



Institut National
Polytechnique de Lorraine



École doctorale IAEM Lorraine
Département de Formation Doctorale en Automatique

Centre de Recherche
en Automatique de Nancy



Estimation d'état et diagnostic de systèmes à commutation par filtrage multi-modèle

THÈSE

présentée et soutenue publiquement le 08 décembre 2006

pour l'obtention du

Doctorat de l'Institut National Polytechnique de Lorraine
(spécialité automatique et traitement du signal)

par

Abdelfettah HOCINE

Composition du jury

<i>Président :</i>	M.OULADSINE	Professeur à l'Université d'Aix-Marseille
<i>Rapporteurs :</i>	D. LEFEBVRE	Professeur à l'Université du Havre
	M. VERGÉ	Professeur à l'ENSAM, Paris
<i>Examineurs :</i>	D. MAQUIN	Professeur à l'INPL, Nancy (directeur de thèse)
	J. RAGOT	Professeur à l'INPL, Nancy (co-directeur)
	D. THEILLIOL	Professeur à l'Université de Nancy 1

Remerciements

Le travail présenté dans ce mémoire a été effectué au sein du Centre de Recherche en Automatique de Nancy (CRAN - UMR 7039) au sein du groupe thématique "Sûreté de Fonctionnement et Diagnostic des Systèmes" sous la direction de Messieurs les professeurs Didier Maquin et José Ragot.

Je tiens à leur témoigner ma profonde gratitude pour l'accueil, le suivi et l'aide précieuse qu'ils m'ont apportés tout au long de ce travail. Je leur suis très reconnaissant pour la confiance qu'ils m'ont témoignée tout au long de mes travaux de recherche.

J'exprime ma gratitude à Messieurs les professeurs Dimitri Lefebvre et Michel Vergé d'avoir accepté de rapporter sur mon mémoire et pour l'intérêt qu'ils ont voulu porter à ce travail. Leur lecture approfondie du mémoire, leurs remarques et interrogations judicieuses m'ont été très précieuses.

Je remercie également Monsieur le professeur Mustapha Ouladsine d'avoir accepté d'examiner ce travail et d'assurer la présidence du jury. Mes remerciements s'adressent aussi à Monsieur Didier Theilliol pour sa participation au jury et pour ses remarques fructueuses.

Je tiens particulièrement à remercier tous les membres du CRAN et tous les doctorants pour leur sympathie et l'ambiance chaleureuse qu'ils ont su entretenir tout au long de mon séjour parmi eux et tout particulièrement Marjorie Schwartz pour sa constante disponibilité.

Enfin, mes remerciements vont à tous ceux qui m'ont soutenu ou qui, d'une manière ou d'une autre, ont contribué à l'élaboration de ce travail.

*Je dédie cette thèse
à mes parents adorés
sans lesquels je ne suis rien
à ma femme qui m'a soutenu et à ma sœur*

Table des matières

Table des figures **ix**

Table des notations

Références personnelles	1
--------------------------------	----------

Introduction générale

Chapitre 1

Méthodes d'estimation d'état par multi-modèle pour les systèmes à commutation
--

1.1	Introduction	9
1.2	Filtre de Kalman	11
1.3	Système à commutation markovienne à temps discret	13
1.3.1	Chaînes de Markov	13
1.3.2	Chaînes de Markov à espace d'états discret	13
1.3.3	Description des systèmes à commutation markovienne	14
1.4	Estimation d'état par multi-modèle	15
1.4.1	Description générale	15
1.4.2	Approche multi-modèle pour le cas non commutant	18
1.4.3	Approche multi-modèle pour les systèmes à commutation	21
1.4.4	Estimateur pseudo-bayésien généralisé du premier ordre	25
1.4.5	Estimateur pseudo-bayésien généralisé du deuxième ordre	29
1.4.6	Estimateur IMM (Interacting Multiple Model Estimator)	33
1.5	Conclusion	39

Chapitre 2

Diagnostic des systèmes à commutation

2.1	Introduction	43
2.2	Principes fondamentaux du diagnostic	44
2.2.1	Surveillance	44
2.2.2	Diagnostic	44
2.2.3	Prise de décision	46
2.2.4	Performance d'une procédure de diagnostic	47
2.3	Estimateurs multi-modèles pour le diagnostic	48
2.3.1	Modèle de défaut	49
2.3.2	Détection de défaut	51
2.3.3	Construction de la matrice de Markov	53
2.3.4	Évaluation des performances de la méthode	54
2.4	Exemple de simulation	55
2.5	Conclusion	60

Chapitre 3

Estimation de la matrice de Markov

3.1	Introduction	64
3.2	Formulation du problème	65
3.3	Fonction de densité de probabilité de la matrice de Markov	67
3.4	Algorithme quasi bayésien	71
3.4.1	Estimation quasi bayésienne des paramètres d'une distribution mélange	72
3.4.2	Estimateur quasi bayésien de la MPT	79
3.5	Algorithme par intégration numérique	84
3.6	Conclusion	87

Chapitre 4

Observateurs à mémoire finie pour les systèmes à commutation markovienne

4.1	Introduction	91
4.2	Observateur à mémoire finie	92
4.2.1	Condition d'existence de l'observateur	94
4.2.2	Taille de l'horizon	94

4.2.3	Propriétés de l'observateur	95
4.3	Observateur à mémoire finie avec entrée inconnue	95
4.4	Observateur à mémoire finie pour les systèmes à commutation markovienne	97
4.4.1	Développement de la méthode	99
4.4.2	Exemple d'application	101
4.5	Extension de la méthode pour les systèmes à entrée inconnue	106
4.6	Conclusion	107

Chapitre 5

Observateur mixte pour les systèmes à commutation markovienne

5.1	Introduction	113
5.2	Observateur à mémoire finie	113
5.3	Observateur mixte	115
5.4	Exemple	117
5.5	Observateur mixte pour les systèmes à commutation markovienne	121
5.5.1	Développement de la méthode	121
5.5.2	Exemple d'application	123
5.6	Conclusion	126

Chapitre 6

Estimateur multi-modèle structuré

6.1	Introduction	129
6.2	Formulation du problème	130
6.3	Matrice des probabilités de transition	130
6.3.1	Matrice des probabilités de transition entre deux sous-ensembles de modèles	131
6.3.2	Matrice des probabilités de transition au sein d'un sous-ensemble de modèles	132
6.4	Procédure d'estimation	133
6.4.1	Sous-estimateur	133
6.4.2	Estimateur global	136
6.4.3	Choix du modèle actif	137
6.5	Exemple	141
6.6	Conclusion	147

Conclusion et perspectives

Bibliographie	153
	165
	167

Table des figures

1.1	Système à commutation markovienne	15
1.2	Estimateur multi-modèle pour les systèmes non commutants	20
1.3	Estimateur multi-modèle optimal pour les systèmes commutants	24
1.4	Estimateur GPB1	27
1.5	Estimateur GPB2	31
1.6	Estimateur IMM	36
2.1	Entrée u du système	57
2.2	Courant i du système	58
2.3	Vitesse angulaire ω	58
2.4	Estimée du courant i du système	58
2.5	Estimée de la vitesse angulaire ω	58
2.6	Probabilités d'activation μ_i	59
3.1	Procédure d'estimation de la matrice de Markov	68
3.2	Densité mélange	76
3.3	Convergence de θ_i	77
3.4	Densité mélange	77
3.5	Convergence de θ_i	78
3.6	Classification des données	78
3.7	Signaux d'entrée-sortie et signaux de commutations	83
3.8	Convergence de π_{ii}	84

3.9	Signaux d'entrée-sortie et signaux de commutation	85
3.10	Convergence de π_{ij}	86
3.11	Signaux d'entrée-sortie et signaux de commutation	87
3.12	Convergence de π_{ij}	88
4.1	Entrées sorties du système	98
4.2	Entrée inconnue estimée en présence de bruit	98
4.3	Entrée inconnue estimée en présence de bruit	99
4.4	Entrées sorties du système	103
4.5	Probabilité d'activation du modèle 1	104
4.6	Probabilité d'activation du modèle 2	104
4.7	Probabilité d'activation du modèle 3	105
4.8	Probabilité d'activation du modèle 1 avec bruit	107
4.9	Probabilité d'activation du modèle 2 avec bruit	108
4.10	Probabilité d'activation du modèle 3 avec bruit	108
4.11	Estimation de l'entrée inconnue en l'absence de bruit	109
4.12	Estimation de l'entrée inconnue en présence de bruit	109
5.1	Évolution du terme $\ D\ \lambda_{max}(P + I)/\sqrt{\lambda_{min}(Q)}$ en fonction du paramètre τ	118
5.2	Erreur d'estimation pour les deux composantes du vecteur d'état et norme de $\ e(k)\ $	119
5.3	Variation du paramètre τ en fonction de la longueur de la fenêtre	120
5.4	Probabilité d'activation du modèle 1	124
5.5	Probabilité d'activation du modèle 2	125
5.6	Probabilité d'activation du modèle 3	125
6.1	Sous-estimateur	135
6.2	Estimateur global	138
6.3	Entrée-sorties du système	143
6.4	Probabilités d'activation (méthode directe)	144
6.5	Probabilités d'activation (méthode hiérarchisée)	145

6.6	Probabilités d'activation méthode mixte	146
-----	---	-----

Table des figures

Table des notations

BDI	bonne détection et isolation
FA	fausse alarme
FDI	fault detection and isolation
GPB	generalized pseudo-Bayesian estimator
GPB1	GPB du premier ordre
GPB2	GPB du deuxième ordre
GPS	global positioning system
IID	isolation incorrecte du défaut
IMM	interacting multiple model estimator
SDH	systèmes dynamiques hybrides
OMF	observateur à mémoire finie
LMI	linear matrix inequalities
MM	multi-modèle
MPT	matrice des probabilités de transition
MV	maximum de vraisemblance
ND	non-détection
SMC	système à commutations markoviennes
SMM	static multiple model
TMD	temps moyen de détection

Références personnelles

Reuves avec comité de lecture

- A. Hocine, D. Maquin et J. Ragot, *Identification de systèmes à commutation*, e-STA, Sciences et Technologies de l'Automatique, revue électronique de la SEE, vol. 2, n° 3, 2005. Article sélectionné des Journées Doctorales, Modélisation, Analyse et Conduite des Systèmes dynamique, JD-MACS, Lyon, France, 5-7 septembre 2005.
- A. Hocine, D. Maquin et J. Ragot, *Observateur à mémoire finie pour les systèmes à commutations : application à la détection de défauts*, e-STA, Sciences et Technologies de l'Automatique, revue électronique de la SEE, vol. 1, n° 4, 2004. Article sélectionné de la Conférence Internationale Francophone d'Automatique, CIFA'2004, Douz, Tunisie, 22-24 novembre 2004.

Congrès internationaux avec comité de lecture et actes

- A. Hocine, M. Chadli, D. Maquin et J. Ragot, *A discrete-time Sliding Window Observer for Markovian Switching System*, 45rd IEEE Conference on Decision and Control, San Diego, CA, USA, December 13-15, 2006.
- A. Hocine, D. Maquin et J. Ragot, *Finite memory observer for switching systems : Application to diagnosis*, 16th IFAC World Congress, Prague, Czech Republic, July 4-8, 2005.

- J. Ragot, A. Hocine et D. Maquin, *Parameter estimation of switching systems*, International Conference on Computational Intelligence for Modelling, Control and Automation, CIMCA'2004, Gold Coast, Australia, July 12-14, 2004.
- A. Hocine, D. Maquin et J. Ragot, *Finite memory observer for switching system : application to diagnosis*, Workshop on Advanced Control and Diagnosis, Karlsruhe, Germany, November 17-18, 2004.

Congrès nationaux avec comité de lecture et actes

- A. Hocine, D. Maquin et J. Ragot, *Utilisation d'une banque de modèles variable pour la détection de changement de régimes*, Journées Doctorales d'Automatique, Valenciennes, France, 25-27 juin 2003.

Introduction générale

Le développement croissant de l'automatisation, au cours de ces deux dernières décennies, a touché tous les secteurs de l'industrie. La complexité des systèmes industriels qui en résulte a rendu leur exploitation plus performante et plus vulnérable.

En effet, l'évolution des technologies a permis l'amélioration des performances des différents dispositifs, mais a aussi entraîné une prise en compte de leur fiabilité, critère qui caractérise leur sûreté de fonctionnement. Il s'ensuit que la réalisation de systèmes sûrs passe par la prise en compte lors de leur conception des préoccupations relatives à leur aptitude à résister aux éventuelles défaillances matérielles (capteurs, actionneurs, etc.), logicielles (par exemple, l'algorithme de commande) et humaines.

Ainsi, cette évolution technologique a contribué au développement de nouvelles procédures de surveillance qui permettent la détection, la localisation et le diagnostic des éventuels défauts. Ces procédures ou algorithmes de surveillance comprennent une étape de génération d'indicateurs de défauts ou résidus, qui caractérise un écart par rapport aux conditions de fonctionnement (plusieurs modes de fonctionnement peuvent être répertoriés). Une phase d'évaluation de ces résidus conduit ensuite à prendre une décision. Or, pour générer ces résidus, on utilise des informations issues d'un modèle analytique du système, afin de les comparer à celles fournies par les instruments de mesure. Une phase préliminaire d'analyse est donc nécessaire : c'est la structuration de la connaissance ou plus exactement la modélisation du système. Si l'on considère qu'un système peut changer de mode de fonctionnement à n'importe quel instant (par exemple passage d'un mode de fonctionnement normal à un mode de fonctionnement en défaut) le système pourra être considéré comme

un système à commutation. Cette catégorie de système est à évolution à la fois continue et événementielle (commutation). Elle appartient à la classe des systèmes dynamiques hybrides SDH (Zaytoon, 2001). En l'absence d'informations concernant les instants de commutation due au caractère aléatoire de l'apparition des défauts, nous avons choisi de modéliser le système réel par un système à commutation markovienne qui est représenté par un ensemble de modèles de fonctionnement (fonctionnements normaux et anormaux) et par une matrice de probabilité de transition de Markov qui contient les probabilités de passage d'un modèle de fonctionnement à un autre. La représentation du système étant fixé, celle-ci offre un cadre idéal à l'application de l'estimation multi-modèle (Magill, 1965) (Bar-Shalom et al., 1989) et (Zhang and Li, 1998). L'intérêt d'utiliser ce type d'estimateurs réside dans le fait qu'en plus de l'estimation de l'état du système, les estimateurs multi-modèle procurent la probabilité d'occurrence ou d'activation de chaque modèle de fonctionnement. Ces probabilités peuvent être utilisées pour la détection de défaut. Dans ce travail, nous avons utilisé les spécificités de l'estimation multi-modèle afin de procéder à la détection et l'isolation des défauts qui peuvent affecter un système linéaire. Cependant, plusieurs améliorations et aménagements ont été apportés à cet estimateur dans le but d'augmenter les performances du diagnostic.

Organisation

Ce mémoire, décomposé en six chapitres, est organisé de la façon suivante :

Chapitre 1

Après un rappel non exhaustif de quelques notions sur le filtre de Kalman ainsi que sur les systèmes à commutation markovienne et leur stabilité, nous introduisons l'estimation d'état par multi-modèle en relatant les différentes formes dont il est fait état dans la littérature spécialisée.

Chapitre 2

Dans ce chapitre, nous utilisons les méthodes d'estimation multi-modèle pour la détection de défaut où le système est représenté par un ensemble de modèles de bon fonctionnement et un ensemble de modèles représentant les défauts. Les probabilités d'activation des modèles issues des estimateurs du chapitre 1 sont utilisées pour la détection, et l'isolation du ou des défauts.

Chapitre 3

Le troisième chapitre est dédié à l'estimation de la matrice de Markov. Nous avons choisi de présenter deux méthodes (méthode quasi bayésienne et méthode par intégration numérique) proposées par (Jilkov and Li, 2004) pour effectuer cette estimation. Ensuite, nous les avons intégré à l'estimation multi-modèle afin d'estimer simultanément la matrice de Markov et l'état du système ainsi que les probabilités d'activation des modèles de fonctionnement du système.

Chapitre 4

Nous proposons dans ce chapitre d'intégrer les observateurs à mémoire finie (OMF) dans le cadre de la procédure de diagnostic et d'estimation multi-modèle. Ce type d'observateur nous permet d'estimer les entrées inconnues du système et d'effectuer la détection de défaut simultanément.

Chapitre 5

Afin d'améliorer les performances de détection et d'estimation de l'estimateur multi-modèle, nous proposons un observateur d'état mixte. Cet observateur est une combinaison de l'observateur à mémoire finie et de l'observateur de Luenberger. Il est intégré à une procédure d'estimation multi-modèle et appliqué à la détection de défaut.

Chapitre 6

Afin de limiter la dispersion des probabilités (difficulté à prendre une décision sur le modèle actif) liée au nombre de filtres fonctionnant en parallèle, nous proposons dans ce chapitre, une approche hiérarchisée. Les modèles sont regroupés en différents sous-ensembles. Un "sous-estimateur" est construit en s'appuyant sur tous les modèles de chaque sous-ensemble. Ces sous-estimateurs constituent ensuite les filtres élémentaires d'un estimateur multi-modèle global. Cette approche hiérarchisée permet de déterminer tout d'abord le sous-ensemble actif (c'est-à-dire le sous-ensemble dont la probabilité de contenir le modèle actif est la plus élevée) puis le modèle actif au sein de ce sous-ensemble.

1

Méthodes d'estimation d'état par multi-modèle pour les systèmes à commutation

Sommaire

1.1	Introduction	9
1.2	Filtre de Kalman	11
1.3	Système à commutation markovienne à temps discret	13
1.3.1	Chaînes de Markov	13
1.3.2	Chaînes de Markov à espace d'états discret	13
1.3.3	Description des systèmes à commutation markovienne	14
1.4	Estimation d'état par multi-modèle	15

1.4.1	Description générale	15
1.4.2	Approche multi-modèle pour le cas non commutant	18
1.4.3	Approche multi-modèle pour les systèmes à commutation	21
1.4.4	Estimateur pseudo-bayésien généralisé du premier ordre	25
1.4.5	Estimateur pseudo-bayésien généralisé du deuxième ordre	29
1.4.6	Estimateur IMM (Interacting Multiple Model Estimator)	33
1.5	Conclusion	39

1.1 Introduction

La complexité croissante des systèmes industriels nécessite le recours à des méthodes de commande de plus en plus performantes. Ces dernières se basent principalement sur l'utilisation de l'état du système. Celui-ci n'est pas toujours mesurable et les mesures disponibles sont souvent entachées de bruits, une procédure d'estimation d'état est donc incontournable.

Le filtre de Kalman est l'un des estimateurs les plus utilisés. Élaboré pour un modèle stochastique, il permet de déterminer d'une manière optimale l'état du système lorsque ce dernier a été au préalable modélisé. Dans le cas d'une erreur de modélisation ou d'un changement significatif du mode de fonctionnement du système, le filtre de Kalman dans sa version de base ne garantit plus des résultats optimaux. Pour résoudre le problème du changement de mode de fonctionnement, une approche consiste à utiliser un groupe de filtres, chaque filtre étant adapté à un mode particulier. Des hypothèses de sélection de modèle et de fusion de données basées sur l'estimation fournie par chaque filtre sont mises en œuvre pour avoir une meilleure estimation d'état, cette approche appartient à la classe des méthodes d'estimation dites *multi-modèle*.

Les systèmes à commutation représentent une catégorie importante des systèmes sujets à des changements dans leur mode de fonctionnement. Dans ce type de systèmes, les changements sont abrupts (commutations) et peuvent dépendre de l'état et/ou entrées du système, de la structure du système ou peuvent être aléatoires.

La motivation de cette approche découle du fait qu'il est souvent difficile de concevoir un modèle qui tient compte de toute la complexité du système étudié. Les systèmes à commutation sont des systèmes à évolution à la fois continue et événementielle (commutation). Ils appartiennent à la classe des systèmes dynamiques hybrides SDH (Zaytoon, 2001) qui permettent de représenter de nombreux systèmes dans des domaines d'application variés tels que les systèmes de transport, les systèmes flexibles de production, la robotique, etc. Les commutations entre les modes de fonctionnement peuvent être représentées par des modèles non stochastiques (automates temporisés, réseaux de Petri temporisés, algèbre min-max, etc.) ou par des modèles stochastiques (chaîne de Markov, réseaux de files d'at-

tentes, réseaux de Petri stochastiques, etc).

Dans le but de déterminer l'état et le mode actif du système, l'approche multi-modèle est appliquée dans le cadre des systèmes à commutation markovienne qui sont représentés par un ensemble de modèles de fonctionnement et par une matrice de probabilités de transition de Markov.

Cette approche permet la prise en compte des commutations entre les modes. L'occurrence des modes de fonctionnement est représentée par les probabilités des modes de fonctionnement (probabilités d'activation) au lieu des fonctions d'activation utilisées dans (Takagi and Sugeno, 1985), (Gasso et al., 2001), (Gasso, 2000). Cette approche peut être appliquée à la fois dans le cas de commutations brutales entre les modes ainsi que dans le cas de passage progressif du fait que les commutations sont représentées par des probabilités d'activation et non par des variables binaires.

Ces dernières années, les méthodes d'estimation basées sur l'approche *multi-modèle* appliquées aux systèmes à commutation markovienne ont été l'objet de nombreux travaux. Elles ont été initiées par *Magill* dans (Magill, 1965), puis améliorées par *Ackerson* dans (Ackerson and Fu, 1970) qui a pris en compte des commutations entre les différents modes. En effet, l'intérêt de cette approche réside dans sa capacité à traiter des problèmes à variations paramétriques et/ou structurelles et de décomposer des problèmes d'estimation complexes en sous-problèmes d'estimation simplifiés. Elle a été utilisée dans de nombreux domaines tel que le suivi de trajectoire d'une cible (Bar-Shalom et al., 1989), (Bar-Shalom, 1990), le GPS (global positioning system) (Burnette, 2001), (Chen and Harigae, 2001) ainsi que la détection de défauts (Zhang and Li, 1998) (Tugnait and Haddad, 1979a).

Ce chapitre sera consacré à la présentation des méthodes d'estimation par multi-modèle appliquées aux systèmes à commutation markovienne pour les utiliser ultérieurement dans le cadre du diagnostic et de la détection de défauts appliqués aux systèmes dynamiques linéaires.

Dans un premier temps, nous commençons par un rappel non exhaustif de quelques notions sur le filtre de Kalman ainsi que sur les systèmes à commutation markovienne et leurs stabilité. Ensuite, nous introduisons l'estimation d'état par multi-modèle en relatant les différentes formes dont il est fait état dans la littérature spécialisée.

1.2 Filtre de Kalman

Le filtre de Kalman résout de façon élégante le problème du filtrage linéaire. Utilisant la représentation d'état du système, le filtre de Kalman se présente sous la forme d'un ensemble d'équations récurrentes faciles à résoudre d'un point de vue numérique. Sa réalisation fournit non seulement l'estimé optimal de l'état du système, mais aussi la variance de l'erreur d'estimation.

L'évolution de l'état du système est décrite par le système d'équations récurrentes :

$$\begin{cases} x(k) = Ax(k-1) + Bu(k-1) + w(k-1) \\ z(k) = Cx(k) + v(k) \end{cases} \quad (1.1)$$

où $x(k) \in \mathbb{R}^n$ est le vecteur d'état à l'instant k , $A \in \mathbb{R}^{n \times n}$ est la matrice d'évolution d'état, $u(k-1) \in \mathbb{R}^p$ est le vecteur de commande à l'instant $k-1$, $B \in \mathbb{R}^{n \times p}$ est la matrice des gains de l'entrée, $C \in \mathbb{R}^{q \times n}$ est la matrice des gains de la sortie, $z(k) \in \mathbb{R}^q$ est la sortie du système à l'instant k , $w(k-1) \in \mathbb{R}^n$ est le bruit sur l'état à l'instant $k-1$ enfin $v(k) \in \mathbb{R}^q$ est le bruit mesure w et v sont blancs, gaussiens et centrés tels que :

$$\begin{aligned} E [w(k)w^T(j)] &= Q(k)\delta_{kj} \\ E [v(k)v^T(j)] &= R(k)\delta_{kj} \\ E [v(k)w^T(j)] &= 0 \quad \forall k, j \end{aligned} \quad (1.2)$$

où $Q(k)$ et $R(k)$ sont les matrices de variance des bruits et δ_{kj} est le symbole de Kronecker

$$\delta_{kj} = \begin{cases} 1 & k = j \\ 0 & k \neq j \end{cases}$$

L'état $x(k)$ et l'observation $z(k)$, qui se déduisent linéairement des bruits $v(k)$, $w(k)$ et des conditions initiales gaussiennes $x(0)$, avec $E[x(0)] = 0$, sont eux aussi gaussiens. L'expression du filtre de Kalman peut s'obtenir de manière itérative. À partir des conditions initiales $\hat{x}(0|0)$ et de la matrice de covariance $P(0|0)$, les équations du filtre de Kalman

discret s'écrivent :

$$\begin{aligned}
 P(k+1|k) &= AP(k|k)A^T + Q(k) \\
 \hat{x}(k+1|k) &= A\hat{x}(k|k) + Bu(k) \\
 K(k+1) &= P(k+1|k)C^T (CP(k+1|k)C^T + R(k+1))^{-1} \\
 \hat{x}(k+1|k+1) &= \hat{x}(k+1|k) + K(k+1)(z(k+1) - C\hat{x}(k+1|k)) \\
 P(k+1|k+1) &= (I - K(k+1)C)P(k+1|k)
 \end{aligned} \tag{1.3}$$

Le résidu $\nu(k+1) = z(k+1) - C\hat{x}(k+1|k)$ est la différence entre la mesure actuelle et la meilleure prédiction de la mesure, il sert à la réadaptation de l'estimée du filtre. Si les hypothèses précédentes (1.2) sur les bruits sont vérifiées, on a :

$$\begin{aligned}
 E[\nu(k)] &= 0 \\
 E[\nu(k)\nu^T(j)] &= 0, \quad k \neq j \\
 E[\nu(k)\nu^T(k)] &= S(k) = CP(k|k-1)C^T + R(k)
 \end{aligned} \tag{1.4}$$

Notons $N(V(k); 0, S(k))$ la distribution gaussienne de $V(k)$ de moyenne nulle et de matrice de variance $S(k)$:

$$N(V(k); 0, S(k)) = |2\pi S(k)|^{-\frac{1}{2}} \exp \left\{ -\frac{1}{2} V^T(k) S^{-1}(k) V(k) \right\} \tag{1.5}$$

La densité de probabilité de $\nu(k)$ est alors donnée par :

$$p(\nu(k)) = N(\nu(k); 0, S(k)) \tag{1.6}$$

Si le résidu n'a pas ces caractéristiques, alors les hypothèses (1.2) ne sont pas vérifiées, cela peut être dû à une mauvaise modélisation du système ou à un défaut sur le système. À partir des équations (1.5), plus la valeur du résidu est proche de 0 avec une variance faible, plus la valeur de la densité de probabilité augmente et plus le modèle du système représente le comportement du système. La valeur de cette densité de probabilité ou fonction de vraisemblance du modèle, sera un élément important dans l'évaluation des modèles dans les méthodes multi-modèle exposées par la suite.

1.3 Système à commutation markovienne à temps discret

Avant d'introduire les systèmes à commutation markovienne, nous allons rappeler quelques notions sur les chaînes de Markov.

1.3.1 Chaînes de Markov

Une chaîne de Markov est un processus stochastique possédant la propriété Markovienne. Dans un tel processus, la prédiction d'un événement futur à partir de l'événement présent ne nécessite pas la connaissance du passé.

1.3.2 Chaînes de Markov à espace d'états discret

Une chaîne de Markov en temps discret est une séquence X_1, X_2, X_3, \dots de variables aléatoires. L'ensemble de leurs valeurs possibles est appelé l'espace d'état, la valeur X_t est l'état du processus à l'instant t .

La distribution de probabilité de X_{t+1} conditionnelle aux états passés est donnée par :

$$P\{X_{t+1} = A | X_1, X_2, X_3, \dots, X_t\} = P\{X_{t+1} = A | X_t\}$$

où A est un état quelconque du processus. L'identité précédente constitue la propriété Markovienne.

On modélise avec une chaîne de Markov l'évolution au cours du temps de l'état X_t qui peut prendre un nombre fini d'états $\{A_1, A_2, \dots, A_n\}$ et qui passe de l'état A_i à l'instant t à l'état A_j à l'instant suivant $t + 1$ avec une probabilité π_{ij} donnée. Les nombres π_{ij} vérifient $\pi_{ij} \in [0, 1]$ et $\sum_{j=0}^n \pi_{ij} = 1$. Si l'espace des états est fini, alors la distribution de probabilité peut être représentée par une matrice appelée matrice de transition, dont le (i, j) ^{ème} élément vaut

$$\pi_{ij} = P\{X_{t+1} = A_j | X_t = A_i\}$$

1.3.3 Description des systèmes à commutation markovienne

Les systèmes physiques peuvent être sujets à des changements brusques dans leurs comportements, par exemple en cas de panne sur un actionneur ou sur un capteur, ainsi le changement est souvent imprévisible. Ce type de comportement peut être modélisé par les systèmes à commutation markovienne. Les systèmes à commutation markovienne appartiennent à la classe des systèmes hybrides, ils contiennent un ensemble de dynamiques régies par un mode de commutation stochastique. Les commutations sont modélisées par une chaîne de Markov à espace d'états fini.

Un système linéaire à commutation markovienne est représenté par un ensemble de modèles $M = \{M_1, M_2, \dots, M_r\}$ où r est le nombre de modèles.

Si le modèle M_j , caractérisé par les matrices A_j, B_j, C_j , est actif à l'instant k , le comportement du système est décrit par :

$$\begin{cases} x(k) = A_j x(k-1) + B_j u(k-1) + w(k-1) \\ z(k) = C_j x(k) + v(k) \end{cases} \quad (1.7)$$

où $x(k) \in \mathbb{R}^n$ est le vecteur d'état à l'instant k , $u(k) \in \mathbb{R}^p$ est le vecteur de commande à l'instant k , $z(k) \in \mathbb{R}^q$ est la sortie du système à l'instant k , enfin $w(k) \in \mathbb{R}^n$ et $v(k) \in \mathbb{R}^q$ sont respectivement le bruit d'état et le bruit de mesure considérés blancs, gaussiens et centrés tels que :

$$\begin{aligned} E [w(k)w^T(j)] &= Q(k)\delta_{kj} \\ E [v(k)v^T(k)] &= R(k)\delta_{kj} \\ E [v(k)w^T(j)] &= 0 \quad \forall k, j \end{aligned} \quad (1.8)$$

où $Q(k)$ et $R(k)$ sont les matrices de variance des bruits. Dorénavant, on notera $M_j(k)$ l'événement correspondant à l'activation du modèle M_j à l'instant k . Les transitions d'un mode à un autre suivent un processus markovien, défini par sa matrice de transition Π

$$\Pi = \begin{bmatrix} \pi_{11} & \cdots & \pi_{1r} \\ \vdots & \ddots & \vdots \\ \pi_{r1} & \cdots & \pi_{rr} \end{bmatrix} \quad (1.9)$$

avec $\sum_{j=1}^r \pi_{ij} = 1$, $i = 1, \dots, r$ et où π_{ij} représente la probabilité de passage du $i^{\text{ème}}$ mode au $j^{\text{ème}}$ mode.

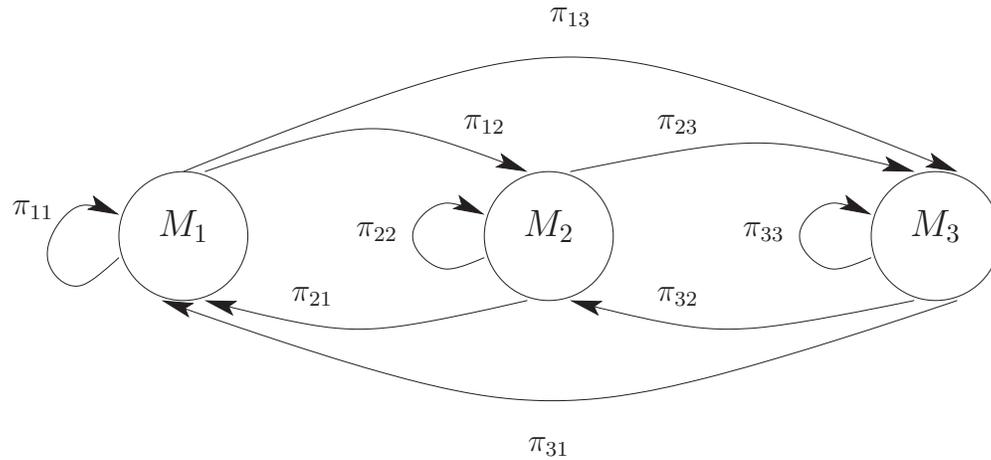


FIG. 1.1 – Système à commutation markovienne

Le système à commutation est représenté sur la figure 1.1 avec $r = 3$. L'utilisation de ce type de système procure un certain avantage dans le cas où les commutations entre les modèles ne sont pas connues a priori et les informations relatives aux commutations sont représentées par la matrice des probabilités de transition. La stabilité de ce type de système hybride peut être étudiée dans le cadre plus général des multi-modèles.

1.4 Estimation d'état par multi-modèle

1.4.1 Description générale

L'idée de base de l'estimation d'état par multi-modèle est de considérer un ensemble de modèles M qui représente le comportement du système. On construit alors un banc de filtres où chaque filtre est calé sur un modèle de l'ensemble M . L'estimée globale est donnée par une combinaison des estimées de ces filtres.

L'approche multi-modèle a été initiée par Magill (Magill, 1965). Les premiers travaux ont traité le cas des systèmes invariants dans le temps, mais représentés par des modèles incertains ou inconnus. Plusieurs applications de cet estimateur multi-modèle peuvent être trouvées dans la littérature sous des noms différents, comme "static multiple-model

(SMM) algorithm" (Li and Zhang, 2000), "multiple model adaptive estimator" (Maybeck, 1982), "parallel processing algorithm" (Anderson and Moore, 1979) et "filter bank method" (Brown, 1983). Ces différentes appellations sont relatives à la structure, aux caractéristiques et aux capacités de cette première génération d'estimateurs multi-modèle (Li, 1996).

Dans la première génération des algorithmes SMM, les filtres élémentaires opèrent individuellement sans aucune interaction entre eux parce qu'il est supposé qu'il n'y a pas de commutation entre les modes de fonctionnement. Par conséquent, cette méthode n'est pas efficace dans le cas des systèmes à commutation fréquentes. Cependant, elle reste efficace dans le cas des systèmes à commutations très rares.

Afin de prendre en compte les commutations entre les modes du système, une deuxième génération d'algorithmes a été développée, notamment l'estimateur "generalized pseudo-Bayesian estimator" (GPB) (Chang and Athans, 1978) et l'estimateur "interacting multiple model" (IMM) (Blom and Bar-Shalom, 1988). Pour ces estimateurs, la commutation d'un mode à un autre est supposée régie par un processus markovien. La principale différence entre ce type d'estimateurs et ceux issus des algorithmes SMM réside dans l'étape de réinitialisation des filtres.

Dans cette section, nous allons traiter, dans un premier temps, le cas d'estimateurs multi-modèles sans commutation et chercher à calculer les probabilités d'activation $\mu_i(k)$ des modèles M_i ainsi que l'estimée globale de l'état. Puis, nous allons considérer le cas de modèles commutant dans le temps, où plusieurs méthodes seront présentées. La différence entre ces méthodes réside dans la manière d'utiliser l'historique des commutations.

En général l'estimation d'état par multi-modèle est mise en œuvre en adoptant la chronologie suivante :

Détermination de l'ensemble des modèles

Les performances d'un estimateur multi-modèle sont largement dépendantes de l'ensemble des modèles utilisé. De ce fait, une des principales étapes dans la construction d'un estimateur multi-modèle est la détermination de l'ensemble des modèles M représentatif du fonctionnement du système. Elle a été étudiée pour les estimateurs multi-modèle dans

Li et al. (2005) où le concept de mode et modèle aléatoire a été introduit. Les auteurs ont représenté les modes de fonctionnement du système par des variables aléatoires afin d'introduire la notion de distance probabiliste entre les modèles. Trois méthodes ont été élaborées en se basant cette notion de distance probabiliste. D'autres résultats sont présentés dans (Li, 2002; Li et al., 2002) où une étude portant sur l'influence du choix de l'ensemble des modèles sur la qualité de l'estimation de l'état a été conduite. D'autres résultats ont été élaborés dans (Caputi, 1995; Li, 1998; Sheldon and Maybeck, 1993).

Sélection du filtre

Plusieurs types de filtre peuvent être employés pour effectuer une estimation d'état sur la base d'un des modèles de l'ensemble M . Chacun de ces filtres sera appelé filtre élémentaire. Le filtre de Kalman est largement utilisé, mais on peut recourir à d'autres types de filtres, le choix sera motivé par l'application et les objectifs de l'estimation. On a proposé dans (Hocine et al., 2005a) l'utilisation d'un observateur à mémoire finie afin de réduire les effets des bruits. Cette méthode sera exposée dans le chapitre 4 ainsi que l'utilisation d'un observateur à fenêtre glissante.

Réinitialisation des filtres

Excepté le cas d'estimateurs multi-modèle pour les systèmes non commutants, les filtres élémentaires ne doivent pas opérer indépendamment les uns des autres. L'entrée de chaque filtre dépend du comportement des autres filtres. La réinitialisation de chaque filtre sert à créer des interactions entre les différents filtres. Plusieurs configurations de réinitialisation sont possibles. Nombre d'entre elles sont présentées dans (Bar-Shalom and Li, 1993; Li, 1998).

Elaboration de l'estimée globale

L'estimée globale est obtenue en fusionnant les estimées issues des filtres élémentaires basés sur l'ensemble des modèles M . Deux approches sont envisageables :

◇ Pas de décision : l'estimée globale est la somme pondérée des estimées $\hat{x}_i(k|k)$ issues

des différents filtres. Dans ce type d'approche, aucune décision relative au choix du mode actif n'est prise. Cela revient à prendre en considération toutes les estimées pondérées par leurs probabilités d'activation.

$$\hat{x}(k|k) = E[x(k)|z(k)] = \sum_{i=1}^r \hat{x}_i(k|k)P\{M_i|z(k)\} \quad (1.10)$$

- ◊ Prise de décision : dans ce cas l'estimée globale est calculée en n'utilisant que les N modèles les plus probables, c-à-d qui ont la probabilité $P\{M_i|z(k)\}$ la plus élevée. Autrement dit l'estimée globale est calculée à partir de la somme des N estimées, issues des filtres relatifs aux N modèles les plus probables, pondérée par leurs probabilités normalisées de façon à ce que la somme des N probabilités soit égale à 1 (Tugnait and Haddad, 1979a). Dans le cas le plus extrême ($N = 1$) l'estimée globale correspond exactement à une des estimées.

1.4.2 Approche multi-modèle pour le cas non commutant

Dans un premier temps, nous considérons que le système évolue selon un unique mode de fonctionnement. Il s'agit de reconnaître le modèle selon lequel le système observé évolue, parmi un ensemble M connu de r modèles $M = \{M_1, \dots, M_r\}$. M_j désigne indifféremment le modèle numéro j ou l'événement (au sens des probabilités) correspondant à son occurrence, c-à-d le fait que le modèle M_j soit effectivement le modèle selon lequel le système fonctionne.

L'approche multi-modèle permet le calcul des probabilités d'occurrence de chaque mode de fonctionnement (*probabilité d'activation*) à l'instant k à partir de cette même probabilité calculée ou connue a priori à l'instant précédent $k - 1$.

La probabilité d'activation a priori du mode M_j est définie par :

$$\mu_j(0) = P\{M_j|Z^0\}, \quad j = 1, \dots, r \quad (1.11)$$

où Z^0 représente l'information a priori et où $\sum_{j=1}^r \mu_j(0) = 1$.

La probabilité d'activation du mode M_j à l'instant k est définie par

$$\mu_j(k) = P\{M_j|Z^k\}, \quad j = 1, \dots, r \quad (1.12)$$

où $Z^k = [z(0), z(1), \dots, z(k)]$ représente le vecteur des mesures cumulées jusqu'à l'instant k avec $\sum_{j=1}^r \mu_j(k) = 1$.

En utilisant la formule de Bayes, la probabilité d'activation du $j^{\text{ème}}$ modèle connaissant les mesures jusqu'à l'instant k , peut être obtenue de manière récursive :

$$\begin{aligned} \mu_j(k) &= P\{M_j | Z^k\} = P\{M_j | z(k), Z^{k-1}\} = \frac{p[z(k) | Z^{k-1}, M_j] P\{M_j | Z^{k-1}\}}{P\{z(k) | Z^{k-1}\}} \\ &= \frac{p[z(k) | Z^{k-1}, M_j] P\{M_j | Z^{k-1}\}}{\sum_{i=1}^r p[z(k) | Z^{k-1}, M_i] P\{M_i | Z^{k-1}\}} \end{aligned}$$

on a alors :

$$\mu_j(k) = \frac{p[z(k) | Z^{k-1}, M_j] \mu_j(k-1)}{\sum_{i=1}^r p[z(k) | Z^{k-1}, M_i] \mu_i(k-1)} \quad j = 1, \dots, r \quad (1.13)$$

Le premier terme à droite de l'équation (1.13) représente la vraisemblance à l'instant k du $j^{\text{ème}}$ mode de fonctionnement. Sous l'hypothèse gaussienne, la vraisemblance peut s'écrire

$$\lambda_j(k) = p[z(k) | Z^{k-1}, M_j] = p[\nu_j(k)] = N[\nu_j(k); 0, S_j(k)] \quad (1.14)$$

$$S_j(k) = C_j A_j P(k-1|k-1) A_j^T C_j^T + R(k)$$

$$\nu_j(k) = z(k) - C_j(A_j \hat{x}_j(k-1|k-1) + B_j u(k-1))$$

où $R(k)$ est la covariance du bruit sur la sortie, $S_j(k)$ est la covariance du résidu $\nu_j(k)$ issu de la différence entre la mesure $z(k)$ et l'estimation $\hat{z}(k)$ calculée par le $j^{\text{ème}}$ filtre représentant le $j^{\text{ème}}$ mode de fonctionnement.

La probabilité d'activation de chaque mode de fonctionnement est calculée par (1.13) en utilisant la vraisemblance (1.14) correspondent au $j^{\text{ème}}$ filtre.

La procédure d'estimation est illustrée par la figure 1.2. Sous l'hypothèse que M_j est le modèle actif, le $j^{\text{ème}}$ filtre donne en sortie l'estimée $\hat{x}_j(k|k)$ et sa variance associée $P_j(k|k)$ ainsi que sa fonction de vraisemblance $\lambda_j(k)$.

Après l'initialisation des états, les filtres fonctionnent d'une manière récursive. A l'instant k , ils utilisent leurs propres estimées calculées à l'instant $k-1$ afin de calculer celles de l'instant présent k . La fonction de vraisemblance relative à chaque modèle est utilisée pour la remise à jour des probabilités d'activation. Ces probabilités d'activation sont utilisées à leur tour pour le calcul de l'estimée globale de d'état et de sa covariance :

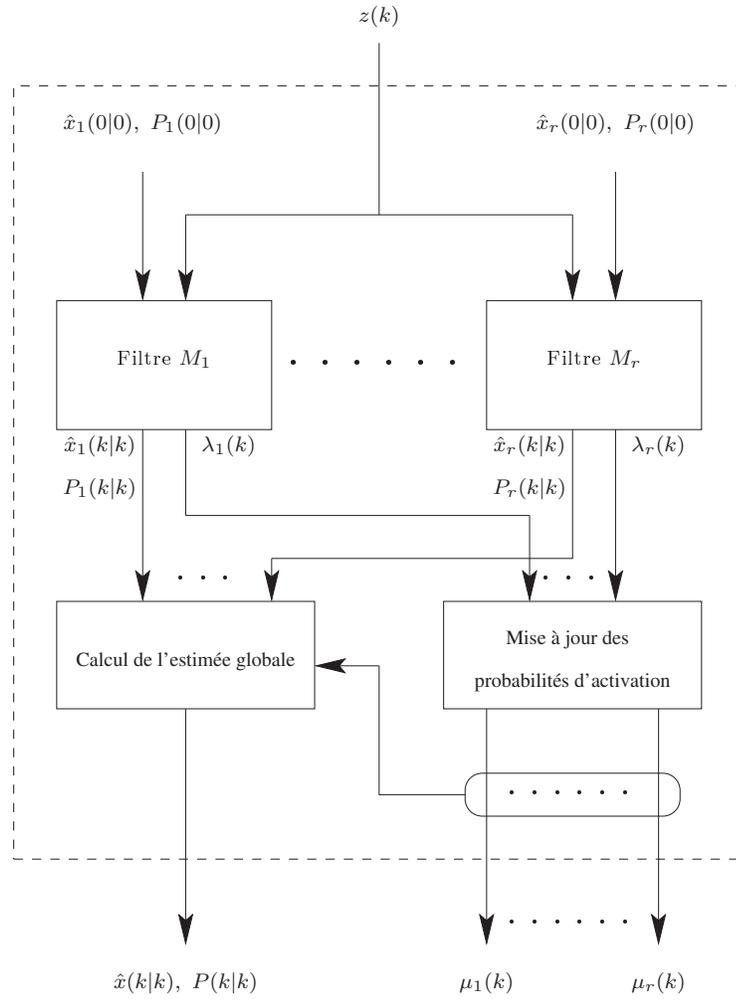


FIG. 1.2 – Estimateur multi-modèle pour les systèmes non commutants

$$\hat{x}(k|k) = \sum_{j=1}^r \mu_j(k) \hat{x}_j(k|k) \quad (1.15)$$

$$P(k|k) = \sum_{j=1}^r \mu_j(k) \left\{ P_j(k|k) + [\hat{x}_j(k|k) - \hat{x}(k|k)] [\hat{x}_j(k|k) - \hat{x}(k|k)]^T \right\} \quad (1.16)$$

Ce résultat est applicable à condition que :

1. Le mode de fonctionnement actif appartient à l'ensemble des modèles M .
2. Un seul mode de fonctionnement est actif pendant toute la durée de l'estimation.

En cas de commutation d'un mode à un autre la deuxième hypothèse n'est plus respectée et une reformulation de l'approche sera nécessaire pour tenir compte des commutations.

1.4.3 Approche multi-modèle pour les systèmes à commutation

Cette fois, à chaque instant, le système peut fonctionner selon l'un des modèles de l'ensemble $M = \{M_1, M_2, \dots, M_r\}$, les commutations d'un modèle à l'autre se produisant de manière aléatoire. Si le modèle M_j , caractérisé par les matrices d'état A_j , B_j et C_j est actif à l'instant k , le comportement du système est décrit par :

$$\begin{cases} x(k) = A_j x(k-1) + B_j u(k-1) + w(k-1) \\ z(k) = C_j x(k) + v(k) \end{cases} \quad (1.17)$$

Du fait de la possibilité de commuter à n'importe quel instant, à l'instant k on a r^k historiques de commutation possibles. Le $l^{\text{ème}}$ historique est représenté par :

$$M_{k,l} = \{M_{i_{1,l}}, \dots, M_{i_{k,l}}\}, \quad l = 1, \dots, r^k \quad (1.18)$$

où $i_{t,l}$ est l'indice du modèle à l'instant t appartenant au $l^{\text{ème}}$ historique. On remarque que le nombre d'historiques r^k est exponentiellement croissant avec le temps. Par exemple pour $r = 2$ et à l'instant $k = 2$, on a $r^k = 4$ séquences possibles

l	i_{1l}	i_{2l}
1	1	1
2	1	2
3	2	1
4	2	2

On suppose que le processus de commutation est un processus markovien régi par une matrice Π . On note $M_j(k)$ l'événement indiquant qu'à l'instant k le $j^{\text{ème}}$ modèle est actif. La probabilité conditionnelle du $l^{\text{ème}}$ historique est définie par :

$$\mu_{k,l} = P\{M_{k,l}|Z^k\} \quad (1.19)$$

La $l^{\text{ème}}$ séquence de modèles jusqu'à l'instant k est représentée par :

$$M_{k,l} = \{M_{k-1,l}, M_j(k)\} \quad (1.20)$$

où $M_{k-1,l}$ est la séquence des $(k-1)$ premier modèles de la $l^{\text{ème}}$ séquence et $M_j(k)$ est le dernier modèle de la séquence à l'instant k . On utilisant la propriété de Markov, on peut écrire :

$$P\{M_j(k)|M_{k-1,l}\} = P\{M_j(k)|M_i(k-1)\} = \pi_{ij} \quad (1.21)$$

où i représente l'indice du dernier modèle de la séquence $M_{k-1,l}$.

L'expression de la densité de probabilité conditionnelle de l'état x à l'instant k est calculée en utilisant le théorème de la probabilité totale. Ce calcul se fait en supposant que l'ensemble M représente tous les comportements possibles du système et qu'un seul modèle est actif à un instant k . On obtient l'expression suivante :

$$p[x(k)|Z^k] = \sum_{l=1}^{r^k} p[x(k)|M_{k,l}, Z^k] P\{M_{k,l}|Z^k\} \quad (1.22)$$

L'utilisation de la densité de probabilité $p[x(k)|Z^k]$ de l'équation (1.22), où $p[x(k)|M_{k,l}, Z^k]$ représente la densité de probabilité de $x(k)$ conditionnelle à l'historique $M_{k,l}$ et $P\{M_{k,l}|Z^k\}$ la probabilité de cet historique, permet de calculer l'espérance de $x(k)$ et par conséquent d'obtenir l'expression de l'estimée globale donnée par :

$$\begin{aligned} \hat{x}(k) &= \sum_{l=1}^{r^k} \hat{x}_{k,l}(k) P\{M_{k,l}|Z^k\} \\ &= \sum_{l=1}^{r^k} \hat{x}_{k,l}(k) \mu_{k,l} \end{aligned} \quad (1.23)$$

où $\hat{x}_{k,l}(k)$ est l'estimée relative au $l^{\text{ème}}$ historique et r^k est le nombre de filtres nécessaires à une estimation optimale à l'instant k . On peut remarquer que ce nombre augmente exponentiellement avec le temps, par conséquent l'approche utilisant tout l'historique est

pratiquement irréalisable. La figure 1.3 représente la procédure d'estimation à l'instant $k = 2$ avec $r = 2$. Elle illustre bien le phénomène d'augmentation du nombre de filtres nécessaire à l'estimation.

La probabilité $\mu_{k,l}$ de l'historique l est développée en utilisant la formule de Bayes de la façon suivante :

$$\begin{aligned}\mu_{k,l} &= P \{ M_{k,l} | Z^k \} \\ &= P \{ M_{k,l} | z(k), Z^{k-1} \}\end{aligned}\quad (1.24)$$

En utilisant le théorème de Bayes, on obtient :

$$\mu_{k,l} = \frac{1}{c} p [z(k) | M_{k,l}, Z^{k-1}] P \{ M_{k,l} | Z^{k-1} \} \quad (1.25)$$

où c est une constante de normalisation. Puis, en tenant compte de (1.20) :

$$\mu_{k,l} = \frac{1}{c} p [z(k) | M_{k,l}, Z^{k-1}] P \{ M_j(k), M_{k-1,l} | Z^{k-1} \} \quad (1.26)$$

Une seconde application du théorème de Bayes, permet d'écrire :

$$\mu_{k,l} = \frac{1}{c} p [z(k) | M_{k,l}, Z^{k-1}] P \{ M_j(k) | M_{k-1,l}, Z^{k-1} \} P \{ M_{k-1,l} | Z^{k-1} \} \quad (1.27)$$

d'où l'expression finale de la probabilité $\mu_{k,l}$:

$$\mu_{k,l} = \frac{1}{c} p [z(k) | M_{k,l}, Z^{k-1}] P \{ M_j(k) | M_{k-1,l}, Z^{k-1} \} \mu_{k-1,l} \quad (1.28)$$

En utilisant la propriété de Markov sur l'indépendance du passé, on obtient :

$$\mu_{k,l} = \frac{1}{c} p [z(k) | M_{k,l}, Z^{k-1}] P \{ M_j(k) | M_i(k-1) \} \mu_{k-1,l} \quad (1.29)$$

ce qui donne :

$$\mu_{k,l} = \frac{1}{c} \lambda_{k,l}(k) \pi_{ij} \mu_{k-1,l} \quad (1.30)$$

où i est l'indice du dernier modèle de la séquence $M_{k-1,l}$ et où $\lambda_{k,l}(k)$ représente la vraisemblance, à l'instant k , de l'historique $M_{k,l}$. Sous l'hypothèse gaussienne, cette vraisemblance peut s'écrire :

$$\lambda_{k,l} = p [z(k) | M_{k,l}, Z^{k-1}] = p [\nu_{k,l}] = N [\nu_{k,l}(k); 0, S_{k,l}] \quad (1.31)$$

$$S_{k,l} = C_j A_j P(k-1|k-1) A_j^T C_j^T + R(k)$$

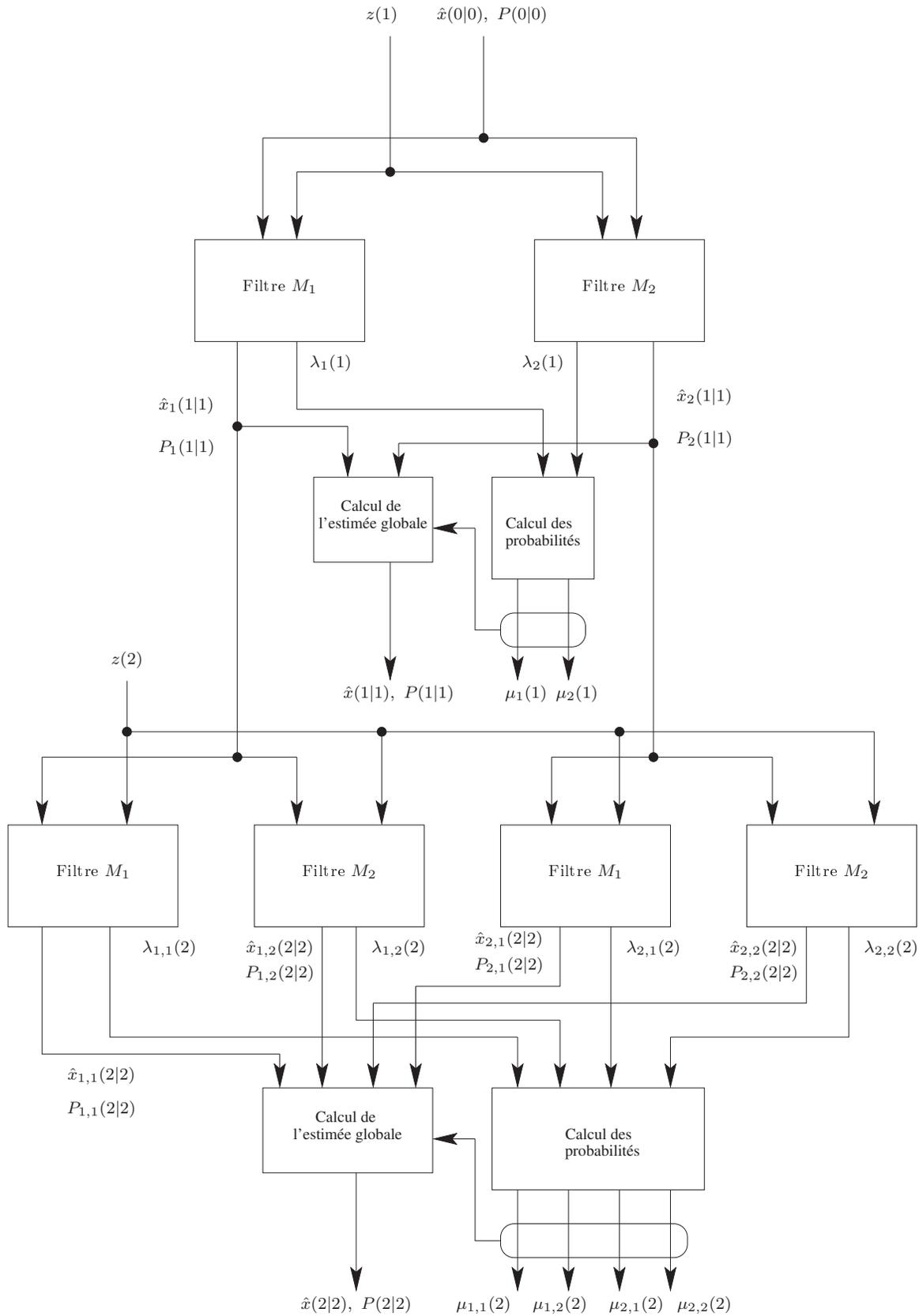


FIG. 1.3 – Estimateur multi-modèle optimal pour les systèmes commutants

$$\nu_{k,l} = z(k) - C_j(A_j \hat{x}_j(k-1|k-1) + B_j u(k-1))$$

où $R(k)$ est la covariance du bruit sur la sortie, $S_{k,l}$ est la covariance du résidu $\nu_{k,l}$ issu de la différence entre la mesure $z(k)$ et l'estimation de la mesure $\hat{z}(k)$ par le $l^{\text{ème}}$ filtre représentant le $l^{\text{ème}}$ historique.

L'équation (1.31) procure les vraisemblances des historiques nécessaires au calcul des probabilités des historiques.

Approche pratique

La seule solution à l'augmentation exponentielle du nombre des historiques, ainsi que du nombre de filtres nécessaires à l'estimation, est l'utilisation de techniques sub-optimales. Une des solutions envisageables consiste à garder, à chaque cycle d'estimation, les N historiques ayant les probabilités les plus élevées et à normaliser les probabilités ainsi retenues de façon à ce que leur somme soit égale à 1.

L'approche pseudo-bayésienne généralisée, en anglais "Generalized Pseudo-Bayesian" (GPB) consiste à utiliser des historiques de longueur finie. Le GPB d'ordre 1 (GPB1) considère seulement les r modèles possibles à l'instant k au lieu de considérer tout l'historique, cela implique l'utilisation de r filtres. Le GPB d'ordre 2 (GPB2) considère, à l'instant k , les modèles possibles aux instants k et $k-1$ et implique l'utilisation de r^2 filtres, cela revient à dire que la longueur des historiques est toujours égale à 2.

L'approche "Interacting Multiple Model" (IMM) est similaire au GPB2 dans le fait qu'elle prend en considération les modèles de l'instant k et $k-1$. Son principal avantage réside dans le fait qu'elle utilise r filtres au lieu de r^2 .

1.4.4 Estimateur pseudo-bayésien généralisé du premier ordre

Dans l'approche GPB1, l'estimation d'état est calculée en considérant r modèles possibles à l'instant k ; il existe donc r possibilités (hypothèses) au lieu des r^k de l'approche optimale.

En appliquant le théorème de la probabilité totale à la densité de probabilité de $x(k)$, on

obtient, à la place de (1.22) :

$$\begin{aligned}
 p[x(k) | Z^k] &= \sum_{j=1}^r p[x(k) | M_j(k), Z^k] P\{M_j(k) | Z^k\} \\
 &= \sum_{j=1}^r p[x(k) | M_j(k), Z^k] \mu_j(k) \\
 &\simeq \sum_{j=1}^r p[x(k) | M_j(k), z(k), \hat{x}(k-1|k-1), P(k-1|k-1)] \mu_j(k)
 \end{aligned} \tag{1.32}$$

L'approximation consiste à remplacer l'information fournie par le vecteur des mesures Z^{k-1} par celle issue de l'estimée $\hat{x}(k-1|k-1)$, issue du filtre de Kalman, de l'instant $k-1$ et de sa covariance. Cette approximation est justifiée intuitivement par le fait que le calcul de l'estimée est basé sur l'utilisation des mesures Z^{k-1} .

À l'instant k , tous les filtres calés sur les différents modèles M_j , $j = 1, 2, \dots, r$, fournissent, à partir de la connaissance de l'estimée $\hat{x}(k-1|k-1)$ et de sa covariance $P(k-1|k-1)$, une estimée $\hat{x}_j(k|k)$.

La densité de probabilité de l'estimée $x(k)$ est un mélange de r densités gaussiennes relatives aux r modèles. L'utilisation des propriétés du mélange gaussien permet de calculer l'estimée $\hat{x}(k|k)$:

$$\hat{x}(k|k) = \sum_{j=1}^r \hat{x}_j(k|k) \mu_j(k) \tag{1.33}$$

et sa covariance

$$P(k|k) = \sum_{j=1}^r \mu_j(k) \{P_j(k|k) + [\hat{x}_j(k|k) - \hat{x}(k|k)][\hat{x}_j(k|k) - \hat{x}(k|k)]^T\} \tag{1.34}$$

Les probabilités d'activation $\mu_j(k)$ de chaque modèle sont mises à jour de façon récurrente en fonction des informations disponibles à chaque instant.

La figure 1.4 représente l'agencement des différentes étapes de la méthode où r filtres en parallèle fournissent des estimées $\hat{x}_j(k|k)$ qui sont combinées en une seule estimée globale. Les probabilités $\mu_j(k)$ des modèles M_j sont mises à jour à chaque cycle de l'estimation. L'enchaînement des différentes opérations est résumé ci-dessous.

Calcul des vraisemblances

À l'instant k , chaque filtre a comme condition initiale $\hat{x}(k-1|k-1)$. Le $j^{\text{ème}}$ filtre fournit en sortie l'estimée $\hat{x}_j(k|k)$ et sa covariance $P_j(k|k)$. La vraisemblance du $j^{\text{ème}}$

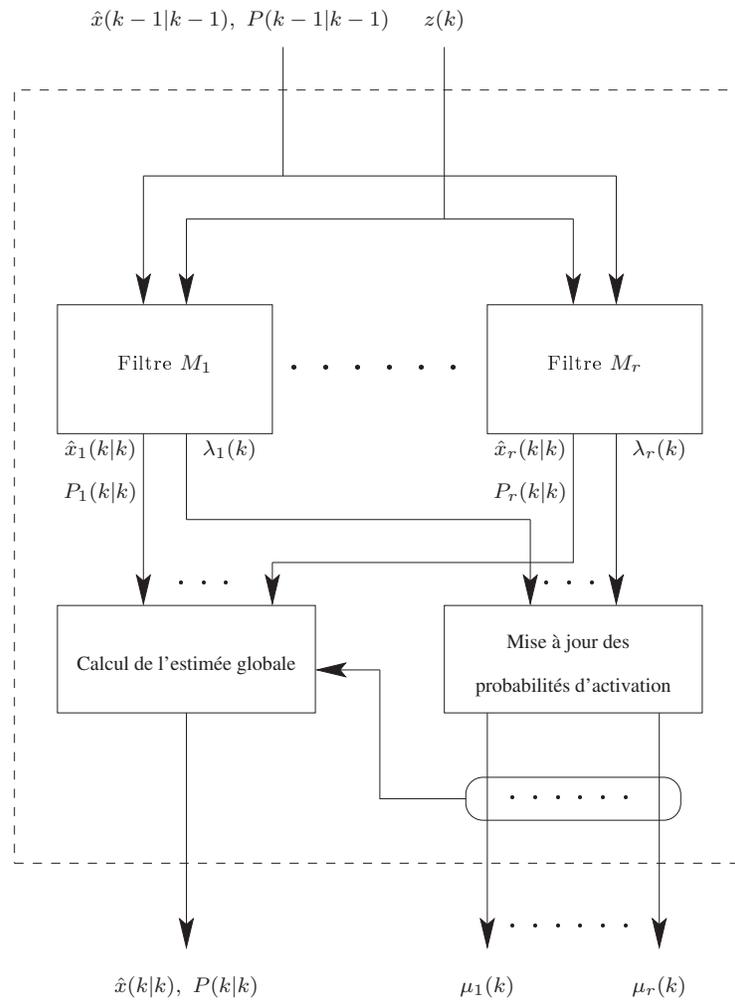


FIG. 1.4 – Estimateur GPB1

modèle $j = 1, \dots, r$ est calculée sous l'hypothèse gaussienne par :

$$\lambda_j(k) = p[z(k) | M_j(k), \hat{x}(k-1|k-1), P(k-1|k-1)] = p[\nu_j(k)] = N[\nu_j(k); 0, S_j(k)] \quad (1.35)$$

$$S_j(k) = C_j A_j P(k-1|k-1) C_j^T A_j^T + R(k)$$

$$\nu_j(k) = z(k) - C_j(A_j \hat{x}(k-1|k-1) + B_j u(k-1))$$

où $R(k)$ est la covariance du bruit sur la sortie, $S_j(k)$ est la covariance du résidu, $\nu_j(k)$ est issu de la différence entre la mesure $z(k)$ et l'estimation $\hat{z}(k)$ calculée par le $j^{\text{ème}}$ filtre représentant le $j^{\text{ème}}$ mode de fonctionnement.

Ces vraisemblances seront utilisées pour le calcul des probabilités d'activation des modèles.

Mise à jour des probabilités d'activation

La probabilité d'activation $\mu_j(k)$ du modèle M_j est évaluée en utilisant la formule de Bayes et en décomposant le vecteur des mesures :

$$\begin{aligned} \mu_j(k) &= P\{M_j(k) | Z^k\} \\ &= P\{M_j(k) | z(k), Z^{k-1}\} \\ &= \frac{1}{c} p[z(k) | M_j(k), Z^{k-1}] P\{M_j(k) | Z^{k-1}\} \end{aligned} \quad (1.36)$$

où c est une constante de normalisation donnée par :

$$c = \sum_{j=1}^r \lambda_j(k) \sum_{i=1}^r \pi_{ij} \mu_i(k-1) \quad (1.37)$$

En utilisant le théorème de la probabilité totale sur $P\{M_j(k) | Z^{k-1}\}$, on obtient :

$$\mu_j(k) = \frac{1}{c} \lambda_j(k) \sum_{i=1}^r P\{M_j(k) | M_i(k-1), Z^{k-1}\} P\{M_i(k-1) | Z^{k-1}\} \quad (1.38)$$

En utilisant la matrice de Markov, l'équation précédente s'écrit :

$$\mu_j(k) = \frac{1}{c} \lambda_j(k) \sum_{i=1}^r \pi_{ij} \mu_i(k-1) \quad (1.39)$$

La formule (1.39) permet un calcul récursif des probabilités d'activation.

Calcul de l'estimée globale

L'estimée globale à l'instant k est une somme des estimées $\hat{x}_j(k|k)$, $j = 1, \dots, r$ pondérée par les probabilités d'activation $\mu_j(k)$, $j = 1, \dots, r$, (voir les équations (1.33), (1.34)). Cette estimée, ainsi que sa vraisemblance, constitueront les données d'entrée des filtres au cycle suivant.

1.4.5 Estimateur pseudo-bayésien généralisé du deuxième ordre

Dans l'algorithme GPB2, l'estimation d'état à l'instant k , est calculée en prenant en compte les r modes possibles à l'instant k ainsi que ceux à l'instant $k - 1$ ce qui représente r^2 possibilités. Une telle procédure nécessite l'utilisation de r^2 filtres.

Comme dans le cas précédent, la densité de probabilité de $x(k)$ s'écrit :

$$p[x(k) | Z^k] = \sum_{j=1}^r p[x(k) | M_j(k), Z^k] P \{ M_j(k) | Z^k \} \quad (1.40)$$

En appliquant le théorème de la probabilité totale :

$$p[x(k) | Z^k] = \sum_{j=1}^r \sum_{i=1}^r p[x(k) | M_j(k), M_i(k-1), Z^k] P \{ M_i(k-1) | M_j(k), Z^k \} P \{ M_j(k) | Z^k \} \quad (1.41)$$

Compte tenu du fait que l'on tient compte d'un historique de longueur égale à 2, on notera l'introduction du terme $P \{ M_i(k-1) | M_j(k), Z^k \}$ qui correspond à la probabilité que le système ait fonctionné selon le modèle M_i à l'instant $k - 1$ sachant qu'il fonctionne selon le modèle M_j à l'instant k et connaissant l'information Z^k . Cette probabilité sera ensuite notée $\mu_{i|j}(k-1|k)$. On peut alors écrire :

$$p[x(k) | Z^k] = \sum_{j=1}^r \sum_{i=1}^r p[x(k) | M_j(k), M_i(k-1), z(k), Z^{k-1}] \mu_{i|j}(k-1|k) \mu_j(k) \quad (1.42)$$

On effectue ensuite une approximation en supposant que l'information disponible à l'instant $k - 1$, $(M_i(k-1), Z^{k-1})$ peut se résumer à la connaissance de l'estimée $\hat{x}_i(k-1|k-1)$ et sa covariance $P_i(k-1|k-1)$:

$$p[x(k) | Z^k] \simeq \sum_{j=1}^r \sum_{i=1}^r p[x(k) | M_j(k), z(k), \hat{x}_i(k-1|k-1), P_i(k-1|k-1)] \mu_{i|j}(k-1|k) \mu_j(k) \quad (1.43)$$

L'organisation de la méthode GPB2 est présentée à la figure 1.5. On peut remarquer qu'à l'instant $k - 1$ on a r estimées $\hat{x}_j(k - 1|k - 1)$, $j = 1, \dots, r$; chacune d'entre elles est utilisée à l'entrée des r filtres représentant les r modes de fonctionnement à l'instant k , ce qui signifie que r^2 filtres sont utilisés à chaque cycle de l'estimation. À la sortie de ces filtres, on obtient r^2 estimées $\hat{x}_{ij}(k|k)$ $\{i, j\} = 1, \dots, r$ chacune étant dédiée à un historique d'ordre deux.

Les estimées correspondant au même modèle à l'instant k sont ensuite combinées avec les poids $\mu_{ij}(k - 1|k)$ pour fournir les r estimées $\hat{x}_i(k|k)$. Ces estimées sont ensuite utilisées à l'entrée du cycle suivant. Par rapport à l'approche GPB1, on remarque qu'on a besoin de r^2 filtres au lieu de r . L'algorithme suit les étapes suivantes :

Calcul des vraisemblances

Connaissant les estimées $\hat{x}_i(k - 1|k - 1)$ relatives aux r modèles ainsi que leurs covariances associées, on peut calculer la vraisemblance de chaque historique $\{i, j\} = 1, \dots, r$:

$$\lambda_{ij}(k) = p[z(k) | M_j(k), \hat{x}_i(k - 1|k - 1), P_i(k - 1|k - 1)] \quad i, j = 1, \dots, r \quad (1.44)$$

Sous l'hypothèse de normalité, on obtient :

$$\lambda_{ij}(k) = p[\nu_{ij}(k)] = N[\nu_{ij}(k); 0, S_{ij}(k)] \quad (1.45)$$

$$S_{ij}(k) = C_j A_j P_i(k - 1|k - 1) C_j^T A_j^T + R(k)$$

$$\nu_{ij}(k) = z(k) - C_j(A_j \hat{x}_i(k - 1|k - 1) + B_j u(k - 1))$$

où $S_{ij}(k)$ est la covariance du résidu $\nu_{ij}(k) = z(k) - C_j(A_j \hat{x}_i(k - 1|k - 1) + B_j u(k - 1))$ qui correspond à l'historique $\{i, j\}$. Ces vraisemblances vont servir au calcul des probabilités des historiques.

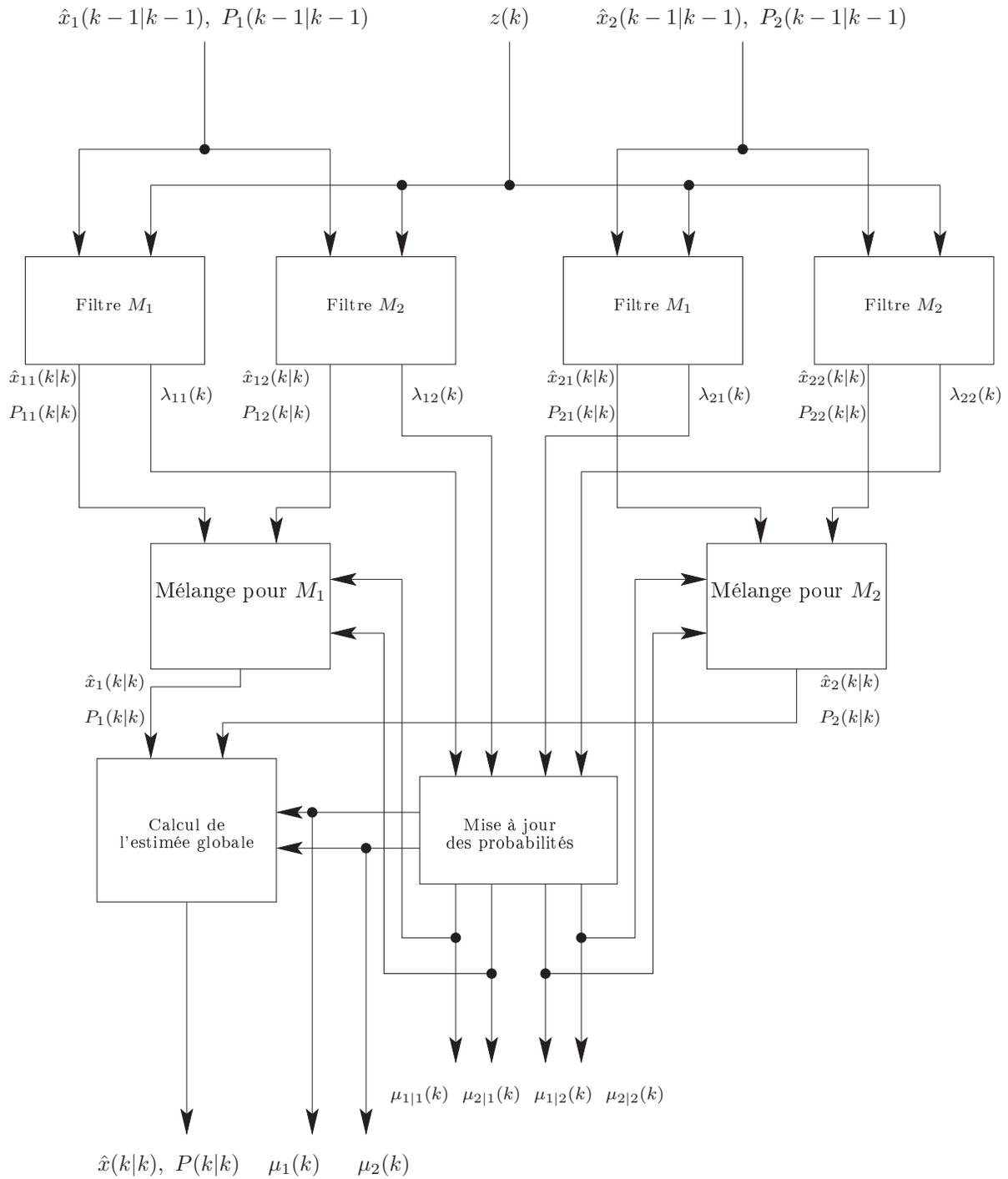


FIG. 1.5 – Estimateur GPB2

Calcul des probabilités de mélange

La probabilité pour que le modèle M_i soit actif à l'instant $k-1$ sachant que le modèle M_j est actif à l'instant k qu'on appellera probabilité de mélange est donnée par :

$$\mu_{i|j}(k-1|k) = P \{ M_i(k-1) | M_j(k), Z^k \} \quad (1.46)$$

Comme dans les calculs précédents, le vecteur des mesures cumulées Z^k a été séparé en Z^{k-1} et $z(k)$

$$\mu_{i|j}(k-1|k) = P \{ M_i(k-1) | z(k), M_j(k), Z^{k-1} \} \quad (1.47)$$

Puis en appliquant la formule de Bayes :

$$\begin{aligned} \mu_{i|j}(k-1|k) &= \frac{1}{c_j} p [z(k), M_j(k) | M_i(k-1), Z^{k-1}] P \{ M_i(k-1) | Z^{k-1} \} \\ &= \frac{1}{c_j} p [z(k) | M_j(k), M_i(k-1), Z^{k-1}] \\ &\quad \times P \{ M_j(k) | M_i(k-1), Z^{k-1} \} P \{ M_i(k-1) | Z^{k-1} \} \end{aligned} \quad (1.48)$$

La formule (1.48) peut être écrite sous la forme :

$$\mu_{i|j}(k-1|k) = \frac{1}{c_j} \lambda_{ij}(k) \pi_{ij} \mu_i(k-1) \quad i, j = 1, \dots, r \quad (1.49)$$

où $\mu_i(k-1)$ est la probabilité du modèle M_i à l'instant $k-1$, π_{ij} étant la probabilité de transition de Markov et c étant donnée par :

$$c_j = \sum_{i=1}^r \lambda_{ij}(k) \pi_{ij} \mu_i(k-1) \quad (1.50)$$

Calcul de l'estimée mélange

L'estimée $\hat{x}_j(k|k)$, correspondant au modèle M_j , est définie par :

$$\hat{x}_j(k|k) = \sum_{i=1}^r \hat{x}_{ij}(k|k) \mu_{i|j}(k-1|k) \quad j = 1, \dots, r \quad (1.51)$$

et sa covariance :

$$P_j(k|k) = \sum_{i=1}^r \mu_{i|j}(k-1|k) \left\{ P_{ij}(k|k) + [\hat{x}_{ij}(k|k) - \hat{x}_j(k|k)] [\hat{x}_{ij}(k|k) - \hat{x}_j(k|k)]^T \right\} \quad (1.52)$$

Mise à jour des probabilités d'activation

La probabilité d'activation $\mu_j(k)$ d'un modèle M_j à l'instant k est donnée par :

$$\begin{aligned}
 \mu_j(k) &= P \{ M_j(k) | z(k), Z^{k-1} \} \\
 &= \frac{1}{c} P \{ z(k), M_j(k) | Z^{k-1} \} \\
 &= \frac{1}{c} \sum_{i=1}^r P \{ z(k), M_j(k) | M_i(k-1), Z^{k-1} \} P \{ M_i(k-1) | Z^{k-1} \} \\
 &= \frac{1}{c} \sum_{i=1}^r p [z(k) | M_i(k-1), Z^{k-1}] P \{ M_j(k) | M_i(k-1), Z^{k-1} \} \mu_i(k-1)
 \end{aligned} \tag{1.53}$$

La mise à jour de la probabilité μ_j est alors donnée par :

$$\mu_j(k) = \frac{1}{c} \sum_{i=1}^r \lambda_{ij}(k) \pi_{ij} \mu_i(k-1) \tag{1.54}$$

où c , constante de normalisation, est définie par :

$$c = \sum_{j=1}^r \sum_{i=1}^r \lambda_{ij} \pi_{ij} \mu_i(k-1) \tag{1.55}$$

Cette expression permet la remise à jour des probabilités d'activation μ_j des r modèles.

Calcul de l'estimée globale

Les estimées issues de chaque filtre sont mélangées pour fournir l'estimée globale à chaque fin de cycle d'estimation. Cette estimée globale, utilisée comme entrée des filtres au cycle suivant, est donnée par :

$$\hat{x}(k|k) = \sum_{i=1}^r \hat{x}_i(k|k) \mu_i(k) \tag{1.56}$$

Sa covariance s'explique :

$$P(k|k) = \sum_{i=1}^r \mu_i(k) \left\{ P_i(k|k) + [\hat{x}_i(k|k) - \hat{x}(k|k)] [\hat{x}_i(k|k) - \hat{x}(k|k)]^T \right\} \tag{1.57}$$

1.4.6 Estimateur IMM (Interacting Multiple Model Estimator)

L'approche IMM permet l'obtention de performances comparables à l'approche précédente GPB2 en n'utilisant que r filtres au lieu de r^2 .

L'estimateur IMM fournit, à l'instant k , une estimée $\hat{x}(k|k)$ calculée en utilisant r filtres fonctionnant en parallèle. À l'entrée de chaque filtre, on utilise une combinaison des estimées calculées à l'instant précédent.

En appliquant le théorème de la probabilité totale, la densité de probabilité de $x(k)$ peut s'écrire :

$$\begin{aligned} p[x(k) | Z^k] &= \sum_{j=1}^r p[x(k) | M_j(k), Z^k] P\{M_j(k) | Z^k\} \\ &= \sum_{j=1}^r p[x(k) | M_j(k), Z^k] \mu_j(k) \end{aligned} \quad (1.58)$$

La densité de probabilité de $x(k)$ conditionnellement au modèle M_j et au vecteur des mesures cumulées Z^k est donnée par :

$$p[x(k) | M_j(k), z(k), Z^{k-1}] = \frac{p[z(k) | M_j(k), x(k)]}{p[z(k) | M_j(k), Z^{k-1}]} p[x(k) | M_j(k), Z^{k-1}] \quad (1.59)$$

Le dernier terme de l'équation (1.59) représente cette même densité à l'instant $k-1$. Le développement de ce terme par le théorème de la probabilité totale donne :

$$\begin{aligned} p[x(k) | M_j(k), Z^{k-1}] &= \\ &= \sum_{i=1}^r p[x(k) | M_j(k), M_i(k-1), Z^{k-1}] P\{M_i(k-1) | M_j(k), Z^{k-1}\} \end{aligned} \quad (1.60)$$

Ce développement permet d'introduire la notion d'historique où on prend en considération les modèles actifs à l'instant $k-1$. On effectue ensuite une approximation en supposant que l'information contenue dans l'historique des mesures Z^{k-1} peut se résumer à la connaissance des estimées $\hat{x}_l(k-1|k-1)$ et leurs covariances $P_l(k-1|k-1)$.

$$\begin{aligned} p[x(k) | M_j(k), Z^{k-1}] &\approx \\ &= \sum_{i=1}^r p[x(k) | M_j(k), M_i(k-1), \{\hat{x}_l(k-1|k-1), P_l(k-1|k-1)\}_{l=1}^r] \mu_{i|j}(k-1|k-1) \end{aligned} \quad (1.61)$$

$$\begin{aligned} p[x(k) | M_j(k), Z^{k-1}] &= \\ &= \sum_{i=1}^r p[x(k) | M_j(k), M_i(k-1), \hat{x}_i(k-1|k-1), P_i(k-1|k-1)] \mu_{i|j}(k-1|k-1) \end{aligned} \quad (1.62)$$

Le terme $\mu_{i|j}(k-1|k-1)$ correspond à la probabilité que le système ait fonctionné selon le modèle M_i à l'instant $k-1$ sachant qu'il fonctionne selon le modèle M_j à l'instant k et connaissant l'information Z^{k-1} .

La densité de probabilité $p [x(k) | M_j(k), Z^{k-1}]$ de l'équation (1.62) peut être représentée par un mélange de gaussiennes. l'équation (1.62) devient :

$$\begin{aligned}
 p [x(k) | M_j(k), Z^{k-1}] &= \\
 &\sum_{i=1}^r N [x(k); E [x(k) | M_j(k), \hat{x}_i(k-1|k-1)], P_i(k-1|k-1)] \mu_{i|j}(k-1|k-1) \\
 &\approx N \left[x(k); \sum_{i=1}^r E [x(k) | M_j(k), \hat{x}_i(k-1|k-1)] \mu_{i|j}(k-1|k-1), P_{0j}(k-1|k-1) \right] \\
 &= N \left[x(k); E \left[x(k) \middle| M_j(k), \sum_{i=1}^r \hat{x}_i(k-1|k-1) \mu_{i|j}(k-1|k-1) \right], P_{0j}(k-1|k-1) \right]
 \end{aligned} \tag{1.63}$$

où $P_i(k-1|k-1)$ est la variance de l'estimée $\hat{x}_i(k-1|k-1)$ et $P_{0j}(k-1|k-1)$ est la variance du mélange $\sum_{i=1}^r \hat{x}_i(k-1|k-1) \mu_{i|j}(k-1|k-1)$ qui sera exprimée par la suite. Cette formule est à la base de l'algorithme IMM. L'entrée de chaque filtre j à l'instant k est issue de la somme des r estimées $\hat{x}_i(k-1|k-1)$ des filtres à l'instant $k-1$ pondérée par les probabilités $\mu_{i|j}(k-1|k-1)$, appelées probabilités de mélange.

Cette façon de procéder permet de considérer un historique $\{M_i(k-1), M_j(k)\}$ tout en n'utilisant que r modèles au lieu des r^2 nécessaires pour la procédure GPB2.

La figure 1.6 visualise les principales étapes de l'algorithme, où r filtres fonctionnent en parallèle. Un mélange des estimées précédentes est réalisé à l'entrée de chaque filtre. Comme pour le GPB, on procède au calcul des probabilités des modèles et de l'estimée globale à l'issue de chaque cycle.

L'algorithme suit les étapes suivantes :

Calcul des probabilités de mélange

La probabilité que le modèle M_i soit actif à l'instant $k-1$, si le modèle M_j est actif à l'instant k , est donnée par :

$$\begin{aligned}
 \mu_{i|j}(k-1|k-1) &= P \{M_i(k-1) | M_j(k), Z^{k-1}\} \\
 &= \frac{1}{\bar{c}_j} P \{M_j(k) | M_i(k-1), Z^{k-1}\} P \{M_i(k-1) | Z^{k-1}\}
 \end{aligned} \tag{1.64}$$

avec

$$\bar{c}_j = \sum_{i=1}^r \pi_{ij} \mu_i(k-1) \quad j = 1, \dots, r \tag{1.65}$$

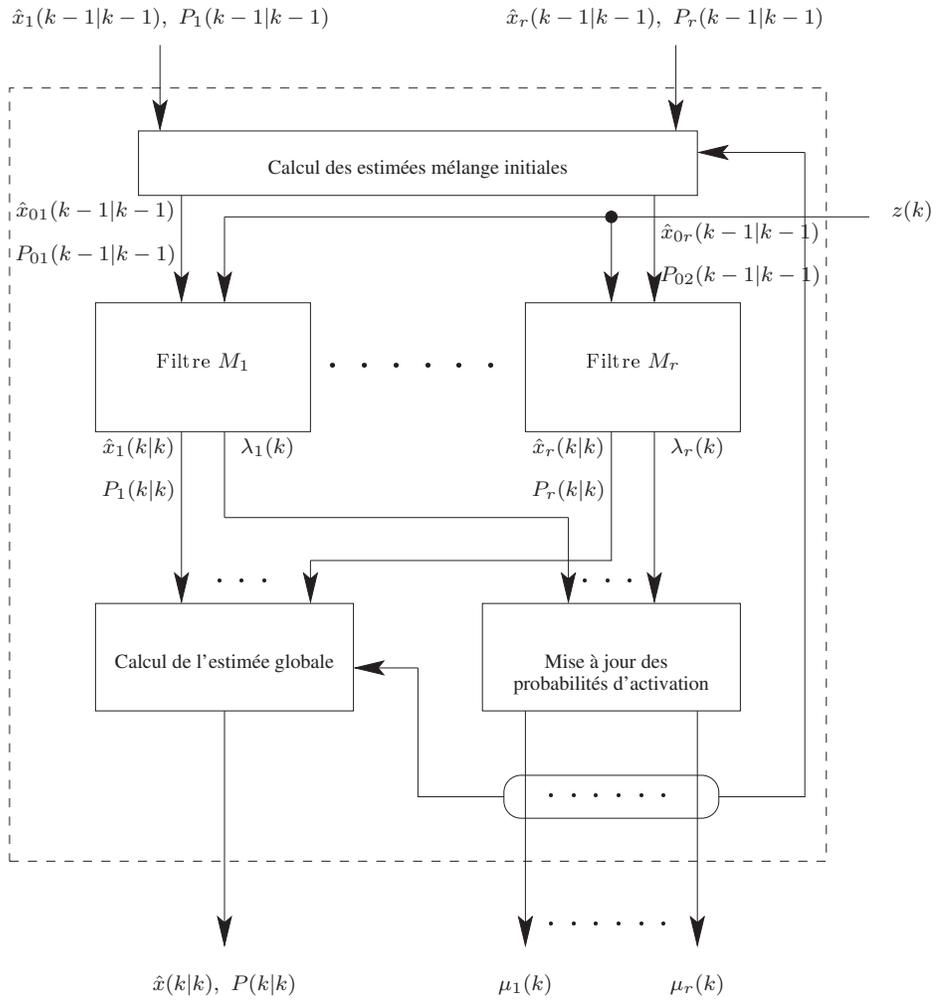


FIG. 1.6 – Estimateur IMM

Cette probabilité peut être réécrite en prenant en considération les probabilités de transition π_{ij} de Markov :

$$\mu_{i|j}(k-1|k-1) = \frac{1}{\tilde{c}_j} \pi_{ij} \mu_i(k-1) \quad i, j = 1, \dots, r \quad (1.66)$$

Les probabilités $\mu_{i|j}(k-1|k-1)$ vont être utilisées pour le calcul des estimées mélanges à l'entrée des filtres.

Calcul des estimées mélanges

Elles sont calculées en combinant les estimées issues des filtres calculées à l'instant précédent $\hat{x}_i(k-1|k-1)$ et les probabilités mélanges $\mu_{i|j}(k-1|k-1)$:

$$\hat{x}_{0j}(k-1|k-1) = \sum_{i=1}^r \hat{x}_i(k-1|k-1) \mu_{i|j}(k-1|k-1) \quad j = 1, \dots, r \quad (1.67)$$

La covariance associée à cette estimée s'écrit :

$$\begin{aligned} P_{0j}(k-1|k-1) &= \sum_{i=1}^r \mu_{i|j}(k-1|k-1) \{P_i(k-1|k-1) \\ &\quad + [\hat{x}_i(k-1|k-1) - \hat{x}_{0j}(k-1|k-1)] \\ &\quad \times [\hat{x}_i(k-1|k-1) - \hat{x}_{0j}(k-1|k-1)]^T\} \\ &\quad j = 1, \dots, r \end{aligned} \quad (1.68)$$

Le calcul des estimées mélanges $\hat{x}_{0j}(k-1|k-1)$ initiales est déduit directement de l'équation (1.63).

L'estimée mélange $\hat{x}_{0j}(k-1|k-1)$ sera utilisée à l'entrée du $j^{\text{ème}}$ filtre qui donnera en sortie l'estimée $\hat{x}_j(k-1|k-1)$ et sa variance $P_j(k-1|k-1)$.

Calcul des vraisemblances

La fonction de vraisemblance d'un modèle M_j à l'instant k est donnée par :

$$\lambda_j(k) = p[z(k) | M_j(k), Z^{k-1}] \quad (1.69)$$

elle est évaluée en utilisant l'estimée mélange $\hat{x}_{0j}(k-1|k-1)$ et sa variance :

$$\lambda_j(k) = p[z(k) | M_j(k), \hat{x}_{0j}(k-1|k-1), P_{0j}(k-1|k-1)] \quad j = 1, \dots, r \quad (1.70)$$

Sous l'hypothèse de la normalité des bruits entachant les mesures et les états du système la vraisemblance peut s'écrire :

$$\lambda_j(k) = p[\nu_j(k)] = N[\nu_j(k); 0, S_j(k)] \quad (1.71)$$

$$S_j(k) = C_j A_j P_j(k-1|k-1) C_j^T A_j^T + R(k)$$

$$\nu_j(k) = z(k) - C_j(A_j \hat{x}_{0j}(k-1|k-1) + B_j u(k-1))$$

où $S_j(k)$ est la variance du résidu $\nu_j(k)$ qui correspond au $j^{\text{ème}}$ modèle. Ces vraisemblances vont servir au calcul des probabilités des modèles.

Mise à jour des probabilités d'activation

La probabilité d'activation $\mu_j(k)$ d'un modèle M_j à l'instant k est donnée par :

$$\begin{aligned} \mu_j(k) &= P\{M_j(k) | Z^k\} \\ &= \frac{1}{c} p[z(k) | M_j(k), Z^{k-1}] P\{M_j(k) | Z^{k-1}\} \\ &= \frac{1}{c} \lambda_j(k) \sum_{i=1}^r P\{M_j(k) | M_i(k-1), Z^{k-1}\} P\{M_i(k-1) | Z^{k-1}\} \\ &= \frac{1}{c} \lambda_j(k) \sum_{i=1}^r \pi_{ij} \mu_i(k-1) \quad j = 1, \dots, r \end{aligned} \quad (1.72)$$

avec

$$c = \sum_{j=1}^r \lambda_j(k) \sum_{i=1}^r \pi_{ij} \mu_i(k-1) \quad (1.73)$$

Ces probabilités d'activation sont utilisées pour le calcul de l'estimée globale ainsi que pour la détection du mode actif à l'instant k comme nous allons le voir dans le chapitre 2.

Calcul de l'estimée globale

Les sorties de chaque filtre sont mélangées en une seule estimée globale à chaque fin de cycle d'estimation :

$$\hat{x}(k|k) = \sum_{i=1}^r \hat{x}_i(k|k) \mu_i(k) \quad (1.74)$$

La variance de cette estimée s'explique :

$$P(k|k) = \sum_{i=1}^r \mu_i(k) \left\{ P_i(k|k) + [\hat{x}_i(k|k) - \hat{x}(k|k)] [\hat{x}_i(k|k) - \hat{x}(k|k)]^T \right\} \quad (1.75)$$

Cette estimée qui constitue la sortie de l'estimateur ne sera pas cependant réutilisée pendant le cycle suivant.

1.5 Conclusion

Dans ce chapitre, nous avons rappelé, dans un premier temps, quelques notions concernant le filtre de Kalman et les systèmes à commutation markovienne, nécessaires à la compréhension et à la mise en oeuvre des estimateurs multi-modèle. Puis, nous avons présenté les différentes approches de l'estimation multi-modèle. La première génération des estimateurs multi-modèle est basée sur l'hypothèse exprimant que le modèle actif ne change pas avec le temps. Cette approche ne permet pas une bonne estimation dans le cas d'un changement fréquent de mode de fonctionnement. Des estimateurs conçus pour les systèmes à commutations markoviennes ont été ensuite définis. Plusieurs algorithmes ont été présentés ; ils diffèrent dans la manière dont ils utilisent l'historique des commutations ainsi que dans le mode de réinitialisation des filtres.

Dans le chapitre suivant, nous présentons une méthode de détection de défaut à base des estimateurs multi-modèle décrits dans ce chapitre.

2

Diagnostic des systèmes à commutation

Sommaire

2.1	Introduction	43
2.2	Principes fondamentaux du diagnostic	44
2.2.1	Surveillance	44
2.2.2	Diagnostic	44
2.2.3	Prise de décision	46
2.2.4	Performance d'une procédure de diagnostic	47
2.3	Estimateurs multi-modèles pour le diagnostic	48
2.3.1	Modèle de défaut	49
2.3.2	Détection de défaut	51
2.3.3	Construction de la matrice de Markov	53
2.3.4	Évaluation des performances de la méthode	54

2.4 Exemple de simulation	55
2.5 Conclusion	60

2.1 Introduction

C'est une évidence de constater que la commande des systèmes devient de plus en plus complexe ; cela est dû à la nature des systèmes, mais aussi à la volonté de contrôler tous les paramètres et toutes les perturbations affectant le système. Dans cette dynamique s'est développée la discipline de la sûreté de fonctionnement.

Pour un grand nombre d'applications, il est nécessaire d'implanter un système de surveillance afin de détecter, isoler, voire identifier tout dysfonctionnement.

Un système de surveillance doit permettre de caractériser le mode de fonctionnement d'un système à partir d'informations préalablement collectées, en reconnaissant et en indiquant les anomalies de comportement. Cette surveillance peut être réalisée en mode exploitation ou en mode hors exploitation, chacun des modes présente un certain nombre d'avantages et d'inconvénients.

Le mode d'implémentation *en exploitation* permet de réagir rapidement en cas de problèmes et s'accompagne souvent d'une procédure de maintenance sur site. Il impose un traitement en temps réel des différents signaux. Le mode d'implémentation *hors exploitation* permet de faire de la maintenance préventive et peut être utilisé en complément du mode d'implémentation en exploitation lorsque celui-ci ne permet pas de préciser la raison du dysfonctionnement du système. Les techniques de FDI sont généralement basées sur l'estimation de l'état de fonctionnement du système et l'analyse de cet état vis-à-vis d'états de référence traduisant le fonctionnement correct du système. En pratique, il est souvent plus simple de générer une estimation de la sortie du système. A partir de cette estimation, on calcule un résidu, différence entre la sortie mesurée et la sortie estimée, et qui a la propriété d'être sensible aux défauts. L'estimation d'état ou de sortie du système est à la base des méthodes de FDI et l'estimation par multi-modèle est un outil adapté pour la détection de défauts qui, elle aussi, se base sur le calcul de résidus.

Dans ce chapitre, nous allons utiliser les méthodes d'estimation multi-modèle pour la détection de défaut où le système sera représenté par un ensemble de modèles de bon fonctionnement et un ensemble de modèles représentant les défauts. Les probabilités d'activation des modèles issues des estimateurs du chapitre 1 vont servir à la détection, l'iso-

lation ainsi que l'identification du ou des défauts.

2.2 Principes fondamentaux du diagnostic

2.2.1 Surveillance

La surveillance d'un système repose sur l'observation des processus (ensemble de phénomènes organisés dans le temps) qui régissent le comportement du système. Cependant, la surveillance n'a pas seulement pour fonction de centraliser et visualiser les informations. Elle a également pour fonction de déterminer le fonctionnement courant du système, de détecter le passage d'un fonctionnement normal vers un fonctionnement anormal et de caractériser ce changement de fonctionnement en déterminant sa cause. Cette seconde fonction est désignée par le terme diagnostic.

2.2.2 Diagnostic

Le diagnostic consiste à détecter un fonctionnement anormal au sein du système et à déterminer sa cause en localisant le ou les composants du système présentant une anomalie de fonctionnement et, éventuellement, en caractérisant l'anomalie (sévérité, instant d'apparition, durée, etc.). Typiquement, le diagnostic débute par la comparaison entre le comportement ou le fonctionnement réel du système (dont une image est fournie par les observations) et le comportement ou le fonctionnement théorique attendu fourni par le modèle.

Modèle

Un système physique étant un ensemble constitué de composants interconnectés, un modèle d'un système est donc une description de sa structure physique et une représentation comportementale et/ou fonctionnelle, abstraite, de chacun de ses composants (Kleer and Williams, 1987). Le diagnostic peut ainsi faire appel à divers niveaux de modélisation, chacun se référant à des degrés de connaissance différents du système, structurelles, comportementales, fonctionnelles. La représentation structurelle décrit les interconnexions

des composants d'un système physique (la structure du système peut être représentée, par exemple, sous forme d'une matrice d'incidence). La représentation comportementale est constituée de relations entre les différents phénomènes qui régissent le comportement du système. Suite à la comparaison entre les comportements observé et attendu, toute incohérence est alors révélatrice d'une ou de plusieurs anomalies de comportement. Les incohérences sont ainsi considérées comme des signes de défauts ou des symptômes de défaillances. Introduisons maintenant la notion de défaillance et de défaut qui sont des termes très utilisés par la suite.

Défaillance

Une défaillance se rapporte à une anomalie fonctionnelle. L'adjectif défaillant est employé pour qualifier un système physique ou un composant dont une ou plusieurs fonctions sont altérées. Par abus, nous qualifions de défaillant un système physique ou un composant qui présente une anomalie fonctionnelle. En cas de cessation fonctionnelle, le terme panne est utilisé, c'est-à-dire l'inaptitude d'un composant à accomplir sa fonction.

Défaut

La notion de défaut, quant à elle, se rapporte à une anomalie de comportement au sein d'un système physique. Dans la littérature, un défaut est souvent défini comme tout écart entre la caractéristique observée sur le dispositif et la caractéristique de référence (Afnor, 1994). La notion de défaut est donc voisine de celle de défaillance. Toutefois, un défaut n'implique pas nécessairement une défaillance car un défaut, lié au comportement, est plus général qu'une défaillance, qui est quant à elle liée aux fonctions qui peuvent être toujours remplies malgré la présence d'un défaut. La description comportementale est plus détaillée que la description fonctionnelle et l'inclut donc. De la même manière, la notion de défaut inclut celle de défaillance ; un défaut n'altère pas nécessairement le fonctionnement d'un système physique mais peut présager d'une défaillance à venir.

Dans la pratique, un processus est composé de trois groupes d'organes : les actionneurs, les capteurs et le processus. Le comportement de l'ensemble dépend de chaque composant. Une défaillance peut donc surgir dans un ou plusieurs de ces organes.

Le diagnostic nécessite alors un ensemble de tests permettant de révéler des incohérences entre les observations et le modèle, signes de défauts ou symptômes de défaillances. L'objectif est ensuite de localiser le ou les composants ayant un comportement ou un fonctionnement anormal à partir des incohérences révélées par la procédure de diagnostic. Pour cela, un ensemble d'hypothèses sur l'origine possible des anomalies est généré. Chacune des hypothèses est formulée en termes de changements dans la structure, le comportement ou le fonctionnement du système. La localisation consiste à rechercher les hypothèses permettant d'expliquer l'ensemble des incohérences observées.

Détection et localisation

La détection consiste à prendre une décision binaire : soit le système fonctionne correctement, soit une panne s'est produite. Le résultat de la procédure de détection est une alarme signifiant que le fonctionnement réel du système ne concorde plus avec le modèle de fonctionnement sain.

Isoler revient à attribuer le défaut au module défectueux du système : capteur, actionneur, processus ou unité de commande.

Diagnostiquer consiste à effectuer la classification des défauts selon certains paramètres qui les caractérisent : instant d'apparition, amplitude. Cette étape consiste également à prévoir l'évolution des défauts et à quantifier leur degré de sévérité.

2.2.3 Prise de décision

Une fois le fonctionnement incorrect du système constaté, il est primordial d'agir de façon à maintenir les performances souhaitées ou à limiter les dégradations sur le système réel. Cette prise de décision permet de choisir entre plusieurs options comme arrêter le système pour faire de la maintenance ou accepter un fonctionnement dégradé. Il peut encore s'agir, quand cela est possible, de reconfigurer ou de réorganiser le système (Beard, 1971). Le rôle de la reconfiguration est de s'affranchir des conséquences du défaut pour conserver les performances initiales lorsque cela est possible ou encore d'assurer un fonctionnement dégradé du système si celui-ci est tolérable. Il est important que le

défaut soit identifié avec exactitude (étendue, amplitude, type, cause) afin d'en permettre sa compensation éventuelle. La reconfiguration peut porter sur le système de régulation (ou une partie de celui-ci), sur la structure de la loi de commande, sur le processus physique (en présence de redondance matérielle, on peut basculer sur les éléments ayant un fonctionnement correct).

2.2.4 Performance d'une procédure de diagnostic

La phase de détection est très importante dans le processus de surveillance du système. Si cette étape n'est pas correctement réalisée, les défauts peuvent être mal ou pas détectés ou de fausses alarmes peuvent apparaître. Les performances souhaitées d'un système de détection ont été décrites par (Patton et al., 1989) :

- la détection de défaut naissant,
- la rapidité de détection,
- l'isolation et la caractérisation des défauts détectés,
- la minimisation du nombre de fausses alarmes,
- la minimisation des mauvaises détections.

Les performances attendues d'une procédure de détection et d'isolation de défauts reposent sur la définition de critères de la méthode de diagnostic. Ils se décomposent en critères à minimiser :

- le retard à la détection,
- le taux de fausse alarme et de mauvaise détection,
- le temps de calcul pour une utilisation en temps réel,

et en critères à maximiser :

- la sensibilité à des défauts de faible amplitude,
- l'insensibilité aux bruits et aux perturbations ainsi qu'aux incertitudes sur les paramètres du modèle du système.

Les critères peuvent être contradictoires, ce qui nécessitera la mise en place d'une optimisation. Certaines méthodes de détection vont avantager certains de ces critères et pénaliser d'autres.

Pour les systèmes sujets à des bruits, les hypothèses sont souvent faites en considérant

ces bruits comme aléatoires et non-corrélés entre eux. Si les bruits sont non-stationnaires et non-gaussiens, alors les performances de la détection en seront affectées.

Une procédure de détection de défauts permet de détecter différents dysfonctionnements du système lorsque ce dernier n'est plus en fonctionnement normal (défauts de capteurs, d'actionneurs, variations de paramètres, changements de structure, présence de bruits, ...). Nous devons donc focaliser l'action de détection sur ce qui nous intéresse en s'affranchissant des autres phénomènes qui seront considérés comme des perturbations. La procédure de détection est alors conçue a priori pour des types particuliers de défauts.

2.3 Estimateurs multi-modèles pour le diagnostic

Dans cette partie, nous appliquerons les méthodes d'estimation d'état présentées au chapitre précédent pour la détection de défaut et le diagnostic des systèmes à commutation.

Les méthodes d'estimation multi-modèle sont élaborées pour des systèmes sujets à des changements structurels et/ou paramétriques. Elles sont naturellement adaptées à la détection de défaut car, dans la réalité, un système peut être souvent représenté par un ou plusieurs modèles de bon fonctionnement, mais également par un ensemble de modèles décrivant les situations de défaut de capteurs, d'actionneurs ou de dommages sur les autres composants du système. Ainsi, le fonctionnement global du système peut être décrit par un ensemble de modèles que l'on peut réunir dans une structure multi-modèle.

Les estimateurs multi-modèles ont été appliqués au diagnostic, dans un premier temps, sous leur version qui ne tient pas en compte des commutations (estimateur multi-modèle non commutant) (Gustafson and al., 1978), (Maybeck, 1982), (Menke and Maybeck, 1995), (Napolitano and Swaim, 1992). Cette approche est appropriée dans le cas d'incertitudes paramétriques et/ou structurelles, mais ne convient pas aux systèmes à commutation. Afin d'adapter ces estimateurs aux systèmes à commutation, plusieurs techniques ont été proposées dans (Gustafson and al., 1978), (Maybeck, 1982), (Menke and Maybeck, 1995) sans pour autant intégrer la notion de commutation markovienne dans l'algorithme d'estimation.

Des méthodes plus récentes, tel que le GPB et l'IMM (Bar-Shalom and Li, 1993), (Blom and Bar-Shalom, 1988), ont permis la prise en compte des commutations entre les différents modes de fonctionnement du système. Ces méthodes sont mieux adaptées à la détection de défaut affectant les systèmes à commutation et ont été appliquées avec succès au diagnostic dans (Tudoroiu and Khorasani, 2005), (Rago et al., 1998), (Hashimoto and Kawashima, 2001), (Fang et al., 1999), (Zhang and Li, 1998), (Zhang and Jiang, 2001). Elle s'appuient sur l'utilisation d'un banc de filtres disposés en parallèle, chaque filtre étant calé sur un modèle local associé à un comportement particulier du processus réel. A partir de l'évaluation des résidus des filtres, on peut détecter le modèle actif et donc détecter les défauts si ceux-ci sont représentés par des modèles appropriés.

2.3.1 Modèle de défaut

L'ensemble M des modèles sur lequel on va s'appuyer pour réaliser l'estimateur multi-modèle sera composé des modèles de *bon fonctionnement* et des modèles de *défaul*. Certains modèles du système en défaut peuvent être déduits des modèles de bon fonctionnement. Par exemple, on peut modéliser un défaut d'actionneur en modifiant la matrice de commande d'un modèle de bon fonctionnement.

La construction de l'ensemble des modèles de défaut et plus particulièrement leur nombre dépend largement de l'application considérée. L'importance de la construction de l'ensemble des modèles a été largement souligné dans la littérature et la difficulté principale dans l'élaboration des estimateurs multi-modèle est la construction de cet ensemble de manière à ce qu'il représente fidèlement le comportement du système. La construction de l'ensemble des modèles pour la détection de défauts nécessite des informations a priori sur le type du défaut.

Un défaut multiplicatif d'actionneur peut être modélisé par la modification d'une colonne de la matrice de commande B . Ainsi, un défaut sur le $i^{\text{ème}}$ actionneur est pris en compte en écrivant le système sous la forme :

$$x(k) = Ax(k-1) + (B + D_{a_i})u(k-1) + Gw(k-1) \quad (2.1)$$

où D_{a_i} est une matrice qui a la même dimension que B et dont toutes les composantes sont nulles sauf celles de la $i^{\text{ème}}$ colonne qui caractérise le défaut multiplicatif sur le $i^{\text{ème}}$ actionneur du fait de sa modification. Dans le cas d'une panne totale sur le $i^{\text{ème}}$ actionneur, on choisit la $i^{\text{ème}}$ colonne de D_{a_i} de manière à ce que sa somme avec la $i^{\text{ème}}$ colonne de B soit nulle.

De la même manière, on peut décrire un défaut multiplicatif de capteur par :

$$z(k) = (C + D_{c_j}) x(k) + v(k) \quad (2.2)$$

où D_{c_j} est une matrice qui a la même dimension que C et dont toutes les composantes sont nulles sauf celles de la $j^{\text{ème}}$ ligne qui caractérise le défaut multiplicatif sur le $j^{\text{ème}}$ capteur. Dans le cas d'une panne totale sur le $j^{\text{ème}}$ capteur, on choisit la $j^{\text{ème}}$ ligne de D_{c_j} de manière à ce que sa somme avec la $j^{\text{ème}}$ ligne de C soit nulle.

Des situations plus complexes peuvent être envisagées, incluant des défauts total ou partiel d'actionneur, des défauts affectant des capteurs ou d'autre composantes du système. Cette multiplicité de situations de défaut motive l'utilisation des multi-modèles et en particulier l'utilisation des systèmes à commutation markovienne où on prend en considération tout les modèles à chaque instant.

Le système considéré est représenté par un système à commutation markovienne défini par l'ensemble de modèles $M = \{M_1, M_2, \dots, M_r\}$ où r est le nombre de modèles. Le comportement du système, pour chaque modèle d'évolution d'état M_j , est caractérisé par les matrices A_j, B_j, C_j, G_j décrivant le système dynamique suivant :

$$\begin{cases} x(k) = A_j x(k-1) + B_j u(k-1) + w(k-1) \\ z(k) = C_j x(k) + v(k) \end{cases}$$

où, à l'instant k , $x(k) \in \mathbb{R}^n$ est le vecteur d'état, $u(k) \in \mathbb{R}^p$ est le vecteur de commande, $z(k) \in \mathbb{R}^q$ est la sortie du système, enfin $w(k) \in \mathbb{R}^n$ et $v(k) \in \mathbb{R}^q$ sont respectivement le bruit d'état et le bruit de mesure considérés blancs, gaussiens et centrés.

Les transitions d'un modèle à un autre suivent un processus markovien, défini par sa

matrice de transition Π :

$$\Pi = \begin{bmatrix} \pi_{11} & \cdots & \pi_{1r} \\ \vdots & \ddots & \vdots \\ \pi_{r1} & \cdots & \pi_{rr} \end{bmatrix}$$

2.3.2 Détection de défaut

L'évolution temporelle de la probabilité d'activation des modèles, calculée par les estimateurs multi-modèles du chapitre 1, donne des informations sur les modifications qui peuvent survenir en cours du fonctionnement du système. En effet, comme on l'a vu précédemment, à l'issue d'un cycle d'estimation, l'estimateur multi-modèle affecte une probabilité $\mu_i(k)$ à chaque modèle M_i . L'analyse de ces probabilités permet de détecter et d'isoler les défauts à partir de règles de décision définies à cet effet. Certaines d'entre elles sont construites de manière à minimiser l'influence des perturbations et à favoriser une prise de décision sans ambiguïté.

Règles de décision

Le choix de la règle de décision et de ses paramètres dépend du système, du niveau de bruit sur les signaux et de la distance entre les modèles. Plusieurs règles de décision peuvent être employées.

- Règle 1 : prise en compte de la probabilité d'activation la plus importante

$$\mu_i(k) = \max_j \mu_j(k) \quad (2.3)$$

- * si $\mu_i(k) \geq \delta$ le modèle i est déclaré actif à l'instant k ,
- * sinon le modèle i est déclaré inactif à l'instant k ainsi que les autres modèles,

La valeur du seuil $\delta \in [0, 1]$ dépend du nombre de modèles, du niveau de bruit et des performances souhaitées. Si l'on augmente δ , on risque de ne plus détecter le modèle actif et si on le diminue, on risque de faire des fausses détections (*fausse alarme*). Un nombre important de modèles implique une diminution de δ à cause

de la dispersion des probabilités (la somme des probabilités reste toujours égale à 1 tandis que leur nombre est plus important).

Une autre règle peut être envisagée afin d'exploiter le contraste entre les probabilités des différents modèles.

- Règle 2 : prise en compte du *contraste* entre les deux probabilités les plus importantes

$$\begin{cases} \mu_{i_1}(k) = \max_j \mu_j(k) \\ \mu_{i_2}(k) = \max_{j \neq i_1} \mu_j(k) \end{cases} \quad (2.4)$$

$$\beta_j(k) = \frac{\mu_{i_1}(k)}{\mu_{i_2}(k)} \quad (2.5)$$

* si $\beta_j(k) \geq \delta$ le modèle i_1 est déclaré actif à l'instant k ,

* sinon le modèle i_1 est déclaré inactif à l'instant k ainsi que les autres modèles.

Cette règle permet de diminuer le risque de fausse alarme. Même si la probabilité μ_{i_1} est faible, si la deuxième probabilité la plus élevée μ_{i_2} est faible proportionnellement à μ_{i_1} , alors l'activation du modèle i_1 est justifiée.

Avantages de la méthode

L'utilisation des méthodes d'estimation multi-modèle présentées dans le chapitre précédent offre de nombreux avantages pour le diagnostic par rapport aux autres méthodes. Ces avantages résident dans le fait que la prise de décision permet non seulement la détection du défaut, mais procure aussi des informations sur le type de défaut (défaut sur un actionneur ou sur un capteur ...), la localisation (quel est l'actionneur ou le capteur en défaut), l'amplitude du défaut ainsi que son instant d'occurrence. Autrement dit, la technique réalise simultanément la détection et le diagnostic.

La détection de défaut est faite en même temps que l'estimation de l'état du système ce qui donne un certain avantage à la méthode. L'estimée globale tient compte des défauts qui peuvent affecter le système et pourra donc être utilisée sans aucune reconfiguration de l'estimateur.

Cette approche peut être associée à une procédure de commande tolérante au défaut afin de permettre une reconfiguration de la commande. Dans (Zhang and Jiang, 2001) l'idée est de considérer un contrôleur issue d'une combinaison de contrôleurs associés aux différents modèles pondérée par les probabilités d'activation des modèles.

2.3.3 Construction de la matrice de Markov

La matrice de Markov est généralement considérée comme un paramètre de réglage de la méthode. Dans la majorité des cas, elle est déterminée préalablement à l'estimation d'état. Cependant, elle peut être également estimée en ligne. Dans le chapitre 3, une procédure d'estimation simultanée de la matrice de Markov et de l'état du système sera exposée.

La détermination a priori de la matrice de Markov pour les multi-modèles a été traitée dans (Bar-Shalom and Li, 1993), (Blackman and Popoli, 1999), (Blair and Watson, 1992), (Bloomer and Gray, 2002) et (Busch and Blackman, 1995). L'approche la plus utilisée consiste à prendre en compte le temps de séjour moyen sur un modèle (Bar-Shalom, 1990) (Li and Bar-Shalom, 1993). La probabilité de se maintenir sur un modèle déterminé est donnée par :

$$\pi_{ii} = \max\{l_i, 1 - \frac{T}{\tau_i}\} \quad (2.6)$$

où τ_i représente le temps de séjour moyen sur le $i^{\text{ème}}$ modèle, π_{ii} est la probabilité de de se maintenir sur le modèle M_i et T est la période d'échantillonnage, l_i représente la borne inférieure de la probabilité de transition. En pratique, la valeur de τ_i est supérieure à celle de la période d'échantillonnage T . Les probabilités de passage du modèle M_i vers les autres modèles sont données par :

$$\pi_{ij} = \frac{1 - \pi_{ii}}{r - 1}, \quad j = 1, \dots, r, \quad i \neq j \quad (2.7)$$

Une mauvaise détermination de la matrice de Markov peut introduire des retards sur la détection du mode et peut aussi augmenter la sensibilité vis-à-vis du bruit. C'est pour cette raison que l'on va exposer deux méthodes d'estimation de la matrice de Markov afin d'avoir une valeur de la matrice proche de la réalité.

2.3.4 Évaluation des performances de la méthode

Indice de performance

Afin d'évaluer la performance du diagnostic par les méthodes multi-modèles proposées dans le chapitre précédent des critères de performance ont été définis. Ces critères prennent en considération différents aspects tel que la détection, la localisation, le délai de détection et les fausses alarmes. Pour cela on définit les indicateurs suivants :

- le taux de fausses alarmes (FA) qui est le ratio entre le nombre de fausses alarmes et le nombre de fois où le système est en bon fonctionnement,
- le taux de non détection (ND) qui est le ratio entre le nombre de fois où la méthode n'a pas abouti à la détection du défaut et le nombre de fois où le système est en défaut,
- le taux de bonne détection est d'isolation (BDI) qui est le ratio entre le nombre de fois où on a détecté le bon mode de fonctionnement et le nombre d'échantillons,
- le taux d'isolation incorrect du défaut (IID) est le ratio entre le nombre de cas où on a une détection de défaut mais le défaut isolé n'est pas le bon et le nombre de fois où le système est en défaut,
- le temps moyen de détection (TMD) qui est donné par la moyenne des retards à la détection.

La règle 2 décrite précédemment est utilisée pour décider du modèle actif. Les différents critères sont calculés en comparant le modèle choisi par la règle 2 avec le vrai mode de fonctionnement. Pour l'évaluation de la méthode, un scénario des commutations est établi préalablement et les instants de commutation sont connus afin de pouvoir les comparer avec les décisions prises.

Scénario pour le test

Pour cela les commutations entre les modes de fonctionnement sont générés en utilisant la matrice de transition de Markov. On peut commuter aléatoirement d'un modèle à l'autre à n'importe quel instant suivant les probabilités de la matrice de Markov. Cependant, pour mieux s'approcher de la réalité, on suppose que le système est semi-markovien ce

qui revient à imposer le temps de séjour minimum t_i sur chaque mode.

D'autres scénarios déterministes ou aléatoires peuvent être élaborés selon le système traité et les performances visées.

Robustesse de la méthode

Les performances d'une procédure de détection peuvent être affectées par des variations ou des erreurs de modélisation des paramètres entrant dans l'élaboration de celle-ci. Plus la procédure de détection est insensible à ces variations et plus la détection s'effectue dans de meilleures conditions. Un des paramètres de la méthode proposée est la matrice de Markov ; il est intéressant d'étudier l'influence des erreurs de modélisation des termes de cette matrice sur la détection de défaut. Ce point a été abordé dans (Li, 1996), (Li and Bar-Shalom, 1993), où il est conclu que la performance de l'estimateur n'est pas très sensible aux erreurs de modélisation de cette matrice. Dans l'exemple qui suit, nous évaluons les performances de la détection par rapport aux changements des termes de la matrice de Markov.

Une procédure de détection est d'autant plus efficace qu'elle n'est pas sensible aux bruits ; pour cette raison, il est intéressant d'étudier aussi les performances de la détection, caractérisées par les indices de performance définis précédemment, par rapport au niveau de bruit.

2.4 Exemple de simulation

Un exemple de simulation est proposé dans cette section pour montrer les performances de la méthode présentée en ce qui concerne la détection de défauts de capteurs et d'actionneurs d'un moteur à courant continu (Park et al., 2000). L'équation linéaire de l'évolution de l'état du système représentant la dynamique du moteur est donnée par :

$$\begin{aligned} x(k) &= \begin{bmatrix} -0.0005 & -0.0084 \\ 0.0517 & 0.8069 \end{bmatrix} x(k-1) + \begin{bmatrix} 0.1815 \\ 1.7902 \end{bmatrix} u(k-1) + w(k-1) \\ y(k) &= \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} x(k) + v(k) \end{aligned} \quad (2.8)$$

Le vecteur d'état $x(k)$ est le suivant :

$$x(k) = \begin{bmatrix} i(k) \\ \omega(k) \end{bmatrix} \quad (2.9)$$

où $i(k)$ représente le courant traversant le moteur et $\omega(k)$ sa vitesse angulaire et où u représente la tension à l'entrée du moteur.

Pour détecter les défauts sur les capteurs et les actionneurs, plusieurs modèles sont générés afin de représenter différentes situations de défaut capteur ou actionneur ou des défauts simultanés de capteur et d'actionneur.

Dans un premier temps, on suppose que l'on ne peut avoir qu'un seul défaut à la fois et que le défaut sera total (panne du composant). Les matrices des différents modèles sont données par :

$$A_i = \begin{bmatrix} -0.0005 & -0.0084 \\ 0.0517 & 0.8069 \end{bmatrix}, \quad i = 1 \dots 3$$

$$B_1 = \begin{bmatrix} 0.1815 & 1.7902 \end{bmatrix}^T, \quad C_1 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}.$$

$$B_2 = \begin{bmatrix} 0.1815 & 1.7902 \end{bmatrix}^T, \quad C_2 = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}.$$

$$B_3 = \begin{bmatrix} 0 & 1.7902 \end{bmatrix}^T, \quad C_3 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}.$$

$$B_4 = \begin{bmatrix} 0.1815 & 1.7902 \end{bmatrix}^T, \quad C_4 = \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}.$$

$$B_5 = \begin{bmatrix} 0.1815 & 0 \end{bmatrix}^T, \quad C_5 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}.$$

Le premier modèle représente le fonctionnement normal du moteur et les quatre autres représentent des situations de pannes d'actionneurs et de capteurs. Les matrices de variances des bruits sur l'état et la sortie sont données par :

$$Q = \begin{bmatrix} 0.001 & 0 \\ 0 & 0.001 \end{bmatrix}.$$

$$R = \begin{bmatrix} 0.01 & 0 \\ 0 & 0.4 \end{bmatrix}.$$

Les transitions d'un modèle à un autre sont régies par la matrice Π :

$$\Pi_1 = \begin{bmatrix} 0.95 & 0.0125 & 0.0125 & 0.0125 & 0.0125 \\ 0.05 & 0.95 & 0 & 0 & 0 \\ 0.05 & 0 & 0.95 & 0 & 0 \\ 0.05 & 0 & 0 & 0.95 & 0 \\ 0.05 & 0 & 0 & 0 & 0.95 \end{bmatrix}$$

L'entrée $u(k)$ choisie pour simuler le modèle (2.8) est présentée sur la figure 2.1 et la sortie $y(k)$ est représentée sur les figures 2.2 et 2.3. Les résultats de l'estimation sont représentés sur les figures 2.4 et 2.5 où on a superposé l'état réel $x(k)$ et son estimée $\hat{x}(k)$. Les probabilités des modèles M_1 , M_2 , M_3 , M_4 et M_5 ainsi que leurs fonctions d'activation réelles sont représentées respectivement sur les figures ??, ??, ??, ?? et 2.6. Le tableau 2.1 donne la valeur des différents critères de performances pour une simulation de la méthode sur 50000 instants d'échantillonnage et pour différents niveaux de bruit de sortie.

De la figure 2.6 on peut remarquer que les probabilités d'activation des modèles sont

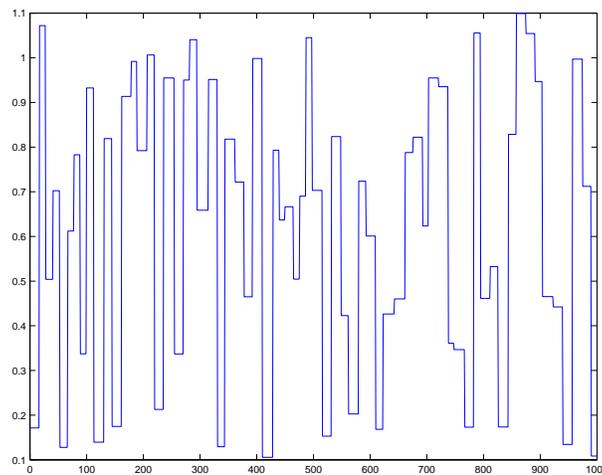


FIG. 2.1 – Entrée u du système

en adéquation avec les vrais occurrences des modèles et elles donnent des informations

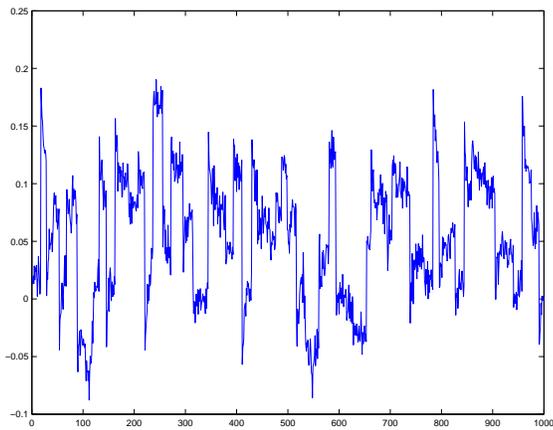


FIG. 2.2 – Courant i du système

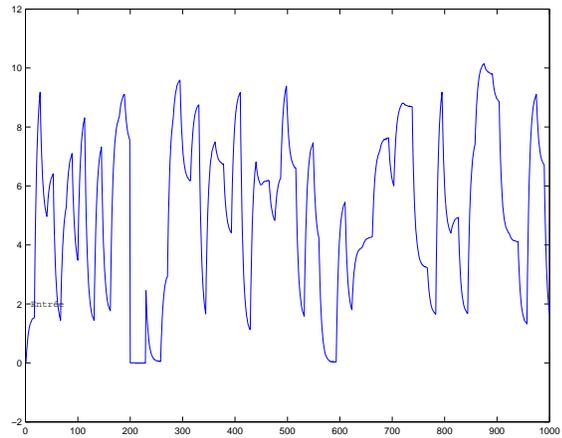


FIG. 2.3 – Vitesse angulaire ω

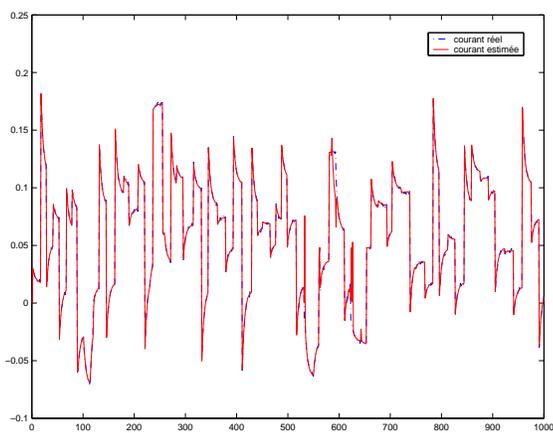


FIG. 2.4 – Estimée du courant i du système

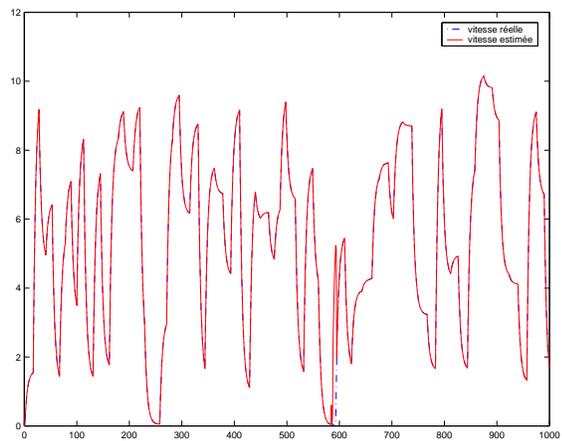
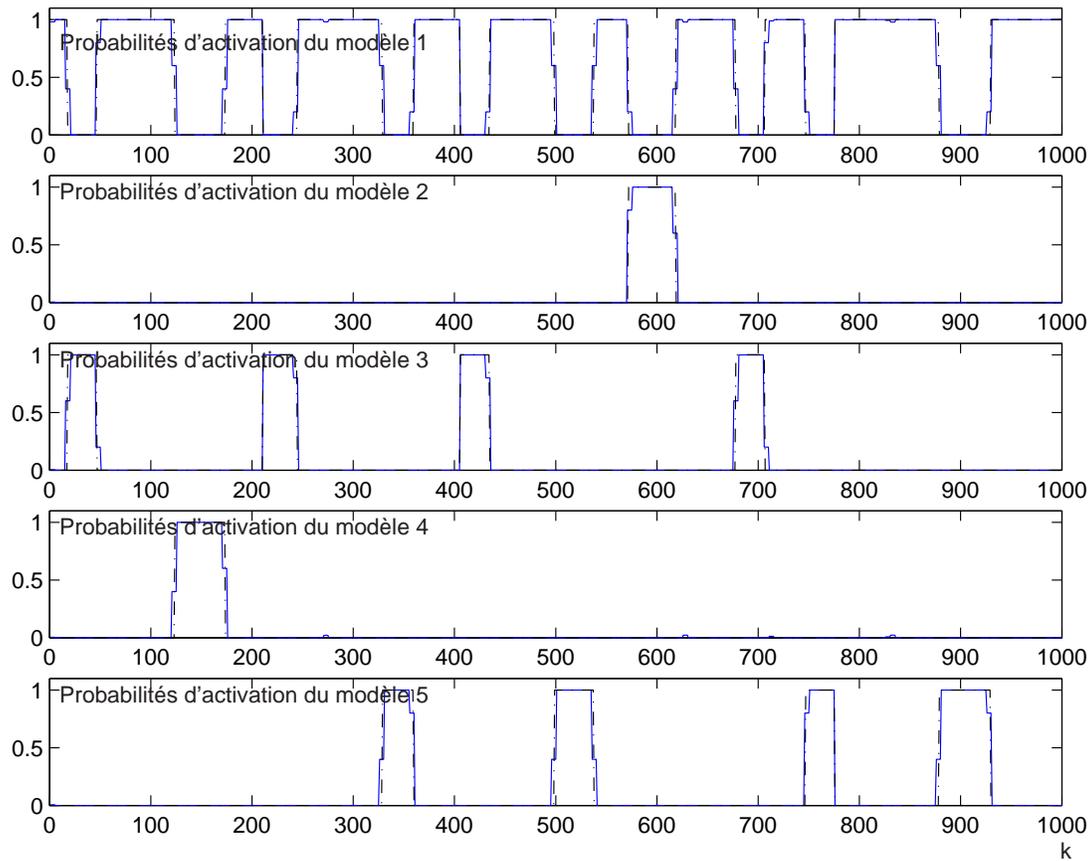


FIG. 2.5 – Estimée de la vitesse angulaire ω

FIG. 2.6 – Probabilités d'activation μ_i

pertinentes sur le "vrai" modèle actif. Afin d'apprécier l'influence des bruits de sortie sur la détection, on a fait plusieurs simulations avec différents niveaux de bruit ; les performances de la méthode de détection sont présentées dans le tableau 2.1 où on peut remarquer la détérioration des performances avec l'augmentation du niveau du bruit. Le tableau 2.2 représente les variations des performances de la détection en fonction des modifications effectuées sur la matrice de Markov. Pour cet exemple, les commutations sont générées en utilisant la matrice Π_1 et en gardant le même niveau de bruit pour toutes les simulations. L'estimation est réalisée quant à elle en utilisant les matrices Π_1 , Π_2 , Π_3 ou Π_4 . On peut remarquer que les performances de la détection se détériorent de plus en plus si l'on s'éloigne de la vraie matrice de transition Π_1 . Cela montre l'influence de cette matrice sur la détection et la nécessité de l'estimer dans le cas où elle n'est pas connue.

	BDI	FA	ND	IID	TMD
niveau de bruit R	0.9120	0.00024	0.1756	0.0066	2.4571
niveau de bruit $2 \times R$	0.8325	0.00078	0.2854	0.0371	4.2541
niveau de bruit $3 \times R$	0.7582	0.0014	0.3377	0.0619	5.7483
niveau de bruit $5 \times R$	0.6124	0.0018	0.4887	0.0869	10.3701

TAB. 2.1 – Performance de la détection en fonction du niveau du bruit de sortie

$$\Pi_2 = \begin{bmatrix} 0.97 & 0.0075 & 0.0075 & 0.0075 & 0.0075 \\ 0.03 & 0.97 & 0 & 0 & 0 \\ 0.03 & 0 & 0.97 & 0 & 0 \\ 0.03 & 0 & 0 & 0.97 & 0 \\ 0.03 & 0 & 0 & 0 & 0.97 \end{bmatrix}$$

$$\Pi_3 = \begin{bmatrix} 0.90 & 0.025 & 0.025 & 0.025 & 0.025 \\ 0.10 & 0.90 & 0 & 0 & 0 \\ 0.10 & 0 & 0.90 & 0 & 0 \\ 0.10 & 0 & 0 & 0.90 & 0 \\ 0.10 & 0 & 0 & 0 & 0.90 \end{bmatrix}$$

$$\Pi_4 = \begin{bmatrix} 0.80 & 0.05 & 0.05 & 0.05 & 0.05 \\ 0.20 & 0.80 & 0 & 0 & 0 \\ 0.20 & 0 & 0.80 & 0 & 0 \\ 0.20 & 0 & 0 & 0.80 & 0 \\ 0.20 & 0 & 0 & 0 & 0.80 \end{bmatrix}$$

2.5 Conclusion

Dans ce chapitre, les méthodes d'estimation multi-modèle ont été appliquées avec succès pour la détection et l'isolation de défaut. Cependant, la méthode de détection proposée, qui offre plusieurs avantages par rapport à d'autres méthodes de détection, reste sensible à plusieurs de ses paramètres (niveau de bruit, matrice des probabilités de

	BDI	FA	ND	IID	TMD
Π_1	0.9120	0.00024	0.1756	0.0066	2.4571
Π_2	0.8305	0.0023	0.1260	0.0846	2.4601
Π_3	0.8040	0.0020	0.1578	0.0772	2.7097
Π_4	0.7760	0.00096	0.1860	0.0669	3.1762

TAB. 2.2 – Performance de la détection en fonction de la précision de la matrice de Markov

transition de Markov (MPT) et nombre de modèles). Afin d'améliorer les performances de cette méthode de détection, nous proposons, dans les chapitres qui suivent, plusieurs améliorations et modifications en agissant à différents niveaux de la méthode d'estimation. Le chapitre 3 est consacré à une procédure d'estimation de la MPT. Cette procédure est indispensable dans le cas où l'on a peu ou pas d'information a priori sur cette matrice.

3

Estimation de la matrice de Markov

Sommaire

3.1	Introduction	64
3.2	Formulation du problème	65
3.3	Fonction de densité de probabilité de la matrice de Markov	67
3.4	Algorithme quasi bayésien	71
3.4.1	Estimation quasi bayésienne des paramètres d'une distribution mélange	72
3.4.2	Estimateur quasi bayésien de la MPT	79
3.5	Algorithme par intégration numérique	84
3.6	Conclusion	87

3.1 Introduction

Les Systèmes à Commutations Markovienne (SCM), commutent comme leur nom l'indique d'un modèle à un autre, suivant un processus de Markov dont les paramètres sont habituellement supposés connus. Dans ce chapitre, nous traitons le problème de l'estimation d'état d'un SCM dans le cas où la Matrice des Probabilités de Transition de la chaîne de Markov régissant les commutations est inconnue, sous l'hypothèse que la matrice de Markov est invariante dans le temps et représentée par une matrice aléatoire.

Dans la pratique, la matrice de Markov est souvent considérée comme un paramètre à calculer préalablement à une procédure d'estimation multi-modèle telle que présenté dans le chapitre 1. Plusieurs techniques de calcul de la matrice de Markov ont été proposées dans divers domaines tel que le suivi de trajectoire d'une cible (Bar-Shalom et al., 1989) ou la détection et l'isolation de défaut (Zhang and Jiang, 2001). Ces techniques sont conçues pour une application où l'on a des informations a priori sur les probabilités de passage d'un modèle à un autre et où l'on n'exploite pas les données en ligne. Cependant, pour beaucoup d'applications, les informations a priori sur la matrice de Markov, tels que la fréquence des pannes sur un système donné et leur durée moyenne, peuvent être insuffisantes ou même erronées. Utiliser une matrice de Markov imprécise, comme on l'a vu dans le chapitre précédent, peut conduire à une mauvaise détection du modèle actif, d'où l'intérêt de l'estimation en ligne de la matrice de Markov.

Dans la littérature, il existe quelques publications consacrées à la résolution du problème de l'estimation de la matrice de Markov dans le cadre des SCM.

Dans (Sawaragi et al., 1973), le problème a été considéré pour le cas simple d'une chaîne de Markov binaire modélisant des défauts de mesure (observations interrompues) où la résolution est basée sur l'approche bayésienne qui permet d'élaborer un estimateur approximatif de la matrice de Markov. Pour le même cas restrictif de chaîne de Markov binaire, (Tugnait and Haddad, 1979b) et (Tugnait and Haddad, 1980) ont développé et analysé la convergence d'un estimateur de Maximum de Vraisemblance (MV) pour la matrice de Markov. Cependant, le temps de calcul de l'algorithme basé sur le MV est exponentiellement croissant en fonction du temps. Pour contourner cette difficulté, un

schéma approximatif a été proposé dans (Tugnait, 1982). L'estimation de la matrice de Markov a aussi été considérée dans (Goutsias and Mendel, 1988), (West and Harrison, 1997), (Pavlovic et al., 1999) et (Ghahramani and Hinton, 2000). Dans (Pavlovic et al., 1999), le SCM a été formulé comme un réseau bayésien dynamique et un estimateur de la matrice de Markov basé sur la méthode MV a été proposé ainsi que dans (Ghahramani and Hinton, 2000).

Nous avons choisi de présenter l'approche traitée par (Jilkov and Li, 2004) pour sa facilité de mise en œuvre et sa possibilité d'intégration dans un contexte d'estimation MM (estimation simultanée de la matrice de Markov et de l'état du MM). L'approche est basée sur l'hypothèse que la matrice de Markov est inconnue mais invariante dans le temps et aléatoire, avec une distribution définie sur un espace continu. Toujours dans un cadre bayésien, nous obtenons une approximation récursive de la fonction de densité de probabilité de la matrice de Markov.

Dans le paragraphe qui suit, une formulation du problème d'estimation de la matrice de Markov est proposée. Ensuite, une approximation de la fonction de densité de probabilité de cette matrice est développée. Puis, un estimateur quasi bayésien de la matrice de Markov est proposé. Enfin, on présente un algorithme numérique d'estimation de la matrice de Markov.

3.2 Formulation du problème

On considère le SCM (1.17) décrit par le modèle d'évolution d'état du système :

$$\begin{cases} x(k) = A_j x(k-1) + B_j u(k-1) + w(k-1) \\ z(k) = C_j x(k) + v(k) \end{cases}$$

où, à l'instant k , $x(k) \in \mathbb{R}^n$ est le vecteur d'état, $u(k) \in \mathbb{R}^p$ est le vecteur de commande, $z(k) \in \mathbb{R}^q$ est la sortie du système, enfin $w(k) \in \mathbb{R}^n$ et $v(k) \in \mathbb{R}^q$ sont respectivement le bruit d'état et le bruit de mesure considérés blancs, gaussiens et centrés. j étant l'indice du modèle actif à l'instant k , on notera que $M_j(k)$ est l'événement correspondant à l'activation du modèle M_j caractérisé par les matrices A_j , B_j , C_j connues. Tous les modèles sont regroupés dans l'ensemble $M = \{M_1, M_2, \dots, M_r\}$ où r est le nombre de modèles.

Les probabilités d'activation initiales des modèles $\mu_j(0)$ et les probabilités de transition d'un modèle à un autre sont données par :

$$\begin{aligned} P\{M_j(0)\} &= \mu_j(0) \\ P\{M_j(k) | M_i(k-1)\} &= P\{M_j(k) | M_i(k-1), Z^{k-1}\} \\ &= \pi_{ij}, \quad i, j = 1, \dots, r \end{aligned} \quad (3.1)$$

où $Z^k = \{z(1), \dots, z(k)\}$ est le vecteur des mesures. Si tous les paramètres $A_j, B_j, C_j, \mu_j(0)$ et π_{ij} des équations (1.17, 3.1) sont connus, on peut procéder à l'estimation d'état pour obtenir $\hat{x}(k) = E[x(k) | Z^k]$ en utilisant les techniques GPB où IMM. Mais, en l'absence de connaissances fiables de la Matrice de Markov, les techniques d'estimation multi-modèle peuvent se révéler inefficaces.

Considérons maintenant le problème d'estimation de l'état x régi par le modèle (1.17) en l'absence d'informations sur la matrice de probabilité de transition Π définie dans (3.1) avec $\Pi = [\pi'_1, \pi'_2, \dots, \pi'_r]'$ et $\pi_i = [\pi_{i1}, \pi_{i2}, \dots, \pi_{ir}]', i = 1, \dots, r$.

Afin d'estimer la matrice de Markov en même temps que l'état du système, il est nécessaire d'estimer la matrice de Markov récursivement où à chaque nouveau cycle d'estimation, l'estimée de la matrice de Markov $\hat{\Pi}(k)$, à l'instant k , est calculée en utilisant l'estimée $\hat{\Pi}(k-1)$ de la matrice de Markov, à l'instant $k-1$, et les nouvelles informations procurées par la mesure $z(k)$, à l'instant k . Le problème d'estimation est formulé de la façon suivante :

- *étape (1)* : à l'instant k , exécuter un algorithme d'estimation multi-modèle (par exemple : GPB ou IMM) en utilisant l'estimée d'état précédente et les probabilités d'activation des modèles $\mu_i(k-1)$ et l'estimée $\hat{\Pi}(k-1)$, à l'instant $k-1$, de la matrice de transition, pour mettre à jour l'estimée de x , les probabilités d'activation des modèles et leurs vraisemblances définies dans le chapitre 1. Les informations issues de cette étape sont les suivantes :

$$\hat{x}(k|k) = E[x(k)|Z^k] \quad (3.2)$$

$$\begin{aligned} \mu_i(k) &= P\{M_i(k) | \hat{\Pi}(k-1), Z^k\} \\ \mu(k) &= [\mu_1(k), \dots, \mu_r(k)]' \end{aligned} \quad (3.3)$$

$$\begin{aligned}\lambda_i(k) &= p \left[z(k) \mid M_i(k), \hat{\Pi}(k-1), Z^{k-1} \right] \\ \Lambda(k) &= [\lambda_1(k), \dots, \lambda_r(k)]'\end{aligned}\quad (3.4)$$

- *étape (2)* : mettre à jour l'estimée $\hat{\Pi}(k)$ de la matrice de Markov on se basant sur $\hat{\Pi}(k-1)$ et sur les informations fournies par l'*étape (1)*.

Notons que cette procédure ne se restreint pas seulement à la méthode IMM où GPB (*étape (1)*) mais peut s'appliquer à d'autres méthodes d'estimation. Pour réaliser l'*étape (2)*, il faut avoir un estimateur récursif de la matrice de Markov afin de calculer $\hat{\Pi}(k)$ en fonction de $\hat{\Pi}(k-1)$.

Les différentes étapes de la procédure d'estimation sont représentées sur la figure 3.1.

La partie qui suit permet d'exprimer la densité de probabilité de la matrice de Markov, à l'instant k , en fonction de celle à l'instant $k-1$. Cette récursion est à la base des deux méthodes d'estimation de la matrice de Markov exposées par la suite.

3.3 Fonction de densité de probabilité de la matrice de Markov

Dans cette partie, une relation récursive est élaborée en vue de la mise à jour de la PDF de la matrice de Markov (Jilkov and Li, 2004). Elle sera exprimée en fonction des probabilités d'activation des modes de fonctionnement et des fonctions de vraisemblance définies dans (3.3, 3.4).

L'objectif est de trouver une relation approximative entre $p[\Pi | Z^k]$ et $p[\Pi | Z^{k-1}]$. En utilisant le théorème de la probabilité totale on a :

$$\begin{aligned}p[z(k) | \Pi, Z^{k-1}] &= \sum_{j=1}^r p[z(k) | M_j(k), \Pi, Z^{k-1}] P\{M_j(k) | \Pi, Z^{k-1}\} \\ &= \sum_{j=1}^r p[z(k) | M_j(k), \Pi, Z^{k-1}] \sum_{i=1}^r P[M_j(k) | M_i(k-1), \Pi, Z^{k-1}] \\ &\quad \times P\{M_i(k-1) | \Pi, Z^{k-1}\}\end{aligned}\quad (3.5)$$

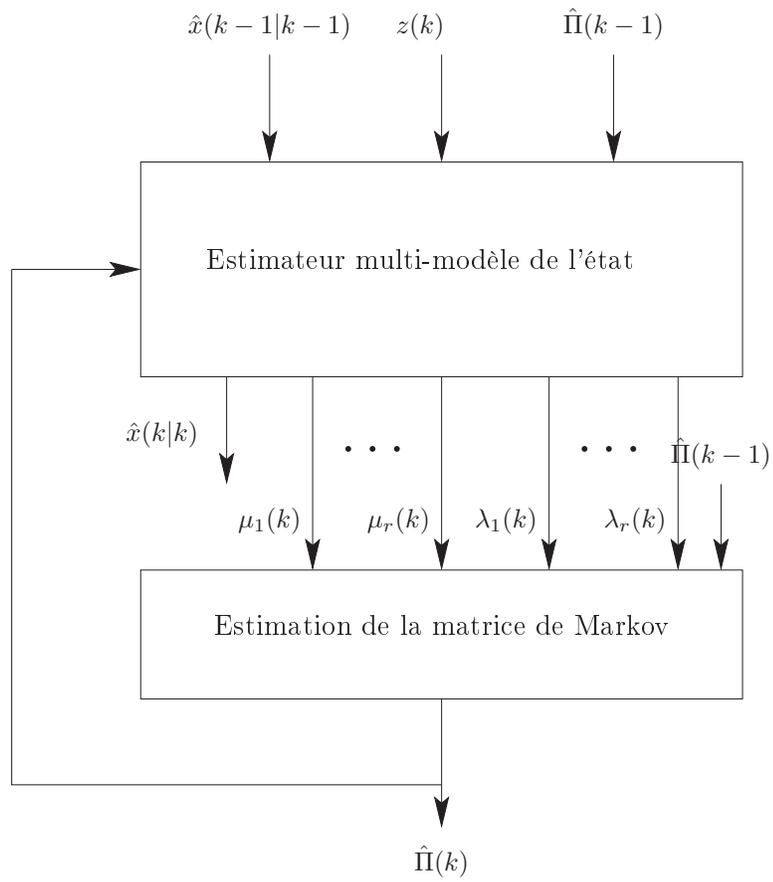


FIG. 3.1 – Procédure d'estimation de la matrice de Markov

La matrice Π n'est pas connue. On utilise alors les informations issues de l'étape (1) de l'algorithme précédent en considérant les approximations suivantes :

$$\begin{aligned} p [z(k) | M_j(k), \Pi, Z^{k-1}] &\approx \lambda_j(k) \\ P \{M_i(k-1) | \Pi, Z^{k-1}\} &\approx \mu_i(k-1) \end{aligned} \quad (3.6)$$

La substitution dans (3.5) conduit à :

$$\begin{aligned} p [z(k) | \Pi, Z^{k-1}] &\approx \sum_{j=1}^r \lambda_j(k) \sum_{i=1}^r \pi_{ij} \mu_i(k-1) \\ &= \Lambda'(k) \Pi' \mu(k-1) \\ &= \mu'(k-1) \Pi \Lambda(k) \end{aligned} \quad (3.7)$$

où Π' représente la matrice transposée de Π .

Notons que les approximations (3.6) consistent à remplacer la matrice inconnue Π par sa meilleure estimée $\hat{\Pi}(k-1)$ obtenue à l'instant $k-1$. Pratiquement, cela revient à approximer la fonction $p [z(k) | \Pi, Z^{k-1}]$ par une fonction linéaire de Π donnée par (3.7).

En considérant l'approximation (3.7), $p [z(k) | Z^{k-1}]$ peut être écrite sous la forme suivante :

$$\begin{aligned} p [z(k) | Z^{k-1}] &= \int p [z(k) | \Pi, Z^{k-1}] p [\Pi | Z^{k-1}] d\Pi \\ &= \int \mu'(k-1) \Pi \Lambda(k) p [\Pi | Z^{k-1}] d\Pi \end{aligned} \quad (3.8)$$

A partir de la définition

$$\hat{\Pi}(k-1) = E[\Pi | Z^{k-1}] = \int \Pi p [\Pi | Z^{k-1}] d\Pi \quad (3.9)$$

de l'estimée de la matrice de Markov, on obtient alors :

$$p [z(k) | Z^{k-1}] = \mu'(k-1) \hat{\Pi}(k-1) \Lambda(k) \quad (3.10)$$

En appliquant la formule de Bayes à $p [\Pi | Z^k]$, puis en remplaçant par les résultats obtenus (3.7) et (3.10) on obtient :

$$\begin{aligned} p [\Pi | Z^k] &= \frac{p [z(k) | \Pi, Z^{k-1}] p [\Pi | Z^{k-1}]}{p [z(k) | Z^{k-1}]} \\ &= \frac{\mu'(k-1) \Pi \Lambda(k)}{\mu'(k-1) \hat{\Pi}(k-1) \Lambda(k)} p [\Pi | Z^{k-1}] \end{aligned} \quad (3.11)$$

L'équation (3.11) est le point de départ des méthodes d'estimation de la MPT proposées dans ce chapitre. Son principal intérêt réside dans le fait qu'elle offre une récurrence sur la densité de probabilité $p[\Pi | Z^k]$.

On peut découpler la densité de probabilité $p[\Pi | Z^k]$ de Π (3.11) afin d'obtenir une densité de probabilité $p[\pi_i | Z^k]$ pour chaque ligne $\pi_i, i = 1, \dots, r$ de la matrice Π dans le but d'estimer chaque ligne indépendamment des autres.

L'hypothèse suivante permet d'obtenir les récurrences entre $p[\pi_i | Z^k]$ et $p[\pi_i | Z^{k-1}]$:

$$\int \pi_l p[\pi_1, \dots, \pi_r | Z^{k-1}] d\pi_1 \dots d\pi_{i-1} d\pi_{i+1} \dots d\pi_r = \hat{\pi}_l(k-1) p[\pi_i | Z^{k-1}] \quad (3.12)$$

où $\hat{\Pi}_l$ est l'estimée de la $l^{\text{ème}}$ ligne de Π à l'instant k et $\forall l \in 1 \dots m$ et $l \neq i$.

Donc l peut être le numéro de n'importe quelle ligne de Π sauf la $i^{\text{ème}}$.

L'intégrale (3.12) revient à dire que la ligne π_l est indépendante de la ligne π_i ce qui implique que $p[\pi_1, \dots, \pi_r | Z^{k-1}] = p[\pi_1, \dots, \pi_{i-1}, \pi_{i+1}, \dots, \pi_r | Z^{k-1}] p[\pi_i | Z^{k-1}]$.

En utilisant (3.12), on obtient la récurrence suivante :

$$p[\pi_i | Z^k] = \{1 + \eta_i(k) [\pi_i - \hat{\pi}_i(k-1)]' \Lambda(k)\} p[\pi_i | Z^{k-1}] \quad (3.13)$$

où

$$\eta_i(k) = \frac{\mu_i(k-1)}{\mu'(k-1) \hat{\Pi}(k-1) \Lambda(k)} \quad (3.14)$$

Pour prouver ce résultat, notons $d\Pi/d\pi_i = d\pi_1 \dots d\pi_{i-1} d\pi_{i+1} \dots d\pi_r$, ce qui permet d'écrire :

$$p[\pi_i | Z^k] = \int p[\Pi | Z^k] d\Pi/d\pi_i \quad (3.15)$$

En utilisant l'équation (3.11), on obtient :

$$p[\pi_i | Z^k] = \frac{A_i}{\mu'(k-1) \hat{\Pi}(k-1) \Lambda(k)} \quad (3.16)$$

avec

$$A_i = \int \mu'(k-1) \Pi \Lambda(k) p[\Pi | Z^{k-1}] d\Pi/d\pi_i \quad (3.17)$$

Compte tenu des définitions de μ et de Π , on a aussi :

$$\begin{aligned}
 A_i &= \int \sum_{l=1}^r \mu_l(k-1) \pi'_l \Lambda(k) p[\Pi | Z^{k-1}] d\Pi / d\pi_i \\
 &= \sum_{l=1}^r \mu_l(k-1) \int \pi'_l p[\Pi | Z^{k-1}] [d\Pi / d\pi_i] \Lambda(k) \\
 &= \sum_{l \neq i}^r \mu_l(k-1) \int \pi'_l p[\Pi | Z^{k-1}] [d\Pi / d\pi_i] \Lambda(k) \\
 &\quad + \mu_i(k-1) \pi'_i \int p[\Pi | Z^{k-1}] [d\Pi / d\pi_i] \Lambda(k)
 \end{aligned} \tag{3.18}$$

et selon (3.12), on a

$$\begin{aligned}
 A_i &= \left[\sum_{l \neq i}^r \mu_l(k-1) \hat{\pi}'_l(k-1) + \mu_i(k-1) \pi'_i \right] \Lambda(k) p[\pi_i | Z^{k-1}] \\
 &= \left[\sum_{l=1}^r \mu_l(k-1) \hat{\pi}'_l(k-1) + \mu_i(k-1) \pi'_i - \mu_i(k-1) \hat{\pi}'_i(k-1) \right] \Lambda(k) p[\pi_i | Z^{k-1}] \\
 &= \left\{ \mu'(k-1) \hat{\Pi}(k-1) \Lambda(k) + \mu_i(k-1) [\pi'_i - \hat{\pi}'_i(k-1)] \Lambda(k) \right\} p[\pi_i | Z^{k-1}]
 \end{aligned} \tag{3.19}$$

Donc à partir de (3.16) et (3.19), on déduit (3.13) ce qui achève la démonstration.

L'intérêt de la formule (3.13) est que $p[\pi_i | Z^k]$ s'exprime en fonction de $p[\pi_i | Z^{k-1}]$.

Cette formule peut être utilisée pour l'estimation de la $i^{\text{ème}}$ ligne de Π . La formule (3.13) sera utile dans le cadre d'un algorithme d'estimation récursif des π_i .

L'objectif étant l'estimation de la MPT, on cherche à obtenir un algorithme de calcul de la moyenne $\hat{\pi}_i(k) = E[\pi_i | Z^k]$, $i = 1, \dots, r$. Pour cela, on propose deux méthodes qui utilisent les expressions récursives (3.11) et (3.13).

3.4 Algorithme quasi bayésien

Afin de résoudre le problème de l'estimation de la MPT, une approche basée sur l'approximation quasi bayésienne (Titterington et al., 1985) est utilisée. Dans cette approche, les auteurs considèrent le problème de l'estimation des paramètres de pondération d'une somme pondérée de distributions connues.

Pour appliquer le résultat de (Titterington et al., 1985), on transforme la vraisemblance $p[z(k) | \pi_i, Z^{k-1}]$ correspondant à chaque ligne π_i , $i = 1, \dots, r$ en se servant de (3.13) et

en utilisant la formule de Bayes afin de la mettre sous forme d'un mélange de vraisemblances :

$$\frac{p[z(k) | \pi_i, Z^{k-1}]}{p[z(k) | Z^{k-1}]} = 1 + \eta_i(k) [\pi_i - \hat{\pi}_i(k-1)]' \Lambda(k) \quad (3.20)$$

En utilisant les définitions de π_i et $\Lambda(k)$, on développe l'équation (3.20) comme suit :

$$\frac{p[z(k) | \pi_i, Z^{k-1}]}{p[z(k) | Z^{k-1}]} = \sum_{j=1}^r \{1 + \eta_i(k) [\lambda_j(k) - \hat{\pi}_i(k-1)]' \Lambda(k)\} \pi_{ij} \quad (3.21)$$

Posons

$$f[z(k) | \pi_i] = p[z(k) | \pi_i, Z^{k-1}] = \sum_{j=1}^r \pi_{ij} g_{ij}(k) \quad (3.22)$$

avec

$$g_{ij}(k) = \tilde{g}_{ij}(k) p[z(k) | Z^{k-1}] \quad (3.23)$$

$$\tilde{g}_{ij}(k) = 1 + \eta_i(k) [\lambda_j(k) - \hat{\pi}_i(k-1)]' \Lambda(k) \quad (3.24)$$

L'équation (3.22), qui est une somme pondérée de densités de probabilités, est de la même forme que celle utilisée dans (Smith and Makov, 1978) et (Titterington et al., 1985). La section suivante présente la méthode utilisée dans (Titterington et al., 1985) pour estimer les paramètres de pondération qui sont dans notre cas les π_{ij} . Cette méthode sera ensuite appliquée pour l'estimation de chaque ligne de la matrice de Markov Π .

3.4.1 Estimation quasi bayésienne des paramètres d'une distribution mélange

Dans cette partie, nous allons considérer le problème d'estimation des pondérations dans le cas d'un mélange de plusieurs fonctions de densités de probabilités. Considérons une série d'observations $X^k = \{x(1), x(2), \dots, x(k)\}$. Chacune de ces observations appartient à une des populations M_1, M_2, \dots, M_r , les populations M_i pouvant être assimilées à des modes de fonctionnement dans le cas d'un système à commutations. L'observation de rang k est classée en se basant sur les observations $x(1), x(2), \dots, x(k)$.

On suppose que la densité de probabilité des observations est donnée par :

$$f[x(k) | \theta] = \theta_1 f_1[x(k)] + \theta_2 f_2[x(k)] + \dots + \theta_r f_r[x(k)], \quad (3.25)$$

On considère le vecteur des pondérations $\theta = (\theta_1, \theta_2, \dots, \theta_r)$ et les fonctions de densité de probabilité f_1, f_2, \dots, f_r relatives à l'appartenance de $x(k)$ à chaque population M_i , où les θ_i sont positifs et leur somme est égale à un. La densité f_i est la distribution de la probabilité d'appartenance de l'observation $x(k)$ à la population M_i , et θ_i est la probabilité que l'observation appartienne à la population M_i . Connaissant les fonctions f_i et les observations $x(k)$, on se propose d'estimer les paramètres θ_i .

On peut remarquer que la densité de l'équation (3.25) est de la même forme que le résultat obtenu dans (3.13).

Recherche d'une formulation bayésienne

Soit $p[\theta]$ la densité de probabilité a priori de θ , la densité a posteriori $p[\theta|X^k] = p[\theta|x(1), x(2), \dots, x(k)]$ de θ avec $x(1), x(2), \dots, x(k)$ connues et $p_i[\theta|X^k]$ la densité a posteriori de θ , conditionnellement à ce que la $k^{\text{ème}}$ observation appartient à M_i .

Par le théorème de Bayes et pour $k \geq 1$ on a

$$p[\theta|X^k] \propto f[x(k)|\theta]p[\theta|X^{k-1}]. \quad (3.26)$$

On définit les variables aléatoires $\tilde{y}(1), \tilde{y}(2), \dots, \tilde{y}(k), \dots$ avec $\tilde{y}(k) = i$ si et seulement si $x(k)$ appartient à M_i , $i = 1, 2, \dots, r$. Alors, par le théorème de la probabilité totale, on a :

$$p[\theta|X^k] = \sum_{i=1}^r P\{\tilde{y}(k) = i|X^k\}p[\theta|\tilde{y}(k) = i, X^k] = \sum_{i=1}^r P\{\tilde{y}(k) = i|X^k\}p_i[\theta|X^k] \quad (3.27)$$

On a aussi par la formule de Bayes

$$P\{\tilde{y}(k) = i|X^k\} = \frac{p[x(k)|\tilde{y}(k) = i, X^{k-1}]P\{\tilde{y}(k) = i|X^{k-1}\}}{\sum_{j=1}^r p[x(k)|\tilde{y}(k) = j, X^{k-1}]P\{\tilde{y}(k) = j|X^{k-1}\}} \quad (3.28)$$

En considérant que la meilleure estimée de $\hat{\theta}_i$, à l'instant $k - 1$, est représentée par $P\{\tilde{y}(k) = i|X^{k-1}\}$ et que l'expression $p[x(k)|\tilde{y}(k) = i, X^{k-1}]$ représente $f_i[x(k)]$ alors l'équation (3.28) peut être formulée :

$$P\{\tilde{y}(k) = i|X^k\} = \frac{f_i[x(k)]\hat{\theta}_i(k-1)}{\sum_{j=1}^r f_j[x(k)]\hat{\theta}_j(k-1)} \quad (3.29)$$

Cette expression permettra le calcul de la probabilité d'appartenance d'une mesure $x(k)$ à une population en connaissant l'estimation de θ à l'instant $k-1$. Les probabilités $P\{\tilde{y}(k) = i|X^k\}$ permettent la classification des mesures en les associant à une des populations M_i en fonction de ces probabilités.

Solution quasi bayésienne

On suppose que $p(\theta)$, la densité a priori de θ , suit une loi de Dirichlet,

$$p[\theta] = D(\theta; \alpha_1(0), \alpha_2(0), \dots, \alpha_r(0)) = \frac{\Gamma(\alpha_1(0) + \alpha_2(0) + \dots + \alpha_r(0))}{\Gamma(\alpha_1(0)) \Gamma(\alpha_2(0)) \dots \Gamma(\alpha_r(0))} \prod_{i=1}^r \theta_i^{\alpha_i(0)-1} \quad (3.30)$$

les $\alpha_i(0)$ sont les paramètres initiaux de la distribution de Dirichlet avec $\alpha_i(0) > 0$, $i = 1, 2, \dots, r$. A partir de l'équation (3.27) et après avoir observé x_1 , on obtient alors :

$$p[\theta|X^1] = \sum_{i=1}^r P\{\tilde{y}(1) = i|X^1\} D(\theta; \alpha_1(0) + \delta_{i1}, \alpha_2(0) + \delta_{i2}, \dots, \alpha_r(0) + \delta_{ir}) \quad (3.31)$$

$\delta_{ij} = 0$ si $i \neq j$, $\delta_{ij} = 1$ si $i = j$. Selon les propriétés de la distribution de Dirichlet, si on sait à quelle population appartient x , alors on obtient :

$$p[\theta|X^1] = D(\theta; \alpha_1(0) + \Delta_{11}, \alpha_2(0) + \Delta_{12}, \dots, \alpha_r(0) + \Delta_{1r}) \quad (3.32)$$

où $\Delta_{1i} = 1$ si $x(1)$ appartient à la population M_i sinon $\Delta_{1i} = 0$. Cependant, on ne sait pas à quelle population $x(1)$ appartient ; par contre, on peut utiliser les probabilités d'appartenance $P\{\tilde{y}(1) = i|X^1\}$ de $x(k)$ aux populations M_i . L'approximation se traduit par le remplacement de Δ_{1i} par $P\{\tilde{y}(1) = i|X^1\}$ qui représente la probabilité d'appartenance de $x(k)$ à la $i^{\text{ème}}$ population, alors l'équation (3.32) devient :

$$p[\theta|X^1] = D(\theta; \alpha_1(0) + P\{\tilde{y}(1) = 1|X^1\}, \alpha_2(0) + P\{\tilde{y}(1) = 2|X^1\}, \dots, \alpha_r(0) + P\{\tilde{y}(1) = r|X^1\}) \quad (3.33)$$

En partant de cette équation et en la remettant à jour à chaque nouvelle donnée, on obtient $p(\theta|X^k)$ qui est représenté par une distribution de Dirichlet dont les paramètres sont donnés par :

$$\alpha_i(k) = \alpha_i(k-1) + P\{\tilde{y}(k) = i|X^k\}, i = 1, 2, \dots, r \quad (3.34)$$

où les $\alpha_i^{(k-1)}$ sont les paramètres de $P[\theta|X^{k-1}]$ et où le calcul de $P\{\tilde{y}(k) = i|X^k\}$ s'effectue grâce à l'équation (3.29).

En considérant que les meilleures estimées des θ_i sont leurs moyennes $E[\theta_i|X^k]$, les $\hat{\theta}_i(k)$ sont directement déduites des propriétés de la distribution de Dirichlet, la propriété de la moyenne de la distribution nous permet d'écrire :

$$\hat{\theta}_i(k) = \frac{\alpha_i(k-1) + P\{\tilde{y}(k) = i|X^k\}}{\alpha_0 + k} = \frac{\alpha_i(k)}{\alpha_0 + k} \quad (3.35)$$

avec $\alpha_0 = \sum_{i=1}^r \alpha_i(0)$. On remarquera que la contrainte $\sum_{i=1}^r \theta_i = 1$ est toujours respectée. L'approche proposée est facile à mettre en œuvre et peut se résumer par les deux étapes suivantes

$$\alpha_i(k) = \alpha_i(k-1) + \frac{f_i[x(k)]\alpha_i(k-1)}{\sum_{j=1}^r f_j[x(k)]\alpha_j(k-1)} \quad (3.36)$$

$$\hat{\theta}_i(k) = \frac{\alpha_i(k)}{\alpha_0 + k} \quad (3.37)$$

Les équations (3.36) et (3.37) permettent d'estimer les paramètres de pondération θ_i en utilisant les données $x(k)$ et les densités $f_i[x(k)]$, l'estimation ne nécessite pas un grand volume de calculs et ces équations vont être utilisées pour estimer la matrice de Markov. La convergence des $\hat{\theta}_i$ vers les θ_i a été étudiée par (Smith and Makov, 1978).

Le calcul des probabilités $P\{\tilde{y}(k) = i|X^k\}$ permet de classer les données $x(k)$. Leurs valeurs sont données par :

$$P\{\tilde{y}(k) = i|X^k\} = \frac{f_i[x(k)]\alpha_i(k-1)}{\sum_{j=1}^r f_j[x(k)]\alpha_j(k-1)} \quad (3.38)$$

où $P\{\tilde{y}(k) = i|X^k\}$ représente la probabilité d'appartenance de la donnée $x(k)$ à la population M_i .

Exemple 1

On considère un ensemble de données X^n qui suivent une loi de distribution mélange de deux lois gaussiennes $f_1[x(k)]$ et $f_2[x(k)]$ avec leurs pondérations θ_1 et θ_2

$$f[x(k)|\theta] = \theta_1 f_1[x(k)] + \theta_2 f_2[x(k)] \quad (3.39)$$

avec $f_1[x(k)] = N(x(k), 0, 1)$, $f_2[x(k)] = N(x(k), 3, 1)$, $\theta_1 = 0.3$ et $\theta_2 = 0.7$ où

$$N(x, \bar{x}, \sigma) = (2\pi\sigma)^{-1/2} e^{-1/2(x-\bar{x})'\sigma^{-1}(x-\bar{x})} \quad (3.40)$$

et $\bar{x} = E[x]$, $\sigma = E[(x - \bar{x})(x - \bar{x})']$.

Notons $N(V(k); \nu, \sigma)$ la distribution gaussienne de $V(k)$ de moyenne ν et de variance σ . Connaissant une série de données X^n , on cherche à estimer les paramètres θ_i par l'algorithme quasi bayésien en posant $\alpha(0) = 2$. On remarque, à la figure 3.3, que l'on

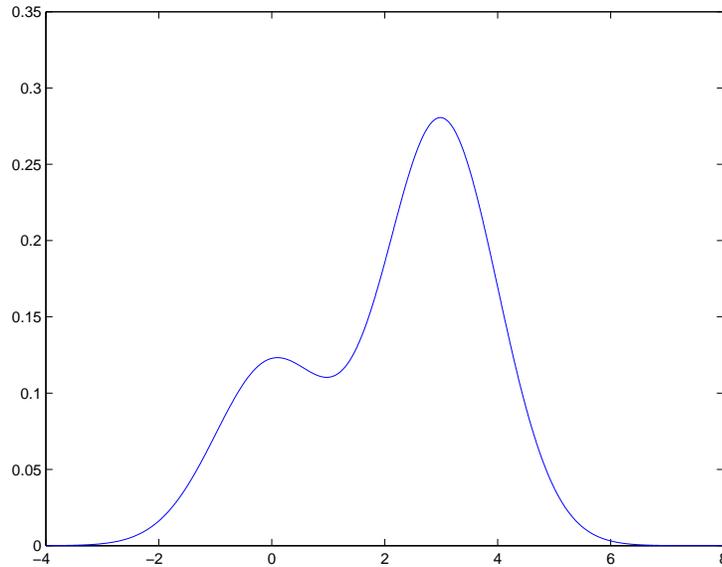


FIG. 3.2 – Densité mélange

obtient une convergence vers les valeurs réelles des paramètres du mélange malgré la proximité des deux densités de probabilité $f_1[x(k)]$ et $f_2[x(k)]$ représentées sur la figure 3.2.

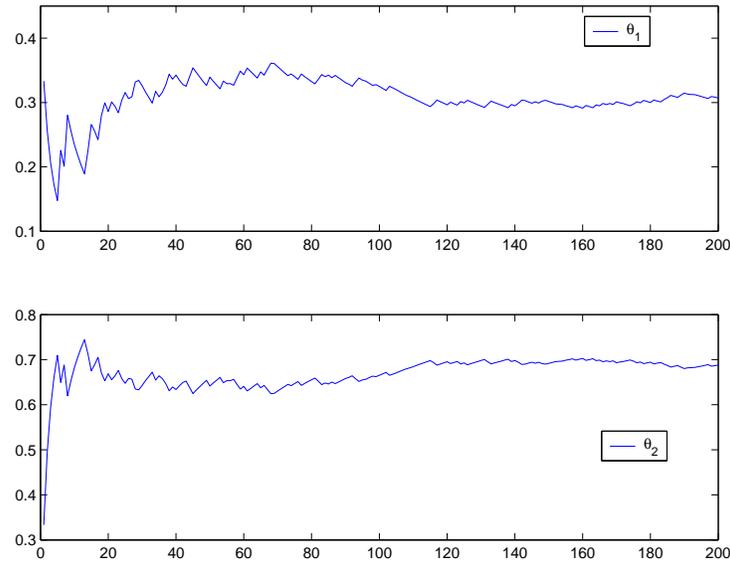
Exemple 2

On considère une variable aléatoire à deux dimensions qui suit une distribution mélange de deux lois gaussiennes $f_1[x(k)]$ et $f_2[x(k)]$ avec leurs pondérations θ_1 et θ_2 :

$$f[x(k)|\theta] = \theta_1 f_1[x(k)] + \theta_2 f_2[x(k)] \quad (3.41)$$

avec $f_1[x(k)] = N(x(k), \bar{x}_1, \sigma_1)$, $f_2[x(k)] = N(x(k), \bar{x}_2, \sigma_2)$, $\theta_1 = 0.3$ et $\theta_2 = 0.7$ où $\bar{x}_1 = [1, 2]$, $\bar{x}_2 = [3, -2]$ et $\sigma_1 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$, $\sigma_2 = \begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix}$.

Par l'algorithme quasi-bayésien, on estime les paramètres du mélange avec $\alpha(0) = 200$.

FIG. 3.3 – Convergence de θ_i

La figure 3.5 représente la convergence des estimées des paramètres θ_i et la figure 3.6 représente la classification des données $x(k)$. Cette dernière utilise les probabilités d'appartenance de chaque donnée aux populations M_i en choisissant à chaque fois la population qui a la probabilité la plus élevée pour un $x(k)$ donné. On observant les figures 3.3 et 3.5

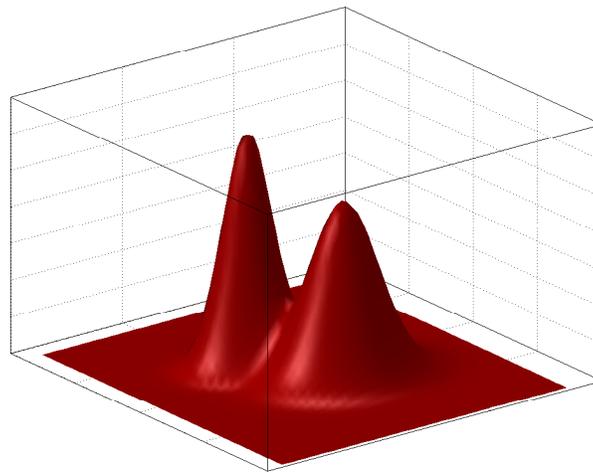


FIG. 3.4 – Densité mélange

on remarque une différence de vitesse de convergence. Cela est dû à la valeur de $\alpha(0)$, en

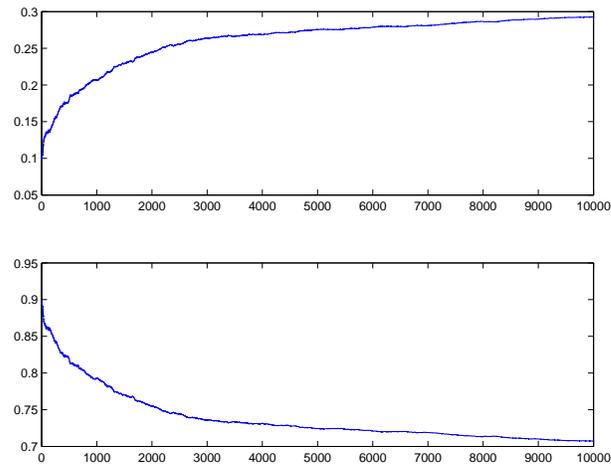


FIG. 3.5 – Convergence de θ_i

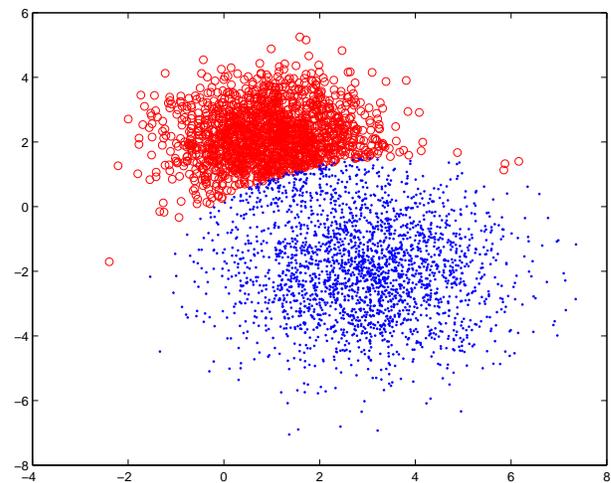


FIG. 3.6 – Classification des données

effet plus sa valeur est élevée plus la convergence est lente.

La méthode proposée dans (Titterington et al., 1985) considère le problème de la classification des données qui suivent une loi de distribution mélange. La loi de distribution de chaque population M_i est supposée connue, mais les paramètres du mélange restent à déterminer. Cette classification s'effectue à l'aide du calcul des probabilités d'appartenance à un groupe, pour chaque donnée, et l'estimation récursive des paramètres du mélange. Cette méthode d'estimation des paramètres du mélange va être adaptée à l'estimation des paramètres de la matrice de Markov en utilisant la formule (3.22).

3.4.2 Estimateur quasi bayésien de la MPT

En appliquant l'algorithme quasi bayésien de (Smith and Makov, 1978), représenté par les équations (3.36) et (3.37), au problème d'estimation de la MPT formulé par (3.22)-(3.24), on cherche à estimer la matrice des probabilités de Markov Π . Comme précisé dans l'étape 1 du paragraphe 3.2, les informations issues de l'estimateur multi-modèle sont les suivantes :

$$\begin{aligned}\hat{x}(k|k) &= E[x(k)|Z^k] \\ \mu_i(k) &= P \left\{ M_i(k) \mid \hat{\Pi}(k-1), Z^k \right\} \\ \mu(k) &= [\mu_1(k), \dots, \mu_r(k)]' \\ \lambda_i(k) &= p \left[z(k) \mid M_i(k), \hat{\Pi}(k-1), Z^{k-1} \right] \\ \Lambda(k) &= [\lambda_1(k), \dots, \lambda_r(k)]'\end{aligned}$$

Ces informations sont utilisées pour calculer les densités de probabilités $g_{ij}(k)$ de l'équation (3.23) afin de pouvoir appliquer la méthode d'estimation des paramètres de pondération d'une densité de probabilité mélange, présentée dans le paragraphe précédent, à l'estimation des π_{ij} . En appliquant la méthode de (Smith and Makov, 1978) à chaque ligne de la matrice Π , on obtient l'algorithme suivant :

Algorithme

- Initialisation :

$$\bar{\alpha}_i(0) = [\alpha_{i1}(0), \alpha_{i2}(0), \dots, \alpha_{ir}(0)]' \quad (3.42)$$

où i est relative à la $i^{\text{ème}}$ ligne de la matrice Π

$$\alpha_i(0) = \sum_{j=1}^r \alpha_{ij}(0), \quad i = 1, 2, \dots, r; \quad \alpha_{ij}(0) \geq 0. \quad (3.43)$$

$$\hat{\pi}_j(0) = \frac{1}{\alpha_i(0)} \bar{\alpha}_i(0), \quad i = 1, 2, \dots, r. \quad (3.44)$$

- Répéter pour $k = 1, 2, \dots$

- Répéter pour $i = 1, 2, \dots, r$

- Répéter pour $j = 1, 2, \dots, r$

$$\eta_i(k) = \frac{\mu_i(k-1)}{\mu'(k-1) \hat{\Pi}(k-1) \Lambda(k)} \quad (3.45)$$

$$\tilde{g}_{ij}(k) = 1 + \eta_i(k) [\lambda_j(k) - \hat{\pi}_i(k-1)' \Lambda(k)] \quad (3.46)$$

$$\alpha_{ij}(k) = \alpha_{ij}(k-1) + \frac{\alpha_{ij}(k-1) \tilde{g}_{ij}(k)}{\sum_{j=1}^r \alpha_{ij}(k-1) \tilde{g}_{ij}(k)} \quad (3.47)$$

$$\hat{\pi}_{ij}(k) = \frac{1}{k + \alpha_i(0)} \alpha_{ij}(k) \quad (3.48)$$

Notons que $\tilde{g}_{ij}(k)$ a été directement utilisé dans (3.47) à la place de $g_{ij}(k)$. Le vecteur des paramètres $\bar{\alpha}_i(0)$ représente les valeurs non normalisées de départ de la MPT $\bar{\pi}_j(0)$ qui sera normalisé par (3.44).

Pour l'application de cette méthode, on utilise l'algorithme IMM comme estimateur pour calculer les vraisemblances et les probabilités d'activation des modèles à l'instant k ((3.3) et (3.4)) qui sont essentielles à la procédure d'estimation de la matrice de Markov. En définitive, la méthode permet une estimation simultanée de l'état du système à commutation et de la matrice de transition de Markov. On considère un système à commutation à r modèles locaux linéaires :

$$\begin{cases} x(k) = A_i x(k-1) + B_i u(k-1) + w(k-1) \\ z(k) = C_i x(k) + v(k) \end{cases} \quad (3.49)$$

$i = 1, 2, \dots, r$

où, à l'instant k , $x(k) \in \mathbb{R}^n$ est le vecteur d'état, $u(k) \in \mathbb{R}^p$ est le vecteur de commande, $z(k) \in \mathbb{R}^q$ est la sortie du système, enfin $w(k) \in \mathbb{R}^q$ et $v(k) \in \mathbb{R}^h$ sont respectivement le bruit d'état et le bruit de mesure considérés blancs, gaussiens et centrés où Q et R sont les matrices de variance des bruits. En utilisant la méthode IMM, on obtient l'algorithme suivant :

Algorithme

- Initialisation :

$$\bar{\alpha}_i(0) = [\alpha_{i1}(0), \alpha_{i2}(0), \dots, \alpha_{ir}(0)]' \quad (3.50)$$

$$\alpha_i(0) = \sum_{j=1}^r \alpha_{ij}(0), \quad i = 1, 2, \dots, r; \quad \alpha_{ij}(0) \geq 0. \quad (3.51)$$

$$x(0), x_i(0), \mu_i(0) \quad (3.52)$$

$$\hat{\pi}_i(0) = \frac{1}{\alpha_i(0)} \bar{\alpha}_i(0), \quad i = 1, 2, \dots, r. \quad (3.53)$$

- Répéter pour $k = 1, 2, \dots$

$$\mu_j(k|k-1) = \sum_{i=1}^r \hat{\pi}_{ij}(k-1) \mu_i(k-1) \quad (3.54)$$

$$\mu_{i|j}(k-1) = \frac{\hat{\pi}_{ij}(k-1) \mu_i(k-1)}{\mu_j(k|k-1)} \quad (3.55)$$

$$\hat{x}_j^0(k-1|k-1) = \sum_{i=1}^r \hat{x}_i(k-1|k-1) \mu_{i|j}(k-1) \quad (3.56)$$

$$\begin{aligned} P_j^0(k-1|k-1) &= \sum_{i=1}^r P_i(k-1|k-1) + (\hat{x}_j^0(k-1|k-1) - \hat{x}_i(k-1|k-1)) \\ &\times (\hat{x}_j^0(k-1|k-1) - \hat{x}_i(k-1|k-1))^T \mu_{i|j}(k-1) \quad j = 1, 2, \dots, r \end{aligned} \quad (3.57)$$

$$\hat{x}_j(k|k-1) = A_j \hat{x}_j^0(k-1|k-1) + B_j u(k-1) \quad (3.58)$$

$$P_j(k|k-1) = A_j P_j^0(k-1|k-1) A_j^T + Q(k-1) \quad (3.59)$$

$$\nu_j(k) = z(k) - C_j \hat{x}_j(k|k-1) \quad (3.60)$$

$$S_j(k) = C_j P_j(k|k-1) C_j^T + R(k) \quad (3.61)$$

$$K_j(k) = P_j(k|k-1) C_j S_j^{-1}(k) \quad (3.62)$$

$$\hat{x}_j(k|k) = \hat{x}_j(k|k-1) + K_j(k) \nu_j(k) \quad (3.63)$$

$$P_j(k|k) = P_j(k|k-1) - K_j(k) S_j(k) K_j^T(k) \quad (3.64)$$

$$\lambda_j(k) = \frac{1}{\sqrt{2\pi |S_j(k)|}} \exp\left(-\frac{1}{2} \nu_j(k) S_j(k) \nu_j^T(k)\right) \quad (3.65)$$

$$\mu_j(k) = \frac{\mu_j(k|k-1) \lambda_j(k)}{\sum_{i=1}^r \mu_i(k|k-1) \lambda_i(k)} \quad (3.66)$$

$$\hat{x}(k|k) = \sum_{i=1}^r \hat{x}_i(k|k) \mu_i(k) \quad (3.67)$$

$$P(k|k) = \sum_{i=1}^r \mu_i(k) \left(P_i(k|k) + (\hat{x}(k|k) - \hat{x}_i(k|k)) (\hat{x}(k|k) - \hat{x}_i(k|k))^T \right) \quad (3.68)$$

- Répéter pour $i = 1, 2, \dots, r$

- Répéter pour $j = 1, 2, \dots, r$

$$\eta_i(k) = \frac{\mu_i(k-1)}{\mu'(k-1) \hat{\Pi}(k-1) \Lambda(k)} \quad (3.69)$$

$$\tilde{g}_{ij}(k) = 1 + \eta_i(k) [\lambda_j(k) - \hat{\pi}'_i(k-1) \Lambda(k)] \quad (3.70)$$

$$\alpha_{ij}(k) = \alpha_{ij}(k-1) + \frac{\alpha_{ij}(k-1) \tilde{g}_{ij}(k)}{\sum_{j=1}^r \alpha_{ij}(k-1) \tilde{g}_{ij}(k)} \quad (3.71)$$

$$\hat{\pi}_{ij}(k) = \frac{1}{k + \alpha_i(0)} \alpha_{ij}(k) \quad (3.72)$$

Exemple

On considère un système commutant entre deux modes de fonctionnement :

$$\begin{cases} x(k) = 0.8x(k-1) + u(k-1) + v(k-1) \\ z(k) = x(k) + w(k) \end{cases} \quad (3.73)$$

$$\begin{cases} x(k) = 0.8x(k-1) + 0.5u(k-1) + v(k-1) \\ z(k) = x(k) + w(k) \end{cases} \quad (3.74)$$

$$\Pi = \begin{bmatrix} 0.5 & 0.5 \\ 0.1 & 0.9 \end{bmatrix} \quad (3.75)$$

L'entrée u appartient à $[0, 1]$, avec des bruits $w(k)$ et $v(k)$ de variance $R = 0.1$ et $Q = 0.01$ et les commutations entre les deux modes de fonctionnement sont régies par la matrice de Markov Π . L'algorithme proposé est appliqué pour 500 simulations. La figure 3.7 représente les signaux entrée-sortie du système ainsi que les commutations et les probabilités d'activation des modèles comparées aux activations réelles. La figure 3.8 représente la convergence de la moyenne des π_{ii} en fonction du temps. Les résultats représentés sur les figures 3.7 et 3.8 montrent la faisabilité de la méthode et la possibilité d'estimer les paramètres de la matrice de Markov simultanément à l'estimation de l'état du système et au calcul des probabilités d'activation des modèles.

L'estimation de la matrice de transition de Markov est calculée pour des paramètres $\bar{\alpha}_i(0) = [20, 20]$, la valeur de $\bar{\alpha}_i(0)$ est liée à la valeur initiale de Π par l'expression

$$\hat{\pi}_j(0) = \frac{1}{\alpha_i(0)} \bar{\alpha}_i(0), \quad i = 1, 2, \dots, r.$$

et au degré de confiance qu'on accorde à cette valeur initiale ; plus on a confiance dans la valeur initiale de Π plus la valeur de $\alpha_i(0)$ est élevée. On trouve une estimation qui est proche de la valeur qui a servi à la simulation.

$$\hat{\Pi} = \begin{bmatrix} 0.4246 & 0.5744 \\ 0.0877 & 0.9113 \end{bmatrix} \quad (3.76)$$

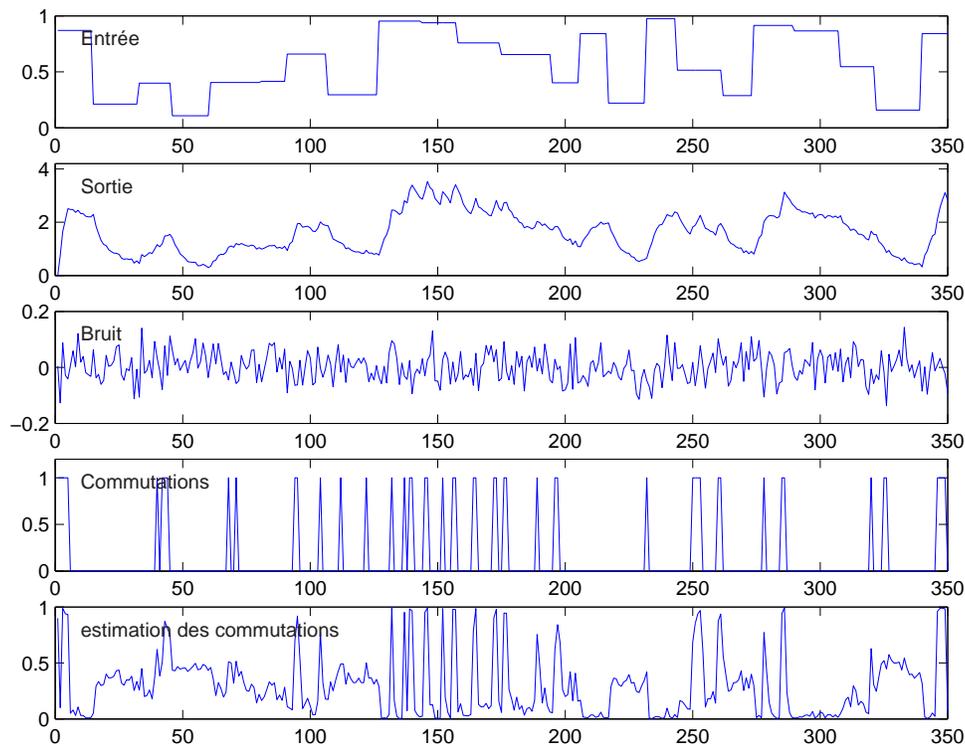


FIG. 3.7 – Signaux d'entrée-sortie et signaux de commutations

En utilisant une autre matrice de Markov pour la génération des commutations, on obtient les résultats représentés sur les figures 3.9 et 3.10 avec les valeurs suivantes :

$$\Pi = \begin{bmatrix} 0.8 & 0.2 \\ 0.3 & 0.7 \end{bmatrix} \quad (3.77)$$

$$\hat{\Pi} = \begin{bmatrix} 0.7681 & 0.2309 \\ 0.2835 & 0.7156 \end{bmatrix} \quad (3.78)$$

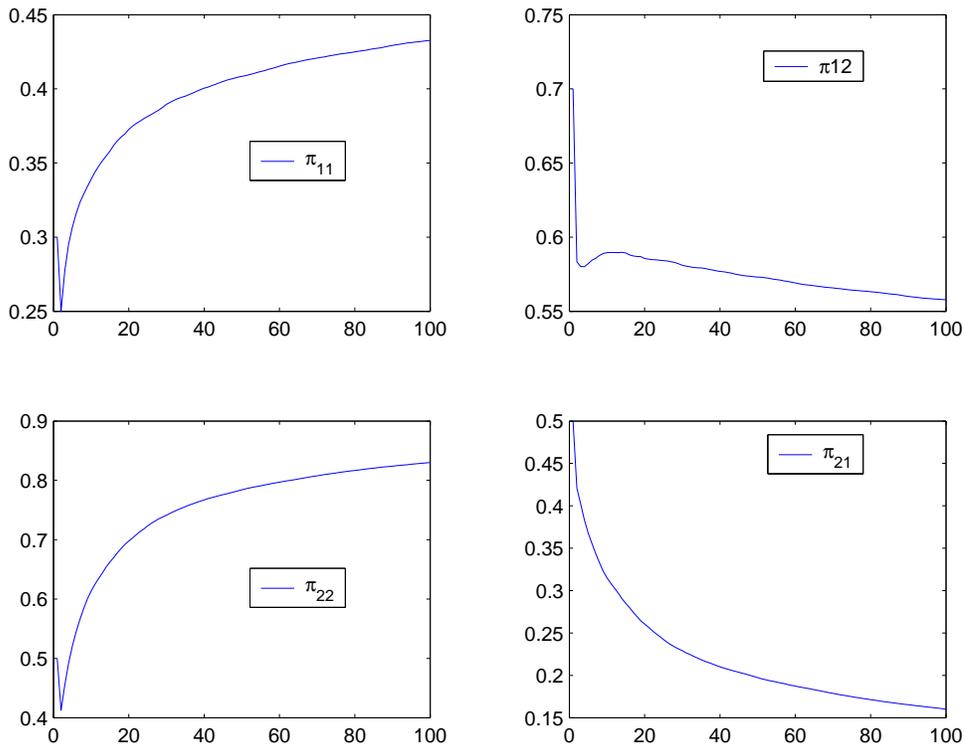


FIG. 3.8 – Convergence de π_{ii}

Les figures 3.8 et 3.10 montrent que l'on arrive à estimer correctement la matrice de Markov utilisée pour générer les commutations. Les figures 3.7 et 3.9 montrent la capacité de la méthode à détecter le modèle actif à chaque instant.

3.5 Algorithme par intégration numérique

On considère cette fois une approche plus directe pour résoudre le problème d'estimation de la MPT (Jilkov and Li, 2004). Cette approche se base sur la discrétisation du domaine possible de la MPT en un ensemble de MPT, $\Pi^{(s)}, s = 1, 2, \dots, N$ où N est le nombre de MPT et où la PDF de chaque $\Pi^{(s)}$ est calculée récursivement par la formule (3.11). Le calcul de l'estimation se fait par intégration numérique en utilisant les PDF des MPTs. On peut présenter l'algorithme de la façon suivante

Algorithme

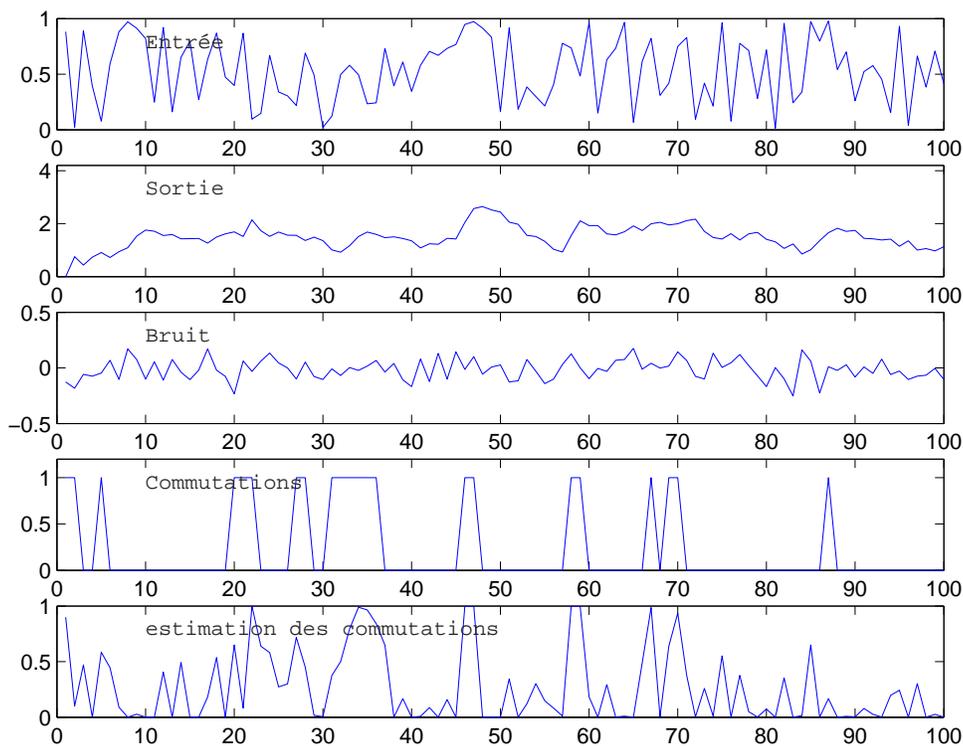


FIG. 3.9 – Signaux d'entrée-sortie et signaux de commutation

- Initialisation :

$$\hat{\Pi}(0) = \frac{1}{N} \sum_{s=1}^N \Pi^{(s)} p^{(s)}(0) \quad , \quad p^{(s)}(0) = P \{ \Pi^{(s)} \}. \quad (3.79)$$

- Répéter pour $k = 1, 2, \dots$

- Répéter pour $s = 1, 2, \dots, N$

$$p^{(s)}(k) = \frac{\mu'(k-1) \Pi^{(s)} \Lambda(k)}{\mu'(k-1) \hat{\Pi}(k-1) \Lambda(k)} p^{(s)}(k-1) \quad (3.80)$$

$$\hat{\Pi}(k) = \frac{1}{N} \sum_{s=1}^N \Pi^{(s)} p^{(s)}(k) \quad (3.81)$$

Cette méthode consiste à définir un ensemble de matrices $\Pi^{(s)}$, $s = 1, 2, \dots, N$ qui sont réparties uniformément dans l'espace des valeurs possibles de la matrice Π , $0 < \pi_{ij}^{(s)} < 1$ et $\sum_{i=1}^r \pi_{ij}^{(s)} = 1$, où r est le nombre de modes de fonctionnement et $\pi_{ij}^{(s)}$ est la $ij^{\text{ème}}$ composante de la matrice $\Pi^{(s)}$. La précision de l'estimée $\hat{\Pi}(k)$ dépend du nombre N , plus N est élevé plus la méthode est précise. Cette méthode est difficile à mettre en œuvre dans

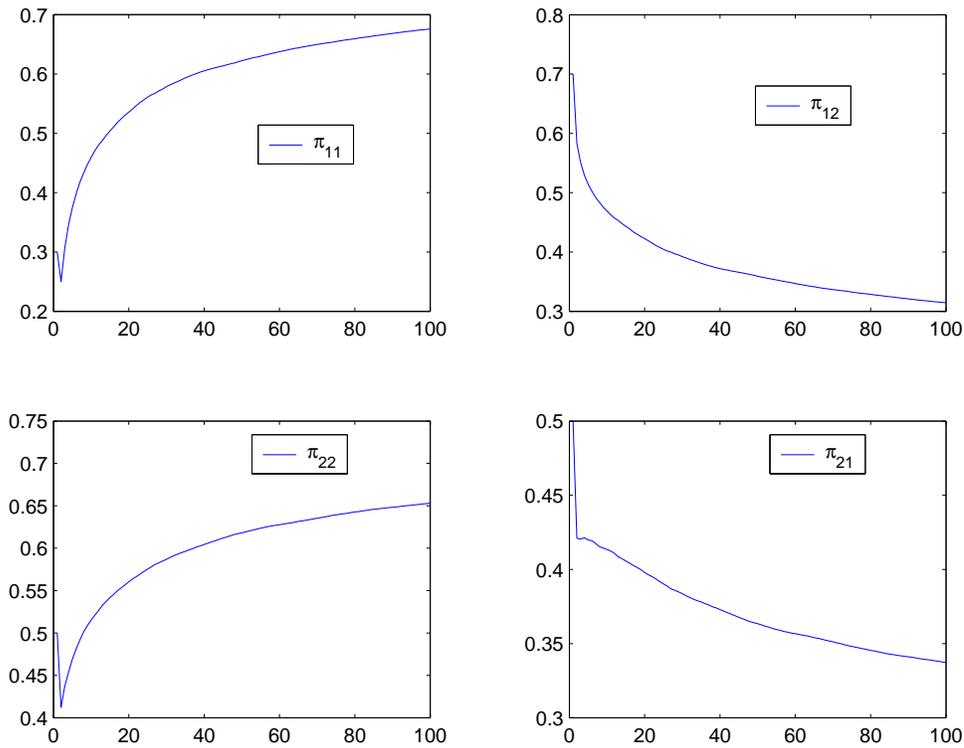


FIG. 3.10 – Convergence de π_{ij}

le cas où la matrice de Markov est de grande dimension, car on est obligé d’augmenter considérablement le nombre N afin de couvrir l’espace des valeurs possibles.

Pour le même exemple que celui utilisé précédemment avec bruit, on obtient les résultats suivants en utilisant un nombre de matrices $\Pi^{(s)}$ égal à $N = 36$, la convergence de la moyenne de l’estimation des π_{ii} est représentée sur la figure 3.12 et l’estimée de Π est donnée par :

$$\hat{\Pi} = \begin{bmatrix} 0.5015 & 0.4985 \\ 0.1 & 0.9 \end{bmatrix} \quad (3.82)$$

La figure 3.11 représente les signaux d’entrée-sortie du système ainsi que les commutations et leurs estimations.

On remarque que l’algorithme par intégration numérique converge rapidement vers la vraie valeur de Π . Cet algorithme est plus rapide et plus précis que l’algorithme quasi bayésien. Cependant, il nécessite un important volume de calcul, car plus on veut être précis, plus on doit augmenter le nombre de matrices $\Pi^{(s)}$ ce qui conduit à l’augmentation

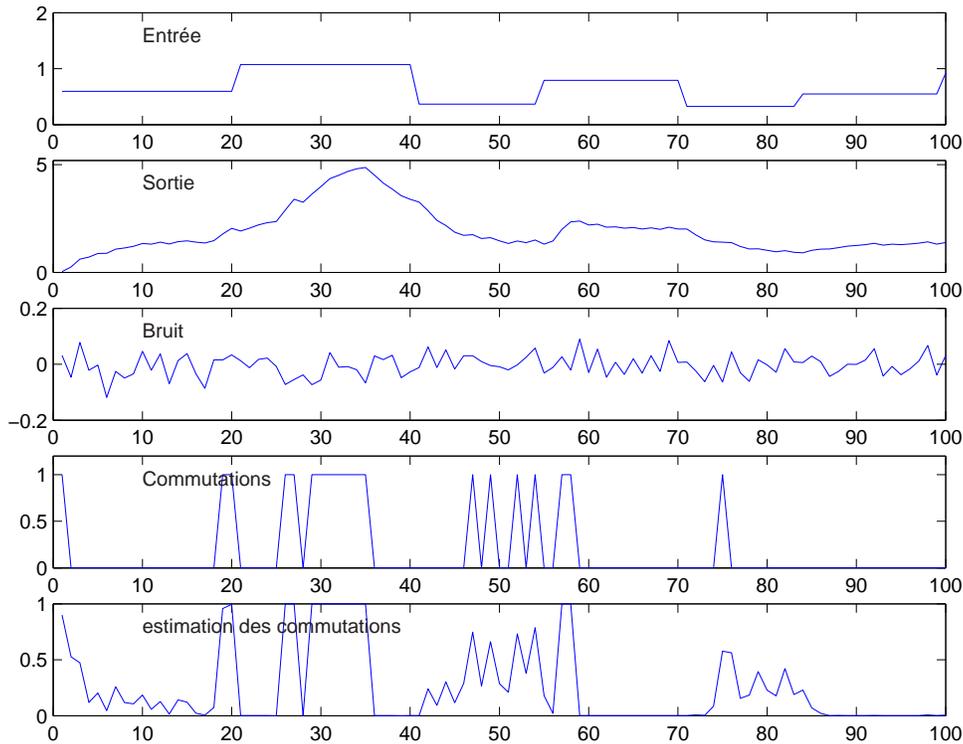


FIG. 3.11 – Signaux d'entrée-sortie et signaux de commutation

du volume de calcul. Ce volume sera d'autant plus élevé que la dimension de la matrice Π est élevée.

3.6 Conclusion

Dans ce chapitre, une forme récursive de la fonction densité de probabilité a été introduite dans le but d'estimer la matrice de probabilités de transition d'un système à commutations markoviennes. Deux algorithmes utilisant cette approximation ont été exposés.

L'estimateur quasi bayésien est simple à implémenter et nécessite un volume de calcul moindre comparé à l'estimateur à intégration numérique. L'estimateur à intégration numérique estime la matrice Π en faisant une mise à jour numérique de la vraisemblance de la matrice de Markov. Il converge plus rapidement avec un volume de calcul fixe (on utilise toujours les mêmes matrices $\Pi^{(s)}$), mais ces performances peuvent être détériorées dans

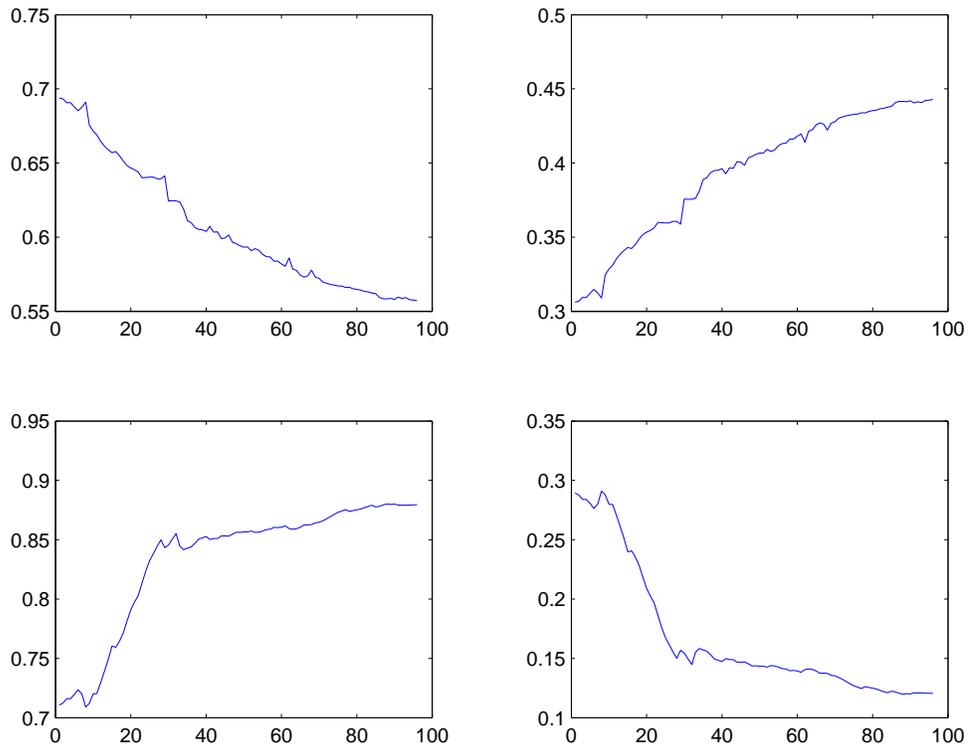


FIG. 3.12 – Convergence de π_{ij}

le cas d'un problème de grande dimension (où on aura un nombre de modes de fonctionnement r important) ce qui veut dire qu'on doit augmenter le nombre N de matrices $\Pi^{(s)}$ afin de pouvoir représenter correctement l'espace des valeurs possibles et par conséquent le volume de calcul.

4

Observateurs à mémoire finie pour les systèmes à commutation markovienne

Sommaire

4.1	Introduction	91
4.2	Observateur à mémoire finie	92
4.2.1	Condition d'existence de l'observateur	94
4.2.2	Taille de l'horizon	94
4.2.3	Propriétés de l'observateur	95
4.3	Observateur à mémoire finie avec entrée inconnue	95
4.4	Observateur à mémoire finie pour les systèmes à commutation markovienne	97
4.4.1	Développement de la méthode	99

4.4.2	Exemple d'application	101
4.5	Extension de la méthode pour les systèmes à entrée inconnue	106
4.6	Conclusion	107

4.1 Introduction

On recense, dans la littérature scientifique, différentes méthodes d'estimation d'état. Parmi celles-ci, on peut citer les méthodes s'appuyant sur tout l'historique des mesures (mémoire infinie) comme l'observateur de Luenberger et le filtre de Kalman. Dans certaines situations, ces estimateurs sont peu sensibles aux mesures récentes (Alessandri, 2000). Une autre classe d'observateurs, appelée observateurs à mémoire finie (Nuninger, 1997) ne se basent que sur les mesures les plus récentes et sont généralement moins sensibles aux bruits sur les états du système (Jazwinski, 1968), (Alessandri et al., 2003) et (Alessandri et al., 2005).

Nous proposons dans ce chapitre d'intégrer les observateurs à mémoire finie (OMF) dans le cadre de la procédure de diagnostic et d'estimation multi-modèle et nous allons voir par la suite que l'utilisation de ce type d'observateur permet d'effectuer simultanément l'estimation des entrées inconnues et la détection de défaut. Des travaux ont déjà été réalisés sur les observateurs à mémoire finie où des méthodes récursives ont été proposées par (Janyene, 1987), (Zasadzinski, 1990) et (Ragot et al., 1992). Dans (Bousghiri, 1994) il a été montré qu'un estimateur généralisé sur horizon glissant de taille finie (qui estime simultanément les entrées et sorties du système) pouvait être utile pour la détection et l'isolation de défauts de capteurs ou d'actionneurs. D'autres auteurs se sont intéressés à ce type d'observateurs comme (Medvedev and Toivonen, 1991) qui se sont intéressés à leur structure et (Kratz, 1991) qui les a utilisés pour la détection de défaut. De plus, (Darouach et al., 1994) ont montré qu'un nombre fini de mesures est suffisant à l'estimation. (Nuninger, 1997) a étudié l'intérêt de la mémoire finie pour augmenter le degré de robustesse de la procédure de diagnostic vis-à-vis des perturbations.

Dans ce chapitre, nous développons une méthode pour la détection de défauts affectant un système dynamique et l'estimation d'entrées inconnues pour des systèmes représentés par des multi-modèles (système à commutation markovienne) où chaque modèle est associé à un régime de fonctionnement particulier. Cette méthode est basée sur l'utilisation des OMF (Hocine et al., 2004b, 2005a) dans le cadre des méthodes d'estimation multi-modèle, présentées dans le chapitre 1, et a pour objectif d'identifier le mode de fonctionnement

du système et d'estimer l'entrée inconnue.

Après avoir présenté l'OMF et ses propriétés, nous adaptons ensuite cet observateur pour l'estimation d'entrées inconnues, puis nous l'intégrons dans le cadre d'une procédure d'estimation multi-modèle à la place des filtres de Kalman utilisés auparavant.

4.2 Observateur à mémoire finie

L'observateur à mémoire finie, comme son nom l'indique, utilise uniquement les mesures dans un intervalle de temps fini appelé horizon. On considère le système discret, invariant dans le temps, suivant :

$$\begin{cases} x(k) = Ax(k-1) + Bu(k-1) + Gw(k-1) \\ z(k) = Cx(k) + v(k) \end{cases} \quad (4.1)$$

où $x(k)$ est le vecteur d'état à l'instant k , A est la matrice d'évolution d'état, $u(k)$ est le vecteur de commande à l'instant k , B est la matrice d'influence de l'entrée, C est la matrice d'influence de la sortie, $w(k)$ et $v(k)$ sont respectivement le bruit d'état et le bruit de mesure, ils sont considérés comme étant gaussiens et centrés. $z(k)$ est la sortie du système à l'instant k .

En écrivant les équations d'observation sur l'horizon $[k-m, k]$ de taille $m+1$, on peut établir l'équation suivante :

$$Z_m^k = L_m x(k-m) + B_m U_m^k + G_m W_m^k + V_m^k \quad (4.2)$$

avec :

$$Y_m^k = [y(k-m)^T \quad y(k-m+1)^T \quad \dots \quad y(k)^T]^T, \quad Y \in \{Z, U, W, V\} \quad (4.3)$$

$$L_m = [C^T \quad (CA)^T \quad \dots \quad (CA^m)^T]^T \quad (4.4)$$

$$B_m = \begin{bmatrix} 0 & 0 & \dots & \dots & 0 \\ CB & 0 & \ddots & \ddots & 0 \\ CAB & CB & \ddots & \ddots & \vdots \\ \vdots & \vdots & \ddots & \ddots & 0 \\ CA^{m-1}B & CA^{m-2}B & \dots & CB & 0 \end{bmatrix} \quad (4.5)$$

$$G_m = \begin{bmatrix} 0 & 0 & \dots & \dots & 0 \\ CG & 0 & \ddots & \ddots & 0 \\ CAG & CG & \ddots & \ddots & \vdots \\ \vdots & \vdots & \ddots & \ddots & 0 \\ CA^{m-1}G & CA^{m-2}G & \dots & CG & 0 \end{bmatrix} \quad (4.6)$$

Les moyennes des bruits w et v étant nulles, l'estimée $\hat{x}(k-m)$ de l'état, à l'instant $k-m$ (premier instant de la fenêtre d'observation), peut être obtenue en minimisant le critère quadratique $J(k) = \|L_m x(k-m) + B_m U_m^k - Z_m^k\|^2$, qui permet de minimiser la somme quadratique des erreurs de mesures de la sortie, par rapport à l'inconnue $x(k-m)$. On choisit ce critère afin de maximiser la vraisemblance de l'erreur d'estimation en minimisant le critère $J(k)$. On obtient :

$$\hat{x}(k-m) = (L_m^T L_m)^{-1} L_m^T (Z_m^k - B_m U_m^k) \quad (4.7)$$

sous réserve que L_m soit de plein rang colonne.

L'estimation de l'état à l'instant terminal k de la fenêtre d'observation s'obtient en intégrant le système (4.1) sans prendre en compte l'influence des bruits :

$$\hat{x}(k) = \Omega Z_m^k - \Phi U_m^k \quad (4.8)$$

avec

$$T_m = [(A^{m-1}B)^T \quad (A^{m-2}B)^T \quad \dots \quad B^T \quad 0]^T \quad (4.9)$$

$$\Omega = A^m (L_m^T L_m)^{-1} L_m^T \quad (4.10)$$

$$\Phi = \Omega B_m - T_m \quad (4.11)$$

4.2.1 Condition d'existence de l'observateur

La condition d'existence de la solution (4.7) est que l'inverse de la matrice $(L_m^T L_m)$ qui représente le grammien d'observabilité existe. La taille minimale $m+1$ assurant l'existence de $(L_m^T L_m)^{-1}$ est donc le plus petit entier pour lequel la matrice L_m est de plein rang colonne (Murthy, 1980); m est l'indice d'observabilité du système. De façon pratique, on étudie l'évolution du rang du grammien d'observabilité quand m augmente afin de vérifier l'existence de l'inverse de la matrice $(L_m^T L_m)$. La recherche de la taille de l'horizon d'observation a été traité dans (Nuninger, 1997) où l'auteur s'appuie sur une étude de sensibilité de l'estimation vis-à-vis des différentes mesures de l'horizon considéré.

4.2.2 Taille de l'horizon

Dans son étude du choix de la taille de l'horizon, (Nuninger, 1997) a constaté que la sensibilité de l'erreur d'estimation par rapport au bruit dépendait de la taille de l'horizon par l'intermédiaire de la matrice $(L_m^T L_m)$. L'étude des propriétés statistiques des erreurs d'estimation a montré que la matrice de variance-covariance de l'erreur d'estimation dépend de $(L_m^T L_m)$.

En l'absence de défauts, il a été rappelé dans (Nuninger, 1997) que l'estimateur assure des erreurs d'estimation statistiquement nulles et dont la variance dépend de la taille de l'horizon. L'étude de la norme de la matrice de variance des erreurs d'estimation lorsque m augmente permet de trouver la valeur de m optimale au-delà de laquelle l'ajout de mesures supplémentaires ne modifie pas significativement la variance de l'erreur d'estimation.

L'horizon adéquat peut donc être choisi simplement en faisant varier m de façon systématique jusqu'à ce que les coefficients de la matrice de variance des erreurs d'estimation n'évoluent plus (les mesures supplémentaires n'apportent pas d'information). L'estimation peut donc être calculée à partir d'un nombre fini de mesures.

4.2.3 Propriétés de l'observateur

Dans cette partie, nous rappelons quelques propriétés de l'observateur à mémoire finie de l'équation (4.8).

Remarque : dans le cas sans bruit, nous avons la relation suivante :

$$\hat{x}(k) = x(k) \quad (4.12)$$

Propriété statistique de l'observateur : en présence de bruit de mesure et de bruit d'état, nous avons la relation suivante :

$$\hat{x}(k) - x(k) = \Gamma W_m^k + \Omega V_m^k \quad (4.13)$$

avec

$$\Gamma = \Omega G_m - F_m \quad (4.14)$$

et

$$F_m = [(A^{m-1}G)^T \quad (A^{m-2}G)^T \quad \dots \quad G^T \quad 0]^T \quad (4.15)$$

Comme les bruits w et v sont centrés, on en déduit

$$E[\hat{x}(k)] = E[x(k)] \quad (4.16)$$

L'expression $E[\hat{x}(k)] = E[x(k)]$ signifie que les estimations sont non biaisées.

4.3 Observateur à mémoire finie avec entrée inconnue

L'observateur à mémoire finie peut être utilisé en présence d'une entrée inconnue en considérant cette dernière comme un état du système. Le système considéré est le suivant :

$$\begin{cases} x(k) = Ax(k-1) + Bu(k-1) + Ed(k-1) + Gw(k-1) \\ z(k) = Cx(k) + v(k) \end{cases} \quad (4.17)$$

où $d(k)$ est l'entrée inconnue à l'instant k et E la matrice d'influence de l'entrée inconnue. Afin d'obtenir un système augmenté pour l'élaboration de l'observateur, on considère que

l'entrée inconnue $d(k)$ suit un processus stochastique de marche aléatoire représenté par :

$$d(k) = d(k-1) + \delta(k-1) \quad (4.18)$$

où $\delta(k)$ est un bruit aléatoire.

Le processus stochastique de marche aléatoire est un outil couramment utilisé pour l'analyse des paramètres inconnus variant avec le temps, il est utilisé en particulier dans le domaine de l'estimation et de l'identification des paramètres (Friedland, 1969), (Ignagni, 1990) et (Ljung, 1999).

On peut alors écrire un système augmenté de la manière suivante :

$$\begin{cases} x'(k) = A_a x'(k-1) + B_a u(k-1) + G_a w'(k-1) \\ z(k) = C_a x'(k) + v(k) \end{cases} \quad (4.19)$$

avec

$$x'(k) = [x(k)^T \ d(k)^T]^T, \quad w'(k) = [w(k)^T \ \delta(k)^T]^T \quad (4.20)$$

$$A_a = \begin{bmatrix} A & E \\ 0 & I \end{bmatrix}, \quad B_a = \begin{bmatrix} B \\ 0 \end{bmatrix} \quad (4.21)$$

$$C_a = [C \ 0] \quad \text{et} \quad G_a = \begin{bmatrix} G & 0 \\ 0 & I \end{bmatrix} \quad (4.22)$$

L'estimation de l'état augmenté $\hat{x}'(k-m)$ peut être effectuée comme dans le cas précédent (4.7) ; on obtient :

$$\hat{x}'(k-m) = (L_{m,a}^T L_{m,a})^{-1} L_{m,a}^T (Z_m^k - B_{m,a} U_m^k) \quad (4.23)$$

où les matrices $L_{m,a}$ et $B_{m,a}$ sont construites comme les matrices L_m et B_m des équations (4.4) et (4.5) en remplaçant les matrices A , B et C respectivement par A_a , B_a et C_a .

La condition d'existence de la solution (4.23) est que l'inverse de la matrice $(L_{m,a}^T L_{m,a})$ qui représente le grammien d'observabilité existe.

Comme précédemment, l'estimation de l'état à l'instant terminal k de la fenêtre d'observation s'obtient en intégrant le système sans prendre en compte les bruits qui ont une espérance mathématique nulle (4.1) :

$$\hat{x}'(k) = A_a^m \hat{x}'(k-m) + T_{m,a} U_m^k \quad (4.24)$$

avec

$$T_{m,a} = [(A_a^{m-1}B_a)^T \quad (A_a^{m-2}B_a)^T \quad \dots \quad B_a^T \quad 0]^T \quad (4.25)$$

Cette formulation permet donc l'obtention simultanée d'une estimation de l'état du système et de l'entrée inconnue.

Exemple :

On considère le système à entrée inconnue suivant :

$$\begin{aligned} x(k) &= \begin{bmatrix} -0.0005 & -0.0084 \\ 0.0517 & 0.8069 \end{bmatrix} x(k-1) + \begin{bmatrix} 0.1815 \\ 1.7902 \end{bmatrix} u(k-1) \\ &+ \begin{bmatrix} 0.0129 \\ -1.2504 \end{bmatrix} d(k-1) + \begin{bmatrix} 1 \\ 10 \end{bmatrix} w(k-1) \\ z(k) &= \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} x(k) + v(k) \end{aligned}$$

Les entrées sorties du système sont représentées sur la figure 4.4. Dans un premier temps, nous considérons une entrée inconnue constante égale à 0.5 dans l'intervalle temporel [100, 300] et nulle le reste du temps. Nous considérons ensuite une entrée inconnue en forme de rampe dans l'intervalle [100, 200], constante égale à 0.5 dans l'intervalle temporel [200, 300] et nulle le reste du temps.

L'examen de la figure 4.2 permet de constater une bonne estimation de l'entrée inconnue avec, cependant, un certain décalage dû à l'horizon d'observation de l'OMF choisi ici égal à 11 ($m = 10$). La figure 4.3 présente des résultats analogues dans le cas où l'entrée inconnue évolue selon une rampe.

4.4 Observateur à mémoire finie pour les systèmes à commutation markovienne

Dans ce paragraphe, nous considérons un système représenté par un ensemble de modèles $M_i, i = 1, \dots, r$, chaque modèle représentant un comportement particulier du système. L'objectif est de détecter, à chaque instant, le modèle qui approche le mieux le

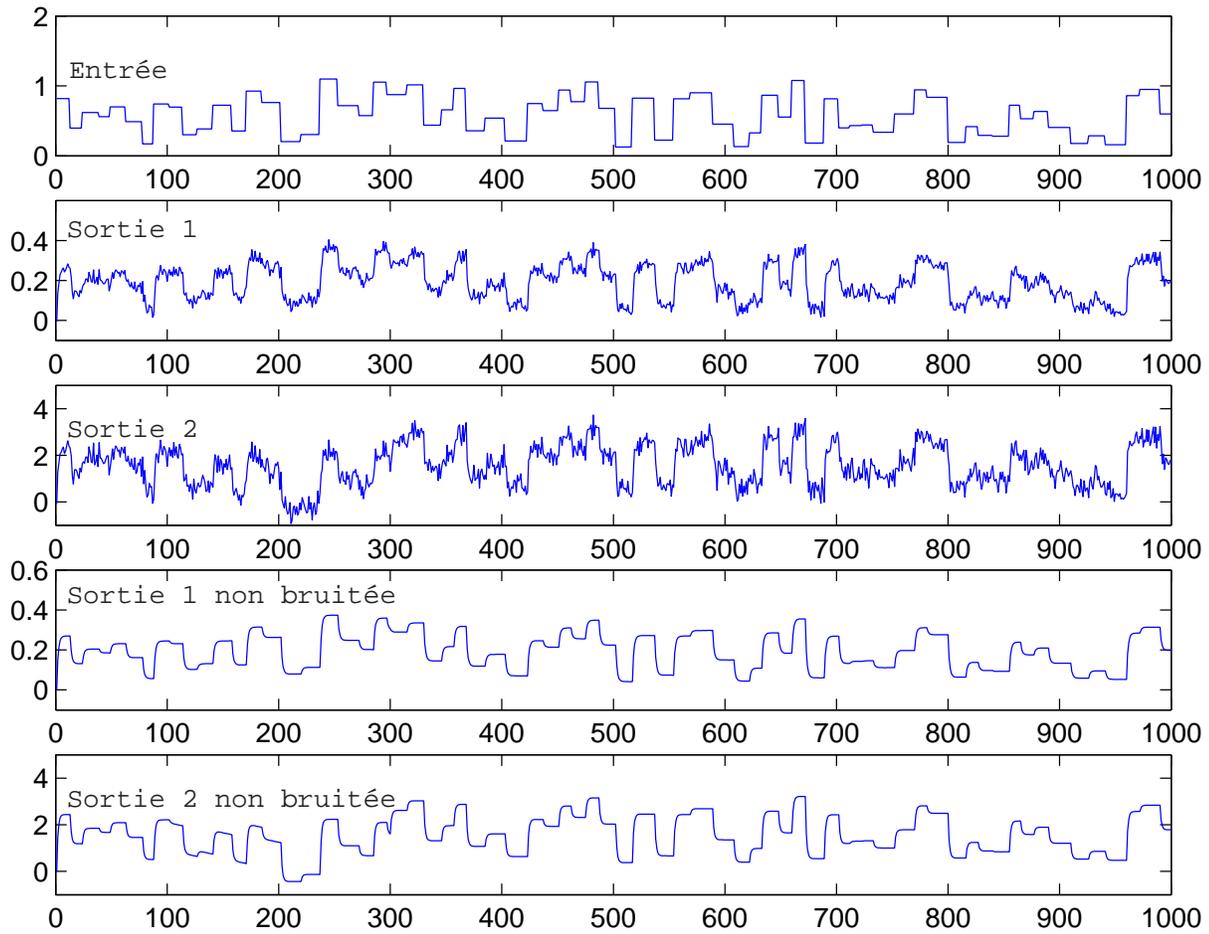


FIG. 4.1 – Entrées sorties du système

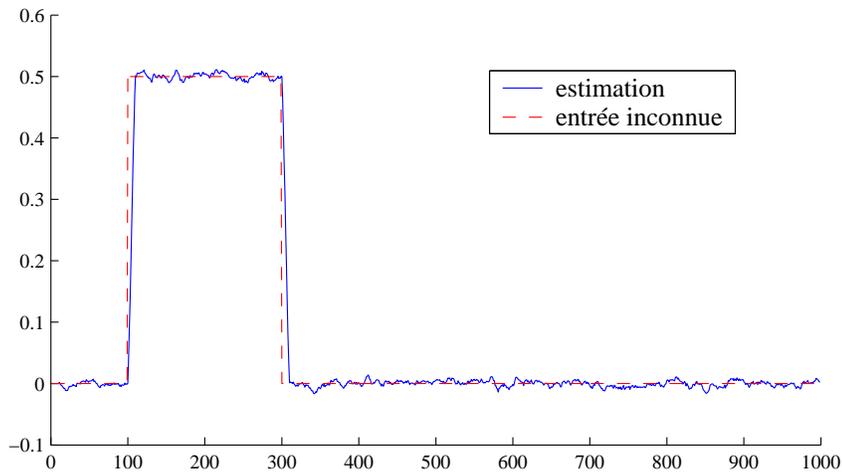


FIG. 4.2 – Entrée inconnue estimée en présence de bruit

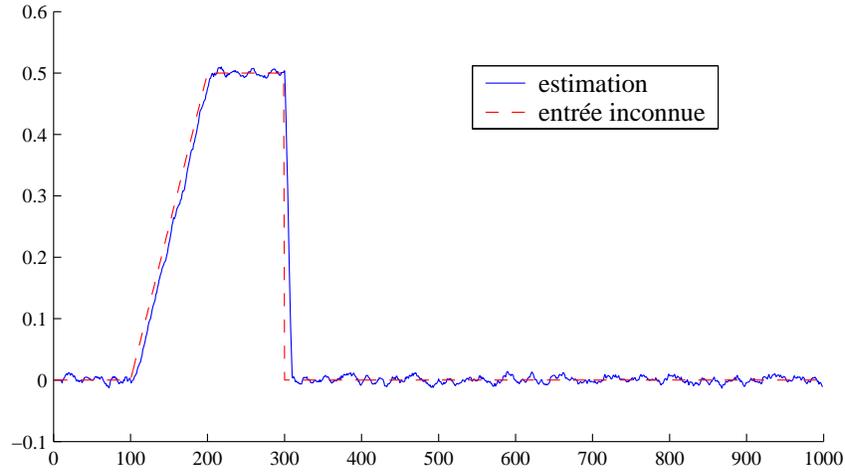


FIG. 4.3 – Entrée inconnue estimée en présence de bruit

comportement du système et, simultanément, d'estimer l'état du système. On suppose que les transitions d'un modèle à l'autre sont décrites par un processus markovien régi par la matrice de transition de Markov Π donnée par :

$$\Pi = \begin{bmatrix} \pi_{11} & \cdots & \pi_{1r} \\ \vdots & \ddots & \vdots \\ \pi_{r1} & \cdots & \pi_{rr} \end{bmatrix}$$

où π_{ij} est la probabilité de transition conditionnelle de passage du modèle M_i vers le modèle M_j ; on note $\mu_j(k)$ la probabilité que le $j^{\text{ème}}$ modèle soit actif à l'instant k .

4.4.1 Développement de la méthode

Considérons le $j^{\text{ème}}$ modèle :

$$\begin{cases} x(k) = A_j x(k-1) + B_j u(k-1) + G_j w(k-1) \\ z(k) = C_j x(k) + v(k) \end{cases} \quad (4.26)$$

L'estimation d'état de ce modèle peut être effectuée à l'aide d'un OMF selon la méthode décrite à la section 4.2 sous l'hypothèse de l'existence de $(L_{j,m}^T L_{j,m})^{-1}$. On obtient ainsi :

$$\hat{x}_j(k-m) = (L_{j,m}^T L_{j,m})^{-1} L_{j,m}^T (Z_m^k - B_{j,m} U_m^k) \quad (4.27)$$

et

$$\hat{x}_j(k) = A_j^m \hat{x}_j(k-m) + T_{j,m} U_m^k \quad (4.28)$$

Les matrices $L_{j,m}$, $B_{j,m}$ et $T_{j,m}$ sont construites en utilisant les définitions (4.4), (4.5) et (4.9) en remplaçant les matrices A , B et C par les matrices A_j , B_j et C_j décrivant le $j^{\text{ème}}$ modèle.

L'estimée de (4.28) est garantie non biaisée en l'absence de commutation sur $[k-m, k]$. L'estimation de l'état $\hat{x}(k)$ du système à commutation est alors calculée comme une somme pondérée des états des différents modèles :

$$\hat{x}(k) = \sum_{j=1}^r \hat{x}_j(k) \mu_j(k) \quad (4.29)$$

En nous inspirant des travaux de (Bar-Shalom, 1990), la probabilité que le système fonctionne selon le $j^{\text{ème}}$ modèle à l'instant k est calculée de la manière suivante :

$$\mu_j(k) = P\{M_j(k) | Z_m^k\} \quad (4.30)$$

Définissons alors \tilde{Z}_m^{k-1} , le vecteur des observations effectuées sur l'horizon $[k-m, k-1]$; on a :

$$Z_m^k = \left[(\tilde{Z}_m^{k-1})^T \quad z(k)^T \right]^T \quad (4.31)$$

L'équation (4.30) peut alors s'écrire :

$$\mu_j(k) = P\{M_j(k) | \tilde{Z}_m^{k-1}, z(k)\} \quad (4.32)$$

puis, en utilisant la formule de Bayes :

$$\mu_j(k) = \frac{p \left[z(k) | M_j(k), \tilde{Z}_m^{k-1} \right] P\{M_j(k) | \tilde{Z}_m^{k-1}\}}{\sum_{l=1}^r p \left[z(k) | M_l(k), \tilde{Z}_m^{k-1} \right] P\{M_l(k) | \tilde{Z}_m^{k-1}\}} \quad (4.33)$$

Afin d'alléger les notations, posons :

$$\lambda_i(k) = p \left[z(k) | M_i(k), \tilde{Z}_m^{k-1} \right] \quad (4.34)$$

Développons également la probabilité d'activation du modèle j à l'instant k , conditionnellement au modèle actif à l'instant $k-1$:

$$P\{M_j(k) | \tilde{Z}_m^{k-1}\} = \sum_{i=1}^r P\{M_j(k) | M_i(k-1), \tilde{z}(k-1)\} P\{M_i(k-1) | \tilde{Z}_m^{k-1}\} \quad (4.35)$$

De manière à élaborer une récurrence sur le calcul des $\mu_j(k)$, on effectue l'approximation suivante :

$$P\{M_i(k-1)|\tilde{Z}_m^{k-1}\} \approx P\{M_i(k-1)|Z_m^{k-1}\} = \mu_j(k-1) \quad (4.36)$$

Cela revient à considérer que l'information apportée par le premier vecteur d'observation $z(k-m-1)$ du vecteur Z_m^{k-1} défini sur l'horizon $[k-m-1, k-1]$ n'est pas très importante et peut être négligée (cela dépend évidemment de l'horizon choisi). Dans ce cas, en considérant les équations (4.33) à (4.36) et en remarquant que, par définition, $P\{M_j(k)|M_i(k-1), \tilde{Z}^k\} = \pi_{ij}$, on obtient alors la récurrence suivante sur la probabilité que le système opère selon le modèle j à l'instant k :

$$\mu_j(k) = \frac{\lambda_j(k) \sum_{i=1}^r \pi_{ij} \mu_j(k-1)}{\sum_{l=1}^r \lambda_l(k) \sum_{i=1}^r \pi_{il} \mu_j(k-1)} \quad (4.37)$$

Cette expression permet de calculer la probabilité d'activation de chaque modèle connaissant les vraisemblances $\lambda_j(k)$ et les probabilités de passage d'un modèle à un autre π_{ij} . A partir de ces probabilités, on peut calculer l'estimée $\hat{x}(k)$ par l'équation (4.29) qui elle-même utilise les estimées $\hat{x}_j(k)$ issues des observateurs (4.28) relatifs aux modèles de fonctionnement.

4.4.2 Exemple d'application

On considère un modèle de fonctionnement normal (A_1, B_1, C_1) , un modèle de défaut d'actionneur (A_2, B_2, C_2) et un modèle de défaut de capteur (A_3, B_3, C_3) , les différentes matrices étant définies par :

$$A_i = \begin{bmatrix} -0.0005 & -0.0084 \\ 0.0517 & 0.8069 \end{bmatrix}, \quad i = 1 \dots 3$$

$$B_1 = [0.1815 \ 1.7902]^T, \quad C_1 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}.$$

$$B_2 = [0.1815 \ 0]^T, \quad C_2 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}.$$

$$B_3 = [0.1815 \ 1.7902]^T, \quad C_3 = \begin{bmatrix} 1 & 0 \\ 0 & 0.1 \end{bmatrix}.$$

Pour tester la méthode, dans un premier temps, le scénario suivant a été établi : initialement le système est en bon fonctionnement (modèle 1), puis à l'instant 100, survient un défaut sur l'actionneur (modèle 2), à l'instant 500, le système revient au mode de bon fonctionnement et, à l'instant 800, un défaut sur un capteur est introduit (modèle 3), les entrées-sorties du système sont représentées sur la figure 4.4.

Les résultats sont présentés aux figures (4.5, 4.6 et 4.7) où sont représentées les probabilités d'activation des différents modèles calculées par la méthode proposée et par la méthode GPB1. On constate clairement les changements de régime ; la probabilité d'activation des modèles, dans leur zones respectives de fonctionnement, fluctue autour de un et donc la détection des défauts est réalisée. On remarque aussi que, sur cet exemple, la méthode proposée donne de meilleurs résultats que la méthode GPB1.

Un deuxième scénario, similaire à celui exposé au deuxième chapitre, est envisagé pour comparer les performances de la méthode proposée avec celle de l'algorithme GPB. Les commutations sont générées en utilisant les probabilités de passage d'un modèle à un autre de la matrice de Markov Π connue donnée par :

$$\Pi = \begin{bmatrix} 0.95 & 0.025 & 0.025 \\ 0.05 & 0.95 & 0 \\ 0.05 & 0 & 0.95 \end{bmatrix}$$

Les variances des bruits sur l'état et sur les mesures sont données par :

$$Q = 0.001$$

$$R = 5 \times \begin{bmatrix} 0.01 & 0 \\ 0 & 0.4 \end{bmatrix}.$$

Les résultats de cette comparaison, qui utilise les critères de performance définis dans le chapitre 2, sont rassemblés dans le tableau 4.1 avec (FA) le taux de fausses alarmes, (ND) le taux de non détection, (BDI) le taux de bonne détection et d'isolation, (IID) le taux d'isolation incorrect du défaut et (TMD) le temps moyen de détection (chapitre 2).

On remarque que les résultats de la méthode proposée sont meilleurs que ceux de la

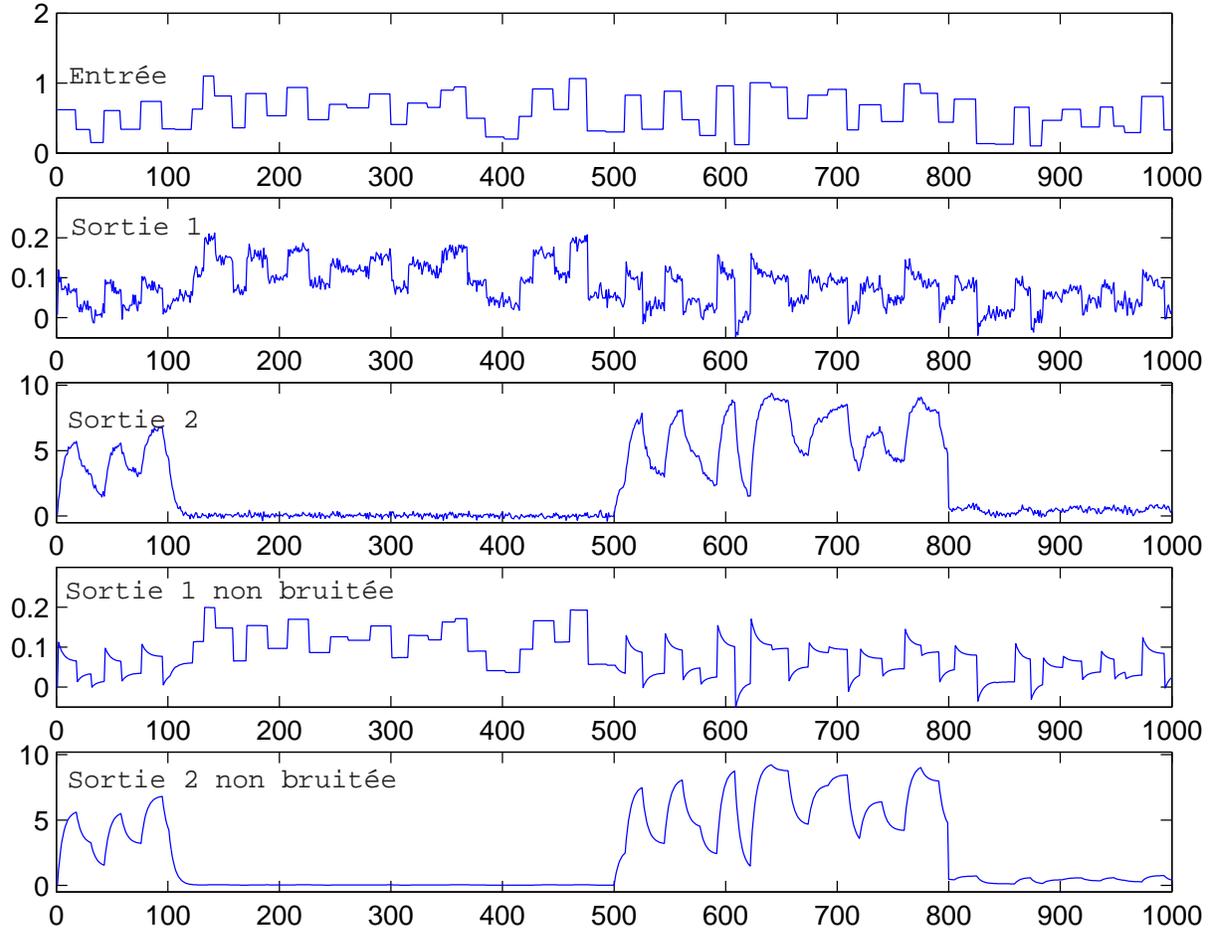


FIG. 4.4 – Entrées sorties du système

	BDI	FA	ND	IID	TMD
Méthode proposée	0.7865	0.00085	0.2774	0.0069	2.4623
Algorithme GPB	0.7100	0.00093	0.2885	0.0978	9.2102

TAB. 4.1 – Performance de la détection pour la méthode proposée et l'algorithme GPB

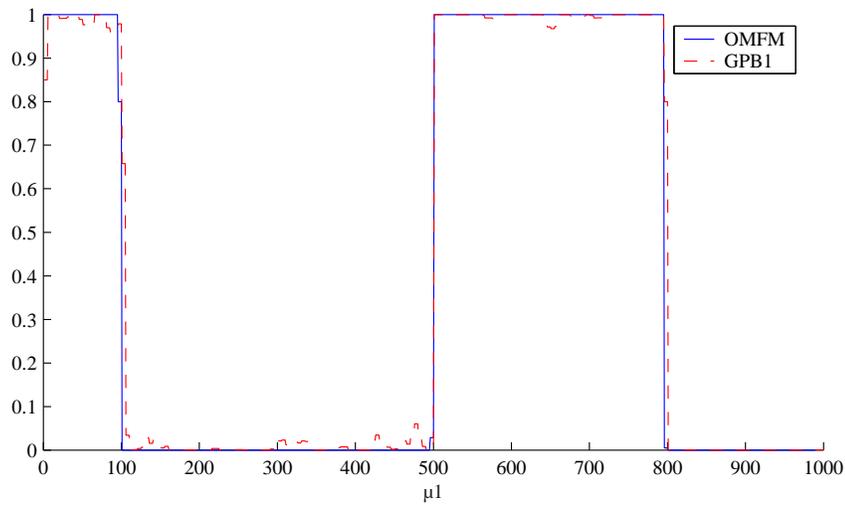


FIG. 4.5 – Probabilité d'activation du modèle 1

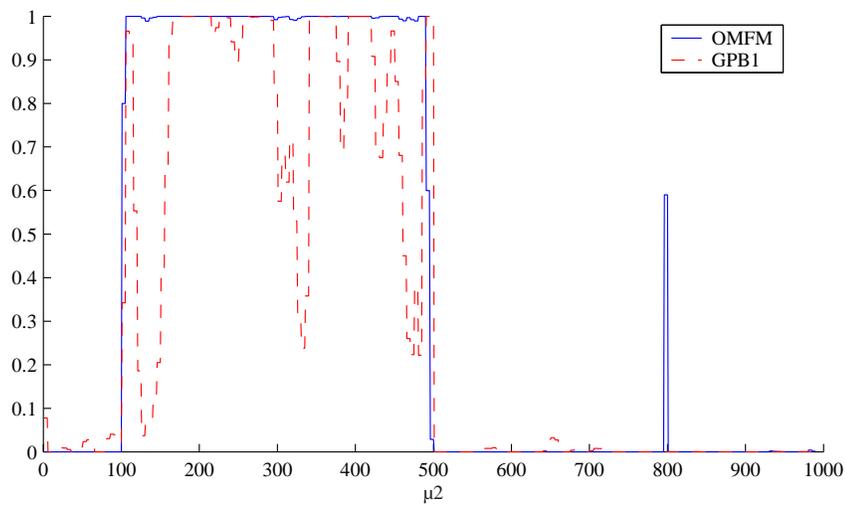


FIG. 4.6 – Probabilité d'activation du modèle 2

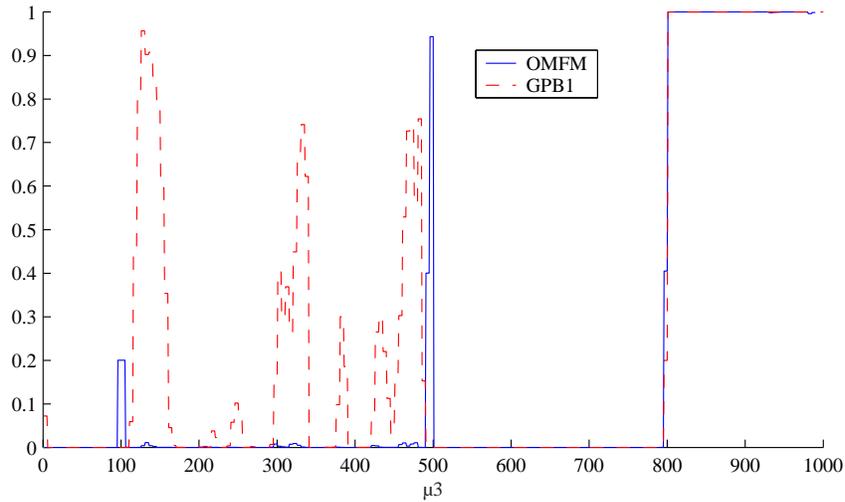


FIG. 4.7 – Probabilité d'activation du modèle 3

	BDI	FA	ND	IID	TMD
$R \times 1$	0.7865	0.00085	0.2774	0.0069	2.4623
$R \times 3$	0.7851	0.00085	0.2785	0.0056	2.4906
$R \times 5$	0.7813	0.0013	0.2763	0.0111	2.4623
$R \times 8$	0.7536	0.0024	0.3089	0.0288	2.4151
$R \times 10$	0.7393	0.0039	0.3164	0.0395	2.4623

TAB. 4.2 – Performance de la détection en fonction du niveau du bruit sur la sortie

méthode GPB1. Cela est dû à la sensibilité de cette dernière au bruit. On conclut donc que l'utilisation d'un observateur à mémoire finie, pour la détection de défaut dans le cadre de systèmes à commutations, donne de bons résultats qui sont moins sensibles au bruit que la méthode GPB classique. Le tableau 4.2 nous permet de constater que les performances de la détection se dégradent avec l'augmentation du niveau du bruit sur la sortie. On a remarqué aussi que, pour cet exemple, la taille de la fenêtre de l'observateur n'a pas une grande influence sur les performances du diagnostic.

4.5 Extension de la méthode pour les systèmes à entrée inconnue

Nous allons appliquer la méthode proposée à la section 4.3 au cas des systèmes à entrée inconnue. Pour cela, l'observateur à mémoire finie est remplacé par un observateur à mémoire finie avec entrée inconnue.

Le modèle M_j s'écrit sous la forme :

$$\begin{cases} x(k) = A_j x(k-1) + B_j u(k-1) + E d(k-1) + G w(k-1) \\ z(k) = C_j x(k) + v(k) \end{cases}$$

où $d(k)$ est l'entrée inconnue à l'instant k et où E est la matrice d'influence de l'entrée inconnue.

En utilisant un modèle augmenté qui prend en compte l'entrée inconnue, qui est considérée comme un état du système, pour chaque modèle M_j , on peut estimer l'état et l'entrée inconnue d . L'estimation de l'état global est ensuite obtenue en suivant la démarche exposée auparavant dans ce chapitre.

Exemple : on considère les mêmes modèles de fonctionnement normal, de défaut d'actionneur et de défaut de capteur que précédemment, auxquels l'entrée inconnue $d(k)$ est ajoutée avec :

$$E = \begin{bmatrix} 0.0129 & -1.2504 \end{bmatrix}^T$$

Pour tester la méthode, le scénario suivant a été établi : initialement le système est en bon fonctionnement, à l'instant 100 intervient une entrée inconnue d'amplitude constante, puis à l'instant 200, survient un défaut sur l'actionneur, à l'instant 300, l'entrée inconnue devient nulle, à l'instant 500, le système revient au mode de bon fonctionnement et, à l'instant 800, un défaut capteur est introduit.

Les résultats des figures 4.8 à 4.10, en présence de bruit, montrent clairement les changements d'un régime à l'autre ce qui permet donc la détection des défauts. La figure 4.11 montre l'estimation de l'entrée inconnue en l'absence de bruit. En présence de bruit, cette estimation est représentée à la figure 4.12. On peut constater que l'utilisation d'un observateur à mémoire finie avec entrée inconnue pour la détection de défaut et l'estimation

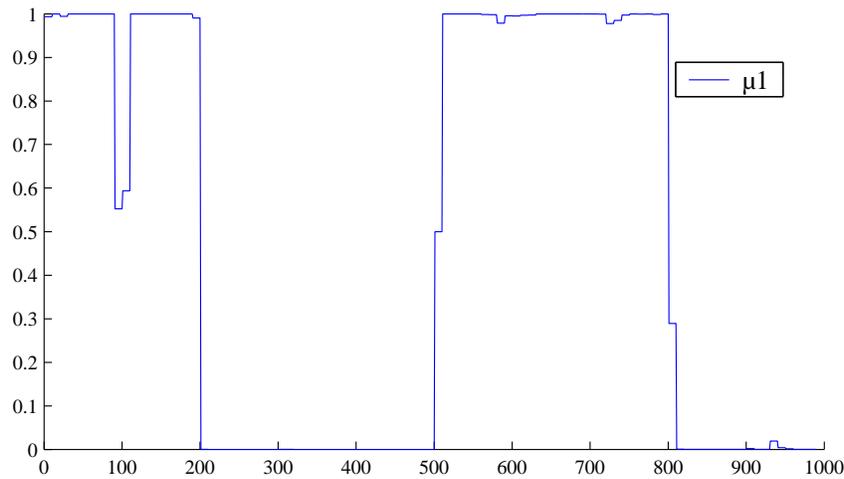


FIG. 4.8 – Probabilité d’activation du modèle 1 avec bruit

de l’entrée inconnue, dans le cadre des systèmes à commutations, donne de bons résultats malgré la présence de bruit.

4.6 Conclusion

Dans ce chapitre, nous avons tout d’abord rappelé la structure d’un observateur à mémoire finie. Nous l’avons ensuite appliqué avec succès à l’estimation d’entrées inconnues. Ce type d’observateur a été utilisé dans le cadre d’un système à commutation markovienne pour lequel on doit également détecter les instants de commutation entre modèles. La comparaison des résultats obtenus avec ceux de la méthode GPB1 a ensuite été effectuée sur un exemple. L’utilisation d’un observateur à mémoire finie avec la méthode GPB1 donne de meilleurs résultats, principalement en présence de bruits.

Finalement, la méthode proposée a été étendue au cas de systèmes soumis à des entrées inconnues. Dans cette situation, la détection des instants de commutation s’effectue simultanément à l’estimation de l’entrée inconnue.

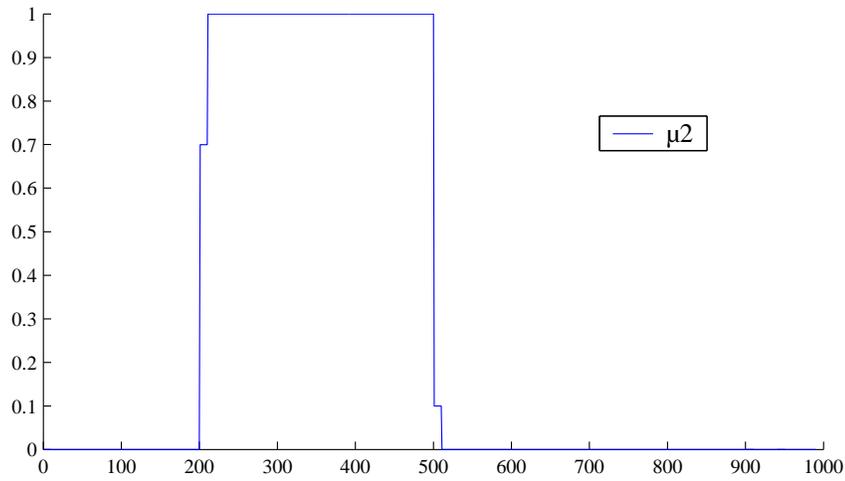


FIG. 4.9 – Probabilité d’activation du modèle 2 avec bruit

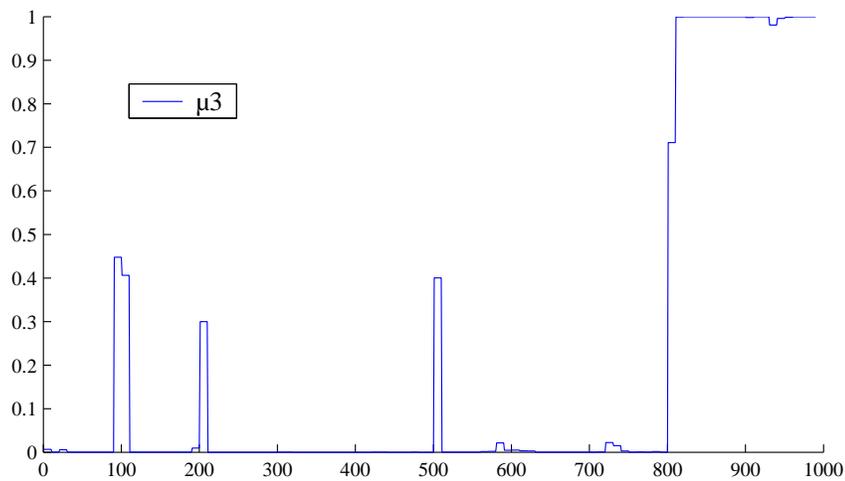


FIG. 4.10 – Probabilité d’activation du modèle 3 avec bruit

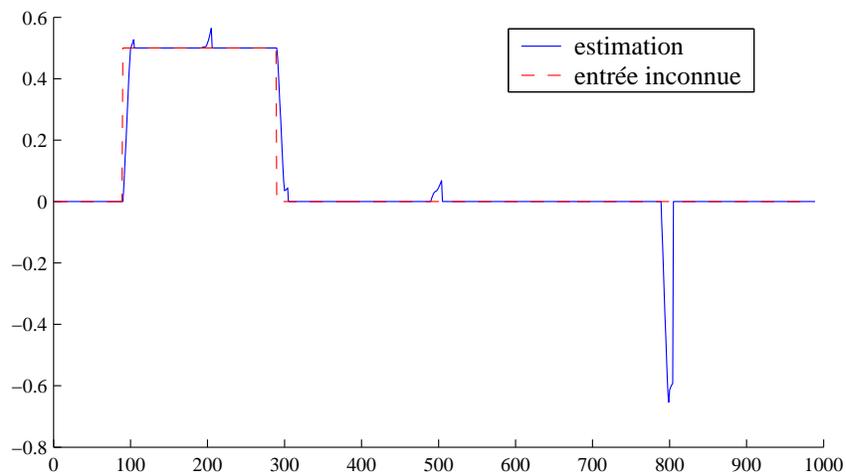


FIG. 4.11 – Estimation de l'entrée inconnue en l'absence de bruit

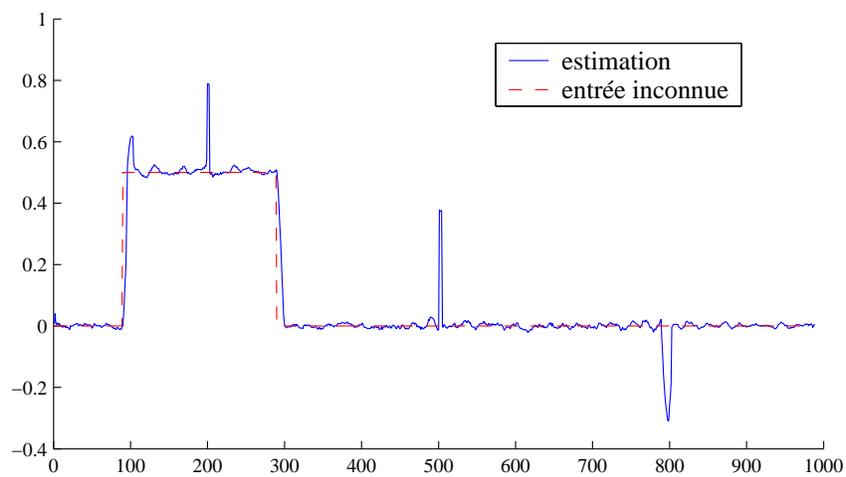


FIG. 4.12 – Estimation de l'entrée inconnue en présence de bruit

5

Observateur mixte pour les systèmes à commutation markovienne

Sommaire

5.1	Introduction	113
5.2	Observateur à mémoire finie	113
5.3	Observateur mixte	115
5.4	Exemple	117
5.5	Observateur mixte pour les systèmes à commutation markovienne	121
5.5.1	Développement de la méthode	121
5.5.2	Exemple d'application	123
5.6	Conclusion	126

5.1 Introduction

Dans le chapitre précédent, nous avons constaté que l'utilisation des observateurs à mémoire finie améliore les performances de la méthode de diagnostic proposée dans le chapitre 2. Cependant, pour améliorer les performances de l'observation et du diagnostic, on propose de combiner l'observateur de Luenberger et l'observateur à mémoire finie à travers un paramètre de pondération τ . Dans un premier temps, on présente la structure de l'observateur à mémoire finie et l'expression des états du système issue de cette formulation. Ensuite, on propose la structure d'un nouvel observateur et on établit ses conditions de stabilité ; celles-ci sont formulées en terme d'inégalités matricielles par l'utilisation d'une fonction de Lyapunov quadratique sous l'hypothèse que les bruits sont bornés. On présente un exemple académique où une comparaison est effectuée avec l'observateur à mémoire finie classique et avec l'observateur de Luenberger.

L'observateur proposé est ensuite utilisé dans le cadre des systèmes à commutation markovienne (Hocine et al., 2006). Les performances obtenues sont comparées à celles des estimateurs multi-modèle classiques et avec l'estimateur multi-modèle proposé dans le chapitre 4.

5.2 Observateur à mémoire finie

Rappelons quelques résultats sur les observateurs à mémoire finie donnés dans le chapitre précédent. On considère le système discret, invariant dans le temps, suivant :

$$\begin{cases} x(k) = Ax(k-1) + Bu(k-1) + Gw(k-1) \\ z(k) = Cx(k) + v(k) \end{cases}$$

où, à l'instant k , $x(k)$ est le vecteur d'état, A est la matrice d'évolution d'état, $u(k)$ est le vecteur de commande, B est la matrice d'influence de l'entrée, C est la matrice d'influence de la sortie, $w(k)$ et $v(k)$ sont respectivement le bruit d'état et le bruit de mesure et $z(k)$ est la sortie du système. On considère l'observateur à mémoire finie proposé dans le chapitre précédent :

$$\hat{x}(k) = \Omega Z_m^k - \Phi U_m^k$$

avec

$$\begin{aligned}\Omega &= A^m(L_m^T L_m)^{-1} \\ \Phi &= \Omega B_m - T_m \\ T_m &= [(A^{m-1}B)^T \quad (A^{m-2}B)^T \quad \dots \quad B^T \quad 0]^T \\ L_m &= [C^T \quad (CA)^T \quad \dots \quad (CA^m)^T]^T\end{aligned}$$

$$B_m = \begin{bmatrix} 0 & 0 & \dots & \dots & 0 \\ CB & 0 & \ddots & \ddots & 0 \\ CAB & CB & \ddots & \ddots & \vdots \\ \vdots & \vdots & \ddots & \ddots & 0 \\ CA^{m-1}B & CA^{m-2}B & \dots & CB & 0 \end{bmatrix}$$

$$G_m = \begin{bmatrix} 0 & 0 & \dots & \dots & 0 \\ CG & 0 & \ddots & \ddots & 0 \\ CAG & CG & \ddots & \ddots & \vdots \\ \vdots & \vdots & \ddots & \ddots & 0 \\ CA^{m-1}G & CA^{m-2}G & \dots & CG & 0 \end{bmatrix}$$

et

$$Z_m^k = L_m x(k-m) + B_m U_m^k + G_m W_m^k + V_m^k$$

En présence de bruit de mesure et de bruit d'état, l'erreur d'estimation d'état s'écrit :

$$\hat{x}(k) - x(k) = \Gamma W_m^k + \Omega V_m^k$$

avec

$$\Gamma = \Omega G_m - F_m$$

et

$$F_m = [(A^{m-1}G)^T \quad (A^{m-2}G)^T \quad \dots \quad G^T \quad 0]^T$$

L'état vrai $x(k)$ peut être alors écrit sous la forme :

$$x(k) = \Omega Z_m^k - \Phi U_m^k - \Gamma W_m^k - \Omega V_m^k \quad (5.1)$$

5.3 Observateur mixte

En combinant les prédictions fournies par les équations (4.1) et (5.1), qui sont relatives au même état, on obtient un état "composite" :

$$\begin{aligned} x(k+1) = & \tau(\Omega Z_m^{k+1} - \Phi U_m^{k+1} - \Gamma W_m^{k+1} - \Omega V_m^{k+1}) \\ & + (1 - \tau)(Ax(k) + Bu(k) + Gw(k)) \end{aligned} \quad (5.2)$$

où τ est un paramètre de pondération entre l'état prédit à un pas et l'état obtenu sur un horizon de taille $m + 1$. Nous examinerons par la suite la façon d'optimiser la valeur de ce paramètre.

Pour le système (5.2), on considère l'observateur suivant :

$$\begin{cases} \hat{x}(k+1) = \tau(\Omega Z_m^{k+1} - \Phi U_m^{k+1}) \\ \quad + (1 - \tau)(A\hat{x}(k) + Bu(k) - K(\hat{z}(k) - z(k))) \\ \hat{z}(k) = C\hat{x}(k) \end{cases} \quad (5.3)$$

où K et τ sont à déterminer.

Pour caractériser l'état fourni par cet observateur, on forme l'équation d'évolution de la dynamique de l'erreur d'estimation $e(k) = \hat{x}(k) - x(k)$:

$$e(k+1) = \bar{A}e(k) + (1 - \tau)Kv(k) + \tau\Omega V_m^{k+1} + \tau\Gamma W_m^{k+1} + (1 - \tau)Gw(k) \quad (5.4)$$

avec

$$\bar{A} = (1 - \tau)(A - KC)$$

L'erreur d'estimation peut être réécrite en concaténant les bruits dans un seul vecteur :

$$e(k+1) = \bar{A}e(k) + DS_m^{k+1} \quad (5.5)$$

avec

$$\begin{aligned} S_m^{k+1} &= \begin{bmatrix} v(k)^T & V_m^{k+1T} & w(k)^T & W_m^{k+1T} \end{bmatrix}^T \\ D &= [(1 - \tau)K \quad \tau\Omega \quad (1 - \tau)G \quad \tau\Gamma] \end{aligned}$$

La conception de l'observateur nécessite la détermination de K et τ .

Proposition : Supposons qu'il existe une matrice de gain K , deux matrices définies

positives $P > 0$ et $Q > 0$, P symétrique, et un scalaire positif τ , vérifiant la contrainte suivante :

$$\begin{aligned} Q - P + (1 - \tau)^2 (A - KC)^T P (A - KC) \\ + (1 - \tau)^2 (A - KC)^T P^2 (A - KC) < 0 \end{aligned} \quad (5.6)$$

alors l'observateur défini en (5.3) a une erreur d'estimation bornée, i.e. il existe une constante positive

$$r^2 = \|D\|^2 \|S_m^{k+1}\|^2 \lambda_{\max}(P + I) / \lambda_{\min}(Q) \quad (5.7)$$

telle que $\Delta V_k(e(k)) < 0$ pour $\|e(k)\| > r$.

Preuve : En considérant la fonction de Lyapunov $V_k(e(k)) = e(k)^T P e(k)$, on obtient :

$$\begin{aligned} V_{k+1}(e(k)) &= e(k+1)^T P e(k+1) \\ &= e(k)^T \bar{A}^T P \bar{A} e(k) + e(k)^T \bar{A}^T P D S_m^{k+1} + S_m^{k+1 T} D^T P \bar{A} e(k) + S_m^{k+1 T} D^T P D S_m^{k+1} \end{aligned} \quad (5.8)$$

La propriété $M^T N + N^T M \leq M^T M + N^T N$ appliquée à des matrices de dimensions appropriées permet d'écrire :

$$e(k)^T \bar{A}^T P D S_m^{k+1} + S_m^{k+1 T} D^T P \bar{A} e(k) \leq e(k)^T \bar{A}^T P P \bar{A} e(k) + S_m^{k+1 T} D^T D S_m^{k+1} \quad (5.9)$$

En substituant (5.9) dans (5.8), on obtient :

$$V_{k+1}(e(k)) \leq e(k)^T \left(\bar{A}^T P \bar{A} + \bar{A}^T P^2 \bar{A} \right) e(k) + S_m^{k+1 T} D^T D S_m^{k+1} + S_m^{k+1 T} D^T P D S_m^{k+1} \quad (5.10)$$

$$V_{k+1}(e(k)) \leq e(k)^T \left(\bar{A}^T P \bar{A} + \bar{A}^T P^2 \bar{A} \right) e_k + S_m^{k+1 T} D^T (P + I) D S_m^{k+1} \quad (5.11)$$

D'où

$$V_{k+1}(e(k)) \leq e(k)^T \left(\bar{A}^T P \bar{A} + \bar{A}^T P^2 \bar{A} \right) e(k) + \|D\|^2 \|S_m^{k+1}\|^2 \lambda_{\max}(P + I) \quad (5.12)$$

Alors s'il existe $P > 0$, $Q > 0$ tel que $\bar{A}^T P \bar{A} + \bar{A}^T P^2 \bar{A} - P < -Q$ (5.6), on a :

$$\begin{aligned} \Delta V_k(e(k)) &\leq -e(k)^T Q e(k) + \|D\|^2 \|S_m^{k+1}\|^2 \lambda_{\max}(P + I) \\ &\leq -\lambda_{\min}(Q) \|e(k)\|^2 + \|D\|^2 \|S_m^{k+1}\|^2 \lambda_{\max}(P + I) \end{aligned} \quad (5.13)$$

et $\Delta V_k(e(k)) < 0$ si $\|e(k)\| \geq r$. Ce qui clos la preuve. Cela revient à dire que la décroissance de l'erreur d'estimation est garantie tant qu'elle est supérieure à r .

Pour résoudre (5.6) on procède par linéarisation (Boyd et al., 1994) en prenant $X = PK$:

$$\begin{bmatrix} (1 - \tau)^{-2}(P - Q) & (PA - XC)^T & (PA - XC)^T \\ PA - XC & P & 0 \\ PA - XC & 0 & I \end{bmatrix} > 0 \quad (5.14)$$

Il s'agit d'une inégalité matricielle linéaire en P , Q et X résolue efficacement par des outils numériques (toolbox LMItool de `Matlab`). Le gain de l'observateur est donné par $K = P^{-1}X$.

Comme on l'a vu dans le paragraphe précédent, la convergence de l'erreur d'estimation est garantie en dehors de l'intervalle $[-r, r]$ (équation (5.7)). Il est évident que si l'on veut améliorer les performances de l'observateur, il faut minimiser la valeur de r . Pour cela, les bornes des bruits étant constantes, on peut minimiser le terme $\|D\|\lambda_{\max}(P+I)/\sqrt{\lambda_{\min}(Q)}$ qui dépend explicitement du paramètre τ et implicitement de τ à travers le gain K et les matrices P et Q . Les matrices K , P et Q sont obtenues en résolvant l'inégalité matricielle de l'équation (5.14) qui dépend directement de τ . On est donc confronté à un problème de minimisation non linéaire. Une des stratégies envisageables est de minimiser numériquement le terme $\|D\|\lambda_{\max}(P+I)/\sqrt{\lambda_{\min}(Q)}$ par rapport au paramètre τ .

5.4 Exemple

Dans cette partie, un exemple académique illustre la mise en œuvre de l'observateur proposé ; des comparaisons avec l'observateur à mémoire finie et l'observateur de Luenberger ont été effectuées, cette comparaison est réalisée en fixant comme critère d'évaluation des performances la somme quadratique E_N des erreurs d'estimation où l'indice N correspond au temps de fin de simulation :

$$E_N = \sum_{k=1}^{k=N} e(k)^T e(k) \quad (5.15)$$

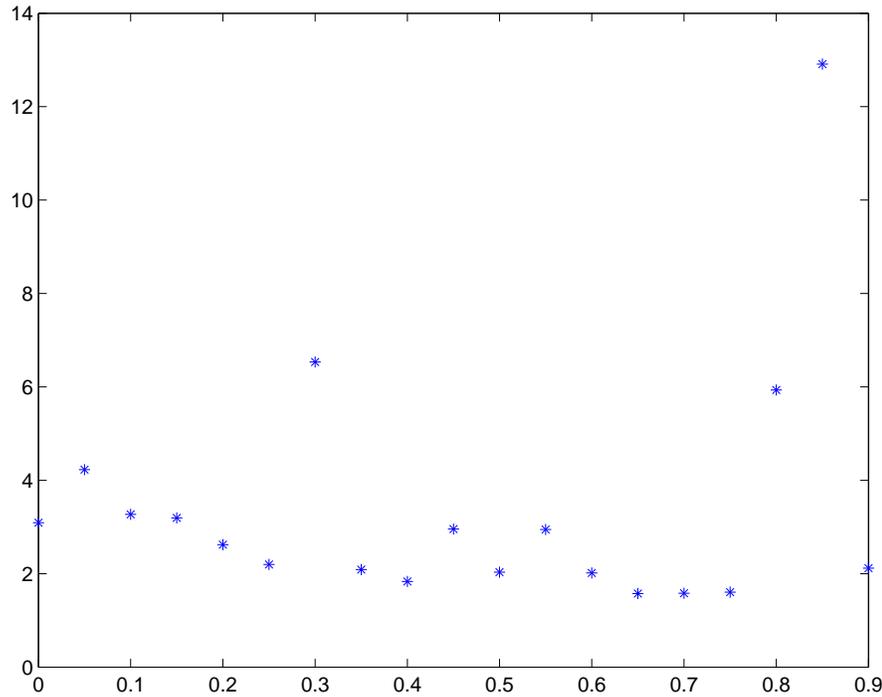


FIG. 5.1 – Évolution du terme $\|D\|\lambda_{\max}(P + I)/\sqrt{\lambda_{\min}(Q)}$ en fonction du paramètre τ

Pour l'exemple, nous considérons le système suivant :

$$x(k) = \begin{bmatrix} 0.85 & 0 \\ 0 & 0.8 \end{bmatrix} x(k-1) + \begin{bmatrix} 0.18 \\ 1.8 \end{bmatrix} u(k-1) + \begin{bmatrix} 0.1 \\ 0.1 \end{bmatrix} w(k-1)$$

$$y(k) = \begin{bmatrix} 1 & 0 \\ 0 & 0.5 \end{bmatrix} x(k) + v(k)$$

où $v(k) \in [-0.1; 0.1]$ et $w(k) \in [-0.2; 0.2]$.

Afin de déterminer le paramètre τ , on minimise le terme $\|D\|\lambda_{\max}(P + I)/\sqrt{\lambda_{\min}(Q)}$ pour $\tau \in [0; 1]$ en utilisant la fonction `Matlab fminbnd`. Cette dernière fonction est basée sur la méthode "Golden Section Search" (Cheney and Kincaid, 1994) et l'interpolation parabolique. Pour ces simulations, on a choisi une longueur de fenêtre $m = 4$. La figure 5.1 montre l'évolution du terme $\|D\|\lambda_{\max}(P + I)/\sqrt{\lambda_{\min}(Q)}$ en fonction du paramètre τ dans l'intervalle $[0; 1]$.

Ce minimum vaut $\tau = 0.6910$ et la résolution des conditions LMI (5.14) aboutit aux

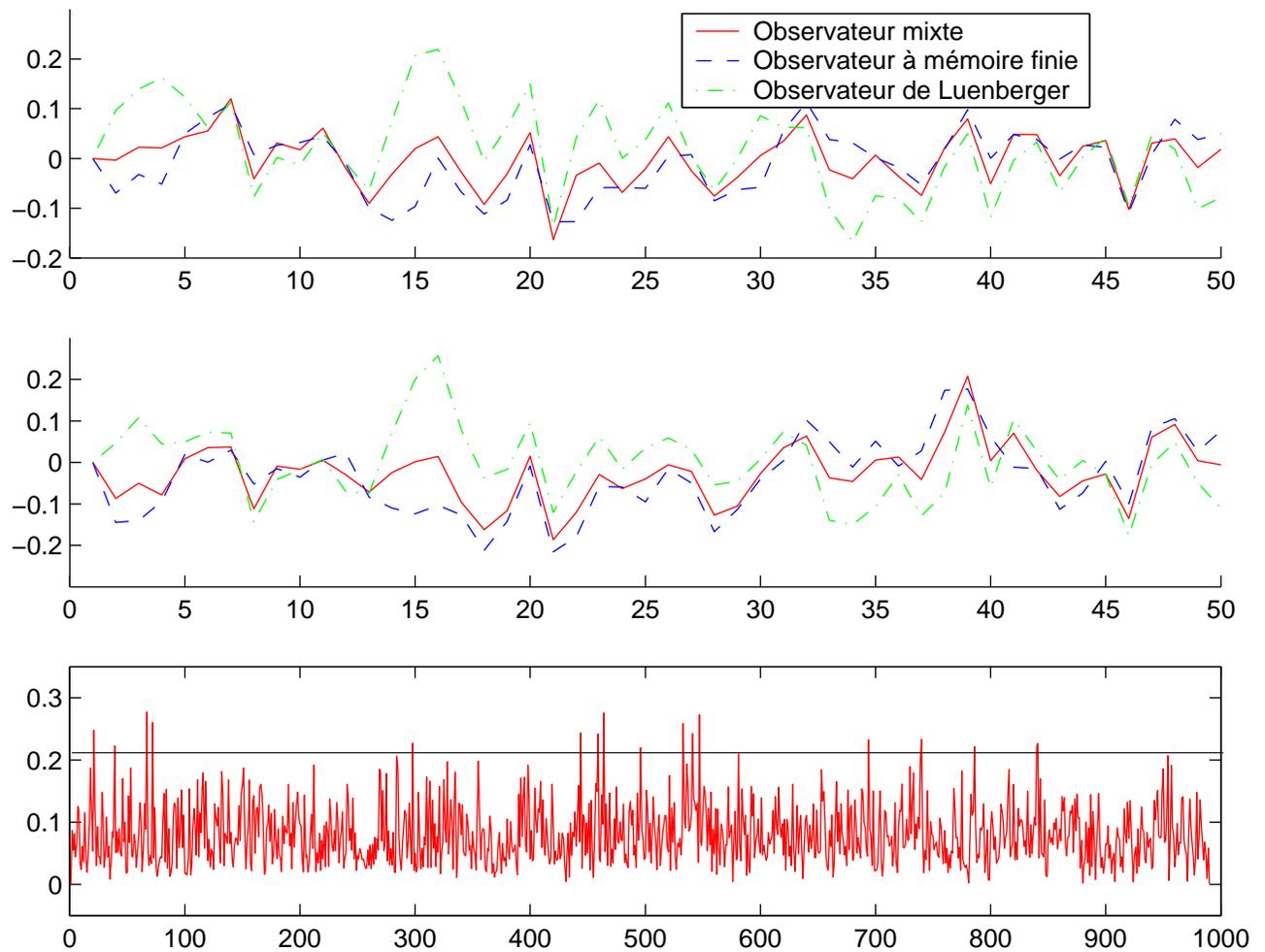


FIG. 5.2 – Erreur d'estimation pour les deux composantes du vecteur d'état et norme de $\|e(k)\|$

résultats suivants :

$$P = \begin{bmatrix} 1.4289 & 0.0001 \\ 0.0001 & 1.4289 \end{bmatrix}$$

$$Q = \begin{bmatrix} 1.2983 & 0.00005 \\ 0.00005 & 1.2983 \end{bmatrix}$$

$$K = \begin{bmatrix} 0.8161 & -0.0194 \\ 0.0008 & 1.5403 \end{bmatrix}$$

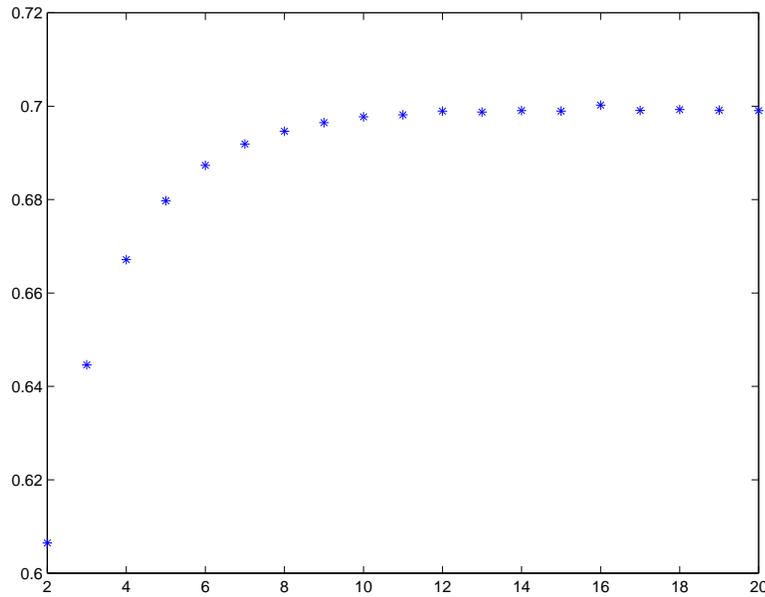


FIG. 5.3 – Variation du paramètre τ en fonction de la longueur de la fenêtre

Les deux premières parties de la figure 5.2 représentent les erreurs d'estimation d'état, calculées par l'observateur proposé, par OMF, de même longueur de fenêtre $m = 7$ que l'observateur proposé, et l'observateur de Luenberger de même gain K que l'observateur proposé. On peut remarquer que l'observateur mixte donne de meilleurs résultats que ceux obtenus avec l'OMF et l'observateur de Luenberger. On peut quantifier cette supériorité par le calcul de la somme quadratique des erreurs d'estimation d'état $e(k)$ des trois observateurs, l'observateur mixte ($E_N^{mixte} = 8.9085$), OMF ($E_N^{OMF} = 14.4150$) et l'observateur de Luenberger ($E_N^{LBO} = 18.5382$). La dernière partie de la figure 5.2 représente la norme des erreurs d'estimation de l'observateur mixte où la ligne droite représente la borne $r = 0.21$. On observe que la norme de l'erreur d'estimation dépasse en quelques points la borne r ce qui est tout à fait normal car r est la borne qui garantit la décroissance de l'erreur d'estimation et non pas la borne de l'erreur d'estimation. La figure 5.3 représente la variation de τ en fonction de la longueur de la fenêtre.

Dans cette partie, on a élaboré un nouveau type d'observateur, sa structure est le résultat d'une pondération entre l'observateur à mémoire finie et l'observateur de Luenberger. En

utilisant une fonction de Lyapunov quadratique, on a formulé les conditions sous forme LMI pour garantir la stabilité de l'observateur. On a procédé au calcul d'une borne garantissant la décroissance de l'erreur d'estimation en présence de bruit sur la sortie et sur le système ; cette borne a ensuite été minimisée en optimisant la valeur du paramètre de pondération τ .

Dans la partie qui suit, l'observateur proposé sera intégré dans le cadre des estimateurs multi-modèle du chapitre 1, de la même manière que les observateurs à mémoire finie.

5.5 Observateur mixte pour les systèmes à commutation markovienne

Afin d'appliquer l'observateur mixte aux systèmes à commutation, nous allons suivre les mêmes étapes que celles du chapitre 4. Nous considérons un système décrit par un ensemble de modèles $M_i, i = 1, \dots, r$, chaque modèle représentant un comportement du système. L'objectif est de détecter, à chaque instant, le modèle qui approche le mieux le comportement du système et simultanément d'estimer l'état du système. On suppose que les transitions d'un modèle à un autre sont décrites par un processus Markovien régi par la matrice de transition de Markov Π .

5.5.1 Développement de la méthode

Considérons le $j^{\text{ème}}$ modèle :

$$\begin{cases} x(k) = A_j x(k-1) + B_j u(k-1) + G_j w(k-1) \\ z(k) = C_j x(k) + v(k) \end{cases} \quad (5.16)$$

L'estimation d'état de ce modèle peut être effectuée à l'aide de l'observateur mixte proposé. On obtient ainsi :

$$\begin{cases} \hat{x}_i(k+1) = \tau_i(\Omega_i Z_m^{k+1} - \Phi_i U_m^{k+1}) \\ \quad + (1 - \tau_i)(A_i \hat{x}_i(k) + B_i u(k) - K_i(\hat{z}_i(k) - z(k))) \\ \hat{z}_i(k) = C_i \hat{x}_i(k) \end{cases} \quad (5.17)$$

Les matrices Ω_i , Φ_i sont construites en remplaçant les matrices A , B et C par les matrices A_j , B_j et C_j décrivant le $j^{\text{ème}}$ modèle. Les paramètres τ_j et K_j sont calculés pour chaque modèle.

L'estimation de l'état $\hat{x}(k)$ du système à commutation est alors calculée comme une somme pondérée des états des différents modèles :

$$\hat{x}(k) = \sum_{j=1}^r \hat{x}_j(k) \mu_j(k) \quad (5.18)$$

Comme dans le chapitre 4, en nous inspirant des travaux de (Bar-Shalom, 1990), la probabilité que le système fonctionne selon le $j^{\text{ème}}$ modèle à l'instant k est calculée de la manière suivante :

$$\mu_j(k) = P\{M_j(k)|Z_m^k\} \quad (5.19)$$

L'équation (5.19) peut alors s'écrire :

$$\mu_j(k) = P\{M_j(k)|Z_m^{k-1}, z(k)\} \quad (5.20)$$

puis, en utilisant la formule de Bayes :

$$\mu_j(k) = \frac{p[z(k)|M_j(k), Z_m^{k-1}] P\{M_j(k)|Z_m^{k-1}\}}{\sum_{l=1}^r p[z(k)|M_l(k), Z_m^{k-1}] P\{M_l(k)|Z_m^{k-1}\}} \quad (5.21)$$

Afin d'alléger les notations, posons :

$$\lambda_i(k) = p[z(k)|M_i(k), Z_m^{k-1}] \quad (5.22)$$

Développons également la probabilité d'activation du modèle j à l'instant k , conditionnellement au modèle actif à l'instant $k-1$:

$$P\{M_j(k)|Z_m^{k-1}\} = \sum_{i=1}^r P\{M_j(k)|M_i(k-1), z(k-1)\} P\{M_i(k-1)|Z_m^{k-1}\} \quad (5.23)$$

On obtient alors la récurrence suivante sur la probabilité que le système opère selon le modèle j à l'instant k :

$$\mu_j(k) = \frac{\lambda_j(k) \sum_{i=1}^r \pi_{ij} \mu_j(k-1)}{\sum_{l=1}^r \lambda_l(k) \sum_{i=1}^r \pi_{il} \mu_j(k-1)} \quad (5.24)$$

Cette expression permet de calculer la probabilité d'activation de chaque modèle connaissant les vraisemblances $\lambda_j(k)$ et les probabilités π_{ij} de passage d'un modèle à un autre.

5.5.2 Exemple d'application

Pour l'application de la méthode proposée, on considère le problème du chapitre 2, décrit par un modèle de fonctionnement normal (A_1, B_1, C_1) , un modèle de défaut d'actionneur (A_2, B_2, C_2) et un modèle de défaut de capteur (A_3, B_3, C_3) ; les différentes matrices étant définies par :

$$A_i = \begin{bmatrix} -0.0005 & -0.0084 \\ 0.0517 & 0.8069 \end{bmatrix}, \quad i = 1 \dots 3$$

$$B_1 = [0.1815 \ 1.7902]^T, \quad C_1 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}.$$

$$B_2 = [0.1815 \ 0]^T, \quad C_2 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}.$$

$$B_3 = [0.1815 \ 1.7902]^T, \quad C_3 = \begin{bmatrix} 1 & 0 \\ 0 & 0.1 \end{bmatrix}.$$

Les gains K_j et les paramètres τ_j relatifs à chaque modèle M_j sont donnés par :

$$K_1 = \begin{bmatrix} 0.1326 & -0.1013 \\ 0.0799 & 0.8428 \end{bmatrix}$$

$$K_2 = \begin{bmatrix} 0.1326 & -0.1013 \\ 0.0799 & 0.8428 \end{bmatrix}$$

$$K_3 = \begin{bmatrix} 0.0470 & -0.0349 \\ 0.0468 & 0.8681 \end{bmatrix}$$

et $\tau_1 = 0.8247$, $\tau_2 = 0.8247$ et $\tau_3 = 0.8451$.

Un scénario similaire à celui du deuxième chapitre est envisagé pour comparer les performances de la méthode proposée avec celle de l'algorithme GPB1 et la méthode du chapitre 4. Les résultats de cette comparaison sont présentés dans le tableau 5.1. Les commutations sont générées en utilisant une matrice de Markov Π donnée par :

$$\Pi = \begin{bmatrix} 0.95 & 0.025 & 0.025 \\ 0.05 & 0.95 & 0 \\ 0.05 & 0 & 0.95 \end{bmatrix}$$

	BDI	FA	ND	IID	TMD/(pe)
Méthode proposée	0.8099	0.0096	0.1040	0.0104	0.3333
Algorithme chapitre 4	0.7666	0.0066	0.1981	0.0104	1.9167
Algorithme GPB	0.6525	0.00003	0.2308	0.0004	9.5000

TAB. 5.1 – Performance de la détection pour la méthode proposée

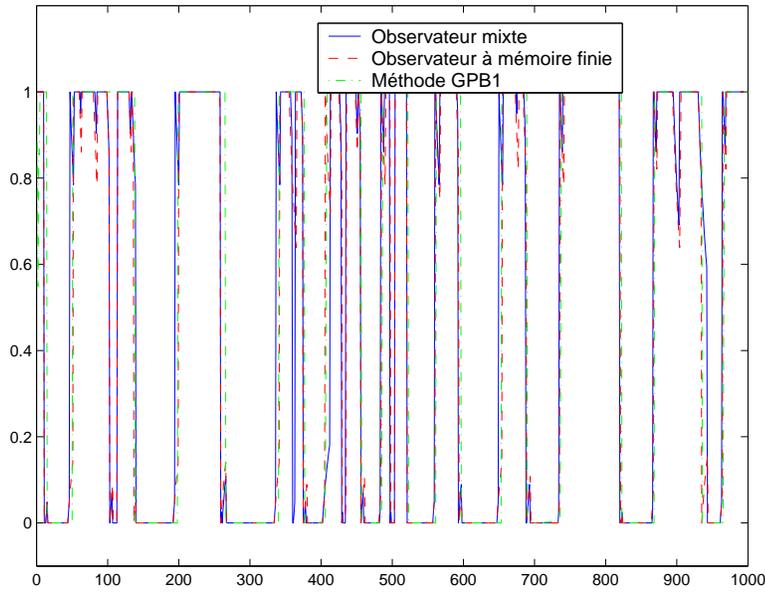


FIG. 5.4 – Probabilité d'activation du modèle 1

Les variances des bruits sont données par :

$$Q = \begin{bmatrix} 0.001 & 0 \\ 0 & 0.001 \end{bmatrix}.$$

$$R = 2 \times \begin{bmatrix} 0.01 & 0 \\ 0 & 0.4 \end{bmatrix}.$$

Les figures 5.4, 5.5 et 5.6 représentent les probabilités d'activation des trois modèles calculées par les différentes méthodes proposées. Le tableau 5.1 représente les différents indices de performance pour le diagnostic présentés au chapitre 2. On remarque, sur cet exemple, que les résultats obtenus avec l'observateur mixte sont meilleurs que ceux de la

5.5. Observateur mixte pour les systèmes à commutation markovienne

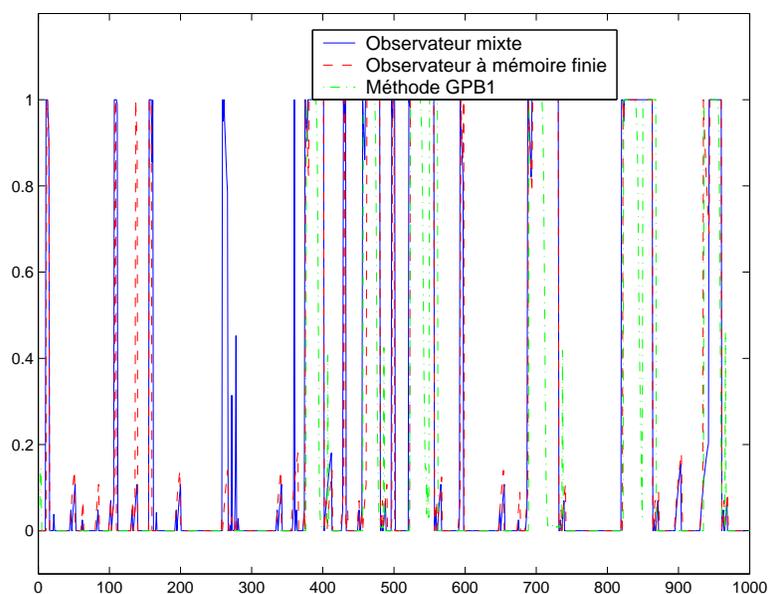


FIG. 5.5 – Probabilité d'activation du modèle 2

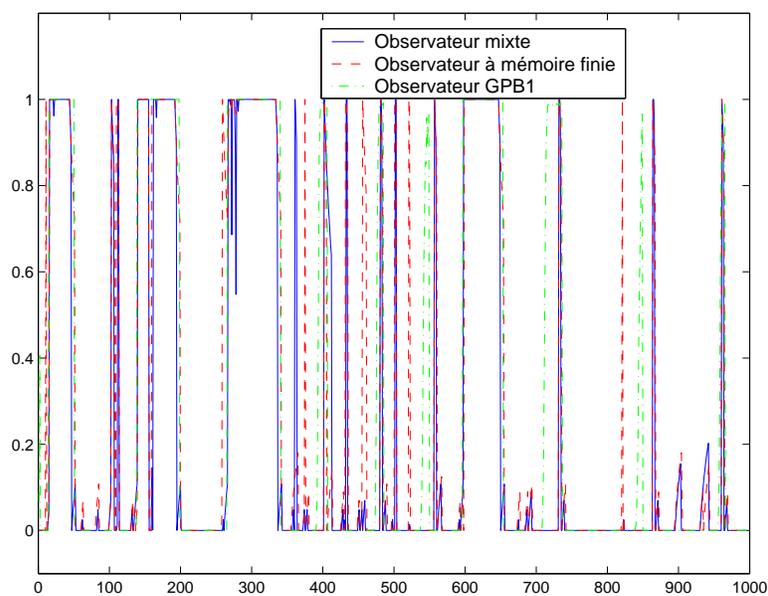


FIG. 5.6 – Probabilité d'activation du modèle 3

méthode GPB1 et de la méthode utilisant des observateurs à mémoire finie proposée dans le chapitre 4.

5.6 Conclusion

Dans ce chapitre, nous avons tout d'abord proposé de combiner l'observateur de Luenberger et l'OMF afin d'améliorer les performances du diagnostic. Les conditions de stabilité de cet observateur mixte ont été formulées en terme d'inégalités matricielles. L'observateur proposé a été ensuite utilisé dans le cadre des systèmes à commutation markovienne où l'on a constaté, sur des exemples de simulation, une amélioration des performances du diagnostic.

6

Estimateur multi-modèle structuré

Sommaire

6.1	Introduction	129
6.2	Formulation du problème	130
6.3	Matrice des probabilités de transition	130
6.3.1	Matrice des probabilités de transition entre deux sous-ensembles de modèles	131
6.3.2	Matrice des probabilités de transition au sein d'un sous-ensemble de modèles	132
6.4	Procédure d'estimation	133
6.4.1	Sous-estimateur	133
6.4.2	Estimateur global	136
6.4.3	Choix du modèle actif	137

6.5 Exemple	141
6.6 Conclusion	147

6.1 Introduction

Dans les méthodes d'estimation multi-modèle proposées dans les chapitres précédents, l'ensemble des modèles est choisi de manière à représenter tous les modes de fonctionnement possibles du système. Le choix des modèles ainsi que leur nombre est une étape importante dans l'élaboration d'un algorithme d'estimation multi-modèle, à cause de son influence sur la qualité de l'estimation d'état du système ainsi que sur la détection du mode actif.

Un nombre important de modèles conduit à une détérioration des résultats caractérisée par une mauvaise détection des instants de commutations et des modes actifs (Li and Bar-Shalom, 1996). Afin de résoudre ce problème, l'utilisation d'un ensemble variable de modèles a été introduite par X. R. Li dans une série d'articles (Li and Bar-Shalom, 1996), (Li et al., 1999a), (Li et al., 1999b) et (Li, 2000) dans le but de réduire la dispersion des probabilités d'activation qui résulte de l'utilisation d'un nombre important de modèles. La méthode proposée se base sur l'utilisation, à chaque instant, d'un sous-ensemble de modèles appartenant à l'ensemble des modèles du système. Elle nécessite la connaissance de toutes les séquences de commutation possibles, ce qui peut être une contrainte lors de la mise en oeuvre. De surcroît, la méthode proposée par X. R. Li n'est pas appropriée dans le cas de commutations arbitraires, car la méthode se base sur une logique de commutation.

Dans ce chapitre, la méthode proposée consiste à subdiviser l'ensemble des modèles en un certain nombre de sous-ensembles. Chaque sous-ensemble est utilisé pour synthétiser un estimateur multi-modèle, appelé sous-estimateur, qui donne une estimée basée sur l'utilisation des modèles du sous-ensemble. On constitue ainsi un banc de sous-estimateurs. Un autre estimateur multi-modèle, appelé estimateur global, est élaboré afin de déterminer le sous-ensemble *actif*, c'est-à-dire celui contenant le modèle actif. Le calcul de l'estimation d'état est effectué en utilisant les estimations fournies par les différents sous-estimateurs. Autrement dit, il s'agit de déterminer, dans un premier temps, le sous-ensemble actif, puis ensuite le modèle actif dans ce sous-ensemble. La méthode permet aussi d'utiliser chaque sous-ensemble pour un type particulier de panne afin de faciliter la prise de décision.

6.2 Formulation du problème

On considère un système représenté par un ensemble de modèles $M = \{M_1, M_2, \dots, M_r\}$, r étant le nombre de modèles. Le comportement du système pour chaque modèle d'évolution d'état M_j est représenté par :

$$\begin{cases} x(k) = A_j x(k-1) + B_j u(k-1) + G_j w(k-1) \\ z(k) = C_j x(k) + v(k) \end{cases} \quad (6.1)$$

Les transitions d'un modèle à un autre suivent un processus Markovien, défini par sa matrice de transition Π :

$$\Pi = \begin{bmatrix} \pi_{11} & \cdots & \pi_{1r} \\ \vdots & \ddots & \vdots \\ \pi_{r1} & \cdots & \pi_{rr} \end{bmatrix} \quad (6.2)$$

Une subdivision de l'ensemble M en sous-ensembles s_i est réalisée afin de construire un ensemble $S = \{s_1, s_2, \dots, s_n\}$ avec n le nombre de sous-ensembles et $s_i \cap s_j = \emptyset, \forall i \neq j \in [1, \dots, n]$ et $\bigcup_{i=1}^n s_i = M$.

La méthode proposée consiste tout d'abord à construire un sous-estimateur pour chaque sous-ensemble $s_j = \{M_{\{s_j, 1\}}, M_{\{s_j, 2\}}, \dots, M_{\{s_j, r_j\}}\}$ avec r_j le nombre de modèles constituant s_j et $\{s_j, i\}$ correspond à l'indice du modèle dans l'ensemble M . Ensuite un estimateur global qui s'appuie sur les estimées issues des sous-estimateurs calcule une estimée globale. Pour pouvoir construire l'estimateur global, on a besoin de calculer la matrice de transition de Markov Π_s qui exprime la probabilité de passage d'un sous-ensemble s_i à un sous-ensemble s_j à partir de la matrice Π qui régit le passage d'un modèle M_i à un modèle M_j .

6.3 Matrice des probabilités de transition

La matrice de transition de Markov est indispensable pour réaliser les différentes étapes de la méthode proposée. A partir de la matrice de transition de Markov Π , supposée connue, du système initial, on construit Π_s la matrice de passage d'un sous-ensemble à un autre et $\Pi^{s_j}, j = 1, \dots, n$ les matrices de passage d'un modèle à un autre appartenant au même sous-ensemble s_j .

6.3.1 Matrice des probabilités de transition entre deux sous-ensembles de modèles

On considère que le processus de commutation stochastique est *stationnaire* ; on peut alors écrire :

$$\gamma = \Pi\gamma, \quad \gamma_j = \sum_{i=1}^r \gamma_i \pi_{ij} \quad \forall j \in [1, \dots, r] \quad (6.3)$$

où γ est le vecteur des probabilités γ_j relatives à chaque modèle M_j avec $\sum_{i=1}^r \gamma_i = 1$. Les γ_j vont permettre de calculer la matrice de transition de Markov d'un sous-ensemble à un autre à partir de la matrice de transition de Markov Π .

Les probabilités de commutation d'un sous-ensemble à un autre peuvent être données par l'équation qui suit :

$$\pi_{s_i|s_j} = \frac{\sum_{l=1}^{r_i} \gamma_l \sum_{h=1}^{r_j} \pi_{\{s_i, l\}\{s_j, h\}}}{\sum_{l=1}^{r_i} \gamma_l} \quad i, j = 1, \dots, n \quad (6.4)$$

Considérons l'exemple d'un ensemble à quatre modèles $M = \{M_1, M_2, M_3, M_4\}$ subdivisé en deux sous-ensembles $S = \{s_1, s_2\}$ avec $s_1 = \{M_1, M_2\}$ et $s_2 = \{M_3, M_4\}$. La matrice de Markov Π est donnée par :

$$\Pi = \begin{bmatrix} \pi_{11} & \pi_{12} & \pi_{13} & \pi_{14} \\ \pi_{21} & \pi_{22} & \pi_{23} & \pi_{24} \\ \pi_{31} & \pi_{23} & \pi_{33} & \pi_{34} \\ \pi_{41} & \pi_{24} & \pi_{43} & \pi_{44} \end{bmatrix} \quad (6.5)$$

On cherche à calculer Π_s

$$\Pi_s = \begin{bmatrix} \pi_{s_1|s_1} & \pi_{s_1|s_2} \\ \pi_{s_2|s_1} & \pi_{s_2|s_2} \end{bmatrix} \quad (6.6)$$

A partir de l'équation (6.3) on obtient le système d'équations :

$$\begin{aligned} \gamma_1 &= \gamma_1 \pi_{11} + \gamma_2 \pi_{12} + \gamma_3 \pi_{13} + \gamma_4 \pi_{14} \\ \gamma_2 &= \gamma_1 \pi_{21} + \gamma_2 \pi_{22} + \gamma_3 \pi_{23} + \gamma_4 \pi_{24} \\ \gamma_3 &= \gamma_1 \pi_{31} + \gamma_2 \pi_{32} + \gamma_3 \pi_{33} + \gamma_4 \pi_{34} \\ \gamma_4 &= \gamma_1 \pi_{41} + \gamma_2 \pi_{42} + \gamma_3 \pi_{43} + \gamma_4 \pi_{44} \end{aligned} \quad (6.7)$$

et en tenant compte de la contrainte $\sum_{i=1}^r \gamma_i = 1$, on obtient les γ_i . Le calcul des $\pi_{s_i|s_j}$ s'effectue de la manière suivante :

$$\begin{aligned}\pi_{s_1|s_1} &= (\gamma_1\pi_{11} + \gamma_1\pi_{12} + \gamma_2\pi_{21} + \gamma_2\pi_{22})/(\gamma_1 + \gamma_2) \\ \pi_{s_1|s_2} &= (\gamma_1\pi_{13} + \gamma_1\pi_{14} + \gamma_2\pi_{23} + \gamma_2\pi_{24})/(\gamma_1 + \gamma_2) \\ \pi_{s_2|s_1} &= (\gamma_3\pi_{31} + \gamma_3\pi_{32} + \gamma_4\pi_{41} + \gamma_4\pi_{42})/(\gamma_3 + \gamma_4) \\ \pi_{s_2|s_2} &= (\gamma_3\pi_{33} + \gamma_3\pi_{34} + \gamma_4\pi_{43} + \gamma_4\pi_{44})/(\gamma_3 + \gamma_4)\end{aligned}\tag{6.8}$$

6.3.2 Matrice des probabilités de transition au sein d'un sous-ensemble de modèles

Afin de pouvoir réaliser les sous-observateurs, il faut construire leurs matrices de transition respectives Π^{s_j} . Elle sont déduites directement de la matrice Π :

$$\pi_{lh}^{s_j} = \frac{\pi_{lh}}{\sum_{i=1}^{r_j} \pi_{\{s_j, l\}\{s_j, i\}}} \quad l, h = 1, \dots, r_j\tag{6.9}$$

Pour illustrer la procédure, on reprend l'exemple précédent. Le calcul des deux matrices Π^{s_1} et Π^{s_2} correspondant aux deux sous-ensembles s_1, s_2 s'effectue de la manière suivante :

$$\Pi^{s_j} = \begin{bmatrix} \pi_{11}^{s_j} & \pi_{12}^{s_j} \\ \pi_{21}^{s_j} & \pi_{22}^{s_j} \end{bmatrix}\tag{6.10}$$

$$\begin{aligned}\pi_{11}^{s_1} &= \frac{\pi_{11}}{\pi_{11} + \pi_{12}} \\ \pi_{12}^{s_1} &= \frac{\pi_{12}}{\pi_{11} + \pi_{12}} \\ \pi_{21}^{s_1} &= \frac{\pi_{21}}{\pi_{21} + \pi_{22}} \\ \pi_{22}^{s_1} &= \frac{\pi_{22}}{\pi_{21} + \pi_{22}}\end{aligned}\tag{6.11}$$

$$\begin{aligned}\pi_{11}^{s_2} &= \frac{\pi_{33}}{\pi_{33} + \pi_{34}} \\ \pi_{12}^{s_2} &= \frac{\pi_{34}}{\pi_{33} + \pi_{34}} \\ \pi_{21}^{s_2} &= \frac{\pi_{43}}{\pi_{43} + \pi_{44}} \\ \pi_{22}^{s_2} &= \frac{\pi_{44}}{\pi_{43} + \pi_{44}}\end{aligned}\tag{6.12}$$

6.4 Procédure d'estimation

Rappelons que la procédure consiste à partitionner l'ensemble M des modèles en sous-ensembles s_j . Un sous-estimateur sera construit pour chaque sous-ensemble s_j . Ce sous-estimateur peut être un estimateur IMM ou GPB ou n'importe quel autre estimateur multi-modèle bayésien. L'estimateur multi-modèle fournira en sortie l'estimation d'état et sa variance ainsi que la probabilité d'activation de chaque modèle appartenant à ce sous-ensemble. Dans ce qui suit, l'élaboration de l'estimateur est basée principalement sur l'algorithme GPB et sur le principe de l'utilisation de deux niveaux d'estimation (sous-estimateur et estimateur global).

6.4.1 Sous-estimateur

A chaque instant k , les entrées du sous-estimateur sont la mesure $z(k)$, l'estimée issue de l'estimateur global $\hat{x}(k-1|k-1)$ et sa covariance $P(k-1|k-1)$. Ces grandeurs sont utilisées à l'entrée des r_j différents filtres de Kalman associés à chaque sous-ensemble $j = 1, \dots, n$. A la sortie de chaque filtre, on obtient les estimées $\hat{x}_i^{s_j}(k|k)$ pour $i = 1, \dots, r_j$ ainsi que leurs covariances associées $P_i^{s_j}(k|k)$.

Calcul des vraisemblances

La fonction de vraisemblance correspondant au $i^{\text{ème}}$ filtre de Kalman du $j^{\text{ème}}$ sous-ensemble s_j est décrite par :

$$\lambda_i^{s_j}(k) = p[z(k) | Z^{k-1}, M_{\{s_j, i\}}] \quad (6.13)$$

avec $i = 1, \dots, r_j$, est calculée en utilisant l'estimée $\hat{x}_i^{s_j}(k-1|k-1)$ et sa covariance :

$$\lambda_i^{s_j}(k) = p[z(k) | M_{\{s_j, i\}}, \hat{x}_i^{s_j}(k-1|k-1), P(k-1|k-1)] \quad (6.14)$$

Sous l'hypothèse de la normalité des bruits entachant les mesures et les états du système, la vraisemblance s'écrit :

$$\lambda_i^{s_j}(k) = \mathbf{N}[\nu_i^{s_j}(k); 0, S_i^{s_j}(k)] \quad (6.15)$$

avec $\nu_i^{s_j}(k) = z(k) - C_i^{s_j}(A_i^{s_j}\hat{x}(k-1|k-1) + B_i^{s_j}u(k-1))$ et $S_i^{s_j}(k)$ sa variance, où $B_i^{s_j}$, et $C_i^{s_j}$ correspondent au $i^{\text{ème}}$ modèle du sous-ensemble s_j .

Mise à jour des probabilités d'activation des modèles

Comme dans l'algorithme GPB1, l'utilisation de la formule de Bayes permet d'écrire :

$$\mu_i^{s_j}(k) = \frac{1}{c_j^{s_j}(k)} \lambda_i^{s_j}(k) \sum_{l=1}^{r_j} \pi_{li}^{s_j} \mu_l^{s_j}(k-1) \quad (6.16)$$

où $\mu_i^{s_j}(k)$ est la probabilité d'activation, à l'instant k , du $i^{\text{ème}}$ modèle appartenant à l'ensemble s_j et $c_j^{s_j}$ étant une constante de normalisation :

$$c_j^{s_j}(k) = \sum_{i=1}^{r_j} \lambda_i^{s_j}(k) \sum_{l=1}^{r_j} \pi_{li}^{s_j} \mu_l^{s_j}(k-1) \quad (6.17)$$

Calcul de l'estimée mélange associée à s_j

L'estimée mélange est calculée de la manière suivante :

$$\hat{x}^{s_j}(k|k) = \sum_{i=1}^{r_j} \mu_i^{s_j}(k) \hat{x}_i^{s_j}(k|k) \quad (6.18)$$

et sa covariance associée :

$$P^{s_j}(k|k) = \sum_{i=1}^{r_j} \mu_i^{s_j}(k) \left\{ P_i^{s_j}(k|k) + [\hat{x}_i^{s_j}(k|k) - \hat{x}^{s_j}(k|k)] [\hat{x}_i^{s_j}(k|k) - \hat{x}^{s_j}(k|k)]' \right\} \quad (6.19)$$

où l'estimée $\hat{x}_i^{s_j}(k|k)$ provient du filtre de Kalman associé au $i^{\text{ème}}$ modèle de l'ensemble s_j . Elle est obtenue en remplaçant dans l'équation (1.3) les A , B , C par les $A_i^{s_j}$, $B_i^{s_j}$, $C_i^{s_j}$.

Chaque sous-estimateur fournit une estimation mélange $\hat{x}^{s_j}(k|k)$, ainsi que les probabilités d'activation $\mu_i^{s_j}(k)$, $i = 1, \dots, r_j$ de chaque modèle appartenant au sous-ensemble s_j .

Les différentes étapes du sous-estimateur sont représentées sur la figure 6.1. On remarque qu'à l'entrée du sous-estimateur, on a la mesure à l'instant k et l'état global $x(k-1|k-1)$ qui sera défini par la suite et en sortie l'estimée $\hat{x}^{s_j}(k|k)$ relative au sous-ensemble s_j .

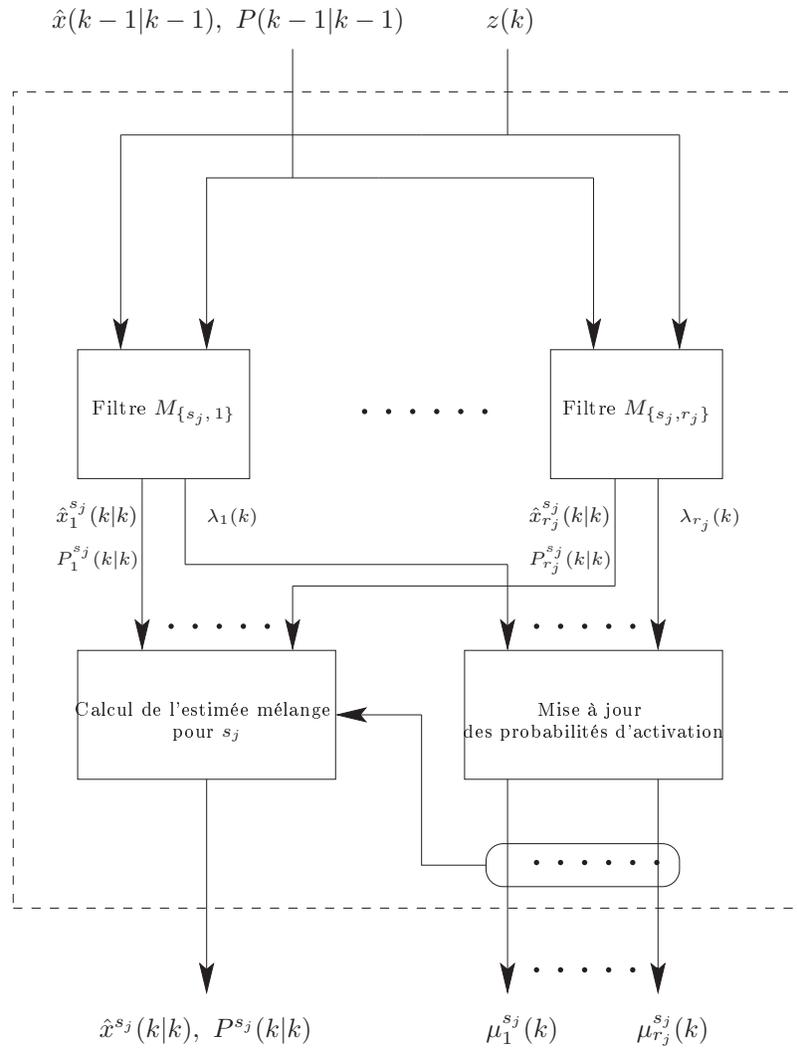


FIG. 6.1 – Sous-estimateur

6.4.2 Estimateur global

La structure de l'estimateur global est celle d'un estimateur de type GPB1. Au lieu d'être des filtres de Kalman ordinaires, les n filtres en parallèle sont les n sous-estimateurs précédents. Les entrées de l'estimateur global sont la mesure $z(k)$ et l'estimée, à l'instant $k-1$, $\hat{x}(k-1|k-1)$ et sa variance $P(k-1|k-1)$. Ces grandeurs sont utilisées à l'entrée de chaque sous-estimateur. L'estimateur global fournit en sortie l'estimée $\hat{x}(k|k)$ à l'instant k et sa covariance $P(k|k)$. Au cours de l'estimation, on procède au calcul des probabilités d'activation de chaque sous-ensemble $\mu^{s_j}(k)$, $j = 1, \dots, n$.

Calcul des vraisemblances

La fonction de vraisemblance correspondant au $j^{\text{ème}}$ sous-estimateur s'écrit :

$$\lambda^{s_j}(k) = p[z(k) | Z^{k-1}, s_j] \quad (6.20)$$

avec $j = 1, \dots, n$. Elle est calculée comme suit :

$$\lambda^{s_j}(k) = p[z(k) | s_j, \hat{x}(k-1|k-1), P(k-1|k-1)] \quad (6.21)$$

Sous hypothèse gaussienne, cette vraisemblance peut s'écrire :

$$\lambda^{s_j}(k) = \mathbf{N}[\nu^{s_j}(k); 0, S^{s_j}(k)] \quad (6.22)$$

avec $\nu^{s_j}(k) = z(k) - C^{s_j}(k)(A^{s_j}(k)\hat{x}(k-1|k-1) + B^{s_j}(k)u(k-1))$ et sa variance $S^{s_j}(k)$.

De la même façon que $\hat{x}^{s_j}(k|k)$ (6.18), $A^{s_j}(k)$, $B^{s_j}(k)$ et $C^{s_j}(k)$ s'explicitent :

$$\begin{aligned} A^{s_j}(k) &= \sum_{i=1}^{r_j} \mu_i^{s_j}(k) A_i^{s_j} \\ B^{s_j}(k) &= \sum_{i=1}^{r_j} \mu_i^{s_j}(k) B_i^{s_j} \\ C^{s_j}(k) &= \sum_{i=1}^{r_j} \mu_i^{s_j}(k) C_i^{s_j} \end{aligned} \quad (6.23)$$

où $A_i^{s_j}$, $B_i^{s_j}$ et $C_i^{s_j}$ sont les matrices caractérisant le $i^{\text{ème}}$ modèle du sous-ensemble s_j , ce qui revient à représenter le sous-ensemble s_j par les matrices $A^{s_j}(k)$, $B^{s_j}(k)$ et $C^{s_j}(k)$. Ces matrices sont des sommes de matrices correspondant aux modèles $M_{s_j, i}$ appartenant au sous-ensemble s_j pondérées par leurs probabilités d'activation $\mu_i^{s_j}(k)$.

Mise à jour des probabilités d'activation des modèles

$$\mu^{s_j}(k) = \frac{1}{c^{s_j}(k)} \lambda^{s_j}(k) \sum_{i=1}^n \pi_{s_i|s_j} \mu^{s_j}(k-1) \quad (6.24)$$

c^{s_j} étant une constante de normalisation :

$$c^{s_j}(k) = \sum_{j=1}^n \lambda^{s_j}(k) \sum_{i=1}^n \pi_{s_i|s_j} \mu^{s_j}(k-1) \quad (6.25)$$

Calcul de l'estimée globale

A partir de (6.24) et de (6.18) on peut calculer :

$$\hat{x}(k|k) = \sum_{j=1}^n \mu^{s_j}(k) \hat{x}^{s_j}(k|k) \quad (6.26)$$

On remarque, à partir de l'équation (6.26), que pour calculer l'estimée globale $\hat{x}(k|k)$, on a besoin des estimées issues des sous-estimateurs s_j avec $j = 1, \dots, n$ et de leurs probabilités associées.

La variance associée à $\hat{x}(k|k)$ s'explique :

$$P(k|k) = \sum_{j=1}^n \mu^{s_j}(k) \{P^{s_j}(k|k) + [\hat{x}^{s_j}(k|k) - \hat{x}(k|k)][\hat{x}^{s_j}(k|k) - \hat{x}(k|k)]'\} \quad (6.27)$$

où $P^{s_j}(k|k)$ est donnée par l'équation (6.19).

En résumé, la méthode d'estimation proposée nécessite l'utilisation de r filtres de Kalman associés aux r modèles de fonctionnement. Ils sont répartis en n sous-estimateurs contenant chacun r_j filtres de Kalman. A la sortie de chaque sous-estimateur, on a l'estimée \hat{x}^{s_j} (6.18) et les probabilités d'activation $\mu_i^{s_j}$ (6.16) des modèles appartenant au sous-ensemble s_j . Les estimées de chaque sous-ensemble sont ensuite utilisées pour le calcul de l'estimée globale $\hat{x}(k|k)$ et des probabilités d'activation $\mu^{s_j}(k)$ associées à chaque sous-ensemble s_j .

6.4.3 Choix du modèle actif

De la procédure d'estimation proposée, résultent deux sortes de probabilités :

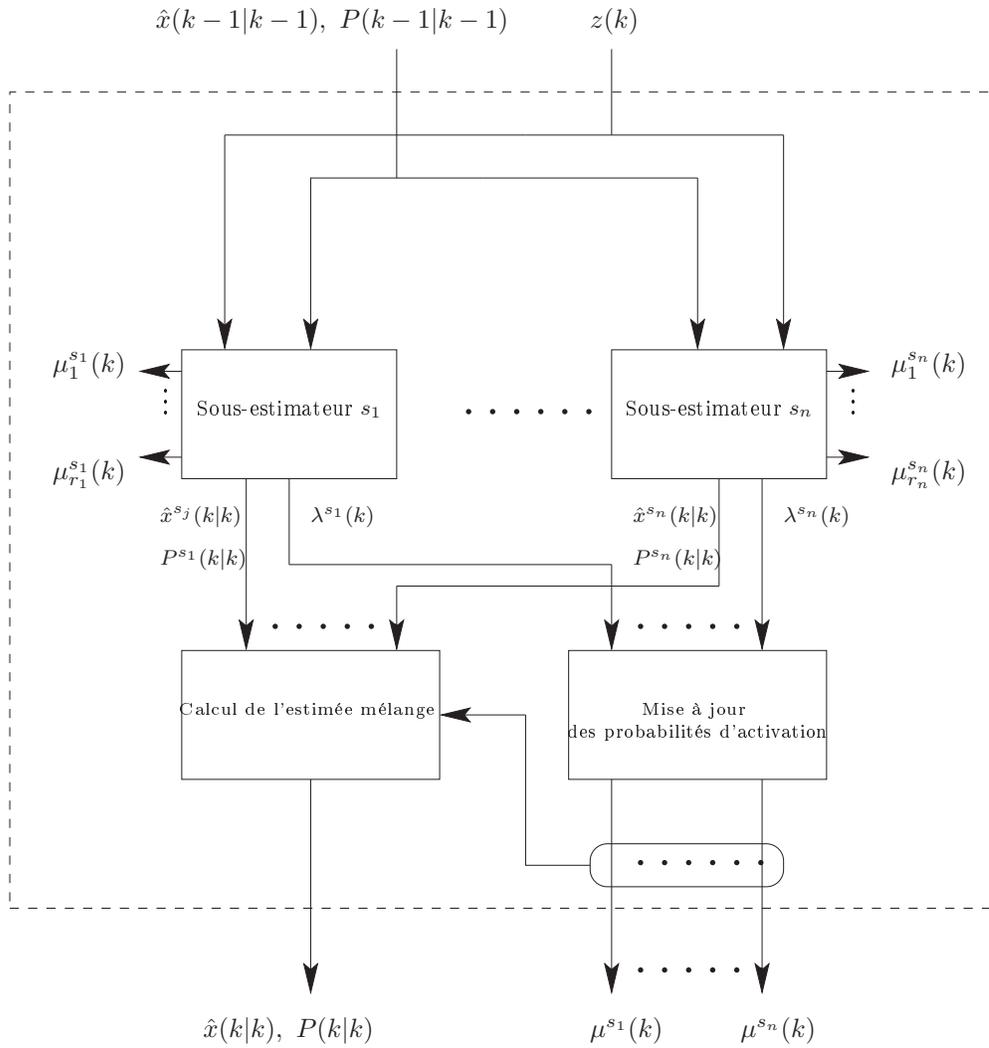


FIG. 6.2 – Estimateur global

- la probabilité d'activation $\mu_i^{s_j}$ du modèle i au sein du sous-ensemble s_j . Elle est calculée en ne prenant en considération que les modèles appartenant au même sous-ensemble.
- la probabilité μ^{s_j} d'activation du sous-ensemble s_j .

Plusieurs stratégies de décision sur le modèle actif peuvent être envisagées. Elles sont basées sur l'utilisation des deux types de probabilités issues de l'algorithme d'estimation que nous avons proposé dans ce chapitre.

Décision par calcul direct

La probabilité d'activation d'un modèle quelconque $M_i = M_l^{s_j}$, $M_i \in s_j$ est calculée directement par le produit de la probabilité d'activation μ^{s_j} du sous-ensemble s_j auquel appartient le modèle avec sa probabilité d'activation au sein de ce sous-ensemble $\mu_l^{s_j}$:

$$\mu_i(k) = \mu^{s_j}(k)\mu_l^{s_j}(k) \quad (6.28)$$

A un instant k donné, la prise de décision pour la détection du modèle actif est directement réalisée en exploitant les $\mu_i(k)$: le modèle déclaré actif est celui ayant la probabilité d'activation $\mu_i(k)$ la plus élevée. Comme dans le chapitre 2, on peut, à un instant k donné, ne pas prendre de décision (c'est-à-dire déclarer qu'aucun modèle n'est actif) si la probabilité $\mu_i(k)$ est inférieure à un seuil donné.

Décision hiérarchisée

On peut envisager, à un instant k donné, une prise de décision à deux niveaux. Pour le premier, on cherche le sous-ensemble contenant le modèle actif en exploitant les probabilités d'activation $\mu^{s_j}(k)$ des sous-ensembles. Puis, pour le deuxième, on détecte le modèle actif en utilisant les probabilités d'activation $\mu_i^{s_j}(k)$ des modèles appartenant au sous-ensemble choisi au cours du premier niveau.

Le premier niveau de décision consiste à trouver le sous-ensemble s_j qui a la probabilité d'activation $\mu^{s_j}(k)$ la plus élevée, de lui attribuer la valeur 1, et de donner la valeur 0 aux probabilités des autres sous-ensembles ; cela peut être écrit :

Si l'on a j tel que $\mu^{s_j}(k) > \mu^{s_i}(k) \quad \forall i = 1, \dots, n$

alors on impose $\mu^{s_j}(k) = 1$ et $\mu^{s_i}(k) = 0$

Puis les probabilités d'activation $\mu_i(k)$ des modèles peuvent être calculées par $\mu_i(k) = \mu^{s_j}(k)\mu_l^{s_j}(k)$ ce qui revient à écrire :

$$\mu_i = \mu_l^{s_j}(k) \text{ pour } M_i \in s_j, \quad M_i = M_l^{s_j}$$

$$\mu_i(k) = 0 \text{ pour } M_i \notin s_j$$

A partir des probabilités $\mu_i(k)$, on prend une deuxième décision pour déterminer le modèle actif. Comme dans la décision par calcul direct, on peut, à un instant k donné, ne pas prendre de décision (c'est-à-dire déclarer qu'aucun modèle n'est actif) si la probabilité $\mu_i(k)$ est inférieure à un seuil donné.

La décision hiérarchisée permet d'éliminer l'influence des modèles appartenant aux autres sous-ensembles qui ont été déclarés non actifs par le premier niveau de décision et permet de ne considérer, au moment du deuxième niveau de décision, qu'un sous-ensemble de modèles au lieu de tout l'ensemble des modèles.

Décision mixte

C'est un couplage entre la décision par calcul direct et la décision hiérarchisée. Un paramètre γ est défini afin de prendre la décision sur l'activation des sous-ensembles. Si la plus élevée des probabilités d'activation des sous-ensembles est supérieure à un seuil γ on force la probabilité d'activation la plus élevée $\mu^{s_j}(k)$ du sous-ensemble s_j à 1 et les probabilités d'activation des autres sous-ensembles à 0. Si la plus élevée des probabilités d'activation des sous-ensembles est inférieure au seuil γ , on procède comme pour la méthode directe, c'est-à-dire la probabilité d'activation $\mu_i(k)$ d'un modèle sera le produit entre la probabilité d'activation $\mu^{s_j}(k)$ du sous-ensemble auquel appartient le modèle et la probabilité d'activation $\mu_l^{s_j}(k)$ de ce modèle dans son sous-ensemble. Les

différentes étapes s'explicitent :

Si l'on a j tel que $\mu^{s_j}(k) > \mu^{s_i}(k) \quad \forall i = 1, \dots, n$

alors si $\mu^{s_j}(k) > \gamma$ on impose

$$\mu^{s_j} = 1 \text{ et } \mu^{s_i} = 0, \quad \forall i = 1, \dots, n$$

$$\text{puis } \mu_i(k) = \mu^{s_j}(k)\mu_l^{s_j}(k), \quad M_i = M_l^{s_j}$$

sinon on calcule directement $\mu_i(k) = \mu^{s_j}(k)\mu_l^{s_j}(k), \quad M_i = M_l^{s_j}$

6.5 Exemple

On considère le fonctionnement d'un système décrit par un ensemble de modèles $M = \{M_1, M_2, M_3, M_4\}$. On choisit de regrouper les modèles deux à deux $S = \{s_1, s_2\}$ avec $s_1 = \{M_1, M_2\}$ et $s_2 = \{M_3, M_4\}$. Les quatre modèles M_1, M_2, M_3, M_4 s'explicitent :

$$\begin{aligned} x(k) &= \begin{bmatrix} 0.5 & 0.1 \\ 0 & 0.6 \end{bmatrix} x(k-1) + \begin{bmatrix} 1 \\ 1 \end{bmatrix} u(k-1) + w(k-1) \\ z(k) &= \begin{bmatrix} 1 & 1 \end{bmatrix} x(k) + v(k) \end{aligned}$$

$$\begin{aligned} x(k) &= \begin{bmatrix} 0.5 & 0.1 \\ 0 & 0.6 \end{bmatrix} x(k-1) + \begin{bmatrix} 1 \\ 1 \end{bmatrix} u(k-1) + w(k-1) \\ z(k) &= \begin{bmatrix} 1.2 & 1 \end{bmatrix} x(k) + v(k) \end{aligned}$$

$$\begin{aligned} x(k) &= \begin{bmatrix} 0.5 & 0.1 \\ 0 & 0.6 \end{bmatrix} x(k-1) + \begin{bmatrix} 1 \\ 1.5 \end{bmatrix} u(k-1) + w(k-1) \\ z(k) &= \begin{bmatrix} 1 & 1 \end{bmatrix} x(k) + v(k) \end{aligned}$$

$$\begin{aligned} x(k) &= \begin{bmatrix} 0.5 & 1 \\ 0 & 0.6 \end{bmatrix} x(k-1) + \begin{bmatrix} 1 \\ 1.3 \end{bmatrix} u(k-1) + w(k-1) \\ z(k) &= \begin{bmatrix} 1 & 1 \end{bmatrix} x(k) + v(k) \end{aligned}$$

Les gains statiques des quatre modèles sont : $g_1 = 5.00, g_2 = 5.50, g_3 = 6.50, g_4 = 5.90$.

Un scénario similaire à celui du deuxième chapitre est envisagé pour comparer les performances de la méthode proposée (avec décision directe, hiérarchisée et mixte) avec

	BDI	FA	ND	IID	TMD/(pe)
Décision directe	0.6860	0.0001	0.2595	0.0172	8.1818
Décision hiérarchisée	0.7420	0.0001	0.1981	0.0430	6.2727
Décision mixte	0.7730	0.0034	0.1031	0.0004	6
Méthode GPB1	0.6610	0.0001	0.3265	0.0017	5.9091

TAB. 6.1 – Performance de la détection pour les méthodes proposées

celle de l’algorithme GPB1. Les résultats de cette comparaison sont présentés dans le tableau 6.1. Les commutations sont générées en utilisant une matrice de Markov Π donnée par :

$$\Pi = \begin{bmatrix} 0.9 & 0.033 & 0.033 & 0.033 \\ 0.033 & 0.9 & 0.33 & 0.033 \\ 0.033 & 0.033 & 0.9 & 0.033 \\ 0.033 & 0.033 & 0.033 & 0.9 \end{bmatrix}$$

Les variances des bruits sont données par :

$$Q = \begin{bmatrix} 0.001 & 0 \\ 0 & 0.001 \end{bmatrix}.$$

$$R = 0.5$$

La figure 6.3 représente les entrées-sorties du système. La figure 6.4 représente les probabilités d’activation de chaque sous-ensemble ainsi que les probabilités d’activation de chaque modèle calculées par la méthode directe ainsi que celles calculées par la méthode GPB1. La figure 6.5 représente les mêmes grandeurs que la figure 6.4, mais les probabilités d’activation des modèles sont calculées par la prise de décision hiérarchisée ainsi que par la méthode GPB1. La figure 6.6 représente les mêmes grandeurs que les figures 6.4 et 6.5, mais obtenues par la méthode mixte.

Les résultats consignés dans le tableau 6.1 montrent que les performances des méthodes de prise de décision proposées sont généralement meilleures que celles de la méthode GPB1 sauf pour le retard à la détection où la méthode GPB1 est légèrement plus performante

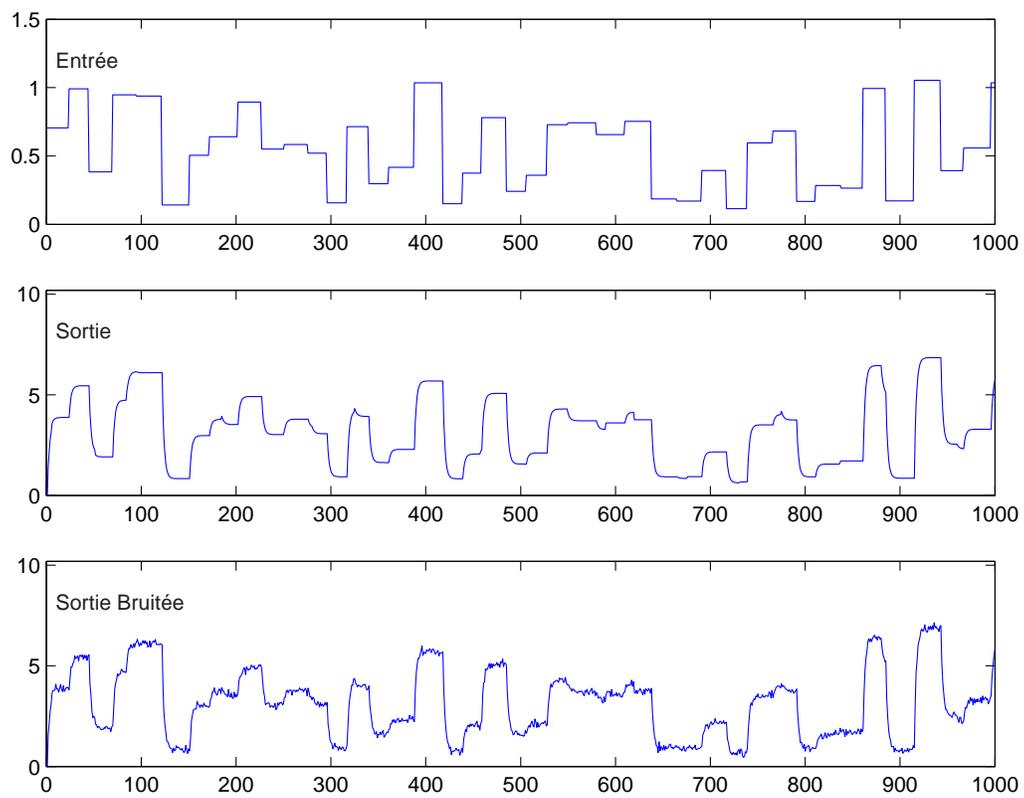


FIG. 6.3 – Entrée-sorties du système

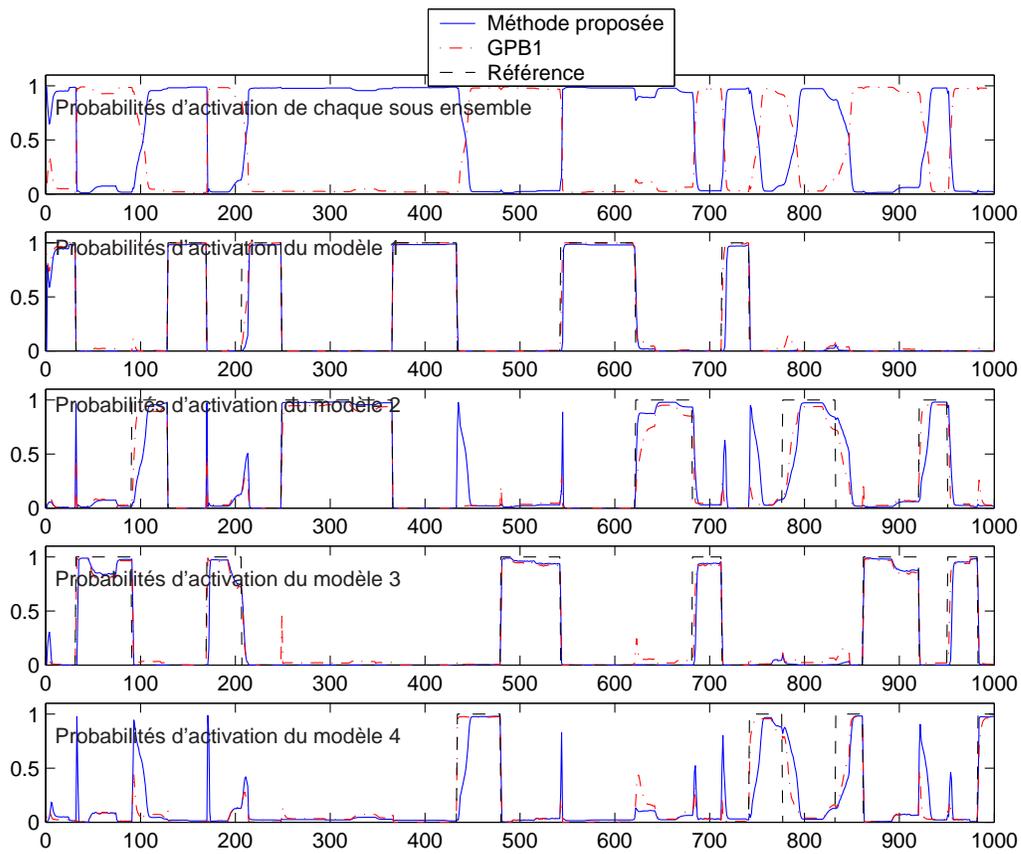


FIG. 6.4 – Probabilités d'activation (méthode directe)

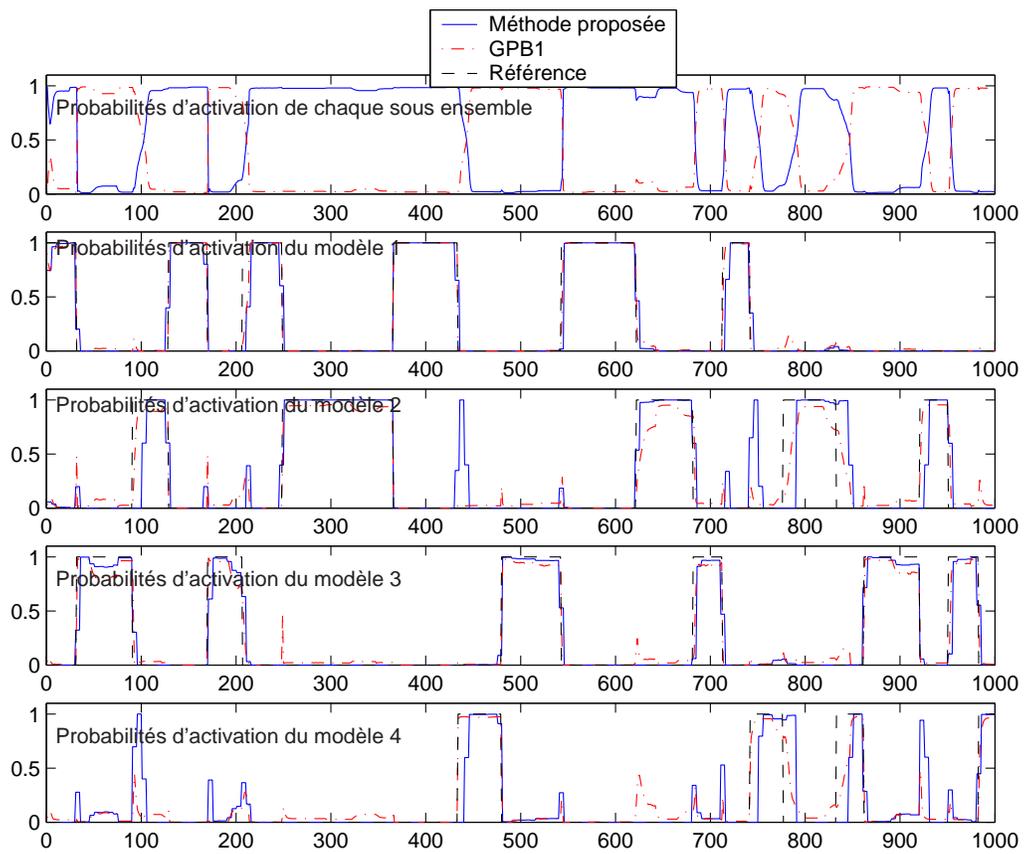


FIG. 6.5 – Probabilités d'activation (méthode hiérarchisée)

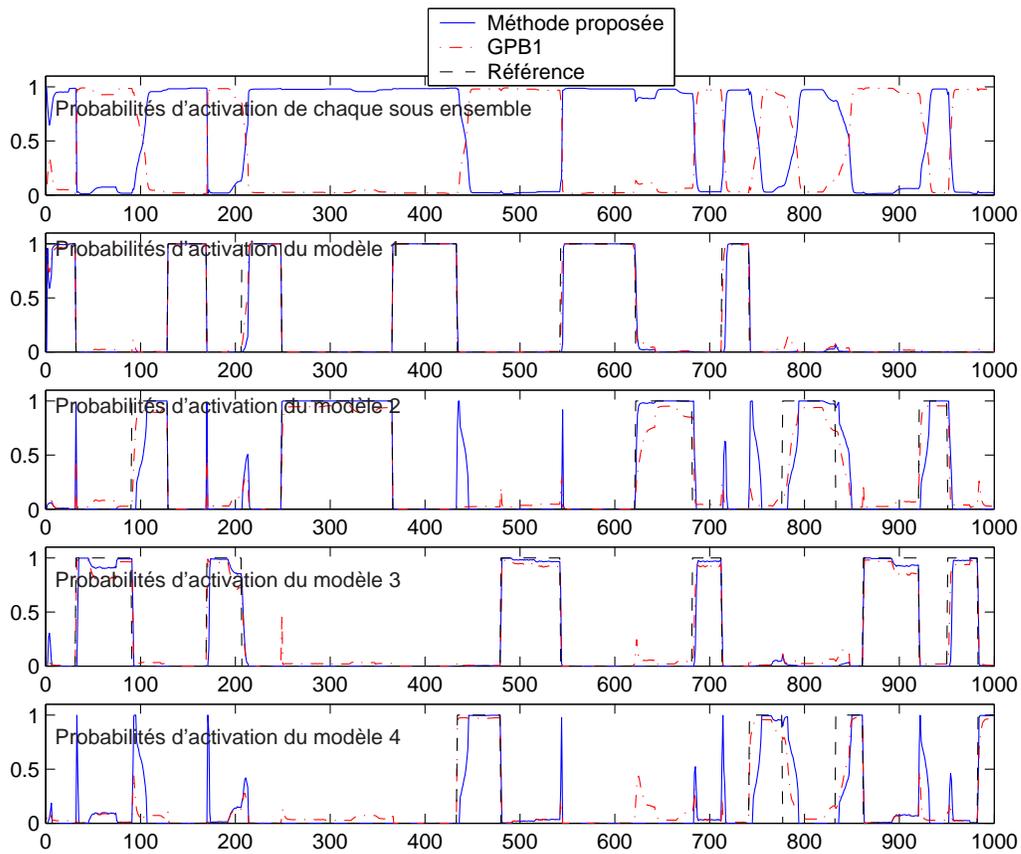


FIG. 6.6 – Probabilités d'activation méthode mixte

	BDI	FA	ND	IID	TMD/(pe)
$\{M_1, M_2\}, \{M_3, M_4\}$	0.7200	0.0038	0.1248	0.0071	5.4000
$\{M_1, M_3\}, \{M_2, M_4\}$	0.6830	0.0015	0.1333	0.0018	5.5556
$\{M_1, M_4\}, \{M_2, M_3\}$	0.7730	0.0034	0.1031	0.0004	5.6400

TAB. 6.2 – Performances de la détection en fonction des regroupements des modèles

que les méthodes proposées dans ce chapitre.

On remarque aussi qu'il y a des différences entre les différentes méthodes de prise de décision. La méthode qui donne les meilleurs résultats est la méthode mixte suivie par la méthode hiérarchisée puis la méthode directe. On a constaté sur d'autres simulations, réalisées avec une autre séquence de commutations, que les résultats, représentés sur le tableau 6.2, varient légèrement si l'on change la manière de regrouper les modèles dans les sous-ensembles, en utilisant pour chaque regroupement la méthode mixte. Cela montre une certaine influence du regroupement des modèles sur les performances du diagnostic.

6.6 Conclusion

Dans cette partie, on a introduit la notion de subdivision de l'ensemble de modèles pour les estimateurs multi-modèles bayésiens et cela en se distinguant de l'approche de X.R. Li. En effet, l'approche proposée considère l'ensemble de tous les modèles en utilisant des sous-estimateurs fonctionnant en parallèle, contrairement à la méthode de X.R. Li qui prend en compte uniquement le sous-ensemble sélectionné à chaque cycle d'exécution. Les résultats obtenus sur plusieurs exemples valident l'approche utilisée par une comparaison avec la méthode GPB.

La méthode structurée nous permet d'agir sur de nouveaux paramètres par rapport aux estimateurs multi-modèles classiques tel que la répartition des modèles dans les sous-ensembles ainsi que le nombre des sous-ensembles. Elle fournit aussi des informations supplémentaires comme la probabilité d'activation des sous-ensembles. La structure de la

méthode proposée permet, dans le cadre d'une procédure de diagnostic, de regrouper les modèles de dysfonctionnement du même type dans le même sous-ensemble et d'utiliser directement la probabilité du sous-ensemble pour détecter le défaut.

Conclusion et perspectives

Les travaux développés dans ce mémoire de thèse constituent une contribution à l'étude des méthodes de détection et de localisation de défauts par filtrage multi-modèle, les apparitions de défauts étant représentés ici par des changements de modèle. Ce type d'estimateurs permet simultanément d'estimer l'état du système et de calculer la probabilité des différents modes de fonctionnement. Le système est caractérisé par un ensemble de modèles de fonctionnements (fonctionnements normaux et anormaux) et par une matrice de transition de Markov caractérisant les passages d'un mode de fonctionnement à un autre mode de fonctionnement.

Tout au long de ce mémoire, notre étude a visé à améliorer les performances des méthodes d'estimation multi-modèle et à les appliquer au diagnostic des systèmes à commutation. Dans le premier chapitre, nous avons introduit l'estimation par multi-modèle sous ses différentes formes présentes dans la littérature spécialisée.

Le deuxième chapitre a été dédié à l'utilisation des méthodes d'estimation multi-modèle pour le diagnostic, où les systèmes en défaut ont été représentés par des modèles de fonctionnement spécifiques. Des règles de prise de décision sur le mode actif à chaque instant ont été définies, en se basant sur les probabilités d'activation des modèles, ainsi que des critères de performance du diagnostic. On a montré sur quelques exemples l'influence du bruit sur les performances de la détection de défauts et l'importance du rôle joué par la matrice de Markov, supposée connue dans un premier temps. Compte tenu de l'influence de cette matrice sur les performances des estimateurs multi-modèles, on a proposé, dans le troisième chapitre, d'estimer la matrice des transitions de Markov. Deux méthodes d'es-

timation ont été proposées, la méthode quasi-bayésienne et la méthode par intégration numérique. Ces méthodes ont été ensuite intégrées dans le contexte d'estimation multi-modèle où l'on estime simultanément l'état du système, les probabilités d'activation des modèles et la matrice de Markov.

Dans le quatrième chapitre, nous avons remplacé le filtre de Kalman, utilisé habituellement dans les méthodes GPB et IMM, par un observateur à mémoire finie qui est construit sur la base d'un critère quadratique qui maximise la vraisemblance du modèle du système. Par extension, cela nous a permis d'estimer les entrées inconnues et d'améliorer les performances du diagnostic.

Nous avons proposé dans le cinquième chapitre un observateur d'état mixte. Cet observateur est une combinaison de l'observateur à mémoire finie et de l'observateur de Luenberger ; la convergence de cet observateur a été étudiée à travers l'utilisation d'inégalités matricielles et d'une fonction de Lyapunov quadratique. Le paramètre de pondération entre les deux observateurs a été optimisé afin de minimiser les erreurs d'estimation. L'estimateur proposé a été intégré dans le cadre de l'estimation multi-modèle en substitution du filtre de Kalman classiquement utilisé. On a pu constater, sur des exemples, l'amélioration des performances du diagnostic.

Afin de minimiser le phénomène de dispersion des probabilités d'activation des modèles lorsque le nombre de modèles est important, on a ensuite élaboré une méthode d'estimation hiérarchisée. Pour cela, les modèles sont regroupés en différents sous-ensembles. Un sous-estimateur est construit en s'appuyant sur tous les modèles de chaque sous-ensemble. Ces sous-estimateurs constituent les filtres élémentaires d'un estimateur multi-modèle global. Cette approche permet une prise de décision à deux niveaux. On détermine d'abord le sous-ensemble actif, puis le modèle actif au sein de ce sous-ensemble.

Ces travaux de recherche ont mis en évidence certains problèmes pouvant faire l'objet de réflexions ultérieures :

- La qualité des résultats obtenus a essentiellement été constaté à travers l'analyse de résultats de simulation. Cependant, des études analytiques peuvent être envisagées dans le futur afin de quantifier l'apport des méthodes proposées par rapport aux méthodes de référence du premier chapitre.

-
- Il serait intéressant de mettre au point une méthode de choix systématique des sous-ensembles de modèles pour les estimateurs multi-modèle structuré du sixième chapitre. En effet, les performances des méthodes proposées reposent en partie sur la nature des regroupements effectués.
 - La forme de l'observateur mixte pourrait être exploitée afin d'intégrer des contraintes sur l'état du système dans l'élaboration de l'observateur.
 - Toutes les méthodes supposent que le système réel fonctionne selon l'un des modèles d'un ensemble prédéfini. En cela, les méthodes proposées s'inscrivent dans un contexte de reconnaissance supervisée. Il serait intéressant, en analysant les probabilités conjointement à l'erreur d'estimation, de détecter l'apparition d'un mode de fonctionnement inconnu non répertorié a priori et d'en effectuer, en ligne, l'identification paramétrique (fonctionnement en mode non supervisé).
 - On pourrait envisager d'étendre la stratégie de détermination du modèle actif employée par l'ensemble des méthodes exposées au cas des systèmes avec incertitudes où les paramètres des modèles de fonctionnement sont incertains est où il faudra utiliser des estimateurs qui tiendront compte de ces incertitudes.
 - Enfin, toutes les méthodes proposées n'ont été testées qu'en simulation. Il serait important d'évaluer leur pertinence sur des systèmes réels en commençant par des pilotes de laboratoire. Le très classique système à trois cuves, pour lequel les modèles sont assez bien maîtrisés, pourrait faire l'objet des premières expérimentations.

Bibliographie

- Ackerson, G. A. and K. Fu (1970). On state estimation in switching environments. *IEEE Transactions on Automatic Control* 15(1), 10–17.
- Afnor (1994). Maintenance - concepts et définitions des activités de maintenance, norme NF x 60-010. *Association Française de Normalisation*, 28.
- Akashi, H. and H. Kumamoto (1977). Random sampling approach to state estimation in switching environments. *Automatica* 13(4), 429–433.
- Alessandri, A. (2000). Design of sliding-mode observers and filters for nonlinear dynamic systems. In *39th IEEE Conference on Decision and Control*, Sydney, NSW, Australie.
- Alessandri, A., M. Baglietto, and G. Battistelli (2003). Receding-horizon estimation for discrete-time linear systems. *IEEE Transactions on Automatic Control* 48(3), 473–478.
- Alessandri, A., M. Baglietto, and G. Battistelli (2005). Receding-horizon estimation for switching discrete-time linear systems. *IEEE Transactions on Automatic Control* 50(11), 1736–1748.
- Anderson, B. and J. Moore (1979). *Optimal filtering*. Prentice Hall.
- Bar-Shalom, Y. (1990). *Multitarget-multisensor tracking : advanced applications*. Artech House.

- Bar-Shalom, Y., K. C. Chang, and H. A. Blom (1989). Tracking a maneuvering target using input estimation versus the interacting multiple model algorithm. *IEEE Transactions on Aerospace and Electronic Systems* 25(2), 296–300.
- Bar-Shalom, Y. and X. Li (1993). *Estimation and tracking : principles, technique and software*. Artech House.
- Beard, R. V. (1971). Failure accommodation in linear systems through self-reorganisation. *Rept. MTV-71-1 Man Vehicle Laboratory, MIT, Cambridge 199*, 1–20.
- Blackman, S. S. and R. F. Popoli (1999). *Design and analysis of modern tracking systems*. Artech House.
- Blair, W. D. and G. A. Watson (1992). Interacting multiple model algorithm with aperiodic data. In *SPIE Symposium on Acquisition, Tracking and Pointing*, Orlando, FL, USA.
- Blom, H. and Y. Bar-Shalom (1988). The interacting multiple model algorithm for systems with markovian switching coefficients. *IEEE Transactions on Automatic Control* 33(8), 780–783.
- Bloomer, L. and J. E. Gray (2002). Are more models better? : The effect of the model transition matrix on the IMM filter. In *34th Southeastern Symposium on System Theory (SSST)*, Huntsville, Alabama, USA.
- Bousghiri, S. (1994). *Diagnostic de fonctionnement de procédés continus par reconciliation d'état généralisé. Application à la détection de pannes de capteurs et d'actionneurs*. Thèse, Université Henri Poincaré - Nancy I.
- Boyd, S., L. E. Ghaoui, E. Feron, and V. Balakrishnan (1994). *Linear matrix inequalities in system and control theory*. Philadelphia : SIAM.
- Brown, R. (1983). *Introduction to random signals and Kalman filtering*. New York : Wiley.

Burnette, D. J. (2001). *Using GPS as reference system to hit a moving target*. Thèse, Air Force Institute of Technology, Wright Patterson AFB, Ohio, USA.

Busch, M. and S. Blackman (1995). Evaluation of IMM filtering for an air defence system application. In *1995 SPIE Conference on Signal and Data Processing of Small Targets*, San Jose, CA, USA.

Caputi, M. (1995). A necessary condition for effective performance of the multiple model adaptive estimator. *IEEE Transactions on Aerospace and Electronic Systems* 31(3), 1132–1139.

Chadli, M. (2002). *Stabilité et commande de systèmes décrits par multi-modèles*. Thèse, Institut National Polytechnique de Lorraine, Vandoeuvre-lès-Nancy.

Chang, C. and M. Athans (1978). State estimation for a discrete system with switching parameters. *IEEE Transactions on Aerospace and Electronic Systems* 14(3), 418–425.

Chen, G. and M. Harigae (2001). Using IMM adaptive estimator in GPS positioning. In *Proceedings of the 40th SICE Annual Conference*, Nagoya, Japon.

Cheney, W. and D. Kincaid (1994). *Numerical mathematics and computing*. Brooks/Cole Publishing Company.

Daafouz, J., P. Riedinger, and C. Iung (2002). Stability analysis and control synthesis for switched systems : A switched Lyapunov function approach. *IEEE Transactions on Automatic Control* 47(11), 1883–1887.

Darouach, M., M. Zasadzinski, and S. T. Xu (1994). Full-order observers for linear systems with unknown inputs. *IEEE Transactions on Automatic Control* 39(3), 606–609.

Ding, Z. and L. A. Hong (1998). A distributed IMM fusion algorithm for multi-platform tracking. *Signal Processing* 64(2), 167–176.

Drummond, O., X. Li, and C. He (1998). Comparison of various static multiple model estimation algorithms. *Signal and Data Processing of Small Targets* 3373, 510–527.

- Fang, T. H., J. W. Choi, and M. Lee (1999). Detection and diagnosis of sensor and actuator failures using perception net. In *Proceedings of the 38th SICE Annual Conference*, Osaka, Japon.
- Ferrari-Trecate, G., D. Mignone, and M. Morari (2002). Moving horizon estimation for hybrid systems. *IEEE Transactions on Automatic Control* 47(10), 1663–1676.
- Friedland, B. (1969). Treatment of bias in recursive filtering. *IEEE Transactions on Automatic Control* 14(4), 359–367.
- Gasso, K. (2000). *Identification des systèmes dynamiques non-linéaire : approche multi-modèle*. Thèse, Institut National Polytechnique de Lorraine, Vandoeuvre-lès-Nancy.
- Gasso, K., G. Mourot, and J. Ragot (2001). Structure identification in multiple model representation : elimination and merging of local models. In *Proceedings of the 40th IEEE Conference on Decision and Control*, Orlando, FL, USA.
- Ghahramani, Z. and G. Hinton (2000). Variational learning for switching state-space models. *Neural Computation* 12, 831–864.
- Goutsias, J. and J. M. Mendel (1988). Optimal simultaneous detection and estimation of filtered discrete semi-markov chains. *IEEE Transactions on Information Theory* 34(3), 296–300.
- Gustafson, P. M. and al. (1978). ECG/VCG rhythm diagnosis using statistical signal analysis, I : Identification of persistent rhythms. *IEEE Transactions on Biomedical Engineering* 25, 344–353.
- Hanlon, P. D. and P. S. Maybeck (1998). Interrelationship of single-filter and multiple-model adaptative algorithms. *IEEE Transactions on Aerospace and Electronic Systems* 34(3), 934–946.
- Hanlon, P. D. and P. S. Maybeck (2000). Characterization of Kalman filter residuals in the presence of mismodeling. *IEEE Transactions on Aerospace and Electronic Systems* 36(1), 114–131.

-
- Hashimoto, M. and H. Kawashima (2001). Sensor fault detection and identification in dead-reckoning system of mobile robot : interacting multiple model approach. In *International Conference on Intelligent Robots and Systems*, Maui, HI, USA.
- Hocine, A., M. Chadli, D. Maquin, and J. Ragot (2006). A discrete-time sliding window observer for markovian switching system. In *45th IEEE Conference on Decision and Control*, San Diego, CA, USA.
- Hocine, A., D. Maquin, and J. Ragot (2003). Utilisation d'une banque de modèles variable pour la détection de changement de régimes. In *Journées Doctorales d'Automatique*, Valenciennes, France.
- Hocine, A., D. Maquin, and J. Ragot (2004a). Finite memory observer for switching system : application to diagnosis. In *Workshop on Advanced Control and Diagnosis*, Karlsruhe, Allemagne.
- Hocine, A., D. Maquin, and J. Ragot (2004b). Observateur à mémoire finie pour les systèmes à commutations : application à la détection de défauts. *e-STA, Sciences et Technologies de l'Automatique, revue électronique de la SEE* 1(4).
- Hocine, A., D. Maquin, and J. Ragot (2004c). Observateur à mémoire finie pour les systèmes à commutations : application à la détection de défauts. In *Conférence Internationale Francophone d'Automatique, CIFA*, Douz, Tunisie.
- Hocine, A., D. Maquin, and J. Ragot (2005a). Finite memory observer for switching systems : application to diagnosis. In *16th IFAC World Congress*, Prague, République Tchèque.
- Hocine, A., D. Maquin, and J. Ragot (2005b). Identification de systèmes à commutation. In *Journées Doctorales, Modélisation, Analyse et Conduite des Systèmes dynamique, JDMACS*, Lyon, France.
- Hocine, A., D. Maquin, and J. Ragot (2005c). Identification de systèmes à commutation. *e-STA, Sciences et Technologies de l'Automatique, revue électronique de la SEE* 2(2).

- Ignagni, M. B. (1990). Separate-bias Kalman estimator with bias state noise. *IEEE Transactions on Automatic Control* 35(3), 338–341.
- Janyene, A. (1987). *Validation de données des systèmes dynamiques linéaires*. Thèse, Université Henri Poincaré, Nancy I.
- Jazwinski, A. H. (1968). Limited memory optimal filtering. *IEEE Transactions on Automatic Control* 13(5), 558–563.
- Jilkov, V. P. and X. R. Li (2004). Online bayesian estimation of transition probabilities for markovian jump systems. *IEEE Transactions on Signal Processing* 52(6), 1620–1630.
- Johansen, T. A. and A. B. Foss (1992). Nonlinear local model representation for adaptive systems. In *IEEE Conference on Intelligent Control and Instrumentation*, Marina, Singapour.
- Kim, H. J. (1998). A review of deadbeat observers and some new result. In *CISL Winter Workshop*, Sung-Woo, Corée.
- Kleer, J. D. and B. C. Williams (1987). Diagnosing multiple faults. *Artificial Intelligence* 32, 97–130.
- Kratz, F. (1991). *Utilisation des techniques de redondances matérielles et analytiques à la détection de pannes de capteurs. Application aux centrales nucléaires*. Thèse, Institut National Polytechnique de Lorraine, Vandoeuvre-lès-Nancy.
- Kwon, W. and P. Kim (2002). A receding horizon unbiased fir filter for discret-time state space models. *Automatica* 38(3), 545–551.
- Li, X. (1996). *Hybrid estimation techniques*. Control and Dynamic Systems : Advances in Theory and Applications. New York : Academic Press.
- Li, X. (1998). Optimal selection of estimate for multiple-model estimation with uncertain parameters. *IEEE Transactions on Aerospace and Electronic Systems* 34(2), 653–657.
- Li, X. (2002). Model-set design for multiple-model method. Part I. In *Proceedings of the Fifth International Conference on Information Fusion*, Annapolis, MD, USA.

-
- Li, X. and Y. Zhang (2000). Numerically robust implementation of multiple-model algorithms. *IEEE Transactions on Aerospace and Electronic Systems* 36(1), 266–278.
- Li, X., Z. Zhao, P. Zhang, and C. He (2002). Model-set design for multiple-model method. Part II. Examples. In *Proceedings of the Fifth International Conference on Information Fusion*, Annapolis, MD, USA.
- Li, X. R. (2000). Multiple-model estimation with variable structure. Part II model set adaptation. *IEEE Transactions on Automatic Control* 45(11), 2047–2060.
- Li, X. R. and Y. Bar-Shalom (1993). Design of an interacting multiple model algorithm for air traffic control tracking. *Special issue on Air Traffic Control. IEEE Transactions on Control Systems Technology* 1(3), 186–194.
- Li, X. R. and Y. Bar-Shalom (1996). Multiple-model estimation with variable structure. *IEEE Transactions on Automatic control* 41(4), 478–493.
- Li, X. R., Y. Zhang, and X. Zhi (1999a). Multiple-model estimation with variable structure. Part III : Model-group switching algorithm. *IEEE Transactions on Aerospace and Electronic Systems* 35(1), 225–241.
- Li, X. R., Y. Zhang, and X. Zhi (1999b). Multiple-model estimation with variable structure. Part IV : desing and evaluation of model-group switching algorithm. *IEEE Transactions on Aerospace and Electronic Systems* 35(1), 242–254.
- Li, X. R., Z. Zhao, and X. B. Li (2005). General model-set design methods for multiple-model approach. *IEEE Transactions on Automatic Control* 50(9), 1260–1276.
- Ljung, L. (1999). *System identification, Theory for the user*. Prentice-Hall.
- Magill, D. (1965). Optimal adaptive estimation of sampled stochastic process. *IEEE Transactions on Automatic Control* 10(4), 434–439.
- Maybeck, D. (1982). *Stochastic models, estimation and control*, Volume II, III. New York : Academic Press.

- Maybeck, P. S. and P. D. Halon (1995). Performance enhancement of a multiple model adaptive estimator. *IEEE Transactions on Aerospace and Electronic Systems* 31(4), 1240–1254.
- Medvedev, A. V. and H. T. Toivonen (1991). Investigation of a finite memory observer structure. Technical report, Process Control Laboratory, Dept. of Chemical Engineering, Abo Akademi.
- Medvedev, A. V. and H. T. Toivonen (1994). Feedforward time-delay structures in state estimation : finite memory smoothing and continuous deadbeat observers. *IEE Proceedings in Control Theory and Application* 141(2), 121–129.
- Menke, T. E. and P. S. Maybeck (1995). Sensor/actuator failure detection in the vista f-16 by multiple model adaptive estimation. *IEEE Transactions on Aerospace and Electronic Systems* 31(4), 1218–1229.
- Murray-Smith, R. and T. A. Johansen (1997). *Multiple model approaches to modelling and control*. Taylor and Francis.
- Murthy, D. N. P. (1980). Observability in discrete time system with faulty sensor. *Automatica* 16(6), 709–710.
- Napolitano, M. R. and R. L. Swaim (1992). Aircraft flight control reconfiguration. *C. T. Leondes, editor, Control and Dynamic Systems*, 155–228.
- Nguyen, H. T. and M. Sugeno (1998). *Fuzzy systems, modelling and control*. Kluwer academic publishers.
- Nuninger, W. (1997). *stratégie de diagnostic robuste a l'aide de la redondance analytique*. Thèse, Institut National Polytechnique de Lorraine, Vandoeuvre-lès-Nancy.
- Oliveira, M. C., J. Bernussou, and J. C. Geromel (1999). A new discrete-time robust stability condition. *System and Control Letters* 37(4), 261–265.
- Park, S. H., P. S. Kim, O. K. Kwon, and W. H. Kwon (2000). Estimation and detection of unknown inputs using optimal fir filter. *Automatica* 36(10), 1481–1488.

Patton, R., P. Frank, and R. Clark (1989). *Fault diagnosis in dynamic systems*. Prentice Hall.

Pavlovic, V., J. Rehg, T. J. Cham, and K. Murphy (1999). A dynamic bayesian network approach to figure tracking using learned dynamic models. In *IEEE International Conference on Computer and Vision*, Kerkyra, Grèce.

Rago, C., R. Prasanth, R. K. Mehra, and R. Fortenbaugh (1998). Failure detection and identification and fault tolerant control. using the IMM-KF with applications to the eagle-eye UAV. In *37th IEEE Conference on Decision and Control*, Tampa, Florida, USA.

Ragot, J., A. Hocine, and D. Maquin (2004). Parameter estimation of switching systems. In *International Conference on Computational Intelligence for Modelling, Control and Automation, CIMCA'2004*, Gold Coast, Australie.

Ragot, J., D. Maquin, and D. Sauter (1992). Data validation using orthogonal filters. *IEEE Proceedings - D : Control Theory and Applications* 139(1), 47–52.

Rao, C. V., J. B. Rawlings, and D. Q. Mayne (2003). Constrained state estimation for nonlinear discrete-time systems : stability and moving horizon approximations. *IEEE Transactions on Automatic Control* 48(2), 246–258.

Sawaragi, Y., T. Katayama, and S. Fujichige (1973). Adaptive estimation for a linear system with interrupted observation. *IEEE Transactions on Automatic control* 18(2), 152–154.

Sheldon, S. and P. Maybeck (1993). An optimizing design strategy for multiple model adaptive estimation and control. *IEEE Transactions on Automatic Control* 38(4), 651–654.

Smith, A. F. M. and U. E. Makov (1978). A quasi-bayesian sequential procedures for mixtures. *Journal of Royal Statistic Society* 40(1), 106–112.

- Sworder, D. and J. Boyd (1999). *Estimation problems in hybrid systems*. Cambridge University Press.
- Takagi, T. and M. Sugeno (1985). Fuzzy identification of systems and application to modelling and control. *IEEE Transactions on Systems Man and Cybernetics* 15, 116–132.
- Titterton, D. M., A. F. M. Smith, and U. E. Makov (1985). *Statistical analysis of finite mixture distributions*. New York : Wiley.
- Tudoroiu, N. and K. Khorasani (2005). Fault detection and diagnosis for satellite's attitude control system (ACS) using an interactive multiple model (IMM) approach. In *IEEE Conference on Control Applications*, Toronto, Ontario, Canada.
- Tugnait, J. and A. Haddad (1979a). A detection-estimation scheme for state estimation in switching environments. *Automatica* 15(4), 477–481.
- Tugnait, J. K. (1982). Adaptive estimation and identification for discrete systems with Markov jump parameters. *IEEE Transactions on Automatic control* 27(5), 1054–1065.
- Tugnait, J. K. and A. H. Haddad (1979b). State estimation under uncertain observations with unknown statistics. *IEEE Transactions on Automatic control* 24(2), 201–210.
- Tugnait, J. K. and A. H. Haddad (1980). Adaptive estimation in linear systems with unknown markovian noise statistics. *IEEE Transactions on Information Theory* 26(1), 66–78.
- Valcher, M. E. and J. C. Willems (1999). Dead beat observer synthesis. *Systems and Control Letters* 37(5), 285–292.
- West, M. and J. F. Harrison (1997). *Bayesian forecasting and dynamic models, 2nd ed.* Springer-Verlag.
- Zasadzinski, M. (1990). *Contribution à l'estimation de l'état des systèmes singuliers. Application à la validation de données des systèmes dynamiques linéaires*. Thèse, Université Henri Poincaré, Nancy I.

Zaytoon, J. (2001). *Systèmes dynamiques hybrides*. Hermes Science.

Zhang, Y. and J. Jiang (2001). Interacting active fault-tolerant control using IMM approach. *IEEE Transactions on Aerospace and Electronic Systems* 37(4), 1221–1235.

Zhang, Y. and X. R. Li (1998). Detection and diagnosis of sensor and actuator failures using IMM estimator. *IEEE Transactions on Aerospace and Electronic Systems* 34, 1293–1313.

Bibliographie

AUTORISATION DE SOUTENANCE DE THESE
DU DOCTORAT DE L'INSTITUT NATIONAL
POLYTECHNIQUE DE LORRAINE

o0o

VU LES RAPPORTS ETABLIS PAR :

Monsieur Michel VERGÉ, Professeur, ENSAM, Paris

Monsieur Dimitri LEFEBVRE, Professeur, Université Le Havre, Le Havre

Le Président de l'Institut National Polytechnique de Lorraine, autorise :

Monsieur HOCINE Abdelfettah

à soutenir devant un jury de l'INSTITUT NATIONAL POLYTECHNIQUE DE LORRAINE,
une thèse intitulée :

"Estimation d'état et diagnostic de systèmes à commutation par filtrage multi-modèle"

NANCY BRABOIS
2, AVENUE DE LA
FORET-DE-HAYE
BOITE POSTALE 3
F - 54501
VANDŒUVRE CEDEX

en vue de l'obtention du titre de :

DOCTEUR DE L'INSTITUT NATIONAL POLYTECHNIQUE DE LORRAINE

Spécialité : « **Automatique et traitement du signal** »

Fait à Vandoeuvre, le 21 novembre 2006

Le Président de l'I.N.P.L.,

L. SCHUFFENECKER



Résumé : Ce travail concerne la détection de défauts sur les systèmes sujets à des changements de mode de fonctionnement. Le système réel est modélisé par un système à commutation markovienne qui est représenté par un ensemble de modèles de fonctionnement (fonctionnements normaux et anormaux) et par une matrice de probabilité de transition de Markov qui contient les probabilités de passage d'un modèle de fonctionnement à un autre. Cette représentation offre un cadre idéal à l'application des méthodes d'estimation multi-modèle. L'intérêt d'utiliser ce type d'estimateurs réside dans le fait qu'en plus de l'estimation de l'état du système, les estimateurs multi-modèles procurent la probabilité d'occurrence ou d'activation de chaque modèle de fonctionnement. Ces probabilités peuvent alors être utilisées pour la détection de défaut. Dans ce travail, nous avons utilisé les spécificités de l'estimation multi-modèle afin de procéder à la détection et l'isolation des défauts qui peuvent affecter un système linéaire. Plusieurs améliorations et aménagements ont été apportés à ce type d'estimateurs dans le but d'augmenter les performances du diagnostic.

Mots-clés : Systèmes à commutation markovienne, Détection et localisation de défauts, Multi-modèle, Estimation d'état, Synthèse d'observateurs.

Abstract : In this thesis, a fault detection method is developed for switching dynamic systems. These systems are represented by several linear models, each of them being associated to a particular operating mode. The proposed method is based on mode probabilities with the aim of finding the system operating mode and estimating the state. The method also uses a priori knowledge information about the mode transition probabilities represented by a Markov chain. This kind of model offers an ideal framework to the application of the multiple model estimation methods. The interest to use this type of estimators lies in the fact that in addition to the state estimation, the multiple model estimators get the probability activation of each model. These probabilities can be used for fault detection purpose. However, several improvements were made to this type of estimators in order to increasing the performances of the diagnosis.

Key words : Markovian switching systems, Fault detection and isolation, Multiple model, State estimation, Observer synthesis.