Nonuniform Interpolation of Noisy Signals Using Support Vector Machines

José Luis Rojo-Álvarez, *Member, IEEE*, Carlos Figuera-Pozuelo,
Carlos Eugenio Martínez-Cruz, *Student Member, IEEE*, Gustavo Camps-Valls, *Member, IEEE*,
Felipe Alonso-Atienza, *Student Member, IEEE*, and Manel Martínez-Ramón, *Senior Member, IEEE*

Abstract—The problem of signal interpolation has been intensively studied in the Information Theory literature, in conditions such as unlimited band, nonuniform sampling, and presence of noise. During the last decade, support vector machines (SVM) have been widely used for approximation problems, including function and signal interpolation. However, the signal structure has not always been taken into account in SVM interpolation. We propose the statement of two novel SVM algorithms for signal interpolation, specifically, the primal and the dual signal model based algorithms. Shift-invariant Mercer's kernels are used as building blocks, according to the requirement of bandlimited signal. The sinc kernel, which has received little attention in the SVM literature, is used for bandlimited reconstruction. Well-known properties of general SVM algorithms (sparseness of the solution, robustness, and regularization) are explored with simulation examples, yielding improved results with respect to standard algorithms, and revealing good characteristics in nonuniform interpolation of noisy signals.

Index Terms—Dual signal model, interpolation, Mercer's kernel, nonuniform sampling, primal signal model, signal, sinc kernel, support vector machine (SVM).

I. INTRODUCTION

SIGNAL interpolation is a widely studied research area [1]–[4]. The interpolation in the information and communication era has its roots on Sampling Theory, and specifically, on the Whittaker-Shannon-Kotel'nikov (WSK) equation, also known as Shannon's sampling theorem [5], [6], which states that a bandlimited, noise free signal can be reconstructed from a uniformly sampled sequence of its values, assumed that the

sampling period is properly chosen according to the signal bandwidth. The nonuniform sampling of a bandlimited signal can also be addressed whenever the average sampling period still fulfills Shannon's sampling theorem, and it is used in a number of applications, such as approximation of geophysical potential fields, tomography, and synthetic aperture radar (SAR) [7]–[9]. Given that noise can often be present, the reconstruction of a bandlimited and noise corrupted signal from its nonuniformly sampled observations becomes a hard problem. According to [10], two strategies have been mainly followed: (1) consideration of shift-invariant spaces, similar to the case of uniform sampling; and (2) definition of new basis functions (or new spaces) that are better suited to the nonuniform structure of the problem. The first one has been studied the most, following the work developed in the late fifties by Yen [11] using the sinc function as an interpolation kernel. Although, in the theory, Yen's interpolator is optimal in the least squares (LS) sense, ill-posing appears when computing the interpolated values numerically [8]. To overcome this limitation, numerical regularization has been widely used [12]. Alternatively, a number of iterative methods have been proposed, including alternating mapping, projections onto convex sets, and conjugate gradient [7], [13]–[15]. Other authors have used noniterative methods, such as filter banks, either to reconstruct the continuous time signal, or to interpolate to uniformly spaced samples [16], [17], [3], but none of these methods is optimal in a LS sense, and thus many approximate forms of the Yen's interpolator have been developed [18], [19]. The previously mentioned methods have addressed the reconstruction of bandlimited signals, but the question of whether a signal that is not strictly bandlimited can be recovered from its samples has emerged. Specifically, a finite set of samples from a continuous-time function can be seen as a duration-limited discrete-time signal in practice, and then it cannot be bandlimited. In this case, the reconstruction of a signal from its samples depends on the *a priori* information that we have, and the classical sinc kernel has been replaced by more general kernels that are not necessarily bandlimited [2], [3]. This issue has been mostly studied in problems of uniformly sampled time series.

As a summary, the following main elements are (either implicitly or explicitly) considered by signal interpolation algorithms: the kind of sampling (uniform or nonuniform), the noise (present or not), the spectral content (bandlimited or not), and the use (or not) of numerical regularization. But, in spite of the great amount of work developed to date, the search for new efficient interpolation procedures is still an active research area.

In this context, we propose the use of support vector machines (SVM) as a signal interpolator in the presence of noise.[1] SVM were originally stated for classification and regression problems. However, the consideration of different signal models (equation that relates the observation and the data according to a given signal structure for both) has allowed to extend the formulation of SVM algorithms to a number of digital signal processing problems, which are in essence very different from a classification and regression model structure [20]–[22]. SVM algorithms exploit the structural risk minimization (SRM) principle to regularize the model, and use the rather old kernel trick to easily build nonlinear models from linear ones [23]. The SRM principle states that a better solution (in terms of generalization capabilities) can be found by minimizing an upper bound of the generalization error. This minimization constitutes a Tikhonov regularization [24] that, in turn, yields the least possible complexity to the resulting machine. As a result, SVM commonly exhibit less overfitting than other classical models developed under the empirical risk minimization (ERM) principle. Also, SVM are robust against outliers and impulse noise, due to its cost function of the residuals [21]. Interestingly, such a procedure produces sparse solutions,[2] which can dramatically reduce the computational burden of the solution in its application stage and, at the same time, it enables to express the solution as a linear combination of the most relevant samples in the problem (the so-called support vectors).

Therefore, understanding the advantages of the above mentioned SVM algorithms gives us the key for making new (robust and sparse) algorithms for signal interpolation. On the one hand, SVM have been previously used for interpolation applications, but key and basic concepts from Information Theory, such as the bandwidth or the kind of sampling, have not been taken into account, so that little connection has been established with the wide existing work in time series interpolation. On the other hand, sparseness and robustness would be extremely useful in the hard problem of interpolation of noisy, possibly nonuniformly sampled, time series. Additionally, the bandlimited nature of the SVM interpolation can be readily controlled by the Mercer's kernel that is being used. The Gaussian [or radial basis function (RBF)] Mercer's kernel is not a bandlimited function, and hence, it is appropriate for interpolation of nonbandlimited time series. Alternatively, it has been proven that the sinc kernel, when adequately expressed, lies in a reproducing kernel Hilbert space (RKHS), and hence, it can be used as a Mercer's kernel in SVM interpolation of bandlimited time series [26], [27]. The RBF kernel has been widely studied in the SVM literature, but this is not the case of the sinc kernel, which has received little attention in this setting. As a result, the study and introduction of SVM methods in this context is well motivated and founded.

In particular, we present here two novel SVM interpolation algorithms. The first algorithm uses a *primal signal model* for-

mulation of the problem, according to the SVM linear framework for digital signal processing presented in [20], [22], in which a robust estimation of model coefficients is indirectly obtained from the SVM Lagrange multipliers. Note that the SVM linear framework tells how to state robust signal processing algorithms in general, but the algorithm for signal interpolation from a primal signal model formulation has not yet been addressed. The second algorithm, or *dual signal model* formulation, uses a nonlinear regression in a RKHS of the time instant corresponding to each observed sample, in such a way that when the solution is expressed in terms of dot products in the RKHS, it can readily be expressed by means of a Mercer's kernel. Though SVM have been previously used for signal reconstruction, the dual signal model has not been explicitly expressed as the nonlinear transformation of the independent variable and analyzed from an Information Theory point of view. For the purpose of comparison with precedent Information Theory based formulations, we develop the SVM algorithms in terms of the sinc kernel, the extension of the notation to RBF kernel being straightforward. Also, in the experiments section we compare the family of SVM methods with well-known methods for signal interpolation, such as Yen's, Jacob's, and minimax algorithms. The SVM methods show good results in a wide range of scenarios.

The scheme of the paper is as follows. Section II briefly introduces Yen's algorithm and some improved (regularized) versions to work in ill-posed problems. Then, we present the formulation of two novel SVM algorithms for signal interpolation. Section III presents the simulation results when comparing with Yen's algorithm and with some representative simplified versions. Finally, in Section IV, discussion and conclusions are given.

II. SVM FOR NONUNIFORM INTERPOLATION

As mentioned in Section I, a wide variety of methods have been proposed for time series interpolation. Among all the available algorithms for reconstruction of nonuniformly sampled time series, the sinc kernel has received special attention, and accordingly, this implies that a bandlimited nature of the time series is assumed. We limit ourselves to briefly present here the basics of Yen's algorithm and some improved versions [11], [28].

Let $x(t)$ be a possibly not bandlimited signal corrupted by Gaussian noise, and let $\{x_n = x(t_n), n = 1, \ldots, N\}$ be a set of N nonuniformly sampled observations. Given $\{t_n, x_n; n = 1, \ldots, N\}$, the interpolation problem consists of finding an approximating function $x^N(t)$ that fits the observations as follows:

$$x(t) = x^N(t) + e(t) = \sum_{i=1}^{N} a_i \mathrm{sinc}(\sigma_0(t - t_i)) + e(t) \quad (1)$$

where $\mathrm{sinc}(t) = \sin(t)/t$; $\sigma_0 = \pi/T_0$ is the bandwidth of the interpolating units (and in general has to be determined from some *a priori* knowledge or search strategy); and $e(t)$ represents the noise. The previous continuous time series model, after

---

[1]Without loss of generality, we will use the problem statement equations for discrete-time series interpolation, though the results can be readily extended to generic real-valued signal approximation problems

[2]Hereafter, sparseness is defined in terms of number of bases associated to training samples rather than with regard to the number of bases in a certain *hidden space* as in the case of neural networks, where good results have been recently obtained [25].

nonuniform sampling, is expressed as the following discrete time model:

$$x_n = x_n^N + e_n = \sum_{i=1}^{N} a_i \mathrm{sinc}(\sigma_0(t_n - t_i)) + e_n. \quad (2)$$

An optimal bandlimited interpolation algorithm, in the least squares (LS) sense, was first proposed by Yen [11]. The problem can be expressed as the minimization of the quadratic loss function, given by

$$\frac{1}{2}\sum_{n=1}^{N}\left(x_n - \sum_{i=1}^{N} a_i \mathrm{sinc}(\sigma_0(t_n - t_i))\right)^2 \quad (3)$$

which, in matrix notation, consists of minimizing

$$\frac{1}{2}\|\boldsymbol{x} - \boldsymbol{S}\boldsymbol{a}\|^2 \quad (4)$$

where $\boldsymbol{a} = [a_1, \ldots, a_N]^T$ is the vector of model coefficients, $\boldsymbol{x} = [x_1, \ldots, x_N]^T$, and $\boldsymbol{S}$ is a square matrix whose elements are

$$\boldsymbol{S}(n, m) = \mathrm{sinc}(\sigma_0(t_n - t_m)). \quad (5)$$

It can be seen that the solution vector is

$$\boldsymbol{a} = \boldsymbol{S}^{-1}\boldsymbol{x}. \quad (6)$$

This is a critically determined problem, as we have as many free parameters as observations, and in the presence of noise this yields an ill-posed problem [12]. In fact, in the presence of even a low level of noise, small perturbations on the coefficient estimations lead to large interpolation errors outside the observed samples. To overcome this limitation, the regularization of the quadratic loss has been proposed [11], thus leading to a different problem that consists of minimizing

$$\frac{1}{2}\|\boldsymbol{x} - \boldsymbol{S}\boldsymbol{a}\|^2 + \frac{\delta}{2}\|\boldsymbol{a}\|^2 \quad (7)$$

where δ is a regularization parameter, which represents the tradeoff between losses and smoothness of the solution. The regularized solution is

$$\boldsymbol{a} = (\boldsymbol{S}^2 + \delta \boldsymbol{I})^{-1}\boldsymbol{S}\boldsymbol{x} \quad (8)$$

where $\boldsymbol{I}$ is the $N \times N$ identity matrix. Note that the bandwidth (in both approaches), as well as the tradeoff parameter (in the second approach), must be previously fixed.

Note that instead of using the sinc function as the interpolation kernel for bandlimited interpolation, this formulation can be extended to other nonbandlimited basis functions, such as the Gaussian kernel, given by

$$g(t_k - t_n) = \exp\left(\frac{-|t_k - t_n|^2}{2\sigma^2}\right) \quad (9)$$

where σ is the kernel free parameter. Also, polynomial functions can be readily used (see, e.g., [29], [30]), which have been subsequently analyzed in terms of Information Theory principles. An excellent review can be found in [4].

A. SVM Robust Cost Function

In this section, we propose to use several SVM approaches for estimating efficiently coefficients $\{a_i\}$ in signal model (2). In the SVM framework for digital signal processing [21], the optimality criterion is a regularized and constrained version of the regularized LS criterion. Residuals $\{e_n\}$ account for the effect of both noise and model approximation errors. In general, SVM algorithms minimize a regularized cost function of the residuals, usually the Vapnik's ε-insensitivity cost function [23]. Alternatively, we introduce in the formulation the ε-Huber robust cost function [21], which is given by

$$\mathcal{L}^{\varepsilon}(e_n) = \begin{cases} 0, & |e_n| \leq \varepsilon \\ \frac{1}{2\gamma}(|e_n| - \varepsilon)^2 & \varepsilon \leq |e_n| \leq e_C \\ C(|e_n| - \varepsilon) - \frac{1}{2}\gamma C^2 & |e_n| \geq e_C. \end{cases} \quad (10)$$

Here, parameter ε is a nonnegative scalar that represents the insensitivity to a low noise level, but more relevant for us, it can provide with a sparse solution, which can be a highly desirable property in the model when working in test mode. Parameters γ and C represent the relevance of the residuals that are in the quadratic or in the linear cost zone, respectively. It can be easily seen that $e_C = \gamma C$ for a residual cost function with continuous first derivative. By an adequate choice of free parameters ε, γ, C, the ε-Huber cost function can be adapted to different kinds of noise while allowing sparse solutions. The function to be minimized by SVM regression consists of a residual cost term plus a regularization term, given by the L_2-norm of the model parameters [23].

For the signal model in (2), there are two possible SVM formulations, which are described next. The first one consists of using signal model (2) as the primal problem in the SVM formulation. The second one consists of considering a generic nonlinear SVM regression, obtaining the SVM dual and solution equations for a generic Mercer's kernel, and finally introducing the sinc kernel as a Mercer's kernel for obtaining signal model (2) in the dual solution.

B. Primal Signal Model Algorithm

The nonuniform signal model in (2) can be used within the SVM linear framework [21]. In this case, the signal model to be considered is given in (2). But instead of following the LS criterion, we consider the minimization of the ε-Huber robust cost in (10), together with the quadratic norm of model coefficients $\{a_i\}$, which can be seen as a regularization term. This is, in this first approach we minimize

$$\sum_{n=1}^{N} \mathcal{L}^{\varepsilon}(e_n) + \frac{1}{2}\|\boldsymbol{a}\|^2. \quad (11)$$

As usual in SVM linear framework, the optimization of this regularized robust cost can be achieved [21] by minimizing

$$\frac{1}{2}\sum_{k=1}^{N} a_k^2 + \frac{1}{2\gamma}\sum_{n \in I_1}\left(\xi_n^2 + \xi_n^{*2}\right)$$

$$+ C\sum_{n \in I_2}\left(\xi_n + \xi_n^*\right) - \frac{\gamma C^2 |I_2|}{2} \quad (12)$$

constrained to

$$x_n - \sum_{i=1}^{N} a_i \mathrm{sinc}(\sigma_0(t_n - t_i)) \leq \varepsilon + \xi_n \qquad (13)$$

$$-x_n + \sum_{i=1}^{N} a_i \mathrm{sinc}(\sigma_0(t_n - t_i)) \leq \varepsilon + \xi_n^* \qquad (14)$$

$$\xi_n, \xi_n^* \geq 0 \qquad (15)$$

for $n = 1, \ldots, N$, and where ξ_n, ξ_n^*, are slack variables or losses; $I_1(I_2)$ are the indices of residuals that are in the quadratic (linear) cost zone; and $|I_2|$ is the cardinality of set I_2. The solution to this optimization problem is given by the saddle point of the following Lagrangian function:

$$\frac{1}{2} \sum_{k=1}^{N} a_k^2 + \frac{1}{2\gamma} \sum_{n \in I_1} \left(\xi_n^2 + \xi_n^{*2}\right) + C \sum_{n \in I_2} (\xi_n + \xi_n^*)$$
$$- \sum_{n=1}^{N} \alpha_n \left(-x_n + \sum_{i=1}^{N} a_i \mathrm{sinc}(\sigma_0(t_n - t_i)) + \varepsilon + \xi_n\right)$$
$$- \sum_{n=1}^{N} \alpha_n^* \left(x_n - \sum_{i=1}^{N} a_i \mathrm{sinc}(\sigma_0(t_n - t_i)) + \varepsilon + \xi_n^*\right)$$
$$- \sum_{n=1}^{N} (\beta_n \xi_n + \beta_n^* \xi_n^*) - \frac{\gamma C^2 |I_2|}{2} \qquad (16)$$

where $\{\alpha_n^{(*)}\}(\{\beta_n^{(*)}\})$ are the Lagrange multipliers corresponding to (13) and (14) (to (15)). Lagrangian duality enables the primal problem to be transformed into its dual one, by taking the derivative of (16) with respect to the primal variables. It is straightforward to show that if we denote

$$T(k, m) = \sum_{n=1}^{N} \mathrm{sinc}(\sigma_0(t_n - t_k)) \mathrm{sinc}(\sigma_0(t_n - t_m)) \qquad (17)$$

then the dual problem consists of maximizing

$$-\frac{1}{2}(\boldsymbol{\alpha} - \boldsymbol{\alpha}^*)^T \boldsymbol{T}(\boldsymbol{\alpha} - \boldsymbol{\alpha}^*) + (\boldsymbol{\alpha} - \boldsymbol{\alpha}^*)^T \boldsymbol{x} - \varepsilon \mathbf{1}^T(\boldsymbol{\alpha} + \boldsymbol{\alpha}^*)$$
$$-\frac{\gamma}{2}(\boldsymbol{\alpha}^T \boldsymbol{I}\boldsymbol{\alpha} + \boldsymbol{\alpha}^* \boldsymbol{I}\boldsymbol{\alpha}^*) \qquad (18)$$

constrained to $0 \leq \alpha_n^{(*)} \leq C$, where $\boldsymbol{\alpha}^{(*)} = [\alpha_1^{(*)}, \ldots, \alpha_N^{(*)}]$, and **1** denotes a column vector of ones. This minimization problem can be solved with quadratic programming techniques [31]. Once dual coefficients $\{\alpha_i^{(*)}\}$ are obtained, primal coefficients $\{a_j\}$ are given by

$$a_j = \sum_{i=1}^{N} (\alpha_i - \alpha_i^*) \mathrm{sinc}(\sigma_0(t_j - t_i)). \qquad (19)$$

Note that coefficients are proportional to the empirical cross correlation of the Lagrange multipliers and a set of sinc basis functions, each centered in time instant t_n (see [21] for a related discussion on generic primal signal models).

C. Dual Signal Model Algorithm

A second SVM-based version of the interpolation function can be obtained by starting with a conventional SVM nonlinear

regression [23]. In this setting, given observations $\{x_n\}$ at time instants $\{t_n\}$, we map these time instants to a higher dimensional (H, possibly infinity) feature space $\mathcal{H}$ by using a nonlinear transformation $\boldsymbol{\phi}$, this is, we consider $\boldsymbol{\phi} : \mathbb{R} \to \mathcal{H}$ that maps $t \in \mathbb{R} \to \boldsymbol{\phi}(t) \in \mathcal{H}$, where a linear approximation to the data can properly fit the observations as follows:

$$x_n = x_n^N + e_n = \langle \boldsymbol{w}, \boldsymbol{\phi}(t_n) \rangle + e_n \qquad (20)$$

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

The optimization criterion is in this case

$$\sum_{n=1}^{N} \mathcal{L}^\varepsilon(e_n) + \frac{1}{2} \|\boldsymbol{w}\|^2. \qquad (21)$$

Note that in this case the regularization term is not referred to the amplitude of the base functions of the model as in (11), but rather to the regression vector in the RKHS. The primal problem consists now of minimizing

$$\frac{1}{2} \sum_{k=1}^{H} w_k^2 + \frac{1}{2\gamma} \sum_{n \in I_1} \left(\xi_n^2 + \xi_n^{*2}\right)$$
$$+ C \sum_{n \in I_2} (\xi_n + \xi_n^*) - \frac{\gamma C^2 |I_2|}{2} \qquad (22)$$

constrained to

$$x_n - \langle \boldsymbol{w}, \boldsymbol{\phi}(t_n) \rangle \leq \varepsilon + \xi_n \qquad (23)$$
$$-x_n + \langle \boldsymbol{w}, \boldsymbol{\phi}(t_n) \rangle \leq \varepsilon + \xi_n^* \qquad (24)$$
$$\xi_n, \xi_n^* \geq 0. \qquad (25)$$

Again, a Lagrange functional can be stated by following a similar procedure to the precedent section. In brief, by taking the gradient, we now obtain

$$\boldsymbol{w} = \sum_{n=1}^{N} (\alpha_n - \alpha_n^*) \boldsymbol{\phi}(t_n). \qquad (26)$$

After substitution of $\boldsymbol{w}$ into the Lagrangian and some simple manipulations, the following Gram matrix can be identified:

$$\boldsymbol{G}(k, m) = \langle \boldsymbol{\phi}(t_k), \boldsymbol{\phi}(t_m) \rangle = K(t_k, t_m) \qquad (27)$$

where $K(t_k, t_m)$ is a Mercer's kernel, which allows to obviate the explicit knowledge of nonlinear mapping $\boldsymbol{\phi}(\cdot)$ [23]. The dual problem consists now of maximizing

$$-\frac{1}{2}(\boldsymbol{\alpha} - \boldsymbol{\alpha}^*)^T \boldsymbol{G}(\boldsymbol{\alpha} - \boldsymbol{\alpha}^*) + (\boldsymbol{\alpha} - \boldsymbol{\alpha}^*)^T \boldsymbol{x} -$$
$$-\varepsilon \mathbf{1}^T(\boldsymbol{\alpha} + \boldsymbol{\alpha}^*) - \frac{\gamma}{2} \left(\boldsymbol{\alpha}^T \boldsymbol{I}\boldsymbol{\alpha} + \boldsymbol{\alpha}^* \boldsymbol{I}\boldsymbol{\alpha}^*\right) \qquad (28)$$

constrained to $0 \leq \boldsymbol{\alpha}^{(*)} \leq C$. Note the similarity with the dual problem of the primal signal model formulation in (18), in which $\boldsymbol{T}$ is replaced with kernel matrix $\boldsymbol{G}$. The final solution is expressed as

$$x_n^N = \sum_{i=1}^{N} (\alpha_i - \alpha_i^*) K(t_i, t_n). \qquad (29)$$

As usual in the SVM framework, by setting $\varepsilon > 0$, we obtain that only a subset of the Lagrange multipliers will be nonzero, thus, providing with a sparse solution. The associated samples are called *support vectors* (SVs) and represent a set of very relevant samples in the data distribution, as the solution is built only in terms of them. Moreover, we can define at this point

$$K(t_k, t_n) = \text{sinc}(\sigma_0(t_k - t_n)) \tag{30}$$

which can be proven to be a Mercer's kernel (see, e.g., [32]). We will call this choice the *sinc Mercer's kernel*. Therefore, when using the sinc Mercer's kernel, (29) is the nonuniform interpolation model given in (2) with $a_j = \alpha_j - \alpha_j^*$.

Note that other Mercer's kernels could be easily used with this approach. Function (9) can be expressed as a valid Mercer's kernel by simply defining it as

$$K(t_k, t_n) = \exp\left(\frac{-|t_k - t_n|^2}{2\sigma^2}\right) \tag{31}$$

thus giving a nonbandlimited (kernel-based) Gaussian interpolator.

As usual in the SVM framework, the free parameter of the kernel and the free parameters of the cost function have to be fixed by some *a priori* knowledge of the problem, or by using some validation set of observations. This issue is analyzed in the experiments section.

Note that there are two kinds of free parameters in the SVM algorithms: kernel parameter σ_0 and residual cost function parameters (ε, δ, C). In SVM, it is not possible to give general and closed expressions for the mentioned free parameters, but rather they have to be either searched by using a validation set, or fixed according to some *a priori* knowledge. However, the other interpolation algorithms described here have also free parameters to be tuned, as (σ_0, δ for Yen's algorithm in (8), which has also to be fixed from some *a priori* knowledge of the problem. We will see in the simulations that SVM interpolation algorithms are robust to moderate deviations of the free parameters from the optimum ones, which is a very convenient property for a signal interpolation algorithm.

D. Comparison Between Primal and Dual Signal Models

In order to qualitatively compare the sinc kernel SVM primal and dual signal models for nonuniform interpolation, note the following expansion of the solution for the primal signal model approach:

$$\begin{aligned}
x_n &= \sum_{i=1}^{N} a_i \text{sinc}(\sigma_0(t_n - t_i)) \\
&= \sum_{i=1}^{N} \left(\sum_{r=1}^{N} (\alpha_r - \alpha_r^*) \text{sinc}(\sigma_0(t_i - t_r)) \right) \\
&\quad \times \text{sinc}(\sigma_0(t_n - t_i)). \tag{32}
\end{aligned}$$

Comparison between (32) and (29) reveals that these are quite different approaches using SVM for solving a similar signal processing problem. For the primal signal model formulation, and according to (19), limiting the value of C will prevent these coefficients from an uncontrolled growing (regularization effect).

For the dual signal model formulation, the SRM principle implicit in the SVM formalism [23] will lead to a reduced number of nonzero coefficients, thus providing with a desirable sparse solution. Also, in this case, it is easy to see that coefficients are bounded.

III. SIMULATIONS AND RESULTS

In this section, we evaluate the performance of the three proposed SVM-based signal interpolators, and they are compared with four standard interpolation techniques.

A. Interpolation Algorithms for Benchmarking

Many nonoptimal interpolation algorithms have been proposed in the literature. The Jacobian weighting [28] uses the following direct interpolation equation:

$$x(t) = x^J(t) + e(t) = \sum_{i=1}^{N} b_i x_i \text{sinc}(\sigma_0(t - t_i)) + e(t) \tag{33}$$

where coefficients b_i are chosen to be the sample spacings, i.e., $b_i = t_{i+1} - t_i$, which corresponds to a Riemann sum approximation to the following integral identity:

$$x(t) = \int_{-\infty}^{+\infty} x(\tau) \text{sinc}(\sigma_0(t - \tau)) d\tau \tag{34}$$

if we assume that σ_0 is the true bandwidth of $x(t)$. This algorithm has poor performance in interpolation, but an extremely reduced computation burden.

Another suboptimal (but rather improved) approach was proposed in [8], where a generalization of the sinc kernel interpolator is presented. The model relies on a minimax optimality criterion as an approximate design strategy, and it yields the following expression for the coefficients:

$$a_i = \frac{\pi}{\sigma_0} \left(\sum_{n=1}^{N} \text{sinc}^2(\sigma_0(t_i - t_n)) \right)^{-1}. \tag{35}$$

Both the performance and the computational burden of this approach are between Yen's and Jacobian sinc kernel interpolators. However, all these approaches exhibit some limitations, such as poor performance in low signal-to-noise scenarios, or in the presence of non-Gaussian noise (in this last case, as a direct consequence of the use of quadratic loss function). In addition, these methods result in nonsparse solutions. These limitations can be alleviated by accommodating the SVM formulation to the nonuniform sampling problem. In Section III-B, two versatile SVM-based algorithms for nonuniformly sampled, noisy time series interpolation, are introduced.

Therefore, the following signal interpolators are considered:
1) Yen's interpolator without regularization (Y1);
2) Yen's interpolator with regularization (Y2);
3) sinc interpolator with uniform weighting (S1);
4) sinc interpolator with Jacobian weighting (S2);
5) sinc interpolator with minimax weighting (S3);
6) primal signal model SVM with sinc kernel (SVM-P);
7) dual signal model SVM with sinc kernel (SVM-D);
8) dual signal model SVM with RBF kernel (SVM-R).

TABLE I
S/E RATIOS (MEAN $\pm$ STD) FOR GAUSSIAN AND IMPULSE NOISE

Method	Gaussian Noise					Impulse Noise				
	No noise	40dB	30dB	20dB	10dB	15dB	10dB	5dB	0dB	-5dB
Y1	47.5±4.1	38.7±1.7	29.6±1.3	19.9±1.2	9.9±1.1	18.9±1.6	17.4±2.3	14.9±3.2	11.4±4.1	7.5±4.8
Y2	53.4±1.9	39.5±1.4	29.8±1.3	20.1±1.2	10.6±1.2	19.2±1.6	17.8±2.2	15.3±3.1	12.1±3.9	8.5±4.3
S1	-0.5±0.7	-0.5±0.7	-0.5±0.7	-0.6±0.8	-1.3±1.0	-0.6±0.8	-0.7±0.8	-0.8±0.9	-1.2±1.2	-2.1±1.8
S2	15.9±3.0	15.9±3.0	15.7±2.9	14.5±2.2	9.8±1.3	14.5±2.2	13.9±2.0	12.6±2.2	10.5±2.8	7.6±3.6
S3	16.9±2.9	16.9±2.9	16.7±2.8	15.4±2.2	10.2±1.3	15.3±2.0	14.7±2.0	13.3±2.3	11.0±3.0	7.9±3.7
SVM-P	49.1±4.3	39.1±1.3	29.9±1.2	20.5±1.1	12.3±1.6	19.8±1.5	18.4±2.0	16.1±2.9	13.3±3.6	11.0±3.9
SVM-D	50.2±3.2	39.5±1.3	29.9±1.2	20.4±1.1	12.4±1.6	19.7±1.5	18.4±2.0	16.0±2.9	13.0±3.5	10.8±3.9
SVM-R	49.8±1.4	39.6±1.1	30.0±1.2	21.4±1.3	13.5±2.0	20.7±1.8	19.2±2.4	16.7±3.2	13.9±3.7	11.8±3.9

B. Training and Validation Signals

In order to compare all these methods, simulations with known solutions were conducted. Our experimental setup is adapted from [8], where a set of signals with stochastic bandlimited spectra were generated, but here we used a signal with deterministic bandlimited spectra instead. In particular, the recovery of a bandlimited signal with relatively lower energy on its high frequency components of the spectrum was chosen, aiming to explore the effect that regularization could have on the details of the signal. The set of signals consisted of the sum of two squared sinc functions, one of them being a lower level, amplitude modulated version of the baseband component, i.e.,

$$x(t) = \text{sinc}^2(\pi t)\left(1 + \frac{1}{2}\sin(2\pi f t)\right) + e(t) \qquad (36)$$

where $f = 0.4$ Hz and $e(t)$ is additive noise. Note that, in spite of the (apparently) structural simplicity of (36), the function cannot be trivially adjusted by a weighted combination of sinc bases. This function was used in a number of simulations, but only the most representative cases are reported in this section.

A set of L samples were used with averaged sampling interval $T[s]$. The sampling instants were obtained by adding uniform noise, in the range $[-0.1T, 0.1T]$, to equally spaced time points $\{t_k = kT\}_{k=1}^L$. Different values of L were taken, changing accordingly averaged sampling interval T, i.e., when $L = m \times 32$ samples were considered, the averaged sampling interval was changed to $T = 0.5/m$ s, with $m = 1, 2, 3, 4$. Different signal-to-noise ratios (SNRs) were explored (no noise, 40, 30, 20, and 10 dB). Sampling intervals falling outside $[0, LT]$ were wrapped inside. A total of 100 realizations were generated for each set of experiments.

The performance of the interpolators was measured by building a validation set consisting of a noise-free, uniformly sampled version of the output signal with sampling interval $T/16$, as an approximation to the continuous time signal, and then comparing it with the predicted interpolator estimations

at the same time instants. The signal to error (S/E) ratio was computed in decibels as

$$\left(\frac{S}{E}\right)_{dB} = 10\log_{10}\left(\frac{E\left\{(x_n^N)^2\right\}}{E\{e_n^2\}}\right) \qquad (37)$$

in the training set. Means and standard deviations of S/E were averaged over 100 realizations.

C. Tuning the Free Parameters

Four free parameters have to be tuned in SVM algorithms, which are cost function parameters (ε, C, γ), and kernel parameter σ_0 (or equivalently, time duration T_0). These free parameters need to be *a priori* fixed, either by theoretical considerations or by cross-validation search with an additional validation data set.

The tuning of the free parameters is, undoubtedly, the most critical issue for all methods in this problem. However, several criteria are available in the SVM literature to tune the free parameters when little or no knowledge about the problem is available [33]–[37]. Note that, in general, σ_0 and γ can be optimized by using the same methodology as in [34], but such an analysis is beyond the scope of the present paper.

In this paper, and for each developed interpolator, the optimal free parameters were searched according to the reconstruction on the validation set. For SVM interpolators, cost function parameters and kernel parameter were optimally adjusted. For the other algorithms, the best kernel width was obtained, and for Y2, the best regularization parameter was determined by using the available validation set.

D. Gaussian Noise and SNR

The left-hand side of Table I shows the performance of the algorithms in the presence of additive, Gaussian noise, as a function of SNR. The poorest performance is noticeably exhibited by S1, and some improvement is observed with S2 and S3. Y1 yields a good performance only for low noise levels, whereas Y2 shows a good performance for all noise levels, according to its theoretical optimality (from a ML point of view) for Gaussian noise. All the SVM approaches remain close to this optimum

TABLE II
RATE [%] OF SVS (MEAN ± STD) WITH SNR

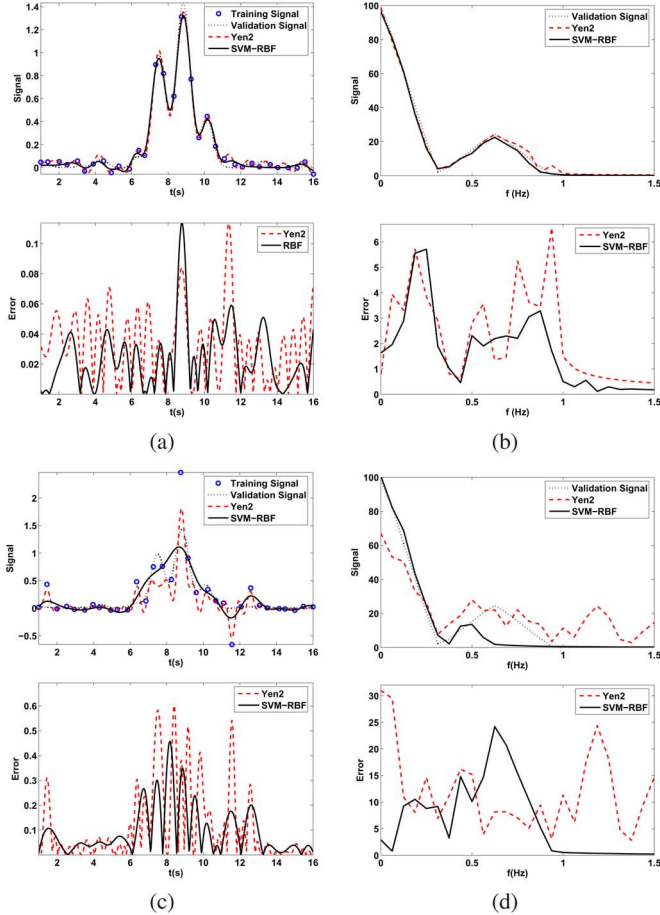| Method	No noise	40dB	30dB	20dB	10dB
SVM-P	96.5±10.8	96.0±14.3	99.3±5.0	92.8±18.8	64.1±18.6
SVM-D	93.9±17.8	95.3±14.4	95.9±14.2	89.1±22.2	59.3±17.3
SVM-R	100.0±0.3	99.9±0.7	99.3±2.5	71.8±16.8	49.5±14.9

Fig. 1. Examples of interpolation in the time (a,c) and frequency (b,d) domains, for Gaussian noise $\text{SNR} = 20$ dB, $L = 32$ samples) (a,b) and BG noise ($\text{SNR} = 20$ dB, $\text{SIR} = 0$ dB, $L = 32$ samples) (c,d).

TABLE III
RATE [%] OF SVS (MEAN ± STD) WITH NUMBER OF SAMPLES

Method	16	32	64	128
SVM-P	97.2±7.8	90.2±21.2	88.6±26.0	76.6±38.8
SVM-D	93.9±12.9	85.1±25.5	81.7±33.0	72.3±40.9
SVM-R	88.7±13.8	69.7±17.5	48.6±12.0	29.6±11.8

least sparse solutions are provided by SVM-R (up to 40% of SV) and SVM-P, respectively. This is an interesting property, which in general is not yielded by conventional Information Theory interpolators or by other previous kernel interpolators.

Fig. 2 shows the sparseness (rate of support vectors SVs[%]) and the S/E obtained as a function of $\varepsilon(\text{SNR} = 20$ dB). With low number of samples ($L = 32$), there is a range of values of ε for which the sparseness can be reduced without significantly modifying the S/E. With increased number of samples ($L = 128$), there is a clear optimum value of ε for SVM-D and SVM-R (dual formulations), for which a notably sparse solution is obtained. Fig. 3 shows two examples (for $L = 32$ and $L = 128$) of the obtained coefficients for the optimum ε. Interestingly, for SVM-P dual coefficients the sparseness is lower, but coefficients a_i, which are obtained by means of dual coefficients, trend to be more sparse than dual coefficients *per se*. The rate of SVs for $L = 128$ samples is dramatically reduced for SVM-R when compared to SVM-D. The high values of standard deviations was due to the presence of bimodality in the distribution of the rate of SVs across experiments, specially present in SVM-P and SVM-D (not shown).

F. Analysis of γ and C

A relevant stage when using SVM algorithms is the selection of the free parameters of the cost function. We studied the effect of changing γ and C for $L = 32$ samples, and the results are shown in Fig. 4. In average, values of C in (10, 1000) yield good performance, but higher values produce high variance in the SE, and γ in $(10^{-5}, 10^{-2})$ are appropriate, while lower values may produce numerical problems due to the lack of regularization.

G. Impulse Noise

In order to test the robustness against impulse noise, similar experiments were conducted. Impulse noise was generated with the Bernoulli-Gaussian (BG) function [38] $n_n^{BG} = v_n \lambda_n$ where v_n is a random process with Gaussian distribution and power σ_{BG}^2, and where λ_n is a random process with probability

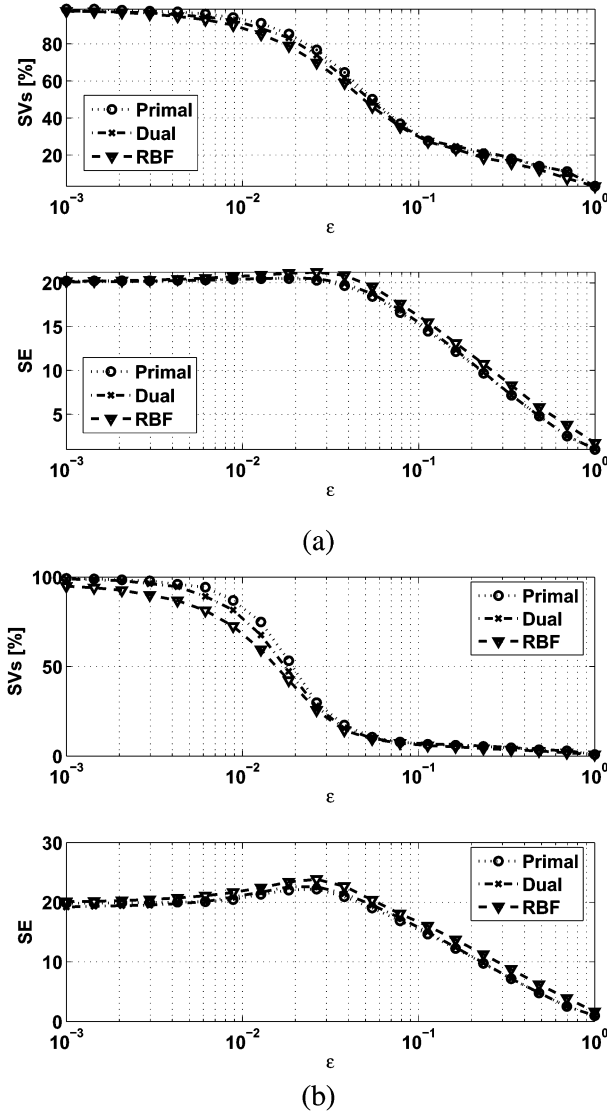$$P_r(\lambda_n) = \begin{cases} p, & \lambda = 1 \\ 1 - p, & \lambda = 0. \end{cases} \tag{38}$$

for high and medium SNR, and even improve around 2 dB for very low SNR.

The top panels in Fig. 1 show a representative example of a modulated sinc signal with $\text{SNR} = 20$ dB, which constitutes a moderate yet more realistic noise level. The interpolation in the time domain is shown to provide a better approximation to the validation signal for Y2 (theoretical optimum) and for the SVM-based methods. Hence, SVM interpolators can yield close-to-optimum performance in the presence of Gaussian noise.

E. Rate of Support Vectors

Sparseness in SVM-based interpolators was also studied as a function of both SNR and signal length, as shown in Table II and Table III. All the SVM methods clearly tend to yield more sparse solutions (in average) with decreasing SNR and with increasing signal lengths. For almost all the situations, the most and the

Fig. 2. Sparseness (upper pannels) and S/E (lower pannels) as a function of ε, for signals with (a) 32 samples, and (b) 128 samples.

Fig. 3. (a) Example of sparseness in SVM coefficients for a signal with $L = 32$ samples. From top to botom: a_i of SVM-P, $\alpha_i - \alpha_i^*$ of SVM-P, SVM-D and SVM-R. (b) The same for an example with $L = 128$ samples.

Accordingly, BG noise with $p = 0.1$ was added to the train signals, with different rates of signal-to-impulse ratio (SIR), defined as

$$\text{SIR}_{dB} = 10\log_{10}\left(\frac{E\left\{\left|x_n - n_n^G - n_n^{BG}\right|^2\right\}}{\sigma_{BG}^2}\right) \qquad (39)$$

where n_n^G is the added Gaussian noise.

The right-hand side of Table I shows the comparison for all the methods in the considered scenario of additive impulse noise. As expected, the SVM algorithms outperform the Y2 algorithm for significantly low SIR values, given that Y2 is no longer the optimum with this noise. Though a fair comparison should take into account M-estimates [39] and its versions for Y2 algorithm, M-estimates can be seen a particular case of SVM with regularized ε-Huber cost function by just taking $\varepsilon = 0$ [21], which was already considered in the search of the free parameters. Tables IV and V show the sparseness of the solution as a function of SIR and of the number of samples.
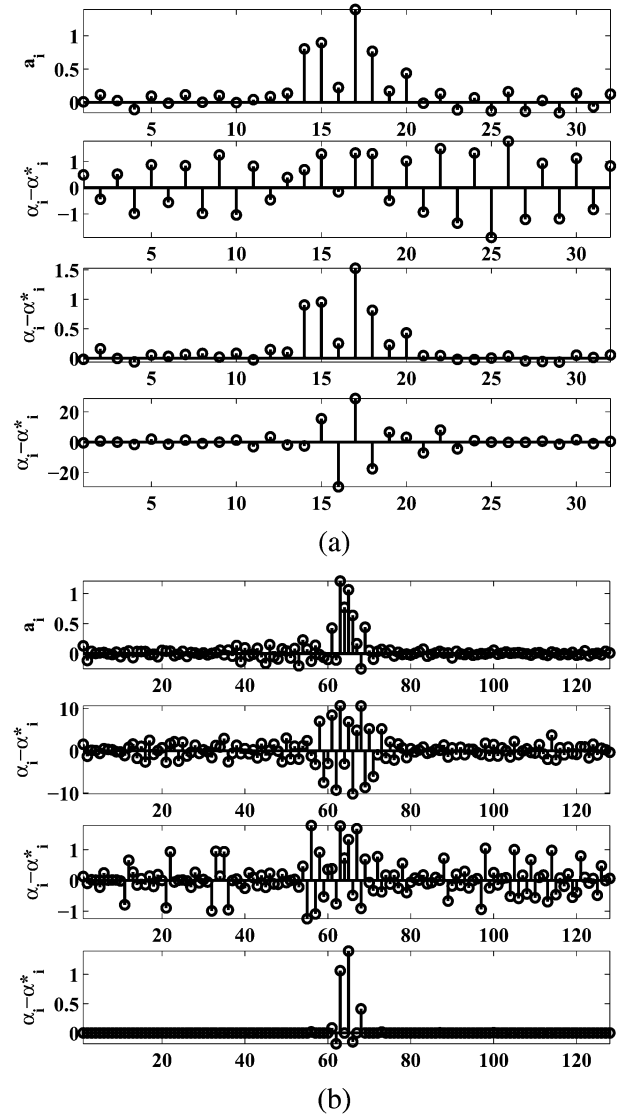
The behavior of SVM algorithms in terms of sparseness in the presence of BG noise was similar than in Gaussian noise, but a slight reduction in sparseness could be observed in terms of a comparable number of samples. Bottom panel of Fig. 1 shows an example of signal reconstruction in BG noise, both in the time and in the frequency domain. Note that SVM algorithms obtain a remarkable improved reconstruction at the high-amplitude spike locations, at the expense of a distorting effect in the band close to the Nyquist frequency of the signal. On the other hand, Y2 algorithm uses a width that often invades the whole spectrum.

IV. CONCLUSION

A new approach to the problem of interpolation of nonuniformly sampled signals has been presented, based on the SVM signal processing framework. Not only the sinc kernel, but also the RBF kernel, popular in the SVM literature, have been shown to be close to Yen's optimal in the presence of Gaussian noise,

TABLE IV
RATE [%] OF SVs (MEAN $\pm$ STD) WITH SIR IN BG NOISE

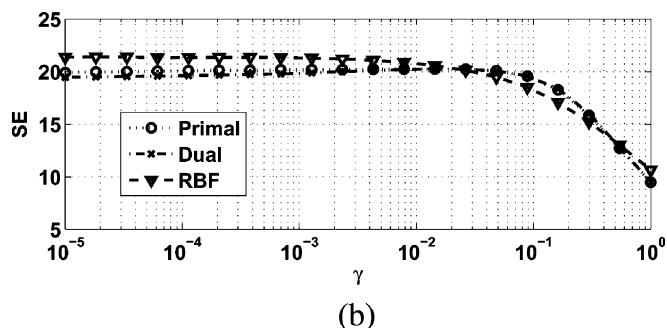| Method	15dB	10dB	5dB	0dB	-5dB
SVM-P	87.2±22.3	79.5±26.0	75.6±25.6	78.9±24.6	85.2±23.7
SVM-D	79.6±27.0	74.5±27.4	67.1±26.9	67.5±27.2	81.3±27.1
SVM-R	67.5±17.1	63.7±17.4	63.2±20.1	71.3±25.3	81.7±26.1

(a)

(b)

Fig. 4. Sparseness (up) and S/E (down), for $L = 32$ samples, as a function of C and γ.

TABLE V
RATE [%] OF SVs (MEAN $\pm$ STD) WITH L IN BG NOISE (SIR $= 0$ dB)

Method	16	32	64	128
SVM-P	92.3±13.1	81.1±24.7	64.1±35.5	44.7±41.1
SVM-D	92.1±14.8	68.7±28.8	55.2±36.8	34.2±37.6
SVM-R	85.7±14.6	63.5±19.7	42.8±16.7	24.0±11.0

Although generated from the same underlying signal model of a weighted sum of sincs basis centered at each sample time, both the approaches coming from Yen's algorithm and the SVM primal and dual signal model formulations have different nature. As a consequence, their particular properties can be exploited in different ways. On the one hand, the SVM algorithm obtained from a primal signal model provides with a robust, yet indirect, estimation of the coefficients, and it is closer in its nature to YenŠs optimal formulation. It can be shown that primal model SVM solution is Y2 for $\varepsilon = 0$ and $C \to +\infty$. For finite values of the parameters and for Gaussian noise, the SVM is biased, while Yen's solution is the optimal unbiased solution through the use of the Maximum Likelihood cost function if the sample size is large enough. If the sample size is small, Yen's solution needs to be regularized, thus becoming biased. On the other hand, dual SVM algorithms, arising from dual signal model formulations, are in fact another form of nonlinear SVM-based regression, and hence, the coefficients are directly obtained as the Lagrange multipliers from the SRM principle.

Some important conclusions can be extracted by comparing the proposed algorithms SVM-D and SVM-R. First, both algorithms produce similar, and noticeably better results than the rest of the standard interpolators in Gaussian or impulse noise environments. Second, a certain tradeoff between sparsity and S/E is observed. In the case of Gaussian noise, the SVM-D outperforms the SVM-R for moderate noise levels (this is, SNR greater than 30 dB), but when higher noise levels are introduced (SNR between 10 and 20 dB), a better interpolated signal is obtained with the SVM-R algorithm. In the case of BG noise, one observes the opposite behavior of the algorithms, thus suggesting that SVM-R is more appropriate for low BG noise levels than the SVM-D algorithm. Third, when the number of training samples was varied, much better results were appreciated for the SVM-R algorithm, specially significant as this number increases. Finally, it is worth noting that there is a clear tradeoff between the sparsity and the S/E curves when comparing primal and dual models. However, and more important, is the fact that

and robust to low number of available samples, outperforming other precedent approximations. Besides, sparse solutions can be obtained. The robustness of SVM algorithms when impulse noise is present has also been explored, with very promising results.

both approaches reveal good robustness capabilities to the selection of the free parameters, and thus good results can be obtained without the need of fine-tuning these parameters.

In conclusion, and given the aforementioned tradeoffs, we can state the algorithm to be used (primal signal model or dual signal model, and RBF or sinc kernel) should be chosen according to the application requirements, such as the amount and nature of noise, or the number of available training samples. Our future research is tied to the study of the sparseness, regularization and robustness capabilities of the proposed methods in real applications of SAR and biomedical signals.

REFERENCES

[1] A. Jerri, "The Shannon sampling theorem-its various extensions and applications: A tutorial review," *Proc IEEE*, vol. 65, pp. 1565–96, 1977.

[2] P. Butzer and R. Stens, "Sampling theory for not necessarily band-limited functions: A historical overview," *SIAM Rev.*, vol. 34, no. 1, pp. 40–53, 1992.

[3] P. Vaidyanathan, "Generalizations of the sampling theorem: Seven decades after Nyquist," *IEEE Trans Circ Syst I*, vol. 48, no. 9, pp. 1094–1109, 2001.

[4] E. Meijering, "A chronology of interpolation: From ancient astronomy to modern signal and image processing," *Proc. IEEE*, vol. 90, pp. 319–342, 2002.

[5] C. Shannon, "Classic paper: Communication in the presence of noise," *Proc. IEEE*, vol. 86, pp. 447–457, 1998.

[6] A. García, "Orthogonal sampling formulas: A unified approach," *SIAM Rev.*, vol. 42, no. 3, pp. 499–512, 2000.

[7] T. Strohmer, "Numerical analysis of the non-uniform sampling problem," *J. Comput. Appl. Math.*, vol. 122, no. 1–2, pp. 297–316, 2000.

[8] H. Choi and D. Munson, "Direct-fourier reconstruction in tomography and synthetic aperture radar," *Int. J. Imag. Syst. Technol.*, vol. 9, no. 1, pp. 1–13, 1998.

[9] J. Jackson, C. Meyer, D. Nishimura, and A. Macovski, "Selection of a convolution function for fourier inversion using gridding," *IEEE Trans. Med. Imag.*, vol. 10, pp. 473–478, 1991.

[10] M. Unser, "Sampling-50 years after Shannon," *Proc. IEEE*, vol. 88, pp. 569–587, 2000.

[11] J. Yen, "On nonuniform sampling of bandwidth-limited signals," *IRE Trans. Circuit Theory*, vol. 3, no. 4, pp. 251–57, 1956.

[12] E. J. Diethron and D. Munson, Jr, "A linear time-varying system framework for noniterative discrete-time band-limited signal extrapolation," *IEEE Trans. Signal Process.*, vol. 39, no. 1, pp. 55–68, 1991.

[13] S. Yeh and H. Stark, "Iterative and one-step reconstruction from nonuniform samples by convex projections," *J. Opt. Soc. Amer.*, vol. 7, no. 3, pp. 491–499, 1990.

[14] C. Cenker, H. Feichtinger, and H. Steier, "Fast iterative and non-iterative reconstruction of band-limited functions from irregular sampling values," in *Proc. ICASSP'91*, Toronto, Canada, 1991, vol. 3, pp. 1773–1776.

[15] F. Marvasti, M. Analoui, and M. Gamshadzahi, "Recovery of signals from nonuniform samples using iterative methods," *IEEE Trans. Signal Process.*, vol. 39, no. 4, pp. 872–878, 1991.

[16] M. Unser and J. Zerubia, "A generalized sampling theory without band-limiting constraints," *IEEE Trans. Circuits Syst. II*, vol. 45, no. 8, pp. 959–969, 1998.

[17] Y. Eldar and A. Oppenheim, "Filterbank reconstruction of bandlimited signals from nonuniform and generalized samples," *IEEE Trans. Signal Process.*, vol. 47, no. 10, pp. 2768–2782, 1999.

[18] G. Kakazu and D. Munson, Jr, "A frequency-domain characterization of interpolation from nonuniformly spaced data," in *IEEE Int. Symp. Circuits Syst.*, Portland, OR, 1989, pp. 288–291.

[19] G. Calvagno and D. C. Munson, Jr, "New results on Yen's approach to interpolation from nonuniformly spaced samples," in *Proc. ICASSP'90*, 1990, pp. 1535–1538.

[20] J. L. Rojo-Álvarez, M. Martínez-Ramón, A. R. Figueiras-Vidal, M. de-Prado Cumplido, and A. Artés-Rodríguez, "Support vector method for ARMA system identification," *IEEE Trans. Signal Process.*, vol. 52, no. 1, pp. 155–164, 2004.

[21] J. L. Rojo-Álvarez, G. Camps-Valls, M. Martínez-Ramón, E. Soria-Olivas, A. N. Vázquez, and A. R. Figueiras-Vidal, "Support vector machines framework for linear signal processing," *Signal Process.*, vol. 85, no. 12, pp. 2316–2326, 2005.

[22] G. Camps-Valls, J. L. Rojo-Álvarez, and M. Martínez-Ramón, *Kernel Methods in Bioengineering, Signal and Image Processing*. Hershey, PA: Idea Group, 2006.

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

[24] A. Tikhonov and V. Arsenen, *Solution to Ill-Posed Problems*. Washington, DC: V.H. Winston, 1977.

[25] G. Huang, Q. Zhu, and C. Siew, "Extreme learning machine: Theory and applications," *Neurocomputing*, 2006, to be published.

[26] K. Yao, "Applications of reproducing kernel Hilbert spaces-bandlimited signal models," *Inf. Control*, vol. 11, pp. 429–444, 1967.

[27] B. Schölkopf and A. Smola, *Learning With Kernels*, 1 ed. Cambridge, MA: MIT Press, 2002.

[28] H. Choi and D. C. Munson, Jr, "Analysis and design of minimax-optimal interpolators," *IEEE Trans Signal Process.*, vol. 46, pp. 1571–1579, 1998.

[29] M. Wand and M. Jones, *Kernel Smoothing*. London, U.K.: Chapman and Hall, 1995.

[30] J. Fan and I. Gijbels, *Local Polynomial Modelling and Its Applications*. London, U.K.: Chapman and Hall, 1995.

[31] J. Platt, *Advances in Kernel Methods: Support Vector Learning*, B. Schölkopf, C. J. C. Burgues, and A. J. Smola, Eds. Cambridge, MA: MIT Press, 1999, ch. Fast training of support vector machines using sequential minimal optimization, pp. 185–208.

[32] L. Zhang, Z. Weida, and L. Jiao, "Wavelet support vector machine," *IEEE Trans. Syst. Man Cybern. B*, vol. 34, pp. 34–39, Feb. 2004.

[33] D. J. C. McKay, "Bayesian interpolation," *Neural Netw.*, vol. 4, pp. 415–447, May 1992.

[34] J. T. Kwok, "The evidence framework applied to support vector machines," *IEEE Trans. Neural Netw.*, vol. 11, pp. 1162–1173, Sep. 2000.

[35] J. T. Kwok and I. W. Tsang, "Linear dependency between ε and the input noise in ε-support vector regression," *IEEE Trans. Neural Netw.*, vol. 14, no. 3, pp. 544–553, 2003.

[36] V. Cherkassky and Y. Ma, "Practical selection of SVM parameters and noise estimation for SVM regression," *Neural Netw.*, vol. 17, no. 1, pp. 113–126, 2004.

[37] A. Chalimourda, B. Schölkopf, and A. Smola, "Experimentally optimal ν in support vector regression for different noise models and parameter settings," *Neural Netw.*, vol. 17, pp. 127–141, 2004.

[38] M. Ghosh, "Analysis of the effect of impulse noise on multicarrier and single carrier QAM systems," *IEEE Trans. Commun.*, vol. 44, no. 2, pp. 145–147, 1996.

[39] P. J. Huber, "Robust statistics: A review," *Ann. Statist.*, vol. 43, pp. 1041–1067, 1972.

José Luis Rojo-Álvarez (M'01) received the telecommunication engineer degree from the University of Vigo, Vigo, Spain, in 1996, and the Ph.D. degree in telecommunication from the Polytechnical University of Madrid, Madrid, Spain, in 2000.

He is an Associate Professor with the Department of Signal Theory and Communications, University Rey Juan Carlos, Madrid. His main research interests include statistical learning theory, digital signal processing, and complex system modeling, with applications both to cardiac signal and image processing, and to digital communications. He has published more than 30 papers and more than 50 international conference communications, on support vector machines (SVMs) and neural networks, robust analysis of time series and images, cardiac arrhythmia mechanisms, and Doppler echocardiographic image for hemodynamic function evaluation.

Carlos Figuera-Pozuelo was born in Madrid, Spain. He received the degree in telecommunication engineering from the Universidad Politécnica of Madrid in 2002.

He has been with the Department of Theory of Signal and Communications, Universidad Carlos III of Madrid, for two years, where he is pursuing the Ph.D. degree. He is currently an assistant professor with the Department of Theory of Signal and Communications, Universidad Rey Juan Carlos, Spain. His research interests include wireless sensor networks, signal processing for wireless communications and theory for wireless *ad hoc* networks.

Carlos Eugenio Martínez-Cruz (S'04) was born in Santa Tecla, El Salvador, in May 1972. He received the M.Sc. degree in electrical engineering from the Universidad de El Salvador in 1996.

He was a Professor with the Universidad de El Salvador until 2004. He is currently pursuing the Ph.D. degree at the Universidad Carlos III de Madrid, Spain. His research interests are in digital signal processing and machine learning.

Gustavo Camps-Valls (M'04) was born in València, Spain, in 1972. He received the Ph.D. degree in physics from the Universitat de València in 2002.

He is currently an Associate Professor with the Department of Electronics Engineering, Universitat de València, where teaches electronics, advanced time-series processing, and digital signal processing. His research interests are neural networks and kernel methods for image processing, remote sensing data analysis, and bioengineering and bioinformatics. He is the author (or coauthor) of 40 journal papers, several books, and more than 50 international conference papers.

Dr. Camps-Valls is a referee of several international journals and has served on the Scientific Committees of several international conferences. For more information: http://www.uv.es/gcamps.

Felipe Alonso-Atienza (S'03) was born in Bilbao, Spain, in May 1979. He received the M.S. degree in telecommunication engineering from the Universidad Carlos III de Madrid, Spain, in 2003.

He is currently pursuing the Ph.D. degree at the Universidad Carlos III de Madrid, where he is doing research on digital signal processing, statistical learning theory, biological system modeling, and feature selection techniques and their application to biomedical problems, with particular attention to cardiac signals.

Manel Martínez-Ramón (SM'04) received the degree from the Universitat Politècnica de Catalunya, Spain, in 1994, and the Ph.D. degree from the Universidad Carlos III de Madrid, Spain, in 1999, both in telecommunications engineering.

He is with the Department of Signal Theory and Communications, Universidad Carlos III de Madrid. His research topics are in applications of the statistical learning to signal processing, with emphasis in communications and brain imaging. He has coauthored 14 papers in international journals and 30 conference papers on these topics. He has written a book on applications of SVMs to antennas and electromagnetics (San Rafael, CA: Morgan and Claypool, 2006) and coauthored several book chapters.