



HAL
open science

Méthodes d'aide à la décision pour la détection et la localisation de défauts dans les entraînements électriques

Christophe Combastel

► **To cite this version:**

Christophe Combastel. Méthodes d'aide à la décision pour la détection et la localisation de défauts dans les entraînements électriques. Automatique / Robotique. Institut National Polytechnique de Grenoble - INPG, 2000. Français. NNT: . tel-00198501

HAL Id: tel-00198501

<https://theses.hal.science/tel-00198501>

Submitted on 17 Dec 2007

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

INSTITUT NATIONAL POLYTECHNIQUE DE GRENOBLE

N° attribué par la bibliothèque

--	--	--	--	--	--	--	--	--	--

THESE

pour obtenir le grade de

DOCTEUR DE L'INPG

Spécialité : Automatique-Productique

préparée au Laboratoire d'Automatique de Grenoble et au Laboratoire d'Electrotechnique de Grenoble
dans le cadre de l'**Ecole Doctorale « Electronique, Electrotechnique, Automatique, Télécommunications, Signal »**

présentée et soutenue publiquement

par

Christophe COMBASTEL

le 27 septembre 2000

**Méthodes d'Aide à la Décision pour la Détection et la Localisation de Défauts
dans les Entraînements Electriques**

—————
Directeur de thèse :

Mme Sylviane GENTIL

—————
JURY

M. Christian ROMBAUT,	Président et rapporteur
M. Marcel STAROSWIECKI,	Rapporteur
Mme Sylviane GENTIL,	Directeur de thèse
M. Jean-Pierre ROGNON,	Co-encadrant
M. Vincent COCQUEMPOT,	Examineur
M. Jean-Luc THOMAS,	Examineur

Avant-Propos

Remerciements

Le travail présenté dans ce mémoire a été réalisé au LAG (Laboratoire d'Automatique de Grenoble) et au LEG (Laboratoire d'Electrotechnique de Grenoble), sous la direction de Madame Sylviane Gentil (LAG) et de Monsieur Jean-Pierre Rognon (LEG), professeurs à l'INPG (Institut National Polytechnique de Grenoble).

J'adresse tout d'abord mes remerciements à tous les membres du jury pour l'honneur qu'ils m'ont fait de participer à l'examen de ma thèse. Je remercie en particulier Monsieur Christian Rombaut, président et rapporteur ; je remercie également Monsieur Marcel Staroswiecki, rapporteur, Monsieur Vincent Cocquempot, examinateur, et Monsieur Jean-Luc Thomas, examinateur.

Je tiens à remercier tout particulièrement mes directeurs de thèse : Madame Sylviane Gentil et Monsieur Jean-Pierre Rognon, aussi bien pour leur aide précieuse que pour leurs qualités humaines. Malgré les nombreuses sollicitations, ils m'ont toujours consacré le temps nécessaire pour être guidé efficacement. Au-delà de leur grande rigueur scientifique et de l'aide technique qu'ils m'ont apportée, leur soutien amical et la reconnaissance dont ils ont fait preuve envers mon travail m'ont aidé à franchir bien des barrières.

Je voudrais aussi souligner l'importance de l'environnement dans lequel j'ai eu plaisir à travailler et remercier vivement Monsieur Bernard Descotes-Genon (mon tuteur dans le cadre du Monitorat), Madame Arlette Cheruy, Monsieur Hassane Alla, Madame Suzanne Lesecq, Madame Mireille Jacomino et Monsieur Daniel Baudois qui m'ont aidé, soutenu ou conseillé à diverses occasions. Je remercie également les secrétaires du laboratoire et l'équipe technique, autant pour leur soutien logistique que pour leur gentillesse.

Les échanges scientifiques ont également une place importante dans le cadre d'une thèse. Je tiens à remercier en particulier Stéphane Ploix : nos discussions m'ont réellement aidé à structurer ma vision du diagnostic. Mes collègues de bureau : Alexandre et Philippe, puis Monica et Siana, ont contribué à créer une ambiance de travail agréable et chaleureuse. J'ai beaucoup apprécié l'aide d'Alex et de Philou lors de mon intégration au sein de l'équipe. Comment ne pas avoir aussi une pensée pour les thésards des deux laboratoires : en particulier Alain, mais aussi Christian, Luc, Mariana, Bertrand, Stéphane...

Je remercie également Vincent, Ginette et Jean-Marie, qui n'ont pas été les derniers à me soutenir.

Enfin, je tiens à adresser une pensée affectueuse à toute ma famille, et en particulier à mes parents, pour leur soutien indéfectible durant toutes ces années d'études. Je suis heureux d'avoir pu lire de la joie et de la fierté dans leur regard un certain 27 septembre 2000. Je sais que ce même regard aurait illuminé le visage de mes grands-parents et de Marthou s'ils avaient pu être présents.

Table des matières

Table des matières

CHAPITRE I : INTRODUCTION	15
I.1. INTRODUCTION GÉNÉRALE	15
<i>I.1.1. Intérêts et objectifs du diagnostic</i>	<i>15</i>
<i>I.1.2. Présentation des entraînements électriques.....</i>	<i>16</i>
I.2. FORMULATION DU PROBLÈME DE DIAGNOSTIC	16
<i>I.2.1. Retour sur quelques définitions</i>	<i>16</i>
I.2.1.1. Supervision	16
I.2.1.2. Anomalie, défaut, défaillance, panne, perturbation	17
I.2.1.3. Diagnostic, détection, localisation, identification	18
I.2.1.4. Observation, mesure, signe, symptôme.....	19
I.2.1.5. Résumé	20
<i>I.2.2. La notion de modèle</i>	<i>20</i>
I.2.2.1. Proposition de définition pour le modèle d'un système	21
I.2.2.2. Simulation d'un modèle.....	22
I.2.2.3. Caractéristiques des modèles (comportementaux)	22
I.2.2.3.1. Granularité.....	22
I.2.2.3.2. (Im)précision et (In)certitude.....	22
I.2.2.3.3. Validité.....	23
I.2.2.3.4. Illustration de quelques caractéristiques d'un modèle	23
I.2.2.3.5. Causalité explicite ou non	24
I.2.2.3.6. Modèle statique ou dynamique.....	24
I.2.2.4. Quelques exemples de modèles	24
I.2.2.4.1. Modèle fonctionnel reposant sur un langage graphique	24
I.2.2.4.2. Modèles fondés sur un langage à la fois graphique et logico-mathématique	25
I.2.2.4.3. Modèle comportemental reposant sur un langage logico-mathématique, distinction entre connaissance a priori et connaissance apprise	26
<i>I.2.3. Validité, cohérence et élaboration d'un modèle</i>	<i>28</i>
<i>I.2.4. Validité, cohérence et diagnostic.....</i>	<i>28</i>
<i>I.2.5. Approches « bon fonctionnement » et « mauvais fonctionnement »</i>	<i>29</i>
I.2.5.1. Cas général	29
I.2.5.2. Approche orientée Bon Fonctionnement (BF)	30
I.2.5.3. Approche orientée Mauvais Fonctionnement (MF)	31
I.2.5.4. Bilan sur les approches BF et MF.....	31
<i>I.2.6. Décomposition du diagnostic</i>	<i>32</i>
I.2.6.1. Associations en parallèle et en série	32
I.2.6.2. Un exemple de diagnostic par redondance analytique	33
<i>I.2.7. Le diagnostic de défauts : découplage direct et indirect</i>	<i>34</i>
I.3. DIFFÉRENTES MÉTHODES POUR LE DIAGNOSTIC	37
<i>I.3.1. Redondance matérielle</i>	<i>37</i>
<i>I.3.2. Diagnostic de défaillances.....</i>	<i>37</i>
<i>I.3.3. Interfaces entre observations de nature différente</i>	<i>38</i>
I.3.3.1. Seuillage	38
I.3.3.2. Fuzzification/Défuzzification	38
<i>I.3.4. Traitement du signal.....</i>	<i>38</i>
I.3.4.1. Filtrage.....	38
I.3.4.2. Tests statistiques	38
<i>I.3.5. Génération de résidus.....</i>	<i>39</i>
I.3.5.1. Exploitation de l'existence de relations entre variables	39
I.3.5.2. Exploitation de la nature des relations entre variables	39
<i>I.3.6. Décision.....</i>	<i>40</i>
I.3.6.1. Classification par coalescence	40
I.3.6.2. Inférences à base de règles.....	41

1.3.7. Problèmes transversaux	42
1.3.7.1. Placement de capteurs	42
1.3.7.2. Prise en compte des incertitudes	42
1.3.7.3. Apprentissage.....	42
1.3.8. Bilan sur les différentes méthodes de diagnostic de défauts	43
I.4. CADRE DE L'ÉTUDE.....	43
CHAPITRE II : ANALYSE CAUSALE	47
II.1. INTRODUCTION	47
II.2. DÉFINITION D'UN MODÈLE « STRUCTUREL ».....	47
II.3. ORDONNANCEMENT CAUSAL : OBTENTION D'UN SCHEMA-BLOC.....	50
II.3.1. Objectif de l'ordonnancement causal	50
II.3.2. Quelques définitions et résultats : rappels et illustrations.....	51
II.3.3. Cas des systèmes dynamiques	53
II.3.4. Algorithme utilisé pour l'ordonnancement causal.....	55
II.3.5. Application à la Machine à Courant Continu (MCC).....	56
II.3.5.1. Avec forçage des équations différentielles	56
II.3.5.2. Sans forçage des équations différentielles	57
II.4. DÉCOMPOSITION DE MODÈLE.....	58
II.4.1. Introduction, définition des modèles locaux	58
II.4.2. Algorithme de construction des modèles locaux.....	59
II.4.3. Application à la MCC	61
II.5. GÉNÉRATION DE RÉSIDUS, ÉTUDE DE LEURS PROPRIÉTÉS	63
II.5.1. Génération des résidus locaux.....	63
II.5.2. Propriétés des résidus locaux	63
II.5.3. Résidus globaux	64
II.5.4. Propriétés des résidus globaux.....	64
II.5.5. Comparaison des propriétés des résidus globaux et locaux.....	65
II.5.6. Génération de résidus par composition de modèles locaux.....	66
II.5.7. Application à la MCC	67
II.5.7.1. Résidus globaux et locaux	67
II.5.7.2. Résidus associés à une composition de modèles locaux	69
II.6. APPLICATION À LA MAS	70
II.7. BILAN	74
II.7.1. Point de vue méthodologique.....	74
II.7.2. Point de vue des applications.....	74
CHAPITRE III : OBSERVATEURS, EQUATIONS DE PARITÉ.....	79
III.1. INTRODUCTION.....	79
III.2. OBJECTIF DE LA GÉNÉRATION DE RÉSIDUS	79
III.3. GÉNÉRATION DE RÉSIDUS POUR LES SYSTÈMES LTI	80
III.3.1. Cas du découplage parfait.....	82
III.3.2. Cas du découplage approximatif.....	83
III.3.3. Liens entre les différentes méthodes	84
III.4. SENSIBILITÉ ET CHOIX DE SEUILS ROBUSTES AUX INCERTITUDES	84
III.4.1. Etude de sensibilité : introduction.....	84
III.4.2. Influence des différentes sources d'incertitude sur un résidu.....	85
III.4.3. Choix des seuils	86
III.5. APPLICATION À LA MCC.....	87
III.5.1. Conditions de découplage parfait : comparaison avec l'analyse causale.....	87
III.5.2. Etude des pertes de rang dans le modèle de la MCC	89
III.5.3. Etude de sensibilité sur des résidus existants	92
III.6. LIENS ENTRE LES RÉSIDUS DE L'ANALYSE CAUSALE ET LES OBSERVATEURS À ENTRÉE INCONNUE	94

III.6.1. Préliminaires.....	94
III.6.1.1. Observateurs d'ordre plein.....	94
III.6.1.2. Observateurs d'ordre réduit.....	95
III.6.2. Représentation du système.....	97
III.6.3. Expression d'un résidu local comme un résidu de Luenberger.....	98
III.6.4. Intérêt du lien établi, perspectives.....	100
III.7. UNE MÉTHODE NUMÉRIQUE DE DÉCOUPLAGE APPROXIMATIF.....	101
III.7.1. La méthode utilisée.....	102
III.7.2. Application à la MCC.....	102
III.8. BILAN.....	105
III.8.1. Point de vue méthodologique.....	105
III.8.2. Point de vue des applications.....	106
CHAPITRE IV : TESTS DE MODÈLES.....	111
IV.1. INTRODUCTION.....	111
IV.2. RETOUR SUR LES PRINCIPES DU DIAGNOSTIC.....	111
IV.2.1. Modélisation des perturbations et des défauts.....	111
IV.2.2. Découplage indirect, interprétation d'un test de modèle.....	113
IV.2.3. Discussion sur le type de découplage.....	114
IV.3. DÉCOUPLAGE INDIRECT D'ENTRÉES ADDITIVES.....	116
IV.3.1. Un exemple de découplage indirect : utilisation du filtre de Kalman.....	116
IV.3.2. Application à la MCC.....	117
IV.3.3. Autres approches pour le découplage indirect.....	121
IV.3.3.1. Espace de parité : notion de direction associée à un défaut.....	121
IV.3.3.2. Le découplage indirect vu comme un problème de commande.....	122
IV.4. DÉCOUPLAGE INDIRECT DE PARAMÈTRES.....	123
IV.4.1. Un exemple de découplage indirect : estimation de paramètres par programmation non linéaire (PNL).....	123
IV.4.2. Condition d'existence du découplage indirect.....	126
IV.4.3. Application à la MAS : préliminaires.....	127
IV.4.3.1. Modèle de l'entraînement asynchrone (sans la commande).....	127
IV.4.3.2. Découplage direct par décomposition de modèle.....	128
IV.4.3.3. Représentation simplifiée : notation complexe (!).....	128
IV.4.4. Application à la MAS : détection robuste à l'échauffement.....	129
IV.4.4.1. Modèle.....	129
IV.4.4.2. Identifiabilité.....	130
IV.4.4.3. Sensibilité.....	130
IV.4.4.4. Découplage indirect par la méthode du gradient.....	132
IV.4.4.5. Interprétation locale du gradient, découplage direct du résidu adapté.....	136
IV.4.5. Application à la MAS : localisation d'un défaut sur le capteur de vitesse.....	139
IV.4.5.1. Modèle.....	139
IV.4.5.2. Identifiabilité.....	139
IV.4.5.3. Sensibilité.....	140
IV.4.5.4. Générateur de résidus.....	142
IV.5. BILAN.....	145
IV.5.1. Point de vue méthodologique.....	145
IV.5.2. Point de vue des applications.....	146
CHAPITRE V : DÉCISION STATIQUE.....	149
V.1. INTRODUCTION.....	149
V.2. SÉLECTION DES RÉSIDUS.....	151
V.2.1. Table de décision booléenne.....	151
V.2.2. Un petit exemple.....	151
V.2.3. Interprétation d'un dépassement de seuil.....	153
V.2.4. Optimisation de fonctions de décision par sélection des résidus.....	155
V.2.4.1. Paramétrage des fonctions de décisions.....	155

V.2.4.2. Evaluation des fonctions de décision.....	156
V.2.4.2.1. Critère de localisation	157
V.2.4.2.2. Critères de sensibilité	157
V.2.5. <i>Algorithme déduit de la formulation du problème</i>	158
V.2.5.1. Analyse.....	158
V.2.5.2. Algorithme de construction d'une fonction de décision	159
V.2.5.3. Application à la machine à courant continu (MCC)	159
V.2.6. <i>Optimisation par un algorithme génétique</i>	163
V.2.6.1. Objectifs et choix d'un algorithme	163
V.2.6.2. Principales caractéristiques d'un algorithme génétique.....	163
V.2.6.3. Application à l'optimisation des fonctions de décision	164
V.2.6.3.1. Codage des fonctions de décision	164
V.2.6.3.2. Choix du critère à optimiser	164
V.2.6.3.3. Résultats.....	164
V.3. ORDRES DE GRANDEUR ABSOLUS ET RELATIFS.....	166
V.3.1. <i>Augmentation de la richesse de l'information donnée par le test d'un résidu</i>	166
V.3.1.1. Quelques éléments de logique floue.....	166
V.3.1.2. Ordres de grandeur absolus qualitatifs : les signes	168
V.3.1.3. Généralisation du modèle « table de décision » et décision floue.....	169
V.3.1.4. Ordres de grandeur relatifs numériques : les directions.....	169
V.3.2. <i>Propriétés directionnelles des résidus</i>	171
V.3.2.1. Cas général, lien avec les tests de modèles basés sur le gradient.....	171
V.3.2.2. Cas des résidus du type « erreur de sortie » pour un modèle LTV	173
V.3.3. <i>Application à la MAS</i>	174
V.3.3.1. Propriétés directionnelles des résidus	174
V.3.3.2. Détection	175
V.3.3.3. Localisation des défauts à direction fixe.....	176
V.3.3.3.1. Défauts de l'onduleur	176
V.3.3.3.2. Défauts des capteurs de courant	178
V.3.3.3.3. Conditions de localisation et identification des défauts à direction fixe.....	178
V.3.3.3.4. Résultats.....	179
V.3.3.4. Localisation des défauts à directions variables.....	180
V.4. BILAN.....	183
V.4.1. <i>Point de vue méthodologique</i>	183
V.4.2. <i>Point de vue des applications</i>	184
CHAPITRE VI : DÉCISION DYNAMIQUE.....	189
VI.1. INTRODUCTION	189
VI.2. UN MODÈLE DYNAMIQUE POUR LA DÉCISION : LES SÉQUENCES FLOUES	190
VI.2.1. <i>Préliminaire : croissance et décroissance d'un signal</i>	190
VI.2.2. <i>Intérêt d'un modèle dynamique pour la décision</i>	192
VI.2.2.1. Exemple statique	192
VI.2.2.2. Première approche : mémorisation des dépassements de seuil.....	194
VI.2.3. <i>Modèle de l'action d'un défaut sur un résidu : séquences floues</i>	195
VI.2.3.1. Le langage des expressions régulières.....	196
VI.2.3.2. Calcul de l'appartenance d'un résidu à une séquence	197
VI.2.3.2.1. Concaténation de plusieurs ensembles flous	197
VI.2.3.2.2. Interprétation de la fin d'une séquence	198
VI.2.3.2.3. Une difficulté liée à la représentation floue	199
VI.2.3.2.4. Bilan sur le calcul de l'appartenance à une séquence.....	199
VI.2.4. <i>Application à la reconnaissance d'un comportement fugace</i>	200
VI.2.5. <i>Application à la MCC</i>	202
VI.2.6. <i>Obtention des tables de décision à partir d'un générateur de résidus</i>	203
VI.2.6.1. Formulation du problème	203
VI.2.6.2. Description qualitative de la réponse indicielle d'un transfert	203
VI.2.6.3. Application aux résidus globaux de la MCC.....	206
VI.3. DÉCISION : APPRENTISSAGE	207
VI.3.1. <i>Introduction</i>	207
VI.3.2. <i>Choix de la topologie du réseau de neurones</i>	207
VI.3.3. <i>Application à la MCC</i>	208

VI.4. BILAN.....	211
VI.4.1. Point de vue méthodologique.....	211
VI.4.2. Point de vue des applications.....	212
CHAPITRE VII : CONCLUSION.....	215
VII.1. BILAN.....	215
VII.1.1. Point de vue méthodologique.....	215
VII.1.2. Point de vue des applications.....	217
VII.2. PERSPECTIVES.....	218
VII.2.1. Point de vue des applications.....	218
VII.2.2. Point de vue méthodologique.....	219
VII.2.2.1. Modélisation.....	219
VII.2.2.2. Améliorer le compromis entre sensibilité et robustesse.....	220
VII.2.2.3. Utiliser au mieux l'information disponible.....	220
VII.2.2.4. Intégration du diagnostic à la supervision et à la maintenance.....	221
BIBLIOGRAPHIE.....	225
ANNEXE A : LEXIQUE.....	237
ANNEXE B : MACHINE À COURANT CONTINU (MCC).....	243
B.1. SCHÉMA SYNOPTIQUE DE L'ENTRAÎNEMENT CONTINU.....	243
B.2. PARAMÈTRES DE LA MCC.....	243
B.3. PARAMÈTRES POUR LES SIMULATIONS.....	244
ANNEXE C : MACHINE ASYNCHRONE (MAS).....	247
C.1. SCHÉMA SYNOPTIQUE DE L'ENTRAÎNEMENT ASYNCHRONE.....	247
C.2. PARAMÈTRES DE LA MAS.....	247
C.3. MODÈLE DE PARK D'UN MOTEUR ASYNCHRONE (SANS DÉFAUT).....	248
C.3.1. Hypothèses.....	248
C.3.2. Equations électriques et mécaniques.....	248
C.3.3. Transformation de Concordia (triphasé→biphasé) et rotation.....	250
C.3.4. Transformation de Park (rotor fictif équivalent fixe).....	251
C.3.5. Obtention d'une représentation d'état dans un repère (dq).....	252
C.4. PRINCIPE DE LA COMMANDE VECTORIELLE.....	255

Chapitre I :

Introduction

Chapitre I :

Introduction

I.1. Introduction générale

I.1.1. Intérêts et objectifs du diagnostic

L'amélioration de la qualité et la réduction des coûts sont l'objet d'une demande croissante de la part des industriels, en particulier pour l'exploitation de systèmes automatisés dont la complexité est croissante. Le concept de qualité n'est pas nouveau : ainsi, le contrôle à partir de méthodes statistiques se développe dans les années 30 aux USA. La gestion de la qualité commence dans les années 70, surtout au Japon, puis se généralise en Europe au cours des années 80. La qualité¹ d'un produit ou d'un service est son aptitude à satisfaire complètement les besoins et les attentes des utilisateurs. Les composantes de la qualité regroupent donc des domaines très variés (Figure I.1).

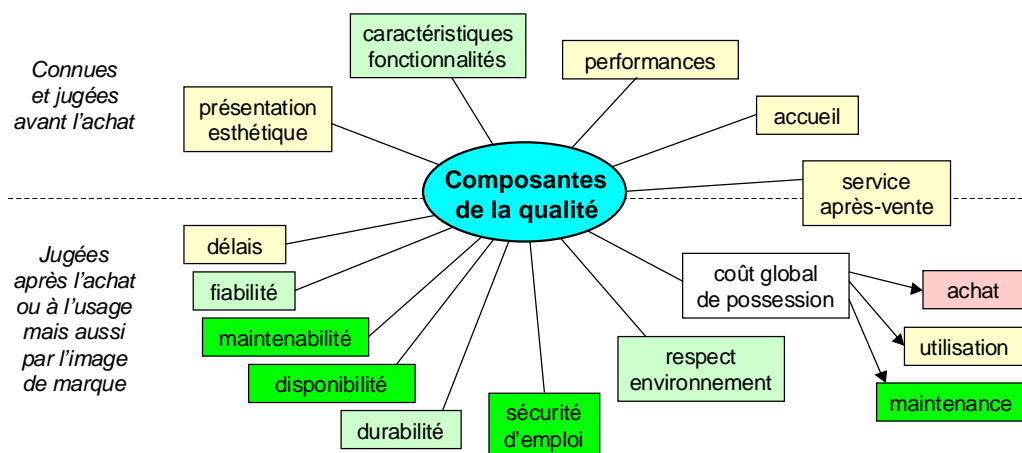


Figure I.1 : Composantes de la qualité d'un produit

Le diagnostic consiste à détecter, à localiser et éventuellement à identifier les défaillances et/ou les défauts qui affectent un système. Le diagnostic s'intègre dans le cadre plus général de la surveillance et de la supervision. Il permet d'améliorer la qualité et de réduire les coûts en intervenant au cours de certaines phases du cycle de vie d'un produit :

¹ Une définition est proposée pour les mots soulignés dans l'Annexe A : Lexique.

- **Conception** : une méthode d'analyse préventive (telle que l'AMDEC¹) peut être utilisée dès les premières étapes d'un projet pour déterminer au mieux les défaillances possibles ainsi que leurs effets : cela conduit à une sorte de diagnostic de la conception. Par ailleurs, la conception d'un système de diagnostic (si elle a lieu) intervient également à ce stade.
- **Production** : certains défauts peuvent être détectés et localisés en cours de fabrication. Le diagnostic du système de production et de ses composants (continus ou discrets) permet donc de corriger ou d'arrêter la fabrication de produits dont on sait qu'ils ne pourront plus satisfaire leur cahier des charges (et par conséquent des contrôles de qualité ultérieurs).
- **Utilisation** : une procédure d'arrêt peut être déclenchée si la sécurité est mise en péril par l'occurrence d'un défaut lors d'une phase d'utilisation. Une localisation précise des défaillances permet d'améliorer la maintenabilité et la disponibilité en indiquant les composants à remplacer. L'intérêt du diagnostic est aussi de fournir des informations permettant de définir une politique de maintenance appropriée (maintenance prédictive, par exemple).

1.1.2. Présentation des entraînements électriques

Les entraînements électriques à vitesse variable représentent actuellement une part importante des actionneurs industriels. Leur gamme de puissance étendue et leur souplesse d'utilisation liées aux progrès de l'électronique de puissance ont contribué à leur succès. Ils se retrouvent ainsi dans nombre de secteurs d'activité (robotique, machine outil, traction ferroviaire, ...).

Bien que n'appartenant pas forcément à des ensembles où la sûreté est aussi cruciale que dans l'aéronautique ou le nucléaire, mais à des procédés dont la continuité de service est un argument commercial ou un élément stratégique de la production, les entraînements électriques sont des systèmes dont la surveillance devient un enjeu majeur.

Souvent considérés comme robustes, ils subissent néanmoins le contrecoup de leur complexité croissante: le fonctionnement à vitesse variable des moteurs nécessite un contrôle de l'énergie par un convertisseur statique qui est très sollicité de par sa fonction; de plus, les entraînements à vitesse variable sont indissociables d'un dispositif de commande qui induit la présence d'éléments sensibles comme les chaînes d'acquisition de mesures et les interfaces de contrôle du convertisseur.

I.2. Formulation du problème de diagnostic

1.2.1. Retour sur quelques définitions

1.2.1.1. Supervision

La supervision est généralement définie comme une tâche de commande et de surveillance de l'exécution d'une opération ou d'un travail accompli par d'autres agents (hommes ou machines), sans rentrer dans les détails de cette exécution. Le concept de supervision s'applique donc dans un cadre hiérarchisé à deux niveaux au moins. Superviser le fonctionnement d'installations industrielles est devenu aujourd'hui le rôle principal des opérateurs humains, les systèmes

¹ AMDEC : Analyse des Modes de Défaillances et de leurs Effets Critiques.

automatiques assurant localement le contrôle et la commande. Des tâches multiples telles que la prédiction, le diagnostic, le pronostic ou le conseil d'action relèvent de la supervision [Travé-Massuyès, 1997]. Le diagnostic, par exemple, peut en fait se décomposer en plusieurs sous-tâches : détection d'une anomalie, localisation du sous-système défaillant et caractérisation d'un défaut ou d'une défaillance.

1.2.1.2. Anomalie, défaut, défaillance, panne, perturbation

Jusqu'ici, nous avons employé différents termes pour évoquer les anomalies pouvant survenir dans un système. Quelques définitions vont préciser le sens accordé à chacun de ces termes :

- **Anomalie** : particularité non conforme à la loi naturelle ou à la logique [Pruvost-Beaurain, 1990].
- **Défaut** : anomalie de comportement au sein d'un système physique [Ploix, 1998].
- **Défaillance** : anomalie fonctionnelle au sein d'un système physique [Ploix, 1998].
- **Panne** : inaptitude d'un dispositif à accomplir sa fonction.

Le terme « anomalie » est générique : il permet de décrire tout ce qui n'est pas « conforme » à une référence. Du point de vue de l'Automatique, une telle référence est définie à l'aide d'un modèle du système à diagnostiquer. Par exemple, un modèle du fonctionnement normal constitue une référence permettant de détecter des anomalies à partir d'observations.

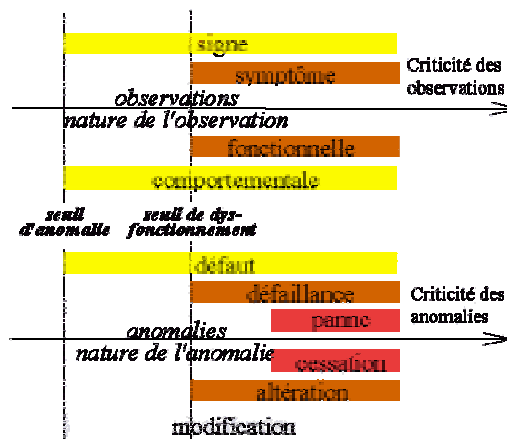


Figure I.2 : Anomalies et Observations classées par criticité croissante (d'après [Ploix, 1998])

Les termes « défaut » et « défaillance » diffèrent par la nature de l'anomalie. En effet, un défaut porte sur un comportement alors qu'une défaillance porte sur une fonction. Par fonction, nous entendons une action réalisée par un composant exprimée sous la forme d'un but à atteindre. Un défaut est une anomalie a priori moins critique qu'une défaillance, puisque certains comportements non conformes à une référence (défauts) peuvent très bien ne pas remettre en question le but à atteindre (fonction). Les défauts peuvent présager des défaillances à venir d'où l'intérêt de les diagnostiquer en vue de définir une politique de maintenance appropriée.

Il existe également une criticité croissante entre une « défaillance » et une « panne ». En effet, d'une non conformité (ou anomalie) dans le cas d'une défaillance, on passe à une inaptitude (à accomplir une fonction) dans le cas d'une panne. La Figure I.2 représente les anomalies classées par criticité croissante.

La définition retenue pour une perturbation est la suivante : écart (entre le comportement d'un système physique et une référence) dont l'origine n'est pas considérée comme une anomalie. Les bruits de mesures sont des exemples de perturbation. De plus, perturbations et défauts sont de même nature ; la différence entre les deux tient au caractère normal ou anormal du comportement du système physique. Le diagnostic devra être aussi robuste que possible vis-à-vis des perturbations et aussi sensible que possible aux défauts. La notion de « normalité » qui distingue défauts et perturbations est assez subjective : elle dépend de l'application et de la criticité des écarts de comportement. Le cas d'un moteur devant assurer une fonction de ventilation est donné à titre d'exemple dans la Table I.1.

<i>Définition illustrée</i>	<i>Événement (exemple)</i>	<i>Ecart au comportement nominal (courant, vitesse)</i>	<i>Aptitude à remplir la fonction de ventilation</i>
Perturbation	Léger échauffement (normal)	Petit	Totale
Défaut	Fort échauffement (anormal)	Moyen	Totale
Défaillance	Déclenchement intermittent d'un relais thermique stoppant le ventilateur jusqu'à ce que la température du moteur redescende à un niveau acceptable.	Grand	Partielle
Panne	Suite aux forts échauffements répétitifs, les isolants sont progressivement endommagés: un court-circuit apparaît; le moteur ne peut plus tourner jusqu'à ce qu'une réparation soit effectuée.	Grand	Nulle

Table I.1 : Illustration des définitions à l'aide d'un moteur de ventilateur.

Contrairement à ce que pourrait laisser penser la Table I.1, les écarts de comportement relatifs à des perturbations ne sont pas nécessairement plus faibles que ceux associés aux défauts. Par exemple, une perturbation telle qu'une variation du couple de charge sur un moteur peut entraîner des écarts très importants par rapport à un comportement de référence, sans que cette situation soit critique. Des écarts de comportement beaucoup plus faibles, résultant par exemple de courts-circuits entre quelques spires d'une même phase, correspondent par contre à une situation plus critique pour laquelle un diagnostic pourra être envisagé.

1.2.1.3. Diagnostic, détection, localisation, identification

L'objectif du diagnostic est de donner à un opérateur ou à un niveau de supervision supérieur des informations sur les anomalies qui se sont produites dans le système surveillé. Traditionnellement, on distingue trois tâches visant à répondre à trois questions portant sur les anomalies :

- **Détection** : Y a-t-il au moins une (nouvelle) anomalie dans le système ?
- **Localisation** : A quelles classes d'anomalies appartiennent les anomalies du système ? (Un ensemble minimal de classes d'anomalies « expliquant » les observations disponibles est en général privilégié).
- **Identification** : Quelles sont les caractéristiques de chacune des anomalies ?

Nous utilisons le terme générique d'« anomalie » de façon à garder dans un premier temps un cadre commun pour le diagnostic des défauts et pour le diagnostic des défaillances.

1.2.1.4. Observation, mesure, signe, symptôme

Pour arriver à déterminer et à caractériser les anomalies d'un système, il est nécessaire d'avoir des informations sur son comportement et/ou sur son fonctionnement. Ces informations peuvent être classées en deux catégories : les connaissances qui ne dépendent pas de l'environnement du système sont fournies par un modèle ; les connaissances qui dépendent de l'environnement du système (changement de consignes, perturbations,...) sont fournies par des observations.

- **Observation** : information (connue) sur le comportement réel ou sur le fonctionnement réel du système et qui dépend a priori de l'environnement de ce dernier.
- **Mesure** : observation représentée par une variable dont le contenu est l'image d'une grandeur physique. L'obtention d'une mesure se fait par l'intermédiaire d'un système physique : le capteur.
- « **Signe** » : caractère distinctif d'un état comportemental anormal [Ploix, 1998].
- **Symptôme** : caractère distinctif d'un état fonctionnel anormal [Ploix, 1998].

Dans l'acception que nous avons choisie, une mesure peut être vue comme une observation élémentaire du fait qu'elle reflète une (et une seule) grandeur physique. Dans le cas des capteurs dits « intelligents », une ou plusieurs mesures sont utilisées par une unité de traitement qui va extraire et fusionner des informations issues de ces mesures. Les informations de plus haut niveau résultant de ce traitement sont toujours des observations, mais ne sont plus nécessairement des mesures puisqu'elles peuvent ne pas être l'image directe des grandeurs physiques mesurées.

Un système de diagnostic peut être vu comme un capteur « intelligent » qui traite des observations (et en particulier les mesures) en vue de fournir des observations de plus haut niveau ciblées sur les anomalies. Par conséquent, le traitement réalisant le diagnostic consiste à extraire les « signes » et les symptômes (Figure I.2) les plus pertinents des observations disponibles (en exploitant la connaissance fournie par un modèle).

Les mesures peuvent être de différentes natures. En effet, les valeurs prises par une mesure appartiennent à un ensemble de modalités qui définissent la nature de cette mesure [Cassar, 1996] :

<i>Nature de la mesure</i>	<i>Ensemble des modalités</i>
Quantitatif	Intervalle défini sur l'ensemble des réels, \mathfrak{R} : la plage de la mesure.
Logique	{vrai, faux} ou {0, 1}
Qualitatif	Ensemble fini de modalités dont chacune est associée à un symbole.

Table I.2 : Exemples de mesures

Les mesures peuvent également être représentées à l'aide de la logique floue [Zadeh, 1965]. Un ensemble flou A dans X est défini comme suit:

$$A = \left\{ (x, \mu_A(x)) \mid x \in X \right\} \quad \text{avec} \quad \mu_A : X \rightarrow [0;1] \quad (\text{I.1})$$

μ_A est la fonction d'appartenance de x à l'ensemble flou A . Suivant l'allure des fonctions d'appartenance, il est possible de caractériser des informations de nature différente [Montmain, 1992], [Frank, 1994a]. Comme indiqué sur la Figure I.3, il est ainsi possible d'unifier les

différents types de mesure rappelés dans la Table I.2. L'intérêt de la logique floue est de permettre un raisonnement graduel.

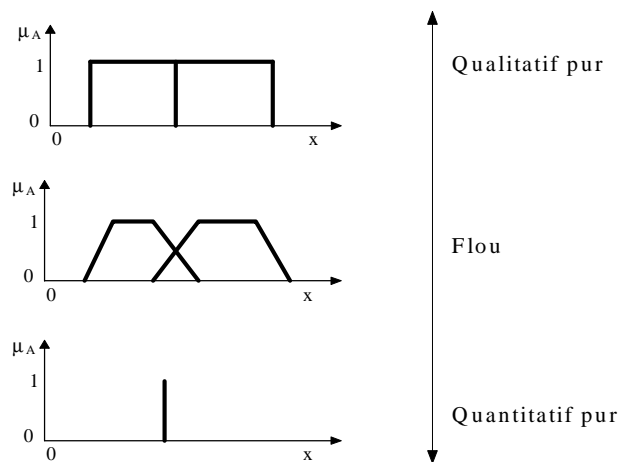


Figure I.3 : Fonctions d'appartenance correspondant à différents types de données

Quelle que soit la représentation choisie, le choix des capteurs et de la nature des observations influe très fortement sur les possibilités de surveillance d'un système.

I.2.1.5. Résumé

anomalies	= {défauts, défaillances, pannes}
environnement	= {consignes, anomalies, perturbations}
connaissance(système, environnement)	= {observations(système, environnement), modèles(système)}
diagnostic	= {détection, localisation, identification}
observations_anomalies	= diagnostic(connaissance(système, environnement))

Table I.3 : Liens entre les différents concepts

I.2.2. La notion de modèle

S'il est un point central dans le domaine de l'Automatique, c'est bien la notion de modèle. Paradoxalement, la notion même de modèle est finalement assez peu discutée dans la branche de la communauté automatique s'intéressant plus spécifiquement au diagnostic. Pourtant, le choix de modèles appropriés s'avère capital pour le diagnostic. En outre, les caractéristiques des modèles utilisés permettent de classer les différentes techniques qui, bien souvent, doivent être combinées afin de profiter de leurs avantages respectifs.

- **Système** : ensemble d'éléments en interaction [Von Bertalanffy, 1993].
- **Modèle** : (objet que l'on imite), reproduction d'un objet, représentation formalisée d'un phénomène [Pruvost-Beaurain, 1990].

La notion de système repose sur des éléments constitutifs, des objets « réels » et sur des interactions entre ces objets. La notion de modèle repose sur l'imitation et la reproduction d'une part, sur la formalisation et l'abstraction de la réalité étudiée d'autre part. Dès lors, il est possible de considérer le modèle d'un système comme un ensemble d'objets « virtuels » qui sont en interaction lorsque le modèle est simulé. Le terme « objet » est à prendre ici au sens informatique du terme, c'est-à-dire comme une entité contenant des données et les méthodes permettant de traiter ces données. Les données décrivent certaines caractéristiques de l'objet et les méthodes

décrivent le comportement de l'objet. Il peut s'agir aussi bien du comportement propre à l'objet que des interactions avec les autres objets.

Dans ce contexte, établir un modèle consiste dans un premier temps à se donner des objets virtuels qui seront autant de « briques » élémentaires utilisées pour la construction de ce modèle. Ce choix est guidé par la nature de la connaissance disponible sur le système et par l'objectif assigné au modèle. Deux types d'ontologies se dégagent en général : l'approche orientée « composants » consiste à décomposer la structure du système, et l'approche orientée « processus » (ou approche comportementale) effectue une décomposition phénoménologique [Travé-Massuyès, 1997]. Feliot propose d'y adjoindre également une approche fonctionnelle [Feliot, 1997]. Trois types de modélisation répertoriés dans la Table I.4 peuvent ainsi être distingués. Suivant celle qui est retenue, le diagnostic ne porte pas nécessairement sur les mêmes anomalies.

<i>Modélisation</i>	<i>Connaissances décrites par le modèle</i>	<i>Diagnostic des...</i>	<i>Forma- lisation</i>	<i>Validité</i>	<i>Ambi- guité</i>
Fonctionnelle	Ce que doit faire le système.	défaillances	faible	forte	forte
Structurelle	Comment le système est fait.	défaillances et défauts	moyenne	moyenne	moyenne
Comportementale	Comment le système fait ce qu'il doit.	défauts	forte	faible	faible

Table I.4 : Type de modélisation

Un modèle n'est qu'un reflet imparfait, simplifié, de la réalité qu'il est censé représenter. Il ne rend donc compte que d'une partie des phénomènes qui s'y déroulent (ceux justement qui sont jugés pertinents compte tenu des objectifs poursuivis). Il existe donc une infinité de modèles d'un même système, chacun ne remplissant que partiellement les fonctions requises (Table I.5). Une solution pour pallier les insuffisances de modèles simples consiste donc à les combiner entre eux en fonction de leurs complémentarités [Feliot, 1997] ¹.

Fonction cognitive	Représenter les connaissances que l'on a du système.
Fonction prévisionnelle	Prévoir l'évolution du système (simulation).
Fonction normative et décisionnelle	Représenter les relations souhaitables entre éléments du système. Cette fonction est nécessaire dans une démarche de conception et d'optimisation. Prendre des décisions d'action sur le système par rapport à des objectifs à atteindre.

Table I.5 : Fonctions d'un modèle

1.2.2.1. Proposition de définition pour le modèle d'un système

L'objectif est de proposer une définition pour un modèle qui soit la plus indépendante possible du type de modélisation. Pour cela, nous revenons sur l'idée selon laquelle un modèle serait construit à partir d'objets virtuels élémentaires au travers d'un exemple ; la description d'un

¹ La combinaison d'une étape de génération de résidus fondée sur un modèle numérique avec une étape de décision reposant sur un modèle logique (table de décision) est une illustration dans le contexte du diagnostic. Ce point sera développé dans la partie I.2.6.2.

système par un ensemble de phrases en français peut déjà être vue comme un modèle de ce système. Ainsi, le mot « moteur » est une représentation abstraite du système physique correspondant. A ce mot, nous associons naturellement un ensemble de fonctions et de comportements qui sont (il faut l'espérer) en rapport avec la réalité du système physique correspondant. En outre, un mot peut être considéré comme un objet du langage français. La richesse de la description d'un système par des phrases vient alors de la combinaison des mots qui obéit aux règles de la grammaire : « Un moteur est constitué d'un stator et d'un rotor » est une phrase dont la syntaxe est correcte ; elle donne de plus une information en rapport avec la réalité physique du moteur. Cette phrase modélise donc au moins partiellement le moteur. Parmi toutes les phrases grammaticalement correctes, seules certaines sont en rapport avec la réalité du système physique. « Un moteur est constitué d'une voiture » est une phrase correcte grammaticalement mais non cohérente avec la réalité physique d'un moteur : elle ne modélise donc pas correctement le moteur.

Pour résumer, nous nous sommes donné tout d'abord un ensemble O d'objets élémentaires (mots) et un langage L (français). La combinaison de ces objets élémentaires en obéissant aux règles du langage (grammaire) a permis d'obtenir une grande richesse descriptive qui s'est traduite par un ensemble d'affirmations possibles $L(O)$ (ensemble des phrases possibles). Le modèle d'un système est un objet formé d'un ensemble d'affirmations de $L(O)$ qui sont jugées en rapport avec la réalité du système à modéliser.

1.2.2.2. Simulation d'un modèle

C'est par le biais de la simulation que la fonction prévisionnelle d'un modèle peut être assurée (Table I.5). La simulation peut être vue comme une mise en interaction des objets virtuels constituant le modèle conformément à leurs comportements respectifs. Or, de même qu'un système réel est soumis aux contraintes de son environnement, la simulation du modèle de ce système dans un contexte réaliste suppose une certaine connaissance de cet environnement. Ainsi, les entrées du modèle constituent une interface entre le modèle et les observations (par exemple : consignes, mesures). De même, une simulation n'aurait pas de sens s'il n'était pas possible d'en récupérer les résultats, d'où l'intérêt de définir les sorties d'un modèle. Les entrées et les sorties font partie des objets élémentaires servant à la définition du modèle ; ce sont généralement des variables¹.

1.2.2.3. Caractéristiques des modèles (comportementaux)

1.2.2.3.1. Granularité

La granularité reflète le nombre de grandeurs utilisées pour décrire un comportement. Par exemple, lorsque les variations de résistance dues à la température sont négligeables, la granularité du modèle d'une résistance peut être diminuée en négligeant la grandeur « température ». La granularité peut également être diminuée en négligeant la dynamique de certains phénomènes par rapport à d'autres.

1.2.2.3.2. (Im)précision et (In)certitude

L'imprécis et l'incertain peuvent être vus comme deux points de vue complémentaires sur une même réalité qu'est l'imperfection de l'information. L'imprécis concerne le contenu de l'information (sa valeur), tandis que l'incertain est relatif à sa vérité (confiance que l'on a sur le

¹ On peut très bien imaginer que les sorties du modèle résultant d'une simulation soient des objets plus complexes que de simples variables.

contenu de l'information) [Dubois, 1994]. La précision reflète le niveau de détail mis pour exprimer les distinctions entre les valeurs des variables décrivant le système.

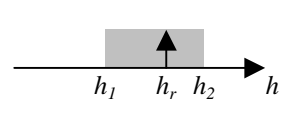
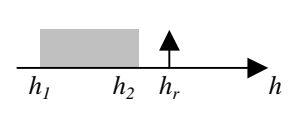
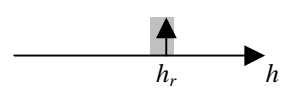
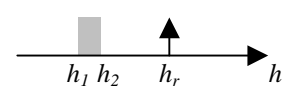
1.2.2.3.3. Validité

Par validité, nous entendons l'adéquation entre les comportements décrits et/ou prédits par le modèle et la réalité du système modélisé. En pratique, la phase de validation consiste à réaliser une série de protocoles expérimentaux pour vérifier cette adéquation.

1.2.2.3.4. Illustration de quelques caractéristiques d'un modèle

Soit h une grandeur représentant la hauteur de liquide dans un bac. La valeur réelle h_r désigne la vraie hauteur de liquide dans le bac. Une information I sur h_r peut être donnée par un intervalle $[h_1; h_2]$ et par la certitude (ou confiance) accordée à la proposition $h_r \in [h_1; h_2]$. Sur la Figure I.4, $[h_1; h_2]$ est représenté par une zone grisée et h_r est représentée par une flèche verticale.

1^{er} cas : l'information I est complètement certaine (confiance totale)
 I peut confirmer ou infirmer la validité du modèle de la hauteur de liquide :

	Validité : Oui	Validité : Non
Faible précision de I		
Forte précision de I		

2^{ème} cas : l'information I est complètement incertaine (confiance nulle)
 I ne peut ni confirmer ni infirmer la validité du modèle de la hauteur de liquide.

Figure I.4 : certitude, précision et validité

La Figure I.5 illustre également quelques propriétés d'un modèle. La courbe noire représente l'évolution temporelle de la sortie réelle du système physique à modéliser. La certitude de l'information donnée par la sortie du modèle est supposée parfaite. La largeur de l'intervalle sur les valeurs de sortie du modèle est décrite par la zone grisée qui reflète alors la précision du modèle. La ligne médiane de cette zone grisée est donnée par la courbe en pointillés.

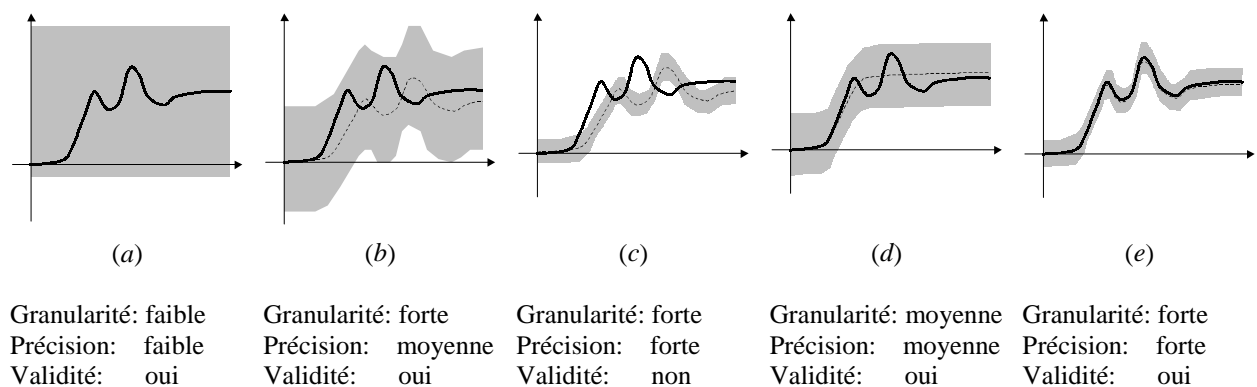


Figure I.5 : Illustration des propriétés d'un modèle

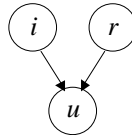
A moins que la précision ne soit un critère indispensable¹, le modèle (*d*) est sans doute le meilleur compromis entre validité, précision et granularité. En comparant (*d*) avec (*c*) et (*e*), il apparaît qu'une précision moins bonne permet de construire un modèle dont la validité dépend moins de petites perturbations. Ce raisonnement peut toutefois trouver ses limites avec (*a*) : même si ce modèle est valide, son imprécision est telle qu'il apporte une information trop pauvre pour déduire un temps de réponse par exemple. En contrepartie, l'information disponible (sortie positive ou non) est suffisamment générale pour correspondre à une grande variété de comportements.

1.2.2.3.5. Causalité explicite ou non

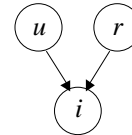
$$u = r.i$$

$$(\Leftrightarrow i = \frac{u}{r})$$

Modèle sans causalité explicite



Modèle avec causalité explicite
u : conséquence de *i* et *r*



Modèle avec causalité explicite
i : conséquence de *u* et *r*

Table I.6 : Modèle avec ou sans causalité explicite : exemple de la loi d'Ohm

Un modèle peut représenter de manière explicite ou non la notion de causalité. La causalité est un concept intéressant pour l'interprétation de phénomènes : en supposant qu'un ordre causal puisse être établi, une série de questions et de réponses du type « Pourquoi ? Parce que ... » permet de remonter itérativement vers l'origine d'un phénomène. L'origine d'un phénomène pouvant être une anomalie, l'intérêt de la causalité pour le diagnostic apparaît clairement. Un modèle avec causalité explicite définit le sens des influences entre variables (Table I.6).

1.2.2.3.6. Modèle statique ou dynamique

Un modèle peut être statique ou dynamique. Contrairement à un modèle statique, un modèle dynamique est caractérisé par la prise en compte de grandeurs mémorisant une partie du passé : l'état. Un état étant donné, un modèle dynamique permet de calculer l'état suivant et ainsi de suite : le comportement global du système se déduit donc de la transition entre deux situations (où états) élémentaires. Cette déduction correspond à une phase dite d'intégration. Pour résumer, un modèle statique décrit un équilibre alors qu'un modèle dynamique décrit une évolution du système.

1.2.2.4. Quelques exemples de modèles

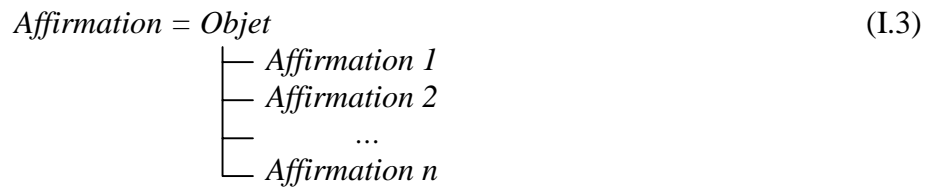
1.2.2.4.1. Modèle fonctionnel reposant sur un langage graphique

Supposons que nous souhaitions réaliser la modélisation fonctionnelle d'une automobile. Un ensemble d'objets élémentaires décrivant les principales fonctions est tout d'abord choisi :

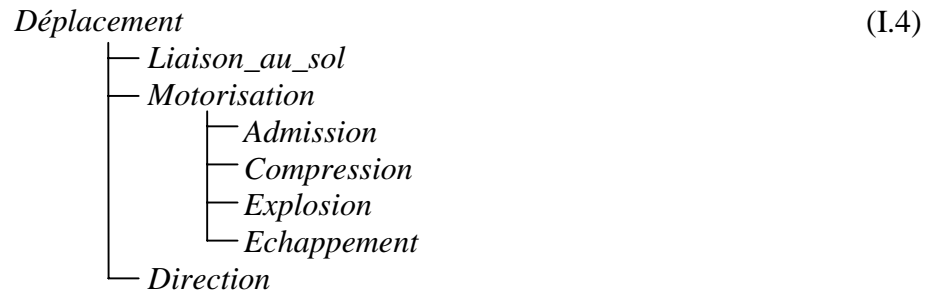
$$O = \{Déplacement, Liaison_au_sol, Motorisation, Admission, Compression, Explosion, Echappement, Direction\} \quad (I.2)$$

La syntaxe du langage *L* retenu est la suivante :

¹ Un degré de précision élevé peut se révéler nécessaire pour arriver à discriminer deux comportements très proches.



Un modèle d'une automobile peut être donné par l'affirmation (I.4).



En supposant que la route soient continuellement en descente et que le point de départ se situe au point le plus haut, un autre modèle pourrait être donné par l'affirmation (I.5).



Toutefois, un modèle (qui ne correspond pas à la réalité) peut très bien être donné par (I.6).



Ces trois modèles mettent en lumière la notion de validité. Par validité, nous entendons ici l'adéquation entre la description donnée par le modèle et la réalité du système modélisé. Dans le premier cas (I.4), le domaine de validité est le meilleur compte tenu des objets et du langage choisi. Il n'en correspond pas moins à une vue très réductrice de la réalité : un objet correspondant à la fonction « *Plaisir_de_conduite* » n'a pas été défini. Or tout passionné d'automobile ne saurait s'en passer ! Dans le deuxième cas (I.5), si l'hypothèse est vérifiée, la validité reste bonne. L'hypothèse permet de restreindre le champ de connaissance que doit décrire le modèle. Enfin, le troisième cas (I.6) correspond à un modèle qui a très peu de chance d'être valide : une voiture dont la fonction serait de polluer et qui polluerait par le simple fait de tourner le volant est difficilement imaginable ! Les modèles que nous avons construits ((I.4) à (I.6)) assurent uniquement une fonction cognitive (Table I.5).

Le langage utilisé pour décrire le modèle n'est plus le français (langage littéraire), mais un langage formé de symboles graphiques assemblés selon des règles d'agencement (langage graphique). Ce langage nous a permis de construire des modèles fonctionnels et hiérarchiques très simples. Plus généralement, le niveau de formalisation correspondant à des langages graphiques est souvent retenu dans des méthodes d'analyse fonctionnelle telle que SADT [Galinier, 1989].

1.2.2.4.2. Modèles fondés sur un langage à la fois graphique et logico-mathématique

Certains modèles reposent sur un langage incluant une composante à la fois graphique et logico-mathématique. Les modèles dérivés des réseaux de Petri ou les modèles du type « Grafcet »

relèvent de cette catégorie [David, 1992]. Les objets élémentaires sont des places et des transitions auxquelles sont associées des comportements décrits à l'aide d'équations logiques¹ permettant une simulation de ces modèles.

La combinaison entre un langage graphique et un langage logico-mathématique est également exploitée par les graphes causaux [Leyval, 1992, 1994]. Les liens entre ces modèles et les réseaux de Petri ont été mis en évidence par Gomaa [Gomaa, 1997]. La possibilité de représenter graphiquement les relations de dépendance (causale) entre les variables facilite l'interprétation du comportement d'une installation par un opérateur humain. Ces propriétés se révèlent particulièrement intéressantes dans le domaine de la supervision pour le développement d'interfaces homme / machine efficaces [Evsukoff, 1998a].

1.2.2.4.3. Modèle comportemental reposant sur un langage logico-mathématique, distinction entre connaissance a priori et connaissance apprise

Les modèles correspondant aux langages logiques et mathématiques sont les plus couramment utilisés en Automatique. Ils sont construits à l'aide d'équations logiques, algébriques, différentielles et permettent de décrire avec précision des comportements complexes.

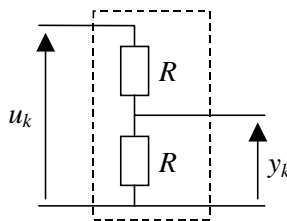


Figure I.6 : Diviseur de tension

La Figure I.6 représente un diviseur de tension dont le comportement va être modélisé à partir d'un langage logico-mathématique. Parmi les objets élémentaires utilisés, O , on peut distinguer plusieurs classes : « Opérateurs », « Constantes » et « Variables ». Parmi les variables, trois classes peuvent être distinguées : « Entrées », « Sorties », « Paramètres ».

$$\text{Opérateurs} = \{=, +, -, \cdot, / \} \quad (\text{I.7})$$

$$\text{Variables} = \text{Entrées} \cup \text{Sorties} \cup \text{Paramètres} \quad (\text{I.8})$$

$$\text{Entrées} = \{u_k, y_k\}, \text{Sorties} = \{\hat{y}_k, r_k\}, \text{Paramètres} = \{R, \theta\}$$

$$O = \text{Opérateurs} \cup \text{Variables} \quad (\text{I.9})$$

Le langage L retenu est celui des équations mathématiques « usuelles » :

$$\text{Equation: } \text{Variable} = \text{Expression} \quad (\text{I.10})$$

$$\text{Expression: } \text{Constante OU Variable OU (Expression Opérateur Expression)}$$

Deux modèles ((I.11) et (I.12)) du diviseur de tension sont donnés sous la forme d'équations (ou affirmations) :

- Le modèle (I.11) assure une fonction cognitive, même si la connaissance n'est pas décrite de manière très « intelligible ». La simulation du modèle (I.11) permet également d'assurer une

¹ Dans le cas des Réseaux de Petri continus, par exemple, les équations peuvent très bien ne plus relever uniquement de la logique booléenne.

fonction prévisionnelle. Par contre, (I.11) n'assure pas une fonction normative ou décisionnelle (Table I.5).

$$\begin{aligned}\theta &= R/(R + R) \\ \hat{y}_k &= \theta.u_k\end{aligned}\tag{I.11}$$

- Le modèle (I.12) assure une fonction cognitive. La simulation de (I.12) n'est pas possible directement car la valeur de θ n'est pas connue a priori. L'équation du résidu r_k assure une fonction normative et décisionnelle (Table I.5).

$$\begin{aligned}\hat{y}_k &= \theta.u_k \\ r_k &= y_k - \hat{y}_k\end{aligned}\tag{I.12}$$

Fonction normative de r_k : partant du modèle (I.12), on cherche une valeur de θ telle que le résidu r_k soit toujours nul, quelle que soit la simulation réalisée à partir d'un jeu d'observations significatives relevées sur le système physique (u_k, y_k). Si une telle valeur de θ peut être trouvée alors \hat{y}_k correspond à la valeur de y_k et la fonction prévisionnelle du modèle est ainsi remplie. La représentation des relations souhaitables entre éléments du système (Critère r_k nul) a permis d'apprendre la valeur de θ à partir d'un jeu d'observations.

Fonction décisionnelle de r_k : en supposant qu'une valeur de θ soit connue (a priori ou apprise), la simulation de (I.12) permet de calculer le résidu r_k . r_k est d'autant plus proche de zéro que le comportement du système physique, connu au travers des observations (u_k, y_k), s'apparente à celui décrit par la première équation de (I.12). La valeur de r_k est donc un indicateur permettant de décider si le comportement réel du système s'apparente ou non à un comportement de référence décrit par le modèle. La fonction décisionnelle d'un modèle est celle prioritairement recherchée dans le contexte du diagnostic.

A travers l'exemple qui vient d'être traité, deux types de connaissance servant à l'élaboration d'un modèle peuvent être distingués :

- **Connaissance a priori :** les lois de la Physique sont des exemples de connaissances a priori utilisées lors de l'élaboration d'un modèle. Ainsi, le modèle (I.11) est obtenu par application de la loi d'Ohm et de la loi des mailles. Les modèles comportementaux construits à partir d'une connaissance a priori sont traditionnellement appelés « modèles de connaissance ».
- **Connaissance apprise :** étant donné une certaine connaissance a priori définissant une famille de modèles (par exemple, $\hat{y}_k = \theta.u_k$, $\theta \in \mathfrak{R}$) et un ensemble d'observations significatives, la connaissance supplémentaire servant à choisir un modèle pertinent dans la famille de modèles peut être obtenue par apprentissage (ou estimation ou, plus généralement, par optimisation). Dans l'exemple, cette connaissance supplémentaire est la valeur du paramètre θ , estimée de façon à minimiser le résidu r_k . Les modèles comportementaux construits à l'aide d'une connaissance apprise sont traditionnellement appelés « modèles de comportement ».

La Figure I.7 rappelle les principaux concepts associés à la notion de modèle. Nous allons revenir plus en détail sur la notion de cohérence, fondamentale pour le diagnostic. Nous pouvons noter qu'elle correspond au résidu r_k pour le modèle (I.12).

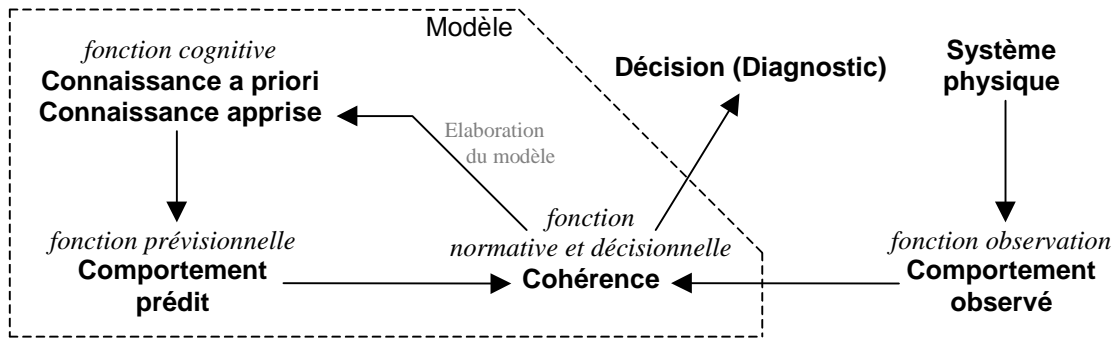


Figure I.7 : Principaux concepts associés à la notion de modèle (modèle comportemental)

I.2.3. Validité, cohérence et élaboration d'un modèle

Deux définitions sont proposées pour les termes « validité » et « cohérence » :

- **Validité** : propriété attribuée à un modèle ayant satisfait une série de protocoles expérimentaux destinés à vérifier la cohérence entre les comportements (et/ou les structures et/ou les fonctionnalités) décrits par ce modèle et la réalité du système modélisé. La validité est associée à un domaine de fonctionnement.
- **Cohérence** : rapport logique entre choses connexes [Pruvost-Beaurain, 1990].

L'élaboration d'un modèle consiste essentiellement à organiser les connaissances disponibles (a priori et apprises) de sorte que la validité de ce modèle sur un certain domaine de fonctionnement soit aussi certaine que possible. Un modèle n'ayant aucune propriété de validité serait en effet totalement inutilisable, puisque les informations qu'il donnerait seraient totalement incertaines (confiance nulle). Pour garantir au mieux la validité d'un modèle, il est nécessaire d'établir un lien entre les comportements (et/ou les structures et/ou les fonctionnalités) décrits par ce modèle et la réalité du système modélisé connue au travers des observations disponibles. Ce lien correspond à la notion de cohérence. En supposant que la cohérence soient évaluée sous la forme d'une grandeur booléenne, on peut poser :

$$\text{validité}(\text{modèle}, \text{domaine}) \Leftrightarrow \forall \text{observation} \in \text{domaine}, \text{cohérence}(\text{modèle}, \text{observation}) \quad (\text{I.13})$$

Bilan : lorsque l'objectif poursuivi est l'élaboration d'un modèle, le système physique est supposé dans un domaine de fonctionnement donné et les différences entre les comportements prédits et observés mettent en évidence les changements à appliquer au modèle (Figure I.7).

Table I.7 : Elaboration d'un modèle

I.2.4. Validité, cohérence et diagnostic

Nous nous sommes attachés à définir la notion de modèle indépendamment de la technique de modélisation utilisée. Notre objectif est maintenant de mettre en évidence les points communs à de très nombreuses méthodes de diagnostic. La Table I.3 est le point de départ de notre raisonnement. La question à laquelle nous cherchons à répondre est la suivante : comment réaliser un diagnostic ? Autrement dit (Table I.3), comment utiliser la connaissance sur le

système et sur son environnement pour obtenir une observation des anomalies ? Cette question peut à nouveau être reformulée (Table I.3) : comment utiliser les observations et le(s) modèle(s) du système pour obtenir une observation des anomalies ?

La seule information disponible sur les anomalies qui se sont effectivement produites est donnée par les observations. Sauf cas tout à fait exceptionnel, les observations élémentaires (valeurs de consigne, mesures,...) ne suffisent pas seules à obtenir une observation des anomalies. La connaissance supplémentaire qui est nécessaire ne peut alors provenir que du(des) modèle(s) disponibles. Or, ces modèles servent à décrire des comportements (et/ou des structures et/ou des fonctionnalités). Il est donc nécessaire d'établir un lien entre les observations et les comportements (et/ou les structures et/ou les fonctionnalités) décrits par les modèles. Comme dans la partie I.2.3, c'est la notion de cohérence qui est utilisée pour établir ce lien. Cela nous amène à énoncer un principe du diagnostic fondé sur la cohérence (Table I.8).

Réaliser une tâche de diagnostic, c'est-à-dire obtenir une observation porteuse d'une information relative aux anomalies et à elles seules consiste à évaluer la cohérence entre des observations et des comportements, structures ou fonctionnalités décrits par un modèle approprié.

Table I.8 : principe du diagnostic

L'énoncé de la Table I.9 permet de préciser ce que l'on entend par « modèle approprié » (Table I.8) dans le cas d'un modèle comportemental :

Bilan : lorsque l'objectif poursuivi est le diagnostic, le modèle est supposé valide sur un domaine de fonctionnement donné et les différences entre les comportements prédits et observés mettent en évidence un fonctionnement du système physique hors du domaine de validité du modèle (Figure I.7).

Table I.9 : diagnostic

La similitude entre les énoncés de la Table I.7 et de la Table I.9 illustre les liens très étroits qui unissent la notion de modèle et le diagnostic. Fonder le diagnostic sur la notion de cohérence n'est pas nouveau ; De Kleer et Williams présentent également le parallèle entre la construction d'un modèle et le diagnostic, mais uniquement dans le cadre d'une approche orientée « bon fonctionnement » [De Kleer, 1987]. Pour Poole, un problème de diagnostic consiste à utiliser un modèle pour chercher des explications satisfaisantes aux symptômes observés [Poole, 1989]. Un critère permettant de juger plus ou moins satisfaisantes certaines explications correspond précisément à la notion de cohérence.

1.2.5. Approches « bon fonctionnement » et « mauvais fonctionnement »

Nous allons maintenant montrer que le principe d'évaluation de la cohérence est très général dans la mesure où il englobe aussi bien les approches orientées « Bon Fonctionnement » (BF) que les approches orientées « Mauvais Fonctionnement » (MF).

1.2.5.1. Cas général

Conformément à l'énoncé de la Table I.9, nous supposons un modèle valide sur un certain domaine (I.14). (I.15) se déduit alors de (I.13) :

$$\text{Hypothèse: } \text{validité}(\text{modèle}, \text{domaine}) \quad (\text{I.14})$$

$$\forall \text{observation} \in \text{domaine}, \text{cohérence}(\text{modèle}, \text{observation}) \quad (\text{I.15})$$

Soit *obs* une observation particulière du comportement du système physique. En supposant que la cohérence soient évaluée sous la forme d'une valeur booléenne, deux cas peuvent se présenter : (I.16) et (I.17).

$$\text{Si } \text{cohérence}(\text{modèle}, \text{obs}) = \text{faux} \text{ Alors } \ll \text{obs} \notin \text{domaine} \gg \text{ est certain} \quad (\text{I.16})$$

$$\text{Si } \text{cohérence}(\text{modèle}, \text{obs}) = \text{vrai} \text{ Alors } \ll \text{obs} \in \text{domaine} \gg \text{ est incertain} \quad (\text{I.17})$$

Bilan : seule la non cohérence donne systématiquement une information certaine.

Remarque 1 : une illustration de (I.17) peut être obtenue en considérant le système décrit par (I.18) où *L*, *R* et ΔL désignent respectivement une inductance, une résistance et un défaut modélisé par une variation d'inductance.

$$\frac{y}{u} = \frac{1}{(L + \Delta L).p + R} \quad (\text{I.18})$$

En régime permanent ($p=0$), même si le comportement entrée/sortie est cohérent avec le fonctionnement normal ($\Delta L=0$), il est néanmoins possible que le défaut ΔL soit présent, simplement « masqué » par le manque d'excitation du système. Bien que la cohérence avec le fonctionnement normal soit vraie, l'appartenance au domaine de fonctionnement « normal » des observations effectuées en régime permanent n'est pas certaine.

Remarque 2 : lorsque l'appartenance des observations à *n* domaines sont testées (*n* modèles) et qu'une incohérence est relevée pour (*n*-1) d'entre eux, tous les domaines sauf un, noté *D*, sont exclus. L'appartenance des observations à *D* est a priori incertaine (I.17). Cependant, en supposant que les *n* domaines représentent l'ensemble des alternatives possibles, la seule possibilité est une appartenance des observations à ce domaine *D*.

Remarque 3 : si l'équivalence (I.19) peut être établie (condition plus forte que (I.15)), la cohérence et la non cohérence donnent toutes les deux une information certaine.

$$\text{observation} \in \text{domaine} \Leftrightarrow \text{cohérence}(\text{modèle}, \text{observation}) \quad (\text{I.19})$$

1.2.5.2. Approche orientée Bon Fonctionnement (BF)

Dans une approche orientée BF, le modèle utilisé décrit le fonctionnement normal (sans défaut) de tout ou partie du système physique à diagnostiquer. L'application du cas général à un modèle BF donne (I.20) et (I.21) (analogie avec (I.14) et (I.16)) :

$$\text{Hypothèse: } \text{validité}(\text{modèle_BF}, \text{domaine_BF}) \quad (\text{I.20})$$

$$\text{Si } \text{cohérence}(\text{modèle_BF}, \text{obs}) = \text{faux} \text{ Alors } \ll \text{obs} \notin \text{domaine_BF} \gg \text{ est certain} \quad (\text{I.21})$$

La non cohérence avec un modèle de bon fonctionnement permet de montrer qu'une observation n'appartient pas au domaine de bon fonctionnement associé au modèle. Cela revient à résoudre un problème de détection sur tout ou partie du système physique à diagnostiquer.

Un modèle BF de l'ensemble du système physique à diagnostiquer permet de résoudre le problème de la détection. Un ensemble de modèles BF relatifs à plusieurs sous-systèmes permet

de détecter la présence de défauts dans chacun des sous-systèmes modélisés, pris indépendamment. Une approche BF multi-modèles permet donc de déterminer des sous-systèmes en défaut, ce qui correspond à un premier niveau de localisation.

Remarque : si un défaut au plus est susceptible d'affecter chacun des sous-systèmes modélisés, alors la localisation des défauts peut être effectuée par une approche BF appliquée à chacun des modèles des sous-systèmes.

I.2.5.3. Approche orientée Mauvais Fonctionnement (MF)

Dans une approche orientée MF, le modèle utilisé décrit un fonctionnement anormal du système physique à diagnostiquer. Par exemple, un modèle décrivant le comportement du système physique en présence d'un défaut particulier est un modèle MF. L'application du cas général à un modèle MF donne (I.22) et (I.23) (analogue à (I.14) et (I.16)) :

$$\text{Hypothèse: } \text{validité}(\text{modèle_MF}, \text{domaine_MF}) \quad (\text{I.22})$$

$$\text{Si } \text{cohérence}(\text{modèle_MF}, \text{obs}) = \text{faux} \text{ Alors } \ll \text{obs} \notin \text{domaine_MF} \gg \text{ est certain} \quad (\text{I.23})$$

La non cohérence avec un modèle de mauvais fonctionnement permet de montrer qu'une observation n'appartient pas au domaine de mauvais fonctionnement associé au modèle. Cela revient à rejeter la possibilité d'occurrence de la ou des anomalie(s) décrite(s) par le modèle MF. Une approche MF s'inscrit donc dans une démarche de localisation. Si un modèle MF est disponible pour chacune des anomalies, une approche MF multi-modèles permet de rejeter sélectivement chacune des anomalies et, par conséquent, de remplir les objectifs de la localisation.

I.2.5.4. Bilan sur les approches BF et MF

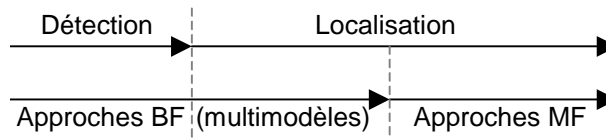


Figure I.8 : Problèmes résolus par les approches BF et MF (d'après [Ploix, 1998])

Même si une approche BF multi-modèles peut résoudre un problème de localisation (Figure I.8), une approche BF (mono-modèle) résout fondamentalement un problème de détection (I.24), alors qu'une approche MF résout un problème de localisation (I.25).

$$\text{observation_détection} = \text{non}(\text{cohérence}(\text{modèle_BF}, \text{observation})) \quad (\text{I.24})$$

$$\text{observation_localisation} = \text{cohérence}(\text{modèle_MF}, \text{observation}) \quad (\text{I.25})$$

(I.24) et (I.25) sont parfaitement conformes au principe du diagnostic de la Table I.8. L'interprétation qui peut être faite des observations relatives aux anomalies, *observation_détection* et *observation_localisation*, est donnée par la Table I.10.

Modèle	Bon Fonctionnement (BF)	Mauvais Fonctionnement (MF)
observation_anomalies	observation_détection (I.24)	observation_localisation (I.25)
Signification de observation_anomalie	« occurrence d'au moins une anomalie dans le système décrit par le modèle »	« occurrence de la ou des anomalie(s) décrite(s) par le modèle »
« Observation_anomalie=vrai » correspond à une ...	Certitude	Possibilité (incertitude)
« Observation_anomalie=faux » correspond à une ...	Possibilité (incertitude)	Certitude

Table I.10 : Comparaison des approches BF et MF

Remarque : la validation d'un modèle BF est souvent plus facile à obtenir que celle d'un modèle MF : des observations du système physique en défaut ne sont généralement pas aisément accessibles.

I.2.6. Décomposition du diagnostic

I.2.6.1. Associations en parallèle et en série

Pour diagnostiquer un système, toute la difficulté repose sur l'obtention d'un modèle approprié. Même pour des systèmes simples, un tel modèle s'avère très complexe ne serait-ce que parce qu'il doit permettre de donner une information sur l'ensemble des anomalies. Heureusement, il est tout à fait possible de décomposer la tâche de diagnostic en plusieurs étapes. D'après la Table I.3, nous pouvons déduire :

$$Observations_anomalies = diagnostic(Modèles, Observations) \quad (I.26)$$

Pour aboutir à un diagnostic global, des diagnostics plus élémentaires peuvent être associés en parallèle ou en série.

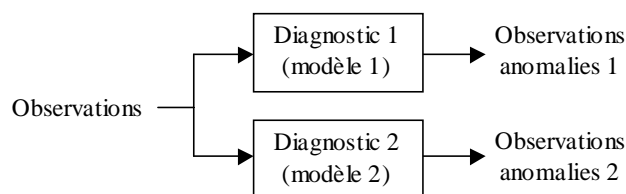


Figure I.9 : association de diagnostics en parallèle

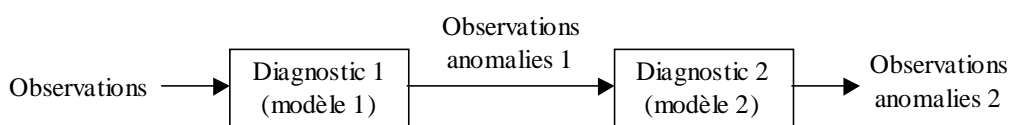


Figure I.10 : association de diagnostics en série

L'association en parallèle (Figure I.9) est utile pour traiter des groupes d'anomalies de façon indépendante, c'est-à-dire avec des modèles distincts. Les distinctions entre modèles peuvent être classées en deux catégories :

- **La manière de décrire un même système est différente** : par exemple, un modèle peut être construit à partir de grandeurs quantitatives alors qu'un autre le sera à partir de grandeurs qualitatives ou floues.
- **Le système décrit par chacun des modèles est différent** : par exemple, une démarche de localisation dans le cadre d'une approche orientée bon fonctionnement repose sur ce type de différence entre modèles : pour un système physique formé de plusieurs composants, le comportement normal de chacun des composants peut être décrit par un modèle particulier.

L'association en série (Figure I.10) permet de combiner des modèles distincts pour diagnostiquer un même groupe d'anomalies : un premier modèle (modèle 1) fournit un pré-traitement des observations élémentaires ; ce pré-traitement revient à extraire certains signes et/ou certains symptômes qui seront ensuite traités par un deuxième modèle (modèle 2). Etape par étape, l'association en série permet de passer d'observations élémentaires à des observations de plus haut niveau, jusqu'à aboutir à un niveau jugé suffisant pour remplir le cahier des charges du diagnostic.

Les associations en série et en parallèle illustrent la possibilité de hiérarchiser la tâche de diagnostic. Une telle hiérarchisation constitue le fondement des méthodes utilisant un arbre de décision. En outre, l'objectif poursuivi en travaillant avec plusieurs modèles est d'obtenir des points de vue complémentaires sur une même réalité : le système physique à diagnostiquer. Des modèles de nature très différente peuvent être associés ; les performances globales d'un diagnostic résident donc au moins autant dans la complémentarité des modèles utilisés que dans les performances atteintes par chacun des modèles élémentaires¹. Nous allons maintenant analyser une classe particulière d'associations de modèles pour le diagnostic.

1.2.6.2. Un exemple de diagnostic par redondance analytique

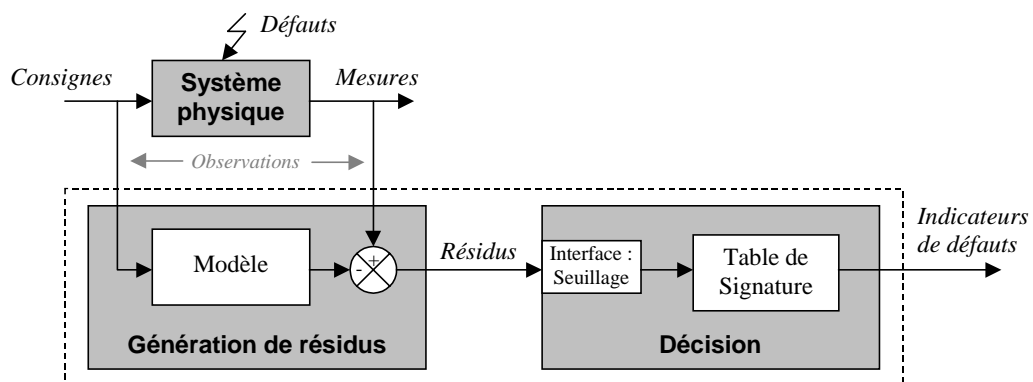


Figure I.11 : Un exemple de diagnostic par redondance analytique

La Figure I.11 montre le cas où un diagnostic global (entouré en pointillés) résulte de l'association en série de deux diagnostics élémentaires :

¹ Pour compléter ce propos, le choix des observations (mesures) est également un critère très important pour les performances finales du diagnostic.

- **Génération de résidus :** la génération de résidus est fondée sur un modèle du système physique. Cette étape consiste à calculer les résidus qui sont des indicateurs de cohérence entre les observations élémentaires (consignes, mesures) et le comportement décrit par le modèle. La génération de résidus est très souvent basée sur des modèles quantitatifs. Sauf prise en compte explicite des incertitudes, ces modèles sont aussi précis que leur domaine de validité est restreint. Les résidus contiennent une information de plus haut niveau sur les défauts que les observations élémentaires ; en particulier, un modèle de bon fonctionnement va contribuer à retirer des résidus les dynamiques de réponse aux consignes et à certaines perturbations ; ces résidus ne sont alors plus sensibles qu'aux autres perturbations et aux défauts.
- **Décision :** les résidus numériques doivent néanmoins être traités pour construire des indicateurs de défaut. C'est l'objet de la décision qui consiste à classer l'infinité des comportements possibles des résidus en plusieurs catégories représentant les différentes situations de défaut. Les modèles (MF) utilisés pour la décision sont très souvent qualitatifs (y compris les modèles flous ou booléens comme les tables de décision). Les grandeurs traitées par ces modèles sont en général peu précises ; en contrepartie, ces modèles sont en général valides sur un large domaine. A condition que les comportements des résidus soient décrits avec une précision suffisante pour discriminer entre eux les différents défauts, les modèles pour la décision permettent de rétablir un domaine de validité pour le diagnostic en prenant en compte les incertitudes (par seuillage, par exemple).

L'association en série de deux classes de modèles pour le diagnostic, la génération de résidus et la décision, montre l'intérêt de combiner des modèles dont les propriétés de précision et de validité sont complémentaires. Le prix à payer pour une combinaison efficace consiste à gérer correctement les interfaces entre les modèles. Ainsi, le choix des seuils dans le cas de la Figure I.11 est une opération délicate dont dépend le domaine de validité du diagnostic. La validité (l'exactitude) du diagnostic est un point crucial : en effet, un indicateur de défaut qui pourrait être lui-même pris en défaut ne servirait à rien.

La génération de résidus et la décision peuvent elles-mêmes résulter de l'association de diagnostics plus élémentaires. Par exemple, plusieurs techniques de génération de résidus peuvent être associées afin d'aboutir à une génération de résidus globale pour le système physique à diagnostiquer.

1.2.7. Le diagnostic de défauts : découplage direct et indirect

Ce paragraphe est ciblé sur la formulation du diagnostic de défauts lorsqu'un modèle comportemental du système S peut s'exprimer sous la forme donnée par (I.27).

$$\begin{aligned} \dot{x} &= g_S(x, u, f, d), & x(0) &= x_0 \\ y &= h_S(x, u, f, d) \end{aligned} \tag{I.27}$$

g_S et h_S représentent alors les équations d'état continues du système. Tout le développement de ce paragraphe pourrait être également écrit avec un modèle discret. g_S et h_S peuvent être construites avec n'importe quel opérateur arithmétique ou logique. (I.27) représente donc la structure d'un modèle décrivant une large classe de systèmes non linéaires où x est l'état (inconnu a priori), u est le vecteur des entrées connues (consignes), y est la sortie mesurée, f représente les défauts et d les perturbations. Compte tenu de la forme de (I.27), f et d peuvent être des entrées non additives. f et d induisent une modification du comportement entrée-sortie du système (entre u et y), ce qui est conforme avec les définitions adoptées pour les défauts et les

perturbations. La détection consiste à décider si le système est ou n'est pas en défaut ($f \neq 0$ ou $f=0$), et ce, indépendamment des perturbations et des changements de point de fonctionnement. Une manière d'atteindre cet objectif consiste à concevoir un indicateur de détection μ qui est souvent booléen ou bien compris dans l'intervalle $[0 ; 1]$.

$$\dot{w} = g_M(w, u, y), \quad w(0) = w_0 \quad (\text{I.28})$$

$$\mu = h_M(w, u, y)$$

$$g_M, h_M \quad / \quad \mu(w, u, y) = 0 \Leftrightarrow f = 0 \quad (\text{I.29})$$

Ne dépendant que de grandeurs connues, μ peut être calculé en ligne conformément à (I.28). L'objectif poursuivi est de concevoir g_M et h_M tels que μ remplisse les objectifs d'une détection « idéale » (I.29); μ représente alors un indicateur de cohérence entre les observations disponibles (u et y) et une situation de défaut quelconque. Pour cela, l'influence des variables inconnues peut être évaluée en substituant y par (I.27) dans (I.28). Les conditions à satisfaire par g_M et h_M , c'est-à-dire par le modèle de diagnostic M permettant une détection « idéale » sont donc les suivantes :

$$g_M, h_M \quad / \quad \mu(w, u, f, x, d) = 0 \Leftrightarrow f = 0 \quad (\text{I.30})$$

(I.30) est difficile à satisfaire en raison de l'influence des grandeurs inconnues x et d sur μ . Pour remplir les objectifs de la détection de défauts, il faut par conséquent trouver g_M et h_M tels que μ soit découplé de x et de d tout en étant aussi sensible que possible à f . Pour cela, deux approches peuvent être utilisées : l'élimination ou découplage direct, l'estimation ou découplage indirect [Zhang, 1998], [Krishnaswami, 1994].

- Le **découplage direct** consiste à extraire du modèle de départ un sous-modèle insensible aux entrées à découpler ; cela revient à n'utiliser le modèle que d'une partie du système, l'autre étant éliminée. La cohérence est alors testée entre le comportement décrit par le sous-modèle considéré et les observations disponibles. On distingue le découplage parfait¹ du découplage approximatif, ce dernier visant à minimiser (et non à annuler) l'influence des variables inconnues sur μ . Un découplage direct peut être vu comme un découplage réalisé en boucle ouverte : techniques de projection pour la génération de résidus, techniques reposant sur des domaines fixés a priori pour la décision...
- Le **découplage indirect** consiste à estimer les entrées à découpler pour compenser leur influence ; le raisonnement se fait alors sur l'ensemble du système considéré, ce qui présente l'avantage de conserver toute l'information disponible sur l'évolution relative des différentes grandeurs mises en jeu. A moins que l'estimation soit parfaite, seul un découplage approximatif peut être obtenu. Un découplage indirect peut être vu comme un découplage réalisé en boucle fermée (l'état du modèle permet d'adapter le découplage en cours de simulation) : techniques d'estimation pour la génération de résidus, techniques d'apprentissage en ligne pour la décision...

Comme nous l'avons évoqué dans la partie I.2.6.2, la conception de g_M et de h_M est souvent décomposée en deux étapes : la génération de résidus et la décision. Cependant, le diagnostic de

¹ Précisons que le qualificatif « parfait » d'un découplage est tout relatif : il ne s'applique vraiment que si le modèle utilisé modélise parfaitement le système physique...

défauts est avant tout un problème de découplage, indépendamment de la manière dont il est décomposé. De plus, on peut souligner que le découplage de grandeurs inconnues peut être effectué soit par la génération de résidus, soit par la décision.

Objectif	Type de variable:		
	Testée	Connue	Inconnue
Détection	f	w, u	x, d
Localisation de f_i	f_i	w, u	$f_{j \neq i}, x, d$

Table I.11 : Détection et Localisation des défauts

Le problème de localisation « idéale » peut être formulé dans les mêmes termes que la détection, sauf que certaines variables modélisant l'influence des défauts doivent également être découplées. Ainsi, la localisation du $i^{\text{ème}}$ défaut, f_i , consiste à concevoir un modèle g_{M_i}, h_{M_i} , tel que l'indicateur de défaut μ_i satisfasse (I.31).

$$g_{M_i}, h_{M_i} / \mu_i(w, u, f, x, d) = 0 \Leftrightarrow f_i = 0 \quad (\text{I.31})$$

μ_i devrait donc être conçu de façon à être découplé de toutes les variables inconnues : $f_{j \neq i}, x$ et d (Table I.11). Ceci est d'autant plus difficile à satisfaire que le nombre de défauts à diagnostiquer augmente. Néanmoins, sous l'hypothèse de défaut simple, toutes les combinaisons impliquant plus d'un défaut à la fois sont supposées impossibles et ne requièrent donc pas de découplage. Dans ces conditions, μ_i doit seulement être découplé de $\{f_j, x, d\}$ pour tous les $j \neq i$ ce qui est nettement moins contraignant.

Remarque : dans ce paragraphe, nous n'avons traité la détection et la localisation que dans un cas « idéal » : arriver à satisfaire strictement (I.30) ou (I.31) est impossible en pratique. En effet, il n'est pas possible de détecter ou de localiser des défauts arbitrairement petits, compte tenu des incertitudes. Ces aspects seront repris plus en détail dans le chapitre V.

La distinction qui est souvent faite entre la détection et la localisation de défauts ne correspond pas à des problèmes différents par nature. Dans les deux cas, il s'agit de combiner des techniques de découplage (direct et/ou indirect) pour satisfaire au mieux le cahier des charges du diagnostic. L'intérêt d'un indicateur de détection fiable et rapide à calculer réside dans le séquençage de la tâche de diagnostic : les calculs de localisation peuvent n'être effectués que lorsqu'un défaut est détecté. Sous certaines conditions, cela peut éviter de sur-dimensionner les calculateurs exécutant les tâches de diagnostic.

I.3.

Différentes méthodes pour le diagnostic

METHODES POUR LE DIAGNOSTIC

- Redondance matérielle
- Diagnostic de défaillances (modèles fonctionnels): AMDEC, arbres de défaillances,...
- Interfaces entre observations de nature différente
 - Seuillage
 - Fuzzification / Défuzzification
- Traitement du signal (modèle comportemental d'un signal(*))
 - Filtrage, en particulier:
 - Analyse fréquentielle ou temps-fréquence (Fourier)
 - Analyse temps-échelle (ondelettes)
 - Tests statistiques
- Génération de résidus (modèle comportemental d'un système(*))
 - Exploitation de l'existence de relations entre variables
 - Relations non orientées a priori : approche dite "structurelle"
 - Relations orientées: graphe causal
 - Exploitation de la nature des relations entre variables
 - Equations de parité
 - Observateurs
 - Estimation paramétrique
- Décision (indépendance vis-à-vis de la nature de la modélisation)
 - Classification par coalescence
 - Inférences à base de règles, en particulier :
 - Arbres de décision
 - Tables de décision
- Problèmes transversaux :
 - Placement de capteurs
 - Prise en compte des incertitudes
 - Apprentissage : techniques d'optimisation (réseaux de neurones, algorithmes génétiques,...)

(*) Un signal ou un système correspondent tout deux à une réalité qui peut être l'objet d'une modélisation. On pourrait ainsi très bien changer "système physique" par "signal" dans la formulation du problème de diagnostic.

Figure I.12 : Représentation hiérarchisée des principales méthodes pour le diagnostic

1.3.1. Redondance matérielle

Cette première méthode consiste à multiplier physiquement les éléments critiques d'une installation. Un traitement des signaux issus des éléments redondants effectue des comparaisons et distingue l'élément défectueux en cas d'incohérence. Cette méthode est pénalisante en termes de poids, puissance consommée, volume et coûts (d'achat et de maintenance). Elle est donc essentiellement réservée aux cas où la continuité de service est obligatoire (avionique). En effet, elle apporte l'avantage, une fois la défaillance détectée et localisée, de pouvoir utiliser la partie de l'équipement encore saine.

1.3.2. Diagnostic de défaillances

Les méthodes basées sur un modèle fonctionnel servent à diagnostiquer les défaillances. Très souvent des approches inductives ou déductives sont retenues : elles reposent sur une modélisation de « toutes » les causes de défaillances ou, au contraire, de leurs conséquences. Les arbres de défaillances ou l'AMDEC (Analyse des Modes de Défaillances et de leurs Effets Critiques) relèvent de cette catégorie de méthodes [Legg, 1978]. Nous nous intéressons maintenant plus spécifiquement au diagnostic de défauts qui correspond au cadre de notre étude.

1.3.3. Interfaces entre observations de nature différente

1.3.3.1. Seuillage

Des interfaces sont nécessaires pour assurer que les grandeurs disponibles soient compatibles avec les modèles utilisés. Très souvent, les résidus sont des grandeurs numériques, alors que la décision repose sur un modèle logique. Un moyen simple de convertir des grandeurs numériques en grandeurs logiques consiste tout simplement à les seuiller. Les seuils sont dits « adaptatifs » lorsqu'ils dépendent du point de fonctionnement [Frank, 1994a].

1.3.3.2. Fuzzification/Défuzzification

L'inconvénient du seuillage net est qu'il conduit à des décisions instables : suivant que la grandeur testée est juste au-dessus ou juste au-dessous du seuil, le résultat du test est complètement remis en question. Pour éviter l'instabilité des décisions due au bruit de mesure, la logique floue peut être utilisée. Outre la progressivité des décisions, l'intérêt de la logique floue est également de permettre une représentation unifiée de la connaissance [Montmain, 1992], [Kiupel, 1993], [Isermann, 1998] : il est ainsi possible de fusionner des informations décrites à l'aide de grandeurs numériques, logiques ou qualitatives (Figure I.3).

1.3.4. Traitement du signal

1.3.4.1. Filtrage

Une première approche du traitement du signal repose sur l'analyse fréquentielle (transformée de Fourier). Elle est bien évidemment très utilisée pour la détection de phénomènes périodiques comme en analyse vibratoire. Le contenu spectral des signaux est utilisé depuis de nombreuses années pour détecter des défauts dans les machines électriques [Cameron, 1986], [Thomson, 1999] tels que les ruptures de barres au rotor des machines asynchrones, la dégradation des roulements, les décentrages, les courts-circuits dans les bobinages. Avec le développement des applications à vitesse variable, les recherches actuelles portent plus particulièrement sur les méthodes adaptées à la caractérisation de signaux non stationnaires : temps-fréquence, temps-échelle (décomposition en ondelettes [Strang, 1996], [Borras, 1999]).

1.3.4.2. Tests statistiques

Une autre catégorie de méthodes de traitement du signal repose sur la détection de changement des caractéristiques statistiques (moyenne, variance) d'une variable aléatoire. Par exemple, le test de somme cumulée CUSUM dit de Page-Hinkley permet de détecter en ligne un changement de moyenne. Cet algorithme est très simple à mettre en œuvre, rapide et robuste au bruit [Basseville, 1988]. D'autres tests existent comme le maximum de vraisemblance généralisé (GLR : Generalized Likelihood Ratio). Le test du GLR, introduit dans [Willsky, 1976], permet de rechercher une forme particulière de rupture dans un signal aléatoire. Il est en général présenté sous une forme nécessitant le calcul de l'innovation d'un filtre de Kalman [Basseville, 1988]. Les résultats des tests d'hypothèses fondés sur des méthodes statistiques étant binaires, ces derniers assurent un rôle d'interface entre les grandeurs numériques d'entrée et les grandeurs logiques de sortie. Le test du GLR utilise un modèle pour représenter la connaissance sur le signal : le calcul de l'innovation du filtre de Kalman peut être vu comme une génération de résidu, ce résidu étant traité parallèlement par des méthodes statistiques pour évaluer le résultat du test.

1.3.5. Génération de résidus

1.3.5.1. Exploitation de l'existence de relations entre variables

Un modèle structurel décrit en général comment est fait un système : liste des composants, connexions entre les composants, ... Un modèle dit « structurel » peut également être vu comme une abstraction des contraintes d'un modèle comportemental [Feliot, 1997]. Il modélise alors l'existence ou l'absence de relations entre variables. Cette information est très utile dans la phase de conception d'un système de surveillance, en particulier pour de grands systèmes [Staroswiecky, 1989], [Cassar, 1994a, 1994b]. Les modèles structurels peuvent jouer un rôle d'interface entre le niveau fonctionnel (défaillances) et le niveau comportemental (défauts).

Un graphe causal ajoute à un modèle structurel la notion d'orientation des relations entre variables. Une telle orientation est un support permettant la représentation de la propagation de phénomènes physiques dans les procédés [Feraud-Beaumont, 1989]. Dans le contexte du diagnostic, cela permet de « suivre » la propagation des défauts et d'implanter des procédures de localisation récursives [Montmain, 1993], [Evsukoff, 1997b].

1.3.5.2. Exploitation de la nature des relations entre variables

La génération de résidus utilise très souvent la connaissance issue d'un modèle quantitatif. Nous allons focaliser notre présentation des méthodes de génération de résidus sur ce cas là, sans oublier que l'utilisation de modèles qualitatifs [Amann, 1997], [Frank, 1996], [Shen, 1993] ou logiques [Valette, 1994]¹ a déjà été proposée dans la littérature.

- **Espace de parité** : les méthodes du type équations de parité [Chow, 1984], [Gertler, 1990], [Gertler, 1997] reposent sur la vérification d'une relation statique ou dynamique qui relie les mesures. L'espace de parité peut être vu comme l'espace des résidus possibles. Il est souvent défini comme l'orthogonal de la matrice d'observabilité, ce qui revient à éliminer l'influence de l'état sur le résidu. Le découplage des perturbations est en général réalisé par une projection ce qui correspond à une technique de découplage direct (Suite à une projection, seule une partie du système de départ est utilisée, l'autre est éliminée).
- **Observateurs** : les méthodes de génération de résidus construites à l'aide d'observateurs reposent sur une estimation de l'état ; elles ont été et sont toujours abondamment traitées dans la littérature [Frank, 1990], [Ding, 1994], [Patton, 1997a]. Un observateur d'ordre réduit revient à ne considérer qu'une partie du système, donc à estimer une partie de l'état et à éliminer l'autre. Par ailleurs, l'élimination d'une partie du système peut être utilisée pour rejeter les perturbations (découplage direct). Les observateurs à entrées inconnues sont basés sur ce principe. Une classe particulière d'observateurs consiste à générer des résidus directionnels : il s'agit des filtres de détection de défaut introduits par [Beard, 1971] et combinés avec les observateurs à entrées inconnues par [Chen, 1996].

Les méthodes du type espace de parité et observateurs sont en général plutôt adaptées au découplage et à la détection de grandeurs additives. Pour remédier à ces insuffisances, des techniques non linéaires sont de plus en plus utilisées ; certains travaux portent sur les observateurs non linéaires [Alcorta Garcia, 1997], [Kabore, 1998] alors que d'autres généralisent le concept des équations de parité, par exemple en utilisant les bases de Gröbner [Zhang, 1997], [Comtet-Varga, 1999]. Cependant, la prise en compte de grandeurs paramétriques (ou

¹ Les modèles logiques utilisés dans cet article sont des Réseaux de Petri ; des éléments de comparaison avec les approches pour les systèmes continus sont donnés.

multiplicatives), conduit bien souvent à appliquer des techniques d'estimation paramétrique. Les capacités de localisation de ces techniques sont en général très bonnes et la possibilité d'identifier les défauts est très intéressante afin de compenser leur influence. Cependant, les contraintes d'excitation sur les signaux d'entrée sont en général importantes [Jiang, 1997], [Weber, 1999a, 1999c].

- **Adaptation** : Les techniques d'adaptation ont été appliquées aussi bien aux équations de parité [Frisk, 1997] qu'aux observateurs. Elles permettent de compenser l'influence de certaines perturbations (découplage indirect). Limiter le nombre de paramètres adaptés est un moyen pour lever partiellement les contraintes sur l'excitation du système [Höfling, 1996].
- **Identification** : L'estimation de paramètres appliquée au diagnostic [Isermann, 1993], [Trigeassou, 1996] consiste en général à évaluer la cohérence entre le système physique et un modèle dans l'espace des paramètres du modèle. Bien souvent, les paramètres estimés ne sont pas homogènes aux paramètres physiques. Lorsque la fonction permettant de remonter aux paramètres physiques peut être établie, la tâche de diagnostic est facilitée [Isermann, 1997b].

Il existe bien évidemment des recoupements entre les différentes approches. Ainsi, de nombreux travaux portent sur les liens existant entre les méthodes de génération de résidus. Les liens entre les équations de parité et les observateurs ont été largement étudiés, entre autres, par [Patton, 1991b], par [Cocquempot, 1993] et, plus récemment, par [Ding, 1999]. Les liens entre équations de parité et estimation paramétrique sont étudiés dans [Gertler, 1995]. D'autres études comparatives portent plus sur les performances des méthodes en fonction du contexte d'application [Jiang, 1997], [Isermann, 1994]. Afin de générer des résidus aussi pertinents que possible, l'optimisation de la génération de résidus fait également l'objet d'études : il s'agit bien souvent d'optimiser des critères de sensibilité [Kinnaert, 1994], [Frank, 1994b], [Ding, 1999]. Toujours dans cette optique, les techniques de robustesse sont de plus en plus utilisées afin de gérer au mieux la présence des incertitudes [Patton, 1997b], [Sadriani, 1997].

1.3.6. Décision

Une introduction aux méthodes d'aide à la décision pour le diagnostic est donnée dans [Leonhardt, 1997]. On peut distinguer les méthodes de classification qui consistent à construire des domaines permettant de discriminer plusieurs classes (de défaut) et les méthodes utilisant une base de règles et un mécanisme d'inférence pour construire la décision. Notons que certaines techniques (neuro-floues par exemple, [Jang, 1997]) permettent parfois de combiner les deux approches.

1.3.6.1. Classification par coalescence

La finalité d'un processus de reconnaissance de formes (classique ou flou) est de pouvoir décider, après avoir observé un objet, à quelle forme typique celui-ci est similaire ou, en d'autres termes, à quelle classe d'objets connus il peut être associé [Masson, 1996], [Dubuisson 1990]. Comme l'indique [Leonhardt, 1997], la classification peut être fondée sur une approche statistique (telle que celles résultant du théorème de Bayes), sur une approche polynomiale ou encore géométrique. Dans ce dernier cas, des algorithmes de coalescence, dont l'un des plus connus est celui des k plus proches voisins, peuvent être utilisés [Frélicot, 1992, 1996]. Les réseaux de neurones sont des outils couramment utilisés en classification : de nombreuses structures de réseaux telles que les cartes de Kohonen ont été développées pour résoudre des problèmes de reconnaissance de formes [Hérault, 1994].

Remarque : la définition d'une classe de défauts par un domaine sur l'espace des observations (qui définit une sorte de « masque ») peut être interprétée comme l'élimination du reste de cet espace. La phase d'apprentissage des classes combine à cette élimination une technique d'estimation (adaptation des domaines).

I.3.6.2. Inférences à base de règles

Les systèmes experts constituent une approche possible pour le diagnostic, en particulier lorsque les observations sont de type symbolique¹. Le principe de fonctionnement d'un système expert s'articule autour de deux éléments principaux : une base de connaissance (base de faits et base de règles) et un moteur d'inférence [Farreny, 1989]. Les observations sont incluses dans la base de faits et les règles traduisent la connaissance sur le système (modèle). Le moteur d'inférence établit de nouveaux faits à partir des faits et règles déjà disponibles et essaye de résoudre les conflits pour proposer une solution au problème spécifié. L'application de systèmes experts simples au diagnostic en ligne soulève néanmoins des problèmes comme la prise en compte du temps-réel, la gestion des informations de manière dynamique et l'évaluation des performances [Gerlinger, 1991].

Ces difficultés justifient l'utilisation en ligne de structures souvent fondées sur des graphes ou des arbres de décision [Chang, 1991], [Mo, 1997]. L'utilisation de bases de règles floues [Frank, 1994a], [Evsukoff, 1998b] ou d'arbres de décision flous [Janikow, 1998], [Isermann, 1998] a également fait l'objet de nombreux développements. Des connaissances analytiques telles que la persistance des signaux, ou des connaissances plus heuristiques telles que la confiance, peuvent être intégrées à la décision [Theilliol, 1997].

Une table de décision peut également être utilisée pour décrire l'information servant à localiser les défauts. En supposant que n défauts doivent être localisés à partir de m résidus, une table de décision booléenne est une matrice S de m lignes et de n colonnes telle que $S_{ij}=1$ si le résidu r_i est affecté par le défaut f_j et $S_{ij}=0$ sinon. Une table de décision est dite « localisante » si toutes les colonnes sont non nulles et distinctes deux à deux [Gertler, 1991]. Sous certaines conditions, une table de décision peut être utilisée pour localiser des défauts multiples [Weber, 1999b].

Dans la mesure où les comportements dynamiques induits par les défauts sur les résidus sont assez mal connus, la décision est très souvent considérée comme un processus statique [Isermann, 1997b]. Certaines études considèrent néanmoins la décision comme un processus dynamique (ou séquentiel) : [Koscielny, 1995] propose ainsi une table des états dynamiques (sorte de table de signature dynamique) ; un réseau de Petri permet de surveiller des contraintes de séquences [Valette, 1994], [Dunlop, 1997] ; la reconnaissance de chroniques pour le diagnostic fait partie des méthodes d'aide à la décision mettant en œuvre un automate [Ramkumar, 1999], [Bibas, 1997].

En dehors des systèmes experts, les mécanismes d'inférence que nous avons évoqués (arbre de décision, tables de décision,...) reviennent en général à associer un domaine de l'espace des observations à chaque classe de défaut, et cela sans contre-réaction possible. Ces méthodes reposent donc sur un découplage direct (élimination) de certains comportements observés. Dès lors, la seule différence avec les méthodes de classification indiquées dans le paragraphe I.3.6.1 tient à la façon de construire les domaines associés aux classes ; la remarque terminant le paragraphe I.3.6.1 continue à s'appliquer.

¹ Ou lorsque les observations sont ramenées à une information de type symbolique par une interface appropriée.

1.3.7. Problèmes transversaux

1.3.7.1. Placement de capteurs

Le choix des capteurs et de la nature des observations influe très fortement sur les possibilités de surveillance d'un système. Tout comme pour un problème de commande, le placement des capteurs est la première étape avant tout autre développement relevant de l'Automatique. Les contraintes à cette étape sont d'ordre techniques et économiques : quelles sont les grandeurs physiques pertinentes ?, existe-t-il un capteur permettant de mesurer ces grandeurs ?, quel est le coût de l'instrumentation par rapport aux performances attendues ? Alors que le choix des capteurs peut sembler assez « naturel » sur des systèmes décrits par un faible nombre de variables, il n'en va pas de même sur des installations complexes (Chimie, nucléaire, ...) [Maquin, 1986], [Luong, 1995].

1.3.7.2. Prise en compte des incertitudes

La prise en compte des incertitudes est un problème transversal qui intervient aussi bien avec des modèles fonctionnels qu'avec des modèles comportementaux, aussi bien lors de la génération de résidus que lors de la décision. La prise en compte des incertitudes conduit à l'utilisation de modèles ou de grandeurs non déterministes :

- **Les approches statistiques :** les filtres de Kalman permettent d'introduire des incertitudes additives sur l'état et les mesures d'un modèle d'état. Par ailleurs, le théorème de Bayes est utilisé en classification (réseaux bayésiens, par exemple).
- **Les approches par les bornes :** la technique la plus simple relevant de cette approche est le seuillage. L'application des techniques de robustesse à la génération des résidus est un autre exemple. D'autres techniques font appel à l'arithmétique des intervalles ou généralisent la notion d'intervalle au cas multi-dimensionnel [Ploix, 1998].
- **Les approches floues :** la logique floue permet de prendre en compte les incertitudes par le caractère graduel de l'appartenance à un ensemble flou.

1.3.7.3. Apprentissage

Dans la mesure où la connaissance d'un expert (connaissance a priori) n'est pas toujours suffisante pour construire un système de diagnostic, les techniques d'apprentissage ont pour but de pallier ces insuffisances. L'objectif poursuivi est d'apprendre des connaissances à partir d'une base d'apprentissage (ensemble significatif d'observations).

En particulier, les réseaux de neurones font partie des techniques d'apprentissage. Dans le cadre du diagnostic de défauts, ils peuvent être utilisés aussi bien au stade de la génération de résidus qu'au stade de leur évaluation [Köppen-Seliger, 1995]. Si la connaissance a priori sur les défauts est parfois délicate à obtenir, il en est souvent de même pour les jeux de données constituant une base d'apprentissage. Dans une approche neuronale, il est donc important de bien choisir la structure du réseau en rapport avec les objectifs poursuivis, de façon à garantir des temps d'apprentissage et des capacités de généralisation raisonnables. Les techniques neuro-floues [Jang, 1995], [Leonhardt, 1997] visent à intégrer au mieux une connaissance a priori décrite par des techniques floues avec une connaissance apprise à l'aide des techniques neuronales. Suivant les structures employées, un des intérêts des techniques neuro-floues est également de garder une capacité d'interprétation des règles apprises.

I.3.8. Bilan sur les différentes méthodes de diagnostic de défauts

Nous venons d'évoquer un certain nombre de méthodes de diagnostic. La présentation que nous en avons faite est plutôt axée sur le diagnostic des défauts. Au-delà de la diversité des méthodes et des techniques envisageables, on peut distinguer :

- les approches « bon fonctionnement » et les approches « mauvais fonctionnement »,
- le type de découplage : direct ou indirect.

Chaque méthode de diagnostic, prise indépendamment, permet de construire un module de diagnostic élémentaire. L'association en série et en parallèle de tels modules permet de construire un diagnostic global.

I.4. Cadre de l'étude

Notre étude porte sur le diagnostic en ligne de défauts dans les entraînements électriques. Les modèles que nous utiliserons seront donc soit des modèles structurels, soit des modèles comportementaux.

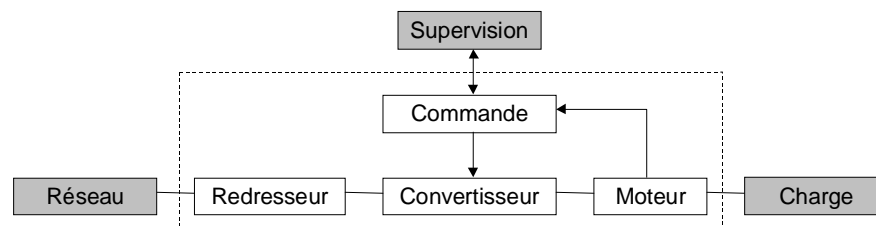


Figure I.13 : Principaux composants d'un entraînement électrique

La structure d'un entraînement électrique est donnée par la Figure I.13. Un entraînement électrique est souvent alimenté en énergie par un réseau électrique. Un redresseur accompagné d'un filtre permet alors de transformer la(les) tension(s) sinusoïdale(s) en une tension continue servant à alimenter en puissance un convertisseur. Cette tension continue peut également être délivrée directement par des batteries. Le convertisseur est un dispositif d'électronique de puissance servant à transférer au moteur l'énergie électrique nécessaire pour entraîner la charge, en fonction des signaux calculés par la commande. La commande échange des informations avec les niveaux de supervision supérieurs.

Le diagnostic que nous cherchons à construire doit satisfaire les impératifs suivants :

- **Indépendance vis-à-vis de la charge** : ne disposant d'aucune connaissance sur la charge, le couple résistant est a priori inconnu. Le diagnostic devra donc être robuste aux variations du couple résistant (ou couple de charge). L'indépendance vis-à-vis de la charge permet de garantir que le système de diagnostic développé pourra être repris à l'identique quelle que soit l'application à laquelle est destiné l'entraînement électrique.
- **Pas d'ajout de capteurs autres que ceux présents pour la commande** : cette contrainte est destinée à limiter le surcoût lié à l'ajout des fonctionnalités de diagnostic. Les problèmes liés au placement des capteurs ne seront donc pas abordés dans notre étude. Combinée à

l'indépendance vis-à-vis de la charge, la contrainte sur l'ajout des capteurs permet d'envisager des variateurs de vitesse intégrant des fonctionnalités de diagnostic avec un surcoût limité.

- **Diagnostic en ligne** : cette contrainte se traduit sur la complexité des algorithmes de diagnostic qui devra rester « raisonnable ». Notons que d'autres stratégies de diagnostic seraient envisageables : les mesures pourraient être exploitées en temps différé (diagnostic hors ligne) ; des algorithmes beaucoup plus complexes que pour le diagnostic en ligne pourraient alors être envisagés.

La validation des méthodes de diagnostic étudiées portent sur deux types d'entraînements électriques. Le premier est construit autour d'un Moteur à Courant Continu (MCC) ; le second est construit autour d'un Moteur ASynchrone (MAS).

Ce document est organisé comme suit :

- **Chapitre 2 (modèle structurel)** : La connaissance « structurelle » et causale qui est utilisée dans ce chapitre permet de guider la décomposition de la tâche de diagnostic. Un premier jeu de résidus peut être déduit de la décomposition d'un modèle global en modèles dits locaux. Les propriétés de ces résidus sont étudiées.
- **Chapitres 3 et 4 (modèle comportemental et génération de résidus)** : le troisième chapitre porte sur la génération de résidus en mettant l'accent sur des techniques de découplage direct pour assurer la robustesse vis-à-vis des perturbations et pour localiser les défauts. Une étude comparative des sensibilités de plusieurs générateurs de résidus est également menée. Le quatrième chapitre porte également sur la génération de résidus mais en mettant plutôt l'accent sur des techniques de découplage indirect (estimation).
- **Chapitres 5 et 6 (modèle comportemental et décision)** : la décision est considérée comme un processus statique dans le cinquième chapitre. L'étude de sensibilité menée dans le troisième chapitre est réutilisée pour optimiser un modèle logique de décision. Nous verrons également que la prise en compte des signes ou, plus largement, des ordres de grandeurs relatifs des résidus permet d'améliorer les performances de la localisation. Le sixième chapitre montre l'intérêt de considérer la décision comme un processus dynamique. La reconnaissance de séquences dans un contexte flou et l'apprentissage d'un modèle de décision à l'aide d'un réseau de neurones récurrent sont étudiés.
- **Conclusion** : les principaux résultats de ce mémoire ainsi que leurs limites sont synthétisés et organisés en adoptant un point de vue lié aux applications d'une part, un point de vue méthodologique sur le diagnostic d'autre part. Enfin, nous dressons un panorama des perspectives qui nous semblent se dégager de ce travail.

Chapitre II :

Analyse causale

Chapitre II :

Analyse causale

II.1. Introduction

L'analyse de la causalité dans les systèmes physiques a donné lieu à plusieurs approches dont les plus connues ont été proposées par [Iwasaki, 1986] et [De Kleer, 1986]. Iwasaki et Simon déduisent l'ordre causal à partir d'une analyse purement structurelle d'un modèle comportemental (décrit par un ensemble d'équations). Avec cette approche, il n'est pas nécessaire de résoudre les équations, contrairement à l'approche proposée par De Kleer et Brown qui consiste à déterminer les chemins de propagation suivis par les perturbations sur les variables en réalisant une propagation de contraintes. De façon plus pragmatique, des modèles causaux peuvent être obtenus à partir d'une description explicite de la causalité, telle qu'elle est perçue par des experts du domaine [Leyval, 1992, 1994].

L'idée directrice de ce chapitre est d'utiliser une connaissance liée à la causalité pour guider la conception d'un système de diagnostic [Montmain, 1992]. Les entraînements électriques étant des systèmes bouclés et peu instrumentés, une description explicite de la causalité selon le point de vue d'un expert se révèle difficile à obtenir [Combastel, 1997a]. Cependant, des modèles de connaissance relativement précis sont disponibles. Compte tenu de nos objectifs et de la connaissance disponible, notre point de vue sur la causalité est celui d'Iwasaki et Simon : la notion de causalité est ainsi assimilée à la notion de calculabilité et repose sur une analyse structurelle. Partant d'un modèle similaire à celui de l'approche structurelle développée au LAIL¹ pour la surveillance des systèmes [Staroswiecki, 1996], nous allons proposer une exploitation sensiblement différente, qui correspond à celle développée au LAG sur les graphes causaux [Montmain, 1992].

II.2. Définition d'un modèle « structurel »

Nous appellerons « modèle structurel » une abstraction des contraintes d'un modèle comportemental². Un tel modèle décrit l'existence ou l'absence de relations entre variables. Les

¹ Laboratoire d'Automatique et d'Informatique Industrielle de Lille.

² D'autres définitions sont possibles. En particulier, la notion de « structure » se réfère souvent à la manière dont sont assemblés les composants physiques d'un système.

objets élémentaires servant à décrire un modèle structurel sont des variables et des relations entre variables (équations). Les variables et les équations sont repérées par un nom mais leur contenu n'a pas besoin d'être spécifié : c'est en effet l'existence et non le contenu de tels objets qui caractérise la connaissance structurelle. Une affirmation est un doublet (*équation, variable*) et signifie que la variable « *variable* » intervient dans l'équation « *équation* ». Une matrice d'incidence indique, parmi toutes les affirmations possibles, celles caractérisant un modèle structurel particulier. Soit $L_E = \{\dots E_i \dots\}$, $i=1\dots n_E$, un ensemble de n_E équations et $L_V = \{\dots V_j \dots\}$, $j=1\dots n_V$, un ensemble de n_V variables. Une matrice d'incidence S de dimension (n_E, n_V) revient à définir un graphe biparti (L_E, L_V, S) ; elle est définie comme suit :

$$\forall (E_i, V_j) \in L_E \times L_V, S_{ij} = \text{'X'}, \text{ si la variable } V_j \text{ intervient dans l'équation } E_i. \quad (\text{II.1})$$

$$= \text{' '}, \text{ sinon.}$$

A titre d'exemple, nous considérons un entraînement électrique à courant continu (commandé en courant) dont un modèle comportemental est donné par (II.2) à (II.10). De nombreux travaux ont été menés au LEG sur cet entraînement [Berendsen, 1993], [Rostaing, 1997], [Courtine, 1997a].

$$PI1 : \quad \dot{X}_I = k_I \cdot (I_{ref} - I_m) \quad (\text{II.2})$$

$$PI2 : \quad \beta = k_p \cdot (I_{ref} - I_m) + X_I \quad (\text{II.3})$$

$$AH : \quad U = \beta_m \cdot U_c \quad (\text{II.4})$$

$$Elec : \quad L \cdot \dot{I} + R \cdot I = V - k \cdot \Omega \quad (\text{II.5})$$

$$Méca : \quad J \cdot \dot{\Omega} + \rho \cdot \Omega = kI - d_T \quad (\text{II.6})$$

$$Déf\beta : \quad \beta_m = \beta + f_\beta \quad (\text{II.7})$$

$$DéfI : \quad I_m = I + f_I \quad (\text{II.8})$$

$$Déf\Omega : \quad \Omega_m = \Omega + f_\Omega \quad (\text{II.9})$$

$$DéfU : \quad V = U + f_U \quad (\text{II.10})$$

I_{ref} est la consigne de courant, I est le courant moteur, β est l'action calculée par un correcteur proportionnel intégral (PI1 et PI2), X_I correspond au terme intégral de l'action. k_p , k_I sont les coefficients proportionnel et intégral du correcteur PI. U_c est la tension filtrée fournie par l'alimentation constituée d'un redresseur, d'un filtre et d'un dispositif de freinage. U est la tension moyenne (sur la période de MLI¹) qui serait appliquée aux bornes du moteur en l'absence d'un défaut modélisé par f_U dans l'ensemble constitué par l'alimentation et le hacheur (convertisseur). V est la tension (moyenne) effectivement appliquée aux bornes du moteur. Ω est la vitesse de rotation du moteur. L , R , k , J , ρ sont les paramètres électriques et mécaniques du moteur : inductance, résistance, constante de couple, inertie, coefficient de frottement. d_T est le couple résistant. *Elec* représente l'équation électrique et *Méca* l'équation mécanique. L'indice m fait référence à une grandeur mesurée. f_β modélise un défaut de l'interface de commande, f_I modélise un défaut du capteur de courant et f_Ω modélise un défaut du capteur de vitesse.

L'approche structurelle (LAIL) de même que les graphes causaux (LAG) pour la surveillance des systèmes sont des approches du diagnostic orientées bon fonctionnement. Il n'est donc pas a priori nécessaire de modéliser l'influence des défauts. Nous avons néanmoins introduit une modélisation des défauts ((II.7) à (II.10)) pour deux raisons :

¹ MLI : Modulation de Largeur d'Impulsion

1. Le principe d'ordonnement causal et de décomposition de modèle n'est en rien modifié.
2. La structure correspondante est le reflet exact du modèle utilisé ultérieurement pour générer des résidus par des approches qui utilisent d'autres connaissances que la seule connaissance structurelle. Des éléments de comparaison pourront donc être plus aisément mis en évidence, en particulier concernant la capacité de localisation des défauts considérés.

	<i>Relations entre variables</i>
<i>Correcteur PI</i>	$\dot{X}_I = PI1(I_{ref}, I_m)$ $PI2(\beta, I_{ref}, I_m, X_I) = 0$
<i>Alimentation + Hacheur</i>	$AH(U, \beta_m) = 0$
<i>Equation électrique</i>	$\dot{I} = Elec(I, V, \Omega)$
<i>Equation mécanique</i>	$\dot{\Omega} = Méca(\Omega, I, d_\Gamma)$
<i>Modèles de défaut</i>	$Déf\beta(\beta, f_\beta, \beta_m) = 0$ $DéfI(I, f_I, I_m) = 0$ $Déf\Omega(\Omega, f_\Omega, \Omega_m) = 0$ $DéfU(U, f_U, V) = 0$

Table II.1 : Connaissance structurelle

La connaissance structurelle (Table II.1) qui correspond à l'abstraction des relations du modèle comportemental décrit par (II.2) à (II.10) peut également être représentée à l'aide de la matrice d'incidence S (Table II.2).

S	I_{ref}	d_Γ	X_I	β	f_β	β_m	I	f_I	I_m	Ω	f_Ω	Ω_m	U	f_U	V
$PI1$	X		X						X						
$PI2$	X		X	X					X						
AH						X							X		
$Elec$							X			X					X
$Méca$		X					X			X					
$Déf\beta$				X	X	X									
$DéfI$							X	X	X						
$Déf\Omega$										X	X	X			
$DéfU$													X	X	X
ENV	X	X			X			X			X			X	

Table II.2 : Matrice d'incidence définissant la structure de la MCC.

Le comportement du système étudié est décrit par l'ensemble des variables, L_V , et par l'ensemble des équations, L_E . Les variables exogènes, L_{Vex} , sont imposées par l'environnement du système (ENV). Les variables différentielles, L_{Vdif} , sont telles que leur dérivée intervient dans les équations différentielles, L_{Edif} .

$$L_E = \{PI1, PI2, AH, Elec, Méca, Déf\beta, DéfI, Déf\Omega, DéfU\} \quad (II.11)$$

$$L_{Edif} = \{PI1, Elec, Méca\} \quad (II.12)$$

$$L_V = \{I_{ref}, d_\Gamma, X_I, \beta, f_\beta, \beta_m, I, f_I, I_m, \Omega, f_\Omega, \Omega_m, U, f_U, V\} \quad (II.13)$$

$$L_{Vex} = \{I_{ref}, d_\Gamma, f_\beta, f_I, f_\Omega, f_U\} \quad (II.14)$$

$$L_{Vdif} = \{X_I, I, \Omega\} \quad (II.15)$$

Dans la Table II.2, les dérivées des variables différentielles n'apparaissent pas : une colonne unique est associée à une variable différentielle (on pourrait aussi imaginer une colonne pour la variable elle-même, une autre colonne pour sa dérivée première,...) ; nous justifierons ce choix par la suite. La connaissance structurelle va désormais être utilisée pour établir un ordonnancement causal des relations entre variables.

II.3. Ordonnancement causal : obtention d'un schéma-bloc.

Partant d'un modèle structurel, nous cherchons à établir un schéma-bloc relatif au système étudié. L'application à la MCC permet d'illustrer la démarche. Cette partie et, plus généralement, ce chapitre, présentent essentiellement un intérêt méthodologique : les entraînements électriques étant des systèmes faisant intervenir peu de variables dans leurs modèles¹, les schémas-blocs peuvent en général être obtenus intuitivement. Néanmoins, une démarche systématique présente l'intérêt de pouvoir être appliquée telle quelle à des systèmes dont les modèles font intervenir un nombre d'équations et de variables beaucoup plus grand.

II.3.1. Objectif de l'ordonnancement causal

Un ordonnancement causal ajoute au graphe biparti associé à un modèle structurel, la notion d'orientation des arcs. La causalité étant assimilée à la calculabilité, l'orientation des arcs coïncide avec un ordonnancement des calculs permettant la simulation du modèle comportemental associé. Ainsi, l'ordonnancement causal utilise la connaissance structurelle pour aboutir à un schéma-bloc. Un schéma-bloc est en effet une représentation possible pour un graphe causal biparti où les variables sont associées aux arcs orientés et où les relations (équations) sont associées aux nœuds représentés par des « boîtes » rectangulaires (Figure II.1).

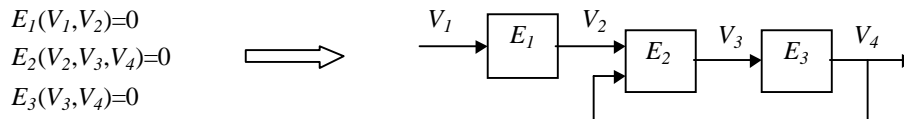


Figure II.1 : Ordonnancement causal (exemple)

Formellement, une manière d'introduire l'orientation des arcs consiste à considérer que les affirmations définissant un modèle structurel causal ne sont plus des doublets (*équation, variable*) mais sont du type (*équation*→*variable*) ou (*variable*→*équation*). La matrice d'incidence S^c définissant un modèle structurel causal n'est donc plus binaire mais ternaire :

$$\forall (E_i, V_j) \in L_E \times L_V, S^c_{ij} = \begin{cases} ' ', & \text{s'il n'y a aucun lien entre } E_i \text{ et } V_j. \\ ' \rightarrow ', & \text{si la variable } V_j \text{ est calculée par l'équation } E_i. \\ & \text{autrement dit, si } V_j \text{ est une sortie du "bloc" } E_i. \\ ' \leftarrow ', & \text{si } V_j \text{ doit être connue pour utiliser } E_i \text{ dans un calcul,} \\ & \text{autrement dit, si } V_j \text{ est une entrée du "bloc" } E_i. \end{cases} \quad (\text{II.16})$$

¹ En tout cas, dans les modèles destinés à réaliser un diagnostic en ligne.

‘ \rightarrow ’ correspond à une orientation *équation*→*variable* (ligne vers colonne de la matrice d’incidence) alors que ‘ \leftarrow ’ correspond à une orientation *variable*→*équation* (colonne vers ligne de la matrice d’incidence). Le graphe (L_E, L_V, S^c) est un graphe causal dont une représentation graphique possible est un schéma-bloc. (II.16) met clairement en évidence le fait que la causalité est assimilée à la calculabilité. Dit simplement, un algorithme d’ordonnement causal consiste à remplacer les ‘X’ de la matrice d’incidence binaire initiale soit par ‘ \rightarrow ’, soit par ‘ \leftarrow ’, en fonction de la calculabilité.

II.3.2. Quelques définitions et résultats : rappels et illustrations

Le point de vue que nous avons retenu sur l’ordonnement causal est celui décrit par Travé-Massuyès et Pons dans [Pons, 1997]. Leur approche, développée pour résoudre des problèmes de prédiction et de changement de mode de fonctionnement, étend celle proposée par Iwasaki et Simon. Nous rappelons tout d’abord quelques définitions et propositions (les démonstrations sont données ou référencées dans [Pons, 1997]).

Définition 1 : Auto-contenance (Self-Containment) : un système statique de n équations algébriques avec n variables est auto-contenu (*self-contained*) si tout sous-ensemble de k ($k \leq n$) équations contient au moins k variables.

Définition 2 : Sous-système complet minimal (SCM) : étant donné un système¹ auto-contenu Σ , un sous-système σ de Σ qui est également auto-contenu et qui ne contient pas de sous-système auto-contenu est un sous-système complet minimal (SCM).

Remarque 1 : L’existence d’un SCM de plus d’une variable se traduit par la présence d’une « boucle », toutes les variables d’un tel SCM étant mutuellement dépendantes. Une boucle correspond à une situation où plusieurs interprétations de la causalité sont possibles. Toutes les interprétations de la causalité dans les boucles étant équivalentes du point de vue de la prédiction (calcul de variables inconnues), la solution proposée dans [Pons, 1997] consiste à effectuer un choix arbitraire. Considérons par exemple un système de deux équations (f, g) avec deux variables non exogènes (x, y) :

$$\begin{aligned} f: x+y &= a \\ g: 2x+y &= b \end{aligned} \quad (\text{II.17})$$

(II.17) est un système de deux équations à deux variables qui est auto-contenu. Il s’agit également d’un SCM à deux variables et les deux variables sont mutuellement dépendantes. La causalité (calculabilité) peut être orientée différemment suivant la manière dont le système est résolu : calcul de x puis déduction de y comme une fonction de x ou l’inverse. Si l’objectif est seulement de calculer x et y , l’orientation de la causalité peut être choisie arbitrairement : la solution du système est identique dans les deux cas. Cela permet de lever une des limitations de l’approche proposée par Iwasaki et Simon pour résoudre des problèmes de prédiction en présence de boucles.

Reformulation des objectifs d’un algorithme d’ordonnement causal : étant donné un système auto-contenu défini par un ensemble de n équations (L_E) reliant les n variables à calculer ($L_V \setminus L_{Vex}$)² et étant donné l’environnement (ENV) du système représenté par l’ensemble des

¹ Le terme « système » correspond ici à un système d’équation, donc à un modèle, et non à un système physique.

² Les variables à calculer sont les variables non exogènes : L_V à l’exclusion de L_{Vex} est noté $L_V \setminus L_{Vex}$.

variables exogènes, L_{Vex} , trouver un ordonnancement causal consiste à déterminer les chemins de dépendance entre les variables. Ces chemins indiquent dans quel ordre chaque équation doit être utilisée pour calculer successivement les n variables supposées inconnues. Le système étant auto-contenu, le problème d'ordonnancement causal se ramène à associer une variable à chaque équation (la variable en question est celle qui sera calculée à partir de l'équation considérée : voir '↔' dans (II.16)), c'est-à-dire à trouver un couplage parfait dans le graphe biparti défini par $(L_E, L_V \setminus L_{Vex}, S(L_E, L_V \setminus L_{Vex}))$. $S(L_E, L_V \setminus L_{Vex})$ est la restriction de la matrice d'incidence binaire S aux colonnes associées aux variables non exogènes. En adoptant le même principe de notation, on a : $S = S(L_E, L_V)$.

Existence d'un couplage parfait : une condition nécessaire et suffisante d'existence d'un couplage parfait sur le graphe biparti $(L_E, L_V \setminus L_{Vex}, S(L_E, L_V \setminus L_{Vex}))$ associé à un système est que ce dernier soit auto-contenu (référence à [Berge, 1957] dans [Staroswiecki, 1996]).

Proposition 1 : chaque SCM d'ordre au moins égal à deux correspond à une composante fortement connexe¹ du graphe biparti orienté (graphe causal).

Proposition 2 : un système décrit par $(L_E, L_V \setminus L_{Vex})$ ne contient pas de SCM de plus d'une variable si et seulement si son graphe biparti $(L_E, L_V \setminus L_{Vex}, S(L_E, L_V \setminus L_{Vex}))$ admet un couplage parfait unique.

Proposition 3 : si le graphe biparti $(L_E, L_V \setminus L_{Vex}, S(L_E, L_V \setminus L_{Vex}))$ admet un couplage parfait unique, le graphe causal correspondant est acyclique, de même que les chemins de dépendance causale reliant toutes les variables.

Pour obtenir un couplage parfait, une démarche possible consiste à permuter les lignes (par exemple) de la matrice d'incidence de manière à maximiser le nombre de 'X' sur la diagonale. Si toute la diagonale est formée de 'X' après la permutation des lignes, le couplage parfait obtenu entre les équations et les variables est donné par les couples (équation, variable) associés aux termes diagonaux. L'exemple de la Figure II.1 est maintenant repris plus en détail.

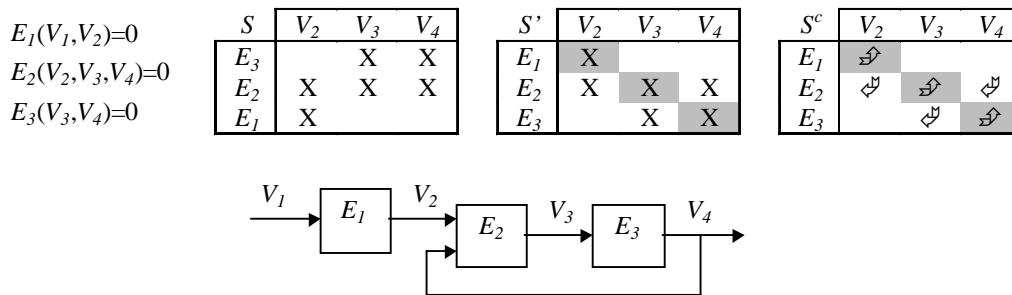


Figure II.2 : Recherche d'un couplage parfait et déduction de l'ordonnancement causal

La structure du système constitué des trois équations E_1, E_2, E_3 et des trois variables non exogènes V_2, V_3, V_4 est représentée par la matrice d'incidence S . Après permutation des lignes 1 et 3, on obtient S' dont tous les éléments diagonaux sont des 'X'. La diagonale de S' définit donc un couplage parfait. La signification d'un couple (équation, variable) appartenant au couplage parfait est que l'équation « équation » peut être utilisée pour calculer la variable « variable ».

¹ Une composante d'un graphe est fortement connexe si, pour tout couple de nœuds n_1 et n_2 appartenant à cette composante, il existe un chemin reliant n_1 et n_2 .

Pour calculer cette variable, il est nécessaire de connaître les autres variables intervenant dans l'équation « *équation* ». Ainsi, la matrice d'incidence causale S^c se déduit en affectant '⇒' aux éléments diagonaux de S' et en affectant '⇐' aux éléments non vides et non diagonaux de S' . Le schéma bloc qui correspond à l'une des représentations possibles d'un graphe causal se déduit directement de S^c . Par ailleurs, le système de la Figure II.2 possède un SCM de deux variables (V_3 et V_4) qui correspond à une composante fortement connexe du graphe causal et qui se traduit par la présence d'une boucle dans le schéma-bloc (liée au terme sur-diagonal de S^c). La présence d'un SCM de plus d'une variable indique que le couplage parfait obtenu n'est pas unique (Proposition 2), chaque couplage parfait correspondant à une interprétation possible de la causalité dans la(les) boucle(s).

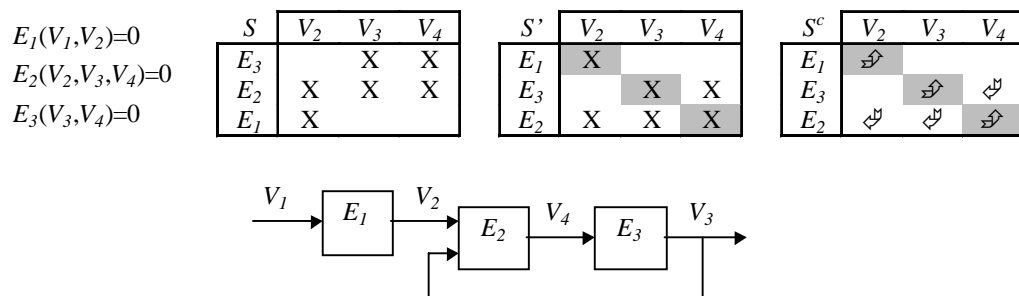


Figure II.3 :une autre interprétation possible de la causalité

Une autre interprétation possible de la causalité est donnée par la Figure II.3. Ainsi, E_3 ne sert plus à calculer V_4 à partir de V_3 mais V_3 à partir V_4 . Ce changement d'orientation de la causalité (calculabilité) ne pose pas de problème pour des relations d'équilibre simples telles que la loi d'Ohm : on peut aussi bien écrire $U=R.I$ que $I=U/R$. Dans le cas de systèmes dynamiques modélisés par des équations différentielles, un tel changement d'orientation peut conduire à des difficultés pratiques. Nous allons maintenant traiter un exemple de système dynamique afin d'illustrer les choix retenus pour la construction d'un ordonnancement causal.

II.3.3. Cas des systèmes dynamiques

Les systèmes dynamiques considérés sont modélisés par des équations différentielles (L_{Vdif}) et par des équations algébriques ($L_V \setminus L_{Vdif}$). Une équation différentielle quelconque peut être ramenée à un ensemble d'équations différentielles sous forme canonique, c'est-à-dire avec une seule dérivée qui est le seul terme du membre de gauche : $dV_j/dt = E_i(V_1, \dots, V_j, \dots, V_n, t)$. Nous supposons que les équations différentielles apparaissent sous cette forme dans L_{Vdif} . En considérant une équation différentielle mise sous forme canonique et en traitant dV_j/dt comme une variable à part entière, il existe également une relation entre dV_j/dt et V_j que nous appellerons dyn_j (pour dynamique sur la variable V_j). Notons que des équations aux différences (récurrentes) pourraient être traitées de manière analogue aux équations différentielles.

Système étudié

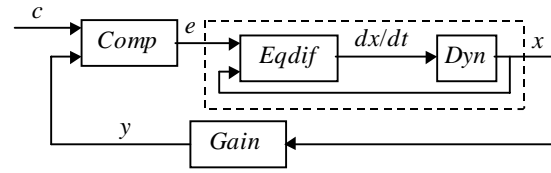
Comp : $e=c-y$

Gain : $y=g.x$

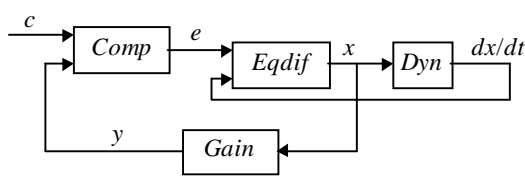
Eqdif : $dx/dt=-x+e$

Dyn : $dyn(dx/dt,x)=0$

	dx/dt	x	e	y
Comp			X ₃	X
Gain		X		X ₄
Eqdif	X ₂	X	X	
Dyn	X	X ₁		



	dx/dt	x	e	y
Comp			X ₃	X
Gain		X		X ₄
Eqdif	X	X ₂	X	
Dyn	X ₁	X		



	dx/dt	x	e	y
Comp			X	X ₄
Gain		X ₃		X
Eqdif	X	X	X ₂	
Dyn	X ₁	X		

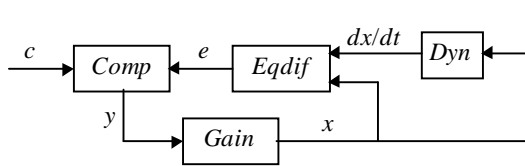


Figure II.4 : Exemple d'ordonnement causal pour un système dynamique

La Figure II.4 représente les ordonnancements causaux possibles pour un système dynamique simple. Les cases grisées correspondent au couplage parfait définissant un ordonnancement causal possible. Les indices numériques dans ces cases donnent, à titre indicatif, l'ordre dans lequel nous avons établi le couplage. Par exemple, dans le cas (a), en supposant que Dyn soit utilisée pour calculer x ('X₁'), dx/dt devra forcément être calculé par Eqdif ('X₂') et ainsi de suite (Figure II.4). Dans les cas (b) et (c), les ordonnancements causaux sont obtenus en considérant que Dyn est utilisée pour calculer dx/dt et non plus x , etc... La Figure II.4 montre que suivant l'ordonnement causal choisi, la relation dynamique Dyn correspond soit à une intégration (a), soit à une dérivation (b), (c).

Le fait d'inclure une étape de dérivation dans une chaîne de calculs peut poser des problèmes pratiques importants, en particulier lorsque des mesures interviennent en amont : bien que le filtrage des bruits de mesure soit possible, il n'est pas toujours suffisant pour permettre le calcul numérique d'une dérivée avec la précision souhaitée. Par conséquent, nous allons privilégier systématiquement les relations dynamiques du type « intégration » par rapport aux relations dynamiques du type « dérivation » dans le choix d'un ordonnancement causal.

$$\underline{V}_j = dyn_j(dV_j/dt) \text{ et } dV_j/dt = E_i(V_1, \dots, V_j, \dots, V_n, t) \tag{II.18}$$

$$\underline{V}_j = dyn_j(E_i(V_1, \dots, V_j, \dots, V_n, t)) \Leftrightarrow E_i(V_1, \dots, \underline{V}_j, \dots, V_n, t) = 0 \tag{II.19}$$

Faire en sorte que dyn_j soit une relation intégrale revient à imposer que dyn_j calcule V_j à partir de dV_j/dt . dV_j/dt devra donc être calculée par une autre relation. Les équations différentielles étant sous forme canonique, la seule équation permettant de calculer dV_j/dt sans dérivation est l'équation différentielle E_i (II.18). Les zones grisées dans (II.18) correspondent au couplage qu'il est nécessaire d'imposer pour éviter les dérivations. Formellement, cela revient à ne pas considérer les dérivées des variables différentielles comme des variables du système et à forcer systématiquement le couplage entre une équation différentielle et sa variable différentielle (II.19) (voir également la zone délimitée par les pointillés sur la Figure II.4).

Remarque 1 : l'absence des dérivées des variables différentielles X_I , I , et Ω dans la matrice d'incidence définissant la structure de la MCC (Table II.2) est justifiée par le choix d'orienter les relations dynamiques dans le sens d'une intégration.

Remarque 2 : à titre illustratif, nous indiquons quelques analogies avec les blocs du logiciel Simulink : la zone délimitée par les pointillés sur la Figure II.4 pourrait correspondre à un bloc « transfert continu ». Le paramétrage d'un tel bloc permet de définir $Eqdif$, et dyn est complètement transparent du point de vue de l'utilisateur : dyn correspond alors à l'intégrateur d'équations différentielles utilisé en interne par Simulink (Runge Kutta, ...). Par contre, en l'absence de bloc dédié, si l'on souhaite simuler une équation différentielle non linéaire, une solution consiste à définir un ou des bloc(s) effectuant le calcul de la dérivée ($Eqdif$) et à ajouter un bloc intégrateur (dyn) qui apparaît alors explicitement. La simulation de modèles non linéaires comme celui d'une machine asynchrone ou la simulation de systèmes à paramètres variant dans le temps (LTV) peuvent être implantées de cette façon.

Une question reste encore en suspens : comment forcer le couplage entre une équation différentielle et sa variable différentielle ? Lorsqu'une équation E et une variable V sont couplées, premièrement : E étant affectée au calcul de V , E ne peut plus être utilisée pour calculer une autre variable, deuxièmement : V étant calculée par E , il n'est plus nécessaire de trouver une autre équation servant à calculer V . Par conséquent, forcer le couplage entre E et V s'obtient en retirant de la matrice d'incidence la ligne associée à E et la colonne associée à V , avant d'appliquer l'algorithme d'ordonnement causal.

Conclusion : Soit $L_E' = L_E \setminus L_{Edif}$ l'ensemble des équations non différentielles (équations algébriques) et $L_V' = L_V \setminus (L_{Vex} \cup L_{Vdif})$ l'ensemble des variables à la fois non exogènes et non différentielles. Un couplage parfait sur le graphe biparti (L_E', L_V', S') où $S' = S(L_E', L_V')$ définit un ordonnancement causal sans relation dynamique du type dérivation.

II.3.4. Algorithme utilisé pour l'ordonnement causal

Deux algorithmes reposant sur la théorie des graphes ont été utilisés en cascade pour obtenir un ordonnancement causal du graphe biparti (L_E', L_V', S') .

Le premier algorithme renvoie une permutation d'une matrice creuse quelconque de façon à obtenir un maximum d'éléments non nuls sur la diagonale [Duff, 1981a, 1981b]. La matrice creuse d'entrée, la matrice d'incidence, est représentée sous la forme d'un graphe biparti pour construire l'algorithme. Suite à l'application de la permutation P à S' , les éléments diagonaux de la matrice d'incidence S_D correspondent à un couplage parfait et donnent ainsi un ordonnancement causal possible.

$$S_D = P.S' \quad (\text{II.20})$$

Le deuxième algorithme est une implantation de l'algorithme de Tarjan [Tarjan, 1972]. Cet algorithme renvoie une permutation symétrique Q qui, appliquée à une matrice donnée, la transforme en une matrice bloc triangulaire inférieure (BLT) [Duff, 1978a, 1978b].

$$S_{BLT} = Q.S_D.Q^T = \begin{bmatrix} A_{11} & & & 0 \\ A_{21} & A_{22} & & \\ \vdots & & \ddots & \\ A_{n1} & \cdots & \cdots & A_{nn} \end{bmatrix} \quad (\text{II.21})$$

Les blocs diagonaux de S_{BLT} sont des composantes fortement connexes du graphe biparti décrit par S_D qui ne peuvent se mettre sous la forme d'une matrice triangulaire inférieure par blocs. Les blocs diagonaux correspondent aux sous-systèmes complets minimaux (SCM) (II.3.2.

proposition 1). Alors que le premier algorithme fournit un ordonnancement causal possible, le deuxième permet de savoir si cet ordonnancement est unique. En effet, si un système ne contient pas de SCM de plus d'une variable, autrement dit, si S_{BLT} est triangulaire inférieure (et non triangulaire inférieure par blocs), le couplage parfait est unique, donc l'ordonnancement causal l'est aussi (II.3.2. proposition 2). Le graphe causal correspondant est alors acyclique, de même que les chemins de dépendance causale reliant toutes les variables (II.3.2. proposition 3 et Figure II.5).

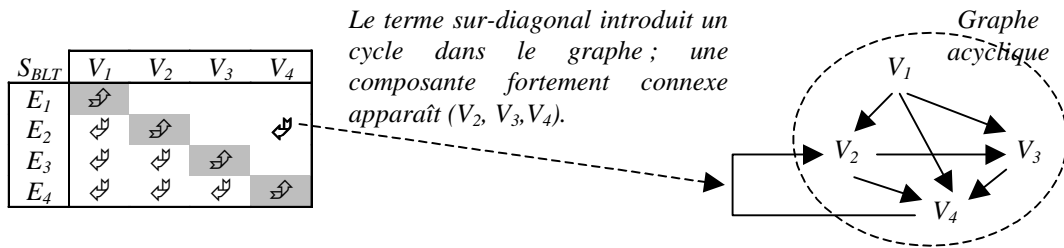


Figure II.5 : Illustration des propriétés d'un graphe causal (graphe orienté)

II.3.5. Application à la Machine à Courant Continu (MCC)

II.3.5.1. Avec forçage des équations différentielles

Partant de la Table II.2 et des ensembles définis par (II.11) à (II.15), l'ordonnancement causal du graphe biparti (L_E', L_V', S') où $L_E' = L_E \setminus L_{Edif}$, $L_V' = L_V \setminus (L_{Vex} \cup L_{Vdif})$ et $S' = S(L_E', L_V')$ donne les résultats de la Table II.3.

$$L_E' = \{DéfI, PI2, Déf\beta, Déf\Omega, AH, DéfU\} \tag{II.22}$$

$$L_V' = \{I_m, \beta, \beta_m, \Omega_m, U, V\} \tag{II.23}$$

S^c	S_{BLT}	Variables non exo. et non diff.					Var. différentielles			Variables exogènes							
		I_m	β	β_m	Ω_m	U	V	X_I	I	Ω	I_{ref}	d_I	f_β	f_I	f_Ω	f_U	
Equations algébriques	DéfI	↗						↗									
	PI2	↘	↗					↘			↘						
	Défβ		↘	↗								↘					
	DéfΩ				↗					↘							↘
	AH			↘		↗											
	DéfU					↘	↗										
Equations diff.	PII	↘						↗			↘						
	Elec						↘		↗	↘							
	Méca							↘	↗		↘						
	ENV										↘	↘	↘	↘	↘	↘	↘

Table II.3 : Ordonnancement causal de la MCC (avec forçage).

De l'ordonnancement causal de la Table II.3, le schéma-bloc de la Figure II.6 peut être déduit.

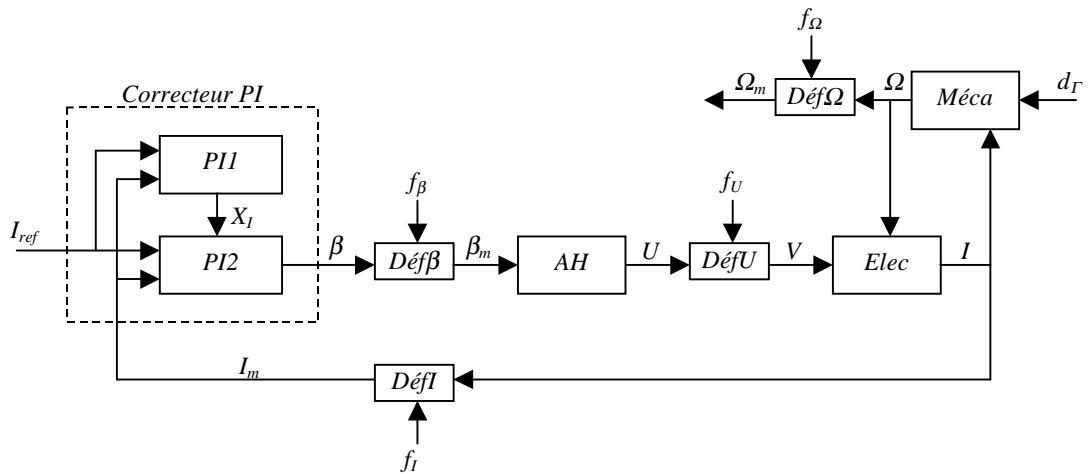


Figure II.6 : Schéma-bloc de la MCC

La matrice d'incidence sur laquelle est effectivement appliqué l'algorithme d'ordonnancement causal est triangulaire inférieure ; elle ne présente pas de blocs de plus d'une variable sur sa diagonale (fond blanc dans la Table II.3). Par conséquent, en privilégiant systématiquement la causalité dans le sens des intégrations (forçage associé aux équations différentielles), l'ordonnancement causal obtenu (Figure II.6) est unique. La présence de cycles dans le graphe causal, autrement dit, la présence de boucles dans le schéma-bloc, est due aux termes sur-diagonaux qui subsistent du fait du forçage des équations différentielles (Table II.3).

II.3.5.2. Sans forçage des équations différentielles

Si on ne cherche pas à imposer une causalité intégrale pour les équations différentielles, un ordonnancement causal du graphe biparti (L_E'', L_V'', S'') , où $L_E'' = L_E \cup L_{Edyn}$, $L_V'' = L_V \setminus L_{Vex}$ et où $S'' = S(L_E'', L_V'')$, est donné par la Table II.4.

$$L_E'' = \{Méca, PII, PI2, DéfI, Elec, DéfU, AH, Déf\beta, Déf\Omega\} \cup \{DynI, Dyn\Omega, DynX_I\} \quad (II.24)$$

$$L_V'' = \{I', \Omega', \Omega, X_I', X_I, I_m, I, V, U, \beta_m, \beta, \Omega_m\} \quad (II.25)$$

S^c	Variables non exogènes											Variables exogènes						
	I'	Ω'	Ω	X_I'	X_I	I_m	I	V	U	β_m	β	Ω_m	I_{ref}	d_Γ	f_β	f_i	f_Ω	f_U
$DynI$	↗						↘											
$Dyn\Omega$		↘	↘															
$Méca$		↘	↘				↘							↘				
PII				↘			↘						↘					
$DynX_I$				↘	↘													
$PI2$					↘	↘					↘		↘					
$DéfI$						↘	↘									↘		
$Elec$	↘		↘				↘	↘										
$DéfU$								↘	↘									↘
AH									↘	↘								
$Déf\beta$										↘	↘				↘			
$Déf\Omega$			↘									↘					↘	
ENV													↘	↘	↘	↘	↘	↘

Table II.4 : Exemple d'ordonnancement causal de la MCC (sans forçage).

La matrice d'incidence fournie par l'algorithme d'ordonnement causal (fond blanc de la Table II.4) n'est pas triangulaire inférieure. Il existe donc plusieurs orientations possibles de la causalité. La Table II.4 décrit un graphe causal dans lequel les relations dynamiques $DynI$ et $Dyn\Omega$ sont orientées dans le sens d'une dérivation alors que la relation $DynX_i$ est orientée dans le sens d'une intégration. En permutant les lignes associées à $Dyn\Omega$ et $Méca$, on obtient une autre interprétation de la causalité pour laquelle la relation $Dyn\Omega$ correspond à une intégration.

II.4. Décomposition de modèle

II.4.1. Introduction, définition des modèles locaux

L'orientation des arcs résultant de l'ordonnement causal est un guide précieux pour générer des séquences de calcul servant à calculer des résidus. « Une relation de redondance analytique est une relation dont la structure ne contient que des variables et des paramètres connus. Elle peut donc être calculée directement. Elle permet de vérifier la cohérence des valeurs des variables et ne peut être vérifiée que lorsque le comportement du système correspond au comportement issu du modèle. Une défaillance sur la mesure d'une variable ou un changement de la valeur d'un paramètre de la relation de redondance analytique entraînera la non-vérification de celle-ci. » [Cassar, 1994a]. Une relation de redondance analytique peut donc être vue comme un modèle dont la simulation est possible et qui permet d'évaluer la cohérence entre un comportement de référence et le comportement du système physique connu au travers de certaines observations. Il s'agit en fait d'un modèle pour le diagnostic tel que nous l'avons décrit dans le premier chapitre.

L'approche structurelle du LAIL repose sur la décomposition canonique du graphe biparti associé à une matrice d'incidence en trois composantes : sous-déterminée, juste déterminée et sur-déterminée, cette dernière permettant justement de générer des relations de redondance analytiques. Il existe une infinité de relations de redondance analytique possibles, en particulier si l'on tient compte de combinaisons linéaires (ou non linéaires...) entre des relations élémentaires. Notre approche concernant le choix de relations de redondance particulières a obéi aux critères suivants :

1. Dans ce chapitre, nous nous interdisons toute combinaison entre relations de redondance analytique autres que celles pouvant être déduites d'une information structurelle et causale.
2. Même si des combinaisons peuvent être envisagées ultérieurement, nous souhaitons aboutir à une décomposition de modèle systématique. En effet, le diagnostic d'une installation dont le modèle fait intervenir un (très) grand nombre de variables¹ peut nécessiter la décomposition du problème de départ en problèmes plus simples, et cela est d'autant plus facile à implanter que la démarche est systématique.

Ces critères nous conduisent à introduire ce que nous appelons les « modèles locaux » : les **modèles locaux** relatifs à un modèle (global) correspondent à une décomposition, à une partition de ce dernier. Un modèle local représente le plus petit sous-système pour lequel la sortie est mesurée et les entrées sont soit mesurées, soit exogènes. Cette définition sous-entend que l'on

¹ Même si les entraînements électriques ne sont pas des systèmes faisant intervenir un grand nombre de variables, notre objectif est ici de proposer une démarche systématique de décomposition de modèle et d'évaluer jusqu'où cette méthodologie peut mener pour diagnostiquer ces systèmes.

soit capable de définir la taille d'un système ; étant donné un ordonnancement causal valide, le nombre d'arcs du sous graphe causal servant à décrire un sous-système (d'équations) sera utilisé pour évaluer la taille de ce sous-système.

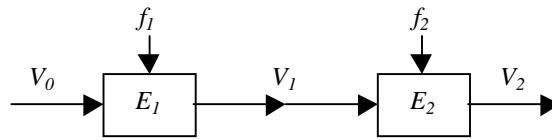


Figure II.7 : Définition des modèles locaux, exemple

La Figure II.7 représente la schéma-bloc (graphe causal) d'un système résultant de l'association en série de deux sous-systèmes décrits respectivement par E_1 et E_2 . V_0 est une variable exogène ; V_1 et V_2 sont des variables parfaitement mesurées (pas de défaut capteur). A chaque variable mesurée, on peut associer un modèle local en « remontant » récursivement dans le graphe causal et en s'arrêtant dès que possible, c'est-à-dire dès que l'on trouve une variable mesurée ou exogène. Ainsi, le modèle local associé à V_1 est $M_1 : V_1 = E_1(V_0, f_1)$, celui associé à V_2 est $M_2 : V_2 = E_2(V_1, f_2)$. Chaque modèle local est de taille plus faible donc de complexité a priori plus faible que le système de départ. En injectant les valeurs de mesure en entrée des modèles locaux et en calculant les erreurs de sortie, on obtient autant de relations de redondance analytique que de modèles locaux :

$$r_1 = V_1 - E_1(V_0, f_1 = 0) \quad (\text{II.26})$$

$$r_2 = V_2 - E_2(V_1, f_2 = 0) \quad (\text{II.27})$$

r_1 (resp. r_2) évalue la cohérence du sous-système décrit par M_1 (resp. M_2) avec le comportement décrit par E_1 (resp. E_2). La modélisation du défaut f_1 (resp. f_2) n'est pas nécessaire ; en effet, la relation de redondance (II.26) (resp. (II.27)) correspond à l'erreur de sortie d'un modèle de bon fonctionnement comme le montre le terme $f_1 = 0$ (resp. $f_2 = 0$). De plus, r_1 (resp. r_2) est sensible aux seuls défauts susceptibles d'affecter le sous-système décrit par M_1 (resp. M_2) : f_1 (resp. f_2). Dans l'exemple de la Figure II.7, les résidus r_1 et r_2 obtenus par analyse causale permettent de localiser f_1 et f_2 . Raisonner sur un sous-système décrit par un modèle local donné revient à éliminer tout le reste du système décrit par un modèle global. Cette élimination permet de découpler les résidus associés aux modèles locaux de façon pertinente en vue de la localisation. Nous allons maintenant reprendre la démarche adoptée pour l'exemple de la Figure II.7 dans le cas d'un système décrit par un graphe biparti causal quelconque. Cette démarche consiste tout d'abord à obtenir les modèles locaux, puis à générer les résidus associés et enfin à étudier leurs propriétés de localisation.

II.4.2. Algorithme de construction des modèles locaux

Une différence importante entre le diagnostic et la commande tient au fait que l'on ne peut pas faire confiance a priori aux mesures. Il est donc très important de distinguer les variables mesurées $L_{V_mé}$ (celles qu'on souhaiterait connaître) et les variables de mesure L_{V_m} (celles qu'on connaît effectivement). Les équations de mesure, L_{E_m} , relient une variable de mesure avec la variable mesurée correspondante et éventuellement des entrées exogènes (servant à modéliser des défauts, par exemple). Les ensembles qui viennent d'être définis sont donnés par (II.28), (II.29) et (II.30) pour la MCC.

$$L_{V_mé} = \{\beta, I, \Omega\} \quad (\text{II.28})$$

$$L_{V_m} = \{\beta_m, I_m, \Omega_m\} \quad (\text{II.29})$$

$$L_{Em} = \{D\acute{e}f\beta, D\acute{e}fI, D\acute{e}f\Omega\} \quad (II.30)$$

Chaque variable de mesure fournit une redondance d'information par rapport à la simulation d'un modèle global. Nous allons donc chercher à obtenir un modèle local M_{V_i} pour chaque variable de mesure $V_i \in L_{Vm}$. M_{V_i} est entièrement défini par un graphe causal $G^c(M_{V_i})$ déduit du graphe causal global $G^c=(L_E, L_V, S^c)$. Afin de simplifier la présentation de l'algorithme de construction des modèles locaux, nous introduisons tout d'abord quelques notations. Etant donné un graphe biparti, ${}^\circ X$ représente l'ensemble des antécédents du nœud X et X° représente l'ensemble des nœuds à l'extrémité des arcs partant de X . Si X est une variable (resp. équation) alors ${}^\circ X$ et X° sont des ensembles d'équations (resp. variables). L'opérateur \rightarrow sert à définir les arcs d'un graphe biparti. Un arc relie soit une équation E à une variable V ($E \rightarrow V$), soit une variable à une équation ($V \rightarrow E$). En généralisant cette notation, $L_V \rightarrow L_E$ indique que toutes les variables de L_V doivent être reliées à toutes les équations de L_E . L'algorithme de construction d'un modèle local M_{V_i} repose sur une « remontée » récursive dans le graphe causal du modèle global $G^c=(L_E, L_V, S^c)$ en partant de la variable de mesure V_i . Comme l'algorithme consiste à parcourir G^c et à construire $G^c(M_{V_i})$, on considère que l'opérateur ${}^\circ$ est appliqué par défaut à G^c et que l'opérateur \rightarrow est appliqué par défaut à $G^c(M_{V_i})$. L'algorithme de construction des modèles locaux est donné et commenté dans la Table II.5.

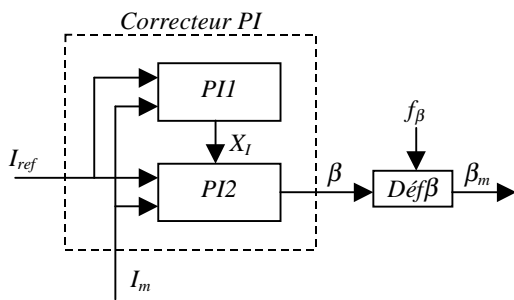
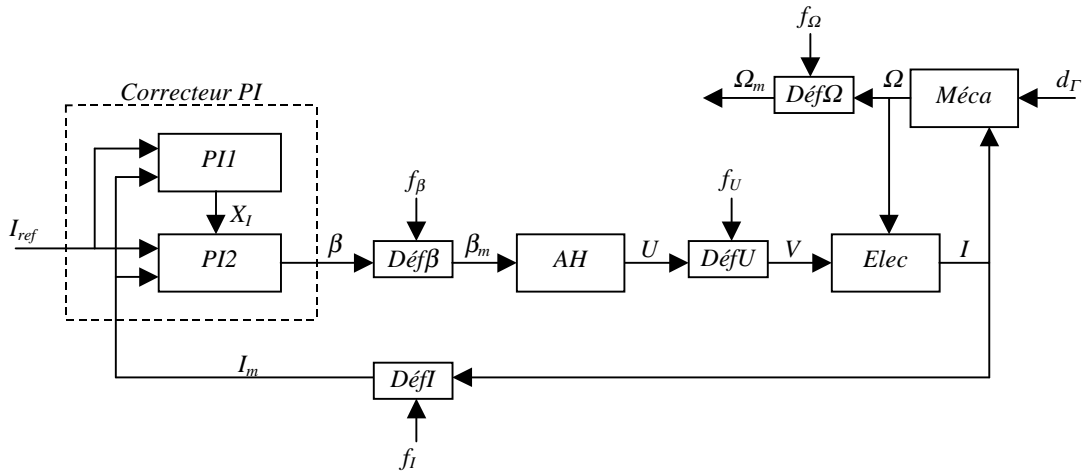
<i>Algorithme</i>	<i>Commentaires</i>
$M_{Vi} = \text{local}(V_i, G^c, M_{Vi}, (L_{Vex}, L_{Vmé}, L_{Vm}, L_{Em}))$ $E = \circ V_i$ $E \rightarrow V_i$ $L_A = \circ E$ Pour $j=1 \dots \text{card}(L_A)$ Si $A_j \in L_{Vex} \cup L_{Vm}$ Alors $A_j \rightarrow E$ Sinon, si $A_j \in L_{Vmé}$ et $M_{Vi} \neq \emptyset$ Alors $E_m = \{ \circ A_j \cup A_j \circ \} \cap L_{Em}$ $E_m \rightarrow A_j$ $\{ \circ E_m \cup E_m \circ \} \setminus A_j \rightarrow E_m$ Sinon $M_{Vi} = \text{local}(A_j, G^c, M_{Vi}, \dots)$ Fin Si Fin Pour	Fonction récursive construisant le modèle local M_{Vi} en partant de la variable de mesure V_i . E : équation permettant le calcul de V_i Création d'un arc entre E et V_i dans $G^c(M_{Vi})$ L_A : ensemble des antécédents de E = ensemble des variables servant à calculer V_i Pour chaque variable $A_j \in L_A$ servant à calculer V_i , ... Si A_j est une variable exogène ou de mesure, alors... Création d'un arc entre A_j et E dans $G^c(M_{Vi})$ Sinon, si A_j est une variable mesurée et si ce n'est pas le premier passage dans "local" ¹ , alors... E_m : équation de mesure associée à A_j Création d'un arc entre E_m et A_j dans $G^c(M_{Vi})$ Création d'arcs entre les variables nécessaires au calcul de A_j par E_m et l'équation E_m Dans tous les autres cas, la remontée récursive se poursuit dans G^c en partant de A_j .
Initialisation de $L = (L_{Vex}, L_{Vmé}, L_{Vm}, L_{Em})$ Initialisation de $G^c = (L_E, L_V, S^c)$ Pour $i=1 \dots \text{card}(L_{Vm})$ $M_{Vi} = \emptyset$ $M_{Vi} = \text{local}(V_i, G^c, M_{Vi}, L)$ Fin Pour	Ensembles des variables exogènes, mesurées, de mesure et des équations de mesure. Obtention de la structure et de l'ordonnancement causal du système étudié. Pour chaque variable de mesure $V_i \in L_{Vm}$, ... Initialisation du modèle local associé à V_i (vide) Construction du modèle local associé à V_i

Table II.5 : Algorithme de construction des modèles locaux

II.4.3. Application à la MCC

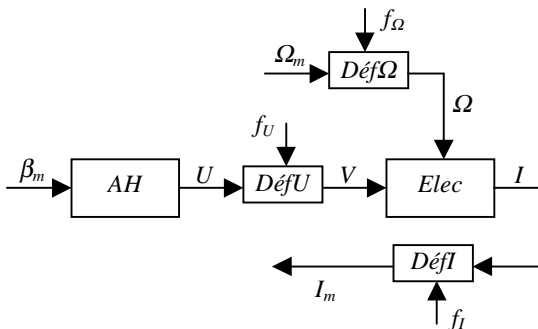
L'initialisation de $L = (L_{Vex}, L_{Vmé}, L_{Vm}, L_{Em})$ et de G^c dans le cas de la machine à courant continu est donnée par (II.11), (II.13), (II.14), (II.28), (II.29), (II.30) et par la Table II.3. Afin d'illustrer visuellement la décomposition de modèle obtenue, le schéma-bloc de la Figure II.6 est tout d'abord rappelé et les modèles locaux associés à β_m , I_m et Ω_m sont donnés en utilisant le même type de représentation (Figure II.8).

¹ Cette condition supplémentaire a pour but d'éviter que le parcours récursif inverse de G^c ne s'arrête à l'équation de mesure associée à V_i .



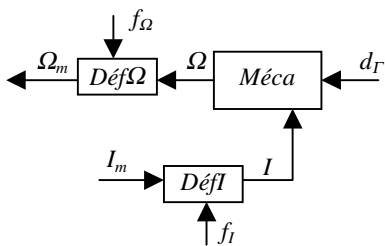
M_{β_m} : modèle local associé à β_m

$$\beta_m = M_{\beta_m}(I_{ref}, I_m, f_{\beta}) \tag{II.31}$$



M_{I_m} : modèle local associé à I_m

$$I_m = M_{I_m}(\beta_m, \Omega_m, f_I, f_{\Omega}, f_U) \tag{II.32}$$



M_{Ω_m} : modèle local associé à Ω_m

$$\Omega_m = M_{\Omega_m}(I_m, d_r, f_I, f_{\Omega}) \tag{II.33}$$

Figure II.8 : Application de la décomposition en modèles locaux au modèle global de la MCC

Le fait d’avoir modélisé les défauts sous la forme d’entrées exogènes permet de voir directement lesquels ont une influence sur chacun des modèles locaux (II.31) à (II.33).

II.5. Génération de résidus, étude de leurs propriétés

II.5.1. Génération des résidus locaux

Les modèles locaux que nous avons obtenus décrivent une partie du système étudié. Ce ne sont pas encore des modèles pour le diagnostic dans la mesure où ils ne permettent pas d'évaluer la cohérence entre le comportement qu'ils décrivent et les observations disponibles. L'objectif d'un générateur de résidu est justement d'évaluer une telle cohérence. Le principe retenu a été le suivant : le modèle local M_{V_i} est utilisé pour calculer une prédiction de V_i , \hat{V}_i . Il est alors possible de générer un résidu local, noté \hat{r}_{V_i} , par différence entre V_i et \hat{V}_i . Ainsi, un résidu local correspond ni plus ni moins à l'erreur de sortie du modèle local associé. Compte tenu de l'algorithme de construction d'un modèle local M_{V_i} , les variables d'entrée de ce dernier sont soit des variables exogènes, $L_{V_{ex}}^{V_i} \subset L_{V_{ex}}$, soit des variables de mesure, $L_{V_m}^{V_i} \subset L_{V_m}$.

$$V_i = M_{V_i}(L_{V_m}^{V_i}, L_{V_{ex}}^{V_i}) \quad (\text{II.34})$$

Par ailleurs, l'ensemble des variables observées dans le cadre de l'application considérée est noté $L_{V_{obs}}$. $L_{V_{nonobs}} = L_V \setminus L_{V_{obs}}$ représente l'ensemble des variables non observées. Notons que toutes les variables de mesure sont des variables observées ($L_{V_m} \subset L_{V_{obs}}$) alors que ce n'est pas forcément le cas pour les variables exogènes : une consigne correspond à une variable exogène connue (observée), alors qu'une perturbation ou un défaut peut être modélisé par une entrée exogène inconnue (non observée). Bien évidemment, la prédiction de V_i ne peut être calculée qu'à partir des observations disponibles ; les variables inconnues sont fixées à une valeur nulle pour rendre la simulation possible.

$$\hat{V}_i = M_{V_i}(L_{V_m}^{V_i}, L_{V_{ex}}^{V_i} \cap L_{V_{obs}}, (L_{V_{ex}}^{V_i} \cap L_{V_{nonobs}}) = 0) \quad (\text{II.35})$$

$$\hat{r}_{V_i} = V_i - \hat{V}_i \quad (\text{II.36})$$

II.5.2. Propriétés des résidus locaux

En reportant (II.34) et (II.35) dans (II.36), on obtient une nouvelle expression de \hat{r}_{V_i} .

$$\begin{aligned} \hat{r}_{V_i} = & M_{V_i}(L_{V_m}^{V_i}, L_{V_{ex}}^{V_i} \cap L_{V_{obs}}, L_{V_{ex}}^{V_i} \cap L_{V_{nonobs}}) \dots \\ & - M_{V_i}(L_{V_m}^{V_i}, L_{V_{ex}}^{V_i} \cap L_{V_{obs}}, (L_{V_{ex}}^{V_i} \cap L_{V_{nonobs}}) = 0) \end{aligned} \quad (\text{II.37})$$

En supposant que les équations modélisent parfaitement le système¹, (II.37) montre que les seules variables pouvant rendre \hat{r}_{V_i} non nul sont les variables exogènes non observées intervenant dans le modèle local M_{V_i} ($L_{V_{ex}}^{V_i} \cap L_{V_{nonobs}}$). Pour être sûr qu'une variable $V \in L_{V_{ex}}^{V_i} \cap L_{V_{nonobs}}$ puissent affecter \hat{r}_{V_i} lorsqu'elle devient non nulle, il est nécessaire de vérifier l'existence d'un chemin entre V et V_i . Un tel chemin existe forcément compte tenu de l'algorithme de construction des modèles locaux. Nous dirons que \hat{r}_{V_i} est causalement sensible à V . Cette formulation vise à distinguer l'existence d'un chemin, information liée à la structure

¹ Si ce n'est pas le cas, on ne peut plus considérer que M_{V_i} dans (II.34) soit strictement identique à M_{V_i} dans (II.35).

causale, de la sensibilité au sens classique du terme. Par exemple, le résidu $r=U.V$ est causalement sensible à V (il existe un chemin entre V et r , autrement dit, une variation de V peut se propager sur r) ; néanmoins, la sensibilité de r à V peut être nulle si, pour une raison ou pour une autre, U s'annule. Contrairement à la notion de résidu « sensible » qui nécessite de connaître le contenu des équations, la notion de résidu « causalement sensible » ne fait appel qu'à l'information structurelle et causale. Etant donné ces définitions, nous pouvons énoncer les propriétés de localisation des résidus locaux (Table II.6).

Propriétés de localisation des résidus locaux : En supposant que les équations modélisent parfaitement le système, le résidu local \hat{r}_{Vi} associé à un modèle local M_{Vi} est causalement sensible aux variables exogènes non observées intervenant dans M_{Vi} ($L_{Vex}^{Vi} \cap L_{Vnonobs}$) et seulement à ces variables là. En supposant qu'il existe des erreurs de modélisation, seules celles qui affectent les équations L_E^{Vi} intervenant dans M_{Vi} peuvent rendre \hat{r}_{Vi} non nul.

Table II.6 : Propriétés de localisation des résidus locaux

L'intérêt de construire les modèles locaux comme les plus petits sous-systèmes pour lesquels la sortie est une variable de mesure et pour lesquels les entrées sont des variables exogènes ou de mesure est ainsi mis en lumière : plus la décomposition du modèle global peut se faire finement, autrement dit, plus les modèles locaux sont nombreux et de petite taille et meilleures seront les propriétés de localisation des résidus locaux correspondants.

II.5.3. Résidus globaux

Un deuxième jeu de résidus peut être calculé pour chaque variable de mesure: le résidu global r_{Vi}^* associé à la variable de mesure $V_i \in L_{Vm}$ correspond à l'erreur de sortie du modèle global M du système étudié :

$$[\dots \ V_i \ \dots]^T = M(L_{Vex}) \quad (\text{II.38})$$

$$[\dots \ V_i^* \ \dots]^T = M(L_{Vex} \cap L_{Vobs}, (L_{Vex} \cap L_{Vnonobs}) = 0) \quad (\text{II.39})$$

$$r^* = [\dots \ r_{Vi}^* \ \dots]^T = [\dots \ V_i \ \dots]^T - [\dots \ V_i^* \ \dots]^T \quad (\text{II.40})$$

(II.38) indique que le modèle global relie toutes les variables exogènes à toutes les variables de mesure. (II.39) indique comment la valeur de référence de V_i , V_i^* , est calculée à partir des variables exogènes observables ; comme en II.5.1, les variables exogènes non observables (modélisant les perturbations, les défauts...) sont fixées à zéro pour rendre la simulation possible. V_i^* peut donc être considérée comme une référence homogène à V_i du comportement sans défaut et sans perturbation du système *global* (de même qu'en II.5.1, \hat{V}_i peut être considérée comme une référence homogène à V_i du comportement sans défaut et sans perturbation du sous-système *local*).

II.5.4. Propriétés des résidus globaux

Le même raisonnement qu'en II.5.2 est appliqué aux résidus globaux : en reportant (II.38) et (II.39) dans (II.40), une nouvelle expression de r^* est obtenue : (II.41).

$$r^* = [\dots \ r_{V_i}^* \ \dots]^T = M(L_{V_m} \cap L_{V_{obs}}, L_{V_{ex}} \cap L_{V_{nonobs}}) \dots - M(L_{V_{ex}} \cap L_{V_{obs}}, (L_{V_{ex}} \cap L_{V_{nonobs}}) = 0) \quad (\text{II.41})$$

En supposant que les équations modélisent parfaitement le système, (II.41) montre que les seules variables susceptibles de rendre les résidus globaux non nuls sont les variables exogènes non observées ($L_{V_{ex}} \cap L_{V_{nonobs}}$). Pour savoir si une variation de $V \in L_{V_{ex}} \cap L_{V_{nonobs}}$ peut se propager sur $r_{V_i}^*$, autrement dit, pour vérifier si $r_{V_i}^*$ est causalement sensible à V , il faut vérifier l'existence d'un chemin entre V et V_i dans le graphe causal du système. Les propriétés de localisation des résidus globaux sont données par la Table II.7.

Propriétés de localisation des résidus globaux : En supposant que les équations modélisent parfaitement le système, un résidu global $r_{V_i}^*$ ne peut être causalement sensible qu'aux variables exogènes non observées. $r_{V_i}^*$ est causalement sensible à une variable $V \in L_{V_{ex}} \cap L_{V_{nonobs}}$ si et seulement si il existe un chemin entre V et V_i dans G^c . En supposant qu'il existe des erreurs de modélisation, toutes celles affectant les équations qui interviennent dans le calcul de V_i à partir des entrées exogènes peuvent rendre $r_{V_i}^*$ non nul.

Table II.7 : Propriétés de localisation des résidus globaux

II.5.5. Comparaison des propriétés des résidus globaux et locaux

La comparaison de la Table II.6 et de la Table II.7 montre que les propriétés de localisation des résidus locaux sont a priori plus fortes que celles des résidus globaux. Cela peut donner lieu à plusieurs interprétations :

1. Dans le cas des résidus locaux, l'existence des chemins entre les variables exogènes non observées et la sortie du modèle local est résolue par la construction même du modèle local.
2. Un modèle local étant le plus petit possible, l'ensemble des variables susceptibles d'affecter le résidu local associé est de cardinal plus faible que pour un résidu global.
3. L'injection des mesures en entrée d'un modèle local (Comparer (II.35) à (II.39)) revient à « couper » l'influence de certaines variables exogènes non observées sur le résidu local. Cela vient du fait que les mesures reflètent les effets de telles propagations.

Une analyse locale permet de donner un autre éclairage sur les propriétés de localisation des résidus globaux et de préciser le point 3. Etant donné un modèle local M_{V_i} , on désigne par $L_{V_m}^{V_i}$ les variables de mesures en entrée de M_{V_i} ($L_{V_m}^{V_i} \subset L_{V_m}$). En outre, (II.39) permet de calculer une valeur de référence pour chaque variable de mesure (L_{V_m}) donc, en particulier, pour chaque variable de $L_{V_m}^{V_i}$. En notant $L_{V^*}^{V_i}$ l'ensemble des valeurs de référence associées aux variables de mesure de $L_{V_m}^{V_i}$, une expression locale d'un résidu global peut être obtenue :

$$V_i^* = M_{V_i}(L_{V^*}^{V_i}, L_{V_{ex}}^{V_i} \cap L_{V_{obs}}, (L_{V_{ex}}^{V_i} \cap L_{V_{nonobs}}) = 0) \quad (\text{II.42})$$

$$r_{V_i}^* = V_i^* - V_i^* \quad (\text{II.43})$$

Reporter (II.34) et (II.42) dans (II.43) permet de mettre en évidence les propriétés de localisation de $r_{V_i}^*$ d'un point de vue local.

$$r_{V_i}^* = M_{V_i}(L_{V_m}^{V_i}, L_{V_{ex}}^{V_i} \cap L_{V_{obs}}^{V_i}, L_{V_{ex}}^{V_i} \cap L_{V_{nonobs}}^{V_i}) \dots \quad (\text{II.44})$$

$$- M_{V_i}(L_{V^*}^{V_i}, L_{V_{ex}}^{V_i} \cap L_{V_{obs}}^{V_i}, (L_{V_{ex}}^{V_i} \cap L_{V_{nonobs}}^{V_i}) = 0)$$

(II.44) montre que $r_{V_i}^*$ est susceptible d'être affecté par les variables exogènes non observées intervenant dans M_{V_i} d'une part, et par toutes les sources d'écart possibles entre les variables de $L_{V^*}^{V_i}$ et de $L_{V_m}^{V_i}$ d'autre part ; l'application des propriétés de localisation des résidus globaux (Table II.7) indique que ces écarts correspondent à une propagation (donc à l'existence d'un chemin) entre les variables exogènes non observées et les entrées de M_{V_i} . En comparant les propriétés de localisation ainsi obtenues avec celles des résidus locaux (Table II.6), cela met en évidence que l'injection des mesures en entrée d'un modèle local revient à « couper » l'influence de certaines variables exogènes non observées sur le résidu local. Parallèlement, l'injection des valeurs de référence en entrée des modèles locaux revient à autoriser la propagation d'écarts sur le résidu par l'intermédiaire des entrées concernées.

Remarque : un des intérêts de l'approche locale réside dans la possibilité d'implanter une démarche de localisation récursive : lorsqu'un écart est détecté sur une variable, c'est-à-dire lorsque le résidu global correspondant est non nul, un test du résidu local permet de déterminer si l'origine de l'écart est locale ou amont¹. En effet, si le résidu local est proche de zéro, cela signifie que le sous-système décrit par le modèle local a un comportement conforme à celui du modèle ; compte tenu de l'existence d'un écart sur le résidu global, on conclut alors que le défaut à l'origine de cet écart agit en amont du sous-système local. En injectant sélectivement les mesures en entrée du modèle local, par exemple, toutes les mesures sauf une, il est possible de déduire par quelle(s) entrée(s) du modèle local les écarts observés se sont propagés. Un raisonnement récursif permet alors de remonter jusqu'au sous-système local à l'origine de l'écart observé [Montmain, 1993]. Une approche récursive pour la localisation ne se justifie vraiment que pour des systèmes fortement instrumentés présentant un grand nombre de variables.

II.5.6. Génération de résidus par composition de modèles locaux

Pour chaque variable de mesure, un modèle local dont la sortie correspond à cette variable a été construit. Par ailleurs, certaines entrées des modèles locaux sont elles-mêmes des variables de mesure. Il est donc possible d'effectuer une composition de plusieurs modèles locaux et de générer de nouveaux résidus. Les propriétés de sensibilité de ces nouveaux résidus diffèrent a priori de celles des résidus locaux et globaux, ce qui peut contribuer à améliorer les capacités de localisation. Soit M_{V_i} et M_{V_j} deux modèles locaux tels que $V_j \in L_{V_m}^{V_i}$.

$$V_i = M_{V_i}(L_{V_m}^{V_i}, L_{V_{ex}}^{V_i}) \quad (\text{II.45})$$

$$V_j = M_{V_j}(L_{V_m}^{V_j}, L_{V_{ex}}^{V_j}) \quad (\text{II.46})$$

Comme $V_j \in L_{V_m}^{V_i}$, il est possible de composer M_{V_i} et M_{V_j} pour former le modèle noté $M_{V_i(V_j)}$:

$$V_i = M_{V_i}(M_{V_j}(L_{V_m}^{V_j}, L_{V_{ex}}^{V_j}), L_{V_m}^{V_i} - \{V_j\}, L_{V_{ex}}^{V_i}) \quad (\text{II.47})$$

$$V_i = M_{V_i(V_j)}(L_{V_m}^{V_j} \cup L_{V_m}^{V_i} - \{V_j\}, L_{V_{ex}}^{V_j} \cup L_{V_{ex}}^{V_i}) \quad (\text{II.48})$$

¹ Le sens accordé à « amont » repose sur l'orientation des arcs déduite de l'ordonnancement causal. Les arcs orientés traduisent les chemins de propagation possibles pour les défauts (entre autres).

Le modèle composé $M_{Vi(Vj)}$ présente la même structure que celle d'un modèle local (comparer (II.48) à (II.45)). Le même raisonnement que celui appliqué aux modèles locaux en II.5.1 et en II.5.2 s'applique. Ainsi, il est possible de générer le résidu $\hat{r}_{Vi(Vj)}$ à partir de ce modèle. De plus, $\hat{r}_{Vi(Vj)}$ est causalement sensible aux variables exogènes non observées intervenant dans M_{Vi} ou M_{Vj} , $(L_{Vex}^{Vj} \cup L_{Vex}^{Vi}) \cap L_{Vnonobs}$, qui n'ont pas été simplifiées du fait de la composition des deux modèles. De telles simplifications peuvent intervenir par exemple au niveau des équations de mesures effectuant la jonction entre deux modèles locaux. La Figure II.9 illustre ce phénomène (v est une variable mesurée, v_m est une variable de mesure, Δv est une variable exogène modélisant un défaut capteur et eq est l'équation de mesure de v).

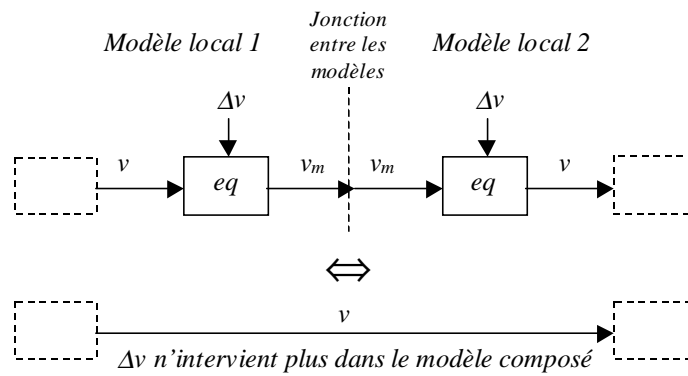


Figure II.9 : Exemple de simplification liée à la composition de deux modèles

Remarque 1 : en l'absence des simplifications liées à la composition des modèles, les nouveaux résidus n'apportent pas plus d'information que les résidus locaux de départ ; la composition de modèles ne présente alors aucun intérêt.

Remarque 2 : nous n'avons présenté que la composition de deux modèles locaux mais la généralisation à plus de deux modèles est évidente. Par exemple, $M_{Vi(Vj,Vk(Vl))}$ représente le modèle issu de la composition de M_{Vi} avec M_{Vj} et M_{Vk} , lui-même composé avec M_{Vl} ; le résidu calculé à partir de $M_{Vi(Vj,Vk(Vl))}$ est noté $\hat{r}_{Vi(Vj,Vk(Vl))}$.

II.5.7. Application à la MCC

II.5.7.1. Résidus globaux et locaux

L'ensemble des variables de mesure ($L_{Vm} = \{\beta_m, I_m, \Omega_m\}$) étant de cardinal trois, trois résidus locaux et trois résidus globaux vont pouvoir être construits, sans compter les résidus résultant de la composition des modèles locaux. L'ensemble des variables observées et l'ensemble des variables non observées sont les suivants :

$$L_{Vobs} = \{I_{ref}, \beta_m, I_m, \Omega_m\} \tag{II.49}$$

$$L_{Vnonobs} = \{f_\beta, f_I, f_\Omega, f_U, d_\Gamma, X_I, \beta, I, \Omega, U, V\} \tag{II.50}$$

$$L_{Vex} \cap L_{Vnonobs} = \{f_\beta, f_I, f_\Omega, f_U, d_\Gamma\} \tag{II.51}$$

A partir des propriétés de localisation des résidus globaux et locaux, il est possible d'établir des tables de signature théoriques indiquant à quels défauts modélisés chacun des résidus est (causalement) sensible.

Exemple 1, résidu global $r_{I_m}^*$: l'application des propriétés de la Table II.7 indique que $r_{I_m}^*$ ne peut être causalement sensible qu'aux variables exogènes non observées (II.51). De plus, le graphe causal de la MCC (Figure II.6) montre qu'il existe un chemin entre f_β et I_m , un chemin entre f_I et I_m , un chemin entre f_U et I_m et un chemin entre d_Γ et I_m . Par contre, il n'existe pas de chemin entre f_Ω et I_m . On peut dès lors conclure que $r_{I_m}^*$ est causalement sensible à f_β, f_I, f_U et d_Γ et que $r_{I_m}^*$ est insensible à f_Ω (Table II.8a).

Exemple 2, résidu local \hat{r}_{I_m} : le modèle local associé à I_m est donné par (II.32). On en déduit que $L_{vm}^{I_m} = \{\beta_m, \Omega_m\}$ et $L_{vex}^{I_m} = \{f_I, f_\Omega, f_U\}$. L'application des propriétés de la Table II.6 indique que \hat{r}_{I_m} est causalement sensible à f_I, f_Ω, f_U et seulement à ces variables là (Table II.8b).

	f_β	f_I	f_Ω	f_U	d_Γ
$r_{\beta_m}^*$	1	1	0	1	1
$r_{I_m}^*$	1	1	0	1	1
$r_{\Omega_m}^*$	1	1	1	1	1

(a) résidus globaux

	f_β	f_I	f_Ω	f_U	d_Γ
\hat{r}_{β_m}	1	0	0	0	0
\hat{r}_{I_m}	0	1	1	1	0
\hat{r}_{Ω_m}	0	1	1	0	1

(b) résidus locaux

Table II.8 : Tables de signature théoriques (résidus globaux et locaux)

La Table II.8 montre l'intérêt des résidus locaux : non seulement ils permettent a priori de discriminer $f_\beta, \{f_I, f_\Omega\}$ et f_U au lieu de $\{f_\beta, f_I, f_U\}$ et f_Ω pour les résidus globaux mais, en plus, deux d'entre eux sont parfaitement découplés de la perturbation due au couple résistant, alors qu'aucun des résidus globaux ne l'est. Logiquement, le résidu local qui reste sensible au couple résistant est celui construit à partir de l'équation mécanique de la MCC.

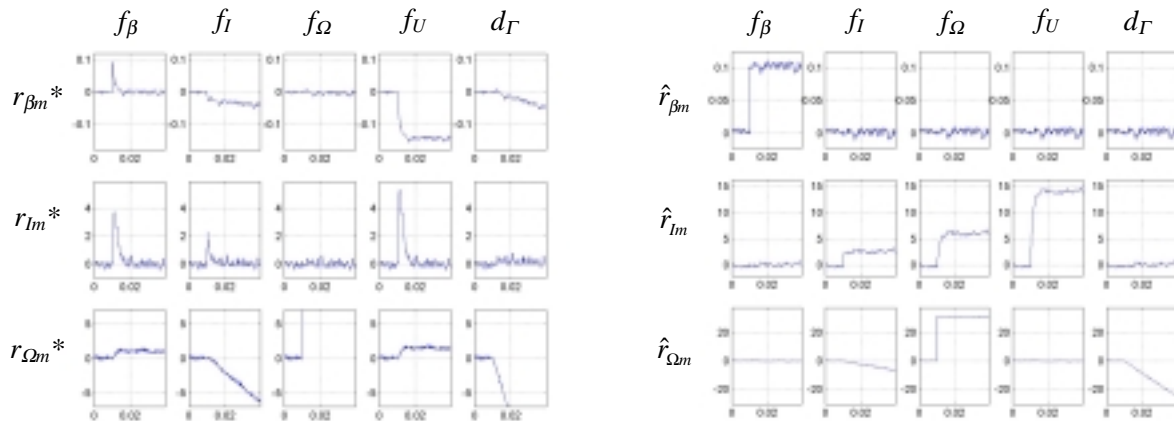


Figure II.10 : Vérification en simulation des tables de signatures théoriques

La Figure II.10 permet de vérifier en simulation les tables de signature théoriques des résidus globaux et locaux (Table II.8) pour des défauts en échelon. C'est au stade de ces simulations (et seulement à partir de là) qu'intervient la connaissance sur le contenu des équations. Certains résidus paraissent répondre à certains défauts en échelon par une rampe ; en fait, il s'agit du début de la réponse correspondant à la constante de temps mécanique de la MCC, beaucoup plus lente que la constante de temps électrique.

II.5.7.2. Résidus associés à une composition de modèles locaux

La composition des modèles locaux est maintenant appliquée à la MCC. L'ensemble des compositions qui ne font pas intervenir plus d'une fois un même modèle local est répertorié dans la Figure II.11. L'arbre correspondant est établi à partir des équations des modèles locaux (II.31), (II.32) et (II.33).

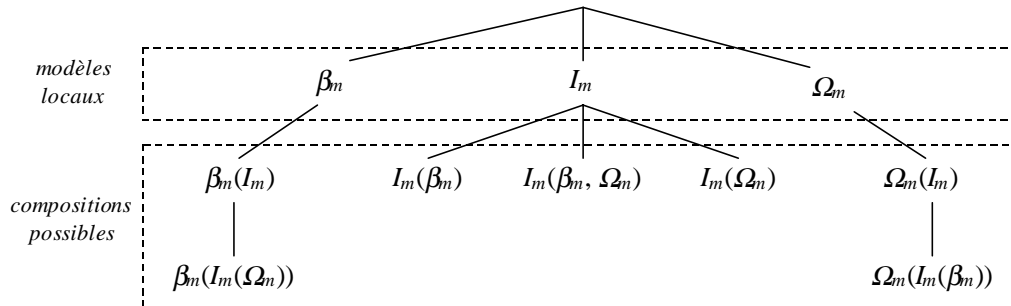


Figure II.11 : Compositions possibles des modèles locaux

Les tables de signatures théoriques déduites de la Figure II.11 sont reportées dans la Table II.9.

	f_β	f_I	f_Ω	f_U	d_Γ
$\hat{r}_{\Omega_m(I_m)}$	0	0	1	1	1
$\hat{r}_{I_m(\Omega_m)}$	0	1	0	1	1
$\hat{r}_{\beta_m(I_m)}$	1	1	1	1	0
$\hat{r}_{I_m(\beta_m)}$	1	1	1	1	0
$\hat{r}_{\beta_m(I_m(\Omega_m))}$	1	1	0	1	1
$\hat{r}_{I_m(\beta_m, \Omega_m)}$	1	1	0	1	1
$\hat{r}_{\Omega_m(I_m(\beta_m))}$	1	0	1	1	1

Table II.9 : Tables de signatures théoriques (Composition des modèles locaux)

Le résidu $\hat{r}_{\Omega_m(I_m)}$ (resp. $\hat{r}_{I_m(\Omega_m)}$) est obtenu en reportant M_{I_m} (resp. M_{Ω_m}) dans M_{Ω_m} (resp. M_{I_m}). La Figure II.8 permet de mettre en évidence la simplification de f_I (resp. f_Ω), comme indiqué sur la Figure II.9. Chacun de ces deux résidus permet de localiser entre eux f_I et f_Ω : la simplification de f_I ou de f_Ω résultant de la composition des deux modèles locaux permet de localiser tous les défauts, ce qui n'était pas possible avec les seuls résidus locaux. Par contre, le couple résistant influence ces deux premiers résidus. La Figure II.12 montre que les réponses de $\hat{r}_{\Omega_m(I_m)}$ et $\hat{r}_{I_m(\Omega_m)}$ à des entrées exogènes non observées en échelon sont conformes à la table de signature théorique (Table II.9).

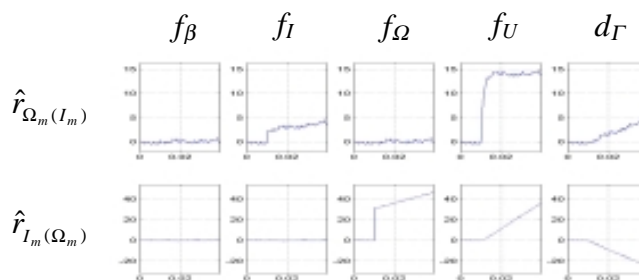


Figure II.12 : Vérification en simulation des signatures théoriques pour $\hat{r}_{\Omega_m(I_m)}$ et $\hat{r}_{I_m(\Omega_m)}$

L'association de M_{β_m} et M_{I_m} ne donne pas lieu à simplification (quel que soit l'ordre de l'association). Les deux résidus ainsi obtenus, $\hat{r}_{\beta_m(I_m)}$ et $\hat{r}_{I_m(\beta_m)}$, n'apportent pas d'information supplémentaire en terme de localisation par rapport aux résidus locaux. Ces deux résidus sont sensibles à tous les défauts et insensibles à la perturbation de couple : ils satisfont les objectifs de la détection.

Les trois dernières lignes de la Table II.9 font référence à trois résidus dont le calcul fait intervenir les trois modèles locaux et, par conséquent, toutes les équations du modèle global. Le premier est homogène à β , le second à I et le troisième à Ω ; une comparaison avec les résidus globaux s'impose. On pourrait s'attendre en effet à retrouver des signatures théoriques similaires. Or, si les influences des entrées exogènes non observées sont identiques pour $\hat{r}_{\beta_m(I_m(\Omega_m))}$ et $r_{\beta_m}^*$ d'une part, pour $\hat{r}_{I_m(\beta_m, \Omega_m)}$ et $r_{I_m}^*$ d'autre part, il n'en va pas de même pour $\hat{r}_{\Omega_m(I_m(\beta_m))}$ et $r_{\Omega_m}^*$. La différence entre les modèles permettant de calculer ces deux derniers résidus vient de l'absence de bouclage dans le cas de la composition des modèles locaux et de la présence de bouclages dans le cas du modèle global du système. Ainsi, les entrées connues du modèle $M_{\Omega_m(I_m(\beta_m))}$ sont I_{ref} , I_m et Ω_m alors que la seule entrée connue du modèle global est I_{ref} .

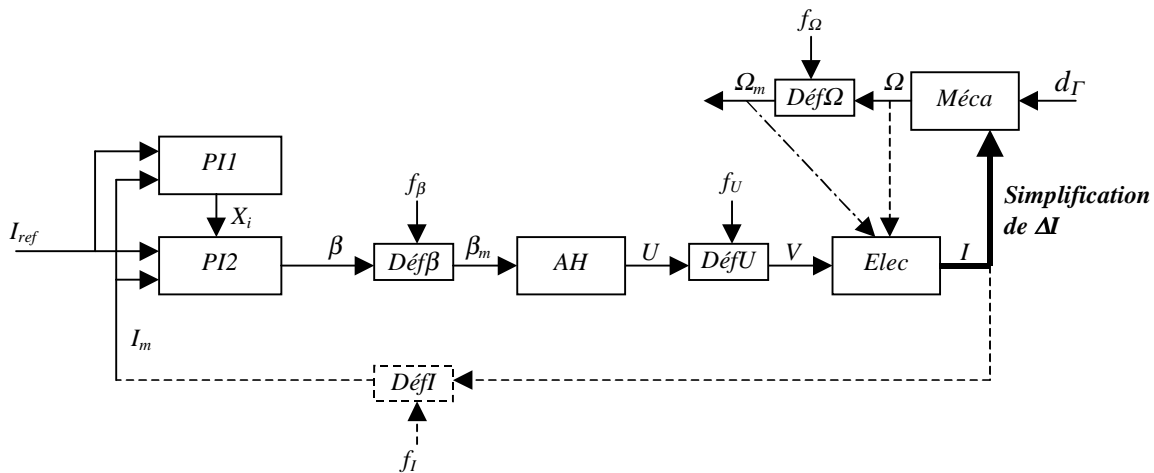


Figure II.13 : $M_{\Omega_m(I_m(\beta_m))}$: différences avec le modèle global (pointillés)

Une étude plus détaillée des différences entre $M_{\Omega_m(I_m(\beta_m))}$ et le modèle global est donnée par la Figure II.13. Les pointillés simples mettent en évidence les liens absents de $M_{\Omega_m(I_m(\beta_m))}$ et présents dans le modèle global (la signification des pointillés mixte est inverse). Alors que f_β , f_Ω , f_U et d_T apparaissent dans les deux modèles, f_I est lié à la boucle de courant qui n'apparaît que dans le modèle global. Ainsi, la présence ou l'absence de certains bouclages nous a permis d'interpréter la différence de sensibilité entre un résidu résultant de la composition de tous les modèles locaux et le résidu global qui lui est homogène.

II.6. Application à la MAS

La décomposition de modèle est appliquée à l'ensemble constitué de l'alimentation, de l'onduleur et du moteur asynchrone. L'alimentation et l'onduleur peuvent être modélisés en première approximation par un gain unitaire. Cela revient à appliquer l'analyse locale au seul

moteur. Le fait de ne pas inclure la commande dans le système étudié ne modifie pas la décomposition de modèle qui sera obtenue. En effet, les entrées et les sorties de la commande étant des grandeurs connues, l'algorithme de construction des modèles locaux s'arrêtera forcément avant que les équations de la commande ne soient atteintes lors du parcours inverse du graphe causal associé au moteur. Contrairement à l'application relative à la MCC, les défauts ne sont pas modélisés explicitement ; l'analyse est donc plus centrée sur la décomposition de modèle et la génération des résidus que sur les propriétés de localisation de ces derniers.

La mise en équation du modèle de Park d'un moteur asynchrone est traitée dans l'annexe C. Rappelons simplement que les équations électriques d'un moteur asynchrone reposent sur les relations (II.52) et (II.53),

$$\Phi = L.I \quad (\text{II.52})$$

$$V = R.I + \frac{d\Phi}{dt} \quad (\text{II.53})$$

où Φ, I, V sont des vecteurs contenant respectivement les flux, les courants et les tensions relatifs au stator et au rotor dans le repère (abc) . L'application de la transformation de Park permet d'exprimer (II.52) et (II.53) dans un repère (dq) quelconque.

$$E_{alg1} : \quad \Phi_{sd} = L_s.I_{sd} + L_m.I_{rd} \quad (\text{II.54})$$

$$E_{alg2} : \quad \Phi_{sq} = L_s.I_{sq} + L_m.I_{rq} \quad (\text{II.55})$$

$$E_{alg3} : \quad \Phi_{rd} = L_m.I_{sd} + L_r.I_{rd} \quad (\text{II.56})$$

$$E_{alg4} : \quad \Phi_{rq} = L_m.I_{sq} + L_r.I_{rq} \quad (\text{II.57})$$

$$E_{dif1} : \quad V_{sd} = R_s.I_{sd} + \frac{d\Phi_{sd}}{dt} - \omega_s.\Phi_{sq} \quad (\text{II.58})$$

$$E_{dif2} : \quad V_{sq} = R_s.I_{sq} + \frac{d\Phi_{sq}}{dt} + \omega_s.\Phi_{sd} \quad (\text{II.59})$$

$$E_{dif3} : \quad 0 = R_r.I_{rd} + \frac{d\Phi_{rd}}{dt} - (\omega_s - p.\Omega).\Phi_{rq} \quad (\text{II.60})$$

$$E_{dif4} : \quad 0 = R_r.I_{rq} + \frac{d\Phi_{rq}}{dt} + (\omega_s - p.\Omega).\Phi_{rd} \quad (\text{II.61})$$

$$Méca : \quad J.\dot{\Omega} + \rho.\Omega = p.(\Phi_{sd}.I_{sq} - \Phi_{sq}.I_{sd}) - d_T \quad (\text{II.62})$$

Les équations algébriques (II.54) à (II.57) sont obtenues à partir de (II.52) et les équations différentielles (II.58) à (II.61) sont obtenues à partir de (II.53). Ainsi, (II.54) à (II.61) représentent les équations électriques alors que (II.62) représente l'équation mécanique du moteur asynchrone. Φ, I, V, Ω, d_T représentent des grandeurs respectivement homogènes à des flux, des courants, des tensions, une vitesse (mécanique), un couple (de charge). L, R, J, ρ, p sont des paramètres électriques et mécaniques respectivement homogènes à des inductances, des résistances, une inertie, un frottement visqueux, un nombre de paires de pôles dans le moteur. L'indice s (resp. r) fait référence à une grandeur statorique (resp. rotorique). Les indices d et q font référence aux axes du repère d'étude (dq) qui tourne par rapport au stator avec une pulsation notée ω_s . Le moteur étant commandé en tension, les ensembles d'équations et de variables à définir pour appliquer la décomposition de modèle sont les suivants :

$$L_E = \{E_{alg1}, E_{alg2}, E_{alg3}, E_{alg4}, E_{dif1}, E_{dif2}, E_{dif3}, E_{dif4}, Méca\} \quad (\text{II.63})$$

$$L_{Edif} = \{E_{dif1}, E_{dif2}, E_{dif3}, E_{dif4}, Méca\} \quad (II.64)$$

$$L_V = \{I_{sd}, I_{rd}, I_{sq}, I_{rq}, \Phi_{sd}, \Phi_{sq}, \Phi_{rd}, \Phi_{rq}, \Omega, V_{sd}, V_{sq}, \omega_s, d_\Gamma\} \quad (II.65)$$

$$L_{Vdif} = \{\Phi_{sd}, \Phi_{sq}, \Phi_{rd}, \Phi_{rq}, \Omega\} \quad (II.66)$$

$$L_{Vex} = \{V_{sd}, V_{sq}, \omega_s, d_\Gamma\} \quad (II.67)$$

$$L_{Vm} = \{I_{sd}, I_{sq}, \Omega\} \quad (II.68)$$

Remarque 1 : tant que l'on ne tient pas compte explicitement des défauts, l'introduction des variables mesurées ($L_{Vmé}$) et des équations de mesure (L_{Em}) n'est pas nécessaire : les seules variables de mesure (L_{Vm}) suffisent.

Remarque 2 : en pratique, les variables de mesure sont les courants I_{sa} et I_{sb} . I_{sd} et I_{sq} se déduisant directement de I_{sa} et I_{sb} par une transformation connue, nous les avons considérés comme des variables de mesure, pour simplifier notre présentation (II.68). Cette simplification ne remet pas en cause les résultats obtenus.

S^c		Var. non exo. et non diff.				Variables différentielles					Variables exogènes				
		S_{BLT}	I_{sd}	I_{rd}	I_{sq}	I_{rq}	Φ_{sd}	Φ_{sq}	Φ_{rd}	Φ_{rq}	Ω	V_{sd}	V_{sq}	ω_s	d_Γ
Equations algébriques	E_{alg3}		↗	↘				↘							
	E_{alg1}		↘	↗		↘									
	E_{alg4}				↗	↘				↘					
	E_{alg2}				↘	↗		↘							
Equations diff.	E_{dif1}		↘			↗	↘				↘		↘		
	E_{dif2}				↘	↘	↗					↘	↘		
	E_{dif3}			↘				↗	↘	↘			↘		
	E_{dif4}					↘		↗	↗	↘			↘		
	$Méca$		↘		↘		↘	↘			↘				↘
	ENV										↗	↗	↗	↗	↗

Table II.10 : Un ordonnancement causal possible pour la MAS

La Table II.10 représente un ordonnancement causal possible pour les équations d'une machine asynchrone (sans sa commande). Pour chacune des variables de mesure, un modèle local peut être établi à partir de cet ordonnancement causal.

$S^c(M_{Isd})$		Var. non exo. et non diff.				Variables différentielles					Variables exogènes				
		S_{BLT}	I_{sd}	I_{rd}	I_{sq}	I_{rq}	Φ_{sd}	Φ_{sq}	Φ_{rd}	Φ_{rq}	Ω	V_{sd}	V_{sq}	ω_s	d_Γ
Equations algébriques	E_{alg3}		↗	↘				↘							
	E_{alg1}		↘	↗		↘									
	E_{alg4}				↗	↘				↘					
	E_{alg2}				↘	↗		↘							
Equations diff.	E_{dif1}		↘			↗	↘				↘		↘		
	E_{dif2}				↘	↘	↗					↘	↘		
	E_{dif3}			↘				↗	↘	↘			↘		
	E_{dif4}					↘		↗	↗	↘			↘		
	$Méca$		↘		↘		↘	↘			↘				↘
	ENV		↗		↗					↗	↗	↗	↗	↗	↗

Table II.11 : M_{Isd} , modèle local associé à I_{sd} .

$S^c(M_{Isq})$		Var. non exo. et non diff.				Variables différentielles					Variables exogènes			
		S_{BLT}	I_{sd}	I_{rd}	I_{sq}	I_{rq}	Φ_{sd}	Φ_{sq}	Φ_{rd}	Φ_{rq}	Ω	V_{sd}	V_{sq}	ω_s
Equations algébriques	E_{alg3}	↔	↔					↔						
	E_{alg1}	↔	↔			↔								
	E_{alg4}			↔	↔				↔					
	E_{alg2}			↔	↔		↔							
Equations diff.	E_{dif1}	↔				↔	↔				↔		↔	
	E_{dif2}			↔		↔	↔					↔	↔	
	E_{dif3}		↔					↔	↔	↔			↔	
	E_{dif4}				↔			↔	↔	↔			↔	
	<i>Méca</i>	↔		↔		↔	↔			↔				↔
	<i>ENV</i>	↔		↔						↔	↔	↔	↔	↔

Table II.12 : M_{Isq} , modèle local associé à I_{sq} .

$S^c(M_\Omega)$		Var. non exo. et non diff.				Variables différentielles					Variables exogènes			
		S_{BLT}	I_{sd}	I_{rd}	I_{sq}	I_{rq}	Φ_{sd}	Φ_{sq}	Φ_{rd}	Φ_{rq}	Ω	V_{sd}	V_{sq}	ω_s
Equations algébriques	E_{alg3}	↔	↔					↔						
	E_{alg1}	↔	↔			↔								
	E_{alg4}			↔	↔				↔					
	E_{alg2}			↔	↔		↔							
Equations diff.	E_{dif1}	↔				↔	↔				↔		↔	
	E_{dif2}			↔		↔	↔					↔	↔	
	E_{dif3}		↔					↔	↔	↔			↔	
	E_{dif4}				↔			↔	↔	↔			↔	
	<i>Méca</i>	↔		↔		↔	↔			↔				↔
	<i>ENV</i>	↔		↔						↔	↔	↔	↔	↔

Table II.13 : M_Ω , modèle local associé à Ω .

La Table II.11 et la Table II.12 mettent en évidence la notion de recouvrement entre modèles locaux : de nombreuses équations et de nombreuses variables interviennent à la fois dans M_{Isd} et M_{Isq} . Ces recouvrements ont pour origine les nombreux couplages intervenant dans le modèle de la MAS. La Figure II.14 illustre cet aspect de la décomposition de modèle par un exemple simple (u_1, y_1, u_2, y_2 sont les mesures).

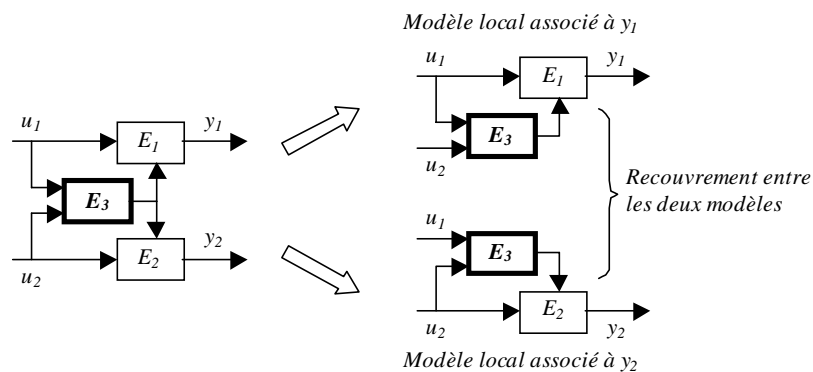


Figure II.14 : Un exemple de recouvrement entre deux modèles locaux

Compte tenu du fait que M_{Isd} et M_{Isq} contiennent chacun toutes les équations électriques sauf une équation algébrique et que les variables et les paramètres de cette équation algébrique se retrouvent tous dans les autres équations électriques, la distinction entre M_{Isd} et M_{Isq} ne se justifie pas du point de vue de la localisation. Ainsi, il est possible de regrouper M_{Isd} et M_{Isq} dans un seul

modèle¹, M_{Isdq} , possédant quatre entrées (V_{sd} , V_{sq} , ω_s et Ω) et deux sorties (I_{sd} et I_{sq}). Deux résidus homogènes aux courants peuvent donc être générés. De plus, comme le couple résistant n'intervient pas dans M_{Isd} et M_{Isq} , il n'intervient pas non plus dans M_{Isdq} , et les résidus ainsi générés sont parfaitement découplés vis-à-vis des variations du couple de charge. Nous retrouvons ici le résultat très classique qui consiste à injecter la mesure de vitesse en entrée du modèle électrique d'un moteur asynchrone pour s'affranchir des effets mécaniques.

Remarque : le modèle local M_Ω (Table II.13) faisant intervenir l'équation mécanique, le résidu local qui pourrait être calculé serait forcément sensible aux variations du couple de charge.

II.7. Bilan

II.7.1. Point de vue méthodologique

La connaissance structurelle utilisée dans ce chapitre repose sur l'existence et non sur la nature des relations entre variables. La contrepartie positive à la relative pauvreté de ce type de connaissance réside dans l'indépendance vis-à-vis de la nature des modèles. Ces derniers peuvent être linéaires ou non, quantitatifs ou qualitatifs,... La causalité étant assimilée à la calculabilité, un ordonnancement causal permet d'orienter les relations entre variables. Cette orientation est ensuite utilisée pour construire une méthode systématique de décomposition de modèle. Les modèles locaux ainsi obtenus permettent de générer des résidus dont les propriétés de localisation peuvent être établies a priori. La construction de tables de signatures théoriques est également un outil permettant de guider la décomposition d'une tâche de diagnostic globale en un ensemble de tâches de diagnostic plus élémentaires. Les découplages résultant de l'analyse causale proposée sont des découplages parfaits aux erreurs de modélisation près. Cependant, pour évaluer l'ordre de grandeur de l'influence de variables non découplées, la seule information structurelle et causale ne suffit plus : il devient alors nécessaire de tenir compte de la nature des relations entre variables et non plus seulement de leur existence. Une autre limitation de l'approche proposée tient au fait que les performances des résidus locaux en terme de localisation sont fortement liées à la finesse de la décomposition de modèle obtenue au préalable (nombre de modèles locaux différents, non recouvrement entre les modèles). Et la décomposition de modèle est elle-même très liée au nombre et à la position des observations disponibles.

II.7.2. Point de vue des applications

Les entraînements électriques étant des systèmes peu instrumentés et présentant de nombreux couplages, l'analyse causale ne permet pas à elle seule de résoudre le problème du diagnostic de ces systèmes. Toutefois, la robustesse au couple résistant est acquise dès lors que le calcul des résidus ne fait pas intervenir l'équation mécanique. Dans le cas de la MCC, l'étude des signatures théoriques montre qu'une détection des défauts considérés de manière robuste aux variations du couple de charge est possible. Sous les mêmes contraintes de robustesse, seuls deux groupes de défauts peuvent être localisés : f_β et $\{f_I, f_\Omega, f_U\}$ (Table II.8 et Table II.9).

L'application de la décomposition de modèle à la MAS peut paraître à certains égards comme l'utilisation d'un « marteau pilon pour écraser une mouche » (!) tant il est vrai que le résultat

¹ Ce modèle peut être vu comme un modèle électrique global (des bouclages font que I_{sd} et I_{sq} ne sont plus des entrées de ce modèle qui n'est pas une simple composition de M_{Isd} et M_{Isq}).

obtenu est classique et intuitif. Néanmoins, cette application met en évidence la notion de recouvrement entre modèles locaux et l'intérêt qu'il peut y avoir à décomposer un modèle global en modèles MIMO¹ et non pas seulement en modèles MISO².

¹ MIMO : multi-entrées, multi-sorties (*Multiple Input and Multiple Output*)

² MISO : multi-entrées, mono-sorties (*Multiple Input and Single Output*)

Chapitre III :
Observateurs, Equations de parité

Chapitre III :

Observateurs, Equations de parité

III.1. Introduction

Suite à l'application de l'analyse causale, tous les défauts ne peuvent pas forcément être localisés de manière robuste vis-à-vis des perturbations. Pour essayer de lever ces points durs, la connaissance sur la nature des relations entre variables va être désormais utilisée pour générer d'autres résidus. Cette connaissance est donnée par un modèle comportemental.

Dans ce chapitre, nous allons nous intéresser plus particulièrement à la génération de résidus à partir d'un modèle LTI (linéaire et invariant dans le temps). Après une synthèse sur des méthodes classiques de génération de résidus et sur les conditions de découplage, l'application de méthodes de découplage parfait et de découplage approximatif à la MCC vont nous permettre de générer un nouvel ensemble de résidus. Une étude de sensibilité destinée à être utilisée ultérieurement pour construire des fonctions de décision sera alors menée. Les liens avec les résidus issus de l'analyse causale seront mis en évidence dans le cas linéaire. Enfin, nous présenterons et appliquerons une méthode numérique de génération de résidus. Les choix arbitraires effectués et l'utilisation des degrés de liberté pour optimiser les sensibilités seront alors discutés.

III.2. Objectif de la génération de résidus

Un résidu est une observation intermédiaire calculée à partir d'observations élémentaires en vue d'aboutir aux observations décrivant les résultats du diagnostic. Les résidus sont donc conçus en vue de faciliter leur exploitation ultérieure par un outil de décision destiné à détecter et à localiser les défauts. Pour cela, deux approches sont possibles¹ [Gertler, 1991] :

- **Génération de résidus directionnels** : les résidus sont conçus de telle sorte que le vecteur des résidus reste confiné dans une direction particulière de l'espace des résidus, en réponse à un défaut particulier.

¹ Une approche unifiée dans le cadre des relations de parité est proposée dans [Gertler, 1997].

- **Génération de résidus structurés** : les résidus sont conçus de façon à répondre à des sous-ensembles de défauts différents, autrement dit (approche duale), les résidus sont conçus de façon à être découplés de sous-ensembles de défauts différents. Ces sous-ensembles de défauts permettent de structurer une table de signature appelée également matrice d'incidence. Gertler distingue trois cas pour une matrice d'incidence :
 - **non localisante** (une colonne est nulle ou deux au moins sont identiques),
 - **faiblement localisante** (les colonnes sont non nulles et distinctes deux à deux),
 - **fortement localisante** (en plus d'être faiblement localisante, aucune colonne ne peut être obtenue à partir d'une autre en remplaçant un '1' par un '0'). Un cas particulier utile, notamment pour la localisation de défauts multiples, correspond à une structure diagonale de la table de signature.

	f_1	f_2	f_3
r_1	1	1	1
r_2	1	1	1
r_3	1	0	0

Non localisante

	f_1	f_2	f_3
r_1	1	1	1
r_2	1	0	1
r_3	1	1	0

Faiblement localisante

	f_1	f_2	f_3
r_1	1	1	0
r_2	1	0	1
r_3	0	1	1

Fortement localisante

	f_1	f_2	f_3
r_1	1	0	0
r_2	0	1	0
r_3	0	0	1

Diagonale

Table III.1 : Illustration des propriétés d'une table de signature

Etant donné une table de signature Sg , les propriétés souhaitées pour les résidus correspondant aux différentes lignes de Sg sont les suivantes [Kinnaert, 1994]:

- La sensibilité doit être importante pour les signaux correspondant à un '1' dans Sg et faible pour les autres signaux. On distingue le cas du découplage parfait qui consiste à annuler les sensibilités associées à un '0' dans Sg (en l'absence d'erreur de modélisation) du cas du découplage approximatif qui consiste à seulement minimiser ces mêmes sensibilités.
- Les propriétés de sensibilité doivent être vérifiées non seulement pour le modèle nominal, mais aussi pour toute la classe des modèles plausibles. Il s'agit d'un problème de robustesse vis-à-vis des incertitudes de modèle.
- Dans le cas de tables de signatures booléennes, on peut également souhaiter que les sensibilités des signaux correspondant à un '1' sur une même ligne de Sg soient équilibrées, afin de pouvoir fixer ultérieurement des seuils sur les résidus sans que les non détections présentent de fortes disparités d'un défaut à l'autre. Nous reviendrons sur ces aspects dans les chapitres consacrés à la décision. Notons simplement l'importance des sensibilités relatives dans les lignes ou dans les colonnes de Sg .

III.3. Génération de résidus pour les systèmes LTI

Soit d et f les entrées modélisant les perturbations et les défauts du système étudié. Un résidu satisfaisant les impératifs de la détection est découplé de d et sensible à f . Etant donné une table de signature désirée, supposons que nous cherchions à construire un résidu r dont les propriétés en terme de sensibilité permettent de structurer une ligne de cette table. Soit f^0 (resp. f^1) l'ensemble des défauts associés à un '0' (resp. à un '1') dans cette ligne. La synthèse de r se ramène à celle d'un résidu pour la détection en effectuant la transformation suivante :

$$d \leftarrow [d \ f^0] \text{ et } f \leftarrow f^1 \quad (\text{III.1})$$

Sans perte de généralité, nous traitons donc maintenant la synthèse de résidus pour la détection. Nous présentons en parallèle trois formulations courantes de la génération de résidus pour les systèmes LTI :

- la première repose sur un observateur à entrées inconnues (OEI),
- la seconde fait appel à l'espace de parité (EP),
- la troisième est une approche par factorisation ou approche fréquentielle (AF).

Afin de présenter ces méthodes aussi bien pour des systèmes continus que pour des systèmes discrets, nous introduisons formellement un opérateur¹ δ (à la différence de [Goodwin, 1986] qui définit un opérateur δ particulier pour établir des liens entre modèles continus et modèles discrets, l'opérateur que nous utilisons est générique).

$$\text{Systèmes continus: } \delta = p, \quad \delta x = \dot{x}(t) = x^{(1)}(t), \quad \delta^n x = x^{(n)}(t) \quad (\text{III.2})$$

$$\text{Systèmes discrets: } \delta = z, \quad \delta x = x(k+1), \quad \delta^n x = x(k+n) \quad (\text{III.3})$$

Le modèle du système étudié est supposé linéaire et invariant dans le temps. Les observateurs à entrée inconnue et les résidus de l'espace de parité résultent d'un modèle d'état, alors que les méthodes de factorisation font appel à un modèle exprimé sous la forme de transferts. u, y, x, d, f représentent respectivement les entrées connues (consignes ou actions), les mesures, les états, les perturbations et les défauts.

$$\text{OEI: } \begin{cases} \delta x = A.x + B.u + E_d.d + E_f.f \\ y = C.x + D.u + F_d.d + F_f.f \end{cases} \quad (\text{III.4})$$

$$\text{EP: } \underline{y} = H_{o,s} \cdot \delta^{(-s)} x + H_{u,s} \cdot \underline{u} + H_{d,s} \cdot \underline{d} + H_{f,s} \cdot \underline{f} \quad (\text{III.5})$$

où:

s est l'ordre de l'espace de parité.

$$\underline{y} = \begin{bmatrix} \delta^{(-s)} y \\ \vdots \\ \delta^{(-1)} y \\ y \end{bmatrix}, \quad H_{o,s} = \begin{bmatrix} C \\ CA \\ \vdots \\ CA^s \end{bmatrix}, \quad H_{u,s} = \begin{bmatrix} D & 0 & \dots & 0 \\ CB & D & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ CA^{s-1}B & \dots & CB & D \end{bmatrix}$$

$\underline{u}, \underline{d}, \underline{f}$ sont définis de manière analogue à \underline{y} .

$H_{d,s}$ (resp. $H_{f,s}$) est défini de manière analogue à $H_{u,s}$: B est remplacé par E_d , (resp. E_f) et D est remplacé F_d (resp. F_f).

$$\text{AF: } y = H_u(\delta).u + H_d(\delta).d + H_f(\delta).f \quad (\text{III.6})$$

d et f étant modélisés sous la forme d'entrées additives, les modèles considérés se prêtent mal à la prise en compte des défauts paramétriques et des erreurs de modèle (perturbations sur les valeurs des paramètres). Les structures des différents générateurs de résidus sont :

¹ Cet opérateur correspond aux relations dynamiques *dyn* introduites dans le chapitre sur l'analyse causale.

$$\text{OEI : } \begin{cases} \delta v = N.v + M.u + G.y \\ r = -P.v - U.u + K.y \end{cases} \text{ avec } v \rightarrow T.x \quad (\text{III.7})$$

$$\text{EP : } r = W_s.(y - H_{u,s}.u) \quad (\text{III.8})$$

$$\text{AF : } r = Q_y(\delta).y + Q_u(\delta).u \quad (\text{III.9})$$

Reporter les expressions de y , (III.4) à (III.6), dans les structures des générateurs de résidus, (III.7) à (III.9), permet de mettre en évidence l'influence des différentes grandeurs sur r .

$$\begin{aligned} \text{OEI : } & e = v - T.x \quad (\text{III.10}) \\ & \begin{cases} \delta e = N.e + (NT - TA + GC).x + (M - TB + GD).u + (GF_d - TE_d).d + (GF_f - TE_f).f \\ r = -P.e + (KC - PT).x + (KD - U).u + KF_d.d + KF_f.f \end{cases} \end{aligned}$$

$$\text{EP : } r = W_s.H_{o,s}.\delta^{(-s)}x + W_s.H_{d,s}.\underline{d} + W_s.H_{f,s}.\underline{f} \quad (\text{III.11})$$

$$\text{AF : } r = (Q_y(\delta).H_u(\delta) + Q_u(\delta)).u + Q_y(\delta).H_d(\delta).d + Q_y(\delta).H_f(\delta).f \quad (\text{III.12})$$

Pour que (III.7) à (III.9) soient effectivement des générateurs de résidus, il faut garantir que r soit exactement nul lorsqu'il n'y a pas de perturbation ($d=0$), pas de défaut ($f=0$) et pas d'erreur de modèle ni de bruit de mesure. Les conditions pour avoir un générateur de résidus se déduisent donc de l'annulation de l'influence de x et de u sur le résidu. Certaines conditions de stabilité doivent également être vérifiées.

$$\begin{aligned} \text{OEI : Conditions de Luenberger:} & \quad (\text{III.13}) \\ & N \text{ stable} \\ & NT = TA - GC \quad PT = KC \\ & M = TB - GD \quad U = KD \end{aligned}$$

$$\begin{aligned} \text{EP : Espace de parité: } P_{o,s} &= \{W_s / W_s.H_{o,s} = 0\}, \quad \gamma = \dim(P_{o,s}) \quad (\text{III.14}) \\ & \text{Condition pour avoir un générateur de résidus: } W_s \in P_{o,s}. \\ & \text{Soit } N_{o,s} \text{ une base de } P_{o,s} \text{ et } \pi \text{ une matrice de pondération des vecteurs de la} \\ & \text{base. } \forall \pi \in \mathfrak{R}^{n_r \times \gamma}, W_s = \pi.N_{o,s} \text{ désigne } n_r \text{ résidus de l'espace de parité}^1. \end{aligned}$$

$$\begin{aligned} \text{AF : Condition pour avoir un générateur de résidu: } Q_u(\delta) &= -Q_y(\delta).H_u(\delta), \quad (\text{III.15}) \\ Q_y(\delta) \text{ stable, } Q_u(\delta) \text{ stable.} & \end{aligned}$$

III.3.1. Cas du découplage parfait

Après avoir découplé r de u et de x , un découplage parfait de r vis-à-vis des perturbations est étudié. Des conditions (nécessaires et suffisantes) d'existence d'un tel découplage et garantissant

¹ Une base de l'espace de parité est donc donnée par une base du noyau de $H_{o,s}^T$. La fonction *null* de Matlab permet d'obtenir directement une telle base : $N_{o,s} = \text{null}(H_{o,s}^T)^T$.

également une sensibilité non nulle à chacun des défauts f_i , $i=1\dots n_f$, sont données par (III.16) et (III.17) [Ding, 1994, 1999].

$$\text{EP} : \forall i=1\dots n_f, \text{rang}([H_{o,s} \ H_{d,s} \ H_{f_i,s}]) > \text{rang}([H_{o,s} \ H_{d,s}]) \quad (\text{III.16})$$

$$\text{AF} : \forall i=1\dots n_f, \text{rang}([H_d(\delta) \ H_{f_i}(\delta)]) = \text{rang}(H_d(\delta)) + 1 \quad (\text{III.17})$$

La condition (III.16) dépend de l'ordre de l'espace de parité s et reste valable lorsque s est égal à l'ordre n du système. En raison des liens existant entre les observateurs à entrée inconnue et les résidus de l'espace de parité, les conditions d'existence d'un découplage parfait sont identiques dans les deux cas ($s=n$) [Ding, 1999]. De plus, un corollaire à (III.17) est : il est impossible de découpler parfaitement plus de n_y-1 perturbations indépendantes, où n_y représente le nombre de mesures (Remarque : n_y correspond au nombre de lignes de $H_d(\delta)$). Lorsque les conditions d'existence sont satisfaites, une solution peut être trouvée aux conditions de découplage parfait données par (III.18) à (III.20).

$$\begin{aligned} \text{OEI} : \quad GF_d - TE_d &= 0 \\ KF_d &= 0 \end{aligned} \quad (\text{III.18})$$

$$\begin{aligned} \text{EP} : \quad P_{d,s} &= \{W_s / W_s \cdot H_{d,s} = 0\} \\ \text{Condition de découplage parfait de } d: \quad W_s &\in P_{d,s}. \end{aligned} \quad (\text{III.19})$$

$$\text{AF} : Q_y(\delta) \cdot H_d(\delta) = 0 \quad (\text{III.20})$$

III.3.2. Cas du découplage approximatif

Du fait des conditions d'existence assez restrictives du découplage parfait, des méthodes de découplage approximatif consistent non plus à annuler mais à minimiser l'influence des perturbations par rapport à l'influence des défauts sur le résidu. La minimisation est réalisée par rapport à un certain critère J . Par exemple :

$$\text{OEI} : \quad J = \min \frac{\|P \cdot (\delta I - N)^{-1} \cdot (T \cdot E_d - G \cdot F_d) + K \cdot F_d\|_\infty}{\|P \cdot (\delta I - N)^{-1} \cdot (T \cdot E_f - G \cdot F_f) + K \cdot F_f\|_\infty} \quad (\text{III.21})$$

$$\text{EP} : \quad J = \min_{\pi} \frac{\pi \cdot N_{o,s} \cdot H_{d,s} \cdot H_{d,s}^T \cdot N_{o,s}^T \cdot \pi^T}{\pi \cdot N_{o,s} \cdot H_{f,s} \cdot H_{f,s}^T \cdot N_{o,s}^T \cdot \pi^T} \quad (\text{III.22})$$

$$\text{AF} : \quad J = \min \frac{\|Q_y(\delta) \cdot H_d(\delta)\|_\infty}{\|Q_y(\delta) \cdot H_f(\delta)\|_\infty} \quad (\text{III.23})$$

Remarque 1 : Soit $G^*(\delta)$ la transposée hermitienne de la matrice de transfert $G(\delta)$: $G^*(\delta) = G^T(-\delta)$. Les valeurs singulières de $G(\delta)$ sont les racines carrées des valeurs propres de la matrice hermitienne $G^*(\delta)G(\delta)$. La plus grande valeur singulière ($\bar{\sigma}$) de $G(\delta)$, permet de définir une norme euclidienne sur l'ensemble des matrices de transfert ayant le même nombre de lignes

et de colonnes que $G(\delta) : \|G(\delta)\|_2 = \bar{\sigma}(G(\delta))$. La plus grande valeur singulière généralise la notion de gain dans la représentation de Bode au cas multivariable. La valeur maximale de $\bar{\sigma}(G(\delta))$ sur la plage des fréquences définit une autre norme, la norme infinie notée $\|G(\delta)\|_\infty$.

Remarque 2 : d'autres choix que ceux donnés par (III.21) à (III.23) sont possibles pour le critère associé à un découplage approximatif. Nous en citerons un, construit à l'aide de $\|G(\delta)\|_-$, valeur minimale sur la plage de fréquence de la plus petite valeur singulière non nulle ($\underline{\sigma}$) de $G(\delta)$.

$$\text{EP: } J = \min_{\pi} \frac{\bar{\sigma}(\pi.N_{o,s}.H_{d,s})}{\underline{\sigma}(\pi.N_{o,s}.H_{f,s})} \quad \text{AF: } J = \min \frac{\|Q_y(\delta).H_d(\delta)\|_\infty}{\|Q_y(\delta).H_f(\delta)\|_-} \quad (\text{III.24})$$

Alors que (III.23) représente la minimisation de l'influence maximale des perturbations sur l'influence maximale des défauts, (III.24) représente la minimisation de l'influence maximale des perturbations sur l'influence minimale des défauts.

Plusieurs méthodes ont été proposées pour résoudre le problème du découplage approximatif : l'une des plus connue est l'approche fréquentielle proposée par [Frank, 1994b]. Les approches robustes et de μ -synthèse relèvent également du découplage approximatif dans la mesure où elles permettent de régler un compromis entre sensibilité et robustesse [Patton, 1997b], [Sadrnia, 1997].

III.3.3. Liens entre les différentes méthodes

Des liens étroits existent entre les différentes méthodes présentées (OEI, EP, AF). Ainsi les résidus obtenus par l'espace de parité sont équivalents à des observateurs à réponse pile [Patton, 1991b], [Cassar, 1996]. Des formules pour passer d'un vecteur de l'espace de parité (de dimension s) à un observateur (d'ordre s) et réciproquement sont données dans [Ding, 1999]. Partant d'un observateur (stable), le transfert entre y et r donne le réglage $Q_y(\delta)$ d'un résidu de l'approche fréquentielle (le transfert entre u et r donne alors un transfert $Q_u(\delta)$ vérifiant (III.15)). Finalement, les différences entre les méthodes tiennent à la manière de paramétrer le résidu et les trois approches sont équivalentes pour le problème du découplage parfait [Ding, 1994]. Cependant, l'approche fréquentielle donne a priori des possibilités de réglage supplémentaires pour la réalisation d'un découplage approximatif du fait que l'ordre du générateur de résidu n'est pas limité à l'ordre du système.

III.4. Sensibilité et choix de seuils robustes aux incertitudes

III.4.1. Etude de sensibilité : introduction

Contrairement à l'analyse causale où la génération de résidu proposée est systématique et conduit à un nombre fini de résidus, les méthodes du type OEI, EP ou AF conduisent à une infinité de résidus : le choix de l'un d'entre eux est en général effectué sur des critères de

sensibilité. Compte tenu de la difficulté d'optimiser globalement la procédure de diagnostic¹, une partie des choix est arbitraire. En effet, lorsque la table de signature est libre a priori, la complexité du problème posé devient telle que sa résolution par un logiciel d'aide à la conception de systèmes de détection et de localisation s'impose [Kinnaert, 1994].

Une solution consiste alors à fixer arbitrairement les découplages à réaliser sur un résidu, en partant des conditions théoriques du découplage parfait, par exemple. Cependant, même si ces conditions sont vérifiées, cela ne signifie pas pour autant que les '0' et les '1' seront valables en pratique :

- **Discussion sur la validité d'un '0' dans une table de signature :** le découplage dit « parfait » ne l'est qu'à la condition que le système physique suive exactement le modèle utilisé pour calculer le découplage. Les inévitables erreurs de modèle conduisent néanmoins le résidu à être sensible même aux grandeurs découplées.
- **Discussion sur la validité des '1' dans une table de signature :** les conditions de découplage parfait ont été construites de telle sorte que les sensibilités aux entrées autres que celles à découpler soit non nulle. Cependant, pour une amplitude de défaut donnée, rien ne garantit que ces sensibilités soient suffisantes pour dépasser le niveau de bruit, l'influence des erreurs de modèles, etc...

Une étude de sensibilité s'avère donc nécessaire pour définir des seuils garantissant un certain degré de robustesse aux perturbations (modélisées ou non). Une telle robustesse étant garantie, l'étude de la sensibilité aux défauts permet également d'enrichir l'interprétation d'un dépassement de seuil.

III.4.2. Influence des différentes sources d'incertitude sur un résidu

Le formalisme de l'approche fréquentielle (AF) est utilisé pour présenter le principe de l'étude de sensibilité. Si l'on tient compte de la présence d'erreurs de modélisation, le modèle d'un système linéaire donné par (III.6) devient :

$$y = \tilde{H}_u(\delta).u + \tilde{H}_f(\delta).f + \tilde{H}_d(\delta).d + \varepsilon \quad (\text{III.25})$$

ε représente les perturbations non modélisées dont le bruit de mesure fait partie. Les paramètres du modèle représentant le système physique, $\theta = [\theta^1, \dots, \theta^{n_\theta}]$ ne sont pas exactement connus et peuvent varier dans le temps. Le modèle utilisé pour la génération de résidus utilise le vecteur des paramètres nominaux, θ_0 . Le comportement sans défaut du système est caractérisé par une variation « acceptable » des paramètres « réels » autour de leur valeur nominale ($\Delta\theta = \theta - \theta_0$). Une fonction de transfert H sujette à des variations de paramètres est indiquée par un tilde.

$$\begin{aligned} \tilde{H}(\delta) &= H(\theta, \delta) \text{ et } H(\delta) = H(\theta_0, \delta) & (\text{III.26}) \\ \Delta H(\delta) &= \tilde{H}(\delta) - H(\delta) \approx \sum_{i=1}^{n_\theta} \left. \frac{\partial H(\theta^i, \delta)}{\partial \theta^i} \right|_{\theta^i = \theta_0^i} \cdot \Delta \theta^i \end{aligned}$$

¹ Par optimisation globale, nous entendons l'optimisation conjointe de la génération des résidus et de la décision en vue d'aboutir au meilleur compromis (par rapport à un critère) entre la sensibilité aux défauts et la robustesse aux perturbations.

L'approximation linéaire permettant d'évaluer $\Delta H(\delta)$ dans (III.26) est à considérer avec précaution ; en effet, les variations des paramètres physiques n'interviennent en général pas de manière linéaire dans la fonction de transfert.

Remarque : Formellement, la dérivée d'un transfert par rapport à un paramètre peut être assez fastidieuse à calculer, en particulier lorsque le modèle est connu sous la forme d'une représentation d'état. Il est toutefois possible d'exprimer la sensibilité du transfert en fonction des sensibilités des matrices formant une représentation d'état (III.27).

$$\text{Si } H(\theta, \delta) = C(\theta).(\delta.I_n - A(\theta))^{-1}.B(\theta) + D(\theta) = [A, B, C, D] \quad (\text{III.27})$$

$$\text{Alors } \left. \frac{\partial H(\theta, \delta)}{\partial \theta^i} \right|_{\theta^i = \theta_0^i} = \left([A, I_n, C, 0]. \frac{\partial A}{\partial \theta^i}. [A, B, I_n, 0] + [A, \frac{\partial B}{\partial \theta^i}, C, 0] + [A, B, \frac{\partial C}{\partial \theta^i}, 0] + \frac{\partial D}{\partial \theta^i} \right) \Big|_{\theta^i = \theta_0^i}$$

En combinant (III.9) et (III.15), n'importe quel générateur de résidus pour un système stable peut se mettre sous la forme (III.28).

$$r = Q_y(\delta).(y - H_u(\delta).u) \quad (\text{III.28})$$

En reportant l'expression du modèle du système réel (III.25) dans (III.28), la sensibilité du résidu aux différentes sources d'incertitudes peut être analysée.

$$r = r_u + r_f + r_d + r_\varepsilon \text{ avec} \quad (\text{III.29})$$

$$r_u = Q_y(\delta).\Delta H_u(\delta).u \quad (\text{III.30})$$

$$r_f = Q_y(\delta).(H_f(\delta) + \Delta H_f(\delta)).f \quad (\text{III.31})$$

$$r_d = Q_y(\delta).(H_d(\delta) + \Delta H_d(\delta)).d \quad (\text{III.32})$$

$$r_\varepsilon = Q_y(\delta).\varepsilon \quad (\text{III.33})$$

Les erreurs paramétriques se répercutent sur le résidu en fonction de l'excitation sur les entrées. Ainsi, l'excitation due aux consignes (u) répercute sur le résidu les erreurs de modélisation représentées par $\Delta H_u(\delta)$. Même si un découplage parfait est réalisé ($Q_y(\delta).H_d(\delta)=0$), les perturbations continuent à influencer le résidu via les erreurs de modèle $\Delta H_d(\delta)$.

III.4.3. Choix des seuils

Le choix de $Q_y(\delta)$, de sorte que r soit le moins sensible possible à r_u , r_d et r_ε et le plus sensible possible à r_f , correspond à une robustesse dite « active » ; le découplage parfait en est un exemple. A contrario, la robustesse « passive » consiste, pour un résidu donné, à fixer un seuil tel que le test du dépassement de ce seuil par le résidu soit robuste aux différentes sources d'incertitude. Autrement dit, le choix du seuil doit être tel que son dépassement par le résidu ne puisse être que la conséquence de l'occurrence d'un défaut de f .

En pratique, le choix des seuils découle de la définition de domaines de valeurs pour les variations de paramètres ($\Delta\theta \in \psi_{\Delta\theta}$), pour les entrées connues ($u \in \psi_u$), pour les perturbations modélisées ($d \in \psi_d$) et pour les perturbations non modélisées ($\varepsilon \in \psi_\varepsilon$). Ces domaines sont choisis en fonction des connaissances sur le domaine d'utilisation du système physique dans le cadre de l'application étudiée. Par exemple, pour un entraînement électrique, les variations de paramètres

peuvent être due à l'échauffement et le couple de charge constitue une perturbation. Dans ces conditions, le seuil λ doit être choisi de façon à vérifier (III.34).

$$\forall(\Delta\theta, u, d, \varepsilon) \in \Psi_{\Delta\theta} \times \Psi_u \times \Psi_d \times \Psi_\varepsilon, \quad |r_u + r_d + r_\varepsilon| < \lambda \quad (\text{III.34})$$

Remarque : en évaluant $\Delta H_u(\delta)$ à l'aide de (III.26) et des valeurs maximales des variations de paramètres ($\max(\Delta\theta)$), un majorant de r_u peut être calculé en ligne et être utilisé pour définir un seuil adaptatif (III.30) [Frank, 1996].

Le seuil étant non nul, le résidu ne peut le dépasser si l'amplitude des défauts est trop faible. La meilleure sensibilité aux défauts est par conséquent obtenue lorsque λ est minimal, tout en satisfaisant (III.34). En théorie, toutes les combinaisons possibles d'incertitudes devraient être prises en compte, avec tous les profils possibles. En pratique, cela est irréalisable. Une solution consiste alors à sommer des seuils élémentaires établis pour chaque incertitude avec un profil donné (échelon,...).

$$\lambda = \underbrace{\sum_{i=1}^{n_\theta} \lambda_u^i}_{|r_u| < \lambda_u} + \underbrace{\sum_{i=1}^{n_d} \lambda_d^i}_{|r_d| < \lambda_d} + \underbrace{\sum_{i=1}^{n_\varepsilon} \lambda_\varepsilon^i}_{|r_\varepsilon| < \lambda_\varepsilon} \quad (\text{III.35})$$

Calculé à partir de la somme des valeurs maximales de l'influence de chacune des sources d'incertitude, λ est un seuil robuste au pire des cas, c'est-à-dire lorsque les influences de toutes les perturbations (additives, paramétriques,...) sont maximales au même instant. Bien évidemment, le prix à payer pour être robuste à ce cas, très peu probable, est une faible sensibilité aux défauts. Cependant, la complexité du choix du seuil est grandement réduite du fait que les influences de chacune des incertitudes sont étudiées séparément.

III.5. Application à la MCC

III.5.1. Conditions de découplage parfait : comparaison avec l'analyse causale

Une première démarche pour guider la conception de générateurs de résidus consiste à étudier de manière systématique les conditions de découplage parfait pour la machine à courant continu. Les résultats obtenus sont alors comparés avec les découplages parfaits issus de l'analyse causale. Etant donné une représentation d'état de la machine à courant continu (III.36), les matrices de transfert correspondantes peuvent être déduites (III.37). Les valeurs des paramètres pour l'application étudiée sont données dans l'annexe B.

$$\begin{cases} \dot{x} = A.x + B.u + E_d.d + E_f.f \\ y = C.x + D.u + F_d.d + F_f.f \end{cases} \quad (\text{III.36})$$

où:

$$u = I_{ref}, \quad x = [X_I \quad I \quad \Omega]^T, \quad y = [\beta_m \quad I_m \quad \Omega_m]^T, \quad d = d_\Gamma, \quad f = [f_\beta \quad f_I \quad f_\Omega \quad f_U]^T,$$

$$A = \begin{bmatrix} 0 & -1 & 0 \\ k_I.U_c/L & -(k_P.U_c + R)/L & -k/L \\ 0 & k/J & \rho/J \end{bmatrix}, \quad B = \begin{bmatrix} 1 \\ k_P.U_c/L \\ 0 \end{bmatrix}, \quad C = \begin{bmatrix} k_I & -k_P & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}, \quad D = \begin{bmatrix} k_P \\ 0 \\ 0 \end{bmatrix}$$

$$E_d = \begin{bmatrix} 0 \\ 0 \\ -1/J \end{bmatrix}, \quad F_d = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}, \quad E_f = \begin{bmatrix} 0 & -1 & 0 & 0 \\ U_c/L & -k_P.U_c/L & 0 & 1/L \\ 0 & 0 & 0 & 0 \end{bmatrix}, \quad F_f = \begin{bmatrix} 1 & -k_P & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix}$$

$$H_u(p) = C.(pI - A)^{-1}.B + D \quad (\text{III.37})$$

$$H_d(p) = C.(pI - A)^{-1}.E_d + F_d$$

$$H_{f_i}(p) = C.(pI - A)^{-1}.E_{f_i} + F_{f_i}, \quad i = 1 \dots n_f$$

Pour l'étude des conditions du découplage parfait, les n_f défauts et les n_d perturbations considérés sont de même nature. Il est donc possible de tester systématiquement les conditions de découplage parfait aux $2^{(n_f+n_d)} - 1$ structures envisageables¹ pour la ligne d'une table de signature. Ainsi, l'ensemble des lignes d'une table de signature réalisables par découplage parfait peut être déterminé. Les calculs de rang nécessaires ont été réalisés à l'aide d'un logiciel de calcul formel. Le résultat obtenu ne tient donc pas compte d'éventuelles singularités liées aux valeurs particulières des paramètres du modèle de la MCC. A la différence de l'analyse causale, la nature des relations (additions, multiplications, etc...) est cependant prise en compte.

	f_β	f_I	f_Ω	f_U	d_Γ	Analyse causale
L_1	0	1	1	1	0	* Exemple: \hat{r}_{I_m}
L_2	1	0	0	0	0	* Exemple: \hat{r}_{β_m}
L_3	1	1	1	1	0	*
L_4	0	0	1	1	1	*
L_5	0	1	0	1	1	*
L_6	0	1	1	0	1	* Exemple: \hat{r}_{Ω_m}
L_7	0	1	1	1	1	(a)
L_8	1	0	1	1	1	*
L_9	1	1	0	1	1	*
L_{10}	1	1	1	0	1	(b)
L_{11}	1	1	1	1	1	*

Table III.2 : Lignes d'une table de signature réalisables par découplage parfait

La Table III.2 contient toutes les lignes d'une table de signature réalisables par découplage parfait. Une étoile (*) indique qu'au moins un résidu ayant des propriétés de découplage similaires a été obtenu à partir de l'analyse causale de la MCC (Table II.8 et Table II.9). Dans deux cas, (a) et (b), aucun résidu issu de l'analyse causale n'a permis de réaliser la structure de découplage parfait correspondante. Néanmoins, une combinaison linéaire des résidus locaux

¹ Une ligne qui ne contiendrait que des zéros est toujours réalisable ($r=0$) mais n'apporte rien pour le diagnostic.

permet de construire deux résidus, r_a et r_b , correspondant aux spécifications des cas (a) : (III.38) et (b) : (III.39).

$$(a): r_a = \alpha_a \cdot \hat{r}_{I_m} + (1 - \alpha_a) \cdot \hat{r}_{\Omega_m} \quad (\text{III.38})$$

$$(b): r_b = \alpha_b \cdot \hat{r}_{\beta_m} + (1 - \alpha_b) \cdot \hat{r}_{\Omega_m} \quad (\text{III.39})$$

Les pondérations α_a et α_b doivent être choisies de façon à éviter une simplification de l'influence d'un même défaut sur les deux résidus locaux intervenant dans une combinaison linéaire. Plus généralement, la pondération α_a permet de régler l'influence relative sur r_a des défauts affectant \hat{r}_{I_m} et \hat{r}_{Ω_m} .

Cette possibilité de régler (moyennant certaines contraintes) les sensibilités des résidus aux défauts conduit à des résidus qui ne sont plus forcément homogènes aux mesures, comme l'étaient les résidus issus de l'analyse causale. Nous montrons maintenant sur un exemple simple (III.40), l'intérêt de tenir compte de la **nature** des relations entre variables et non plus seulement de l'**existence** de telles relations.

$$\begin{bmatrix} y_1 \\ \dots \\ y_2 \end{bmatrix} = \begin{bmatrix} H_{f,11}(p) & H_{f,12}(p) \\ \dots & \dots \\ H_{f,21}(p) & H_{f,22}(p) \end{bmatrix} \begin{bmatrix} f_1 \\ \dots \\ f_2 \end{bmatrix} \quad (\text{III.40})$$

Dans le cas du système (III.40), la seule prise en compte de l'existence des relations entre variables ne pourra conduire à une conclusion autre que « f_1 et f_2 sont couplés » ; l'analyse causale conduit en effet à deux modèles locaux sensibles chacun à f_1 ET à f_2 (l'un des modèles locaux est associé à y_1 , l'autre à y_2). Pourtant, la prise en compte de la nature des relations entre variables permet de construire par combinaison linéaire un résidu r insensible à f_1 et a priori sensible à f_2 .

$$r = \begin{bmatrix} H_{f,21}(p) & -H_{f,11}(p) \end{bmatrix} \begin{bmatrix} y_1 \\ \dots \\ y_2 \end{bmatrix} = (H_{f,21}(p) \cdot H_{f,12}(p) - H_{f,11}(p) \cdot H_{f,22}(p)) \cdot f_2 \quad (\text{III.41})$$

La combinaison linéaire (III.41) correspond en fait à une projection de $[y_1 \ y_2]^T$ dans l'orthogonal de la trajectoire induite par f_1 , ce qui réalise le découplage parfait de f_1 .

Cet exemple montre que l'information sur la nature des relations entre variables autorise des découplages que la seule existence des relations entre variables (information structurelle et causale) ne permet pas.

III.5.2. Etude des pertes de rang dans le modèle de la MCC

Comme nous l'avons vu dans le cas particulier des conditions de découplage parfait, les possibilités de réglage des sensibilités des résidus aux différents défauts sont limitées par les relations de dépendance linéaire au sein du modèle. Or, toute relation de dépendance linéaire dans un système se traduit par une perte de rang (le rang du système n'est plus maximum). Dans le cas de la MCC, une étude du rang des 2^5 combinaisons possibles pour les vecteurs colonne $H_{f_\beta}(p)$, $H_{f_I}(p)$, $H_{f_\Omega}(p)$, $H_{f_U}(p)$ et $H_{d_T}(p)$ montre que toutes les pertes de rang observées peuvent s'expliquer par les trois points suivants :

1. Le rang ne peut excéder le nombre de mesures, $n_y=3$, correspondant au nombre de lignes des vecteurs de transfert.
2. $\text{rang}([H_{f_I}(p) \ H_{f_U}(p) \ H_{d_T}(p)])=2$ (perte de rang = 1)
3. $\text{rang}([H_{f_\Omega}(p) \ H_{f_U}(p) \ H_{d_T}(p)])=2$ (perte de rang = 1)

Ainsi, deux relations de dépendance linéaire simples (perte de rang égale à un) suffisent à expliquer toutes les relations de dépendance linéaire au sein du modèle. Le point 2 peut s'interpréter de la manière suivante : pour toute variation de f_I (resp. f_U), il existe une variation simultanée de d_T et de f_U (resp. f_I) induisant le même comportement sur les mesures donc sur les résidus. Ainsi, une localisation robuste au couple, de f_I et f_U d'une part (point 2), de f_Ω et f_U d'autre part (point 3), n'est pas possible à moins d'ajouter un supplément d'information. Par exemple, ce supplément d'information peut se traduire par des hypothèses (valides si possible...) sur la forme des défauts : échelons, rampes, ... Les limites d'une localisation robuste au couple que nous venons d'évoquer sont conformes au contenu de la Table III.2 (lignes associées à un '0' pour d_T).

En se rappelant que le découplage parfait ne permet pas de découpler plus de $n_y-1=2$ entrées indépendantes, il est possible de retrouver, à partir des relations de dépendance linéaire élémentaires (points 1., 2. et 3.), l'ensemble des découplages parfaits possibles (Table III.2). Le découplage parfait apparaît alors comme un cas particulier de réglage des sensibilités consistant à en annuler certaines :

- **Aucun découplage parfait** (ligne L_{11} de la Table III.2) : aucune dimension n'étant « consommée » par le découplage parfait, il reste trois dimensions indépendantes pour régler les sensibilités. Cependant les sensibilités à f_I, f_U, d_T d'une part (point 2), à f_Ω, f_U, d_T d'autre part (point 3) et, par conséquent, les sensibilités à f_I, f_Ω, f_U et à f_I, f_Ω, d_T sont liées entre elles.

Enfin, en agissant indépendamment sur les sensibilités associées à deux entrées parmi f_I, f_Ω, f_U et d_T , les sensibilités aux deux autres sont nécessairement fixées.

- **Découplage parfait de deux entrées indépendantes** : la conclusion que nous venons d'établir s'applique à nouveau : si deux entrées parmi f_I, f_Ω, f_U, d_T sont découplées parfaitement (sensibilités fixées à zéro sur deux entrées indépendantes), alors les quatre entrées sont nécessairement découplées (ligne L_2 de la Table III.2). Lorsque l'une des deux entrées découplées est f_β , les entrées autres que celles découplées ont une sensibilité non nulle (lignes L_1, L_4, L_5, L_6 de la Table III.2). Par ailleurs, comme le nombre de dimensions indépendantes ne peut excéder le nombre de mesures ($n_y=3$) et comme deux dimensions indépendantes sont « consommées » par le découplage parfait, il ne reste plus qu'une seule dimension pour régler les sensibilités du résidu aux entrées non découplées ; ces sensibilités seront donc nécessairement liées.
- **Découplage parfait d'une seule entrée** : Comme deux entrées quelconques sont toujours indépendantes, le découplage parfait d'une entrée n'en induit pas d'autres (lignes L_3, L_7 à L_{10} de la Table III.2). Si l'entrée à découpler est parmi f_I, f_Ω, f_U, d_T , il ne reste plus qu'une dimension indépendante pour fixer les sensibilités du résidu aux trois autres. Si l'entrée à découpler est f_β , deux dimensions indépendantes permettent de régler les sensibilités du résidu à f_I, f_Ω, f_U et d_T .

Bilan : les pertes de rang sont porteuses d'une information plus riche que les conditions de découplage parfait : elles permettent d'étudier le nombre de dimensions indépendantes pour régler les sensibilités ; le découplage parfait apparaît alors comme un cas particulier correspondant au réglage à zéro de certaines sensibilités (sans se soucier des autres...).

Le fait que les sensibilités d'un résidu à deux défauts soient liées ne signifie pas pour autant qu'il n'y ait plus moyen d'agir sur l'influence relative des deux défauts. Nous allons illustrer cela à travers l'exemple donné par (III.42).

$$r = \begin{bmatrix} H(p) & g(p).H(p) \end{bmatrix} \begin{bmatrix} f_1 \\ f_2 \end{bmatrix} \quad (III.42)$$

r est un vecteur colonne (résidus), $H(p)$ est un vecteur de transfert et $g(p)$ est un transfert simple (mono entrée, mono sortie). Dans (III.42), les sensibilités des résidus de r à f_1 et à f_2 sont manifestement liées. Il est donc impossible de construire un résidu parfaitement découplé de f_1 sans que ce résidu soit également découplé parfaitement de f_2 . Supposons maintenant que la bande passante des filtres de $H(p)$ soit très large et que $g(p)$ soit un filtre passe-bas (PB). L'effet de f_2 sur les résidus se traduit nécessairement par des composantes basse fréquence, alors que l'effet de f_1 peut se traduire aussi bien par des hautes que par des basses fréquences, en fonction de l'excitation liée à la forme de f_1 . r est néanmoins aussi bien sensible à f_1 qu'à f_2 . De plus, il n'existe qu'une seule dimension pour régler la sensibilité aux défauts ; autrement dit, lorsque la sensibilité à un défaut est fixée, la sensibilité à l'autre l'est également. Un vecteur ligne de transferts, $Q(p)$ permet d'effectuer un tel réglage de sensibilité :

$$r_Q = Q(p).r = \begin{bmatrix} \overset{\text{PH}}{\swarrow} Q(p).H(p) & \underbrace{Q(p).g(p).H(p)}_{=0} \overset{\text{PB}}{\swarrow} \end{bmatrix} \begin{bmatrix} f_1 \\ f_2 \end{bmatrix} \quad (III.43)$$

Cependant, en choisissant pour $Q(p)$ un vecteur ligne de filtres passe-haut (PH), le nouveau résidu obtenu, r_Q , est pratiquement insensible à f_2 tout en restant sensible à f_1 , lorsque f_1 excite au moins en partie les hautes fréquences.

Moyennant une hypothèse sur l'excitation des entrées modélisant les défauts (ou, en généralisant, les perturbations), un découplage (approximatif) a été réalisé.

Cet exemple illustre les limites d'une localisation robuste au couple pour la MCC. En effet, le couple résistant agit sur les autres grandeurs par l'intermédiaire de l'équation mécanique $1/(J.p+\rho)$ ((II.6) et Figure II.6) qui filtre les dynamiques rapides (dynamiques électriques). Or, l'influence du couple résistant (basse fréquence) est liée à celle des défauts f_I et f_U , d'une part (point 2.), f_Ω et f_U , d'autre part (point 3.). Afin d'espérer localiser entre eux les défauts f_I, f_Ω, f_U le plus indépendamment possible du couple résistant, il est donc nécessaire d'atténuer les basses fréquences (filtrage passe-haut). La contrepartie à cette atténuation est que les défauts f_I, f_Ω et f_U doivent exciter les hautes fréquences pour sensibiliser les résidus correspondants. Si la localisation de défauts en échelon demeure possible, nous vérifierons par la suite que la localisation de défauts en rampe est beaucoup plus délicate. Ces difficultés de localisation robuste ne sont pas seulement liées aux limites inhérentes à toute méthode de génération de résidus, mais également à la nature même du système « machine à courant continu ».

III.5.3. Etude de sensibilité sur des résidus existants

Une étude de sensibilité commune à plusieurs jeux de résidus est désormais appliquée à la MCC. La démarche retenue pour cette étude est celle décrite dans la partie III.4. Dix huit résidus répartis en cinq familles ont été étudiés (Table III.3).

- **Modèle global (MG), r_1 à r_3** : ces résidus sont les résidus globaux définis au chapitre II. $r_1 = r_{\beta_m}^*$, $r_2 = r_{I_m}^*$, $r_3 = r_{\Omega_m}^*$.
- **Modèles locaux (ML), r_4 à r_8** : ces résidus sont construits à partir des modèles locaux et ont été définis au chapitre II. $r_4 = \hat{r}_{\beta_m}$, $r_5 = \hat{r}_{I_m}$, $r_6 = \hat{r}_{I_m(\beta_m)}$, $r_7 = \hat{r}_{I_m(\Omega_m)}$, $r_8 = \hat{r}_{\Omega_m}$.
- **Observateurs de Luenberger (OL), r_9 à r_{11}** : ces résidus sont construits sur le même modèle que les observateurs à entrée inconnue mais sans appliquer les conditions du découplage parfait. La synthèse de ces résidus est présentée dans [Courtine, 1997b].
- **Observateurs à entrée inconnue (OEI), r_{12} à r_{14}** : l'application des observateurs à entrée inconnue à la génération de résidus pour la MCC est détaillée dans [Courtine, 1997a]. Les objectifs poursuivis en terme de découplage sont indiqués dans la Table III.3 : un fond grisé (resp. blanc) correspond à '1' (resp. '0') dans la table de signature désirée. Les découplages souhaités font tous partie des découplages parfaits possibles indiqués dans la Table III.2 (r_{12} , r_{13} , r_{14} correspondent respectivement aux spécifications L_3 , L_4 et L_6 de la Table III.2).
- **Approche fréquentielle (AF), r_{15} à r_{18}** : ces résidus résultent d'une synthèse fréquentielle reposant sur un découplage approximatif optimal. La méthode de génération de résidus employée est présentée dans [Frank, 1994b]. L'application à la MCC est traitée dans [Courtine, 1997a]. Les critères d'optimisation utilisés pour construire r_{15} à r_{18} , similaires à (III.23), visaient à se rapprocher le plus possible d'une structure diagonale pour la table de signature associée à ces résidus (diagonale grisée dans la Table III.3).

α		$f_1 = f_\beta$	$f_2 = f_I$	$f_3 = f_\Omega$	$f_4 = f_U$	Déc. Parf. d_T
MG	$r_1 = r_{MG}^1$	27.4%	57.1%	>>	15.1%	Non
	$r_2 = r_{MG}^2$	36.3%	52.8%	>>	23.0%	Non
	$r_3 = r_{MG}^3$	>>	>>	17.2%	>>	Non
ML	$r_4 = r_{ML}^1$	1.3%	>>	>>	>>	Oui
	$r_5 = r_{ML}^2$	>>	39.8%	16.1%	6.7%	Oui
	$r_6 = r_{ML}^3$	41.9%	59.4%	25.1%	27.1%	Oui
	$r_7 = r_{ML}^4$	>>	59.2%	>>	14.7%	Non
	$r_8 = r_{ML}^5$	>>	>>	17.8%	>>	Non
OL	$r_9 = r_{OL}^1$	>>	21.4%	9.5%	4.2%	Oui
	$r_{10} = r_{OL}^2$	5.9%	>>	>>	>>	Oui
	$r_{11} = r_{OL}^3$	>>	73.5%	0.8%	>>	Non
OEI	$r_{12} = r_{OEI}^1$	23.3%	31.5%	13.5%	5.8%	Oui
	$r_{13} = r_{OEI}^2$	>>	>>	1.3%	15.9%	Non
	$r_{14} = r_{OEI}^3$	>>	68.3%	1.1%	>>	Non
AF	$r_{15} = r_{AF}^1$	6.0%	>>	>>	>>	Oui
	$r_{16} = r_{AF}^2$	31.1%	13.6%	>>	19.4%	Non
	$r_{17} = r_{AF}^3$	>>	>>	0.3%	32.1%	Non
	$r_{18} = r_{AF}^4$	>>	20.2%	4.6%	3.8%	Non

Table III.3 : Etude de sensibilité appliquée à la MCC

Pour chacun des résidus, l'influence de différentes sources d'incertitude est évaluée en simulation (voir III.4.2). Les valeurs maximales des réponses temporelles associées à ces

influences sont utilisées pour fixer des seuils conformément à (III.35) (dans le paragraphe III.4.3). Le seuil associé à un résidu r_i est noté λ_i . Les critères de robustesse retenus sont les suivants : les profils considérés pour les défauts, les écarts de paramètres et les perturbations sont des échelons ; la perturbation de couple a une amplitude maximale de 2Nm ; les paramètres électriques et mécaniques L , R , k , J et f varient chacun de 10% par rapport à leur valeur nominale, puis un échelon de 5A est appliqué sur la consigne de courant I_{ref} (n'oublions pas qu'en l'absence d'excitation sur les entrées, les écarts de paramètres ne se répercutent pas sur les résidus (III.30)).

A titre d'exemple, la Figure III.1 représente l'étude de sensibilité menée pour le résidu r_{11} . Chaque barre correspond à la valeur maximale de l'influence d'un défaut, d'une perturbation, etc... Le seuil λ obtenu par sommation est indiqué par une ligne horizontale, permettant ainsi d'évaluer visuellement la sensibilité du test de dépassement de seuil aux différents défauts. Manifestement, la sensibilité de r_{11} à f_Ω est bien supérieure à celle relative à f_j .

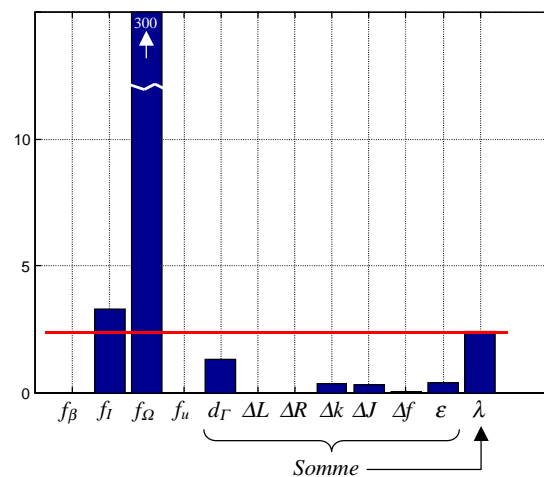


Figure III.1 : Etude de sensibilité pour le résidu r_{11}

Pour chacun des quatre défauts f_j , une amplitude maximale est définie, $\max(f_j)$. L'étude de sensibilité permet de calculer les amplitudes minimales des défauts (en échelon) pour que chaque résidu dépasse son seuil. En normalisant ces amplitudes minimales par $\max(f_j)$, on obtient les pourcentages reportés dans la Table III.3. Le symbole '>>' représente un pourcentage supérieur à 100%. Cela signifie que l'amplitude du défaut (f_j) conduisant à un dépassement de seuil est supérieure à $\max(f_j)$; autrement dit, le résidu ne dépassera jamais son seuil en réponse à f_j , quelle que soit l'amplitude de ce dernier ; le résidu peut donc être considéré comme insensible au défaut lorsque le symbole '>>' apparaît dans la Table III.3.

Un symbole '>>' dans une case blanche signifie que les objectifs initiaux de découplage du résidu ont bien été remplis. Un symbole '>>' dans une case grisée signifie que la sensibilisation d'un résidu à un défaut a échoué. Cette situation se rencontre uniquement avec les résidus issus de l'analyse causale. Ce constat est à rapprocher du fait que ces résidus sont construits sans possibilité de réglage de la sensibilité ; seule l'existence de relations entre variables est en effet prise en compte pour les générer.

Moyennant une hypothèse sur la forme des défauts (échelons), l'étude de sensibilité permet d'enrichir l'information donnée par un dépassement de seuil, en associant à ce dernier une amplitude minimale pour chacun des défauts.

Les résultats de l'étude de sensibilité (Table III.3) seront repris dans le chapitre V pour optimiser des fonctions de décision.

III.6. Liens entre les résidus de l'analyse causale et les observateurs à entrée inconnue

Le modèle d'un système est un cas particulier d'observateur à gain nul. Ainsi, les résidus globaux, correspondant à l'erreur de sortie d'un modèle parallèle, peuvent être vus comme des résidus issus d'un observateur à gain nul. Les résidus locaux sont quant à eux obtenus en injectant les mesures disponibles en entrée des modèles locaux : la structure de ces générateurs de résidus ressemble à s'y méprendre à celle d'un observateur (Figure III.2).

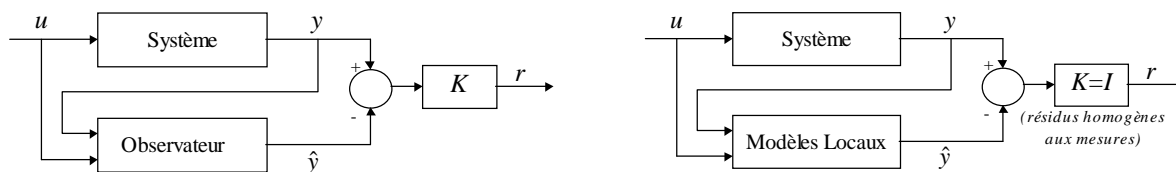


Figure III.2 : Comparaison des structures de générateurs de résidus

Compte tenu de cette analogie, nous cherchons à établir une condition pour qu'un résidu local corresponde à un résidu de Luenberger¹ pour le système global. Notre objectif est également d'explicitier les réglages d'un observateur du système **global** équivalent à un générateur de résidu **local**. Enfin, le réglage d'un observateur du système global conduisant à un résidu équivalent à la pondération de deux résidus locaux est donné. Le système étudié est supposé linéaire.

III.6.1. Préliminaires

Deux préliminaires vont nous permettre d'établir des résultats utiles à notre démarche. Le premier porte sur les observateurs de Luenberger d'ordre plein et le second, sur les liens entre les observateurs d'ordre réduit et les observateurs d'ordre plein. Même si les résultats de ces préliminaires sont assez classiques, il ne nous semble pas inutile de rappeler la (une) démarche conduisant à leur élaboration.

III.6.1.1. Observateurs d'ordre plein

La Figure III.3 rappelle la forme d'un observateur de Luenberger. L représente le gain d'adaptation et K , une combinaison linéaire de l'erreur de sortie de l'observateur.

¹ « résidu de Luenberger » : contraction de « résidu construit à l'aide d'un observateur de Luenberger »

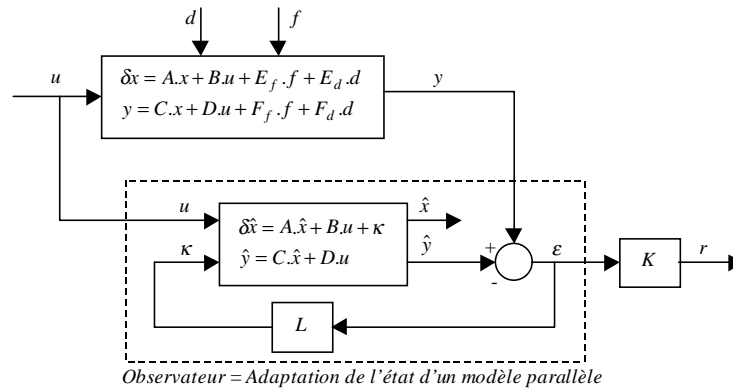


Figure III.3 : Observateur de Luenberger d'ordre plein

Remarque 1 : une adaptation différente d'un simple gain est parfois utilisée pour construire un observateur. Ainsi, le gain peut être remplacé par une structure proportionnelle intégrale (PI). Une application au diagnostic de tels observateurs est présentée dans [Linder, 1998].

Remarque 2 : dans l'approche fréquentielle proposée par Frank, les degrés de liberté supplémentaires par rapport à l'observateur de la Figure III.3 viennent du fait que K n'est plus un gain mais un transfert.

Un changement de base dans la représentation d'état de l'observateur de la Figure III.3 ne change pas les transferts décrivant le générateur de résidu associé.

$$\begin{aligned}
 \delta \hat{x} &= A.\hat{x} + B.u + L.(y - \hat{y}) & v &= T.\hat{x}, \quad T \text{ inversible} & (III.44) \\
 \hat{y} &= C.\hat{x} + D.u & \Leftrightarrow \delta v &= T(A - LC)T^{-1}.v + T(B - LD).u + TL.y \\
 r &= K.(y - \hat{y}) & r &= -KCT^{-1}.v \quad -KD.u \quad + K.y
 \end{aligned}$$

En identifiant le membre de droite de (III.44) avec la forme générale d'un générateur de résidus donnée par (III.7), on retrouve les conditions de Luenberger données par (III.13) pour la synthèse d'un générateur de résidus ($G=TL$). Ainsi,

Les seuls degrés de liberté effectifs pour la synthèse d'un générateur de résidus d'ordre plein à partir des conditions de Luenberger (III.13) sont la matrice de gain L d'un observateur (d'ordre plein) et la combinaison linéaire sur l'erreur de sortie (K).

III.6.1.2. Observateurs d'ordre réduit

Soit $T \in \mathcal{R}^{p \times n}$ une matrice de rang plein en lignes ($p \leq n$ où n est l'ordre du système global) et T^* la matrice complémentaire canonique de T . T^* permet de compléter les lignes de T de façon à aboutir à une matrice de changement de base pour un observateur d'ordre plein (matrice carrée et inversible de dimension n); de plus, T^* est choisie telle que $T^*T^*T = I_{n-p}$. T^* est entièrement déterminée par :

$$\begin{bmatrix} T \\ T^* \end{bmatrix}^{-1} = \begin{bmatrix} T^+ & T^{*T} \end{bmatrix}, \text{ où } T^+ \text{ est l'inverse à droite de } T: T^+ = T^T (TT^T)^{-1} \quad (III.45)$$

T^* peut être obtenue par une décomposition en valeurs singulières de T :

$$T = U_T \cdot [S_T \ 0] \cdot V_T^T \rightarrow T^+ = V_T \cdot [S_T^{-1} \ 0]^T \cdot U_T^T, \quad T^* = [0 \ I_{n-p}] \cdot V_T^T \quad (\text{III.46})$$

U_T et V_T sont des matrices orthogonales : $U_T^T \cdot U_T = U_T \cdot U_T^T = I$ (idem pour V_T). T étant de rang plein en lignes, S_T est une matrice diagonale à éléments diagonaux non nuls (valeurs singulières). Compte tenu de (III.45), plusieurs relations relient T , T^+ et T^* :

$$\begin{bmatrix} T \cdot T^+ & T \cdot T^{*T} \\ T^* \cdot T^+ & T^* \cdot T^{*T} \end{bmatrix} = \begin{bmatrix} I_p & 0 \\ 0 & I_{n-p} \end{bmatrix} \quad \text{et} \quad T^+ \cdot T + T^{*T} \cdot T^* = I_n \quad (\text{III.47})$$

Un changement de base construit à partir de T et de sa matrice complémentaire canonique est désormais appliqué à l'observateur d'ordre plein de la Figure III.3 :

$$\begin{aligned} \text{Changement de base: } \begin{bmatrix} v \\ \tilde{v} \end{bmatrix} &= \begin{bmatrix} T \\ T^* \end{bmatrix} \cdot \hat{x} & (\text{III.48}) \\ \begin{bmatrix} \delta v \\ \delta \tilde{v} \end{bmatrix} &= \begin{bmatrix} T(A-LC)T^+ & T(A-LC)T^{*T} \\ T^*(A-LC)T^+ & T^*(A-LC)T^{*T} \end{bmatrix} \cdot \begin{bmatrix} \delta v \\ \delta \tilde{v} \end{bmatrix} + \begin{bmatrix} T(B-LD) \\ T^*(B-LD) \end{bmatrix} \cdot u + \begin{bmatrix} TL \\ T^*L \end{bmatrix} \cdot y \\ r &= \begin{bmatrix} -KCT^+ & -KCT^{*T} \end{bmatrix} \cdot \begin{bmatrix} \delta v \\ \delta \tilde{v} \end{bmatrix} - KD \cdot u + K \cdot y \end{aligned}$$

Nous utilisons maintenant ces résultats pour établir un lien entre les observateurs d'ordre réduit et les observateurs d'ordre plein. Soit un observateur d'ordre réduit satisfaisant les conditions de Luenberger (III.13) :

$$\begin{cases} \delta w = N \cdot w + M \cdot u + G \cdot y \\ r = -P \cdot w - U \cdot u + K \cdot y \end{cases} \quad \text{avec } w \rightarrow T \cdot x \quad (\text{III.49})$$

T est une matrice de rang plein en ligne. On pose :

$$L = T^+ \cdot G + T^{*T} \cdot \tilde{G} \quad (\text{III.50})$$

\tilde{G} est une matrice de dimension appropriée dont les éléments peuvent être fixés à une valeur quelconque. En reportant (III.50) dans l'observateur d'ordre plein (III.48) et en tenant compte à la fois des conditions de Luenberger (III.13) satisfaites par l'observateur d'ordre réduit et des relations données par (III.47), il vient :

$$\begin{aligned} \begin{bmatrix} \delta v \\ \delta \tilde{v} \end{bmatrix} &= \begin{bmatrix} N & 0 \\ (T^*A - \tilde{G}C)T^+ & (T^*A - \tilde{G}C)T^{*T} \end{bmatrix} \cdot \begin{bmatrix} v \\ \tilde{v} \end{bmatrix} + \begin{bmatrix} M \\ (T^*B - \tilde{G}D) \end{bmatrix} \cdot u + \begin{bmatrix} G \\ \tilde{G} \end{bmatrix} \cdot y & (\text{III.51}) \\ r &= \begin{bmatrix} -P & 0 \end{bmatrix} \cdot \begin{bmatrix} \delta v \\ \delta \tilde{v} \end{bmatrix} - U \cdot u + K \cdot y \end{aligned}$$

Il apparaît clairement que (III.51) et (III.49) sont équivalents en terme de fonction de transfert.

Bilan : pour tout générateur de résidus d'ordre réduit (III.49) satisfaisant les conditions de Luenberger (III.13), il existe une infinité d'observateurs de Luenberger d'ordre plein (Figure III.3) conduisant à un générateur de résidus équivalent en terme de transfert.

Choisir $L = T^+ . G + T^{*T} . \tilde{G}$ où les éléments de \tilde{G} sont fixés arbitrairement permet de synthétiser un tel observateur. L'observateur étant réglé, la combinaison linéaire sur les erreurs de sortie conduisant au résidu est choisie identique à celle du générateur de résidu d'ordre réduit (K).

Corollaire : dans l'optique d'une optimisation de la génération de résidus, il est possible de choisir un observateur de Luenberger d'ordre plein comme structure pour le générateur de résidu (Figure III.3). Les gains matriciels L et K étant choisis comme paramètres de réglage, tous les générateurs de résidus de Luenberger (d'ordre plein et d'ordre réduit) font partie de l'espace des solutions. En particulier, tous les observateurs à entrée inconnue font également partie de l'espace des solutions. La difficulté d'une telle optimisation vient à la fois du choix du critère et du nombre de paramètres à optimiser (qui croît très rapidement avec l'ordre du système et le nombre de mesures).

III.6.2. Représentation du système

Ces préliminaires étant donnés, nous revenons maintenant à notre objectif initial qui était d'explicitier les réglages d'un observateur du système global correspondant aux générateurs de résidus locaux. Afin de refléter la décomposition de modèle issue de l'analyse causale, le modèle (linéaire) du système global est décomposé en deux parties : un modèle global et un modèle du reste du système (Figure III.4).

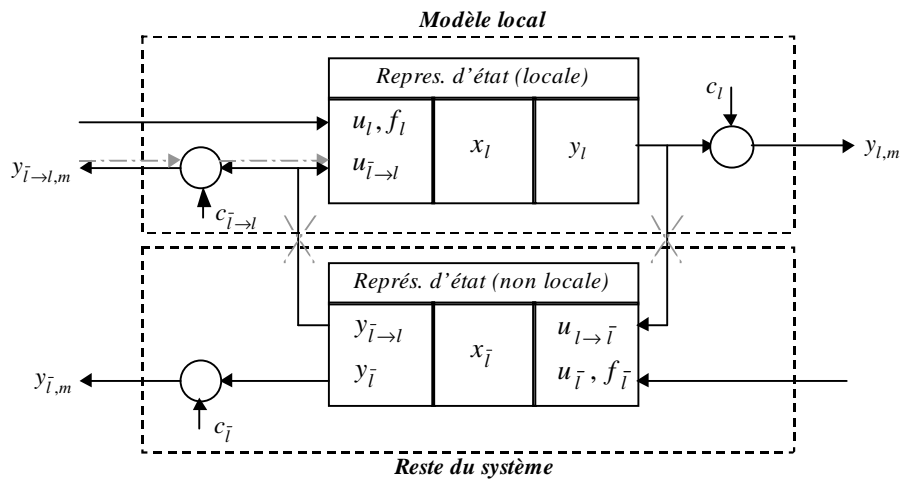


Figure III.4 : Représentation du système

Les notations suivantes sont utilisées dans la représentation du système (Figure III.4) : les vecteurs u, c, f, x et y représentent respectivement les entrées, les défauts capteurs non propagés, les autres défauts, les états et les sorties d'un sous-système. L'indice m correspond à une grandeur mesurée. L'indice l fait référence à un modèle local particulier, alors que l'indice \bar{l} fait référence au reste du système (non local). L'hypothèse de linéarité permet de trouver une

$$\begin{cases} \delta v = N.v + M.u + G.y \\ r = -P.v - U.u + K.y \end{cases}, r = \hat{r} \text{ lorsque :} \quad (\text{III.56})$$

$$\begin{aligned} T &= [I \quad 0] \text{ d'où } v \rightarrow x_l \\ N &= A_l \quad M = [B_{l1} \quad 0] \quad G = [0 \quad [B_{l2} \quad 0]] \\ P &= C_l \quad U = [D_{l1} \quad 0] \quad K = [1 \quad [-D_{l2} \quad 0]] \end{aligned}$$

Une condition suffisante pour que (III.56) satisfasse les conditions de Luenberger (III.13) est donnée par (III.57).

$$\begin{cases} A_l \text{ est stable (i.e. le modèle local est stable)} \\ (D_{l2} \cdot D_{l11} = 0) \text{ et } (D_{l11} \cdot D_{l2} = 0) \end{cases} \quad (\text{III.57})$$

En pratique, la condition (III.57) est peu restrictive dans la mesure où le second terme de la condition est vrai si l'un des sous-systèmes 'local' ou 'reste du système' est strictement propre. Intuitivement, le second terme de (III.57) correspond à une situation où il n'y a pas simultanément une action directe de l sur \bar{l} et de \bar{l} sur l . Lorsque (III.57) est satisfaite, (III.56) définit un résidu (au sens des conditions de Luenberger) équivalent au résidu local associé à l . Ce résidu est construit avec un observateur d'ordre réduit, qui prend en compte seulement la dynamique du modèle local. La condition (III.57) est supposée vérifiée dans la suite.

Le générateur du résidu local pouvant être considéré comme un générateur de résidu de Luenberger d'ordre réduit, il existe une infinité de réglages d'un observateur de Luenberger d'ordre plein conduisant au même résidu. En appliquant les résultats du préliminaire sur les observateurs d'ordre réduit III.6.1.2 à (III.56), le réglage du générateur de résidu d'ordre plein dont la structure est donnée par la Figure III.3 peut être choisi comme suit :

$$L = \begin{bmatrix} 0 & [B_{l2} & 0] \\ * & [* & *] \end{bmatrix} \text{ et } K = [1 \quad [-D_{l2} \quad 0]] \quad (\text{III.58})$$

Il est intéressant de noter que le réglage de l'observateur d'ordre plein sur le système global (III.58) se construit directement à partir de la représentation d'état du modèle local¹. Les étoiles correspondent à des éléments pouvant être fixés arbitrairement, sans modification sur le résidu en terme de fonction de transfert. Le réglage (III.58) résume assez explicitement le principe de construction des résidus locaux : la combinaison linéaire K fournit un résidu homogène à la sortie du modèle local (élément égal à 1 dans K). Les '0' de L signifient que seules les mesures associées aux entrées du modèle local (B_{l2}) doivent être injectées dans la partie « locale » du système qui s'apparente à un observateur d'ordre réduit. L'analyse causale a montré que cela permettait de rendre le résidu indépendant des influences sur le reste du système ; ainsi, les éléments de L injectant des mesures dans la partie « reste du système » peuvent être fixés arbitrairement (étoiles *). De plus, si le résidu est découplé des influences sur le reste du système, les conditions de découplage des observateurs à entrée inconnue ont toutes les chances

¹ Cela s'explique par la représentation que nous avons choisie pour le système, dans laquelle les états du modèle local sont homogènes aux états correspondants du système global. Cette correspondance est obtenue implicitement du fait que le modèle global résulte de la connexion des sous-systèmes « local » et « reste du système ».

d'être vérifiées. En choisissant $f = f_l$, $d = f_{\bar{l}}$, il est en effet possible de montrer que les conditions de découplage des OEI (III.18) sont vérifiées. Par conséquent, le résidu local \hat{r} associé au modèle local l est en fait un résidu construit à l'aide d'un observateur à entrée inconnue qui réalise le découplage des entrées $f_{\bar{l}}$. \hat{r} n'est donc pas sensible à $f_{\bar{l}}$ (défauts sur le reste du système) et \hat{r} est a priori sensible à f_l (défauts sur le sous-système local) car $T.E_f - G.F_f \neq 0$ et/ou $K.F_f \neq 0$. En ce qui concerne les défauts capteurs non propagés, les conditions de découplage conduisent à (III.59) : \hat{r} est donc parfaitement découplé de $c_{\bar{l}}$ et sensible à c_l (défaut capteur sur la sortie du modèle local) et à $c_{\bar{l} \rightarrow l}$ (défaut capteur affectant les mesures en entrée du modèle local).

$$\begin{aligned} T.0 - G.I &= -G = \begin{bmatrix} 0 & [* & 0] \\ K.I & = K = \begin{bmatrix} * & [* & 0] \end{bmatrix} \end{bmatrix} \end{aligned} \quad (\text{III.59})$$

Ainsi, les propriétés de découplage des résidus locaux établies dans le chapitre II sur l'analyse causale sont également mises en évidence et vérifiées dans le cadre des observateurs à entrée inconnue : \hat{r} est sensible aux entrées exogènes non observées du modèle local (f_l , c_l , $c_{\bar{l} \rightarrow l}$) et est parfaitement découplé vis-à-vis des autres ($f_{\bar{l}}$, $c_{\bar{l}}$).

III.6.4. Intérêt du lien établi, perspectives

Le lien, que nous venons d'établir entre les résidus locaux et les résidus issus des observateurs à entrée inconnue, donne la structure d'un observateur d'ordre plein équivalent au générateur de résidu local. Même s'ils présentent des propriétés de découplage intéressantes, les résidus locaux sont limités par l'impossibilité d'effectuer un réglage de sensibilité. Moyennant des pondérations entre les résidus locaux, nous avons vu que certains réglages de sensibilité pouvaient être envisagés. En outre, ces pondérations permettent d'obtenir l'ensemble des structures possibles de découplage parfait dans le cas particulier de la machine à courant continu (partie III.5.1).

Parallèlement, une optimisation numérique de L et de K peut se révéler délicate compte tenu du nombre de paramètres impliqués : pour générer un résidu scalaire avec un générateur d'ordre n appliqué à un système dans lequel n_y grandeurs sont mesurées, le paramétrage de L et de K fait intervenir $(n+1).n_y$ paramètres (soit déjà 12 pour un petit système faiblement instrumenté comme une MCC ($n=n_y=3$)).

De ces contraintes vient l'idée de limiter le nombre de paramètres à optimiser en utilisant les résultats établis pour imposer une certaine structure à L et à K . Cela revient en quelque sorte à utiliser les résidus locaux comme une base de résidus élémentaires, ces résidus étant, par construction, liés à la structure du système.

A titre d'illustration, considérons un système décomposé en trois parties : modèle local l_a (d'ordre n_a), modèle local l_b (d'ordre n_b) et reste du système \bar{l} (d'ordre $n-n_a-n_b$). En supposant que les recouvrements entre l_a et l_b ne fassent intervenir aucune variable d'état, un observateur d'ordre plein peut être associé à chaque résidu local :

$$\text{Pour } l_a : T_a = \begin{bmatrix} I_{n_a \times n_a} & 0_{n_a \times n_b} & 0_{n_a \times (n-n_a-n_b)} \end{bmatrix}, \quad L = \begin{bmatrix} G_a \\ * \\ * \end{bmatrix}, \quad K = K_a \quad (\text{III.60})$$

$$\text{Pour } l_b : T_b = \begin{bmatrix} 0_{n_b \times n_a} & I_{n_b \times n_b} & 0_{n_b \times (n - n_a - n_b)} \end{bmatrix}, \quad L = \begin{bmatrix} * \\ G_b \\ * \end{bmatrix}, \quad K = K_b$$

Compte tenu du découplage entre les états induit par le réglage du gain associé à un observateur d'ordre réduit (III.51), compte tenu du fait que les éléments * peuvent être fixés librement, une combinaison linéaire des résidus locaux peut être obtenue à partir d'un seul observateur d'ordre plein réglé comme suit :

$$L = \begin{bmatrix} G_a \\ G_b \\ * \end{bmatrix} \text{ et } K = \alpha.K_a + (1-\alpha).K_b \Rightarrow r = \alpha.\hat{r}_a + (1-\alpha).\hat{r}_b \quad (\text{III.61})$$

Rappelons toutefois que (III.61) n'est valable que s'il n'y a pas de variable d'état commune à l_a et l_b (recouvrement des modèles locaux). Dans le cas de la machine à courant continu, n'importe quelle combinaison de découplage parfait peut désormais être obtenue par un réglage de L et de K déduit d'une décomposition de modèle. Un réglage de sensibilité faisant intervenir un petit nombre de paramètres est désormais possible. Ce réglage est toutefois limité à des pondérations scalaires : la pondération de (III.61) ne repose pas sur un filtrage comme dans une approche fréquentielle.

Même si de nombreux points restent à approfondir (recouvrement des modèles locaux faisant intervenir une ou plusieurs variables d'état), l'intérêt de ce travail est d'établir un lien entre la structure du système (au sens de l'analyse causale) et le réglage des observateurs. Les résultats obtenus pourraient sans doute être généralisés en traitant la génération de résidus par une approche géométrique. L'idée directrice consisterait à figer aussi judicieusement que possible une partie de la structure du générateur de résidus pour pouvoir appliquer une optimisation numérique dans de bonnes conditions (peu de paramètres à optimiser) et avec des critères plus diversifiés que les rapports de normes infinies.

III.7. Une méthode numérique de découplage approximatif

Le découplage approximatif pose la génération de résidus comme un problème d'optimisation visant à maximiser l'influence des défauts et à minimiser l'influence des perturbations. Si le principe est élégant d'un point de vue théorique, d'autant plus qu'il englobe le cas du découplage parfait, il n'est pas pour autant trivial à mettre en pratique. Ainsi, l'application d'une méthode de factorisation matricielle, telle que celle proposée par [Frank, 1994b] et appliquée à la MCC par [Courtine, 1997a], est assez lourde à mettre en œuvre¹. Dans son travail, Courtine a choisi d'optimiser les résidus de façon à approcher le plus possible une table de signature à structure diagonale (résidus r_{15} à r_{18} de la Table III.3). Or, compte tenu de l'étude que nous avons menée sur les pertes de rang, une telle structure ne peut être atteinte par découplage parfait mais, surtout, il s'agit de la structure la plus contraignante en terme de réglage de sensibilité². Les

¹ Ces difficultés sont mentionnées dans [Shen, 1998] (Un algorithme basé sur des opérations matricielles élémentaires est proposé dans cet article).

² Malgré ces contraintes, une structure diagonale présente l'avantage d'autoriser le diagnostic de défauts multiples.

objectifs poursuivis dans cette partie sont multiples : appliquer une méthode numérique relativement simple de découplage approximatif à la MCC, profiter de cette relative simplicité afin de varier les structures désirées pour la table de signature et, pour finir, discuter les choix arbitraires effectués lors de la génération des résidus.

III.7.1. La méthode utilisée

La méthode que nous avons retenue pour le découplage approximatif est celle proposée par Ding [Ding, 1999]. Même s'il est possible de revenir à une expression récursive du générateur de résidu (observateur), la représentation utilisée pour la synthèse des résidus est celle de l'espace de parité. d représentant les entrées à découpler et f , les entrées à sensibiliser, le critère d'optimisation sur lequel repose le découplage approximatif est le suivant:

$$J = \min_{\pi} \frac{\overline{\sigma}(\pi.N_{o,s}.H_{d,s})}{\underline{\sigma}(\pi.N_{o,s}.H_{f,s})} \quad (\text{III.62})$$

L'algorithme utilisé pour la génération de résidus repose sur le théorème suivant [Ding, 1999]:

Si $\gamma = \text{rang}(N_{o,s}) = \text{rang}(N_{o,s}.H_{d,s})$ et si $N_{o,s}.H_{d,s}$ est décomposée en valeurs singulières¹ sous la forme $N_{o,s}.H_{d,s} = U.[S \ 0].V^T$ où S est une matrice diagonale avec γ éléments diagonaux non nuls (les valeurs singulières),

Alors $\pi^* = S^{-1}.U^T$ résoud le problème d'optimisation (III.62).

Remarque 1: la condition sur le rang signifie que les entrées à découpler ont une influence sur tout l'espace de parité.

Remarque 2: S étant une matrice carrée d'ordre γ , π^* est une matrice dont chacune des γ lignes correspond à un résidu particulier.

Remarque 3: les résidus de l'espace de parité, équivalents à ceux d'un observateur à réponse pile, peuvent être filtrés. Aucune indication n'est donnée pour réaliser un tel filtrage.

III.7.2. Application à la MCC

Après des essais nombreux mais faciles à mettre en œuvre, cinq résidus ont été retenus pour diagnostiquer les défauts de la MCC ; l'objectif de robustesse au couple résistant a été constamment recherché, et les réglages effectués visent à faciliter la localisation :

	Choix du critère					s	l	z_p
	d_{Γ}	f_{β}	f_I	f_{Ω}	f_U			
r_1	0	1	0	0	0	3	8	0.99
r_2	0	0	0	1	1	3	3	0.97
r_3	0	0	1	0	1	3	8	0.9
r_4	0	0	0	1	0	3	7	0.9
r_5	0	0	1	0	0	3	2	-0.995

Table III.4 : Résidus synthétisés par découplage approximatif pour la MCC

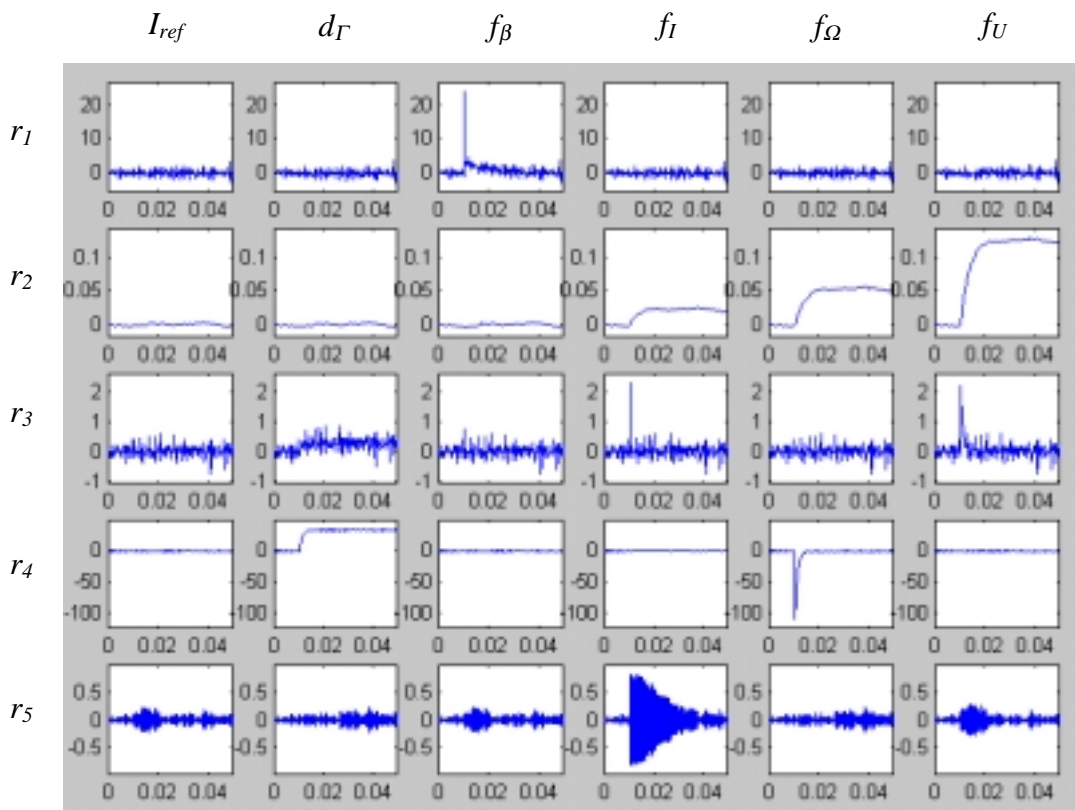
¹ Fonction *svd* du logiciel Matlab.

La Table III.4 rappelle les critères de synthèse pour chacun des cinq résidus retenus. L'ordre de l'espace de parité est choisi égal à l'ordre du système ($s=3$) afin de ne pas se priver de degrés de liberté dans les réglages. Même sur un petit système comme la MCC, des problèmes numériques peuvent apparaître. Ainsi, l'étude des conditions de découplage parfait (par l'analyse causale ou par une méthode formelle basée sur (III.17)) montre que la structure de r_l est réalisable ; pourtant, un calcul numérique du rang appliqué à (III.16) ne permet pas de retrouver l'existence de ce découplage parfait (La structure de r_l induit : $H_{o,s} \in \mathcal{R}^{12 \times 4}$, $H_{d,s} \in \mathcal{R}^{12 \times 16}$, $H_{f,s} \in \mathcal{R}^{12 \times 4}$).

Dans la Table III.4, l désigne la ligne de la matrice de pondération optimale, π^* , qui correspond au résidu sélectionné. Enfin, z_p représente le pôle discret d'un filtre $F(z)$ appliqué au résidu de parité.

$$F(z) = \frac{1 - z_p}{z - z_p} \tag{III.63}$$

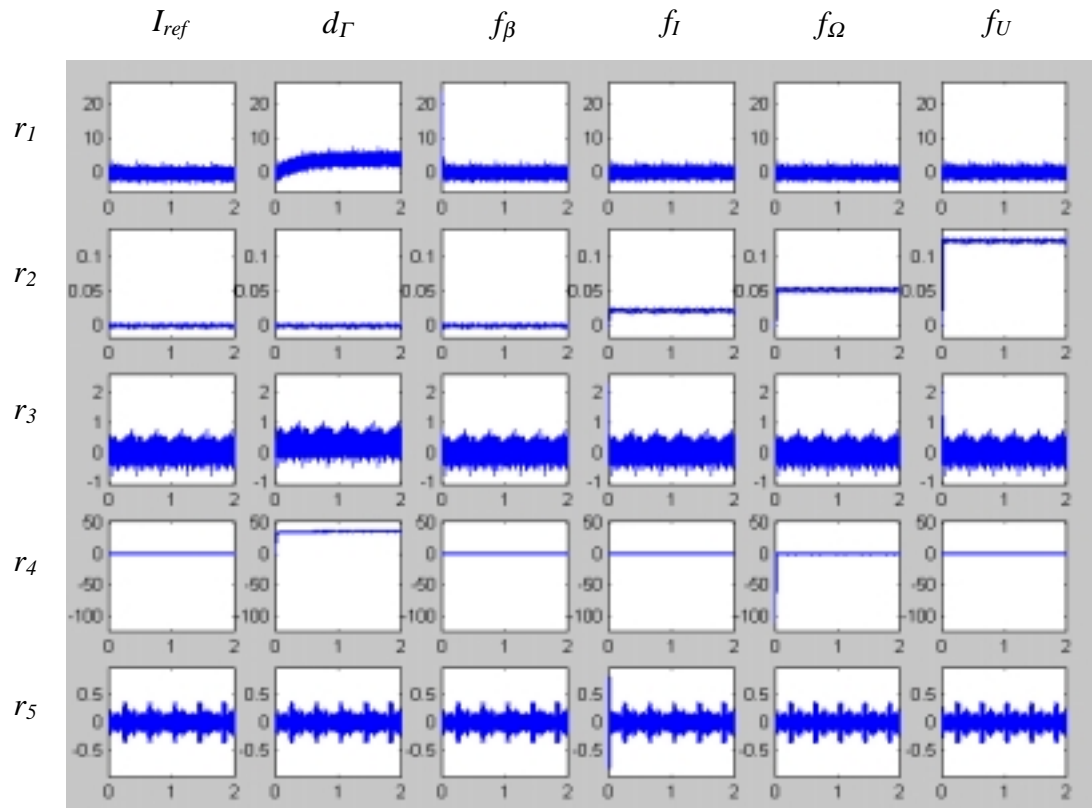
Bien que le réglage de ce filtre ait une influence parfois considérable sur les sensibilités relatives du résidu aux différentes entrées, rien ne permet de guider son réglage dans la méthode proposée par Ding. Notons que c'est toujours le cas lors d'une génération de résidus dans l'espace de parité (équivalence avec un observateur à réponse pile) et que certaines méthodes de synthèse des observateurs à entrée inconnue [Courtine, 1997a] consistent à fixer, avant tout autre réglage, les pôles de l'observateur.



A $t=0.01s$, I_{ref} : échelon 5A, d_Γ : échelon 20Nm, f_i : échelon 10%.max(f_i)

Figure III.5 : Découplage approximatif (défauts en échelon, échelle de temps courte)

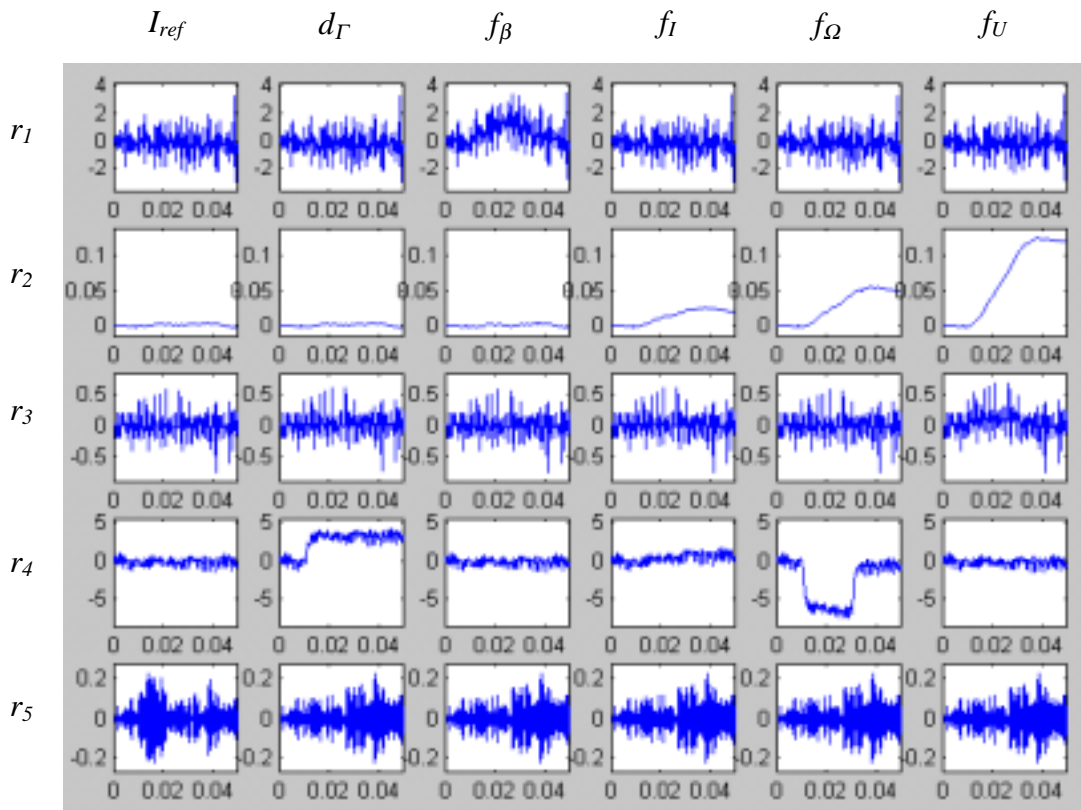
La Figure III.5 met en évidence la possibilité d'une localisation des défauts en échelon, tout en garantissant une robustesse intéressante aux variations du couple de charge (20Nm). Le seul réglage qui nous ait permis de rendre r_5 presque exclusivement sensible à f_I a consisté à choisir un pôle discret alterné pour le filtre du résidu de parité (Table III.4). Cela souligne une fois de plus l'importance du choix des pôles pour le réglage des sensibilités relatives.



A $t=0.01s$, I_{ref} : échelon 5A, d_{Γ} : échelon 20Nm, f_i : échelon $10\% \cdot \max(f_i)$

Figure III.6 : Découplage approximatif (défauts en échelon, échelle de temps longue)

Nous avons constaté deux dynamiques très différentes sur les résidus issus de l'analyse causale : une dynamique électrique (rapide) et une dynamique mécanique (très lente par rapport à la première) ; les résidus issus de l'analyse causale ont en effet les mêmes dynamiques que celles du système, du fait qu'ils résultent de l'erreur de sortie d'un modèle parallèle. La Figure III.6 permet de vérifier que les résultats de la Figure III.5 ne sont pas remis en cause sur une échelle de temps longue (les réponses transitoires de la Figure III.5 sont complètement écrasées sur l'axe des ordonnées de la Figure III.6 compte tenu de l'échelle de temps).



A $t=0.01s$, I_{ref} : échelon 5A, d_T : échelon 2Nm, f_i : rampe montant à 10%.max(f_i) en 0.02s

Figure III.7 : Découplage approximatif (défauts en échelon, échelle de temps longue)

Si la Figure III.5 et la Figure III.6 avaient pu nous donner quelques espoirs sur les possibilités d'une localisation robuste au couple de charge, force est de constater que la Figure III.7 nous ramène à la dure réalité ! L'étude des pertes de rang dans le modèle de la MCC (partie III.5.2) avait mis en évidence une possible difficulté de localisation pour des formes de défaut excitant peu les hautes fréquences. La Figure III.7 montre une situation où la localisation de tous les défauts n'est plus possible : cette situation correspond pourtant à une situation où le couple résistant a été réduit de 20Nm à 2Nm et où les rampes modélisant les défauts n'ont rien de dérivées lentes (variations de 10% sur une durée de seulement 0.02s).

Même si certains aspects de la méthode employée sont discutables (choix arbitraire des pôles), les difficultés de localisation robuste au couple de charge sont en partie liées à la nature même de la MCC. Ce constat est confirmé par le fait que les autres méthodes de génération de résidus employées (y compris l'approche fréquentielle de Frank) conduisent à une situation similaire.

III.8. Bilan

III.8.1. Point de vue méthodologique

La décomposition de modèle du chapitre II permet de guider le choix de sous-systèmes sur lesquels la prise en compte de la nature des relations entre variables peut être utilisée pour générer des indicateurs de défauts (résidus) plus performants que ceux déduits de l'analyse causale. Lorsque le modèle d'un tel sous-système est linéaire et que les défauts et les

perturbations peuvent être modélisés sous la forme d'entrées additives, les méthodes de génération de résidus traitées dans ce chapitre s'appliquent. Malgré des conditions d'application assez restrictives, les degrés de liberté gagnés par rapport aux résidus de l'analyse causale permettent de régler les sensibilités des résidus et de réaliser des découplages (parfaits ou approximatifs) inaccessibles auparavant. De plus, les dynamiques de réponse des résidus ne sont plus nécessairement celles du système étudié : cela est particulièrement utile pour une exploitation ultérieure dans la décision, délicate si des dynamiques très différentes influent sur un même résidu (par exemple, dynamique électrique et dynamique mécanique). Dans ce chapitre, nous avons également établi des liens entre les résidus locaux et les réglages des observateurs et nous avons défini le cadre dans lequel pourrait s'inscrire une optimisation de la génération de résidus à partir d'une base de résidus élémentaires liés à la structure du système. Enfin, l'optimisation de la génération de résidus a été appliquée sous la forme plus classique d'une méthode de découplage approximatif. Au-delà des spécificités de la méthode employée, cette application met en évidence le nombre important de choix arbitraires au cours de la phase de génération des résidus : choix a priori de l'ordre et des pôles¹ du générateur de résidus, choix a priori de la structure désirée pour la table de signature. Or, lorsque la table de signature est libre a priori, la complexité du problème posé devient telle que sa résolution par un logiciel d'aide à la conception de systèmes de détection et de localisation s'impose ; l'interaction entre le concepteur humain et la machine est nécessaire pour pouvoir déterminer les compromis acceptables compte tenu de l'application envisagée [Kinnaert, 1994]. Ces propos illustrent, s'il y en avait besoin, les nombreux développements encore nécessaires pour aboutir à une optimisation globale dans la synthèse d'un système de diagnostic. En effet, les performances du diagnostic tiennent au moins autant à l'homogénéité de la combinaison entre différents modules (générations de résidus, décisions, ...) que dans les performances propres à chaque module. Sans négliger les difficultés d'une telle approche, cela nous conduit à penser que le choix de critères pertinents pour l'optimisation de la génération de résidus ne peut se faire sans considérer la procédure de décision qui leur sera appliquée.

III.8.2. Point de vue des applications

Plusieurs méthodes de génération de résidus ont été appliquées à la machine à courant continu (MCC) et ont été comparées au travers d'une étude de sensibilité qui servira dans le chapitre V à optimiser des fonctions de décision. Les limitations d'une localisation robuste au couple de charge ont été mises en évidence.

En ce qui concerne la machine asynchrone (MAS), l'analyse causale a montré que la robustesse au couple de charge pouvait être obtenue en injectant la mesure de vitesse dans les équations électriques. La vitesse (mécanique) variant lentement par rapport aux grandeurs électriques, les équations électriques peuvent être considérées en première approximation comme linéaires, même si elles doivent être adaptées avec la mesure de vitesse. La mesure de vitesse est ainsi « consommée » pour garantir la robustesse au couple de charge. Il reste donc les mesures de deux courants statoriques ($n_y=2$). Il est donc impossible de découpler parfaitement plus de $(n_y-1)=1$ perturbation (ou défaut). Sachant que certaines perturbations sont paramétriques (variations des résistances dues à l'échauffement), l'application à la MAS des techniques linéaires de découplage semblent donc fort compromise. Ce constat est d'ailleurs confirmé par les études des découplages possibles réalisées parallèlement par le LEG et le LAG d'une part, le

¹ Certaines méthodes de génération de résidus permettent cependant de fixer des contraintes sur le lieu des pôles [Liu, 1998].

LAP¹ d'autre part, dans le cadre de l'action incitative sur la MAS à laquelle nous avons participé [Cocquempot, 2000]. La génération de résidus pour la machine asynchrone repose donc essentiellement sur des techniques non linéaires : généralisation au cas non linéaire des techniques de l'espace de parité par l'utilisation des bases de Gröbner [Comtet Varga, 1999], observateurs grand gain (approche retenue par le CEGELY² dans le cadre de l'action incitative), techniques d'estimation paramétriques [Moreau, 1999]. Etant donné la difficulté de régler la sensibilité des résidus lors de l'application de méthodes de découplage direct, et ce, même dans le contexte linéaire, nous allons centrer notre étude sur les méthodes de découplage indirect dans le prochain chapitre.

¹ LAP : Laboratoire d'Automatique et de Productique (Bordeaux)

² Centre de génie électrique de Lyon.

Chapitre IV :

Tests de modèles

Chapitre IV :

Tests de modèles

IV.1. Introduction

Les conditions d'application restrictives des méthodes de génération de résidus étudiées au chapitre III nous conduisent à revenir sur les fondements du diagnostic afin d'explorer de nouvelles pistes pour la génération de résidus. Après quelques indications sur la modélisation des défauts et après avoir précisé la notion de découplage indirect dans un cadre général, l'application à des défauts modélisés sous la forme d'entrées additives est réalisée sur la machine à courant continu (MCC). L'application à des perturbations et à des défauts modélisés sous la forme de variations de paramètres est utilisée pour le diagnostic de la machine asynchrone (MAS).

IV.2. Retour sur les principes du diagnostic

IV.2.1. Modélisation des perturbations et des défauts

Les approches de diagnostic orientées bon fonctionnement présentent l'avantage de ne pas nécessiter une modélisation des défauts : le diagnostic consiste alors à tester le (bon) comportement de plusieurs sous-systèmes ; les performances de la localisation sont alors très liées à la décomposition du système de départ (chapitre II). Lorsque les couplages sont nombreux et/ou lorsque l'instrumentation est limitée, ce qui est le cas pour les entraînements électriques, il est bien souvent nécessaire de modéliser l'influence des perturbations et des défauts afin d'affiner les performances de la détection et de la localisation.

La modélisation du système peut alors aussi bien résulter d'un modèle de connaissance (modèle physique) que d'un modèle de comportement (boîte noire) ou d'une approche mixte (boîte grise). Notre étude porte plus sur le développement de méthodes d'exploitation de modèles existants que sur la modélisation du système proprement dite. Le peu de place que nous accordons à la modélisation des systèmes ne doit en aucun être interprété comme une minimisation de l'importance de cette étape qui est cruciale pour aboutir à une bonne localisation des défauts.

Une approche de modélisation mixte nous semble bien adaptée pour établir des modèles de diagnostic des entraînements électriques utilisables en ligne. Alors qu'un modèle physique du comportement normal est disponible (modèle de Park pour la MAS, par exemple), les influences

de perturbations telles que l'échauffement ou les défauts machine (courts-circuits de spires, ruptures de barres) sont beaucoup plus délicates à modéliser. Dans ces conditions, il peut être intéressant de construire un modèle décrivant l'évolution (le comportement) des paramètres du modèle physique, lorsque les différents défauts ou perturbations ont lieu. Un tel modèle peut résulter de plusieurs campagnes d'identification. Nous sommes alors dans le cadre d'une approche mixte : le modèle de connaissance du comportement normal est combiné à un modèle de comportement des défauts. Deux cas sont envisageables en fonction du type de connaissance disponible :

- Si des jeux de mesures correspondant à différentes situations de défaut sont disponibles, la méthode proposée s'applique directement (l'observation d'un « vrai » défaut étant souvent difficile à obtenir, il est parfois possible de recréer expérimentalement une situation proche d'un défaut réel et d'effectuer les relevés de mesure correspondants).
- Un modèle physique modélisant finement la machine, y compris les situations de défaut, (modèle du type éléments finis, par exemple) peut être utilisé à la place des relevés de mesure pour construire un modèle de diagnostic dont la complexité le rend utilisable en ligne [Schaeffer, 1999].

En partant des limites d'un modèle LTI¹ pour le diagnostic, nous allons présenter une démarche de découplage indirect adaptée à une classe de modèles beaucoup plus large. Une structure de modèle LTI peut en effet présenter un domaine de validité trop restreint pour de nombreux systèmes.

$$\begin{aligned} \dot{x} &= A.x + B.u + E_d.d + E_f.f \\ y &= C.x + D.u + F_d.d + F_f.f \end{aligned} \quad (\text{IV.1})$$

Les perturbations et les défauts étant modélisés par des entrées additives (d, f), la structure du modèle (IV.1) ne permet pas de modéliser simplement des perturbations et des défauts qui modifieraient la dynamique entrée-sortie (y/u) du système. Or, ce genre de situation est courant en pratique : la définition même d'une perturbation ou d'un défaut indique qu'il s'agit d'un écart² de comportement du système. La connaissance sur un comportement pouvant être décrite à l'aide d'un modèle, modéliser la classe des défauts du type f_i revient à définir une famille de modèles entrée-sortie (y/u). De manière générale, une famille de modèles peut être décrite par un modèle, $M(\theta)$, dans lequel certains paramètres, θ , sont libres sur un certain domaine Ψ . Ainsi, la famille de modèles associée à f_i peut être décrite par un modèle $M(\theta_i)$, $\theta_i \in \Psi_i$, θ_i est un paramètre (vectoriel) libre qui permet de décrire (parcourir) l'ensemble des comportements entrée-sortie (y/u) possibles, lorsqu'un défaut du type f_i , éventuellement combiné à une perturbation d , se produit. Le choix que nous avons retenu pour définir θ_i est : $\theta_i = [d, f_i]$. A chaque valeur particulière de θ_i correspond donc un comportement entrée-sortie (y/u) relatif à une perturbation d particulière et à un défaut f_i particulier.

Remarque : *une* famille de modèles peut représenter une infinité de comportements possibles. Ainsi, *un* modèle $M(\theta_i)$ peut décrire une famille infinie de modèles entrée-sortie (y/u).

¹ LTI : Linéaire et Invariant dans le Temps.

² Ecart considéré comme « normal » dans le cas d'une perturbation et comme « anormal » dans le cas d'un défaut. Pour plus de précisions, voir les définitions données dans le chapitre I.

IV.2.2. Découplage indirect, interprétation d'un test de modèle

Pour déterminer si le défaut f_i a pu se produire, on peut imaginer de tester la cohérence de chacun des modèles de la famille $M(\theta_i)$, $\theta_i \in \Psi_i$, avec les observations disponibles, $Y=[u \ y]$. Si la modélisation est valide, l'existence d'au moins un modèle de la famille cohérent avec les observations signifie que l'occurrence du seul défaut f_i est possible. Il s'agit bien d'une possibilité et non d'une certitude : un autre défaut ou une combinaison d'autres défauts peut en effet avoir un comportement identique à celui décrit par l'un des modèles de la famille associée à f_i . En supposant qu'un critère d'évaluation de la cohérence de $M(\theta_i)$ avec Y soit donné par $J(\theta_i, Y) \geq 0$ et que ce critère soit nul si la cohérence est parfaite (J est donc analogue à un résidu), on a :

$$\text{Modélisation OK} \Rightarrow ((f_i \text{ seul est possible}) \Leftrightarrow \exists \theta_i \in \Psi_i / J(\theta_i, Y)=0) \quad (\text{IV.2})$$

Il est évidemment impossible de tester la cohérence des observations avec l'infinité de modèles appartenant à une famille donnée. Ce constat s'applique a fortiori dans le contexte du diagnostic en ligne. En supposant que $J(\theta_i, Y)$ présente certaines propriétés (conditions de découplage indirect), une solution consiste à chercher à se « caler » sur le modèle de la famille le plus cohérent avec les observations disponibles. On est ainsi ramené à un problème d'optimisation :

$$\theta_i^* = \min_{\theta_i \in \Psi_i} J(\theta_i, Y) \quad (\text{IV.3})$$

$$\text{Modélisation OK} \Rightarrow ((f_i \text{ seul est possible}) \Leftrightarrow J(\theta_i^*, Y)=0)$$

Comme θ_i^* n'est pas connu a priori, la réalisation en ligne de ce que l'on peut appeler le test du modèle associé au défaut f_i se ramène d'une part au calcul d'une estimation $\hat{\theta}_i$ de θ_i^* pour minimiser $J(\theta_i, Y)$ et d'autre part à un test de la nullité de J (étape décisionnelle du test).

$$(\text{Modélisation OK}) \text{ et } (\hat{\theta}_i = \theta_i^*) \Rightarrow ((f_i \text{ seul est possible}) \Leftrightarrow J(\hat{\theta}_i, Y)=0) \quad (\text{IV.4})$$

Le test du critère est désormais exprimé en fonction de grandeurs connues ($\hat{\theta}_i, Y$). (IV.4) indique que l'interprétation que l'on peut faire d'un test de modèle est soumise à deux conditions qui doivent être simultanément remplies : la validité du modèle et la convergence de l'algorithme d'estimation. Ces conditions étant supposées remplies, l'interprétation du résultat du test est réalisée comme suit :

$$J(\hat{\theta}_i, Y)=0 \Rightarrow f_i \text{ seul est possible} \quad (\text{IV.5})$$

$$J(\hat{\theta}_i, Y) \neq 0 \Rightarrow f_i \text{ seul est impossible} \quad (\text{IV.6})$$

Remarque 1 : lorsque $J(\hat{\theta}_i, Y)=0$, on est certain d'être à l'optimum. Comme la famille de modèles associée à f_i contient en général tous les modèles décrivant le comportement normal du système ($\theta_i=[d, f_i=0]$), $\hat{f}_i=0$ correspond à une situation de comportement normal. Plus généralement, lorsque $\theta_i=[d, f_i]$ est identifiable et que $J(\hat{\theta}_i, Y)=0$, $[\hat{d}, \hat{f}_i]$ correspond aux valeurs réelles de d et f_i sous l'hypothèse d'occurrence du défaut f_i seul. Non seulement la détection et la localisation, mais aussi l'identification des défauts devient alors possible.

Remarque 2 : la situation « f_i seul est impossible » exclut le comportement normal (qui correspond à f_i seul avec $f_i=0$). Cette situation correspond donc à l'occurrence d'un seul défaut autre que f_i ou à une combinaison de plusieurs défauts. Sous l'hypothèse de défaut simple (HDS), toute combinaison impliquant simultanément plusieurs défauts est exclue.

Remarque 3 : pour utiliser $J(\hat{\theta}_i, Y) \neq 0$ comme un test de détection pour le défaut $f_{j \neq i}$ (entre autres), il faut garantir que l'occurrence de n'importe quel défaut du type f_j (en présence de n'importe quelle perturbation d), conduit à un comportement incohérent avec tous les modèles de la famille associée à f_i : $f_j \neq 0 \Rightarrow J(\hat{\theta}_i, Y) \neq 0$, ce qui est équivalent à $J(\hat{\theta}_i, Y) = 0 \Rightarrow f_j = 0$.

Hypothèses:	
<ul style="list-style-type: none"> • Validité des modèles, • Convergence de l'algorithme d'estimation, • $\theta_i = [d, f_i]$ est identifiable • Hypothèse de défaut simple (HDS) 	
Situation	Interprétations
$J(\hat{\theta}_i, Y) = 0$ et $\hat{f}_i = 0$	Comportement normal
$J(\hat{\theta}_i, Y) = 0$ et $\hat{f}_i \neq 0$	L'occurrence du défaut f_i (seul) est possible si l'occurrence de f_i a eu lieu, $[d, f_i] = [\hat{d}, \hat{f}_i]$ (identification de d et f_i) ($f_j \neq 0 \Rightarrow J(\hat{\theta}_i, Y) \neq 0$) \Leftrightarrow l'occurrence du défaut $f_{j \neq i}$ est impossible
$J(\hat{\theta}_i, Y) \neq 0$	L'occurrence du défaut f_i (seul) est impossible L'occurrence d' un défaut $f_{j \neq i}$ est certaine. ($f_j \neq 0 \Rightarrow J(\hat{\theta}_i, Y) \neq 0$) \Leftrightarrow l'occurrence du défaut $f_{j \neq i}$ est possible

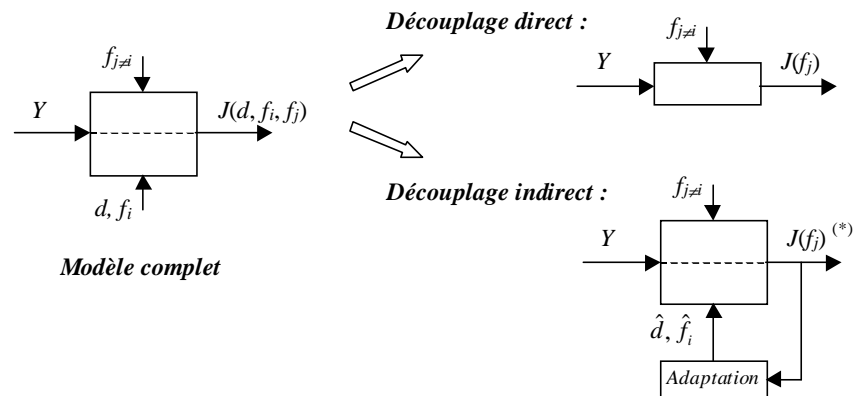
Table IV.1 : interprétation du test de modèle associé au défaut f_i .

La prise en compte des remarques 1, 2, 3 dans l'interprétation d'un test de modèle est résumée par la Table IV.1.

IV.2.3. Discussion sur le type de découplage

De par la manière dont il a été construit, le critère d'évaluation de la cohérence $J(\hat{\theta}_i, Y)$ reste nul lorsque le défaut f_i se produit, et ce, même en présence de perturbations d . Ce découplage vis-à-vis de $\theta_i = [d, f_i]$ est indirect : il résulte en effet d'une adaptation (estimation) de $\hat{\theta}_i = [\hat{d}, \hat{f}_i]$. $J(\hat{\theta}_i, Y)$ est par contre sensible a priori à tous les défauts $f_{j \neq i}$; cette sensibilité a priori doit néanmoins être vérifiée pour garantir que les défauts $f_{j \neq i}$ sensibilisent effectivement $J(\hat{\theta}_i, Y)$. Cette situation est similaire à celle du découplage direct décrite dans le chapitre III : une fois les conditions de découplage remplies, il fallait encore vérifier la sensibilité des résidus aux grandeurs non découplées. Un résidu et un critère du type $J(\hat{\theta}_i, Y)$ sont homogènes dans la mesure où ils servent chacun à évaluer la cohérence d'un modèle avec des observations.

Intuitivement, la famille de modèles associée à f_i peut être vue comme l'ensemble des comportements qu'il aurait fallu extraire du modèle complet pour aboutir à un sous-modèle induisant un découplage direct équivalent au découplage indirect de f_i . Lorsque les influences de f_i et de $f_{j \neq i}$ sont totalement couplées, il devient impossible de retirer l'influence de l'un sans retirer une partie au moins de l'influence de l'autre : le découplage direct ne peut plus être parfait. Un découplage direct approximatif résultant de l'optimisation d'un critère est alors recherché. Cette optimisation a lieu hors ligne, au stade de la conception, contrairement au découplage indirect, pour lequel l'optimisation a lieu en ligne et permet donc de « suivre » l'évolution système physique.



(*) En toute rigueur, on a : $J(d, f_i, f_j=0, Y)=0$ et $J(d, f_i=0, f_j, Y) \neq 0$ si f_j n'agit pas comme une combinaison de d et f_i . En remarquant qu'il est nécessaire d'avoir $f_j \neq 0$ pour que $J \neq 0$, nous avons simplifié les notations.

Figure IV.1 : Comparaison des découplages direct et indirect

Un modèle du système étant donné, il suffit qu'une procédure d'estimation puisse être mise en œuvre pour pouvoir réaliser un découplage indirect. En particulier, dès que le gradient du critère $J(\hat{\theta}_i, Y)$ peut être calculé, des méthodes de programmation non linéaire s'appliquent. Le découplage indirect reste donc valable pour une large classe de modèles.

L'inconvénient majeur des techniques d'estimation est qu'elles requièrent en général une excitation importante sur l'entrée pour assurer la convergence des algorithmes. Cela peut réellement limiter les possibilités d'implantation en ligne à moins de réduire considérablement le nombre de paramètres à identifier. Lors de la réalisation du découplage indirect d'un défaut spécifique, le nombre de paramètres à estimer est relativement restreint (le défaut et les principales perturbations). Ainsi, la difficulté supplémentaire de modélisation par rapport aux techniques reposant sur l'estimation de nombreux paramètres, non nécessairement représentatifs des perturbations et des défauts, trouve une contrepartie dans l'excitation requise.

Il est tout à fait possible de combiner les découplages direct et indirect. Deux types de combinaisons élémentaires peuvent être envisagés pour la réalisation des découplages :

- **Direct puis Indirect** : l'estimation du découplage indirect porte alors sur le modèle d'un sous-système obtenu par découplage direct.
- **Indirect puis Direct** : étant donné un générateur de résidus obtenu par découplage indirect, une technique de découplage direct (projection, par exemple) peut être appliquée à ces résidus pour en construire de nouveaux.

IV.3. Découplage indirect d'entrées additives

IV.3.1. Un exemple de découplage indirect : utilisation du filtre de Kalman

Une approche pour construire un estimateur consiste à implanter un filtre de Kalman. Le modèle du système est (dans un premier temps) un modèle LTI :

$$\begin{cases} \dot{x} = A.x + B.u + E_d.d + E_{f_i}.f_i + E_{f_a}.f_a \\ y = C.x + D.u + F_d.d + F_{f_i}.f_i + F_{f_a}.f_a \end{cases} \quad (\text{IV.7})$$

d désigne les perturbations, f_i l'hypothèse de défaut à tester et f_a les défauts autres que f_i . La famille des modèles décrivant les relations entrées / sorties (y/u) en présence de d et de f_i est donnée par :

$$M(\theta_i), \theta_i = [d, f_i] : \begin{cases} \dot{x} = A.x + B.u + E_{\theta_i}.\theta_i \\ y = C.x + D.u + F_{\theta_i}.\theta_i \end{cases} \quad \text{avec} \quad \begin{matrix} E_{\theta_i} = [E_d & E_{f_i}] \\ F_{\theta_i} = [F_d & F_{f_i}] \end{matrix} \quad (\text{IV.8})$$

Le découplage indirect de θ_i est réalisé par une estimation de cette grandeur qui peut être obtenue en étendant l'état du modèle (IV.8) et en construisant un filtre de Kalman.

$$\begin{cases} \begin{bmatrix} \dot{x} \\ \dot{\theta}_i \end{bmatrix} = \begin{bmatrix} A & E_{\theta_i} \\ 0 & 0 \end{bmatrix} \begin{bmatrix} x \\ \theta_i \end{bmatrix} + \begin{bmatrix} B \\ 0 \end{bmatrix} u + \varepsilon \\ y = \begin{bmatrix} C & F_{\theta_i} \end{bmatrix} \begin{bmatrix} x \\ \theta_i \end{bmatrix} + D.u + \varepsilon_y \end{cases} \quad (\text{IV.9})$$

$$E\{\varepsilon\} = 0, E\{\varepsilon.\varepsilon^T\} = Q = \begin{bmatrix} q_x.I & 0 \\ 0 & q_{\theta_i}.I \end{bmatrix}, E\{\varepsilon_y\} = 0, E\{\varepsilon_y.\varepsilon_y^T\} = R, E\{\varepsilon.\varepsilon_y^T\} = 0$$

(IV.8) indique la représentation adoptée pour décrire la famille de modèles associée à f_i en vue de l'implantation du filtre de Kalman. ε_y modélise le bruit de mesure dont la variance est réglée par une matrice diagonale, R . Deux paramètres, q_x et q_{θ_i} , servent à régler la variance du bruit sur l'état étendu, ε . Le choix de ces deux paramètres est capital pour respecter la « philosophie » du test de modèle : on cherche en effet à savoir si une variation de θ_i suffit à expliquer le comportement du système physique. La reconstruction de x est donc secondaire par rapport à celle de θ_i . Régler $q_{\theta_i} \gg q_x$ signifie que l'incertitude est beaucoup plus forte sur θ_i que sur x . Dans ces conditions, le filtre de Kalman, qui vise à minimiser l'incertitude sur l'état, agit en priorité sur θ_i ; la connaissance sur x est alors pratiquement déduite de la connaissance sur θ_i puisque l'incertitude sur x est choisie très faible. Ce réglage est donc conforme à l'idée d'expliquer le comportement du système physique majoritairement par les variations de θ_i .

Remarque 1 : une condition pour pouvoir reconstruire θ_i est que le système décrit par (IV.8) soit observable.

Remarque 2 : il n'y a rien de choquant à inclure des défauts et des perturbations dans l'état du système, bien au contraire : l'objectif du diagnostic est en effet d'obtenir un maximum

d'information sur l'état de défaut du système qui est inconnu a priori. La technique proposée permet de tester une structure particulière pour l'état de défaut du système (test de modèle).

Le générateur de résidu résultant de l'implantation du filtre de Kalman est donné par (IV.9).

$$\begin{cases} \begin{bmatrix} \dot{\hat{x}} \\ \dot{\hat{\theta}}_i \end{bmatrix} = \begin{bmatrix} A & E_{\theta_i} \\ 0 & 0 \end{bmatrix} \begin{bmatrix} \hat{x} \\ \hat{\theta}_i \end{bmatrix} + \begin{bmatrix} B \\ 0 \end{bmatrix} u + L.(y - \hat{y}) \\ \hat{y} = \begin{bmatrix} C & F_{\theta_i} \end{bmatrix} \begin{bmatrix} \hat{x} \\ \hat{\theta}_i \end{bmatrix} + D.u \end{cases}, \quad r = \frac{\zeta}{1 + \tau.s} |y - \hat{y}| \quad (\text{IV.10})$$

L représente le gain du filtre de Kalman en régime permanent, obtenu par la résolution d'une équation de Riccati¹. Le résidu scalaire r est obtenu par un filtrage passe-bas de l'innovation $(y - \hat{y})$. ζ est un vecteur ligne qui compense les différences d'ordre de grandeur entre les mesures. Le filtrage passe-bas permet de cumuler, avec un facteur d'oubli, les écarts sur l'innovation.

IV.3.2. Application à la MCC

Le modèle de la MCC exprimé sous la forme d'une représentation d'état (III.36) a la même structure que (IV.7). Dans un premier temps, une étude d'observabilité permet de déterminer les découplages indirects (impliquant un défaut) réalisables par la méthode décrite dans le paragraphe précédent. Rappelons que $d = d_r$ (couple résistant) et $f = [f_\beta \ f_I \ f_\Omega \ f_U]^T$ pour la MCC.

Découplage de:	Rang ² de la matrice d'observabilité de (IV.9)	
$\theta_i = [d_r \ f_\beta]$	$3 < 5$	Rang plein: NON
$\theta_i = [d_r \ f_I]$	$4 < 5$	Rang plein: NON
$\theta_i = [d_r \ f_\Omega]$	$4 < 5$	Rang plein: NON
$\theta_i = [d_r \ f_U]$	$4 < 5$	Rang plein: NON
$\theta_i = f_\beta$	$3 < 4$	Rang plein: NON
$\theta_i = f_I$	4	Rang plein: OUI
$\theta_i = f_\Omega$	4	Rang plein: OUI
$\theta_i = f_U$	4	Rang plein: OUI

Table IV.2 : Existence des découplages indirects par observateurs étendus pour la MCC

La Table IV.2 montre que la reconstruction simultanée du couple résistant et d'un défaut n'est pas possible. Le découplage indirect ne s'applique qu'aux défauts seuls, à l'exception de f_β , et sans garantir la robustesse au couple résistant.

Remarque 1 : l'observabilité est une condition suffisante d'existence du découplage indirect. Elle n'est pas forcément nécessaire : l'observabilité garantit en effet que l'état étendu, $[x \ d \ f_i]$ pourra être totalement reconstruit, et ce, de manière unique. S'il est en effet souhaitable de reconstruire la grandeur physique associée à f_i pour identifier ce défaut, la reconstruction de x et

¹ Fonction « kalman » du logiciel Matlab.

² Les calculs de rang ont été réalisés numériquement avec les valeurs des paramètres de notre application.

de d n'est pas a priori nécessaire. Une perte d'observabilité peut parfois conduire à ce qu'il existe une infinité de valeurs possibles pour (x, d) sans que la reconstruction de f_i soit affectée ; or, l'important est seulement de savoir si le critère de cohérence J est proche de zéro ou non. Parmi l'infinité de valeurs de (x, d) minimisant J , il suffit donc d'en avoir trouvé une quelconque.

Remarque 2 : une analyse de la perte d'observabilité pour le cas « $\theta_i = f_\beta$ » montre que cette dernière vient du fait que le défaut f_β agit de la même façon que l'état X_I sur le système (III.36). Il existe donc une infinité de couples (X_I, f_β) minimisant le critère de cohérence pour une situation réelle donnée. On peut également remarquer que X_I , état de l'intégrateur du correcteur PI, peut être parfaitement connu si la commande est intégrée dans le calculateur affecté au diagnostic de la MCC. En supposant que X_I soit parfaitement connu, il n'est plus nécessaire de le reconstruire, et toute l'adaptation peut porter sur f_β . La parfaite connaissance de X_I retire le degré de liberté qui conduisait à une infinité de couples (X_I, f_β) ; f_β est alors correctement estimé. Cette remarque est à mettre en relation à la fois avec la remarque précédente (perte d'observabilité) et avec le principe retenu pour le réglage du filtre de Kalman ($q_{\theta_i} \gg q_x$).

L'application à la MCC du découplage indirect de chacun des défauts seul permet de construire un jeu de résidu : r_β pour $\theta_i=f_\beta$, r_I pour $\theta_i=f_I$, r_Ω pour $\theta_i=f_\Omega$ et r_U pour $\theta_i=f_U$. La Figure IV.2 montre l'évolution temporelle de ces quatre résidus (dans un même système d'axe) pour différents scénarios de défauts (en échelon) et lorsqu'aucun défaut n'affecte le système. Pour un scénario de défaut donné, le résidu réalisant le découplage indirect de ce défaut est le plus faible des quatre (après un court transitoire) ; il rejoint rapidement la même valeur que celle qu'il aurait dans la situation « aucun défaut ». Cela illustre la propriété de découplage indirect : par exemple, le résidu r_Ω est très peu sensible à f_Ω et beaucoup plus sensible aux autres défauts. La présence d'un transitoire sur le résidu qui devrait être découplé est liée à la forme abrupte du défaut (échelon) : le découplage indirect résultant de l'estimation du défaut, la durée du transitoire correspond au temps de convergence de l'estimation. Ce point peut être vérifié sur la Figure IV.3.

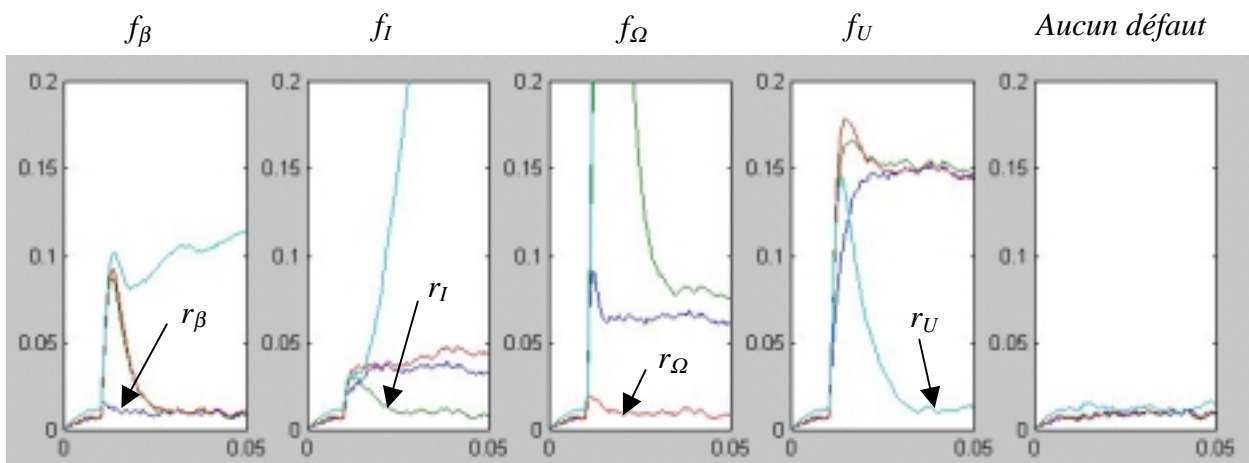


Figure IV.2 : Evolution des résidus en l'absence de couple résistant et avec $q_{\theta_i} \gg q_x$

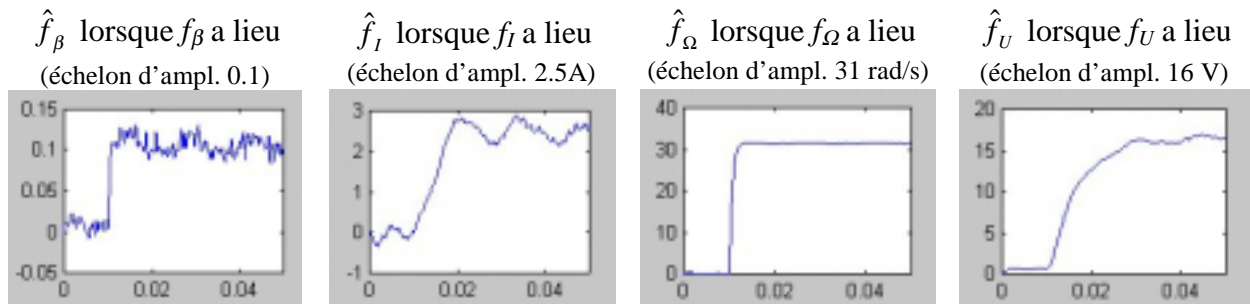


Figure IV.3 : Identification des défauts en l'absence de couple résistant et avec $q_{\theta_i} \gg q_x$

La Figure IV.3 montre qu'en l'absence de couple résistant et avec un réglage approprié du filtre de Kalman ($q_{\theta_i} = 10000 \gg q_x = 1$), les défauts sont correctement estimés. Même si f_β semble correctement reconstruit, aucune garantie n'existe sur la justesse de cette estimation (remarque 2). En présence d'un faible couple résistant (0.2 Nm), les propriétés de découplage indirect sont conservées pour tous les résidus, sauf pour r_U (Figure IV.4). La structure de la MCC est en effet telle qu'une robustesse parfaite au couple est impossible, ce qui, après l'étude des pertes de rang au chapitre III, est à nouveau confirmé par l'étude de l'observabilité (Table IV.2).

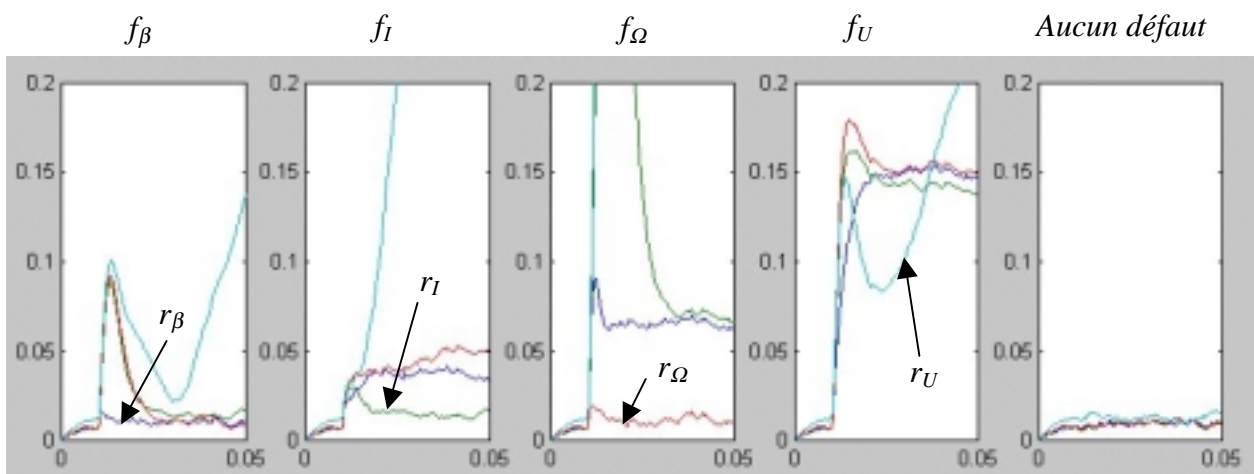


Figure IV.4 : Evolution des résidus en présence d'un couple résistant (0.2 Nm) et avec $q_{\theta_i} \gg q_x$

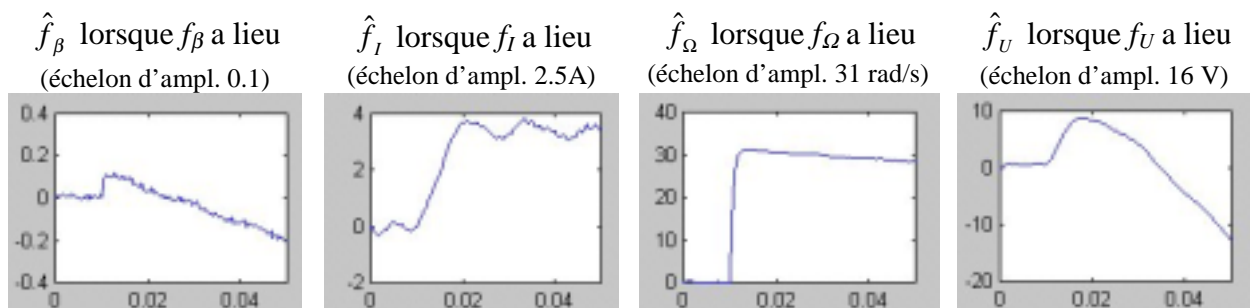


Figure IV.5 : Identification des défauts en présence d'un couple résistant (0.2 Nm) et avec $q_{\theta_i} \gg q_x$

La Figure IV.5 montre que les défauts ne sont pas reconstruits correctement en présence de couple : dans la mesure où l'estimation simultanée de d_I et d'un défaut n'est pas possible (Table IV.2), un seul défaut est estimé à la fois. En présence de couple, le découplage indirect essaie d'expliquer au mieux le comportement du système réel, y compris l'action du couple, en agissant presque exclusivement sur l'estimation du défaut ($q_{\theta_i} \gg q_x$). Il est important de noter que même si la reconstruction du défaut n'est pas parfaite, cela ne signifie pas pour autant que le découplage souhaité sur le résidu soit en échec : si le couple a une action sur le système proche de celle du défaut à découpler, l'estimation du défaut inclut certes l'action du couple, mais les propriétés de découplage attendues pour le résidu restent très bonnes ; toute action assimilable à celle du défaut à découpler est en effet compensée. On se retrouve alors dans une situation assimilable à un découplage approximatif (mais pas forcément optimal).

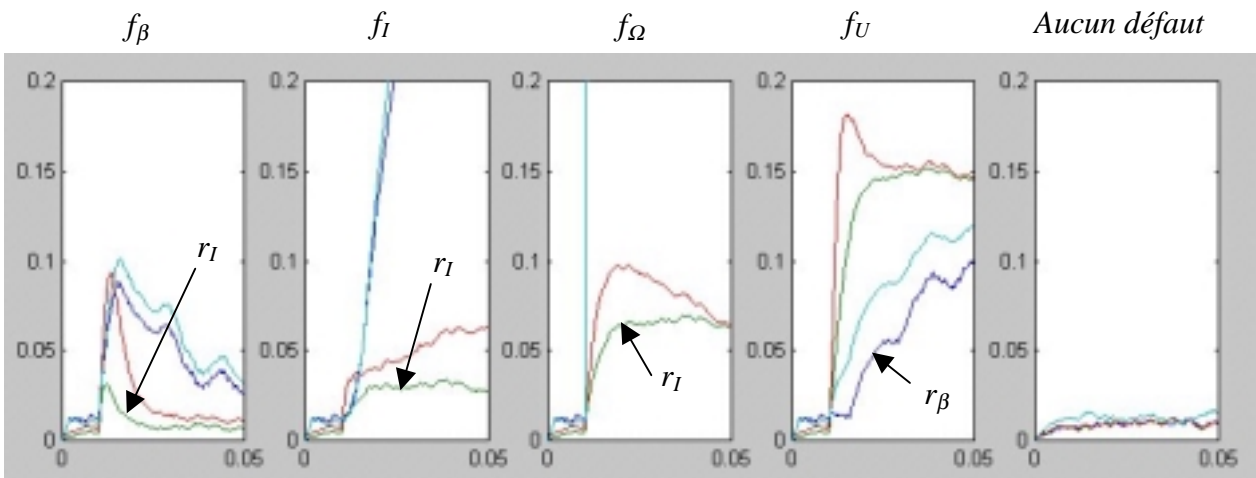


Figure IV.6 : Evolution des résidus en l'absence de couple résistant et avec $q_{\theta_i} \ll q_x$

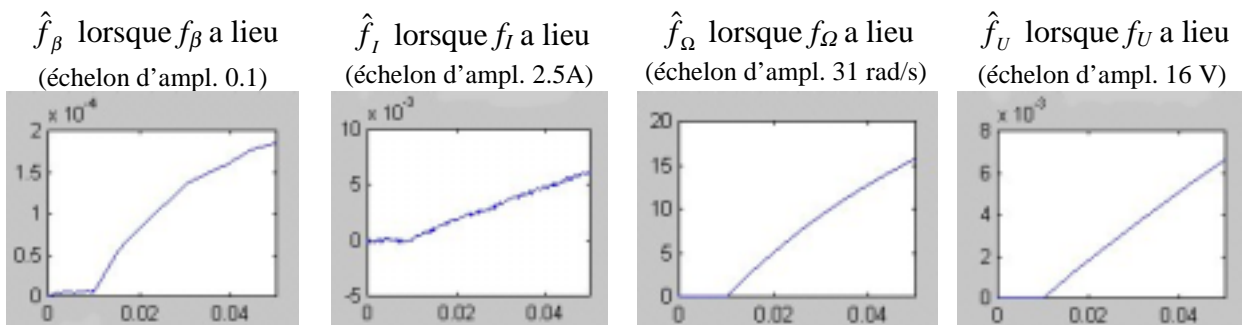


Figure IV.7 : Identification des défauts en l'absence de couple résistant et avec $q_{\theta_i} \ll q_x$

La Figure IV.6 et la Figure IV.7 illustrent l'influence du réglage du filtre de Kalman. Si l'estimation de l'état est privilégiée par rapport à celle du défaut à découpler ($q_{\theta_i} \ll q_x$ au lieu de $q_{\theta_i} \gg q_x$ précédemment), la Figure IV.6 montre que les propriétés de découplage attendues ne sont plus du tout vérifiées (sur une échelle de temps courte). Comme l'incertitude sur le défaut estimé est très faible, l'estimation de ce défaut est remise en cause très lentement comme l'atteste la Figure IV.7. Dans la mesure où l'estimation du défaut à découpler est incapable de « suivre » dans des temps raisonnables l'échelon modélisant le défaut réel, le découplage indirect de ce dernier échoue (Figure IV.6). Le mauvais réglage du filtre de Kalman amplifie donc considérablement les effets transitoires dus au temps de convergence déjà évoqués.

Bilan : Même si la robustesse au couple de charge reste une difficulté dans le cadre de l'application à la MCC, une méthode particulière de découplage indirect a été testée et a montré la possibilité non seulement de détecter et de localiser, mais aussi d'identifier les défauts.

IV.3.3. Autres approches pour le découplage indirect

IV.3.3.1. Espace de parité : notion de direction associée à un défaut

En reprenant des notations similaires à celles du chapitre III ((III.5), (III.8), (III.11), (III.14)), le modèle (IV.7) discrétisé ($\delta=z$) peut être exprimé en utilisant le formalisme de l'espace de parité (d'ordre s) :

$$\underline{y} = H_{o,s} \cdot x(k-s) + H_{u,s} \cdot \underline{u} + H_{d,s} \cdot \underline{d} + H_{f_i,s} \cdot \underline{f}_i + H_{f_a,s} \cdot \underline{f}_a \quad (\text{IV.11})$$

avec $\underline{y} = [y(k-s) \ \dots \ y(k)]^T$, idem pour \underline{u} , \underline{d} , \underline{f}_i , \underline{f}_a .

Un résidu quelconque r est obtenu par projection de \underline{y} dans l'espace de parité, ce qui revient à effectuer un découplage direct de l'état. π est le vecteur de pondération associé au résidu r :

$$r = \pi \cdot N_{o,s} \cdot (\underline{y} - H_{u,s} \cdot \underline{u}) = \pi \cdot N_{o,s} \cdot H_{d,s} \cdot \underline{d} + \pi \cdot N_{o,s} \cdot H_{f_i,s} \cdot \underline{f}_i + \pi \cdot N_{o,s} \cdot H_{f_a,s} \cdot \underline{f}_a \quad (\text{IV.12})$$

Remarque : plutôt que d'éliminer $x(k-s)$ (découplage direct), on pourrait également chercher à l'estimer (découplage indirect).

Pour réaliser un découplage indirect, une hypothèse simplificatrice consiste à supposer une variation lente des entrées à estimer (cette hypothèse a d'ailleurs été faite dans la construction des filtres de Kalman (IV.9)). Dans le cas du découplage indirect de d et f_i , (IV.12) conduit à (IV.13). Pour simplifier les notations, on suppose que d correspond à une seule perturbation et f_i à un seul défaut (la généralisation à plusieurs perturbations et à plusieurs défauts est immédiate). L'hypothèse de variation lente est traduite par une valeur constante de d et de f_i sur les $(s+1)$ échantillons correspondant à l'espace de parité ; ainsi, tous les éléments du vecteur \underline{d} sont égaux à d et tous les éléments du vecteur \underline{f}_i sont égaux à f_i .

$$r = \pi \cdot N_{o,s} \cdot (\underline{y} - H_{u,s} \cdot \underline{u}) = \pi \cdot N_{o,s} \cdot h_{d,s} \cdot d + \pi \cdot N_{o,s} \cdot h_{f_i,s} \cdot f_i + \pi \cdot N_{o,s} \cdot H_{f_a,s} \cdot \underline{f}_a \quad (\text{IV.13})$$

$$d \text{ et } f_i \text{ sont scalaires, } h_{d,s} = H_{d,s} \cdot \begin{bmatrix} 1 \\ \vdots \\ 1 \end{bmatrix}, \quad h_{f_i,s} = H_{f_i,s} \cdot \begin{bmatrix} 1 \\ \vdots \\ 1 \end{bmatrix}$$

Dans le cas d'un résidu vectoriel r , chaque colonne de $\pi \cdot N_{o,s} \cdot h_{d,s}$ et de $\pi \cdot N_{o,s} \cdot h_{f_i,s}$ désigne la direction dans laquelle une perturbation d (supposée constante) ou un défaut f_i (supposé constant) agit sur le résidu. Plus généralement, la direction dans laquelle le défaut f_i agit sur le résidu r est donnée (en continu) par $\partial r / \partial f_i$.

Remarque : supposer les entrées à pente constante et non à valeur constante sur la fenêtre d'ordre s permet de limiter l'erreur induite sur le résidu par l'hypothèse de régularité des entrées à découpler (Figure IV.8). A chaque instant et pour chaque entrée à découpler, on cherche à estimer une pente; une intégration donne alors une estimation de l'entrée.

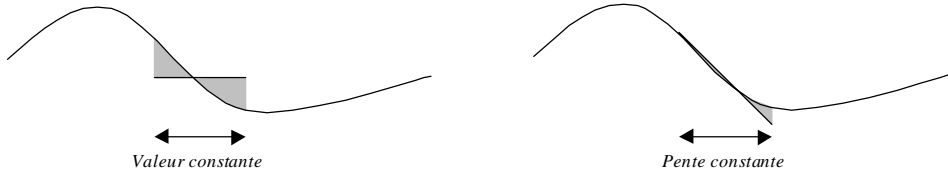


Figure IV.8 : Erreur dans l'évaluation de la cohérence suivant l'hypothèse de régularité

Lorsque π est la matrice identité, le vecteur des résidus r correspond à une base des résidus qu'il est possible de générer dans l'espace de parité. Des conditions de découplage indirect de d et de f_i sous l'hypothèse de variations lentes de ces entrées se déduisent de (IV.13).

$$\text{rang}([N_{o,s}.h_{d,s} \quad N_{o,s}.h_{f_i,s}]) = n_d + n_{f_i} \quad (\text{IV.14})$$

$$\forall j \neq i, \text{rang}([N_{o,s}.h_{d,s} \quad N_{o,s}.h_{f_i,s} \quad N_{o,s}.H_{f_j,s}]) > n_d + n_{f_i} \quad (\text{IV.15})$$

(IV.14) garantit que les n_d perturbations et les n_{f_i} défauts à découpler pourront être reconstruits à partir d'un résidu choisi correctement dans l'espace de parité (par exemple, $\pi=I$). (IV.15) garantit l'existence d'un résidu (dans l'espace de parité) qui est sensible à tous les défauts non découplés. Comme dans le cas du découplage direct (chapitre III), reste à vérifier pour un résidu donné si cette sensibilité est suffisante pour dépasser un seuil robuste aux incertitudes non modélisées.

Remarque 1 : il serait intéressant de comparer les conditions (IV.14) et (IV.15) avec les conditions du découplage parfait et avec la condition d'observabilité (l'observabilité s'apparente à (IV.14) mais ne tient pas compte des impératifs de (IV.15)).

Remarque 2 : si l'état n'est pas éliminé mais estimé, des conditions d'existence du découplage indirect analogues à (IV.14) et à (IV.15) sont données par (IV.16) et par (IV.17) où n est l'ordre du système :

$$\text{rang}([H_{o,s} \quad h_{d,s} \quad h_{f_i,s}]) = n + n_d + n_{f_i} \quad (\text{IV.16})$$

$$\forall j \neq i, \text{rang}([H_{o,s} \quad h_{d,s} \quad h_{f_i,s} \quad H_{f_j,s}]) > n + n_d + n_{f_i} \quad (\text{IV.17})$$

IV.3.3.2. Le découplage indirect vu comme un problème de commande

Dans la manière de réaliser le découplage indirect, nous avons fait une hypothèse de variations lentes des entrées à reconstruire. Nous nous sommes placés ainsi dans le contexte de l'estimation paramétrique. Mais le principe général du découplage indirect peut tout aussi bien se ramener à un problème de commande (Figure IV.9).

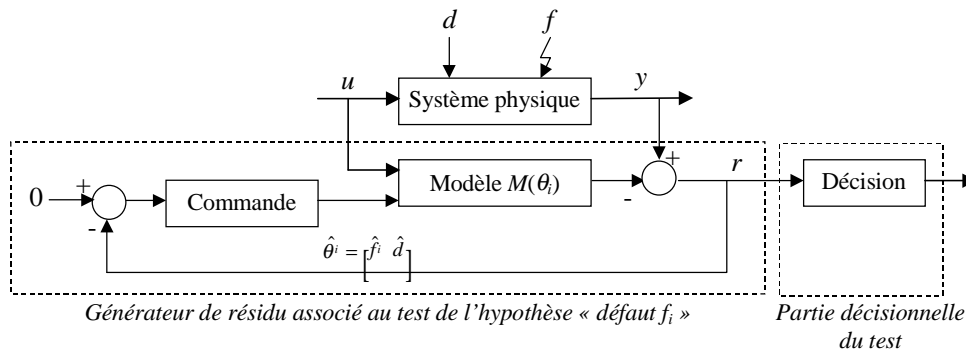


Figure IV.9 : Découplage indirect vu comme un problème de commande

Sur la Figure IV.9, la minimisation du critère de cohérence (le résidu r) entre le comportement connu du système physique (y) et la famille de modèles $M(\theta_i)$ peut être formulée comme un problème de poursuite d'une consigne nulle. Toutes les perturbations ne sont pas forcément modélisées et ne peuvent pas toujours être reconstruites: dans ce cas, la formulation du découplage indirect comme un problème de commande permet d'appliquer tous les outils de prise en compte des incertitudes développés dans ce cadre là: techniques robustes H_∞ , μ synthèse,...

IV.4. Découplage indirect de paramètres

IV.4.1. Un exemple de découplage indirect : estimation de paramètres par programmation non linéaire (PNL)

La famille des modèles décrivant les comportements à découpler est supposée décrite par un modèle de la forme (IV.18).

$$M(\theta_i) : \begin{cases} \dot{x} = g(x, u, \theta_i) \\ y = h(x, u, \theta_i) \end{cases} \quad (\text{IV.18})$$

Nous supposons que les variations de θ_i sont suffisamment lentes pour que l'algorithme d'estimation soit toujours près de l'optimum ($\hat{\theta}_i \approx \theta_i^*$). Les perturbations et les défauts à découpler sont alors considérés comme des paramètres; notons que l'hypothèse de variation lente est le seul critère permettant de distinguer un paramètre d'une entrée dans un modèle comme (IV.18). Afin de tester la cohérence entre le comportement du système physique, connu partiellement grâce aux mesures, et le comportement décrit par $M(\hat{\theta}_i)$, le résidu correspondant à l'erreur de sortie est retenu :

$$\begin{cases} \dot{\hat{x}} = g(\hat{x}, u, \hat{\theta}_i) \\ r = y - h(\hat{x}, u, \hat{\theta}_i) \end{cases} \quad (\text{IV.19})$$

Remarque : un observateur peut très bien être choisi à la place du modèle (IV.19). Il suffit pour cela de remplacer u par le couple (u, y) . Le résidu r correspond alors non plus à une erreur de sortie mais à une erreur de prédiction. Toute la démarche et les justifications qui suivent s'appliquent également au cas d'un observateur.

Le découplage indirect des perturbations et des défauts modélisés par θ_i est obtenu en minimisant r par une adaptation de $\hat{\theta}_i$. Compte tenu de la forme générale du modèle (IV.19), l'éventail des techniques permettant de réaliser une telle adaptation relève pour l'essentiel de la programmation non linéaire (PNL). Dans ce cadre, l'un des algorithmes les moins coûteux en calculs est celui du gradient en ligne : à chaque période d'échantillonnage k , $\hat{\theta}_i$ est adapté dans le sens opposé à la dérivée du critère $J=r^T r$:

$$\hat{\theta}_{i,k+1} = \hat{\theta}_{i,k} - \eta \left(\frac{\partial r}{\partial \hat{\theta}_i} \right)_k^T \cdot r_k \quad (\text{IV.20})$$

Afin de pouvoir calculer (IV.20), une expression du gradient $\partial r / \partial \hat{\theta}_i$ est d'abord établie en continu à partir de (IV.19) :

$$\frac{\partial r}{\partial \hat{\theta}_i} = -\frac{\partial h}{\partial \hat{x}} \cdot \frac{\partial \hat{x}}{\partial \hat{\theta}_i} - \frac{\partial h}{\partial \hat{\theta}_i} \quad (\text{IV.21})$$

$$\left(\frac{\partial \dot{\hat{x}}}{\partial \hat{\theta}_i} \right) = \frac{\partial(\dot{\hat{x}})}{\partial \hat{\theta}_i} = \frac{\partial g}{\partial \hat{x}} \cdot \frac{\partial \hat{x}}{\partial \hat{\theta}_i} + \frac{\partial g}{\partial \hat{\theta}_i} \quad (\text{IV.22})$$

Les équations (IV.21) et (IV.22) doivent être justifiées ; en effet, rien ne garantit a priori que :

$$\frac{\partial y}{\partial \hat{\theta}_i} = 0, \quad \frac{\partial u}{\partial \hat{\theta}_i} = 0 \quad \text{et} \quad \left(\frac{\partial \dot{\hat{x}}}{\partial \hat{\theta}_i} \right) = \frac{\partial(\dot{\hat{x}})}{\partial \hat{\theta}_i}.$$

Justification:

$$\boxed{\dot{\hat{x}} = g(\hat{x}, u, \hat{\theta}_i)} \quad (a)$$

(a) est une équation différentielle dont la solution dépend de t , u , $\hat{\theta}_i$ et d'une condition initiale que nous omettons pour simplifier les notations. Par conséquent :

$$d\hat{x} = \frac{\partial \hat{x}}{\partial t} dt + \frac{\partial \hat{x}}{\partial u} du + \frac{\partial \hat{x}}{\partial \hat{\theta}_i} d\hat{\theta}_i, \quad \text{soit encore: } \dot{\hat{x}} = \frac{d\hat{x}}{dt} = \frac{\partial \hat{x}}{\partial t} + \frac{\partial \hat{x}}{\partial u} \dot{u} + \frac{\partial \hat{x}}{\partial \hat{\theta}_i} \dot{\hat{\theta}}_i$$

On en déduit alors que :

$$\boxed{\frac{\partial \dot{\hat{x}}}{\partial \hat{\theta}_i} = \frac{\partial g(\hat{x}, u, \hat{\theta}_i)}{\partial \hat{\theta}_i} = \frac{\partial^2 \hat{x}}{\partial t \partial \hat{\theta}_i} + \frac{\partial^2 \hat{x}}{\partial u \partial \hat{\theta}_i} \dot{u} + \frac{\partial \hat{x}}{\partial u} \cdot \frac{\partial \dot{u}}{\partial \hat{\theta}_i} + \frac{\partial^2 \hat{x}}{\partial \hat{\theta}_i^2} \dot{\hat{\theta}}_i + \frac{\partial \hat{x}}{\partial \hat{\theta}_i} \cdot \frac{\partial \dot{\hat{\theta}}_i}{\partial \hat{\theta}_i}} \quad (b)$$

Par ailleurs, comme \hat{x} dépend de t , u , $\hat{\theta}_i$, il en est de même pour $\frac{\partial \hat{x}}{\partial \hat{\theta}_i}$. Par conséquent :

$$\boxed{\left(\frac{\partial \dot{\hat{x}}}{\partial \hat{\theta}_i} \right) = \frac{d \left(\frac{\partial \hat{x}}{\partial \hat{\theta}_i} \right)}{dt} = \frac{\partial^2 \hat{x}}{\partial \hat{\theta}_i \partial t} + \frac{\partial^2 \hat{x}}{\partial \hat{\theta}_i \partial u} \dot{u} + \frac{\partial^2 \hat{x}}{\partial \hat{\theta}_i^2} \dot{\hat{\theta}}_i} \quad (c)$$

Etant donné que les dérivées secondes croisées sont égales (théorème de Schwarz), la combinaison de (b) et de (c) donne :

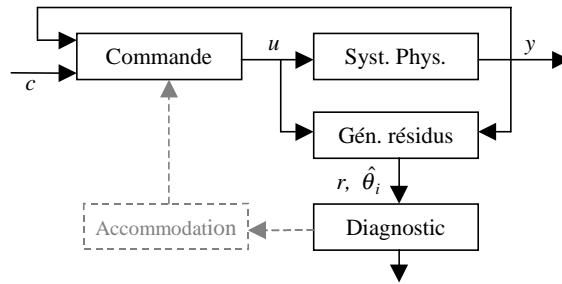
$$\boxed{\left(\frac{\partial \dot{\hat{x}}}{\partial \hat{\theta}_i} \right) - \frac{\partial \hat{x}}{\partial u} \cdot \frac{\partial \dot{u}}{\partial \hat{\theta}_i} - \frac{\partial \hat{x}}{\partial \hat{\theta}_i} \cdot \frac{\partial \dot{\hat{\theta}}_i}{\partial \hat{\theta}_i} = \frac{\partial \dot{\hat{x}}}{\partial \hat{\theta}_i} = \frac{\partial g(\hat{x}, u, \hat{\theta}_i)}{\partial \hat{\theta}_i}} \quad (d)$$

Ayant supposé une variation lente des paramètres θ_i , si l'algorithme d'estimation est correctement paramétré, alors $\hat{\theta}_i$ est également soumis à des variations lentes. Dans ces conditions, une approximation linéaire est localement très bonne, donc $\dot{\hat{\theta}}_i \approx \text{constante}$, ce qui conduit à :

$$\boxed{\frac{\partial \dot{\hat{\theta}}_i}{\partial \hat{\theta}_i} \approx 0} \quad (e)$$

\dot{u} dépend des mêmes variables que u . Nous supposons que u est la commande appliquée au système physique et non une consigne extérieure (ce qui sera le cas dans l'application à la machine

asynchrone). Il est alors important de bien évaluer les grandeurs influençant u :



Dans le cas d'une application de diagnostic simple (sans accommodation), u dépend uniquement du temps t , des consignes c et des mesures y . Un diagnostic simple est réalisé en boucle ouverte. Par contre, si un mécanisme d'accommodation est envisagé, le générateur de résidus n'est plus positionné en boucle ouverte mais en boucle fermée par rapport au système physique. Il est alors tout à fait possible que u et y dépendent de $\hat{\theta}_i$. Ces réserves étant faites, comme le cadre de notre étude concerne le seul diagnostic, nous considérons que u et y sont indépendants de $\hat{\theta}_i$, d'où:

$$\left[\frac{\partial y}{\partial \hat{\theta}_i} = 0 \right] \quad (f) \quad \left[\frac{\partial u}{\partial \hat{\theta}_i} = 0 \right] \quad (g) \quad \left[\frac{\partial \dot{u}}{\partial \hat{\theta}_i} = 0 \right] \quad (h)$$

En tenant compte de (e) et de (h) dans (d), on obtient (i), ce qui achève la justification:

$$\left[\begin{matrix} \dot{\hat{x}} \\ \frac{\partial \hat{x}}{\partial \hat{\theta}_i} \end{matrix} \right] \approx \frac{\partial(\hat{x})}{\partial \hat{\theta}_i} \quad (i)$$

En réordonnant les équations (IV.19) à (IV.22), on obtient la représentation d'état du générateur de résidus réalisant le découplage indirect de $\hat{\theta}_i$ par la méthode du gradient :

$$\text{Etats continus:} \quad \left| \begin{array}{l} \dot{\hat{x}} = g(\hat{x}, u, \hat{\theta}_i) \\ \left(\frac{\partial \hat{x}}{\partial \hat{\theta}_i} \right) = \frac{\partial g(\hat{x}, u, \hat{\theta}_i)}{\partial \hat{x}} \cdot \left(\frac{\partial \hat{x}}{\partial \hat{\theta}_i} \right) + \frac{\partial g(\hat{x}, u, \hat{\theta}_i)}{\partial \hat{\theta}_i} \end{array} \right. \quad (IV.23)$$

$$\text{Sorties:} \quad \left| \begin{array}{l} r = y - h(\hat{x}, u, \hat{\theta}_i) \\ \left(\frac{\partial r}{\partial \hat{\theta}_i} \right) = - \frac{\partial h(\hat{x}, u, \hat{\theta}_i)}{\partial \hat{x}} \cdot \left(\frac{\partial \hat{x}}{\partial \hat{\theta}_i} \right) - \frac{\partial h(\hat{x}, u, \hat{\theta}_i)}{\partial \hat{\theta}_i} \end{array} \right. \quad (IV.24)$$

$$\text{Etats discrets:} \quad \left| \begin{array}{l} \text{(Adaptation)} \quad \hat{\theta}_{i,k+1} = \hat{\theta}_{i,k} - \eta \cdot \left(\frac{\partial r}{\partial \hat{\theta}_i} \right)_k^T \cdot r_k \end{array} \right. \quad (IV.25)$$

Du point de vue de l'implantation, la partie continue du générateur de résidus doit être discrétisée. Même si la méthode proposée nécessite une quantité importante de calcul en ligne (augmentation importante du nombre d'états par rapport au système initial), elle présente néanmoins l'avantage d'être adaptée à une large classe de modèles. De plus, la possibilité d'estimer directement des « paramètres physiques » simplifie considérablement l'étape ultérieure de décision, à condition que les « paramètres physiques » estimés modélisent correctement l'influence des défauts à diagnostiquer.

Remarque 1 : les méthodes du type « moindres carrés » sont couramment employées pour l'estimation de paramètres. Elles ne sont malheureusement applicables qu'à une classe de modèles très restrictive ; comme il est rare que les paramètres physiques interviennent linéairement dans un modèle suffisamment précis, une approche du type « moindres carrés » n'a pas été retenue.

Remarque 2 : une méthode couramment employée pour estimer à la fois l'état et certains paramètres consiste à implanter un filtre de Kalman étendu. Le modèle (IV.18) doit alors être linéarisé autour de l'état étendu estimé ($\Delta x = x - \hat{x}$, ...) :

$$\begin{cases} \begin{bmatrix} \dot{\Delta x} \\ \Delta \theta_i \end{bmatrix} = \begin{bmatrix} \partial g(x,u,\theta_i)/\partial x & \partial g(x,u,\theta_i)/\partial \theta_i \\ 0 & 0 \end{bmatrix} \begin{bmatrix} \Delta x \\ \Delta \theta_i \end{bmatrix} + \begin{bmatrix} \partial g(x,u,\theta_i)/\partial u \\ 0 \end{bmatrix} \Delta u \\ \Delta y = \begin{bmatrix} \partial h(x,u,\theta_i)/\partial x & \partial h(x,u,\theta_i)/\partial \theta_i \end{bmatrix} \begin{bmatrix} \Delta x \\ \Delta \theta_i \end{bmatrix} + \begin{bmatrix} \partial h(x,u,\theta_i)/\partial u \end{bmatrix} \Delta u \end{cases} \quad (\text{IV.26})$$

Nous avons essayé d'appliquer cette méthode à l'estimation de l'état et d'un paramètre pour le diagnostic de la machine asynchrone. Le réglage du filtre de Kalman s'est révélé délicat et ne nous a pas permis d'aboutir à des résultats satisfaisants dans des temps raisonnables. Ces difficultés sont mentionnées dans plusieurs publications relatives au diagnostic de la MAS [Lapeyre, 1997], [Eldem, 1999]. En comparant (IV.23) et (IV.26), on peut remarquer que l'algorithme du gradient limite la linéarisation au seul calcul du gradient (l'équation d'état modélisant l'évolution de \hat{x} n'est pas linéarisée). De plus, le gradient donne par nature une information sur la direction dans laquelle agit le paramètre estimé et cette direction fait partie de l'état de l'estimateur. Par comparaison, l'état d'un filtre de Kalman étendu classique ne contient qu'une information sur la valeur du paramètre estimé.

IV.4.2. Condition d'existence du découplage indirect

L'existence du découplage indirect est liée à l'identifiabilité des paramètres à estimer. Nous utilisons les résultats de [Ljung, 1994] qui montrent l'équivalence entre l'identifiabilité globale et la possibilité d'exprimer l'estimation des paramètres comme une régression linéaire. Ces résultats sont applicables à une large classe de modèles dont (IV.18) fait partie¹. Par conséquent, on peut affirmer que la « structure »² $M(\theta_i)$ est globalement identifiable dans l'espace des paramètres θ_i ($\theta_i = [\theta_{i1}, \dots, \theta_{im}]$) si et seulement si il est possible de réarranger les équations de $M(\theta_i)$ de manière à obtenir, paramètre par paramètre, une régression linéaire de la forme (IV.27) où $p_j(u,y)$ et $q_j(u,y)$ sont deux polynômes, fonctions des variables u , y et de leurs dérivées successives.

$$p_j(u,y)\theta_{ij} - q_j(u,y) = 0, \quad j=1 \dots m \quad (\text{IV.27})$$

Remarque : un algorithme explicite pour tester l'identifiabilité globale et basé sur l'algèbre différentielle est proposé dans [Ljung, 1994].

¹ Le résultat de Ljung est établi pour des systèmes à paramètres invariants dans le temps. Nous supposons que l'hypothèse de variations lentes des paramètres est suffisante pour que ce résultat puisse être appliqué.

² Le terme « structure » est à prendre ici au sens de l'identification. Il correspond à la famille de modèles $M(\theta_i)$ et s'écarte donc de la définition du chapitre II (connaissance de l'existence de relations entre variables).

Même si la famille de modèles $M(\theta_i)$ est globalement identifiable, il se peut très bien que pour certaines valeurs de l'entrée u et de la sortie correspondante y , on ait : $p_f(u,y)=0$. Cette situation met en évidence la nécessité d'une excitation suffisante sur l'entrée pour estimer correctement les paramètres. Il est évident que les contraintes sur l'excitation de l'entrée sont d'autant plus fortes que le nombre de paramètres à estimer est important.

IV.4.3. Application à la MAS : préliminaires

IV.4.3.1. Modèle de l'entraînement asynchrone (sans la commande)

Partant des équations électriques d'une MAS équilibrée ((II.54) à (II.62)), une représentation d'état peut être obtenue dans un repère $(\alpha\beta)$ fixe par rapport au stator (les équations dans ce repère se déduisent de celles établies dans un repère tournant (dq) quelconque en posant $\omega_s=0$) (voir l'annexe C). Cette représentation d'état est complétée pour introduire une modélisation des différents défauts et perturbations étudiées.

- **Equations électriques:** (IV.28)

$$\begin{aligned}\dot{x}_e &= A(x_m, f_p, d_T) \cdot x_e + B(f_p) \cdot K_{23} \cdot (u + f_v) \\ y_e &= K_{22}^{-1} \cdot C \cdot x_e + f_I\end{aligned}$$

- **Equations mécaniques:** (IV.29)

$$\begin{aligned}\dot{x}_m &= \Lambda(x_m, x_e, d_\Gamma) \\ \Omega &= x_m + f_\Omega\end{aligned}$$

- **Entrées, états, mesures:** (IV.30)

$$u = \begin{bmatrix} V_{sa} \\ V_{sb} \\ V_{sc} \end{bmatrix}, f = \begin{bmatrix} f_I \\ f_v \\ f_\Omega \\ f_p \end{bmatrix}, d = \begin{bmatrix} d_\Gamma \\ d_T \end{bmatrix}, x = \begin{bmatrix} x_e \\ x_m \end{bmatrix}, y = \begin{bmatrix} y_e \\ \Omega \end{bmatrix} \text{ où } x_e = \begin{bmatrix} \tilde{\Phi}_{s\alpha} \\ \tilde{\Phi}_{s\beta} \\ \tilde{I}_{s\alpha} \\ \tilde{I}_{s\beta} \end{bmatrix}, y_e = \begin{bmatrix} I_{sa} \\ I_{sb} \end{bmatrix}$$

- **Changements de repère:** (IV.31)

K_{23} : (abc) vers $(\alpha\beta)$, K_{22} : (ab) vers $(\alpha\beta)$ en supposant abc équilibrés.

$$K_{23} = \sqrt{\frac{2}{3}} \cdot \begin{bmatrix} 1 & -\frac{1}{2} & -\frac{1}{2} \\ 0 & \frac{\sqrt{3}}{2} & -\frac{\sqrt{3}}{2} \end{bmatrix}, K_{22} = \sqrt{2} \cdot \begin{bmatrix} \frac{\sqrt{3}}{2} & 0 \\ \frac{1}{2} & 1 \end{bmatrix}$$

$$\begin{bmatrix} V_{s\alpha} \\ V_{s\beta} \end{bmatrix} = K_{23} \cdot (u + f_v) \quad \begin{bmatrix} I_{s\alpha} \\ I_{s\beta} \end{bmatrix} = K_{22} \cdot y_e$$

- **Matrices de la représentation d'état (partie électrique):** (IV.32)

$$A(x_m, f_p, d_T) = \begin{bmatrix} 0 & 0 & -R_s & 0 \\ 0 & 0 & 0 & -R_s \\ R_r/\sigma & p \cdot x_m \cdot L_r/\sigma & -\xi/\sigma & -p \cdot x_m \\ -p \cdot x_m \cdot L_r/\sigma & R_r/\sigma & p \cdot x_m & -\xi/\sigma \end{bmatrix} B(f_p) = \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ L_r/\sigma & 0 \\ 0 & L_r/\sigma \end{bmatrix}$$

$$C = \begin{bmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}, \sigma = L_s \cdot L_r - L_m^2, \xi = R_r \cdot L_s + R_s \cdot L_r$$

R_s et R_r dépendent de d_T . R_s, R_r, L_s, L_r, L_m dépendent de f_p .

x_e représente l'état électrique (indice e) et x_m représente l'état mécanique (indice m). $R_s, R_r, L_s, L_r, L_m, p$ sont les paramètres électriques usuels du modèle de Park d'un moteur asynchrone. I, Φ, V représentent respectivement des courants, des flux et des tensions. Les mesures y sont celles utilisées usuellement en commande : les courants stator (I_{sa}, I_{sb}) et la vitesse de rotation du moteur Ω . Les perturbations considérées d sont le couple de charge, d_T , et les variations des paramètres dues à l'échauffement, d_T . Les défauts capteurs, f_I et f_Ω , sont modélisés par des entrées additives. f_V modélise les défauts dans l'onduleur (assimilé à un gain unitaire en fonctionnement normal). f_P représente les défauts du moteur (court-circuit de spires, rupture de barre) modélisés sous la forme de variations des paramètres électriques du modèle.

IV.4.3.2. Découplage direct par décomposition de modèle

L'entraînement électrique asynchrone est complètement modélisé par les équations électriques et mécaniques (IV.28), (IV.29), et par les relations correspondant à la commande du moteur. Dans le contexte de la décomposition de modèle (chapitre II), la commande correspond à un modèle local particulier puisque ses entrées et ses sorties sont connues. Nous choisissons d'éliminer ce modèle local afin que le diagnostic soit indépendant du type de commande utilisé ; il existe en effet une grande variété de commandes pour les entraînements électriques asynchrones (commandes scalaires, commandes vectorielles,...).

L'analyse causale appliquée au chapitre II à l'entraînement asynchrone (sans sa commande) a mis en évidence que le modèle électrique de la MAS pouvait être parfaitement découplé du couple de charge en utilisant la mesure de vitesse en « entrée » de ce modèle. On élimine ainsi l'équation mécanique, et le diagnostic porte uniquement sur le modèle électrique de la MAS (IV.28).

Pour remplir les objectifs de la détection, il est nécessaire de construire un indicateur découplé de l'état x et des perturbations d (IV.30). Le découplage direct résultant de l'analyse causale permet de satisfaire partiellement ces objectifs puisqu'il conduit au découplage de l'état mécanique x_m et du couple de charge d_T . Il faut encore réaliser le découplage de x_e et de d_T et vérifier la sensibilité aux défauts f pour réaliser une détection robuste aux perturbations. Pour satisfaire les objectifs de localisation, le découplage devra également porter sur les défauts.

IV.4.3.3. Représentation simplifiée : notation complexe (!)

Une représentation simplifiée des équations du modèle est proposée afin de simplifier les notations pour l'étude de l'identifiabilité de certains paramètres. Cette représentation sert également à introduire une représentation graphique des propriétés du modèle de la MAS. On pose :

$$\underline{V} = V_{s\alpha} + j.V_{s\beta}, \quad \underline{\Phi} = \tilde{\Phi}_{s\alpha} + j.\tilde{\Phi}_{s\beta}, \quad \underline{I} = I_{s\alpha} + j.I_{s\beta} \quad \text{où } j = \sqrt{-1} \quad (\text{IV.33})$$

Compte tenu de la structure par blocs 2×2 du type $\begin{bmatrix} a & -b \\ b & a \end{bmatrix}$ des matrices A, B, C du modèle de la MAS (IV.32), les équations électriques (IV.28) sont équivalentes à (IV.34) et (IV.35) lorsque $f_V=0$ et $f_I=0$. $\underline{a}, \underline{b}$ et \underline{c} sont des fonctions à valeurs complexes :

$$\dot{\underline{\Phi}} = \underline{V} - R_s(d_T).\underline{I} \quad (\text{IV.34})$$

$$\dot{\underline{I}} = \underline{a}(\Omega - f_\Omega, f_P, d_T).\underline{\Phi} + \underline{b}(\Omega - f_\Omega, f_P, d_T).\underline{I} + \underline{c}.\underline{V} \quad (\text{IV.35})$$

où:

$$\underline{a}(\Omega - f_\Omega, f_P, d_T) = \frac{R_r - j.p.(\Omega - f_\Omega).L_r}{\sigma}, \quad \underline{b}(\Omega - f_\Omega, f_P, d_T) = \frac{-\xi}{\sigma} + j.p.(\Omega - f_\Omega), \quad \underline{c} = \frac{L_r}{\sigma}$$

En considérant que la vitesse est un paramètre du modèle électrique évoluant lentement, la dérivation de (IV.35) et la substitution de $\underline{\Phi}$ par (IV.34) donnent :

$$\underline{\ddot{I}} = \underline{a}(\Omega - f_\Omega, f_P, d_T).(\underline{V} - R_s(d_T).\underline{I}) + \underline{b}(\Omega - f_\Omega, f_P, d_T).\underline{\dot{I}} + \underline{c}.\underline{\dot{V}} \quad (\text{IV.36})$$

Par ailleurs, on peut également montrer que :

$$\underline{I}(j\omega) = \underline{h}(j\omega).\underline{V}(j\omega) \quad (\text{IV.37})$$

avec $\underline{h}(j\omega) = [1 \quad 0]C.(j\omega.I - A)^{-1}.B. \begin{bmatrix} 1 \\ -j \end{bmatrix}$

A une fréquence ω donnée, le transfert entre les deux tensions et les deux courants de la MAS se ramène donc à un simple nombre complexe. En utilisant cette propriété, des représentations graphiques en trois dimensions (ω , $\Re(\underline{h}(j\omega))$, $\Im(\underline{h}(j\omega))$)¹ vont nous permettre de visualiser les déformations du transfert électrique de la MAS induites par différents types de variations de paramètres.

IV.4.4. Application à la MAS : détection robuste à l'échauffement

IV.4.4.1. Modèle

Les effets thermiques peuvent avoir une influence considérable sur le comportement de la MAS. Ils se traduisent essentiellement par des variations des résistances ; ainsi, la résistance rotorique R_r peut facilement varier de plus de 50% à cause de l'échauffement. Afin de pouvoir détecter avec une sensibilité correcte des défauts machines (f_P) qui se traduisent, entre autres², par des petites variations des paramètres électriques (inductances et résistances), il est donc nécessaire de découpler les résidus des effets de l'échauffement (d_T). La correction en température d'une résistance étant communément obtenue par une loi linéaire [Grellet, 1997], l'échauffement sera modélisé comme indiqué par (IV.38).

$$\begin{aligned} R_s &= R_{s0}.(1 + k_s.d_T) \\ R_r &= R_{r0}.(1 + k_r.d_T) \end{aligned} \quad (\text{IV.38})$$

d_T est **un** paramètre (scalaire) qui correspond à **un** phénomène physique : l'échauffement. Il peut donc être considéré comme un paramètre physique qui décrit l'évolution relative des paramètres du modèle de Park en réponse à l'échauffement. k_s et k_r peuvent être évalués par une

¹ $\Re(\cdot)$ et $\Im(\cdot)$ désignent respectivement les parties réelles et imaginaires d'un argument complexe.

² Les défauts machines peuvent également se traduire par un déséquilibre des grandeurs triphasées. Le modèle de Park ne permet pas de modéliser la composante homopolaire (celle qui traduit un déséquilibre). Ainsi, le diagnostic que nous établissons pour la MAS porte uniquement sur la projection du modèle (abc) dans le plan ($\alpha\beta$) (ou dq). Ce plan étant orthogonal à l'influence du déséquilibre, il n'est pas tenu compte d'un éventuel déséquilibre dans le diagnostic proposé. Cela ne signifie pas pour autant que le diagnostic sera faux, mais simplement qu'il n'utilise pas toute la connaissance possible : comme tout modèle, le modèle de Park ne représente qu'une partie de la réalité du système physique.

identification à froid et à chaud des paramètres électriques. Toutefois, le modèle (IV.38) ne modélise que très imparfaitement le phénomène d'échauffement : pour garantir une précision correcte à un modèle thermique pour la MAS, un modèle à deux corps (inducteur et induit) s'impose [Grellet, 1997]. Il traduit des dynamiques d'échauffement différentes entre le stator et le rotor, ce que la relation statique (IV.38) ne peut représenter. Toutefois, (IV.38) est plus générale que la seule adaptation de R_r traditionnellement utilisée en commande (il suffit de poser $k_s=0$ pour s'y ramener). Surtout, l'utilisation d'une méthode de PNL pour adapter d_T laisse une grande liberté dans le choix de la famille de modèles décrivant l'échauffement. Ainsi, un modèle dynamique peut très bien être adopté à la place de (IV.38). Reste alors à trouver un compromis acceptable entre la précision du modèle et sa complexité. Notons que la possibilité d'identifier correctement les paramètres d'un tel modèle est un point clé pour aboutir à un bon compromis.

En tenant compte de (IV.38), la famille de modèles $M(d_T)$ décrivant les comportements normaux ($f=0$) de la partie électrique de la MAS en fonction de l'échauffement est donnée par (IV.39).

$$M(d_T) : \begin{cases} \dot{x}_e = A(\Omega, d_T).x_e + B.K_{23}.u \\ y_e = K_{22}^{-1}.C.x_e \end{cases} \quad (IV.39)$$

La méthode du gradient va être appliquée pour réaliser le découplage indirect de d_T et générer un résidu qui remplit les objectifs de la détection de défauts.

IV.4.4.2. Identifiabilité

Le découplage indirect de d_T reposant sur l'estimation de ce paramètre, son identifiabilité est étudiée. En premier lieu, l'échauffement est un phénomène très lent par rapport aux phénomènes électriques et mécaniques. L'hypothèse de variation lente de d_T est donc fondée.

Afin de prouver l'identifiabilité globale de d_T , nous cherchons à manipuler les équations de (IV.39) de façon à mettre en évidence une régression linéaire en d_T conforme à (IV.27). En tenant compte de la linéarité des résistances par rapport à d_T (IV.38), la représentation complexe des équations électriques (IV.36) peut se mettre sous la forme d'un trinôme à coefficients complexes.

$$\underline{m}.d_T^2 + \underline{p}.d_T + \underline{q} = 0 \Leftrightarrow \begin{cases} \Re(\underline{m}).d_T^2 + \Re(\underline{p}).d_T + \Re(\underline{q}) = 0 \\ \Im(\underline{m}).d_T^2 + \Im(\underline{p}).d_T + \Im(\underline{q}) = 0 \end{cases} \quad (IV.40)$$

\underline{m} , \underline{p} et \underline{q} sont des polynômes à valeurs complexes qui dépendent des mesures et de leurs dérivées. En combinant les parties réelles et imaginaires de (IV.40) de manière à éliminer d_T^2 , on obtient :

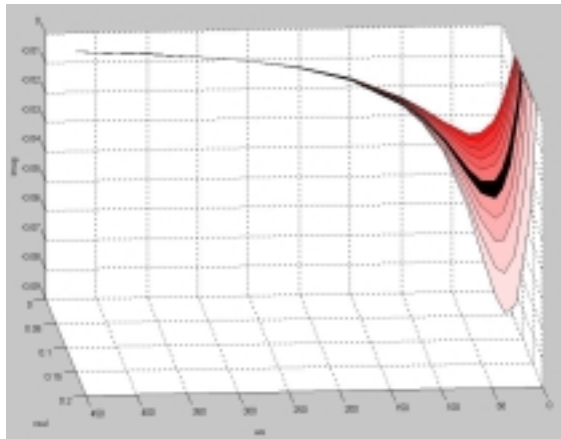
$$\left(\Im(\underline{m}).\Re(\underline{p}) - \Re(\underline{m}).\Im(\underline{p})\right)d_T + \left(\Im(\underline{m}).\Re(\underline{q}) - \Re(\underline{m}).\Im(\underline{q})\right) = 0 \quad (IV.41)$$

(IV.41) est une expression linéaire en d_T ayant la forme de (IV.27), ce qui prouve l'identifiabilité globale de d_T lorsque d_T est le seul paramètre estimé.

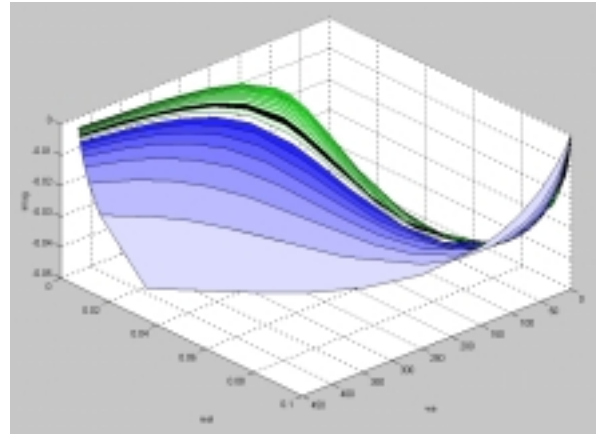
IV.4.4.3. Sensibilité

Une étude graphique de la sensibilité des équations électriques de la MAS à plusieurs types de variations de paramètres a été menée. En l'absence d'information précise sur l'évolution relative des paramètres du modèle de Park décrivant un échauffement ou un défaut machine, des

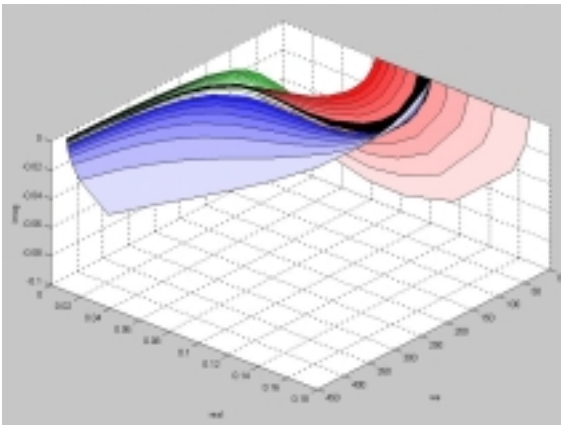
variations du même ordre de grandeur des paramètres relatifs au stator et au rotor ont été étudiées. Cette position est celle retenue par le groupe de travail « méthodes de redondance analytique » dans le cadre de l'action incitative¹ sur le diagnostic de la MAS. Dans un premier temps, des variations de R_s et de R_r du même ordre de grandeur ($k_s=k_r$) sont comparées à des variations de L_s et de L_r du même ordre de grandeur.



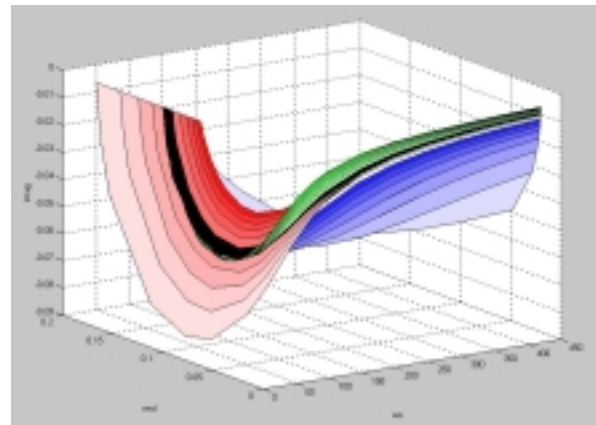
(a) Variation des résistances (échauffement)



(b) Variation des inductances



(c) Superposition de (a) et (b)



(d) Superposition de (a) et (b), (bis)

Figure IV.10 : Etude de l'influence des variations des résistances et des inductances sur le transfert des équations électriques de la MAS (à glissement nul)

Nous avons vu que le nombre complexe $\underline{h}(j\omega)$ permet de décrire le transfert entre les tensions et les courants (IV.37). La courbe en trois dimensions paramétrée par $(\omega, \Re(\underline{h}(j\omega)), \Im(\underline{h}(j\omega)))$ décrit complètement ce transfert. Pour les valeurs nominales des paramètres, cette courbe est tracée en noir sur la Figure IV.10. Pour un type particulier de variation des paramètres, cette courbe se déforme et l'on obtient ainsi une surface décrivant le lieu du transfert. Par exemple, la surface rouge (Figure IV.10) décrit l'évolution du transfert en fonction de l'échauffement : la couleur rouge est d'autant plus foncée que l'échauffement est fort. Cette surface correspond à la famille de modèles $M(d_T)$ lorsque le glissement est nul. Dans les mêmes conditions, une autre surface décrit l'évolution du transfert lorsque les inductances varient. La partie verte de cette surface

¹ Action commune aux GdR Automatique et Electrotechnique regroupant plusieurs équipes de recherche à l'échelle nationale.

décrit un domaine dans lequel le transfert est stable alors que la partie bleue décrit un domaine dans lequel le transfert est instable. La courbe correspondant aux valeurs nominales fait (heureusement) partie du domaine de stabilité. Cette courbe (noire) est nécessairement incluse dans l'intersection entre les surfaces associées aux résistances (rouge) et aux inductances relatives à un transfert stable (vert). On retrouve sur la Figure IV.10 certains comportements qualitativement similaires à un transfert monovariable de la forme $1/(L.p+R)$: les variations des résistances ont une influence forte aux basses fréquences ; cette influence diminue au fur et à mesure que la fréquence augmente. Les variations des inductances se traduisent par une évolution inverse en fonction de la fréquence. Surtout, les surfaces rouge et verte ont tendance à se croiser et non à être tangentes l'une à l'autre. Cela signifie que le découplage vis-à-vis des variations des résistances n'induit pas une désensibilisation trop forte du résidu vis-à-vis des variations des inductances. Cette situation permet de construire un indicateur de défaut robuste aux variations de paramètres dues à l'échauffement et sensible aux autres variations de paramètres (associées à un défaut dans le moteur). Le résidu ne permet pas seulement de détecter les variations des inductances, mais toute variation des paramètres autre que celles explicables par un échauffement. Si un défaut moteur se traduit à la fois par une variation des résistances, assimilable à un échauffement, et par d'autres variations de paramètres (en particulier, sur les inductances), c'est sur ces autres variations de paramètres que peut se faire la détection, car elles seules sensibilisent le résidu découplé de d_T . Pour qu'une variation des paramètres puisse sensibiliser un tel résidu, elle doit en effet se traduire par une composante orthogonale à la surface rouge (échauffement) non nulle.

IV.4.4.4. Découplage indirect par la méthode du gradient

L'application de la méthode du gradient à la famille de modèles $M(d_T)$ conduit au générateur de résidus donné par (IV.42) à (IV.44). L'expression du résidu r_T , exprimé dans le repère $(\alpha\beta)$, résulte de l'application de (IV.18), (IV.23), (IV.24) et (IV.25) à (IV.39):

$$\begin{bmatrix} \dot{\hat{x}}_e \\ \frac{\partial \hat{x}_e}{\partial \hat{d}_T} \end{bmatrix} = \underbrace{\begin{bmatrix} A(\Omega, \hat{d}_T) & 0 \\ \frac{\partial A(\Omega, \hat{d}_T)}{\partial \hat{d}_T} & A(\Omega, \hat{d}_T) \end{bmatrix}}_{\tilde{A}_c(\Omega, \hat{d}_T)} \begin{bmatrix} \hat{x}_e \\ \frac{\partial \hat{x}_e}{\partial \hat{d}_T} \end{bmatrix} + \underbrace{\begin{bmatrix} B \\ 0 \end{bmatrix}}_{\tilde{B}_c} \cdot K_{23} \cdot u \quad (\text{IV.42})$$

$$r_T = K_{22} \cdot y_e - C \cdot \hat{x}_e \quad (\text{IV.43})$$

$$\delta_T = \frac{\partial r_T}{\partial \hat{d}_T} = -C \cdot \frac{\partial \hat{x}_e}{\partial \hat{d}_T}$$

$$\hat{d}_{T,k+1} = \hat{d}_{T,k} - \eta_T \cdot \delta_{T,k}^T \cdot r_{T,k} \quad (\text{IV.44})$$

Le réglage retenu pour le gain d'adaptation est $\eta_T = 10^{-5}$. La discrétisation du générateur de résidus continu est obtenue en développant jusqu'à l'ordre deux l'exponentielle de matrice associée. Les matrices de la représentation d'état discrète (indice d) sont donc obtenues à partir des matrices de la représentation continue (indice c) comme l'indique (IV.45) :

$$\tilde{A}_{d,k} = I + \tilde{A}_c(\Omega_k, \hat{d}_{T,k}) \cdot T_e + \frac{1}{2} (\tilde{A}_c(\Omega_k, \hat{d}_{T,k}) \cdot T_e)^2 \quad (\text{IV.45})$$

$$\tilde{B}_{d,k} = T_e \left(I + \frac{1}{2} \tilde{A}_c(\Omega_k, \hat{d}_{T,k}) T_e \right) \tilde{B}_c$$

L'interprétation du comportement de r_T est réalisée en transposant les résultats de la Table IV.1 dans le cas où seules des perturbations sont découplées. Plus précisément, si r_T est proche de zéro, cela signifie que le comportement du système physique est explicable par une variation des paramètres due à l'échauffement. Si r_T est différent de zéro, un autre type de comportement s'est produit. r_T étant découplé du couple de charge (découplage direct) et de l'échauffement (découplage indirect), cet autre type de comportement est significatif de l'occurrence d'un défaut. Pour garantir que r_T remplit les objectifs de la détection, il faut vérifier que chaque type de défaut agit différemment de l'échauffement sur le système. Pour les variations des inductances, ce point a déjà été étudié graphiquement. Nous présentons maintenant quelques résultats de simulation. Pour cela, deux types de scénarios de base (*ScSi* et *ScCo*) définissent le profil de la consigne de vitesse et le profil du couple de charge (Figure IV.11). A ces scénarios de base sont ensuite ajoutés différents scénarios de défaut pour étudier le comportement de r_T .

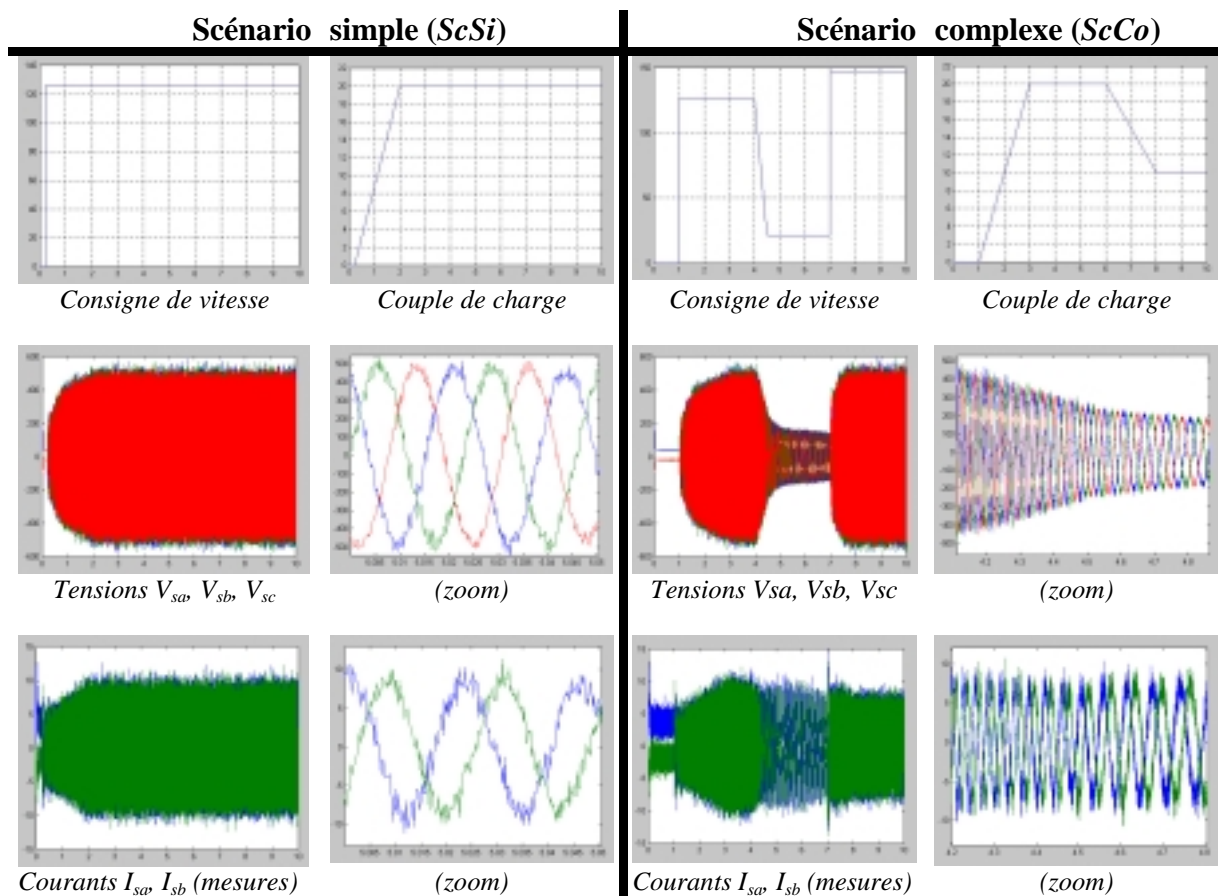


Figure IV.11 : Définition des scénarios simple et complexe pour la consigne de vitesse et pour le couple de charge, tensions et courants correspondants en fonctionnement normal.

Le scénario simple (*ScSi*) est tel que les défauts affecteront le système en régime établi, contrairement au scénario complexe (*ScCo*). La Figure IV.11 met en évidence que l'excitation

induite par les tensions de commande est très localisée en fréquence autour du fondamental¹. Des scénarios secondaires complètent (*ScSi*) et (*ScCo*) pour décrire les autres perturbations et les défauts (Table IV.3).

Notation	Descriptif	Nature	$t_{début}$	t_{fin}	Valeur de référence (100%)
R	Variations identiques de R_s et de R_r (perturbation assimilée à l'échauffement du moteur)	Rampe	2s	7s	Valeurs nominales
L	Variations identiques de L_s et de L_r (On suppose qu'un défaut moteur se traduit au moins en partie par L)	Rampe	2s	7s	Valeurs nominales
f_{Ω}	Défaut du capteur de vitesse	Rampe	5s	6s	125 rad/s
$f_V = [f_{Vsa} f_{Vsb} f_{Vsc}]^T$	Modélisation des défauts du type interrupteur ouvert dans l'onduleur	Echelon	4s	-	500 V
$f_I = [f_{Isa} f_{Isb}]^T$	Défaut des capteurs de courant	Echelon	4s	-	10 A

Table IV.3 : définition des scénarios étudiés

Un filtrage des résidus est souvent nécessaire pour augmenter le rapport signal sur bruit. $F_v(r)$ désigne le résultat du filtrage de r par le filtre discret passe-bas du premier ordre $(1-v)/(z-v)$. Tout d'abord, l'allure de r_T est présentée en l'absence d'adaptation ($\hat{d}_T = 0$), ce qui donne une référence pour juger de l'utilité de cette adaptation (Figure IV.12). Le filtre $F_{0.75}(r_T)$ sert uniquement à améliorer le rapport signal sur bruit (les fréquences autour du fondamental ne sont pratiquement pas atténuées, y compris à haute vitesse).

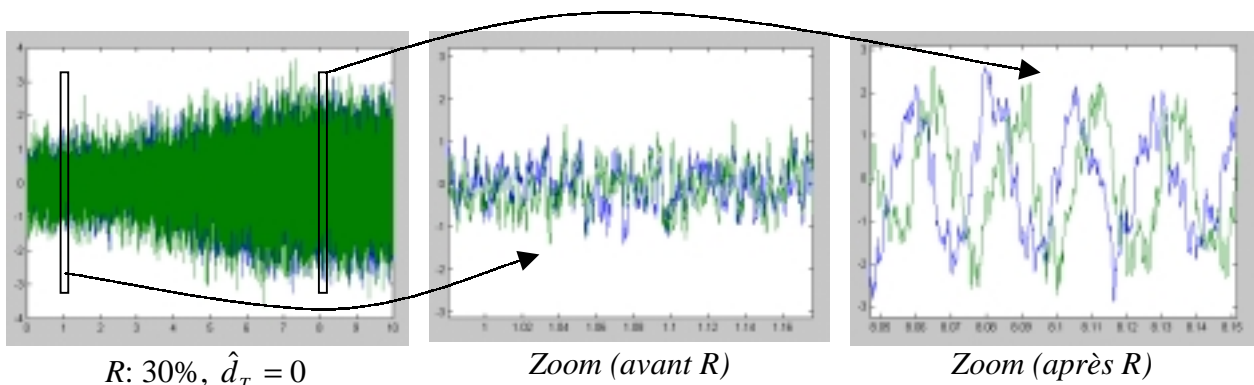


Figure IV.12 : Résidu $F_{0.75}(r_T)$ pour des variations des résistances, en l'absence d'adaptation

L'influence sur r_T de chacun des scénarios de la Table IV.3 est donnée par la Figure IV.13 et par la Figure IV.14 (zoom temporel). En comparant les amplitudes de r_T pour $R : 30\%$ sur la Figure IV.12 d'une part, sur la Figure IV.13 et sur la Figure IV.14 d'autre part, on peut constater que l'adaptation de \hat{d}_T a bien contribué à découpler r_T des variations des résistances modélisées par d_T . Compte tenu de l'excitation sinusoïdale induite par les tensions, les défauts se traduisent par une composante sinusoïdale sur les résidus. Cette composante est même la seule dans le cas de L et de f_{Ω} : en effet, ces défauts sont modélisés par des variations de paramètres dans les équations électriques : ils ne peuvent donc être mis en évidence qu'au travers de l'excitation appliquée à l'entrée du système. Les défauts de l'onduleur (par exemple, f_{Vsa}) et les défauts des capteurs de

¹ la Modulation de Largeur d'Impulsions (MLI) associée à l'onduleur se traduit par des composantes à hautes fréquences dont l'influence sur les résidus s'est révélée totalement négligeable (étude en simulation).

courant (par exemple, f_{Isa}) induisent une excitation qui se superpose à celle des tensions u . Le gain statique du système étant non nul, une composante continue apparaît sur r_T . **Cette composante continue sert à détecter et à localiser f_V et f_I des autres défauts.** Elle est estimée par un filtrage passe-bas de r_T : $F_{0,999}(r_T)$ (Figure IV.15). Nous pouvons déjà noter que l'ordre de grandeur relatif entre les composantes continues de deux éléments de r_T est différent pour f_{Vsa} et f_{Isa} . Ce point sera généralisé au chapitre V pour localiser les défauts du type f_V et du type f_I .

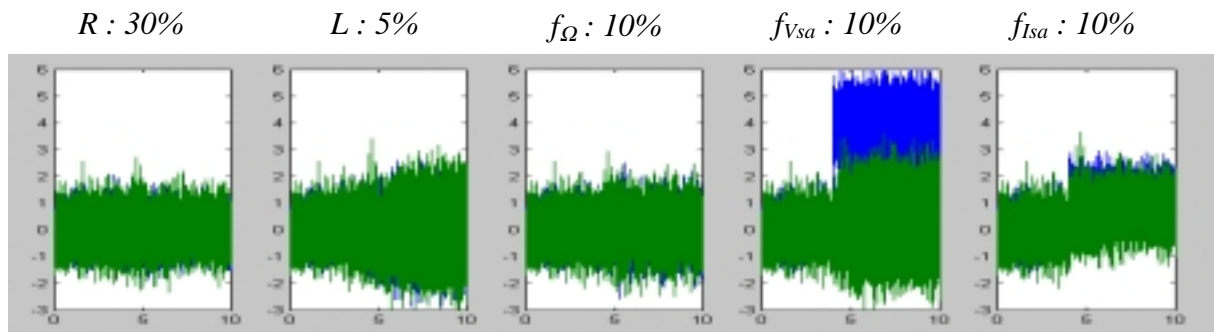


Figure IV.13 : Résidu $F_{0.75}(r_T)$ pour différents scénarios simples (ScSi)

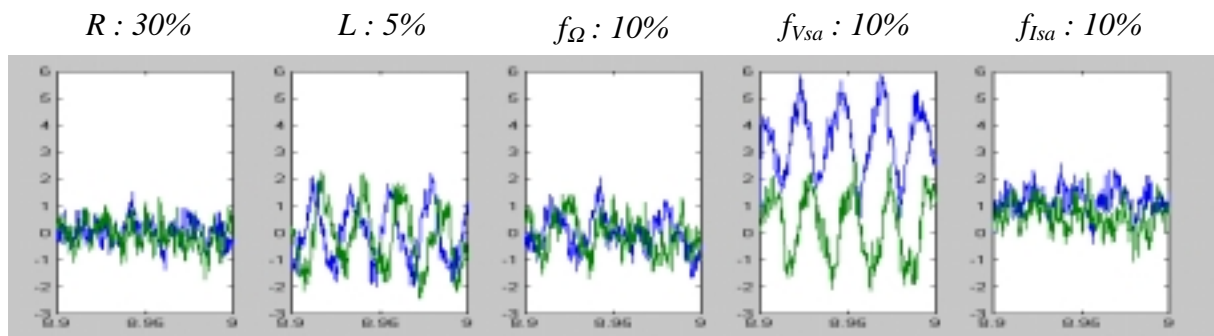


Figure IV.14 : Résidu $F_{0.75}(r_T)$ pour différents scénarios simples (ScSi), zoom temporel

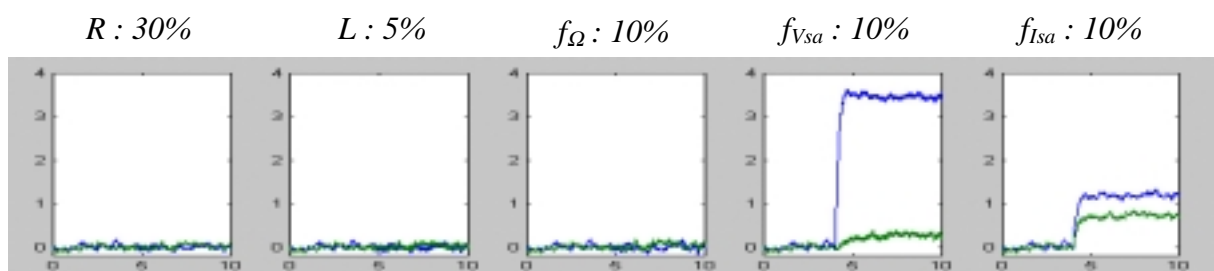


Figure IV.15 : Résidu $F_{0.999}(r_T)$ pour différents scénarios simples (ScSi)

Les défauts L et f_Ω se traduisant par une composante sinusoïdale très bruitée sur le résidu, leur détection est difficile directement à partir de r_T (Figure IV.14) : un filtrage simple se révèle insuffisant ; en effet, une atténuation forte du bruit entraîne également une atténuation importante du signal (la composante sinusoïdale). Le supplément d'information permettant de réaliser la détection va être donné par le gradient de r_T déjà calculé par le générateur de résidus.

IV.4.4.5. Interprétation locale du gradient, découplage direct du résidu adapté

L'adaptation à l'échauffement du résidu r_T repose sur le calcul du gradient $\delta_T = \partial r_T / \partial \hat{d}_T$ ((IV.42) et (IV.43)). $\delta_{T,k}$ représente, localement et à l'instant k , la direction dans laquelle r_T se déplace lorsque \hat{d}_T augmente. En supposant que les erreurs de modélisation soient négligeables et que $\hat{d}_T \approx d_T$, δ_T est une très bonne approximation de la direction dans laquelle d_T (et non plus \hat{d}_T) agit sur r_T . La composante de r_T parallèle à δ_T , notée r_T^{\parallel} , est celle qui peut s'expliquer par des variations de d_T . Si l'adaptation de \hat{d}_T était parfaite, cette composante serait nulle ; or, en pratique, il existe toujours un décalage entre une variation de d_T et la convergence de \hat{d}_T . La composante de r_T orthogonale à δ_T , notée r_T^{\perp} , est celle qui ne peut pas s'expliquer par des variations de d_T .

$$r_{T,k}^{\parallel} = \frac{\delta_{T,k}^T \cdot r_{T,k}}{\|\delta_{T,k}\|}, \quad r_{T,k}^{\perp} = \frac{\delta_{T,k}^{\perp T} \cdot r_{T,k}}{\|\delta_{T,k}^{\perp}\|} \quad (\text{IV.46})$$

Remarque : dans (IV.46), la direction orthogonale à celle du gradient, δ_T^{\perp} , est unique car δ_T est un vecteur de dimension 2. La généralisation à une dimension supérieure à 2 est donnée par :

$$r_{T,k}^{\parallel} = \frac{\delta_{T,k}^T \cdot r_{T,k} \cdot \delta_{T,k}}{\delta_{T,k}^T \cdot \delta_{T,k}}, \quad r_{T,k}^{\perp} = r_{T,k} - r_{T,k}^{\parallel} \quad (\text{IV.47})$$

La Figure IV.16 illustre les positions relatives de r_T et de δ_T dans deux situations : dans le premier cas, les variations des paramètres du système sont assimilables à une variation de d_T ; dans le second cas, elles ne le sont pas.

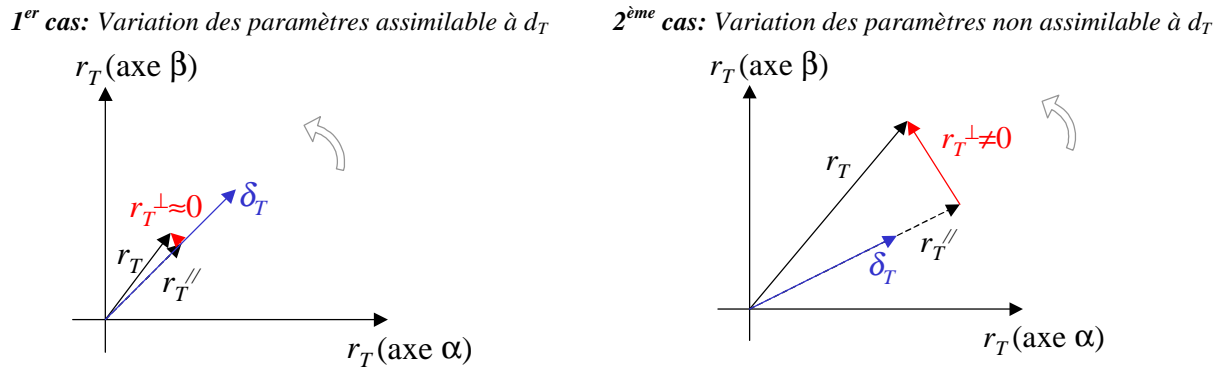


Figure IV.16 : Interprétation locale du gradient

Le résidu r_T^{\perp} est peu sensible aux transitoires liés à la convergence de l'estimation et reste proche de zéro lorsque les variations des paramètres sont assimilables à d_T (1^{er} cas). Il est par contre sensible à toute variation des paramètres non assimilable à d_T (2^{ème} cas) : ce résidu satisfait donc les objectifs fixés pour la détection. De plus, le résidu r_T et son gradient δ_T sont soumis à la même excitation (sinusoïdale) de l'entrée ((IV.42) et (IV.43)). Ainsi, r_T et δ_T sont

deux vecteurs tournant à la même vitesse¹ dans le plan $(\alpha\beta)$. Par conséquent, le résidu r_T^\perp présente une composante continue non nulle lorsqu'une variation des paramètres non assimilable à d_T se produit (et nulle dans le cas contraire). La présence de cette composante continue permet d'utiliser un simple filtre passe-bas pour augmenter le rapport signal sur bruit et améliorer ainsi la sensibilité de la détection. La Figure IV.17 montre que les défauts L et f_Ω peuvent être aisément détectés à partir de r_T^\perp ; de plus, r_T^\perp est très peu sensible à R .

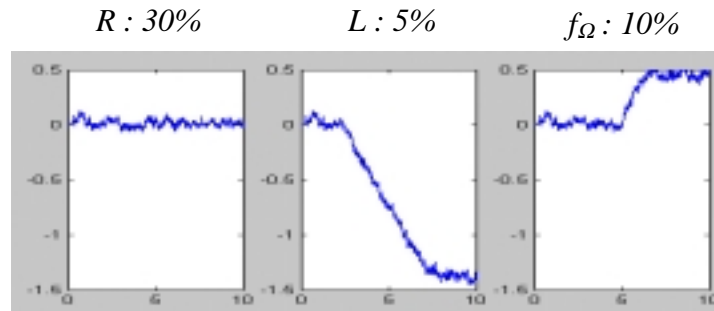


Figure IV.17 : Résidu $F_{0,999}(r_T^\perp)$ pour différents scénarios simples (ScSi)

Le transitoire de faible amplitude sur r_T^{\parallel} pour le scénario R (échauffement) illustre le décalage entre \hat{d}_T et la valeur d_T « réelle » (Figure IV.18). Pendant toute la durée de la rampe (de 2s à 7s), l'algorithme du gradient « court » après la valeur « réelle » de d_T sans y parvenir complètement. Toutefois, l'approximation linéaire étant localement très bonne, r_T^\perp reste pratiquement nul dans ces conditions (Figure IV.17).

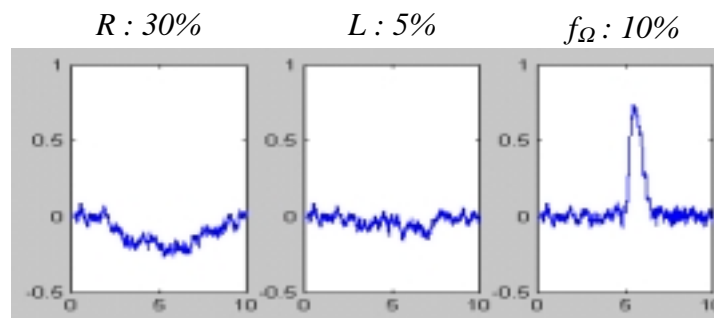


Figure IV.18 : Résidu $F_{0,999}(r_T^{\parallel})$ pour différents scénarios simples (ScSi)

Le résidu r_T^{\parallel} est très faible pour le scénario L (Figure IV.18). Cela signifie que les variations des inductances ressemblent très peu aux variations de résistances. Ce résultat confirme celui de l'étude graphique de la sensibilité, dans laquelle la surface associée aux variations des résistances se croisait nettement avec celle associée aux variations des inductances. Par contre, l'action de f_Ω est partiellement assimilable à des variations de résistances, surtout lorsque f_Ω apparaît (entre 5s et 6s). Cette ressemblance va être confirmée par une nouvelle étude graphique de la sensibilité menée dans le prochain paragraphe (IV.4.5).

¹ Cette vitesse peut très bien varier dans le temps : fonctionnement à vitesse variable de la MAS (scénarios (ScCo))

Afin d'étudier le comportement de r_T^\perp dans des conditions plus réalistes, des scénarios complexes (*ScCo*) ont été simulés et les défauts apparaissent simultanément avec des variations des résistances (échauffement). De plus, le modèle des variations des résistances en fonction de l'échauffement (IV.38) n'est pas parfaitement connu (mention « Approx. » sur la Figure IV.19) : alors que la simulation du système physique est réalisée avec $k_s=k_r$, le générateur de résidus est calculé pour $k_s=0.8*k_r$, soit un écart de 20% sur l'ordre de grandeur relatif des variations de R_s et de R_r .

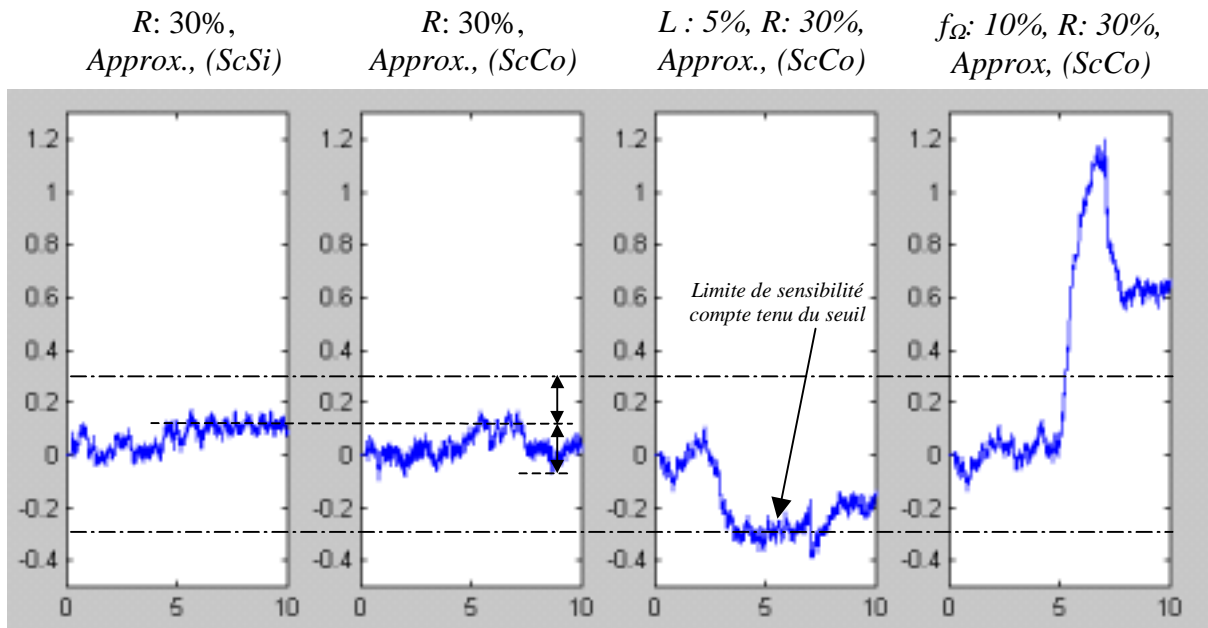


Figure IV.19 : Résidu $F_{0,999}(r_T^\perp)$ avec des incertitudes sur le modèle d'échauffement (Approx.) et pour des scénarios simples (*ScSi*) ou complexes (*ScCo*)

Des simulations, qui ne sont pas reportées dans ce mémoire, montrent que l'effet d'« Approx. », d'une part, l'occurrence simultanée des défauts et de R , d'autre part, ne modifie pas de manière très significative les résultats de la Figure IV.17. Le passage d'un scénario simple (*ScSi*) à un scénario complexe (*ScCo*) influe plus nettement sur la sensibilité aux défauts et, en particulier, sur la sensibilité aux variations des inductances : des variations de 5% correspondent alors à la limite de sensibilité (Figure IV.19).

Remarque 1 : une interprétation possible pour la faible influence d'« Approx. » est la suivante : même si le modèle de l'échauffement n'est pas très précis, il modélise ce dernier par des variations des résistances (et pas des autres paramètres) et il indique que ces variations se font dans le même sens (lorsque k_s et k_r sont de même signe). Sans tenir compte de l'ordre de grandeur relatif des variations des résistances, ce type d'information est déjà relativement riche.

Remarque 2 : en l'absence de défaut, \hat{d}_T donne une indication sur l'état thermique du moteur. Notons qu'une valeur anormalement élevée de \hat{d}_T peut être utilisée pour décider d'un arrêt du moteur.

Bilan : Le résidu r_T et sa composante orthogonale au gradient, r_T^\perp , nous ont permis de remplir les objectifs d'une détection robuste à la fois au couple de charge et aux variations des résistances induites par un échauffement. La composante continue de r_T permet de détecter les défauts du type interrupteur ouvert dans l'onduleur (f_V) et du type capteur de courant (f_I) et de localiser ce groupe de défauts. Les défauts paramétriques (défauts moteur, f_P , et défaut du capteur de vitesse, f_Ω) sont quant à eux détectés par l'intermédiaire de r_T^\perp .

IV.4.5. Application à la MAS : localisation d'un défaut sur le capteur de vitesse

IV.4.5.1. Modèle

Le bilan de la partie précédente (IV.4.4) montre qu'il reste, d'une part, à localiser entre eux les défauts de f_I et de f_V et, d'autre part, à localiser le défaut du capteur de vitesse f_Ω des autres défauts susceptibles d'affecter le moteur. Ce deuxième point porte sur des défauts paramétriques. Une manière de le résoudre consiste à construire un résidu découplé de f_Ω . Comme précédemment, la robustesse au couple de charge est obtenue en utilisant les seules équations électriques. Ainsi, la famille de modèles qui devrait être utilisée est décrite par (IV.48) :

$$M(f_\Omega, d_T) : \begin{cases} \dot{x}_e = A(\Omega - f_\Omega, d_T).x_e + B.K_{23}.u \\ y_e = K_{22}^{-1}.C.x_e \end{cases} \quad (\text{IV.48})$$

Choisir $M(f_\Omega, d_T)$ comme famille de modèles permet a priori de générer un résidu découplé de f_Ω et robuste à l'échauffement. Nous n'avons pas retenu cette famille de modèles car elle nécessite l'estimation de deux paramètres simultanément. En dehors des limites propres à l'algorithme du gradient¹ et de la charge de calcul importante dans l'optique d'une implantation en ligne, l'estimation de deux paramètres risque de poser problème compte tenu de l'excitation du système, presque parfaitement sinusoïdale. Bien que la MLI induise des raies à haute fréquence, l'excitation est en fait très pauvre. Afin de ne pas perturber les spécifications de la commande ou de rejet des harmoniques, nous avons choisi de ne pas appliquer une excitation supplémentaire en entrée du système et de n'estimer qu'un seul paramètre, f_Ω . La famille de modèles effectivement retenue est donc $M(f_\Omega, d_T=0)$. La méthode du gradient va être appliquée à nouveau pour générer un résidu r_Ω découplé de f_Ω .

Remarque : le fait de ne pas estimer d_T peut être en partie compensé : d_T est déjà estimé par le résidu de détection r_T . Avant qu'un défaut apparaisse, autrement dit, tant que r_T (et r_T^\perp) est proche de zéro, la valeur de d_T estimée pour la détection est valide et peut être utilisée à la place de la valeur « 0 » pour le calcul de r_Ω .

IV.4.5.2. Identifiabilité

Le défaut du capteur de vitesse f_Ω apparaît sous la forme d'un paramètre dans les équations électriques (IV.48). Cependant, rien ne garantit que l'hypothèse de variations lentes soit satisfaite pour ce paramètre : une déconnexion du capteur de vitesse peut se traduire par un défaut abrupt

¹ L'algorithme du gradient peut amener à des paramètres fortement corrélés si le nombre de grandeurs estimées est trop grand [Grellet, 1997].

(échelon) sur f_Ω . Nous allons néanmoins poursuivre notre étude avec l'hypothèse de variations lentes de f_Ω et étudier a posteriori les modifications de comportement induites par un défaut abrupt.

Pour établir un résultat sur l'identifiabilité globale de f_Ω dans ce contexte, nous cherchons à manipuler les équations de $M(f_\Omega, d_T=0)$ (IV.48), de façon à mettre en évidence une régression linéaire en f_Ω conforme à (IV.27). En tenant compte de la linéarité de \underline{a} et de \underline{b} par rapport à f_Ω (IV.35), la représentation complexe des équations électriques (IV.36) est de la forme :

$$\underline{p} \cdot f_\Omega + \underline{q} = 0 \Leftrightarrow \begin{cases} \Re(\underline{p}) \cdot f_\Omega + \Re(\underline{q}) = 0 \\ \Im(\underline{p}) \cdot f_\Omega + \Im(\underline{q}) = 0 \end{cases} \quad (\text{IV.49})$$

\underline{p} et \underline{q} sont des polynômes à valeurs complexes qui dépendent des mesures et de leurs dérivées. La partie réelle (ou imaginaire) de (IV.49) est une expression linéaire en f_Ω ayant la forme de (IV.27), ce qui prouve l'identifiabilité globale de f_Ω lorsque f_Ω est le seul paramètre estimé.

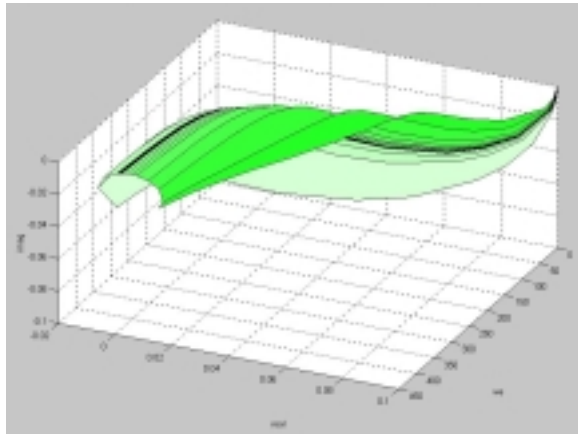
IV.4.5.3. Sensibilité

Une étude graphique des sensibilités du transfert $\underline{h}(j\omega)$ a été menée pour évaluer qualitativement le comportement attendu du résidu r_Ω . Pour éviter les contraintes inhérentes à la dimension 4 (!)¹, nous avons choisi de représenter l'évolution de $\underline{h}(j\omega)$ en fonction du glissement gl .

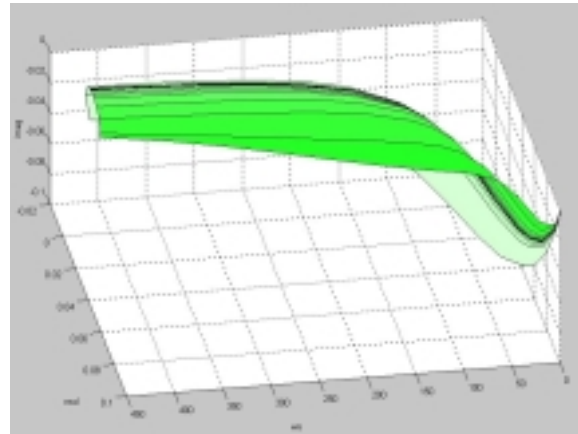
$$gl = \frac{\omega - p \cdot (\Omega - f_\Omega)}{\omega} \quad (\text{IV.50})$$

ω représente alors la pulsation des grandeurs électriques (excitation sinusoïdale). p désigne le nombre de paires de pôles du moteur. gl est de l'ordre de quelques pour-cent en régime établi. La courbe noire correspond à une situation où le glissement serait nul (Figure IV.20, (a) et (b)). Il est intéressant de noter que pour des fréquences proches de zéro, $\underline{h}(j\omega)$ est très peu sensible à gl (et même plus du tout en régime statique). Ce constat est à rapprocher des difficultés de la commande sans capteur (de vitesse) de la MAS à très basse vitesse.

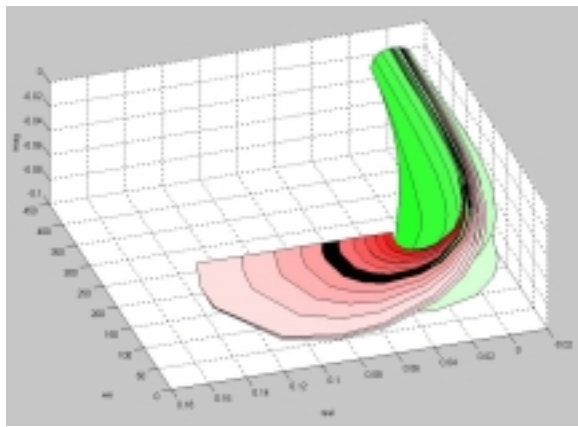
¹ Le transfert $\underline{h}(j\omega)$ est une fonction de Ω et de f_Ω . Pour tracer la surface correspondant aux variations de f_Ω , quatre axes sont donc a priori nécessaires : $(\omega, \Re(\underline{h}(j\omega)), \Im(\underline{h}(j\omega)), \Omega)$



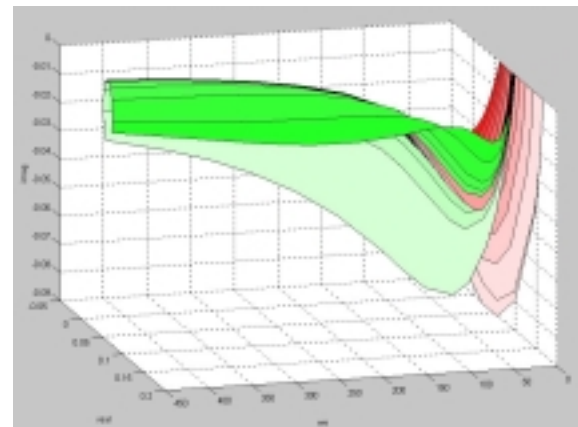
(a) Variation du glissement



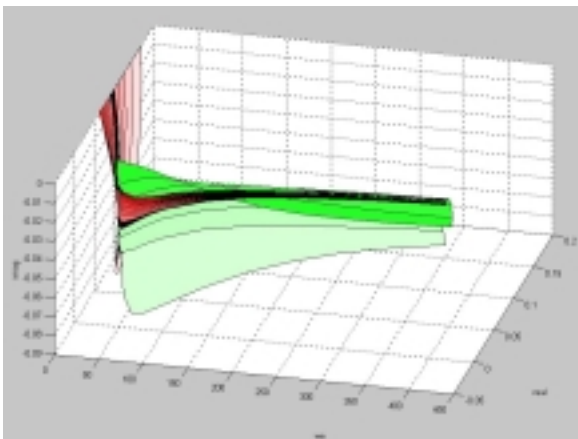
(b) Variation du glissement (bis)



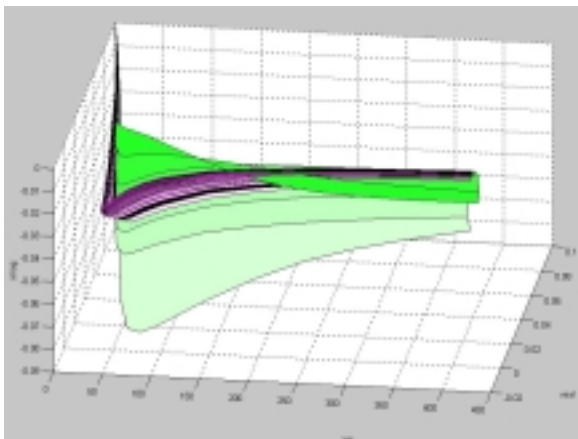
(c) Variation des résistances (rouge) et du glissement (vert)



(d) Variation des résistances (rouge) et du glissement (vert), (bis)



(e) Variation des résistances (rouge) et du glissement (vert), (ter)



(f) Variation des inductances (violet) et du glissement (vert),

Figure IV.20 : Etude de l'influence des variations des résistances et des inductances par rapport à celles du glissement sur le transfert des équations électriques de la MAS

La Figure IV.20 montre qu'en dehors des basses fréquences, la surface associée aux variations des résistances est à peu près tangente à celle associée au glissement (c), (d), (e). Comme, de plus, l'influence des variations des résistances est plus faible que celle du glissement (largeur relative des surfaces rouges et vertes), cela laisse supposer que :

- Le découplage des variations des résistances (d_T) conduit à un résidu (r_T) qui reste sensible à f_Ω . Ceci a déjà été vérifié précédemment par les simulations de r_T (Figure IV.17).
- Le découplage de f_Ω conduit à un résidu (r_Ω) peu sensible aux variations des résistances dues à l'échauffement (d_T). Ce dernier point (qui sera vérifié en simulation : Figure IV.21) fournit une justification supplémentaire pour le choix de la famille de modèles $M(f_\Omega, d_T=0)$ conduisant à l'estimation d'un unique paramètre.

On peut également remarquer que la surface associée aux variations des inductances (violet) croise beaucoup plus nettement la surface associée au glissement (vert) que la surface associée aux variations des résistances (rouge) (comparer (e) et (f) sur la Figure IV.20). Cela laisse supposer que le résidu r_Ω sera nettement plus sensible aux variations des inductances qu'aux variations des résistances. Par chance, cette situation correspond à celle souhaitée pour remplir les objectifs de localisation robuste (à l'échauffement) des défauts paramétriques.

IV.4.5.4. Générateur de résidus

L'application de la méthode du gradient à la famille de modèles $M(f_\Omega, d_T=0)$ conduit au générateur de résidus donné par (IV.51) à (IV.53). L'expression du résidu r_Ω , exprimé dans le repère $(\alpha\beta)$, résulte de l'application de (IV.18), (IV.23), (IV.24) et (IV.25) à (IV.48) avec $d_T=0$:

$$\begin{bmatrix} \dot{\hat{x}}_e \\ \frac{\partial \hat{x}}{\partial \hat{f}_\Omega} \end{bmatrix} = \begin{bmatrix} A(\Omega - \hat{f}_\Omega) & 0 \\ \frac{\partial A(\Omega - \hat{f}_\Omega)}{\partial \hat{f}_\Omega} & A(\Omega - \hat{f}_\Omega) \end{bmatrix} \begin{bmatrix} \hat{x}_e \\ \frac{\partial \hat{x}_e}{\partial \hat{f}_\Omega} \end{bmatrix} + \begin{bmatrix} B \\ 0 \end{bmatrix} \cdot K_{23} \cdot u \quad (IV.51)$$

$$\begin{aligned} r_\Omega &= K_{22} \cdot y_e - C \cdot \hat{x}_e \\ \delta_\Omega &= \frac{\partial r_\Omega}{\partial \hat{f}_\Omega} = -C \cdot \frac{\partial \hat{x}_e}{\partial \hat{f}_\Omega} \end{aligned} \quad (IV.52)$$

$$\hat{f}_{\Omega, k+1} = \hat{f}_{\Omega, k} - \eta_\Omega \cdot \delta_{\Omega, k}^T \cdot r_{\Omega, k} \quad (IV.53)$$

Le réglage retenu pour le gain d'adaptation est $\eta_\Omega=0.05$. La discrétisation du générateur de résidus continu est obtenue en développant jusqu'à l'ordre deux l'exponentielle de matrice correspondante (transposition directe à (IV.51) de la discrétisation de (IV.42) par (IV.45)).

Pour les mêmes raisons que celles évoquées dans la partie IV.4.4.5, nous construisons les composantes de r_Ω dont l'une est parallèle au gradient et l'autre, orthogonale (c'est surtout la composante orthogonale qui nous intéresse) :

$$r_{\Omega, k}^{\parallel} = \frac{\delta_{\Omega, k}^T \cdot r_{\Omega, k}}{\|\delta_{\Omega, k}\|}, \quad r_{\Omega, k}^\perp = \frac{\delta_{\Omega, k}^\perp{}^T \cdot r_{\Omega, k}}{\|\delta_{\Omega, k}^\perp\|} \quad (IV.54)$$

Les notations relatives aux conditions de simulation sont celles données dans la Figure IV.11 et dans la Table IV.3.

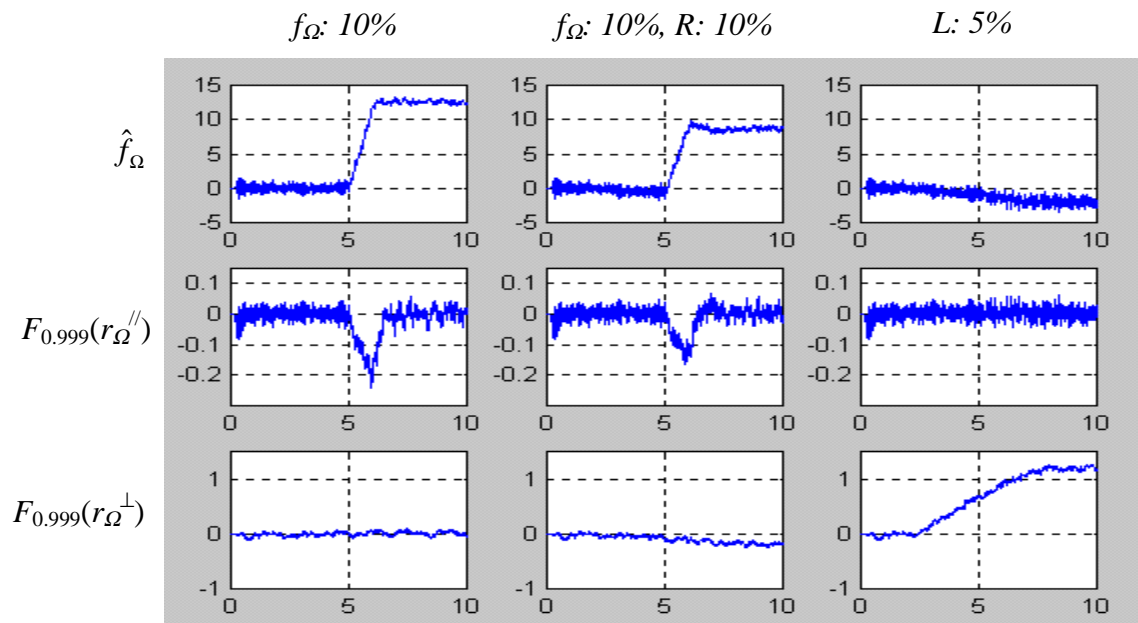


Figure IV.21 : Découplage indirect de f_{Ω} pour des scénarios simples (ScSi)

Les propriétés mises en évidence lors de l'étude de sensibilité sont confirmées par les résultats de la Figure IV.21. Lorsque f_{Ω} se produit (seul), le résidu r_{Ω}^{\perp} reste très proche de zéro, ce qui n'est pas le cas lors d'une variation des inductances. r_{Ω}^{\perp} permet donc de localiser entre eux ces deux situations de défaut. Le découplage des variations des résistances dues à l'échauffement n'est pas parfait mais ce type de variations des paramètres affecte peu r_{Ω} . Le fait que \hat{f}_{Ω} et $r_{\Omega}^{//}$ restent proches de zéro lorsque L se produit confirme les observations faites lors de l'étude de sensibilité : les surfaces associées à L et gl se croisent de manière plutôt orthogonale (Figure IV.20). Par contre, les surfaces associées à R et à gl étant plutôt tangentes, les variations des résistances peuvent être interprétées en partie par une variation de f_{Ω} . Ainsi, bien que \hat{f}_{Ω} soit correctement reconstruit dans le cas où seul f_{Ω} se produit (première colonne de la Figure IV.22), \hat{f}_{Ω} est légèrement biaisé (10 rad/s au lieu de 12.6 rad/s) lorsque des variations des résistances se produisent en même temps que f_{Ω} . Les résultats de la Figure IV.21, obtenus pour des scénarios simples (ScSi), sont confirmés lors de l'application de scénarios complexes (Figure IV.22).

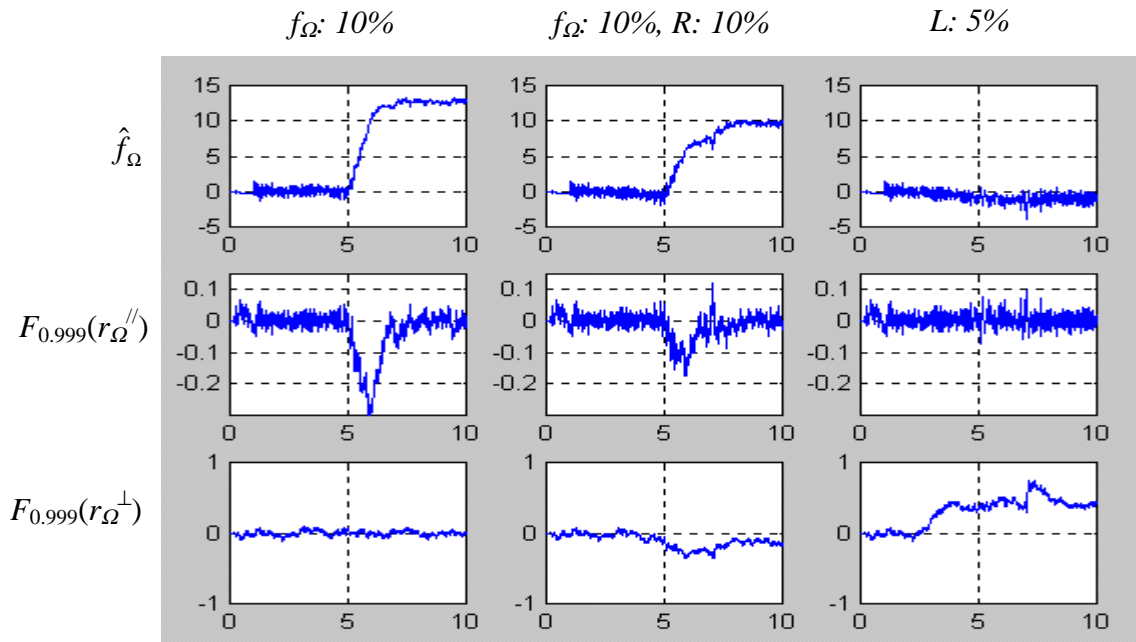


Figure IV.22 : Découplage indirect de f_{Ω} pour des scénarios complexes (*ScCo*)

La Figure IV.21 et la Figure IV.22 montrent que l'hypothèse de variations lentes de f_{Ω} est valide pour une rampe montant de 10% de la vitesse nominale en 1s. Afin de mieux cerner les limites de l'algorithme, des défauts en échelon du capteur de vitesse ont été appliqués à la MAS (Figure IV.23). Que le scénario soit simple (*ScSi*) ou complexe (*ScCo*), un défaut en échelon de 10% n'affecte pratiquement pas r_{Ω}^{\perp} . Par contre, un défaut en échelon de -75% induit un transitoire sur r_{Ω}^{\perp} correspondant au temps nécessaire à l'algorithme du gradient pour converger. Les défauts f_{Ω} en échelon sont en outre correctement estimés (Figure IV.23).

Remarque : La simulation du scénario correspondant à un échelon de -75% sur f_{Ω} a été obtenue dans des conditions particulières. En effet, indépendamment du diagnostic, la simulation de l'entraînement asynchrone en défaut pose problème : la mesure de vitesse défaillante désoriente complètement le repère tournant de la commande vectorielle qui, dans ces conditions, « décroche ». La solution que nous avons adoptée pour s'affranchir de ces problèmes liés à la commande a consisté à injecter la vitesse réelle du moteur dans le calcul de la pulsation du repère tournant. La mesure de vitesse défaillante a toutefois été conservée en entrée du correcteur de la boucle de vitesse. Pour rendre la commande moins sensible aux défauts du capteur de vitesse, une solution pourrait consister à compenser f_{Ω} par son estimation, \hat{f}_{Ω} , en espérant que le tout converge...

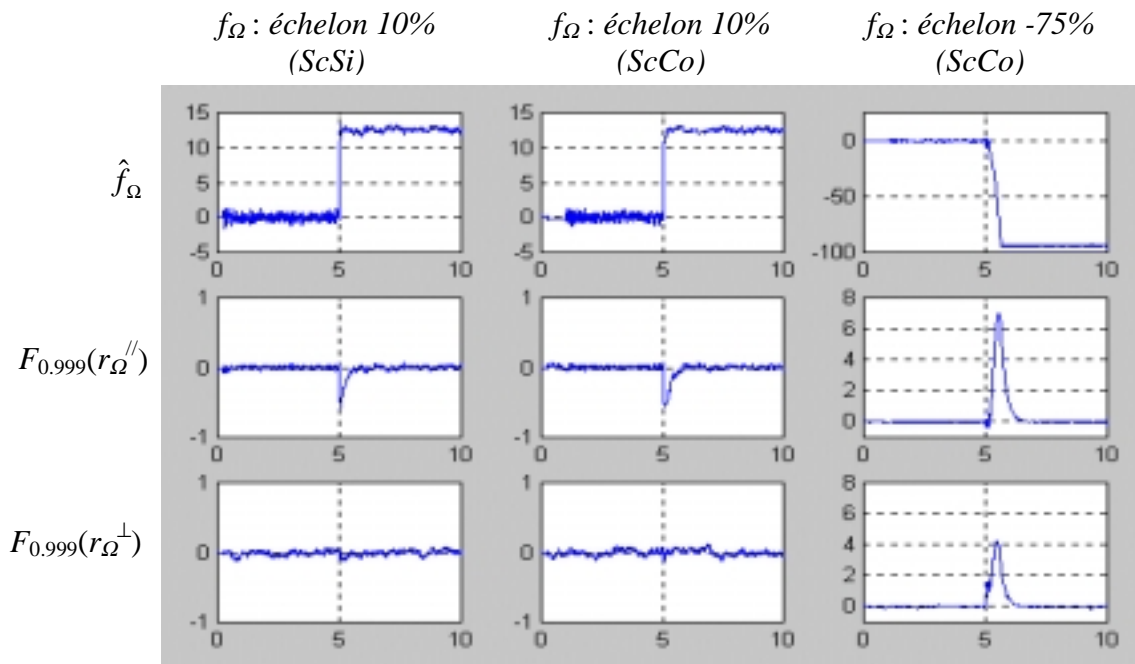


Figure IV.23 : Validité de l'hypothèse de variations lentes pour f_{Ω} .

Bilan : la présence d'un transitoire sur r_{Ω}^{\perp} induit par un fort échelon de f_{Ω} pouvant être traitée par une décision appropriée (chapitre VI), le résidu r_{Ω}^{\perp} permet de localiser entre eux le défaut du capteur de vitesse et les défauts du moteur qui se traduisent par des variations des inductances.

IV.5. Bilan

IV.5.1. Point de vue méthodologique

Les méthodes usuelles de génération des résidus (observateurs, équations de parité) étudiées au chapitre III conduisent à des résidus dont l'implantation en ligne est aisée. Ces méthodes sont pourtant bien souvent limitées par les restrictions sur la nature du modèle et par l'impossibilité d'identifier les défauts. Cette étape d'identification est cependant presque indispensable pour envisager une accommodation aux défauts ou même seulement pour réinitialiser le diagnostic suite à l'occurrence d'un défaut, afin de pouvoir en diagnostiquer un nouveau.

Afin de lever au moins partiellement ces contraintes, nous avons étudié des méthodes de découplage indirect. Ces méthodes consistent à tester la cohérence du comportement du système physique avec une famille de modèles (autrement dit, avec une famille de comportements entrée-sortie (y/u)) par la minimisation d'un critère (lié au résidu). Une technique fondée sur un filtre de Kalman a été utilisée pour découpler et estimer des défauts additifs. Une autre technique, fondée sur un algorithme de programmation non linéaire (gradient), a permis de s'affranchir en grande partie des contraintes liées à la nature du modèle. Même si les contraintes en vue d'une implantation en ligne sont non négligeables, le réglage se révèle beaucoup plus simple que celui d'un filtre de Kalman étendu, puisqu'un seul paramètre doit être réglé lorsqu'on estime un paramètre (gain d'adaptation).

L'estimation d'un nombre restreint de paramètres (voire un seul) dans un générateur de résidus réduit considérablement les contraintes liées à l'excitation du système. En particulier,

l'adaptation d'un unique paramètre a été réalisée sans recourir à une excitation autre que celle qui est « naturellement » appliquée au système. Dans ces conditions, l'adaptation est applicable en ligne sans altérer les spécifications de la commande.

De plus, l'interprétation locale du gradient comme une direction associée à un défaut ou à une perturbation permet, par une combinaison entre un découplage indirect (estimation) et un découplage direct (projection), de limiter l'effet des transitoires liés à la convergence de l'estimation. Si les performances des méthodes de découplage indirect se révèlent intéressantes pour la détection et la localisation, elles dépendent néanmoins de la qualité (validité, précision,...) des modèles de défauts et de perturbations qui peuvent être établis.

IV.5.2. Point de vue des applications

La méthode de découplage indirect appliquée à la MCC repose sur l'utilisation d'un filtre de Kalman. L'influence du réglage de ce filtre a été étudiée. Malgré les contraintes liées au couple de charge (inhérentes à la nature de la MCC), l'application du découplage indirect a mis en évidence la possibilité de détecter, de localiser et d'identifier des défauts additifs.

L'application à la MAS repose sur une combinaison de techniques de découplages directs et indirects. La décomposition de modèle obtenue au chapitre II permet de découpler les effets de la partie mécanique et, en particulier, l'influence du couple de charge. Sur la partie non éliminée (équations électriques), une méthode de découplage indirect a été appliquée : l'adaptation fondée sur un calcul du gradient a permis de remplir les objectifs suivants:

- La détection des défauts est robuste à la fois au couple de charge (perturbation additive) et aux variations des résistances induites par un échauffement (perturbation paramétrique). Un simple filtrage permet également de localiser deux groupes de défauts : les défauts additifs (f_I et f_V) et les défauts paramétriques (f_P et f_Ω).
- La distinction entre un défaut du capteur de vitesse (f_Ω) et les autres défauts paramétriques est également obtenue par l'application de la méthode du gradient.

En l'absence de modèles validés décrivant l'évolution des paramètres du modèle de Park en fonction de l'échauffement et des défauts, nous nous sommes placés dans le cadre fixé par le groupe de travail « action incitative sur la MAS ». Cependant, des variations du même ordre de grandeur pour les résistances d'une part, pour les inductances d'autre part, risquent de se révéler insuffisantes en pratique ; aussi avons nous veillé à développer des outils et des méthodes suffisamment généraux pour être transposés, soit au modèle de Park, lorsque l'évolution des paramètres en fonction de l'échauffement et en fonction des différents défauts sera connue, soit à d'autres modèles de défauts.

Chapitre V :
Décision statique

Chapitre V :

Décision statique

V.1. Introduction

L'objectif du diagnostic est de détecter, de localiser (et éventuellement d'identifier) les défauts. Détecter consiste à déterminer si le système présente ou non des défauts et localiser consiste à préciser quels défauts se sont produits. La conception d'un système de diagnostic (localisation) se ramène donc en général à construire un indicateur D_j pour chaque défaut f_j ($j=1 \dots n_f$) précisant si ce dernier est survenu ou non. Le premier niveau d'observation correspondant aux résidus r_i , $i=1 \dots n_r$, ne remplit bien souvent pas directement les objectifs de la localisation : il faudrait pour cela avoir réussi à structurer les résidus sous la forme d'une table de signature diagonale. Dès lors, la décision pour le diagnostic consiste à résoudre le problème de localisation des défauts à partir des observations disponibles : mesures ou observations plus élaborées telles que les résidus. Dans le contexte des entraînements électriques, cette deuxième alternative nous semble la plus appropriée¹.

Nous avons vu au chapitre I que le principe général d'un module de diagnostic consistait à construire des observations porteuses d'une information relative aux défauts et à eux seuls en évaluant la cohérence entre des observations élémentaires et des comportements décrits par un modèle approprié (Table I.8). La génération de résidus, d'une part, la décision par table de décision, d'autre part, sont des cas particuliers de modules de diagnostic (Table V.1), généralement associés en série (Figure I.11).

¹ Il serait en effet dommage de se priver de la connaissance fournie par les modèles numériques classiques, à la fois simples et relativement précis... à moins de trouver mieux...

<i>Principe du diagnostic</i>	<i>Génération de résidus</i>	<i>Décision</i>
<p>Partant de la connaissance fournie par un modèle:</p> <p>Un modèle pour le diagnostic peut être synthétisé:</p> <p>La simulation du modèle pour le diagnostic permet d'évaluer la cohérence d'observations élémentaires :</p> <p>Avec certains comportements décrits par le modèle initial :</p> <p>La grandeur correspondant à l'évaluation de la cohérence est conçue pour être une observation porteuse d'une information sur les défauts et sur eux seuls:</p>	<p>Modèle du système</p> <p>Générateur de résidus</p> <p>Entrées (u), mesures (y)</p> <p>Comportements entrée / sortie normaux (approche bon fonctionnement) ou défaillants (approche mauvais fonctionnement)</p> <p>Résidus (r)</p>	<p>Modèle pour la décision</p> <p>Fonctions de décision¹</p> <p>Résidus (r)</p> <p>Comportements des résidus en fonctionnement normal (détection) ou en réponse aux défauts (localisation)</p> <p>Indicateurs de défaut (D)</p>

Table V.1 : Génération de résidus et Décision: deux modules de diagnostic construits avec des modèles différents.

Il existe une grande variété de modèles pour la décision. Une approche possible consiste à utiliser des tables de décision. Ces dernières sont des modèles de mauvais fonctionnement (MF), dans la mesure où elles décrivent le comportement de chacun des résidus en réponse à chacun des défauts. Cette approche présente en général l'avantage de la simplicité en vue d'une implantation en ligne ; les contraintes temps-réel sont en effet importantes dans le contexte du diagnostic des entraînements électriques. Un modèle pour la décision peut être défini à l'aide d'une connaissance a priori (connaissance experte) ou résulter d'un apprentissage.

En ce qui concerne les fonctions de décision, « comme les comportements temporels (dynamique) [...] ne sont pas connus en général, les relations sont supposées statiques » [Isermann, 1997b]. L'idée directrice de ce chapitre, portant sur une décision statique, consiste à proposer différentes approches pour enrichir le modèle « table de décision » par rapport au cas booléen. Deux principales approches sont proposées : l'une repose sur la sélection de résidus pertinents, compte tenu de critères d'évaluation des fonctions de décision ; l'autre approche repose sur la prise en compte des ordres de grandeur relatifs des résidus.

¹ Par exemple, ces fonctions peuvent reposer sur une mise en forme des résidus (interfaces correspondant à un seuillage, fuzzification,...) et sur une ou plusieurs équations logiques.

V.2. Sélection des résidus

V.2.1. Table de décision booléenne

Une table de décision est définie par une matrice Sg de n_r lignes et de n_f colonnes où n_r est le nombre de résidus et n_f le nombre de défauts. Elle modélise le comportement des résidus en réponse à chacun des défauts ; le modèle « table de décision » conduit donc à une approche du diagnostic orientée mauvais fonctionnement : on cherche à savoir si le comportement réel du système, connu au travers des résidus, est cohérent ou non avec une classe de défaut donnée (et non avec un comportement normal).

Sg	...	f_j	...
...
r_i	...	Sg_{ij}	...
...

Table V.2 : Table de décision

Une table de décision est dite « booléenne » lorsque les éléments de Sg sont à valeur dans $\{0;1\}$. $Sg_{ij}=0$ (resp. $Sg_{ij}=1$) signifie que le résidu r_i est insensible (resp. sensible) au défaut f_j .

Remarque : les tables de **décision** booléennes sont similaires aux tables de **signatures** booléennes présentées au début du chapitre III (partie III.2) mais leur objectif est différent. Alors qu'une table de signature spécifie les objectifs de la génération de résidus, une table de décision est un modèle pour la décision. Nous continuerons à appeler « signature » une colonne d'une table de décision.

Pour savoir si un résidu est ou n'est pas sensibilisé à un instant donné, il est nécessaire de tester la valeur de ce résidu. Dans un contexte booléen, le test de r_i résulte d'une comparaison avec le seuil λ_i qui lui est associé. T_i désigne le résultat du test :

$$T_i = (|r_i| > \lambda_i) \quad (\text{V.1})$$

Le choix du seuil est réalisé de sorte que T_i reste nul (faux) en l'absence de défaut mais aussi en présence de bruit, de perturbations, d'erreurs de modélisation. Ce point a déjà été traité dans la partie III.4. Une manière usuelle de construire la fonction de décision D_j associée à f_j est :

$$D_j = \left(\bigwedge_{i/Sg_{ij}=0} \bar{T}_i \right) \wedge \left(\bigwedge_{i/Sg_{ij}=1} T_i \right) \quad (\text{V.2})$$

où \wedge désigne un ET logique et la barre horizontale, une négation logique. Un OU logique sera noté \vee . Une première manière d'enrichir le contenu d'une table de décision consiste à utiliser les résultats d'une étude de sensibilité des résidus.

V.2.2. Un petit exemple

Afin de mettre en évidence certaines limites d'une approche purement booléenne, un exemple est tout d'abord étudié. Le point de départ de cet exemple est une étude de sensibilité des résidus (en

réponse à des défauts en échelon) telle que celle menée dans la partie III.5.3 pour la machine à courant continu (MCC).

α	f_1	f_2
r_1	>>	40%
r_2	10%	20%
r_3	30%	20%

(a) Etude de sensibilité

Sg	f_1	f_2
r_1	0	1
r_2	1	1
r_3	1	1

(b) Table de décision booléenne

Table V.3 : Un exemple avec deux défauts et trois résidus

La Table V.3(a) est construite sur le même modèle que la Table III.3 : $\alpha_{ij} \in [0;1] \cup \{ '>>' \}$. Lorsque α_{ij} est un nombre compris entre 0 et 1, α_{ij} désigne l'amplitude minimale du défaut f_j pour que r_i dépasse son seuil en réponse à f_j . L'amplitude du défaut f_j est normalisée sous la forme d'un pourcentage par rapport à $\max(f_j)$. Lorsque $\alpha_{ij} = '>>'$, cela correspond à un pourcentage supérieur à 100%, autrement dit, quel que soit $|f_j| < \max(f_j)$, r_i ne dépasse jamais son seuil en réponse à f_j .

Remarque 1 : α_{ij} est inversement proportionnel à la sensibilité de r_i à f_j .

Remarque 2 : $Sg_{ij} = 0$ si $\alpha_{ij} = '>>'$ et $Sg_{ij} = 1$ si $\alpha_{ij} \neq '>>'$ (Table V.3)

Une approche très classique pour localiser les défauts f_1 et f_2 consiste à utiliser les équations logiques déduites de l'application de (V.2) à la table de décision donnée par la Table V.3.

$$D_1 = \overline{T_1} \wedge T_2 \wedge T_3 \quad (\text{V.3})$$

$$D_2 = T_1 \wedge T_2 \wedge T_3 \quad (\text{V.4})$$

Une condition nécessaire pour que D_1 indique l'occurrence de f_1 ($D_1=1$) est que r_2 et r_3 dépassent simultanément leur seuil. De plus, la Table V.3(a) montre que r_2 dépasse son seuil ($T_2=1$) seulement lorsque $|f_1| > 10\%$, alors que r_3 ne dépasse le sien que lorsque $|f_1| > 30\%$ (r_2 est plus sensible à f_1 que r_3). Par conséquent, la décision calculée par (V.3) est seulement sensible à des défauts f_1 dont l'amplitude est supérieure à 30%. Ainsi, la sensibilité à f_j d'une fonction de décision D_j déduite directement d'une table de décision booléenne par (V.2) est égale à la plus petite sensibilité parmi les résidus associés à un '1' dans la signature de f_j . Lorsque certains résidus n'interviennent pas dans la décision ($D_1 = \overline{T_1} \wedge T_2$, par exemple), la sensibilité de la décision peut, dans certains cas, être améliorée : en éliminant r_3 de D_1 , D_1 devient sensible à des défauts $|f_1| > 10\%$ (et non plus seulement $|f_1| > 30\%$).

La sensibilité du diagnostic au défaut f_j peut être améliorée lorsque D_j est calculée en utilisant seulement les résidus les plus sensibles associés à un '1' dans la table de décision.

En supprimant certains résidus de l'expression de D_j , il est cependant nécessaire de vérifier que D_j continue à satisfaire des propriétés de localisation des défauts. Jusqu'à maintenant, nous avons porté notre attention sur la sensibilité des résidus pour un défaut donné (f_1). Cependant, des sensibilités des résidus différentes pour différents défauts peuvent avoir une influence considérable sur les propriétés de localisation des fonctions de décision. Sur la Table V.3, on peut remarquer que $[T_1 T_2 T_3] = [0 1 1]$ lorsque $|f_2|$ est compris entre 20% et 40%. $[T_1 T_2 T_3]$ coïncide alors avec la signature de f_1 et non avec celle de f_2 . Par conséquent, les fonctions de décision (V.3) et (V.4) ne permettent pas une localisation correcte des défauts.

Des sensibilités différentes des résidus peuvent induire des fausses décisions.

Etant donné un jeu de résidus pour lesquels une étude de sensibilité a été menée, nous allons chercher à concevoir des fonctions de décision présentant à la fois de bonnes propriétés de sensibilité et de localisation (au sens de critères qui seront explicités).

V.2.3. Interprétation d'un dépassement de seuil

Dans cette partie, nous explicitons certaines hypothèses faites dans les interprétations de l'exemple de la Table V.3. Ces hypothèses seront utilisées pour généraliser le cas de l'exemple. Dans le cadre que nous nous sommes fixé, la seule information disponible sur le comportement réel du système est donnée par les tests de dépassement de seuil sur les résidus : T_i , $i=1 \dots n_r$. Parallèlement, les sorties des fonctions de décision doivent donner une information sur les défauts (objectif du diagnostic). Dans un premier temps, nous cherchons à interpréter le (non-)dépassement de seuil d'un résidu en terme de connaissance sur les défauts, avant de combiner entre eux les tests sur les résidus pour construire des fonctions de décision.

Soit r un résidu sensible (linéairement) à une perturbation d et à un défaut f . $r=r_d+r_f$ où r_d (resp. r_f) représente l'influence de d (resp. f) sur r . α désigne l'amplitude minimale de f pour que r dépasse son seuil λ en réponse à f seul. Le seuil λ associé à r est choisi de sorte que le test du résidu $T=(|r|>\lambda)$ soit toujours faux en l'absence de défaut et quelle que soit la perturbation, autrement dit, $r_d \in [-\lambda; \lambda]$. Notre objectif étant d'interpréter la valeur de T en terme de connaissance sur les défauts, plusieurs situations caractéristiques sont envisageables (Table V.4).

r_d	Signe de r_f	Condition Nécessaire et Suffisante sur $ f $ pour que $ r >\lambda$ ($=T$) soit vrai
$-\lambda$	-	$ f >0$
$-\lambda$	+	$ f >2\alpha$
0	-	$ f >\alpha$
0	+	$ f >\alpha$
λ	-	$ f >2\alpha$
λ	+	$ f >0$

Table V.4 : Plusieurs situations caractéristiques relatives à un dépassement de seuil

Afin d'élargir la portée de l'exemple de la Table V.4, les notations suivantes sont introduites : $\varepsilon=0$, $\sigma=2\alpha$. Toujours dans un but de généralisation, $|f|$, amplitude de f , peut être vue comme une mesure de la criticité du défaut f . Deux tables de vérités découlent de la Table V.4. Ces deux tables sont indépendantes des perturbations et du signe de l'influence des défauts, qui sont inconnus a priori ; chacune des deux tables établit un lien entre T et la criticité de f :

$ f > \sigma$	$ r > \lambda (=T)$	$ f \leq \varepsilon$	$ r \leq \lambda (= \bar{T})$
0	0 ou 1	0	0 ou 1
1	1	1	1
\Leftrightarrow		\Leftrightarrow	
$ f > \sigma \Rightarrow r > \lambda$		$ f \leq \varepsilon \Rightarrow r \leq \lambda$ (V.5)	
\Leftrightarrow		\Leftrightarrow	
$ r \leq \lambda \Rightarrow f \leq \sigma$		$ r > \lambda \Rightarrow f > \varepsilon$ (V.6)	

Table V.5 : Tables de vérité et équations logiques reliant le résultat du test d'un résidu (dépassement de seuil) à une connaissance sur la criticité d'un défaut.

Les implications de (V.6) traduisent la connaissance sur la criticité de f déduite de la valeur (connue) de T . (V.5) et (V.6) mettent en évidence les significations de ε et de σ :

L'interprétation de ε et de σ déduite de (V.5) est :

ε représente la criticité maximale du défaut f en dessous de laquelle il est certain que T est faux.
 σ représente la criticité minimale du défaut f au dessus de laquelle il est certain que T est vrai.

L'interprétation de ε et de σ déduite de (V.6) est :

ε représente la criticité minimale du défaut f qu'il est possible de détecter avec T .
 σ représente la criticité maximale du défaut f qu'il est possible de ne pas détecter avec T .

Les résultats de la Table V.5 et, en particulier, les implications de (V.5) et de (V.6) peuvent être représentés graphiquement (Figure V.1) :

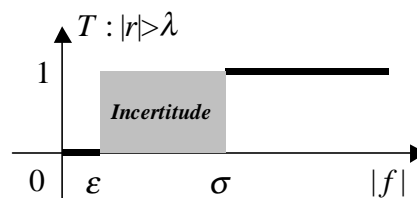


Figure V.1 : Relation entre le test d'un résidu et la criticité d'un défaut

Le sens des implications de (V.5) correspond à celui utilisé pour évaluer les performances d'un générateur de résidu et de son seuil. σ est une mesure de performance du générateur de résidu en terme de sensibilité aux défauts ($\sigma=2\alpha$ dans l'exemple de la Table V.4). ε résulte d'une marge de sécurité dans le choix du seuil par rapport à l'influence maximale des perturbations ($\varepsilon=0$ dans l'exemple de la Table V.4). Le sens des implications de (V.6) correspond à celui utilisé dans le contexte de la décision où l'objectif est de traduire les résultats des tests sur les résidus en terme de connaissance sur les défauts.

La Figure V.1 montre qu'une grandeur certaine, la valeur booléenne de T , se traduit par une connaissance incertaine sur $|f|$. L'apparition de cette incertitude est liée à la présence des perturbations inconnues. La modélisation de cette incertitude est un point clé pour l'interprétation des dépassements de seuil en terme de connaissance sur les défauts. Cette interprétation est au cœur de l'interface entre la génération des résidus **numériques** et l'évaluation des résidus (ou décision) par des fonctions **logiques**.

En remarquant, d'une part, que « T est faux » est d'autant plus certain que $|f|$ est proche de ε et, d'autre part, que « T est vrai » est d'autant plus certain que $|f|$ est proche de σ , la zone d'incertitude de la Figure V.1 peut être « modélisée » de plusieurs manières :

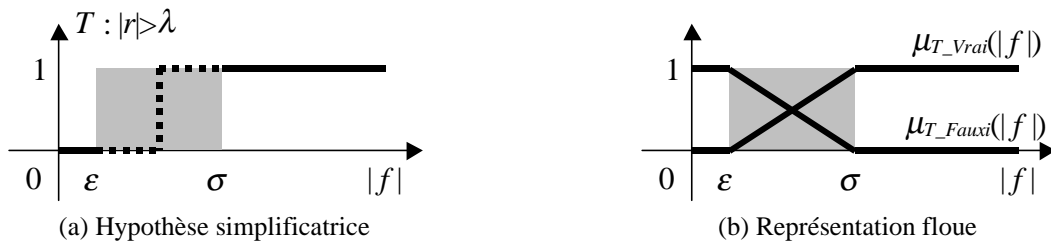


Figure V.2 : Deux modèles de la relation entre le test d'un résidu et la criticité d'un défaut

Même si la représentation floue nous semble la mieux adaptée (Figure V.2), nous n'avons pas retenu cette alternative, faute du temps nécessaire pour la développer. L'hypothèse simplificatrice (Figure V.2a) est assez discutable dans la mesure où elle nie la nature incertaine de la relation entre T et $|f|$. Elle constitue néanmoins un premier pas dans la prise en compte de la sensibilité des résidus pour optimiser le choix des fonctions de décision. L'hypothèse simplificatrice (Figure V.2a) se traduit par la relation (V.7) dans l'exemple de la Table V.4 (un seul résidu : r , un seul défaut : f , $\varepsilon=0$, $\sigma=2\alpha$).

$$|r| > \lambda \Leftrightarrow |f| > \alpha \tag{V.7}$$

(V.7) est une interprétation du dépassement de son seuil par r . La généralisation au cas où plusieurs défauts sont susceptibles d'affecter un résidu r_i est donnée par (V.8), sous l'hypothèse de défaut simple (HDS).

$$T_i = (|r_i| > \lambda_i) = \bigvee_{k=1 \dots n_f} (|f_k| > \alpha_{ik}) \tag{V.8}$$

(V.8) signifie que T_i est vrai si et seulement si l'un des défauts f_k dépasse l'amplitude minimale α_{ik} résultant d'une étude de sensibilité. Notons qu'une combinaison des influences de deux défauts (ou plus) est exclue sous l'hypothèse de défaut simple.

V.2.4. Optimisation de fonctions de décision par sélection des résidus

Afin d'optimiser des fonctions de décision, nous allons commencer par paramétrer un ensemble de fonctions de décision, puis définir des critères permettant d'évaluer leurs performances. Deux méthodes de construction des fonctions de décision sont alors proposées pour satisfaire au mieux les critères de performance : l'une repose sur un algorithme déduit de la formulation du problème, l'autre repose sur une optimisation par un algorithme génétique.

V.2.4.1. Paramétrage des fonctions de décisions

Chaque fonction de décision est définie par deux ensembles disjoints de résidus repérés par leurs indices : $I_0 \subset \{1 \dots n_r\}$, $I_1 \subset \{1 \dots n_r\}$. (V.9) définit la fonction de décision $D(I_0, I_1)$ paramétrée par I_0 et I_1 :

$$D(I_0, I_1) = \left(\bigwedge_{i \in I_0} \bar{T}_i \right) \wedge \left(\bigwedge_{i \in I_1} T_i \right) \text{ où } I_0 \cap I_1 = \emptyset \tag{V.9}$$

Les ensembles de résidus définissant la fonction de décision D_j utilisée pour localiser le défaut f_j sont notés I_0^j et I_1^j :

$$D_j = D(I_0^j, I_1^j) \quad (\text{V.10})$$

Intuitivement, I_0^j (resp. I_1^j) correspond à un sous-ensemble de résidus qui sont peu ou pas affectés (resp. affectés) par l'occurrence de f_j . $I_{0,\max}^j$ (resp. $I_{1,\max}^j$) désigne tous les résidus associés à un '0' (resp. '1') dans la signature booléenne de f_j . On a donc :

$$I_{0,\max}^j = \{i / \alpha_{ij} = ' >> ' \}, \quad I_{1,\max}^j = \{i / \alpha_{ij} \neq ' >> ' \} \quad (\text{V.11})$$

Une construction usuelle de D_j , à partir d'une table de décision booléenne, est donc donnée par $D_j = D(I_{0,\max}^j, I_{1,\max}^j)$, ce qui est équivalent à (V.2). C'est également de cette manière que (V.3) et (V.4) sont déduites de la Table V.3.

V.2.4.2. Evaluation des fonctions de décision

(V.9) définit une fonction de décision à partir des tests des résidus T_i . (V.9) indique donc comment la décision est calculée en ligne. Nous cherchons maintenant à répondre à la question suivante : quelle information sur les amplitudes des défauts peut être déduite à partir du calcul en ligne de $D(I_0, I_1)$? Pour cela, T_i est substitué par l'interprétation d'un dépassement de seuil (V.8) dans l'expression de $D(I_0, I_1)$. Afin de simplifier les équations logiques, $D(I_0, I_1)$ est tout d'abord reformulée (V.13) en tenant compte de l'hypothèse de défaut simple¹ : $Sf=1$ (V.12) (Sf_k correspond à l'occurrence possible du défaut f_k seul).

Hypothèse de défaut simple: $Sf=1$

$$Sf = \bigvee_{k=1 \dots n_f} Sf_k = 1 \text{ avec } Sf_k = \bigwedge_{k \neq l} (f_l = 0) \quad (\text{V.12})$$

Expression de $D(I_0, I_1)$ tenant compte de $Sf=1$:

$$D(I_0, I_1) = D(I_0, I_1) \wedge Sf = \bigvee_k \left[\left(\bigwedge_{i \in I_0} (\overline{T_i} \wedge Sf_k) \right) \wedge \left(\bigwedge_{i \in I_1} (T_i \wedge Sf_k) \right) \right] \quad (\text{V.13})$$

Substitution de l'interprétation de T_i en terme de connaissance sur les défauts:

$$D(I_0, I_1) = \bigvee_k \left(\tilde{f}_k(I_1) < |f_k| \leq \hat{f}_k(I_0) \right) \quad (\text{V.14})$$

où $\tilde{f}_k(I_1) = \max_{i \in I_1} (\alpha_{ik})$, $\hat{f}_k(I_0) = \min_{i \in I_0} (\alpha_{ik})$

Dans (V.14), l'intervalle $[\tilde{f}_k(I_1); \hat{f}_k(I_0)]$ est associé à chaque défaut f_k . Plus précisément, (V.14) établit un lien entre la valeur d'une fonction de décision quelconque, $D(I_0, I_1)$, et une connaissance sur les amplitudes des défauts ; cette connaissance est exprimée sous la forme d'intervalles dont les bornes dépendent de I_0 , de I_1 et de la sensibilité des résidus aux défauts. L'application de (V.9) et de (V.14) avec $I_0 = I_{0,\max}^1 = \{1\}$ et $I_1 = I_{1,\max}^1 = \{2,3\}$ à l'exemple de la Table V.3 donne:

¹ L'hypothèse de défaut simple revient à supposer qu'un défaut au plus peut se produire

$$D(I_0, I_1) = D_1 = \bar{T}_1 \wedge T_2 \wedge T_3 = (30\% < |f_1| \leq' \ll') \vee (20\% < |f_2| \leq 40\%) \quad (\text{V.15})$$

Le symbole ' \ll ' dans (V.15) peut être considéré comme arbitrairement grand. (V.15) indique donc que D_1 devient vrai lorsque $|f_1| > 30\%$ ou bien lorsque $|f_2|$ est compris entre 20% et 40%. Si l'objectif assigné à D_1 est de localiser f_1 , la sensibilité de D_1 à f_2 peut conduire à une fausse décision. Par l'application de (V.14), on retrouve ainsi de manière systématique les constats établis dans la partie V.2.2 sur l'exemple de la Table V.3.

Partant de l'interprétation de $D(I_0, I_1)$ en terme de connaissance sur les défauts, il reste encore à définir des critères de performance destinés à évaluer si une fonction de décision est « bonne » ou « mauvaise ». Une « bonne » fonction de décision D_j permet de localiser f_j des autres défauts tout en étant aussi sensible que possible à f_j .

V.2.4.2.1. Critère de localisation

Dans l'exemple donné par (V.15), il apparaît que D_1 permettrait de localiser f_1 (autrement dit, D_1 serait sensible à f_1 et seulement à f_1) si l'intervalle associé à f_2 était vide. En utilisant (V.14), cette remarque est généralisée afin d'obtenir un critère sur I_0^j et sur I_1^j pour que $D_j = D(I_0^j, I_1^j)$ localise f_j . Le critère de localisation de f_j , $J_{LOC(j)}$, est vérifié si et seulement si, pour tout défaut autre que f_j , l'intervalle des amplitudes qui conduisent D_j à prendre la valeur « vraie » est vide :

$$J_{LOC(j)} = 1 \text{ si } \forall k \neq j, \tilde{f}_k(I_1^j) > \hat{f}_k(I_0^j) \quad (\Leftrightarrow \max_{i \in I_1^j}(\alpha_{ik}) > \min_{i \in I_0^j}(\alpha_{ik})) \quad (\text{V.16})$$

$$J_{LOC(j)} = 0, \text{ sinon}$$

Annuler l'intervalle associé aux défauts $f_{k \neq j}$ revient à découpler $D_j = D(I_0^j, I_1^j)$ des défauts autre que f_j (V.14). Comme nous l'avons mentionné au chapitre I, le diagnostic apparaît avant tout comme un problème de découplage, ce dernier pouvant être réalisé aussi bien avec des modèles numériques (génération de résidus) que logiques (décision).

Remarque : la localisation est d'autant plus robuste aux différentes sources d'incertitude que $\tilde{f}_k(I_1^j)$ est plus grand que $\hat{f}_k(I_0^j)$ dans (V.16).

V.2.4.2.2. Critères de sensibilité

Le critère de localisation porte sur la (non-)sensibilité de D_j aux défauts autres que f_j . Des critères portant sur la sensibilité de D_j au défaut f_j sont maintenant établis. Cette sensibilité est d'autant plus forte que l'intervalle associé à f_j dans D_j , $[\tilde{f}_j(I_1^j); \hat{f}_j(I_0^j)]$, est grand. La borne inférieure doit être minimisée pour assurer une bonne Sensibilité aux Petits Défauts (SPD) ; la borne supérieure doit être maximisée pour assurer une bonne Sensibilité aux Grands Défauts (SGD). Les critères $J_{SPD(j)}$ et $J_{SGD(j)}$ sont définis en tenant compte de la normalisation entre 0 et 1 des coefficients α_{ij} . Ils doivent être maximisés lors de la synthèse de la fonction de décision D_j .

$$J_{SPD(j)} = 1 - \tilde{f}_j(I_1^j) = 1 - \max_{i \in I_1^j}(\alpha_{ij}) \quad (\text{V.17})$$

$$J_{SGD(j)} = \hat{f}_j(I_0^j) = \min_{i \in I_0^j}(\alpha_{ij}) \quad (\text{V.18})$$

V.2.5. Algorithmes déduits de la formulation du problème

V.2.5.1. Analyse

En nous limitant à une forme particulière pour les fonctions de décision (V.9), la synthèse d'une fonction de décision $D_j = D(I_0^j, I_1^j)$ pour le défaut f_j se ramène donc à trouver les ensembles de résidus I_0^j et I_1^j qui maximisent simultanément un critère de localisation de f_j (V.16), un critère de sensibilité aux petits défauts f_j (V.17) et un critère de sensibilité aux grands défauts f_j (V.18). Dans un premier temps, une analyse de ce problème d'optimisation est menée ; cette analyse est développée en trois points :

Analyse 1 : une première solution consiste à parcourir systématiquement l'espace des solutions. Comme chaque résidu peut être soit dans I_0^j , soit dans I_1^j , soit dans le complément de $I_0^j \cup I_1^j$, un balayage de l'ensemble des solutions conduirait à calculer les critères 3^n fois. Dans l'étude de sensibilité de la MCC, 18 résidus ont été générés. $3^{18} \approx 3.9 \cdot 10^8$. Un parcours systématique de l'espace des solutions est donc écarté.

Analyse 2 : en considérant qu'une limite supérieure de la sensibilité de D_j à f_j doit être écartée, la maximisation du critère de sensibilité aux grands défauts impose que $I_0^j \subset I_{0,\max}^j$ ((V.18) et (V.11)). Dans ces conditions, choisir $I_0^j = I_{0,\max}^j$ est optimal pour satisfaire le critère de localisation (V.16) même si ce n'est pas nécessairement le seul choix possible. En effet,

$$\min_{i \in I_{0,\max}^j} (\alpha_{ik}) \leq \min_{i \in I_0^j \subset I_{0,\max}^j} (\alpha_{ik}).$$

Analyse 3 : de manière similaire, comme $\max_{i \in I_{1,\max}^j} (\alpha_{ik}) \geq \max_{i \in I_1^j \subset I_{1,\max}^j} (\alpha_{ik})$, choisir $I_1^j = I_{1,\max}^j$ est optimal pour satisfaire le critère de localisation (V.16) même si ce n'est pas nécessairement le seul choix possible. Néanmoins, ce choix conduit à la plus mauvaise sensibilité aux petits défauts f_j (V.17). Il est intéressant de noter que l'on retrouve un compromis entre sensibilité et robustesse : sensibilité au défaut f_j (V.17) et robustesse de la localisation (voir remarque de la partie V.2.4.2.1). Comme nous le verrons dans l'application à la MCC, lorsque $I_1^j = I_{1,\max}^j$, la sensibilité de D_j au défaut f_j peut rendre D_j presque inutilisable en pratique ; nous cherchons donc à rééquilibrer le compromis entre sensibilité et robustesse au profit de la sensibilité à f_j .

Pour cela, une solution consiste à construire progressivement I_1^j en choisissant les résidus de $I_{1,\max}^j$ du plus sensible à f_j au moins sensible et en arrêtant la construction de I_1^j dès que le critère de localisation est satisfait (V.16).

Corollaire de Analyse 3 : si le critère de localisation (V.16) n'est pas satisfait lorsque $I_0^j = I_{0,\max}^j$ et $I_1^j = I_{1,\max}^j$, alors aucune fonction de décision de la forme (V.9) ne permet de localiser f_j au sens du critère (V.16). Dans ce cas, deux approches sont envisageables : un retour sur la génération de résidus, en vue de modifier les sensibilités des résidus existants ou en vue d'en créer de nouveaux, est une alternative. Une deuxième alternative consisterait à élargir l'ensemble des fonctions de décision possibles (par exemple, en généralisant (V.9) avec l'utilisation d'opérateurs du type OU logique). Toute la démarche qui a été développée sur l'évaluation des fonctions de décision devrait alors être reprise et adaptée en conséquence.

V.2.5.2. Algorithme de construction d'une fonction de décision

Dans le cadre que nous nous sommes fixé, l'analyse du problème d'optimisation des fonctions de décision nous conduit à proposer l'algorithme de la Table V.6.

<i>Algorithme</i>	<i>Commentaires</i>
$I_0^j = I_{0,\max}^j$ $I_1^j = \emptyset$ Tant que $J_{LOC(j)}=0$ et $I_1^j \neq I_{1,\max}^j$ $I_1^j = I_1^j \cup \left\{ i \text{ telque } \alpha_{ij} = \min_{i \in I_{1,\max}^j - I_1^j} \alpha_{ij} \right\}$ Fin Tant que Si $J_{LOC(j)}=0$ Alors Retour à la génération de résidus Etude de sensibilité => nouvel α Reprendre l'algorithme au début Sinon Pour $i=1 \dots n_r$ Si $i \in I_0^j \cup I_1^j$ et si retirer i de I_0^j ou de I_1^j ne fait pas varier $J_{LOC(j)}$, $J_{SPD(j)}$ et $J_{SGD(j)}$, Alors Retirer i de I_0^j si $i \in I_0^j$ Retirer i de I_1^j si $i \in I_1^j$ Fin Si Fin Pour Fin Si $D_j = D(I_0^j, I_1^j)$	Initialisation de I_0^j (voir analyse 2) Initialisation en vue d'une construction itérative de I_1^j Tant que le critère de localisation (V.16) n'est pas satisfait et qu'il reste des résidus à sélectionner dans $I_{1,\max}^j$ Sélection du résidu de $I_{1,\max}^j$ le plus sensible à f_j et non sélectionné précédemment. Si la localisation de f_j (V.16) n'a pu être remplie D'autres résidus doivent être générés Sinon on élimine les résidus sélectionnés inutilement (décision avec un nombre minimal de résidus) Pour chaque résidu Construire la fonction de décision D_j avec I_0^j et I_1^j

Table V.6 : Algorithme pour construire la fonction de décision D_j la plus sensible possible à f_j , satisfaisant le critère de localisation et utilisant un nombre minimal de résidus

V.2.5.3. Application à la machine à courant continu (MCC)

L'application à la MCC est basée sur l'étude de sensibilité menée au chapitre III (Table III.3). Partant de cette table, les fonctions de décision faisant intervenir tous les résidus ($D_j = D(I_{0,\max}^j, I_{1,\max}^j)$) comme dans (V.2) satisfont le critère de localisation mais leur sensibilité au défaut testé est plutôt mauvaise (Table V.7).

Défaut	Sensibilité
$f_1 = f_\beta$	$f_\beta > 41.9\%$
$f_2 = f_I$	$f_I > 73.5\%$
$f_3 = f_\Omega$	$f_\Omega > 25.1\%$
$f_4 = f_U$	$f_U > 32.1\%$

Table V.7 : Sensibilité des fonctions de décision faisant intervenir tous les résidus

Défaut	Fonction de décision	Sensibilité
$f_1 = f_\beta$	$D_1 = T_4$	$f_\beta > 1.3\%$
$f_2 = f_I$	$D_2 = \overline{T_{13}} \wedge \overline{T_{15}} \wedge T_{16}$	$f_I > 13.6\%$
$f_3 = f_\Omega$	$D_3 = \overline{T_{16}} \wedge T_{17}$	$f_\Omega > 0.3\%$
$f_4 = f_U$	$D_4 = T_{13} \wedge \overline{T_{14}}$	$f_U > 15.9\%$

Table V.8 : Application de l'algorithme de construction des fonctions de décision

La mauvaise sensibilité de certains résidus de la Table III.3 se répercute directement sur les résultats de la Table V.7 : en effet, la sensibilité de la décision faisant intervenir tous les résidus est celle du résidu le moins sensible. L'application de l'algorithme conduit à des fonctions de décision dont la sensibilité aux défauts est correcte (Table V.8) tout en satisfaisant le critère de localisation.

On peut remarquer que les résidus 13 à 17 sont ceux retenus pour calculer les fonctions logiques D_2 à D_4 . Ces résidus sont obtenus dans le cadre des observateurs à entrée inconnue (OEI) ou par une approche fréquentielle (AF). Ces méthodes de génération de résidus procurent en effet plus de degrés de liberté pour régler les sensibilités des résidus que la synthèse des résidus globaux (RG) ou locaux (RL).

La sensibilité de la fonction de décision associée à f_β est celle du résidu le plus sensible à f_β . Une situation similaire peut être constatée pour f_I et f_Ω . Cependant, la sensibilité de D_4 est $f_U > 15.9\%$, alors que le résidu le plus sensible à f_U est r_{18} avec $f_U > 3.8\%$ (Table III.3). Contrairement aux autres défauts, plusieurs itérations dans la boucle « tant que » de l'algorithme (Table V.6) sont nécessaires avant de satisfaire le critère de localisation (V.16). Le résidu le plus sensible à f_U qui doit être inclus dans I_1^4 pour satisfaire (V.16) est r_{13} . Par conséquent, la sensibilité de D_4 à f_U est celle de r_{13} ($f_U > 15.9\%$).

Remarque : pour réduire à cinq le nombre de résidus impliqués dans les fonctions de décision de la Table V.8, il est possible de choisir $D_1 = T_{15}$. La sensibilité de D_1 à f_β est alors : $f_\beta > 6\%$.

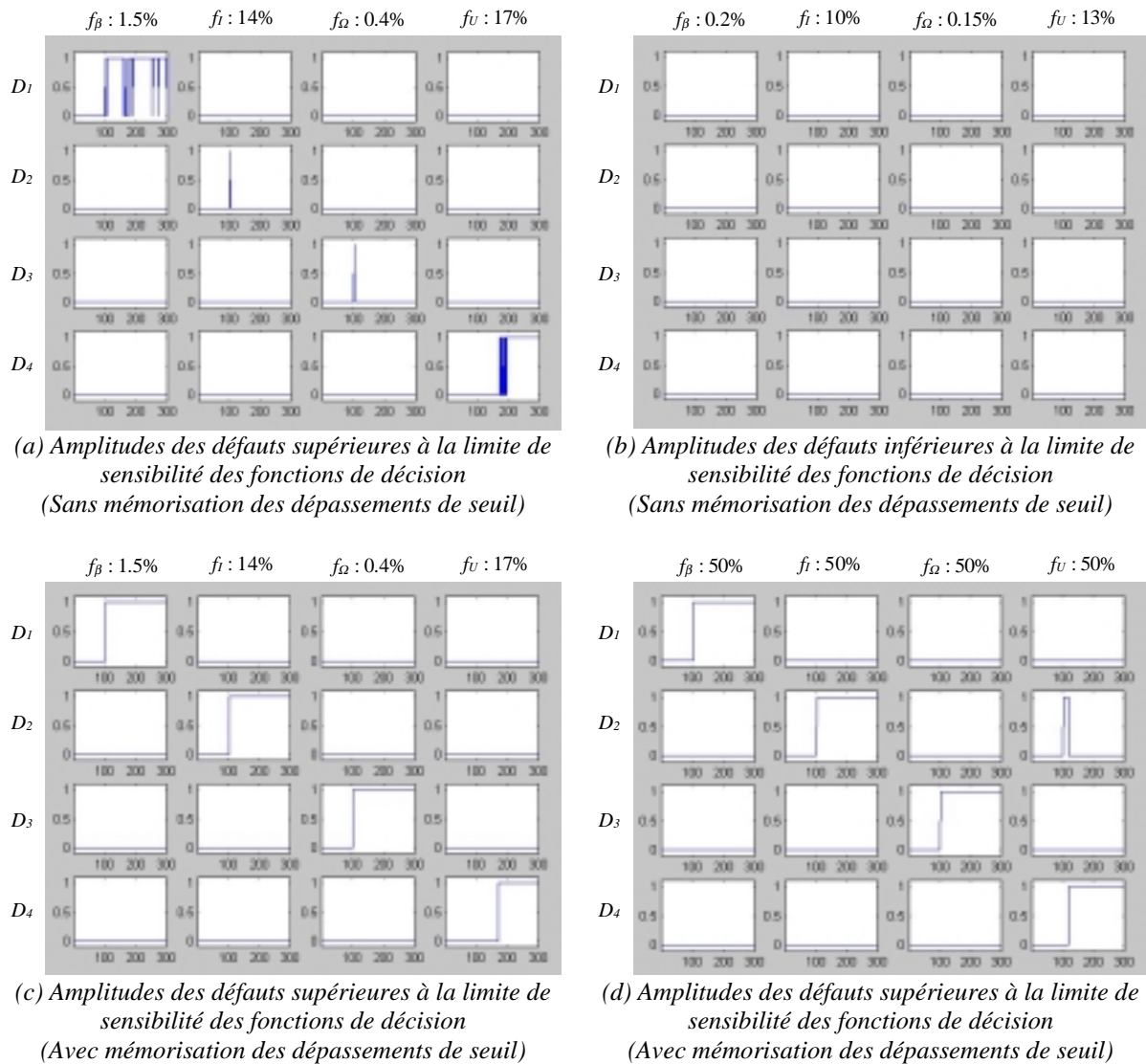


Figure V.3 : Simulation des fonctions de décision obtenues

Les simulations des fonctions de décision (Figure V.3a et Figure V.3b) confirment les limites de sensibilité de la Table V.8. Les instabilités des décisions D_1 et D_4 sur la Figure V.3a sont liées au bruit de mesure ; une simulation sans bruit nous a permis de vérifier que les valeurs transitoirement non nulles de D_2 et D_3 ne sont pas liées à la présence du bruit de mesure. Le fait que D_2 et D_3 prennent la valeur vraie pendant une courte durée s'explique par le dépassement de seuil nécessairement bref de résidus en limite de sensibilité et ayant un comportement transitoire (Figure V.4). Une solution pour pallier en partie ce problème consiste à mémoriser les dépassements de seuils (Figure V.3c).

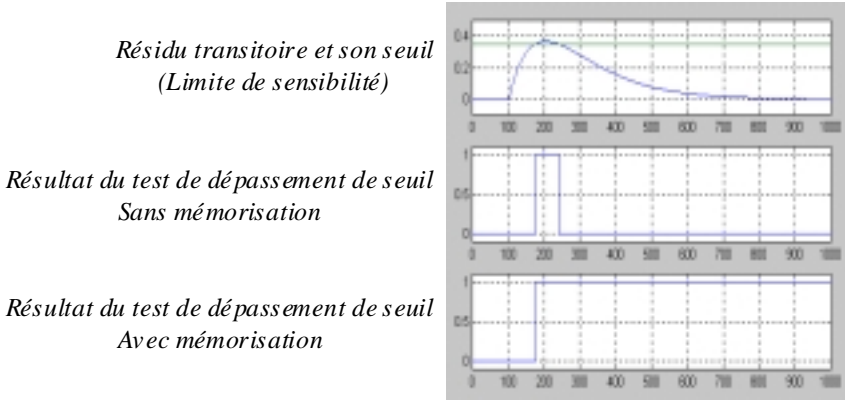


Figure V.4 : Tests de dépassement de seuil sans et avec mémorisation

Même si les dépassements de seuil sont mémorisés, comme les aspects temporels ne sont pas pris en compte, l'ordre des dépassements de seuil peut être à l'origine de fausses décisions (Figure V.3d). Afin de vérifier cette interprétation, les résultats des tests de dépassement de seuil sont indiqués sur la Figure V.5. On peut vérifier que les dépassements de seuil correspondent parfaitement aux résultats attendus compte tenu de l'étude de sensibilité (Table III.3). La Figure V.5 permet de constater que l'origine de la fausse décision sur la Figure V.3d vient du retard avec lequel r_{13} dépasse son seuil par rapport à r_{16} , en réponse f_U .

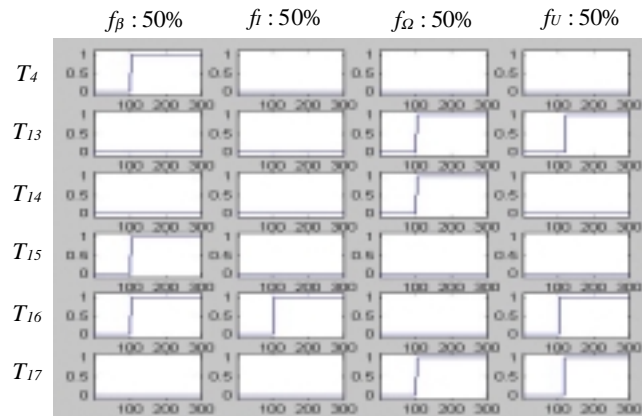


Figure V.5 : Tests de dépassement de seuil avec mémorisation pour des défauts de 50%

Remarque : la mémorisation des dépassements de seuil correspond en fait à une décision dynamique et non plus à une décision statique au sens strict.

La décision proposée présente certaines limites liées à la non prise en compte des aspects temporels. Notons que l'information fournie par l'étude de sensibilité réalisée n'est pas suffisante pour une telle prise en compte. Une autre limite vient de l'hypothèse sur la forme des défauts (échelon) et, plus généralement, de la manière dont sont prises en compte les incertitudes liées à l'interprétation des dépassements de seuil (Figure V.2).

V.2.6. Optimisation par un algorithme génétique

V.2.6.1. Objectifs et choix d'un algorithme

Même si un algorithme déduit de la formulation du problème peut donner des résultats intéressants, cette approche est difficilement transposable : modifier une seule hypothèse impose de refaire toute la démarche d'analyse avant d'espérer aboutir à un algorithme dédié. Cette démarche sera d'autant plus délicate à mettre en œuvre que la structure des fonctions de décision sera complexe et que les hypothèses de travail seront limitées. Cependant, la formulation du problème adoptée pour la synthèse des fonctions de décision se prête naturellement à l'optimisation. Il suffit pour cela d'avoir défini la forme des fonctions de décision et d'être capable d'évaluer leurs performances (critères de localisation, de sensibilité,...).

Les algorithmes destinés à résoudre (éventuellement de manière sous-optimale) un problème fortement combinatoire peuvent être classés en trois catégories [Rebreymend, 1999] :

- Les algorithmes exacts,
- Les heuristiques qui sont des algorithmes spécifiques à un problème donnant un résultat approché,
- Les méta-heuristiques qui sont des méthodes générales.

Comme nous cherchons un algorithme peu dépendant du problème traité, nous nous intéressons aux méta-heuristiques. Parmi elles, la descente du gradient est une heuristique donnant comme résultat une solution qui est un minimum (ou un maximum) local. Une approche locale paraissant difficilement applicable à l'optimisation des fonctions de décision, un algorithme génétique a été appliqué : ces algorithmes cherchent en effet une solution optimale (ou sous-optimale) parmi un vaste espace de solutions potentielles [Goldberg, 1989].

V.2.6.2. Principales caractéristiques d'un algorithme génétique

Les principales caractéristiques d'un algorithme génétique sont maintenant rappelées. Les solutions (individus) sont codées sous forme de chaînes (chromosomes), définies sur un alphabet fini (un codage binaire est souvent utilisé, par exemple : 10010111). Un critère permettant d'évaluer une chaîne quelconque doit être défini : il traduit l'objectif de l'optimisation. Une population (ensemble d'individus) de départ est tout d'abord générée aléatoirement (première génération). Chaque itération (ou génération) de l'algorithme produit une population de solutions possibles basées sur la population de la génération précédente. Trois opérateurs sont appliqués avec une certaine probabilité aux individus d'une génération pour passer à la génération suivante :

- **Sélection** : les solutions conduisant à une bonne valeur du critère d'optimisation survivent d'une génération à l'autre avec une meilleure probabilité que les autres.
- **Croisement** : deux solutions choisies par la sélection sont combinées par croisement (crossover) : une position dans les chaînes (chromosomes) est choisie au hasard et les parties des deux chaînes en amont de cette position sont échangées pour produire deux nouvelles chaînes.
- **Mutation** : les chaînes codant les solutions peuvent être soumises à des changements spontanés (avec une probabilité assez faible) afin de produire de nouvelles chaînes situées dans différentes parties de l'espace des solutions (diversité génétique).

L'efficacité d'un algorithme génétique par rapport à son besoin en calculs est parfois moyenne. Toutefois, l'optimisation des fonctions de décision intervenant au stade de la conception, le temps de calcul n'est pas nécessairement un problème critique (tout dépend de la complexité du problème à résoudre). En outre, des heuristiques peuvent être adoptées : par exemple, conserver systématiquement les n meilleures solutions d'une génération dans la suivante, ou retirer les n plus mauvaises. Notre choix a consisté à conserver systématiquement la meilleure solution.

V.2.6.3. Application à l'optimisation des fonctions de décision

V.2.6.3.1. Codage des fonctions de décision

Les fonctions de décision étudiées sont définies par deux ensembles disjoints de résidus, I_0 et I_1 (V.9). Pour définir I_0 et I_1 , il suffit donc d'associer à chaque résidu un code (gène) précisant si ce dernier appartient à I_0 (code : 0), à I_1 (code : 1) ou au complément de $I_0 \cup I_1$ (code : 2). Ainsi, le code 2 associé à un résidu signifie que ce dernier n'intervient pas dans la fonction de décision correspondante (V.9). La chaîne (le chromosome) $Ch(I_0, I_1)$, décrivant de manière unique la fonction $D(I_0, I_1)$, résulte des codes (gènes) associés à chacun des résidus (V.19).

$$Ch(I_0, I_1) = [\dots \ g_i \ \dots], \quad i = 1 \dots n_r, \quad \begin{cases} g_i = 0 & \text{si } i \in I_0 \\ g_i = 1 & \text{si } i \in I_1 \\ g_i = 2 & \text{sinon} \end{cases} \quad (\text{V.19})$$

Exemple : en supposant que $n_r=5$, le chromosome $[1 \ 2 \ 0 \ 2 \ 0]$ correspond à $I_0=\{3;5\}$ et $I_1=\{1\}$; ce chromosome code donc la fonction de décision $D(I_0, I_1) = T_1 \wedge \overline{T_3} \wedge \overline{T_5}$.

V.2.6.3.2. Choix du critère à optimiser

L'optimisation d'une fonction de décision D_j , destinée à localiser f_j , est une optimisation multi-critères. On cherche en effet à maximiser simultanément le critère de localisation, $J_{LOC(j)}$ (V.16), le critère de sensibilité aux petits défauts, $J_{SPD(j)}$ (V.17), et le critère de sensibilité aux grands défauts, $J_{SGD(j)}$ (V.18). Le critère global $J_{Global(j)}$, retenu pour l'application de l'algorithme génétique, résulte de la pondération des trois critères élémentaires :

$$J_{Global(j)} = w_{LOC} \cdot J_{LOC(j)} + w_{SPD} \cdot J_{SPD(j)} + w_{SGD} \cdot J_{SGD(j)} \quad (\text{V.20})$$

V.2.6.3.3. Résultats

L'algorithme génétique est appliqué à l'étude de sensibilité effectuée au chapitre III (Table III.3.). Une comparaison avec les résultats de l'algorithme de la partie V.2.5 est donc possible. Pour obtenir les résultats de la Table V.9, les pondérations et les probabilités de (V.21) ont été utilisées.

$$\begin{aligned} [w_{LOC} \ w_{SPD} \ w_{SGD}] &= [30 \ 30 \ 100] \\ \text{Probabilité de Croisement} &= 0,6 \\ \text{Probabilité de Mutation} &= 0,06 \end{aligned} \quad (\text{V.21})$$

Les poids relatifs des trois critères élémentaires ont été obtenus par essais successifs (en analysant les résultats de chaque optimisation). Les probabilités de croisement et de mutation correspondent à des ordres de grandeur couramment utilisés dans les applications des algorithmes génétiques. L'algorithme est stoppé lorsqu'un certain nombre de générations a été atteint (Table V.9). Le nombre d'individus est constant d'une génération à l'autre. La même population de départ est utilisée pour générer chacune des quatre fonctions de décision (modifier

le paramètre associé au générateur de nombres aléatoires¹ revient, entre autres, à modifier la population de départ). Une fonction de décision associée au défaut f_j résulte de l'optimisation du critère $J_{Global(j)}$. La Table V.9 présente les résultats de l'optimisation des fonctions de décision.

Arrêt après n_g générations, $n_g =$	Nombre d'individus	Générateur de nombres aléatoires	Critère $J_{global(j)}$ $j=$	I_0, I_1 de la "meilleure" fonction de décision dans la population de la dernière génération	Sensibilité de la "meilleure" décision
15	40	1	1	$I_0=\{5,7,8,11,13,17,18\}$ $I_1=\{4,10,15\}$	$6\% < f_\beta$
			2	$I_0=\{4,8\}$ $I_1=\{6,12,14,18\}$	$68.3\% < f_i$
			3	$I_0=\{4,16\}$ $I_1=\{5,11,13,14,18\}$	$16.1\% < f_\Omega$
			4	$I_0=\{1,2,4,5,6,7,10,11,14,15,16,17\}$ $I_1=\{9,12,18\}$	$5.8\% < f_U < 6.7\%$
15	40	2	1	$I_0=\{6,7,8,14,17,18\}$ $I_1=\{4\}$	$1.3\% < f_\beta < 41.9\%$
			2	$I_0=\{1,5,6,8,10\}$ $I_1=\{16\}$	$13.6\% < f_i < 39.8\%$
			3	$I_0=\{2,5,6,7,8,9,12\}$ $I_1=\{17\}$	$0.3\% < f_\Omega$
			4	$I_0=\{8,14\}$ $I_1=\{5,7,9,13,18\}$	$15.9\% < f_U$
100	40	1	1	$I_0=\{3,7,8,9,11,17,18\}$ $I_1=\{4\}$	$1.3\% < f_\beta$
			2	$I_0=\{3,4,8,13\}$ $I_1=\{16\}$	$13.6\% < f_i$
			3	$I_0=\{4,16\}$ $I_1=\{11\}$	$0.8\% < f_\Omega$
			4	$I_0=\{4,8,14\}$ $I_1=\{1,9,12,13,18\}$	$15.9\% < f_U$
100	40	2	1	$I_0=\{7,8,14,17,18\}$ $I_1=\{4\}$	$1.3\% < f_\beta$
			2	$I_0=\{10,13,17\}$ $I_1=\{9,16\}$	$21.4\% < f_i$
			3	$I_0=\{1,2,7,10\}$ $I_1=\{17\}$	$0.3\% < f_\Omega$
			4	$I_0=\{8,14\}$ $I_1=\{5,7,9,13,18\}$	$15.9\% < f_U$

Table V.9 : Résultats de l'optimisation des fonctions de décision par un algorithme génétique

Les sensibilités des décisions obtenues dans la Table V.9 sont à comparer à celles obtenues avec l'algorithme déduit de la formulation du problème (Table V.8). Au bout de 15 générations, les résultats de la Table V.9 sont encore assez contrastés : certaines décisions sont bonnes alors que d'autres présentent une mauvaise sensibilité (aux petits ou aux grands défauts, suivant les cas). La population initiale influence fortement les fonctions de décision obtenues. Au bout de 100 générations, la population initiale influence beaucoup moins les résultats : les critères de localisation et de sensibilité aux grands défauts sont totalement satisfaits et la sensibilité aux petits défauts, si elle n'est pas optimale (Table V.8), n'en est pas moins très bonne.

Remarque : une pénalité liée au nombre de résidus sélectionnés pourrait être ajoutée au critère global $J_{Global(j)}$ afin de converger vers des fonctions de décision faisant intervenir le moins de résidus possible.

Etant donné que la population d'une génération est constituée de 40 individus, 4000 fonctions de décision sont évaluées au bout de 100 générations. Le nombre 4000 est à comparer aux $3^{18} \approx 3.9 \cdot 10^8$ éléments de l'espace des solutions. La méta-heuristique génétique permet de s'affranchir au moins partiellement du caractère combinatoire lié à l'optimisation des fonctions de décision. Même si l'algorithme dédié développé dans la partie V.2.5 est beaucoup plus efficace en terme de quantité de calculs, la moindre modification dans la formulation du problème réduit à néant les efforts consentis pour son analyse ; par contre, un algorithme génétique a de fortes chances d'être à nouveau applicable pour optimiser des fonctions de décision. Dans l'exemple traité, nous disposons des résultats de l'algorithme dédié pour évaluer

¹ Si le tirage des nombres aléatoires était complètement aléatoire, les populations obtenues lors de deux exécutions successives du même programme seraient a priori différentes. Le paramètre associé au générateur de nombres aléatoires sélectionne une suite aléatoire particulière et, par conséquent, permet de retrouver la même population de départ lors de deux exécutions successives de l'algorithme génétique.

ceux de l'algorithme génétique. Cette comparaison n'est généralement pas possible, mais les solutions obtenues peuvent toujours être évaluées à l'aide du critère d'optimisation. En outre, une propriété remarquable des algorithmes génétiques est une certaine robustesse par rapport aux différentes instances du problème traité [Rebreyend, 1999] : quel que soit le défaut à localiser, le même algorithme génétique (avec les mêmes paramètres) converge en effet vers une « bonne » fonction de décision au sens du critère choisi.

Deux techniques (algorithme dédié et algorithme génétique) ont été utilisées dans un même but : sélectionner les résidus les plus pertinents au sens de certains critères pour construire des fonctions de décision. Par rapport à une simple table de décision booléenne, l'information disponible a été enrichie par les résultats d'une étude de sensibilité des résidus. Par contre, les tests sur les résidus sont restés élémentaires : la comparaison de chaque résidu à un seuil fournit une grandeur booléenne utilisée (ou non) par les fonctions de décision.

V.3. Ordres de grandeur absolus et relatifs

V.3.1. Augmentation de la richesse de l'information donnée par le test d'un résidu

Si la simplicité des tests sur les résidus utilisés précédemment (comparaison à un seuil) a facilité leur interprétation en terme de connaissance sur les défauts, il n'en reste pas moins que cette approche présente deux limites importantes : le résultat de la comparaison à un seuil est en tout ou rien et les ordres de grandeur relatifs des résidus ne peuvent être testés directement.

V.3.1.1. Quelques éléments de logique floue

Lorsque le résultat de la comparaison à un seuil est booléen (tout ou rien), la décision peut être complètement différente suivant que le résidu est juste au-dessous ou juste au-dessus du seuil qui a été fixé. L'utilisation de seuils « nets » favorise l'instabilité des décisions. Par comparaison, l'utilisation de la logique floue permet de construire une décision progressive (Figure V.6).

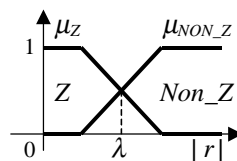


Figure V.6 : Un exemple de fuzzification du résidu r

L'exemple de la fuzzification d'un résidu donné par la Figure V.6 consiste à calculer l'appartenance de $|r|$ aux ensembles flous désignés par leur terme linguistique : Z (pour « Zéro ») et Non_Z (pour « Non Zéro »). En raisonnant sur de tels termes linguistiques, facilement intelligibles, un raisonnement logique graduel peut être construit. Pour cela, il existe plusieurs types d'opérateurs qui peuvent être classés en trois catégories (Figure V.7).

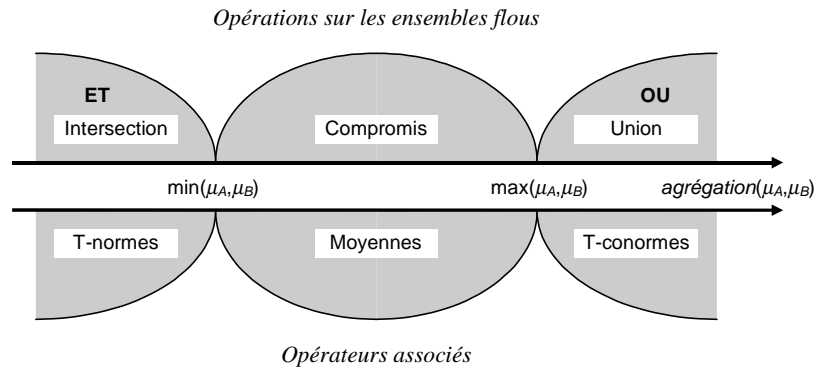


Figure V.7 : Classement des opérateurs flous élémentaires

Etant donné deux ensembles flous A et B , un troisième peut être obtenu par agrégation des deux premiers. Une agrégation conjonctive représente une intersection (ET logique) ; une agrégation disjonctive représente une union (OU logique) ; une agrégation du type « compromis » est un intermédiaire entre l'intersection et l'union. Il existe toute une famille d'opérateurs susceptibles d'être utilisés pour décrire une intersection (resp. une union) : les T-normes (resp. les T-conormes) [Dubois, 1985]. La fonction min est la plus grande des T-normes et la fonction max , la plus petite des T-conormes (Figure V.7). Les moyennes servent à décrire une agrégation du type « Compromis ».

	<i>T-norme</i>	<i>T-conorme</i>
<i>Zadeh (z)</i>	$\mu_A \overset{z}{\wedge} \mu_B = \min(\mu_A, \mu_B)$	$\mu_A \overset{z}{\vee} \mu_B = \max(\mu_A, \mu_B)$
<i>Lukasiewicz (l)</i>	$\mu_A \overset{l}{\wedge} \mu_B = \max(0, \mu_A + \mu_B - 1)$	$\mu_A \overset{l}{\vee} \mu_B = \min(1, \mu_A + \mu_B)$
<i>Probabiliste (p)</i>	$\mu_A \overset{p}{\wedge} \mu_B = \mu_A \cdot \mu_B$	$\mu_A \overset{p}{\vee} \mu_B = \mu_A + \mu_B - \mu_A \cdot \mu_B$

<i>Négation floue</i>	<i>Moyenne</i>
$Non(\mu_A) = 1 - \mu_A$	$Moy(\mu_A, \mu_B) = (\mu_A + \mu_B) / 2$

Table V.10 : Quelques opérateurs flous usuels

Afin de simplifier les notations, l'opérateur \wedge (resp. \vee) servira à désigner une T-norme (resp. T-conorme) quelconque. \wedge (resp. \vee) permet donc de modéliser une intersection (resp. union). Ces notations sont homogènes à celles déjà utilisées où \wedge (resp. \vee) désignait un ET (resp. OU) logique. Dès lors, la transcription des fonctions de décision de la partie V.2, dans un contexte de logique floue, est immédiate :

$$\text{Exemple: } D_1 = \bar{T}_1 \wedge T_2 \wedge T_3 \text{ devient } \mu_{D(f_1)} = \mu_Z(r_1) \wedge \mu_{Non_Z}(r_2) \wedge \mu_{Non_Z}(r_3) \tag{V.22}$$

$\mu_{D(f_1)}$ désigne la valeur d'appartenance à l'ensemble flou $D(f_1)$ (« Décision défaut f_1 »). Cette valeur d'appartenance décrit la cohérence du comportement des résidus avec le comportement attendu lorsque le défaut f_1 se produit. La table de décision décrivant la connaissance à partir de laquelle $\mu_{D(f_1)}$ peut être construite est toujours booléenne : $Sg_{ij} \in \{Z; Non_Z\}$ (Table V.2).

V.3.1.2. Ordres de grandeur absolus qualitatifs : les signes

Le nombre d'ensembles flous utilisés pour décrire un résidu donné n'est pas limité : la granularité de la description floue d'un résidu peut être augmentée en tenant compte, par exemple, de l'information sur le signe de ce résidu (Figure V.8).

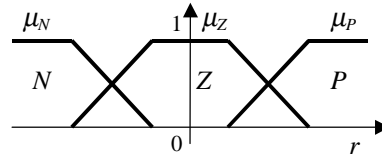


Figure V.8 : Fuzzification d'un résidu avec prise en compte du signe

Il devient désormais possible de décrire avec plus de précision le comportement des résidus en utilisant une table de décision dont les éléments Sg_{ij} sont à valeur dans $\{N,Z,P\}$ (N pour « Négatif », Z pour « Zéro » et P pour « Positif »). Ainsi, la prise en compte des signes pour décrire le comportement des résidus globaux de la MCC (Figure II.10) conduit à transformer la table de décision booléenne Table II.8a en la Table V.11.

	$f_{\beta} > 0$	$f_I > 0$	$f_{\Omega} > 0$	$f_U > 0$
$r_{\beta_m}^*$	P	N	Z	N
$r_{I_m}^*$	P	P	Z	P
$r_{\Omega_m}^*$	P	N	P	P

(a) défauts positifs

	$f_{\beta} < 0$	$f_I < 0$	$f_{\Omega} < 0$	$f_U < 0$
$r_{\beta_m}^*$	N	P	Z	P
$r_{I_m}^*$	N	N	Z	N
$r_{\Omega_m}^*$	N	P	N	N

(b) défauts négatifs

Table V.11 : Tables de décision avec prise en compte des signes

Sur l'exemple de la Table V.11, alors que la table de décision booléenne (zone grisée) n'est pas localisante, la prise en compte des signes permet de localiser les défauts considérés.

Une description plus précise du comportement des résidus a permis d'accroître la richesse descriptive de la table de décision, rendant ainsi possible la localisation de certains défauts alors qu'elle ne l'était pas auparavant.

L'exemple de la Table V.11 illustre également le compromis entre précision et validité d'un modèle : la Table V.11a (resp. Table V.11b) n'est valide que pour des défauts positifs¹ (resp. négatifs). La Table V.11 ne permet donc pas de décrire le comportement des résidus en réponse à des défauts dont le signe varie ; cette contrainte n'existait pas avec une table de décision booléenne. Même si le gain en précision a permis d'améliorer les capacités de localisation, il s'est également traduit par une restriction sur le domaine de validité de la décision. La décision floue destinée à localiser un défaut, par exemple f_I , est construite sur le modèle suivant :

$$\mu_{D(f_I > 0)} = \mu_N(r_{\beta_m}^*) \wedge \mu_P(r_{I_m}^*) \wedge \mu_N(r_{\Omega_m}^*) \tag{V.23}$$

$$\mu_{D(f_I < 0)} = \mu_P(r_{\beta_m}^*) \wedge \mu_N(r_{I_m}^*) \wedge \mu_P(r_{\Omega_m}^*) \tag{V.24}$$

¹ Le signe d'un défaut est défini comme le signe de la variable d'entrée qui modélise ce défaut (en supposant qu'une telle variable ait été préalablement définie).

$$\mu_{D(f_j)} = \mu_{D(f_j>0)} \vee \mu_{D(f_j<0)} \tag{V.25}$$

(V.25) est une agrégation disjonctive représentant un OU logique entre deux hypothèses associées à un défaut : défaut positif OU défaut négatif. La prise en compte de telles hypothèses nous amène à généraliser le modèle « table de décision ».

V.3.1.3. Généralisation du modèle « table de décision » et décision floue

Supposons que n_h hypothèses aient été faites pour chaque classe de défauts initiale ($f_j, j=1 \dots n_f$). Par exemple : défaut positif, défaut négatif, défaut en échelon, défaut en rampe, ... Pour chaque hypothèse $h \in 1 \dots n_h$, une table de décision Sg^h permet de décrire le comportement des résidus ($r_i, i=1 \dots n_r$) en réponse à chacun des défauts f_j^h où f_j^h désigne le défaut f_j sous l'hypothèse h .

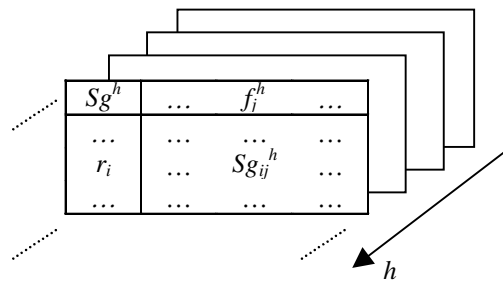


Table V.12 : Table de décision généralisée

Soit $L_i^l, l=1 \dots n_l$, les termes linguistiques désignant les ensembles flous servant à décrire le comportement du résidu r_i . Par exemple, la prise en compte des signes a été obtenue dans la partie V.3.1.2 en choisissant $\{L_i^l, l=1 \dots n_l\} = \{N, Z, P\}$. Les valeurs d'appartenance aux ensembles flous désignés par L_i^l peuvent résulter directement de la fuzzification de r_i ou de traitements plus complexes (voir Chapitre VI). L'objectif d'une table de décision est de décrire le comportement des résidus en réponse aux défauts ; or, L_i^l désigne un comportement particulier de r_i qui peut être testé en calculant l'appartenance à l'ensemble flou correspondant. On choisit donc :

$$\forall i, \forall j, \forall h, Sg_{ij}^h \in \{L_i^l, l=1 \dots n_l\} \tag{V.26}$$

La décision floue déduite du modèle de la Table V.12 est alors construite comme suit :

$$\mu_{D(f_j^h)} = \bigwedge_{i=1 \dots n_r} \mu_{Sg_{ij}^h}(r_i) \tag{V.27}$$

$$\mu_{D(f_j)} = \bigvee_{h=1 \dots n_h} \mu_{D(f_j^h)} \tag{V.28}$$

Nous avons ramené le calcul de la décision à une série d'agrégations floues directement déduites d'une table de décision. Des relations floues peuvent également être utilisées pour calculer la décision. Le lien entre deux approches de ce type est explicité dans [Evsukoff, 1999].

V.3.1.4. Ordres de grandeur relatifs numériques : les directions

Les ensembles flous élémentaires servant à décrire les comportements des résidus n'ont jusqu'ici porté que sur un seul résidu. Ainsi, les fuzzifications de la Figure V.6 et de la Figure V.8 définissent des ensembles flous décrivant l'ordre de grandeur **absolu** d'un résidu. Ce n'est qu'a

posteriori que les comportements des résidus sont reliés entre eux (V.27). Il est parfois très utile de tester directement l'ordre de grandeur **relatif** d'un (sous-)ensemble de résidus.

Par exemple, deux résidus, r_1 et r_2 , sont supposés sensibles à deux défauts, f_1 et f_2 . Si r_1 et r_2 sont tout les deux positifs lorsque $f_1 > 0$ ou lorsque $f_2 > 0$, l'information sur les signes ne permet pas de localiser f_1 et f_2 . Un supplément d'information peut provenir de l'ordre de grandeur relatif de r_1 et r_2 : supposons que r_1 soit trois fois plus grand que r_2 en réponse à f_1 alors que r_2 est deux fois plus grand que r_1 en réponse à f_2 . Cela signifie que le vecteur $[r_1 \ r_2]^T$ est dans la direction du vecteur $[3 \ 1]^T$ en réponse à f_1 et dans la direction du vecteur $[1 \ 2]^T$ en réponse à f_2 .

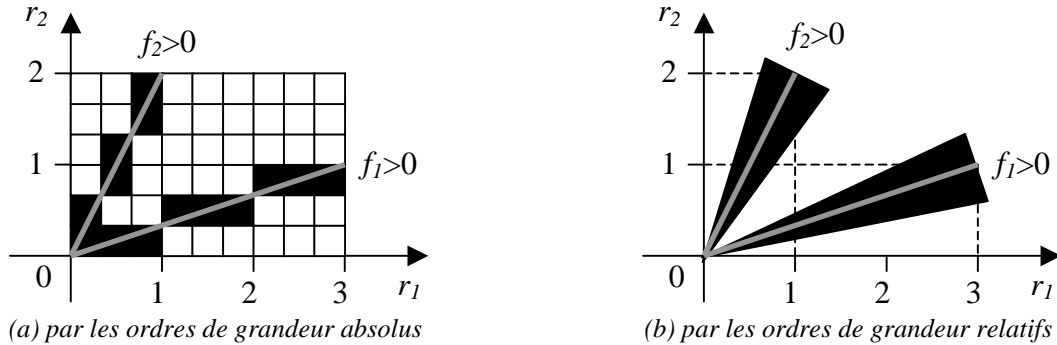


Figure V.9 : Modélisation d'une direction

Modéliser avec précision de tels comportements en testant les ordres de grandeur absolus des résidus accroît considérablement la complexité de la décision. Sur la Figure V.9a, r_1 est décrit par 9 ensembles flous et r_2 par 6. Dans ce contexte, les ordres de grandeur absolus ne sont pas des critères pertinents. Par contre, un secteur angulaire modélise simplement un ensemble de comportements relatifs de r_1 et r_2 autour d'une certaine direction (Figure V.9b) : ainsi, l'angle entre $[3 \ 1]^T$ et $[r_1 \ r_2]^T$ est proche de zéro si r_1 est à peu près trois fois plus grand que r_2 .

Fuzzifier l'angle entre une direction donnée et un vecteur de résidus revient à définir un ensemble flou dont la valeur d'appartenance permet d'évaluer les ordres de grandeur relatifs des résidus considérés.

L'angle entre une direction δ et un vecteur de résidus r est noté $\delta \angle r$. L'appartenance de r à un ensemble flou représenté par un secteur angulaire autour de δ est donc calculée par (V.29), où μ_z désigne une fonction d'appartenance centrée sur zéro.

$$\mu_\delta(r) = \mu_z(\delta \angle r) \quad \text{avec} \quad \delta \angle r = \arccos\left(\frac{\delta^T \cdot r}{\|\delta\| \|r\|}\right) \quad (\text{V.29})$$

$\mu_\delta(r)$ représente l'appartenance de r à une direction **orientée**. L'appartenance de r à une direction **non orientée** est donnée par $\mu_\delta(r) \vee \mu_{-\delta}(r)$: direction sens positif OU direction sens négatif. Afin d'éviter toute fausse décision lorsque les résidus sont proches de zéro (comportement normal), $\mu_\delta(r)$ est systématiquement combiné avec un indicateur de détection. Par exemple : $\mu_\delta(r) \wedge \mu_{\text{Non}_Z}(|r|)$ exclut un disque centré sur l'origine.

Remarque 1 : dans le cas limite où r et δ sont scalaires, on retrouve une situation correspondant exactement à la prise en compte des signes lorsque la direction est orientée (Figure V.8) :

$$\mu_N(r) = \mu_\delta(r) \wedge \mu_{Non_z}(|r|), \quad \forall \delta < 0 \quad (\text{V.30})$$

$$\mu_P(r) = \mu_\delta(r) \wedge \mu_{Non_z}(|r|), \quad \forall \delta > 0 \quad (\text{V.31})$$

On retrouve une situation identique à celle de la Figure V.6 lorsque la direction est non orientée :

$$\mu_{Non_z}(|r|) = \underbrace{(\mu_\delta(r) \vee \mu_{-\delta}(r))}_{=1} \wedge \mu_{Non_z}(|r|), \quad \forall \delta \neq 0 \quad (\text{V.32})$$

Remarque 2 : l'angle par rapport à une direction n'est pas la seule alternative pour décrire des ordres de grandeur relatifs. Une autre approche couramment utilisée consiste à fuzzifier le rapport entre deux grandeurs. Par exemple : $\mu_{Petit}(b/a)$, $\mu_{Grand}(c/a)$,... Nous n'avons pas retenu ce point de vue qui impose de choisir une référence.

V.3.2. Propriétés directionnelles des résidus

L'appartenance d'un résidu (vectoriel) à une direction fournit une information sur les ordres de grandeur relatifs des résidus. Le choix des directions à tester nous amène donc à étudier les propriétés directionnelles des résidus. Rappelons que certaines méthodes de génération de résidus sont spécifiquement dédiées à la synthèse de résidus directionnels [Chen, 1995].

V.3.2.1. Cas général, lien avec les tests de modèles basés sur le gradient

Afin d'éviter toute confusion entre l'opérateur δ introduit au chapitre III et les directions également notées δ , notre étude est réalisée avec un modèle discret ; l'opérateur dynamique discret est noté z . (V.33) est la forme générale d'un modèle discret du système physique ou d'une partie du système physique. Ce modèle est supposé découplé des perturbations :

$$h(z, u, y, f) = 0 \quad (\text{V.33})$$

Le résidu donné par (V.34) permet de tester la cohérence des observations disponibles avec les comportements décrits par le modèle (V.33), lorsque $f=0$.

$$r = -h(z, u, y, 0) \quad (\text{V.34})$$

$$\text{soit encore: } r = h(z, u, y, f) - h(z, u, y, 0)$$

(V.34) peut décrire une large classe de générateurs de résidus. L'influence des défauts sur les résidus peut être évaluée par un développement au premier ordre.

$$r \approx \sum_j \delta_j(z, u, y, 0) \cdot f_j \quad \text{où } \delta_j(z, u, y, 0) = \left. \frac{\partial r}{\partial f_j} \right|_{f=0} = \left. \frac{\partial h(z, u, y, f)}{\partial f_j} \right|_{f=0} \quad (\text{V.35})$$

δ_j représente la direction de r en réponse à un petit défaut f_j seul. δ_j est a priori variable dans le temps et dépend de l'excitation du système. δ_j n'est rien d'autre que le gradient de r par rapport à f_j en $f=0$. A moins que h ne soit une fonction linéaire des défauts (les directions δ_j sont alors fixes), le développement au premier ordre autour de $f=0$ n'est valable que pour des défauts de faible amplitude.

Une première idée pour lever cette difficulté pourrait consister à estimer f . Il serait alors possible de calculer les directions associées aux défauts autour de $f = \hat{f}$, et non plus autour de $f = 0$. Si l'estimation simultanée de tous les défauts était réalisée avec l'excitation « naturelle » du système, la tâche de diagnostic serait totalement remplie : tous les défauts seraient en effet identifiés, même dans le cas de défauts multiples. Cette solution « idéale » se révèle bien souvent inapplicable en pratique.

La solution retenue au chapitre IV a consisté à estimer un seul défaut et à vérifier si le comportement des résidus pouvait s'expliquer par ce seul défaut (test de modèle) :

$$r \parallel \delta_j(z, u, y, [0 \ \hat{f}_j \ 0]) ? \tag{V.36}$$

Dans (V.36), δ_j correspond au gradient calculé au chapitre IV autour de la grandeur estimée (IV.43), (IV.52). Le test de colinéarité (V.36) réalisé au chapitre IV a consisté à calculer r^\perp , la composante de r orthogonale à δ_j , et à vérifier si elle était proche de zéro (Figure IV.16). Ce test de colinéarité est à rapprocher de la valeur d'appartenance $\mu_{\delta_j}(r) \vee \mu_{-\delta_j}(r)$ qui permet également d'évaluer si r est ou n'est pas dans la direction δ_j (Figure V.10).

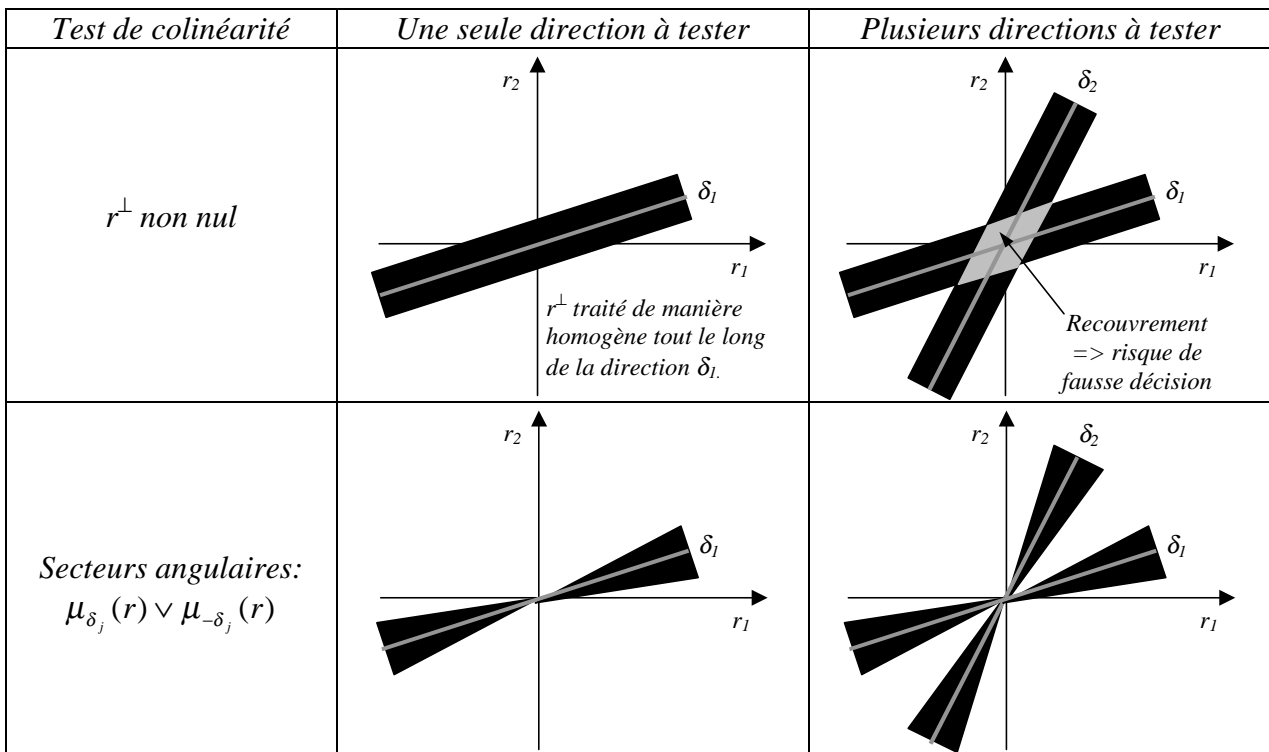


Figure V.10 : Comparaison de deux tests de colinéarité

L'intérêt des secteurs angulaires est d'éviter tout recouvrement entre les noyaux des fonctions d'appartenance (lorsque plusieurs directions doivent être testées simultanément), mais la composante orthogonale à la direction testée n'est pas traitée de manière homogène suivant l'amplitude de r (Figure V.10).

V.3.2.2. Cas des résidus du type « erreur de sortie » pour un modèle LTV

On se restreint désormais au cas où (V.33) a la forme d'un modèle linéaire variant dans le temps (LTV) (V.37).

$$G(z, \theta, f_\theta) \cdot (u + f_u) - (y - f_y) = 0 \quad (\text{V.37})$$

G désigne un transfert discret dépendant des paramètres θ et des défauts paramétriques, f_θ , f_y et f_u désignent respectivement des défauts sur les capteurs et sur les actionneurs. Le résidu construit sur le modèle de (V.34) est un résidu du type « erreur de sortie »:

$$r = y - G(z, \theta, 0) \cdot u \quad (\text{V.38})$$

$$\text{soit encore: } r = f_y + G(z, \theta, f_\theta) \cdot (u + f_u) - G(z, \theta, 0) \cdot u$$

Sur le modèle de (V.35), le développement¹ de r au premier ordre par rapport aux défauts va permettre d'étudier ses propriétés directionnelles.

$$r \approx \left[I \quad \left| \begin{array}{c} G(z, \theta, 0) \\ \frac{\partial G(z, \theta, f_\theta)}{\partial f_\theta} \Big|_{f_\theta=0} \end{array} \right. \right] \cdot u \cdot \begin{bmatrix} f_y \\ f_u \\ f_\theta \end{bmatrix} \quad (\text{V.39})$$

Pour un modèle linéaire, (V.39) permet d'analyser les propriétés directionnelles de résidus « erreur de sortie » en réponse à différents types de défauts:

- **Un défaut capteur**, f_y , agit sur r dans une direction fixe et indépendante des paramètres. Le résidu n'est excité que par f_y . Lorsque f_y est un défaut du type « offset », l'excitation de f_y se traduit par une composante continue non nulle sur le résidu (vectoriel).
- **Un défaut actionneur**, f_u , agit sur r dans une direction variable et dépendant a priori des paramètres. Comme son évolution temporelle dépend de l'excitation liée à f_u , inconnue a priori, cette direction ne peut être calculée en ligne sans hypothèse supplémentaire. Pour des défauts f_u en rampe (bornée) ou en échelon, f_u se stabilise à une valeur constante au bout d'un certain temps. Par conséquent, après un transitoire, le régime statique est atteint :

$$G(z, \theta, 0) \approx G(1, \theta, 0) \quad (\text{V.40})$$

Moyennant l'hypothèse (V.40), les directions associées aux défauts actionneurs deviennent fixes et sont par conséquent aisément calculables. Elles dépendent toujours a priori des paramètres. Si le gain statique $G(1, \theta, 0)$ est non nul, un défaut actionneur se traduit par une composante continue non nulle sur le résidu (vectoriel)

- **Un défaut paramétrique**, f_θ , agit sur r dans une direction variable et dépendant des paramètres. En négligeant la dynamique induite par f_θ (paramètre), son évolution temporelle dépend de la seule excitation sur l'entrée u . Cette direction peut donc être calculée en ligne (principe de la méthode du gradient au chapitre IV).

¹ En négligeant $\partial u / \partial f_\theta$ (action du défaut paramétrique sur l'action u par l'intermédiaire de la commande)

V.3.3. Application à la MAS

V.3.3.1. Propriétés directionnelles des résidus

Le résidu r_T , calculé au chapitre IV, correspond à l'erreur de sortie dans le repère $(\alpha\beta)$ du modèle électrique adaptatif de la MAS (IV.28), (IV.39), (IV.42), (IV.43). La discrétisation de (IV.28) donne le modèle LTV (V.41), à rapprocher de (V.37) :

$$G(z, \Omega - f_\Omega, f_P, d_T) \cdot (u + f_V) - (y_e - f_I) = 0 \quad (\text{V.41})$$

où:

$$f_I = [f_{I_{sa}} \quad f_{I_{sb}}]^T \text{ et } f_V = [f_{V_{sa}} \quad f_{V_{sb}} \quad f_{V_{sc}}]^T$$

f_I, f_V, f_Ω, f_P modélisent respectivement : les défauts des capteurs de courant, les défauts du type « interrupteur ouvert » dans l'onduleur, le défaut du capteur de vitesse et les défauts du moteur modélisés sous la forme de variations de paramètres. Le générateur de résidus r_T discrétisé ((IV.42) à (IV.45)) se traduit, outre le mécanisme d'adaptation du terme modélisant l'échauffement, \hat{d}_T , par la relation¹ (V.42), à rapprocher de (V.38) :

$$r'_T = K_{22}^{-1} \cdot r_T = y_e - G(z, \Omega, f_P, \hat{d}_T) \cdot u \quad (\text{V.42})$$

$r'_T = K_{22}^{-1} \cdot r_T$ représente le résidu r_T exprimé dans le repère (ab) . Le développement au premier ordre de (V.42) est l'analogue de (V.39). En tenant compte de l'hypothèse (V.40), on obtient² (V.43) :

$$r'_T \approx \left[\begin{array}{cc|c} 1 & 0 & \frac{1}{3R_s} \begin{bmatrix} 2 & -1 & -1 \\ -1 & 2 & -1 \end{bmatrix} \\ 0 & 1 & \end{array} \right] \left[\begin{array}{c} \frac{\partial G(z, \Omega - f_\Omega, 0, \hat{d}_T)}{\partial f_\Omega} \cdot u \\ \frac{\partial G(z, \Omega, f_P, \hat{d}_T)}{\partial f_P} \cdot u \end{array} \right] \cdot \begin{bmatrix} f_I \\ f_V \\ f_\Omega \\ f_P \end{bmatrix} \quad (\text{V.43})$$

Les propriétés établies dans la partie V.3.2.2 s'appliquent et vont être utilisées pour réaliser l'étape décisionnelle du diagnostic de la MAS. Par rapport à ces propriétés générales, un point remarquable concerne les défauts actionneurs : les directions associées à f_V sont indépendantes des paramètres du modèle. Tout comme pour les défauts f_I , une décision robuste aux perturbations paramétriques est donc possible. De plus, l'indépendance vis-à-vis des paramètres du modèle fait que les propriétés directionnelles de f_I et de f_V sont conservées, que le générateur de résidus soit adaptatif ou non, et quels que soient le ou les paramètres adaptés.

¹ On peut vérifier que la discrétisation (IV.45) n'introduit pas de couplage supplémentaire entre les équations électriques et le mécanisme d'adaptation. Le seul couplage du point de vue des équations électriques provient donc du paramètre adapté, \hat{d}_T .

² En régime statique ($d(\dots)/dt=0$) et dans un repère fixe par rapport au stator ($(dq)=(\alpha\beta)$ et $\omega_s=0$), les équations (II.58) et (II.59) de la MAS montrent que $[V_{sa} \quad V_{sb}]^T = R_s \cdot [I_{sa} \quad I_{sb}]^T$. D'où : $G(1, \Omega, f_P, \hat{d}_T) = (1/R_s) \cdot K_{22}^{-1} \cdot K_{23} = \frac{1}{3R_s} \begin{bmatrix} 2 & -1 & -1 \\ -1 & 2 & -1 \end{bmatrix}$.

Remarque 1 : un défaut d'« offset » sur l'un des capteurs de courant se traduit par l'apparition d'une composante continue sur les résidus (V.39). Compte tenu de la forme sinusoïdale des courants, il n'en va pas de même pour un défaut sur le gain d'un capteur. Cependant, la direction fixe associée à un défaut capteur reste valable.

Remarque 2 : le raisonnement mené dans le repère (ab) , avec r'_T , peut également être mené dans le repère $(\alpha\beta)$, avec r_T .

V.3.3.2. Détection

Les défauts des capteurs de courant (du type « offset »), f_I , et les défauts du type « interrupteur ouvert dans l'onduleur », f_V , se traduisent par l'apparition d'une composante continue sur r'_T (ou r_T). Cette propriété a déjà été mentionnée au chapitre IV (Figure IV.15). L'excitation de l'entrée u étant presque sinusoïdale, un filtrage passe-bas de r_T ($F_{0,999}(r_T)$) permet, d'une part, de détecter f_I et f_V , d'autre part, de les distinguer des autres défauts (V.43).

$$\mu_{Detect_cont} = 1 - \mu_Z(|F_{0,999}(r_T)|) \quad (V.44)$$

De plus, le filtrage de r_T^\perp permet de détecter, entre autres, le défaut du capteur de vitesse, f_Ω , et les défauts du moteur modélisés par des variations de paramètres, f_P (Voir le bilan de la partie IV.4.4).

$$\mu_{Detect_orth} = 1 - \mu_Z(F_{0,999}(r_T^\perp)) \quad (V.45)$$

En combinant les deux indicateurs (V.44) et (V.45), un indicateur de détection d'un défaut dans l'entraînement asynchrone est donné par (V.46) :

$$\mu_{Detect} = \mu_{Detect_cont} \vee \mu_{Detect_orth} \quad (V.46)$$

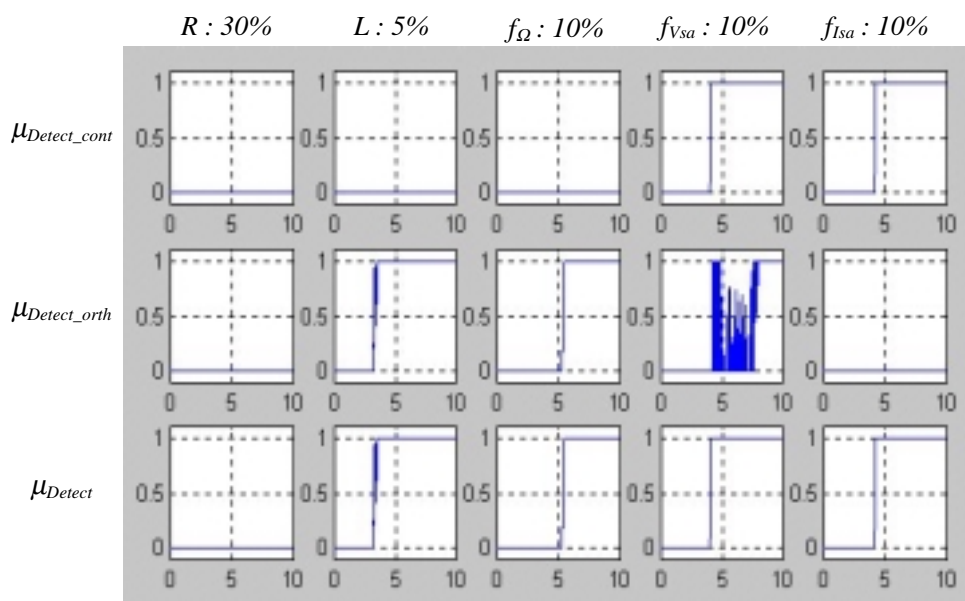


Figure V.11 : Détection des défauts appliquée à la MAS (ScCo)

La Figure V.11 montre les résultats de la détection des défauts appliquée à la MAS. Parmi les défauts de f_V (resp. f_I), seul $f_{Vsa} > 0$ (resp. f_{Isa}) est représenté, afin de simplifier la présentation des résultats. Les scénarios de défaut sont ceux de la Table IV.3 (avec le scénario complexe (*ScCo*) de la Figure IV.11). Tous les défauts sont détectés. De plus, l'indicateur μ_{Detect_cont} permet un premier niveau de localisation en discriminant les défauts f_I et f_V des autres défauts.

V.3.3.3. Localisation des défauts à direction fixe

V.3.3.3.1. Défauts de l'onduleur

A plusieurs reprises, nous avons mentionné que les défauts $f_V = [f_{Vsa} \ f_{Vsb} \ f_{Vsc}]^T$ (IV.28) modélisaient des défauts du type « interrupteur ouvert » dans l'onduleur. Avant de poursuivre le développement sur la décision appliquée au diagnostic de la MAS, quelques précisions sur la modélisation des défauts de l'onduleur sont données. La Figure V.12a rappelle de façon très simplifiée la structure d'un onduleur de tension triphasé. E est la tension d'alimentation. Le comportement des composants d'électronique de puissance formant les « bras » de l'onduleur (IGBT¹, diodes) peut être modélisé en première approximation par des « interrupteurs » : S_a^+ et S_a^- pour le bras de l'onduleur relié à la phase a du stator, par exemple. La commande de ces « interrupteurs » résulte d'une MLI (Modulation de Largeur d'Impulsion) et vise à appliquer au moteur les tensions désirées. Deux principaux types de défauts peuvent affecter l'onduleur :

- **Les défauts du type « interrupteur fermé » :** au lieu de commuter normalement, un interrupteur reste à l'état « fermé », ce qui a pour conséquence de court-circuiter l'alimentation ; la protection matérielle (fusibles) désactive alors immédiatement le système. A moins d'envisager un diagnostic prédictif des composants de puissance, le principe retenu pour la décision est le suivant : Si les fusibles ont sauté, Alors vérifier les composants de puissance de l'onduleur... et ce qu'il y a autour !
- **Les défauts du type « interrupteur ouvert » :** au lieu de commuter normalement, un interrupteur reste à l'état « ouvert ». Le système ne s'arrête pas brutalement mais fonctionne en marche dégradée. Afin de diagnostiquer ces défauts, un modèle simple de leur influence a été établi (Figure V.12b et Figure V.12c). Par exemple, si S_a^+ reste ouvert alors qu'il devait être normalement passant, la tension de la phase a est égale à $-E/2$ au lieu de $+E/2$. En moyenne sur une période de MLI², la tension V_{sa}^{moteur} est donc plus faible que la tension V_{sa} calculée par la commande. Comme, de plus, le défaut « S_a^+ ouvert » n'influe pas sur les autres phases, ce défaut est modélisé par un échelon négatif sur f_{Vsa} . Le même raisonnement s'applique aux autres défauts du même type (Figure V.12).

¹ IGBT : Insulated Gate Bipolar Transistor.

² MLI : Modulation de Largeur d'Impulsion

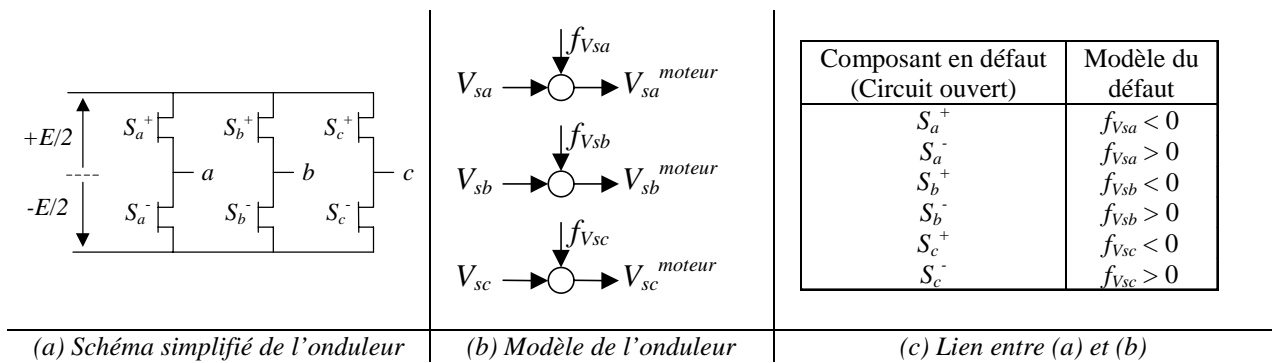


Figure V.12 : Modèle de l'onduleur en défaut

Le modèle de l'onduleur de la Figure V.12(b et c) est celui à partir duquel les propriétés directionnelles des résidus ont été étudiées (V.43). Sa validité a été étudiée en simulation, par comparaison au modèle d'un entraînement asynchrone développé au LEG. Dans ce dernier modèle, les composants de puissance de l'onduleur (éventuellement en défaut) sont modélisés par des « interrupteurs qui commutent ».

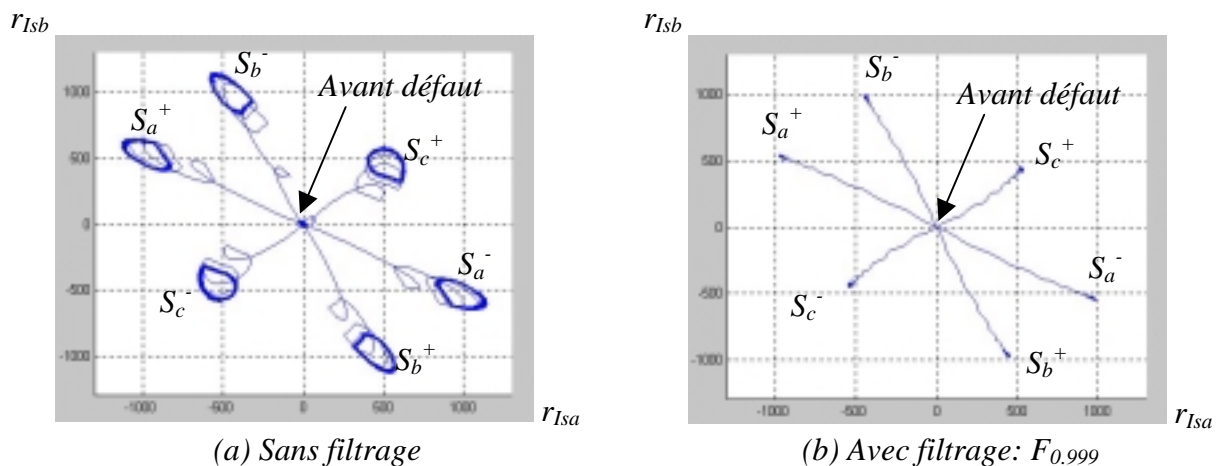


Figure V.13 : Résultats de la simulation avec un modèle « fin » de l'onduleur

Remarque 1 : les paramètres du moteur asynchrone utilisés pour les simulations de la Figure V.13 ne sont pas ceux du moteur utilisé dans le cadre de notre étude (ce qui ne modifie pas pour autant les directions associées aux défauts).

Remarque 2 : les simulations correspondant à chacun des défauts du type « interrupteur ouvert » sont superposées.

Deux résidus homogènes respectivement au courant stator de la phase a et de la phase b , r_{Isa} et r_{Isb} , ont été générés. Après un filtrage destiné à extraire la composante continue (et même avant), les propriétés directionnelles de (V.43) sont satisfaites (Figure V.13). Par exemple, lorsque S_a^+ reste ouvert, le vecteur $[r_{Isa} \ r_{Isb}]^T$ « part » dans la direction orientée $[-2 \ 1]^T$. Or, S_a^+ est modélisé par un échelon $f_{Vsa} < 0$ (Figure V.12c). (V.43) indique bien qu'un défaut $f_{Vsa} < 0$ se traduit en régime statique (V.40) par un vecteur des résidus dans la direction $[-2 \ -1]^T$. Le même raisonnement s'applique aux autres défauts.

La localisation des défauts de l'onduleur repose sur des directions orientées : une direction non orientée permet de déterminer la phase du composant en défaut ; le sens (orientation) permet de localiser le composant défaillant parmi les deux reliés à une phase donnée (Table V.13).

Composant en défaut (Circuit ouvert)	Modèle du défaut	Décision: valeurs d'appartenance aux défauts (d'après (V.43) et (V.29))	
S_a^+	$f_{Vsa} < 0$	$\mu_{D(f_{Vsa}<0)} = \mu_{-[-2 \ -1]^T} (F_{0.999} (r'_T)) \wedge \mu_{Detect_cont}$	(V.47)
S_a^-	$f_{Vsa} > 0$	$\mu_{D(f_{Vsa}>0)} = \mu_{+[2 \ -1]^T} (F_{0.999} (r'_T)) \wedge \mu_{Detect_cont}$	(V.48)
S_b^+	$f_{Vsb} < 0$	$\mu_{D(f_{Vsb}<0)} = \mu_{-[-1 \ 2]^T} (F_{0.999} (r'_T)) \wedge \mu_{Detect_cont}$	(V.49)
S_b^-	$f_{Vsb} > 0$	$\mu_{D(f_{Vsb}>0)} = \mu_{+[-1 \ 2]^T} (F_{0.999} (r'_T)) \wedge \mu_{Detect_cont}$	(V.50)
S_c^+	$f_{Vsc} < 0$	$\mu_{D(f_{Vsc}<0)} = \mu_{-[-1 \ -1]^T} (F_{0.999} (r'_T)) \wedge \mu_{Detect_cont}$	(V.51)
S_c^-	$f_{Vsc} > 0$	$\mu_{D(f_{Vsc}>0)} = \mu_{+[-1 \ -1]^T} (F_{0.999} (r'_T)) \wedge \mu_{Detect_cont}$	(V.52)

Table V.13 : Localisation des défauts du type « interrupteur ouvert » dans l'onduleur

V.3.3.3.2. Défauts des capteurs de courant

La localisation des défauts des capteurs de courant, $f_I = [f_{Isa} \ f_{Isb}]^T$, repose sur des directions non orientées : le sens dans lequel évolue le résidu (vectoriel) reflète le signe du défaut mais ne correspond pas à une classe de défaut différente (Figure V.14). Les décisions pour les défauts de f_I sont donc calculées comme suit :

$$\mu_{D(f_{Isa})} = (\mu_{[1 \ 0]^T} (F_{0.999} (r'_T)) \vee \mu_{-[-1 \ 0]^T} (F_{0.999} (r'_T))) \wedge \mu_{Detect_cont} \quad (V.53)$$

$$\mu_{D(f_{Isb})} = (\mu_{[0 \ 1]^T} (F_{0.999} (r'_T)) \vee \mu_{-[0 \ 1]^T} (F_{0.999} (r'_T))) \wedge \mu_{Detect_cont} \quad (V.54)$$

V.3.3.3.3. Conditions de localisation et identification des défauts à direction fixe

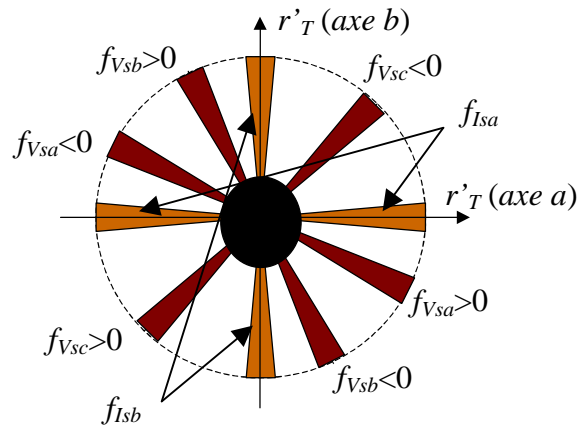


Figure V.14 : Décision pour les défauts à direction fixe

La Figure V.14 montre que les directions associées aux défauts de f_V et de f_I sont assez bien réparties dans l'espace des résidus : cette situation est intéressante pour limiter l'ambiguïté entre les décisions. Le fait que toutes les directions soient distinctes montre que les défauts de f_V et de f_I sont tous localisables entre eux sous l'hypothèse de défaut simple (HDS). Sans cette hypothèse, les conditions garantissant que les défauts puissent être localisés sont beaucoup plus restrictives : n'importe quel point du plan (ab) peut en effet être décrit par une combinaison de

deux défauts quelconques ; il suffit que les directions associées à ces deux défauts soient distinctes pour former une base du plan. Par conséquent, sans l'hypothèse de défaut simple (et en se limitant à l'étude des comportements statiques), la localisation dans le plan des résidus est impossible si plus de deux défauts sont susceptibles d'affecter le système étudié. La Table V.14 généralise ces résultats pour n_r résidus, $r=[\dots r_i \dots]$ et n_f défauts $f=[\dots f_j \dots]$ agissant dans n_f directions ($\delta=[\dots \delta_j \dots]$) de l'espace des résidus : $r=\delta.f$.

$r=\delta.f$	Avec hypothèse de défaut simple (HDS)	Défauts multiples
Condition pour pouvoir localiser les défauts:	Toutes les directions sont différentes : $\forall j \neq h, (\delta_j \angle \delta_h) \neq 0 [\pi]$	Toutes les directions sont indépendantes : $\text{Rang}(\delta)=n_f$
Nombre minimal de résidus nécessaires pour localiser:	$n_r=2 \quad (\forall n_f > 1)$	$n_r=n_f$
Identification du ou des défauts	Identification du défaut: $\hat{f}_j = \delta_j^+ . r$	Identification des défauts: $\hat{f} = \delta^+ . r$

δ^+ désigne dans cette table l'inverse à gauche de δ (de rang plein en colonnes): $\delta^+ = (\delta^T . \delta)^{-1} . \delta^T$

Table V.14 : Conditions pour pouvoir localiser des défauts conduisant à des résidus directionnels, Identification des défauts (cas statique)

Dans l'application à la MAS, l'identification d'un défaut dans l'onduleur pourrait servir à compenser la composante continue induite par ce défaut, mais ne donne pas une information supplémentaire sur le défaut qui est en tout ou rien. L'identification d'un défaut affectant un capteur de courant est à la fois utile et triviale avec des résidus du type « erreur de sortie » :

$$\begin{aligned} \hat{f}_{Isa} &= [1 \ 0].r'_T, \quad \text{si } f_{Isa} \text{ seul} \\ \hat{f}_{Isb} &= [0 \ 1].r'_T, \quad \text{si } f_{Isb} \text{ seul} \end{aligned} \tag{V.55}$$

V.3.3.3.4. Résultats

Chaque figure à l'intérieur de la Figure V.15 correspond à la simulation d'un scénario de défaut particulier (Table IV.3 et scénario complexe (ScCo) de la Figure IV.11). Pour chacune de ces situations, l'évolution temporelle des huit valeurs d'appartenance décrivant la décision est représentée ((V.47) à (V.54)). Les défauts considérés sont correctement localisés (Figure V.15).

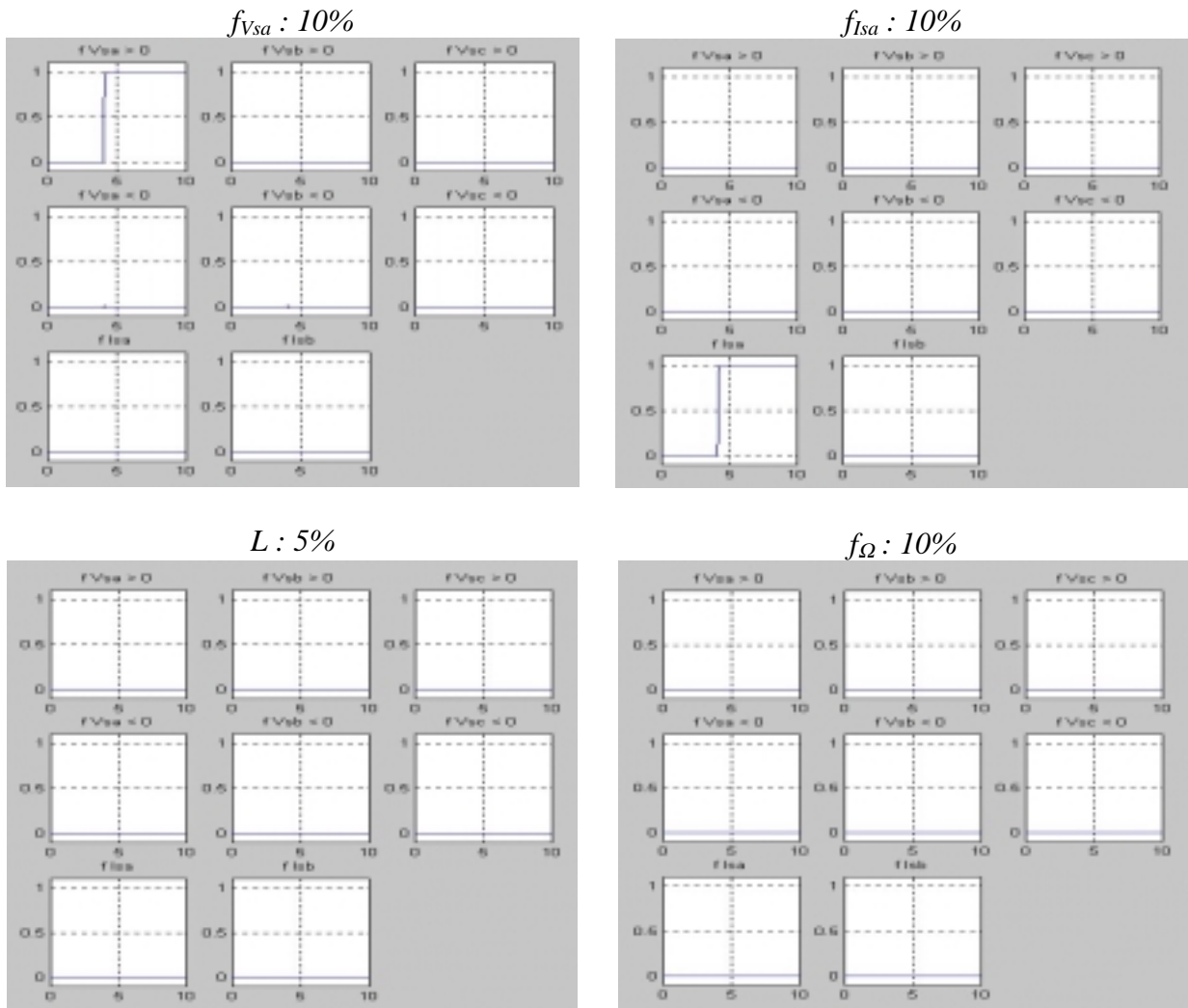


Figure V.15 : Localisation des défauts à direction fixe (défauts des capteurs de courant, f_b , et défauts du type « interrupteur ouvert » dans l'onduleur, f_V) (ScCo)

V.3.3.4. Localisation des défauts à directions variables

La localisation des défauts à direction variable est maintenant traitée. Ces défauts interviennent sous la forme de paramètres dans les équations électriques de la MAS : il s'agit du défaut du capteur de vitesse, f_Ω , et des défauts du moteur modélisés sous la forme de variations de paramètres, f_p . Le générateur de résidus réalisant le découplage indirect de f_Ω est celui utilisé pour localiser entre eux f_p et f_Ω (partie IV.4.5). Afin de cibler la décision sur ces seuls défauts, un ensemble flou décrivant si l'un des autres défauts (f_I , f_V) a eu lieu est défini : $\mu_{Detect} \wedge (1 - \mu_{Detect_cont})$ est proche de « 1 » si un défaut a été détecté et que ce dernier ne se traduit pas par l'apparition d'une composante continue, ce qui exclut f_I et f_V .

Remarque : une manière plus sélective de réaliser cette exclusion consiste à utiliser $\mu_{Detect} \wedge (1 - \mu_{D(\{f_I, f_V\})})$ où $\mu_{D(\{f_I, f_V\})}$ est défini par (V.56).

$$\mu_{D(\{f_I, f_V\})} = \mu_{D(I_{sa})} \vee \mu_{D(I_{sb})} \vee \mu_{D(f_{Vsa} < 0)} \vee \mu_{D(f_{Vsa} > 0)} \vee \mu_{D(f_{Vsb} < 0)} \vee \mu_{D(f_{Vsb} > 0)} \vee \mu_{D(f_{Vsc} < 0)} \vee \mu_{D(f_{Vsc} > 0)} \quad (V.56)$$

La localisation de f_P et de f_Ω consiste pour l'essentiel à tester r_Ω^\perp ((IV.54) et le bilan de la partie IV.4.5). En supposant que la convergence de l'estimation \hat{f}_Ω soit assurée, une valeur significativement non nulle de r_Ω^\perp (testée par $1 - \mu_Z(F_{0,999}(r_\Omega^\perp))$) signifie que le défaut f_Ω est impossible (Table IV.1). Cela permet de construire une valeur d'appartenance, $\mu_{D(f_P)}$, qui devient proche de « 1 » pour toute variation des paramètres autre que celle de f_Ω (et celles induites par l'échauffement en tenant compte de la remarque de la partie IV.4.5.1 (V.57).

$$\mu_{D(f_P)} = (1 - \mu_Z(F_{0,999}(r_\Omega^\perp))) \wedge (\mu_{Detect} \wedge (1 - \mu_{Detect_cont})) \quad (V.57)$$

Même si nous ne l'avons testé que dans des situations où les inductances L_s et L_r varient dans les mêmes proportions, $\mu_{D(f_P)}$ est a priori sensible à bien d'autres types de variations de paramètres (avec un degré de sensibilité à étudier). $\mu_{D(f_P)}$ est donc l'indicateur que nous associons aux défauts du moteur, tout en restant conscient que la modélisation de ce type de défauts par des variations des paramètres du modèle de Park n'est qu'une vue de l'esprit : des courts-circuits de spires ou des ruptures de barres se traduisent par un déséquilibre de la machine que le modèle de Park ne peut représenter, d'où la nécessité d'un important travail de modélisation [Schaeffer, 1999].

Lorsque r_Ω^\perp est proche de zéro, ce qui est testé par $\mu_Z(F_{0,999}(r_\Omega^\perp))$, le comportement de la MAS est assimilable à celui induit par f_Ω . Pour conclure à un défaut f_Ω , il est donc nécessaire de « prouver » que les autres hypothèses de défaut sont exclues (Table IV.1). Pour f_i et f_v , cela est obtenu par le terme $\mu_{Detect} \wedge (1 - \mu_{Detect_cont})$. Pour f_P , nous supposons que les défauts du moteur se traduisent par un comportement distinct de celui du défaut f_Ω . Ceci a été vérifié au chapitre IV dans le cas où les inductances varient dans les mêmes proportions (partie IV.4.5). Par conséquent, $\mu_Z(F_{0,999}(r_\Omega^\perp)) \wedge \mu_{Detect} \wedge (1 - \mu_{Detect_cont})$ indique que le comportement de la MAS est celui lié à la présence de f_Ω , y compris le bon fonctionnement ($f_\Omega=0$). Dans ces conditions, \hat{f}_Ω estime effectivement le défaut du capteur de vitesse (Table IV.1). L'indicateur de présence du défaut f_Ω , $\mu_{D(f_\Omega)}$, est donc obtenu en excluant le comportement normal par le terme $1 - \mu_Z(\hat{f}_\Omega)$:

$$\mu_{D(f_\Omega)} = \mu_Z(F_{0,999}(r_\Omega^\perp)) \wedge (1 - \mu_Z(\hat{f}_\Omega)) \wedge (\mu_{Detect} \wedge (1 - \mu_{Detect_cont})) \quad (V.58)$$

La Figure V.16 montre l'évolution temporelle de $\mu_{D(f_P)}$ et de $\mu_{D(f_\Omega)}$ pour différents scénarios de défaut (Table IV.3 et scénario simple (*ScSi*) ou complexe (*ScCo*) de la Figure IV.11). Les défauts considérés sont correctement localisés.

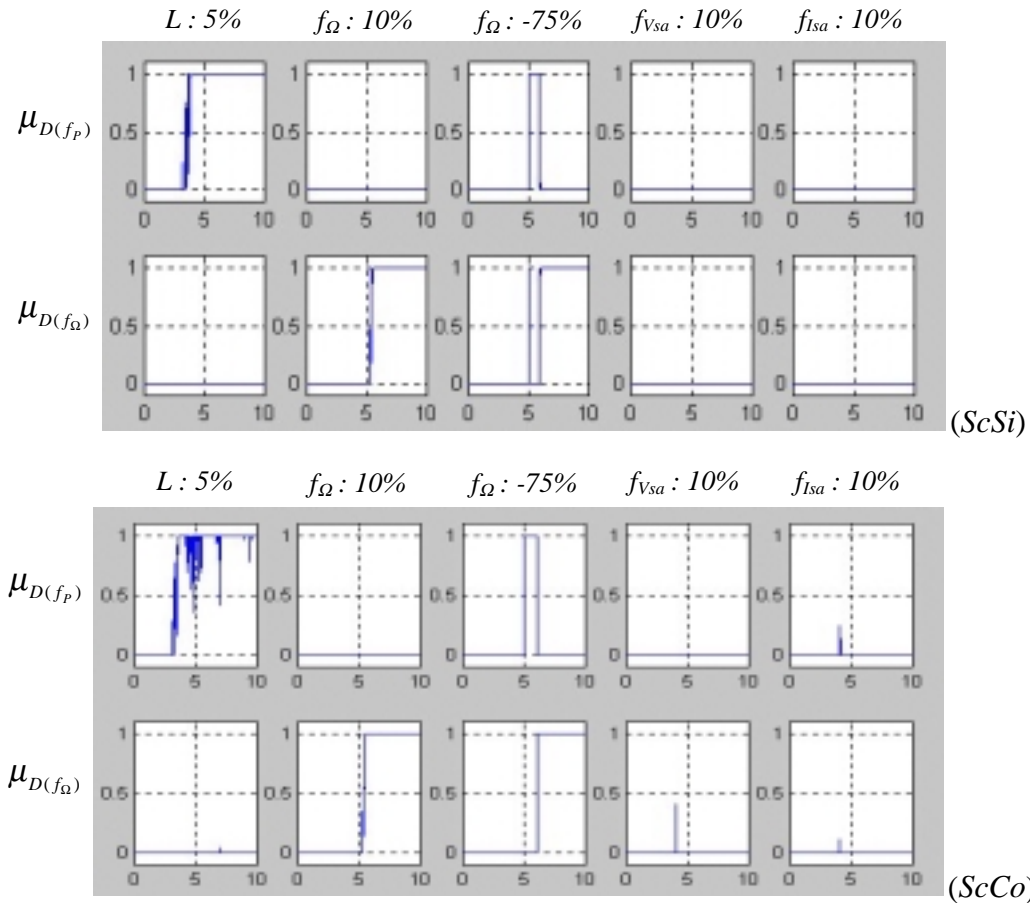


Figure V.16 : Localisation des défauts à direction variable (défauts du capteur de vitesse, f_Ω , et défauts du moteur modélisés sous la forme d'une variation des paramètres, f_P)

Dans le cas des scénarios complexes (ScCo), la variation des inductances de 5% correspond à une situation proche de la limite de sensibilité, déjà évoquée au chapitre IV (Figure IV.19). Les petits pics qui apparaissent en réponse à f_{Vsa} et à f_{Isa} s'expliquent par un léger retard à la détection de la composante continue (retard lié à la constante de temps du filtre passe-bas utilisé) ; ainsi, le terme $(1 - \mu_{Detect_cont})$ de (V.57) et de (V.58) ne s'annule pas immédiatement et rend possible l'apparition d'une fausse décision pendant une courte durée. Le cas d'un défaut en échelon de -75% sur f_Ω (obtenu dans des conditions particulières indiquées au chapitre IV) est à rapprocher de l'interprétation de la (Figure IV.23). L'hypothèse de variations lentes de f_Ω n'étant pas vérifiée, l'adaptation de \hat{f}_Ω ne converge pas immédiatement. Cela se traduit par la présence d'un transitoire qui se retrouve au niveau de la décision : tant que l'algorithme d'estimation n'a pas convergé, le comportement du système n'est pas assimilé à un défaut f_Ω , induisant ainsi transitoirement une fausse décision. Dans les limites inhérentes à nos hypothèses de travail, la détection et la localisation des défauts considérés ont été réalisées (Figure V.11, Figure V.15, Figure V.16).

V.4. Bilan

V.4.1. Point de vue méthodologique

Le principe du diagnostic, évaluer la cohérence entre des observations et un modèle, s'applique à l'étape décision : lorsque la décision est construite à partir de résidus, un modèle de l'évolution des résidus en réponse aux défauts est nécessaire pour construire des fonctions de décision. Ce modèle peut résulter directement d'une table de signature (qui décrit les spécifications lors de la génération de résidus) ou bien il peut résulter d'une analyse a posteriori des résidus.

En particulier, les résultats d'une analyse de sensibilité donnent une information beaucoup plus riche qu'une simple table de décision booléenne. Nous avons utilisé cette information pour optimiser des fonctions de décision par sélection des résidus les plus pertinents. Il apparaît que l'interface entre un module **numérique** de diagnostic et un module **logique** est cruciale pour garantir voire optimiser les performances du diagnostic global. De plus, le parallèle entre la génération des résidus et la synthèse des fonctions de décision nous semble particulièrement intéressant : les fonctions de décision sont construites, d'une part, pour être découplées vis-à-vis de tous les défauts sauf un (critère de localisation garantissant la robustesse de la décision), d'autre part, pour être aussi sensibles que possible au défaut non découplé. On retrouve donc le traditionnel compromis entre sensibilité et robustesse, mais du côté de la décision. Le parti que nous avons pris a alors consisté à privilégier la sensibilité tout en garantissant une robustesse minimale (localisation). Afin de synthétiser des fonctions de décision appropriées, deux approches ont été utilisées : l'une repose sur un algorithme dédié (efficace mais non réutilisable dès que la formulation du problème change), l'autre repose sur un algorithme génétique (nécessitant plus de calculs mais très facilement transposable à d'autres problèmes). La pondération entre les critères de localisation et de sensibilité dans le critère global à optimiser laisse entrevoir la possibilité d'agir sur le compromis entre sensibilité et robustesse dans la synthèse des fonctions de décision. Malgré son intérêt sur le plan conceptuel, cette approche de la décision a été réalisée dans un cadre contraignant afin de faciliter la modélisation du lien entre les défauts et les indicateurs de décision : l'étude de sensibilité n'a été effectuée que pour des défauts en échelon, l'incertitude dans l'interprétation des dépassements de seuil n'est pas prise en compte et les tests sur les résidus sont booléens. Ce dernier point peut favoriser l'instabilité des décisions et a justifié l'introduction de la logique floue.

En gardant comme ligne directrice l'idée d'enrichir l'information contenue dans le modèle utilisé pour la décision, un test des ordres de grandeur relatifs des résidus a été introduit. La possibilité d'appliquer ce test repose sur les propriétés directionnelles des résidus qui ont également été étudiées. Un autre « pont » entre la génération des résidus et la décision est alors apparu : la méthode du gradient développée au chapitre IV repose finalement sur les mêmes principes que la décision basée sur les propriétés directionnelles des résidus. En effet, le gradient représente ni plus ni moins que la direction, variable au cours du temps, dans laquelle un résidu (vectoriel) évolue sous une hypothèse donnée : hypothèse de perturbation (échauffement) ou hypothèse de défaut (défaut du capteur de vitesse). Que ce soit pour la génération de résidus ou pour la décision, toute la difficulté consiste alors à déterminer à quelles conditions une direction est valide (influence de l'excitation du système, des perturbations, des erreurs de modélisation,...). Il est alors possible de tester si un résidu (vectoriel) est ou n'est pas dans la direction attendue. Le diagnostic est avant tout un problème de découplage : par exemple, les découplages peuvent être réalisés par projection (parallèlement ou orthogonalement à une direction), ou par la définition d'un ensemble flou sur un domaine donné (la valeur d'appartenance correspondante masque (découple) tous les éléments n'appartenant pas du tout à l'ensemble considéré). Dans ce

sens, réaliser un découplage revient à spécifier un certain domaine, et de nombreuses techniques, numériques et/ou logiques, permettent de remplir cet objectif.

La combinaison usuelle entre la génération des résidus et la décision permet de profiter des avantages respectifs de deux types de modèles pour construire de tels domaines. Les modèles numériques sont très précis mais, en contrepartie, la description de domaines est ardue alors que c'est bien souvent l'inverse pour les modèles logiques. Les modèles numériques servent donc souvent à définir une référence précise autour de laquelle les domaines sont définis par un modèle logique. Par exemple, $\mu_z(|y - G(z)u|)$ définit un domaine désignant un ensemble de comportements relatifs possibles pour u et y (entrées et sorties du système), autour d'un comportement de référence modélisé de façon très précise par $G(z)$. $G(z)$ est tellement précis qu'il n'a pratiquement aucune chance d'être valide : les résidus $r = y - G(z)u$ ne sont jamais exactement nuls à cause du bruit, des perturbations, des erreurs de modélisation, ... Un moyen de restaurer la validité consiste alors à définir le domaine μ_z autour du comportement de référence ($r=0$). La validité des modèles est cruciale pour le diagnostic car toute la confiance dans les indicateurs de défaut repose sur elle. Malgré l'intérêt indéniable de la combinaison des modèles numériques et logiques, la combinaison de grandeurs non homogènes fait qu'il est bien difficile de garantir la validité du diagnostic global, d'où l'importance de bien modéliser les interfaces numériques/logiques (Interprétation des dépassements de seuil, ...) et d'approfondir l'exploitation des propriétés directionnelles.

V.4.2. Point de vue des applications

L'application à la MCC de l'optimisation des fonctions de décision permet de vérifier sur un cas concret les développements théoriques. Des développements ou, en tout cas, des vérifications supplémentaires sont toutefois nécessaires pour déboucher sur une réelle application ; ces vérifications sont liées aux hypothèses de travail qui ont été faites (en particulier, défauts en échelon) et à l'influence de l'ordre des dépassements de seuil. Ce dernier point est difficile à modéliser car il dépend de la « forme » du défaut (échelon, rampe, ...). Notons que l'ordre des dépassements de seuil est lié aux sensibilités relatives des résidus et, par conséquent, à l'évolution de la direction du vecteur des résidus au cours du temps.

Dans l'application à la MAS, un modèle simple de l'onduleur en défaut a été proposé. Ce modèle suffit pour modéliser la direction de la composante continue des résidus suite aux défauts du type « interrupteur ouvert ». Le filtrage passe-bas ($F_{0,999}$) des résidus déduits des équations électriques permet de découpler les défauts dont l'excitation présente une composante continue non nulle (défauts des capteurs de courants, défauts de l'onduleur) des défauts paramétriques (défauts du capteur de vitesse, défauts du moteur). Ces défauts paramétriques ne sont « visibles » qu'à travers l'excitation sinusoïdale de l'entrée. Le filtrage passe-bas $F_{0,999}$ est donc un cas particulier de découplage fréquentiel qui est presque parfait, compte tenu des excitations très différentes des termes découplés (sauf à très basse vitesse). L'utilisation des propriétés directionnelles a permis de localiser 8 défauts (défauts des 6 interrupteurs de l'onduleur (f_V) et des 2 capteurs de courant (f_I)) à partir de seulement 2 résidus. Ce résultat a été facilité par les « bonnes » propriétés du modèle de l'entraînement : les directions en réponse aux défauts f_V sont fixes (du moins, les composantes à basse-fréquence) et, surtout, indépendantes des paramètres du modèle. Les défauts paramétriques se traduisent par des résidus dont la direction varie au cours du temps. En se plaçant dans un repère lié à la direction δ_Ω variable de f_Ω (méthode du gradient au chapitre IV), d'autres résidus peuvent être générés (r_Ω'' et r_Ω^\perp , Figure IV.16, (IV.54)). Une décision résultant de l'interprétation des tests de modèle (Table IV.1) a alors été proposée pour ces défauts. Les objectifs fixés en terme de détection et de localisation des défauts ont été remplis et

la plupart des défauts (f_I , f_Ω voire f_V) peuvent également être identifiés. En fonctionnement normal, un terme lié à l'échauffement, \hat{d}_T , est également identifié. Les principales limites proviennent de l'hypothèse de défaut simple, de l'hypothèse de variations lentes de f_Ω et de la modélisation des perturbations (échauffement, pertes fer,...) et des défauts du moteur.

Remarque : pour être détectés, les défauts du moteur n'ont pas besoin d'être modélisés. Même si des simulations avec des variations des inductances ont été réalisées (pour « valider »), le système de diagnostic de la MAS a été développé sans avoir recours à un modèle des défauts du moteur : ainsi, tout ce qui n'est pas interprétable par f_I , f_V ou f_Ω est considéré comme un défaut du moteur. Un modèle déséquilibré de la MAS n'est utile que si l'on souhaite localiser entre eux les défauts du moteur (qui doivent alors être modélisés) ou si l'on souhaite continuer à diagnostiquer f_I , f_V et f_Ω , alors qu'un défaut moteur est présent. Ce dernier cas présente un réel intérêt pratique, puisqu'un moteur peut fonctionner en marche dégradée pendant très longtemps malgré la présence de courts-circuits de spires, de ruptures de barres,... Cependant, le diagnostic à construire pour remplir de tels objectifs ne peut plus être basé sur la tout aussi sympathique que limitative hypothèse de défaut simple.

Chapitre VI :
Décision dynamique

Chapitre VI :

Décision dynamique

VI.1. Introduction

Une table de décision est un modèle décrivant le comportement des résidus en réponse aux défauts. Plusieurs manières d'enrichir l'information d'une telle table, par rapport à une table de décision booléenne, ont été étudiées au chapitre V. En particulier, la prise en compte des signes ou des ordres de grandeur relatifs permet d'accroître la précision du modèle utilisé pour la synthèse des fonctions de décisions. Jusqu'ici, le modèle utilisé pour la décision était statique ; autrement dit, la décision à un instant donné ne tenait compte que de la valeur des résidus¹ à cet instant. La prise en compte d'une évolution passée dans la décision peut néanmoins enrichir utilement l'information utilisée pour le diagnostic. La décision n'est alors plus statique mais dynamique.

Cependant, les comportements temporels des résidus sont en général mal connus. Ce constat est d'ailleurs celui qui nous avait conduit à focaliser notre attention sur une décision statique au chapitre V. Une manière de lever ce manque de connaissance peut consister à décrire qualitativement l'évolution temporelle des résidus, par exemple à l'aide de séquences. Dans ce cas, l'évolution temporelle n'est certes pas décrite très précisément, mais le modèle décrivant cette évolution reste valide pour un large éventail de situations.

La connaissance sur le comportement (statique ou dynamique) des résidus en réponse aux défauts étant souvent délicate à obtenir, une solution consiste à utiliser des méthodes d'apprentissage pour la décision. Dans la deuxième partie de ce chapitre, une décision calculée à partir d'un réseau de neurones récurrent est donc proposée.

¹ Pour les raisons indiquées dans l'introduction du chapitre V, la décision que nous construisons porte sur des résidus et non directement sur les mesures.

VI.2. Un modèle dynamique pour la décision : les séquences floues

VI.2.1. Préliminaire : croissance et décroissance d'un signal

Nous cherchons tout d'abord à recueillir une information sur la croissance (ou la décroissance) d'un signal au cours du temps. Le résultat de cette partie est très général et n'est pas propre au modèle dynamique pour la décision qui sera construit par la suite. Aussi, la construction d'un indicateur de croissance et de décroissance d'un signal est traitée en préliminaire. Le cahier des charges pour cet indicateur est de détecter une phase de croissance (ou de décroissance) dans un signal **indépendamment de l'amplitude et de la durée** de cette phase et de manière **robuste au bruit**. Ainsi, un tel indicateur doit être capable de détecter aussi bien une croissance forte sur une courte durée qu'une croissance lente sur une longue durée.

Une des premières idées qui vient à l'esprit est de calculer la pente du signal et d'étudier son signe. Soit u le signal étudié, T_e la période d'échantillonnage et u_k la valeur du signal à l'instant k . La pente $p_k = (u_k - u_{k-1})/T_e$ doit tout d'abord être filtrée car p_k est très sensible au bruit de mesure.

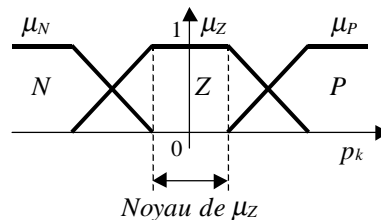


Figure VI.1 : Fuzzification de la pente

Suivant que le signal est peu ou fortement croissant, la pente est faible ou forte ; par conséquent, l'indicateur de croissance de u_k reste très majoritairement¹ Z , si u_k croît lentement. Dans ces conditions, une phase de croissance induisant des écarts importants sur u_k (croissance lente pendant longtemps) peut passer complètement inaperçue. Le cahier des charges n'est donc pas satisfait. Une autre approche consiste à fuzzifier $u(k) - u(k-n)$. Le choix de n revient à fixer arbitrairement l'échelle de temps sur laquelle la croissance de u_k est étudiée². Le cahier des charges n'est toujours pas rempli : par exemple, si n est choisi trop grand, des variations rapides peuvent passer complètement inaperçues (Figure VI.2).

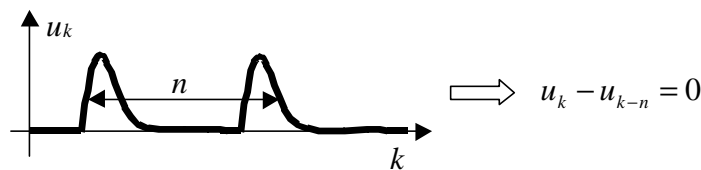


Figure VI.2 : Limites de la fuzzification de $u_k - u_{k-n}$

Une fuzzification de $u_k - u_{k-n}$ pour plusieurs valeurs de n (autrement dit, pour plusieurs échelles de temps) est une approche possible [Evsukoff, 1998b]. Nous ne l'avons pas retenue, car elle conduit à une multitude d'indicateurs qui doivent être ensuite agrégés avec précaution. De plus, la sensibilité au bruit doit être traitée pour des petites valeurs de n .

¹ Voire totalement si le noyau de μ_Z ne se réduit pas au singleton $\{0\}$ (Figure VI.1).

² Le choix de n s'apparente au choix de la constante de temps pour le filtrage de p_k

Pour tester si un signal est croissant ou décroissant, il faut de toute façon comparer le « présent » et le « passé proche » du signal. La notion de « passé proche » dépend fortement des dynamiques caractérisant l'évolution temporelle de u_k . Soit x un état décrivant le « passé proche » du signal u . x_k peut être interprété comme une référence dans le passé de u_k à partir de laquelle la croissance de u_k est étudiée : si u_k est supérieur (resp. inférieur) à x_k alors u est croissant (resp. décroissant) à l'instant k ; si u_k est proche de x_k alors u est constant à l'instant k (Figure VI.3).

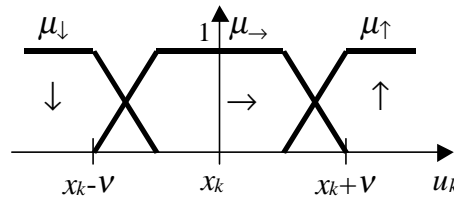


Figure VI.3 : Fuzzification permettant de comparer u_k à la référence x_k

\downarrow , \rightarrow et \uparrow désignent les ensembles flous associés respectivement à la décroissance, la « constance » et la croissance de u_k (Figure VI.3). Le paramètre v est choisi en fonction du niveau de bruit dans le signal u . Il reste à définir l'évolution de la référence x_k au cours du temps. Nous avons opté pour la solution suivante : x_k n'est mis à jour que si l'on est sûr (valeur d'appartenance égale à un) de la croissance ou de la décroissance de u_k . Autrement dit, le recalage de la référence sur le passé de u_k n'a lieu que si une variation significative de u_k s'est produite ; par « variation significative », on attend une variation supérieure à la tolérance v définissant le niveau de bruit (VI.1).

$$\begin{aligned} x_{k+1} &= u_k - v, & \text{si } \mu_{\uparrow} &= 1 \\ x_{k+1} &= u_k + v, & \text{si } \mu_{\downarrow} &= 1 \\ x_{k+1} &= x_k, & \text{sinon} \end{aligned} \tag{VI.1}$$

Soit $y_k = [\mu_{\uparrow} \quad \mu_{\downarrow}]^T$. Les équations des indicateurs de croissance et de décroissance de u peuvent être mises sous la forme de (VI.2).

$$\begin{aligned} x_{k+1} &= g(x_k, u_k, y_k) \\ y_k &= h(x_k, u_k) \end{aligned} \tag{VI.2}$$

(VI.2) est une représentation d'état dans laquelle l'équation d'état g correspond à (VI.1) et dans laquelle l'équation de mesure correspond à la fuzzification de la Figure VI.3. Il est assez amusant de constater que le système dynamique retenu pour calculer les indicateurs de croissance et de décroissance d'un signal a la forme d'un observateur (l'évolution de l'état dépend des mesures), contrairement aux systèmes correspondant aux alternatives non retenues (Figure VI.1 et Figure VI.2). (VI.2) n'est pas un système purement logique, puisque x_k n'est pas une grandeur logique. (VI.2) fait partie d'une classe très large de systèmes dynamiques, pour lesquels les équations sont décrites en combinant des opérateurs et des grandeurs qui peuvent être arithmétiques et logiques (systèmes hybrides).

La Figure VI.4 montre l'évolution de x_k et de $y_k = [\mu_{\uparrow} \quad \mu_{\downarrow}]^T$ sur un signal de test u_k pour deux valeurs de la tolérance au bruit v . En dépit de la présence de bruit, les phases de croissance et de décroissance (au sens large¹) sont correctement indiquées. La valeur de x_k change d'autant plus souvent que la pente de u_k est forte et que la tolérance au bruit est faible. Plus v est faible, plus le bruit influe sur les indicateurs de croissance et de décroissance, qui peuvent devenir non nuls à des instants rapprochés (voir les zones entourées sur la Figure VI.4). L'état x_k « suit » le signal u_k .

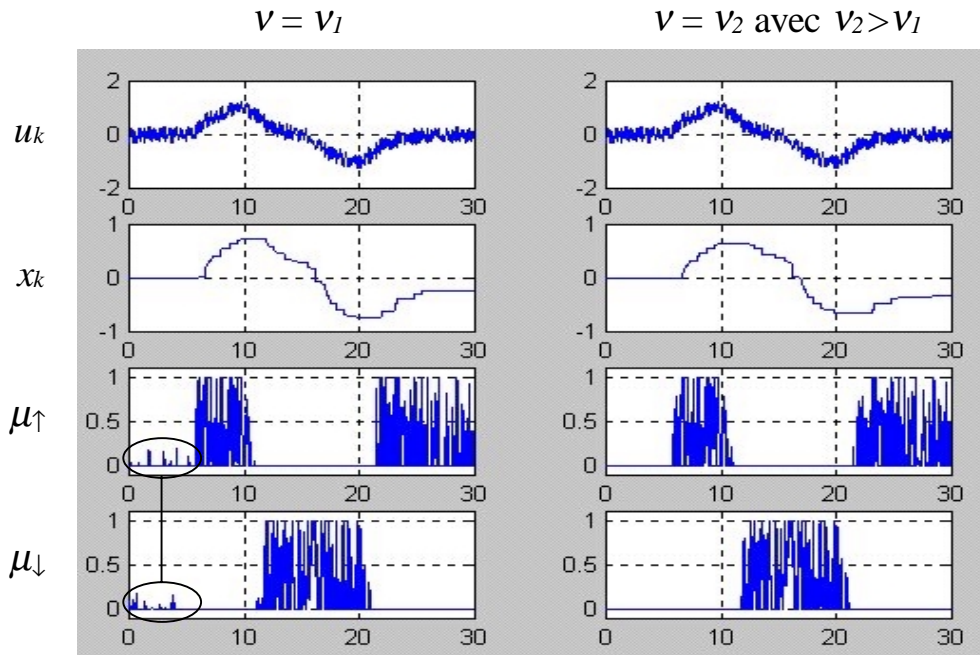


Figure VI.4 : Simulation des indicateurs de croissance et de décroissance d'un signal

Un indicateur de croissance (et de décroissance) d'un signal a été construit. Il dépend très peu de l'amplitude et de la durée des phases de croissance et de décroissance du signal tout en étant robuste au bruit et facile à régler (v peut être choisi proportionnel à l'écart type du bruit). En contrepartie, cet indicateur ne donne pas une information précise sur l'amplitude de la croissance ou de la décroissance du signal.

En considérant implicitement que $x_0 = u_0$, l'appartenance de u_k à l'ensemble flou « croissant », \uparrow , (resp. décroissant, \downarrow) sera notée par la suite $\mu_{\uparrow}(u_k)$ (resp. $\mu_{\downarrow}(u_k)$).

VI.2.2. Intérêt d'un modèle dynamique pour la décision

VI.2.2.1. Exemple statique

Afin d'illustrer l'intérêt d'un modèle dynamique pour la décision, la décision statique avec prise en compte des signes proposée au chapitre V (partie V.3.1.2) est tout d'abord appliquée aux résidus globaux de la machine à courant continu (MCC). Le couple résistant est supposé connu. La table de décision utilisée pour synthétiser les fonctions de décision est la Table V.11. Un exemple de fonction de décision est donné par (V.23) à (V.25).

¹ Par opposition à « strictement croissant » et à « strictement décroissant » qui excluent les phases où le signal est constant.

La Figure VI.5 montre les évolutions temporelles des résidus en réponse à des défauts modélisés par des échelons (positifs) et les décisions correspondantes. Deux représentations sont utilisées : l'une est classique, l'autre est fondée sur un code de couleur. Pour les résidus, la couleur verte représente une appartenance forte à Z (Zéro), alors que la couleur rouge (resp. bleue) représente une appartenance forte à P (Positif) (resp. N (Négatif)). Pour les décisions, la couleur verte représente une valeur proche de « zéro » et la couleur rouge une valeur proche de « un ». Cette représentation favorise une interprétation qualitative des résultats obtenus.

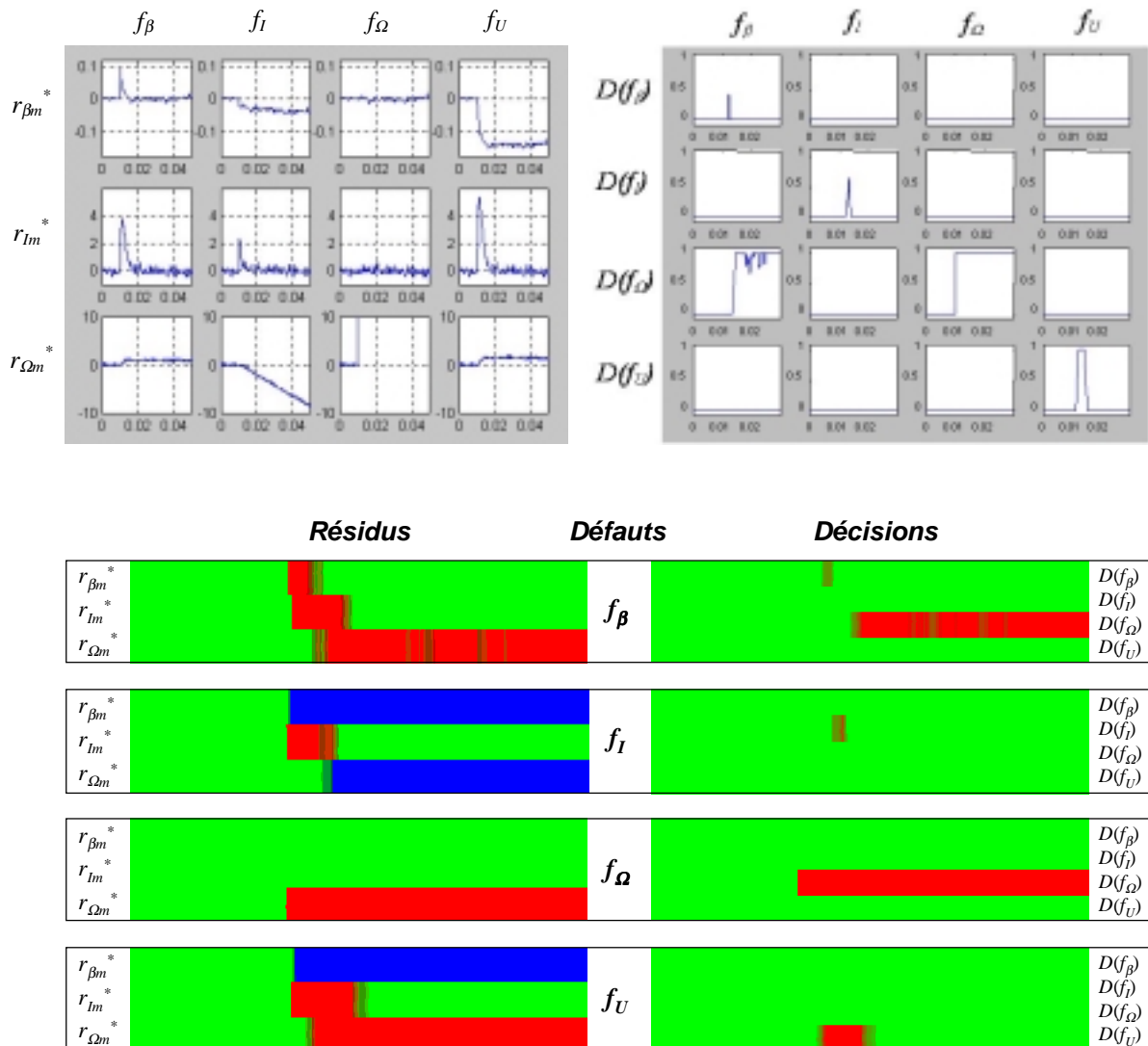


Figure VI.5 : Evolution temporelle des résidus globaux et des décisions statiques avec prise en compte des signes (sur l'intervalle de temps [0 ; 0.03s])

Les décisions associées à f_I et à f_U sont fugaces alors que le défaut est toujours présent. En particulier, la décision pour le défaut f_I est non seulement fugace, mais dépasse à peine la valeur 0.5. En observant le comportement qualitatif des résidus en réponse à f_I , on peut remarquer que les trois résidus ne correspondent à la table de décision qualitative (Table V.11) que pendant un très court instant : en effet, une appartenance forte de r_{Im}^* à P n'a pas lieu simultanément avec une appartenance forte de $r_{\Omega m}^*$ à N . Plus grave, après un comportement fugace qui s'explique de

la même manière que pour f_i , la décision correspondant au scénario f_β est une fausse décision : les résidus $r_{\beta m}^*$ et r_{Im}^* ayant un comportement fugace en réponse à f_β , lorsqu'ils « retombent » à zéro, seul $r_{\Omega m}^*$ reste positif, ce qui correspond exactement à la signature du défaut f_Ω . Cela explique l'apparition de la fausse décision qui conclut à un défaut f_Ω , alors que f_β s'est produit.

Pour éviter des fausses décisions ou des pertes de sensibilité de la décision, qui sont dues à des différences dans les dynamiques de réponse des résidus aux défauts, il est nécessaire de tenir compte du passé dans la décision à l'instant courant.

VI.2.2.2. Première approche : mémorisation des dépassements de seuil

Une première approche pour tenir compte du passé consiste à mémoriser les dépassements de seuil. Ce point a déjà été évoqué au chapitre V (Figure V.4). Ainsi, le test T du résidu r par rapport à son seuil λ découle d'une relation dynamique traduisant un « effet mémoire » :

$$T_{k+1} = T_k \vee (|r_k| > \lambda), \quad T_0 = 0 \quad (\text{VI.3})$$

(VI.3) s'interprète comme suit : le test de r à l'instant $k+1$ prend la valeur « vrai » si et seulement si ce test était précédemment « vrai » ou si r dépasse son seuil. T est donc « vrai » si r a dépassé **au moins une fois** son seuil par le passé. La négation de T vaut « vrai » si r n'a jamais dépassé son seuil par le passé, autrement dit, si r est **toujours** resté inférieur à son seuil.

La généralisation à la logique floue est directe : les ensembles flous \underline{N} , \underline{P} et \underline{Z} signifient respectivement : « r a été **au moins une fois** N (négatif) par le passé », « r a été **au moins une fois** P (positif) par le passé » et « r a **toujours** été Z (zéro) par le passé ».

$$\mu_{\underline{N},k+1} = \mu_{\underline{N},k} \vee \mu_N(r_k), \quad \mu_{\underline{N},0} = 0 \quad (\text{VI.4})$$

$$\mu_{\underline{P},k+1} = \mu_{\underline{P},k} \vee \mu_P(r_k), \quad \mu_{\underline{P},0} = 0 \quad (\text{VI.5})$$

$$\mu_{\underline{Z},k+1} = \mu_{\underline{Z},k} \wedge \mu_Z(r_k), \quad \mu_{\underline{Z},0} = 1 \quad (\text{VI.6})$$

La décision construite à partir de la Table V.11, où les fonctions de décision sont construites sur le modèle de (V.23) à (V.25) en remplaçant μ_N , μ_P et μ_Z par $\mu_{\underline{N}}$, $\mu_{\underline{P}}$ et $\mu_{\underline{Z}}$, est représentée sur la Figure VI.6. Les fausses décisions et les décisions fugaces sont résolues par l'introduction d'un « effet mémoire » dans la décision, montrant ainsi l'intérêt d'une décision non pas statique mais dynamique.

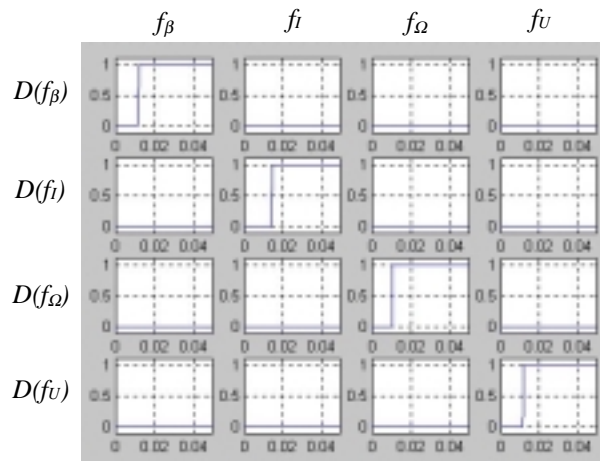


Figure VI.6 : Décision avec « effet mémoire »

VI.2.3. Modèle de l'action d'un défaut sur un résidu : séquences floues

Si la mémorisation des dépassements de seuil et son extension à la logique floue ont permis de résoudre le problème de la décision pour l'application à la MCC (Figure VI.5 et Figure VI.6), cette approche élémentaire d'une décision dynamique présente certaines limites. En particulier, il n'est pas possible de discriminer un comportement fugace d'un comportement non fugace pour un résidu (Figure VI.7). Arriver à distinguer ces deux situations présente un intérêt, entre autres, pour éviter que les comportements fugaces liés au temps de convergence d'une estimation réalisant un découplage indirect ne perturbent la décision ; cette situation peut être rencontrée pour le résidu r_{Ω}^{\perp} dans l'application à la MAS (Figure IV.21, Figure IV.23 et Figure V.16).

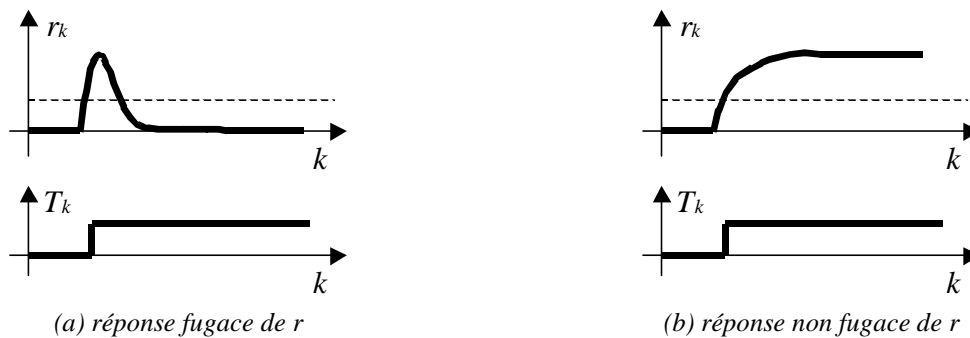


Figure VI.7 : Deux réponses types d'un résidu r à un défaut et valeur du test de dépassement de seuil avec mémorisation, T , résultant de (VI.3)

Une manière simple de décrire le comportement du résidu de la Figure VI.7a est la suivante : « r_k est initialement Z (zéro), puis il devient P (positif), puis il devient Z et reste Z ». De même, le comportement du résidu de la Figure VI.7b peut être décrit comme suit : « r_k est initialement Z , puis il devient P et reste P ». De manière plus synthétique, le comportement (a) correspond à la séquence ZPZ , alors que le comportement (b) correspond à la séquence ZP (Figure VI.7). La description qualitative du comportement du résidu, sous la forme de séquences, est suffisamment précise pour discriminer (a) et (b), puisque les séquences ZPZ et ZP sont distinctes. Nous allons donc chercher à calculer l'appartenance d'un résidu à une séquence et construire ainsi une décision dynamique.

Remarque : calculer l'appartenance d'un résidu à une séquence revient à évaluer le degré de cohérence du comportement réel du résidu, connu au travers de r_k , avec un ensemble de comportements (dynamiques) de référence décrits par la séquence.

VI.2.3.1. Le langage des expressions régulières

Pour décrire des séquences, le langage des expressions régulières peut être utilisé [David, 1992]. Un alphabet doit tout d'abord être défini : l'alphabet utilisé est l'ensemble des n_l termes linguistiques désignant les ensembles flous élémentaires qui servent à décrire le comportement des résidus $\Xi = \{L^l, l=1 \dots n_l\}$.

Exemple 1 : dans le cas où seuls les signes sont pris en compte et où les séquences ne font intervenir qu'un seul résidu¹, l'alphabet utilisé est $\{N, Z, P\}$.

Exemple 2 : les termes linguistiques désignant des ensembles flous associés à une direction dans l'espace des résidus (orientée ou non) peuvent également faire partie de l'alphabet utilisé. Ainsi, la séquence $\delta_1 \delta_2 \delta_3$ peut décrire le comportement d'un résidu vectoriel r : r est d'abord dans la direction δ_1 , puis dans la direction δ_2 et enfin dans la direction δ_3 .

\emptyset désigne la séquence nulle. Etant donné un alphabet, trois opérateurs sont définis (Table VI.1).

Opérateur (Notation)	Exemple	Signification
Concaténation (.)	$C=A.B=AB$	C a lieu si A a lieu puis B a lieu
Union (+)	$C=A+B$	C a lieu si A a lieu ou si B a lieu
Itération (*)	C^*	Répétition de C un nombre fini quelconque de fois ² .

Table VI.1 : Opérateurs dans le langage des expressions régulières

Pour tout triplet (A, B, C) d'expressions licites dans le langage considéré, les propriétés suivantes sont vérifiées :

$$A.(B+C)=AB+AC \quad (\text{VI.7})$$

$$(A+B).C=AC+BC \quad (\text{VI.8})$$

Exemple: partant de l'alphabet $\{N, Z, P\}$ et en remarquant que la réponse continue d'un résidu ne peut passer de N à P ou de P à N sans passer par Z , l'expression régulière $Z.((P+N).Z)^*.(Z+P+N)$ désigne l'ensemble des séquences possibles pour un résidu initialement nul. Parmi toutes ces séquences, celles de longueur inférieure ou égale à trois sont : Z, ZN, ZP, ZNZ, ZPZ . La propriété (VI.7), par exemple, permet de factoriser une expression régulière :

$$Z+ZN+ZP+ZNZ+ZPZ = Z.(Z+N).(Z+Z)+P.(Z+Z) \quad (\text{VI.9})$$

¹ Ce qui n'est pas une obligation : la séquence « résidu r_1 négatif puis résidu r_2 positif » peut très bien être définie.

² y compris zéro : $C^* = \emptyset + C + CC + CCC + \dots$

VI.2.3.2. Calcul de l'appartenance d'un résidu à une séquence

Dans le but de décrire l'appartenance à une séquence quelconque, les opérateurs de la Table VI.1 doivent être redéfinis dans le contexte de la logique floue. L'union de deux ensembles flous est une opération élémentaire qui peut être réalisée par n'importe quelle T-conorme (Figure V.7 et Table V.10). En se limitant à des séquences de faible longueur, l'opérateur d'itération n'est pas indispensable. Il reste donc à définir la concaténation de plusieurs ensembles flous. C'est à ce stade qu'est introduite une dynamique dans la décision, puisque la concaténation exprime une succession temporelle.

VI.2.3.2.1. Concaténation de plusieurs ensembles flous

Soit L^1, \dots, L^n les éléments de l'alphabet décrivant une séquence S de longueur n_e :

$$S = L^1 \dots L^e \dots L^{n_e} \tag{VI.10}$$

La séquence S est constituée de n_e étapes, chacune étant repérée par une valeur particulière de l'indice e ($e \in 1 \dots n_e$). A chaque étape de la séquence sont associés deux ensembles flous :

- Le premier, L^e , décrit la condition de **transition** pour passer de l'étape $(e-1)$ à l'étape e . L_k^e est l'appartenance à l'ensemble flou L^e calculée à l'instant k .
- Le second, X^e , décrit l'**état** d'avancement dans la séquence : l'appartenance à X^e à l'instant k , X_k^e , est d'autant plus forte que la condition L^1 puis L^2 puis ... puis L^e est satisfaite. X^e représente donc la concaténation des ensembles flous L^1 à L^e : $X_k^e = (L^1 \dots L^e)_k$.

Ces définitions étant données, notre objectif est de décrire l'évolution des états en fonction des transitions. La notion de succession temporelle asynchrone, sur laquelle reposent les séquences que nous avons définies, peut être traduite en ces termes : premièrement, il faut être arrivé à l'étape $(e-1)$ pour pouvoir passer à l'étape e ; d'où : $X_{k+1}^e = X_k^{e-1} \wedge \dots$. Deuxièmement, en supposant que l'étape $(e-1)$ a été complètement réalisée ($X_k^{e-1} = 1$), l'étape e est réalisée si et seulement si la condition de transition décrite par L^e a été remplie au moins une fois ; ce deuxième point correspond à une mémorisation similaire à celles de la partie VI.2.2.2. L'évolution des états en fonction des transitions est donc donnée par (VI.11).

$$X_{k+1}^e = (X_k^e \vee L_k^e) \wedge X_k^{e-1} \tag{VI.11}$$

avec comme conditions initiales: $X_k^0 = 1$, $X_0^1 = L_0^1$, $X_0^{e \neq 1} = 0$

(VI.11) se ramène très simplement à l'équation logique d'une bascule RS (floue) à R prioritaire :

$$\begin{array}{|c|c|} \hline S & Q \\ \hline R & \overline{Q} \\ \hline \end{array} \quad Q_{k+1} = (Q_k \vee S_k) \wedge \overline{R_k} \tag{VI.12}$$

Partant de ce constat, une représentation graphique du type « logique câblée » peut être utilisée pour représenter le système dynamique servant à calculer les appartenances aux ensembles flous qui résultent de la concaténation de plusieurs ensembles flous, $(L^1 \dots L^e)_k$, $e=1 \dots n_e$ (Figure VI.8).

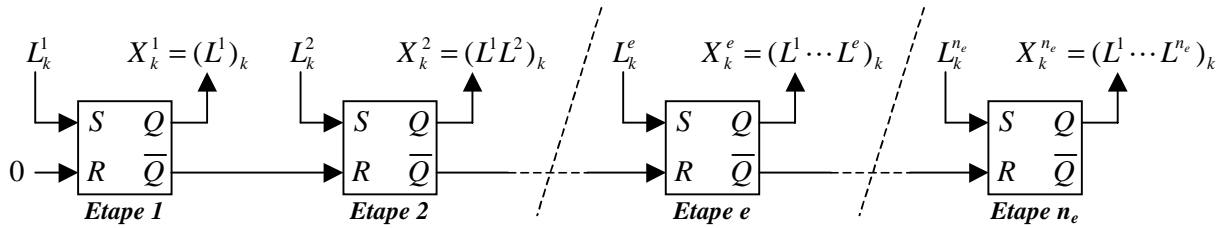


Figure VI.8 : Concaténation de plusieurs ensembles flous : représentation graphique des équations dynamiques

VI.2.3.2.2. Interprétation de la fin d'une séquence

L'appartenance S_k à la séquence S (VI.10) pourrait être définie par $S_k = (L^1 \dots L^{n_e})_k$. Cette définition ne permet pas de remplir les objectifs préalablement fixés qui consistent, entre autres, à distinguer un comportement fugace d'un comportement non fugace. En effet, S_k devient grand dès que les conditions L^1 puis L^2 puis ... puis L^{n_e} ont été satisfaites, et ce, indépendamment de ce qui se passe après. Par exemple, l'appartenance d'un résidu r à la séquence Z puis P , $(ZP)_k$, devient grande (et reste grande) dès que r a été successivement Z puis P ... et même s'il redevient Z . Par conséquent, choisir $S_k = (L^1 \dots L^{n_e})_k$ ne permet pas de discriminer un comportement fugace (ZPZ) d'un comportement non fugace (ZP) . La signification que l'on donne à la séquence S doit donc être précisée : « l'appartenance S_k à la séquence S (VI.10) est d'autant plus forte que L^1 puis L^2 puis ... puis L^{n_e} ont été successivement remplies et que L^{n_e} reste vérifié ». Une façon de garantir que L^{n_e} reste vérifié consiste à tester si une situation autre que L^{n_e} a été rencontrée ($\overline{L^{n_e}}$ avec $\overline{L_k^{n_e}} = 1 - L_k^{n_e}$) et d'annuler l'appartenance à S lorsque cette alternative se produit (VI.13).

$$S_k = (L^1 \dots L^e \dots L^{n_e})_k \wedge \overline{(L^1 \dots L^e \dots L^{n_e} \overline{L^{n_e}})_k} \quad (\text{VI.13})$$

L'utilisation de (VI.13) permet de garantir que $(ZP)_k$ retombe à zéro dès que r_k redevient Z (après avoir été Z puis P). Simultanément, $(ZPZ)_k$ se rapproche de un.

Remarque : une solution plus « élégante » consisterait à désactiver l'étape $(e-1)$ au fur et à mesure que l'étape e devient active. Un tel mécanisme d'activation/désactivation lié au franchissement de transitions est typiquement celui du grafcet (Figure VI.9). En logique classique, cette solution est sans doute la plus appropriée. Nous ne l'avons pas retenue car, dans un contexte de logique floue, les équations du grafcet ne se généralisent pas facilement : elles conduisent à des situations d'équilibre dans les appartenances aux étapes qui se révèlent préjudiciables en pratique. Le choix des opérateurs flous prend alors une importance considérable : les opérateurs de Lukasiewicz (Table V.10) sont alors bien adaptés à cette situation ; ils permettent en effet de désactiver complètement l'étape amont et d'activer complètement l'étape avale lorsque la transition est complètement validée (et franchissable). Une étude approfondie de la bibliographie concernant les systèmes flous du type « grafcet » ou « réseaux de Petri » nous semble donc utile pour continuer à développer l'approche proposée.

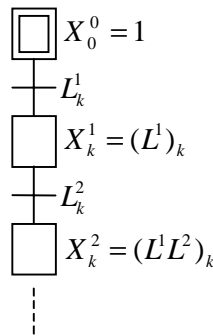


Figure VI.9 : Implantation sous la forme d'un grafcet

VI.2.3.2.3. Une difficulté liée à la représentation floue

Supposons qu'un résidu r , fuzzifié conformément à la Figure V.8, soit initialement nul, puis devienne progressivement positif et reste positif : la Table VI.2 montre l'évolution des appartenances à ZP et à ZPZ avec les opérateurs de Zadeh (Table V.10).

Z_k	P_k	$(ZP)_k$	$(ZPZ)_k$
1	0	0	0
0.7	0.3	0.3	0.3
0.5	0.5	0.5	0.5
0.3	0.7	0.7	0.5
0	1	1	0.5

Table VI.2 : Exemple d'évolution de l'appartenance aux séquences lorsque r passe de Z à P

La Table VI.2 montre que l'appartenance à $(ZPZ)_k$ devient non nulle alors que r est resté constamment croissant en partant de zéro. Ce phénomène s'explique par le recouvrement des ensembles flous Z et P . En effet, r peut être à la fois Z et P à un instant donné ; par exemple : $Z_k=0.7$ et $P_k=0.3$. Pour reconnaître correctement les séquences (ZP) et (ZPZ) , il faut donc arriver à différencier le passage de Z à P du passage de P à Z . Dans le premier cas, Z_k décroît en même temps que P_k croît, alors que les sens de variation sont inversés dans le deuxième cas. Pour résoudre le problème lié au recouvrement des ensembles flous, il suffit de choisir comme condition de transition de (ZP) à (ZPZ) : Z_k est fort et Z_k est croissant. La notion de croissance exprime la transition vers Z (et non depuis Z).

La solution adoptée dans l'exemple de la Table VI.2 se généralise en remplaçant les conditions de transition L_k^e , $e=1 \dots n_e$, de la partie VI.2.3.2.1 (voir Figure VI.8) par \tilde{L}_k^e :

$$\tilde{L}_k^e = L_k^e \wedge \mu_{\uparrow}(L_k^e) \tag{VI.14}$$

Où μ_{\uparrow} est l'indicateur de croissance d'un signal défini dans la partie VI.2.1 (Figure VI.3 et (VI.1)).

VI.2.3.2.4. Bilan sur le calcul de l'appartenance à une séquence

Finalement, le schéma logique décrivant le calcul de l'appartenance S_k à la séquence $S = L^1 \dots L^e \dots L^{n_e}$ se déduit de la Figure VI.8, en tenant compte des remarques qui ont conduit à poser (VI.13) et (VI.14) (Figure VI.10).

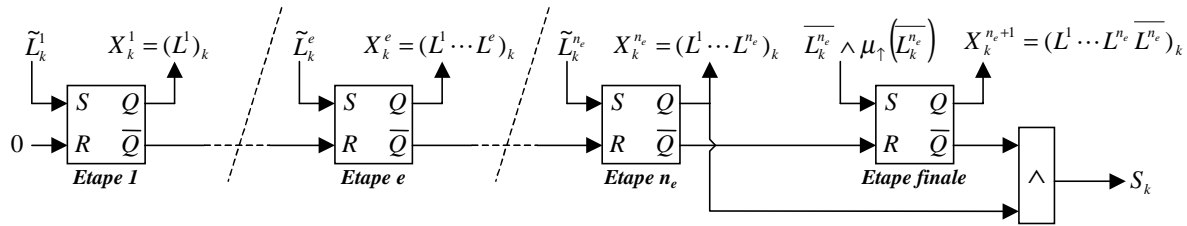


Figure VI.10 : Calcul de l'appartenance à une séquence

Remarque : si l'appartenance à la séquence S a été préalablement calculée, le calcul de l'appartenance à la séquence $SL = L^1 \dots L^{n_e} L$ ne nécessite pas de redéfinir entièrement une chaîne de bascules RS comme sur la Figure VI.10 : les bascules ayant servi au calcul de S peuvent en effet être réutilisées. Lorsque les appartenances à plusieurs séquences doivent être calculées, il est donc utile d'identifier les préfixes communs à plusieurs séquences. Cela peut être réalisé en factorisant l'expression régulière qui décrit les séquences à l'aide de la propriété (VI.7). L'exemple de factorisation donné par (VI.9) correspond au calcul de l'appartenance d'un résidu aux séquences Z, ZN, ZP, ZNZ et ZPZ .

Notation : $\mu_{(S)}(r)$ désigne l'appartenance de r à la séquence S ($\mu_{(S)}(r_k) = S_k$). Les termes linguistiques définissant la séquence S font alors référence à des ensembles flous qui décrivent le comportement de r .

VI.2.4. Application à la reconnaissance d'un comportement fugace

Un des objectifs à l'origine de l'étude des séquences était de distinguer un comportement fugace d'un comportement non fugace. Deux signaux correspondant à ces deux situations sont fuzzifiés de la même manière et les appartenances aux séquences $(Z), (ZP)$ et (ZPZ) sont calculées. La Figure VI.11 et la Figure VI.12 montrent les résultats obtenus. L'appartenance à la séquence (ZPZ) est pratiquement nulle pour le signal non fugace, alors qu'elle est forte pour le signal fugace. La discrimination des deux situations caractéristiques est donc réalisée. L'appartenance à (ZPZ) ne devient forte qu'avec un certain retard, et pour cause : il faut attendre que le signal « retombe » à zéro pour mettre en évidence la présence d'un comportement fugace.

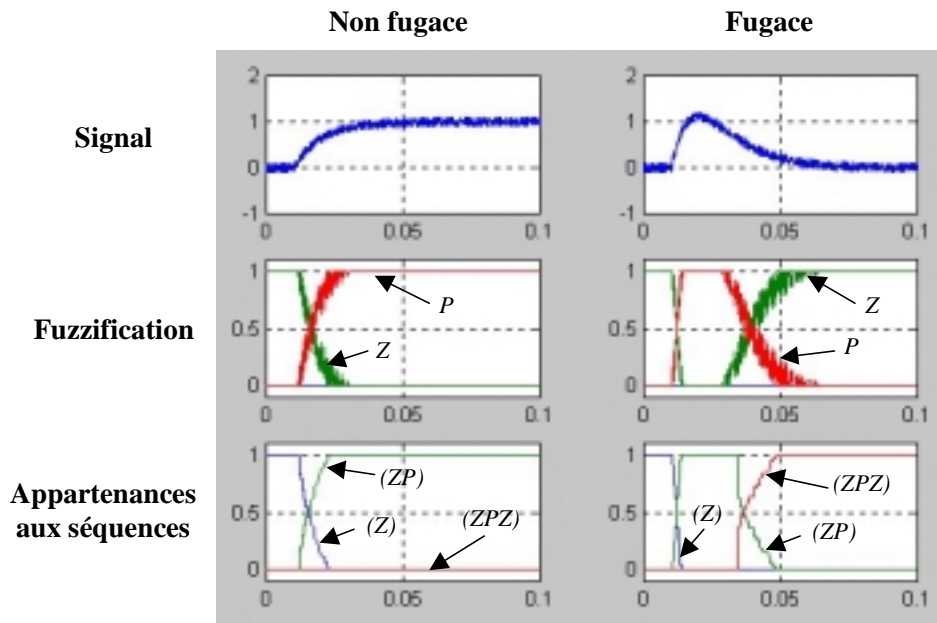


Figure VI.11 : Illustration des séquences floues sur deux signaux (Amplitudes fortes)

La Figure VI.12 montre les résultats obtenus avec deux signaux de faible amplitude : l'appartenance à P (resp. Z) n'atteint jamais la valeur 1 (resp. 0). La Figure VI.12 confirme la progressivité des décisions induites par la logique floue. L'appartenance d'un signal à une séquence représente bien la « ressemblance » plus au moins forte du comportement du signal aux comportements décrits par la séquence qualitative.

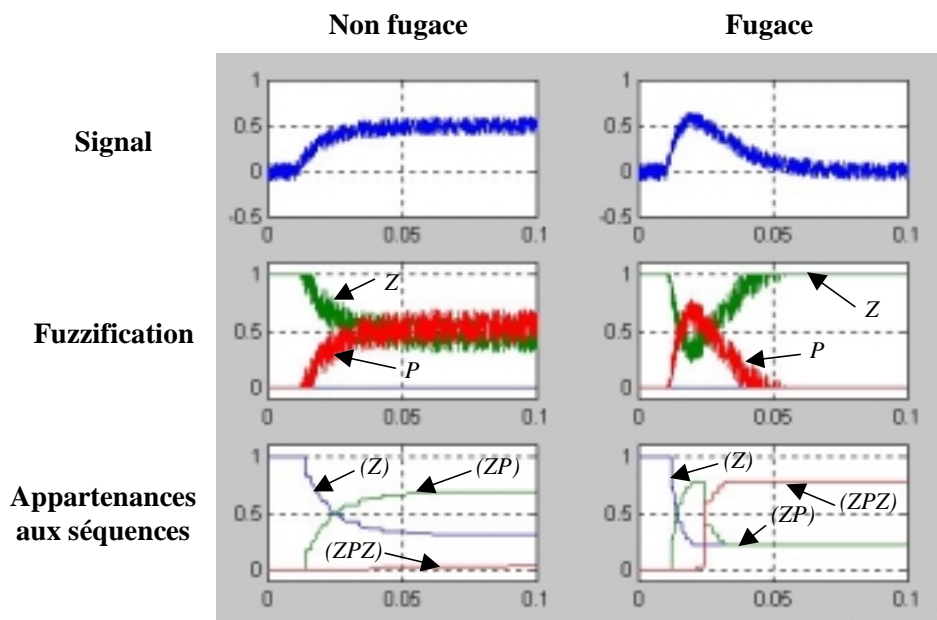


Figure VI.12 : Illustration des séquences floues sur deux signaux (Amplitudes faibles)

VI.2.5. Application à la MCC

Le comportement des résidus de la Figure VI.5 en réponse aux défauts peut être décrit par des séquences. Une table de décision est ainsi obtenue (Table VI.3).

	$f_\beta > 0$	$f_I > 0$	$f_\Omega > 0$	$f_U > 0$
$r_{\beta_m}^*$	ZPZ	ZN	Z	ZN
$r_{I_m}^*$	ZPZ	ZPZ	Z	ZPZ
$r_{\Omega_m}^*$	ZP	ZN	ZP	ZP

(a) défauts positifs

	$f_\beta < 0$	$f_I < 0$	$f_\Omega < 0$	$f_U < 0$
$r_{\beta_m}^*$	ZNZ	ZP	Z	ZP
$r_{I_m}^*$	ZNZ	ZNZ	Z	ZNZ
$r_{\Omega_m}^*$	ZN	ZP	ZN	ZN

(b) défauts négatifs

Table VI.3 : Table de décision décrivant le comportement des résidus par des séquences

Comme pour la décision statique avec prise en compte des signes (partie V.3.1.2), la Table VI.3 met en évidence la nécessité de distinguer deux hypothèses pour les défauts : défauts positifs (a) et défauts négatifs (b). Des fonctions de décision sont déduites de la Table VI.3 en considérant cette dernière comme une table de décision généralisée (partie V.3.1.3, Table V.12). Par exemple, la décision pour le défaut f_I découle de l'application de (V.27) et (V.28) :

$$\mu_{D(f_I > 0)} = \mu_{(ZN)}(r_{\beta_m}^*) \wedge \mu_{(ZPZ)}(r_{I_m}^*) \wedge \mu_{(ZN)}(r_{\Omega_m}^*) \quad (\text{VI.15})$$

$$\mu_{D(f_I < 0)} = \mu_{(ZP)}(r_{\beta_m}^*) \wedge \mu_{(ZNZ)}(r_{I_m}^*) \wedge \mu_{(ZP)}(r_{\Omega_m}^*) \quad (\text{VI.16})$$

$$\mu_{D(f_I)} = \mu_{D(f_I > 0)} \vee \mu_{D(f_I < 0)} \quad (\text{VI.17})$$

La Figure VI.13 montre les résultats obtenus. Les défauts sont correctement localisés. On peut vérifier que pour une amplitude supérieure du défaut f_I , l'appartenance à $D(f_I)$ monte jusqu'à un. Une étude détaillée des appartenances des résidus à chacune des séquences a montré que ces appartenances étaient conformes au comportement qualitatif des résidus. Seule $\mu_{(ZPZ)}(r_{I_m}^*)$ n'atteint pas la valeur un en réponse à f_I , ce qui s'explique par le comportement fugace de faible amplitude de $r_{I_m}^*$ et par le niveau de bruit qui impose de choisir une tolérance ν relativement élevée afin de détecter les phases de croissance de l'appartenance à P (Figure VI.5 et Figure VI.3).

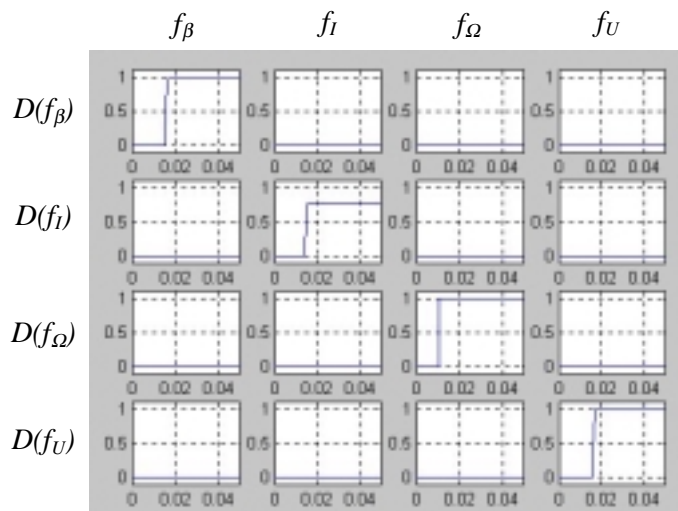


Figure VI.13 : Application à la MCC de la décision fondée sur les séquences floues

Remarque : d'un point de vue pratique, cette application à la MCC n'apporte rien de plus que la décision avec « effet mémoire » (Figure VI.6). Néanmoins, cette application illustre le fait que des comportements plus précis des résidus ont pu être reconnus dans les mêmes conditions que celles de la Figure VI.6. De manière générale, la possibilité de décrire et de reconnaître des comportements plus précis des résidus est un atout pour améliorer les capacités de localisation.

VI.2.6. Obtention des tables de décision à partir d'un générateur de résidus

VI.2.6.1. Formulation du problème

Dans l'application des séquences floues à la MCC, la table de décision (Table VI.3) a été obtenue à partir d'une analyse du comportement des résidus en simulation. Dans la mesure où cette table de décision décrit le comportement des résidus en réponse aux défauts, il semble légitime d'essayer de l'obtenir directement à partir du générateur de résidus et du modèle du système.

Lorsque le modèle du système et le générateur de résidus sont LTI (linéaires et invariants dans le temps), en négligeant les erreurs de modélisation et en supposant les perturbations parfaitement découplées, le résidu r est égal à l'influence des défauts sur le résidu ($r=r_f$). Pour des systèmes continus, l'opérateur δ est l'opérateur de Laplace, p , et (III.28) à (III.33) donnent (VI.18).

$$r = G(p).f \quad \text{avec} \quad G(p) = Q_y(p).H_f(p) \tag{VI.18}$$

$G(p)$ est une matrice de transfert qui décrit le comportement dynamique entre les défauts et les résidus. Cette matrice de transfert peut être vue comme une table de décision (Table VI.4).

$G(p)$...	f_j	...
...
r_i	...	$G_{ij}(p)$...
...

Table VI.4: Interprétation de $G(p)$ comme une table de décision

Si l'information contenue dans la table de décision $G(p)$ est très riche, elle n'en est pas moins délicate à exploiter, dans la mesure où la forme et l'amplitude des défauts sont a priori inconnues. Plusieurs approches sont possibles. Moyennant une hypothèse de défauts en échelon, Weber propose une localisation fondée sur une technique d'estimation paramétrique [Weber, 1999c]. L'approche que nous avons retenue consiste à décrire qualitativement la réponse des résidus à des défauts en échelon sous la forme d'une séquence. La perte de précision liée à une telle description qualitative trouve une contrepartie dans une dépendance moins forte à l'hypothèse de défauts en échelon. On retrouve ici le compromis entre précision et domaine de validité d'un modèle : si un modèle qualitatif est a priori peu précis par rapport à un modèle numérique, il permet par contre de décrire tout un ensemble de comportements similaires.

Pour chaque transfert $G_{ij}(p)$, on cherche donc à décrire la réponse indicielle de ce transfert à l'aide d'une séquence définie sur l'alphabet N, Z, P (négatif, zéro, positif).

VI.2.6.2. Description qualitative de la réponse indicielle d'un transfert

En supposant que $H(p)$ soit un transfert monovarié à pôles simples, réels et stables, il peut être décomposé en éléments simples sous la forme (VI.19).

$$H(p) = \sum_{m=0}^n \frac{g_m}{1 + \tau_m \cdot p} \quad \text{avec } \tau_0=0 \text{ et } \forall m, \tau_m < \tau_{m+1} \quad (\text{VI.19})$$

Les éléments simples associés à chacun des modes m sont classés par constantes de temps croissantes. Le mode $m=0$ peut être vu comme un mode dégénéré qui correspond à une constante de temps nulle : il subsiste donc un simple gain, g_0 . n est l'ordre du système. La réponse indicielle de $H(p)$ est la somme des réponses indicielles des transferts du premier ordre associés à chaque mode (Figure VI.14).

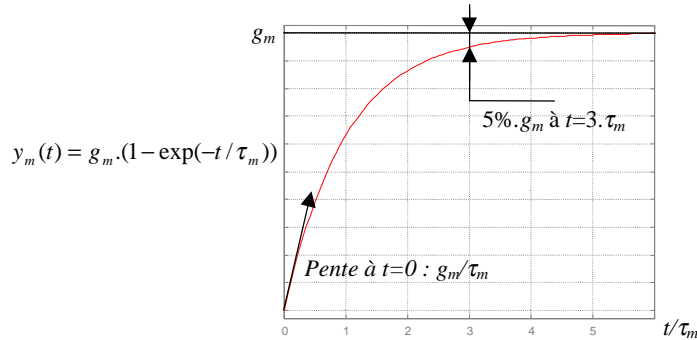


Figure VI.14 : Réponse indicielle d'un système du premier ordre

Soit s_k la somme des gains associés aux k premiers modes :

$$s_k = \sum_{m=0}^k g_m \quad (\text{VI.20})$$

Si toutes les constantes de temps étaient très différentes, $\forall m, \tau_m \ll \tau_{m+1}$, alors le mode m serait pratiquement en régime permanent vers $t=3\tau_m$ (de même que les modes d'indice inférieur à m) alors que la réponse aux modes d'indice supérieur à m serait presque nulle. Dans ces conditions, la séquence associée à la réponse indicielle de $H(p)$ serait : zéro puis signe¹ de s_0 puis signe de s_1 puis ... puis signe de s_n . Les situations dans lesquelles ce résultat s'applique étant rares en pratique, une solution consiste à regrouper les modes ayant des réponses qui ne peuvent pas être dissociées. Pour effectuer ces regroupements, un critère est défini :

$$d_m = \begin{cases} \left(|g_{m+1} \cdot (1 - \exp(-3\tau_m / \tau_{m+1}))| < \frac{\min(|s_m|, |g_{m+1}|)}{M} \right), & \text{pour } 0 \leq m < n \quad (M=10) \\ d_n = 1 \end{cases} \quad (\text{VI.21})$$

d_m vaut 1 si l'influence du mode $(m+1)$ est encore négligeable en absolu (comparaison à g_{m+1}) et en relatif (par rapport aux modes plus rapides : comparaison à s_m), alors que le mode m a pratiquement fini de répondre ($t=3\tau_m$). Sinon, d_m vaut 0. Lorsque le critère d_m est vérifié, la constante de temps associée au mode m est considérée comme négligeable par rapport à celle associée mode $(m+1)$:

¹ une valeur proche de zéro est assimilée Z.

$$d_m = 1 \Leftrightarrow \tau_m \ll \tau_{m+1} \quad (\text{VI.22})$$

Soit m_1 et m_2 deux entiers tels que les modes $m \in [m_1; m_2]$ soient non séparables au sens du critère (VI.21). m_1 est un entier tel que $m_1=0$ ou tel que $\tau_{m_1-1} \ll \tau_{m_1}$ ($d_{m_1-1} = 1$). m_2 est alors le plus petit entier supérieur ou égal à m_1 tel que $\tau_{m_2} \ll \tau_{m_2+1}$ ($d_{m_2} = 1$). La réponse indicielle (échelon unitaire) de $H(p)$ étant notée $y(t)$, on a :

$$\text{Si } \tau_{m_1-1} \ll \tau_{m_1} \leq \tau_{m_2} \ll \tau_{m_2+1} \text{ Alors } y(3\tau_{m_1-1}) \approx s_{m_1-1} \text{ et } y(3\tau_{m_2}) \approx s_{m_2} \quad (\text{VI.23})$$

Il reste encore à préciser l'évolution de $y(t)$ pour $t \in [3\tau_{m_1-1}; 3\tau_{m_2}]$. Deux cas sont facilement résolus :

1^{er} cas : si tous les g_m , $m \in [m_1; m_2]$, sont de même signe, $y(t)$ peut être considérée comme monotone pour $t \in [3\tau_{m_1-1}; 3\tau_{m_2}]$.

2^{ème} cas : si l'un des g_m , $m \in [m_1; m_2]$, est beaucoup plus grand¹ que tous les autres, son influence l'emporte. Les modes $[m_1; m_2] \setminus \{m\}$ peuvent être négligés et $y(t)$ peut être considérée comme monotone pour $t \in [3\tau_{m_1-1}; 3\tau_{m_2}]$.

Si aucun des deux cas n'est vérifié, une solution pour lever certaines ambiguïtés consiste à étudier la dérivée de l'influence des modes $m \in [m_1; m_2]$ sur $y(t)$. Soit $y_m(t)$ l'influence du mode m sur $y(t)$.

$$y_m(t) = g_m \cdot (1 - \exp(-t/\tau_m)) \quad (\text{VI.24})$$

$$y_{[m_1; m_2]}(t) = \sum_{m=m_1}^{m_2} y_m(t) \quad (\text{VI.25})$$

$$\dot{y}_{[m_1; m_2]}(t) = \sum_{m=m_1}^{m_2} \frac{g_m}{\tau_m} \cdot \exp(-t/\tau_m) \quad (\text{VI.26})$$

Comme $\tau_m < \tau_{m+1}$, on a : $\forall t > 0$, $\exp(-t/\tau_m) < \exp(-t/\tau_{m+1})$. Cela justifie la propriété (VI.27) :

$$\text{Si } \left| \frac{g_{m_2}}{\tau_{m_2}} \right| > \sum_{m=m_1}^{m_2-1} \left| \frac{g_m}{\tau_m} \right| \text{ Alors } \text{signe}(\dot{y}_{[m_1; m_2]}(t)) = \text{signe}(\dot{y}_{[m_1; m_2]}(0)) \quad (\text{VI.27})$$

Lorsque la condition de (VI.27) est vérifiée, $\dot{y}_{[m_1; m_2]}(t)$ est de signe constant, donc $y(t)$ peut être considérée comme monotone sur $t \in [3\tau_{m_1-1}; 3\tau_{m_2}]$. Ce résultat peut contribuer à lever l'ambiguïté sur le comportement de $y(t)$, lorsque les deux cas étudiés ci-dessus ne sont pas vérifiés. Si la condition de (VI.27) n'est pas vérifiée, l'ambiguïté n'est pas résolue. Cependant, $\dot{y}_{[m_1; m_2]}(0)$, calculable directement à partir de g_m et de τ_m , donne une indication sur le comportement de $y(t)$.

¹ pour l'application au modèle de la MCC, « beaucoup » plus grand a été réglé à 100 fois plus grand.

VI.2.6.3. Application aux résidus globaux de la MCC

La description qualitative de la réponse indicielle d'un transfert est appliquée à chaque transfert entre un défaut de la MCC et un résidu global obtenu en supposant le couple de charge parfaitement connu. Dans ces conditions, $G_{ij}(p)$ ((VI.18) et Table VI.4) correspond au transfert entre le défaut f_j et la mesure y_i dans le modèle de la MCC ((III.36) et (III.37)). Les pôles du modèle linéaire de la MCC sont simples, réels et stables, ce qui permet d'appliquer les résultats de la partie VI.2.6.2. La MCC est un système d'ordre $n=3$. Deux constantes de temps ont le même ordre de grandeur correspondant à des phénomènes électriques. La troisième constante de temps, beaucoup plus grande, peut être associée aux phénomènes mécaniques.

Pour un transfert $G_{ij}(p)$, chaque indication entre parenthèses présente les résultats obtenus pour un ensemble de modes non séparables $[m_1; m_2]$ (des plus rapides aux moins rapides). Si le 1^{er} cas ou le 2^{ème} cas est vérifié (voir partie VI.2.6.2), seule la valeur qualitative de la réponse indicielle à $t = 3\tau_{m_2}$ est donnée. Exemple : ($y=P$). Si aucun des deux cas n'est vérifié, les résultats sont présentés comme suit : ($dy(0)=P$, $dy=?$, $y=Z$) où $dy(0)$ représente la dérivée à l'origine ($\dot{y}_{[m_1; m_2]}(0)$) et où dy représente la dérivée à un instant quelconque ($\dot{y}_{[m_1; m_2]}(t)$). $dy=?$ si le signe de dy n'est pas nécessairement constant. Pour chaque transfert, la séquence qualitative déduite est donnée en indiquant entre crochets l'influence liée à la constante de temps mécanique (Table VI.5).

$G(p)$	$f_\beta > 0$	$f_I > 0$	$f_\Omega > 0$	$f_U > 0$
$r_{\beta_m} *$	($y=P$), ($dy=N$, $y=Z$), ($y=Z$) <i>Bilan: ZPZ[Z]</i>	($y=N$), ($y=N$), ($y=N$) <i>Bilan: ZN[N]</i>	($y=Z$) <i>Bilan: Z[Z]</i>	($dy=N$, $y=N$), ($y=N$) <i>Bilan: ZN[N]</i>
$r_{I_m} *$	($dy(0)=P$, $dy=?$, $y=Z$), ($y=Z$) <i>Bilan: ZPZ[Z] (?)</i>	($y=P$), ($dy=N$, $y=Z$), ($y=Z$) <i>Bilan: ZPZ[Z]</i>	($y=Z$) <i>Bilan: Z[Z]</i>	($dy(0)=P$, $dy=?$, $y=Z$), ($y=Z$) <i>Bilan: ZPZ[Z] (?)</i>
$r_{\Omega_m} *$	($dy=P$, $y=P$), ($y=Z$) <i>Bilan: ZP[Z]</i>	($y=N$) <i>Bilan: ZN[N]</i>	($y=P$) <i>Bilan: ZP[P]</i>	($dy=P$, $y=P$), ($y=Z$) <i>Bilan: ZP[Z]</i>

Table VI.5 : Obtention des séquences qualitatives à partir d'un générateur de résidus

Dans l'application à la MCC reportée sur la Table VI.5, les séquences, qui décrivent de manière qualitative le comportement des résidus globaux en réponse à des défauts en échelon (positif), ont été déduites à partir de raisonnements systématiques sur les gains et les constantes de temps qui résultent d'une décomposition en élément simple des transferts de $G(p)$. Cependant, certaines ambiguïtés n'ont pas pu être levées dans la réponse de $r_{I_m} *$ à f_β et à f_U : à chaque fois, la solution la plus simple a été privilégiée. En excluant les dynamiques lentes dues aux « effets mécaniques » (entre crochets), les séquences de la Table VI.5 correspondent bien à celles de la Table VI.3, déduites de la simulation représentée sur la Figure VI.5. Si l'on cherchait à reconnaître la séquence ZPZ pour $r_{\Omega_m} *$ en réponse à f_β (ou à f_U), la décision serait fortement retardée, puisqu'il faudrait attendre que la constante de temps mécanique ait complètement répondu. C'est pourquoi seules les dynamiques rapides sont prises en compte dans la Table VI.3 et dans les fonctions de décision associées.

VI.3. Décision : Apprentissage

VI.3.1. Introduction

L'obtention d'un modèle est un (sinon le) point clé pour l'étape de décision. Tous les modèles du comportement des résidus en réponse aux défauts que nous avons utilisés jusque là ont été établis par une approche cognitive : ainsi, les tables de décision utilisées résultaient soit d'une connaissance experte (prise en compte des signes, déduction des séquences en simulation), soit de l'étude du générateur de résidus (étude de sensibilité, déduction de séquences qualitatives à partir d'un transfert). Si l'utilisation d'une connaissance experte peut se révéler très utile, elle n'est pas toujours disponible ou suffisante, et l'étude systématique des propriétés d'un générateur de résidus repose bien souvent sur des hypothèses restrictives. Cependant, le modèle d'un système (dynamique) peut certes résulter d'une connaissance a priori (modèle de connaissance) mais aussi d'une procédure d'identification ou d'apprentissage (modèle de comportement). Notre objectif est donc de construire, par apprentissage, un modèle dynamique pour la décision.

VI.3.2. Choix de la topologie du réseau de neurones

D'un point de vue théorique, les réseaux de neurones permettent d'approcher n'importe quelle fonction continue. Ils ont donc été largement utilisés pour l'identification ou la commande de systèmes non linéaires. De nombreuses architectures de réseaux de neurones ont été développées [Hérault, 1994]. Un point capital lors de l'utilisation de réseaux de neurones consiste à choisir une structure appropriée à l'application traitée. Par structure, nous entendons ici le nombre de couches, le nombre de neurones par couche et, plus généralement, la manière dont les différents neurones constituant le réseau sont connectés entre eux. A titre d'exemple, la Figure VI.15 donne la structure « classique » d'un perceptron multicouche.

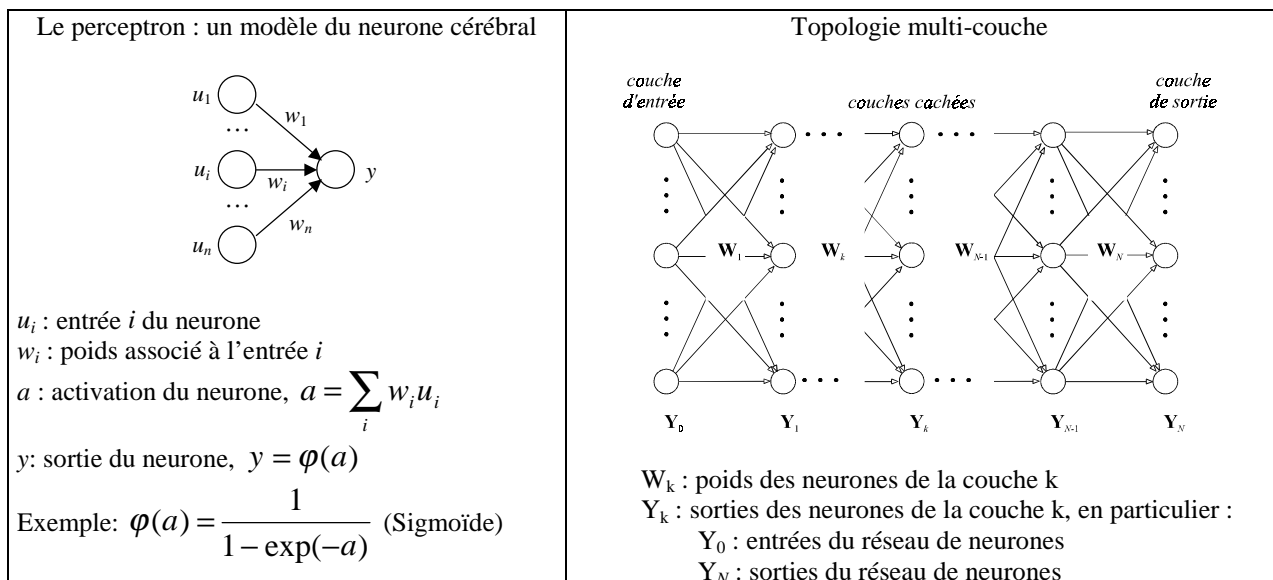


Figure VI.15 : Un exemple de réseau de neurones : le perceptron multicouches.

La topologie présentée sur la Figure VI.15 est celle d'un réseau multicouches unidirectionnel : les unités d'une couche ne dépendent en effet que des sorties de la couche précédente. Dans les réseaux récurrents, les unités d'une couche peuvent également dépendre des sorties des couches

suivantes (décalées dans le temps par un ou plusieurs retards). Ainsi, les réseaux unidirectionnels représentent des modèles non linéaires statiques, alors que les réseaux récurrents correspondent à des modèles dynamiques. Plus généralement, les topologies récurrentes imposent en contrepartie une complexité supplémentaire pour l'adaptation des poids du réseau [Pearlmutter, 1995]. Certaines topologies consistent à fixer à un le poids des connexions récurrentes. De plus, lorsque les caractéristiques dynamiques sont obtenues uniquement par l'utilisation, à l'entrée du réseau, de sorties décalées dans le temps, l'adaptation des poids peut reposer sur des algorithmes similaires à ceux utilisés dans le cas unidirectionnels. Les réseaux du type NARX¹ généralisent l'idée d'utiliser en entrée les sorties décalées dans le temps. Ces réseaux présentent de bonnes performances pour l'estimation de modèles dynamiques [Lin, 1996]. La forme générale d'un modèle NARX discret d'entrée $u(t)$ et de sortie $y(t)$ est donnée par (VI.28).

$$y(t) = \psi(y(t-1), \dots, y(t-n_y), u(t), u(t-1), \dots, u(t-n_u)) \quad (\text{VI.28})$$

n_u et n_y désignent respectivement les ordres de la récurrence sur les entrées et sur les sorties. ψ est une fonction non linéaire. Lorsque cette fonction est approchée par un perceptron multicouches (Figure VI.15), le système correspondant est un réseau de neurones du type NARX.

Il est largement reconnu que l'utilisation d'une connaissance a priori, si elle n'est pas nécessaire d'un point de vue théorique, se révèle néanmoins très précieuse pour choisir une architecture initiale du réseau adéquate. Par exemple, si l'information sur les signes des résidus semble un critère pertinent pour la décision, les valeurs d'appartenance aux ensembles flous N , Z , P pour chacun des résidus peuvent être utilisées comme entrées du réseau de neurones. Il s'agit là d'une association neuro-floue élémentaire. Les associations neuro-floues sont largement étudiées dans la littérature (voir [Leonhardt, 1997] pour une synthèse sur ces approches). L'intérêt porté aux associations neuro-floues s'explique par les potentialités de l'union des deux domaines : la capacité d'interprétation des systèmes flous et les possibilités d'apprentissage des réseaux de neurones.

VI.3.3. Application à la MCC

Un réseau de neurones NARX est utilisé pour calculer la décision du diagnostic à partir des résidus globaux fuzzifiés. L'importance de la prise en compte du passé a déjà été illustrée par la Figure VI.5. Une manière simple d'introduire une dynamique dans la décision consiste à choisir un réseau NARX d'ordre $n_u=0$ sur l'entrée et d'ordre $n_y=1$ sur la sortie (VI.28).

¹ NARX : Nonlinear AutoRegressive models with eXogenous inputs.

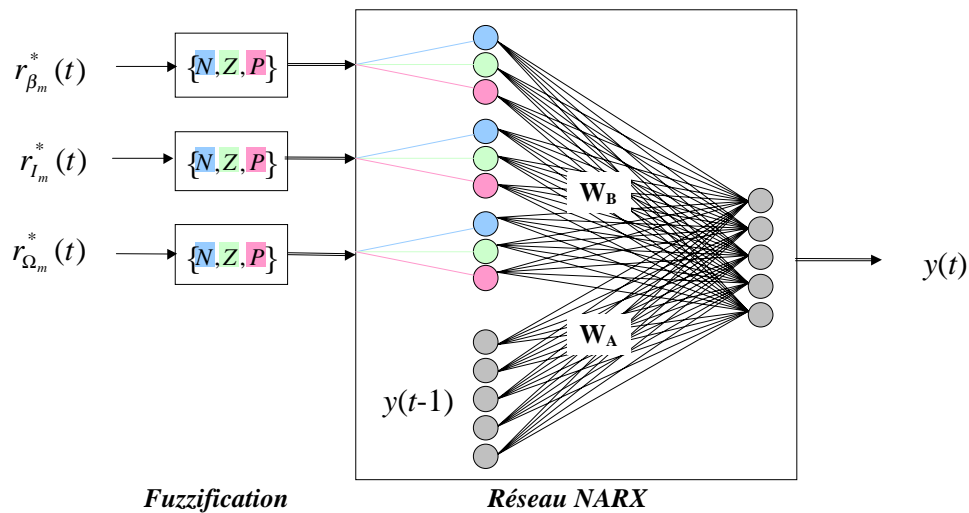


Figure VI.16 : Réseau NARX utilisé pour l'application à la MCC

Le réseau utilisé est constitué de deux couches (Figure VI.16) : la couche d'entrée contient les appartenances à N, Z, P de chacun des résidus globaux et la sortie retardée (VI.29). La couche de sortie contient les indicateurs de défaut représentant la décision du diagnostic : une sortie est affectée à un indicateur de détection et chacune des quatre autres sorties est affectée à un défaut particulier : $f_{\beta}, f_I, f_{\Omega}$ et f_U (VI.30).

$$u(t) = [\mu_N(r_{\beta_m}^*(t)) \ \mu_Z(r_{\beta_m}^*(t)) \ \mu_P(r_{\beta_m}^*(t)) \ \mu_N(r_{I_m}^*(t)) \ \mu_Z(r_{I_m}^*(t)) \ \mu_P(r_{I_m}^*(t)) \ \dots \ \mu_N(r_{\Omega_m}^*(t)) \ \mu_Z(r_{\Omega_m}^*(t)) \ \mu_P(r_{\Omega_m}^*(t))] \quad (VI.29)$$

$$y(t) = [\mu_{Normal}(t) \ \mu_{D(f_{\beta})}(t) \ \mu_{D(f_I)}(t) \ \mu_{D(f_{\Omega})}(t) \ \mu_{D(f_U)}(t)] \quad (VI.30)$$

Conformément à la Figure VI.16, la fonction non linéaire ψ (VI.28) est décrite par (VI.31) :

$$y(t) = \varphi(W_A \cdot y(t-1) + W_B \cdot u(t) + b) \quad (VI.31)$$

φ représente la fonction « sigmoïde » appliquée à chaque élément du vecteur d'entrée. b est un vecteur de poids représentant un biais sur les activités des neurones. W_B désigne les poids relatifs aux connexions entre les entrées et les activations des neurones de sortie. W_A désigne les poids relatifs aux connexions entre les sorties précédentes et les activations des neurones de sortie.

La décision à un instant donné tient donc compte de la description floue des résidus à cet instant et de la décision à l'instant précédent (prise en compte du passé).

Une base d'apprentissage a été réalisée pour l'adaptation des poids du réseau. Cette base est constituée de l'évolution temporelle de $[u(t) \ y_d(t)]$ pour différents scénarios de défaut. Des échelons d'une amplitude de 10% ont été considérés. Les sorties désirées, $y_d(t)$, sont obtenues par la réponse d'un système du premier ordre à un échelon unitaire pour la sortie associée au scénario de défaut et correspondent à une réponse nulle pour les sorties associées aux autres défauts. L'introduction d'une dynamique (rapide¹) sur les sorties désirées vise à faciliter l'apprentissage en évitant que le réseau apprenne « trop » $y(t)=y(t-1)$. La sortie désirée associée

¹ Cette dynamique reste de l'ordre de grandeur de la constante de temps électrique (trois fois plus rapide).

au comportement normal a été définie de sorte que la somme des sorties du réseau soit toujours égale à un.

L'adaptation des poids a été réalisée par l'algorithme RTBP (Real Time Back Propagation). Le mécanisme d'adaptation est détaillé dans [Evsukoff, 1998b]. Cet algorithme repose sur un calcul récursif du gradient¹ du critère par rapport aux poids de W_A . Les poids de W_B sont également adaptés, mais cela ne nécessite pas un calcul récursif du gradient, dans la mesure où ces poids ne sont pas associés à des connexions récurrentes.

A la fin de l'adaptation, les réponses des sorties associées à chaque défaut et obtenues par simulation du réseau sur les données d'apprentissage, sont montrées sur la Figure VI.17. La couleur gris clair représente l'état normal et la couleur gris foncé, l'occurrence d'un défaut.

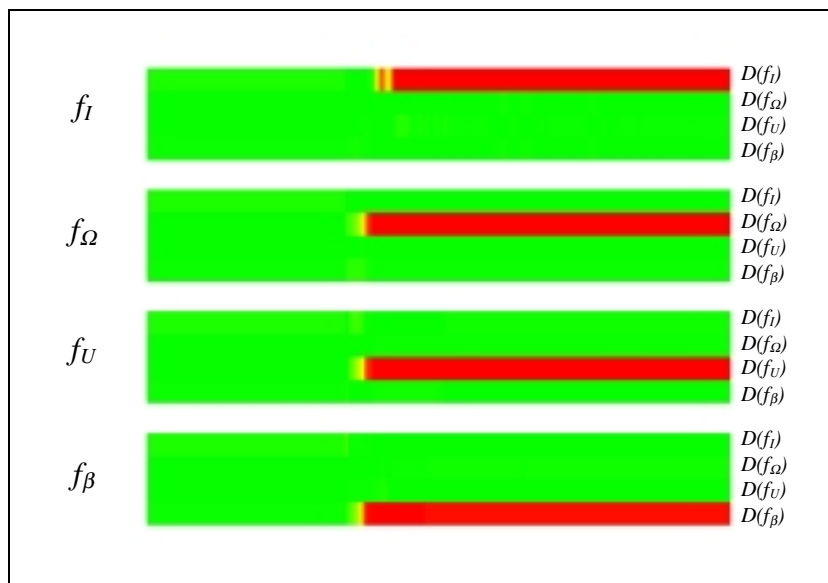


Figure VI.17 : Evolution temporelle des sorties du réseau NARX sur les données de la base d'apprentissage ($t \in [0; 0.03s]$)

La Figure VI.17 montre que la structure choisie pour le réseau de neurones est suffisante pour apprendre les décisions désirées sur les données de la base d'apprentissage. Le réseau NARX a ensuite été testé avec succès pour des défauts positifs de différentes amplitudes et de différentes formes. Par exemple, la Figure VI.18 représente les sorties du réseau pour des défauts positifs en forme de rampe qui atteignent une amplitude de 10% en une centaine d'échantillons (0.01s). Malgré un comportement des résidus sensiblement différent de celui des données d'apprentissage, la décision reste correcte, même si elle intervient avec un certain retard. Cela illustre les capacités de généralisation du réseau de neurones et justifie le choix de la structure récurrente la plus simple ($n_u=0$ et $n_y=1$ dans (VI.28)). Ce résultat est à rapprocher du fait qu'un simple « effet mémoire » suffit à surmonter les inconvénients d'une décision statique (Figure VI.6). Le réseau nous a permis de construire une décision dynamique tout en s'affranchissant de la connaissance d'une table de décision telle que la Table V.11. Dans l'application à la MCC,

¹ Nous avons utilisé le calcul récursif du gradient à deux reprises ; à chaque fois, l'objectif poursuivi était d'adapter les paramètres d'un modèle dynamique : méthode de découplage indirect proposée au chapitre IV ((IV.23) à (IV.25)), et adaptation des poids d'un réseau récurrent.

seuls les défauts positifs ont été étudiés. Nous pouvons néanmoins espérer des résultats similaires pour les défauts négatifs et combiner les décisions obtenues.

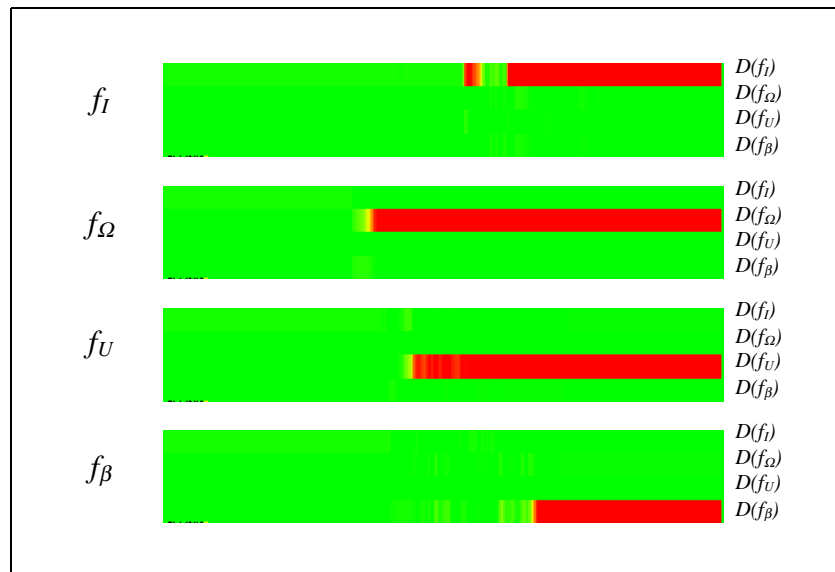


Figure VI.18 : Généralisation du réseau NARX pour des défauts en rampe ($t \in [0; 0.03s]$)

VI.4. Bilan

VI.4.1. Point de vue méthodologique

Partant du constat qu'une décision statique peut se révéler insuffisante, un outil générique est proposé pour construire une décision dynamique : les séquences floues. Cet outil permet de décrire le comportement temporel des indicateurs de défaut sous la forme d'une suite d'événements. Le caractère asynchrone de cette suite d'événements permet de décrire de façon similaire des comportements sensiblement différents, ce qui est utile dans la mesure où la forme des défauts n'est pas connue a priori. Cependant, rien n'empêche d'inclure des contraintes temporelles dans les transitions entre étapes au sein d'une séquence. De plus, les séquences étant un outil générique, il est tout à fait possible de décrire une séquence de directions pour un vecteur de résidus, ce qui revient à décrire l'évolution temporelle des ordres de grandeur relatifs des résidus. Le système dynamique représentant les fonctions de décision peut être ramené sous la forme classique d'une représentation d'état, illustrant ainsi le caractère séquentiel et non plus strictement combinatoire de la logique de décision.

$$\begin{aligned} x_{k+1} &= g(x_k, r_k) \\ D_k &= h(x_k, r_k) \end{aligned} \quad (\text{VI.32})$$

r_k, x_k, D_k représentent les résidus, l'état de la décision, et la décision à l'instant k . g et h désignent les équations d'état et de sortie de la décision ; ces équations peuvent être des équations logiques (logique séquentielle pour g , logique combinatoire pour h) mais cela n'a aucun caractère obligatoire. Par exemple, les calculs d'angle par rapport à une direction ne sont pas logiques mais numériques. Si la précision du modèle utilisé pour la décision est importante pour garantir

de bonnes capacités de localisation, le domaine de validité du modèle utilisé ne doit pas être négligé ; le paramétrage du modèle doit rester simple pour pouvoir intégrer facilement la connaissance d'un expert.

Lorsque la connaissance d'un expert est insuffisante, deux approches peuvent être envisagées : l'une repose sur l'exploitation des modèles numériques (modèle du système, générateur de résidus) pour déduire les modèles de décision. Ce lien a été établi sous des hypothèses restrictives dans la partie VI.2.6, où des séquences qualitatives ont été déduites de modèles numériques. Le problème de la modélisation de l'interface numérique/logique est de nouveau posé, mais dans un contexte dynamique et non plus statique comme au chapitre V. Une autre approche pour pallier un manque de connaissance a priori consiste à identifier un modèle à partir d'une base d'apprentissage. L'apprentissage d'un réseau de neurones NARX nous a ainsi permis d'apprendre une décision dynamique.

VI.4.2. Point de vue des applications

Ce chapitre est plus orienté sur les méthodes que sur les applications. Cependant, l'application de différentes approches pour la décision aux résidus globaux de la machine à courant continu (MCC) a montré qu'une exploitation dynamique et qualitative des résidus permettait de localiser les défauts considérés, avec des résidus correspondant pourtant à une table de signature non localisante. Dans les applications à la MCC, le couple de charge a été systématiquement supposé connu. En effet, les modèles pour la décision devant avoir un domaine de validité aussi large que possible, ce résultat est obtenu au prix d'une précision moindre par rapport à un modèle numérique. Or, il est nécessaire de décrire très précisément les comportements dynamiques pour espérer, sous certaines conditions (voir chapitre III), découpler le couple de charge. Ce découplage nous semble donc devoir être réalisé essentiellement au stade de la génération des résidus (si possible...) et non au stade de la décision. Cela justifie l'hypothèse « couple de charge connu » utilisée au cours de ce chapitre dans l'application à la MCC, et illustre l'intérêt de combiner des modèles numériques et logiques aux propriétés complémentaires.

Chapitre VII :

Conclusion

Chapitre VII :

Conclusion

VII.1. Bilan

VII.1.1. Point de vue méthodologique

Chapitre I. Le principe du diagnostic repose sur l'évaluation de la cohérence entre des observations et des modèles. Les observations apportent une information sur le comportement réel du système physique à diagnostiquer. Les modèles permettent de décrire des comportements de référence. Pour que l'évaluation de la cohérence entre observations et modèles fournisse des observations relatives aux seules anomalies, ce qui traduit un objectif de diagnostic, des modèles appropriés doivent être définis. A chacun de ces modèles correspond un module de diagnostic. Les modules de diagnostic peuvent être associés en parallèle et en série pour combiner leurs avantages respectifs. Les performances du diagnostic global tiennent alors au moins autant à la complémentarité des modules utilisés qu'aux performances des modules pris séparément. Dans ce cadre général, on peut distinguer les approches orientées bon et mauvais fonctionnement d'une part, les techniques de découplage direct et indirect d'autre part. Notre travail a consisté à étudier différentes méthodes pour construire les modules de diagnostic en fonction de l'information disponible et des objectifs poursuivis : détection, localisation voire identification des défauts.

Chapitre II. Partant d'une information relativement pauvre, l'existence de relations entre les variables d'un modèle, un ordonnancement causal permet tout d'abord de construire un schéma-bloc. L'orientation des arcs ainsi obtenue est à la base d'une méthode systématique de décomposition de modèle permettant de générer un premier jeu de résidus (résidus globaux et locaux). D'un point de vue méthodologique, l'intérêt d'une telle approche est de décomposer, de manière systématique, le problème de départ en plusieurs problèmes plus simples. L'utilisation de la seule information sur la structure du système rend la méthode indépendante de la nature du(des) modèle(s) (linéaires ou non, numériques ou qualitatifs,...). Par contre, les capacités de découplage des défauts sont fortement liées à la position et au nombre de capteurs. En outre, la contrepartie du caractère systématique de cette approche est l'impossibilité de régler la sensibilité des résidus aux défauts. Une fois la décomposition de modèle appliquée, il peut ainsi rester des points durs concernant la surveillance ; d'autres méthodes ont donc été étudiées.

Chapitre III. La nature et non plus seulement l'existence des relations entre variables sont alors prises en compte. Dans le cas de systèmes LTI (linéaires et invariants dans le temps) et lorsque

les défauts sont modélisés par des entrées additives, des méthodes classiques de génération de résidus, fondées sur des découplages parfaits ou approximatifs (observateurs à entrée inconnue, équations de parité, approche fréquentielle), ont été utilisées. Une étude de sensibilité a été menée dans l'optique d'enrichir l'information utilisée pour la synthèse des fonctions de décision. Lorsque les conditions d'application (restrictives) des méthodes linéaires de génération de résidus sont vérifiées, les possibilités de découplage sont meilleures que celles de la décomposition de modèle. La charge de calcul pour une implantation en ligne est également un point fort de ces approches. Cependant, la synthèse de ces générateurs de résidus s'avère délicate en pratique : de nombreux choix sont arbitraires (entrées à découpler pour structurer une table de signature, pôles du filtre ou de l'observateur qui ont une forte influence sur la réponse des résidus). De plus, les performances de ces méthodes (qui réalisent un découplage direct) peuvent se dégrader fortement lorsque les défauts ne sont plus abrupts, mais apparaissent sous la forme de dérives lentes.

Chapitre IV. Toutes ces restrictions nous ont conduit à explorer des méthodes de découplage indirect et à les combiner avec des méthodes de découplage direct (décomposition de modèle du chapitre II, projection orthogonalement à une direction variable dans le temps). L'idée directrice est de tester si les observations sont cohérentes avec l'un des comportements décrits par une famille de modèles. Une famille de modèles est décrite par un modèle dont certains paramètres¹ sont libres. Ces paramètres à découpler sont estimés dans le but de tester (localement) la cohérence avec les observations. En testant ainsi parallèlement plusieurs familles de modèles, il est ainsi possible de construire un diagnostic. L'implantation pratique est réalisée à l'aide d'un modèle (ou d'un observateur) adaptatif. L'adaptation d'un seul paramètre permet de lever des contraintes sur l'excitation du système, souvent incompatibles avec les objectifs de la commande. Une méthode basée sur un calcul récursif du gradient a été utilisée ; le gradient s'interprète comme une direction variable dans le temps et une projection orthogonalement à cette direction permet d'atténuer les transitoires liés à la convergence de l'estimation. Le découplage indirect a été appliqué aussi bien dans le cadre d'une approche orientée « bon fonctionnement » (détection par découplage indirect de perturbations paramétriques) que dans le cadre d'une approche orientée « mauvais fonctionnement » (localisation par découplage indirect de défauts paramétriques). Même si les contraintes en terme de modélisation sont non négligeables et même si la charge de calcul croît rapidement avec le nombre de tests, l'intérêt de ces méthodes est de permettre un diagnostic complet : détection, localisation et identification des défauts.

Chapitre V. La décision est un module de diagnostic construit à partir d'un modèle (par exemple, une table de décision) dont l'objectif est de détecter et de localiser les défauts, à partir de divers indicateurs (les résidus, dans notre travail). Il s'agit d'un problème de classification, pour lequel les critères correspondent aux résidus et les classes correspondent à différents scénarios de défaut. A partir de l'étude de sensibilité menée au chapitre III, des fonctions de décision satisfaisant à la fois un critère de sensibilité et un critère de robustesse (critère de localisation dans le contexte de la décision) ont été synthétisées par deux approches : l'une repose sur un algorithme dédié, très dépendant de la formulation du problème ; l'autre, plus coûteuse en calculs mais moins dépendante de la formulation du problème, repose sur une optimisation par algorithme génétique. Cette étude a mis en évidence un parallèle entre la démarche de génération de résidus et la démarche de synthèse de fonctions de décision, d'une

¹ Si les grandeurs « libres » décrivant une famille de modèles sont des entrées et non des paramètres, le découplage indirect devient un problème de commande et non plus un problème d'estimation paramétrique. Sous l'hypothèse d'une variation lente de ces entrées, une méthode d'estimation paramétrique peut néanmoins être appliquée.

part, l'importance de la modélisation des interfaces numérique/logique où apparaît la notion d'incertitude, d'autre part. La logique floue, qui conduit à un raisonnement graduel, a été utilisée afin de tenir compte des incertitudes, tout en évitant l'instabilité des décisions liées à l'utilisation de seuils nets. Par la prise en compte des signes (négatif, zéro, positif), une description qualitative du comportement des résidus a alors permis d'enrichir l'information utilisée pour la décision. L'importance du compromis entre précision et domaine de validité des modèles utilisés pour le diagnostic a également été soulignée. Une autre manière d'enrichir l'information utilisée pour la décision a consisté à prendre en compte les ordres de grandeur relatifs des résidus en testant l'angle d'un résidu vectoriel avec une direction donnée. L'étude des propriétés directionnelles des résidus a mis en évidence des liens étroits avec la méthode du gradient développée au chapitre IV. Le gradient s'interprète en effet comme une direction variable dans le temps et la projection orthogonalement au gradient est analogue à un test de direction.

Chapitre VI. Partant du constat qu'une décision statique peut se révéler insuffisante, plusieurs méthodes pour construire une décision dynamique (décision tenant compte du passé) ont été proposées. L'une, élémentaire, consiste à mémoriser les dépassements de seuils. Une autre, plus générale, repose sur la description du comportement temporel des résidus par des séquences (suites asynchrones d'événements) ; les séquences permettent d'enrichir considérablement le modèle utilisé pour la décision. La reconnaissance de séquences floues a été traitée et conduit à la mise en œuvre d'une logique à la fois séquentielle et floue. Dans le but d'automatiser la synthèse de la décision, d'autant plus délicate que les comportements doivent être décrits avec plus de précision, une méthode d'apprentissage a été étudiée. L'originalité de cette approche est de combiner une méthode à base de modèle pour générer les résidus et une méthode d'apprentissage basée sur des critères qualitatifs. L'intérêt est de réduire la taille et, par conséquent, la difficulté d'apprentissage du réseau de neurones utilisé, tout en préservant de bonnes capacités de généralisation. La topologie récurrente du réseau de neurones NARX a permis d'apprendre une décision non pas statique mais dynamique.

VII.1.2. Point de vue des applications

Machine à courant continu (MCC). Un entraînement commandé en courant a été considéré. Il peut être utilisé tel quel dans les applications de contrôle de couple ou avec une boucle externe dans le cas d'un contrôle de vitesse. Les défauts étudiés étaient les suivants : défaut de l'interface de commande, défaut du capteur de courant, défaut du capteur de vitesse et défaut de l'ensemble constitué de l'alimentation et du convertisseur. En supposant le couple de charge connu, plusieurs approches peuvent être utilisées pour détecter et localiser ces défauts : décomposition de modèle (résidus locaux et globaux), méthodes de génération de résidus appliquées aux systèmes linéaires (observateurs à entrée inconnue, équations de parité, approche fréquentielle), découplage indirect par filtrage de Kalman, décision qualitative avec effet mémoire (sur des résidus a priori non localisants, les résidus globaux), décision dynamique résultant de l'apprentissage d'un réseau de neurones récurrent. Pour l'application à la MCC, la principale difficulté consiste à garantir la robustesse aux variations du couple de charge. Dans ce contexte, les meilleurs résultats ont été obtenus par l'application des méthodes de découplage approximatif (Chapitre III). Cependant, l'étude du modèle de la MCC montre que l'atténuation de l'influence du couple de charge revient à filtrer les basses fréquences : par conséquent, il faut choisir entre un diagnostic robuste au couple de charge et un diagnostic sensible aux défauts en dérives lentes. La seule manière de satisfaire simultanément ces deux contraintes consiste à mesurer le couple résistant directement (capteur) ou indirectement (modélisation de la charge).

Machine asynchrone (MAS). Un entraînement asynchrone à vitesse variable a été considéré. Les défauts étudiés étaient les suivants : défauts des deux capteurs de courant (f_i), six défauts du

type « interrupteur ouvert » dans l'onduleur (f_V), défaut du capteur de vitesse (f_Ω), autres défauts (défauts du moteur : f_P). L'objectif consistant à établir un diagnostic robuste à la fois au couple de charge et à des variations de paramètres dues à l'échauffement a été en grande partie atteint. Le diagnostic global résulte de la combinaison de plusieurs approches. L'application de la décomposition de modèle permet de garantir la robustesse au couple de charge et, plus généralement, de découpler le diagnostic des effets mécaniques. Contrairement à la MCC, la mesure de deux courants permet de raisonner sur leurs ordres de grandeur relatifs¹ et, par conséquent, d'établir un diagnostic sans tenir compte de l'équation mécanique. De plus, l'excitation sinusoïdale de la MAS permet de découpler, sauf à très basse vitesse, les défauts additifs (f_I, f_V) induisant une composante continue sur les résidus (excitation induite par le défaut lui-même), des défauts paramétriques (f_Ω, f_P), qui n'excitent les résidus qu'au travers de l'excitation de l'entrée. Même si le modèle de la MAS rend difficile l'application des méthodes usuelles de génération de résidus (chapitre III), les particularités liées à l'excitation facilitent certains découplages. A l'exception des influences de f_I et de f_V , « tout tourne avec la même pulsation dans la MAS » ; dès lors, les grandeurs pertinentes pour réaliser des découplages ne sont plus nécessairement des gains fréquentiels mais des déphasages (angles, directions). L'analyse des propriétés directionnelles des résidus et la prise en compte des ordres de grandeur relatifs dans la décision ont permis de localiser f_I et f_V . Une méthode de découplage indirect, basée sur un algorithme du gradient, a été appliquée de deux façons : une approche orientée « bon fonctionnement » réalise le découplage des variations de paramètres dues à l'échauffement (un modèle de ces variations de paramètres est supposé connu) et permet ainsi de remplir les objectifs de la détection ; une approche orientée « mauvais fonctionnement » permet de distinguer le défaut f_Ω des autres défauts paramétriques (f_P). L'interprétation du comportement des résidus issus d'un découplage indirect sert à construire des fonctions de décision permettant de localiser l'ensemble des défauts considérés. Les résultats présentés dans ce travail sont obtenus en simulation pour des scénarios simples (apparition des défauts lors d'un fonctionnement à vitesse et à couple constants) ou complexes (apparition des défauts lors d'un fonctionnement à vitesse et à couple variables). Sous l'hypothèse de défaut simple, les défauts f_I du type « offset », les défauts f_V et f_Ω sont non seulement détectés et localisés mais également identifiés. Les autres défauts (f_P) sont détectés et localisés par rapport à f_I, f_V et f_Ω .

VII.2. Perspectives

VII.2.1. Point de vue des applications

Pour l'application à la MCC, la principale difficulté concerne le réglage du compromis entre la sensibilité aux défauts et la robustesse à la perturbation liée au couple de charge.

Pour l'application à la MAS, le diagnostic de défauts abrupts de forte amplitude sur le capteur de vitesse met en évidence la nécessité de travailler sur la commande pour qu'elle tolère ce type de défauts. L'estimation obtenue par l'intermédiaire du découplage indirect pourrait être utilisée pour compenser l'influence du défaut, mais il faut alors garantir que le tout converge. Le découplage indirect du défaut du capteur de vitesse a été réalisé en assimilant ce dernier à un défaut paramétrique. Or, l'hypothèse de variations lentes d'un paramètre n'est plus valide pour

¹ Plus précisément, les ordres de grandeur relatifs pris en compte sont ceux des résidus calculés à partir des équations électriques (résidus homogènes aux courants).

des défauts abrupts et de forte amplitude. Le découplage indirect devient alors un problème de commande non linéaire et non plus seulement d'estimation paramétrique.

Les défauts des capteurs de courant que nous avons considérés pour la MAS sont du type « offset ». Ils se traduisent par un déplacement du vecteur des résidus dans une direction fixe et par l'apparition d'une composante continue. Les courants étant sinusoïdaux, un défaut sur le gain d'un capteur ne se traduit plus par une composante continue, mais la direction fixe du défaut du type « offset » reste valable. En distinguant les défauts à direction fixe et à direction variable, la localisation reste a priori possible mais requiert des développements spécifiques.

Les résidus considérés pour le diagnostic de la MAS sont du type « erreur de sortie » des équations électriques. Rien n'empêche a priori d'utiliser une erreur d'équation. Le découplage indirect ne serait alors plus réalisé à l'aide d'un modèle adaptatif mais à l'aide d'un observateur adaptatif. Une question délicate concerne le réglage le plus approprié pour un tel observateur : un réglage trop lent peut nuire au suivi du système par l'observateur, un réglage trop rapide risque de conduire au report systématique de l'erreur sur l'estimation de l'état et non plus sur l'adaptation des paramètres.

Un important travail de modélisation des défauts du moteur asynchrone est nécessaire pour réaliser leur diagnostic en ligne (courts-circuits de spires, ruptures de barre,...). En particulier, une question, qui reste ouverte, concerne l'utilisation du modèle de Park comme base pour l'élaboration de modèles des défauts du moteur.

S'il fallait retenir une perspective du point de vue des applications, ce serait certainement une validation expérimentale des résultats obtenus. Pour l'application à la MAS, la principale difficulté à lever nous semble provenir de la compensation des décalages temporels dus aux temps de garde dans l'onduleur. Il s'agit d'un problème qui doit également être traité lors de l'implantation d'un observateur, dans le contexte de la commande. Même si la validation expérimentale peut être réalisée dans un premier temps avec des mesures des tensions de phase, le transfert de l'expérience acquise pour la commande semble possible. Il faudra alors évaluer les répercussions sur les sensibilités du diagnostic aux différents défauts.

VII.2.2. Point de vue méthodologique

VII.2.2.1. Modélisation

L'obtention d'un ou de plusieurs modèles est un point central en automatique. Du point de vue du diagnostic, trois types d'objectifs nécessitant l'élaboration de modèles peuvent être distingués :

- Modélisation du système physique, y compris en présence de défauts, pour valider en simulation le système de diagnostic. Ces modèles peuvent être relativement complexes ; l'objectif poursuivi est d'obtenir une image aussi fidèle que possible des observations réelles (Exemple : modélisation de l'onduleur par des « interrupteurs » qui commutent)
- Modélisation du système physique pour la synthèse des modules de diagnostic. Ces modèles doivent être suffisamment précis pour permettre la localisation des défauts et rester suffisamment simples, notamment dans le contexte du diagnostic en ligne (Exemples : modélisation de l'onduleur en défaut par des entrées additives, tables de décision)
- Modélisation associée aux modules de diagnostic. Les générateurs de résidus, les fonctions de décision sont des modèles dont la simulation permet de calculer le résultat du diagnostic. Ils

résultent plus ou moins directement de l'application de techniques de découplage aux modèles du système physique.

Le premier point relève des techniques classiques de modélisation ; le troisième point fait appel aux différentes techniques de diagnostic. Il nous semble que le deuxième point, peu abordé en dehors du contexte lié à une application particulière, pourrait utilement faire l'objet de développements méthodologiques. Il s'agirait de répondre à la question suivante : comment recueillir une information pertinente pour modéliser l'influence des défauts, non avec une précision maximale, mais avec une précision minimale pour pouvoir les localiser ?

VII.2.2.2. Améliorer le compromis entre sensibilité et robustesse

Indépendamment de la nature des modèles utilisés (numériques, logiques, ...), le diagnostic est avant tout un problème consistant à réaliser de multiples découplages. Il s'agit donc de construire des grandeurs sensibles à certaines influences et insensibles (ou robustes) à d'autres. Toute possibilité de découplage supplémentaire ou, plus généralement, toute amélioration du compromis entre sensibilité et robustesse se répercute donc directement sur les performances du diagnostic. Plusieurs perspectives s'inscrivent dans ce contexte :

- L'information structurelle n'a été utilisée au chapitre II que dans le cadre d'une approche orientée « bon fonctionnement ». La construction de modèles de « mauvais fonctionnement » peut également être envisagée, par exemple en calculant de deux façons différentes la variable modélisant un défaut. Les possibilités de découplage supplémentaires sont alors à étudier.
- Dans le cadre des méthodes de découplage approximatif, le compromis entre sensibilité et robustesse est souvent défini par un rapport entre deux gains (rapport de normes infinies, par exemple). Au regard de l'application à la MAS, on peut se demander si ce type de critères est toujours bien adapté. La principale difficulté n'est pas forcément de proposer un critère, mais surtout de résoudre le problème d'optimisation pour un critère donné...
- Dans l'application du découplage indirect à la MAS au chapitre IV, des surfaces représentant les « déformations » du transfert du système, par rapport à certains paramètres, ont été étudiées. La possibilité de « modeler » ces surfaces constitue une approche pour améliorer le compromis entre sensibilité et robustesse.
- Le critère de localisation utilisé pour la synthèse de fonctions de décision au chapitre V est en tout ou rien. En le redéfinissant de sorte qu'il soit d'autant plus grand que la robustesse de la localisation est grande (voir remarque dans la partie V.2.4.2.1), la pondération des critères utilisés par l'algorithme génétique (V.20) donne un moyen pour agir sur le compromis entre sensibilité et robustesse lors de la synthèse des fonctions de décision.

VII.2.2.3. Utiliser au mieux l'information disponible

Pour l'application du découplage indirect, nous nous sommes limités à l'estimation d'un seul paramètre. Cela nous a permis de nous affranchir des contraintes liées à l'excitation du système. De manière générale, lorsque n paramètres doivent être découplés indirectement, ces n paramètres doivent être estimés. L'excitation naturelle du système peut se révéler insuffisante pour satisfaire cet objectif. Ajouter une excitation supplémentaire est une solution possible mais conduit bien souvent à une remise en cause des spécifications de la commande. Une autre approche consisterait à n'adapter les paramètres qu'en fonction de l'excitation disponible (en privilégiant l'adaptation de certains paramètres par rapport à d'autres, par exemple). Une prise en compte plus fine des propriétés directionnelles des résidus est une piste pour atteindre cet objectif.

Utiliser au mieux l'information disponible peut également se traduire par une modélisation des incertitudes. Ce point a été mis en évidence lors de l'interprétation des dépassements de seuil au chapitre V. La modélisation des incertitudes dans les interfaces numérique/logique et leur prise en compte dans la synthèse des fonctions de décision nous semble être un champ d'étude à approfondir.

VII.2.2.4. Intégration du diagnostic à la supervision et à la maintenance

En faisant abstraction des difficultés inhérentes à l'utilisation de la logique floue, les séquences utilisées au chapitre VI peuvent être reconnues à l'aide de grafjets ou de réseaux de Petri. L'utilisation de ces outils, pour construire la décision du diagnostic d'un composant (entraînement électrique, par exemple), peut être intéressante car elle conduit à une représentation homogène avec celles classiquement employées pour la commande et la supervision des systèmes de production. Cela permet en effet d'envisager une intégration des diagnostics de bas niveau (composants) avec les diagnostics de niveau supérieur (ensembles de composants, systèmes de production).

Un problème finalement peu traité dans la littérature est celui de la reconfiguration d'un système de diagnostic, une fois qu'un défaut simple a été détecté et localisé, dans le but d'en diagnostiquer un nouveau. Le diagnostic de défauts successifs est nettement moins contraignant que celui de défauts multiples, tout en étant d'un intérêt pratique certain : pour de nombreux systèmes (en particulier, les entraînements électriques), il n'est pas raisonnable d'envisager l'intervention d'un opérateur pour effectuer une réinitialisation dès qu'un capteur a quelques états d'âme ! L'intérêt du découplage indirect est de suivre constamment l'état du système, y compris en situation de défaut : ainsi, lorsqu'un défaut a été détecté et localisé, l'état du générateur de résidus ayant réussi à expliquer le comportement du système peut être utilisé pour « recalculer » les autres générateurs de résidus.

Un système de diagnostic capable de construire un historique des défauts peut alors être envisagé. L'exploitation de tels historiques donnerait une information utile à la définition d'une politique de maintenance, dans laquelle le côté prédictif pourrait tenir une place importante. Ne dit-on pas souvent, avec sagesse, qu'il vaut mieux prévenir que guérir ?

Bibliographie

Bibliographie

- [**Alcorta Garcia, 1997**] : Alcorta Garcia E. and P.M. Frank, Deterministic Nonlinear Observer-based Approaches to Fault Diagnosis: A Survey. *Control Engineering Practice*, Vol. 5, No. 5, pp. 663-670, 1997.
- * [**Alexandru, 2000**] : Alexandru M., Combastel C., Gentil S., Diagnostic Decision using Recurrent Neural Networks. *IFAC Safeprocess'2000*, Budapest, Hungary, June 14-16, 2000.
- [**Amann, 1997**] : Amann P., Frank P.M., On fuzzy model-building in observers for fault diagnosis. *15th IMACS World Congress on Scientific Computation, Modelling and Applied Mathematics*, Berlin, August 24 - 29, 1997.
- [**Basseville, 1988**] : Basseville M., Detecting changes in signals and systems – a survey. *Automatica*, Vol. 24, No. 3, pp. 309-326, 1988.
- [**Beard, 1971**] : Beard R.V., Failure accommodation in linear systems through self reorganization. PhD thesis, Department of Aeronautics and Astronautics, MIT, Cambridge, Massachusetts, USA, 1971.
- [**Berendsen, 1993**] : Berendsen C.S., Méthodes de diagnostic dans la partie commande des associations convertisseurs-machines tournantes. Thèse de doctorat, Institut National Polytechnique de Grenoble, 1993.
- [**Berge, 1957**] : Berge C., Two theorems in graph theory. *Proceedings of the National Academy of Science*, Vol. 43, pp. 842-844, 1957.
- [**Bibas, 1997**] : Bibas S., GASPAR : simulateur à base de modèles pour la supervision de réseaux de télécommunication. Thèse de Doctorat de l'Université Paris Nord , 1997.
- [**Borras, 1999**] : Borras D., Castilla M., Moreno N., Montano J.C., Wavelet and neural structure: a new tool for diagnostic of power system disturbances. *IEEE SDEMPED'99, International Symposium on Diagnostics for Electrical Machines, Power Electronics and Drives*, pp. 375-380, Gijón, Spain, Sept. 1-3, 1999.
- [**Cameron, 1986**] : Cameron J.R., Thomson W.T., Vibration and Current Monitoring for Detecting Airgap Eccentricity in Large Induction Motors. *IEEE Transactions on Energy Conversion*, Vol. 10, pp. 603-614, 1986.
- [**Cassar, 1994a**] : Cassar J.P., Litwak R.G., Cocquempot V., Staroswiecki M., Approche structurelle de la conception de systèmes de surveillance pour des procédés industriels complexes. *Revue Européenne Diagnostic et Sûreté de Fonctionnement*, Vol. 2, pp. 179-202, 1994.

- [**Cassar, 1994b**] : Cassar J.P., Staroswiecki M., Advanced Design of the Decision Procedure in Failure Detection and Isolation Systems. IFAC Safeprocess'1994, Espoo, Finland, 1994.
- [**Cassar, 1996**] : Cassar J.P., Staroswiecki M., Pour une approche unifiée de la surveillance. Ecole d'été d'Automatique de Grenoble: surveillance des systèmes continus, Tome 1, 1996.
- [**Chang, 1991**] : Chang J. S., Di Cesare F., Goldbogen G., Failure propagation trees for diagnosis in manufacturing systems. IEEE Transactions on Systems, Man, and Cybernetics, Vol. 21, pp. 767-776, 1991.
- [**Chen, 1995**] : Chen J., Patton R.J., Zhang H.Y., Design of Robust Structured and Directional Residuals for Fault Isolation via Unknown Input Observers. Proceedings of 3rd European Control Conference, Rome, Italy, September 1995.
- [**Chen, 1996**] : Chen J., Patton R.J., Zhang, H.Y., Design of unknown input observers and robust fault detection filters. International Journal of Control, Vol. 63, No. 1, pp. 85-105, 1996.
- [**Chow, 1984**] : Chow E.Y., Willsky A.S., Analytical redundancy and the design of robust failure detection systems. IEEE transactions on Automatic Control, Vol. 29, pp. 603-614, 1984.
- [**Cocquempot, 1993**] : Cocquempot V., Surveillance des processus industriels complexes: génération et optimisation des relations de redondance analytique. Thèse de Doctorat de l'Université des Sciences et Technologies de Lille. Villeneuve d'Ascq, France, 1993.
- * [**Cocquempot, 2000**] : Cocquempot V., *et al*, Projet "Méthodes de Redondance Analytique pour la Surveillance de la Machine Asynchrone", synthèse rédigée par V. Cocquempot (LAIL), à partir des contributions des équipes du CEGELY, du LAG, du LAIL, du LAP, et du LEG. Journées SDSE (Sûreté et Disponibilité des Systèmes Electrotechniques), Villeurbanne, 20-21 janvier 2000.
- * [**Combastel, 1997a**] : Combastel C., Application des Méthodes d'Aide à la Décision Floue à la Détection de Défaillances dans un Entraînement Electrique. DEA Automatique-Productique, LAG/INPG, Grenoble, France, Septembre 1997.
- * [**Combastel, 1997b**] : Combastel C., Gentil S., Rognon J.-P., Application des Méthodes d'Aide à la Décision Floue à la Détection de Défaillances dans un Entraînement Electrique. IAR Workshop Fuzzy and Diagnostic 97, pp19-24, Mulhouse, France, 24 oct. 1997.
- * [**Combastel, 1998**] : Combastel C., Gentil S., Rognon J.-P., A Symbolic Reasoning Approach to Fault Detection and Isolation Applied to Electrical Machines. IEEE-CCA'98, Conference on Control Applications, pp 475-479, Trieste, Italy, Sept. 1-4, 1998.
- * [**Combastel, 1999a**] : Combastel C., Gentil S., Rognon J.P., Fault Detection and Isolation using Local Models, Comparison with Unknown Input Observers. ECC'99, European Control Conference, Karlsruhe, Germany, Aug. 31-Sept. 3, 1999.
- * [**Combastel, 1999b**] : Combastel C., Gentil S., Rognon J.P., Sensitivity of Diagnostic Residuals to Modelling Errors. IEEE SDEMPED'99, International Symposium on Diagnostics for Electrical Machines, Power Electronics and Drives, pp. 297-302, Gijón, Spain, Sept. 1-3, 1999.
- * [**Combastel, 2000a**] : Combastel C., Gentil S., Rognon J.P., Fault Diagnosis of an Induction Drive using Residual Directional Properties. IFAC Safeprocess'2000, Budapest, Hungary, June 14-16, 2000.

- * **[Combastel, 2000b]** : Combastel C., Petropol S., Gentil S., Lesecq S., Model-based and Wavelet Approaches to On-line Fault Detection: Application to an Induction Drive. IFAC/IFIP/IEEE MCPL'2000, Conference on Management and Control of Production and Logistics, Grenoble, France, July 5-8, 2000.
- [Comtet-Varga, 1999]** : Comtet-Varga, G., Christophe C., Cocquempot V., Staroswiecki M., FDI for the Induction Motor using Elimination Theory. ECC'99, European Control Conference, Karlsruhe, Germany, 1999.
- [Courtine, 1997a]** : Courtine S., Détection et localisation de défauts dans les entraînements électriques. Thèse de doctorat, Institut National Polytechnique de Grenoble, 1997.
- [Courtine, 1997b]** : Courtine S., Sename O., Rognon J.P., Robust Fault Detection and Isolation in a DC motor: Comparison of Observer-Based Methods. IFAC Safeprocess'97, pp444-449, Kingston Upon Hull, United Kingdom, August 26-28, 1997.
- [David, 1992]** : David R., Alla H., Petri Nets and Grafcet: Tools for Modelling Discrete Event Systems. Prentice Hall, 1992.
- [De Kleer, 1986]** : De Kleer, J. et J. S. Brown, Theories of causal ordering. Artificial Intelligence, Vol. 29, pp. 33-62, 1986.
- [De Kleer, 1987]** : De Kleer J., Williams B.C., Diagnosing Multiple Faults. Artificial Intelligence 32, pp. 97-130, 1987.
- [Ding, 1994]** : Ding X., Frank P.M., Comparison of Observer-based Fault Detection Approaches. IFAC Safeprocess'94, pp556-561, Helsinki University of Technology, Espoo, Finland, June 13-16, 1994.
- [Ding, 1999]** : Ding S.X., Ding E.L., Jeinsch T., An approach to analysis and design of observer and parity relation based FDI systems. IFAC 14th Triennial World Congress, Vol. P, pp. 37-42, Beijing, China, 1999.
- [Dubois, 1985]** : Dubois, D. et H. Prade : A review of fuzzy set aggregation connectives. Information Sciences, Vol. 36, pp. 85-121, 1985.
- [Dubois, 1994]** : Dubois D., Prade H., Possibility Theory and Data Fusion in Poorly Informed Environments. IFAC Control Engineering and Practice, Vol. 2, No.5, pp. 811-823, 1994.
- [Dubuisson, 1990]** : Dubuisson B., Diagnostic et Reconnaissance de Formes. Hermès, Paris, 1990.
- [Duff, 1981a]** : Duff I.S., On algorithms for obtaining a maximum transversal. ACM Association for Computing Machinery, Transactions on Mathematical Software, Vol. 7, No. 3, pp. 315-330, September 1981.
- [Duff, 1981b]** : Duff I.S., Algorithm 575, Permutations for a zero-free diagonal. ACM Association for Computing Machinery, Transactions on Mathematical Software, Vol. 7, No. 3, pp. 387-390, September 1981.
- [Duff, 1978a]** : Duff I.S., Reid J.K., An implementation of Tarjan's algorithm for the block triangularization of a matrix. ACM Association for Computing Machinery, Transactions on Mathematical Software, Vol. 4, No. 2, pp. 137-147, June 1978.
- [Duff, 1978b]** : Duff I.S., Reid J.K., Algorithm 529, Permutations to block triangular form. ACM Association for Computing Machinery, Transactions on Mathematical Software, Vol. 4, No. 2, pp. 189-192, June 1978.

- [**Dunlop, 1997**] : Dunlop S., Jones B, McFarlane D., Petri Net Based Design and Diagnosis for an Assembly Cell. IFAC Safeprocess'97, pp. 1108-1113, Hull, United Kingdom, Aug. 26-28, 1997.
- [**Eldem, 1999**] : Eldem V., Wetherilt A.J., Parmaksiz S., Albas E., Ceçen H., Duyar A., On the identification of three phase induction motors. ECC'99, European Control Conference, Karlsruhe, Germany, 1999.
- [**Evsukoff, 1997a**] : Evsukoff A., Montmain J., Gentil S., Les systèmes flous pour l'aide en ligne aux opérateurs. 6^{ème} Congrès Français de Génie des Procédés, Paris, 24-26 septembre 1997.
- [**Evsukoff, 1997b**] : Evsukoff A., Montmain J., Gentil S., Dynamic Model and Causal Knowledge-based Fault Detection and Isolation. IFAC Safeprocess'97, pp699-704, Kingston Upon Hull, United Kingdom, 1997.
- [**Evsukoff, 1998a**] : Evsukoff A., Montmain J., Gentil S., Causal model based supervising and training. IFAC Workshop on Supervision of Chemical Industries, Lyon, 4-5 juin 1998.
- [**Evsukoff, 1998b**] : Evsukoff A., Le raisonnement approché pour la surveillance de procédés. Thèse de doctorat, Institut National Polytechnique de Grenoble, 1998.
- * [**Evsukoff, 1999**] : Evsukoff A., Combastel C., Gentil S., Qualitative and Neural Decision for Fault Detection and Isolation. IFAC'99, 14th World Congress of International Federation of Automatic Control, Beijing, China, July 5-9, 1999.
- [**Fanchon, 1994**] : Fanchon J.L., Guide des sciences et technologies industrielles. Nathan, France, 1994.
- [**Farreny, 1989**] : Farreny H., Les systèmes experts – principes et exemples. Cépadués, 1989.
- [**Feliot, 1997**] : Feliot C., Modélisation de systèmes complexes: intégration et formalisation de modèles. Thèse de doctorat de l'Université de Lille 1, LAIL, LIA, 1997.
- [**Feray-Beaumont, 1989**] : Feray-Beaumont S., Modèle qualitatif de comportement pour un système d'aide à la supervision des procédés. Thèse de doctorat de l'Institut National Polytechnique de Grenoble, 1989.
- * [**Foulloy, 1997**] : Foulloy L., Galichet S., Combastel C., Gentil S., Fuzzy Processing for Electrical Machines. IEEE-SDEMPED'97 International Symposium on Diagnostics for Electrical Machines, Power Electronics and Drives, Carry-le-Rouet, France, Sept. 1-3, 1997.
- [**Frank, 1990**] : Frank P.M., Fault Diagnosis in Dynamic Systems Using Analytical and Knowledge-based Redundancy – A Survey and Some New Results. Automatica, Vol. 26, No. 3, pp 459-474, 1990.
- [**Frank, 1994a**] : Frank P.M., Application of Fuzzy Logic to Process Supervision and Fault Diagnosis. IFAC Safeprocess'94, Vol. 2, pp. 531-538, Espoo, Finland, 1994.
- [**Frank, 1994b**] : Frank P.M., Ding X., Frequency domain approach to optimally robust residual generation and evaluation for model-based fault diagnosis. Automatica, Vol. 30, pp. 789-804, 1994.
- [**Frank, 1996**] : Frank P.M., Analytical and Qualitative Model-based Fault Diagnosis – A Survey and Some New Results. European Journal of Control, Vol. 2, pp. 6-28, 1996.

- [Frélicot, 1992] : Frélicot C., Un système de diagnostic prédictif par reconnaissance des formes floue. Thèse de Doctorat de l'Université de Technologie de Compiègne, 1992.
- [Frélicot, 1996] : Frélicot C., A fuzzy-based pronostic adaptive system. RAIRO-API-JESA, Vol. 30, n° 2, pp. 281-299, 1996.
- [Frisk, 1997] : Frisk E., Nyberg M., Nielsen L., FDI with adaptive residual generation applied to a DC-Servo. IFAC Safeprocess'97, Vol. 1, pp. 438-443, Kingston Upon Hull, UK, Aug. 26-28, 1997.
- [Galinier, 1989] : Galinier M., "SADT, un langage pour communiquer". Eyrolles, France, 1989.
- [Gerlinger, 1991] : Gerlinger G., Diagnostic en temps-réel par système expert – application à un système de diagnostic embarqué sur automobile, Thèse de doctorat, Université de Technologie de Compiègne, 1991.
- [Gertler, 1990] : Gertler J., Singer D., A new structural framework for parity equation-based failure detection and isolation. Automatica, Vol. 26, No. 2, pp. 381-388, 1990.
- [Gertler, 1991] : Gertler J., Analytical redundancy methods in fault detection and isolation: survey and synthesis. IFAC Safeprocess'91, pp. 9-21, Baden-Baden, Germany, 1991.
- [Gertler, 1995] : Gertler J., Diagnosing parametric faults: from parameter estimation to parity relations. Proceedings of ACC, pp. 1615-1620, 1995.
- [Gertler, 1997] : Gertler J., Fault Detection and Isolation using Parity Relations. Control Engineering Practice, Vol. 5, No. 5, pp. 653-661, 1997.
- [Goldberg, 1989] : Goldberg D.E., Genetic algorithms in search, optimisation and machine learning. Addison-Wesley, 1989.
- [Gomaa, 1997] : Gomaa M., Représentation et Supervision des Systèmes Hybrides par Réseaux de Petri. Thèse de doctorat, Institut National Polytechnique de Grenoble, 1997.
- [Goodwin, 1986] : Goodwin G.C., Lozano Leal R., Maynes D.Q., Middleton R.H., Rapprochement between continuous and discrete model reference adaptive control. Automatica, Vol. 22, No. 2, pp 199-207, 1986.
- [Grellet, 1997] : Grellet G., Clerc G., "Actionneurs électriques". Eyrolles, France, 1997.
- [Hérault, 1994] : Hérault J., Jutten C., Réseaux neuronaux et traitement du signal. Hermès, Paris, 1994.
- [Höfling, 1996] : Höfling T., Isermann R., Fault Detection based on Adaptive Parity Equations and Single Parameter Tracking. Control Engineering Practice, Vol. 4, No. 10, pp.1361-1369, 1996.
- [Isermann, 1993] : Isermann R., Fault diagnosis of machines via parameter estimation and knowledge processing - Tutorial paper. Automatica, Vol. 29, No. 4, pp. 815-835, 1993.
- [Isermann, 1994] : Isermann R., Integration of fault detection and diagnosis methods. IFAC Safeprocess'94, Vol. 2, pp. 597- 612, Helsinki, Finland, Jul. 13-16, 1994.
- [Isermann, 1997a] : Isermann R., Balle P., Trends in the application of model-based fault detection and diagnosis of technical processes. Control Engineering Practice, Vol. 5, No. 5, pp. 709-719, 1997.

- [**Isermann, 1997b**] : Isermann R., Supervision, fault-detection and fault-diagnosis methods – An introduction. Control Engineering Practice, Vol. 5, No. 5, pp. 639-652, 1997.
- [**Isermann, 1998**] : Isermann R., On Fuzzy Logic Applications for Automatic Control, Supervision and Fault Diagnosis. IEEE Transactions on Systems, Man, and Cybernetics – Part A: Systems and Humans, Vol. 28, No.2, March 1998.
- [**Iwasaki, 1986**] : Iwasaki Y, Simon H.A., Causality in Device Behaviour. Artificial Intelligence, Vol. 29, No. 1, pp3-33, 1986.
- [**Jang, 1995**] : Jang J.-R. S., Sun C.-T., Neuro-Fuzzy Modeling and Control. Proceedings of the IEEE, Vol. 83, No. 3, March 1995.
- [**Jang, 1997**] : Jang, J.-R. S., Sun C.-T., Mizutani E., Neuro-Fuzzy and Soft Computing – A Computational Approach to Learning and Machine Intelligence. Prentice Hall, 1997.
- [**Janikow, 1998**] : Janikow C. Z., Fuzzy Decision Trees: Issues and Methods. IEEE Transactions on Systems, Man, and Cybernetics – Part B : Cybernetics, Vol. 28, No. 1, February 1998.
- [**Jiang, 1997**] : Jiang J., Zhao Q., Should we use Parameter Estimation or State Estimation Based FDI Algorithms ?. IFAC Safeprocess'97, pp.474-479, Hull, United Kingdom, 1997.
- [**Kabore, 1998**] : Kabore P., Contribution à la détection et l'isolation de pannes dans les systèmes non linéaires. Thèse de doctorat de l'Université Claude Bernard – Lyon I, 1998.
- [**Kinnaert, 1994**] : Kinnaert M., Optimisation de la génération de résidus, Actes des journées nationales S3 : Sûreté, Surveillance, Supervision. GdR Automatique, Paris, 17-18 novembre 1994.
- [**Kiupel, 1993**] : Kiupel N., Frank P.M., Process supervision with the aid of fuzzy logic. IEEE transactions on Systems, Man, and Cybernetics, Vol. 2, pp. 409 – 414, Le Touquet, France, 17 – 20 octobre 1993.
- [**Köppen-Seliger, 1995**] : Köppen-Seliger B., Frank P.M., Fault detection and isolation in technical processes with neural networks. Conference on Decision and Control, New Orleans, 1995.
- [**Koscielny, 1995**] : Koscielny J.M., Fault isolation in industrial processes by the dynamic table of states method. Automatica, Vol. 31, No. 5, pp 747-753, 1995.
- [**Krishnaswami, 1994**] : Krishnaswami V., Rizzoni G., Nonlinear Parity Equation Residual Generation for Fault Detection and Isolation. IFAC Safeprocess'94, pp.317-322, Espoo, Finland, 1994.
- [**Lapeyre, 1997**] : Lapeyre F., Habmelouk N., Zolghadri A., Monsion M., Fault detection in induction motors via parameter estimation techniques. IFAC Safeprocess'97, Vol. 1, pp. 270-275, Kingston upon Hull, United Kingdom, August 26-28, 1997.
- [**Legg, 1978**] : Legg J.M., Computerized Approach for Matrix-Form FMEA. IEEE Transactions on Reliability, Vol. 27, No.1, pp 154-157, 1978.
- [**Leonhardt, 1997**] : Leonhardt S., Ayoubi M., Methods of fault diagnosis. Control Engineering Practice, Vol. 5, No. 5, pp. 683-692, 1997.
- [**Leyval, 1992**] : Leyval L., Raisonnement causal pour la simulation de procédés industriels continus. Thèse de doctorat, Institut National Polytechnique de Grenoble, 1992.

- [**Leyval, 1994**] : Leyval L., Gentil S., Feray-Beaumont S., Model-based Causal Reasoning for Process Supervision. *Automatica*, Vol. 30, No. 8, pp. 1295-1306, 1994.
- [**Lin, 1996**] : Lin, T. ; B. G. Horne ; P. Tino et L. Giles : Learning long-term dependencies in NARX recurrent neural networks. *IEEE Transactions on Neural Networks*, Vol. 7, No. 6, pp. 1329-1338, 1996.
- [**Linder, 1998**] : Linder S.P., Shafai B., Saif M., Estimating and Accommodating Unknown Actuator Faults with PI Observers. *IEEE CCA'98, Conference on Control Applications*, pp. 461-465, Trieste, Italy, Sept. 1-4, 1998.
- [**Liu, 1998**] : Liu J.H., Frank P.M., Robust H_∞ detection filter with circular region pole constraint design for uncertain linear systems. *IAR scientific report*, pp. 49-55, 1998.
- [**Ljung, 1994**] : Ljung L., Torkel G., On global identifiability for arbitrary model parametrizations. *Automatica*, Vol. 30, No. 2, pp. 265-276, 1994.
- [**Luong, 1995**] : Luong M., Paris J.M., Maquin D., Ragot J., Observability, reliability and sensor positioning. *AIChE spring national meeting, Houston, USA, March 19-23, 1995*.
- [**Maquin, 1986**] : Maquin D., Fayolle J., Darouach M., Ragot J., Le positionnement de capteurs dans les systèmes industriels de grande dimension. *Proceedings of IMACS congress, Villeneuve d'Ascq, 1986*.
- [**Masson, 1996**] : Masson M.H., Dubuisson B., Frelicot C., Conception d'un module de décision pour le diagnostic par reconnaissance des formes floues. *RAIRO APII JESA*, Vol. 30, No. 2, pp. 319-341, 1996.
- [**Mo, 1997**] : Mo K. J., Lee G., Nam D.S., Yoon Y.H., Yoon E.S., Robust fault diagnosis based on clustered symptom trees. *Control Engineering Practice*, Vol. 5, No. 2, pp. 199-208, 1997.
- [**Montmain, 1992**] : Montmain J., Interprétation qualitative de simulations pour le diagnostic en ligne de procédés continus. *Thèse de Doctorat, Institut National Polytechnique de Grenoble, 1992*.
- [**Montmain, 1993**] : Montmain J., Gentil S., Interprétation qualitative pour le diagnostic en ligne. *Revue Européenne Diagnostic et Sûreté de Fonctionnement*, Vol. 3, No.1, pp. 23-45, 1993.
- [**Moreau, 1999**] : Moreau S., Trigeassou J.C., Champenois G., Gaubert J.P., Diagnosis of induction machines: a procedure for electrical fault detection and localization. *IEEE SDEMPED'99*, pp. 225-229, Gijón, Spain, Sept. 1-3, 1999.
- [**Patton, 1991a**] : Patton R.J. and Chen J., A review of parity space approaches to fault diagnosis. *IFAC Safeprocess'91*, Vol. 1, pp. 239-255, Baden-Baden, Germany, 1991.
- [**Patton, 1991b**] : Patton R.J. and Chen J., A re-examination of the relationship between parity space and observer-based approaches in fault diagnosis. *Revue Européenne de Diagnostic et de Sûreté de Fonctionnement*, Vol. 1, pp. 183-200, 1991.
- [**Patton, 1997a**] : Patton R.J. and Chen J., Observer-based Fault Detection and Isolation: Robustness and Applications. *Control Engineering and Practice*, Vol. 5, No. 5, pp. 671-682, 1997.

- [**Patton, 1997b**] : Patton R.J., Hou M., H_∞ estimation and robust fault detection. ECC'97, European Control Conference, Brussels, Belgium, July 1-4, 1997.
- [**Pearlmutter, 1995**] : Pearlmutter, B. A. : Gradient calculations for dynamic recurrent neural network : a survey. IEEE Transactions on Neural Networks, Vol. 6, No. 5, pp. 1212-1228, 1995.
- [**Ploix, 1998**] : Ploix S., Diagnostic des Systèmes Incertains: Approche Bornante. Thèse de Doctorat, Université Henri Poincaré, CRAN, 1998.
- [**Pons, 1997**] : Pons R., Travé-Massuyès L., Causal ordering for multiple mode systems. In 11th International Workshop on Qualitative Reasoning, 1997.
- [**Poole, 1989**] : Poole D., Normality and faults in logic-based Diagnosis. International Joint Conference on Artificial Intelligence, Detroit, Michigan USA, pp1304-1310, August 20-25, 1989.
- [**Pruvost-Beaurain, 1990**] : Pruvost-Beaurain J.M., *et al.*, Dictionnaire actuel de la langue française. Flammarion, France, 1990.
- [**Ramkumar, 1999**] : Ramkumar K.B., Philips P., Ho W.K., Presig H.A., Lim K.W., A real-time realization of fault detection and diagnosis using finite state automaton. IFAC'99, 14th World Congress of International Federation of Automatic Control, pp.211-216, Beijing, China, July 5-9, 1999.
- [**Rebreyend, 1999**] : Rebreyend P., Algorithmes génétiques hybrides en optimisation combinatoire. Thèse de doctorat, Ecole Normale Supérieure de Lyon, LIP (Laboratoire d'Informatique du Parallélisme), 1999.
- [**Rostaing, 1997**] : Rostaing G., Diagnostic de défauts dans les entraînements électriques. Thèse de doctorat, Institut National Polytechnique de Grenoble, 1997.
- [**Sadrnia, 1997**] : Sadrnia M.A., Patton R.J., Chen J., Robust H_∞ / μ fault diagnosis observer design. ECC'97, European Control Conference, Brussels, Belgium, July 1-4, 1997.
- [**Schaeffer, 1999**] : Schaeffer E., Diagnostic des machines asynchrones : modèles et outils paramétriques dédiés à la simulation et à la détection de défauts. Thèse de Doctorat, Ecole Centrale de Nantes, 1999.
- [**Shen, 1993**] : Shen Q., Leitch R., Fuzzy qualitative simulation. IEEE Transactions on Systems, Man, and Cybernetics, Vol. 23, No. 4, pp. 1038-1064, 1993.
- [**Shen, 1998**] : Shen, L.C. Hsu P.L., Robust design of fault isolation observers. Automatica, Vol. 34, No. 11, pp. 1421-1429, 1998.
- [**Staroswiecki, 1989**] : Staroswiecki M., Declerk P., Analytical Redundancy in non linear Interconnected Systems by Means of Structural Analysis. IFAC IMACS IFORS AIPAC, 1989.
- [**Staroswiecki, 1996**] : Staroswiecki M., Cassar J.P., Approche structurelle pour la conception des systèmes de surveillance. Ecole d'été d'Automatique de Grenoble: surveillance des systèmes continus, Tome 1, 1996.
- [**Strang, 1996**] : Strang G., Nguyen T., Wavelets and Filter Banks. Wellesley-Cambridge Press, 1996.

- [**Tarjan, 1972**] : Tarjan R., Depth-first search and linear graph algorithms. SIAM J. Comput., Vol. 1, No.2, pp. 146-160, June 1972.
- [**Theilliol, 1997**] : Theilliol D., Sauter D., Vela Valdes L.G., Integration of qualitative and quantitative methods for fault detection and isolation. IFAC Safeprocess'97, pp. 687-692, Hull, United Kingdom, Aug. 26-28, 1997.
- [**Thomson, 1999**] : Thomson W.T., A review of on-line condition monitoring techniques for three-phase squirrel-cage induction motors – Past, present and future. Proceedings of IEEE SDEMPED'99, pp3-18, Gijón, Spain, 1999.
- [**Travé-Massuyès, 1997**] : Travé-Massuyès L., Dague P., Guerrin F., Le raisonnement qualitatif. Hermès, France, 1997.
- [**Trigeassou, 1996**] : Trigeassou J.C., Estimation paramétrique de modèles continus. Ecole d'été d'automatique de Grenoble, Surveillance des systèmes continus, Tome 1, Grenoble, France, 2-6 Sept. 1996.
- [**Valette, 1994**] : Valette R., Künzle L.A., Réseaux de Petri pour la détection et le diagnostic. Actes des journées nationales S3 : Sécurité, Surveillance, Supervision, GdR Automatique, Paris, 17-18 novembre 1994.
- [**Von Bertalanffy, 1993**] : Von Bertalanffy L., Théorie générale des systèmes, trad. J-B. Chabrol. Dunod, Paris, 1993.
- [**Von Raumer, 1994**] : Von Raumer T., Commande adaptative non linéaire de machines asynchrones. Thèse de Doctorat, Institut National Polytechnique de Grenoble, 1994.
- [**Weber, 1999a**] : Weber P., Gentil S., A parameter estimation method for the diagnosis of sensor or actuator abrupt faults. ECC'99 European Control Conference, Duisburg, Germany, Sept., 1999.
- [**Weber, 1999b**] : Weber P., Gentil S., Ripoll P., Foulloy L., Multiple fault detection and isolation using fuzzy decision. IFAC World Congress, Beijing, China, 1999.
- [**Weber, 1999c**] : Weber P., Diagnostic de procédé par l'analyse des estimations paramétriques de modèles de représentation à temps discret. Thèse de doctorat, Institut National Polytechnique de Grenoble, 1999.
- [**Willsky, 1976**] : Willsky A.S., A survey of design methods for failure detection in dynamic systems. Automatica, Vol. 12, pp. 601-611, 1976.
- [**Zadeh, 1965**] : Zadeh L.A., Fuzzy sets. Information and Control, No. 8, pp. 338-353, 1965.
- [**Zhang, 1997**] : Zhang, Q., Basseville M., Benveniste A., Monitoring Dynamic Systems with Polynomial Nonlinearities. IFAC Safeprocess'97, pp.1015-1020, Hull, United Kingdom, 1997.
- [**Zhang, 1998**] : Zhang Q., Basseville M., Benveniste A., Fault Detection and Isolation in Nonlinear Dynamic Systems: A Combined Input-Output and Local Approach. Automatica, Vol. 34, No. 11, pp.1359-1373, 1998.

Annexe A :
Lexique

Annexe A :

Lexique

Anomalie	Particularité non conforme à la loi naturelle ou à la logique. Plus précisément: particularité non conforme à une référence comportementale ou fonctionnelle. Exemple : les défauts, les défaillances ou les pannes sont des anomalies.
Capteur	Support physique servant à obtenir une mesure.
Compensation	Fonction permettant d'assurer la continuité du service en présence d'un défaut ou d'une défaillance.
Défaillance	Anomalie fonctionnelle au sein d'un système physique. Plusieurs types de défaillances peuvent être distinguées : <ul style="list-style-type: none">• Défaillance critique : nécessite une intervention d'urgence,• Défaillance significative : nécessite un processus de traitement,• Défaillance absorbable : peut être ignorée dans un premier temps.
Défaut	Anomalie de comportement au sein d'un système physique. Plusieurs types de défauts peuvent être distingués : <ul style="list-style-type: none">• Défaut abrupt : apparaît instantanément (sur une durée négligeable),• Défaut lent : apparaît très progressivement,• Défaut intermittent : apparaît et disparaît successivement.
Détection	Premier niveau de diagnostic consistant à déterminer de façon rapide et fiable l'existence d'une anomalie.
Diagnostic	Fonction visant à fournir des informations sur les anomalies au sein d'un système physique. Remarque : on distingue traditionnellement plusieurs niveaux de diagnostic : détection, localisation et identification des anomalies.

Disponibilité	<p>Aptitude d'un système à être en état de fonctionner normalement. Un indicateur de disponibilité: la disponibilité opérationnelle (<i>D</i>):</p> $D = \frac{MTBF}{MTBF + MTTR + MTL}$ <p>La disponibilité peut être améliorée par l'intermédiaire de la fiabilité (<i>MTBF</i> augmente) et/ou de la maintenabilité (<i>MTTR</i> diminue) et/ou de la logistique de maintenance¹ (<i>MTL</i> diminue).</p>
Durabilité	<p>Durée de fonctionnement probable ou attendue d'un produit dans des conditions normales d'utilisation (incluant la maintenance). Remarque: la durabilité est liée à la notion de durée de vie d'un produit.</p>
Fiabilité	<p>Aptitude d'un système à fonctionner sans incident. Un indicateur de fiabilité : le temps moyen entre défaillances (<i>MTBF</i>).</p>
Fonction	<p>Action réalisée par un composant exprimée sous forme d'un but à atteindre. Remarque : Les fonctions sont indépendantes des solutions qui les réalisent.</p>
Identification	<p>Troisième niveau de diagnostic consistant à caractériser précisément les anomalies qui se sont produites. Exemple : l'identification d'un défaut consiste à déterminer sa taille (amplitude) et son évolution (forme,...).</p>
Localisation	<p>Deuxième niveau de diagnostic, déclenché par une procédure de détection et consistant à déterminer les anomalies qui se sont produites (sans les caractériser précisément).</p>
Maintenabilité	<p>Aptitude d'un système à être remis en état de fonctionnement normal. Un indicateur de maintenabilité : le temps moyen de réparation (<i>MTTR</i>).</p>
Maintenance	<p>Action consistant à faire revenir ou à maintenir un système dans son état de fonctionnement normal. Plusieurs types de maintenance sont possibles:</p> <ul style="list-style-type: none"> • Maintenance corrective: effectuée après une défaillance • Maintenance préventive: effectuée avant une défaillance en vue de réduire sa probabilité d'occurrence. • Maintenance systématique: maintenance préventive effectuée selon un calendrier préétabli. • Maintenance conditionnelle: maintenance préventive effectuée suite à une mesure révélatrice de l'état de dégradation du système. • Maintenance prédictive : maintenance préventive et conditionnelle.
Mesure	<p>Observation représentée par une variable dont le contenu est l'image d'une grandeur physique. L'obtention d'une mesure se fait par l'intermédiaire d'un système physique : le capteur.</p>

¹ La logistique de maintenance tient compte de la maintenance préventive, des délais logistiques (attentes de fournitures...) et des délais administratifs. Elle est prise en compte par la moyenne des temps logistiques de maintenance (*MTL*).

Modèle	(Objet que l'on imite), reproduction d'un objet, représentation formalisée d'un phénomène.
Observation	Information (connue) sur le comportement réel ou sur le fonctionnement réel du système et qui dépend a priori de l'environnement de ce dernier (changements de consignes, présence de perturbations,...). Remarque: une observation peut résulter du traitement d'observations plus élémentaires.
Panne	Inaptitude d'un dispositif à accomplir sa fonction.
Perturbation	Ecart (entre le comportement d'un système physique et une référence) dont l'origine n'est pas considérée comme une anomalie. Remarque 1: perturbations et défauts sont de même nature. La différence entre les deux tient au caractère normal (perturbation) ou anormal (défaut) de l'écart de comportement. Remarque 2 : si une référence comportementale parfaite du système physique pouvait être trouvée, tous les écarts de comportement seraient des anomalies. Ainsi, la notion de perturbation n'a de sens que dans la mesure où toute référence comportementale (modèle) est imparfaite.
Processus	Ensemble de phénomènes organisés dans le temps rapportés à un même système physique.
Pronostic	Détermination des conséquences des défaillances sur le fonctionnement futur du système physique.
Qualité	Aptitude d'un produit ou d'un service à satisfaire complètement les besoins et les attentes des utilisateurs.
Reprise	Exécution d'une solution curative permettant de revenir au fonctionnement normal.
Résidu	Signal conçu comme indicateur d'anomalies comportementales (défauts) ou fonctionnelles (défaillances). Un résidu est donc un signal porteur de signes et de symptômes.
Signe	Caractère distinctif d'un état comportemental anormal.
Supervision	Tâche de commande et de surveillance de l'exécution d'une opération ou d'un travail effectué par d'autres sans rentrer dans les détails de cette exécution. Remarque : le concept de supervision s'applique dans un cadre hiérarchisé à deux niveaux au moins.
Surveillance	Tâche continue en temps-réel dont le but est de caractériser le mode de fonctionnement du système physique, en enregistrant des informations, en reconnaissant et en indiquant les anomalies de comportement.
Symptôme	Caractère distinctif d'un état fonctionnel anormal.

Système Ensemble déterminé d'éléments (composants, constituants) en interaction.

Remarque : nous avons établi ce lexique à partir des références suivantes : [Ploix 1998], [Isermann 1997a], [Pruvost-Beaurain 1990], [Fanchon 1994] et, plus généralement, à partir des références citées dans le chapitre I.

Annexe B :
Machine à courant continu (MCC)

Annexe B :

Machine à courant continu (MCC)

B.1. Schéma synoptique de l'entraînement continu

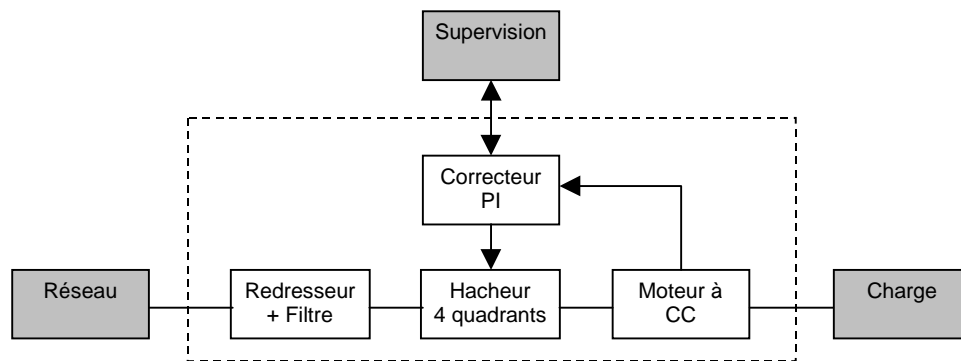


Schéma synoptique de l'entraînement continu

B.2. Paramètres de la MCC

Moteur (1 kW)

Résistance d'induit	$R = 1.15 \Omega$
Inductance d'induit	$L = 1.264 \times 10^{-3} \text{ H}$
Constante de couple	$k = 0.216$
Inertie	$J = 2.267 \times 10^{-3} \text{ kg.m}^2$
Frottement visqueux	$\rho = 7.5 \times 10^{-3} \text{ N.m.s}$

Redresseur + Filtre

Tension d'alimentation	$U_c = 111.5 \text{ V}$
------------------------	-------------------------

Correcteur PI (continu)

Constante Proportionnelle	$k_p = 0.0095$
Constante Intégrale	$k_i = 8.47$

Echantillonnage

Fréquence	$F_{ech} = 10.1 \text{ kHz}$
-----------	------------------------------

B.3. Paramètres pour les simulations

Valeurs “maximales”

Consigne de courant	$\max(I_{ref}) = 25 \text{ A}$
Défaut de l'interface de commande	$\max(f_\beta) = 1$
Défaut du capteur de courant	$\max(f_i) = 25 \text{ A}$
Défaut du capteur de vitesse	$\max(f_\Omega) = 314 \text{ rad/s} (= 3000 \text{ tours/min})$
Défaut alimentation ou hacheur	$\max(f_U) = 160 \text{ V}$

Etat initial

Intégrateur	$X_I(t=0) = 0.3305/k_I$
Courant	$I(t=0) = 5 \text{ A}$
Vitesse	$\Omega(t=0) = 143.99 \text{ rad/s} (= 1375.7 \text{ tours/min})$

Remarque 1 : Une représentation d'état de la MCC est donnée par (III.36) et (III.37). Les états initiaux sont donnés pour cette représentation d'état.

Remarque 2 : Le bruit de mesure utilisé a été obtenu en retirant la composante continue d'un enregistrement en régime permanent effectué sur le banc d'essai « MCC ».

Annexe C :
Machine asynchrone (MAS)

Annexe C :

Machine asynchrone (MAS)

C.1. Schéma synoptique de l'entraînement asynchrone

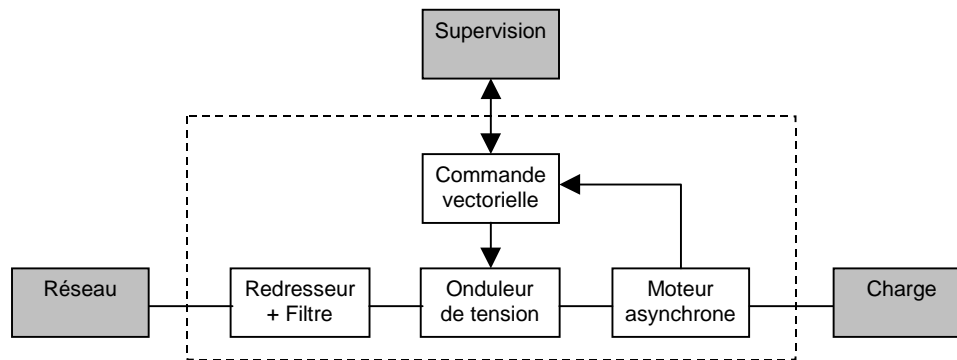


Schéma synoptique de l'entraînement asynchrone

C.2. Paramètres de la MAS

Moteur (1 kW), paramètres du modèle de Park :

Résistance statorique	$R_s = 10 \Omega$
Inductance statorique	$L_s = 0.38 \text{ H}$
Résistance rotorique	$R_r = 3.5 \Omega$
Inductance rotorique	$L_s = 0.3 \text{ H}$
Inductance mutuelle stator/rotor	$L_m = 0.3 \text{ H}$
Nombre de paires de pôles	$p = 2$
Vitesse nominale	$\Omega_n = 157 \text{ rad/s}$ (= 1500 tours/min)
Inertie	$J = 0.02 \text{ kg.m}^2$
Frottement visqueux	$\rho = 0.04 \text{ N.m.s}$

Redresseur + Filtre, Onduleur

Tension d'alimentation	$E = 800 \text{ V}$
Fréquence de la porteuse	$F_p = 500 \text{ Hz}$

Echantillonnage

Fréquence	$F_{ech} = 5 \text{ kHz}$
-----------	---------------------------

C.3. Modèle de Park d'un moteur asynchrone (sans défaut)

L'objectif de cette partie est de rappeler les principales hypothèses et étapes du raisonnement aboutissant aux représentations d'état associées au modèle de Park d'un moteur asynchrone. Nous nous sommes attachés à adopter des notations faisant ressortir les étapes clé de la démarche sans omettre celles établissant le lien entre les paramètres du modèle de Park et les paramètres « physiques » du moteur. Pour plus de détails, les références [Grellet, 1997] et [Von Raumer, 1994] pourront être consultées.

C.3.1. Hypothèses

Plusieurs hypothèses simplificatrices permettent la mise en place du modèle de Park d'un moteur asynchrone :

- Entrefer parfaitement lisse,
- Pertes fer négligeables,
- Saturation dans le circuit magnétique négligeable,
- Harmoniques d'espace négligeables: les enroulements du stator et du rotor sont à répartition sinusoïdale de sorte que les inductances mutuelles entre le stator et le rotor sont des fonctions sinusoïdales de la position mécanique du rotor par rapport au stator.

C.3.2. Equations électriques et mécaniques

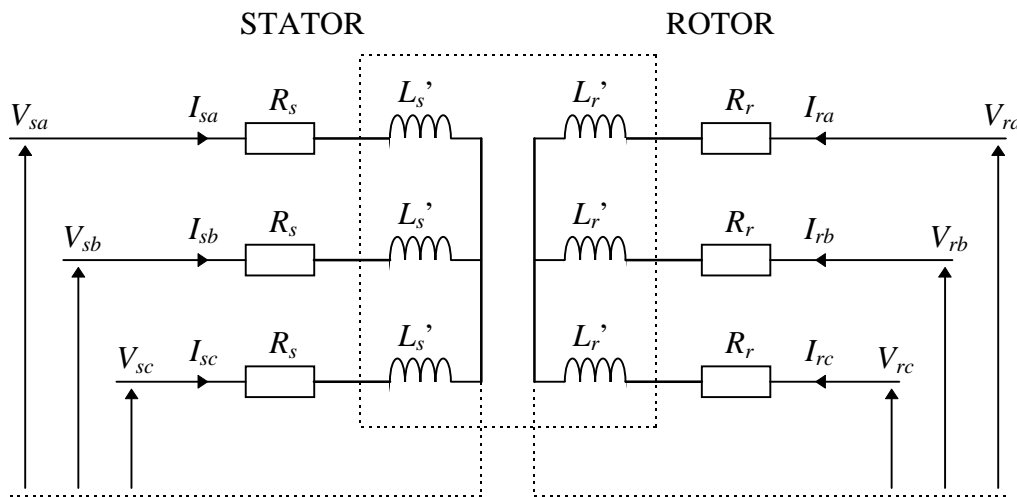


Schéma électrique d'un moteur asynchrone

Equation électrique :

$$V = R.I + \frac{d\Phi}{dt} \text{ et } \Phi = L.I \quad (\text{C.1})$$

$$V = \begin{bmatrix} V_s \\ V_r \end{bmatrix}, I = \begin{bmatrix} I_s \\ I_r \end{bmatrix}, \Phi = \begin{bmatrix} \Phi_s \\ \Phi_r \end{bmatrix}, R = \begin{bmatrix} R_{ss} & 0 \\ 0 & R_{rr} \end{bmatrix}, L = \begin{bmatrix} L_{ss} & M_{sr}(\theta) \\ M_{rs}(\theta) & L_{rr} \end{bmatrix} \quad (\text{C.2})$$

Stator :

$$V_s = \begin{bmatrix} V_{sa} \\ V_{sb} \\ V_{sc} \end{bmatrix}, I_s = \begin{bmatrix} I_{sa} \\ I_{sb} \\ I_{sc} \end{bmatrix}, \Phi_s = \begin{bmatrix} \Phi_{sa} \\ \Phi_{sb} \\ \Phi_{sc} \end{bmatrix}, R_{ss} = \begin{bmatrix} R_s & 0 & 0 \\ 0 & R_s & 0 \\ 0 & 0 & R_s \end{bmatrix}, L_{ss} = \begin{bmatrix} L_s' & M_s' & M_s' \\ M_s' & L_s' & M_s' \\ M_s' & M_s' & L_s' \end{bmatrix} \quad (C.3)$$

Rotor :

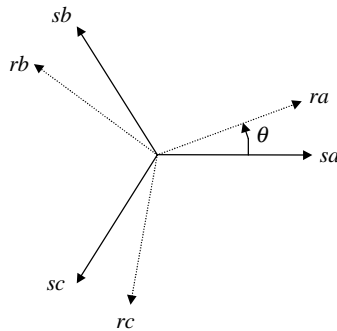
$$V_r = \begin{bmatrix} V_{ra} \\ V_{rb} \\ V_{rc} \end{bmatrix}, I_r = \begin{bmatrix} I_{ra} \\ I_{rb} \\ I_{rc} \end{bmatrix}, \Phi_r = \begin{bmatrix} \Phi_{ra} \\ \Phi_{rb} \\ \Phi_{rc} \end{bmatrix}, R_{rr} = \begin{bmatrix} R_r & 0 & 0 \\ 0 & R_r & 0 \\ 0 & 0 & R_r \end{bmatrix}, L_{rr} = \begin{bmatrix} L_r' & M_r' & M_r' \\ M_r' & L_r' & M_r' \\ M_r' & M_r' & L_r' \end{bmatrix} \quad (C.4)$$

Couplage Stator/Rotor :

$$M_{sr}(\theta) = M' \begin{bmatrix} \cos(\theta) & \cos(\theta - 120^\circ) & \cos(\theta + 120^\circ) \\ \cos(\theta + 120^\circ) & \cos(\theta) & \cos(\theta - 120^\circ) \\ \cos(\theta - 120^\circ) & \cos(\theta + 120^\circ) & \cos(\theta) \end{bmatrix} \quad (C.5)$$

$$M_{rs}(\theta) = M_{sr}(\theta)^T = M_{sr}(-\theta) \quad (C.6)$$

θ désigne l'angle entre le stator et le rotor :



Axes du repère (abc) associé au stator (sa, sb, sc) et du repère associé au rotor (ra, rb, rc)

Bilan de puissance :

$$P = I^T \cdot V \quad (C.7)$$

$$P = I^T \cdot R \cdot I + \frac{d}{dt} \left(\frac{1}{2} \cdot I^T \cdot L \cdot I \right) + \frac{1}{2} \cdot I^T \cdot \frac{dL}{dt} \cdot I \quad (C.8)$$

La puissance instantanée se décompose en trois termes:

- Puissance dissipée par effet Joule
- Variation totale d'énergie électromagnétique
- Puissance transformée en puissance mécanique ($P_{méca}$)

Puissance mécanique :

$$P_{méca} = \Gamma \cdot \Omega = \frac{1}{2} \cdot I^T \cdot \frac{dL}{d\theta} \cdot I \cdot p\Omega \quad \text{où} \quad p\Omega = \frac{d\theta}{dt} \quad (\text{C.9})$$

Couple moteur :

$$\Gamma = p \cdot \frac{1}{2} \cdot I^T \cdot \frac{dL}{d\theta} \cdot I \quad (\text{C.10})$$

Equation mécanique :

$$J \cdot \frac{d\Omega}{dt} + p \cdot \Omega = \Gamma - d_\Gamma \quad (\text{C.11})$$

C.3.3. Transformation de Concordia (triphase \rightarrow biphasé) et rotation

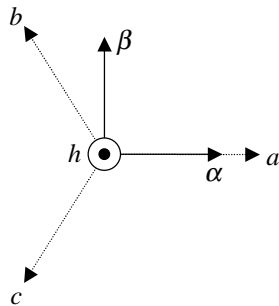
Un vecteur x de grandeurs triphasées peut être décrit par les signaux sur chacune des trois phases (abc) : $x_{abc} = [x_a \ x_b \ x_c]^T$. Un changement de repère approprié (abc) \rightarrow ($\alpha\beta h$) permet de réduire la complexité du système sous certaines hypothèses. Supposons que l'axe h soit orienté suivant la composante homopolaire du système triphasé (c'est-à-dire la somme des trois signaux de phase). Lorsque le système triphasé est équilibré, cette composante est nulle, ce qui permet de réduire le système triphasé à un système biphasé (axes α, β).

Sans prise en compte de la composante homopolaire (hypothèse : $x_a + x_b + x_c = 0$) : (C.12)

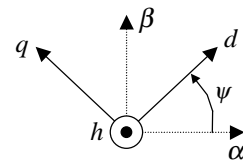
$$x_{\alpha\beta} = K_{23} \cdot x_{abc} \quad \text{avec} \quad K_{23} = \sqrt{\frac{2}{3}} \begin{bmatrix} 1 & -\frac{1}{2} & -\frac{1}{2} \\ 0 & \frac{\sqrt{3}}{2} & -\frac{\sqrt{3}}{2} \end{bmatrix}$$

Avec prise en compte de la composante homopolaire :

$$x_{\alpha\beta h} = K_{33} \cdot x_{abc} \quad \text{avec} \quad K_{33} = \sqrt{\frac{2}{3}} \begin{bmatrix} 1 & -\frac{1}{2} & -\frac{1}{2} \\ 0 & \frac{\sqrt{3}}{2} & -\frac{\sqrt{3}}{2} \\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{bmatrix} \quad (\text{C.13})$$



Transformation de Concordia



Rotation

Un repère (dq) quelconque (ou (dqh)) s'obtient par rotation d'un angle ψ du repère $(\alpha\beta)$ (ou $(\alpha\beta h)$). d signifie « direct » et q signifie « quadrature ».

Sans prise en compte de la composante homopolaire : (C.14)

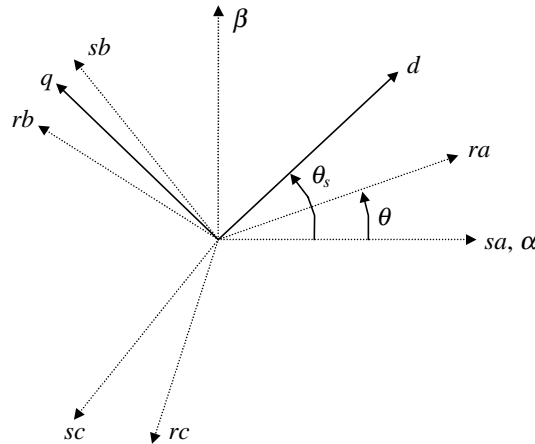
$$x_{dq} = R(\psi).x_{\alpha\beta} \text{ avec } R(\psi) = \begin{bmatrix} \cos(\psi) & \sin(\psi) \\ -\sin(\psi) & \cos(\psi) \end{bmatrix}$$

Avec prise en compte de la composante homopolaire : (C.15)

$$x_{dqh} = \tilde{R}(\psi).x_{\alpha\beta h} \text{ avec } \tilde{R}(\psi) = \begin{bmatrix} \cos(\psi) & \sin(\psi) & 0 \\ -\sin(\psi) & \cos(\psi) & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

Les changements de repère de ce paragraphe s'appliquent à n'importe quel vecteur de grandeurs triphasées. En particulier, ils s'appliquent aussi bien aux grandeurs triphasées relatives au stator qu'à celles relatives au rotor d'une machine asynchrone.

C.3.4. Transformation de Park (rotor fictif équivalent fixe)



Repères triphasés fixes par rapport au stator (sa, sb, sc) et au rotor (ra, rb, rc) ,
repère (dq) formant un angle θ_s quelconque par rapport au stator

Pour éliminer l'influence de θ dans les équations électriques ((C.1), (C.2) et (C.5)), Park a proposé de remplacer le rotor tournant par un rotor équivalent « fixe ». Cela revient à exprimer les grandeurs rotoriques dans le repère (dq) défini par rapport au stator.

Transformation de Park :

Sans prise en compte de la composante homopolaire (hypothèse : MAS équilibrée) : (C.16)

$$X_{dq} = \begin{bmatrix} T(\theta_s) & 0 \\ 0 & T(\theta_s - \theta) \end{bmatrix}.X_{abc} \text{ avec:}$$

$$X_{dq} = \begin{bmatrix} X_{s,dq} \\ X_{r,dq} \end{bmatrix}, X_{abc} = \begin{bmatrix} X_{s,abc} \\ X_{r,abc} \end{bmatrix} \text{ et } T(\psi) = R(\psi).K_{23}$$

Avec prise en compte de la composante homopolaire :

(C.17)

$$X_{dqh} = \begin{bmatrix} \tilde{T}(\theta_s) & 0 \\ 0 & \tilde{T}(\theta_s - \theta) \end{bmatrix} \cdot X_{abc} \text{ avec:}$$

$$X_{dqh} = \begin{bmatrix} X_{s,dqh} \\ X_{r,dqh} \end{bmatrix}, X_{abc} = \begin{bmatrix} X_{s,abc} \\ X_{r,abc} \end{bmatrix} \text{ et } \tilde{T}(\psi) = \tilde{R}(\psi) \cdot K_{33}$$

Remarque 1 : \tilde{T} est orthonormée (\tilde{T} est une matrice unité d'ordre 3: $\tilde{T} \cdot \tilde{T}^T = I_3$) ce qui permet de conserver la puissance instantanée et d'opérer la même transformation sur les tensions que sur les courants. Les axes du repère (dqh) sont orthogonaux.

Remarque 2 : $T(\psi) = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix} \tilde{T}(\psi)$

$\omega_s = \frac{d\theta_s}{dt}$ représente la pulsation du repère (dq) par rapport au stator.

C.3.5. Obtention d'une représentation d'état dans un repère (dq)

Préliminaires :

$$T(\psi) \cdot L_{ss} = L_s \cdot T(\psi) \quad \text{avec } L_s = L_s' - M_s' \quad (C.18)$$

$$T(\psi) \cdot L_{rr} = L_r \cdot T(\psi) \quad \text{avec } L_r = L_r' - M_r' \quad (C.19)$$

$$T(\psi) \cdot R_{ss} = R_s \cdot T(\psi) \quad (C.20)$$

$$T(\psi) \cdot R_{rr} = R_r \cdot T(\psi) \quad (C.21)$$

$$T(\psi) \cdot M_{sr}(\theta) = L_m \cdot T(\psi - \theta) \quad \text{avec } L_m = \frac{3}{2} \cdot M' \quad (C.22)$$

$$\tilde{T}(\psi) \cdot M_{sr}(\theta) = \begin{bmatrix} L_m & 0 & 0 \\ 0 & L_m & 0 \\ 0 & 0 & 0 \end{bmatrix} \tilde{T}(\psi - \theta) \quad (C.23)$$

$$\frac{dT(\psi)}{dt} = \frac{d\psi}{dt} \cdot R\left(\frac{\pi}{2}\right) T(\psi) \quad \text{et } R\left(\frac{\pi}{2}\right) = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} \quad (C.24)$$

$$\tilde{T}(\varphi) \cdot \tilde{T}(\psi)^{-1} = \begin{bmatrix} R(\varphi - \psi) & 0 \\ 0 & 1 \end{bmatrix} \quad (\text{Rappel : } \tilde{T}(\psi)^{-1} = \tilde{T}(\psi)^T) \quad (C.25)$$

Expression des flux statoriques et rotoriques dans un repère (dq) :

$$\Phi_{dq} = \begin{bmatrix} T(\theta_s) & 0 \\ 0 & T(\theta_s - \theta) \end{bmatrix} \cdot \Phi_{abc} \text{ avec } \Phi_{abc} = \begin{bmatrix} L_{ss} & M_{sr}(\theta) \\ M_{rs}(\theta) & L_{rr} \end{bmatrix} \begin{bmatrix} I_{s,abc} \\ I_{r,abc} \end{bmatrix} \quad (\text{C.26})$$

donne :

$$\begin{bmatrix} \Phi_{sd} \\ \Phi_{sq} \\ \Phi_{rd} \\ \Phi_{rq} \end{bmatrix} = \begin{bmatrix} L_s & 0 & M & 0 \\ 0 & L_s & 0 & M \\ M & 0 & L_r & 0 \\ 0 & M & 0 & L_r \end{bmatrix} \begin{bmatrix} I_{sd} \\ I_{sq} \\ I_{rd} \\ I_{rq} \end{bmatrix} \quad (\text{C.27})$$

Expression des tensions statoriques et rotoriques dans un repère (dq) :

$$V_{dq} = \begin{bmatrix} T(\theta_s) & 0 \\ 0 & T(\theta_s - \theta) \end{bmatrix} V_{abc} \text{ avec } V = R \cdot I + \frac{d\Phi}{dt} \text{ dans le repère } (abc) \quad (\text{C.28})$$

donne (en utilisant, entre autres, (C.24)) :

$$V_{sd} = R_s \cdot I_{sd} + \frac{d\Phi_{sd}}{dt} - \omega_s \cdot \Phi_{sq} \quad (\text{C.29})$$

$$V_{sq} = R_s \cdot I_{sq} + \frac{d\Phi_{sq}}{dt} + \omega_s \cdot \Phi_{sd} \quad (\text{C.30})$$

$$V_{rd} = 0 = R_r \cdot I_{rd} + \frac{d\Phi_{rd}}{dt} - (\omega_s - p\Omega) \cdot \Phi_{rq} \quad (\text{C.31})$$

$$V_{rq} = 0 = R_r \cdot I_{rq} + \frac{d\Phi_{rq}}{dt} + (\omega_s - p\Omega) \cdot \Phi_{rd} \quad (\text{C.32})$$

Dans le cas d'un moteur asynchrone à cage, les tensions rotoriques sont nulles.

Expression du couple moteur à partir de grandeurs électriques dans un repère (dq) :

$$\Gamma = p \cdot \frac{1}{2} \cdot I_{abc}^T \cdot \left(\frac{dL}{d\theta} \right)_{abc} \cdot I_{abc} = p \cdot \frac{1}{2} \cdot I_{dqh}^T \cdot \left(\frac{dL}{d\theta} \right)_{dqh} \cdot I_{dqh} \quad (\text{C.33})$$

A partir de :

$$\left(\frac{dL}{d\theta} \right)_{dqh} = \begin{bmatrix} \tilde{T}(\theta_s) & 0 \\ 0 & \tilde{T}(\theta_s - \theta) \end{bmatrix} \left(\frac{dL}{d\theta} \right)_{abc} \begin{bmatrix} \tilde{T}(\theta_s) & 0 \\ 0 & \tilde{T}(\theta_s - \theta) \end{bmatrix}^{-1} \quad (\text{C.34})$$

et en remarquant que :

$$\frac{dM_{sr}(\theta)}{d\theta} = M_{sr} \left(\theta + \frac{\pi}{2} \right) \quad (\text{C.35})$$

on montre (en utilisant, entre autres, (C.25)) :

$$\Gamma = p.L_m.(I_{sq}.I_{rd} - I_{sd}.I_{rq}) \quad (\text{C.36})$$

En utilisant (C.27), le couple moteur peut être exprimé avec d'autres grandeurs électriques.
Par exemple :

$$\Gamma = \frac{p.L_m}{L_r}.(\Phi_{rd}.I_{sq} - \Phi_{rq}.I_{sd}) \quad (\text{C.37})$$

$$\Gamma = p.(\Phi_{sd}.I_{sq} - \Phi_{sq}.I_{sd}) \quad (\text{C.38})$$

Représentations d'état dans un repère (dq) :

En combinant (C.27) avec (C.29) à (C.32), plusieurs représentations d'état peuvent être obtenues. Nous en donnons deux :

Etat électrique constitué des courants statoriques et des flux rotoriques :

Equation électrique: (C.39)

$$X = \begin{bmatrix} I_{sd} & I_{sq} & \Phi_{rd} & \Phi_{rq} \end{bmatrix}^T, U = \begin{bmatrix} V_{sd} & V_{sq} \end{bmatrix}^T$$

$$\frac{dX}{dt} = A.X + B.U \text{ avec:}$$

$$A = \begin{bmatrix} -\gamma & \omega_s & \kappa/\tau_r & p\Omega\kappa \\ -\omega_s & -\gamma & -p\Omega\kappa & \kappa/\tau_r \\ L_m/\tau_r & 0 & -1/\tau_r & (\omega_s - p\Omega) \\ 0 & L_m/\tau_r & -(\omega_s - p\Omega) & -1/\tau_r \end{bmatrix} \quad B = \begin{bmatrix} 1/(\tilde{\sigma}.L_s) & 0 \\ 0 & 1/(\tilde{\sigma}.L_s) \\ 0 & 0 \\ 0 & 0 \end{bmatrix}$$

$$\tau_s = \frac{L_s}{R_s}, \tau_r = \frac{L_r}{R_r}, \tilde{\sigma} = 1 - \frac{L_m^2}{L_s.L_r}, \kappa = \frac{1}{L_m} \cdot \frac{1 - \tilde{\sigma}}{\tilde{\sigma}}, \gamma = \frac{1}{\tau_s.\tilde{\sigma}} + \frac{1}{\tau_r} \cdot \frac{1 - \tilde{\sigma}}{\tilde{\sigma}}$$

Equation mécanique: (C.40)

$$\frac{d\Omega}{dt} = \frac{p.L_m}{J.L_r}.(\Phi_{rd}.I_{sq} - \Phi_{rq}.I_{sd}) - \frac{\rho}{J}.\Omega - \frac{d_\Gamma}{J}$$

Etat électrique constitué des flux statoriques et des courants statoriques :

Equation électrique:

(C.41)

$$X = [\Phi_{sd} \quad \Phi_{sq} \quad I_{sd} \quad I_{sq}]^T, U = [V_{sd} \quad V_{sq}]^T$$

$$\frac{dX}{dt} = A.X + B.U \text{ avec:}$$

$$A = \begin{bmatrix} 0 & \omega_s & -R_s & 0 \\ -\omega_s & 0 & 0 & -R_s \\ R_r/\sigma & p\Omega.L_r/\sigma & -\xi/\sigma & (\omega_s - p\Omega) \\ -p\Omega.L_r/\sigma & R_r/\sigma & -(\omega_s - p\Omega) & -\xi/\sigma \end{bmatrix} \quad B = \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ L_r/\sigma & 0 \\ 0 & L_r/\sigma \end{bmatrix}$$

$$\sigma = L_s.L_r - L_m^2, \xi = R_r.L_s + R_s.L_r$$

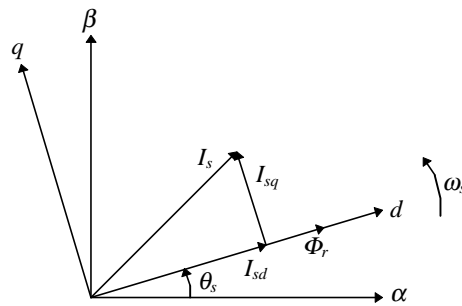
Equation mécanique:

(C.42)

$$\frac{d\Omega}{dt} = \frac{p}{J} \cdot (\Phi_{sd} \cdot I_{sq} - \Phi_{sq} \cdot I_{sd}) - \frac{\rho}{J} \cdot \Omega - \frac{d\Gamma}{J}$$

C.4. Principe de la commande vectorielle

Le contrôle porte sur la valeur instantanée des grandeurs d'état (donc sur le module et sur la phase). La commande par orientation de flux consiste à régler le flux par une composante du courant et le couple par l'autre composante (d'où le nom de commande vectorielle). Pour cela, il faut choisir un système d'axe (d, q) et une loi de commande assurant le découplage du couple et du flux.



Principe de la commande vectorielle

Si le flux rotorique est orienté sur l'axe d d'un repère lié au champ tournant, alors $\Phi_{rq} = 0$.

On a alors :

$$\Gamma = \frac{p.L_m}{L_r} \cdot (\Phi_{rd} \cdot I_{sq}) \quad (C.43)$$

$$\tau_r \cdot \frac{d\Phi_{rd}}{dt} + \Phi_{rd} = L_m \cdot I_{sd} \quad (C.44)$$

Ainsi, I_{sd} fixe le flux et I_{sq} le couple. On retrouve un découplage entre inducteur et induit similaire à celui d'une machine à courant continu (le courant de l'inducteur fixe le flux, le courant d'induit fixe le couple pour un flux donné).

En tenant compte de $\Phi_{rq} = 0$, les quatre équations d'état électrique (C.39) de la machine deviennent (C.45) à (C.48).

$$V_{sd} = \tilde{\sigma} \cdot L_s \left(\frac{dI_{sd}}{dt} + \gamma \cdot I_{sd} - \omega_s \cdot I_{sq} - \frac{\kappa}{\tau_r} \cdot \Phi_{rd} \right) \quad (\text{C.45})$$

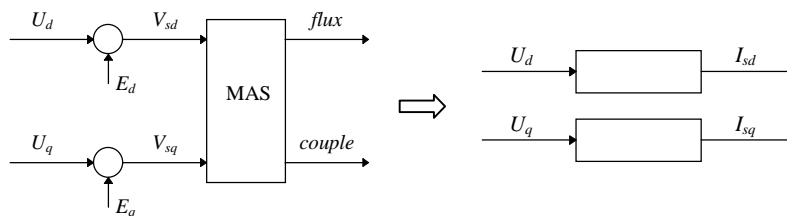
$$V_{sq} = \tilde{\sigma} \cdot L_s \left(\frac{dI_{sq}}{dt} + \omega_s \cdot I_{sd} + \gamma \cdot I_{sq} + \omega \cdot \kappa \cdot \Phi_{rd} \right) \quad (\text{C.46})$$

$$\tau_r \cdot \frac{d\Phi_{rd}}{dt} + \Phi_{rd} = L_m \cdot I_{sd} \quad (\text{C.47})$$

$$\omega_s = \omega + \frac{L_m}{\tau_r} \cdot \frac{I_{sq}}{\Phi_{rd}} \quad (\text{C.48})$$

(ω_s : pulsation du repère tournant induite par $\Phi_{rq} = 0$)

Ces équations montrent que V_{sd} influe à la fois sur I_{sd} et I_{sq} , donc sur le flux et le couple. Il en est de même pour V_{sq} . On est alors amené à réaliser un découplage ; deux nouvelles variables, U_d et U_q , sont définies telles que U_d n'agit que sur I_{sd} (donc que sur le flux) et U_q n'agit que sur I_{sq} (donc que sur le couple). Une méthode de découplage possible consiste à compenser les termes de couplage à l'aide de leurs estimées (E_d et E_q) :



Un régulateur de flux produit l'action U_d et un régulateur de couple produit l'action U_q . La consigne de flux est souvent une loi fonction de la vitesse.

Dans une application de contrôle de vitesse, une boucle de commande supplémentaire agit sur la consigne de couple pour réguler la vitesse.

Pour plus de détails sur la synthèse et le réglage d'une commande vectorielle, l'ouvrage de Grellet et Clerc sur les actionneurs électriques pourra être consulté [Grellet, 1997].

RESUME

Le diagnostic des entraînements électriques permet d'envisager une amélioration de la disponibilité et de la politique de maintenance dans les systèmes de production. L'objectif de la thèse est de détecter et de localiser en ligne des défauts survenant dans un entraînement électrique.

Une analyse causale sert tout d'abord à décomposer systématiquement un modèle global en modèles locaux, indépendamment de la nature du modèle. Un premier jeu de résidus (indicateurs de défaut) est ainsi obtenu. Des méthodes classiques de génération de résidus, fondées sur des modèles numériques, sont également utilisées: observateurs, équations de parité. Dans ce cadre, des techniques de découplage parfait et approximatif sont appliquées. Une étude de sensibilité des résidus est menée. Certaines de ces méthodes sont appliquées à une machine à courant continu et d'autres le sont à une machine asynchrone. Des observateurs adaptatifs réalisant des "tests de modèles" sont utilisés pour le diagnostic d'une machine asynchrone.

L'étape de décision (classification) repose sur un ensemble de critères liés aux résidus. La décision indique quelles classes de défauts sont les plus cohérentes avec les observations disponibles. Les sensibilités des résidus peuvent servir à sélectionner les plus pertinents (au sens d'un critère de performance) soit par l'application de règles élémentaires, soit par une optimisation (algorithme génétique). Les ordres de grandeur relatifs des résidus fournissent également une information plus riche que celle d'une table de signature booléenne et permettent d'améliorer ainsi les performances de la localisation. Bien souvent considérée comme statique, la décision peut également être construite comme un processus dynamique : la reconnaissance de séquences dans un contexte de logique floue illustre cet aspect. Enfin, des méthodes d'apprentissage fondées sur des réseaux de neurones constituent une alternative pour pallier un manque de connaissance a priori.

MOTS-CLES : Diagnostic, Défauts, Détection, Localisation, Génération de résidus, Décision, Machine à courant continu, Machine asynchrone.

TITLE : Decision Methods for Fault Detection and Isolation in Electrical Drives

SUMMARY

The diagnosis of electrical drives allows to think of improving the availability and the maintenance policy in manufacturing plants. The aim of the thesis is online detection and isolation of faults occurring in an electrical drive.

Firstly, a causal analysis is used to systematically decompose a global model into multiple local models, independently from the kind of model. A first set of residuals (fault indicators) is so obtained. Classical residual generation methods based on numerical models are also used: observers, parity equations. In this framework, perfect decoupling and approximate decoupling techniques are applied. The residual sensitivities are studied. Some of these methods are applied to a DC drive and some others are applied to an induction drive. Adaptive observers performing "model tests" are used to diagnose an induction drive.

The decision step (classification) is based on a set of criteria deduced from the residuals. The decision indicates which fault classes are the more coherent with the available observations. The residual sensitivities can be used to select the more pertinent residuals (according to a performance criterion) either by the application of elementary rules or by optimisation (genetic algorithm). The residual relative orders of magnitude also provide a richer information than the one given by a Boolean signature table, allowing thus an improvement of the isolation ability. Often considered as a static process, the decision can also be built as a dynamic process: sequence recognition in a fuzzy logic context illustrates this aspect. Finally, learning methods based on neural networks represent an alternative to bypass a lack of a priori knowledge.

KEYWORDS: Diagnosis, Fault, Detection, Isolation, Residual Generation, Decision, DC Drive, Induction Drive.

DISCIPLINES : Automatique-Productique, Génie électrique.

U.F.R. (Unité de Formation et de Recherche)

ENSIEG (Ecole Nationale Supérieure d'Ingénieurs Electriciens de Grenoble)
rue de la Houille Blanche – BP 46
38402 Saint Martin d'Hères Cedex

LABORATOIRES

LAG : Laboratoire d'Automatique de Grenoble (même adresse que l'UFR)
LEG : Laboratoire d'Electrotechnique de Grenoble (même adresse que l'UFR)
