STATE ESTIMATION OF NONLINEAR DISCRETE-TIME SYSTEMS BASED ON THE DECOUPLED MULTIPLE MODEL APPROACH

Rodolfo Orjuela, Benoît Marx, José Ragot and Didier Maquin
Centre de Recherche en Automatique de Nancy, UMR 7039, Nancy-Université, CNRS
2, Avenue de la Forêtr de Haye, 54 516 Vandœuvre-lès-Nancy, France
{rodolfo.orjuela, benoit.marx, jose.ragot, didier.maquin}@ensem.inpl-nancy.fr

Keywords: State estimation, nonlinear discrete-time systems, multiple model approach, decoupled multiple model.

Abstract: Multiple model approach is a powerful tool for modelling nonlinear systems. Two structures of multiple models can be distinguished. The first structure is characterised by decoupled submodels, i.e. with no common state (decoupled multiple model), in opposition to the second one where the submodels share the same state (Takagi-Sugeno multiple model). A wide number of research works investigate the state estimation of nonlinear systems represented by a classic Takagi-Sugeno multiple model. On the other hand, to our knowledge, the state estimation of the decoupled multiple model has not been investigated extensively. This paper deals with the state estimation of nonlinear systems represented by a decoupled multiple model. Conditions for ensuring the convergence of the estimation error are formulated in terms of a set of Linear Matrix Inequalities (LMIs) employing the Lyapunov direct method.

1 INTRODUCTION

Highly nonlinear processes are commonly encountered in practical engineering problems (chemistry, mechanic, hydraulic, electrotechnics, etc). An accurate model with a simple structure, preferably linear, is often necessary for designing a control law or setting up a diagnosis strategy using conventional control tools. Building only one model, valid in whole operating space of the system, is not always possible due, for example, to the change of the dynamic behaviour in the operating space. Hence, the operating space of the system is often limited before the identification stage (local modelling).

New techniques of identification have been developed for modelling the overall behaviour of the process (global modelling). One of these techniques is based on the decomposition of the operating space of the system into a finite number of operating zones. Each operating zone is characterised by a submodel that has a simple structure. According to the zone where the nonlinear system evolves, the output $y_i$ of each submodel is more or less requested in order to describe the global behaviour $y$ of the nonlinear system, that is to say:

$$y(k) = \sum_{i=1}^{L} \mu_i(k)y_i(k),$$  \hspace{1cm} (1)

where the $i$th submodel contribution depends on the weighting function $\mu_i$. A wide number of identification techniques based on this same principle can be distinguished: piecewise linear model, radial basis function networks, fuzzy models, multiple models, etc.

In this communication, we tackle the multiple model approach. Classically, the multiple model is built using linear submodels associated with weighting functions that ensure a smooth blend between the submodels. It is important to note that the multiple models are considered as an universal approximation tool of nonlinear systems (Johansen et al., 2000). Hence, it is possible to apply the available tools for linear systems to nonlinear systems represented by a multiple model.

In (Filev, 1991) two possible interpretations of equation (1) have been investigated in a fuzzy modelling framework (these interpretations will be directly related to multiple model). In the first interpretation, the submodels are decoupled and their state vector is different (decoupled multiple model); in the second one, the submodels have the same state vector (Takagi-Sugeno multiple model).

The second interpretation has been widely popularized and many works deal with the identification and analysis (control, state estimation, diagnosis, etc.) of nonlinear systems represented by this class of mul-
multiple model.

By comparison with the Takagi-Sugeno multiple model, the decoupled multiple model has been less investigated. Some works in control domain (Gawthrop, 1995; Gatzke and Doyle III, 1999; Gregoric and Lightbody, 2000) and in identification (Venkat et al., 2003) of nonlinear systems have employed successfully this structure and shown its relevance. However, to our knowledge the state estimation problem has not been investigated.

In this paper, a new method for designing a state estimator of nonlinear discrete-time systems represented by a decoupled multiple model is presented. The paper starts with section 2 that introduces two multiple model structures according to the selected interpretation. Stability of decoupled multiple model is investigated in section 3. In section 4, sufficient conditions (in LMIs terms) are established in order to ensure the asymptotic convergence of the estimation error. Finally, section 5 presents an academic example of state estimation of a decoupled multiple model.

2 MULTIPLE MODEL STRUCTURES

The interconnection of the submodels can be performed with various structures in order to generate the global output of the multiple model. Two essential structures of multiple models can be distinguished whether the same state vector appears in all submodels or not.

Concerning the identification step, there exists different techniques (linearisation, parametric optimisation) for the parameter estimation of the submodels for a particular multiple model structure. See (Murray-Smith and Johansen, 1997; Gasso et al., 2001; Venkat et al., 2003) and the references therein for further information about these techniques.

2.1 Takagi-Sugeno Multiple Model

The Takagi-Sugeno multiple model structure is conventionally employed in multiple model analysis and synthesis (Murray-Smith and Johansen, 1997). This multiple model has the following structure:

\[ x_i(k+1) = A_i x_i(k) + B_i u(k), \]
\[ x(k+1) = \sum_{i=1}^{L} \mu_i(\xi(k)) x_i(k+1), \quad (2) \]
\[ y(k) = \sum_{i=1}^{L} \mu_i(\xi(k)) C_i x_i(k), \]

where \( x \in \mathbb{R}^n \) is the state vector, \( u \in \mathbb{R}^m \) the input and \( y \in \mathbb{R}^p \) the output vector. For the \( i \)-th submodel, \( A_i \in \mathbb{R}^{n \times n} \) is the system matrix, \( B_i \in \mathbb{R}^{n \times m} \) the input matrix and \( C_i \in \mathbb{R}^{p \times n} \) the output matrix. The \( \mu_i \) are the weighting functions with the following properties:

\[ \sum_{i=1}^{L} \mu_i(\xi(k)) = 1, \quad \forall k \] (3a)
\[ 0 \leq \mu_i(\xi(k)) \leq 1 \quad \forall i = 1...L, \forall k \] (3b)
\[ \xi \] is the decision variable that depends, for example, on the measurable state variable and/or input or output of the system.

From equation (2), one can see that in the Takagi-Sugeno multiple model there is a common state that couples all submodel states \( x_i \). Therefore the dimension of the state vectors must be identical for all the submodels.

2.2 Decoupled Multiple Model

Another possible structure using a parallel interconnection of the submodels is proposed in (Filev, 1991). Here, this structure is slightly modified using a state representation as follows:

\[ x_i(k+1) = A_i x_i(k) + B_i u(k), \]
\[ y_i(k) = C_i x_i(k), \]
\[ y(k) = \sum_{i=1}^{L} \mu_i(\xi(k)) y_i(k), \quad (4) \]

where \( x_i \in \mathbb{R}^n \) and \( y_i \in \mathbb{R}^p \) are, respectively, the state vector and the output vector for the \( i \)-th submodel and where \( u, y, \xi, A_i, B_i, C_i \in \mathbb{R}^{n \times n}, \mathbb{R}^{n \times m} \) et \( \mathbb{R}^{p \times n} \) have been defined in the previous section.

It should be noted that the global output of the multiple model is given by a weighted sum of the submodel outputs. The blending between the submodels is made through the static equation. Therefore each submodel evolves independently in its own state space according to the input control and its initial state.

It is obvious that the principal interest of this structure is the decoupling between the submodels. Indeed, in contrast to the Takagi-Sugeno multiple model, in the decoupled multiple model the dimension of the state vector \( x_i \) of each submodel can be different (of course the output vector dimension must be identical). Therefore, this structure is well adapted for modelling strongly nonlinear systems whose structure varies with the operating zone.

**Notation:** The following notations will be used all along this paper. \( P > 0 \) \((P < 0)\) means \( P \) is a positive (negative) definite matrix; \( P^T \) denotes the transpose of \( P \). We shall simply write \( \mu_i(\xi(k)) = \mu_i(k) \).
\section{Stability Analysis}

It is possible to rewrite the equations (4) using an augmented state vector as follows:
\begin{align}
    x(k+1) &= \hat{A}x(k) + \hat{B}u(k), \\
    y(k) &= \hat{C}(k)x(k),
\end{align}
where:
\[
    \hat{A} = \begin{bmatrix}
        0 & 0 & 0 & 0 \\
        0 & 0 & 0 & 0 \\
        0 & 0 & 0 & 0
    \end{bmatrix},
\]
\[
    \hat{B} = \begin{bmatrix}
        B_1 \\
        \vdots \\
        B_L
    \end{bmatrix},
\]
\[
    \hat{C}(k) = \begin{bmatrix}
        \mu_1(k)C_1 \\
        \vdots \\
        \mu_L(k)C_L
    \end{bmatrix}^T.
\]

The matrices $\hat{A}$ and $\hat{B}$ are partitioned block matrices. The output matrix $\hat{C}(k)$ is a partitioned block matrix whose parameters vary with time. Indeed, the weighting functions $\mu_i(k)$ only affect the submodel outputs.

The stability of a decoupled multiple model can be easily established by analysing the eigenvalues of the matrix $\hat{A}$. Notice that the matrix $\hat{A}$ is a block diagonal matrix. Therefore, all eigenvalues of this matrix are inside the unit circle if and only if all eigenvalues of every matrices $A_i$ are inside the unit circle.

To sum up, a decoupled multiple model is stable if and only if all submodels are stable, in contrast to Takagi-Sugeno multiple model where the stability depends not only on the stability of the submodels but also on the weighting function values. In the sequel, the multiple model is assumed to be stable.

\section{State Estimation}

State estimation of Takagi-Sugeno multiple model has been widely investigated in a stabilisation law control design perspective (Tanaka and Sugeno, 1990; Feng et al., 1997; Chadli et al., 2003; Guerra and Vermeiren, 2004). Indeed, most of the used control techniques need the state vector knowledge which is not in general fully measurable.

The classically used state estimator is an extension of the proportional (Luenberger) observer. However, some other classes of state estimators have been developed, for example, sliding mode observers (Palm and Bergstern, 2000) and unknown input observers (Akhenak et al., 2004).

The Lyapunov second method is typically employed in order to establish the sufficient convergence conditions of the estimation error in terms of a set of Linear Matrix Inequalities (LMIs) (Boyd et al., 1994) which are solved using standard convex optimization algorithms.

State estimation of decoupled multiple model has been partially explored in a self-tuning control law design perspective in (Gawthrop, 1995). Thanks to the decoupling between the submodels, the observer gains can be classically calculated under the assumption that the submodel outputs are known. However, in our case, it is assumed that this information is not available. Therefore, this approach cannot be used here.

The aim of this section is to investigate the state estimation of a decoupled multiple model, using only the measurable signals i.e. the input and the output of the system (the outputs of the submodels are not available). It is important to notice that the design of the observer must take into account the blending between the outputs of the submodels in order to guarantee the convergence of the estimation error.

\subsection{Observer Structure}

The proportional gain observer for the decoupled multiple model is given by:
\begin{align}
    \hat{x}_i(k+1) &= A_i\hat{x}_i(k) + B_iu(k) + K_i(y(k) - \hat{y}(k)), \\
    \hat{y}_i(k) &= C_i\hat{x}_i(k), \\
    \hat{y}(k) &= \sum_{i=1}^{L} \mu_i(k)\hat{y}_i(k),
\end{align}
where $\hat{x}_i \in \mathbb{R}^{n_i}$ is the state estimation for the $i^{th}$ submodel, $y(k)$ the output of the multiple model, $\hat{y}(k)$ the output estimation and $K_i \in \mathbb{R}^{n_i \times p}$ the gain of the $i^{th}$ observer. Equation (6) can be written in a compact form using the partitioned matrices (5):
\begin{align}
    \hat{x}(k+1) &= A_{obs}(k)\hat{x}(k) + \hat{B}u(k) + \hat{K}y(k), \\
    \hat{y}(k) &= \hat{C}(k)\hat{x}(k),
\end{align}
where $\hat{K} = [K_1 \ldots K_L]^T$, (8)
\[
    A_{obs}(k) = \hat{A} - \hat{K}\hat{C}(k).
\]

Note that the matrix $A_{obs}(k)$ may be decomposed as follows:
\begin{equation}
    A_{obs}(k) = \sum_{i=1}^{L} \mu_i(k)\Phi_i, 
\end{equation}
(10)
\[
    \Phi_i = \hat{A} - \hat{K}\hat{C}_i,
\]
(11)
where $\hat{C}_i$ is the following partitioned block matrix:
\[
    \hat{C}_i = \begin{bmatrix}
        0 & \ldots & C_i & \ldots & 0 \\
    \end{bmatrix}.
\]

The design of the observer consists in determining the gain $\hat{K}$ such that the estimation error given by:
\begin{equation}
    e(k) = x(k) - \hat{x}(k),
\end{equation}
(13)
converges asymptotically to zero for an arbitrary blending between the submodel outputs.
4.2 Estimation Error Convergence

Here, the second Lyapunov method is used to investigate the estimation error convergence by means of a quadratic Lyapunov function. It is clear that other Lyapunov functions can be considered (see section 4.4). The following Theorem gives a sufficient condition for ensuring the estimation error convergence.

**Theorem 1.** Consider the decoupled multiple model (4) and the observer (6). The asymptotic convergence towards zero of the estimation error is guaranteed if there exists a symmetric and positive definite matrix \( P \) and a matrix \( G \) such that:

\[
\begin{bmatrix}
P & \Phi_i^T P \\
\Phi_i P & G
\end{bmatrix} > 0, \quad i = 1 \ldots L,
\]

where the observer gain is deduced from \( \hat{K} = P^{-1}G \).

**Proof.** Let us consider the following quadratic Lyapunov function:

\[ V(e(k)) = e^T(k)Pe(k), \quad P = P^T \text{ and } P > 0. \]  

(15)

The variation of the above function is given by:

\[ \Delta V(e(k)) = V(e(k + 1)) - V(e(k)), \]  

(16)

\[ \Delta V(e(k)) \text{ must be negative in order to ensure its decrease and the asymptotic error convergence also.} \]

Considering the dynamics of the estimation error given by:

\[ e(k + 1) = A_{\text{obs}}(k)e(k), \]  

(17)

and substituting (15) and (17) into (16), then \( \Delta V(e(k)) \) becomes:

\[ \Delta V(e(k)) = e^T(k)\left( A_{\text{obs}}^T(k)PA_{\text{obs}}(k) - P \right)e(k), \]  

(18)

that is a quadratic form in \( e(k) \). Therefore, a necessary and sufficient condition for ensuring \( \Delta V(e(k)) < 0 \) is:

\[ A_{\text{obs}}^T(k)PA_{\text{obs}}(k) - P < 0, \quad \forall k. \]  

(19)

By considering (10), the inequality (19) can be rewritten as:

\[
\sum_{j=1}^{L} \mu_j(k)\Phi_j^T PP^{-1}P \sum_{i=1}^{L} \mu_i(k)\Phi_i - P < 0.
\]

(20)

Combining the Schur complement with property (3a) of the weighting functions, it is possible to write:

\[
\sum_{i=1}^{L} \mu_i(k) \begin{bmatrix} P & \Phi_i^T P \end{bmatrix} > 0.
\]

(21)

The inequality (21) can be upper bounded using the property (3b) of the weighting functions. Finally, a sufficient condition that ensures the error convergence is given by:

\[
\begin{bmatrix} P & \Phi_i^T P \end{bmatrix} > 0, \quad i = 1 \ldots L
\]

(22)

and substituting (11) for \( \Phi_i \), we obtain:

\[
\begin{bmatrix} P & (\hat{A} - \hat{K}\hat{C}_i)^T P \end{bmatrix} > 0, \quad i = 1 \ldots L.
\]

(23)

These matrix inequalities are nonlinear in \( \hat{K} \) and \( P \). Therefore, it is not possible to solve them directly using classical LMI tools. The following change of variables \( G = PK \) allows the linearisation of this problem and ends the demonstration of Theorem 1.

4.3 Eigenvalue Placement

In order to enforce dynamic performances of the observer (for example, the damping and the estimation error decay rate) the eigenvalue placement of the observer must be investigated. In (Chilali and Gahinet, 1996) a general characterization for eigenvalues clustering in subregions of the complex plane in terms of LMIs is proposed.

The eigenvalues of the matrix \( X \) are placed inside the circle with radius \( R \) and centred at \((q,0)\) in the \( z \) plane if the following LMI is feasible:

\[
\begin{bmatrix}
-RP \\
-qP + (XP)^T & -RP
\end{bmatrix} < 0,
\]

(24)

where \( P \) is a symmetric and positive definite matrix.

Let us notice that if \( R = 1 \) and \( q = 0 \) then we obtain the stability condition for linear discrete-time systems.

In order to place the eigenvalues of the observer, the LMIs of Theorem 1 are modified as follows.

**Theorem 2.** Consider the decoupled multiple model (4) and the observer (6). The eigenvalues of the observer are placed inside the circle with radius \( R \) and centred at \((q,0)\) if there exists a symmetric and positive definite matrix \( P \) and a matrix \( G \) such that:

\[
\begin{bmatrix}
-RP \\
-qP + (XP)^T & -RP
\end{bmatrix} < 0,
\]

(25)

for \( i = 1 \ldots L \), where the observer gain is given by \( \hat{K} = P^{-1}G \).

It is clear that this Theorem coincides with Theorem 1 if \( R = 1 \) and \( q = 0 \). In order to avoid strong oscillations of the estimation error, the real part of the eigenvalues of the observer are placed in the positive zone of the unit circle and their imaginary part must be reduced. A judicious choice of the radius \( R \) and the centre \((q,0)\), for example \( q = 0.5 \) and \( R = 0.45 \), allows an appropriate placement of the eigenvalues of the observer.
4.4 Relaxed Convergence Conditions

The asymptotic convergence conditions of the estimation error, presented in the previous section, depend on the existence of the common matrix $P$ which satisfies a set of LMIs. In general, when the multiple model has a large number of submodels, the matrix $P$ cannot be found.

In order to reduce the conservatism of the conditions obtained with a quadratic Lyapunov function, new candidate Lyapunov functions called nonquadratic functions have been proposed. A wide number of published works show the efficient relaxation of the stability conditions provided by this class of functions for a continuous time Takagi-Sugeno multiple model (Jadbabaie, 1999; Rhee and Won, 2006) and also in the discrete time case (Guerra and Vermeiren, 2004). The following Theorem gives a sufficient condition for ensuring the estimation error convergence using a nonquadratic function.

**Theorem 3.** Consider the decoupled multiple model (4) and the observer (6). The asymptotic convergence towards zero of the estimation error is guaranteed if there exists symmetric and positive definite matrices $P_i$ and $P_j$ and a some matrix $M$ and $G$ such that:

$$
\begin{bmatrix}
  P_i \\
  M - M^T - P_j
\end{bmatrix} \begin{bmatrix}
  (M \tilde{A} - GC_i)^T \\
  M + M^T - P_j
\end{bmatrix} > 0 \forall i, j = 1...L,
$$

where the observer gain is deduced from $\hat{K} = M^{-1}G$.

*Proof.* The considered nonquadratic Lyapunov function is given by:

$$
V(e(k)) = e^T(k) \sum_{i=1}^{L} \mu_i(k) P_i e(k) = e^T(k) P(k) e(k),
$$

where $P_i = P_i^T$ and $P_i > 0$. The convergence error analysis is performed as in the previous case. A necessary and sufficient condition in order to ensure the error convergence is given by:

$$
A_{obs}(k)P(k+1)A_{obs}(k) - P(k) < 0.
$$

Introducing (10) and using the Schur complement, the above inequality becomes:

$$
\sum_{j=1}^{L} \mu_j(k) \mu_j(k+1) \begin{bmatrix}
  P_j \\
  P_j \Phi_j
\end{bmatrix} > 0.
$$

Using property (3b) of the weighting functions and substituting (11) for $\Phi_j$, one obtains the following sufficient condition that ensures the asymptotic convergence of the estimation error:

$$
\begin{bmatrix}
  P_i \\
  P_j (\tilde{A} - \hat{K} C_i)
\end{bmatrix} \begin{bmatrix}
  (\tilde{A} - \hat{K} C_i)^T P_j \\
  P_j
\end{bmatrix} > 0, \quad i, j = 1...L.
$$

Inequalities (30) are nonlinear matrix inequalities in $\hat{K}, P_i$ and $P_j$. In contrast to the quadratic case, there is not variable change that allows the direct linearisation of this problem. However, the results coming from (De Oliveira et al., 1999) (Theorem 2) help to rewrite the inequalities (30) as follows:

$$
\begin{bmatrix}
  P_i \\
  M(\tilde{A} - \hat{K} C_i)
\end{bmatrix} \begin{bmatrix}
  (M(\tilde{A} - \hat{K} C_i))^T \\
  M + M^T - P_j
\end{bmatrix} > 0, \quad i, j = 1...L,
$$

where $M$ is not constrained to be symmetric ($M \neq M^T$). After this transformation, the linearisation of the above inequalities can be effectively yielded by using the change of variables $G = M \hat{K}$. Hence, the proof of Theorem 3 is completed.

Notice that Theorem 1 is encompassed by Theorem 3. Indeed, if one sets $P_j = P_j = M = P$ then Theorem 3 coincides with the Theorem 1. Therefore, the previous result is less conservative than the condition obtained with a conventional quadratic function.

5 EXAMPLE

Let us consider the state estimation of the decoupled multiple model with $L = 3$ submodels. The numerical matrices $A_1, B_1$ and $C_1$ are:

$$
\begin{align*}
A_1 &= \begin{bmatrix} 0.8 & 0 \\ 0.4 & 0.1 \end{bmatrix}, & A_2 &= \begin{bmatrix} -0.3 & -0.5 & 0.2 \\ -0.2 & -0.1 & -0.7 \end{bmatrix}, & A_3 &= \begin{bmatrix} -0.5 & 0.1 \\ -0.6 & -0.5 \end{bmatrix}, \\
B_1 &= \begin{bmatrix} 0.2 & -0.4 \\ 0.7 & -0.5 \end{bmatrix}, & B_2 &= \begin{bmatrix} 0.7 & -0.5 & 0.3 \\ 0.1 & -0.2 & 0 \end{bmatrix}, & B_3 &= \begin{bmatrix} -0.2 & 0 \\ -0.9 & 0.3 \end{bmatrix}, \\
C_1 &= \begin{bmatrix} 0.7 & 0.2 \\ 0.5 & 0.8 \end{bmatrix}, & C_2 &= \begin{bmatrix} 0.7 & 0.2 & 0.1 \\ 0.1 & -0.2 & 0.1 \\ 0.2 & -0.1 & 0.1 \\ 0.3 & -0.2 & 0.1 \end{bmatrix}, & C_3 &= \begin{bmatrix} 0.7 & 0.3 \\ 0.5 & 0.8 \end{bmatrix}.
\end{align*}
$$

Here, the decision variable $\xi$ is the input signal $u(k) \in [0, 1]$. The weighting functions are obtained from normalised Gaussian function:

$$
\begin{align*}
\mu_i(u(k)) &= \omega_i(u(k)) / \sum_{j=1}^{L} \omega_j(\xi(k)), \\
\omega_i(\xi(k)) &= \exp \left( - (u(k) - c_i) / \sigma^2 \right),
\end{align*}
$$

with the standard deviation $\sigma = 0.4$ and the centre $c_i = [0.1, 0.5, 0.9]$. The eigenvalues of the matrix $\hat{A}$ are inside the unit circle, thus the multiple model is stable. Using Theorem 3, we obtain the following observer gain:

$$
\hat{K} = \begin{bmatrix}
0.041 & 0.020 & 0.160 & 0.190 & 0.221 & -0.090 & -0.181 \\
0.194 & 0.113 & -0.299 & -0.204 & -0.701 & 0.172 & 0.268
\end{bmatrix}^T.
$$

As can be seen in figures 1 and 2, the suggested observer provides a good output estimation. The error around the origin time is due to the different initial conditions of the multiple model and the observer.
6 CONCLUSION

A decoupled discrete time multiple observer has been presented in order to proceed to the state estimation of a class of nonlinear systems. The proposed observer is an extension of the proportional observer used in the linear observer theory.

Sufficient conditions that guarantee the asymptotic convergence of the estimation error are given in terms of a set of LMIs using a quadratic Lyapunov function. Less conservative conditions are also proposed thanks to a nonquadratic Lyapunov function. In order to illustrate the performances of the proposed observer an academic example is presented.

There are interesting prospects in control and diagnosis of nonlinear systems using this class of multiple model and observer. In particular, this observer class may be useful for setting up a diagnosis strategy for example. This task can be done with a bank of the proposed observers that produce a set of residual signals useful for sensor fault detection and isolation. In future work, the proposed approach will be extended to other observer classes as proportional integral observer or unknown input observer.

![Figure 1: Output $y_1$ of the multiple model (solid line) and its estimated (dashed line).](image1)

![Figure 2: Output $y_2$ of the multiple model (solid line) and its estimated (dashed line).](image2)

REFERENCES


