Proceeding of the 2004 American Control Conference
Boston, Massachusetts June 30 - July 2, 2004

Basis-Function Optimization for Subspace-Based Nonlinear Identification of Systems with Measured-Input Nonlinearities

Harish J. Palanthandalam-Madapusi, Jesse B. Hoagg, and Dennis S. Bernstein

Abstract— **For nonlinear systems with measured-input nonlinearities, a subspace identification algorithm is used to identify the linear dynamics with the nonlinear mappings represented as a linear combination of basis functions. A selective-refinement technique and a quasi-Newton optimization algorithm are used to iteratively improve the representation of the system nonlinearity. For both methods, polynomials, splines, sigmoids, wavelets, sines and cosines, or radial basis functions can be used as basis functions. Both approaches can be used to identify nonlinear maps with multiple arguments and with multiple outputs.**

1. INTRODUCTION

Nonlinear identification is an increasingly active research area. Among the various approaches that have been developed are nonparametric methods, which are primarily frequency-domain based. These methods include techniques for identifying Volterra kernels, which characterize input-output response by means of a sum of multi-frequency convolutions. For time-domain simulation, however, it is convenient to construct nonlinear state-space realizations of these maps, which may be difficult in practice [11]. On the other hand, parametric methods have been developed based on structured and unstructured time-domain models. Unstructured or black-box models [7, 12] often rely on neural network models to exploit their function approximation properties. Structured or grey-box models [3–5, 10] are based on the interconnection of linear and nonlinear subsystems.

The most common model structures are the Hammerstein, Wiener, nonlinear feedback, and combined Hammerstein/nonlinear feedback models shown in Figure 1, Figure 2, Figure 3, and Figure 4. These models involve the interconnection of a single linear block and a single nonlinear block. Identification with these model structures has been widely considered, see, for example, [1, 2, 6, 14]. A point that has not been stressed is the fact that nonlinear identification with the Wiener model structure is significantly more difficult than identification with the Hammerstein structure. The reason for this difficulty is the fact that identification of a nonlinear map is more tractable when a measurement of the input to the map is available.

The Hammerstein nonlinear/feedback model can be viewed as a realization of a nonlinear system. Accordingly, the representation of the nonlinearities is not unique. It has been shown in [14] that it is not possible to uniquely identify the nonlinear maps. The identified maps could differ by a scaling coefficient or a bias from the real nonlinearity.

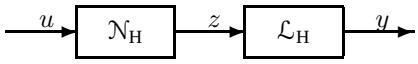

Fig. 1. Hammerstein Model

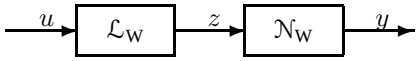

Fig. 2. Wiener Model

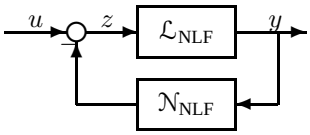

Fig. 3. Nonlinear Feedback Model

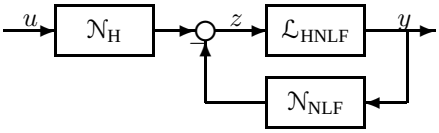

Fig. 4. Hammerstein Nonlinear Feedback Model

In view of the fact that nonlinear identification is facilitated by the availability of inputs, a general formulation of nonlinear identification in this setting was considered in [8] using a subspace identification algorithm [9, 13] along with a given basis expansion for the nonlinear maps. The function expansion was chosen to be linear in the parameters, which allows the nonlinear identification problem to be recast as a linear identification problem with generalized inputs. The multivariable capability of subspace identification algorithms is essential to this approach by allowing an arbitrary number of generalized inputs.

The approach of [8] requires a fixed set of basis functions to represent the nonlinear mapping. For nonlinear mappings of several inputs, the curse of dimensionality requires an excessively large number of basis functions. In addition, the subspace identification algorithm identifies only the coefficients of the basis functions, not the basis functions themselves. Hence, without prior knowledge of the form of the nonlinear mappings, it may be necessary to employ a large number of basis functions, rendering the problem numerically intractable.

This research was supported in part by the Air Force Office of Scientific Research under grant F49620—01–1–0094 and the National Science Foundation Information Technology Research initiative, through Grant ATM-0325332.

The authors are with the Department of Aerospace Engineering, University of Michigan, Ann Arbor, MI 48109-2140, USA

In view of these issues, the present paper develops two techniques for iteratively refining the basis function representation of nonlinear mappings that are functions of measured inputs. As in [8], both techniques use a subspace identification algorithm to identify the linear dynamics for a set of basis functions representing the nonlinear functions.

The first technique uses selective refinement to improve the representation of the nonlinear functions. By applying a singular value decomposition to the input matrix, the dominant nonlinearities are identified for the chosen set of basis functions. Next, a random collection of basis functions is introduced to improve the representation of the dominant nonlinearities. Iteration of these steps constitutes the selective refinement process.

The second algorithm optimizes a fixed set of basis functions by means of a BFGS quasi-Newton optimization code. The representation of the nonlinear map is systematically improved by modifying the basis functions rather than by adding additional basis functions. A subspace identification algorithm is used to identify the linear dynamics for a chosen set of basis functions representing the nonlinear functions. For that particular set of state space matrices, the basis functions are then optimized using a quasi-Newton optimization algorithm.

Both techniques are flexible in their implementation. For example, arbitrary basis functions such as polynomials, splines, sigmoids, sinusoids, or radial basis functions can be used. In fact, a multi-layer neural network can also be used to represent the nonlinear mapping. Furthermore, the inputs to the nonlinear mapping can consist of either measurements of exogenous signals or measurements of system outputs that are fed back to the system through a nonlinear mapping. Both approaches can be used to identify nonlinear maps of multiple arguments and with multiple outputs.

2. Problem Formulation

Consider the nonlinear discrete-time system

$$x_{k+1} = Ax_k + F(u_k, y_k), \tag{2.1}$$
$$y_k = Cx_k + G(u_k), \tag{2.2}$$

where $x_k \in \mathbb{R}^n$, $u_k \in \mathbb{R}^m$, $y_k \in \mathbb{R}^p$, $A \in \mathbb{R}^{n\times n}$, $C \in \mathbb{R}^{p\times n}$, $F\colon\ \mathbb{R}^m \times \mathbb{R}^p \to \mathbb{R}^n$, and $G\colon\ \mathbb{R}^m \to \mathbb{R}^p$. The functions F and G can be written in terms of their scalar-valued components as

$$F(u,y) = \left[\begin{array}{c} F_1(u,y) \\ \vdots \\ F_n(u,y) \end{array}\right], \quad G(u) = \left[\begin{array}{c} G_1(u) \\ \vdots \\ G_p(u) \end{array}\right], \tag{2.3}$$

where, for all $i = 1, \ldots, n$, $F_i\colon\ \mathbb{R}^m \times \mathbb{R}^p \to \mathbb{R}$ and, for all $i = 1, \ldots, p$, $G_i\colon\ \mathbb{R}^m \to \mathbb{R}$. By defining

$$z \triangleq \mathcal{N}(u,y) \triangleq \left[\begin{array}{c} F(u,y) \\ G(u) \end{array}\right], \tag{2.4}$$

the system (2.1), (2.2) can be illustrated as in Figure 5, where $\mathcal{N}\colon\ \mathbb{R}^m \times \mathbb{R}^p \to \mathbb{R}^{n+p}$ and $\mathfrak{L}$ represents the linear system

$$x_{k+1} = Ax_k + \left[\begin{array}{cc} I_n & 0 \end{array}\right] z_k, \tag{2.5}$$
$$y_k = Cx_k + \left[\begin{array}{cc} 0 & I_p \end{array}\right] z_k, \tag{2.6}$$

where $z_k \triangleq \mathcal{N}(u_k, y_k)$ is viewed as an unmeasured, exogenous input to $\mathfrak{L}$.

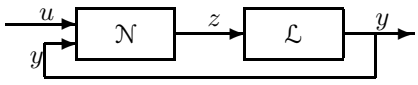

Fig. 5. Nonlinear System with Measured-Input Nonlinearities

The main feature of the model (2.5), (2.6) is the fact that all of the inputs to $\mathcal{N}$ are measured. Therefore, the model (2.1), (2.2) includes the Hammerstein and nonlinear feedback models shown in Figure 1 and Figure 3. However, (2.1), (2.2) does not encompass the Wiener system shown in Figure 2.

Next, we assume that the components F_i and G_i can be expanded in terms of basis functions $f_1(u,y), \ldots, f_q(u,y)$, $g_1(u), \ldots, g_r(u)$, and $h_1(u), \ldots, h_s(u)$ as

$$F(u,y) = \left[\begin{array}{c} \sum_{i=1}^q b_{f1i} f_i(u,y) + \sum_{i=1}^s b_{h1i} h_i(u) \\ \vdots \\ \sum_{i=1}^q b_{fni} f_i(u,y) + \sum_{i=1}^s b_{hni} h_i(u) \end{array}\right], \tag{2.7}$$

$$G(u) = \left[\begin{array}{c} \sum_{i=1}^r d_{g1i} g_i(u) + \sum_{i=1}^s d_{h1i} h_i(u) \\ \vdots \\ \sum_{i=1}^r d_{gpi} g_i(u) + \sum_{i=1}^s d_{hpi} h_i(u) \end{array}\right]. \tag{2.8}$$

The functions h_i are the basis functions that are common to both F and G. Defining $f\colon\ \mathbb{R}^m \times \mathbb{R}^p \to \mathbb{R}^q$, $g\colon\ \mathbb{R}^m \to \mathbb{R}^r$, and $h\colon\ \mathbb{R}^m \to \mathbb{R}^s$ by

$$f(u,y) = \left[\begin{array}{c} f_1(u,y) \\ \vdots \\ f_q(u,y) \end{array}\right], \ g(u) = \left[\begin{array}{c} g_1(u) \\ \vdots \\ g_r(u) \end{array}\right], \ h(u) = \left[\begin{array}{c} h_1(u) \\ \vdots \\ h_s(u) \end{array}\right],$$

it follows from (2.7) and (2.8) that

$$F(u,y) = B_f f(u,y) + B_h h(u), \tag{2.9}$$
$$G(u) = D_g g(u) + D_h h(u), \tag{2.10}$$

where $B_f \triangleq [b_{fij}] \in \mathbb{R}^{n\times q}$, $B_h \triangleq [b_{hij}] \in \mathbb{R}^{n\times s}$, $D_g \triangleq [d_{gij}] \in \mathbb{R}^{p\times r}$, and $D_h \triangleq [d_{hij}] \in \mathbb{R}^{p\times s}$. Thus (2.1), (2.2) can be written as

$$x_{k+1} = Ax_k + B_f f(u_k, y_k) + B_h h(u_k), \tag{2.11}$$
$$y_k = Cx_k + D_g g(u_k) + D_h h(u_k), \tag{2.12}$$

or more compactly as

$$x_{k+1} = Ax_k + B \left[\begin{array}{c} f(u_k,y_k) \\ g(u_k) \\ h(u_k) \end{array}\right], \tag{2.13}$$

$$y_k = Cx_k + D \left[\begin{array}{c} f(u_k,y_k) \\ g(u_k) \\ h(u_k) \end{array}\right], \tag{2.14}$$

where

$$B \triangleq \left[\begin{array}{ccc} B_f & 0 & B_h \end{array}\right], \quad D \triangleq \left[\begin{array}{ccc} 0 & D_g & D_h \end{array}\right]. \tag{2.15}$$

As a special case of the system shown in Figure 5, we can consider the Hammerstein system

$$x_{k+1} = Ax_k + F(u_k), \tag{2.16}$$
$$y_k = Cx_k + G(u_k), \tag{2.17}$$

where now the function F depends only on the input u. In the case that F and G are represented by a common set of basis function $h_1, \ldots, h_s$, it follows that

$$z = \begin{bmatrix} F(u) \\ G(u) \end{bmatrix} = \begin{bmatrix} B \\ D \end{bmatrix} h(u), \tag{2.18}$$

where $B = B_h \in \mathbb{R}^{n \times s}$ and $D = D_h \in \mathbb{R}^{p \times s}$. Hence (2.16), (2.17) become

$$x_{k+1} = Ax_k + Bh(u_k), \tag{2.19}$$

$$y_k = Cx_k + Dh(u_k). \tag{2.20}$$

The goal of the nonlinear identification problem is to construct models of both $\mathcal{L}$ and $\mathcal{N}$ given measurements of (u_k, y_k) over the interval $0 \leq k \leq \ell$. The signal z is assumed to be unavailable. However, when $h(u)$ is approximated by $\hat{h}(u)$ and B, D are approximated by $\hat{B}, \hat{D}$ then, the computed signal

$$\hat{z} \triangleq \begin{bmatrix} \hat{B} \\ \hat{D} \end{bmatrix} h(u), \tag{2.21}$$

is available as the input to $\mathcal{L}$.

3. Identification Algorithms

With the basis functions $f_i(u, y)$, $g_i(u)$, $h_i(u)$ specified, subspace identification algorithms [9, 13] can be applied directly to the system (2.5), (2.6) with the computed signal $\hat{z}$ playing the role of the exogenous input. This is the approach developed in [8]. However, the choice of basis functions remains the main difficulty. The type of basis functions chosen (for example, polynomial or splines) will, in general, have a strong effect on the number of basis functions needed to achieve a satisfactory approximation of the nonlinear mappings in a particular application. Unfortunately, the subspace identification algorithm yields the coefficient matrices (A, B, C, D) but not the basis functions per se.

To address this difficulty, we consider the following two approaches, which we illustrate for the Hammerstein case (2.19), (2.20).

3.1. Selective Refinement Algorithm

To begin, consider an initial set of basis functions $\hat{h}_1, \ldots, \hat{h}_{\hat{s}}$ with $\hat{h} \triangleq \begin{bmatrix} \hat{h}_1 & \cdots & \hat{h}_{\hat{s}} \end{bmatrix}^{\mathrm{T}}$ and let $(\hat{A}, \hat{B}, \hat{C}, \hat{D})$ denote an estimate of (A, B, C, D) provided by the subspace identification algorithm. Next, consider the singular value decomposition of $\begin{bmatrix} \hat{B} \\ \hat{D} \end{bmatrix}$ written in standard notation as

$$\begin{bmatrix} \hat{B} \\ \hat{D} \end{bmatrix} = \hat{U}\hat{\Sigma}\hat{V}. \tag{3.1}$$

Then, we retain the ν largest singular values in $\hat{\Sigma}$ to obtain the approximation $\hat{\Sigma} \approx \hat{\Sigma}_0 = \hat{L}_0\hat{R}_0$, where $\operatorname{rank} \hat{\Sigma}_0 = \nu$ and the matrices $\hat{L}_0 \in \mathbb{R}^{(n+p)\times\nu}$ and $\hat{R}_0 \in \mathbb{R}^{\nu \times s}$ have full column rank and full row rank, respectively. The retained ν largest singular values can be incorporated into either $\hat{L}_0$ or $\hat{R}_0$, yielding the approximation

$$\begin{aligned} \begin{bmatrix} \hat{B} \\ \hat{D} \end{bmatrix} \hat{h}(u) &= \hat{U}\hat{\Sigma}\hat{V}\hat{h}(u) \approx \hat{U}\hat{\Sigma}_0\hat{V}\hat{h}(u) \\ &= \hat{U}\hat{L}_0\hat{R}_0\hat{V}\hat{h}(u) = \begin{bmatrix} \hat{B}_0 \\ \hat{D}_0 \end{bmatrix} \hat{h}_0(u), \end{aligned} \tag{3.2}$$

where the matrix $\begin{bmatrix} \hat{B}_0 \\ \hat{D}_0 \end{bmatrix} \triangleq \hat{U}\hat{L}_0 \in \mathbb{R}^{(n+p)\times\nu}$ and $\hat{h}_0(u) \triangleq \hat{R}_0\hat{V}\hat{h}(u)$ satisfying $\hat{h}_0\colon\ \mathbb{R}^m \to \mathbb{R}^\nu$ is a column vector consisting of ν scalar-valued nonlinear functions. The motivation for this procedure is to retain only ν scalar-valued nonlinear functions each of which is a linear combination of $\hat{s}$ basis functions. Since $\nu << \hat{s}$, the ν scalar-valued components of $\hat{h}_0$ can be viewed as *dominant nonlinearities*, while the choice of ν reflects the *rank* of the nonlinear mapping $\left[\begin{smallmatrix} F \\ G \end{smallmatrix}\right]$. Hence the number of dominant nonlinearities is effectively the rank of the nonlinear map.

To refine the mapping $\hat{h}_0$ we repeat the above procedure with a new set of basis functions $\hat{h}'_1, \ldots \hat{h}'_{\hat{s}'}$ with $\hat{h}' \triangleq \begin{bmatrix} \hat{h}'_1 & \cdots & \hat{h}'_{\hat{s}'} \end{bmatrix}^{\mathrm{T}}$, where $\hat{h}'_1, \ldots, \hat{h}'_\nu$ are chosen to be the ν components of $\hat{h}_0$, and $\hat{h}'_{\nu+1}, \ldots \hat{h}'_{\hat{s}'}$ are chosen randomly. Repeating the above procedure yields a new estimate $\hat{h}'_0$ and the approximation

$$\begin{bmatrix} \hat{B}' \\ \hat{D}' \end{bmatrix} \hat{h}'(u) \approx \begin{bmatrix} \hat{B}'_0 \\ \hat{D}'_0 \end{bmatrix} \hat{h}'_0(u), \tag{3.3}$$

where $\hat{B}'$, $\hat{D}'$ are the estimates of B and D obtained from the subspace identification algorithm at the current iteration. Note that the components of the dominant nonlinearity $\hat{h}'_0$ are now linear combinations of $\hat{s} + \hat{s}'$ basis functions. However, the number of scalar components is fixed at ν.

The selective refinement algorithm is implemented with several options. Specifically the number s of additional basis functions that are introduced at each iteration is manually chosen by the user. In addition, the number ν of dominant nonlinearities retained at each step is also a critical parameter. This parameter can be manually specified by the user or automatically set by numerical criteria. The random selection of additional basis functions is automated for radial basis functions as is the bookkeeping needed to keep track of the accumulated basis functions that contribute to the dominant nonlinearities.

3.2. Basis Function Optimization Algorithm

In the basis function optimization algorithm we optimize a fixed set of basis functions instead of introducing additional basis functions. A convenient choice of basis functions is radial basis functions because of the ease of programming and the ability to handle multi-dimensional inputs. Radial basis functions are of the form

$$f(u) = e^{-\alpha\|u-c\|_2^2}, \tag{3.4}$$

where α determines the spread of the function and c decides the center of the function. For nonlinear identification, we optimize a set of radial basis functions with respect to the parameters α and c and identify the linear dynamics using a subspace identification algorithm. By optimizing a fixed set of basis functions, a more accurate representation of the nonlinear mapping is obtainable with a smaller number of basis functions than is possible with the selective refinement algorithm.

The identification error is defined to be the mean square error at the output y_k given by

$$E(\alpha, c) = \frac{1}{2}\sum_{k=1}^{l}(y_k - \hat{y}_k)^2, \tag{3.5}$$

where y_k and $\hat{y}_k$ are the desired and actual outputs of the identified Hammerstein system, and l is the length of the data set.

Now, writing $\hat{y}_k$ in terms of $\hat{A}, \hat{B}, \hat{C}$ and $\hat{D}$, (3.5) becomes

$$\begin{aligned} E(\alpha,c) &= \frac{1}{2}\sum_{k=1}^{l}(y_k - \hat{C}\hat{A}^k\hat{x}_0 \\ &\quad - \sum_{i=0}^{k-1}\hat{C}\hat{A}^{k-i-1}\hat{B}\hat{h}(u_i) - \hat{D}\hat{h}(u_k))^2. \end{aligned} \tag{3.6}$$

Using a set of s radial basis functions for $\hat{h}(u)$ equation (3.6) becomes

$$\begin{aligned} E(\alpha,c) &= \frac{1}{2}\sum_{k=1}^{l}\left(y_k - \hat{C}\hat{A}^k\hat{x}_0 - \sum_{i=0}^{k-1}\hat{C}\hat{A}^{k-i-1}\hat{B}\begin{bmatrix} e^{-\alpha_1\|u_i-c_1\|_2^2} \\ \vdots \\ e^{-\alpha_s\|u_i-c_s\|_2^2} \end{bmatrix} \right. \\ &\quad \left. -\hat{D}\begin{bmatrix} e^{-\alpha_1\|u_k-c_1\|_2^2} \\ \vdots \\ e^{-\alpha_s\|u_k-c_s\|_2^2} \end{bmatrix} \right)^2. \end{aligned} \tag{3.7}$$

The gradient of $E(\alpha, c)$ with respect to the parameters α_j and c_j can be calculated as

$$\begin{aligned} \frac{\partial E}{\partial \alpha_j} &= (y_k - \hat{y}_k)\sum_{k=1}^{l}\left(-\sum_{i=0}^{k-1}\hat{C}\hat{A}^{k-i-1}\hat{B}\begin{bmatrix} 0 \\ \vdots \\ \frac{\partial}{\partial\alpha_j}e^{-\alpha_j\|u_i-c_j\|_2^2} \\ \vdots \\ 0 \end{bmatrix} \right. \\ &\quad \left. -\hat{D}\begin{bmatrix} 0 \\ \vdots \\ \frac{\partial}{\partial\alpha_j}e^{-\alpha_j\|u_k-c_j\|_2^2} \\ \vdots \\ 0 \end{bmatrix} \right) \end{aligned} \tag{3.8}$$

and

$$\begin{aligned} \frac{\partial E}{\partial c_j} &= (y_k - \hat{y}_k)\sum_{k=1}^{l}\left(-\sum_{i=0}^{k-1}\hat{C}\hat{A}^{k-i-1}\hat{B}\begin{bmatrix} 0 \\ \vdots \\ \frac{\partial}{\partial c_j}e^{-\alpha_j\|u_i-c_j\|_2^2} \\ \vdots \\ 0 \end{bmatrix} \right. \\ &\quad \left. -\hat{D}\begin{bmatrix} 0 \\ \vdots \\ \frac{\partial}{\partial c_j}e^{-\alpha_j\|u_k-c_j\|_2^2} \\ \vdots \\ 0 \end{bmatrix} \right). \end{aligned} \tag{3.9}$$

Since

$$\begin{aligned} \frac{\partial}{\partial\alpha_j}e^{-\alpha_j\|u-c_j\|_2^2} &= -e^{-\alpha_j\|u-c_j\|_2^2}\|u-c_j\|_2^2, \\ \frac{\partial}{\partial c_j}e^{-\alpha_j\|u-c_j\|_2^2} &= -e^{-\alpha_j\|u-c_j\|_2^2}\alpha_j\left[2c_j^T - 2u^T\right], \end{aligned}$$

the gradients (3.8) and (3.9) can be evaluated.

By computing these gradients, a BFGS quasi-Newton optimization code is used to optimize the basis function parameters. Since the state space matrices and the basis function parameters cannot be estimated simultaneously, basis-function optimization and state space model identification are done alternately. First, an initial set of basis functions is chosen, and then the linear dynamics are identified using a subspace identification algorithm. Once the state space matrices are available, the set of basis functions is optimized. For the optimized set of basis functions, the linear dynamics is identified again, and so on.

4. EXAMPLES

4.1. Example 1: Hammerstein System with Scalar Input

For this example we consider a Hammerstein system whose linear dynamics are given by the discrete-time simple harmonic oscillator

$$\begin{aligned} A &= \begin{bmatrix} 0 & 1 \\ -(\omega^2 T_s^2 + 1) & 2 \end{bmatrix}, \quad B = \begin{bmatrix} 0 \\ T_s^2 \end{bmatrix}, \\ C &= \begin{bmatrix} 1 & 0 \end{bmatrix}, \quad D = 0, \end{aligned}$$

where $\omega = 0.7$ and $T_s = 0.1$, with input nonlinearity $\mathcal{N}(u) = u^2$. A total of 1000 data points are used for the identification, and both algorithms were used.

For the selective refinement algorithm, we choose 11 radial basis functions to initialize the algorithm and include 10 random radial basis functions at each subsequent iteration. A total of 2000 iterations are performed, of which 12 are accepted as determined by the data fit decrement. The dominant nonlinearity, involving 131 radial basis functions, is shown in Figure 6.

The basis functions optimization algorithm is employed with sines and cosines and 15 of each are used. The linear system order is specified as 2, and a single dominant nonlinearity is retained at each iteration. The data fit is shown in Figure 7 and the dominant identified nonlinearity is shown in Figure 8.

4.2. Example 2: Hammerstein System with Scalar Input and Rank-2 Nonlinearity

This example is a Hammerstein system based on the discrete-time simple harmonic oscillator of Example 1 cascaded with the low pass filter $\frac{1+z^{-1}}{230-228z^{-1}}$. The scalar input is taken to be a white noise signal with input nonlinearities

$$f(u) = \begin{bmatrix} u^3 & e^{-u} \end{bmatrix}.$$

The 3rd-order system has the realization

$$\begin{aligned} A &= \begin{bmatrix} 0.9913 & 1 & 0 \\ 0 & 0 & 0.9901 \\ 0 & -0.995 & 1.9802 \end{bmatrix}, \quad B = \begin{bmatrix} 0 & 0 \\ 0 & 1 \\ 0.01 & 0 \end{bmatrix}, \\ C &= \begin{bmatrix} 0.0087 & 0.0043 & 0 \end{bmatrix}, \quad D = \begin{bmatrix} 0 & 0 \end{bmatrix}. \end{aligned}$$

A total of 1000 data points are used for the identification, and the basis function optimization algorithm is implemented with radial basis functions. 15 radial basis functions are used to initialize the algorithm and a total of 10 iterations are performed. The subspace identification algorithm identified a 3rd-order system.

The data fit is shown in Figure 9, and the corresponding dominant nonlinearities involving 15 radial basis functions are shown Figure 10 and Figure 11. These nonlinearities provide estimates of the input nonlinearities u^3 and e^{-u}, respectively.

4.3. Example 3: Nonlinear Feedback System with Scalar Input

For this system we consider the nonlinear feedback system

$$y(k+1) = \tfrac{1}{2}\mathrm{sat}(y_k) + u_k,$$

with u_k chosen to be a white noise signal for $k = 1, \ldots, 1000$. Using 11 radial basis functions to initialize the identification and with 5 additional radial basis functions introduced at each iteration, 100 selective refinement iterations are performed of which 5 are accepted based on fit error decrement. The final estimated nonlinearity in Figure 12 is thus a linear combination of 36 radial basis functions.

4.4. Example 4: Hammerstein Model of Space Weather System

This example is based on data used for space weather prediction. The input data set was measured by the NASA Advanced Composition Explorer (ACE) spacecraft and includes the three components of the magnetic field vector, the solar wind speed, solar wind proton density, and temperature. The system output is the cross polar cap potential, which is derived from 85 magnetometers located in Greenland, Canada, Scandinavia, Alaska, and Russia.

For the data fit shown in Figure 13, 3 of the 7 inputs (one component of the solar wind velocity and two components of the magnetic field) are used to construct a rank 2 input nonlinearity, that is, a Hammerstein system with two dominant nonlinearities. No linear input functions are used.

A total of 20 iterations of the selective refinement algorithm are run, with 3 updates accepted as determined by fit improvement. A total of 27 radial basis functions are used on the first iteration, with 27 radial basis functions introduced at each iteration. The final rank 2 input nonlinearity is thus a combination of 108 radial basis functions. The linear dynamics identified by the subspace algorithm are second order. Figure 13 shows the best fit obtained, occurring on the 18th iteration.

5. Conclusions

We developed a nonlinear identification technique for a class of systems with measured-input nonlinearities. A subspace identification algorithm was used to identify the linear dynamics with the nonlinear mappings represented as a linear combination of basis functions. A selective-refinement technique and a quasi-Newton optimization algorithm were used to iteratively improve the representation of the system nonlinearity. In the first approach, a singular value decomposition of the input matrix is used to identify the dominant nonlinearities for an initial set of basis functions. Random basis functions were then introduced to improve the representation of the dominant nonlinearities through selective refinement. In the second approach, the basis function parameters were optimized using a BFGS quasi-Newton optimization algorithm. In this case, the representation of the nonlinearity is refined by optimizing a fixed number of basis functions of a chosen type. Splines, sinusoids, and radial basis functions were used as basis functions.

The selective refinement algorithm is inefficient as would be expected, but provides a simple baseline technique for comparison with alternative methods. While the use of a BFGS quasi-Newton algorithm limits convergence to local minima, the algorithm worked well for the examples that we considered. From a function approximation point of view, the use of quasi-Newton optimization is more efficient than purely gradient-based optimization.

6. Acknowledgments

We wish to thank Robert Clauer, Aaron Ridley, and Tamas Gombosi of the University of Michigan Space Physics Laboratory for providing the space weather data.

References

[1] E.-W. Bai. An Optimal Two-Stage Identification Algorithm for Hammerstein-Wiener Nonlinear Systems. *Automatica*, 34(3):333–338, 1998.

[2] C. H. Chen and S. D. Fassois. Maximum Likelihood Identification of Stochastic Wiener-Hammerstein-Type Non-Linear Systems. *Mech. Sys. Sig. Proc.*, 6(2):135–153, 1992.

[3] H. Diaz and A. A. Desrochers. Modeling of Nonlinear Discrete-Time Systems from Input-Output Data. *Automatica*, 24:629–641, 1988.

[4] R. Haber and L. Keviczky. *Nonlinear System Identification – Input-Output Modeling Approach*, volume 1: Nonlinear System Parameter Identification. Kluwer Academic Publishers, 1999.

[5] R. Haber and L. Keviczky. *Nonlinear System Identification – Input-Output Modeling Approach*, volume 2: Nonlinear System Structure Identification. Kluwer Academic Publishers, 1999.

[6] Z. Hasiewicz. Identification of a Linear System Observed Through Zero-Memory Non-Linearity. *Int. J. Sys. Sci.*, 18(9):1595–1607, 1987.

[7] A. Juditsky et al. Nonlinear Black-Box Models in System Identification: Mathematical Foundations. *Automatica*, 31(12):1725–1750, 1995.

[8] S. L. Lacy and D. S. Bernstein. Subspace Identification for Nonlinear Systems That Are Linear in Unmeasured States. In *Proc. Conf. Dec. Contr.*, pages 3518–3523, Orlando, Florida, December 2001.

[9] M. Moonen, B. De Moor, L. Vandenberghe, and J. Vandewalle. On- and Off-Line Identification of Linear State-Space Models. *Int. J. Contr.*, 49(1):219–232, 1989.

[10] R. K. Pearson. *Discrete-Time Dynamic Models*. Oxford University Press, 1999.

[11] C. A. Schwarz and B. W. Dickinson. On Finite Dimensional Realization Theory of Discrete Time Nonlinear Systems. *Sys. Contr. Lett.*, 7:117–123, 1986.

[12] J. Sjoberg et al. Nonlinear Black-Box Modeling in System Identification: A Unified Overview. *Automatica*, 31(12):1691–1724, 1995.

[13] P. Van Overschee and B. De Moor. *Subspace Identification for Linear Systems: Theory, Implementation, Applications*. Kluwer, 1996.

[14] T. Van Pelt and D. S. Bernstein. Nonlinear System Identification Using Hammerstein and Nonlinear Feedback Models with Piecewise Linear Static Maps. *Int. J. Contr.*, 74:1807–1823, 2001.

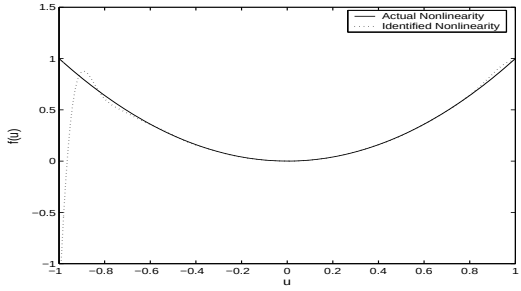

Fig. 6. Example 1: True (u^2) and identified input nonlinearities for second-order Hammerstein system with scalar input and rank 1 nonlinearity using selectively refined radial basis functions.

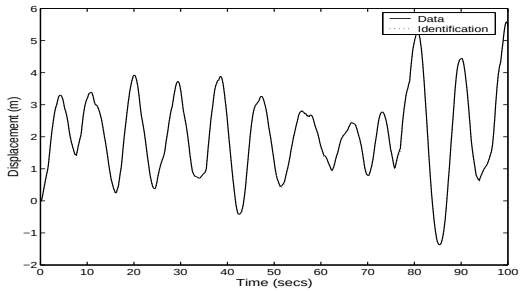

Fig. 7. Example 1: Data fit for second-order Hammerstein system with scalar input and rank 1 nonlinearity using optimized sines and cosines.

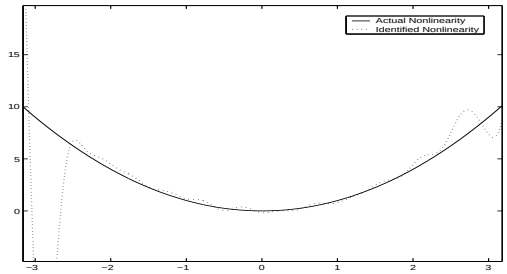

Fig. 8. Example 1: True (u^2) and identified input nonlinearity for second-order Hammerstein system with scalar input and rank 1 nonlinearity using optimized sines and cosines.

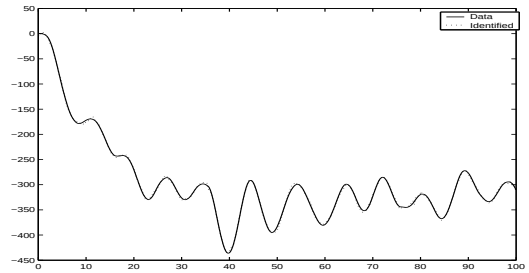

Fig. 9. Example 2: Data fit for third-order Hammerstein system with scalar input and rank 2 nonlinearity using optimized radial basis functions.

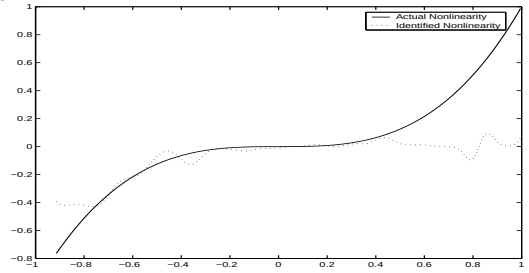

Fig. 10. Example 2: True (u^3) and identified input nonlinearities for third-order Hammerstein system with scalar input and rank 2 nonlinearity using optimized radial basis functions.

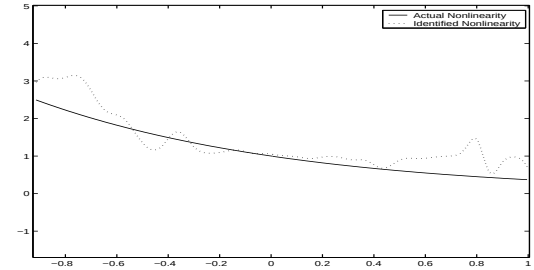

Fig. 11. Example 2: True (e^{-u}) and identified input nonlinearities for third-order Hammerstein system with scalar input and rank 2 nonlinearity using optimized radial basis functions.

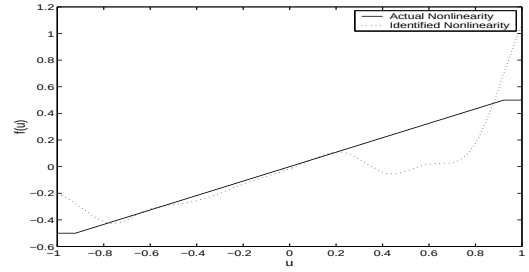

Fig. 12. Example 3: Identified feedback nonlinearity for first-order nonlinear feedback system using selective refinement with radial basis functions.

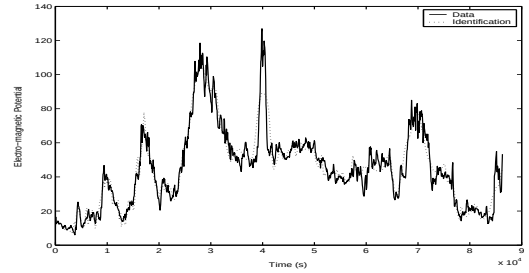

Fig. 13. Example 4: Data fit for space weather system with three inputs and a rank 2 nonlinearity.