

# NON-PARAMETRIC IDENTIFICATION OF STATIC NONLINEARITIES IN A GENERAL INTERCONNECTED SYSTEM

Kenneth Hsu<sup>†</sup>, Mareike Claassen<sup>°</sup>, Carlo Novara<sup>◇</sup>  
Pramod Khargonekar<sup>‡</sup>, Mario Milanese<sup>△</sup>, Kameshwar Poolla<sup>‡</sup>

Abstract: We are concerned with the identification of static nonlinear maps in a structured interconnected system. Structural information is often neglected in nonlinear system identification methods. We take advantage of *a priori* structural information and employ a non-parametric method of identification. We focus on the case where the linear part of the interconnection is known and only the static nonlinear components require identification. We propose an identification algorithm and explore its convergence properties.

Keywords: system identification, nonlinear systems, structured systems, convergence, non-parametric

## 1. INTRODUCTION

This paper is concerned with identification problems in interconnected nonlinear systems. These problems are of considerable importance in the context of control, simulation, and design of complex systems.

There is available limited past work on the identification of such systems on a case-by-case basis. These include studies of Hammerstein and Wiener systems (Billings and Fakhouri, 1978), (Narendra and Gallman, 1966), (Pawlak, 1991). However, many of the simplest problems here remain open. For instance, the systematic inclusion of *a priori* structural information has been limited by the lack of a paradigm that is sufficiently general to incorporate such information.

We believe that the development of generalizations such as linear fractional transformations (LFT's) in the control systems literature (Packard and Doyle, 1993), (Safonov, 1982), together with the advent of powerful, inexpensive computational resources offer the promise of significant advances in system identification for complex nonlinear systems.

Much of the available literature treats nonlinear system identification problems in extreme generality, for example using Volterra kernel expansions, neural networks, or radial basis function expansions (Billings and Fakhouri, 1978), (Boutayeb *et al.*, 1993), (Johansen, 1996), (Sjoberg *et al.*, 1995). However, it is our conviction that a completely general theory of nonlinear system identification will have little material impact on the many practical problems that are of interest. We believe that it is more beneficial to study specific classes of nonlinear system identification problems, devise appropriate *systematic* algorithms, and to study the behavior of these algorithms. This experience can then be collated together with experimental studies to develop a broader theory.

We offer a systematic framework based on linear fractional transformations to incorporate known structural information about the interconnected system. While there is occasional work that does incorporate *a priori* structural information (Narendra and Gallman, 1966), (Stoica, 1981), (Vandersteen and Schoukens, 1997), this is not commonly the case.

---

<sup>1</sup> Supported in part by NSF under Grant ECS 03-02554.

<sup>†</sup> Dept. of Mechanical Engineering, University of California, Berkeley, CA 94720. e-mail: ken@jagger.me.berkeley.edu

<sup>°</sup> Math Department, Fullerton College, 321 E. Chapman Ave, Fullerton, CA 92835. e-mail: mclaassen@fullcoll.edu

<sup>◇</sup> Dip. di Automatica e Informatica, Politecnico di Torino, Italy. e-mail: carlo.novara@polito.it

<sup>‡</sup> College of Engineering, University of Florida, Gainesville, FL 32611. e-mail: pramod@ufl.edu

<sup>△</sup> Dip. di Automatica e Informatica, Politecnico di Torino, Italy. e-mail: mario.milanese@polito.it

<sup>‡</sup> Corresponding author. Dept. of Mechanical Engineering, University of California, Berkeley, CA 94720. e-mail: poolla@jagger.me.berkeley.edu

In this paper, we are concerned only with the identification of the static nonlinear components in interconnected systems. We will assume that the linear components of the interconnection are known. In particular, we will be concerned with problems in which the static nonlinear elements to be identified are *non-parametric*.

Non-parametric elements in an interconnected system are often handled by coercing them to assume artificial parameterizations, and then to invoke a parameter estimation procedure. For example, static nonlinearities may be parameterized by radial basis functions, polynomial expansions, or (finite) Fourier series (Greblicki, 1989; Kryzak, 1989).

We submit that this technique of coercing arbitrary “parameterizations” on the inherently non-parametric elements in an interconnected model is at best inefficient, and at worst, can yield substantially biased model estimates. In this paper, we propose a non-parametric identification algorithm that systematically incorporates *a priori* known structural information.

The remainder of this paper is organized as follows. In Section 2, we define the class of model structures under consideration. In Section 3, we motivate and present our identification algorithm. Section 4 introduces the dispersion function. In Section 5 we consider identifiability issues. We briefly discuss convergence aspects of our algorithm in Section 6. In Section 7, we investigate computational aspects of our identification algorithm. In Section 8, we offer an illustrative example. The proofs of our main results can be obtained by contacting the first author.

## NOTATION

$\mathbb{R}^n$	standard Euclidean space
$u, y, w, \dots$	vector-valued discrete-time signals (finite or infinite)
$y_t$	value of signal $y$ at time $t$
$L$	length of data record
$\Pi_L$	$L$ sample truncation operator
$e$	noise signal
LTI	linear time-invariant
$\mathcal{L}$	linear time-invariant operator
$\mathcal{N}$	static nonlinear operator
$z, w$	input and output signals of $\mathcal{N}$
$\mathbb{N}$	$\{\mathcal{N} : \ \mathcal{N}(\zeta_2) - \mathcal{N}(\zeta_1)\  \leq \gamma \ \zeta_2 - \zeta_1\ \}$
$\hat{\mathcal{N}}$	estimate of $\mathcal{N}$
$\mathcal{N}^{true}$	“true” nonlinear map
$\mathcal{D}^{[L]}(z, w)$	dispersion function
$\Omega_i \subseteq \mathbb{R}^{m_i}$	domain over which $\mathcal{N}^{[i]}$ is to be identified
$\Omega \subseteq \mathbb{R}^m$	$\Omega_1 \times \Omega_2 \times \dots \times \Omega_b$

## 2. PROBLEM FORMULATION

We are concerned with the identification of static nonlinear maps in general structured interconnected

systems. An example of the class of “structured” systems we consider is shown in Figure 1. Here, the static nonlinearities  $\mathcal{N}_1$  and  $\mathcal{N}_2$  are to be identified. The (possibly unstable) LTI systems  $\mathcal{L}_1$  and  $\mathcal{L}_2$  are known. We have access to the (noisy) input-output data  $\mathcal{D}^L = \{u_t, y_t\}_0^{L-1}$  and  $e$  is a noise signal.

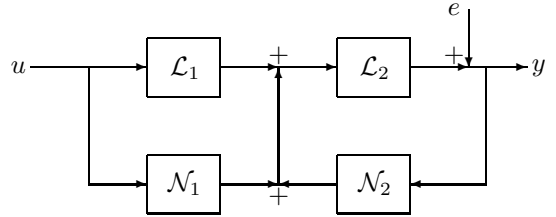


Fig. 1. Example of structured interconnected system

Any general interconnected nonlinear system may be represented through a linear fractional transformation (LFT) framework as shown in Figure 2. The LFT framework allows us to separate the LTI dynamics from the static nonlinearities in an interconnected system. The signals  $u, y$  are measured, and the signals  $z, w$  will denote the inputs and outputs of the static nonlinear block  $\mathcal{N}$ , respectively. The signal  $e$  is a zero-mean white noise process.

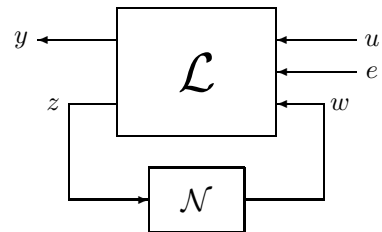


Fig. 2. LFT Model Structure

We gather all the nonlinearities of the interconnection into the multi-input multi-output block  $\mathcal{N}$ , which is to be identified. In general, the static nonlinear block  $\mathcal{N}$  has block diagonal structure (Claassen, 2001).

$$\mathcal{N} = \begin{bmatrix} \mathcal{N}^{[1]} & & \\ & \ddots & \\ & & \mathcal{N}^{[m]} \end{bmatrix}$$

We partition the inputs  $z$  and outputs  $w$  of  $\mathcal{N}$  conformably with its structure. More generally, the nonlinear block  $\mathcal{N}$  may have *repeated* components. This situation arises when a particular nonlinearity appears more than once in the dynamical equations describing the interconnected system. Without loss of generality, all the component nonlinearities in  $\mathcal{N}$  can be assumed to be *single output*. This can be always arranged by introducing redundant copies of the signal  $z$ .

The frequently studied Hammerstein and Wiener systems are special cases of our formulation. However, under our additional assumption that the linear components of the interconnection are known,

the identification of these classes of systems becomes trivial. Indeed, it is important to note that the class of problems we wish to identify involve *complex* interconnections. For example, consider the system depicted in Figure 3.

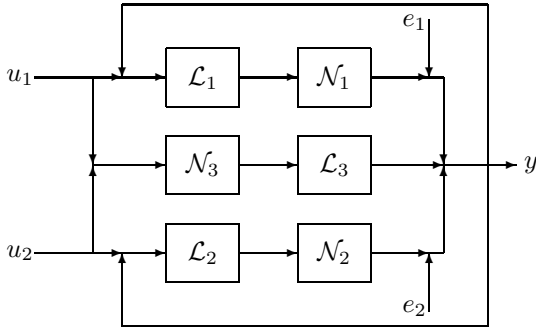


Fig. 3. Complexity in an interconnected system.

Here, the feedback interconnection along with the presence of numerous multivariable linear and nonlinear blocks suggests the complexity of the system. Figure 3 also illustrates a situation where the measured output is the sum of the outputs of several nonlinearities, as well as other signals. In order to develop a systematic approach for the identification of these systems, we use the LFT to collect all such systems under a common framework for analysis.

We will refer to the interconnected system of Figure 2 as the *LFT model structure*. We assume that the LTI block  $\mathcal{L}$  and the dimensions of all signals is known. The components of the non-parametric nonlinear block  $\mathcal{N}$  are to be identified. For this, we have available measured (bounded) input-output data  $\{u_k, y_k\}_{k=0}^{L-1}$ .

Let us partition  $\mathcal{L}$  conformably and realize it as

$$\mathcal{L} = \begin{bmatrix} \mathcal{L}_{yu} & \mathcal{L}_{ye} & \mathcal{L}_{yw} \\ \mathcal{L}_{zu} & \mathcal{L}_{ze} & \mathcal{L}_{zw} \end{bmatrix} \sim \left[ \begin{array}{c|ccc} A & B_u & B_e & B_w \\ \hline C_y & D_{yu} & D_{ye} & D_{yw} \\ C_z & D_{zu} & D_{ze} & D_{zw} \end{array} \right] \quad (1)$$

We summarize our principal assumptions below.

A.1 The realization of  $\mathcal{L}$  in (1) is stabilizable and detectable.

A.2  $D_{zw} = 0$ .

A.3 Measurability of  $z$ , i.e., there exists an LTI system  $\Psi_M$  such that

$$\begin{bmatrix} \mathcal{L}_{ze} & \mathcal{L}_{zw} \end{bmatrix} = \Psi_M \begin{bmatrix} \mathcal{L}_{ye} & \mathcal{L}_{yw} \end{bmatrix} \quad (2)$$

A.4 Co-measurability of  $z$ , i.e., there exists an LTI system  $\Psi_C$  such that

$$\begin{bmatrix} \mathcal{L}_{ze} & \mathcal{L}_{zw} \end{bmatrix} = \mathcal{L}_{zu} \Psi_C \quad (3)$$

A.5 There is no undermodelling. That is, there exists  $\mathcal{N}^{true} \in \mathbb{N}$  that generated the input-output data  $\{u_k, y_k\}_{k=0}^{L-1}$ .

A.6 The class of static nonlinearities we consider is Lipschitz continuous.

$$\mathbb{N} = \{ \mathcal{N} : \|\mathcal{N}(\zeta_2) - \mathcal{N}(\zeta_1)\| \leq \gamma \|\zeta_2 - \zeta_1\| \}$$

We now make several comments regarding these assumptions.

R.1 Note that we do not require  $\mathcal{L}$  to be stable. Indeed,  $\mathcal{L}$  may be stabilized by the feedback nonlinearity  $\mathcal{N}$ .

R.2 This is made for well-posedness of the interconnected system of Figure 2.

R.3 Assumption A.3 is critical to our needs. Observe that

$$\begin{aligned} z &= \mathcal{L}_{zu}u + \mathcal{L}_{ze}e + \mathcal{L}_{zw}w \\ &= \mathcal{L}_{zu}u + \Psi_M \mathcal{L}_{ye}e + \Psi_M \mathcal{L}_{zw}w \\ &= \mathcal{L}_{zu}u + \Psi_M(y - \mathcal{L}_{yu}u) \end{aligned}$$

This is equivalent to requiring that  $z$  be measured, i.e.,  $z$  can be inferred from  $u, y$  and  $\mathcal{L}$ .

R.4 Assumption A.4 is the dual of assumption A.2. We do not require this assumption for our identification procedure. We require this only for our analysis on persistence of excitation.

R.5 Assumption A.5 ensures that there exists  $\mathcal{N}^{true} \in \mathbb{N}$  that generated the input-output data set  $\{u_k, y_k\}_{k=0}^{L-1}$ .

R.6 Assumption A.6 is made to restrict  $\mathcal{N}$  to this specific class of static nonlinearities. While our identification algorithm does not require that  $\mathcal{N}$  be Lipschitz continuous, we require this for our analysis regarding convergence of our estimate.

### 3. THE IDENTIFICATION ALGORITHM

In this section, we describe our proposed identification algorithm for general structured nonlinear systems. In subsequent sections, we analyze this candidate algorithm, address computational issues, and offer an illustrative example.

We draw the reader's attention to Figure 2 that captures the system identification problem we consider. Here,  $\mathcal{L}$  is a *known* LTI discrete-time system, and  $\mathcal{N}$  is a non-parametric static nonlinear map with *known* structure. The problem is to identify the components of the static nonlinear map  $\mathcal{N}$ .

We partition  $\mathcal{L}$  conformably as

$$\mathcal{L} \sim \begin{bmatrix} \mathcal{L}_{yu} & \mathcal{L}_{ye} & \mathcal{L}_{yw} \\ \mathcal{L}_{zu} & \mathcal{L}_{ze} & \mathcal{L}_{zw} \end{bmatrix}$$

Notice that the observed input-output data imposes an affine constraint on the signals  $w$  and  $e$ . We can therefore parameterize the set of all signals  $(w, e)$  consistent with the input-output data  $(u, y)$  as

$$\begin{bmatrix} e \\ w \end{bmatrix} = \begin{bmatrix} e^o \\ w^o \end{bmatrix} + \begin{bmatrix} \mathcal{B}_e \\ \mathcal{B}_w \end{bmatrix} f$$

$$\text{where } \begin{bmatrix} \mathcal{L}_{ye} & \mathcal{L}_{yw} \end{bmatrix} \begin{bmatrix} \mathcal{B}_e \\ \mathcal{B}_w \end{bmatrix} = 0$$

Here  $f$  is a free signal and  $(w^o, e^o)$  is a particular solution.

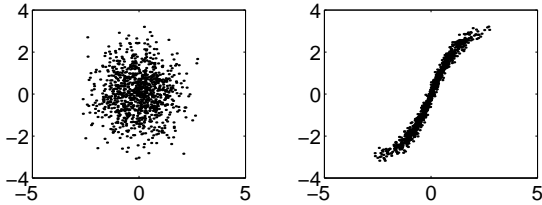


Fig. 4. Illustrating the dispersion function

The problem is then to select the free signal  $f$ . There are two competing criteria that must be considered in this selection. We should require that the graphs from the (vector-valued) signals  $z$  to  $w$  appear to be static nonlinear maps that are consistent with the structure of the block  $\mathcal{N}$ . In addition, we should insist that this choice of  $f$  results in a signal  $e$  which could likely be a sample path produced from a noise process, consistent with any a priori statistical information we may have.

To navigate these requirements we propose to select  $f$  as follows. Imagine we have a suitable function  $\mathcal{D}^{[L]}(z, w)$  that processes the data  $\{z_k, w_k\}_{k=0}^{L-1}$ . This function returns a measure of how well we can interpolate this data by a *smooth, static nonlinearity*. Define the cost function

$$J^{[L]}(f) = \mathcal{D}^{[L]}(z, w) + \eta \|\Pi_L e\|^2$$

Here,  $L$  is the number of samples in the data record being processed. Notice that the signals  $w, z, e$  are affine in  $f$ . We select the free signal  $f$  by solving

$$f^{opt} = \arg \min_f J^{[L]}(f) \quad (4)$$

We then calculate  $w^{opt} = w^o + \mathcal{B}_w f^{opt}$ . Our estimate  $\hat{\mathcal{N}}$  of the static nonlinearity is offered non-parametrically by specifying the points on its graph  $\{w_k^{opt}, z_k^{opt}\}_{k=0}^{L-1}$ . We could interpolate through these points by fitting piecewise linear interpolants or smooth functions using any of a variety of methods. This final step injects bias in our estimates.

## 4. THE DISPERSION FUNCTION

In this section we develop and analyze our data-driven measure of *staticness*  $\mathcal{D}^{[L]}(z, w)$  alluded to in the previous section.

We begin by asking a very simple question. Given input-output data  $\{z_k, w_k\}_{k=0}^{L-1}$ , when can we reasonably assert that this data came from a static nonlinear map  $w = \mathcal{N}(z)$ ? Equivalently, we seek some “measure of staticness” for the map from  $z$  to  $w$ .

To illustrate our thinking, consider the input-output data sets shown in Figure 4. Each data set has 1024 samples. Our native intuition immediately suggests that there is a static nonlinear input-output relationship revealed in the second data set. On the other hand, for the first data set we might reasonably conclude that there is no static nonlinear function relating the input and the output. In this section, we mathematically capture this intuition.

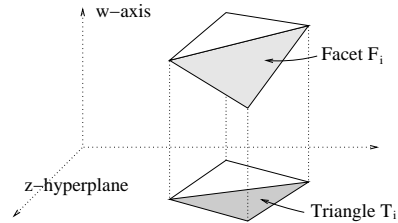


Fig. 5. Triangulation and corresponding facets.

### 4.1 Triangulation

Consider a set of points  $\mathcal{Z} = \{\xi_k\}_{k=0}^{L-1}$ ,  $\xi_k \in \mathbb{R}^2$ . A triangulation partitions the convex hull of  $\mathcal{Z}$  into a disjoint set of triangles whose vertices are points in  $\mathcal{Z}$ . More generally, for a set of points  $\mathcal{Z} = \{\xi_k\}_{k=0}^{L-1}$ ,  $\xi_k \in \mathbb{R}^n$ , a triangulation partitions the convex hull of  $\mathcal{Z}$  into a set of simplices. We will also refer to these simplices as “triangles”.

Note that a triangulation is not *recursive*, i.e., the addition a data point may alter the entire triangulation. While triangulations are not unique, we consider the Delaunay triangulation (Barber *et al.*, 1996) due to its attractive geometric and computational properties. Our subsequent arguments, however, do not rely on this choice of triangulation.

### 4.2 The Faceted Interpolant

Suppose an input-output data set  $\{z_k, w_k\}_{k=0}^{L-1}$  came from a static nonlinear map  $w = \mathcal{N}(z)$ . The Delaunay triangulation partitions the convex hull of the input data points  $\{z_k\}_{k=0}^{L-1}$  into a set of  $N$  triangles  $T = \{T_1, \dots, T_N\}$ . Corresponding to each triangle  $T_i$  is a facet  $F_i$  as shown in Figure 5. These facets define a piecewise linear surface which is the graph of a static nonlinearity  $\hat{\mathcal{N}}$  defined on the convex hull of  $\{z_k\}_{k=0}^{L-1}$ .

We call  $\hat{\mathcal{N}}$  the *faceted interpolant*. The faceted interpolant is an estimate of the nonlinear map  $\mathcal{N}$ .

### 4.3 The Dispersion Function

We propose the use of the dispersion function as a measure of staticness for nonlinear maps. The dispersion function is essentially a measure of how “smooth” the faceted interpolant is.

*Definition 1.* Let  $\{z_k, w_k\}_{k=0}^{L-1}$  be a set of input-output data. Suppose a Delaunay triangulation on  $\{z_k\}_{k=0}^{L-1}$  results in  $N$  facets  $F_i$  with areas  $A_{F_i}$ . The dispersion function is defined to be

$$\mathcal{D}^{[L]}(z, w) = \sum_{i=1}^N A_{F_i}^2 \quad (5)$$

### 4.4 Properties

We first examine the functional dependence of the dispersion function  $\mathcal{D}^{[L]}(z, w)$  on the signal  $w$ . This

is important as our identification algorithm involves minimizing the dispersion function subject to an affine constraint on  $w$ . We have the following theorem.

*Theorem 2.* The dispersion function  $\mathcal{D}^{[L]}(z, w)$  is quadratic in  $w$ , i.e., there exist matrices  $Q, r$  (dependent on  $z$ ) such that

$$\mathcal{D}^{[L]}(z, w) = w^* Q w + r^* w$$

The dispersion function is a measure of the total variation in the interpolated graph of a data set  $\{z_k, w_k\}_{k=0}^{L-1}$ . In the one dimensional case, the dispersion function is the discretization of  $\int (dz^2 + dw^2) = \int ds^2$ , where  $s$  is the arc length of the map  $\mathcal{N}$ . Note that for a continuous function,  $\int ds^2 = 0$ , so if the data set  $\{z_k, w_k\}_{k=0}^{L-1}$  is sufficiently dense, we expect that the dispersion function will converge to zero. This property is captured by the following theorem.

*Theorem 3.* Suppose  $\mathcal{Z}$  is dense on  $\Omega$ . Suppose the input-output data  $\{z_k, w_k\}_{k=0}^{L-1}$  came from a static nonlinearity  $\mathcal{N} \in \mathbb{N}$ . Then,

$$\lim_{L \rightarrow \infty} \mathcal{D}^{[L]}(z, w) = 0$$

We also suspect, but cannot yet prove, that the converse of Theorem 3 holds. Intuitively we suggest that if the dispersion function  $\mathcal{D}^{[L]}(z, w) \rightarrow 0$ , then the facted interpolant  $\hat{\mathcal{N}}^{[L]}$  converges to a Lipschitz continuous static nonlinearity. More precisely, we offer the following conjecture.

*Conjecture 4.* Consider the infinite horizon input-output data  $\{z_k, w_k\}_{k=0}^{L-1}$ . Suppose  $\{z_k\}_{k=0}^{\infty}$  is dense on  $\Omega$ . Suppose

$$\lim_{L \rightarrow \infty} \mathcal{D}^{[L]}(z, w) = 0$$

Let  $\hat{\mathcal{N}}^{[L]}$  be the facted interpolant as described in Section 4.2. Then,

$$\hat{\mathcal{N}}^{[L]} \rightarrow \mathcal{N}^{[\infty]}$$

pointwise on  $\Omega$ . Further,  $\mathcal{N}^{[\infty]} \in \mathbb{N}$ .

This conjecture (or a result analogous to it) will prove central in establishing convergence of our identification algorithm (see Section 6).

## 5. IDENTIFIABILITY

In this section, we treat the issue of identifiability. Loosely speaking, the static nonlinear maps in the LFT model structure of Figure 2 are identifiable if it is possible to determine them uniquely on the basis of input-output experiments. As is well known, identifiability concepts are of fundamental importance in system identification (Ljung and Glad, 1994).

We consider, in this section, the situation where the noise channel  $e$  is absent. Let us denote the

input-output behavior of the LFT model structure as  $y = \Omega(\mathcal{L}, \mathcal{N})u$ .

*Definition 5.* Suppose  $\mathcal{N}^\circ \in \mathbb{N}$ . The LFT model structure is *identifiable at  $\mathcal{N}^\circ$*  if for any  $\mathcal{N}^1 \in \mathbb{N}$  with  $\mathcal{N}^\circ \neq \mathcal{N}^1$  we have

$$\Omega(\mathcal{L}, \mathcal{N}^\circ) \neq \Omega(\mathcal{L}, \mathcal{N}^1)$$

The LFT Model Structure is *identifiable everywhere* if it is identifiable at all  $\mathcal{N}^\circ \in \mathbb{N}$ .

*Remark 6.* We note that these are *global* notions of identifiability.

Partition  $\mathcal{L}$  conformably with its inputs and outputs as

$$\mathcal{L} = \begin{bmatrix} \mathcal{L}_{yu} & \mathcal{L}_{yw} \\ \mathcal{L}_{zu} & \mathcal{L}_{zw} \end{bmatrix}.$$

Define the set of structured matrices

$$\mathbb{X} = \{X = \text{blk-diag}(X^{[1]}, \dots, X^{[m]})\}$$

where  $X^{[i]}$  has the same input-output dimensions as  $\mathcal{N}^{[i]}$ .

We have the following result:

*Theorem 7.* The LFT Model Structure of Figure 2 is not identifiable everywhere if and only if there exists  $0 \neq X \in \mathbb{X}$  such that

$$\mathcal{L}_{yw} X \mathcal{L}_{zu} = 0 \quad (6)$$

## 6. CONVERGENCE

We cannot, at present, prove convergence of our identification algorithm. We will, however, offer a plausibility argument.

Consider the situation where there is no undermodelling (i.e., there is some “true” static nonlinearity  $\mathcal{N}^{true} \in \mathbb{N}$  from which the input-output data are generated). Assume further that there is no noise ( $e = 0$ ). Let  $e^{true}$  be the “true” noise sample path, and let  $w^{true}, z^{true}$ , be the corresponding loop signals.

We can re-parameterize the set of all signals  $w$  consistent with the input-output data as

$$w = w^{true} + \mathcal{B}_w f$$

Then, the cost function  $J^{[L]}(f)$  may be written as

$$J^{[L]}(f) = \mathcal{D}^{[L]}(z^{true}, w^{true} + \mathcal{B}_w f)$$

Next, observe that

$$0 \leq \inf_f J^{[L]}(f) \leq J^{[L]}(0) = \mathcal{D}^{[L]}(z^{true}, w^{true})$$

Note that  $w^{true} = \mathcal{N}^{true}(z^{true})$ . It then follows from Theorem 3 that

$$0 \leq \lim_{L \rightarrow \infty} \inf_f J^{[L]}(f) \leq \lim_{L \rightarrow \infty} \mathcal{D}^{[L]}(z^{true}, w^{true}) = 0$$

As a consequence,

$$J^{[L]}(f^{opt}) = \mathcal{D}^{[L]}(z^{true}, w^{true} + \mathcal{B}_w f) \longrightarrow 0$$

asymptotically as  $L \rightarrow \infty$ .

Assume  $z$  is dense on  $\Omega$ . This will serve as our “persistence of excitation” condition. If Conjecture 4 holds, we conclude that the sequence of faceted interpolants  $\hat{\mathcal{N}}^{[L]} \rightarrow \hat{\mathcal{N}} \in \mathbb{N}$ , and  $\hat{\mathcal{N}}^{[L]}$  is consistent with the infinite horizon input-output data record. Under the further hypothesis of identifiability, there is a unique nonlinearity  $\mathcal{N}^{true} \in \mathbb{N}$  with this property, forcing  $\lim_{L \rightarrow \infty} \hat{\mathcal{N}}^{[L]} = \mathcal{N}^{true}$  pointwise on  $\Omega$ . Of course, this argument rests crucially on Conjecture 4, which remains open.

## 7. COMPUTATION

The computations that arise in our candidate identification algorithm can be efficiently conducted using state-space methods. A critical fact that we wish to stress is that the core optimization problem (4) in our identification procedure reduces to a (large) least squares problem. This is because the objective function  $J^{[L]}(f)$  is jointly quadratic in  $e$  and  $w$ , which are linear in the decision variable  $f$ . The details of the computational algorithm can be obtained by contacting the first author.

## 8. EXAMPLE

Consider the the LFT model structure in Figure 2. Here,  $\mathcal{L}$  is known and the two nonlinear components of  $\mathcal{N}$  are to be identified. Our input  $u$  is a random sequence and  $e$  is a white gaussian process. The length of the data sample is  $L = 1024$ . In Figure 8, the solid lines are the true nonlinearities. The dots are the estimates based on the 1024 samples.

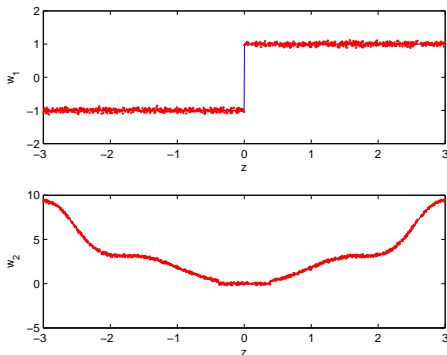


Fig. 6. Estimate of nonlinearities.

## REFERENCES

Barber, C.B., D.P. Dobkin and H.T. Huhdanpaa (1996). The quickhull algorithm for convex hulls. *ACM Transactions on Mathematical Software*.

- Billings, S.A. and S.Y. Fakhouri (1978). Identification of a class of nonlinear systems using correlation analysis. *Proc. of the IEEE* **125**, 691–697.
- Boutayeb, M., M. Darouach, H. Rafaralahy and G. Krzakala (1993). A new technique for identification of miso hammerstein model. *Proc. of the ACC* **2**, 1991–1992.
- Claassen, M. (2001). System identification for structured nonlinear systems. *Ph.D Dissertation, University of California at Berkeley*.
- Greblicki, W. (1989). Nonparametric orthogonal series identification of hammerstein systems. *International Journal of Systems Science* **20**, 2355–2367.
- Johansen, T.A. (1996). Identification of nonlinear systems using empirical data and prior knowledge—an optimization approach. *Automatica* **32**, 337–356.
- Kryzak, A. (1989). Identification of discrete hammerstein systems by the fourier series regression estimate. *International Journal of Systems Science* **20**, 1729–1744.
- Ljung, L. and T. Glad (1994). On global identifiability for arbitrary model parameterizations. *Automatica* **30**, 265–276.
- Narendra, K.S. and P.G. Gallman (1966). An iterative method for the identification of nonlinear systems using the hammerstein model. *IEEE TAC* **11**, 546–550.
- Packard, A. and J.C. Doyle (1993). The complex structured singular value. *Automatica* **29**, 71–110.
- Pawlak, M. (1991). On the series expansion approach to the identification of hammerstein systems. *IEEE TAC* **36**, 763–767.
- Safonov, M.B. (1982). Stability margins of diagonally perturbed multivariable feedback systems. *IEEE Proc.* **129**, 251–256.
- Sjoberg, J., Q. Zhang, L. Ljung, A. Benveniste, B. Delyon, P.Y. Glorennec, h. Hjalmarsson and A. Juditsdy (1995). Nonlinear black-box modeling in system identification: A unified overview. *Automatica* **31**, 1691–1724.
- Stoica, P. (1981). On the convergence of an iterative algorithm used for hammerstein system identification. *IEEE TAC* **26**, 967–969.
- Vandersteen, G. and J. Schoukens (1997). Measurement and identification of nonlinear systems consisting of linear dynamic blocks and one static nonlinearity. *IEEE Instrumentation and Measurement Technology Conference* **2**, 853–858.