BLIND EQUALIZATION OF MULTILEVEL SIGNALS USING SUPPORT VECTOR MACHINES

Marcelino Lázaro[1], *Ignacio Santamaría*[2], *Javier Vía*[2] *and Deniz Erdogmus*[3]

[1]Departamento de Teoría de la Señal y Comunicaciones. Universidad Carlos III de Madrid
Leganés, Madrid, Spain. Email: marce@ieee.org
[2] Departamento de Ingeniería de Comunicaciones. Universidad de Cantabria
Santander, Cantabria, Spain. Email: {nacho,jvia}@gtas.unican.es
[3] Computational NeuroEngineering Laboratory. University of Florida.
Gainesville, Florida, USA. Email: derdogmus@ieee.org

ABSTRACT

The support vector machine (SVM) has been recently proposed for blind equalization of constant modulus signals. In this paper we extend this previous work in two directions: first, the high computational cost of the original procedure is significantly reduced by transforming the original quadratic programming (QP) problem into an equivalent least squares problem. Secondly, the penalty term of the SVM is now a Godard-like error function; therefore, the proposed procedure allows the equalization of multilevel signals. A dual mode algorithm is also proposed: once convergence is achieved, the Godard-like penalty term is switched to a radius directed-like error function, which reduces the final intersymbol interference (ISI) level. Simulation experiments show that the proposed SVM equalization method performs better than cumulant-based methods: it requires a lower number of data samples to achieve the same equalization level and convergence ratio.

1. INTRODUCTION

In most digital communication systems the channel introduces intersymbol interference (ISI), which distorts the original transmitted sequence and can make it unrecoverable. In this case, channel equalization is typically employed to reduce, or ideally to completely remove, the ISI. When a reference sequence is not available, blind equalization techniques are necessary. These methods rely on the knowledge about the probabilistic or statistical properties of the transmitted signals [1, 2].

In this paper we will focus on algorithms working at symbol rate. In this case, two broad classes of algorithms can be identified. On one hand, the so called on-line techniques, which are typically based on the maximization/minimization of a cost function by means of stochastic gradient descent (SGD) methods. The most representative method of this class is the well known CMA algorithm [3, 4]. In the other hand, batch techniques collect a block of data and iteratively maximize/minimize a cost function, commonly based on cumulants. To this class of methods belongs the super-exponential algorithm, proposed by Shalvi and Weinstein [5].

Support Vector Machines (SVM) are state-of-the-art tools for linear and nonlinear input-output knowledge discovery [6]. SVM's have been successfully applied to linear and nonlinear supervised equalization problems. Recently, this framework has been used to formulate the blind equalization of constant modulus signals [7, 8] (see therein references to SVM-based equalization methods). This method is implemented by means of an iterative re-weighted quadratic programming (IRWQP) technique, which requires a high computational burden.

In this paper we propose a new SVM-based method to solve the blind equalization problem. The cost function incorporates a Godard or CMA-like error function as the penalty term and the solution is obtained by means of an iterative re-weighted least squares algorithm (IRWLS). For constant modulus signals, this solution provides basically the same results as the method in [8] with a notably lower computational burden. Moreover, like CMA, the error function allows to extend the method to multilevel modulations. In this case, a dual mode algorithm is proposed. Dual mode equalization techniques are commonly used in communication systems working with multilevel signals. Practical blind algorithms for multilevel modulation are able to open the eye of the constellation but they usually exhibit a high residual error. In a dual mode scheme, once the eye is opened by the blind algorithm, the system switches to another algorithm, which is able to obtain a lower residual error under a suitable initial ISI level. The most common choices are decision directed equalization [9] and radius directed equalization [10].

The paper is organized as follows. In Section 2, the blind equalization problem is formulated and CMA and the super-exponential methods are outlined. The proposed algorithm is presented in Section 3. Section 4 shows the performance of the method by means of some experimental results. Finally, Section 5 includes some concluding remarks.

2. PROBLEM FORMULATION

The usual baseband representation of the blind equalization problem in a digital communication system is shown in Fig. 1.

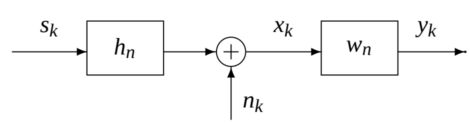

Figure 1: Block diagram for blind equalization.

The transmitted data is modeled as a sequence of i.i.d. symbols, s_k, belonging to the alphabet of the corresponding modulation. This sequence is transmitted through a linear time-invariant channel with impulse response coefficients h_n. Considering a baud rate system, the output of the channel is given by

$$x_k = \sum_{n=0}^{L_h-1} h_n s_{k-n} + n_k,$$

where L_h is the channel length and n_k is a zero-mean white Gaussian noise sequence.

The blind equalizer will operate on the channel output to reduce the ISI introduced by the channel. In this paper a linear equalizer will be implemented by means of an FIR filter of length L_w. In this case, the equalizer output is given by

$$y_k = \sum_{n=0}^{L_w-1} w_n x_{k-n} = \mathbf{x}_k^T \mathbf{w}, \quad (1)$$

This work has been partially supported by Spanish Ministry of Science and Technology (MCYT) through grant TIC2001-0751-C04-03.

where $\mathbf{w}$ is the vector of filter coefficients to be adapted by the blind equalization algorithm. The Godard algorithms [3] adapt the equalizer to minimize the following cost function

$$J_G(\mathbf{w}) = E\left[(|y_k|^p - R_p)^2\right].$$

The ratio R_p contains the *a priori* knowledge about the current modulation,

$$R_p = \frac{E\left[|s_k|^{2p}\right]}{E\left[|s_k|^p\right]}. \tag{2}$$

CMA is the Godard algorithm for $p = 2$. The proposed method introduces a penalty term inspired by the CMA cost function.

For comparison purposes we use the super-exponential method, proposed by Shalvi and Weinstein [5]. This cumulant-based algorithm maximizes $|K_y|$, the modulus of the kurtosis of the equalizer output, y_k, where the kurtosis is defined as

$$K_y = E[|y_k|^4] - 2\left(E\left[|y_k|^2\right]\right) - \left|E\left[y_k^2\right]\right|^2,$$

subject to $E[|y_k|^2] = E[|s_k|^2]$.

3. MULTILEVEL SVM-BASED BLIND EQUALIZATION

In this section we formulate two variants of the proposed method: a blind algorithm and a radius directed algorithm. For multilevel modulations, these two algorithms can be used in a dual mode equalization scheme: the blind algorithm is used until convergence is obtained; then, a switch to the radius directed algorithm will allow to reduce the residual ISI.

3.1 Blind algorithm

Given a data block of N symbols, the proposed algorithm minimizes the following SVM-based cost function

$$L_P(\mathbf{w}) = \frac{1}{2}||\mathbf{w}||^2 + C\sum_{i=1}^{N} L_\varepsilon(u_i), \tag{3}$$

where

$$L_\varepsilon(u) = \begin{cases} 0, & u < \varepsilon \\ u^2 - 2u\varepsilon + \varepsilon^2, & u \geq \varepsilon \end{cases}$$

is a ε-insensitive quadratic loss function modified to guarantee a continuous derivative. Continuity of the derivative is necessary for the numerical stability of the algorithm. To apply this cost function to the problem of blind equalization, a suitable penalization term, u_i, has to be selected. Here, we propose to use $u_i = |e_i|$ with the error term e_i being

$$e_i = |y_i|^2 - R_2 = y_i y_i^* - R_2. \tag{4}$$

R_2 is the Godard constant (2) for $p = 2$ and superindex * denotes the complex conjugate.

To optimize the proposed cost function an iterative re-weighted least square (IRWLS) procedure is employed. This procedure has been successfully applied to solve SVM's [11] and it has recently proven to converge to the SVM solution [12]. To obtain the IRWLS algorithm, a first order Taylor expansion of $L_\varepsilon(u)$ leads to the cost function

$$L_P^{'}(\mathbf{w}) = \frac{1}{2}\|\mathbf{w}\|^2 + C\left(\sum_{i=1}^{N} L_\varepsilon(u_i^k) + \left.\frac{dL_\varepsilon(u)}{du}\right|_{u_i^k}[u_i - u_i^k]\right),$$

where $u_i^k = |e_i^k|$ and $e_i^k = |\mathbf{x}_i^T\mathbf{w}^k|^2 - R_2$ is the error term after the k-th iteration. Then, a quadratic approximation is constructed as follows:

$$\begin{aligned} L_P^{''}(\mathbf{w}) &= \frac{1}{2}\|\mathbf{w}\|^2 + C\left(\sum_{i=1}^{N} L_\varepsilon(u_i^k) + \left.\frac{dL_\varepsilon(u)}{du}\right|_{u_i^k}\frac{(u_i)^2 - (u_i^k)^2}{2u_i^k}\right) \\ &= \frac{1}{2}\|\mathbf{w}\|^2 + \frac{1}{2}\sum_{i=1}^{N} a_i|e_i|^2 + CTE. \end{aligned} \tag{5}$$

CTE represents constant terms that do not depend on $\mathbf{w}$, and the weights a_i are

$$a_i = \frac{C}{u_i^k}\left.\frac{dL_\varepsilon(u)}{du}\right|_{u_i^k} = \begin{cases} 0, & u_i^k < \varepsilon \\ \frac{2C(u_i^k - \varepsilon)}{u_i^k}. & u_i^k \geq \varepsilon \end{cases}. \tag{6}$$

The functional $L_P^{''}(\mathbf{w})$ is a quadratic approximation to $L_P(\mathbf{w})$ in (3) that presents the same value $L_P^{''}(\mathbf{w}^k) = L_P(\mathbf{w}^k)$ and gradient $\nabla_\mathbf{w} L_P^{''}(\mathbf{w}^k) = \nabla_\mathbf{w} L_P(\mathbf{w}^k)$ for $\mathbf{w} = \mathbf{w}^k$. Therefore, we can define $\mathbf{p}^k = \mathbf{w}^s - \mathbf{w}^k$ as a descending direction for $L_P(\mathbf{w})$, where $\mathbf{w}^s$ is the least square solution to (5), and we can use it to construct a line search method [13], i.e. $\mathbf{w}^{k+1} = \mathbf{w}^k + \eta^k\mathbf{p}^k$. The value of η^k can be computed using a backtracking line search [13], in which η^k is initially set to 1 and if $L_P(\mathbf{w}^{k+1}) \geq L_p(\mathbf{w}^k)$, it is iteratively reduced until a strict decrease in the functional in (3) is observed. To obtain the solution to $L_P^{''}(\mathbf{w})$, its gradient is equated to zero

$$\nabla_\mathbf{w} L_P^{''}(\mathbf{w}) = \mathbf{w} + 2\sum_{i=1}^{N} a_i\left(|\mathbf{x}_i^T\mathbf{w}|^2 - R_2\right)\mathbf{x}_i^T\mathbf{w}\mathbf{x}_i^* = \mathbf{0}.$$

This is a nonlinear function of $\mathbf{w}$. To circumvent this nonlinearity, the equalizer output y_i is considered fixed, which leads to

$$\nabla_\mathbf{w} L_P^{''}(\mathbf{w}) = \mathbf{w} + 2\sum_{i=1}^{N} a_i\left(\mathbf{x}_i^T\mathbf{w}y_i^* - R_2\right)y_i\mathbf{x}_i^* = \mathbf{0}.$$

This equation can be expressed in matrix notation

$$[2\mathbf{X}^H\mathbf{D_a}\mathbf{D_{|y|^2}}\mathbf{X} + \mathbf{I}]\mathbf{w} = 2R_2\mathbf{X}^H\mathbf{D_a}\mathbf{Y}, \tag{7}$$

where $\mathbf{X}^T = [\mathbf{x}_1, \mathbf{x}_2, \cdots, \mathbf{x}_N]$, $\mathbf{D_a}$ is a diagonal matrix with diagonal elements a_i and $\mathbf{D_{|y|^2}}$ is another diagonal matrix with diagonal elements $|y_i|^2$. Superindex ^{H} denotes the hermitian operator.

3.1.1 *Implementation details*

The equalizer is initialized with a tap-centered scheme. This initialization has experimentally demonstrated a good convergence behavior under a number of different kind of channels (see results under random channels in Section 4).

With respect to parameters C and ε, although further research is necessary to determine their optimal values, simulations have shown that the algorithm is not very sensitive to its choice. Typically, values of $C = 10$ and $\varepsilon = 0.01$ produce suitable results under a wide range of channels and signal to noise ratios.

Finally, the IRWLS procedure is summarized in Table 1.

3.2 Radius directed algorithm

The radius directed algorithm is formulated by replacing R_2 in the blind algorithm by the radius R_{ki}, which is defined as follows

$$R_{ki} = \min_{R_k}(||y_i|^2 - R_k|).$$

Here, R_k are the different values of $|s_i|^2$ in the underlying signal constellation. For instance, for a 16QAM with levels ± 3, ± 1 in

1. Initialization: initialize $\mathbf{w}^0$ with tap-centered strategy, obtain y_i by (1), e_i by (4), calculate $u_i = |e_i|$ and compute a_i from (6). Set $k = 0$.
2. Compute $\mathbf{w}^s$ by solving (7) and set $\eta^k = 1$.
3. Set $\mathbf{w}^{k+1} = \mathbf{w}^k + \eta^k[\mathbf{w}^s - \mathbf{w}^k]$. If $L_P(\mathbf{w}^{k+1}) < L_P(\mathbf{w}^k)$ go to Step 5.
4. Set $\eta^k = \rho\eta^k$ with $0 < \rho < 1$ and go to Step 3.
5. Recompute e_i, u_i and a_i, set $k = k+1$ and go to Step 2 until convergence.

Table 1: IRWLS pseudocode

both the phase and quadrature components, $R_k = \{2, 10, 18\}$. With this simple modification, the error term to be penalized is

$$e_i = |y_i|^2 - R_{ki}, \tag{8}$$

and the matrix system to obtain the solution $\mathbf{w}^s$ becomes

$$[2\mathbf{X}^H\mathbf{D_a}\mathbf{D}_{|\mathbf{y}|^2}\mathbf{X} + \mathbf{I}]\mathbf{w} = 2\mathbf{X}^H\mathbf{D_a}\mathbf{D_R}\mathbf{Y}. \tag{9}$$

$\mathbf{D_R}$ is a diagonal matrix with diagonal elements R_{ki}. The corresponding IRWLS algorithm is the same one summarized in Table 1 with the following differences:

- Step 1: since this algorithm is used as the second step of a dual-mode algorithm, $\mathbf{w}^0$ is the value of $\mathbf{w}$ provided by the blind algorithm.
- Steps 1 and 5: e_i is evaluated by (8), instead of (4).
- Step 2: $\mathbf{w}^s$ is obtained by solving (9), instead of (7).

4. SIMULATION RESULTS

In this section, simulation results for constant modulus signals and multilevel signals are presented. As a figure of merit we use the ISI defined as

$$ISI = 10\log_{10}\frac{\sum_n|\theta_n|^2 - \max_n|\theta_n|^2}{\max_n|\theta_n|^2},$$

where $\theta = \mathbf{h} * \mathbf{w}$ is the combined channel-equalizer impulse response.

4.1 Constant modulus signals

In the first example, the proposed blind algorithm, labeled Blind-SVM in the following, is tested with a QPSK modulation ($s_k = \{\pm 1 \pm j\}/\sqrt{2}$) and the following channel

$$H_1(z) = \frac{0.7 - z^{-1}}{1 - 0.7z^{-1}}e^{j\pi/4}.$$

The proposed method is compared with the super-exponential algorithm proposed by Shalvi and Weinstein [5], labeled SW-algorithm in the following. We used an equalizer of length $L_w = 17$ taps, with tap-centered initialization. Parameters $C = 10$ and $\varepsilon = 0.01$ are selected. Signal to noise ratios (SNR) of 30 dB and 10 dB are considered. Figs. 2 and 3 show the ISI and the probability of convergence for different data block sizes, respectively. For each data block size, both algorithms were tested in 200 Monte Carlo trials.

The Blind-SVM method provides better results than the SW-algorithm, specially for short data blocks. These results are almost equal to the results provided by the IRWQP algorithm presented in [8] (they are not included in the figures because they basically overlap the Blind-SVM results). In this case, the proposed method, which solves the SVM by a IRWLS algorithm instead of a IRWQP algorithm, has the advantage of requiring a notably lower computational burden since the original quadratic programming (QP) problem is now formulated as a least squares problem.

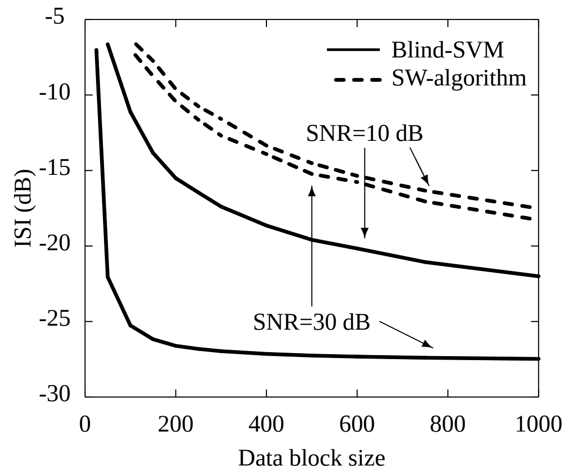

Figure 2: Mean ISI level after convergence (final ISI under -5 dB) vs. data block size for the proposed Blind-SVM algorithm and the SW-algorithm. Channel $H_1(z)$ and QPSK input.

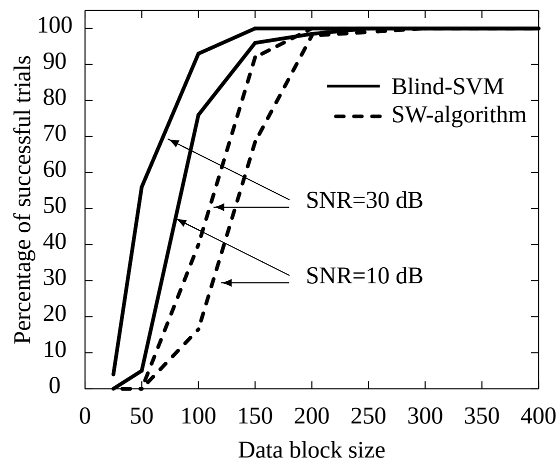

Figure 3: Percentage of convergence vs. data block size for the proposed Blind-SVM algorithm and the SW-algorithm. Channel $H_1(z)$ and QPSK input.

4.2 Multilevel signals

A 16-QAM modulation, with $s_k = \{\{\pm 3, \pm 1\} + j\{\pm 3, \pm 1\}\}$, will be used to test the performance of the proposed methods for multilevel signals. In this second example, the proposed Blind-SVM and the dual mode algorithm, which starts with the blind algorithm and after convergence is achieved switches to the proposed radius directed SVM algorithm, are compared again with the SW-algorithm. The dual mode algorithm will be labeled DM-SVM in the following. 1000 Monte Carlo simulation with random channels of length $L_h = 7$ and a SNR=30 dB have been averaged. Again, $L_w = 17$, $C = 10$ and $\varepsilon = 0.01$ are used. Figs. 4 and 5 compare the mean residual ISI and the percentage of convergence of these methods. It is necessary to remark that the DM-SVM algorithm has the same percentage of convergence as the Blind-SVM algorithm because this is the initial stage, responsible of the convergence, of the DM-SVM algorithm.

Again, the proposed methods outperform the SW-algorithm. The SVM-based methods exhibit a higher percentage of convergence, obtaining good convergence rates even for very short data block sizes. Moreover, the residual ISI is lower than the one obtained by the SW-algorithm. As expected, the DM-SVM method produces a better final solution than the Blind-SVM algorithm, as the radius directed algorithm achieves a better final ISI under a suitable initialization.

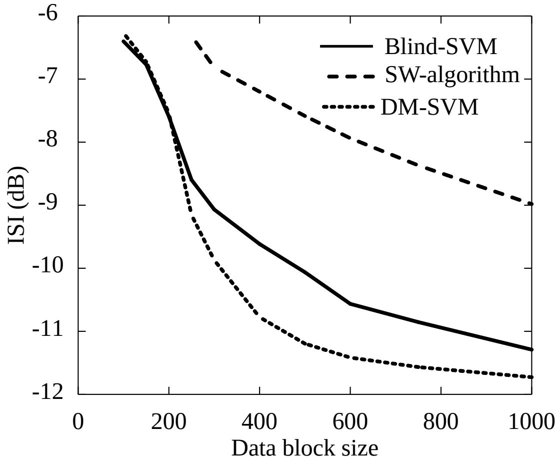

Figure 4: Mean ISI level after convergence (final ISI under -5 dB) vs. data block size for the proposed Blind-SVM and DM-SVM algorithms and the SW-algorithm. 16-QAM input and 1000 random channels with a SNR=30 dB.

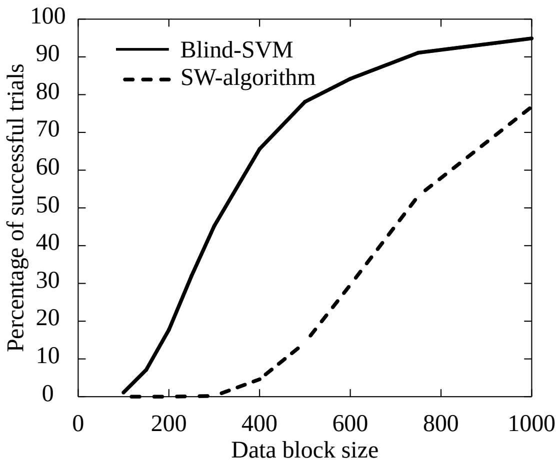

Figure 5: Percentage of convergence vs. data block size for the proposed Blind-SVM) algorithm (and consequently for the DM-SVM algorithm) and the SW-algorithm. 16QAM input and 1000 random channels with a SNR=30 dB.

Finally, to illustrate the convergence of the IRWLS algorithm, Fig. 6 shows the ISI versus the number of iterations of the IRWLS algorithm used in the Blind-SVM method for channel $H_1(z)$, a 16QAM input, a SNR=30 dB and a data block size $N = 750$ (the percentage of convergence for this value is 100%). It can be seen that all experiments have converged in less than 10 iterations.

5. CONCLUSION

A new SVM-based method to address blind equalization of constant modulus and multilevel signals has been presented. The error function to be penalized is defined as the difference between the equalizer output and the R_p parameter of the CMA algorithm. An IRWLS algorithm is employed to obtain the equalizer minimizing the proposed cost function. For constant modulus constellations, this algorithm provides basically the same results as the IRWQP method in [8] with a lower computational burden. For multilevel signals, a dual mode algorithm is proposed. In this case, after initial convergence of the blind algorithm is achieved, a switch to a radius directed error function is performed. The proposed method performs better than the SW-algorithm for constant modulus as well as for multilevel input signals. In both cases, the SVM-based methods present a higher percentage of convergence for a given data block

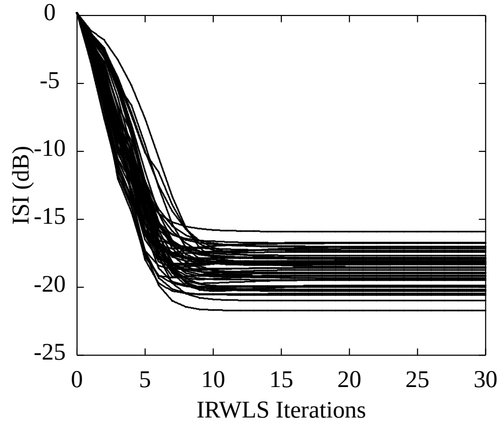

Figure 6: ISI vs. number of iterations of the IRWLS algorithm for the Blind-SVM method using a data block size $N = 750$. Channel $H_1(z)$, 16QAM input and SNR=30 dB.

and they reach a lower residual intersymbol interference.

REFERENCES

[1] S. Haykin, *Blind Deconvolution*, Prentice Hall, Englewood, 1994.

[2] Z. Ding and Y. Li, *Blind Equalization and Identification*, Marcel Dekker, New York, 2001.

[3] D. N. Godard, “Self-recovering equalization and carrier tracking in two dimensional data communication systems,” *IEEE Transactions on Communications*, vol. 28, pp. 1867–1875, Nov. 1980.

[4] J. R. Treichler and B.G. Agee, “A new approach to multipath correction of constant modulus signals,” *IEEE Transactions on Acoustics, Speech, and Signal Proccesing*, vol. 31, pp. 349–372, Nov. 1983.

[5] O. Shalvi and E. Weinstein, “Super-exponential methods for blind deconvolution,” *IEEE Transactions on Information Theory*, vol. 39, no. 2, pp. 504–519, Mar. 1993.

[6] Bernhard Schölkopf and Alex Smola, *Learning with Kernels*, MIT Press, Cambridge, MA, 2002.

[7] I. Santamaría, J. Ibáñez, L.Vielva, and C. Pantaleón, “Blind equalization of constant modulus signals via support vector regression,” in *Proceedings of the International Conference on Acoustics, Speech, and Signal Proccesing*, Hong Kong, China, Apr. 2003, vol. II, pp. 737–740.

[8] I. Santamaría, C. Pantaleón, L. Vielva, and J. Ibáñez, “Blind equalization of constant modulus signals using support vector machines,” *IEEE Trans. on Signal Processing*, In Press, 2004.

[9] Albert Benveniste and Maurice Goursat, “Blind equalizers,” *IEEE Transactions on Communications*, , no. 32, pp. 871–883, Aug. 1984.

[10] M. Ready and R. Gooch, “Blind equalization based on radius directed adaptation,” in *Proceedings of the International Conference on Acoustics, Speech, and Signal Proccesing*, Albuquerque, 1990, pp. 1699–1702.

[11] F. Pérez-Cruz, A. Navia-Vázquez, P. Alarcón-Diana, and A. Artés-Rodríguez, “An IRWLS procedure for SVR,” in *Proceedings of the EUSIPCO’00*, Tampere, Finland, Sept. 2000.

[12] F. Pérez-Cruz, C. Bousoño-Calzón, C.-H. Lin, and A. Artés-Rodríguez, “Convergence of the IRWLS procedure to the support vector machine solution,” *IEEE Transactions on Neural Networks*, Submitted, 2003.

[13] J. Nocedal and S. J. Wright, *Numerical Optimization*, Springer, 1999.