in *IEEE Int. Conf. Acoust. Speech Sig. Proc. (ICASSP'04)*, volume IV, pp. 441–444, May 17-21, 2004, Montreal.

# BLIND EQUALIZERS BASED ON POLYNOMIAL CRITERIA

*Ludwig ROTA, Pierre COMON*

I3S, Les Algorithmes/Euclide B - 2000 route des Lucioles,
BP 121, 06903 Sophia-Antipolis Cedex, France.

**ABSTRACT**

We describe 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) criterion. Algorithms based on several polynomial criteria have been implemented in block form (including CPA and APF), as well as the CMA and the Kurtosis Maximization (KMA). 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. Computational complexities of APF, KMA and CMA are eventually compared, as well as their performances for various digitally modulated inputs.

## 1. INTRODUCTION

Blind equalization schemes have been the subject of intense interest since the work of Sato [9] and Godard [7]. 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.

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][11]; 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 computational complexities and performances of APF, *Kurtosis Maximization* (KMA) [10] and *Constant Modulus Algorithm* (CMA) [11] 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 the following 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 $\rho$ 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:

**Hypothesis H1**: *Source $x(n)$ is a zero-mean random process, with unit variance.*

**Hypothesis 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.

**Hypothesis 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 $n$.*

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

$a_0 = 5625/256, a_1 = 12529/16, a_2 = -221/8,$
$a_3 = 17, a_4 = 1.$

**Table 1**. Polynomials characterizing PSK-q and QAM16.

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]:*

- ***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 separated, 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*

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

*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 $\sigma$, not necessarily injective, from $\{1, \dots, K\}$ to $\{1, \dots, N\}$, only one component $c_k$ is non zero.*

The proof of lemma 1 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 $k$th 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 evry 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 $\lambda$ 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:

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

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 $\mu$ 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 $\mu$ since $J_{APF}(\boldsymbol{f}(k) + \mu \boldsymbol{g}(k))$ describe 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

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

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 $\mu$ 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 $\mu$:

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

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

Algorithms have been implemented in block form for QPSK, PSK-8 and QAM16 modulations. In this section, we compare the computational complexities of the CMA and KMA with APF algorithms. Then we compare performances of theses algorithms with different modulations and for values of Signal to Noise Ratio (SNR).

### 5.1. Computational Complexities

Let us remind the Constant Modulus criterion [11]

$$J_{CM}(z) = \mathrm{E}\{(1 - z^2)^2\} \tag{11}$$

and the Kurtosis Maximization criterion [10]

$$J_{KM} = \frac{\mathrm{E}\{|z|^4\} - \mathrm{E}\{z^2\}^2}{\mathrm{E}\{|z|^2\}} - 2. \tag{12}$$

We note that they are almost equivalent, because related by a monotonous decreasing function as mentioned in [6].

We evaluate the computational complexity of an algorithm by counting the number of floating-point operations (flops) as described in [8]. Hence, each multiplication, division, addition, and subtractions are counted as one flop. We have two different stages in theses algorithms. The first, which is not present in our implementation of APF algorithms, is the *initialization* stage. The second corresponds to the *main loop* of algorithms, *i.e.* the iterative descent.

- **Initialization**: CMA and KMA are based on cumulants values. Theses algorithms have to compute tensors of cumulants before running the main loop. Table 2 summaries the number of flops for both algorithms. As mentioned above, our APF algorithms do not need any initialization calculus.

| **Algorithm** | **Complexity** (Flops) |
|---|---|
| **CMA** | $L^2[14(N-L+2)-3$ $+L^2(26(N-L)+45)]$ |
| **KMA** | $L^2[28(N-L+2)-6$ $+L^2(26(N-L)+45)]$ |

$L$ : length of the equalizer
$N$ : length of the block of observations

**Table 2**. Computational complexities of *Initialization* stage.

- **Main loop**: after initialization, the algorithms loop on a certain number of sweeps. Table 3 shows the theoretical number of flops for one loop of each algorithm. From this table, APF algorithms seem to be more complex than CMA or KMA but we remind that APF do not have any initialization stage. The value $RootsComplexity$ depends on the degree of the polynomial $Q(z)$. Table 5.1 gives the estimated number of FLOPS for BPSK, QPSK, PSK-8, and QAM16 modulations. Values have been obtained thanks to command `roots(.)` of Matlab.

| **Algorithm** | **Complexity** (Flops) |
|---|---|
| **CMA** | $358L^4 + 8L^3 + 70L^2 + 36L + 901$ |
| **KMA** | $8L^6 + 392L^4 + 8L^3 + 134L^2 + 76L$ $+2150$ |
| **APF** | $(N-L-1)(32L + 8l_q^2 + 82l_q - 28$ $+23\sum_{k=1}^{l_q-3} k) + 68l_q + 22L - 37$ $+RootsComplexity$ |

$l_q$ : length of polynomial $Q(z)$.

**Table 3**. Computational complexities of one iteration.

| $l_q$ | $RootsComplexity$ |
|---|---|
| 2 (BPSK) | 57 |
| 4 (QPSK) | 10457 |
| 8 (8-PSK) | 177846 |
| 16 (16-QAM) | 1844165 |

**Table 4**. Typical computational complexities for finding roots of $Q(z)$.

Finally, we can compute the global number of flops for the three algorithms for different values of $L$ and $N$. For instance, the length of the observations is $N = 1000$ symbols, and we consider that algorithms need 20 sweeps for estimating the equalizer vector taps. The number of flops is depicted in figure 1. From this figure, we

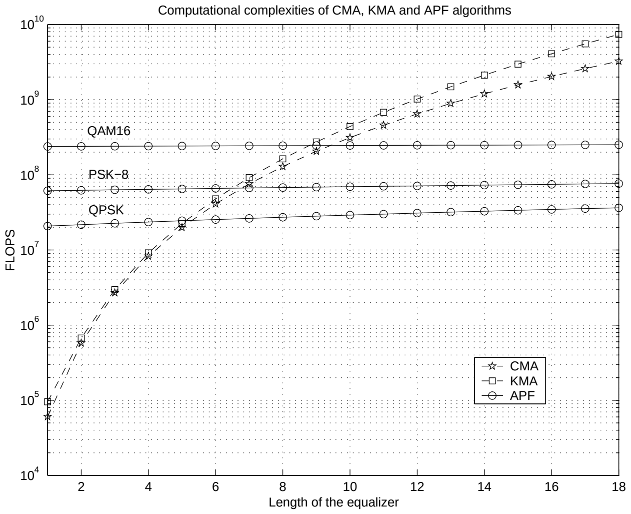


**Fig. 1**. Computational complexity for $N = 1000$ symbols and 20 loops.

see that computational complexities of APF algorithms are approximatively constant when the length of the equalizer increases, contrary to CMA and KMA which grow exponentially. For instance, for a QAM16 modulation, the APF algorithm is very attractive when $L \geq 8$. However, for $L \leq 5$, CMA and KMA require less flops than APF algorithms and then, they are less attractive for small lengths $L$.

### 5.2. Performances

The previous APF algorithms have been tested on complex channels of length $K = 3$, with unit variance QPSK and QAM16 white processes. For each randomly generated channel, blocks of noisy observations are filtered like in (1). We have tested CMA and APF algorithms on random channels with data block length of 1000 symbols. The length-6 equalizers returned by algorithms are then tested with the same sequence in order to compute the Symbol Error Rate (SER). Figure 2 shows median results for QPSK and QAM16 signals. This figure shows that the APF algorithm implemented for QAM16 signals works better than CMA when noise is greater than 10dB. For QPSK signals, the SER is zero for SNR greater than 10dB. For the two modulations considered, APF yields a zero median for a SNR of 20dB.

## 6. CONCLUDING REMARKS

Throughout this paper, a family of criteria based on Alphabet Polynomial Fitting has been introduced for blind SISO equalizers. Then, from theoretical results, numerical algorithms based on several polynomial criteria have been implemented in block form. Next, the comparison of CMA, KMA and APF computational complexities shows that APF algorithms are very attractive

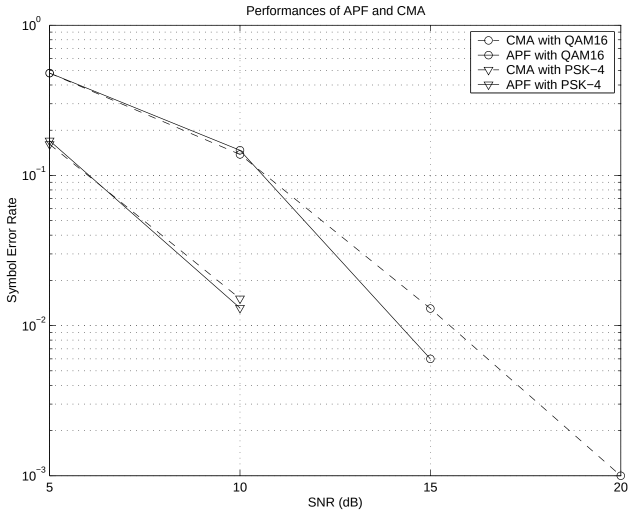


**Fig. 2**. Performances of APF and CMA with QPSK and QAM16 modulated signals.

for long equalizer vector taps, typically from length greater than 8. Moreover, simulations show that the improvement is relative to the modulation of the signal. Open issues currently being addressed include the robustness of APF algorithms in the presence of carrier residual.

## 7. 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] G. H. GOLUB and C. F. VAN LOAN. *Matrix Computation*. The John Hopkins University Press, 3rd edition, 1996.

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

[10] O. SHALVI and E. WEINSTEIN. New criteria for blind deconvolution of nonminimum phase systems. *IEEE Trans. on Information Theory*, 36(2):312–321, Mar. 1990.

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