LABORATOIRE

INFORMATIQUE, SIGNAUX ET SYSTÈMES
DE SOPHIA ANTIPOLIS
UMR 6070

BLIND AND SEMI-BLIND EQUALIZATION BASED ON THE CONSTANT POWER CRITERION

Vicente ZARZOSO, Pierre COMON

Projet ASTRE

Rapport de recherche
ISRN I3S/RR–2004-34–FR

Octobre 2004

LABORATOIRE I3S: Les Algorithmes / Euclide B – 2000 route des Lucioles – B.P. 121 –
06903 Sophia-Antipolis Cedex, France – Tél. (33) 492 942 701 – Télécopie : (33) 492 942 898
http://www.i3s.unice.fr/I3S/FR/

Blind and Semi-Blind Equalization Based on the Constant Power Criterion†

Vicente Zarzoso[1*] and Pierre Comon[2]

[1] Department of Electrical Engineering and Electronics, The University of Liverpool, Liverpool L69 3GJ, UK
vicente@liv.ac.uk

[2] Laboratoire I3S, Les Algorithmes – Euclide-B, BP 121, 06903, Sophia Antipolis, France
comon@i3s.unice.fr

Abstract

This paper focuses on the constant power (CP) criterion for blind linear equalization of digital communication channels. This recently proposed criterion is specially designed for the extraction of q-ary phase shift keying (q-PSK) signals using finite impulse response equalizers. When zero-forcing equalizers exist, the CP cost function accepts exact analytic solutions which are unaffected by undesired local extrema and spare costly iterative optimization. A subspace-based method exploiting the Toeplitz-like structure of the solution space is put forward to recover the minimum-length equalizer impulse response from the overestimated-length solutions. The proposed method is more robust to the equalizer vector configuration than existing techniques. In less ideal scenarios where the analytic solutions are only approximate minimizers of the criterion, a gradient-descent algorithm is proposed to minimize the cost function. To reduce the detrimental effects of spurious equilibria and accelerate convergence, the iterative algorithm is initialized with the approximate closed-form solution and an optimal step size is incorporated into its updating rule. This optimal step size, which globally minimizes the cost function along the search direction, can be computed algebraically. A semi-blind implementation, useful when training data are available, further reduces the impact of local extrema and enhances the convergence characteristics (particularly the robustness to the equalizer initialization) of the iterative algorithm from just a few pilot symbols. All these beneficial features are demonstrated with an experimental study of the proposed CP-based methods in a variety of channels and simulation conditions.

† Accepted in *IEEE Transactions on Signal Processing* (subject to minor revisions).

* Supported by a Post-Doctoral Research Fellowship awarded by the Royal Academy of Engineering of the UK.

I. Introduction

A. Background

In digital communications, transmission effects such as multipath propagation and limited bandwidth produce linear distortion in the emitted signal, causing intersymbol interference (ISI) at the receive sensor output. To enable the recovery of the input symbols, channel equalization aims to compensate these distorting effects [1]. Since the late 70's, the drawbacks of training-based methods [1], [2] have aroused considerable research interest in the so-called *blind* equalization techniques, which spare the use of bandwidth-consuming pilot sequences and prove especially attractive in broadcast and non-cooperative scenarios. In the fundamental single-input single-output (SISO) scenario, non-minimum phase (NMP) channels cannot be blindly identified using only second-order statistics (SOS); hence, the need for blind SISO equalizers to rely (explicitly or not) on higher-order statistics (HOS) [3]–[5]. Most blind methods are essentially property restoral techniques: the equalizer filter is updated so as to produce an output signal that recovers an *a priori* known property of the input signal, such as the finite alphabet or constant modulus of its data symbols.

The constant modulus (CM) criterion [4], [5] — which can be considered as a particular member of the more general family of Godard's methods [4] — is arguably the most widespread blind equalization principle. Although Godard's methods were proven to be globally convergent in the combined channel-equalizer parameter space, they were later shown to generally present spurious equilibria in the equalizer parameter space [6]. Spurious equilibria are those associated with filter tap settings which cannot sufficiently open the eye pattern of equalizer output signal, so that the detecting device is then unable to extract the transmitted symbols with a reasonably low probability of error. Often, these local extrema lie close to minimum mean square error (MMSE) solutions for equalization delays with high MSE. This shortcoming renders the performance of gradient-based implementations of Godard's criterion very dependent on the initial value of the equalizer impulse response. As discussed in [6], the misconvergence problems of iterative blind SISO equalization methods calls for the design of suitable initialization schemes and, perhaps, additional strategies to keep the equalizer tap trajectories away from undesired local equilibria. An alternative is to develop globally-convergent algorithms free from spurious extrema.

One such globally-convergent method is developed in [7], which obtains a closed-form solution for the CM equalizer. The CM criterion is posed as a nonlinear least squares (LS) problem. Through an appropriate mapping of the equalizer parameter space, the nonlinear setting is transformed into a linear LS problem subject to a constraint on the solution structure. Recovering the right structure of the solution space is particularly important when multiple zero-forcing (ZF) solutions exist; for

instance, in all-pole channels with overparameterized equalizers, different ZF equalization delays are possible. From a matrix algebra perspective, imposing this structure can be considered as a matrix diagonalization problem, in which the matrix performing the diagonalization of the unstructured solution matrix is composed of the equalizers' tap vectors. After obtaining a non-structured LS solution via pseudoinversion, the minimum-length equalizer is extracted by a subspace-based approach or two other simpler structuring procedures. LMS and RLS algorithms are also designed to solve the linear LS problem; hence, they still require structuring after convergence. Alternatively, the linearized LMS algorithm can be modified to partially impose the appropriate structure. However, the introduction of nonlinear constraints precludes the formulation of a closed-form solution.

The blind equalization method of [7] is strongly related to the analytical CM algorithm (ACMA) of [8] for blind source separation, a related but somewhat different problem. ACMA provides, in the noiseless case, exact closed-form solutions for the spatial filters which extract the source signals from their observed instantaneous linear mixtures. Interestingly, recovering the separating spatial filters from a basis of the solution space turns out to be tantamount to the joint diagonalization of the corresponding matrices. This joint diagonalization can be achieved through a QZ iteration for which convergence proof has not yet been found. Whether for source separation or equalization, ACMA requires special modifications to handle input signals with a one-dimensional (i.e., binary) alphabet [7]–[9]. These modifications give rise to the so-called real ACMA (RACMA) method [9].

Multichannel (fractionally-spaced) implementations are also able to avoid some of the deficiencies of SISO equalizers. In the first place, single-input multiple-output (SIMO) channels can be blindly identified using only SOS, regardless of their phase characteristics. Also, finite impulse response (FIR) SIMO channels can be perfectly equalized, in the absence of noise, by FIR filters. Seminal methods are presented in [10]–[12]. However, in certain practical scenarios it may not be possible to achieve the required degree of spatio-temporal diversity, due to lack of excess bandwidth or to hardware constraints limiting the number of receiving sensors (e.g., antennas in a mobile handset). This paper is mainly concerned with, but not restricted to, the basic SISO model.

B. Contribution and Outline

The present contribution studies a novel criterion for the blind equalization of digital channels excited by input signals with q-ary phase shift keying (q-PSK) modulations, for arbitrary $q \geqslant 2$. The criterion can be considered as a modification on the original Godard's family of blind equalizers, with a power value q matched to the signal constellation; thus the suitable name of *constant power (CP)* criterion. It is shown that if multiple ZF solutions exist — e.g., when the noiseless SISO channel follows a pure autoregressive (AR) model and the equalizer filter is of sufficient length — the criterion

accepts, much in ACMA's fashion, an exact solution which can be computed analytically, i.e., without iterative optimization. The minimum-length equalizer impulse response can then be obtained from a joint decomposition of qth-order tensors, the so-called rank-1 combination problem [13]. Since no effective tool has yet been developed for this task, an approximate solution is proposed in the form of a subspace-based method, which exploits the particular structure of the tensors associated with satisfactory equalization solutions. As opposed to [7], the subspace method proposed herein takes into account a whole basis of the solution space. This use of extra information is expected to increase the algorithm's robustness to the minimum-length equalizer structure. In addition, our closed-form blind equalization method naturally deals with binary inputs (e.g., BPSK, MSK) without further modification.

In additive noise or less ideal channel-equalizer conditions, the CP cost function can be minimized through a gradient-descent algorithm. The impact of non-equalizing extrema are considerably reduced by initializing the algorithm with the approximate closed-form solution. In computationally-limited systems, however, simple initializations may be preferred to more sophisticated, and thus more complex, alternatives. Whatever the option, the value of the step size (adaption coefficient) that globally minimizes the cost function along the search direction can be computed analytically at each iteration. This optimal step size provides remarkable benefits in convergence speed and avoidance of spurious local extrema, even with conventional (e.g., center-tap) initializations. The CP criterion is easily modified to operate in semi-blind mode, relevant in typical real scenarios where training sequences are available. The optimal step size can also be algebraically computed in pilot-assisted operation. Using just a few pilot symbols, this semi-blind optimal step size algorithm shows an outstanding robustness to the equalizer filter initialization.

The material is organized as follows. A brief explanation of the problem and the signal model is given in Section II. After presenting the CP criterion in Section III, its closed-form solutions are found in Section IV with the aid of a subspace-based algorithm for recovering the minimum-length equalizer. Iterative implementations are the focus of Section V, featuring the optimal step-size gradient-descent algorithm. Semi-blind solutions, in block and iterative operation, are put forward in Section VI. An experimental study is reported in Section VII. Finally, the summary and concluding remarks of Section VIII bring the paper to an end. For the sake of clarity, proofs and other mathematical derivations are postponed to the Appendix.

C. Notations

In the following, scalars, vectors, and tensors (of which matrices are assumed a particular case) will usually be denoted by plain lowercase (a), boldface lowercase ($\mathbf{a}$) and boldface uppercase ($\mathbf{A}$)

symbols, respectively, the only exceptions being the structures derived from Kronecker tensorial products, as explained below. $\mathbf{I}_n$ refers to the $(n \times n)$ identity matrix, whereas $\mathbf{0}_n$ is the length-n zero vector; $(\cdot)^{\mathrm{T}}$, $(\cdot)^{\mathrm{H}}$ and $(\cdot)^{-1}$ indicate the transpose, Hermitian (conjugate-transpose) and inverse matrix operators, respectively; $\|\cdot\|$ is the conventional 2-norm. $(\mathbf{A})_{i_1 i_2 \ldots i_q}$ denotes the entry located in position $(i_1, i_2, \ldots, i_q)$ of qth-order tensor $\mathbf{A}$. $\mathbb{C}$ is the set of complex numbers; $\mathbb{R}\mathrm{e}(\cdot)$ and $\mathbb{I}\mathrm{m}(\cdot)$ denote the real and imaginary part, respectively, of their complex argument; $\mathrm{E}\{\cdot\}$ represents the mathematical expectation. Symbol $*$ denotes the convolution operator, whereas $\otimes$, $\odot$ and $\circ$ stand for the Kronecker, elementwise and outer products, respectively. Given a vector $\mathbf{a} \in \mathbb{C}^L$ we define its qth-order rank-1 Kronecker tensor product as $\mathbf{a}^{\otimes q} = \underbrace{\mathbf{a} \circ \cdots \circ \mathbf{a}}_{q}$ (e.g., $\mathbf{a}^{\otimes 2} = \mathbf{a} \circ \mathbf{a} = \mathbf{a}\mathbf{a}^{\mathrm{T}}$). A symmetric tensor $\mathbf{A}$ of order q and dimension L can be stored in a vector $\mathbf{vecs}\{\mathbf{A}\}$, which contains only the $L_q = \binom{L+q-1}{q}$ distinct entries of $\mathbf{A}$, scaled by the square root of the number of times they appear so that the Frobenius norm is preserved [13]. In particular, we denote $\mathbf{a}^{\oslash q} = \mathbf{vecs}\{\mathbf{a}^{\otimes q}\}$. Similarly, given a vector $\mathbf{b}$ of dimension L_q, $\mathbf{unvecs}_q\{\mathbf{b}\}$ denotes the symmetric qth-order tensor constructed from its entries.

II. Problem Statement and Signal Model

The problem of channel equalization can simply be posed as follows. A digital signal $s(t) = \sum_n s_n \delta(t - nT)$ is transmitted at a known baud-rate $1/T$ through a time dispersive channel with impulse response $h(t)$. The channel is assumed linear and time-invariant (at least over the observation window), with a stable, causal and possibly non-minimum phase transfer function. The continuous-time baseband signal at the receive sensor output is given by $x(t) = r(t) + v(t)$, where $r(t) = h(t) * s(t)$ denotes the noiseless observation and $v(t)$ the additive noise. Assuming perfect synchronization and carrier-residual elimination, baud-rate sampling produces the discrete-time output

$$x_n = r_n + v_n = \sum_k h_k s_{n-k} + v_n \tag{1}$$

in which $x_n = x(nT)$, and analogous definitions hold for h_k, s_n and v_n. Each observed sample consists of a noisy linear mixture of the original data symbols, an undesired phenomenon known as ISI. Our goal is to recover the original data symbols from the received signal corrupted by ISI and noise. To this end, a baud-spaced linear equalizer with impulse response taps $\mathbf{f} = [f_1, \ldots, f_L]^{\mathrm{T}} \in \mathbb{C}^L$ is sought so that the equalizer output $y_n = \mathbf{f}^{\mathrm{H}} \mathbf{x}_n$ is a close estimate of the source symbols s_n, where $\mathbf{x}_n = [x_n, x_{n-1}, \ldots, x_{n-L+1}]^{\mathrm{T}}$.

In this paper, the data symbols are assumed to belong to a q-ary phase shift keying (q-PSK) constellation $\mathcal{A}_q = \{a_k\}_{k=0}^{q-1}$, with $a_k = a^k$, in which $a^q = d$ depends on the actual constellation; for instance, $(q, d) = (2, 1)$ for BPSK and $(q, d) = (4, -1)$ for QPSK. By allowing a time-varying d,

the above definitions are readily extended to encompass other non-PSK modulations such as MSK [14], modelled with $(q,\, d_n) = \big(2,\, (-1)^n\big)$. Note that set $\mathcal{A}_q$ is an Abelian group under complex multiplication.

III. A Blind Equalization Criterion for PSK Modulations

A. The Constant Power Criterion

Since $s_n \in \mathcal{A}_q$ it follows that $s_n^q = d_n$. Thus, a somewhat natural cost function to measure the closeness of the equalizer output to the original data symbols is given by the *constant power (CP)* criterion:

$$J_q(\mathbf{f}) = \mathrm{E}\{|y_n^q - d_n|^2\} = \mathrm{E}\{\big|(\mathbf{f}^{\mathrm{H}}\mathbf{x}_n)^q - d_n\big|^2\}. \tag{2}$$

Cost function (2) is a particular case of the more general class of alphabet polynomial fitting (APF) criteria, where the equalizer output constellation is matched to the source alphabet, characterized by the complex roots of a specific polynomial [15], [16]. In the context of blind source separation, criterion (2) is shown to be equivalent, for sufficiently low noise levels, to the maximum a posteriori (MAP) principle [17], [18]. In addition, it is proved in [15] that, when the total channel-equalizer impulse response is of finite length and the input signal sufficiently exciting, the global minima in the combined noiseless channel and equalizer parameter space of the sample estimate of criterion (2) correspond to ZF equalization solutions. However, this result does not assure that the desired solutions can always be reached or that undesired (non-equalizing) equilibria do not exist when the cost function is observed from the actual equalizer parameter space, as noted in [6], [19] for Godard's criterion. The existence of local extrema in the CP criterion will be illustrated with a few simple experiments in Section VII.

A sufficient excitation condition for input sequence s_n consists of observing all possible q^K states of the K-uplet $[s_n,\, s_{n-1},\, \ldots,\, s_{n-K+1}]^{\mathrm{T}}$, where K denotes the total length of the cascaded channel-equalizer impulse response [15]. The group structure of the PSK alphabet enables the reduction of the sufficiently-exciting observation length to q^{K-1} (Appendix A).

B. Connections with Existing Criteria

CP functional (2) bears close resemblance to Godard's class of cost functions [4], which in the PSK case shows the general form:

$$J_q^{\mathrm{G}}(\mathbf{f}) = \mathrm{E}\{(|y_n|^q - |s_n|^q)^2\} = \mathrm{E}\big\{\big(|\mathbf{f}^{\mathrm{H}}\mathbf{x}_n|^q - 1\big)^2\big\}. \tag{3}$$

For $q = 2$, the above function corresponds to the CM criterion [4], [5]. For BPSK sources and real-valued channel and equalizer, the CP and CM criteria are identical; in such a case, we anticipate that

the closed-form treatment of the CP minimization (Section IV) is equivalent to that of the specialized ACMA for binary modulations [7], [9]. All PSK constellations being CM, the CM principle is not discriminant over the set of PSK constellations. Similarly, it is not clear, at least at first glance, how the more general criterion (3) could privilege a particular PSK modulation. By contrast, criterion (2) explicitly takes into account the discrete nature of PSK-type alphabets, so that it should exhibit enhanced discriminating properties among the CM constellations.

If d_n is substituted by the available training symbols s_n^{t}, the CP cost function (2) reduces, with $q = 1$, to the supervised MMSE equalization principle. This fact will be revisited when designing the semi-blind methods of Section VI.

IV. Blind Closed-form Solutions

When the channel accepts a noiseless AR model and the FIR equalizer is sufficiently long, perfect ZF SISO equalization is possible. In particular, the CP criterion (2) can be perfectly minimized (zeroed) and an exact global minimum can be computed in closed-form, that is, without iterative optimization. This analytic solution can be considered as an extension of the ACMA algorithm [8] to the CP principle. Consequently, the method may be called *analytical constant power algorithm (ACPA)*.

A. Obtaining a Basis of the Solution Space

The perfect minimizers of (2) are given by the solutions to the set of equations:

$$(\mathbf{f}^{\mathrm{H}}\mathbf{x}_n)^q = d_n, \qquad n = 0,\, 1,\, \ldots,\, N-1 \tag{4}$$

where $N = (N_d - L + 1)$ and N_d denotes the observation length in number of samples. This non-linear system can be linearized by taking into account that $(\mathbf{f}^{\mathrm{H}}\mathbf{x}_n)^q = \mathbf{f}^{\oslash q\mathrm{H}}\mathbf{x}_n^{\oslash q}$ (Appendix A), and can be compactly expressed as

$$\mathbf{X}^{q\mathrm{H}}\mathbf{w} = \mathbf{d} \tag{5}$$

where $\mathbf{X}^q = [\mathbf{x}_0^{\oslash q},\, \mathbf{x}_1^{\oslash q},\, \ldots,\, \mathbf{x}_{N-1}^{\oslash q}]$ and $\mathbf{d} = [d_0,\, d_1,\, \ldots,\, d_{N-1}]^{\mathrm{H}}$. Eqn. (5) is to be solved under the structural constraint that $\mathbf{w}$ be written as $\mathbf{w} = \mathbf{f}^{\oslash q}$, for certain $\mathbf{f} \in \mathbb{C}^L$ (Appendix A).

Let us assume an all-pole channel with AR-model order of M. Such a channel can be equalized with a minimum-length FIR filter $\mathbf{f}_0$ composed of $L_0 = (M+1)$ taps. Assume the equalizer filter is overparameterized, that is, the equalizer length L has been overestimated, $L \geqslant L_0$. Then $P = (L - L_0 + 1)$ ZF solutions exist, each of them corresponding to a different equalization delay:

$$\mathbf{f}_p = [\mathbf{0}_{p-1}^{\mathrm{T}},\, \mathbf{f}_0^{\mathrm{T}},\, \mathbf{0}_{P-p}^{\mathrm{T}}]^{\mathrm{T}}, \qquad 1 \leqslant p \leqslant P. \tag{6}$$

Since there are P linearly independent solutions, the dimension of the null space of $\mathbf{X}^{q\mathrm{H}}$ is equal to $(P-1)$. Hence, the solutions to (5) can be written as an affine space of the form $\mathbf{w} = \mathbf{w}_0 + \sum_{p=1}^{P-1} \alpha_p \mathbf{w}_p$, where $\mathbf{w}_0$ is a particular solution to the non-homogeneous system and $\mathbf{w}_p \in \ker(\mathbf{X}^{q\mathrm{H}})$, $1 \leqslant p \leqslant (P-1)$.

As in [8], we find it more convenient to work in a fully linear subspace, which is obtained through a $(N \times N)$ unitary transformation $\mathbf{Q}$ such that $\mathbf{Q}\mathbf{d} = [\sqrt{N}, \mathbf{0}_{N-1}^{\mathrm{T}}]^{\mathrm{T}}$. For instance, $\mathbf{Q}$ can be a Householder transformation [20] or, if $\mathbf{d}$ is composed of N equal values, an N-point DFT matrix. Then,

$$\mathbf{Q}\mathbf{X}^{q\mathrm{H}} = \begin{bmatrix} \mathbf{r}^{\mathrm{H}} \\ \mathbf{R} \end{bmatrix} \tag{7}$$

and system (5) reduces to:

$$\begin{cases} \mathbf{r}^{\mathrm{H}}\mathbf{w} = \sqrt{N} \\ \mathbf{R}\mathbf{w} = \mathbf{0}_{N-1}. \end{cases} \tag{8}$$

subject to the constraint $\mathbf{w} = \mathbf{f}^{\oslash q}$. Along the lines of [8, Lemma 4], it can be proved (Appendix A) that this problem is equivalent to solving

$$\begin{cases} \mathbf{R}\mathbf{w} = \mathbf{0}_{N-1} \\ \mathbf{w} = \mathbf{f}^{\oslash q} \end{cases} \tag{9}$$

and then scaling the solution to impose

$$\mathbf{c}^{\mathrm{H}}\mathbf{w} = 1, \qquad \text{with} \quad \mathbf{c} = \frac{1}{\|\mathbf{d}\|^2} \sum_{n=0}^{N-1} d_n \mathbf{x}_n^{\oslash q} \tag{10}$$

or, equivalently,

$$\frac{1}{\|\mathbf{d}\|^2} \sum_{n=0}^{N-1} d_n (\mathbf{f}^{\mathrm{H}}\mathbf{x}_n)^q = 1. \tag{11}$$

If $\dim\ker(\mathbf{X}^{q\mathrm{H}}) = (P-1)$ and

$$N_d \geqslant L_q + L_0 - 1 \tag{12}$$

(or $N > L_q - P$), then $\dim\ker(\mathbf{R}) = P$ (Appendix A). Hence, all solutions to $\mathbf{R}\mathbf{w} = \mathbf{0}$ are linearly spanned by a basis $\{\mathbf{w}_k\}_{k=1}^{P}$ of $\ker(\mathbf{R})$. This basis can be computed from the SVD of $\mathbf{R}$ by taking its P least significant right singular vectors. The structured solutions $\{\mathbf{f}_p^{\oslash q}\}_{p=1}^{P}$ are also a basis of the same subspace and, therefore, a set of scalars $\{\alpha_{pk}\}_{p,k=1}^{P}$ exists such that

$$\mathbf{f}_p^{\oslash q} = \sum_{k=1}^{P} \alpha_{pk} \mathbf{w}_k, \qquad 1 \leqslant p \leqslant P \tag{13}$$

where matrix $(\mathbf{A})_{kp} = \alpha_{pk}$ is full rank. The problem of structuring the solution to the linearized system (5) consists of imposing the rank-1 symmetric Kronecker structure to the basis $\{\mathbf{w}_k\}_{k=1}^{P}$, which, in

turn, yields $\{\mathbf{f}_p\}_{p=1}^{P}$. This is a particular subspace-fitting problem with structural constraints. In terms of q-order tensors, eqn. (13) can be expressed as

$$\mathbf{f}_p^{\otimes q} = \sum_{k=1}^{P} \alpha_{pk} \mathbf{W}_k, \qquad 1 \leqslant p \leqslant P \tag{14}$$

where $\mathbf{W}_k = \mathbf{unvecs}_q\{\mathbf{w}_k\}$. This is the rank-1 combination problem: given the set $\{\mathbf{W}_k\}$, find the scalars producing tensors of rank one. The obtained rank-1 tensors will precisely correspond to $\{\mathbf{f}_p^{\otimes q}\}$. Such a tensor decomposition is, in general, a notoriously non-trivial task (see, e.g., [13], [21] and references therein).

Before continuing, it is worth remarking that sample-size bound (12) is too restrictive. In practice, satisfactory closed-form equalization usually requires shorter observation windows, as will be demonstrated in the experiments of Section VII.

B. Solution Structuring: A Subspace-Based Approach

A subspace-based method, reminiscent of [12], can be used to recover the minimum-length equalizer impulse response $\mathbf{f}_0$ from a basis of (generally) unstructured solutions $\{\mathbf{w}_k\}_{k=1}^{P}$. The subspace-fitting problem (13) can be compactly written as $\mathbf{WA} = \mathbf{F}$, with $\mathbf{W} = [\mathbf{w}_1, \ldots, \mathbf{w}_P]$ and $\mathbf{F} = [\mathbf{f}_1^{\varnothing q}, \ldots, \mathbf{f}_P^{\varnothing q}]$. Since $\mathbf{A}$ is full rank, matrices $\mathbf{W}$ and $\mathbf{F}$ span the same column space: range($\mathbf{W}$) = range($\mathbf{F}$). In particular, $\forall \mathbf{u}_i \in \ker(\mathbf{W}^{\mathrm{H}})$, $\mathbf{u}_i^{\mathrm{H}}\mathbf{F} = \mathbf{0}_P^{\mathrm{T}}$. There are $\dim\ker(\mathbf{W}^{\mathrm{H}}) = (L_q - P)$ such linearly independent vectors.

Now, since equalization solutions are of the form (6), the corresponding columns of $\mathbf{F}$ have a particular structure whereby the elements not associated with the minimum-length equalizer $\mathbf{f}_0$ are all zero. The remaining $L_{0q} = \binom{L_0+q-1}{q}$ entries form $\mathbf{f}_0^{\varnothing q}$. Denote by σ_p the set of L_{0q} positions of $\mathbf{f}_0^{\varnothing q}$ in $\mathbf{f}_p^{\varnothing q}$, that is, $\sigma_p = \left\{j_1 + L(j_2 - 1) + \cdots + L^{q-1}(j_q - 1)\right\}$, with $j_k \in [p,\, p + L_0 - 1]$, $k = 1, \ldots, q$, and $j_1 \geqslant j_2 \geqslant \ldots \geqslant j_q$. Accordingly, $(\mathbf{u}_i)_{\sigma_p} \in \mathbb{C}^{L_{0q}}$ is the subvector composed of the elements $\mathbf{u}_i$ in positions σ_p. Let $\mathbf{U}_i = [(\mathbf{u}_i)_{\sigma_1}, \ldots, (\mathbf{u}_i)_{\sigma_P}] \in \mathbb{C}^{L_{0q} \times P}$. Hence,

$$\mathbf{u}_i^{\mathrm{H}}\mathbf{F} = \mathbf{0}_P^{\mathrm{T}} \quad \Leftrightarrow \quad \mathbf{U}_i^{\mathrm{H}}\mathbf{f}_0^{\varnothing q} = \mathbf{0}_P. \tag{15}$$

In total, the above equalities define a set of $P(L_q - P)$ linear equations, characterized by matrix $\mathbf{U} = [\mathbf{U}_1, \ldots, \mathbf{U}_{L_q - P}] \in \mathbb{C}^{L_{0q} \times P(L_q - P)}$, on the entries of $\mathbf{f}_0^{\varnothing q}$. As long as $L > L_0$, this linear system determines, up to a scale, the properly structured $\mathbf{f}_0^{\varnothing q}$; its scale can later be set via (11) from a zero-padded version (any $\mathbf{f}_p$) of the estimated $\mathbf{f}_0$. In practice, $\mathbf{f}_0^{\varnothing q}$ can be estimated as the least significant left singular vector of matrix $\mathbf{U}$. Once matrix $\mathbf{F}$ has been reconstructed, an LS estimate of coefficients $\{\alpha_{kp}\}$ can be obtained as $\hat{\mathbf{A}}_{\mathrm{LS}} = (\mathbf{W}^{\mathrm{H}}\mathbf{W})^{-1}\mathbf{W}^{\mathrm{H}}\mathbf{F} = \mathbf{W}^{\dagger}\mathbf{F}$. These coefficients relate qth-order tensors $\{\mathbf{W}_k\}$ with their rank-1 symmetric tensor decomposition (14). Hence, the elements of $\hat{\mathbf{A}}_{\mathrm{LS}}$ solve the rank-1 combination problem.

To recover the equalizer impulse response $\mathbf{f}_0$ from its symmetric Kronecker vectorization $\mathbf{f}_0^{\oslash q}$, one can resort to the SVD of a matrix unfolding of $\mathbf{f}_0^{\otimes q} = \mathbf{unvecs}_q\{\mathbf{f}_0^{\oslash q}\}$ [22], [23]. Let matrix $\mathbf{F}_0 \in \mathbb{C}^{L_0 \times L_0^{q-1}}$ such that $(\mathbf{F}_0)_{i_1, i_2+L_0(i_3-1)+\ldots L_0^{q-2}(i_q-1)} = (\mathbf{f}_0^{\otimes q})_{i_1 i_2 i_3 \ldots i_q}$. Then, $\mathbf{F}_0 = \mathbf{f}_0 \bar{\mathbf{f}}_0^{\mathrm{T}}$, with $(\bar{\mathbf{f}}_0)_{i_2+L_0(i_3-1)+\ldots L_0^{q-2}(i_q-1)} = (\mathbf{f}_0)_{i_2}(\mathbf{f}_0)_{i_3} \ldots (\mathbf{f}_0)_{i_q}$. Therefore, $\mathbf{f}_0$ can be estimated (up to a scale) as the most significant left singular vector of the rank-1 matrix unfolding $\mathbf{F}_0$. In the presence of noise, it will generally be impossible to express the estimated $\hat{\mathbf{f}}_0^{\oslash q}$ as the symmetric vectorization of a rank-1 tensor; that is, a vector $\mathbf{f}_0$ cannot be found such that $\hat{\mathbf{f}}_0^{\oslash q} = \mathbf{vecs}_q\{\mathbf{f}_0^{\otimes q}\}$ holds exactly. As a result, the matrix unfolding will not be of rank one, and the above SVD-based procedure will yield inaccuracies which may ultimately limit the equalization performance.

C. Other Structuring Methods

In the context of the CM criterion, a similar subspace-based structuring method was proposed in [7, Section III.C], which operates on a single (LS) unstructured solution (see also [24]). Such structure-forcing procedure can be interpreted as the diagonalization of the matrix associated to the unstructured solution. By contrast, our approach takes advantage of a full basis of the solution subspace, which should lead to a subsequent increase in robustness, especially for large P. The method of [22] and [7, Section III.B] is based on the observation that the top L entries of a solution $\mathbf{w}_k$ are equal to $\alpha_{k1} f_1^{q-1}[f_1,\, \sqrt{q} f_2,\, \ldots,\, \sqrt{q} f_{L_0-1},\, \sqrt{q} f_{L_0},\, \mathbf{0}_{P-1}^{\mathrm{T}}]^{\mathrm{T}}$, from which $\mathbf{f}_0$ can be extracted. This ingenious simple method is bound to be inaccurate when either the coefficient α_{k1} or the equalizer leading tap f_1 are small relative to the noise level.

To circumvent this drawback, one may notice that the L entries at the bottom of $\mathbf{w}_k$ are equal to $\alpha_{kP} f_{L_0}^{q-1}[\mathbf{0}_{P-1}^{\mathrm{T}},\, \ldots,\, \sqrt{q} f_1,\, \sqrt{q} f_2,\, \ldots,\, \sqrt{q} f_{L_0-1},\, f_{L_0}]^{\mathrm{T}}$ [7, Section III.B]. This second option can provide, properly combined with the estimate from the first L entries, an improved estimate of $\mathbf{f}_0$. In the experiments of Section VII, we use the following (still suboptimal) LS linear combination. Assume that the filter estimate from the top and bottom non-overlapping entries of an unstructured solution are, respectively, $\hat{\mathbf{f}}_1 = \beta_1 \tilde{\mathbf{f}}_0$ and $\hat{\mathbf{f}}_2 = \beta_2 \tilde{\mathbf{f}}_0$, with $\tilde{\mathbf{f}}_0 = \mathbf{f}_0 / \|\mathbf{f}_0\|$. Then, the unit-norm minimum-length equalizer LS estimate is given by $\hat{\tilde{\mathbf{f}}}_0 = [\hat{\mathbf{f}}_1,\, \hat{\mathbf{f}}_2]\boldsymbol{\gamma}$, with $\boldsymbol{\gamma} = \boldsymbol{\beta}^* / \|\boldsymbol{\beta}\|^2$, $\boldsymbol{\beta} = [\beta_1,\, \beta_2]^{\mathrm{T}}$. The coefficients in $\boldsymbol{\beta}$ are simply estimated as $\beta_i = \|\hat{\mathbf{f}}_i\|$, $i = 1, 2$. This kind of maximal-ratio combining (MRC)is reminiscent of the RAKE receiver and the matched filter [25]. Robustness can be further enhanced by exploiting a whole set $\{\mathbf{w}_k\}$ instead of just one solution.

D. Approximate Solution in the Presence of Noise

In the presence of additive noise at the receive sensor output, the exact solution to (4) may no longer exist. An approximate solution in the LS sense can be reached by minimizing $\|\mathbf{X}^{q\mathrm{H}}\mathbf{w} - \mathbf{d}\|^2$,

always subject to the structural constraint $\mathbf{w} = \mathbf{f}^{\oslash q}$. This minimization generally requires an iterative method, as will be detailed in the next section.

Nevertheless, the guidelines to obtain the exact solution in the noiseless case may still provide a judicious initialization for the iterative search. After applying transformation $\mathbf{Q}$ the LS problem turns out to be equivalent to the minimization of $|\mathbf{c}^{\mathrm{H}}\mathbf{w}-1|^2+\|\mathbf{R}\mathbf{w}\|^2$. To find a basis of the (approximate) solution space, we look for a set of vectors which minimize $\|\mathbf{R}\mathbf{w}\|^2$ (e.g., the P least significant right singular vectors of $\mathbf{R}$), then structure them as in Section IV-B, and finally normalize the solution to fulfil $\mathbf{c}^{\mathrm{H}}\mathbf{w} = 1$ [eqns. (10)–(11)].

V. Blind Iterative Solutions

A. Gradient-Based Algorithm

In practice, exact ZF equalization may not be feasible, due to the presence of noise, the existence of an FIR SISO channel, or just because the equalizer length is insufficient. In such cases, the CP cost function must be iteratively minimized, e.g., via a gradient-descent algorithm. The gradient of function (2) with respect to $\mathbf{f}$ is given by $\nabla J_q(\mathbf{f}) = \nabla_{\mathbb{R}\mathrm{e}(\mathbf{f})} J(\mathbf{f}) + j\nabla_{\mathbb{I}\mathrm{m}(\mathbf{f})} J_q(\mathbf{f})$ and can be expressed as:

$$\nabla J_q(\mathbf{f}) = \mathrm{E}\big\{(\mathbf{f}^{\mathrm{H}}\mathbf{x}_n)^{q-1}[(\mathbf{f}^{\mathrm{H}}\mathbf{x}_n)^q - d_n]^*\mathbf{x}_n\big\}. \tag{16}$$

We refer to the resulting iterative method as *constant power algorithm (CPA)*. As a sensible initialization, one can use the equalizer vector provided by an ACPA method, such as the approximate structured solution described in Section IV-D or the unstructured LS solution to the linearized problem (5), $\hat{\mathbf{f}}_{\mathrm{LS}} = (\mathbf{X}^{q\mathrm{H}})^{\dagger}\mathbf{d}$. At each iteration, the equalizer vector is updated in the LMS fashion as

$$\mathbf{f}_{k+1} = \mathbf{f}_k - \mu\nabla J_q(\mathbf{f}_k), \qquad k = 0,\, 1,\, \ldots \tag{17}$$

The iterations are terminated when

$$\frac{\|\mathbf{f}_{k+1} - \mathbf{f}_k\|}{\|\mathbf{f}_k\|} < \eta/N \tag{18}$$

where η is a small positive constant.

We advocate the use of block (or 'windowed') iterative implementations, as opposed to stochastic algorithms. The latter methods approximate the gradient by using a one-sample estimate, which is tantamount to dropping the expectation operator. This simplification generally leads to extremely slow convergence and poor misadjustment. By contrast, the former methods approximate the gradient by its sample estimate from a block of channel output samples. This more precise gradient estimate improves convergence speed and accuracy [18], [26]. In addition, tracking capabilities are not necessarily

sacrificed, since good performance can be obtained from rather small block sizes; it suffices that the channel be stationary over the (short) observation window.

It is well known that gradient-based optimization algorithms, though simple, are plagued with a number of drawbacks, such as convergence to spurious local extrema, lack of robustness to initialization, and slow convergence [6], [19], [27]. These problems persist in block implementations, although convergence is often faster. When the function to be optimized is quadratic in the unknowns, more elaborate approaches such as conjugate-direction algorithms alleviate these shortcomings [28]. However, the fact that function (2) is not quadratic leads us to seek alternative optimization strategies.

B. Closed-Form Steepest Descent

Steepest descent (or exact line search) methods look for the value of the step size which minimizes the cost function along the search direction:

$$\mu_{\text{opt}} = \arg\min_{\mu} J_q(\mathbf{f} - \mu\mathbf{g}). \tag{19}$$

A sensible search direction is the gradient, $\mathbf{g} = \nabla J_q(\mathbf{f})$. These algorithms are generally unattractive due to their complexity, for the one-dimensional minimization must usually be performed using costly numerical methods. Another drawback is the orthogonality of consecutive gradient vectors, which, depending on the initialization and the shape of the cost-function surface, may slow down convergence [28].

However, it is observed in [15], [23], [29] that the CP cost $J_q(\mathbf{f} - \mu\mathbf{g})$ is a rational function in the step size μ, so that μ_{opt} can be found in closed form. This fact allows the *global* line minimization of the cost function while reducing complexity. In effect, μ_{opt} can be found among the roots of the $(2q-1)$th-degree polynomial $p(\mu) = \mathbb{R}\text{e}\big(\sum_{m=1}^{2q-1} b_m \mu^m\big)$ where

$$b_m = \begin{cases} \sum_{p=0}^{m} (m+1-p)\mathrm{E}\{a^*_{m+1-p} a_p\} - (m+1)\mathrm{E}\{a^*_{m+1} d_n\}, & 0 \leqslant m \leqslant q-1 \\ \sum_{p=m+1-q}^{q} (m+1-p)\mathrm{E}\{a^*_{m+1-p} a_p\}, & q \leqslant m \leqslant 2q-1 \end{cases} \tag{20}$$

with $a_p = (-1)^p \binom{q}{p} (\mathbf{g}^{\mathrm{H}}\mathbf{x}_n)^p (\mathbf{f}^{\mathrm{H}}\mathbf{x}_n)^{q-p}$, $0 \leqslant p \leqslant q$ (Appendix A). The cost function can then be evaluated at the candidate roots in order to find the global minimum along direction $\mathbf{g}$. Numerical conditioning is improved by normalizing vector $\mathbf{g}$ before evaluating (20).

Although undesired equilibria (especially those lying near flat areas) are not avoided in all cases, our experiments indicate that this *optimal step-size CPA (OS-CPA)* converges much faster and more accurately than the CPA with constant adaption coefficient. In addition, the frequency of misconvergence to spurious non-equalizing solutions is remarkably diminished. These benefits will be demonstrated in Section VII. An analogous optimal step-size algorithm for the CM criterion (OS-CMA) is developed in [30].

VI. Semi-Blind Equalization

A. Semi-Blind CP-Based Criterion

The previous sections have developed CP-based equalization algorithms in the fully blind case. However, practical communication systems typically feature pilot sequences to aid synchronization and channel equalization. For example, the second-generation GSM wireless system uses 26 out of the 148 bits in its data frame for training. Exploiting this available information can notably improve equalization performance. In order to take advantage of these benefits, the CP criterion can be easily modified to incorporate training symbols, resulting in a semi-blind equalization method. The minimization of the following hybrid cost function constitutes a semi-blind CP-MMSE criterion:

$$J_{\mathrm{SB}}(\mathbf{f}) = \lambda J_{\mathrm{MMSE}}(\mathbf{f}) + (1-\lambda) J_q(\mathbf{f}) \tag{21}$$

where $J_{\mathrm{MMSE}}(\mathbf{f}) = \mathrm{E}\{|y_n - s_n^{\mathrm{t}}|^2\}$ is the pilot-based MMSE cost function, and $\{s_n^{\mathrm{t}}\}$ denote the available training symbols. Parameter λ is a real constant in the interval $[0,\ 1]$ which can be considered as the relative degree of confidence between the blind- and the training-based parts of the criterion. By looking at expression (2), it turns out that J_{MMSE} can be derived from J_q by setting $q = 1$ and substituting s_n^{t} for d_n. This equivalence will be useful in simplifying some of the following mathematical derivations. As in the blind scenario, closed-form and iterative solutions for this semi-blind CP-based criterion exist and are developed next.

B. Semi-Blind Closed-Form Solutions

Assume N_t training symbols are transmitted and are known to the receiver. We are looking for the simultaneous solution of the compound system

$$\mathbf{X}^{\mathrm{H}}\mathbf{f} = \mathbf{s} \tag{22}$$

$$\mathbf{X}^{q\mathrm{H}}\mathbf{w} = \mathbf{d} \tag{23}$$

subject to $\mathbf{w} = \mathbf{f}^{\oslash q}$, with $\mathbf{X} = [\mathbf{x}_0,\, \mathbf{x}_1,\, \ldots,\, \mathbf{x}_{N_t-1}]$, $\mathbf{X}^q = [\mathbf{x}_{N_t}^{\oslash q},\, \mathbf{x}_{N_t+1}^{\oslash q},\, \ldots,\, \mathbf{x}_{N-1}^{\oslash q}]$, $\mathbf{s} = [s_0,\, s_1,\, \ldots,\, s_{N_t-1}]^{\mathrm{H}}$, and $\mathbf{d} = [d_{N_t},\, d_{N_t+1},\, \ldots,\, d_{N-1}]^{\mathrm{H}}$.

Firstly, let us consider the case of a possibly noisy AR-channel with a sufficiently long equalizer. An approximate suboptimal solution can be found by combining the solutions computed separately for both systems. Let $\hat{\mathbf{f}}_{\mathrm{MMSE}}$ be the solution to (22), and $\hat{\mathbf{f}}_{\mathrm{CP}}^{\oslash q}$ the same delay solution to (23), computed as in Section IV. Unfold $\mathbf{unvecs}_q\{\hat{\mathbf{f}}_{\mathrm{CP}}^{\oslash q}\}$ into an $(L \times L^{q-1})$ matrix $\mathbf{F}_{\mathrm{CP}}$ as described at the end of Section IV-B. Then, the joint solution to (22)–(23) can be approximated as the left singular vector of matrix $\mathbf{F}_{\mathrm{SB}} = [\lambda\hat{\mathbf{f}}_{\mathrm{MMSE}},\, (1-\lambda)\mathbf{F}_{\mathrm{CP}}]$. In the noiseless case, solutions $\hat{\mathbf{f}}_{\mathrm{MMSE}}$ and $\hat{\mathbf{f}}_{\mathrm{CP}}$ are

exact, identical, and equal to the left singular vector of rank-1 matrix $\mathbf{F}_{\mathrm{SB}}$; an iterative search is not necessary.

In the case of an FIR channel, no exact solution to (22)–(23) exists even in the absence of noise. Still, the systems can be solved separately in the LS sense, and their respective solutions combined by the SVD-based procedure just described. The combined solution can initialize an iterative search aiming to refine the approximate closed-form result.

C. Semi-Blind Iterative Solutions

As in the fully-blind case, cost function (21) can be iteratively minimized using a steepest-descent gradient-based algorithm in which the the optimal step size can be algebraically computed at each iteration. The equalizer impulse response is updated as:

$$\mathbf{f}_{k+1} = \mathbf{f}_k - \mu \nabla J_{\mathrm{SB}}(\mathbf{f}_k), \qquad k = 0,\, 1,\, \ldots \tag{24}$$

where $\nabla J_{\mathrm{SB}}(\mathbf{f}) = \lambda \nabla J_{\mathrm{MMSE}}(\mathbf{f}) + (1-\lambda)\nabla J_q(\mathbf{f})$. Criterion (18) still remains valid for checking convergence. Due to the relationship between the CP and the MMSE cost functions, gradient ∇J_{MMSE} can readily be computed by setting $q = 1$ and substituting s_n^{t} for d_n in expression (16). By virtue of the same relationship, the step size which minimizes function J_{SB} along direction $\mathbf{g} = \nabla J_{\mathrm{SB}}(\mathbf{f})$ can be found among the roots of the composite polynomial $\mathbb{R}\mathrm{e}\big(\lambda p_{\mathrm{MMSE}}(\mu) + (1-\lambda) p_{\mathrm{CP}}(\mu)\big)$, where p_{CP} and p_{MMSE} are obtained as in expansion (20) from the appropriate values of q and d_n. Note that for $\lambda = 1$ the above iterative procedure reduces to the well-known least mean squares (LMS) algorithm for supervised MMSE equalization.

VII. Experimental Results

This section reports some computer simulations to evaluate the performance of the CP-based methods elaborated in this paper.

Blind ACPA solutions. The first experiment compares the performance of the closed-form blind equalization methods of Section IV. The methods compared are the direct LS solution without structuring ('LS, no struct'); the structuring method of [22] from the top non-overlapping sections of the LS solution ('LS, top'); *idem*, from the bottom sections ('LS, bottom'); the MRC of the top and bottom parts as explained in Section IV-C ('LS, top+bottom'); *idem*, from the whole basis of solutions ('basis, top+bottom'); and the subspace method of Section IV-B ('basis, subspace'). The performance of the supervised MMSE receiver is also computed as a reference. In the first simulation

set-up, a QPSK signal ($q = 4$) excites a simple AR-1 channel

$$H_1(z) = \frac{1}{1 - 0.5z^{-1}}, \qquad |z| > 0.5 \tag{25}$$

with pole at $\rho = 0.5$, well approximated by an order-50 FIR truncation. ISI is perfectly removed by the equalizer $\mathbf{f}_0 = [1,\, -0.5]^{\mathrm{T}}$, which presents a dominant leading tap. The equalizer minimum length is $L_0 = 2$, but an overestimated length of $L = 5$ is chosen, yielding $P = 4$ possible ZF solutions, which are just delayed versions of each other [as in (6)]. Additive white complex circular Gaussian noise is present at the channel output, with signal-to-noise ratio (SNR) given by $\mathrm{E}\{|r|^2\}/\mathrm{E}\{|v|^2\}$. Blocks of $N_d = 100$ symbol periods are observed, and performance parameters are averaged over ν independent Monte Carlo (MC) runs, with $\nu N_d \geqslant 10^5$. Fig. 1 plots the symbol error rate (SER) obtained by the different analytic methods as a function of the SNR. The performance of direct LS solution makes apparent the need for structuring. Using the bottom part of the LS solution exhibits similarly poor results, with a rather low noise tolerance. By contrast, the other methods present a superior performance, just 2–3 dB above the MMSE bound. Interestingly, taking the top part of the LS solution proves best for moderate SNR values in this scenario. This superiority depends, however, on the equalizer tap configuration, as demonstrated in the next example.

We repeat the above experiment but moving the AR channel pole to $\rho = 2$, and taking a stable causal implementation of the channel transfer function

$$H_2(z) = \frac{1}{1 - 2z^{-1}}, \qquad |z| < 2 \tag{26}$$

by shifting the truncated impulse response. The minimum-length equalizer now becomes $\mathbf{f}_0 = [1,\, -2]^{\mathrm{T}}$, with dominant trailing tap. Fig. 2 shows the closed-form blind equalization results. The performance of the LS-top method considerably degrades, being very similar to that of the LS-bottom method in the previous experiment. The performance of the subspace structuring method remains almost the same as in the simulation of Fig. 1, thus showing its robustness to the relative weights of the equalizer coefficients.

Fig. 3 evaluates the sample size requirements of the closed-form solutions, under the general conditions of the first experiment and SNR $= 15$ dB. Satisfactory equalization from a basis of the solution space is achieved even below the limit imposed by (12) for this simulation example, $N_d \geqslant 71$. The subspace approach provides the most efficient results for short observation windows.

CPA solutions — basins of attraction. The next experiments assess the CP-based iterative methods, both in blind (Section V) and semi-blind (Section VI-C) operation. We observe $N_d = 200$ symbols with SNR $= 10$ dB at the output of channel $H_1(z)$ excited by a BPSK input. The contour lines (in the equalizer parameter space) of the logarithm of the blind CP criterion (2) calculated from the

data are plotted in Fig. 4a. The solid lines display the trajectories of the equalizer taps updated with the CPA (17), from 16 different initial configurations and $\eta = 10^{-5}$ in termination criterion (18). Convergence points are marked with a cross. A step size $\mu = 10^{-2}$ was chosen for fastest convergence without compromising stability. The plot also represents the delay zero and one MMSE solutions $\mathbf{f}_{\mathrm{MMSE},0} = [0.85, -0.38]^{\mathrm{T}}$ and $\mathbf{f}_{\mathrm{MMSE},1} = [0, 0.70]^{\mathrm{T}}$, which provide an output MSE of -8.66 and -4.98 dB, respectively. From most of the initial points the algorithm converges to the desired solutions, close to the optimal-delay MMSE equalizer. However, the algorithm gets sometimes stuck at spurious stable extrema located at $\pm[0, 0.58]$, near the suboptimal MMSE equalizer. The basins of attraction of these undesired equilibria are not negligible, and may have a significant negative impact on equalization performance. The method requires, on average, about half thousand iterations to converge (Table I).

The spurious convergence points of the CPA correspond to the theoretical values obtained in [19, Section III.C] for the CM criterion, $\pm[0, 0.65]$. Indeed, as already pointed out in Section III-B, for $q = 2$ and real-valued source and filters the CM and CP criteria coincide.

Under identical conditions and the same observed data, the tap trajectories for the OS-CPA (Section V-B) are obtained as in Fig. 4b. Not only are undesired solutions avoided, but convergence is notably accelerated relative to the previous case: just over ten iterations suffice (Table I).

Using $N_t = 10$ pilot symbols and a confidence parameter $\lambda = 0.5$, the contour lines of the semi-blind CP criterion (21) follow the shape displayed in Fig. 5a. The introduction of training data alters the CP cost function by emphasizing the minimum near the MMSE solution while naturally vanishing the previously acceptable equilibrium across the origin. The use of the optimal step size still leads to good equalization solutions (Fig. 5b) and, again, remarkably speeds up convergence (Table I).

Non-minimum phase channel. We now evaluate performance on the non-minimum phase channel of [7], given by

$$H_3(z) = (-0.033 + 0.014j) + (0.085 - 0.039j)z^{-1} - (0.232 - 0.136j)z^{-2} + (0.634 - 0.445j)z^{-3} + (0.070 - 0.233j)z^{-4} - (0.027 + 0.071j)z^{-5} - (0.023 + 0.012j)z^{-6}. \quad (27)$$

This order-6 FIR channel can be well equalized with a length-3 FIR filter ($L_0 = 3$), but we choose $L = 5$. From a data block of $N_d = 100$ symbols and using several structuring procedures, the blind closed-form CP methods display the SER performance shown in the dashed lines of Fig. 6. The closed-form solutions are then used to initialize the OS-CPA described in Section V, yielding the solid curves in Fig. 6. The gradient iterations refine the analytical estimates, approaching the MMSE bound.

The performance of the semi-blind CP methods is summarized in Fig. 7, for the same simulation setting with $N_t = 10$ pilot symbols and $\lambda = 0.5$. Depending on the window length employed to calculate the MMSE solution, two MMSE curves are obtained as a reference: using just the pilot sequence, as would occur in a conventional receiver, and using the whole data block (MMSE bound). The benefits of the semi-blind approach are noteworthy. Firstly, the performance of the analytic solutions is considerably enhanced compared to blind operation. Secondly, the semi-blind OS-CPA shows identical performance irrespective of initialization, following quite closely the MMSE bound. The exploitation of 'blind symbols' in addition to the training period improves the conventional receiver, and nearly reaches the MMSE bound while considerably increasing the effective data throughput. Also, the convergence rate is improved relative to the fully-blind case, particularly at low SNR, as depicted in Fig. 8.

Influence of pilot-sequence length. Next, we evaluate the CP criterion performance as a function of the proportion of data block symbols used for training. In the previous scenario ($\lambda = 0.5$), two blind ACPA methods are combined with the MMSE solution to generate respective closed-form estimates: the direct LS solution (without structuring), and the subspace-based structuring procedure from a basis of solutions. The OS-CPA is initialized with the center-tap filter and the analytical subspace-based estimate. Results are displayed in Fig. 9. The subspace-based structuring, as expected, worsens as less data are considered in the blind part of the criterion [cf. eqn. (12)]. The performance and convergence speed of the semi-blind OS-CPA seem independent of initialization, although the subspace approach slightly improves the center-tap initial filter in the blind case (0% of training). Note that the performance of a given conventional receiver with up to 30% of pilot symbols can be attained by operating in semi-blind mode with a shorter training preamble, and hence a higher spectral efficiency. A peak in SER and convergence time is observed around 90% of training symbols (Fig. 10). This outcome could be due to the fact that the few symbols in the blind part of the criterion hinder the convergence to the MMSE solution imposed by the pilot symbols. Nevertheless, both performance indices drop to the MMSE limit when the whole observed block is used for training.

Influence of parameter λ. The performance of the semi-blind methods as a function of confidence parameter λ is illustrated in Figs. 11–12, obtained in the same scenario with $N_t = 10$ pilot symbols. As expected, performance improves as more weight is laid on the known data. For $\lambda = 1$, the blind part of the criterion is dropped altogether and the equalization entirely relies on just a few training symbols, thus the increase in SER up to the conventional MMSE receiver level. Accordingly, this increase is not observed in larger training windows. Over a wide range of λ (roughly in the interval $[0.3,\, 0.9]$), the influence of initialization on the performance and convergence speed of the semi-blind OS-CPA seems unimportant, and for practically any $\lambda \in]0,\, 1[$ the semi-blind iterative methods

improve the conventional equalizer. Fig. 12 also shows that a value of the confidence parameter exists ($\lambda \approx 0.7$) for which the cost-function surface is best adapted to the operation of the optimal step-size gradient-descent algorithm, so that convergence is achieved in the lowest number of iterations. This optimal value of λ will generally depend on the specific system conditions, sample size and SNR.

Comparison with CM criterion. A final experiment makes an illustrative comparison between the CP and CM criteria in semi-blind operation (10% training). A co-channel interferer with the same modulation as the desired signal (QPSK) and a given signal-to-interference ratio (SIR) is added at the output of channel $H_3(z)$. The respective top-structuring analytic solutions are first obtained, and then used as initial points for the optimal-step size iterations. Figs. 13–14 show that, although the ACPA solution is poorer than ACMA's in this particular scenario, the OS-CPA improves its CM counterpart with half the number of iterations.

VIII. Summary and Conclusions

The present work has focused on the CP criterion for blind linear equalization of digital communication channels excited by PSK signals. When exact ZF solutions exist (as in all-pole SISO channels), the global minima can be reached in closed-form. These non-iterative solutions are unaffected by the existence of spurious (non-equalizing) local extrema in the cost-function surface. Through an appropriate transformation the non-linear criterion can be linearized; then the structure of the solution must be restored. The algebraic treatment is similar to ACMA's, but the analytic solutions to the CP criterion (ACPA) do not need to be specialized to handle binary modulations. Obtaining a basis of the solution space allows the design of more robust structure-forcing methods to recover the minimum-length equalizer from the solutions to the linearized problem. In simulations, the proposed subspace-based approach has effectively proven more robust than simpler structuring methods. Algebraically, the subspace method solves a particular instance of the rank-1 tensor combination problem. In simulations, the blind analytic solutions show a restricted tolerance to noise, especially for long equalizers. The key issue limiting performance is probably the SVD-based procedure described in Sections IV-B and VI-B for extracting the equalizer vector from the estimated symmetric tensor.

When the algebraic solution is only an approximation (e.g., when no exact ZF equalizer exist), or when it is too costly to compute, iterative techniques are necessary to seek the global minima of the criterion; an iterative method can also be used to refine a good algebraic guess. An exact line search gradient-descent block algorithm has been proposed in which the optimal step size is computed analytically at each iteration. This algorithm (OS-CPA) shows a very fast convergence and is able to avoid spurious local extrema.

The CP criterion is easily modified to include training information. Indeed, the conventional

supervised MMSE principle can be seen as a special case of CP equalization. With just a few pilot symbols, the analytic solutions' noise tolerance is ameliorated. The semi-blind OS-CPA performs near the MMSE bound at a fraction of the bandwidth cost and is very robust to the equalizer-filter initialization.

In short, the CP criterion has been endowed with a number of strategies aiming to reduce the impact of spurious local minima and slow convergence in iterative blind equalizers:

1) informed initialization with closed-form solutions,
2) block iterative operation,
3) global line minimization with algebraically-computed optimal step size, and
4) incorporation of training data.

These strategies are not exclusive to the CP principle but can also benefit other equalization criteria.

Further lines of inquiry could include the theoretical study of spurious extrema in the CP criterion; the improvement of the SVD-based procedure for recovering the equalizer vector; the robust automatic detection of the number of ZF solutions and extraction of the optimum-delay equalizer [31]; the optimal choice of pilot-confidence parameter λ (e.g., based on an asymptotic analysis of variance); the impact of carrier residual [23], [29]; and a thorough theoretical and experimental comparison of the CP principle with other equalization schemes such as the CM criterion.

APPENDIX A: PROOFS

Section III-A:

• Number of input states for PSK modulations.

Let $[t_0, \ldots, t_{K-1}]^{\mathrm{T}}$ be the cascaded equalizer-channel impulse response vector, assuming a total length of K taps. The number of different equations $y_n^q = d_n$ is determined by the possible states of vector $[s_n, \ldots, s_{n-K+1}]^{\mathrm{T}}$. Such K-uplets can take q^K different states. However, we have that $y_n^q = (\sum_{k=0}^{K-1} t_k s_{n-k})^q = \big(s_n(t_0 + \sum_{k=1}^{K-1} t_k \tilde{s}_{n-k})\big)^q = (t_0 + \sum_{k=1}^{K-1} t_k \tilde{s}_{n-k})^q$, where $\tilde{s}_{n-k} = s_n^{-1} s_{n-k}$. By virtue of the group property of the PSK alphabet, $\tilde{s}_{n-k}$ also belongs to the input signal constellation. Hence, the number of equations is actually determined by the different values of the $(K-1)$-uplet $[\tilde{s}_{n-1}, \ldots, \tilde{s}_{n-K+1}]^{\mathrm{T}}$, which can take at most q^{K-1} states. □

Section IV-A:

• $(\mathbf{f}^{\mathrm{H}}\mathbf{x}_n)^q = \mathbf{f}^{\oslash q\,\mathrm{H}}\mathbf{x}_n^{\oslash q}$.

$(\mathbf{f}^{\mathrm{H}}\mathbf{x}_n)^q = \big(\sum_i (\mathbf{f})_i^* (\mathbf{x}_n)_i\big)^q = \sum_{i_1,\ldots i_q} (\mathbf{f})_{i_1}^* \ldots (\mathbf{f})_{i_q}^* (\mathbf{x}_n)_{i_1} \ldots (\mathbf{x}_n)_{i_q}$ can be expressed as the sum of all terms of tensor $\mathbf{f}^{\otimes q^*} \odot \mathbf{x}_n^{\otimes q}$ or, equivalently, of vector $\mathbf{f}^{\oslash q^*} \odot \mathbf{x}_n^{\oslash q}$. This sum is the same as $\mathbf{f}^{\oslash q\,\mathrm{H}}\mathbf{x}_n^{\oslash q}$. □

• Problem (5) equivalent to problem (4).

We need to prove that the set of solutions $\{\mathbf{w}_k\}_{k=1}^P$ of the form $\mathbf{w}_k = \mathbf{f}_k^{\oslash q}$ is linearly independent if and only if (iff) the set $\{\mathbf{f}_k\}_{k=1}^P$ is linearly independent. This can be done along the lines of [8, Proof of Lemma 3] by considering the matrix unfolding of $\mathbf{f}_k^{\otimes q} = \mathbf{unvecs}_q\{\mathbf{f}_k^{\oslash q}\}$, $\mathbf{F}_k \in \mathbb{C}^{L \times L^{q-1}}$, defined as $(\mathbf{F}_k)_{i_1, i_2+L(i_3-1)+\ldots L^{q-2}(i_q-1)} = (\mathbf{f}_k^{\otimes q})_{i_1 i_2 i_3 \ldots i_q}$. This matrix can be expressed as the rank-one product $\mathbf{F}_k = \mathbf{f}_k \bar{\mathbf{f}}_k^{\mathrm{T}}$, with $(\bar{\mathbf{f}}_k)_{i_2+L(i_3-1)+\ldots L^{q-2}(i_q-1)} = (\mathbf{f}_k)_{i_2}(\mathbf{f}_k)_{i_3} \ldots (\mathbf{f}_k)_{i_q}$. Now, vectors $\{\mathbf{f}_k^{\oslash q}\}$ are linearly independent iff $\sum_{k=1}^P \alpha_k \mathbf{f}_k^{\oslash q}$ implies $\alpha_k = 0$, $k = 1, \ldots, P$. That linear combination vanishes iff $\sum_{k=1}^P \alpha_k \mathbf{f}_k^{\otimes q}$ is the null tensor or, equivalently, $\sum_{k=1}^P \alpha_k \mathbf{F}_k$ the zero matrix. Due to the structure of matrices $\{\mathbf{F}_k\}$, this latter condition necessarily implies that $\{\alpha_k\}$ be zero iff rank($[\mathbf{f}_1, \ldots \mathbf{f}_P]$) $= P$, i.e., $\{\mathbf{f}_k\}$ form a linearly independent set. □

• Problem (8) equivalent to problem (9) with scale constraint (10)–(11).

We only need to show that $\mathbf{r}^{\mathrm{H}}\mathbf{w} = \sqrt{N}$ and (10)–(11) are equivalent. Vector $\mathbf{r}^{\mathrm{H}}$ is given by the product of the first row of $\mathbf{Q}$, say $\mathbf{q}^{\mathrm{H}}$, and matrix $\mathbf{X}^{q\mathrm{H}}$. Since $\mathbf{Q}\mathbf{d} = [\sqrt{N},\, \mathbf{0}_{N-1}^{\mathrm{T}}]^{\mathrm{T}}$, it follows that the rest of the rows of $\mathbf{Q}$ are orthogonal to vector $\mathbf{d}$. Also, $\mathbf{Q}$ is unitary, so that $\mathbf{q}$ must be parallel to $\mathbf{d}$; specifically, $\mathbf{q} = \sqrt{N}\mathbf{d}/\|\mathbf{d}\|^2$. Then, $\mathbf{q}^{\mathrm{H}}\mathbf{X}^{q\mathrm{H}} = \frac{\sqrt{N}}{\|\mathbf{d}\|^2} \sum_{n=0}^{N-1} d_n^* \mathbf{x}_n^{\oslash q\mathrm{H}}$. The scale constraint becomes $\frac{\mathbf{w}^{\mathrm{H}}}{\|\mathbf{d}\|^2} \sum_{n=0}^{N-1} d_n \mathbf{x}_n^{\oslash q} = 1$, which reduces to $\frac{1}{\|\mathbf{d}\|^2} \sum_{n=0}^{N-1} d_n (\mathbf{f}^{\mathrm{H}}\mathbf{x}_n)^q = 1$ when $\mathbf{w} = \mathbf{f}^{\oslash q}$. □

• If $\dim\ker(\mathbf{X}^{q\mathrm{H}}) = (P-1)$ and $N > (L_q - P)$ $\quad\Rightarrow\quad$ $\dim\ker(\mathbf{R}) = P$.

Since $\dim\ker(\mathbf{R}) = (L_q - \mathrm{rank}(\mathbf{R}))$ and $\mathrm{rank}(\mathbf{R}) \leqslant (N-1)$, it follows that $\dim\ker(\mathbf{R}) \geqslant (L_q - N + 1)$. Hence, a necessary condition for $\dim\ker(\mathbf{R}) = P$ is that $N > (L_q - P)$ or, in terms of the observed sample size, $N_d \geqslant (L_q + L_0 - 1)$. On the other hand, the number of linearly independent constraints that $\mathbf{X}^{q\mathrm{H}}$ can introduce is equal to $\mathrm{rank}(\mathbf{X}^{q\mathrm{H}}) = (L_q - P + 1)$. Due to the unitarity of $\mathbf{Q}$, such constraints are equivalent to those introduced by $\mathbf{Q}\mathbf{X}^{q\mathrm{H}}$. Since $\mathbf{r}^{\mathrm{H}}\mathbf{w}$ imposes one such constraints, $\mathbf{R}\mathbf{w}$ must impose the remaining $(L_q - P)$. This is precisely the rank of $\mathbf{R}$, and proves that $\dim\ker(\mathbf{R}) = P$. □

Section V-B:

• Optimal step-size polynomial.

$J_q(\mathbf{f} - \mu\mathbf{g}) = \mathrm{E}\{|\varepsilon_n|^2\}$, with $\varepsilon_n = \xi(\mu) - d_n$, where $\xi(\mu) = \big((\mathbf{f} - \mu\mathbf{g})^{\mathrm{H}}\mathbf{x}_n\big)^q$. This latter polynomial in μ can be expanded as $\xi(\mu) = \sum_{p=0}^q a_p \mu^p$, where $a_p = (-1)^p \binom{q}{p} (\mathbf{g}^{\mathrm{H}}\mathbf{x}_n)^p (\mathbf{f}^{\mathrm{H}}\mathbf{x}_n)^{q-p}$. Since $\partial J_q(\mathbf{f} - \mu\mathbf{g})/\partial\mu = 2\mathrm{E}\{\mathbb{R}\mathrm{e}\big(\xi'^* \varepsilon_n\big)\} \stackrel{\triangle}{=} 2p(\mu)$, the first-order necessary condition $\partial J_q/\partial\mu = 0$ reduces to finding the zeros of $p(\mu)$. It remains to prove that such polynomial accepts the expansion of eqn. (20). Now, $p(\mu) = \mathrm{E}\{\mathbb{R}\mathrm{e}\big(\xi'^*(\xi - d_n)\big)\} = \mathbb{R}\mathrm{e}\big(\mathrm{E}\{\xi_1 - \xi_2\}\big)$, with $\xi_1(\mu) = \xi'^*\xi$ and $\xi_2(\mu) = \xi'^* d_n$. As $\xi'(\mu) = \sum_{p=1}^q p a_p \mu^{p-1}$, the coefficients of $(2q-1)$th-degree polynomial $\xi_1(\mu)$ are given by the

convolution $[qa_q^*,\ (q-1)a_{q-1}^*,\ \ldots,\ a_1^*] * [a_q,\ a_{q-1},\ \ldots,\ a_0]$, which produces

$$b_m^{(1)} = \begin{cases} \sum_{p=0}^{m}(m+1-p)a_{m+1-p}^* a_p, & 0 \leqslant m \leqslant q-1 \\ \sum_{p=m+1-q}^{q}(m+1-p)a_{m+1-p}^* a_p, & q \leqslant m \leqslant 2q-1. \end{cases} \tag{28}$$

Similarly, the coefficients of qth-degree polynomial $\xi_2(\mu)$ are simply $b_m^{(2)} = (m+1)a_{m+1}^* d_n$, $0 \leqslant m \leqslant (q-1)$. The combination of these two sets of coefficients and the expectation operator leads to expansion (20). □

ACKNOWLEDGEMENTS

V. Zarzoso carried out part of this work while on leave at the Laboratoire I3S. He thanks Dr. Comon for his kind hospitality and interesting discussions.

REFERENCES

[1] J. K. Tugnait, L. Tong, and Z. Ding, "Single-user channel estimation and equalization," *IEEE Signal Processing Magazine*, vol. 17, no. 3, pp. 16–28, May 2000.

[2] A.-J. van der Veen, S. Talwar, and A. Paulraj, "A subspace approach to blind space-time signal processing for wireless communication systems," *IEEE Transactions on Signal Processing*, vol. 45, no. 1, pp. 173–190, Jan. 1997.

[3] Y. Sato, "A method of self-recovering equalization for multi-level amplitude modulation," *IEEE Transactions on Communications*, vol. 23, pp. 679–682, June 1975.

[4] D. N. Godard, "Self-recovering equalization and carrier tracking in two-dimensional data communication systems," *IEEE Transactions on Communications*, vol. 28, no. 11, pp. 1867–1875, Nov. 1980.

[5] J. R. Treichler and B. G. Agee, "A new approach to multipath correction of constant modulus signals," *IEEE Transactions on Acoustics, Speech and Signal Processing*, vol. 31, no. 2, pp. 459–472, Apr. 1983.

[6] Z. Ding, C. R. Johnson, and R. A. Kennedy, "On the (non)existence of undesirable equilibria of Godard blind equalizers," *IEEE Transactions on Signal Processing*, vol. 40, no. 10, pp. 2425–2432, Oct. 1992.

[7] K. Doğançay and R. A. Kennedy, "Least squares approach to blind channel equalization," *IEEE Transactions on Signal Processing*, vol. 47, no. 11, pp. 1678–1687, Nov. 1999.

[8] A.-J. van der Veen and A. Paulraj, "An analytical constant modulus algorithm," *IEEE Transactions on Signal Processing*, vol. 44, no. 5, pp. 1136–1155, May 1996.

[9] A.-J. van der Veen, "Analytical method for blind binary signal separation," *IEEE Transactions on Signal Processing*, vol. 45, no. 4, pp. 1078–1082, Apr. 1997.

[10] D. T. M. Slock, "Blind fractionally-spaced equalization, perfect-reconstruction filter banks and multichannel linear prediction," in *Proc. ICASSP-94, 19th International Conference on Acoustics, Speech and Signal Processing*, vol. IV, Adelaide, Australia, Apr. 19–22, 1994, pp. 585–588.

[11] L. Tong, G. Xu, and T. Kailath, "Blind identification and equalization based on second-order statistics: a time domain approach," *IEEE Transactions on Information Theory*, vol. 40, no. 2, pp. 340–349, Mar. 1994.

[12] E. Moulines, P. Duhamel, J.-F. Cardoso, and S. Mayrargue, "Subspace methods for the blind identification of multichannel FIR filters," *IEEE Transactions on Signal Processing*, vol. 43, no. 2, pp. 516–525, Feb. 1995.

[13] P. Comon, "Tensor decompositions: state of the art and applications," in *Mathematics in Signal Processing V*, J. G. McWhirter and I. K. Proudler, Eds. Oxford, UK: Clarendon Press, 2002, pp. 1–24.

[14] O. Grellier and P. Comon, "Blind equalization and source separation with MSK inputs," *Proc. SPIE Conference on Advances in Signal Processing.*, pp. 26–34, July 19–24, 1998.

[15] P. Comon, "Contrasts, independent component analysis, and blind deconvolution," *International Journal of Adaptive Control and Signal Processing (Special Issue on Blind Signal Separation)*, vol. 18, no. 3, pp. 225–243, Apr. 2004.

[16] L. Rota and P. Comon, "Blind equalizers based on polynomial criteria," in *Proc. ICASSP-2004, 29th International Conference on Acoustics, Speech and Signal Processing*, Montreal, Canada, May 17–21, 2004, accepted.

[17] O. Grellier and P. Comon, "Blind separation of discrete sources," *IEEE Signal Processing Letters*, vol. 5, no. 8, pp. 212–214, Aug. 1998.

[18] P. Comon, "Block methods for channel identification and source separation," in *Proc. IEEE Symposium on Adaptive Systems for Signal Processing, Communications and Control*, Alberta, Canada, Oct. 1–4, 2000, pp. 87–92.

[19] Z. Ding, R. A. Kennedy, B. D. O. Anderson, and C. R. Johnson, "Ill-convergence of Godard blind equalizers in data communication systems," *IEEE Transactions on Communications*, vol. 39, no. 9, pp. 1313–1327, Sept. 1991.

[20] G. H. Golub and C. F. Van Loan, *Matrix Computations*, 3rd ed. Baltimore, MD: The John Hopkins University Press, 1996.

[21] P. Comon and B. Mourrain, "Decomposition of quantics in sums of powers of linear forms," *Signal Processing (Special Issue on Higher-Order Statistics)*, vol. 53, no. 2, pp. 93–107, Sept. 1996.

[22] O. Grellier and P. Comon, "Closed-form equalization," *Proc. SPAWC-99, 2nd IEEE Workshop on Signal Processing Advances in Wireless Communications*, pp. 219–222, May 9–12, 1999.

[23] P. Comon, "Blind equalization with discrete inputs in the presence of carrier residual," in *Proc. 2nd IEEE International Symposium on Signal Processing and Information Theory*, Marrakech, Morocco, Dec. 2002.

[24] K. Doğançay, K. Abed-Meraim, and Y. Hua, "Convex optimization for blind equalization," in *Proc. ICOTA-98, 4th International Conference on Optimization: Techniques and Applications*, Perth, Australia, July 3-5, 1998, pp. 1017–1023.

[25] J. G. Proakis, *Digital Communications*, 4th ed. New York: McGraw-Hill, 2000.

[26] P. A. Regalia, "A finite-interval constant modulus algorithm," in *Proc. ICASSP-2002, 27th International Conference on Acoustics, Speech and Signal Processing*, vol. III, Orlando, FL, May 13–17, 2002, pp. 2285–2288.

[27] C. R. Johnson, P. Schniter, I. Fijalkow, L. Tong, *et al.*, "The core of FSE-CMA behavior theory," in *Unsupervised Adaptive Filtering, Vol. II: Blind Deconvolution*, S. S. Haykin, Ed. New York: John Wiley & Sons, 2000, ch. 2, pp. 13–112.

[28] W. H. Press, S. A. Teukolsky, W. T. Vetterling, and B. P. Flannery, *Numerical Recipes in C. The Art of Scientific Computing*, 2nd ed. Cambridge, UK: Cambridge University Press, 1992.

[29] P. Comon, "Independent component analysis, contrasts, and convolutive mixtures," in *Proc. 2nd IMA International Conference on Mathematics in Communications*, University of Lancaster, UK, Dec. 16–18, 2002.

[30] V. Zarzoso and P. Comon, "Optimal step-size constant modulus algorithm," *IEEE Communications Letters*, July 2004, submitted.

[31] V. Zarzoso and A. K. Nandi, "Blind MIMO equalization with optimum delay using independent component analysis," *International Journal of Adaptive Control and Signal Processing (Special Issue on Blind Signal Separation)*, vol. 18, no. 3, pp. 245–263, Apr. 2004.

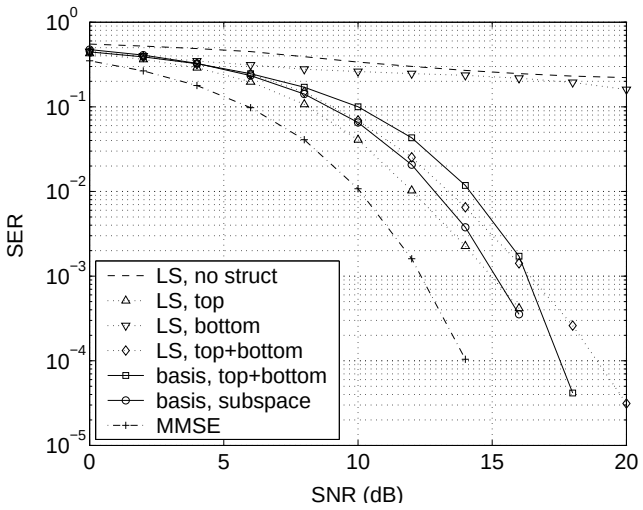

Fig. 1. Closed-form blind equalization based on the CP criterion for several structuring methods. Channel $H_1(z)$, QPSK input ($q = 4$), $N_d = 100$ symbol periods, 1000 MC runs.

Step size	Blind	Semi-blind
constant	422	363
optimal	11	9

TABLE I

AVERAGE NUMBER OF ITERATIONS FOR CONVERGENCE IN THE EXPERIMENTS OF FIGS. 4–5.

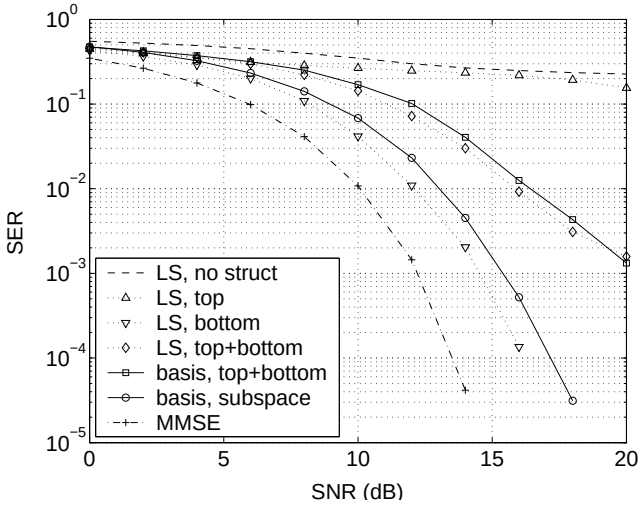

Fig. 2. Closed-form blind equalization based on the CP criterion for several structuring methods. Channel $H_2(z)$, QPSK input ($q = 4$), $N_d = 100$ symbol periods, 1000 MC runs.

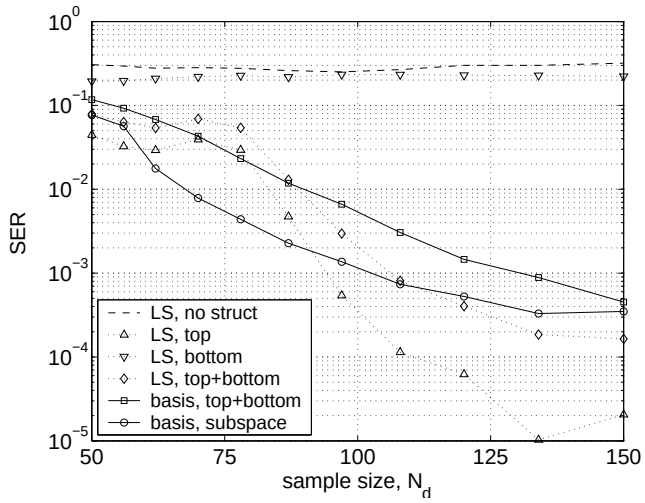

Fig. 3. Closed-form blind equalization based on the CP criterion for several structuring methods. Channel $H_1(z)$, QPSK input ($q = 4$), SNR = 15 dB, ν MC runs, with $\nu N_d \geqslant 10^5$.

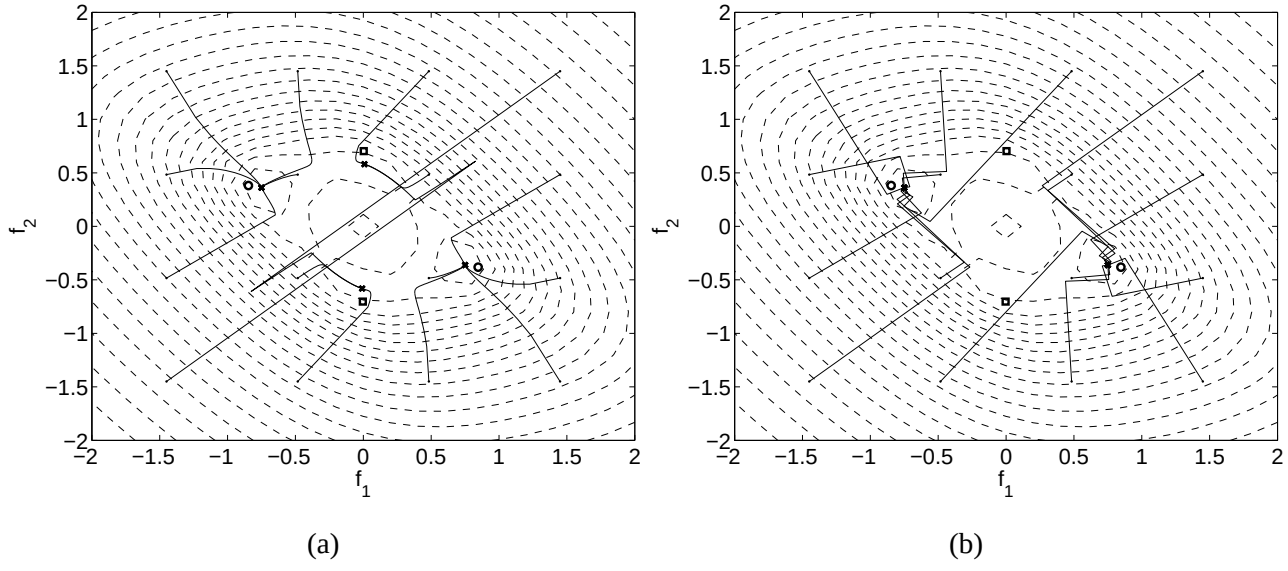

Fig. 4. Blind CP cost function contour lines (dashed) and CPA equalizer tap trajectories (solid lines): (a) constant step size, (b) optimal step size. Channel $H_1(z)$, BPSK input ($q = 2$), $N_d = 200$ symbol periods, SNR = 10 dB. '$\bullet$': initial point; '$\times$': final point; '$\circ$': optimal-delay MMSE solution; '$\square$': suboptimal-delay MMSE solution.

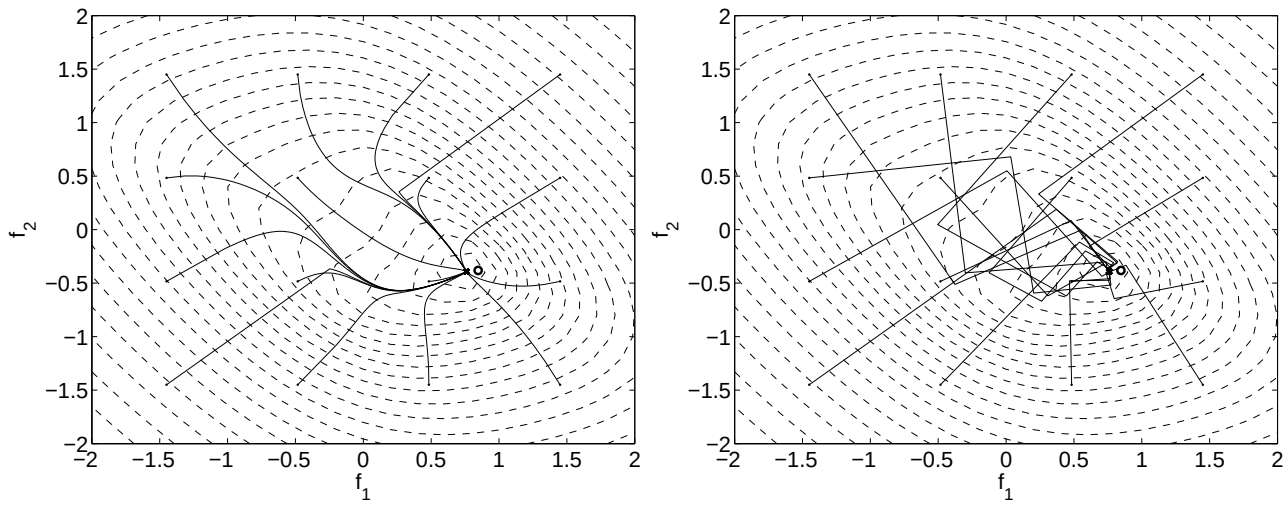

Fig. 5. Semi-blind CP cost function contour lines (dashed) and OS-CPA equalizer tap trajectories (solid line): (a) constant step size, (b) optimal step size. Channel $H_1(z)$, BPSK input ($q = 2$), $N_d = 200$ symbol periods, $N_t = 10$ pilot symbols, SNR = 10 dB, $\lambda = 0.5$. '$\bullet$': initial point; '$\times$': final point; '$\circ$': optimum-delay MMSE solution.

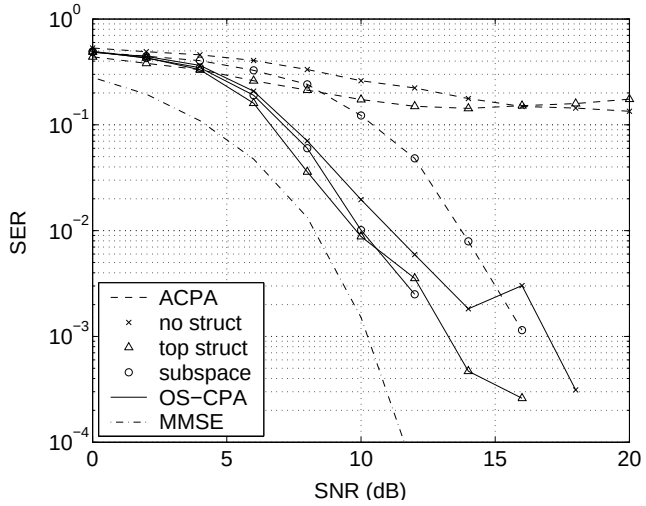

Fig. 6. Blind CP equalization. The OS-CPA is initialized with the corresponding ACPA solution. Channel $H_3(z)$, QPSK input ($q = 4$), $N_d = 100$ symbol periods, 200 MC runs.

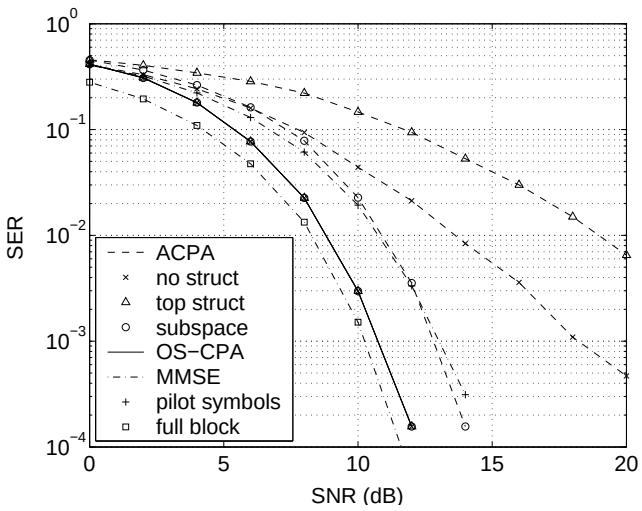

Fig. 7. Semi-blind CP equalization in the simulation of Fig. 6 with $N_t = 10$ pilot symbols and $\lambda = 0.5$. The OS-CPA is initialized with the corresponding ACPA solution.

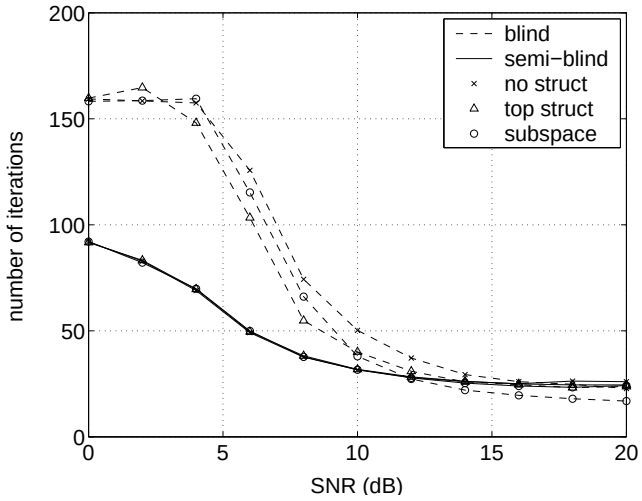

Fig. 8. Average number of iterations for the three initializations of the OS-CPA in the experiments of Figs. 6–7.

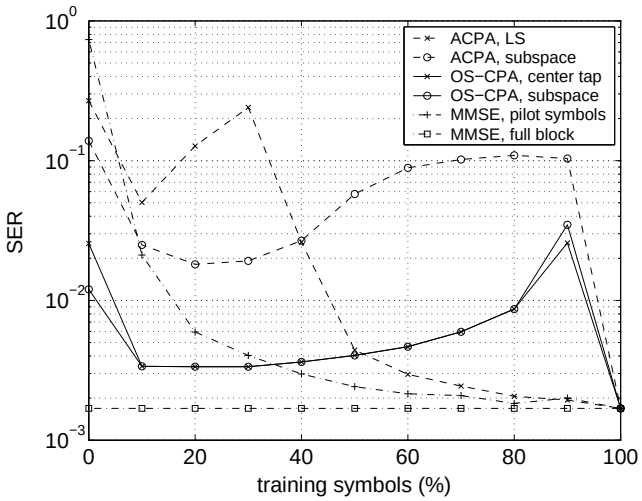

Fig. 9. Impact of the training window length on the performance of the semi-blind CP methods. Channel $H_3(z)$, QPSK input ($q = 4$), $N_d = 100$ symbol periods, SNR = 10 dB , $\lambda = 0.5$, 500 MC runs.

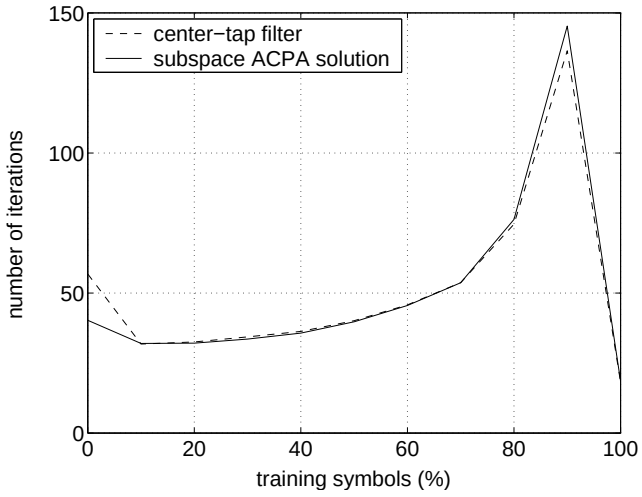

Fig. 10. Average number of iterations for the two initializations of the OS-CPA in the experiment of Fig. 9.

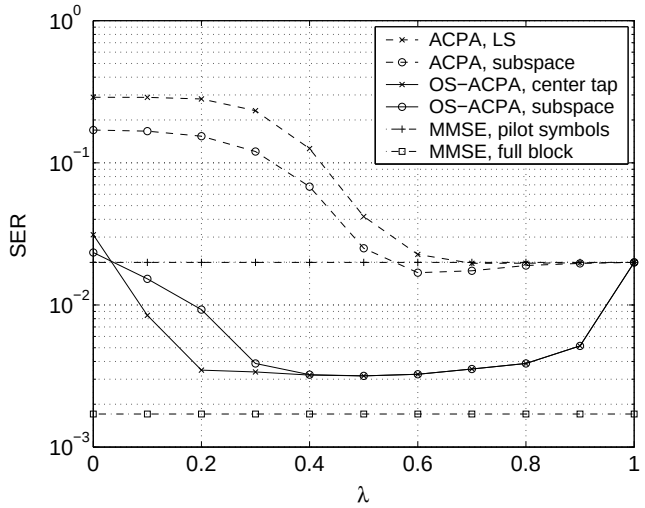

Fig. 11. Impact of confidence parameter λ on the performance of the semi-blind CP methods. Channel $H_3(z)$, QPSK input ($q = 4$), $N_d = 100$ symbol periods, $N_t = 10$ pilot symbols, SNR = 10 dB, 500 MC runs.

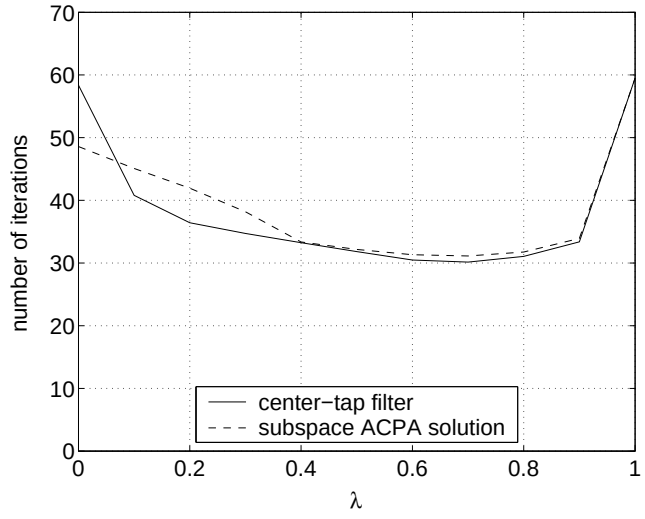

Fig. 12. Average number of iterations for the two initializations of the OS-CPA in the experiment of Fig. 11.

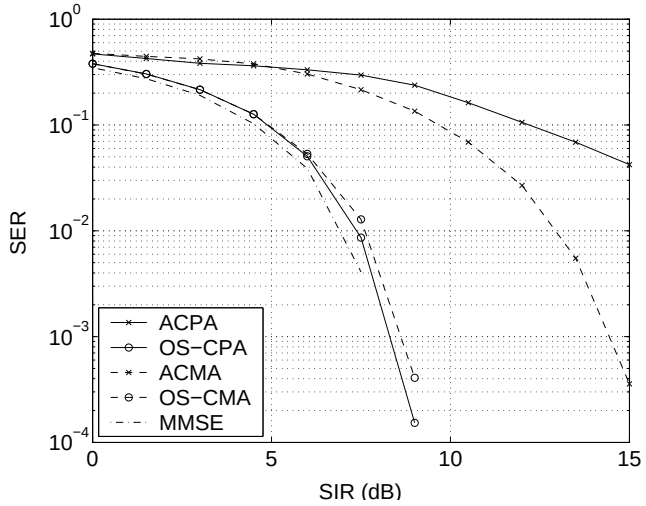

Fig. 13. Semi-blind equalization with the CP and CM criteria. The analytic solutions are obtained using the top structuring method. Channel $H_3(z)$, QPSK input ($q = 4$), QPSK co-channel interferer, $N_d = 200$ symbol periods, $N_t = 20$ pilot symbols, 100 MC runs.

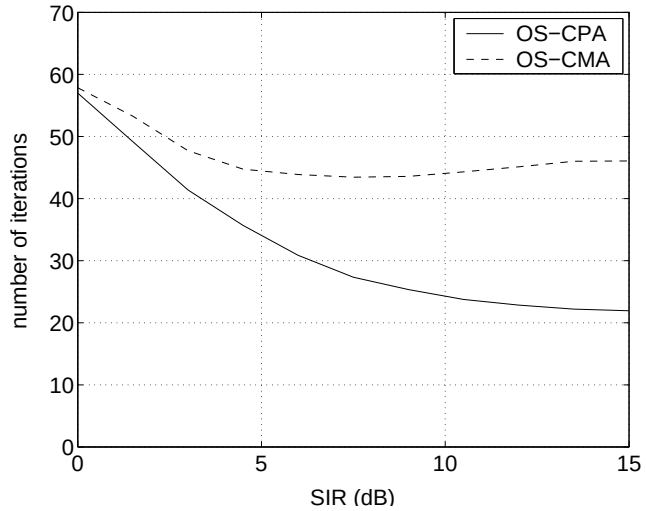

Fig. 14. Average number of iterations in the experiment of Fig. 13.