On the Interpretation and Identification of Dynamic Takagi-Sugeno Fuzzy Models

T. A. Johansen[*] and R. Shorten[†] and R. Murray-Smith[‡]

September 30, 1999

Abstract

Dynamic Takagi-Sugeno fuzzy models are not always easy to interpret, in particular when they are identified from experimental data. Ideally, it is desirable that a dynamic Takagi-Sugeno fuzzy model should give accurate global nonlinear prediction, and at the same time that its local models are close approximations to the local linearizations of the nonlinear dynamic system. The latter is important in many applications where the constituent local models are used individually, and aids considerable validation and interpretation of the model. This defines a multi-objective identification problem; namely, the construction of a dynamic model that is a good approximation of both local and global dynamics of the underlying system. While these objectives are often conflicting, it is shown that there exists a close relationship between dynamic Takagi-Sugeno fuzzy models and dynamic linearization when using affine local model structures, which suggests that a solution to the multi-objective identification problem exists. However, it is also shown that the affine local model structure is a highly sensitive parameterization when applied in transient operating regimes, i.e. far away from equilibrium. The reason is essentially that the constant term in the affine local model tend to dominate over the linear term during transients. In addition, it is inherently more difficult to design informative experiments in transient regions compared to near-equilibrium regions. Due to the multi-objective nature of the identification problem studied here, special considerations must be made during model structure selection, experiment design and identification in order to meet both objectives. Some guidelines for experiment design are suggested and some robust nonlinear

[*]Norwegian University of Science and Technology, Department of Engineering Cybernetics, N-7491 Trondheim, Norway.

[†]Department of Computer Science, NUI, Maynooth, Co. Kildare, Ireland

[‡]Department of Computing Science, Glasgow University, Glasgow G12 8QQ, Scotland.

identifcation algorithms are studied. These include constrained and regularized identification and locally weighted identification. Their usefulness in the present context is illustrated by examples.

1 Introduction

Consider the nonlinear dynamic system

$$\dot{x} = f(x,u) \tag{1}$$

A Takagi-Sugeno fuzzy model will approximate this system by smoothly interpolating affine local models (Takagi and Sugeno 1985). Each local model contributes to the global model in a fuzzy subset of the space where (x,u) belongs. This fuzzy set is characterized by a membership function μ_i, and leads to the Takagi-Sugeno fuzzy model

$$\dot{x} = \sum_{i=1}^{N} (A_i x + B_i u + d_i) w_i(x,u) \tag{2}$$

where the weighting functions are given by the fuzzy inference

$$w_i(x,u) = \frac{\mu_i(x,u)}{\sum_{j=1}^{N} \mu_j(x,u)} \tag{3}$$

This equation follows by using center-of-gravity defuzzification, and it has been assumed that $\sum_j \mu_j(x,u) > 0$ for all (x,u).

The Takagi-Sugeno fuzzy model has recently found wide applicability in fuzzy model based control, e.g. (Driankov and Palm 1998, Passino and Yurkovich 1998, Johansen 1994, Wang *et al.* 1996, Wang 1993). The reason for this is that its consequence part is an affine dynamic model rather than a fuzzy set or constant value, which has several advantages:

- From a control engineering perspective the use of local affine (or affine linear) models bridges the gap between fuzzy control and conventional control. Many existing tools and theories in linear systems theory can be partially applied to Takagi-Sugeno fuzzy models and controllers. The relationships to gain scheduling, e.g. (Driankov *et al.* 1996, Hunt and Johansen 1997), and piecewise linear systems, are evident.
- The relatively complex consequence part allows the number of fuzzy rules (local models) to be quite small in many applications. Consequently, the Takagi-Sugeno fuzzy model is less prone to the curse of dimensionality than other fuzzy models.
- The model structure (partitioning of the state-space and local model structure) and local model properties can, in some applications, be easily related to the physics of the system. This simplifies model development and validation.

One typically attempts to select the local state space parameters (A_i, B_i), the constant term d_i, and the membership functions such that the fuzzy model (2) is a good approximation to the nonlinear dynamic system (1) in some appropriate sense. For many applications it is important that the global behaviour of the nonlinear model (2) is similar to the global behaviour of the nonlinear system (1). For example, this is typically the case when the global model is used for nonlinear prediction or when the global model is used as an internal model in a controller as in e.g. (Wang 1993, Johansen 1994). On the other hand, it is sometimes required and often desireable that the local linear models in (2) are accurate approximations to the local linearizations of (1). This is the case when the dynamic Takagi-Sugeno fuzzy model is used as a basis for a fuzzy gain-scheduled controller since the local linear models are used to design local linear controllers (Driankov *et al.* 1996, Hunt and Johansen 1997). This property is also very useful when validating and analysing the dynamic model (2), (Shorten *et al.* 1999).

It has widely been observed that it is often non-trivial to identify dynamic local models which are close approximations to linearizations of the nonlinear system (Shorten *et al.* 1999, Murray-Smith and Johansen 1997, Johansen *et al.* 1998*a*, Yen *et al.* 1998). It has been argued that this problem is a consequence of excessive degrees of freedom in the affine local model structure when it is applied in transient operating regimes (Shorten *et al.* 1999). Another important reason is that the algorithm of choice for many practicioners, the least squares algorithm, is often employed with the explicit objective of selecting the local model parameters in order to optimize global prediction performance. This is often achieved with local models that are significantly different from the local linearizations (Murray-Smith and Johansen 1997, Yen *et al.* 1998). The problems are, in most practical applications, amplified by constraints on the experiment design which restrict the amount of informatation in the transient data, as exemplified in (Johansen *et al.* 1998*a*). A consequence is that one may quite easily determine a Takagi-Sugeno fuzzy model which provides a good global nonlinear model of the nonlinear system, but with local models that have little in common with local linearizations. Furthermore, it is in general more difficult to understand the meaning and properties of the off-equilibrium local models than equilibrium local models in the Takagi-Sugeno fuzzy model (Shorten *et al.* 1999). Most of the above mentioned problems are unique to dynamic modelling and identification problems (as opposted to modelling and identification of static mappings), although the tradeoff between local and global approximation accuracy also appears in static modelling problems considered in (Murray-Smith and Johansen 1997, Yen *et al.* 1998). The present paper explicitly focuses on dynamic systems, which are considerably more challenging than static function approximation.

The present paper continues the work of Shorten *et al.* (1999). The main message in (Shorten *et al.*

1999) is that local affine dynamic models in Takagi-Sugeno fuzzy dynamic models contain excessive degrees of freedom and must be interpreted carefully. Here we provide a theoretical foundation for understanding how identified Takagi-Sugeno fuzzy models might be interpreted by relating them to dynamic linearization about trajectories. Furthermore, the interpretation and identification problems are illustrated by simple, yet practically relevant and illustrative examples. Finally, guidelines for experiment design and robust system identification are given in order to improve the interpretability and accuracy of the identified fuzzy model. The suggested system identification methods are based on well known ideas originally developed for the purpose of robust nonlinear system identification without having the particular problems related to transient operating regimes or interpretability in mind (Murray-Smith and Johansen 1997, Johansen 1996, Johansen 1997*b*).

The aim of the present paper is to present an *overview* of some problems and solutions related to identification and interpretation of dynamic Takagi-Sugeno fuzzy models. The outline of this paper is as follows: Dynamic linearization of a general nonlinear system is considered in section 2. In section 3 it is shown that a dynamic Takagi-Sugeno fuzzy model can form an approximation to dynamic linearization about arbitrary trajectories. In section 4 it is illustrated that the Takagi-Sugeno fuzzy model structure, with local affine models, contains excessive degrees of freedom when transient dynamics are considered. The different properties of selecting local constant, linear and affine models are discussed in section 5, and the tradeoffs related to identifiability and interpretability are explored. Furthermore, the effect of minimizing the number of premise variables and the role of the membership functions on the interpretability are outlined in sections 6 and 7, respectively. Identification is discussed in section 8, with particular emphasis given to the distinction between equilibrium and transient local models. Experiment design guidelines, and modifications of standard identification algorithms, are suggested to improve the robustness of the identification procedure. A simplified simulation example considering dynamic modelling of the nonlinear longitudinal dynamics of a vehicle is used to illustrate the ideas throughout the paper. Some additional aspects are illustrated by a pendulum dynamics modelling problem. Conclusions are given in section 9.

2 Dynamic linearization

The dynamic Takagi-Sugeno fuzzy model (2) is composed of multiple local affine dynamic models. It would be desirable for the purpose of interpretation, analysis and application that these local affine models can be related to linearizations of the nonlinear system. In order to gain some understanding of the relationship between the local affine models and linearizations we continue with a brief review of dynamic linearization about trajectories of nonlinear systems of the quite general form (1), where f is assumed to be smooth.

At an equilibrium point (x_0', u_0') (an equilibrium point satisfies $\dot{x} = 0$, i.e. $f(x_0', u_0') = 0$) the local linearization of (1) is given by

$$\dot{x} \;=\; \frac{\partial f}{\partial x}(x_0', u_0')(x - x_0') + \frac{\partial f}{\partial u}(x_0', u_0')(u - u_0') + \text{h.o.t.} \tag{4}$$

by neglecting the higher-order terms (denote h.o.t.) in the Taylor-series expansion (4). In deviation coordinates $\Delta x = x - x_0', \Delta u = u - u_0'$, eq. (4) becomes

$$\dot{\Delta x} \;=\; A(x_0', u_0')\Delta x + B(x_0', u_0')\Delta u + \text{h.o.t.} \tag{5}$$

where we have defined

$$\begin{aligned} A(x_0', u_0') &= \frac{\partial f}{\partial x}(x_0', u_0') \\ B(x_0', u_0') &= \frac{\partial f}{\partial u}(x_0', u_0') \end{aligned}$$

This local linearization describes the *linearized dynamics* of the nonlinear system subject to small perturbations near the equilibrium point (x_0', u_0').

Next, consider the more general case when the linearization of (1) is made about an arbitrary point (x_0, u_0) that need not be an equilbrium point. This corresponds to dynamic linearization about a point on a trajectory. The trajectory $(x_0(t), u_0(t))$ is defined by an initial state $x_0(0)$ and satisfies $\dot{x}_0 = f(x_0, u_0)$. The dynamic linearization of (1) about the (time-varying) point (x_0, u_0) on some arbitrary trajectory is given by

$$\dot{x} \;=\; f(x_0, u_0) + A(x_0, u_0)(x - x_0) + B(x_0, u_0)(u - u_0) + \text{h.o.t.} \tag{6}$$

Introducing deviation coordinates $\Delta x = x - x_0, \Delta u = u - u_0$ we get the *linearized dynamics*

$$\dot{\Delta x} \;=\; A(x_0, u_0)\Delta x + B(x_0, u_0)\Delta u + \text{h.o.t} \tag{7}$$

that describes the response to small perturbations about a point (x_0, u_0) on the nominal trajectory $(x_0(t), u_0(t))$. In addition, the nominal trajectory itself (*trend*) is locally approximated by the equation

$$\dot{x} \;=\; f(x_0, u_0) + \text{h.o.t.} \tag{8}$$

which approximates the flow of the state by a constant vector near the point (x_0, u_0) on the nominal trajectory $(x_0(t), u_0(t))$.

3 The Takagi-Sugeno fuzzy models as an approximation to dynamic linearization of nonlinear systems

The ability of the Takagi-Sugeno fuzzy model to approximate arbitrary smooth static functions on compact domains to arbitrary accuracy are well known. Local constant models was considered in (Wang and Mendel 1992), while the theory was developed for local polynomial models (including affine) in (Johansen and Foss 1993). Finally, in (Rovatti 1996) it was shown that the parameters of local affine model can be selected to guarantee that the derivative of the right-hand side of the Takagi-Sugeno fuzzy model (2) could be an arbitrarily good approximation to the derivative of the nonlinear system function f in (1). This property is closely related to the ability of local affine models to represent the small-signal dynamics. However, notice that the approximation result of (Rovatti 1996) concerns the accuracy of the linearization of the nonlinear Takagi-Sugeno fuzzy model rather than the accuracy of its constituent local affine models.

It was shown in (Johansen *et al.* 1998*b*) that the finite set of linearizations about a finite number of points (equilibria and transient points) can be used to accurately approximate dynamic linearization about arbitrary trajectories using an interpolated multiple model structure with local affine dynamic models. The result of (Johansen *et al.* 1998*b*) considers autonomous systems, and an extension to systems with exogeneous inputs is given below.

Assume a continuous trajectory (x_0, u_0) that satisfies $(x_0(t), u_0(t)) \in X \times U$ is given, $X \subset R^n$ is a compact subset of the state-space, $U \subset R^r$ is a compact subset of the input-space, and $\dot{x}_0(t) = f(x_0(t), u_0(t))$, $x_0(0) = \overline{x}$. The dynamic linearization about the trajectory (x_0, u_0) is now the linear time-varying (LTV) system (following from (6) by neglecting the higher-order dynamics)

$$\dot{\xi}_1(t) = f(x_0(t), u_0(t)) + A(x_0(t), u_0(t))(\xi_1(t) - x_0(t)) + B(x_0(t), u_0(t))(u(t) - u_0(t)) \quad (9)$$

$$\xi_1(0) = \overline{x} \quad (10)$$

An approximation to the dynamic linearisation is given by the Takagi-Sugeno fuzzy system scheduled on the state/input, cf. (2):

$$\dot{\xi}_2(t) = \sum_{i=1}^{N} (f_i(x_i, u_i) + A(x_i, u_i)(\xi_2(t) - x_i) + B(x_i, u_i)(u(t) - u_i)) w_i(\xi_2(t), u(t)) \quad (11)$$

$$\xi_2(0) = \overline{x} \quad (12)$$

We will continue with developing an approximation result that shows the closeness between (9) and (11). Notice that a similar approximation result can be derived when scheduling on the nominal trajectory:

$$\dot{\xi}_3(t) = \sum_{i=1}^{N} (f_i(x_i, u_i) + A(x_i, u_i)(\xi_3(t) - x_i) + B(x_i, u_i)(u(t) - u_i)) w_i(x_0(t), u_0(t)) \quad (13)$$

$$\xi_3(0) = \overline{x} \tag{14}$$

Scheduling on a nominal (reference) trajectory is applied in e.g. (Driankov *et al.* 1996) in the context of fuzzy gain scheduled control.

Let $Z^0 = \{(x_1, u_1), (x_2, u_2), ..., (x_N, u_N)\} \subset Z = X \times U$ be the set of linearization points and define deviations from the nominal trajectory (x_0, u_0) by $\Delta\xi_1(t) = \xi_1(t) - x_0(t)$, $\Delta\xi_2(t) = \xi_2(t) - x_0(t)$, and $\Delta u(t) = u(t) - u_0(t)$. Intuitively, one would expect that the system (11) approximates (9) when "Z^0 covers Z densely", for example in the sense that the maximum distance between neighbouring elements of Z^0 is sufficiently small:

$$\delta = \max_{i \in \{1,2,\ldots,N\}} \sup_{z \in \mathrm{supp}(w_i)} ||z - z_i||_2 \tag{15}$$

where $\mathrm{supp}(w_i)$ denotes the support of the function w_i, i.e. $\mathrm{supp}(w_i) = \{z \in Z \mid w_i(z) > 0\}$. Thus, notice that $w_i(z) = 0$ if $||z - z_i||_2 > \delta$. This is formalised in the following theorem:

Theorem 1. *Assume $X \times U$ is a compact set, f is smooth with bounded Lipschitz constant L, and (x_0, u_0) is a given trajectory of (1) in $X \times U$. Then*

$$\dot{\Delta\xi}_1(t) = A(x_0(t), u_0(t))\Delta\xi_1(t) + B(x_0(t), u_0(t))\Delta u(t) \tag{16}$$

$$\dot{\Delta\xi}_2(t) = A(x_0(t), u_0(t))\Delta\xi_2(t) + B(x_0(t), u_0(t))\Delta u(t) + \varepsilon(\xi_2(t), u(t), x_0(t), u_0(t)) \tag{17}$$

and ε satisfies

$$\sup_{(\xi_2, u), (x_0, u_0) \in X \times U} ||\varepsilon(\xi_2, u, x_0, u_0)||_2 \le 2L\delta + O(\delta^2) \tag{18}$$

where $O(\epsilon)$ means that the limit of $O(\epsilon)/\epsilon$ exists (is finite) when $\epsilon \to 0$.

□

The proof is given in the Appendix.

The above theorem shows that the dynamic Takagi-Sugeno fuzzy model where the local affine dynamic models are local linearizations (including off-equilibrium linearisations) leads to an arbitrary close approximation to the LTV dynamic system resulting from dynamic linearisation about the trajectory, because when the number of rules $N \to \infty$ (and the local models are sensibly located in $X \times U$ and with sensible membership functions with compact support), then $\delta \to 0$. The result also suggests that the local affine time-invariant dynamic models in the Takagi-Sugeno fuzzy model can be interpreted as local approximations to the frozen time local affine time-varying models of the dynamic linearization about any trajectory passing through the respective regions of validity of the Takagi-Sugeno local affine models.

Off-equilibrium local models have by definition no equilibrium point within their region of validity. Such local models still have equilibrium points, but the local models are not valid at their equilibrium points. Such equilibrium points are called "virtual equilibrium points" in (Shorten *et al.* 1999) because they need not have anything to do with the equilbrium points of the nonlinear system. Consequently, the most common linear system analysis tools are not directly suitable for off-equilibrium local affine models since they focus on characterizing the dynamic behaviour in the neighbourhood of equilibria. For example, the eigenvalues of the A_i matrix will contain information about the stability of the virtual eqilibrium point. But this is of little interest since the virtual equilibrium point has no meaning for the underlying nonlinear system. We therefore conclude that the eigenvalues of off-equilibrium local models need not contain any information about local stability of equilibrium points of the nonlinear system.

Stability of equilibrium points is a special case of stability of trajectories (the special case when the trajectory stays at equilibrium at all time).[1] Stability of a trajectory essentially means that a trajectory that can be considered as a small perturbation from the nominal trajectory will stay close to the nominal trajectory (or approach the nominal trajectory as time goes to infinity in the case of asymptotic stability). In the local model context, the notion of stability means that if a perturbed trajectory will stay in the region of validity of the local model for a long time, then the perturbed trajectories will approach the nominal trajectory. However, no trajectory will stay in a transient region for a long time, so such a local analysis of stability of trajectories need not contain any information about the system. Even if the trajectory moves through a sequence of operating regions with different (but stable) local models, it is still not possible to argue anything about stability of trajectories, cf. the theory on stability of time-varying linear systems, e.g. (Kailath 1980). Stronger conditions are necessary, see for example (Lohmiller and Slotine 1998) for some contraction conditions that ensure that neighbouring trajectories of nonlinear systems are attracted towards each other. Also, if a local off-equilibrium model has unstable modes, a perturbed trajectory passing through its region of validity may still approach the nominal trajectory during some part of the transient. Notice that in transient regions with local models $\dot{x} = A_i(x - x_i) + B_i(u - u_i) + d_i$, the constant term d_i may dominate the system behaviour, and both the $A_i(x-x_i)$ and $B_i(u-u_i)$ terms may be of secondary importance. Considering A_i and B_i alone should therefore not be expected to give much information. To conclude, it is difficult to make any generally valid statements about the qualitative behaviour of the system from off-equilibrium affine local model analysis.

Example: Unforced pendulum, dynamic lineariziation.

To illustrate the effects in a simple system we consider an unforced pendulum of unit mass, on a

[1] The concept of stability of trajectories and paths, denoted Lyapunov stability and Poincare stability respectively, is treated in (Jordan and Smith 1987).

rigid string of length l, with an angle θ and angular velocity $\dot{\theta} = \omega$, leading to state-equations,

$$\begin{aligned} \dot{\theta} &= \omega && (19) \\ \dot{\omega} &= -\frac{g}{l}\sin(\theta) - \frac{d}{l}\omega, && (20) \end{aligned}$$

which are simple yet provide some features of interest for this investigation. Here $l = 1$, the friction coefficient $d = 0.5$, and the gravitational constant $g = 9.81$.

Linearization of the nonlinear equations (19)-(20) about an arbitrary point (θ^o, ω^o) leads to the following linearized system

$$\begin{aligned} \dot{\theta} &= \omega^o + (\omega - \omega^o) && (21) \\ \dot{\omega} &= -\frac{g}{l}\sin(\theta^o) - \frac{d}{l}\omega^o - \frac{g}{l}\cos(\theta^o)(\theta - \theta^o) - \frac{d}{l}(\omega - \omega^o) + \text{h.o.t.} && (22) \end{aligned}$$

which can be simplified and written in the matrix form

$$\begin{pmatrix} \dot{\theta} \\ \dot{\omega} \end{pmatrix} = d(\theta^o) + A(\theta^o)\begin{pmatrix} \theta \\ \omega \end{pmatrix} + \text{h.o.t.} \qquad (23)$$

where

$$\begin{aligned} d(\theta^o) &= \begin{pmatrix} 0 \\ 9.81(-\sin(\theta^o) + \cos(\theta)^o\theta^o) \end{pmatrix} && (24) \\ A(\theta^o) &= \begin{pmatrix} 0 & 1 \\ -9.81\cos(\theta^o) & -0.5 \end{pmatrix} && (25) \end{aligned}$$

The flow of the nonlinear pendulum system is illustrated in Figure 1 (the flow is the velocity of the state-vector at each state, i.e. the vector field f defined by the r.h.s. of the differential equation for the system (1)). Three different trajectories for the system are shown in the figure, starting from somewhat different initial conditions. We observe that two of these trajectories are attracted towards each other and a common equilibrium point as time increases, while the third trajectory diverges and ends up in a different equilibrium point (the pendulum starts rotating the opposite way due to lower initial kinetic energy). In the part of the illustrated state space where $\pi/2 < \theta < 3\pi/2$, the local linearizations has an $A(\theta^o)$ matrix with at least one positive eigenvalue. In the context of dynamic linearization, this means that trajectories with slightly different intial conditions may diverge. Considering the trajectory starting at $(4, -3)$ as the nominal trajectory, we see that the other two trajectories both diverge from this one initially. However, while the trajectory starting at $(4, -2)$ continues to diverge from the nominal trajectory, the trajectory starting at $(4, -4)$ eventually converges towards the nominal trajectory and later moves into the region where $-\pi/2 < \theta < \pi/2$. Thus, knowledge of the eigenvalues of $A(x, u)$ does not allow us to make any prediction on the qualitative behaviour of the system in transient operating regimes.

□

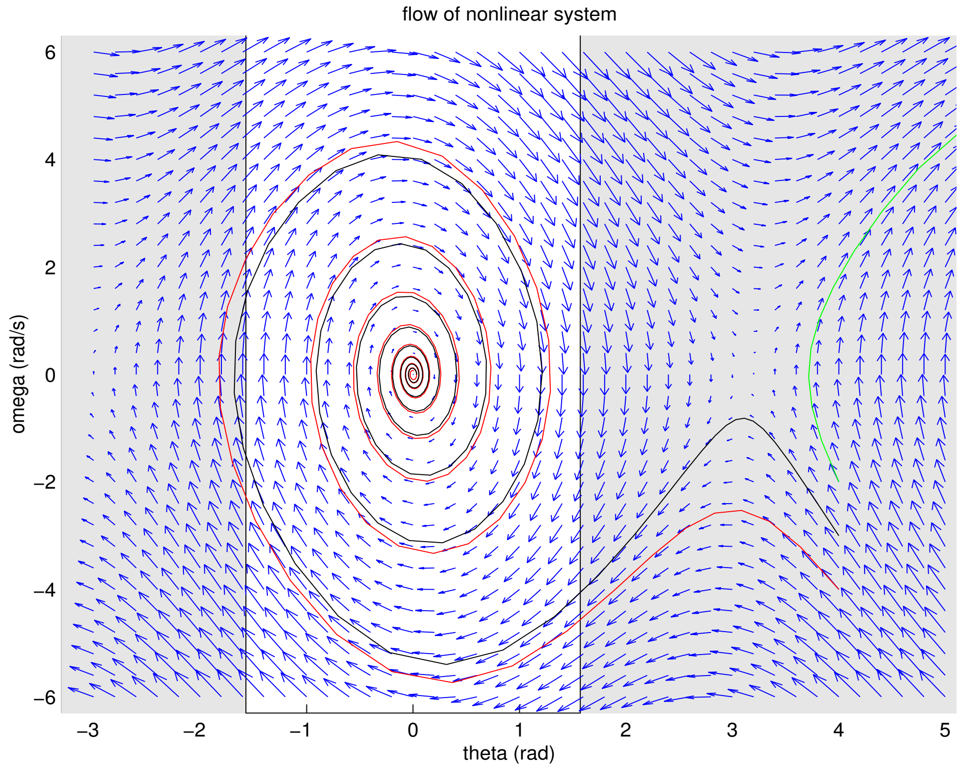

Figure 1: The flow of the nonlinear pendulum system together with 3 trajectories starting a somewhat different initial conditions.

4 Identifiability of the Takagi-Sugeno fuzzy model

In section 3 we argued that under general conditions there exist Takagi-Sugeno model parameters such that,

- the global fuzzy model (2) accurately approximates the global behaviour of the nonlinear system (1), and at the same time,
- the local affine models of the Takagi-Sugeno fuzzy model admit valid interpretation as local linearizations of the nonlinear system about points on trajectories as described in section 2.

In other words, the Takagi-Sugeno model might be both accurate and with a useful interpretation of the local models as local linearizations. However, in this section we argue that the local affine model structure applied in transient operating regimes (where no equilibrium exist) contains excessive degrees of freedom and may be poorly identifiable in the sense that large perturbations of some combinations of affine local model parameters may only have a small effect on the identification criterion.[2] This has serious consequences in particular for the interpretation, applicability and

[2]Lack of identifiability is characterized by non-uniqueness of the model structure, i.e. two different parameter vectors yields the same input/output behaviour. Poor identifiability is (informally) defined as a somewhat relaxed property, namely that two significantly different parameter vectors give very similar input/output behaviour.

accuracy of the individual local affine models when identified from experimental data, but also for the accuracy of the global nonlinear fuzzy model in some cases. [3] Furthermore, the poor identifiability of the local models gives rise to an ill-conditioned system identification problem which we shall discuss in depth in section 8.

Suppose we seek local affine models of the form

$$\dot{x} \quad = \quad A_i(x - x_i) + B_i(u - u_i) + d_i \tag{26}$$

to be approximately valid in a small neighbourhood of a point (x_i, u_i). The structure (26) sometimes contain excessive degrees of freedom when the point (x_i, u_i) is far away from equilibrium. The reason for this is that in many cases the constant term d_i (trend) will dominate (26), while varying some elements of A_i and B_i may only have a minor effect on the local model accuracy. In order to motivate this claim, suppose

$$d_i \quad = f(x_i, u_i)$$

which is an exact model of (1) at the point (x_i, u_i) and a reasonable approximation in a small neighbourhood of this point especially if it is far from equilibrium. The additional degrees of freedom available in the parameters A_i and B_i can be used in different ways:

- A_i and B_i can be selected to accurately represent the linearized dynamics, i.e. $A_i \approx A(x_i, u_i)$ and $B_i \approx B(x_i, u_i)$. This is advantageous in terms of interpretation, analysis and applicability of the model in control systems design, but has the disadvantage that it may lead to a smaller region of validity of the local model. Furthermore, the problem of estimating the poorly identifiable parameters $A(x_i, u_i)$ and $B(x_i, u_i)$ remains difficult.

- A_i and B_i can be selected to increase the region of validity of the local affine approximation (26) and/or to improve the accuracy of the global model (2). In this case these parameters may be completely different from the true linearization parameters $A(x_i, u_i)$ and $B(x_i, u_i)$ and serves only the purpose of providing a richer class of function approximators. Consequently, the local affine model cannot always be interpreted in terms of a local linearization in this case. It is well known that $A(x_i, u_i)$ and $B(x_i, u_i)$ are often sub-optimal choices for A_i and B_i when only considering global approximation accuracy (Murray-Smith and Johansen 1997), see also (Babuska and Verbruggen 1997)

Poor identifiability is a problem closely linked with *off-equilibrium* affine local models. The reason for this is simply that at equilibrium the constant trend term d_i vanishes, and the dynamics must

[3] The reason why we may insist on using locally affine models even though they are poorly identifiable is that the interpretation in terms of linearizations is useful in terms of system analysis and local control design for example in gain-scheduled control.

be fully captured by the (A_i, B_i)-parameters. Thus, near equilibrium the dynamics are captured by the (A_i, B_i) parameters and there is no problem.

Example: Unforced pendulum, poor identifiability.

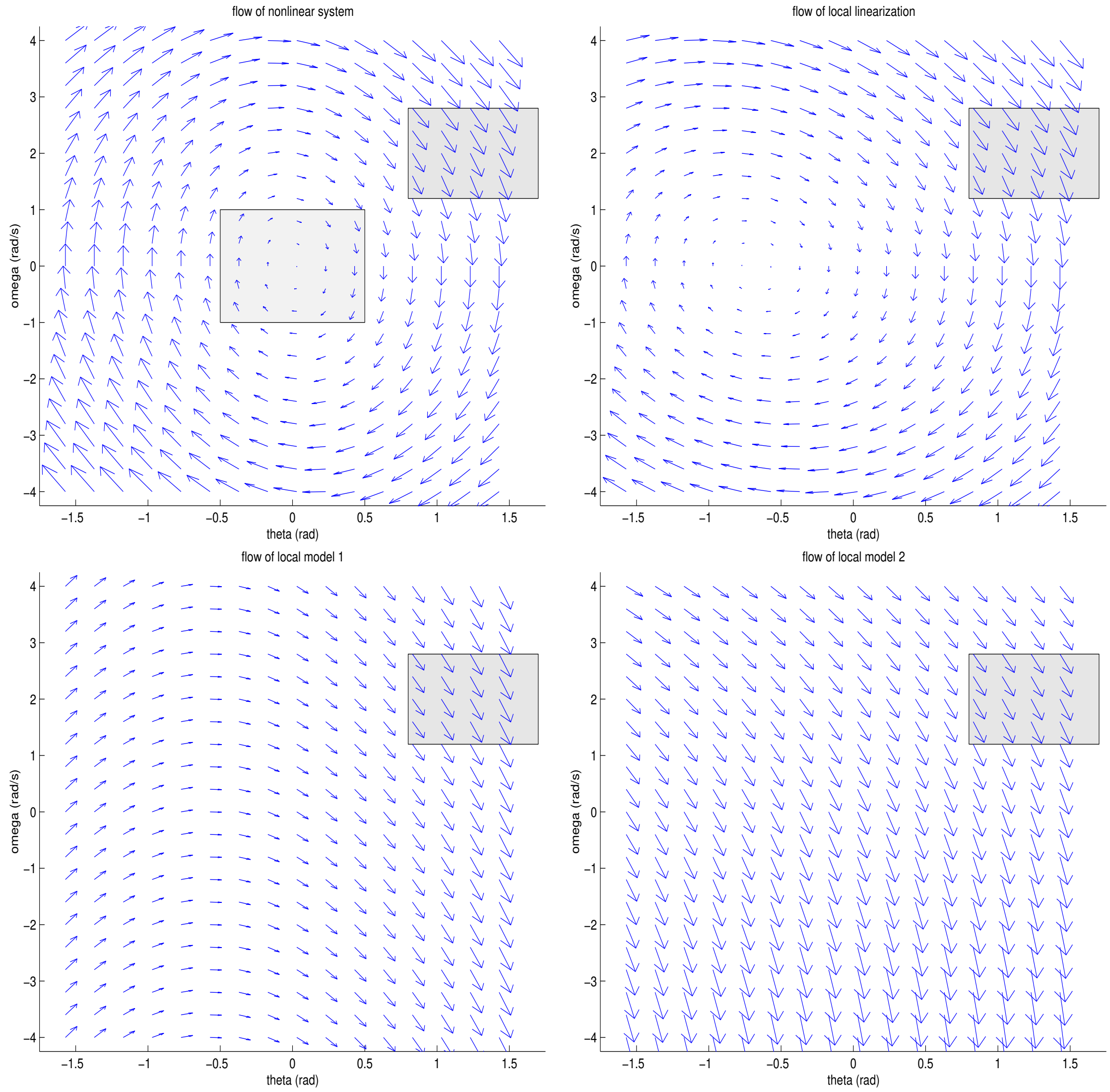

Figure 2: The flow of the nonlinear pendulum system, local linearization and two accurate local affine models.

Consider the point $(\theta, \omega) = (\pi/3, 2)$, which is a transient state of the autonomous system. Figure 2 shows the flow of the nonlinear system (upper left), the flow of the local linearization about $(\pi/3, 2)$ (upper right), and the flow of two alternative local affine models (lower) that were selected manually by trial and error to match the local dynamics near $(\pi/3, 2)$. We observe that the local linearization and both the two local affine models are quite accurate models of the nonlinear system in their assumed region of validity (shaded region in the figure) with very similar flow field, while outside this region they are all more or less invalid models and their flow fields differ considerably.

Hence, when applied as local models in a dynamic Takagi-Sugeno fuzzy model, they would all lead to more or less equally accurate global approximations to the nonlinear dynamic system.

The numerical values of the local linearization and the two local affine models are

$$A(\pi/3,2) = \begin{pmatrix} 0.00 & 1.00 \\ -0.40 & -0.50 \end{pmatrix} \qquad d(\pi/3,2) = \begin{pmatrix} 0.0 \\ -3.36 \end{pmatrix} \tag{27}$$

$$A_1 = \begin{pmatrix} -0.20 & 0.20 \\ -1.00 & 1.00 \end{pmatrix} \qquad d_1 = \begin{pmatrix} 1.75 \\ -10.00 \end{pmatrix} \tag{28}$$

$$A_2 = \begin{pmatrix} 0.00 & 0.00 \\ -1.00 & 0.00 \end{pmatrix} \qquad d_2 = \begin{pmatrix} 2.00 \\ -3.00 \end{pmatrix} \tag{29}$$

respectively. The eigenvalues of $A(\pi/3,2)$ are $-0.25 \pm 0.58j$, which corresponds to a stable underdamped linear system. The eigenvalues of A_1 are 0 and 0.8, which corresponds to an unstable linear system. Both the eigenvalues of A_2 are 0, which also corresponds to an unstable linear system. The example illustrates that local affine models with very different parameters and structural properties can lead to dynamic Takagi-Sugeno fuzzy models with very similar global properties at this transient point. This is made possible by excessive degrees of freedom in the affine local model.

Alternatively, consider the equilibrium point $x = 0$. At this point the drift term d_i must be close to zero in order for the local model to have an equilibrium point close to zero. There are therefore no additional degrees of freedom available since in order for the local affine model to be accurate, the parameters in A_i must necessarily be similar to the parameters of the local linearization.

□

More examples of non-unique off-equilibrium local models can be found in (Shorten *et al.* 1999).

We believe that the sensitive parameterisation of off-equilibrium dynamics has profound consequences for both control system design, and the simulation of non-linear systems. We note that in off-equilibrium regions, the objective of control design, is to manipulate the direction and speed of the local system (i.e. the direction and size of the arrows representing the flow vector field in Figure 2), so that the local system approximates some desired vector field. Many linear control design techniques, such as conventional pole placement, are concerned only with the speed of the modes of the system in the neighbourhood of equilibria. The direction of the closed loop vector field is considered to be of secondary importance, and is determined as consequence of the desired pole locations. To make matters worse, the qualitative interpretation of the A_i and B_i matrices (e.g. stability), which is often used as a design parameter in linear control, is meaningless without the d_i vector, and the size of the region of validity. Hence, some linear control techniques do not necessarily provide a valid basis for off-equilibrium design, and one must approach the off-equilibrium design problem with extreme caution. The discussion in (Leith and Leithead 1999) is interesting in this context.

Finally we note that the excessive degrees of freedom can be exploited to parameterise entire off-equilibrium regions in some desired manner. For example, entire regions could be possibly parameterised using only Hurwitz (the eigenvalues of the A_i matrices have negative real parts) local models, resulting in a system model that is robust from a simulation viewpoint. This observation opens a number of interesting future research directions.

5 Local models: Constant, linear or affine?

The most common choices of dynamic Takagi-Sugeno fuzzy models, are local linear models, affine linear models, and local constants. In this section we examine the consequences of model choice for model interpretability and identifiability.

Example: Longitudinal speed dynamics, different local model structures.

As an example, consider the longitudinal speed dynamics of a vehicle with mass m and speed v. The vehicle is powered by an engine which generates a longitudinal force $g_e(v, u)$ where u is the throttle angle. The vehicle is subject to a disturbance force g_d. A simple first order model of the vehicle is given by the force balance,

$$m\dot{v} \quad = \quad g_e(v, u) - g_d, \tag{30}$$

which can be written

$$\dot{v} \quad = \quad f(v, u) \quad = \quad \frac{g_e(v, u) - g_d}{m}. \tag{31}$$

In the example, we set $g_d = 1000$ N, $m = 1000$ kg and the engine characteristic is given by

$$g_e(v, u) \quad = \quad 500(1 + 3u)(1 + \arctan(6u^2 - 0.4v + 1.2))$$

(see Figure 3). With this characteristic engine curve (which corresponds to a fixed gear ratio), the engine operates in a speed interval between 2 and 20 m/s. [4]

Linearization of the engine model (31) leads to the following characteristic parameters: Pole $\frac{\partial f}{\partial v}(v, u)$, gain $\frac{\partial f}{\partial u}(v, u)$ and trend $f(v, u)$. These parameters are illustrated in Figure 4.

[4] This example is motivated by the experimental vehicle speed control problem considered in (Johansen *et al.* 1998*b*). The model is simplified, but contains the relevant aspects of the experimental vehicle in order to illustrate the main ideas.

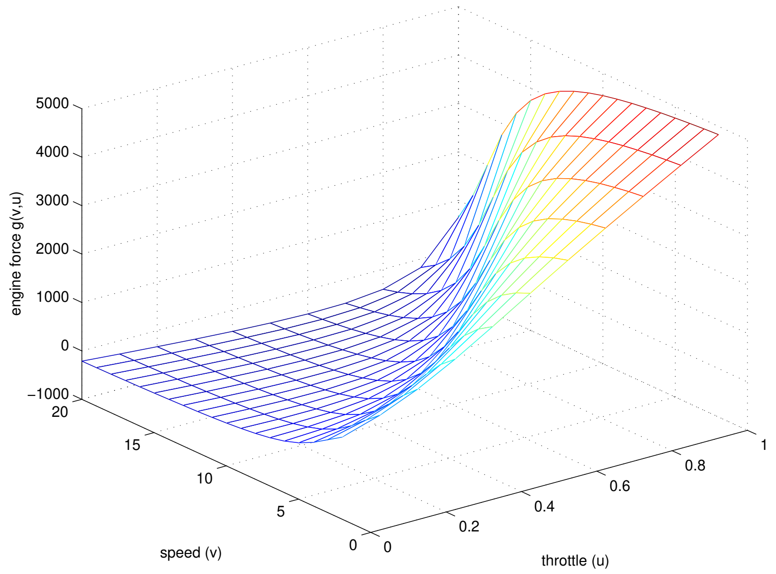

Figure 3: Engine force.

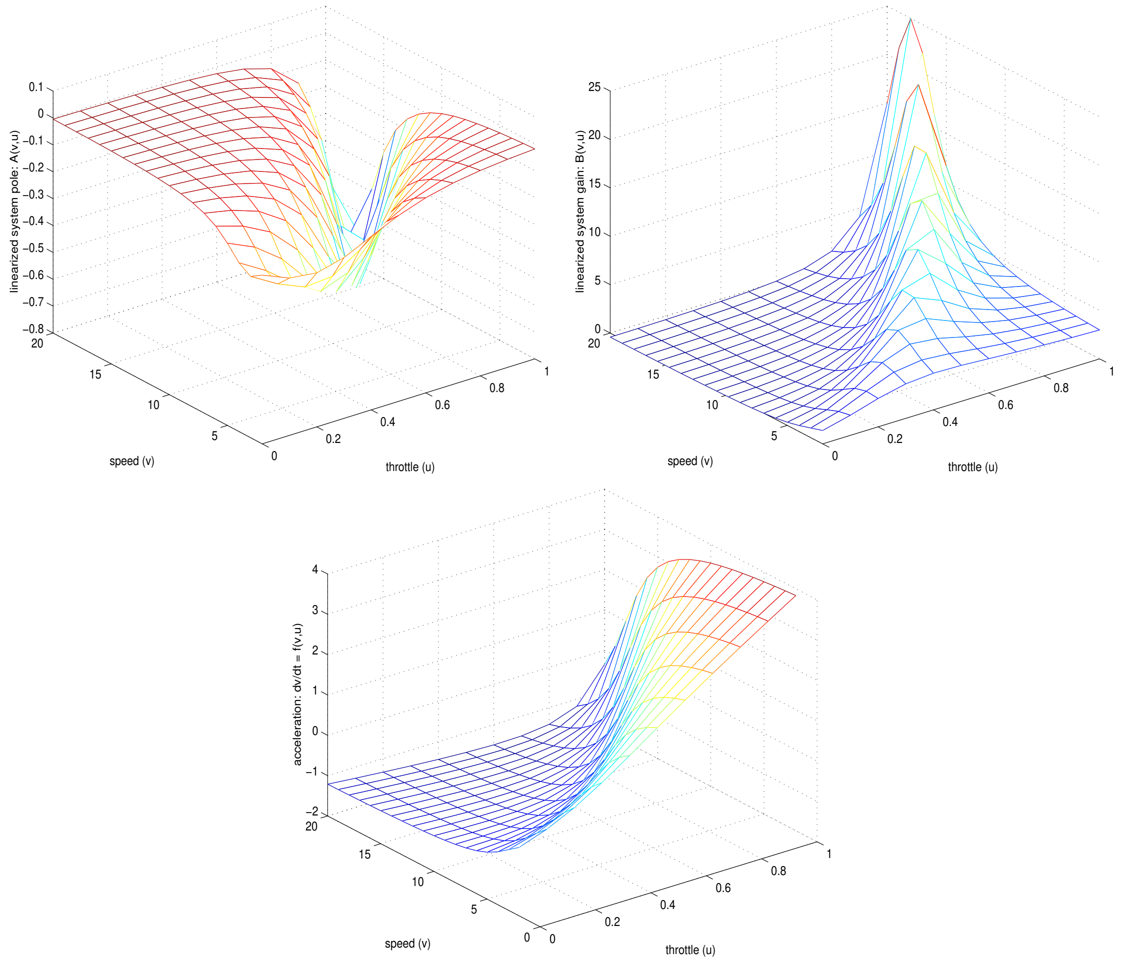

Figure 4: Linearized engine model (pole, gain, trend).

Assume that the state-space partitioning of the Takagi-Sugeno fuzzy model is such that a valid local model in a neighbourhood of the point $v_1 = 10$, $u_1 = 0.1$ is required. Consider the following alternative local model structures:

$$\dot{v} = A_1 v + B_1 u + d_1 \tag{32}$$

$$\dot{v} = A_1 v + B_1 u \tag{33}$$

$$\dot{v} = d_1 \tag{34}$$

Local models with these structures that are exact models at the point $(v, u) = (10, 0.1)$, are shown in Figure 5. The local linear and affine models have equilibria at $(v, u) = (0.2434u, u)$ and $(v, u) = (13.0943u - 6.6959, u)$ respectively. All these equilibrum points are located far outside the region of validity of their corresponding local models, and are therefore not equilibrium points for the nonlinear model. We can thus conclude that the point $(v, u) = (10, 0.1)$ is a transient point, which can be confirmed simply by observing that the acceleration is non-zero at this point. The location of the equilibrium points of the local linear and affine models has therefore no relevance when the local model is a transient model. Moreover, it is clear that only the local affine model results in an accurate local model with a valid interpretation as a local linearization (in the sense that the local model parameters A_i and B_i corresponds approximately to the pole and gain of the linearization of the nonlinear system).

Notice that with the local linear model (where the plane is restricted to go through the origin and cannot therefore always be a tangent plane), it is possible to select the A_i and B_i parameters in different ways. In Figure 5 the pole A_i is chosen equal to the pole of the local linearization, which leads to $B_i = -4.0516$ which is different from the gain of the local linearization $\partial f / \partial u(10, 0.1) = 0.8960$ because we require that the local model is exact at the point $(v, u) = (10, 0.1)$. Hence, although it cannot be interpreted as a local linearization, it is still an accurate model of the trend near the point $(v, u) = (10, 0.1)$. Another alternative is to select the A_i and B_i parameters of the local linear model exactly equal to the pole and gain of the local linearization. In this case the local model parameters can be interpreted as a local linearization about a trajectory passing through $(v, u) = (10, 0.1)$, but because the constant term is zero, the model cannot be used to give accurate global nonlinear predictions. Hence, with local linear models, one can achieve accurate approximation of either the linearized dynamics or the trend, but not both simultaneously. With the local affine model structure there are no such limitations, while the local constant model structure, by its nature, contains trend information but no information on the linearization dynamics. Note that we only discuss the information contained in the constituent local models when interpreted *individually*. Some results on the ability of the *global* Takagi-Sugeno fuzzy model with constituent local linear models to approximate local linearizations can be found in (Fantuzzi and Rovatti 1996).

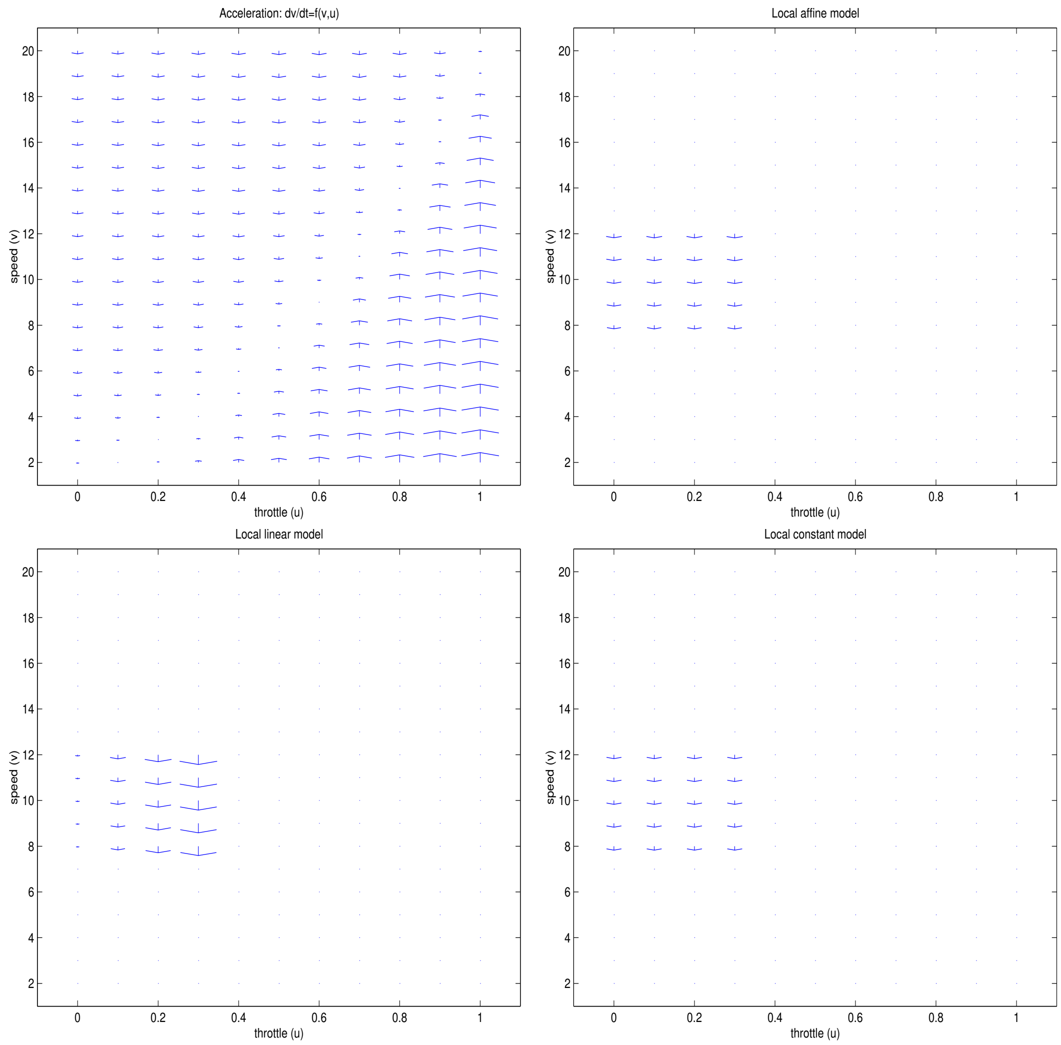

Figure 5: The flow of the nonlinear system and local models (affine, linear and constant) at the point $v = 10$, $u = 0.1$.

It is particularly difficult to interpret local models that correspond to transient operating regimes (where no equilibrium points exist), as opposed to local models corresponding to equilibrium operating regimes. This is true even if the transient local model has a valid interpretation as a local linearization. Consider for example the point $v = 20$, $u = 0$ of the vehicle modelling example. The local linearization is given by

$$\dot{v} = -0.0085(v-20) - 0.2978u - 1.4248 \tag{35}$$

The model gain is given by -0.2978. Naive control design based upon this model, might interpret the model gain as meaning that in order to decrease the speed, the throttle angle must be increased. This is an incorrect interpretation. The cause of the misinterpretation is that the trend (constant -1.4248) has not been taken into account. At this point the vehicle will have an acceleration of -1.4248 with zero throttle angle, and this is in fact the dominating dynamics in this region since a throttle angle varying in the interval, $0 \leq u \leq 1$, leads to a force corresponding to an acceleration in the interval, $-0.2978 \leq \dot{v} \leq 0$. The gain is so small that any perturbation of the throttle angle is of minor importance compared to the trend. Hence, it is clear that little information about the local dynamics can be inferred from the linearization parameters $\frac{\partial f}{\partial v}(v,u)$ and $\frac{\partial f}{\partial u}(v,u)$ alone without taking into consideration the trend. This concerns only off-equilibrium local models since equilibrium local models by definition have zero trend and are in general much easier to interpret. The above example illustrates that interpretation problems exist, even when the transient local model has a completely valid interpretation as a local linearization. The interpretation problems are even more severe when this interpretation is invalid.

□

So far we have argued by means of theoretical constructs, and practical examples, that Takagi-Sugeno fuzzy models, with all three local model structures, are capable of accurately representing global nonlinear dynamics. However, as we have seen, even with a locally affine model structure, not all of these local affine models have a valid interpretation as a local linearization. If the local model parameters are identified from data (rather than resulting from an explicit linearization of a nonlinear model), one has no guarantee that the obtained affine local models have a valid interpretation as local linearizations. If this is important, special care must be taken during experiment design and identification to achieve this.

It is also evident that the identifiability problems are due to the use of affine local models in transient operating regions. If local linear or constant local models are applied, the problems due to excessive degrees of freedom may be avoided, but such local models will not be valid local linearizations of the system.

6 Effect of minimizing the number of premise variables

A common modelling objective is to obtain a parsimonious parameterisation of the system dynamics. In the context of local model structures, parsimonious representations of the system are sometimes difficult to obtain due to the curse of dimensionality. Consequently, to reduce the complexity of the Takagi-Sugeno fuzzy model, it is common (where possible) to restrict the membership functions to depend on a subset of the variables (x, u). If f depends affinely on some of these variables, it is known that it is not necessary to partition along these axes (Johansen and Foss 1993). In cases where the nonlinearities are not too strong, one tends to minimize the number of premise variables in order to keep the model complexity to a minimum. Similar methods are employed in gain scheduled control where it is common to restrict the number of scheduling variables to the number that is necessary in order to characterize the equilibrium manifold, e.g. (Rugh 1991, Lawrence and Rugh 1995) (even though it has been argued that this may restrict the transient performance of the control system (Johansen *et al.* 1998*b*)). In any case, practical considerations usually necessitate keeping the number of scheduling/premise variables as low as possible to reduce the effects of the curse of dimensionality.

Example: Exploiting the affine functional form of f.

Consider the very simple nonlinear system

$$\dot{x} = -xu \tag{36}$$

At the point (x_i, u_i) the local linearization is

$$\dot{x} = -u_i x - x_i u + x_i u_i \tag{37}$$

Observing that the system (36) is bilinear, it is clear that it is sufficient to use either x or u in the premise of the Takagi-Sugeno fuzzy model (Johansen and Foss 1993). Selecting x as the premise variable, we get an alternative local model

$$\dot{x} = -x_i u \tag{38}$$

valid in some neighbourhood of x_i, i.e. $(x, u) \in (x_i - \delta, x_i + \delta) \times R$ for some small $\delta > 0$. The Takagi-Sugeno fuzzy model

$$\dot{x} = -\left(\sum_{i=1}^{N} x_i w_i(x) \right) u \tag{39}$$

is an arbitrary close approximation to the system (36) on any compact set by appropriate selection of the points $x_1, x_2, ..., x_N$.

Clearly, the local models (38) leads to an accurate global model, but with $A_i = 0$ and $B_i = -x_i$ they do not have a valid interpretation as local linearization. The reason for this is that we have left out the input variable u from the premise due to the bilinearity. Hence, the number of premise (scheduling) variables is reduced without reducing the accuracy, but by sacrificing the interpretability of the local model as local linearizations. Notice that at equilibrium $x = 0$ or $u = 0$, both approaches to construction of local models lead to similar results, emphasizing that they have excessive degrees of freedom only in transient operating regions.

□

Another consequence of reduction of the number of scheduling/premise variables is that a single local model may be used both in transient and equilbrium operating regimes. If the dynamics are significantly different in the transient and equilibrium operating regimes corresponding to a single local model, this obviously leads to difficulties.

Example: Longitudinal speed dynamics, cont'd.

In the longitudinal vehicle speed model, the equilibrium points are given by the curve in Figure 6. This curve can be parameterized by a single variable, for example either speed or throttle angle. If one wants to restrict the number of premise variables to only one, it is natural to select either speed or throttle angle. Suppose we select speed. For a nominal speed $v = 12$ m/s the corresponding *assumed* region of validity is the shaded region in Figure 6. When the system is operating far away from equilibrium, but with a speed in the interval $10 \leq v \leq 14$, this local model contributes to the Takagi-Sugeno fuzzy model, cf. Figure 6. Within the shaded region corresponding to both equilibrium and transient operating conditions, $-0.6630 \leq A(v,u) \leq -0.0098$ and $-0.3473 \leq B(v,u) \leq 17.9217$. When only considering equilibrium operating conditions corresponding to $10 \leq v \leq 14$, we get $-0.6630 \leq A(v,u) \leq -0.6081$ and $13.9088 \leq B(v,u) \leq 17.9271$. It is clear that the local affine model is reasonably accurate at the equilibrium operating conditions corresponding to the interval $10 \leq v \leq 14$, but when the transient operating conditions corresponding to this interval are included, the dynamics are too varied over the operating range for a single local affine model to be valid.

□

Although this example may seem trivial, it is included because this problem has been experienced to be relevant in practical applications, especially when no special attention has been given to the distinction between equilibrium and off-equilibrium operating conditions during experiment design and model structure selection.

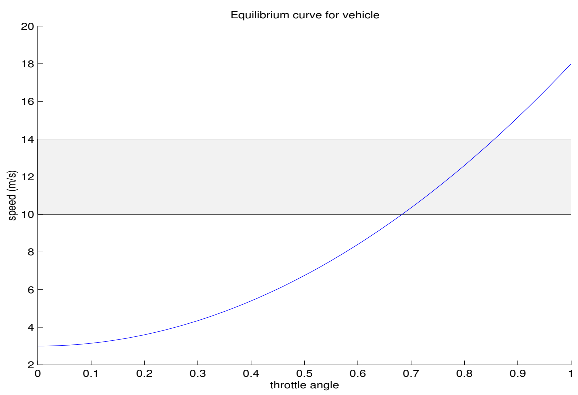

Figure 6: Equilibrium curve for vehicle model.

7 The role of the membership functions

We have argued that the global Takagi-Sugeno fuzzy model can provide a satisfactory approximation of the underlying non-linear system, even when the constituent affine local models are not conventional linearizations. In practice Takagi-Sugeno models are constructed by interpolating the parameters of the constituent local models using fuzzy inference (see also (Babuska and Verbruggen 1997)). The choice of membership functions is of crucial importance in this procedure. To illustrate some of the issues which arise when choosing membership functions, we consider the following example. Let

$$\dot{x} = A_1(x - x_1) \tag{40}$$

be the linearization of some autonomous nonlinear system about the equilibrium point x_1. Similarly, let

$$\dot{x} = A_2(x - x_2) \tag{41}$$

be the linearization of the same nonlinear system about the equilibrium point x_2. The Takagi-Sugeno model of the non-linear system constructed from these local models defined above is of the form,

$$\begin{aligned} \dot{x} &= (w_1(x)A_1 + w_2(x)A_2)x + w_1(x)d_1 + w_2(x)d_2 && (42)\\ &= A(x)x + d(x) && (43) \end{aligned}$$

where $d_1 = -A_1x_1$, and where $d_2 = -A_2x_2$, and where $w_1(x)$ and $w_2(x)$ are interpolation functions. If these linearizations are sufficiently close together it can be assumed that this model approximates the nonlinear system in some fashion between these equilibrium points. The model predicts a manifold of equilibria given by

$$A(x_0)x_0 + d(x_0) = 0. \tag{44}$$

Furthermore, if we assume that the model also describes the transitions between the equilibrium points, then model linearizations along (44) are given by:

$$\dot{x} \quad = \quad \tilde{A}(x_0)(x - x_0) + \text{h.o.t.} \tag{45}$$

where

$$\tilde{A}(x) \quad = \quad \sum_{i=1}^{2} \left(A_i w_i(x) + (A_i x + d_i) \left(\frac{\partial w_i}{\partial x}(x) \right)^T \right) \tag{46}$$

Let us consider the meaning of the above equations. Equation (44) suggests that the manifold of equilibria, is described as a function of the interpolation functions, the matrices A_1 and A_2, and the endpoints x_1 and x_2. This may or may not be a reasonable assumption. Furthermore, (46) suggests that the system linearization along this manifold cannot be obtained by simple interpolation of the matrices A_1 and A_2; the membership functions themselves contribute to the linearization. In fact, (46) contains two terms; a term which is the interpolation of the matrices A_1 and A_2 and a term that is a function of the derivatives of the interpolation functions. The interpolation functions appear in both of these terms. It is clear that the membership functions should be chosen to provide as accurate a representation of the system dynamics as possible. However, membership functions chosen to approximate the global dynamics may be inadequate when the model is linearised. In many applications this may not be an issue. However, for control applications requiring linearization of the plant model, model fidelity in the neighbourhood of plant equilibria is of paramount importance. In our context we must ensure that (46) provides as accurate an approximation to the system linearization as possible. However, there are several difficulties associated with the approximation method proposed in (46). Firstly, there is no guarantee that the qualitative nature of the locii of eigenvalues constructed by interpolating between the matrices A_1 and A_2 is consistent with the qualitative nature of these matrices. In addition, we must also select our membership functions whoose first derivatives not only exist, but also satisfy certain properties. From these observations it can be seen that the choice of membership functions is by no means a trivial matter, affecting not only the global model dynamics, but also the manifold of equilibria, and the system linearizations along the equilibria.

In general, for a globally accurate fuzzy model consisting of constant, linear or affine local models, a locally valid linearization may be generated by linearizing the equations (2):

$$\dot{x} \quad = \quad \tilde{A}(x_0, u_0)(x - x_0) + \tilde{B}(x_0, u_0)(u - u_0) + \tilde{d}(x_0, u_0) + \text{h.o.t.} \tag{47}$$

where

$$\tilde{A}(x, u) \quad = \quad \sum_{i=1}^{N} \left(A_i w_i(x, u) + (A_i x + B_i u + d_i) \left(\frac{\partial w_i}{\partial x}(x, u) \right)^T \right) \tag{48}$$

$$\tilde{B}(x,u) = \sum_{i=1}^{N} \left(B_i w_i(x,u) + (A_i x + B_i u + d_i) \left(\frac{\partial w_i}{\partial u}(x,u) \right)^T \right) \tag{49}$$

$$\tilde{d}(x,u) = \sum_{i=1}^{N} (A_i x + B_i u + d_i) w_i(x,u) \tag{50}$$

As discussed above, the linearization will depend on the membership functions, and it usually assumed that they provide an accurate smooth interpolation of the local models. In practice, verification of this assumption is carried out experimentally.

8 Identification of Takagi-Sugeno fuzzy models

The purpose of this section is to illustrate the poor results one might get when identifying off-equilibrium affine local models using standard identification methods such as least squares, and suggest alternative methods with better robustness and performance.

Example: Longitudinal vehicle dynamics, least squares identification.

For the identification of a Takagi-Sugeno fuzzy model of the vehicle we have generated by simulation a data sequence with 1000 noise-free samples with a sampling interval $T = 1$ s. The experiment design takes the vehicle to several equilibrium states through some transient states and perturbs the system by a Pseudo-Random Binary Signal (PRBS) signal at each equilibrium, cf. Figure 7.

The operating region $(v,u) \in [2,20] \times [0,1]$ is then partitioned into 5 equilibrium regions and 4 transient regions, each of which is characterized by a membership function. The corresponding affine local models are identified from data sequence 1. A discrete-time dynamic Takagi-Sugeno fuzzy model with constituent local models of the form

$$v(k+1) = -a_{i,1} v(k) + b_{i,1} u(k) + \delta_i \tag{51}$$

is first identified using the least squares algorithm. Using exact discretization of the continuous time first order local model $\dot{x} = A_i x + B_i u + d_i$ and zero-order-hold of the input signal

$$x(t+T) = \exp(A_i T) x(t) + \frac{1}{A_i} (\exp(A_i T) - 1)(B_i u(t) + d_i) \tag{52}$$

this discrete-time model is converted into a continous-time model by the following parameter transformation (derived from (51) and (52):

$$A_i = \ln(-a_{i,1})/T \tag{53}$$

$$B_i = \frac{\ln(-a_{i,1}) b_{i,1}}{-a_{i,1} - 1} \tag{54}$$

$$d_i = \frac{\ln(-a_{i,1}) \delta_i}{-a_{i,1} - 1} \tag{55}$$

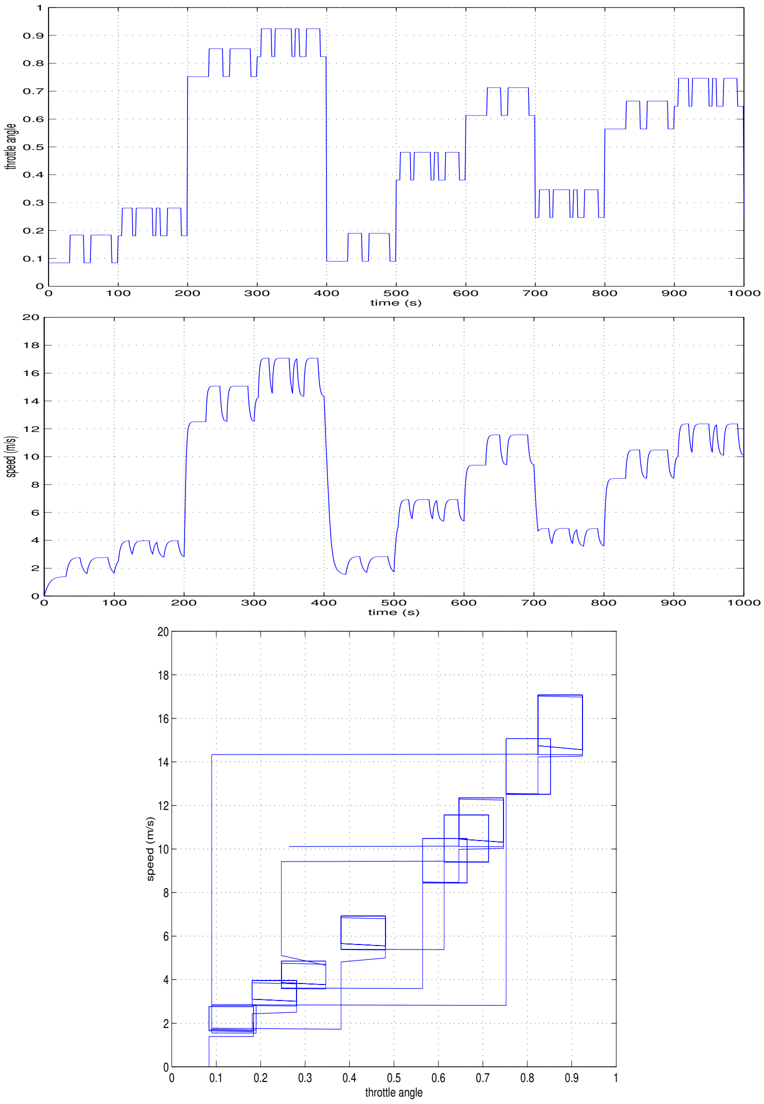

Figure 7: Identification data sequence 1.

Let us define the pole, zero and trend that results when we interpolate the local model parametes:

$$\overline{A}(v,u) = \sum_{i=1}^{N} A_i w_i(v,u) \tag{56}$$

$$\overline{B}(v,u) = \sum_{i=1}^{N} B_i w_i(v,u) \tag{57}$$

$$\overline{d}(v,u) = \sum_{i=1}^{N} d_i w_i(v,u) \tag{58}$$

If the local model parameters are valid linearizations, then the interpolated local model parameters should be similar to the parameters resulting from dynamic linearization. The plots of $\overline{A}(v,u), \overline{B}(v,u)$ and $\overline{d}(v,u)$ in Figure 8 shows that this is not the case. First, there is some bias because the model structure is not sufficiently rich to exactly model the system. But the most interesting feature of the identified model is the large difference between the identified parameters and the parameters of the linearized system in the off-equilibrium regions. In this case, this is due to both poor identifiability of the off-equilibrium local models and because off-equilibrium identification data are relatively sparse.

□

Identification of the parameters of (26), using for example, a standard least-squares criterion, and some experimental data, will only treat local models as approximators and thus not necessarily lead to local model parameters A_i and B_i with a valid interpretation as local linearizations. This was illustrated above, and also in some application studies (Shorten *et al.* 1999, Johansen *et al.* 1998*a*). The core of the problem is poor identifiability. These problems are amplified when there is very sparse information about the response to perturbations in transient operating regimes available in the data, which is a very typical situation in real-world applications. The reasons for this are diverse: The system typically spends little time in transient conditions compared to stationary operating conditions. Further, in transient regions, the trend component in the data will have larger amplitude and therefore more influence in the identification criterion than perturbation components. Furthermore, in such regions, operational constraints often limit the amount of perturbations and the time the system should spend in transient states. Hence, our practical experience is that carefully planned experiments are required in order to get even a modest amount of perturbation information in transient operating regimes. Furthermore, prior knowledge should be applied as far as possible in the identification in order to make it more robust. Next, we consider experiment design and continue with a discussion of various identification methods that can help improve the interpretability and accuracy of the identified model.

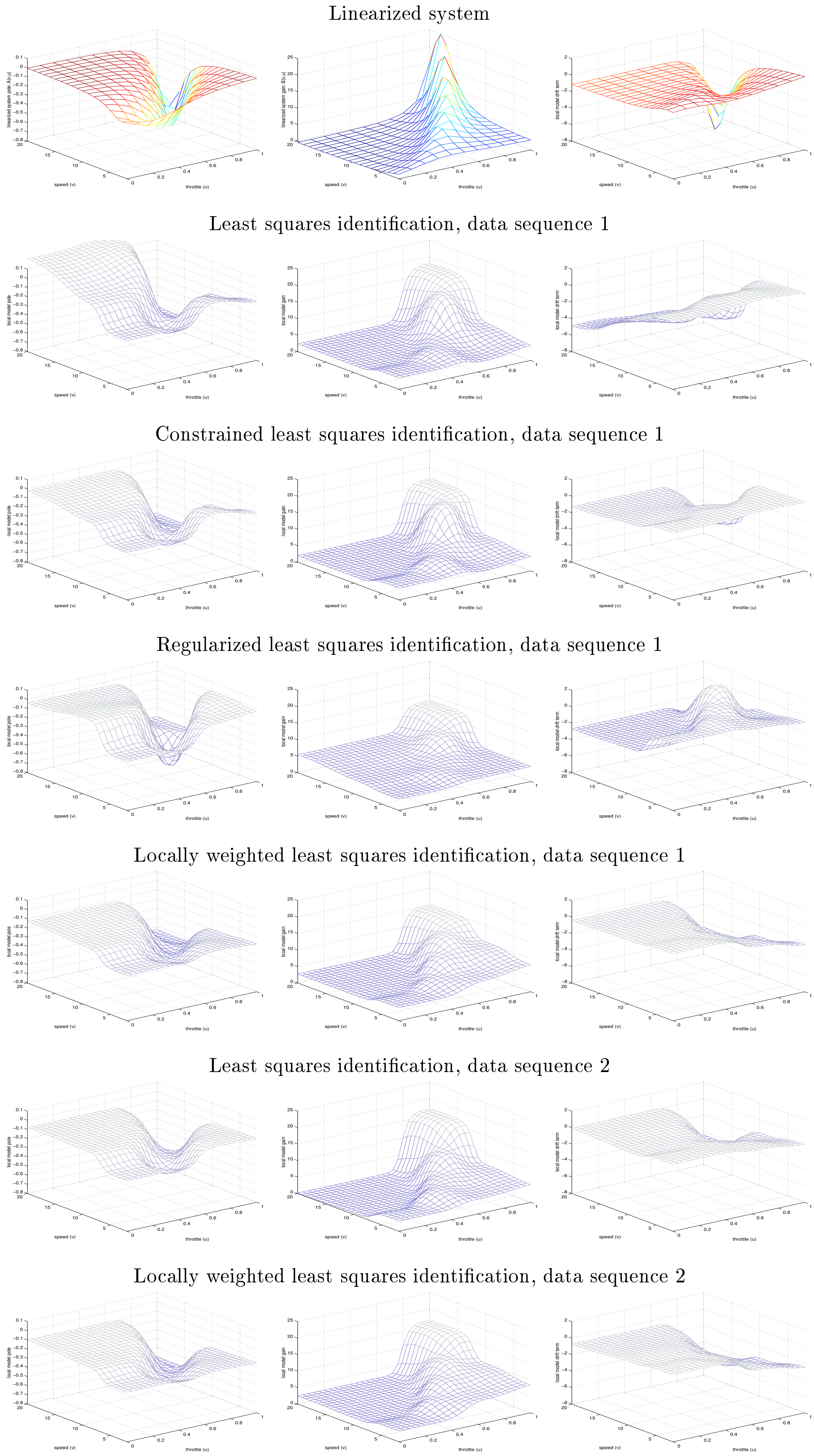

Figure 8: Comparison of the pole/gain/trend of the linearized system and the interpolated local model pole/gain/trend of the Takagi-Sugeno fuzzy models identified using various identification methods and data sequences.

8.1 Experiment design

In nonlinear system identification, both the amplitude and frequency contents of the input signals are of major importance. For identifying Takagi-Sugeno fuzzy models containing both equilibrium and off-equilibrium local affine models, it is our experience that input signals designed according to the following guidelines are useful if the model is intended for control design:

- The system should be brought through a sequence of equilibria that includes the equilibria of the local models. At each equilibrium the system should be excited by superpositioned small-amplitude PRBS signals. The PRBS signals should have a frequency content that covers an interval from the inverse rise-time to above the bandwidth of the closed-loop system.

- Several transient trajectories should be generated for each off-equilibrium local model. The corresponding input signals should contain both large-amplitude steps and perturbations in order to determine both the trend and perturbation dynamics of the off-equilibrium local models. The frequency contents should typically be higher compared to the frequency content of the equilibrium data to prevent the system from settling at some equilibrium.

Of course, there are practical constraints that will often limit the number of transitions, length of the experiment, frequency content and amplitudes. The requirements in terms of accuracy of the off-equilibrium local models will depend strongly on the application. Sometimes, equilibrium local models can be extrapolated into transient operating regions without significant loss of accuracy.

Example: Longitudinal vehicle dynamics, improved experiment design.

The data sequence 2 shown in Figure 9 is generated according to the above guidelines. The first 1000 samples of this data sequence is identical to data sequence 1 which excitates the equilibrium behaviour of the system. The second 1000 samples contains a high-frequency large-amplitude input signal that will take the system through its transient states.

Identification of the Takagi-Sugeno fuzzy model, with exactly the same structure as the models considered above, is now carried out using the least squares algorithm and this data sequence. The results are shown in Figure 8. Clearly, these depict a more accurate model, both in terms of global accuracy, and the closeness of the local affine models to the linearization of the system.

□

8.2 Constrained and regularized identification

Poor identifiability leads to an ill-conditioned identification problem where certain parameters (or combinations of parameters) can be chosen more or less arbitrarily. Constrained and regularized

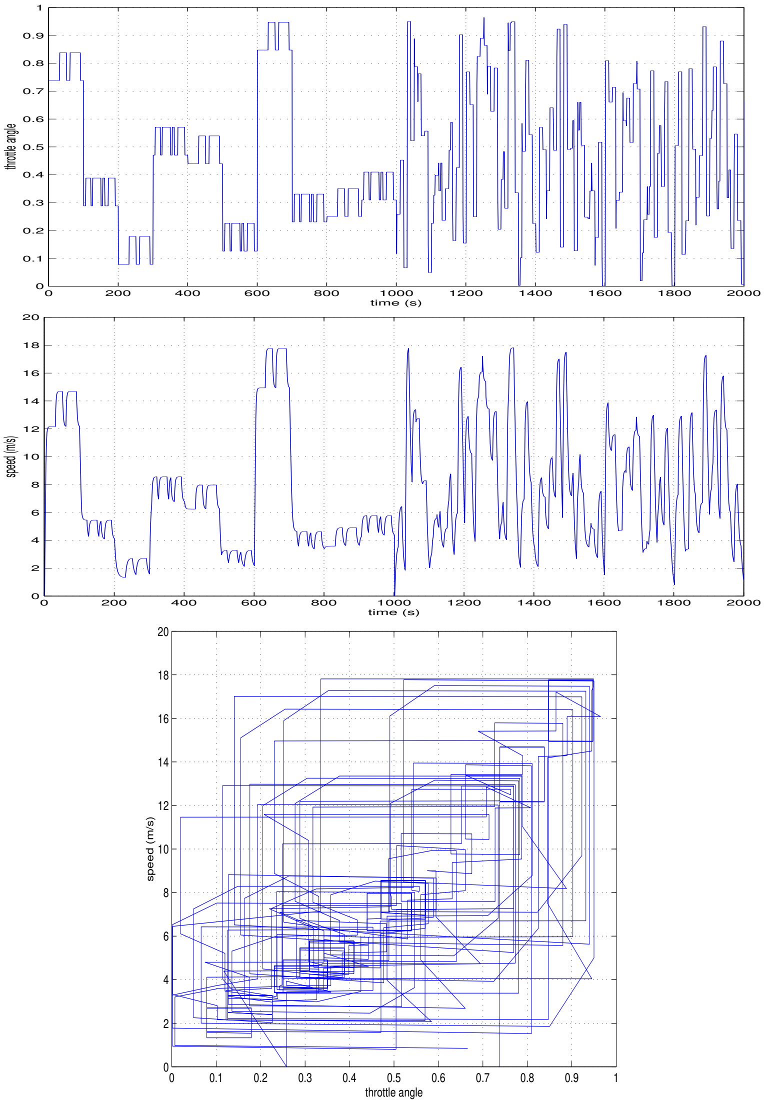

Figure 9: Identification data sequence 2.

system identification is a general method for improving the robustness and accuracy of the system identification algorithm when the model structure is poorly identifiable (Sjöberg and Ljung 1992, Sjöberg *et al.* 1993, Johansen 1997*b*, Johansen 1997*a*). The general idea is to explicitly constrain the model parameters using hard or soft constraints such that the model is consistent with some prior knowledge or desired properties. Regularization (which can be seen as soft constraints) may be implemented by adding a penalty function to the least squares criterion. This penalty will attract the model parameters towards some set of *a priori* desirable parameters.

It is assumed that the identification is carried out using discrete-time data and a discrete-time dynamic Takagi-Sugeno fuzzy model that admits the one-step-ahead predictor to be written in the form $\hat{y}(k|k-1) = \theta^T \varphi(k-1)$ where θ is a vector that contains the local model parameters $\theta_1, ..., \theta_N$ and $\varphi(k)$ is an information vector that may depend on the current state, as well as the current and past inputs and outputs.

The common least squares identification method will minimize the least squares criterion

$$V_{LS}(\theta) = \sum_{k=1}^{l} ||y(k) - \theta^T \varphi(k-1)||^2 \tag{59}$$

where l is the length of the data sequence. Constraints are normally added in terms of a set of linear equalites and inequalities that incorporates some form of prior knowledge:

$$H\theta = h \tag{60}$$

$$K\theta \leq k \tag{61}$$

where H, K, h and k are constant matrices and vectors. Minimizing the least squares criterion (59) subject to the constraints (60) and (61) is a convex optimization problem that can be solved using quadratic programming (Luenberger 1984).

We have found that an approximate Tikhonov regularization stabilizer may be useful for improving both the accuracy and interpretability of the identified local models. This stabilizer (penalty function) is given by (Johansen 1996)

$$\Omega_{\text{TIK}}^{\text{approx}}(\theta) = \sum_{i=1}^{N} \sum_{j=1}^{i-1} (\theta_i - \theta_j)^T W_{ij} (\theta_i - \theta_j) \rho_{i,j} \tag{62}$$

where $\rho_{i,j}$ is a measure of how close the fuzzy sets with indices i and j are. Hence, $\rho_{i,j}$ is close to one for a neighboring pair of fuzzy sets and close to zero for a distant pair of fuzzy sets. The matrix W_{ij} is a positive definite diagonal weighting matrix that ensures that the different local model parameters have sensible relative weights in the stabilizer. The interpretation of (62) is that the local model parameters θ_i and θ_j corresponding to neighbouring regions are expected to be similar, while there should not be any such constraint on local model parameters of regions that are far apart. The definition of the weighting factor $\rho_{i,j}$ should depend on the parameterization of the membership functions for the fuzzy sets. Let us give an example of how it may be defined, using tensor product Gaussian membership functions:

$$\rho_{i,j} = \exp\left(-\frac{1}{2}\sum_{k=1}^{n}\frac{(c_{i,k}-c_{j,k})^2}{\sigma_{i,k}\sigma_{j,k}}\right)$$

where $c_i = (c_{i,1}, ..., c_{i,n})$ is the center point of the fuzzy set with index i, and $\sigma_{i,k}$ is the average "standard deviation" of the membership functions defining the fuzzy sets along the kth dimension.

Combining the approximate Tikhonov stabilizer with the least squares criterion and the constraints leads to a quadratic programming problem of minimizing

$$V(\theta) = V_{LS}(\theta) + \gamma\Omega_{\text{TIK}}^{\text{approx}}(\theta) \tag{63}$$

subject to the constraints (60) and (61). The scalar regularization parameter $\gamma > 0$ defines the weight on the penalty function relative to the least squares penalty on the prediction error. This parameter can be selected on the basis of both subjective and objective criteria. Further details on some objective statistical criteria can be found in (Johansen 1997a).

Example: Longitudinal vehicle dynamics, constrained identification.

Suppose we restrict the gains of the local models to be non-negative, i.e. $B_i \geq 0$, and the poles of the local models to be non-positive, i.e. $A_i \leq 0$ and use the constrained least squared identification algorithm with data sequence 1. Eqs. (53)-(55) are used to convert between continuous and discrete time model parameterizations. The results are shown in Figure 8. The equilibrium local models are almost unchanged since their poles and gains were consistent with the constraints also in the unconstrained identification experiment. The off-equilibrium local models are now more accurate, which one should expect due to the introduction of prior knowledge that was violated in the unconstrained identification experiment.

□

Example: Longitudinal vehicle dynamics, regularization.

Next, we apply the approximate Tikhonov stabilizer (62) without any constraints. The identification results using data sequence 1 are shown in Figure 8. The equilibrium local models are almost unchanged since they were easily identifiable also with the pure least squares algorithm. The off-equilibrium local models are now somewhat more accurate, even though the Tikhonov stabilizer was based on the invalid prior assumption that the local off-equilibrium models were similar to neighbouring local equilibrium models. In this particular example, this assumption is clearly violated (notice that the gain and pole are significantly larger near the equilibrium manifold than in transient regions). The reason why better results are achieved is that the increased bias introduced by the incorrect prior assumption is less than the reduction in variance that always accompanies the regularization. A study of bias and variance as a result of regularization can be found in (Johansen 1997*b*), see also (Sjöberg and Ljung 1992).

□

The structure identification problem becomes somewhat more complicated when constraints and regularization are introduced. However, well know criteria such as the Final Prediction Error criterion (FPE) has been generalized to these cases (Larsen and Hansen 1994, Johansen 1997*a*).

8.3 Locally weighted identification

The least squares objective (59) explicitly aims at determining the parameters of the local models that gives the best global model prediction. As we have already discussed, this objective will often be in conflict with the objective of determining local models that are accurate linearizations of the nonlinear system. This motivates the use of parametric locally weighted identification where this is the explicit objective of the identification (Murray-Smith and Johansen 1997). Individual criteria are defined for each local model:

$$V^i_{\mathrm{LWLS}}(\theta_i) \quad = \quad \sum_{k=1}^{l} ||y(k) - \theta_i^T \varphi_i(k-1)||^2 w_i(z(k-1) \tag{64}$$

where $z(k)$ is a vector containing the premise variables. The data are weighted according to their relevance to the local model with index i, and this relevance is measured by the interpolation function w_i. The one-step-ahead predictor based on the local model with index i is assumed to be given in the form $\hat{y}(k|k-1;i) = \theta_i^T \varphi_i(k-1)$. The resulting identification problem is still linear in the parameters and can be solved explicitly. The implicit regularization effect of locally weighted least squares has been studied in detail in (Murray-Smith and Johansen 1997). Notice that locally weighted least squares can be combined with explicit regularization and constraints, still leading to a convex quadratic program for the local parametric identification problem. In (Yen *et al.* 1998) it was suggested to combine the local criteria (64) with the global criterion (59) in order to address the trade-off between local accuracy and global performance.

Example: Longitudinal vehicle dynamics, locally weighted identification.

Parameter identification results using locally weighted least squares are shown in Figure 8 for the cases when data sequences 1 and 2 are applied. It can be seen that the interpolated pole/gain/trend of the identified local models are close approximations to the linearization of the nonlinear system, even with data sequence 1 in the off-equilibrium regions. The reduction in global prediction performance compared to the global least squares is not significant.

□

Recently, there has been some interest in the use of locally weighted least squares identification (Fan and Gijbels 1996) to determine local linear dynamic models (Stenman *et al.* 1996, Bontempi *et al.* 1999). In these approaches the local models are determined online at each sample, and there is no underlying nonlinear model structure. On the other hand, it has also been proposed to use nonparametric offline identification methods based on a probabilistic multivariate Gaussian process prior model of the underlying nonlinear function (Murray-Smith *et al.* 1999). This approach has a high computational load, but appears to be very robust with respect to poor identifiability. Local affine models which are close to linearizations can easily be derived analytically from the nonparametric model, with analytic estimates of their variance.

9 Conclusions

We have illustrated fundamental interpretability and identifiability limitations of dynamic Takagi-Sugeno fuzzy models with constituent affine local models. The major problem is that in transient operating regimes (where the nonlinear system has no equilibrium points) the local affine model is poorly identifiable. Further, practical limitations exist that restrict the design of informative experiments for many systems. A theoretical understanding of the problems is developed by relating the dynamic Takagi-Sugeno fuzzy model with local affine models to dynamic linearization. Differences between local constants, linear and affine models are pointed out, the effect of minimizing the number of variables is the premise of the fuzzy rules are investigated, and the role of the membership functions are studied. Finally, suggestions are presented, that potentially alleviate or reduce some of these problems through suitable experiment design and use of robust identification methods. In particular, it is shown how constrained and regularized identification methods may improve the interpretability of the constituent local models as local linearizations. Furthermore, locally weighted least squares identification is shown to explicitly address the tradeoff between local and global accuracy of the Takagi-Sugeno fuzzy model.

The practical importance of the problems are illustrated by very simple examples. Still, one may expect that the problems related to interpretability and identifiability will be much more pronounced when more complex higher-order and multi-variable examples are considered.

The present results are not only relevant to dynamic Takagi-Sugeno fuzzy models, but also related model representation with constituent local models, such as Local Model Networks (e.g. (Johansen and Foss 1993, Murray-Smith and Johansen 1997, Hunt and Johansen 1997)) and piecewise linear models (e.g. (Billings and Voon 1987)).

Acknowlegements

The work of the first author was sponsored by the European Comission under the ESPRIT Long Term Research project 28104 H^2C. Roderick Murray-Smith gratefully acknowledges the support of Marie Curie TMR grant FMBICT961369.

Appendix A: Proof of Theorem 1

$$\begin{aligned}
\Delta\dot{\xi}_2(t) &= \dot{\xi}_2(t) - \dot{x}_0(t) &(65)\\
&= \sum_{i=1}^{N} f_i(x_i,u_i)w_i(\xi_2(t),u(t)) - f(x_0(t),u_0(t)) \\
&\quad + \sum_{i=1}^{N} \left(A(x_i,u_i)(\xi_2(t)-x_i) + B(x_i,u_i)(u(t)-u_i)\right) w_i(\xi_2(t),u(t)) &(66)\\
&= f(\xi_2(t),u(t)) - f(x_0(t),u_0(t)) \\
&\quad + \sum_{i=1}^{N} \left(f_i(x_i,u_i)w_i(\xi_2(t),u(t)) - f(\xi_2(t),u(t))\right) \\
&\quad + \sum_{i=1}^{N} \left(A(x_i,u_i)(\xi_2(t)-x_i) + B(x_i,u_i)(u(t)-u_i)\right) w_i(\xi_2(t),u(t)) &(67)\\
&= A(x_0(t),u_0(t))\Delta\xi_2(t) + B(x_0(t),u_0(t))\Delta u(t) \\
&\quad + \sum_{i=1}^{N} \left(f_i(x_i,u_i)w_i(\xi_2(t),u(t)) - f(\xi_2(t),u(t))\right) \\
&\quad + \sum_{i=1}^{N} \left(A(x_i,u_i)(\xi_2(t)-x_i) + B(x_i,u_i)(u(t)-u_i)\right) w_i(\xi_2(t),u(t)) \\
&\quad + \text{h.o.t.} &(68)\\
&= A(x_0(t),u_0(t))\Delta\xi_2(t) + B(x_0(t),u_0(t))\Delta u(t) + \varepsilon(\xi_2(t),u(t),x_0(t),u_0(t)) &(69)
\end{aligned}$$

where

$$\begin{aligned}
\varepsilon(\xi_2(t),u(t),x_0(t),u_0(t)) &= \sum_{i=1}^{N} \left(f_i(x_i,u_i)w_i(\xi_2(t),u(t)) - f(\xi_2(t),u(t))\right) \\
&\quad + \sum_{i=1}^{N} \left(A(x_i,u_i)(\xi_2(t)-x_i) + B(x_i,u_i)(u(t)-u_i)\right) w_i(\xi_2(t),u(t)) \\
&\quad + \text{h.o.t.} &(70)
\end{aligned}$$

From the definition of δ is follows immediately that $w_i(\xi_2,u) = 0$ if $||(\xi_2,u) - (x_i,u_i)||_2 > \delta$. Application of the triangle inequality on (70) leads to

$$||\varepsilon(\xi_2,u,x_i,u_i)||_2 \leq L\delta + L\delta + O(\delta^2) = 2L\delta + O(\delta^2) \tag{71}$$

References

Babuska, R. and H. B. Verbruggen (1997). Fuzzy modelling and identification. In: *Multiple Model Approaches to Modelling and Control* (R. Murray-Smith and T. A. Johansen, Eds.). Taylor and Francis Ltd., London.

Billings, S. A. and W. S. G. Voon (1987). Piecewise linear identification of non-linear systems. *Int. J. Control* **46**, 215–235.

Bontempi, G., M. Birattari and H. Bersini (1999). Lazy learning for modeling and control design. *Int. J. Control* **72**, 643–658.

Driankov, D. and Palm, R., Eds.) (1998). *Advances in Fuzzy Control.* Physica-Verlag, Heidelberg.

Driankov, D., R. Palm and U. Rehfuss (1996). A Takagi-Sugeno fuzzy gain-scheduler. In: *Proc. IEEE Conf. Fuzzy Systems, New Orleans.* pp. 1053–1059.

Fan, J. and I. Gijbels (1996). *Local Polynomial Modelling and Its Applications.* Chapman and Hall.

Fantuzzi, C. and R. Rovatti (1996). On the approximation capabilities of the homogeneous Takagi-Sugeno model. In: *Proc. IEEE Conf. Fuzzy Systems, New Orleans.* pp. 1067–1072.

Hunt, K. J. and T. A. Johansen (1997). Design and analysis of gain-scheduled local controller networks. *Int. J. Control* **66**, 619–651.

Johansen, T. A. (1994). Fuzzy model based control: Stability, robustness, and performance issues. *IEEE Trans. Fuzzy Systems* **2**, 221–234.

Johansen, T. A. (1996). Robust identification of Takagi-Sugeno-Kang fuzzy models using regularization. In: *Proc. IEEE Conf. Fuzzy Systems, New Orleans.* pp. 180–186.

Johansen, T. A. (1997*a*). Constrained and regularized system identification. In: *Preprints IFAC Symposium on System Identification, Kitakyushu, Japan.* pp. 1467–1472.

Johansen, T. A. (1997*b*). On Tikhonov regularization, bias and variance in nonlinear system identification. *Automatica* **33**, 441–446.

Johansen, T. A. and B. A. Foss (1993). Constructing NARMAX models using ARMAX models. *Int. J. Control* **58**, 1125–1153.

Johansen, T. A., K. J. Hunt and H. Fritz (1998*a*). A software environment for gain scheduled local controller network design. *IEEE Control Systems Magazine* **18**(2), 48–60.

Johansen, T. A., K. J. Hunt, P. J. Gawthrop and H. Fritz (1998*b*). Off-equilibrium linearisation and design of gain scheduled control with application to vehicle speed control. *Control Engineering Practice* **6**, 167–180.

Jordan, D. W. and P. Smith (1987). *Nonlinear Ordinary Differential Equations*. Oxford University Press, Oxford, UK.

Kailath, T. (1980). *Linear Systems*. Prentice Hall, Englewood Cliffs, N.J.

Larsen, J. and L. K. Hansen (1994). Generalization performance of regularized neural network models. In: *Proc. IEEE Workshop on Neural Networks for Signal Processing, Ermioni, Greece*.

Lawrence, D. A. and W. J. Rugh (1995). Gain scheduling dynamic linear controllers for a nonlinear plant. *Automatica* **31**, 381–390.

Leith, D. J. and W. E. Leithead (1999). Analytic framework for blended multiple model systems using local linear models. *Int. J. Control* **72**, 605–619.

Lohmiller, W. and J.-J. E. Slotine (1998). On contraction analysis for nonlinear systems. *Automatica* **6**, 683–696.

Luenberger, D. G. (1984). *Introduction to Linear and Nonlinear Programming, 2nd Ed.*. Addison-Wesley, Inc., Reading, MA.

Murray-Smith, R. and T. A. Johansen (1997). Local learning in local model networks. In: *Multiple Model Approaches to Modelling and Control* (R. Murray-Smith and T. A. Johansen, Eds.). Taylor and Francis, London.

Murray-Smith, R., T. A. Johansen and R. Shorten (1999). On transient dynamics, off-equilibrium behaviour and identification in blended multiple model structures. In: *Proceedings of the Eurpoean Control Conference, Karlsruhe*. pp. F1020–7.

Passino, K. M. and S. Yurkovich (1998). *Fuzzy Control*. Addison Wesley Longman, Menlo Park CA.

Rovatti, R. (1996). Takagi-Sugeno models as approximators in Sobolev norms - the SISO case. In: *Proc. IEEE Conf. Fuzzy Systems, New Orleans*. pp. 1060–1066.

Rugh, W. J. (1991). Analytical framework for gain scheduling. *IEEE Control Systems Magazine* **11**(1), 79–84.

Shorten, R., R. Murray-Smith, R. Bjørgan and H. Gollee (1999). On the interpretation of local models in blended multiple model structures. *Int. J. Control* **72**, 620–628.

Sjöberg, J. and L. Ljung (1992). Overtraining, regularization, and searching for minimum in neural networks. In: *Preprints IFAC Symposium on Adaptive Systems in Control and Signal Processing, Grenoble, France*. pp. 669–674.

Sjöberg, J., T. McKelvey and L. Ljung (1993). On the use of regularization in system identification. In: *Preprints 12th IFAC World Congress, Sydney*. Vol. 7. pp. 381–386.

Stenman, A., F. Gustafsson and L. Ljung (1996). Just in time models for dynamical systems. In: *Proc. IEEE Conf. Decision and Control, Kobe, Japan*. pp. 1115–1120.

Takagi, T. and M. Sugeno (1985). Fuzzy identification of systems and its application to modeling and control. *IEEE Trans. Systems, Man, and Cybernetics* **15**, 116–132.

Wang, H. O., K. Tanaka and M. F. Griffin (1996). An approach to fuzzy control of nonlinear systems: Stability and design issues. *IEEE Trans. Fuzzy Systems* **4**, 14–23.

Wang, L.-X. (1993). Stable adaptive fuzzy control of nonlinear systems. *IEEE Trans. Fuzzy Systems* **1**, 146–155.

Wang, L.-X. and J. M. Mendel (1992). Fuzzy basis functions, universal approximation, and orthogonal least-squares learning. *IEEE Trans. Neural Networks* **3**, 807–814.

Yen, J., L. Wang and C. W. Gillespie (1998). Improving the interpretability of TSK fuzzy models by combining global learning and local learning. *IEEE Trans. Fuzzy Systems* **6**, 530–537.