Signal Processing 143 (2018) 298–302

Contents lists available at ScienceDirect

Signal Processing

journal homepage: www.elsevier.com/locate/sigpro

Short communication

A theoretical note on the generalized ML optimality of constant modulus equalizers

F. Benedetto[a,*], G. Giunta[a], L. Vandendorpe[b]

[a] University of ROMA TRE, via Vito Volterra 62, 00146, Rome, Italy
[b] Université catholique de Louvain, B-1348, Louvain-la-Neuve, Belgium

ARTICLE INFO

Article history:
Received 27 March 2017
Revised 5 September 2017
Accepted 16 September 2017
Available online 19 September 2017

Index Terms:
Blind equalization
Constant-modulus algorithm
Maximum likelihood principle
Optimality
Wireless communications

ABSTRACT

In this work, we derive the optimum equalizer according to the General Maximum Likelihood (GML) principle and show the optimality of the constant-modulus algorithm (CMA) according to the GML principle. This reported discussion illustrates why CMA works well and hence is so popular. Moreover, we show that the minimization of normalized variance algorithm (MNVA) previously introduced by the authors, as much as the asymptotically equivalent Kurtosis maximization algorithm and "Rayleigh-ness" test criteria, are asymptotically optimum according to the GML criterion.

1. Introduction

Blind (non-data-aided) methods employing pre- and post-correlation techniques have been widespread investigated in many communication systems [1–3]. In particular, blind adaptive multi-user detection based on the constant modulus algorithm (CMA) has received a lot of attention [4–5]. CMA is also a promising technique for peak-to-average-power-ratio (PAPR) reduction in Orthogonal Frequency Division Multiplexing (OFDM)-based communications, such as in Long Term Evolution (LTE) systems [6]. In CMA equalizers, we are dealing with the maximization problem of a likelihood objective function, depending on a number of unknowns to be estimated (the equalizer coefficients), but also characterized by one more random parameter of the system under observation (i.e. an amplitude factor compensating for the gain of the whole chain). In the seminal book of Van Trees [7], the problem of the maximum likelihood (ML) estimation of unknown parameters in the presence of further random parameters was effectively addressed. According to Van Trees' approach, three solving ways must be followed when the ambiguity function depends on one parameter. If we may assume a statistical model for that parameter, the likelihood function is averaged over the parameter distribution to define an unconditioned ML detection and estimation. Otherwise, if no statistical information is available (i.e. no distribution can be a priori assumed for the unknown parameter), a possible solution is to consider the ML solution conditioned to a given set of values for the unknown parameters (conditioned ML criterion). The operating problem still remains since one should wonder which values would have to be considered as the most likely. Conversely, the only a posteriori knowledge can be taken into account by considering the ML estimate for the unknown parameter to be substituted in the ambiguity function to be maximized. This latter solution, called generalized ML (GML) criterion [7], is equivalent to define a particular ML detection and estimation conditioned to the ML estimate of the parameter. In our case, the GML criterion reduces to the ML criterion itself. As reported in the following of our work, the maximization of the likelihood function should bring to the same formal solution, since both the criteria convey to the same principal scheme. We know that this result is not general at all (e.g. see the sufficient conditions and the counterexamples in hypothesis testing [8] or the discussion on its optimality for signal detection in [9]), while such an academic discussion is out of the scope of the paper.

A new blind equalizer has been introduced in literature namely MNVA, i.e. minimization of normalized variance algorithm [10]. Authors showed that the MNVA is asymptotically equivalent to the well-known CMA, as well as to the Kurtosis maximization algorithm (KMA) [11], because they pursue the same "Ricianity" criterion already employed in the "Rayleigh-ness" test for spread-spectrum code acquisition [12]. Moreover, in [10] authors showed the asymptotic equivalence between all the CMA schemes (e.g. normalized CMA (N-CMA) [13,14], the signed error CMA (SE-CMA)

* Corresponding author.
E-mail address: francesco.benedetto@uniroma3.it (F. Benedetto).

http://dx.doi.org/10.1016/j.sigpro.2017.09.018

[15–17], the signed regressor CMA (SR-CMA) [18,19], and the signed error signed regressor CMA (SS-CMA)) [20,21].

In this note, we theoretically explain why they are optimum according to the GML criterion. We derive the optimum equalizer according to the GML principle, illustrating why CMA is so popular. The remainder of this correspondence is organized as follows. In Section 2, we first derive the GML optimality for the simpler case of coherent data-aided equalization. The derivation of the optimality of the blind equalizers will be approached in the second half of Section 2, before our conclusions finally drawn in Section 3.

2. Post-despreading application of GML criterion

The ML criterion (both conventional and general version) requires to statistically modelling the received signal for any set of values of the parameters to be estimated. The model may vary in significant way whether the operating unknowns' sets are near or far from the optimum ones. In practice, either *local* or *global* models have to be assumed, respectively within the *neighbourhood* of the actual maximum or far away. Nevertheless, the goal of a maximization procedure consists in reaching the absolute maximum of the objective function. In such a case, the absolute maximum of the optimality function happens while the *local* model asymptotically applies. Therefore, we emphasize that the statistical model employed in the following maximization is based on the validity of the assumption of the *quasi-synchronous condition* (i.e. the variance of the output of the matched filter can be considered as a *constant*). In a non-equalized environment, the above asymptotic approximation may be not more valid. In practice, the *local* objective function may apply after a pre-equalizer has brought the system in the vicinity of the actual maximum. Our paper is focused on non-coherent blind equalization. Nevertheless, to derive the GML optimality of all the constant modulus equalizers (CMA, MNVA, KMA, and others), it can be useful to start from the simpler academic case of coherent data-aided equalization. The derivation of the optimality of the blind equalizers will be approached just further, modifying the basic scheme of the coherent (non-blind) estimation.

2.1. Data aided equalization

We consider here K independent successive symbols with rate T. Because we aim to model the *in-sync* situation to implement *local* ML detection, the perfect orthogonality between signal waveforms will be assumed after the equalizer. In fact, our objective is just to model the system in the vicinity of the optimum operating point. On the other side, we aim to estimate the equalizer's coefficients to reach the ML condition. Therefore, we neglect in our analysis the windowing effects between contiguous symbols, so that the simpler *one-shot* receiver can be equivalently considered. In fact, the equalizer perfectly compensates for distortions due to the channel and the receiver becomes a one-shot receiver based on symbol-detection in independent noise. The principle scheme of the (locally optimum) ML coherent receiver of the k-th symbol, under the *in-sync* condition, is depicted in Fig. 1a. For sake of simplicity, all the signals are expressed in terms of their *representative vectors*. The signal $r(t)$ represented by the vector $\boldsymbol{r}$ is received from the channel (with unknown attenuation factor). The signal $\boldsymbol{r}$ includes both additive noise and interference effects. It is correlated with $\boldsymbol{h}_k$, which accounts for spreading, chip shaping, and data symbol b_k, with $k=1, \ldots, K$. The symbols are here assumed to belong to a (unitary) constant-modulus constellation (e.g. QPSK). The signal $g(t)$, denoted by $\boldsymbol{g}$, is then (near perfectly) equalized (and synchronized) by the linear array with the equalizer's vector coefficients $\boldsymbol{w}$. The result of the scalar product $x_k=\boldsymbol{w}^H\times\boldsymbol{g}$ is multiplied by a factor A, to compensate for the unknown channel attenuation and the equalizer's gain. The operator $\{\cdot\}^H$ means complex conjugate transpose. In such a way, we can exploit the (unitary) constant-modulus constraint of the possible data symbols (i.e. $|b_k|=1$) by assuming a unitary output in the ideal absence of noise and interference. Let us assume that the multi-user interference is represented by the sequence $\{e_k\}$. This sequence can be considered as an uncorrelated sequence because, as happens in many communication networks, the effects of multi-users interference can be assumed to be greater than the thermal noise component (that can be discarded in our discussion without loss of generality).

Therefore, the effect of noise and interference on the output sample can be regarded as an additive random variable affecting the constant output (say $y_k=1+e_k$). The output sample y_k finally goes to the ML computation device. Its maximization brings to a feedback tuning of the system's parameters to be set, namely the vector of equalizer's coefficients $\boldsymbol{w}$ and the scalar gain factor A. In practice, the more likely set of $\{\boldsymbol{w}, A\}$ that produces the more constant output y_k around unity is chosen. The likelihood function of the output signal y_k for $k=1, \ldots, K$ requires expressing its joint probability density function (PDF). The output y_k can be asymptotically modelled as a non-zero complex Gaussian random variable, due to the central limit's theorem. It assumes a near constant value under an ideal synchronization condition (*in-sync*). We are then assuming the asymptotic Gaussianity of the random series $\boldsymbol{y}=\{y_k\}$, characterized by the following mean and variance (non-varying under the *in-sync* case):

$$E[y_k] = E[1+e_k] = b_k^* b_k = 1, \operatorname{var}[y_k] = \operatorname{var}[e_k] = \sigma_e^2. \tag{1}$$

The PDF of a K-dimensional complex Gaussian random independent and identically distributed (i.i.d.) series $\boldsymbol{y}=\{y_k\}$ conditioned to the parameters' values $\{\boldsymbol{w}, A\}$ can be expressed as:

$$p(\mathbf{y}|\mathbf{w}, A) = \frac{1}{\left(\pi\sigma_e^2\right)^K} \exp\left\{-\frac{1}{\sigma_e^2}\sum_{k=1}^{K}|y_k-1|^2\right\}. \tag{2}$$

According to the scheme in Fig. 1a, the variable y_k depends on both the parameters $\boldsymbol{w}$ and A, while the variable x_k is function of the equalizer's coefficients $\boldsymbol{w}$ only. For sake of compactness of the algebraic expressions, such explicit dependence will be omitted in the following derivation. The ML detector looks for the more likely value of $\{\boldsymbol{w}, A\}$ that satisfies the K-dimensional complex Gaussian statistical model of $\{y_k\}$, whose mean and variance have been defined before. In fact, it maximizes the following function:

$$\begin{aligned}\{\mathbf{w}_{ML}, A_{ML}\} &= \{\mathbf{w}, A\} : MAX_{\mathbf{w},A}\{p(\mathbf{y})|\mathbf{w}, A\} \\ &= MIN_{\mathbf{w},A}\left\{\frac{1}{\sigma_e^2}\sum_{k=1}^{K}|y_k-1|^2\right\} \\ &= MIN_{\mathbf{w},A}\left\{\frac{1}{\sigma_e^2}\sum_{k=1}^{K}|A\cdot x_k-1|^2\right\}\end{aligned} \tag{3}$$

The relationship (3) is a function of the unknowns $\boldsymbol{w}$ and A. Let us first impose that the partial derivative on the unknown A is zero to obtain its ML estimate A_{ML}, then obtaining:

$$A\cdot\sum_{k=1}^{K}|x_k|^2 = 1\cdot\sum_{k=1}^{K}x_k^* \text{ that is } A_{ML} = \sum_{k=1}^{K}x_k^* \Big/ \sum_{k=1}^{K}|x_k|^2 \tag{4}$$

As a consequence, we can rewrite the ML equation, formerly (3) function of $\boldsymbol{w}$ and A, for $A=A_{ML}$ to be finally maximized versus $\boldsymbol{w}$ according to the GML criterion [6] as follows:

$$\mathbf{w}_{ML} : MIN_{\mathbf{w}}\{\,p(\mathbf{y})|\mathbf{w},\ A_{ML}\} = MIN_{\mathbf{w}}\left\{\sum_{k=1}^{K}|A_{ML}-x_k-1|^2\right\}$$