SVM-Based Blind Beamforming of Constant Modulus Signals

Ignacio Santamaría
Dept. of Communications Engineering
University of Cantabria
39005 Santander, Spain
E-mail: nacho@gtas.dicom.unican.es

Javier Vía
Dept. of Communications Engineering
University of Cantabria
39005 Santander, Spain
E-mail: jvia@gtas.dicom.unican.es

Javier Merino
Dept. of Communications Engineering
University of Cantabria
39005 Santander, Spain
E-mail: merino@gtas.dicom.unican.es

***Abstract*— Recent work has shown how the support vector machine (SVM) framework can be used for blind equalization of constant modulus (CM) signals. The basic idea consists of exploiting the CM property of the input signals to reformulate the blind equalization problem as a regression problem. In this paper, we extend this idea to encompass the problem of separating and estimating multiple CM signals mixed through an unknown matrix (i.e., blind beamforming). The quadratic inequalities derived from the CM property are transformed into linear ones, thus yielding a quadratic programming (QP) problem. Then an iterative reweighted procedure is proposed to blindly restore the CM property. Once a signal is recovered, its contribution to the original observations is removed and the iterative procedure can be applied again to extract another CM signal. Simulation results show that this SVM-based algorithm offers better performance than the algebraic constant modulus algorithm (ACMA), mainly when only a small number of snapshots is available.**

I. Introduction

In this work we consider the problem of separating and estimating multiple CM signals (e.g., QPSK) mixed through an unknown matrix. This is a common problem in wireless communications, where an array of antennas receives a number of signals from distinct locations at the same frequency and at the same time (blind beamforming) [1].

In the context of blind beamforming, the constant modulus algorithm (CMA) has been applied to train a set of linear filters (beamformers) to restore the constant modulus property of the sources [2]. Different variations and implementations (block or iterative) of the CMA for beamforming have been proposed [3], [4], [5]. In particular, the analytical constant modulus algorithm (ACMA) [6] is a block technique which finds the solution for P beamformers by solving a generalized eigenvalues problem. The ACMA is a robust algorithm in the presence of noise; however, its perfomance degrades substantially whit rank-deficient or ill-conditioned covariance matrices. This is a typical situation, for instance, in passive sonar, where large aperture arrays that require larger duration snapshots are used [7]. Also, in wireless communications, the use of blind beamforming techniques which are able to work with only a few snapshots, can reduce unnecessary delays and increase the throughput.

To solve this drawback, in this paper we propose a new technique for blind beamforming based on regression via support vector machines (SVM). The SVM-based learning approach has been recently applied to several unsupervised digital communication problems, such as blind equalization of CM signals [8], [9], or blind identifications of single-input-multiple-output (SIMO) channels [10]. This technique is extended here to the case of multiple CM signals mixed through an unknown matrix. Specifically, we consider the blind beamforming problem due to its interest in wireless communications. Nevertheless, the proposed technique can be applied to any mixing matrix (blind source separation). Although with a high computational cost, the proposed technique benefits from the advantages of SVM in regression problems. In particular, as it was expected, simulation results show that the proposed SVM-based blind beamforming technique requires a smaller number of snapshots than the ACMA.

II. Problem formulation

A set of L signals that simultaneously impinges on a linear array of M antennas is considered. Observations at the output of the array can be modeled as

$$\mathbf{X} = \mathbf{AS} + \mathbf{N},$$

where $\mathbf{X}$ is a matrix of dimensions $M \times N$ that contains N samples of the signals (baseband) received by each of the M antennas, $\mathbf{S}$ is an $L \times N$ matrix with the CM source signals, and $\mathbf{A}$ is an $M \times L$ matrix that represents the array response. Finally, $\mathbf{N}$ is an $M \times N$ matrix that takes into account the additive noise present at the observations, which is modeled as spatially white and Gaussian. For an uniform linear array of omnidirectional antennas, matrix $\mathbf{A}$ is given by

$$\mathbf{A} = \begin{bmatrix} 1 & \cdots & 1 \\ e^{-j\phi_1} & \cdots & e^{-j\phi_L} \\ \vdots & \vdots & \vdots \\ e^{-j(M-1)\phi_1} & \cdots & e^{-j(M-1)\phi_L} \end{bmatrix},$$

where

$$\phi_i = 2\pi \frac{d}{\lambda} \sin\theta_i, \quad \text{for} \quad i = 1, 2, \ldots, L;$$

where d is the antenna separation, λ is the wavelength of the signals and θ_i is the angle of arrival of the i-th signal [1].

In our problem, we suppose that a number P (known) from the L incident signals are constant modulus. The problem consists in finding the P beamformers $\mathbf{w}_j$ which provide the estimates of the original signals,

$$y_j[n] = \sum_{i=1}^{M} w_{i,j} x_i[n] = \mathbf{w}_j^{\mathrm{T}} \mathbf{x}[n],$$

for $j = 1, \ldots, P$ and $n = 0, \ldots, N-1$; where $\mathbf{x}[n]$ is the n-th column of $\mathbf{X}$.

III. SVM-BASED BLIND BEAMFORMING

In this section we consider the situation where only one of the incident signals has constant modulus ($P = 1$). The goal of the beamformer is to restore the CM property of this digital communications signal, i.e., $\|y[n]\|^2 = \left\|\mathbf{w}^{\mathrm{T}}\mathbf{x}[n]\right\|^2 = 1$, for $n = 0, \ldots, N-1$.

The support vector machine (SVM) is a powerful learning technique for solving classification, regression and estimation problems that has received considerable attention in recent years [11], [12]. In the context of SVM regression [11], the blind beamforming problem amounts to minimize the cost function

$$J(\mathbf{w}) = \frac{1}{2}\|\mathbf{w}\|^2 + C \sum_{n=0}^{N-1} \left|1 - \|\mathbf{w}^{\mathrm{T}}\mathbf{x}[n]\|^2\right|_{\epsilon},$$

where

$$\left|1 - \|\mathbf{w}^{\mathrm{T}}\mathbf{x}[n]\|^2\right|_{\epsilon} = \max\{0, \left|1 - \|\mathbf{w}^{\mathrm{T}}\mathbf{x}[n]\|^2\right| - \epsilon\}$$

is the Vapnik's ϵ-insensitive loss function, and C is the regularization parameter that penalizes errors larger than ϵ. This problem is equivalent to minimizing the function

$$L(\mathbf{w}, \boldsymbol{\xi}, \tilde{\boldsymbol{\xi}}) = \frac{1}{2}\|\mathbf{w}\|^2 + C \sum_{n=0}^{N-1} \left(\xi_n + \tilde{\xi}_n\right),$$

subject to

$$\begin{aligned} \|\mathbf{w}^{\mathrm{T}}\mathbf{x}[n]\|^2 - 1 \quad &\leq \epsilon + \xi_n, && (1) \\ 1 - \|\mathbf{w}^{\mathrm{T}}\mathbf{x}[n]\|^2 \quad &\leq \epsilon + \tilde{\xi}_n, && (2) \\ \xi_n, \tilde{\xi}_n \quad &\geq 0, \end{aligned}$$

for $n = 0, \ldots, N-1$.

In the conventional SVM formulation, the constraints are linear with respect to the unknowns $\mathbf{w}$. Then, a quadratic programming (QP) problem results that can be efficiently solved [13]. However, inequalities (1) and (2) are quadratic with respect to the beamformer weights. The method proposed in [8], [9] is used to solve this problem. In particular, the squared modulus at the beamformer output can be rewritten as

$$\|y[n]\|^2 = \|\mathbf{w}^{\mathrm{T}}\mathbf{x}[n]\|^2 = \widetilde{\mathbf{w}}^{\mathrm{T}}\widetilde{\mathbf{g}}[n],$$

where we have defined

$$\widetilde{\mathbf{w}} = \begin{bmatrix} \mathrm{Re}(\mathbf{w}) \\ \mathrm{Im}(\mathbf{w}) \end{bmatrix}, \qquad \widetilde{\mathbf{g}}[n] = \begin{bmatrix} \mathrm{Re}(y[n]\mathbf{x}[n]^*) \\ \mathrm{Im}(y[n]\mathbf{x}[n]^*) \end{bmatrix}, \tag{3}$$

and $(\cdot)^*$ denotes complex conjugated.

It is important to note that in (3) we have considered the beamformer output $y[n]$ fixed. In this way, inequalities (1) and (2) become

$$\begin{aligned} \widetilde{\mathbf{w}}^{\mathrm{T}}\widetilde{\mathbf{g}}[n] - 1 \quad &\leq \epsilon + \xi_n, && (4) \\ 1 - \widetilde{\mathbf{w}}^{\mathrm{T}}\widetilde{\mathbf{g}}[n] \quad &\leq \epsilon + \tilde{\xi}_n, && (5) \end{aligned}$$

which are linear with respect to $\widetilde{\mathbf{w}}$. Now, the problem can be written as follows: to maximize

$$\begin{aligned} W(\alpha, \tilde{\alpha}) = \sum_{n=0}^{N-1} (\tilde{\alpha}_n - \alpha_n) - \epsilon \sum_{n=0}^{N-1} (\alpha_n + \tilde{\alpha}_n) - \quad (6) \\ -\frac{1}{2} \sum_{n,m=0}^{N-1} (\tilde{\alpha}_n - \alpha_n)(\tilde{\alpha}_m - \alpha_m) \left\langle \widetilde{\mathbf{g}}[n], \widetilde{\mathbf{g}}[m] \right\rangle, \end{aligned}$$

subject to $0 \leq \alpha_n, \tilde{\alpha}_n \leq C$; and where $\langle \widetilde{\mathbf{g}}[n], \widetilde{\mathbf{g}}[m] \rangle$ denotes the inner product.

In summary, by considering the weighted input patterns, we have shown that the problem can be reformulated as a conventional QP problem with real variables. Its solution is given by

$$\widetilde{\mathbf{w}} = \sum_{n=0}^{N-1} (\tilde{\alpha}_n - \alpha_n) \widetilde{\mathbf{g}}[n].$$

Finally, it is straightforward to show that the beamformer coefficients can be expanded as

$$\mathbf{w} = \sum_{n=0}^{N-1} (\tilde{\alpha}_n - \alpha_n) y[n] \mathbf{x}[n]^* = \sum_{n=0}^{N-1} \beta_n \mathbf{x}[n]^*, \tag{7}$$

where a new set of *weighted* Lagrange multipliers have been defined as

$$\beta_n = (\tilde{\alpha}_n - \alpha_n) y[n]. \tag{8}$$

Typically, the linear regressor in a SVM framework includes a bias term: $y[n] = \mathbf{w}^{\mathrm{T}}\mathbf{x}[n] + b$. For this particular problem, however, $b = 0$ is needed; otherwise, the trivial solution $\mathbf{w} = \mathbf{0}$ and $b = 1$ would be always obtained.

IV. ITERATIVE REWEIGHTED QP

The problem formulated in the previous section cannot be directly solved in a single step, because the weighted Lagrange multipliers depend on $y[n]$ (see (8)). Consequently, an iterative procedure must be applied to find the solution. Here we use the algorithm introduced in [8]: it is called the Iterative Reweighted Quadratic Programming (IRWQP) algorithm, because its similarity with the Iterative Reweighted Least Squares (IRWLS) algorithm used in some approximation and regression problems [14].

The IRWQP method is summarized in three steps:

1) Solve the QP problem (6), assuming $y[n]$ fixed.
2) Obtain new beamformer coefficients with (7) and update the output $y[n]$.
3) Repeat until convergence.

The algorithm is completed with a smoothing of the beamformer coefficients, iteration from iteration. In this way, we avoid a possible oscillation of the output values between

$y[n]$ and $1/y[n]$. In particular, the beamformer coefficients at iteration k are obtained as

$$\mathbf{w}_k = \lambda \mathbf{w}_{k-1} + (1-\lambda)\mathbf{w}_{\text{QP}},$$

where $\mathbf{w}_{\text{QP}}$ are the coefficients obtained from the QP problem at the k-th iteration, and λ is a smoothing parameter.

V. Blind separation of multiple CM signals

In this section the simultaneous separation of P CM signals is considered. The algorithm consists of two stages; in the *initialization* stage each beamformer roughly extracts a different CM source. In the *convergence* stage each beamformer continues iterating as in the single source method discussed in Section IV, refining the solutions obtained in the previous stage.

For a correct initialization, the signals extracted by the previous beamformers $1,\ldots,k-1$ are subtracted from the observations at the input of the k-th beamformer. In this way, the k-th beamformer cannot extract a CM source already extracted by any of the previous $1,\ldots,k-1$ beamformers. In particular, denoting the output at the k-th beamformer as

$$\mathbf{y}_k = \left[y_k[0], y_k[1], \ldots, y_k[N-1]\right]^{\mathrm{T}},$$

the input data matrix for the $(k+1)$-th beamformer (during the initialization stage) is the orthogonal projection of the original data matrix onto the complementary subspace of the extracted CM sources,

$$\mathbf{X}_{k+1} = \mathbf{X}\left[\mathbf{I} - \mathbf{Y}_k^H \left(\mathbf{Y}_k \mathbf{Y}_k^H\right)^{-1} \mathbf{Y}_k\right], \tag{9}$$

where

$$\mathbf{Y}_k = \left[\mathbf{y}_1, \mathbf{y}_2, \ldots, \mathbf{y}_k\right]^{\mathrm{T}}.$$

After a number of iterations of the initialization stage, each beamformer roughly extracts one of the CM signals. The convergence stage begins at this point. In this final stage the signal cancelling mechanism is not used, and each beamformer carries out a number of iterations independently from each other.

Finally, the proposed method for the simultaneous separation of P CM signals is summarized in Algorithm 1.

Initialize C, ϵ, λ and $\mathbf{w}_{0,k}$.
Initialization stage
for $p = 1, 2, \ldots,$ niter1 **do**
  for $k = 1, 2, \ldots, P$ **do**
    Compute $\mathbf{y}_k$ for $\mathbf{w}_{p-1,k}$ and $\mathbf{X}_k$.
    Solve QP problem (6) and obtain $\mathbf{w}_{\text{QP}}$.
    $\mathbf{w}_{p,k} = \lambda \mathbf{w}_{p-1,k} + (1-\lambda)\mathbf{w}_{\text{QP}}$.
    Compute $\mathbf{y}_k$ for $\mathbf{w}_{p,k}$, and $\mathbf{X}_{k+1}$ with (9).
  end for
end for
Convergence stage
for $p = 1, 2, \ldots,$ niter2 **do**
  for $k = 1, 2, \ldots, P$ **do**
    Compute $\mathbf{y}_k$ for $\mathbf{w}_{p-1,k}$ and $\mathbf{X}$.
    Solve QP problem (6) and obtain $\mathbf{w}_{\text{QP}}$.
    $\mathbf{w}_{p,k} = \lambda \mathbf{w}_{p-1,k} + (1-\lambda)\mathbf{w}_{\text{QP}}$.
  end for
end for

Algorithm 1: Summary of the SVM-based blind beamforming algorithm.

VI. Simulation results

In this section the proposed method is compared with the ACMA [6]. An example with four signals is considered. Three of the signals are QPSK (constant-modulus) and the other is Gaussian distributed. The signals impinge on an uniform array of omnidirectional antennas. The angles of arrival are: $0°$, $30°$ and $60°$ for the QPSK signals, and $-20°$ for the Gaussian signal. The signal to noise ratio is 15 dB for every signal. Examples with a different number of snapshots are considered, $N = 8, 10, 13, 16, 25, 50$. For each number of collected snapshots the results of 300 independent simulations are averaged.

For the beamforming method based on SVM, the following values are chosen: $\epsilon = 0.01$ (Vapnik's function parameter) and $\lambda = 0.3$ (smoothing factor). The regularization parameter C is estimated according to

$$C = \frac{1}{M}\sum_{m=1}^{M}\left(\overline{g_m[n]} + 3\sigma_{g_m}\right),$$

where $g_m[n] = \|x_m[n]\|^2$ and $\overline{g_m[n]}$ denotes the mean. As discussed in [15], this choice increases the robustness of the regression method. Finally, a maximum of 5 iterations of the IRWQP procedure described in Section IV, were applied for the initialization and convergence stages.

As a figure of merit for the extracted signals, we have used the Average Modulus Error (AME), which measures the deviation over the ideal CM property of the sources and is defined as

$$\left[\frac{1}{N}\sum_{n=0}^{N-1}\left(\|y_k[n]\| - 1\right)^2\right]^{1/2}.$$

Figure 1 shows the evolution of the AME for each source (in logarithmic scale) versus the number of iterations for one realization of the IRWQP procedure. The first 5 iterations correspond to the initialization stage, whereas the last 5 correspond to the convergence stage: the final AME of the second and third sources is reduced during the convergence stage.

Although we have not been able to theoretically prove the convergence of the proposed procedure yet, in all the examples the algorithm always converged to a solution for which the AME was a minimum. Note, however, that a minimum of the AME does not necessarily mean that one of the original sources has been extracted. This point is further clarified in Fig. 2 that shows the probability of correct signal extraction as a function of the number of snapshots for the blind SVM-based method and the ACMA. For both methods an extracted signal is considered correct if the quadratic error to signal power ratio is

$$\frac{\|\mathbf{s_k} - \tilde{\mathbf{y}}_\mathbf{k}\|_2^2}{\|\mathbf{s_k}\|_2^2} < 0.5,$$

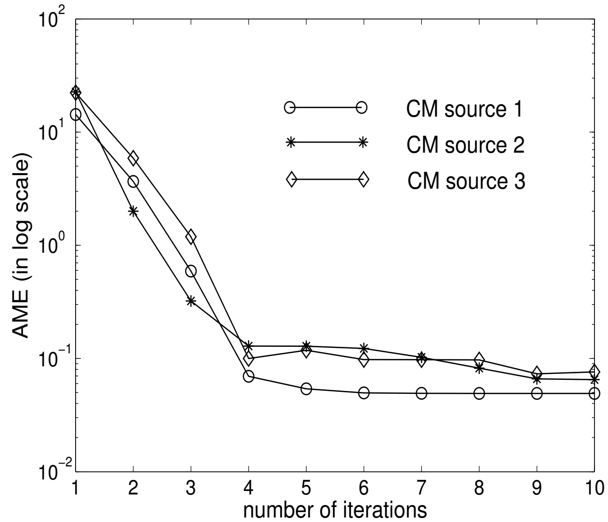

Fig. 1. AME vs. number of iterations for the IRWQP procedure. N=50, SNR=15dB.

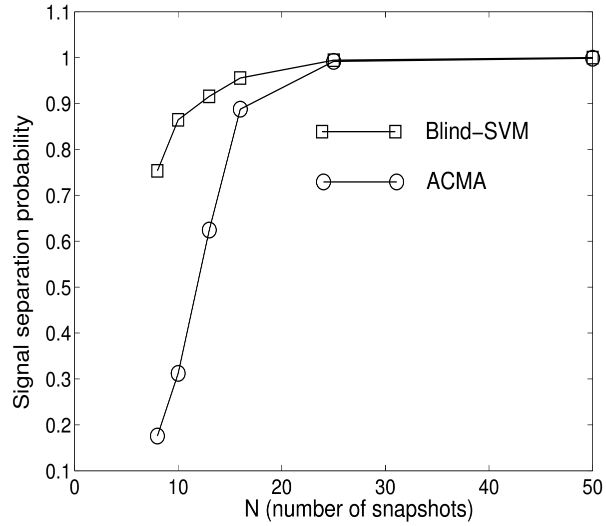

Fig. 2. Correct extraction probability in an example with three QPSK signals and one with non-constant modulus, SNR = 15dB.

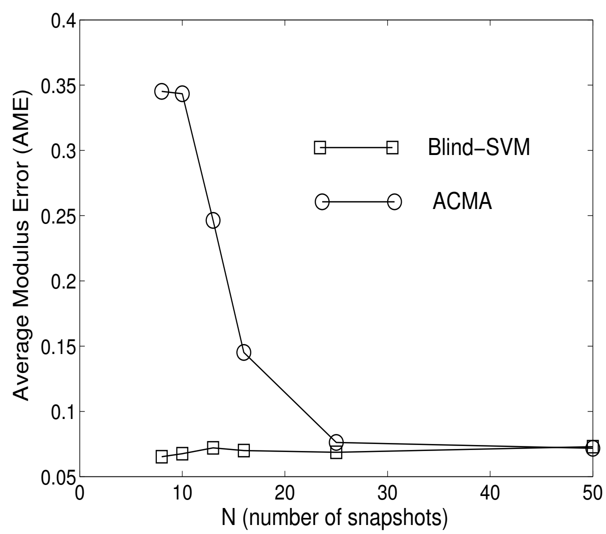

Fig. 3. Average modulus error (AME) in an example with three QPSK signals and one non-constant modulus signal, SNR = 15dB; as a function of the number of snapshots N.

where $\mathbf{s_k}$ is the original QPSK signal, and $\tilde{\mathbf{y}}_\mathbf{k}$ is a scaled version of $\mathbf{y_k}$ that minimizes the quadratic error with respect to $\mathbf{s_k}$. The scale factor applied to $\mathbf{y_k}$ is needed due to the phase ambiguity (i.e. a rotation of the constellation) inherent to blind equalization and beamforming problems.

From Fig. 2 we can see that the proposed blind-SVM method provides a better extraction probability than the ACMA, mainly when the number of snapshots is small ($N < 25$, which is a moderate number for array processing applications).

Finally, Fig. 3 depicts the mean value of the final AME for the three sources versus the number of snapshots for the ACMA and blind-SVM approaches: again, for examples with very few snapshots the proposed method achieves a much lower AME than the ACMA.

It must be noticed, however, that these improvements over the ACMA are obtained in exchange for an increase in computational cost. To elaborate on this point, let us remark that the ACMA requires two SVD's, the first one of an $M \times N$ matrix and the second one of an $(N-1) \times L^2$ matrix. Moreover, the ACMA solves a simultaneous diagonalization problem of P matrices of dimension $L \times L$ [6][1]. On the other hand, the proposed method requires to solve $(niter1 + niter2) \times P$ a quadratic programming problem of size N, where $niter1$ and $niter2$ are the number of iterations of the IRWQ procedure for the initialization and convergence stages, respectively ($niter1 + niter2 = 10$ for this example). Using the Matlab SVM toolbox available in [13] to solve the QP problem at each iteration, the proposed approach requires roughly five more times of execution time (without any code optimization) than the ACMA. The application of some efficient techniques recently proposed to solve the QP problem [16], [17], [18] can be useful to alleviate this drawback .

[1]Remember that M is the number of antennas, N is the number of snapshots, L is the number of sources and $P \leq L$ is the number of CM signals

VII. Conclusions

In this paper, blind beamforming for CM signals problem has been formulated as a regression problem and a SVM based technique has been applied to solve it. An iterative algorithm has been proposed; it converges to one of the constant-modulus signals present at the observations. This method has been used to elaborate a new method with a number of beamformers working in parallel for the extraction of multiple CM signals.

Simulation results show that the proposed method offers a better performance than the ACMA, mainly in the cases with a small number of snapshots, which is of interest in wireless as well as in passive sonar applications. The proposed algorithm, however, has a high computational cost. The application of recently proposed efficient techniques for solving QP problems [16], [17], or the iterative reweighted least squares procedure described in [18] can be useful for reducing this drawback .

Finally, another advantage of the proposed procedure is that it can readily be extended to nonlinear blind beamforming by performing the linear regression in another space of higher dimension by the so-called kernel trick [12]. The nonlinear beamforming problem, as well as the development of "on-line" adaptive versions of the proposed batch procedure are currently under investigation.

VIII. Acknowledgements

This work was supported in part by MCYT (Ministerio de Ciencia y Tecnología) under grant TIC2001-0751-C04-03.

References

[1] B. D. Van Veen and K. M. Buckley, "Beamforming: A versatile approach to spatial filtering," *IEEE Signal Proc. Magazine*, vol.7, no. 3, pp. 4-24, April 1988.

[2] J. R. Treichler and B. G. Agee, "A new approach to multipath correction of constant modulus Signals," *IEEE Trans. on Acoust., Speech, Signal Processing*, vol. ASSP-31, pp. 459-471, April 1983.

[3] R. Gooch and J. Lundell, "The CM array: An adaptive beamformer for constant modulus signals," in *Proc. ICASSP'1986*, pp. 2523-2526, Tokyo, 1986.

[4] J. Shynk and R. Gooch, "The constant modulus array for cochannel signal copy and direction finding," *IEEE Trans. on Signal Processing*, vol. 44, no. 3, pp. 652-660, March 1996.

[5] B. G. Agee, "The least-squares CMA: A new technique for rapid correction of constant modulus signals," in *Proc. ICASSP'1986*, pp. 953-956, Tokyo, 1986.

[6] A. J. Van der Veen and A. Paulraj, "An analytical constant modulus algorithm," *IEEE Trans. on Signal Processing*, vol. 44, no. 5, May 1996.

[7] H. Song, W. A. Kuperman, W. S. Hodgkiss, P. Gerstoft, J. S. Kim, "Null broadening with snapshot-deficient covariance matrices in passive sonar," *IEEE J. Oceanic Eng.*, vol. 28, no. 2, April 2003.

[8] I. Santamaría, C. Pantaleón, L. Vielva and J. Ibáñez, "Blind equalization of constant modulus signals via support vector regression," in *Proc. ICASSP'2003*, pp. 737-740, Hong Kong, China, May 2003.

[9] I. Santamaría, C. Pantaleón, L. Vielva and J. Ibáñez, "Blind equalization of constant modulus signals using support vector machines," to be published in *IEEE Trans. on Signal Processing*, 2004.

[10] I. Santamaría, J. Vía, C.C. Gaudes, "Robust blind identification of SIMO channels: A support vector regression approach," in *Proc. ICASSP'2004*, Montreal, Canada, May 2004.

[11] V. Vapnik, *The Nature of Statistical Learning Theory*, Springer Verlag, New York, 1995.

[12] B. Schölkopf, A. Smola, *Learning with Kernels*, Cambridge, MA: The MIT Press, 2002.

[13] S. R. Gunn, *Matlab Support Vector Machine Toolbox*, available at `http://www.isis.ecs.soton.ac.uk/isystems/kernel/`, University of Southampton, Image Speech and Intelligent Systems Research Group, UK, 1998.

[14] P. W. Holland and R. E. Welch, "Robust regression using iterative reweighted least squares," *Commun. Statist. Theory Methods*, vol. A 6, n 9, pp. 813-827, 1997.

[15] V. Cherkassky and Y. Ma, "Selection of meta-parameters for support vector regression," in J. R. Dorronsoro (Ed.), *Proc. of ICANN'2002*, pp. 687-693, Springer-Verlag, Berlin, 2002.

[16] J. Platt, "Fast training of support vector machines using sequential minimal optimization," in B. Schlkopf, C. J. C Burges and A. J. Smola, (Eds.), *Advances in Kernel Methods: Support Vector Learning*, pp. 185-208, The MIT Press, Cambridge, MA, 1999.

[17] T. Joachims, "Making large-scale support vector machine learning practical," in B. Schlkopf, C. J. C. Burges and A. J. Smola, (Eds.), *Advances in Kernel Methods: Support Vector Learning*, pp. 169-184, The MIT Press, Cambridge, MA, 1999.

[18] F. Pérez-Cruz, A. Navia-Vázquez, P. Alarcón-Diana, A. Artés-Rodríguez, "An IRWLS procedure for SVR", in *Proc. EUSIPCO 2000*, Tampere, Finland, Sept. 2000.