D.3.2.1 - Hybrid state estimation theory

WP3.2 Hybrid systems

WP3 State estimation and system monitoring

MODRIO (11004)

Version final report
Date April 2016

Author Qinghua Zhang Inria

Executive summary

In this study we consider hybrid systems consisting of a finite number of modes, each exhibiting a time invariant stochastic linear dynamics, with mode transitions governed by a Markov model. In Chapter 1 each mode of a hybrid system is described by a discrete time state-space model, typically derived by discretizing continuous time state equations in the form of ordinary differential equations (ODE). Joint state-parameter estimation is studied in this framework, in particular for actuator fault diagnosis. In Chapter 2, in order to study hybrid systems with modes characterized by differential algebraic equations (DAE), descriptor mode equations are considered at the place of state-space equations. While mode models are assumed available in Chapters 1 and 2, the purpose of Chapter 3 is to estimate them from data. By collecting input-output data during a known sequence of mode transitions, the state-space model of each mode is estimated with a common state-space basis. Simulation results are reported in each chapter illustrating the developed algorithms.

Summary

The term "hybrid systems" in this study means dynamical systems that combine continuous dynamics modeled by differential equations (usually discretized in time) with discrete dynamics characterized by transitions among a finite number of discrete states (modes). The purpose of this study is to estimate both continuous and discrete states of a hybrid system from input-output observations, and possibly to estimate some model parameters.

While in principle continuous and discrete states in a hybrid system can be combined in an arbitrary manner, most existing studies assume a hierarchical structure: at the upper level the system is governed by transition rules among a finite number of modes, and at the lower level the system in each mode is described by differential equations. In other words, at each time instant, the system has a continuous dynamics corresponding to one of the modes, and from time to time the continuous dynamics changes to another one corresponding to another mode.

In this report we consider hybrid systems consisting of a finite number of modes, each exhibiting a time invariant stochastic linear dynamics, with mode transitions governed by a Markov model. In Chapter 1 each mode of a hybrid system is characterized by a discrete time state-space model, typically derived by discretizing continuous time state equations in the form of ordinary differential equations (ODE). Joint state-parameter estimation is studied in this framework, in particular for actuator fault diagnosis. In Chapter 2, in order to study hybrid systems with modes characterized by differential algebraic equations (DAE), descriptor mode equations are considered at the place of state-space equations. While mode models are assumed available in Chapters 1 and 2, the purpose of Chapter 3 is to estimate them from data. By collecting input-output data during a known sequence of mode transitions, the state-space model of each mode is estimated with a common state-space basis. Simulation results are reported in each chapter illustrating the developed algorithms.

Chapter 1

Joint state-parameter estimation of state-space hybrid systems

1.1 Introduction

In order to improve the safety, the reliability and the performance of complex industrial systems, the topic of fault diagnosis is attracting more and more researchers during the last decades (Basseville and Nikiforov, 1993; Gertler, 1998; Chen and Patton, 1999; Ding, 2008; Isermann, 2006; Korbicz, 2004; Blanke et al., 2003; Simani et al., 2003). Most of these studies are model-based, assuming that the considered systems are described by differential equations or by their discrete time counterpart. However, complex industrial systems may have behaviors that cannot be described by a single set of differential equations, because of working mode changes. Typically, each of the working modes is modeled by a different set of differential equations. Such systems are known as *hybrid systems*. It is thus necessary to develop methods for hybrid system fault diagnosis in order to cope with complex industrial systems.

In general a hybrid system involves both continuous dynamics and discrete events that may be combined in a more or less complex manner (Blom and Lygeros, 2006; Lunze and Lamnabhi-Lagarrigue, 2009). In this chapter it is assumed that each considered hybrid system has a finite number of working modes described by discrete time stochastic state-space models subject to parameter changes, and that at every time instant one of the working modes is active. If the active mode is known all the time, then the hybrid system can be treated as a (discontinuous) time varying system, and in this case some fault diagnosis methods designed for time varying systems can be applied (Chung and Speyer, 1998; Chen and Speyer, 2000; Chen et al., 2003; Zhong et al., 2010; Zhang and Basseville, 2014). It is more challenging to study hybrid systems whose active mode is unknown. For hybrid systems with a deterministic mode switching mechanism, fault diagnosis has been studied in (Bemporad et al., 1999) through the mixed logic dynamic formalism (see also (Ferrari-Trecate et al., 2002)), in (Belkhiat et al., 2011) with state observers, and in (Wang et al., 2013) with a bond graph-based approach. When the deterministic mode switching mechanism can be determined from observations within a finite time, a hybrid observer-based method has been proposed in (Wang et al., 2007). Stochastically switching hybrid systems are considered in (Cinquemani et al., 2004), with each monitored fault modeled as one of the modes of the hybrid system. More generally, particle filters can be applied to stochastic hybrid systems for fault diagnosis (Guo et al., 2013), with numerically intensive algorithms.

For the hybrid systems considered in this chapter, each mode is described by a state-space model subject to Gaussian noises, and the mode switching mechanism is characterized by a Markov model. Such systems exhibit stochastic behaviors both in each mode and during mode transitions. In this framework, actuator faults are modeled as parameter changes, typically the gain losses of actuators. The magnitude of each parameter change, when it occurs, is an *unknown* real value, possibly belonging to some bounded interval. It is thus impossible to model such faults as modes of a hybrid system, as each fault would correspond to a particular value of the parameter change magnitude and infinitely many modes would have to be considered.

In the *fault-free case*, the hybrid systems considered in this chapter have been well studied for the problem of *state estimation* (Blom and Bar-Shalom, 1988; Bar-Shalom et al., 2001; Hwang et al., 2006), notably with the well-known Interacting Multiple Model (IMM) estimator (Bar-Shalom et al., 2001). Such results cannot be directly applied to the systems considered in this chapter for state estimation, because of unknown parameter changes. On the other hand, for time varying state-space systems, efficient joint state-parameter estimation methods exist, for instance the *adaptive Kalman filter*, also known as adaptive observer for continuous time systems (Zhang, 2002; Li et al., 2011). The *main idea* of this chapter is to combine the IMM estimator and the adaptive Kalman filter, in order to design an *adaptive IMM estimator* for joint state-parameter estimation of hybrid systems. Then the resulting algorithm can be directly applied to the actuator fault diagnosis problem considered in this chapter.

1.2 Problem statement

The stochastic hybrid system considered in this chapter is modeled at two levels. At the top level, the system has a finite number of working modes. At each time instant, one of the modes is active, and random transitions between different modes are characterized by a Markov model. At the bottom level, each mode of the system is described by a stochastic linear state-space model subject to actuator faults formulated as parameter changes.

1.2.1 Markov transition model

Assume that a hybrid system has r working *modes*, labeled by $M_1, M_2, \ldots, M_r$. At the initial time instant $k = 0$, the *prior probability* that (the mode labeled by[1]) M_j is active is

$$P\{M_j\} = \mu_j(0) \tag{1.1}$$

with known probabilities $\mu_1(0), \mu_2(0), \ldots, \mu_r(0)$ satisfying

$$\sum_{j=1}^{r} \mu_j(0) = 1.$$

The mode switching mechanism is characterized by a Markov process with the transition probabilities

$$P\{M_j(k)|M_i(k-1)\} = p_{i,j}. \tag{1.2}$$

where k is the *discrete time instant*, and $p_{i,j}$ are known transition probabilities independent of k and satisfying

$$\sum_{j=1}^{r} p_{i,j} = 1 \quad \text{for} \quad i = 1, 2, \ldots, r.$$

[1]For shorter statements, "the mode labeled by M_j" is often simply written as "M_j" in this chapter.

1.2.2 Stochastic state-space mode model

In each of the possible r modes, say M_j, the considered hybrid system is described by the linear state-space model

$$x(k) = A(M_j)x(k-1) + B(M_j)u(k-1) + w(k) + \Phi(k-1; M_j)\theta \tag{1.3a}$$

$$y(k) = C(M_j)x(k) + v(k) \tag{1.3b}$$

where $x(k) \in \mathbb{R}^n$ is the state, $u(k) \in \mathbb{R}^l$ the input, $y(k) \in \mathbb{R}^m$ the output, $A(M_j), B(M_j), C(M_j)$ are mode-dependent matrices of appropriate sizes, $w(k) \in \mathbb{R}^n, v(k) \in \mathbb{R}^m$ are mutually independent white Gaussian noises of covariance matrices $Q(M_j) \in \mathbb{R}^{n\times n}$ and $R(M_j) \in \mathbb{R}^{m\times m}$ respectively, and the term $\Phi(k; M_j)\theta$ represents actuator faults with a known matrix sequence $\Phi(k; M_j) \in \mathbb{R}^{n\times p}$ and a constant (or piecewise constant) vector $\theta \in \mathbb{R}^p$. At the initial time instant $k = 0$, the initial state $x(0)$, under the assumption of each possible mode M_j, is assumed to be a Gaussian random vector

$$x(0) \sim \mathcal{N}(\hat{x}^j(0|0), P^j(0|0)). \tag{1.4}$$

A typical example of actuator faults represented by $\Phi(k; M_j)\theta$ is in the case where θ corresponds to actuator gain loss coefficients. Assume that each of the l actuators, say the one corresponding to the q-th component of the vector $u(k)$, is affected by a gain loss represented by a coefficient $(1-\theta_q)$, where θ_q is the q-th component of θ. Then the input term changes from its nominal form $B(M_j)u(k)$ to the faulty form

$$B(M_j)(I_l - \mathrm{diag}(\theta))u(k) = B(M_j)u(k) - B(M_j)\mathrm{diag}(u(k))\theta$$

where I_l is the $l \times l$ identity matrix. In this case, $p = l$ and

$$\Phi(k; M_j) = -B(M_j)\mathrm{diag}(u(k)). \tag{1.5}$$

1.2.3 Actuator fault diagnosis

The problem of actuator fault diagnosis considered in this chapter is to characterize actuator parameter changes (typically gain loss coefficients associated with $\Phi(k; M_j)$ as expressed in (1.5)) by estimating them from the mode prior probabilities $\mu_j(0)$, the mode transition probabilities $p_{i,j}$, the mode-dependent matrices $A(M_j), B(M_j), C(M_j)$, $Q(M_j)$, $R(M_j)$, $\Phi_k(M_j)$, and the input-output data sequences $u(k), y(k)$.

In this considered framework, the actual active mode M_j at each time instant k is *unknown*. This is the main cause of difficulty for stochastic hybrid system fault diagnosis compared to the case of single mode (non hybrid) systems. As a matter of fact, if the active mode was known at each time instant k, then the considered hybrid system would be equivalent to a *time varying* state-space system, for which the actuator fault diagnosis problem, similar to the one formulated in this chapter, has already been studied. See, for instance, (Zhang and Basseville, 2014).

1.3 Interacting multiple model estimator for state estimation

In this section, one of the two basic elements for developing the proposed fault diagnosis method, the well-known *Interacting Multiple Model* (IMM) estimator for hybrid system state estimation, is shortly recalled.

The problem of state estimation, for the considered *fault-free* ($\theta = 0$ in (1.3)) stochastic hybrid system, is to characterize the probability distribution of the state vector $x(k)$, from the mode prior probabilities $\mu_j(0)$, the mode transition probabilities $p_{i,j}$, the mode-dependent matrices $A(M_j), B(M_j), C(M_j), Q(M_j), R(M_j)$, and the input-output data sequences $u(k), y(k)$.

1.3.1 Optimal multiple model estimator and simplifications

In general, mode transitions can happen at every time instant k. To ease the introduction of multiple model estimators, let us first consider a much simpler case: the considered system always remains in one of the r possible modes, but the actual mode is unknown. This is not really a hybrid system and will be referred to as a *static unknown mode system*. In this case, as the actual mode is unknown, we have to try each of them. Under the assumption of each possible mode, say M_j, the system is characterized by the faut-free ($\theta = 0$) stochastic state-space model (1.3), to which the Kalman filter can be applied for stat estimation. Therefore, at every time instant, r Kalman filters are run in parallel, each assuming a particular working mode. Each of these Kalman filters provides a state estimation. The overall state estimation can be made from a weighted average of the r state estimates with weighting coefficients equal to the posterior probabilities of the r possible modes given input-output observations. This algorithm is known as the *static multiple model estimator* (Bar-Shalom et al., 2001).

Now let us consider true hybrid systems with mode transitions that may happen at each time instant k. In this case, in principle it is no longer sufficient to run r Kalman filters. As mode transitions can happen at every time instant, *all the possible mode sequences* up to the current instant k should be considered. Within the instants from 1 to k, there are r^k different possible mode sequences. In principle, for the *optimal dynamic multiple model estimator*, r^k Kalman filters should be run in parallel, each corresponding to one of the possible mode sequences. Following this approach, *the number of Kalman filters increases exponentially* with k. In practice it is not reasonable to implement such solutions, it is then necessary to make simplifications, leading to heuristic solutions, notably the IMM estimator.

1.3.2 The IMM estimator

At each time instant $k = 1, 2, 3, \ldots$, the IMM estimator performs one iteration composed of the same computation steps based on the input-output data, on the mode-dependent system matrices, and for $k = 1$ on the initialization data *or* for $k > 1$ on the results of the last iteration.

The computation steps at iteration k are as follows.

- Compute the mixing probabilities for mixing the state estimates and covariances of the last iteration.
- Compute the mixed state estimates and covariance matrices from the last iteration of the r parallel Kalman filters weighted by the mixing probabilities.
- For each of the r assumed active modes at instant k, perform a Kalman filter iteration delivering a state estimate and a covariance matrix.
- Update mode probabilities from the likelihoods of the r Kalman filters.
- Deliver the algorithm output of the current iteration, by averaging the estimates of the r Kalman filters weighted with the updated mode probabilities.

See (Bar-Shalom et al., 2001) for more details about the IMM estimator.

1.4 Time varying system adaptive Kalman filter

In this section, the second basic element for developing the proposed fault diagnosis method, the *adaptive Kalman filter*, is shortly introduced.

Consider time varying state-space systems in the form of

$$
\begin{aligned}
x(k) &= A(k)x(k-1) + B(k)u(k-1) + w(k) \\
&\quad + \Phi(k-1)\theta && (1.6a) \\
y(k) &= C(k)x(k) + v(k). && (1.6b)
\end{aligned}
$$

Here it is assumed that the *time varying* system matrices $A(k), B(k), C(k), \Phi(k-1)$ and noise covariance matrices $Q(k), R(k)$ are known at every time instant k. The stochastic hybrid systems formulated in Section 1.2 are also state-space systems with system and covariance matrices evolving in time, but the evolution characterized by a Markov model is *not* known exactly at each time instant.

It is also assumed for system (1.6) that the initial state $x(0)$ is a random vector following the Gaussian distribution $\mathcal{N}(\hat{x}(0|0), P(0|0))$, and an initial guess of the unknown parameter vector θ, namely $\hat{\theta}(0)$, is given.

The adaptive Kalman filter is designed for joint estimation of the state vector $x(k)$ and the parameter vector θ of system (1.6). If the parameter vector $\theta \in \mathbb{R}^p$ was known in (1.6), then the basic Kalman filter would be applicable to its state estimation. In order to perform joint state-parameter estimation, the adaptive Kalman filter presented below incorporates a parameter estimation mechanism, complimenting the part originating from the basic Kalman filter.In addition to the variables originating from the basic Kalman filter, the new algorithm involves new matrices $\Upsilon(k) \in \mathbb{R}^{n\times p}, \Omega(k) \in \mathbb{R}^{m\times p}, S(k) \in \mathbb{R}^{p\times p}, \Gamma(k) \in \mathbb{R}^{p\times m}$ and a forgetting factor $\lambda \in (0, 1)$. In particular, the recursively computed $\Upsilon(k)$ and $S(k)$ need initializations $\Upsilon(0) = 0$ (the $n \times p$ zero matrix) and $S(0) = \alpha I_p$ with some $\alpha > 0$.

At each time instant k, the adaptive Kalman filter performs the following computations.

$$
\begin{aligned}
P(k|k-1) &= A(k)P(k-1|k-1)A^T(k) + Q(k) && (1.7a) \\
\Sigma(k) &= C(k)P(k|k-1)C^T(k) + R(k) && (1.7b) \\
K(k) &= P(k|k-1)C^T(k)\,[\Sigma(k)]^{-1} && (1.7c) \\
P(k|k) &= (I_n - K(k)C(k))P(k|k-1) && (1.7d)
\end{aligned}
$$

$$
\begin{aligned}
\Upsilon(k) &= (I_n - K(k)C(k))A(k)\Upsilon(k-1) \\
&\quad + (I_n - K(k)C(k))\Phi(k-1) && (1.7e)
\end{aligned}
$$

$$
\begin{aligned}
\Omega(k) &= C(k)A(k)\Upsilon(k-1) + C(k)\Phi(k-1) &\quad (1.7f)\\
S(k) &= \frac{1}{\lambda}S(k-1) - \frac{1}{\lambda}S(k-1)\Omega^T(k)\big(\lambda\Sigma(k)\\
&\quad + \Omega(k)S(k-1)\Omega^T(k)\big)^{-1}\Omega(k)S(k-1) &\quad (1.7g)\\
\Gamma(k) &= S(k)\Omega^T(k)\left(\lambda\Sigma(k) + \Omega(k)S(k)\Omega^T(k)\right)^{-1} &\quad (1.7h)\\
\tilde{y}(k) &= y(k) - C(k)\big(A(k)\hat{x}(k-1|k-1)\\
&\quad - B(k)u(k-1) - \Phi(k-1)\hat{\theta}(k-1)\big) &\quad (1.7i)\\
\hat{\theta}(k) &= \hat{\theta}(k-1) + \Gamma(k)\tilde{y}(k) &\quad (1.7j)\\
\hat{x}(k|k) &= A(k)\hat{x}(k-1|k-1) + B(k)u(k-1)\\
&\quad + \Phi(k-1)\hat{\theta}(k-1) + K(k)\tilde{y}(k)\\
&\quad + \Upsilon(k)(\hat{\theta}(k) - \hat{\theta}(k-1)). &\quad (1.7k)
\end{aligned}
$$

Similar algorithms corresponding to the continuous-time counterpart of this adaptive Kalman filter are more frequently studied (Zhang, 2002; Li et al., 2011). Because of the noises in the state and output equations, the state and parameter estimation errors do not converge to zero when k tends to infinity. If the noises are ignored, the estimation errors converge exponentially to zero (Guyader and Zhang, 2003). In the presence of the noises, by taking mathematical expectations of all the terms in the error equations, the mathematical expectations of the estimation errors converge exponentially to zero.

1.5 Adaptive IMM estimator for hybrid system actuator fault diagnosis

Now let us go back to the problem of hybrid system actuator fault diagnosis formulated in Section 1.2 through equations (1.1)-(1.4).

If in (1.3) the mode-dependent matrices $A(M_j), B(M_j), C(M_j), Q(M_j), R(M_j), \Phi(k; M_j)$ were known at each time instant (in other words, *if the active mode was known* at every time instant), then the system would fit into the framework of time varying systems formulated in (1.6), and it would be possible to directly apply the adaptive Kalman filter. Of course, in the presently considered case of stochastic hybrid systems, the active mode is unknown. Like in the case of optimal state estimator discussed in Section 1.3.1 for fault-free hybrid systems, in principle it is possible to try *all the possible mode sequences* up to the current instant k and to combine somehow all the resulting state and parameter estimates. Again the number of possible mode sequences (r^k) increases exponentially with time.

Let us follow the same ideas as in the basic IMM estimator to avoid the exponentially increasing complexity of the optimal estimator. At each time instant, instead of considering all the possible past mode sequences, summarize the past estimates with weighted averages by using *mixing probabilities* $\mu_{i|j}(k-1|k-1)$, and run r parallel adaptive Kalman filters, each assuming a different currently active mode M_j. The state and parameter estimates of the hybrid system are then obtained by taking the weighted average of the r estimates delivered by the r adaptive Kalman filters, with weights equal to the posterior probabilities of the corresponding modes.

At each instant k, the *adaptive IMM estimator* for joint state-parameter estimation consists of the following steps.

Calculation of mixing probabilities

The mixing probabilities for mixing results of the last iteration (of instant $k-1$) are computed as

$$\mu_{i|j}(k-1|k-1) = \frac{1}{\bar{c}_j} p_{i,j} \mu_j(k-1) \tag{1.8}$$

where $\mu_j(0)$ (for instant $k = 1$) are mode prior probabilities, $\mu_j(k-1)$ (for $k > 1$) are mode probabilities updated at instant $k-1$ (see the *Mode probability update* step below), and

$$\bar{c}_j = \sum_{i=1}^{r} p_{i,j} \mu_i(k-1).$$

Intermediate results mixing

During the last iteration (at instant $k-1$), r adaptive Kalman filters were run in parallel, each assuming a different active mode at instant $k-1$, yielding state estimates $\hat{x}^i(k-1|k-1)$, parameter estimates $\hat{\theta}^i(k-1)$, the matrices $P^i(k-1|k-1)$, $S^i(k-1)$, $\Upsilon^i(k-1)$, all indexed by the corresponding assumed mode M_i.

The mixed quantities are then computed as

$$\begin{aligned}
\check{x}^j(k-1|k-1) &= \sum_{i=1}^{r} \hat{x}^i(k-1|k-1)\mu_{i|j}(k-1|k-1) \\
\check{\theta}^j(k-1) &= \sum_{i=1}^{r} \hat{\theta}^i(k-1)\mu_{i|j}(k-1|k-1) \\
\check{S}^j(k-1) &= \sum_{i=1}^{r} S^i(k-1)\mu_{i|j}(k-1|k-1) \\
\check{\Upsilon}^j(k-1) &= \sum_{i=1}^{r} \Upsilon^i(k-1)\mu_{i|j}(k-1|k-1)
\end{aligned}$$

$$\begin{aligned}
\check{P}^j(k-1|k-1) = \sum_{i=1}^{r} \mu_{i|j}(k-1|k-1)\{P^i(k-1|k-1) \\
+ [\hat{x}^i(k-1|k-1) - \check{x}^j(k-1|k-1)] \\
\cdot [\hat{x}^i(k-1|k-1) - \check{x}^j(k-1|k-1)]^T\}.
\end{aligned}$$

Mode-matched filtering

For each of the possible modes M_j at instant k, an adaptive Kalman filter is implemented as follows.

$$
\begin{aligned}
P^j(k|k-1) &= A(M_j)\check{P}^j(k-1|k-1)A^T(M_j)+Q(M_j)\\
\Sigma^j(k) &= C(M_j)P^j(k|k-1)C^T(M_j)+R(M_j)\\
K^j(k) &= P^j(k|k-1)C^T(M_j)\left[\Sigma^j(k)\right]^{-1}\\
P^j(k|k) &= (I_n - K^j(k)C(M_j))P^j(k|k-1)\\
\Upsilon^j(k) &= (I_n - K^j(k)C(M_j))A(M_j)\check{\Upsilon}^j(k-1)\\
&\quad + (I_n - K^j(k)C(M_j))\Phi(k-1;M_j)\\
\Omega^j(k) &= C(M_j)A(M_j)\check{\Upsilon}^j(k-1)\\
&\quad + C(M_j)\Phi(k-1;M_j)
\end{aligned}
$$

$$
\begin{aligned}
S^j(k) &= \frac{1}{\lambda}\check{S}^j(k-1) - \frac{1}{\lambda}\check{S}^j(k-1)[\Omega^j(k)]^T\\
&\quad \cdot \left(\lambda\Sigma^j(k) + \Omega^j(k)\check{S}^j(k-1)[\Omega^j(k)]^T\right)^{-1}\\
&\quad \cdot \Omega^j(k)\check{S}^j(k-1)\\
\Gamma^j(k) &= S^j(k)[\Omega^j(k)]^T\left(\lambda\Sigma^j(k)\right.\\
&\quad \left. + \Omega^j(k)S^j(k)[\Omega^j(k)]^T\right)^{-1}\\
\tilde{y}^j(k) &= y(k) - C(M_j)\big(A(M_j)\check{x}^j(k-1|k-1)\\
&\quad - B(M_j)u(k-1) - \Phi(k-1;M_j)\check{\theta}^j(k-1)\big)\\
\hat{\theta}^j(k) &= \check{\theta}^j(k-1) + \Gamma^j(k)\tilde{y}^j(k)\\
\hat{x}^j(k|k) &= A(M_j)\check{x}^j(k-1|k-1) + B(M_j)u(k-1)\\
&\quad + \Phi(k-1;M_j)\check{\theta}^j(k-1) + K^j(k)\tilde{y}^j(k)\\
&\quad + \Upsilon^j(k)(\hat{\theta}^j(k) - \check{\theta}^j(k-1)).
\end{aligned}
$$

The likelihood of the mode M_j, given the input-output data up to instant k, is evaluated through the innovation $\tilde{y}^j(k)$

$$
\begin{aligned}
\Lambda_j(k) &= \frac{1}{\sqrt{(2\pi)^m \det(\Sigma^j(k))}}\\
&\quad \cdot \exp\left(-\frac{1}{2}(\tilde{y}^j(k))^T(\Sigma^j(k))^{-1}\tilde{y}^j(k)\right).
\end{aligned}
$$

Mode probability update
The mode probabilities are updated as

$$\mu_j(k) = \frac{1}{c}\Lambda_j(k)\bar{c}_j$$

with

$$
\begin{aligned}
\bar{c}_j &= \sum_{i=1}^{r} p_{i,j}\mu_i(k-1)\\
c &= \sum_{j=1}^{r} \Lambda_j(k)\bar{c}_j.
\end{aligned}
$$

Algorithm outputs

The outputs of the adaptive IMM estimator at instant k are the state estimate

$$\hat{x}(k|k) = \sum_{j=1}^{r} \hat{x}^j(k|k)\mu_j(k)$$

and the parameter estimate

$$\hat{\theta}(k|k) = \sum_{j=1}^{r} \hat{\theta}^j(k|k)\mu_j(k).$$

Like the basic IMM estimator, this adaptive algorithm is also based on heuristic simplifications of the optimal estimator to avoid the exponentially increasing complexity.

1.6 Numerical example

In order to illustrate the proposed solution for actuator fault diagnosis in hybrid systems, let us consider a system with 4 modes ($r = 4$). Each mode is described by a third order state-space model ($n = 3$) subject to white Gaussian noises, with one input ($l = 1$) and 2 outputs ($m = 2$). The mode-dependent system matrices $A(M_j), B(M_j), C(M_j)$ are randomly generated such that each mode is stable (the eigenvalues of $A(M_j)$ are inside the unit cycle), observable and controllable. The noise covariances matrices are chosen as $Q(M_j) = 0.1I_3$ and $R(M_j) = 0.05I_2$ for all the 4 modes. The mode transition probabilities $p_{i,j}$ are randomly generated. During each simulation trial, a gain loss of 50%, corresponding to a jump of θ (a scalar parameter, i.e., $p = 1$) from 0 to 0.5 at the time instant $k = 500$ is simulated, and the simulation runs till $k = 1000$. The adaptive IMM estimator proposed in this chapter is then applied to the simulated system for joint state-parameter estimation. The result of parameter estimation for one of the trials is presented in Fig. 1.1, and that of state estimation in Fig. 1.2.

In these figures, the results obtained with the adaptive Kalman filter (see Section 1.4) are also presented as a reference for the purpose of comparison. *In practice the adaptive Kalman filter is not applicable*, as its computations require the true mode transition sequence, which is unknown in practice. Because the adaptive IMM estimator uses less information, it cannot perform as well as the adaptive Kalman filter, but yet the results are quite similar.

In order to statistically evaluate the performance of the proposed method, 1000 simulated trials are performed, each corresponding to a different random realization of mode transition probabilities, mode-dependent system matrices, state and output noises. At each time instant k, the histogram of the parameter estimation error based on the 1000 simulated trials is generated, and all the histograms are depicted as a 3D illustration in Fig. 1.3 for the adaptive IMM estimator, and in Fig, 1.4 for the adaptive Kalman filter, again for the purpose of comparison. The histograms are normalized so that they are similar to probability density functions.

1.7 Conclusion

Based on the existing basic IMM estimator for hybrid system state estimation and on the adaptive Kalman filter for time varying system joint state-parameter estimation, a new algorithm, the *adaptive IMM estimator* has been proposed in this chapter for actuator fault diagnosis in stochastic hybrid systems. In the presented numerical illustrations, the results of the adaptive

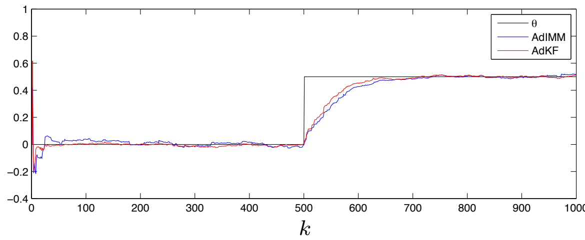

Figure 1.1: The simulated "true" parameter (θ, black), parameter estimates by the adaptive IMM estimator (adIMM, blue) and by the adaptive Kalman filter (AdKF, red).

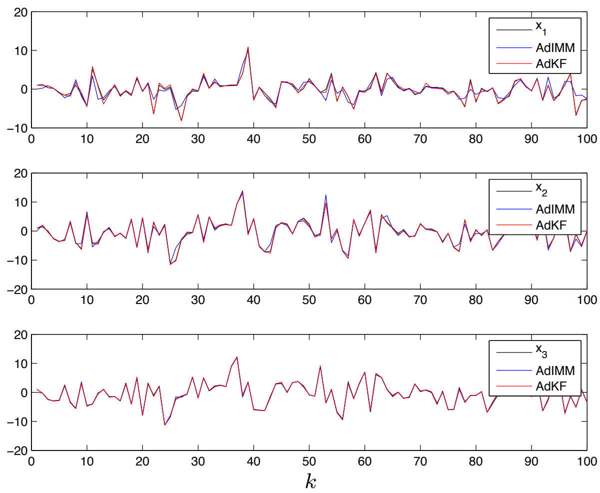

Figure 1.2: State estimates by the adaptive IMM estimator (adIMM) and by the adaptive Kalman filter (AdKF), zoomed for $0 \leq k \leq 100$.

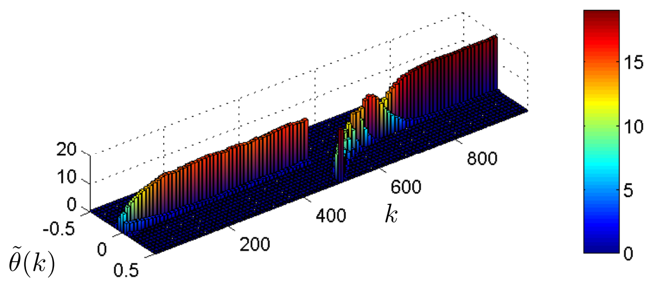

Figure 1.3: Histogram per instant k of the parameter estimation error of the adaptive IMM estimator.

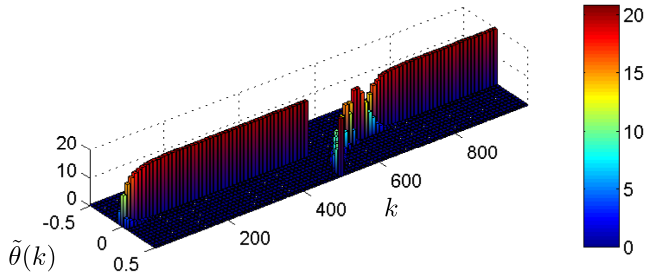

Figure 1.4: Histogram per instant k of the parameter estimation error of the adaptive Kalman filter.

IMM estimator are quite close to those of the adaptive Kalman filter, which represents a performance upper bound that cannot be attained in practice. The computational burden of the adaptive IMM estimator is essentially equal to that of the adaptive Kalman filter multiplied by the number of modes.

Chapter 2

State estimation of descriptor hybrid systems

2.1 Introduction

Differential equations have been widely used in the study of dynamic systems, in particular, finite dimensional state space equations of the form

$$\dot{x}(t) = f(x(t), u(t)) \tag{2.1}$$

are often used for modeling engineering systems, with $x(t)$ and $u(t)$ denoting respectively the state vector and the input vector of the considered system, $\dot{x}(t) = dx(t)/dt$, and f is a function characterizing the dynamic behavior of the system. Despite the large success of such theories in engineering practice, some complex systems cannot be appropriately described in this framework. Such exceptions include differential-algebraic systems and hybrid systems.

Algebraic constraints in engineering systems typically result from singularities of differential equations. For example, a train following a turning track is constrained by the geometrical form of the tack if the mass of the earth is considered infinitely large compared to the mass of the train. Such a system can be described by differential-algebraic equations (DAE) of the form

$$g(\dot{x}(t), x(t), u(t))$$

which can represent a wider class of systems than the classical state space systems of the form (2.1). After linearization and discretization in time, such equations can be approximated by implicit discrete time state space equations of the form

$$E_{k+1}x(k+1) = A_k x(k) + B_k u(k) + \upsilon(k), \tag{2.2}$$

where $x(k), u(k)$ and $\upsilon(k)$ are respectively the discrete time state, input and the modeling errors indexed by $k = 1, 2 \ldots,$ and E_{k+1}, A_k, B_k are time varying matrices of appropriate sizes. With a possibly rank deficient matrix E_{k+1}, refer to (2.2) which is known as a *descriptor* equation (Nikoukhah et al., 1992).

On the other hand, some complex systems have different working modes, for example, the starting mode, the normal working mode, or some reduced regime mode in case of component failures. If in each of these modes the system is described by some differential equations, the over-all functioning of the system, including the mode switching mechanism and the behavior within each mode, is usually modeled as a *hybrid* system (Bar-Shalom et al., 2001).

The purpose of this chapter is to study state estimation for *hybrid* systems with working modes described by *descriptor* equations of the form (2.2).

The study of descriptor systems, notably about state estimation, has a rich literature (Yeu and Kawaji, 2001; Nikoukhah et al., 1992; Marx et al., 2004; Koenig and Mammar, 2002; Gao and Wang, 2006) (for more information see references therein). Besides this topic, state estimation for hybrid systems has been largely studied with various applications such as target tracking (Bar-Shalom and Fortmann, 1988; Blom and Bar-Shalom, 1988), signal processing (Doucet et al., 2001), and fault diagnosis (Hanlon and Maybeck, 2000; Koutsoukos et al., 2002). These reported results concern hybrid systems involving modes described by state space equations. To our knowledge, no study has been reported about hybrid systems with working modes described by *descriptor* equations.

In the case of stochastic hybrid systems with working modes described by classical linear Gaussian state space equations, it is known that the complexity of the optimal state estimator increases exponentially with time. In practice, heuristic algorithms of reduced complexity are used, without convergence proof. Among such algorithms, is the *interacting multiple model* (IMM for short) estimator (Bar-Shalom et al., 2001), (Seah and Hwang, 2009). It is based on multiple Kalman filters, each assuming a particular working mode. At each iteration, in addition to the classical prediction and update steps of the Kalman filter, all the Kalman filters are modified based on the knowledge about the mode transition mechanism modeled as a Markov chain.

In this chapter, the algorithm proposed for hybrid descriptor system state estimation follows also the IMM approach. The main novelties in this algorithm reside in the replacement of the classical state space system Kalman filter by the descriptor system Kalman filter, and in a new method for the evaluation of the likelihood of each Kalman filter at each iteration. Of course, like the classical IMM estimator for state-space hybrid systems, no convergence proof of the algorithm is known.

This chapter is organized as follows: In Section 2.2 we present a formal formulation of the estimation problem for a stochastic linear hybrid descriptor system. In Section 2.3, we make a comparison between the Kalman filter for classical situation and the descriptor one. We state the IMM algorithm in the framework of descriptor system in 2.4. In Section 2.6, we provide numerical examples. Finally, conclusions are presented in Section 2.7.

2.2 Problem formulation

The stochastic hybrid systems considered in this chapter are modeled at two levels. At the top level, each system has a finite number of working modes. At each time instant, one of the modes is active, and random transitions between different modes are characterized by a Markov model. At the bottom level, each mode of the system is described by a stochastic descriptor system model.

2.2.1 Top level Markov transition model

Let $\{M_1, M_2, \ldots, M_r\}$ be the set of r possible modes of a hybrid system, and $m(k) \in \{1, 2, \ldots, r\}$ denote the index of the mode at time instant $k = 0, 1, 2, \ldots$. The mode evolution, corresponding to the sequence $m(0), m(1), \cdots$, is a Markov chain described by the mode transition matrix

$$\Pi = \{p_{ij}\}_{i,j=1,\cdots,r},$$

where

$$p_{i,j} = P\left[m(k+1) = j | \, m(k) = i\right],$$

with known transition probabilities $p_{i,j}$ independent of k and satisfying

$$\sum_{j=1}^{r} p_{i,j} = 1, \qquad i = 1, 2, \ldots, r.$$

For the initial time instant $k = 0$, the *prior probability* that M_j is active is

$$P[m(0) = j] = \mu_j(0)$$

with known probabilities $\mu_1(0)$, $\mu_2(0), \ldots, \mu_r(0)$ satisfying

$$\sum_{j=1}^{r} \mu_j(0) = 1.$$

2.2.2 Bottom level stochastic descriptor system model

At the bottom level, the state $x(k) \in \mathbb{R}^n$, the input $u(k) \in \mathbb{R}^q$ and the output $y(k) \in \mathbb{R}^p$ of the hybrid system satisfy the stochastic descriptor equations

$$\begin{aligned} E_{M_j} x(k+1) &= A_{M_j} x(k) \\ &\quad + B_{M_j} u(k) + W_{M_j}(k), \qquad (2.3) \\ y(k) &= C_{M_j} x(k) \\ &\quad + D_{M_j} u(k) + V_{M_j}(k), \qquad (2.4) \end{aligned}$$

where the time index $k = 0, 1, 2, \cdots$., the descriptor system matrices $E_{M_j} \in \mathbb{R}^{l\times n}$, $A_{M_j} \in \mathbb{R}^{l\times n}$, $B_{M_j} \in \mathbb{R}^{l\times q}$, $C_{M_j} \in \mathbb{R}^{p\times n}$, $D_{M_j} \in \mathbb{R}^{p\times q}$, and the state and output white noises $W_{M_j}(k) \in \mathbb{R}^l$, $V_{M_j}(k) \in \mathbb{R}^p$, with $W_{M_j}(k) \sim \mathcal{N}(0, Q_{M_j})$, $V_{M_j}(k) \sim \mathcal{N}(0, S_{M_j})$ which are mutually independent to each other.

When the mode of the hybrid system stays unchanged, say at M_j, the mode index $m(k) = j$, then the system described by equation (2.3) behaves like a stochastic descriptor system characterized by the matrices A_j, B_j, C_j, D_j, E_j, Q_j, S_j. Within each mode, it is assumed that descriptor system (2.3)-(2.4) is observable and controllable by the state noise. See (Nikoukhah et al., 1992) for the definitions of descriptor system observability and controllability.

Other other hand, the mode M_j may evolve, so that the mode index sequence $m(0), m(1), \cdots$ forms a Markov chain, as described at the top level. In particular, during every mode change, say from $M_{m(k)} = M_j$ to $M_{m(k+1)} = M_i$, the evolution of the state $x(k+1)$ and the output $y(k+1)$ are also described by equation (2.4), given a realization of the mode sequence Markov chain.

Let $Z^k = [u(0), y\,(0)\,, u(1), y(1), \ldots, u(k), y(k)]$ denote the measurements up to time k.

The following assumptions are made in this chapter.

(A1): Assume that $x(0) \sim \mathcal{N}(0, P_0)$ is independent of $W_{M_j}(k)$ and $V_{M_j}(k)$;

(A2): For $j \in \{1, 2, \ldots, r\}$, $H_{M_j} = \begin{bmatrix} E_{M_j} \\ C_{M_j} \end{bmatrix}$ is full column rank;

(A3): For $j \in \{1, 2, \ldots, r\}$, the matrices Q_{M_j} and S_{M_j} are symmetric positive definite.

2.3 Kalman filter for descriptor systems

In this section, let us consider the single-mode descriptor system

$$\begin{aligned} E_{k+1}x(k+1) &= A_k x(k) + B_k u(k) + W(k), && (2.5) \\ y(k+1) &= C_{k+1}x(k+1) + D_{k+1}u(k+1) \\ &\quad +V(k+1). && (2.6) \end{aligned}$$

First recall the classical Kalman filter for state space systems corresponding to the case $l = n$ and $E_{k+1} = I_{n \times n}$.

$$\begin{cases} \hat{x}(0|0) = \mathbb{E}(x_0), \;\; P(0,0) = Var(x_0), \\ P(k+1|k) = A(k)P(k|k)A^T(k) + Q(k), \\ G(k+1) = P(k+1|k)C^T(k+1) \\ \quad \cdot (C(k+1)P(k+1|k)C^T(k+1) + S(k+1))^{-1}, \\ \hat{x}(k+1|k) = A(k)\hat{x}(k|k) + B(k)u(k), \\ \hat{x}(k+1|k+1) = \hat{x}(k+1|k) + G(k+1)(y(k+1) \\ \quad -D(k+1)u(k+1) - C\,(k+1)\,\hat{x}(k+1|k)), \end{cases}$$

where the innovation process, also known as the prediction error, $v(k+1) = y(k+1) - \hat{y}(k+1|k)$ with

$$\hat{y}(k+1|k) = C(k+1)\hat{x}(k+1|k) + D(k+1)u(k+1),$$

is a white Gaussian sequence. The associated likelihood function can be computed by using the innovation $v(k+1)$.

Now let us consider the more general descriptor system (5)-(6). Its Kalman filter (Nikoukhah et al., 1992) writes ("d" in subscript for "descriptor system"),

$$
\begin{cases}
\hat{x}_d(0|0) = \mathbb{E}(x_0), \\
\Sigma_d(k) = A(k)P_d(k|k)A^T(k) + Q_d(k), \\
G_d(k+1) \\
= (E^T(k+1)\Sigma_d^{-1}(k)E(k+1))^{-1}C^T(k+1) \\
\quad \cdot(C(k+1)(E^T(k+1)\Sigma_d^{-1}(k)E(k+1))^{-1} \\
\quad \cdot C^T(k+1) + S_d(k))^{-1}, \\
L_d(k+1) = (E^T(k+1)\Sigma_d^{-1}(k)E(k+1))^{-1} \\
\quad \cdot E^T(k+1)\Sigma_d^{-1}(k) - G_d(k+1)C(k+1) \\
\quad \cdot(E^T(k+1)\Sigma_d^{-1}(k)E(k+1))^{-1} \\
\quad \cdot E^T(k+1)\Sigma_d^{-1}(k), \\
\hat{x}_d(k+1|k+1) = L_d(k+1)A(k)\hat{x}_d(k|k) \\
\quad +L_d(k+1)B(k)u(k) \\
\quad -G_d(k+1)D(k+1)u(k+1) \\
\quad +G_d(k+1)y_d(k+1), \\
P_d(k+1|k+1) = L_d(k+1)A(k)P_d(k|k)A^T(k) \\
\quad \cdot L_d^T(k+1) + L_d(k+1)Q_d(k)L_d^T(k+1) \\
\quad +G_d(k+1)S_d(k)G_d^T(k+1).
\end{cases}
\tag{2.7}
$$

In (Nikoukhah et al., 1992; Moussa Ali and Zhang, 2014) the gain matrices $G_d(k+1)$ and $L_d(k+1)$ were given in implicit forms. Their explicit forms presented here will be proved in Section 2.5. Unlike the classical state space system Kalman filter which was formulated in two steps known as prediction and update, computing respectively $\hat{x}(k+1|k)$ and $\hat{x}(k+1|k+1)$. Here for descriptor systems the filter is written in a single step. It is thus not obvious if the computed state estimate $\hat{x}_d(k+1|k+1)$ in every iteration is the predicted state or the filtered state. In Section 2.5 it will be shown that $\hat{x}_d(k+1|k+1)$ is indeed the filtered state. Consequently, the output estimation error $\bar{v}_d(k+1) = y_d(k+1) - \hat{y}_d(k+1|k+1)$ where

$$
\hat{y}_d(k+1|k+1) = C(k+1)\hat{x}_d(k+1|k+1) + D(k+1)u(k+1).
$$

is not the innovation sequence in the sense of prediction error as defined in the classical Kalman filter. Usually the innovation sequence is used for likelihood evaluation in the IMM approach to hybrid system estimation. Fortunately, like the innovation sequence, this estimation error $v_d(k+1)$ is a white Gaussain sequence in (Moussa Ali and Zhang, 2014). This result is important for the new IMM estimator presented in this chapter for hybrid systems, at the step of likelihood computation.

2.4 Hybrid descriptor system IMM estimator

The algorithm presented below for hybrid descriptor systems is quite similar to the IMM estimator for classical hybrid state space systems (Bar-Shalom et al., 2001), the main differences reside in the Kalman filter for descriptor systems and in the evaluation of the likelihood at each iteration.

The key feature of IMM we point out is that it consists of r interacting filters operating in parallel.

The algorithm consists of the following steps.

1. **Calculation of the mixing probabilities (i, $j = 1, \ldots, r$).** The probability that mode M_i was in effect at k given that M_j is in effect at $k+1$ conditioned on Z^{k+1} is

$$\begin{aligned}\mu_{i|j}(k|k+1) &= P\{M_i(k)\,|M_j(k+1), Z^{k+1}\}\\ &= \frac{1}{\bar{c}_j} p_{i,j}\mu_j(k),\end{aligned}$$

where

$$\bar{c}_j = \sum_{i=1}^{r} p_{i,j}\mu_i(k).$$

2. **Intermediate results mixing ($j = 1, \ldots, r$).**
During the last iteration, r descriptor system Kalman filters were run in parallel, each assuming a different active mode at instant k, yielding r state estimates $\hat{x}(k|k)$ and r covariance matrices $P^i(k|k)$. Based on these results, the mixed state estimates and covariance matrices are

$$\begin{aligned}\hat{x}^{0,j}(k|k) &= \sum_{i=1}^{r} \hat{x}(k|k)\mu_{i|j}(k|k),\\ P^{0,j}(k|k) &= \sum_{i=1}^{r} \mu_{i|j}(k|k)\{P^i(k|k)\\ &\quad +[\hat{x}(k|k) - \hat{x}^{0,j}(k|k)]\\ &\quad \cdot[\hat{x}(k|k) - \hat{x}^{0,j}(k|k)]^T\}.\end{aligned}$$

3. **Mode-matched filtering ($j = 1, \ldots, r$).**
For each of the r assumed active modes at instant $k+1$, say M_j, a descriptor system Kalman filter delivers a state estimate $\hat{x}^j(k+1|k+1)$ as follows:

$$\begin{aligned}\hat{x}^j(0|0) &= \mathbb{E}(x_0),\\ \Sigma^j(k) &= A_{M_j}P^{0,j}(k|k)A_{M_j}^T + Q_{M_j},\\ G_{M_j}(k+1) &= (E_{M_j}^T(\Sigma^j(k))^{-1}E_{M_j})^{-1}C_{M_j}^T\\ &\quad \cdot(C_{M_j}(E_{M_j}^T(\Sigma^j(k))^{-1}E_{M_j})^{-1}C_{M_j}^T\\ &\quad +S_{M_j})^{-1},\\ L_{M_j}(k+1) &= (E_{M_j}^T(\Sigma^j(k))^{-1}E_{M_j})^{-1}\\ &\quad \cdot E_{M_j}^T(\Sigma^j(k))^{-1} - G_{M_j}(k+1)C_{M_j}\\ &\quad \cdot(E_{M_j}^T(\Sigma^j(k))^{-1}E_{M_j})^{-1}E_{M_j}^T(\Sigma^j(k))^{-1},\\ \hat{x}^j(k+1|k+1) &= L_{M_j}(k+1)A_{M_j}\hat{x}^{0,j}(k|k)\\ &\quad +L_{M_j}(k+1)B_{M_j}u(k)\\ &\quad -G_{M_j}(k+1)D_{M_j}u(k+1)\\ &\quad +G_{M_j}(k+1)y(k+1),\\ P^j(k+1|k+1) &= L_{M_j}(k+1)A_{M_j}P^{0,j}(k|k)A_{M_j}^T\\ &\quad \cdot L_{M_j}^T(k+1)\\ &\quad +L_{M_j}(k+1)Q_{M_j}L_{M_j}^T(k+1)\\ &\quad +G_{M_j}(k+1)S_{M_j}G_{M_j}^T(k+1).\end{aligned}$$

The likelihood of the mode M_j, given the input-output data up to instant $k+1$, is evaluated through the "innovation"

$$\begin{aligned}
\bar{v}^j(k+1) &= y(k+1) - \hat{y}^j(k+1|k+1) \\
&= y(k+1) - D_{M_j} u(k+1) \\
&\quad -C_{M_j} \hat{x}^j(k+1|k+1), \\
\Lambda_j(k+1) &= \frac{1}{\sqrt{(2\pi)^p \det(\Sigma^j(k+1))}} \\
&\quad \cdot \exp(\Upsilon^j(k+1)),
\end{aligned}$$

where

$$\Upsilon^j(k+1) = -\frac{1}{2} \left(\bar{v}^j(k+1)\right)^T \left(\Sigma^j(k+1)\right)^{-1} \bar{v}^j(k+1).$$

4. **Mode probability update.**
   The mode probabilities are updated as

$$\mu_j(k+1) = \frac{1}{c} \Lambda_j(k+1) \bar{c}_j$$

with

$$\bar{c}_j = \sum_{i=1}^{r} p_{i,j} \mu_i(k), \qquad c = \sum_{j=1}^{r} \Lambda_j(k+1) \bar{c}_j.$$

5. **Estimate and covariance combination.**
   Combination of the model-conditioned estimates and covariances is completed according to the mixture equations

$$\hat{x}(k+1|k+1) = \sum_{j=1}^{r} \hat{x}^j(k+1|k+1) \mu_j(k+1)$$

and

$$\begin{aligned}
& P(k+1|k+1) \\
=& \sum_{j=1}^{r} \mu_j(k+1) \Big\{ P^j(k+1|k+1) \\
& + \left[\hat{x}^j(k+1|k+1) - \hat{x}(k+1|k+1)\right] \\
& \cdot \left[\hat{x}^j(k+1|k+1) - \hat{x}(k+1|k+1)\right]^T \Big\}.
\end{aligned}$$

2.5 Gain matrices and estimated states

In this section, we will check that the state estimation $\hat{x}_d(k+1|k+1)$ computed in (2.7) is indeed the filtered state, not the predicted state, by showing that, in the particular case of $E_{k+1} = I$, it coincides with the filtered state of the classical Kalman filter. The explicit forms of the gain matrices $G_d(k+1)$ and $L_d(k+1)$ in (2.7) are also derived in this section. For simplicity of notation, we omit $M_i(k+1)$, and set

$$[L_{k+1}, K_{k+1}] = \left[L_{M_i(k+1)}, K_{M_i(k+1)}\right], \text{ etc,}$$

and $B_k = D_k = 0$. We are going to compute explicitly L_k, and K_k which were defined in (Moussa Ali and Zhang, 2014) as

$$[L_{k+1}, K_{k+1}] = (H_{k+1}^T R_k^{-1} H_{k+1})^{-1} H_{k+1}^T R_k^{-1}, \tag{2.8}$$

where

$$R_k = \begin{bmatrix} \Sigma_k & O \\ O & S_k \end{bmatrix}$$

and

$$\Sigma_k = A_k P_k A_k^T + Q_k.$$

Then, we have

$$\begin{aligned} & H_{k+1}^T R_k^{-1} H_{k+1} \\ = & [E_{k+1}^T, C_{k+1}^T] \begin{bmatrix} \Sigma_k^{-1} & O \\ O & S_k^{-1} \end{bmatrix} \begin{bmatrix} E_{k+1} \\ C_{k+1} \end{bmatrix} \\ = & E_{k+1}^T \Sigma_k^{-1} E_{k+1} + C_{k+1}^T S_k^{-1} C_{k+1}. \end{aligned}$$

Recall the matrix inverse formula,

$$\begin{aligned} (A + BCD)^{-1} &= A^{-1} - A^{-1}B \\ &\quad \cdot (DA^{-1}B + C^{-1})^{-1} DA^{-1} \end{aligned}$$

where A^{-1}, C^{-1}, and $(DA^{-1}B + C^{-1})$ are assumed to exist.

Using the matrix inverse formula, we obtain

$$\begin{aligned} & (H_{k+1}^T R_k^{-1} H_{k+1})^{-1} \\ = & (E_{k+1}^T \Sigma_k^{-1} E_{k+1} + C_{k+1}^T S_k^{-1} C_{k+1})^{-1} \\ = & (E_{k+1}^T \Sigma_k^{-1} E_{k+1})^{-1} \\ & -(E_{k+1}^T \Sigma_k^{-1} E_{k+1})^{-1} C_{k+1}^T \\ & \cdot \left(C_{k+1} \left(E_{k+1}^T \Sigma_k^{-1} E_{k+1} \right)^{-1} C_{k+1}^T + S_k \right)^{-1} \\ & \cdot C_{k+1} (E_{k+1}^T \Sigma_k^{-1} E_{k+1})^{-1}. \end{aligned}$$

Let us define

$$\begin{aligned} G_{k+1} &= (E_{k+1}^T \Sigma_k^{-1} E_{k+1})^{-1} C_{k+1}^T \\ &\quad \cdot (C_{k+1} (E_{k+1}^T \Sigma_k^{-1} E_{k+1})^{-1} C_{k+1}^T \\ &\quad + S_k)^{-1}. \end{aligned} \tag{2.9}$$

Further,

$$
\begin{aligned}
& (H_{k+1}^T R_k^{-1} H_{k+1})^{-1} H_{k+1}^T R_k^{-1} \\
= \; & \left[(E_{k+1}^T \Sigma_k^{-1} E_{k+1})^{-1} \right. \\
& \left. -G_{k+1} C_{k+1} (E_{k+1}^T \Sigma_k^{-1} E_{k+1})^{-1} \right] \\
& \cdot \left[E_{k+1}^T, C_{k+1}^T \right] \begin{bmatrix} \Sigma_k^{-1} & O \\ O & S_k^{-1} \end{bmatrix} \\
= \; & \left[(E_{k+1}^T \Sigma_k^{-1} E_{k+1})^{-1} E_{k+1}^T \Sigma_k^{-1} \right. \\
& -G_{k+1} C_{k+1} (E_{k+1}^T \Sigma_k^{-1} E_{k+1})^{-1} E_{k+1}^T \Sigma_k^{-1}, \\
& (E_{k+1}^T \Sigma_k^{-1} E_{k+1})^{-1} C_{k+1}^T S_k^{-1} \\
& \left. -G_{k+1} C_{k+1} (E_{k+1}^T \Sigma_k^{-1} E_{k+1})^{-1} C_{k+1}^T S_k^{-1} \right].
\end{aligned}
$$

Immediately, from (2.8), we get

$$
\begin{aligned}
L_{k+1} \; = \; & (E_{k+1}^T \Sigma_k^{-1} E_{k+1})^{-1} E_{k+1}^T \Sigma_k^{-1} \\
& -G_{k+1} C_{k+1} (E_{k+1}^T \Sigma_k^{-1} E_{k+1})^{-1} E_{k+1}^T \Sigma_k^{-1},
\end{aligned}
$$

and

$$
\begin{aligned}
K_{k+1} \; = \; & (E_{k+1}^T \Sigma_k^{-1} E_{k+1})^{-1} C_{k+1}^T S_k^{-1} \\
& -G_{k+1} C_{k+1} (E_{k+1}^T \Sigma_k^{-1} E_{k+1})^{-1} \\
& \cdot C_{k+1}^T S_k^{-1}.
\end{aligned} \tag{2.10}
$$

Substituting (2.9) into (2.10), we have

$$
\begin{aligned}
K_{k+1} \; = \; & (E_{k+1}^T \Sigma_k^{-1} E_{k+1})^{-1} C_{k+1}^T \\
& \cdot (C_{k+1} (E_{k+1}^T \Sigma_k^{-1} E_{k+1})^{-1} C_{k+1}^T \\
& + S_k)^{-1},
\end{aligned}
$$

which is exactly equals to G_{k+1} expressed by (2.9). To obtain this result, the following equalities are used

$$
\begin{aligned}
& \left(C_{k+1} (E_{k+1}^T \Sigma_k^{-1} E_{k+1})^{-1} C_{k+1}^T + S_k \right)^{-1} \\
= \; & S_k^{-1} - (C_{k+1} (E_{k+1}^T \Sigma_k^{-1} E_{k+1})^{-1} C_{k+1}^T + S_k)^{-1} \\
& \cdot C_{k+1} (E_{k+1}^T \Sigma_k^{-1} E_{k+1})^{-1} C_{k+1}^T S_k^{-1}.
\end{aligned} \tag{2.11}
$$

Indeed,

$$
\begin{aligned}
& (C_{k+1} (E_{k+1}^T \Sigma_k^{-1} E_{k+1})^{-1} C_{k+1}^T + S_k) \cdot \\
& \left((S_k^{-1} - (C_{k+1} (E_{k+1}^T \Sigma_k^{-1} E_{k+1})^{-1} C_{k+1}^T + S_k)^{-1} \right. \\
& \left. \cdot C_{k+1} (E_{k+1}^T \Sigma_k^{-1} E_{k+1})^{-1} C_{k+1}^T S_k^{-1} \right) \\
= \; & C_{k+1} \left(E_{k+1}^T \Sigma_k^{-1} E_{k+1} \right)^{-1} C_{k+1}^T S_k^{-1} + I \\
& -C_{k+1} \left(E_{k+1}^T \Sigma_k^{-1} E_{k+1} \right)^{-1} C_{k+1}^T S_k^{-1} \\
= \; & I.
\end{aligned}
$$

Hence, in particular, if we set $E_{k+1} = I$, one can get

$$\begin{aligned} L_{k+1} &= I - G_{k+1}C_{k+1}, \\ G_{k+1} &= \Sigma_k C_{k+1}^T (C_{k+1}\Sigma_k C_{k+1}^T + S_k)^{-1}, \end{aligned}$$

which is just the Kalman gain. At last, we have

$$\begin{aligned} \hat{x}(k+1) &= L_{k+1}A_k\hat{x}(k) + G_{k+1}y(k+1) \\ &= A_k\hat{x}(k) + G_{k+1}[y(k+1) \\ &\quad -C_{k+1}A_{k+1}\hat{x}(k)]. \end{aligned}$$

The associated variance can be verified similarly.

2.6 Numerical examples

Consider a hybrid descriptor stochastic system as formulated in Section 2, with (2.3) and (2.4), where $r = 3$, the mode dependent system matrices A_i, B_i, C_i, D_i, E_i, Q_i, S_i, Π and the control u are as follows:

$$\begin{aligned} A_1 &= \begin{bmatrix} 0 & 1 & 0 \\ 0.5 & 0 & 1 \\ 0 & 0 & -0.1 \end{bmatrix}, \ A_2 = \begin{bmatrix} 0.1 & 0 & 0.3 \\ 0 & 0.6 & 0 \\ 0 & 0.3 & 0.9 \end{bmatrix}, \\ A_3 &= \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ -0.05 & -0.1 & 1 \end{bmatrix}, \ B_i = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}, \ D_i = \begin{bmatrix} 0 \\ 0 \end{bmatrix}; \end{aligned}$$

$$C_1 = \begin{bmatrix} 0 & 0 & 1 \\ 1 & 1 & 0 \end{bmatrix}, \ C_2 = \begin{bmatrix} 1 & 0 & 1 \\ 1 & 1 & 0 \end{bmatrix}, \ C_3 = \begin{bmatrix} 0 & 1 & 1 \\ 1 & 1 & 0 \end{bmatrix};$$

$$E_1 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 0 \end{bmatrix}, \ E_2 = \begin{bmatrix} 2 & 0 & 3 \\ 0 & 2 & 4 \\ 0 & 0 & 0 \end{bmatrix}, \ E_3 = \begin{bmatrix} 3 & 0 & 0 \\ 0 & 4 & 0 \\ 0 & 0 & 0 \end{bmatrix};$$

$$Q_i = 0.001 \cdot \begin{bmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 3 \end{bmatrix}, \ S_i = 10 \cdot \begin{bmatrix} 4 & 0 \\ 0 & 1 \end{bmatrix}, \ i = 1, 2, 3;$$

$$\Pi = \begin{bmatrix} 0.5 & 0.1 & 0.4 \\ 0.3 & 0.6 & 0.1 \\ 0.1 & 0.3 & 0.5 \end{bmatrix}, \ u(k) = \begin{cases} 1, & \begin{cases} \text{if, } 21 \le k \le 60; \\ \text{if, } 101 \le k \le 140; \\ \text{if, } 181 \le k \le 220; \end{cases} \\ 0, & \text{otherwise.} \end{cases}$$

We will denote $x_i, i = 1, 2, 3, y_1, y_2$ as the components of vector x and y, respectively in the above figures.

The simulation runs from $k = 1$ till $k = 300$. The IMM estimator for hybrid descriptor systems proposed in this chapter is applied to the simulated system for state estimation. The Markov mode sequence can be seen in Fig. 2.1. System states and outputs are presented in Fig. 2.2, and the state estimates in Fig. 2.3.

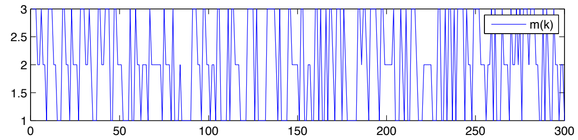

Figure 2.1: Markov mode sequence.

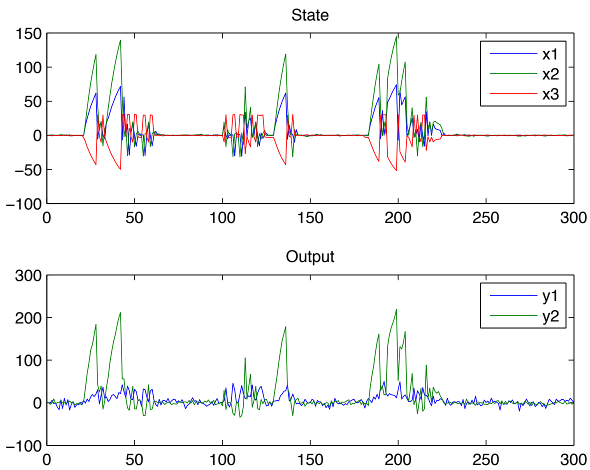

Figure 2.2: System states and outputs.

In Fig. 2.3, the results attained by the time varying descriptor system for Kalman filter (Moussa Ali and Zhang, 2014; Nikoukhah et al., 1992) are also presented in the aim of comparison. This descriptor system Kalman filter assumes that the actual mode sequence is known, which is not the case of the IMM estimator.

The results shown in Fig. 2.3 are based on one random realization of the simulated stochastic system. In order to illustrate the statistic properties of the proposed IMM estimator, 1000 random realizations have been made, and the histograms of the state estimation errors at instants $k = 50$, 140, 200, are shown in Figs. 2.4, 2.5, 2.6.

2.7 Conclusion

In this chapter the existing interacting multiple model estimator has been extended to the case of stochastic hybrid systems with modes described by descriptor equations. We have presented a numerical example to illustrate the performance of our method.

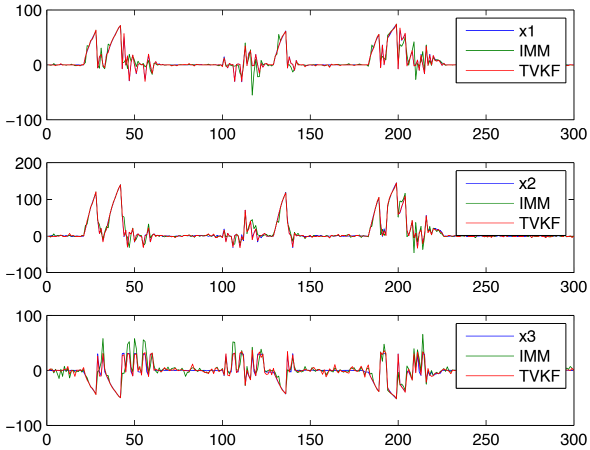

Figure 2.3: States estimates by descriptor IMM estimator and by descriptor Kalman filter (TVKF for short).

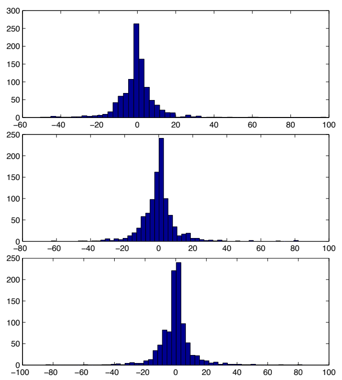

Figure 2.4: Histogram for $k = 50, 140, 200$ of the state estimation error of x_1, respectively.

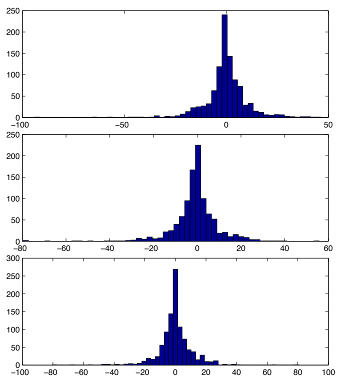

Figure 2.5: Histogram for $k = 50, 140, 200$ of the state estimation error of x_2, respectively.

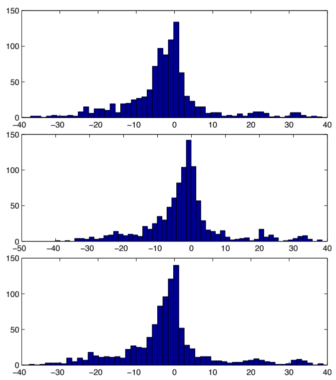

Figure 2.6: Histogram for $k = 50, 140, 200$ of the state estimation error of x_3, respectively.

Chapter 3

Hybrid system identification

3.1 Introduction

In the previous chapters it was assumed that, in a considered hybrid system, the mode models (the equations governing each mode) were known. Such mode models may be based on physical laws of the considered system, but quite often the available physical knowledge is not sufficient to build mode models. Instead, sensor data may be used to estimate mode models. Building models essentially from data is known as *system identification* in the field of automatic control. In the framework of hybrid systems with stochastic linear mode models, if the label referring to each mode is viewed as a parameter, then such hybrid systems are also known as *linear parameter varying* (LPV) systems. Assume that a sequence of input-output data and the corresponding sequence of mode transitions are known, the corresponding LPV system identification problem is studied in this chapter.

Besides the framework of hybrid systems, LPV models provide also an effective approach to handling nonlinear control systems (Toth, 2010; Mohammadpour and Scherer, 2012; Lopes dos Santos et al., 2012; Sename et al., 2013). Some successful methods for LPV system identification have been reported recently (Van Wingerden and Verhaegen, 2009; Mercere et al., 2011; Lopes dos Santos et al., 2011; Toth et al., 2012; Zhao et al., 2012; Piga et al., 2015), with various assumptions about LPV model structures. As a matter of fact, many variant LPV model structures have been proposed, and the methods developed for LPV system identification strongly depend on the particular model structures.

This chapter is focused on the identification of LPV systems in the case where the scheduling variable $p(t)$ takes values from a finite set, say $P = \{p_1, \ldots, p_m\}$. Typically each $p_i \in P$ corresponds to a working point of the considered system. It is assumed that input-output data are collected at different working points, and that the resulting LPV model will be used at the same working points. In other words, the problem of local model interpolation between working points, as studied in (De Caigny et al., 2011; De Caigny et al., 2014), is not considered in this chapter.

With the scheduling variable $p(t)$ restricted to a finite set P, the considered particular class of LPV systems will be referred to as *linear finite parameter varying* (LFPV) systems. When $p(t)$ is fixed to a particular value, the related LFPV system behaves like a linear time invariant (LTI) system. An LFPV system is thus characterized by a collection of *local LTI models*.

Given an input-output data sample and the corresponding scheduling variable sequence, in principle the LFPV system identification problem can be solved by the prediction error method

(PEM) after having chosen some parametrization of the local LTI models. However, such a solution implies solving a large optimization problem, in terms of the number of unknowns and the amount of data to be processed as a whole. Moreover, it is not easy to make a good initial guess of the model parameters before applying the PEM. Methods following such a global approach often assume some global parametric structure of the LPV system (Toth, 2010).

Alternatively, it is possible to separately estimate each local LTI model from the data collected at the related working point corresponding to a particular scheduling value $p_i \in P$. Such a local approach is attractive in practice, because the whole LFPV model can be built and completed progressively, by limiting the requirement on computing resources. Unless physical models are used, each local state-space LTI model is estimated up to an arbitrary similarity transformation. If each estimated local model is only used at the related working point (in this case the concept of LFPV model is of little interest), the indetermination of the arbitrary similarity transformation is not a problem at all. However, if the estimated collection of local LTI models is used as a global LFPV model spanning different working points, the estimated local LTI models must be *coherent*, in the sense that they are related to the true LFPV system by the *same* similarity transformation.

Global structural assumptions about the LFPV system can help to make the estimated local LTI models coherent (see examples in Section 3.3.2). Such assumptions should be based on physical insights about the considered system, otherwise they may excessively restrict the flexibility of the resulting model structure. In practice, local LTI models are often reduced order approximation a complex system, then such models have a strong black-box nature. In this case it is difficult to make the estimated local LTI models coherent.

It will be pointed out in this chapter that, *without global structural assumptions*, the local LTI models themselves do *not* contain the necessary information about the similarity transformations making them coherent. Nevertheless, it is possible to estimate these similarity transformations from input-output data under some excitation conditions, through the estimation of the states around the scheduling variable jump instants, with a method initially introduced in the framework of piecewise linear hybrid systems (Verdult and Verhaegen, 2004).

3.2 Problem statement

Let $u(t) \in \mathbb{R}^{n_u}$ and $y(t) \in \mathbb{R}^{n_y}$ be respectively the input and the output of a dynamic system at discrete time instants $t \in \mathbb{N}^* = \{0, 1, 2, \dots\}$. Assume that there exists a scheduling variable $p(t)$ defined by $p : \mathbb{N}^* \to P$, where P is a finite set

$$P = \{p_1, p_2, \dots, p_m\}, \tag{3.1}$$

such that the considered dynamic system is described by the finite dimensional state-space model

$$x(t+1) = A(p(t))x(t) + B(p(t))u(t) + w(t) \tag{3.2a}$$

$$y(t) = C(p(t))x(t) + D(p(t))u(t) + v(t) \tag{3.2b}$$

where $x(t) \in \mathbb{R}^{n_x}$ is the state vector, $A(p(t)), B(p(t)), C(p(t)), D(p(t))$ are matrices of appropriate sizes depending on $p(t) \in P$, and $w(t) \in \mathbb{R}^{n_x}, v(t) \in \mathbb{R}^{n_y}$ are white Gaussian noises with covariance matrices $Q(p(t)), R(p(t))$. Given a scheduling sequence $p(t)$, an input sequence $u(t)$, a realization of the random noises $w(t), v(t)$, and an initial state $x(0) = x_0 \in \mathbb{R}^{n_x}$, the state $x(t)$ and output $y(t)$ of the system are then fully determined.

Consider a set of consecutive time instants

$$\mathbb{T}_l^k = \{l, l+1, l+2, \dots, k\} \tag{3.3}$$

within which the value of $p(t)$ is fixed, say $p(t) = p_i$ for all $t \in \mathbb{T}_l^k$. Within these time instants, the system formulated by (3.2) is characterized by constant matrices $A(p(t)), B(p(t)), C(p(t)), D(p(t)), Q(p(t)), R(p(t))$, say

$$A_i = A(p_i),\ \ B_i = B(p_i), C_i = C(p_i),\ \text{etc.}, \tag{3.4}$$

corresponding to a linear time invariant (LTI) system

$$x(t+1) = A_i x(t) + B_i u(t) + w(t) \tag{3.5a}$$
$$y(t) = C_i x(t) + D_i u(t) + v(t) \tag{3.5b}$$

for $t \in \mathbb{T}_l^k$. Such a model will be referred to as a *local LTI model*, as it is valid only at the working point specified by $p(t) = p_i$. The notation

$$\text{LTI}_i \triangleq (A_i, B_i, C_i, D_i, Q_i, R_i) \tag{3.6}$$

will be used to denote the local LTI system model indexed by i.

In addition to the local LTI behavior of the system when the value of $p(t)$ remains constant, the model formulated in (3.2) specifies also the transition at every jump of $p(t)$. Assume that $p(k) = p_i \neq p(k+1) = p_j$. During this transition the system is no longer LTI, because of the changes of the matrices $A(p(t)), B(p(t))$ etc.. Nevertheless, according to (3.2), the value of $x(k+1)$ is determined by

$$x(k+1) = A_i x(k) + B_i u(k) + w(k). \tag{3.7}$$

The whole trajectory of $x(t)$ is thus well defined for all $t \in \mathbb{N}^*$.

As the scheduling variable $p(t)$ evolves within the finite set P, the matrix function $A(p)$ takes also values within a finite set of matrices, say $\{A_1, \dots, A_m\}$, and so do similarly the other system matrices. The characteristics of a system as formulated in (3.2) are thus fully specified by the finite sets of matrices $\{A_1, \dots, A_m\}$, $\{B_1, \dots, B_m\}$, etc., without requiring any structural assumption about the matrix functions $A(p), B(p)$, etc..

Based on the above comments, the global system described by (3.2) with $p(t) \in P$ will be referred to as a *linear finite parameter varying* (LFPV) system (the word "finite" refers to the fact that P is a finite set).

The LFPV system identification problem considered in this chapter is to estimate the finite sets of matrices $\{A_1, \dots, A_m\}$, $\{B_1, \dots, B_m\}$, etc., *solely* from the scheduling sequence $p(t)$, the input-output data $u(t), y(t)$ for $t = 0, 1, \dots, N$, and the known model order n_x.

When the local models LTI_i are separately estimated without using any global structural information about the LFPV system, each local model LTI_i can only be estimated *up to a similarity transformation*. However, to ensure consistent *global* simulation with an estimated LFPV model, the estimated local LTI models must be related to the true LFPV system by *the same* similarity transformation. This fact motivates the following definition.

Definition 1 *The set of local LTI models*

$$\{(\tilde{A}_i, \tilde{B}_i, \tilde{C}_i, \tilde{D}_i, \tilde{Q}_i, \tilde{R}_i) : \ i = 1, 2, \dots, m\} \tag{3.8}$$

constitutes a coherent *representation of the LFPV system* (3.2) *composed of a set of local LTI systems as formulated in* (3.5) *and characterized by* $(A_i, B_i, C_i, D_i, Q_i, R_i)$*, if there exits an invertible transformation matrix* $T \in \mathbb{R}^{n_x \times n_x}$ *such that, for all* $i = 1, \ldots, m$*,*

$$\tilde{A}_i = TA_iT^{-1}, \ \tilde{B}_i = TB_i, \tag{3.9a}$$
$$\tilde{C}_i = C_iT^{-1}, \ \tilde{D}_i = D_i, \tag{3.9b}$$
$$\tilde{Q}_i = TQ_iT^T, \ \tilde{R}_i = R_i. \tag{3.9c}$$

□

In practice, when a set of local LTI models are estimated from a finite data sample subject to random uncertainties, the definition of coherent local models should be understood in an approximative sense.

If some global structural assumptions of the matrix functions $A(p), B(p)$, etc. were assumed, then it would be relatively easy to make estimated local LTI models coherent, as given $A(p_i) = A_i, B(p_i) = B_i$, etc. for any $i = 1, 2, \ldots, m$, the other $A(p_j) = A_j, B(p_j) = B_j$, etc. would be partly or fully determined. Unless based on particular physical insights, such global structural assumptions may excessively restrict the flexibility of the LFPV model structure. In this chapter, the LFPV system identification problem is considered with a fully flexible LFPV model structure characterized by a set of *independent local LTI models*, as defined below.

Definition 2 *A set of local LTI models composing an LFPV system, as formulated in* (3.5)*, are* independent *if their parametrizations are such that the matrices* A_i, B_i, C_i, Q_i *characterizing* LTI$_i$ *do not imply any information about the matrices* A_j, B_j, C_j, Q_j *characterizing* LTI$_j$*, for all* $i \neq j$*, both belonging to* $\{1, 2, \ldots, m\}$*.* □

Remark 1 *In principle, it is possible to apply the prediction error method (PEM) (Ljung, 1999) to simultaneously estimate all the finite sets of matrices* $\{A_1, \ldots, A_m\}$*,* $\{B_1, \ldots, B_m\}$*, etc. by processing all the available data as a whole, but such a global approach amounts to processing a large set of data as a whole, and requires reasonably coherent initial guesses of the local LTI models. Alternatively, this chapter follows a local approach, by processing the available data in pieces segmented according to the value of* $p(t)$*, without requiring initial guesses of local LTI models.*

Remark 2 *In this chapter, the absence of any* global structural assumption *implies the LFPV model structure composed of* independent *local LTI models in the sense of Definition 2. This is a major difference from most LPV system identification methods which are based on some particular global structural assumptions, typically with a affine structure and a reduced set of parameters (Toth, 2010; Lopes dos Santos et al., 2012).*

3.3 Attempts to making local LTI estimates coherent

Compared to the global approach, the advantages of the local approach have been mentioned in Remark 1 of the previous section. However, the local approach has also a serious problem: as the local models estimated from input-output data are expressed in arbitrary state bases, the resulting local models LTI$_i$ are in general not coherent in the sense of Definition 1. A set of incoherent local LTI models cannot be used together as a whole LFPV model.

3.3.1 A general fact about independent local LTI models

Proposition 1 *Assume that the local systems* LTI_i *as formulated in* (3.5) *composing a true LFPV system are* independent *in the sense of Definition 2. Given a set of LTI models* ($\hat{A}_i, \hat{B}_i$, *etc.) such that*

$$\hat{A}_i = \hat{T}_i A_i \hat{T}_i^{-1}, \;\; \hat{B}_i = \hat{T}_i B_i, \, etc. \tag{3.10}$$

for $i = 1, \ldots, m$, *where* $\hat{T}_i \in \mathbb{R}^{n_x \times n_x}$ *are* arbitrary unknown *invertible matrices, then it is* impossible *to determine similarity transformation matrices* $\tilde{T}_i \in \mathbb{R}^{n_x \times n_x}$ *for* $i = 1, \ldots, m$, solely *based on the given LTI models* ($\hat{A}_i, \hat{B}_i$, *etc.) themselves, so that the transformed LTI models characterized by*

$$\tilde{A}_i = \tilde{T}_i \hat{A}_i \tilde{T}_i^{-1}, \;\; \tilde{B}_i = \tilde{T}_i \hat{B}_i, \, etc., \tag{3.11}$$

are coherent in the sense of Definition 1; or in other words, the involved unknowns are under-determined *by the available equations.* □

Proof. The proof consists in counting the unknowns and the available equations.

All the relevant equations are, for $i = 1, \ldots, m$,

$$\tilde{A}_i = \tilde{T}_i \hat{A}_i \tilde{T}_i^{-1}, \;\; \tilde{B}_i = \tilde{T}_i \hat{B}_i, \;\; \tilde{C}_i = \hat{C}_i \tilde{T}_i^{-1}, \;\; \tilde{Q}_i = \tilde{T}_i \hat{Q}_i \tilde{T}_i^T$$
$$\tilde{A}_i = T A_i T^{-1}, \;\; \tilde{B}_i = T B_i, \;\; \tilde{C}_i = C_i T^{-1}, \;\; \tilde{Q}_i = T Q_i T^T.$$

where $\hat{A}_i, \hat{B}_i, \hat{C}_i, \hat{Q}_i$ are known, whereas all the other involved quantities are unknowns.

To take into account the fact that there is no need to uniquely determine the transformation matrix T common to all $i = 1, \ldots, m$, eliminate $\tilde{A}_i, \tilde{B}_i, \tilde{C}_i, \tilde{Q}_i$ from these equations and replace the unknowns $\tilde{T}_i, T$ by $\breve{T}_i \triangleq T^{-1} \tilde{T}_i$ for every $i = 1, \ldots, m$. Then the remaining equations are

$$A_i = \breve{T}_i \hat{A}_i \breve{T}_i^{-1}, \;\; B_i = \breve{T}_i \hat{B}_i, \;\; C_i = \hat{C}_i \breve{T}_i^{-1}, \;\; Q_i = \breve{T}_i \hat{Q}_i \breve{T}_i^T$$

with the unknowns A_i, B_i, C_i, Q_i and $\breve{T}_i$ for $i = 1, \ldots, m$.

If $\breve{T}_i$ were known, then there would be exactly the same number of equations as the unknowns A_i, B_i, C_i, Q_i, either counted in the matrix sense or in the scalar sense. Because of the extra unknowns $\breve{T}_i$, the entire unknowns are then clearly *underdetermined* by the whole set equations. Proposition 1 is thus proved. □

3.3.2 Examples using global structural assumptions

The result of Proposition 1 may seem in contradiction with some known publications proposing methods for making local LTI models coherent. In fact, each of these existing methods assumes, explicitly or implicitly, some particular structure of the matrix-valued functions $A(p), B(p)$ etc., therefore they do not cover the case studied in this chapter. To better clarify the situations, some examples of the published methods are recalled in this subsection, by pointing out their particular global structural assumptions.

Coherent LTI models based on canonical forms

In order to make local LTI models coherent, a natural method is to find the similarity transformations leading to some canonical state-space form of the LTI systems, typically the controllable or the observable form. The idea behind this method is that the local models should be coherent when they are all transformed into the same canonical form.

For the sake of presentation simplicity, consider a single-input-single-output (SISO) LFPV system. In the controllable form, the m local models involve, for $i = 1, 2, \ldots, m$,

$$\tilde{A}_i = \begin{bmatrix} 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \cdots & \vdots \\ a_i^{(1)} & a_i^{(2)} & a_i^{(3)} & \cdots & a_i^{(n)} \end{bmatrix}, \quad \tilde{B}_i = \begin{bmatrix} 0 \\ 0 \\ 0 \\ \vdots \\ 1 \end{bmatrix}. \tag{3.12}$$

The assumption that the local LTI models in their canonical form are coherent implies that there exists a single invertible matrix $T \in \mathbb{R}^{n_x \times n_x}$ such that, for all $i = 1, 2, \ldots, m$, $\tilde{A}_i = TA_iT^{-1}$. Let $S(M)$ denote the sub matrix of M excluding its last row, then

$$S\left(TA_iT^{-1}\right) = \begin{bmatrix} 0 & 1 & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \cdots & \vdots \\ 0 & 0 & 0 & \cdots & 1 \end{bmatrix}, \quad \forall i = 1, 2, \ldots, m. \tag{3.13}$$

The assumption that $S\left(TA_iT^{-1}\right)$ are equal to the same particular matrix for all $i = 1, 2, \ldots, m$ is indeed a strong global structural assumption about $A(p)$.

Coherent LTI models based on the observability matrix

In (De Caigny et al., 2014) another method is proposed to make local LTI models coherent. The class of systems considered in (De Caigny et al., 2014) is more general than that of this chapter. When applied to the LFPV systems as considered in this chapter, this method consists in finding different similarity transformations so that the m transformed LTI models all have the same observability matrix.

This method is based on the assumption that the local LTI systems composing the considered LFPV system all have the same observability matrix, or more explicitly,

$$C_iA_i^s = C_jA_j^s \tag{3.14}$$

for all $i, j \in \{1, 2, \ldots, m\}$ and $s \in \{0, 1, \ldots, n_x - 1\}$. This is clearly also a global structural assumption about $A(p), C(p)$.

Notice that this observability matrix-based method is incompatible with the previously presented canonical form-based method, as in general the local LTI models in their canonical form do not have the same observability matrix. This incompatibility between the two "natural" methods confirms the fact that there is no generally natural global structural assumption for making estimated local LTI models coherent.

3.4 Data-based transformations for coherent LTI models

It was shown in the previous section that, in the local approach to LFPV system identification, it is impossible to determine similarity transformations to make independent local LTI models coherent *solely from the estimated local LTI models themselves*. It is thus necessary to make use of other information, not contained in the local LTI models. By excluding global structural assumptions, the only information left seems the available input-output data and the scheduling variable sequence. These data have already been used for the estimation of local LTI models, but they can also provide more information, for making the local models coherent.

3.4.1 The case of $m = 2$

For the ease of presentation, let us first consider the case of a LFPV system with only 2 local LTI systems ($m = 2$). The more general case will be considered later.

Assume that $p(t)$ changes from p_1 to p_2 at $t = k$, or more accurately, $p(t) = p_1$ for $t \in \mathbb{T}_l^k = \{l, l+1, l+2, \ldots, k\}$ and $p(t) = p_2$ for $t \in \mathbb{T}_{k+1}^q = \{k+1, k+2, \ldots, q\}$. Assume further that the two corresponding input-output data segments are informative enough so that two local LTI models of order n_x can be estimated from them, with any of the classical LTI system identification methods (Ljung, 1999). These estimated local models will be denoted by

$$\mathrm{LTI}_1 = (\hat{A}_1, \hat{B}_1, \mathrm{etc.}) \text{ and } \mathrm{LTI}_2 = (\hat{A}_2, \hat{B}_2, \mathrm{etc.}). \tag{3.15}$$

With the two estimated LTI models and the input-output data, the state sequence $x(t)$ for $t \in \mathbb{T}_l^k$ and $t \in \mathbb{T}_{k+1}^q$ can be estimated with different approaches. For some subspace identification methods, the state sequence estimate $\hat{x}(t)$ is a co-product (Van Overschee and De Moor, 1996). With the PEM (Ljung, 1999), the initial state estimate (for each of $\mathbb{T}_l^k, \mathbb{T}_{k+1}^q$) is also provided with the estimated LTI model, then the whole state sequence, separately for $t \in \mathbb{T}_l^k$ and $t \in \mathbb{T}_{k+1}^q$, can be estimated by the Kalman filter.

Here the final state estimate with LTI_1 for $t \in \mathbb{T}_k^k$, namely $\hat{x}_1(k)$, and the initial state estimate with LTI_2 for $t \in \mathbb{T}_{k+1}^q$, namely $\hat{x}_2(k+1)$, are of particular interests.

If LTI_1 and LTI_2 were coherent, then according to (3.7), the equality

$$\hat{x}_2(k+1) = \hat{A}_1 \hat{x}_1(k) + \hat{B}_1 u(k) \tag{3.16}$$

would hold, up to random estimation errors of $\hat{x}_1(k)$ and $\hat{x}_2(k+1)$ and the noise $w(k)$. In general, of course, LTI_1 and LTI_2 are not coherent and equality (3.16) does not hold.

Assume that after applying some similarity transformation $T_{1,2}$ to LTI_1, the transformed LTI model will be coherent with LTI_2. *The remaining part of this subsection* is for the purpose of introducing a method for the estimation of $T_{1,2}$.

As the application of the transformation matrix $T_{1,2}$ makes LTI_1 coherent with LTI_2, the incorrect equality (3.16) should be replaced by

$$\hat{x}_2(k+1) = T_{1,2} \hat{A}_1 T_{1,2}^{-1} T_{1,2} \hat{x}_1(k) + T_{1,2} \hat{B}_1 u(k) \tag{3.17}$$

$$= T_{1,2} \hat{A}_1 \hat{x}_1(k) + T_{1,2} \hat{B}_1 u(k). \tag{3.18}$$

Define

$$\hat{x}_1(k+1) \triangleq \hat{A}_1 \hat{x}_1(k) + \hat{B}_1 u(k) \tag{3.19}$$

as an estimate of $x(k+1)$ from LTI_1, then (3.18) is rewritten as

$$\hat{x}_2(k+1) = T_{1,2}\hat{x}_1(k+1). \tag{3.20}$$

As $\hat{x}_2(k+1)$ and $\hat{x}_1(k+1)$ can be both estimated from available data with the estimated local LTI models, they provide information about $T_{1,2}$ through (3.20).

The matrix $T_{1,2}$ has $n_x \times n_x$ unknown entries, but (3.20) contains only n_x scalar equations. It is not yet sufficient to determine $T_{1,2}$ in order to make the two local LTI models coherent.

Now assume that in the available data set there are more than one jumps of $p(t)$ within $P = \{p_1, p_2\}$. If each of these jumps leads to an equation on $T_{1,2}$ similar to (3.20), then it is possible to determine $T_{1,2}$ from these equations.

Assume that $k^{(0)}, k^{(1)}, k^{(2)}, \dots, k^{(s+1)}$ are jump instants interlacing $p(t) = p_1$ and $p(t) = p_2$, such that for every

$$t \in \mathbb{T}_{k^{(j)}+1}^{k^{(j+1)}} = \{k^{(j)}+1, k^{(j)}+2, \dots, k^{(j+1)}\}, \tag{3.21}$$

$p(t) = p_1$ if j is an odd number, and $p(t) = p_2$ if j is an even number.

If it was chosen to estimate separately one local LTI model from the data within each $\mathbb{T}_{k^{(j)}+1}^{k^{(j+1)}}$, those corresponding to odd numbers j would in principle all describe the same local LTI model, but in different state bases. To avoid this trouble, a simple idea is to treat all these data for $t \in \mathbb{T}_{k^{(j)}+1}^{k^{(j+1)}}$ with odd numbers j as a whole *multi-experiment* data set (Ljung, 2014), composed of different experiments corresponding to different $\mathbb{T}_{k^{(j)}+1}^{k^{(j+1)}}$. A single LTI model, namely LTI_1, is then estimated from this multi-experiment data set. Similarly, a single model LTI_2 is estimated from the multi-experiment data set corresponding to even numbers j. A similar method was proposed in (Verdult and Verhaegen, 2004) in the framework of subspace system identification.

With the two estimated LTI models $\mathrm{LTI}_1, \mathrm{LTI}_2$ and the available data subsets, after every jump instant $k^{(j)}$ for $j = 1, 2 \dots, s$, two state estimates $\hat{x}_1(k^{(j)}+1)$ and $\hat{x}_2(k^{(j)}+1)$ are computed, respectively with LTI_1 and LTI_2.

When j is an even number, $\hat{x}_1(k^{(j)}+1)$ is computed from the final state estimate within $\mathbb{T}_{k^{(j-1)}+1}^{k^{(j)}}$ in a way similar to (3.19), whereas $\hat{x}_2(k^{(j)}+1)$ is simply the initial state estimate within $\mathbb{T}_{k^{(j)}+1}^{k^{(j+1)}}$.

When j is an odd number, $\hat{x}_1(k^{(j)}+1)$ is simply the initial state estimate within $\mathbb{T}_{k^{(j)}+1}^{k^{(j+1)}}$, whereas the computation of $\hat{x}_2(k^{(j)}+1)$ is made in a way similar to (3.19).

With these results, the vector equation (3.20) is generalized to the matrix equation

$$\begin{aligned}&[\hat{x}_2(k^{(1)}+1), \ \cdots, \ \hat{x}_2(k^{(s)}+1)] \\ &= T_{1,2}[\hat{x}_1(k^{(1)}+1), \ \cdots, \ \hat{x}_1(k^{(s)}+1)].\end{aligned} \tag{3.22}$$

Assume that the matrix $\hat{X}_1 \triangleq [\hat{x}_1(k^{(1)}+1), \ \cdots, \ \hat{x}_1(k^{(s)}+1)]$ has full row rank, then $T_{1,2}$ can be estimated, by simply inverting $\hat{X}_1$ if it is a square matrix, otherwise by solving (3.22) for $T_{1,2}$ in the least squares sense.

Remark that, for estimating two LTI models from the two multi-experiment data sets (Ljung, 2014) formed from the available data set, it is sufficient to assume that each of these two multi-experiment data sets are informative enough, instead of assuming that each of the data segment corresponding to $\mathbb{T}_{k^{(j)}+1}^{k^{(j+1)}}$ is informative enough.

3.4.2 The case of $m \geq 2$

Now consider the case of $m \geq 2$. The m local models LTI_i are first estimated from the available data, so is the whole state sequence.

For each pair of indexes $i \neq j$, both belonging to $\{1, 2, \dots, m\}$, it is possible to estimate a transformation matrix $T_{i,j}$ to make the estimated LTI_i and LTI_j coherent, by applying the method described in the previous subsection to appropriately selected data segments.

If a transformation matrix was estimated for each pair of the m estimated local LTI models, there would be $m(m+1)/2$ such estimated transformation matrices, but they are not all necessary.

For example, assume that the available data set allows the estimation of the $m-1$ transformation matrices $T_{1,2}, T_{2,3}, \dots, T_{m-1,m}$, then by applying the matrix product $T_{m-1,m} \cdot T_{m-2,m-1} \cdots T_{2,3} \cdot T_{1,2}$ as a single transformation matrix to LTI_1, the transformed model is coherent with LTI_m. Similarly, the other local models $\mathrm{LTI}_2, \dots, \mathrm{LTI}_{m-1}$ are then also made coherent with LTI_m. Therefore, $m-1$ estimated transformation matrices are sufficient to make the m local LTI models coherent.

Generally, this method is based on the following assumptions.

A1. For each $p_i \in P$, the data subset

$$Z_i = \{(u(t), y(t), p(t)) : \ t = 0, 1, \dots, N, \ p(t) = p_i\}$$

forms a multi-experiment data set sufficiently informative for the estimation of a state-space LTI model of order n_x, namely LTI_i.

A2. There are sufficient jumps in the scheduling sequence $p(t)$ such that pairs of adjacent data segments allow the estimation of $m-1$ transformation matrices $T_{i,j}$, each linking two local models LTI_i and LTI_j, so that the entire m local LTI models are linked together directly or indirectly by the of $m-1$ transformation matrices.

The complete algorithm for estimating the $m-1$ transformation matrices is well described in (Verdult and Verhaegen, 2004).

3.5 Numerical examples

Let us consider the case of an LFPV system composed of 5 single input-single output local LTI systems ($m = 5$). Each local LTI system has two conjugate complex poles with modulus randomly drown in the interval $[0.8, 0.9]$, and a real zero randomly drawn in $[-0.8, 0.8]$. The models of the local LTI system are converted to the state-space form, each with an arbitrary state basis, before being linked together to form an LFPV system model, which is then used to data simulation.

The *estimation data set* is generated with a piecewise-constant scheduling variable sequence $p(t) \in P = \{1, 2, 3, 4, 5\}$ as shown in Figure 3.1, and an independent random input sequence uniformly distributed in $[0, 1]$. The output is then simulated with the randomly generated LFPV model as described above. No state noise is added during the simulation, but a white Gaussian noise with standard deviation 0.01 is added to the simulated output.

From one randomly generated estimation data set, 5 independent local LTI models of second order ($n_x = 2$) are first estimated with the PEM (Ljung, 1999), which are then transformed into a coherent set of LTI models with the method presented in this chapter. The LFPV model composed of the 5 coherent local LTI models is then tested on an evaluation data set, which is generated with an *independent* random scheduling variable sequence $p(t)$ equally distributed within $P = \{1, 2, 3, 4, 5\}$, which is *radically different* from the one used in the estimation data generation, as shown in Figure 3.2. The output simulated with the estimated LFPV model is then compared with the true output in Figure 3.2, where the two curves (blue and green) are hardly distinguishable. The model fit (percentage of the output variance explained by the estimated model) in this example is 0.9537.

The above result is only based on one randomly generated estimation data set and one validation set. The same simulation is then randomly repeated 1000 times, with different random realizations of the LFPV system composed of local LTI models and different random noise realizations. Based on these results, the empirical mean of the model fit is 0.9639, and standard deviation 0.0306.

In order to illustrate the importance of making estimated local LTI models coherent, for the same random realization as shown in Figure 3.2, the output simulated with the estimated local LTI models before their "coherentization" is compared to the true output in Figure 3.3. The benefit of coherent models is then clear.

3.6 Conclusion

For the purpose of LPV system identification, a data-based method has been proposed in this chapter to make independently estimated local LTI models coherent, without making any global structural assumption about the LPV system, based on an algorithm initially introduced in (Verdult and Verhaegen, 2004).

As a final note, because the local model interpolation problem, as studied in (De Caigny et al., 2011; De Caigny et al., 2014), is not considered in this chapter, it does not matter if the elements p_i of the finite set P are scalar real values or any other mathematical objects. Throughout this chapter, the values p_i could have been replaced by their indexes i, defined in any arbitrary order.

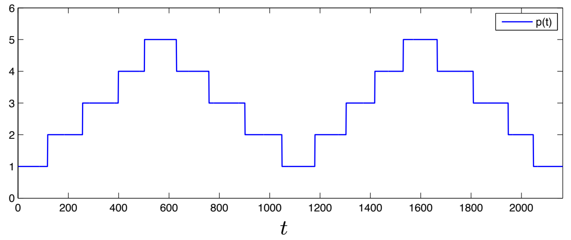

Figure 3.1: Scheduling sequence $p(t)$ of the estimation data set.

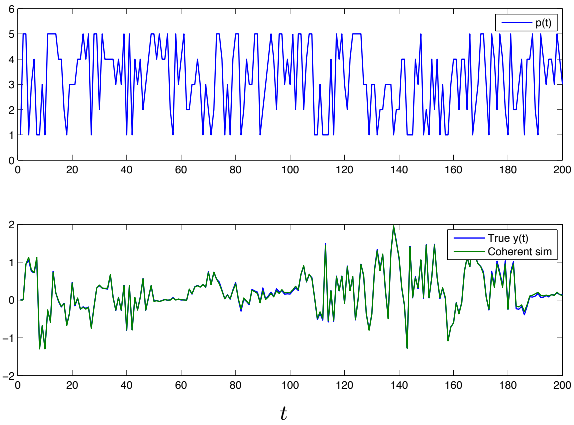

Figure 3.2: Top: scheduling sequence $p(t)$ of the validation data set. Bottom: comparison between the true output (blue) and the output simulated with *coherent* local LTI models (green).

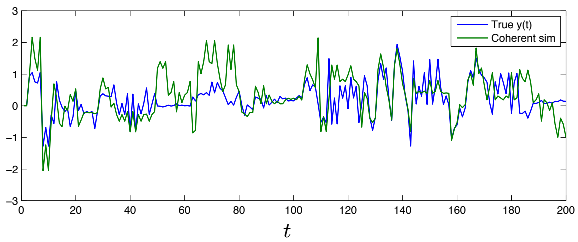

Figure 3.3: Comparison between the true output (blue) and the output simulated with *non coherent* local LTI models (green).

References

Bar-Shalom, Y. and Fortmann, T. (1988). *Tracking and data association*. Academic Press.

Bar-Shalom, Y., Li, X. R., and Kirubarajan, T. (2001). *Estimation with Applications to Tracking and Navigation*. Wiley.

Basseville, M. and Nikiforov, I. (1993). *Detection of abrupt changes: theory and application*. Prentice Hall.

Belkhiat, D. E., Messai, N., and Manamanni, N. (2011). Design of robust fault detection observers for discrete-time linear switched systems. In *IFAC World Congress*, Milano.

Bemporad, A., Mignone, D., and Morari, M. (1999). Moving horizon estimation for hybrid systems and fault detection. In *American Control Conference*, San Diego.

Blanke, M., Kinnaert, M., Lunze, J., and Staroswiecki, J. (2003). *Diagnosis and Fault-Tolerant Control*. Springer.

Blom, H. and Bar-Shalom, Y. (1988). The interacting multiple model algorithm for systems with markovian switching coefficients. *IEEE Transactions on Automatic Control*, 33(8):780–783.

Blom, H. A. and Lygeros, J., editors (2006). *Stochastic Hybrid Systems – Theory and Safety Critical Applications*. Springer, Berlin, Heidelberg, New York.

Chen, J. and Patton, R. J. (1999). *Robust Model-Based Fault Diagnosis for Dynamic Systems*. Kluwer, Boston, USA.

Chen, R., Mingori, D., and Speyer, J. (2003). Optimal stochastic fault detection filter. *Automatica*, 39(3):377–390.

Chen, R. and Speyer, J. (2000). A generalized least-squares fault detection filter. *Int. Jal Adaptive Control Signal Processing*, 14(7):747–757.

Chung, W. and Speyer, J. (1998). A game theoretic fault detection filter. *IEEE Trans. Automatic Control*, 43(2):143–161.

Cinquemani, E., Micheli, M., and Picci, G. (2004). Fault detection in a class of stochastic hybrid systems. In *Conference on Decision and Control*, Atlantis.

De Caigny, J., Camino, J. F., and Swevers, J. (2011). Interpolation-based modeling of MIMO LPV systems. *IEEE Trans. on Control System Technology*, 19(1):46–63.

De Caigny, J., Pintelon, R., Camino, J., and Swevers, J. (2014). Interpolated modeling of LPV systems. *IEEE Trans. on Control System Technology*, 22(6):2232 – 2246.

Ding, S. X. (2008). *Model-Based Fault Diagnosis Techniques - Design Schemes Algorithms and tools*. Springer-Verlag.

Doucet, A., Gordon, N., and Krishnamurthy, V. (2001). Particle filters for state estimation of jump markov linear systems. *IEEE Transactions on Signal Processing*, 49:613–624.

Ferrari-Trecate, G., Mignone, D., and Morari, M. (2002). Moving horizon estimation for hybrid systems. *IEEE Transactions on Automatic Control*, 47(10):1663–1676.

Gao, Z. and Wang, H. (2006). Descriptor observer approaches for multivariable systems with measurement noises and application in fault detection and diagnosis. *Systems & Control Letters*, 55(4):304–313.

Gertler, J. J. (1998). *Fault detection and diagnosis in engineering systems*. Marcel Dekker, Inc., New York, Basel, Hong Kong.

Guo, J.-B., Ji, D.-F., Du, S.-H., Zeng, S.-K., and Sun, B. (2013). Fault diagnosis of hybrid systems using particle filter based hybrid estimation algorithm. *Chemical Engineering Transactions*, 33:145–150.

Guyader, A. and Zhang, Q. (2003). Adaptive observer for discrete time linear time varying systems. In *13th IFAC/IFORS Symposium on System Identification (SYSID)*, Rotterdam.

Hanlon, P. D. and Maybeck, P. S. (2000). Multiple-model adaptive estimation using a residual correlation kalman filter bank. *IEEE Transactions on Aerospace and Electronic Systems*, 36(2):393–406.

Hwang, I., Balakrishnan, H., and Tomlin, C. (2006). State estimation for hybrid systems: applications to aircraft tracking. *IEE Proceedings – Control Theory and Applications*, 153(5):556–566.

Isermann, R. (2006). *Fault-Diagnosis System*. Springer, Berlin.

Koenig, D. and Mammar, S. (2002). Design of proportional-integral observer for unknown input descriptor systems. *IEEE Trans. on Automatic Control*, 47(12):2057–2062.

Korbicz, J. (2004). *Fault Diagnosis: Models, Artificial Intelligence, Applications*. Springer-Verlag GmbH.

Koutsoukos, X., Kurien, J., and Zhao, F. (2002). Monitoring and diagnosis of hybrid systems using particle filtering method. In *Symposium on Mathematical Theory of Networks and Systems*.

Li, X., Zhang, Q., and Su, H. (2011). An adaptive observer for joint estimation of states and parameters in both state and output equations. *International Journal of Adaptive Control and Signal Processing*, 25(9):765–854.

Ljung, L. (1999). *System Identification – Theory for the User*. Prentice-Hall, 2nd edition edition.

Ljung, L. (2014). System identification toolbox User's Guide. MathWorks Matlab documentation. R2014b.

Lopes dos Santos, P., Azevedo-Perdicoúlis, T.-P., Ramos, J., Martins de Carvalho, J., Jank, G., and Milhinhos, J. (2011). An LPV modeling and identification approach to leakage detection in high pressure natural gas transportation networks. *IEEE Trans. on Control System Technology*, 19(1):77–92.

Lopes dos Santos, P., Perdicoulis, T. P. A., Novara, C., Ramos, J. A., and Rivera, D. E., editors (2012). *Linear Parameter-Varying System Identification*. World Scientific Publishing Company, New Jersey.

Lunze, J. and Lamnabhi-Lagarrigue, F., editors (2009). *Handbook of Hybrid Systems Control – Theory, Tools, Applications*. Cambridge University Press, Cambridge, New York, Melbourne.

Marx, B., Koenig, D., and Georges, D. (2004). Robust fault-tolerant control for descriptor systems. 49(10):1869–1876.

Mercere, G., Palsson, H., and Poinot, T. (2011). Continuous-time linear parameter-varying identification of a cross: A local approach. *IEEE Trans. on Control System Technology*, 19(1):64–76.

Mohammadpour, J. and Scherer, C. W., editors (2012). *Control of Linear Parameter Varying Systems with Applications*. Springer, New York.

Moussa Ali, A. and Zhang, Q. (2014). An innovations approach to fault diagnosis in linear time-varying descriptor systems. In *European Control Conference*, Strasbourg, France.

Nikoukhah, R., Willsky, A., and Levy, B. (1992). Kalman filtering and riccati equations for descriptor systems. *IEEE Trans. on Automatic Control*, 37(9):1325 –1342.

Piga, D., Cox, P. B., Toth, R., and Laurain, V. (2015). Identification of LPV models from noise-corrupted scheduling parameter and output observations. *Automatica*, To appear.

Seah, C. E. and Hwang, I. (2009). State estimation for stochastic linear hybrid systems with continuous-state-dependent transitions: An imm approach. *IEEE Transactions on Aerospace and Electronic Systems*, 5(1):376–392.

Sename, O., Gaspar, P., and Bokor, J., editors (2013). *Robust Control and Linear Parameter Varying Approaches*. Springer, Berlin.

Simani, S., Fantuzzi, C., Patton, R., and Patton, R. (2003). *Model-based Fault Diagnosis in Dynamic Systems Using Identification Techniques*. Springer.

Toth, R. (2010). *Modeling and Identification of Linear Parameter-Varying Systems*. Springer.

Toth, R., Laurain, V., Gilson, M., and Garnier, H. (2012). Instrumental variable scheme for closed-loop LPV model identification. *Automatica*, 48(9):2314–2320.

Van Overschee, P. and De Moor, B. (1996). *Subspace Identification for Linear Systems: Theory, Implementation Applications*. Springer-Verlag.

Van Wingerden, J.-W. and Verhaegen, M. (2009). Subspace identification of bilinear and LPV systems for open- and closed-loop data. *Automatica*, 45(2):372–381.

Verdult, V. and Verhaegen, M. (2004). Subspace identification of piecewise linear systems. In *43rd IEEE Conference on Decision and Control*, Atlantis, Paradise Island, Bahamas.

Wang, D., Yu, M., Low, C. B., and Arogeti, S. (2013). *Model-based Health Monitoring of Hybrid Systems*. Springer, New York, Heidelberg.

Wang, W., Wang, L., Zhou, D., and Zhou, K. (2007). Robust state estimation and fault diagnosis for uncertain hybrid nonlinear systems. *Nonlinear Analysis: Hybrid Systems*, 1(1):2–15.

Yeu, T. and Kawaji, S. (2001). Fault detection and isolation for descriptor systems using sliding mode observer. In *40th IEEE Conference on Decision and Control (CDC)Control*, pages 596–597.

Zhang, Q. (2002). Adaptive observer for multiple-input-multiple-output (MIMO) linear time varying systems. *IEEE Trans. on Automatic Control*, 47(3):525–529.

Zhang, Q. and Basseville, M. (2014). Statistical detection and isolation of additive faults in linear time-varying systems. *Automatica*, 50(10):2527–2538.

Zhao, Y., Huang, B., Su, H., and Chu, J. (2012). Prediction error method for identification of LPV models. *Journal of Process Control*, 22(1):180–193.

Zhong, M., Ding, S., and Ding, E. (2010). Optimal fault detection for linear discrete time-varying systems. *Automatica*, 46(8):1395–1400.