16th IFAC Symposium on System Identification
The International Federation of Automatic Control
Brussels, Belgium. July 11-13, 2012

A Combined SVM/OB-based Wiener Model Identification Method

Juan C. Gómez * **Enrique Baeyens** **

* *Laboratory for System Dynamics and Signal Processing*
FCEIA, Universidad Nacional de Rosario
CIFASIS, CONICET
Riobamba 245 Bis, 2000 Rosario, Argentina
e-mail: jcgomez@fceia.unr.edu.ar

** *Instituto de las Tecnologías Avanzadas de la Producción*
Universidad de Valladolid
Paseo del Cauce 59, 47011 Valladolid, Spain
e-mail: enrbae@eis.uva.es

Abstract: A novel method, combining Support Vector Machines and Least Squares Prediction Error techniques, for the identification of the linear and nonlinear blocks in a Wiener model is presented in this paper. The identification is carried out by minimizing an augmented cost function defined as the sum of the standard structural risk function appearing in Support Vector Regression and the quadratic criterion on the prediction errors associated to Least Squares estimation methods. The properties of the proposed method are illustrated through simulation examples.

Keywords: System Identification; Wiener Models; Support Vector Machines; Orthonormal Bases Expansions; Nonlinear Kernel Expansions

1. INTRODUCTION

Most dynamical systems can be better represented by nonlinear models, which are able to describe the global behavior of the system over the whole operating range, rather than by linear ones that are only able to approximate the system around a given operating point. One of the most frequently studied classes of nonlinear models are the so-called *block-oriented* models, which consist of the interconnection of Linear Time Invariant (LTI) systems and static (memoryless) nonlinearities. Within this class, Wiener models, which consist of the cascade connection of an LTI block followed by a static nonlinearity, have the capability of approximating, with arbitrary accuracy, any fading memory time-invariant system, [Boyd and Chua, 1985]. Wiener models have been successfully used to represent dynamical systems in different application areas such as chemical processes [Gómez and Baeyens, 2004], biological processes, signal processing, and control [Norquay et al., 1998].

Several methods have been proposed in the literature for the identification of Wiener models from input-output data. See for instance [Bai, 1998] and [Gómez and Baeyens, 2004], for techniques based on over-parametrization and least squares estimation, [Westwick and Verhaegen, 1996] and [Gómez and Baeyens, 2005], for subspace methods, and [Hagenblad et al., 2008] and the references therein, for maximum likelihood estimation techniques.

In recent years, Support Vector Machines (SVM) regression techniques [Vapnik, 1998] have been proposed for the identification of nonlinear models represented by expansions in terms of nonlinear mappings of the model inputs. These mappings are expressed in terms of associated kernels, so that they do not need to be explicitly computed. The estimates are obtained as the solution of a (convex) Quadratic Programming (QP) problem. Good overviews on Support Vector Regression techniques can be found in [Schölkopf and Smola, 2002], and the recent tutorial paper in [Suykens, 2009]. Several techniques based on Least Squares SVM have been proposed for the identification of Hammerstein models [Goethals et al., 2005a], [Goethals et al., 2005b], and Wiener-Hammerstein models [Falck et al., 2009].

In this paper, a new Wiener model identification method combining SVM regression and least squares estimation approaches is proposed. The method delivers a kernel representation of the system, but also allows for the separate identification of the linear and the nonlinear blocks in the Wiener model. The method is formulated as the minimization of an augmented cost function, resulting in a convex QP problem for which convergence to the global solution can be guaranteed.

The rest of the paper is organized as follows. In section 2, a parameterized model of the Wiener system is derived. The SVM-based identification method is presented in section 3, while the method based on Orthonormal Bases and least squares estimation, introduced in [Gómez and Baeyens, 2004], is briefly described in section 4. The combined identification method is presented in section 5. Simulation examples illustrating the properties of the proposed method are presented in section 6. Finally, some concluding remarks are included in section 7.

978-3-902823-06-9/12/$20.00 © 2012 IFAC

10.3182/20120711-3-BE-2027.00154

2. PROBLEM FORMULATION

In this paper, discrete time Wiener models consisting of the cascade connection of a Linear Time Invariant (LTI) block followed by a static nonlinear block, are considered. The Wiener model is schematically depicted in Fig. 1, where $u \in \mathbb{R}$ is the scalar input signal, $y \in \mathbb{R}$ is the scalar measured output signal, $\nu \in \mathbb{R}$ is additive noise, $v \in \mathbb{R}$ is the intermediate variable (output of the LTI block), $\tilde{y}$ is the output of the nonlinear block, $\mathcal{N}(\cdot)$ is the nonlinear mapping representing the static nonlinearity, and $G(q^{-1})$ is the transfer function (in the backward-shift operator[1] q^{-1}) of the LTI block.

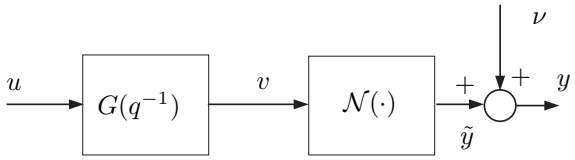

Fig. 1. Wiener Model.

The model can be described as follows:

$$\begin{aligned} v(n) &= G(q^{-1})u(n), \\ \tilde{y}(n) &= \mathcal{N}(v(n)), \\ y(n) &= \tilde{y}(n) + \nu(n). \end{aligned} \tag{1}$$

It is assumed that the LTI block is represented using rational orthonormal bases as follows

$$G(q^{-1}) = \sum_{\ell=1}^{p} b_\ell \mathcal{B}_\ell(q^{-1}), \tag{2}$$

where $b_\ell \in \mathbb{R}$ are unknown parameters, and $\{\mathcal{B}_\ell(q^{-1})\}_{\ell=1}^{\infty}$ are rational orthonormal bases[2] on $H_2(\mathbb{T})$, the space of square integrable functions on the unit circle $\mathbb{T}$, which are analytic outside the unit disk.

In this paper, the rational Orthonormal Bases with Fixed Poles (OBFP) studied in [Ninness and Gustafsson, 1997], that have the more common FIR, Laguerre [Wahlberg, 1991], and Kautz bases as special cases, are considered. The bases are defined as

$$\mathcal{B}_\ell(q) = \left(\frac{\sqrt{1-|\xi_\ell|^2}}{q-\xi_\ell} \right) \prod_{i=1}^{\ell-1} \left(\frac{1-\xi_i q}{q-\xi_i} \right), \qquad \ell \geq 2 \tag{3}$$

$$\mathcal{B}_1(q) = \frac{\sqrt{1-|\xi_1|^2}}{q-\xi_1}, \tag{4}$$

and they allow prior knowledge about an arbitrary number of system modes to be incorporated in the identification process. By choosing the poles of the bases $(\xi_1, \xi_2, \cdots, \xi_p)$, close to the (approximately known) dominant system poles, the accuracy of the estimation can be considerably improved with respect to the case of using FIR, Laguerre or Kautz bases, where the poles need all to be at the same fixed location.

With this parametrization for the LTI block, the Wiener model can be represented as in Fig. 2.

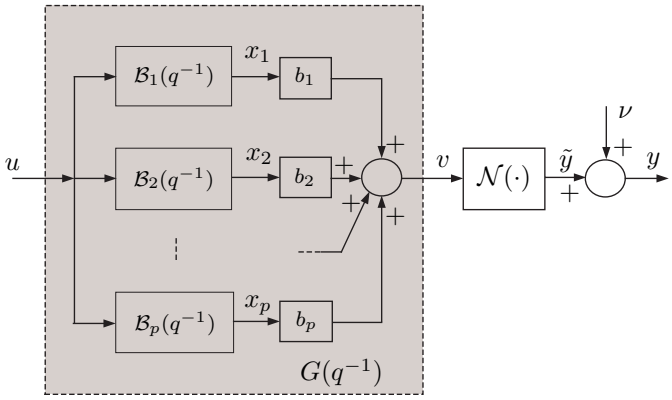

Fig. 2. Parameterized Wiener Model.

The nonlinear block in Fig. 2 can also be parameterized using basis functions in the form

$$\mathcal{N}(v(n)) = \sum_{i=1}^{r} a_i g_i(v(n)), \tag{5}$$

where $a_i, i = 1, 2, \cdots, r$, are unknown parameters, and $g_i(\cdot), i = 1, 2, \cdots, r$ are nonlinear basis functions.

An input-output equivalent representation of the parameterized Wiener model in Fig. 2 is depicted in Fig. 3, where now all the unknowns are concentrated in the nonlinear static Multi-Input Single-Output (MISO) block $\tilde{\mathcal{N}}(\cdot)$. The inputs $x_1, x_2, \cdots, x_p$ of this block are computed by filtering the actual input u with the basis functions $\mathcal{B}_1(q^{-1}), \mathcal{B}_2(q^{-1}), \cdots, \mathcal{B}_p(q^{-1})$ used to represent the LTI block.

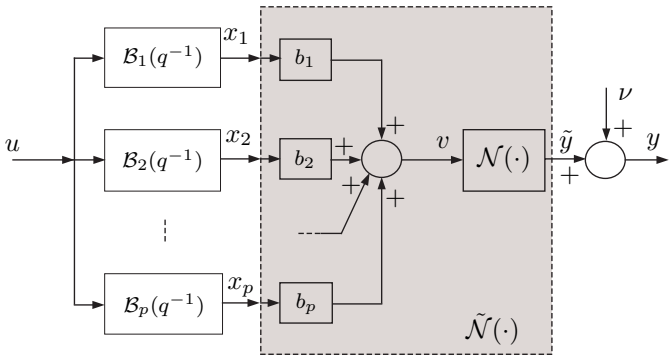

Fig. 3. Input-Output Equivalent Parameterized Wiener Model.

The static nonlinearity can then be described as:

$$\tilde{y}(n) = \tilde{\mathcal{N}}(\mathbf{x}(n)), \tag{6}$$

[1] The backward-shift operator q^{-1} is defined as: $q^{-1}x(n) \triangleq x(n-1)$.

[2] The bases are orthonormal in the sense that

$$\langle \mathcal{B}_\ell, \mathcal{B}_k \rangle = \delta_{\ell k},$$

where $\delta_{\ell k}$ is the Kronecker delta, and $\langle \cdot, \cdot \rangle$ is the standard inner product in $L_2(\mathbb{T})$, defined as

$$\langle \mathcal{B}_\ell, \mathcal{B}_k \rangle \triangleq \frac{1}{2\pi} \int_{-\pi}^{\pi} \mathcal{B}_\ell(e^{j\omega}) \mathcal{B}_k(e^{j\omega}) d\omega.$$

where $\mathbf{x}(n) \triangleq [x_1(n), x_2(n), \cdots, x_p(n)]^T \in \mathbb{R}^p$. Taking into account (5), equation (6) can be written as

$$\tilde{y}(n) = \tilde{\mathcal{N}}(\mathbf{x}(n)), \tag{7}$$
$$= \mathbf{a}^T \mathbf{g}(\mathbf{b}^T \mathbf{x}(n)), \tag{8}$$
$$= \mathbf{a}^T \tilde{\mathbf{g}}(\mathbf{x}(n)), \tag{9}$$

where

$$\mathbf{a} \triangleq [a_1, a_2, \cdots, a_r]^T \in \mathbb{R}^r, \tag{10}$$
$$\mathbf{b} \triangleq [b_1, b_2, \cdots, b_p]^T \in \mathbb{R}^p, \tag{11}$$
$$\mathbf{g}(\cdot) \triangleq [g_1(\cdot), g_2(\cdot), \cdots, g_r(\cdot)]^T : \mathbb{R} \to \mathbb{R}^r, \tag{12}$$
$$\tilde{\mathbf{g}}(\mathbf{x}(n)) \triangleq \mathbf{g}(\mathbf{b}^T \mathbf{x}(n)) : \mathbb{R}^p \to \mathbb{R}^r. \tag{13}$$

Finally, the measured output signal of the parameterized Wiener model in Fig. 3 becomes

$$y(n) = \mathbf{a}^T \tilde{\mathbf{g}}(\mathbf{x}(n)) + \nu(n) \tag{14}$$

In the next section, SVM methods will be used to identify the static nonlinear function $\tilde{\mathcal{N}}(\cdot)$ in terms of the nonlinear basis functions $\tilde{\mathbf{g}}(\mathbf{x}(n))$. Note the reader that the nonlinear functions $\tilde{\mathbf{g}}(\mathbf{x}(n))$ need not be explicitly known, but instead they can be implicitly defined in terms of an associated kernel function (the so-called *kernel trick* in the machine learning literature), [Vapnik, 1998]. A similar problem is considered in [Tötterman and Toivonen, 2009].

3. SVM-BASED IDENTIFICATION

Equation (14) will be the departure point for the formulation of the estimation problem within the framework of the Support Vector Machine methodology. The identification problem, in the so-called *primal weight space*, can be formulated as follows:

Primal Formulation: Given a data set of measured inputs and outputs $\{u(n), y(n)\}_{n=1}^N$, the goal is to estimate a model of the form

$$y(n) = \mathbf{a}^T \tilde{\mathbf{g}}(\mathbf{x}(n)) + c + \nu(n), \tag{15}$$

where c is a *bias* term, and $\{\nu(n)\}$ is an i.i.d. random process with zero mean and finite variance. The unknowns in the model are $\mathbf{a} \in \mathbb{R}^r$, $c \in \mathbb{R}$, and the order r. □

Note the reader that, as pointed out in the previous section, since the filters $\{\mathcal{B}_i(q^{-1})\}_{i=1}^p$ are specified, the internal variable $\mathbf{x}(n)$ associated with the input sequence can be computed straightforwardly by filtering the input with the above mentioned filters.

It is well known that the unknowns $\mathbf{a}$ and c can be determined by solving the following constrained optimization problem [Vapnik, 1998]

$$\begin{aligned} \min_{\mathbf{a},c,\nu} \quad & \frac{1}{2}\mathbf{a}^T\mathbf{a} + \gamma \sum_{n=1}^N L_\epsilon(\nu(n)) \\ \text{subject to} \quad & y(n) - \mathbf{a}^T \tilde{\mathbf{g}}(\mathbf{x}(n)) - c - \nu(n) = 0, \\ & n = 1, \cdots, N \end{aligned} \tag{16}$$

where $\gamma > 0$ is a regularization constant providing a tradeoff between model complexity (penalized by the first term in (16)) and fitting accuracy to the experimental data (penalized by the second term in (16)), and $L_\epsilon(\nu(n))$ is Vapnik's ϵ-insensitivity loss function, defined as

$$L_\epsilon(\nu(n)) = \begin{cases} |\nu(n)| - \epsilon & \text{if } |\nu(n)| \geq \epsilon \\ 0 & \text{otherwise} \end{cases} \tag{17}$$

The optimization problem states that the measured data are contained within an ϵ-tube, by requiring $|y(n) - \mathbf{a}^T \tilde{\mathbf{g}}(\mathbf{x}(n)) - c| \leq \epsilon$, for a given ϵ-accuracy. Usually, to allow for the possibility that some of the data points could be located outside the ϵ-tube, slack variables ξ_n and ξ_n^* are introduced, so that the optimization problem in (16) can be re-stated as follows

$$\begin{aligned} \min_{\mathbf{a},c,\xi_n,\xi_n^*} \quad & \frac{1}{2}\mathbf{a}^T\mathbf{a} + \gamma \sum_{n=1}^N (\xi_n + \xi_n^*) \\ \text{subject to} \quad & y(n) - \mathbf{a}^T \tilde{\mathbf{g}}(\mathbf{x}(n)) - c \leq \epsilon + \xi_n, \\ & -y(n) + \mathbf{a}^T \tilde{\mathbf{g}}(\mathbf{x}(n)) + c \leq \epsilon + \xi_n^*, \\ & \xi_n, \xi_n^* \geq 0, \quad n = 1, \cdots, N \end{aligned} \tag{18}$$

In most cases, the optimization problem (18) can be solved more easily in its dual formulation using Lagrange multipliers, [Schölkopf and Smola, 2002]. The associated Lagrangian is given by

$$\begin{aligned} L(\mathbf{a}, \xi_n, \xi_n^*, \alpha_n, \alpha_n^*) = & \frac{1}{2}\mathbf{a}^T\mathbf{a} + \gamma \sum_{n=1}^N (\xi_n + \xi_n^*) - \sum_{n=1}^N (\eta_n \xi_n + \eta_n^* \xi_n^*) \\ & + \sum_{n=1}^N \alpha_n \left(y(n) - \mathbf{a}^T \tilde{\mathbf{g}}(\mathbf{x}(n)) - c - \epsilon - \xi_n \right) \\ & + \sum_{n=1}^N \alpha_n^* \left(-y(n) + \mathbf{a}^T \tilde{\mathbf{g}}(\mathbf{x}(n)) + c - \epsilon - \xi_n^* \right), \end{aligned} \tag{19}$$

where α_n and α_n^* are the Lagrange multipliers associated with the first two sets of constraints in (18), and η_n and η_n^* are the ones associated with the third set of constraints in (18). The solution of the *primal* optimization problem in (18) is given by the saddle point of the Lagrangian in (19), which should be minimized with respect to $\mathbf{a}$, ξ_n and ξ_n^*, and maximized with respect to the Lagrange multipliers $\alpha_n, \alpha_n^* \geq 0$ and $\eta_n, \eta_n^* \geq 0$.

Introducing the positive definite kernels [Vapnik, 1998]

$$K(\mathbf{x}(n), \mathbf{x}(k)) \triangleq \tilde{\mathbf{g}}^T(\mathbf{x}(n)) \tilde{\mathbf{g}}(\mathbf{x}(k)) \tag{20}$$

associated to the functions $\tilde{\mathbf{g}}(\mathbf{x}(n))$, the *dual* problem in the Lagrange multipliers can be formulated as follows:

$$\begin{aligned} \max_{\alpha_n, \alpha_n^*} \quad & -\frac{1}{2} \sum_{n,k=1}^N (\alpha_n - \alpha_n^*)(\alpha_k - \alpha_k^*) K(\mathbf{x}(n), \mathbf{x}(k)) \\ & -\epsilon \sum_{n=1}^N (\alpha_n + \alpha_n^*) + \sum_{n=1}^N y(n)(\alpha_n - \alpha_n^*) \\ \text{subject to} \quad & \sum_{n=1}^N (\alpha_n - \alpha_n^*) = 0 \\ & \alpha_n, \alpha_n^* \in [0, \gamma], \quad n = 1, \cdots, N \end{aligned}$$

which is a quadratic programming (QP) problem with box constraints, [Suykens, 2009]. The *dual* model representation is then given by

$$y(n) = \sum_{k=1}^{N} (\alpha_k - \alpha_k^*) K(\mathbf{x}(n), \mathbf{x}(k)) + c \qquad (21)$$

Although the number of terms in the representation (21) equals the number of data points N, only a reduced number of terms, corresponding to a small number of vectors $\mathbf{x}(k)$, will have non vanishing coefficients $(\alpha_k - \alpha_k^*)$. These vectors are the so-called *support vectors*. The number of support vectors will depend on the chosen values for ϵ and γ, and on the chosen kernel function. Commonly used kernels are Gaussian Radial Basis Functions (RBF), polynomial kernels, and MultiLayer Perceptrons (MLP) [Schölkopf and Smola, 2002], [Suykens, 2009]. In this paper, Gaussian RBF kernels defined as:

$$K(\mathbf{x}(n), \mathbf{x}(k)) \triangleq \exp(-\|\mathbf{x}(n) - \mathbf{x}(k)\|_2^2 / \sigma^2), \qquad (22)$$

where σ is the (user defined) kernel bandwidth, will be considered for the simulation example in section 6.

4. IDENTIFICATION USING OBFP

The Wiener model identification method based on orthonormal bases introduced in [Gómez and Baeyens, 2004] is briefly described in this section, since it will be combined in the following section with the method described in section 3.

In [Gómez and Baeyens, 2004], the transfer function of the linear block in the Wiener structure of Fig. 1 is parameterized using orthonormal basis functions as in (2). On the other hand, the inverse of the static nonlinearity $\mathcal{N}(\cdot)$ is represented using nonlinear basis functions as follows:

$$\mathcal{N}^{-1}(y(n)) = \sum_{i=1}^{r} d_i f_i(y(n)), \qquad (23)$$

where $d_i, i = 1, 2, \cdots, r$, are unknown parameters, and $f_i(\cdot), i = 1, 2, \cdots, r$ are nonlinear basis functions. Without loss of generality it is assumed that $d_1 = 1$.

By equating the values of the intermediate variable $(v(n))$ computed from the input and from the output in Fig. 1, the following linear regressor equation is obtained

$$f_1(y(n)) = -\sum_{i=2}^{r} d_i f_i(y(n)) + \sum_{\ell=1}^{p} b_\ell \mathcal{B}_\ell(q^{-1}) u(n), \qquad (24)$$

which can be written in matrix form as

$$f_1(y(n)) = \phi^T(n)\theta, \qquad (25)$$

where

$$\theta \triangleq [d_2, d_3, \cdots, d_r, b_1, b_2, \cdots, b_p]^T, \qquad (26)$$

$$\phi(n) \triangleq [-f_2(y(n)), \cdots, -f_r(y(n)), \mathcal{B}_1(q^{-1})u(n), \cdots, \mathcal{B}_p(q^{-1})u(n)]^T. \qquad (27)$$

Given an N-point data set of measured inputs and outputs $\{u(n), y(n)\}_{n=1}^{N}$, an estimate of the parameter vector θ can be computed by minimizing a quadratic criterion on the prediction errors $e(n) = f_1(y(n)) - \phi^T(n)\theta$. This is the well known least squares estimate, which is given by

$$\widehat{\theta} = (\Phi\Phi^T)^{-1}\Phi\mathbf{f}, \qquad (28)$$

provided the indicated inverse exists, and where

$$\Phi \triangleq \left[\phi^T(1); \phi^T(2); \cdots; \phi^T(N)\right]^T, \qquad (29)$$

$$\mathbf{f} \triangleq [f_1(y(1)), f_1(y(2)), \cdots, f_1(y(N))]^T. \qquad (30)$$

The first $(r-1)$ components of vector $\widehat{\theta}$ correspond to estimates of the parameters $d_i, i = 2, \cdots, r$, used to represent the inverse of the static nonlinearity $\mathcal{N}(\cdot)$, while the last p components correspond to an estimate of vector $\mathbf{b}$ used to represent the LTI block in the Wiener model.

5. A COMBINED METHOD

The SVM-based identification method described in section 3 and the OB-based method of section 4 can be combined by defining an augmented cost function as the sum of the one appearing in (18) and the quadratic criterion associated to the least squares estimation in section 4. The optimization problem of the combined method can then be formulated as follows:

$$\min_{\mathbf{a}, c, \xi_n, \xi_n^*, \theta} \left\{ \underbrace{\frac{1}{2}\mathbf{a}^T\mathbf{a} + \gamma \sum_{n=1}^{N} (\xi_n + \xi_n^*)}_{(I)} + \underbrace{\frac{1}{2}\left(\mathbf{f} - \Phi^T\theta\right)^T\left(\mathbf{f} - \Phi^T\theta\right)}_{(II)} \right\} \qquad (31)$$

$$\text{subject to} \quad y(n) - \mathbf{a}^T\tilde{\mathbf{g}}(\mathbf{x}(n)) - c \le \epsilon + \xi_n,$$
$$-y(n) + \mathbf{a}^T\tilde{\mathbf{g}}(\mathbf{x}(n)) + c \le \epsilon + \xi_n^*,$$
$$\xi_n, \xi_n^* \ge 0, \quad n = 1, \cdots, N$$

Note the reader that since term (I) in (31) depends on $\mathbf{a}, c, \xi_n, \xi_n^*$, but does not depend on θ, and term (II) depends on θ but does not depend on $\mathbf{a}, c, \xi_n, \xi_n^*$, then each term can be minimized separately.

In this way, the solution of the optimization problem (31) allows not only to obtain the model kernel representation (21), but also an estimate of vector $\mathbf{b}$ parameterizing the Z-domain transfer function of the linear block in the Wiener model.

The support vectors $\mathbf{x}(k)$ in (21) can then be mapped forward to the corresponding intermediate variable $v(k)$ (which, with some abuse of terminology, will also be called support vectors) in the original parametrization in Fig. 2, using the estimate $\widehat{\mathbf{b}}$ of $\mathbf{b}$, as follows:

$$v(k) = \mathbf{x}^T(k)\widehat{\mathbf{b}}. \qquad (32)$$

Now, with the support vectors $v(k)$ and $y(k)$ associated to $\mathbf{x}(k)$, an explicit model of the nonlinear function $\mathcal{N}(\cdot)$ can be obtained by, for instance, polynomial fitting techniques[3]. This will be illustrated in the following section.

6. SIMULATION EXAMPLE

The properties of the identification methods introduced in Sections 3, 4 and 5 are illustrated in this section through a simulation example.

It is assumed that the real system has a Wiener structure like the one in Fig. 1, where the linear block is represented by a Z-domain transfer function of the form

[3] Any smooth function in an interval can be represented with arbitrary accuracy by a polynomial of sufficiently high order.

$$G(z) = \frac{0.5z^2 + 0.35z - 0.75}{z^3 + 0.9z^2 + 0.15z + 0.002}, \quad (33)$$

and the nonlinear static block is characterized by a saturation-type function given by

$$\mathcal{N}(v) = \tanh(v). \quad (34)$$

For the purposes of identification the system was excited with a Gaussian distributed, zero mean, unit variance random signal, while the output was corrupted with additive zero mean colored noise with (energy density) spectrum

$$\Phi_\nu(\omega) = \frac{0.64 \times 0.0025}{1.04 - 0.4\cos\omega}. \quad (35)$$

The identification algorithm introduced in Section 3 was employed to estimate a Wiener model using the first ($N = 1000$) samples of observed input-output data. The remaining 1000 samples were used for validation purposes.

To evaluate the influence that the location of the poles of the basis functions used to describe the linear block in the Wiener model, have on the prediction accuracy of the estimated models, identification experiments with different pole locations were carried out. The results are summarized in Table 1, where the prediction accuracy of the estimated models is quantified by the Best FIT [4] between the measured and estimated outputs (validation data). The number of support vectors (SV) for each experiment is also shown in the table. For these experiments, the design parameters of the SVM method were set to: $\epsilon = 0.08$, $\gamma = 2000$, $\sigma^2 = 1$. As can be observed from Table 1, as the poles of the bases approach the actual poles of the plant, the estimation accuracy is improved, and the number of support vectors is decreased for the same model order of the linear block. The best performance corresponds to the last row (in bold) of the table, where the poles of the bases exactly match the actual poles.

Table 1. Prediction accuracy and number of support vectors SV, for different bases, pole locations and model orders of the linear block. (SVM-based identification).

Bases	Pole location	p	FIT [%]	SV
Laguerre	{-0.02}	3	29.1022	822
Laguerre	{-0.23}	3	63.5778	609
Laguerre	{-0.7}	4	63.6714	656
Laguerre	{-0.7}	3	63.7273	632
OBFP (app.)	{-0.7, -0.23, -0.02}	3	89.5159	121
OBFP (ex.)	**{-0.6854, -0.2, -0.0146}**	**3**	**90.4881**	**72**

To investigate how does the size of the ϵ-tube affect the estimation accuracy and the resulting model complexity (number of support vectors), identification experiments with different values of ϵ in the range $[0.01, 0.2]$, were carried out. The results are shown in Fig. 4, where the best FIT as a function of ϵ (top plot), and the number of Support Vectors as a function of ϵ (bottom plot) are displayed. As expected, both the estimation accuracy and the model complexity decrease as ϵ increases. For these experiments, the remaining design parameters of the SVM-based identification method were set to: $\gamma = 2000$, $\sigma^2 = 1$, and poles of the bases exactly matching the actual poles.

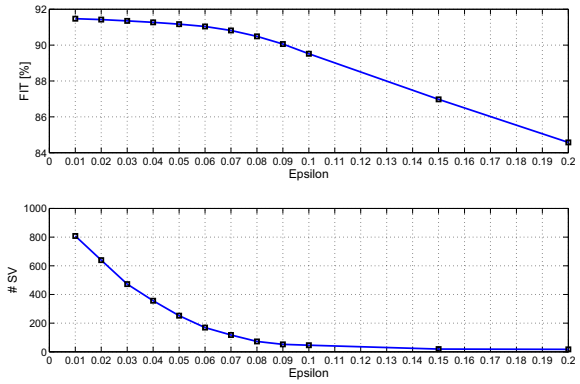

Fig. 4. Top plot: Best FIT vs. ϵ. Bottom plot: Number of SV vs. ϵ.

To investigate how does the regularization parameter γ affect the estimation accuracy and the resulting model complexity, identification experiments with different values of γ in the range $[0.001, 10000]$, were carried out. The results are shown in Fig. 5, where the best FIT as a function of γ (top plot), and the number of Support Vectors as a function of γ (bottom plot) are displayed. In this case, the estimation accuracy increases and the model complexity decreases as γ increases.

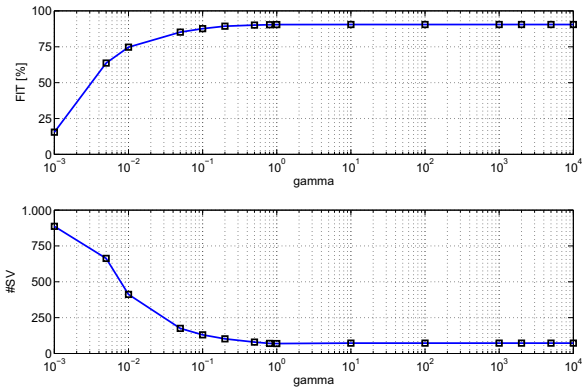

Fig. 5. Top plot: Best FIT vs. γ. Bottom plot: Number of SV vs. γ. ($\epsilon = 0.08, \sigma^2 = 1$, poles of the bases matching actual poles)

The same input-output data were used to estimate Wiener models using the identification algorithm based on orthonormal bases described in section 4. Identification experiments using the same bases, pole locations and model orders as the ones listed in Table 1 were performed, and the prediction accuracy of the estimated models was quantified by means of the Best FIT between the measured and estimated outputs (validation data). The results are summarized in Table 2. The number of estimated parameters (#P) is also shown in Table 2. In this case, the nonlinear basis functions $f_i(y)$ in (23) were chosen as the polynomials: $f_1(y) = y$, $f_2(y) = y^3$ and $f_3(y) = y^5$, $(r = 3)$.

Also here, and as it can be observed from Table 2, as the poles of the bases approach the actual poles of the

[4] The Best Fit is defined as

$$FIT = \left(1 - \frac{\|Y - Y_v\|}{\|Y_v - y_{mean}\|}\right) \times 100,$$

where Y is a vector with the output of the model when excited with the validation input data, Y_v is a vector with the validation output data, and y_{mean} is the mean value of the validation output.

Table 2. Prediction accuracy and number of estimated parameters (#P), for different bases, pole locations and model orders of the linear block. (OB-based identification).

Bases	Pole location	p	FIT [%]	#P
Laguerre	{-0.02}	3	-51.8749	5
Laguerre	{-0.23}	3	50.4458	5
Laguerre	{-0.7}	4	49.7030	6
Laguerre	{-0.7}	3	50.4786	5
OBFP (app.)	{-0.7, -0.23, -0.02}	3	88.0209	5
OBFP (ex.)	**{-0.6854, -0.2, -0.0146}**	**3**	**89.9626**	**5**

plant, the estimation accuracy is improved, for the same model order of the linear block. The best performance corresponds to the last row (in bold) of the table, where the poles of the bases exactly match the actual poles. This performance is comparable to the one obtained with the SVM-based method (last row in Table 1). Comparing the results in Tables 1 and 2, it can be observed that the SVM-based method seems to be more robust to changes in the location of the poles than the OB-based method.

The true (blue solid line) and the OB-based Estimated (black-dashed line) nonlinearities are plotted in Fig. 6, where a good agreement between them can be observed. Also included in Fig. 6 are the estimated nonlinearity using the combined method described in section 5 (green dash-dotted line), and the corresponding support vectors (red crosses). The estimates in Fig. 6 were obtained with the design parameters corresponding to the last row in Tables 1 and 2. The estimated transfer function of the linear block is given by (compare to (33)):

$$\widehat{G}(z) = \frac{0.3944z^2 + 0.2767z - 0.5929}{z^3 + 0.9z^2 + 0.15z + 0.002}. \qquad (36)$$

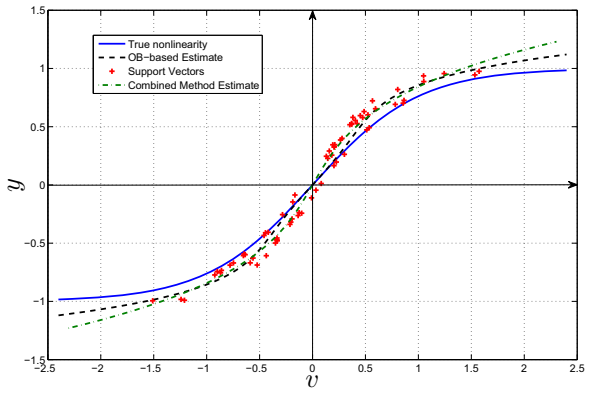

Fig. 6. True nonlinearity (blue solid line), OB-based Estimate (black dashed line), Combined Method Estimate (green dash-dotted line) and the corresponding Support Vectors (red crosses).

7. CONCLUSIONS

A new method for the identification of the linear and nonlinear blocks in a Wiener model has been presented in this paper. The identification is carried out by minimizing an augmented cost function defined as the sum of the standard structural risk function appearing in Support Vector Regression and the quadratic criterion on the prediction errors associated to Least Squares estimation methods. The resulting optimization problem is a convex QP problem for which the convergence to the global solution is guaranteed. The proposed algorithm not only delivers a kernel model of the system but also separate models for the linear and nonlinear blocks in the Wiener structure. The simulation examples show that the proposed method has a performance comparable to that of other state-of-the-art techniques.

REFERENCES

E. Bai. An optimal two-stage identification algorithm for Hammerstein-Wiener nonlinear systems. *Automatica*, 34(3):333–338, 1998.

S. Boyd and L. Chua. Fading memory and the problem of approximating nonlinear operators with Volterra series. *IEEE Trans. on Circuits and Systems*, CAS-32(11): 1150–1161, 1985.

T. Falck, K. Pelckmans, J.A.K. Suykens, and B. De Moor. Identification of Wiener-Hammerstein systems using LS-SVMs. In *Proc. of the 15th IFAC SYSID*, pages 820–825, Saint-Malo, France, July 2009.

I. Goethals, K. Pelckmans, J. Suykens, and B. De Moor. Identification of MIMO Hammerstein models using least squares support vector machines. *Automatica*, 41:1263–1272, 2005a.

I. Goethals, K. Pelckmans, J. Suykens, and B. De Moor. Subspace identification of Hammerstein systems using least squares support vector machines. *IEEE Trans. on Autom. Control*, 50(10):1509–1519, 2005b.

J.C. Gómez and E. Baeyens. Identification of block-oriented nonlinear systems using orthonormal bases. *Journal of Process Control*, 14(6):685–697, 2004.

J.C. Gómez and E. Baeyens. Subspace identification of multivariable Hammerstein and Wiener models. *European Journal of Control*, 11(2):127–136, 2005.

A. Hagenblad, L. Ljung, and A. Wills. Maximum likelihood identification of Wiener models. *Automatica*, 44: 2697–2705, 2008.

B. Ninness and F. Gustafsson. A unifying construction of orthonormal bases for system identification. *IEEE Trans. on Autom. Control*, AC-42(4):515–521, 1997.

S. Norquay, A. Palazoglu, and J. Romagnoli. Model predictive control based on Wiener models. *Chemical Engineering Science*, 53(1):75–84, 1998.

B. Schölkopf and A.J. Smola. *Learning with Kernels*. MIT Press, Cambridge, MA, 2002.

J.A.K. Suykens. Support vector machines and kernel-based learning for dynamical systems modelling. In *Proc. of the 15th IFAC SYSID*, pages 1029–1037, Saint-Malo, France, July 2009.

S. Tötterman and H. Toivonen. Support vector method for identification of Wiener models. *Journal of Process Control*, 19:1174–1181, 2009.

V. Vapnik. *Statistical Learning Theory*. John Wiley & Sons, New York, 1998.

B. Wahlberg. System identification using Laguerre models. *IEEE Trans. on Autom. Control*, AC-36(5):551–562, 1991.

D. Westwick and M. Verhaegen. Identifying MIMO Wiener systems using subspace model identification methods. *Signal Processing*, 52:235–258, 1996.