Blind Signal Processing based on Information Theoretic Learning with Kernel-size Modification for Impulsive Noise Channel Equalization

NAMYONG KIM*
*Dep. of Information & Communication Engineering
Kangwon National University
Gyo-dong Samcheok Kangwon Do
KOREA
namyong@kangwon.ac.kr

Abstract: - This paper presents a new performance enhancement method of information-theoretic learning (ITL) based blind equalizer algorithms for ISI communication channel environments with a mixture of AWGN and impulsive noise. The Gaussian kernel of Euclidian distance (ED) minimizing blind algorithm using a set of evenly generated symbols has the net effect of reducing the contribution of samples that are far away from the mean value of the error distribution. The process of ED minimization between desired probability density function (PDF) and output PDF is considered as a harmonious force interaction on PDF shaping between concentrating force and spreading force. The spreading force is composed of the difference between output sample values themselves, and is directly related with the output information potential and the output entropy that leads to the output distribution spreading out. The proposed kernel-size modification scheme is to impose loose discipline on the spreading force by employing larger kernel-size so that the long distance between two outputs which are correct symbol-related and far-located is less likely to be treated as impulsive noise. From the simulation results, the proposed kernel-size modified blind algorithm not only outperforms correntropy blind algorithm, but also significantly enhance robustness against impulsive noise.

Key-Words: - Impulsive noise, blind equalization, kernel-size modification, source symbol assignment, information-theoretic learning, PDF, correntropy, CMA.

1 Introduction

In communication systems such as broadcast and multipoint networks, blind equalizers to counteract multipath effects of channel are very useful since they do not require a training sequence [1][2]. Multipath channels induce severe inter-symbol interference (ISI) and are contaminated by noise; not only Gaussian noise but also impulsive noise from a variety of impulse noise sources [3][4]. Impulsive noise induces large instantaneous system output and error which makes the system fail to produce desirable performance.

Recently introduced information theoretic learning (ITL) method has shown superior performance as an alternative to mean squared error (MSE) method in equalization applications [5][6]. The most commonly used and MSE-based blind equalization algorithm is constant modulus algorithm (CMA), which is designed to minimize the average of the constant modulus error between equalizer output power and constant modulus [7]. However, CMA often fails to converge in impulsive noise environment. The correntropy blind method [8] based on ITL has shown better results in partial response systems with impulsive noise environments compared to CMA, but has not yielded satisfying results in PAM systems [9].

Our initial approach to blind equalization for improved impulsive-noise resistance was to minimize the Euclidian distance of probability density function (PDF) between output PDF and desired PDF by using a set of evenly generated source symbol values at the receiver according to the modulation scheme [9]. The process of ED minimization between desired PDF and output PDF is considered as a harmonious force interaction on PDF shaping between spreading force and concentrating force. The spreading force is composed of the difference between output sample values themselves. Based on this concept, we propose a kernel-size modified version of our initial blind algorithm in order to reduce the spreading force so that the long distance between the positively largest output and negatively largest output has less chance of being considered as impulsive noise.

This paper is organized as follows. In Section 2, we describe the impulsive noise model composed of the background Gaussian noise and impulse noise. The constant modulus blind equalizer algorithm which is based on MSE criterion is briefly introduced in Section 3. In Section 4, a recent blind method for partial response system based on correntropy for impulsive noise environments is described. Our initial blind equalization algorithm for impulsive-noise resistance is explained in Section 5. In Section 6 we propose a kernel-size modified blind algorithm, and Section 7 reports simulation results and discussions. Finally, concluding remarks are presented in Section 8.

2 Impulsive Noise Model

The impulsive channel noise model in this paper is composed of the background Gaussian noise and impulse noise. The background noise is AWGN of which variance is σ_{GN}^2. The impulsive noise occurs according to a Poisson process and the average number of Poisson occurrence impulses per information symbol duration is defined as ε. The amplitude distribution of impulsive noise has a Gaussian with variance σ_{IN}^2. This noise model is widely used as an impulsive noise model in [8][10].

The PDF of the background AWGN is expressed as

$$f_{GN}(\xi) = \frac{1}{\sigma_{GN}\sqrt{2\pi}} \exp[\frac{-\xi^2}{2\sigma_{GN}^2}] \quad (1)$$

The impulsive noise with Gaussian amplitude has the PDF expression as

$$f_{IN}(\xi) = (1-\varepsilon)\cdot\delta(\xi) + \frac{\varepsilon}{\sigma_{IN}\sqrt{2\pi}} \exp[\frac{-\xi^2}{2\sigma_{IN}^2}] \quad (2)$$

The total noise is a sum of the two random processes and the PDF form of the total noise is obtained by taking the convolution of (1) and (2). From the convolution process we obtain the following total noise PDF expression.

$$f_{NOISE}(\xi) = \frac{1-\varepsilon}{\sigma_1\sqrt{2\pi}} \exp[\frac{-\xi^2}{2\sigma_1^2}] + \frac{\varepsilon}{\sigma_2\sqrt{2\pi}} \exp[\frac{-\xi^2}{2\sigma_2^2}] \quad (3)$$

where $\varepsilon < 1$, $\sigma_1 = \sigma_{GN}$, $\sigma_2 = \sqrt{\sigma_{GN}^2 + \sigma_{IN}^2}$, and $\sigma_1^2 << \sigma_2^2$.

3 CMA based on MSE Criterion

For a tapped delay line (TDL) equalizer with weight vector W of L elements in training-aided equalization, error sample e_k between the desired training symbol d_k and output y_k are produced by $e_k = d_k - y_k = d_k - W_k^T X_k$ at time k where the equalizer input vector is $X_k = [x_k, x_{k-1}, x_{k-2}, ..., x_{k-L+1}]^T$. Channel equalization without the aid of a training sequence d_k is referred to as blind channel equalization. One of the well known blind equalization algorithms is CMA which minimizes the CMEs based on MSE criterion.

$$P_{CMA} = E[(|y_k|^2 - R_2)^2] \quad (4)$$

where the constant modulus $R_2 = E[|d_k|^4]/E[|d_k|^2]$.

For Gaussian noise cases, averaging CMEs taken from different time instants discards the effects of the Gaussian noise, but a single large, impulsive noise sample can dominate these sums and defeat the averaging.

According to the steepest descent method with the step-size parameter μ_{CMA}, we obtain the following CMA for adjusting the blind equalization [7]:

$$W_{k+1} = W_k - 2\mu_{CMA} \cdot X_k^* \cdot y_k \cdot (|y_k|^2 - R_2) \quad (5)$$

4 Correntropy Blind Algorithm

Since inner products are a measure of similarity, the pair-wise interaction of the feature vectors, that is, inner product of vectors in a kernel feature space where feature vectors are separated by a certain time delay in input space can be another measure of similarity that can be utilized in signal processing applications.

Kernel algorithms transform the data X_i from the input space to a high dimensional feature space of vectors $\Phi(X_i)$, where the inner products $\langle \cdot,\cdot \rangle$ can be computed using a positive definite kernel function satisfying Mercer's conditions [8]:

$$K(X_i, X_j) = \langle \Phi(X_i), \Phi(X_j) \rangle \tag{6}$$

This makes it possible to obtain nonlinear versions of any linear algorithm expressed in terms of inner products. In this case the knowledge of the exact mapping function Φ is no longer needed.

Let $\{X(t), t \in T\}$ be a stochastic process with being an index set and the nonlinear mapping Φ induced by the Gaussian kernel maps the data into the feature space F, where F is an infinite dimensional reproducing kernel Hilbert space so the following equation holds

$$G_\sigma((X(t) - X(s)) = \langle \Phi(X(t)), \Phi(X(s)) \rangle_F \tag{7}$$

where $G_\sigma(\cdot)$ is a zero-mean Gaussian kernel with standard deviation σ. Then the auto correntropy function $V_X(t,s)$ [8] is defined as

$$\begin{aligned} V_X(t,s) &= E[\langle \Phi(X(t)), \Phi(X(s)) \rangle_F] \\ &= E[G_\sigma(X(t) - X(s))] \end{aligned} \tag{8}$$

where $E[\cdot]$ denotes statistical expectation. This similarity measure has the analogy with the autocorrelation of two random processes and the property that its average over the lags [11]. For a discrete-time stationary stochastic process, the correntropy function with the time lag m is defined as

$$V_X[m] = E[G_\sigma(X_k - X_{k-m})] \tag{9}$$

And through the sample mean, it can be estimated using N samples as

$$V_X[m] = \frac{1}{N-m+1} \sum_{k=m}^{N} G_\sigma(X_k - X_{k-m}) \tag{10}$$

Since the correntropy function conveys information about the PDF and correlation of the signal, the authors in [8] proposed the following cost function.

$$P_{CE} = \sum_{m=1}^{M} (V_S[m] - V_Y[m])^2 \tag{11}$$

where $V_S[m]$ is the source correntropy, $V_Y[m]$ is the equalizer output correntropy, and M is the number of lags. The cost function can be minimized by using a gradient descent approach, and then the correntropy blind algorithm is obtained as

$$\begin{aligned} W_{k+1} = W_k - \mu_{CE} \frac{1}{(N-m+1)\sigma^2} \sum_{m=1}^{M} \sum_{i=k-N+m}^{k} (V_S[m] \\ -V_Y[m]) \cdot G_\sigma(y_i - y_{i-m}) \cdot (y_i - y_{i-m})(X_i - X_{i-m}) \end{aligned} \tag{12}$$

To use correntropy algorithm (12) for blind equalizer weight update, first we must obtain the theoretical correntropy function value for a given source signal. For the 4 PAM i.i.d. source signal $\{-3,-1,+1,+3\}$ as an example, the theoretical source correntropy $V_S[m]$ is calculated as

$$V_S[m] = \begin{cases} G_\sigma(0) & ; m = 0 \\ \frac{1}{4}G_\sigma(0) + \frac{3}{8}G_\sigma(2) + \frac{1}{4}G_\sigma(4) + \frac{1}{8}G_\sigma(6) & ; |m| \geq 1 \end{cases} \tag{13}$$

5 PDF-Distance Minimizing Blind Algorithm using A Set of Assigned Symbol Values

The Euclidian distance [5] between the source (transmitted) symbol PDF $f_D(\xi)$ and the equalizer output PDF $f_Y(\xi)$ is defined as

$$\begin{aligned} &ED[f_D(\xi), f_Y(\xi)] = \int [f_D(\xi) - f_Y(\xi)]^2 d\xi \\ &= \int f_D^2(\xi) d\xi - 2\int f_D(\xi) f_Y(\xi) d\xi + \int f_Y^2(\xi) d\xi \end{aligned} \tag{13}$$

Minimization of (13) leads to PDF matching between the two PDFs, $f_D(\xi)$ and $f_Y(\xi)$. Using Parzen window method $f_X(\xi) = \frac{1}{N} \sum_{i=1}^{N} G_\sigma(\xi - x_i)$ can be used to estimate the PDF with N samples [5]. We note that the first term of (13) is not a function of weight. So the cost function can be composed of the following two terms:

$$\int f_D(\xi) f_Y(\xi) d\xi = \frac{1}{N^2} \sum_{i=1}^{N} \sum_{j=1}^{N} G_{\sigma\sqrt{2}}(d_j - y_i) \tag{14}$$

$$\int f_Y^2(\xi)d\xi = \frac{1}{N^2}\sum_{i=1}^{N}\sum_{j=1}^{N} G_{\sigma\sqrt{2}}(y_j - y_i) \qquad (15)$$

The desired symbols are unknown in blind equalization, so we propose to assign transmitted source levels to N desired symbols evenly without knowing the exact desired values. We assume that Q-ary PAM signaling systems are employed and the all Q levels are equally likely to be transmitted a priori with a probability $1/Q$, and the transmitted levels A_q takes the following discrete values

$$A_q = 2q - 1 - Q,\ q = 1,2,...,Q. \qquad (16)$$

The level value of A_1 is assigned to N/Q symbols of d_1, the level value of A_2 is assigned to N/Q symbols of d_2, and so on. Now a sliding-window for output samples and a gradient descent method for the minimization of the cost function (13) can be applied to obtain the following our initial algorithm utilizing evenly assigned level values in place of desired symbols as in [9].

$$\begin{aligned} W_{k+1} = W_k + \mu[&\frac{1}{N^2\sigma^2}\sum_{i=k-N+1,}^{k}\sum_{j=1}^{N}(d_j - y_i) \cdot G_{\sigma\sqrt{2}}(d_j - y_i)\cdot X_i \\ &- \frac{1}{2N^2\sigma^2}\sum_{i=k-N+1,}^{k}\sum_{j=k-N+1}^{k}(y_j - y_i) \cdot G_{\sigma\sqrt{2}}(y_j - y_i)\cdot(X_i - X_j)] \end{aligned} \qquad (17)$$

For convenience sake, this algorithm will be referred to in this paper as initial algorithm.

6 PDF-Distance Minimizing Blind Algorithm with Kernel-size Modification

In the previous section, in order to match output PDF and source PDF without knowing the exact source symbol values, we constructed the desired symbol PDF by evenly assigning transmitted levels to N desired symbols which are required in Parzen window method. In the process of minimizing ED, the two terms (14) and (15) play the role of cost function. In this section we investigate the specific role of each term and search for ways to enhance robustness against impulsive noise and ISI cancellation performance.

We note that (15) is referred to as information potential (IP) of output [5][12]. To understand the effect of increasing or decreasing output information potential, we need to mention relationship between information potential and entropy.

Entropy is a scalar quantity that provides a measure of the average information contained in a given PDF [13]. When output entropy is maximized, the output distribution of adaptive systems gets spread out. Renyi's quadratic entropy [14][15] as a useful tool for entropy calculation is defined as

$$E(x) = -\log(\int f_X^2(\xi)d\xi) \qquad (18)$$

Substituting output information potential $IP_{YY} = \int f_Y^2(\xi)d\xi$ into (18), we obtain

$$E(y) = -\log(IP_{YY}) \qquad (19)$$

Obviously, minimizing the output information potential IP_{YY} is equivalent to maximizing the output entropy $E(y)$. This leads to the output distribution spreading out.

Since minimization of (13) is equal to maximization of $\int f_D(\xi) f_Y(\xi)d\xi$ and minimization of IP_{YY} simultaneously, PDF-distance minimizing process between $f_D(\xi)$ and $f_Y(\xi)$ can be considered as a harmonious force interaction on the PDF shape between spreading force and concentrating force, where spreading force comes from IP_{YY} minimization and concentrating force from maximization of $\int f_D(\xi) f_Y(\xi)d\xi$.

Noting the term $IP_{YY} = \int f_Y^2(\xi)d\xi$ in (15) is related with $G_{\sigma\sqrt{2}}(y_j - y_i)$, we can notice that the long distance between the two output points most far away from each other, induced by A_1 and A_H, can be treated as impulsive noise and cut out passing through the Gaussian kernel. Based on this observation, we can come up with the idea of reducing the effect of cutting

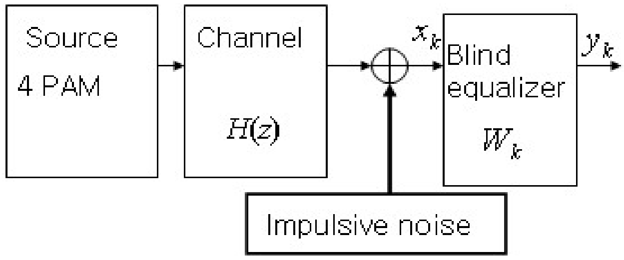

Fig. 1. System model with impulsive noise

out the difference between extreme outputs by means of employing a larger kernel size. This approach also has the effect of diminishing the effect of the spreading force on the PDF matching process.

There can be many approaches to increasing the kernel size in (15), in this paper, empirically-based, we propose to replace $G_{\sigma\sqrt{2}}(y_j - y_i)$ in (15) with $G_{2\sigma}(y_j - y_i)$ as expressed in (20).

$$\int f_Y^2(\xi)d\xi = \frac{1}{N^2}\sum_{i=1}^{N}\sum_{j=1}^{N} G_{2\sigma}(y_j - y_i) \qquad (20)$$

We chose 2σ in (20) as the best kernel size after testing the performance with various kernel sizes, which will be shown in Section 7.

Then the new PDF-distance minimizing blind algorithm with kernel-size modification becomes

$$\begin{aligned} W_{k+1} = W_k + \mu[\frac{1}{N^2\sigma^2} \sum_{i=k-N+1,}^{k} \sum_{j=1}^{N} (d_j - y_i) \\ \cdot G_{\sigma\sqrt{2}}(d_j - y_i) \cdot X_i \\ - \frac{1}{2N^2\sigma^2} \sum_{i=k-N+1,}^{k} \sum_{j=k-N+1}^{k} (y_j - y_i) \\ \cdot G_{2\sigma}(y_j - y_i) \cdot (X_i - X_j)] \qquad (21) \end{aligned}$$

where $\{d_i\}$ is a set of self-created and evenly-assigned source symbol values according to modulation schemes and $\{y_i\}$ is a sliding-window block of equalizer output samples. We will refer to (21) as kernel-modified algorithm for the sake of convenience.

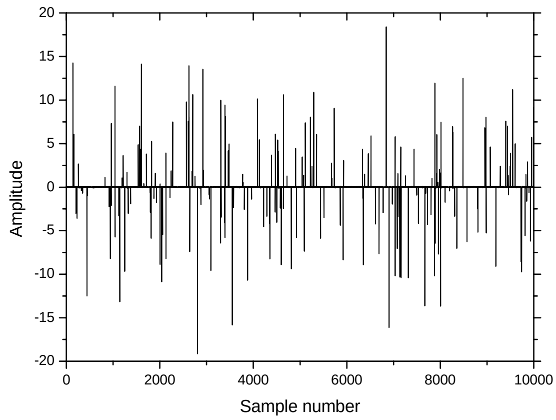

Fig. 2. Impulsive noise for the simulation.

7 Results and Discussion

In this section, we test the CMA, blind correntropy, the initially proposed algorithm and the kernel-size modified algorithm in linear radio channels and compare their MSE convergence and error distribution performance.

As depicted in Fig. 1, we consider 4 level PAM $\{-3,-1,+1,+3\}$ source signal distorted by a radio channel model as

$$H_1(z) = 0.26 + 0.93z^{-1} + 0.26z^{-2} \qquad (22)$$

Then the channel output signal is added with a zero-mean white impulsive noise n_k, generated according to the following Gaussian mixture model as described in previous section:

$$f_{NOISE}(n_k) = \frac{1-\varepsilon}{\sigma_{GN}\sqrt{2\pi}} \exp[\frac{-n_k^2}{2\sigma_{GN}^2}] + \frac{\varepsilon}{\sigma_2\sqrt{2\pi}} \exp[\frac{-n_k^2}{2\sigma_2^2}] \qquad (23)$$

where we used $\varepsilon = 0.03$, $\sigma_{GN}^2 = 0.001$, and $\sigma_2^2 = \sigma_{GN}^2 + \sigma_{IN}^2 = 50.001$. The impulsive noise signal with those parameters is depicted in Fig. 2. We can observe large impulses with above the magnitude of 19 voltages that will be added to the equalizer input signal.

Fig. 3 shows the MSE convergence curves. A 11-tap TDL equalizer is used and initialized with the center weight set to unity and the rest to zero. The step-size for initially proposed and kernel-modified

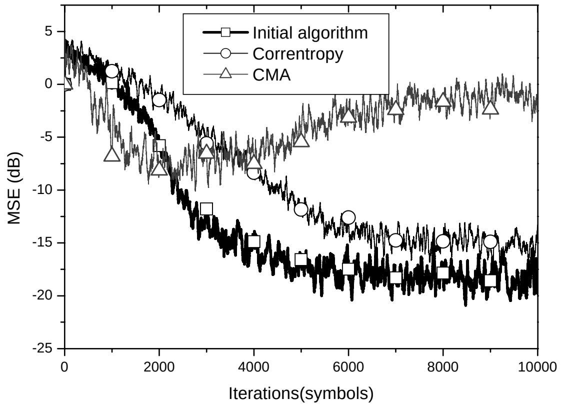

Fig. 3. MSE convergence performance under impulsive noise.

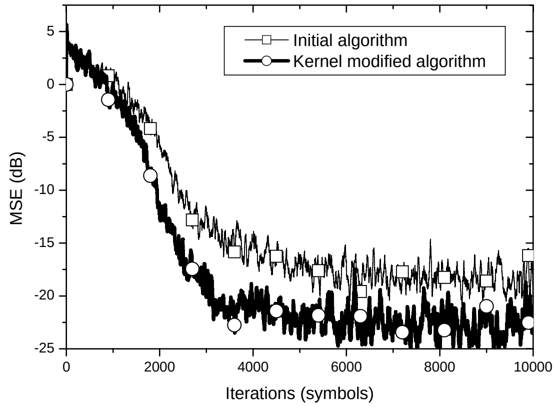

Fig. 5. MSE performance comparison between the initially proposed algorithm and the kernel-modified algorithm.

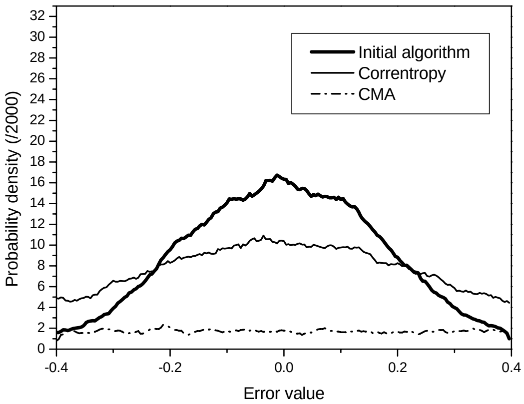

Fig. 4. Probability density for errors under impulsive noise.

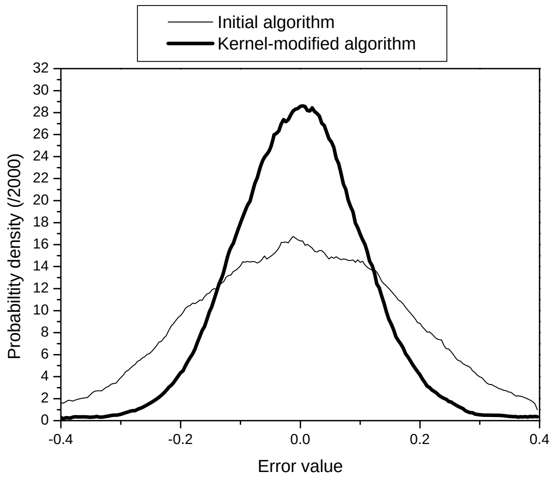

Fig. 6. Performance comparison of probability density for errors between the initially proposed algorithm and the kernel-modified algorithm.

algorithm is commonly set to 0.007. The step-size for blind correntropy algorithm and CMA is $\mu_{CE} = 0.01$ and $\mu_{CMA} = 0.000001$, respectively.

The number of lags is $M = 20$ and data-block size $N = 32$. And the numerical kernel size σ for the initially proposed and kernel-modified algorithm is commonly 0.6 and that for correntropy is 2.8. We see that CMA fails to converge even for the small step-size. On the other hand, the blind algorithms based on ITL converge well.

Compared to the correntropy algorithm, the MSE curve for the initially proposed algorithm reaches lower steady state MSE, and the minimum MSE performance enhancement is above 3 dB. Figure 4 depicts the estimated error probability densities of the algorithms. Their performance differences are shown more clearly. The error values of CMA appear not to gather well around zero, but correntropy and initially proposed algorithm produce error distributions still concentrated around zero. Clearly, the latter algorithm yields better equalizer-error PDF performance.

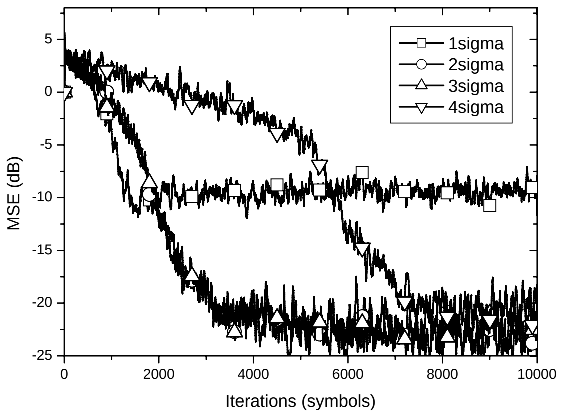

Fig. 7. MSE performance comparison of kernel-modified algorithm for four different sizes; 1sigma= σ , 2sigma= 2σ , 3sigma= 3σ and 4sigma= 4σ , respectively.

To investigate the effect of kernel-size modification, we compare the initial algorithm with the kernel-modified algorithm under the same simulation environment. In Fig. 5, the MSE curve for the kernel-modified algorithm reaches about -24 dB of steady state MSE and the initial algorithm does about -18.5 dB. The kernel modification method produces MSE performance improvement of about 5.5 dB.

Besides the performance enhancement of steady-state MSE, the learning speed is also improved significantly. The initial algorithm is considered to converge in about 5000 iterations, and the kernel-modified algorithm converges in around 3500 iterations.

In Fig. 6, the estimated error probability densities of those two algorithms are depicted. We can see their performance differences more clearly. The error values of the initially proposed algorithm gather well around zero, and moreover the kernel-modified algorithm produces superior equalizer-error PDF performance to the initially proposed algorithm.

Now we compare the performance difference when we vary the kernel size of (20). We test the performance for four different sizes; σ, 2σ, 3σ and 4σ, and depicted the results in Fig. 7.

In Fig. 7, the smaller kernel size σ than that of the initial algorithm, $\sigma\sqrt{2}$, produces very poor MSE convergence and it keeps the minimum MSE staying above -10 dB. This proves that our approach of increasing the kernel size of the spreading force is

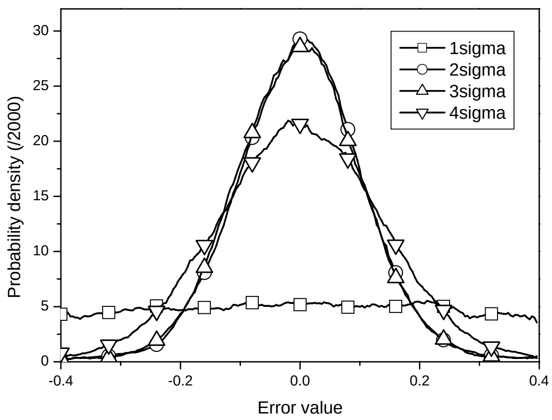

Fig. 8. Error PDF comparison of kernel-modified algorithm for four different sizes; 1sigma= σ , 2sigma= 2σ , 3sigma= 3σ and 4sigma= 4σ , respectively.

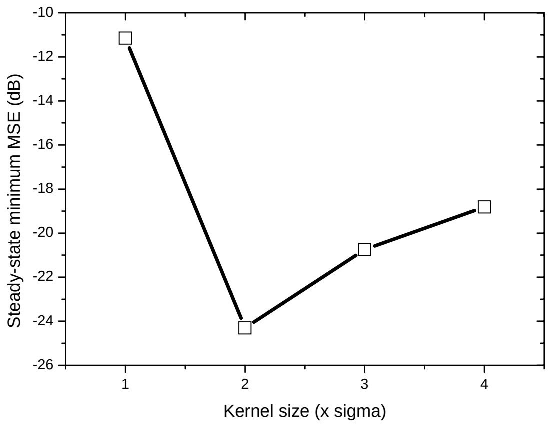

Fig. 9. Steady-state minimum MSE versus kernel-sizes; in the horizontal axis, 1= σ , 2= 2σ , 3= 3σ and 4= 4σ , respectively.

reasonable. However, increasing the kernel size of the spreading force indefinitely does not guarantee good performance since impulsive-noise induced large difference between two outputs is more likely to be counted in the spreading-force related Gaussian kernel function with a far larger kernel size than an acceptable one. From the modified size 4σ , the kernel

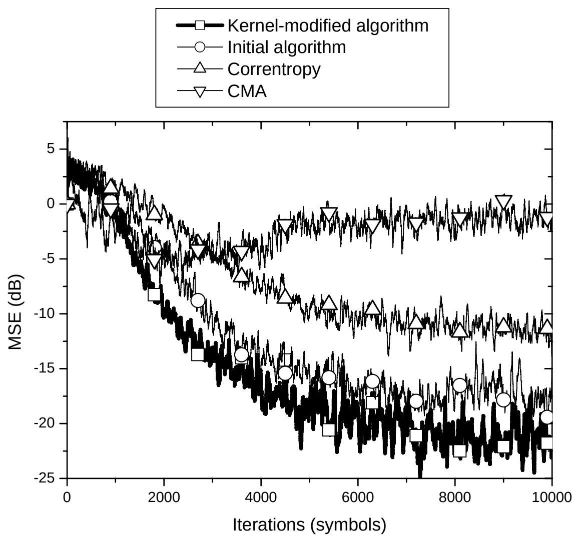

Fig. 10. MSE convergence performance in the channel model $H_2(z)$ under impulsive noise.

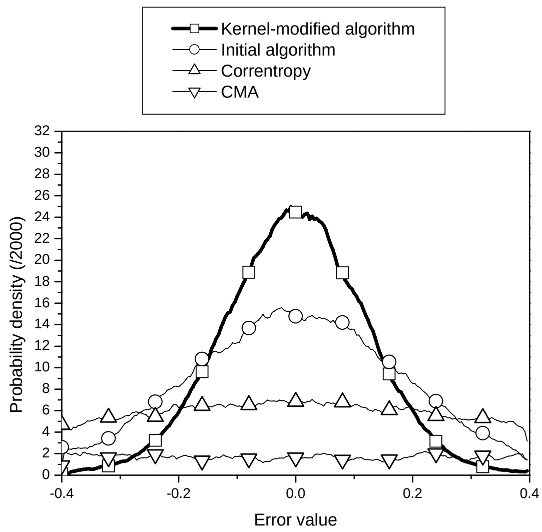

Fig. 11. Probability density for errors in the channel model $H_2(z)$ under impulsive noise.

size degrades the performance producing slow convergence and increased steady state MSE. The kernel size 2σ or 3σ yield better performance, and we can find 2σ the best for the simulation environment by examining the error PDF performance as depicted in Fig. 8. The case of 2σ is slightly better than 3σ, but 4σ results in downgraded error PDF performance as observed in Fig. 8. This result is verified more clearly in Fig. 9 where steady-state minimum MSE values with various kernel sizes are presented. Compared to correntropy algorithm in Fig. 3, the case of smaller kernel size, σ than that of the initial algorithm brings about much worse performance though it produces better performance than CMA. This implies that the Gaussian kernel of spreading force cuts out excessively the difference between outputs, that is, the long distance between the two furthest output points, induced by A_1 and A_H, is more likely to be treated as impulse noise. This is considered to yield worse performance.

Opposite to small kernel size for the spreading force, large kernel size is considered to make the Gaussian kernel function pass some extent of impulsive noise as well as large output differences. This causes the equalizer to lose the function of cutting out outliers. According to this analysis, we observe slow convergence and deteriorated MSE performance as the kernel size gets bigger than 2σ in Fig. 7 and 9.

In order to investigate their performance for different channel models, we perform the same experiment in a severer channel model $H_2(z)$ as

$$H_2(z) = 0.304 + 0.903z^{-1} + 0.304z^{-2} \qquad (24)$$

While the eigen-value spread ratio of the channel model $H_2(z)$ is 21, the previous channel model $H_1(z)$ in (22) has an eigen-value spread ratio of 11. The eigen-value spread (ES) indicates the amount of multipath in the channel, ranging from large values for nearly line-of-sight channels to lower values for channels with richer multipath. Eigen-value spread represents the eigen-value mean and implies how much severe the channel model is.

In Fig. 10, we obtained similar MSE convergence and error PDF performance to the case of $H_1(z)$ but a little slower convergence and higher minimum MSE. The three ITL-type algorithms, Correntroy, initial algorithm and the proposed kernel-modified algorithm converge in about 7000 iterations and still the kernel-modified algorithm shows better performance. While the initial algorithm and correntropy algorithm have the steady-state minimum MSE of -12 dB and -17 dB, respectively, the proposed algorithm reaches the steady-state minimum MSE of -22 dB. We can

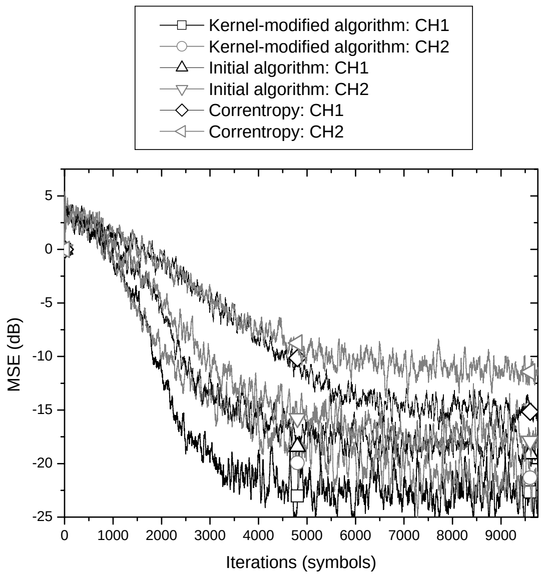

Fig. 12. MSE performance comparison between algorithms in the channel model $H_1(z)$ (CH1) and in the channel model $H_2(z)$ (CH2) under impulsive noise.

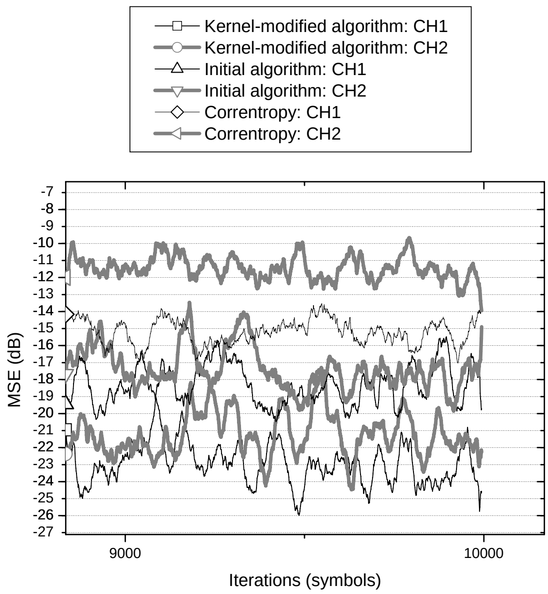

Fig. 13. Steady-state minimum MSE comparison between algorithms in the channel model $H_1(z)$ (CH1) and in the channel model $H_2(z)$ (CH2) under impulsive noise.

observe the superiority of the proposed kernel-modified algorithm even in the worse channel environment more clearly in the error PDF comparison depicted in Fig. 11. While the error PDFs of correntropy approaches almost flat PDF shape, the initial algorithm keeps its PDF in a bell-shape and more interestingly the proposed kernel-modified algorithm shows a more zero-concentrated PDF shape. This indicates that error from the proposed equalizer is highly concentrated around zero so that it is less likely to be produced as large values. As we will investigate more the robustness against channel eigen-spread variations through the comparison of minimum MSE performance in the two channel models in Fig. 12, the comparison of error PDF performance shown in Fig. 8 and 11 reveals that the proposed kernel-modified algorithm is very robust to severe channel conditions.

In Fig. 12, the MSE convergence comparison for two channel models $H_1(z)$ (described as CH1) and $H_2(z)$ (described as CH2) is shown. The initial algorithm does not show much difference in convergence speed, but the convergence speed of the proposed kernel-modified algorithm depends on the channel conditions. The correntropy algorithm reveals more dependency on the channel conditions in convergence speed and minimum MSE as well. To compare their minimum MSE performance upon the channel conditions in more detail, the steady-state MSE part after convergence for the two channel models is shown in Fig. 13. We see in Fig. 13 that the minimum MSEs of correntropy algorithm are -15.5 for $H_1(z)$ and -11.5 for $H_2(z)$. The difference of minimum MSE is about 4 dB. On the other hand, the minimum MSEs of initial algorithm are -18.5 for $H_1(z)$ and -17.5 for $H_2(z)$. The difference of minimum MSE for the initial algorithm is only about 1 dB. the minimum MSEs of initial algorithm are -24 dB for $H_1(z)$ and -22 dB for $H_2(z)$. The difference of minimum MSE for the proposed kernel-modified algorithm is about 2 dB. From this observation, we can judge that correntropy algorithm is considered sensitive to channel conditions but the initial algorithm and the proposed kernel-modified algorithm are very robust against the variation of channel characteristics.

8 Conclusion

This paper presented a new performance enhancement method of ITL-based blind equalizer algorithms for ISI communication channel environments with a mixture of AWGN and impulsive noise. MSE-based

CMA fails in impulsive noise environment but the ITL methods using Gaussian kernel have the net effect of reducing the contribution of samples that are far away from the mean value of the error distribution. The correntropy blind method based on ITL has shown better results in PAM communication system model with impulsive noise environments compared to CMA, but the ED minimization blind equalizer algorithms produce significant performance enhancement.

PDF-distance minimizing process between desired PDF $f_D(\xi)$ and output PDF $f_Y(\xi)$ can be considered as a harmonious force interaction on the PDF shape between spreading force and concentrating force. Since the spreading force is related with $G_{\sigma\sqrt{2}}(y_j - y_i)$, we noticed that the long distance between the two output points that are not contaminated by impulsive noise but located most far away from each other can be treated as impulsive noise and cut out passing through the Gaussian kernel. As a way of reducing the effect of cutting out the difference between extreme outputs, equivalently diminishing the effect of the spreading force on the PDF matching process, the approach of modifying kernel size of the Gaussian kernel in the spreading force yielded superior equalization performance in impulsive noise environment. Through the experiment for worse channel conditions, we observed that correntropy algorithm is sensitive to channel conditions but the initial algorithm and the proposed kernel-modified algorithm are very robust against the variation of channel characteristics.

From these results and observations, we conclude that the kernel modification scheme applied to the ED minimization-blind algorithm can not only outperform the CMA and correntropy blind algorithm, but also significantly enhance robustness against impulsive noise and channel inferiority.

References:

[1] L. M. Garth, A dynamic convergence analysis of blind equalization algorithms, *IEEE Trans. Comm.*, Vol. 49, April 2001, pp. 624-634.

[2] F. Mazzenga, Channel estimation and equalization for M-QAM transmission with a hidden pilot sequence, *IEEE Trans. Broadcasting*, Vol. 46, June 2000, pp. 170-176.

[3]H. Sedarat and K. Fisher, Multicarrier communication in presence of biased-Gaussian noise sources, *Signal Processing*, Vol. 88, July 2008, pp. 1627-1635.

[4] S. Unawong, S. Miyamoto, and N. Morinaga, A novel receiver design for DS-CDMA systems under impulsive radio noise environments, *IEICE Trans. Comm.*, Vol. E82-B, June 1999, pp. 936 -943.

[5] J.C. Principe, D. Xu and J. Fisher, Information Theoretic Learning, in: S. Haykin, *Unsupervised Adaptive Filtering*, Wiley, New York, Vol. I, 2000, pp. 265-319.

[6] D. Erdogmus, Y. Rao and J. C. Principe, Supervised training of adaptive systems with partially labeled data, *Proceedings of the International Conference on ASSP*, Apr. 2005. pp. v321-v324.

[7] J. R. Treichler and B. Agee, A new approach to multipath correction of constant modulus signals, *IEEE Trans. ASSP*., Vol. ASSP-31, Nov. 1983, pp. 349-372.

[8] I. Santamaria, P. P. Pokharel, and J. C. Principe, Generalized correlation function: Definition, properties, and application to blind equalization, *IEEE Trans. Signal Processing*, Vol.54, June 2006, pp. 2187-2197.

[9] N. Kim and H. Byun, Blind equalization based on information theoretic learning for impulsive noise, *Proceedings of the 4th European Computing Conference (ECC'10)*, Apr. 2010. pp. 69-72.

[10] K. Koike and H. Ogiwara, Application of turbo TCM codes for impulsive noise channel, *IEICE Trans. Fundamentals,* Vol. E81-A, Oct. 1998, pp. 2032-2039.

[11] Weifeng Liu, P. P. Pokharel, and J. C. Principe, Correntropy: properties and applications in non-Gaussian signal Processing, *IEEE Trans. Signal Processing*, vol. 55, no. 11, Nov. 2007, pp. 5286-5298.

[12] D. Erdogmus and J. C. Principe, Generalized information potential criterion for adaptive system training, *Trans. on Neural Networks*, Vol. 13, Sept. 2002, pp. 1035-1044.

[13] C. E. Shannon, A mathematical theory of communication, *Bell Sys. Tech. J*, Vol. 27, 1948, pp. 379-423, 623-653.

[14] D. Erdogmus, and J.C. Principe, An entropy minimization algorithm for supervised training of nonlinear systems, *IEEE Trans. Signal Processing*, Vol. 50, July, 2002, pp. 1780-1786.

[15] A. Renyi, On measures of entropy and information, *Selected Papers of Alfred Renyi*,

Akademia Kiado, Budapest, Vol. 2. 1976, pp. 565-580.

[16] L. Wasserman, *All of Statistics: A Concise Course in Statistical Inference*, Springer Texts in Statistics, 2005.