Research Article Vol. 33, No. 3 / 10 Feb 2025 / Optics Express 4582
Optics EXPRESS

Improvement of least mean square adaptive algorithm for non-monotonic systems

ZHENGYANG ZHANG,[1] SIWEI LIU,[1,*] STEFAN M. GOETZ,[1] AND OZGUR B. AKAN[1,2]

[1]Department of Engineering, University of Cambridge, Trumpington Street, Cambridge CB2 1PZ, United Kingdom
[2]Centre for neXt-generation Communications (CXC), Koç University, 34450 Istanbul, Turkey
[*]sl2049@cam.ac.uk

Abstract: This paper introduces the zlLMS algorithm, an improvement over the traditional least mean square (LMS) algorithm, addressing its limitations in handling nonlinear and non-monotonic transfer functions commonly encountered in engineering systems. The proposed method replaces the LMS algorithm's error inputs with a function monotonically correlated with the controllable signal. Through mathematical derivations and simulations, the zlLMS algorithm demonstrates superior performance in nonlinear adaptive filtering scenarios, such as the raised-cosine transfer function of Mach-Zehnder modulators and the Lorentz transfer function of diode lasers. Simulation results reveal that the amplitude difference between the ideal and zlLMS-equalized signals is reduced to 1/1000th of the original input signal's amplitude, underscoring its effectiveness in optics and terahertz technologies. These findings highlight the algorithm's robustness and potential for broad applications in engineering, especially where nonlinear dynamics are involved.

Published by Optica Publishing Group under the terms of the Creative Commons Attribution 4.0 License. Further distribution of this work must maintain attribution to the author(s) and the published article's title, journal citation, and DOI.

1. Introduction

The foundation for adaptive filtering, which would eventually lead to the development of the Least Mean Squares (LMS) algorithm, was established during the 1940s and 1950s. This period saw significant advancements in theories related to adaptive control systems and statistical signal processing [1]. In 1960, Bernard Widrow and Marcian E. Hoff at Stanford University introduced the LMS algorithm in the context of their work on adaptive neural networks known as "ADALINE" (Adaptive Linear Neuron) [2]. This groundbreaking work detailed the application of the LMS algorithm in minimizing the mean square error between a desired signal and the output of an adaptive filter. Later, Ted Hoff, co-inventor of the LMS algorithm, played a pivotal role in the development of the first microprocessor at Intel [3].

The LMS algorithm is fundamentally based on the principles of stochastic gradient descent method. The algorithm updates the filter weights using the following formula [4,5]:

$$\boldsymbol{w}_{k+1} = \boldsymbol{w}_k + 2\mu \cdot e_k \cdot \boldsymbol{x}_k \tag{1}$$

where:

- $\boldsymbol{w}_k$ represents the filter coefficient vector at iteration k,
- μ denotes the step size parameter,
- e_k is the error signal, defined as the difference between the desired output and the filter output, and

#546882 https://doi.org/10.1364/OE.546882
Journal © 2025 Received 18 Nov 2024; revised 5 Jan 2025; accepted 23 Jan 2025; published 29 Jan 2025

Optics EXPRESS

- $\boldsymbol{x}_k$ is the input signal.

The error signal e_k in (1) is given by:

$$e_k = d_k - y_k$$

where d_k is the desired output, and y_k is the filter output, which can be calculated as:

$$y_k = \boldsymbol{w}_k^T \boldsymbol{x}_k$$

Therefore, the error signal e_k can be expressed as:

$$e_k = d_k - \boldsymbol{w}_k^T \boldsymbol{x}_k$$

Note that the function describing the relationship between the input signal $\boldsymbol{x}_k$ and the output signal y_k is called the transfer function.

During the 1960s and 1970s, the LMS algorithm gained significant traction across various applications, including echo cancellation in telephony, adaptive noise canceling, and system identification [6–8]. The 1980s witnessed the development of several variants of the LMS algorithm, such as the Normalized LMS (NLMS) and Recursive Least Squares (RLS), which aimed to enhance convergence speed and stability [9,10].

From the 1990s to the present day, the LMS algorithm and its variants have continued to be widely utilized in fields such as digital signal processing, communications, control systems, and machine learning [11–14].

Advancements in computational power have significantly improved the efficiency of LMS algorithm implementations, enabling real-time applications. The principles underlying the LMS algorithm are foundational in many machine learning algorithms, particularly in linear regression and neural networks [15]. Ongoing research continues to develop new variants and enhancements of the LMS algorithm to address challenges posed by nonlinear systems, non-stationary environments, and large datasets [16–19]. Nowadays, the LMS algorithm remains a cornerstone in adaptive filtering and signal processing, underscoring its enduring relevance and adaptability to emerging technological challenges.

The LMS can reach the best performance when it is utilized to equalize the input signal ($\boldsymbol{x}_k$) that has a linear relation with the output signal of LMS (y_k). However, when the transfer function between $\boldsymbol{x}_k$ and y_k is nonlinear (or non-monotonic), the effectiveness of LMS reduces. In the worst case, the LMS is not able to equalize the signal correctly when the transfer function is non-monotonic.

As most engineering applications have nonlinear (even non-monotonic) transfer functions, the implementation of LMS is limited. For example, the transfer function in the Mach Zehnder Modulator (MZM) is non-monotonic (raised cosine function). Moreover, the non-monotonic Lorentz function also exists as a transfer function in a number of devices, such as diode lasers [20]. In addition, in some worse cases, the output signal y_k is not always available.

To tackle these disadvantages of LMS, we propose a novel way to implement the LMS algorithm. In our new method (named by zlLMS algorithm in the following), we do not use $\boldsymbol{w}_k^T \boldsymbol{x}_k$ and d_k for the error calculation of LMS. Instead, we will find an increasing function f and use $f(\boldsymbol{w}_k^T \boldsymbol{x}_k)$ and $f(d_k)$ to calculate the error in our new algorithm. In this case, $f(d_k)$ will become the desired output signal, and the algorithm becomes

$$\boldsymbol{w}_{k+1} = \boldsymbol{w}_k + 2\mu \boldsymbol{x}_k (f(d_k) - f(\boldsymbol{x}_k^T \boldsymbol{w}_k)).$$

The remaining sections of this paper are structured as follows: in Section 2, zlLMS is mathematically proved. We will show the convergence of our new algorithm. Moreover, we will apply our new algorithm to the signal processing of MZM in Section 3, showing that our

Optics EXPRESS

zlLMS algorithm can find the optimal weights in a more efficient way than the traditional LMS algorithm. A simulation will be provided for the MZM system equalized by zlLMS, followed by results and further discussions. In Section 4, diode lasers, which have the transfer function of Lorentz function, will be simulated to prove that zlLMS has generality and can be applied to all transfer functions with non-monotonic properties. Section 5 concludes the paper.

2. zlLMS algorithm

Notation 1 *In all the following proofs,* $\prod_{i=k_1}^{k_2} \boldsymbol{Q}_i = \boldsymbol{Q}_{k_1} \cdot \boldsymbol{Q}_{k_1-1} \cdots \boldsymbol{Q}_{k_2}$ *if* $k_1 > k_2$ *and* $k_1, k_2 \in \mathbb{Z}$ *for any series of matrix* $\{\boldsymbol{Q}_i\}$.

Throughout the following analysis, we assume that the input signal of our system is stationary. As we mentioned before, in some cases, the filter output signal $y_k = \boldsymbol{w}_k^T \boldsymbol{x}_k$ is not available, or the transfer function is non-monotonic. As a consequence, the traditional LMS algorithm will not generate correct results [21]. In this case, we can introduce an increasing function $f(x) : \mathbb{R} \to \mathbb{R}$ which is monotonic in the input signal's operation region and replace the error signal e_k by

$$e_k = f(d_k) - f(y_k). \tag{2}$$

Note that $f(d_k)$ represents the ideal signal in this case. Then, the whole algorithm becomes

$$\boldsymbol{w}_{k+1} = \boldsymbol{w}_k + 2\mu \boldsymbol{x}_k (f(d_k) - f(\boldsymbol{x}_k^T \boldsymbol{w}_k)). \tag{3}$$

We will show the convergence of the zlLMS algorithm step by step. Firstly, we will demonstrate that the traditional LMS algorithm, when applied with a bounded and random step size sequence, is convergent.

Lemma 2.1 *For any two positive numbers m and M that satisfy* $0<m<M$*, there exists a constant* $\gamma>0$ *such that if the step size of the traditional LMS algorithm is replaced by a bounded and random step size sequence* $\{\mu_1, \mu_2, \mu_3, \dots\}$ *with* $\mu_k \in [m\gamma, M\gamma]$ *for all* $k \in \mathbb{Z}^+$*, the weight sequence* $\{\boldsymbol{w}_1, \boldsymbol{w}_2, \boldsymbol{w}_3, \dots\}$ *generated from the LMS algorithm with the new step size sequence*

$$\boldsymbol{w}_{k+1} = \boldsymbol{w}_k + 2\mu_k \cdot (d_k - \boldsymbol{w}_k^T \boldsymbol{x}_k) \cdot \boldsymbol{x}_k$$

converges almost surely (a.s.) to $\boldsymbol{w}_0$*, where* $\boldsymbol{w}_0 = \boldsymbol{R}^{-1}\boldsymbol{p}$ *is the Wiener solution that the weight sequence of the LMS algorithm with a constant step size converges to. Following the notation in [4],* $\boldsymbol{R} = E[\boldsymbol{x}_k \boldsymbol{x}_k^T]$ *and* $\boldsymbol{p} = E[d_k \boldsymbol{x}_k]$*. Moreover,* $\boldsymbol{x}_k$ *is the* k^{th} *sample of the input signal, and* d_k *is the* k^{th} *sample of the ideal output signal.*

Proof. Following the notations in Section 3.3.2 of [4], we still denote $\boldsymbol{R} = \boldsymbol{Q}^T \boldsymbol{\Lambda} \boldsymbol{Q}$ and $\Delta \boldsymbol{w'}_k = \boldsymbol{Q}^T (\boldsymbol{w}_k - \boldsymbol{w}_0)$. Following the same ideas in Section 3.3.2 of [4], (3.18) in [4] becomes

$$E\left[\Delta \boldsymbol{w'_{k+1}}\right] = \prod_{i=k}^{1} (\boldsymbol{I} - \mu_i \boldsymbol{\Lambda}) E\left[\Delta \boldsymbol{w'_1}\right], \quad \mu_i \in [m\gamma, M\gamma], \quad \forall i \in \mathbb{N}^+.$$

When γ is positive but small, we know that

$$\left\|(\boldsymbol{I} - M\gamma\boldsymbol{\Lambda})^k E\left[\Delta \boldsymbol{w'_1}\right]\right\| \le E\left[\Delta \boldsymbol{w'_{k+1}}\right] \le \left\|(\boldsymbol{I} - m\gamma\boldsymbol{\Lambda})^k E\left[\Delta \boldsymbol{w'_1}\right]\right\|.$$

Moreover,

$$\lim_{k\to\infty} \left\|(\boldsymbol{I} - M\gamma\boldsymbol{\Lambda})^k E\left[\Delta \boldsymbol{w'_1}\right]\right\| = \lim_{k\to\infty} \left\|(\boldsymbol{I} - m\gamma\boldsymbol{\Lambda})^k E\left[\Delta \boldsymbol{w'_1}\right]\right\| = 0$$

if

$$0<\gamma<\frac{1}{M\lambda_{\max}},$$

Optics EXPRESS

where $\lambda_{\max}$ is the largest eigenvalue of $\boldsymbol{R}$. Therefore,

$$\lim_{k\to\infty} E\left[\Delta \boldsymbol{w}'_{\boldsymbol{k+1}}\right] = 0.$$

The weight sequence generated from the LMS with random and bounded step size is convergent in mean if $0<\gamma<1/M\lambda_{\max}$.

Then, we will prove this weight sequence is also convergent in variance. Following the same idea in Section 3.3.3 of [4], we can derive

$$0<\mu_k<\frac{1}{2\lambda_{\max}+\sum_{j=1}^{D}\lambda_j}$$

like (3.30) in [4], where $\lambda_1, \lambda_2, \ldots, \lambda_D$ are eigenvalues of $\boldsymbol{R}$. If this condition can be satisfied for all μ_k, $k \in \mathbb{N}^+$, the weight sequence will be convergent in variance. Note that $\mu_k \in [m\gamma, M\gamma]$, $\forall k \in \mathbb{N}^+$. Therefore, a sufficient condition for the variance of the weight sequence to converge is

$$0<\gamma<\frac{1}{M(2\lambda_{\max}+\sum_{j=1}^{D}\lambda_j)}.$$

Therefore, the weight sequence $\{\boldsymbol{w}_1, \boldsymbol{w}_2, \boldsymbol{w}_3, \ldots\}$ generated from the LMS with random and bounded step size sequence converges a.s. to $\boldsymbol{w}_0$ if

$$0<\gamma<\frac{1}{M(2\lambda_{\max}+\sum_{j=1}^{D}\lambda_j)}.$$

□

Then, we will prove the convergence of zlLMS with the above Lemma.

Theorem 2.2 *zlLMS is convergent. That is,*

$$\boldsymbol{w}_{k+1} = \boldsymbol{w}_k + 2\mu\boldsymbol{x}_k(f(d_k) - f(\boldsymbol{x}_k^T\boldsymbol{w}_k))$$

will generate a weight sequence $\{\boldsymbol{w}_1, \boldsymbol{w}_2, \boldsymbol{w}_3, \ldots\}$ *that a.s. converges to* $\boldsymbol{w}_0$*, where* $\boldsymbol{w}_0$ *is the Wiener solution of the traditional LMS with the constant step size. Moreover,* $f(x)$ *is a continuous and strictly increasing function satisfying* $0<m<f'(x)<M$ *for some positive constants m and M. Let* $\boldsymbol{x}_k$ *denote the* k^{th} *sample of the input signal,* μ *the positive constant step size, and* d_k *the* k^{th} *sample of the ideal output signal.*

Proof. By the mean value theorem,

$$\begin{aligned}\boldsymbol{w}_{k+1} &= \boldsymbol{w}_k + 2\mu\boldsymbol{x}_k(f(d_k) - f(\boldsymbol{x}_k^T\boldsymbol{w}_k))\\ &= \boldsymbol{w}_k + 2\mu f'(u_k)\boldsymbol{x}_k(d_k - \boldsymbol{x}_k^T\boldsymbol{w}_k)\\ &= \boldsymbol{w}_k + 2\mu_k^z\boldsymbol{x}_k(d_k - \boldsymbol{x}_k^T\boldsymbol{w}_k)\end{aligned}$$

if we denote $\mu_k^z = \mu f'(u_k)$ and u_k^z is a number lies between d_k and $\boldsymbol{x}_k^T\boldsymbol{w}_k$. By writing the formula of zlLMS into this form, we can see that for any weight sequence generated by zlLMS, there must exist another a.s. convergent weight sequence generated by the LMS with a bounded and random step size sequence that has the same initial weight, input signal, and step size sequence if

$$0<\mu<\frac{1}{M(2\lambda_{\max}+\sum_{j=1}^{D}\lambda_j)},$$

where $\lambda_1, \lambda_2, \ldots, \lambda_D$ are eigenvalues of $\boldsymbol{R} = E[\boldsymbol{x}_k\boldsymbol{x}_k^T]$ and $\lambda_{\max}$ is the largest eigenvalue of $\boldsymbol{R}$. Note that the μ in this theorem just likes the γ in Lemma 2.1 since $\mu_k^z = \mu f'(u_k) \in [m\mu, M\mu]$.

Optics EXPRESS

The distribution of each entry $\boldsymbol{w}_k$ of these two weight sequences is the same. By Lemma 2.1, we know that the weight sequence generated by the LMS with a bounded and random step size sequence converges to $\boldsymbol{w}_0$ a.s. Therefore, the weight sequence generated by zlLMS also converges to $\boldsymbol{w}_0$ a.s. □

Remark 1 *If $f(x)$ is a continuous and strictly decreasing function satisfying $0>-m>f'(x)>-M$ for some positive numbers m and M, the zlLMS is also convergent a.s. by changing the iteration formula into*

$$\boldsymbol{w}_{k+1} = \boldsymbol{w}_k - 2\mu \boldsymbol{x}_k (f(d_k) - f(\boldsymbol{x}_k^T \boldsymbol{w}_k)).$$

The idea of proof is exactly the same as before.

3. Application of zlLMS in MZM

All the simulations of this paper are performed in MATLAB R2016b (9.1.0.441655) 64-bit (win64).

3.1. Background

In this Section, we will take the application of the Mach-Zehnder modulator (MZM) in processing the Nyquist pulse train with a bias as an example to show the efficiency of the zlLMS algorithm. Previous works have shown that the traditional LMS is not compatible with the non-monotonic transfer function of MZM [20].

Nyquist pulse trains are widely used in various digital communication systems such as:

- **Digital Modulation**: In schemes like Quadrature Amplitude Modulation, Phase Shift Keying, and others where efficient use of bandwidth and minimizing ISI are crucial [22].
- **Data Transmission**: In technologies like Ethernet, Wi-Fi, and mobile communications, which rely on clear and reliable signal transmission [23].

MZM is a pivotal component in optical communication systems, serving as an electro-optic device that modulates the intensity, phase, or frequency of light. Its operation is based on the principles of interference and the electro-optic effect, making it a critical tool for encoding information onto an optical carrier signal [24].

The input signal to an MZM is primarily an optical carrier signal, typically provided by a continuous wave laser source. This optical carrier is characterized by its wavelength and power, and it is usually a coherent light source with a stable phase and amplitude. The optical signal is fed into the MZM through an optical fiber or integrated waveguide [25]. Moreover, the electrical signal can also be the input of MZM. This electrical signal serves as the modulating signal and can be either analog or digital. For digital modulation, the electrical signal typically consists of binary data representing the information to be transmitted. For analog modulation, the electrical signal could be a radio-frequency (RF) signal or any other analog waveform [26,27].

The output signal from the MZM is a modulated optical signal, where the information from the electrical input signal is encoded onto the optical carrier [28]. The modulated optical output is then transmitted through an optical fiber to the receiving end of the communication system, where it is demodulated to retrieve the original information [29,30].

3.2. Signaling in MZM

Firstly, we need to generate our ideal input signal $N_0(t)$—Nyquist pulse train with a bias (DC offset) from the single Nyquist pulse $h(t)$, that is,

$$h(t) = \frac{\sin\left(\frac{\pi t}{T_s}\right)\cos\left(\frac{\pi\beta t}{T_s}\right)}{\left(\frac{\pi t}{T_s}\right)\left(1-\left(\frac{2\beta t}{T_s}\right)^2\right)}, \tag{4}$$

Optics EXPRESS

where $T_s = 2.5$ ps is the pulse duration between zero crossings and $\beta = 0.5$ is the roll-off factor. The Dirac comb $\amalg_P(t)$ is then employed to convolve with the single Nyquist pulse $h(t)$ to produce a Nyquist pulse train $h_P(t)$ [21,31]. This process is described by the following equations:

$$h_P(t) = \amalg_P(t) * h(t), \tag{5}$$

$$\amalg_P(t) \triangleq \sum_{k=-\infty}^{\infty} \delta(t - kP) = \frac{1}{P} \amalg \left(\frac{t}{P}\right), \tag{6}$$

where $P = 10$ ps denotes the period of the Dirac comb, which is also the period of the Nyquist pulse train, and $\delta(t)$ is the Dirac delta function. The amplitude of the Nyquist pulse train $h_P(t)$ lies in $(-0.13, 1]$. In the following analysis, we will denote the value of the (normalized) highest input voltage of the Nyquist pulse train $h_P(t)$ by $V_\pi/2$. Therefore, $V_\pi = 2$. The ideal input signal for the transfer function (i.e., the ideal input for MZM), denoted as $N_0(t)$, is given by the sum of the signal $h_P(t)$ and the bias V_π, so that MZM can work as a pulse compressor to further narrow the width of the pulse's peak. Mathematically, this is expressed as:

$$N_0(t) = h_P(t) + V_\pi. \tag{7}$$

The Nyquist pulse train with a bias $N_0(t)$ generated by this process is shown in Fig. 1(bottom left). The amplitude of $N_0(t)$ lies in $(1.87, 3]$.

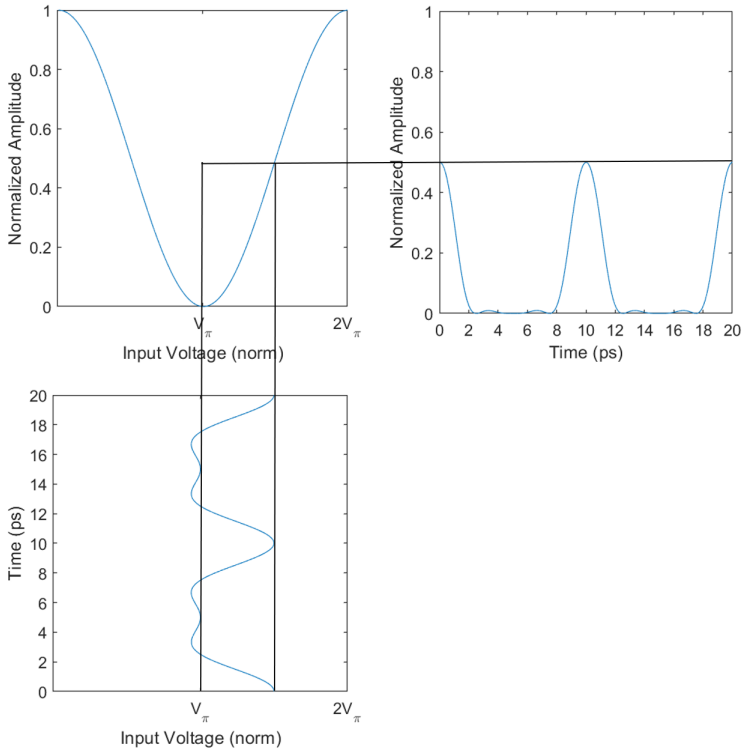

Fig. 1. Transfer function of MZM with $V_\pi = 2$ (top left) and its ideal input signal $N_0(t)$—Nyquist pulse train with a bias of V_π (bottom left) and the corresponding ideal output signal $S_0(t)$ (top right).

Optics EXPRESS

The transfer function of MZM between its input and output signals is a raised cosine function, that is,

$$T(t) = \frac{1}{2}\cos\frac{\pi t}{V_\pi} + \frac{1}{2}$$

(shown in Fig. 1(top left)), where $2V_\pi$ is the period of the transfer function (V_π is different between MZMs). The amplitude of the transfer function is normalized for convenience. Then, the ideal input signal $N_0(t)$ can pass through the transfer function in MZM and generate the ideal output signal $S_0(t)$ (shown in Fig. 1(top right)). Note that the transfer function of MZM is non-monotonic on its whole operating region. It is decreasing on $(V_\pi/2, V_\pi)$ and increasing on $(V_\pi,\ 3V_\pi/2)$. One way to solve this problem is to use zlLMS to replace the traditional LMS. Then, we need to find the increasing function f in zlLMS algorithm.

3.3. Function f of the zlLMS on MZM

Signal transmission in the optic-electronic system is inherently lossy. To more accurately reproduce the scene, we let the Nyquist pulse train with a bias $N_0(t)$ (ideal input signal) pass through a low-pass filter $LP(t)$ to produce the actual input signal $N(t)$ for the MZM. The signal $N(t)$ also serves as the input $\boldsymbol{x}_k$ for the zlLMS algorithm. Then,

$$N(t) = N_0(t) * LP(t).$$

Our simulation shows that the range of $N(t)$ is narrower than that of $N_0(t)$. The DC (constant) component of $N_0(t)$, whose amplitude is V_π, can pass entirely through the low-pass filter. We denote the actual output signal of the MZM as $S(t)$. Then, the signaling relationships in MZM are as follows,

$$\begin{aligned} S(t) &= T(N(t)), \\ T(t) &= \frac{1}{2}\cos\frac{\pi t}{V_\pi} + \frac{1}{2}. \end{aligned} \tag{8}$$

Then, we will show that the function f between $N(t)$ and $S'(t)/N'(t)$ is monotonic.

Theorem 3.1 *For the signal processing in MZM, the function f between $N(t)$ and $S'(t)/N'(t)$ is monotonic. $S(t)$ is the actual output signal of MZM and $N(t)$ is the actual input signal of MZM.*

Proof.

$$\begin{aligned} S(t) &= T\left(N(t)\right) \\ T(t) &= \frac{1}{2}\cos\frac{\pi t}{V_\pi} + \frac{1}{2} \\ S'(t) &= -\frac{\pi}{2V_\pi}\sin\left(\frac{\pi N(t)}{V_\pi}\right)\cdot N'(t) \\ \frac{S'(t)}{N'(t)} &= -\frac{\pi}{2V_\pi}\sin\left(\frac{\pi N(t)}{V_\pi}\right) \\ &= f\left(N(t)\right) \end{aligned}$$

Note that the DC component of $N_0(t)$, whose amplitude is V_π, can pass entirely through the low-pass filter. Since the amplitude of $N_0(t)$ lies between $V_\pi/2$ and $3V_\pi/2$ and the low-pass filter can narrow the range of the signal, the amplitude range of $N(t)$ lies strictly between $V_\pi/2$ and $3V_\pi/2$. Since $T'(t) = f(t) = -\pi/2V_\pi \sin(\pi t/V_\pi)$ is monotonic in the range of $N(t)$, $N(t)$ is monotonic with $S'(t)/N'(t)$. □

Therefore, by Theorem 2.2 and 3.1, we can use $S'(t)/N'(t)$ as $f(\boldsymbol{x}_k^T\boldsymbol{w}_k)$ and $S_0'(t)/N_0'(t)$ as $f(d_k)$ in (3) to calculate the error term e_k in the zlLMS algorithm and find the optimal weights. Note that $T'(t) = f(t)$ have a positive and bounded derivative in the range of $N(t)$ since the highest (normalized) input voltage of $N(t)$ cannot reach $3V_\pi/2$.

Optics EXPRESS

3.4. Implementation of zlLMS iterations

In our simulation, the sampling frequency is set to 100 THz. Additionally, we utilize an 11-tap system, where both $\boldsymbol{w}_k$ and $\boldsymbol{x}_k$ are 11-dimensional vectors for all k. The initial weight vector, $\boldsymbol{w}_1$, is initialized with a value of 1 at the first tap (first term) and 0 for all other taps.

The zlLMS algorithm is executed in the simulation as follows (see Fig. 2). First, we generate the ideal input signal $N_0(t)$, a Nyquist pulse train with a bias, and pass it through a raised cosine transfer function to obtain the ideal output signal $S_0(t)$. These two signals are then transmitted to implement the zlLMS algorithm. Simultaneously, $N_0(t)$ is passed through a 'Low-pass Filter' to produce the actual input signal $N(t)$ for the MZM. The signal $N(t)$ is subsequently processed by an 'Adaptive Filter' and then sent to the 'MZM' module to generate the actual output signal $S(t)$. Both $S(t)$ and $N(t)$ are then transmitted to the module executing the zlLMS algorithm. It is important to note that we assume all signal loss in this process–including losses in the 'MZM' and during wire transmission–occurs exclusively in the 'Low-pass Filter' module. As a result, the 'MZM' module is treated as a lossless signal converter characterized by a raised cosine transfer function.

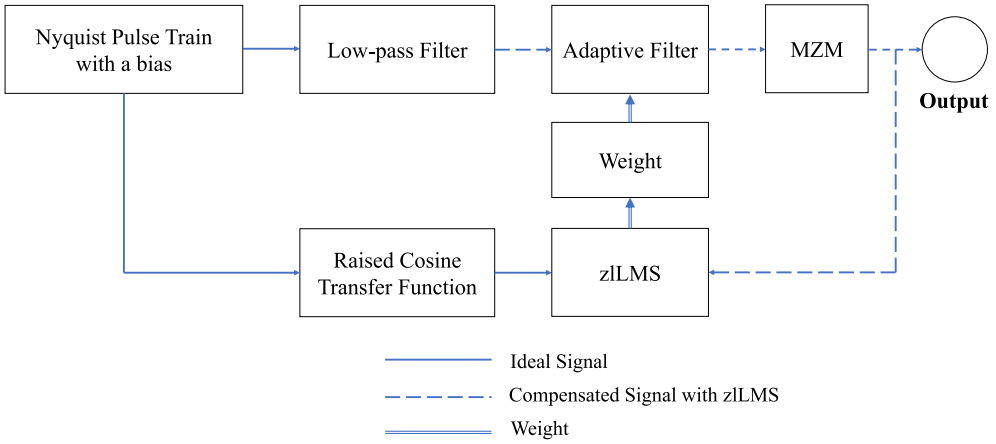

Fig. 2. Signal processing procedures in the simulation.

Then, the zlLMS algorithm is implemented based on (3). The sample of $\boldsymbol{x}_k$ is derived from the $N(t)$ signal, while the sample of $f(d_k)$ is obtained from the $S_0'(t)/N_0'(t)$ signal. Similarly, the sample of $f(\boldsymbol{x}_k^T\boldsymbol{w}_k)$ is derived from the $S'(t)/N'(t)$ signal. The 'zlLMS' module then updates the weights accordingly, sends the newly updated weights to the 'Adaptive Filter', and repeats the process iteratively to refine the weights.

3.5. Results and discussion

The simulation result of using zlLMS to approximate the ideal signal is shown in Fig. 3.

The amplitude difference between the ideal signal and the approximating signal is 1000 times smaller than the amplitude of the ideal signals. Thus, the two signals can be seen as identical from an engineering perspective.

To compare the performance of the two algorithms, we first use the traditional LMS to approximate the ideal output signal. That is, we use $S(t)$ as the signal, instead of $S'(t)/N'(t)$ proposed by zlLMS method, to calculate the error. With traditional LMS, the fit of the algorithm output signal to the ideal signal will become worse and worse, especially after 100 iterations, which forces the training iterations to stop. The trial signals in between 99th iteration and 101st

Optics EXPRESS

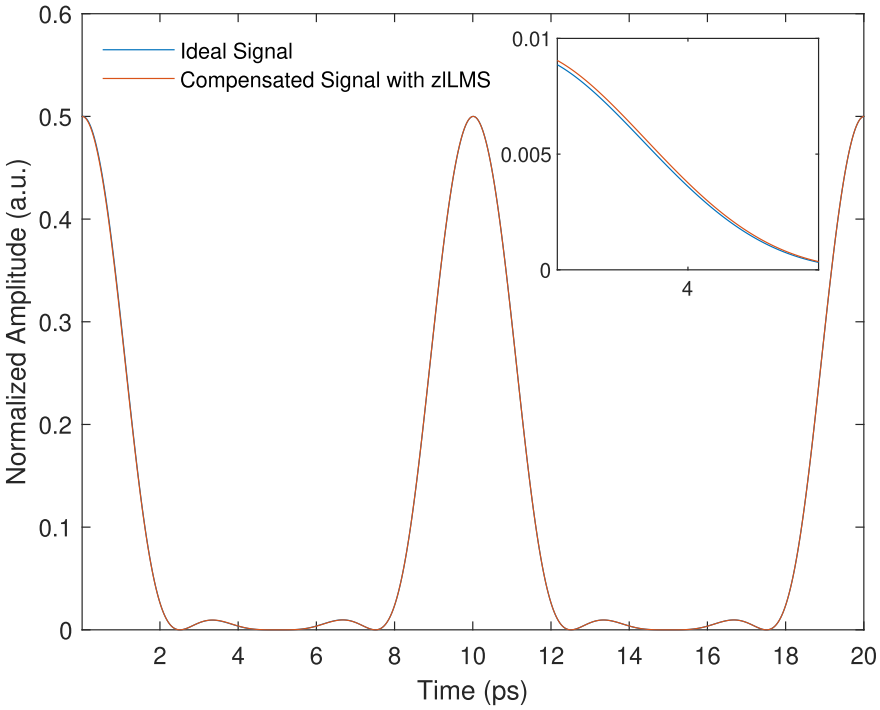

Fig. 3. Ideal MZM output signal $S_0(t)$ (blue) and actual signal after compensation $S(t)$ (orange) when using zlLMS to approximate $S_0(t) = T(N_0(t))$.

iteration are plotted and inspected in Fig. 4. Inside this figure, we can see that the iteration of signals is irregular.

When observing curves between 3 ps and 7 ps, the error between target signal and algorithm output signal increases after iterations of training by traditional LMS algorithm. Hence, all in all, zlLMS shows much better performance compared to the traditional LMS in the case study of MZM.

Optics EXPRESS

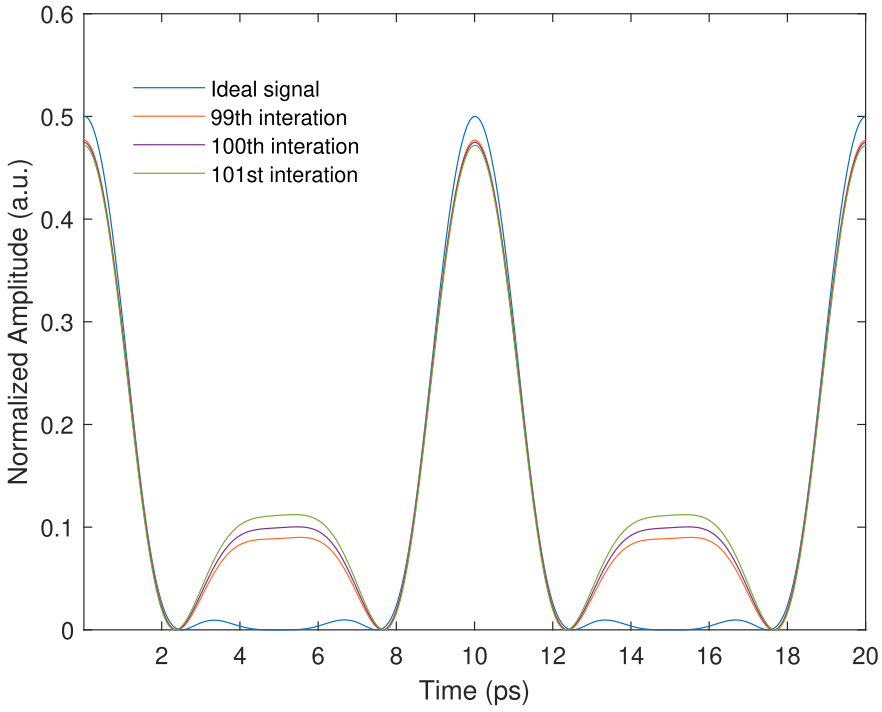

Fig. 4. Performance of the traditional LMS.

4. Application of zILMS in diode lasers

Another non-monotonic transfer function commonly implemented is the Lorentz function, which is a common transfer function in diode lasers. The standard form of the Lorentz function is given by:

$$L(x) = \frac{1}{\pi} \cdot \frac{\gamma}{(x - x_0)^2 + \gamma^2}, \tag{9}$$

where:

- x_0 is the location of the peak of the function.
- γ is the half-width at half-maximum, which describes the width of the peak.

In this simulation, the Lorentz function utilizes parameters $x_0 = 0$ and $\gamma = 6$. The simulation steps and assumptions in this Section are almost the same as in Section 3 and Fig. 2. The only difference is that the transfer function is replaced by the Lorentz function, and MZM is replaced by the diode lasers. The input signal to be processed is still $N_0(t)$—the Nyquist pulse train with a bias V_π generated in (7).

Corollary 4.1 *When the transfer function is replaced from the raised cosine function in (8) into the Lorentz function in (9), the function f between $N(t)$ and $S'(t)/N'(t)$ is still monotonic like Theorem 3.1. $S(t)$ is the actual output signal of the diode laser, and $N(t)$ is the actual input signal of the diode laser.*

Optics EXPRESS

Proof.

$$
\begin{aligned}
S(t) &= L\left(N(t)\right) \\
L(t) &= \frac{1}{\pi} \cdot \frac{6}{t^2+36} (\text{by setting } x_0 = 0 \text{ and } \gamma = 6) \\
S'(t) &= L'\left(N(t)\right) \cdot N'(t) \\
\frac{S'(t)}{N'(t)} &= L'\left(N(t)\right) \\
&= f\left(N(t)\right).
\end{aligned}
$$

We know that

$$L'(t) = -\frac{12t}{\pi(t^2+36)^2}$$

is monotonic in $[-2\sqrt{3}, 2\sqrt{3}]$. Moreover, the range of $N(t)$ lies strictly between $V_\pi/2$ and $3V_\pi/2$ (from Theorem 3.1). Hence, we can conclude that the function f (which is L') describing the relationship between $N(t)$ and $S'(t)/N'(t)$ is monotonic in the signal's operating range. □

Therefore, like the case of MZM, we can use $S'(t)/N'(t)$ as $f(\boldsymbol{x}_k^T \boldsymbol{w}_k)$ and $S_0'(t)/N_0'(t)$ as $f(d_k)$ in (3) to calculate the error term e_k in the zlLMS algorithm and find the optimal weights since $L'(t) = f(t)$ have a positive and bounded derivative in the range of $N(t)$. Results in Fig. 5 demonstrate that zlLMS can also be implemented with a Lorentz transfer function–an approach that poses significant challenges for traditional LMS algorithms. This outcome further confirms the broad applicability of zlLMS.

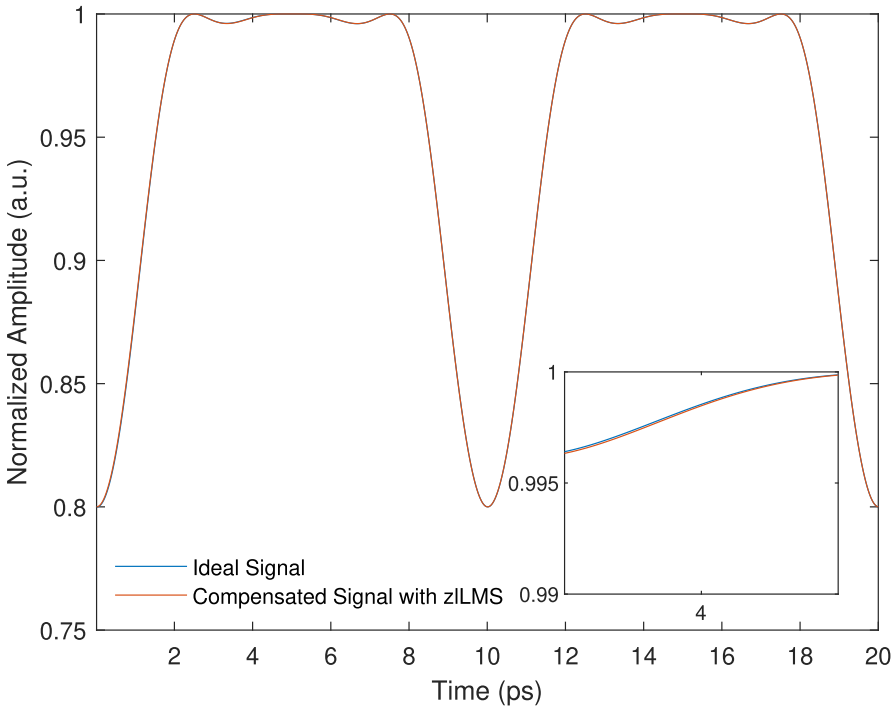

Fig. 5. Ideal diode lasers output signal $S_0(t)$ (blue) and actual signal after compensation $S(t)$ (orange) when using zlLMS to approximate $S_0(t) = L(N_0(t))$.

Optics EXPRESS

5. Conclusion

In this paper, we introduce the zlLMS algorithm as a novel improvement over the traditional LMS algorithm, specifically tailored to address the challenges posed by non-monotonic transfer functions such as the raised cosine function in MZMs and diode lasers with Lorentz transfer functions. zlLMS introduces a novel way to calculate the error term in LMS. We also prove the convergence of zlLMS. Through extensive simulations, we demonstrate that, compared to the traditional LMS algorithm, the zlLMS algorithm not only achieves effective convergence but also significantly minimizes the error between the desired and actual signals. Specifically, it reduces the amplitude difference by a factor of 1000 relative to the original input signal's amplitude in the MZM. This performance underscores the robustness and efficiency of the zlLMS algorithm in handling nonlinear and non-monotonic systems. Moreover, the generality of the zlLMS algorithm has been validated by its successful application to systems governed by different non-monotonic transfer functions, illustrating its potential for broad applicability in various engineering domains.

Future work will focus on experimental validation and further optimization of the zlLMS algorithm for real-world implementations. We are planning to test the algorithm and its future variants on hardware. The convergence speed of the LMS algorithm will also be an interesting objective that will be included in our future research. According to [4], the algorithm's convergence speed depends on the choice of the learning rate, where a higher learning rate accelerates convergence but may cause instability, while a lower learning rate enhances stability but slows down convergence. Additionally, the convergence speed is influenced by the statistical properties of the input signal. Various improvements, such as normalized LMS and variable step-size LMS, can optimize convergence performance under different conditions.

Funding. University of Cambridge.

Acknowledgment. We would like to express our sincere gratitude to the AXA Research Fund for the generous support provided through the AXA Chair for Internet of Everything at Koç University.

Disclosures. The authors declare no conflicts of interest.

Data availability. Data underlying the results presented in this paper are not publicly available at this time but may be obtained from the authors upon reasonable request.

References

1. A. Zaknich, "Adaptive filtering," in *Principles of Adaptive Filters and Self-learning Systems*, International Society for Optics and Photonics pp. 3–23 (2005).
2. B. Widrow and M. E. Hoff, "Adaptive switching circuits," in *Neurocomputing: Foundations of Research* (The MIT Press, 1988), pp. 123–134.
3. R. Kline, "An overview of twenty-five years of electrical and electronics engineering in the proceedings of the ieee, 1963-1987," Proc. IEEE **93**(12), 2170–2187 (2005).
4. P. S. R. Diniz, *Adaptive Filtering: Algorithms and Practical Implementation* (Springer, 2020).
5. S. Liu, K. Ma, and S. M. Goetz, "A robbins–monro sequence that can exploit prior information for faster convergence," arXiv (2024).
6. J. Radecki, Z. Zilic, and K. Radecka, "Echo cancellation in ip networks," in *45th Midwest Symposium on Circuits and Systems*, vol. 2 (IEEE, 2002), pp.II.
7. A. Mendiratta and D. Jha, "Adaptive noise cancelling for audio signals using least mean square algorithm," in *International conference on electronics, communication and instrumentation* (IEEE, 2014), pp.1–4.
8. M. L. Wang and F. Wu, "Structural system identification using least mean square (lms) adaptive technique," Soil Dynamics and Earthquake Engineering **14**(6), 409–418 (1995).
9. D. T. Slock, "On the convergence behavior of the lms and the normalized lms algorithms," IEEE Trans. Signal Process. **41**(9), 2811–2825 (1993).
10. S. A. U. Islam and D. S. Bernstein, "Recursive least squares for real-time implementation [lecture notes]," IEEE Control Syst. **39**(3), 82–85 (2019).
11. C. Safarian, T. Ogunfunmi, W. J. Kozacky, *et al.*, "Fpga implementation of lms-based fir adaptive filter for real time digital signal processing applications," *International Conference on Digital Signal Processing* (IEEE, 2015), pp.1251–1255.
12. A. S. Abdygalievich, A. Barlybayev, and K. B. Amanzholovich, "Quality evaluation fuzzy method of automated control systems on the lms example," IEEE Access **7**, 138000–138010 (2019).

Optics EXPRESS

13. A. Nicolson and K. K. Paliwal, "Deep learning for minimum mean-square error approaches to speech enhancement," Speech Communication **111**, 44–55 (2019).
14. Y. Fan, L. Zhang, K. Li, *et al.*," Deep learning-based emi and iemi classification in 5g-r high-speed rail wireless communications," in *99th Vehicular Technology Conference* (IEEE, 2024), pp.1–6.
15. Y. Sai, R. Jinxia, and L. Zhongxia, "Learning of neural networks based on weighted mean squares error function," in *Second International Symposium on Computational Intelligence and Design*, vol. 1 (IEEE, 2009), pp.241–244.
16. P. V. Kumar, K. Prabhu, and D. Das, "Block filtered-s least mean square algorithm for active control of non-linear noise systems," IET Signal Process. **4**(2), 168–180 (2010).
17. J. Chen, C. Richard, J.-C. M. Bermudez, *et al.*, "Non-stationary analysis of the convergence of the non-negative least-mean-square algorithm," in *21st European Signal Processing Conference* (IEEE, 2013), pp.1–5.
18. D. A. Cartes, L. R. Ray, and R. D. Collier, "Experimental evaluation of leaky least-mean-square algorithms for active noise reduction in communication headsets," J. Acoust. Soc. Am. **111**(4), 1758–1771 (2002).
19. A. Maalouf, I. Jubran, and D. Feldman, "Fast and accurate least-mean-squares solvers," Advances in Neural Information Processing Systems **32** (2019).
20. V. Vercesi, D. Onori, J. Davies, *et al.*, "Electronically synthesized nyquist pulses for photonic sampling of microwave signals," Opt. Express **25**(23), 29249–29259 (2017).
21. Z. Zhang, "Novel improvement for nonlinear compatibility of least mean square adaptive algorithm," arXiv (2022).
22. E. P. da Silva, R. Borkowski, S. Preußler, *et al.*, "Combined optical and electrical spectrum shaping for high-baud-rate nyquist-wdm transceivers," IEEE Photonics J. **8**(1), 1–11 (2016).
23. M. Nakazawa, T. Hirooka, and P. Ruan, "Ultrahigh-speed orthogonal tdm transmission with an optical nyquist pulse train," Opt. Express **20**(2), 1129–1140 (2012).
24. R. Pradeep and N. Vijayakumar, "Performance analysis of mach-zehnder modulator in radio over fiber systems," Int. J. Adv. Res. Eng. Technol **7**, 45–52 (2016).
25. X. Li, J. Yu, J. Xiao, *et al.*, "Pdm-qpsk vector signal generation by mzm-based optical carrier suppression and direct detection," Opt. Commun. **355**, 538–542 (2015).
26. Y. Fu, X. Zhang, B. Hraimel, *et al.*, "Mach-zehnder: a review of bias control techniques for mach-zehnder modulators in photonic analog links," IEEE Microwave **14**(7), 102–107 (2013).
27. M. Lu, M. Qin, W. Mu, *et al.*, "A hybrid gallium-nitride–silicon direct-injection universal power flow and quality control circuit with reduced magnetics," IEEE Trans. Ind. Electron. **71**(11), 14161–14174 (2024).
28. X. Li, J. Yu, J. Zhang, *et al.*, "Qam vector signal generation by optical carrier suppression and precoding techniques," IEEE Photon. Technol. Lett. **27**(18), 1977–1980 (2015).
29. G. P. Agrawal, *Fiber-optic Communication Systems* (John Wiley & Sons, 2012).
30. S. Liu, C. Lu, and G. He, "Distributed electric bicycle batteries for subway station energy management as a virtual power plant," Appl. Energy **370**, 123094 (2024).
31. R. S. Strichartz, *A guide to distribution theory and Fourier transforms* (World Scientific Publishing Company, 2003).