

Article

# A General Zero Attraction Proportionate Normalized Maximum Correntropy Criterion Algorithm for Sparse System Identification

Yingsong Li <sup>1,2,\*</sup> , Yanyan Wang <sup>1</sup>, Felix Albu <sup>3</sup>  and Jingshan Jiang <sup>2</sup>

<sup>1</sup> College of Information and Communication Engineering, Harbin Engineering University, Harbin 150001, China; wangyanyan@hrbeu.edu.cn

<sup>2</sup> National Space Science Center, Chinese Academy of Sciences, Beijing 100190, China; jiangjingshan@mirslab.cn

<sup>3</sup> Department of Electronics, Valahia University of Targoviste, 130082 Targoviste, Romania; felix.albu@valahia.ro

\* Correspondence: liyingsong@ieee.org; Tel.: +86-451-82519808

Received: 18 September 2017; Accepted: 6 October 2017; Published: 15 October 2017

**Abstract:** A general zero attraction (GZA) proportionate normalized maximum correntropy criterion (GZA-PNMCC) algorithm is devised and presented on the basis of the proportionate-type adaptive filter techniques and zero attracting theory to highly improve the sparse system estimation behavior of the classical MCC algorithm within the framework of the sparse system identifications. The newly-developed GZA-PNMCC algorithm is carried out by introducing a parameter adjusting function into the cost function of the typical proportionate normalized maximum correntropy criterion (PNMCC) to create a zero attraction term. The developed optimization framework unifies the derivation of the zero attraction-based PNMCC algorithms. The developed GZA-PNMCC algorithm further exploits the impulsive response sparsity in comparison with the proportionate-type-based NMCC algorithm due to the GZA zero attraction. The superior performance of the GZA-PNMCC algorithm for estimating a sparse system in a non-Gaussian noise environment is proven by simulations.

**Keywords:** adaptive filter; normalized maximum correntropy criterion; normalized least mean square (NLMS); proportionate NLMS (PNLMS); zero attracting; impulsive noise

## 1. Introduction

Adaptive filtering techniques have been extensively exploited, and they have experienced a challenging problem as the impulse response (IR) of the unknown systems is sparse [1–3]. As a result, a concept for developing sparse adaptive filters to handle such problem is becoming a hot topic for all researchers [4–12]. For sparse systems, most of their coefficients take the values of zero or near-zeros, while only a few coefficients have significant values [13–16]. Such sparse systems are commonly encountered in our in-nature world and plenty of real-world engineering applications such as the multi-path channels in wireless communications [17,18], underwater communications [19,20] and acoustic channels [21]. For example, for the signal transmissions in wireless communication in hilly-terrain environments, the channel response always exhibits a multi-path fading selecting phenomenon due to the delays of the different propagation paths [22]. Thus, the measurement result shows a sparse characteristic for such channels [23]. Furthermore, this phenomenon happens over wireless voice transmission systems and networks, in which the echo paths are demonstrated as active ones and some unknown bulk delays occur because of the propagation of the network, buffer delays and encoding processing [24]. Usually, these bulk delay intervals have little energy, which makes

the impulse response sparse [25]. For dealing with this class of sparse impulse responses, traditional adaptive filters might converge slowly since all the system coefficients have the same step sizes, such as the normalized least mean square (NLMS) algorithm [26–28]. To cancel out this drawback, proportionate-type techniques have been presented to exploit the natural sparseness characteristics of these impulse responses [11,29,30].

The proportionate NLMS (PNLMS) is a popular sparse signal processing algorithm for handling these sparse systems, which is realized by proportionately updating the filter coefficients in accordance with the magnitudes of these coefficients [11]. As a result, the PNLMS converges faster at its initial stage compared with the NLMS algorithm [29–31]. However, after the initial convergence, the PNLMS's convergence is getting bad, even its convergence is worse than the NLMS algorithm. Several improved PNLMS algorithms have been reported to treat the medium sparsity systems, including the improved PNLMS (IPNLMS) and PNLMS++ algorithms [30,31]. However, these algorithms still cannot give the expected convergence achieved from the PNLMS at the beginning iteration stage for high sparseness impulse response applications. Recently, an information theoretic quantity has been presented to build the required cost function for adaptive signal processing applications [32–38]. To calculate the quantities, the entropy estimator was reported in detail. Then, a minimum error entropy (MEE) has been developed in the context of the adaptive filtering framework by using the entropy as a cost function to justify the signal structure well [32]. From the estimation performance of the MEE, we found that it was robust for processing the impulsive noise signals [39]. However, it has high computational complexity compared with the least mean square (LMS) algorithm, which is posing a problem for practical applications [39,40]. Moreover, the correntropy can be used as an efficient optimality criterion in robust and sparse adaptive filtering [37]. Recently, a maximum correntropy criterion (MCC) algorithm was reported based on a cost function with localized similarity [40]. Furthermore, the MCC has a similar robustness characteristics to the MEE algorithm, and it possesses a LMS-like complexity, making it suitable for practical engineering applications in impulsive noise environments. Additionally, Chen et al. proposed a novel similarity measure called generalized correntropy, which when used in robust adaptive filtering, can significantly outperform the existing methods [37]. However, these MCCs cannot utilize the sparseness characteristics of the unknown systems. Motivated by the zero attracting theory, zero attracting MCC (ZA-MCC), reweighted ZA-MCC (RZA-MCC), group-constrained MCC (GC-MCC) and reweighted GC-MCC (RGC-MCC) algorithms [41,42] have been reported for estimating sparse channels. However, these algorithms are realized based on the MCC algorithm to exploit the sparsity property of the multi-path channels, which might be affected by the input scalings.

In this paper, a general zero attraction proportionate normalized MCC (GZA-PNMCC) algorithm is developed for the application in the sparse system identifications. The devised GZA-PNMCC algorithm is proposed based on the proportionate adaptive filtering scheme and the zero attracting techniques. In order to get the presented GZA-PNMCC, the following steps are considered. Firstly, the normalized MCC (NMCC) is proposed, followed by the introduction of the proportionate-type technique into the NMCC to construct the PNMCC algorithm. Secondly, a parameter adjustment function is integrated into the PNMCC algorithm to create the proposed GZA-PNMCC algorithm, and the zero attracting can be well adjusted for meeting various applications. The proposed GZA-PNMCC algorithm is analyzed, and the influence of different parameters and sparsity levels is investigated. The achieved results demonstrate that the GZA-PNMCC algorithm provides a faster convergence and better sparse system estimation performance in a mixed Gaussian noise environment. Different from the GC-MCC and RGC-MCC, the proposed GZA-PNMCC is exploited based on the PNMCC that is a proportionate-type NMCC algorithm. The GZA-PNMCC algorithm utilizes the parameter adjustment function to adjust the penalties on the channel coefficients to approximate various norms, which is also different with the GC and RGC techniques that can only use the  $l_0$ -norm penalty on the large group and the  $l_1$ -norm penalty on the small group. The proposed GZA-PNMCC can exert various penalties on the channel coefficients. Since the proposed GZA-PNMCC and the previous GC-MCC algorithms are presented based on different basic algorithms and different penalties are used in these algorithms,

we believe that the proposed GZA-PNMCC algorithm can further exploit the inherent property of the multi-path channels.

The paper consists of the following sections. The NMCC algorithm and its proportionate form, the PNMCC algorithm, are presented in Section 2. In Section 3, the GZA-PNMCC algorithm is mathematically derived within the framework of the sparse system identifications. The influence of key parameters and sparsity levels on the estimation performance of GZA-PNMCC is investigated in Section 4. Finally, conclusions are summarized in Section 5.

## 2. Past Works on NMCC and PNMCC Algorithms

### 2.1. NMCC Algorithm

We herein consider a sparse unknown system whose impulse responses are  $\mathbf{w}_o(n) = [w_0, w_1, w_2, \dots, w_{N-2}, w_{N-1}]^T$ . We use the NMCC algorithm to get an estimation  $\hat{\mathbf{w}}(n)$  of the unknown system  $\mathbf{w}_o(n)$  by using a training signal  $\mathbf{x}(n) = [x(n), x(n-1), x(n-2), \dots, x(n-N+2), x(n-N+1)]^T$ , and the estimated error is  $e(n)$ . Here,  $e(n)$  denotes the differences between the expected signal  $d(n)$  and the estimator output signal  $y(n)$ . Based on the adaptive filtering theory, we have:

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

$$y(n) = \mathbf{x}^T(n)\hat{\mathbf{w}}(n), \quad (2)$$

and:

$$e(n) = d(n) - \mathbf{x}^T(n)\hat{\mathbf{w}}(n), \quad (3)$$

where  $v(n)$  denotes as an impulsive noise. NMCC-based system identification tries to find out a solution of the following problem [40]:

$$\begin{aligned} & \min \frac{1}{2} \|\hat{\mathbf{w}}(n+1) - \hat{\mathbf{w}}(n)\|^2, \\ & \text{subject to } \hat{e}(n) = \left[ 1 - \xi \exp\left(-\frac{e^2(n)}{2\sigma^2}\right) \right] e(n) \end{aligned} \quad (4)$$

where  $\hat{e}(n) = d(n) - \mathbf{x}^T(n)\hat{\mathbf{w}}(n+1)$ ,  $\|\cdot\|^2$  denotes the Euclidean vector norm and  $\sigma > 0$  is used to give a tradeoff effect between the convergence and the MSE,  $\xi = \chi \|\mathbf{x}(n)\|^2$ . By using the Lagrange multiplier method (LMM), we can get NMCC's updating equation, which is [40]:

$$\hat{\mathbf{w}}(n+1) = \hat{\mathbf{w}}(n) + \chi \frac{\exp\left(-\frac{e^2(n)}{2\sigma^2}\right)}{\|\mathbf{x}(n)\|^2} e(n)\mathbf{x}(n). \quad (5)$$

From the derivation of the NMCC algorithm, we find that the updating equation in (5) of the NMCC algorithm is similar to the NLMS algorithm. However, there is an exponential term in (5), which renders the NMCC algorithm more powerful for preventing impulsive-like noises.

### 2.2. PNMCC Algorithm

A gain assignment matrix  $\mathbf{G}(n)$  similar to that of the known PNLMS algorithm is used to devise the PNMCC algorithm. By introducing  $\mathbf{G}(n)$  into (5), the PNMCC's updating equation is obtained [11,43]:

$$\hat{\mathbf{w}}(n+1) = \hat{\mathbf{w}}(n) + \chi \frac{\mathbf{G}(n) \exp\left(-\frac{e^2(n)}{2\sigma^2}\right)}{\mathbf{x}^T(n)\mathbf{G}(n)\mathbf{x}(n) + \vartheta} e(n)\mathbf{x}(n), \quad (6)$$

where  $\vartheta > 0$  denotes a small constant, which is to prevent the denominator from being divided by zero, and it can provide a stable solution. Similar to the PNLMS algorithm,  $\mathbf{G}(n)$  denotes a diagonal matrix,

which is used for modifying the step size of the NMCC algorithm under a certain rule. Generally speaking,  $\mathbf{G}(n)$  is [11]:

$$\mathbf{G}(n) = \text{diag}(g_0(n), g_1(n), \dots, g_{N-1}(n)), \quad (7)$$

where the elements  $g_i(n)$  is described as:

$$g_i(n) = \frac{\kappa_i(n)}{\sum_{i=0}^{N-1} \kappa_i(n)}, \quad 0 \leq i \leq N-1 \quad (8)$$

with:

$$\kappa_i(n) = \max[\gamma_g \max[\rho_p, |\hat{w}_0(n)|, |\hat{w}_1(n)|, \dots, |\hat{w}_{N-1}(n)|], |\hat{w}_i(n)|], \quad (9)$$

where  $\rho_p > 0$  and  $\gamma_g >$  usually have typical values of  $\rho_p = 0.01$  and  $\gamma_g = 5/N$ .  $\rho_p$  aims to avoid the weight updating's stalling at the initial iteration stage when all the system coefficients are set to be zero, and  $\gamma_g$  aims to prevent  $\hat{w}_i(n)$  from stalling when it is much smaller than the largest coefficient.

### 3. Proposed GZA-PNMCC Algorithm

The GZA-PNMCC algorithm is based on a parameter adjustment function constraint and the PNMCC algorithm to enhance the convergence characteristic of the PNMCC algorithm after the initial stage. To realize the GZA-PNMCC algorithm, a parameter adjustment function is incorporated into PNMCC's cost function, where the parameter adjustment function is given as follows:

$$S_\beta(\hat{\mathbf{w}}(n)) = (1 + \beta^{-1})(1 - e^{-\beta|\hat{\mathbf{w}}(n)|}), \quad (10)$$

where  $\beta > 0$  represents a regulation parameter to model the desired general zero attraction. The behavior of the parameter adjustment function is well investigated with different  $\beta$  and is shown in Figure 1. It can be seen that the parameter adjustment function can be used for approximating an  $l_0$ -norm when a large  $\beta$  is used. If we choose a very small  $\beta$ , the parameter adjustment function has a characteristic that is similar to the  $l_1$ -norm. Therefore, a proper  $\beta$  should be selected to develop a flexible norm-like penalty. Based on the analysis of the performance of the parameter adjustment function, we can say that it can be used for implementing  $l_1$ -norm-like,  $l_0$ -norm-like, even  $l_p$ -norm-like penalties. Since the parameter adjustment function is flexible and useful for implementing the desired norms, it is used to derive the general zero attraction PNMCC algorithm. The proposed parameter adjustment function is integrated into the PNMCC to continue to use the sparseness for system identification. The GZA-PNMCC algorithm aims to solve the following problem:

$$\begin{aligned} & (\hat{\mathbf{w}}(n+1) - \hat{\mathbf{w}}(n))^T \mathbf{G}^{-1}(n) (\hat{\mathbf{w}}(n+1) - \hat{\mathbf{w}}(n)) + \gamma_{\text{GZA}} \mathbf{G}^{-1}(n) S_\beta(\hat{\mathbf{w}}(n+1)) \\ & \text{subject to } \hat{e}(n) = \left[ 1 - \zeta \exp\left(-\frac{e^2(n)}{2\sigma^2}\right) \right] e(n) \end{aligned} \quad (11)$$

In Equation (11),  $\mathbf{G}^{-1}(n)$  represents the inverse form of the gain assignment matrix  $\mathbf{G}(n)$ .  $\gamma_{\text{GZA}}$  denotes a small positive regularization to give a tradeoff between the convergence and the sparseness development. It can be seen that the GZA-PNMCC algorithm utilizes a  $\gamma_{\text{GZA}} \mathbf{G}^{-1}(n) S_\beta(\hat{\mathbf{w}}(n+1))$  to create an expected zero attraction term, which makes it different from the PNMCC algorithm. Additionally,  $\mathbf{G}^{-1}(n)$  is integrated into Equation (11).

Then, we employ the LMM to find out the solution of Equation (11). As a result, we write the proposed GZA-PNMCC'S cost function as:

$$\begin{aligned} J_{\text{GZA}}(n+1) &= (\hat{\mathbf{w}}(n+1) - \hat{\mathbf{w}}(n))^T \mathbf{G}^{-1}(n) (\hat{\mathbf{w}}(n+1) - \hat{\mathbf{w}}(n)) \\ &+ \gamma_{\text{GZA}} \mathbf{G}^{-1}(n) S_\beta(\hat{\mathbf{w}}(n+1)) + \lambda \left( \hat{e}(n) - \left[ 1 - \zeta \exp\left(-\frac{e^2(n)}{2\sigma^2}\right) \right] e(n) \right). \end{aligned} \quad (12)$$

where  $\lambda$  represents the Lagrange multiplier. The gradients of  $J_{GZA}(n+1)$  with respect to  $\hat{\mathbf{w}}(n+1)$  and  $\lambda$  are obtained as follows:

$$\frac{\partial J_{GZA}(n+1)}{\partial \hat{\mathbf{w}}(n+1)} = 0 \text{ and } \frac{\partial J_{GZA}(n+1)}{\partial \lambda} = 0, \tag{13}$$

and:

$$\hat{\mathbf{w}}(n+1) = \hat{\mathbf{w}}(n) + \lambda \mathbf{G}(n)\mathbf{x}(n) - \gamma_{GZA} S'_\beta(\hat{\mathbf{w}}(n+1)), \tag{14}$$

where  $S'_\beta(\hat{\mathbf{w}}(n+1))$  is the derivative of the parameter adjustment function, which is:

$$s'_{i,\beta}(\hat{w}_i(n+1)) = (\beta + 1)e^{(-\beta|\hat{w}_i(n+1)|)} \text{sgn}(w_i(n+1)), \tag{15}$$

and its vector form is:

$$S'_\beta(\hat{\mathbf{w}}(n+1)) = (\beta + 1)e^{(-\beta|\hat{\mathbf{w}}(n+1)|)} \text{sgn}(\hat{\mathbf{w}}(n+1)). \tag{16}$$

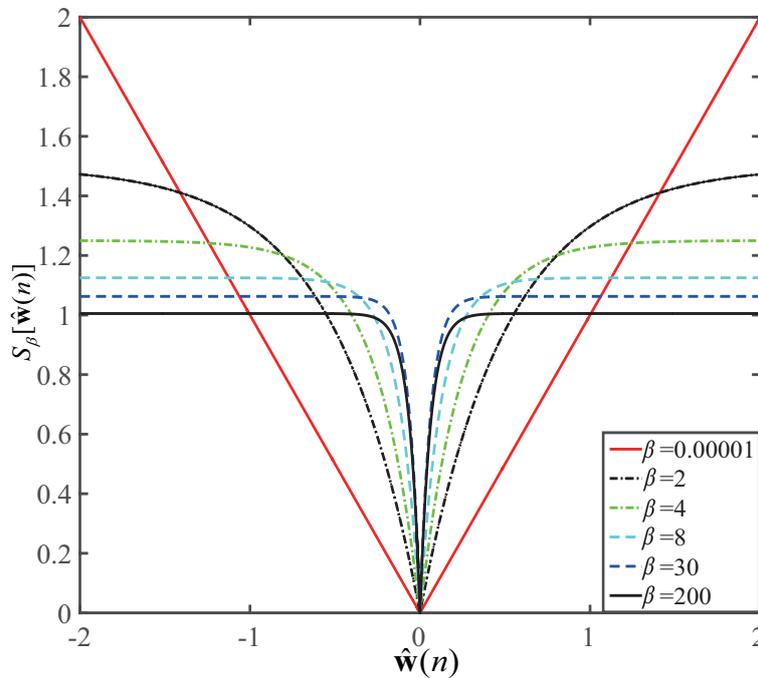


Figure 1. Behavior of the parameter adjustment function with various  $\beta$ .

By left multiplying by  $\mathbf{x}^T(n)$  on both sides of Equation (14), we can obtain:

$$\mathbf{x}^T(n)\hat{\mathbf{w}}(n+1) = \mathbf{x}^T(n)\hat{\mathbf{w}}(n) + \lambda \mathbf{x}^T(n)\mathbf{G}(n)\mathbf{x}(n) - \gamma_{GZA}\mathbf{x}^T(n)S'_\beta(\hat{\mathbf{w}}(n+1)). \tag{17}$$

From (13), we can get:

$$\hat{e}(n) = \left[ 1 - \zeta \exp\left(-\frac{e^2(n)}{2\sigma^2}\right) \right] e(n). \tag{18}$$

Thus, Lagrange multiplier  $\lambda$  is:

$$\lambda = \frac{\zeta \exp\left(-\frac{e^2(n)}{2\sigma^2}\right) e(n) + \gamma_{GZA}\mathbf{x}^T(n)S'_\beta(\hat{\mathbf{w}}(n+1))}{\mathbf{x}^T(n)\mathbf{G}(n)\mathbf{x}(n)} \tag{19}$$

By substituting (19) into (14), we can get:

$$\begin{aligned}
 & \hat{\mathbf{w}}(n+1) \\
 &= \hat{\mathbf{w}}(n) + \frac{\xi \exp\left(-\frac{e^2(n)}{2\sigma^2}\right) e(n) + \gamma_{GZA} \mathbf{x}^T(n) S'_\beta(\hat{\mathbf{w}}(n+1))}{\mathbf{x}^T(n) \mathbf{G}(n) \mathbf{x}(n)} \mathbf{G}(n) \mathbf{x}(n+1) \\
 & \quad - \gamma_{GZA} S'_\beta(\hat{\mathbf{w}}(n)) \\
 &= \hat{\mathbf{w}}(n) + \frac{\xi e(n) \exp\left(-\frac{e^2(n)}{2\sigma^2}\right) \mathbf{G}(n) \mathbf{x}(n)}{\mathbf{x}^T(n) \mathbf{G}(n) \mathbf{x}(n)} \\
 & \quad - \gamma_{GZA} \left\{ \mathbf{I} - \frac{\mathbf{G}(n) \mathbf{x}(n) \mathbf{x}^T(n)}{\mathbf{x}^T(n) \mathbf{G}(n) \mathbf{x}(n)} \right\} S'_\beta(\hat{\mathbf{w}}(n)) \\
 &= \hat{\mathbf{w}}(n) + \frac{\xi e(n) \exp\left(-\frac{e^2(n)}{2\sigma^2}\right) \mathbf{G}(n) \mathbf{x}(n)}{\mathbf{x}^T(n) \mathbf{G}(n) \mathbf{x}(n)} \\
 & \quad - \gamma_{GZA} \left\{ \mathbf{I} - \frac{\mathbf{G}(n) \mathbf{x}(n) \mathbf{x}^T(n)}{\mathbf{x}^T(n) \mathbf{G}(n) \mathbf{x}(n)} \right\} (\beta + 1) e^{(-\beta|\hat{\mathbf{w}}(n)|)} \text{sgn}(\hat{\mathbf{w}}(n))
 \end{aligned} \tag{20}$$

From the updating Equation (20), we found that the elements in  $\mathbf{G}(n) \mathbf{x}(n) \mathbf{x}^T(n) \{\mathbf{x}^T(n) \mathbf{G}(n) \mathbf{x}(n)\}^{-1}$  were smaller than one; therefore, they can be ignored [29]. Thereby, we rewrite the updating equation of the developed GZA-PNMCC algorithm as:

$$\hat{\mathbf{w}}(n+1) = \hat{\mathbf{w}}(n) + \frac{\xi e(n) \exp\left(-\frac{e^2(n)}{2\sigma^2}\right) \mathbf{G}(n) \mathbf{x}(n)}{\mathbf{x}^T(n) \mathbf{G}(n) \mathbf{x}(n)} - \gamma_{GZA} (\beta + 1) e^{(-\beta|\hat{\mathbf{w}}(n)|)} \text{sgn}(\hat{\mathbf{w}}(n)) \tag{21}$$

To better control the convergence, a step size  $\mu$  is introduced into the updating Equation (21), which is similar to the PNMCC algorithm. Additionally, we also employ a small positive constant to avoid dividing by zero. As a result, the updated equation of the developed GZA-PNMCC algorithm is modified to be:

$$\hat{\mathbf{w}}(n+1) = \hat{\mathbf{w}}(n) + \chi_1 \frac{\exp\left(-\frac{e^2(n)}{2\sigma^2}\right) \mathbf{G}(n) e(n) \mathbf{x}(n)}{\mathbf{x}^T(n) \mathbf{G}(n) \mathbf{x}(n) + \varepsilon_{GZA}} - \rho_{GZA} (\beta + 1) e^{(-\beta|\hat{\mathbf{w}}(n)|)} \text{sgn}(\hat{\mathbf{w}}(n)) \tag{22}$$

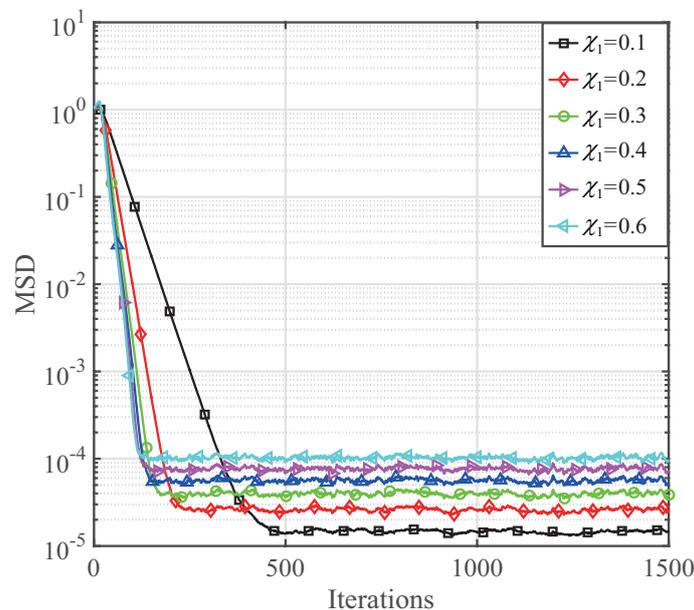
where  $\chi_1 = \xi\mu$ , and  $\rho_{GZA} = \mu\gamma_{GZA}$  is a regularization parameter that controls the ability of the zero attraction. The zero attraction term  $-\rho_{GZA} (\beta + 1) e^{(-\beta|\hat{\mathbf{w}}(n)|)} \text{sgn}(\hat{\mathbf{w}}(n))$  in Equation (22) is named as an expected zero attractor. In the developed GZA-PNMCC algorithm, a proper  $\beta$  can be chosen to produce different penalties. The GZA zero attraction can give a high probability zero attraction on the zero coefficients or near-zero coefficients. What is more, a gain enhancement matrix is used for further assigning the large step-size to dominant coefficients. In fact, our developed GZA-PNMCC algorithm provides fast convergence for large coefficients by using the gain matrix scheme, while it speeds up the convergence for the small coefficients by using the proposed GZA zero attraction.

#### 4. Behavior of the Proposed GZA-PNMCC Algorithm

We construct two examples to discuss the performance of the developed GZA-PNMCC algorithm within the framework of the sparse system identifications. We compare its performance with the MCC [39,40], NMCC [40], ZA-MCC [41], RZA-MCC [41] and PNLMS [11] algorithms. We use mean square deviation (MSD) to investigate the estimation behavior of the developed GZA-PNMCC algorithm. The definition of the MSD is:

$$\text{MSD}(\hat{\mathbf{w}}(n)) = E[\|\mathbf{w}_o(n) - \hat{\mathbf{w}}(n)\|^2]. \quad (23)$$

We find that the parameter  $\chi_1$  has an important effect on the estimation behavior of the GZA-PNMCC algorithm. Therefore, we create an experiment to give an illustration of the performance of the GZA-PNMCC in the presence of the impulsive noise. We use  $(1 - \theta)N(\iota_1, \nu_1^2) + \theta N(\iota_2, \nu_2^2) = (0, 0.01, 0, 20, 0.05)$  to model the desired impulsive noise, where  $N(\iota_i, \nu_i^2)$  ( $i = 1, 2$ ) are Gaussian distributions with means of  $\iota_i$  and variances of  $\nu_i^2$  [41].  $\theta$  denotes a mixture parameter, which is to control the mixed performance of the two noises. In this paper, we assume  $\mathbf{x}(n)$  is independent with  $v(n)$ . Furthermore, we use a sparse system whose length is  $N = 16$ . There is only one dominant coefficient, which is randomly distributed within the impulse response of the system. The parameters  $\vartheta = 0.01$  and  $\sigma = 1000$  are used to investigate the effects of  $\chi_1$ , and the computer simulation results are shown in Figure 2. It is noted from Figure 2 that the developed GZA-PNMCC algorithm has a fast convergence when  $\chi_1$  is large. For smaller  $\chi_1$ , the GZA-PNMCC algorithm converges slower in comparison with larger  $\chi_1$ . However, we can achieve lower estimation misalignment when we choose small  $\chi_1$ .



**Figure 2.** Effects of  $\chi_1$  on the estimation behavior.

Furthermore,  $\rho_{\text{GZA}}$  plays an important role in the GZA-PNMCC algorithm since it controls the zero-attraction ability. We discuss the effects of  $\rho_{\text{GZA}}$  for estimation of a sparse system, which is the same as that in the  $\chi_1$  investigation. The performance of  $\rho_{\text{GZA}}$  is given in Figure 3. It can be seen that  $\rho_{\text{GZA}} = 9 \times 10^{-5}$  provides the fastest convergence speed rate and the best system estimation behavior. Generally speaking, the estimation misalignment becomes smaller when  $\rho_{\text{GZA}}$  decreases from  $9 \times 10^{-4}$  to  $9 \times 10^{-5}$ . If we continue to reduce  $\rho_{\text{GZA}}$ , the estimation misalignment is changed in the opposite direction, meaning that the estimation misalignment becomes large when  $\rho_{\text{GZA}}$  decreases from  $5 \times 10^{-5}$  to  $5 \times 10^{-7}$ . Therefore, we should carefully select  $\rho_{\text{GZA}}$  and  $\chi_1$  in order to get a good performance and a balance between the convergence speed rate and estimation misalignment.

Based on the key parameter investigations, we create an experiment to discuss the convergence characteristics of the GZA-PNMCC algorithm and compare its convergence with the PNLMS, PNMCC, NMCC, MCC, ZA-MCC and RZA-MCC algorithms. To compare the convergence at the same estimation misalignment level, we set all the simulation parameters as  $\chi_{\text{MCC}} = 0.0052$ ,  $\chi_{\text{NMCC}} = 0.085$ ,  $\mu_{\text{ZA}} = 0.01$ ,  $\mu_{\text{RZA}} = 0.015$ ,  $\rho_{\text{ZA}} = 3 \times 10^{-5}$ ,  $\rho_{\text{RZA}} = 7 \times 10^{-5}$ ,  $\mu_{\text{PNLMS}} = 0.072$ ,  $\chi = 0.088$ ,  $\chi_1 = 0.35$ ,  $\rho_{\text{GZA}} = 8 \times 10^{-5}$ , where  $\chi_{\text{MCC}}$ ,  $\chi_{\text{NMCC}}$ ,  $\mu_{\text{ZA}}$ ,  $\mu_{\text{RZA}}$ ,  $\mu_{\text{PNLMS}}$  are the step sizes for the MCC, NMCC,

ZA-MCC, RZA-MCC and PNLMS algorithms.  $\rho_{ZA}$  and  $\rho_{RZA}$  are the zero attraction parameters for ZA-MCC and RZA-MCC algorithms. The convergence of the GZA-PNMCC algorithm is described in Figure 4. Comparing with the previously presented sparse adaptive filtering ZA-MCC, RZA-MCC, PNMCC and PNLMS algorithms, our GZA-PNMCC algorithm achieves the fastest convergence speed rate because it integrates a flexible GZA zero attractor that can accelerate the convergence of zero or near-zero coefficients.

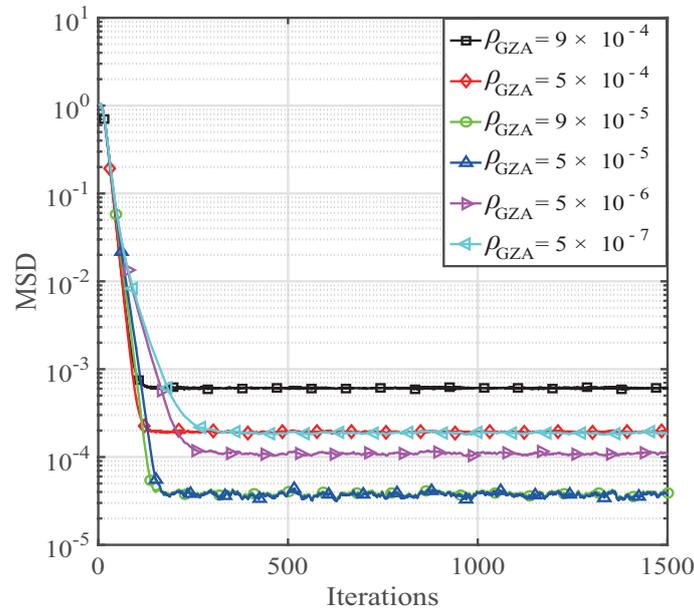


Figure 3. Effects of  $\rho_{GZA}$  on the estimation behavior.

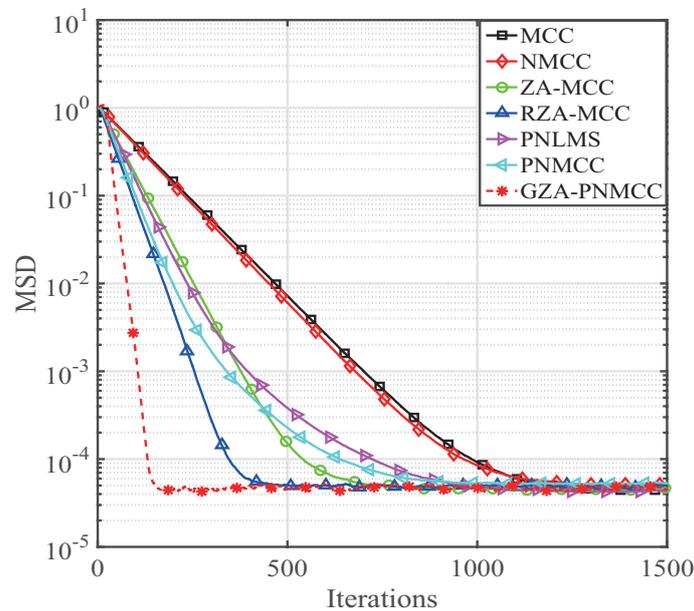
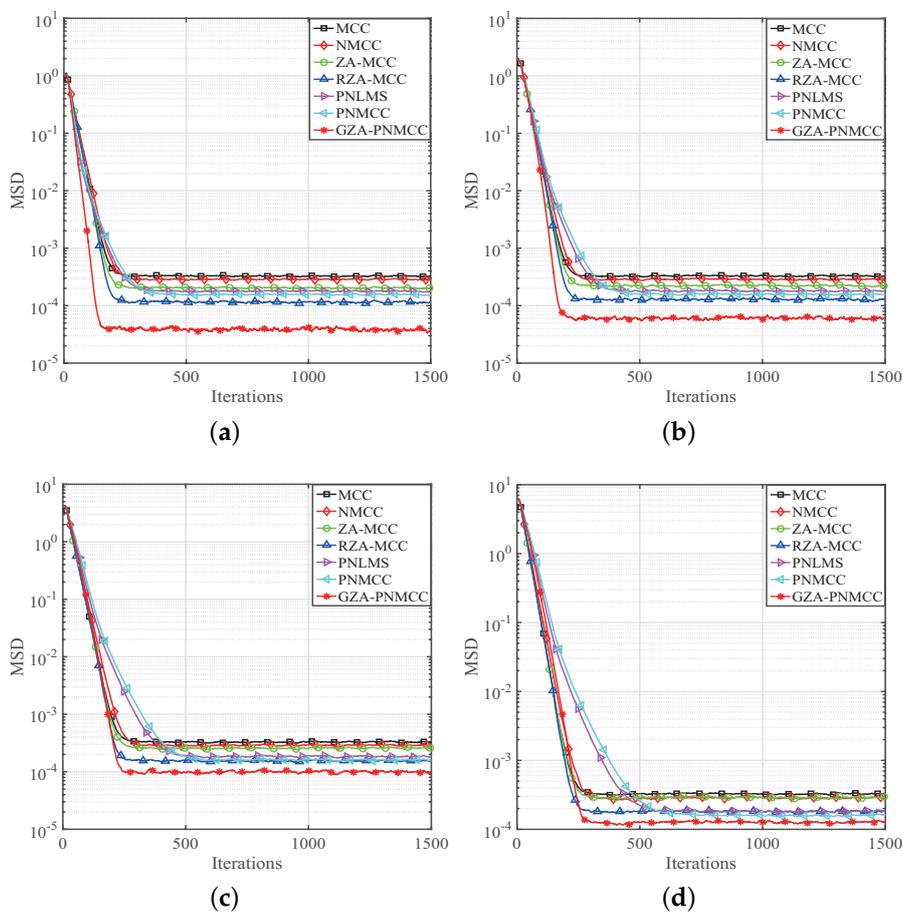


Figure 4. Convergence of the proposed general zero attraction proportionate normalized maximum correntropy criterion (GZA-PNMCC) algorithm.

Then, an experiment is set up to investigate the system estimation performance of the GZA-PNMCC algorithm in the context of the sparse system identifications with different effects of the sparsity levels  $K$ , where  $K$  denotes the number of the dominant coefficients in the impulse response of the unknown sparse system. Assuming that the IR of the unknown sparse system has a

length of  $N = 16$ . To achieve the same convergence at the early stage, the simulation parameters are set as  $\chi_{\text{MCC}} = 0.03$ ,  $\chi_{\text{NMCC}} = 0.4$ ,  $\mu_{\text{ZA}} = \mu_{\text{RZA}} = 0.03$ ,  $\rho_{\text{ZA}} = 8 \times 10^{-5}$ ,  $\rho_{\text{RZA}} = 2 \times 10^{-4}$ ,  $\mu_{\text{PNLMS}} = 0.27$ ,  $\chi = 0.24$ ,  $\chi_1 = 0.3$ ,  $\rho_{\text{GZA}} = 9 \times 10^{-5}$ . The GZA-PNMCC's estimation behavior is given in Figure 5, and its performance is compared with those of the MCC, NMCC, ZA-MCC, RZA-MCC and PNLMS algorithms. In this experiment, we choose  $K = 1, 2, 4, 6$  to discuss the system estimation behavior of the GZA-PNMCC. From Figure 5a, we can see that the GZA-PNMCC algorithm has the fastest convergence and achieves the smallest estimation misalignment. The GZA-PNMCC algorithm provides much more gain than that of the PNMCC algorithm. Comparing to the previously reported ZA-MCC and RZA-MCC, our GZA-PNMCC algorithm has clear advantages with respect to both convergence speed rate and steady-state MSD. As  $K$  increase from one to six, it is found that the estimation misalignment becomes larger because of the reduced sparsity. However, the developed GZA-PNMCC is still better than the MCC, NMCC, ZA-MCC, RZA-MCC and PNLMS algorithms in terms of the MSD.



**Figure 5.** Estimation behaviors of the GZA-PNMCC algorithm with different sparsity level  $K$ . (a)  $K = 1$ ; (b)  $K = 2$ ; (c)  $K = 4$ ; (d)  $K = 6$ .

Finally, we set up an experiment to study the tracking behavior of our GZA-PNMCC algorithm for estimating a long-tap echo channel with two different sparsity levels and a length of 256. The sparsity measurement of the echo channel is  $\zeta_{12}(\mathbf{w}_0) = \frac{N}{N - \sqrt{N}} \left( 1 - \frac{\|\mathbf{w}_0\|_1}{\sqrt{N}\|\mathbf{w}_0\|_2} \right)$  [44–48]. A typical echo channel is described in Figure 6. In this experiment, the sparsity is  $\zeta_{12}(\mathbf{w}_0) = 0.8222$  for the first 8000 iterations, while it is changed to be  $\zeta_{12}(\mathbf{w}_0) = 0.7362$  for the last 8000 iterations. Here,  $\zeta_{12}(\mathbf{w}_0)$  can choose different values to get different channels. For  $\zeta_{12}(\mathbf{w}_0) = 0.8222$ , the channel is sparser than that of  $\zeta_{12}(\mathbf{w}_0) = 0.7362$ . The parameters are  $\chi_{\text{MCC}} = 0.0055$ ,  $\chi_{\text{NMCC}} = 1.3$ ,  $\mu_{\text{ZA}} = \mu_{\text{RZA}} = 0.0055$ ,

$\rho_{ZA} = 4 \times 10^{-6}$ ,  $\rho_{RZA} = 1 \times 10^{-5}$ ,  $\mu_{PNLMS} = 1$ ,  $\chi = 0.9$ ,  $\chi_1 = 0.8$ ,  $\rho_{GZA} = 1 \times 10^{-6}$ . The behavior of the GZA-PNMCC algorithm for estimating an echo channel is shown in Figure 7. We note that our GZA-PNMCC algorithm provides the fastest convergence speed rate and smallest MSD for  $\zeta_{12}(\mathbf{w}_o) = 0.8222$ . Furthermore, it can be seen that there are two convergence stages, which are caused by the proportionate and the zero attraction schemes, respectively. Even when  $\zeta_{12}(\mathbf{w}_o) = 0.7362$ , the GZA-PNMCC algorithm still outperforms the PNMCC, PNLMS, MCC, NMCC, ZA-MCC and RZA-MCC algorithms. The GZA-PNMCC algorithm is less affected by the sparseness of the unknown system  $\mathbf{w}_o$ , which renders it effective and useful for sparse system identification. In addition, the proposed GZA-PNMCC algorithm has a moderate computational complexity, including  $(4N + 5)$  additions,  $(7N + 4)$  multiplications,  $(N)$  divisions and  $(N + 1)$  exponential calculations.

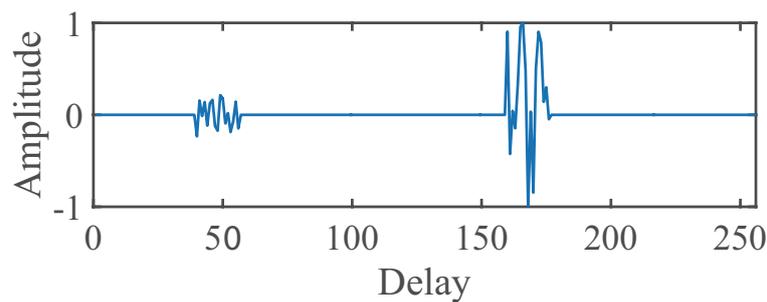


Figure 6. A typical echo channel with a length of 256.

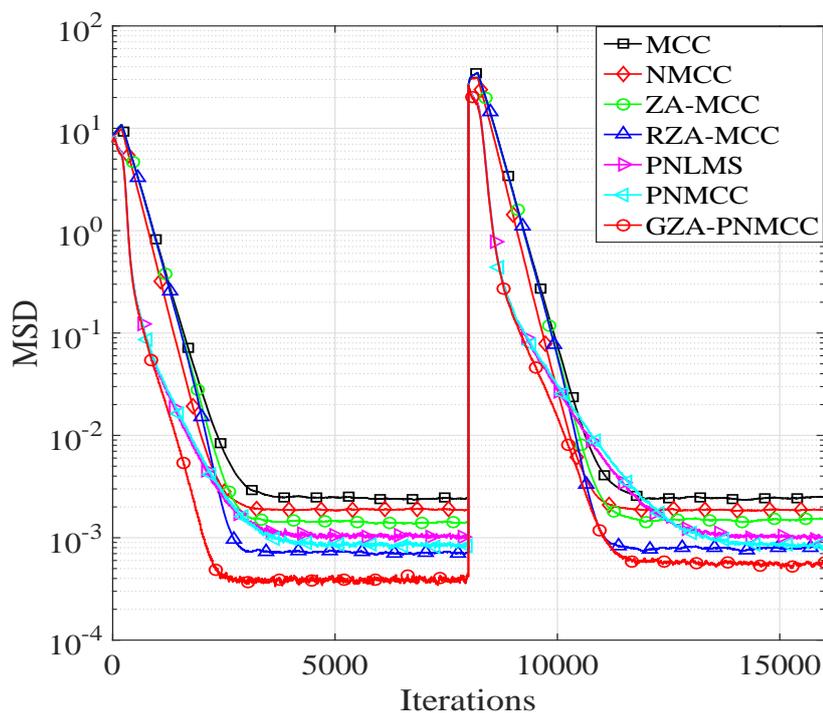


Figure 7. Tracking behavior of our GZA-PNMCC algorithm.

## 5. Conclusions

A GZA-PNMCC algorithm with zero attraction has been developed for estimating sparse system identifications. The derivation and the analysis of the GZA-PNMCC algorithm is achieved by the LMM. The GZA-PNMCC algorithm uses a zero attraction scheme to improve the convergence at the subsequent stage by quickly forcing the zero system coefficients or near-zero system coefficients to

zero. The simulation results demonstrated that the developed GZA-PNMCC algorithm is effective and useful for sparse system identifications because it provides superior performance to that of the PNLMS and PNMCC algorithms. Thereby, our developed GZA-PNMCC is suitable for sparse signal processing in practical engineering applications such as sparse channel estimation and echo cancellation.

**Acknowledgments:** This work was partially supported by the National Key Research and Development Program of the China-Government Corporation Special Program (2016YFE0111100), the Science and Technology innovative Talents Foundation of Harbin (2016RAXXJ044), the Projects for the Selected Returned Overseas Chinese Scholars of Heilongjiang Province and MOHRSS of China and the PhD Student Research and Innovation Fund of the Fundamental Research Funds for the Central Universities (HEUGIP201707).

**Author Contributions:** Yingsong Li wrote the draft of the paper and wrote the code. Yanyan Wang did the simulations of this paper. Felix Albu helped modify the paper and checked the grammar. Jingshan Jiang gave the analysis of the paper. All the authors wrote this paper together, and they have read and approved the final manuscript.

**Conflicts of Interest:** The authors declare no conflict of interest.

## References

1. Khong, A.W.H.; Naylor, P.A. Efficient use of sparse adaptive filters. In Proceedings of the Fortieth Asilomar Conference on Signals, Systems and Computers (ACSSC '06), Pacific Grove, CA, USA, 29 October–1 November 2006; pp. 1375–1379.
2. Paleologu, C.; Benesty, J.; Ciochina, S. *Sparse Adaptive Filters for Echo Cancellation*; Morgan & Claypool: San Rafael, CA, USA, 2010.
3. Rodger, J.A. Toward reducing failure risk in an integrated vehicle health maintenance system: A fuzzy multi-sensor data fusion Kalman filter approach for IVHMS. *Expert Syst. Appl.* **2012**, *139*, 9821–9836.
4. Murakami, Y.; Yamagishi, M.; Yukawa, M.; Yamada, I. A sparse adaptive filtering using time-varying soft-thresholding techniques. In Proceedings of the 2010 IEEE International Conference on Acoustics Speech and Signal Processing (ICASSP), Dallas, TX, USA, 14–19 March 2010; pp. 3734–3737.
5. Li, Y.; Hamamura, M. Zero-attracting variable-step-size least mean square algorithms for adaptive sparse channel estimation. *Int. J. Adapt. Control Signal Process.* **2015**, *29*, 1189–1206.
6. Chen, Y.; Gu, Y.; Hero, A.O., III. Sparse LMS for system identification. In Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP '09), Taipei, Taiwan, 19–24 April 2009; pp. 3125–3128.
7. Wang, Y.; Li, Y.; Yang, R. Sparse adaptive channel estimation based on mixed controlled  $l_2$  and  $l_p$ -norm error criterion. *J. Frankl. Inst.* **2017**, *354*, 7215–7239.
8. Li, Y.; Wang, Y.; Jiang, T. Sparse least mean mixed-norm adaptive filtering algorithms for sparse channel estimation applications. *Int. J. Commun. Syst.* **2017**, *30*, 1–16.
9. Wang, Y.; Li, Y. Sparse multi-path channel estimation using norm combination constrained set-membership NLMS algorithms. *Wirel. Commun. Mob. Comput.* **2017**, *2017*, 8140702.
10. Li, Y.; Jin, Z.; Wang, Y. Adaptive channel estimation based on an improved norm constrained set-membership normalized least mean square algorithm. *Wirel. Commun. Mob. Comput.* **2017**, *2017*, 8056126.
11. Duttweiler, D.L. Proportionate normalized least-mean-squares adaptation in echo cancelers. *IEEE Trans. Speech Audio Process.* **2000**, *8*, 508–518.
12. Naylor, P.A.; Cui, J.; Brookes, M. Adaptive algorithms for sparse echo cancellation. *Signal Process.* **2009**, *86*, 1182–1192.
13. Cotter, S.F.; Rao, B.D. Sparse channel estimation via matching pursuit with application to equalization. *IEEE Trans. Commun.* **2002**, *50*, 374–377.
14. Gui, G.; Peng, W.; Adachi, F. Improved adaptive sparse channel estimation based on the least mean square algorithm. In Proceedings of the 2013 IEEE Wireless Communications and Networking Conference (WCNC), Shanghai, China, 7–10 April 2013; pp. 3105–3109.
15. Arenas, J.; Figueiras-Vidal, A.R. Adaptive combination of proportionate filters for sparse echo cancellation. *IEEE Trans. Audio Speech Lang. Process.* **2009**, *17*, 1087–1098.
16. Nekui, M.; Atarodi, M. A fast converging algorithm for network echo cancellation. *IEEE Signal Process. Lett.* **2004**, *11*, 427–430.

17. Li, Y.; Wang, Y.; Jiang, T. Norm-adaption penalized least mean square/fourth algorithm for sparse channel estimation. *Signal Process.* **2016**, *128*, 243–251.
18. Li, Y.; Zhang, C.; Wang, S. Low-complexity non-uniform penalized affine projection algorithm for sparse system identification. *Circuits Syst. Signal Process.* **2016**, *35*, 1611–1624.
19. Stojanovic, M.; Freitag, L.; Johnson, M. Channel-estimation-based adaptive equalization of underwater acoustic signals. In Proceedings of the OCEANS '99 MTS/IEEE, Riding the Crest into the 21st Century, Seattle, WA, USA, 13–16 September 1999; pp. 985–990.
20. Pelekanakis, K.; Chitre, M. Comparison of sparse adaptive filters for underwater acoustic channel equalization/estimation. In Proceedings of the 2010 IEEE International Conference on Communication Systems (ICCS), Singapore, 17–19 November 2010; pp. 395–399.
21. Li, Y.; Wang, Y.; Jiang, T. Sparse-aware set-membership NLMS algorithms and their application for sparse channel estimation and echo cancelation. *AEU Int. J. Electron. Commun.* **2016**, *70*, 895–902.
22. Gui, G.; Mehbodniya, A.; Adachi, F. Least mean square/fourth algorithm for adaptive sparse channel estimation. In Proceedings of the 2013 IEEE 24th International Symposium on Personal Indoor and Mobile Radio Communications (PIMRC), London, UK, 8–11 September 2013; pp. 296–300.
23. Vuokko, L.; Kolmonen, V.M.; Salo, J.; Vainikainen, P. Measurement of large-scale cluster power characteristics for geometric channel models. *IEEE Trans. Antennas Propagat.* **2007**, *55*, 3361–3365.
24. Radecki, J.; Zilic, Z.; Radecka, K. Echo cancellation in IP networks. In Proceedings of the 45th Midwest Symposium on Circuits and Systems, Tulsa, OK, USA, 4–7 August 2002; pp. 219–222.
25. Cui, J.; Naylor, P.A.; Brown, D.T. An improved IPNLMS algorithm for echo cancellation in packet-switched networks. In Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP '04), Montreal, QC, Canada, 17–21 May 2004.
26. Widrow, B.; Stearns, S.D. *Adaptive Signal Processing*; Prentice Hall: Upper Saddle River, NJ, USA, 1985.
27. Wang, Y.; Li, Y. Norm penalized joint-optimization NLMS algorithms for broadband sparse adaptive channel estimation. *Symmetry* **2017**, *9*, 133.
28. Haykin, S. *Adaptive Filter Theory*; Prentice Hall: Upper Saddle River, NJ, USA, 1991.
29. Li, Y.; Hamamura, M. An improved proportionate normalized least-mean-square algorithm for broadband multipath channel estimation. *Sci. World J.* **2014**, *2014*, 572969.
30. Benesty, J.; Gay, S.L. An improved PNLMS algorithm. In Proceedings of the 2002 IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP), Orlando, FL, USA, 13–17 May 2002; Volume II, pp. 1881–1884.
31. Gay, S.L. An efficient, fast converging adaptive filter for network echo cancellation. In Proceedings of the 32nd Asilomar Conference on Signals and System for Computing, Pacific Grove, CA, USA, 1–4 November 1998; Volume 1, pp. 394–398.
32. Liu, W.F.; Pokharel, P.P.; Principe, J.C. Correntropy: Properties and applications in non-Gaussian signal processing. *IEEE Trans. Signal Process.* **2007**, *55*, 5286–5298.
33. Li, Y.; Wang, Y. Sparse SM-NLMS algorithm based on correntropy criterion. *Electron. Lett.* **2016**, *52*, 1461–1463.
34. Chen, B.; Principe, J.C. Maximum correntropy estimation is a smoothed MAP estimation. *IEEE Signal Process. Lett.* **2012**, *19*, 491–494.
35. Chen, B.; Xing, L.; Liang, J.; Zheng, N.; Principe, J.C. Steady-state mean-square error analysis for adaptive filtering under the maximum correntropy criterion. *IEEE Signal Process. Lett.* **2014**, *21*, 880–884.
36. Zhao, S.; Chen, B.; Principe, J.C. Kernel adaptive filtering with maximum correntropy criterion. In Proceedings of the 2011 International Joint Conference on Neural Networks (IJCNN), San Jose, CA, USA, 31 July–5 August 2011; pp. 2012–2017.
37. Chen, B.; Xing, L.; Zhao, H.; Zheng, N.; Principe, J.C. Generalized correntropy for robust adaptive filtering. *IEEE Trans. Signal Process.* **2016**, *64*, 3376–3387.
38. Chen, B.; Wang, J.; Zhao, H.; Zheng, N.; Principe, J.C. Convergence of a fixed-point algorithm under maximum correntropy criterion. *IEEE Signal Process. Lett.* **2015**, *22*, 1723–1727.
39. Singh, A.; Principe, J.C. Using correntropy as a cost function in linear adaptive filters. In Proceedings of the International Joint Conference on Neural Networks, Atlanta, GA, USA, 14–19 June 2009; pp. 2950–2955.

40. Haddad, D.B.; Petraglia, M.R.; Petraglia, A. A unified approach for sparsity-aware and maximum correntropy adaptive filters. In Proceedings of the 24th European Signal Processing Conference (EUSIPCO), Budapest, Hungary, 29 August–2 September 2016; pp. 170–174.
41. Ma, W.; Qu, H.; Gui, G.; Xu, L.; Zhao, J.; Chen, B. Maximum correntropy criterion based sparse adaptive filtering algorithms for robust channel estimation under non-Gaussian environments. *J. Frankl. Inst.* **2015**, *352*, 2708–2727.
42. Wang, Y.; Li, Y.; Albu, F.; Yang, R. Group-constrained maximum correntropy criterion algorithms for estimating sparse mix-noised channels. *Entropy* **2017**, *19*, 432.
43. Wu, Z.; Peng, S.; Chen, B.; Zhao, H.; Principe, J.C. Proportionate minimum error entropy algorithm for sparse system identification. *Entropy* **2015**, *17*, 5995–6006.
44. Salman, M.S. Sparse leaky-LMS algorithm for system identification and its convergence analysis. *Int. J. Adapt. Control Signal Process.* **2014**, *28*, 1065–1072.
45. Li, Y.; Wang, Y.; Yang, R.; Albu, F. A soft parameter function penalized normalized maximum correntropy criterion algorithm for sparse system identification. *Entropy* **2017**, *19*, 45.
46. Wang, Y.; Li, Y.; Yang, R. A sparsity-aware proportionate normalized maximum correntropy criterion algorithm for sparse system identification in non-gaussian environment. In Proceedings of the 25th European Signal Processing Conference (EUSIPCO), Kos Island, Greece, 28 August–2 September 2017; pp. 246–250.
47. Hoyer, P.O. Non-negative matrix factorization with sparseness constraints. *J. Mach. Learn. Res.* **2001**, *49*, 1208–1215.
48. Huang, Y.; Benesty, J.; Chen, J. *Acoustic MIMO Signal Processing*; Springer: Berlin, Germany, 2006.



© 2017 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).