11-2003

Kalman Filtering for Fuzzy Discrete Time Dynamic Systems

Daniel J. Simon
Cleveland State University, d.j.simon@csuohio.edu

Original Citation

D. Simon. (2003). Kalman Filtering for Fuzzy Discrete Time Dynamic Systems. Applied Soft Computing, 3(3), 191-207.

Repository Citation

Kalman filtering for fuzzy discrete time dynamic systems

Dan Simon*

Department of Electrical Engineering, Cleveland State University, Stilwell Hall Room 332, 1960 East 24th Street, Cleveland, OH 44115, USA

1. Introduction

We represent a nonlinear system with a Takagi–Sugeno (T–S) type fuzzy model. The T–S fuzzy model is based on the observation that a modeling problem can be broken up into local approximations. The local approximations are then smoothly interpolated to obtain the global model [13,14,28]. This is not unique to fuzzy systems, but is a specific example of the general approach of combining local representations to represent nonlinear dynamics [20]. Some T–S model parameter identification results can be found in [5,13]. Some studies on the universal approximation capabilities of T–S models can be found in [2,29,33–36].

Automatic control via fuzzy logic has attracted a lot of attention during the past couple of decades, from both the academic and industrial communities. Fuzzy control offers a promising alternative for the control of complex nonlinear systems. It generally offers the advantages of multi-objective control, and the realization of expert and robust control [22].

But before we can control a system we first need a good state estimate. Fuzzy state estimation is a topic that has received very little attention. There have been a few papers published recently on fuzzy observer design; however, these papers usually deal with the noise-free case. That is, fuzzy observers are designed for systems that are not affected by noise [11,18,27]. In addition, they require a common solution to a set of Ricatti equations, which may be difficult or impossible to obtain [3,6].

Kalman filters have had a long and illustrious experience in the estimation of system states. Kalman filters are attractive theoretically due to their optimality properties [1,12], and they also are easy to implement and give good results in many practical systems. State estimation is often interesting in its own right;

for instance, if someone wants to track a vehicle, or if someone wants to estimate the health of an engineering system (which can be inferred from state values). In addition, state estimation is often necessary in order to implement state feedback control systems. This paper is motivated by the practical importance of state estimation and the growing use of T–S models for the representation of nonlinear systems.

Fuzzy Kalman filtering (FKF) [10] is a recently proposed method for extending Kalman filtering to the case where the linear system parameters are fuzzy variables within intervals. FKF is based on interval Kalman filtering (IKF) [9], in which the system parameters are completely unknown within intervals. IKF can also be modified for the case where the parameters' uncertainties within their intervals are given in terms of possibility distributions [19]. IKF can also be combined with evolutionary programming to find optimal state estimates at every iteration [30]. The primary difference between the present work and IKF methods is that IKF methods deal with linear systems with unknown parameters, whereas the present paper deals with T–S models.

There is some existing literature on T–S fuzzy models that does take noise into account. For instance, [8] focuses on H_∞ disturbance rejection for T–S models. Like the previously mentioned observer results, it requires a common solution to a set of Ricatti equations. Similarly [16] presents an H_∞ controller for T–S models with time delays. The present work differs from [8,16] in that this paper focuses on H_2 disturbance rejection, and the result is a set of steady state estimators that can be found via independent solutions of an uncoupled set of Ricatti equations. The steady state estimators are then combined to obtain a global estimator.

The fuzzy separation property developed in [18] offers additional guidance in the area of state estimation. This property says that (for T–S type systems) the fuzzy controller and the fuzzy state estimator can be designed independently. This is similar to the separation property in standard non-fuzzy linear systems theory [7]. The fuzzy separation property holds only if the premise variables are independent of the state. In general the premise variables of a T–S model are functions of the state or control. However, they are sometimes independent of the state and control, as shown in the first simulation example in the present paper.

One of the important areas of fuzzy control has been the theoretical investigation of stability [3,11,17,23,25,27,28]. If stability cannot be guaranteed for a controller then practitioners will be reluctant to implement it, especially in areas that involve complicated, sensitive, or dangerous applications (such as aerospace or biomedical applications) [24]. The same can be said for fuzzy estimation. If stability cannot be guaranteed for an estimator then practitioners will be reluctant to implement it. The fuzzy estimator presented in this paper is guaranteed (under certain conditions) to be stable.

Another requirement for many control systems is optimality [31,32]. If optimality cannot be guaranteed for a control system, then practitioners will look for a better controller. Again, the same can be said for estimation. If optimality cannot be guaranteed for an estimator then practitioners will look for a better estimator. The fuzzy estimator presented in this paper is guaranteed (under certain conditions) to be optimal (in a well-defined sense).

The idea presented in this paper for fuzzy state estimation is analogous to a widely adopted approach taken for fuzzy control [11,18]. First, we represent the fuzzy system as a family of local linear state space systems. Second, we design a state estimator for each local state space model. Third, we construct a global state estimator by combining the local state estimators. This can be viewed as a decomposition principle; the design of a fuzzy control system can be decomposed into the design of a set of subsystems. Each subsystem controller is designed independently, and the individual solutions are combined to obtain a solution for the global problem [6]. Although a T–S fuzzy model can be shown to be a linear time-varying system, each of its local constituent models are time-invariant, so steady state Kalman filters can be designed for each local model. Then the local models can be combined to derive a state estimator for the global system.

The state estimation problem presented here is demonstrated on a simulated backing up truck–trailer system, a nonlinear system first presented in [21] and subsequently used by many researchers [6,17,23,26,31,32].

Section 2 presents the state estimation problem for a T–S fuzzy model. Section 3 solves the state estimation problem for each local system in the T–S model and discusses some of the local estimators' properties.

Section 4 solves the global estimation problem and explores some of the properties of the solution. Section 5 presents some simulation results, and Section 6 offers some concluding remarks. Lemma and theorem proofs are provided in Appendix A at the end of the paper.

2. Problem statement

Nonlinear systems can be approximated as locally linear systems in much the same way that nonlinear functions can be approximated as piecewise linear functions. Nonlinear systems can be represented by fuzzy linear models of the following form [4,6,8,11,23,25,28]:

$$\text{if } z_1[k] \text{ is } F_{i1} \text{ and} \ldots \text{and } z_g[k] \text{ is } F_{ig} \text{ then}$$
$$x[k+1] = A_i x[k] + B_i u[k] + G_i w[k],$$
$$y[k] = C_i x[k] + v[k] \quad (i = 1, \ldots, L) \tag{1}$$

This is referred to as a Takagi–Sugeno (T–S) fuzzy model. The z_j are premise variables, k is the time index, F_{ij} are fuzzy sets, $x[k] \in R^n$ is the state vector, $u[k] \in R^m$ is the deterministic input, $w[k]$ is the process noise, $y[k] \in R^r$ is the measured output, and $v[k]$ is the measurement noise. We assume that the process noise $w[k]$ is white with power spectral density (PSD) S_w, the measurement noise $v[k]$ is white with PSD S_v, and the process noise and measurement noise are uncorrelated. Each of the L local models of (1) is a linear time-invariant model. The fuzzy combination of these local models results in the global model:

$$x[k+1] = \sum_{i=1}^{\mathrm{L}} h_i(z[k])\{A_i x[k] + B_i u[k] + G_i w[k]\},$$
$$y[k] = \sum_{i=1}^{L} h_i(z[k]) C_i x[k] + v[k] \tag{2}$$

where the membership grades $h_i(z[k])$ are defined as:

$$h_i(z[k]) = \frac{\mu_i(z[k])}{\mu[k]} \tag{3}$$

$$\mu_i(z[k]) = \prod_{j=1}^{g} F_{ij} = (z_j[k]) \tag{4}$$

$$\mu[k] = \sum_{i=1}^{L} \mu_i(z[k]) \tag{5}$$

$$z[k] = [z_1[k] \cdots z_g[k]] \tag{6}$$

$F_{ij}(z_j[k])$ is the membership grade of $z_j[k]$ in F_{ij}. Note that $h_i(z[k]) \in [0, 1]$. From (3) and (5) we can see that:

$$\sum_{i=1}^{L} h_i(z[k]) = 1 \tag{7}$$

From (2) we can derive:

$$x[k+1] = A[k]x[k] + B[k]u[k] + G[k]w[k],$$
$$y[k] = C[k]x[k] + v[k] \tag{8}$$

where $A[k]$, $B[k]$ $C[k]$, and $G[k]$ are given as:

$$A[k] = \sum_{i=1}^{L} h_i(z[k]) A_i, \quad B[k] = \sum_{i=1}^{L} h_i(z[k]) B_i,$$
$$C[k] = \sum_{i=1}^{L} h_i(z[k]) C_i, \quad G[k] = \sum_{i=1}^{L} h_i(z[k]) G_i \tag{9}$$

In other words, the global model, which is a fuzzy combination of L local linear time-invariant models, can be represented as a time-varying model. If the premise variables $z[k]$ are functions of the state or control, then the model is nonlinear. However, if the premise variables are independent of the state and control, then the model is linear. Now we define L discrete time signals $x_i[k]$ and L discrete time signals $y_i[k]$ as:

$$x_i[k] = h_i(z[k])x[k], \qquad y_i[k] = h_i(z[k])y[k] \tag{10}$$

From these definitions and (7) it can be seen that:

$$x[k] = \sum_{i=1}^{L} x_i[k], \qquad y[k] = \sum_{i=1}^{L} y_i[k] \tag{11}$$

The dynamic behavior of the $x_i[k]$ and $y_i[k]$ signals is presented in the following lemma.

Lemma 1.

$$x_i[k+1] = A_i x_i[k] + h_i(z[k]) B_i u[k]$$
$$+ h_i(z[k]) G_i w[k],$$
$$y_i[k] = C_i x_i[k] + h_i(z[k]) v[k] \quad (i = 1, \ldots, L)$$

Proof. See Appendix A. □

3. Kalman filtering

Kalman's solution to the state estimation problem can be found in many texts, such as [1,12]. In this section we modify the Kalman filter for the system given by (12). Suppose we are given an n-dimensional linear discrete time system of the form:

$$x[k+1] = Ax[k] + h[k]Bu[k] + h[k]Gw[k],$$
$$y[k] = Cx[k] + h[k]v[k] \tag{13}$$

where the scalar $h[k] \in [0, 1]$, the process noise $w[k]$ is white with PSD S_w, the measurement noise $v[k]$ is white with PSD S_v, and the process noise and measurement noise are uncorrelated. Although the A, B, and C matrices are constant, the system is time-varying because of the time-varying scalar $h[k]$. If the premise variables are functions of the state or control, then the system is also nonlinear because $h[k]$ is a function of the state or control. The state x of the system can be estimated by the Kalman filter, which can be derived by assuming a recursive estimator of the form:

$$\hat{x}^+[k] = M[k]\hat{x}^-[k] + K[k]y[k],$$
$$\hat{x}^-[k+1] = A\hat{x}^+[k] + h[k]Bu[k] \tag{14}$$

$M[k]$ and $K[k]$ are matrices to be determined. In general, we use the "−" superscript to indicate a quantity before the measurement is taken into account, and we use the "+" superscript to indicate a quantity after the measurement is taken into account. So $\hat{x}^-[k]$ is the state estimate at time k before the measurement $y[k]$ is taken into account, and $\hat{x}^+[k]$ is the state estimate at time k after the measurement $y[k]$ is taken into account. Requiring the state estimate to be unbiased results in the constraint [12]:

$$M[k] = I - K[k]C \tag{15}$$

where I is the appropriately dimensioned identity matrix. We define the estimation error $\tilde{x}$ and its covariance P as:

$$\tilde{x} = \hat{x} - x, \qquad P = E(\tilde{x}\tilde{x}^{\mathrm{T}}) \tag{16}$$

where $E(\cdot)$ is the expected value operator. Then, if $h[k]$ is independent of x, it can be shown that the covariance is propagated as follows:

$$P^+[k] = (I - K[k]C)P^-[k](I - K[k]C)^{\mathrm{T}} + h^2[k]K[k]S_vK^{\mathrm{T}}[k] \tag{17}$$

We can find the optimal value of $K[k]$ by taking the partial derivative of the trace of $P^+[k]$ with respect to $K[k]$ and setting it equal to zero, which gives:

$$(K[k]C - I)P^-[k]C^{\mathrm{T}} + h^2[k]K[k]S_v = 0 \tag{18}$$

3.1. Minimizing the average covariance

At this point we could solve (18) for $K[k]$, but because of the time-varying $h[k]$, that would result in a time-varying filter with no steady state solution. If we want to derive a time-invariant filter we can use the fact that $h[k] \in [0, 1]$ and treat $h[k]$ as a random variable that is uniformly distributed on [0, 1]. We can take the partial derivative of the expected value of the trace of $P^+[k]$ in (17) with respect to $K[k]$. That is, we can compute the expected value of (17), where $E(h^2[k]) = 1/3$, to obtain:

$$\bar{P}^+[k] = (I - K[k]C)\bar{P}^-[k](I - K[k]C)^{\mathrm{T}} + \tfrac{1}{3}K[k]S_vK^{\mathrm{T}}[k] \tag{19}$$

where $\bar{(\cdot)}$ indicates the expected value operator. We can then find the optimal value of $K[k]$ by setting the partial derivative of the trace of $\bar{P}^+[k]$ with respect to $K[k]$ equal to zero and then solving for $K[k]$, which gives:

$$K[k] = \bar{P}^-[k]C^{\mathrm{T}}\left(C\bar{P}^-[k]C^{\mathrm{T}} + \tfrac{1}{3}S_v\right)^{-1} \tag{20}$$

We can use (13), (16), and our assumption that $E(h^2[k]) = 1/3$, to obtain:

$$\bar{P}^-[k] = A\bar{P}^+[k-1]A^{\mathrm{T}} + \tfrac{1}{3}GS_wG^{\mathrm{T}} \tag{21}$$

In order to find the steady-state solution to the Kalman filter we assume that $\bar{P}^+[k-1] = \bar{P}^+[k]$, which means we can substitute (20) for $K[k]$ in (19), and then substitute the right side of (19) for $\bar{P}^+[k-1]$ in (21). This gives the steady state solution:

$$\bar{P}^- = A(\bar{P}^- - \bar{P}^-C^{\mathrm{T}}(C\bar{P}^-C^{\mathrm{T}} + \tfrac{1}{3}S_v)^{-1}C\bar{P}^-)A^{\mathrm{T}} + \tfrac{1}{3}GS_wG^{\mathrm{T}} \tag{22}$$

This is an algebraic Ricatti equation that can be solved for $\bar{P}^-$, if (A, C) is detectable and (A, GH) is stabilizable for any H that satisfies $HH^{\mathrm{T}} = S_w$ [1]. The steady state Kalman gain K is then the time-invariant matrix given by (20), with $\bar{P}^-[k]$ in (20) replaced with (22).

The steady state covariance and gain matrices, which we will refer to as $P^{(2)}$ and $K^{(2)}$, are given as:

$$P^{(2)} = A(P^{(2)} - K^{(2)}CP^{(2)})A^{\mathrm{T}} + \tfrac{1}{3}GS_wG^{\mathrm{T}},$$
$$K^{(2)} = P^{(2)}C^{\mathrm{T}}(CP^{(2)}C^{\mathrm{T}} + \tfrac{1}{3}S_v) \tag{23}$$

The state estimate is then given by:

$$\hat{x}^+[k] = (I - K^{(2)}C)\hat{x}^-[k] + K^{(2)}y[k],$$
$$\hat{x}^-[k+1] = A\hat{x}^+[k] + h[k]Bu[k] \tag{24}$$

3.2. Minimizing the worst case covariance

In the above development we minimized the expected value of the trace of the estimation error covariance. If we want to be more conservative we can solve the problem under worst case noise assumptions. That is, we can minimize the trace of the estimation error covariance under the assumption that $h[k] = 1$ in (13). The development in the preceding subsection can then be repeated with the change that $E(h^2[k]) = 1$. That gives the standard and well known steady state Kalman filter. We will refer to these covariance and gain matrices as $P^{(\infty)}$ and $K^{(\infty)}$, which are given as:

$$P^{(\infty)} = A(P^{(\infty)} - K^{(\infty)}CP^{(\infty)})A^{\mathrm{T}} + GS_wG^{\mathrm{T}},$$
$$K^{(\infty)} = P^{(\infty)}C^{\mathrm{T}}(CP^{(\infty)}C^{\mathrm{T}} + S_v)^{-1} \tag{25}$$

The state estimate is still given by (24) (except that $K^{(2)}$ is replaced with $K^{(\infty)}$). The following interesting relationship can be shown to exist between the steady state solution given here and that given in the preceding subsection.

Lemma 2.

$$P^{(\infty)} = 3P^{(2)}, \quad K^{(\infty)} = K^{(2)}$$

Proof. See Appendix A. This lemma shows that it does not matter if we try to minimize the estimation error covariance under worst case noise assumptions, or if we try to minimize the expected value of the estimation error covariance. In either case we arrive at the same Kalman gain matrix and hence the same steady state estimator. □

The above lemma can be explained intuitively. The filter in this subsection uses S_w and S_v as the noise covariance matrices. The filter in the previous subsection is identical except that it uses $(1/3)S_w$ and $(1/3)S_v$ as the covariance matrices. But the Kalman gain is a measure of the confidence that we have in the measurement relative to the system dynamics. So if the measurement noise and process noise are both scaled by the same factor, then it stands to reason that the Kalman gain does not change. Note that this holds true for any scale factor that is applied to S_w and S_v, not just the special scale factor of 1/3 that is used in this paper.

4. A state estimator for the T–S fuzzy model

In this section we combine the Kalman filters for the local systems given in (12) to obtain a state estimator for the T–S fuzzy model given in (1). We show that our resultant state estimator is unbiased and, under certain assumptions, stable and minimum variance.

The steady state Kalman filter presented in the preceding section can be used to estimate the states of each of the L dynamic systems given in (12). This will give us L local steady state estimates as follows:

$$\begin{aligned}
P_i^-[k+1] &= A_i(P_i^-[k] - K_i[k]C_iP_i^-[k])A_i^{\mathrm{T}} \\
&\quad + G_iS_wG_i^{\mathrm{T}}, \\
K_i[k] &= P_i^-[k]C_i^{\mathrm{T}}(C_iP_i^-[k]C_i^{\mathrm{T}} + S_v)^{-1}, \\
\hat{x}_i^+[k] &= (I - K_i[k]C_i)\hat{x}_i^-[k] + K_i[k]y_i[k], \\
\hat{x}_i^-[k+1] &= A_i\hat{x}_i^+[k] + h_i[k]B_iu[k] \quad (i = 1, \ldots, L)
\end{aligned} \tag{27}$$

Note that S_w and S_v in the above equations can be replaced with $(1/3)S_w$ and $(1/3)S_v$, respectively. This will result in different P_i matrices but the same K_i matrices (see Lemma 2). Since we know from (11) that $x[k] = \sum_{i=1}^{L} x_i[k]$, we can combine the local state estimates in (27) to estimate the state of the T–S fuzzy model (1) as:

$$\hat{x}[k] = \sum_{i=1}^{L} \hat{x}_i[k] \tag{28}$$

Theorem 1. *The state estimate given by* (27) *and* (28) *is an unbiased estimate of the true state of the T–S fuzzy model given by* (1).

Proof. See Appendix A. Note that the global estimate in (28) is unbiased regardless of whether S_w and S_v are used in (27), or whether $(1/3)S_w$ and $(1/3)S_v$ are used in (27). □

Theorem 2. *Consider the A_i, C_i, G_i, and S_w matrices of the L dynamic systems in* (1). *If all of the (A_i, C_i) pairs are detectable ($i = 1, \ldots, L$), and all of the (A_i, G_iH) pairs are stabilizable for any H that satisfies $HH^{\mathrm{T}} = S_w$ ($i = 1, \ldots, L$), then the state estimator given by* (27) *and* (28) *is stable.*

Proof. See Appendix A. Note the condition given in the theorem is a sufficient but not necessary condition. Also note that if the (A_i, G_iH) pairs are stabilizable for any H that satisfies $HH^{\mathrm{T}} = S_w$, then the (A_i, G_iH) pairs are also stabilizable for any H that satisfies $HH^{\mathrm{T}} = (1/3)S_w$. It therefore follows that the global estimate given by (27) and (28) is stable regardless of whether S_w and S_v are used in (27), or whether $(1/3)S_w$ and $(1/3)S_v$ are used in (27). □

The next three lemmas are intermediate results that will be used to prove the minimum variance property of the state estimator given by (27) and (28).

Lemma 3. *Consider the ith and jth local linear systems in* (1), *where $i \neq j$. Assume that the states of the ith local linear system are uncorrelated from each other so that P_i is diagonal, and the states of the jth local linear system are uncorrelated from each other so that P_j is diagonal. Further assume that for every $m \in [1, n]$ either the mth column of C_i contains all zeros or the mth column of C_j contains all zeros. Then the Kalman gains K_i and K_j satisfy the equality*:

$$K_i^{\mathrm{T}} K_j = 0 \tag{29}$$

Proof. See Appendix A. The condition on P_i and P_j is equivalent to decoupling the states of the ith and jth local Kalman filters, respectively. This is an approximation that is sometimes used to reduce the computational expense of the Kalman filter [12]. The condition on C_i and C_j is equivalent to the mth component of the state vector directly appearing in the output of either the ith local linear system or the jth local linear system, but not in both. □

Lemma 4. *Consider the ith and jth local linear systems in* (1), *where $i \neq j$, and where the process noise covariance S_w is diagonal. Assume that the conditions of* Lemma 3 *hold. Also assume that the initial states of the ith and jth local linear systems are uncorrelated random variables, and that the local Kalman filters are initialized such that $\hat{x}_i^-[0] = E(x_i[0])$ and $\hat{x}_j^-[0] = E(x_j[0])$. Further assume that $G_i S_w G_j^{\mathrm{T}} = 0$ for all $i \neq j$. Then the estimation errors of the ith and jth local Kalman filters satisfy*:

$$E(\tilde{x}_i^{\mathrm{T}} \tilde{x}_j) = 0 \tag{30}$$

Proof. See Appendix A. The condition $G_i S_w G_j^{\mathrm{T}} = 0$ for all $i \neq j$ can be satisfied one of two ways. One way is for $S_w = 0$, which means that there is not any process noise in the system. The other way is for every column m either the mth column of G_i contains all zeros or the mth column of G_j contains all zeros. This is equivalent to stating that each component of the noise vector w appears in the state equation of either the ith local linear system or the jth local linear system, but not in both. Note that (30) can be equivalently stated as: □

$$\mathrm{Trace}[E(\tilde{x}_i \tilde{x}_j^{\mathrm{T}})] = 0 \tag{31}$$

Lemma 5. *Consider the ith and jth local linear systems in* (1), *where $i \neq j$. Assume that the conditions of* Lemma 4 *hold. Then the estimate of the ith local Kalman filter and the estimation error of the jth local Kalman filter satisfy*:

$$E(\hat{x}_i^{\mathrm{T}} \tilde{x}_j) = 0 \tag{32}$$

Proof. See Appendix A. Note that (32) can be equivalently stated as: □

$$\mathrm{Trace}[E(\hat{x}_i \tilde{x}_j^{\mathrm{T}})] = 0 \tag{33}$$

Theorem 3. *Assume that the conditions of* Lemma 5 *hold. Consider the set of all global state estimators of*

the form:

$$\hat{x}[k] = \sum_{i=1}^{L} g_i \hat{x}_i[k] \tag{34}$$

where the g_i are constants to be determined, and the local estimates $\hat{x}_i[k]$ are given in (27). *Of all estimators that are in the form of* (34), *the following global state estimator*:

$$\hat{x}[k] = \sum_{i=1}^{L} \hat{x}_i[k] \tag{35}$$

minimizes the expected value of the trace of the covariance of the global estimation error. It also minimizes the trace of the covariance of the global estimation error under worst case noise assumptions. This global state estimator is the same as that postulated in (28).

Proof. See Appendix A. The conditions for this theorem are restrictive and will not be fulfilled in most problems of practical interest. But the simulation results presented in the next section demonstrate that Kalman filters designed in this way may operate well even when these conditions are not satisfied. This is conceptually similar to the stringent stability conditions of the standard Kalman filter (i.e. complete observability, complete controllability, and exact knowledge of the system model and noise statistics). Although the stability conditions of the standard Kalman filter are rarely satisfied in applications, this does not prevent its successful implementation in many practical cases [15] ([12], p. 132). □

5. Simulation results

In this section we consider state estimation for a simple vehicle tracking problem, and also for a discrete time model of a truck–trailer system. For the vehicle tracking problem, the assumptions of this paper are satisfied. For the truck–trailer system, the assumptions are not satisfied, but the estimation results are nevertheless satisfactory.

5.1. Vehicle tracking

Consider a simple vehicle tracking problem. The east component of the vehicle position is x_1, the north component is x_2, the known commanded acceleration is u, and the known steering angle (measured counterclockwise from due east) is θ. For purposes of illustration we will assume that $0 < \theta < \pi/2$. The vehicle position is measured on the vehicle via two radio transponders, one (labeled R_e) located in the due east direction and the other (labeled R_n) located in the due north direction. However, the vehicle itself has only one transmitter/receiver pair. If the vehicle is pointing due east, then the transmission from the vehicle reaches R_e but not R_n, and the measurement is therefore equal to x_1 (plus measurement noise). If the vehicle is pointing due north, then the transmission from the vehicle reaches R_n but not R_e, and the measurement is therefore equal to x_2 (plus measurement noise). If the vehicle is pointing some direction between due east and due north, then the measurement is some combination of x_1 and x_2. With this description in mind, we can formulate the dynamic system as follows:

$$\begin{aligned} x[k+1] &= x[k] + \begin{bmatrix} T\cos\theta \\ T\sin\theta \end{bmatrix} u, \\ y[k] &= \begin{bmatrix} \cos\theta & \sin\theta \end{bmatrix} x[k] + v[k] \end{aligned} \tag{36}$$

where T is the sample time and $v[k]$ is the measurement noise. Now consider two subsystems.

The first subsystem is as follows:

$$\begin{aligned} x_1[k+1] &= x_1[k] + h_1 \begin{bmatrix} \dfrac{T}{\cos\theta} \\ 0 \end{bmatrix} u, \\ y_1[k] &= h_1 \begin{bmatrix} \dfrac{1}{\cos\theta} & 0 \end{bmatrix} x_1[k] + h_1 v[k] \end{aligned} \tag{37}$$

where $h_1 = \cos^2\theta$. The second subsystem is given as:

$$\begin{aligned} x_2[k+1] &= x_2[k] + h_2 \begin{bmatrix} 0 \\ \dfrac{T}{\sin\theta} \end{bmatrix} u, \\ y_2[k] &= h_2 \begin{bmatrix} 0 & \dfrac{1}{\sin\theta} \end{bmatrix} x_2[k] + h_2 v[k] \end{aligned} \tag{38}$$

where $h_2 = \sin^2\theta$. It can be seen that $h_1 + h_2 = 1$ and that the combination of these two subsystems in the manner given in (2) results in the dynamic system model shown in (36). These two subsystems satisfy

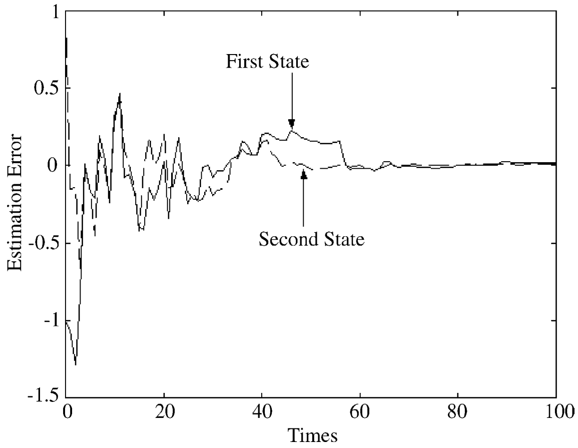


Fig. 1. Vehicle tracking estimation errors.

all the assumptions of the lemmas and theorems in Section 4 so the combined Kalman filter discussed in this paper can be used with confidence. The two local state vectors of (12) are estimated according to (27) and are then combined according to (28) to obtain the global state estimate. The system and the Kalman filter equations were simulated using Matlab with initial estimation errors of 1 and with white Gaussian unity-variance measurement noise. The estimation errors are shown in Fig. 1. It is seen from the figure that the Kalman filter works well and provides state estimates that converge to zero.

In this example the T–S system matrices are time-varying. Most T–S model formulations have constant system matrices. However, there is no explicit requirement in T–S modeling that the system matrices be constant. Also, for this example there is not really any need to use the Kalman filter proposed in this paper. A standard Kalman filter could be directly applied to the system given in (36) without the added complication of the approach proposed in this paper. However, this simple example serves to illustrate the theory. The next example may be a more realistic application of the theory.

5.2. *A truck–trailer system*

A noise-free representation of a truck–trailer system can be described as [23]:

$$\begin{aligned}
\alpha[k+1] &= \alpha[k] + \frac{VT}{l}\tan(u[k]),\\
\beta[k+1] &= \beta[k] + \frac{VT}{L}\sin(\alpha[k]),\\
N[k+1] &= N[k] + VT\cos(\alpha[k])\sin\left(\frac{\beta[k+1]+\beta[k]}{2}\right),\\
E[k+1] &= E[k] + VT\cos(\alpha[k])\cos\left(\frac{\beta[\mathrm{k}+1]+\beta[\mathrm{k}]}{2}\right)
\end{aligned} \tag{39}$$

where α is the angle of the truck (measured counterclockwise from due east), β the angle of the trailer (measured counterclockwise from due east), N the northerly position of the rear of the trailer, and E the easterly position of the rear of the trailer. l is the length of the truck, L the length of the trailer, T the sampling time, V the constant speed of backward movement of the truck, and u is the controlled steering angle (measured counterclockwise with respect to the truck orientation). The following noisy fuzzy model, adapted from [6,23], can be used to represent the above system:

$$\begin{aligned}
&\text{if } z[k] \text{ is } F_1 \text{ then } x[k+1] = A_1x[k] + B_1u[k] + G_1w[k],\\
&y[k] = C_1x[k] + v[k];\\
&\text{if } z[k] \text{ is } F_2 \text{ then } x[k+1] = A_2x[k] + B_2u[k] + G_2w[k],\\
&y[k] = C_2x[k] + v[k]
\end{aligned} \tag{40}$$

The state of the above model is comprised of α, β, and N. The premise variable $z[k]$ is given as:

$$z[k] = \beta[k] + \frac{\alpha[k]VT}{(2/L)} \tag{41}$$

The membership functions in (40) are defined as $F_1 = \{\text{about } 0\}$ and $F_2 = \{\text{about } \pm\pi\}$.

The membership grades h_1 and h_2 are therefore chosen as:

$$h_1 = \left(1 - \frac{1}{1+\exp(-3(z-\pi/2))}\right) \times \left(\frac{1}{1+\exp(-3(z+\pi/2))}\right), \qquad h_2 = 1 - h_1 \tag{42}$$

These membership grade functions are shown in Fig. 2. The A_i, B_i, C_i, and G_i matrices are given by:

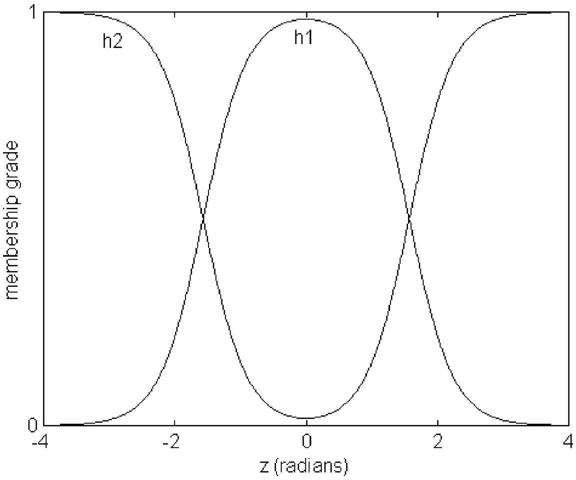


Fig. 2. Truck–trailer membership functions.

$$A_1 = \begin{bmatrix} 1 - VT/L & 0 & 0 \\ VT/L & 1 & 0 \\ (VT)^2/(2/L) & V/T & 1 \end{bmatrix},$$

$$A_2 = \begin{bmatrix} 1 - VT/L & 0 & 0 \\ VT/L & 1 & 0 \\ (VT)^2/(2/L)(\pi/100) & V/(\pi/100) & 1 \end{bmatrix},$$

$$B_1 = B_2 = \begin{bmatrix} VT/l \\ 0 \\ 0 \end{bmatrix},$$

$$C_1 = C_2 \quad \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}, \quad G_1 = G_2 \quad \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \tag{43}$$

We will use the following matrices for the process noise and measurement noise covariances:

$$S_w = \begin{bmatrix} 0.05 & 0 & 0 \\ 0 & 0.05 & 0 \\ 0 & 0 & 0.25 \end{bmatrix}^2,$$

$$S_v = \begin{bmatrix} 0.2 & 0 & 0 \\ 0 & 0.2 & 0 \\ 0 & 0 & 1 \end{bmatrix}^2 \tag{44}$$

We use the following system parameters:

$$l = 2.8\,\text{m}, \qquad L = 5.5\,\text{m},$$
$$V = -1\,\text{m/s}, \qquad T = 0.5\,\text{s} \tag{45}$$

With the above parameters, Matlab gives the algebraic Ricatti equation solutions to (25) (assuming worst case noise in the system) for P_1 and P_2 as follows:

Table 1
Average estimation and measurement errors for various initial conditions, based on 20 Monte Carlo simulations for each set of initial conditions

Initial conditions			Truck angle error (°)		Trailer angle error (°)		Trailer position error (m)	
$\alpha[0]$	$\beta[0]$	$N[0]$	Estimated	Measured	Estimated	Measured	Estimated	Measured
−45	−45	−5	0.94	2.24	0.97	2.35	0.37	1.00
−45	45	−5	1.00	2.30	0.99	2.31	0.37	1.01
45	−45	−5	0.99	2.27	0.96	2.25	0.35	0.98
45	45	−5	0.92	2.21	1.00	2.33	0.37	1.00

Table 2
Average estimation errors for various initial conditions, based on 20 Monte Carlo simulations for each set of initial conditions

Initial conditions			Truck angle error (°)		Trailer angle error (°)		Trailer position error (m)	
$\alpha[0]$	$\beta[0]$	$N[0]$	SS	Optimal	SS	Optimal	SS	Optimal
−45	−45	−5	0.94	0.99	0.97	0.95	0.37	0.37
−45	45	−5	1.00	1.02	0.99	0.94	0.37	0.35
45	−45	−5	0.99	1.06	0.96	0.97	0.35	0.36
45	45	−5	0.92	0.98	1.00	0.96	0.37	0.37

The optimal estimator requires 304 more floating point operations per time step than the steady state estimator.

$$P_1 = \begin{bmatrix} 0.01601890922659 & -0.00281321170625 & 0.00188286372510 \\ -0.00281321170625 & 0.01190698795888 & -0.01147510636123 \\ 0.00188286372510 & -0.01147510636123 & 0.30207909309150 \end{bmatrix},$$

$$P_2 = \begin{bmatrix} 0.01602068339457 & -0.00281805591335 & 0.00000600023562 \\ -0.00281805591335 & 0.01205531607052 & -0.00003764513178 \\ 0.00000600023562 & -0.00003764513178 & 0.28319575494677 \end{bmatrix} \tag{46}$$

If our objective is instead to minimize the expected value of the trace of the estimation error covariance, then Matlab gives the algebraic Ricatti equation solutions to (23) for P_1 and P_2 as follows:

$$P_1 = \begin{bmatrix} 0.00533963640886 & -0.00093773723542 & 0.00062762124170 \\ -0.00093773723542 & 0.00396899598629 & -0.00382503545374 \\ 0.00062762124170 & -0.00382503545374 & 0.10069303103050 \end{bmatrix},$$

$$P_2 = \begin{bmatrix} 0.00534022779819 & -0.00093935197112 & 0.00000200007854 \\ -0.00093935197112 & 0.00401843869017 & -0.00001254837726 \\ 0.00000200007854 & -0.00001254837726 & 0.09439858498226 \end{bmatrix} \tag{47}$$

Note that the P_i matrices in (46) are equal to three times the P_i matrices in (47), in accordance with Lemma 2. Either of the pairs of P_i matrices above lead to the following Kalman gain matrices (again, see Lemma 2):

$$K_1 = \begin{bmatrix} 0.28399097507240 & -0.03865207356457 & 0.00069474332969 \\ -0.03865207356457 & 0.22580001638107 & -0.00676706247449 \\ 0.01736858324213 & -0.16917656186220 & 0.23048144538012 \end{bmatrix},$$

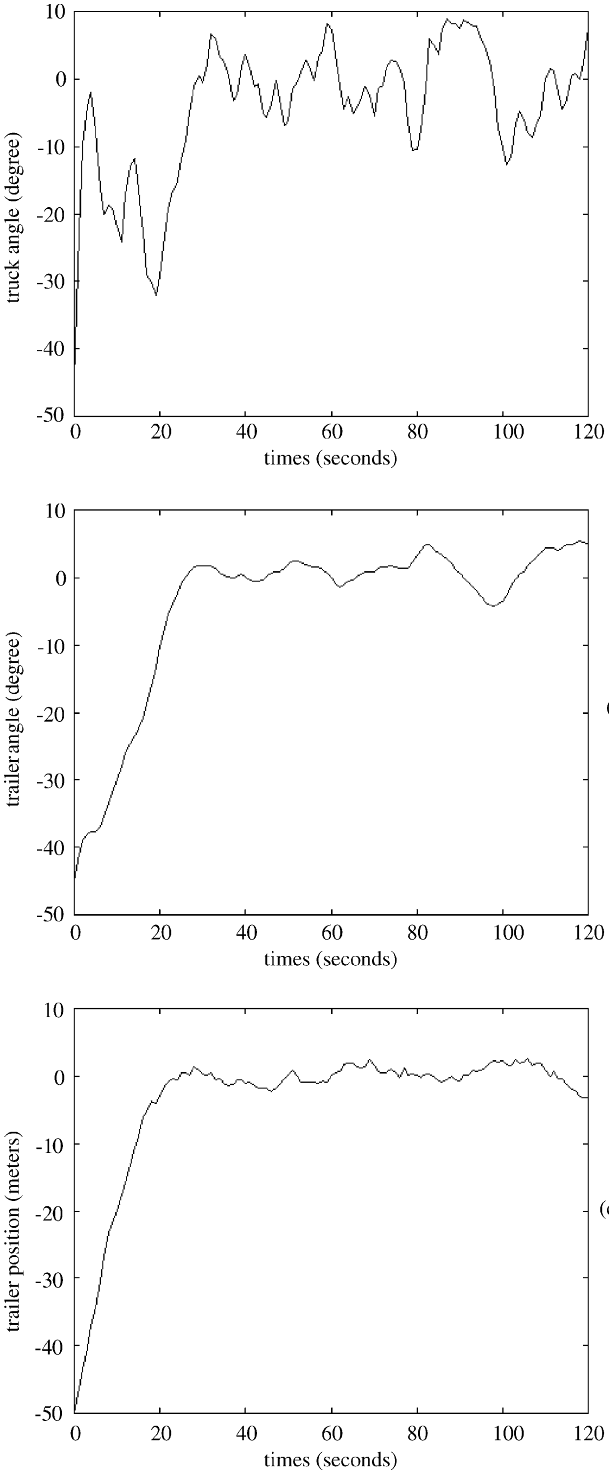
$$K_2 = \begin{bmatrix} 0.28402824492063 & -0.03875969847601 & 0.00000221079696 \\ -0.03875969847601 & 0.22948834299108 & -0.00002242327052 \\ 0.00005526992391 & -0.00056058176294 & 0.22069565958291 \end{bmatrix} \tag{48}$$

The two local state vectors of (12) are estimated according to (27) using the K_i matrices above, and are then combined according to (28) to obtain the global state estimate. The nonlinear system was simulated using Matlab, starting with various poor initial conditions. The control $u[k]$ that was used was based on the fuzzy infinite horizon optimal control described in [31].

Table 1 shows the average estimation error and measurement error that resulted with various initial conditions. It can be seen that the fuzzy Kalman filter improved the state estimate by a significant amount for all of the initial conditions that were considered. However, since the Kalman filter and optimal controller are based on a linearization of the nonlinear system, neither the filter nor the controller will work well if the initial conditions are too extreme.

Note that the restrictive requirements for stability and optimality are not satisfied in this simple example. For example, C_1 and C_2 clearly do not satisfy Lemma 3, and G_1 and G_2 do not satisfy Lemma 4. In spite of this, the Kalman filter still works well.

It was noted at the beginning of Section 3.1 that a time-varying Kalman filter could provide a theoretically exact state estimation solution to the T–S fuzzy model approximation of a nonlinear system. The time-varying filter was implemented for the T–S model described in this section. Table 2 shows the average estimation error that resulted from the use of the optimal time-varying Kalman filter, and the steady state Kalman filter. It can be seen that the performance of the two filters is nearly identical. The advantage of the steady state filter lies in its computational expense. The steady state filter in this example requires 304 fewer coating point operations per iteration, and



(a) Truck Angle

(b) Trailer Angle

(c) Trailer Position

Fig. 3. Typical simulation results using infinite time optimal control.

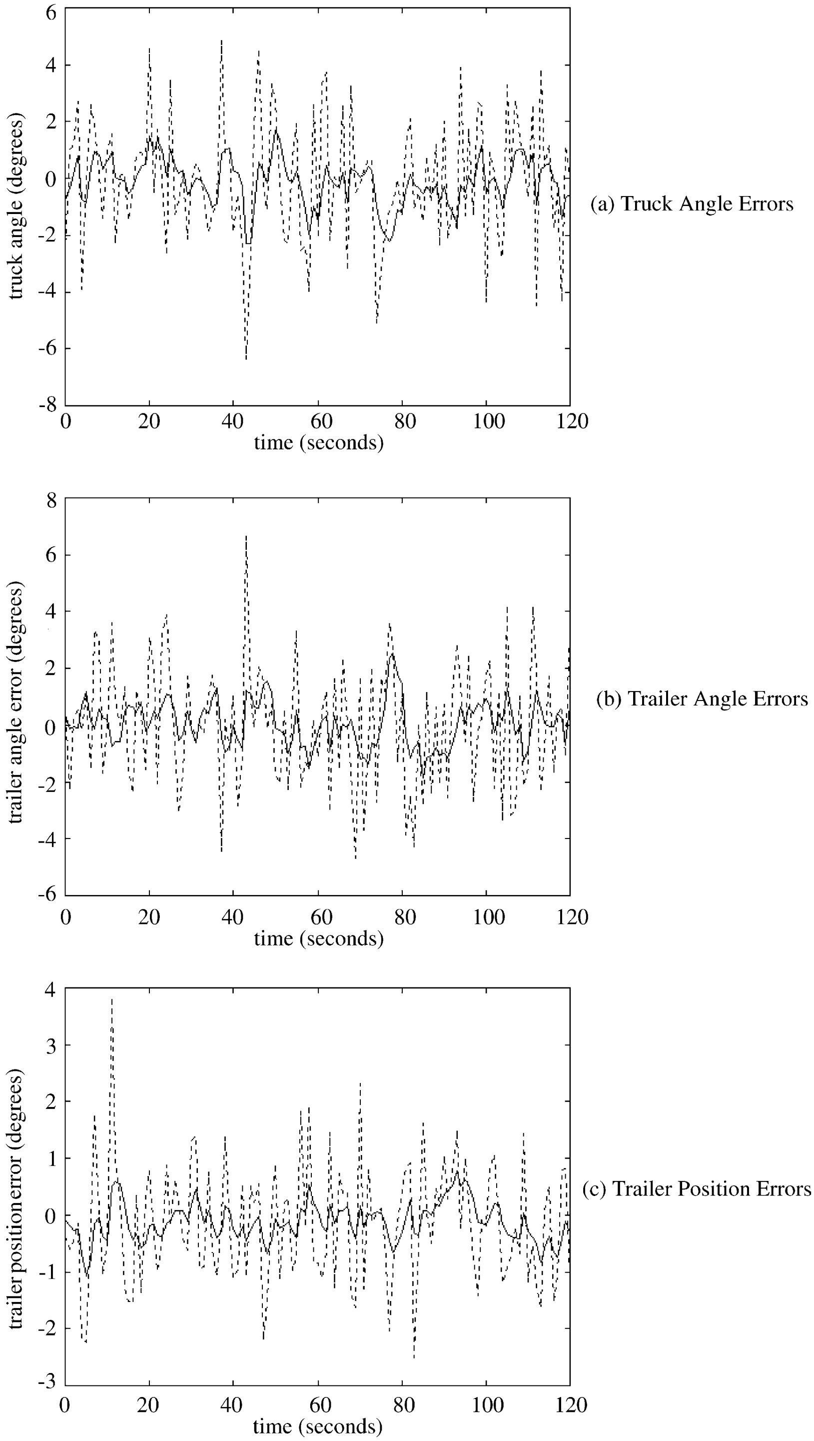


Fig. 4. Typical errors. The dotted lines are measurement errors and the solid lines are estimation errors.

this is only a third order system. For higher order systems the difference would be more extreme since the computational effort of the time-varying Kalman filter is on the order of n^3, where n is the number of states. This could be a significant consideration for a real time implementation.

Fig. 3 shows the truck angle, trailer angle, and trailer position for a typical simulation with the initial conditions $\alpha[0] = -45^\circ$, $\beta[0] = -45^\circ$, and $N[0] = -5\,\text{m}$. Fig. 4 shows close-ups of the error of the measurement and estimation of the truck angle, trailer angle, and trailer position. The Matlab m-files that were used to produce these simulation results can be downloaded from the World Wide Web page http://academic.csuohio.edu/simond/kalmanfuzzy/.

6. Conclusion

State estimation is often required for effective control. In addition, it is often interesting for its own sake. With this motivation, a linear state estimator has been presented for noisy T–S type fuzzy systems, which can approximate noisy nonlinear systems. The state estimator is based on Kalman filter theory. Steady state Kalman filters are designed for each of the local systems of the T–S model, and the local filters are then combined to obtain the global estimator. We showed that the estimator is unbiased. We also showed, under certain conditions, that the estimator is stable and minimum variance. The estimator not only minimizes the expected value of the estimation variance, but it also minimizes the estimation variance under worst case noise assumptions. Simulation results have been presented for a nonlinear system showing the effectiveness of this scheme for state estimation.

It was shown that a standard time-varying Kalman filter can be used to directly estimate the states of a T–S system. However, this results in a high level of computational effort due to the time-varying characteristic of the filter and the resultant need for matrix inversion at each time step. The simulation results in Section 5.2 showed that the state estimator in this paper provides performance that is comparable to a time-varying Kalman filter, but with much less computational effort.

The theoretical results of this paper are restricted to T–S models where the premise variables are independent of the state variables. This results in a linear time-varying system, in which case a standard time-varying Kalman filter can be used for state estimation. However, in many implementations the computational cost of a time-varying Kalman filter will be prohibitive. The new T–S Kalman filter presented in this paper shows how to approximate the time-varying Kalman filter with a time-varying linear combination of steady state Kalman filters. This achieves state estimation performance on par with the time-varying filter while drastically reducing the computational effort. The simulations results presented in this paper showed that the use of the T–S Kalman filter resulted in an insignificant loss in estimation performance (relative to the time-varying Kalman filter). But the T–S Kalman filter showed a computational savings of 304 floating point operations per time step for a third order filter.

In many practical T–S models (including one of the examples presented in this paper) the premise variables are functions of the state variables. The initial simulation results presented in this paper indicate that the T–S Kalman filter operates well even when the required theoretical conditions are not satisfied. This indicates that the T–S Kalman filter may have some robustness properties that could be investigated theoretically. Further research is needed to explore the effect that the required conditions have on the formulation of the T–S Kalman filter, and on its stability and optimality properties.

The focus of this paper has been on discrete time systems because of their prevalence in real world applications. It is expected that similar results could be shown for continuous time systems. This would be academically fruitful, although the practical benefits of such an extension may not be readily apparent.

Appendix A

In this Appendix A we provide proofs for the various lemmas and theorems that are presented in the paper.

Proof of Lemma 1. We approach this proof by showing that (12) implies (8), which in turn implies that (12) does indeed describe the dynamic behavior of x_i

and y_i. From (10), (11), and (12) we obtain:

$$\begin{aligned}x[k+1] &= \sum_{i=1}^{L} x_i[k+1] \\ &= \sum_{i=1}^{L} \{A_i x_i[k] + h_i(z[k]) B_i u[k] \\ &\quad + h_i(z[k]) G_i w[k]\} \\ &= \sum_{i=1}^{L} A_i h_i(z[k]) x[k] + \sum_{i=1}^{L} h_i(z[k]) B_i u[k] \\ &\quad + \sum_{i=1}^{L} h_i(z[k]) G_i w[k] \end{aligned} \tag{49}$$

Now we can use (9) to obtain:

$$x[k+1] = A[k]x[k] + B[k]u[k] + G[k]w[k] \tag{50}$$

where the $A[k]$, $B[k]$, and $G[k]$ matrices are given in (9). This is exactly the dynamic behavior of the global system as described in (8), which shows that (8) does indeed describe the dynamic behavior of x_i. A similar method can be used to show that the premises of the lemma also result in:

$$y[k] = C[k]x[k] + v[k] \tag{51}$$

which completes the proof. □

Proof of Lemma 2. We will assume that $P^{(\infty)} = 3P^{(2)}$. We will then show that this leads to a consistent equation, which will therefore verify our assumption.

If $P^{(\infty)} = 3P^{(2)}$, then from (25) we obtain:

$$\begin{aligned}P^{(\infty)} &= A(3P^{(2)} - 3P^{(2)}C^{\mathrm{T}}(C3P^{(2)}C^{\mathrm{T}} + S_v)^{-1} \\ &\quad \times C3P^{(2)})A^{\mathrm{T}} + GS_wG^{\mathrm{T}} \\ &= 3[A(P^{(2)} - P^{(2)}C^{\mathrm{T}}(CP^{(2)}C^{\mathrm{T}} + \tfrac{1}{3}S_v)^{-1}CP^{(2)})A^{\mathrm{T}} \\ &\quad + G\tfrac{1}{3}S_wG^{\mathrm{T}}] \\ &= 3P^{(2)}\end{aligned} \tag{52}$$

where the last equality comes from (23) and verifies our original assumption. Now since $P^{(\infty)} = 3P^{(2)}$, then (25) tells us that:

$$\begin{aligned}K^{(\infty)} &= P^{(\infty)}C^{\mathrm{T}}(CP^{(\infty)}C^{\mathrm{T}} + S_v)^{-1} \\ &= P^{(2)}C^{\mathrm{T}}(CP^{(2)}C^{\mathrm{T}} + \tfrac{1}{3}S_v)^{-1} = K^{(2)}\end{aligned} \tag{53}$$

where the last equality follows from (23). □

Theorem 1 Proof. In the following development we drop the time index for ease of notation. We can use (11) and (28) to derive the error in the state estimate as:

$$\tilde{x} = \hat{x} - x = \sum_{i=1}^{L} \hat{x}_i - \sum_{i=1}^{L} x_i = \sum_{i=1}^{L} \hat{x}_i \tag{54}$$

Therefore, knowing from Section 3 that $E(\tilde{x}_i) = 0$, we obtain:

$$E(\tilde{x}) = 0 \tag{55}$$

Theorem 2 Proof. Assume that the premise of the theorem is true. That is, given the A_i, C_i, G_i, and S_w matrices of the L dynamic systems in (1), all of the (A_i, C_i) pairs are detectable ($i = 1, \ldots, L$), and all of the (A_i, G_iH) pairs are stabilizable for any H that satisfies $HH^{\mathrm{T}} = S_w$ ($i = 1, \ldots, L$). Then we know that each of the L local estimators in (27) are stable [1]. So if the estimators are unforced (i.e. $y[k] = 0$ for all k) then for any initial state estimate $\hat{x}_i[0]$ we have:

$$\lim_{k\to\infty} \hat{x}_i[k] = 0 \quad (i = 1, \ldots, L) \tag{56}$$

So if the state estimator of (28) is unforced then:

$$\lim_{k\to\infty} \hat{x}[k] = \sum_{i=1}^{L} \lim_{k\to\infty} \hat{x}_i[k] = 0$$

This shows that the state estimator of (28) is stable.

Proof of Lemma 3. From (27) we have the Kalman gain of the ith local linear system as:

$$K_i = P_iC_i^{\mathrm{T}}(C_iP_iC_i^{\mathrm{T}} + S_v)^{-1} \tag{57}$$

Therefore we obtain:

$$\begin{aligned}K_i^{\mathrm{T}}K_j &= (C_iP_iC_i^{\mathrm{T}} + S_v)^{-1}C_iP_iP_jC_j^{\mathrm{T}} \\ &\quad \times (C_jP_jC_j^{\mathrm{T}} + S_v)^{-1}\end{aligned} \tag{58}$$

when $i \neq j$. If the states of the ith local linear system are uncorrelated from each other so that P_i is diagonal, and the states of the jth local linear system are uncorrelated from each other so that P_j is diagonal, we can write:

$$\begin{aligned}P_i &= \mathrm{diag}(p_{i1}, \ldots, p_{in}), \\ P_j &= \mathrm{diag}(p_{j1}, \ldots, p_{jn})\end{aligned} \tag{59}$$

So the middle expression on the right hand side of (58) can be written as:

$$C_i P_i P_j C_j^{\mathrm{T}} = \sum_{m=1}^{n} p_{im} p_{jm} C_{im} (C_{jm})^{\mathrm{T}} \tag{60}$$

where C_{im} and C_{jm} are the mth columns of C_i and C_j, respectively. But if, for every column $m \in [1, n]$, either the mth column of C_i contains all zeros or the mth column of C_j contains all zeros, then for every m either $C_{im} = 0$ or $C_{jm} = 0$. Therefore,

$$C_i P_i P_j C_i^{\mathrm{T}} = 0 \tag{61}$$

which, when substituted into (58) gives:

$$K_i^{\mathrm{T}} K_j = 0$$

□

Proof of Lemma 4. The $\hat{x}$ and $\tilde{x}$ quantities in this proof are taken before the measurement is processed, but the "−" superscript will be omitted for ease of notation. The estimation error at the $(k+1)$st time step of the ith local Kalman filter is given by:

$$\tilde{x}_i[k+1] = \hat{x}_i[k+1] - x_i[k+1] \tag{62}$$

This can be related to variables at the kth time step by using (12) and (27) to obtain:

$$\begin{aligned} \tilde{x}_i[k+1] = {} & A_i \hat{x}_i[k] + A_i K_i (C_i x_i[k] + v[k] - C_i \hat{x}_i[k]) \\ & + h_i[k] B_i u[k] - A_i x_i[k] - h_i[k] B_i u[k] \\ & - G_i w[k] \end{aligned} \tag{63}$$

From this equation, we can use the fact that $\hat{x}_i[k]$ and $x_i[k]$ are both uncorrelated with $w[k]$ and $v[k]$, and $E(w[k]) = E(v[k]) = 0$, to obtain:

$$\begin{aligned} & E(\tilde{x}_i^{\mathrm{T}}[k+1]\tilde{x}_j[k+1]) \\ & \quad = E(\tilde{x}_i^{\mathrm{T}}[k](A_i - K_i C_i)^{\mathrm{T}} (A_i - K_j C_j) \tilde{x}_j[k]) \\ & \qquad + E(v^{\mathrm{T}}[k] K_i^{\mathrm{T}} K_j v[k]) + G_i S_w G_j^{\mathrm{T}} \end{aligned} \tag{64}$$

Now we will show via induction that $E(\tilde{x}_i^{\mathrm{T}}[k]\tilde{x}_j[k]) = 0$ for all k. The conditions of Lemma 3 show that $E(v^{\mathrm{T}}[k] K_i^{\mathrm{T}} K_j v[k]) = 0$, and the conditions of this present lemma show that $G_i S_w G_j^{\mathrm{T}} = 0$. We therefore obtain:

$$\begin{aligned} E(\tilde{x}_i^{\mathrm{T}}[1]\tilde{x}_j[1]) = {} & E(\tilde{x}_i^{\mathrm{T}}[0](A_i - K_i C_i)^{\mathrm{T}} \\ & \times (A_j - K_j C_j) \tilde{x}_j[0]) \end{aligned} \tag{65}$$

But since (from the premises of the lemma) $\tilde{x}_i[0]$ and $\tilde{x}_j[0]$ are uncorrelated, and $E(\tilde{x}_i[0]) = E(\tilde{x}_j[0]) = 0$, we obtain:

$$\begin{aligned} & E(\tilde{x}_i^{\mathrm{T}}[0](A_i - K_i C_i)^{\mathrm{T}} (A_j - K_j C_j) \tilde{x}_j[0]) \\ & \quad = E(\tilde{x}_i^{\mathrm{T}}[0])(A_i - K_i C_i)^{\mathrm{T}} (A_j - K_j C_j) \\ & \qquad \times E(\tilde{x}_j[0]) = 0 \end{aligned} \tag{66}$$

Therefore (65) becomes:

$$E(\tilde{x}_i^{\mathrm{T}}[1]\tilde{x}_j[1]) = 0 \tag{67}$$

We conclude by induction that $E(\tilde{x}_i^{\mathrm{T}}[k]\tilde{x}_j[k]) = 0$ for all k. □

Proof of Lemma 5. The $\hat{x}$ and $\tilde{x}$ quantities in this proof are taken before the measurement is processed, but the "−" superscript will be omitted for ease of notation. From (12), (16), and (27) we obtain:

$$\begin{aligned} \hat{x}_i[k+1] = {} & A_i \hat{x}_i[k] + A_i K_i (C_i x_i[k] \\ & + v[k] - C_i \hat{x}_i[k]) + h_i[k] B_i u[k], \\ \tilde{x}_j[k+1] = {} & A_j \hat{x}_j[k] + A_j K_j (C_j x_j[k] + v[k] \\ & - C_j \hat{x}_j[k]) + h_j[k] B_j u[k] - A_j x_j[k] \\ & - h_j[k] B_j u[k] - G_j w[k] \end{aligned} \tag{68}$$

From this we can use the fact that $\hat{x}_i[k]$ and $x_i[k]$ are both uncorrelated with $w[k]$ and $v[k]$, and $E(w[k]) = E(v[k]) = 0$, to obtain:

$$\begin{aligned} & E(\hat{x}_i^{\mathrm{T}}[k+1]\tilde{x}_j[k+1]) \\ & \quad = E(\hat{x}_i^{\mathrm{T}}[k] A_i^{\mathrm{T}} (K_j C_j - A_j) \tilde{x}_j[k]) \\ & \qquad + E(\tilde{x}_i^{\mathrm{T}}[k] C_i^{\mathrm{T}} K_i^{\mathrm{T}} (K_j C_j - A_j) \tilde{x}_j[k]) \\ & \qquad + E(v[k] K_i^{\mathrm{T}} K_j v[k]) \end{aligned} \tag{69}$$

We can use Lemma 4 to write the second term on the right side of the above equation as:

$$\begin{aligned} & E(\tilde{x}_i^{\mathrm{T}}[k] C_i^{\mathrm{T}} K_i^{\mathrm{T}} (K_j C_j - A_j) \tilde{x}_j[k]) \\ & \quad = E(\tilde{x}_i^{\mathrm{T}}[k]) C_i^{\mathrm{T}} K_i^{\mathrm{T}} (K_j C_j - A_j) E(\tilde{x}_j[k]) = 0 \end{aligned} \tag{70}$$

We can use Lemma 3 to write the third term on the right side of (69) as:

$$E(v[k] K_i^{\mathrm{T}} K_j v[k]) = 0 \tag{71}$$

So (69) simplifies to:

$$E(\hat{x}_i^{\mathrm{T}}[k+1]\tilde{x}_j[k+1]) \\ = E(\hat{x}_i^{\mathrm{T}}[k]A_i^{\mathrm{T}}(K_jC_j - A_j)\tilde{x}_j[k]) \tag{72}$$

Now we know that $\hat{x}_i[0]$ and $\tilde{x}_j[0]$ are uncorrelated. So substituting $k = 1$ into the above equation results in:

$$\begin{aligned} E(\hat{x}_i^{\mathrm{T}}[1]\tilde{x}_j[1]) &= E(\hat{x}_i^{\mathrm{T}}[0]A_i^{\mathrm{T}}(K_jC_j - A_j)\tilde{x}_j[0]) \\ &= \hat{x}_i^{\mathrm{T}}[0]A_i^{\mathrm{T}}(K_jC_j - A_j)E(\tilde{x}_j[0]) = 0 \end{aligned} \tag{73}$$

$$E(\hat{x}_i^{\mathrm{T}}[1]\tilde{x}_j[1]) = \hat{x}_i^{\mathrm{T}}[0]A_i^{\mathrm{T}}(K_jC_j - A_j)E(\tilde{x}_j[0]) = 0 \tag{74}$$

We conclude by induction that $E(\hat{x}_i^{\mathrm{T}}[k]\tilde{x}_j[k]) = 0$ for all k. □

Theorem 3 Proof. In this proof we omit the time index k for ease of notation. From (11) we can write:

$$x = \chi 1_L \tag{75}$$

where 1_L is the $L \times 1$ vector containing all 1s, and χ is the $n \times L$ matrix given by:

$$\chi = [x_1 \cdots x_L] \tag{76}$$

If the global state estimate is formed as a linear combination of the local state estimates (27) then we can write:

$$\hat{x}[k] = \sum_{i=1}^{L} g_i \hat{x}_i[k] = \hat{\chi} g \tag{77}$$

where $\hat{\chi}$ is defined in an analogous manner to χ, and g is the $L \times 1$ vector consisting of the g_i constants, which are yet to be determined. Then we can write the global estimation error as:

$$\tilde{x} = \hat{x} - x = \hat{\chi} g - \chi 1_L \tag{78}$$

The covariance of the estimation error can be written as:

$$\begin{aligned} P &= E(\tilde{x}\tilde{x}^{\mathrm{T}}) \\ &= E(\hat{\chi} g g^{\mathrm{T}} \hat{\chi}^{\mathrm{T}} - \hat{\chi} g 1_L^{\mathrm{T}} \chi^{\mathrm{T}} - \chi 1_L g^{\mathrm{T}} \hat{\chi}^{\mathrm{T}} + \chi 1_L 1_L^{\mathrm{T}} \chi^{\mathrm{T}}) \end{aligned} \tag{79}$$

From the above expression we can write the trace of the error covariance as:

$$\begin{aligned} \mathrm{Trace}(P) = E(&g^{\mathrm{T}} \hat{\chi}^{\mathrm{T}} \hat{\chi} g - 1_L^{\mathrm{T}} \chi^{\mathrm{T}} \hat{\chi} g - g^{\mathrm{T}} \hat{\chi}^{\mathrm{T}} \chi 1_L \\ &+ 1_L^{\mathrm{T}} \chi^{\mathrm{T}} \chi 1_L) \end{aligned} \tag{80}$$

To minimize the trace of P with respect to g we compute the partial derivative of the above equation with respect to g, which gives:

$$\begin{aligned} \frac{\partial \mathrm{Trace}(P)}{\partial g} &= 2E(\hat{\chi}^{\mathrm{T}} \hat{X}) g - 2E(\hat{\chi}^{\mathrm{T}} \chi) 1_L \\ &= 2E(\hat{\chi}^{\mathrm{T}} \hat{\chi}) g - 2E(\hat{\chi}^{\mathrm{T}} \hat{\chi}) 1_L + 2E(\hat{\chi}^{\mathrm{T}} \tilde{\chi}) 1_L \end{aligned} \tag{81}$$

where $\tilde{\chi}$ is defined analogously to χ. But the last term in the above equation can be written as:

$$E(\hat{\chi}^{\mathrm{T}} \tilde{\chi}) = \begin{bmatrix} E(\hat{x}_1^{\mathrm{T}} \tilde{x}_1) & \cdots & E(\hat{x}_1^{\mathrm{T}} \tilde{x}_L) \\ \vdots & & \vdots \\ E(\hat{x}_L^{\mathrm{T}} \tilde{x}_1) & \cdots & E(\hat{x}_L^{\mathrm{T}} \tilde{x}_L) \end{bmatrix} \tag{82}$$

We know from standard Kalman filtering theory that $E(\hat{x}_i^{\mathrm{T}} \tilde{x}_i) = 0$ [12], so the diagonal terms in the above matrix are zero. We know from Lemma 5 that $E(\hat{x}_i^{\mathrm{T}} \tilde{x}_j) = 0$ for $i \neq j$, therefore the off diagonal terms in the above matrix are also zero. So (81) can be simplified to:

$$\frac{\partial \mathrm{Trace}(P)}{\partial g} = 2E(\hat{\chi}^{\mathrm{T}} \hat{\chi}) g - 2E(\hat{\chi}^{\mathrm{T}} \hat{\chi}) 1_L \tag{83}$$

We need to set the above partial derivative equal to zero to minimize the trace of the error covariance. This results in the solution $g = 1_L$ and concludes the proof.

References

[1] B. Anderson, J. Moore, Optimal Filtering, Prentice-Hall, Englewood Cliffs, New Jersey, 1979.

[2] J. Buckley, Sugeno type controllers are universal controllers, Fuzzy Sets Syst. 53 (1993) 299–303.

[3] S. Cao, N. Rees, G. Feng, Analysis and design of fuzzy control systems using dynamic fuzzy global models, Fuzzy Sets Syst. 75 (1995) 47–62.

[4] S. Cao, N. Rees, G. Feng, H-infinity control of nonlinear continuous time systems based on dynamical fuzzy models, Int. J. Syst. Sci. 27 (1996) 821–830.

[5] S. Cao, N. Rees, G. Feng, Analysis and design for a class of complex control systems. Part I: Fuzzy modelling and identification, Automatica 33 (6) (1997) 1017–1028.

[6] S. Cao, N. Rees, G. Feng, Analysis and design for a class of complex control systems. Part II: Fuzzy controller design, Automatica 33 (6) (1997) 1029–1039.
[7] C. Chen, Linear System Theory and Design, Oxford University Press, New York, 1999.
[8] B. Chen, C. Tseng, H. Uang, Mixed H_2/H_∞ fuzzy output feedback control design for nonlinear dynamic systems: an LMI approach, IEEE Trans. Fuzzy Syst. 8 (3) (2000) 249–265.
[9] G. Chen, J. Wang, L. Shieh, Interval Kalman filtering, IEEE Trans. Aerosp. Electron. Syst. 33 (1997) 250–259.
[10] G. Chen, Q. Xie, L. Shieh, Fuzzy Kalman filtering, Inf. Sci. 109 (1998) 197–209.
[11] G. Feng, S. Cao, N. Rees, C. Chak, Design of fuzzy control systems with guaranteed stability, Fuzzy Sets Syst. 85 (1997) 1–10.
[12] A. Gelb, Applied Optimal Estimation, MIT Press, Cambridge, Massachusetts, 1974.
[13] T. Johansen, Fuzzy model based control: stability, robustness, and performance issues, IEEE Trans. Fuzzy Syst. 2 (3) (1994.) 221–234.
[14] T. Johansen, R. Shorten, R. Murray-Smith, On the interpretation and identification of dynamic Takagi-Sugeno fuzzy models, IEEE Trans. Fuzzy Syst. 8 (3) (2000) 297–313.
[15] R. Kalman, R. Bucy, New results in linear filtering and prediction, J. Basic Eng. (ASME) 83D (1961) 95–108.
[16] K. Lee, J. Kim, E. Jeung, H. Park, Output feedback robust H^∞ control of uncertain fuzzy dynamic systems with time-varying delay, IEEE Trans. Fuzzy Syst. 8 (2000) 657–664.
[17] J. Lo, Y. Chen, Stability issues on Takagi-Sugeno fuzzy model–parametric approach, IEEE Trans. Fuzzy Syst. 7 (5) (1999) 597–607.
[18] X. Ma, Z. Sun, Y. He, Analysis and design of fuzzy controller and fuzzy observer, IEEE Trans. Fuzzy Syst. 6 (1) (1998) 41–51.
[19] M. Oussalah, J. DeSchutter, Possibilistic Kalman filtering for radar 2D tracking, Inf. Sci. 130 (2000) 85–107.
[20] R. Shorten, R. Murray-Smith, R. Bjorgan, H. Gollee, On the interpretation of local models in blended multiple model structures, Int. J. Control 72 (7–8) (1999) 620–628.
[21] T. Takagi, M. Sugeno, Fuzzy identification of systems and its applications to modeling and control, IEEE Trans. Syst. Man Cybernetics SMC-15 (1) (1985) 116–132.
[22] K. Tanaka, M. Sugeno, Stability analysis and design of fuzzy control systems, Fuzzy Sets Syst. 45 (1992) 135–156.
[23] K. Tanaka, M. Sano, A robust stabilization problem of fuzzy control systems and its application to backing up control of a truck-trailer, IEEE Trans. Fuzzy Syst. 2 (2) (1994) 119–134.
[24] K. Tanaka, M. Sano, Trajectory stabilization of a model car via fuzzy control, Fuzzy Sets Syst. 70 (1995) 155–170.
[25] K. Tanaka, T. Ikeda, H. Wang, Robust stabilization of a class of uncertain nonlinear systems via fuzzy control: quadratic stabilizability, H^∞ control theory, and linear matrix inequalities, IEEE Trans. Fuzzy Syst. 4 (1) (1996) 1–13.
[26] K. Tanaka, T. Kosaki, Design of a stable fuzzy controller for an articulated vehicle, IEEE Trans. Syst. Man Cybernetics Part B: Cybernetics 27 (3) (1997) 552–558.
[27] K. Tanaka, T. Ikeda, H. Wang, Fuzzy regulators and fuzzy observers: relaxed stability conditions and LMI-based designs, IEEE Trans. Fuzzy Syst. 6 (2) (1998) 250–265.
[28] H. Wang, K. Tanaka, M. Griffin, An approach to fuzzy control of nonlinear systems: stability and design issues, IEEE Trans. Fuzzy Syst. 4 (1) (1996) 14–23.
[29] L. Wang, Universal approximation by hierarchical fuzzy systems, Fuzzy Sets Syst. 93 (1998) 223–230.
[30] Z. Weng, G. Chen, L. Shieh, J. Larsson, Evolutionary programming Kalman filter, Inf. Sci. 129 (2000) 197–210.
[31] S. Wu, C. Lin, Optimal fuzzy controller design: local concept approach, IEEE Trans. Fuzzy Syst. 8 (2) (2000) 171–185.
[32] S. Wu, C. Lin, Optimal fuzzy controller design in continuous fuzzy system: global concept approach, IEEE Trans. Fuzzy Syst. 8 (6) (2000) 713–729.
[33] H. Ying, General SISO Takagi-Sugeno fuzzy systems with linear rule consequent are universal approximators, IEEE Trans. Fuzzy Syst. 6 (1998) 582–587.
[34] H. Ying, General Takagi-Sugeno fuzzy systems with simplified linear rule consequent are universal controllers, J. Inf. Sci. 108 (1998) 91–107.
[35] H. Ying, Sufficient conditions on uniform approximation of multivariate functions by general Takagi-Sugeno fuzzy systems with linear rule consequent, IEEE Trans. Syst. Man Cybernetics 28 (1998) 515–520.
[36] K. Zeng, N. Zhang, W. Xu, A comparative study on sufficient conditions for Takagi-Sugeno fuzzy systems as universal approximators, IEEE Trans. Fuzzy Syst. 8 (6) (2000) 773–780.