Nonminimum-phase Channel Equalization using All-Pass CMA

Koen C.H. Blom*, Marco E.T. Gerards, André B.J. Kokkeler, and Gerard J.M. Smit
*Dep. of Electrical Engineering, Mathematics and Computer Science
University of Twente, P.O. Box 217, 7500 AE, Enschede, The Netherlands
Email: k.c.h.blom@utwente.nl

***Abstract*—A nonminimum-phase channel can always be decomposed into a minimum-phase part and an all-pass part. In our approach, called all-pass CMA, the dimensionality of the CMA algorithm has been reduced to improve blind equalization of a nonminimum-phase channel's all-pass part. The dimensionality reduction has been performed by parameterizing the CMA cost function in terms of the nonminimum-phase zero location of the all-pass part to be compensated. Currently, all-pass CMA can only compensate a single nonminimum-phase zero. However, compared to CMA, it typically provides a faster and more accurate compensation of this zero.**

I. Introduction

Equalization to deal with the effects of multipath propagation is a major challenge in the design of digital communication systems. Multipath components add up constructively and destructively producing frequency-selective channel distortion. In various channels, e.g., underwater acoustic communication, these multipath components are also changing over time resulting in a time-varying frequency selective channel. Various techniques exist for mitigation of frequency-selective channel distortion, e.g., Spread Spectrum (SS) modulation, (blind) adaptive equalization and multi-carrier modulation [1]. The main objective of this work is accelerating blind adaptive equalization for a subclass of frequency-selective channels, namely nonminimum-phase channels.

A physical channel response can be approximated by a finite length impulse response. Ideal equalization of Finite Impulse Response (FIR) channels requires filters of infinite length. Depending on the position of the channel zeros, finite length approximations may be used. Long FIR equalizers, having a large computational complexity, may be needed for channels with zeros close to the unit circle.

In general, adaptive equalizers compensate for time-varying channel distortion in the received signal by inserting a filter, that in combination with the channel response approximates the identity operation. Blind adaptive equalizers use statistical and geometrical properties of the received signal to estimate the response of the communication channel.

Realistic channel responses can be minimum-phase or nonminimum-phase. A channel is characterized as nonminimum-phase if some or all z-plane zeros lie outside the unit circle. For all channels with a similar magnitude response, the minimum-phase system has (most of) its energy concentrated near the start of the impulse response [2]. Given only the magnitude response of a nonminimum-phase channel, its phase response cannot uniquely be identified.

Since many blind equalization techniques cannot determine the correct phase response of nonmimimum-phase channels, equalization of these channels is more involved than equalization of minimum phase channels. This apparent 'phase blindness' is caused by the fact that a lot of equalizers are based on Second-Order Statistics (SOS) (e.g., autocorrelation) and reasoning in the autocorrelation domain suppresses phase information. SOS based equalization applied to a nonminimum phase channel produces an equalizer for the Spectrally Equivalent Minimum Phase (SEMP) version of the channel, but it leaves the phase response uncorrected. To equalize the phase response of nonminimum-phase channels, methods based on Higher-Order Statistics (HOS), as well as techniques exploiting cyclostationarity can be employed [3]. Our work incorporates the Constant Modulus (CM) blind equalization criterion. This criterion is nonlinear and therefore implicitly uses HOS [4].

Blind equalization of the magnitude and phase response can be performed separately because a nonminimum-phase equalizer can always be decomposed into a minimum-phase and an all-pass part. To reduce the computational complexity of a nonminimum-phase channel equalizer, Da Roche et al. [5] proposed an equalizer based on a cascade of an (less computationally complex) Infinite Impulse Response (IIR) filter and an all-pass FIR filter. Herein, the IIR filter is responsible for compensating the channel's magnitude response, whereas the all-pass filter equalizes the phase. The all-pass filter weights are estimated by minimizing a Direct Decision (DD) error criterion. The work of Lambotharan and Chambers [6] further reduces the net equalizer computational complexity by splitting up the all-pass filter in an IIR and a FIR filter. In their approach the FIR filter weights are estimated using a normalized version of the Constant Modulus Algorithm (CMA) algorithm.

Another important consideration for application of an adaptive equalizer is convergence speed. The convergence time should be small with respect to the coherence time of the channel. Our work focuses on accelerating and improving blind estimation of the all-pass filter weights required to mitigate the channel's phase response. We exploit prior knowledge of the mathematical structure of the possible equalizer weights

Supported through STW project: SeaSTAR (10552).

978-1-4577-1348-4/13/$31.00 ©2013 IEEE

to reduce the dimensionality of this blind estimation problem. Since our technique is based on CMA, but confines its solution space to merely all-pass filters, it is called all-pass CMA. To fully focus on the algorithm's convergence behaviour instead of (potential) instability, we only employ FIR filters.

An introduction to blind nonminimum-phase channel equalization using a cascade of a magnitude and a phase equalizer is given in Section II. Section III focuses on our proposed dimension reduction of conventional CMA to confine its solution space. The actual derivation of all-pass CMA, using Wirtinger calculus, is discussed in Section IV. A comparison of conventional CMA and all-pass CMA for equalization of the channel's phase response can be found in Section V. Conclusions and ideas for future work are discussed in Section VI.

II. Problem Definition

Consider the structure of the blind adaptive equalizer shown in Figure 1. The equalizer $W(z)$ mitigates time-varying frequency-selective distortion $H(z)$ in the received signal $x[n]$ based on structural and statistical properties of the non-Gaussian source signal $s[n]$. Additive White Gaussian Noise (AWGN) is represented by $v[n]$. In further analysis $v[n]$ will be ignored, because usually the Intersymbol Interference (ISI) caused by $H(z)$ dominates the degradation of the received signal [7]. Ideal equalization is attained if the aggregate

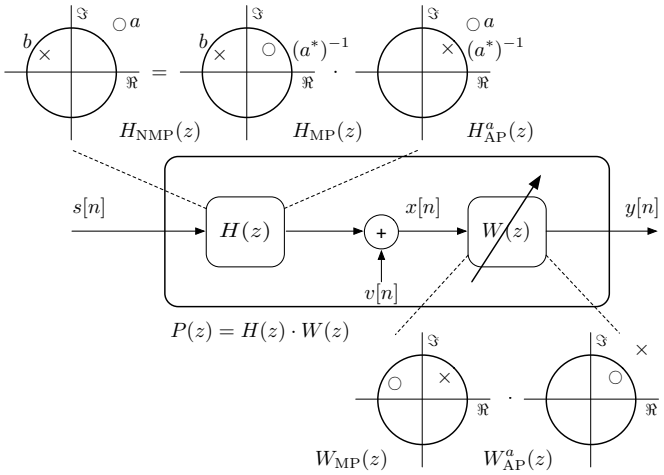

Fig. 1. Structure of a blind adaptive equalizer and the decomposition of a (single-pole single-zero) nonminimum phase channel.

channel response $P(z)$ approximates the identity operation. Nevertheless, the equalizer is allowed to impose a constant delay d, a gain c and a phase shift θ to the output signal $y[n]$ [3]:

$$Z^{-1}\{P(z)\} = p[n] = \delta[n-d] \cdot c \cdot e^{j\theta} \tag{1}$$

A nonminimum-phase channel has all or some of its zeros outside the unit circle and can always be decomposed into a minimum phase and an all-pass part [2]. For a single-pole single-zero nonminimum-phase channel, this decomposition is shown in Figure 1. Herein, $H_{\text{NMP}}(z)$ is factored into $H_{\text{MP}}(z)$ and $H^a_{\text{AP}}(z)$. $H^a_{\text{AP}}(z)$ is the all-pass channel with a zero at location a. Mathematically, the decomposition is given as:

$$\begin{aligned} H_{\text{NMP}}(z) &= n_0 \cdot \frac{(z-a)}{(z-b)} \\ &= n_0 \cdot \frac{(z-\frac{1}{a^*})}{(z-b)} \cdot \frac{(z-a)}{(z-\frac{1}{a^*})} = H_{\text{MP}}(z) \cdot H^a_{\text{AP}}(z). \end{aligned}$$

Herein, a and b are the zero and pole location of the nonminimum phase channel. Furthermore, n_0 is a gain factor used to normalize the magnitude response.

The minimum phase transfer $H_{\text{MP}}(z)$ has a stable and causal inverse $W_{\text{MP}}(z)$. Therefore, the time domain analog of $W_{\text{MP}}(z)$, the (causal) impulse response $w_{\text{MP}}[k]$, can be used to equalize the magnitude response of the channel. The inverse z-transform of $W^a_{\text{AP}}(z)$ is only stable if the Region of Convergence (ROC) is chosen to extend inward from the outermost pole, which results in the anti-causal time-domain characterization $w^a_{\text{AP}}[k]$. For phase equalization a block-based time reverser [5] [6] or an anti-causal FIR can be used.

Phase equalization always requires an all-pass magnitude response. Therefore, in our work the search space of CMA is reduced to (single-pole single-zero) all-pass filters through a dimensionality reduction of the parameters to be estimated.

III. Dimensionality Reduction of CMA

The time-domain expression of a single-pole single-zero all-pass equalizer is discussed in Section III-A. In Section III-B this expression is used to reduce the dimensionality of conventional CMA. Section III-C discusses the error-performance surface of all-pass CMA.

A. All-pass Channel and Equalizer Characterization

The unit magnitude (use $n_0 = 1/a^*$) and causal z-domain characterization of the all-pass part $H^a_{\text{AP}}(z)$ can be found as follows (ROC: $|z| > |1/a^*|$):

$$H^a_{\text{AP}}(z) = \frac{1}{a^*}\frac{(1-az^{-1})}{(1-\frac{1}{a^*}z^{-1})} = \frac{(1-az^{-1})}{(a^*-z^{-1})}$$

The Kth-order time domain approximation $H^a_{\text{AP}}(z)$ is the discrete-time impulse response $\boldsymbol{h}^a_{\text{AP}}$:

$$\boldsymbol{h}^a_{\text{AP}} = a \cdot \delta[k] + (\frac{1}{a^*} - a)(\frac{1}{a^*})^k u[k], \quad k = 0, \ldots, (K-1).$$

Herein, $\delta[k]$ is the Kronecker delta and $u[k]$ is the discrete-time unit step function.

The Kth-order anti-causal time domain approximation $\boldsymbol{w}^a_{\text{AP}}$ of the (single-pole single-zero) all-pass equalizer $W^a_{\text{AP}}(z)$ is given as (ROC: $|z| < |a|$):

$$\begin{aligned} \boldsymbol{w}^a_{\text{AP}} &= Z^{-1}\left\{\frac{(a^*-z^{-1})}{(1-az^{-1})}\right\} \\ &= \frac{\delta[k]}{a} - (a^* - \frac{1}{a})a^k u[-k-1], \quad k = (-K+1), \ldots, 0. \end{aligned} \tag{2}$$

Since we are only dealing with all-pass transfers in the rest of this work, for notational convenience and clarity, $\boldsymbol{h}^a_{\text{AP}}$ and $\boldsymbol{w}^a_{\text{AP}}$ will be written as $\boldsymbol{h}(a)$ and $\boldsymbol{w}(a)$.

B. Single- pole single-zero All-Pass CMA

The CM cost criterion $J_{\text{CMA}}(\boldsymbol{w})$ is a positive measure of the average deviation of the equalizer output $y = \boldsymbol{w}^T\boldsymbol{x}$ from a constant modulus [7]:

$$J_{\text{CMA}}(\boldsymbol{w}) = E\left\{\left[\left|\boldsymbol{w}^T\boldsymbol{x}\right|^2 - R_2\right]^2\right\}, \boldsymbol{x} = x[n,\ldots,n{+}N{-}1]^T. \tag{3}$$

R_2 is called the Godard parameter and is defined as:

$$R_2 = \frac{E[|s[n]|^4]}{E[|s[n]|^2]}.$$

In this work, we use Quadrature Phase-Shift Keying (QPSK) symbols as non-Gaussian source signal $s[n]$. Because of this QPSK source, we set R_2 to one such that the derivative of J_{CMA} w.r.t. $\boldsymbol{w}$ is zero when minimum costs are reached.

Conventional CMA tries to minimize deviation from a constant modulus by choosing an appropriate weight vector $\boldsymbol{w}$. In this work, the dimensionality of CMA is reduced by substituting $\boldsymbol{w}$ for $\boldsymbol{w}(\hat{a})$. Note that this subsitution is only valid if CMA is used for (single-pole single-zero) phase equalization. Effectively, J becomes a function of the zero position $\hat{a}$ compensated by the equalizer $\boldsymbol{w}(\hat{a})$:

$$J(\hat{a}) = E\left\{\left[\left|\boldsymbol{w}(\hat{a})^T x\right|^2 - R_2\right]^2\right\} \quad , \mathbf{x} = x[n\ldots n{+}N{-}1]^T. \tag{4}$$

In contrast to an N-tap conventional CMA equalizer the dimensionality of the optimization problem is now reduced from N complex weights to a single complex weight ($\mathbb{C}^N \rightarrow \mathbb{C}$).

C. Error-performance Surface

The dependence of the cost function J on the (equalizer) weight parameter $\hat{a}$ is called the error-performance surface. For all-pass CMA holds $J : \mathbb{C} \rightarrow \mathbb{R}$, therefore this dependence can be visualized in a three-dimensional surface plot. Mathematically, the error-performance surface can be given as follows:

$$J^a(\hat{a}) = E\left\{\left[|s[n] \otimes \boldsymbol{h}(a) \otimes \boldsymbol{w}(\hat{a})|^2 - R_2\right]^2\right\} \tag{5}$$

Herein, $\otimes$ is the convolution operator and $s[n]$ is a sequence of uniform independent and identically distributed (i.i.d.) zero-mean QPSK symbols. a is the zero position in the nonmimimum phase channel and $\hat{a}$ is the zero position compensated by the equalizer.

The error-performance surface for the channel $\boldsymbol{h}(1{+}1j)$ is shown in Figure 2. Both $\boldsymbol{h}(1{+}1j)$ and $\boldsymbol{w}(\hat{a})$ are implemented using a eight-order FIR filter. It can clearly be seen that the costs are minimized if the (ideal) equalizer $\boldsymbol{w}(1+1j)$ is used for compensating the channel's response. Note that, values of $|\hat{a}| < 1$ are not shown because these result in an unbounded equalizer impulse response. By choosing appropriate initial conditions for $\hat{a}$, e.g. a gradient descent optimization would never converge to such an unbounded response.

The error-performance surface of our CM based nonlinearity is multimodal. However, no other zero-memory nonlinear cost function is known, which would always result in global convergence [7]. The basin of attraction to a global minimum becomes larger if the zero of the non-minimum phase channel moves further away from the unit circle. The latter can be seen in Figure 3, wherein a is set to $2 + 2j$.

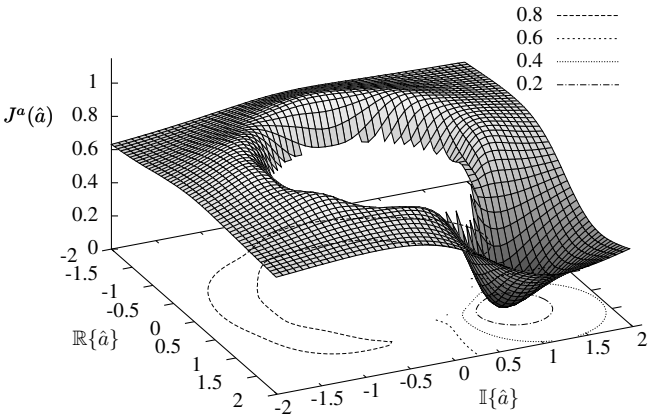

Fig. 2. Error-performance surface $J^{1+1j}(\hat{a})$ (K=8).

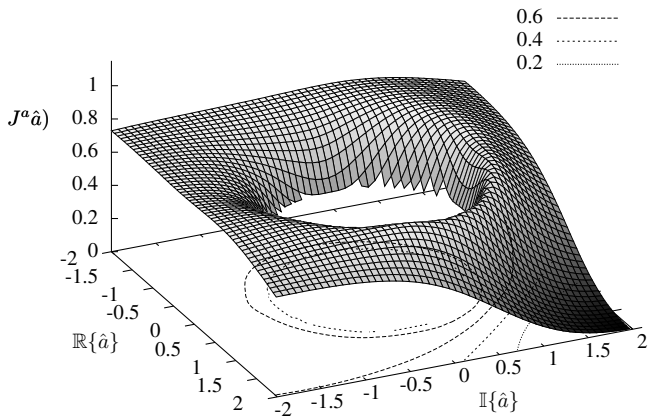

Fig. 3. Error-performance surface $J^{2+2j}(\hat{a})$ (K=8).

IV. Derivation of All-Pass CMA

To derive a gradient-descent minimizer for all-pass CMA, Wirtinger calculus is used. Wirtinger calculus is discussed in Section IV-A. The stochastic gradient-descent minimizer of all-pass CMA is given in Section IV-B.

A. Wirtinger Calculus

Our all-pass CMA cost function J is of type $\mathbb{C} \rightarrow \mathbb{R}$. For minimization of J, we use a stochastic gradient-descent method. Unfortunately, the function J is not complex differentiable, since it does not satisfy the Cauchy-Riemann equations. An elegant approach to relax complex differentiability is to work in the Wirtinger calculus. The latter provides a means to overcome the hassle of decomposing the complex cost function into real and imaginary parts and differentiating both parts separately. In the Wirtinger calculus the cost function is decomposed into two complex conjugate variables and these variables are treated as they are independent from each other. The work by Brandwood [8] states:

Theorem 1 (Wirtinger Calculus): Let $f : \mathbb{C}^N \rightarrow \mathbb{R}$ be a real-valued scalar function of a complex vector $\boldsymbol{z}$. Let $f(\boldsymbol{z}) = g(\boldsymbol{z}, \boldsymbol{z}^*)$, where $g : \mathbb{C}^N \times \mathbb{C}^N \rightarrow \mathbb{R}$ is a real-valued function of two complex variables and g is analytic with respect to each element in $\boldsymbol{z}$ and $\boldsymbol{z}^*$. Then either of the conditions $\nabla_{\boldsymbol{z}} g = 0$ or $\nabla_{\boldsymbol{z}^*} g = 0$ is necessary and sufficient to determine a stationary point of f. The gradient $\nabla_{\boldsymbol{z}^*} g$ defines the direction of the maximum rate of change of f with respect to $\boldsymbol{z}$.

B. Minimizer Derivation

Corollary 1 (All-pass CMA minimizer): An estimate for the channel's nonminimum phase zero $\hat{a}$ can be found using the following gradient-descent minimization:

$$\hat{a}[n+1] = \hat{a}[n] - \mu \cdot [\left|(\boldsymbol{w}(\hat{a})^T \boldsymbol{x}\right|^2 - 1] \cdot \left\{ \boldsymbol{w}_1(\hat{a}) \boldsymbol{x}^* \boldsymbol{x}^T \boldsymbol{w}(\hat{a}) + \boldsymbol{w}(\hat{a})^H \boldsymbol{x}^* \boldsymbol{x}^T \boldsymbol{w}_2(\hat{a}) \right\} \quad (6)$$

where,

$$\boldsymbol{w}_1(\hat{a}) = -\hat{a} k (\hat{a}^*)^{k-1} + (k-1)(\hat{a}^*)^{k-2}, \quad k = (-K+1), \ldots, 0$$
$$\boldsymbol{w}_2(\hat{a}) = -\hat{a}^k u[-k-1], \quad k = (-K+1), \ldots, 0.$$

Herein, μ is the convergence factor and K is the order of the blind equalizer.

Proof: First we rewrite $J(\hat{a})$ in terms of $\hat{a}$ and $\hat{a}^*$ by regarding $\boldsymbol{w}(\hat{a})$ as a function of $\hat{a}$ and $\hat{a}^*$:

$$J(\hat{a}) = G(\hat{a}, \hat{a}^*) = E\left\{ [\left|(\boldsymbol{w}(\hat{a}, \hat{a}^*)^T \boldsymbol{x}\right|^2 - 1]^2 \right\}.$$

Note that, function $G(\hat{a}, \hat{a}^*)$ is analytic with respect to $\hat{a}$ and $\hat{a}^*$. Therefore, according to Theorem 1, we may partially differentiate $G(\hat{a}, \hat{a}^*)$ with respect to $\hat{a}^*$ to find the direction of the maximum rate of change of J with respect to $\hat{a}$. Start partial differentiation by expanding $G(\hat{a}, \hat{a}^*)$ using the chain rule:

$$\frac{\delta G(\hat{a}, \hat{a}^*)}{\delta \hat{a}^*} = 2E\left\{ [\left|\boldsymbol{w}(\hat{a}, \hat{a}^*)^T \boldsymbol{x}\right|^2 - 1] \frac{\delta}{\delta \hat{a}^*} \left\{ (\boldsymbol{w}(\hat{a}, \hat{a}^*)^H \boldsymbol{x}^* \boldsymbol{x}^T \boldsymbol{w}(\hat{a}, \hat{a}^*) \right\} \right\}.$$

Thereafter, the partial derivative of the second part is found by application of the product rule:

$$\frac{\delta G(\hat{a}, \hat{a}^*)}{\delta \hat{a}^*} = 2E\left\{ [\left|\boldsymbol{w}(\hat{a}, \hat{a}^*)^T \boldsymbol{x}\right|^2 - 1] \left(\frac{\delta}{\delta \hat{a}^*} \boldsymbol{w}(\hat{a}, \hat{a}^*)^H\right) \boldsymbol{x}^* \boldsymbol{x}^T \boldsymbol{w}(\hat{a}, \hat{a}^*) + \boldsymbol{w}(\hat{a}, \hat{a}^*)^H \boldsymbol{x}^* \boldsymbol{x}^T \frac{\delta}{\delta \hat{a}^*} \boldsymbol{w}(\hat{a}, \hat{a}^*) \right\}$$

having,

$$\frac{\delta}{\delta \hat{a}^*} \boldsymbol{w}(\hat{a}, \hat{a}^*)^H = -a k (a^*)^{k-1} + (k-1)(a^*)^{k-2}, \quad k = (-K+1), \ldots, 0$$
$$\frac{\delta}{\delta \hat{a}^*} \boldsymbol{w}(\hat{a}, \hat{a}^*) = -a^k u[-k-1], \quad k = (-K+1), \ldots, 0.$$

The derivative $\frac{\delta}{\delta \hat{a}^*} G(\hat{a}, \hat{a}^*)$ is used in an instantaneous gradient-descent estimate to find the minimizer shown in Equation 6. ■

V. Simulations

To evaluate all-pass CMA, we illustrate its convergence behaviour in Section V-A. Thereafter, all-pass CMA is compared to conventional CMA in terms of ISI in Section V-B.

For simulations, the following real single-pole single-zero all-pass channel is used [9]:

$$H(z) = \frac{0.7 - z^{-1}}{1 - 0.7 z^{-1}}$$

In literature, channel $H(z)$ also includes a constant phase offset. This phase offset has been removed, to facilitate analysis, because CMA and all-pass CMA do not correct rotations of the symbol constellation. In the following simulations, both the all-pass channel transfer $H(z)$ and the equalizer are implemented using an eight-tap FIR filter.

The convergence factor of CMA is set to 5^{-3}. A factor in this order of magnitude typically results in decent equalization performance [10]. To initialize CMA, the fourth tap of the weight vector $\boldsymbol{w}$ is set to one and all others to zero. Such an initial weight vector is common practice [3].

The convergence factor of all-pass CMA is set to 1.5^{-1}. Choosing an initial estimate $\hat{a}[1]$ for the zero position is discussed in Section V-A.

A. Convergence Behaviour Analysis

The error-performance surface of all-pass CMA can be visualized in a contour plot. In such a plot, the gradient-descent paths of all-pass CMA can be included to study the algorithm's convergence behaviour. This relatively simple representation of the search space provides a means to make informed decisions on, e.g., initial conditions for the all-pass CMA algorithm. For example, the descent paths of our algorithm for the channel $H(z)$ and two initial values of the zero position ($\hat{a}[1] = 5 + 0j$ and $\hat{a}[1] = 2 + 2j$) are shown in Figure 4.

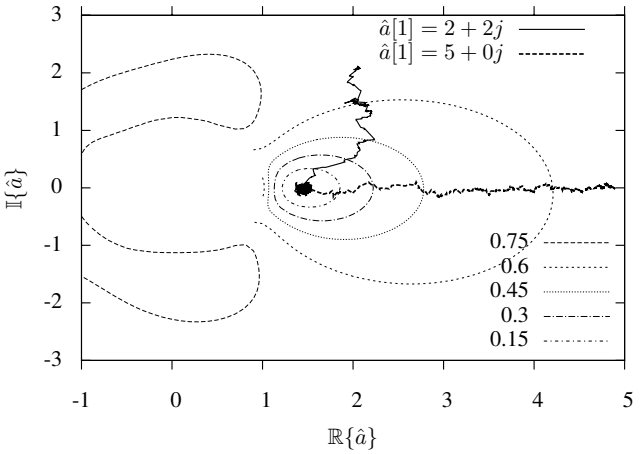

Fig. 4. Contour plot of $J^{\frac{1}{0.7}+0j}(\hat{a})$ annoted with the gradient-descent estimates $\hat{a}[n]$ of all-pass CMA for two different starting positions.

The values of $\hat{a}[1]$, being used in the simulations of Figure 4, have equal (all-pass CMA) costs. Clearly, both gradient-descent paths find the correct zero position ($\frac{1}{0.7}+0j$). However, in the contour plot we recognize a steeper gradient in the

imaginary direction than in the real direction. Therefore, one can conclude that, for channels having real zeros, our adaptive algorithm can possibly benefit from having a non-zero imaginary component in the initial estimate.

Although not shown in Figure 4, simulations with 16dB AWGN still result in correct convergence. However, the convergence factor μ should be chosen smaller in the case of large additive noise levels.

B. Equalization Performance Comparison

The equalization performance of all-pass CMA is studied in terms of the ISI level of the aggregate channel. For the aggregate channel response $\boldsymbol{p} = \boldsymbol{h} \otimes \boldsymbol{w}(\hat{a})$, a numerical measure of the ISI is found using [3]:

$$\text{ISI} = \frac{\sum |p[n]|^2 - |p[n_{\max}]|^2}{|p[n_{\max}]|^2}.$$

Herein, $n_{\max}$ is the index of the impulse response component having the largest magnitude.

The ISI levels of CMA and all-pass CMA, during adaptive equalization of channel $H(z)$, are shown in Figure 5. We clearly recognize a more accurate compensation of the nonminimum-phase zero by all-pass CMA compared to conventional CMA. For eight-tap all-pass CMA equalizers, the ISI levels after convergence are approximately 10dB smaller. Such an ISI reduction enables the use of higher-order modulation schemes and is advantageous for channels corrupted by high levels of AWGN.

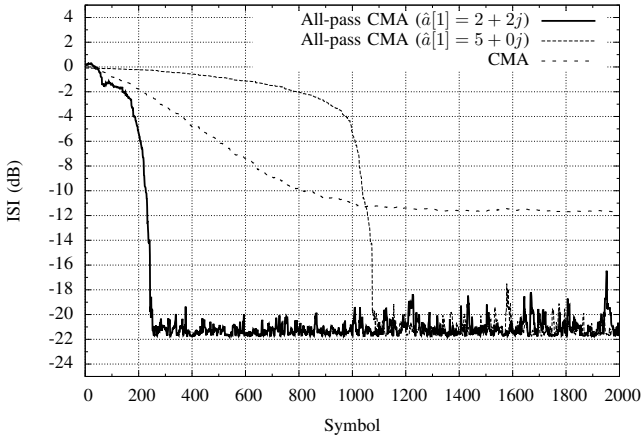

Fig. 5. Aggregate channel ISI levels for CMA and all-pass CMA.

The values of $\hat{a}[1]$ and the initial weight vector of CMA, being used in the simulations of Figure 5, result in equal levels of ISI for the initial aggregate channel response. All-pass CMA typically provides similar or faster convergence than conventional CMA. However, the convergence speed of both algorithms strongly depends on their initial weights and the convergence factor μ. As explained in Section V-A, being able to visualize to search space of all-pass CMA, can be valuable to find appropriate initial weights. A theoretical analysis of the convergence speed of all-pass CMA is out of the scope of this paper.

VI. Conclusions and Future Work

A nonminimum-phase channel can always be decomposed into a cascade of a minimum-phase response and an all-pass response. In general, blind equalization of this all-pass response is more involved than blind equalization of the minimum-phase response, since HOS are required to estimate the nonminimum-phase zero positions of the all-pass response. In our method, named all-pass CMA, the search space of a popular blind nonminimum-phase equalization technique (CMA) is reduced to merely all-pass responses to improve blind equalization of the all-pass part of a nonminimum-phase channel.

Our dimensionality reduction has been performed by parameterizing the CMA cost function in terms of the nonminimum-phase zero location of the all-pass response to be compensated. Equalization performance and convergence behaviour of all-pass CMA have been compared to CMA. All-pass CMA provides a far more accurate compensation of the all-pass response, while having a similar or improved convergence rate. E.g., the ISI level of an eight-tap all-pass CMA equalizer is 10 dB smaller than for conventional CMA.

Currently, all-pass CMA can only be used for compensation of a single-pole single-zero all-pass part of a nonminimum-phase channel. However, multiple all-pass CMA equalizers with different initial conditions can be cascaded to enable for compensation of multiple nonminimum-phase zeros. A similar derivation for higher-order all-pass channels also seems promising.

References

[1] B. Sklar, "Rayleigh fading channels in mobile digital communication systems part ii: Mitigation," *Communications Magazine, IEEE*, vol. 35, no. 9, pp. 148 –155, sept. 1997.

[2] J. G. Proakis and D. K. Manolakis, *Digital Signal Processing: Principles, Algorithms, and Applications*. Prentice Hall, 1996.

[3] A. K. Nandi, *Blind Estimation using Higher-Order Statistics*. Kluwer Academic Publishers, 2010.

[4] J.-T. Yuan and T.-C. Lin, "Equalization and carrier phase recovery of cma and mma in blind adaptive receivers," *Signal Processing, IEEE Transactions on*, vol. 58, no. 6, pp. 3206–3217, 2010.

[5] C. Da Rocha, O. Macchi, and J. Romano, "An adaptive nonlinear iir filter for self-learning equalization," in *Internat. Telecom. Conf. Rio de Janeiro, Brasil*, 1994, pp. 6–10.

[6] S. Lambotharan and J. A. Chambers, "A new blind equalization structure for deep-null communication channels," *Circuits and Systems II: Analog and Digital Signal Processing, IEEE Transactions on*, vol. 45, no. 1, pp. 108–114, 1998.

[7] S. Haykin, *Adaptive Filter Theory*, 4th ed. Prentice Hall, 2001.

[8] D. Brandwood, "A complex gradient operator and its application in adaptive array theory," in *IEE Proceedings H (Microwaves, Optics and Antennas)*, vol. 130, no. 1. IET, 1983, pp. 11–16.

[9] R. A. Kennedy and Z. Ding, "Blind adaptive equalizers for quadrature amplitude modulated communication systems based on convex cost functions," *Optical Engineering*, vol. 31, no. 6, pp. 1189–1199, 1992.

[10] K. Blom, M. van de Burgwal, K. Rovers, A. Kokkeler, and G. Smit, "DVB-S signal tracking techniques for mobile phased arrays," in *Vehicular Technology Conference Fall (VTC 2010-Fall), 2010 IEEE 72nd*, 2010, pp. 1–5.