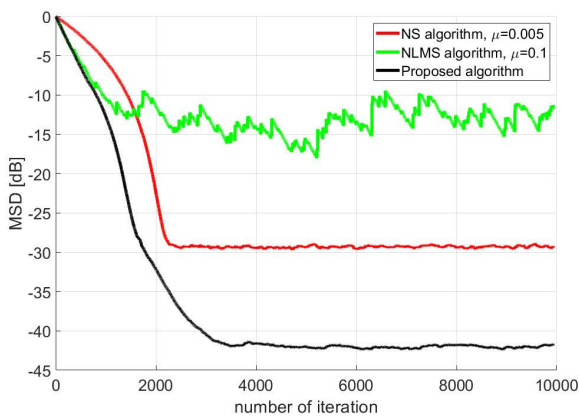
FIGURE 3. Comparisons of various MSD estimations (a) About NS algorithm (b) About NLMS algorithm.

at each iteration, it has a faster convergence rate than the NS and the NLMS algorithms and has a low steady-state error of the NLMS algorithm. Fig. 4(b) shows the MSD curves for the NS, the NLMS algorithms and the proposed algorithm in the impulsive noise environments. The proposed algorithm has the robustness to the impulsive noises of the NS algorithm and low steady-state misalignment of the NLMS algorithm. It operates in \mathcal{L}_1 mode when the impulsive noises occur. Therefrom, the weight coefficient of the adaptive filter does not diverge to the abnormal one and stably converges to the steady-state misalignment value of the NLMS algorithm obtained in non-impulsive noise environments.

In Figs 5 and 6, the MSD curves of the proposed algorithm and other algorithms are shown for comparison. The proposed algorithm is compared with NS, NLMS, LAD, RMN, VSS-CMPN and GVSS-CMPN. For the RMN algorithm, $N_w = 10$ as suggested in [20]. For the GVSS-CMPN, $\kappa = 0.1$ and $\theta = -2$ as described in [23]. In Fig. 5, $\mu_1 = 0.005$ and $\mu_2 = 0.1$ for the proposed algorithm and $\mu = 0.00015$ for the VSS-CMPN. In Fig. 6, $\mu_1 = 0.005$ and $\mu_2 = 0.2$ for the proposed algorithm and $\mu = 0.0003$ for the VSS-CMPN. In Figs 5(a) and 6(a), the white Gaussian noise is used as inputs. It is shown that the proposed algorithm has better



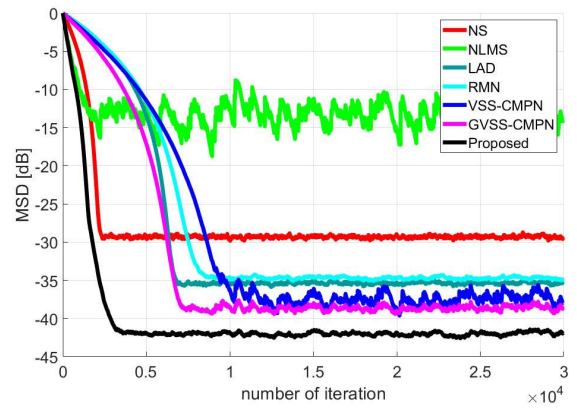
(a)



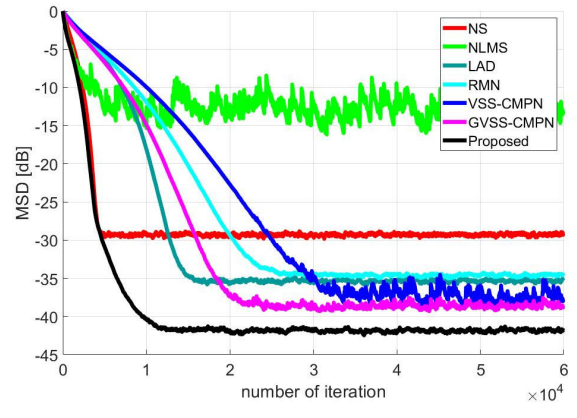
(b)

FIGURE 4. MSD curves of NS, NLMS and proposed algorithm for the white Gaussian noise input (a) In non-impulsive noise environments (b) In impulsive noise environments.

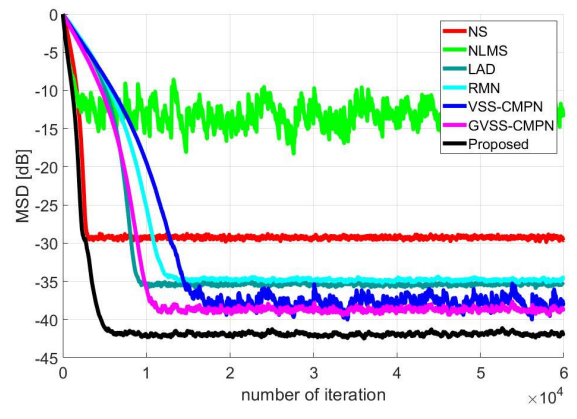
performance than the other conventional algorithms including the mixed-norm algorithms in terms of the convergence rate and the steady-state misalignment. In Figs 5(b) and 6(b), AR(1) process generated using $G_1(z)$ is used as inputs. In Figs 5(c) and 6(c), ARMA process generated using $G_2(z)$ are used as the inputs. It is also shown that the proposed algorithm has the great improvement compared to the conventional algorithms including the mixed-norm algorithms in terms of the convergence rate and the steady-state misalignment when the inputs are correlated. The mixed-norm adaptive filters [20], [22], [23] are derived from the convex combination of the L_1 norm and L_2 norm or the continuous mixed norm. The performance of these mixed-norm adaptive filters partially deteriorates as the L_2 norm-based term in them is badly affected by the large value of the impulsive noises. However, the proposed algorithm selects L_1 mode to prevent the wrong updates when the impulsive noises occur, otherwise, it selects L_2 mode to get a fast convergence rate. Therefore, we can say that the proposed algorithm has improved performance in both impulsive and non-impulsive noise environments and we can see that in simulation results.



(a)



(b)

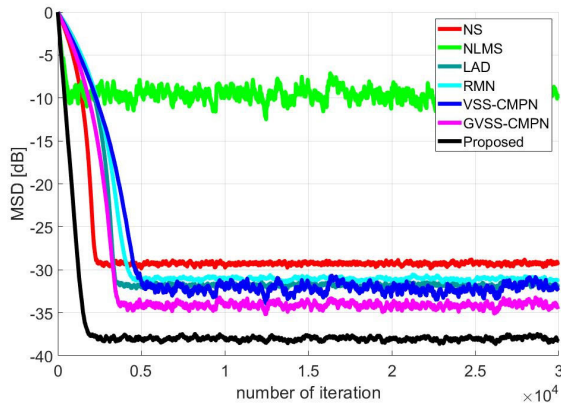


(c)

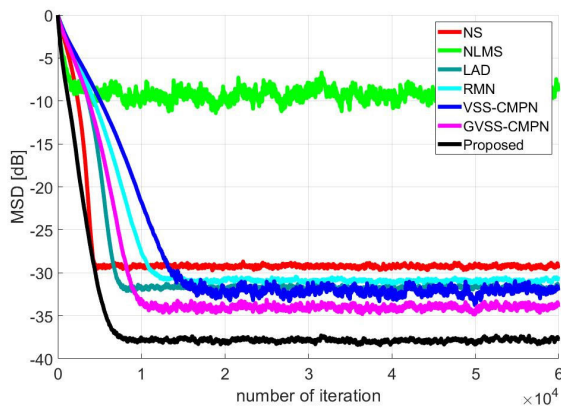
FIGURE 5. MSD curves of various algorithms in impulsive noise environments with $pr = 0.001$ and $SIR = -30\text{dB}$ (a) Inputs: white Gaussian noise (b) Inputs: correlated input ($G_1(z)$) (c) Inputs: correlated input ($G_2(z)$).

C. SYSTEM IDENTIFICATION WITH CHANGED ENVIRONMENTS

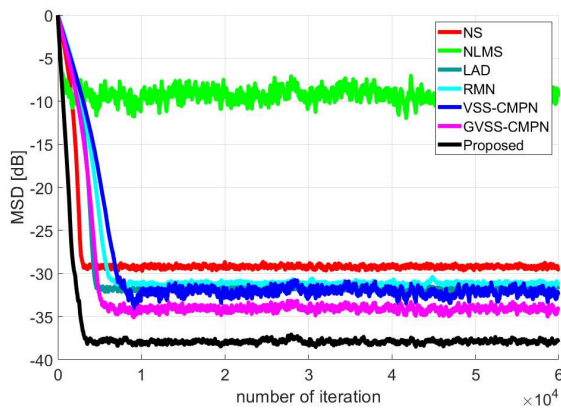
In Figs 7 and 8, the impulsive noises exist in the first and third quarter of total iterations and it does not exist in the second and fourth quarter. The tuning parameters in Fig. 7 are the same as in Fig. 5, and the tuning parameters in Fig. 8 are the same as in Fig. 6. In Figs 7(a) and 8(a), the input signals are white Gaussian process. We can see that the proposed algorithm has faster convergence rate and



(a)



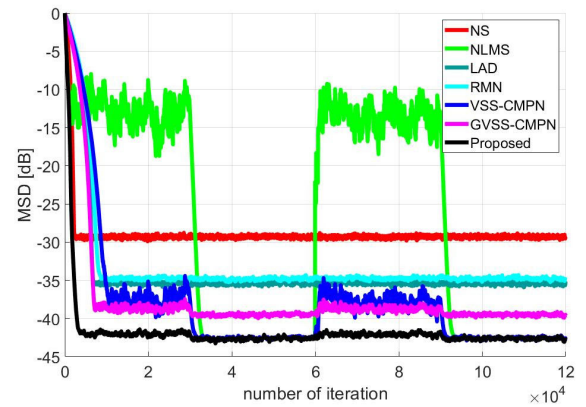
(b)



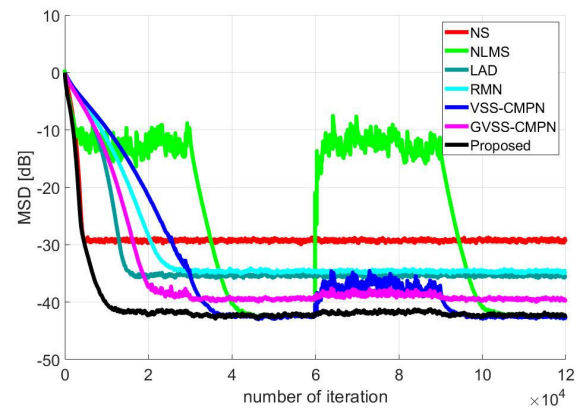
(c)

FIGURE 6. MSD curves of various algorithms in impulsive noise environments with $pr = 0.01$ and $SIR = -20\text{dB}$ (a) Inputs: white Gaussian noise (b) Inputs: correlated input ($G_1(z)$) (c) Inputs: correlated input ($G_2(z)$).

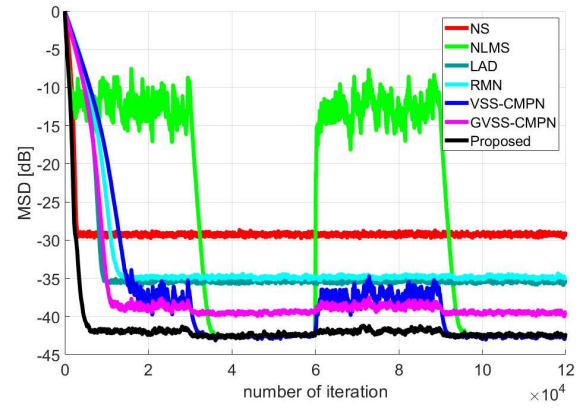
lower steady-state misalignment than the other algorithms including the mixed-norm algorithms. At the steady-state, the curves of the conventional algorithms are fluctuated due to the steady-state performance degradation in the impulsive noise environments, but that of the proposed algorithm remains at an optimal value. In Figs 7(b) and 8(b), AR(1) process generated using $G_1(z)$ are used as the inputs. In Figs 7(c) and 8(c), ARMA process generated using $G_2(z)$



(a)



(b)



(c)

FIGURE 7. MSD curves of various algorithms in on/off impulsive noise environments with $pr = 0.001$ and $SIR = -30\text{dB}$ (a) Inputs: white Gaussian noise (b) Inputs: correlated input ($G_1(z)$) (c) Inputs: correlated input ($G_2(z)$).

are used as the inputs. We can also see that the proposed algorithm has faster convergence rate and lower steady-state misalignment than the other algorithms including the mixed-norm algorithms when the inputs are correlated. At the steady-state, the performance degradation of the NLMS and VSS-CMPN algorithms are noticeable due to the slow convergence rates of them caused by correlated inputs. However, the steady-state misalignment of the proposed algorithm remains stable.

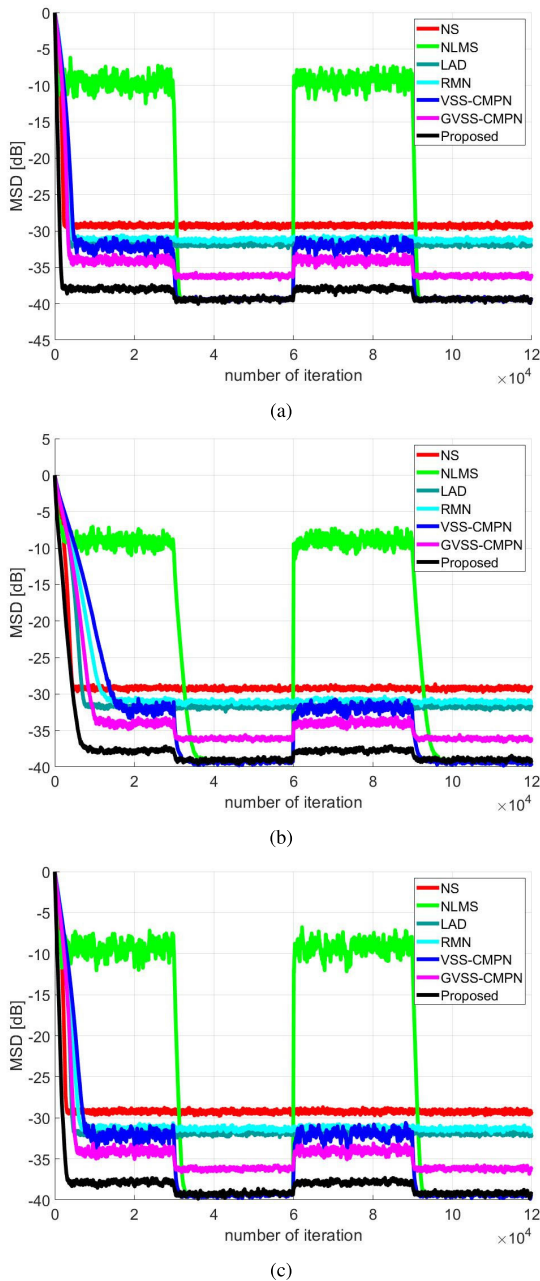


FIGURE 8. MSD curves of various algorithms in on/off impulsive noise environments with $pr = 0.01$ and $SIR = -20\text{dB}$ (a) Inputs: white Gaussian noise (b) Inputs: correlated input ($G_1(z)$) (c) Inputs: correlated input ($G_2(z)$).

In Fig. 9, the unknown system is abruptly changed to $-w_o$ at 60000th iteration. For the reset algorithm, $V_T = M/2$, $V_D = [0.2 \times V_T]$, $\epsilon = 10^{-6}$ and $\zeta = 0.001$, where $\lfloor \cdot \rfloor$ stands for the round-off function. The other tuning parameters used in Figs 9(a) and 9(b) are the same as Figs 7(b) and 8(b), respectively. After the unknown system changes, the convergence rate of the proposed algorithm is faster than those of the other algorithms and its steady-state misalignment is lower than those of the other algorithms. Therefore, the proposed algorithm is more robust to unknown system abrupt changes.

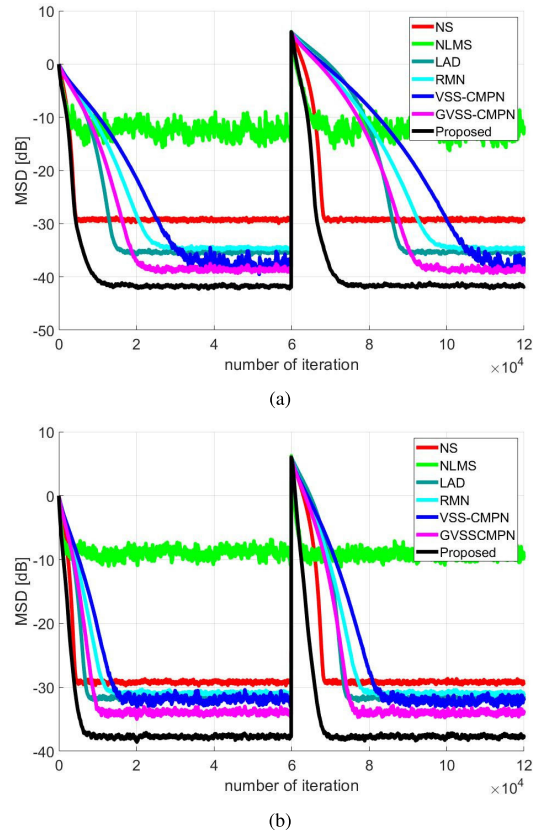


FIGURE 9. MSD curves of various algorithms for correlated input ($G_1(z)$) with system abrupt change (a) $pr = 0.001$ and $SIR = -30\text{dB}$ (b) $pr = 0.01$ and $SIR = -20\text{dB}$.

V. CONCLUSION

This paper proposed an $\mathcal{L}_1/\mathcal{L}_2$ -mode switching adaptive filter algorithm by adopting mode-switching scheme. The mode-switching scheme compared the MSD estimations of each mode and selected the better mode in the sense that the forthcoming MSD of the selected mode is lower than that of the other mode. To estimate MSDs precisely, the MSD estimations for the NS and the NLMS algorithms were analyzed in the novel methods and verified in the system identification simulation. The proposed algorithm was conducted in various environments and the simulation results showed that the proposed algorithm outperformed other conventional algorithms in terms of the convergence rate and steady-state misalignment.

REFERENCES

- [1] S. S. Haykin, *Adaptive Filter Theory*. Noida, India: Pearson Education, 2008.
- [2] A. H. Sayed, *Fundamentals of Adaptive Filtering*. Hoboken, NJ, USA: Wiley, 2003.
- [3] S. M. Kuo and D. Morgan, *Active Noise Control Systems: Algorithms and DSP Implementations*. Hoboken, NJ, USA: Wiley, 1995.
- [4] P. Gibbs and H. H. Asada, "Reducing motion artifact in wearable biosensors using mems accelerometers for active noise cancellation," in *Proc. Amer. Control Conf.*, 2005, pp. 1581–1586.
- [5] H.-C. Shin, A. H. Sayed, and W.-J. Song, "Variable step-size NLMS and affine projection algorithms," *IEEE Signal Process. Lett.*, vol. 11, no. 2, pp. 132–135, Feb. 2004.

- [6] K. Wagner and M. Doroslovacki, "Proportionate-type normalized least mean square algorithms with gain allocation motivated by mean-square-error minimization for white input," *IEEE Trans. Signal Process.*, vol. 59, no. 5, pp. 2410–2415, May 2011.
- [7] W. Gao and J. Chen, "Performance analysis of diffusion LMS for cyclostationary white non-Gaussian inputs," *IEEE Access*, vol. 7, pp. 91243–91252, 2019.
- [8] Z. Jin, L. Guo, and Y. Li, "The bias-compensated proportionate NLMS algorithm with sparse penalty constraint," *IEEE Access*, vol. 8, pp. 4954–4962, 2020.
- [9] T. Park, M. Lee, and P. Park, "Scheduled step-size subband adaptive filter algorithm with implemental consideration," *IEEE Access*, vol. 8, pp. 199025–199033, 2020.
- [10] P. Petrus, "Robust huber adaptive filter," *IEEE Trans. Signal Process.*, vol. 47, no. 4, pp. 1129–1133, Apr. 1999.
- [11] S.-C. Chan and Y.-X. Zou, "A recursive least M-estimate algorithm for robust adaptive filtering in impulsive noise: Fast algorithm and convergence performance analysis," *IEEE Trans. Signal Process.*, vol. 52, no. 4, pp. 975–991, Apr. 2004.
- [12] S. M. Jung and P. Park, "Normalised least-mean-square algorithm for adaptive filtering of impulsive measurement noises and noisy inputs," *Electron. Lett.*, vol. 49, no. 20, pp. 1270–1272, Sep. 2013.
- [13] P. Park, M. Chang, and N. Kong, "Scheduled-stepsize NLMS algorithm," *IEEE Signal Process. Lett.*, vol. 16, no. 12, pp. 1055–1058, Dec. 2009.
- [14] P. H. Prajapati and A. D. Darji, "FPGA implementation of MRMN with step-size scaler adaptive filter for impulsive noise reduction," *Circuits, Syst., Signal Process.*, vol. 39, pp. 3682–3710, Jan. 2020.
- [15] O. Arikan, A. Enis Cetin, and E. Erzincan, "Adaptive filtering for non-Gaussian stable processes," *IEEE Signal Process. Lett.*, vol. 1, no. 11, pp. 163–165, Nov. 1994.
- [16] C. Kwong, "Dual sign algorithm for adaptive filtering," *IEEE Trans. Commun.*, vol. 34, no. 12, pp. 1272–1275, Dec. 1986.
- [17] T. Shao, Y. R. Zheng, and J. Benesty, "An affine projection sign algorithm robust against impulsive interferences," *IEEE Signal Process. Lett.*, vol. 17, no. 4, pp. 327–330, Apr. 2010.
- [18] J. Shin, J. Yoo, and P. Park, "Variable step-size affine projection sign algorithm," *Electron. Lett.*, vol. 48, no. 9, pp. 483–485, 2012.
- [19] J. Yoo, J. Shin, and P. Park, "Variable step-size affine projection sign algorithm," *IEEE Trans. Circuits Syst. II, Exp. Briefs*, vol. 61, no. 4, pp. 274–278, Apr. 2014.
- [20] 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.
- [21] 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.
- [22] H. Zayyani, "Continuous mixed p -norm adaptive algorithm for system identification," *IEEE Signal Process. Lett.*, vol. 21, no. 9, pp. 1108–1110, Sep. 2014.
- [23] L. Shi, H. Zhao, and Y. Zakharov, "Generalized variable step size continuous mixed p -norm adaptive filtering algorithm," *IEEE Trans. Circuits Syst. II, Exp. Briefs*, vol. 66, no. 6, pp. 1078–1082, Jun. 2019.
- [24] Y.-R. Chien and W.-J. Tseng, "Switching-based variable step-size approach for partial update LMS algorithms," *Electron. Lett.*, vol. 49, no. 17, pp. 1081–1083, Aug. 2013.
- [25] N. J. Bershad and J. C. M. Bermudez, "A switched variable step size NLMS adaptive filter," *Digit. Signal Process.*, vol. 101, Jun. 2020, Art. no. 102730.
- [26] D. P. Mandic, "A generalized normalized gradient descent algorithm," *IEEE Signal Process. Lett.*, vol. 11, no. 2, pp. 115–118, Feb. 2004.
- [27] I. Song, P. Park, and R. W. Newcomb, "A normalized least mean squares algorithm with a step-size scaler against impulsive measurement noise," *IEEE Trans. Circuits Syst. II, Exp. Briefs*, vol. 60, no. 7, pp. 442–445, Jul. 2013.
- [28] P. Park, J.-H. Seo, and N. Kong, "Variable matrix-type step-size affine projection algorithm with orthogonalized input vectors," *Signal Process.*, vol. 98, pp. 135–142, May 2014.
- [29] S. M. Jung and P. Park, "Stabilization of a bias-compensated normalized Least-Mean-Square algorithm for noisy inputs," *IEEE Trans. Signal Process.*, vol. 65, no. 11, pp. 2949–2961, Jun. 2017.
- [30] L. Lu and H. Zhao, "Adaptive volterra filter with continuous l_p -norm using a logarithmic cost for nonlinear active noise control," *J. Sound Vib.*, vol. 364, pp. 14–29, Mar. 2016.
- [31] Z. Zheng, Z. Liu, and H. Zhao, "Bias-compensated normalized least-mean fourth algorithm for noisy input," *Circuits, Syst., Signal Process.*, vol. 36, no. 9, pp. 3864–3873, Sep. 2017.
- [32] S. J. M. de Almeida, J. C. M. Bermudez, N. J. Bershad, and M. H. Costa, "A statistical analysis of the affine projection algorithm for unity step size and autoregressive inputs," *IEEE Trans. Circuits Syst. I, Reg. Papers*, vol. 52, no. 7, pp. 1394–1405, Jul. 2005.
- [33] N. J. Bershad, J. C. M. Bermudez, and J.-Y. Tournet, "An affine combination of two LMS adaptive filters' transient mean-square analysis," *IEEE Trans. Signal Process.*, vol. 56, no. 5, pp. 1853–1864, Apr. 2008.
- [34] K. D. S. Olinto, D. B. Haddad, and M. R. Petraglia, "Transient analysis of 10-LMS and 10-NLMS algorithms," *Signal Process.*, vol. 127, pp. 217–226, Oct. 2016.
- [35] Y. Yu and H. Zhao, "Novel sign subband adaptive filter algorithms with individual weighting factors," *Signal Process.*, vol. 122, pp. 14–23, May 2016.
- [36] S. M. Jung, P. Park, and J.-H. Seo, "Efficient variable step-size diffusion normalised least-mean-square algorithm," *Electron. Lett.*, vol. 51, no. 5, pp. 395–397, Mar. 2015.



MINHO LEE (Graduate Student Member, IEEE) received the B.S. degree in electronics engineering from Kyungpook National University, Daegu, South Korea, in 2015, and the M.S. degree in electrical engineering from the Pohang University of Science and Technology, Pohang, South Korea, in 2018, where he is currently pursuing the Ph.D. degree. His research interests include signal processing, artificial intelligence, and PID control.



TAESU PARK (Student Member, IEEE) received the B.S. degree in electronics engineering from Kyungpook National University, Daegu, South Korea, in 2015, and the M.S. degree in electrical engineering from the Pohang University of Science and Technology, Pohang, South Korea, in 2017, where he is currently pursuing the Ph.D. degree. His research interests include signal processing and reinforcement learning in artificial intelligence.



JUNWOONG HUR received the B.S. degree in electrical engineering, the M.S. degree in electrical engineering, and the Ph.D. degree from the Pohang University of Science and Technology, Pohang, South Korea, in 2012, 2014, and 2019, respectively. His research interests include signal processing and artificial intelligence.



POOGYEON PARK (Senior Member, IEEE) received the B.S. and M.S. degrees in control and instrumentation engineering from Seoul National University, Seoul, South Korea, in 1988 and 1990, respectively, and the Ph.D. degree from the Stanford University, Stanford, CA, USA, in 1995. From 1996 to 2000, he was an Assistant Professor with the Pohang University of Science and Technology. Since 2006, he has been a Professor with the Department of Electronic and Electrical Engineering, Pohang University of Science and Technology. He has authored over 170 articles and the total citation for his articles is 10400. His research interest includes control and signal processing.