The Bias-Compensated Proportionate NLMS Algorithm With Sparse Penalty Constraint

ZHAN JIN\textsuperscript{1,2}, LONGXIANG GUO\textsuperscript{3}, AND YINGSONG LI\textsuperscript{1}, (Senior Member, IEEE)

\textsuperscript{1}College of Information and Communication Engineering, Harbin Engineering University, Harbin 150001, China
\textsuperscript{2}College of Communication and Electronic Engineering, Qiqihar University, Qiqihar 161006, China
\textsuperscript{3}Acoustic Science and Technology Laboratory, Harbin Engineering University, Harbin 150001, China

Corresponding author: Longxiang Guo (heu503@hrbeu.edu.cn)

This work was supported in part by the National Science Foundation of China under Grant 61571149, in part by the Project Plan of Science Foundation of Heilongjiang Province of China under Grant QC2018045, in part by the Fundamental Research Funds in Heilongjiang Provincial Universities under Grant 135309456 and Grant 135309342, in part by the Education Reform Project of Heilongjiang Province under Grant SGY20190718, in part by the Ph.D. Student Research and Innovation Fund of the Fundamental Research Funds for the Central Universities under Grant HEUGIP201811, in part by the National Key Research and Development Program of China under Grant 2016YFE111100, in part by the Key Research and Development Program of Heilongjiang under Grant GX17A016, and in part by the China Postdoctoral Science Foundation under Grant 2017M620918 and Grant 2019T120134.

ABSTRACT For compensating the bias caused by the noisy input which is always ignored by ordinary algorithms, two novel algorithms with zero-attraction (ZA) penalties are proposed in this paper. The first one constructs a bias-compensated term in the updating recursion of the zero-attraction proportionate normalized least mean square (PNLMS) algorithm which is named BC-ZA-PNLMS algorithm. The second one employs the bias-compensated term and the correntropy induced metric (CIM) constraint to renew the updating recursion of the PNLMS algorithm which is named BC-CIM-PNLMS algorithm. Both of these two algorithms are derived on the basis of unbiased criterion. Simulation examples are carried out, and the results indicate that the two newly developed unbiased algorithms outperform the related algorithms previously presented in other literatures for combating noisy input and measurement noises.

INDEX TERMS Unbiased adaptive filtering, zero-attraction (ZA), correntropy induced metric (CIM), PNLMS algorithm.

I. INTRODUCTION

Normalized least-mean-square (NLMS) algorithm is one of the popular adaptive filtering algorithms which are widely used in signal processing field, such as echo cancellation, system identification and linear prediction [1], [2]. Because of the simplicity and excellent performance of the NLMS, the family of the NLMS algorithms are extensively and deeply studied. It is well known that the channel responses of several practical systems such as the wireless communication system and echo path, are sparse, which means that most of the parameters to be estimated are insignificant and negligible, only small part of the parameters are remarkable and considerable [3]–[5]. For the purpose of taking advantages of the sparse characteristics in these systems, the proportionate type algorithms which obtain fast convergence speed by assigning an independent step-size to each coefficient have been developed [6]–[11]. Among these algorithms, the most popular one is the proportionate normalized least mean square (PNLMS) algorithm [6]. Inspired by the compressed sensing, a type of zero attraction algorithm has been developed by introducing a penalty term into the updating recursion of the original adaptive filter algorithm [12]–[19]. The role of the sparse penalty is to attract the small value coefficients to zero, consequently, faster convergence speed can be obtained than that of original algorithms and even than that of the proportionate type algorithms. Among the zero attraction algorithms, the zero attraction NLMS (ZA-NLMS) algorithm is the most famous one [14]. Besides the traditional zero-attractors obtained from different norm penalty, the correntropy induced metric (CIM) method has been considered to construct a zero-attractor in [20]–[26], which is to measure the similarity between the two different variables. If one variable is zero, the CIM tends to be $l_0$-norm. Thus, the CIM can be used to approximate the $l_0$-norm to exploit the sparsity of the system.

Although the two categories algorithms which fall into proportionate type one and zero attraction type one can get smaller steady state error and faster convergence, which only focus on the output noise and do not consider the input noise. However, the input noise caused by modeling and instrument

Received November 16, 2019, accepted December 24, 2019, date of publication December 30, 2019, date of current version January 8, 2020.

Digital Object Identifier 10.1109/ACCESS.2019.2962861
error is unavoidable, and the neglect of input noise may introduce estimation bias which decreases the accuracy accordingly. Using higher order model and increasing the signal-to-noise ratio (SNR) can mitigate the adverse impact of input noise in some degree but can never be eliminated completely [27]. To overcome this drawback, some improved algorithms for unbiased estimation have been investigated [28]–[39]. The total least squares (TLS) and the bias-compensated least squares (TLS) methods can handle the bias issue but suffer to computational inefficiency [40]. Therefore, the bias-compensated algorithms under least mean square criterion obtain more attention relay on their easy implementation and low complexity. In [28], the unbiased criterion for steady state is utilized to provide a simple approach of bias compensation via the statistical property of the input noise. Since then, unbiased algorithms deriving from the unbiased criterion have been developed. Among the existing ideas, the classical adaptive filter algorithms like the NLMS [1], [2], [41], PNLMs [6], constrained least-mean-square (CLMS) [42], least mean fourth (LMF) [43], [44], normalized subband adaptive filter (NSAF) [45], [46], proportionate least-mean-square/fourth (PLMS/F) [47]–[49], affine-projection-like adaptive filter (NSAF) [45], [46], proportionate least-mean-square (FONLMS) [55]–[58] are improved by reducing the impacts of the input noise via appending a bias-compensated term into the updating equation of the primal algorithms. Furthermore, the stabilization for the bias-compensated NLMS (BC-NLMS) algorithm based on unbiased criterion was also researched carefully in [59].

Inspired by the idea of bias-compensation and motivated by utilizing the sparsity of system impulse response, the novel algorithm is proposed to utilize both of the bias-compensated and sparsity to construct new updating recursion of the PNLMs algorithm in this paper. The bias-compensated term is derived from the unbiased criterion and incorporated with zero-attraction scheme. Besides the traditional zero attractor vector with $\lambda$, which is noise-free vector. The desired signal is described as

$$d(k) = v^T(k)w(k) + n(k). \quad (1)$$

$n(k)$ represents the noise signal from the environment. The system estimation error is

$$e(k) = d(k) - v^T(k)\hat{w}(k) \quad (2)$$

with an estimation $\hat{w}(k)$.

The PNLMs’s update recursion is

$$\hat{w}(k+1) = \hat{w}(k) + \mu \frac{v(k)Q(k)e(k)}{v^T(k)Q(k)v(k) + \epsilon} \quad (3)$$

with a step-size of $\mu$, where $\epsilon > 0$ is to prevent the denominator from zero. $Q(k)$ is a weight assignment matrix which assigns individual step-size to each tap corresponding to its magnitude, and it is described as

$$Q(k) = \text{diag} \{q_1(k), q_2(k), \ldots, q_L(k)\}. \quad (4)$$

The element $q_i$ in the matrix $Q(k)$ is

$$q_i(k) = \frac{\alpha_i(k)}{\sum_{i=1}^{L} \alpha_i(k)}, \quad 1 \leq i \leq L, \quad (5)$$

where

$$\alpha_i(k) = \max \left\{\rho\max \left\{|\hat{w}_1(k)|, |\hat{w}_2(k)|, \ldots, |\hat{w}_L(k)|\right\}, \left|\hat{w}_i(k)\right|\right\}. \quad (6)$$

The parameter $\rho$ in (6) is used to keep the iteration going on when $|\hat{w}_i(k)|$ is much smaller than the previous item, and its value is always $\frac{1}{L} \sim \frac{5}{L}$. The parameter $\delta$ in (6) is a small constant to promote the iteration running at beginning when all the coefficients are zeros.

The PNLMs can achieve fast convergence speed at initial stage of iteration which benefits from the step-size individual assignment of each coefficient. However, the advantage cannot be maintained at the later stage, which suffers from the slow convergence speed of the small value coefficients that are assigned small step-sizes. To accelerate the convergent of the small coefficients, the $l_1$-norm regularization constraint is introduced to the optimization of the PNLMs algorithm to force the small coefficients to approach zeros rapidly. Then, the resulting updating recursion of the ZA-PNLMs algorithm is obtained as

$$\hat{w}(k+1) = \hat{w}(k) + \mu \frac{v(k)Q(k)e(k)}{v^T(k)Q(k)v(k) + \epsilon} - \gamma \text{sgn}(\hat{w}(k)). \quad (7)$$

It is observed that the first two terms of the iteration are the same as those of the PNLMs algorithm, and the last term of the sign function with a zero attraction strength control parameter $\gamma$ is the constructed zero attractor.

**II. REVIEW OF THE PNLMs ALGORITHM WITH SPARSE PENALTY CONSTRAINT**

**A. THE PNLMs ALGORITHM WITH ZERO ATTRACTOR**

Consider a finite impulse response (FIR) system with $L$ taps, and the weight vector $w(k)$ denotes a column-vector with $L$ rows. The system input signal is $v(k) = \{v(k), v(k-1), v(k-2), \ldots, v(k-(L+1))\}^T$, which is noise-free vector. The desired signal is described as

$$d(k) = v^T(k)w(k) + n(k). \quad (1)$$

$n(k)$ represents the noise signal from the environment. The system estimation error is

$$e(k) = d(k) - v^T(k)\hat{w}(k) \quad (2)$$

with an estimation $\hat{w}(k)$.

The PNLMs’s update recursion is

$$\hat{w}(k+1) = \hat{w}(k) + \mu \frac{v(k)Q(k)e(k)}{v^T(k)Q(k)v(k) + \epsilon} \quad (3)$$

with a step-size of $\mu$, where $\epsilon > 0$ is to prevent the denominator from zero. $Q(k)$ is a weight assignment matrix which assigns individual step-size to each tap corresponding to its magnitude, and it is described as

$$Q(k) = \text{diag} \{q_1(k), q_2(k), \ldots, q_L(k)\}. \quad (4)$$

The element $q_i$ in the matrix $Q(k)$ is

$$q_i(k) = \frac{\alpha_i(k)}{\sum_{i=1}^{L} \alpha_i(k)}, \quad 1 \leq i \leq L, \quad (5)$$

where

$$\alpha_i(k) = \max \left\{\rho\max \left\{|\hat{w}_1(k)|, |\hat{w}_2(k)|, \ldots, |\hat{w}_L(k)|\right\}, \left|\hat{w}_i(k)\right|\right\}. \quad (6)$$

The parameter $\rho$ in (6) is used to keep the iteration going on when $|\hat{w}_i(k)|$ is much smaller than the previous item, and its value is always $\frac{1}{L} \sim \frac{5}{L}$. The parameter $\delta$ in (6) is a small constant to promote the iteration running at beginning when all the coefficients are zeros.

The PNLMs can achieve fast convergence speed at initial stage of iteration which benefits from the step-size individual assignment of each coefficient. However, the advantage cannot be maintained at the later stage, which suffers from the slow convergence speed of the small value coefficients that are assigned small step-sizes. To accelerate the convergent of the small coefficients, the $l_1$-norm regularization constraint is introduced to the optimization of the PNLMs algorithm to force the small coefficients to approach zeros rapidly. Then, the resulting updating recursion of the ZA-PNLMs algorithm is obtained as

$$\hat{w}(k+1) = \hat{w}(k) + \mu \frac{v(k)Q(k)e(k)}{v^T(k)Q(k)v(k) + \epsilon} - \gamma \text{sgn}(\hat{w}(k)). \quad (7)$$

It is observed that the first two terms of the iteration are the same as those of the PNLMs algorithm, and the last term of the sign function with a zero attraction strength control parameter $\gamma$ is the constructed zero attractor.

**B. REVIEW OF THE CIM**

Recently, the CIM is introduced to the adaptive filter which acts as an approximation of $l_0$-norm [20]–[26]. Compared with the global measurement of the mean square error (MSE), the correntropy focuses on local statistics. To describe the similarity of two random vectors, the correntropy of two vectors $W$ and $Z$ is defined as

$$V(W, Z) = \frac{1}{M} \sum_{j=1}^{M} k(w_j, z_j). \quad (8)$$
where $\kappa(.)$ represents the kernel used to satisfy the Mercer’s theorem. Among the various kernel functions, the Gaussian kernel is the most popular one, and it is described as

$$\kappa(w, z) = \frac{1}{\sqrt{2\pi} \sigma} \exp\left(-\frac{||w - z||^2}{2\sigma^2}\right),$$

(9)

where the kernel width is $\sigma$, and its size is equal to the range of a special set. The CIM is

$$\text{CIM}(W, Z) = \sqrt{\kappa(0)} - V(W, Z),$$

(10)

which is a nonlinear metric derived from the correntropy. Herein, make $W$ to $\hat{w}(k)$ and make $Z$ to zero, then, the degree to which the coefficient approaches zero can be obtained. This is lucky coincidence to the role of the zero attractor. Thus, the CIM can be chosen to be a sparse penalty to exploit the sparsity of the original adaptive algorithm. To simplify the expression, the CIM usually takes squared rather than the form of square root which is shown as

$$\text{CIM}^2(\hat{w}(k), 0) = \frac{1}{L \sigma \sqrt{2\pi}} \sum_{i=1}^{L} \left(1 - \exp\left(-\frac{\hat{w}_i^2(k)}{2\sigma^2}\right)\right).$$

(11)

Taking the derivation of $\text{CIM}^2(\hat{w}(k), 0)$ with respect to $\hat{w}(k)$, we can get

$$\frac{\partial \text{CIM}^2(\hat{w}(k), 0)}{\partial \hat{w}(k)} = \frac{1}{L \sigma^3 \sqrt{2\pi}} \hat{w}(k) \exp\left(-\frac{\hat{w}_i^2(k)}{2\sigma^2}\right).$$

(12)

The updating recursion of the PNLMS algorithm with CIM penalty can be described as

$$\hat{w}(k + 1) = \hat{w}(k) + \mu \frac{x^{T}(k)Q(k)e(k)}{x^{T}(k)Q(k)x(k) + \varepsilon} \hat{w}(k) + \varepsilon \gamma_{\text{CIM}} \frac{1}{L \sigma^3 \sqrt{2\pi}} \hat{w}(k) \exp\left(-\frac{\hat{w}_i^2(k)}{2\sigma^2}\right)$$

(13)

Comparing to (7), we can find that $\gamma_{\text{CIM}} \frac{1}{L \sigma^3 \sqrt{2\pi}} \hat{w}(k) \exp\left(-\frac{\hat{w}_i^2(k)}{2\sigma^2}\right)$ is regarded as a zero attractor, and the algorithm with CIM constraint is also can be considered as a zero attraction type algorithm.

### III. THE BIAS-COMPENSATED ZERO ATTRACTION ALGORITHMS WITH INPUT NOISE

The algorithms mentioned in the previous section can both exploit the system’s sparse characteristic and further accelerate the convergence rate, nevertheless, the input noise of the filter which is unavoidable in practice is not taken into account. To mitigate the adverse impacts of the noisy input, an unbiased term is incorporated into the updating equation of the traditional algorithm.

Input signal containing noises is described as

$$x(k) = v(k) + n_{in}(k)$$

(14)

with $n_{in}(k) = [n_{in}(k), n_{in}(k-1), n_{in}(k-2) \ldots, n_{in}(k-L+1)]^T$ is the input noise (see Fig.1) that is assumed to be white Gaussian noise (WGN) and independent of $v(k)$. The resulting error signal will be replaced by

$$e_{in}(k) = d(k) - x^{T}(k)\hat{w}(k) = d(k) - (v(k) + n_{in}(k))^T \hat{w}(k) = e(k) - n_{in}^T(k)\hat{w}(k).$$

(15)

As can be seen from (15), the considering input noise causes additional bias compared with $e(k)$ which is expressed in (2). Therefore, the updating equation of the $\hat{w}(k)$ shown in (7) is improper. To make the estimation accurate, a bias-compensation term is introduced and is written as $W_{BC}$. Then, the new updating equation of the BC-ZA-PNLMS algorithm turns to

$$\hat{w}(k + 1) = \hat{w}(k) + \mu \frac{x(k)Q(k)e(k)}{x^{T}(k)Q(k)x(k) + \varepsilon} \hat{w}(k) + \gamma_{\text{BC}} sgn(\hat{w}(k)) + W_{BC}(k).$$

(16)

In order to figure out the concrete expression of $W_{BC}(k)$, the unbiased criterion proposed in [28] is employed which is described as

$$E\left[w^*(k + 1)|x(k)\right] = 0$$

whenever $E\left[w^*(k)|x(k)\right] = 0$,  

(17)

where

$$w^* = w - \hat{w}.$$  

(18)

The unbiased criterion is only satisfied in steady state, and by this time the $w^*$ tends to zero. According to (16) and (18), we can get

$$w^*(k + 1) = w^*(k) - \mu \frac{x(k)Q(k)e_{in}(k)}{x^{T}(k)Q(k)x(k) + \varepsilon} + \gamma_{\text{BC}} sgn(\hat{w}(k)) - W_{BC}(k).$$

(19)

Taking the conditional expectation of both sides of (19) to get the same form in (17). It is found that the third term on the right-hand side of the equation can be omitted in the following operation, avoiding introducing extra bias and being unnecessary for input noise elimination [35]. Then, the result of taking conditional expectation will be

$$E\left[w^*(k + 1)|x(k)\right] = E\left[w^*(k)|x(k)\right]$$

$$- \mu E\left[\frac{x(k)Q(k)e_{in}(k)}{x^{T}(k)Q(k)x(k) + \varepsilon} | x(k)\right]$$

$$- E\left[W_{BC}(k)|x(k)\right].$$

(20)
In steady state, the equation will become

\[
E \left[ W_{BC}(k) | x(k) \right] = -\mu E \left[ \frac{x(k)Q(k)e_{in}(k)}{x^T(k)Q(k)x(k) + \epsilon} \right] x(k) ,
\]

(21)

where

\[
E \left[ \frac{x(k)Q(k)e_{in}(k)}{x^T(k)Q(k)x(k) + \epsilon} \right] x(k) = \frac{Q(k)E \left[ x(k)e_{in}(k) | x(k) \right]}{x^T(k)Q(k)x(k) + \epsilon} .
\]

(22)

Substituting (14) and (15) into (22), we will obtain

\[
E \left[ x(k)e_{in}(k) | x(k) \right] = E \left[ \left( \mathbf{v}(k) + n(k) \right) e(k) \right] x(k) = E \left[ \mathbf{v}(k)e(k) + n(k)e(k) \right] x(k)
\]

\[
- \mathbf{n}_m^T(k) \mathbf{w}(k) x(k) - E \left[ \mathbf{n}_m(k) n_m^T(k) \mathbf{w}(k) | x(k) \right] .
\]

(23)

To make the calculation tractable, some assumptions are given. First of all, the input noise \( n(k) \) is independent of \( \mathbf{v}(k) \) with mean and variance of zero and \( \delta_n^2 \), respectively. Next, the observation noise \( n(k) \) is independent of \( \mathbf{n}_m(k) \) and \( \mathbf{v}(k) \) with \( N(0,1) \) distribution. In addition, both of \( x(k) \) and \( \mathbf{v}(k) \) are uncorrelated to \( \mathbf{w}^*(k) \). Based on these assumptions, (23) can be rewritten as

\[
E \left[ x(k)e_{in}(k) | x(k) \right] = E \left[ \mathbf{v}(k)e(k) | x(k) \right] + E \left[ n_m(k)e(k) | x(k) \right]
\]

\[
+ E \left[ n_m(k) n_m^T(k) | x(k) \right] .
\]

(24)

where

\[
E \left[ \mathbf{v}(k)e(k) | x(k) \right] = 0 ,
\]

(25)

\[
E \left[ n_m(k)e(k) | x(k) \right] = 0 ,
\]

(26)

\[
E \left[ \mathbf{v}(k)n_m(k)^T | x(k) \right] = 0 ,
\]

(27)

\[
E \left[ n_m(k)n_m^T(k) | x(k) \right] = \delta_n^2 E \left[ \mathbf{w}(k) | x(k) \right] .
\]

(28)

Substituting the result into (21), yields

\[
E \left[ W_{BC}(k) | x(k) \right] = \mu \delta_n^2 \frac{Q(k)\mathbf{w}(k)}{x^T(k)Q(k)x(k) + \epsilon} .
\]

(29)

Then, the bias-compensation term can be calculated via (29)

\[
W_{BC}(k) = \mu \delta_n^2 \frac{Q(k)\mathbf{w}(k)}{x^T(k)Q(k)x(k) + \epsilon} .
\]

(30)

Substituting this result into (16), the updating equation of the bias-compensation ZA-PNLMS (BC-ZA-PNLMS) algorithm is obtained as

\[
\hat{w}(k+1) = \hat{w}(k) + \mu \frac{x(k)Q(k)e_{in}(k)}{x^T(k)Q(k)x(k) + \epsilon} \hat{w}(k) + \mu \delta_n^2 \frac{Q(k)\mathbf{w}(k)}{x^T(k)Q(k)x(k) + \epsilon} - \gamma \text{sgn}(\mathbf{w}(k)) .
\]

(31)

From (31), it can be found that \( \hat{w}(k+1) = \hat{w}(k) + \mu \frac{x(k)Q(k)e_{in}(k)}{x^T(k)Q(k)x(k) + \epsilon} \hat{w}(k) \) is the regular expression of PNLMS.
The key parameter $\delta^2_{in}$ in (31) and (32) is the variance of the input noise which usually cannot be obtained directly in practice. So the methods of $\delta^2_{in}$ estimation are investigated in several literatures and are summarized in [59]. Herein, $\delta^2_{in}$ is replaced by $\delta^2_{in}(k)$ which is the instant value at $k$th iteration, and $\delta^2_{in}(k)$ can be estimated by following equations based on time ergodicity.

$$\delta^2_{in}(k) = \frac{1}{L\hat{\sigma}^2_w(k) + \eta} \left[ \frac{L}{\hat{\delta}^2_{in}(k)} + \frac{L}{x^T(k)x(k)} \right],$$

where

$$\delta^2_{in}(k) = (1 - f)\frac{1}{L}\hat{w}^T(k)\hat{w}(k) + f\delta^2_{in}(k - 1).$$

$$\delta^2_{in}(k) = (1 - f)e^2_{in}(k) + f\delta^2_{in}(k - 1).$$

The parameter $f$ in (34) and (35) is a forgetting parameter, and $\eta$ is the pre-known input-output noise-ratio.

From the derivation of the two algorithms, we can conclude that the BC-ZA-PNLMS and BC-CIM-PNLMS have both of fast convergence rate and low estimation error which benefits from the inserting of the bias-compensated term and sparse penalty constraint. However, the computational complexity is a little bit higher than the original algorithm. The recursion of the PNLMS requires $2L^2 + 2L - 1$ additions, $2L^2 + 3L + 2$ multiplications and $L + 1$ divisions. The BC-ZA-PNLMS requires $3L^2 + 3L - 1$ additions, $3L^2 + 5L + 3$ multiplications and $L + 1$ divisions. The BC-CIM-PNLMS requires $3L^2 + 4L - 2$ additions, $3L^2 + 7L + 7$ multiplications, $L + 3$ divisions and $L$ exponentiation.

**IV. SIMULATION AND RESULT ANALYSIS**

To investigate the behaviors of the developed algorithms, several numerical examples will be carried out in this part. The behaviors of the BC-ZA-PNLMS and BC-CIM-PNLMS algorithms with different sparsities are compared with those of the original ZA-PNLMS algorithm and other bias-compensated algorithms. Then, the performance of the two new algorithms are studied with different SNRs, different input signals and different channel responses, respectively. Besides, the parameter $\gamma$ in BC-ZA-PNLMS and $\sigma$ in BC-CIM-PNLMS are analyzed. At last, the directly comparisons between the real channel response and the estimation result by the proposed methods are made. All of the numerical examples are conducted by 4000 trials for 200 Monte runs to get an average MSD defined by

$$\text{MSD} = E \left[ \left| w - \hat{w}(k) \right|^2 \right].$$

**Example 1:** The estimation behaviors of BC-ZA-PNLMS and BC-CIM-PNLMS with different sparsities of 1, 4 and 8 are studied compared with the BC-ZA-NLMS, BC-NLMS, ZA-PNLMS, BC-PNLMS algorithms. Herein, the sparsity
(represented by $S$) denotes the number of the nonzero coefficient and other coefficients are assumed to be zero. The length of the estimated channel is 64. The input signal, input noise and output noise are Gaussian white signal with distribution of $N(0, 1)$. The forgetting factors for all the algorithms involved in simulation are 0.6. To obtain the same initial convergence rate, the step-sizes for the BC-NLMS, BC-ZA-NLMS, ZA-PNLMS are set as 0.4, and the step-sizes for the BC-PNLMS, BC-ZA-PNLMS, BC-CIM-PNLMS are 0.12. The parameter $\varepsilon$ in all the algorithms are set to 0.001. The $\gamma$ and $\gamma_{CIM}$ in (31) and (32) are 0.0001 and 0.00001, while both of the corresponding parameters in other ZA algorithms such as the BC-ZA-NLMS and ZA-PNLMS are 0.0001. All of the parameters $\rho$ and $\delta$ in the proportionate algorithms are set to $\frac{5}{6}$ and 0.1, respectively. $\sigma$ is 0.007. $\eta$ is 0.0001. The learning lines with sparsities of 1, 4 and 8 are shown in Figure 2.

We can conclude that the BC-ZA-PNLMS and BC-CIM-PNLMS algorithms have superior performance respect with estimation precision under different sparsities. In general, the algorithms with bias-compensation term behave better than those algorithms without bias-compensation term which is shown in Figure 2(b) and Figure 2(c). However, if the channel response is extremely sparse, the proportional algorithms show their superiority comparing with the non-proportional algorithms which is shown in Figure 2(a).

**Example 2:** The performance of the two developed algorithms for different SNRs are studied. The experimental environment and parameter settings are the same as those in Example 1. Herein, the SNR denotes the input signal variance and input noise ratio variance whose values are 0dB, 10dB, 20dB and 30dB, respectively. The MSD lines of the BC-ZA-PNLMS and BC-CIM-PNLMS algorithms are shown in Figure 3. It is evident from the result that the lower the SNR is, the worse the algorithm performs.

**Example 3:** The behaviors of BC-ZA-PNLMS and BC-CIM-PNLMS with different input signals are investigated. The input signals include WGN signal with distribution of $N(0, 1)$, colored signal that is generated by WGN through a first order filter with a pole of 0.8, and the speech signal which lasts 8 seconds and is sampled by 8kHz, respectively. The SNR is set to 10dB and the sparsity is 4. The step-sizes of BC-ZA-NLMS and BC-NLMS are 0.4, and the step-sizes for other algorithms are 0.12. $f$ is 0.1. The estimation behaviors of the algorithms with WGN, colored signal and speech signal are shown in Figure 4.

As shown from the results, the developed algorithms can still maintain the superior performance for dealing with the WGN, colored signal and speech signal. However, the learning line of speech signal which lasts 8 seconds is not smooth since it is time-varying. As a result, the results obtained from the speech signal do not look smooth enough [30], [36], [39], [59].

**Example 4:** For the sake of verifying the stability of the developed algorithms, different channel responses are considered and simulated. An echo path which is shown in Figure 5(a) has 256 taps. The performance is shown in Figure 5(b). The underwater communication channel shown in Figure 6(a) with 222 taps is considered [60]. The performance is shown in Figure 6(b). The simulation result validates the efficiency of the proposed algorithms for different applications.
Example 5: In this example, the values of the key parameters in the BC-ZA-PNLMS and BC-CIM-PNLMS algorithms are discussed. The appropriate value of zero attraction parameter $\gamma$ in the BC-ZA-PNLMS algorithm is discussed, and the result is shown in Figure 7. It is observed that $\gamma = 0.0001$ is the best choice. In addition, the value of kernel width $\sigma$ in the BC-CIM-PNLMS algorithm is discussed, and the result is shown in Figure 8. As can be seen from the result, both of the values $\sigma = 0.1$ and $\sigma = 1$ are the best ones.

Example 6: To illustrate the performance of the proposed algorithms in channel estimation more intuitively, the real response which is shown in Figure 6(a) is compared with the estimation one. Consider that the input signal is WGN and the SNR is 0dB which conforms the SNR of underwater communication channel. $\gamma$ and $\gamma_{\text{CIM}}$ are 0.000001, and $\sigma$ is 0.7. The comparison results are shown in Figure 9.

From the simulation results, we can find that both of the two novel algorithms can achieve good performance in estimating and tracking even at low SNR and complex environment.

V. CONCLUSION

The PNLMS algorithms with zero attraction scheme and bias-compensated term are developed in this paper. The bias-compensation term reduces the adverse effect of the noisy input which is always not considered by the traditional algorithms. The zero attractors incorporated into the PNLMS can accelerate the returning to zeros for the close-to-zero coefficients. The BC-ZA-PNLMS and BC-CIM-PNLMS algorithms have superior performance on both estimation accuracy and convergence speed. The derivations of the two algorithms are proposed and discussed on the basis of unbiased criterion, and the behaviors of the new algorithms are studied in a comprehensive way. Simulation results of numerical examples demonstrate the validity of bias compensation for suppressing the noisy input and the excellent performance of the two developed algorithms.

REFERENCES

[1] P. S. R. Diniz, Adaptive Filtering: Algorithms and Practical Implementation, 4th ed. New York, NY, USA: Springer, 2013.
[2] A. H. Sayed, Fundamentals of Adaptive Filtering. New York, NY, USA, Wiley, 2003.
[3] J. Benesty, T. Gaensler, D. R. Morgan, M. M. Sondhi, and S. L. Gay, Advances in Network and Acoustic Echo Cancellation. Berlin, Germany: Springer, 2001.
[4] L. Vuoikko, V.-M. Kolmonen, J. Salo, and P. Vainikainen, “Measurement of large-scale cluster power characteristics for geometric channel models,” IEEE Trans. Antennas Propag., vol. 55, no. 11, pp. 3361–3365, Nov. 2007.
[5] W. Shi, Y. Li, L. Zhao, and X. Liu, “Controllable sparse antenna array for adaptive beamforming,” IEEE Access, vol. 7, pp. 6412–6423, 2019.
[6] D. L. Duttweiler, “Proportionate normalized least-mean-squares adaptation in echo cancelers,” IEEE Trans. Speech Audio Process., vol. 8, no. 5, pp. 508–518, Sep. 2000.
[7] J. Benesty and S. L. Gay, “An improved PNLMS algorithm,” in Proc. IEEE Int. Conf. Acoust. Speech Signal Process. Speech Audio Process., vol. 12, no. 3, pp. 181–184, Mar. 2005.
[8] H. Deng and M. Doroslovacki, “Improving convergence of the PNLMS algorithm for sparse impulse response identification,” IEEE Signal Process. Lett., vol. 12, no. 3, pp. 181–184, Mar. 2005.
[9] Y. Dong and H. Zhao, “A new proportionate least-mean-square algorithm for high measurement noise,” in Proc. IEEE Int. Conf. Signal Process., Commun. Comput. (ICSPCC), Ningbo, China, Sep. 2015.
[10] J. Benesty and Y. Huang, “The LMS, PNLMS, and exponentiated gradient algorithms,” in Proc. Eur. Signal Process. Conf., Sep. 2004, pp. 721–724.
B. Chen, L. Xing, H. Zhao, N. Zheng, and J. C. Príncipe, “Generalized correntropy criterion algorithm for cluster-sparse system identification,” IEEE Trans. Signal Process., vol. 64, no. 13, pp. 3376–3387, Jul. 2016.

J. Zhao, H. Zhang, G. Wang, and X. Liao, “Enhancing sparsity by reweighted 11-minimization,” J. Fourier Anal. Appl., vol. 15, nos. 5–6, pp. 877–905, 2008.

J. Jin, Y. Gu, and S. Mei, “Inherent decorrelating and least perturbation properties of the normalized subband adaptive filter,” IEEE Trans. Signal Process., vol. 54, no. 11, pp. 4475–4480, Nov. 2006.

Y. Li, Y. Wang, and T. Jiang, “Sparse-aware set-membership NLMS algorithms and their application for sparse channel estimation and echo cancelation,” AEU-Int. J. Electron. Commun., vol. 70, no. 7, pp. 895–902, Jul. 2016.

E. J. Candès, M. B. Wakin, and S. P. Boyd, “Enhancing sparsity by reweighted 11-minimization,” J. Fourier Anal. Appl., vol. 15, nos. 5–6, pp. 877–905, 2008.

Z. Jin, Y. Gu, and A. O. Hero, “Sparse LMS for system identification,” IEEE Trans. Signal Process., vol. 64, no. 13, pp. 3376–3387, Jul. 2016.

J. Jin, Y. Gu, and S. Mei, “A stochastic gradient approach on compressive sensing signal reconstruction based on adaptive filtering framework,” IEEE J. Sel. Topics Signal Process., vol. 4, no. 2, pp. 409–420, Apr. 2010.

W. Liu, P. P. Pokharel, and J. C. Príncipe, “Correntropy: Properties and applications in non-Gaussian signal processing,” IEEE Trans. Signal Process., vol. 55, no. 11, pp. 5286–5298, Nov. 2007.

B. Chen, L. Xing, J. Liang, N. Zheng, and J. C. Príncipe, “Steady-state mean-square error analysis for adaptive filtering under the maximum correntropy criterion,” IEEE Trans. Signal Process., vol. 60, no. 5, pp. 2223–2235, May 2012.

Y. Li, Y. Wang, and T. Jiang, “Norm-adaption penalized least mean square/fourth order algorithm for sparse channel estimation,” Signal Process., vol. 128, pp. 243–251, Nov. 2016.

S. Seth and J. C. Príncipe, “Compressed signal reconstruction using the correntropy induced metric,” in Proc. IEEE Int. Conf. Acoust., Speech Signal Process., Las Vegas, NV, USA, Apr. 2008, pp. 3843–3846.

A. Singh and J. C. Príncipe, “Using Correntropy as a cost function in linear adaptive filters,” in Proc. Int. Joint Conf. Neural Netw., Atlanta, GA, USA, Jun. 2009, pp. 2950–2955.

B. Chen, L. Xing, H. Zhao, N. Zheng, and J. C. Príncipe, “Generalized correntropy for robust adaptive filtering,” IEEE Trans. Signal Process., vol. 64, no. 13, pp. 3376–3387, Jul. 2016.

Y. Li, Y. Wang, and T. Jiang, “Sparse-averaged adaptive filtering algorithms using the maximum correntropy criterion for system identification with noisy input,” Signal Process., vol. 152, pp. 160–164, Nov. 2018.

S. Seth and J. C. Príncipe, “Compressed signal reconstruction using the correntropy induced metric,” in Proc. IEEE Int. Conf. Acoust., Speech Signal Process., Las Vegas, NV, USA, Apr. 2008, pp. 3843–3846.

A. Singh and J. C. Príncipe, “Using Correntropy as a cost function in linear adaptive filters,” in Proc. Int. Joint Conf. Neural Netw., Atlanta, GA, USA, Jun. 2009, pp. 2950–2955.
ZHAN JIN received the B.S. degree in electrical and information engineering from Heilongjiang University, Harbin, China, in 2005, and the M.S. degree in information and communication engineering from Harbin Engineering University, Harbin, in 2009, where she is currently pursuing the Ph.D. degree in information and communication engineering. She has been working with the College of Communication and Electronic Engineering, Qiqihar University, Qiqihar, China, since 2009. She is currently an Associate Professor. Her research interests include signal processing and sparse adaptive filtering.

LONGXIANG GUO received the B.E., M.S., and Ph.D. degrees from the College of Underwater Acoustic Engineering (UAE), Harbin Engineering University (HEU), in 1998, 2001, and 2006, respectively. From 2013 to 2015, he was a Visiting Scholar with the Laboratories of Image, Signal Processing and Acoustics, Université Libre de Bruxelles (ULB), Brussels, Belgium. He is currently an Associate Professor with the College of Underwater Acoustic Engineering, Harbin Engineering University. His main interests cover acoustic signal processing, image processing, and acoustical oceanography. He is currently involved in research related to acoustic array processing and sensor information fusion. His awards and honors include the Second-Class Reward of Heilongjiang Science and Technology, and the Second-Class Reward of the China State Shipbuilding Company, Ltd., (CSSC) Science and Technology Progress Award.

YINGSONG LI (Senior Member, IEEE) received the B.S. degree in electrical and information engineering, in 2006, the M.S. degree in electromagnetic field and microwave technology from Harbin Engineering University, China, in 2011, and the joint Ph.D. degree from the Kochi University of Technology (KUT), Japan, and Harbin Engineering University, in 2014. He was a Visiting Scholar with the University of California, Davis, from March 2016 to March 2017. He has been a Full Professor with Harbin Engineering University, since July 2014. He has also been a Visiting Professor with Far Eastern Federal University (FEFU) and KUT, since 2018. His recent research interests are mainly in remote sensing, underwater communications, signal processing, radar, SAR imaging, and compressed sensing and antennas. He is also a senior member of the Chinese Institute of Electronics (CIE). He also serves as a Reviewer for more than 20 journals. He is also an Associate Editor of IEEE ACCESS and the AEU-International Journal of Electronics and Communications.