New Algorithms for Blind Equalization:
The Constant Norm Algorithm Family

Alban Goupil and Jacques Palicot

Abstract—In this paper, a new, efficient class of blind equalization algorithms is proposed for use in high-order, two-dimensional digital communication systems. We have called this family: the Constant Norm Algorithms (CNA). This family is derived in the context of Bussgang techniques. Therefore, the resulting algorithms are very simple. We show that some well-known blind algorithms such as "Sato's algorithm" or the Constant Modulus Algorithm (CMA) are particular cases in our CNA family. In addition, from this class, a new cost function, named Constant sQuare Algorithm (CQA), is derived, which is well designed for QAM. It results in a lower algorithm noise without increasing the complexity. Another advantage of this class lies in the possibility of creating new norms by combining several existing norms in order to benefit from the advantages of each original norm. For example, we present the norm resulting from the combination of the two algorithms, CMA and CQA. Moreover, we highlight that, with regard to the excess mean-square error performance, there is an optimal norm for each constellation, i.e., each modulation, in order to equalize it blindly.

Index Terms—Adaptive equalizers, algorithms, equalizers.

I. INTRODUCTION

IN any wireless communication system, most of the introduced intersymbol interference (ISI) is caused by the so-called *multipath propagation phenomenon*, which consists of the reception of multiple signals originating from a single transmitted signal. If mobility is considered, the above situation is more complex. In addition, , it should be emphasized that the channel characteristics are time varying, mainly due to the movement of the mobile terminal. Moreover, if the transmission is carried out in bursts, and assuming that the channel may change significantly during the interburst period, we conclude that the equalization system in question must be able to equalize the channel successfully within each separate burst. Classically, a training period is inserted in order to start the convergence of the equalizer. However, bandwidth is a precious resource. Therefore, the need for training sequence reduction is imperative. Conventional training-based channel estimation methods are thus inappropriate, especially when the channel span is large, as in high-rate applications. To improve the overall throughput of a transmission system, we should avoid the use of a training period. In other words, we should perform blind equalization on the receiver side. Thus, the whole system is not burdened by the training overhead and a considerable saving in bandwidth results is achieved. Over the last two decades, a lot of work has been carried out on blind equalization schemes [1], [2]. In this paper, we discuss both direct equalization (deconvolution) and equalization based on channel estimation. In fact, our class could be applied for both situations. The first work dealing with the blind equalization problem dates back to Sato [3] in 1975. This was followed by works by Godard [4] in 1980, Treichler and Agee [5] in 1983, and Bellini [6] in 1986. Each of these algorithms was proven to be optimal for a specific constellation and therefore generated a lot of noise—or even worse, diverged—when applied to another constellation.

The purpose of our work was to decrease the noise of the algorithm when it is applied to another constellation, as when we apply Constant Modulus Algorithm (CMA) to QAM for example. Some authors apply a CMA to each component in quadrature to decrease this noise. This algorithm is known as MMA [7]. But, in our point of view, this solution was not sufficiently general to be applied to any constellation. It is effectively clear that CMA-type algorithms are constellation dependent. This is obvious when considering the constant value involved in the algorithm. In fact, we aimed to find a class of cost functions, which could be applied to any constellation. This is exactly what we obtained with the Constant Norm Algorithm class. This family is a simple generalization of the CMA where the modulus is substituted by any norm over the plane $\mathbb{C}$.

The paper is organized as follows. In Section II, the problem is formulated. In Section III, the new Constant Norm Algorithm (CNA) family is described, and we highlight the combination of norms in Subsection III-B before describing particular cases of CNA in Subsection III-C. We then analyze the performance in Section IV, due to the computation of the excess mean-square error (EMSE). In addition, we have derived EMSE equations for both CMA and CQA cases. A discussion about the possibility of obtaining an optimal algorithm in the CNA family for each modulation takes place in Section V in light of the analysis of Section IV. Section VI provides simulation results to support this discussion. Finally, the paper is concluded in Section VII.

II. PROBLEM FORMULATION

Some notations are needed to aid the problem formulation. We consider that all signals are complex. The complex conjugate of z is noted $\bar{z}$. The uppercase letters such as S denote streams of complex numbers whose kth element is noted s_k. Multiplication of streams means the convolution of them. Bold capital letters, such as $\mathbf{S}$, is used for row or column vector of a predefined length.
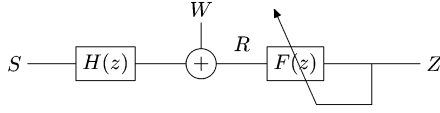
Fig. 1. Blind equalization scheme.

In blind equalization, we consider that the emitter sends a stream S of symbols drawn from a known finite alphabet, called a *constellation*. The source S is supposed to be composed of independent and identically distributed (i.i.d.) symbols, whose distribution is uniform over the constellation. The effects of the environment on the message are modeled by a channel given by an FIR filter H and by white Gaussian noise W. This system is described in Fig. 1. The aim of blind equalization is to recover the source stream S only from the received sequence R. In its simpler form, the blind equalizer is also a finite-inpulse-response (FIR) filter denoted F whose output is Z, thus equal to $Z = FR = F(HS + W)$.

The idea behind the blind Bussgang equalization is to minimize, through the coefficients of the equalizer filter F, a certain function depending on the output Z. In fact, in a noiseless case, it is possible to show that if Z satisfies particular criteria (e.g., the independence of its components), then the channel is perfectly equalized up to some inherent indeterminate such as delay or amplitude [8]. To describe this kind of blind equalization precisely, the Bussgang algorithms try to solve the following optimization problem:

$$F_{\text{opt}} = \underset{F}{\arg\min} \, \mathcal{J}(Z) \quad \text{with} \quad \mathcal{J} = \mathbb{E}J. \tag{1}$$

This kind of optimization problem allows the blind equalization problem to be split into two smaller problems: the search for the proper cost function J and the implementation of the optimization algorithm. Due to this separation, we could consider that the minimization is carried out online with a simple stochastic gradient algorithm. If we consider $\mathbf{F}_k$ to be the coefficient vector of the equalizer F, the equalizer's coefficient updating algorithm is

$$\mathbf{F}_{k+1} = \mathbf{F}_k + \mu \bar{\mathbf{R}}_k \phi(z_k) \quad \text{with} \quad \phi(z) = \partial_z J(z) \tag{2}$$

where $\mathbf{R}_k = [r_k, r_{k-1}, \ldots, r_{k-l}]$ is a vector containing the samples of the received sequence and z_k the equalizer's output at time k, and ∂_z is a well-chosen derivation over the complex z (see [9] for details on the subject). The parameter μ is a step governing the speed of convergence and the level of the steady-state period. In order to keep the description simple, this step is kept constant during the optimization process.

Algorithms of the form (2) need approximately the same amount of computations. Indeed, the difference in complexity between two of them takes place in the computation of their related ϕ functions. In fact, the computational burden of one update (2) is mainly concerned by the vector addition and their scalar multiplication by $\mu\phi(z_k)$; this scalar is computed only once by iteration.

There are several cost functions J in the blind equalization literature. The simplest one is the decision-directed algorithm whose cost function could be written

$$J(z) = \frac{1}{2}|z - \text{dec}(z)|^2 \quad \text{and} \quad \phi(z) = z - \text{dec}(z) \tag{3}$$

where $\text{dec}(z)$ denotes the decision related to the constellation from the output z. However, due to its poor performance, many authors consider it as nonblind. The first blind criterion is Sato's cost function [3], which is the pioneering function in the blind equalization field:

$$J(z) = \frac{1}{2}|z - \gamma \, \text{sgn}(z)|^2 \quad \text{and} \quad \phi(z) = z - \gamma \, \text{sgn}(z) \tag{4}$$

where γ is a constant that depends on the constellation. This function shows that the blind equalization could be carried out by comparing the output of the filter with a reference that is not as selective as the decision. Many other cost functions were presented in the literature such as in the paper by Benveniste *et al.* [8] or such as Picchi and Prati's Stop-and-Go [10], whose algorithms consist of a special combination of the above-described decision-directed algorithm (3) and Sato's algorithm (4). However, one of the most frequently studied and used blind cost functions is the CMA presented in [4] by Godard. The present paper discusses a generalization of this function, and this will be presented in the next section.

III. Cost Functions

A. Constant Norm Algorithms

The idea behind the CNA is the comparison of a function of the equalizer's output with a constant. The functions considered in this paper belong to the wide class of norm functions. These norms are defined on $\mathbb{C}$ viewed as the plane $\mathbb{R}^2$, and should respect relations (5) to (7), and are noted $n(z)$.

$$\forall z \in \mathbb{C}, \quad n(z) \geq 0 \tag{5}$$

$$\forall z \in \mathbb{C}, \quad \forall \alpha \in \mathbb{R} \quad n(\alpha z) = |\alpha| n(z) \tag{6}$$

$$\forall a, b \in \mathbb{C}^2, \quad n(a + b) \leq n(a) + n(b). \tag{7}$$

Several examples of the norm, which is useful for blind equalization, will be given below. Once given a norm, the cost function of CNA for blind equalization could be written as

$$\mathcal{J}(z) = \frac{1}{ab}\mathbb{E}|n^a(z) - \gamma|^b \tag{8}$$

where γ is a constant that depends only on the constellation. Commonly, this constant is noted R for the CMA in the literature. The value of this constant will be derived thereafter. The two parameters a and b are two degrees of freedom of the algorithm and put there to keep the generality. However, the most frequently used values are $a = b = 2$.

Even without a thorough analysis of the algorithm, the constant γ could be easily derived. In fact, in the noiseless case the equalizer should converge to the optimum, which in the case of

the double infinite equalizer filter is the inverse of the channel filter.

As this is true for most channel filters, it should be true for the simplest one: the one-tap channel. We then derive the value of γ in this case and, as this constant cannot depend on the channel, we shall obtain its global value. So, we consider that the channel is constant: $H(z) = h_0$, and use a one-tap equalizer $F(z) = f_0$. The output of the equalizer is then $Z = \alpha S$ with $\alpha = h_0 f_0$, the combined system. To further simplify, we assume that α is real. The problem involves only real values as the range of the function J is $\mathbb{R}$. If the cost function is good enough, the optimization problem (1) should lead to $|\alpha| = 1$. Without loss of generality, we can assume that α is strictly positive, and we obtain the relation

$$\left.\frac{\partial \mathcal{J}(\alpha S)}{\partial \alpha}\right|_{\alpha=1} = 0. \tag{9}$$

However, from (8) and (6), the cost function is

$$\mathcal{J}(\alpha S) = \frac{1}{ab}\mathbb{E}|n^a(\alpha S) - \gamma|^b \tag{10}$$

$$= \frac{1}{ab}\mathbb{E}|\alpha^a n^a(S) - \gamma|^b \tag{11}$$

which gives

$$\left.\frac{\partial \mathcal{J}(\alpha S)}{\partial \alpha}\right|_{\alpha=1} = \mathbb{E}\{(n^a(S) - \gamma)^{b-1} n^a(S)\}. \tag{12}$$

The right-hand side of (12) should be zero, and developing the term in brackets gives a relation that γ should verify, that is

$$\gamma \in \left\{ \text{roots of } \sum_{k=0}^{b-1} \binom{b-1}{k} \mathbb{E}\left[n^{a(b-k)}(S)\right] \gamma^k \right\} \cap \mathbb{R}^+. \tag{13}$$

It should be noted that, as we planned, the constant γ depends only on the statistics of the source symbols S, and not on other system parameters. The last formula appears complex. However, in the most frequently used case given by $b = 2$, the (12) becomes $\mathbb{E}[(n^a(S) - \gamma)n^a(S)] = \mathbb{E}[n^{2a}(S) - \gamma n^a(S)] = \mathbb{E}n^{2a}(S) - \gamma \mathbb{E}n^a(S)$. The value of γ can be derived, in this case, by resolving (9); that gives

$$\gamma = \frac{\mathbb{E}n^{2a}(S)}{\mathbb{E}n^a(S)}. \tag{14}$$

It can be seen that this derivation does not imply that the cost function is good enough to ensure that the equalizer will converge to the solution of (1). But the value of γ is nevertheless a necessary condition.

B. Combination of Two Norms

One simple way to build a new norm is to combine two existing norms. In fact, if n_1 and n_2 are two norms, then $\lambda n_1 + (1 - \lambda)n_2$ is also a norm for $0 \leq \lambda \leq 1$. The parameter λ is a proportion given the dominant norm. This kind of combination allows new algorithms to be built whose proportion parameter λ varies given the overall environment. The aim of a such a combination is to obtain a new norm, which combines the advantages of the two norms. This last point will be clearer with the example of Subsection III-C-6). Due to parameter λ, the obtained algorithm can track the channel variations and can use the best
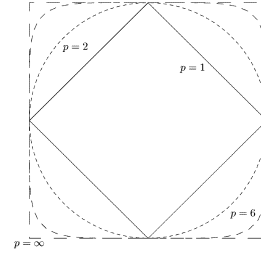

Fig. 2.   Unit ball for different norms.

adapted norm for the considered real-time situation. A second advantage of this combination is that the resulting norm could better match the constellation of the modulation.

C. Particular Cases of CNA

The above description of the CNA is general enough to investigate several particular cases. The most important parameter is the choice of norm. That is why, thereafter, we consider the parameters a and b to be set to 2. In this section, we shall see five choices for the norm and a combination of two of them. One of the choices is the classical CMA, which belongs to the CNA family.

1) Sato's Algorithm: When amplitude modulations are concerned, the real line $\mathbb{R}$ takes the place of the plane $\mathbb{C}$. Moreover, the absolute value is the only one norm over $\mathbb{R}$ up to scaling. Thus, the CNA family consists, in this case, of all cost functions with the form

$$J(z) = \frac{1}{ab}||z|^a - \gamma|^b \quad \text{where } z \in \mathbb{R}. \tag{15}$$

If the parameters a and b equal 1 and 2, respectively, we get back the Sato's algorithm (4) because $|z| = z\,\text{sgn}(z)$ in $\mathbb{R}$.

2) CNA-p, p-Norm Family: When presenting norms, one of the first examples is the class of p-norm whose expression is given in the plane $\mathbb{C}$, by

$$||z||_p = \sqrt[p]{|\text{Re}\,z|^p + |\text{Im}\,z|^p}. \tag{16}$$

The unit ball (loci of points satisfying $n(z) = 1$) of some of these norms are drawn in Fig. 2, where $p = 1, 2,$ and 6. The case $p = \infty$ is special and will be presented later. However, we see that for p, increasing the unit ball seems to tend to the unit square.

3) CNA-2, Constant Modulus Algorithm: The CMA was developed by Godard [4] for constant modulus modulations [like phase shift keying (PSK)]. This is one of the most widely studied algorithms. The cost function can be written as

$$\mathcal{J}(z) = \frac{1}{ab}\mathbb{E}||z|^a - \gamma|^b. \tag{17}$$

With our previous simplification ($a = b = 2$), the algorithm takes the simple form

$$\mathbf{F}_{k+1} = \mathbf{F}_k - \mu(|z_k|^2 - \gamma)z_k\bar{\mathbf{R}}_k \tag{18}$$

and the constant γ is chosen so that the inverse of the channel is a minimum of CMA in a noiseless environment and for a doubly infinite length equalizer. This is then found to be equal to $\mathbb{E}|a|^4/\mathbb{E}|a|^2$ in accordance with (14). The fact that this cost

function, which was conceived for the PSK modulation, also works for QAM is quite surprising. However, in this case, the descent algorithm (18) generates a significant amount of noise.

4) CNA-6: We shall see in the optimization process that the CNA-6 plays a central role. That is why we develop it further here. The cost function is in the case $a = b = 2$

$$\mathcal{J}(z) = \frac{1}{2}\mathbb{E}\left|\|z\|_6^2 - \gamma\right|^2. \tag{19}$$

This cost function leads to the following algorithm:

$$\mathbf{F}_{k+1} = \mathbf{F}_k - \mu\left(\|z_k\|_6^2 - \gamma\right)\frac{(\operatorname{Re} z_k)^5 + i(\operatorname{Im} z_k)^5}{\|z_k\|_6^4}\bar{\mathbf{R}}_k. \tag{20}$$

This algorithm should be implemented carefully because of the division by $\|z\|_6$ when z is near the origin. However, this numerical difficulty could be avoided thanks to the series expansion of the fraction. CNA-p cost functions for $p \neq 2$ always give the same kind of algorithm and attention should also be paid when z is close to the origin.

5) CNA-∞, Constant Square Algorithm: On studying the QAM constellation, the first impression is its square aspect. Whilst the CMA was first designed for the PSK modulation, it also performs quite well on QAM. However, the idea of the square aspect of the QAM leads to the use of a norm with a square aspect: the infinite norm, which is simply defined by

$$\|z\|_\infty = \max(|\operatorname{Re} z|, |\operatorname{Im} z|). \tag{21}$$

Strictly speaking, this norm does not belong to the p-norm family, but can be roughly considered as the norm obtained when p tends to ∞. Notice also that theoretical results derived for the infinite norm cannot be obtained by a limit process and need their own derivation most of the time. That is why the algorithm derived from it has its own name: Constant sQuare Algorithm

$$\mathbf{F}_{k+1} = \mathbf{F}_k - \mu\left(\|z\|_\infty^2 - \gamma\right)F(z_k)\bar{\mathbf{R}}_k \tag{22}$$

where the function $F(\cdot)$ will be obtained in Section IV-B. This algorithm shows the appropriateness between a constellation and a norm. Consider Fig. 3, which illustrates the QAM-16 constellation and the unit ball of the modulus and the infinite norm. We shall show that the steady-state performances of the algorithms depend roughly on the mean distance ℓ (defined by the norm considered) between the symbols of the constellation and the ball of radius related to γ. In the case of the CQA, this mean distance is smaller than that of the CMA. So, the performance level of the CQA for QAM should be better, as the simulation results show.

The MMA algorithm of [7] is updated in just the same way as here. In fact, the CQA update at each iteration considering
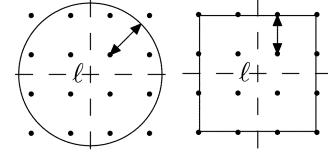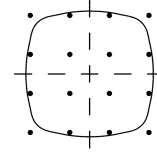

Fig. 3. Principle of CMA and CQA.

Fig. 4. Principle of CMA/CQA combination for $\lambda = 0.5$.

only the in-phase or the in-quadrature part of the filter output though the MMA consider both parts at the same time. Thus, the performances of the MMA and the CQA should be the same in average if we consider that the step μ of the CQA is twice that of the MMA. However, the MMA does not fit into our family of algorithms.

6) CMA/CQA Dynamical Combination: As discussed above, we can combine two norms to build another one. We could obtain a good algorithm by using the classical modulus of the CMA and the infinite norm of CQA. One of the disadvantages of CQA is its phase sensitivity. The CMA however does not suffer from this. But the CQA, as shown below, generates less noise during the steady state. A combination of norms is a solution for getting the best of both algorithms. We would like to start the convergence with the CMA and to use the CQA once the phase has been recovered sufficiently. Consider the following norm:

$$\|z\|_\lambda = \alpha\lambda\|z\|_\infty + (1 - \lambda)|z| \tag{23}$$

where λ gives the proportion between the two norms and α is a fixed proportion added to have a new degree of freedom for the system design. When $\lambda = 0$, the norm is the modulus and the algorithm is the CMA, and when $\lambda = 1$, we recover the CQA. For example, the scheme of the unit ball and the 16-QAM constellation is plotted in Fig. 4 with $\lambda = 0.5$. As above, for ease of presentation, we only consider the cost function of the algorithm where $a = b = 2$:

$$\mathcal{J}(z) = \frac{1}{4}\mathbb{E}\left|\|z\|_\lambda^2 - \gamma(\lambda)\right|^2. \tag{24}$$

Naturally, the parameter γ becomes dependent on the parameter λ. However, its expression is simply a rational fraction whose computation gives (25), shown at the bottom of the page.

As the norm depends on a new parameter, the algorithm is separated into two parts. The first part, as usual, updates the

$$\gamma(\lambda) = \frac{\sum_{i=0}^4 \binom{4}{i}\alpha^i\lambda^i(1-\lambda)^{4-i}\mathbb{E}\left[\|a\|_\infty^i|a|^{4-i}\right]}{\alpha^2\lambda^2\mathbb{E}\|a\|_\infty^2 + (1-\lambda)^2\mathbb{E}|a|^2 + 2\alpha\lambda(1-\lambda)\mathbb{E}[\|a\|_\infty|a|]}. \tag{25}$$

equalizer's coefficients $\mathbf{F}_k$, and the second part of the algorithm updates the parameter λ automatically, as follows:

$$
\mathbf{F}_{k+1} = \mathbf{F}_k - \mu_F \left(\|z_k\|_{\lambda_k}^2 - \gamma(\lambda_k) \right) \|z_k\|_{\lambda_k}
$$
$$
\times (\alpha \lambda_k F(z_k) + (1-\lambda_k)\mathrm{sgn}(z_k))\bar{\mathbf{R}}_k \quad (26)
$$
$$
\lambda_{k+1} = \lambda_k + \mu_\lambda \left(\|z_k\|_{\lambda_k}^2 - \gamma(\lambda_k) \right)
$$
$$
\times (2\|z_k\|_{\lambda_k}(\alpha\|z_k\|_\infty - |z_k|) - R'(\lambda_k)) \quad (27)
$$

where for complex number $\mathrm{sgn}\, z = z/|z|$.

The part of the algorithm involving the update rule of λ is designed using the optimization problem similar to (1), as follows:

$$
\lambda_{\mathrm{opt}} = \underset{\lambda}{\mathrm{argmin}}\, \mathcal{J}(Z). \quad (28)
$$

And thus, as for F, λ is updated using the stochastic gradient descent algorithm

$$
\lambda_{k+1} = \lambda_k - \mu_\lambda \partial_\lambda J(z_k). \quad (29)
$$

After some algebra, we obtain (27) from (29).

IV. PERFORMANCE ANALYSIS

A. Generalities

Performance analysis of the blind equalization algorithm requires two studies. The first concerns the transient phase and the second deals with the steady-state noise level. In the following, we shall focus on the latter aspect. The difficulty here is to get general enough results, which are applicable to the CNA family. To this aim, we shall derive a general formula given the excess mean square error for the stochastic gradient algorithm in the Bussgang blind equalization field considering noiseless channels. We shall then apply it to several algorithms in the CNA family. Using results published in [12] based on the "energy conservation relation" of Sayed [13], the EMSE of the Bussgang algorithm could be approximated by

$$
\mathrm{EMSE} \approx \mu P_R \frac{\mathbb{E}|\phi(z_{\mathrm{opt}})|^2}{2\mathbb{E}\partial_{z\bar{z}}\{\mathrm{Re}\phi(z)(\overline{z - z_{\mathrm{opt}}})|_{z_{\mathrm{opt}}}\}}. \quad (30)
$$

In this approximation, z_{opt} corresponds to the output of the optimal filter with respect to the minimization of (1), and ϕ is defined in (2). The derivation $\partial_{z\bar{z}}$ should be understood from Brandwood's operator [9]. The approximation (30) is better when μP_R is small enough where $P_R = \mathbb{E}\|\mathbf{R}_k\|^2$ is the power of the input of the equalizer.

Application of (30) for our purpose is based on two assumptions: the first is the noiseless environment, and for the second, we assume that the equalizer has converged to the optimal equalizer with regard to the ZF criterion.

In order to apply this general formula derived from the geometrical interpretation of the stochastic gradient algorithm to the CNA family cost functions, we then only need to compute the derivative of the denominator of the fraction (30). These applications are given below.

B. Applications

Under the assumptions made to derive (30), we could derive EMSE for different algorithms in order to classify them. Even if this criterion is more or less adequate, it should be suitable enough to provide us with some insights. In Section V, we shall see that the computation of EMSE provides a nice way of obtaining the best norm for the CNA, given the constellation of a modulation. First of all, in the following subsections, we describe the computation of this criterion for a different member of the CNA class. In the particular case of the CMA, (30) gives the same results as those found in the literature.

Thereafter, the operator $\mathbb{E}f(a)$ deals with the mean of f over the constellation points a of the stream S of the source.

1) For CNA: Let $n(\cdot)$ be a norm regular enough to not be concerned by the problem with its derivative. Then, the general EMSE expression is, using (30), found in (31), shown at the bottom of the page, where the norm and its derivatives are taken at points a of the constellation.

2) For CNA-p: If the norm used is a p-norm, then we could compute precisely the EMSE given by (31). In this case, and for $p \neq \infty$, the value of the derivatives are

$$
|\partial_z n|^2 = \left(\frac{\|a\|_p}{\|a\|_{p-1}} \right)^{2p-2} \quad (32)
$$
$$
\partial_{z\bar{z}} n = \frac{1-p}{\|a\|_p^p} \left(\|a\|_{p-1}^{p-1} + \|a\|_p \|a\|_{p-2}^{p-2} \right). \quad (33)
$$

What is interesting about these complex relations is that we can use them to optimize the parametrization of the norm in order to make it fit the modulation we want to use. We shall see this use below for QAM-16.

3) For CMA, Same as CNA-2: In the case of the CMA, for ϕ and $\mathrm{Re}\phi(z)(\overline{z-a}) = g(z)$, we obtained the expression

$$
\phi(z) = (|z|^2 - \gamma) \quad (34)
$$
$$
g(z) = (|z|^2 - \gamma)(2|z|^2 - z\bar{a} - \bar{z}a). \quad (35)
$$

The expression of the derivative of $g(z)$ is then

$$
\partial_{z\bar{z}} g(z)|_{z=a} = 4|a|^2 - 2R. \quad (36)
$$

Substitution of (34) and (36) into (30) gives the following approximation of the EMSE of the CMA:

$$
\mathrm{EMSE} \approx \mu P_R \frac{\mathbb{E}|a|^6 - 2\gamma\mathbb{E}|a|^4 + \gamma^2\mathbb{E}|a|^2}{4\mathbb{E}|a|^2 - 2R}. \quad (37)
$$

This relation was already found in [14]. But the derivation carried out there was specific to the CMA and not based on a

$$
\mathrm{EMSE} \approx \mu P_R \times \frac{\mathbb{E}n^6|\partial_{\bar{z}}n|^2 - 2R\mathbb{E}n^4|\partial_{\bar{z}}n|^2 + R^2\mathbb{E}n^2|\partial_{\bar{z}}n|^2}{2\mathbb{E}n^3\partial_{z\bar{z}}n - 2R\mathbb{E}n\partial_{z\bar{z}}n + 6\mathbb{E}n^2|\partial_z n|^2 - 2R\mathbb{E}|\partial_z n|^2} \quad (31)
$$

general relation such as (30), which confirms that (30) runs correctly.

4) For CNA-6: As the results given in Section IV-B-2) seem to be complex, the particular case of the CNA-6 is developed thereafter. In order to simplify the expression of the EMSE, let x and y be the real and imaginary parts of the symbols a of the constellation and n be the norm-6. Once all the computations are carried out, the expression of the EMSE of the CNA-6 is given by (38), shown at the bottom of the page.

5) For CQA Same as CNA-∞: The computation of (30) needs some derivatives of J. In the case of the CQA, this function is not well behaved according to the derivation at some point because of the use of absolute value involved. However, the computation of the EMSE can be carried out due to (30) by splitting the $\mathbb{C}$ plane into domains, where ϕ is derivable. For the CQA algorithm, we get

$$\phi(z) = \frac{1}{2}\left(\|z\|_\infty^2 - \gamma\right)\|z\|_\infty F(z) \tag{39}$$

$$F(z) = \begin{cases} \operatorname{Re} z, & \text{if } \operatorname{Re} z > \operatorname{Im} z \\ \operatorname{Im} z, & \text{otherwise} \end{cases}. \tag{40}$$

Once all the computations are carried out, the expression of the EMSE gives

$$\text{EMSE} \approx \mu P_R \frac{\mathbb{E}\|a\|_\infty^6 - 2R\mathbb{E}\|a\|_\infty^4 + \gamma^2\mathbb{E}\|a\|_\infty^2}{6\mathbb{E}\|a\|_\infty^2 - 2R}. \tag{41}$$

A direct computation of both values shows that for a QAM-16, the EMSE of the CQA is less than that of the CMA.

6) For CMA/CQA Combination: As seen before in Section III-C-6), the combination of two norms could be advantageous in order to get the best of both. In this case, the EMSE could also be computed and will depend on the parameter λ.

However, the result is not necessarily of a great interest. What is important is that, when the parameter λ is close to 1, the EMSE is close to the CQA algorithm's noise power; and when λ is close to 0, the EMSE of the CQA is close to the EMSE value of the CMA.

On the one hand, the CMA is insensitive to the phase residue and the CQA is sensitive to it. On the other hand the CMA generates more noise during the steady-state than the CQA. Thus, the combination between both algorithms permits the introduction of a tradeoff between phase sensitivity and noise power.

V. For Each Modulation, Its Own Norm

A. Generalities

As we already shown in the previous sections, the optimal norm, in the EMSE sense, depends on the constellation of the modulation. We recall here that the blind equalization studied in
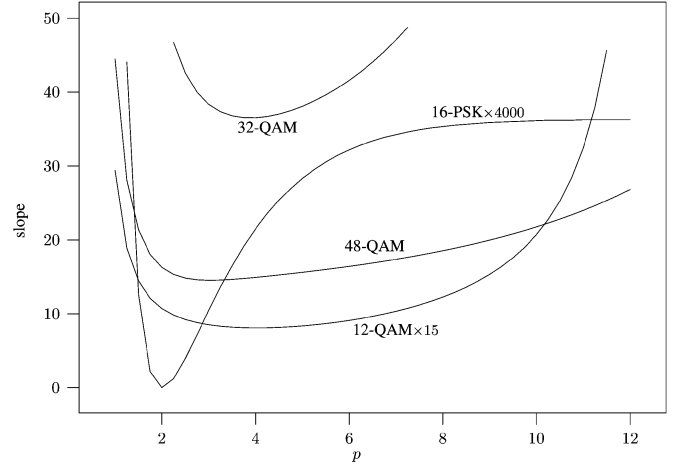

Fig. 5. Excess MSE slope w.r.t. p of CNA-p for a "nonsquare" constellation.

this paper should be interpreted in the collaborative sense. This implies that the receiver knows some statistics of the transmitted signal such as the constellation. In this section, we intend to show that there is an optimal norm for blind equalizing a specific modulation. It is clear from the description of Section III that we obtain a large family of cost functions. Our aim is to find the one best adapted for a specific mapping. As can be seen in Fig. 2, the unit ball of different norms can match a particular constellation. In the following subsection, we shall provide some examples of matched norms for the particular mapping in the EMSE sense. Moreover, we shall once again find some well-known results, such as Sato's algorithm for one-dimensional modulation and CMA for PSK modulation.

B. Examples

1) ASK Modulation: Amplitude shift keying (ASK) modulations are, by definition, one-dimensional modulation. Only one carrier is modulated by several different amplitude levels. Consequently, we should find a norm for the real situation. As explained in Subsection III-C-1), the only norm on $\mathbb{R}$ is the absolute value. For ASK modulations, the CNA family reduces to the cost functions given by (15), and the choice is restricted to the parameters a and b. We find back Sato's algorithm when $a = 1$ and $b = 2$.

2) 2^n-PSK Modulation: PSK modulations, whatever their constellation size, have their points on a circle. In this last situation, the best-adapted norm is the norm 2. This is true because the EMSE of the CMA in a noiseless environment and with PSK modulations is zero. In Fig. 5, the slope of the EMSE from (31)–(33) is plotted with regard to the value p of the p-norm. This slope is defined by $\text{EMSE}/\mu P_R$ and does not depend on the step size nor on the input power. Therefore, the slope is a good characterization of the algorithm's noise. Effectively for 16-PSK, the minimum is attained for $p = 2$, i.e., the CMA.

$$\text{EMSE} \approx \mu P_R \times \frac{\mathbb{E}(n^2 - \gamma)^2(x^{10} + y^{10})/n^8}{\mathbb{E}[(4R - 2n^2)(x^{10} + y^{10}) + 5n^6(n^2 - \gamma)(x^4 + y^4)]/n^{10}}. \tag{38}$$
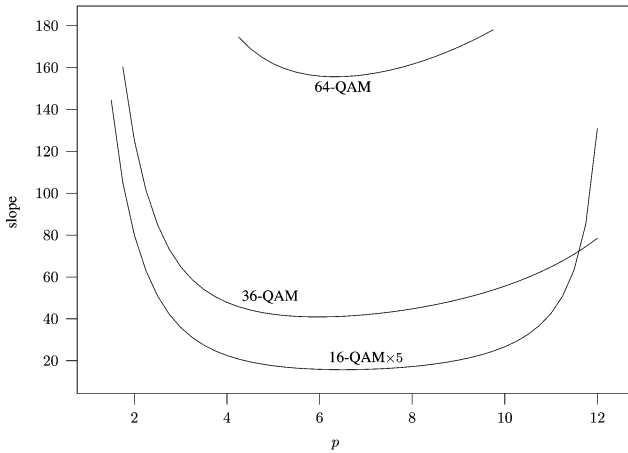
Fig. 6. Excess MSE slope w.r.t. p of CNA-p for a "square" QAM.

Fig. 7. CNA performance on Proakis 1 channel.

3) 2^N-*QAM Modulation:* In Figs. 5 and 6, the slope of the EMSE of CNA-p is plotted for several QAM constellations. They are split into two classes: a "square" constellation in Fig. 6 and "nonsquare" in Fig. 5. It can be seen that some curves are scaled on the Y axis. The EMSE for the CQA is not drawn on these figures. But notice that it is not correct to extrapolate its value from the curves given, because, as mentioned before, the infinite norm is a member of the p-norm family by convention only.

The examples of "square" constellations are 16-QAM, 36-QAM, and 64-QAM. In these three examples, the optimal p-norm is roughly $p = 6$. In this case, the algorithm seems to be equivalent to the CQA as shown in Fig. 7. In this figure, the EMSE of the CQA and the CNA-6 observed are at the same level, while their starting periods also match.

Naturally, the slope of the EMSE becomes greater as the size of the constellation increases. Indeed, that means that the noise of the algorithm growths with this constellation's size.

To illustrate that the optimal p-norm depends on the shape of the constellation, the slope of three "nonsquare" QAMs is depicted in Fig. 5. Two of them, 32 and 12-QAM, correspond to "square" QAMs where the four symbols in the corner are removed. As the shape of this kind of modulation is no longer square, the optimal p of the p-norm tends to decrease to $p = 4$. This result is confirmed with the 48-QAM, where 16 symbols in the corner of a 64-QAM are removed. Then, the resulting constellation is more "round" than the previous one and the optimal p decreases to 3.

VI. SIMULATION RESULTS

(The different algorithms seen before were tested by simulations. In all the simulations, the channel used was the one described in Proakis [11] whose coefficients are given by the vector $[4, -5, 7, -21, -50, 72, 36, 21, 3, 7]$.

The criterion used to compare the algorithms is widely used in this type of case and is called the ISI measure. This is computed by considering the combined system $C(z) = H(z)F(z)$ with the formula

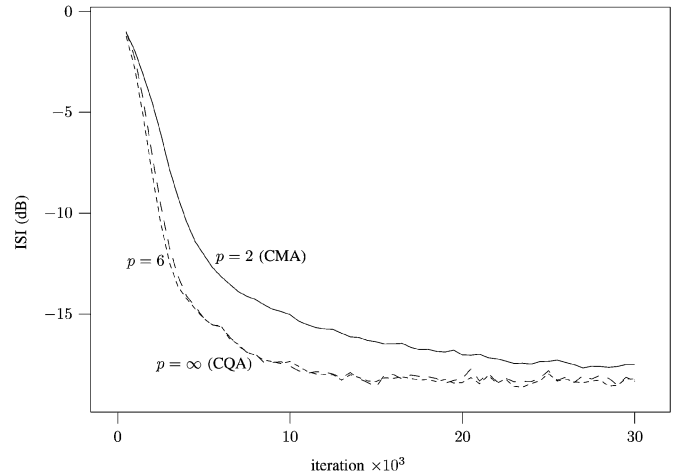$$\text{ISI measure} = \frac{\sum_k |c_k| - \max_k |c_k|}{\max_k |c_k|}. \tag{42}$$

It can be seen that the measure tends to 0 when the equalizer tends to the inverse of the channel. The measure is given on the results in decibels. The ISI's insensitivity to phase recovery allows us to compare the proposed algorithms fairly because some of them recover the phase. The mean-square error (MSE) does not have this property and is not used thereafter. For the noise level considered (40 dB), this measure has the same behavior as the classical MSE measure. So, there is no need to compare the trajectories of the algorithms for these two measurements. In all simulations, the modulation of the source chosen is QAM-16, as it is for this modulation, since the generalization of the CMA was conceived with this modulation in mind.

A. CMA/CQA/CNA-6 Comparison

The CMA, the CQA, and also the CNA-6 were simulated on the channel described above with a signal to noise ratio of 40 dB. The steps of the stochastic gradient algorithms were all chosen in order to have the same ISI measure during the steady state. The equalizers were composed of 61 coefficients all initialized to 0 except for the central one, which was set to 1. The trajectories of the algorithms are plotted in Fig. 7. First, we might say that the CNA-6 and the CQA perform better than the classical CMA. This could be due to the better adaptation of the cost function of CNA-6 and CQA to the modulation QAM-16. Thus, the steady-state ISI measure is lower for these algorithms than for the CMA with the same algorithm step μ. It then allows us to choose a greater value for μ in order to speed up the convergence. The second remark we can outline from the results is that the trajectories of the CNA-6 and the CQA are virtually identical, and it is difficult to distinguish between them. This fact is explained, in Section IV, by the study of the steady-state of the two algorithms. Also, an examination of the unit ball of these two norms provides us with another intuitive explanation due to their similarities.

B. CMA/CQA Combination Results

In order to see the interest for the CMA/CQA combination in comparison to the CMA and the CQA, we have simulated it under the same conditions as those previously mentioned. In these simulations, we call this combination Constant Dynamic
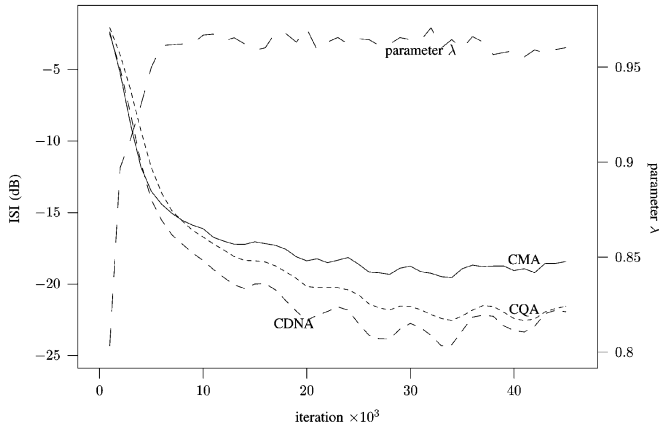
Fig. 8. CDNA performance on Proakis 1 channel.

Fig. 9. Tracking performance of the CDNA.

Fig. 10. CDNA performance on Proakis 1 channel with phase noise burst.

Norm Algorithm (CDNA). For the CDNA, the proportion parameter α is set to 1.18, and the step μ_λ is fixed at $2 \cdot 10^{-2}$. The CDNA and the CMA have the same step size μ in order to have the same initial behavior. Moreover, step μ of the CQA is chosen in order to obtain the same steady state as the CDNA's.

We notice in Fig. 8 that the CDNA starts like the CMA, as expected, then converges more rapidly because it takes into account the contribution of the CQA. Finally, it reaches performances equivalent to the CQA. Latter, once the steady state is reached, the CDNA's steady state is slightly higher than that of the CQA because of the contribution of the CMA part of the algorithm. In fact, if the λ parameter tends to one, it is not exactly equal to one and therefore there is still a contribution by the CMA. At that state of convergence, this contribution is the drawback of this new norm and could be considered as additional noise for the CQA. This first result confirms that the CDNA tends to choose at each iteration the best algorithm between the CMA and the CQA, therefore the CDNA corresponds well to the original intuitive idea.

As we explained in Subsection III-B, the parameter λ gives the proportion between the two norms of this combination at each iteration. Furthermore, we define it as a dynamic parameter. In the case of CDNA, this property is given by (23). It is this tracking issue that we would like to illustrate in Fig. 9. We simulate a perturbation that is an addition of noise between the 8 000 and 10 000 iterations. In this range, the SNR goes from 20 to 9.6 dB. We notice that when the noise increases, the CDNA has a tendency to return to the CMA mode. In fact, parameter λ decreases very quickly from about 0.95 to about 0.5, so the CMA contribution increases in the combination. The CDNA is all the more remarkable in that it converges more rapidly than the CMA.

Another series of simulations were performed to analyze the behavior of the different algorithms on a channel that changes the phase of the constellation. For this, we use the same parameters as before, but we add a rotation in the channel. It is multiplied by a constant, which is equal to $e^{j0.2\pi}$. The resulting performances are plotted in Fig. 10. As the cost function of the CMA depends only on the modulus of the received signal, it is not disturbed by this sudden modification of the
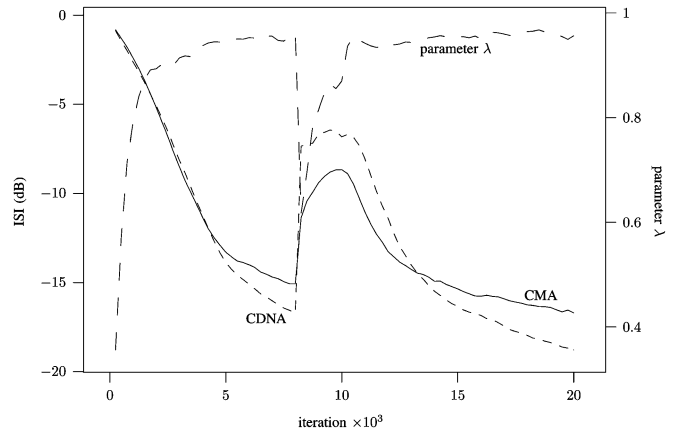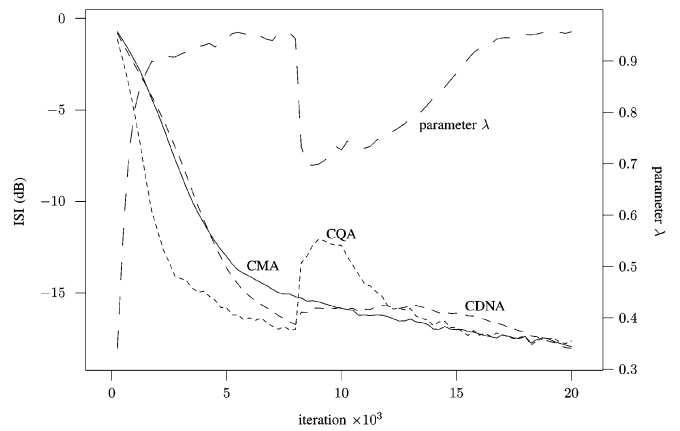
propagation characteristics. On the contrary, the CQA is affected by this modification. Because the CDNA had previously converged to the CQA solution, it is also affected. We notice that the CDNA reacts correctly to this new perturbation by again giving a great value to the CMA contribution in the combination. In fact, parameter λ decreases very quickly from around 0.95 to around 0.7, so the CMA contribution increases in the combination. Then after having recovered the phase, the CDNA again reaches the equivalent CQA solution. It should be noted that the behavior of the CDNA is very efficient because the rotation introduced in the channel ($\approx \pi/4$) is quite difficult for the CQA cost function.

VII. CONCLUSION

In this paper, we have proposed a new blind class of algorithms for both equalization and estimation purposes. This class uses the general description of a norm on both real and complex dimensions. The main contribution of this work is the observation that the CMA blind cost function generally does not match the constellation perfectly (except for PSK constellations). On the contrary, we can find in our new class a particular norm that better matches (in the EMSE sense) the constellation. The new algorithm is very simple to implement; it works well and, with respect to the EMSE, gives the best results, as proven by our simulation results. These simulations were performed in an

single-input single-output environment, but it should be possible to extend the idea developed in this paper to mutliple-input multiple-output channels. Another point that should be studied is the convergence rate of the CNA, particularly of the CDNA. These are topics currently under investigation. The performance of the new algorithms was verified through extensive simulations.

REFERENCES

[1] P. Duhamel, "Blind equalization," presented at the Tutorials of the Int. Conf. Acoustics, Speech, Signal Processing (ICASSP'95) Detroit, MI, 1995.

[2] C. R. Johnson, Jr., P. Schniter, T. J. Endres, J. D. Behm, D. R. Brown, and R. A. Casas, "Blind equalization using the constant modulus criterion: A review," *Proc. IEEE*, vol. 86, no. 10, pp. 1927–1950, Oct. 1998.

[3] Y. Sato, "A method of self recovering equalization for multilevel amplitude-modulation systems," *IEEE Trans. Commun.*, vol. 23, no. 6, pp. 679–682, Jun. 1975.

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

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

[6] S. Bellini, "Bussgang techniques for blind equalization," in *Proc. IEEE GLOBECOM'86*, Houston, TX, Dec. 1–4, 1986, pp. 1634–1640.

[7] J. Yang, J.-J. Werner, and G. A. Durmont, "The multimodulus blind equalization algorithm," in *Proc. 13th Int. Conf. Digital Signal Processing*, Santorini, Greece, Jul. 1997, pp. 127–130.

[8] A. Benveniste and M. Goursat, "Blind equalizers," *IEEE Trans. Commun.*, vol. COM-32, no. 8, pp. 871–883, Aug. 1984.

[9] D. H. Brandwood, "A complex gradient operator and its application in adaptive array theory," *Proc. Inst. Elect. Eng.*, vol. 1, no. 130, pp. 11–16, Feb. 1983.

[10] G. Picchi and G. Prati, "Blind equalization and carrier recovery using a 'stop-and-go' decision directed algorithm," *IEEE Trans. Commun.*, vol. COM-35, no. 9, pp. 877–887, Sep. 1987.

[11] J. G. Proakis, *Digital Communications*. New York: McGraw-Hill, 1989.

[12] A. Goupil and J. Palicot, "A geometrical derivation of the excess mean square error for Bussgang algorithms in a noiseless environment," *Signal Process.*, vol. 84, no. 2, pp. 311–315, Feb. 2004.

[13] A. H. Sayed, *Fundamentals of Adaptive Filtering*. New York: Wiley, 2003.

[14] I. Fijalkow, C. E. Manlove, and C. R. Johnson Jr, "Adaptive fractionally spaced blind CMA equalization: Excess MSE," *IEEE Trans. Signal Process.*, vol. 46, no. 1, pp. 227–231, Jan. 1998.

Alban Goupil was born in Nancy, France, in March 1976. He received the Engineer degree in telecommunications from the École Nationale Supérieure des Télécommunications de Bretagne (ENST Bretagne), Bretagne, France, and the Diplôme d'Études Approfondies and Ph.D. degrees in signal processing from the University of Rennes, France, in 2000 and 2004, respectively.

He was with France Télécom R&D, Rennes, from March 2000 to September 2003. He was with CreSTIC Laboratory, University of Reims Champagne-Ardenne, from February 2004 to February 2006. He joined ETIS laboratory, ENSEA, from March to August 2006. Currently, he is an Assistant Professor in University of Reims Champagne-Ardenne, where he is involved with coding theory and sensor networks.

Jacques Palicot received the Ph.D. degree in signal processing from the University of Rennes, France, in 1983.

Since 1988, he has been involved in studies about equalization techniques applied to digital transmissions and new analog TV systems. Since 1991, he has been involved mainly in studies concerning the digital communications area and automatic measurement techniques. He has taken an active part in various international bodies (EBU, CCIR), and in European projects. He has published various scientific articles, notably on equalization techniques, echo cancellation, neural techniques, and hierarchical modulations. He is currently involved in adaptive signal processing and in new techniques as software radio and Cognitive Radio techniques. Since October 2003, he has been with Supelec, Campus de Rennes, Cesson-Sévigné, France.