ON THE OPTIMAL CONSTANT NORM ALGORITHM, IN RESPECT TO THE EMSE, FOR BLIND QAM EQUALIZATION

Alban Goupil[1] *and Jacques Palicot*[2]

[1]FranceTelecom R&D
DMR/DDH Laboratory
BP59 Rue du clos Courtel
35512 Cesson-Sévigné – France

[2]Supelec
Avenue de la boulaie
BP 81127
35511 Cesson-Sévigné – France
jacques.palicot@supelec.fr

ABSTRACT

During the last Eusipco conference [1] we proposed a new class of algorithm, called Constant Norm Algorithm (CNA), which contains the well-known CMA. From this class, two new cost functions well designed for QAM modulation, were derived. The first, named CQA for Constant sQuare Algorithm, is better adapted for QAM than the classical CMA. It results in a lower algorithm's noise without an increase of complexity. This algorithm was derived thanks to the infinite norm which was intuitively better adapted for square mapping modulation than the norm 2 of the CMA. In the same period [4] we proposed a geometrical derivation for computing the Excess Mean Square Error for Bussgang algorithm. Then in respect to this derivation, we prove in this paper that the optimal norm which minimizes the EMSE for QAM modulation is not the infinite norm (even it gives a lower EMSE than the norm 2) but it is the norm 6.

1. INTRODUCTION

Nowadays, due to socio-economic and technological reasons, the demand for wireless access has been increasing rapidly. It is estimated that this trend will continue in the coming years. Over the last few years there has been an explosion in the number of standards, of networks, of services on these networks and finally an exponential increase in bit-rate. This situation explains, partly, why optimized use of different systems and standards as well as efficient spectrum utilization have become vital issues. To the above ends, many different ways are currently studied in the telecommunications area. Among them, we are specifically interested in this paper on modulation schemes which exhibit high spectrum efficiency, such as QAM modulation. This type of mapping is more sensitive to the channel perturbations and therefore efficient equalization schemes are needed. To improve the overall throughput of a transmission system, we should avoid the use of training period, in other words, we should perform at the receiver side blind equalization. There were a lot of work on blind equalization schemes. In this paper we start from the Constant Norm Algorithm family already presented in the last Eusipco conference [1]. We shown that the infinite norm belonging to this family offers better performances than the classical norm 2 (the CMA). We are going further, in this paper, showing that, in respect to the Excess Mean Square Error, the optimal norm for 16-QAM modulation is the norm 6.

The paper is organized as follows. In Section 2 the problem is formulated. In Section 3 the Constant Norm Algorithm (CNA) family is described then we analyze the convergence in section 4 thanks to the computation of the EMSE. In addition we derived EMSE equations for the CNA case (then implicitly for both CMA and CQA cases). Then in section 6 some simulation results are provided. These results concern EMSE for CNA family algorithms. Finally the work is concluded in section 7.

2. PROBLEM FORMULATION AND PRELIMINARIES

Before proceeding any further let us define the notation used throughout the paper. The vectors are denoted by bold upper case letter and the superscript ^{H} stands for Hermitian transpose. The notation $\langle \mathbf{X}, \mathbf{Y} \rangle$ represents the dot product of the two vectors $\mathbf{X}$ and $\mathbf{Y}$. The conjugate of a complex $z = \mathrm{Re}\, z + \sqrt{-1}\, \mathrm{Im}\, z$ is denoted by $\bar{z}$, and its norm by $\|z\|$ where z is considered as a point in the real plane. Finally, the expectation of the random variable X is denoted by $\mathbb{E}X$.

We are studying Bussgang algorithms, which obey the classical stochastic gradient algorithm. We can derive the general formula (with the notations of figure 1):

$$\mathbf{W}_{k+1} = \mathbf{W}_k - \mu\, \phi(z_k)\, \overline{\mathbf{X}}_k \tag{1}$$

These techniques attempt to find the source data a_n (supposedly i.i.d.), in the most efficient way according to a certain criterion (like MMSE, ZF, . . .), from an observation x_n, which is the result of the convolution of a_n by a finite impulse response channel H and disturbed by a white additive Gaussian noise b_n. In the framework of blind equalization, also called unsupervised or self-learning, the only available *a priori* knowledge is the statistics of the data a_n.

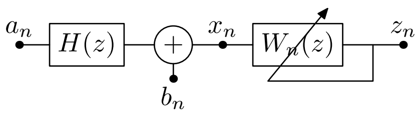

Figure 1: Blind equalization scheme

The resolution to this problem can be made by filtering the received data through a filter W. This filter is optimized in order to minimize a certain cost function $\mathscr{J}$ that depends only on the output z_n. This minimization is made, for example, by a stochastic gradient algorithm. This kind of algorithms contains the well known CMA. Furthermore, we derived a generalization of the algorithm better suited for other modulations than constant modulus ones, like QAM. These algorithms produce a smaller algorithm's noise and the gain could reach in some particular conditions 6 dB in Mean Square Error (MSE).

With a view to simplifying the notations, the time indexes will often be left out. The aim, therefore, is to find a cost function $\mathscr{J}$ such as the perfect equalizer W_{opt} is the global minimum. We will now limit ourselves to cost functions which verify:

$$\begin{cases} \mathbf{W}_{\text{opt}} = \operatorname{argmin}_{\mathbf{W}} \mathscr{J}(z) \\ \mathscr{J}(z) = \mathbb{E} J(z) \end{cases} \tag{2}$$

where $\mathbb{E}$ indicates expectation. From this cost function, it is possible to develop a stochastic gradient algorithm:

$$\begin{cases} \mathbf{W}_{n+1} = \mathbf{W}_n - \mu \phi(z) \overline{\mathbf{X}}_n, \\ \phi(z) = \frac{\partial J(z)}{\partial \bar{z}}. \end{cases} \tag{3}$$

The next section 3 presents a class of cost functions called CNA (Constant Norm Algorithm) and derives from there some particular cases included the CMA in 3.2.1, the CQA (Constant sQuare Algorithm) in 3.2.2 designed for the QAM, and the p-norm CNA. Thanks to a derivation in sections 4 and 5 of the Excess Mean Square Error (EMSE) which measures the algorithm's noise, we optimize the p-norm algorithmsin respect to this criterion in section 6.

3. COST FUNCTIONS

3.1 Constant Norm Algorithm

We defined a new algorithm family. We called it the Constant Norm Algorithm (CNA). This denomination is due to the replacement of the modulus of the CMA by $n(\cdot)$ which is a norm on $\mathbb{R}^2$. Then we can write a new cost function as

$$\mathscr{J}(z) = \frac{1}{pq} \mathbb{E} \left| n^p(z) - R \right|^q. \tag{4}$$

We will show further that the classical CMA cost function is a particular case of (4). In order to respect (2), we can derive the constant R as presented in [1]. Actually, R is fixed in such a way that the perfect equalizer, in the sense of the ZF criterion, is at least a local minimum of $\mathscr{J}$ in a noiseless environment. In the particular case of the $\text{CNA}^{\text{p},2}$, this constant is given by

$$R = \frac{\mathbb{E}\, n^{2p}(a)}{\mathbb{E}\, n^p(a)}, \tag{5}$$

where we recognize the case of the $\text{CMA}^{\text{p},2}$ but where the modulus is changed into the general norm.

3.2 Particular cases of CNA's cost functions

3.2.1 Constant Modulus Algorithm

The $\text{CMA}^{\text{p,q}}$ has been developed by Godard [2] for constant modulus modulations (like the PSK). This is one of the most widely studied algorithms. The cost function can be written as

$$\mathscr{J}(z) = \frac{1}{pq} \mathbb{E} \left| |z|^p - R \right|^q. \tag{6}$$

The $\text{CMA}^{\text{p,q}}$ is then a particular case of the CNA by taking the modulus as the norm $n(\cdot)$.

With our previous simplification, the algorithm takes the simple form

$$\mathbf{W}_{n+1} = \mathbf{W}_n - \mu \left(|z|^2 - R \right) z \overline{\mathbf{X}}_n; \tag{7}$$

where the constant R 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 $\mathbb{E}|a|^4 / \mathbb{E}|a|^2$.

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 (7) generates a significant amount of noise.

3.2.2 Constant sQuare Algorithm

If we look at the constellation of the QAM modulation (fig. 2), we see that it presents more a "square" aspect than a "round" one. As the CMA uses the module norm the "round" aspect is taken into account but not the "square" one. As the CNA class is pretty simple and nice, we would like to derive a particular case satisfying the intuitive idea. This principle was already developed in the multi-modulus algorithm (MMA) [3]. But it is more a matter of the decomposition of the CMA on the phase and quadrature channels, than the generalization of the CMA to QAM. Moreover this algorithm does not belongs to the CNA class.

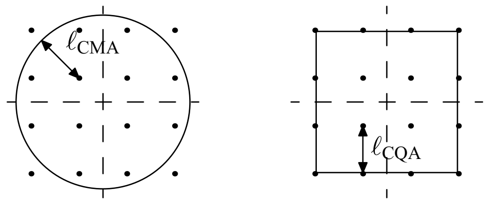

Figure 2: Principle of CMA and CQA.

We remark also on the figure 2 that the distance ℓ_{CMA} between the symbol and the circle is, in average, bigger than the distance ℓ_{CQA} between the symbol and the square. From this fact, we expect that the noise of the algorithms is lower for the square than for the circle.

In order to build a CNA which corresponds to the intuitive description above, a norm should be chosen carefully. In fact, if we look the field $\mathbb{C}$ as the real plane $\mathbb{R}^2$ we could define a norm whose ball is a square: the infinite norm, defined by

$$\|z\|_\infty = \max\left(|\text{Re}\, z|, |\text{Im}\, z| \right), \tag{8}$$

Thus taking this norm and plugging it into the cost function (4) gives us the algorithm called CQA for Constant sQuare Algorithm. This algorithms penalizes the output of the filter W which are far away from the square of radius R. The idea is then the same than the CMA which penalizes the outputs away from the circle.

The parameter R is also given by (5) which depends on the constellation. Of course, the value of R for the CQA is different for the one for the CMA.

Therefore, the CQA cost function is given by

$$\mathscr{J}(z) = \frac{1}{4} \mathbb{E} \left| \|z\|_\infty^2 - R \right|^2, \tag{9}$$

where we restrict ourselves to the more usable case $p = q = 2$. The pseudo-error function, $\phi(z)$, used in the descent algorithm (3), for the CQA becomes:

$$\begin{cases} \phi(z) = \left(\|z\|_\infty^2 - R\right) F(z) \\ F(z) = \begin{cases} \mathrm{Re}\, z & \text{if } |\mathrm{Re}\, z| > |\mathrm{Im}\, z|, \\ \sqrt{-1}\cdot \mathrm{Im}\, z & \text{otherwise.} \end{cases} \end{cases} \tag{10}$$

Even if the algorithm of the CQA given by (10) seems to be more expensive than the CMA, the complexity is of the same order.

3.2.3 p-norm CNA

In order to simplify the presentation, the two parameter p and q of (4) will be set to 2. Moreover we will focus on the norms which are called p-norm defined by

$$\|z\|_p = \sqrt[p]{|\mathrm{Re}\, z|^p + |\mathrm{Im}\, z|^p}. \tag{11}$$

With this norm the cost function becomes

$$\mathscr{J}(z) = \frac{1}{4}\mathbb{E}\left|\|z\|_p^2 - R\right|^2. \tag{12}$$

The interest of these algorithms is to permits to optimize the norm thanks to a criterion. In the following, we will derive a general form given the Excess Mean Square Error. This parameter which measure the noise of the algorithm will allow us to optimize the norm. Moreover, as the modulus corresponds to the 2-norm, the CMA is included into the p-norm.

4. CONVERGENCE ANALYSIS

In a recent paper [4], we proposed a geometrical derivation of the Excess Mean Square Error (EMSE) for Bussgang Algorithms in a noiseless environment.In that article, we applied the Pythagoras theorem and we found a general form of the EMSE. In this form the cost function appears explicitly in the equation. As an example, we applied this method to CMA and we found the results previously obtained in the literature [5, 6]. Now, we are going to apply this result to CNA algorithm.

5. PYTHAGORAS AND LMS

First of all, we should define several points of the figure 3. z_{opt} is the output of the optimal filter and z_k is the output of the filter at time k and z_{k+1} the output of the filter after the update (1). Mathematically speaking, we have $z_{\mathrm{opt}} = \langle \mathbf{W}_{\mathrm{opt}}, \mathbf{X}_k\rangle$, $z_k = \langle \mathbf{W}_k, \mathbf{X}_k\rangle$ and $z_{k+1} = \langle \mathbf{W}_{k+1}, \mathbf{X}_k\rangle$.

As classically done, we define two errors: e_a the *a priori* error, and e_p the *a posteriori* error. These errors represent the difference between z_{opt} and z_k or z_{k+1}. Noting $\Delta\mathbf{W}_k = \mathbf{W}_k - \mathbf{W}_{\mathrm{opt}}$, we have $e_a = \langle \Delta\mathbf{W}_k, \mathbf{X}_k\rangle$ and $e_p = \langle \Delta\mathbf{W}_{k+1}, \mathbf{X}_k\rangle$. To further simplify the notation, we introduce the vector $\mathbf{u} = \frac{\overline{\mathbf{X}}_k}{\|\mathbf{X}_k\|}$. And in order to put the points z_{opt}, z_k and z_{k+1} on the figure, we defined α_{opt}, α_k, and α_{k+1} as $\alpha_i = \frac{z_i}{\|\mathbf{X}_k\|}\mathbf{u}$.

Now if we apply the Pythagoras theorem in both triangles $\left(\alpha_{\mathrm{opt}}, \alpha_k, \mathbf{W}_k\right)$ and $\left(\alpha_{\mathrm{opt}}, \alpha_{k+1}, \mathbf{W}_{k+1}\right)$, taking into account, as it is obvious and could be seen in the figure 3, that the

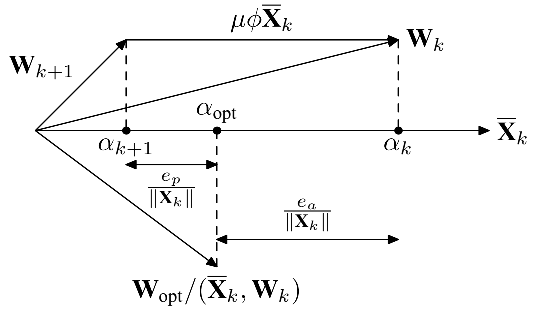

Figure 3: Geometrical interpretation of LMS

norm of the projection of $\mathbf{W}_k$ on the orthogonal space of $\overline{\mathbf{X}}_k$ is constant, we find

$$\|\mathbf{X}_k\|^2\|\Delta\mathbf{W}_{k+1}\|^2 - |e_p|^2 = \|\mathbf{X}_k\|^2\|\Delta\mathbf{W}_k\|^2 - |e_a|^2. \tag{13}$$

We search now a relation between e_a and e_p. As it could be seen in figure 3 and using (1), e_p could be rewritten as

$$|e_p|^2 = |e_a|^2 + \mu^2|\phi|^2\|\mathbf{X}_k\|^4 - 2\mu\|\mathbf{X}_k\|^2\,\mathrm{Re}\,\phi\overline{e_a}. \tag{14}$$

For the ease of computation, we define the following quantity,

$$\Delta e \triangleq \frac{|e_p|^2 - |e_a|^2}{\mu\|\mathbf{X}_k\|^2} = \mu\|\mathbf{X}_k\|^2|\phi|^2 - 2\,\mathrm{Re}\,\phi\overline{e_a}. \tag{15}$$

5.1 EMSE computation

The following remarks will be used to derive the EMSE expression. During the Steady State phase we remark that:

$$\lim_{k\to\inf}\mathbb{E}\|\tilde{\mathbf{W}}_{k+1}\|^2 = \lim_{k\to\inf}\mathbb{E}\|\tilde{\mathbf{W}}_k\|^2, \tag{16}$$

then we obtain:

$$\mathbb{E}|e_a|^2 = \mathbb{E}|e_p|^2. \tag{17}$$

Therefore, with the classical independence assumption between the data and the error, we obtain $\mathbb{E}\Delta e \approx 0$. The main idea of the following is to develop an approximation of the cost function near its optimum z_{opt}. As $z_k - z_{\mathrm{opt}}$ is the *a priori* error e_a, we will find a simple relation between e_a and the value of the function ϕ at the optimum. Thanks to this relation, the EMSE will be computed as $\mathbb{E}|e_a|^2$.

Some assumptions are needed to derive the EMSE form. The *a priori* error e_a is assumed centered and independent of $\|\mathbf{X}_k\|^2$. For the real case, e_a is supposed to be such that $\mathbb{E}e_a^3 \approx 0$ and in complex case, e_a is circular (*i.e.* $\mathbb{E}e_a^2 = 0$). The power of the input vector will thereafter be noted $P_X = \mathbb{E}\|\mathbf{X}_k\|^2$.

The complete development could be found in [4]. In the complex case we find

$$\mathrm{EMSE} \approx \mu P_X \frac{\mathbb{E}|\phi(z_{\mathrm{opt}})|^2}{2\,\mathbb{E}\,\partial_{z\bar{z}}\left\{\mathrm{Re}\,\phi(z)\overline{(z - z_{\mathrm{opt}})}\right\}\Big|_{z_{\mathrm{opt}}}}. \tag{18}$$

As the EMSE is a linear function of μP_X, comparing algorithms could then be done through the comparison of the slope of the EMSE given by the ratio of (18).

5.1.1 Constant Norm Algorithm Application

We consider thereafter that the different algorithms converge to the optimal ZF equalizer. In this case, the z_{opt} variables is a point a of the constellation. With this assumptions, and once the maths are done, the EMSE of the general CNA is given by

$$\text{EMSE}_{\text{CNA}} \approx \mu P_X \cdot \frac{\mathbb{E}n^6(a)|\partial_{\bar{z}}n|^2 - 2R\,\mathbb{E}n^4(a)|\partial_{\bar{z}}n|^2 + R^2\,\mathbb{E}n^2(a)|\partial_{\bar{z}}n|^2}{2\,\mathbb{E}n^3(a)\partial_{z\bar{z}}n - 2R\,\mathbb{E}n(a)\partial_{z\bar{z}}n + 6\,\mathbb{E}n^2(a)|\partial_z n|^2 - 2R\,\mathbb{E}|\partial_z n|^2}. \tag{19}$$

The EMSE of the p-norm CNA is then given by putting the following relations (20) and (21) in (19),

$$|\partial_z n|^2|_a = \left(\frac{\|a\|_p}{\|a\|_{p-1}}\right)^{2(p-1)}; \tag{20}$$

$$\partial_{z\bar{z}}n|_a = \frac{1-p}{\|a\|_p^p}\left(\|a\|_{p-1}^{p-1} + \|a\|_p\|a\|_{p-2}^{p-2}\right). \tag{21}$$

The EMSE of the CMA is then given by computing the previous equations with $p = 2$ which gives

$$\text{EMSE}_{\text{CMA}} \approx \mu P_X \frac{\mathbb{E}|a|^6 - 2R\,\mathbb{E}|a|^4 + R^2\,\mathbb{E}|a|^2}{2\left(2\,\mathbb{E}|a|^2 - R\right)}. \tag{22}$$

This equation is exactly the equation found in [5, 6].

Mathematically speaking, it is not possible to apply directly equation (19) with the infinite norm instead of p. Then to compute the EMSE of the CQA we should start from the equation (9) using the equation (18), and after some maths, we found the EMSE of the CQA which is given by

$$\text{MSE}_{\text{CQA}} \approx \mu P_X \frac{\mathbb{E}\|a\|_\infty^6 - 2R\,\mathbb{E}\|a\|_\infty^4 + R^2\,\mathbb{E}\|a\|_\infty^2}{3\,\mathbb{E}\|a\|_\infty^2 - R}. \tag{23}$$

6. RESULTS AND COMPARISONS

Firstly, we drew on the figure 4 the slopes of equation (19) with respect to the parameter p of the p-norm. The EMSE is dependent of the point a of the constellation, we choose a 16QAM constellation for this result. This graph show that the optimal value of p in respect to the EMSE seems to be around 6. The slope of the CQA drawn on the figure 4 should be taken carefully because the mathematical derivation is not exact. In fact, if we take into account the bias introduced by the derivation, the CNA-6 becomes better than the CQA.

To check these results, we simulate the algorithms and compute the EMSE.We performed these simulations in a SIMO context, with a 16-QAM modulation. We also draw the theoretical results given by (19). The results are given on the figure 5. The dotted lines represent the simulation results and the solid ones the theory. It could be seen that there is a good agreement between the simulation and the theoretical results. These results also confirms that the EMSE of the CNA for $p = 6$ is much lower that the one of the CMA (for the same step size).

7. CONCLUSION

In this article, we have shown that the best norm, belonging to the CNA family,in order to equalize QAM modulation, in a blind manner is the norm 6.

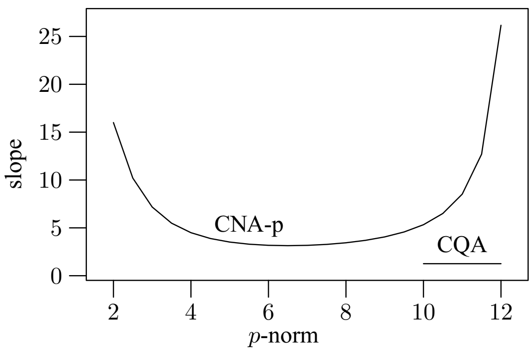

Figure 4: Slopes of the EMSE of p-norm CNA wrt p

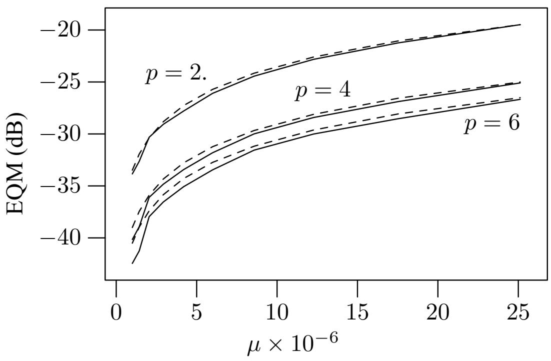

Figure 5: SIMO simulation for a 16-QAM

This result should be considered in respect and (only in respect) of the EMSE of the algorithms.

REFERENCES

[1] Alban Goupil and Jacques Palicot. Constant norm algorithms class. In *Proceedings of EUSIPCO'02*, Toulouse, France, September 2002.

[2] Dominique N. Godard. Self-recovering equalization and carrier tracking in two-dimensional data communication systems. *IEEE Trans. Commun.*, COM-28(11):1867–1875, November 1980.

[3] Jian Yang, Jean-Jacques Werner, and Guy A. Durmont. The multimodulus blind equalization algorithm. In *Proceedings of 13th Int. Conf. Digital Sig. Processing*, pages 127–130, Santorini, Greece, July 1997.

[4] Alban Goupil and Jacques Palicot. A geometrical derivation of the excess mean square error for Bussgang algorithms in a noiseless environment. *Signal Processing*, 84(2):311–315, February 2004.

[5] Inbar Fijalkow, C. E. Manlove, and C. R. Johnson, Jr. Adaptive fractionally spaced blind CMA equalization: excess MSE. *IEEE Trans. Signal Processing*, 46(1):227–231, January 1998.

[6] Junyu Mai and Ali H. Sayed. A feedback approach to the steady-state performance of fractionally spaced blind adaptive equalizers. *IEEE Trans. Signal Processing*, 48(1):80–91, January 2000.