

Semi-blind methods for communications

15

V. Zarzoso, P. Comon, and D. Slock

15.1 INTRODUCTION

15.1.1 Blind source separation and channel equalization

The problem of BSS arises in a wide variety of real-life applications, which helps explain the intense interest this research area has attracted over the last years. A typical example is encountered in multi-user communication systems, where several mobile users sharing the transmission medium over the same time-frequency-code slot cause *co-channel interference (CCI)* to each other. CCI is due to signals from different spatial origins interfering with the signal of interest, giving rise to *spatial mixtures* observed at the receiving end. Hence, CCI cancellation can naturally be set out as a problem of blind source separation (BSS) in instantaneous linear mixtures, where the signal transmitted by each co-channel user represents one of the sources. These transmission scenarios are also known as instantaneous or static *multi-input multi-output (MIMO)* channels.

In digital communication systems, transmission effects such as multipath propagation and limited bandwidth produce linear distortion over the transmitted signals, causing *intersymbol interference (ISI)* at the receiver output, even if the channel is excited by a single input. Such distortions become more significant as the transmission rate and the user mobility (in the context of wireless communications) increase. ISI arises when a transmitted signal gets corrupted by time-delayed versions of itself, thus generating at the receiver end what could be described as *temporal mixtures*. The problem of ISI suppression is referred to as *deconvolution* or *channel equalization*. Classically, equalization is based on *training* or *pilot symbols* known by the receiver, which leads to a reduced bandwidth utilization. *Blind* equalization methods spare the use of training information [31,65,69,78], with the subsequent gain in bandwidth efficiency. If the transmitted symbols are temporally statistically independent (e.g., iid sequences), channel equalization can be formulated as a BSS problem of independent sources in instantaneous linear mixtures, that is, it accepts an ICA model [7,89,99]. Note that this model still holds if each symbol sequence is a linear process instead of an iid process. In the ICA formulation, the mixing matrix exhibits a Toeplitz structure fully characterized by the channel impulse response. In time-dispersive multi-input channels typically associated with high-data rate multi-user wireless communication systems, both ISI and CCI need

594 CHAPTER 15 Semi-blind methods for communications

to be tackled simultaneously, which calls for *spatio-temporal equalization* techniques or, in the BSS/ICA jargon, for BSS methods in convolutive mixtures. These transmission scenarios are also referred to as convolutive or dynamic MIMO systems.

On account of the above connections, generic methods for BSS used in other applications could also be employed to perform digital channel equalization. However, digital communication channels present particular features that can be capitalized on to improve the source recovery. Firstly, digital modulations have finite support or, in other words, they contain only a small number of possible complex amplitudes. Criteria such as the *constant modulus (CM)* or the *constant power (CP)* are specifically adapted to the blind estimation of signals with such modulations and, as shown in Chapter 3, constitute valid contrasts for the separation and extraction of these signals in linear mixtures, either instantaneous or convolutive. The CM has long been used in blind equalization [31,65,78], whereas the CP criterion has been recently proposed for inputs with *q-ary phase shift keying (q-PSK)* modulation, for an arbitrary integer $q \geq 2$ [14]. These principles can be considered as *quasi-deterministic* rather than statistical criteria, in the sense that signals with adapted modulations cancel exactly (in the absence of noise) the sample version of the contrasts for any data length. As a result, these contrasts offer the potential of achieving good performance even for short sample size. A related benefit is that constellation-adapted criteria spare the input statistical independence assumption [14]. Secondly, training symbols known by the receiver can be incorporated into the transmitted signal to assist the equalization process, thus giving rise to *semi-blind methods*. By appropriately combining pilot information with blind criteria, semi-blind methods can outperform traditional training-based techniques at a fraction of the bandwidth utilization and with just a moderate increase in computational cost.

15.1.2 Goals and organization of the chapter

The present chapter studies a number of strategies for semi-blind equalization of digital communication channels. Only direct equalization, i.e., without previous channel identification, is addressed. For completeness, semi-blind channel estimation will also be treated, though briefly, at the end of the chapter. Our focus is on digital-modulation based contrasts like the CM and CP and, more particularly, semi-blind criteria that can be derived from them. In addition, we develop two other strategies aiming at improving the deficiencies of blind equalizers, which can also be employed in conjunction with semi-blind criteria: equalizer initialization by means of algebraic solutions and iterative search based on optimal step size computation. Algebraic methods are associated with challenging matrix and tensor decomposition problems analogous to those found in BSS based on statistical independence (ICA problem). Efficient iterative equalization techniques can be developed by the optimal step size approach presented at the end of Chapter 6. As shown there for the kurtosis contrast, the CM and CP criteria also admit algebraic solutions for the step size *globally* optimizing the contrast function along the search direction at each iteration. As demonstrated throughout the chapter, the combination of these three strategies, namely semi-blind contrasts, algebraic initialization and

15.2 Training-based and blind equalization 595

optimal step-size iterative search, leads to equalizers with increased robustness, high convergence speed and modest complexity. For reasons of space, our attention is restrained to the basic *single-input single-output (SISO)* systems characterized by a single user without space-time diversity. However, these results are readily extended to the multi-channel case, including *single-input multi-output (SIMO)* systems [95] and MIMO systems typical of multi-user environments [96–98], more directly related to the BSS problem.

We begin the exposition by briefly reviewing the basic concepts of training-based and blind equalization in section 15.2; some strategies for improving their limitations are then summarized in section 15.3. The signal model and notational conventions that will be employed throughout the rest of the chapter are presented in section 15.4. Semi-blind equalization criteria are put forward in section 15.5. Algebraic solutions to the corresponding contrast optimization problems are then studied in section 15.6, whereas iterative methods based on algebraic optimal step-size optimization (introduced on the kurtosis contrast in the fully blind case in Chapter 6) are the topic of section 15.7. A thorough experimental study aiming to illustrate the performance of the presented techniques is reported in section 15.8. Finally, section 15.9 comments on the related problem of semi-blind channel identification. For the sake of clarity, the reader is referred to references [92–95] for details, proofs and other mathematical derivations.

15.2 TRAINING-BASED AND BLIND EQUALIZATION

15.2.1 Training-based or supervised equalization

Traditional equalization techniques are based on a sequence of symbols known by the receiver, the so-called pilot or training sequence, incorporated to the transmitted signal frame. The supervised equalizer is simply obtained by *optimal Wiener filtering* of the received signal using the pilot sequence as desired output. In this context, Wiener filters are also known as *minimum mean square error (MMSE)* equalizers and, in practical settings involving finite data, are usually obtained as the solution to the associated *least squares (LS)* problem. Due to their simplicity and robustness, supervised methods are employed by most of the current wireless communication systems; for instance, the second-generation GSM standard dedicates 20% of each burst to training [73]. The periodic transmission of pilot sequences reduces the useful data rate, and proves particularly ineffective in *broadcast, multicast* or non-cooperative environments, where synchronization is difficult [80,84]. Bandwidth utilization can be improved by reducing the pilot sequence length, but the sequence may become too short relative to the channel delay spread and additional information is thus necessary to render the supervised approach more robust in these conditions. The use of additional information in this context can sometimes be considered as a regularization [43].

15.2.2 Blind equalization

In the late 1970s, the drawbacks of training-based methods spurred a rising interest in blind techniques. The papers by Sato [65], Godard [31] and Treichler [78] are the

596 CHAPTER 15 Semi-blind methods for communications

pioneering contributions to the blind approach. By sparing the pilot sequence, blind equalization techniques increase the effective transmission rate and alleviate the need for synchronization.

In the fundamental SISO case, *non minimum phase* systems cannot be identified using circular second-order statistics (SOS) only [44]. It is thus necessary to resort, explicitly or otherwise, to *higher-order statistics (HOS)*, or to non-circular SOS [36]. Essentially, most blind methods aim to restore at the equalizer output a known property of the input signal, such as a modulation alphabet with a finite number of symbols or with constant modulus. Under certain conditions, the use of these properties through the minimization of an appropriate cost function, or the maximization of a contrast function, guarantees the extraction of the signal of interest.

Despite their improved bandwidth utilization and versatility, blind methods present a number of important shortcomings:

- indeterminacy of the amplitude and/or the phase of the equalized signal;
- multi-modality, that is, existence of local extrema in the cost function to be optimized;
- increased computational complexity;
- larger data volume (block size) than supervised techniques required for the same equalization quality;
- in certain situations, slow convergence or tracking of the variations of the system parameters.

The last two drawbacks are mainly due to HOS estimation errors, which are typically more important than those of SOS for the same sample size.

15.2.3 A classical blind criterion: the constant modulus

The CM criterion [78] – a particular member of the more general family of Godard methods [31] – is perhaps the most widespread blind equalization principle, probably due to its simplicity and flexibility. Indeed, the CM criterion is easy to implement in an iterative fashion and can also deal with non-CM modulations at the expense of an increased estimation error due to constellation mismatch. The CP criterion, also studied in this chapter, can be considered as a modification of Godard’s family, with a power parameter adapted to the number of symbols in the constellation; hence the name “constant power”. Although Godard methods are globally convergent in the combined channel-equalizer space, they present suboptimal equilibrium points in the equalizer space [21,22]. Such points correspond to stable local extrema associated with filters unable to open sufficiently the eye diagram at the equalizer output, so that the detection device cannot extract the transmitted symbols with a reasonably low probability of error. Suboptimal equilibria are often called *spurious solutions* in the literature. Yet these solutions are typically very close to Wiener equalizers, which may question the term “spurious”. In any case, the presence of suboptimal attractors renders the performance of gradient-type iterative algorithms based on Godard criteria very dependent on the initial value of the equalizer impulse response.

15.3 OVERCOMING THE LIMITATIONS OF BLIND METHODS

As discussed in [21,22], among other works, the convergence problems of iterative blind SISO equalizers requires *ad hoc* strategies for suitable filter tap *initialization*, and even for maintaining the tap trajectories far from spurious attractor basins. Three of such strategies are algebraic solutions, multi-channel systems and semi-blind approaches.

15.3.1 Algebraic solutions

Algebraic methods (sometimes called analytic) provide an equalization solution in a finite number of operations, and can always be employed as judicious initializations to iterative equalizers. An algebraic CM solution is obtained in [25], where the CM criterion is formulated as a nonlinear *least squares (LS)* problem. Through an appropriate transformation of the equalizer parameter space, the nonlinear system becomes a linear LS problem subject to certain constraints on the solution structure. Recovering of the correct structure is particularly important when multiple *zero forcing (ZF)* solutions exist; for instance, in all-pole channels with over-parameterized *finite impulse response (FIR)* equalizers, several ZF equalization delays are possible. From a matrix algebra perspective, enforcing this structure can be considered as a matrix diagonalization problem, where the resulting matrix is composed of the equalizer vectors. Once a non-structured solution has been obtained via pseudo-inversion, the minimum-length equalizer can be extracted by a subspace-based approach or other simple procedures for structure restoration.

The blind equalization method of [25] has strong connections with the *analytical CM algorithm (ACMA)* of [82] for BSS. ACMA yields, in the noiseless case, exact algebraic solutions for the spatial filters extracting the sources from observed instantaneous linear mixtures. It is interesting to note that the recovery of separating spatial filters from a basis of the solution space is equivalent to the joint diagonalization of the corresponding matrices. This joint diagonalization can be performed by the generalized *Schur decomposition* [32] of several (more than two) matrices, for which a convergence proof has not yet been found. Either for source separation or channel equalization, ACMA requires special modifications to treat signals with one-dimensional alphabets (e.g., binary) [25,81,82]. Such modifications give rise to the *real ACMA (RACMA)* method [81].

Other solutions aiming at estimating algebraically the best SISO equalizer, or to identify the SISO channel, when the input belongs to a known alphabet have been proposed in [1,18,19,28,33,36,42,46,48,76,83,86,90]. The *discrete alphabet* hypothesis is then crucial, and replaces the assumption of statistical independence between symbols [14], which is no longer necessary. The alphabet-based CP criterion also admits algebraic solutions, which, as reviewed in section 15.6, can be considered as a generalization of the algebraic CM solutions. Algebraic CP solutions are linked to challenging tensor decomposition problems. For a q -symbol constellation, the minimum-length equalizer can be determined from the joint decomposition of q th-order tensors, which, in turn, is linked to the rank-1 linear combination problem in the tensor case. To surmount the lack of effective tools for performing this task, approximate solutions can be proposed in the form

598 CHAPTER 15 Semi-blind methods for communications

of a subspace method exploiting the particular structure of the tensors associated with satisfactory equalization solutions. As opposed to [25], the subspace method proposed here takes into account a complete basis of the solution space. The use of this additional information allows one to increase the robustness of the algorithm with respect to the structure of the minimum-length equalizer. Moreover, the proposed blind algebraic solution deals naturally with binary inputs (BPSK, MSK) without any modifications.

15.3.2 Multi-channel systems

Multi-channel implementations, enabled by time oversampling or the use of multiple sensors, can avoid some of the deficiencies of blind SISO equalizers. Indeed, SIMO channels can be identified blindly by using SOS, regardless of their phase (minimal or non-minimal). Moreover, FIR SIMO channels can be perfectly equalized in the absence of noise by FIR filters [55,71,77]. However, the channel must verify strict diversity conditions, and a good number of these methods do not work when the channel length is overestimated [16]. In any case, the indeterminacy problems remarked in section 15.2.2 remain in the blind context.

Similarly, Godard SIMO equalizers do not present suboptimal minima for noiseless channels satisfying certain length and zero conditions [49]. All minima are indeed global, and coincide with MMSE solutions associated with achievable equalization delays. This feature results in iterative blind equalizers with improved performance. In the presence of noise, however, some of these minima become local and the respective equalizers provide different MSE performance [40]. Depending on its performance, a local minimum can also lead to a suboptimal solution. Consequently, the need for strategies to avoid local extrema remains pertinent in the multi-channel context. In some practical scenarios, it is not possible to attain the degree of spatio-temporal diversity required for a SIMO formulation, due to an insufficient *excess bandwidth* or hardware constraints limiting the number of receiving sensors (consider, for instance, the reduced spatial diversity available in a mobile phone). These are among the reasons for which this chapter is mainly focused on the SISO case, even if principles extend beyond that case.

15.3.3 Semi-blind approach

The combination of a training-based and a blind criterion can avoid their respective drawbacks while preserving their advantages. Indeed, it has been shown that any channel (SISO or SIMO) is identifiable from a small number of known symbols. Thanks to the use of a blind criterion, the pilot sequence necessary to estimate a channel of given length can become shorter relative to the training-only solution; spectral efficiency can thus be increased for a fixed estimation quality. As a result, semi-blind techniques often outperform supervised and blind techniques. When they fail, their semi-blind association can succeed [16]. On the one hand, the semi-blind approach can be interpreted as the regularization of the conventional supervised approach, avoiding its performance degradation for insufficient pilot length. On the other hand, the incorporation of a

few pilot symbols “softens” the cost function, suppressing local minima, accelerating convergence and eliminating the indeterminacies of fully blind criteria. These features will be illustrated throughout the chapter.

The performance and robustness of the semi-blind approach justify the interest in this kind of techniques. The fact that many of the current as well as future communication systems include pilot sequences in their definition standards (in particular to assist synchronization) provides a strong additional motivation for semi-blind equalization techniques. Nevertheless, their use in currently available commercial products is rather limited.

In the context of algebraic methods, it has been recalled above that ACMA requires a joint diagonalization stage (a costly QZ iteration) in the general case where multiple solutions exist [82], although its complexity can be alleviated if the different solutions are delayed versions of each other [25]. The *semi-blind ACMA (SB-ACMA)* proposed in [74] avoids the costly joint diagonalization step of its blind version by constraining the spatial filter or *beamformer* to lie in a certain subspace associated with the pilot symbol vector. Nevertheless, the uniqueness of this semi-blind solution as well as its performance in the presence of noise remain to be ascertained in more detail.

15.4 MATHEMATICAL FORMULATION

15.4.1 Signal model

A digital signal $s(t) = \sum_n s_n \delta(t - nT)$ is transmitted at a known symbol rate $1/T$ through a dispersive channel with impulse response $h(t)$. The channel is linear and time invariant (at least over the observation window), and has a stable causal possibly non-minimum phase transfer function. The baseband signal at the receiver output is given by $x(t) = r(t) + v(t)$, where $r(t) = h(t) * s(t)$ denotes the noiseless observation and $v(t)$ an additive noise independent of $s(t)$. Assuming perfect synchronization and carrier residual elimination, symbol-rate sampling produces the discrete-time output:

$$x_n = r_n + v_n = \sum_k h_k s_{n-k} + v_n \quad (15.1)$$

where $x_n = x(nT)$ while h_k , s_n and v_n can be defined similarly. Each observed sample consists of a noisy linear mixture of time-delay versions of the original data, a phenomenon known as ISI. The goal of channel equalization or deconvolution is to recover the original data from the signal corrupted by convolutive channel effects (ISI) and noise. To this end, we seek a baud-rate FIR discrete-time equalizer with coefficients $\mathbf{f} = [f_1, \dots, f_L]^T \in \mathbb{C}^L$. The equalizer vector is sought so that the equalizer output

$$y_n = \mathbf{f}^H \mathbf{x}_n$$

is an accurate estimate of the source symbols s_n , where:

$$\mathbf{x}_n = [x_n, x_{n-1}, \dots, x_{n-L+1}]^T.$$

600 CHAPTER 15 Semi-blind methods for communications

Signal blocks composed of N_d symbol periods are observed at the channel output. These samples can be stored in a Toeplitz matrix:

$$\mathbf{X} = [\mathbf{x}_{L-1}, \mathbf{x}_L, \dots, \mathbf{x}_{N_d-1}] \quad (15.2)$$

with dimensions $L \times N$, where $N = (N_d - L + 1)$. A similar signal model holds if the channel output is sampled at an integer multiple of the symbol rate (fractional sampling), if there exist multiple spatially separated sensors at reception (spatial oversampling), or if several signal sources transmit simultaneously, giving rise to additional CCI (multi-input system).

To enable the semi-blind mode of operation, we further assume that the transmitted block includes a pilot or training sequence composed of N_t symbols, denoted by $\check{\mathbf{s}} = [\check{s}_0, \check{s}_1, \dots, \check{s}_{N_t-1}]^H$. For the sake of simplicity, the training symbols are assumed to appear, without loss of generality, at the beginning of each block. It has been proven that, as far as channel estimation is concerned, the location of the pilot sequence at the beginning of the block is generally suboptimal [16]; this result probably applies as well to the direct equalization problem under study here. Nevertheless, the following results can be easily extended to an arbitrary location of the pilot sequence, including the optimal placement analyzed in [16].

15.4.2 Notations

Scalars, vectors and tensors (of which matrices are considered as particular cases) are denoted by lowercase (a), bold lowercase (\mathbf{a}) and bold uppercase (\mathbf{A}) symbols, respectively, except structures derived from Kronecker tensorial products, as detailed below. As will be explained in section 15.6, tensor structures will be employed in the derivation of algebraic solutions to the CP contrast. \mathbf{I}_n refers to the identity matrix with dimensions $(n \times n)$, whereas $\mathbf{0}_n$ represents the vector with n zeroes; $\|\cdot\|$ is the conventional L^2 norm. $(\mathbf{A})_{i_1 i_2 \dots i_q}$ stands for the (i_1, i_2, \dots, i_q) -element of q th-order tensor \mathbf{A} . $\Re(\cdot)$ et $\Im(\cdot)$ represent the real and imaginary parts, respectively, of their complex argument; $j = \sqrt{-1}$ is the imaginary unit. Symbols \otimes , \odot and \oslash denote, respectively, the Kronecker product, the element-wise product and the outer product. Given a vector $\mathbf{a} \in \mathbb{C}^L$, we define its q th-order tensor product as the following rank-1 tensor: $\mathbf{a}^{\otimes q} = \underbrace{\mathbf{a} \otimes \dots \otimes \mathbf{a}}_q$. For instance, matrix $\mathbf{a}^{\otimes 2} = \mathbf{a} \otimes \mathbf{a}$ can also be written as $\mathbf{a}\mathbf{a}^T$. Note that

on an appropriate basis, tensor $\mathbf{a}^{\otimes q}$ will have vector $\mathbf{a} \otimes \mathbf{a} \otimes \dots \otimes \mathbf{a}$ as coordinates¹; but this representation does not take into account the reduced dimension of the associated space due to symmetries. Indeed, a symmetric tensor of order q and dimension L can be stored in a vector $\mathbf{vecs}\{\mathbf{A}\}$ that contains only the $L_q = \binom{L+q-1}{q}$ different components of \mathbf{A} . Moreover, they can be normalized by the number of times they appear, so as to preserve the Frobenius norm [13]. In particular, we will write $\mathbf{a}^{\otimes q} = \mathbf{vecs}\{\mathbf{a}^{\otimes q}\}$. Similarly,

¹Recall that \otimes denotes the Kronecker product, whereas \oslash denotes the tensor (outer) product.

$\text{unvecs}_q\{\mathbf{b}\}$ denotes the symmetric q th-order tensor made up of the elements of vector \mathbf{b} with dimension L_q .

15.5 CHANNEL EQUALIZATION CRITERIA

15.5.1 Supervised, blind and semi-blind criteria

The supervised MMSE criterion aims at the minimization of the cost function:

$$\Upsilon_{\text{MMSE}}(\mathbf{f}) = \mathbb{E}\{|y_n - \check{s}_{n-\tau}|^2\}. \quad (15.3)$$

Symbol τ represents the equalization delay, on which performance strongly depends, and \check{s} denotes the pilot sequence, as introduced before.

Concerning blind approaches, the CM criterion is defined by:

$$\Upsilon_{\text{CM}}(\mathbf{f}) = \mathbb{E}\{(|y_n|^2 - \gamma)^2\} \quad (15.4)$$

where $\gamma = \mathbb{E}\{|s_n|^4\}/\mathbb{E}\{|s_n|^2\}^2$ is an alphabet-dependent constant.

A widely used criterion is the standardized cumulant [44], whose most elegant introduction is due to Donoho [26]. At fourth order, the standardized cumulant is called *kurtosis* (see also Chapters 3 and 6), and is given by

$$\Upsilon_{\text{KM}}(\mathbf{f}) = \frac{\text{cum}_4\{y_n\}}{\text{cum}_2^2\{y_n\}} \quad (15.5)$$

where $\text{cum}_2\{y_n\}$ is the variance of y_n and $\text{cum}_4\{y_n\} = \text{cum}\{y_n, y_n^*, y_n, y_n^*\}$ can be defined as:

$$\text{cum}_4\{y_n\} = \mathbb{E}\{|y_n|^4\} - 2\mathbb{E}\{|y_n|^2\}^2 - |\mathbb{E}\{y_n^2\}|^2.$$

When the source sequence s_n is white, the channel performs a mixture of independent random variables, so that the observation x_n is more Gaussian than s_n . The maximization of this criterion, or its square modulus when its sign is unknown, renders the equalizer output as non-Gaussian as possible. This *kurtosis maximization (KM) criterion* has widely been used for SISO and MIMO channel equalization, as well as for source separation in instantaneous linear mixtures (static MIMO channels). More generally, it is shown by Proposition 3.11, page 84, that if the r th-order source cumulant $\text{cum}_r\{s_n\}$ is not null, one can maximize the normalized r th-order cumulant of the equalizer output, $\text{cum}_r\{y_n\}/\text{cum}_2^{r/2}\{y_n\}$.

For the application of the blind *constant power (CP)* approach, the transmitted symbols are assumed to belong to a q -PSK digital modulation, represented by the finite alphabet $\mathcal{A}_q = \{a^k\}_{k=0}^{q-1}$, where $a^q = d$ depends on the constellation; for instance, $(q, d) = (2, 1)$ for BPSK and $(q, d) = (4, 1)$ for QPSK sources. In addition, allowing a time-varying d , the above definitions are directly extended to modulations other than PSK, such as

602 CHAPTER 15 Semi-blind methods for communications

MSK [33], which can be described by $(q, d_n) = (2, (-1)^n)$. As $s_n \in \mathcal{A}_q$, it follows that $s_n^q = d_n$. Consequently, a rather natural way to measure the proximity of the equalizer output to the original symbols is through the criterion:

$$\Upsilon_{\text{CP}}(\mathbf{f}) = \mathbb{E}\{|y_n^q - d_n|^2\}. \quad (15.6)$$

This function is a particular case of the more general class of *alphabet polynomial fitting (APF)* criteria, where the equalizer output constellation is matched to that of the source, characterized by the complex roots of a specific polynomial [14,62]. In the context of BSS, the criterion is equivalent, for a sufficiently low noise level, to the *maximum a posteriori (MAP)* principle [10,34]. Moreover, it has been proved in [14] that, when the global channel-equalizer impulse response is of finite length and the input signal is sufficiently exciting, the global minima of the sample average of (15.6) in the combined channel-equalizer space correspond to ZF solutions. Nevertheless, this result does not guarantee that the desired solutions can always be attained. Indeed, spurious extrema can appear when the cost function is observed from the equalizer parameter space, due to the finite equalizer length, as remarked in [21,22] for Godard criteria. The existence of suboptimal extrema in the CP criterion will be illustrated by some simple experiments in section 15.8.

The linear combination of the above cost functions provides in a quite natural fashion the *semi-blind CM-MMSE (SB-CM-MMSE) and CP-MMSE (SB-CP-MMSE) criteria*:

$$\Upsilon_{\text{SB-CM}}(\mathbf{f}) = \lambda \Upsilon_{\text{MMSE}}(\mathbf{f}) + (1 - \lambda) \Upsilon_{\text{CM}}(\mathbf{f}) \quad (15.7)$$

$$\Upsilon_{\text{SB-CP}}(\mathbf{f}) = \lambda \Upsilon_{\text{MMSE}}(\mathbf{f}) + (1 - \lambda) \Upsilon_{\text{CP}}(\mathbf{f}). \quad (15.8)$$

Parameter λ is a real-valued constant in the interval $[0, 1]$. It can be considered as the relative degree of confidence in the blind and pilot-based parts of the criterion.

Remark that, in practice, mathematical expectations are replaced by sample averages over the data available in the observed signal block.

15.5.2 Relationships between equalization criteria

The CP criterion (15.6) bears close resemblance to the Godard class [31], which in the PSK case becomes:

$$\Upsilon_{\text{G}}^{(q,2)}(\mathbf{f}) = \mathbb{E}\{(|y_n|^q - |s_n|^q)^2\} = \mathbb{E}\{(|y_n|^q - |d_n|)^2\}. \quad (15.9)$$

For $q = 2$, this function corresponds to the CM criterion [31,78]. For BPSK sources and real-valued channels and equalizers, the CP and CM criteria are identical; in this case, we anticipate that the algebraic treatment of CP minimization (section 15.6) is also equivalent to that of ACMA for binary modulations (RACMA) [25,81]. The parallelism between the CM and CP cost functions suggests the existence of local extrema for the latter, even in the case $q > 2$.

15.5 Channel equalization criteria 603

The phase insensitivity of the CM criterion is one of its main interests, as it allows the independent operation of the equalization and carrier recovery stages [31,78]. However, for the same reason the carrier residual cannot be detected or identified by using this criterion. By contrast, the CP criterion can incorporate an appropriate carrier residual compensation mechanism into the algorithm or, otherwise, it requires the previous suppression of the residual before applying the algorithm. On the other hand, all PSK constellations being CM, the CM criterion does not make the difference between PSK modulations; similarly, the more general criterion (15.9) cannot privilege a particular PSK modulation. In contrast, criterion (15.6) explicitly takes into account the discrete nature of PSK alphabets, so that it may exhibit better discriminating properties among CM constellations.

If d_n is replaced by the available pilot symbols \check{s}_n , the CP cost function (15.6) reduces, with $q = 1$, to the MMSE supervised equalization principle (15.3). This fact will be exploited when designing semi-blind iterative methods in section 15.7.

We have seen that the KM criterion (15.5) maximizes non-Gaussianity at the equalizer output. This is also what the CM criterion does by forcing the equalizer output to approach the unit circle. To realize this connection, it is interesting to compare the CM and KM criteria. First, note that the KM criterion, as opposed to the CM, is insensitive to scale. Hence, we can write $y_n = \rho \bar{y}_n$, where ρ is a positive scale factor, and also $\mathbf{f} = \rho \bar{\mathbf{f}}$, where $\bar{\mathbf{f}}$ has fixed norm. This normalization is equivalent to fixing the variance of \bar{y}_n . Let us minimize $\Upsilon_{\text{CM}}(\rho \bar{\mathbf{f}})$ with respect to ρ . We obtain $\rho_0^2 = \gamma^2 \mu_2 / \mu_4$, denoting $\mu_r = \mathbb{E}\{\bar{y}_n^r\}$. For this optimal value of ρ , we have:

$$\frac{1}{\gamma^2} \Upsilon_{\text{CM}}(\rho_0 \bar{\mathbf{f}}) = 1 - \frac{\mu_2^2}{\mu_4}.$$

The CM criterion can then be linked to the KM criterion, provided that the source has a distribution with second-order circular symmetry, that is, $\mathbb{E}\{s_n^2\} = 0$. Under this assumption, $\Upsilon_{\text{KM}}(\bar{\mathbf{f}}) + 2 = \mu_4 / \mu_2^2$ and we have the following simple relationship:

$$\frac{1}{\gamma^2} \Upsilon_{\text{CM}}(\rho_0 \bar{\mathbf{f}}) = 1 - \frac{1}{\Upsilon_{\text{KM}}(\bar{\mathbf{f}}) + 2} \quad (15.10)$$

which shows that both criteria have the same stationary points in $\bar{\mathbf{f}}$. Hence, this equivalence, already established in [23, Chapter 4] [59–61], applies to most complex alphabets. However, the equivalence does not hold any more when the channel is complex-valued and the source real-valued (e.g., PAM modulations); indeed, in such a case the observation is no longer second-order circular.

Semi-blind CM-MMSE criterion (15.7) was initially proposed in [43], but using the so-called “CMA 1-2” cost instead of the “CMA 2-2” cost (15.7). The originality was to surmount the deficiencies of the LS solution to (15.3) (see the next section) when the pilot sequence is not long enough, an improvement known by its regularization capabilities. In addition, it has been proven that the incorporation of pilot symbols is

604 CHAPTER 15 Semi-blind methods for communications

capable of reducing the probability of converging towards spurious solutions due to the non-convexity of the CM cost function. The techniques presented in the following sections further reduce the impact of local extrema on equalization performance while accelerating convergence. As far as iterative techniques are concerned (section 15.7), we propose to minimize the hybrid criteria CM-MMSE (15.7) and CP-MMSE (15.8) through an efficient gradient algorithm where the step size is determined algebraically at each iteration by computing exhaustively all roots of a low degree polynomial. As demonstrated in the numerical experiments of section 15.8, this optimal step size accelerates convergence and makes performance attain the MMSE bound, even from a reduced number of pilot symbols. In addition, these optimal step-size iterative techniques can be judiciously initialized with the aid of the algebraic solutions presented next.

15.6 ALGEBRAIC EQUALIZERS

Perfect ZF equalization of a SISO channel is possible when both of the following conditions hold:

- C1. The channel admits a noiseless M th-order auto-regressive (AR) model.
- C2. The FIR equalizer length is sufficient, $L \geq L_0$, with $L_0 = (M + 1)$.

Indeed, a channel satisfying C1 can be equalized by an FIR filter \mathbf{f}_0 with minimum length L_0 . If the equalizer filter is over-parameterized, i.e., its length verifies $L > L_0$, there exist $P = (L - L_0 + 1)$ exact ZF solutions, each one corresponding to a different equalization delay:

$$\mathbf{f}_p = [\mathbf{0}_{p-1}^T, \mathbf{f}_0^T, \mathbf{0}_{p-p}^T]^T, \quad 1 \leq p \leq P. \quad (15.11)$$

As will be seen in the following, under these conditions the MMSE, CM and CP criteria can be perfectly minimized (even cancelled if the sources verify the conditions of each criterion), and the global minimum can be computed algebraically, that is, without iterative optimization. The algebraic solution to the CP criterion (section 15.6.2) can be considered as an extension of the ACMA algorithm [82] to the CP principle; consequently, it can be referred to as *algebraic constant power algorithm (ACPA)*. The algebraic solutions to the supervised and blind criteria are later combined (section 15.6.3), giving rise to algebraic semi-blind equalizers. In practice, even if conditions C1–C2 are not satisfied, algebraic solutions can be used as judicious initializations for iterative equalizers (section 15.7).

15.6.1 Algebraic MMSE equalizer

It is well known that the MMSE criterion (15.3) is minimized by the Wiener-Hopf solution:

$$\mathbf{f}_{\tau\text{MMSE}} = \mathbf{R}_x^{-1} \mathbf{p}_\tau, \quad \text{with } \mathbf{R}_x = \mathbb{E}\{\mathbf{x}_n \mathbf{x}_n^H\} \text{ and } \mathbf{p}_\tau = \mathbb{E}\{\mathbf{x}_n \tilde{s}_{n-\tau}^*\}.$$

15.6 Algebraic equalizers 605

Assuming that the source signal is normalized (i.e., it has zero mean and unit variance), the mean square error (MSE) of the MMSE solution with delay τ is given by:

$$\text{MSE}_\tau = 1 - \mathbf{p}_\tau^H \mathbf{R}_x^{-1} \mathbf{p}_\tau.$$

Let the observed block associated with the pilot symbols at delay τ be denoted by:

$$\check{\mathbf{X}}_\tau = [\mathbf{x}_\tau, \mathbf{x}_{\tau+1}, \dots, \mathbf{x}_{\tau+N_t-1}]$$

with $N_t \geq L$. Cancelling the criterion (15.3) would be tantamount to solving the linear system:

$$\check{\mathbf{X}}_\tau^H \mathbf{f} = \check{\mathbf{s}}. \quad (15.12)$$

However, such a system does not generally have an exact solution, as it consists of more equations than unknowns. Its LS solution is given by:

$$\mathbf{f}_{\tau\text{LS}} = (\check{\mathbf{X}}_\tau \check{\mathbf{X}}_\tau^H)^{-1} \check{\mathbf{X}}_\tau \check{\mathbf{s}}, \quad (15.13)$$

which we consider here as the algebraic solution to the MMSE criterion (15.3). This solution exists and is unique as long as matrix $\check{\mathbf{X}}_\tau$ is full rank, which is the case in the presence of noise. In the noiseless case, the whole observation matrix \mathbf{X} given in (15.2) has rank L_0 , so that it exists an infinite number of solutions to system (15.12) as soon as $L > L_0$. Under conditions C1–C2, the minimum-norm solution is given by $\mathbf{f}_{\tau\text{LS}} = \check{\mathbf{X}}_\tau^\dagger \check{\mathbf{s}}$, where $(\cdot)^\dagger$ denotes the Moore-Penrose pseudo-inverse. This solution corresponds to one of the exact ZF equalizers (15.11), which are identical up to a delay. In the presence of noise, the impact of delay on equalization performance may become important. The optimal delay in the MMSE sense, τ_{opt} , can be determined by comparing the MSE of the different equalization delays according to (15.13):

$$\tau_{\text{opt}} = \arg \min_{\tau} \text{MSE}_\tau = \arg \max_{\tau} \mathbf{p}_\tau^H \mathbf{R}_x^{-1} \mathbf{p}_\tau.$$

15.6.2 Algebraic blind equalizers

The algebraic solution to the CM criterion has been developed at length in [25,81,82]. Hence, we only describe in this section the solution to the CP criterion, that we naturally refer to as ACPA. We will see that the search for such solutions can be associated with interesting tensor decomposition problems.

15.6.2.1 Determining a basis of the solution space

The exact minimizers of (15.6) are the solutions to the system of equations:

$$(\mathbf{f}^H \mathbf{x}_{n+L-1})^q = d_n, \quad n = 0, 1, \dots, N-1. \quad (15.14)$$

606 CHAPTER 15 Semi-blind methods for communications

This nonlinear system can be linearized by taking into account that $(\mathbf{f}^H \mathbf{x}_n)^q = \mathbf{f}^{\otimes q H} \mathbf{x}_n^{\otimes q}$, and can be compactly expressed as:

$$\mathbf{X}^{qH} \mathbf{w} = \mathbf{d} \quad (15.15)$$

where $\mathbf{X}^q = [\mathbf{x}_{L-1}^{\otimes q}, \mathbf{x}_L^{\otimes q}, \dots, \mathbf{x}_{N_d-1}^{\otimes q}]$ and $\mathbf{d} = [d_0, d_1, \dots, d_{N-1}]^H$. Equation (15.15) must be solved under the following structural constraint: $\mathbf{w} \in \mathbb{C}^{L_q}$ must be of the form $\mathbf{w} = \mathbf{f}^{\otimes q}$, for certain vector $\mathbf{f} \in \mathbb{C}^L$.

Under conditions C1–C2, there must be P linearly independent solutions. Consequently, the dimension of the null space of \mathbf{X}^{qH} , denoted $\ker(\mathbf{X}^{qH})$, is $(P-1)$, and the solutions of (15.15) can be expressed 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 (15.15) and $\mathbf{w}_p \in \ker(\mathbf{X}^{qH})$, for $1 \leq p \leq (P-1)$.

As in [82], we find more convenient to work in a vector space, obtained through a unitary transformation \mathbf{Q} with dimensions $(N \times N)$, such that $\mathbf{Q}\mathbf{d} = [\sqrt{N}, \mathbf{0}_{N-1}^T]^T$. For instance, \mathbf{Q} can be a Householder transformation [32] or, if \mathbf{d} is composed of N equal values (as is the case for PSK sources), an N -point DFT matrix. Then, denoting:

$$\mathbf{Q}\mathbf{X}^{qH} = \begin{bmatrix} \mathbf{r}^H \\ \mathbf{R} \end{bmatrix},$$

system (15.15) reduces to:

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

under the constraint $\mathbf{w} = \mathbf{f}^{\otimes q}$. Similarly to [82, Lemma 4], it is possible to prove that this problem is equivalent to the solution of:

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

followed by a scaling factor to enforce:

$$\mathbf{c}^H \mathbf{w} = 1, \quad \text{with } \mathbf{c} = \frac{1}{\|\mathbf{d}\|^2} \sum_{n=0}^{N-1} d_n \mathbf{x}_n^{\otimes q} \quad (15.16)$$

or, equivalently:

$$\frac{1}{\|\mathbf{d}\|^2} \sum_{n=0}^{N-1} d_n (\mathbf{f}^H \mathbf{x}_n)^q = 1. \quad (15.17)$$

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

$$N_d \geq L_q + L_0 - 1 \quad (15.18)$$

15.6 Algebraic equalizers 607

(or $N > L_q - P$), then $\dim \ker(\mathbf{R}) = P$, where L_0 is defined in condition C2 above. Hence, all the solutions of $\mathbf{R}\mathbf{w} = \mathbf{0}$ are linearly spanned by a basis $\{\mathbf{w}_k\}_{k=1}^P$ of $\ker(\mathbf{R})$. Such a basis can be computed from the *singular value decomposition (SVD)* of matrix \mathbf{R} by taking its P least significant right singular vectors. The structured solutions $\{\mathbf{f}_p^{\otimes q}\}_{p=1}^P$ are also a basis of this subspace and, as a result, there exists a set of scalars $\{\alpha_{pk}\}_{p,k=1}^P$ such that:

$$\mathbf{f}_p^{\otimes q} = \sum_{k=1}^P \alpha_{pk} \mathbf{w}_k, \quad 1 \leq p \leq P \quad (15.19)$$

where matrix $(\mathbf{A})_{kp} = \alpha_{pk}$ is full rank. The problem of finding structured solutions to the linearized problem (15.15) is hence a particular *subspace fitting* problem with structural constraints. In terms of q th-order tensors, Eq. (15.19) can be rewritten as:

$$\mathbf{f}_p^{\otimes q} = \sum_{k=1}^P \alpha_{pk} \mathbf{W}_k, \quad 1 \leq p \leq P \quad (15.20)$$

where $\mathbf{W}_k = \text{unvecs}_q\{\mathbf{w}_k\}$. This is the tensorial *rank-1 linear combination problem*, which can be stated as follows: given a set of q th-order tensors $\{\mathbf{W}_k\}$, find the scalars $\{\alpha_{pk}\}$ in Eq. (15.20) yielding rank-1 tensors. The obtained rank-1 tensors correspond to $\{\mathbf{f}_p^{\otimes q}\}$. This tensor decomposition is generally a non-trivial task [13,15].

Before resuming our search for algebraic solutions to the CP contrast, it is interesting to remark that satisfactory algebraic equalization can be achieved in practice with observation windows shorter than the sample size bound (15.18), as illustrated by the numerical analysis of section 15.8.

15.6.2.2 Structuring the solutions: a subspace approach

A subspace method reminiscent of [55] can be used to recover the minimum-length equalizer \mathbf{f}_0 from a basis of (generally) unstructured solutions $\{\mathbf{w}_k\}_{k=1}^P$. The subspace fitting problem (15.19) can be written in compact form as $\mathbf{W}\mathbf{A} = \mathbf{F}$, with $\mathbf{W} = [\mathbf{w}_1, \dots, \mathbf{w}_p]$ and $\mathbf{F} = [\mathbf{f}_1^{\otimes q}, \dots, \mathbf{f}_p^{\otimes q}]$. Since \mathbf{A} is full rank, matrices \mathbf{W} and \mathbf{F} span the same column space, denoted by $\text{range}(\mathbf{W}) = \text{range}(\mathbf{F})$. In particular, $\forall \mathbf{u}_i \in \ker(\mathbf{W}^H)$, $\mathbf{u}_i^H \mathbf{F} = \mathbf{0}_p^T$. There are $\dim \ker(\mathbf{W}^H) = (L_q - P)$ such linearly independent vectors.

Since equalization solutions have the form (15.11), the corresponding columns of \mathbf{F} have a particular structure whereby the elements non associated with the minimum-length equalizer \mathbf{f}_0 are all zero. The remaining $L_{0q} = \binom{L_0+q-1}{q}$ elements form $\mathbf{f}_0^{\otimes q}$. Let σ_p describe the set of L_{0q} positions of $\mathbf{f}_0^{\otimes q}$ in $\mathbf{f}_p^{\otimes q}$, that is, $\sigma_p = \{j_1 + L(j_2 - 1) + \dots + L^{q-1}(j_q - 1)\}$, with $j_k \in [p, p + L_0 - 1]$, $k = 1, \dots, q$, and $j_1 \geq j_2 \geq \dots \geq j_q$. Similarly, $(\mathbf{u}_i)_{\sigma_p} \in \mathbb{C}^{L_{0q}}$ is the sub-vector composed of the elements of \mathbf{u}_i in positions σ_p . Denote by $\mathbf{U}_i = [(\mathbf{u}_i)_{\sigma_1}, \dots, (\mathbf{u}_i)_{\sigma_p}] \in \mathbb{C}^{L_{0q} \times P}$. Hence:

$$\mathbf{u}_i^H \mathbf{F} = \mathbf{0}_p^T \quad \Leftrightarrow \quad \mathbf{U}_i^H \mathbf{f}_0^{\otimes q} = \mathbf{0}_p.$$

608 CHAPTER 15 Semi-blind methods for communications

The above equalities define a set of $P(L_q - P)$ linear equations, characterized by matrix $\mathbf{U} = [\mathbf{U}_1, \dots, \mathbf{U}_{L_q - P}] \in \mathbb{C}^{L_{0q} \times P(L_q - P)}$, where the unknowns are the components of $\mathbf{f}_0^{\otimes q}$. As long as $L > L_0$, this linear system determines, up to a scale factor, the well-structured vector $\mathbf{f}_0^{\otimes q}$; its amplitude can then be set via (15.17) from a zero-padded version of the estimate of \mathbf{f}_0 , yielding one of the solutions \mathbf{f}_p in Eq. (15.11). In practice, we minimize the quadratic form $\|\mathbf{U}^H \mathbf{f}_0^{\otimes q}\|^2 = \mathbf{f}_0^{\otimes qH} \mathbf{U} \mathbf{U}^H \mathbf{f}_0^{\otimes q}$, which leads to the estimation of $\mathbf{f}_0^{\otimes q}$ by 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 $\hat{\mathbf{A}}_{\text{LS}} = (\mathbf{W}^H \mathbf{W})^{-1} \mathbf{W}^H \mathbf{F} = \mathbf{W}^\dagger \mathbf{F}$. The elements of $\hat{\mathbf{A}}_{\text{LS}}$ provide a solution to the rank-1 linear combination problem.

15.6.2.3 Recovering the equalizer vector from its symmetric tensor

In order to recover the equalizer impulse response \mathbf{f}_0 from its symmetric vectorization $\mathbf{f}_0^{\otimes q}$, it is possible to perform the SVD of a matrix unfolding of $\mathbf{f}_0^{\otimes q} = \text{unvecs}_q\{\mathbf{f}_0^{\otimes q}\}$ [11, 35]. Denote by $\mathbf{F}_0 \in \mathbb{C}^{L_0 \times L_0^{q-1}}$ the matrix with elements $(\mathbf{F}_0)_{i_1, i_2 + L_0(i_3 - 1) + \dots + L_0^{q-2}(i_q - 1)} = (\mathbf{f}_0^{\otimes q})_{i_1 i_2 i_3 \dots i_q}$. Then, $\mathbf{F}_0 = \mathbf{f}_0 \mathbf{f}_0^{\text{T}}$, with $(\bar{\mathbf{f}}_0)_{i_2 + L_0(i_3 - 1) + \dots + L_0^{q-2}(i_q - 1)} = (\mathbf{f}_0)_{i_2} (\mathbf{f}_0)_{i_3} \dots (\mathbf{f}_0)_{i_q}$. Hence, \mathbf{f}_0 can be estimated (up to a scale factor) as the dominant left singular vector of the matrix unfolding of \mathbf{F}_0 . This matrix has rank one in the absence of noise.

In the presence of noise, it is generally no longer possible to express the estimated vector $\hat{\mathbf{f}}_0^{\otimes q}$ as the symmetric vectorization of a rank-1 tensor. In other words, no vector \mathbf{f}_0 exists such that $\hat{\mathbf{f}}_0^{\otimes q} = \text{vecs}_q\{\mathbf{f}_0^{\otimes q}\}$ is verified exactly. Consequently, the matrix unfolding will not be of rank one, and the SVD-based procedure explained above will only yield an approximate solution. One is actually facing the problem of the *rank-1 approximation* to the symmetric tensor $\hat{\mathbf{f}}_0^{\otimes q}$. To date, only iterative solutions, e.g., inspired on the *iterative power method* [17, 41], have been proposed to solve this problem. However, our experiments reveal that the solution previously described for the noiseless case is an excellent initialization.

15.6.2.4 Other structuring methods

In the context of the CM criterion, a subspace method similar to that of section 15.6.2.2 was proposed in [25, section III.C], operating on a single non-structured (LS) solution (see also [24]). This structure forcing procedure can be interpreted as the diagonalization of the matrix associated with the non-structured solution. By contrast, our approach takes advantage of a whole basis of the solution subspace, which should lead to an improved robustness, particularly for large values of P . The method of [35] and [25, section III.B] is based on the observation that the first L components of a solution \mathbf{w}_k are equal to $\tilde{\alpha}_{k1} f_1^{q-1} [f_1, \sqrt{q} f_2, \dots, \sqrt{q} f_{L_0-1}, \sqrt{q} f_{L_0}, \mathbf{0}_{P-1}^T]^T$, $\tilde{\alpha}_{kp} = (\mathbf{A}^{-1})_{pk}$, from which \mathbf{f}_0 can be extracted. This method is simple and ingenious, but inaccurate when coefficient $\tilde{\alpha}_{k1}$ or the first term f_1 of the equalizer are small relative to the noise level.

To surmount this limitation, we can note that the last components of \mathbf{w}_k are equal to [25, section III.B]:

$$\tilde{\alpha}_{kP} f_{L_0}^{q-1} [\mathbf{0}_{P-1}^T, \dots, \sqrt{q}f_1, \sqrt{q}f_2, \dots, \sqrt{q}f_{L_0-1}, f_{L_0}]^T.$$

Appropriately combined with the estimation carried out from the first L components, this second option can provide an improved estimation of \mathbf{f}_0 . In the simulation study of section 15.8, we employ the following heuristic (suboptimal) linear combination. Let us suppose that the filters estimated from the first and the last non-overlapping components of a non-structure 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 $\hat{\mathbf{f}}_0 = [\hat{\mathbf{f}}_1, \hat{\mathbf{f}}_2] \boldsymbol{\gamma}$, with $\boldsymbol{\gamma} = \boldsymbol{\beta}^* / \|\boldsymbol{\beta}\|^2$, $\boldsymbol{\beta} = [\beta_1, \beta_2]^T$. The coefficients of $\boldsymbol{\beta}$ can simply be estimated from the equation $\beta_i = \|\hat{\mathbf{f}}_i\|$, $i = 1, 2$. This type of linear *maximal ratio combining* is reminiscent of the *RAKE* receiver and the matching filter [58]. Robustness can still be improved by exploiting a whole set $\{\mathbf{w}_k\}$ instead of a single solution, as explained above.

15.6.2.5 Approximate solution in the presence of noise

In the presence of additive noise at the sensor output, conditions C1–C2 are no longer satisfied, and an exact solution of (15.14) may not exist. An approximate solution in the LS sense can be obtained by minimizing $\|\mathbf{X}^q \mathbf{w} - \mathbf{d}\|^2$, under the structural constraint $\mathbf{w} = \mathbf{f}^{\otimes q}$. This minimization generally requires an iterative method, as detailed in section 15.7.

Nevertheless, the guidelines for determining an exact solution in the noiseless case can still provide a sensible initialization to an iterative equalizer in the noisy case. After applying the transformation \mathbf{Q} , the LS problem proves equivalent to the minimization of the quadratic form:

$$|\mathbf{c}^H \mathbf{w} - 1|^2 + \|\mathbf{R} \mathbf{w}\|^2. \quad (15.21)$$

To find a basis of the solution space, we seek a set of vectors minimizing $\|\mathbf{R} \mathbf{w}\|^2$ (for instance, the least significant P right singular vectors of \mathbf{R}), then structure them as in section 15.6.2.2 and finally normalize the solution to satisfy $\mathbf{c}^H \mathbf{w} = 1$ [cf. Eqs (15.16)–(15.17)]. Although suboptimal, this solution will be tested in the experimental study of section 15.8.

A more accurate solution can be determined by realizing that expression (15.21) represents a non-negative quadratic form in vector $[\mathbf{w}^T, 1]^T$. Formulating the problem in the projective space, we can look for the least significant eigenvector \mathbf{v}_m of matrix:

$$\begin{bmatrix} \mathbf{R} \mathbf{R}^H + \mathbf{c} \mathbf{c}^H & -\mathbf{c} \\ -\mathbf{c}^H & 1 \end{bmatrix}$$

and take as an approximate estimation of \mathbf{w} the first $\dim(\mathbf{w})$ components of \mathbf{v}_m normalized by the first one.

15.6.3 Algebraic semi-blind equalizers

By extending the above ideas, we can also develop algebraic solutions to the semi-blind criterion CP-MMSE (15.8), the solutions to criterion CM-MMSE (15.7) being obtained in a totally analogous manner. To minimize algebraically the CP-MMSE criterion, we seek the simultaneous solution of systems (15.12) and (15.15):

$$\check{\mathbf{X}}_{\tau}^H \mathbf{f} = \check{\mathbf{s}} \quad (15.22)$$

$$\mathbf{X}^q H \mathbf{w} = \mathbf{d} \quad (15.23)$$

under the structural constraint $\mathbf{w} = \mathbf{f}^{\otimes q}$, where now

$$\mathbf{X}^q = [\mathbf{x}_{\tau+N_i}^{\otimes q}, \mathbf{x}_{\tau+N_i+1}^{\otimes q}, \dots, \mathbf{x}_{N_d-1}^{\otimes q}]$$

and $\mathbf{d} = [d_{N_i}, d_{N_i+1}, \dots, d_{N_d-\tau-1}]^H$. Note that only the symbols not employed in the supervised part contribute to the blind part of the criterion.

The case where conditions C1–C2 are verified is trivial, since both solutions of the composite system are exact and identical. Hence, let us first consider the case of a noisy AR channel with a sufficiently long equalizer. A suboptimal solution can be obtained by combining the solutions computed separately for the two sub-systems [11,35]. Let $\hat{\mathbf{f}}_{\text{MMSE}}$ denote the solution of (15.22) and $\hat{\mathbf{f}}_{\text{CP}}^{\otimes q}$ that of (15.23) associated with the same equalization delay τ ; these solutions are computed as explained in sections 15.6.1 and 15.6.2, respectively. Let us unfold $\text{unvecs}_q\{\hat{\mathbf{f}}_{\text{CP}}^{\otimes q}\}$ into a matrix \mathbf{F}_{CP} with dimensions $(L \times L^{q-1})$, as described in section 15.6.2.3. Then, the joint solution to (15.22)–(15.23) can be approximated by the dominant left singular vector of matrix $\mathbf{F}_{\text{SB}} = [\lambda \hat{\mathbf{f}}_{\text{MMSE}}, (1 - \lambda) \mathbf{F}_{\text{CP}}]$. In the noiseless case, solutions $\hat{\mathbf{f}}_{\text{MMSE}}$ and $\hat{\mathbf{f}}_{\text{CP}}$ coincide with the dominant left singular vector of the rank-1 matrix \mathbf{F}_{SB} ; an iterative search is not necessary.

In the case of an FIR channel, no exact solution to system (15.22)–(15.23) exists, even in the absence of noise. However, the two sub-systems can be solved separately in the LS sense and the respective solutions can then be combined according to the above SVD-based procedure. We refer to this method as *semi-blind algebraic constant power algorithm (SB-ACPA)*.

The combined solution just described can initialize an iterative minimization algorithm aiming to refine this algebraic approximate solution.

15.7 ITERATIVE EQUALIZERS

15.7.1 Conventional gradient-descent algorithms

In practice, exact ZF equalization may not be feasible, due to noise or just to an insufficient equalizer length. In such cases, the cost function must be minimized iteratively,

15.7 Iterative equalizers 611

for instance, via a gradient-descent or a Newton algorithm. We describe here gradient-descent methods; these results can easily be extended to Newton implementations. If a good initialization has been obtained, only a few iterations will usually be necessary for convergence.

We define the complex gradient of a generic real-valued function $\Upsilon(\mathbf{f})$ with respect to complex variable \mathbf{f} as:

$$\nabla\Upsilon(\mathbf{f}) = \nabla_{\mathbf{f}_r} \Upsilon(\mathbf{f}) + j \nabla_{\mathbf{f}_i} \Upsilon(\mathbf{f})$$

where $\mathbf{f}_r = \Re(\mathbf{f})$ and $\mathbf{f}_i = \Im(\mathbf{f})$ represent the real and imaginary parts, respectively, of vector \mathbf{f} . Up to an inconsequential scale factor, this definition corresponds to Brandwood's complex gradient [3]. Accordingly, the gradients of the CM (15.4) and CP (15.6) criteria can be expressed as:

$$\nabla\Upsilon_{\text{CM}}(\mathbf{f}) = 4\mathbb{E}\{(\mathbf{f}^H \mathbf{x}_n)^* [|\mathbf{f}^H \mathbf{x}_n|^2 - \gamma] \mathbf{x}_n\} \quad (15.24)$$

$$\nabla\Upsilon_{\text{CP}}(\mathbf{f}) = 2q\mathbb{E}\{(\mathbf{f}^H \mathbf{x}_n)^{q-1} [(\mathbf{f}^H \mathbf{x}_n)^q - d_n]^* \mathbf{x}_n\}. \quad (15.25)$$

From the relationships remarked in section 15.5.2, the gradient of MMSE criterion (15.3) can be computed from that of the CP criterion by setting $q = 1$ and replacing \check{s}_n by d_n in expression (15.25). This yields:

$$\nabla\Upsilon_{\text{MMSE}}(\mathbf{f}) = 2\mathbb{E}\{[(\mathbf{f}^H \mathbf{x}_n) - \check{s}_n]^* \mathbf{x}_n\}. \quad (15.26)$$

The gradients of the semi-blind CM-MMSE and CP-MMSE criteria are simply obtained by linear combination of (15.24)–(15.26) according to (15.7)–(15.8). We refer to the resulting iterative methods as *constant modulus algorithm (CMA)* and *constant power algorithm (CPA)*; and their semi-blind versions as *semi-blind constant modulus algorithm (SB-CMA)* and *semi-blind constant power algorithm (SB-CPA)*.

As a judicious initialization in the blind case, we can employ the equalizer vector provided by an algebraic method, such as the direct LS (generally non-structured) solution of the linearized problem (15.15), $\hat{\mathbf{f}}_{\text{LS}} = (\mathbf{X}^q)^{\dagger} \mathbf{d}$, or the structured solution described in section 15.6.2. In the semi-blind case, the algebraic solution of section 15.6.3 becomes applicable as initialization. At each iteration, the equalizer vector can be adjusted by means of a gradient-based update:

$$\mathbf{f}^+ = \mathbf{f} - \mu \nabla\Upsilon(\mathbf{f}). \quad (15.27)$$

Iterations are stopped when

$$\frac{\|\mathbf{f}^+ - \mathbf{f}\|}{\|\mathbf{f}\|} < \eta/N \quad (15.28)$$

where η is a small positive constant.

612 CHAPTER 15 Semi-blind methods for communications

We advocate the use of *block* or *batch* implementations [10], also known as fixed-window methods [61], rather than stochastic algorithms. The latter approximate the gradient by a single-sample estimate, which may be seen as dropping the expectation operator in the gradient expression. This simplification, which in the case of the CM criterion gives rise to the stochastic-gradient CMA, generally leads to a slow convergence and a poor final accuracy. Indeed, a single parameter, μ , must control at the same time the step size in the search trajectory and the implicit statistical average; this is a difficult balance. Stochastic algorithms found justification when the available computer power was rather limited. Nowadays, computer power is no longer the limiting factor of equalization performance, but the algorithms that are implemented, or the operating conditions (e.g., channel non-stationarity).

By contrast, batch methods estimate the gradient from a whole block of channel output samples, using the same data block at each iteration. This gradient estimate is more accurate and thus improves the convergence speed and equalization quality of the resulting algorithm [10,61]. Moreover, tracking capabilities are not necessarily sacrificed, since good performance can be achieved from small data blocks; it suffices that the channel be stationary over the (short) observation window. Block methods are particularly suited to burst transmission systems (e.g., TDMA). The possibility of combining batch and stochastic operation in iterative optimization methods is discussed in section 6.4.3, Chapter 6.

It is well known that gradient-based algorithms for blind equalization, despite their simplicity, present numerous drawbacks such as lack of robustness to local extrema, dependence on initialization and slow convergence [21,22,40]. These problems persist in block implementations, even though convergence is often accelerated. When the function to be optimized is convex in the unknowns, this problem can be alleviated with more elaborate approaches such as the conjugate gradient [57]. Nevertheless, the blind and semi-blind functions based on the CM and CP criteria (section 15.5) are not convex. This leads us to consider alternative optimization strategies without compromising the simplicity and numerical convenience of the implementation.

15.7.2 Algorithms based on algebraic optimal step size

15.7.2.1 Step-size polynomials

Exact *global* line search aims at finding the step size minimizing the cost function along the search direction:

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

A possible search direction is simply the gradient, $\mathbf{g} = \nabla \Upsilon(\mathbf{f})$. These algorithms are generally unattractive due to their complexity, because the one-dimensional minimization must typically be carried out by costly numerical techniques. Another drawback is the orthogonality between successive gradient vectors (see section 15.7.2.3), which, depending on the initialization and the shape of the cost-function surface, can slow down convergence [57].

15.7 Iterative equalizers 613

However, it has been observed in [14,34] that, for a number of equalization criteria, including the CM, the CP and their semi-blind versions studied herein, functional $\Upsilon(\mathbf{f} - \mu \mathbf{g})$ is a rational function in the step size μ . This allows us to find μ_{opt} algebraically, so that it is possible to *globally* minimize the cost function in the descent direction while reducing complexity. Indeed, for the CM criterion (15.4), some algebraic manipulations show that the derivative of $\Upsilon_{\text{CM}}(\mathbf{f} - \mu \mathbf{g})$ with respect to μ is the following cubic:

$$p(\mu) = b_3 \mu^3 + b_2 \mu^2 + b_1 \mu + b_0. \quad (15.29)$$

Its real coefficients are given by [93]:

$$\begin{aligned} b_3 &= 2\mathbb{E}\{a_n^2\}, & b_2 &= 3\mathbb{E}\{a_n b_n\} \\ b_1 &= \mathbb{E}\{2a_n c_n + b_n^2\}, & b_0 &= \mathbb{E}\{b_n c_n\} \end{aligned}$$

where $a_n = |g_n|^2$, $b_n = -2\text{Re}(y_n g_n^*)$, and $c_n = (|y_n|^2 - \gamma)$, with $g_n = \mathbf{g}^H \mathbf{x}_n$. Similarly, for the CP criterion (15.6), the optimal step size μ_{opt} is found among the roots of the $(2q - 1)$ th-degree polynomial [92]:

$$p_{\text{CP}}(\mu) = \sum_{m=0}^{2q-1} \text{Re}(b_m) \mu^m \quad (15.30)$$

where

$$b_m = \begin{cases} \sum_{p=0}^m (m+1-p) \mathbb{E}\{a_{m+1-p}^* a_p\} - (m+1) \mathbb{E}\{a_{m+1}^* d_n\}, & 0 \leq m \leq q-1 \\ \sum_{p=m+1-q}^q (m+1-p) \mathbb{E}\{a_{m+1-p}^* a_p\}, & q \leq m \leq 2q-1 \end{cases}$$

with

$$a_p = (-1)^p \binom{q}{p} g_n^p y_n^{q-p}, \quad 0 \leq p \leq q.$$

The step-size polynomial of the MMSE criterion is easily determined by taking into account the link between the MMSE and CP criteria observed in section 15.5.2, which leads to:

$$p_{\text{MMSE}}(\mu) = b_1 \mu + b_0 \quad (15.31)$$

$$b_1 = \mathbb{E}\{|g_n|^2\}, \quad b_0 = -\text{Re}(\mathbb{E}\{g_n^*(y_n - \check{s}_n)\}). \quad (15.32)$$

The polynomials defining the optimal step size for the semi-blind CM-MMSE and CP-MMSE criteria are made up of polynomials (15.29), (15.30) and (15.31) according to the linear combinations of the respective cost functions (15.7)–(15.8). A similar polynomial (a quartic) is obtained for the kurtosis contrast, leading to the *RobustICA* algorithm described in section 6.11.2 of Chapter 6.

614 CHAPTER 15 Semi-blind methods for communications

Once the coefficients have been determined, the roots of the optimal step-size polynomial can be obtained as explained in section 15.7.2.2. The optimal step size corresponds to the root attaining the minimal value of the cost function, thus leading to the *global* minimization of $\Upsilon(\cdot)$ in the descent direction. After determining μ_{opt} , the equalizer vector coefficients are updated as in (15.27), and the process is repeated with the new equalizer and gradient vectors, until convergence, which is tested with (15.28). We refer to this technique as *optimal step-size (OS) algorithm*, which gives rise, in particular, to the blind *OS-CMA* and *OS-CPA* algorithms and to the semi-blind *OS-SB-CMA* and *OS-SB-CPA* algorithms.

To improve numerical conditioning in the determination of μ_{opt} , it is useful to normalize the gradient vector \mathbf{g} beforehand. Since the pertinent parameter is the search direction $\hat{\mathbf{g}} = \mathbf{g}/\|\mathbf{g}\|$, this normalization does not cause any inconvenience. As a consequence, vector \mathbf{g} is replaced by $\hat{\mathbf{g}}$ when computing the optimal step-size polynomial coefficients, as well as in the update rule (15.27).

15.7.2.2 Root extraction

Standard procedures such as *Cardan's formula*, or often less costly iterative methods [27, 45], are available to extract the roots of cubics (15.29) and (15.30) with $q = 2$; an efficient MatlabTM implementation, valid for polynomials with real or complex coefficients, is given in [57] (see also [95]). For solving quartics, elementary algebra textbooks present the method developed by *Ferrari*, Cardan's student, in the 16th century. Since the end of the 18th century, we know that polynomials of degree higher than four cannot be solved by radicals; one thus needs to resort to iterative methods.

Concerning the roots of cubics (15.29) and (15.30) with $q = 2$, two options are possible: either all three roots are real-valued, or one is real and the two other form a complex-conjugate pair. In the first case, one just needs to verify which one provides the smallest value of $\Upsilon(\mathbf{f} - \mu\mathbf{g})$. In our computer experiments, when a complex-conjugate pair exists, it is the real root that typically minimizes the cost function. Even when the real root does not produce the minimal value of $\Upsilon(\cdot)$, it often provides lower MSE at the equalizer output than the complex roots. Real roots are thus preferred. This observation is also applicable to polynomials of higher degree; for instance, (15.30) with $q > 2$. Another possibility, employed in the RobustICA method based on the kurtosis contrast (section 6.11.2 of this book) is to consider only the real parts of the roots.

15.7.2.3 Convergence of optimal step-size algorithms

By construction of exact line search algorithms, gradient vectors of consecutive iterations are orthogonal, which, depending on initialization and the shape of the cost-function surface, can slow down convergence [57]. Gradient orthogonality is mathematically expressed as $\Re(\mathbf{g}^H \mathbf{g}^+) = 0$, with $\mathbf{g}^+ = \nabla \Upsilon(\mathbf{f}^+)$. This relationship can easily be derived by taking into account that

$$\frac{\partial \Upsilon(\mathbf{f} - \mu\mathbf{g})}{\partial \mu} = -\Re(\mathbf{g}^H \nabla \Upsilon(\mathbf{f} - \mu\mathbf{g})) = 0.$$

In our numerical experiments, the optimal step-size algorithms have always converged in fewer iterations than its fixed step-size counterparts [95]. Fast convergence and improved stability have also been reported in [87]. Moreover, the probability of converging to local extrema is decreased with the optimal step-size strategy, as shown empirically in [95] and section 15.8.

15.7.2.4 Variants

The coefficients of OS-CMA (15.29) and OS-CPA (15.30) cubics, the latter with $q = 2$, can also be determined from the sensor-output statistics, computed before starting the iterations [95,96]. This alternative requires the previous computation of the covariance matrix and the whole fourth-order cumulant tensor of the channel output. Indeed, the equalizer-output statistics can be deduced by multi-linearity; this way of computing the cumulants is called *deductive estimation* in [9]. Both alternatives are equivalent regarding equalization performance and convergence speed measured in terms of iterations. The only difference lies in their computational cost in terms of number of operations (section 15.7.2.5).

The algebraic optimal step-size technique can also be applied to other equalization criteria. For instance, the *kurtosis maximization (KM)*, also called Shalvi-Weinstein criterion [69] in the context of blind SISO equalization, can also be globally optimized along a given direction by rooting a fourth-degree polynomial in μ ; all stationary points can be computed by Ferrari's formula for quartics. This naturally gives rise to the OS-KMA, developed in the context of BSS and referred to as RobustICA algorithm in [96] (see also section 6.11.2 of this book), with a complexity per iteration similar to OS-CMA's. On the other hand, the optimal step-size technique remains applicable if data are prewhitened, for instance through a QR decomposition of the observation matrix, as in the QR-CMA method of [61]. Prewhitening improves conditioning and can accelerate convergence under the hypothesis of iid inputs. Finally, by using the Hessian of the cost function, the optimal step-size technique can easily be combined with Newton optimization, as well as with any other method constructing successive search directions $\{\mathbf{g}_k\}$.

15.7.2.5 Computational complexity

The complexity of the optimal step-size technique is dominated by the computation of polynomial coefficients [Eqs (15.29), (15.30), etc.]. In practice, mathematical expectations are replaced by sample averages over the observed signal block. The cost of these averages for (15.29) is of order $O(LN)$ per iteration, for data blocks composed of N vectors \mathbf{x}_n . For the alternative procedure based on the previous computation of the second- and fourth-order moments of the sensor output (section 15.7.2.4), the cost per iteration is approximately of the order of $O(L^4)$, with an additional initial cost of $O(L^4N)$ operations. Depending on the number of iterations needed for convergence and the relative values of N and L , this initial burden can render the second method (that we refer to as OS-CMA-2) more costly than the first one (OS-CMA-1) [93,95]. Similar alternatives are possible for the OS-CPA and OS-CPA algorithms.

Table 15.1 sums up the computational cost of different optimal step-size techniques in terms of number of real-valued *floating point operations (flops)*; a flop represents a

616 CHAPTER 15 Semi-blind methods for communications

Table 15.1 Computational cost in terms of number of flops for different iterative equalization algorithms in the case of real-valued signals and filters. L : number of equalizer filter coefficients; N : number of data vectors in the observed data burst

| | Initialization | Per iteration |
|---------------------|--|--------------------|
| OS-CMA-1 | – | $(3L + 10)N$ |
| OS-CMA-2 | $\left[\binom{L+3}{4} + \binom{L+1}{2} \right] N$ | $6L^4 + 3L^2 + 2L$ |
| OS-CPA | – | $[3L + q(q + 4)]N$ |
| OS-KMA | – | $(5L + 12)N$ |
| SG-CMA | – | $2(L + 1)$ |
| CMA | – | $2(L + 1)N$ |
| QR-CMA [61] | L^2N | $(2L + 3)N$ |
| RLS-CMA [5] | – | $2L(2L + 3)$ |
| AAF-CMA [70] | – | $6L$ |

product or a division followed by an addition, and typically corresponds to a multiply-and-accumulate (MAC) cycle in a digital signal processor (DSP). Also considered are other representative equalization techniques, specially those based on the CM criterion: the stochastic CMA (SG-CMA), the QR-CMA of [61], the recursive least squares CMA (RLS-CMA) of [5] and the accelerating adaptive filtering CMA (AAF-CMA) of [70]. Only dominant terms in the pertinent parameters (L, N) are retained in the flop counting, under the hypothesis of real-valued signals and filters. In the complex case, the cost is around four times that of the real case with the same parameters. Remark that the cost of the optimal step-size polynomial root extraction is independent of (L, N) and can thus be considered as negligible (see section 15.7.2.2).

The complexity per iteration of the OS-CPA and OS-CMA-1 is of the same order of magnitude, for moderate alphabet size relative to the equalizer length; both algorithms present practically the same cost for BPSK sources ($q = 2$). Finally, the complexity per iteration of the semi-blind techniques is essentially the same as that of their blind counterparts.

15.8 PERFORMANCE ANALYSIS

By means of a detailed empirical analysis, this section evaluates the performance of the different methods studied in this chapter.

15.8.1 Performance of algebraic blind equalizers

We begin by comparing the performance of the algebraic blind equalizers based on the CP criterion developed in section 15.6. The methods considered are: the direct non-structured LS solution of (15.15) (“LS, no struct”); the structuring method of [35]

15.8 Performance analysis 617

from the first non-overlapping components of the LS solution (“LS, top”); *idem*, from the last components (“LS, bottom”); the maximal ratio combination of the first and last components (“LS, top+bottom”); *idem*, from a whole basis of solutions (“basis, top+bottom”); and the subspace method of section 15.6.2.2 (“basis, subspace”). The “LS, top”, “LS, bottom”, “LS, top+bottom” and “basis, top+bottom” solutions are explained in section 15.6.2.4. After estimating the symmetric Kronecker vectorization in the direct LS and subspace solutions, the respective equalizer vectors are obtained from the SVD-based rank-1 tensor approximation described in section 15.6.2.3. The performance of the supervised MMSE receiver (15.13) is also computed as a reference. In the first simulation example, a QPSK signal ($q = 4$) excites the AR-1 channel:

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

with a pole located at $z_p = 0.5$. The impulse response of this channel is well approximated by an order-50 FIR filter. ISI is perfectly cancelled for the equalizer $\mathbf{f}_0 = [1, -0.5]^T$, with a dominant first coefficient. The minimal equalizer length is thus $L_0 = 2$, but we suppose its length has been overestimated as $L = 5$, generating $P = 4$ possible ZF solutions, which are just delayed versions of each other [as in (15.11)]. Complex circular additive white Gaussian noise corrupts the channel output, with a *signal-to-noise ratio (SNR)* given by $\mathbb{E}\{|r|^2\}/\mathbb{E}\{|v|^2\}$. Blocks of size $N_d = 100$ symbol periods are observed, and performance indices are averaged over ν independent *Monte Carlo (MC)* iterations, with $\nu N_d \geq 10^5$. Figure 15.1a shows the *symbol error rate (SER)* obtained by the algebraic equalizers as a function of the SNR. The performance of the direct LS solution stresses the need for structuring. Yet structuring from only the last components of the LS solution (“LS, bottom”) also offers poor results. By contrast, the other methods present a superior performance, just 2–4 dB over the MMSE bound. It is interesting to note that the first components of the LS solution provide the best results for moderate SNR in this scenario. This superiority, however, depends on the optimal equalizer configuration, as shown by the next example.

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

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

by delaying the truncated impulse response. The minimum-length equalizer is now $\mathbf{f}_0 = [1, -2]^T$, with a dominant last coefficient. Figure 15.1b shows the algebraic equalization results. The performance of the “LS-top” method degrades considerably, and becomes similar to that of the “LS-bottom” method in the previous experiment. The performance of the subspace-based structuring method remains practically unchanged compared to the simulation of Fig. 15.1a, thus showing its robustness to the relative weight of the equalizer coefficients.

618 CHAPTER 15 Semi-blind methods for communications

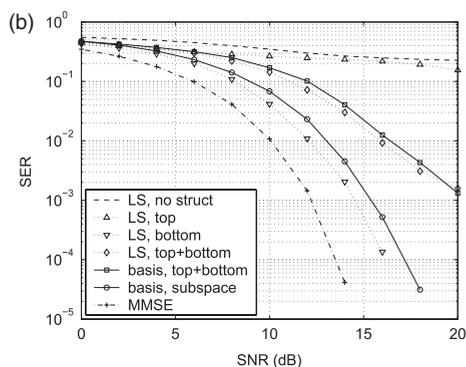
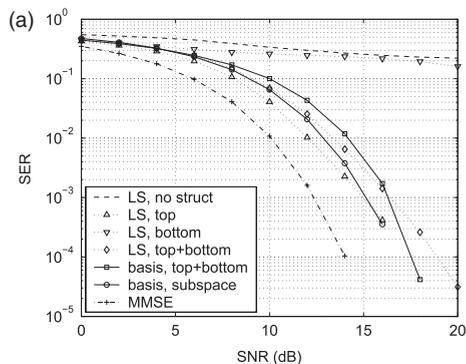


FIGURE 15.1

Algebraic blind equalization based on the CP criterion for different structuring methods, with a QPSK input ($q = 4$), $N_d = 100$ symbol periods, $L = 5$ ($L_0 = 2$), 1000 MC iterations: (a) channel $H_1(z)$; (b) channel $H_2(z)$.

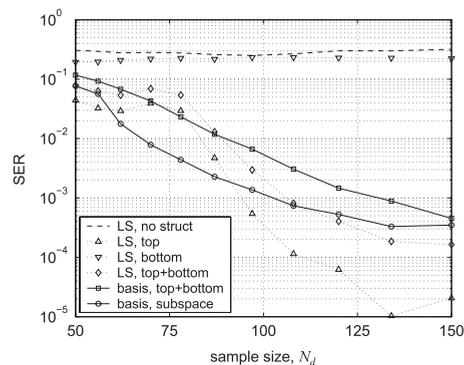


FIGURE 15.2

Algebraic blind equalization based on the CP criterion. Channel $H_1(z)$, QPSK input ($q = 4$), $L = 5$ ($L_0 = 2$), SNR = 15 dB, ν MC iterations, with $\nu N_d \geq 10^5$.

Figure 15.2 assesses the *sample size* needs of algebraic solutions, under the general conditions of the first experiment and with SNR = 15 dB. Satisfactory equalization from a basis of the solution space is obtained even under the bound imposed by (15.18) for this simulation example, $N_d \geq 71$. The subspace approach provides better results for short observation windows. However, the simplified procedure combining the first and last components of a single non-structured (LS) solution seems to yield reasonable results for a sufficient sample size.

The algebraic semi-blind solutions will be evaluated from section 15.8.4.

15.8.2 Attraction basins of blind and semi-blind CP equalizers

The following experiment evaluates the iterative methods based on the CP criterion, in blind as well as semi-blind operation (section 15.7). In particular, we aim at illustrating

Table 15.2 Average number of iterations for convergence in the experiments of Figs 15.3–15.4

| Step size | Blind | Semi-blind |
|-----------|-------|------------|
| Fixed | 422 | 363 |
| Optimal | 11 | 9 |

the ability of the optimal step-size technique in escaping from *spurious solutions* and that of the training sequence in eliminating them.

We observe a burst of $N_d = 200$ symbols with SNR = 10 dB at the output of channel $H_1(z)$ excited by a BPSK input. Figure 15.3a shows the contour lines (in the equalizer parameter space) of the logarithm of the CP criterion (15.6) for $L = L_0 = 2$, computed from the data. Solid lines represent the trajectories of the equalizer coefficients updated by the CPA (section 15.7.1) from 16 different initial configurations (marked by “+”) and $\eta = 10^{-5}$ in termination criterion (15.28); convergence points are marked by “×”. A fixed step size $\mu = 10^{-2}$ is chosen to obtain the fastest convergence without compromising stability. The plot also shows the MMSE solutions with delays zero and one, $\mathbf{f}_{\text{MMSE},0} = [0.85, -0.38]^T$ and $\mathbf{f}_{\text{MMSE},1} = [0, 0.70]^T$, yielding an output MSE of -8.66 and -4.98 dB, respectively. From most starting points, the algorithm converges to the desired solutions, near the optimal-delay MMSE equalizer. However, the trajectories get trapped in stable extrema located at $\pm[0.01, 0.58]$, near the suboptimal-delay MMSE equalizer. The attraction basins of these spurious solutions are not negligible and can have a significant negative impact on equalization performance. CPA requires, on average, around 500 iterations to converge (Table 15.2).

Under identical conditions, and operating on the same observed data, the trajectories of the OS-CPA equalizer (section 15.7.2) are plotted in Fig. 15.3b. Not only are undesired solutions avoided, but also convergence is considerably accelerated relative to the previous case: just about 10 iterations suffice (Table 15.2).

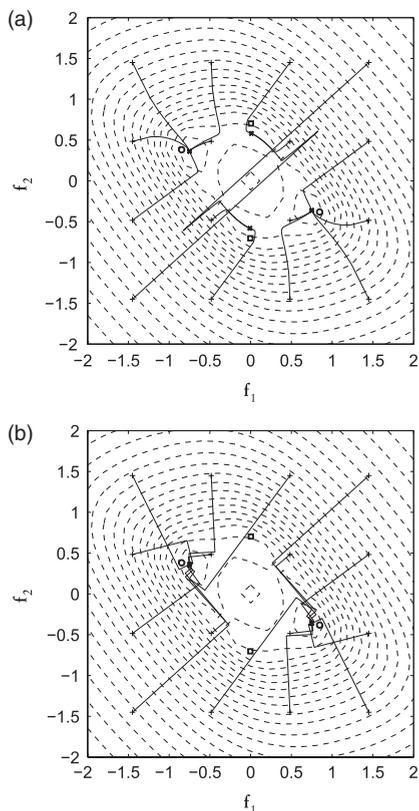
Using $N_t = 10$ pilot symbols and a confidence parameter $\lambda = 0.5$, the contour lines of semi-blind CP-MMSE criterion (15.8) have the shape shown in Fig. 15.4a. The introduction of training data modifies the CP cost function by stressing the global minimum near the optimal MMSE solution while suppressing the previously admissible equilibrium points symmetrically located across the origin. The optimal step size still leads to good equalization solutions (Fig. 15.4b) and, again, notably accelerates convergence (Table 15.2).

Similar results for the CM criterion are reported in [95].

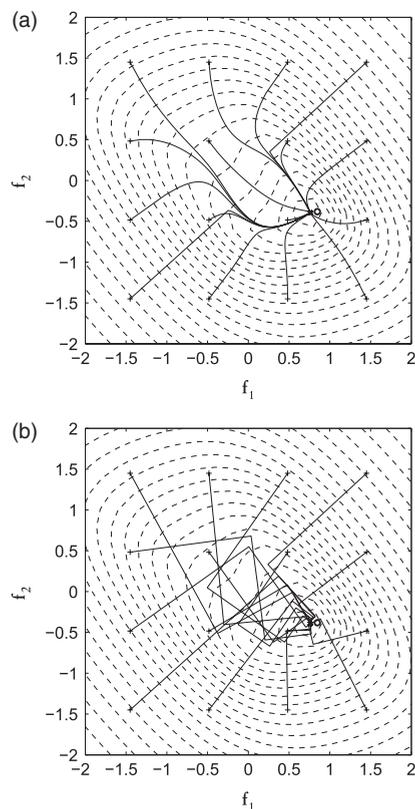
15.8.3 Robustness of optimal step-size CM equalizers to local extrema

The following experiment demonstrates the faster convergence speed of the OS-CMA compared with the fixed step-size CMA and the RLS-CMA of [5], as well as its ability to escape the attraction basins of undesired equilibria in the CM cost surface.

620 CHAPTER 15 Semi-blind methods for communications

**FIGURE 15.3**

(Dashed lines) Blind CP criterion contour lines. (Solid lines) Iterative equalizer trajectories: (a) CPA with $\mu = 10^{-2}$; (b) OS-CPA. Channel $H_1(z)$, BPSK input ($q = 2$), $N_d = 200$ symbol periods, $L = L_0 = 2$, SNR = 10 dB. “+”: initial point; “x”: final point; “o”: optimal-delay MMSE solution; “□”: suboptimal-delay MMSE solution.

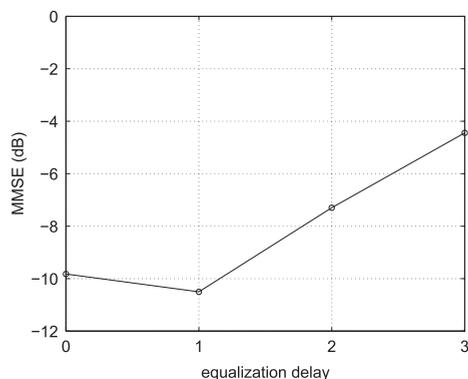
**FIGURE 15.4**

(Dashed lines) Semi-blind CP-MMSE criterion contour lines. (Solid lines) Iterative equalizer trajectories: (a) SB-CPA with $\mu = 10^{-2}$, (b) OS-SB-CPA. Same conditions as in Fig. 15.3, with $N_t = 10$ pilot symbols and $\lambda = 0.5$. “+”: initial point; “x”: final point; “o”: optimal-delay MMSE solution.

Bursts of $N_d = 200$ baud periods are observed at the output of a channel oversampled at twice the symbol rate (fractionally-spaced SIMO system) excited by a BPSK source ($\gamma = 1$) and corrupted by additive white Gaussian noise with 10-dB SNR. We choose the channel with impulse response:

$$\{0.7571, -0.2175, 0.1010, 0.4185, 0.4038, 0.1762\}$$

corresponding to the second example of [40, section 2.4, pp. 82–83], and a 4-tap equalizer. This system presents the theoretical output MMSE against equalization delay profile of

**FIGURE 15.5**

Theoretical output MMSE as a function of the equalization delay, for the experiment of section 15.8.3.

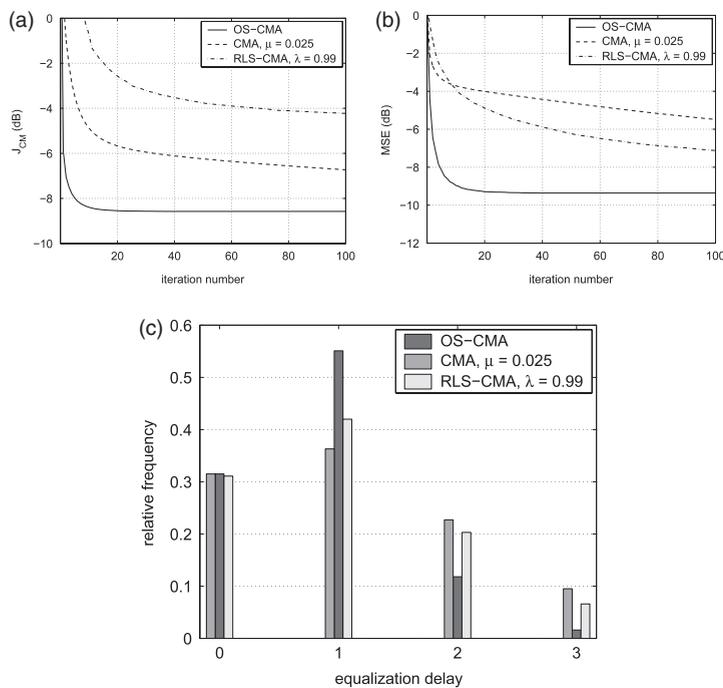
Fig. 15.5: delay 1 provides the best MMSE performance, closely followed by delay 0; the worst performance is obtained by delay 3. The initial equalizer coefficients are drawn randomly from a normalized Gaussian distribution before processing each signal block. The same initialization is used for all methods. A fixed step size $\mu = 0.025$ is found to prevent the divergence of the CMA. Following the guidelines given for the RLS-CMA in [5], we use the typical forgetting factor $\lambda_{\text{RLS}} = 0.99$ and an inverse covariance matrix initialized at the identity ($\delta = 1$). The samples of the observed signal block are re-used as many times as required. Iterations are stopped as soon as Eq. (15.28) is satisfied, with $\eta = 0.1\mu = 0.0025$. An upper bound of 1000 iterations is also set.

The evolution of the CM cost and the equalizer output MSE, averaged over 1000 independent signal blocks, are plotted in Fig. 15.6a–b. The normalized histogram of equalization delays obtained by the three methods appears in Fig. 15.6c, while Table 15.3 summarizes their computational cost. The CMA and the RLS-CMA often achieve the same suboptimal equalization delays. In contrast, the OS-CMA converges more frequently near the optimal-delay MMSE equalizer setting, and requires around an order of magnitude fewer iterations. Indeed, the CMA, the OS-CMA and the RLS-CMA converge to one of the two best equalization delays (0 or 1) with a probability of 67.8%, 86.6% et 73.1%, respectively. The OS-CMA obtains the best performance with an affordable complexity, which, thanks to its fast convergence, is always below that of the classical fixed step-size CMA.

15.8.4 CP equalizers for a non-minimum phase channel

We evaluate now the performance of algebraic and iterative solutions for the blind CP and semi-blind CP-MMSE criteria on the non-minimum phase channel of [25, section V],

622 CHAPTER 15 Semi-blind methods for communications

**FIGURE 15.6**

Iterative CM equalizers. Performance of the classical CMA ($\mu = 0.025$), the OS-CMA and the RLS-CMA ($\lambda_{\text{RLS}} = 0.99$) for the SIMO system of section 15.8.3 and a 4-tap equalizer with random Gaussian initialization. (a) Evolution of CM cost function; (b) evolution of equalizer output MSE; (c) normalized histogram of equalization delay. Results are averaged over 1000 signal realizations.

Table 15.3 Average computational cost for convergence of the CM-based iterative equalizers in the experiment of Figs 15.5–15.6

| COST | CMA | OS-CMA | | RLS-CMA |
|-------------------------------|--------|----------|----------|---------|
| | | OS-CMA-1 | OS-CMA-2 | |
| Iterations | 565 | 38 | | 286 |
| Total flops ($\times 10^3$) | 1124.4 | 166.4 | 69.5 | 26.3 |

given by:

$$\begin{aligned}
 H_3(z) = & (-0.033 + 0.014j) + (0.085 - 0.039j)z^{-1} - (0.232 - 0.136j)z^{-2} \\
 & + (0.634 - 0.445j)z^{-3} + (0.070 - 0.233j)z^{-4} \\
 & - (0.027 + 0.071j)z^{-5} - (0.023 + 0.012j)z^{-6}.
 \end{aligned} \tag{15.33}$$

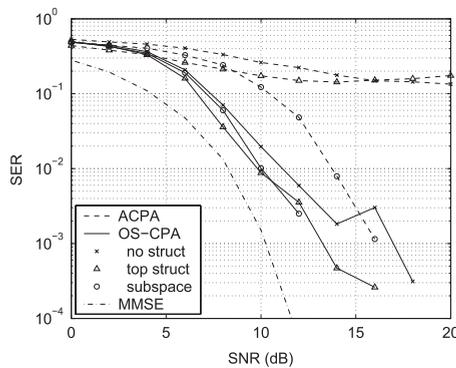


FIGURE 15.7

Blind CP equalization. The OS-CPA is initialized with different ACPA solutions. Channel $H_3(z)$, QPSK input ($q = 4$), $N_d = 100$ symbol periods, $L = 5$ ($L_0 = 3$), 200 MC iterations.

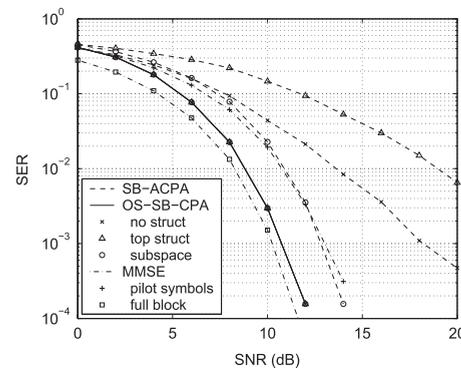


FIGURE 15.8

Semi-blind CP-MMSE equalization, in the same conditions as Fig. 15.7, with $N_t = 10$ pilot symbols and $\lambda = 0.5$. The OS-SB-CPA is initialized with different SB-ACPA solutions.

This 6th-order FIR channel can be perfectly equalized by an FIR filter with $L_0 = 3$ coefficients, but we choose $L = 5$. From a data block of $N_d = 100$ symbols and using several structuring procedures, the algebraic solutions to the blind CP criterion (section 15.6) yield the dotted-line curves shown in Fig. 15.7. These algebraic solutions are then employed to initialize the OS-CPA described in section 15.7.2, generating the results displayed by the solid lines in Fig. 15.7. The gradient-descent iterations refine the algebraic estimates, approaching the MMSE bound.

The performance of the semi-blind CP-MMSE methods are summarized in Fig. 15.8, for the same scenario with $N_t = 10$ pilot symbols and $\lambda = 0.5$. Algebraic estimates are first determined by combining the blind and supervised solutions as explained in section 15.6.3 (dotted lines), and used then to initialize the OS-SB-CPA of section 15.7.2 (solid lines). MMSE performance bounds (dashed lines) are determined by computing the LS solution (15.13) to the MMSE criterion. Two MMSE curves are obtained: using only the pilot sequence, as in a conventional receiver, and using the whole data block (MMSE bound); this bound is obviously unreachable in practice since the whole bandwidth would be used for training.

The advantages of the semi-blind approach are remarkable. In the first place, the performance of algebraic solutions are improved compared to the purely blind case. In the second place, the OS-SB-CPA exhibits identical performances, regardless of initialization, and nearly reaches the MMSE bound. The exploitation of “blind symbols” in addition to the training sequence improves the conventional receiver and almost attains the MMSE bound. Moreover, convergence speed is increased relative to the fully blind case, specially for low SNR, as shown in Fig. 15.9.

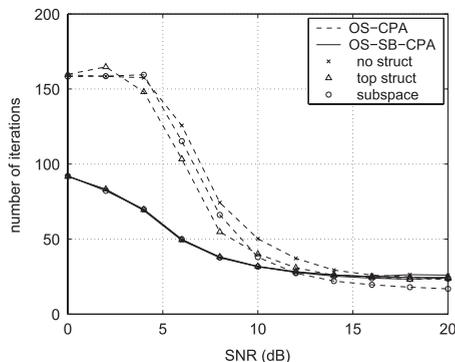


FIGURE 15.9

Average number of iterations for the three initializations of the OS-CPA (blind) and OS-SB-CPA (semi-blind) in the experiment of Figs 15.7–15.8.

15.8.5 Blind CM and semi-blind CM-MMSE equalizers

A zero-mean unit-variance QPSK-modulated input excites the same non-minimum phase channel $H_3(z)$ [Eq. (15.33)] of [25, section V], whose output is corrupted by complex circular additive white Gaussian noise. An FIR filter of length $L = 5$ is used to equalize the channel, aiming at the optimal-MMSE delay ($\tau_{\text{opt}} = 6$ at 20-dB SNR). Bursts of $N_d = 100$ symbols are observed at the channel output, generating a total of $N = 96$ channel-output vectors. We choose $\lambda = 0.5$ and $\mu = 10^{-3}$ for the fixed step-size algorithms. Iterations are stopped when (15.28) is verified, with $\eta = 0.1\mu$. Equalization quality is measured in terms of SER, which is estimated by averaging over 500 independent bursts.

We first compare several fully blind criteria ($N_t = 0$). The algebraic solution of [25, section II-B] is called “DK-top”, and corresponds to the structuring method described in section 15.6.2.4 based over the first elements of the non-structured LS solution to the CM criterion. Iterative solutions are obtained by the fixed step-size CMA (section 15.7.1) with three different initializations: *first-tap* filter, *center-tap* filter and DK-top solution. MMSE receiver and bound curves are also plotted for reference. Figure 15.10 shows that the algebraic solution is only useful as an initial point for the iterative blind receiver, whose performance depends strongly on the initialization used.

In the same scenario, the performance of the fixed step-size SB-CMA (section 15.7.1) is summarized in Fig. 15.11. The algebraic SB-ACMA solution of [74] is also considered; the semi-blind approach to the DK-top method (SB-DK-top) is enabled by the SVD-based procedure described in section 15.6.3. Although the inclusion of training information improves the DK-top method compared with the blind case (Fig. 15.10), the SB-ACMA proves superior and outperforms the conventional receiver for sufficient SNR. Nevertheless, the SB-ACMA can still be improved if used as initialization for the iterative SB-CMA, whose performance becomes practically independent of initialization for a low to moderate SNR. A performance flooring effect is observed for higher SNR. As

15.8 Performance analysis 625

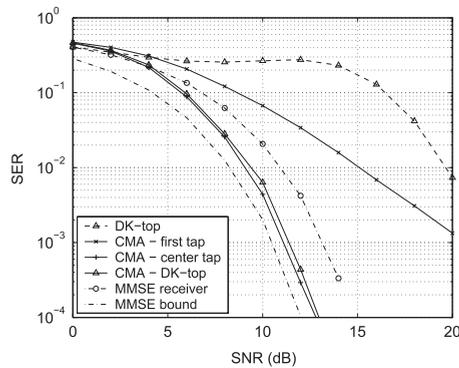


FIGURE 15.10

Blind CM equalization. Channel $H_3(z)$, QPSK input, $N_d = 100$ symbol periods, $L = 5$ ($L_0 = 3$), 500 MC iterations. Solid lines: CMA with fixed step size $\mu = 10^{-3}$ and different initializations.

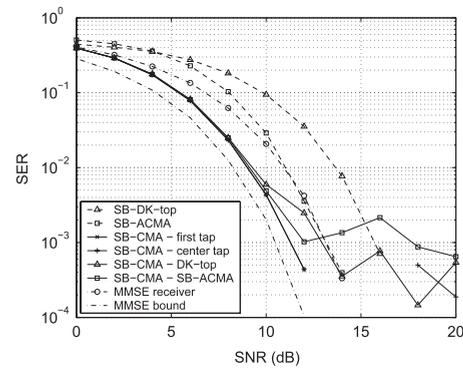


FIGURE 15.11

Semi-blind CM-MMSE equalization, under the conditions of Fig. 15.10 and 10% of pilot symbols. Solid lines: fixed step-size SB-CMA with different initializations.

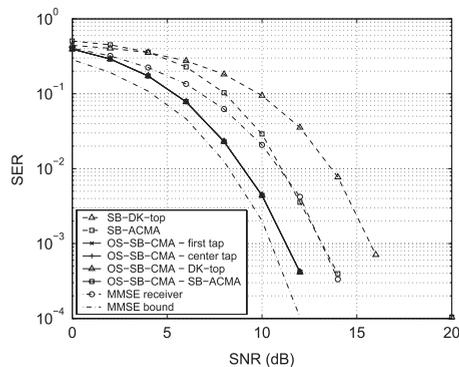


FIGURE 15.12

Semi-blind CM-MMSE equalization, under the conditions of Figs 15.10–15.11. Solid lines: OS-SB-CMA with different initializations.

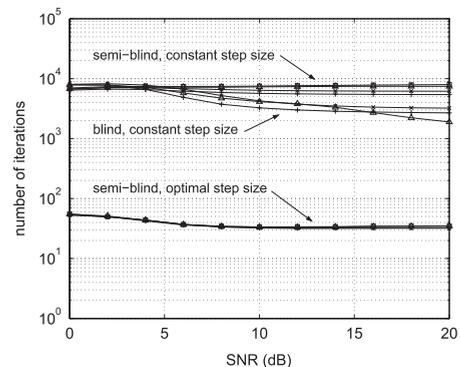


FIGURE 15.13

Average number of iterations for the iterative equalizers in the experiments of Figs 15.10–15.12.

highlighted by Fig. 15.13, the number of iterations required for convergence increases relative to the blind scenario. This increase is probably due to the flattening of the CM cost function when incorporating training data. A similar effect has been observed for the CP criterion in section 15.8.2.

Figures 15.12–15.13 show that the performance of the OS-SB-CMA (section 15.7.2) is virtually independent of initialization, and its iteration count is reduced by around two orders of magnitude relative to the constant step-size techniques. Furthermore,

626 CHAPTER 15 Semi-blind methods for communications

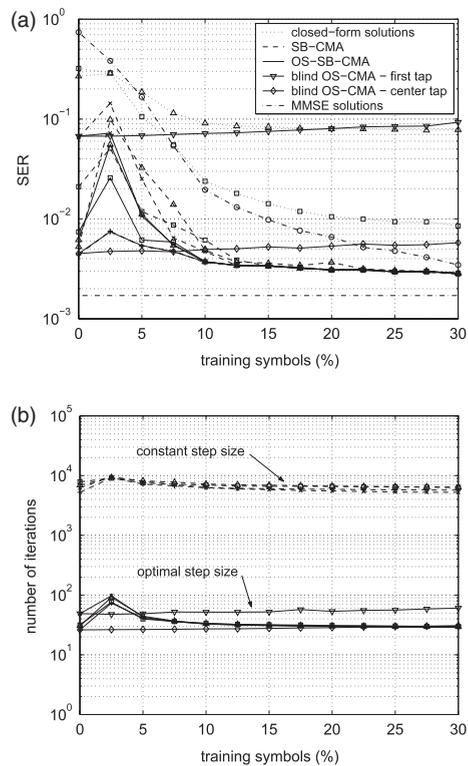


FIGURE 15.14

CM equalization with a variable number of pilot symbols in the transmitted burst, under the conditions of Figs 15.10--15.12. Line markers are as in such figures. (a) SER performance. (b) Average number of iterations for the iterative equalizers.

the performance flooring observed for the fixed step-size SB-CMA at high SNR now disappears.

15.8.6 Influence of pilot-sequence length

Under the same previous conditions, Fig. 15.14 illustrates the performance of semi-blind techniques as a function of the percentage of symbols in the transmitted block used for training, computed as $N_t/N \times 100\%$, for a 10-dB SNR. The OS-CMA only using the “blind symbols” is also tested for two different initializations. The SB-ACMA equalizer only outperforms the conventional receiver for short pilot sequences, and always benefits from gradient-descent iterations. The OS-SB-CMA slightly improves the SB-CMA for short training and for all initializations (\times : first tap; $+$: center tap;

Δ : SB-DK-top; \square : SB-ACMA), while maintaining its reduced complexity over the whole training-length range. For reasonable pilot-sequence sizes, semi-blind methods are capable of attaining the conventional MMSE receiver performance while improving spectral efficiency (decreasing the pilot-sequence length), thus increasing the effective transmission rate. With the appropriate initialization, fully blind processing outperforms the semi-blind methods for short pilot sequences, as if the use of too few training symbols could somehow confuse the blind receiver; a similar effect is observed for sufficient training, where “blind symbols” appear to divert the conventional receiver from its satisfactory solution. Yet the performance of the blind OS-CMA in this scenario depends strongly on initialization, although, as shown in sections 15.8.2–15.8.3, the optimal step-size approach provides certain immunity to local extrema.

A very similar behavior of CP-based equalizers against the pilot-sequence length has been reported in [92].

15.8.7 Influence of the relative weight between blind and supervised criteria

The performance of the semi-blind CP-MMSE methods as a function of confidence parameter λ are illustrated in Fig. 15.15, obtained in the same scenario as in section 15.8.4 with $N_t = 10$ pilot symbols. Equalization results are gradually improved as more weight is laid on the known data. Performance then deteriorates as the blind part of the criterion is neglected and the equalization is left to entirely depend on just a few training symbols; hence the SER increase up to the conventional MMSE receiver level when λ approaches 1. Consequently, this severe increase is not observed with longer training windows. Over a wider range of λ (roughly, in the interval $[0.3, 0.9]$), the influence of initialization on equalization quality and convergence speed of the OS-SB-CPA does not seem to be significant and, for practically all $\lambda \in]0, 1[$, the iterative semi-blind methods outperform the conventional equalizer.

Figure 15.15(b) also shows that, for certain value of confidence parameter ($\lambda \approx 0.7$), the cost-function surface seems best adapted to the execution of the optimal step-size algorithm, so that convergence is obtained in the minimum number of iterations. This optimal value of λ will generally depend on the specific system conditions, the sample size and the SNR.

15.8.8 Comparison between the CM and CP criteria

A final experiment makes a brief 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. Figure 15.16 show that, although the SB-ACPA solution is poorer than SB-ACMA's in this particular scenario, the OS-SB-CPA improves its CM counterpart with half the number of iterations.

628 CHAPTER 15 Semi-blind methods for communications

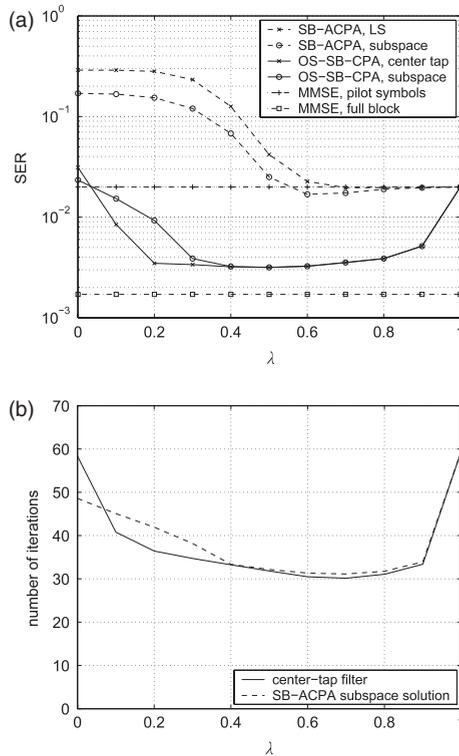


FIGURE 15.15

Impact of confidence parameter λ on the performance of semi-blind CP-MMSE methods. Channel $H_3(z)$, QPSK input ($q = 4$), $N_d = 100$ symbol periods, $N_t = 10$ pilot symbols, $L = 5$ ($L_0 = 3$), SNR = 10 dB, 500 MC runs. (a) SER performance. (b) Average number of iterations for the two initializations of the OS-SB-CPA.

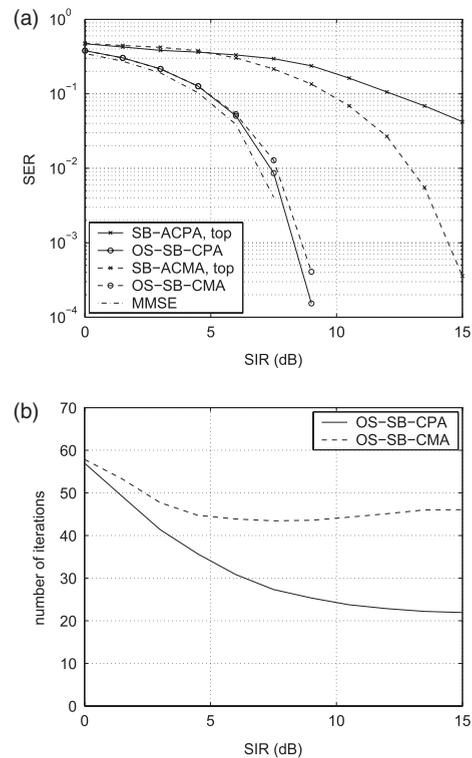


FIGURE 15.16

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. (a) SER performance. (b) Average number of iterations for the iterative equalizers.

15.9 SEMI-BLIND CHANNEL ESTIMATION

We conclude this chapter by discussing the indirect approach to channel equalization. This approach consists of two stages: the channel is estimated in the first stage; equalization is then performed in a second stage. Whereas the direct approach is usually limited to linear equalizers as described in the preceding sections, the indirect approach allows the exploitation nonlinear equalization techniques such as the Viterbi algorithm.

As explained throughout the chapter, the basic semi-blind approach refers to the simultaneous exploitation of known pilot (or training) sequences and blind information.

15.9 Semi-blind channel estimation 629

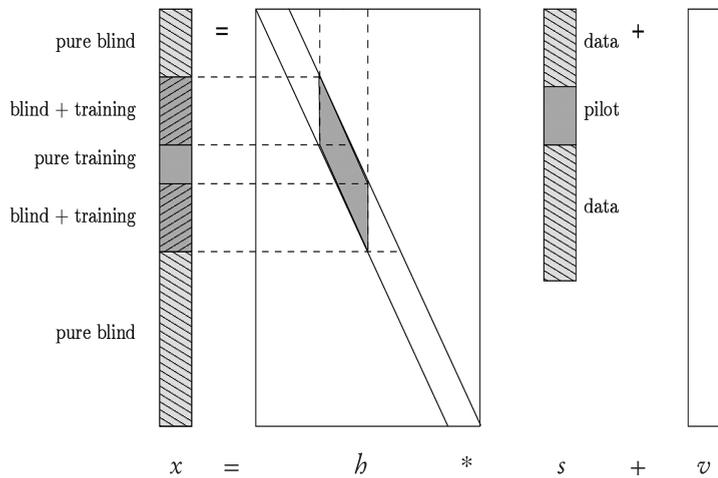


FIGURE 15.17

Received signal structure for a frequency-selective channel.

The blind information can arise from multichannel structures (SIMO, MIMO), possibly obtained after oversampling (as in, e.g., CDMA, where the spreading factor can be viewed as an oversampling factor), or from non-Gaussianity of the transmitted signals (constant modulus, finite alphabet, etc.) The channel can be time-invariant or time-varying, in which case it can be modeled with a Basis Expansion Model (BEM). Due to various forms of memory (delay spread in the time domain, Doppler spread in the frequency domain), the received signal components often contain a mixture of known and unknown symbols. An example of a signal in the time domain passing through a channel with memory is illustrated in Fig. 15.17 [cf. Eq. (15.1)]. Due to the channel memory, some received signal samples generally contain pure training symbols, others contain purely unknown data symbols and still others contain a mixture of both.

In this scenario, the optimal approach is to jointly estimate all unknown quantities (symbols and channel) with an optimal criterion such as maximum likelihood, if possible by incorporating also prior information (if available). Various forms of simpler suboptimal approaches have been proposed in the literature. A class of suboptimal approaches aim to optimize a weighted sum of a training-based and a blind channel estimation criterion; see [16] for an extensive discussion of these techniques for the SIMO case and [53] for the MIMO case.

With perfect Channel State Information at the Receiver (CSIR), no CSI at the Transmitter (CSIT) and iid channel elements, the optimal input signal is a zero-mean spatio-temporally white Gaussian noise. Any deviation from this (side information) will lower the perfect CSIR channel capacity. However, there is usually no CSIR, so that any such deviation may allow channel estimation, leading to an increase in actual channel

630 CHAPTER 15 Semi-blind methods for communications

capacity (see [100] for optimal input distributions in absence of CSIR). Possible forms of side information are enumerated and briefly discussed below:

- Higher-order statistics of data symbols [4,8].
- Finite alphabet (FA) of unknown symbols (see also the algebraic methods presented earlier in the chapter). Note that the use of FA symbols instead of Gaussian inputs constitutes an introduction of side information and thus a reduction of full CSIR channel capacity. The FA can be exploited through iterative channel estimation and data detection; see, e.g., [20,50,67,75,101], or [91] with two-level Kalman filtering. In [63], it is shown that when constraints on the input symbols such as those based on the FA property only leave a discrete ambiguity, then the CRB (which is a local bound) for channel estimation is the same as if the unknown symbols were known.
- Channel coding in unknown symbols, exploited, e.g., through turbo detection and estimation. In [68], a channel estimation CRB is provided when data symbol channel coding is exploited, involving the minimum distance amplification introduced by the channel code. As the SNR increases from low to high values, this CRB moves from the case of the data symbols are unknown and Gaussian to the case of known pilot symbols.
- Partial FA knowledge, such as constant modulus (e.g., the 8-PSK modulation used in the EDGE standard) [37,51,64].
- Some training/pilot symbols, only enough to allow iterative joint data detection/channel estimation to converge.
- Symbol modulus variation pattern, which is a particular form of transmitter-induced cyclostationarity. Some of the techniques proposed to exploit this property lead to wide-sense cyclostationarity [79], without consistency in SNR. The technique proposed in [47], though, is deterministic.
- Space-time coding redundancies through reduced rate linear precoding, introducing subspaces in the transmitted signal covariance, e.g., Alamouti or other orthogonal space-time coding schemes [6,51,56].
- Guard intervals in time or frequency, as in [66,88], or cyclic prefix structure.
- Symbol stream color. In [39], it is shown that colored inputs can be separated if their spectra are linearly independent. Correlation can be introduced by linear convolutive precoding, e.g., in the form of MIMO prefiltering. In [47], an example of low rate precoding appears since the same symbol sequence gets distributed over all transmit antennas; see [53] for a more detailed discussion.
- Known pulse shape, which can be exploited when the received signal is oversampled with respect to the symbol rate (temporal oversampling).
- The spreading codes in CDMA. Direct sequence spectrum spreading (DS-SS) is a special case in which the oversampling factor corresponds to the spreading factor, a sample is called a chip, and a memoryless SIMO prefilter corresponds to an instantaneous multiplication with the spreading code (which can be time-varying in the case of long/aperiodic/pseudo-random codes or time-invariant as in the case of short/periodic/deterministic codes). Of course, CDMA can be combined with oversampling with respect to the chip rate and exploitation of a chip pulse shape.

15.9 Semi-blind channel estimation 631

The use of different spreading codes for different inputs allows for fairly robust blind source separation and channel estimation; see [29,30,38,51].

- Transmitter-induced non-zero mean, also known as superimposed training. Besides the use of time multiplexed (TM) pilots as in the semi-blind techniques presented in this chapter, a recent twist (which is actually not so recent) on the training paradigm is the appearance of superimposed (SI, also called embedded) pilots. SI pilots are actually classical in CDMA standards, which use a pilot signal, sometimes combined with TM pilots. In [101], SI pilot based channel estimates are used to initialize an iterative receiver. In [2], optimization of a mixture of TM and SI pilots is considered. The continuous SI pilots form actually a pilot signal and their large duration leads to quasi-orthogonality with the data. It is found that for large enough and equivalent pilot power, both pilot forms lead to similar performance. Only the channel estimation (CRB) is considered though as performance indicator. In [85], the effect of both types of pilots on the throughput is considered and TM pilots appear to be favored. Indeed, pilots not only allow channel estimation but also influence the data detection. The presence of TM pilots leads to reduced ISI in frequency-selective channels with time-domain transmission. Semi-blind channel estimation and detection with SI pilots is considered in [54]. An important question here is: is orthogonality of pilots and data desirable? The answer may depend on how mixed information (pilot/data) is used and combined.
- Spatial multiplexing schemes that achieve the optimal rate-diversity trade-off typically do not introduce any blind information (other than that provided by Gaussian white inputs) for the channel estimation. In [52], for instance, a previously introduced linear prefiltering scheme was shown to attain this optimal trade-off. Since the prefilter is a MIMO all-pass filter, it leaves the white vector input white. However, perturbations of optimal trade-off achieving schemes can be derived that introduce side information.

Hence, some questions that so far have only been very partially answered are: what is the optimal amount of side information to maximize capacity, as more side information reduces capacity with CSIR but reduces also channel estimation error and hence increases capacity? More importantly, what is the optimal distribution of side information over the various forms? Note that, strictly speaking, blind approaches are based on just exploiting second-order information and/or subspaces. The exploitation of any form of side information mentioned above should be called a semi-blind approach.

Some other research avenues include:

- Multiuser case. In this scenario, the number of unknowns per received sample increases further. Whereas spatial multiplexing is the cooperative case of multi-input, the multi-user case corresponds to the non-cooperative version. Differentiation of users at the level of SOS can be obtained through coloring (e.g., CDMA) as mentioned earlier. In [88], a semi-blind multiuser scenario is considered.
- Non-coherent approaches.
- Channel estimation for the transmitter. Questions that arise here involve not only channel estimation but also its possible quantization and (digital or analog)

632 CHAPTER 15 Semi-blind methods for communications

retransmission. A key issue here also is the degree of reciprocity of the channel or, e.g., its pathwise parameters (direction, delay, Doppler shift, power). Another issue is the effect of sensor array design on channel estimation and reciprocity, e.g., beamspace (beam selection should be reciprocal).

- Description of channel variation in terms of user mobility.
- Bayesian (semi-)blind: from deterministic unknown channels to fading random channels [72].

15.10 SUMMARY, CONCLUSIONS AND OUTLOOK

The present chapter has addressed the problem of channel equalization, which consists in recovering the information emitted through a time-dispersive propagation medium. Source separation and channel equalization can be considered as dual problems whose goal is to unravel, respectively, spatial and temporal mixtures of the source(s). Yet the particularities of digital communication systems allow the design of more specific source recovery techniques, some of which have been presented in this chapter.

Our focus has been on SISO channel equalization. Several semi-blind criteria have easily been defined by combining purely blind criteria based on the finite alphabet of digital signals, such as the CM and CP principles, and the conventional training-based MMSE equalizer. Under certain conditions (essentially, the existence of exact ZF equalizers and input signals adapted to the blind part of the criteria), the global minima of such semi-blind principles can be attained algebraically. These non-iterative solutions are unaffected by the presence of local extrema on the cost-function surface. The algebraic treatment of the CP criterion, resulting in the ACPA equalizer, is similar to that of ACMA, but does not require special modifications to treat binary modulations. Algebraically, the proposed subspace method provides a particular solution to the challenging rank-1 tensor linear combination problem. In our numerical study, this subspace approach proves more robust than other structuring methods, but the blind algebraic solutions offer a low tolerance to noise, particularly for long equalizers. This tolerance can be slightly improved by semi-blind techniques from just a few pilot symbols. The key point limiting performance is probably the SVD-based rank-1 tensor approximation employed to extract the equalizer vector from the estimated symmetric tensor, as described in section 15.6.2.3 and 15.6.3. A refinement of this rank-1 tensor approximation, such as that obtained by the power method [17,41], could alleviate this limitation.

In general, algebraic solutions can only approximate a good equalization setting, and iterative techniques are generally necessary to find the global minimum of the criterion; an iterative method can also be used to refine an algebraic solution. An exact *global* line search technique based on block iterations has been proposed, allowing an optimal algebraic adaptation of the step size at each iteration; this adaptation only involves the roots of a polynomial that can be solved by radicals. The optimal step-size iterative algorithm offers a very fast convergence and, in semi-blind mode, yields equalization results very close to the MMSE bound while increasing the useful transmission rate

and the robustness to the equalizer vector initialization. These benefits have been demonstrated by the numerical experiments presented in the last part of the chapter.

In summary, the impact of local minima and slow convergence typical of blind equalizers can be limited with the incorporation of training sequences, giving rise to semi-blind criteria. To further alleviate these drawbacks, the present chapter has endowed semi-blind criteria with a number of additional strategies:

- judicious initialization with *algebraic solutions*;
- iterative updates operating on signal *blocks* (or bursts);
- *one-dimensional global minimization* (exact line search) with an optimal step size computed algebraically.

These strategies are not exclusive of the equalization principles considered in this chapter (CM, CP), but can also profit other criteria such as the KM [96] (see also Chapter 6) or those of [34].

Avenues of further research could include the following aspects, which have been left aside in our analysis: a theoretical study of the local extrema of the CP criterion; the improvement of the SVD-based technique to recover the equalizer vector; the robust automatic detection of the number of ZF solutions and the optimal equalization delay [97]; the theoretical optimal choice of confidence parameter λ (e.g., based on an asymptotic variance analysis); the evaluation and reduction of carrier-residual effects on CP equalizers [11,12] (although the inclusion of pilot information may already play an important role in their compensation); and an exhaustive comparison, both theoretical and experimental, of other equalization principles with those presented in this chapter. Other challenging open issues in the related topic of semi-blind channel estimation have also been discussed in section 15.9.

References

- [1] A. Belouchrani, J.-F. Cardoso, Maximum likelihood source separation for discrete sources, in: Proc. EUSIPCO-94, VII European Signal Processing Conference, Edinburgh, UK, September 13–16, 1994, pp. 768–771.
- [2] L. Berriche, K. Abed-Meraim, J.-C. Belfiore, Cramer-Rao bounds for MIMO channel estimation, in: Proc. IEEE ICASSP, Montreal, Canada, May 17–21, 2004.
- [3] D.H. Brandwood, A complex gradient operator and its application in adaptive array theory, IEE Proceedings F: Communications Radar and Signal Processing 130 (1983) 11–16.
- [4] J.-F. Cardoso, Blind signal separation: Statistical principles, Proceedings of the IEEE 86 (1998) 2009–2025.
- [5] Y. Chen, T. Le-Ngoc, B. Champagne, C. Xu, Recursive least squares constant modulus algorithm for blind adaptive array, IEEE Transactions on Signal Processing 52 (2004) 1452–1456.
- [6] J. Choi, Equalization and semi-blind channel estimation for space-time block coded signals over a frequency-selective fading channel, IEEE Transactions on Signal Processing 52 (2004) 774–785.
- [7] S. Choi, A. Cichocki, Blind equalisation using approximate maximum likelihood source separation, Electronics Letters 37 (2001) 61–62.
- [8] A. Cichocki, S. Amari, Adaptive Blind Signal and Image Processing: Learning Algorithms and Applications, John Wiley & Sons, Inc., Chichester, UK, 2002.

634 CHAPTER 15 Semi-blind methods for communications

- [9] P. Comon, Analyse en Composantes Indépendantes et identification aveugle, *Traitement du Signal* 7 (1990) 435–450. Numero special non lineaire et non gaussien.
- [10] P. Comon, Block methods for channel identification and source separation, in: *Proc. IEEE Symposium on Adaptive Systems for Signal Processing, Communications and Control*, Lake Louise, Alberta, Canada, October 1–4, 2000, pp. 87–92.
- [11] 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, December 2002.
- [12] P. Comon, Independent component analysis, contrasts, and convolutive mixtures, in: *Proc. 2nd IMA International Conference on Mathematics in Communications*, University of Lancaster, UK, December 16–18, 2002.
- [13] P. Comon, Tensor decompositions: State of the art and applications, in: J.G. McWhirter, I.K. Proudler (Eds.), *Mathematics in Signal Processing V*, Clarendon Press, Oxford, UK, 2002, pp. 1–24.
- [14] P. Comon, Contrasts, independent component analysis, and blind deconvolution, *International Journal of Adaptive Control and Signal Processing* 18 (2004) 225–243 (Special issue on Blind Signal Separation).
- [15] P. Comon, B. Mourrain, Decomposition of quantics in sums of powers of linear forms, *Signal Processing* 53 (1996) 93–107 (Special issue on Higher-Order Statistics).
- [16] E. De Carvalho, D. Slock, Semi-blind methods for FIR multichannel estimation, in: G.B. Giannakis, Y. Hua, P. Stoica, L. Tong (Eds.), *Signal Processing Advances in Wireless and Mobile Communications. Vol. 1: Trends in Channel Estimation and Equalization*, Prentice Hall, Upper Saddle River, NJ, 2001, pp. 211–254. (Ch. 7).
- [17] L. De Lathauwer, P. Comon, et al. Higher-order power method, application in Independent Component Analysis, in: *NOLTA Conference*, vol. 1, Las Vegas, 10–14 December 1995, pp. 91–96.
- [18] J. P. Delmas, P. Comon, Y. Meurisse, Identifiability of BPSK, MSK, and QPSK FIR SISO channels from modified second-order statistics, in: *IEEE Spawc'06*, Cannes, July 2–5 2006.
- [19] J.P. Delmas, Y. Meurisse, P. Comon, Performance limits of alphabet diversities for FIR SISO channel identification, *IEEE Transactions on Signal Processing* 57 (2008) 73–82.
- [20] R. Demo Souza, J. Garcia-Frias, A.M. Haimovich, A semi-blind receiver for iterative data detection and decoding of space-time coded data, in: *Proc. IEEE-COM WCNC Conf.*, 2004.
- [21] Z. Ding, C.R. Johnson, R.A. Kennedy, On the (non)existence of undesirable equilibria of Godard blind equalizers, *IEEE Transactions on Signal Processing* 40 (1992) 2425–2432.
- [22] Z. Ding, R.A. Kennedy, B.D.O. Anderson, C.R. Johnson, Ill-convergence of Godard blind equalizers in data communication systems, *IEEE Transactions on Communications* 39 (1991) 1313–1327.
- [23] Z. Ding, Y. Li, *Blind Equalization and Identification*, Dekker, New York, 2001.
- [24] K. Dogançay, K. Abed-Meraim, 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] K. Dogançay, R.A. Kennedy, Least squares approach to blind channel equalization, *IEEE Transactions on Signal Processing* 47 (1999) 1678–1687.
- [26] D. Donoho, On minimum entropy deconvolution, in: *Applied Time-Series Analysis II*, Academic Press, 1981, pp. 565–609.
- [27] E. Durand, *Solutions numériques des équations algébriques*, vol. I, Masson, Paris, 1960.
- [28] E. Gassiat, F. Gamboa, Source separation when the input sources are discrete or have constant modulus, *IEEE Transactions on Signal Processing* 45 (1997) 3062–3072.
- [29] I. Ghauri, D.T.M. Slock, Blind and semi-blind single user receiver techniques for asynchronous CDMA in multipath channels, in: *Proc. IEEE Globecom*, Sydney, Australia, Nov. 1998.
- [30] I. Ghauri, D.T.M. Slock, MMSE-ZF receiver and blind adaptation for multirate CDMA, in: *Proc. IEEE Vehicular Technology Conference*, Amsterdam, The Netherlands, September 1999.

- [31] D.N. Godard, Self-recovering equalization and carrier tracking in two-dimensional data communication systems, *IEEE Transactions on Communications* 28 (1980) 1867–1875.
- [32] G.H. Golub, C.F. Van Loan, *Matrix Computations*, 3rd ed., The John Hopkins University Press, Baltimore, MD, 1996.
- [33] O. Grellier, P. Comon, Blind equalization and source separation with MSK inputs, in: *Proc. SPIE Conference on Advances in Signal Processing*, San Diego, CA, July 19–24, 1998, pp. 26–34.
- [34] O. Grellier, P. Comon, Blind separation of discrete sources, *IEEE Signal Processing Letters* 5 (1998) 212–214.
- [35] O. Grellier, P. Comon, Closed-form equalization, in: *Proc. SPAWC-99, 2nd IEEE Workshop on Signal Processing Advances in Wireless Communications*, Annapolis, MD, May 9–12, 1999, pp. 219–222.
- [36] O. Grellier, P. Comon, B. Mourrain, P. Trebuchet, Analytical blind channel identification, *IEEE Transactions on Signal Processing* 50 (2002) 2196–2207.
- [37] M.A.S. Hassan, B.S. Sharif, W. Woo, S. Jimaa, Semiblind estimation of time varying STFBC-OFDM channels using Kalman filter and CMA, in: *Proc. IEEE ISCC*, 2004.
- [38] B. Hochwald, T.L. Marzetta, C.B. Papadias, A transmitter diversity scheme for wideband CDMA systems based on space-time spreading, *IEEE Journal of Selected Areas in Communications* 19 (2001).
- [39] Y. Hua, J.K. Tugnait, Blind identifiability of FIR-MIMO systems with colored input using second-order statistics, *IEEE Signal Processing Letters* (2000).
- [40] C.R. Johnson, P. Schniter, I. Fijalkow, L. Tong, et al., The core of FSE-CMA behavior theory, in: S.S. Haykin (Ed.), *Unsupervised Adaptive Filtering, Vol. II: Blind Deconvolution*, John Wiley & Sons, New York, 2000, pp. 13–112. (Ch. 2).
- [41] E. Kofidis, P.A. Regalia, On the best rank-1 approximation of higher-order supersymmetric tensors, *SIAM Journal on Matrix Analysis and Applications* 23 (2002) 863–884.
- [42] M. Kristensson, B. Ottersten, D.T.M. Slock, Blind subspace identification of a BPSK communication channel, in: *Proc. 30th Asilomar Conf. Sig. Syst. Comp.*, Pacific Grove, CA, 1996.
- [43] A.M. Kuzminskiy, L. Féty, P. Foster, S. Mayrargue, Regularized semi-blind estimation of spatio-temporal filter coefficients for mobile radio communications, in: *Proc. XVIème Colloque GRETSI*, Grenoble, France, September 15–19, 1997, pp. 127–130.
- [44] J.L. Lacoume, P.O. Amblard, P. Comon, *Statistiques d'ordre supérieur pour le traitement du signal*, Collection Sciences de l'Ingénieur, Masson, 1997. téléchargeable gratuitement à www.i3s.unice.fr/~comon/livreSOS.html.
- [45] C. Lanczos, *Applied Analysis*, Dover, New York, 1988.
- [46] J. Lebrun, P. Comon, A linear algebra approach to systems of polynomial equations with application to digital communications, in: *Proc. EUSIPCO-2004, XII European Signal Processing Conference*, Vienna, Austria, September 6–10, 2004.
- [47] G. Leus, P. Vandaele, M. Moonen, Deterministic blind modulation-induced source separation for digital wireless communications, *IEEE Transactions on Signal Processing* 49 (2001) 219–227.
- [48] T.H. Li, K. Mbarek, A blind equalizer for nonstationary discrete-valued signals, *IEEE Transactions on Signal Processing* 45 (1997) 247–254. Special issue on communications.
- [49] Y. Li, Z. Ding, Global convergence of fractionally spaced Godard (CMA) adaptive equalizers, *IEEE Transactions on Signal Processing* 44 (1996) 818–826.
- [50] Y. Li, L. Yang, Semi-blind MIMO channel identification based on error adjustment, in: *Proc. IEEE Conf. Neural Netw. Sig. Proc.*, Nanjing, December 2003.
- [51] Z. Liu, G.B. Giannakis, S. Barbarossa, A. Scaglione, Transmit-antennae space-time block coding for generalized OFDM in the presence of unknown multipath, *IEEE Journal of Selected Areas in Communications* 19 (2001) 1352–1364.
- [52] A. Medles, D. Slock, Achieving the optimal diversity-vs-multiplexing tradeoff for MIMO flat channels with QAM space-time spreading and dfe equalization, *IEEE Transactions on Information Theory* 52 (2006).

636 CHAPTER 15 Semi-blind methods for communications

- [53] A. Medles, D.T.M. Slock, Blind and semiblind mimo channel estimation, in: H. Bölcskei, D. Gesbert, C. Papadias, A.-J. van der Veen (Eds.), *Space-Time Wireless Systems, From Array Processing to MIMO Communications*, Cambridge University Press, 2006.
- [54] X. Meng, J.K. Tugnait, Semi-blind channel estimation and detection using superimposed training, in: *Proc. IEEE ICASSP, Montreal, Canada, May 17–21, 2004*.
- [55] E. Moulines, P. Duhamel, J.-F. Cardoso, S. Mayrargue, Subspace methods for the blind identification of multichannel FIR filters, *IEEE Transactions on Signal Processing* 43 (1995) 516–525.
- [56] H.J. Pérez-Iglesias, J.A. García-Naya, A. Dapena, L. Castedo, V. Zarzoso, Blind channel identification in Alamouti coded systems: A comparative study of eigendecomposition methods in indoor transmissions at 2.4 GHz, *European Transactions on Telecommunications* 19 (2008) 751–759.
- [57] W.H. Press, S.A. Teukolsky, W.T. Vetterling, B.P. Flannery, *Numerical Recipes in C. The Art of Scientific Computing*, 2nd ed., Cambridge University Press, Cambridge, UK, 1992.
- [58] J.G. Proakis, *Digital Communications*, 4th ed., McGraw-Hill, New York, 2000.
- [59] P.A. Regalia, On the equivalence between the Godard and Shalvi-Weinstein schemes of blind equalization, *Signal Processing* 73 (1999) 185–190.
- [60] P.A. Regalia, Blind deconvolution and source separation, in: J.G. McWhirter, I.K. Proudler (Eds.), *Mathematics in Signal Processing V*, Clarendon Press, Oxford, UK, 2002, pp. 25–35.
- [61] P.A. Regalia, A finite-interval constant modulus algorithm, in: *Proc. ICASSP, 27th International Conference on Acoustics, Speech and Signal Processing*, vol. III, Orlando, FL, May 13–17, 2002, pp. 2285–2288.
- [62] L. Rota, P. Comon, Blind equalizers based on polynomial criteria, in: *Proc. ICASSP-2004, 29th International Conference on Acoustics, Speech and Signal Processing*, vol. IV, Montreal, Canada, May 17–21, 2004, pp. 441–444.
- [63] B. Sadler, M. Kozik, T. Moore, Bounds on bearing and symbol estimation with side information, *IEEE Transactions on Signal Processing* 49 (2001) 822–834.
- [64] A. Safavi, K. Abed-Meraim, Blind channel identification robust to order overestimation: A constant modulus approach, in: *Proc. IEEE ICASSP, 2003*.
- [65] Y. Sato, A method of self-recovering equalization for multi-level amplitude modulation, *IEEE Transactions on Communications* 23 (1975) 679–682.
- [66] A. Scaglione, G. Giannakis, S. Barbarossa, Redundant filterbank precoders and equalizers, part II: Blind channel estimation, synchronization, and direct equalization, *IEEE Transactions on Signal Processing* 47 (1999) 2007–2022.
- [67] A. Scaglione, A. Vosoughi, Turbo estimation of channel and symbols in precoded MIMO systems, in: *Proc. IEEE ICASSP, Montreal, Canada, May 17–21, 2004*.
- [68] A. Scherb, V. Kühn, K.-D. Kammeyer, Cramer-Rao lower bound for semiblind channel estimation with respect to coded and uncoded finite-alphabet signals, in: *Proc. 38th Asilomar Conf. Signals, Systems & Computers*, Pacific Grove, CA, November 2004.
- [69] O. Shalvi, E. Weinstein, New criteria for blind deconvolution of nonminimum phase systems (channels), *IEEE Transactions on Information Theory* 36 (1990) 312–321.
- [70] M.T.M. Silva, M. Gerken, M.D. Miranda, An accelerated constant modulus algorithm for space-time blind equalization, in: *Proc. EUSIPCO-2004, XII European Signal Processing Conference*, Vienna, Austria, September 6–10, 2004, pp. 1853–1856.
- [71] 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.
- [72] D.T.M. Slock, Bayesian blind and semiblind channel estimation, in: *Proc. IEEE-SP Workshop on Sensor Array and Multichannel Signal Processing*, Spain, July 2004.
- [73] R. Steele, L. Hanzo (Eds.), *Mobile Radio Communications*, 2nd ed., John Wiley & Sons, New York, 1999.
- [74] A.L. Swindlehurst, A semi-blind algebraic constant modulus algorithm, in: *Proc. ICASSP-2004, 29th International Conference on Acoustics, Speech and Signal Processing*, vol. IV, Montreal, Canada, May 17–21, 2004, pp. 445–448.

- [75] S. Talwar, A. Paulraj, Blind separation of synchronous co-channel digital signals using an antenna array — Part II: Performance analysis, *IEEE Transactions on Signal Processing* 45 (1997) 706–718.
- [76] S. Talwar, M. Viberg, A. Paulraj, Blind estimation of multiple co-channel digital signals arriving at an antenna array: Part I, algorithms, *IEEE Transactions on Signal Processing* (1996) 1184–1197.
- [77] L. Tong, G. Xu, T. Kailath, Blind identification and equalization based on second-order statistics: A time domain approach, *IEEE Transactions on Information Theory* 40 (1994) 340–349.
- [78] J.R. Treichler, B.G. Agee, A new approach to multipath correction of constant modulus signals, *IEEE Transactions on Acoustics, Speech and Signal Processing* 31 (1983) 459–472.
- [79] M. Tsatsanis, G. Giannakis, Transmitter induced cyclostationarity for blind channel estimation, *IEEE Transactions on Signal Processing* 45 (1997) 1785–1794.
- [80] J.K. Tugnait, L. Tong, Z. Ding, Single-user channel estimation and equalization, *IEEE Signal Processing Magazine* 17 (2000) 16–28.
- [81] A.-J. van der Veen, Analytical method for blind binary signal separation, *IEEE Transactions on Signal Processing* 45 (1997) 1078–1082.
- [82] A.-J. van der Veen, A. Paulraj, An analytical constant modulus algorithm, *IEEE Transactions on Signal Processing* 44 (1996) 1136–1155.
- [83] A.-J. van der Veen, S. Talwar, A. Paulraj, Blind estimation of multiple digital signals transmitted over FIR channels, *IEEE Signal Processing Letters* 2 (1995) 99–102.
- [84] A.-J. van der Veen, S. Talwar, A. Paulraj, A subspace approach to blind space-time signal processing for wireless communication systems, *IEEE Transactions on Signal Processing* 45 (1997) 173–190.
- [85] A. Vosoughi, A. Scaglione, On the effect of channel estimation error with superimposed training upon information rates, in: *Proc. IEEE International Symposium on Information Theory, Montreal, Canada, June–July 2004*.
- [86] T. Wigren, Avoiding ill-convergence of finite dimensional blind adaptation schemes excited by discrete symbol sequences, *Signal Processing* 62 (1997) 121–162. Elsevier.
- [87] C. Xu, J. Li, A batch processing constant modulus algorithm, *IEEE Communications Letters* 8 (2004) 582–584.
- [88] W. Y., T.-S. Ng, A semi-blind channel estimation method for multiuser multiantenna OFDM systems, *IEEE Transactions on Signal Processing* 52 (2004).
- [89] H. Yang, On-line blind equalization via on-line blind separation, *Signal Processing* 68 (1998) 271–281.
- [90] D. Yellin, B. Porat, Blind identification of FIR systems excited by discrete-alphabet inputs, *IEEE Transactions on Signal Processing* 41 (1993) 1331–1339.
- [91] J. Yue, K.J. Kim, T. Reid, J. Gibson, Joint semi-blind channel estimation and data detection for MIMO-OFDM systems, in: *Proc. IEEE-CAS Symposium on Emerging Technologies: Mobile and Wireless Communications, Shanghai, China, May 2004*.
- [92] V. Zarzoso, P. Comon, Blind and semi-blind equalization based on the constant power criterion, *IEEE Transactions on Signal Processing* 53 (2005) 4363–4375.
- [93] V. Zarzoso, P. Comon, Blind channel equalization with algebraic optimal step size, in: *EUSIPCO-2005, XIII European Signal Processing Conference, Antalya, Turkey, September 4–8, 2005*.
- [94] V. Zarzoso, P. Comon, Semi-blind constant modulus equalization with optimal step size, in: *ICASSP-2005, 30th International Conference on Acoustics, Speech and Signal Processing, vol. III, Philadelphia, PA, March 18–23, 2005, pp. 577–580*.
- [95] V. Zarzoso, P. Comon, Optimal step-size constant modulus algorithm, *IEEE Transactions on Communications* 56 (2008) 10–13.
- [96] V. Zarzoso, P. Comon, Robust independent component analysis by iterative maximization of the kurtosis contrast with algebraic optimal step size, *IEEE Transactions on Neural Networks*, (2009) (in press) (<http://www.i3s.unice.fr/~mh/RR/2009/RR-09.02-V.ZARZOSO.pdf>).
- [97] V. Zarzoso, A.K. Nandi, Blind MIMO equalization with optimum delay using independent component analysis, *International Journal of Adaptive Control and Signal Processing* 18 (2004) 245–263. Special issue on Blind Signal Separation.

638 **CHAPTER 15** Semi-blind methods for communications

- [98] V. Zarzoso, A.K. Nandi, Exploiting non-Gaussianity in blind identification and equalization of MIMO FIR channels, *IEE Proceedings — Vision, Image and Signal Processing* 151 (2004) 69–75. Special issue on Non-Linear and Non-Gaussian Signal Processing.
- [99] Y. Zhang, S.A. Kassam, Blind separation and equalization using fractional sampling of digital communications signals, *Signal Processing* 81 (2001) 2591–2608.
- [100] L. Zheng, D. Tse, Communication on the Grassmann manifold: A geometric approach to the non-coherent multiple antenna channel, *IEEE Transactions on Information Theory* 48 (2002) 359–383.
- [101] H. Zhu, B. Farhang-Boroujeny, C. Schlegel, Pilot embedding for joint channel estimation and data detection in MIMO communication systems, *IEEE Transactions on Signal Processing* 7 (2003) 30–32.

Author Queries

Q1 (Page 601)

Please check whether the new page reference as per the current page number given here is correct or not.

Q2 (Page 637)

Please check the author names given in Ref. [88].
