A novel normalized sign algorithm for system identification under impulsive noise interference

Lu Lu and Haiquan Zhao

Abstract To overcome the performance degradation of adaptive filtering algorithm in presence of impulsive noise, a novel normalized sign algorithm (NSA) based on the convex combination strategy is proposed in this paper, which is an adaptive combination of two NSA filters with the different step-size, called NSA-NSA. The proposed approach is capable of solving the conflicting requirement of fast convergence rate and low steady-state error for the single NSA filter. To further improve the robustness against impulsive noise, a mixing parameter updating formula based on a sign cost function is derived. Moreover, a tracking weight transfer scheme of coefficients from a fast NSA filter to a slow NSA filter is proposed for speed up convergence rate. Finally, the proposed algorithm is verified by theoretical analysis and simulation studies.

Keywords: Adaptive filtering, Convex combination, Normalized sign algorithm, System identification, Impulsive noise.

1 Introduction

The performance of adaptive filtering algorithms would degrade when the scenario was contaminated with the impulsive noise. To overcome the limitation, several algorithms based on clipping of the estimation error were proposed [1-3]. Particularly, the sign algorithm (SA) was successfully applied to identification when the system corrupted by impulsive noise. However, its convergence rate was still slow [1]. As the normalized least mean square (NLMS) algorithm derived from LMS algorithm, the normalized versions of SA can be easily developed the normalized SA (NSA) [4]. Moreover, several variants were proposed with an aim of improving the original convergence characteristics [5-13]. Particularly, in [11], a dual SA (DSA) with the variable step-size (VSS) scheme was presented, but it has a local divergence problem when large disparity between the step sizes. In [12], an interesting trial was attempted to obtain a better stability and convergence property by inserting a third step-size. It should be noted that the above-mentioned efforts have all been made for single adaptive filter.

On the other hand, to improve the filtering performance under the impulsive noise, some stochastic algorithms based on mixed error norms were introduced [14-18]. In [17], a robust mixed-norm (RMN) algorithm was introduced whose cost function combined the different error norms based on robust statistics. Later, a novel VSS RMN (NRMN) algorithm by using time-varying learning rate was proposed by Papoulis et al. [18-19]. It circumvents the disadvantage of slow convergence for RMN to some extent.

The convex combination algorithm was another effective way to balance the tradeoff between the convergence rate and steady-state error. In 2006, a combination least mean square (LMS)
(CLMS) algorithm was developed [20], which utilized two LMS filters with different step-size to obtain fast convergence and low misadjustment. Nevertheless, when the scenario was contaminated with the impulsive noise, algorithms in [21] and [22] may fail to work. To achieve the improved performance, a NLMS-NSA algorithm was developed where a combination scheme was used to switch between the NLMS and NSA algorithms [23]. Regrettably, in initial convergence stage, the high misadjustment was achieved by the NLMS algorithm under a relatively high background noise environment. Moreover, the mixing parameter of NLMS-NSA was unsuitable for impulsive noise so that the algorithm failed to achieve the best result.

In this work, to solve the above-mentioned problem, a NSA-NSA algorithm is proposed which offers a robust performance by adaptive combining of two independent NSA filters with large and small step-size. To further enhance the robustness against impulsive noise, the mixing parameter is adjusted by a sign cost function. And, a tracking weight transfer of coefficients is proposed to obtain fast convergence during the transition period. In particular, our main contributions are listed as follows: (1) to develop a NSA-NSA algorithm that is well suited for the system identification under impulsive noise, (2) to modify the update manner of mixing parameter. And, the analyzed of this manner is conducted in this paper, (3) to propose a novel weight transfer scheme that is low computational complexity and effectiveness for increase convergence rate.

The paper is organized in the following manner. In Section 2, the proposed NSA-NSA is described with explanation and the novel weight transfer scheme made for the proposed algorithm. In Section 3, simulation results in different impulsive noise environments are provided. Finally, in Section 4, concluding remarks are provided.

2 Adaptive combination of NSA algorithms

2.1. The proposed algorithm

![Fig. 1. Diagram of the proposed algorithm.](image)

The diagram of the adaptive combination scheme of two NSA filters is illustrated in Fig. 1, where \( x(n) \) and \( y(n) \) are the filter input and output signals, \( d(n) \) is the desired signal, \( y_1(n) \) and \( y_2(n) \) are symbols of two component filters, \( v(n) \) is the impulsive noise, \( w_0 \) is the weight vector of unknown system, respectively. The overall error of the filter is given by
\( e(n) = d(n) - y(n) \). To achieve a better overall performance, these filters are combined with a scalar mixing parameter \( \lambda(n) \):

\[
y(n) = \lambda(n)y_1(n) + [1 - \lambda(n)]y_2(n) \tag{1}
\]
\[
e(n) = \lambda(n)e_1(n) + [1 - \lambda(n)]e_2(n) \tag{2}
\]

where \( \lambda(n) \in [0,1] \) is the mixing parameter, which is defined by using a sigmoidal activation function with the auxiliary parameter \( a(n) \)

\[
\lambda(n) = \frac{1}{1 + e^{-a(n)}}. \tag{3}
\]

A gradient descent adaptation of \( a(n) \) is given as

\[
a(n + 1) = a(n) - \frac{\nu_a \cdot \hat{e} \cdot \hat{e}^T(n)}{\sigma(a(n))}. \tag{4}
\]

where \( \nu_a \) is the step-size of the auxiliary parameter \( a(n) \). This adaptation approach is derived by the cost function \( J(n) = e(n)^2 \) [20]. To adapt the normalized sign algorithm and improve the robustness to impulsive noise, the new cost function is defined as \( J_s(n) = |e(n)| \) based on the concept of the classical sign-error LMS algorithm [1]. Therefore, the adaptation of \( a(n) \) is derived by minimizing the cost function \( J_s(n) \) as follows:

\[
a(n + 1) = a(n) - \frac{\mu_a \cdot \hat{J}_s(n)}{\sigma(a(n))} \tag{5}
\]

where \( \mu_a \) is the step-size.

Using the chain rule, the gradient of the cost function \( J_s(n) \) can be calculated as follows:

\[
a(n + 1) = a(n) - \frac{\mu_a \cdot \hat{J}_s(n)}{\sigma(a(n))} \frac{\partial \lambda(n)}{\partial a(n)} \tag{6}
\]

\[
= a(n) + \rho_s \cdot \text{sign}(e(n)) \cdot [y_1(n) - y_2(n)] \lambda(n) [1 - \lambda(n)]
\]

where \( \rho_s \) is the positive constant, and \( \text{sign}(\cdot) \) is expressed by

\[
\text{sign}(x) = \frac{x}{\|x\|_2} = \begin{cases} 
1, & \text{if } x > 0 \\
0, & \text{if } x = 0 \\
-1, & \text{if } x < 0
\end{cases}
\tag{7}
\]

At each iteration cycle, the weight adaptation of NSA-NSA takes the form [1]

\[
\mathbf{w}_i(n + 1) = \mathbf{w}_i(n) + \mu_i \cdot \mathbf{x}(n) \text{sign}[e_i(n)] \cdot e_i + \| \mathbf{x}(n) \|_2^2 \tag{i = 1, 2}
\tag{8}
\]

where \( \mathbf{w}_i(n) \) being the weight vectors with length \( M \), \( \mu_i \) is the constant step-size, \( e_i > 0 \) is a regularization constant close to zero, and \( \| \cdot \|_2 \) represents Euclidian 2-norm. As a result, the combined filter are obtained using the following convex combination scheme

\[
\mathbf{w}(n) = \lambda(n)\mathbf{w}_1(n) + [1 - \lambda(n)]\mathbf{w}_2(n) \tag{9}
\]
Table 1. Summary of the proposed algorithm

| Initialize $N_0, a^*, \rho_1, \mu_2$ |
|-----------------|-----------------|
| Loop $n=1$ to end do |
| $y_i(n) = w_i^n(n)x(n)$ \ $(i = 1, 2)$ |
| $e_i(n) = d(n) - y_i(n)$ |
| $y(n) = \lambda(n)y_1(n) + (1 - \lambda(n))y_2(n)$ |
| $w_1(n+1) = w_1(n) + \rho_1 \frac{\text{sign}[e_1(n)]x(n)}{\varepsilon_e + \|x(n)\|^2}$ |
| $w_2(n+1) = w_2(n) + \rho_2 \frac{\text{sign}[e_2(n)]x(n)}{\varepsilon_e + \|x(n)\|^2}$ |
| $a(n+1) = a(n) + \rho_1 \text{sign}[e_1(n)]$ \ \ \ $[(y_1(n) - y_2(n))\lambda(n)[1 - \lambda(n)]$ |
| $\lambda(n+1) = 1/((1 + e^{-\alpha})$ |
| % Tracking weight transfer scheme (the proposed method) |
| if (mod($n-1, N_0$) equal to zero) |
| if $a(n+1) < -a^*$ |
| $a(n+1) = -a^*$ |
| $\lambda(n+1) = 0$ |
| endif |
| if $a(n+1) \geq a^*$ |
| $a(n+1) = a^*$ |
| $\lambda(n+1) = 1$ |
| $w_2(n+1) = w_2(n+1)$ |
| endif |
| Let $n=n+1$ |

2.2. Proposed weight transfer scheme

Inspired by the instantaneous transfer scheme from [24], a tracking weight transfer scheme is proposed, as shown in Table. 1. By using a sliding widow approach, the proposed scheme involves few parameters and remains the robustness against impulsive noise with low-cost. Like the instantaneous transfer scheme in [24], the parameter of tracking weight transfer strategy is not sensitive to the choice. This scheme can speed up convergence property of the overall filter, especially during the period of convergence transition. Define $N_0$ is a window length, then, if $n-1 \mod N_0$ equal to zero, implement following operations. It is well known that the standard convex combination scheme needs to check if $a(n+1) = a^*$, so the only additional operation is $n \mod N_0$ operation. The operation $a(n+1) \geq a^*$ denotes the stage that the fast filter (filter with large step-size) switch to slow filter (filter with small step-size) at transient stage. The mixing parameter $\lambda(n+1) = 0$ and $\lambda(n+1) = 1$ is the limitations for $a(n+1) < -a^*$ and $a(n+1) \geq a^*$, respectively. The operation $w_2(n+1) = w_2(n+1)$ denotes the transfer of coefficients, which is only applied in transient stage. By applying the weight transfer, the adaptation of $\mu_2$ NSA filter, as a consequence, the convergence of NSA-NSA filter. The cost of this weight
transfer scheme is actually smaller than that of original combination, because only one filter is adapted.

![Fig. 2](image.png)

**Fig. 2.** Comparison of EMSE of NSA-NSA for Gaussian input in example 1.

![Fig. 3](image.png)

**Fig. 3.** Comparison of EMSE of NRMN algorithm, NSA, and NSA-NSA for Gaussian input in example 2.

The excess means-square error (EMSE) resulting by comparing the tracking weight transfer scheme and no transfer scheme are shown in Figs. 2 and 3 in different situations (both the mixing parameter adjusted based on (6), and the same step-size are chosen). As can be seen, the slow filter of tracking weight transfer moves up compared to the slow filter with no transfer scheme. Therefore, the overall performance of the filter bank is elevated by transfer scheme. It reveals from these figures that the proposed weight transfer scheme exhibits faster convergence than that of no transfer scheme.

In summary, the proposed algorithm is summarized by using pseudocode as Table 1.

### 2.3. Computational complexity

Table. 2 summarizes the computational complexity of the Basic-CLMS [20], NLMS-NSA [23] and the NSA-NSA algorithms. Since the basic-CLMS combines two LMS algorithms, it requires $4M^2+2$ multiplications for the adaptation of the component filters. The NLMS-NSA algorithm
provides additional insensitivity to the input signal level by combining the NLMS and NSA, it requires $6M+1$ multiplications for the adaptation of the component filters. In contrast, the proposed algorithm uses NSA as the fast filter to replace the NLMS filter, which reduces the computational burden and the impact from impulsive noise. From (1) and (4), the basic-CLMS and NLMS-NSA algorithms need 6 more multiplications to compute the filter output and to update $a(n)$. In contrast, the proposed algorithm needs 5 multiplications for update $a(n)$ (see (1) and (6), respectively). According to (9), all the algorithms demand $2M$ additional multiplications to calculate the explicit weight vector. Moreover, due to using the slide window of tracking weight transfer scheme, the NSA-NSA algorithm can further reduce the computation operations. Consequently, these would lead to significant computational efficiency.

Table 2. Summary of the computational complexity.

| Algorithms      | Component filter adaptation | Basic combination | Explicit weight calculation | Weight transfer |
|-----------------|-----------------------------|-------------------|-----------------------------|-----------------|
| Basic-CLMS [20] | $4M+2$                      | 6                 | $2M$                        | $2M$            |
| NLMS-NSA [23]   | $6M+1$                      | 6                 | $2M$                        | $M+3$           |
| NSA-NSA         | $6M$                        | 5(using (6))      | $2M$                        | No              |

2.4. Steady-state performance of the proposed algorithm

To measure the stationary performance, the EMSEs of the filters are expressed by [20]

$$J_{x,i}(\infty) = \lim_{n \to \infty} E[e_{x,i}^2(n)], \quad i = 1, 2$$

(17)

$$J_{x}(\infty) = \lim_{n \to \infty} E[e_{x}^2(n)]$$

(18)

$$J_{x,x2}(\infty) = \lim_{n \to \infty} E[e_{x,1}(n)e_{x,2}(n)]$$

(19)

where $E[\cdot]$ denotes expected value, $J_{x,i}(\infty)$ represents the EMSE of individual filter, $J_{x}(\infty)$ is cross-EMSE of combination of filters, $J_{x,x2}(\infty)$ the steady-state correlation between the a priori errors of the elements of the combination, $e_{x,i}(n)$ and $e_{x}(n)$ are priori error, defined by

$$e_{x,i}(n) = (w_{i0} - w_{i}(n))^T x(n) = \zeta_{i}^T(n)x(n)$$

(20)

$$e_{x}(n) = (w_{0} - w(n))x(n) = \zeta^T(n)x(n)$$

(21)

where $\zeta_{i}(n)$ is the weight error vector of individual filter, and $\zeta(n)$ is the weight error vector of the overall filter.

In addition, for the modified combination (1), $J_{x,x}(\infty)$ is defined in terms of $\lambda(n)$, i.e.,

$$J_{x,x}(\infty) = \lim_{n \to \infty} E[\lambda^2(n)e_{x,1}^2(n) + [1 - \lambda(n)]^2 e_{x,2}^2(n) + 2\lambda(n)(1 - \lambda(n))e_{x,1}(n)e_{x,2}(n)]$$

(22)

where

$$\lambda(n) = \begin{cases} 1 & a(n) \geq a^* - \varepsilon \\ \tilde{\lambda}(n) & a^* - \varepsilon > a(n) > -a^* + \varepsilon \\ 0 & a(n) \leq -a^* + \varepsilon \end{cases}$$

(23)

and $\varepsilon$ is a small positive constant.
Taking expectations of both sides of (6) and using \( y_1(n) - y_2(n) = e_{a,2}(n) - e_{a,1}(n) \), we arrive at:

\[
E[a(n + 1)] = E[a(n)] + \mu_a E[\text{sign}(e(n))[e_{a,2}(n) - e_{a,1}(n)]\lambda(n)[1 - \lambda(n)]]
\]

(24)

According to the Price theorem [25,26]

\[
E[\text{sign}(e(n))\theta(n)] \approx \sqrt{\frac{2}{\pi}} \frac{1}{\chi_{e,n}} E[e(n)\theta(n)]
\]

(25)

where \( \chi_{e,n} \) is the variance of \( e(n) \) assuming the error has zero mean. According to (25), \( \theta(n) \) can be defined as

\[
\theta(n) = [e_{a,2}(n) - e_{a,1}(n)]\lambda(n)[1 - \lambda(n)].
\]

Therefore, (24) becomes

\[
E[a(n + 1)] = E[a(n)] + \phi_a E[e(n)[e_{a,2}(n) - e_{a,1}(n)]\lambda(n)[1 - \lambda(n)]].
\]

(26)

For convenience, introducing the constant \( \phi_a = \mu_a \sqrt{\frac{2}{\pi}} \frac{1}{\chi_{e,n}} \), (26) is approximately equal to:

\[
E[a(n + 1)] \approx [E[a(n)] + \phi_a E[\lambda(n)[1 - \lambda(n)]]
\]

(27)

\[
\phi_a E[e_{a,2}(n) - e_{a,1}(n)]\lambda^2(n)[1 - \lambda(n)]v_{e,n}.
\]

Assume \( \lambda(n) \) is independent of \( e_{a,n}(n) \) in steady state, under this assumption, \( E[a(n + 1)] \) is governed by

\[
E[a(n + 1)] \approx [E[a(n)] + \phi_a E[\lambda(n)[1 - \lambda(n)]]
\]

(28)

\[
- \phi_a E[\lambda^2(n)[1 - \lambda(n)]\Delta J_n]v_{e,n}.
\]

where \( \Delta J_i = J_{\alpha,i}(\infty) - J_{\alpha,i2}(\infty), i = 1,2 \). Suppose the NSA-NSA converges, the optimal mean combination weights under convex constraint are given by [20], which is discussed in three situations as follows:

1) If \( J_{\alpha,1}(\infty) \leq J_{\alpha,2}(\infty) \leq J_{\alpha,2}(\infty), \) we have \( \Delta J_1 \leq 0, \Delta J_2 \geq 0. \) Since \( a(n) \) and \( \lambda(n) \) are limited in effective range, an assume can be expressed as

\[
E[a(n + 1)] \geq [E[a(n)] + C]v_{e,n} \quad \text{as } n \to \infty
\]

(29)

where \( C = \lambda^+ (1 - \lambda^+)^2 (\Delta J_2 - \Delta J_1) \) is a positive constant. A conclude can be obtained in this case that

\[
\begin{align*}
J_{\alpha,1}(\infty) & = J_{\alpha,2,1}(\infty) \\
J_{\alpha,1}(\infty) & = J_{\alpha,2,2}(\infty) \cdot
\end{align*}
\]

(30)

Therefore, (30) shows that the performance of NSA-NSA is as well as individual filters.

2) if \( J_{\alpha,1}(\infty) \geq J_{\alpha,2}(\infty) \geq J_{\alpha,1}(\infty), \Delta J_1 \geq 0 \) and \( \Delta J_2 \leq 0 \) are given. Then, (28) can be rewritten as

\[
E[a(n + 1)] \leq [E[a(n)] - C]v_{e,n} \quad \text{as } n \to \infty.
\]

(31)

For a positive constant \( C = \lambda^- (1 - \lambda^-)^2 (\Delta J_1 - \Delta J_2) \) and

\[
\begin{align*}
J_{\alpha,1}(\infty) & = J_{\alpha,2,1}(\infty) \\
J_{\alpha,1}(\infty) & = J_{\alpha,2,2}(\infty) \cdot
\end{align*}
\]

(32)

As seen to be from (32), the performance of the overall filter is approximately equal to their best component filters.
3) if \( J_{\alpha,12}(\infty) < J_{\alpha,1}(\infty), i = 1, 2 \), we have \( \Delta J_1 > 0 \) and \( \Delta J_2 > 0 \).

Assuming that \( \lambda(n) \to 0 \) when \( n \to \infty \), (33) is obtained

\[
[1 - \lambda(\infty)]\Delta J_2 = \lambda(\infty)\Delta J_1
\]

(33)

\( \lambda(\infty) \) is given by

\[
\lambda(\infty) = \left[ -\frac{\Delta J_2}{\Delta J_1 + \Delta J_2} \right]^{1/\Delta J_1}.
\]

(34)

Hence, it can be concluded from (34) that: if \( J_{\alpha,1}(\infty) < J_{\alpha,2}(\infty) \), then \( \lambda^* \geq \lambda(\infty) > 0.5 \); if \( J_{\alpha,1}(\infty) > J_{\alpha,2}(\infty) \), so \( 0.5 \leq \lambda(\infty) > 1 - \lambda^* \).

Consider the following formula

\[
J_{\alpha}(\infty) = \lambda^2(\infty)J_{\alpha,1}(\infty) + \left[1 - \lambda(\infty)\right]^2 J_{\alpha,2}(\infty) + 2\lambda(\infty)\left[1 - \lambda(\infty)\right]J_{\alpha,12}(\infty)
\]

(35)

\[
J_{\alpha,12}(\infty) = \lambda^2(\infty)J_{\alpha,12}(\infty) + \left[1 - \lambda(\infty)\right]^2 J_{\alpha,2}(\infty) + 2\lambda(\infty)\left[1 - \lambda(\infty)\right]J_{\alpha,12}(\infty)
\]

(36)

Rearranging terms of (35), i.e.,

\[
J_{\alpha}(\infty) = \lambda(\infty)\left[J_{\alpha,1}(\infty) + \left[1 - \lambda(\infty)\right]J_{\alpha,12}(\infty)\right] + \left[1 - \lambda(\infty)\right]\left[\left[1 - \lambda(\infty)\right]J_{\alpha,2}(\infty) + \lambda(\infty)J_{\alpha,12}(\infty)\right]
\]

(37)

Then, (34) is used to rewrite (37) as

\[
J_{\alpha}(\infty) = \lambda(\infty)\left[J_{\alpha,1}(\infty) + \left[1 - \lambda(\infty)\right]J_{\alpha,12}(\infty)\right] + \left[1 - \lambda(\infty)\right]\left[J_{\alpha,1}(\infty) + \left[1 - \lambda(\infty)\right]J_{\alpha,2}(\infty)\right]
\]

(38)

Since \( \lambda(\infty) = \Delta J_2 / (\Delta J_1 + \Delta J_2) \) and \( 1 - \lambda(\infty) = \Delta J_1 / (\Delta J_1 + \Delta J_2) \), giving

\[
J_{\alpha}(\infty) = \lambda(\infty)\left[J_{\alpha,1}(\infty) + \frac{\Delta J_2}{\Delta J_1 + \Delta J_2}\right] + \left[1 - \lambda(\infty)\right]\left[J_{\alpha,1}(\infty) + \frac{\Delta J_1}{\Delta J_1 + \Delta J_2}\right]
\]

(39)

Hence

\[
J_{\alpha}(\infty) = J_{\alpha,12}(\infty) = J_{\alpha,12}(\infty) + \frac{\Delta J_1\Delta J_2}{\Delta J_1 + \Delta J_2}
\]

(40)

Since \( \lambda(\infty) \in (1 - \lambda^*, \lambda^*) \), the following bound hold:

\[
J_{\alpha}(\infty) = J_{\alpha,12}(\infty) = J_{\alpha,12}(\infty) + \lambda(\infty)\Delta J_1 < J_{\alpha,1}(\infty)
\]

(41)

\[
J_{\alpha}(\infty) = J_{\alpha,12}(\infty) = J_{\alpha,12}(\infty) + \lambda(\infty)\Delta J_2 < J_{\alpha,2}(\infty)
\]

(42)

That is

\[
\begin{align*}
J_{\alpha}(\infty) &< \min\{J_{\alpha,1}(\infty), J_{\alpha,2}(\infty)\} \\
J_{\alpha,12}(\infty) &< \min\{J_{\alpha,1}(\infty), J_{\alpha,2}(\infty)\}
\end{align*}
\]

(43)

From above three situations, it is clearly concluded that the performance of the proposed NSA-NSA filter is better than the best component filter.
3 Simulation Results

To evaluate the performance of the proposed algorithm, two examples of system identification are carried out in this simulation. The results presented here are obtained from 200 independent Monte Carlo trials. To measure the performance of the algorithms, the EMSE using logarithmic coordinates is applied, which is defined as:

\[
\text{EMSE} = 10 \log_{10} \{e^2(n)\} .
\] (44)

The unknown system was a ten-tap FIR transversal filter given by random. As input, White Gaussian noise (WGN) with zero mean and variance 1 is used. The system is corrupted by a WGN and an impulsive noise. The impulsive noise \(v(n)\) is generated from the Bernoulli-Gaussian (BG) distribution \([17-19,27]\)

\[
v(n) = A(n)I(n)
\] (45)

where \(A(n)\) is a binary independent identically distributed occurrence process with \(p\{A(n) = 1\} = c, p\{A(n) = 0\} = 1 - c\), and \(c=0.01\) denotes the probability of the occurrence of the impulsive interference \(I(n)\). For \(v(n)\) to be a realistic model of the impulsive noise, \(\sigma_i^2 = \text{var}\{I(n)\}\). Assuming that \(I(n)\) is uncorrelated with \(A(n)\) and have a symmetric amplitude distribution. The variance of \(v(n)\) can be calculated by

\[
\text{var}\{v(n)\} = c\sigma_i^2 .
\] (46)

3.1. Example 1

In this example, \(\sigma_i^2\) is set as \(\sigma_i^2 = 10^4/12\) \([17-19,27]\), and the WGN with SNR=10dB. At \(n=10000\), the plant changes abruptly.

![Fig. 4. The choice of parameter \(N_0\) in example 1.](image-url)
Fig. 5. The choice of parameter $\rho_\alpha$ in example 1 (the mixing parameter $a(n)$).

Fig. 4 and 5 show the different values of $N_0$ and $\rho_\alpha$ of proposed algorithm. The filter values of the NSA with $\mu_1 = 0.05$, $\mu_2 = 0.005$ (which satisfies the stability condition), $a(0) = 0$, $\lambda(0) = 0.5$, and $\epsilon_1 = \epsilon_2 = 0.0001$. During the period of the weight transfer, $a^+$ is fixed at 4. Consider the stability of evolution of the mixing parameter and the convergence rate, the best choice is $N_0 = 2$. In addition, it is observed from Fig. 5 that the best choice is $\rho_\alpha = 10$.

Figs. 6 and 7 display the evolution of the mixing coefficients parameters $\lambda(n)$ and $a(n)$ in NSA-NSA, two curves have similar trends. Run 1 is the no transfer scheme [20]. The Run 2 and Run 3 represent the mixing parameters based on the tracking weight transfer scheme are updated according to (4) and (6), respectively. It demonstrate that the proposed method achieve rapidly changed towards mixing parameter in terms of the convergence rate, also improves its robustness in the impulsive noise environment. Moreover, simulation results from Figs. 6 and 7 illustrate that adjust the mixing parameter $a(n)$ by using (6) (Run3) has better stability than other methods.
Fig. 7. Evolution of the mixing parameter $a(n)$ of NSA-NSA.

Fig. 8. Comparison of EMSE of NLMS-NSA algorithm and NSA-NSA for Gaussian input when 1% impulsive noises are added.

Fig. 9. Comparison of EMSE of NRMN, NSA, VSS-NSA, VSS-APSA algorithms and NSA-NSA for Gaussian input when 1% impulsive noises are added.
To demonstrate the performance clearly, Fig. 8 depicts the learning curves of NLMS-NSA and NSA-NSA algorithms. This figure verifies performance of the proposed algorithm is at least as good as the best element in the mixture, and the same as the theoretical analysis in Section 2.4. Both algorithms keep the same misadjustment due to the values of $\mu_i$ are the same. However, the fast filter of the NLMS-NSA is the NLMS algorithm, which reaches the high misadjustment in high background noise environment. As a result, the NLMS-NSA algorithm suffer from relatively high misadjustment of in initial convergence stage. Fig. 9 plots a comparisons of NRMN, NSA, VSS-NSA, VSS-APSA and the proposed algorithm. As can be seen, the NSA has the conflicting requirement of fast convergence rate and low EMSE. In contrast, the proposed algorithm has well balance between steady-state error and convergence rate.

3.2. Example 2

Compared to example 1, this example relates to the level of the impulsive noise decreases and the level of the Gaussian noise increases. The parameter $\sigma_i^2$ is set as $\sigma_i^2 = 10^4 / 20$, SNR=5dB, respectively. The desired signal rapidly changes direction at $n=10000$.

In this example, the step-size of NSA-NSA filter is selected at $\mu_1 = 0.05$, $\mu_2 = 0.008$, $\varepsilon_1 = \varepsilon_2 = 0.0001$, $\alpha(0) = 0$, and $\lambda(0) = 0.5$. This selection of the parameters can ensure good performance of the convergence property and the misadjustment of the algorithm. Fig. 10 displays the choice of the parameter $N_o$ in example 2. As we can see, the method is not sensitive to this selection, but it turns out that the best option is $N_o = 2$. Fig. 11 shows the selection of mixing parameters $\rho_a$. The mixing parameter $\rho_a = 10$ for algorithm is selected to guarantee the stability.

Figs. 12 and 13 show the time evolution of the mixing coefficients, where Run 1 represents no weight transfer scheme [20], and the Run 2 and Run 3 represent the mixing parameters based on the tracking weight transfer scheme are updated according to (4) and (6), respectively. Obviously, it can be observed from these figures that the best selection is Run 3. The robust performance in the presence of impulsive noise is also improved by using (9) to update mixing parameter.

![Fig. 10. The choice of parameter $N_o$ in example 2.](image-url)
Fig. 11. The choice of parameter $\rho_\alpha$ in example 2. (the mixing parameter $a(n)$).

Fig. 12. Evolution of the mixing parameter $\lambda(n)$ of NSA-NSA.

Fig. 13. Evolution of the mixing parameter $a(n)$ of NSA-NSA.
Fig. 14. Comparison of EMSE of NLMS-NSA algorithm and NSA-NSA for Gaussian input when 1% impulsive noises are added.

Fig. 15. Comparison of EMSE of NRMN, NSA, VSS-NSA, VSS-APSA algorithms and NSA-NSA for Gaussian input when 1% impulsive noises are added.

Fig. 14 plots the comparison of NLMS-NSA and the proposed algorithms. For fair comparison, the same parameters are selected for NLMS-NSA algorithm. We see that the EMSE of NSA-NSA is agree with the theoretical analysis again. Both algorithms reach quite similar steady-state error, but the proposed algorithm has the lower misadjustment in initial convergence stage. This is due to the fact that the NLMS algorithm is not suitable for the impulsive noise and high SNR environment. Fig. 15 shows a comparison of learning curves from NRMN, NSA, VSS-NSA [13], VSS-APSA [28] and NSA-NSA for high Gaussian noise and low impulsive noise environment. Again, it is observed that the proposed algorithm achieve the improved performance in presence of the impulsive noise.

4 Conclusions

A novel NSA-NSA algorithm was presented to improve the performance of NSA for system identification under the impulsive noise. The presented scheme which was adaptive convex
combine of one fast and one slow NSA filters constitutes a reasonable approach to get both speed and steady-state error. Moreover, a sign cost function scheme to adjust mixing parameter was developed in this paper, which improves the robustness of the algorithm in impulsive noise. To further accelerate the initial convergence rate, a tracking weight transfer scheme was applied in NSA-NSA. Simulation results demonstrated that the proposed algorithm has better performance in terms of convergence rate and steady-state error.

Acknowledgements. This work was partially supported by: (1) National Science Foundation of P.R. China (Grant: 61271340, 61071183, and U1134205), (2) the Sichuan Provincial Youth Science and Technology Fund (Grant: 2012JQ0046), (3) the Fundamental Research Funds for the Central Universities (Grant: SWJTU12CX026 and SWJTU11ZT07).

References
1. Sayed, A.H.: Fundamentals of Adaptive Filtering. Wiley IEEE Press (2003)
2. Eweda, E.: Analysis and design of signed regressor LMS algorithm for stationary and nonstationary adaptive filtering with correlated Gaussian data. IEEE Transactions on Circuits and Systems 37, 1367–1374. (1990)
3. Jun, B.E., Park, D.J., Kim, Y.W.: Convergence analysis of sign-sign LMS algorithm for adaptive filters with correlated Gaussian data. In: IEEE International Conference on Acoustics, Speech, and Signal Processing, pp. 1380–1383. (1995)
4. Koike, S.: Variable step size normalized sign algorithm for fast convergent adaptive filters with robustness against impulsive noise. NEC research & development 41, 278–28. (2000)
5. Yuvapoositanon, P., Chambers, J.: An adaptive step-size code-constrained minimum output energy receiver for nonstationary CDMA channels. In: IEEE International Conference on Acoustics, Speech, and Signal Processing, pp. 465–468. (2003)
6. Soo, J., Pang, K. K.: A multi step size frequency domain adaptive filter. IEEE Transactions on Signal Processing 39, 115–121 (1991)
7. Mathews, V.J., Xie, Z.: A stochastic gradient adaptive filter with gradient adaptive step size. IEEE Transactions on Signal Processing 41, 2075–2087 (1993)
8. Kwong, R.H., Johnston, E.W.: A variable step size LMS algorithm. IEEE Transactions on Signal Processing 40, 1633–1642 (1992)
9. Douglas, S.C.: A family of normalized LMS algorithms. IEEE Signal Processing Letters 1, 49–51 (1994).
10. Douglas, S.C.: Analysis and implementation of the max-NLMS adaptive filter. In PRoc. 29th Asilomar Conf. on Signals, Systems, and Computers, pp. 659–663. (1995)
11. Kwong, C.P.: Dual sign algorithm for adaptive filtering. IEEE Transactions on communications 34, 1272–1275 (1986)
12. Jebara, S.B., Besbes, H.: Variable step size filtered sign algorithm for acoustic echo cancellation. IET Electronics Letters,39, 936–938 (2003)
13. Shao, T., Zheng, Y.R., Benesty, J.: A variable step-size normalized sign algorithm for acoustic echo cancellation. In: IEEE International Conference on Acoustics, Speech, and Signal Processing, pp. 333–336. (2010)
14. Chambers, J.A., Tanrikulu, O., Constantinides, A.G.: Least mean mixed-norm adaptive filtering. IET Electronics Letters 30, 1574–1575 (1994)
15. Pazaitis, D.I., Constantinides, A.G.: LMS+F algorithm. IET Electronics Letters 31, 1423–1424 (1995)
16. Tanrikulu, O., Chambers, J.A.: Convergence and steady-state properties of the least-mean mixed-norm (LMMN) adaptive algorithm. IEE Proceeding Vision, Image and Signal Processing 143, 137–142 (1996)
17. Chambers, J., Avlonitis, A.: A robust mixed-norm adaptive filter algorithm. IEEE Signal Processing Letters 4, 46–48 (1997)
18. Mandic, D., Papoulis, E., Boukis, C.: A normalized mixed-norm adaptive filtering algorithm robust under impulsive noise interference. In: IEEE International Conference on Acoustics, Speech, and Signal Processing, pp. 333–336. (2003)
19. Papoulis, E.V., Stathaki, T.: A normalized robust mixed-norm adaptive algorithm for system identification. IEEE Signal Processing Letters 11, 56–59 (2004)
20. Arenas-Garcia, J., Figueiras-Vidal, A.R., Sayed, A.H.: Mean-square performance of a convex combination of two adaptive filters. IEEE Transactions on Signal Processing 54, 1078–1090 (2006)
21. Arenas-Garcia, J., Ramon, M.M., Vazquez, A.N., Figueiras-Vidal, A.R.: Plant identification via adaptive combination of transversal filters. Signal Processing 86, 2430–2438 (2006)
22. Parsons, J.D.: The mobile radio propagation channel, New York: Wiley. (2000)
23. Arenas-Garcia, J., Figueiras-Vidal, A.R.: Adaptive combination of normalised filters for robust system identification. Electronics Letters 41, 874–875 (2005)
24. Nascimento, V.H., de Lamare, R.C.: A low-complexity strategy for speeding up the convergence of convex combinations of adaptive filters. In: IEEE International Conference on Acoustics, Speech and Signal Processing, pp. 3553–3556. (2012)
25. Price, R.: A useful theorem for nonlinear devices having Gaussian inputs. IRE Transactions on Information Theory 4, 69–72 (1958)
26. Mathews, V.J., Cho, S.H.: Improved convergence analysis of stochastic gradient adaptive filters using the sign algorithm. IEEE Transactions on Acoustics, Speech and Signal Processing 35, 450–454 (1987)
27. Shao, T., Zheng, Y.R., Benesty, J.: An affine projection sign algorithm robust against impulsive interferences. IEEE Signal Processing Letters 17, 327-330 (2010)
28. Shin, J., Yoo, J., Park, P.: Variable step-size affine projection sign algorithm. IET Electronics Letters 48, 483–485 (2012)