

---

## Extending the Kalman filter for structured identification of linear and nonlinear systems

---

Matthew C. Best\* and Karol Bogdanski

Department of Aeronautical and Automotive Engineering,  
Loughborough University,  
Ashby Road, Loughborough, LE11 3TU, UK  
Email: m.c.best@lboro.ac.uk  
Email: k.bogdanski@lboro.ac.uk

\*Corresponding author

**Abstract:** This paper considers a novel approach to system identification which allows accurate models to be created for both linear and nonlinear multi-input/output systems. In addition to conventional system identification applications, the method can also be used as a black-box tool for model order reduction. A nonlinear Kalman filter is extended to include slow-varying parameter states in a canonical model structure. Interestingly, in spite of all model parameters being unknown at the start, the filter is able to evolve parameter estimates to achieve 100% accuracy in noise-free test cases, and is also proven to be robust to noise in the measurements. The canonical structure ensures a well-conditioned model which simultaneously provides valuable dynamic information to the engineer. After extensive testing of a linear example, the model structure is extended to a generalised nonlinear form, which is shown to accurately identify the handling response of a full vehicle model.

**Keywords:** system identification; model order reduction; Kalman filter; linear systems; nonlinear systems.

**Reference** to this paper should be made as follows: Best, M.C. and Bogdanski, K. (2017) 'Extending the Kalman filter for structured identification of linear and nonlinear systems', *Int. J. Modelling, Identification and Control*, Vol. 27, No. 2, pp.114–124.

**Biographical notes:** Matthew C. Best received his BEng and PhD from Loughborough University, UK. He is a Senior Lecturer in Vehicle Dynamics and Control in the Department of Aeronautical and Automotive Engineering at Loughborough. He has published over 50 papers on system identification, estimation and control of vehicles and drivers, and has a particular interest in the use of Kalman filters for both state and parameter estimation.

Karol Bogdanski received his BSc in Mechanical Engineering from the University of Padua, Italy and MSc in Automotive Systems Engineering from Loughborough University, UK. He is currently pursuing his PhD in System Identification and Kalman Filtering at Loughborough, as part of the Jaguar Land Rover/EPSC funded Programme for Simulation Innovation (PSi).

---

### 1 Introduction

Several publications explore system identification, most of which can be categorised by method, employing neural networks and/or genetic algorithms, frequency domain methods, probabilistic approaches and in a few cases Kalman filtering and recursive least squares. Numerous examples successfully employ neural networks to achieve excellent identified model performance, e.g., Ahmad (2015), but this is in a black-box format which offers no insight into the plant dynamics. They also typically employ high numbers of tuned parameters, which raises concern over parameter conditioning, repeatability and the range of inputs that will guarantee accuracy.

Classical alternatives can achieve excellent conditioning by employing smaller parameter sets, and indeed those mathematically proven to be minimal. These include the

well-known references on system identification (e.g., Soderstrom and Stoica, 1989) and also papers which use statistical probability methods (e.g., Aguero et al., 2012; Behzad et al., 2012; Hafayed et al., 2016). Though their effectiveness in solving both complex theoretical and practical identification cases is not questioned, these techniques are often too complicated for most engineers to apply in practice. Simpler techniques are available, e.g., using least-squares methods (Guo et al., 2015) or by identification in the frequency domain (e.g., Polifke, 2014; Ahn et al., 2003; Wang et al., 2015). The latter can be effective at replicating particular system resonances, particularly if they are well separated in frequency. However, they rely on sequential identification of SISO models and the combination of these into accurate multi input/output time domain models is not always easily achieved.

The well-known Kalman filter has also been applied to system identification in previous publications, but in most cases has only been effective in identifying a subset of unknown parameters in an existing known model structure. Examples include (Kallapur et al., 2008; Hassani et al., 2009; Best, 2007; Best et al., 2000). In this paper, we re-visit the Kalman filter, but address many of the concerns above, identifying a MIMO model of unknown structure and minimum parameter set using an iterative time-domain approach. One existing publication (Ding et al., 2012) operates in a similar way to that considered here, except that it employs a recursive least-squares method and assumes a discrete model of known order. In the new method, the most appropriate model order is also identified, and we also see that the same Identifying Extended Kalman Filter (IEKF) can be configured to identify both linear and nonlinear systems. Ultimately, the best combination of good parameter conditioning and lowest order can be identified, and a further advantage lies in the modal canonical structure of the resulting model, which has the advantage of revealing the most significant system eigenvalues.

## 2 The IEKF

The conventional extended Kalman filter (EKF) employs a nonlinear system model  $\mathbf{f}$  and sensor model  $\mathbf{h}$ , based on system state vector  $\mathbf{x}$ , to predict sensor set  $\mathbf{y}$  in response to known input set  $\mathbf{u}$  and model parameters  $\boldsymbol{\theta}$ . At a given time instant  $k$ :

$$\dot{\mathbf{x}}_k = \mathbf{f}(\mathbf{x}_k, \mathbf{u}_k, \boldsymbol{\theta}_k) + \boldsymbol{\omega}_k \quad (1)$$

$$\mathbf{y}_k = \mathbf{h}(\mathbf{x}_k, \mathbf{u}_k, \boldsymbol{\theta}_k) + \mathbf{v}_k \quad (2)$$

An optimal filter, typically employed to provide accurate state estimates and/or filter the outputs, can be found from estimates of the error covariance matrices,

$$\mathbf{Q}_k = E(\boldsymbol{\omega}_k \boldsymbol{\omega}_k^T), \mathbf{S}_k = E(\mathbf{v}_k \mathbf{v}_k^T) = \mathbf{0}, \mathbf{R}_k = E(\mathbf{v}_k \mathbf{v}_k^T) \quad (3)$$

where the sequence  $\boldsymbol{\omega}$  accounts for modelling errors, and  $\mathbf{v}$  is the sensor model error and noise. The EKF also needs state derivative matrices (Jacobians) to be available at each time step, defined

$$\mathbf{F}_k = \left. \frac{\partial \mathbf{f}(\mathbf{x}, \mathbf{u}_k, \boldsymbol{\theta}_k)}{\partial \mathbf{x}} \right|_{\mathbf{x}=\hat{\mathbf{x}}_k} \quad (4)$$

$$\mathbf{H}_k = \left. \frac{\partial \mathbf{h}(\mathbf{x}, \mathbf{u}_k, \boldsymbol{\theta}_k)}{\partial \mathbf{x}} \right|_{\mathbf{x}=\hat{\mathbf{x}}_k}$$

Examples which use the EKF for both state and parameter estimation are given in Best (2007) and Best et al. (2000).

Here, we start by identifying an unknown linear model from time-histories of one or more input(s)  $\mathbf{u}$  and one or more output(s)  $\mathbf{y}$  only. We do not know the physics of the model so we do not know the state vector  $\mathbf{x}$ .

$$\begin{aligned} \dot{\mathbf{x}}_k &= \mathbf{A}\mathbf{x}_k + \mathbf{B}\mathbf{u}_k \\ \mathbf{y}_k &= \mathbf{C}\mathbf{x}_k + \mathbf{D}\mathbf{u}_k \end{aligned} \quad (5)$$

We start by imposing a structure to the system ( $\mathbf{A}$ ) matrix. This could be any structure which suitably minimises the number of parameters in the model. Modal and companion canonical forms are valid, and we choose the modal form here, as this allows a smaller range in the magnitude of parameters in the model, giving advantages in conditioning; we will see later that the EKF assumes equal parameter variation (error) statistics. The modal form defines  $\mathbf{A}$  in terms of its eigenvalues which also has advantages in providing information to the engineer on the system's resonance frequencies and damping. An appropriate structure for a five state model is

$$\mathbf{A} = \begin{bmatrix} \sigma_1 & \omega_1 & 0 & 0 & 0 \\ -\omega_1 & \sigma_1 & 0 & 0 & 0 \\ 0 & 0 & \sigma_2 & \omega_2 & 0 \\ 0 & 0 & -\omega_2 & \sigma_2 & 0 \\ 0 & 0 & 0 & 0 & \sigma_3 \end{bmatrix} \quad \mathbf{B} = \begin{bmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \\ b_{31} & b_{32} \\ b_{41} & b_{42} \\ b_{51} & b_{52} \end{bmatrix} \quad (6)$$

$$\mathbf{C} = \begin{bmatrix} c_{11} & c_{12} & c_{13} & c_{14} & c_{15} \\ c_{21} & c_{22} & c_{23} & c_{24} & c_{25} \end{bmatrix} \quad \mathbf{D} = \begin{bmatrix} d_{11} & d_{12} \\ d_{21} & d_{22} \end{bmatrix}$$

This is easily expanded or contracted depending on the known number of inputs and outputs and for different estimates of model order. The eigenvalues are initially set in complex conjugate pairs,  $\lambda = \sigma + j\omega$ ,  $\lambda = \sigma - j\omega$  as far as possible. This is a first assumption which is easily corrected to include more single (real) poles if necessary; as the filter identifies the model,  $\omega$  estimates will reduce to zero if the modes are non-oscillatory, at which point the structure can be modified in an obvious way.

A further constraint is then needed; in the form above, each state pair is excited by the inputs scaled by a pair of  $\mathbf{B}$  matrix parameters, and each output is then scaled again by a parameter pair in the  $\mathbf{C}$  matrix. If we consider a SISO case, only the  $\mathbf{B}$  or  $\mathbf{C}$  parameters need to be set to fully constrain the model; five of these parameters could be set arbitrarily. It happens that the required constraint for any  $n$  state model is to fix  $n$  of the  $c$  or  $b$  parameters however, many inputs or outputs there are. It is sensible to constrain only elements in the  $\mathbf{C}$  matrix, however; we need to allow some parameters to be at or approaching zero in  $\mathbf{B}$ , as any given input will not generally excite all states. Setting  $\mathbf{C} = [1 \ 1 \ 1 \ \dots]$  is appropriate in single output cases. For multiple outputs it is appropriate to 'connect' particular eigenvalue pairs to specific outputs; for example, in equation (6), choosing  $c_{11} = c_{12} = c_{23} = c_{24} = c_{25} = 1$  associates the first output most strongly with the first eigenvalue pair, and the second to the remaining eigenvalues. Since the order of the eigenvalues (and hence states) is not fixed in advance, this allows the identified parameters to emerge naturally in the order which best represents the dynamic information in the outputs; we will see in Section 4 that this approach minimises the risk of poor conditioning.



assumptions about system resonance frequencies. The 25 Hz low pass filter is not required for correct function of the IEKF, but it improves the speed of convergence by allowing the tuning parameter  $\rho$  to be set at a higher level. The offset on each input provides a zero frequency element in the I/O data which assists in the separate identification of the **D** matrix. One potential symptom of poor conditioning in the identification is propagation of one or more  $\sigma$  parameters to high (negative) values; unless very fast dynamics exist in the model, this approximates direct feed-through from input to output via **A**, **B**, **C** rather than **D**. Offset of the input signals helps to prevent this by promoting early convergence of **D** parameters.

No significant prior knowledge of the system has been assumed here; the 25 Hz filter frequency just needs to be ‘well above’ the highest dynamic frequencies of interest, and the appropriate maximum can easily be found for any plant by observation of output PSD in response to high bandwidth white noise inputs.

Output data is obtained by simulation of model equation (15) from zero initial conditions and with outputs sampled at the same rate as the inputs. In general, it is also wise to normalise **u** and **y** prior to identification as this maximises conditioning in the identified model parameters, but this is unnecessary in the given example as the outputs naturally emerge with similar magnitudes to the inputs.

The IEKF is initialised with ‘true’ states  $z_{1-n} = 0$ , with eigenvalue parameter states,  $z_{(n+1)-2n} = -1$  and with most of the remaining parameter states  $z_{(2n+1)-m} = 0$ . Constrained *c* parameters within  $z_{(2n+1)-m}$  are set to 1. This ensures the identified model initialises in a stable form but with no assumptions about the unknown parameters. Alternative starting conditions, using randomised stable initial parameter sets have been tested without variation of the converged model. A high setting of  $\rho = 1$  is used for the majority of the results discussed below and **P** can sensibly be initialised as **P**<sub>0</sub> = **Q**.

The available data is repeatedly applied to the IEKF. At the start of each iteration, the true states are (appropriately) reset to zero, but all the parameter states, **K**<sub>k</sub> and **P**<sub>k</sub> matrices are carried over from the last step of the previous iteration. This has the effect of ‘rinsing’ the parameters with the available I/O data, with the innovations slowly varying them to improve the fit and reduce expected parameter error **P**<sub>k</sub>.

## 4 Linear identification results

### 4.1 Convergence

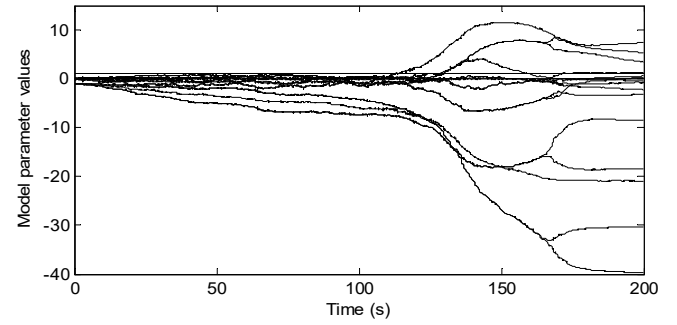
Consider a simple first test, identifying model equations (15) from the two inputs to just the first output. The first two iterations of the filter are illustrated in Figure 1, which shows all the parameters rapidly diverging before the majority start to converge.

A commonly used measure of performance, ‘percentage explanation’ can be used to quantify quality of fit in the output from any given model:

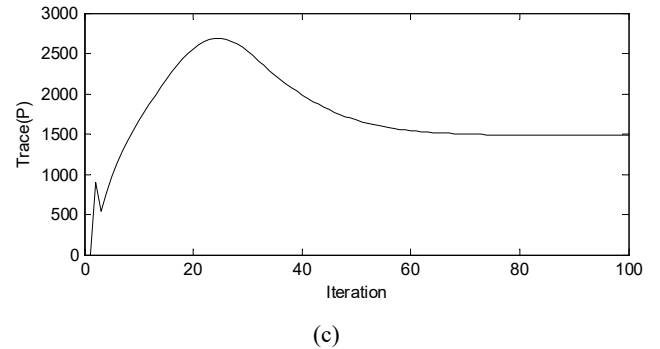
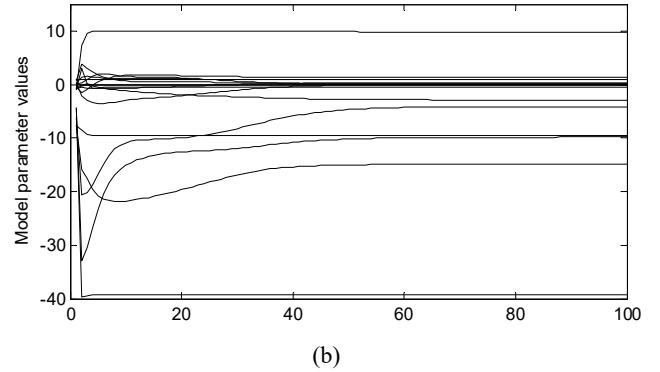
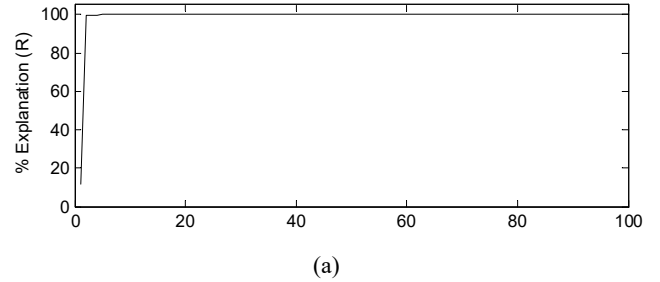
$$R = \left( 1 - \frac{\sum_k (y_k - \hat{y}_k)^2}{\sum_k y_k^2} \right) \times 100 \quad (17)$$

Taking the set of parameters at the end of iteration 2 (at 200 seconds above) as the (fixed) model provides output accuracy  $R = 99.4\%$ .

**Figure 1** Development of parameter values in the first two iterations

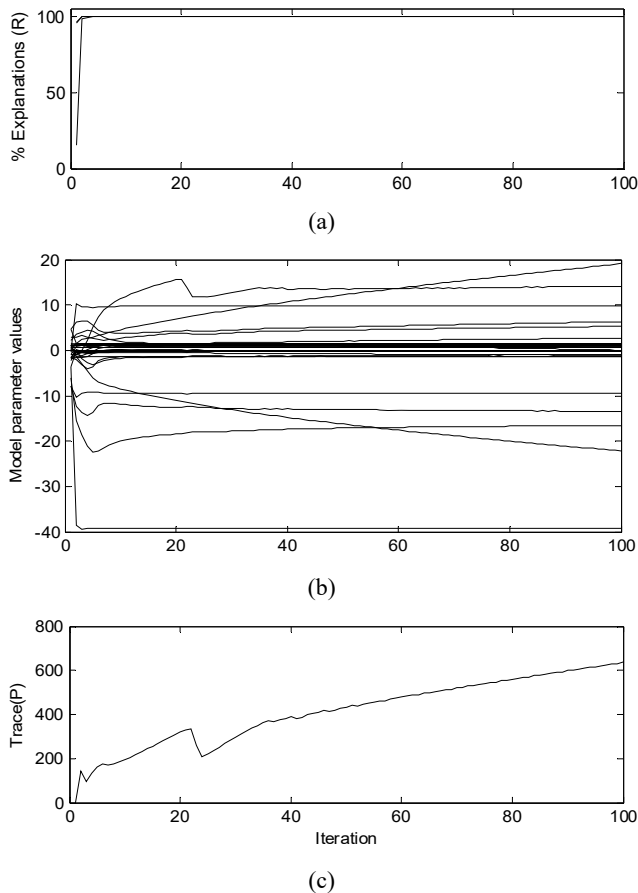


**Figure 2** Explanation and parameter convergence metrics (to one output)



A full optimisation of this case over many iterations is illustrated in Figure 2. Output explanation quickly approaches 100% (plot a) and the parameters take around 100 iterations to converge finally to a settled, final value (plot b). On the mid-range PC used to conduct these tests, 100 iterations takes around 4.4 minutes, so the process is not overly time consuming. The  $\mathbf{P}$  matrix is the covariance of expected error in the states, so when all parameter states converge, so do their expected error. The trace of  $\mathbf{P}$  provides a simple single value which can be used to confirm convergence (plot c), though alternatively a sum of the parameter states could be used for the same purpose. Note how different combinations of parameters achieve very close to 100% accuracy through iterations 2–100, yet the final parameter values conform very accurately to the original model (Table 1). Interestingly, some of the parameters prove to be insensitive in their influence on the output; parameter  $b_{51}$  is not accurate, and one of the eigenvalues has converged with real part  $-9.5$  rather than  $-8$ . This is also the case in most of the identification results shown later.

**Figure 3** Explanation and parameter convergence metrics (to three outputs)



If we repeat the exercise with the filter applied to all three outputs, the influence of individual parameter combinations adapting unexpectedly is illustrated further. Figure 3 shows parameter and trace( $\mathbf{P}$ ) divergence in spite of 100% accuracy, and Table 2 shows some unexpected results (highlighted), including a coupled pair of ill-conditioned

and relatively insensitive modes. Also note however that any model taken beyond around the 10th iteration is accurate and successful.

**Table 1** Final identified parameters (to one output)

Identified parameters (source model parameters)						
-4.09 (-4)	0	0	0	0	0.104 (0.1)	0.106 (0.1)
0	-14.89 (-15)	9.83 (10)	0	0	0.307 (0.3)	-2.87 (-3)
0	-9.83 (-10)	-14.89 (-15)	0	0	1.45 (1.5)	0.019 (0)
0	0	0	-9.51 (-8)	39.32 (40)	9.86 (10)	-0.413 (-0.5)
0	0	0	-39.32 (-40)	-9.51 (-8)	-0.096 (0.7)	1.02 (1)
1	1	1	1	1	0.0123 (0)	-0.002 (0)

**Table 2** Final identified parameters (to three outputs)

Identified parameters (source model parameters)						
-15.95 (-4)	0	0	0	0	0.029 (0.1)	-37.6 (0.1)
0	-13.94 (-15)	0.57 (10)	0	0	4.39 (0.3)	41.55 (-3)
0	-0.57 (-10)	-13.94 (-15)	0	0	-2.19 (1.5)	-6.94 (0)
0	0	0	-9.46 (-8)	39.32 (40)	9.75 (10)	-0.36 (-0.5)
0	0	0	-39.32 (-40)	-9.46 (-8)	-0.17 (0.7)	1.00 (1)
1	1	1	1	1	0.0108 (0)	-0.002 (0)
-0.367 (2)	-0.606 (0.3)	-1.55 (0.35)	-1.36 (-1.35)	-0.054 (-0.06)	1.49 (1.5)	0.670 (0.67)
7.25 (-2.75)	9.45 (-1.3)	16.84 (3.6)	1.41 (1.4)	0.756 (0.77)	-0.182 (-0.2)	1.40 (1.4)

Thus far, the  $c$  parameters have been constrained only in the first output, to allow easier comparison of identified parameters. However, the suggested method prescribes sharing the  $c$  constraints across all three outputs. It was suggested that constraining as

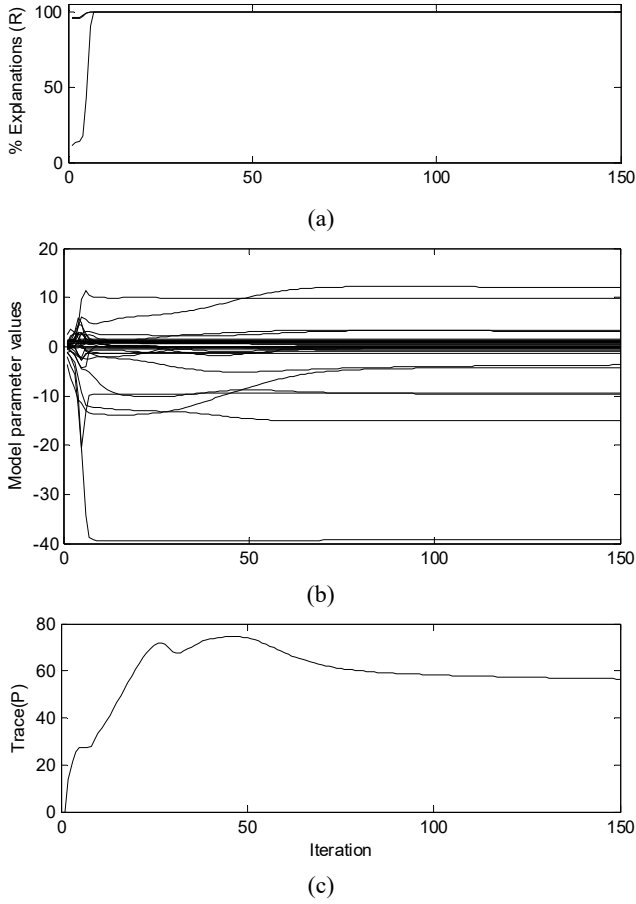
$$\mathbf{C} = \begin{bmatrix} 1 & 1 & c_{13} & c_{14} & c_{15} \\ c_{21} & c_{22} & 1 & 1 & c_{25} \\ c_{31} & c_{32} & c_{33} & c_{34} & 1 \end{bmatrix} \quad (18)$$

would have permitted a better conditioned model to evolve.

Indeed this is true; Figure 4 shows results with  $\mathbf{C}$  constrained as in equation (18), and Table 3 shows the final model. The revised constraints make many variables in the  $\mathbf{C}$  and  $\mathbf{B}$  matrices more difficult to compare. Also note that the eigenvalues (and hence states) are, in practice identified in a different order to that of equation (15). They have been re-ordered before presentation in the table, which explains

why the  $\mathbf{C}$  constraints do not appear the same in Table 3 as in equation (18). However, with this configuration of constraint, we see better convergence in  $\text{trace}(\mathbf{P})$ , the eigenvalues are now all correctly identified, and as in Table 1, only parameter  $b_{51}$  is now ‘incorrect’.

**Figure 4** Explanation and parameter convergence metrics (with better constraints to three outputs)



**Table 3** Final identified parameters (with better constraints to three outputs)

<i>Identified parameters (source model parameters only given here where comparison is meaningful)</i>									
-4.10 (-4)	0	0	0	0	-0.285	-0.290			
0	-14.93 (-15)	9.67 (10)	0	0	0.129	-0.923			
0	-9.67 (-10)	-14.93 (-15)	0	0	0.452	0.07			
0	0	0	-9.51 (-8)	39.32 (40)	9.86 (10)	-0.41 (-0.5)			
0	0	0	-39.32 (-40)	-9.51 (-8)	-0.103 (0.7)	1.02 (1)			
-0.360	3.33	2.98	1	1	0.0123 (0)	-0.002 (0)			
0.73	1	1	-1.35 (-1.35)	-0.060 (-0.06)	1.49 (1.5)	0.670 (0.67)			
1	-3.44	11.7	1.40 (1.4)	0.770 (0.77)	-0.181 (-0.2)	1.40 (1.4)			

## 4.2 Robustness to noise

Considering the single output case again, zero mean white noise is now added to the output, prior to identification in proportion to its RMS in ratios 10, 30, 50 and 70%. In all cases, the IEKF converged as previously, though with predictably lower performance (Table 4,  $R$  values). However, most of the error in these values is due to the added noise; when the converged models are compared with noise-free original data ( $R_0$  in the table) explanation is above 99% even in the worst case. Not shown here, the identified model parameters do vary slightly as noise increases, but the technique and ultimate performance of the models remains robust. Similar results were also seen from models identified for all three outputs. Of course any systematic bias in the noise would, unavoidably, result in a model with parameters identified to explain that bias.

**Table 4** Converged model explanations under the influence of noise in the output

<i>Explanation</i>	<i>Added noise ratio</i>			
	10%	30%	50%	70%
$R$ (of noisy data)	98.97	91.75	79.45	66.84
$R_0$ (of clean data)	99.96	99.97	99.60	99.90

## 4.3 Reduced and higher order models

In practice, we would not know the required order of the plant to be identified. However, we can reasonably assume that the desired model is that which delivers very high output explanations from the lowest possible order; this is certainly the objective in the identification of reduced-order models. Since the IEKF identification process is not overly time-consuming, we can establish the best order choice by running a series of identifications with incremental  $n$ . Consider for example, identifications from all three outputs of model equation (15) as follows:

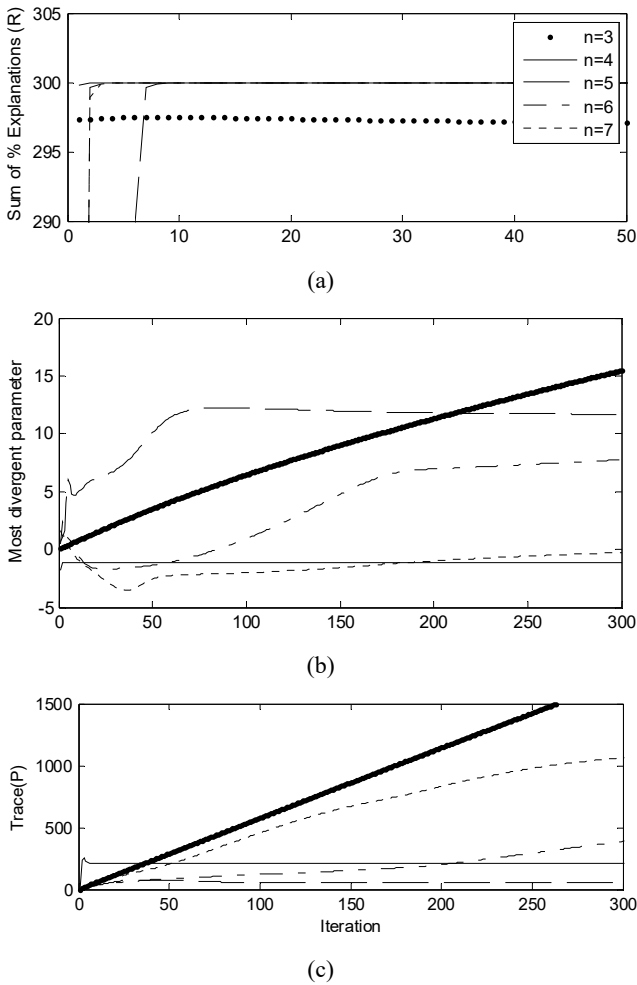
$$\begin{aligned}
 n=3, \mathbf{C} &= \begin{bmatrix} 1 & 1 & c \\ c & c & 1 \\ c & c & c \end{bmatrix} & n=4, \mathbf{C} &= \begin{bmatrix} 1 & 1 & c & c \\ c & c & 1 & 1 \\ c & c & c & c \end{bmatrix} \\
 n=6, \mathbf{C} &= \begin{bmatrix} 1 & 1 & c & c & c & c \\ c & c & 1 & 1 & c & c \\ c & c & c & c & 1 & 1 \end{bmatrix} & & \\
 n=7, \mathbf{C} &= \begin{bmatrix} 1 & 1 & c & c & c & c & c \\ c & c & 1 & 1 & c & c & c \\ c & c & c & c & 1 & 1 & 1 \end{bmatrix} & & 
 \end{aligned} \tag{19}$$

Accuracy,  $\text{trace}(\mathbf{P})$  and parameter convergence results are given for all five optimised models from 3rd to 7th order, in Figure 5. Some results in plots b and c have been scaled, since the intention is to illustrate their divergent or convergent nature, and for the same reason plot b illustrates only the most divergent parameter in the set.

These results provide further insight on the full output 5th order results seen earlier. Here, all order choices apart

from  $n = 4$  and  $n = 5$  result in divergent  $\mathbf{P}$  with some divergent or slowly varying parameters; conversely we see very fast convergence in the 4th order result. All results with  $n$  above three produce excellent explanations, and as we would expect, final accuracy increases with model order; clearly any well-posed model with more parameters will be able to fit the data with greater accuracy. The result at  $n = 3$  would be rejected due to poor performance, but results at  $n = 6$  or  $7$  are also undesirable, since  $\mathbf{P}$  diverges markedly, and the parameters continue to drift. This behaviour illustrates poor conditioning, with more parameters than are required to explain the outputs.

**Figure 5** Explanation and parameter convergence metrics (with varying  $n$ )



We know the original model has five eigenvalues, but the best trade-off between accuracy and a consistent (well-conditioned) identified model is achieved with four states. This is because the dynamic influence of the source model’s eigenvalue at  $-4$  is very weak in combination with the other modes; although it was derived arbitrarily, the source model itself has been defined with relatively poor conditioning. Identifications with five states and more than one output can yield unreliable estimates for that single eigenvalue, along with corresponding divergent parameters associated with it (in the  $\mathbf{B}$  and  $\mathbf{C}$  matrix).

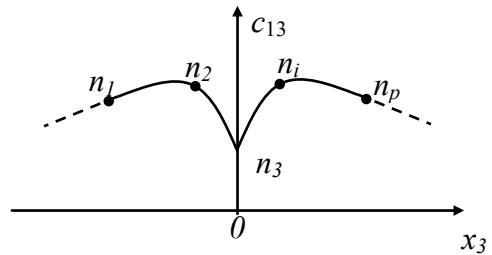
As the latter results show however, provided the objective of any real identification is extraction of the best possible model to describe the outputs, the four-state model here provides an excellent solution, regardless of the fact that it does not replicate the original model parameters. For comparison, its eigenvalues are identified at  $\lambda_{1,2} = -9.57 \pm 39.38j$  and  $\lambda_{1,2} = -12.68 \pm 10.27j$ .

**5 Extension to nonlinear model identification**

The EKF is naturally adapted to estimation of states and parameters of nonlinear models, and several authors have already used it for parameter or combined state and parameter estimation of structured nonlinear models (Kallapur et al., 2008; Hassani et al., 2009). The new IEKF is consequently also well-suited to identification of nonlinear models in the form of equations (1) and (2), and this can be achieved quite generally; for any smooth functions  $\mathbf{f}$ ,  $\mathbf{h}$ , it can operate using the same set of equations (4), (11) to (14). The only challenge is to prescribe a form that allows sufficient freedom to map the nonlinearity while ensuring adequate conditioning. General polynomial combinations (terms) of  $\mathbf{x}$ ,  $\mathbf{u}$  could be used, as could any differentiable function of them. However, the problem arises that (to retain conditioning) we cannot introduce a large number of terms at the outset, so a form of iterative term selection would be needed. Correlation methods could be used to select most significant terms, but with no known state structure at the outset, these methods cannot easily be used to decide the best candidates for inclusion in  $\mathbf{f}$  or  $\mathbf{h}$ . Also, as model order increases the number of potential terms increases disproportionately, and with an entirely free state structure the resulting model can no longer be interpreted modally.

Without dismissing the potential of an entirely free structure (if a suitable process for term selection can be developed) the nonlinearity considered here is applied on to the state space framework used earlier. Using cubic splines, each of the previously scalar parameters within the  $\mathbf{A}$ ,  $\mathbf{B}$ ,  $\mathbf{C}$  and  $\mathbf{D}$  matrices is allowed to vary across the domain of the state or input it multiplies. For example, parameter  $c_{13}$  might vary as illustrated in Figure 6.

**Figure 6** Nonlinear spline-based variation of example parameter  $c_{13}$



One advantage of this structure is that the splines are uniquely determined by a few node points. Introducing  $(\chi_i, \gamma_i)$  as the coordinate position of node  $i$ , if we fix the  $\chi$  ordinates such that the  $p$  nodes are spread evenly in  $x$ , each

of the model coefficients can be described by just  $p$  nodal  $\gamma$  ordinates. Also, although the number of parameters in the model then increases by a factor of  $p$ , each parameter remains well-conditioned as it represents a unique subspace in the range of  $\mathbf{x}$ ,  $\mathbf{u}$ . Note that in most cases the parameter will vary with state or input magnitude, as illustrated in Figure 6, but we do not constrain the nodes to reflect exactly in the  $\gamma$  axis. The resulting spline set has  $p + 1$  regions with a unique cubic function of  $x$  (or  $u$ ) describing the output from each. To allow predictable extrapolation the first and last regions are simplified to linear functions (dashed in the figure). For example,

$$c_{13} = (\eta_0 x + \delta_0) L_0 + (\alpha_1 x^3 + \beta_1 x^2 + \eta_1 x + \delta_1) L_1 + \dots + (\alpha_i x^3 + \beta_i x^2 + \eta_i x + \delta_i) L_i + \dots + (\eta_p x + \delta_p) L_p \quad (20)$$

where the  $L$  are logical operators,

$$\begin{aligned} L_0 &= x < \chi_1 \\ L_1 &= \chi_1 \leq x < \chi_2 \\ &\dots\dots \\ L_p &= \chi_p \leq x \end{aligned} \quad (21)$$

The coefficient set

$$\theta = [\gamma_0 \ \delta_0 \ \alpha_1 \ \beta_1 \ \eta_1 \ \delta_1 \ \dots \ \alpha_i \ \beta_i \ \eta_i \ \delta_i \ \dots \ \eta_p \ \delta_p]^T$$

can be written in terms of the  $\gamma$  ordinates of the nodes using constraint equations achieving continuity of value, rate and curvature at each node except the central node, which is constrained by value and with zero curvature. E.g., for  $p = 5$  the set of coefficients for negative  $x$  (nodes 1–3) is

$$\begin{aligned} \theta_{(-)} &= [\eta_0 \ \delta_0 \ \alpha_1 \ \beta_1 \ \eta_1 \ \delta_1 \ \alpha_2 \ \beta_2 \ \eta_2 \ \delta_2]^T \\ &= U^{-1} \gamma_{(-)} \end{aligned}$$

with

$$U = \begin{bmatrix} \chi_1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ -1 & 0 & 3\chi_1^2 & 2\chi_1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 3\chi_1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ \hline & & \chi_1^3 & \chi_1^2 & \chi_1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ \underline{0} & & \chi_2^3 & \chi_2^2 & \chi_2 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ & & -3\chi_2^2 & -2\chi_2 & -1 & 0 & 3\chi_2^2 & 2\chi_2 & 1 & 0 & 0 & 0 \\ & & -3\chi_2 & -1 & 0 & 0 & 3\chi_2 & 1 & 0 & 0 & 0 & 0 \\ \hline & & & & & & \chi_3^3 & \chi_3^2 & \chi_3 & 1 & & \\ \underline{0} & & \underline{0} & & & & \chi_3^3 & \chi_3^2 & \chi_3 & 1 & & \\ & & & & & & 2\chi_3 & 1 & 0 & 0 & & \end{bmatrix}, \gamma_{(-)} = \begin{bmatrix} \gamma_1 \\ 0 \\ \gamma_1 \\ \gamma_2 \\ 0 \\ \gamma_2 \\ 0 \\ \gamma_3 \\ 0 \end{bmatrix} \quad (22)$$

Coefficients for positive  $x$  and corresponding sets for larger  $p$  are found by obvious repetition and expansion respectively of equations (22).

By first fixing a set of node  $\chi$  ordinates, expressing the model symbolically in the form of equation (6) with each component coefficient in the form of equation (20) and solving for the cubic coefficients using equation (22), a final

set of equations for  $\mathbf{f}$  and  $\mathbf{h}$  emerges, together with a set of logical operators  $L$ , for each  $x$  and  $u$ . The extended set of states  $\mathbf{z}$  comprises true states and  $\gamma$  ordinates of the nodes only for each coefficient. Similar to equation (7), here

$$\mathbf{z} = [x_1 \ \dots \ x_n \ \gamma_0(\sigma_1) \ \dots \ \gamma_p(\sigma_1) \ \gamma_0(\omega_1) \ \dots \ \gamma_p(\omega_1) \ \dots]^T$$

And the required Jacobians  $\mathbf{F}$  and  $\mathbf{H}$  are also easily symbolically computed from these. The most convenient way to create executable code for the whole IEKF is to define a template function with core equations (11) to (14) and write strings of code for  $L$ ,  $\mathbf{f}$ ,  $\mathbf{h}$ ,  $\mathbf{F}$  and  $\mathbf{H}$  into it when the symbolic functions are evaluated in a build process. The symbolic strings become large, but remain manageable within MATLAB up to  $p = 9$  nodes.

## 6 Nonlinear handling model identification

The splined IEKF is tested by identification of the handling response of a full vehicle model. The model comprises a rigid vehicle body which is free to move in all six degrees of freedom and is suspended above four massless tyre nodes on a flat road surface by quarter-car suspensions. Lateral and longitudinal tyre forces are created using a load dependent, combined-slip Pacejka tyre model which employs wheel-spin degrees of freedom along with first-order relaxation lags. In the tests conducted here, torques are applied to the (front wheel) driven wheels from a PI controller maintaining a constant forward speed of 20 m/s. The source model is highly, but smoothly nonlinear and comprises 23 states.

Three outputs, lateral acceleration  $a_y$ , yaw angular rate  $r$  and roll angle  $\phi$ , are recorded in response to input steer angle  $\delta$ . The identification input signal adopts the principles discussed earlier for equation (16), but adds a sinusoidal sweep over 30 seconds in order to cover the full range of nonlinearity in the output – in this case  $-7 \text{ m/s}^2 < a_y < 7 \text{ m/s}^2$ .

$$\begin{aligned} \delta &= N\left(0, \frac{5\pi}{180}\right) + \frac{0.1\pi}{180} + \frac{3\pi}{180} \sin\left(\frac{2\pi}{30} t\right) \\ &(\text{rad}, 0 - 25 \text{ Hz}) \end{aligned} \quad (23)$$

With this nonlinear structure, constraints in the  $\mathbf{B}$  or  $\mathbf{C}$  parameters cannot be applied as in Section 2, but it turns out these are no longer required; the identification is automatically self-constraining due to the setting of the  $x$  ordinate node positions,  $\chi_i$ . These are first distributed in the range  $-1 < \chi_i < 1$  for all true states and inputs, and since the input is now normalised, the node spacings relevant to  $\mathbf{B}$  and  $\mathbf{D}$  terms stay fixed. Those for  $\mathbf{A}$  and  $\mathbf{C}$  are simply redistributed (and the filter rebuilt) periodically, as the filter iterates, in order to cover the range that evolves in each modal state or state pair. With the output also normalised, the states self-regulate, also approximately in the range  $-1 < x < 1$ , provided the number of states has been chosen appropriately.

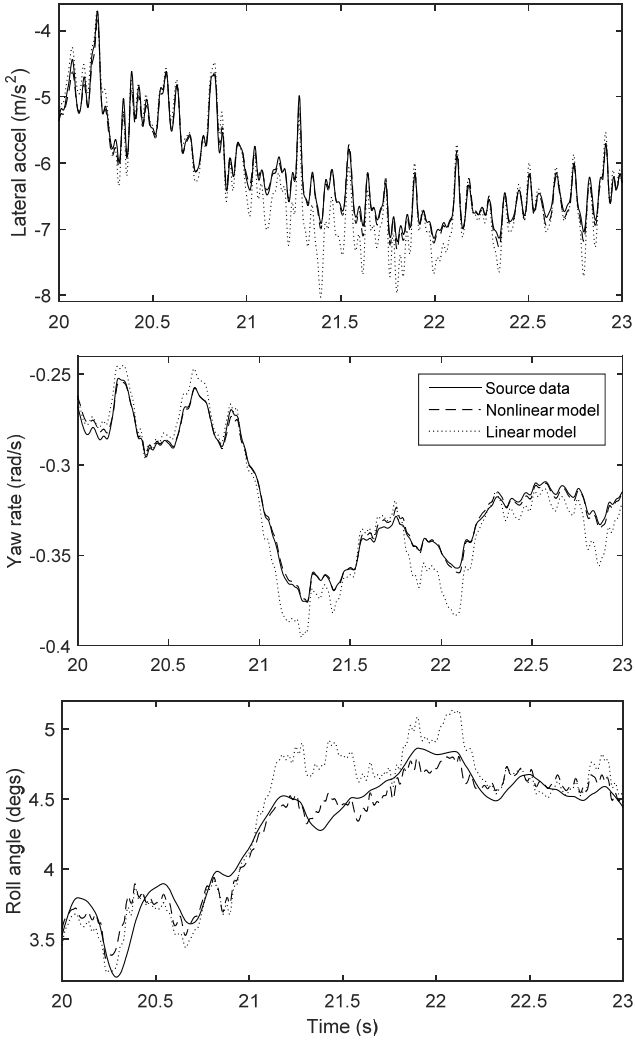


The filter is further modified to take full advantage of known error covariance information; although  $\mathbf{R} = \mathbf{I}$  achieves correct identification, we can improve the speed of convergence by substituting the known error covariance,

$$\mathbf{R} = \text{cov}(\mathbf{y} - \mathbf{h}) = \frac{1}{N} \sum_{k=1}^N (\mathbf{y}_k - \mathbf{h}_k)(\mathbf{y}_k - \mathbf{h}_k)^T$$

with  $\mathbf{h}$  re-evaluated at all  $N$  time points using a fixed parameter set (model), e.g., the model at time  $N$ , and with  $\mathbf{R}$  recomputed at each iteration of the ‘rinsing’ process.

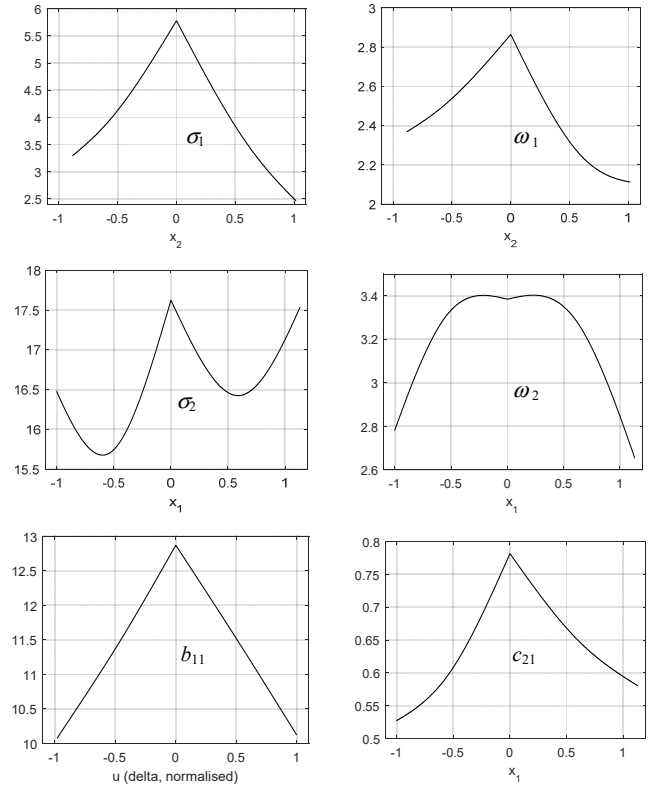
**Figure 7** Identified model fits to identification data



The nonlinear IEKF is initialised using values identified from a linear model identification on the same data; this linear model also provides a useful comparator to examine model performance. It is run with much lower  $\rho = 10^{-6}$ , to ensure that adaptation of  $\mathbf{P}_k$  and  $\mathbf{K}_k$  happens slowly, taking account of the whole 30 second data set; this is necessary since different parameters in the model are now active at different times in the data. A summary of results for increasing order choice is given in Table 5; the best-fit linear model eigenvalues are given, together with performance figures for the identification data and also for a validation test considering various magnitudes of step

response (see Figure 9 later). Tabled performance figures show the last percentage in the average  $R$  value across the three outputs, so 80 represents average  $R = 99.80\%$ .

**Figure 8** Splines for nonlinear variation of  $\lambda$  and two other typical splined coefficients



The best compromise of performance and conditioning occurs at the (highlighted) 4th order, where we can also see the linear eigenvalues reflecting an expected 0.4 Hz yaw-sideslip vehicle handling mode combined with a 1.5 Hz mode which correlates with expectations of the roll degree of freedom. Figure 7 shows the detailed dynamic fit comparison of linear vs. nonlinear models for a section of the response time history at high lateral acceleration and this more clearly illustrates the capability of the nonlinear model. Figure 8 shows the splines mapping the nonlinear eigenvalue parameters along with typical splines for components in the  $\mathbf{B}$  and  $\mathbf{C}$  matrices. Note particularly the change in the first mode which shows reducing frequency and also reducing damping as state magnitude increases; both are expected yaw-sideslip behaviours, which the fixed linear  $\lambda$  cannot achieve.

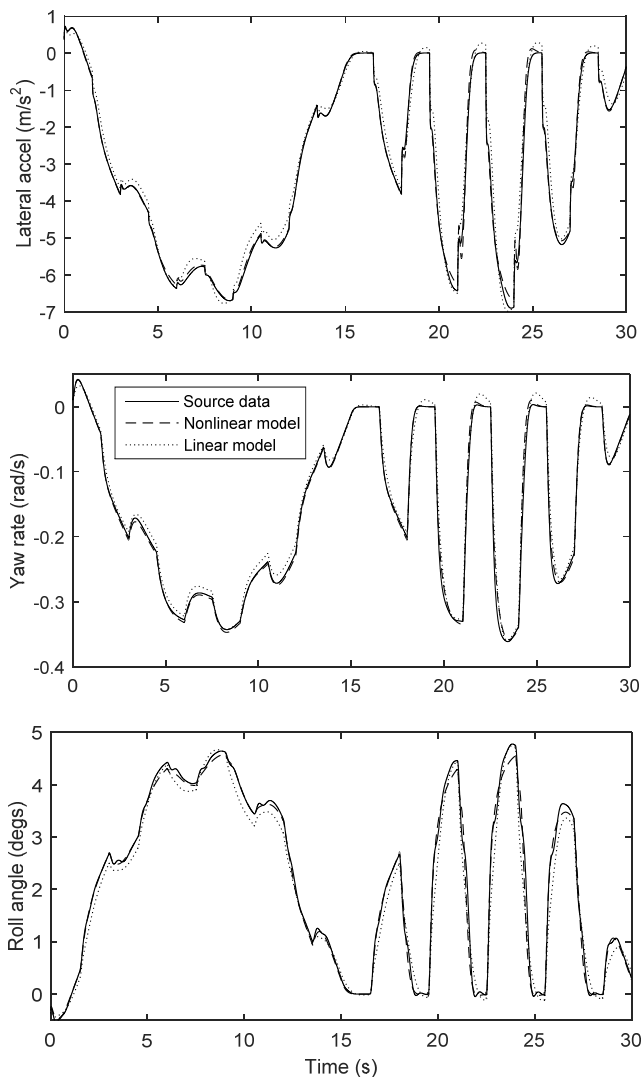
Finally, the validation test attempts to fully test performance using a different type of input; in the first half of this test, steps are added to a sweep across the range of nonlinearity and in the second half a range of higher magnitude steps is applied. Figure 9 shows how the nonlinear model matches the source model more accurately on this test, both dynamically and across magnitude.

**Table 5** Linear and nonlinear identified model performance

Linear $\lambda$ and linear and nonlinear model performance (average output R = 99.xx%)						
$n$	Linear model eigenvalues		Linear identn	Nonlinear identn	Linear validation	Nonlinear validation
2	$-2.3 \pm 1.7j$		60	73	29	56
3	$-9.3 \pm 2.5j$	$-3.3$	72	92	19	59
4	$-2.4 \pm 2.5j$	$-13.0 \pm 9.3j$	79	96	37	80
5	$-5.3 \pm 5.4j$	$-12.1 \pm 8.7j$	81	98	35	53
6	$-2.5 \pm 2.5j$	$-3.7 \pm 10.0j$	82	99	40	88

The results here are not claimed as the best possible reduced order nonlinear handling model; by constraining the nonlinearity to operate effectively as a variable linear model, nonlinear inter-relationships between states in the source model are not able to be replicated. The results do provide an effective illustration of the flexibility of extension of the IEKF method to a nonlinear structure though. Further research will consider how more general state combinations can be explored without predefined state definition or structure. All indications suggest that any smooth state nonlinearity can be identified using the IEKF.

**Figure 9** Identified model fits to validation data



## 7 Conclusions

A novel method for system identification and/or model order reduction has been successfully demonstrated on both linear and nonlinear multi input/output systems. The IEKF is capable of defining and self-ordering a state-space model from input/output data alone, without any prior knowledge of the system dynamics. The method provides rapid convergence of model parameters in a uniquely constrained modal canonical form and hence also provides information on the system dynamics in terms of the identified eigenvalues. The proposed method has just one easily tuned parameter, which governs speed of convergence without influencing accuracy in the identified model. Clear convergence criteria have been demonstrated, and these can also be used to determine the most appropriate model order for an accurate yet also well-conditioned reduced-order model.

Linear model tests show the technique is robust to high levels of noise, provided the noise is not correlated with the system dynamics. A nonlinear model structure using spline interpolation of coefficient variations with states and inputs has also been tested. This exhibits expected eigenvalue migration and delivers an accurate reduced order model of vehicle handling behaviour.

## Acknowledgements

This work was supported by Jaguar Land Rover and the UK-EPSC grant EP/K014102/1 as part of the jointly funded Programme for Simulation Innovation (PSi).

## References

Aguero, J.C., Rojas, C.R., Hjalmarsson, H. and Goodwin, G.C. (2012) ‘Accuracy of linear multiple-input multiple-output (MIMO) models obtained by maximum likelihood estimation’, *Automatica*, Vol. 48, No. 4, pp.632–637.

Ahmad, S.M. (2015) ‘Linear and nonlinear system identification techniques for modelling of a remotely operated underwater vehicle’, *International Journal of Modelling, Identification and Control*, Vol. 24, No. 1, pp.75–87.

Ahn, H-J., Lee, S-W., Lee, S-H. and Han, D-C. (2003) ‘Frequency domain control-relevant identification of MIMO AMB rigid rotor’, *Automatica*, Vol. 39, No. 2, pp.299–307.

Behzad, H., Shandiz, H., Toossian, N.A. and Abrishami, T. (2012) ‘Robot identification using fractional subspace method’, *Proceedings 2011 2nd International Conference on Control, Instrumentation and Automation*, pp.1193–1199.

- Best, M.C., Gordon, T.J. and Dixon, P.J. (2000) 'An extended adaptive Kalman filter for real-time state estimation of vehicle handling dynamics', *Vehicle System Dynamics*, Vol. 34, No. 1, pp.57–75.
- Best, M.C. (2007) 'Parametric identification of vehicle handling using an extended Kalman filter', *International Journal of Vehicle Autonomous Systems*, Vol. 5, Nos. 3/4, pp.256–273.
- Ding, F., Liu, Y. and Bao, B. (2012) 'Gradient-based and least-squares-based iterative estimation algorithms for multi-input multi-output systems', *Proceedings of the Institution of Mechanical Engineers, Part I: Journal of Systems and Control Engineering*, Vol. 226, No. 1, pp.43–55.
- Guo, Y., Guo, L.Z., Billings, S.A. and Wei, H-L. (2015) 'Identification of nonlinear systems with non-persistent excitation using an iterative forward orthogonal least squares regression algorithm', *International Journal of Modelling, Identification and Control*, Vol. 23, No. 1, pp.1–7.
- Hafayed, M., Abba, A. and Boukaf, S. (2016) 'On Zhou's maximum principle for near-optimal control of mean-field forward-backward stochastic systems with jumps and its applications', *International Journal of Modelling, Identification and Control*, Vol. 25, No. 1, pp.1–16.
- Hassani, V., Aguiar, A.P., Athans, M. and Pascoal, A.M. (2009) 'Multiple model adaptive estimation and model identification using a minimum energy criterion', *Proceedings American Control Conference*, pp.518–523.
- Kallapur, A., Samal, M., Puttige, V., Anavatti, S. and Garratt, M. (2008) 'A UKF-NN framework for system identification of small unmanned aerial vehicles', *Proceedings IEEE Conference on Intelligent Transportation Systems*, pp.1021–1026.
- Polifke, W. (2014) 'Black-box system identification for reduced order model construction', *Annals of Nuclear Energy*, Vol. 67, No. 1, pp.109–128.
- Soderstrom, T. and Stoica, P. (1989) *System Identification*, Prentice Hall International, Cambridge, UK.
- Wang, Z., Chen, S., Yu, Y. and Zhu, J. (2015) 'Frequency domain system identification of a micro unmanned aircraft', *International Journal of Modelling, Identification and Control*, Vol. 24, No. 3, pp.187–195.