ALPHABET POLYNOMIAL FITTING CRITERIA FOR BLIND EQUALIZATION

Ludwig ROTA, Pierre COMON

Laboratoire I3S, 2000 route des Lucioles, BP 121, 06903 Sophia-Antipolis Cedex, France.
Tel: +33-4 9294 2717; fax: +33-4 9294 2896,
e-mail: {rota,comon}@i3s.unice.fr
web: www.i3s.unice.fr/~{rota,comon}

ABSTRACT

This paper introduces a family of criteria dedicated to blind SISO equalizers. These criteria are based on *Alphabet Polynomial Fitting* (APF), and remind the well-known Constant Modulus Algorithm (CMA) criterion, and encompass the Constant Power Algorithm (CPA). Two APF-based algorithms have been implemented in block form for QPSK and QAM16 modulated inputs. Block implementations are indeed more efficient for short data records, and allow the direct computation of the optimal step size in a gradient descent, as shown in the paper. Finally, APF-QPSK and APF-QAM16 performances are compared with MMSE solutions for various data block lengths.

1. INTRODUCTION

One of the main advantages of blind techniques is that training sequences are not required. By deleting pilot sequences, one can thus increase the transmission rate. That's why Blind Equalization schemes have been the subject of intense interest since the work of Sato [9] and Godard [7].

Our paper is dedicated to Single-Input Single-Output (SISO) equalizers. This is not restrictive, *i.e.* the same criteria can be used with MIMO channels, since sources can be extracted one by one with a deflation approach [5][10]; this also allows to write a descent algorithm as a fixed point search [1].

This paper is organized as follows. In section 2, we introduce the SISO blind equalization problem; model and notations are also included. Then, in section 3, we describe the family of criteria based on *Alphabet Polynomial Fitting* (APF); assumptions and definition of *contrast* criteria are also given in this section. Practical algorithms, using an optimal step size gradient descent, are implemented in section 4. Finally, comparisons of performances of APF equalizers with Minimum Mean-Square Error linear equalizers (MMSE-LE) are presented in section 5.

2. MODEL AND NOTATIONS

Throughout the paper, ($^{\mathsf{T}}$) stands for transposition, ($^{\mathsf{H}}$) for conjugate transposition, (*) for complex conjugation, and $\jmath = \sqrt{-1}$. Vectors and matrices are denoted with bold lowercase and bold uppercase letters respectively, $\boldsymbol{I}$ stands for identity matrix. Moreover, let $\mathcal{H}$ be a set of filters, $\mathcal{S}$ the set of processes and $\mathcal{T}$ the subset of $\mathcal{H}$ of trivial filters [2].

In the field of digital communications, we consider baseband SISO observation model

$$y(n) = \sum_{k=1}^{K} c_k x(n-k+1) + \rho w(n) \tag{1}$$

where $x(n)$ denotes the useful unknown sequence, c_k the channel impulse response, $y(n)$ the received sequence, $w(n)$ the unit variance additive noise and ρ a parameter introduced in order to control the Signal to Noise Ratio. The blind equalization problem consists of finding a LTI filter, $\boldsymbol{f} = [f_1, \ldots, f_L]^{\mathsf{T}}$, in order to retrieve the input sequence solely from the observation of the output sequence of the unknown LTI channel $\boldsymbol{c} = [c_1, \ldots, c_K]^{\mathsf{T}}$. In other words, we search f_l, with $1 \leq l \leq L$, such that

$$z(n) = \sum_{\ell=1}^{L} f_l y(n-l+1) \tag{2}$$

yields a good estimate of the input sequence $x(n)$. The signal recovered can be delayed by a filter $\boldsymbol{\lambda}$, so that $\boldsymbol{c} \star \boldsymbol{f} = \boldsymbol{\lambda}$, where $\star$ is the convolution operator. When $\boldsymbol{\lambda}$ is of the form

$$\boldsymbol{\lambda} = [\underbrace{0, \ldots, 0}_{p-1}, \lambda, \underbrace{0, \ldots, 0}_{L+K-1-p}]^{\mathsf{T}} \tag{3}$$

then it belongs to the set of trivial filters [2], *i.e.* $\boldsymbol{\lambda} \in \mathcal{T}$.

3. POLYNOMIAL CRITERIA

The main assumption in blind equalization is the independence between successive symbols. Thus, we consider the following hypotheses:

- **H1**: Source $x(n)$ is a zero-mean random process, with unit variance.
- **H2**: Source $x(n)$ belongs to a known finite alphabet $\mathcal{A}$ characterized by the d distinct complex roots of a polynomial $Q(x) = 0$. For instance, a discrete PSK-q input is characterized by roots of $Q(x) = x^q - 1$. Table 1 gives polynomials $Q(x)$ for PSK-q and QAM16 modulations.
- **H3**: Source $x(n)$ is stationary up to order $r, r \geq q-1$: the order-r marginal cumulants,

$$\mathcal{C}_p^s(x(n)) = Cum\{\underbrace{x(n), \ldots, x(n)}_{p}, \underbrace{x^*(n), \ldots, x^*(n)}_{s=r-p}\} \tag{4}$$

do not depend on n.

Moreover, for PSK-q modulations, elements of the complex constellation satisfy $x^q = 1$. As a consequence, $E\{x^q\} = 1$ but $E\{x^m\} = 0, \forall m < q$. We shall say that x is circular up to order $q-1$, but non circular at order q.

Now, let us remind the definition of contrast criteria:

Definition 1: *An optimization criterion, $J(\boldsymbol{f}; z)$, is referred to as a contrast, defined on $\mathcal{H} \times \mathcal{H} \cdot \mathcal{S}$, if it enjoys the three properties below [2]:*

Modulation	$\mathcal{A}$	$Q(x)$
PSK-q	$\{e^{\jmath 2k\pi/q}\}\vert_{k\in 0,\dots,q-1}$	$x^q - 1$
QAM16	$[\{\pm 1, \pm 3\}, \{\pm\jmath, \pm 3\jmath\}]$	$\sum_{k=0}^{4} \alpha_k x^{4k}$

$\alpha_0 = 5625/256, \alpha_1 = 12529/16, \alpha_2 = -221/8,$
$\alpha_3 = 17, \alpha_4 = 1.$

Table 1: Polynomials characterizing PSK-q and QAM16.

***P1*. Invariance:** *The contrast should not change within the set of acceptable solutions, which means that* $\forall z \in \mathcal{H}\cdot\mathcal{S}$, $\forall \boldsymbol{f} \in \mathcal{T}$ *then* $J(\boldsymbol{f}; z) = J(\boldsymbol{I}; z)$.

***P2*. Domination:** *If sources are already equalized, any filter should decrease the contrast. In other words,* $\forall z \in \mathcal{S}$, $\forall \boldsymbol{f} \in \mathcal{H}$, *then* $J(\boldsymbol{f}; z) \leq J(\boldsymbol{I}; z)$.

***P3*. Discrimination:** *The maximum contrast should be reached only for filters linked to each other via trivial filters:* $\forall z \in \mathcal{S}, J(\boldsymbol{f}; z) = J(\boldsymbol{I}; z) \Rightarrow \boldsymbol{f} \in \mathcal{T}$.

Considering discrete inputs and SISO channel, one can blindly equalize it thanks to the polynomial criterion below:

Theorem 1: *The criterion*

$$J_{APF}(\boldsymbol{f}, z) = -\sum_n |Q(z(n))|^2 \quad (5)$$

*is a contrast under hypotheses **H2** and **H3**.*

The proof of the theorem needs the following lemma:

Lemma 2: *Let* $\mathcal{A} = \{x_n, 1 \leq n \leq N\}$ *be a given finite set of complex numbers not reduced to* $\{0\}$, *and* $\{c_k, 1 \leq k \leq K\}$ *non zero complex coefficients. Then, if* $\sum_{k=1}^{K} c_k x_{\sigma(k)} \in \mathcal{A}$, *for all mappings* σ, *not necessarily injective, from* $\{1, \dots, K\}$ *to* $\{1, \dots, N\}$, *only one component* c_k *is non zero.*

The proof of lemma 2 is rather long [4] and is not given due to lack of space. In a few words, $\boldsymbol{c}$ is shown to be trivial. The idea is to prove that a non trivial vector $\boldsymbol{c}$ generates symbols that may lie outside the convex hull of alphabet $\mathcal{A}$.

Now, let us prove that J_{APF} enjoys the three properties of a contrast:

Proof.

- *Property **P1***: for any trivial filter $\boldsymbol{\lambda} \in \mathcal{T}$, we have $-J_{APF}(\boldsymbol{\lambda}; z) = \sum_n |Q(\lambda z(n+\tau))|^2$, with $\tau \in \mathbb{Z}$, $\lambda \in \mathbb{C}$. Because of the sums, this can also be simply written as $-J_{APF}(\boldsymbol{\lambda}; z) = \sum_m |Q(\lambda z(m))|^2$. If z is in $\mathcal{S}$, then $z(m)$ belongs to $\mathcal{A}$, and so is $\lambda z(m)$. Thus $Q(\lambda z(m)) = 0$.
- *Property **P2***: since $\sum_n |Q(y(n))|^2 \geq 0$, J_{APF} is larger than or equal to $\sum_n |Q(x(n))|^2$, because the latter is null when $x(n) \in \mathcal{S}$. We have indeed $-J_{APF}(\boldsymbol{f}; x) \geq -J_{APF}(\boldsymbol{I}, x)$.
- *Property **P3***: we must show that if we have the equality $\sum_n |Q(y(n))|^2 = 0$, then $\boldsymbol{\lambda}$ is trivial. Denote $y(n) = \sum_k c_k x(n-k)$, with $x(n) \in \mathcal{A}$, and where c_k define the kth component of filter $\boldsymbol{c}$. Then we have $\forall n, Q(y(n)) = 0$. We thus have that $Q(\sum_k c_k z(n-k)) = 0$. We are under the conditions of lemma 2, and we may conclude that a single c_k is non zero. In addition, this c_k is necessarily in $\mathbb{C}$ since $c_k z$ must be in $\mathcal{A}$ for any $z \in \mathcal{A}$. By proceeding in the same way for every $y(n)$, we end up with an impulse response $\boldsymbol{c}$ having only one non zero entry.

◇

Criterion (5), also named *Alphabet Polynomial Fitting*, is based only on the modulation used for the transmission of the input sequence. Hence, we obtain a set of polynomial criteria dedicated to each modulation, in the presence of a perfect synchronization.

As mentioned in section 1, it is possible to use a deflation approach for equalizing mixtures from outputs of a MIMO channel. If all signals transmitted use different modulations, then it could be interesting to extract only one signal of the mixture thanks to the knowledge of its alphabet. For this, one can apply an APF criterion on the observations in order to extract the suitable signal.

If PSK modulations are used in the transmission scheme, then criteria J_{APF} are similar to the Constant Power Algorithm (CPA) described in [3] since they are reduced to the form $J(\boldsymbol{f}) = \|z(n)^q - d(n)\|^2$. In fact, all PSK-$q$ modulations can be characterized with $d(n)$ and q as mentioned in [3]. Nevertheless, contrary to APF algorithms, CPA is not able to equalize signals with amplitude modulations like QAM16. Moreover, one can combine criteria thanks to a simple theorem:

Theorem 3: *If* $J_k(z)$ *are contrasts defined on* $\mathcal{H}\cdot\mathcal{S}_k$, *and* $\{a_k\}$ *are strictly positive numbers, then* $J(z) = \sum_k a_k J_k(z)$ *is a contrast on* $\mathcal{H}\cdot\bigcup_k \mathcal{S}_k$.

Proof. Property **P2** is obtained immediately, because all terms are positive: $J(z) = \sum_k a_k J_k(z) \leq \sum_k a_k J_k(x) = J(x)$. If equality holds, then $\sum_k a_k[J_k(x) - J_k(z)] = 0$, which is possible only if every term vanishes because they are all positive. Thus $J_k(z) = J_k(x), \forall k$. But $x \in \mathcal{S}_k$ for some k, by hypothesis. And since J_k is a contrast, one can conclude that $z = \lambda \star x$, for some trivial filter λ of $\mathcal{H}$. This proves the theorem. ◇

Thus, by combining J_{APF} and J_{CM}, one obtain new contrast criteria.

4. OPTIMAL STEP SIZE DESCENT

The usual practice in SISO and deflation cases, is to run a gradient descent:

$$\boldsymbol{v} = \boldsymbol{f}(k) + \mu \boldsymbol{g}(k); \boldsymbol{f}(k+1) = \boldsymbol{v}/\|\boldsymbol{v}\| \quad (6)$$

where $\boldsymbol{g}(k)$ denotes the equalizer tap vector at iteration k, $\boldsymbol{g}(k)$ the gradient of J_{APF} calculated at $\boldsymbol{f}(k)$, and μ the step size. Most iterative algorithms run with a fixed step, which performs poorly when the criterion contains many saddle points. Even if the step size is adjusted like in quasi-Newton algorithm, it does not improve anything since the iterations can stay a long time in the neighborhood of a saddle point and then suddenly burst out far away from the attraction basin. One can improve significantly the convergence time with an optimal step size calculation. In fact, criterion J_{APF} is a rational function in the f_l's. It is also a rational function in variable μ since $J_{APF}(\boldsymbol{f}(k) + \mu \boldsymbol{g}(k))$ describes the same criterion. As a consequence, all its stationary points can be explicitly computed as roots of a polynomial in a single variable.

Now, we can rewrite (2) in a compact form

$$z(n) = \boldsymbol{f}^{\mathsf{T}} \boldsymbol{y}_n \quad (7)$$

where $\boldsymbol{y}_n = [y(n), y(n-1), \ldots, y(n-L+1)]^{\mathsf{T}}$ denotes the observation vector. Hence, we obtain the criterion

$$J_{APF}(\boldsymbol{f}) = -\sum_n Q(\boldsymbol{f}^{\mathsf{T}}\boldsymbol{y}_n)Q({\boldsymbol{y}_n}^{\mathsf{H}}\boldsymbol{f}^*). \tag{8}$$

Then, the gradient vector $\boldsymbol{g}$ is

$$\boldsymbol{g} = -\sum_n \boldsymbol{y}_n Q'(\boldsymbol{f}^{\mathsf{T}}\boldsymbol{y}_n)Q({\boldsymbol{y}_n}^{\mathsf{H}}\boldsymbol{f}^*) \tag{9}$$

where the function $Q'(z)$ denotes the derivative of the polynomial function $Q(z)$.

Now, we consider J_{APF} as a rational function of μ by substituting $z(n) = (\boldsymbol{f}+\mu\boldsymbol{g})^{\mathsf{T}}\boldsymbol{y}_n$ in (8):

$$J_{APF}(\mu) = -\sum_n Q((\boldsymbol{f}+\mu\boldsymbol{g})^{\mathsf{T}}\boldsymbol{y}_n)Q({\boldsymbol{y}_n}^{\mathsf{H}}(\boldsymbol{f}^*+\mu\boldsymbol{g}^*)). \tag{10}$$

Then, take its derivative with respect to variable μ:

$$\frac{\partial J_{APF}(\mu)}{\partial \mu} = -\sum_n \frac{\partial Q((\boldsymbol{f}+\mu\boldsymbol{g})^{\mathsf{T}}\boldsymbol{y}_n)}{\partial \mu} Q({\boldsymbol{y}_n}^{\mathsf{H}}(\boldsymbol{f}^*+\mu\boldsymbol{g}^*)) - \sum_n \frac{\partial Q({\boldsymbol{y}_n}^{\mathsf{H}}(\boldsymbol{f}^*+\mu\boldsymbol{g}^*))}{\partial \mu} Q((\boldsymbol{f}+\mu\boldsymbol{g})^{\mathsf{T}}\boldsymbol{y}_n).$$

It suffices to eventually plug back the roots of this derivative into criterion $J_{APF}(\mu)$, and to pick up the optimal step size $\mu(k)$ for the gradient descent, *i.e.* the root that maximizes criterion $J_{APF}(\mu)$. Of course, all this also applies to J_{KMA} ans J_{CMA}.

5. NUMERICAL ALGORITHMS

Two algorithms have been implemented in block form for QPSK and QAM16 modulated signals. Their performances are then compared with Minimum Mean-Square Error linear equalizers' (MMSE-LE). The mean-squared error (MSE) for such equalizers is given by

$$\sigma^2_{MMSE-LE} = \min_{\boldsymbol{f}} \mathrm{E}\{|x(n) - z(n)|^2\} \tag{11}$$

where $\mathrm{E}\{.\}$ denotes the expectation. We use the MSE criterion because it does not ignore noise enhancement. Indeed, the optimization of MMSE-LE filters compromises between eliminating intersymbol interference (ISI) and increasing noise power. From criterion (11), we obtain the finite length impulse response (FIR) of the MMSE linear equalizer [8]

$$\boldsymbol{f}_{MMSE} = \arg\min_{\boldsymbol{f}} \|\boldsymbol{f}^{\mathsf{T}}\boldsymbol{y}_n - x(n)\|^2 \tag{12}$$

with closed-form solution $\boldsymbol{f}_{MMSE} = \boldsymbol{R}^*_{xy}\boldsymbol{R}^{-1}_{yy}$, where the output covariance matrix is defined as $\boldsymbol{R}_{yy} = \mathrm{E}\{\boldsymbol{y}_n{\boldsymbol{y}_n}^{\mathsf{H}}\}$, and the cross-correlation matrix as $\boldsymbol{R}_{xy} = \mathrm{E}\{x_{n-\Delta}{\boldsymbol{y}_n}^{\mathsf{H}}\}$. Since Δ has little consequence when the equalizer length L is long, we can fixed it to $\Delta = \left\lceil \frac{L+K-1}{2} \right\rceil$, *i.e.* center of the impulse response of the global system (channel and equalizer). Correlation matrices have more specific expressions

$$\boldsymbol{R}_{xy} = \boldsymbol{C}[\underbrace{0,\ldots,0}_{\Delta-1}, \sigma^2_x, \underbrace{0,\ldots,0}_{L+K-1-\Delta}] \tag{13}$$

and

$$\boldsymbol{R}_{yy} = \sigma^2_x \boldsymbol{C}\boldsymbol{C}^{\mathsf{H}} + \rho^2 \boldsymbol{I} \tag{14}$$

where σ^2_x and ρ^2 are variances of input sequence and gaussian noise respectively, $\boldsymbol{I}$ denotes the $L \times L$ identity matrix and $\boldsymbol{C}$ the $L \times (L+K-1)$ block Toeplitz matrix

$$\boldsymbol{C} = \begin{bmatrix} c_1 & c_2 & \cdots & c_K & \cdots & 0 \\ 0 & c_1 & \ddots & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & \ddots & \vdots \\ 0 & 0 & \cdots & \cdots & c_{K-1} & c_K \end{bmatrix}. \tag{15}$$

In the following, we compare performances of APF-QPSK and APF-QAM16 algorithms for different Signal to Noise Ratio (SNR) and with various data block lengths.

5.1 APF-QPSK algorithm

The APF-QPSK algorithm has been tested on complex channels of length $K = 5$, with unit variance QPSK white processes. For each randomly generated channel, blocks of noisy observations are filtered according to in (1). We have tested APF-QPSK on random channels with data block lengths of 400, 800, 1200 and 1600 symbols. Since channels' coefficients are Gaussian distributed, we compute the MSE obtained with MMSE-LE in order to choose candidate channels for testing APF algorithms, *i.e.* invertible channels. The length-20 equalizers returned by the algorithm are then tested with a sequence of 1600 symbols in order to compute the Symbol Error Rate (SER). With 75 random channels, the minimal resolution is $(1600 * 75)^{-1} = 8,3.10^{-6}$. Figure 1

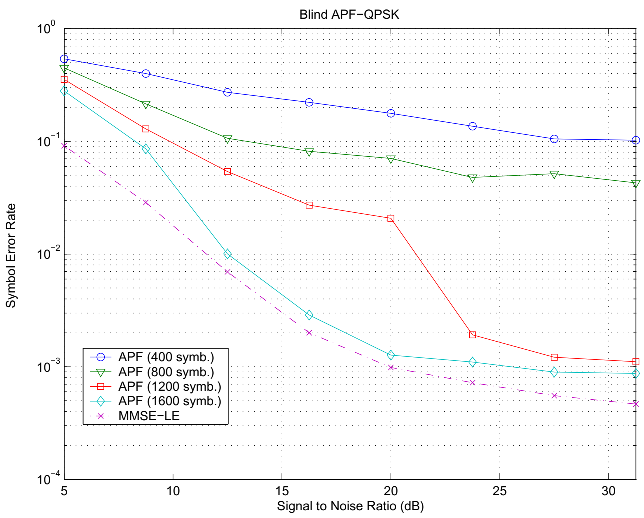

Figure 1: Performances of blind APF-QPSK equalizers with 400, 800, 1200, and 1600 symbols with K=5 and L=20.

shows the average of the 75 trials. This figure shows that above 23dB, APF-QPSK algorithm works well with 1200 symbols, whereas the SER obtained with 1600 symbols is close to the SER of MMSE-LE over the whole SNR range.

5.2 APF-QAM16 algorithm

We have tested the APF-QAM16 algorithm like in the previous section but with some modifications: we generate length-4 complex channels ($K = 4$) and the algorithm search for a length-16 equalizer ($L = 16$). Random channels have been tested with data block lengths of 1600, 2400 and 3200 symbols. The equalizers returned by APF-QAM16 are then tested with a sequence of 3200 symbols. The number of trials changed to 45 and the minimal resolution is now $(3200 * 45)^{-1} \approx 7.10^{-6}$. Figure 2 shows the average SER of the 45 trials. We can note that APF-QAM16 algorithm

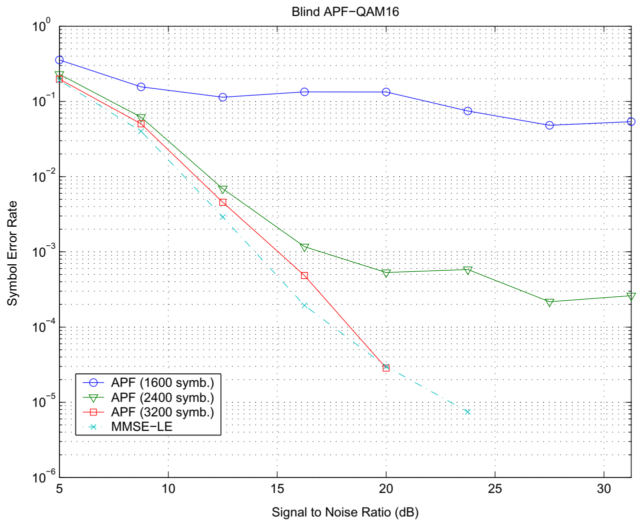

Figure 2: Performances of blind APF-QAM16 equalizers with 1600, 2400, and 3200 symbols with K=4 and L=16.

needs more symbols for equalizing the system. Indeed, we obtain good SER's for data block lengths greater than 2400 symbols. Moreover, with 3200 symbols, APF-QAM16 has approximatively the same behavior than the MMSE-LE. Indeed, the SER obtained with this size and for 20dB of SNR is below $0,03\%$.

5.3 Comparison with CMA

Finally, we compare performances of the CMA with APF algorithms. Let us remind the Constant Modulus criterion [10][6] $J_{CM}(z) = \mathrm{E}\{(1 - z^2)^2\}$. We have tested CMA and APF algorithms on length-3 channels with data block length of 1200 symbols. The SER returned by the length-10 equalizers are computed from another sequence of 2000 symbols. Figure 3 shows that APF-QPSK works better than CMA when noise is greater than 10dB. Moreover, APF-QAM16 is always below the SER obtained with CMA. This shows the good behavior of blind APF algorithms.

6. CONCLUDING REMARKS

A set of criteria based on Alphabet Polynomial Fitting has been introduced for blind SISO equalizers. Numerical algorithms based on two polynomial criteria have been implemented in block form for QPSK and QAM16 modulated inputs. These algorithms have been tested with different data block lengths and for various SNR's. Simulations show that the improvement is relative to the modulation of the signal and the data block length used. Open issues currently being addressed include the robustness of APF algorithms in the presence of carrier residual (*e.g.* via joint estimation of carrier offset) and the extraction of known alphabet signals from outputs of MIMO channels.

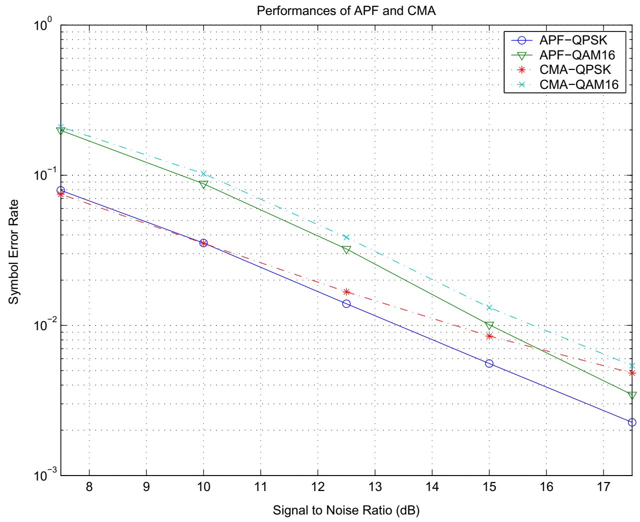

Figure 3: Performances of APF-QPSK, APF-QAM16, and CMA with K=3 and L=10.

REFERENCES

[1] A. HYVÄRINEN and J. KARHUNEN and E. OJA. *Independent Component Analysis*. Willey, 2001.

[2] P. COMON. Contrasts for Multichannel Blind Deconvolution. *IEEE Signal Processing Letters*, 3(7):209–211, July 1996.

[3] P. COMON. Blind Equalization with Discrete Inputs in the Presence of Carrier Residual. In *In Second IEEE Int. Symp. Sig. Proc. Inf. Theory*, Marrakech, Morocco, December 2002.

[4] P. COMON. Contrasts for Independent Component Analysis and Blind Deconvolution. I3s-cnrs research report 2003-06-fr, www.i3s.unice.fr, March 2003.

[5] N. DELFOSSE and P. LOUBATON. Adaptive blind separation of independent Sources: a deflation approach. *Signal Processing*, 45:59–83, 1995.

[6] Z. DING and Y. LI. *Blind Equalization and Identification*. Dekker, New York, 2001.

[7] D. GODARD. Self recovering equalization and carrier tracking in two dimensional data communication systems. *IEEE Trans. on Signal Processing*, 28(11):1867–1875, Nov. 1980.

[8] J. G. PROAKIS. *Digital Communications*. Series in Electrical Engineering and Computer Science. McGraw-Hill. 4th Edition.

[9] Y. SATO. A method of self recovering equalization for multilevel amplitude-modulation systems. *IEEE Trans. on Com.*, 23:679–682, June 1975.

[10] J.R. TREICHLER and M.G. LARIMORE. New processing techniques based on the constant modulus algorithm. *IEEE Trans. on Acoust. Speech Sig. Proc.*, 33(2):420–431, April 1985.