Proceedings of the 7th Mediterranean Conference on Control and Automation (MED99) Haifa, Israel - June 28-30, 1999

Nonlinear Identification of Automobile Vibration Dynamics *

David T. Westwick† Koshy George‡
Michel Verhaegen§
Systems and Control Engineering Group
Faculty of Information Technology and Systems
Delft Technical University

Abstract

The identification of nonlinear state-space systems from input-output measurements is considered. The system is separated into a linear state-space system with a static nonlinearity, driven by the state and input, in feedback. Initially, the contribution from the nonlinearity is treated as an unknown system input driving an otherwise linear plant. A neural network is then used to model the feedback nonlinearity. A realistic simulation of a nonlinear automobile suspension is used to demonstrate the application of the identification algorithm.

1 Introduction

Fatigue and vibration testing of automobiles is often performed using a 4-post "test-rig" (De Cuyper *et al.*, 1998), where the test vehicle sits on 4 hydraulic actuators which are used to replicate motions recorded during a previous test-drive. Typically, recordings are made of the axle and chassis accelerations and suspension displacement at each of the 4 corners of the vehicle. The hydraulic actuators are then driven such that the suspension displacements, and axle and chassis accelerations track those measured in the field. Thus, realistic body motions can be reproduced under controlled conditions, and, for fatigue testing, over extended periods of time.

The current solution to this "mission reproduction" involves an iterative, off-line procedure. First, an identification experiment is performed by exciting the system with a relatively broad-band noise input. A linear Frequency Response Function (FRF, essentially a non-parametric transfer function), is estimated from the test data. The target outputs are then filtered using the (frequency by frequency) inverse of the FRF, producing an initial input sequence.

If, as is often the case, the initial input sequence does not cause the system to track the target outputs adequately, the inputs are refined. The inverse FRF is applied to the tracking error, and the result is added to the test input. This off-line refinement process iterates until sufficient tracking accuracy is obtained.

The overall goal is to replace this off-line iterative tuning process with an online controller. The currently envisioned controller will include a feedback component, for disturbance rejection, and a feed-forward (i.e. system inversion) element. Clearly, the design of the feed-forward element will require an accurate model of the car/test-rig system, which includes any significant nonlinearities present in the

*This research was supported by the EC Brite-Euram programme through project nr.BE-97-4186 (SCOOP)
†Email: *D.Westwick@its.tudelft.nl*
‡Email: *K.George@its.tudelft.nl*
§Email: *M.Verhaegen@its.tudelft.nl*

system. Since this model will depend on the characteristics of the particular car attached to the test rig, it will have to be identified from experimental data.

In this paper, we will consider the identification of such a car/test-rig model. Since the model will be incorporated into an online controller, a relatively small, parametric form is desirable. However, since it will be incorporated into a feed-forward controller, it must be capable of performing "free run" simulations, rather than simply generating one-step-ahead predictions, as is the case with input-output models such as the NARMAX model class (Chen and Billings, 1989). Thus, we will consider the identification of a class of nonlinear state-space systems.

1.1 Nonlinear State Space Models

Consider the following nonlinear state-space model,

$$x_{k+1} = F(x_k, u_k) \tag{1}$$
$$y_k = Cx_k + Du_k + v_k \tag{2}$$

where $x_k \in \mathbb{R}^{\mathrm{n}}$ is the state, $y_k \in \mathbb{R}^{\mathrm{p}}$ is the output, and the input, $u_k \in \mathbb{R}^{\mathrm{m}}$. The measurement noise, $v_k \in \mathbb{R}^{\mathrm{p}}$, is assumed to be a white-noise sequence that is independent of the input, u_k, and the initial state, x_0. With the exception of the (linear) direct transmission term, this is exactly the nonlinear state-space model considered by Haykin (1999), albeit in a neural-network setting. While a model with the structure Eq. (1 – 2) can be trained using a suitably designed iterative optimization, such as the back propagation through time algorithm, the resulting error surface may be very complex, including many suboptimal local minima. Our goal is to develop a non-iterative procedure for the identification of nonlinear state-space models, or at least to limit identification to well defined optimizations with relatively simple cost-functions. Ideally, the model should be sufficiently accurate to be used without further refinement. However, the results may also be used to initialize an iterative optimization procedures with a nearly optimal model, thus minimizing the possibility of converging to a sub-optimal local minimum.

In its general form, the nonlinear state-space model can be used to represent a wide variety of nonlinear systems ranging from the voltages and currents in a nerve cell (Hodgkin and Huxley, 1952), to stiction and friction forces in a robot arm (Kim *et al.*, 1997). In this paper, we will develop an algorithm for identifying the dynamics of a subset of this class of models, the restrictions on the sub-class being determined by several of the procedures used in the identification. We note that similar approaches have been used in the design of nonlinear observers (Kim *et al.*, 1997; Zhang *et al.*, 1998) for a restricted class of SISO nonlinear systems.

2 Identification of a Class of Nonlinear State Space Systems

In this section, we describe the identification approach, and present the theoretical foundations for the procedures that are employed.

2.1 Linear/Nonlinear Decomposition

First, we extract a linear subsystem from the nonlinear state-space model. Let A and B be constant matrices of appropriate dimensions. Then, we may write Eq. (1 – 2) as,

$$\begin{aligned} x_{k+1} &= Ax_k + Bu_k + \tilde{F}(x_k, u_k, A, B) \\ y_k &= Cx_k + Du_k + v_k \end{aligned} \tag{3}$$

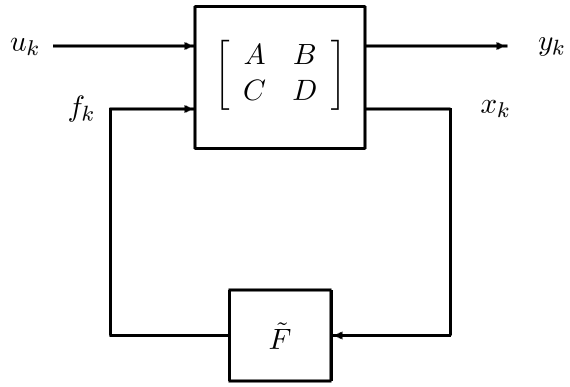

Figure 1: Linear/Nonlinear decomposition of a nonlinear state-space model

Clearly, $F(x_k, u_k) = Ax_k + Bu_k + \tilde{F}(x_k, u_k, A, B)$, so that the choice of A and B is in some sense arbitrary. In the sequel, we will focus on one particular choice for A and B.

Compute the RQ factorization of a matrix whose k'th column is a stack consisting of the input u_k, state, x_k, and subsequent state, x_{k+1},

$$Z_k = \begin{bmatrix} u_1 & u_2 & \dots & u_{N-1} \\ x_1 & x_2 & \dots & x_{N-1} \\ x_2 & x_3 & \dots & x_N \end{bmatrix} = \begin{bmatrix} r_{11} & 0 & 0 \\ r_{21} & r_{22} & 0 \\ r_{31} & r_{32} & r_{33} \end{bmatrix} \begin{bmatrix} q_1 \\ q_2 \\ q_3 \end{bmatrix} \tag{4}$$

By construction, we have

$$x_{k+1} = r_{32} r_{22}^{-1} x_k + \left(r_{31} - r_{32} r_{22}^{-1} r_{21} \right) r_{11}^{-1} u_k + r_{33} q_3$$

where q_3 is orthogonal to the current input and state sequences.

Hence by choosing

$$\begin{aligned} A &= r_{32} r_{22}^{-1} \\ B &= \left(r_{31} - r_{32} r_{22}^{-1} r_{21} \right) r_{11}^{-1} \end{aligned}$$

in (3) we see that the contribution of the term $\tilde{F}(x_k, u_k, A, B) = r_{33} q_3$ will be uncorrelated (but clearly not independent of) the input and state. Thus, we can re-express the nonlinear state-space model (1 – 2) as a linear system with an auxiliary input,

$$\begin{aligned} x_{k+1} &= Ax_k + Bu_k + \tilde{B} f(x_k, u_k) \\ y_k &= Cx_k + Du_k + v_k \end{aligned} \tag{5}$$

where $\tilde{B} f_k$ is the $k'th$ column of $r_{33} q_3$.

2.2 overview of identification approach

Equation (5) can be viewed as a multiple-input linear system, between $[u_k^T f_k^T]^T$ and y_k. The first m inputs, u_k, are the measured system inputs. The second group of inputs, the "auxiliary" inputs, f_k, are generated by the nonlinearities in the system. In section 2.3, subspace methods will be used to identify the linear dynamics, treating the "auxiliary" inputs as if they were process noise. This will be shown to lead to a bias in the identified linear dynamics.

In section 2.4, the conditions necessary for f_k to and $\tilde{B}$ to be estimated from the linear residuals will be investigated, as well as a procedure for estimating $\tilde{B}$. Of critical importance is the requirement that the auxiliary input be a static nonlinear function of the current system state, and input. Section 2.5 will present a robust method for system inversion, which will be used to reconstruct the auxiliary input, f_k. Finally, in section 2.6, a neural network will be used to fit a static nonlinearity between the system state and the auxiliary input. The overall algorithm will be summarized in section 2.7.

2.3 Identification of the Linearized Model

Given records of u_k and y_k, one might attempt to estimated the quadruple (A, B, C, D) using linear system identification techniques. Direct, non-iterative estimation of the system matrices can be accomplished using subspace methods (Verhaegen and DeWilde, 1992; Verhaegen, 1993, 1994; Viberg, 1995).

The basic subspace schemes are unable to cope with either measurement noise or process noise, and are hence not suitable to this identification. Verhaegen (Verhaegen, 1993, 1994) has proposed schemes that use instrumental variables to remove the bias caused by measurement and process noise. However, even these schemes may be biased by the presence of the state nonlinearity.

Consider the MOESP family of subspace methods for identification (Verhaegen and DeWilde, 1992; Verhaegen, 1993, 1994). First, the input and output are placed in Hankel matrices,

$$\mathbf{U}_{1,s,N} = \begin{bmatrix} u_1 & u_2 & \dots & u_{N-s+1} \\ u_2 & u_3 & \dots & u_{N-s+2} \\ \vdots & \vdots & & \vdots \\ u_s & u_{s+1} & \dots & u_N \end{bmatrix} \tag{6}$$

where the subscripts refer to the index of the first data sample, the number of rows in the matrix, and the index of the final data sample in the matrix. The number of rows, s, in the matrix is user-specified *over-dimension* parameter which is chosen to be greater than the system order, n. The output Hankel matrix, $\mathbf{Y}_{1,s,N}$ is defined analogously. Then we observe that for a *linear* system, we could write,

$$\mathbf{Y} = \Gamma X + H\mathbf{U} + \mathbf{V} \tag{7}$$

where $\mathbf{V}$ is a Hankel matrix constructed from the output noise sequence, v_k, Γ is the extended observability matrix, defined as

$$\Gamma = \begin{bmatrix} C \\ CA \\ \vdots \\ CA^{s-1} \end{bmatrix} \tag{8}$$

the matrix $X = \begin{bmatrix} x_1 & x_2 & \dots & x_{N-s+1} \end{bmatrix}$, and H is a lower triangular matrix containing the system's Markov parameters,

$$H = \begin{bmatrix} D & 0 & 0 & \dots & 0 \\ CB & D & 0 & \dots & 0 \\ CAB & CB & D & & 0 \\ \vdots & \ddots & \ddots & & \vdots \\ CA^{s-2}B & \dots & \dots & CB & D \end{bmatrix} \tag{9}$$

For the nonlinear system (5), the input, output and auxiliary input may be related by a similar equation,

$$\mathbf{Y} = \Gamma X + H\mathbf{U} + \tilde{H}\mathbf{F} + \mathbf{V} \tag{10}$$

where $\mathbf{F}$ is a Hankel matrix constructed from samples of the auxiliary input, f_k, and $\tilde{H}$ is defined analogously to (9), replacing B with $\tilde{B}$, and D with 0.

Consider the following RQ factorization,

$$\begin{bmatrix} \mathbf{U}_{s+1,s,N} \\ \mathbf{U}_{1,s,N-s} \\ \mathbf{Y}_{s+1,s,N} \end{bmatrix} = \begin{bmatrix} R_{11} & 0 & 0 \\ R_{21} & R_{22} & 0 \\ R_{31} & R_{32} & R_{33} \end{bmatrix} \begin{bmatrix} Q_1 \\ Q_2 \\ Q_3 \end{bmatrix} \tag{11}$$

and note that the factor $R_{32} = \mathbf{Y}_{s+1,s,N} Q_2^T$ is given by

$$R_{32} = \Gamma X Q_2^T + \tilde{H}\mathbf{F} Q_2^T + \mathbf{V} Q_2^T$$

and note that $\mathbf{V}Q_2^T$ will vanish as $N \to \infty$, since the measurement noise is assumed to be independent of the input. However, the term $\tilde{H}\mathbf{F}Q_2^T$ cannot, in general, be expected to vanish. For a *linear* system, $\mathbf{F}$ would be zero, and R_{32} would contain an unbiased estimate of the column space of Γ. Although using past inputs as instrumental variables removes the bias caused by measurement noise (and process noise, since it is independent of the input), they do not remove the bias due to the nonlinearity.

The usual subspace procedure (Verhaegen, 1993) is then followed. Thus Γ is extracted from R_{32} using a SVD, and then used to construct estimates of A and C. The remaining matrices, B and D can be obtained from a least-squares regression (Westwick and Verhaegen, 1996).

Remark 1 *To minimize the bias due to the nonlinearity, we must ensure that f_k is a zero-mean sequence. This can easily be accomplished by adding a constant row to the top of the data matrix (11).*

Remark 2 *Once an estimate, $\hat{f}_k$, of the auxiliary signal, f_k, is available, the linear identification may be repeated using the extended input $[u_k^T \, \hat{f}_k^T]^T$, to reduce the bias.*

2.4 Tracking the Contribution due to the Nonlinearity

Once an estimate of the linear system, A, B, C, D, is available, the next task is to estimate the "auxiliary" signal generated by the nonlinearity, f_k, which so far has been treated as a process disturbance. Consider the output of the identified linear model

$$\begin{aligned} x_{l,k+1} &= Ax_{l,k} + Bu_k \\ \hat{y}_{l,k} &= Cx_{l,k} + Du_k \end{aligned} \tag{12}$$

where the subscript l indicates the state and output of the *linearized* model. The linear residual sequence is

$$z_{f,k} = y_k - \hat{y}_{l,k} \tag{13}$$

$$\phantom{z_{f,k}} = y_{f,k} + v_k \tag{14}$$

Our task is to generate an input sequence, $\hat{f}_k$, and input matrix, $\tilde{B}$, such that $\hat{y}_{f,k}$, given by

$$\begin{aligned} \hat{x}_{f,k+1} &= A\hat{x}_{f,k} + \tilde{B}\hat{f}_k \\ \hat{y}_{f,k} &= C\hat{x}_{f,k} \end{aligned} \tag{15}$$

tracks the nonlinear output component, $y_{f,k}$. First, we must determine the conditions under which the system (15) can track the ideal residual sequence, $y_{f,k}$, given noise corrupted measurements $z_{f,k}$.

Definition 1 *A sequence $\{\vartheta_i\}_{i=0}^N$ in $\mathrm{I\!R}^{\mathrm{p}}$ is said to be output trackable if there exists an input sequence $\{\hat{f}_k\}_{i=0}^N$ such that the output sequence $\{\hat{y}_{f,k}\}_{i=0}^N$ of the system (15) satisfies $\|\vartheta - Y\|_2 = 0$ where*

$$
\begin{aligned}
\vartheta &\triangleq \begin{pmatrix} \vartheta_0' & \vartheta_1' & \cdots & \vartheta_N' \end{pmatrix}' \\
Y &\triangleq \begin{pmatrix} y_{f,0}' & y_{f,1}' & \cdots y_{f,N}' \end{pmatrix}'
\end{aligned}
$$

■

For the system (15) let

$$
\tilde{H} \triangleq \begin{pmatrix} 0 & 0 & 0 & \cdots & 0 \\ C\tilde{B} & 0 & 0 & \cdots & 0 \\ CA\tilde{B} & C\tilde{B} & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ CA^{N-1}\tilde{B} & CA^{N-2}\tilde{B} & CA^{N-3}\tilde{B} & \cdots & 0 \end{pmatrix} \tag{16}
$$

Lemma 3 *An arbitrary sequence $\{\vartheta_i\}_{i=0}^N$ in $\mathrm{I\!R}^{\mathrm{p}}$ is output trackable if, and only if, $\vartheta \in \mathrm{Im}\,\tilde{H}$.*

■

The proof of this result readily follows from the definitions of output trackability and $\tilde{H}$.

In particular, consider a square plant ($\tilde{B}$ in $\mathrm{I\!R}^{\mathrm{n}\times\mathrm{p}}$). If c_i denotes the ith row of C, then the system is said to have a well-defined relative degree $r \triangleq \begin{pmatrix} r_1 & r_2 & \cdots & r_p \end{pmatrix}'$ if

$$
c_i A^l \tilde{B} = 0 \quad \forall\, l < r_i - 1, \quad 1 \le i \le p
$$

and the matrix

$$
\begin{pmatrix} c_1 A^{r_1 - 1}\tilde{B} \\ c_2 A^{r_2 - 1}\tilde{B} \\ \vdots \\ c_p A^{r_p - 1}\tilde{B} \end{pmatrix}
$$

is nonsingular.

Corollary 3.1 *For a square system with a well-defined relative degree r, any arbitrary sequence $\{\vartheta_i\}_{i=0}^N$ in $\mathrm{I\!R}^{\mathrm{p}}$ is output trackable if, and only if,*

$$
\vartheta_{j,i} = 0, \quad 0 \le i \le r_j - 1, \quad 1 \le j \le p
$$

■

Note, however, that $\tilde{B}$ has not yet been estimated. For an arbitrary $\tilde{B} \in \mathrm{I\!R}^{\mathrm{n}\times\mathrm{p}}$, (provided that results in system with well defined relative degree $r = 1$), there will exist a p-dimensional sequence $\hat{f}_k$ that results in perfect tracking of $z_{f,k}$. Hence, since both $\tilde{B}$ and f_k are unknown, tracking ability alone is not sufficient to guarantee their correct identification. Thus, given a system with p outputs, we can uniquely identify the contributions due to at most $p-1$ nonlinearities in the state. For such a system, $\tilde{B}$ can be determined by by minimizing the tracking error

$$
\hat{B} = \min_{B} \| (I - H_T^{\dagger}) Y_T \|_2 \tag{17}
$$

where the subscript T indicates that matrix has been truncated by removing its top row. While the computations are burdensome, we note that fast orthogonalization techniques (Korenberg, 1988) can be used to reduce the computational and storage requirements dramatically. Furthermore, this is a separable least squares optimization (Golub and Pereyra, 1973; Ruhe and Wedin, 1980).

2.5 Stable Dynamic Inversion (SDI)

Given an estimate of $\tilde{B}$, we now describe an approach that inverts the model of the plant whilst overcoming the main drawbacks of existing techniques, namely, that of handling either non-minimum phase zeros, or zeros on the unit circle, and noise in the desired signal. Since the desired signal is the linear residual, its SNR will be much lower than that of the raw measurements. Thus, noise performance is of utmost importance in this application.

Recall that the linear residuals, $y_{f,k}$, are assumed to originate from the system (15), and that we have measurements, $z_{f,k}$ that also contain an additive white noise process, v_k, with covariance R_v, that is independent of the input and initial condition.

Problem 4 Given the above state space model of a plant, a target time history $\{z_{f,k}\}_{k=1}^{N}$, determine $\left\{\hat{f}_k\right\}_{k=1}^{N}$ such that

$$\hat{y}_{f,k} = y_{f,k} \quad k = 1, \ldots, N$$

where $\hat{y}_{f,k}$ is the response of the system (15) to the input $\hat{f}_k$.

This is illustrated in Fig. 2. In this figure, the 'System' is represented by eqn. (15). Our objective is to generate a desired input sequence $\hat{f}_k$ by obtaining a suitable 'inverse system' that relates this input sequence to the given measurement sequence $z_{f,k}$. With the input sequence $\hat{f}_k$ thus generated, the output $\hat{y}_{f,k}$ of the given system (15) should reasonably represent the desired output sequence $y_{f,k}$.

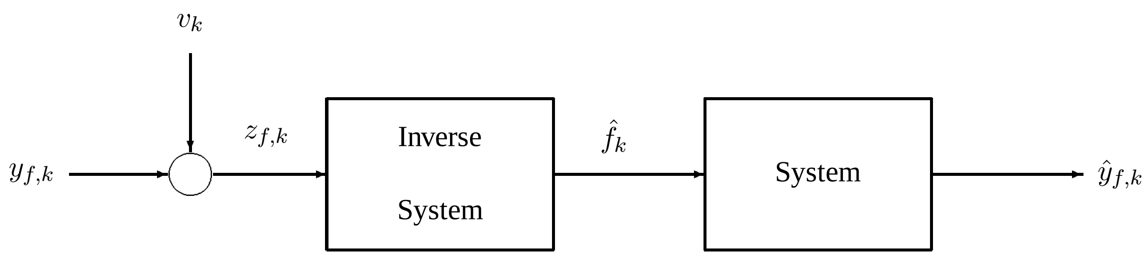

Figure 2: Stable dynamic model inversion.

The SDI technique is based on augmenting the given state space model (15) by a reasonable model for the input sequence and then designing a Kalman filter to provide an estimate of the input sequence from the measurements $z_{f,k}$. From intuitive and physical reasoning, it seems realistic to model the input signal $\hat{f}_k$ as follows

$$\hat{f}_{k+1} = \hat{f}_k + \eta_k \tag{18}$$

for some η_k. For simplicity, we assume that η_k is white noise with covariance Q_η, and uncorrelated with v_k. The resulting augmented system is as follows:

$$\begin{pmatrix} x_{k+1} \\ f_{k+1} \end{pmatrix} = \begin{pmatrix} A & \tilde{B} \\ 0 & I \end{pmatrix} \begin{pmatrix} x_k \\ f_k \end{pmatrix} + \begin{pmatrix} 0 & 0 \\ 0 & I \end{pmatrix} \begin{pmatrix} v_k \\ \eta_k \end{pmatrix} \tag{19}$$

$$z_k = \begin{pmatrix} C & 0 \end{pmatrix} \begin{pmatrix} x_k \\ f_k \end{pmatrix} + \begin{pmatrix} I & 0 \end{pmatrix} \begin{pmatrix} v_k \\ \eta_k \end{pmatrix} \tag{20}$$

In compact form, we have,

$$\begin{aligned} x_{a,k+1} &= A_a x_{a,k} + G_a w_{a,k} \\ z_k &= C_a x_{a,k} + H_a w_{a,k} \end{aligned}$$

where the definition of each of the individual matrices and vectors is rather obvious. The following result shows that if the original system is observable, then the augmented system is observable for almost all points in the complex plane; the proof of the result readily follows from the definitions of observability and the zeros of a system.

Lemma 5 *Suppose the pair (C, A) is observable. The pair $\left(\begin{pmatrix} C & 0 \end{pmatrix}, \begin{pmatrix} A & B \\ 0 & I \end{pmatrix} \right)$ is observable if, and only if, $z = 1$ is not a zero of the system $(A, B, C, 0)$.*

■

Therefore, under the conditions for observability of the augmented system given in Lemma 5 we can set up a Kalman filter to estimate the input signal:

$$\hat{x}_{a,k+1|k} = A_a \hat{x}_{a,k|k-1} - K \left(C_a \hat{x}_{a,k|k-1} - z_k \right)$$

where $K \triangleq \begin{pmatrix} K_1 \\ K_2 \end{pmatrix} \triangleq A_a P C_a' \left(C_a P C_a' + R_v \right)^{-1}$ and P satisfies the following Riccati equation:

$$P = A_a P A_a' - A_a P C_a' \left(C_a P C_a' + R_v \right)^{-1} C_a P A_a' + \begin{pmatrix} 00 & 0 \\ 0 & Q_\eta \end{pmatrix}$$

Here the Kalman gains K_1 and K_2 respectively correspond to the state x_k and the input f_k. We can easily see that $\hat{u}_{k+1|k} = \begin{pmatrix} 0 & I \end{pmatrix} \hat{x}_{a,k+1|k}$, and hence the transfer function from the measurements z_k to the estimate of the input $\hat{f}_{k|k-1}$ (i.e., the inverse system) has the following realization:

$$\text{Inverse System} = \begin{pmatrix} 0 & I \end{pmatrix} \left(zI - \begin{pmatrix} A - K_1 C & B \\ -K_2 C & I \end{pmatrix} \right)^{-1} K \qquad (21)$$

We note that the poles of this inverse system (i.e., the eigenvalues of $A_a - KC_a$) are the relocated eigenvalues of the system matrix A_a. Moreover, we have the following result on the zeros of the inverse system; its proof is rather straightforward from the definition of zeros:

Lemma 6 *Every eigenvalue of the system matrix A is a zero of the inverse system (21).*

■

Evidently, the augmented system will be observable at almost all points on the complex plane provided $\tilde{B}$ is chosen such that the LTI quadruple $(A, \tilde{B}, C, 0)$ has no zeros at $z = 1$. Thus, this technique is more general than the earlier schemes in that it does not restrict the presence of zeros to the stable and unstable regions in the strict sense (Devasia *et al.*, 1996; George *et al.*, 1999), or to the stable region (Hou and Patton, 1998). Moreover, by suitably designing the Kalman filter, we can easily take into account the presence of noise in target time histories. We recall that the method provided in (Devasia *et al.*, 1996) does not handle noise in target time histories, and the states corresponding to the zero dynamics are poorly re-constructed in the technique presented in (Hou and Patton, 1998).

2.6 Identification of the Nonlinearity

Thus far, the the linear dynamics of the separated model (3) have been identified, and the excitation provided by the remaining nonlinearity, $\tilde{F}(x_k, u_k, A, B) = \tilde{B} f_k$, has been estimated. In order to use this model in "free run", we need to be able to estimate the nonlinearity contribution, f_k, given the current state, x_k, and input, u_k.

First, the state of the system must be reconstructed, by simulating the system, using the extended input, $[u_k^T \hat{f}_k^T]^T$. Then, a traditional feed-forward sigmoidal neural network can be fitted from the input and estimated state to $\hat{f}_k$. This can be accomplished efficiently using the Levenberg Marquardt optimization procedure in a separable least squares framework (Sjöberg and Viberg, 1997).

Note, however, that in a free run simulation, the sequence $\hat{f}_k$ will be generated by the neural network, which will, in turn, influence the state. Thus, re-estimating $\tilde{B}$, and the neural network, using the state sequence, and auxiliary input resulting from a free run simulation may be beneficial.

2.7 Algorithmic Summary

Given length N records of input-output data from a system with m inputs and p outputs, and an upper bound, s, on the system order, n, we require the following additional assumptions,
Assumptions:

1. The input, u_k is persistently exciting of order at least $2s$, where s is the *over-dimension* parameter used in the subspace identification. Note that s must be chosen to be larger than the system order.

2. The system contains at most $p-1$ nonlinearities, where p is the number of outputs. Thus, it must be possible to find a representation of the state, x, such that at most $p-1$ components of the state have non-constant derivatives with respect to x and u. Note that this limit can be increase to p, if the system can be placed in a canonical form where $\tilde{B}$ is known *a priori*.

3. The measurement noise is independent of the input and state.

Algorithm

1. Fit a linear system between u_k and y_k using subspace methods, where the past inputs are used as instrumental variables (Verhaegen, 1993; Haverkamp and Verhaegen, 1997)

2. Estimate $\tilde{B}$, by solving the separable least squares minimization (17).

3. Use SDI, as described in Sec 2.5, to produce an auxiliary input, f_k, which causes the system $(A, \tilde{B}, CD)$ to track the linear residuals, $y_{f,k}$.

4. Fit a feed-forward NN which generates the auxiliary input, f_k, from the current state and input $[x_k u_k]$.

5. If the model is satisfactory, stop. Otherwise, re-estimate the linear subsystem using $[u_k^T \hat{f}(\hat{x}_k, u_k)^T]^T$ as the input. Note that because $\hat{f}_k$ is included in the extended input, this will also provide an updated estimate of $\tilde{B}$. Continue from step 3.

3 Simulations

3.1 Benchmark Example: A Quarter Car

The quarter-car model is the simplest model that includes a proper representation of the problem of controlling wheel load variations. It contains no representation of geometric effects of having four wheels and offers no possibility of studying longitudinal interactions or the use of front suspension state information to improve the performance at the rear. Moreover, it cannot describe problems related to handling. However, it does appear to contain the most basic features of the real problem and gives rise to design thinking which accords with experience. Thus, to demonstrate the application of our identification algorithm on a manageable, but nonetheless realistic, example, we will use a continuous-time simulation of a quarter-car model.

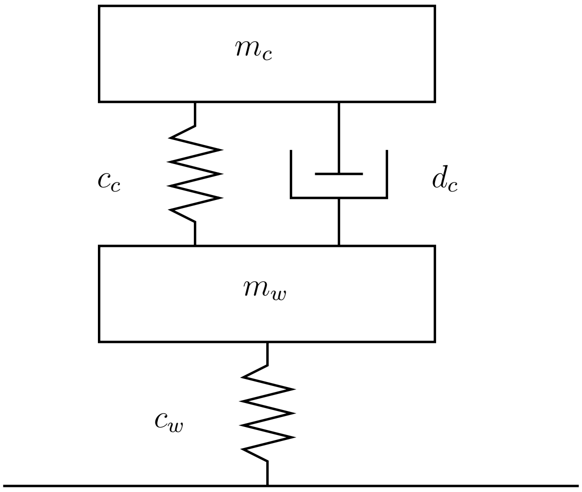

Figure 3: The quarter car model. Note that the suspension spring, c_c is nonlinear since it stiffens as it is removed from its equilibrium position.

The quarter car can be expressed by a two-mass model as shown in Fig. 3. It contains two vertical degrees of freedom: the displacement of the unsprung mass (or, axle), x_w, and displacement of the sprung mass (or, body), x_c. The road displacement input is denoted by x_b. The differential equations for a two degrees of freedom model of a quarter car are

$$\begin{aligned} m_c\ddot{x}_c &= -F_c(x_c - x_w) - d_c\,(\dot{x}_c - \dot{x}_w) \\ m_w\ddot{x}_w &= -c_w(x_w - x_b) + F_c(x_c - x_w) + d_c\,(\dot{x}_c - \dot{x}_w) \\ F_c(x) &= c_c\left(x + \alpha x^3 + \beta\tan(\pi x/2x_{max})\right) \end{aligned}$$

where the subscripts c and w respectively denote the body and the wheel of the car. Thus, m_c and m_w respectively represent the sprung mass and the unsprung mass of the car; c_c and c_w the suspension and tire stiffness. The only nonlinearity in this system is provided by the suspension spring, whose characteristic is given by F_c. Note that the tangent term in F_c causes the suspension travel to be limited to x_{max}. The system outputs are the acceleration of the masses, $\ddot{x}_c$ and $\ddot{x}_w$, and the suspension length, $x_c - x_w$.

The quarter-car model was driven by a band-limited (10 Hz) input, presented through a zero-order hold at 1000 Hz. Integration was performed using the fourth-order Runge-Kutta algorithm. The data was then re-sampled to 200 Hz for analysis.

The algorithm outlined in Sec. 2.7 was then used to fit a nonlinear state-space model to the input-output data. The accuracy of the nonlinear state-space model, as well as that of the initial linear model, was validated by comparing the system and model responses to a separate input signal. Model accuracy was expressed as the "Percent Variance Accounted For" (% VAF), defined as

$$\text{VAF} = \left(1 - \frac{\sum_{k=1}^{N}(y_k - \hat{y}_k)^2}{\sum_{k=1}^{N} y_k^2}\right) \times 100$$

Signal	Linear Model	Nonlinear Model
Suspension Displacement	75.0	95.4
Chassis Acceleration	74.2	97.6
Axle Acceleration	72.6	98.8

Table 1: Accuracy of the linear and non-linear state-space models obtained from free-run simulations on cross-validation data. Accuracies are expressed as % VAF

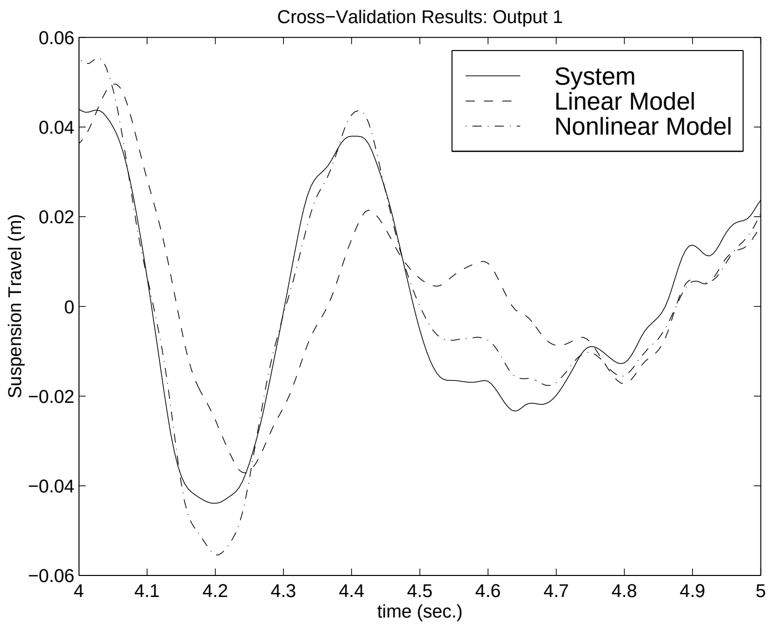

Figure 4: One second segment of the suspension displacement from the cross-validation data set (solid line). The output of the linear model (dashed line) and nonlinear state-space model (dash-dotted line) are superimposed.

where $\hat{y}_k$ is an estimate of the sequence $\hat{y}$. The results of these free run simulations are summarized in Table 1.

Figure 4, shows a 1 second segment of the suspension displacement (due to the validation input), superimposed on the outputs of the linear and nonlinear models. Figures 5 and 6 show the chassis and axle accelerations, respectively, for the same 1 second segment of the validation trial. Note that the model outputs (both linear and nonlinear models) are "free run" simulations, *not* one-step-ahead predictions. It is interesting to note that there are segments in the data (t between 4.7 and 5.0 seconds in Figs 4 – 6 for example), where the linear model performs extremely well. However, there are other instances, (t between 4.0 and 4.4 seconds) where the linear model fails completely. Thus, it is clear that the system is being occasionally driven out of its linear range.

4 Conclusions

We have developed a technique for identifying a restricted class of nonlinear state-space systems, a structure which can be used to model a wide variety of systems. The application considered in this research

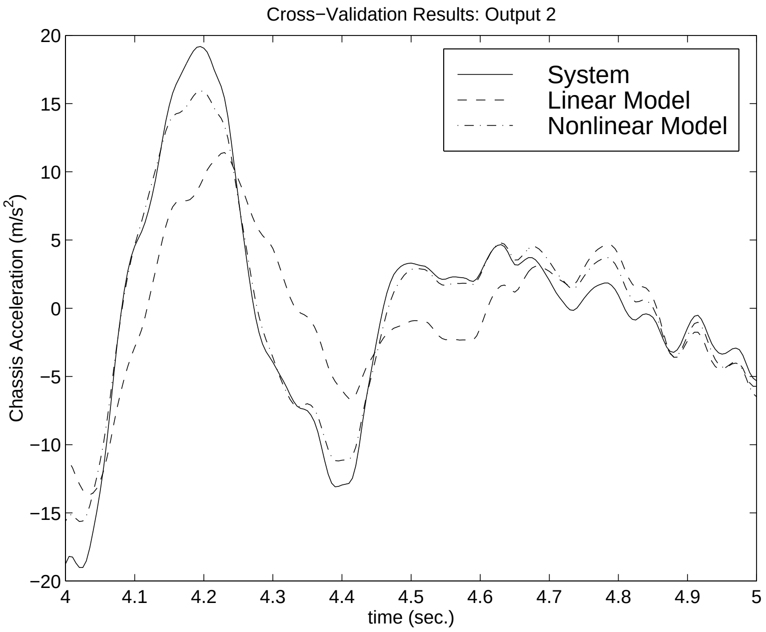

Figure 5: One second segment of the chassis acceleration from the cross-validation data set (solid line). The output of the linear model (dashed line) and nonlinear state-space model (dash-dotted line) are also shown.

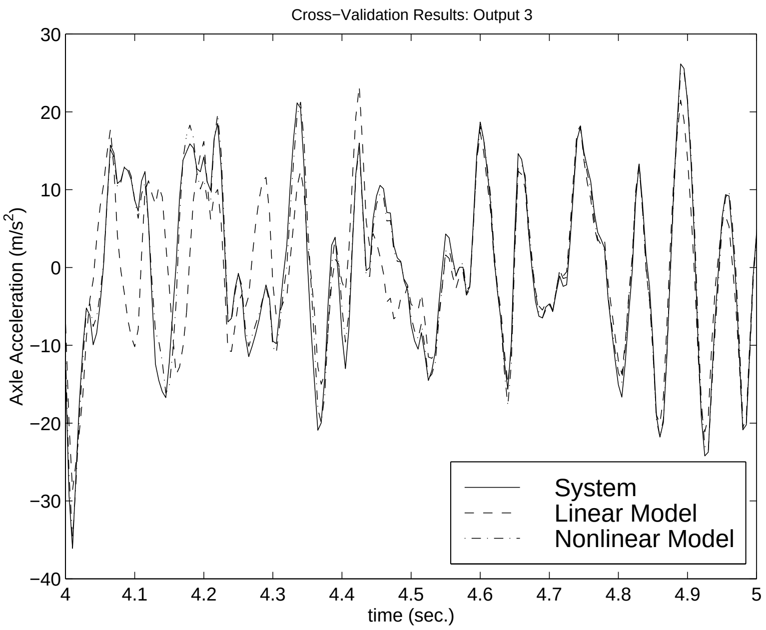

Figure 6: One second segment of the axle acceleration from the cross-validation data set (solid line). The output of the linear model (dashed line) and nonlinear state-space model (dash-dotted line) are superimposed.

was the identification of the dynamics of a car attached to a vibration test-rig. Here, the suspension springs, which are designed to stiffen as they approach the limits of the the suspension travel, are thought to provide the most significant nonlinearity in the system. In particular, it is a nonlinearity which affects the state directly, making the nonlinear state-space model an obvious model structure to employ.

The applicability of the identification algorithm was demonstrated using data from a continuous-time simulation model. The identified model was able to reproduce validation data sets with much greater accuracy than is possible with a linear model. We must emphasize that these were free run simulations, and not simply one-step-ahead predictions.

At present, the key restriction on the model class is that the "number" of nonlinearities must be less than the number of outputs, p. This constraint can be expressed more precisely in a variety of ways: the rank of the matrix r_{33} in (4) must be strictly less than p, for example. This restriction is due to the necessity to simultaneously estimate *both* the input matrix for the nonlinearities, $\tilde{B}$, and the corresponding auxiliary input sequence, f_k. Methods for determining $\tilde{B}$, when it is in $\mathbb{R}^{\mathrm{n} \times \mathrm{p}}$ are under development.

References

Chen, S. and S. Billings (1989). "Representation of non-linear systems: the NARMAX model," *International Journal of Control*, **49**, pp. 1013–1032.

De Cuyper, J., D. Coppens, C. Liefooghe, J. Swevers, and M. Verhaegen (1998). "Advanced drive file development methods for improved service load simulation on multi axial durability test rigs," in *Proceedings of the Asian Conference on Acoustics and Vibration*.

Devasia, S., D. Chen, and B. Paden (1996). "Nonlinear inversion-based output tracking," *IEEE Transactions on Automatic Control*, **41**, no. 7, pp. 930–942.

George, K., M. Verhaegen, and J. M. A. Scherpen (1999). "Stable inversion of MIMO linear discrete time non-minimum phase systems," in *The 7th IEEE Mediterranean Conference on Control and Automation*, Haifa, Israel.

Golub, G. and V. Pereyra (1973). "The differentiation of pseudo-inverses and nonlinear least squares problems whose variables separate," *SIAM Journal of Numerical Analysis*, **10**, no. 2, pp. 413–432.

Haverkamp, B. R. J. and M. G. H. Verhaegen (1997). *SMI Toolbox, State Space Model Identification Software for Multivariable Dynamical System*, Delft University of Technology, 1st edn.

Haykin, S. (1999). *Neural Networks a Comprehensive Foundation*, Prentice Hall, 2nd edn.

Hodgkin, A. and A. Huxley (1952). "A quantitative description of membrane current and its application to conduction and excitation in nerve," *Journal of Physiology*, **117**, pp. 500–544.

Hou, M. and R. J. Patton (1998). "Optimal filtering for systems with unknown inputs," *IEEE Transactions on Automatic Control*, **43**, no. 3, pp. 445–449.

Kim, Y., F. Lewis, and C. Abdallah (1997). "A dynamic recurrent neural-network-based adaptive observer for a class of nonlinear systems," *Automatica*, **33**, no. 8, pp. 1539–1543.

Korenberg, M. (1988). "Identifying nonlinear difference equation and functional expansion representations: The fast orthogonal algorithm," *Annals of Biomedical Engineering*, **16**, pp. 123–142.

Ruhe, A. and P. Wedin (1980). "Algorithms for separable nonlinear least squares problems," *SIAM Review*, **22**, no. 3, pp. 318–337.

Sjöberg, J. and M. Viberg (1997). "Separable non-linear least squares minimization – possible improvements for neural net fitting," in *IEEE Workshop on Neural Networks for Signal Processing*, vol. 7, pp. 345–354.

Verhaegen, M. (1993). “Subspace model identification part 3. analysis of the ordinary output-error state-space model identification algorithm,” *International Journal of Control*, **58**, no. 3, pp. 555–586.

Verhaegen, M. (1994). “Identification of the deterministic part of MIMO state space models given in innovations form from input-output data,” *Automatica*, **30**, no. 1, pp. 61–74.

Verhaegen, M. and P. DeWilde (1992). “Subspace model identification part 1. the output-error state-space model identification class of algorithms,” *International Journal of Control*, **56**, no. 5, pp. 1187–1210.

Viberg, M. (1995). “Subspace-based methods for the identification of linear time-invariant systems,” *Automatica*, **31**, no. 12, pp. 1835–1851.

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

Zhang, H., C. Chan, and H. Cheung (1998). “Nonlinear observer design with unknown nonlinearity via B-spline network approach,” in *Proceedings of the Americal Control Conference*, pp. 2339–2343.