

# ESTIMATING INPUTS OF NONLINEAR DYNAMICAL SYSTEMS USING DIFFERENTIAL ALGEBRA TECHNIQUES

Gianluigi Pillonetto\* Maria Pia Saccomani\*

*\* Department of Information Engineering  
University of Padova, Italy*

Abstract: The paper considers the problem of estimating the unknown input of a nonlinear dynamical system, described by polynomial or rational differential equations, from a finite set of noisy output samples. Without additional information this problem is ill-posed since it is assumed that the unknown function belongs to an infinite-dimensional space. We tackle this difficulty by designing a novel class of fast regularization algorithms that relies upon differential algebra techniques. Monte Carlo studies are used to demonstrate the goodness of the new approach.  
*Copyright ©2005 IFAC*

Keywords: inverse problems, characteristic set, input-output relation, Tikhonov regularization, regularization parameter, Kalman smoothing filter, Sobolev spaces

## 1. INTRODUCTION

At the beginning of the past century ill-posed problems were considered mathematical anomalies, unable to describe physical systems. This belief was wrong and, in particular, algorithms able to reconstruct functions, starting from noisy and indirect samples, have now many applications. For instance, in medicine they may allow one to estimate inputs of physiological systems that are not accessible to direct measurements, e.g. hormone secretion, from output data, e.g. hormone concentration in plasma. Crucial information on the intimate function of a bio-system may thus be provided (DeNicolao *et al.*, 1997; Pillonetto and Bell, 2004). Other applications regard e.g. imaging, artificial vision, numerical differentiation and seismology, see e.g. (Bertero, 1989; Mendel, 1983; Thompson and Tapia, 1990).

In recent years many methods have been proposed to solve ill-posed problems, see e.g. (Bertero and Boccacci, 1998) for an excellent review. The most powerful and employed technique is Tikhonov regularization. It provides the function estimate as

the solution of a variational problem containing a term accounting for the adherence to experimental data and another one that privileges smooth solutions. A central question for this theory is the determination of the so-called regularization parameter that has to establish the right trade-off between these two contrasting components. When the input-output relationship is linear, a number of robust criteria have been derived to automatically determine this key parameter, e.g. generalized cross validation (GCV) and maximum likelihood (ML) (Wahba, 1990). In nonlinear inverse theory the problem is instead much more involved since it becomes difficult to define rigorous statistical regularization criteria. In practice, sub-optimal extensions of the above-mentioned ones are employed. Of note, studies report that they may work well in many situations, see e.g. (Gu, 1993; Kristensen *et al.*, 2004; Bell and Pillonetto, 2004) but in general turn out difficult to implement. In fact, they typically rely on iterative algorithms that do often not give guarantee of reaching a global minimum. In addition, they may

turn out computationally demanding, e.g. when every objective evaluation needs the integration of complex nonlinear differential equations.

In this paper we address the attention to input reconstruction of a rather general type of nonlinear systems described by polynomial or rational differential equations. Then, we design a novel numerical scheme that circumvents the above-described difficulties related to function estimation by exploiting some concepts of differential algebra (Audoly *et al.*, 2001; Saccomani *et al.*, 2003). So far, these methods have been mainly used to study a priori identifiability of unknown system parameters. Here, instead, we employ them to define a new class of fast regularization algorithms able to recover an accurate input reconstruction without the need of estimating any kind of regularization parameter contained in a nonlinear regularization problem.

The paper is organized as follows. In Section 2, we report the mathematical statement of the problem and an example system involving Michaelis-Menten dynamics. In Section 3 the new algorithm is introduced. We first design a class of Kalman smoothing filters able to estimate a continuous-time profile of the system output and of its derivatives, until the desired order. Then, we show how to reconstruct the unknown input from the smoothed output estimates. In Section 4 the dynamic system described in Section 2 is used to test the new approach via Monte Carlo studies. Conclusions are offered in Section 5. Finally, one can find in Appendix a brief recall of Sobolev spaces and the notation used in this paper.

## 2. STATEMENT OF THE PROBLEM

We are given a discrete set of noisy observations modelled as the  $r$ -dimensional output  $\mathbf{y}(t)$  of a nonlinear dynamical system

$$\begin{cases} \dot{\mathbf{x}}(t) = \mathbf{f}[\mathbf{x}(t)] + \mathbf{g}[\mathbf{x}(t)]u(t) \\ \mathbf{y}(t) = \mathbf{h}[\mathbf{x}(t), u(t)] \end{cases} \quad (1)$$

where  $\mathbf{x}$  is the  $n$ -dimensional state variable, representing e.g. masses, concentrations etc., and  $u$  is the input, e.g. time histories of drug injections. The system is assumed to be scalar-input for simplicity. The general procedure of this paper can be applied to a general MIMO system, but our scope here is mainly to convey the basic idea.

If initial conditions are specified, the relevant equation  $\mathbf{x}(t_0) = \mathbf{x}_0$  is added to the system.

Noisy data are assumed collected at discrete instants  $\{t_k\}$  and denoted

$$\mathbf{z}_k = \mathbf{y}(t_k) + \nu_k \quad (2)$$

with  $k = 1, 2, \dots, d$ , where  $\nu_k$  is a sequence of uncorrelated Gaussian random errors of known

covariance superimposed to the output at time  $t_k$ . This output error term, either a disturbance or a measurement error term, affects additively the measured output.

The essential assumptions on system (1) are that

- the admissible scalar input functions  $t \rightarrow u(t)$  are smooth functions of time,
- there is no feedback, i.e.  $u$  is a free variable, in particular not allowed to depend on  $\mathbf{x}$ ,
- the entries of  $\mathbf{f}$ ,  $\mathbf{g}$  and  $\mathbf{h}$  are *polynomial or rational functions* of their arguments.

Our goal is to reconstruct the input function  $u$  from the noisy measurements of  $\mathbf{y}$ . Since the system designer does often not have sufficient information to restrict  $u$  to a finite-dimensional model this typically results in a nonlinear ill-posed problem.

We shall use differential algebra techniques to describe the input-output relation of the model (1) in terms of a nonlinear static combination of input and output derivatives (Audoly *et al.*, 2001; Saccomani *et al.*, 2003). For the fundamentals of differential algebra, the reader is referred to (Ritt, 1950). Here we only recall that a differential polynomial  $f(\mathbf{x})$  is a polynomial in the variable  $\mathbf{x}$  and a finite number of its derivatives  $\mathbf{x}^{(1)}(t), \dots, \mathbf{x}^{(m-1)}(t)$ . The dynamic system (1) can be looked upon as a set of  $n + r$  differential polynomials which are the generators of a differential ideal  $I$  in a differential ring. The characteristic set of the ideal  $I$  is a finite set of  $n + r$  nonlinear differential equations which describes the same solution set of the original system (Ljung and Glad, 1994). In particular, the first  $r$  differential polynomial equations of the characteristic set have the following form:

$$\hat{L}_j[\mathbf{y}, u] = 0 \quad j = 1, \dots, r \quad (3)$$

These polynomials in the variables  $u, u^{(1)}, \dots, y, y^{(1)}, \dots$  are obtained after elimination of the state variables  $\mathbf{x}$  and hence represent exactly the pairs  $(\mathbf{y}, u)$  which are described by the original system. We will refer to them as the *input-output relation* of the system. We have developed a differential algebra algorithm (Audoly *et al.*, 2001) to find efficiently the characteristic set of the polynomials defining the model.

It will be convenient to rewrite the input-output relations (3) of model (1) in the following form:

$$L_{j0}[\mathbf{y}] = L_j[\mathbf{y}, u] \quad j = 1, \dots, r \quad (4)$$

where, for every  $j$ ,  $L_{j0}[\mathbf{y}]$  is of degree 0 in  $u$  (and assumed non identically null in the sequel only for simplicity) and  $L_j[\mathbf{y}, u]$  is the polynomial term of the input-output differential equation that involves derivatives and powers of the observable variables  $\mathbf{y}$  and  $u$ .

Having brought the  $u$ -dependent part of the relation to the right-hand side, eq.(4) can be seen as a regression problem. In principle, from knowledge of  $\mathbf{y}, \mathbf{y}^{(1)}, \dots, \mathbf{y}^{(m-1)}$  at times  $t_k$ , one could try to obtain  $u, u^{(1)}, \dots, u^{(m-1)}$  at times  $t_k$  by solving the above algebraic relation. However, in practice, we have some problems:

- (1) since the derivative terms in equations (4) are not available from measurements, they must be estimated from the output data records,
- (2) there will be unavoidable approximation errors involved in the process. So, if (4) is to be used for estimating  $u$ , it would be worth introducing a statistical model of these errors,
- (3) it would be important to reconstruct a continuous-time profile of  $u$  and not only estimates of  $u(t_k)$ , i.e. at the discrete times  $t_k$ .

Our plans to solve the above problems is presented in the following. After introducing a system example, we will attack the problems in the given order.

*Example 1.* Consider a two-compartment model which describes the kinetics of a drug in the human body. The drug is injected into the blood where it exchanges linearly with the tissues; the drug is irreversibly removed with a nonlinear saturative characteristic from the blood and with a linear one from the tissue. The system is

$$\begin{cases} \dot{x}_1(t) = -(\theta_1 + \frac{\theta_3}{\theta_4 + x_1(t)})x_1(t) + \theta_2 x_2(t) + u(t) \\ \dot{x}_2(t) = \theta_1 x_1(t) - \theta_2 x_2(t) \\ y(t) = x_1(t) \end{cases} \quad (5)$$

where  $x_1, x_2$  are drug masses in compartment 1 and 2, with initial conditions  $x_1(0) = x_2(0) = 0$ ;  $u$  is the drug input;  $y$  is the measured drug output;  $\theta_1, \theta_2$  are the (constant) rate parameters;  $\theta_3, \theta_4$  are the classical Michaelis-Menten parameters. The system is accessible from every point. Using differential algebra techniques, in particular Ritt's algorithm to eliminate the state variables (Ritt, 1950) one ends up with the following input-output differential equation of the model, as seen from the input-output terminals,

$$\begin{aligned} y^2 \ddot{y} + \theta_1 \theta_2 \dot{y}(y + \theta_4)^2 + \theta_2 \theta_3 y(y + \theta_4) + \\ \theta_4 \ddot{y}(2y + \theta_4) + \theta_3 \theta_4 \dot{y} = (\dot{u} + u \theta_2)(y + \theta_4)^2 \end{aligned} \quad (6)$$

### 3. THE ALGORITHM

#### 3.1 Estimating the output function and its derivatives

For the sake of simplicity, from now we will restrict our attention to single-input single-output dynamical systems, i.e. in eq. (4)  $r$  is set to 1. Thus, in the following  $L_0 = L_{10}$  and  $L = L_1$ ,

even if generalizations are possible and will be discussed in a future paper. In this Section we face the problem of estimating, from  $\{z_k\}$ , the signals  $y(t), y^{(1)}(t), \dots, y^{(m-1)}(t)$  present in eq. (4). To this aim, we use a Tikhonov-type estimator that recovers the estimate of  $y(t)$  as the solution of the following variational problem

$$\arg \min_{y \in W^m} \sum_{k=1}^d \frac{(y(t_k) - z_k)^2}{\sigma_k^2} + \frac{\|P_0[y]\|_{W^m}^2}{\lambda_1^2} \quad (7)$$

where  $\sigma_k^2$  denotes the variance of  $\nu_k$  while  $1/\lambda_1^2$  is the unknown regularization parameter. In order to derive a criterion to determine  $\lambda_1$  as well as an efficient computational scheme implementing it, we interpret Problem (7) in a Bayesian setting. We assume that the curve  $y(t)$  is the realization of a continuous-time stochastic Gaussian process defined by the following stochastic Ito's integral

$$y(t) = \int_0^1 G_m(t, s) \lambda_1 d\beta(s) + \sum_{i=0}^{m-1} \xi_i \psi_i(t) \quad (8)$$

where  $\beta(s)$  is a Brownian motion (independent from  $\{\nu_k\}$ ) and the scalars  $\{\xi_i\}$  may be unknown. It can be derived from (Wahba, 1990) that, given  $\lambda_1$  and  $\{z_k\}$ , Problem (7) provides the minimum variance estimate of  $y$ . Now, let  $\mathbf{e}_i$  the  $m$ -dimensional row vector with  $i$ -th component equal to 1 and all the other ones equal to zero while, by definition, the components of  $\mathbf{e}_0$  are all set to zero. A state-space representation of the Gaussian process  $y(t)$  and of the observations  $\{z_k\}$  is then given by

$$\begin{cases} \dot{\mathbf{x}}_y(t) dt = \mathbf{A}_y \mathbf{x}_y(t) dt + \mathbf{B}_y \lambda_1 d\beta(t) \\ \mathbf{x}_y(0) \sim N(\mathbf{x}_y^0, \mathbf{V}_0) \\ z_k = \mathbf{C}_y \mathbf{x}_y(t_k) + \nu_k \end{cases} \quad (9)$$

where  $\mathbf{x}_y(0) = [\xi_{m-1} \ \xi_{m-2} \ \dots \ \xi_0]^T$ , with  $\mathbf{x}_y^0$  and  $\mathbf{V}_0$  representing the available information on  $\{\xi_i\}$  prior to measurements, and

$$\mathbf{A}_y = \begin{pmatrix} \mathbf{e}_0 \\ \mathbf{e}_1 \\ \dots \\ \mathbf{e}_{m-1} \end{pmatrix} \quad \mathbf{B}_y = (\mathbf{e}_1^T) \quad \mathbf{C}_y = (\mathbf{e}_m) \quad (10)$$

One can thus see that the last component of the state vector  $\mathbf{x}_y(t) \in \mathfrak{R}^m$  corresponds to  $y(t)$  while the other ones represent its derivatives until order  $m - 1$ . The estimation problem thus becomes a continuous-time fixed-interval linear smoothing problem with discrete data and an unknown system parameter  $\lambda_1$ . A maximum likelihood estimate of the regularization parameter can be obtained by maximizing the likelihood function given by

$$l(\mathbf{z}, \lambda_1) = p(z_1 | \lambda_1) \prod_{k=2}^d p(z_k | \mathbf{Z}_{k-1}, \lambda_1) \quad (11)$$

where  $\mathbf{Z}_k = [z_k \ z_{k-1} \ \dots \ z_1]$ . Optimization of the likelihood just requires an univariate search, where the objective can be evaluated at every  $\lambda_1$  value by a forward Kalman filtering. In fact, following (Kitagawa and Gersch, 1985) and denoting with  $\varsigma_k$  and  $r_k$  the innovations and the observation variance conditioned on the observations up to time  $t_{k-1}$ , with  $t_0 = 0$ , we obtain

$$-\log(l(\mathbf{z}, \lambda_1)) = K + \frac{1}{2} \sum_{k=1}^d \log r_k + \frac{1}{2} \sum_{k=1}^d \frac{\varsigma_k^2}{r_k} \quad (12)$$

where  $K$  is a constant we are not concerned with. It thus becomes clear that, in order to obtain an algorithm that scales with the number of the observed data, it is important to determine the sampled version of our state-space model at  $\{t_k\}$ . The next Proposition will provide a computable representation for it. It can be proved after some simple computations that exploit Ito's isometry and the simple structure of matrices  $\mathbf{A}_y$  and  $\mathbf{B}_y$ .

*Proposition 2.* Assume that  $\mathbf{x}_y(t)$  satisfies equation (9). The sequence  $\{\mathbf{x}_y(t_k)\}$  satisfies the following discrete stochastic difference equation

$$\begin{cases} \mathbf{x}_y(t_{k+1}) = e^{\mathbf{A}\Delta_k} \mathbf{x}_y(t_k) + \lambda_1 \omega_k \\ \omega_k \sim \mathbf{N}(\mathbf{0}, \mathbf{Q}^k) \end{cases} \quad (13)$$

where  $\{\omega_k\}$  are mutually independent,  $\Delta_k = t_k - t_{k-1}$ ,  $e^{\mathbf{A}\Delta_k}$  is an  $m \times m$  lower-triangular Toeplitz matrix with  $i$ -th component of its first column given by  $\Delta_k^{i-1}/(i-1)!$  while  $\mathbf{Q}^k$  is an  $m \times m$  matrix whose  $(i, j)$ -entry is given by

$$\mathbf{Q}^k(i, j) = \frac{\Delta_k^{i+j-1}}{(i-1)!(j-1)!(i+j-1)} \quad (14)$$

Finally, once estimated  $\lambda_1$ , a Kalman smoothing filter can be used to obtain estimates of the state vector at any desired temporal instant as well as the covariance matrix of the error affecting them (hereby denoted  $\mathbf{V}_s^k$  when referred to instant  $t_k$ ).

## 3.2 Estimating the unknown input

*3.2.1. Input-output relation linear-in- $u$*  We start considering the case where  $L$  turns out to be a linear operator as function of  $u, u^{(1)}, \dots$ . Interestingly, we have noticed that many dynamic biological models induce this kind of input-output relationship, see (Audoly *et al.*, 2001). A relevant example has been also provided in Section 2. Let  $L^k[y, u]$  and  $L_0^k[y]$  the differential operators  $L$  and  $L_0$  evaluated at  $t_k$ . Let instead

$w(t), w^{(1)}(t), \dots, w^{(m-1)}(t)$  be continuous-time (Gaussian) error terms that are superimposed to the estimate of  $y(t)$ , denoted  $\hat{y}(t)$ , and to its derivatives and that represent the reconstruction error of the Kalman smoothing filter (note that for every  $i$  the statistics of  $w^{(i)}(t_k)$  can be recovered by  $\mathbf{V}_s^k$ ). We define  $\mathbf{b}_k$  the  $m$ -dimensional row vector whose entries are given by

$$\mathbf{b}_k(i) = \frac{\partial L_0^k[\hat{y}, \hat{y}^{(1)}, \dots, \hat{y}^{(m-i)} + w^{(m-i)}, \dots]}{\partial w^{(m-i)}} \Big|_{w^{(m-i)}=0} \quad (15)$$

The following approximated relationship, where the influence of  $w(t), w^{(1)}(t), \dots$  on  $L$  is neglected, is then derived

$$L_0^k[\hat{y}] = L^k[\hat{y}, u] + \epsilon_k \quad (16)$$

where  $\{\epsilon_k\}$  are assumed Gaussian and mutually independent with variance  $\rho_k^2$  defined by

$$\rho_k^2 = \mathbf{b}_k \mathbf{V}_s^k \mathbf{b}_k^T \quad (17)$$

Then, the estimate of  $u$  is defined as the solution of the following variational problem

$$\arg \min_{u \in W^v} \sum_{k=1}^d \frac{(L_0^k[\hat{y}] - L^k[\hat{y}, u])^2}{\rho_k^2} + \frac{\|P_0[u]\|_{W^v}^2}{\lambda_2^2} \quad (18)$$

We note that in order to have a well-defined solution it is important to choose  $v$  so as to make  $L^k : W^v \mapsto \mathfrak{R}$  a bounded functional for every  $k$ . This holds only if  $L^k[R(\cdot, t)]$  turns out a well defined function in  $W^v$  (see Section 11 of (Aronszajn, 1950)). Proposition 3, reported in Appendix, allows the conclusion that such condition is satisfied for values of  $v$  larger than the maximum differential order of the operator  $L$ . As concerns the numerical determination of the solution of Problem (18), we move along the same line developed in the previous Section. Let the scalars  $\{\zeta_i^k\}$  be defined by

$$L^k(\hat{y}, u) = \sum_{i=1}^v \zeta_i^k u^{(v-i)}(t_k) \quad (19)$$

We then obtain the following state-space model for the measurements  $\{L_0^k[\hat{y}]\}$

$$\begin{cases} \dot{\mathbf{x}}_u(t) dt = \mathbf{A}_u \mathbf{x}_u(t) dt + \mathbf{B}_u \lambda_2 d\beta(t) \\ \mathbf{x}_u(0) \sim N(\mathbf{x}_u^0, \mathbf{V}_0) \\ L_0^k[\hat{y}] = \mathbf{C}_u^k \mathbf{x}_u(t_k) + \epsilon_k \end{cases} \quad (20)$$

where, defining the  $v$ -dimensional vector  $\mathbf{e}_i$  as in the previous sub-section, we have

$$\mathbf{A}_u = \begin{pmatrix} \mathbf{e}_0 \\ \mathbf{e}_1 \\ \dots \\ \mathbf{e}_{v-1} \end{pmatrix} \quad \mathbf{B}_u = (\mathbf{e}_1^T) \quad \mathbf{C}_u^k = (\zeta_1^k \dots \zeta_v^k) \quad (21)$$

In (20), the last component of the state vector  $\mathbf{x}_u(t) \in \mathbb{R}^v$  represents  $u(t)$ . Given  $\mathbf{x}_u(0)$ , it is modeled as a Gaussian process with mean determined by  $\mathbf{x}_u(0)$  and deviation from the mean described as the  $(v - 1)$ -fold integration of the Wiener process. As concerns  $\mathbf{x}_u^0$  and  $\mathbf{V}_0$ , such quantities can be defined starting from the prior information on  $u$  at time zero combined with the knowledge of the initial conditions of the system (2) and with the estimates of  $y(0), \dots, y^{(m-1)}(0)$ . Using the stochastic interpretation of Tikhonov regularization, the minimum variance estimate of the  $\nu$ -th component of  $\mathbf{x}_u(\mathbf{t})$ , given  $\{L_0^k[\hat{y}(t)]\}$  and  $\lambda_2$ , corresponds to the solution of Problem (18) (except for a penalty term deriving from the Gaussian prior on  $\mathbf{x}_u(0)$ ). Such estimate can thus be efficiently obtained by exploiting the same numerical procedure developed in the previous Section. We first achieve  $\lambda_2$  by maximum likelihood. Then a Kalman smoothing filter is used to obtain a continuous-time estimate of the input.

*3.2.2. The general case* The second strategy we describe to rapidly recover the unknown input does not necessarily require the linearity of  $L$  in  $u$ . It provides a continuous-time and smooth reconstruction of  $u$  just solving a regularized estimate of the differential equation (4) conditioned on the true output, obtained by replacing  $y(t)$  with the estimate  $\hat{y}(t)$  and assumed to have only one solution in  $u$ . It is worth noticing that, differently from the estimator of eq.(18), this approach does not account for the uncertainty of the estimates of  $y$ . Thus, when  $L$  is linear, if the noise model of eq.(16) is able to well describe such uncertainty, the method reported in the previous sub-section could provide better results.

Finally, it can be proven that the estimator here described is statistically consistent, under some technical conditions on the dynamic system under study. The proof of this result is left to a future extended version of this paper.

#### 4. NUMERICAL EXPERIMENTS

The algorithm is here applied to reconstruct, from a finite set of noisy output samples, two signals modeled as the unknown input of the system (5) with initial conditions set to zero. In both the cases we consider a Monte Carlo study where at every of the 1000 Monte Carlo runs a new realization of the output noise is generated and the system parameters are drawn from a uniform distribution on  $[0, 10]^4$ . The value  $u(0)$  is assumed known.

As concerns the first case study, the unknown input is  $1 - e^{-6t}$ , with  $t \in [0, 1]$ . The function is displayed in the top panel of Figure 1 (solid line)

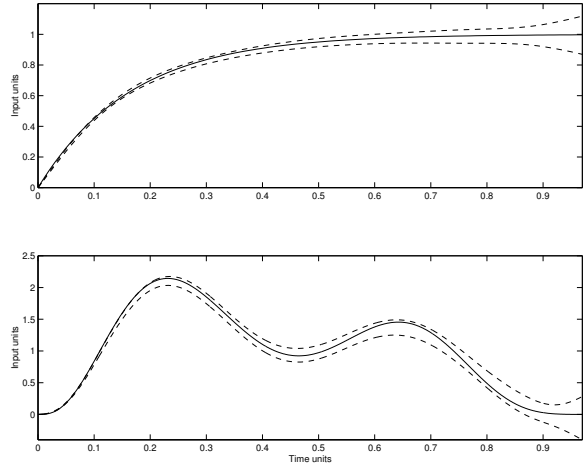


Fig. 1. Monte Carlo simulation *Top and bottom panel* True function (solid line) and mean of the estimates  $\pm$  one SD (dashed lines)

and has to be reconstructed from 40 output samples uniformly collected on  $[0, 1]$  and corrupted by a zero-mean Gaussian noise with a constant percentage of variation (CV%) equal to 5. At every Monte Carlo run we first reconstruct the output function and its derivatives. Then, we use estimator (18) to recover the input estimate. Figure 1 (top panel) plots the mean of the resulting 1000 estimates plus/minus one SD (dashed lines). One can see that the performance of the estimator is very good. The estimate is not biased and the variance of the estimates is low during all the experimental interval. The average root mean square error (AR) turns out 0.043. We also estimated the input without resorting to the estimator of eq. (18), as described in Section (3.2.2). In this case AR turns out 0.052. This result shows that the additional regularization step is able to improve the estimation process but it also demonstrates that the quality of the estimates is similar.

As concerns the second case study, we use a classical test function taken from (Craven and Wahba, 1979) and displayed in the bottom panel of Figure 1 (solid line). The unknown input has to be reconstructed from 100 noisy output samples, uniformly collected on  $[0, 1]$  and corrupted by a zero-mean Gaussian noise with a CV% equal to 10. Figure 1 (bottom panel) plots the mean of the 1000 estimates  $\pm$  one SD (dashed lines). It is apparent that the algorithm still performs well. In this case AR turns out 0.127 while avoiding the additional regularization step provided by estimator (18) it increases to 0.145. Thus, comments similar to those regarding the previous case still hold.

Finally, it is worth mentioning that the entire algorithm has been coded using Matlab 6.5, running under a Pentium IV 2.4 Ghz. On average, solving a single nonlinear inverse problem takes around one second, a result that appears remarkable also considering that the code is not optimized.

## 5. CONCLUSIONS

In this paper it has been shown that the input of a rather general class of nonlinear dynamical systems may be estimated in a stable and efficient way without directly solving any kind of nonlinear regularization problem. Numerical experiments have been used to illustrate the validity of the new approach by using a well-known model involving Michaelis-Menten dynamics. In the near future we plan to further test the method by employing mathematical descriptions of other relevant physical phenomena.

### Appendix. Sobolev spaces

We denote  $W^m(0, 1)$ , or simply  $W^m$  with  $m > 0$  an integer, the Sobolev space of continuous functions  $f$  defined on  $[0, 1]$  such that

- $f$  has continuous derivatives  $f^{(k)}$  of order  $k < m$
- $f^{(m-1)}$  is absolutely continuous and its almost everywhere defined derivative  $f^{(m)}$  lies in the classical Lebesgue space  $L^2(0, 1)$ .

Let  $\|\cdot\|_{L^2}$  be the norm on  $L^2(0, 1)$ . The norm on  $W^m$  is instead denoted by  $\|\cdot\|_{W^m}$  and defined by

$$\|f\|_{W^m}^2 = \sum_{k=0}^{m-1} [f^{(k)}(0)]^2 + \|f^{(m)}\|_{L^2}^2$$

Now, let  $\psi_k(t) = t^k/k!$ , for  $k = 0, 1, 2, \dots, m-1$ , and let  $G_m$  be the Green's function given by

$$G_m(s, t) = \begin{cases} 0 & \text{if } s \leq t \\ 1 & \text{if } s > t \text{ and } m = 1 \\ (s-t)^{m-1}/(m-1)! & \text{otherwise} \end{cases}$$

Then, we remark that  $W^m$  is a reproducing kernel Hilbert space (RKHS) with reproducing kernel given by

$$R(s, t) = \sum_{k=0}^{m-1} \psi_k(s)\psi_k(t) + \int_0^1 G_m(s, u)G_m(t, u)du$$

The above equation also shows that  $W^m$  can be interpreted as the direct sum of two RKHS, see (Wahba, 1990). The first one is spanned by  $\{\psi_k\}$ . The second one admits the same definition of  $W^m$  except that its functions also satisfy the boundary conditions  $f^{(k)} = 0, k = 0, 1, \dots, m-1$ . The projection of  $f$  on such sub-space is denoted  $P_0[f]$ .

The following proposition can be finally proved with arguments analogue to those used in the proof of Theorem 23 in (Bell and Pillonetto, 2004)

*Proposition 3.* For every  $s$ ,  $R(s, t)$ , as function of  $t$ , belongs exactly to  $C^{2m-2}$ , i.e. it is continuous exactly until its  $(2m-2)$ -th derivative

## REFERENCES

- Aronszajn, N. (1950). Theory of reproducing kernels. *Transactions of the American Mathematical Society* **68**, 337–404.
- Audoly, S., L. D'Angiò, M.P. Saccomani and C. Cobelli (2001). Global identifiability of nonlinear models of biological systems. *IEEE Trans. Biomedical Eng.* **48**(1), 55–65.
- Bell, B.M. and G. Pillonetto (2004). Estimating parameters and stochastic functions of one variable using nonlinear measurements models. *Inverse Problems* **20**, 627–646.
- Bertero, M. (1989). Linear inverse and ill-posed problems. *Advances in Electronics and Electron Physics* **75**, 1–120.
- Bertero, M. and P. Boccacci (1998). *Introduction to Inverse Problems in Imaging*. IOP Publishing.
- Craven, P. and G. Wahba (1979). Estimating the correct degree of smoothing by the method of generalized cross validation. *Numerische Mathematik* **31**, 377–403.
- DeNicolao, G., G. Sparacino and C. Cobelli (1997). Nonparametric input estimation in physiological systems: problems, methods and case studies. *Automatica* **33**, 851–870.
- Gu, C. (1993). Smoothing spline density estimation: A dimensionless automatic algorithm. *Journal of the American Statistical Association* **88**, 495–504.
- Kitagawa, G. and W. Gersch (1985). A smoothness priors time-varying ar coefficient modeling of nonstationary covariance time series. *IEEE Transactions on automatic control* **30**, 48–56.
- Kristensen, N.R., H. Madsen and S.B. Jorgensen (2004). Parameter estimation in stochastic grey-box models. *Automatica* **40**, 225–237.
- Ljung, L. and S.T. Glad (1994). On global identifiability for arbitrary model parameterizations. *Automatica* **30**(2), 265–276.
- Mendel, J.M. (1983). *Optimal seismic deconvolution: an estimation-based approach*. Academic Press, New York.
- Pillonetto, G. and B. Bell (2004). Deconvolution of nonstationary physical signals: a smooth variance model for insulin secretion rate. *Inverse Problems* **20**, 367–383.
- Ritt, J.F. (1950). *Differential Algebra*. RI: American Mathematical Society. Providence.
- Saccomani, M.P., S. Audoly and L. D'Angiò (2003). Parameter identifiability of nonlinear systems: the role of initial conditions. *Automatica* **39**, 619–632.
- Thompson, J.R. and R.A. Tapia (1990). *Nonparametric function estimation, modelling and simulation*. Philadelphia, PA. SIAM.
- Wahba, G. (1990). *Spline models for observational data*. SIAM, Philadelphia.