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

Tor A. Johansen, *Member, IEEE*, Robert Shorten, *Member, IEEE*, and Roderick Murray-Smith

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 validation and interpretation of the model considerably. 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 tends 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 identification algorithms are studied. These include constrained and regularized identification and locally weighted identification. Their usefulness in the present context is illustrated by examples.

Index Terms—Dynamic analysis, fuzzy models, linearization, system identification, transient dynamics.

I. INTRODUCTION

CONSIDER the nonlinear dynamic system

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

where $x \in X \subset R^n$ and $u \in U \subset R^r$. A Takagi–Sugeno fuzzy model will approximate this system by smoothly interpolating affine local models [34]. Each local model contributes to the global model in a fuzzy subset of $X \times U$. This fuzzy set is characterized by a membership function $\mu_i : X \times U \to [0, 1]$, 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 $w_i : X \times U \to [0, 1]$ 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 assumes that $\sum_j \mu_j(x, u) > 0$ for all $(x, u) \in X \times U$.

The Takagi–Sugeno fuzzy model has recently found wide applicability in fuzzy model based control, e.g., [4], [9], [27], [35], [36]. 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 local 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., [5], [8], 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 behavior of the nonlinear model (2) is similar to the global behavior 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, [9], [36]. On the other hand, it is sometimes required and often desirable 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

T. A. Johansen is with the Department of Engineering Cybernetics, Norwegian University of Science and Technology, N-7491 Trondheim, Norway. He is also with SINTEF Electronics and Cybernetics, N-7491 Trondheim, Norway (e-mail: Tor.Arne.Johansen@itk.ntnu.no).

R. Shorten is with the Department of Computer Science, National University of Ireland, Maynooth, County Kildare, Ireland (e-mail: rshorten@cs.may.ie).

R. Murray-Smith is with the Department of Computing Science, Glasgow University, Glasgow G12 8QQ, Scotland (e-mail: rod@dcs.gla.ac.uk).

Publisher Item Identifier S 1063-6706(00)05117-1.

basis for a fuzzy gain-scheduled controller since the local linear models are used to design local linear controllers [5], [8]. This property is also very useful when validating and analyzing the dynamic model (2), [30].

It has widely been observed that it is often nontrivial to identify dynamic local models which are close approximations to linearizations of the nonlinear system [14], [24], [25], [30], [38]. 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 to transient operating regimes [30]. Another important reason is that the algorithm of choice for many practitioners (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 [24], [25], [38]. The problems are, in most practical applications, amplified by constraints on the experiment design which restrict the amount of information in the transient data, as exemplified in [14]. 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 generally more difficult to understand the meaning and properties of the off-equilibrium local models than equilibrium local models in the Takagi–Sugeno fuzzy model [30]. Most of the above mentioned problems are unique to dynamic modeling and identification problems (as opposed to modeling and identification of static mappings) although the tradeoff between local and global approximation accuracy also appears in static modeling problems considered in [25] and [38]. The present paper explicitly focuses on dynamic systems, which are considerably more challenging than static function approximation. The aim is to present an *overview* of some problems and solutions related to identification and interpretation of dynamic Takagi–Sugeno fuzzy models, and this paper continues the work of [30]. The main message in [30] is that local affine dynamic models in Takagi–Sugeno fuzzy dynamic models contain excessive degrees of freedom and must be interpreted carefully. The interpretability problem is also discussed in [20], where a velocity-based linearization approach is suggested as an alternative. 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 [10], [12], [25].

The outline of this paper is as follows. In Section II, it is shown that a dynamic Takagi–Sugeno fuzzy model can form an approximation to dynamic linearization about arbitrary trajectories. In Section III, it is illustrated that there are various interpretability problems associated with the Takagi–Sugeno fuzzy model structure. In particular, when the constituent local models are identified from experimental data there is a risk of missing interpretability for the sake of accuracy. Identification algorithms that combine high accuracy with high interpretability are discussed in Section IV. A simplified simulation example considering dynamic modeling 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 modeling problem. Conclusions are given in Section V.

II. EXISTENCE OF INTERPRETABLE TAKAGI–SUGENO FUZZY MODELS

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 were considered in [37], while the theory was developed for local polynomial models (including affine) in [13]. Finally, in [28] 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 linearized dynamics. However, notice that the approximation result of [28] 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 [15] 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 [15] considers autonomous systems, and an extension to systems with exogenous inputs is given below.

A. 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 (denoted by h.o.t.) in the Taylor series expansion (4). In deviation coordinates $\Delta x = x - x_0'$, $\Delta u = u - u_0'$, (4) becomes

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

where we have defined the functions A and B as follows:

$$A(x_0', u_0') = \frac{\partial f}{\partial x}(x_0', u_0') \tag{6}$$

$$B(x_0', u_0') = \frac{\partial f}{\partial u}(x_0', u_0'). \tag{7}$$

The local linearization (5) of the nonlinear system (1) 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 a point (x_0, u_0) on some arbitrary trajectory. Hence, it need not be an equilibrium point. 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

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

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

$$\dot{\Delta x} = A(x_0, u_0)\Delta x + B(x_0, u_0)\Delta u + \text{h.o.t.} \tag{9}$$

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

$$\dot{x} = f(x_0, u_0) + \text{h.o.t.} \tag{10}$$

which approximates the flow of the state by a constant vector near the point (x_0, u_0) on the nominal trajectory $(x_0(t), u_0(t))$. Obviously, the higher order terms in (10) and (8) are of different order.

B. Approximation Theorem

Assume a continuous trajectory (x_0, u_0) that satisfies $(x_0(t), u_0(t)) \in X \times U$ is given, and $\dot{x}_0(t) = f(x_0(t), u_0(t))$, $x_0(0) = \bar{x}$. By neglecting the higher order dynamics, the dynamic linearization about the trajectory (x_0, u_0) is now the linear time-varying (LTV) system

$$\dot{\xi}_1(t) = f(x_0(t), u_0(t)) + A(x_0(t), u_0(t))(\xi_1(t) - x_0(t))$$
$$+ B(x_0(t), u_0(t))(u(t) - u_0(t)) \tag{11}$$
$$\xi_1(0) = \bar{x} \tag{12}$$

where the functions A and B are defined in (6) and (7). An approximation to the dynamic linearization is given by the Takagi–Sugeno fuzzy system scheduled on the state/input, c.f. (2)

$$\dot{\xi}_2(t) = \sum_{i=1}^{N} (f(x_i, u_i) + A(x_i, u_i)(\xi_2(t) - x_i)$$
$$+ B(x_i, u_i)(u(t) - u_i))w_i(\xi_2(t), u(t)) \tag{13}$$
$$\xi_2(0) = \bar{x}. \tag{14}$$

We will continue with developing an approximation result that shows the closeness between (11) and (13). Let $Z^0 = \{(x_1, u_1), (x_2, u_2), \cdots, (x_N, u_N)\} \subset Z = X \times U$ 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 (13) approximates (11) when "Z^0 covers Z densely," for example, in the sense that the maximum distance between neighboring elements of Z^0

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

is sufficiently small and $\text{supp}(w_i)$ denotes the support of the function w_i, i.e. $\text{supp}(w_i) = \{z \in Z \,|\, w_i(z) > 0\}$. Thus, notice that $w_i(z) = 0$ if $\|z - z_i\|_2 > \delta$. This is formalized 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

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

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

where ε satisfies

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

□

$O(\epsilon)$ 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 linearizations about transient points) leads to an arbitrary close approximation to the LTV dynamic system resulting from dynamic linearization 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 discussion in [20] provides additional insight into the relationship between the constituent local linear models and the dynamic linearization.

III. INTERPRETABILITY PROBLEMS IN TAKAGI–SUGENO FUZZY MODELS

In this section, it is illustrated that it is not straightforward to identify constituent local models of Takagi–Sugeno fuzzy models that can be interpreted as local linearizations of the nonlinear system even though we proved in Theorem 1 that such local models exist. Furthermore, we show that even if local models that closely approximate local linearizations exist, they are still not easy to interpret in transient operating regimes.

A. Off-Equilibrium Local Models and Stability

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 [30] because they need not have anything to do with the equilibrium 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 behavior in the neighborhood of equilibria. For example, the eigenvalues of the $A(x_0, u_0)$ matrix will provide information about the stability of the virtual equilibrium 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 provide 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 provide any relevant 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., [17]. Stronger conditions are necessary (see, for example, [22] for some contraction conditions that ensure that neighboring trajectories of nonlinear systems are attracted toward 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 behavior 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 behavior 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 rigid string of length l, with an angle θ and angular velocity $\dot{\theta} = \omega$, leading to state-equations

$$\dot{\theta} = \omega \tag{19}$$

$$\dot{\omega} = -\frac{g}{l}\sin(\theta) - \frac{d}{l}\omega \tag{20}$$

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) and (20) about an arbitrary point (θ^o, ω^o) leads to the following linearized system:

$$\dot{\theta} = \omega^o + (\omega - \omega^o) \tag{21}$$

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

which can be simplified and written in the matrix form

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

where

$$d(\theta^o) = \begin{pmatrix} 0 \\ 9.81(-\sin(\theta^o) + \cos(\theta)^o\theta^o) \end{pmatrix} \tag{24}$$

$$A(\theta^o) = \begin{pmatrix} 0 & 1 \\ -9.81\cos(\theta^o) & -0.5 \end{pmatrix}. \tag{25}$$

The flow of the nonlinear pendulum system is illustrated in Fig. 1 [the flow is the velocity of the state-vector at each state, i.e., the vector field f defined by the right-hand side (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 toward 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 initial 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 toward 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 behavior of the system in transient operating regimes. $\square$

B. Loss of Identifiability

In Section II-B, we argued that under general conditions there exist Takagi–Sugeno model parameters such that

- the global fuzzy model (2) accurately approximates the global behavior of the nonlinear system (1);
- 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 II-A.

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

---

[1]The concept of stability of trajectories and paths, denoted Lyapunov stability and Poincare stability, respectively, is treated in [16].

[2]Lack of identifiability is characterized by nonuniqueness of the model structure, i.e., two different parameter vectors yields the same input/output behavior. Poor identifiability is (informally) defined as a somewhat relaxed property, namely, that two significantly different parameter vectors give very similar input/output behavior.
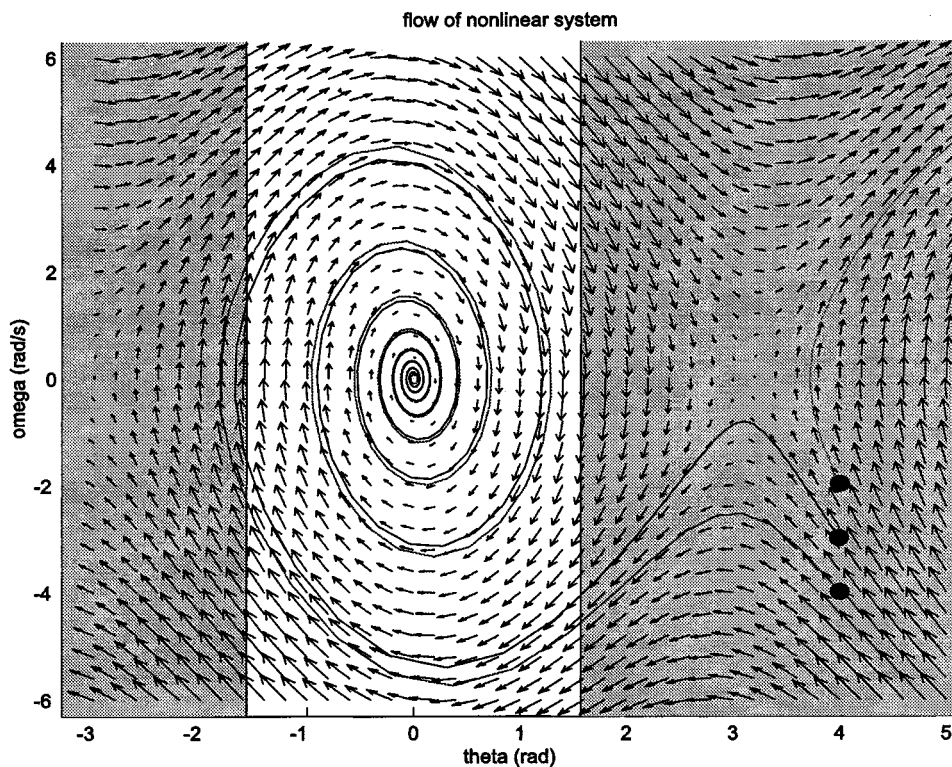
Fig. 1. The flow of the nonlinear pendulum system together with three trajectories starting a somewhat different initial conditions: $(4, -2), (4, -3), (4, -4)$.

interpretation, applicability and 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 IV.

Suppose we seek local affine models of the form

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

to be approximately valid in a small neighborhood 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

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

which is an exact model of (1) at the point (x_i, u_i) and a reasonable approximation in a small neighborhood 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

[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.

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 suboptimal choices for A_i and B_i when only considering global approximation accuracy [1], [24], [25].

Poor identifiability is a problem that is particularly pronounced 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 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: Consider the point $(\theta, \omega) = (\pi/3, 2)$, which is a transient state of the autonomous system. Fig. 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
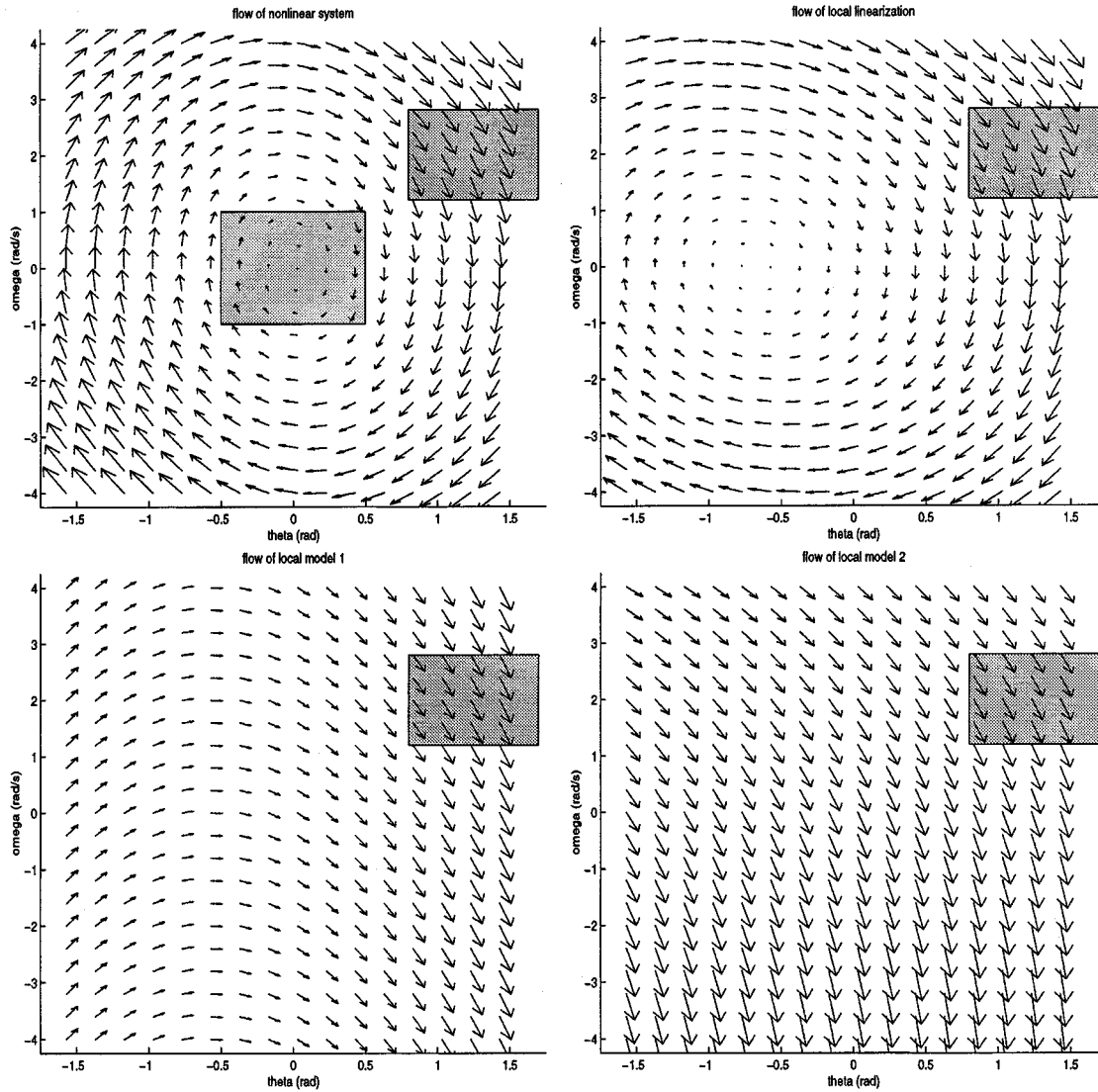
Fig. 2. The flow of the nonlinear pendulum system, local linearization, and two accurate local affine models.

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

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

$$A_1 = \begin{pmatrix} -0.20 & 0.20 \\ -1.00 & 1.00 \end{pmatrix} \quad d_1 = \begin{pmatrix} 1.75 \\ -10.00 \end{pmatrix}$$
(28)

$$A_2 = \begin{pmatrix} 0.00 & 0.00 \\ -1.00 & 0.00 \end{pmatrix} \quad d_2 = \begin{pmatrix} 2.00 \\ -3.00 \end{pmatrix}$$
(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 correspond to an unstable linear system. Both the eigenvalues of A_2 are 0, which also corresponds to an unstable linear system. The example il-

lustrates 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 nonunique off-equilibrium local models can be found in [30].

C. 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.
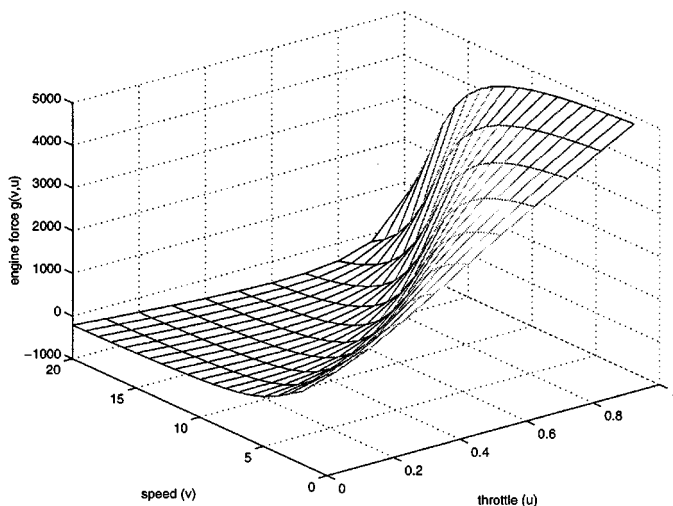
Fig. 3.   Engine force.

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

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

which can be written

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

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

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

(see Fig. 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 $(\partial f/\partial v)(v, u)$, gain $(\partial f/\partial u)(v, u)$ and trend $f(v, u)$. These parameters are illustrated in Fig. 4.

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

$$\dot{v} = A_1 v + B_1 u + d_1 \qquad (32)$$
$$\dot{v} = A_1 v + B_1 u \qquad (33)$$
$$\dot{v} = d_1. \qquad (34)$$

Local models with these structures that are exact models at the point $(v, u) = (10, 0.1)$, are shown in Fig. 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 equilibrium points are located far outside the region of validity of their corresponding local models and are, therefore, not equilib-

---

[4]This example is motivated by the experimental vehicle speed control problem considered in [15]. The model is simplified, but contains the relevant aspects of the experimental vehicle in order to illustrate the main ideas.

rium 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 nonzero at this point. The location of the equilibrium points of the local linear and affine models therefore has 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 Fig. 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 [7].

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 modeling example. The local linearization is given by

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

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 $0 \le u \le 1$, leading to a force corresponding to an acceleration in the interval $-0.2978 \le \dot{v} \le 0$. The gain is so small that any perturbation of the throttle angle is of minor importance compared to the
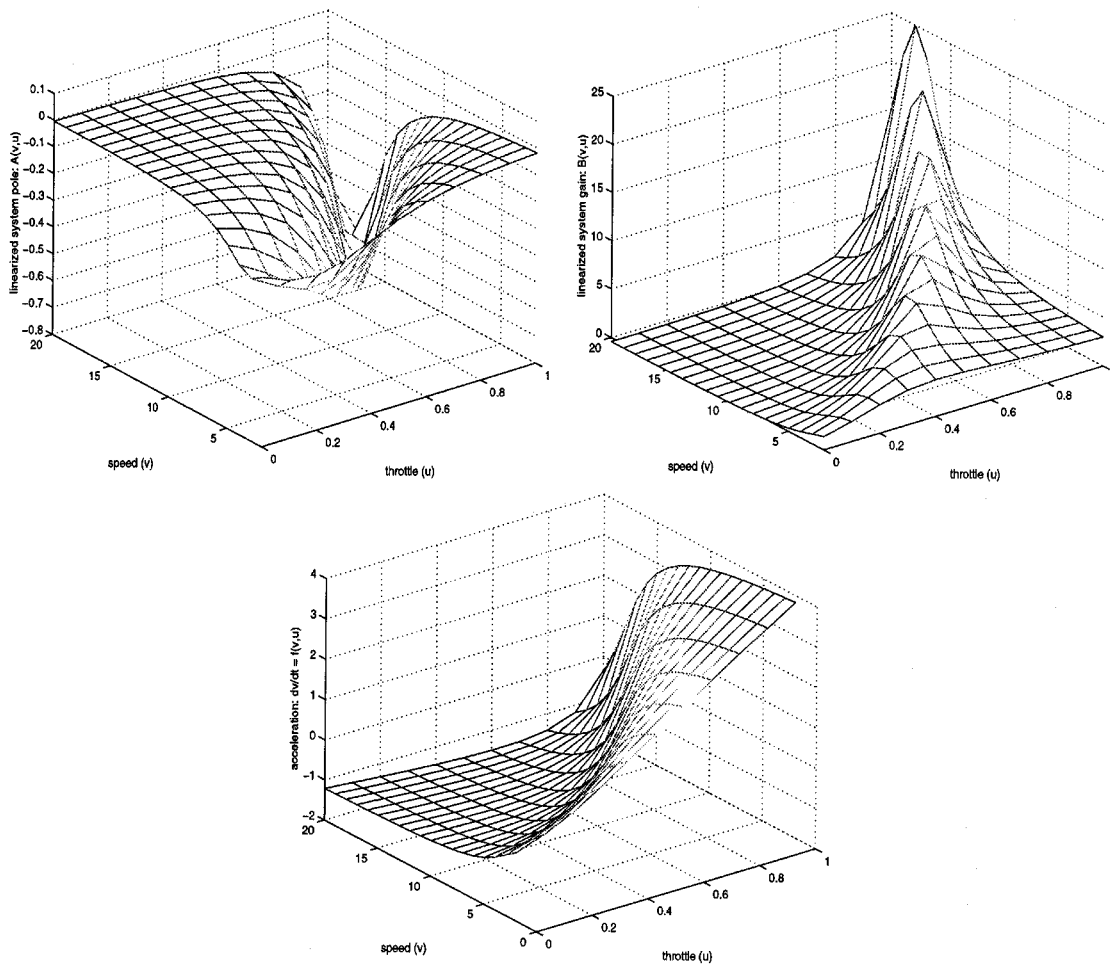
Fig. 4.   Linearized engine model (pole, gain, trend).

trend. Hence, it is clear that little information about the local dynamics can be inferred from the linearization parameters $(\partial f/\partial v)(v, u)$ and $(\partial f/\partial u)(v, u)$ 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.

D. Minimizing the Number of Premise Variables

A common modeling objective is to obtain a parsimonious parameterization 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 [13]. 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., [29], [19] (even though it has been argued that this may restrict the transient performance of the control system [15]). 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.
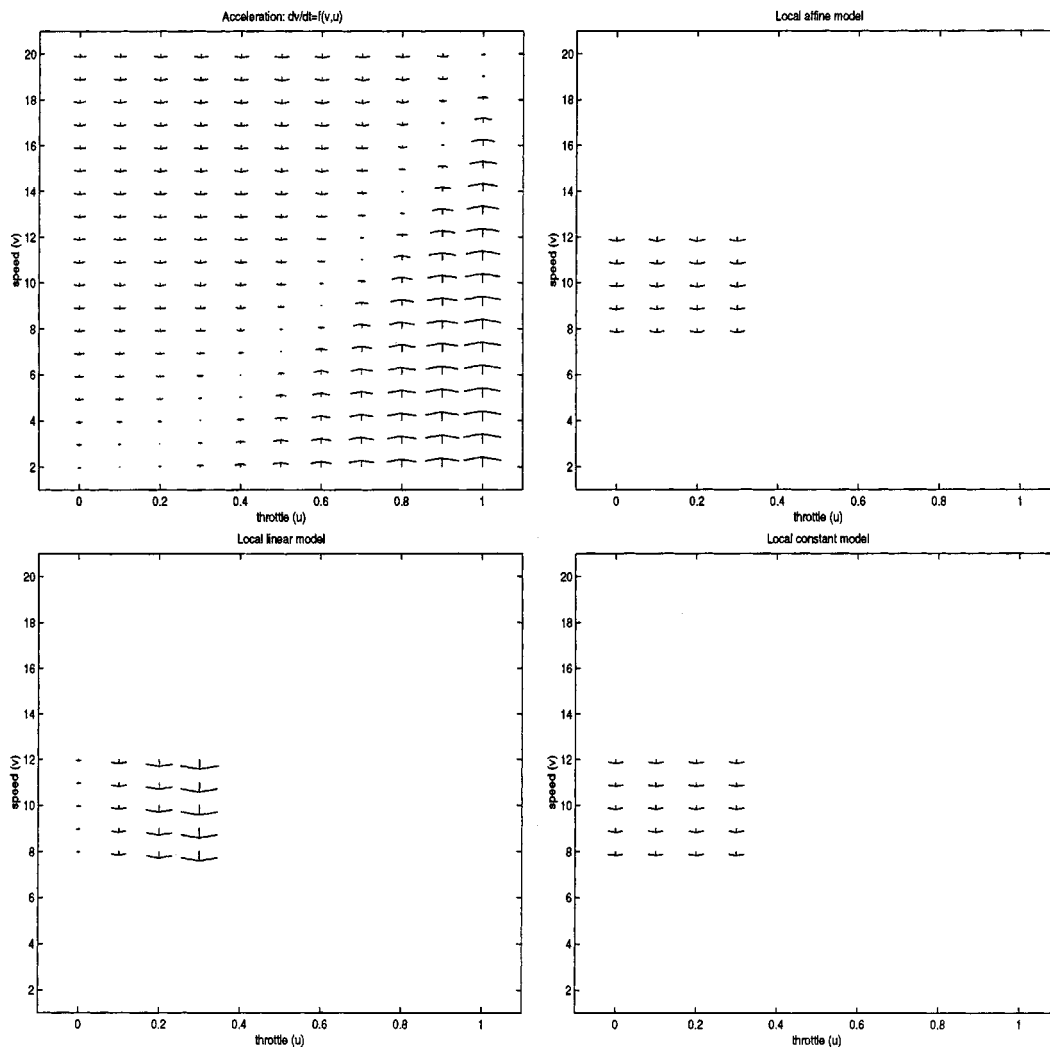
Fig. 5. The flow of the nonlinear system and local models (affine, linear, and constant) at the point $v = 10$, $u = 0.1$.

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 [13]. Selecting x as the premise variable, we get an alternative local model

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

valid in some neighborhood 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 $x_1, x_2, \cdots, x_N$.

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 equilibrium 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—Continued: In the longitudinal vehicle speed model, the equilibrium points are given by the curve in Fig. 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
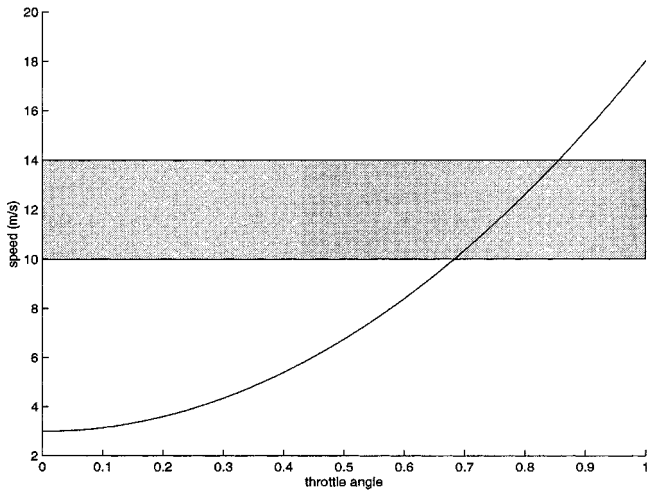
Fig. 6. Equilibrium curve for vehicle model.

in Fig. 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, c.f. Fig. 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.

E. The Role of the Membership Functions

We have argued that the global Takagi–Sugeno fuzzy model can provide a satisfactory approximation of the underlying nonlinear 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, [1]). The choice of membership functions is of crucial importance in this procedure [20]. 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 nonlinear

system constructed from these local models defined above is of the form

$$\dot{x} = (w_1(x)A_1 + w_2(x)A_2)x + w_1(x)\, d_1 + w_2(x)\, d_2 \tag{42}$$
$$= A(x)x + d(x) \tag{43}$$

where $d_1 = -A_1 x_1$, and where $d_2 = -A_2 x_2$, 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

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

where

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

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 linearized. In many applications this may not be an issue. However, for control applications requiring linearization of the plant model, model fidelity in the neighborhood 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 in (46). First, 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 whose 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 (2)

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

where

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

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

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

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.

IV. 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 pseudorandom binary signal (PRBS) signal at each equilibrium, c.f. Fig. 7.

The operating region $(v, u) \in [2, 20] \times [0, 1]$ is then partitioned into five equilibrium regions and four 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

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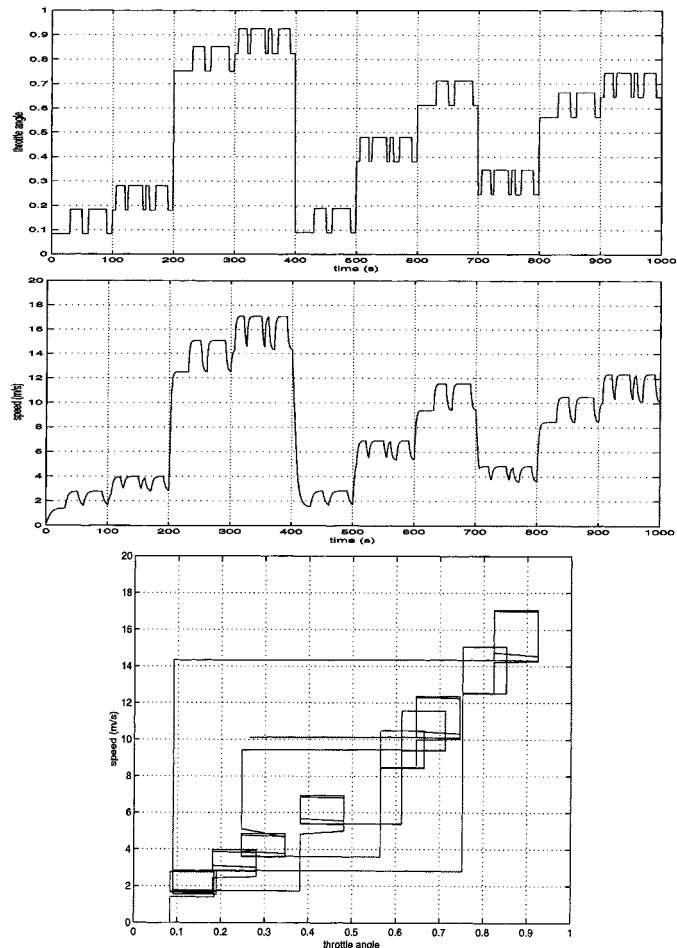$$+ \frac{1}{A_i} (\exp(A_i T) - 1)(B_i u(t) + d_i) \tag{52}$$

Fig. 7. Identification data sequence 1.

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

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

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

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

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

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

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

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

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 $\bar{A}(v, u)$, $\bar{B}(v, u)$, and $\bar{d}(v, u)$ in Fig. 8 shows that this is not the case. First, there is some bias because the model structure is not

Linearized system

Least squares identification, data sequence 1

Constrained least squares identification, data sequence 1

Regularized least squares identification, data sequence 1

Locally weighted least squares identification, data sequence 1

Least squares identification, data sequence 2

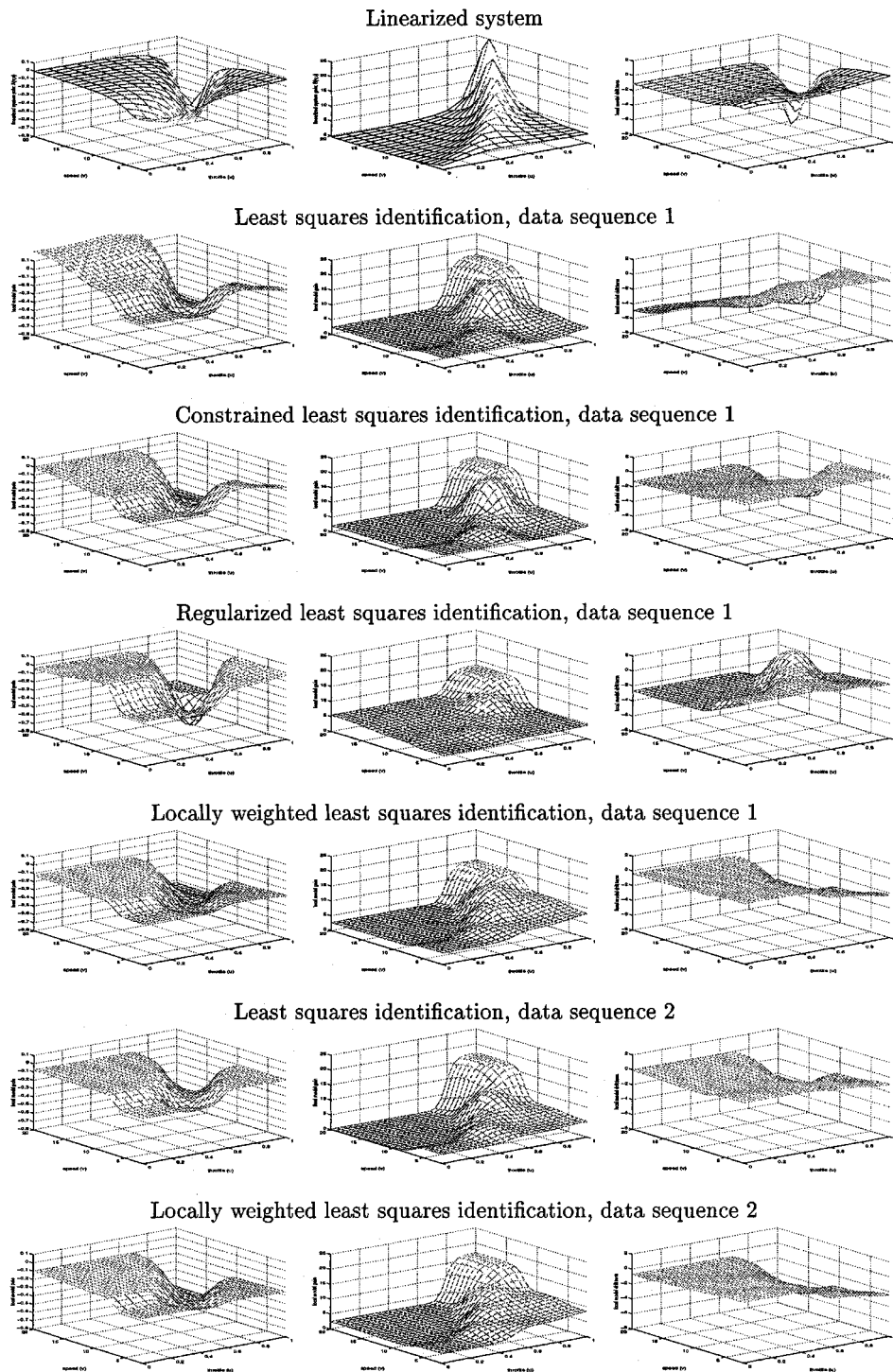Locally weighted least squares identification, data sequence 2

Fig. 8.　Comparison of the pole (left column), gain (middle column) and trend (right column) as a function of throttle and speed 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.

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 ([30], [14]). 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.

A. 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 Fig. 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 behavior 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 Fig. 8. Clearly, these depict a more accurate model, both in terms of global accuracy and the
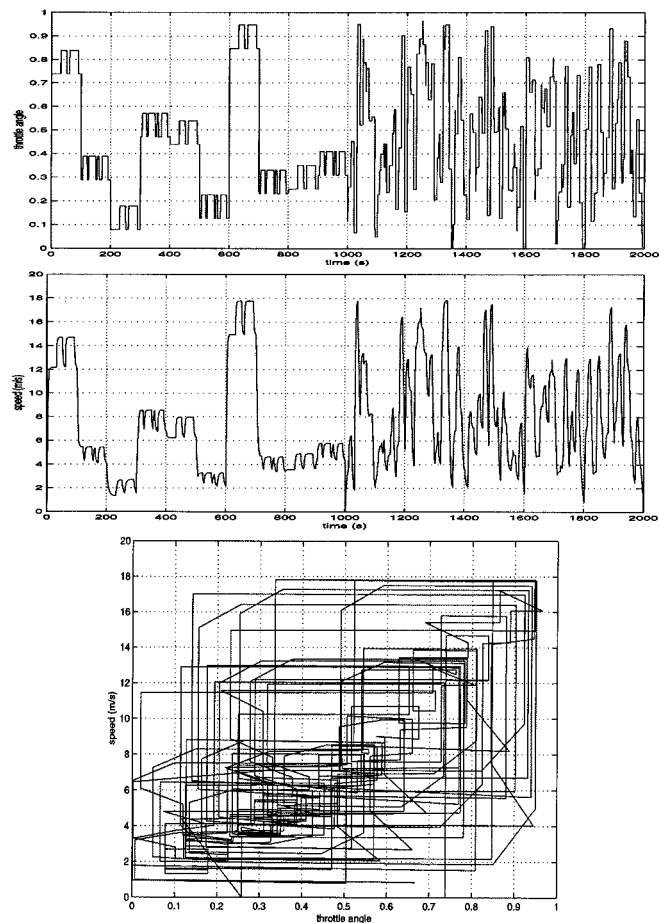

Fig. 9. Identification data sequence 2.

closeness of the local affine models to the linearization of the system. □

B. 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 system identification is a general method for improving the robustness and accuracy of the system identification algorithm when the model structure is poorly identifiable ([11], [12], [31], [32]). 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 toward some set of *a priori* desirable parameters.

It is assumed that the identification is carried out using discrete-time input/output data and a discrete-time dynamic Takagi–Sugeno fuzzy model that admits the one-step ahead predictor to be written in the form

$$\hat{y}(k \mid k-1) = g(\varphi(k-1), \theta) \tag{59}$$

where θ is a vector that contains the local model parameters $\theta_1, \cdots, \theta_N$ and $\varphi(k)$ is an information vector that may depend on the current and past inputs and outputs. Note that similar to

the example above, one may transform (at least approximately) a discrete-time model into a continuous-time state-space model, which is being assumed in previous sections of this paper. Also note that only in special cases (e.g., when all states are measured or the state space realization is chosen in a convenient way) it is natural to formulate the one-step ahead predictor (59) as a linear function of the parameter vector θ, admitting a simple least squares normal solution. In the general case, the predictor might contain a state observer that whose parameters will typically depend in a nonlinear manner on the model parameters, e.g., [21].

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

$$V_{\text{LS}}(\theta) = \sum_{k=1}^{l} \|y(k) - g(\varphi(k-1), \theta)\|^2 \qquad (60)$$

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

$$H\theta = h \qquad (61)$$
$$K\theta \le k \qquad (62)$$

where H, K, h, and k are constant matrices and vectors. When the predictor (59) is a linear function of the parameters θ, minimizing the least squares criterion (60) subject to the constraints (61) and (62) is a convex optimization problem that can be solved using quadratic programming [23]. In the general case, it is a nonlinear programming problem.

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 [10]

$$\Omega_{\text{TIK}}^{\text{approx}}(\theta) = \sum_{i=1}^{N} \sum_{j=1}^{i-1} (\theta_i - \theta_j)^T W_{ij}(\theta_i - \theta_j)\rho_{i,j} \qquad (63)$$

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 (63) is that the local model parameters θ_i and θ_j corresponding to neighboring 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 $c_i = (c_{i,1}, \cdots, c_{i,n})$ 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 problem of minimizing

$$V(\theta) = V_{\text{LS}}(\theta) + \gamma\Omega_{\text{TIK}}^{\text{approx}}(\theta) \qquad (64)$$

subject to the constraints (61) and (62). 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 [11].

Example: Longitudinal Vehicle Dynamics, Constrained Identification: Suppose we restrict the gains of the local models to be nonnegative, i.e., $B_i \ge 0$, and the poles of the local models to be nonpositive, i.e., $A_i \le 0$ and use the constrained least squared identification algorithm with data sequence 1. Equations (53)–(55) are used to convert between continuous and discrete time model parameterizations and a quadratic programming algorithm is applied to compute the estimate. The results are shown in Fig. 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 (63) without any constraints. The identification results using data sequence 1 are shown in Fig. 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 neighboring 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 [12] (see also [31]). □

The structure identification problem becomes somewhat more complicated when constraints and regularization are introduced. However, well-known criteria such as the final prediction error criterion (FPE) have been generalized to these cases [11], [18].

C. Locally Weighted Identification

The least squares objective (60) 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 [24], [25]. Individual criteria are defined for each local model

$$V_{\text{LWLS}}^i(\theta_i) = \sum_{k=1}^l \|y(k) - g_i(\varphi_i(k-1), \theta_i)\|^2 w_i(z(k-1))$$

(65)

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 $\hat{y}(k \mid k-1; i) = g_i(\varphi_i(k-1), \theta_i)$. The resulting identification problem is linear in the parameters and can be solved explicitly if the predictor g_i is linear with respect to θ_i. The implicit regularization effect of locally weighted least squares has been studied in detail in [25]. 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 with a linearly parameterized predictor. In [38] it was suggested to combine the local criteria (65) with the global criterion (60) in order to address the tradeoff between local accuracy and global performance.

Example: Longitudinal Vehicle Dynamics, Locally Weighted Identification: Parameter identification results using locally weighted least squares are shown in Fig. 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 [6] to determine local linear dynamic models ([3], [33]). 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 [26]. 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.

V. 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 is 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 multivariable 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., [13], [25]) and piecewise linear models (e.g., [2]).

APPENDIX A
PROOF OF THEOREM 1

From (13) and the property that $\sum_i w_i(x, u) = 1$ for all $(x, u) \in X \times U$ it follows that:

$$\Delta \dot{\xi}_2(t)$$
$$= \dot{\xi}_2(t) - \dot{x}_0(t) \tag{66}$$
$$= \sum_{i=1}^N f(x_i, u_i) w_i(\xi_2(t), u(t)) - f(x_0(t), u_0(t))$$
$$+ \sum_{i=1}^N (A(x_i, u_i)(\xi_2(t) - x_i) + B(x_i, u_i)(u(t) - u_i))$$
$$\times w_i(\xi_2(t), u(t)) \tag{67}$$
$$= f(\xi_2(t), u(t)) - f(x_0(t), u_0(t))$$
$$+ \sum_{i=1}^N (f(x_i, u_i) w_i(\xi_2(t), u(t)) - f(\xi_2(t), u(t)))$$
$$+ \sum_{i=1}^N (A(x_i, u_i)(\xi_2(t) - x_i) + B(x_i, u_i)(u(t) - u_i))$$
$$\times w_i(\xi_2(t), u(t)) \tag{68}$$
$$= A(x_0(t), u_0(t)) \Delta \xi_2(t) + B(x_0(t), u_0(t)) \Delta u(t)$$
$$+ \sum_{i=1}^N (f(x_i, u_i) - f(\xi_2(t), u(t))) w_i(\xi_2(t), u(t))$$
$$+ \sum_{i=1}^N (A(x_i, u_i)(\xi_2(t) - x_i) + B(x_i, u_i)(u(t) - u_i))$$
$$\times w_i(\xi_2(t), u(t)) + \text{h.o.t.} \tag{69}$$
$$= A(x_0(t), u_0(t)) \Delta \xi_2(t) + B(x_0(t), u_0(t)) \Delta u(t)$$
$$+ \varepsilon(\xi_2(t), u(t), x_0(t), u_0(t)) \tag{70}$$

where

$$
\varepsilon(\xi_2(t), u(t), x_0(t), u_0(t))
$$
$$
= \sum_{i=1}^{N}(f(x_i, u_i) - f(\xi_2(t), u(t)))w_i(\xi_2(t), u(t))
$$
$$
+ \sum_{i=1}^{N}(A(x_i, u_i)(\xi_2(t) - x_i) + B(x_i, u_i)(u(t) - u_i))
$$
$$
\times w_i(\xi_2(t), u(t)) + \text{h.o.t.} \tag{71}
$$

Application of the triangle inequality on (71) leads to

$$
\varepsilon(\xi_2(t), u(t), x_0(t), u_0(t))
$$
$$
\leq \sum_{i=1}^{N}\|f(x_i, u_i) - f(\xi_2(t), u(t))\|_2 w_i(\xi_2(t), u(t))
$$
$$
+ \sum_{i=1}^{N}\|A(x_i, u_i)(\xi_2(t) - x_i) + B(x_i, u_i)(u(t) - u_i)\|
$$
$$
\times w_i(\xi_2(t), u(t)) + \text{h.o.t.} \tag{72}
$$

For the point $(x_i, u_i) \in Z^0$ that is closest to $\xi_2, u)$ it holds that $w_i(\xi_2, u) > 0$ and for points $(x_i, u_i) \in Z^0$ that are more than a distance δ away from (ξ_2, u) it holds that $w_i(\xi_2, u) = 0$, i.e., $w_i(\xi_2, u) = 0$ if $\|(\xi_2, u) - (x_i, u_i)\|_2 > \delta$. Since there is always a $(x_i, u_i) \in Z^0$ within a distance δ from any $(\xi_2, u) \in X \times U$, only terms with $\|(\xi_2, u) - (x_i, u_i)\|_2 \leq \delta$ contributes to (72). Due to the Lipschitz property

$$
\|\varepsilon(\xi_2, u, x_i, u_i)\|_2 \leq L\delta + L\delta + O(\delta^2) = 2L\delta + O(\delta^2) \tag{73}
$$

and the result follows.

REFERENCES

[1] R. Babuska and H. B. Verbruggen, "Fuzzy modeling and identification," in *Multiple Model Approaches to Modeling and Control*, R. Murray-Smith and T. A. Johansen, Eds.   London, U.K.: Taylor Francis, 1997.

[2] S. A. Billings and W. S. G. Voon, "Piecewise linear identification of nonlinear systems," *Int. J. Control*, vol. 46, pp. 215–235, 1987.

[3] G. Bontempi, M. Birattari, and H. Bersini, "Lazy learning for modeling and control design," *Int. J. Control*, vol. 72, pp. 643–658, 1999.

[4] D. Driankov and R. Palm, Eds., *Advances in Fuzzy Control*.   Heidelberg, Germany: Physica-Verlag, 1998.

[5] D. Driankov, R. Palm, and U. Rehfuss, "A Takagi-Sugeno fuzzy gain-scheduler," in *Proc. IEEE Conf. Fuzzy Syst.*, New Orleans, LA, Sept. 1996, pp. 1053–1059.

[6] J. Fan and I. Gijbels, *Local Polynomial Modeling and Its Applications*.   London, U.K.: Chapman Hall, 1996.

[7] C. Fantuzzi and R. Rovatti, "On the approximation capabilities of the homogeneous Takagi–Sugeno model," in *Proc. IEEE Conf. Fuzzy Syst.*, New Orleans, LA, Sept. 1996, pp. 1067–1072.

[8] K. J. Hunt and T. A. Johansen, "Design and analysis of gain-scheduled local controller networks," *Int. J. Control*, vol. 66, pp. 619–651, 1997.

[9] T. A. Johansen, "Fuzzy model based control: Stability, robustness, and performance issues," *IEEE Trans. Fuzzy Syst.*, vol. 2, pp. 221–234, Aug. 1994.

[10] ——, "Robust identification of Takagi–Sugeno–Kang fuzzy models using regularization," in *Proc. IEEE Conf. Fuzzy Syst.*, New Orleans, LA, Sept. 1996, pp. 180–186.

[11] ——, "Constrained and regularized system identification," in *IFAC Symp. Syst. Identificat.*, Kitakyushu, Japan, July 1997, pp. 1467–1472.

[12] ——, "On Tikhonov regularization, bias and variance in nonlinear system identification," *Automatica*, vol. 33, pp. 441–446, 1997.

[13] T. A. Johansen and B. A. Foss, "Constructing NARMAX models using ARMAX models," *Int. J. Contr.*, vol. 58, pp. 1125–1153, 1993.

[14] T. A. Johansen, K. J. Hunt, and H. Fritz, "A software environment for gain scheduled local controller network design," *IEEE Contr. Syst. Mag.*, vol. 18, no. 2, pp. 48–60, 1998.

[15] T. A. Johansen, K. J. Hunt, P. J. Gawthrop, and H. Fritz, "Off-equilibrium linearization and design of gain scheduled control with application to vehicle speed control," *Contr. Eng. Practice*, vol. 6, pp. 167–180, 1998.

[16] D. W. Jordan and P. Smith, *Nonlinear Ordinary Differential Equations*, Oxford, U.K.: Oxford Univ. Press, 1987.

[17] T. Kailath, *Linear Systems*.   Englewood Cliffs, NJ: Prentice-Hall, 1980.

[18] J. Larsen and L. K. Hansen, "Generalization performance of regularized neural network models," in *Proc. IEEE Workshop Neural Networks Signal Processing*, Ermioni, Greece, 1994.

[19] D. A. Lawrence and W. J. Rugh, "Gain scheduling dynamic linear controllers for a nonlinear plant," *Automatica*, vol. 31, pp. 381–390, 1995.

[20] D. J. Leith and W. E. Leithead, "Analytic framework for blended multiple model systems using local linear models," *Int. J. Contr.*, vol. 72, pp. 605–619, 1999.

[21] L. Ljung, *System Identification: Theory for the User*.   Englewood Cliffs, NJ: Prentice-Hall, 1999.

[22] W. Lohmiller and J.-J. E. Slotine, "On contraction analysis for nonlinear systems," *Automatica*, vol. 6, pp. 683–696, 1998.

[23] D. G. Luenberger, *Introduction to Linear and Nonlinear Programming*, 2nd ed.   Reading, MA: Addison-Wesley, 1984.

[24] R. Murray-Smith and T. A. Johansen, "Local learning for local model networks," in *Proc. 4th Inst. Elect. Eng. Int. Conf. Artif. Neural Networks*, Cambridge, U.K., 1995, pp. 40–46.

[25] ——, "Local learning in local model networks," in *Multiple Model Approaches to Modeling and Control*, R. Murray-Smith and T. A. Johansen, Eds.   London, U.K.: Taylor Francis, 1997.

[26] R. Murray-Smith, T. A. Johansen, and R. Shorten, "On transient dynamics, off-equilibrium behavior and identification in blended multiple model structures," in *Proc. Eur. Contr. Conf.*, Karlsruhe, Germany, 1999, pp. F1020–7.

[27] K. M. Passino and S. Yurkovich, *Fuzzy Control*.   Menlo Park, CA: Addison-Wesley, 1998.

[28] R. Rovatti, "Takagi–Sugeno models as approximators in Sobolev norms—The SISO case," in *Proc. IEEE Conf. Fuzzy Systems*, New Orleans, LA, Sept. 1996, pp. 1060–1066.

[29] W. J. Rugh, "Analytical framework for gain scheduling," *IEEE Contr. Syst. Mag.*, vol. 11, pp. 79–84, Feb. 1991.

[30] R. Shorten, R. Murray-Smith, R. Bjørgan, and H. Gollee, "On the interpretation of local models in blended multiple model structures," *Int. J. Contr.*, vol. 72, pp. 620–628, 1999.

[31] J. Sjöberg and L. Ljung, "Overtraining, regularization, and searching for minimum in neural networks," in *IFAC Symp. Adaptive Syst. Contr. Signal Processing*, Grenoble, France, 1992, pp. 669–674.

[32] J. Sjöberg, T. McKelvey, and L. Ljung, "On the use of regularization in system identification," in *12th IFAC World Congress*, vol. 7, Sydney, Australia, 1993, pp. 381–386.

[33] A. Stenman, F. Gustafsson, and L. Ljung, "Just in time models for dynamical systems," in *Proc. IEEE Conf. Decision Contr.*, Kobe, Japan, 1996, pp. 1115–1120.

[34] T. Takagi and M. Sugeno, "Fuzzy identification of systems and its application to modeling and control," *IEEE Trans. Syst., Man, Cybern.*, vol. 15, pp. 116–132, 1985.

[35] H. O. Wang, K. Tanaka, and M. F. Griffin, "An approach to fuzzy control of nonlinear systems: Stability and design issues," *IEEE Trans. Fuzzy Syst.*, vol. 4, pp. 14–23, Feb. 1996.

[36] L.-X. Wang, "Stable adaptive fuzzy control of nonlinear systems," *IEEE Trans. Fuzzy Syst.*, vol. 1, pp. 146–155, 1993.

[37] L.-X. Wang and J. M. Mendel, "Fuzzy basis functions, universal approximation, and orthogonal least-squares learning," *IEEE Trans. Neural Networks*, vol. 3, pp. 807–814, 1992.

[38] J. Yen, L. Wang, and C. W. Gillespie, "Improving the interpretability of TSK fuzzy models by combining global learning and local learning," *IEEE Trans. Fuzzy Syst.*, vol. 6, pp. 530–537, Nov. 1998.

Tor A. Johansen (M'98) received the Dr.Ing. (Ph.D.) degree in electrical and computer engineering from the Norwegian University of Science and Technology, Trondheim, Norway, in 1994.

He is an Associate Professor at the Department of Engineering Cybernetics, Norwegian University of Science and Technology, Trondheim. In 1990 he was with the Norwegian Defence Research Establishment at Kjeller. During 1992 he was a Research Visitor at the University of Southern California, Los Angeles. From 1995 to 1997 he was a Research Engineer with SINTEF Electronics and Cybernetics, Trondheim. He serves as an associate editor of *Automatica* . His research interests include hybrid control, constrained control, optimization based control, multiple model methods, fuzzy control, nonlinear system identification, and industrial applications of systems engineering and real-time control.

Dr. Johansen is an associate editor of IEEE TRANSACTIONS ON FUZZY SYSTEMS. He is a member of the IEEE Technical Commitee on Fuzzy Systems and the IFAC Technical Commitee on Neural and Fuzzy Systems.

Roderick Murray-Smith received the B.Eng. and Ph.D. degrees from Strathclyde University, Glasgow, U.K., in 1990 and 1994, respectively.

From 1990 to 1997, he was with the research labs of Daimler-Benz, Berlin, Germany. He was then with the Institute of Mathematical Modeling, Technical University of Denmark, Lyngby, from 1997 to 1999. He is currently a Lecturer of computing science at the University of Glasgow, U.K. His current research interests include multiple model approaches to empirical modeling, human operator modeling and cooperative control systems, probabilistic approaches to intelligent systems, and neural networks.

Robert Shorten (M'98) received the B.E. degree in electronic engineering (first class honors) from the University College Dublin, Ireland, in 1990, and the M.Eng.Sc. and Ph.D. degrees in electrical engineering from the National University of Ireland, Mayna, in 1993 and 1996, respectively.

From 1993 to 1997, he was the holder of a Marie Curie Fellowship to conduct applied research at a Daimler-Benz Research Institute, Berlin, Germany. The results of this research were presented in the form of a Ph.D. thesis, academic publications, and in the form of an industrial patent with Daimler-Benz. From April 1996 to July 1996, he worked as a Visiting Fellow with a research group at the Center for Systems Science, Yale University, New Haven, CT, working on the area of hybrid systems. Since 1998 he has been a Lecturer in the Department of Computer Science, National University of Ireland, Maynooth.

Dr. Shorten was awarded a European Presidential Fellowship to return to Ireland in 1997. He is a member of the Institute of Engineers of Ireland.