Automatica 49 (2013) 3661–3669

Contents lists available at ScienceDirect

Automatica

journal homepage: www.elsevier.com/locate/automatica

Brief paper

# Mixed parametric/non-parametric identification of systems with discontinuous nonlinearities[☆]

Tyrone L. Vincent[a,1], Carlo Novara[b]

[a] Department of Electrical Engineering and Computer Science, Colorado School of Mines, USA

[b] Dipartimento di Automatica e Informatica, Politecnico di Torino, Italy

ARTICLE INFO

*Article history:*
Received 20 November 2012
Received in revised form
8 July 2013
Accepted 6 September 2013
Available online 4 October 2013

*Keywords:*
System identification
Nonlinear systems
Discontinuities
Regularization

ABSTRACT

The subject of this paper is identification of discrete time nonlinear dynamical systems when the system dynamics are defined by a discontinuous nonlinear function with the location of the discontinuity unknown. By representing the nonlinear function using both a parametric term to capture the continuous part and a non-parametric term to capture the discontinuous part, we present an identification algorithm along with conditions for recovery of the true nonlinearity.



## 1. Introduction

In a large number of system identification schemes for nonlinear systems, the problem reduces to the estimation of a set of nonlinear functions (Hsu, Poolla, & Vincent, 2008; Novara, Vincent, Hsu, Milanese, & Poolla, 2011). A very general model structure, known in the literature as Linear Fractional Transformation (LFT), is of the form

$$y = \mathcal{L}_{yu}u + \mathcal{L}_{ye}e + \mathcal{L}_{yw}w$$
$$z = \mathcal{L}_{zu}u + \mathcal{L}_{ze}e + \mathcal{L}_{zw}w \qquad (1)$$
$$w_k = \mathcal{F}(z_k)$$

where the measured input and output are $u$ and $y$ respectively, $e$ is an unmeasured input, $\mathcal{F}$ is a static nonlinear function to be identified and $\mathcal{L}_{**}$ represents a known linear time invariant (LTI) dynamic system that changes depending on the application. The dimensions of $\mathcal{L}$ and $\mathcal{F}$ are assumed to be compatible with these signals. Nonlinear Auto-regressive with eXogenous input (NARX) and Nonlinear Auto-regressive and Moving Average with eXogenous input (NARMAX) are examples of model structures that can be utilized in such an identification scheme. For example, to represent the NARX model structure

$$y_k = \mathcal{F}(y_{k-1}, y_{k-2}, u_{k-1}, u_{k-2}) + e_k,$$

we may choose

$$\mathcal{L}_{yu} = 0, \quad \mathcal{L}_{ye} = 1, \quad \mathcal{L}_{yw} = 1 \quad \mathcal{L}_{ze} = 0,$$
$$\mathcal{L}_{zu} = \begin{bmatrix} 0 & 0 & q^{-1} & q^{-2} \end{bmatrix}^T,$$
$$\mathcal{L}_{zw} = \begin{bmatrix} q^{-1} & q^{-2} & 0 & 0 \end{bmatrix}^T,$$

where $q^{-1}$ is the unit delay operator. Similarly, NARMAX and other model structures can be represented in this way. Note that by restricting part of $\mathcal{F}$ to be a linear function, block structured systems such as Hammerstein and Wiener systems can also be represented. Other systems may have a physical structure that suggests a particular choice for $\mathcal{L}_{**}$. Examples such as roll to roll physical vapor deposition and an automobile suspension system are discussed in Vincent, Novara, Hsu, and Poolla (2010). Other motivating examples are also discussed in Vincent et al. (2010) and include thin film deposition using multi-zone co-evaporation, thermal network models for building control, and electrical circuits with nonlinear components. Another example that will be used to illustrate the results of this paper is a drill-string, which will be described in more detail below.

In the LFT representation (1), the LTI systems $\mathcal{L}_{**}$ are assumed known, while the static nonlinear function $\mathcal{F}$ has to be identified

☆ Partially supported by NSF Grant CNS-0931748. The material in this paper was partially presented at the 2013 American Control Conference (ACC13), June 17–19, 2013, Washington, DC, USA. This paper was recommended for publication in revised form by Associate Editor Wei Xing Zheng under the direction of Editor Torsten Söderström.

*E-mail addresses:* tvincent@mines.edu (T.L. Vincent), carlo.novara@polito.it (C. Novara).

[1] Tel.: +1 303 273 3641; fax: +1 303 273 3602.

0005-1098/$ – see front matter © 2013 Elsevier Ltd. All rights reserved.
http://dx.doi.org/10.1016/j.automatica.2013.09.022

from the measured input and output signals. The identification of the unknown function has been the subject of many papers, both for specific model structures (see e.g. Sjöberg et al. (1995) and Nowak (2002)) and the generalized structure of (1) (see Hsu et al. (2008) and Novara et al. (2011)). However, the majority of work has focused on the case when either $\mathcal{F}$ is smooth, or is represented by a (small) set of known basis functions, with notable exceptions in the case of Wiener and Hammerstein block structures (Chen, 2006; Vörös, 1997, 2001).

The subject of this paper will be the identification of $\mathcal{F}$ when this function is non-smooth over a part of its domain, but the locations of the discontinuities are unknown *a-priori*. The approach will be a mixed parametric/non-parametric method, in that $\mathcal{F}$ will be represented as

$$\mathcal{F}(z_k) = \sum_{i=1}^{N} \theta_i \phi^{[i]}(z_k) + \eta(z_k), \tag{2}$$

where $\phi^{[i]} : \mathbf{R}^{n_z} \to \mathbf{R}^{n_w}$ are known (smooth) basis functions, $\theta_i$ are parameters, and $\eta : \mathbf{R}^{n_z} \to \mathbf{R}^{n_w}$ is a discontinuous function to be identified. Note that $\eta$ may be characterized by multiple discontinuities. The system identification problem is to select parameters $\theta_k$ and function $\eta$. The intent is for the parametric basis to represent the smooth part of the function $\mathcal{F}$, while $\eta$ will represent the non-smooth part. The identification approach will be to represent $\eta$ using a complete, possibly non-smooth basis. This estimate is called non-parametric because the number of basis elements will be the same as the number of values of $\eta$ to be estimated. Note that, in general, this is an *over-parameterization* of $\mathcal{F}$ (sampled at $z$). Thus, it is necessary to apply additional *a-priori* assumptions about $\eta$. However, smoothness assumptions are not useful, as it is expected that $\eta(z)$ is non-smooth. Instead, it is assumed that the true nonlinear function $\eta$ can be represented with a *small* number of basis elements, although the exact elements are unknown. By applying a regularization term that is the sum of the absolute values of the coefficients $x_k$, i.e. $\|x\|_1$, we will be able to show that the true nonlinear function $\mathcal{F}$ can be recovered with an error that is proportional to the size of the signal $e$.

This approach of signal recovery using an over-parameterized basis is closely related to basis-pursuit for denoising (Chen, Donoho, & Saunders, 2001) and compressive sensing (Candès, Eldar, Needell, & Randall, 2011; Candès, Romberg, & Tao, 2006). There is also a close connection to regression methods that utilize total variation regularization terms (Chan & Tai, 2003; Park, Park, Ahn, & Lee, 2007) in the case that $\eta$ is represented using the gradient basis that is discussed below. However, there are unique aspects of the identification problem that motivate an analysis specific to this case. In particular, the fact that the basis for $\eta$ depends on the signal $z_k$ provides a new twist, and, indeed, recovery depends on the specific realization of $z_k$. The conditions for recovery will provide an implicit persistence of excitation condition that does not occur in standard basis pursuit or compressive sensing problems. Development of the recovery conditions and the study of their implications on the choice of basis for $\eta$ are the primary contributions of this paper.

A preliminary version of this paper was published in Vincent and Novara (2013).

### 1.1. Notation

A variable without subscript (i.e. $z$) will represent a signal, while a variable with subscript (i.e. $z_k$) is the value of the signal at time $k$. Throughout this paper, we will assume an experiment length of $L$. Signal $u$ has dimension $n_u$ and similarly for $e$, $w$, $y$, and $z$. Thus, for example, $u$ represents the signal $u = \{u_k\}_{k=1}^{L}$ where $u_k \in \mathbf{R}^{n_u}$. The $i$th component of a signal at time $k$ will be denoted by $[y_k]_i$, that is, this is the $i$th output at time $k$. When applied to a causal linear operator, such as in the expression $\mathcal{L}_{yw} w$, the result is another signal of time length $L$, which is the truncated output of $\mathcal{L}_{yw}$ with input $w$ with prior zeros appended. In this case, by choosing a basis for $w$ and $y$, $\mathcal{L}_{yw}$ can be represented as an $\mathbf{R}^{n_y L \times n_w L}$ matrix. The notation $w = \mathcal{F}(z)$ indicates that $w$ is the result of applying the nonlinearity pointwise in time to the signal $z$. So, for example, if $\mathcal{F} : \mathbf{R}^{n_z} \to \mathbf{R}^{n_w}$ is a nonlinear mapping, $w$ then is a length $L$ signal with $w_k = \mathcal{F}(z_k)$.

When explicit expressions are necessary, we will use a channel by channel representation for a signal. That is, when e.g. $w$ is written as a vector, it becomes

$$w = \begin{bmatrix} [w_1]_1 & \cdots & [w_L]_1 & [w_1]_2 & \cdots & [w_L]_{n_w} \end{bmatrix}'.$$

Given a matrix $M$ the notation $M_\Lambda$ denotes the matrix made up of the columns of $M$ in the index set $\Lambda$. Similarly, $x_\Lambda$ is the vector made up of the elements of $x$ in the index set $\Lambda$. The vector norms $\|x\|_2 \doteq \sqrt{x'x}$ and $\|x\|_1 \doteq \sum_i |[x]_i|$ are standard. Given a matrix $M$, the operator norm $\|M\|_{p,q}$ is defined to be

$$\|M\|_{p,q} \doteq \sup_{x \neq 0} \frac{\|Mx\|_q}{\|x\|_p}.$$

## 2. Mixed parametric/non-parametric identification

### 2.1. Identification algorithm

Our identification approach will be either via a single optimization or iterative, depending on the following property of the model structure.

**Definition 1.** The signal $z$ is *measurable* if there exists an LTI system $\mathcal{G}_m$ such that

$$\begin{bmatrix} \mathcal{L}_{ze} & \mathcal{L}_{zw} \end{bmatrix} = \mathcal{G}_m \begin{bmatrix} \mathcal{L}_{ye} & \mathcal{L}_{yw} \end{bmatrix}.$$

When $z$ is measurable, $z$ can be determined from knowledge of $y$, $u$ and $\mathcal{L}$ alone, which implies that it is available for the identification algorithm given below. In the case that $z$ is not measurable, an iterative approach can be taken, where an initial guess for $z$ is made, the identification performed, and then using the known $\mathcal{L}$ and estimated $\mathcal{F}$, a new guess for $z$ is obtained. See Novara et al. (2011) for the application of such an iterative approach to a problem of practical interest, related to the identification of semi-active suspension systems.

The proposed mixed parametric/non-parametric identification algorithm can now be stated. We will assume that the system $\mathcal{L}_{ye}$ is invertible, which is standard when this system is used to model the spectrum of a stationary random process. As discussed in the introduction, the nonlinear function $\mathcal{F}$ is represented using expansion (2). It is assumed that $\phi^{[i]}$ contains the constant function. The nonlinearity $\eta(z)$ will be represented using a complete basis that spans the space orthogonal to constant functions, i.e.

$$\eta(z_k) = \sum_{i=1}^{n_w \times (L-1)} x_i \psi^{[i]}(z_k),$$

where $\psi^{[i]}$ sampled at $z_k$ are basis functions satisfying the structural assumptions on $\mathcal{F}$ and $x_i$ are parameters. By structural assumptions, we mean that the appropriate input/output relationships are maintained, so that if a basis element is used to represent a particular output of $\mathcal{F}$, it only depends on the elements of $z_k$ appropriate for that output. Note that this basis function expansion allows us to represent functions with multiple discontinuities. Specific examples of bases suitable for the representation of non-smooth functions will be provided in Section 4.