Institut National
Polytechnique de Lorraine

Centre de Recherche
en Automatique de Nancy

École doctorale IAEM Lorraine
Département de Formation Doctorale en Automatique

Diagnostic des Systèmes à Changement de Régime de Fonctionnement

THÈSE

présentée et soutenue publiquement le 6 octobre 2006

pour l'obtention du

Doctorat de l'Institut National Polytechnique de Lorraine

(spécialité automatique et traitement du signal)

par

Elom Ayih DOMLAN

Composition du jury

Président :	M. DE MATHELIN	Professeur à l'Université Louis Pasteur
Rapporteurs :	V. COCQUEMPOT	Maître de Conférences HDR à l'Université de Lille 1
	J. ZAYTOON	Professeur à l'Université de Reims
Examinateurs :	M. DE MATHELIN	Professeur à l'Université Louis Pasteur
	L. DUGARD	Directeur de Recherche LAG – CNRS
	D. MAQUIN	Professeur à l'INPL
	J. RAGOT	Professeur à l'INPL

Centre de Recherche en Automatique de Nancy – UMR 7039 – Nancy-Université, CNRS
2, Avenue de la Forêt de Haye 54516 Vandœuvre-Lès-Nancy
Tél.+33 (0)3 83 59 59 59 Fax +33 (0)3 83 59 56 44

Mis en page avec la classe thloria.

Remerciements

Les travaux présentés dans ce mémoire ont été effectués au sein de l'équipe Sûreté de Fonctionnement et Diagnostic des systèmes (SURFDIAG) du Centre de Recherche en Automatique de Nancy (CRAN), sous la direction de Monsieur Didier MAQUIN, Professeur à l'Institut National Polytechnique de Lorraine, et la codirection de Monsieur José RAGOT, également Professeur à l'Institut National Polytechnique de Lorraine.

Je tiens à exprimer mes vifs et sincères remerciements à Messieurs les Professeurs Didier MAQUIN et José RAGOT pour leur constante disponibilité, leur aide et encouragements qu'ils m'ont toujours prodigués et aussi pour m'avoir fait bénéficier amplement de leur rigueur scientifique, de leurs critiques objectives et de leurs conseils avisés.

J'exprime ma sincère reconnaissance à Monsieur Janan ZAYTOON, Professeur à l'Université de Reims, ainsi qu'à Monsieur Vincent COCQUEMPOT, Maître de Conférences HDR à l'Université de Lille I, pour avoir accepté de juger ce travail en qualité de rapporteurs au sein du jury.

Mes remerciements vont également à l'endroit de Monsieur Michel DE MATHELIN, Professeur à l'Université Louis Pasteur, ainsi qu'à Monsieur Luc DUGARD, Directeur de Recherche LAG - CNRS, pour leur participation à ce jury de thèse.

Je remercie mes amis et collègues de laboratoire de l'équipe SURFDIAG du CRAN, pour l'ambiance conviviale qu'ils ont contribuée à entretenir, les bons moments passés en leur compagnie et leur sympathie.

Enfin, je ne pourrais terminer ces remerciements sans mentionner la plus formidable de toutes les secrétaires, Marjorie SCHWARTZ. Merci pour ta disponibilité et ton efficacité enrobée d'une couche permanente de jovialité.

À mon papa André,
À ma maman Thérèse,
À mes sœurs Justine et Edwige.

Table des matières

Acronymes

Références personnelles

Introduction générale

Chapitre 1
Les Systèmes à changement de régime de fonctionnement **15**

1.1 Les modèles de régression issus de l'économétrie 18
1.1.1 Les modèles à rupture de structures 19
1.1.1.1 Les modèles autorégressifs à seuils (TAR) 20
1.1.1.2 Les modèles MS-AR 22
1.1.2 Les modèles à changement progressif de régimes 25
1.1.2.1 Les modèles STAR 26
1.2 Les modèles utilisés en automatique 28
1.2.1 Les représentations sous forme d'espace d'état 31
1.2.1.1 Les modèles linéaires à sauts 32
1.2.1.2 Les modèles linéaires à sauts markoviens 33
1.2.1.3 Les modèles affines par morceaux 33
1.2.1.4 Les modèles MLD 35
1.2.1.5 Les modèles MMPS 37
1.2.1.6 Les modèles LC 38
1.2.1.7 Les modèles ELC 39
1.2.1.8 Équivalence des modèles PWA, MLD, MMPS, LC et ELC 40
1.2.2 Les représentations sous forme de régression 41

1.2.2.1 Les modèles PWARX 41
1.2.2.2 Les modèles HH 45
1.2.2.3 Représentation canonique de Chua 47
1.2.2.4 Les modèles Hammerstein/Wiener PWARX 48
1.3 Représentation d'état et représentation sous forme de régression . . 49
1.4 Identification des systèmes à changement de régime 53
1.5 Liens entre multi-modèles et systèmes à changement de régime . . . 57
1.6 Stabilité des systèmes à changement de régime 61

Chapitre 2
Diagnostic à base de modèle

2.1 Définitions et généralités 68
2.1.1 Termes généraux 68
2.1.2 Fonctions 69
2.2 La détection 70
2.3 La localisation 72
2.3.1 Résidus structurés 73
2.3.2 Résidus directionnels 74
2.4 Méthodes de génération des résidus 76
2.4.1 L'approche par espace de parité 76
2.4.1.1 Espace de parité – cas statique 76
2.4.1.2 Espace de parité – cas dynamique 78
2.4.2 L'approche par observateurs 80
2.4.3 L'approche par estimation paramétrique 83
2.4.3.1 Minimisation de l'erreur d'équation 84
2.4.3.2 Minimisation de l'erreur de sortie 86
2.5 Diagnostic des systèmes à commutation 86
2.6 Observabilité des systèmes à commutation 91
2.6.1 Généralités 91
2.6.2 Observabilité dans le cas où l'évolution des modes est connue 94
2.6.3 Observabilité dans le cas où l'évolution des modes est inconnue 96
2.6.3.1 Observabilité des modes 96
2.6.3.2 Observabilité de l'état 99

Chapitre 3
Recherche du Mode Actif

3.1 Position du problème . . . 104
3.2 Reconnaissance du mode actif . . . 105
3.2.1 Détection du chemin actif : première formulation . . . 105
3.2.2 Détection du chemin actif : seconde formulation . . . 108
3.2.3 Sur le nombre de chemins . . . 110
3.2.4 Raffinement de la procédure de détection du chemin actif en présence de bruit de mesure . . . 111
3.2.4.1 Cas d'un bruit borné . . . 111
3.3 Discernabilité des chemins . . . 114
3.3.1 Exemple d'un système statique à commutation . . . 115
3.3.2 Cas d'un système dynamique à commutation : absence de bruit de mesure . . . 120
3.3.3 Cas d'un système dynamique à commutation : présence d'un bruit de mesure borné . . . 124

Chapitre 4
Identification du mécanisme de commutation

4.1 Position du problème . . . 129
4.2 Recherche d'un hyperplan séparateur linéaire . . . 130
4.2.1 Recherche d'un domaine de solutions sous forme simple . . . 133
4.2.2 Recherche d'une valeur unique . . . 135
4.2.3 Classes non linéairement séparables . . . 139
4.3 Recherche d'un hyperplan séparateur linéaire par morceaux . . . 143
4.3.1 Approche *one-against-all* . . . 144
4.3.2 Approche *all-together* . . . 145

Chapitre 5
Etude de cas : exemple de simulation

5.1 Reconnaissance du mode actif . . . 150
5.1.1 Cas déterministe . . . 150
5.1.2 Estimation de l'état du système . . . 154
5.1.3 Influence d'un bruit . . . 156

5.1.4 Taille de l'horizon d'observation 161
5.2 Identification du mécanisme de commutation 164
5.2.1 Recherche d'un domaine de solutions de forme simple 165
5.2.2 Recherche d'une solution particulière 166

Conclusion générale et perspectives

Bibliographie **173**

Autorisation de soutenance

Acronymes

ARX	*AutoRegressive Exogenous*
ELC	*Extended Linear Complementary*
HH	*Hinging Hyperplane*
HHARX	*Hinging Hyperplane AutoRegressive Exogenous*
H-PWARX	*Hammerstein PieceWise Affine autoRegressive eXogenous*
JL	*Jump Linear*
JML	*Jump Markov Linear*
LC	*Linear Complementary*
MS-AR	*Markov Switching AutoRegressive*
MLD	*Mixed Logical Dynamical*
MMPS	*Max-Min-Plus-Scaling*
PLS	*Piecewise Linear System*
PWA	*PieceWise Affine*
PWARX	*PieceWise Affine autoRegressive eXogenous*
SAC	Système À Commutation
SETAR	*Self-Exciting AutoRegressive*
STAR	*Smooth Transition AutoRegressive*
STR	*Smooth Transition Regression*
TAR	*Threshold AutoRegressive*
W-PWARX	*Wiener PieceWise Affine AutoRegressive Exogenous*

Table des figures

1.1 Système à changement abrupt de régime 20
1.2 Système TAR 22
1.3 Modèle autorégressif à changement de régime markovien 26
1.4 Modèle STAR 28
1.5 Processus de montage de composants électroniques 31
1.6 Partionnement polyédrique en dimension deux 34
1.7 Représentation graphique des liens entres les classes de modèles PWA, MLD, MMPS, LC et ELC 40
1.8 Fonctions affines par morceaux $f_1(\cdot)$ et $f_2(\cdot)$ 43
1.9 Exemple de modèle PWARX 44
1.10 Evolution de $u(\cdot)$ et de $y(\cdot)$ en fonction du temps 44
1.11 Hyperplans $y = \varphi_k^T\theta^-$ et $y = \varphi_k^T\theta^+$ et la fonction HH correspondante $y = \max\{\varphi_k^T\theta^+, \varphi_k^T\theta^-\}$ 46
1.12 Exemple de modèle HH 47
1.13 Exemple de non-linéarités affines 48
1.14 Les modèles H-PWARX et W-PWARX 49
1.15 Evolution de l'entrée et de la sortie 57
1.16 Paramètre a et plan $(u(k), y(k))$ 57
1.17 Fonction coût 57
1.18 Structure d'un multimodèle à modèles locaux couplés 59
1.19 Structure d'un multimodèle à modèles locaux découplés 59
1.20 Approximation de la fonction $f(\cdot)$ par un multi-modèle 61

2.1 Structure de diagnostic 71
2.2 Résidus directionnels 75
2.3 Localisation à partir des résidus directionnels 75
2.4 Méthode de l'espace de parité 80
2.5 Génération de résidus à partir d'un observateur 82
2.6 Observateur à entrées inconnues (DOS) pour la détection de défauts d'actionneurs 82
2.7 Observateur à entrées inconnues (GOS) pour la détection de défauts d'actionneurs 83
2.8 Observateur à entrées inconnues (DOS) pour la détection de défauts de capteurs 83
2.9 Observateur à entrées inconnues (GOS) pour la détection de défauts de capteurs 84
2.10 Estimation paramétrique par minimisation de l'erreur d'équation . . 85
2.11 Estimation paramétrique par minimisation de l'erreur de sortie . . . 86
2.12 Commande par retour d'état 88

3.1 Résidus ambigus 114
3.2 Discernabilité dans le plan $\{u, y\}$ 119

4.1 Domaines admissibles pour $\bar{h}_{ij}$ et g_{ij} 132
4.2 Exemple de zone de commutation $\nu_k = \text{signe}\,(y_{k-1} - 5)$ 134
4.3 Séparation linéaire de deux classes dans le plan $(y(\cdot), u(\cdot))$ 136
4.4 Partitionnement incomplète de l'espace des régresseurs 138
4.5 Deux classes non linéairement séparables 140
4.6 Approche *one-against-all* 144
4.7 Séparateur linéaire par morceaux pour trois classes 145

5.1 Entrée, sortie, état, loi de commutation 151
5.2 Résidus associés aux chemins de longueur 2 153
5.3 Reconnaissance des modes 155
5.4 Résidus associés aux chemins $(1 \cdot 1)$, $(2 \cdot 2)$ et $(3 \cdot 3)$ 156
5.5 Etats réels et estimés 157

5.6 Entrée, sortie, état et modes du système 158
5.7 Amplitude du bruit sur l'amplitude de la sortie en pourcentage . . . 158
5.8 Evolution des résidus 159
5.9 Reconnaissance du chemin actif 160
5.10 Etats réels et estimés 161
5.11 Estimation des bornes de l'état 162
5.12 Résidus intervalle de tous les chemins 163
5.13 Evolution du taux de détection en fonction de la taille de la fenêtre d'observation 163
5.14 Influence de la variance du bruit sur la taille optimale de l'horizon d'observation 164
5.15 Classes $\mathcal{C}_1$, $\mathcal{C}_2$ et $\mathcal{C}_3$ 165
5.16 Validation des paramètres identifiés 167
5.17 Recherche d'une solution particulière 168

Références personnelles

Ragot, J., D. Maquin et E. A. Domlan (2003). Switching time estimation of piecewise linear system. Application to diagnosis. In : *5th IFAC Symposium on Fault Detection, Supervision and Safety of Technical Processes, SAFEPROCESS'2003*. Washington, D.C., USA.

Domlan, E. A., J. Ragot et D. Maquin (2003). Switching time estimation of piecewise linear system. In : *Workshop on Advanced Control and Diagnosis, IARACC 2003*. Duisburg, Germany.

Domlan, E. A., D. Maquin et J. Ragot (2004). Diagnostic des systèmes à commutation, approche par la méthode de l'espace de parité. In : *Conference Internationale Francophone d'Automatique, CIFA'2004*. Douz, Tunisie.

Domlan, E. A., J. Ragot et D. Maquin (2004). Distinguishability of switching system for diagnosis. In : *Workshop on Advanced Control and Diagnosis, IARACC 2004*. Karlsruhe, Germany.

Domlan, E. A., J. Ragot et D. Maquin (2005). Diagnostic des systèmes à commutation : identification de la loi de commutation. In : *Journées Doctorales Modélisation, Analyse et Conduite des Systèmes dynamiques, JDMACS 2005*. Lyon, France.

Domlan, E. A., J. Ragot et D. Maquin (2006). Systèmes à commutation : diagnostic de fonctionnement et identification de la loi de commutation. In : *Conférence Internationale Francophone d'Automatique, CIFA'2006*. Bordeaux, France.

Domlan, E. A., J. Ragot et D. Maquin (2006). Systèmes à commutation : diagnostic de fonctionnement et identification de la loi de commutation. *e-STA, Sciences et Technologies de l'Automatique*.

Domlan, E. A., J. Ragot et D. Maquin (2006). Systèmes à commutation : recherche du mode actif, identification de la loi de commutation. *JESA, Journal*

Européen des Systèmes Automatisés.

Introduction générale

De façon générale, la modélisation de systèmes complexes conduit à l'obtention de modèles non-linéaires dont la dynamique est influencée à la fois par des évènements discrets et des dynamiques continues. Le principe de « simplicité » conduit, le plus souvent, à une démarche de modélisation s'appuyant sur l'utilisation d'un ensemble de modèles de structures simples (par exemple linéaires), chaque modèle décrivant le comportement du système dans une zone de fonctionnement particulière. Le comportement global du système est alors décrit à l'aide des modèles locaux et d'un mécanisme de gestion des transitions d'un modèle à un autre. En fonction de la forme des fonctions associées à ces transitions, on obtient différentes classes de modèles, notamment les « multi-modèles » dans le cas de transitions lisses (interpolation entre les modèles locaux) et les modèles dits « hybrides » dans le cas de transitions abruptes. Cette dernière classe de modèles, les modèles hybrides, constituent l'objet de cette thèse.

Les modèles hybrides caractérisent des processus physiques régis par des équations différentielles continues et des variables discrètes. Le processus modélisé est ainsi décrit par plusieurs modes de fonctionnement et le passage (commutation) d'un mode de fonctionnement à un autre est régi par l'évolution de variables internes du système ou par une loi externe non maîtrisée. Lorsque la représentation des modes de fonctionnement est faite à l'aide de modèles linéaires invariants, le modèle obtenu appartient à la classe des modèles linéaires par morceaux ou encore à celle des systèmes à commutation. Cette classe de modèles est largement utilisée, car les outils d'analyse et de commande pour les systèmes linéaires sont très développés et aussi parce qu'une grande partie des processus réels peut être représentée

par des modèles issus de cette classe.

Les recherches sur les systèmes à commutation sont essentiellement concentrées dans les domaines de l'identification, de la synthèse de loi de commande et de l'analyse de la stabilité. Un point crucial, nécessaire pour la mise en oeuvre de tous ces outils ou qui en simplifierait l'usage, est *la reconnaissance du mode de fonctionnement actif* à un instant particulier à partir de jeux de mesures. Elle est évidemment plus aisée si le mécanisme de commutation (ou loi de commutation) est maîtrisé. Ceci justifie l'importance de *l'identification du mécanisme de gestion des transitions* d'un mode de fonctionnement à un autre.

Cette thèse aborde les deux problèmes évoqués précédemment : la reconnaissance du mode de fonctionnement actif à un instant particulier à partir de jeux de mesures et l'identification du mécanisme de gestion des transitions d'un mode de fonctionnement à un autre. Les outils du diagnostic à base de modèle sont utilisés pour la résolution du premier problème, le deuxième étant traité à l'aide des méthodes de reconnaissance de forme et de classification.

Le premier chapitre aborde quelques concepts généraux sur les modèles connus sous la dénomination générique de modèles hybrides. Différents modèles servant à représenter des systèmes présentant plusieurs modes de fonctionnement seront présentés. Le second chapitre est consacré aux concepts fondamentaux du diagnostic de fonctionnement des systèmes à bases de modèle. Un lien sera établi entre le diagnostic et l'usage des modèles hybrides. Il sera également question de l'extension de la notion d'observabilité aux modèles hybrides. Les troisième et quatrième chapitres portent respectivement sur la reconnaissance du mode de fonctionnement actif et sur l'identification du mécanisme de commutation. Il s'agit dans un premier temps de proposer des méthodes permettant, à partir de la connaissance des signaux d'entrée et de sortie, de retrouver le mode de fonctionnement dans lequel se trouve le système, ceci sans savoir comment est géré le passage d'un mode de fonctionnement à un autre. Dans un second temps, on formulera une hypothèse sur la structure du mécanisme de commutation et l'objectif sera de déterminer, une fois la reconnaissance du mode actif effectué, les valeurs des paramètres du mécanisme de commutation. Un exemple académique sera mis en œuvre dans le dernier chapitre afin d'illustrer les méthodes proposées.

1

Les Systèmes à changement de régime de fonctionnement

Sommaire

1.1 Les modèles de régression issus de l'économétrie 18

1.1.1 Les modèles à rupture de structures 19

1.1.2 Les modèles à changement progressif de régimes 25

1.2 Les modèles utilisés en automatique 28

1.2.1 Les représentations sous forme d'espace d'état 31

1.2.2 Les représentations sous forme de régression 41

1.3 Représentation d'état et représentation sous forme de régression 49

1.4 Identification des systèmes à changement de régime . . 53

1.5 Liens entre multi-modèles et systèmes à changement de régime 57

1.6 Stabilité des systèmes à changement de régime 61

Introduction

L'économétrie est sans aucun doute le premier domaine où a été introduite la notion de changement de régime de fonctionnement ou de commutation au sein de modèles. Dans le passé, les fluctuations macro économiques ont été largement explorées en utilisant des séries temporelles linéaires. Très vite, l'insuffisance de ce type de modèle à appréhender la tendance très changeante de certaines variables économiques a été relevée. Pour une meilleure compréhension des phénomènes économiques, il fallait donc développer de nouveaux outils capables de prendre en compte les non-linéarités présentes dans les séries temporelles ainsi que le changement de mode de fonctionnement. Une des solutions qui fût apportée à ce problème a été d'envisager la possibilité que les paramètres d'une série temporelle puissent changer dans le temps, introduisant ainsi des changements paramétriques au sein du modèle. La série temporelle peut alors être scindée en plusieurs régimes avec des paramètres différents propres à chaque régime.

Quandt [1958] fut à l'origine des premiers principes de la représentation par des séries temporelles à changement de régime. S'en suivirent plusieurs travaux qui complétèrent les travaux de Quandt et posèrent un formalisme complet d'une telle représentation [Goldfeld et Quandt, 1973; Hamilton, 1989, 1990; Hudson, 1966]. De façon globale, on peut regrouper les modèles de séries temporelles à changement de régime en trois grands groupes en fonction de la nature du mécanisme conduisant au changement de régime :

- les modèles où le changement de régime est déterminé par une variable observée. On peut citer dans cette catégorie les modèles TAR (*Threshold Auto-Regressive models*) [Fair et Jafeee, 1972], les modèles SETAR (*Self-Exciting Threshold AutoRegressive models*) [Potter, 1995; Tong, 1983]
- les modèles où le changement de régime ne peut être directement observé. Dans ce cas, le mécanisme générateur des changements de régime est inconnu et est modélisé indépendamment des variables observées. On peut citer ici par exemple les modèles MS-AR (*Markov-Switching AutoRegression*) [Hamilton, 1989].
- enfin les modèles combinant les deux principes énoncés précédemment.

En automatique, l'idée de changement de régime fut reprise à des fins de modélisation, voire de commande. Ainsi, on parlera de systèmes hybrides, de systèmes à commutation, de systèmes affines par morceaux.
Partant de l'idée de changement de régime, on peut aller encore plus loin dans les capacités d'approximation de ces modèles en ne procédant non plus à des passages brusques d'un régime de fonctionnement à l'autre, mais plutôt à une véritable interpolation des modèles se traduisant par une transition douce d'un régime de fonctionnement à un autre. On peut imaginer aisément la force de ce type de modélisation qui permet de prendre en compte un grand nombre de dynamiques non-linéaires tout en autorisant, sous certaines conditions, la transposition de bon nombre de méthodes associées aux modèles linéaires [Gasso *et al.*, 2002; Johansen et Foss, 1993; Murray-Smith et Johansen, 1997].

1.1 Les modèles de régression issus de l'économétrie

Bien que l'importance de la notion de changement de régime soit depuis fort longtemps reconnue, il n'existe aucune théorie établie suggérant une approche unique afin de spécifier des modèles économétriques qui incluent des changements de régime. Le concept de base qui sous-tend la conception des modèles incluant des changements de régime est que le processus à modéliser est supposé être une fonction d'une variable s_k indicatrice du régime en cours à l'instant k ($s_k = i$ transcrit le fait qu'à l'instant k le système opère dans le $i^{\text{ème}}$ régime de fonctionnement). Le modèle caractérise ainsi un processus non linéaire générant des données linéaires par morceaux. Chaque régime de fonctionnement est donc naturellement représenté par un modèle linéaire.
Il existe une panoplie assez vaste de modèles qui entrent dans la classe des modèles à changement de régime. D'un modèle à l'autre, la différence provient des hypothèses formulées sur le mécanisme générateur du changement de régime. Les changements de régime peuvent, dans certaines situations, être considérés comme des évènements déterministes et dans d'autres, on peut envisager qu'ils sont régis

par un processus stochastique exogène.
Sans avoir la prétention de couvrir l'ensemble des modèles existants, nous présentons ici quelques modèles couramment utilisés pour prendre en compte les changements de régime. Dans tout ce qui suivra, la variable k représentera le temps et nous ferons indifféremment usage de l'écriture compacte x_k ou celle plus longue $x(k)$ pour indiquer la dépendance par rapport au temps de la variable $x(\cdot)$.

1.1.1 Les modèles à rupture de structures

Connu dans la littérature anglo-saxonne sous le nom de *Structural Break Model*, ce type de modèle est caractérisé par un changement de structure à un instant $k = \tau$. Le modèle s'explicite comme suit :

$$y_k = \begin{cases} \nu_1 + \sum_{i=1}^{p} \alpha_{1,i} y_{k-i} + \varepsilon_k & \text{si } k < \tau \\ \nu_2 + \sum_{i=1}^{p} \alpha_{2,i} y_{k-i} + \varepsilon_k & \text{si } k \geq \tau \end{cases} \tag{1.1}$$

où ν_1 et ν_2 sont des constantes et $\varepsilon_{(\cdot)}$ est un terme d'erreur ou de perturbation supposé être indépendant et identiquement distribué (IID), de moyenne nulle et de variance σ^2 : $\varepsilon_k \sim \text{IID}(0, \sigma^2)$. y_k est la k-ième observation de la variable aléatoire $y_{(\cdot)}$. En introduisant la fonction indicatrice $I(k, \tau)$:

$$I(k, \tau) = \begin{cases} 1 & \text{si } k < \tau \\ 0 & \text{si } k \geq \tau \end{cases} \tag{1.2}$$

le modèle (1.1) peut être réécrit sous la forme compacte :

$$y_k = \left(\nu_1 + \sum_{i=1}^{p} \alpha_{1,i} y_{k-i} \right) (1 - I(k, \tau)) + \left(\nu_2 + \sum_{i=1}^{p} \alpha_{2,i} y_{k-i} \right) I(k, \tau) + \varepsilon_k \tag{1.3}$$

Lorsque l'instant de changement de régime τ est connu, on parle de rupture déterministe. Dans le cas contraire on parle de rupture stochastique.

Exemple 1.1 (Structural Break Model) *La figure 1.1 illustre le modèle de l'équation (1.3) pour les valeurs suivantes :*

$$\begin{cases} p = 2 \\ \nu_1 = -0.17 \quad \alpha_{1,1} = 0.34 \quad \alpha_{1,2} = -0.21 \\ \nu_2 = 0.30 \quad \alpha_{2,1} = 0.25 \quad \alpha_{2,2} = -0.11 \end{cases}$$

Le changement de régime est intervenu à l'instant $\tau = 50$ et est indiqué sur la

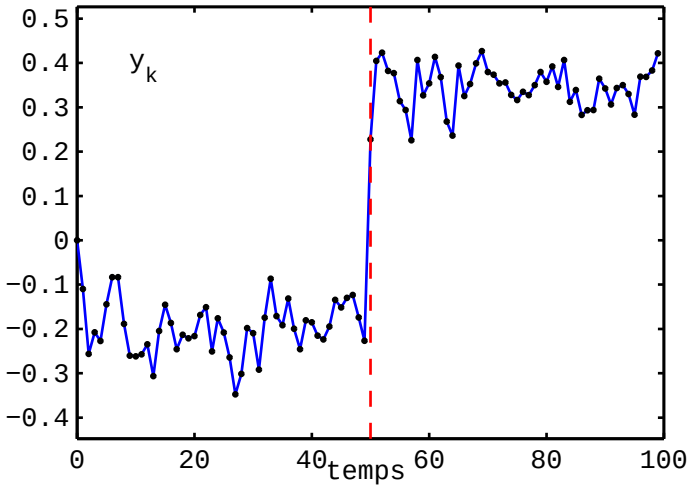


Fig. 1.1: Système à changement abrupt de régime

figure par la ligne verticale discontinue. Dans cet exemple précis, le changement de régime est aisément détectable en analysant l'évolution des données représentées. Cette évolution est caractérisée par un saut à l'instant de changement de régime. Toutefois, on peut imaginer que, dans certaines situations, ce changement peut ne pas être facile à détecter selon l'allure des différents régimes. Ceci pose déjà la question de la distinction entre une légère variation paramétrique d'un système et un vrai changement de régime.

1.1.1.1 Les modèles autorégressifs à seuils (TAR)

Les modèles autorégressifs à seuil ou modèle TAR (*Threshold AutoRegressive*) exhibent un comportement incorporant des changements de régime liés au fran-

chissement d'un seuil c par une variable de transition exogène observée $x(\cdot)$:

$$y_k = \left(\nu_1 + \sum_{i=1}^{p} \alpha_{1,i} y_{k-i}\right)(1 - I(x_k, c)) + \left(\nu_2 + \sum_{i=1}^{p} \alpha_{2,i} y_{k-i}\right) I(x_k, c) + \varepsilon_k \quad (1.4)$$

où $\varepsilon_k \sim \text{IID}(0, \sigma^2)$. La fonction indicatrice $I(x_k, c)$ est définie par :

$$I(x_k, c) = \begin{cases} 1 & \text{si } x(k-1) > c \\ 0 & \text{si } x(k-1) \leq c \end{cases} \quad (1.5)$$

L'idée sous-jacente dans la modélisation TAR est d'appréhender le caractère non linéaire d'une régression en ayant recours à un modèle linéaire par morceaux. Chaque « morceau » correspond à un régime auquel est étiqueté un modèle autorégressif linéaire. Un seul régime est actif à chaque instant k.

Dans le cas où la variable $x(\cdot)$ qui pilote le changement de régime est identique à la variable temporelle k ($x(k) = k$), le modèle TAR obtenu présente une rupture structurelle à l'instant $k = c$.

Une variante de modèles TAR est définie en indexant le changement de régime sur la variable observée $y(\cdot)$ retardée de d pas d'échantillonnage. Le modèle résultant est alors :

$$y_k = \left(\nu_1 + \sum_{i=1}^{p} \alpha_{1,i} y_{k-i}\right)(1 - I(y_{k-d}, c)) + \left(\nu_2 + \sum_{i=1}^{p} \alpha_{2,i} y_{k-i}\right) I(y_{k-d}, c) + \varepsilon_k \quad (1.6)$$

Le modèle (1.6) est connu sous le nom de modèle SETAR (*Self-Exciting AutoRegressive model*). Le paramètre c est la constante qui fixe le seuil à partir duquel intervient le changement de régime. La différence avec le modèle TAR réside dans le processus générateur du changement de régime qui n'est plus lié à une variable exogène, mais dépend directement de la variable de régression observée (d'où le nom de *Self-Exciting*). On tient compte ici des valeurs passées du processus afin de déterminer le régime à appliquer.

Exemple 1.2 (Modèle autorégressif à seuil TAR) *La figure 1.2 montre un exemple de modèle autorégressif à seuil. Les valeurs retenues pour les paramètres*

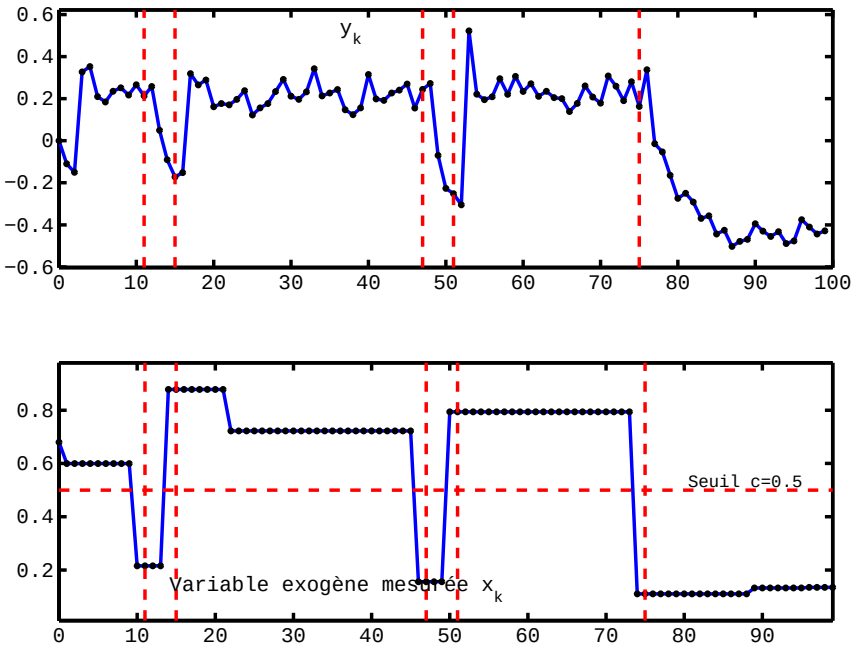


Fig. 1.2: Système TAR

du modèle (1.4) sont les mêmes que celles de l'exemple 1.1. Sur le deuxième graphique de la figure 1.2 est présentée l'évolution de la variable de transition $x(\cdot)$. Le seuil c définissant l'instant de changement de régime a été fixé à 0.5. Le premier graphique montre l'évolution de la sortie mesurée $y(\cdot)$ du processus. Les traits verticaux discontinus marquent les instants de changement de régime.

1.1.1.2 Les modèles MS-AR

Les modèles MS-AR (*Markov-Switching AutoRegressive*) introduisent une hypothèse probabiliste sur le passage d'un régime à un autre. L'évolution de la variable discrète s_k indicatrice du régime en cours est supposée dépendre d'une chaîne de Markov cachée (décrite par une variable non observée) à états finis, homogène et ergodique. Il convient de donner ici la définition d'une chaîne de Markov.

Chaîne de Markov

Généralement, un processus stochastique (un processus stochastique est une collection de variables aléatoires) est une suite d'expériences dont le résultat dépend du hasard. Pour décrire l'évolution temporelle d'un système dynamique, la méthodologie consiste à définir un espace d'état (l'espace dans lequel une variable

aléatoire prend ses valeurs) dans lequel se promène aléatoirement le système. En admettant qu'à chaque instant, le système peut se trouver dans l'un des états d'une collection finie d'états possibles, l'observation du système peut ainsi être considérée comme une expérience dont le résultat (aléatoire) est l'état dans lequel se trouve le système. La théorie des processus stochastiques permet alors de calculer les probabilités d'état stationnaires. Ces probabilités d'état peuvent être vues comme la probabilité que le système se trouve dans un état donné à un instant choisi « aléatoirement »loin dans le futur. Elles peuvent également être vues comme la proportion de temps que l'on a passé dans cet état au cours d'une très longue observation du système.

Définition 1.1 (Chaîne de Markov homogène)

Un processus stochastique à temps discret et à espace d'état discret $x(k)$, $k \in \mathbb{N}$, à valeurs dans E, est une chaîne de Markov homogène si et seulement si :

- *Propriété de Markov : pour tout $k \in \mathbb{N}$ et tout $k+2$-uplet $(j_0, j_1, \ldots, j_{k+1})$ d'éléments de E, on a*

$$Pr[x(k+1) = j_{k+1} | x(k) = j_k, \ldots, x(0) = j_0] = Pr[x(k+1) = j_{k+1} | x(k) = j_k] \tag{1.7}$$

- *Homogénéité : pour tout $k \in \mathbb{N}$ et toute paire (i, j) de E, on a*

$$Pr[x(k+1) = j | x(k) = i] = Pr(i, j) \tag{1.8}$$

  indépendamment de k.

En d'autres termes, une chaîne de Markov possède la propriété que son évolution (passage de $x(k)$ à $x(k+1)$) ne dépend que de l'état courant ($x(k) = j_k$) et pas de son passé. Les nombres $\Pr(i, j)$ sont les probabilités de transition de la chaîne. Ainsi, $\Pr(i, j)$ est la probabilité d'aller à l'étape j sachant qu'on se trouve à l'étape i. La matrice P dont les éléments sont donnés par les $\Pr(i, j)$, pour $(i, j) \in E \times E$, est la matrice de transition de la chaîne de Markov. Cette chaîne est homogène si son évolution ne dépend pas de l'instant k, mais seulement des états concernés.
On introduit la probabilité de transition de passer de l'état i à l'état j en m étapes,

notée $\Pr^m(i,j)$:

$$\Pr^m(i,j) = \Pr[x(k+m) = j | x(k) = i] \ \ \forall k \in \mathbb{N} \tag{1.9}$$

Définition 1.2 (Irréductibilité)

On dit qu'une chaîne de Markov est irréductible si tout état est atteignable en un nombre fini d'étapes à partir de tout autre état :

$$\forall i, j \in E, \exists m > 1 \ / \ Pr^m(i,j) \neq 0 \tag{1.10}$$

Définition 1.3 (Périodicité)

Un état i est périodique si on ne peut y revenir qu'après un nombre d'étapes multiple de $k > 1$:

$$\forall k > 1 \ / \ Pr^m(i,i) = 0 \ \ pour \ m \ non \ multiple \ de \ k \tag{1.11}$$

La période d'une chaîne de Markov est le plus grand commun diviseur (PGCD) de la période de chacun de ses états. Une chaîne de Markov est dite périodique si sa période est supérieure à 1. Dans le cas contraire, elle est dite apériodique.

Définition 1.4 (Ergodicité)

Une chaîne de Markov qui est irréductible et apériodique est dite ergodique.

Exemple 1.3 (Marche aléatoire (random walk) sur $\mathbb{Z}$) *Une particule se déplace sur l'ensemble des entiers relatifs $\mathbb{Z}$. A l'instant 0, la particule se trouve à l'origine. A chaque unité de temps, la particule se déplace vers un des deux sites voisins, vers la droite avec probabilité p et vers la gauche avec probabilité $1-p$. Si on pose : x_k = la position de la particule après k déplacements, alors la suite $x = (x_k; k \geq 0)$ est une chaîne de Markov sur $\mathbb{Z}$, issue de l'origine, dont les éléments $P_{i,j}$ de la matrice de transition P sont définis par :*

$$P_{i,j} = \begin{cases} p & \text{si } j = i+1 \\ 1-p & \text{si } j = i-1 \\ 0 & \text{si } j \notin \{i-1, i+1\} \end{cases}$$

Cette chaîne de Markov s'appelle une marche aléatoire sur $\mathbb{Z}$.

Ayant défini ce qu'est une chaîne de Markov, la définition d'un modèle autorégressif à changement de régime markovien en est facilitée.
Un modèle autorégressif à changement de régime markovien (*Markov Switching AutoRegressive model*) est un processus bivarié $\{s_k, y_k\}$ dans lequel s_k, variable indicatrice des changements de régime, est modélisée par une chaîne de Markov à états finis, homogène et ergodique et l'évolution de la variable observée y_k suit un modèle autorégressif linéaire. Le modèle obtenu est le suivant :

$$\begin{aligned} y_k &= \begin{cases} \nu_1 + \sum_{i=1}^{p} \alpha_{1,i} y_i + \varepsilon_k, \quad \varepsilon_k \,|s_k \sim NID\,(0, \sigma_1) \quad \text{si } s_k = 1 \\ \vdots \\ \nu_M + \sum_{i=1}^{p} \alpha_{M,i} y_i + \varepsilon_k, \quad \varepsilon_k \,|s_k \sim NID\,(0, \sigma_M) \quad \text{si } s_k = M \end{cases} \\ &\Pr\left(s_k = j \,|s_{k-1} = i\right) = p_{ij}, \ \sum_{i=1}^{p} p_{ij} = 1 \quad \forall i, j \in \{1, \cdots, M\} \end{aligned} \tag{1.12}$$

où $s_k \sim NID\left(0, \sigma_{(.)}\right)$ signifie que s_k est un bruit blanc gaussien de moyenne nulle et de variance $\sigma_{(.)}$.

Exemple 1.4 (Modèle autorégressif à changement de régime markovien)

La figure 1.3 représente un système autorégressif à changement de régime markovien avec deux régimes. Le dernier graphique présente des réalisations du processus stochastique qui régit les changements de régime du système. Ce processus est modélisé par une chaîne de Markov, ce qui conduit à l'obtention du second graphique sur lequel est représenté le régime actif à chaque instant. Ce régime est indexé par la variable $s_{(.)}$ qui prend soit la valeur 1 (indiquant que c'est le régime 1 qui est actif) ou 2 (indiquant que c'est le régime 2 qui est actif). La sortie mesurée $y_{(.)}$ du système est tracée sur le premier graphique.

1.1.2 Les modèles à changement progressif de régimes

Ce type de modèle, introduit par Granger et Terasvirta [1993], gère les changements de régimes en attribuant des poids aux différents régimes. Le changement de

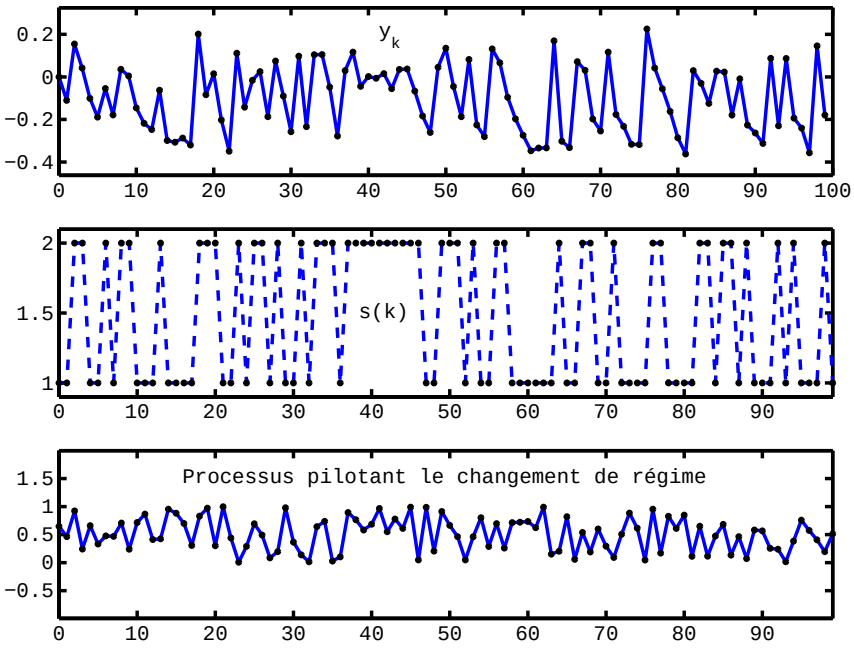


Fig. 1.3: Modèle autorégressif à changement de régime markovien

régime n'intervient plus de manière brusque, mais on passe plutôt progressivement d'un régime à un autre (ceci n'est pas sans rappeler l'approche multi-modèle qui sera présentée brièvement par la suite).

1.1.2.1 Les modèles STAR

Les modèles STAR (*Smooth Transition AutoRegressive*) sont largement utilisés en économétrie. Ils découlent d'une modification apportée aux modèles SETAR par Chan et Tong [1986] de façon à ce que le changement de régime ne se fasse pas brutalement, mais progressivement selon une fonction continuement dérivable (origine du terme Smooth). Ils s'explicitent par :

$$y_k = \left(\nu_1 + \sum_{i=1}^{p} \alpha_{1,i} y_{k-i}\right)(1 - G(z_k, \gamma, c)) + \left(\nu_2 + \sum_{i=1}^{p} \alpha_{2,i} y_{k-i}\right) G(z_k, \gamma, c) + \varepsilon_k \tag{1.13}$$

où $\varepsilon_k \sim \text{IID}(0, \sigma^2)$.

La fonction de transition $G(z_k, \gamma, c)$ est une fonction continuement dérivable définissant le poids attribué à chacun des régimes et est généralement à valeur comprise entre 0 et 1. La variable de transition z_k peut être aussi bien une variable endo-

gène (par exemple $z_k = y_{k-d}$, pour $d > 0$) qu'une variable exogène (par exemple $z_k = x_k$).

Pour $z_k = k$, on obtient un modèle à paramètres variant progressivement. La constante c est alors le seuil définissant le changement de régime et γ est un paramètre réglant la vitesse de passage d'un régime à l'autre.

Selon le choix de la fonction de transition $G(z_k, \gamma, c)$, on peut faire des liens entre les modèles STAR et les modèles à rupture de structures et même les modèles linéaires.

Pour $G(z_k, \gamma, c) = 1/(1 + \exp(-\gamma(z_k - c)))$, on remarquera que lorsque γ tend vers l'infini on a $G(z_k, \gamma, c) = I(z_k > c)$. On obtient ainsi un modèle SETAR. De même lorsque γ tend vers zéro, on a $G(z_k, \gamma, c) = 0.5$ et on obtient un modèle linéaire autorégressif.

En choisissant $G(z_k, \gamma, c) = 1 - \exp\left(-\gamma(z_k - c)^2\right)$, on a, lorsque γ tend vers l'infini ou vers zéro, $G(z_k, \gamma, c) = 0$. On obtient également un modèle linéaire autorégressif.

Enfin, pour $G(z_k, \gamma, c) = 1/(1 + \exp(-\gamma(z_k - c_1)(z_k - c_2)))$, le modèle STAR est équivalent à un modèle SETAR à trois régimes lorsque γ tend vers l'infini et à un modèle autorégressif linéaire lorsque γ tend vers zéro.

Exemple 1.5 (Modèle autorégressif à changement de régime markovien)

La figure 1.4 représente un modèle STAR combinant deux modèles autorégressifs dont les paramètres ont les mêmes valeurs que ceux de l'exemple 1.1. La fonction de transition $G(\cdot)$ est tracée sur le dernier graphique. Il a été choisi $G(z_k, \gamma, c) = 1/(1 + \exp(-\gamma(z_k - c)))$ avec $\gamma = 0.15$ et $c = 0.25$. Le premier graphique montre le résultat de l'interpolation des deux modèles autorégressifs de base par le biais de la fonction $G(\cdot)$.

Il existe d'autres extensions et modifications du modèle STAR, par exemple les modèles STR (*Smooth Transition Regression*) [Granger et Ramanathan, 1984].

Tous les modèles présentés existent également en version multivariable et sous forme de représentation d'état. Ce type de représentation sera mis en exergue par la suite lorsqu'il sera question des modèles couramment utilisés en Automatique.

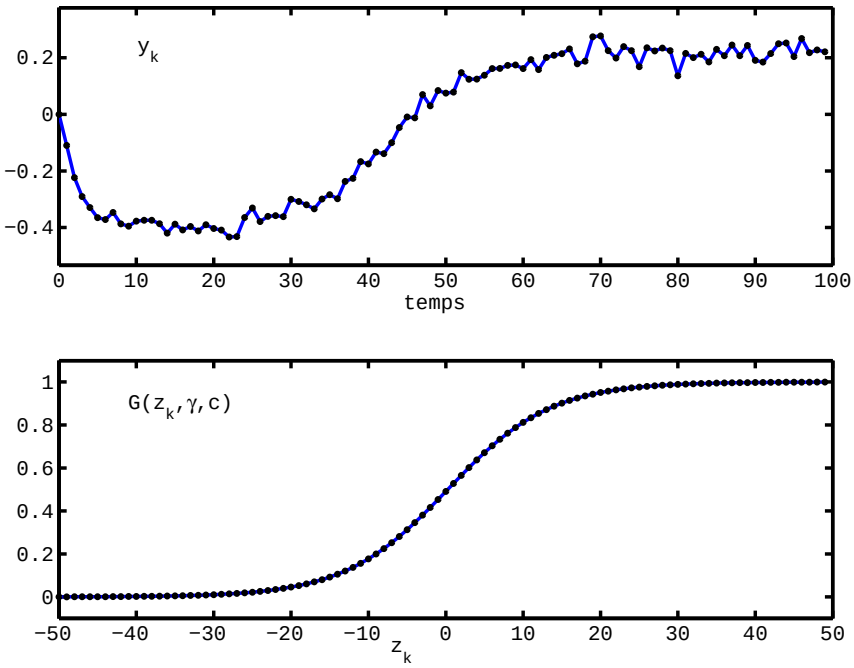


Fig. 1.4: Modèle STAR

1.2 Les modèles utilisés en automatique

En économétrie, l'objectif poursuivi en analysant des séries temporelles est de découvrir certaines régularités de comportement afin de pouvoir établir une prévision. Ce même objectif de modélisation se retrouve dans presque toutes les communautés scientifiques, plus particulièrement en automatique. En effet, pour que l'automatisation d'un système soit effective, il est indispensable d'avoir des informations quant au comportement et aux réactions de ce système en présence de diverses actions. Il vient alors naturellement le besoin indéniable de disposer d'un modèle pour décrire le comportement du système.

En Automatique, on définit de façon informelle un système comme étant un ensemble de composants interconnectés entre eux en vue de remplir une tâche bien précise. L'interaction entre les différents composants du système est réalisée à travers un certain nombre de canaux d'entrée et de sortie plus ou moins bien définis, canaux caractérisés par des flux d'énergie ou d'informations. L'ensemble des relations explicatives du fonctionnement du système constitue un modèle de fonctionnement.

Un modèle idéal doit posséder trois propriétés fondamentales : la simplicité, la pré-

cision et la généralité. La simplicité du modèle implique que ce dernier doit être aisé à comprendre et à interpréter, facile à utiliser et conduire à une implémentation simple quand il est utilisé à des fins de prédiction. La précision implique que le modèle décrit correctement le système et procure des prédictions exactes auxquelles on peut se fier. Enfin, la généralité du modèle lui confère la capacité de prendre en compte un nombre varié de situations différentes. Malheureusement, il existe un conflit inhérent entre ces trois propriétés idéales. Il est nécessaire le plus souvent d'effectuer un compromis car on dégrade généralement l'une des propriétés en voulant réaliser strictement une autre.

De la même façon qu'en économétrie où des modifications permanentes ou temporaires de la façon dont est générée une variable existent, on rencontre également en automatique des systèmes qui exhibent naturellement des comportements différents selon les conditions de fonctionnement. Par exemple un avion est modélisé par différentes équations dynamiques selon qu'il est en phase de décollage, d'atterrissage ou en plein vol.

De plus, la présence de régimes de fonctionnement multiples pour un même système répond aussi à une volonté de réduction de la complexité des modèles obtenus pour les systèmes réels. Au lieu d'avoir un modèle global qui couvre différentes situations et ainsi probablement plus complexe, une alternative est d'utiliser des modèles locaux pour chaque situation et une loi de passage d'un modèle à un autre. De cette façon, on obtient un modèle beaucoup plus simple à manipuler tout en gardant la capacité de fournir des prédictions exactes dans le domaine d'activation de chaque modèle local. Un problème complexe est ainsi partitionné d'une façon ou d'une autre en un certain nombre de sous-problèmes relativement plus simples qui peuvent être résolus indépendamment et dont les solutions individuelles conduisent globalement à la solution du problème complexe de départ.

Il existe une variété impressionnante de noms pour désigner en automatique des systèmes présentant des changements de régime. On citera par exemple les systèmes linéaires par morceaux, les systèmes affines par morceaux, les systèmes à commutation, les systèmes à sauts markoviens, etc... Il est devenu d'usage de regrouper toutes ces classes de modèles sous l'appellation générique de systèmes hybrides qui

englobe en réalité plusieurs classes de modèles.

Exemple 1.6 (Exemples de systèmes réels à changement de régime) *Un exemple fort simple de système à changement de régime est la diode. En effet, cette dernière est non passante lorsque la tension à ses bornes est inférieure à un seuil dit tension seuil et passante lorsque la tension appliquée dépasse ce seuil. On distingue ainsi deux régimes de fonctionnement distincts.*

Un autre exemple, extrait de la vie courante, est le chauffage d'une pièce. Pour contrôler la température dans une pièce, on dispose dans les maisons d'un thermostat qui enclenche ou arrête un radiateur selon que la température dans la pièce est en dessus ou au dessous d'un certain seuil fixé par l'utilisateur. De ce fait, on peut distinguer deux régimes de fonctionnement pour le processus selon que le radiateur est allumé ou éteint.

Un exemple, un peu plus complexe, est celui d'un processus de montage de composants électroniques [Juloski et al., 2004, 2005]. Le processus (voir figure 1.5) est constitué d'une tête de montage qui déplace un composant électronique et le place au dessus d'un circuit imprimé. La tête de montage applique ensuite une pression sur le composant jusqu'à ce que ce dernier soit en contact avec le circuit imprimé. Il est important de disposer d'un modèle adéquat pour ce genre de processus afin de pouvoir bien contrôler la course de la tête de montage de façon à ne pas endommager le composant électronique et le circuit imprimé lors du montage. Quatre modes de fonctionnement ont été mis en exergue pour ce système. Dans le premier mode, appelée mode libre, la tête de montage se déplace sans contrainte en n'étant pas en contact avec la surface d'impact sur le circuit imprimé. Le second mode est le mode d'impact. Dans ce cas, la tête de montage se déplace en étant en contact avec la surface d'impact sur le circuit imprimé. Les troisième et quatrième modes, dénommés respectivement mode de saturation haute et mode de saturation basse, correspondent aux situations où la tête de montage est incapable de se déplacer respectivement vers le haut ou le bas, ceci à cause de contraintes physiques sur le système. Juloski et al. [2004] ont modélisé ce processus en décrivant chaque mode du système par un modèle autorégressif.

Fig. 1.5: Processus de montage de composants électroniques

1.2.1 Les représentations sous forme d'espace d'état

Comme mentionné plus haut, les modèles associés aux systèmes à changement de régime peuvent se présenter, comme cela est le cas pour presque tout système, sous la forme d'une représentation d'état ou sous la forme d'une régression. Les modèles sous la forme d'une représentation d'état sont présentés ici. Suivront ensuite les modèles sous forme de régression. Toutes les descriptions à venir sont des modèles à temps discret. Par souci de simplicité des notations, les modèles considérés ne tiennent pas compte de la présence de bruit. Ce dernier sera introduit par la suite sous la forme d'un terme additif.

De façon générale, la modélisation d'un système présentant plusieurs régimes de fonctionnement sous la forme d'une représentation d'état nécessite de modéliser chacun des régimes du système par une représentation d'état. Ensuite, on se sert d'une procédure de gestion des changements de régimes afin de déterminer, à chaque instant, le régime actif et de basculer si nécessaire d'un régime à un autre.

Le modèle global associé au système est décrit par les équations suivantes :

$$\begin{cases} x(k+1) = A_{\mu_k} x(k) + B_{\mu_k} u(k) \\ y(k) = C_{\mu_k} x(k) \end{cases} \tag{1.14}$$

L'entrée, la sortie et l'état du système sont respectivement représentés par les variables $u(\cdot) \in \mathbb{R}^p$, $y(\cdot) \in \mathbb{R}^q$ et $x(\cdot) \in \mathbb{R}^n$. Les variables $u(\cdot)$, $y(\cdot)$ et $x(\cdot)$ sont relatives à la partie continue du système. La variable $\mu_{(\cdot)}$, souvent appelée loi ou fonction de commutation, indexe le modèle ou régime actif à chaque instant k et prend ses valeurs dans un ensemble fini $S = \{1, \ldots, s\}$, où s représente le nombre de régimes de fonctionnement du système. Les matrices A_i, B_i et C_i, $i = 1, \ldots, s$ sont des matrices réelles de dimensions appropriées qui décrivent l'évolution du système dans chacun de ses régimes de fonctionnement. Le modèle (1.14) peut être vu comme une famille de modèles linéaires reliés entre eux par des commutations indexées par la variable $\mu_{(\cdot)}$.

L'évolution de la loi de commutation peut être envisagée de diverses manières. En effet $\mu_{(\cdot)}$ peut être une fonction du temps k, de variables mesurées du système comme l'entrée $u(\cdot)$ et la sortie $y(\cdot)$, de variables à estimer comme l'état $x(\cdot)$ dans le cas où il serait inconnu, ou encore de variables exogènes dont la maîtrise fait défaut. A partir de la forme générale du modèle (1.14), diverses représentations des systèmes à changement de régime sont obtenues selon la description faite de la loi de commutation $\mu_{(\cdot)}$.

1.2.1.1 Les modèles linéaires à sauts

Les modèles linéaires à sauts ou modèles JL (*Jump Linear*) traduisent une situation extrême de connaissance sur l'évolution de la loi de commutation $\mu_{(\cdot)}$. En effet, $\mu_{(\cdot)}$ est considérée comme étant une entrée déterministe, mais inconnue, constante par morceaux. De ce fait, il n'existe aucune information *a priori* * sur l'évolution de la loi de commutation du système. Le système est modélisé par l'équation (1.14), sans que la forme de $\mu_{(\cdot)}$ ne soit explicitement donnée. La valeur

*Il serait plus correct de dire, pour être plus précis, qu'il existe très peu d'informations sur l'évolution de la loi de commutation car l'hypothèse selon laquelle $\mu_{(\cdot)}$ est constante par morceaux est déjà une information importante

prise par $\mu_{(\cdot)}$ à chaque instant est donc inconnue. Cependant, de façon indirecte, l'influence de $\mu_{(\cdot)}$ est visible sur la sortie mesurée $y(\cdot)$ du système.

1.2.1.2 Les modèles linéaires à sauts markoviens

Les modèles linéaires à sauts markoviens ou modèles JML (*Jump Markov Linear*) s'apparentent beaucoup aux modèles MS-AR présentés dans le cadre des modèles économétriques. Ils introduisent, comparativement aux modèles JL, une connaissance supplémentaire sur l'évolution de la loi de commutation $\mu_{(\cdot)}$. On modélise ici la loi de commutation μ_k par une chaîne de Markov homogène à états finis (s états) en temps discret avec des probabilités de transition p_{ij} définies par $p_{ij} = \Pr(\mu_{k+1} = j \,|\mu_k = i)$ pour tout $i, j \in S = \{0, 1, \ldots, s\}$. La matrice des probabilités de transition P (matrice de passage) est donc une matrice $s \times s$ dont les éléments satisfont $p_{ij} \geq 0$ et $\sum_{j=1}^{s} p_{ij} = 1$, pour tout $i \in S$. Le couple $(x_k; \mu_k)$ représente l'état du système à l'instant k. En particulier, x_k est l'état continu du système et μ_k représente l'état discret.
Les modèles JML ont de nombreuses applications dans divers domaines tels que la poursuite de cibles où ils servent à modéliser le comportement réel en poursuite [Evans et Evans, 1999; Krishnamurthy et Elliott, 1997]. En télécommunications, ils sont également utilisés pour modéliser des systèmes de télécommunication par étalement de spectres utilisant la modulation CDMA (*Code Division Multiple Access*) [Krishnamurthy *et al.*, 1999].

1.2.1.3 Les modèles affines par morceaux

Les modèles affines par morceaux ou modèles PWA (*PieceWise Affine*) [Bemporad *et al.*, 2000; Sontag, 1981, 1996] résultent d'un partitionnement de l'espace entrée/état du système en polyèdres et de l'affectation à chacun des polyèdres obtenus d'une fonction affine de mise à jour de l'état et d'une fonction de sortie définissant les mesures effectuées sur le système.

Définition 1.5 (Polyèdre)

Un ensemble convexe $\mathcal{P} \in \mathbb{R}^n$ défini par l'intersection d'un nombre fini s de demi-plans est appelé polyèdre. $\mathcal{P}$ s'explicite alors par $\mathcal{P} = \{x \in \mathbb{R}^n \,|Hx \leq G\}$, $H \in \mathbb{R}^{q\times n}$, $G \in \mathbb{R}^q$.

Exemple 1.7 (Exemple de partition polyédrique en dimension deux)

Sur la figure 1.6, les ensembles $\mathcal{P}_1$, $\mathcal{P}_2$ et $\mathcal{P}_3$ définissent une partition polyédrique dans un espace à deux dimensions.

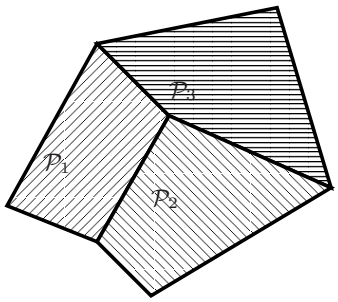


Fig. 1.6: Partionnement polyédrique en dimension deux

Un système représenté par un modèle PWA est décrit par :

$$\begin{cases} x\,(k+1) = A_{\mu_k} x\,(k) + B_{\mu_k} u\,(k) + b_{\mu_k} \\ y\,(k) = C_{\mu_k} x\,(k) + d_{\mu_k} \end{cases} \tag{1.15}$$

$$\mu_k = i \;\text{ si }\; (x_k; u_k) \in \Omega_i, \quad i = 1, \ldots, s \tag{1.16}$$

$$\Omega_i = \left\{(x; u) \in \mathbb{R}^n \times \mathbb{R}^p \,\middle|\, \bar{H}_i x + \bar{J}_i u \leq g_i \right\} \tag{1.17}$$

$$\bigcup_{i=1}^{s} \Omega_i = \Omega, \;\; \Omega \subseteq \mathbb{R}^n \times \mathbb{R}^p \tag{1.18}$$

$$\Omega_i \cap \Omega_j = \emptyset, \quad \text{pour } i \neq j \tag{1.19}$$

où $\bar{H}_i \in \mathbb{R}^{q_i\times n}$, $\bar{J}_i \in \mathbb{R}^{q_i\times p}$ et $g_i \in \mathbb{R}^{q_i}$, $\;i = 1, \ldots, s$.

L'équation (1.16) définit le mécanisme permettant le choix d'un régime particulier du système à un instant k. Ce choix est effectué en vérifiant l'appartenance du couple $(x_k; u_k)$ à l'un des polyèdres Ω_i. Les ensembles Ω_i forment une partition complète de l'espace état/entrée Ω. L'équation (1.19) impose aux ensembles Ω_i

d'être disjoints. Cette contrainte est nécessaire pour assurer l'unicité de la dynamique associée au système pour toutes les valeurs de l'état.
Dans (1.15), si les vecteurs $b_{\mu_{(\cdot)}}$ et $d_{\mu_{(\cdot)}}$ sont nuls pour toutes les valeurs de $\mu(\cdot)$, on parle alors de systèmes linéaires par morceaux ou PLS (*Piecewise Linear Systems*).

1.2.1.4 Les modèles MLD

Les modèles MLD (*Mixed Logic Dynamical*) se différencient un peu des formes de modélisation énumérées jusqu'à présent par la structure particulière de l'équation modélisant les différents régimes de fonctionnement du système. Le point fort des modèles MLD réside dans leur capacité à modéliser les parties logiques d'un processus (commutateurs « Marche/Arrêt », réseaux de mécanismes, réseaux de logiques combinatoires et séquentielles discrètes) et la représentation de la connaissance heuristique disponible sur le processus sous la forme d'inégalités linéaires de variables entières.
Dans un modèle MLD, la notion de régime de fonctionnement, comme explicitée par l'équation (1.14), n'apparaît plus vraiment. Introduite par Bemporad et Morari [1999], l'approche MLD permet d'inclure directement des contraintes dans le modèle et d'affecter des priorités aux contraintes. On peut séparer les contraintes en des contraintes fortes qui sont à respecter impérativement, et donc prioritaires, et en des contraintes faibles qui peuvent ne pas être respectées de façon passagère sans que cela n'affecte le fonctionnement du processus. La forme générale d'un modèle MLD est la suivante :

$$x(k+1) = Ax(k) + B_1 u(k) + B_2 \delta(k) + B_3 z(k) \tag{1.20a}$$

$$y(k) = Cx(k) + D_1 u(k) + D_2 \delta(k) + D_3 z(k) \tag{1.20b}$$

$$E_2 \delta(k) + E_3 z(k) \leq E_1 u(k) + E_4 x(k) + E_5 \tag{1.20c}$$

où $x(\cdot)$, $y(\cdot)$ et $u(\cdot)$ représentent respectivement l'état, la sortie et l'entrée du processus. Les variables $x(\cdot)$, $y(\cdot)$ et $u(\cdot)$ se décomposent en une partie binaire (partie

discrète du système) et en une partie continue (partie continue du système) :

$$x = \begin{bmatrix} x_c \\ x_l \end{bmatrix}, \; x_c \in \mathbb{R}^{n_c}, \; x_l \in \{0,1\}^{n_l} \tag{1.21}$$

$$y = \begin{bmatrix} y_c \\ y_l \end{bmatrix}, \; y_c \in \mathbb{R}^{q_c}, \; y_l \in \{0,1\}^{q_l} \tag{1.22}$$

$$u = \begin{bmatrix} u_c \\ u_l \end{bmatrix}, \; u_c \in \mathbb{R}^{p_c}, \; u_l \in \{0,1\}^{p_l} \tag{1.23}$$

La variable $\delta(\cdot)$ est une variable binaire auxiliaire : $\delta \in \{0,1\}^{r_l}$. $z(\cdot)$ est une variable continue auxiliaire : $z \in \mathbb{R}^{r_c}$. L'introduction de $\delta(\cdot)$ et $z(\cdot)$ vient de la conversion de propositions logiques en des inégalités linéaires. Toutes les contraintes du système sont regroupées dans (1.20c).

Exemple 1.8 (Modèle MLD) *Considérons le système PWA suivant :*

$$x(k+1) = \begin{cases} 0.3x(k) + u(k) & \text{si } x(k) \geq 0 \\ -0.8x(k) + u(k) & \text{si } x(k) < 0 \end{cases} \tag{1.24}$$

où $x(k) \in [-10, 10]$ *et* $u(k) \in [-1, 1]$.
Pour passer au modèle MLD, on associe à la variable binaire $\delta(k)$ *la condition* $x(k) \geq 0$ *de façon à ce que* $\delta(k) = 1 \Leftrightarrow x(k) \geq 0$. *On peut ainsi récrire l'équation (1.24) sous la forme suivante :*

$$x(k+1) = 1.1\delta(k)\,x(k) - 0.8x(k) + u(k) \tag{1.25}$$

Dans [Bemporad et Morari, 1999], il est démontré que pour une fonction $f(\cdot)$ *telle que* $f : \mathbb{R}^n \mapsto \mathbb{R}$, *le produit* $\delta f(x)$, δ *étant une variable binaire et* x *élément d'un*

domaine borné $\mathcal{H}$, est équivalent à :

$$\begin{cases} y \leq M\delta \\ y \geq m\delta \\ y \leq f(x) - m(1-\delta) \\ y \geq f(x) - M(1-\delta) \end{cases} \tag{1.26}$$

où y est une variable auxiliaire définie par $y = \delta f(x)$, $M = \max_{x\in\mathcal{H}} f(x)$, $m = \min_{x\in\mathcal{H}} f(x)$. En se servant de (1.26), on définit une variable auxiliaire $z(\cdot)$ de manière à ce que $z(k) = \delta(k)x(k)$ et on obtient les inégalités suivantes :

$$\begin{cases} z(k) \leq 10\delta(k) \\ z(k) \geq -10\delta(k) \\ z(k) \leq x(k) + 10(1-\delta(k)) \\ z(k) \geq x(k) - 10(1-\delta(k)) \end{cases} \tag{1.27}$$

Les inégalités (1.27) servent à respecter la contrainte $x(k) \in [-10, 10]$.
Finalement, le modèle MLD obtenu est :

$$\begin{aligned} x(k+1) &= 1.1z(k) - 0.8x(k) + u(k) \\ &\begin{cases} z(k) \leq 10\delta(k) \\ z(k) \geq -10\delta(k) \\ z(k) \leq x(k) + 10(1-\delta(k)) \\ z(k) \geq x(k) - 10(1-\delta(k)) \end{cases} \end{aligned} \tag{1.28}$$

1.2.1.5 Les modèles MMPS

Les modèles MMPS (*Max-Min-Plus-Scaling*) constituent une classe de modèles basée sur l'usage d'expressions MMPS, expressions qui combinent des opérateurs de maximisation, de minimisation, d'addition et de produit scalaire [De Schutter et Van Den Boom, 2001]. Par exemple l'expression $x-1-5x_2+\min(4x_1, \max(x_1, x_2))$ est une expression MMPS. Les modèles MMPS sont obtenus en utilisant des ex-

pressions MMPS dans le modèle du système :

$$\begin{cases} x(k+1) = \mathcal{M}_x(x(k), u(k), d(k)) \\ y(k) = \mathcal{M}_y(x(k), u(k), d(k)) \\ \mathcal{M}_c(x(k), u(k), d(k)) \leq c \end{cases} \tag{1.29}$$

où $\mathcal{M}_x$, $\mathcal{M}_y$ et $\mathcal{M}_c$ sont des expressions MMPS en $x(\cdot)$, $u(\cdot)$, $y(\cdot)$ et $d(\cdot)$ (variable auxiliaire à valeurs réelles).

Exemple 1.9 (Modèle MMPS) *On considère l'équation d'un intégrateur avec saturation :*

$$\begin{aligned} x(k+1) &= \begin{cases} x(k) + u(k) & \text{si } x(k) + u(k) \leq 1 \\ 1 & \text{si } x(k) + u(k) \geq 1 \end{cases} \\ y(k) &= x(k) \end{aligned} \tag{1.30}$$

L'équation (1.30) peut être récrite sous la forme MMPS :

$$\begin{cases} x(k+1) = \min(x(k) + u(k), 1) \\ y(k) = x(k) \end{cases} \tag{1.31}$$

De façon similaire, le modèle PWA de l'exemple 1.8, représenté par l'équation (1.24), peut être mis sous la forme d'un modèle MMPS :

$$x(k+1) = -0.8x(k) + 1.1\max(0, x(k)) + u(k) \tag{1.32}$$

1.2.1.6 Les modèles LC

Les modèles LC (*Linear Complementarity*) ont été étudiés dans [Heemels *et al.*, 2000; Van der Schaft et Schumacher, 1998]. Leur domaine d'application est essentiellement les systèmes mécaniques sous contraintes, les réseaux électriques avec des diodes idéales ou tout autre système dynamique présentant des relations linéaires

par morceaux, les systèmes à structures variables, les problèmes de commande sous contraintes et bien d'autres encore. Ces modèles sont décrits par :

$$\begin{cases} x(k+1) = Ax(k) + B_1 u(k) + B_2 w(k) \\ y(k) = Cx(k) + D_1 u(k) + D_2 w(k) \\ v(k) = E_1 x(k) + E_2 u(k) + E_3 w(k) + g_4 \\ v(k) \geq 0,\ w(k) \geq 0,\ v(k) \perp w(k) \end{cases} \tag{1.33}$$

où $x(\cdot) \in \mathbb{R}^n$, $u(\cdot) \in \mathbb{R}^p$ et $y(\cdot) \in \mathbb{R}^q$, $v(\cdot) \in \mathbb{R}^m$ et $w(\cdot) \in \mathbb{R}^m$. Le symbole $\perp$ indique l'orthogonalité des vecteurs : $v(k) \perp w(k) \Leftrightarrow (v(k))^T w(k) = 0$. Les variables $v(\cdot)$ et $w(\cdot)$ sont appelées variables complémentaires.

1.2.1.7 Les modèles ELC

Les modèles ELC (*Extended Linear Complementary*) sont décrits par [De Schutter, 2000; De Schutter et Van Den Boom, 2001] :

$$x(k+1) = Ax(k) + B_1 u(k) + B_2 d(k) \tag{1.34a}$$

$$y(k) = Cx(k) + D_1 u(k) + D_2 d(k) \tag{1.34b}$$

$$E_1 x(k) + E_2 u(k) + E_3 d(k) \leq g_4 \tag{1.34c}$$

$$\sum_{i=1}^{p} \prod_{j \in \phi_i} (g_4 - E_1 x(k) - E_2 u(k) - E_3 d(k))_j = 0 \tag{1.34d}$$

où $d(\cdot)$ est une variable auxiliaire. Les $\Phi_i, i \in \{1, 2, \ldots, p\}$ sont des sous-ensembles de $\{1, 2, \ldots, l\}$, l étant le nombre de lignes de la matrice E_1.
A cause de (1.34c), on a : $g_4 - E_1 x(k) - E_2 u(k) - E_3 d(k) = 0$. La condition (1.34d) est ainsi équivalente à :

$$\forall i \in \{1, 2, \ldots, p\}, \exists j \in \Phi_i \text{ tel que } (g_4 - E_1 x(k) - E_2 u(k) - E_3 d(k))_j = 0 \tag{1.35}$$

Les équations (1.34c) et (1.34d) peuvent donc être envisagées comme un système d'inégalités linéaires (les inégalités larges introduites par (1.34c)), dans lequel il y a p groupes d'inégalités linéaires (un groupe pour chaque ensemble d'indices Φ_i)), et dans chaque groupe au moins une des inégalités vérifie la condition d'égalité.

1.2.1.8 Équivalence des modèles PWA, MLD, MMPS, LC et ELC

Il a été montré dans [Heemels *et al.*, 2000] l'équivalence des classes de modèles cités précédemment. Étant donné deux classes de modèle $\mathcal{A}$ et $\mathcal{B}$, pour tout modèle élément de la classe $\mathcal{A}$, il existe un modèle appartenant à la classe $\mathcal{B}$ avec le même comportement entrée/sortie et vice versa. La représentation graphique de la figure 1.7 illustre le passage d'une classe à une autre. Pour plus de détails sur les conditions de passage d'une classe à l'autre, le lecteur peut se référer à [Bemporad *et al.*, 2000; Heemels *et al.*, 2000].

L'équivalence entre les différentes classes de modèles permet de choisir le modèle le plus adapté aux besoins de l'analyse à conduire et du système considéré. Toutefois, chaque classe possède ses propres propriétés :

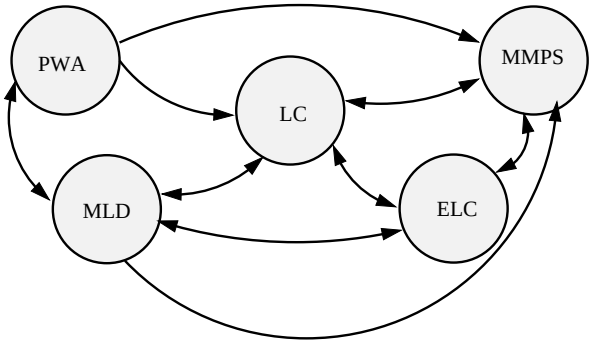


Fig. 1.7: Représentation graphique des liens entres les classes de modèles PWA, MLD, MMPS, LC et ELC

- existence de critères de stabilité pour les modèles PWA
- méthodes de vérification et de commande pour les modèles MLD
- méthodes de commande pour les modèles MMPS
- conditions d'existence et d'unicité des solutions pour les modèles LC et les modèles ELC

Remarque 1.1 *La liste présentée des classes de modèle pour les systèmes à changement de régime n'est aucunement exhaustive. En effet, dans la communauté des systèmes à évènements discrets, on retrouve d'autres formes de modélisation comme les automates temporisés, les réseaux de Petri temporisés dont il n'est pas fait mention ici.*

1.2.2 Les représentations sous forme de régression

De façon générale, les systèmes à changement de régime écrits sous forme de régression admettent la représentation suivante :

$$y(k) = \varphi_k^T \theta_{\mu_k} \tag{1.36}$$

où $\varphi_{(\cdot)}$ est le vecteur de régression. Le régime actif à l'instant k est indexé par la variable $\mu_{(\cdot)}$, $\mu_k \in \{1, \ldots, s\}$. θ_i, $i = 1, \ldots, s$ sont les vecteurs de paramètres associés aux différents modes de fonctionnement du système.
Le vecteur de régression peut être en principe n'importe quelle fonction des entrées $u(\cdot)$ et sorties $y(\cdot)$ passées du système. Nous nous limiterons ici au cas où la structure de $\varphi_{(\cdot)}$ est :

$$\varphi_k = [y_{k-1} \ldots y_{k-n_a} \; u_{k-1} \ldots u_{k-n_b} \; 1]^T \tag{1.37}$$

La dernière valeur du vecteur $\varphi_{(\cdot)}$ est égale à 1 afin d'autoriser la présence d'un terme constant dans (1.36). On peut aussi écrire $\varphi_k = [Y_{k-1,k-n_a} \; U_{k-1,k-n_b} \; 1]^T$ où $Y_{k-1,k-n_a}$ (respectivement $U_{k-1,k-n_b}$) représente l'empilement de la sortie (respectivement de l'entrée) du système de l'instant $k-1$ jusqu'à l'instant $k-n_a$ (respectivement $k-n_b$). En posant $X_k = [Y_{k-1,k-n_a} \; U_{k-1,k-n_b}]$, on a : $\varphi_k = [X_k \; 1]^T$.
Comme dans le cas des modèles sous forme de représentation d'état, il existe diverses manières de modéliser l'évolution du mécanisme générateur du mode actif $\mu_{(\cdot)}$.

1.2.2.1 Les modèles PWARX

Les modèles PWARX (*PieceWise AutoRegressive eXogenous*) modélisent le passage d'un régime de fonctionnement à un autre en utilisant une partition polyédrique d'un sous-ensemble $\Omega \subseteq \mathbb{R}^n$ sur lequel la relation (1.36) est vérifiée :

$$\mu_k = i \;\text{ si } X_k \in \Omega_i, \quad i = 1, \ldots, s \tag{1.38}$$

où $\bigcup_{i=1}^{s} \Omega_i = \Omega$, est une partition complète de l'ensemble des régresseurs Ω. Chaque région Ω_i est un polyèdre convexe :

$$\Omega_i = \left\{ X \in \mathbb{R}^n \,\middle|\, \bar{H}_i X + g_i \leq 0 \right\} \tag{1.39}$$

avec $\bar{H}_i \in \mathbb{R}^{q_i \times n}$, $g_i \in \mathbb{R}^{q_i}$, $i = 1, \ldots, s$.
En posant $H_i = [\bar{H}_i \; g_i]$ et en définissant une fonction affine par morceaux $f \; : \; \Omega_i \rightarrow \mathbb{R}$:

$$f\left(X_k\right) = \begin{cases} \varphi_k^T \theta_1 \; si \; H_1 \varphi_k \leq 0 \\ \vdots \qquad\qquad \vdots \\ \varphi_k^T \theta_s \; si \; H_s \varphi_k \leq 0 \end{cases}, \; \varphi_k = \begin{bmatrix} X_k \; 1 \end{bmatrix}^T \tag{1.40}$$

on peut récrire (1.36) sous la forme :

$$y\left(k\right) = f\left(X_k\right) \tag{1.41}$$

En résumé, un modèle PWARX est une famille de modèles ARX (*AutoRegressive eXogeneous*) reliés par des commutations qui sont gérées par une partition polyédrique de l'espace des régresseurs. Les modèles PWARX sont, pour les systèmes à changement de régime modélisés sous forme de régression, ce que sont les modèles PWA pour les systèmes à changement de régime modélisés sous forme de représentation d'état.

Exemple 1.10 (Fonction affine par morceaux) *La figure 1.8 montre deux fonctions affines par morceaux $f_1(\cdot)$ et $f_2(\cdot)$. La première fonction, celle du graphique de gauche, est continue sur les frontières communes aux polyèdres formant la partition polyédrique. La seconde fonction par contre, celle du graphique de droite, présente une discontinuité sur les frontières communes des polyèdres. Cette discontinuité vient du fait que sur les frontières communes aux polyèdres, la fonction affine $f_2(\cdot)$ admet plusieurs valeurs.*

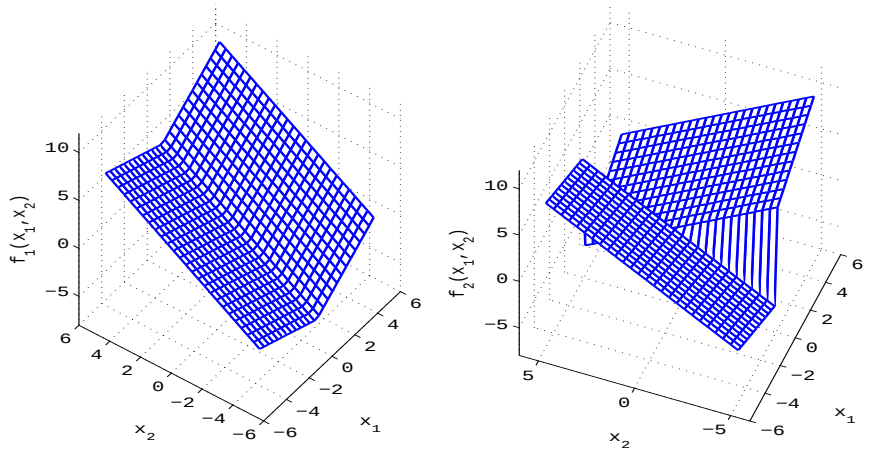


FIG. 1.8: Fonctions affines par morceaux $f_1(\cdot)$ et $f_2(\cdot)$

Exemple 1.11 (Modèle PWARX) *Considérons le système décrit par le modèle PWARX (1.42) :*

$$
\begin{aligned}
y(k) &= \begin{cases} \varphi_k^T\theta_1 \text{ si } H_1\varphi_k \leq 0 \text{ et } H_2\varphi_k \leq 0 \\ \varphi_k^T\theta_2 \text{ si } H_3\varphi_k \leq 0 \text{ et } H_4\varphi_k \leq 0 \\ \varphi_k^T\theta_3 \text{ si } H_5\varphi_k \leq 0 \text{ et } H_6\varphi_k \leq 0 \end{cases} \\
\varphi_k^T &= [y(k-1) \quad u(k-1) \quad 1] \\
\theta_1 &= [0 \ \ -1 \ \ 0], \ \theta_2 = [0 \ \ 1 \ \ 0], \ \theta_3 = [0 \ \ 3 \ \ -2], \\
H_1 &= [0 \ \ -1 \ \ -6], \ H_2 = [0 \ 1 \ 1], \\
H_3 &= [0 \ \ -1 \ \ -1], \ H_4 = [0 \ 1 \ \ -3], \\
H_5 &= [0 \ 1 \ \ -3], \ H_6 = [0 \ \ 1 \ \ -6]
\end{aligned}
\tag{1.42}
$$

Ce système est présenté à la figure 1.9. Sur cette figure sont superposés le tracé global du système ainsi que des points représentatifs du système qui pourraient éventuellement servir dans une phase d'identification du système. Le graphique de gauche est tracé dans le plan (u_{k-1}, y_k) alors que celui de droite est réalisé dans le plan (u_{k-1}, y_{k-1}, y_k).

La figure 1.10 représente l'évolution de l'entrée $u(\cdot)$ (graphique de la première ligne) et de la sortie $y(\cdot)$ (graphique de la seconde ligne) en fonction du temps. Contrairement aux graphiques de la figure 1.9, les différents régimes de fonctionnement

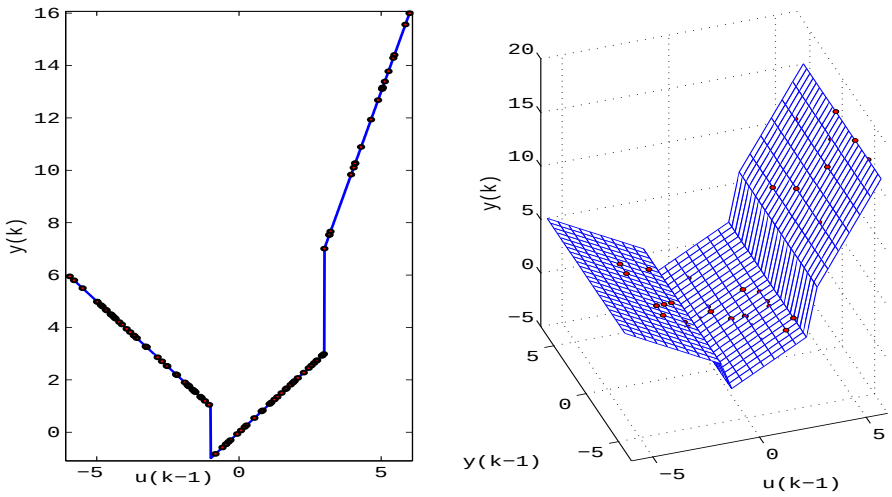


Fig. 1.9: Exemple de modèle PWARX

du système ne sont plus mis en exergue par les tracés de l'évolution temporelle de l'entrée et de la sortie du système. Ceci illustre l'apport de la connaissance du vecteur de régression φ_k lors de la phase d'identification du système.

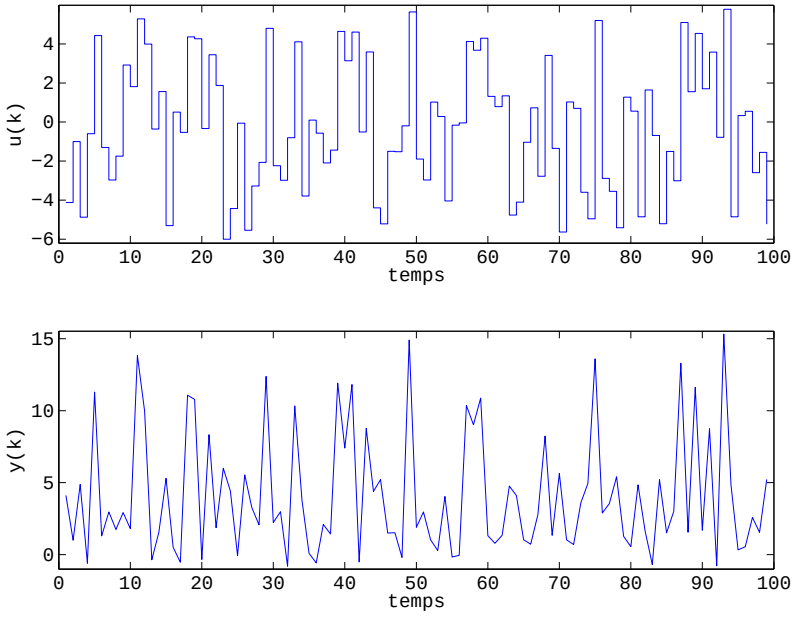


Fig. 1.10: Evolution de $u(\cdot)$ et de $y(\cdot)$ en fonction du temps

Les modèles PWARX englobent un grand nombre de systèmes affines ou linéaires par morceaux couramment usités. Parmi ces systèmes, nous allons présenter dans

les sections suivantes les modèles HH (*Hinging Hyperplane*), la représentation canonique de Chua et les modèles Hammerstein/Wiener PWARX (H/W-PWARX).

1.2.2.2 Les modèles HH

Les modèles HH (*Hinging Hyperplane*) sont basés sur les fonctions HH introduites par Breiman [1993]. L'avantage des modèles HH est que les fonctions affines par morceaux qui les définissent sont toujours continues sur les frontières des polyèdres définissant la partition polyédrique de l'espace des régresseurs. Utilisées pour l'approximation de fonctions non-linéaires et pour la classification, les fonctions HH sont formées en effectuant une somme de fonctions $h_i(\cdot),\ i = 1, \dots M$, appelées fonctions charnières. Chaque fonction $h_i(\cdot)$ est formée de deux demi-hyperplans continus à leur frontière commune (voir figure 1.11) :

$$\begin{cases} f(x) = \sum_{i=1}^{M} h_i(x) \\ h_i(x) = \pm \max\left\{\varphi_k^T \theta_i^+, \varphi_k^T \theta_i^-\right\} \end{cases} \tag{1.43}$$

où $f(\cdot)$ est la fonction HH obtenue en effectuant la somme des foncions charnières $h_i(\cdot),\ i = 1, \dots M$.
Le signe $\pm$ devant la fonction $\max(\cdot)$ permet la représentation de fonctions convexes ou non convexes.
On obtient, à partir de la définition de la fonction HH, le modèle HH (1.44) :

$$y(k) = \sum_{i=1}^{M} \pm \max\left\{\varphi_k^T \theta_i^+, \varphi_k^T \theta_i^-\right\} \tag{1.44}$$

où $y(\cdot) \in \mathbb{R}$ est la sortie du système. Le modèle (1.44) peut être également mis sous la forme :

$$y(k) = \varphi_k^T \theta_0 + \sum_{i=1}^{M^+} \max\left\{\varphi_k^T \theta_i, 0\right\} - \sum_{i=M^+ +1}^{M} \max\left\{\varphi_k^T \theta_i, 0\right\} \tag{1.45}$$

Ici la fonction affine par morceaux est explicitée comme la somme d'un terme affine $(\varphi_k^T \theta_0)$ et de M fonctions max positives et négatives. Les fonctions max positives sont regroupées dans un terme et celles qui sont négatives dans un autre terme, ce

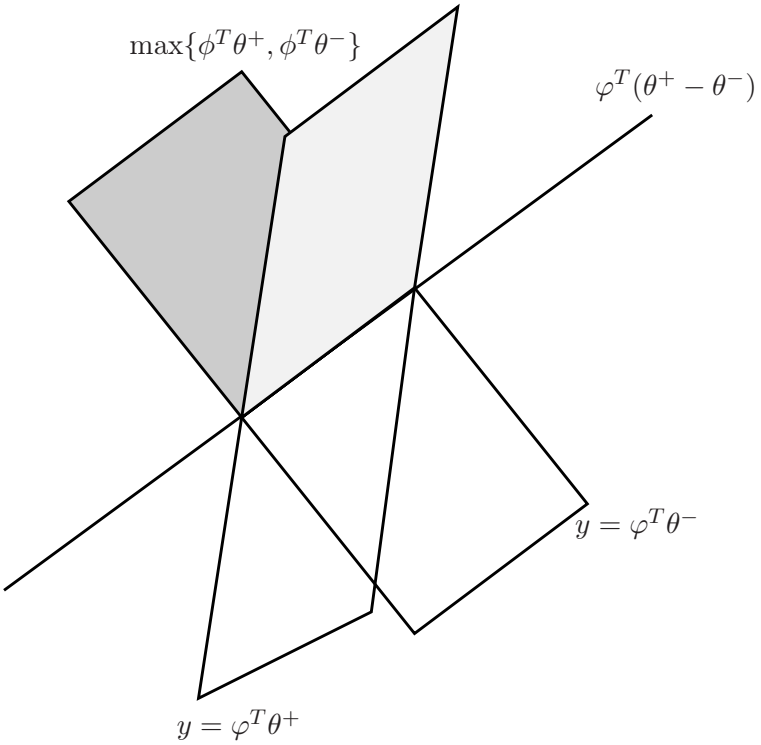


FIG. 1.11: Hyperplans $y = \varphi_k^T\theta^-$ et $y = \varphi_k^T\theta^+$ et la fonction HH correspondante $y = \max\{\varphi_k^T\theta^+, \varphi_k^T\theta^-\}$

qui justifie la disparition du signe $\pm$. Cette forme de représentation porte le nom de modèle HHARX (*Hinging Hyperplane AutoRegressive eXogenous*).

Exemple 1.12 (Modèle HH) *La figure 1.12 montre le modèle HH (1.46) :*

$$\begin{aligned} y(k) &= 0.3y(k-1) - 0.21u(k-1) \\ &\quad + \max\{-2y(k-1) + 3u(k-1), 0\} \\ &\quad + \max\{0.5y(k-1) + 1.5u(k-1), 0\} \end{aligned} \tag{1.46}$$

Les graphiques de la première ligne représentent respectivement, de la gauche vers la droite, les termes $0.1y(k-1) - 0.2u(k-1)$, $\max\{y(k-1) + u(k-1), 0\}$ *et* $\max\{-y(k-1) + 3u(k-1), 0\}$. *La sortie* $y(\cdot)$ *du système, représentée à la seconde ligne, est obtenue en sommant les graphiques de la première ligne.*

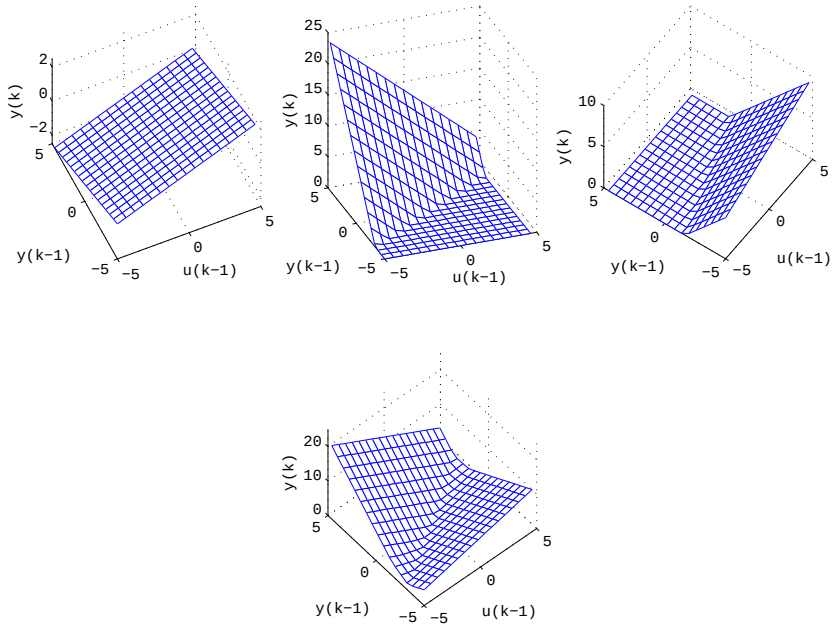


Fig. 1.12: Exemple de modèle HH

1.2.2.3 Représentation canonique de Chua

La représentation canonique de Chua est une classe de modèles affines par morceaux proposée par Chua et Kang [1977]; Kang et Chua [1978] et qui s'apparente beaucoup aux modèles HH :

$$y(k) = \varphi_k^T \alpha_0 + \sum_{i=1}^{M} c_i \left| \varphi_k^T \alpha_i \right| \tag{1.47}$$

Il existe une équivalence entre la représentation canonique de Chua et les modèles HH qui est aisée à démontrer. En effet, il est possible de faire la transformation suivante :

$$\begin{aligned}
&\max\left\{\varphi_k^T\theta_i^+, \varphi_k^T\theta_i^-\right\} \\
&\quad= \frac{1}{2}\varphi_k^T\left(\theta_i^+ + \theta_i^-\right) + \max\left\{\varphi_k^T\left(\theta_i^+ - \frac{1}{2}\left(\theta_i^+ + \theta_i^-\right)\right), \varphi_k^T\left(\theta_i^- - \frac{1}{2}\left(\theta_i^+ + \theta_i^-\right)\right)\right\} \\
&\quad= \frac{1}{2}\varphi_k^T\left(\theta_i^+ + \theta_i^-\right) + \frac{1}{2}\max\left\{\varphi_k^T\left(\theta_i^+ + \theta_i^-\right), -\varphi_k^T\left(\theta_i^+ - \theta_i^-\right)\right\} \\
&\quad= \frac{1}{2}\varphi_k^T\left(\theta_i^+ + \theta_i^-\right) + \frac{1}{2}\left|\varphi_k^T\left(\theta_i^+ - \theta_i^-\right)\right|
\end{aligned}$$

L'équivalence entre (1.44) et (1.47) est ainsi établie.

1.2.2.4 Les modèles Hammerstein/Wiener PWARX

Cette classe de modèles, qui admet plusieurs applications pratiques, résulte de l'interconnection d'un modèle linéaire dynamique et d'une non linéarité statique. On distingue deux cas : lorsque la non linéarité précède le modèle linéaire, on parle de modèle de Hammerstein et lorsqu'elle le suit on parle de modèle de Wiener. Si de plus, la non linéarité est affine par morceaux, on parle alors de modèles Hammerstein PWARX (H-PWARX) et de modèles Wiener PWARX (W-PWARX). De façon générale, les modèles H-PWARX (respectivement les modèles W-PWARX) sont obtenus par la connexion en cascade d'une non linéarité statique affine par morceaux suivie (respectivement précédée) d'un modèle ARX (*AutoRegressive eXogenous*). La figure 1.13 montre quelques exemples typiques de non linéarités statiques affines par morceaux. Le graphique de gauche illustre le phénomène de saturation, celui du milieu représente une zone morte et enfin le dernier montre un frottement visqueux de Coulomb.

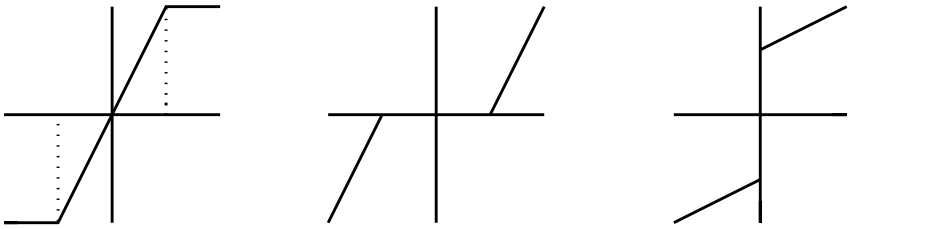


Fig. 1.13: Exemple de non-linéarités affines

Un modèle H-PWARX est donné par :

$$\begin{cases} x(k) = g(u(k)) \\ y(k) = -\sum_{i=1}^{n_a} a_i y(k-i) + \sum_{j=1}^{n_b} b_j x(k-j) \end{cases} \tag{1.48}$$

$u(\cdot)$ et $y(\cdot)$ sont respectivement l'entrée et la sortie du système. $x(\cdot)$ est une variable interne non mesurée qui constitue l'entrée du système ARX. $a_i \in \mathbb{R}$, $i = 1, \ldots, n_a$ et $b_j \in \mathbb{R}$, $j = 1, \ldots, n_b$ sont les paramètres du modèle ARX. La fonction $g(\cdot)$ est

affine par morceaux.
De façon similaire, le modèle W-PWARX se présente sous la forme :

$$\begin{cases} x(k) = -\sum_{i=1}^{n_a} a_i x(k-i) + \sum_{j=1}^{n_b} b_j u(k-j) \\ y(k) = g(x(k)) \end{cases} \tag{1.49}$$

Dans (1.49), la variable interne $x(\cdot)$ est la sortie du modèle ARX.

La figure 1.14 montre le schéma bloc des modèles H/W-PWARX.

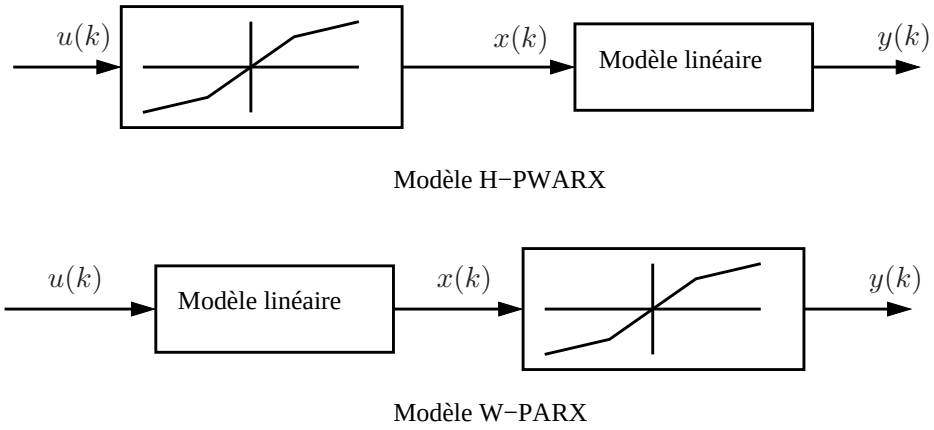


Fig. 1.14: Les modèles H-PWARX et W-PWARX

1.3 Liens entre la représentation d'état et la représentation sous forme de régression

Il est souvent admis dans la littérature que le passage d'une représentation sous forme de modèle d'état à une représentation sous forme de régression (et vice versa) est aisé à effectuer et conduit à l'obtention de structures qui modélisent de façon équivalente le système considéré. Ceci reste valable pour les systèmes linéaires invariants. La situation est légèrement différente pour les systèmes à changements brusques de régimes. La réciprocité du passage modèle d'état/modèle de régression n'est établie que pour certaines formes particulières de modèles. De façon générale, on peut passer pour un système à changements brusques de régimes d'un modèle d'état à un modèle de régression, mais la structure des modèles obtenus subit des

modifications lors du transfert.
Considérons le modèle d'état (1.50) représentatif d'un système linéaire :

$$\begin{cases} x(k+1) = Ax(k) + Bu(k) \\ y(k) = Cx(k) \end{cases} \tag{1.50}$$

Pour obtenir un modèle de régression équivalent à partir (1.50), on peut se servir des propriétés de l'opérateur retard q défini par : $x(k-1) = q^{-1}x(k)$. On obtient ainsi un modèle de régression équivalent à (1.50) :

$$y(k) = C\left(I - Aq^{-1}\right)^{-1} Bq^{-1}u(k) \tag{1.51}$$

Considérons maintenant un système à changements de régime modélisé par la représentation d'état (1.14) :

$$\begin{cases} x(k+1) = A_{\mu_k}x(k) + B_{\mu_k}u(k) \\ y(k) = C_{\mu_k}x(k) \end{cases}$$

A partir de l'équation d'état, on obtient :

$$x(k) = \left(I - A_{\mu_{k-1}}q^{-1}\right)^{-1} B_{\mu_{k-1}}q^{-1}u(k)$$

On a alors :

$$y(k) = C_{\mu_k}\left(I - A_{\mu_{k-1}}q^{-1}\right)^{-1} B_{\mu_{k-1}}q^{-1}u(k) \tag{1.52}$$

Il faut constater ici que l'expression (1.52) fait apparaître des valeurs de la loi de commutation à des instants différents, notamment $k-1$ et k. Ceci signifie que l'intuition première selon laquelle il aurait suffit de procéder au passage en modèle de régression de chacun des modèles d'état modélisant les différents régimes du système est fausse. En fait le modèle de régression équivalent obtenu comprend tout d'abord plus de régimes de fonctionnement que le modèle d'état de départ. Ensuite, ce modèle de régression comprend des régimes dont les paramètres sont des combinaisons des modèles d'états locaux de la représentation d'état de départ. Rosenqvist et Karlström [2005] ont montré que pour certaines formes de modèles

d'état, il est possible d'obtenir un modèle de régression dont les paramètres de chacun des régimes sont associés à un seul des modes de la représentation d'état.

Proposition 1.1

On suppose que la loi de commutation $\mu(\cdot)$ du système à changement brusque de régimes dépend exclusivement de l'entrée $u(\cdot)$ et de la sortie $y(\cdot)$. Si le modèle d'état (1.14) associé au système est de la forme :

$$\begin{aligned} A_{\mu_k} &= \begin{bmatrix} -a_{1,\mu_k} & \dots & -a_{n-1,\mu_k} & -a_{n,\mu_k} \\ 1 & \dots & 0 & 0 \\ \vdots & \ddots & \vdots & \vdots \\ 0 & \dots & 1 & 0 \end{bmatrix} \\ B_{\mu_k} &= \begin{bmatrix} 1 & \dots & 0 & 0 \end{bmatrix}^T \\ C_{\mu_k} &= \begin{bmatrix} b_{1,\mu_k} & \dots & b_{n-1,\mu_k} & b_{n,\mu_k} \end{bmatrix}, \end{aligned} \tag{1.53}$$

il existe alors un modèle de régression entrée/sortie correspondant à ce modèle d'état et dans lequel chaque paramètre est associé à un seul mode :

$$y(k) = \frac{\sum_{j=1}^{n} b_{j,\mu_{k-1}} q^{-j}}{1 + \sum_{i=1}^{n} a_{i,\mu_{k-1}} q^{-i}} u(k). \tag{1.54}$$

De même si le modèle d'état (1.14) est sous la forme :

$$\begin{aligned} A_{\mu_k} &= \begin{bmatrix} -a_{1,\mu_k} & 1 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ -a_{n-1,\mu_k} & 0 & \dots & 1 \\ -a_{n,\mu_k} & 0 & \dots & 0 \end{bmatrix} \\ B_{\mu_k} &= \begin{bmatrix} b_{1,\mu_k} & \dots & b_{n-1,\mu_k} & b_{n,\mu_k} \end{bmatrix}^T \\ C_{\mu_k} &= \begin{bmatrix} 1 & \dots & 0 & 0 \end{bmatrix}, \end{aligned} \tag{1.55}$$

il existe un modèle de régression entrée/sortie correspondant dans lequel chaque paramètre est associé à un seul mode :

$$y\left(k\right)=\frac{\sum_{j=1}^{n} b_{j,\mu_{k-1}} q^{-j}}{1+\sum_{i=1}^{n} a_{i,\mu_{k-1}} q^{-i}} u\left(k\right). \tag{1.56}$$

La proposition 1.1 montre que le passage d'un modèle d'état à un modèle de régression dans lequel les paramètres associés aux différents régimes de fonctionnement ne dépendent que d'un seul mode de la représentation d'état n'est possible qu'à condition que les modèles locaux du modèle d'état soient sous forme compagne d'observabilité ou de commandabilité. La démonstration de cette proposition peut être retrouvée dans [Rosenqvist et Karlström, 2005].

Le passage d'une représentation sous forme de régression à une représentation d'état est beaucoup plus aisé et se fait en posant tout simplement :

$$x(k)=\begin{bmatrix} y(k-1)\ \ldots\ y(k-n_a)\ \ u(k-1)\ \ldots\ u(k-n_b) \end{bmatrix}^T \tag{1.57}$$

Pour le cas particulier où on a :

$$y\left(k\right)=\sum_{i=1}^{n_a} a_{i,\mu_k} y\left(k-i\right)+u\left(k-1\right), \tag{1.58}$$

on obtient, en posant $x(k)=\begin{bmatrix} y(k-1)\ \ldots\ y(k-n_a) \end{bmatrix}^T$, une représentation d'état équivalente dont les matrices d'état A_{μ_k}, de commande B_{μ_k} et de sortie C_{μ_k} sont données par :

$$\begin{aligned} A_{\mu_k} &= \begin{bmatrix} a_{1,\mu_k} & \ldots & a_{n_a-1,\mu_k} & a_{n_a,\mu_k} \\ 1 & \ldots & 0 & 0 \\ \vdots & \ddots & \vdots & \vdots \\ 0 & \ldots & 1 & 0 \end{bmatrix} \\ B_{\mu_k} &= \begin{bmatrix} 1 \ldots 0\ 0 \end{bmatrix}^T \\ C_{\mu_k} &= \begin{bmatrix} a_{1,\mu_k} & \ldots & a_{n_a-1,\mu_k} & a_{n_a,\mu_k} \end{bmatrix}, \end{aligned} \tag{1.59}$$

la sortie du modèle d'état étant définie par $y(k)-u(k)$.

1.4 Identification des systèmes à changement de régime

Nous considérons ici un système à changement de régime modélisé sous forme de régression :

$$y\left(k\right) = \varphi_k^T \theta_{\mu_k} + \varepsilon(k) \tag{1.60}$$

où $\mu_k \in \{1, \ldots, s\}$, $\varphi_k = [Y_{k-1,k-n_a}\ U_{k-1,k-n_b}\ 1]^T$ et $\varepsilon(\cdot)$ est un terme d'erreur.
L'objectif est de procéder à l'identification des paramètres du modèle (1.60) à partir d'un jeu de mesures $\mathcal{D}$ de N points entrée/sortie du système :

$$\mathcal{D} = \{(y_k, X_k)\,,\ k = 1, \ldots, N\} \tag{1.61}$$

Dans (1.61), $y_{(\cdot)} \in \mathbb{R}$ est la sortie mesurée du système ; $X_{(\cdot)} \in \mathbb{R}^n$ est le vecteur de régression défini par $X_k = [Y_{k-1,k-n_a}\ U_{k-1,k-n_b}]$, n_a et n_b étant les ordres de régression.
De façon générale, l'identification de la structure (1.60) nécessite la détermination de trois quantités :

- le nombre s de régimes de fonctionnement du système,
- les paramètres $\theta_i,\ i = 1, \ldots, s$ associés à chacun des régimes de fonctionnement du système,
- les coefficients $H_i,\ i = 1, \ldots, s$ des hyperplans définissant le partitionnement de l'ensemble des régresseurs.

La première difficulté inhérente à l'identification d'un système à changement de régime est le choix du nombre s de régimes de fonctionnement. Par exemple, on obtiendrait un modèle qui expliquerait parfaitement le jeu de données $\mathcal{D}$ en choisissant autant de régimes de fonctionnement que de données : $s = N$. Il est évident que cette solution est à proscrire car elle introduit une surparamétrisation du modèle obtenu. Il est nécessaire d'imposer des contraintes sur s de façon à minimiser le nombre de régimes de fonctionnement tout en garantissant une erreur d'identification faible.
Dans la situation où le nombre de régimes de fonctionnement du système est connu, le problème d'identification est équivalent à la recherche des paramètres d'une fonc-

tion affine à partir du jeu de données $\mathcal{D}$. On procède alors à la minimisation du critère (1.62) par rapport à $\theta = (\theta_1, \ldots, \theta_s)$ et $H_i,\ i = 1, \ldots, s$:

$$V_N(\theta, H_i) = \frac{1}{N} \sum_{k=1}^{N} \ell\left(y_k - \varphi_k^T \theta_{\mu_k}\right) \tag{1.62}$$

où ℓ est une fonction non négative, par exemple $\ell(\varepsilon) = \varepsilon^2$, ou $\ell(\varepsilon) = |\varepsilon|$.
La minimisation du critère (1.62) par rapport à θ conduit généralement à un problème fortement non convexe avec des minima locaux.
La difficulté sous-jacente lors de l'identification d'un système à changement de régime est le couplage existant entre la classification des données et l'identification des paramètres. En effet, chaque élément du jeu de données $\mathcal{D}$ est associé à une région Ω_i de l'espace des régresseurs définie en (1.39) et au régime de fonctionnement correspondant à cette région. Pour estimer les paramètres d'un régime de fonctionnement, il faut donc déjà connaître les données qui caractérisent ce régime (classification).
En introduisant les quantités :

$$\omega_i(k) = \begin{cases} 1 \text{ si } X_k \in \Omega_i \\ 0 \text{ tous les autres cas} \end{cases}, \tag{1.63}$$

la minimisation du critère (1.62) peut alors être récrite de la façon suivante :

$$\min_{\theta} \left(\frac{1}{N} \sum_{k=1}^{N} \sum_{i=1}^{s} \ell\left(y(k) - \varphi_k^T \theta_i\right) \omega_i(k) \right) \tag{1.64}$$

On peut scinder de façon générale l'ensemble des méthodes existantes pour l'identification des systèmes à changement de régime en deux grandes catégories [Paoletti, 2004; Roll, 2003].
La première catégorie utilise une approche nécessitant la définition d'un partitionnement *a priori* de l'espace des régresseurs [Billings et Voon, 1987; Julián *et al.*, 1999]. L'avantage de cette approche est que l'estimation des paramètres des régimes de fonctionnement est simplifiée car on peut, connaissant le partitionnement, utiliser des méthodes classiques d'identification pour estimer les paramètres de chaque

régime de fonctionnement. Toutefois, l'existence de nombreux minima locaux, la complexité calculatoire de la méthode et la croissance exponentielle du nombre de régions rend ce type d'approche difficilement exploitable, notamment pour les systèmes d'ordre élevé.

Le seconde catégorie de méthodes procède à l'estimation simultanée ou itérative des régimes de fonctionnement et des régions du partionnement de l'espace des régresseurs. En fonction de la façon dont le partitionnement est effectué, on distingue dans cette catégorie quatre groupes :

- le premier groupe englobe les méthodes procédant à la minimisation à l'aide de méthodes numériques (méthode du gradient par exemple) d'un critère adéquat permettant l'estimation simultanée des paramètres des régimes de fonctionnement et des coefficients associés aux hyperplans définissant les régions du partitionnement [Batruni, 1991; Chan et Tong, 1986; Gad *et al.*, 2000; Julián *et al.*, 1998; Pucar et Sjöberg, 1998]. Il demeure possible, lors de la minimisation, de converger vers des minima locaux, mais cette approche a l'avantage d'offrir une formulation directe du problème d'identification.
- le second groupe de méthodes, qui est en fait une extension du premier groupe, opère par une identification simultanée des paramètres des régimes de fonctionnement et des régions du partitionnement à partir d'un partitionnement de base très simple de l'espace des régresseurs [Breiman, 1993; Ernst, 1998; Heredia et Arce, 1996; Hush et Horne, 1998; Julián *et al.*, 1998; Pucar et Sjöberg, 1998]. Dans le cas où le résultat de l'identification n'est pas satisfaisant, de nouveaux régimes de fonctionnement ou de nouvelles régions sont ajoutés afin d'améliorer les résultats. Le risque d'existence de minima locaux reste présent et il existe également un risque de surparamétrisation du modèle lors de l'ajout de nouveaux régimes de fonctionnement ou de nouvelles régions. L'avantage obtenu ici est que le problème d'identification est divisé en plusieurs sous-problèmes d'identification beaucoup plus aisés à résoudre que le problème de départ.
- les méthodes du troisième groupe identifient de façon itérative les paramètres des régimes de fonctionnement et les régions du partitionnement [Bemporad *et al.*, 2003; Ferrari-Trecate *et al.*, 2003; Medeiros *et al.*, 2002; Münz et Krebs,

2002; Ragot *et al.*, 2003*b*; Skeppstedt *et al.*, 1992; Vidal *et al.*, 2003]. A chaque étape, on prend en compte soit les régimes de fonctionnement, soit les régions du partitionnement.

– le dernier groupe estime les régions du partitionnement en utilisant exclusivement l'information sur la distribution des vecteurs de régression [Choi et Choi, 1994; Strömberg *et al.*, 1991]. Il arrive très souvent, dans ce cas, qu'un ensemble de données qui devraient en principe appartenir à la même région soit scindé en plusieurs régions distinctes.

Il est à noter que quelle que soit la méthode utilisée, le problème de l'existence de minima locaux reste présent et il est impossible de dire laquelle de toutes ces méthodes est la meilleure, chaque méthode présentant ses avantages et ses inconvénients.

Exemple 1.13 (Système linéaire statique à deux modes de fonctionnements) *Considérons le système linéaire statique à deux modes de fonctionnement modélisé par :*

$$\begin{aligned} y(k) &= a u(k) \\ a &= \begin{cases} a_1 = 0.85 & \text{si } \nu(k) - 0.5 \geq 0 \\ a_2 = -0.70 & \text{si } \nu(k) - 0.5 < 0 \end{cases} \end{aligned} \tag{1.65}$$

où $\nu(k)$ est la réalisation d'une variable aléatoire ayant une distribution uniforme sur l'intervalle $[0,\ 1]$.

La figure 1.15 montre l'évolution temporelle de l'entrée $u(\cdot)$ et de la sortie $y(\cdot)$ du système.

On peut, à partir de l'entrée et de la sortie du système, visualiser l'évolution temporelle du paramètre a du modèle (1.65) en calculant la quantité $y(k)/u(k)$. C'est ce que montre le premier graphique de la figure 1.16. On y note la présence de commutations et de deux valeurs distinctes pour le paramètre a. En se positionnant dans le plan $(u(k), y(k))$, on met encore plus en évidence l'existence de deux valeurs distinctes pour le paramètre a. Le deuxième graphique de la figure 1.16 illustre cet état de chose. Par contre, le paramètre temps étant masqué, les commutations n'apparaissent plus cette fois-ci.

En définissant la fonction coût suivante :

$$\Phi(a) = \frac{1}{2}\sum_{k=1}^{N} \frac{1}{\tau + (y(k) - au(k))^2},$$

on peut discriminer les deux modes de fonctionnement en maximisant cette fonction par rapport à a, N étant le nombre de points de mesure et τ une constante assez petite. La figure 1.17 montre l'évolution de la fonction coût $\Phi(\cdot)$ vis-à-vis de a, pour $\tau = 0.0001$. On y voit distinctement deux pics correspondant aux deux valeurs prises par a.

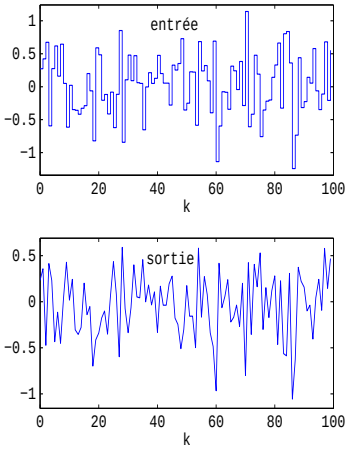


Fig. 1.15: Evolution de l'entrée et de la sortie

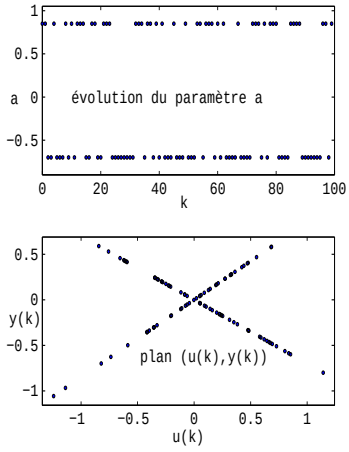


Fig. 1.16: Paramètre a et plan $(u(k), y(k))$

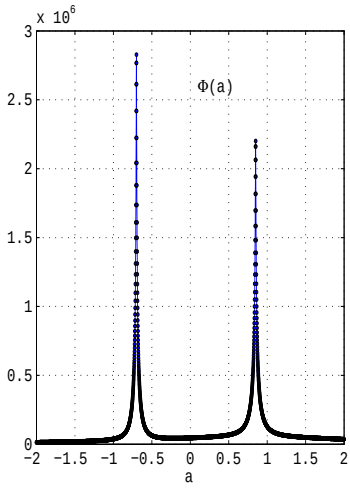


Fig. 1.17: Fonction coût

1.5 Liens entre multi-modèles et systèmes à changement de régime

Lors de l'élaboration d'un multi-modèle, l'objectif recherché est la représentation du comportement d'un système complexe (souvent non linéaire) par un ensemble de sous-modèles, chacun des sous-modèles étant valable dans une des régions complémentaires de l'espace défini par les grandeurs permettant d'expliquer la sortie du système. En d'autres termes, il s'agit de trouver des modèles de structures assez simples (souvent linéaires), chacun représentant le comportement du système dans une zone de fonctionnement bien spécifique. Le modèle global du

processus résulte alors de la combinaison des différents sous-modèles par le biais d'un mécanisme d'agrégation. Ces sous-modèles sont qualifiés de modèles locaux puisque n'étant valables que dans une zone donnée de fonctionnement ou autour d'un point de fonctionnement donné. Un multi-modèle est donc une structure résultant de l'agrégation de modèles locaux.
Pour caractériser un multi-modèle, il faut spécifier le nombre de modèles locaux, la structure de chaque modèle local ainsi que la nature de la fonction d'agrégation ou d'activation. On recense essentiellement deux structures de multi-modèles : les structures à modèles locaux couplés [Johansen et Foss, 1993] et les structures à modèles locaux découplés [Filev, 1991].

La structure à modèles locaux couplés suppose que le multi-modèle possède un état $x(\cdot)$ unique et global (voir figure 1.18). Celui-ci est défini de la manière suivante :

$$x\left(k+1\right)=\sum_{i=1}^{s}\mu_i\left(\xi\left(k\right)\right)\left(A_i x\left(k\right)+B_i u\left(k\right)+D_i\right) \tag{1.66}$$

où $\mu_i(\xi(k)), i \in \{1, ..., s\}$ sont les fonctions d'activation et $\xi(k)$ est le vecteur des variables de décision dépendant des grandeurs mesurables du système et éventuellement de grandeurs estimables.

Dans le cas des structures à modèles locaux découplés, le multi-modèle est défini par une somme pondérée de modèles locaux dont les états évoluent indépendamment les uns des autres (voir figure 1.19). Le modèle global est alors donné par :

$$\begin{aligned} x_i\left(k+1\right) &= A_i x_i\left(k\right)+B_i u\left(k\right)+D_i \qquad i=1,\ldots,s \\ x\left(k+1\right) &= \sum_{i=1}^{s}\mu_i\left(\xi\left(k\right)\right)x_i\left(k+1\right) \end{aligned} \tag{1.67}$$

Cette structure modélise la mise en parallèle de s modèles affines pondérés par leurs poids respectifs.

De façon plus générale, un multi-modèle permet de traduire une plus grande variété de comportements non linéaires comparativement aux systèmes à régimes multiples présentés jusqu'ici. En jouant sur l'allure de la fonction d'agrégation, il est possible d'approximer toute forme de non-linéarités, mêmes celles se traduisant par un passage brusque d'un régime de fonctionnement à un autre comme cela est le cas avec les systèmes à changement abrupt de régimes. De ce fait, les systèmes

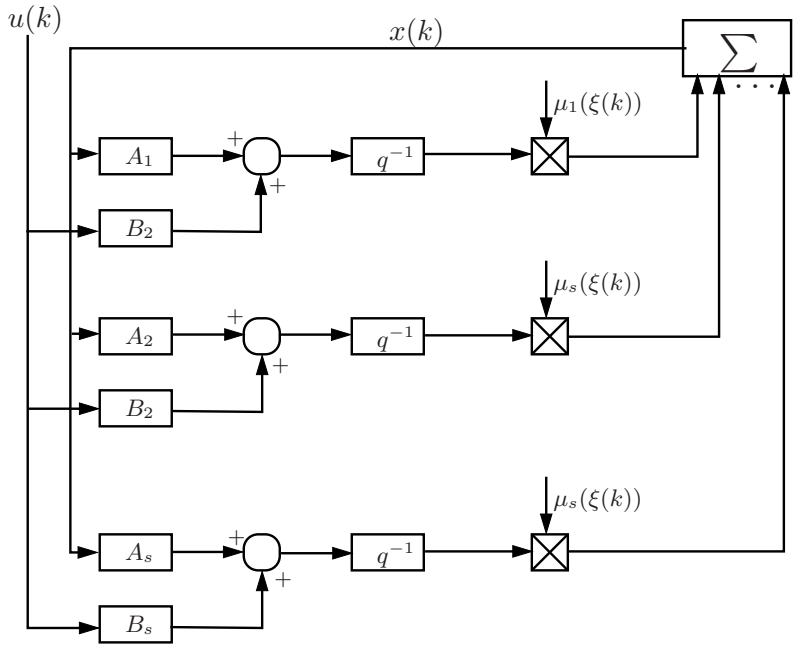


Fig. 1.18: Structure d'un multimodèle à modèles locaux couplés

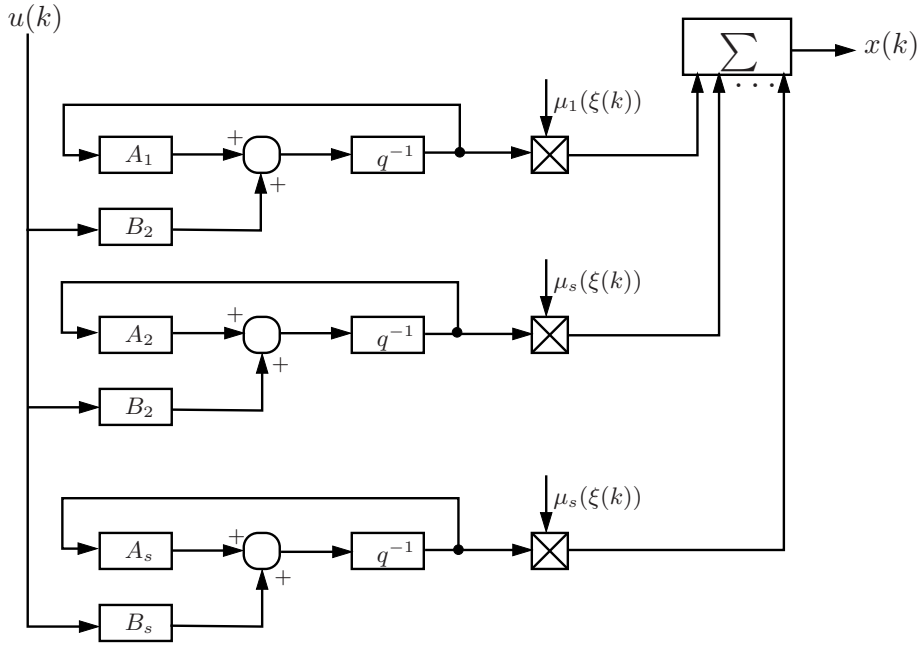


Fig. 1.19: Structure d'un multimodèle à modèles locaux découplés

à changement abrupt de régimes peuvent être vus comme un cas particulier de système multi-modèle.

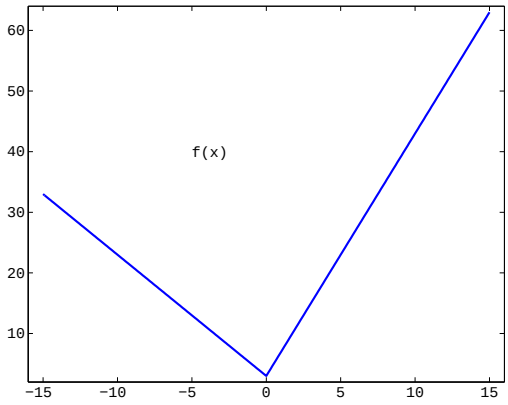
Exemple 1.14 (Du multi-modèle vers le modèle à changement de régime)

Cet exemple montre les capacités d'approximation des multi-modèles, de même que le lien existant entre les multi-modèles et les modèles à changement de régime.

On considère la fonction affine par morceaux définie par :

$$f(x) = \begin{cases} -2x + 3 & \text{si } x \leq 0 \\ 4x + 3 & \text{si } x > 0 \end{cases}$$

La fonction $f(\cdot)$ est représentée ci-contre.



On approxime ensuite cette fonction en se servant d'un multi-modèle à deux modèles locaux :

$$\begin{cases} f_1(x) = -2x + 3 \\ f_2(x) = 4x + 3 \end{cases}$$

Les fonctions d'agrégation des modèles locaux sont choisies sous forme de fonctions tangentes hyperboliques définies par :

$$\mu_i(x) = \frac{1}{2}\left(1 + \tanh\left(\frac{x - c_i}{\sigma_i}\right)\right), \quad i = 1, 2$$

Le paramètre σ_i permet de définir la pente à l'origine de la fonction $\tanh(\cdot)$. Il règle le mélange entre les modèles locaux. Dans notre exemple, on a choisi $\sigma_1 = \sigma_2 = \sigma$ et $c_1 = c_2 = c$.

Le modèle réalisant l'approximation est :

$$\hat{f}(x) = \mu_1(x) f_1(x) + \mu_2(x) f_2(x)$$

la qualité de l'appromixation pouvant être définie par :

$$\Phi = \left\| f(x) - \hat{f}(x) \right\|$$

où représente $\|\cdot\|$ par exemple la norme euclidienne.

La figure 1.20 montre l'approximation obtenue pour diverses valeurs de σ_i. On

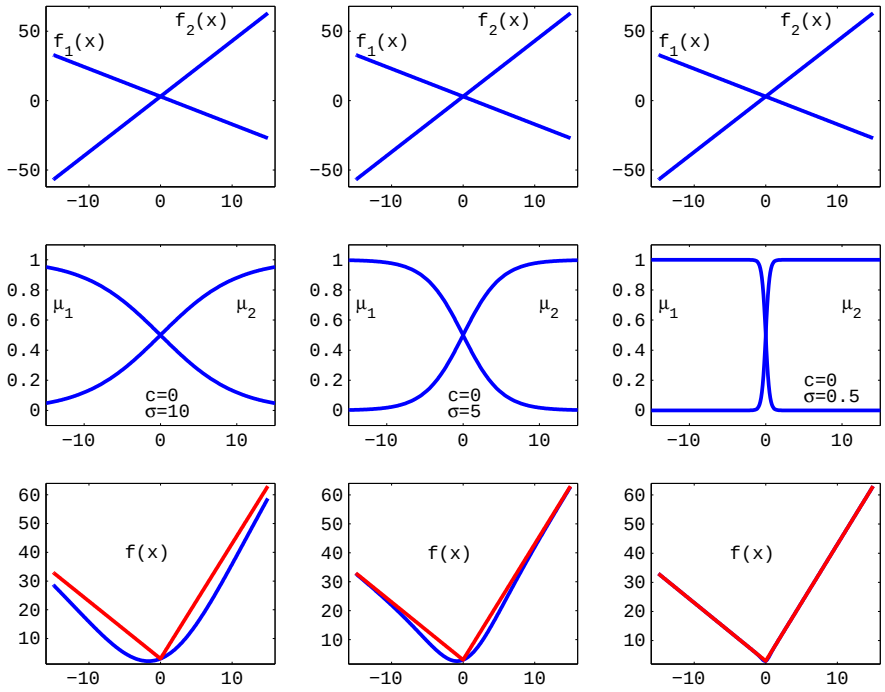


Fig. 1.20: Approximation de la fonction $f(\cdot)$ par un multi-modèle

peut constater que Φ diminue lorsque σ diminue. Il est possible de déterminer σ pour que la quantité Φ soit inférieure à un seuil aussi petit que l'on souhaite.

1.6 Stabilité des systèmes à changement de régime

La stabilité des systèmes à changement de régime est essentiellement étudiée pour les systèmes sous forme de représentation d'état (modèle PWA, JL, PL). Contrairement aux cas des systèmes linéaires invariants où la stabilité peut être aisément vérifiée en calculant les valeurs propres de la matrice d'état, la situation est plus compliquée lorsqu'on doit tenir compte de divers régimes de fonctionnement. Il est inutile d'espérer tirer des informations sur la stabilité ou l'instabilité d'un système à changement de régime en s'intéressant uniquement à la stabilité ou à l'instabilité de ses modes.

L'étude de la stabilité des systèmes à changements de régimes a connu récemment un regain d'attention. Ceci essentiellement parce qu'elle permet la synthèse de loi

de commande par rétroaction explicite et de faible complexité [Grieder et Morari, 2003; Grieder *et al.*, 2004, 2003]. La majeure partie des méthodes existantes procède à une extension de la méthode de Lyapunov aux systèmes à commutation.
Un éventail de méthodes avec des degrés variables de conservatisme a été développé pour analyser la stabilité des systèmes à changement de régime à partir de la théorie de Lyapunov. Puisqu'il n'y a aucune méthode standard pour construire des fonctions de Lyapunov, des algorithmes permettant le calcul de fonctions de Lyapunov pour une large classe de fonctions candidates ont été proposés. Il s'agit du calcul de fonctions de Lyapunov communes sous forme d'un polynôme de degré 2 (*Quadratic Lyapunov function*) ou 4 (*Quartic Lyapunov function*), de fonctions de Lyapunov affines par morceaux (*PieceWise Affine Lyapunov function*), de fonctions de Lyapunov polynomiales par morceaux, chaque morceau de la fonction étant représenté par un polynôme de degré 2 (*PieceWise Quadratic Lyapunov function*) ou de degré 4 (*PieceWise Quartic Lyapunov function*). Ces fonctions sont efficacement calculées à l'aide de la programmation linéaire ou de la programmation semi-définie. Dans [Biswas *et al.*, 2005], il est présenté un aperçu général des techniques permettant de trouver des fonctions de Lyapunov candidates. Une conséquence directe de ces études de stabilité est la conception d'observateurs d'état stables.

Remarque 1.2 *Dans la suite de ce document, nous nous intéresserons aux systèmes à changement de régime pour lesquels le passage d'un régime à un autre est brusque. Il n'y a donc pas de phase de transition d'un régime de fonctionnement à un autre. Nous désignerons ce type de systèmes par l'appellation générique de systèmes à commutations (SAC) et les différents régimes de fonctionnement du système seront appelés modes. De plus, les SAC considérés seront sujets à des changements de modes ou commutations de fréquences relativement moyennes. Ceci exclut de facto les systèmes susceptibles de commuter à chaque instant comme les hacheurs, par exemple, en électronique de puissance.*

Conclusion

Ce premier chapitre traite des principaux modèles utilisés dans la littérature pour la représentation des systèmes à changement de régime de fonctionnement.

On peut regrouper ces modèles en deux grandes classes : les modèles utilisant une représentation d'état et ceux s'appuyant sur une représentation sous forme de régression. Des passerelles existent, sous certaines conditions, entre ces deux grandes classes et également au sein des classes elles-mêmes. Un aperçu sur les méthodes d'identification et d'étude de stabilité est également donné.

2

Diagnostic de fonctionnement des systèmes à base de modèle – concepts fondamentaux

Sommaire

2.1 Définitions et généralités 68

2.1.1 Termes généraux 68

2.1.2 Fonctions 69

2.2 La détection 70

2.3 La localisation 72

2.3.1 Résidus structurés 73

2.3.2 Résidus directionnels 74

2.4 Méthodes de génération des résidus 76

2.4.1 L'approche par espace de parité 76

2.4.2 L'approche par observateurs 80
2.4.3 L'approche par estimation paramétrique 83
2.5 Diagnostic des systèmes à commutation 86
2.6 Observabilité des systèmes à commutation 91
2.6.1 Généralités 91
2.6.2 Observabilité dans le cas où l'évolution des modes est connue 94
2.6.3 Observabilité dans le cas où l'évolution des modes est inconnue 96

Introduction

Pour atteindre les objectifs d'automatisation des processus technologiques, on fait appel à des méthodes qui deviennent de plus en plus sophistiquées. La finalité de cette complexité croissante est l'augmentation de la performance, de la fiabilité, de la disponibilité et de la sûreté de fonctionnement de ces processus. Le besoin de sûreté de fonctionnement et de fiabilité est encore plus crucial lorsqu'il s'agit de systèmes sensibles pour lesquels une fausse manoeuvre peut être coûteuse aussi bien humainement que financièrement. Ceci est le cas, par exemple, des usines de produits chimiques, des réacteurs nucléaires, des systèmes de transport à grande vitesse, des systèmes aéronautiques et bien d'autres encore. En vue de remplir ces objectifs de performance, de sécurité et de disponibilité des processus technologiques, on leur associe des modules de diagnostic servant à détecter tout écart de comportement par rapport au comportement souhaité et même dans certaines situations à reconfigurer le fonctionnement du système.

Un module de diagnostic procède en trois étapes fondamentales. Dans un premier temps, il s'agit de déterminer si le processus considéré se comporte ou opère normalement, c'est-à-dire s'il remplit le cahier des charges qu'a fixé l'ingénieur lors de la conception du système en question. C'est l'étape dite de détection. Arrive ensuite une seconde étape qui a pour objet la recherche de certaines caractéristiques du défaut comme son instant d'apparition, son amplitude, sa gravité. Enfin, la dernière étape, qui découle des deux étapes précédentes, permet de décider de l'action à entreprendre sur le système. Il peut s'agir de maintenir le système dans le même mode opératoire, de corriger son fonctionnement ou encore de l'arrêter complètement.

En consultant la très abondante littérature existante sur le diagnostic, on se rend assez vite compte de la non unicité de la terminologie dans ce domaine. Le comité technique IFAC (*International Federation of Automatic Control*) SAFEPROCESS a tenté de normaliser certaines définitions généralement acceptées par l'ensemble de la communauté de l'Automatique. Nous rappellerons dans ce qui suit les définitions données à certains principaux termes propres au domaine du diagnostic.

2.1 Définitions et généralités

Les définitions qui suivent sont les fruits des travaux du Groupement pour la Recherche en Productique (*http ://www.laas.fr/ 7Ecombacau/SPSF-sursup.html*) dans le cadre des réflexions sur la terminologie en Surveillance et Supervision.

2.1.1 Termes généraux

Faute : action, volontaire ou non, dont le résultat est la non prise en compte correcte d'une directive, d'une contrainte exprimée par le cahier des charges.

Défaut : écart existant entre la valeur réelle d'une caractéristique du système et sa valeur nominale.

Défaillance : interruption permanente de la capacité d'un système à assurer une fonction requise dans des conditions opérationnelles spécifiées.

Erreur : partie du système ne correspondant pas ou correspondant incomplètement au cahier des charges. En toute logique, une erreur est la conséquence d'une faute.

Erreur latente : l'erreur est latente tant que la partie erronée du système n'est pas sollicitée. Elle devient effective au moment de la sollicitation de la partie erronée.

Dysfonctionnement : exécution d'une fonction du système au cours de laquelle le service rendu n'est pas délivré ou est délivré de manière incomplète.

Panne : état d'un système incapable d'assurer le service spécifié à la suite d'une défaillance.

Symptôme : événement ou ensemble de données au travers duquel le système de détection identifie le passage du procédé dans un fonctionnement anormal. C'est le seul élément dont a connaissance le système de surveillance au moment de la détection d'une anomalie.

Résidu : signal conçu comme indicateur d'anomalies fonctionnelles ou comportementales.

2.1.2 Fonctions

Acquisition : collecte des données en provenance du procédé.

Détection : caractérise le fonctionnement du système de normal ou d'anormal. On peut distinguer deux grandes classes d'anomalies :

- la première regroupe les situations pour lesquelles le comportement du système devient anormal par rapport à ses caractéristiques intrinsèques, la seconde regroupe les situations dans lesquelles le comportement est anormal par rapport à la loi de commande appliquée. A l'évidence, la deuxième classe est incluse dans la première.
- la deuxième classe recouvre les anomalies de fabrication mises en évidence par des contrôles (métrologie) de qualité du produit en fabrication.

Suivi : fonction maintenant en permanence un historique des traitements effectués par le système de commande/supervision et une trace des événements que perçoit le système.

Diagnostic : établit un lien de cause à effet entre un symptôme observé et la défaillance qui est survenue, ses causes et ses conséquences. On distingue classiquement trois étapes :

- localisation : détermine le sous-système fonctionnel à l'origine de l'anomalie et progressivement affine cette détermination pour désigner l'organe ou le dispositif élémentaire défectueux.
- identification : détermine les causes qui ont engendré la défaillance constatée.
- explication : justifie les conclusions du diagnostic.

Reconfiguration : fonction consistant à changer la commande envoyée au système ou la disposition matérielle du système pour éviter (ou faire face à) une panne.

Les systèmes pour lesquels la sûreté de fonctionnement est très cruciale sont le plus souvent conçus en se basant fortement sur la redondance matérielle. Par exemple deux (système duplex) ou trois capteurs (système triplex) sont souvent utilisés pour mesurer la même grandeur. Il s'agit de multiplier les chaînes de mesure afin d'obtenir plusieurs mesures de la même grandeur. Un détecteur et un

voteur permettent ensuite de donner la valeur la plus vraisemblable de la quantité mesurée. L'approche par redondance matérielle, bien qu'étant très efficace, présente l'inconvénient d'augmenter le coût (nécessité d'installer un plus grand nombre de capteurs) et l'encombrement (espace aditionnel requis pour l'installation des capteurs supplémentaires) des processus. Dans certaines applications, dans le domaine de l'aérospatial par exemple, le poids des équipements ainsi que leur coût constituent une contrainte critique du cahier des charges. Il est donc nécessaire, pour ce type d'application, de trouver une alternative à l'approche utilisant la redondance matérielle. L'alternative proposée consiste à utiliser les relations qui existent entre les mesures de grandeurs dépendantes qu'elles soient ou non de même nature : c'est la redondance analytique.

La redondance analytique exploite la corrélation existant entre les différents signaux mesurés. Sa mise en œuvre nécessite l'existence d'un modèle statique ou dynamique, linéaire ou non linéaire, déterministe ou stochastique du système reliant les entrées et les sorties mesurées. L'idée principale est de vérifier la cohérence des mesures de différentes grandeurs effectuées sur le processus et le comportement prévu pour ce processus à partir du modèle analytique qui lui est associé. L'approche utilisant la redondance analytique se compose essentiellement de deux phases distinctes : une phase de génération de résidus et une phase de prise de décision dont l'objet est la détection et éventuellement la localisation d'un composant défaillant. La figure 2.1 résume les principales étapes lors de la mise en œuvre d'un outil de diagnostic utilisant l'approche par redondance analytique.

2.2 La détection

La procédure de détection a pour objectif de déterminer l'apparition et l'instant d'occurence d'un défaut. Pour parvenir à cet objectif, on utilise des résidus qui sont obtenus en comparant le comportement du modèle du système à celui du système réel. Les résidus sont représentatifs des écarts entre le comportement observé du système et le comportement de référence attendu lorsque le système fonctionne « normalement ». Ces résidus sont généralement à moyenne nulle et ont une variance déterminée en l'absence de défauts de fonctionnement.

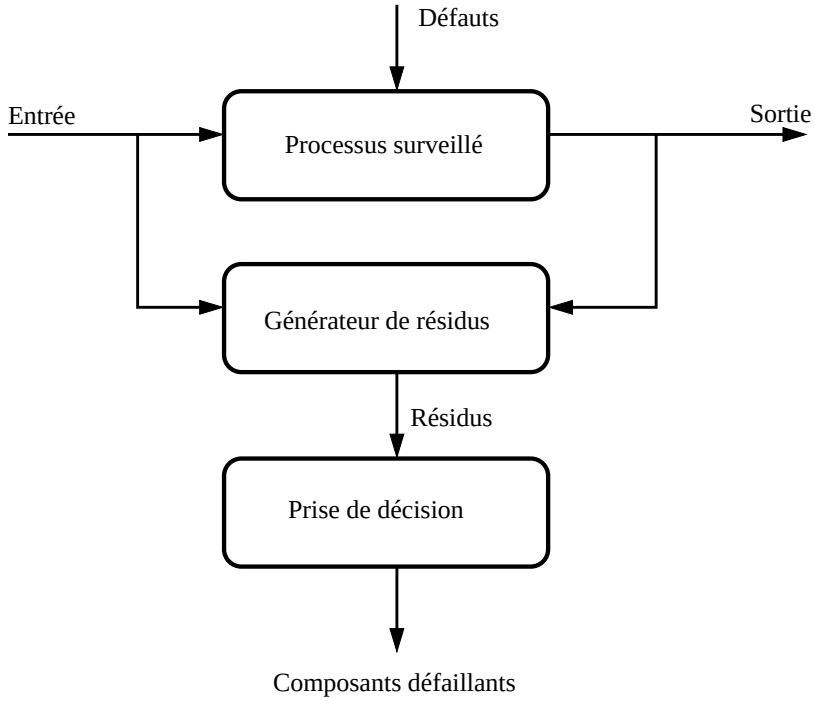


Fig. 2.1: Structure de diagnostic

Un moyen générique de construire un résidu est d'estimer le vecteur de sortie $y(\cdot)$ du système. L'estimé $\hat{y}(\cdot)$ est alors soustrait du signal de sortie $y(\cdot)$ afin de former le vecteur de résidus $r(\cdot)$ suivant :

$$r\left(k\right) = y\left(k\right) - \hat{y}\left(k\right) \tag{2.1}$$

En présence de défauts, le signal $r(\cdot)$ ainsi formé s'écartera notablement de la valeur zéro et sera identique à zéro lorsque le système fonctionne « normalement ».
Dans la pratique, le résidu n'a pas exactement une valeur nulle en l'absence de défauts car, lors de la phase de modélisation, plusieurs hypothèses simplificatrices sont introduites conduisant à un modèle qui ne reflète pas fidèlement le système réel. De plus, les mesures effectuées sur le système sont le plus souvent entachées de bruits de mesure. Le vecteur de résidus s'écrit alors :

$$r\left(k\right) = y_m\left(k\right) - \hat{y}\left(k\right) \tag{2.2}$$

où $y_m(\cdot)$ est la sortie mesurée du système qui est composée, en plus de la sortie réelle $y(\cdot)$, de bruits de diverses natures relatifs à l'instrumentation et aux incertitudes de modélisation. Dans cette situation, une méthode de détection élémentaire consiste à comparer la valeur du résidu à un seuil prédéfini ε (fonction des erreurs de modélisation). Une alarme est déclenchée à chaque franchissement de ce seuil :

$$\begin{cases} r(k) \leq \varepsilon \Longleftrightarrow d(k) = 0 \\ r(k) > \varepsilon \Longleftrightarrow d(k) \neq 0 \end{cases} \tag{2.3}$$

où $d(\cdot)$ représente le vecteur des défauts.
On peut également modéliser le résidu comme une variable aléatoire distribuée selon une loi normale. On met ainsi en œuvre à ce niveau des tests statistiques permettant de détecter des changements des caractéristiques statistiques du résidu. Une des méthodes les plus utilisées pour la détection de changement brusque d'une caractéristique statistique du résidu est la méthode dite CUSUM (*CUmulative SUM*). On peut trouver dans [Basseville et Nikiforov, 1993] plus de détails sur la théorie générale de la détection.

2.3 La localisation

Après avoir détecté la présence d'un défaut, il est important de situer l'élément affecté par ce défaut. Cette opération porte le nom de localisation ou d'isolation de défauts. Pour la réaliser, on procède à une structuration de l'ensemble des résidus générés de manière à assurer la localisation du défaut à partir des résidus affectés par ce défaut.
De façon générale, on construit en premier lieu un ensemble de résidus $r_i(\cdot)$ qui dépendent *a priori* de tous les défauts. Ces résidus sont appelés résidus de base. A partir de ces résidus de base, on forme ensuite des résidus plus « évolués » en rendant les résidus de base insensibles à certains défauts. On obtient ainsi deux types de résidus [Gertler, 1998, 1995; Patton, 1994; Patton *et al.*, 1989] : des résidus structurés [Ben-Haim, 1980; Chen *et al.*, 1995; Gertler, 1992, 1995; Patton, 1994] et des résidus de directions privilégiées [Beard, 1971; Chen *et al.*, 1995; Gertler et Monajemy, 1993; Hamelin *et al.*, 1994].

Dans le cas des résidus structurés, seul un ensemble spécifique de résidus sera sensible en présence d'un défaut. On peut, à titre d'exemple, imaginer que la structuration des résidus soit faite pour qu'un défaut $d_i(\cdot)$ agisse sur toutes les composantes du vecteur de résidus sauf le $i^{\text{ème}}$. Quant aux résidus de directions privilégiées, en présence de chaque défaut, le vecteur de résidus s'oriente dans une direction particulière. C'est donc la direction prise par le vecteur de résidus qui représente, dans ce cas, la signature du défaut.

2.3.1 Résidus structurés

Les résidus structurés sont construits de façon à être chacun affecté par un sous-ensemble de défauts et à être insensible aux autres défauts. Ainsi, pour un défaut donné, seule une partie des résidus réagit, c'est-à-dire s'écarte notablement de la valeur zéro, pour indiquer la présence de ce défaut.
La conception de tels résidus passe par deux étapes. Tout d'abord, il est nécessaire de définir les sensibilités ou robustesses désirées des résidus par rapport aux défauts à détecter ou à ne pas détecter. Puis, selon ces contraintes, il faut concevoir le générateur de résidus approprié.
Les informations de sensibilité et de robustesse souhaitées pour les résidus sont répertoriées dans une table binaire, appelée table des signatures théoriques. Pour construire cette table, on affecte, lorsque le $i^{\text{ème}}$ résidu doit être sensible (respectivement robuste) au $j^{\text{ème}}$ défaut, la valeur binaire 1 (respectivement 0) à la ligne et à la colonne correspondante de la table des signatures théoriques.
Une fois la table des signatures théoriques construite, on applique à chaque résidu une procédure de détection afin d'obtenir la signature réelle des résidus à chaque instant. Si cette signature réelle est nulle, alors le système est exempt de tout défaut et est donc déclaré sain. Lorsqu'intervient un défaut, au moins un des résidus générés est sensible à ce défaut et la signature réelle devient alors non nulle. La procédure de localisation consiste ensuite à faire la correspondance entre la signature réelle obtenue et les signatures présentes dans la table des signatures théoriques.

Exemple 2.1 (Tables des signatures théoriques) *On considère trois résidus r_1, r_2 et r_3 ainsi que trois défauts d_1, d_2 et d_3. La table 2.1 présente quatre tables*

de signatures théoriques ayant des propriétés de localisation différentes.
Dans la table 2.1a, les défauts d_1 et d_2 ne sont pas isolables car ils possèdent tous deux la même signature. Ainsi, il est impossible de faire la distinction entre ces deux défauts lorsqu'ils interviennent.
Dans la table 2.1b, tous les défauts sont isolables. Il faut constater que les signatures de d_1 et de d_2 ne diffèrent que d'un bit. On dira que les défauts sont isolables d'ordre 1.
Les tables 2.1c et 2.1d montrent des défauts isolables d'ordre 2. *Lors de la mise en place d'une procédure de diagnostic, la table 2.1d sera la plus simple à traiter car de manière générale, lorsque deux tables de signatures théoriques ont le même ordre d'isolabilité, la table contenant le plus de zéros est systématiquement retenue.*
La table 2.1a est qualifiée de table non localisante alors que la table 2.1b est dite faiblement localisante. Les tables 2.1c et 2.1d sont dites fortement localisantes.

	d_1	d_2	d_3
r_1	1	1	0
r_2	1	1	1
r_3	0	0	1

a)

	d_1	d_2	d_3
r_1	1	1	0
r_2	1	0	1
r_3	0	0	1

b)

	d_1	d_2	d_3
r_1	1	1	0
r_2	1	0	1
r_3	0	1	1

c)

	d_1	d_2	d_3
r_1	1	0	0
r_2	0	1	0
r_3	0	0	1

d)

Tab. 2.1: Tables de signatures théoriques

2.3.2 Résidus directionnels

Les résidus directionnels représentent une alternative aux résidus structurés. Ils sont construits de façon à ce que, en réponse à un défaut donné, le vecteur des résidus soit orienté suivant une direction privilégiée de l'espace des résidus (voir figure 2.2).

Le vecteur de résidus directionnels $\vec{r}(k)$, en réponse à un défaut $d_i(\cdot)$, s'explicite sous la forme :

$$\vec{r}\,(k\,|d_i\,) = \alpha_i\,(k)\,\vec{l}_i, \qquad i = \{1, \ldots, m\} \tag{2.4}$$

où $\vec{l}_i$ est un vecteur constant appelé signature directionnelle du défaut i dans l'espace des résidus et α_i est une fonction scalaire qui dépend de l'amplitude et de la dynamique du défaut.

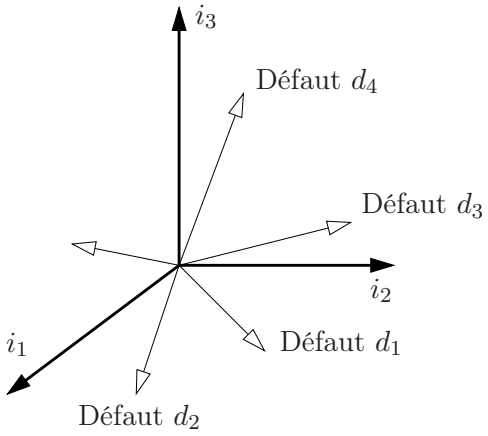


Fig. 2.2: Résidus directionnels

La localisation des défauts est effectuée en déterminant la signature directionnelle théorique la plus proche de la signature directionnelle obtenue par le calcul du vecteur des résidus. L'isolation des défauts est ici liée à la proximité des directions privilégiées de résidus. De fait, l'isolation des défauts ne sera possible que pour de grandes variations des projections du vecteur des résidus sur les directions privilégiées. La figure 2.3 illustre ce problème d'isolation de défauts. Les trois vecteurs en pointillé représentent les signatures directionnelles théoriques. Les vecteurs en trait plein représentent des signatures réelles du résidu à des instants différents. La signature réelle r_1 est très proche de la signature théorique du défaut 3. Il est donc probable que ce défaut soit présent sur le système à l'instant auquel r_1 a été calculé. Par contre, le résidu r_2 est plus difficile à évaluer car étant aussi proche de $\vec{l_1}$ que de $\vec{l_2}$.

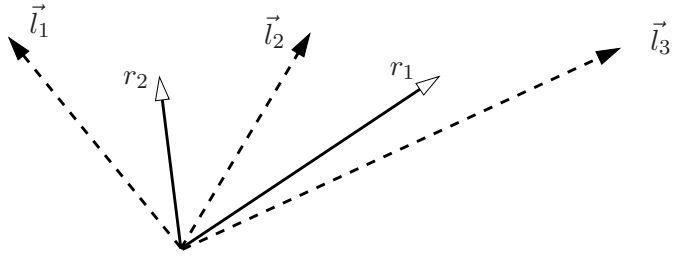


Fig. 2.3: Localisation à partir des résidus directionnels

2.4 Méthodes de génération des résidus

La génération de résidus caractéristiques du fonctionnement du système constitue le problème fondamental du diagnostic à base de modèle. Il existe dans la littérature une grande variété de méthodes pour la génération de résidus. Nous présentons ici quelques concepts de base à savoir l'approche par espace de parité, l'approche par observateurs et enfin l'approche par estimation paramétrique. Toutes ces méthodes reposent sur l'usage d'un modèle supposé exact du système et génèrent des résidus qui sont les écarts entre les signaux de sortie mesurés et leur estimation.

2.4.1 L'approche par espace de parité

L'approche par espace de parité repose sur l'utilisation de la redondance entre les entrées et les sorties du système indépendamment des états du système [Maquin, 2005]. Potter et Suman [1977] ont tout d'abord développé cette méthode pour les systèmes statiques. Ensuite, les travaux de Chow et Willsky au début des années 80 [Chow, 1980; Chow et Willsky, 1984] ont permis de généraliser cette approche aux systèmes dynamiques en utilisant les relations temporelles entre les sorties et les entrées du système dans le but de générer des résidus.

2.4.1.1 Espace de parité – cas statique

On considère une équation de mesure à l'instant k :

$$y_k = Cx_k + \varepsilon_k + Fd_k \tag{2.5}$$

$$x_{(\cdot)} \in \mathbb{R}^n, y_{(\cdot)} \in \mathbb{R}^m, d_{(\cdot)} \in \mathbb{R}^p, \varepsilon_{(\cdot)} \in \mathbb{R}^n$$
$$C \in \mathbb{R}^{m\times n}, F \in \mathbb{R}^{m\times p}$$

où $y_{(\cdot)}$ est le vecteur de mesure, $x_{(\cdot)}$ le vecteur des variables à mesurer, $d_{(\cdot)}$ le vecteur des défauts éventuels sur certains capteurs et $\varepsilon_{(\cdot)}$ le vecteur des bruits de mesure. La matrice d'observation C caractérise le système de mesure et la matrice F traduit la direction des défauts. On suppose que la matrice C est de rang n et que le nombre de mesures m est supérieur au nombre de variables n.

On définit le vecteur parité r_k projection du vecteur des mesures y_k :

$$r_k = W y_k \tag{2.6}$$

où W est une matrice de projection telle que : $WC = 0$.
L'orthogonalité de la matrice de projection W avec C conduit à :

$$r_k = W\varepsilon_k + WFd_k \tag{2.7}$$

L'expression (2.6), la forme de calcul du vecteur parité, permet le calcul numérique du vecteur parité à partir des mesures disponibles y_k tandis que l'expression (2.7), forme d'évaluation du vecteur parité, permet d'expliquer l'impact des erreurs de mesure et des défauts sur le vecteur parité. Il faut noter que dans le cas idéal (absence d'erreurs de mesure $\varepsilon_{(.)}$ et de défauts $d_{(.)}$) le vecteur parité est nul. Lors d'une défaillance d'un capteur, l'amplitude du vecteur parité évolue et s'oriente dans la « direction de défaillance » associée au capteur concerné. L'équation (2.8) traduit l'ensemble des relations de redondance qui lient les mesures y_k :

$$Wy_k = 0 \tag{2.8}$$

De nombreuses méthodes peuvent être employées pour la détermination de la matrice de projection W. On peut, par exemple, effectuer une élimination directe par substitution des inconnues. La matrice C, de rang n, peut être décomposée sous la forme :

$$C = \begin{pmatrix} C_1 \\ C_2 \end{pmatrix} \tag{2.9}$$

où C_1 est régulière. Une matrice orthogonale à C s'écrit alors simplement :

$$W = \begin{pmatrix} C_2 C_1^{-1} & -I \end{pmatrix} \tag{2.10}$$

où I est la matrice identité.
La démarche générale présentée ici s'étend aisément au cas de systèmes de me-

sure dont les variables sont contraintes, cette situation étant celle d'un processus caractérisé par un modèle et une équation de mesure.

2.4.1.2 Espace de parité – cas dynamique

On considère le modèle déterministe (2.11) :

$$\begin{cases} x(k+1) = Ax(k) + Bu(k) + F_1 d(k) \\ y(k) = Cx(k) + F_2 d(k) \end{cases} \tag{2.11}$$

$$x(\cdot) \in \mathbb{R}^n, y(\cdot) \in \mathbb{R}^m, u(\cdot) \in \mathbb{R}^r, d(\cdot) \in \mathbb{R}^p$$
$$A \in \mathbb{R}^{n\times n}, B \in \mathbb{R}^{n\times r}, C \in \mathbb{R}^{m\times n}, F_1 \in \mathbb{R}^{n\times p}, F_2 \in \mathbb{R}^{m\times p}$$

où $x(\cdot)$ est le vecteur d'état inconnu, $u(\cdot)$ et $y(\cdot)$ les vecteurs des entrées et sorties connus. On suppose, sans atteinte à la généralité, que les mesures $y(\cdot)$ dépendent seulement de l'état $x(\cdot)$ et ne font pas intervenir l'entrée $u(\cdot)$.
Sur un horizon d'observation $[k, k+h]$, les équations du système peuvent être regroupées sous la forme :

$$Y_{k,k+h} - T_h U_{k,k+h} = \mathcal{O}_h x(k) + F_h D_{k,k+h} \tag{2.12}$$

où les vecteurs $W_{k,k+h}$ avec $W \in \{Y, U, D\}$ et la matrice O_h sont définis par :

$$W_{k,k+h} = \begin{pmatrix} w(k) \\ w(k+1) \\ \vdots \\ w(k+h) \end{pmatrix} \quad \mathcal{O}_h = \begin{pmatrix} C \\ CA \\ \vdots \\ CA^h \end{pmatrix}$$

et avec les définitions suivantes :

$$T_h = \begin{pmatrix} 0 & 0 & \dots & 0 & 0 \\ CB & 0 & \dots & 0 & 0 \\ CAB & CB & \dots & 0 & 0 \\ \vdots & \vdots & \ddots & 0 & 0 \\ CA^{h-1}B & CA^{h-2}B & \dots & CB & 0 \end{pmatrix}$$

$$F_h = \begin{pmatrix} F_2 & 0 & \dots & 0 & 0 \\ CF_1 & F_2 & \dots & 0 & 0 \\ CAF_1 & CF_1 & \dots & 0 & 0 \\ \vdots & \vdots & \ddots & F_2 & 0 \\ CA^{h-1}F_1 & CA^{h-2}F_1 & \dots & CF_1 & F_2 \end{pmatrix}$$

L'entrée $u(\cdot)$ et la sortie $y(\cdot)$ du système étant connues, la seule inconnue dans l'équation (2.12) est l'état $x(k)$ du système. Afin de générer des relations de redondance entre l'entrée et la sortie du système, il est nécessaire que l'état inconnu $x(k)$ soit éliminé. Les équations de redondance qui lient $Y_{k,k+h}$ et $U_{k,k+h}$ indépendamment de $x(k)$ sont obtenues en multipliant l'équation (2.12) par une matrice W, appelée matrice de parité, orthogonale à $\mathcal{O}_h$:

$$W\mathcal{O}_h = 0 \tag{2.13}$$

dont l'existence est liée à la condition d'observabilité de l'état du système (2.11). Le vecteur de parité s'explicite donc en fonction des grandeurs connues (forme de calcul ou forme externe) :

$$r(k) = W(Y_{k,k+h} - T_h U_{k,k+h}) \tag{2.14}$$

ou sous forme interne en fonction des défauts :

$$r(k) = WF_h D_{k,k+h} \tag{2.15}$$

Le vecteur $r(k)$, appelé vecteur de parité généralisé, caractérise toutes les relations existant entre les entrées et les sorties du système. Il a une valeur nulle (en l'absence de bruit de mesure) si aucun défaut n'existe sur le système. En présence d'une défaillance d'un capteur ou d'un actionneur, le vecteur de parité devient différent de zéro et s'oriente dans une direction privilégiée en fonction du défaut, l'ensemble des directions étant constitué des colonnes de la matrice WF_h.

La recherche des équations de redondance peut être affinée en recherchant tout d'abord les équations de redondance pour chaque sortie prise isolément (auto-redondance), puis ensuite les relations de redondance entre différentes sorties (inter-

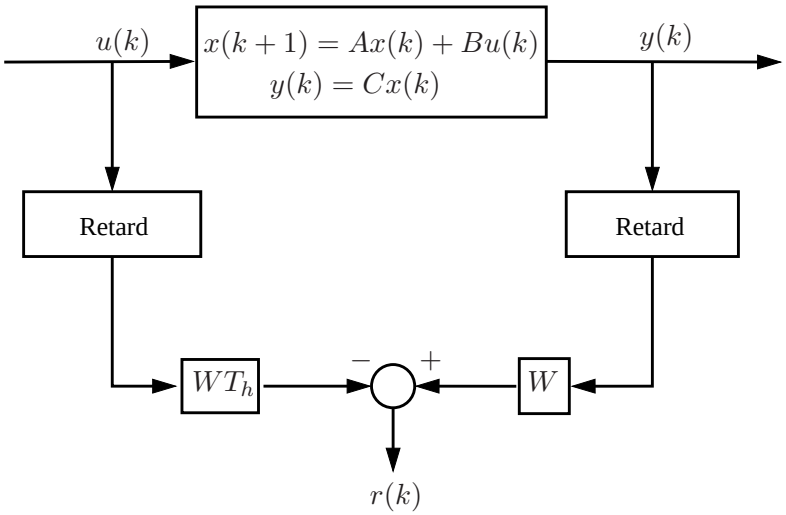


Fig. 2.4: Méthode de l'espace de parité

redondance). Cette hiérarchisation peut être mise à profit dans l'étape d'isolation des défauts affectant capteurs et actionneurs.

2.4.2 L'approche par observateurs

On considère le système (2.11) en l'absence de défaut :

$$\begin{cases} x\left(k+1\right) = Ax\left(k\right) + Bu\left(k\right) \\ y\left(k\right) = Cx\left(k\right) \end{cases} \tag{2.16}$$

$$x(\cdot) \in \mathbb{R}^n, y(\cdot) \in \mathbb{R}^m, u(\cdot) \in \mathbb{R}^r$$
$$A \in \mathbb{R}^{n\times n}, B \in \mathbb{R}^{n\times r}, C \in \mathbb{R}^{m\times n}$$

Les matrices A, B et C étant connues, on peut synthétiser un observateur afin de reconstruire l'état du système à partir des grandeurs connues $u(\cdot)$ et $y(\cdot)$. On utilise alors l'observateur (2.17) :

$$\begin{cases} \hat{x}\left(k+1\right) = A\hat{x}\left(k\right) + Bu\left(k\right) + Ke\left(k\right) \\ e\left(k\right) = y\left(k\right) - C\hat{x}\left(k\right) \end{cases} \tag{2.17}$$

où $\hat{x}(\cdot)$ est l'estimation de l'état $x(\cdot)$ du système, $e(\cdot)$ est l'erreur de reconstruction de la sortie.

L'évolution de l'erreur d'estimation d'état $e_x(\cdot)$ est régie par les équations (2.18) :

$$\begin{cases} e_x(k) = x(k) - \hat{x}(k) \\ e_x(k+1) = (A - KC)\, e_x(k) \end{cases} \tag{2.18}$$

Par un choix approprié de la matrice de gain K, on peut faire tendre asymptotiquement l'erreur d'estimation d'état vers zéro : $\lim_{k\to\infty} e_x(k) = 0$. Il suffit pour cela que la matrice de gain K soit choisie de manière à ce que la matrice $A - KC$ soit stable.

En présence de défauts, le modèle (2.16) devient :

$$\begin{cases} x(k+1) = Ax(k) + Bu(k) + F_1 d(k) \\ y(k) = Cx(k) + F_2 d(k) \end{cases} \tag{2.19}$$

$$x(\cdot) \in \mathbb{R}^n, y(\cdot) \in \mathbb{R}^m, u(\cdot) \in \mathbb{R}^r, d(\cdot) \in \mathbb{R}^p$$
$$A \in \mathbb{R}^{n\times n}, B \in \mathbb{R}^{n\times r}, C \in \mathbb{R}^{m\times n}, F_1 \in \mathbb{R}^{n\times p}, F_2 \in \mathbb{R}^{m\times p}$$

où $d(\cdot)$ est le vecteur des défauts.
L'erreur d'estimation d'état et celle de reconstruction de la sortie deviennent alors :

$$\begin{cases} e_x(k+1) = (A - KC)\, e_x(k) + (F_1 - KF_2)\, d(k) \\ e(k) = Ce_x(k) + F_2 d(k) \end{cases} \tag{2.20}$$

Ainsi, en présence d'un défaut, si la condition $F_1 - KF_2 \neq 0$ est vérifiée, l'erreur d'estimation d'état, et du coup l'erreur de reconstruction de la sortie, dévie notablement de la valeur zéro. On peut donc choisir $e(\cdot)$ comme signal de résidu.

En présence de bruits, on peut utiliser un filtre de Kalman [Jazwinski, 1970] à la place des observateurs classiques tel que celui présenté précédemment. On introduit à ce moment des hypothèses sur les propriétés statistiques des bruits.

Pour résoudre le problème de l'isolation des défauts, il convient de structurer les résidus, la situation idéale étant qu'un résidu soit sensible à un défaut particulier [Ragot, 2005]. Cette structuration, qui correspond à un découplage, peut être effectuée de différentes façons, soit en réglant le gain de l'observateur soit en utilisant des formes particulières d'observateurs construits à partir d'une partie

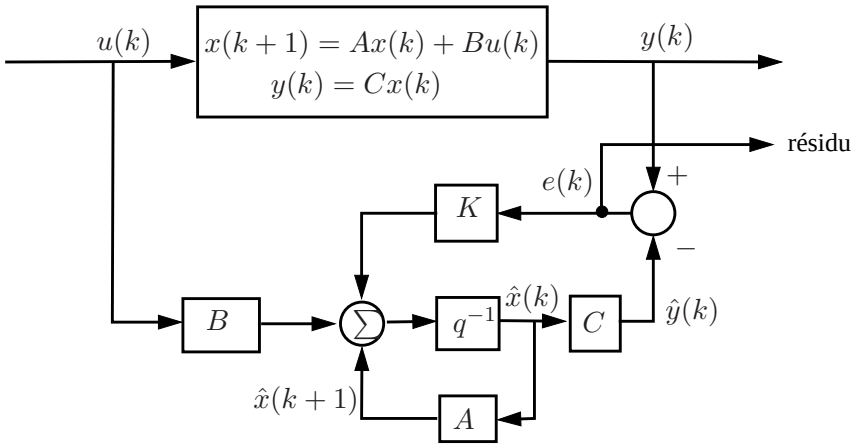


Fig. 2.5: Génération de résidus à partir d'un observateur

seulement des entrées et sorties du système. On trouve à cet effet les observateurs utilisant une partie des entrées et des sorties selon que l'on souhaite détecter des défauts d'actionneurs ou de capteurs. N'utilisant alors qu'une partie des informations disponibles, ces observateurs entrent dans la classe des observateurs à entrées inconnues.

La figure 2.6 illustre le principe de la détection de défaut d'actionneurs par observateurs dédiés (*Dedicated Observer Scheme* ou DOS). Le $i^{\text{ème}}$ observateur est piloté par la $i^{\text{ème}}$ entrée et toutes les sorties ; les $m-1$ autres entrées restantes sont considérées comme inconnues et la sortie de ce $i^{\text{ème}}$ observateur est insensible aux défauts des entrées non utilisées. La figure 2.7, *Generalized Observer Scheme*,

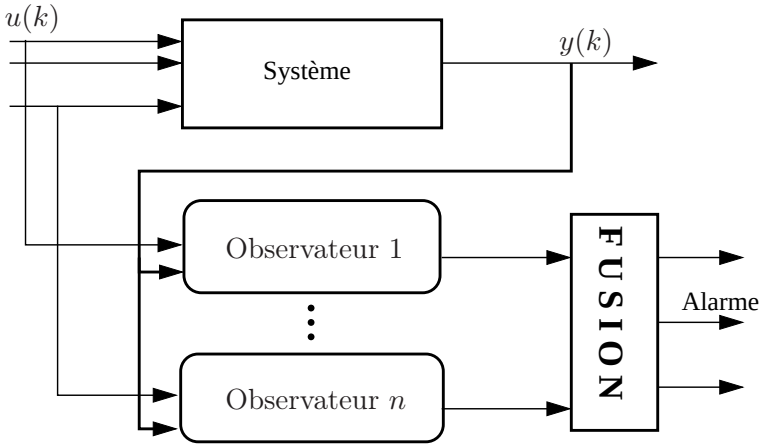


Fig. 2.6: Observateur à entrées inconnues (DOS) pour la détection de défauts d'actionneurs

est relative à une structure où le $i^{\text{ème}}$ observateur est piloté par toutes les entrées

sauf la $i^{\text{ème}}$ et toutes les sorties. La sortie de cet observateur est donc sensible aux défauts de toutes les entrées sauf ceux de la $i^{\text{ème}}$. Les figures 2.8 et 2.9 obéissent

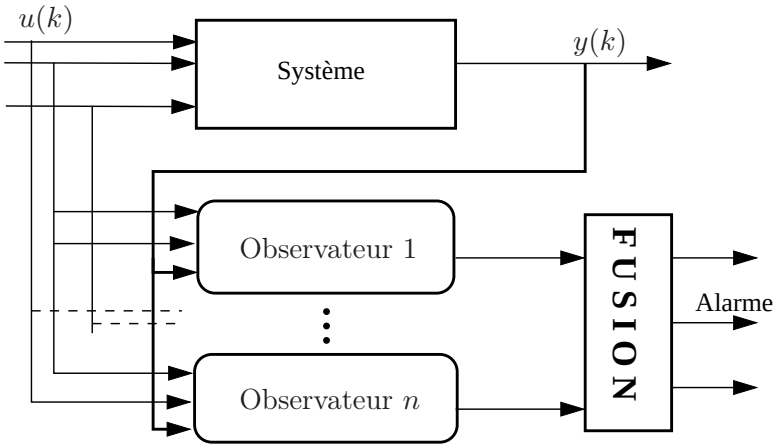


Fig. 2.7: Observateur à entrées inconnues (GOS) pour la détection de défauts d'actionneurs

aux mêmes principes de construction mais vis-à-vis des sorties.

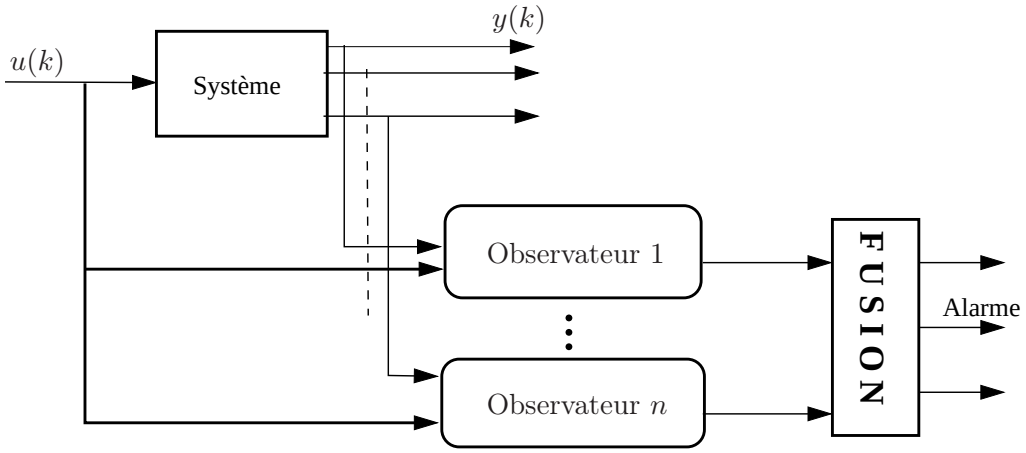


Fig. 2.8: Observateur à entrées inconnues (DOS) pour la détection de défauts de capteurs

2.4.3 L'approche par estimation paramétrique

Cette approche repose sur le principe selon lequel les effets de l'apparition d'un défaut sur le système se répercutent sur ses paramètres physiques [Simani *et al.*, 2003]. Il s'agit de procéder à une estimation en ligne des paramètres physiques du système surveillé en utilisant des méthodes classiques d'estimation et de comparer

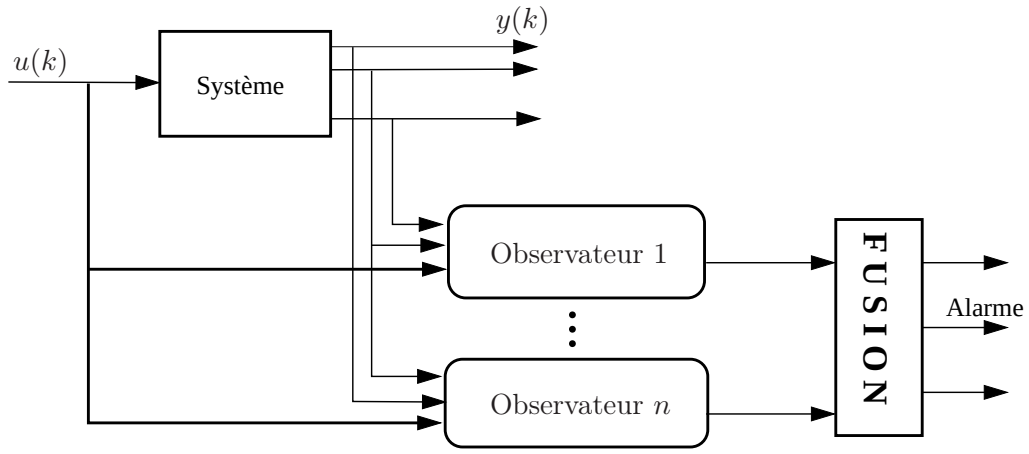


Fig. 2.9: Observateur à entrées inconnues (GOS) pour la détection de défauts de capteurs

ensuite cette estimation aux paramètres physiques réels du modèle obtenus initialement lors de son fonctionnement normal. Tout écart entre les paramètres physiques réels et estimés en ligne est considéré comme le symptôme de l'apparition d'un défaut.

Nous présentons maintenant deux méthodes de modélisation utilisées lors de la génération de résidus par estimation paramétrique : l'estimation paramétrique par minimisation de l'erreur d'équation et l'estimation paramétrique par minimisation de l'erreur de sortie.

2.4.3.1 Minimisation de l'erreur d'équation

Le modèle du processus est écrit sous la forme :

$$y\left(k\right) = \Psi^T \Theta \tag{2.21}$$

Θ et Ψ désignant les vecteurs de paramètres et des régresseurs :
$\Theta^T = [a_1 \dots a_n, b_1 \dots b_n]$, $\Psi^T = [y(k-1) \dots y(k-n)\ u(k-1) \dots u(k-n)]$.

On introduit l'erreur d'équation $e(\cdot)$ pour l'estimation paramétrique (voir figure 2.10). En définissant la fonction de transfert discrète du système par :

$$\frac{y\left(z\right)}{u\left(z\right)} = \frac{B\left(z\right)}{A\left(z\right)} = \frac{\sum_{i=1}^{n_b} b_i z^{-i}}{1 - \sum_{i=1}^{n_a} a_i z^{-i}}, \tag{2.22}$$

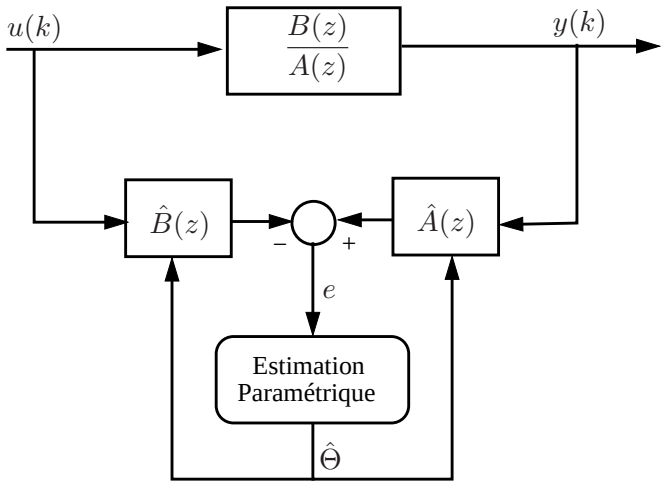


Fig. 2.10: Estimation paramétrique par minimisation de l'erreur d'équation

l'erreur d'équation est définie par :

$$e(z) = \hat{B}(z)u(z) - \hat{A}(z)y(z), \tag{2.23}$$

où $\hat{A}(z)$ et $\hat{B}(z)$ sont les estimations de $A(z)$ et $B(z)$.
On peut fournir une forme récursive de l'estimé $\hat{\Theta}$ en utilisant l'algorithme des moindres carrés récursifs [Isermann, 1992; Patton *et al.*, 2000] :

$$\hat{\Theta}(k+1) = \hat{\Theta}(k) + \gamma(k)e(k+1), \tag{2.24}$$

avec

$$\begin{cases} \gamma\left(k\right) = \dfrac{1}{\Psi^T\left(k+1\right)P\left(k\right)\Psi\left(k+1\right)+1}P\left(k\right)\Psi\left(k+1\right) \\ P\left(k+1\right) = \left[I - \gamma\left(k\right)\Psi^T\left(k+1\right)\right]P\left(k\right) \\ e\left(k\right) = y\left(k\right) - \Psi^T\left(k\right)\hat{\Theta}\left(k-1\right) \end{cases}$$

où I est la matrice identité.

2.4.3.2 Minimisation de l'erreur de sortie

Ici, au lieu de calculer comme précédemment l'erreur d'équation, on évalue plutôt l'erreur de sortie (voir figure 2.11) :

$$e(k) = y(k) - \hat{y}(\Theta, k) \tag{2.25}$$

où

$$\hat{y}(\Theta, z) = \frac{\hat{B}(z)}{\hat{A}(z)} u(z) \tag{2.26}$$

est la sortie du modèle.
L'erreur de sortie $e(\cdot)$ étant dans ce cas une fonction non linéaire des paramètres à estimer, on a alors recours à des méthodes numériques d'optimisation pour minimiser l'erreur $e(k)$ par rapport aux paramètres Θ.

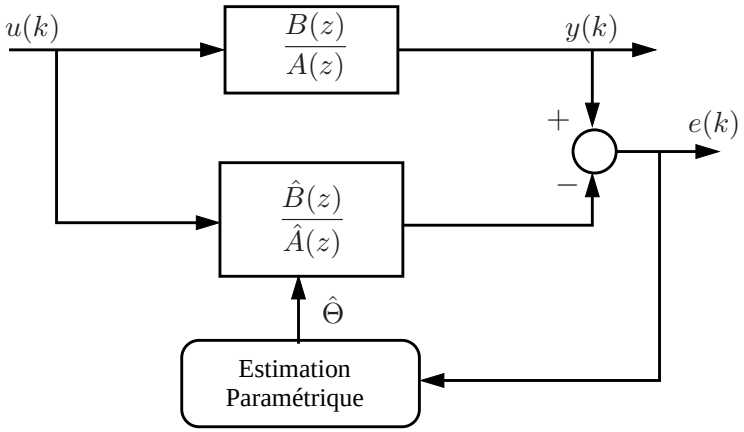


Fig. 2.11: Estimation paramétrique par minimisation de l'erreur de sortie

2.5 Diagnostic des systèmes à commutation

Considérons le système dynamique suivant :

$$\begin{cases} x(k+1) = Ax(k) + B_1 u_1(k) + B_2 u_2(k) \\ y(k) = Cx(k) \end{cases} \tag{2.27}$$

$$x(\cdot) \in \mathbb{R}^n, y(\cdot) \in \mathbb{R}^m, u_1(\cdot) \in \mathbb{R}, u_2(\cdot) \in \mathbb{R}$$
$$A \in \mathbb{R}^{n\times n},\ B_1 \in \mathbb{R}^{n\times 1},\ B_2 \in \mathbb{R}^{n\times 1} \text{ et } C \in \mathbb{R}^{m\times n}$$

Les actionneurs sont supposés être soit complètement opérationnels ou soit complètement non fonctionnels. A un instant quelconque k, on suppose que l'un des actionneurs du système est susceptible d'être sujet à l'apparition d'un défaut. La dynamique du système à tout instant peut alors être décrite par le système à commutation (SAC) suivant :

$$x(k+1) = Ax(k) + [B_1\ B_2]u, \quad y(k) = Cx(k) \tag{2.28a}$$
$$x(k+1) = Ax(k) + [0\ \ B_2]u, \quad y(k) = Cx(k) \tag{2.28b}$$
$$x(k+1) = Ax(k) + [B_1\ \ 0]u, \quad y(k) = Cx(k) \tag{2.28c}$$

où $u = [u_1\ \ u_2]^T$.
L'équation (2.28a) modélise le système en fonctionnement normal c'est-à-dire lorsque tous les actionneurs sont complètement opérationnels. Lorsque l'actionneur associé à la commande $u_1(\cdot)$ devient non fonctionnel, le système est modélisé par (2.28b). Enfin, lors de la présence d'un défaut sur l'actionneur délivrant à la commande u_2, la dynamique du système est décrite par (2.28c). Cet exemple montre que les modes d'un SAC peuvent représenter aussi bien des régimes de fonctionnement sain d'un système que des régimes de fonctionnement en présence d'un défaut. Ceci implique qu'avoir la connaissance du mode de fonctionnement dans lequel se trouve un SAC est indispensable à la détection de défaut sur le système.

Considérons maintenant l'exemple de la figure 2.12 relatif à une commande par retour d'état d'un système linéaire. Une contrainte est imposée sur l'entrée $u(\cdot)$ du système linéaire : $|u(k)| \leq 1$. Le système peut être représenté par le SAC suivant :

$$x\,(k+1) = (A - BL)\,x\,(k) + Bv\,(k)\,, \quad y\,(k) = x\,(k) \tag{2.29a}$$
$$x\,(k+1) = Ax\,(k) + B, \quad y\,(k) = x\,(k) \tag{2.29b}$$
$$x\,(k+1) = Ax\,(k) - B, \quad y\,(k) = x\,(k) \tag{2.29c}$$

L'équation (2.29a) modélise le système lorsque $|u(k)| \leq 1$. En saturation haute, $u(k) > 1$, le système est modélisé par (2.29b). Enfin, en saturation basse, $u(k) < 1$,

la dynamique du système est décrit par (2.29c). Contrairement à l'exemple précédent, tous les modes du SAC sont ici relatifs à des régimes de fonctionnement sans défaut sur le système. Si le mode dans lequel se trouve le système à chaque instant est connu, il est possible de concevoir un observateur afin de former des résidus dans le but de procéder à la détection de défauts sur le système.

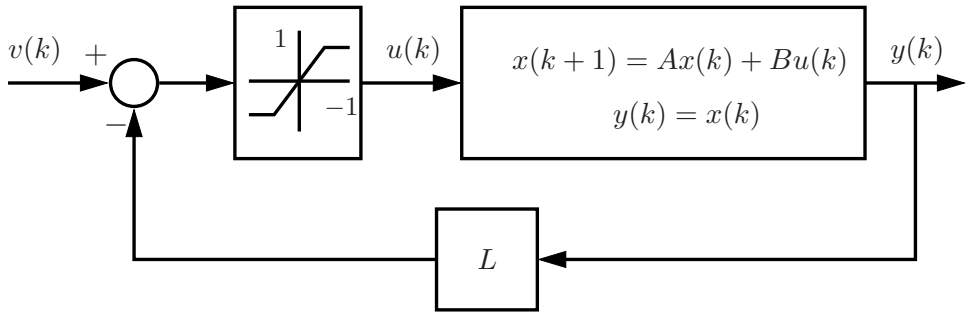


Fig. 2.12: Commande par retour d'état

Une analyse des deux exemples précédents conduit à relever une légère différence entre le diagnostic des systèmes linéaires à temps invariant et le diagnostic des SAC. En effet, la démarche générale est semblable dans les deux cas, mais ce que l'on recherche dans la phase de diagnostic différencie les deux classes de systèmes. Pour les SAC, on peut en effet distinguer deux situations selon que l'évolution des modes (où le mode dans lequel se trouve le système à l'instant courant) est connue ou non.

Lorsqu'on a la connaissance de l'évolution des modes, la procédure de diagnostic rejoint la démarche adoptée pour les systèmes linéaires invariants en utilisant l'approche par observateur. On synthétise un observateur d'état afin de reconstruire l'état et, partant de cette reconstrution, on construit des résidus en comparant la sortie mesurée du système à sa reconstruction à partir de l'état estimé.

Dans le cas où l'évolution des modes du système est inconnue, le système se trouve à chaque instant dans un mode inconnu, qui toutefois appartient à l'ensemble des modes du système qui a été au préalable répertorié. Le diagnostic nécessite alors tout d'abord l'implémentation d'une méthode permettant de retrouver à tout instant le mode actif du système.

La synthèse d'observateurs pour les SAC a fait, et continue de faire, l'objet

d'une attention considérable. La littérature est très abondante sur le sujet et la majorité des approches de synthèse existantes procède par extensions des observateurs développés pour les systèmes linéaires invariants. La première contribution importante portant sur l'estimation d'état des SAC est apparue dans [Ackerson et Fu, 1970] sous la forme d'un problème d'estimation d'état en milieu bruité. Dans cet article, les auteurs proposent une extension du filtre de Kalman aux situations où le bruit influençant le système n'est pas de nature gaussienne mais est plutôt modélisé par un ensemble de distributions gaussiennes avec des moyennes et variances différentes. Une seule des sources de bruit perturbe à chaque instant le système et le passage d'une source de bruit à l'autre est déterminé par une matrice de transition Markovienne. Ackerson et Fu [1970] utilisent une estimation par la méthode de Bayes pour filtrer les mesures et produire une estimation de l'état du système. Il est proposé également des algorithmes sous-optimaux qui donnent une bonne estimation de l'état du système tout en nécessitant des temps de calcul raisonnables. Dans [Tugnait, 1982], on retrouve un aperçu de l'état de l'art en 1982 quant au problème d'estimation d'état des JMLS. L'article présente et compare en simulation des algorithmes sous-optimaux d'estimation d'état, les algorithmes optimaux étant trop coûteux en temps et puissance de calcul. Toujours pour les systèmes à saut markovien (JML systems), Doucet *et al.* [2001] propose un algorithme appelé « particle filtering » permettant de procéder à l'estimation d'état. Dans [Alessandri et Coletta, 2001*a*,*b*, 2003; Bara *et al.*, 2000], l'estimation de l'état du système à commutation est effectuée sur la base de la connaissance du mode courant du système à chaque instant. L'observateur proposé dérive de l'observateur de Luenberger et garantit la stabilité asymptotique globale de l'erreur d'estimation, ceci évidemment dans la situation où la synthèse d'un tel observateur est réalisable. Balluchi *et al.* [2001] suppriment l'hypothèse de la connaissance *a priori* du mode courant du système à chaque instant. Il est alors proposé un observateur hybride composé de deux parties : une première partie destinée à l'estimation du mode courant du système à chaque instant et une seconde partie, qui utilise l'information provenant de la partie précédente, servant à estimer l'état du système. Balluchi *et al.* [2002] ont essayé de mettre sur pied une méthodologie pour la conception d'observateurs dynamiques pour les systèmes hybrides. L'observateur hybride mis

en œuvre est constitué de deux parties, comme dans [Balluchi *et al.*, 2001], à savoir une partie destinée à la reconnaissance du mode actif à partir des signaux mesurés sur le système et une seconde partie qui procède à l'estimation de l'état du système. Dans la même ligne de pensée, Juloski *et al.* [2003] présentent deux approches pour l'estimation d'état des systèmes à commutation. La première approche est une extension de l'observateur de Luenberger et la seconde approche est une extension de l'algorithme « Particle Filtering » aux SAC. Toujours pour des objectifs d'estimation d'état, une extension des observateurs à mémoire finie aux SAC est présentée dans [Kratz et Aubry, 2003] .

Une autre approche reposant sur une estimation sur un horizon glissant est proposée dans [Bemporad *et al.*, 1999; Ferrari-Trecate *et al.*, 2002] pour les modèles du type MLD. Cette approche, bien qu'étant applicable aux systèmes hybrides de façon générale, reste toutefois très gourmande en temps et puissance de calcul, ce qui rend difficile sa mise en œuvre.

Récemment, certains auteurs se sont intéressés plus particulièrement au problème de la reconnaissance du mode actif pour les SAC [Babaali et Egerstedt, 2005; Cocquempot *et al.*, 2003; Ragot *et al.*, 2003*a*]. Les approches proposées s'inspirent essentiellement des méthodes de diagnostic à base de modèle. Dans [Cocquempot *et al.*, 2003], il est proposé une méthode de reconnaissance du mode actif à chaque instant d'un SAC en utilisant des relations de redondance analytique. Babaali et Egerstedt [2005] reprennent le détecteur de mode synthétisé dans [Ragot *et al.*, 2003*a*] et proposent sur la base de ce détecteur, un observeur d'état garantissant la convergence asymptotique de l'estimation d'état sous certaines conditions.

Pour que toutes les méthodes de synthèse d'observateur ou de reconnaissance de mode fonctionnent correctement, il faut que le système en question soit observable. Pour les SAC, l'observabilité concerne aussi bien les modes du système que l'état du système.

2.6 Observabilité des systèmes à commutation

2.6.1 Généralités

Dans le cadre classique des systèmes linéaires invariants, un système est dit observable s'il existe un horizon temporel de taille h tel que les h observations $\{y(1), \ldots, y(h)\}$ déterminent de façon unique l'état initial $x(1)$. L'observabilité d'un système est déduite du calcul du rang de la matrice d'observabilité de Kalman : $\mathcal{O} = \left[C \; CA \cdots CA^{h-1} \right]^T$. Si le système est observable alors on a : $rang(\mathcal{O}) = n$, où n représente l'ordre du système.

En ce qui concerne les SAC, l'apparition des changements de mode augmente la complexité de l'analyse de l'observabilité du système. De manière analogue à d'autres champs d'étude tels que la stabilité et la synthèse de loi de commande [Blondel et Tsitsiklis, 1999; Lagarias et Wang, 1995; Liberzon et Morse, 1999; Liberzon *et al.*, 1999], l'étude de l'observabilité des SAC a fait récemment l'objet de nombreuses publications.

Dans [Hwang *et al.*, 2003; Ji et Chizeck, 1988; Mariton, 1986; West et Haddad, 1994], l'observabilité des systèmes à commutation est étudiée dans un cadre stochastique sur la base de la synthèse de lois de commande optimale et d'observateurs optimaux. Dans [Goshen-Meskin et Bar-Itzhack, 1992], partant de la connaissance du mode actif à chaque instant, une condition est formulée pour que le système soit observable. Bemporad *et al.* [2000] traitent le cas des systèmes PWA. Ils proposent des tests numériques permettant de vérifier en pratique l'observabilité du système.

La première tentative de généralisation de la notion d'observabilité des SAC a été présentée dans [Vidal *et al.*, 2002] où l'on trouve des conditions algébriques permettant de caractériser de façon systématique l'observabilité d'un SAC. Les conditions d'observabilité énoncées reposent sur le principe d'« indistingabilité » que nous allons rappeler. Pour cela, considérons le système autonome (absence d'entrée de commande) suivant :

$$\begin{cases} x(k+1) = & A_{\mu_k} x(k) + v(k) \\ y(k) = & C_{\mu_k} x(k) + w(k) \end{cases} \tag{2.30}$$

$$A_{\mu_k} \in \mathbb{R}^{n\times n}, C_{\mu_k} \in \mathbb{R}^{p\times n}$$

où $v(\cdot) \sim \mathcal{N}(0, Q_{\mu_k})$ et $w(\cdot) \sim \mathcal{N}(0, R_{\mu_k})$ sont des bruits de nature gaussienne. Dans (2.30), rappelons que la variable $\mu_{(.)}$ décrit l'évolution des modes du système à commutation.

Définition 2.1 (Indistingabilité)

Les états $\{x(k_0), \mu_{k_0}, \ldots, \mu_{k_0+h}\}$ *et* $\{\bar{x}(k_0), \bar{\mu}_{k_0}, \ldots, \bar{\mu}_{k_0+h}\}$ *sont indistingables sur l'horizon* $[k_0,\ k_0+h]$ *si les sorties* $\{y(k_0), \ldots, y(k_0+h)\}$ *et* $\{\bar{y}(k_0), \ldots, \bar{y}(k_0+h)\}$ *obtenues sur l'horizon d'observation en partant respectivement des états initiaux* $x(k_0)$ *et* $\bar{x}(k_0)$*, l'évolution des modes étant décrite respectivement par les séquences* $\{\mu_{k_0}, \ldots, \mu_{k_0+h}\}$ *et* $\{\bar{\mu}_{k_0}, \ldots, \bar{\mu}_{k_0+h}\}$*, sont identiques.*

L'ensemble des états indistingables de l'état $\{x(k_0), \mu_{k_0}, \ldots, \mu_{k_0+h}\}$ *est noté* $\mathcal{I}\{x(k_0), \mu_{k_0}, \ldots, \mu_{k_0+h}\}$.

Définition 2.2 (Observabilité)

Un état $\{x(k_0), \mu_{k_0}, \ldots, \mu_{k_0+h}\}$ *est observable si l'ensemble des états distingables de cet état se résume à l'état lui même :*

$\mathcal{I}\{x(k_0), \mu_{k_0}, \ldots, \mu_{k_0+h}\} = \{x(k_0), \mu_{k_0}, \ldots, \mu_{k_0+h}\}$.

Si tous les états du modèle (2.30) sont observables alors on dit que le modèle (2.30) est lui-même observable.

En désignant par $\{k_t,\ t \in \mathbb{N}^*\}$ l'ensemble des instants de commutation (changement de mode) du modèle (2.30), on définit $\tau_t = k_{t+1} - k_t$ comme le temps séparant deux commutations successives survenues aux instants k_t et k_{t+1}. On introduit la matrice d'observabilité jointe $\mathcal{O}_j(i,i') = [\mathcal{O}_j(i)\ \mathcal{O}_j(i')]$ avec :

$$\mathcal{O}_j(i) = \left[C_i^T\ (C_iA_i)^T\ \ldots\ \left(C_iA_i^{j-1}\right)^T \right]^T.$$

L'indice d'observabilité jointe des modes associés aux matrices d'observabilité $\mathcal{O}_j(i)$ et $\mathcal{O}_j(i')$ est défini comme étant le plus petit entier $\nu(i,i')$ tel que le rang de la matrice d'observabilité jointe $\mathcal{O}_j(i,i')$ cesse de croître lorsqu'on incrémente j. Soit ν le plus grand de tous les indices d'observabilité jointe : $\nu = \max\limits_{i\neq i'} \{\nu(i,i')\}$.

Théorème 2.1

Si pour tout $l \in \mathbb{N}^*$ *on a :*

$$\tau_l \geq 2\nu,$$

et pour tout $i \neq i' \in \{1, \ldots, s\},\ j' = 0, \ldots, j-1$ *et* $j \leq \nu$ *on a :*

$$\begin{aligned} &rang([\mathcal{O}_\nu(i) \quad \mathcal{O}_\nu(i')]) = 2n \\ &rang\left((\mathcal{O}_\nu\,(i) - \mathcal{O}_\nu\,(i'))\,A_i\right) = n \\ &rang(A_{i'}^{j} - A_{i'}^{j'} A_i^{j-j'}) = n, \end{aligned}$$

alors $\{x(k_0), \mu_{k_0}, \mu_{k_0+1}, \ldots, \mu_{k_0+h}\}$ *est observable sur l'horizon* $[k_0,\ k_0+h]$.

Les travaux de Vidal *et al.* [2002] ont été approfondis par la suite par Babaali et Egerstedt [2004] qui distinguent deux types d'observabilité : l'observabilité des modes et l'observabilité de l'état. Nous présentons dans la suite les résultats principaux des travaux de Babaali et Egerstedt [2004] qui formalisent bien le problème de l'observabilité des systèmes à commutation.

Considérons le système à commutation modélisé par l'équation (2.31) :

$$\begin{cases} x\,(k+1) = & A_{\mu_k} x\,(k) \\ y\,(k) = & C_{\mu_k} x\,(k) \end{cases} \tag{2.31}$$

On définit un chemin μ comme étant une séquence finie de modes c'est-à-dire :
$\mu = (\mu_1 \cdot \mu_2 \cdot \ldots \cdot \mu_N)$. On introduit ainsi les notations suivantes :

- la longueur du chemin μ, est notée $|\mu|$. Pour $\mu = (\mu_1 \cdot \mu_2 \cdot \ldots \cdot \mu_N)$, on a $|\mu| = N$.
- Θ_N est l'ensemble des chemins de longueur N.
- la concaténation de deux chemins μ et μ' est notée $\mu\mu'$
- le chemin μ^0 est un préfixe du chemin μ s'il existe un chemin μ^1 tel que $\mu = \mu^0\mu^1$.
- $\mu_{[i,j]}$ est l'infixe du chemin μ entre i et j, c'est-à-dire $\mu_{[i,j]} = (\mu_i \cdot \mu_{i+1} \cdot \ldots \cdot \mu_j)$
- $\Phi(\mu) = A_{\mu_N} \ldots A_{\mu_1}$ représente la matrice de transition associée au chemin μ.

– la matrice d'observabilité $\mathcal{O}_\mu$ associée au chemin μ est donnée par :

$$\mathcal{O}_\mu = \begin{pmatrix} C_{\mu_1} \\ \vdots \\ C_{\mu_N} A_{\mu_{N-1}} \dots A_{\mu_1} \end{pmatrix} \tag{2.32}$$

En supposant dans (2.31) que $x = x(1)$ et $\mu = (\mu_1 \cdot \mu_2 \cdot \ldots \cdot \mu_N)$, on a :

$$Y(\mu, x) = \mathcal{O}_\mu x(1) \tag{2.33}$$

avec $Y(\mu, x) = (y_1^T \dots y_N^T)^T$. La relation (2.33) sert à décrire de façon plus compacte le système autonome (2.31).

2.6.2 Observabilité dans le cas où l'évolution des modes est connue

On suppose dans cette section que les chemins μ suivis par les modes du système lors de son fonctionnement sont connus. On connaît donc à chaque instant le mode qui caractérise le système. Les mesures $Y(\mu, x)$ sont également disponibles. Il ne reste alors plus, dans ce cas, qu'à définir les conditions sous lesquelles il est possible de retrouver de façon unique l'état du système. La relation (2.33) montre qu'il suffit pour cela que $\mathcal{O}_\mu$ soit de plein rang colonne c'est-à-dire rang $(\mathcal{O}_\mu) = n$, n étant l'ordre du système. On définit ainsi la notion d'observabilité par chemin.
Pour les démonstrations des divers théorèmes et propositions énoncés, le lecteur peut se référer à [Babaali et Egerstedt, 2004].

Définition 2.3 (Observabilité par chemin)
Le système à commutation (2.31) est observable par chemin s'il existe un entier N tel que tout chemin de longueur N satisfasse la condition : rang $(\mathcal{O}_\mu) = n$, n étant l'ordre du système. Le plus petit des entiers N est appelé indice d'observabilité par chemin.

Exemple 2.2 (Observabilité par chemin) *Considérons le système modélisé par (2.31) avec :*

$$\begin{aligned} &\mu_k \in \{1,2\} \\ &A_1 = A_2 = \begin{pmatrix} 0 & 0.1 \\ 0 & 0.1 \end{pmatrix} \\ &C_1 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \qquad C_2 = \begin{pmatrix} 0 & 0 \\ -1 & 0 \end{pmatrix} \end{aligned} \tag{2.34}$$

Les matrices d'observabilité correspondant aux chemins (1) *et* (2) *de longueur* 1 *ne sont pas toutes de rang égal à l'ordre du système (voir tableau 2.2). Par*

chemin μ	matrice d'observabilité $\mathcal{O}_\mu$	rang$(\mathcal{O}_\mu)$
(1)	$\mathcal{O}_{(1)} = C_1$	2
(2)	$\mathcal{O}_{(2)} = C_2$	1

Tab. 2.2: Chemins, matrices d'observabilité et rangs

contre tous les chemins μ de longueur 2 *ont des matrices d'observabilité dont le rang est égal à l'ordre du système (voir tableau 2.3). On dira donc que le système est observable sur un chemin de longueur* 2.

chemin	matrice d'observabilité	rang
$(1 \cdot 1)$	$\mathcal{O}_{(1\cdot 1)} = \begin{pmatrix} C_1 \\ C_1A_1 \end{pmatrix}$	2
$(1 \cdot 2)$	$\mathcal{O}_{(1\cdot 2)} = \begin{pmatrix} C_1 \\ C_2A_1 \end{pmatrix}$	2
$(2 \cdot 1)$	$\mathcal{O}_{(2\cdot 1)} = \begin{pmatrix} C_2 \\ C_1A_2 \end{pmatrix}$	2
$(2 \cdot 2)$	$\mathcal{O}_{(2\cdot 2)} = \begin{pmatrix} C_2 \\ C_2A_2 \end{pmatrix}$	2

Tab. 2.3: Chemins, matrices d'observabilité et rangs

Ainsi, la reconstruction de l'état du système à commutation (2.31) est régie par l'existence d'un entier N tel que le système soit observable pour tout chemin de

longueur N. Babaali et Egerstedt [2004] ont prouvé qu'il existe une borne supérieure $N(s,n)$ pour l'indice d'observabilité par chemin. Cette borne est fonction du nombre de modes s et de l'ordre n du système.

2.6.3 Observabilité dans le cas où l'évolution des modes est inconnue

On suppose maintenant que seules les mesures $Y(\mu,x)$ sont disponibles. De ce fait, il est nécessaire de reconstuire simultanément l'évolution des modes, ce qui revient à retrouver les chemins, et l'état du système. Deux types d'observabilité sont donc étudiés : l'observabilité des modes et l'observabilité de l'état.

2.6.3.1 Observabilité des modes

Il s'agit ici de voir sous quelles conditions il est possible de retrouver un préfixe d'un chemin μ (c'est-à-dire $\mu_{[1,N']}$ pour $N^{'} < |\mu|$) à partir seulement des mesures $Y(\mu,x)$.

Définition 2.4 (Observabilité des modes)
Les modes du système à commutation (2.31) sont observables sur un chemin de longueur N s'il existe un entier $N^{'}$ tel que pour tout chemin $\mu \in \Theta_{N+N'}$ et pour presque tous les $x \in \mathbb{R}^n$,

$$\mu_{[1,N]} \neq \mu^{'}_{[1,N]} \Rightarrow Y\left(\mu,x\right) \neq Y\left(\mu^{'},x^{'}\right) \ \forall\, x^{'} \in \mathbb{R}^n \tag{2.35}$$

Le plus petit des entiers $N^{'}$ est appelé indice d'observabilité des modes pour N.

Exemple 2.3 *Considérons le système modélisé par (2.31) avec $\mu_k \in \{1,2\}$. Supposons que les modes de ce système soient observables sur un chemin de longueur 1 avec un indice d'observabilité des modes de 1. D'après la définition 2.4, pour vérifier cette hypothèse, il faut commencer par dénombrer tous les chemins de longueur $1+1=2$. L'ensemble de ces chemins est présenté à la table 2.4. Ensuite, on regarde si pour deux chemins différents dont les modes $\mu_1 = \mu_{[1,1]}$ sont également différents, les sorties correspondantes à l'activation de ces chemins sont différentes.*

μ^1	μ^2	μ^3	μ^4
$(1 \cdot 1)$	$(1 \cdot 2)$	$(2 \cdot 1)$	$(2 \cdot 2)$

Tab. 2.4: Ensemble des chemins de longueur 2

Il s'agit en fait de tester l'ensemble de relations (2.36) :

$$\begin{cases} Y(\mu_1, x) \neq Y(\mu_3, x') \\ Y(\mu_1, x) \neq Y(\mu_4, x') \\ Y(\mu_2, x) \neq Y(\mu_3, x') \\ Y(\mu_2, x) \neq Y(\mu_4, x') \end{cases} \tag{2.36}$$

où $x \in \mathbb{R}^n$ *et* $x' \in \mathbb{R}^n$.

En clair, on cherche à retrouver les N premiers modes (c'est-à-dire le chemin $\theta_{[1,N]}$) à partir de $N + N'$ mesures disponibles (c'est-à-dire $Y(\mu, x)$) et ceci pour presque tous les x. Il faut noter que la reconstruction des modes est envisagée pour presque tous les x et non pour tous les x. Ceci parce qu'il existe des valeurs de x pour lesquelles il est naturellement impossible de reconnaitre le chemin ayant conduit à la génération des mesures $Y(\mu, x)$. Par exemple, pour $x = 0$, il est impossible de faire une distinction entre les divers chemins à partir des mesures car quel que soit le chemin considéré, on obtient à partir de (2.33) : $Y(\mu, x) = 0$.
D'après (2.33), la reconstruction unique des modes est possible si on a : $\mu \neq \mu' \Rightarrow Y(\mu, x) \notin \mathcal{R}(\mathcal{O}'_\mu)$, où $\mathcal{R}(\cdot)$ est l'opérateur permettant d'évaluer l'espace engendré par les colonnes d'une matrice. On introduit ainsi le concept de discernabilité des chemins :

Définition 2.5 (Discernabilité des chemins)
Un chemin μ *est discernable d'un autre chemin* μ' *de même longueur si*

$$rang\left(\begin{bmatrix}\mathcal{O}_\mu & \mathcal{O}_{\mu'}\end{bmatrix}\right) > rang\left(\mathcal{O}_{\mu'}\right), \tag{2.37}$$

Le degré de discernabilité est défini par :

$$d = rang\left(\begin{bmatrix}\mathcal{O}_\mu & \mathcal{O}_{\mu'}\end{bmatrix}\right) - rang\left(\mathcal{O}_{\mu'}\right) \tag{2.38}$$

Le chemin μ est alors dit d-discernable du chemin $\mu^{'}$.

Exemple 2.4 (Discernabilité des chemins) *Considérons le système de l'exemple 2.3. Pour les chemins $(1 \cdot 1)$ et $(2 \cdot 2)$, on a rang $\left(\left[\mathcal{O}_{(1\cdot 1)} \quad \mathcal{O}_{(2\cdot 2)}\right]\right) = 4$ et donc rang $\left(\left[\mathcal{O}_{(1\cdot 1)} \quad \mathcal{O}_{(2\cdot 2)}\right]\right) >$ rang $\left(\mathcal{O}_{(2\cdot 2)}\right)$. Le chemin $(1 \cdot 1)$ est dit 2-discernable du chemin $(2 \cdot 2)$.*

Proposition 2.1

$Y(\mu, x) \notin \mathcal{R}\left(\mathcal{O}_{\mu'}\right)$ pour presque tous les $x \in \mathbb{R}$ si et seulement si μ est discernable de $\mu^{'}$.

Pour aller plus loin dans la discernabilité des chemins, il est proposé d'utiliser la notion de discernabilité-avant.

Définition 2.6 (Discernabilité-avant)

Etant donné un entier $d > 0$, un chemin μ est d-discernable-avant d'un autre chemin $\mu^{'}$ de même longueur s'il existe un entier N_d tel que pour toute paire de chemin λ et $\lambda^{'}$ de longueur N_d, $\mu\lambda$ et $\mu^{'}\lambda^{'}$ sont au moins d-discernables. Le plus petit des entiers N_d est l'indice de discernabilité-avant de μ par rapport à $\mu^{'}$.

Exemple 2.5 (Discernabilité-avant) *On vérifiera que pour le système de l'exemple 2.3, le chemin $\mu = (1)$ est 1-discernable-avant du chemin $\mu^{'} = (2)$. En effet, quels que soient les chemins λ et $\lambda^{'}$ de longueur 1, on montre aisément que rang $\left(\left[\mathcal{O}_{\mu\lambda} \quad \mathcal{O}_{\mu'\lambda')}\right]\right) \geq 3$.*

Proposition 2.2

$Y(\mu\lambda, x) \notin \mathcal{R}\left(\mathcal{O}_{\mu'\lambda'}\right)$ pour tout λ, $\lambda^{'} \in \Theta_{N'}$ et presque tous les $x \in \mathbb{R}$ si et seulement si μ est discernable-avant (ou 1-discernable-avant) de $\mu^{'}$ avec un index de discernabilité-avant inférieur à $N^{'}$.

Afin d'énoncer le théorème principal caractérisant l'observabilité des modes, on introduit la définition suivante :

Définition 2.7 (Discernabilité-avant complète)

Etant donné un entier $d > 0$, un chemin μ est dit complètement d-discernable-avant si μ est d-discernable-avant de tout autre chemin $\mu^{'}$ de même longueur. Le plus grand des indices de d-discernabilité de μ par rapport à tous les $\mu^{'} \neq \mu$ est l'indice de complète discernabilité-avant de μ.

Exemple 2.6 (Discernabilité-avant complète) *Pour le système de l'exemple 2.3, le chemin $\mu = (1)$ est complètement 1-discernable-avant. L'ensemble des chemins de longueur 1 étant constitué des chemins (1) et (2), et comme le chemin (1) est 1-discernable-avant du chemin (2), alors le chemin (1) est complètement 1-discernable-avant.*

Théorème 2.2

Les modes du système à commutation (2.31) sont observables sur un horizon de longueur N si et seulement si tout chemin de longueur N est complètement discernable-avant (ou complètement 1-discernable-avant).

L'indice d'observabilité des modes est le plus grand des indices de complète discernabilité-avant obtenu pour tous les chemins de longueur N.

2.6.3.2 Observabilité de l'état

L'observabilité de l'état met en exergue les conditions garantissant l'unicité de la reconstruction d'état.

Définition 2.8 (Observabilité de l'état)

L'état du système à commutation (2.31) est observable s'il existe un entier N (le plus petit étant l'indice d'observabilité de l'état du système) tel que $\forall x \in \mathbb{R}^n$ et $\forall \mu \in \Theta_N$,

$$x \neq x^{'} \Rightarrow Y(\mu, x) \neq Y\left(\mu^{'}, x^{'}\right) \quad \forall \mu^{'} \in \Theta_N \tag{2.39}$$

En d'autres termes, l'état du système (2.31) est observable si toute séquence de N mesures consécutives $Y(\mu, x)$ permet de reconstruire de façon unique l'état x du système, ceci sans la connaissance du chemin μ ayant permi de générer les mesures. La fonction qui au couple (x, μ) associe les mesures $Y(\mu, x)$ est donc injective en

la variable x. Une condition suffisante pour que cette fonction soit injective est la suivante :

Proposition 2.3

Si un système est observable par chemin avec un indice d'observabilité par chemin de N_o et si tout chemin de longueur N_o est complètement n-discernable-avant alors l'état du système est observable.

Pour approfondir l'étude de l'observabilité de l'état, on introduit le concept d'observabilité conjointe.

Définition 2.9 (Observabilité conjointe)

Deux chemins μ et $\mu^{'}$ de même longueur sont conjointement observables s'ils sont tous les deux observables et si la condition suivante est respectée :

$$\left(\mathcal{O}_{\mu}{}^{\{1\}} - \mathcal{O}_{\mu'}{}^{\{1\}}\right) P_{C\left(\mu,\mu'\right)} = 0 \tag{2.40}$$

$C\left(\mu,\mu^{'}\right)$ est défini par : $C\left(\mu,\mu^{'}\right) = \mathcal{R}\left(\mathcal{O}_{\mu}\right) \cap \mathcal{R}\left(\mathcal{O}_{\mu'}\right)$. La matrice $P_{C\left(\mu,\mu'\right)}$ est la matrice associée à la projection linéaire sur $C\left(\mu,\mu^{'}\right)$. La matrice $\mathcal{O}^{\{1\}}$ est la 1-inverse de $\mathcal{O}$ définie par $\mathcal{O}\mathcal{O}^{\{1\}}\mathcal{O} = \mathcal{O}$

Proposition 2.4

Deux chemins μ et $\mu^{'}$ sont conjointement observables pour tout x, $x^{'} \in \mathbb{R}^n$ si et seulement si :

$$x \neq x^{'} \Rightarrow Y\left(\mu,x\right) \neq Y\left(\mu^{'},x^{'}\right) \tag{2.41}$$

Définition 2.10 (Observabilité-avant conjointe)

Deux chemins μ et $\mu^{'}$ de même longueur sont conjointement observable-avant s'il existe un entier N tel que pour tout chemin λ et $\lambda^{'}$ de longueur N, les chemins $\mu\lambda$ et $\mu^{'}\lambda^{'}$ sont conjointement observables. L'indice d'observabilité-avant conjointe est le plus petit des entiers N.

Théorème 2.3

Les propositions suivantes sont équivalentes.

1. *L'état du système à commutation (2.31) est observable.*

2. *Le système à commutation (2.31) est observable par chemin avec un indice d'observabilité par chemin de N_o et toute paire de chemins différents de longueur N_o est conjointement observable-avant.*
3. *Le système à commutation (2.31) est observable par chemin et tout chemin minimal observable (c'est-à-dire un chemin sans aucun préfixe observable) est conjointement observable-avant avec tout autre chemin observable de même longueur.*

L'analyse menée jusqu'ici porte uniquement sur les systèmes autonomes. En ce qui concerne le cas des systèmes non autonomes (présence d'une entrée de commande dans le modèle du système), il est traité en procédant à une extension de l'analyse précédente. La nouveauté dans ce cas est la prise en compte de l'influence de l'entrée de commande sur l'observabilité (aussi bien des modes que de l'état) du système. Ceci se traduit par l'introduction de nouveaux concepts comme celui d'« entrée de discernabilité » [Babaali et Egerstedt, 2004].

Conclusion

Ce chapitre a été consacré aux principes fondamentaux du diagnostic de fonctionnement des systèmes à base de modèle. Diverses méthodes de génération de résidus ont été présentées. Le problème du diagnostic des systèmes à commutation a été également abordé. Ceci a conduit à l'étude de l'observabilité des systèmes à commutation qui, contrairement à celui des systèmes linéaires invariants, révèle quelques particularités.

3

Recherche du Mode Actif

Sommaire

3.1 Position du problème 104

3.2 Reconnaissance du mode actif 105

3.2.1 Détection du chemin actif : première formulation 105

3.2.2 Détection du chemin actif : seconde formulation 108

3.2.3 Sur le nombre de chemins 110

3.2.4 Raffinement de la procédure de détection du chemin actif en présence de bruit de mesure 111

3.3 Discernabilité des chemins 114

3.3.1 Exemple d'un système statique à commutation 115

3.3.2 Cas d'un système dynamique à commutation : absence de bruit de mesure 120

3.3.3 Cas d'un système dynamique à commutation : présence d'un bruit de mesure borné 124

Introduction

On trouve dans la littérature, une vaste panoplie d'outils de synthèse de loi de commande et d'analyse de stabilité pour les systèmes à commutation. Ces outils sont fondés sur la connaissance du mode de fonctionnement dans lequel se trouve le système à tout instant. La disponibilité de cette information est donc un point clé dans la mise en œuvre de ces outils ou du moins elle en simplifie l'implémentation. Ce chapitre traite du problème de la reconnaissance, à un instant particulier, du mode de fonctionnement d'un système à commutation (ce mode étant alors qualifié de mode actif), à partir d'un jeu de mesures formé de grandeurs mesurables du système, notamment son entrée et sa sortie. Il s'agit donc de concevoir un observateur permettant de reconstruire l'évolution des modes de fonctionnement du système sur un horizon de taille finie et de détecter ainsi les instants de commutation, c'est-à-dire les instants de passage d'un mode i à un mode j.
Nous conservons ici les définitions de Babaali et Egerstedt [2004] qui ont été introduites à la page 93 § 2.6.1.

3.1 Position du problème

Dans ce chapitre, le modèle considéré est celui de l'équation (3.1) :

$$\begin{cases} x(k+1) = A_{\mu_k} x(k) + Bu(k) \\ y(k) = Cx(k) \end{cases} \tag{3.1}$$

L'entrée, la sortie et l'état du système sont respectivement représentés par les variables $u(\cdot) \in \mathbb{R}^p$, $y(\cdot) \in \mathbb{R}^q$ et $x(\cdot) \in \mathbb{R}^n$. La variable $\mu_{(\cdot)}$, indexe toujours le mode actif à chaque instant k et prend ses valeurs dans un ensemble fini $S = \{1, \ldots, s\}$, où s représente le nombre de modes de fonctionnement du système. Sans atteinte à la généralité, les changements de mode sont traduits par des changements de valeurs de la matrice d'état du système, l'étude réalisée pouvant être étendue aux cas où les matrices B et C changent également de valeurs. Le problème posé est de retrouver le mode actif courant (ou la valeur prise par $\mu_{(\cdot)}$ à l'instant courant) à partir des mesures de l'entrée $u(\cdot)$ et de la sortie $y(\cdot)$ sur une fenêtre temporelle de

taille finie. Ce problème est équivalent à la détermination du chemin actif sur une fenêtre d'observation du système. Une fois le chemin actif retrouvé, la reconnaissance du mode actif est immédiate, de même que celle des instants de commutation du système.

3.2 Reconnaissance du mode actif

Comme cela a été mentionné précédemment, le problème de la reconnaissance du mode actif à tout instant peut être résolu en s'intéressant à la reconstruction du chemin actif sur un horizon d'observation de taille finie. Pour cette raison, on s'intéressera plus à la reconnaissance du chemin actif qu'à la reconnaissance du mode actif, le premier concept conduisant directement au second.

3.2.1 Détection du chemin actif : première formulation

Dans le cadre des systèmes linéaires invariants, il a été mis en exergue (page 78 § 2.4.1.2), en présence de défauts, l'existence d'une relation liant les mesures de l'entrée et de la sortie du système collectées sur un horizon d'observation de taille $h+1$ au vecteur des défauts et à l'état initial du système sur la fenêtre d'observation. En l'absence de défauts, cette relation s'explicite sur une fenêtre temporelle $[k-h,\ k]$ par :

$$Y_{k-h,k} - T_h U_{k-h,k} = \mathcal{O}_h x(k-h) \tag{3.2}$$

où les vecteurs $W_{k-h,k}$ avec $W \in \{Y, U\}$ et la matrice $\mathcal{O}_h$ sont définis par :

$$W_{k-h,k} = \begin{pmatrix} w(k) \\ w(k-1) \\ \vdots \\ w(k-h) \end{pmatrix} \quad \mathcal{O}_h = \begin{pmatrix} C \\ CA \\ \vdots \\ CA^h \end{pmatrix}$$

et avec les définitions suivantes :

$$T_h = \begin{pmatrix} 0 & 0 & \dots & 0 & 0 \\ CB & 0 & \dots & 0 & 0 \\ CAB & CB & \dots & 0 & 0 \\ \vdots & \vdots & \ddots & 0 & 0 \\ CA^{h-1}B & CA^{h-2}B & \dots & CB & 0 \end{pmatrix}$$

La relation (3.2) lie les entrées et sorties du système relevées sur la fenêtre d'observation à l'état initial $x(k-h)$, c'est-à dire l'état au début de la fenêtre d'observation. A partir de la définition d'une matrice de projection W telle que $W\mathcal{O}_h = 0$, on peut alors former un résidu $r(\cdot)$, indépendant de l'état initial $x(k-h)$, explicité par :

$$r(k) = W(Y_{k-h,k} - T_h U_{k-h,k}) \tag{3.3}$$

En l'absence de bruits et de perturbations, le résidu (3.3) a la propriété d'être nul à chaque instant.

Afin d'adapter la méthode de génération de résidus utilisée dans le cadre des systèmes linéaires au modèle (3.1), on procède à la réécriture des matrices $\mathcal{O}_h$ et T_h définies précédemment en incluant le fait que le système peut passer à chaque instant d'un mode de fonctionnement à un autre :

$$\mathcal{O}_{\mu,h} = \begin{pmatrix} C \\ CA_{\mu_{k-h}} \\ \vdots \\ C\underbrace{A_{\mu_{k-1}}A_{\mu_{k-2}}\cdots A_{\mu_{k-h}}}_{h} \end{pmatrix} \tag{3.4}$$

$$T_{\mu,h} = \begin{pmatrix} 0 & 0 & \cdots & 0 & 0 \\ CB & 0 & & 0 & 0 \\ \vdots & & & & \vdots \\ C\underbrace{A_{\mu_{k-1}}\cdots A_{\mu_{k-h+1}}}_{h-1}B & C\underbrace{A_{\mu_{k-1}}\cdots A_{\mu_{k-h+2}}}_{h-2}B & \cdots & CB & 0 \end{pmatrix} \tag{3.5}$$

La valeur de l'indice μ, h des matrices d'état dans les équations (3.4) et (3.5) est liée au chemin considéré sur la fenêtre d'observation et à la taille de cette fenêtre.

Exemple 3.1 *Considérons le modèle (3.1) avec $s = 2$ et pour $h = 3$, la fenêtre d'observation est $[k-3,\ k]$ à l'instant k. L'ensemble Θ_3 des chemins de longueur 3 sur cet horizon correspond à l'ensemble des huit chemins présentés à la table 3.1. Le*

Tab. 3.1: Ensemble des chemins de longueur 3

N° du chemin	1	2	3	4	5	6	7	8
μ_1	1	1	1	1	2	2	2	2
μ_2	1	2	1	2	1	2	1	2
μ_3	1	1	2	2	1	1	2	2

chemin n°4 de la table 3.1 (cinquième colonne de la table 3.1) indique qu'à l'instant $k-3$, la matrice d'état prend la valeur A_1 et qu'ensuite aux instants $k-2$ et $k-1$, elle prend la valeur A_2. Dans cette configuration, le chemin est $\mu = (1 \cdot 2 \cdot 2)$. Dans le cas général, on peut dénombrer s^h séquences différentes.

On énonce la proposition suivante.

Proposition 3.1

Les matrices d'observabilité $\mathcal{O}_{\mu,h}$ associées aux divers chemins de longueur h générés sur la fenêtre d'observation $[k-h,\ k]$ sont de plein rang : rang $(\mathcal{O}_{\mu,h}) = \dim(x) = n, \forall h \geq n$.

Grâce aux matrices d'observabilité et de Toeplitz définies respectivement par les équations (3.4) et (3.5), la relation (3.2) peut être modifiée sous la forme suivante :

$$Y_{k-h,k} - T_{\mu,h} U_{k-h,k} = \mathcal{O}_{\mu,h} x(k-h) \tag{3.6}$$

On définit alors le résidu :

$$r_{\mu,h}(k) = W_{\mu,h} \left(Y_{k-h,k} - T_{\mu,h} U_{k-h,k} \right), \tag{3.7}$$

la matrice $W_{\mu,h}$ vérifiant $W_{\mu,h}\mathcal{O}_{\mu,h} = 0$. Le calcul de ces résidus nécessite donc la détermination préalable de toutes les matrices $W_{\mu,h}$.

Théorème 3.1 (Détection du chemin actif sur une fenêtre d'observation) *Le chemin μ^*, dit chemin actif, caractérisant l'évolution des modes du système sur l'horizon $[k-h,\ k]$ est celui satisfaisant l'équation :*

$$r_{\mu^*,h}(k) = W_{\mu^*,h}\left(Y_{k-h,k} - T_{\mu^*,h}U_{k-h,k}\right) = 0 \tag{3.8}$$

Pour détecter ou reconnaître le chemin actif μ^* à partir des mesures, on doit procéder de la façon suivante :

- on construit sur l'horizon $[k-h,\ k]$ tous les chemins de longueur h possibles et l'on en déduit l'ensemble des matrices $\mathcal{O}_{\mu,h}$ correspondantes.
- on détermine ensuite à partir des matrices $\mathcal{O}_{\mu,h}$ les matrices de projection $W_{\mu,h}$.
- les matrices $\mathcal{O}_{\mu,h}$ et $W_{\mu,h}$ étant connues, on peut alors former les résidus $r_{\mu,h}(k)$ à partir des mesures faites sur le système.
- la séquence de commutation est détectée en testant l'ensemble des résidus $r_{\mu,h}(k)$, de manière à relever le résidu ayant une valeur nulle. Ce résidu correspond alors au chemin actif sur la fenêtre temporelle considérée.

3.2.2 Détection du chemin actif : seconde formulation

La seconde formulation de la solution au problème de la reconnaissance du chemin actif sur un horizon de taille $h+1$ passe par l'alternative de Fredholm.

Lemme 3.1 (Alternative de Fredholm)
Étant donné une matrice $A \in \mathbb{R}^{m\times n}$ et un vecteur $b \in \mathbb{R}^m$, il existe un vecteur $x \in \mathbb{R}^n$ tel que $Ax = b$ si et seulement si l'égalité (3.9) est vérifiée :

$$\left(I - AA^+\right) b = 0 \tag{3.9}$$

où A^+ est la pseudo-inverse (inverse de Moore-Penrose) de A définie par $AA^+A = A$.

A partir des équations (3.4), (3.5) et (3.2), le lemme 3.1 permet d'énoncer le théorème suivant :

Théorème 3.2 (Détection du chemin actif sur une fenêtre d'observation) *Le chemin μ^*, dit chemin actif, caractérisant l'évolution des modes du système sur l'horizon $[k-h,\ k]$ est celui satisfaisant l'équation :*

$$\left(I - \mathcal{O}_{\mu^*,h}\left(\mathcal{O}_{\mu^*,h}\right)^+\right)\left(Y_{k-h,k} - T_{\mu^*,h}U_{k-h,k}\right) = 0 \tag{3.10}$$

où I désigne la matrice identité.

L'algorithme 3.1 résume la méthode de détection du chemin actif. Cet algorithme comporte deux étapes principales : une première étape servant à la détermination des paramètres d'initialisation de l'algorithme et une seconde étape faite d'une procédure de calcul et d'analyse de résidus à chaque instant.

Algorithme 3.1

Étape 1. *Initialisation*
Collecter les données d'entrée et de sortie
Choisir la taille h de la fenêtre d'observation
Pour $\mu \in \Theta_h$:
Calculer les matrices d'observabilité $\mathcal{O}_{\mu,h}$ et leur inverse $(\mathcal{O}_{\mu,h})^+$
Calculer les matrices de Toeplitz $T_{\mu,h}$
Étape 2. *Génération des résidus*
A chaque instant k
Pour $\mu \in \Theta_h$:
Calculer $r_{\mu,h}(k) = \left(I - \mathcal{O}_{\mu,h}\left(\mathcal{O}_{\mu,h}\right)^+\right)\left(Y_{k-h,k} - T_{\mu,h}U_{k-h,k}\right)$
Analyser les résidus $r_{\mu,h}(k)$
Si $r_{\mu^*,h}(k) = 0$ alors le chemin μ^* est actif

Remarque 3.1 *Les conditions (3.8) et (3.10) des théorèmes 3.1 et 3.2 peuvent être exprimées sous la forme d'un test de rang de matrice. En effet à partir de (3.6), on peut écrire :*

$$R_{\mu,h}(k)\begin{pmatrix} -1 \\ x\left(k-h\right) \end{pmatrix} = 0 \tag{3.11}$$

avec $R_{\mu,h}(k) = \begin{bmatrix} Y_{k-h,k} - T_{\mu,h}U_{k-h,k} & \mathcal{O}_{\mu,h} \end{bmatrix}$.
La recherche du chemin actif peut être alors conduite en testant parmi les matrices

$R_{\mu,h}(\cdot)$ celle dont le rang est égal à la dimension du vecteur d'état : rang $(R_{\mu^*,h}(k)) = n$.

3.2.3 Sur le nombre de chemins

On se rend bien compte que le dénombrement des chemins possibles sur une fenêtre temporelle introduit un problème d'explosion combinatoire lié à la dimension des matrices d'état et à la taille de l'horizon d'observation. En effet, le nombre de résidus à calculer augmente avec la taille de l'horizon d'observation du système et avec le nombre de modes du système. Toutefois, lorsqu'à un instant k_0 quelconque, le chemin actif sur une fenêtre d'observation $[k_0 - h, k_0]$ a été identifié, les s^h résidus ne sont plus tous à tester à l'instant suivant $k_0 + 1$. On ne considère plus à l'instant $k_0 + 1$ que les chemins μ dont l'infixe $\mu_{[k_0+1-h,k_0-1]}$ est identique à l'infixe $\mu^*_{[k_0+1-h,k_0-1]}$ du chemin μ^* détecté précédemment à l'instant k_0.

Exemple 3.2 (Réduction du nombre de chemins) *Considérons le tableau 3.1 de l'exemple 3.1 qui présente l'ensemble des chemins de longueur 3 pour un système à commutation à deux modes et pour $h = 3$. Si, à un instant k_0, le chemin identifié est le chemin $\mu^* = (1 \cdot 2 \cdot 1)$ (chemin n° 2) alors les chemins à considérer à l'instant suivant k_0+1 sont les chemins $\mu = (2 \cdot 1 \cdot 1)$ (chemin n° 5) et $\mu = (2 \cdot 1 \cdot 2)$ (chemin n° 7).*

Le nombre de chemins à considérer à chaque instant peut être également réduit en formulant l'hypothèse d'une seule commutation par fenêtre d'observation. En d'autres termes, on suppose que la fréquence de commutation d'un mode à un autre est telle que la durée minimale séparant deux commutations successives est supérieure à la taille de la fenêtre d'observation. L'application de cette hypothèse à l'exemple présenté dans la table 3.1 permet de passer de 8 séquences de commutations à 6 car les séquences $\mu = (1 \cdot 2 \cdot 1)$ (chemin n° 2) et $\mu = (2 \cdot 1 \cdot 2)$ (chemin n° 7) deviennent irréalisables. Pour aller plus loin, on peut également limiter le nombre de chemins à prendre en compte en n'utilisant que les chemins décrivant le système lorsque ce dernier reste dans le même mode sur toute la durée de la fenêtre d'observation. Il s'agirait ainsi, pour l'exemple présenté dans la table 3.1,

de ne considérer que les séquences $\mu = (1 \cdot 1 \cdot 1)$ (chemin n° 1) et $\mu = (2 \cdot 2 \cdot 2)$ (chemin n° 8). Néanmoins dans ce cas, la réduction du nombre de résidus à générer s'obtient au prix d'un retard à la détection de changement de mode puisque tous les chemins de longueur h possibles sur la fenêtre d'observation ne sont pas considérés. La reconnaissance d'un chemin décrivant le système lorsqu'il reste dans le même mode sur tout l'horizon d'observation ne peut alors avoir lieu tant que l'instant de commutation se situe dans l'horizon d'observation. On observe donc un retard maximal à la détection de changement de mode égal à la taille de la fenêtre d'observation.

La connaissance d'informations *a priori* sur le processus peut également permettre de limiter le nombre de résidus à générer. Il s'agit d'informations comme les séquences de commutations « interdites » ou impossibles.

3.2.4 Raffinement de la procédure de détection du chemin actif en présence de bruit de mesure

La reconnaissance du chemin actif à tout instant a été effectuée jusqu'ici dans un cadre déterministe, exempt de tout bruit de mesure. On suppose maintenant que la sortie du système est entachée d'un bruit de mesure borné. On suppose que la seule information disponible est la borne maximale de l'amplitude du bruit de mesure. Aucune hypothèse probabiliste n'est formulée quant à la distribution du bruit de mesure.

3.2.4.1 Cas d'un bruit borné

On suppose la présence d'un bruit de mesure borné $n(\cdot)$ sur la sortie du système décrit par (3.1) :

$$\begin{cases} x(k+1) = A_{\mu_k} x(k) + Bu(k) \\ y(k) = Cx(k) + n(k) \\ \forall k, |n(k)| \leq \delta, \quad \delta > 0 \end{cases} \tag{3.12}$$

où δ est l'amplitude maximale du bruit de mesure $n(\cdot)$.

Dans ce cas, le résidu $r_{\mu^*,h}(\cdot)$, défini en (3.8) et correspondant au chemin μ^* actif sur la fenêtre temporelle $[k-h,\ k]$, ne prend plus une valeur nulle. En effet, l'expression

du résidu $r_{\mu^*,h}(\cdot)$ à partir de (3.8) devient :

$$r_{\mu^*,h}(k) = W_{\mu^*,h}\left(Y_{k-h,k} - T_{\mu^*,h}U_{k-h,k} + N_{k-h,k}\right) \tag{3.13}$$

où $N_{k-h,k}$ est le vecteur correspondant à l'empilement du bruit de mesure sur la fenêtre d'observation $[k-h,\ k]$. Comme $W_{\mu^*,h}\left(Y_{k-h,k} - T_{\mu^*,h}U_{k-h,k}\right) = 0$, on peut écrire :

$$r_{\mu^*,h}(k) = W_{\mu^*,h}N_{k-h,k} \tag{3.14}$$

En utilisant la bornitude de l'amplitude du bruit de mesure, on définit un résidu intervalle $[r_{\mu^*,h}(k)]$ [Adrot, 2000] :

$$[r_{\mu^*,h}(k)] = [\underline{r}_{\mu^*,h}\ ,\ \bar{r}_{\mu^*,h}] \tag{3.15}$$

où les bornes $\underline{r}_{\mu^*,h}$ et $\bar{r}_{\mu^*,h}$ dépendent de la borne δ du bruit de mesure et sont données par : $\underline{r}_{\mu^*,h} = -\left|W_{\mu^*,h}\right|\mathbb{U}\delta$ et $\bar{r}_{\mu^*,h} = \left|W_{\mu^*,h}\right|\mathbb{U}\delta$ avec $\mathbb{U}$ un vecteur colonne de dimension égale au nombre de colonnes de $W_{\mu^*,h}$ et dont tous les éléments sont égaux à 1.
Dans un contexte intervalle, la détermination du chemin actif revient à rechercher lequel des chemins $\mu \in \Theta_h$ correspond à un résidu intervalle $[r_{\mu,h}(\cdot)]$ incluant la valeur zéro. Ce test peut être conduit en calculant le signe du produit des bornes inférieure et supérieure de chacun des résidus intervalle $[r_{\mu,h}(\cdot)]$. Le résidu intervalle $[r_{\mu,h}(\cdot)]$ correspondant au chemin actif μ^* est celui pour lequel ce produit est négatif.

En fonction des dynamiques des différents modes de fonctionnement du système, il peut arriver que l'on s'écarte de la situation idéale dans laquelle un seul des résidus intervalle $[r_{\mu,h}(\cdot)]$ contient la valeur zéro et être confronté à la présence de plus d'un résidu intervalle contenant la valeur zéro[†]. Dans cette situation, on s'abstient de prendre une décision quant au chemin actif et éventuellement on énumère l'ensemble des chemins candidats. Pour aller plus loin dans l'analyse, on peut reconsidérer les situations d'indécision en assortissant le résultat d'un indice témoignant de la confiance qu'on peut avoir quant à l'exactitude de la reconnaissance

[†]Cette situation est liée à la discernabilité des chemins μ de longueur h sur l'horizon $[k-h,\ k]$. L'étude des conditions de discernabilité des chemins, aussi bien dans le cas déterministe qu'en présence de bruit de mesure, fera l'objet de la section 3.3

de chemin effectuée. Pour cela, il faut examiner la position relative des résidus intervalle ambigus par rapport à des résidus intervalle de référence définis dans les situations où leur chemin correspondant est actif sur l'horizon considéré.

Supposons qu'à un instant k quelconque, l'ambiguïté sur le chemin actif porte sur deux chemins μ^1 et μ^2. On peut tracer dans le plan des résidus $r_{\mu^1,h}(\cdot)$ et $r_{\mu^2,h}(\cdot)$ les valeurs prises par ces résidus lorsque leur chemin correspondant, μ^1 et μ^2, est actif sur l'horizon d'observation. Ce domaine, appelé domaine de référence, est un rectangle centré sur l'origine du plan des résidus. Il correspond à l'ensemble des situations conduisant à une ambiguïté portant sur les deux chemins μ^1 et μ^2. Le domaine de référence est obtenu en effectuant le produit cartésien des intervalles $[r_{\mu^1,h}]$ et $[r_{\mu^2,h}]$, dénommés respectivement intervalle de référence associé au résidu $r_{\mu^1,h}(\cdot)$ et intervalle de référence associé au résidu $r_{\mu^2,h}(\cdot)$. L'intervalle de référence $[r_{\mu^i,h}]$, $i \in \{1,2\}$, représente l'ensemble des valeurs prises par le résidu $r_{\mu^i,h}(\cdot)$ lorsque μ^i est actif. Sur les graphiques de la figure 3.1, les régions hachurées sont les domaines de référence et contiennent l'ensemble des valeurs prises par les résidus $r_{\mu^1,h}(\cdot)$ et $r_{\mu^2,h}(\cdot)$ lorsque leur chemin correspondant est actif sur la fenêtre d'observation.

Partant du domaine de référence, il convient de tester le positionnement des résidus intervalle par rapport à ce domaine. La figure 3.1 montre, pour deux chemins μ^1 et μ^2, deux situations où les résidus intervalle correspondants, $[r_{\mu^1,h}(\cdot)]$ et $[r_{\mu^2,h}(\cdot)]$, contiennent tous deux la valeur zéro. Le domaine hachuré correspond, comme mentionné plus haut, au domaine de référence. La région grisée est le domaine défini par les résidus intervalle $[r_{\mu^1,h}(\cdot)]$ et $[r_{\mu^2,h}(\cdot)]$ à un instant k_0 quelconque. Sur les deux graphiques de la figure 3.1, on peut noter la présence d'une zone de recouvrement entre le domaine de référence et les domaines correspondant aux résidus intervalle à l'instant k_0. Le symbole "o" pointe sur le centre des domaines associés aux résidus intervalle calculés à l'instant k_0.

En fonction de la taille de la zone de recouvrement, on assortit la décision prise quant au chemin actif d'un degré d'incertitude. Le calcul du degré d'incertitude associé à un résidu intervalle $[r_{\mu,h}(\cdot)]$ s'effectue de la manière suivante :

$$\begin{cases} d_{r_{\mu,h}(k)} = \dfrac{w\left([r_{\mu,h}(k)]\right) - w\left([r_{\mu,h}(k)] \cap [r_{\mu,h}]\right)}{w\left([r_{\mu,h}(k)]\right)} \\ \text{avec pour } \ [a] = [\underline{a}\ ,\ \overline{a}]\,, \quad w\left([a]\right) = \overline{a} - \underline{a} \end{cases} \tag{3.16}$$

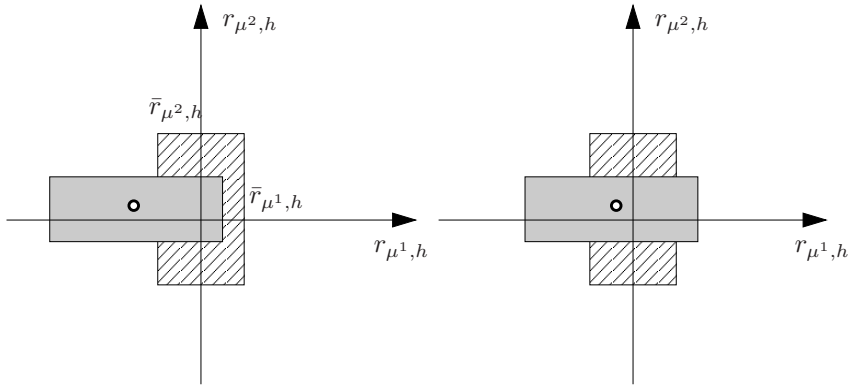


Fig. 3.1: Résidus ambigus

Dans (3.16), $[r_{\mu,h}]$ représente l'intervalle de référence associé au résidu $r_{\mu,h}(\cdot)$. C'est donc l'ensemble des valeurs pouvant être prises par $r_{\mu,h}(k)$ lorsque le chemin μ est actif et cet ensemble est défini par :

$$[r_{\mu,h}] = \left[\underline{r}_{\mu,h}\ ,\ \bar{r}_{\mu,h}\right] \tag{3.17}$$

Le degré d'incertitude $d_{r_{\mu,h}(k)}$ varie entre 0 et 1. Si $d_{r_{\mu,h}(k)} = 0$ alors toutes les valeurs pouvant être prises par le résidu $r_{\mu,h}(k)$ sont comprises dans le domaine décrit par le résidu de référence $[r_{\mu,h}]$. De ce fait, le chemin μ est inévitablement le chemin actif. De façon générale, on retiendra de tous les résidus intervalle conduisant à une situation d'ambiguïté, celui dont le degré d'incertitude est le plus faible et on associera à cette décision un degré d'incertitude synonyme de la confiance que l'on peut avoir en cette décision.

3.3 Discernabilité des chemins

La discernabilité des chemins concerne les conditions sous lesquelles il est possible de retrouver de façon unique et sans ambiguïté le chemin actif sur un horizon de taille donnée à partir de la connaissance de l'entrée et de la sortie du système. Le terme « unique » signifie qu'il est impossible d'avoir deux chemins différents qui induisent le système dans la même dynamique et « sans ambiguïté » dans le sens où il est garanti que le chemin identifié est effectivement le chemin actif sur l'horizon d'observation. Les conditions de discernabilité garantissent l'unicité du

chemin actif μ^* sur l'horizon d'observation lors de la phase de reconnaissance du mode actif.

Dans le chapitre 2, page 96 § 2.6.3, le problème de discernabilité des chemins, dans le cas où l'évolution des modes est inconnue, a été abordé en analysant les conditions sous lesquelles des chemins différents peuvent induire le système dans la même dynamique, notamment à travers la sortie du système. La reconnaissance du chemin actif étant effectuée ici à partir de l'analyse de résidus associés aux différents chemins générés sur un horizon d'observation, la discernabilité des chemins est étudiée en cherchant à caractériser les situations pouvant conduire à l'obtention de résidus identiques (de valeur nulle) pour des chemins différents. L'étude de la discernabilité des chemins commencera par un exemple illustratif portant sur le cas d'un système statique à commutation, en l'absence et en présence d'un bruit de mesure sur la sortie, et sera ensuite étendue aux systèmes dynamiques à commutation. Dans ce dernier cas, deux situations seront également considérées. Dans la première, le système est supposé fonctionner en l'absence de tout bruit de mesure et dans la seconde on suppose l'existence d'un bruit de mesure borné sur la sortie du système.

3.3.1 Exemple d'un système statique à commutation

On considère un système statique à commutation représenté par le modèle de l'équation (3.18) :

$$y(k) = K_i u(k) \tag{3.18}$$

où $K_i \in \{K_1, K_2\}$, $u(\cdot) \in \mathbb{R}$, et $y(\cdot) \in \mathbb{R}$. Les variables $u(\cdot)$ et $y(\cdot)$ représentent respectivement l'entrée et la sortie du système statique. Le gain K_i du système prend ses valeurs dans un ensemble *a priori* connu et composé des gains K_1 et K_2. Le système à commutation présente donc deux modes de fonctionnement.

La reconnaissance du mode actif à chaque instant ne nécessite plus ici de générer des résidus sur une fenêtre d'observation, les résidus pouvant être directement formés à partir d'un seul point de mesure, et donc le problème de la reconnaissance du mode actif peut être résolu sans avoir à transiter par la reconnaissance du chemin actif sur un horizon d'observation. Une autre façon de formuler le problème serait de dire

qu'on ne considère dans cette situation que les chemins de la forme $(i), i \in \{1, 2\}$. Pour cela, on forme, à partir de (3.18), deux résidus $r_1(\cdot)$ et $r_2(\cdot)$ définis par :

$$\begin{cases} r_1(k) = y(k) - K_1 u(k) \\ r_2(k) = y(k) - K_2 u(k) \end{cases} \tag{3.19}$$

Le résidu $r_i(k)$ prend la valeur zéro lorsque le mode i est actif à l'instant k. L'étude de la discernabilité des modes consiste à trouver les conditions sous lesquelles les résidus $r_1(\cdot)$ et $r_2(\cdot)$ prennent simultanément la valeur zéro, rendant impossible la reconnaissance du mode actif.
A partir de (3.19), les résidus $r_1(\cdot)$ et $r_2(\cdot)$ sont simultanément nuls si :

$$\begin{cases} y(k) = K_1 u(k) \\ y(k) = K_2 u(k) \end{cases} \tag{3.20}$$

De (3.20), on obtient :

$$\begin{cases} K_1 = K_2 \\ u(k) = 0 \end{cases} \tag{3.21}$$

En résumé, les deux modes du système sont discernables à tout instant si $K_1 \neq K_2$ et $u(k) \neq 0, \forall k \in \mathbb{N}$. On a ainsi obtenu une condition relative à la structure du système et à l'excitation fournie à ce système.

Considérons maintenant la présence d'un bruit borné sur la sortie du système modélisé par (3.18) :

$$\begin{cases} y(k) = K_i u(k) \\ y_m(k) = y(k) + n(k) \end{cases} \tag{3.22}$$

où $n(\cdot) \in \mathbb{R}$ est un bruit borné tel que $|n(k)| < \delta, \forall k \in \mathbb{N}$, $\delta > 0$. La variable $y_m(\cdot)$ est la sortie mesurée du système qui est égale à la somme de la sortie réelle $y(\cdot)$ et le bruit de mesure $n(\cdot)$. Le symbole $|\cdot|$ représente la valeur absolue.
A partir de l'entrée $u(\cdot)$ connue, on peut générer à chaque instant les sorties relatives aux deux modes du système (3.22) :

$$\begin{cases} y_1(k) = K_1 u(k) \\ y_2(k) = K_2 u(k) \end{cases} \tag{3.23}$$

où $y_i, i \in \{1,2\}$ est la sortie du système lorsque le mode i est actif.
Les sorties $y_1(\cdot)$ et $y_2(\cdot)$ définissent deux domaines auxquels doivent appartenir les mesures $y_m(k)$. Compte tenu de (3.22), ces domaines sont définis par :

$$D_1 = \{y \in \mathbb{R} / y = K_1 u(k) + n(k)\} \tag{3.24}$$

$$D_2 = \{y \in \mathbb{R} / y = K_2 u(k) + n(k)\} \tag{3.25}$$

Comme le bruit de mesure $n(\cdot)$ est borné, les domaines D_i sont des bandes du plan $\{u, y\}$ limitées par les droites parallèles d'équation :

$$D_i^+ = \{y \in \mathbb{R} / y = K_i u(k) + \delta\} \tag{3.26}$$

$$D_i^- = \{y \in \mathbb{R} / y = K_i u(k) - \delta\} \tag{3.27}$$

A un instant k_0 quelconque, on peut retrouver, à partir d'un couple de mesure $(u(k_0); y_m(k_0))$, l'ensemble des entrées pouvant générer la sortie $y_m(k_0)$, ceci pour chaque mode du système. Il suffit pour cela de projeter l'intersection de la droite d'équation $y_m = y_m(k_0)$ avec les domaines D_i sur l'axe des entrées du système. On obtient ainsi deux intervalles I_1 et I_2 définis respectivement pour les modes 1 et 2 :

$$\begin{cases} I_1 = \left[\dfrac{y_m(k_0) - \delta}{K_1}; \dfrac{y_m(k_0) + \delta}{K_1}\right] \\ I_2 = \left[\dfrac{y_m(k_0) - \delta}{K_2}; \dfrac{y_m(k_0) + \delta}{K_2}\right] \end{cases} \tag{3.28}$$

Les modes 1 et 2 sont discernables à l'instant k_0 si, pour la sortie mesurée $y_m(k_0)$, les intervalles I_1 et I_2 trouvés sont disjoints. Ceci s'explique par le fait que l'intersection des intervalles I_1 et I_2 représente l'ensemble des entrées qui appliquées au système donnent une sortie $y_m(k_0)$ appartenant à la fois à la bande D_1 et à la bande D_2. De ce fait, en cas d'existence d'une intersection entre les intervalles I_1 et I_2, il est nécessaire de vérifier en plus si $u(k_0) \in (I_1 \cap I_2)$ avant de conclure à la non discernabilité des modes 1 et 2. Dans le cas où $u(k_0) \notin (I_1 \cap I_2)$, la sortie $y_m(k_0)$ correspondante appartient à une et une seule des bandes D_i. En résumé, la

discernabilité des modes du système à l'instant k_0 n'est donc pas assurée si :

$$\begin{cases} (I_1 \cap I_2) \neq \emptyset \\ \text{et} \\ u\,(k_0) \in (I_1 \cap I_2) \end{cases} \tag{3.29}$$

Supposons $K_1 > K_2$. Dans ce cas, si $I_1 \cap I_2 \neq \emptyset$, on a :

$$(I_1 \cap I_2) = \left[\frac{y_m\,(k_0) + \delta}{K_1}; \frac{y_m\,(k_0) - \delta}{K_2} \right] \tag{3.30}$$

On a $u(k_0) \in (I_1 \cap I_2)$ si et seulement si :

$$\begin{cases} \dfrac{y_m\,(k_0) + \delta}{K_1} - u\,(k_0) < 0 \\ \dfrac{y_m\,(k_0) - \delta}{K_2} - u\,(k_0) > 0 \end{cases} \tag{3.31}$$

Soit :

$$\begin{cases} \frac{y_m(k_0)}{K_1} - u\,(k_0) + \frac{\delta}{K_1} < 0 \\ \dfrac{y_m\,(k_0)}{K_2} - u\,(k_0) - \dfrac{\delta}{K_2} > 0 \end{cases} \tag{3.32}$$

En reprenant le raisonnement précédent avec $K_1 < K_2$, on obtient deux autres inégalités :

$$\begin{cases} \dfrac{y_m\,(k_0)}{K_1} - u\,(k_0) - \dfrac{\delta}{K_1} > 0 \\ \dfrac{y_m\,(k_0)}{K_2} - u\,(k_0) + \dfrac{\delta}{K_2} < 0 \end{cases} \tag{3.33}$$

On déduit de (3.32) et (3.32) que le domaine dans lequel les modes 1 et 2 ne sont pas discernables est caractérisé dans le plan $\{u, y\}$ par l'ensemble d'inégalités linéaires

(3.34) :

$$\begin{cases} \dfrac{y_m(k_0)}{K_1} - u(k_0) + \dfrac{\delta}{K_1} < 0 \\ \dfrac{y_m(k_0)}{K_1} - u(k_0) - \dfrac{\delta}{K_1} > 0 \\ \dfrac{y_m(k_0)}{K_2} - u(k_0) - \dfrac{\delta}{K_2} > 0 \\ \dfrac{y_m(k_0)}{K_2} - u(k_0) + \dfrac{\delta}{K_2} < 0 \end{cases} \tag{3.34}$$

Exemple 3.3 (Discernabilité des modes dans le cas statique en présence d'un bruit de mesure borné) *Considérons un système modélisé par (3.22) avec $K_1 = 6.21$, $K_2 = 4.13$ et $\delta = 1$. La figure 3.2 illustre la discernabilité des modes*

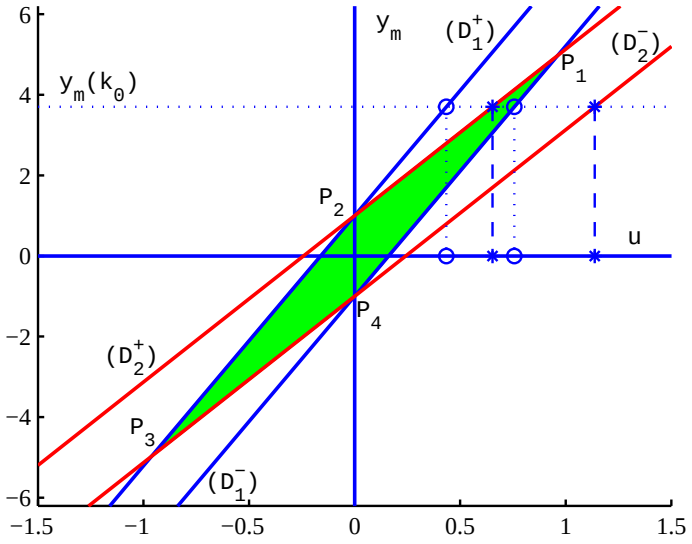


Fig. 3.2: Discernabilité dans le plan $\{u, y\}$

dans le plan $\{u, y\}$. La zone dans laquelle les deux modes 1 et 2 ne sont pas discernables est déterminée à partir de l'ensemble d'inégalités (3.34). Elle correspond au parallélogramme $(P_1P_2P_3P_4)$.

Le concept de discernabilité va être étendu dans les sections suivantes aux systèmes dynamiques présentant plusieurs modes de fonctionnement. L'objectif reste de mettre en place des tests permettant de se prononcer quant à la discernabilité des chemins sur un horizon d'observation.

3.3.2 Cas d'un système dynamique à commutation : absence de bruit de mesure

En préliminaire à tout développement, on donne la définition de la discernabilité des chemins d'un système à commutation.

Définition 3.1 (Chemins discernables : cas déterministe)
En l'absence de bruit de mesure, deux chemins μ^1 et μ^2 sont discernables sur un horizon d'observation $[k-h,\ k]$ si les résidus $r_{\mu^1,h}(k)$ et $r_{\mu^2,h}(k)$ correspondant aux deux chemins ne sont pas simultanément nuls lorsqu'un des deux chemins est actif sur l'horizon d'observation considéré.

Afin d'établir les conditions de discernabilité de deux chemins quelconques, on considère, à titre d'exemple, deux chemins μ^1 et μ^2 de longueur h sur l'horizon $[k-h,\ k]$. On note $Y^{\mu^1}_{k-h,k}$ (respectivement $Y^{\mu^2}_{k-h,k}$) le vecteur d'observation relatif au chemin μ^1 (respectivement μ^2). On suppose par exemple que pour ce système, à un instant k, le chemin actif sur la fenêtre d'observation est le chemin μ^1. Cette information n'étant pas connue, il convient d'analyser les possibilités que le chemin μ^1 ou que le chemin μ^2 décrivent le fonctionnement du système. Pour cela, à partir de (3.8), les expressions des résidus $r_{\mu^1,h}(\cdot)$ et $r_{\mu^2,h}(\cdot)$ sont données par les équations :

$$\begin{cases} r_{\mu^1,h}(k) = W_{\mu^1,h}\left(Y_{k-h,k} - T_{\mu^1,h}U_{k-h,k}\right) \\ r_{\mu^2,h}(k) = W_{\mu^2,h}\left(Y_{k-h,k} - T_{\mu^2,h}U_{k-h,k}\right) \end{cases} \tag{3.35}$$

Comme μ^1 est le chemin actif, l'équation (3.35) peut être alors mise sous la forme :

$$\begin{cases} r_{\mu^1,h}(k) = W_{\mu^1,h}\left(Y^{\mu^1}_{k-h,k} - T_{\mu^1,h}U_{k-h,k}\right) \\ r_{\mu^2,h}(k) = W_{\mu^2,h}\left(Y^{\mu^1}_{k-h,k} - T_{\mu^2,h}U_{k-h,k}\right) \end{cases} \tag{3.36}$$

et l'on a également $W_{\mu^1,h}\left(Y^{\mu^1}_{k-h,k} - T_{\mu^1,h}U_{k-h,k}\right) = 0$. D'où :

$$\begin{cases} r_{\mu^1,h}(k)(k) = 0 \\ r_{\mu^2,h}(k) = W_{\mu^2,h}\left(Y^{\mu^1}_{k-h,k} - T_{\mu^2,h}U_{k-h,k}\right) \end{cases} \tag{3.37}$$

En ajoutant et en retranchant $Y^{\mu^2}_{k-h,k}$ dans l'expression de $r_{\mu^2,h}(\cdot)$ on obtient :

$$\begin{cases} r_{\mu^1,h}(k)(k) = 0 \\ r_{\mu^2,h}(k) = W_{\mu^2,h}\left(Y^{\mu^1}_{k-h,k} - Y^{\mu^2}_{k-h,k} + Y^{\mu^2}_{k-h,k} - T_{\mu^2,h}U_{k-h,k}\right) \end{cases} \tag{3.38}$$

Par définition $W_{\mu^2,h}\left(Y^{\mu^2}_{k-h,k} - T_{\mu^2,h}U_{k-h,k}\right) = 0$, on a donc :

$$\begin{cases} r_{\mu^1,h}(k)(k) = 0 \\ r_{\mu^2,h}(k) = W_{\mu^2,h}\left(Y^{\mu^1}_{k-h,k} - Y^{\mu^2}_{k-h,k}\right) \end{cases} \tag{3.39}$$

Il apparaît clairement dans l'équation (3.39) que le résidu calculé pour le chemin μ^2 (chemin non actif sur l'horizon d'observation) est directement lié à la différence entre les sorties obtenues pour le système lorsque ses modes évoluent selon les deux chemins μ^1 et μ^2, le système étant soumis aux mêmes entrées dans les deux cas. L'écart entre $r_{\mu^1,h}(k)$ et $r_{\mu^2,h}(k)$ provient de $W_{\mu^2,h}\left(Y^{\mu^1}_{k-h,k} - Y^{\mu^2}_{k-h,k}\right)$. Pour que les deux résidus soient distincts, il faut satisfaire aux deux conditions suivantes :

- $$Y^{\mu^1}_{k-h,k} - Y^{\mu^2}_{k-h,k} \notin \mathcal{N}_r(W_{\mu^2,h}) \tag{3.40}$$
- $$Y^{\mu^1}_{k-h,k} - Y^{\mu^2}_{k-h,k} \neq 0 \tag{3.41}$$

où $\mathcal{N}_r$ désigne l'opérateur « espace nul à droite ».
Nous allons donc analyser séparément chacune des deux conditions (3.40) et (3.41). Commençons par examiner la première (3.40). On a, d'après (3.6) :

$$Y^{\mu^1}_{k-h,k} - Y^{\mu^2}_{k-h,k} = (\mathcal{O}_{\mu^1,h} - \mathcal{O}_{\mu^2,h})\, x(k-h) + (T_{\mu^1,h} - T_{\mu^2,h})\, U_{k-h,k} \tag{3.42}$$

où $x(k-h)$ est la valeur de l'état du système à l'instant initial de la fenêtre d'observation. On déduit de (3.42) après multiplication à gauche par $W_{\mu^2,h}$:

$$W_{\mu^2,h}(Y^{\mu^1}_{k-h,k} - Y^{\mu^2}_{k-h,k}) = W_{\mu^2,h}\mathcal{O}_{\mu^1,h}x(k-h) + W_{\mu^2,h}\,(T_{\mu^1,h} - T_{\mu^2,h})\,U_{k-h,k} \tag{3.43}$$

Si $Y^{\mu^1}_{k-h,k} - Y^{\mu^2}_{k-h,k}$ appartient à l'espace nul à droite de la matrice $W_{\mu^2,h}$, on obtient :

$$W_{\mu^2,h}\mathcal{O}_{\mu^1,h}x(k-h) + W_{\mu^2,h}\left(T_{\mu^1,h} - T_{\mu^2,h}\right)U_{k-h,k} = 0 \tag{3.44}$$

La matrice $W_{\mu^2,h}$ étant de plein rang colonne, deux situations peuvent se présenter selon la valeur de l'état initial $x(k-h)$. Dans la première situation, on a $x(k-h) = 0$. Ceci implique que :

$$W_{\mu^2,h}\left(T_{\mu^1,h} - T_{\mu^2,h}\right)U_{k-h,k} = 0, \tag{3.45}$$

et donc que le vecteur des entrées $U_{k-h,k}$ appartient à l'espace nul à droite de la matrice $W_{\mu^2,h}\left(T_{\mu^1,h} - T_{\mu^2,h}\right)$.
La deuxième situation est celle où $x(k-h) \neq 0$. Selon (3.44) et comme $W_{\mu^2,h} \neq 0$, pour que, quelque soit $x(k-h)$, on ait $Y^{\mu^1}_{k-h,k} - Y^{\mu^2}_{k-h,k}$ élément de l'espace nul à droite de $W_{\mu^2,h}$, il faut que :

$$\mathcal{O}_{\mu^1,h}x(k-h) + \left(T_{\mu^1,h} - T_{\mu^2,h}\right)U_{k-h,k} = 0 \tag{3.46}$$

$$\mathcal{O}_{\mu^1,h}x(k-h) = \left(T_{\mu^2,h} - T_{\mu^1,h}\right)U_{k-h,k} \tag{3.47}$$

Pour s'affranchir de l'état initial, multiplions les deux termes de (3.47) par $W_{\mu^1,h}$. On a :

$$W_{\mu^1,h}\left(T_{\mu^2,h} - T_{\mu^1,h}\right)U_{k-h,k} = 0 \tag{3.48}$$

Pour satisfaire cette équation, le vecteur des entrées $U_{k-h,k}$ doit donc appartenir à l'espace nul à droite de la matrice $W_{\mu^1,h}\left(T_{\mu^2,h} - T_{\mu^1,h}\right)$.
Finalement, compte tenu de (3.45) et (3.48), la condition (3.40) est satisfaite si l'entrée $U_{k-h,k}$ n'appartient ni à l'espace nul à droite de $W_{\mu^2,h}\left(T_{\mu^1,h} - T_{\mu^2,h}\right)$, ni à l'espace nul à droite de $W_{\mu^1,h}\left(T_{\mu^2,h} - T_{\mu^1,h}\right)$.

La seconde condition (3.41) assurant la distinction des résidus $r_{\mu^1,h}(\cdot)$ et $r_{\mu^2,h}(\cdot)$, à savoir la condition $Y^{\mu^1}_{k-h,k} - Y^{\mu^2}_{k-h,k} \neq 0$, est vérifiée dès qu'au moins un élément de $Y^{\mu^1}_{k-h,k} - Y^{\mu^2}_{k-h,k}$ est différent de zéro. En exprimant, à partir de l'état à l'instant k, la sortie du système pour les deux chemins μ^1 et μ^2 à un instant quelconque $k+1$

sur l'horizon d'observation, on a :

$$\begin{cases} y_{\mu_k^1}(k+1) = CA_{\mu_k^1}x(k) + CBu(k) \\ y_{\mu_k^2}(k+1) = CA_{\mu_k^2}x(k) + CBu(k) \end{cases} \tag{3.49}$$

D'où :

$$y_{\mu_k^1}(k+1) - y_{\mu_k^2}(k+1) = C\left(A_{\mu_k^1} - A_{\mu_k^2}\right)x(k) \tag{3.50}$$

Comme $\mu^1 \neq \mu^2$, il existe toujours sur l'horizon d'observation $[k-h,\ k]$, un instant k_0 tel que $A_{\mu_{k_0}^1} \neq A_{\mu_{k_0}^2}$. De ce fait, pour que, sur une fenêtre d'observation, on ait au moins un élément de $Y_{k-h,k}^{\mu^1} - Y_{k-h,k}^{\mu^2}$ qui soit différent de zéro indépendamment de l'état $x(k)$, il suffit que les lignes de la matrice C ne soient pas orthogonales aux colonnes de la matrice $A_{\mu_{k_0}^1} - A_{\mu_{k_0}^2}$ et qu'au moins une valeur de l'état du système sur l'horizon d'observation soit différent de zéro (condition qui est toujours vérifiée). Donc, pour qu'on ait $Y_{k-h,k}^{\mu^1} - Y_{k-h,k}^{\mu^2} \neq 0$, il suffit que la matrice C n'appartienne pas à l'espace nul à gauche de la matrice $A_i - A_j$, $i \neq j$ et $i, j \in S_{1,2} = \{1, \ldots, s_{\mu^{1,2}}\}$ où l'ensemble $S_{1,2}$ recense tous les modes intervenant dans la construction des chemins μ^1 et μ^2.

Théorème 3.3 (Condition de discernabilité des chemins)
Deux chemins μ^1 et μ^2 d'un système à commutation sont discernables sur un horizon d'observation $[k-h, k]$, si :

$$U_{k-h,k} \notin \mathcal{N}_r\left(W_{\mu^1,h}\left(T_{\mu^2,h} - T_{\mu^1,h}\right)\right) \tag{3.51}$$

$$U_{k-h,k} \notin \mathcal{N}_r\left(W_{\mu^2,h}\left(T_{\mu^1,h} - T_{\mu^2,h}\right)\right) \tag{3.52}$$

$$C \notin \mathcal{N}_l\left(A_i - A_j\right), i \neq j,\ i, j \in S_{1,2} = \{1, \ldots, s_{\mu^{1,2}}\} \tag{3.53}$$

$\mathcal{N}_r$ (respectivement $\mathcal{N}_l$) désigne l'opérateur « espace nul à droite » (respectivement « espace nul à gauche ») et $U_{k-h,k}$ est le vecteur des entrées empilées sur l'horizon d'observation. L'ensemble $S_{1,2}$ recense tous les modes intervenant dans la construction des chemins μ^1 et μ^2.

La démonstration de ce théorème découle directement des remarques précédentes. La relation (3.53) est identique à la condition de discernabilité des modes mise en

exergue dans Cocquempot *et al.* [2003] ; il s'agit d'une condition de discernabilité structurelle qui ne dépend que des différentes matrices décrivant le système. Les relations (3.51) et (3.52) viennent compléter cette condition de discernabilité des modes en prenant en compte la nature de l'excitation fournie au système.

3.3.3 Cas d'un système dynamique à commutation : présence d'un bruit de mesure borné

Il est tentant de trouver, en présence d'un bruit de mesure, des conditions de discernabilité des chemins semblables à celles énoncées dans le théorème 3.3. On considère ici le cas d'un bruit de mesure borné.
En présence de bruit de mesure, la reconnaissance du chemin actif étant basée sur l'examen de résidus intervalle, il convient de reformuler la définition 3.1 afin de l'adapter à ce cas précis.

Définition 3.2 (Chemins discernables : présence d'un bruit de mesure) *En présence d'un bruit de mesure borné, deux chemins μ^1 et μ^2 sont discernables sur un horizon d'observation $[k-h,\ k]$ si les résidus intervalle $[r_{\mu^1,h}(k)]$ et $[r_{\mu^2,h}(k)]$ correspondant aux deux chemins ne contiennent pas simultanément la valeur zéro lorsqu'un des deux chemins est actif sur l'horizon d'observation considéré.*

La présence d'un bruit de mesure borné sur la sortie du système conduit au calcul de résidus intervalle sous la forme (3.15). L'étude de la discernabilité des chemins consiste ici à trouver les conditions sous lesquelles deux résidus intervalle, calculés pour des chemins différents, peuvent simultanément contenir tous les deux la valeur zéro. Pour cela, on considère deux chemins μ^1 et μ^2 de longueur h sur une fenêtre temporelle $[k-h,\ k]$ et on suppose que le chemin μ^1 est le chemin actif sur cet horizon. Les expressions des résidus intervalle $[r_{\mu^1,h}(\cdot)]$ et $[r_{\mu^2,h}(\cdot)]$ associés respectivement aux chemins μ^1 et μ^2 sont :

$$\begin{aligned}[r_{\mu^1,h}(k)] &= [\underline{r}_{\mu^1,h}\ ,\ \bar{r}_{\mu^1,h}] \\ [r_{\mu^2,h}(k)] &= [\underline{r}_{\mu^2,h}\ ,\ \bar{r}_{\mu^2,h}] + W_{\mu^2,h}\left(Y^{\mu^1}_{k-h,k} - Y^{\mu^2}_{k-h,k}\right)\end{aligned} \tag{3.54}$$

où les bornes $\underline{r}_{\mu^i,h}$ et $\bar{r}_{\mu^i,h}$, $i \in \{1,2\}$ sont données par : $\underline{r}_{\mu^i,h} = -\left|W_{\mu^i,h}\right| \mathbb{U}\delta$ et $\bar{r}_{\mu^i,h} = \left|W_{\mu^i,h}\right| \mathbb{U}\delta$ avec δ la borne maximale de l'amplitude du bruit de mesure et $\mathbb{U}$ un vecteur colonne de dimension égale au nombre de colonnes de $W_{\mu^i,h}$ et dont tous les éléments sont égaux à 1.
Le résidu intervalle $[r_{\mu^1,h}(\cdot)]$ contient par définition la valeur zéro. La valeur zéro n'est pas incluse dans le résidu intervalle $[r_{\mu^2,h}(\cdot)]$ qu'à condition que :

$$\left|W_{\mu^2,h}\left(Y^{\mu^1}_{k-h,k} - Y^{\mu^2}_{k-h,k}\right)\right| > \left|W_{\mu^2,h}\right| \mathbb{U}\delta \tag{3.55}$$

Comme $W_{\mu_2,h}Y^{\mu_2}_{k-h,k} = -W_{\mu_2,h}T_{\mu_2,h}U_{k-h,k}$ et $Y^{\mu_1}_{k-h,k} = \mathcal{O}_{\mu_1,h}x\,(k-h) + T_{\mu_1,h}U_{k-h,k}$, l'inégalité (3.55) conduit à :

$$\left|W_{\mu_2,h}\left(\mathcal{O}_{\mu_1,h}x\,(k-h) + \left(T_{\mu_1,h} + T_{\mu_2,h}\right)U_{k-h,k}\right)\right| > \left|W_{\mu_2,h}\right| \mathbb{U}\delta \tag{3.56}$$

L'inégalité (3.56) contient une inconnue à savoir $x(k-h)$. Cette condition ne peut alors être testée. Il faut donc se contenter de l'analyse de l'appartenance de la valeur zéro aux résidus intervalle générés pour décider de la discernabilité des chemins en présence d'un bruit de mesure borné.

Conclusion et discussion

Il a été abordé ici le problème de la détermination du mode actif à chaque instant d'un système à commutation ainsi que des instants de commutation. La méthode repose sur l'analyse de résidus générés à partir des différents chemins générés sur une fenêtre d'observation $[k-h\ k]$ du système. Des conditions garantissant la discernabilité des différents chemins générés sur la fenêtre d'observation ont été établies mises en place aussi bien dans un cadre déterministe qu'en présence d'un bruit de mesure borné. La démarche adoptée en présence d'un bruit de mesure borné peut être élargie à un bruit modélisé par une distribution gaussienne. Il suffit dans ce cas de définir des intervalles de confiance pour les résidus et de tester ensuite l'appartenance des résidus calculés à ces intervalles de confiance.

Un point à développer est la taille de l'horizon d'observation assurant une meilleure discernabilité des chemins. Des simulations numériques, qui seront ex-

posées plus tard, ont montré l'influence de la taille de cet horizon sur la qualité de la reconnaissance des différents chemins. Toutefois, une expression analytique de cette taille n'a pu être exhibée. Il est clair que cette taille dépend de la condition d'observabilité de la proposition 3.1 et également des conditions de discernabilité du théorème 3.3. Il pourrait éventuellement exister également un lien avec la condition de discernabilité-avant de Babaali et Egerstedt [2004].

Un autre point à approfondir est le fonctionnement du système en mode non supervisé. On suppose, dans ce cas, que l'on ne dispose pas d'une connaissance complète de tous les régimes de fonctionnement du système. Il s'agit alors de procéder, au cours du fonctionnement du système, à l'identification simultanée des modes de fonctionnement non répertoriés, c'est-à-dire des matrices d'état associées à ces modes. La faisabilité de cette démarche sera illustrée plus tard par un exemple de simulation.

Notons enfin qu'il est tout à fait envisageable de procéder à l'estimation de l'état $x(\cdot)$ du système à commutation une fois la reconnaissance du mode actif effectuée. Pour cela, et afin de conserver une cohérence dans la démarche proposée, on peut mettre en œuvre un observateur à mémoire finie.

4

Identification du mécanisme de commutation

Sommaire

4.1 Position du problème 129

4.2 Recherche d'un hyperplan séparateur linéaire 130

4.2.1 Recherche d'un domaine de solutions sous forme simple 133

4.2.2 Recherche d'une valeur unique 135

4.2.3 Classes non linéairement séparables 139

4.3 Recherche d'un hyperplan séparateur linéaire par morceaux 143

4.3.1 Approche *one-against-all* 144

4.3.2 Approche *all-together* 145

Introduction

Dans le chapitre précédent, il n'a été fait mention d'aucune hypothèse quant à la structure du mécanisme régissant les changements de mode du système à commutation. La loi de commutation a été donc considérée comme un processus dont la maîtrise fait défaut. Toutefois, la logique voudrait, pour aller dans la continuité de ce qui a été fait jusqu'à présent, qu'une fois le mode actif à chaque instant détecté, l'on tente de décrire le mécanisme permettant le choix d'un mode particulier du système à tout instant. Pour cela, on s'intéresse, dans ce chapitre, à une classe particulière de structure pour la loi de commutation. Il est supposé que la loi de commutation du système est obtenue en partitionnant un espace associé à des grandeurs connues du système. L'ensemble des grandeurs connues est restreint ici à l'entrée et à la sortie du système comme dans le cas des modèles PWARX. Il faut noter que, dans la littérature, les grandeurs dont dépend la loi de commutation incluent le plus souvent l'état du système ou des combinaisons linéaires de l'état. Par rapport à cette remarque, on peut recadrer le problème posé dans ce chapitre sous trois angles :

- le mécanisme de commutation dépend de grandeurs mesurables. Les méthodes proposées dans ce chapitre s'inscrivent dans ce contexte.
- le mécanisme de commutation dépend de grandeurs non mesurables mais pouvant être estimées. Dans cette situation, il est nécessaire de procéder au préalable à l'estimation de ces grandeurs. Une fois l'estimation des grandeurs inconnues effectuées, les méthodes mises en œuvre dans ce chapitre peuvent alors être utilisées.
- le mécanisme de commutation dépend de grandeurs non mesurables et ne pouvant être estimées. Cette dernière situation reste un problème ouvert.

La structure de la loi de commutation étant fixée, il faut procéder à son identification, c'est-à-dire à l'estimation de ses paramètres. Ce problème est abordé ici avec deux approches différentes. La première approche met en œuvre une vision ensembliste du problème d'estimation paramétrique en recherchant une expression intervalle des paramètres de la loi de commutation. La seconde approche recherche les valeurs optimales des paramètres de la loi de commutation par rapport à des cri-

tères de distance en utilisant des techniques issues du domaine de la reconnaissance de forme.

4.1 Position du problème

On considère le modèle (3.1) du chapitre 3 en lui adjoignant une information supplémentaire sur la loi de commutation :

$$\begin{cases} x(k+1) = A_{\mu_k} x(k) + Bu(k) \\ y(k) = Cx(k) \end{cases} \qquad (4.1)$$
$$\mu_k = i \text{ si } X_k \in \Omega_i, \quad i = 1, \ldots, s$$

où $X_k = [Y_{k-1,k-n_a} \;\; U_{k-1,k-n_b}]^T$, $\bigcup_{i=1}^{s} \Omega_i = \Omega$ est une partition complète de l'ensemble des régresseurs $\Omega \subseteq \mathbb{R}^{n_a+n_b}$. Chaque région Ω_i est un polyèdre convexe :

$$\Omega_i = \left\{ X \in \mathbb{R}^{n_a+n_b} \,\middle|\, \bar{H}_i X + g_i \leq 0 \right\} \qquad (4.2)$$

avec $\bar{H}_i \in \mathbb{R}^{q_i \times (n_a+n_b)}$, $g_i \in \mathbb{R}^{q_i}$, $i = 1, \ldots, s$.
On suppose qu'à partir du modèle (4.1), on a obtenu un jeu de données $\mathcal{D} = \{X_k, \; k = 1, \ldots, N\}$ sur lequel on a procédé au préalable à la reconnaissance du mode actif à chaque instant. Connaissant les domaines temporels où chaque mode est actif ainsi que les instants de commutation d'un mode à un autre, le problème posé est d'estimer les paramètres $\bar{H}_i$ et g_i, $i = 1 \ldots, s$ des polyèdres convexes définissant les régions d'activation de chaque mode du système. Les polyèdres Ω_i, $i = 1, \ldots, s$ définissent la loi de commutation du système.
Le mode actif à chaque instant étant connu, on peut scinder le jeu de données $\mathcal{D}$ en s classes $\mathcal{C}_i$, $i = 1, \ldots, s$ en effectuant l'affectation des données X_k aux classes $\mathcal{C}_i$ de la façon suivante :

$$X_k \in \mathcal{C}_i \quad \text{si } \mu_k = i \qquad (4.3)$$

A partir des classes $\mathcal{C}_i$, $i = 1 \ldots, s$, le problème de l'estimation des paramètres $\bar{H}_i$ et g_i, $i = 1 \ldots, s$ revient à séparer les classes $\mathcal{C}_i$, $i = 1 \ldots, s$ au moyen de

séparateurs linéaires, notamment des hyperplans. Ce problème a été étudié dans plusieurs domaines et deux approches sont retenues en général pour le résoudre :

- dans la première approche, on construit un hyperplan séparateur linéaire en considérant deux à deux les classes $\mathcal{C}_i,\ i = 1\ldots, s$. Il s'agit dans ce cas de résoudre $s(s-1)/2$ problèmes de séparation linéaire de deux classes.
- la seconde approche recherche un hyperplan séparateur linéaire par morceaux et considère donc simultanément l'ensemble des points des classes $\mathcal{C}_i,\ i = 1\ldots, s$.

La première approche, connue dans la littérature anglo-saxonne sous le nom d'approche *one-against-one*, est intéressante car peu gourmande en charge calculatoire. Toutefois, elle présente l'inconvénient de ne pas garantir que les régions estimées forment une partition complète de l'ensemble des régresseurs Ω. La seconde approche permet de s'affranchir de cet inconvénient au prix d'une charge de calcul élevée.

Les méthodes de séparation linéaire seront par la suite utilisées afin de trouver les paramètres $\bar{H}_i$ et $g_i,\ i = 1\ldots, s$.

4.2 Estimation du partitionnement de l'espace des régresseurs en considérant les classes deux à deux

Le jeu de données $\mathcal{D}$ ayant été scindé en s classes $\mathcal{C}_i,\ i = 1\ldots, s$, on recherche ici les paramètres $\bar{H}_i$ et $g_i,\ i = 1\ldots, s$ en prenant en compte deux à deux les classes $\mathcal{C}_i,\ i = 1\ldots, s$. Pour cela, considérons deux classes $\mathcal{C}_i$ et $\mathcal{C}_j$, $i \neq j$. Séparer linéairement ces deux classes revient à déterminer un hyperplan $\mathcal{H}_{ij} = \{X_k \in \mathbb{R}^{n_a+n_b} \mid \bar{h}_{ij}X_k + \mathrm{g}_{ij} = 0\}$ tel que :

$$\begin{cases} \bar{h}_{ij}X_k + \mathrm{g}_{ij} > 0 \text{ si } X_k \in \mathcal{C}_i \\ \bar{h}_{ij}X_k + \mathrm{g}_{ij} < 0 \text{ si } X_k \in \mathcal{C}_j \end{cases} \tag{4.4}$$

où $\bar{h}_{ij} \in \mathbb{R}^{n_a+n_b}$ et $\mathrm{g}_{ij} \in \mathbb{R}$.

La détermination des hyperplans $\mathcal{H}_{ij},\ i = 1, \ldots, s-1,\ j = i+1, \ldots, s$ permet de retrouver les régions d'activation $\Omega_i,\ i = 1, \ldots, s$ des modes du système à commutation (4.1) qui sont ainsi définies par :

$$\begin{cases} h_{ji}X_k + \mathrm{g}_{ji} \leq 0 & j = 1, \ldots, i-1 \\ h_{ij}X_k + \mathrm{g}_{ij} \geq 0 & j = i+1, \ldots, s \end{cases} \tag{4.5}$$

Pour un mode $i^* \in \{1, \ldots, s\}$ du système, les paramètres $\bar{H}_i^*$ et g_i^* de la région d'activation de ce mode sont donc donnés par :

$$\begin{cases} \bar{H}_{i^*} = \begin{bmatrix} \bar{h}_{1i^*} & \ldots & \bar{h}_{(i^*-1)i^*} & -\bar{h}_{i^*(i^*+1)} & \ldots & -\bar{h}_{i^*s} \end{bmatrix}^T \\ g_{i^*} = \begin{bmatrix} \mathrm{g}_{1i^*} & \ldots & \mathrm{g}_{(i^*-1)i^*} & -\mathrm{g}_{i^*(i^*+1)} & \ldots & -\mathrm{g}_{i^*s} \end{bmatrix}^T \end{cases} \tag{4.6}$$

A partir de (4.4), on peut écrire l'inégalité (4.7) pour un élément quelconque X_k de la classe $\mathcal{C}_i$ ou $\mathcal{C}_j$:

$$\nu_k \left(\bar{h}_{ij}X_k + \mathrm{g}_{ij} \right) > 0 \tag{4.7}$$

où $\nu_k = \text{signe}\left(\bar{h}_{ij}X_k + \mathrm{g}_{ij} \right)$.

Pour l'ensemble des N_{ij} points X_k d'observation (N_{ij} est égal à la somme du cardinal de $\mathcal{C}_i$ et du cardinal de $\mathcal{C}_j$), on obtient un ensemble d'inégalités pouvant s'écrire sous la forme d'une inégalité matricielle linéaire en les paramètres $\bar{h}_{ij}$ et g_{ij} :

$$-\begin{bmatrix} \nu_1 X_1^T & \nu_1 \\ \vdots & \vdots \\ \nu_{N_{ij}} X_{N_{ij}}^T & \nu_{N_{ij}} \end{bmatrix} \begin{bmatrix} \bar{h}_{ij}^T \\ \mathrm{g}_{ij} \end{bmatrix} < \begin{bmatrix} 0 \\ \vdots \\ 0 \end{bmatrix} \tag{4.8}$$

La résolution de (4.8) donne un domaine auquel appartiennent les paramètres $\bar{h}_{ij}$ et g_{ij} et donc un ensemble de solutions admissibles. La précision de l'estimation de ce domaine dépend de la capacité de l'ensemble des données entrée/sortie à appréhender les différents modes du système. Dans le cas général, la résolution de (4.8) peut conduire à un domaine « complexe », c'est-à-dire décrit par un ensemble important de sommets. Il peut être souhaitable de réduire la complexité de ce domaine en déterminant un polytope de forme plus simple (parallélotope ou zonotope).

A titre d'exemple, le premier graphique de la figure 4.1 représente la projection dans $\mathbb{R}^2$ du domaine trouvé pour l'ensemble des paramètres admissibles pour le jeu de mesures de la table 4.1 avec $X_k = [y_{k-1}, u_{k-1}]^T$, $\bar{h}_{ij} = [h_{ij,1}\ h_{ij,2}]$. Ce domaine est tracé dans le plan $(h_{ij,2}, \mathrm{g}_{ij})$. Tout point de ce domaine représente une solution particulière. Le symbole « o » pointe une solution particulière. Le second graphique de la figure 4.1 présente une solution sous-optimale (rectangle gris) envisageable qui simplifie la description du domaine trouvé sous forme d'inégalités indépendantes.

u_{k-1}	0	0	-1	-2
y_{k-1}	-1	-2	0	0
ν_k	1	-1	1	-1

Tab. 4.1: Jeu de mesures

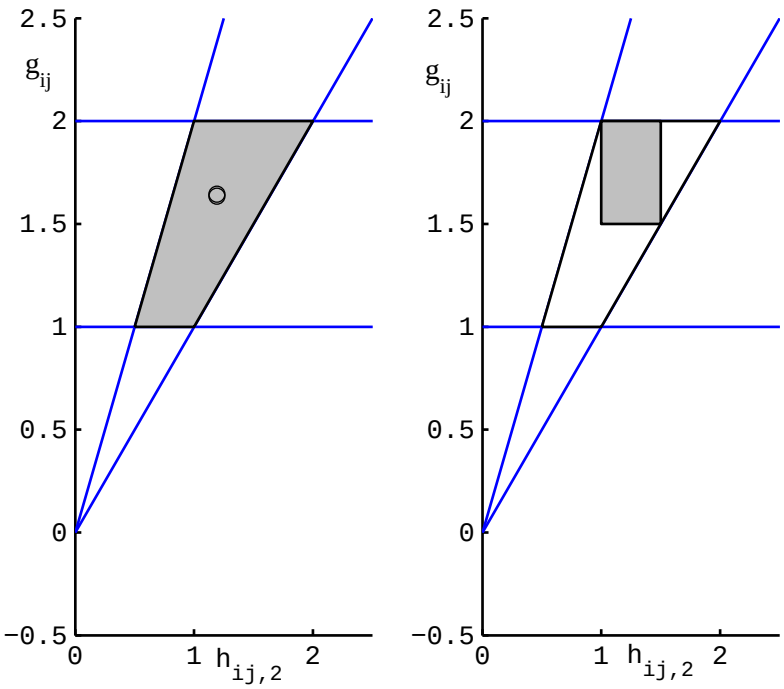


Fig. 4.1: Domaines admissibles pour $\bar{h}_{ij}$ et g_{ij}

On peut donc envisager la recherche des paramètres $\bar{h}_{ij}$ et g_{ij} de l'hyperplan séparateur $\mathcal{H}_{ij}$ des deux classes $\mathcal{C}_i$ et $\mathcal{C}_j$ sous trois points de vue différents :

- la recherche du domaine de l'ensemble des solutions admissibles pour $\bar{h}_{ij}$ et g_{ij}
- la recherche d'un domaine de solutions admissibles de forme simple pour $\bar{h}_{ij}$ et g_{ij}

– la recherche d'une valeur unique pour $\bar{h}_{ij}$ et g_{ij}

Le domaine correspondant à l'ensemble des solutions admissibles pour $\bar{h}_{ij}$ et g_{ij} est le domaine décrit par l'inégalité linéaire matricielle (4.8). La recherche d'un domaine de solutions admissibles de forme simple se fera au moyen d'une description intervalle des paramètres recherchés. Enfin, la recherche d'une valeur unique pour $\bar{h}_{ij}$ et g_{ij} consistera à maximiser une distance entre les classes $\mathcal{C}_i$ et $\mathcal{C}_j$, notamment la marge de séparation entre les deux classes.

4.2.1 Recherche d'un domaine de solutions sous forme simple

La non-exhaustivité de l'échantillon de mesures rend difficile la détermination d'une valeur unique des paramètres de la loi de commutation. C'est notamment le cas lorsqu'on dispose de peu de mesures au voisinage de l'hyperplan séparateur $\mathcal{H}_{ij}$ des classes $\mathcal{C}_i$ et $\mathcal{C}_j$. L'idée est alors de rechercher non pas une valeur unique pour les paramètres $\bar{h}_{ij}$ et g_{ij} de l'hyperplan $\mathcal{H}_{ij}$, mais plutôt un domaine auquel ces paramètres sont susceptibles d'appartenir. On ne parle alors plus de seuil de commutation, mais plutôt de zone de commutation.

A titre d'exemple, sur la figure 4.2 sont représentés des points $(y_{k-1}; \nu_k)$ lorsque le mécanisme de commutation assurant la séparation entre les classes $\mathcal{C}_i$ et $\mathcal{C}_j$ est définie par :

$$\begin{cases} y_k \in \mathcal{C}_i \text{ si } \nu_k = \text{signe}\,(y_{k-1} - 5) = 1 \\ y_k \in \mathcal{C}_j \text{ si } \nu_k = \text{signe}\,(y_{k-1} - 5) = -1 \end{cases} \tag{4.9}$$

La zone de commutation correspond à la partie grisée sur la figure 4.2 et est égale à l'intervalle de centre 5 et de demi-largeur 1. Il a été souligné précédemment que la résolution de (4.8) donne un domaine décrivant l'ensemble des valeurs admissibles pour les paramètres à identifier. Toutefois, les domaines obtenus, généralement des polytopes, ne sont pas aisés à manipuler à cause de leur forme géométrique. Pour contourner cette complexité, il est proposé de représenter de tels domaines par des orthotopes alignés aux axes. Cela revient à déterminer les paramètres $\bar{h}_{ij}$ et g_{ij} sous forme intervalle. Plusieurs critères d'optimisation peuvent alors être choisis. Par exemple, on peut imposer que les demi-largeurs des intervalles à déterminer soient maximales tout en respectant les contraintes du système. En posant $h_{ij} = [\bar{h}_{ij} \;\; g_{ij}]$,

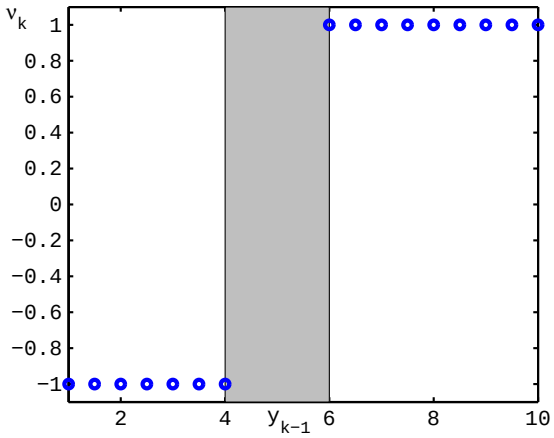


FIG. 4.2: Exemple de zone de commutation $\nu_k = \text{signe}\,(y_{k-1} - 5)$

on obtient :

$$\begin{cases} h_{ij} = h_{ij_0} + \lambda_{h_{ij}} \otimes r_{h_{ij}} \\ r_{h_{ij}} > 0 \\ \left\|\lambda_{h_{ij}}\right\|_\infty \leq 1 \end{cases} \tag{4.10}$$

où h_{ij_0} est un vecteur dont les composantes sont les centres des intervalles à déterminer, le vecteur $r_{h_{ij}}$ représente les demi-largeurs de ces intervalles et les variables $\lambda_{(.)}$ sont des variables bornées et normalisées permettant de prendre en compte l'ensemble des valeurs à l'intérieur des intervalles définis. L'opérateur $\otimes$ effectue la multiplication composante par composante de deux vecteurs. On rappelle que pour tout vecteur $e \in \mathbb{R}^n$, on a $\|e\|_\infty = \max\limits_{1 \leq i \leq n} |e_i|$, e_i désignant la $i^{\text{ème}}$ composante du vecteur e.

Compte tenu de l'expression intervalle de h_{ij} (4.10), l'équation (4.7) peut être récrite sous la forme :

$$\begin{cases} r_{h_{i,j}} > 0 \\ \left\|\lambda_{h_{ij}}\right\|_\infty \leq 1 \\ \nu_k \left(h_{ij_0} + \lambda_{h_{ij}} \otimes r_{h_{ij}}\right) \varphi_k \geq 0,\ k = 1, \ldots, N_{ij} \end{cases} \tag{4.11}$$

où $\varphi_k = [X_k \;\; 1]$.
On obtient ainsi :

$$\begin{cases} r_{h_{i,j}} > 0 \\ \nu_k \left(h_{ij_0} + r_{h_{ij}} \right) \varphi_k \geq 0, \; k = 1, \ldots, N_{ij} \\ \nu_k \left(h_{ij_0} - r_{h_{ij}} \right) \varphi_k \geq 0, \; k = 1, \ldots, N_{ij} \end{cases} \tag{4.12}$$

Finalement, pour trouver h_{ij_0} et $r_{h_{ij}}$, il faut rechercher les intervalles de demi-largeur maximale tout en respectant les contraintes (4.12). Un choix naturel peut consister à maximiser l'aire de l'orthotope aligné aux axes, inscrit dans le polytope des contraintes. On obtient dans ce cas le problème d'optimisation sous contrainte suivant :

$$\begin{cases} \max\limits_{h_{ij_0}, r_{h_{ij}}} \prod\limits_{m=1}^{n_a+n_b+1} r_{h_{ij},m} \\ \text{sous les contraintes (4.12)} \end{cases} \tag{4.13}$$

On peut alors faire appel à des algorithmes classiques d'optimisation sous contraintes Bonnans *et al.* [2002] pour la résolution de (4.13).

4.2.2 Recherche d'une valeur unique

Afin de trouver une valeur unique pour les paramètres $\bar{h}_{ij}$ et g_{ij}, il est nécessaire de fixer un critère permettant de définir l'optimalité des valeurs trouvées. Plusieurs choix s'avèrent possibles. Par exemple, on peut envisager de retenir simplement le centre des intervalles trouvés lors de la recherche d'un domaine de solutions sous forme simple, c'est-à-dire h_{ij_0}. En fait, les résultats obtenus en utilisant un critère ou un autre sont fonction de la forme du domaine correspondant à l'ensemble des solutions admissibles et des connaissances *a priori* portant sur les valeurs des paramètres à déterminer (plage de variation des paramètres,...).
Comme cela a été dit plus haut, la détermination d'une valeur unique pour les paramètres $\bar{h}_{ij}$ et g_{ij} est équivalente à la recherche d'un hyperplan séparateur linéaire séparant les classes $\mathcal{C}_i$ et $\mathcal{C}_j$ et se résume donc à un problème de séparation linéaire Bennett et Bredensteiner [1993].
La reconnaissance du mode actif à chaque instant ayant été effectuée, les valeurs

prises par $\mu_{(\cdot)}$ sont alors connues et les classes $\mathcal{C}_i$ et $\mathcal{C}_j$ sont linéairement séparables dans un contexte déterministe, exempt de tout bruit de mesure. La figure 4.3 montre, dans un espace à deux dimensions, deux classes $\mathcal{C}_1$ et $\mathcal{C}_2$ qui sont linéairement séparables. L'hyperplan séparateur est réduit, dans ce cas, à une droite, notamment la droite (D) sur la figure 4.3. Les classes $\mathcal{C}_i$ et $\mathcal{C}_j$ étant linéairement sé-

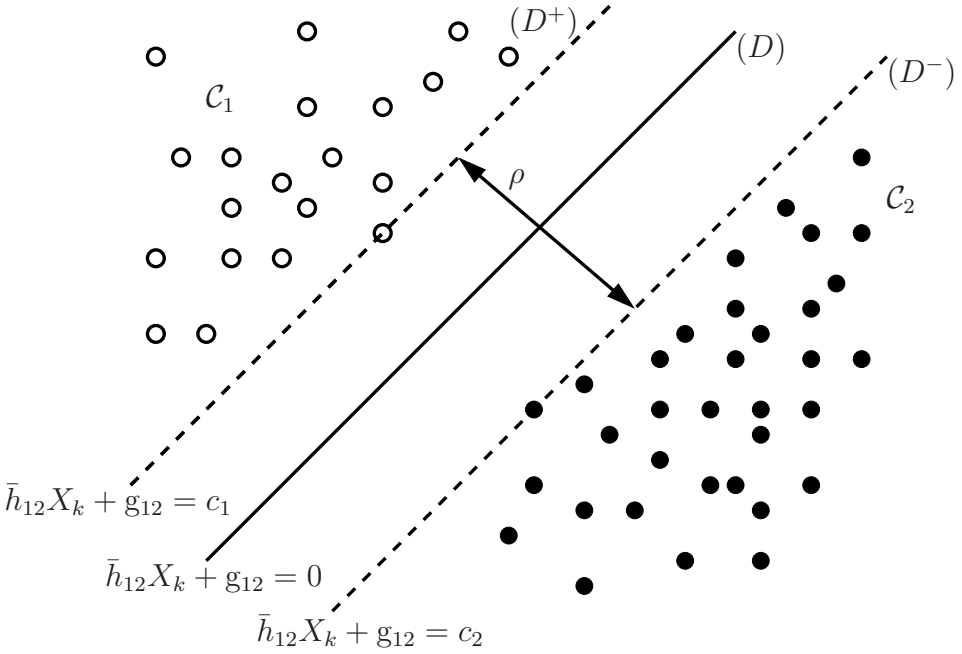


Fig. 4.3: Séparation linéaire de deux classes dans le plan $(y(\cdot), u(\cdot))$

parables, il existe alors une infinité d'hyperplans permettant de les séparer Bennett et Bredensteiner [1993]. On choisit de rechercher l'hyperplan $(\mathcal{H}_{ij})$ qui maximise la marge de séparation entre $\mathcal{C}_i$ et $\mathcal{C}_j$. Sur la figure 4.3, la marge de séparation entre les classes $\mathcal{C}_1$ et $\mathcal{C}_2$ est noté ρ. Cette marge de séparation est calculée en évaluant la distance entre les deux droites parallèles (D^+) et (D^-).

La maximisation de la marge de séparation permet de trouver l'hyperplan séparateur tel que la distance du point du jeu de données le plus proche de cet hyperplan est maximale. Afin d'écrire l'expression générale de la marge de séparation, on introduit la définition de la distance $d(X_k; \mathcal{H}_{ij})$ d'un point X_k à l'hyperplan $\mathcal{H}_{ij}$:

$$d(X_k; \mathcal{H}_{ij}) = \min_{X \in \mathcal{H}_{ij}} \|X - X_k\| \tag{4.14}$$

où $\|\cdot\|$ représente la norme euclidienne.
La marge de séparation ρ entre $\mathcal{C}_i$ et $\mathcal{C}_j$ est alors définie par :

$$\rho = \min_{X_k \in \mathcal{C}_i} d(X_k; \mathcal{H}_{ij}) + \min_{X_k \in \mathcal{C}_j} d(X_k; \mathcal{H}_{ij}) \tag{4.15}$$

Dans [Mangasarian, 1999], on montre que la définition (4.14) est équivalente à :

$$d(X_k; \mathcal{H}_{ij}) = \frac{\left|\bar{h}_{ij} X_k + \mathrm{g}_{ij}\right|}{\left\|\bar{h}_{ij}\right\|} \tag{4.16}$$

Il faut noter en plus que les paramètres $\bar{h}_{ij}$ et g_{ij} de l'hyperplan séparateur des classes $\mathcal{C}_i$ et $\mathcal{C}_j$ sont définis à un coefficient positif multiplicatif près. En effet, si l'hyperplan d'équation $\bar{h}^*_{ij} X_k + \mathrm{g}^*_{ij} = 0$ est une solution au problème de séparation des classes $\mathcal{C}_i$ et $\mathcal{C}_j$ alors tout hyperplan d'équation $\chi \bar{h}^*_{ij} X_k + \chi \mathrm{g}^*_{ij} = 0$, χ étant une constante strictement positive, est également une solution. On peut alors, sans atteinte à la généralité, restreindre la recherche à la classe des hyperplans sous la forme canonique de Vapnik [Vapnik, 1995].

Définition 4.1 (Hyperplan sous forme canonique [Vapnik, 1995])
Soit $\mathcal{H}$ un hyperplan de l'espace euclidien caractérisé par le couple $(\bar{h}; \mathrm{g})$: $(\mathcal{H})$: $\bar{h} X_k + \mathrm{g} = 0$. Soit $\mathcal{S}$ un ensemble de points de cet espace tel que :

$$\min_{X_k \in \mathcal{S}} \left(\left|\bar{h} X_k + \mathrm{g}\right|\right) > 0 \tag{4.17}$$

L'hyperplan $\mathcal{H}_*$ sous forme canonique par rapport à l'ensemble $\mathcal{S}$ correspond au couple $(\bar{h}_*; \mathrm{g}_*)$ (égal au couple $(\bar{h}; \mathrm{g})$ à un coefficient multiplicatif près) tel que :

$$\min_{X_k \in \mathcal{S}} \left(\left|\bar{h}_* X_k + \mathrm{g}_*\right|\right) = 1 \tag{4.18}$$

A partir de la définition de la marge de séparation en (4.15) et de (4.16), il apparaît qu'un hyperplan sous forme canonique possède la propriété que la distance qui le sépare de l'ensemble de points $\mathcal{S}$ est égale à $1/\left\|\bar{h}_*\right\|$. En d'autres termes, la distance séparant l'hyperplan sous forme canonique au point de l'ensemble $\mathcal{S}$ le plus proche de cet hyperplan est égale à $1/\left\|\bar{h}_*\right\|$.
Maximiser la marge de séparation entre les classes $\mathcal{C}_i$ et $\mathcal{C}_j$, en restant dans la

classe des hyperplans sous forme canonique de Vapnik, revient donc à minimiser la norme euclidienne du vecteur $\bar{h}_{ij}$. Il faut donc, pour cela, résoudre le problème d'optimisation sous contraintes (4.19) [Bennett et Bredensteiner, 1993] :

$$\begin{cases} \min\limits_{\bar{h}_{ij},\mathrm{g}_{ij}} \left(\dfrac{1}{2} \left\| \bar{h}_{ij} \right\|^2 \right) \\ \text{sous les contraintes : } \nu_k \left(\bar{h}_{ij} X_k + \mathrm{g}_{ij} \right) \geq 1, \ k = 1, \ldots, N_{ij} \end{cases} \tag{4.19}$$

Remarque 4.1 *En considérant les classes $\mathcal{C}_i$, $i = 1, \ldots, s$ deux à deux et en utilisant (4.6) pour retrouver le partitionnement de l'ensemble des régresseurs Ω, les régions Ω_i, $i = 1, \ldots, s$ obtenues peuvent ne pas former une partition complète de l'espace des régresseurs Ω. Il existe en effet des points X_k pour lesquels $X_k \notin \Omega_i$, $\forall i = 1, \ldots, s$. Sur le graphique de gauche de la figure 4.4, la région grisée correspond à une zone que l'on ne peut associer à aucune des classes $\mathcal{C}_i$, $i = 1, \ldots, 3$. Pour que le partitionnement de l'ensemble des régresseurs soit*

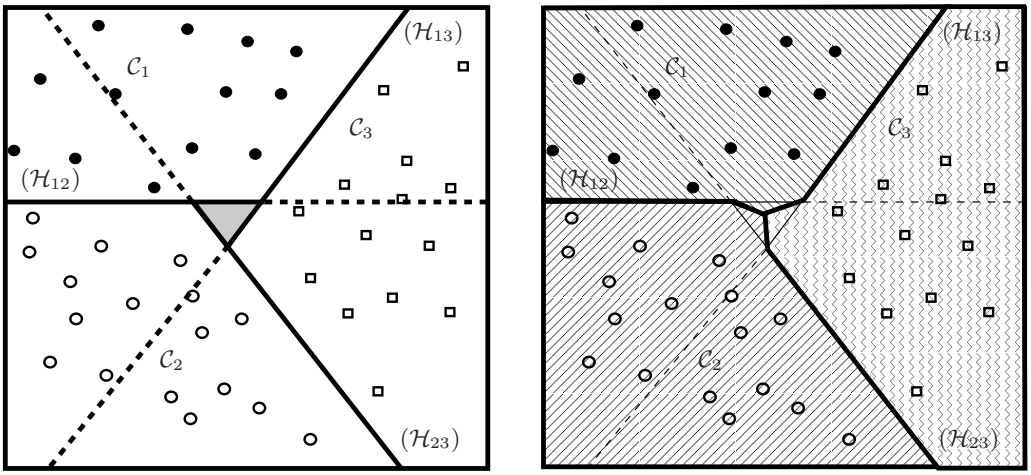


Fig. 4.4: Partitionnement incomplète de l'espace des régresseurs

complet, on peut définir une règle d'affectation qui consisterait à affecter les points $X_k \notin \Omega_i$, $\forall i = 1, \ldots, s$ à la région Ω_i, $i \in \{1, \ldots, s\}$ la plus proche de ce point. Le graphique de droite de la figure 4.4 illustre l'application de cette règle dans le cas de la séparation linéaire des trois classes du graphique de gauche.

Remarque 4.2 *Dans le cadre de la recherche d'un domaine de solutions sous forme simple effectuée à la section 4.2.1, on peut également se restreindre à la classe des*

hyperplans sous forme canonique de Vapnik. Il suffit pour cela de modifier les contraintes du problème d'optimisation (4.13) en prenant en compte la condition (4.18). Dans cette configuration, on obtient le problème d'optimisation suivant :

$$\begin{cases} \max\limits_{h_{ij_0}, r_{h_{ij}}} \prod\limits_{m=1}^{n_a+n_b+1} r_{h_{ij},m} \\ \text{sous les contraintes : } r_{h_{i,j}} > 0 \\ \qquad\qquad \nu_k \left(h_{ij_0} + r_{h_{ij}}\right) \varphi_k \geq 1,\ k = 1, \ldots, N_{ij} \\ \qquad\qquad \nu_k \left(h_{ij_0} - r_{h_{ij}}\right) \varphi_k \geq 1,\ k = 1, \ldots, N_{ij} \end{cases} \tag{4.20}$$

où $\varphi_k = [X_k \ \ 1]$, $h_{ij} = [\bar{h}_{ij} \ \ \mathrm{g}_{ij}]$ *et* $h_{ij} = h_{ij_0} + \lambda_{h_{ij}} \otimes r_{h_{ij}}$.

4.2.3 Estimation du partitionnement de l'espace des régresseurs en présence de bruit de mesure

Lors de la phase de reconnaissance du mode actif à chaque instant, on a constaté qu'en présence d'un bruit de mesure, une ambiguïté pouvait porter sur le mode ou le chemin actif détecté. Cette situation correspond à un cas où il existe des risques de mauvaises affectations des points X_k aux classes $\mathcal{C}_i,\ i = 1, \ldots, s$ en utilisant la procédure décrite en (4.3). Les mauvaises affectations conduisent à des classes $\mathcal{C}_i$ et $\mathcal{C}_j,\ i, j \in \{1, \ldots, s\}$ qui ne sont plus linéairement séparables car leur enveloppe convexe s'intersecte. La figure 4.5 illustre, dans un espace à deux dimensions, le problème de la non séparabilité linéaire de deux classes $\mathcal{C}_1$ et $\mathcal{C}_2$. Les enveloppes convexes des deux classes sont tracées en traits discontinus. On peut constater que les deux enveloppes convexes présentent une intersection et que la séparation obtenue à partir de la droite (D) fait ressortir des points mal classés.

Dans le cadre de la recherche d'un domaine de solutions sous forme simple, la non séparabilité linéaire des classes $\mathcal{C}_i$ et $\mathcal{C}_j,\ i, j \in \{1, \ldots, s\}$, influe sous la forme du domaine correspondant à l'ensemble des solutions possibles, domaine défini par l'ensemble d'inégalités linéaires (4.8). Les résultats obtenus en résolvant le problème d'optimisation (4.13), étant fonction de la forme du domaine (4.8), sont donc affectés par le problème de la non séparabilité linéaire des classes $\mathcal{C}_i$ et $\mathcal{C}_j,\ i, j \in \{1, \ldots, s\}$. Toutefois, il faut se rappeler que l'objectif poursuivi en résol-

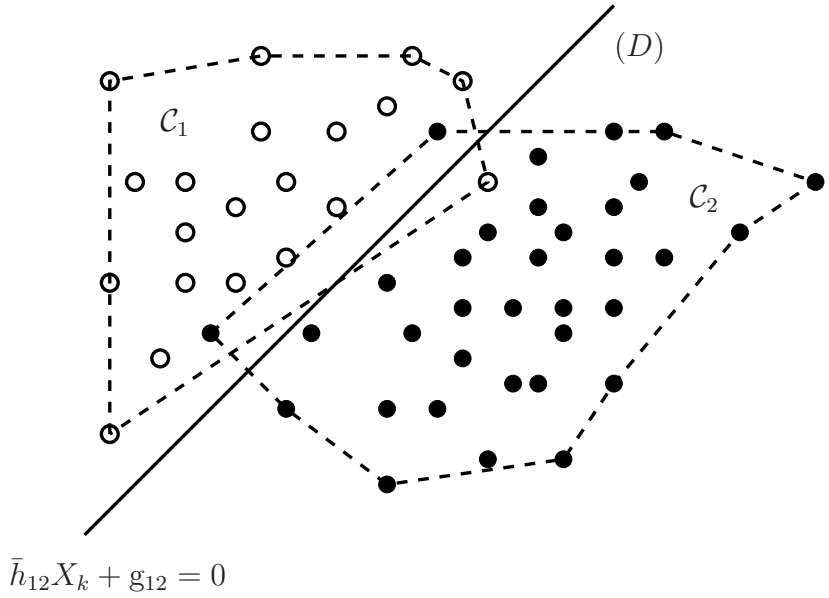


Fig. 4.5: Deux classes non linéairement séparables

vant le problème (4.13) est de décrire le domaine (4.8) par une forme géométrique simple et quelle que soit l'amplitude du bruit de mesure la résolution du problème (4.13) permet d'atteindre cet objectif.

En ce qui concerne la recherche d'une valeur unique pour les paramètres h_{ij} et g_{ij} effectuée à la section 4.2.2, il est possible d'améliorer la qualité de l'estimation des paramètres h_{ij} et g_{ij} en procédant à des modifications du problème (4.19) afin de prendre en compte l'influence du bruit de mesure.

Une première approche, pour la recherche d'un hyperplan séparant les classes $\mathcal{C}_i$ et $\mathcal{C}_j$ lorsqu'elles ne sont pas linéairement séparables, consiste à procéder à une réorganisation du jeux de données $\mathcal{D}$. On rappelle que, lors de la phase de reconnaissance du mode ou du chemin actif, les conditions de discernabilité des chemins permettent d'identifier les instants auxquelles la décision prise quant au chemin actif souffre d'une ambiguïté. On peut donc enlever les points X_k correspondant à ces instants du jeu de données $\mathcal{D}$. Ceci permet d'obtenir des classes C_i et C_j linéairement séparables. Cette façon de procéder, quoique très incisive, permet de réutiliser les méthodes de recherche d'hyperplan séparateur pour des classes linéairement séparables, méthodes exposée aux sections 4.2.1 et 4.2.2.

Dans le cas où les classes $\mathcal{C}_i$ et $\mathcal{C}_j$ ne sont pas linéairement séparables, une autre approche, permettant de trouver un hyperplan séparateur, consiste à définir un cri-

tère permettant d'évaluer l'optimalité de l'hyperplan trouvé par rapport aux points mal classés. Cette idée conduit à la formulation du problème d'optimisation sous contrainte (4.21) :

$$\begin{cases} \min\limits_{\bar{h}_{ij},\mathrm{g}_{ij},\varepsilon_k} \left(\sum\limits_{l=1}^{N_{ij}} c_k \varepsilon_k \right) \\ \text{sous les contraintes : } \nu_k \left(\bar{h}_{ij} X_k + \mathrm{g}_{ij} \right) \geq 1 - \varepsilon_k, k = 1, \ldots, N_{ij} \\ \qquad\qquad \varepsilon_k \geq 0, k = 1, \ldots, N_{ij} \end{cases} \tag{4.21}$$

où les variables $\varepsilon_k,\ k = 1, \ldots, N_{ij}$ sont des variables dites de relaxation qui permettent de tolérer des violations des contraintes $\nu_k \left(\bar{h}_{ij} X_k + \mathrm{g}_{ij} \right) \geq 1,\ k = 1, \ldots, N_{ij}$. $c_k > 0$ est donc une pondération des erreurs dues à la violation des contraintes.
Le problème (4.21) permet de trouver un hyperplan séparateur qui minimise une somme pondérée des erreurs associées aux mauvaises classifications. Il faut noter que la marge de séparation n'intervient pas dans la formulation du critère d'optimisation du problème (4.21). Pour pallier cela, on procède à la modification de (4.21) en recherchant un hyperplan qui minimise la somme des erreurs de mauvaise classification tout en maximisant la marge de séparation entre les classes $\mathcal{C}_i$ et $\mathcal{C}_j$ [Vapnik, 1995]. On obtient, dans ce cas, le problème d'optimisation suivant :

$$\begin{cases} \min\limits_{\bar{h}_{ij},\mathrm{g}_{ij},\varepsilon_k} \left(\frac{1}{2} \left\| \bar{h}_{ij} \right\|^2 + c \sum\limits_{l=1}^{N_{ij}} \varepsilon_k \right) \\ \text{sous les contraintes : } \nu_k \left(\bar{h}_{ij} X_k + \mathrm{g}_{ij} \right) \geq 1 - \varepsilon_k, k = 1, \ldots, N_{ij} \\ \qquad\qquad \varepsilon_k \geq 0, k = 1, \ldots, N_{ij} \end{cases} \tag{4.22}$$

Dans (4.22), les variables $\varepsilon_k,\ k = 1, \ldots, N_{ij}$ demeurent les variables de relaxation permettant de tolérer des violations des contraintes $\nu_k \left(\bar{h}_{ij} X_k + g_{ij} \right) \geq 1,\ k = 1, \ldots, N_{ij}$. Le terme $\sum\limits_{k=1}^{N_{ij}} \varepsilon_l$ permet de quantifier les violations de contraintes. La constante $c, c > 0$, est un paramètre de régularisation qui permet d'effectuer un compromis entre la marge de séparation et le coût des violations de contraintes. Lorsque $c \to \infty$, le problème (4.22) donne plus de poids à la recherche d'un hyperplan minimisant la somme des erreurs de mauvaise classification alors que lorsque

$c \to 0$, (4.22) privilégie la recherche d'un hyperplan minimisant la marge de séparation.

On peut reformuler le problème d'optimisation (4.22) en prenant en compte l'information portant sur les points issus de situation d'ambiguïté quant au chemin ou mode actif. En effet, au lieu de les enlever tout simplement du jeu de données, comme cela a été suggéré plus haut dans cette section, il est possible de les utiliser afin d'avoir une information supplémentaire sur la position de l'hyperplan séparateur des classes $\mathcal{C}_i$ et $\mathcal{C}_j$. Il faut noter que, lorsque les conditions de discernabilité des chemins en absence de bruit de mesure sont respectées, l'ambiguïté portant sur le chemin actif dans un cadre de fonctionnement bruité provient de l'amplitude du bruit de mesure. Ceci pour dire que les points du jeu de données $\mathcal{D}$, qui conduisent à l'obtention de classes non linéairement séparables, sont des points situés au voisinage des hyperplans séparant les domaines d'activation des différents modes du système. A partir de cette remarque, il convient de rechercher l'hyperplan séparateur qui, tout en maximisant la marge de séparation des classes obtenus en enlevant les points provenant de situations ambiguës lors de la reconnaissance du mode ou du chemin actif, minimise la somme des distances de ces points à l'hyperplan. En toute logique, en désignant par X_k^* les points issus de situation ambiguë, il faut donc maximiser la quantité $1/\left\|\bar{h}_{ij}\right\|$, relative à la marge de séparation, et minimiser aussi la quantité $\sum_{k=1}^{N_{ij}^*}\left(\frac{\left|\bar{h}_{ij}X_k^*+\mathrm{g}_{ij}\right|}{\left\|\bar{h}_{ij}\right\|}\right)$, relative à la somme de la distance des points X_k^* à l'hyperplan séparateur, N_{ij}^* étant le nombre de points X_k^*. L'optimisation simultanée des deux quantités précédentes est irréalisable. Pour cela, il convient de redéfinir un critère unique qui prend en compte ces deux quantités. On choisit le critère suivant :

$$\Phi = \left(\frac{1}{\left\|\bar{h}_{ij}\right\|}\right) \Bigg/ \sum_{k=1}^{N_{ij}^*}\left(\frac{\left|\bar{h}_{ij}X_k^* + \mathrm{g}_{ij}\right|}{\left\|\bar{h}_{ij}\right\|}\right), \tag{4.23}$$

qui conduit à :

$$\Phi = \frac{1}{\sum_{k=1}^{N_{ij}^*}\left(\left|\bar{h}_{ij}X_k^* + \mathrm{g}_{ij}\right|\right)} \tag{4.24}$$

Il faut donc maximiser le critère Φ en respectant les contraintes obtenues à partir de l'ensemble des points des classes $\mathcal{C}_i$ et $\mathcal{C}_j$ privé des points $X_k^*,\ k = 1, \ldots, N_{ij}^*$.

Le problème d'optimisation à résoudre est donc le suivant :

$$\begin{cases} \min\limits_{\bar{h}_{ij},\mathrm{g}_{ij},\varepsilon_k} \left(\sum\limits_{k=1}^{N_{ij}^*} \left(\left| \bar{h}_{ij} X_k^* + \mathrm{g}_{ij} \right| \right) \right) \\ \text{sous les contraintes : } \nu_k \left(\bar{h}_{ij} X_k + \mathrm{g}_{ij} \right) \geq 1 - \varepsilon_k, k = 1, \ldots, N_{ij} - N_{ij}^* \\ \qquad\qquad \varepsilon_k \geq 0, k = 1, \ldots, N_{ij} - N_{ij}^* \end{cases} \tag{4.25}$$

Il est clair que la faisabilité et l'efficacité de cette approche dépend du nombre N_{ij}^* de points X_k^* disponibles.

4.3 Estimation du partitionnement de l'espace des régresseurs en recherchant un hyperplan séparateur linéaire par morceaux

On considère ici simultanément l'ensemble des points du jeu de données $\mathcal{D} = \{X_k,\ k = 1, \ldots, N\}$ et on cherche à retrouver les paramètres $\bar{H}_i$ et $g_i,\ i = 1 \ldots, s$ des polyèdres convexes définissant les régions d'activation de chaque mode du système. Pour ce faire, on utilise une démarche issue du domaine de la classification multiclasse. En effet, afin de classer des points provenant de s classes $\mathcal{C}_i,\ i = 1, \ldots, s$, on se sert d'un classificateur linéaire par morceaux. Le classificateur linéaire par morceaux est défini comme étant le maximum de s classificateurs linéaires :

$$\begin{cases} f_i\left(X_k\right) = \bar{h}_i X_k + \mathrm{g}_i \\ f\left(X_k\right) = \max\limits_{i=1,\ldots,s} f_i\left(X_k\right) \end{cases} \tag{4.26}$$

où $\bar{h}_i$ et g_i sont les paramètres correspondant à la classe $\mathcal{C}_\rangle$.
Le classificateur $f(\cdot)$ est en fait une règle d'affectation qui affecte un point X_k à la classe $\mathcal{C}_i$ pour laquelle la quantité $f_i\left(X_k\right),\ i = 1 \ldots, s$ est maximale.
Cette démarche générale de classification conduit à deux types d'approche permettant de trouver des hyperplans séparateurs entre s classes $\mathcal{C}_i,\ i = 1, \ldots, s$.

4.3.1 Approche *one-against-all*

La première approche, qualifiée dans la littérature anglo-saxonne d'approche *one-against-all*, procède en recherchant tout d'abord les paramètres $\bar{h}_i$ et g_i de l'hyperplan, associé à la classe $\mathcal{C}_i$, séparant la classe $\mathcal{C}_i$, des $s-1$ classes restantes [Bredensteiner et Bennett, 1999]. L'opération est effectuée pour toutes les classes et nécessite donc la résolution de s problèmes de séparation linéaire de deux classes. La recherche de cet hyperplan est effectuée en utilisant la méthode présentée à la section 4.2.2. On obtient ainsi dans un premier temps s hyperplans $\mathcal{H}_i$, $i = 1\ldots,s$, l'hyperplan $\mathcal{H}_i$ étant obtenu en séparant la classe $\mathcal{C}_i$ de l'union des $s-1$ classes restantes. La figure 4.6 illustre cette méthode pour le cas de trois classes. Le graphique de gauche montre les hyperplans (droites) séparateurs obtenus en considérant chaque classe avec les deux classes restantes. Chaque hyperplan est obtenu en maximisant la marge de séparation entre chaque classe et l'union des deux autres classes restantes.

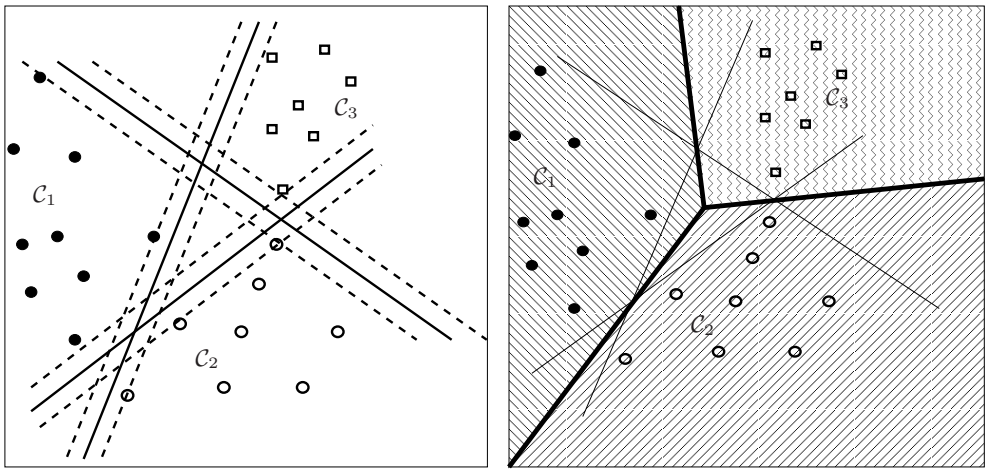


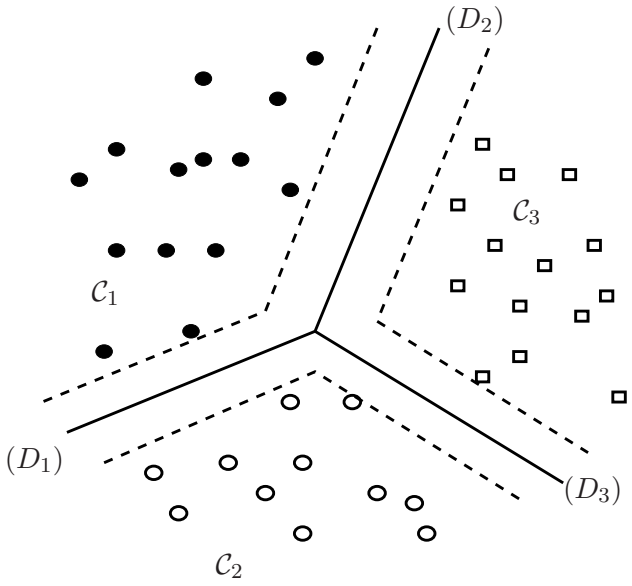
Fig. 4.6: Approche *one-against-all*

A partir des hyperplans $\mathcal{H}_i$, $i = 1\ldots,s$ obtenus, on utilise une règle de décision pour obtenir un séparateur linéaire par morceaux du jeu de données $\mathcal{D}$. La règle de décision est définie par :

$$X_k \in \mathcal{C}_{i^*} \text{ si } i^* = \arg \max_{i=1,\ldots,s} \left(\bar{h}_i X_k + \mathrm{g}_i\right) \tag{4.27}$$

Le graphique de droite de la figure 4.6 montre le séparateur linéaire par morceaux obtenu en appliquant la règle de décision (4.27).
Il faut noter que l'application de cette méthode nécessite que chaque classe $\mathcal{C}_i,\ i \in \{1,\ldots,s\}$ soit linéairement séparable de l'union des $s-1$ classes restantes. Dans la situation où la séparabilité linéaire n'est pas effective, on utilise le problème d'optimisation (4.22) afin de trouver les hyperplans séparateurs $\mathcal{H}_i,\ i = 1\ldots,s$ et on applique ensuite la règle de décision (4.27).

4.3.2 Approche *all-together*



$$(D_1) : (\bar{h}_1 - \bar{h}_2)X_k + (\mathrm{g}_1 - \mathrm{g}_2) = 0$$
$$(D_2) : (\bar{h}_1 - \bar{h}_3)X_k + (\mathrm{g}_1 - \mathrm{g}_3) = 0$$
$$(D_3) : (\bar{h}_2 - \bar{h}_3)X_k + (\mathrm{g}_2 - \mathrm{g}_3) = 0$$

Fig. 4.7: Séparateur linéaire par morceaux pour trois classes

La seconde approche, connue dans la littérature anglo-saxonne sous l'appellation d'approche *all-together*, permettant de séparer les classes $\mathcal{C}_i,\ i = 1\ldots,s$ opère en considérant simultanément l'ensemble des classes et en recherchant directement un hyperplan séparateur linéaire par morceaux [Bennett et Mangasarian, 1994]. On est ainsi ramené à la résolution d'un seul problème d'optimisation sous contrainte mais le coût calculatoire induit est très élevé. Le problème d'optimisation à résoudre

dans ce cas est le suivant [Bredensteiner et Bennett, 1999] :

$$\begin{cases} \min\limits_{\bar{h}_i, \mathrm{g}_i, \varepsilon_k} \left(\dfrac{1}{2} \displaystyle\sum_{i=1}^{s} \sum_{j=1}^{i-1} \left\| \bar{h}_i - \bar{h}_j \right\|^2 + \dfrac{1}{2} \sum_{i=1}^{s} \left\| \bar{h}_i \right\|^2 \right) \\ \text{sous les contraintes : } P_i \left(\bar{h}_i - \bar{h}_j \right) + \left(\mathrm{g}_i - \mathrm{g}_j \right) \mathbb{U} \geq 1 \\ \qquad\qquad\qquad\qquad\qquad\qquad i, j = 1, \ldots, s, \ j \neq i \end{cases} \tag{4.28}$$

où $P_i, \ i \in \{1, \ldots, s\}$ est une matrice élément de $\mathbb{R}^{N_i \times (n_a + n_b)}$ dont les lignes représentent des points de la classe $\mathcal{C}_i, \ i \in \{1, \ldots, s\}$, N_i étant le nombre de points de la classe $\mathcal{C}_i$. $\mathbb{U}$ est un vecteur dont toutes les composantes sont égales à 1.
Dans le cas où, les classes $\mathcal{C}_i, \ i \in \{1, \ldots, s\}$ ne sont pas linéairement séparables par morceaux, le problème d'optimisation (4.28) peut être modifié sous la forme (4.29) afin d'inclure un terme lié aux erreurs de mauvaise classification [Bredensteiner et Bennett, 1999] :

$$\begin{cases} \min\limits_{\bar{h}_i, \mathrm{g}_i, \xi_{ij}} \left(\dfrac{1}{2} \displaystyle\sum_{i=1}^{s} \sum_{j=1}^{i-1} \left\| \bar{h}_i - \bar{h}_j \right\|^2 + \dfrac{1}{2} \sum_{i=1}^{s} \left\| \bar{h}_i \right\|^2 + C \sum_{i=1}^{s} \sum_{\substack{j=1 \\ j \neq i}}^{s} 1^T \xi_{ij} \right) \\ \text{sous les contraintes : } P_i \left(\bar{h}_i - \bar{h}_j \right) + \left(\mathrm{g}_i - \mathrm{g}_j \right) \mathbb{U} \geq \mathbb{U} - \xi_{ij} \\ \qquad\qquad\qquad\qquad\qquad\qquad i, j = 1, \ldots, s, \ i \neq j \\ \qquad\qquad\qquad\qquad\qquad \xi_{ij} \geq 0 \quad i, j = 1, \ldots, s, \ i \neq j \end{cases} \tag{4.29}$$

Conclusion et discussions

Ce chapitre a été consacré à la détermination du mécanisme gérant les commutations d'un mode de fonctionnement à un autre. On a supposé que ce mécanisme est défini par un partitionnement polyédrique de l'ensemble des régresseurs. Deux approches ont été examinées.

La première approche utilise une représentation intervalle des paramètres du mécanisme de commutation pour rechercher un ensemble de valeurs admissibles de ses paramètres, l'objectif étant l'obtention d'un domaine de solution pouvant être décrit par une forme géométrique simple. Cette approche procède en considérant deux à deux les ensembles de données associées aux différents modes du système et présente donc l'avantage de ne pas imposer une charge de calcul trop

importante. L'inconvénient majeure de cette méthode réside dans le fait qu'en présence de bruit de mesure, les résultats obtenus dépendent de la forme du domaine (4.8). La modification de ce domaine, en prenant en compte l'influence du bruit de mesure, est actuellement est un des points de recherche en cours. Il s'agirait de reconnaître parmi l'ensemble des contraintes constituant le domaine (4.8), celles dont, à cause de la présence d'un bruit de mesure, la quantité $(\bar{h}_{ij}X_k + g_{ij})$ est susceptible d'avoir changé de signe. Une fois ces contraintes identifiées, il faudrait les éliminer de l'ensemble des contraintes (4.8) afin d'obtenir un domaine de solutions formés de contraintes « sûres ».

La seconde approche recherche les valeurs optimales des paramètres du mécanisme de commutation au sens d'un critère de distance. La distance considérée dérive de la définition de la distance d'un point à un hyperplan. On distingue dans cette approche deux types de démarche en fonction de la manière dont est utilisé l'ensemble des points associés aux différents modes du système.
Une première démarche considère deux à deux les ensembles de points associés aux divers modes du système, comme dans le cas de la recherche d'un domaine de forme géométrique simple. L'inconvénient de cette démarche est que le partitionnement de l'ensemble des régresseurs obtenus peut ne pas être complet, c'est-à-dire que certaines régions de l'ensemble des régresseurs ne sont associées à aucun mode à la fin de la recherche des paramètres du mécanisme de commutation. Une solution consiste à affecter les points appartenant à ces régions au mode dont la région de fonctionnement est la plus proche.
La seconde démarche met en œuvre un critère qui prend en compte simultanément tous les ensembles de points correspondant aux différents modes du système. On évite ainsi l'obtention d'un partitionnement incomplet de l'ensemble des régresseurs. Le prix à payer à ce niveau est la charge de calcul élevée du fait que tous les ensembles de points associés aux divers modes du système sont simultanément utilisés. Cette démarche est applicable lorsque le jeu de données disponible sur le système ne contient pas un très grand nombre de points.

Il faut noter que, dans les différentes méthodes mises en œuvre, on a le plus souvent considéré la norme euclidienne lors de la formulation des divers problèmes d'optimisation obtenus. Il est légitime de se poser la question de savoir l'impact de

l'utilisation d'une autre norme (norme 1 $\|\cdot\|_1$, norme infini $\|\cdot\|_\infty$) sur le problème d'optimisation. Il est évident qu'en utilisant une autre norme que la norme euclidienne, les domaines obtenus pour le partitionnement de l'ensemble des régresseurs sont différents mais il est, à l'heure actuelle, impossible de se prononcer quant à la norme donnant le meilleur résultat. Il serait à ce niveau intéressant de mettre sur pied un critère permettant de comparer, de façon expérimentale, les capacités de chacune des normes.

Un dernier point à développer est la richesse du jeu de données du système. Si lors de la phase de reconnaissance de mode, les excitations fournies au système conduisent ce dernier à fonctionner majoritairement dans un mode spécifique alors le résultat de la recherche des domaines d'activation de chaque mode sera fortement détérioré. Il faudrait pouvoir garantir l'exhaustivité du jeu de données ou sa capacité à représenter l'ensemble des modes du système avant de procéder à l'identification du mécanisme de commutation. En outre, les points du jeu de données situés aux voisinages des frontières séparant les régions d'activation des modes sont porteurs d'une information importante car ils renseignent sur le positionnement exact de ces frontières. La multiplicité de ces points dans le jeu de données peut permettre l'obtention de meilleurs résultats lors de la recherche du partitionnement de l'ensemble des régresseurs.

5

Etude de cas : exemple de simulation

Sommaire

5.1 Reconnaissance du mode actif 150

5.1.1 Cas déterministe 150

5.1.2 Estimation de l'état du système 154

5.1.3 Influence d'un bruit 156

5.1.4 Taille de l'horizon d'observation 161

5.2 Identification du mécanisme de commutation 164

5.2.1 Recherche d'un domaine de solutions de forme simple . . 165

5.2.2 Recherche d'une solution particulière 166

Introduction

Des exemples académiques sont présentés dans ce chapitre afin d'évaluer les performances et les limites des diverses méthodes développées dans les chapitres 3 et 4 pour la reconnaissance du mode actif à un instant particulier et l'identification des paramètres du mécanisme régissant le passage d'un mode de fonctionnement à un autre.

5.1 Reconnaissance du mode actif

L'exemple académique présenté ici est celui d'un système à commutation comportant trois modes de fonctionnement et modélisé par l'équation (3.1) avec :

$$A_1 = \begin{pmatrix} -0.211 & 0 \\ 0 & 0.521 \end{pmatrix} A_2 = \begin{pmatrix} 0.691 & 0 \\ 0 & -0.310 \end{pmatrix} A_3 = \begin{pmatrix} 0.153 & 0 \\ 0 & 0.410 \end{pmatrix} \tag{5.1}$$

$$B = \begin{pmatrix} 2 & -1 \end{pmatrix}^T C = \begin{pmatrix} 1 & 2 \end{pmatrix}$$

5.1.1 Cas déterministe

La simulation est faite ici dans un contexte déterministe en absence de tout bruit de mesure.
La figure 5.1 indique l'évolution temporelle de l'entrée $u(\cdot)$, de la sortie $y(\cdot)$, de l'état $x(\cdot)$ et des modes $\mu_{(\cdot)}$ du système. Les lignes verticales en pointillés sur le troisième graphique de cette figure indiquent les instants de commutation ou de changement de mode de fonctionnement. Le quatrième graphique retrace l'évolution des modes du systèmes au cours du temps. On peut y voir, par exemple, que sur les fenêtres temporelles $[1,\ 8]$ et $[9,\ 17]$, le système évolue respectivement dans les modes 1 et 2.

La condition (3.53) du théorème 3.3 portant sur la discernabilité des chemins est vérifiée car $C \notin \mathcal{N}_l\left(A_i - A_j\right) = (0\ \ 0)$, $i \neq j \in \{1, 2, 3\}$. Les conditions (3.51) et (3.52) sont testées à chaque instant. Dans le cas où elles ne sont pas satisfaites, aucune décision quant à la reconnaissance du chemin actif n'est prise.
Afin de reconnaître le mode actif à chaque instant à partir des signaux d'entrée

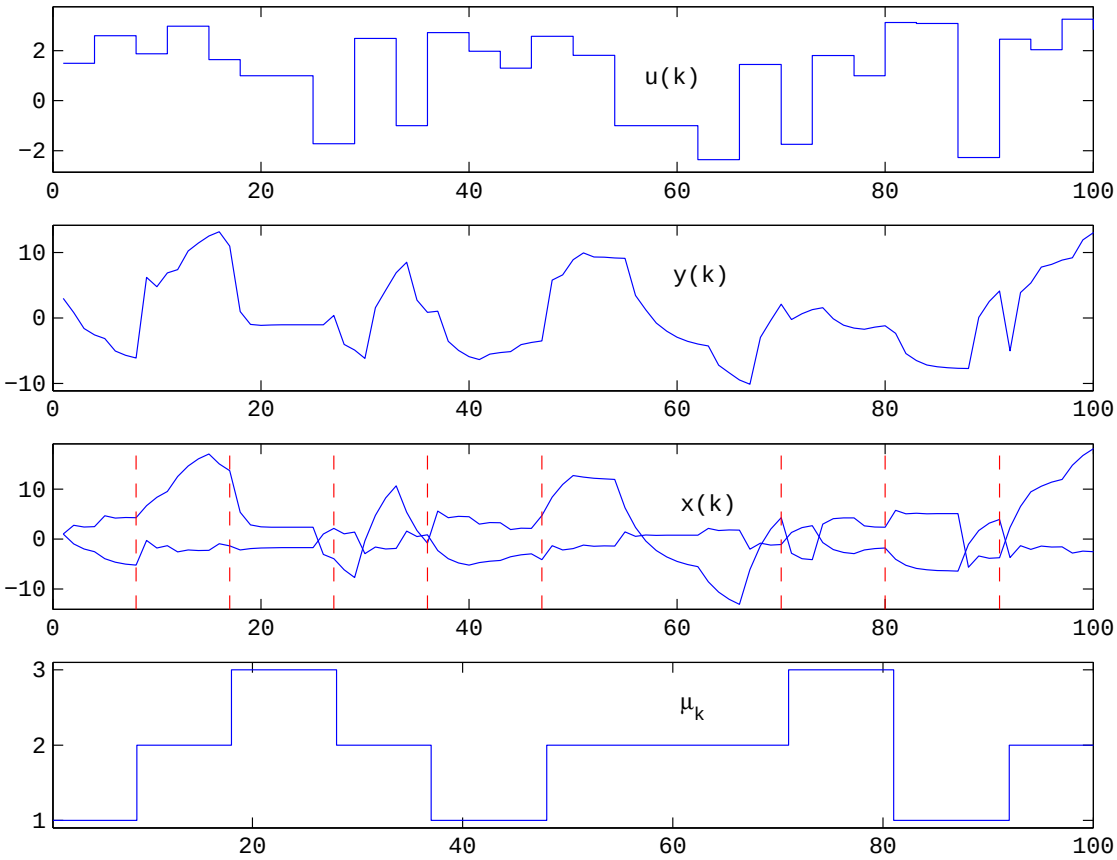


Fig. 5.1: Entrée, sortie, état, loi de commutation

et de sortie du système, on applique la méthode proposée dans le chapitre 3. Pour ce faire, on considère une fenêtre d'observation de taille $h = 2$. L'ensemble Θ_2 des chemins de longueur 2 sur cet horizon correspond à l'ensemble des neuf chemins présentés à la table 5.1.

Tab. 5.1: Ensemble des chemins de longueur 2

N° du chemin	1	2	3	4	5	6	7	8	9
μ_1	1	1	1	2	2	2	3	3	3
μ_2	1	2	3	1	2	3	1	2	3

Les matrices d'observabilité $\mathcal{O}_{\mu,h}$ calculées sont données par :

$$\begin{cases} \mathcal{O}_{(1\cdot 1),h} = \begin{pmatrix} C \\ CA_1 \\ CA_1A_1 \end{pmatrix} & \mathcal{O}_{(1\cdot 2),h} = \begin{pmatrix} C \\ CA_2 \\ CA_1A_2 \end{pmatrix} & \mathcal{O}_{(1\cdot 3),h} = \begin{pmatrix} C \\ CA_3 \\ CA_1A_3 \end{pmatrix} \\ \mathcal{O}_{(2\cdot 1),h} = \begin{pmatrix} C \\ CA_1 \\ CA_2A_1 \end{pmatrix} & \mathcal{O}_{(2\cdot 2),h} = \begin{pmatrix} C \\ CA_2 \\ CA_2A_2 \end{pmatrix} & \mathcal{O}_{(2\cdot 3),h} = \begin{pmatrix} C \\ CA_3 \\ CA_2A_3 \end{pmatrix} \\ O_{(3\cdot 1),h} = \begin{pmatrix} C \\ CA_1 \\ CA_3A_1 \end{pmatrix} & \mathcal{O}_{(3\cdot 2),h} = \begin{pmatrix} C \\ CA_2 \\ CA_3A_2 \end{pmatrix} & \mathcal{O}_{(3\cdot 3),h} = \begin{pmatrix} C \\ CA_1 \\ CA_3A_3 \end{pmatrix} \end{cases}$$

La figure 5.2 présente l'évolution des différents résidus calculés. Il s'agit des résidus $r_{(i\cdot j),h}(\cdot)$, $i,j \in \{1,2,3\}$ associés aux chemins de longueur 2 de la table 5.1. On constate qu'un seul des neufs résidus est nul à chaque instant et l'indice $(i \cdot j)$ de ce résidu correspond au chemin actif sur l'horizon temporel considéré. Par exemple, de l'instant $k = 1$ jusqu'à l'instant $k = 6$, le résidu $r_{(1\cdot 1),h}(\cdot)$ (première ligne et première colonne de la figure 5.2) a une valeur nulle, montrant que le système est resté dans le mode 1 sur l'horizon $[1,\ 6]$. A l'instant $k = 7$, seul le résidu $r_{(1\cdot 2),h}(\cdot)$ est nul, montrant ainsi la reconnaissance du chemin $(1 \cdot 2)$ et donc une commutation du mode 1 vers le mode 2 à l'instant $k = 9$.

Le tableau précise de façon numérique les résultats de la figure 5.2 entre les instants $k = 1$ et $k = 20$. En colonnes 1 à 6 sont indiqués respectivement le temps et les neuf résidus. On peut noter les commutations aux instants $k = 7$ et $k = 16$ sont parfaitement détectés. Les deux commutations sont marquées par des annulations respectives des résidus $r_{(1\cdot 2),h}$ à l'instant $k = 7$ et $r_{(2\cdot 3),h}$ à l'instant $k = 16$.

La figure 5.3 montre l'évolution des modes du système ainsi que les modes détectés à partir de l'analyse des résidus. On peut y voir que le mode actif à chaque instant est parfaitement détecté.

Sur la figure 5.2, les graphiques se trouvant sur la diagonale sont relatifs aux résidus obtenus pour les chemins $(i \cdot i)$, $i \in \{1,2,3\}$ alors que les autres graphiques sont relatifs aux résidus obtenus pour les chemins $(i \cdot j)$, $i \neq j \in \{1,2,3\}$. De ce fait, les résidus $r_{(i\cdot j),h}(\cdot)$, $i \neq j \in \{1,2,3\}$ sont fort utiles lorsque les commutations du

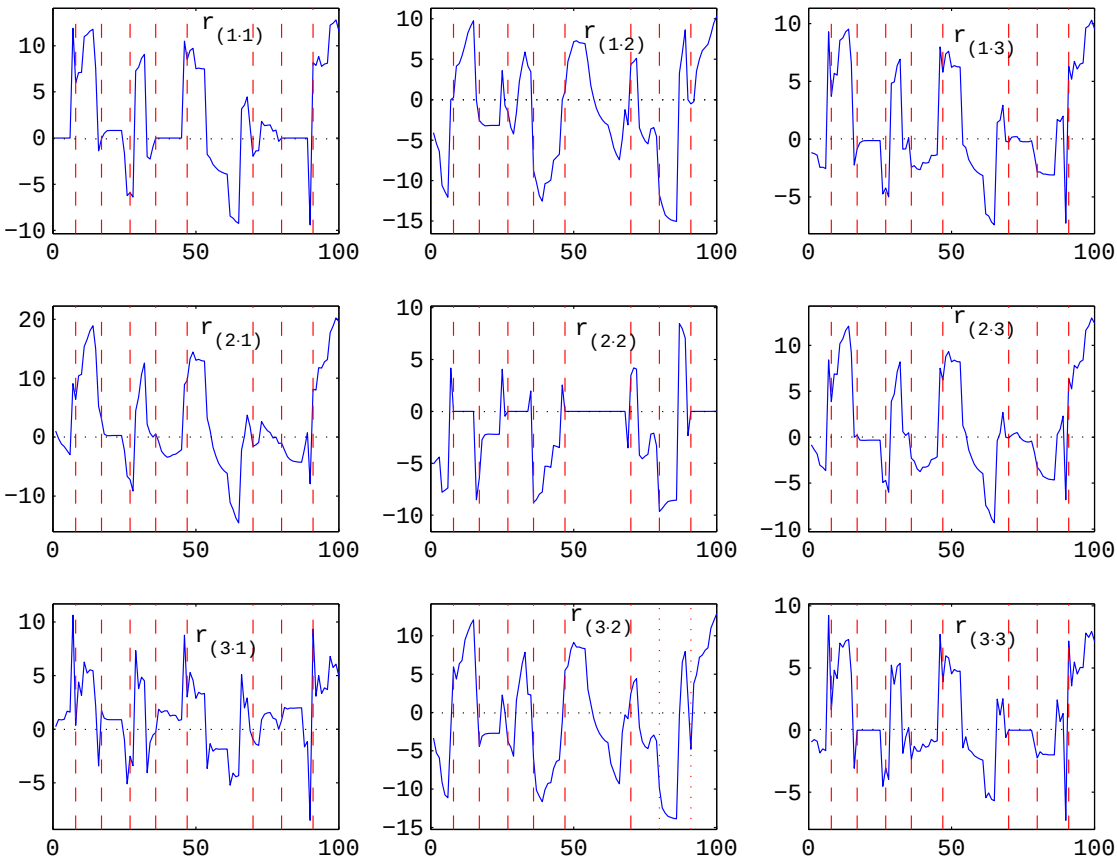


Fig. 5.2: Résidus associés aux chemins de longueur 2

Tab. 5.2: Valeurs des résidus

k	$r_{(1\cdot1),h}$	$r_{(1\cdot2),h}$	$r_{(1\cdot3),h}$	$r_{(2\cdot1),h}$	$r_{(2\cdot2),h}$	$r_{(2\cdot3),h}$	$r_{(3\cdot1),h}$	$r_{(3\cdot2),h}$	$r_{(3\cdot3),h}$
1	0.00	-4.06	-1.16	1.03	-5.07	-0.83	0.23	-3.32	-0.96
2	0.00	-5.41	-1.24	-0.24	-4.71	-1.36	0.91	-5.265	-0.76
3	0.00	-6.36	-1.38	-1.17	-4.41	-1.82	0.87	-5.905	-0.90
4	0.00	-10.60	-2.43	-1.60	-7.79	-3.05	0.93	-9.113	-1.84
5	0.00	-11.37	-2.42	-2.30	-7.61	-3.28	1.69	-10.747	-1.51
6	0.00	-12.07	-2.53	-2.99	-7.39	-3.64	1.60	-11.140	-1.65
7	11.89	0.00	9.29	9.08	4.17	8.40	10.63	-2.31	9.24
8	5.97	0.253	3.70	6.39	0.00	3.86	0.36	5.94	1.64
9	7.09	4.13	5.69	10.40	0.00	6.89	4.41	4.36	4.83
10	7.10	4.47	5.55	10.65	0.00	6.83	3.15	6.36	4.16
11	10.98	5.43	8.55	15.38	0.00	10.15	6.27	6.79	7.00
12	11.22	6.74	8.76	16.60	0.00	10.70	5.24	9.46	6.67
13	11.58	8.25	9.28	18.07	0.00	11.61	5.55	10.52	7.17
14	11.77	9.10	9.53	18.89	0.00	12.08	5.45	11.49	7.29
15	7.52	9.75	6.55	14.95	0.00	9.19	2.16	12.09	4.51
16	-1.39	-0.077	-2.24	5.21	-8.52	0.00	-3.417	3.235	-3.06
17	-0.08	-2.55	-0.84	3.09	-6.43	0.24	1.750	-4.41	0.00
18	0.50	-2.96	-0.31	0.50	-2.72	-0.32	1.038	-3.10	0.00
19	0.79	-3.22	-0.19	0.24	-2.25	-0.35	0.917	-2.82	-0.00
20	0.82	-3.21	-0.13	0.25	-2.20	-0.34	0.894	-2.74	0.00

système sont assez rapprochées puisqu'ils marquent plus précisement les instants de changement de mode du système. Comme cela a été expliqué dans la section 3.2.3, on peut ne pas prendre en compte ces résidus dans la phase de reconnaissance du chemin actif pour des systèmes avec des fréquences de commutation relativement faibles.

La figure 5.4 effectue un zoom sur les trois graphiques se situant sur la diagonale de la figure 5.2 (résidus $r_{(1\cdot1),h}(\cdot)$, $r_{(2\cdot2),h}(\cdot)$ et $r_{(3\cdot3),h}(\cdot)$)). On peut y voir qu'un seul des trois résidus est nul à chaque instant, excepté dans un voisinage des instants de commutation. La non nullité des résidus dans ce voisinage provient du fait qu'aucun des trois chemins $(1\cdot1)$, $(2\cdot2)$ et $(3\cdot3)$ ne correspond au chemin actif sur les horizons temporels considérés dans ce voisinage.

5.1.2 Estimation de l'état du système

L'évolution des modes du système étant connue, on peut envisager de procéder à l'estimation de l'état du système. L'analyse étant effectuée sur un horizon fini, on ne peut mettre en œuvre que des observateurs à mémoire finie afin de conserver une cohérence dans la démarche proposée. Considérons à nouveau l'équation (3.6)

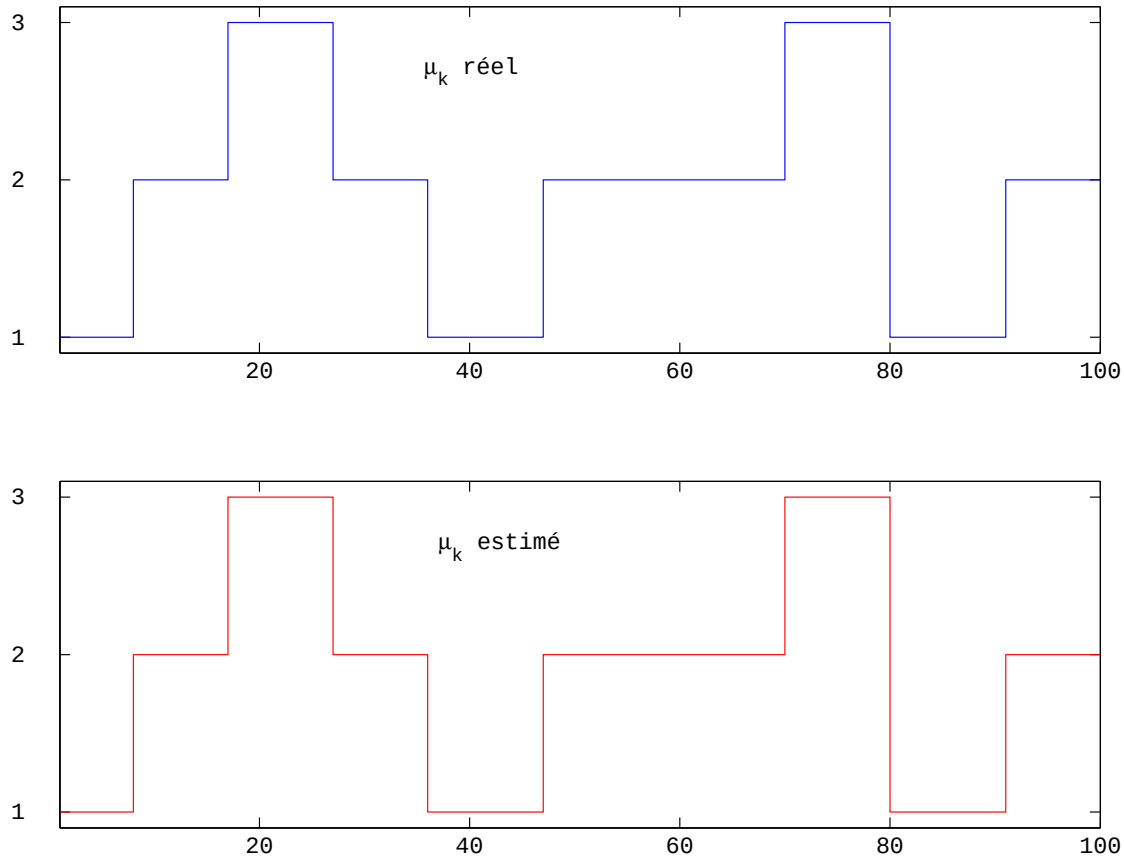


Fig. 5.3: Reconnaissance des modes

décrivant l'observation du système sur une fenêtre temporelle $[k,\ k-h]$:

$$Y_{k-h,k} - T_{\mu,h} U_{k-h,k} = \mathcal{O}_{\mu,h} x(k-h)$$

Comme les matrices $\mathcal{O}_{\mu,h}$ sont, par construction, de plein rang colonne, les matrices $\mathcal{O}_{\mu,h}\mathcal{O}_{\mu,h}^T$ sont donc inversibles. L'état du système peut être estimé à l'aide de l'équation d'observation (3.6) et l'état estimé $\hat{x}(\cdot)$ est donné par :

$$\hat{x}\,(k-h) = \left(\mathcal{O}_{\mu,h}^T \mathcal{O}_{\mu,h}\right)^{-1} \mathcal{O}_{\mu,h}^T \left(Y_{k-h,k} - T_{\mu,h} U_{k-h,k}\right)$$

L'état estimé s'exprime donc linéairement en fonction des mesures de l'entrée et de la sortie du système disponibles sur l'horizon d'observation.
Les états réels et reconstruits sont présentés à la figure 5.5 pour l'exemple précédent. On peut constater une superposition parfaite des états réels et reconstruits.

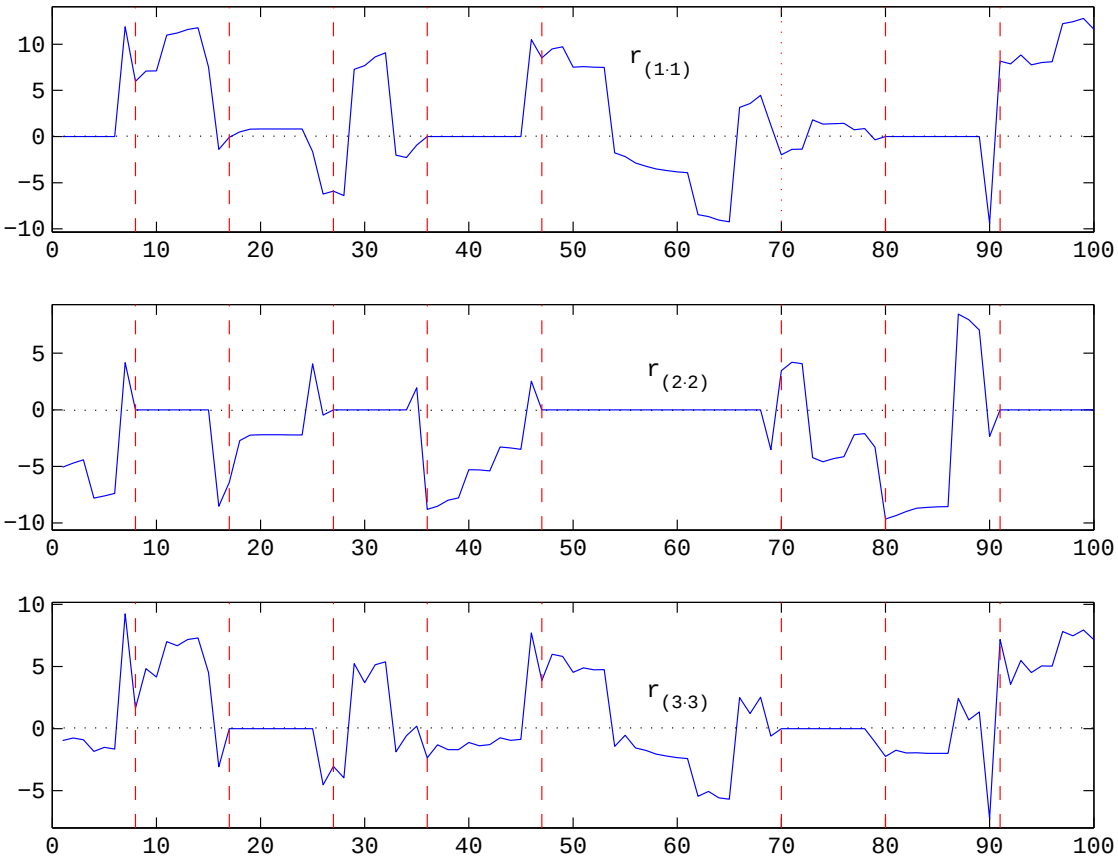


Fig. 5.4: Résidus associés aux chemins $(1 \cdot 1)$, $(2 \cdot 2)$ et $(3 \cdot 3)$

5.1.3 Influence d'un bruit

A partir du système modélisé par (3.1) et caractérisé par les matrices (5.1), on montre la reconnaissance du chemin ou du mode actif en présence d'un bruit de mesure borné. La figure 5.6 est relative à l'entrée, la sortie, l'état et l'évolution des modes du système. Comme précédemment, les lignes verticales en pointillés indiquent les instants de commutation du système.

La figure 5.7 montre l'évolution au cours du temps du rapport de l'amplitude du bruit de mesure sur celle de la sortie du système en termes de pourcentage. Ce rapport est au maximum approximativement égal à 25%.

La reconnaissance du chemin actif est effectuée en générant des résidus intervalle. Seuls les chemins $(i \cdot i)$, $i \in \{1, 2, 3\}$ sont considérés. La figure 5.8 illustre l'évolution temporelle des différents résidus intervalles (en traits discontinus) obtenus. En présence de bruit, on constate qu'effectivement les résidus (en traits continus)

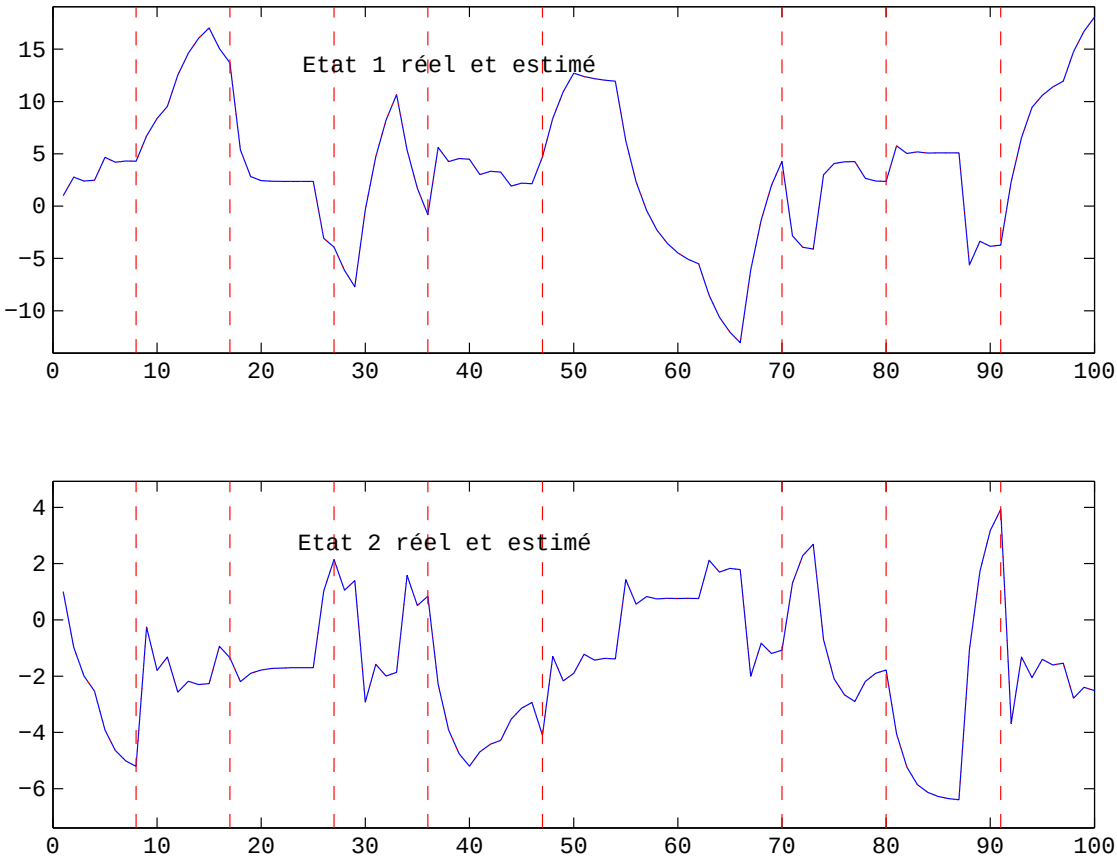


Fig. 5.5: Etats réels et estimés

ne prennent pas une valeur nulle lorsque leur chemin correspondant est actif mais plutôt une valeur proche de zéro et un seul des trois résidus intervalle contient à chaque instant la valeur zéro. Ce résidu correspond au chemin actif sur l'horizon considéré.

Le second et le troisième graphique de la figure 5.9 montrent les résultats de la reconnaissance du chemin actif suite à l'analyse des résidus intervalle générés. Le second graphique montre les modes détectés en testant l'appartenance de la valeur zéro aux résidus intervalles calculés. Bien que les modes soient assez bien détectés dans l'ensemble, on note, comme cela a été dit précédemment, la présence d'instants où il a été impossible de fournir une estimation de $\mu_{(\cdot)}$ à cause du fait que plus d'un résidu intervalle contenait la valeur zéro ou aucun résidu intervalle ne contenait la valeur zéro. La présence de ces instants, indiqués sur le second graphique de la figure 5.9 par des points d'ordonnée nulle, provient de deux raisons. La première raison est que tous les chemins possibles sur l'horizon d'observation n'ont

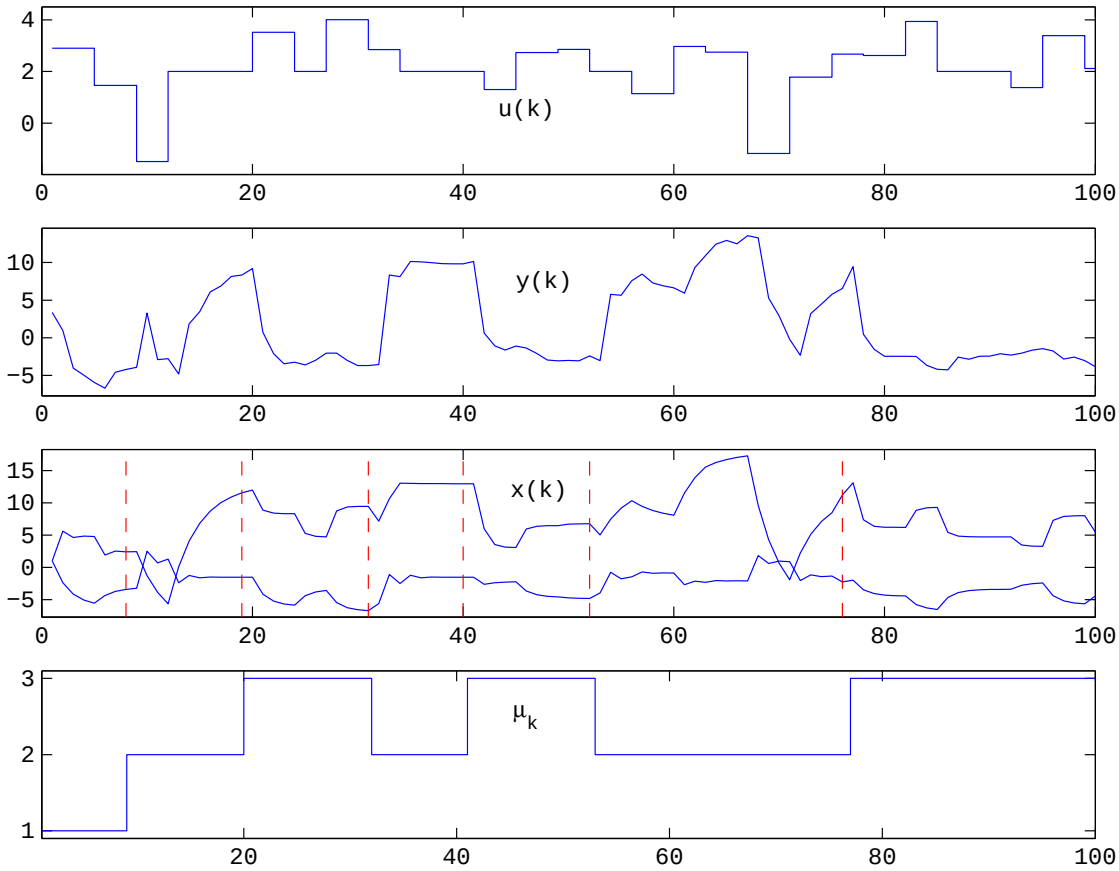


Fig. 5.6: Entrée, sortie, état et modes du système

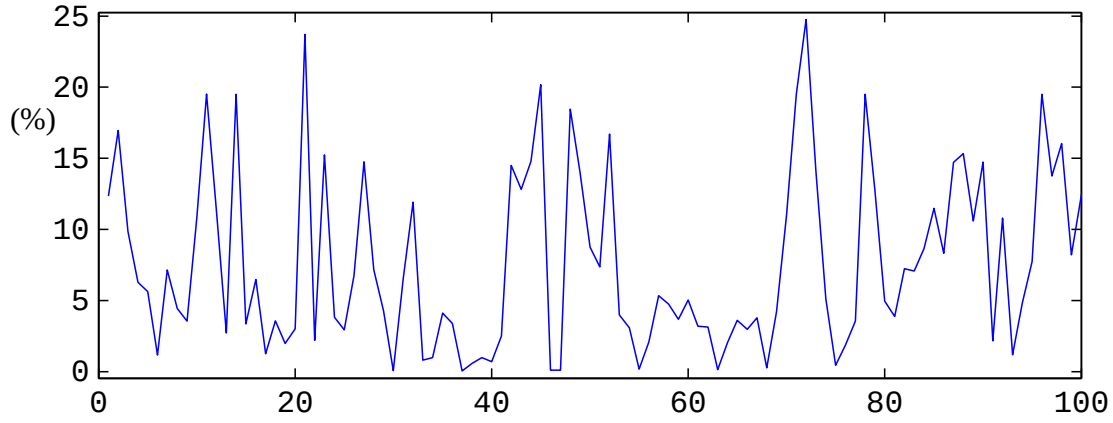


Fig. 5.7: Amplitude du bruit sur l'amplitude de la sortie en pourcentage

pas été considérés. En fait, au voisinage des instants de commutation, aucun des chemins considérés ne correspond au chemin actif sur la fenêtre d'observation. La seconde raison est liée à l'influence du bruit de mesure qui est de nature à modifier le positionnement des résidus intervalle par rapport à la valeur zéro. Le troisième

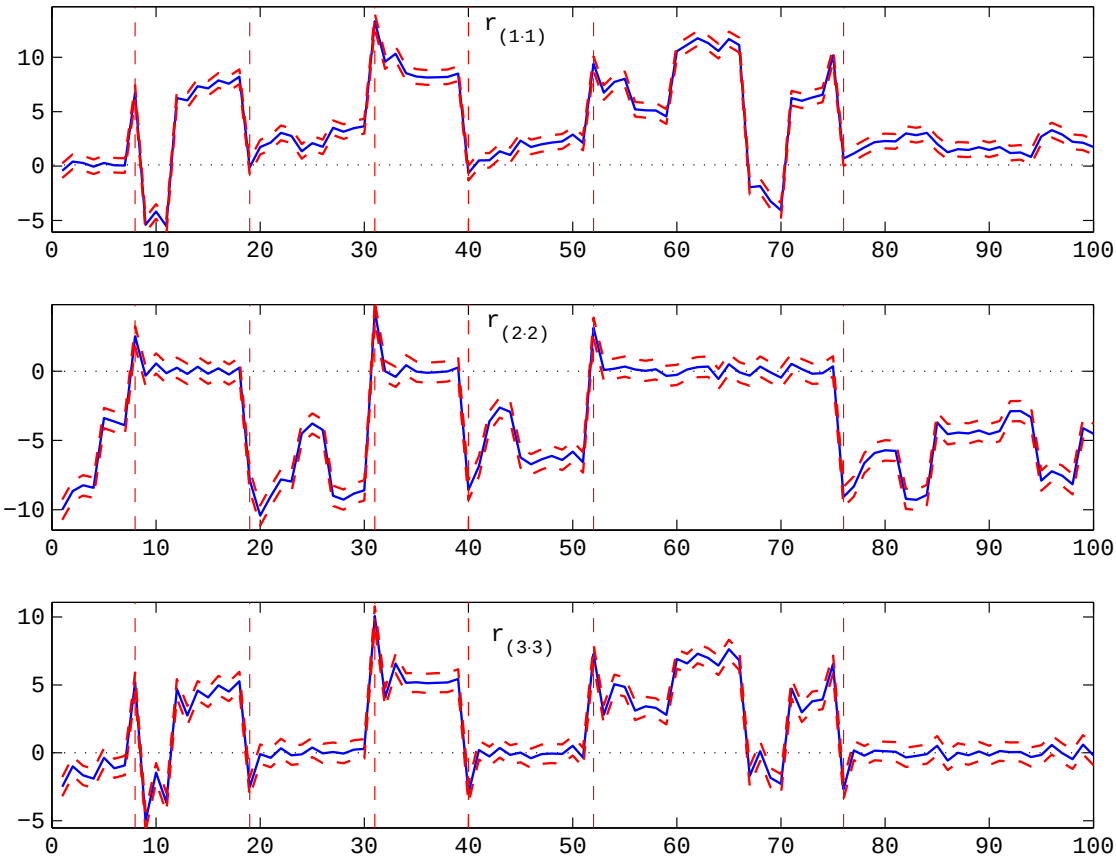


FIG. 5.8: Evolution des résidus

graphique de la figure 5.9 est obtenu en procédant à une analyse de continuité des séquences comme indiqué dans la section 3.2.3. Cette analyse est conduite en calculant les résidus $r_{(i\cdot j),h}(\cdot)$, $i \neq j \in \{1,2,3\}$ à l'instant précédant l'instant auquel la reconnaissance du mode actif n'a pas été possible compte tenu de la non discernabilité des chemins considérés sur la fenêtre d'observation. On utilise ensuite la continuité des chemins pour discriminer les chemins non discernables. Il s'agit de tester la cohérence dans la succession des chemins détectés à des instants qui se suivent. Pour cela, lorsqu'à un instant quelconque k_0, le chemin actif μ^{1^*} a été identifié sur une fenêtre d'observation $[k_0-h, k_0]$, à l'instant suivant k_0+1 on vérifiera si l'infixe du chemin μ^{2^*} détecté est identique à l'infixe du chemin μ^{1^*} détecté précédemment à l'instant k_0 : $\mu^{1^*}_{[k_0+1-h,k_0-1]} = \mu^{2^*}_{[k_0+1-h,k_0-1]}$.

Bien que le troisième graphique de la figure 5.9 montre une parfaite reconstruction de l'évolution des modes du système, il faut garder en esprit que, sur une fenêtre d'observation, l'appartenance de la valeur zéro à plus d'un résidu intervalle indique

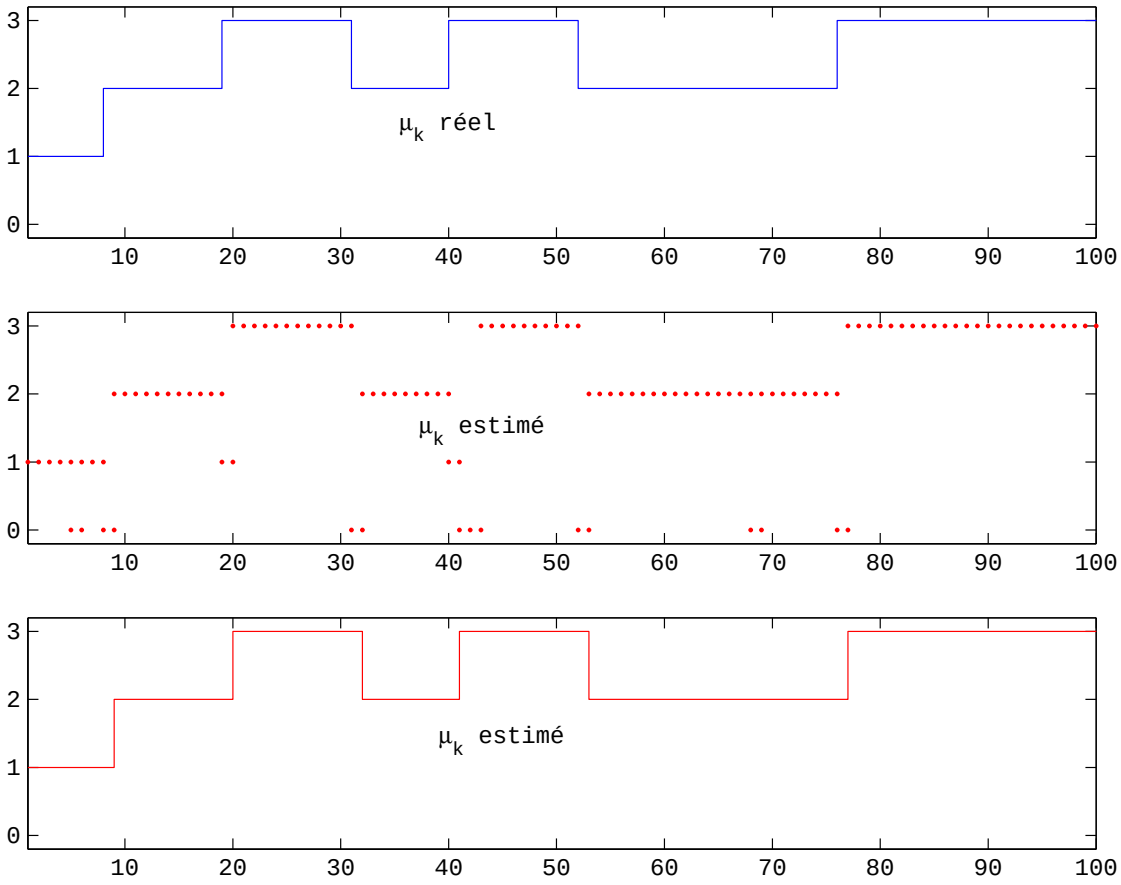


Fig. 5.9: Reconnaissance du chemin actif

une situation de non discernabilité des chemins considérés sur la fenêtre d'observation. Pour cela, les résultats obtenus en utilisant le principe de continuité des chemins sont fonction de l'amplitude du bruit. Dans le cas où le temps minimal de séjour dans un mode est spécifié, on peut également utiliser cette information pour de meilleurs résultats.

Comme dans le cas sans bruit, il est également possible, une fois l'évolution des modes reconstruite, d'estimer l'état du système à l'aide d'un observateur à mémoire fini. La figure 5.10 visualise les estimations d'état (en traits discontinus) comparées aux états réels (en traits continus). L'état du système est globalement bien reconstruit.

Pour aller plus loin, étant donné que le bruit considéré est borné, on peut à l'aide des outils de l'arithmétique intervalle étendre l'observateur à mémoire finie à l'estimation des bornes de l'état du système. La figure 5.11 présente les résultats de

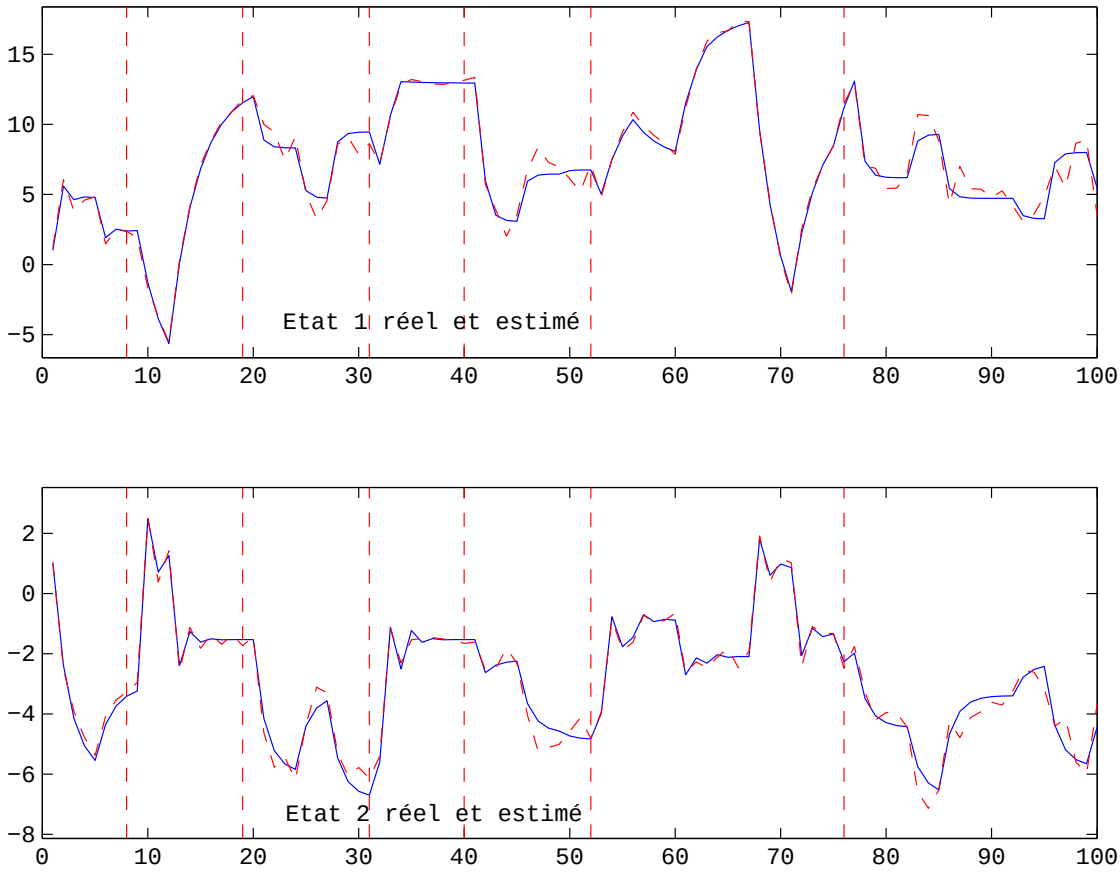


Fig. 5.10: Etats réels et estimés

l'estimation. L'état réel est toujours inclus dans les bornes estimées.

On présente, à titre purement indicatif, à la figure 5.12 les résidus intervalle obtenus si tous les chemins possibles sont considérés sur la fenêtre d'observation.

5.1.4 Taille de l'horizon d'observation

Dans un cadre de fonctionnement sans bruit, la taille de la fenêtre d'observation $[k-h,\ k]$ est obtenue en considérant la plus petite valeur de h pour laquelle les matrices d'observabilité $\mathcal{O}_{\mu,h}$ générées pour les différents chemins possibles sont toutes de plein rang colonne.

En présence de bruit, on a constaté que l'augmentation de la taille de la fenêtre d'observation induisait un effet de filtrage sur les résidus, les rendant ainsi moins sensibles au bruit de mesure. Cette hypothèse a été testée de façon numérique. Le test est conduit en considérant différentes valeurs de h. Pour chaque valeur de h,

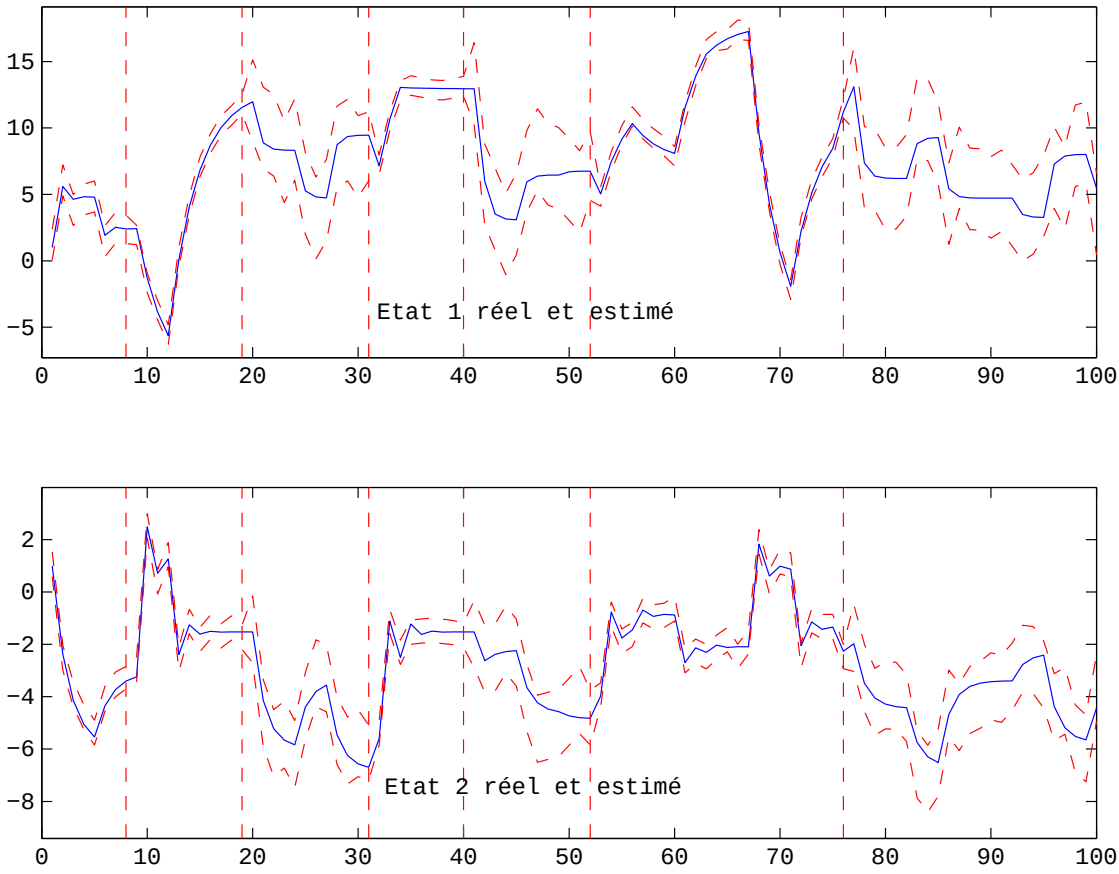


Fig. 5.11: Estimation des bornes de l'état

1000 simulations sont effectuées en vue de procéder à la reconnaissance du mode ou du chemin actif à chaque instant. Dans chaque cas de simulation, on calcule le taux de bonne détection des modes à partir de l'analyse des résidus. A partir du taux de bonne détection du mode actif, on calcule finalement le taux moyen de bonne détection pour l'ensemble des 1000 simulations. La figure 5.13 montre l'évolution du taux moyen de bonne détection en fonction de la taille de la fenêtre d'observation. On peut y voir une augmentation du taux de bonne détection jusqu'à une certaine valeur h_{opt} à partir de laquelle ce taux commence par chuter. La valeur h_{opt} semble correspondre à la taille optimale de l'horizon d'observation.

La courbe de variation du taux de bonne détection en fonction de h a été tracée pour diverses valeurs de la borne δ du bruit de mesure sur la sortie du système. Les résultats obtenus sont représentés à la figure5.13. On note, comme à la figure 5.13, la présence d'un horizon de taille optimale qui permet un meilleur taux de détection.

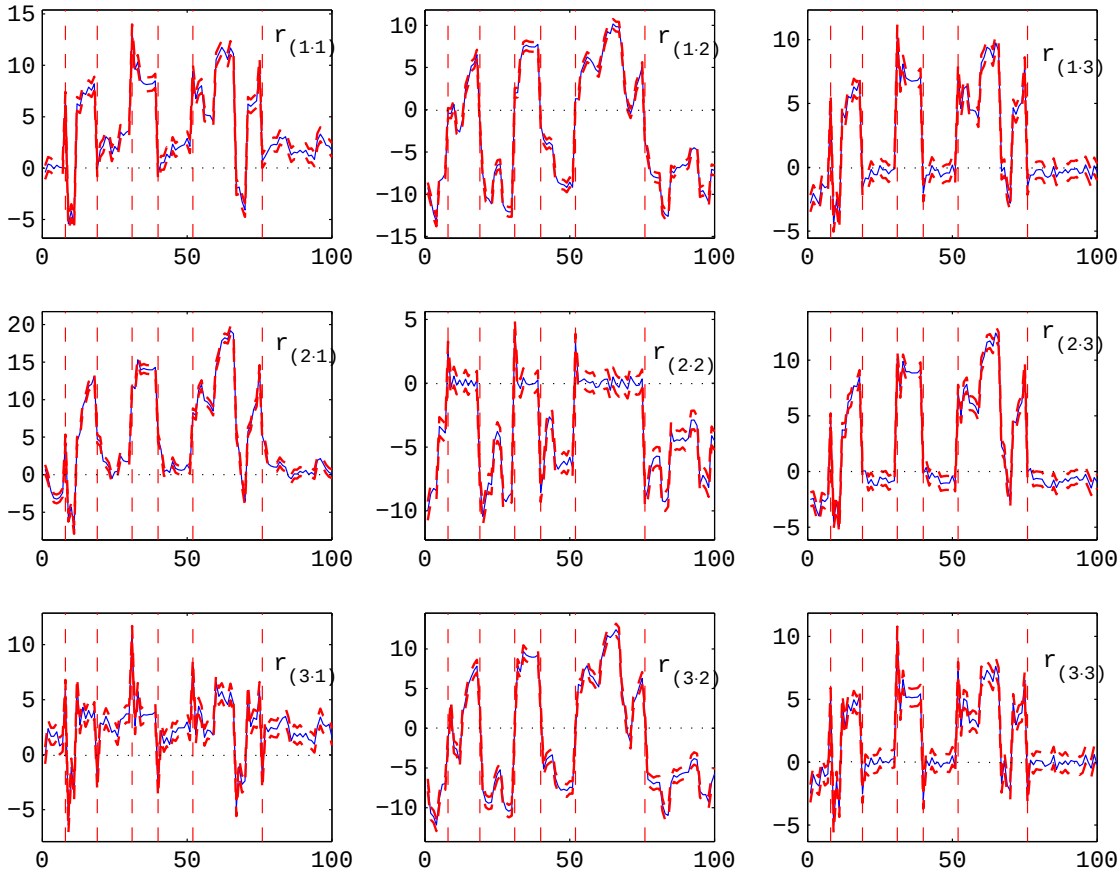


Fig. 5.12: Résidus intervalle de tous les chemins

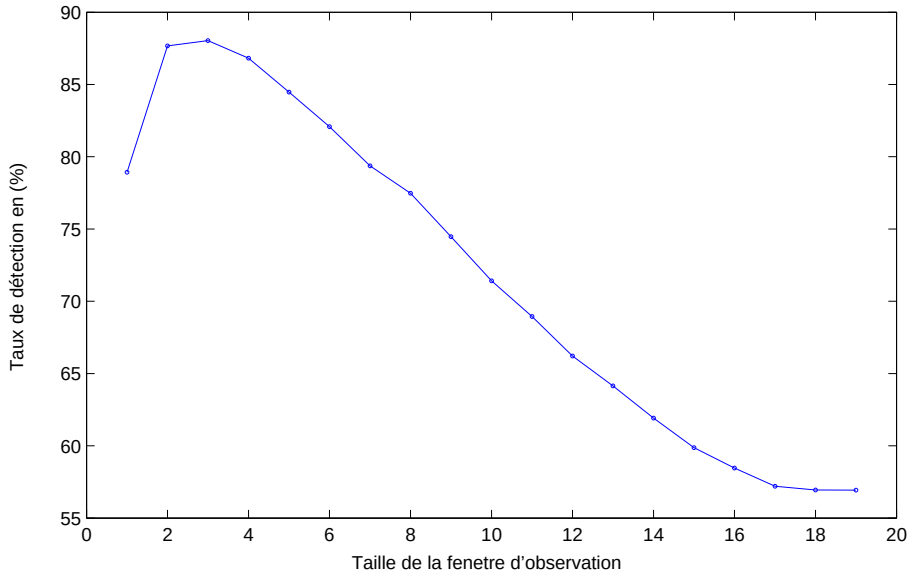


Fig. 5.13: Evolution du taux de détection en fonction de la taille de la fenêtre d'observation

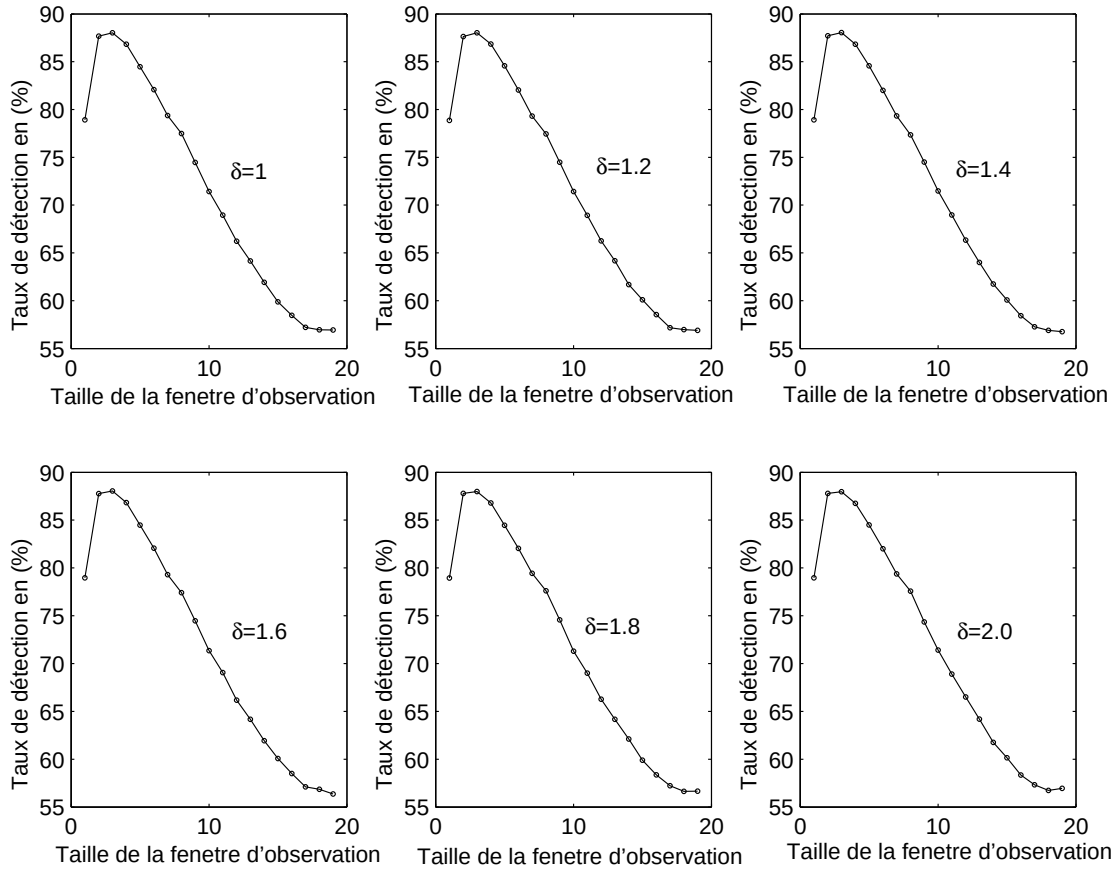


Fig. 5.14: Influence de la variance du bruit sur la taille optimale de l'horizon d'observation

5.2 Identification du mécanisme de commutation

On considère à présent le modèle (3.1), les matrices caractérisant les modèles associés aux différents modes du système étant données par (5.1). On modélise ensuite le mécanisme de commutation en partitionnant l'espace $\{y_{k-1}; u_{k-1}\}$ en trois régions définies de la manière suivante :

$$\begin{cases} \mu_k = 1 \text{ si } \left(h_1 \begin{bmatrix} y_{k-1} & u_{k-1} & 1 \end{bmatrix}^T\right) \geq 0 \text{ et } \left(h_2 \begin{bmatrix} y_{k-1} & u_{k-1} & 1 \end{bmatrix}^T\right) \geq 0 \\ \mu_k = 2 \text{ si } \left(h_1 \begin{bmatrix} y_{k-1} & u_{k-1} & 1 \end{bmatrix}^T\right) < 0 \text{ et } \left(h_3 \begin{bmatrix} y_{k-1} & u_{k-1} & 1 \end{bmatrix}^T\right) < 0 \\ \mu_k = 3 \text{ si } \left(h_3 \begin{bmatrix} y_{k-1} & u_{k-1} & 1 \end{bmatrix}^T\right) \geq 0 \text{ et } \left(h_2 \begin{bmatrix} y_{k-1} & u_{k-1} & 1 \end{bmatrix}^T\right) < 0 \end{cases} \quad (5.2)$$

avec $h_1 = \begin{bmatrix} 1 & 0.51 & 0 \end{bmatrix}$, $h_2 = \begin{bmatrix} 0 & 1 & 0 \end{bmatrix}$, $h_3 = \begin{bmatrix} 1 & -0.29 & 0 \end{bmatrix}$.

5.2.1 Recherche d'un domaine de solutions de forme simple

A partir des données entrée/sortie collectées, on procède dans un premier temps à la reconnaissance du mode actif à chaque instant, ce qui permet de savoir la valeur prise par $\mu_{(\cdot)}$ à chaque instant. On peut alors ensuite former en fonction des valeurs de $\mu_{(\cdot)}$ trois classes $\mathcal{C}_1$, $\mathcal{C}_2$ et $\mathcal{C}_3$ de points $(y_{k-1}; u_{k-1})$. Ces trois classes sont présentées à la figure 5.15. Les droites séparant les différentes classes sont aussi représentées sur cette figure. Dans le jeu de données généré, la classe $\mathcal{C}_1$ contient 41 points, $\mathcal{C}_2$ en contient 14 et on en dénombre 45 pour $\mathcal{C}_3$.

Comme mentionné dans le chapitre 4, page 133 § 4.2.1, on procède à la recherche

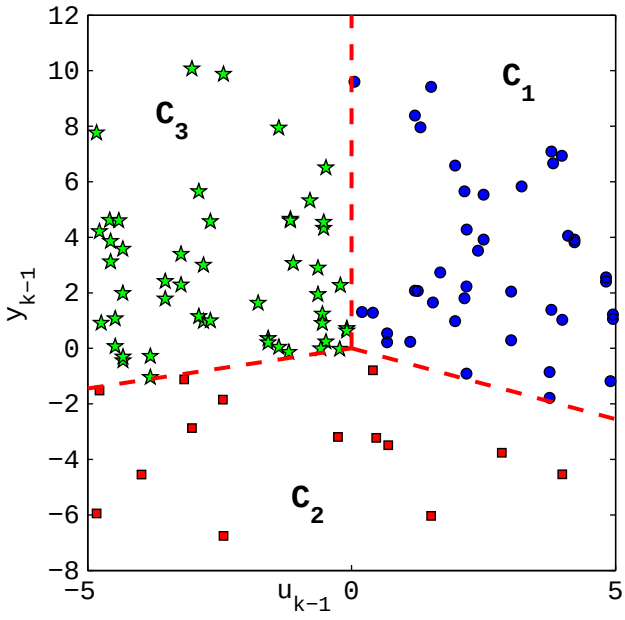


Fig. 5.15: Classes $\mathcal{C}_1$, $\mathcal{C}_2$ et $\mathcal{C}_3$

d'un domaine de solution sous la forme d'un zonotope, ce qui est équivalent à la recherche des paramètres h_1, h_2 et h_3 du mécanisme de commutation sous une forme intervalle. Pour cela, on résout de façon itérative le problème d'optimisation sous contraintes (4.13). Les résultats sont présentés dans la table 5.3.

Dans la table 5.3, les centres des intervalles trouvés sont indiqués par $h_{.0}$ et les

h_1	h_{1_0}	r_{h_1}
1	1.013	0.526
0.51	0.694	0.215
0	0.001	0.000

h_2	h_{2_0}	r_{h_2}
0	0.013	0.180
1	1.052	0.381
0	0.011	0.021

h_3	h_{3_0}	r_{h_3}
1	1.112	0.281
-0.29	-0.326	0.197
0	0.007	0.001

Tab. 5.3: Recherche des paramètres bornés

rayons sont notés $r_{h_.}$.

Afin de valider les paramètres identifiés, on a retenu le centre des intervalles trouvés, h_{1_0}, h_{2_0} et h_{3_0}, comme valeur des paramètres du mécanisme de commutation. Le premier graphique de la figure 5.16 montre la sortie du système obtenue à partir des paramètres réels du mécanisme de commutation et la sortie obtenue en utilisant le centre des intervalles trouvés. On peut noter que les deux courbes sont pratiquement superposées. Le second et le troisième graphique de la figure 5.16 montrent l'évolution de la séquence des modes du système pour les paramètres réels du mécanisme de commutation et pour les valeurs de paramètres correspondant aux centres des intervalles obtenus. Dans les deux cas, l'évolution de la séquence des modes est identique. Il faut toutefois garder en esprit que comme dans tout problème d'identification, les résultats obtenus sont fonction de la capacité du jeu de données disponible à représenter les divers modes de fonctionnement du système.

5.2.2 Recherche d'une solution particulière

On cherche maintenant à déterminer pour le même modèle qu'en (5.2) une valeur unique pour les paramètres h_1, h_2 et h_3 du mécanisme de commutation. Le nombre de points du jeux de données de travail n'étant pas très élevée, on met en œuvre l'algorithme résultant de la résolution du problème d'optimisation sous contraintes (4.28) exposé à la page 145 § 4.3.2. Toutes les classes sont ainsi simultanément prises en compte. La figure 5.17 montre le résultat obtenu en résolvant le problème d'optimisation sous contraintes (4.28). Sur cette figure, on peut voir la superposition du partitionnement réel de l'espace des régresseurs (en traits discontinus) et le partitionnement obtenu à partir des paramètres estimés (en traits continus). Les valeurs estimées $\hat{h}_1$, $\hat{h}_2$ et $\hat{h}_3$ des paramètres h_1, h_2 et h_3 sont pré-

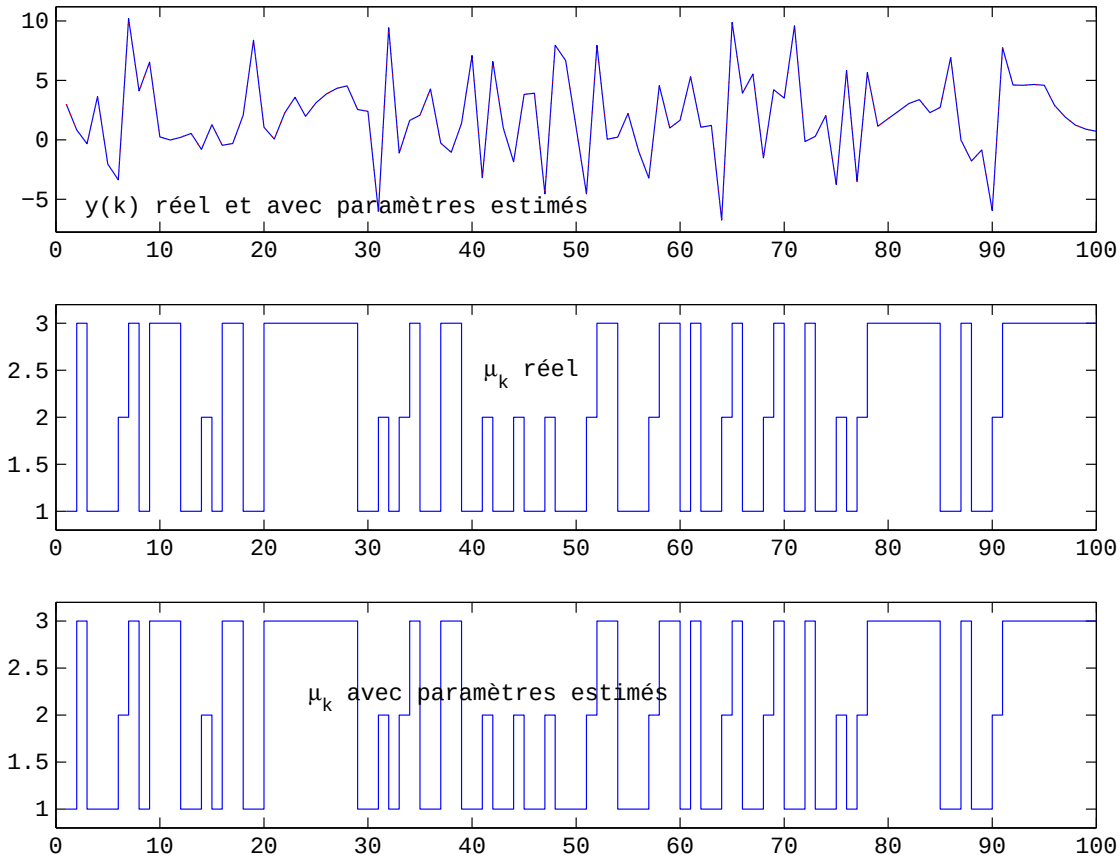


Fig. 5.16: Validation des paramètres identifiés

sentées dans la table 5.4. Les valeurs présentées dans les troisièmes colonnes sont obtenues après normalisation.

h_1	$\hat{h}_1$	$\hat{h}_1$
1	-17.893	1
0.51	-13.382	0.747
0	0.037	-0.002

h_2	$\hat{h}_2$	$\hat{h}_2$
0	0.000	0.000
1	-17.771	1
0	0.048	0.000

h_3	$\hat{h}_3$	$\hat{h}_3$
1	17.771	1
-0.29	-4.395	-0.246
0	0.012	0.000

Tab. 5.4: Recherche d'une solution particulière

Conclusion

Dans ce chapitre, les différentes méthodes proposées dans les chapitres 3 et 4 ont été mises en œuvre sur des exemples académiques. Il s'agit des méthodes de reconnaissance du chemin actif et d'identification du mécanisme de commutation. On a ainsi pu vérifier la faisabilité de ces méthodes, ceci tout en soulevant les

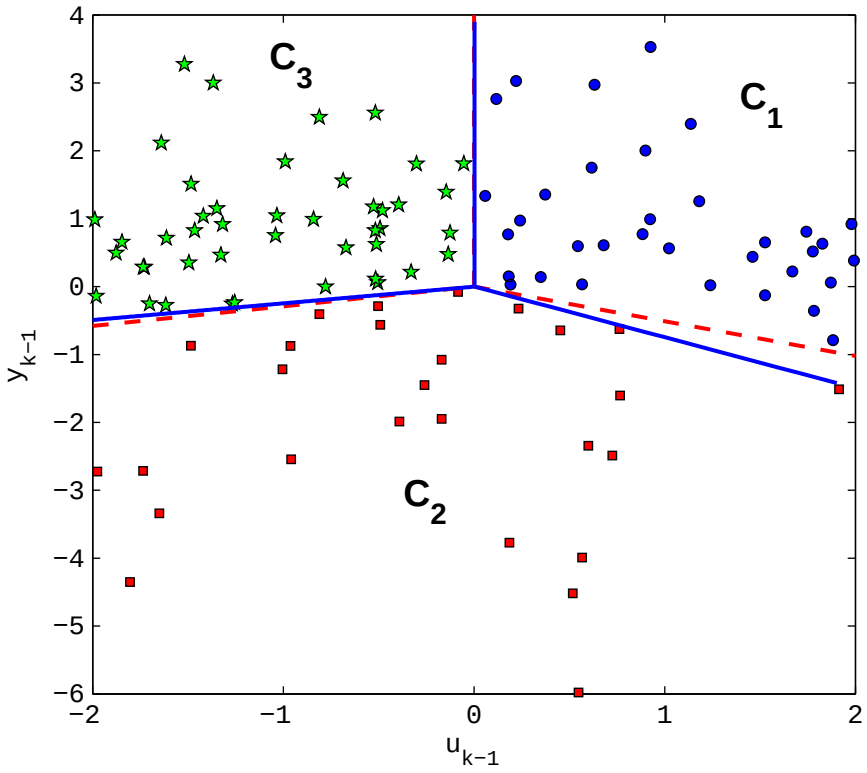


Fig. 5.17: Recherche d'une solution particulière

difficultés de mise en œuvre pouvant exister. Il serait intéressant de procéder par la suite à l'implémentation des approches développées sur un cas pratique, à savoir celui d'un système réel.

Conclusion générale et perspectives

Les systèmes à commutation constituent une classe particulière de systèmes hybrides. Dans cette thèse, nous avons abordé les problèmes liés à la reconnaissance de l'état discret ou du mode actif à chaque instant d'un système à commutation. Il s'agit de procéder à la reconnaissance du mode de fonctionnement actif chaque instant à partir de jeux de mesures et d'identifier ensuite les paramètres du mécanisme de gestion des transitions d'un mode de fonctionnement à un autre, ceci sous certaines hypothèses quant à la structure de ce mécanisme.

Les chapitres 1 et 2 constituent des chapitres introductifs qui donnent une vue d'ensemble sur les systèmes à commutation et le diagnostic, tout en essayant d'établir des passerelles entres ces deux thématiques. Le premier chapitre présente les modèles hybrides dans leur globalité, en mettant l'accent sur les modèles à commutation. Le second chapitre constitue une présentation sommaire du diagnostic à base de modèle et de son application aux modèles à commutation. Les notions d'observabilité pour les modèles à commutation y sont présentées.

Les troisième et quatrième chapitres portent sur la reconnaissance du mode actif et l'identification du mécanisme de commutation. Le problème posé dans le chapitre 3 est celui de la détermination à chaque instant du mode dans lequel se trouve le système, ce qui implique directement la détermination des instants de commutation ou de passage d'un mode de fonctionnement à un autre. Le modèle du système est connu, c'est-à-dire que les modèles linéaires décrivant les différents modes de fonctionnement sont connus. On dispose également de mesures de l'entrée et de la sortie du système. L'approche adoptée pour résoudre le problème consiste à décomposer l'horizon de mesures disponibles en des périodes de taille plus petite et effectuer

ensuite une recherche locale du mode de fonctionnement actif à chaque instant en s'inspirant des méthodes du diagnostic à base de modèle. On analyse ensuite des résidus générés sur une fenêtre d'observation du système, à partir des différentes séquences de commutation, afin de déterminer la séquence de commutation active et, du coup, le mode actif à chaque instant. La reconnaissance du mode actif étant effectuée en se servant des signaux d'entrée et de sortie, on peut se poser la question de savoir les conditions sous lesquelles des séquences de commutation différentes peuvent conduire le système à avoir le même comportement entrée/sortie, entraînant ainsi une impossibilité à retrouver le mode dans lequel se trouve le système. Ceci pose le problème de la discernabilité des séquences de commutation. La discernabilité des séquences de commutation est étudiée en cherchant à caractériser les situations pouvant conduire à l'obtention de résidus identiques (de valeur nulle) pour des séquences de commutation différentes. On obtient ainsi des conditions de discernabilité des séquences de commutation relatives à la structure du système et à la nature de l'entrée appliquée au système. Il faut noter qu'une fois la reconnaissance du mode actif effectuée, on peut procéder à l'estimation de l'état du système à commutation. Pour cela, et afin de conserver une cohérence dans la démarche proposée, on peut mettre en œuvre un observateur à mémoire finie.

Dans le chapitre 4, le mécanisme de commutation est considéré comme un processus dépendant des signaux d'entrée et de sortie du système. Dans ce cadre, on suppose que la loi de commutation du système est obtenue en partitionnant un espace associé à des grandeurs connues du système et en associant ensuite un mode de fonctionnement à chacun des domaines obtenus. L'ensemble des grandeurs connues est restreint ici à l'entrée et à la sortie du système comme dans le cas des modèles PWARX (*PieceWise Affine autoRegressive eXogenous* model). La structure du mécanisme de commutation étant fixée, il faut procéder à son identification, c'est-à-dire à l'estimation de ses paramètres. Ce problème est abordée avec deux approches différentes. La première approche met en œuvre une vision ensembliste du problème d'estimation paramétrique en recherchant une expression intervalle des paramètres de la loi de commutation. La seconde approche recherche les valeurs optimales des paramètres de la loi de commutation par rapport à des critères de distance en utilisant des techniques issues du domaine de la reconnaissance de

forme. Le chapitre 5 illustre la mise en œuvre des méthodes proposées à travers un exemple académique.

Certains points techniques restent à analyser dans les travaux futurs. Un point à développer est la taille de l'horizon d'observation assurant une meilleure discernabilité des séquences de commutation. Des simulations numériques ont montré l'influence de la taille de cet horizon sur la qualité de la reconnaissance des différents chemins. Toutefois, une expression analytique de cette taille n'a pu être exhibée. La prise en compte d'incertitude au sein des modèles décrivant les divers modes du système doit être approfondie. Un autre point à approfondir est le fonctionnement du système en mode non supervisé. On suppose, dans ce cas, que l'on ne dispose pas d'une connaissance complète de tous les régimes de fonctionnement du système. Il s'agit alors de procéder, au cours du fonctionnement du système, à l'identification simultanée des modes de fonctionnement non répertoriés, c'est-à-dire des modèles associés à ces modes. En ce qui concerne l'identification du mécanisme de commutation, il faut noter que, les différentes méthodes mises en œuvre conduisent à des problèmes d'optimisation sous contraintes dans lesquels la norme euclidienne est souvent utilisée. Il faudrait pouvoir évaluer l'impact de l'utilisation d'autres normes sur le problème d'optimisation. De façon plus générale, le problème de l'identification des systèmes à commutation pourra être considéré ainsi que celui de l'estimation d'état. Il serait également intéressant d'étudier l'extension des méthodes proposées à d'autres classes de modèle hybride.

Bibliographie

Ackerson, G. A. et K. S. Fu (1970). On state estimation in switching environnements. *IEEE Transactions on Automatic Control* **15**(1), pp. 10–17.

Adrot, Olivier (2000). Diagnostic à base de modèles incertains utilisant l'analyse par intervalle : l'approche bornante. PhD thesis. Institut National Polytechnique de Lorraine. Nancy, France.

Alessandri, A. et P. Coletta (2001*a*). Design of Luenberger observers for a class of hybrid linear systems. In : *Proceedings of Hybrid Systems : Computation and Control*. Rome, Italy.

Alessandri, A. et P. Coletta (2001*b*). Switching observers for continuous-time and discrete-time linear systems. In : *Proceedings of the American Control Conference*. Arlington, USA. pp. 2516–2521.

Alessandri, A. et P. Coletta (2003). Design of observers for switched discrete-time linear systems. In : *Proceedings of the American Control Conference*. Denver, USA. pp. 2785–2790.

Babaali, M et M. Egerstedt (2004). *Hybrid Systems : Computation and Control*. Chap. Observability of switched linear systems, pp. 48–63. Lecture Notes in Computer Science. Springer-Verlag.

Babaali, M. et M. Egerstedt (2005). Asymptotic observers for discrete-time switched linear systems. In : *Proceedings of 16th IFAC World Congress*. Prague, The Czech Republic.

Balluchi, A., L. Benvenuti, M. D. Di Benedetto et A. L. Sangiovanni-Vincentelli (2001). A hybrid observer for the driveline dynamics. In : *Proceedings of the European Control Conference*. Porto, Portugal. pp. 618–623.

Balluchi, A., L. Benvenuti, M. D. Di Benedetto et A. L. Sangiovanni-Vincentelli (2002). Design of observers for hybrid systems. In : *Proceedings of Hybrid Systems : Computation and Control*. Berlin, Germany. pp. 76–89.

Bara, I. G., J. Daafouz, F. Kratz et C. Iung (2000). State estimation for a class of hybrid systems. In : *Proceedings of 4th International Conference on Automation of Mixed Processes : Hybrid Dynamic Systems*. Dortmund, Germany. pp. 313–316.

Basseville, M. et I.V. Nikiforov (1993). *Detection of abrupt changes - Theory and application*. Information and System Sciences Series. Prentice Hall, Englewood Cliffs. New Jersey, USA.

Batruni, R. (1991). A multilayer network with piecewise-linear structure and back-propagation learning. *IEEE Transactions on Neural Networks* **2**(3), pp. 395–403.

Beard, R. V. (1971). Failure accomodation in linear system trhough self-reorganisation. PhD thesis. Dept. of Aero. and Astro., MIT, Cambridge, USA.

Bemporad, A., A. Garulli, S. Paoletti et A. Vicino (2003). *Hybrid Systems : Computation and Control*. Chap. A greedy approach to identification to identification of piecewise affine models. Lecture Notes in Computer Science. Springer-Verlag.

Bemporad, A., D. Mignone et M. Morari (1999). Fault detection and state estimation for hybrid system. In : *Proceedings of the American Control Conference*. Chicago, USA. pp. 2471–2475.

Bemporad, A. et M. Morari (1999). Control of systems integrating logic, dynamics and constraints. *Automatica* **35**(3), pp. 407–427.

Bemporad, A., G. Ferrari-Trecate et M. Morari (2000). Observability and controllability of piecewise affine and hybrid systems. *IEEE Transactions on Automatic Control* **45**, pp. 1864–1876.

Ben-Haim, Y. (1980). An algorithm for failure location in complex network. *Nuclear Science and Engineering* **75**, pp. 191–199.

Bennett, K. P. et E. J. Bredensteiner (1993). *Geometry at work*. Chap. Geometry in learning. Mathematical Association of America Press. Washington DC, USA.

Bennett, K. P. et O. L. Mangasarian (1994). Multicategory discrimination via linear programming. *Optimization Methods and Software* **3**, pp. 27–39.

Billings, S. A. et W. S. F. Voon (1987). Piecewise linear identification of nonlinear systems. *International Journal of Control* **46**(1), pp. 215–235.

Biswas, P., P. Grieder, J. Löfberg et M. Morari (2005). A survey on stability analysis of discrete-time piecewise affine systems. In : *Proceedings of the* 16*th IFAC World Congress*. Prague, Czech Republic.

Blondel, V. et J. N. Tsitsiklis (1999). Complexity of stability and controllability of elementary hybrid systems. *Automatica* **35**, pp. 479–489.

Bonnans, J. F., J. Ch. Gilbert, C. Lemarchal et C. A. Sagastizbal (2002). *Numerical optimization : theoretical and practical aspects*.

Bredensteiner, E. J. et K. P. Bennett (1999). Multicategory classification by support vector machines. *Computational Optimization and Applications* **12**, pp. 53–79.

Breiman, L. (1993). Hinging hyperplanes for regression, classification and function approximation. *IEEE Transactions on Information Theory* **39**(3), pp. 999–1013.

Chan, K. S. et H. Tong (1986). On estimating thresholds in autoregressive models. *Journal of Time Series Analysis* **7**(3), pp. 179–190.

Chen, J., R. J. Patton et H. Y. Zhang (1995). Design of robust structured and directional residuals for fault isolation via unknow input observers. In : *Proceedings of the 3rd European Control Conference*. Rome, Italy. pp. 348–353.

Choi, C.-H. et J. Y. Choi (1994). Constructive neural networks with piecewise interpolation capabilities for function approximations. *IEEE Transactions on Neural Networks* **5**(6), pp. 936–944.

Chow, E.Y. (1980). A failure detection system design methodology. PhD thesis. Department of Electrical Engineering and Computer Science, Massachusetts Institute of Technology. Cambridge, Massachusetts.

Chow, E.Y. et A.S. Willsky (1984). Analytical redundancy and the design of robust failure detection system. *IEEE Transactions on Automatic Control* **AC-29**(7), pp. 603–614.

Chua, L. O. et S. M. Kang (1977). Section-Wise piecewise-linear functions : Canonical representation, properties and applications. In : *Proceedings of the IEEE*. Vol. 65. pp. 915–929.

Cocquempot, V., M. Staroswiecki et T. El Mezyani (2003). Switching time estimation and fault detection for hybrid systems using structured parity residuals. In : *Proceedings of the 5th IFAC Symposium on Fault Detection, Supervision andSafety of Technical Processes*. Washington DC, USA. pp. 681–686.

De Schutter, B. (2000). Optimal control of a class of linear hybrid system with saturation. *SIAM Journal on Control and Optimization* **39**(3), pp. 835–851.

De Schutter, B. et T. J. J. Van Den Boom (2001). Model predictive control for max-min-plus-scaling systems. In : *Proceedings of the American Control Conference*. Arlington, VA. pp. 319–324.

Doucet, A, A. Logothetis et V. Krishnamurthy (2001). Particle Filters for State Estimation of Jump Markov Linear Systems. *IEEE Transactions on Signal Processing* **49**(3), pp. 613–624.

Ernst, S. (1998). Hinging hyperplanes trees for approximation and identification. In : *Proceedings of the 37th IEEE Conference on Decision and Control*. Vol. 2. Tampa, FL. pp. 1266–1271.

Evans, J. S. et R. J. Evans (1999). Image-enhanced multiple model tracking. *Automatica* **35**(11), pp. 1769–1786.

Fair, R. C. et D. M. Jafeee (1972). Methods of estimation for markets in disequilibrium. *Econometrica* **40**, pp. 299–308.

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

Ferrari-Trecate, G., M. Muselli, D. Liberati et M. Morari (2003). A clustering technique for the identification of piecewise affine systems. *Automatica* **39**(2), pp. 205–217.

Filev, D. (1991). Modelling of complex systems. *International Journal of Approximate Reasoning* **5**(5), pp. 281–290.

Gad, E. F., A. F. Atiya, S. Shaheen et A. El-Dessouki (2000). A new algorithm for learning in piecewise piecewise-linear neural networks. *Neural Networks* **13**(4-5), pp. 485–505.

Gasso, K., G. Mourot et J. Ragot (2002). Structure identification of multiple models with output error local models. In : *Proceedings of 15th IFAC World Congress on Automatic Control*. Barcelona, Spain.

Gertler, J. (1998). *Fault detection and diagnosis in engineering systems*. Marcel Dekker. New York, USA.

Gertler, J. J. (1992). Structured residuals for fault isolation, disturbance decoupling and modeling error robustness. In : *Proceedings of the IFAC symposium on-line Fault Detection and Supervision in the Chemical Processes Industries*. Newark, USA. pp. 111–119.

Gertler, J. J. (1995). Optimal residual decoupling for robust fault diagnosis. *International Journal of Control* **61**(2), pp. 395–421.

Gertler, J. J. et R. Monajemy (1993). Generating directional residuals by dynamic parity equations. In : *Proceedings of the 12th IFAC World Congress*. Vol. 5. Sidney, Australia. pp. 931–940.

Goldfeld, S. M. et R.E. Quandt (1973). A markov model for switching regressions. *Journal of Econometrics* **1**, pp. 3–16.

Goshen-Meskin, D. et I. Y. Bar-Itzhack (1992). Observability analysis of piecewise constant systems – part I : Theory. *IEEE Transactions on Aerospace & Electronic Systems* **28**(4), pp. 1056–1067.

Granger, C. W. J. et R. Ramanathan (1984). Improved methods of combining forecast. *Journal of Forecasting* **3**, pp. 197–204.

Granger, C. W. J. et T. Terasvirta (1993). *Modelling nonlinear economic relationships*. Oxford University Press. Oxford.

Grieder, P. et M. Morari (2003). Complexity reduction of receding horizon control. In : *Proceedings of the 42th IEEE Conference on Decision and Control*. Hawaii, USA.

Grieder, P., M. Kvasnica, M. Baoticét M. Morari (2004). Low complexity control of piecewise affine systems with stability guarantee. In : *Proceedings of the American Control Conference*. Boston, USA.

Grieder, P., P. Parillo et M. Morari (2003). Robust horizon receding control - analyis and synthesis. In : *Proceedings of the 42th IEEE Conference on Decison and Control*. Hawaii, USA.

Hamelin, F., D. Sauter et M. Aubrun (1994). Fault diagnosis in systems using directional residuals. In : *Proceedings of the 33rd IEEE Conference on Decison and Control*. Vol. 3. Buena Vista Palace, USA. pp. 3040–3045.

Hamilton, J. D. (1989). A new approach to the economic analysis of nonstationary time series and the business cycle. *Econometrica* **57**, pp. 357–384.

Hamilton, J. D. (1990). Analysis of time series subject to changes in regime. *Journal of Econometrics* **45**, pp. 39–70.

Heemels, W. P. M. H., J. M. Schumacher et S. Weiland (2000). Linear complementary systems. *SIAM Journal of Applied Mathematics* **60**(4), pp. 1234–1269.

Heredia, E. A. et G. R. Arce (1996). Piecewise linear system modeling based on a continuous threshold decomposition. *IEEE Transactions on Signal Processing* **44**(6), pp. 1440–1453.

Hudson, D. J. (1966). Finding segmented curves whose join points have to be estimated. *Journal of the American Statistical Association* **61**, pp. 1097–1129.

Hush, D. R. et B. Horne (1998). Efficient algorithms for function approximation with piecewise linear sigmoidal networks. *IEEE Transactions on Signal Processing* **9**(6), pp. 1129–1141.

Hwang, I., H. Balakrishnan et C. Tomlin (2003). Observability criteria and estimator design for stochastic linear hybrid systems. In : *Proceedings of the IEE European Control Conference*. Cambridge, UK.

Isermann, R. (1992). Estimation of physical parameters for dynamic processes with application to an industrial robot. *International Journal of Control* **55**, pp. 1287–1298.

Jazwinski, A. H. (1970). *Stochastic processes and filtering theory*. Academic Press. New York, USA.

Ji, Y. et H. Chizeck (1988). Controllability, observability and discrete-time jump linear quadratic control. *International Journal of Control* **48**(2), pp. 481–498.

Johansen, T. A. et A. B. Foss (1993). Constructing NARMAX using ARMAX. *International Journal of Control* **58**(5), pp. 1125–1153.

Julián, P., A. Desages et O. Amgamennoni (1999). High level canonical piecewise linear representation using a simplicial partition. *IEEE Transactions on Circuits and Systems — I : Fundamental Theory and Applications* **46**(4), pp. 463–480.

Julián, P., M. Jordan et A. Desages (1998). Canonical piecewise-linear approximation of smooth functions. *IEEE Transactions on Circuits and Systems — I : Fundamental Theory and Applications* **45**(5), pp. 567–571.

Juloski, A., M. Heemels et G. Ferrari-Trecate (2004). Data Based Hybrid Modeling of the Component Placement Process in Pick-and-Place Machines. *Control Engineering Practice* **12**, pp. 1241–1252.

Juloski, A., M. Heemels, G. Ferrari-Trecate, R. Vidal, S. Paoletti et H. Niessen (2005). *Hybrid Systems : Computation and Control*. Chap. Comparison of Four

Procedures for Identification of Hybrid Systems. Vol. 3141 of *Lecture Notes on Computer Science*. Springer-Verlag. Zurich, Switzerland.

Juloski, A., M. Heemels, Y. Boers et F. Veschure (2003). Two approaches to state estimation for a class of piecewise affine systems. In : *Proceedings of the 42nd IEEE Conference on Decision and Control*. Maui, Hawaii. pp. 143–148.

Kang, M. A. et L. O. Chua (1978). A global representation of multidimensional piecewise-linear functions. *IEEE Transactions on Circuits and Systems* **25**, pp. 938–940.

Kratz, F. et D. Aubry (2003). Finite memory observer for state estimation of hybrid system. In : *Proceedings of the 5th IFAC Symposium on Fault Detection, Supervision and Safety of Technical Processes*. Washington DC, USA. pp. 687–691.

Krishnamurthy, V. et R. J. Elliott (1997). Filters for estimating Markov modulated poisson processes and image based tracking. *Automatica* **33**(5), pp. 821–833.

Krishnamurthy, V., F. Vazquez Abad et K. Martin (1999). Adaptive nonlinear filters for narrowband interference suppression in spread spectrum CDMA systems. *IEEE Transactions on Communications* **47**(5), pp. 742–753.

Lagarias, J. C. et Y. Wang (1995). The finiteness conjecture for the generalized spectral radius of a set of matrices. *Linear Algebra and its Applications* **214**, pp. 17–42.

Liberzon, D. et A. S. Morse (1999). Basic problems in stability and design of switched systems. *IEEE Conrol Systems Magazine* **37**, pp. 59–70.

Liberzon, D., J. P. Hespanha et A. S. Morse (1999). Stability of switched systems : a lie-algebric condition. *Systems & Control Letters* **37**, pp. 117–122.

Mangasarian, O. L. (1999). Arbitrary-norm separating plane. *Operations Research Letters* **24**(1-2), pp. 15–23.

Maquin, D. (2005). *Cours de Master Automatique, Diagnostic, Signal, Bio-Imagerie : Surveillance des Processus*. Institut National Polytechnique de Lorraine, Centre de Recherche en Automatique de Nancy. Vandoeuvre-lès-Nancy, France.

Mariton, M. (1986). Stochastic observability of linear systems with markovian jump. In : *Proceedings of the 25th IEEE Conference on Decison and Control*. Athens, Greece.

Medeiros, M. C., A. Veiga et M. G. C. Resende (2002). A combinatorial approach to piecewise linear time series analysis. *Journal of Computational and graphical Statistics* **11**(1), pp. 236–258.

Münz, E. et V. Krebs (2002). Identification of hybrid systems using a priori knowledge. In : *Proceedings of the 15th IFAC World Congress*. Barcelona, Spain.

Murray-Smith, R. et T. Johansen (1997). *Multiple model approaches to modelling and control*. Taylor & Francis. London, United Kingdom.

Paoletti, S. (2004). Identification of Piecewise Affine Models. PhD thesis. Department of Information Engineering, University of Siena. Siena, Italy.

Patton, R. J. (1994). Robust model-based fault detection : the state of art. In : *Proceedings of IFAC Symposium on Fault Detection, Supervision and Safety for-Technical Processes, SAFEPROCESS'94*. Helsinki, Finland. pp. 359–367.

Patton, R. J., Frank, P. M. et Clark, R. N., Eds. (2000). *Issues of Fault Diagnosis for Dynamic Systems*. Springer-Verlag. London, Great Britain.

Patton, R.J., P.M. Frank et R.N. Clark (1989). *Fault diagnosis in dynamic systems : theory and application*. International Series in Systems and Control Engineering. Prentice Hall, Englewood Cliffs. New Jersey, USA.

Potter, J. E. et M. C. Suman (1977). Thresholdless redundancy management with array of skewed instruments. Technical report. Integrity in Electronic Flight Control Systems, AGARDOGRAPH-224.

Potter, S. M. (1995). A nonlinear approach to US GNP. *Journal of Applied Econometrics* **10**, pp. 109–125.

Pucar, P. et J. Sjöberg (1998). On the hinge-finding algorithm for hinging hyperplanes. *IEEE Transactions on Information Theory* **44**(3), pp. 1310–1319.

Quandt, R. E. (1958). The estimation of the parameters of a linear regression obeying two separate regime. *Journal of the American Statistical Association* **53**, pp. 873–880.

Ragot, J. (2005). *Cours de Master Automatique, Diagnostic, Signal, Bio-Imagerie : Estimation d'état. Application au diagnostic*. Institut National Polytechnique de Lorraine, Centre de Recherche en Automatique de Nancy. Vandoeuvre-lès-Nancy, France.

Ragot, J., D. Maquin et E. A. Domlan (2003*a*). Switching time estimation of piecewise linear system. Application to diagnosis. In : *Proceedings of the 5th IFAC Symposium on Fault Detection, Supervision and Safety of Technical Processes*. Washington DC, USA. pp. 669–704.

Ragot, J., G. Mourot et D. Maquin (2003*b*). Parameter estimation of switching piecewise linear systems. In : *Proceedings of the 42nd Conference on Decision and Control*. Maui, Hawaii. pp. 5783–5788.

Roll, J. (2003). Local and Piecewise affine Approaches to System Identification. PhD thesis. Departement of Electrical engineering, Linköping University. Linköping, Sweden.

Rosenqvist, F. et A. Karlström (2005). Realisation and Estimation of Piecewise-Linear Output-Error Models. *Automatica* **41**(3), pp. 545–551.

Simani, S., C. Fantuzzi et R. J. Patton (2003). *Model-based Fault Diagnosis in Dynamic Systems Using Identification Techniques*. Springer-Verlag. London, Great Britain.

Skeppstedt, A., L. Ljung, A. Benveniste, B. Deylon, P. Glorennec, H. Hjalmarsson et A. Juditsky (1992). Construction of composite models from observed data. *International Journal of Control* **55**(1), pp. 141–152.

Sontag, E. D. (1981). Nonlinear regulation : the piecewise linear approach. *IEEE Transactions on Automatic Control* **26**(2), pp. 346–358.

Sontag, E. D. (1996). *Hybrid Systems III – Verification and Control*. Chap. Interconnected automata and linear systems : a theoretical framework in discrete-time, pp. 436–448. Vol. 1066 of *Lecture Notes in Computer Science*. Springer-Verlag.

Strömberg, J.-E., F. Gustafsson et L. Ljung (1991). Trees as black-box model structures for dynamical systems. In : *Proceedings of the European Control Conference*. Grenoble, France. pp. 1175–1180.

Tong, H. (1983). *Threshold models in non-linear time series analysis*. Springer-Verlag. New York.

Tugnait, J. K. (1982). Detection and estimation for abruptly changing systems. *Automatica* **18**(5), pp. 607–615.

Van der Schaft, A. J. et J. M. Schumacher (1998). Complementary modelling of hybrid systems. *IEEE Transactions on Automatic Control* **43**(4), pp. 483–490.

Vapnik, V. N. (1995). *The nature of statistical learning theory*. Springer-Verlag.

Vidal, R., A. Chiuso et S. Soatto (2002). Observability and identifiabilty of jump linear systems. In : *Proceedings of the 41st IEEE Conference on Decision and Control*. Las Vegas, NV. pp. 3614–3619.

Vidal, R., S. Soatto, Y. Ma et S. Sastry (2003). An algebraic approach to the identification of a class of linear hybrid systems. In : *Proceedings of the 42nd IEEE Conference on Control and Decision*. Maui, Hawaii. pp. 167–172.

West, P. D. et A. H. Haddad (1994). On the observability of linear stochastic switching systems. In : *Proceedings of the 1994 American Control Conference*. Baltimore, MD. pp. 1846–1847.

Autorisation de soutenance

INSTITUT NATIONAL
POLYTECHNIQUE
DE LORRAINE

AUTORISATION DE SOUTENANCE DE THESE
DU DOCTORAT DE L'INSTITUT NATIONAL
POLYTECHNIQUE DE LORRAINE

o0o

VU LES RAPPORTS ETABLIS PAR :
Monsieur Vincent COCQUEMPOT, Maître de Conférences, LAGIS, Ecole Centre de Lille, Villeneuve d'Ascq

Monsieur Janan ZAYTOON, Professeur, CRESTIC, Université de Reims

Le Président de l'Institut National Polytechnique de Lorraine, autorise :

Monsieur DOMLAN Elom Ayih

NANCY BRABOIS
2, AVENUE DE LA
FORET-DE-HAYE
BOITE POSTALE 3
F - 54501
VANDŒUVRE CEDEX

à soutenir devant un jury de l'INSTITUT NATIONAL POLYTECHNIQUE DE LORRAINE, une thèse intitulée :

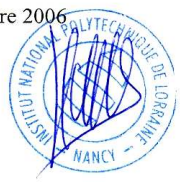
"Diagnostic des systèmes à changement de régime"

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 15 septembre 2006
Le Président de l'I.N.P.L.,
L. SCHUFFENECKER



TEL. 33/03.83.59.59.59
FAX. 33/03.83.59.59.55

Résumé : Les systèmes à commutation représentent une classe particulière de systèmes hybrides. Ils sont décrits par plusieurs modèles de fonctionnement et chaque modèle, définissant un mode du système, est actif sous certaines conditions opératoires particulières. Lorsque la loi de commutation régissant le passage d'un modèle de fonctionnement à l'autre est parfaitement connue, il est aisé de manipuler de tels systèmes car le mode actif peut être connu à chaque instant. Par contre, dans la situation où aucune information n'est disponible sur l'évolution de la loi de commutation, il est plus ardu de procéder au diagnostic ou encore de synthétiser une loi de commande sur ces systèmes. Il est abordé ici le problème de la reconnaissance du mode actif sur la base d'observations de l'entrée et de la sortie du système. L'identification des paramètres de la loi de commutation est ensuite étudiée sous l'hypothèse de la connaissance de la structure de la loi de commutation.

Mots-clés : système hybride, système à commutation, diagnostic de fonctionnement, discernabilité, analyse intervalle, identification, reconnaissance de forme.

Abstract : Switching systems are a particular class of hybrid systems. They are described by several operating regimes, each of them being active under certain particular conditions. When the switching mechanism is perfectly known, it is easy to handle such systems because the active regime can be known at every moment. On the other hand, in the case where no information is available on the switching mechanism, the situation is more complicated. In this situation, it is difficult to carry out a fault diagnosis scheme or to synthesize a control law. This thesis tackles the problem of the determination of the active mode on the basis of the system's input/output data. The issue of the identification of the switching mechanism's parameters is also considered assuming that its structure is known.

Keywords : hybrid system, switching system, diagnosis, discernability, interval analysis, identification, pattern recognition.