Adaptive Blind Equalizers with Whitening Filters

Wei-Chieh Chang
Graduate Institute of Applied Science and Engineering
Fu Jen Catholic University, Taiwan, R.O.C.
402068076@mail.fju.edu.tw

Jenq-Tay Yuan
Department of Electrical Engineering
Fu Jen Catholic University, Taiwan, R.O.C.
yuan@ee.fju.edu.tw

***Abstract*—Owing to the multimodality of the constant modulus (CM) cost surfaces, poor initializations of the adaptive blind equalizer (BE) employing the constant modulus algorithm (CMA) may converge to undesired local minima, thus preventing the adaptive BE from opening eye. This paper conducts theoretical analyses via the CM cost surfaces and demonstrates that the use of the BE in tandem with a whitening filter (WF) may still open eye even if the BE based on the CM criterion converges to an undesired local minimum, whereas in the same situation the use of the BE alone may not achieve eye opening. Computer simulations were also conducted to verify our results.**

Keywords—Blind equalization, constant modulus algorithm (CMA), minimum output energy (MOE) criterion, whitening filter (WF)

I. Introduction

Constant modulus (CM) criterion is the most widely used blind equalization criterion which can successfully mitigate the effect of channel by employing the knowledge of the known constellation of the transmitted signal [1], [2]. Labat, Macchi and Loat [3] proposed an adaptive blind decision feedback equalizer (DFE) employing an infinite impulse response (IIR) all-pole whitening filter (APWF) followed by a blind equalizer (BE). Consequently, the whitening filter (WF) based on the minimum output energy (MOE) criterion may produce an uncorrelated input sequence to the BE to increase the convergence speed of the BE based on the CM criterion. Lim, Kennedy and Abhayapala [4] then proposed an adaptive blind DFE employing a finite impulse response (FIR) all-zero whitening filter (AZWF) to replace the APWF in [3] in order to avoid some of the potential problems with adaptation of the IIR WFs in [3]. More recently, an adaptive blind DFE employing an IIR zero-pole whitening filter (ZPWF) has been proposed in [5] to replace the APWF in [3] and the AZWF in [4] since the ZPWF may not only reduce computational complexities but also improve performance of the adaptive blind DFE in terms of both the rate of convergence and the steady-state mean-squared error (MSE).

In order to analyze both the transient behavior and the steady-state performance of filters using the stochastic gradient descent (SGD) algorithms to minimize the cost functions, the surface plots of the cost versus the coefficients of the filters are widely used, such as the output error surface in the system identification field [6]-[8], and the CM and MSE cost surfaces in the channel equalization field [2]. This paper employs the CM cost surfaces to analyze the BE employing the WF as a pre-processing filter. Analyses of the BE based on the CM criterion in the previous papers [9]-[11] were conducted under the condition of the input sequence of the BE being white owing to the use of the WF followed by the BE. However, in practice, the output of the WF may not be perfectly white unless the chosen WF matches perfectly the underlying channel, which is usually unknown. Moreover, since the CM surface is multimodal, the choice of initialization of the BE affects both the convergence speed and the steady-state performance [2]. If the adaptive BE chooses an inappropriate initialization, it may be temporarily captured by saddle points and/or may converge to an undesired local minimum. To our knowledge, this paper demonstrates for the first time that the use of the BE in tandem with the WF may still very likely open eye even if the adaptive BE converges to an undesired local minimum. Moreover, this paper analytically verifies that the WF followed by the BE has better steady-state performance than the use of the BE alone from the CM surfaces without assuming that the input sequence of the BE is white. Computer simulations were also conducted to verify that the WF followed by the BE has a faster convergence speed than that of the use of the BE alone.

II. Background

A. System Model

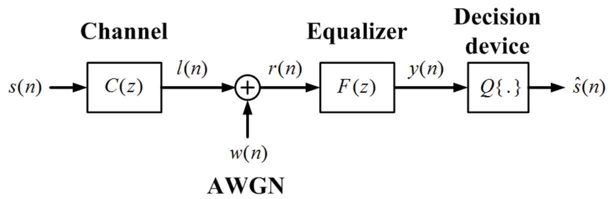

Fig. 1. Baseband model of SISO communication systems.

The baseband single input and single output (SISO) communication system model is depicted in Fig. 1 in which the transmitted source symbol, $\{s(n)\}$, is assumed to be independent and identically distributed (i.i.d.) with variance σ_s^2. The channel noise, $\{w(n)\}$, is assumed to be the additive white Gaussian noises (AWGN) with variance σ_w^2, and the source symbol is assumed to be uncorrelated with the channel noise. The z-transform of the non-minimum phase IIR autoregressive-moving-average (ARMA) channel model can be expressed as $C(z)$ such that [12]

$$C(z) = \frac{C_Z(z)}{C_P(z)} = \frac{\sum_{i=0}^{M_Z} c_{Z,i} \cdot z^{-i}}{1+\sum_{j=1}^{M_P} c_{P,j} \cdot z^{-j}} \tag{1}$$

where $c_{Z,i}$, $i = 0,\ldots,M_Z$, represent the coefficients of the M_Z^{th}-order numerator, and $c_{P,j}$, $j = 1,\ldots,M_P$, represent the coefficients of the M_P^{th}-order denominator. The ARMA channel model in (1) can be reduced to the moving-average (MA) channel model and the autoregressive (AR) channel model by setting $M_P = 0$ and $M_Z = 0$, respectively. As depicted in Fig. 1, the received input, $r(n)$, can then be expressed as

$$r(n)=\left[\sum_{i=0}^{M_Z} c_{Z,i}\cdot s(n-i)-\sum_{j=1}^{M_P} c_{P,j}\cdot l(n-j)\right]+w(n) \quad (2)$$

where $l(n)$ is the channel output. The output of the equalizer, $F(z)$, is $y(n)$, which is the input to the decision device, $Q\{.\}$, which, in turn, generates the estimated symbol of $s(n)$, $\hat{s}(n)$.

It has been demonstrated in [3], [5] that the ARMA channel in (1) can be decomposed into a real-valued attenuation factor (η), amplitude distortion ($C_{AD}(z)$) along with phase distortion ($C_{PD}(z)$). Consequently, (1) may be rewritten by

$$\begin{aligned} C(z) &= \underbrace{\eta}_{\text{attenuation}} \cdot \underbrace{\left[\frac{\prod_{j=1}^{M_{Z1}}(1-c_{Zi,j}\cdot z^{-1})\cdot\prod_{j=1}^{M_{Z2}}(c_{Zo,j}^{*}-z^{-1})}{\prod_{j=1}^{M_P}(1-c_{Pi,j}\cdot z^{-1})}\right]}_{\text{amplitude distortion}} \\ &\cdot \underbrace{\left[\frac{\prod_{j=1}^{M_{Z2}}(1-c_{Zo,j}\cdot z^{-1})}{\prod_{j=1}^{M_{Z2}}(c_{Zo,j}^{*}-z^{-1})}\right]}_{\text{phase distortion}} = \eta\cdot C_{AD}(z)\cdot C_{PD}(z) \end{aligned} \quad (3)$$

where $M_Z=M_{Z1}+M_{Z2}$; $c_{Zi,j}$ and $c_{Zo,j}$ denote the zeros inside and outside the unit circle, respectively, such that $\left|c_{Zi,j}\right|<1$ and $\left|c_{Zo,j}\right|>1$; c_{Pi} denotes the pole inside the unit circle such that $\left|c_{Pi,j}\right|<1$. Accordingly, the amplitude distortion ($C_{AD}(z)$) is a minimum-phase system, and the phase distortion ($C_{PD}(z)$) is an all-pass system such that $\left|C_{PD}(z)\right|=1$.

B. Equalizer $F(z)$

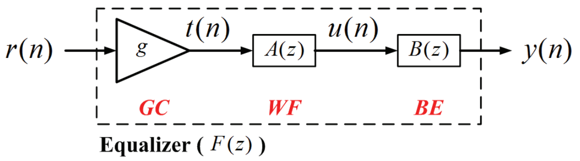

Fig. 2. Structure of the equalizer, $F(z)$, with WF.

The attenuation (η), amplitude distortion ($C_{AD}(z)$) and phase distortion ($C_{PD}(z)$) described in (3) can be compensated by a gain controller (GC), WF and BE, respectively. The equalizer ($F(z)$) in Fig. 1 is composed of GC, WF and BE as shown in Fig. 2. The GC is a filter with one single real-valued tap weight, and its output can be expressed as

$$t(n)=g\cdot r(n) \quad (4)$$

The general form of the WF is a ZPWF [5]

$$A(z)=\frac{1+A_Z(z)}{1+A_P(z)}=\frac{1+\sum_{i=1}^{N_{AZ}} a_{Z,i}\cdot z^{-i}}{1+\sum_{j=1}^{N_{AP}} a_{P,j}\cdot z^{-j}} \quad (5)$$

whose output can be expressed as

$$u(n)=t(n)+\sum_{i=1}^{N_{AZ}} a_{Z,i}\cdot t(n-i)-\sum_{j=1}^{N_{AP}} a_{P,j}\cdot u(n-j) \quad (6)$$

where $a_{Z,i}$, $i=1\sim N_{AZ}$, and $a_{P,j}$, $j=1\sim N_{AP}$, are complex-valued tap weights of the ZPWF. When $N_{AZ}>0$ and $N_{AP}=0$, (5) is reduced to the AZWF, $A(z)=1+A_Z(z)$, proposed in [4]. When $N_{AZ}=0$ and $N_{AP}>0$, (5) is reduced to the APWF, $A(z)=1/[1+A_P(z)]$, proposed in [3]. The BE is an FIR equalizer whose output can be expressed as

$$y(n)=\sum_{i=0}^{N_B} b_i\cdot u(n-i)=\mathbf{b}^T\mathbf{u}(n) \quad (7)$$

where $\mathbf{b}=\left[b_0 \ \cdots \ b_{N_B}\right]^T$ denotes the vector-valued tap weights of the BE, and $\mathbf{u}(n)=\left[u(n) \ \cdots \ u(n-N_B)\right]^T$.

III. Optimal WF Based On MOE Criterion

The MOE criterion is based on the second-order statistics such that the cost function can be expressed as

$$\begin{aligned} E\left\{|u(n)|^2\right\} &= \left[E\{|s(n)|^2\}\cdot\frac{1}{2\pi j}\oint_U |C(z)\cdot g\cdot A(z)|^2\frac{dz}{z}\right] \\ &+\left[E\{|w(n)|^2\}\cdot\frac{1}{2\pi j}\oint_U |g\cdot A(z)|^2\frac{dz}{z}\right] \\ &=\left[\sigma_s^2\cdot\frac{1}{2\pi j}\oint_U |g\cdot C(z)\cdot A(z)|^2\frac{dz}{z}\right] \\ &+\left[\sigma_w^2\cdot\frac{1}{2\pi j}\oint_U |g\cdot A(z)|^2\frac{dz}{z}\right] \end{aligned} \quad (8)$$

where the integration is taken in the counterclockwise direction along the contour U which denotes a closed contour within the region of convergence (ROC). For a system to be stable, the ROC of $E\{|u(n)|^2\}$ must satisfy $|z|<1$ implying that all the poles of $C(z)$ and $A(z)$ are inside the unit circle. Substituting (3) into (8) yields

$$\begin{aligned} E\left\{|u(n)|^2\right\} &= \left[\sigma_s^2\cdot\frac{1}{2\pi j}\oint_U |g\cdot\eta\cdot C_{AD}(z)\cdot C_{PD}(z)\cdot A(z)|^2\frac{dz}{z}\right] \\ &+\left[\sigma_w^2\cdot\frac{1}{2\pi j}\oint_U |g\cdot A(z)|^2\frac{dz}{z}\right] \\ &=\sigma_s^2 g^2\eta^2\cdot\left[\frac{1}{2\pi j}\oint_U |C_{AD}(z)A(z)|^2\frac{dz}{z}\right] \\ &+\sigma_w^2 g^2\cdot\left[\frac{1}{2\pi j}\oint_U |A(z)|^2\frac{dz}{z}\right] \end{aligned} \quad (9)$$

where $\left|C_{PD}(z)\right|=1$ has been used in deriving (9). Dividing σ_s^2 on both sides of (9) yields a source-power-normalized output energy (OE)

$$\begin{aligned} J_{OE} &= g^2\eta^2\cdot\left[\frac{1}{2\pi j}\oint_U |C_{AD}(z)A(z)|^2\frac{dz}{z}\right] \\ &+\lambda\cdot g^2\cdot\left[\frac{1}{2\pi j}\oint_U |A(z)|^2\frac{dz}{z}\right] \end{aligned} \quad (10)$$

where $\lambda=\sigma_w^2/\sigma_s^2$. Owing to the fact that the GC and the WF can be designed separately, and, furthermore, the WF is the main focus of this paper, the following assumption is made to simplify the derivations of J_{OE}.

Assumption (A): *The GC has been perfectly designed such that* $g = 1/\eta$, *which utterly compensates for the attenuation* (η) *from the channel in (3)*

Based on **Assumption (A)**, (10) can be rewritten as

$$J_{OE} = \left[\frac{1}{2\pi j}\oint_U |C_{AD}(z)A(z)|^2 \frac{dz}{z}\right] + \underbrace{\frac{\lambda}{\eta^2}\cdot\left[\frac{1}{2\pi j}\oint_U |A(z)|^2 \frac{dz}{z}\right]}_{\text{noise enhancement}} \quad (11)$$

where the second term on the right-hand side of (11) represents the noise enhanced by the GC and the WF. Applying the *Residual theorem* [13] to (11) yields

$$J_{OE} = \sum_{i=1}^{L_1} \text{Res}\left[|C_{AD}(z)A(z)|^2 \cdot z^{-1} \quad \text{at pole } p_i\right] + (\lambda/\eta^2)\cdot\sum_{j=1}^{L_2} \text{Res}\left[|A(z)|^2 \cdot z^{-1} \quad \text{at pole } q_j\right] \quad (12)$$

where $\text{Res}[X(z) \text{ at pole } x]$ denotes the residue at x ; p_i, $i = 1 \sim L_1$, and q_j, $j = 1 \sim L_2$, denote, respectively, the poles of $\left[|C_{AD}(z)A(z)|^2 z^{-1}\right]$ and the poles of $\left[|A(z)|^2 z^{-1}\right]$, all of which are inside the ROC.

In the noiseless case, since the WF is used to compensate for the amplitude distortion caused by the channel, it would not be difficult to show that the optimal WF will be exactly the inverse of the amplitude distortion caused by the channel. However, it would not be so obvious to obtain the optimal WF *in the presence of channel noise*. The following example demonstrates how the optimal two-tap ZPWF in the presence of channel noise may be obtained by using the Residual theorem in (12).

Example 1. The following non-minimum phase ARMA channels is adopted whose z-transform is

$$C(z) = \frac{1}{5}\cdot\frac{(1+5\cdot z^{-1})}{(1+0.6\cdot z^{-1})} = \frac{1}{\eta}\cdot\underbrace{\frac{(1+0.2\cdot z^{-1})}{(1+0.6\cdot z^{-1})}}_{C_{AD}(z)}\cdot\underbrace{\frac{(1+5\cdot z^{-1})}{(5+z^{-1})}}_{C_{PD}(z)} \quad (13)$$

which may be decomposed into η, $C_{AD}(z)$ and $C_{PD}(z)$ (see (3)). The two-tap ZPWF is used to compensate for the ARMA channel whose z-transform is $A(z) = (1 + a_{Z,1}\cdot z^{-1})/(1 + a_{P,1}\cdot z^{-1})$ where the tap weights, $a_{Z,1}$ and $a_{P,1}$, are both real-valued. The signal-to-noise ratio (SNR) is 20 dB such that $\lambda = 0.01$. The J_{OE} surface turns out to be an unimodal with only one global minimum identified by a "▲" at $(a_{Z,1}, a_{P,1}) = (0.6, 0.204)$ yielding $J_{OE} = 1.0117$, and its two-dimensional contour plot is depicted in Fig. 3(a). Thus, the z-transform of the optimal two-tap ZPWF in the presence of channel noise is

$$A^{\dagger}(z) = (1+0.6\cdot z^{-1})/(1+0.204\cdot z^{-1}) \quad (14)$$

This is in contrast to the noiseless case in which the optimal two-tap ZPWF is $A(z) = C_{AD}^{-1}(z) = (1+0.6\cdot z^{-1})/(1+0.2\cdot z^{-1})$, which is the zero-forcing version of the ZPWF.

Computer simulations were conducted using a binary phase-shift keying (BPSK) transmitted source to verify (14) and $J_{OE} = 1.0117$ by using the algorithm proposed in [5] to recursively update the tap weights of the ZPWF by

$$\begin{cases} a_{Z,1}(n+1) = a_{Z,1}(n) - \mu_A\cdot u(n)\cdot t(n-1) \\ a_{P,1}(n+1) = a_{P,1}(n) + \mu_A\cdot u(n)\cdot u(n-1) \end{cases} \quad (15)$$

where μ_A is a step-size parameter, and both $a_{Z,1}(n)$ and $a_{P,1}(n)$ were initialized by $[a_{Z,1}(0),\ a_{P,1}(0)] = [0,\ 0]$. As shown in Fig. 3(a), the ensemble-averaged trajectory of the adaptive tap weights $(a_{Z,1}, a_{P,1})$ of the ZPWF using $\mu_A = 0.0005$ (green line) over 100 independent runs asymptotically converges to the global minimum at $(a_{Z,1}, a_{P,1}) = (0.6, 0.204)$, and its corresponding ensemble-averaged OE asymptotically approaches $J_{OE} = 1.0117$ as shown in Fig. 3(b).

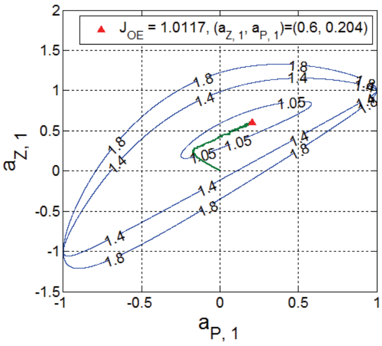

Fig. 3(a). The ensemble-averaged trajectory of adaptive tap weights $(a_{Z,1}, a_{P,1})$ over 100 independent runs on the two-dimensional contour plot of J_{OE} surface.

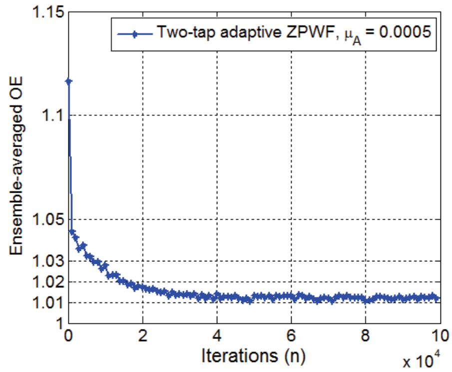

Fig. 3(b). The resulting ensemble-averaged OE over 100 independent runs using the adaptive ZPWF corresponding to Fig. 3(a).

IV. Preformance of Equalizers With or Without WF

The BE employs the CM criterion [1], [2] to minimize the cost function

$$J_{CM} = E\{[|y(n)|^2 - R]^2\},\ R \triangleq E\{|s(n)|^4\}/E\{|s(n)|^2\} \quad (16)$$

In order to express the equalizer output $y(n)$ in (16) in terms of $\{s(n)\}$, it is necessary to define the channel convolutional matrix and vector-valued impulse response of equalizer $F(z)$ first. If the WF has been used as the pre-processing filter of the BE as demonstrated in Fig. 2, which is referred to herein as the *combined WF-BE system*, whose z-transform may be approximated by using an N_D^{th}-order MA model

$$\tilde{D}(z) = \sum_{k=0}^{N_D} \tilde{d}_k \cdot z^{-k} \approx A(z)\cdot B(z) \quad (17)$$

whose vector-valued impulse response may be defined as

$$\tilde{\mathbf{d}} = \begin{bmatrix} \tilde{d}_0 & \cdots & \tilde{d}_{N_D} \end{bmatrix}^T \tag{18}$$

Similarly an arbitrary ARMA channel model in (1) may be approximated by an M_C^{th}-order MA model

$$\tilde{C}(z) = \sum_{k=0}^{M_C} \tilde{c}_k \cdot z^{-k} \approx C(z) \tag{19}$$

whose vector-valued impulse response may be defined as $\tilde{\mathbf{c}} = \begin{bmatrix} \tilde{c}_0 & \cdots & \tilde{c}_{M_C} \end{bmatrix}^T$. Then, a $(M_C + N_D + 1) \times (N_D + 1)$ convolutional matrix $\tilde{\mathbf{C}}$ of the linear system relating $s(n)$ to $l(n)$ in Fig. 1 can be defined by using (19) as follows

$$\tilde{\mathbf{C}} = \begin{bmatrix} c_0 & 0 & \cdots & 0 \\ \vdots & \ddots & \ddots & \vdots \\ c_{M_C} & \ddots & \ddots & 0 \\ 0 & \ddots & \ddots & c_0 \\ \vdots & \ddots & \ddots & \vdots \\ 0 & \cdots & 0 & c_{M_C} \end{bmatrix} \tag{20}$$

The equalizer output $y(n)$ can then be expressed by using (4), (18) and (20) as

$$y(n) = \mathbf{s}^T(n) \cdot \tilde{\mathbf{C}} \cdot (g \cdot \tilde{\mathbf{d}}) + \mathbf{w}^T(n) \cdot (g \cdot \tilde{\mathbf{d}}) \tag{21}$$

where $\mathbf{s}(n) = [s(n) \ \cdots \ s(n - M_C - N_D)]^T$ and $\mathbf{w}(n) = [w(n) \ \cdots \ w(n - M_C - N_D)]^T$. Assuming that the channels, sources, and noise are all real-valued, the expansion of J_{CM} can then be obtained by substituting (21) into (16)

$$\begin{aligned} J_{CM} &= \sigma_s^4(\kappa_s - 3)\left\|\tilde{\mathbf{C}} \cdot (g \cdot \tilde{\mathbf{d}})\right\|_4^4 + 3\sigma_s^4\left\|\tilde{\mathbf{C}} \cdot (g \cdot \tilde{\mathbf{d}})\right\|_2^4 \\ &+ \sigma_w^4(\kappa_w - 3)\left\|g \cdot \tilde{\mathbf{d}}\right\|_4^4 + 3\sigma_w^4\left\|g \cdot \tilde{\mathbf{d}}\right\|_2^4 \\ &+ 6\sigma_s^2\sigma_w^2\left\|\tilde{\mathbf{C}} \cdot (g \cdot \tilde{\mathbf{d}})\right\|_2^2\left\|g \cdot \tilde{\mathbf{d}}\right\|_2^2 \\ &- 2\sigma_s^2\kappa_s\left(\sigma_s^2\left\|\tilde{\mathbf{C}} \cdot (g \cdot \tilde{\mathbf{d}})\right\|_2^2 + \sigma_w^2\left\|g \cdot \tilde{\mathbf{d}}\right\|_2^2\right) + \sigma_s^4\kappa_s^2 \end{aligned} \tag{22}$$

where $\kappa_s = E\{|s(n)|^4\}/\sigma_s^4$ and $\kappa_w = E\{|w(n)|^4\}/\sigma_s^4$; the operations $\|\mathbf{x}\|_2$ and $\|\mathbf{x}\|_4$ denote the l_2-norm and l_4-norm of vector $\mathbf{x}$, respectively, such that $\|\mathbf{x}\|_2 = \sqrt{\sum_k |x_k|^2}$ and $\|\mathbf{x}\|_4 = \sqrt[4]{\sum_k |x_k|^4}$. Based on **Assumption (A)**, (22) can be rewritten as

$$\begin{aligned} J_{CM} &= \sigma_s^4(\kappa_s - 3) \cdot \frac{1}{\eta^4} \cdot \left\|\tilde{\mathbf{C}}\tilde{\mathbf{d}}\right\|_4^4 + 3\sigma_s^4 \cdot \frac{1}{\eta^4} \cdot \left\|\tilde{\mathbf{C}}\tilde{\mathbf{d}}\right\|_2^4 \\ &+ \sigma_w^4(\kappa_w - 3) \cdot \frac{1}{\eta^4} \cdot \left\|\tilde{\mathbf{d}}\right\|_4^4 + 3\sigma_w^4 \cdot \frac{1}{\eta^4} \cdot \left\|\tilde{\mathbf{d}}\right\|_2^4 \\ &+ 6\sigma_s^2\sigma_w^2 \cdot \frac{1}{\eta^4} \cdot \left\|\tilde{\mathbf{C}}\tilde{\mathbf{d}}\right\|_2^2\left\|\tilde{\mathbf{d}}\right\|_2^2 \\ &- 2\sigma_s^2\kappa_s \cdot \frac{1}{\eta^2} \cdot \left(\sigma_s^2\left\|\tilde{\mathbf{C}}\tilde{\mathbf{d}}\right\|_2^2 + \sigma_w^2\left\|\tilde{\mathbf{d}}\right\|_2^2\right) + \sigma_s^4\kappa_s^2 \end{aligned} \tag{23}$$

If the WF is not employed as the pre-processing filter as demonstrated in Fig. 4, the cost function J_{CM} in (23) can be written by replacing vector $\tilde{\mathbf{d}}$ directly with $\mathbf{b} = \begin{bmatrix} b_0 & \cdots & b_{N_B} \end{bmatrix}^T$ and by changing the dimension of the convolutional matrix $\tilde{\mathbf{C}}$ from $(M_C + N_D + 1) \times (N_D + 1)$ to $(M_C + N_B + 1) \times (N_B + 1)$.

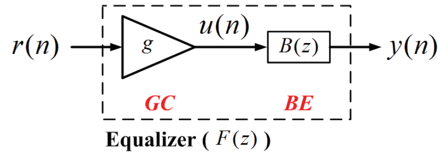

Fig. 4. Structure of the equalizer, $F(z)$, without WF (or BE alone).

Since the CM surface plotted by using (23), in general, is known to be multimodal, the adaptive BE may be temporarily attracted by a saddle point and/or may converge to an undesired local minimum if an inappropriate initialization is chosen. The following example analytically compares the combined WF-BE system with the BE alone system (without WF) in terms of the convergence speed and the steady-state performance.

Example 2. Consider the non-minimum phase ARMA channel adopted in Example 1 in which $\sigma_s^2 = 1$ and $\sigma_w^2 = 0.01$ such that $SNR = 20\ dB$. The z-transform of the two-tap BE may be expressed as

$$B(z) = b_0 + b_1 \cdot z^{-1} \tag{24}$$

whose vector-valued impulse response is $\mathbf{b} = [b_0, \ b_1]^T$. The two-tap ZPWF is used as the pre-processing filter of the BE, which is referred to herein as the "ZPWF-BE." Notably, the two-tap ZPWF is assumed to achieve the optimal solution in (14) when comparing the ZPWF-BE with the BE alone. Figures 5(a) and 6(a) show that the *CM surfaces* of using the BE alone and the ZPWF-BE, respectively, in terms of the tap weights of the BE, while Figs. 5(b) and 6(b) show their corresponding *contours*.

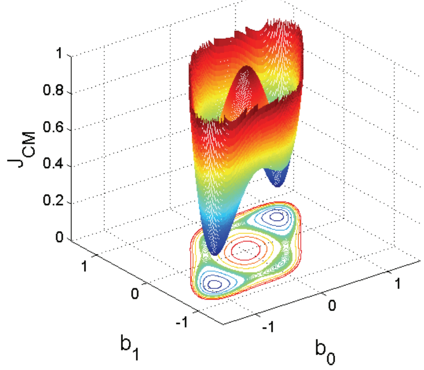

Fig. 5(a). J_{CM} surface plot for $C(z)$ with $SNR = 20\ dB$ using the BE alone.

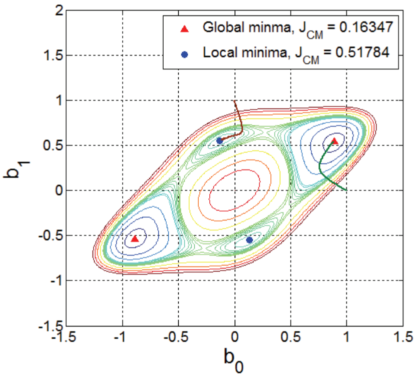

Fig. 5(b). J_{CM} contour plot for $C(z)$ with $SNR = 20\ dB$ using the BE alone along with the ensemble-averaged simulation trajectories.

As demonstrated in Figs. 5(b) and 6(b), the CM costs at the global minima (identified by "▲") and at the local minima (identified by "●") by using the ZPWF-BE are both smaller than those of their counterparts by using the BE alone since the ZPWF helps the BE compensate for the amplitude distortion of the channel. Consequently, the steady-state performance of the ZPWF-BE is expected to outperform that of using the BE alone. More importantly, the contours of the J_{CM} surfaces shown in Figs. 5(b) and 6(b) demonstrate that even if the BE in the ZPWF-BE converges to the local minima owing to a poor choice of initialization, the ZPWF-BE may still open eye simply because $J_{CM} = 0.16451$ at local minimum when using the ZPWF-BE is close to $J_{CM} = 0.16347$ at the global minimum when using the BE alone. However, it is very unlikely for the BE alone to open eye when it converges to a local minimum.

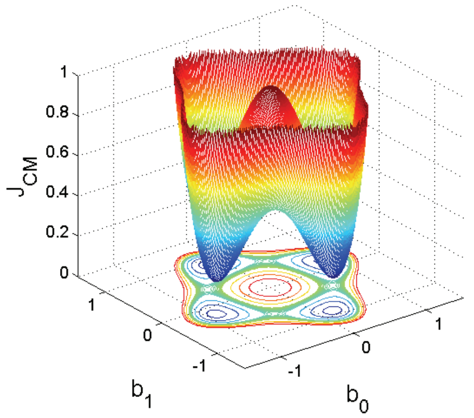

Fig. 6(a). J_{CM} surface plot for $C(z)$ with $SNR = 20\,dB$ using the ZPWF-BE.

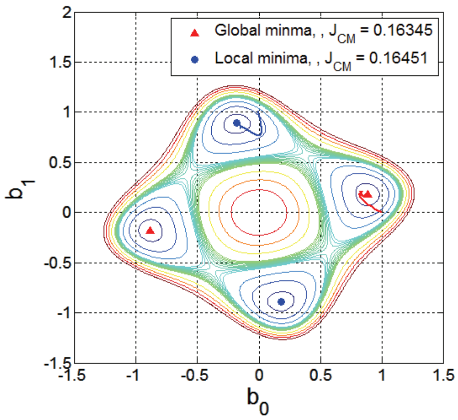

Fig. 6(b). J_{CM} contour plot for $C(z)$ with $SNR = 20\,dB$ using the ZPWF-BE along with the ensemble-averaged simulation trajectories.

Computer simulations have been conducted using a BPSK source to verify the abovementioned results in which the two-tap ZPWF is recursively updated by using the algorithm in (15), and the two-tap BE is recursively updated by using the CMA such that

$$\begin{bmatrix} b_0(n+1) \\ b_1(n+1) \end{bmatrix} = \begin{bmatrix} b_0(n) \\ b_1(n) \end{bmatrix} - \mu_B \cdot (|y(n)|^2 - R_2) \cdot y(n) \cdot \begin{bmatrix} u(n) \\ u(n-1) \end{bmatrix} \quad (25)$$

where μ_B is a step-size parameter. As depicted in Fig. 5(b), the trajectories of using the *BE alone* by initially setting $[b_0(0),\ b_1(0)] = [1,\ 0]$ (green line, $\mu_B = 0.01$) and $[b_0(0),\ b_1(0)] = [0,\ 1]$ (brown line, $\mu_B = 0.01$), respectively, asymptotically converge to the global minimum at $[b_0,\ b_1] = [0.89,\ 0.54]$ and the undesired local minimum at $[b_0,\ b_1] = [-0.13,\ 0.55]$, whereas the trajectories of using the *BE in the ZPWF-BE* by initially setting $[b_0(0),\ b_1(0)] = [1,\ 0]$ (red line, $\mu_A = 0.01$ and $\mu_B = 0.01$) and $[b_0(0),\ b_1(0)] = [0,\ 1]$ (blue line, $\mu_A = 0.01$ and $\mu_B = 0.01$), respectively, asymptotically converge to the global minimum at $[b_0,\ b_1] = [0.89,\ 0.19]$ and the undesired local minimum at $[b_0,\ b_1] = [-0.18,\ 0.89]$, as depicted in Fig. 6(b). Figure 7 demonstrates the ensemble-averaged symbol error rates (SERs) corresponding to the four ensemble-averaged trajectories over 1000 independent runs shown in Figs. 5(b) and 6(b). Figure 7 confirms that the ZPWF-BE was indeed capable of opening eye even when the BE converged to the undesired local minimum (see also Fig. 6(b)) owing to a poor choice of initialization, where eye opening herein is defined as $SER \le 1\%$ (or $\log_{10}(SER) \le -2$) [14], though it took a little longer to do so as compared with its convergence to the global minimum with an appropriate initialization. However, the use of the BE alone converging to the undesired local minimum was unable to open eye as depicted in Fig. 7. Furthermore, the rate of convergence of the BE in the ZPWF-BE converging to the global minimum is faster than that of using the BE alone as it converges to the global minimum.

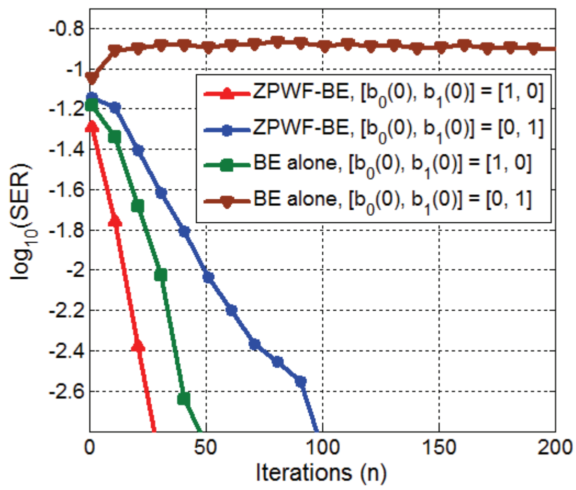

Fig. 7. Comparison between using the BE alone and using the ZPWF-BE in terms of ensemble-averaged SER with two different choices of initialization for the BE.

V. CONCLUSION

Since the CM surface is known to be multimodal, the adaptive BE with poor choices of initialization may converge to undesired local minima. Consequently, the use of the BE alone may result in poor transient and steady-state performances both of which may be ameliorated by pre-whitening the input signal by using a WF. This paper conducts analyses via CM surfaces to demonstrate that even if the BE based on the CM criterion converges to an undesired local minimum owing to a poor choice of initialization, the use of the BE along with a WF may still achieve eye opening, whereas the use of the BE alone may not. Computer simulations verify that the WF may not only help achieve eye opening but also increase the convergence speed of the BE.

REFERENCES

[1] D. N. Godard, "Self-recovering equalization and carrier tracking in two-dimensional data communication system," *IEEE Trans. Commun.*, vol. COM-28, pp. 1867-1875, Nov. 1980.

[2] C. R. Johnson, Jr., P. Schniter, T. J. Endres, J. D. Behm, D. R. Brown, and R. A. Casas, "Blind equalization using the constant Modulus criterion: A review," *Proc. IEEE*, vol. 86, no. 10, pp. 1927-1950, Oct. 1998.

[3] J. Labat, O. Macchi, and C. Laot, "Adaptive decision feedback equalization: Can you skip the training period?" *IEEE Trans. Commun.*, vol. 46, no. 7, pp. 921-930, Jul. 1998.

[4] W. G. Lim, R. A. Kennedy, and T. D. Abhayapala., "Adaptive decision feedback equalization under parallel adaptation," in *Proc. IEEE Int. Conf. Commun. (ICC 2008)*, pp. 495-499, May 2008.

[5] W.-C. Chang, S.-H. Chuang and J.-T. Yuan," A Zero-Pole Whitening Filter in Adaptive Blind Decision Feedback Equalizers," in *Proc. IEEE Int. Conf. Syst., Man, Cybern*, pp. 3259-3264, Oct. 2014.

[6] C. R. Johnson, Jr., and M. G. Larimore, "Comments on and additions to 'An adaptive recursive LMS filter,'" *Proc. IEEE*, vol. 65, no. 9, pp. 1399-1402, Sep. 1977.

[7] Stearns, S., "Error surfaces of recursive adaptive filters," *IEEE Trans. Acoust., Speech, Signal Process.*, vol. 29, No. 3, pp. 763-766, Jun 1981.

[8] J. J. Shynk, "Adaptive IIR filtering," *IEEE ASSP Mag.*, vol. 6, no. 2, pp. 4-21, Apr. 1989.

[9] J. P. LeBlanc and I. Fijalkow, "Blind adapted, pre-whitened constant modulus algorithm," in *Proc. Int. Conf. Commun. (ICC'01)*, vol. 8, pp. 2438-2442, Jun. 2001.

[10] R. López-Valcarce and F. Perez-Gonzalez, "Efficient Reinitialization of the Prewhitened Constant Modulus Algorithm," *IEEE Commun. Lett.*, vol. 5, no. 12, pp. 488-490, Dec. 2001.

[11] R. López-Valcarce, and F. Perez-Gonzalez, "Analysis of the prewhitened constant modulus cost function," in *Proc. IEEE Int. Conf. Acoust., Speech, and Signal Process. (IEEE ICASSP'02)*, vol. 3, pp. III-2641-III-2644, May 2002.

[12] M. Sternad and A. Ahlen, "The structure and design of realizable decision feedback equalizers for IIR channels with colored noise," *IEEE Trans. Inf. Theory*, vol. 36, no. 4, pp. 848-858, Jul. 1990.

[13] B. Widrow and S. D. Stearns, *Adaptive signal processing*, Englewood Cliffs, NJ: Prentice-Hall, 1985.

[14] J. Yang, J.-J. Werner, and G. A. Dumont, "The multimodulus blind equalization and its generalized algorithms," *IEEE J. Select. Areas Commun.*, vol. 20, no. 5, pp. 997-1015, June 2002.