IEICE TRANSACTIONS

on Fundamentals of Electronics, Communications and Computer Sciences

VOL. E105-A NO. 2
FEBRUARY 2022

A PUBLICATION OF THE ENGINEERING SCIENCES SOCIETY
The Institute of Electronics, Information and Communication Engineers
Kikai-Shinko-Kaikan Bldg., 5-8, Shibakoen 3 chome, Minato-ku, TOKYO, 105-0011 JAPAN

LETTER

Impulse-Noise-Tolerant Data-Selective LMS Algorithm

Ying-Ren CHIEN[†a)], ***Senior Member and*** **Chih-Hsiang YU**[††], ***Nonmember***

SUMMARY Exponential growth in data volumes has promoted widespread interest in data-selective adaptive algorithms. In a pioneering work, Diniz developed the data-selective least mean square (DS-LMS) algorithm, which is able to reduce specific quantities of computation data without compromising performance. Note however that the existing framework fails to consider the issue of impulse noise (IN), which can greatly undermine the benefits of reduced computation. In this letter, we present an error-based IN detection algorithm for implementation in conjunction with the DS-LMS algorithm. Numerical evaluations confirm the effectiveness of our proposed IN-tolerant DS-LMS algorithm.

key words: *data-selective least mean square (DS-LMS), impulse noise (IN)*

1. Introduction

Data-selective adaptive filtering algorithms have been attracting interest for their ability to reduce computational overhead by avoiding weight updates for adaptive filters in cases where the magnitude of the error signals is too small or too large. Diniz presented the basic framework of this approach in the form of data-selective least-mean-square (DS-LMS) algorithms [1]. Under the assumption of white Gaussian background measurement noise in conventional system identification problems, the existing framework allows users to select a prescribed weight update probability (P_{up}) and threshold parameter ($\tau_{\max}$), such that only innovative input data is used to update the weights of adaptive filters. Under this framework, adaptive filters are updated in according with a prescribed updating constraint, which allows operations with a tolerable degree of misadjustment.

Unfortunately, the existing scheme is prone to failure in situations where the measurement noise is non-Gaussian impulse noise (IN) rather than additive white Gaussian noise (AWGN). A number of robust adaptive filtering algorithms have been proposed to deal with IN; however, none of those works have been implemented with data-selective updating schemes aimed at reducing computational complexity. In [2], the authors presented a prefiltered observation-based adaptive filtering algorithm in which the weights of the adaptive filters are frozen when IN is detected. Note however that this approach suffers from a low IN detection rate, particularly when the input is a colored Gaussian signal. To the best of our knowledge, this is the first study to address the data-selective algorithm within the context of IN. In this letter, we first outline the method used to detect the presence of IN based on an error signal. We then outline our modification of the DS-LMS algorithm, which takes into consideration the potential effects of IN. Numerical simulations demonstrated that our modified DS-LMS implementation outperforms the original DS-LMS algorithm in the presence of IN.

2. System Models

Under a scenario of system identification, we denote the impulse response of the unknown system as $\mathbf{w}(n)$ and the inputs to this system as $x(n)$. The observable outputs $d(n)$ can be expressed as

$$d(n) = \mathbf{w}^T(n)\mathbf{x}(n) + v(n) + \eta(n), \tag{1}$$

where $\mathbf{w}(n) := [w_0(n), w_1(n), \ldots, w_{L-1}(n)]^T \in \mathbb{R}^{L\times 1}$ denotes a real weight vector of length L; superscript T denotes the transpose operation; $\mathbf{x}(n) := [x(n), x(n-1), \ldots, x(n-L+1)]^T \in \mathbb{R}^{L\times 1}$ denotes the real input regressor vector of length L; and n denotes the time index. This system model considers two additive noise sources: background white Gaussian noise $v(n)$ with zero-mean and variance σ_v^2, and IN $\eta(n)$. IN can be described using the Bernoulli-Gaussian (BG) IN model as follows:

$$\eta(n) = b(n) \cdot g(n), \tag{2}$$

where $b(n)$ denotes the Bernoulli process with a probability of occurrence of IN p and $g(n)$ is the white Gaussian process with zero-mean and variance of σ_g^2. Two parameters are used to represent the intensity of the IN, i.e., the occurrence probability of IN p and the Gaussian-to-IN ratio (GINR) $\Gamma := \sigma_v^2/\sigma_g^2$. Higher values of p and smaller values of Γ indicate a situation with stronger additive IN.

3. Proposed IN Tolerable DS-LMS Algorithm

3.1 Review of Diniz DS-LMS Algorithm

Without loss of generality, the information of the length of the unknown system (i.e., L) is assumed to be a piece of given information. The recursive updating of the original Diniz DS-LMS algorithm is as follows:

Manuscript received May 29, 2021.
Manuscript revised July 5, 2021.
Manuscript publicized August 2, 2021.

[†]The author is with Department of Electrical Engineering, National Ilan University, Yilan City, 26041, Taiwan, R.O.C.

[††]The author is with Graduate Institute of Communication Engineering, National Taiwan University, Taipei City, 10617, Taiwan, R.O.C.

a) E-mail: yrchien@niu.edu.tw (Corresponding author)

DOI: 10.1587/transfun.2021EAL2046

$$\widehat{\mathbf{w}}(n+1) = \widehat{\mathbf{w}}(n) + \mu \cdot \delta(n) \cdot \mathbf{x}(n) \cdot e(n), \tag{3}$$

where $\widehat{\mathbf{w}}(n) \in \mathbb{R}^{L\times 1}$ is the weight vector of the adaptive filter; μ is the step-size; $e(n) = d(n) - \widehat{y}(n)$ is the error signal; $\widehat{y}(n) = \widehat{\mathbf{w}}^T(n)\mathbf{x}(n)$ is the output of the adaptive filter; and $\delta(n)$ is the selection factor, which is defined as follows:

$$\delta(n) = \begin{cases} 0, & \text{if } \frac{e^2(n)}{\sigma_v^2} \le \tau(n) \\ 0, & \text{if } \frac{e^2(n)}{\sigma_v^2} > \tau_{\max} \\ 1, & \text{otherwise} \end{cases}. \tag{4}$$

Note that the DS-LMS algorithm freezes the weight vector when the error signal is either too small or too large.

For a given probability of updating P_{up}, the original derives threshold τ under the assumption of no outliers (i.e., $\tau_{\max} \approx \infty$), as follows:

$$\tau = (1+\alpha)\left(Q^{-1}\left(P_{up}/2\right)\right)^2, \tag{5}$$

where α is selected such that $\sigma_e^2 \approx (1+\alpha)\sigma_v^2$; σ_e^2 represents the variance of the error signal calculated within a specified observation window; and $Q^{-1}\{\cdot\}$ denotes the inverse of the complementary Gaussian cumulative distribution function. Note that the DS-LMS algorithm expects the weight updating $\widehat{\mathbf{w}}$ to occur with probability P_{up}, while the variance of error is maintained at a prescribed value $\sigma_e^2 \approx (1+\alpha)\sigma_v^2$. However, when the observed signal is corrupted by strong outliers, (i.e., IN), this approach does not provide a systematic means to determine the value of $\tau_{\max}$.

3.2 Proposed Error-Based IN Detection

We first derive the bound of $e^2(n)$ in the absence of IN. The square of the instantaneous error signal can be expressed as follows:

$$e^2(n) = \tilde{e}^2(n) + v^2(n) + 2\tilde{e}(n)v(n), \tag{6}$$

where $\tilde{e}(n) := y(n) - \widehat{y}(n) = \mathbf{x}^T(n)\widetilde{\mathbf{w}}(n)$ and $\widetilde{\mathbf{w}}(n) := \mathbf{w}(n) - \widehat{\mathbf{w}}(n)$. Leverage of the Peter-Paul inequality [3, p.105] for the products of any two real numbers a and b means the following inequality holds for all $\epsilon > 0$:

$$2ab \le \frac{a^2}{\epsilon} + \epsilon b^2. \tag{7}$$

After some simple manipulations, we can rewrite (6) as follows:

$$e^2(n) \le \left(1 + \frac{1}{\epsilon}\right)\tilde{e}^2(n) + (1+\epsilon)v^2(n). \tag{8}$$

Moreover, $\tilde{e}^2(n) \approx ||\mathbf{x}(n)||_2^2||\widetilde{\mathbf{w}}(n)||_2^2$ and $||\mathbf{x}(n)||_2^2 \approx L\sigma_x^2$ [4]; therefore, we obtain

$$\begin{aligned}\tilde{e}^2(n) &\approx L\sigma_x^2||\widetilde{\mathbf{w}}(n)||_2^2 \\ &\approx L\left(\sigma_e^2 - \sigma_v^2\right).\end{aligned} \tag{9}$$

By substituting (9) into (8) and selecting L for the value of ϵ, we obtain the following:

$$\begin{aligned} e^2(n) &\le \left(\frac{L+1}{L}\right)L\left(\sigma_e^2 - \sigma_v^2\right) + (1+L)v^2(n) \\ &\approx (1+L)\sigma_e^2(n), \end{aligned} \tag{10}$$

where we assume that $v^2(n) \approx \sigma_v^2(n)$.

Next, a median filter [5] is used to estimate the value of $\sigma_e^2(n)$ for the case of environments involving IN as follows:

$$\sigma_e^2(n) = \gamma\sigma_e^2(n-1) + C_1(1-\gamma)\mathrm{med}(\sigma_{\mathbf{e}}^2(n)), \tag{11}$$

where γ is a positive forgetting factor that is smaller than but close to one [6]. The larger value of γ implies that the estimation of $\sigma_e^2(n)$ depends heavier on the estimation of $\sigma_e^2(n-1)$ and depends less on the output of the median filter. $\sigma_{\mathbf{e}}^2(n) = \left[e^2(n), e^2(n-1), \ldots, e^2(n-N_w+1)\right]$ with N_w as the observation window; med$(\cdot)$ denotes the sample median operation; and $C_1 = 1.483/(1+5/(N_w-1))$ is the finite sample correction factor [7]. Practically, the value of N_w can be selected to lie between 5 and 11 so that the computational complexity incurred by (11), which is $O(N_w \log N_w)$ [8], is affordable [9].

Finally, we claim that the IN is detected if (10) does not hold. Therefore, we modify the selection factor in (3) as follows:

$$\delta(n) = \begin{cases} 0, & \text{if } \frac{e^2(n)}{\sigma_v^2} \le \tau(n) \\ 0, & \text{if } e^2(n) > (L+1)\sigma_e^2(n) \\ 1, & \text{otherwise} \end{cases}. \tag{12}$$

4. Simulation Results

The effectiveness of the IN tolerat DS-LMS algorithm was evaluated by conducting numerical simulations of the system identification problem with a target $P_{up} = 30\%$. Referring to [1], we set the impulse response of the unknown system as follows:

$$[0.1010 \;\; 0.3030 \;\; 0 \;\; -0.2020 \;\; -0.4040 \;\; -0.7071 \;\; -0.4040 \;\; -0.2020].$$

Moreover, we assumed that the unknown system had undergone an abrupt change in the 5000th iteration. The step-size $\mu = 1/\left(\nu(L+1)\sigma_x^2\right)$ with $\nu = 5$, and $\alpha = P_{up}/\left(\nu - P_{up}\right)$ [1, Table IV]. Note that the forgetting factor and the length of the observation windows in (11) are $\gamma = 0.9$ and $N_w = 11$, respectively.

Variance in the additive background Gaussian noise is determined by the SNR, which is given as

$$\mathrm{SNR} := 10\log_{10}\left(\sigma_y^2/\sigma_v^2\right), \tag{13}$$

where σ_y^2 is the variance of the outputs of the unknown system. In our simulation, the SNR value was set at 30 dB. The other additive noise was BG IN. We considered three IN environments: 1) weakly disturbed (p = 0.01 and Γ = 0.1); 2) moderately disturbed (p = 0.05 and Γ = 0.01); and 3) strongly disturbed (p = 0.08 and $\Gamma = 5 \times 10^{-3}$) [10].

We considered four cases. The first case was the ideal case, labeled "Diniz DS-LMS (no IN)". In this case, there is no IN and the only additive noise is AWGN noise. The second case was labeled "Diniz DS-LMS", in which it was assumed that $\tau_{\max} = \infty$. The third case was labeled "Jeong IN Detection", in which we selected $\tau_{\max}$ in accordance with prefiltered observation data to detect outliers (i.e., IN [2]). The fourth case was labeled "Proposed method" using the selection factor defined in (12).

The performance metric in this study was the normalized mean-square deviation (NMSD), which can be expressed as follows:

$$\mathrm{NMSD}(n) := 10 \log_{10} \left(\frac{||\mathbf{w}(n) - \widehat{\mathbf{w}}(n)||_2^2}{||\mathbf{w}(n)||_2^2} \right). \qquad (14)$$

We also compared the averaged update probability P_{up}, the detection rate, and the false alarm of IN detection. All results were obtained by averaging the results from 300 independent trials.

4.1 White Gaussian Inputs

The input signals were white Gaussian signals with zero-mean and unit variance. Figure 1 illustrates the NMSE learning curves under various IN disturbance levels. We can see that the NMSE learning curve of the proposed method is close to that of the ideal case. To better emphasize the difference between the Diniz DS-LMS (no IN) and our proposed method, we show the zoom-in plots for the number of iterations $n \in [180, 820]$. Note that our proposed method is comparable to the ideal case when the measurement noises contained weak or moderate IN. However, our method causes few spikes in the resulted NMSE learning curves for the strong IN case. By contrast, the original Diniz method and the Jeong method presented a higher degree of misadjustment.

4.2 Auto-Regressive (AR) Inputs

In accordance with [1], we considered fourth-order AR(4) processes, by filtering the white Gaussian signal with zero-mean and unit variance using the following Z-transfer function:

$$H(z) = \frac{1}{1 + b_1 z^{-1} + b_2 z^{-2} + b_3 z^{-3} + b_4 z^{-4}},$$

where b_1 = 0.55, b_2 = 0.221, b_3 = 0.49955, and b_4 = 0.4536 to obtain the AR(4) signals. As shown in Fig. 2, the eigenvalue spread associated with the AR(4) signals was larger than that associated with white Gaussian inputs, which slowed the rate of convergence to below that of white Gaussian inputs. Other behaviors were consistent with the case of white Gaussian inputs.

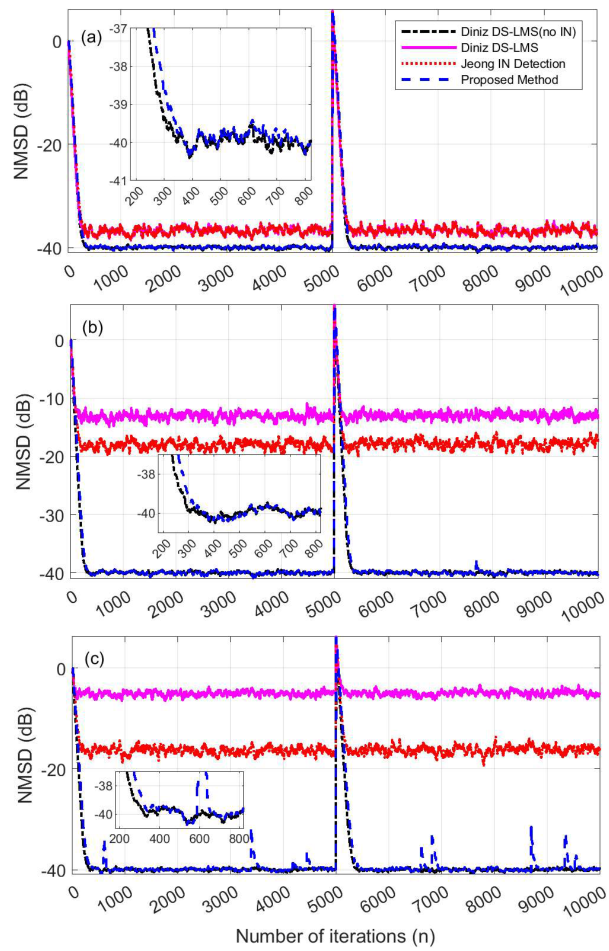

Fig. 1 Learning curves of NMSD for white Gaussian inputs under various disturbance levels: (a) weak IN; (b) moderate IN; and (c) strong IN.

Table 1 presents a summary of performance comparisons under various scenarios. Note that Methods 1 to 4 respectively refer to "Diniz DS-LMS (no IN)", "Diniz DS-LMS", "Jeong IN Detection", and "Proposed Method".

1. Comparison of averaged update probability
   Our simulation results revealed that the averaged update probability P_{up} was affected by the type of input signals as well as the types of IN. Method 1 refers to the ideal case without IN. Among the four methods, the average P_{up} of the proposed method was closest to the target (P_{up} = 0.3). When the input type was non-white Gaussian, some of the assumptions outlined in Sect. 3.2 did not necessarily hold. This led to a slight increase in the average P_{up} when the input type was AR(4). Under these conditions, the original Diniz DS-LMS method (Method 2) and Jeong IN detection method (Method 3) performed poorly, resulting in an average P_{up} that was far higher than the target P_{up}.
2. Comparison of detection and false alarm rates
   Note that only Methods 3 and 4 take into account the issue of IN. Jeong IN detection was unable to detect the occurrence of IN, particularly when the IN was

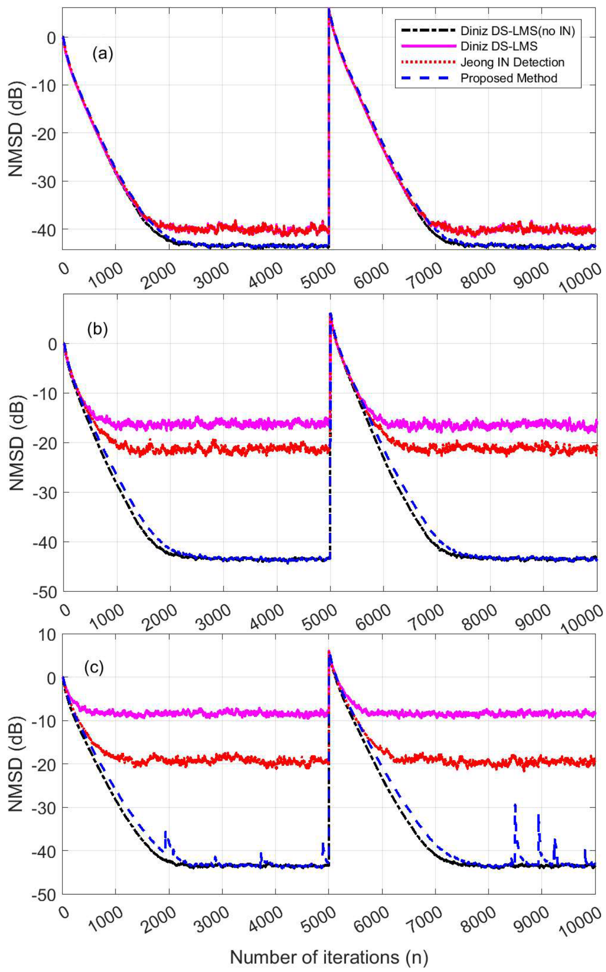

Fig. 2 Learning curves of NMSD for AR(4) inputs under various disturbance levels: (a) weak IN; (b) moderate IN; and (c) strong IN.

Table 1 Performance comparison under various scenarios with target P_{up} = 30%.

Input Type		White Gaussian			AR(4)		
IN Type	Method	Weak	Moderate	Strong	Weak	Moderate	Strong
Averaged Update Probability	1	0.3271	0.3265	0.3268	0.4023	0.4010	0.4011
	2	0.3507	0.8157	0.9340	0.4221	0.8227	0.9326
	3	0.3497	0.6812	0.6943	0.4209	0.6996	0.7131
	4	0.3222	0.3101	0.3024	0.3999	0.3869	0.3789
Detection Rate	3	0.0069	0.2971	0.5639	0.0063	0.3043	0.5645
	4	0.7073	0.9581	0.9761	0.6402	0.9354	0.9628
False Alarm Rate	3	2.61E-03	3.56E-03	2.93E-03	2.44E-03	2.97E-03	2.31E-03
	4	6.12E-03	4.63E-03	3.74E-03	5.95E-03	4.54E-03	3.77E-03

weak. The proposed method was able to detect 70% of the IN in the case of weak IN, and more than 90% IN detection rate when the IN is moderate or strong disturbance. Note that the false alarm rate of both methods was neglectable.

Simulation results have confirmed that our proposed error-based IN detection method is superior to the Jeong IN detector, which used the desired signals to detect the IN. When no IN presence, condition $\sigma_e^2(n) < \sigma_d^2(n)$ holds. Thus, our detector is more sensitive than the Jeong method when the IN presents. Furthermore, the Jeong method may fail to detect the IN because the instantaneous power of IN is not always significantly more prominent than $(L+1)\sigma_d^2(n)$. Table 1 has shown that the detection rate is getting higher as the disturbance levels are getting higher. Even the sensitivity of our IN detection is high; however, the resulted false alarm rate is still kept at a tolerable level.

5. Conclusions

This letter presents an error-based IN detection algorithm,which can be used in conjunction with the original DS-LMS algorithm. Our IN-tolerant DS-LMS algorithm provides a means of dynamically determining a threshold for the detection of outliers in observed data. The modified DS-LMS algorithm is able to reduce computation overhead in accordance with a prescribed update probability, P_{up}, even in an environment of high impulse noise levels, without sacrificing the performance in terms of steady-state errors.

Acknowledgments

This work was supported in part by the Ministry of Science and Technology of Taiwan under Grant 109-2221-E-197-026.

References

[1] P.S.R. Diniz, "On data-selective adaptive filtering," IEEE Trans. Signal Process., vol.66, no.16, pp.4239–4252, 2018.
[2] J. Jeong and S. Kim, "Robust adaptive filter algorithms against impulsive noise," Circuits Syst. Signal Process., vol.38, pp.5651–5664, 2019.
[3] H. Sira-Ramírez, Sliding Mode Control: The Delta-Sigma Modulation Approach, Birkhauser-Springer, 2015.
[4] J. Benesty, H. Rey, L.R. Vega, and S. Tressens, "A nonparametric VSS NLMS algorithm," IEEE Signal Process. Lett., vol.13, no.10, pp.581–584, Oct. 2006.
[5] M.Z.A. Bhotto and A. Antoniou, "Robust quasi-Newton adaptive filtering algorithms," IEEE Trans. Circuits Syst. II: Exp. Briefs, vol.58, no.8, pp.537–541, 2011.
[6] Y. Zhou, S.C. Chan, and K.L. Ho, "New sequential partial-update least mean M-estimate algorithms for robust adaptive system identification in impulsive noise," IEEE Trans. Ind. Electron., vol.58, no.9, pp.4455–4470, 2011.
[7] J.J. Jeong, "A robust affine projection algorithm against impulsive noise," IEEE Signal Process. Lett., vol.27, pp.1530–1534, 2020.
[8] Y. Zou, S. Chan, and T. Ng, "A recursive least M-estimate (RLM) adaptive filter for robust filtering in impulse noise," IEEE Signal Process. Lett., vol.7, no.11, pp.324–326, 2000.
[9] S.C. Chan, Y. Wen, and K.L. Ho, "A robust past algorithm for subspace tracking in impulsive noise," IEEE Trans. Signal Process., vol.54, no.1, pp.105–116, 2006.
[10] Y.R. Chien, "Iterative channel estimation and impulsive noise mitigation algorithm for OFDM-based receivers with application to power-line communications," IEEE Trans. Power Del., vol.30, no.6, pp.2435–2442, 2015.