



HAL
open science

Diagnostic des systèmes incertains : l'approche bornée

Stéphane Ploix

► **To cite this version:**

Stéphane Ploix. Diagnostic des systèmes incertains : l'approche bornée. Sciences de l'ingénieur [physics]. Institut National Polytechnique de Lorraine, 1998. Français. NNT : 1998INPL140N . tel-01776088

HAL Id: tel-01776088

<https://hal.univ-lorraine.fr/tel-01776088>

Submitted on 24 Apr 2018

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.



AVERTISSEMENT

Ce document est le fruit d'un long travail approuvé par le jury de soutenance et mis à disposition de l'ensemble de la communauté universitaire élargie.

Il est soumis à la propriété intellectuelle de l'auteur. Ceci implique une obligation de citation et de référencement lors de l'utilisation de ce document.

D'autre part, toute contrefaçon, plagiat, reproduction illicite encourt une poursuite pénale.

Contact : ddoc-theses-contact@univ-lorraine.fr

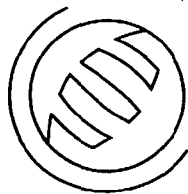
LIENS

Code de la Propriété Intellectuelle. articles L 122. 4

Code de la Propriété Intellectuelle. articles L 335.2- L 335.10

http://www.cfcopies.com/V2/leg/leg_droi.php

<http://www.culture.gouv.fr/culture/infos-pratiques/droits/protection.htm>



inpl
nancy

[M] 13438 PLOIX, S.

Service Commun de la Documentation
INPL
Nancy-Brabois

Thèse

Présentée pour l'obtention du

Doctorat de l'Institut National Polytechnique de Lorraine Spécialité Automatique et Traitement Numérique du Signal

par

Stéphane PLOIX

Diagnostic des systèmes incertains : l'approche bornante

soutenue publiquement le 23 décembre 1998 devant la commission d'examen :

J. FAUCHER	Rapporteur
M. GUGLIELMI	Rapporteur
M. STAROSWIECKI	Rapporteur
J.L. GREFFE	Examineur
F. KRATZ	Codirecteur de thèse
J. RAGOT	Directeur de thèse
E. WALTER	Président du Jury

Remerciements

Cette thèse a été réalisée au laboratoire, dirigé par Monsieur le professeur José Ragot, du Centre de Recherche en Automatique de Nancy à l'Institut National Polytechnique de Lorraine. Ce travail n'aurait pu être mené à bien sans l'aide de nombreuses personnes. Je tiens à remercier ici, succinctement, les protagonistes et m'excuse auprès de ceux que j'ai omis de citer.

Je remercie José Ragot pour ses conseils, sa disponibilité et son engouement pour les questions scientifiques.

Je remercie Frédéric Kratz pour son soutien et sa confiance.

Je remercie Olivier Adrot pour sa collaboration et son aide.

Je remercie Stéphane Giuliani pour sa collaboration et son écoute.

Je remercie les participants du Projet de Recherches Concertées intitulé *diagnostic des machines asynchrones* et en particulier Emmanuel Schaeffer, qui, contre vents et marées, m'ont permis de mieux concevoir les difficultés du diagnostic des systèmes physiques réels.

Je remercie mes camarades du laboratoire, pour l'ambiance conviviale qu'ils ont contribué à entretenir ainsi que Marjorie Schwartz pour son aide et sa vitalité.

Je remercie Eric Walter pour s'être intéressé à mon travail, pour la pertinence de ses remarques et surtout, pour sa sympathie.

Je remercie les rapporteurs et les examinateurs de mon travail pour m'avoir consacré un peu de leur temps et pour leurs corrections.

Je remercie Catherine et Erik pour leur soutien et leurs sacrifices durant toutes ces longues journées de rédaction.

Je remercie mon frère pour sa contribution tant technique que scientifique et je n'oublie pas de remercier mes parents sans qui je n'aurais pu mener à bien mes études.

On dit souvent qu'il faut expérimenter sans idée préconçue. Cela n'est pas possible ; non seulement ce serait rendre toute expérience stérile, mais on le voudrait qu'on ne le pourrait pas. Chacun porte en soi sa conception du monde dont il ne peut se défaire si aisément. Il faut bien, par exemple, que nous nous servions du langage, et notre langage n'est pétri que d'idées préconçues et ne peut l'être d'autre chose. Seulement ce sont des idées préconçues inconscientes, mille fois plus dangereuses que les autres.

Dirons-nous que si nous en faisons intervenir d'autres, dont nous aurons pleine conscience, nous ne ferons qu'aggraver le mal ! je ne le crois pas ; j'estime plutôt qu'elles se serviront mutuellement de contre-poids, j'allais dire d'antidote ; elles entreront en conflit les unes avec les autres et par elles nous forceront à envisager les choses sous différents aspects. C'est assez pour nous affranchir ; on n'est plus esclave quand on peut choisir son maître.

Henri Poincaré
(*La science et l'hypothèse, chapitre IX*)

Avant-propos

Suite à un stage de fin d'étude d'ingénieur dans un laboratoire industriel, j'ai souhaité m'orienter vers la recherche fondamentale. J'ai donc posé ma candidature à différents D.E.A. et, pour sa sympathie et la passion qu'il suscitait pour la recherche, j'ai choisi d'intégrer le laboratoire dirigé par José Ragot. Pour poursuivre un travail amorcé par Besma Gaddouna-Ouladsine, j'ai commencé par m'intéresser aux observateurs d'état, et notamment au découplage à partir de matrices complémentaires. Après un morne service militaire, durant lequel j'ai tout de même pu, bon an mal an, poursuivre mes recherches, j'ai réintégré le laboratoire. A l'instigation de José, avec l'intention de me laisser prolonger mes précédents travaux, le sujet de ma thèse a été arrêtée sur le *diagnostic des systèmes incertains*.

Pour nous confronter aux problèmes réels, nous avons participé au Projet de Recherches Concertées intitulé *Diagnostic des Machines Asynchrones*. L'expérience fut très enrichissante ; le contact avec les électrotechniciens nous a conduit à réviser complètement notre approche du problème. Après une réflexion sur les fondements même du diagnostic, que le lecteur trouvera dans le premier chapitre de ce mémoire, nous avons eu la conviction que les approches par découplage, qu'il soit total ou partiel, ne nous permettraient pas d'atteindre les objectifs que nous visions. Nous avons alors cherché les moyens de prendre en compte les incertitudes dans une opération de diagnostic et nous en sommes venus à développer une approche bornante ou ensembliste du diagnostic.

Certains chercheurs avec lesquels nous avons collaboré pourraient s'étonner de ne trouver dans ce manuscrit que brièvement évoqué les observateurs d'état et les approches par découplage et aucunement la machine asynchrone. Nous avons en effet préféré la machine à courant continu, plus simple, dans la mesure où elle nous permettait d'aborder plus progressivement cette nouvelle approche du diagnostic. Le lecteur pourra s'étonner de ne trouver, dans la bibliographie, que peu de nos publications portant sur les aspects développés dans ce manuscrit, la plupart portent sur nos premiers travaux de recherche. Mais les méandres d'une recherche sont nombreux, surtout lorsque l'on s'astreint à résoudre un problème précis. Qui peut prétendre, lorsque rien n'est balisé, atteindre un but sans détours ? La procédure de publication des travaux présentés dans ce mémoire est en cours et, très certainement, le lecteur intéressé trouvera des compléments dans la littérature scientifique internationale durant les années suivant la soutenance de cette thèse.

Table des matières générale

Chapitre 1 – terminologie et principes du diagnostic **11**

I.	Introduction	13
II.	Terminologie	14
II.1.	Système (<i>system</i>)	15
II.2.	Processus (<i>process</i>)	15
II.3.	Modèle (<i>model</i>)	16
II.4.	Défaillance (<i>failure</i>)	16
II.5.	Défaut (<i>fault</i>)	17
II.6.	Symptôme (<i>symptom</i>)	18
II.7.	Perturbation (<i>disturbance, perturbation</i>)	20
II.8.	Diagnostic (<i>diagnosis, fault diagnosis</i>)	20
III.	Principes du diagnostic à base de modèles analytiques	23
III.1.	Notion de modèle analytique pour le diagnostic	23
III.2.	Procédures de Diagnostic à base de modèles analytiques	26
III.3.	Principes essentiels du diagnostic à base de modèles analytiques	29
IV.	Diagnostic des systèmes incertains	35
IV.1.	Représentation et constitution de la réalité	36
IV.2.	Représentation des systèmes incertains	41
IV.3.	Approches fondamentales du diagnostic des systèmes incertains	49
V.	Conclusion	52

Chapitre 2 – les variables abstraites **55**

I.	Introduction	57
II.	Description probabiliste	58
II.1.	Spécificités des variables abstraites stochastiques	61
II.2.	Loi de Gauss-Laplace	63
II.3.	Loi de distribution uniforme	64
III.	Aspects théoriques de la description par les bornes	67
III.1.	Propriétés des domaines abstraits monodimensionnels	68
III.2.	Propriétés des domaines abstraits multidimensionnels	71
IV.	Evaluation numérique des domaines abstraits	88
IV.1.	Evaluation d'un domaine par balayage	88
IV.2.	Evaluation d'un domaine par pavage	90
V.	Conclusion	92

Chapitre 3 – les métamodèles	95
<hr/>	
I. Introduction	97
II. Structures particulières de métamodèles incertains	98
II.1. Structures additives des incertitudes	100
II.2. Structures multiplicatives des incertitudes	102
II.3. Invariance et dépendance temporelle	105
II.4. Influence de la structure du métamodèle incertain	108
III. Caractérisation des variables abstraites bornées	111
III.1. Exemple introductif	113
III.2. Cas général sans connaissance a priori des incertitudes	115
III.3. Exemple physique	123
III.4. Quelques extensions	126
IV. Conclusion	130
Chapitre 4 – les hypomodèles	133
<hr/>	
I. Introduction	135
II. Les hypomodèles certains	136
II.1. Formulation générique des hypomodèles certains	138
II.2. Propriétés génériques des hypomodèles certains	146
II.3. Conditions d'existence des hypomodèles certains	152
II.4. Spécificités des hypomodèles certains dans le cadre du diagnostic	159
III. Les hypomodèles incertains	167
III.1. Formulation générique des métamodèles dynamiques incertains	168
III.2. Les hypomodèles incertains à horizon infini	174
III.3. Les hypomodèles incertains à horizon fini	192
IV. Conclusion	208
Chapitre 5 – les procédures	209
<hr/>	
I. Introduction	211
II. Procédure orientée bon fonctionnement	212
II.1. Métamodèles de la machine à courant continu	213
II.2. Environnement et simulation	218
II.3. Diagnostic global	224
II.4. Diagnostics locaux	230
III. Procédure orientée mauvais fonctionnement	236
III.1. Métamodèle de la machine à courant continu avec un court-circuit	236
III.2. Spécificités des tests de cohérence	240
IV. Conclusion	247

Conclusion

249

Bibliographie

253

Introduction

Depuis une vingtaine d'années, des communications scientifiques consacrées au diagnostic des systèmes physiques abondent dans la littérature internationale¹. Au-delà de l'aspect scientifique, ces travaux présentent un enjeu commercial certain en ingénierie : un système de diagnostic automatisé accroîtrait autant l'autonomie que la fiabilité d'une installation industrielle. Au fil du temps, l'étude théorique du diagnostic s'est décomposée en deux phases : la mise en évidence de symptômes suivie d'une phase de décision. Cependant, quant on sait que la première phase est principalement étudiée par la communauté *automatique* et que la seconde intéresse la communauté *traitement du signal*, on ne s'étonnera plus du fait que ce découpage ait contribué à faire passer les objectifs initiaux au second plan. Confiez à un chercheur, la tâche d'élaborer des stéthoscopes et à un médecin celle de les exploiter pour diagnostiquer des maladies et, rapidement, verront le jour deux théories : l'une sur les différentes manières de réaliser des stéthoscopes et l'autre, sur les différentes façons de détecter des particularités dans les signaux, mais probablement rien de tangible sur le ou les stéthoscopes qui permettraient à coup sûr de diagnostiquer l'état d'un patient. Pour éviter cet écueil, tout au moins dans le cas des systèmes physiques, nous sommes revenus sur la problématique du diagnostic de sorte à traiter le problème dans sa globalité. Notre objectif est de parvenir à réaliser des procédures, exploitant autant les observations que les modèles disponibles d'un système physique, qui conduiraient à l'objectif final de diagnostic de défauts. La tâche serait relativement facile si les systèmes physiques admettaient toujours des modèles déterministes simples qui soient précis. Or, la simplicité engendre généralement l'imprécision qui se traduit par des incertitudes de modélisation. Pourtant, hormis les approches stochastiques, la plupart des travaux publiés s'inscrivent dans un contexte déterministe ne permettant pas prendre en compte l'incertain. Nous avons étudié comment les incertitudes d'un modèle peuvent être prises en compte par une procédure de diagnostic, particulièrement, avec les approches bornantes qui nous ont semblé les mieux adaptées à la situation. En outre, notre démarche ne postule pas que ces approches doivent être utilisées ; tout au contraire, nous sommes partis de la problématique du diagnostic et, c'est en adoptant une démarche discursive que nous sommes parvenus à nos conclusions. Le mémoire est construit suivant ce cheminement.

Dans un chapitre introductif, nous commençons par développer un raisonnement conceptuel dont le dessein est d'établir de manière univoque la problématique du diagnostic pour en déduire les objectifs à poursuivre. Grâce à cette démarche, nous précisons certains concepts fondamentaux sur lesquels nous nous sommes appuyés pour mettre en évidence les principes basiques du diagnostic. Parmi les solutions théoriques, nous justifierons l'utilité du découpage

¹ Parmi les plus récentes, citons *IFAC symposium on fault detection, supervision and safety for technical processes, SAFEPROCESS'97*, Hull, UK, 26-28 août 1997 ainsi que des sessions consacrées au diagnostic dans : *American Control Conference*, Philadelphia, Pennsylvania, USA, 24-26 juin 1998 ; *Workshop on on-line-fault detection and supervision in the chemical process industries*, Lyon, France, 4-5 juin 1998 et *IEEE conference and decision and control*, San Diego, California, USA, 10-12 décembre 1998.

en deux grands types d'approches possibles : les procédures orientées bon fonctionnement et celles qui sont orientées mauvais fonctionnement. Nous verrons également l'importance de la notion de *test de cohérence*. En discutant des moyens de réalisation de ces tests, nous rechercherons les limites de l'emploi de modèles déterministes. Nous débattrons alors du choix des modèles faisant intervenir des variables stochastiques. Enfin, nous en viendrons aux modèles comportant des incertitudes bornées et nous montrerons leur intérêt théorique.

Le second chapitre commence par faire apparaître les limites de l'usage de variables stochastiques car s'il s'avérait qu'elles soient d'une application facile, rien ne justifierait le recours à l'approche bornante. Or, le titre de la thèse laisse deviner nos conclusions. Alors, en effectuant un parallèle avec les variables stochastiques, nous proposons un formalisme permettant d'appréhender les propriétés des variables définies uniquement par le domaine auquel elles peuvent appartenir. Pour progressivement parvenir aux procédures de diagnostic dans un contexte borné, nous présentons des méthodes de calcul des bornes d'un domaine défini par un champ vectoriel.

Pour mieux comprendre l'influence des incertitudes dans la représentation de systèmes statiques et dynamiques, il faut étudier l'ensemble des comportements possibles associés à différentes structures d'incertitudes. C'est l'objet du troisième chapitre. Nous avons étudié les structures additives et multiplicatives, puis mis en évidence les différences engendrées par la nature temporelle des variables bornées, selon qu'elles soient invariantes ou non. Par ailleurs, il peut arriver que les bornes des incertitudes soient inconnues. Pour les déterminer, nous présentons un algorithme de caractérisation, qui est aux variables dépendantes du temps ce que l'identification est aux variables invariantes.

Le quatrième chapitre est le chapitre clef du mémoire. En partant des modèles déterministes et en nous fondant sur des formulations génériques, nous commençons par mettre en évidence l'influence de la nature d'un modèle sur l'objectif de diagnostic. Les déductions s'étendent aux modèles comportant des incertitudes. Enfin, nous abordons les moyens de tester la cohérence entre des observations et un modèle contenant des variables bornées. Deux situations peuvent se présenter : soit le modèle est intégré, soit il est utilisé comme une relation de parité. Des méthodes permettant d'obtenir et de tester la cohérence à partir de ces modèles ont été mises au point.

Le dernier chapitre rejoint le premier puisqu'il fait le lien entre les résultats fondamentaux qui y figurent et les outils mathématiques qui ont été développés dans les chapitres intermédiaires. Il s'agit d'un chapitre de synthèse illustrant, au moyen de l'exemple de la machine à courant continu, comment des incertitudes de modélisation peuvent être prises en compte par une procédure complète de diagnostic.

Ce mémoire s'achève par une conclusion générale où quelques perspectives de futures recherches sont évoquées.

Chapitre 1

terminologie et principes du diagnostic

TABLE DES MATIÈRES

I.	Introduction	13
II.	Terminologie	14
II.1.	Système (<i>system</i>)	15
II.2.	Processus (<i>process</i>).....	15
II.3.	Modèle (<i>model</i>).....	16
II.4.	Défaillance (<i>failure</i>).....	16
II.5.	Défaut (<i>fault</i>)	17
II.6.	Symptôme (<i>symptom</i>)	18
II.7.	Perturbation (<i>disturbance, perturbation</i>)	20
II.8.	Diagnostic (<i>diagnosis, fault diagnosis</i>).....	20
III.	Principes du diagnostic à base de modèles analytiques	23
III.1.	Notion de modèle analytique pour le diagnostic	23
III.2.	Procédures de Diagnostic à base de modèles analytiques.....	26
III.3.	Principes essentiels du diagnostic à base de modèles analytiques	29
IV.	Diagnostic des systèmes incertains	35
IV.1.	Représentation et constitution de la réalité	36
	Représentation de la réalité physique	37
	Constitution de la réalité physique	38
IV.2.	Représentation des systèmes incertains	41
	Modèle complet pour le diagnostic	41
	Du déterminisme à l'indéterminisme	42
	Nature des modèles des systèmes incertains	44
IV.3.	Approches fondamentales du diagnostic des systèmes incertains.....	49
	Approches a posteriori du diagnostic des systèmes incertains	49
	Approches a priori du diagnostic des systèmes dynamiques incertains	50
V.	Conclusion.....	52

I. Introduction

Alléguer que la résolution d'un problème est directement tributaire de sa formulation n'a rien de surprenant. Ce serait en effet une gageure que d'escompter une solution juste à un problème dont la formulation est erronée. Ainsi, pour espérer définir correctement des procédures de diagnostic de systèmes physiques, il est crucial de se pencher sérieusement sur la formulation des problèmes avant d'en envisager la résolution. Très souvent, dans la littérature, la formulation d'un problème de diagnostic se résume à un modèle d'état comprenant des incertitudes représentant des perturbations et des défauts possibles. Or, ce que nous voulons démontrer dans ce premier chapitre, c'est que cette formulation est généralement incomplète dans la mesure où elle occulte la nature véritable du problème.

Pour parvenir à nos objectifs, il nous a semblé nécessaire de définir les concepts fondamentaux du diagnostic. Certains trouveront cette tâche superflue et ne verront là que redites et manière d'étoffer un manuscrit ; ils auraient tort. Pour des notions aussi fondamentales que *défaut* ou même *diagnostic*, nous avons trouvé de nombreuses définitions très souvent différentes les unes des autres ; certaines, flattant l'opinion générale, se cantonnent derrière un flou sémantique, tandis que d'autres, plus précises certes, ne respectent pas le sens générique des notions. Rares sont les notions définies avec précision et convenant à la description de la pensée relative au diagnostic. Comment expliquer cet état de fait ? Nous répondrons, et nous nous efforcerons de le montrer, que ces carences viennent de ce qui est en amont des mathématiques et qui n'a malheureusement que peu fait l'objet de réflexions : l'axiomatisation de la physique, qui ne peut se concevoir que par les concepts d'une langue naturelle [Poincaré,02]. Rien n'est moins surprenant donc que la formulation d'axiomes selon la mathématique dépende éminemment de l'appréhension verbale d'une problématique. Cette réflexion alimentera la deuxième partie de ce chapitre.

Si la formulation d'un problème de commande est presque immédiate, le diagnostic, qui n'est que l'acte de déterminer quels sont les organes défaillants d'un système physique repose sur l'aperception du réel qui n'est nullement a priori mathématique. Un certain nombre de principes ont déjà fait l'objet de publications dans la littérature propre à différentes communautés scientifiques, citons notamment, l'intelligence artificielle, le diagnostic (de la communauté automatique) et le raisonnement qualitatif. Il est alors indispensable à celui qui a l'ambition de contribuer aux diagnostics des systèmes physiques d'étudier tout autant que de critiquer ces travaux pour parvenir à en faire la synthèse. C'est ce à quoi s'attache la troisième partie de ce chapitre.

Malgré diverses réflexions émanant de chercheurs soit de la communauté automatique, soit de la communauté informatique, nous ne sommes pas parvenus à caractériser les entités conceptuelles sur lesquelles repose la notion de modèle pour le diagnostic à tel point qu'il

nous est apparu difficile d'argumenter contre les a priori mécanistes¹. Or, il est fondamental de se départir de cette conception lorsqu'il s'agit de représenter l'incertain. En effet, ne concevoir les processus que par l'existence de modèles immanents pose le problème rédhibitoire de l'existence des processus dont les modèles contiennent des incertitudes. Nous sommes alors revenus aux principes premiers de la représentation du réel promulgués, dans une large mesure, par Kant [Kant,87] dans la *critique de la raison pure* (1787) pour définir les entités sémantiques nécessaires à la conception d'un modèle pour le diagnostic des systèmes incertains. Cette réflexion nous a permis d'éclaircir la problématique du diagnostic des systèmes physiques ainsi que de réviser certains réflexes insidieux. Nous avons pris la liberté de redéfinir un vocabulaire en rapport avec la notion de modèle à la fois plus générique et plus proche de la physique. Enfin, en nous appuyant sur ces résultats, nous nous efforcerons de recenser les solutions possibles aux problèmes du diagnostic des systèmes incertains en déterminant leurs caractéristiques fondamentales.

II. Terminologie

Les notions fondamentales propres au diagnostic nous ont semblé mériter une étude rigoureuse car elles conditionnent l'adéquation de la physique avec la formulation mathématique. Nous avons parfois proposé de nouvelles définitions lorsque celles que nous avons rencontrées nous étaient parues insuffisantes, en prenant garde toutefois de nous justifier et de nous situer par rapport aux autres propositions. Nous avons voulu exprimer les spécificités essentielles propres au diagnostic de systèmes physiques tout en restant conformes aux définitions génériques. Puisqu'il fallait se limiter, seuls les termes permettant d'appréhender, dans une seconde partie, les principes fondamentaux du diagnostic ont fait l'objet d'études. Notre travail a donc consisté à étudier les principaux écrits s'y rapportant, qui s'étendent aux communautés *traitement du signal*, *automatique* jusqu'à *l'intelligence artificielle* en passant par la *médecine*. Pour effectuer une synthèse des définitions recueillies, souvent en désaccord, nous avons respecté les principes suivants :

- être cohérent avec les principes du diagnostic,
- le spécifique doit s'inscrire dans le cadre du générique,
- l'ensemble des propositions doit être pertinent dans le sens où leur fonction est de décrire les entités intervenant dans la synthèse d'une procédure de diagnostic.

¹ Propre à la théorie philosophique qui admet que la totalité des phénomènes peut être ramenée à une combinaison de mouvements physiques (par extension, à des relations mathématiques déterministes).

Chaque notion est suivie des définitions les plus intéressantes que nous ayons recensées² ; celle qui nous a semblé la plus adaptée à notre dessein apparaît en italique. Pour justifier ce choix, nous commentons les différentes propositions recueillies et, emprunts d'un souci de complétude, nous avons quelquefois ajouté les définitions de certaines notions voisines. Nous nous sommes aussi permis de proposer d'autres définitions lorsque celles que nous avons collectionnées ne nous avaient pas satisfait.

II.1. Système (*system*)

Ensemble structuré d'éléments naturels de même espèce ou de même fonction. Appareil, dispositif formé par la réunion d'organes, d'éléments analogues constituant un ensemble cohérent [Le Petit Robert,94].

Ensemble d'éléments en interaction [Von Bertalanffy,93].

Ensemble déterminé d'éléments discrets (composants, constituants) interconnectés ou en interaction [Vesely et al,81].

La notion de *système*, telle qu'elle est utilisée en automatique, est issue de la théorie générale des systèmes datant de la première moitié du XX^e siècle. Il ne s'agissait non plus d'étudier les interactions entre éléments mais de raisonner sur le comportement d'ensembles : les systèmes. Bien que cette notion, telle qu'elle est entendue par Von Bertalanffy, s'applique particulièrement aux ensembles physiques, il n'en demeure pas moins qu'elle peut réunir des éléments plus abstraits tels que des équations. Nous avons préféré la définition proposée par l'U.S. Nuclear Regulatory Commission (Vesely et al), conforme à celle proposée par Von Bertalanffy, qui a l'avantage d'être plus précise. La notion de système reste pourtant imprécise puisqu'elle s'applique autant aux systèmes physiques, aux systèmes de phénomènes (processus) qu'aux systèmes d'équations pouvant constituer les modèles analytiques des processus. Pour lever cette ambiguïté, De Kleer et al (1987) ont introduit le terme anglo-saxon *artifact* ou *artefact* qui n'a pas d'équivalent français pour distinguer les systèmes physiques des autres systèmes.

II.2. Processus (*process*)

Ensemble de phénomènes³, conçu comme actif et organisé dans le temps [Le Petit Robert,94].

(processus industriel) installation complexe assurant un objectif fonctionnel de haut niveau (production de biens ou de services) [Zwingelstein,95].

² Lorsque les définitions étaient formulées en langue anglaise, nous les avons traduites et, conscient des approximations inhérentes à une telle opération, nous avons fait figurer le texte original en note de bas de page.

³ Tout ce qui se manifeste à la conscience (par l'intermédiaire des sens, ici).

Ensemble de phénomènes organisés dans le temps rapportés à un même système physique.

Un processus est, dans son acception la plus générale, un ensemble de phénomènes conçus comme organisés dans le temps. Ainsi, si l'expression *établir un modèle d'un processus* est légitime, confondre système physique et processus est un non-sens pourtant fréquent (définition de Zwingelstein). En effet, un ensemble de phénomènes ne peut s'identifier au système physique dont il est issu.

II.3. Modèle (*model*)

Schéma, c'est-à-dire description mentale (intériorisée) ou figurée (diagrammes, formules mathématiques, ...) qui, pour un champ de questions, est pris comme représentation abstraite d'une classe de phénomènes, plus ou moins habilement dégagés de leur contexte par un observateur pour servir de support à l'investigation et/ou la communication [AFCET,88].

Le modèle d'un système physique (artifact en anglais) est une description (N.D.A. schématique) de sa structure physique avec en plus les modèles (N.D.A. comportementaux ou fonctionnels) de chacun de ses constituants [De Kleer et al,87].⁴

Bien que la définition proposée par l'AFCET brille par sa précision, nous avons toutefois préféré la définition de De Kleer, plus spécifique au diagnostic des systèmes physiques, qui a le mérite d'être plus concise. La qualité d'un modèle réside moins dans son aptitude à décrire des phénomènes que dans sa conformité avec les objectifs pour lesquels il a été conçu [Penalva,90].

II.4. Défaillance (*failure*)

Etat d'incapacité à réaliser une fonction normale de façon adéquate [Webster's,96].⁵

Altération ou cessation de l'aptitude d'un ensemble à accomplir sa ou ses fonctions requises avec les performances définies dans les spécifications techniques [AFNOR,94].

La cessation de la capacité d'un système physique ou d'un dispositif à accomplir ses fonctions [IEEE,88].⁶

Interruption permanente de la capacité d'un système à assurer une fonction requise dans des conditions opérationnelles spécifiées [Isermann et al,97].⁷

⁴ The model for the artifact is a description of its physical structure, plus models for each of its constituents.

⁵ A state of inability to perform a normal function adequately

⁶ The termination of the ability of an item or equipment to perform its required function.

⁷ A permanent interruption of a system's ability to perform a required function under specified operating conditions.

Anomalie fonctionnelle au sein d'un système physique.

Puisque la notion de *défaillance* est fondamentale en diagnostic, elle réclame une grande rigueur lors de sa définition ; or, les définitions apparaissant dans la littérature manquent singulièrement de précision. La *défaillance* fait souvent référence, tout au moins ce cas est inclus, à l'état d'un système physique ne remplissant plus les fonctions pour lesquelles il a été conçu. Identifier *défaillance* à un état ou à une altération d'aptitude soulève le problème de l'unité : bien que la notion admette un pluriel, rien ne permet de distinguer une *défaillance*, entendue comme un état, de la présence simultanée de plusieurs. Pour remédier à ce problème, il est indispensable de ne rapporter *défaillance* qu'à une fonction anormalement remplie. Remarquons qu'il aurait été possible de trouver une unité telle qu'elle est donnée à maladie, c'est-à-dire considérer qu'une *défaillance* est une entité définissable caractérisée par une évolution dont le terme serait une panne. Néanmoins, aucune des définitions que nous ayons recensées accreditée à cette conception. Ainsi la notion de *défaillance* semble bien liée à la notion de fonction, c'est-à-dire qu'elle suppose l'existence d'une connaissance téléologique⁸ du système physique auquel elle s'applique.

L'adjectif *défaillant* est employé pour qualifier un système physique ou un composant dont une ou plusieurs fonctions sont altérées. Par abus, nous qualifierons aussi de *défaillant* un système physique ou un composant auquel peut être rapporté une anomalie comportementale (cf. *défaut*).

En cas de cessation fonctionnelle, il est permis d'utiliser le terme :

Panne (*break-down*)

Inaptitude d'un dispositif à accomplir sa fonction requise [Zwingelstein,95].

Remarquons aussi qu'une altération fonctionnelle trouve une origine au-delà de la fonction, il s'agit de la cause de *défaillance*.

Cause de *défaillance*

Circonstances liées à la conception, la fabrication ou à l'emploi et qui ont entraîné la *défaillance* [Zwingelstein,95].

II.5. Défaut (*fault*)

Tout écart entre la caractéristique observée sur le dispositif et la caractéristique de référence, lorsque celui-ci est en dehors des spécifications [AFNOR,94].

⁸ Connaissance relative à la finalité, à ce pour quoi un système physique est conçu.

N'importe quel état indésirable d'un composant ou d'un système. Un défaut n'implique pas nécessairement une défaillance [IEEE,88].⁹

Déviations non permises d'au moins une propriété ou un paramètre caractéristique du système des conditions acceptables ou (et) standards [Isermann et al,97].¹⁰

Anomalie de comportement au sein d'un système physique.

La définition de l'AFNOR rattache la notion de défaut à celle de déviance d'une caractéristique d'un phénomène, subordonnant ainsi cette notion à l'existence d'une référence absolue. Cependant, dans le cas général, il y a plusieurs modèles de référence, s'exprimant en terme de relations entre des phénomènes, pouvant représenter un même phénomène. La notion de défaut ne serait donc plus à rattacher à celle de phénomène, mais à celle de modèle ce qui n'est pas conforme aux autres définitions. Celles données par le dictionnaire IEEE et par Isermann et al rapprochent bien *défaut* de comportement ; cependant, les qualificatifs *indésirable* ou *non permis* appartiennent au jugement ; en conséquence, nous avons préféré utiliser *anomalie* qui procède d'une détermination arbitraire. La notion de *défaut* est donc voisine de celle de *défaillance* mais comme le souligne le dictionnaire IEEE, un défaut n'implique pas nécessairement une défaillance. *Défaut*, lié au comportement, est plus général que *défaillance*, liée aux fonctions puisqu'elles sont une abstraction du comportement telle qu'il est conçu téléologiquement. La description comportementale est plus détaillée que la description fonctionnelle et l'inclut donc. De la même manière, la notion de *défaut* inclut celle de *défaillance* ; un défaut n'altère pas nécessairement le fonctionnement d'un système physique mais peut présager d'une défaillance à venir.

II.6. Symptôme (*symptom*)

Phénomène, caractère perceptible ou observable lié à un état ou à une évolution qu'il permet de déceler. Ce qui manifeste, révèle ou permet de prévoir (un état, une évolution) [Le Petit Robert,94].

Un changement dans un organisme indiquant une maladie ou une anomalie physique [Webster's,96].¹¹

Variation anormale d'une quantité observable [Isermann et al,97].¹²

En médecine, phénomène qui révèle un trouble fonctionnel. Dans le cadre du diagnostic industriel, le trouble fonctionnel est une défaillance [Chatain,93].

⁹ Any undesired state of a component or system. A fault does not necessarily require failure.

¹⁰ An unpermitted deviation of at least one characteristic property or parameter of the system from acceptable / usual / standard condition.

¹¹ A change in an organism indicative of disease or physical abnormality.

¹² A change of an observable quantity from normal behaviour.

Appelé aussi mode de défaillance, effet par lequel une défaillance est observée [Zwingelstein,95].

Intuitivement, un symptôme est toute différence entre une prédiction faite par une procédure d'inférence et une observation [De Kleer et al,87].¹³

Caractère distinctif d'un état fonctionnel anormal.

Un symptôme est caractérisé par un ensemble d'observations en provenance d'un système physique ; elles peuvent être brutes ou conditionnées en une information plus facilement exploitable. Leonhard et al (1997) baptise ce conditionnement *extraction de symptômes*. Quoi qu'il en soit, un symptôme est un signe distinctif d'un état fonctionnel anormal : il révèle donc la présence de défaillances. La définition proposée par Isermann omet ce caractère révélateur d'une ou de plusieurs défaillances. D'autre part, bien que la définition de Zwingelstein évite cet impair, elle n'est pas pleinement satisfaisante puisqu'elle suppose un ordonnancement causal des phénomènes qui n'est pas acceptable dans un cas général [Iwasaki,94]. Quant à la définition intuitive de De Kleer, en plus d'omettre le caractère anormal et révélateur d'une défaillance, il subordonne la notion à des méthodes reposant sur la synthèse de prédictions.

Notons que la notion de signe englobe celle de symptôme. Elle peut être associée à la manifestation d'un défaut tandis que symptôme l'était à une défaillance :

Signe (*sign*)

Élément ou caractère (de quelqu'un ou de quelque chose) qui permet de distinguer, de reconnaître [Le Petit Robert, 94].

(signe clinique, médical) caractéristique distinctive observable d'un phénomène.

Caractère distinctif d'un état comportemental anormal.

Si le résultat d'un test de cohérence entre des observations et des modèles se présente sous la forme d'un signal, il se nomme *résidu*. Un résidu est un signal porteur de signes et de symptômes :

Résidu (*residual*)

Indicateur de défauts basé sur la différence entre des mesures et des résultats de calcul à partir d'équations de modèle [Isermann et al,97].¹⁴

Signal conçu comme indicateur d'anomalies fonctionnelles ou comportementales

¹³ Intuitively, a symptom is any difference between a prediction made by the inference procedure and an observation.

¹⁴ A fault indicator, based on a deviation between measurements and model-equation-based computations.

II.7. Perturbation (*disturbance, perturbation*)

(*disturbance*) Une entrée inconnue agissant sur un système [Isermann et al,97].¹⁵

(*perturbation*) Une entrée agissant sur un système, qui provoque un changement temporaire de l'état courant [Isermann et al,97].¹⁶

Tout phénomène conçu comme normal influençant un processus, non ou mal représenté par un modèle de référence.

La notion de *perturbation* est subordonnée à l'existence d'une référence. Cependant, l'usage d'un terme tel qu'*entrée* présuppose un ordonnancement des phénomènes d'un processus et exclut ainsi les modèles de type relations de parité. Nous avons préféré définir *perturbation* à partir de la notion de phénomène.

II.8. Diagnostic (*diagnosis, fault diagnosis*)

Détermination d'un état d'après ses symptômes [Le Petit Robert,94].

L'art ou l'acte d'identifier une maladie (N.D.A. : un état défaillant) à partir de ses signes et symptômes ou conclusion atteinte par le diagnostic [Webster's,96].¹⁷

Détermination de la (des) cause(s) d'un trouble fonctionnel (N.D.A. : un état défaillant), à partir des symptômes et/ou observations [Chatain,93].

Identification de la cause probable de la (ou des) défaillance à l'aide d'un raisonnement logique fondé sur un ensemble d'informations provenant d'une inspection d'un contrôle ou d'un test [AFNOR,94].

(*fault diagnosis*) Détermination du type, de la taille, de l'endroit et de l'instant d'apparition d'un défaut. Suit la détection de défaut. Inclut la localisation et l'identification d'un défaut [Isermann et al,97].¹⁸

Détermination des composants ou organes défaillants d'un système physique.

La définition de Chatain, en dépit du flou inhérent à l'emploi de *cause*, est spécifique aux systèmes industriels. Isermann ne s'intéresse qu'à la caractérisation des défauts, perdant ainsi en conformité avec les deux premières définitions génériques. La définition proposée par l'AFNOR semble être une extension de la définition précédente. Cependant, en se référant à la médecine, cette extension relève de l'étiologie (étude des causes des maladies) et non plus du diagnostic. Grosso modo, la plupart des définitions s'accordent ; aussi, nous en sommes nous

¹⁵ An unknown (and uncontrolled) input acting on a system.

¹⁶ An input acting on a system, which results in a temporary departure from the current state.

¹⁷ The art or act of identifying a disease from its signs and symptoms.

¹⁸ Determination of the kind, size, location and time of a fault. Follows fault detection. Includes fault isolation and identification.

inspirés pour proposer une définition qui ne soit ni trop générique, ni trop spécifique. Dans la littérature scientifique, nous avons rencontré autant les expressions *diagnostic de défauts* (*fault diagnosis* en anglais) que *diagnostic de défaillance* (*failure diagnosis* en anglais). S'il est légitime qu'une procédure de diagnostic s'intéresse aux défaillances, il n'est pas interdit d'élaborer une procédure recherchant des défauts pour formuler alors des diagnostics prédictifs de défaillances. La détermination des défauts apporte plus d'informations que la détermination des défaillances pour remonter aux éléments défaillants d'un système physique. L'opération de diagnostic est souvent décomposée en plusieurs étapes [Willisky,76], [Frank,90] :

Détection des défauts (*fault detection, failure detection*)

(*fault detection*) Détermination des défauts présents dans un système et de leur instant d'apparition [Isermann et al,97].¹⁹

L'étape de détection consiste à identifier le mode de fonctionnement de l'artefact (système physique) et décider s'il est bon ou mauvais [Chatain,93].

Un ou plusieurs tests effectués pour déterminer si un mauvais fonctionnement ou un défaut est présent dans le système [IEEE,88].²⁰

Premier niveau de diagnostic consistant à déterminer de façon rapide et fiable un état comportemental anormal [Horak et al,90].²¹

Localisation des défauts (*fault isolation, failure isolation*)

(*fault isolation*) Détermination de la nature, de l'endroit et de l'instant d'apparition d'un défaut. Suit la détection [Isermann et al,97].²²

L'étape de localisation/diagnostic consiste à chercher la cause probable de la défaillance, détectée dans l'étape précédente [Chatain,93].

Deuxième niveau de diagnostic, déclenché par une procédure de détection, consistant à déterminer de manière plus approfondie les composants défaillants [Horak,90].²¹

¹⁹ Determination of the faults present in a system and the time of detection.

²⁰ One or more tests performed to determine if any malfunctions or faults are present in a unit.

²¹ The goal of the failure detection stage is to determine that something has failed quickly and reliably, in the presence of modeling errors and noise. The goal of the failure isolation stage is to determine details of the failure. While the detection process is independent of the type of the failure, the isolation process requires different algorithms for the handling of different failures types. We structure the FDI system so that one failure detection algorithm, optimized to perform in the presence of modeling errors and noise, serves as a trigger for a multiple-algorithm failure isolation stage.

²² Determination of the kind, location and time of detection of a fault. Follows fault detection.

Identification des défauts (*fault identification*)

Détermination de la taille et de l'évolution d'un défaut [Isermann et al,97].²³

Achevons cette partie en ordonnant les notions caractérisant les observations et les anomalies d'un système physique ; la figure I est la représentation graphique d'un ordonnancement des notions de base du diagnostic selon leurs criticités. Il apparaît nettement que le divers du concept *signe* est plus étendu que celui de *symptôme*. Un symptôme suppose une altération ou une cessation fonctionnelle ; il révèle donc des défaillances et des pannes. L'étendue du concept *défaut* est à comparée à celle de *signe* ; un défaut se manifeste par des signes et non pas exclusivement par des symptômes comme cela était le cas de la défaillance.

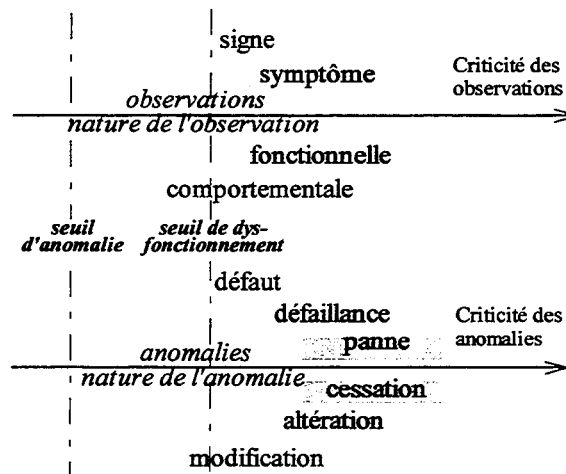


figure I - Ordonnancement des notions selon leur criticité

Un signe révèle une modification qui n'a pas nécessairement d'impact, tout au moins dans l'immédiat, sur le fonctionnement du système physique. Une vibration anormale d'une machine électrique n'altère généralement pas l'exécution de la fonction qui lui est assignée, il s'agirait là d'un signe révélateur d'un défaut et non pas d'un symptôme. Ce signe ne révèle donc pas de défaillance mais il peut présager d'une défaillance à venir. En conséquence, les notions de *symptôme* et de *défaillance* dépendent de la téléologie d'un système physique ; si, pour revenir à l'exemple de la machine électrique, en plus de sa fonction d'entraînement était ajouté : *sans vibration* alors l'anomalie ne serait plus uniquement comportementale ; elle serait fonctionnelle, et les termes *symptôme* et *défaillance* seraient alors justifiés. Dans la suite de ce travail, nous nous restreindrons aux termes de *signes* et de *défaut* parce que nous ne nous intéressons pas à la téléologie des systèmes physiques à ausculter.

²³ Determination of the size and time-variant behaviour of a fault. Follows fault isolation.

III. Principes du diagnostic à base de modèles analytiques

Cette partie résume une étude transversale des principes du diagnostic à base de modèles analytiques ; en effet, le diagnostic concerne plusieurs communautés scientifiques. Citons notamment l'informatique, l'automatique et le traitement du signal sans oublier la médecine bien qu'elle se distingue quelque peu parce qu'elle ne traite pas des systèmes physiques tels que nous les concevons. En médecine, un diagnostic se traduit souvent en termes de *syndrome* ou de *maladie* qui ne sont pas directement comparables aux notions de *défaute* ou de *défaillance* puisqu'ils ne trouvent leur unité que dans leur définition et leur évolution, et non dans l'anomalie elle-même. D'autre part, les modèles analytiques sont généralement inexistantes et les observations non quantifiables. Les seuls rapprochements qu'il est possible d'effectuer avec la médecine se résument au lexique. Les trois autres communautés traitent effectivement de problèmes de diagnostic à base de modèles analytiques à l'exception peut-être de la communauté *traitement de signal* dont les modèles sont souvent locaux (moyenne, variance,... d'un signal) : ils ne représentent pas le comportement global des systèmes physiques. Ainsi, le diagnostic à base de modèles analytiques, sous entendu du comportement global des systèmes physiques, est essentiellement abordé par l'informatique et l'automatique sous les libellés de thèmes tels que diagnostic, intelligence artificielle, raisonnement qualitatif, observateurs d'état, relations de parité, identification... Cette partie est une synthèse des principes fondamentaux du diagnostic à base de modèles analytiques apparaissant dans la littérature des différentes communautés scientifiques précédemment mentionnées. Pour éviter l'obscurité de propos éthérés, nous nous sommes restreints au diagnostic tel qu'il a été défini précédemment en excluant tout ce qui pouvait en dépasser le cadre c'est-à-dire ce qui avait trait au pronostic, à la reconfiguration, à la sûreté de fonctionnement, à la validation de données ou même à l'étiologie.

III.1. Notion de modèle analytique pour le diagnostic

Avant d'aborder les procédures de diagnostic à base de modèles analytiques, il est nécessaire d'éclaircir la notion de modèle telle qu'elle est utilisée lors des opérations de diagnostic des systèmes physiques (artefact). De Kleer décompose la notion de modèle en deux niveaux fondamentaux de représentation : le niveau structurel et le niveau comportemental. La représentation comportementale est constituée de relations entre différents phénomènes du système physique tandis que la représentation structurelle évoque les interconnexions des éléments d'un système physique. Milne [Milne,87] propose deux niveaux de représentation supplémentaires (figure II).

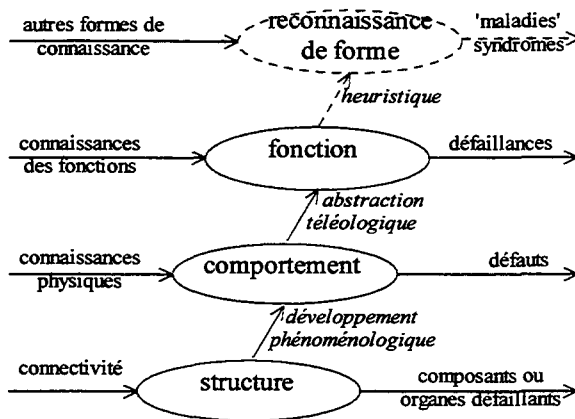


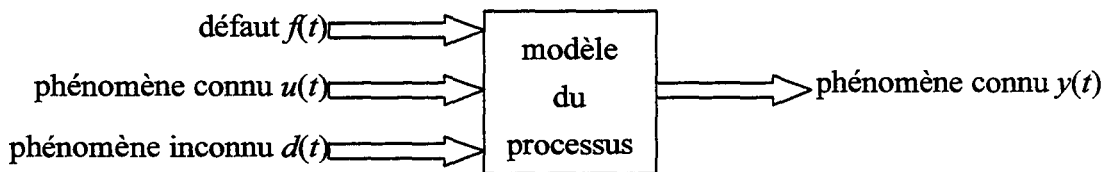
figure II - Niveaux de raisonnement d'un diagnostic, d'après Milne (1987)

Le niveau fonctionnel s'obtient par un raisonnement téléologique ; il est constitué de relations de même nature que celle de la représentation comportementale, mais plus globales, représentant uniquement les fonctions présumées d'un système physique, par exemple la fonction amplification et non pas le comportement des éléments constitutifs de l'amplificateur. Milne ajoute un niveau de représentation supplémentaire : celui de la reconnaissance de forme qui se fonde sur une connaissance heuristique et pourrait identifier des *syndromes* ou des "*maladies*", permettant de présager de l'évolution d'un état. Il est aussi possible d'imaginer que la connaissance heuristique permette de déterminer des causes de défaillance ; cependant, ce niveau de représentation découle généralement sur des techniques d'intelligence artificielle qui sortent du cadre du diagnostic au sens strict. Sur la figure II, est indiqué, en entrée, à gauche, la connaissance sur laquelle chaque niveau de représentation repose et en sortie, à droite, les informations que le niveau permet de déduire. Les flèches rejoignant les niveaux entre eux désignent la démarche qu'il faut suivre pour passer d'un niveau à l'autre.

Les trois niveaux de représentation qui nous intéressent sont structurel, comportemental et fonctionnel. Or, les relations du niveau fonctionnel sont de même nature que les relations comportementales, mais plus abstraites : une relation fonctionnelle est un ensemble de relations comportementales plus ou moins simplifiées. Il faut remarquer que les phénomènes sont a priori non-ordonnés pour un diagnostic [De Kleer et al,87] puisque la causalité est déjà une représentation [Iwasaki et al,94]. Le niveau *comportemental et fonctionnel* comprend des relations entre des phénomènes ou des caractéristiques de phénomène. Le niveau structurel décrit les interconnexions entre les différents sous-systèmes physiques. Les relations entre les éléments structurels et les relations comportementales et fonctionnelles permettent de localiser l'organe défaillant lors d'un diagnostic tandis que les relations entre les phénomènes et les relations comportementales et fonctionnelles permettent de mettre en évidence la présence de défauts et de défaillances.

Une représentation théorique usuelle d'un modèle comportemental (ou fonctionnel) apparaît sur la figure III. Le modèle d'un processus associé à un système physique est souvent supposé,

tout du moins en théorie, linéaire et constitué d'équations récurrentes ou algèbraïco-différentielles ordinaires où chaque variable représente un phénomène. Les variables contenues dans les vecteurs $d(t)$ et $f(t)$ représentent des phénomènes dont les caractéristiques sont mal connues voire complètement inconnues, respectivement les perturbations considérées comme normales et les défauts qui sont théoriquement supposés, conformément au modèle représenté par la figure III, se représenter par des variables exogènes. Puisque les défauts apparaissent dans le modèle, il s'agit d'un modèle de mauvais fonctionnement (MF en abrégé). D'autre part, les variables représentant les caractéristiques connues des phénomènes figurent dans les vecteurs $u(t)$ et $y(t)$.



$$\text{Système linéaire en temps continu : } \begin{cases} \frac{dx(t)}{dt} = Ax(t) + Bu(t) + Ed(t) + Kf(t) \\ y(t) = Cx(t) + Fd(t) + Gf(t) \end{cases}$$

figure III - Modèle MF courant en diagnostic d'après Frank (1990), Patton (1994), ...

Lorsque Frank [Frank,90] et Patton [Patton,94] définissent ce modèle, ils affectent de manière générale les phénomènes commandés à $u(t)$ et les mesures à $y(t)$. Cependant, cette représentation semble plus adaptée à la formulation des problèmes de commande où la distinction faite entre les phénomènes connus est naturelle alors qu'elle demeure pour le moins discutable lorsqu'il s'agit de diagnostic.

Illustrons ceci par l'exemple simple d'un asservissement en vitesse d'un moteur à courant continu commandé en courant et couplé à une charge se comportant comme un frottement visqueux. Supposons que les phénomènes tension aux bornes de la machine $u(t)$, courant $i(t)$ et vitesse de rotation $\Omega(t)$ soient mesurés. Un modèle pour la commande de ce moteur s'écrit :

$$\frac{d\Omega(t)}{dt} = -\frac{f}{J}\Omega(t) + \frac{\Phi}{J}i(t) \quad (1)$$

*f représente le coefficient de frottement, J l'inertie de la partie tournante
 Φ est proportionnel au flux d'excitation.*

L'équation électrique suivante n'intervient pas dans le modèle pour la commande²⁴ :

²⁴ Nous supposons ici qu'aucun contrôle du courant n'est effectué.

$$\frac{di(t)}{dt} = -\frac{R}{L}i(t) - \frac{\Phi}{L}\Omega(t) + \frac{1}{L}u(t) \quad (2)$$

R représente la résistance équivalente et *L* l'inductance.

Le modèle de commande (1) a pour entrée le courant et pour sortie la vitesse de rotation de la machine. Ce modèle, du point de vue du diagnostic, permet de tester l'état de la partie mécanique. Les entrées du modèle (2) sont la vitesse et la tension, sa sortie : le courant. Il permet, lui, de tester l'état de la partie électrique. Notons aussi que l'agrégation des modèles (1) et (2) permet de tester l'ensemble de la machine. Ce nouveau modèle (3) a pour entrée la tension et pour sortie le courant et la vitesse ; il s'écrit :

$$\frac{d}{dt} \begin{bmatrix} i(t) \\ \Omega(t) \end{bmatrix} = \begin{bmatrix} -\frac{R}{L} & -\frac{\Phi}{L} \\ \frac{\Phi}{J} & -\frac{f}{J} \end{bmatrix} \begin{bmatrix} i(t) \\ \Omega(t) \end{bmatrix} + \begin{bmatrix} \frac{1}{L} \\ 0 \end{bmatrix} u(t) \quad (3)$$

modèle	entrée(s)	sortie(s)	partie(s) testée(s)
(1)	$i(t)$	$\Omega(t)$	mécanique
(2)	$u(t), \Omega(t)$	$i(t)$	électrique
(3)	$u(t)$	$i(t), \Omega(t)$	mécanique et électrique

Tableau I - Comparaison des modèles de la machine à courant continu

Les résultats précédents peuvent être résumés dans le tableau ci-contre.

Il apparaît donc nettement à travers cet exemple tout d'abord qu'il n'y a aucune de raison d'ordonner les variables représentant les phénomènes connus

lorsqu'il s'agit de diagnostic, et surtout que le choix d'un modèle détermine la partie d'un système physique qui peut être auscultée. Il est donc incomplet de fonder la formulation générale du diagnostic sur un modèle unique où les phénomènes connus sont ordonnancés a priori.

III.2. Procédures de Diagnostic à base de modèles analytiques

Nous avons emprunté à la médecine la formulation générale d'un problème de diagnostic. Le médecin commence par ausculter un patient et constate un certain nombre de signes cliniques à partir desquels il formule son diagnostic. La figure IV inspirée d'une représentation de Leonhard [Leonhardt et al,97] illustre la démarche.

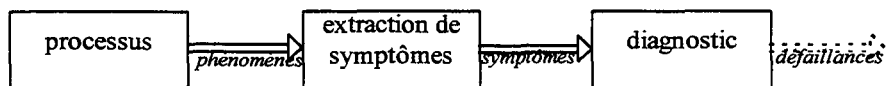


figure IV - Principe général du diagnostic d'après [Leonhardt,97]

Il est alors possible d'effectuer un parallèle avec les approches du diagnostic des systèmes physiques. Lorsqu'un médecin soupçonne une maladie, il tente d'en reconnaître les symptômes parmi les signes qu'il observe. Puisque le médecin connaît les symptômes de la maladie, le chercheur en diagnostic des systèmes physiques parlera d'une approche mauvais fonctionnement (seule la connaissance de la maladie est requise). Par contre, si le médecin

recherche les fonctions vitales altérées et qu'il en déduit l'organe malade, alors son approche sera qualifiée de bon fonctionnement (la seule connaissance du bon fonctionnement des différents organes est nécessaire). Pratiquement, le médecin recourt évidemment aux deux approches.

De Kleer représente l'opération de diagnostic orienté bon fonctionnement par la figure V ; elle montre clairement que les différences structurelles se déduisent des différences de comportements.

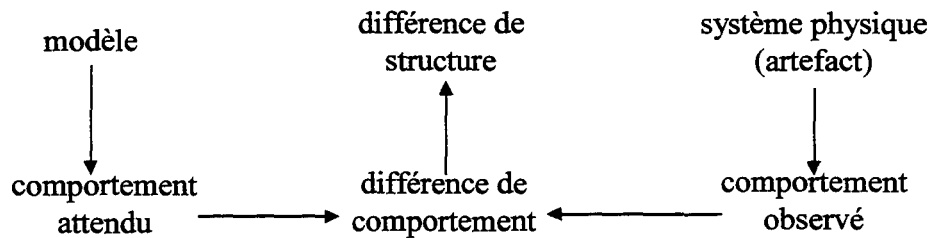


figure V - Procédure de diagnostic orientée Bon Fonctionnement d'après De Kleer (1987)

Il décrit les étapes du diagnostic *bon fonctionnement* en terme de *conflit* et de *candidat*. Un conflit apparaît lorsqu'un symptôme est reconnu c'est-à-dire lorsqu'une incohérence entre le fonctionnement d'un ensemble d'éléments et leur fonctionnement attendu est constatée. Le conflit est la réunion des éléments de cet ensemble qui ne peuvent pas tous fonctionner correctement à la fois tandis qu'un candidat est un ensemble d'hypothèses sur le mauvais fonctionnement des éléments d'un système physique. Chatain décompose ces procédures de diagnostic en deux étapes [Chatain,93]. La première consiste à extraire les symptômes des observations : il s'agit d'une opération de décision. Chow et Willsky se sont intéressés à ce problème d'extraction à partir de modèles analytiques et ont décomposé cette étape en deux fonctions : la génération de résidus et la prise de décision (figure VI).

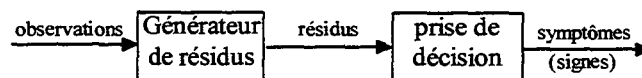


figure VI - Algorithme de diagnostic utilisant la redondance analytique d'après [Chow et al,84]

Mais, à proprement parler, le diagnostic réside dans la seconde étape. Il s'agit d'identifier les éléments défaillants d'un système physique c'est-à-dire qu'il faut commencer par recenser les conflits en recherchant tous les conflits minimaux possibles. Cela revient à remonter des symptômes aux éléments physiques en utilisant la topologie et les connaissances fonctionnelles des organes. Il faut alors générer des candidats minimaux cohérents avec l'ensemble des conflits ; autrement dit, déterminer quelles sont les hypothèses de mauvais fonctionnement d'éléments physiques minimum qui expliquent les conflits observés. Pour Reiter [Reiter,87], il s'agit de *déterminer les constituants du système qui, en les supposant défaillants, expliqueraient les contradictions entre les éléments observés et corrects*. Une fois

la fonction défaillante déterminée, il ne reste plus qu'à exploiter l'information contenue dans la représentation structurelle pour reconnaître l'élément physique défaillant.

Les procédures de diagnostic orientées mauvais fonctionnement sont nécessairement dédiées aux défauts dont un modèle a pu être développé, ne serait-ce qu'en termes de changement de caractéristique d'un résidu ou, d'une façon plus détaillée, sous forme de modèle analytique. Il est en effet difficilement envisageable de disposer des modèles de toutes les défaillances pouvant affecter un système physique ; aussi, soit une telle procédure est précédée d'une procédure orientée bon fonctionnement, soit elle est dédiée à une défaillance particulière. L'étape de diagnostic est alors immédiate puisque la détection de symptômes renseigne directement sur la défaillance. Remarquons que, puisque le modèle de comportement intègre la défaillance, l'inférence vers le composant ou l'organe défaillant peut être établie au moment de l'étude. La représentation du niveau structurel n'est plus à intégrer à la procédure de diagnostic. Pour clore ces développements, nous avons résumé, sur les figures VIII et IX, les différentes étapes des procédures de diagnostic orientées bon et mauvais fonctionnement basées sur des modèles analytiques. Chaque bloc correspond à une étape de procédure ; si l'étape est facultative, alors le bloc apparaît en pointillé.

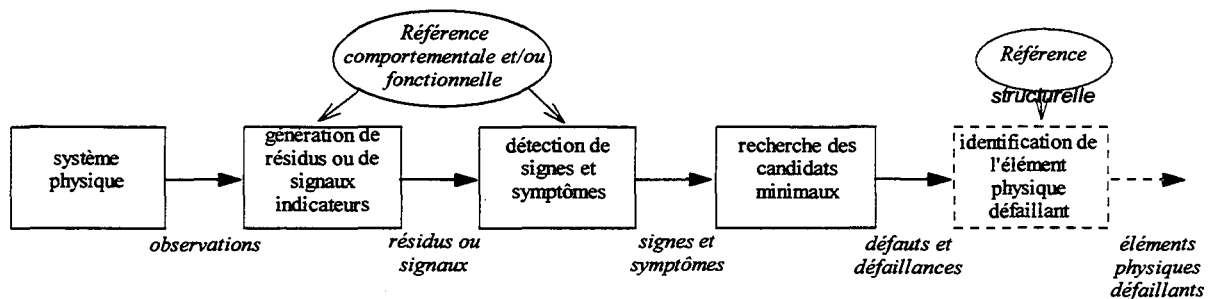


figure VIII- Procédures de diagnostic à base de modèles analytiques orientées bon fonctionnement

Les procédures orientées bon fonctionnement reposent traditionnellement soit sur la définition de générateurs de résidus rejetant les perturbations, soit sur l'identification des paramètres du modèle de bon fonctionnement. Une incohérence comportementale est révélée lorsque les résidus sortent d'un voisinage de zéro ou que les paramètres identifiés du modèle de bon fonctionnement dérivent anormalement. Puis advient l'étape de détection qui consiste à décider si les observations doivent être ou non interprétées comme des signes ou des symptômes. Puisque cette étape dépend directement de la connaissance disponible sur les perturbations, elle dépend, autant que la précédente, de la référence comportementale et/ou structurelle. Suit alors l'étape de diagnostic proprement dit qui se résume à rechercher les candidats minimaux i.e. les hypothèses minimales sur les éléments en dysfonctionnement.

Les procédures orientées mauvais fonctionnement se développent traditionnellement selon deux voies distinctes. La voie baptisée FDI de l'anglais *fault detection and identification* et la voie, moins usitée, qui procède de l'identification directe des caractéristiques des défauts.

Toutes les étapes de ces procédures reposent sur la référence comportementale et/ou structurelle parce qu'elle est sensée contenir toute l'information utile sur le système physique. La voie FDI repose sur la génération de résidus dédiés à des défauts, c'est-à-dire qu'un résidu n'est, idéalement, sensible qu'à un seul défaut malgré les perturbations et les autres défauts possibles. Cette étape est naturellement suivie d'une étape de détection décidant si les résidus révèlent ou non des signes d'anormalités et, éventuellement, d'une étape d'identification des caractéristiques associées aux défauts. L'autre voie se décompose en deux étapes. Il s'agit tout d'abord d'identifier directement les caractéristiques d'un défaut et de décider de sa présence ou non selon les résultats obtenus.

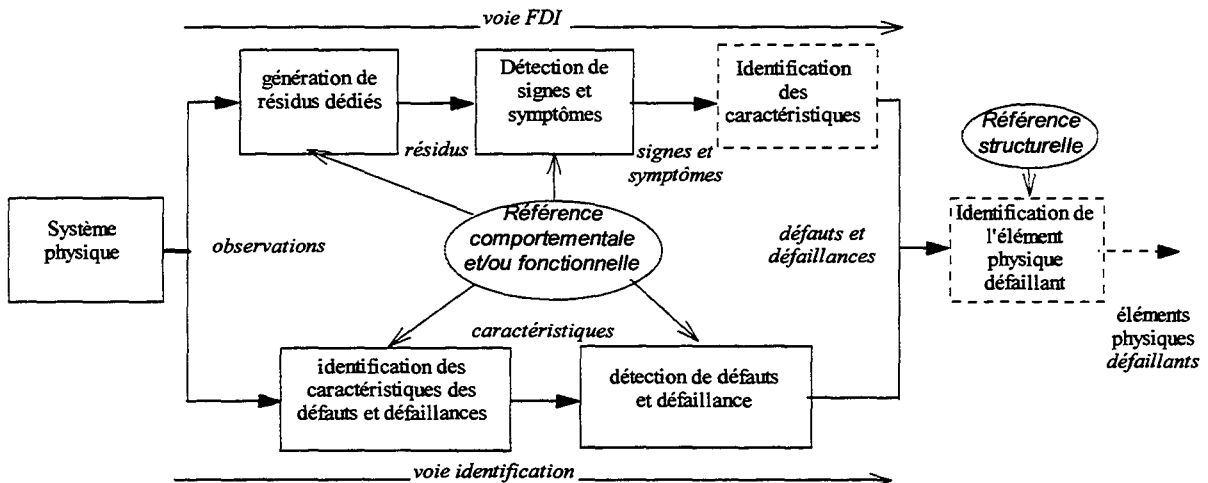


figure IX - Procédures de diagnostic à base de modèles analytiques orientées mauvais fonctionnement

Pour parachever le diagnostic, ces procédures peuvent être éventuellement suivies d'une étape de détermination de l'élément physique défaillant.

III.3. Principes essentiels du diagnostic à base de modèles analytiques

La sous-partie précédente s'attache à relater les concepts et principes du diagnostic à base de modèles analytiques des systèmes physiques qui ressortent de la littérature scientifique. Cependant, pour que la synthèse soit complète, il manque d'en extraire quelques principes fondamentaux. En effet, la résolution d'un problème de diagnostic ne peut s'obtenir de la résolution de sous-problèmes sans qu'eux-mêmes s'inscrivent dans le schéma d'une résolution globale. C'est pourtant là un écueil de spécialistes. Le diagnostic des systèmes incertains en est un bon exemple : l'automaticien propose des générateurs de résidus toujours plus complexes, résumant le diagnostic à un problème de détection ; l'expression détection et localisation de défauts s'est d'ailleurs substituée à celle de diagnostic, légitimant ainsi le découplage des problèmes et, plus particulièrement, de générer des résidus et de les traiter sans considérer le diagnostic dans son ensemble. Le chercheur en traitement du signal avalise cette décomposition voyant là matière à appliquer ses outils tandis que le chercheur en intelligence artificielle se désintéresse du problème affirmant avec sagesse que ces

développements ne dépassent pas le cadre de la théorie. Que faut-il entendre par détecter ? Bien que la confusion soit courante, il ne s'agit certainement pas de détecter un défaut car ce serait aussi le localiser : savoir qu'une partie d'un système physique se comporte anormalement (détecter) implique, a fortiori, la localisation de cette partie. Ce qui est souvent entendu par détecter, c'est l'action de générer des résidus insensibles aux perturbations et sensibles aux défauts puis, par filtrage et seuillage, de décider si le fonctionnement s'apparente ou non à un fonctionnement normal. Le problème serait simple si la réalité s'obstinait à ne pas vérifier les modèles qui lui sont rapportés ; aussi, l'automaticien développe des méthodes tendant à effacer des résidus les comportements déviant des systèmes physiques en accordant peu de considération à ce qui ne peut plus être détecté. On nous pardonnera cette satire qui nous semble pourtant illustrer, sommairement certes, de nombreux travaux s'inscrivant dans le cadre du diagnostic des systèmes dynamiques incertains. C'est pour éviter de reproduire de tels sophismes que nous avons choisi de commencer par établir les principes essentiels du diagnostic à base de modèles analytiques des systèmes dynamiques incertains.

Quelle que soit l'approche adoptée, le principe fondamental du diagnostic des systèmes physiques unifiant ce qui a été développé dans la partie précédente s'énonce ainsi :

principe de cohérence

Toute procédure de diagnostic procède nécessairement de tests de cohérence entre des signaux issus d'observations et des modèles de comportement.

La partie IV montrera l'évidence de ce principe. D. T. Horak [Horak et al,90] affirmait déjà en 1990 que la détection et la localisation de défaillances dans les systèmes dynamiques s'obtenaient grâce à un test de cohérence du système au moyen de redondance matérielle et/ou analytique²⁵. Ce principe s'applique donc autant aux procédures exploitant une redondance matérielle (plusieurs capteurs pour une même grandeur physique) et particulièrement adaptées aux défauts de capteurs, qu'aux procédures à base de modèles analytiques. Nous avons distingué deux grands types de procédures de diagnostic ; le principe doit donc s'appliquer indifféremment à l'un comme à l'autre. Avec une approche orientée bon fonctionnement, le modèle est une référence du comportement normal, alors qu'en suivant une démarche mauvais fonctionnement, le modèle est une référence d'un comportement anormal. Notons, car cela peut surprendre, qu'il y a de l'arbitraire dans les adjectifs *normal* et *anormal*, mais cela n'a rien d'étonnant quand on considère que ce qui est qualifié de bon fonctionnement est un ensemble de comportements acceptables dans le sens où ils n'ont pas à provoquer d'alarme ; l'*anormal* se définit alors par opposition au *normal*. Malgré la clarté de la représentation du principe des procédures orientées bon fonctionnement (BF) proposée par De Kleer (§III.2, figure V), elle restreint le test de cohérence à une différence, excluant ainsi les approches ne procédant pas

²⁵ Failure detection and isolation (FDI) in dynamic systems is accomplished by testing the consistency of a system via hardware and/or analytical redundancy.

par différence. Comment en effet inclure des approches dites par relations de parité où il ne s'agit que de vérifier si des observations satisfont à des relations algébriques ? Qu'en est-il alors de l'application du principe de cohérence aux approches de mauvais fonctionnement ? L'approche *mauvais fonctionnement* (MF) usuelle consiste à exploiter un modèle de fonctionnement anormal pour générer des résidus à la fois sensibles à ce qui manifeste l'anormal et insensibles au reste. Nous ne reviendrons pas ici sur la pertinence de cette approche qui sera étudiée plus avant dans ce chapitre, nous nous contenterons de remarquer qu'elle consiste à générer des signes (résidus) mesurant la cohérence entre le modèle de fonctionnement anormal et des observations. Nous énoncerons d'ores et déjà, et nous nous étayerons de développements ultérieurs, que conformément à l'évidence théorique, toute approche MF procède aussi d'un test de cohérence entre des observations et des modèles de fonctionnement anormal.

Ainsi, par rapport au principe de cohérence, que l'approche du diagnostic à base de modèles analytiques soit orientée bon fonctionnement ou mauvais fonctionnement, cela ne fait pas de différence. Est-ce à conclure que la distinction approche BF et approche MF est dépourvue d'intérêt ? En fait, non, car la différence vient du résultat de ce test. Nous ne considérons dans cette sous-partie que les systèmes certains pour lesquels les modèles de bon fonctionnement sont supposés *infiniment* précis et justes et dont, d'autre part, les seuls incertitudes figurant dans les modèles de mauvais fonctionnement ont traits aux occurrences des défauts.

Dans le cas d'une approche orientée bon fonctionnement, la cohérence entre des observations et des modèles analytiques ne prouve rien quant à l'absence de défauts. Il faut tout d'abord conserver à l'esprit qu'un modèle n'a pratiquement qu'un domaine de validité limité, généralement dans le domaine des phénomènes basses fréquences pour les modèles analytiques ; aussi se peut-il qu'un défaut se traduise par des changements de comportement qui n'apparaissent pas dans ce domaine. D'autre part, un modèle analytique n'est rien moins qu'une relation analytique caractérisant le comportement normal. Or, il est possible qu'en certains points de fonctionnement, un comportement défaillant s'apparente à un comportement normal. Ce n'est donc pas parce qu'un comportement s'apparente au comportement normal qu'il est possible de conclure à l'absence de défauts. En conséquence, seule l'incohérence entre un modèle BF et des observations permet de conclure avec certitude à la présence de défauts. Il est cependant possible de contourner cette difficulté en limitant le diagnostic à la constatation d'une nature apparente de comportement ; conclure que tel sous-modèle génère un conflit au sens de De Kleer, c'est-à-dire que l'une des relations comportementales contenues par le sous-modèle est ou n'est pas vérifiée. Une approche BF sera d'autant plus efficace que le nombre de variables physiques observées sera important puisqu'elles

permettront d'examiner plus finement le comportement²⁶. Or, il est réaliste de donner pour fini le nombre d'observations, ainsi, le diagnostic issu d'une procédure BF s'exprime en termes de candidats : un certain nombre de relations comportementales (sous-modèles) BF dont une au moins n'est pas satisfaite.

Cohérence (dépendante du point de fonctionnement)	comportement normal	comportement anormal
Incohérence (indépendante du point de fonctionnement)	comportement anormal	autre comportement

Tableau II - Tests de cohérence du diagnostic

Dans le cas d'une approche orientée mauvais fonctionnement, c'est l'incohérence entre un modèle et des observations qui a un intérêt moindre ; cela prouve uniquement que le comportement ne s'apparente pas au modèle MF exploité par la procédure. Une procédure de diagnostic complète repose sur un modèle BF et sur plusieurs modèles MF représentant par essence un nombre incomplet de comportements anormaux. Ainsi, il se peut très bien qu'aucun modèle MF ne coïncide avec un fonctionnement défaillant et ceci illustre bien le fait que l'incohérence avec des modèles MF a moins d'intérêt que dans un contexte de bon fonctionnement. Contrairement à l'approche BF, la cohérence entre des observations et un modèle MF est plus intéressante. Néanmoins, le problème est ici identique au bon fonctionnement, il se peut qu'un point de fonctionnement s'apparente à la fois à un comportement normal et à un comportement anormal. Mais cela ne constitue pas un dilemme puisqu'il est naturel de privilégier la cohérence par rapport au bon fonctionnement. En revanche, il est possible que le comportement anormal résulte d'un défaut pour lequel aucun modèle MF n'existe : dans ce cas, la conclusion serait imprécise. Elle prédirait un défaut qui n'appartiendrait pas aux défauts représentés par les modèles de mauvais fonctionnement. Pour remédier à ces problèmes, il est nécessaire de s'assurer que le diagnostic demeure valable pour de nombreux points de fonctionnement. Le tableau II résume ces résultats. Contrairement aux approches BF, les approches MF requièrent peu de variables physiques observées mais beaucoup de modèles MF. Là aussi, il est réaliste de considérer que seul un nombre fini de modèle MF est disponible alors qu'il existe un nombre a priori supérieur et non fini de défauts possibles.

²⁶ Nous reviendrons ultérieurement sur cet aspect.

En guise de synthèse, une procédure BF repose sur un certain nombre de relations comportementales caractérisant le bon fonctionnement, associées à des domaines de validité qui s'expriment autant en terme de bande passante qu'en terme de point de fonctionnement. Puisque chaque défaut provoque un certain nombre d'incohérences entre relations comportementales BF et observations, une procédure BF permet de dégrossir un diagnostic en éliminant a priori un certain nombre d'hypothèse de défauts possibles. Par ailleurs, il y a autant de procédures MF que de modèles de comportement défaillant ; il est donc intéressant de disposer des résultats d'une procédure BF pour n'amorcer qu'un nombre minimal de procédures MF qui permettront d'affiner le diagnostic.

Illustrons ces principes fondamentaux par un exemple simple montrant comment des procédures MF peuvent être enchaînées à des procédures BF. Soit un modèle de bon fonctionnement constitué de cinq relations comportementales élémentaires notées : \mathcal{M}_i^{BF} ($i \in \{1...5\}$). Supposons qu'étant données les observations disponibles, les relations élémentaires doivent être agrégées pour former trois modèles pouvant faire l'objet de tests de cohérence :

$$(1) \mathcal{M}_{14}^{BF} \supset (\mathcal{M}_1^{BF} \cup \mathcal{M}_4^{BF})$$

$$(2) \mathcal{M}_{123}^{BF} \supset (\mathcal{M}_1^{BF} \cup \mathcal{M}_2^{BF} \cup \mathcal{M}_3^{BF})$$

$$(3) \mathcal{M}_5^{BF}$$

Par ailleurs, le système physique peut être affecté de défauts dont trois, notés a , b et c , peuvent être modélisés par les modèles de mauvais fonctionnement \mathcal{M}_a^{MF} , \mathcal{M}_b^{MF} et \mathcal{M}_c^{MF} ; chacun d'eux peut faire l'objet d'un test de cohérence.

Tableau d' enchaînement	\mathcal{M}_{14}^{BF}	\mathcal{M}_{123}^{BF}	\mathcal{M}_5^{BF}
défaut a	0	1	1
défaut b	1	0	0
défaut c	1	1	0
test	1	1	0

Tableau III - Tableau de cohérence de l'exemple

Chaque modèle agrégé de bon fonctionnement représente un ensemble de propriétés du système physique ; or, lorsqu'un défaut survient, un certain nombre de propriétés s'invalident tandis que les autres perdurent. Chacune des trois premières lignes du tableau d'enchaînement III vaut pour un défaut alors que les colonnes valent pour les propriétés représentées par les trois modèles agrégés de bon fonctionnement. A l'intersection d'une ligne et d'une colonne, figure un '1' si la propriété représentée par le modèle BF de la colonne demeure valide en présence du défaut correspondant, un '0' apparaît sinon. Le tableau III, qui n'a qu'un but d'illustration, a été rempli aléatoirement.

Conformément à ce que nous avons vu précédemment, les procédures de diagnostic procédant aux tests de cohérence entre les observations et les modèles BF agrégés sont amorcées les premières. Supposons, pour l'exemple, que les résultats suivants soient obtenus :

- \mathcal{M}_{14}^{BF} cohérent avec les observations
- \mathcal{M}_{123}^{BF} cohérent avec les observations
- \mathcal{M}_5^{BF} incohérent avec les observations

Puisqu'il existe une incohérence, nous concluons à la présence d'un défaut. L'incohérence portant sur le modèle \mathcal{M}_5^{BF} conduit d'emblée à éliminer le défaut a des possibilités. En considérant le tableau III, nous constatons que la présence du défaut c implique la cohérence des observations avec les modèles \mathcal{M}_{14}^{BF} et \mathcal{M}_{123}^{BF} . Ainsi, pouvons-nous émettre l'hypothèse de la présence du défaut c et la vérifier en amorçant la procédure testant la cohérence entre le modèle MF correspondant : \mathcal{M}_3^{MF} . Si le test conclut à une cohérence, alors le défaut c est reconnu et le diagnostic est parachevé en déterminant quels sont les organes physiques correspondant à ce défaut. Si le test conclut à une incohérence alors la procédure reposant sur le modèle \mathcal{M}_2^{MF} peut néanmoins être amorcée. En effet, nous avons établi qu'en certains points de fonctionnement, une incohérence peut s'apparenter à une cohérence. Ainsi, les cohérences portant sur les modèles \mathcal{M}_{14}^{BF} et \mathcal{M}_{123}^{BF} peuvent cacher des incohérences. Par conséquent, si le défaut n'a pas été reconnu comme étant le défaut c , il est nécessaire de s'assurer de l'absence du défaut b , en amorçant une procédure reposant sur le modèle \mathcal{M}_b^{MF} , avant de conclure qu'il existe un défaut qui ne correspond à aucun des défauts représentés. Voilà un exemple théorique de la façon d'enchaîner des procédures MF à des procédures BF.

Nous baptiserons de manière globale la procédure qui consiste à enchaîner des procédures MF à des procédures BF une procédure BMF pour procédure bon/mauvais fonctionnement. Il est intéressant de faire le parallèle avec la procédure FDI. FDI suppose souvent une première *détection* globale qui n'est rien d'autre qu'une étape d'une approche BF : seul le modèle global du système physique est utilisé pour formuler le premier diagnostic. La conclusion ne peut porter que sur le comportement global : normal ou anormal. La seconde étape de *localisation*, qui peut comporter des sous-procédures orientées bon ou mauvais fonctionnement, consiste à déterminer plus précisément les défauts présents. Ainsi la décomposition FDI est une décomposition différente de celle dénommée BMF ; elle suppose une première procédure BF sommaire, et généralement globale, suivie d'une seconde dite de *localisation* reposant sur des procédures BF et MF. Une définition intéressante de la démarche FDI est donnée par D. T. Horak [Horak et al,90] : *le but de l'étape de détection de défaillances est de déterminer rapidement et avec fiabilité, malgré la présence d'erreurs de modélisation et de bruits, si quelque chose est défaillant. Le but de l'étape de localisation est de déterminer des détails sur les défaillances. Tandis que la procédure de détection est indépendante du type de défaillance, la procédure de localisation requiert différents algorithmes pour l'appréhension*

des différents types de défaillances. Nous structurons le système FDI de telle sorte qu'un algorithme de détection, optimisé pour travailler en présence d'erreurs de modélisation et de bruits, serve à déclencher un niveau multi-algorithme de localisation des défaillances²¹. La figure X compare la conception usuelle de l'approche FDI avec l'approche BMF.

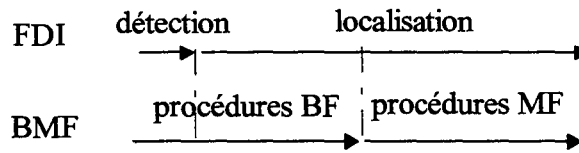


Figure X - Comparaison entre BMF et FDI

L'approche FDI, tout autant que l'approche BMF, recouvre les différentes étapes du diagnostic, la différence tient au découpage des procédures de diagnostic. Le découpage FDI, en séparant ce qui est rapide de ce qui est plus lent, présente un plus grand intérêt pratique puisque les procédures rapides peuvent être testées de façon systématique alors que les procédures lentes ne sont à déclencher que dans les cas où le premier niveau de diagnostic conclut à la présence d'anomalies comportementales. Le découpage BMF, que nous privilégierons dans ce document, est par contre plus intéressant pour le théoricien puisque toutes les procédures orientées BF s'élaborent de la même façon, de même que les procédures MF présentent des similitudes théoriques évidentes.

IV. Diagnostic des systèmes incertains

Les principes du diagnostic que nous avons développés précédemment ont été envisagés dans le cadre des systèmes certains. Or, la plupart des systèmes physiques ne peuvent être représentés précisément. Une solution approximative consiste à se ramener aux systèmes certains en négligeant les perturbations dans le raisonnement et en exploitant autant que possible les degrés de liberté disponibles pour le réglage des générateurs de résidus afin d'atténuer l'effet des perturbations sur les résidus. Cependant, formuler un diagnostic en se fondant sur un raisonnement qui repose sur des hypothèses non vérifiées ne peut mener qu'à des solutions approximatives. Nous montrerons dans cette partie l'universalité du principe de cohérence selon lequel toute procédure de diagnostic résulte de tests de cohérence entre des observations et des modèles. Ainsi, si les modèles ne caractérisent pas l'ensemble des comportements possibles mais se restreignent aux comportements moyens, les cohérences ne seront testées que par rapport à ces comportements moyens. Les perturbations provoquent des incohérences qui n'ont pas lieu d'être et la conséquence en est qu'une incohérence ne peut plus être interprétée de façon stricte. Les modèles déterministes moyens ne suffisent pas, d'un point de vue théorique, au diagnostic des systèmes dynamiques incertains : les modèles

doivent contenir des variables non déterministes²⁷ afin de représenter un ensemble de comportements possibles. Le problème du diagnostic des systèmes incertains est donc en premier lieu un problème de représentation de l'incertain.

Pour établir une formulation mathématique correcte du diagnostic des systèmes physiques, une réflexion sur ce qu'est la réalité physique et sur les moyens de la représenter s'impose. La notion de système physique incertain est déjà obscure ; un processus n'est-il pas toujours certain en tenant compte du fait que la réalité physique semble pouvoir se passer de modèle ? Le problème est de savoir s'il faut attribuer l'incertain aux systèmes physiques, aux processus ou aux modèles. Nous avons jusqu'ici employé les termes de phénomènes et de relations comportementales sans nous justifier, cette partie tente d'éclairer ces entités et d'en établir un ordonnancement en distinguant les entités qui prévalent de celles découlant des autres. Autant il est simple de concevoir la notion de modèle dans un cas déterministe où un certain nombre de phénomènes satisfont les lois qu'il rassemble, autant dans le cas de modèles incertains, la notion de modèle se complexifie. Un modèle ne peut plus être qualifié de vrai ou d'exact. Il est important de réfléchir sur les caractéristiques des modèles au sens large pour distinguer les différentes entités intervenant lors de la synthèse d'une procédure de diagnostic.

La notion de modèle est très imprécise du fait de l'étendue de ce qu'elle représente. Une relation comportementale élémentaire est un modèle au même titre que l'ensemble des relations comportementales représentant le comportement d'un processus. Un modèle de comportement a des propriétés très différentes d'un observateur d'état qui n'est rien d'autre qu'un modèle de simulation ; pourtant nulle distinction n'est faite. Par ailleurs, rien d'autre que des équations casuelles ne permettent d'exprimer les particularités des observateurs d'état. Qu'ont-ils de différent ? Autant de problèmes fondamentaux à résoudre, autant de questions auxquelles il va nous falloir répondre dans cette partie.

IV.1. Représentation et constitution de la réalité

De nombreux travaux théoriques fondent le problème du diagnostic sur un modèle sous forme d'équations différentielles ordinaires (cf. §III, figure III) et n'associent les anomalies comportementales qu'à des erreurs sur des variables physiques, résumant ainsi le diagnostic à un problème de générations de résidus. A contrario, nous avons affirmé que les procédures de diagnostic devaient être conçues comme des tests de cohérence entre des modèles et des observations. La différence entre ces deux conceptions vient de ce que l'une tient le modèle pour donné tandis que l'autre tient les variables physiques pour données. Dès lors, pour justifier nos assertions, il nous faut déterminer lesquels des modèles ou des variables physiques priment. Une telle réflexion dépasse le cadre des mathématiques et nous renvoie

²⁷ C'est-à-dire stochastiques, bornées... Nous développerons cet aspect plus loin.

aux fondements même de la physique qui prennent leurs sources dans la métaphysique sur laquelle depuis longtemps les philosophes se penchent. Pour étayer nos assertions, nous avons commencé par étudier comment les différents philosophes avaient envisagé la représentation de la réalité physique afin de mettre en évidence les rapports qu'entretenaient chacune des deux conceptions avec les différents modes de penser la réalité physique. De la sorte, nous pourrions déterminer les implications de chacun des modes et débattre de leur validité. Dans un second temps, nous prolongerons l'étude au problème de la constitution du réel sur lequel repose l'approche que nous proposons.

Représentation de la réalité physique

Dès l'origine de la philosophie, comme le souligne A. Koyré [Koyré,66], le problème de la représentation de la réalité physique apparaît. La pensée d'Aristote se caractérise par une décomposition de la nature en natures élémentaires qui respectent des principes d'ordre en vertu desquels l'ensemble des êtres réels forme un tout (le cosmos) naturellement bien ordonné. La pensée d'Aristote est empiriste car l'ordre n'y est révélé que par l'expérience ; mais, les principes ne dépassent pas le cadre du qualitatif. Galilée, physicien et géomètre, affirme que ces principes ou lois sont géométriques, il énonce : *le grand livre de la nature est écrit en caractères géométriques*. Malgré la modernité de la pensée Galiléenne, elle n'aborde cependant pas le thème de la constitution du réel ; elle se cantonne aux faits et, en ce sens, elle est aussi empiriste. Descartes (1637) prolonge les travaux de Galilée en conservant l'idée de lois mathématiques (mécaniques) et aborde le problème de la constitution du réel. Par la pratique systématique du doute, et notamment en ce qui concerne les sens, il en vient à formuler l'axiome apodictique selon lequel *tout ce que mon esprit conçoit clairement et distinctement de réel et de vrai, et qui contient en soi quelque perfection, est contenu et renfermé tout entier dans cette idée (Dieu)*. Descartes désire donc ordonner le réel et lui conférer une unité. La physique cartésienne va donc nécessairement devoir procéder de manière déductive et reconstruire l'ensemble de la réalité physique dans un enchaînement universel de causes où les expériences ne viendront que pour combler les lacunes de la théorie : orientant la déduction, elles permettent de séparer le réel du possible, de choisir, quelquefois, entre plusieurs déductions cohérentes [Tournandre,82]. C'est en ce sens que Descartes n'est pas un empiriste puisque la cause première de la nature est Dieu qui est au-delà de toute expérience tout en étant en toute expérience : Dieu est en tout et tout est en Dieu. Les lois mathématiques trouvent leur unité en Dieu et ne sont que les manifestations de la loi divine : la loi est donc a priori. La pensée scientifique de Descartes pose un problème à Newton ; elle se veut universelle, mathématique et déductive et c'est précisément ce dernier aspect qui lui paraît problématique. Toute science suppose en effet des arguments supportant la critique ; or, Descartes donne pour argument Dieu dont la contradiction ne conduit qu'à des

polémiques aporétiques impropres à la science. Newton revient donc à l'empirisme de Galilée en affirmant *les connaissances ne peuvent provenir que de l'expérience*. Il refuse la constitution de la réalité physique de Descartes en rejetant la métaphysique, l'occulte et même la mécanique cartésienne de la représentation du réel ; pour Newton, les modèles sont mathématiques et s'étendent jusqu'à des entités corpusculaires. Il développe une conception phénoméniste de la réalité physique conforme à son empirisme [Parrochia,91]. Naturellement, la raison première des phénomènes échappe à sa physique qui ne s'occupe que des lois.

Constitution de la réalité physique

C'est avec la parution en 1781 de la *Critique de la raison pure* que Kant apporte une contribution majeure au problème de la constitution du réel. Kant adopte une approche scientifique en s'appuyant sur des arguments qui éludent toute contradiction. Se souvenant de la pratique systématique du doute de Descartes qui tenait les objets pour donnés, il annonce : *on a admis jusqu'ici que toutes nos connaissances devaient se régler sur les objets ; mais, dans cette hypothèse, tous nos efforts pour établir à l'égard de ces objets quelque jugement a priori et par concept qui étendit notre connaissance n'aboutit à rien. Que l'on cherche une fois si nous ne serions pas plus heureux dans les problèmes de la métaphysique, en supposant que les objets se règlent sur notre connaissance, ce qui s'accorde déjà mieux avec ce que nous désirons démontrer, à savoir la possibilité d'une connaissance a priori de ces objets qui établisse quelque chose à leur égard avant même qu'ils nous soient donnés*.

Kant se range du côté de Newton et le dépasse même en affirmant que la réalité physique n'est faite que de connaissances alimentées par l'expérience. C'est-à-dire qu'en plus de donner l'expérience comme seul moyen de connaître la réalité, il restreint la réalité physique à la connaissance issue de l'expérience. Or, *quand nous affirmons qu'il n'y a de connaissance que des phénomènes (par l'expérience), c'est-à-dire des êtres sensibles, des êtres tels qu'ils apparaissent à notre sensibilité, nous ne pouvons nous empêcher de penser soit à la nature de ces êtres telle qu'elle est indépendamment de notre façon de percevoir, soit à d'autres êtres possibles qui ne seraient pas des objets de nos sens. Certes l'entendement n'a pas le pouvoir de dépasser la sensibilité, en ce qui concerne la connaissance, mais ce que l'on ne peut connaître peut cependant être pensé, et c'est ainsi que nous cherchons autre chose que l'expérience [Pascal,95]. Les noumènes sont précisément ces objets que nous pensons au-delà des phénomènes, et ces objets nous les considérons comme des objets simplement conçus par l'entendement, et nous les appelons êtres intelligibles (Noumena)*. Mais il faut prendre garde à ne pas considérer ces objets de pensée comme des objets de connaissance et à ne pas faire du monde intelligible un monde que nous connaîtrions avec notre entendement. Le passage du sensible à l'intelligible, dans le domaine de la connaissance, ne nous élève pas au-dessus des objets des sens : *ce que nous appelons noumène ne doit donc être entendu qu'au sens négatif*.

Le problème de l'ordonnement de la réalité physique n'est pas encore résolu. Il s'agit de déterminer comment s'ordonne cette connaissance. En premier lieu, Kant pose le phénomène qu'il ne conçoit pas, contrairement à Descartes, comme une apparence. Il lui confère, au contraire, une réalité, mais une réalité qu'il faut chercher dans le rapport de l'objet au sujet et non dans l'objet lui-même : *le phénomène est quelque chose qu'il ne faut pas chercher dans l'objet lui-même, mais toujours dans le rapport de cet objet au sujet et qui est inséparable de la représentation que nous en avons. Si toute connaissance commence avec l'expérience, il n'en résulte pas qu'elle dérive toute de l'expérience ; en effet, nous ne connaissons a priori des choses que ce que nous y mettons nous-mêmes.* Kant décompose la connaissance en deux grandes modalités : l'une sensitive, l'autre intelligible elle-même rapportée à l'entendement et à la raison ; chacune d'elle est constituée d'une forme a priori, c'est-à-dire nécessaire, et d'une matière a posteriori. La forme n'est donc rien moins qu'une matrice a priori, réceptacle de toute information, et c'est là l'explication que propose Kant pour justifier la cohérence de nos connaissances. Il associe les intuitions à la fraction sensitive de la connaissance et les concepts de l'entendement (le divers d'un concept appartient au domaine sensitif) ainsi que les lois de la raison (le divers d'une loi appartient au domaine des concepts) à la fraction intelligible : *Intuitions et concepts, tels sont donc les éléments de toute notre connaissance, de telle sorte que ni les concepts sans une intuition qui leur corresponde de quelque manière, ni une intuition sans les concepts ne peuvent fournir aucune connaissance.*

Nous ne débattons pas plus de ce qui, dans la connaissance, est a priori et de ce qui est a posteriori ; ce problème appartient à la métaphysique et dépasse le cadre de ce mémoire. Nous nous restreindrons au problème de la constitution de la réalité physique de manière à faire ressortir les entités fondamentales de la représentation. Le système physique tel que nous l'avons défini précédemment est, dans un langage Kantien, une noumène ; en ce sens, il n'appartient pas au champ de la connaissance mais à celui de la croyance. Il n'existe qu'en tant qu'entité intelligible et n'a l'intérêt que de conférer une unité à un système de phénomènes interdépendant ou processus. La fraction sensitive est le niveau premier de la connaissance, c'est-à-dire celui sur lequel repose tous les autres. Les entités qui se rapportent à ce niveau sont les phénomènes ou, dans les cas qui nous intéressent, les observations quantifiables, qui peuvent être rassemblées en processus. Il est important de remarquer que ce niveau est indépendant de ceux qui lui succèdent, autrement dit, que les phénomènes se passent de concepts autant que de lois ou de principes. En second lieu est la fraction intelligible rapportée à l'entendement de la connaissance à laquelle Kant rattache les concepts, que nous appellerons, dans le cas des systèmes physiques, les variables physiques. Ces variables physiques se rapportent nécessairement à un divers sensitif, c'est-à-dire à des phénomènes et non pas à des variables abstraites intervenant dans des lois. Les variables stochastiques représentant l'imprécision d'une loi ne sont, par exemple, pas à classer parmi les variables

physiques. Enfin, en dernier lieu, est la fraction intelligible rapportée à la raison dont les entités constitutives sont les lois ou les principes. Nous avons préféré l'expression *relation comportementale* parce qu'elle précise celle de loi lorsqu'il s'agit d'étudier les systèmes physiques. On rencontre bien plus fréquemment le terme de modèle mais nous préférons l'éviter pour un usage systématique parce qu'un pouvoir explicatif nuisant à la généralité lui est souvent adjoint. Ces relations mathématiques en ce qui nous concerne sont constituées d'opérateurs mathématiques, de paramètres et de variables abstraites et, évidemment de variables physiques. La constitution de la réalité physique appliquée au cas des systèmes physiques telle que nous venons de la décrire est synthétisée par la figure XI.

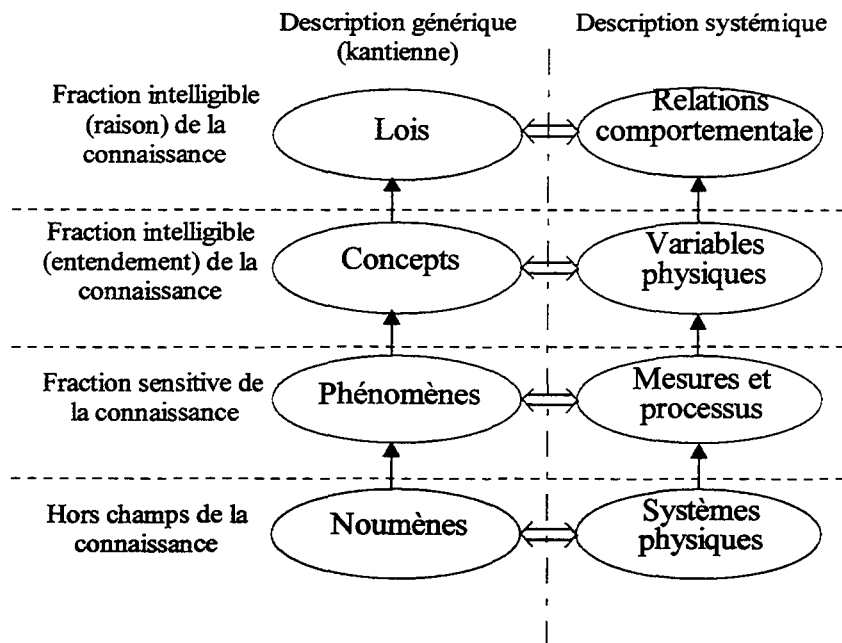


Figure XI - Constitution de la connaissance de la réalité physique

Le XX^{ème} siècle a été marqué par l'apparition de la théorie de la relativité et de la mécanique quantique qui ont remis en cause un certain nombre de points de la philosophie kantienne. Einstein a montré, dans la théorie de la relativité, que le temps et l'espace ne pouvaient pas toujours être considérés comme indépendants ; mettant ainsi en défaut les formes a priori de la sensibilité. La théorie des quanta montre que la causalité n'est pas une forme a priori de l'entendement et énonce que la validité de certains modèles est fonction des conditions d'observation des phénomènes auxquels ils se rapportent. Ainsi, si la distribution kantienne entre la forme (a priori) et la matière (a posteriori) de la connaissance de la réalité physique a été remise en cause, la constitution de la réalité physique que nous avons évoquée précédemment conserve aujourd'hui toute sa pertinence et, elle est même corroborée par les théories modernes qui font un pas de plus, passant de l'absolutisme kantien au relativisme en associant aux lois certains domaines de validité. En 1962, Heisenberg écrit : *la conception de la réalité objective des particules élémentaires s'est donc étrangement dissoute non pas dans le brouillard d'une nouvelle conception de la réalité, obscure et mal comprise, mais dans la*

clarté transparente d'une mathématique qui ne représente plus le comportement de la particule élémentaire mais la connaissance que nous en possédons [Heisenberg,62].

En nous référant à la métaphysique, nous avons établi que les phénomènes qui, par définition, sont les objets d'observations potentielles, primaient sur les variables physiques qui n'en sont que la représentation ainsi que sur les modèles ou relations comportementales qui formalisent des rapports entre phénomènes. Ainsi, il apparaît de façon immédiate et sans équivoque que la détermination des anomalies comportementales ne peut résulter que d'un rapprochement entre une fraction sensitive et une fraction intelligible de la connaissance, autrement dit, d'un test de cohérence entre des observations et des relations comportementales. Cette assertion ne peut être niée et, parce qu'elle est incontournable, nous concluons à son caractère universel.

IV.2. Représentation des systèmes incertains

Un système incertain est un système physique dont le comportement peut être décrit par des relations comportementales parfois imprécises et dont les phénomènes sur lesquels elles portent, sont quantifiables. Cette sous-partie est consacrée à l'étude des particularités des modèles de ces systèmes qui interviennent dans la description des procédures de diagnostic.

Modèle complet pour le diagnostic

Travé-Massuyès et al [Travé et al,97] soulignent qu'il existe deux types d'*ontologies* (concepts élémentaires) propres à la modélisation des systèmes physiques : les composants et les phénomènes. Or, un composant appartient aux noumènes de Kant c'est-à-dire que c'est une entité idéale dont l'unité est donnée par le rattachement à un domaine spatial fini d'un ensemble de phénomènes interdépendants. Ainsi, chaque composant est donné par un ou plusieurs processus auxquels se rattachent des relations comportementales. En conséquence, chaque relation comportementale peut être associée à un ou plusieurs composants. L'intérêt de cette association provient de ce que les composants sont interchangeables ; il est intéressant, après avoir recensé les comportements anormaux, de déterminer les composants à changer pour recouvrer un comportement normal. Ce niveau de représentation, que l'on pourrait qualifier de niveau composant intervient à un niveau supérieur du diagnostic qui ne procède que d'une simple inférence des relations comportementales anormales vers les composants associés. Les entités fondamentales d'un modèle pour le diagnostic sont représentées sur la figure XII. En rapprochant les niveaux de raisonnement du diagnostic d'après Milne (§III.1 figure II) de la figure XII, on constate un certain parallélisme entre les niveaux structurel et comportemental. On peut néanmoins s'étonner de ne pas retrouver le niveau fonctionnel. La raison en est que nous ne distinguerons pas les défauts des défaillances ou le comportement du fonctionnement puisqu'une fonction est aussi une relation comportementale abstraite selon la téléologie d'un système physique. S'il avait fallu

représenter les fonctions sur la figure XII, elles auraient été représentées par un carré parmi les relations comportementales.

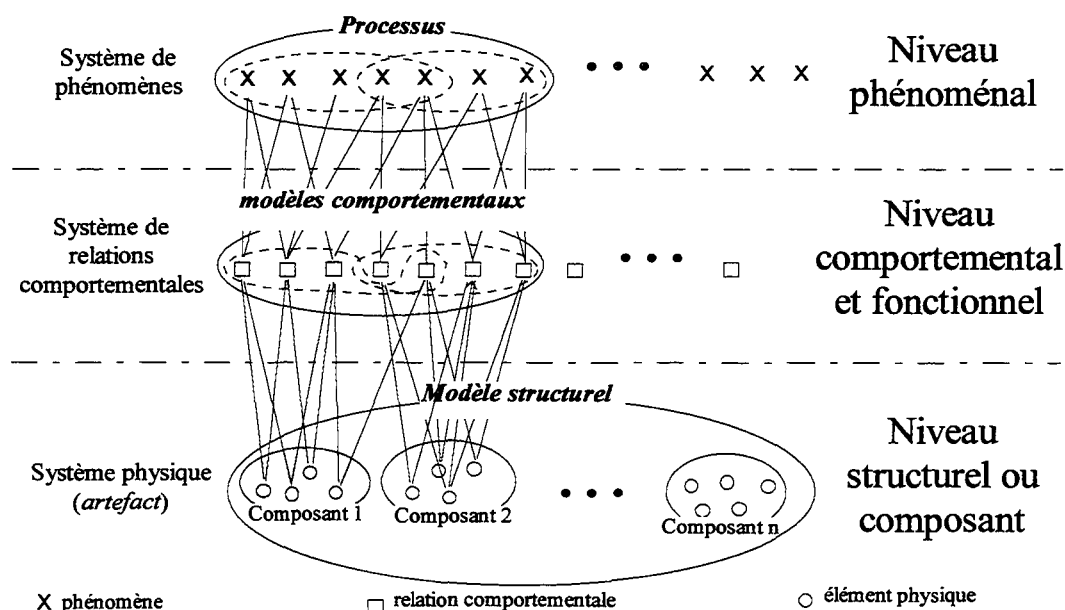


figure XII - Les différents niveaux de représentation d'un modèle complet pour le diagnostic

Du déterminisme à l'indéterminisme

Les systèmes incertains sont caractérisés par le fait qu'il n'existe pas de relations comportementales déterministes suffisamment précises pour représenter exactement les comportements possibles. Ces cas sont pratiquement très fréquents parce qu'il est rarement possible d'établir des modèles de connaissance précis des systèmes physiques et, dans le cas contraire, ils sont souvent trop complexes et on leur préfère des modèles simplifiés : une trop grande complexité conduit à une sur-paramétrisation qui pose des problèmes souvent insolubles d'identification ainsi que de temps de calcul. Grâce au calcul de champs, des modèles très précis des machines électriques peuvent être obtenus (bien qu'ils ne représentent pas encore la totalité des phénomènes présents) ; cependant, la complexité de tels modèles ainsi que leur nombre colossal de paramètres en regard avec le peu de variables physiques mesurables en interdit l'usage pour le diagnostic. Les modèles très précis tendent à descendre à l'échelle microscopique pour expliquer l'échelle macroscopique ; ils n'en demeurent pas moins intéressants pour ce qui est de la compréhension des systèmes physiques mais les modèles macroscopiques simplifiés, dont le nombre de paramètres est adapté aux variables physiques mesurables, est bien plus utile au diagnostic. Une donnée pratique du diagnostic est que les relations comportementales exploitables n'auront généralement qu'une précision finie.

La partie III nous a donné l'occasion de montrer à quel point le déterminisme²⁸ s'avérait inadapté à la représentation des systèmes incertains. Dans ce contexte, il est erroné de qualifier une relation comportementale de vraie ou de fausse. Il est nécessaire de substituer à ces attributs ceux de précis et d'imprécis²⁹ conduisant à remplacer une conception dichotomique de la représentation par une conception nuancée. L'usage des probabilités a permis de formaliser ces nuances de précision. En effet, une variable stochastique est déterminée par une fonction de répartition donnant les probabilités d'occurrence des valeurs qui lui sont associées. Ainsi, plus la variance ou l'écart-type d'une variable stochastique est faible, et plus sa détermination est précise. Il faut citer les techniques de calcul à variance minimale dont est issue la technique de filtrage de Kalman qui reposent sur des modèles sous forme d'équations différentielles ordinaires ou récurrentes et qui, grâce à un terme additif supplémentaire contenant des variables stochastiques, permettent de déduire non seulement les valeurs moyennes de certaines variables physiques, mais aussi d'en évaluer la précision.

Pour apprécier l'intérêt d'indices de précision pour le diagnostic, revenons au principe de cohérence : toute procédure de diagnostic résulte de tests de cohérence entre des modèles et des observations. Or, la conclusion de tels tests est par nature dichotomique : cohérent ou incohérent. Par conséquent, l'information contenue dans la nuance d'un indice de précision est a priori trop riche pour le diagnostic : une information dichotomique suffit. Nous avons emprunté à l'instrumentation l'attribut *juste* et son contraire *faux*. Si un capteur est donné avec une précision de $\pm 10\%$ et si la mesure d'étalons conduit à des valeurs comprises dans l'intervalle de précision donné, le capteur sera considéré comme juste sinon comme faux. Nous voyons ici que d'un indice nuancé de précision, il est possible d'extraire l'information dichotomique de justesse essentielle au diagnostic. Si l'adéquation d'un modèle avec des observations est juste, on conclura à la cohérence, sinon à l'incohérence. Struss [Struss,92] a commencé à formaliser cet attribut ; il propose de représenter la loi d'Ohm de la façon suivante :

$$\left\{ (u, i, r) \in \mathbb{R}^3 : \frac{u}{i} \in [0.9r, 1.1r] \right\}$$

La justesse est caractérisée par l'intervalle $[0.9r, 1.1r]$; si le rapport de la tension sur le courant appartient à cet intervalle, alors le test de cohérence mène à un résultat positif. Cet exemple montre que la notion de justesse peut être formalisée directement sans avoir recours à la nuance d'un indice de précision. La représentation de Struss est cependant limitée à des cas triviaux où un seul intervalle apparaît dans une relation comportementale. La notion de

²⁸ Principe scientifique suivant lequel les conditions d'existence d'un phénomène sont déterminées, fixées absolument de telle façon que, ces conditions étant posées, le phénomène ne peut pas ne pas se produire

²⁹ Confère théorie de l'instrumentation.

variable abstraite qui s'applique autant à des variables bornées³⁰ qu'à des variables stochastiques, et dont la vocation est la représentation de tout type d'indétermination, permet d'intégrer les incertitudes à la représentation des systèmes physiques. Signalons que certaines variables abstraites ont un statut particulier : il s'agit des variables abstraites d'occurrence propres aux modèles MF. La nullité d'une telle variable correspond à une absence de défaut. Si θ est une variable abstraite d'occurrence du $i^{\text{ème}}$ modèle MF alors $\theta=0$ représente l'absence du $i^{\text{ème}}$ défaut. Nous ne développerons pas plus ce sujet car il fait l'objet des chapitres suivants de ce mémoire.

Nature des modèles des systèmes incertains

De nombreux types de modèles apparaissent dans le diagnostic des systèmes incertains, qu'il s'agisse de variables physiques, de relations comportementales, de modèles réunissant un ensemble de relations comportementales associées à un processus ou de modèles utilisés pour des tests de cohérence. Pour mieux comprendre les imbrications existant entre ces différentes notions, nous proposons d'introduire de nouveaux termes caractérisant chacun des types de modèle.

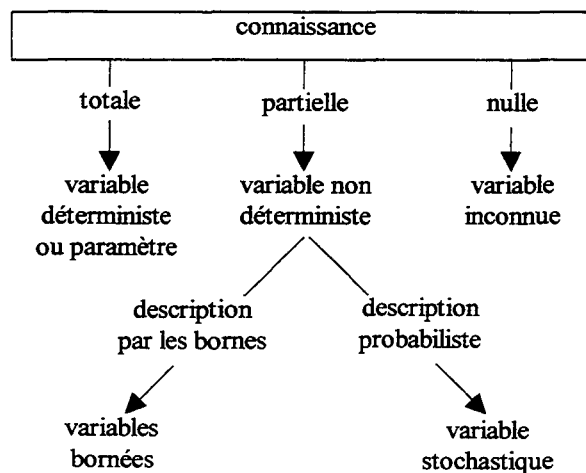


Figure XIII - Types de connaissance et natures des variables associées

Toutes les grandeurs intervenant dans les modèles n'ont pas la même nature ; il est utile de les distinguer car elles peuvent avoir des rôles très différents. Trois natures distinctes peuvent être isolées pour caractériser différents types de connaissance. Dans le cas commun au déterminisme d'une connaissance totale, nous emploierons les expressions de *variable déterministe* ou de *paramètre* lorsque la connaissance représentée est constante. Par ailleurs, une *variable non déterministe* peut être, sans restriction, une variable stochastique ou une variable bornée ; elle représente assurément une connaissance partielle. Il ne nous reste plus qu'à aborder le cas de la représentation d'une absence de connaissance auquel nous réservons

³⁰ variables non déterministes définies sur un domaine borné connu

l'expression *variable inconnue*. Un lecteur assidu pourrait s'étonner, à juste titre, du rôle de telles variables dans la modélisation, nous pourrions néanmoins le rassurer en lui assurant que la représentation d'une connaissance ignorée consiste déjà en une forme de connaissance. Plus pratiquement, nous verrons un peu plus loin comment des variables physiques peuvent conduire à des variables inconnues. La figure XIII ordonne les différentes natures de variables que nous venons de décrire.

Les rapports entre les phénomènes sont représentés, au niveau le plus élémentaire, par des *relations comportementales élémentaires (relations CE)* provenant généralement de lois de la physique. Ces relations comportementales contiennent, dans le cas des systèmes incertains, des variables abstraites en plus des variables physiques ; elles constituent les éléments de base de la modélisation. Les variables physiques seront représentées entre crochets afin de manifester la nature formelle de telles variables. En effet, une variable physique ne représente rien de plus que la connaissance potentielle d'un phénomène. Les seuls moyens d'accéder effectivement à une telle connaissance seraient soit l'observation, soit la déduction, et ce ne serait qu'à ce stade que la connaissance du phénomène deviendrait effective. La connaissance potentielle d'une tension pourrait être notée $[u]$, son observation de nature nécessairement déterministe $\tilde{u}(t)$ et sa déduction $\hat{u}(t)$ qui peut conduire à une connaissance totale ou partielle du phénomène.

Il nous reste à éclaircir ce que nous entendons par niveau le plus élémentaire de représentation. Prenons l'exemple d'un circuit électrique constitué d'une résistance r et d'une inductance l en série (figure XIV).

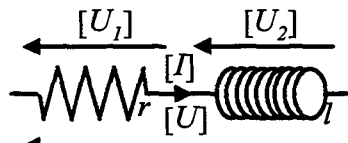


Figure XIV - Circuit RL série

S'il est intéressant de faire apparaître les variables physiques $[U_1]$ et $[U_2]$, le processus associé au circuit peut être représenté par trois relations comportementales élémentaires issues des lois de l'électricité :

$$\left\{ [U_1] = r[I], [U_2] = l \frac{d}{dt}[I], [U] = [U_1] + [U_2] \right\}$$

Les variables physiques sont naturellement des fonctions du temps bien que nous ne le représentions pas.

Lorsque les variables physiques $[U_1]$ et $[U_2]$ ne peuvent être mesurées séparément. Les relations CE précédentes peuvent être agrégées en une seule relation CE :

$$[U] = r[I] + l \frac{d}{dt}[I].$$

Un cas particulier intéressant est celui des équations d'observation représentant la mesure d'une variable physique par un capteur. Une relation CE particulière permet de représenter une mesure avec, éventuellement, ses erreurs inhérentes. Pour représenter un capteur mesurant la tension $[U]$ avec une précision de 10%, on pourra utiliser la relation CE suivante :

$$\tilde{U} = (1 + \theta)[U] : \theta \in [-0.1, 0.1]$$

\tilde{U} représente la mesure de la tension $[U]$

Il nous a semblé utile d'introduire la notion de *modèle formel* ou de *métamodèle* résultant de la concaténation de la particule grecque *meta*, exprimant dans son acception philosophique l'idée de *ce qui dépasse et englobe*, et de *modèle*. Nous définissons le métamodèle comme étant la réunion des relations comportementales élémentaires participant à la représentation du comportement global d'un processus. Il comprend, en plus des relations CE représentant les liens entre les phénomènes d'un processus, les relations CE d'observation. Reprenons l'exemple du moteur à courant continu décrit par les équations (1) et (2) et pour lequel la tension est connue à 5% près et la vitesse de rotation est mesurée avec une précision égale à 10%. Le métamodèle peut s'écrire de la façon suivante :

$$\left\{ \begin{array}{l} \Gamma \frac{d[\Omega]}{dt} = -f[\Omega] + \Phi[i], L \frac{d[i]}{dt} = -R[i] - \Phi[\Omega] + [u], \\ \tilde{\Omega}(t) = (1 + \theta_{\Omega})[\Omega] : \theta_{\Omega} \in [-0.1, 0.1], \tilde{u}(t) = (1 + \theta_u)[u] : \theta_u \in [-0.05, 0.05] \end{array} \right\} \quad (4)$$

Il faut remarquer que le métamodèle se distingue du modèle au sens usuel par le fait qu'il contient des variables de nature formelle dénotant une connaissance potentielle mais non effective des phénomènes ; aussi, sous cette forme, il n'admet nulle réification³¹ et ne peut donc être utilisé tel quel dans une procédure de diagnostic.

Il est alors naturel d'introduire la notion de *modèle réifié* ou d'*hypomodèle* qui vaut pour tout modèle réifiable, c'est-à-dire ne contenant plus de variables formelles : elles doivent être considérées comme inconnues. Le préfixe *hypo* exprime l'idée de diminution et d'infériorité qui vient de ce qu'un hypomodèle ne peut se déduire que par élimination des composantes inconnues d'un métamodèle. L'*hypomodèle* n'a donc pas pour vocation de représenter formellement une réalité physique mais d'en exprimer les propriétés vérifiables. C'est en ce sens qu'ils sont essentiels à la synthèse de tests de cohérence et donc, au diagnostic. Dans le cas des systèmes certains, un hypomodèle se déduit facilement d'un métamodèle, mais la déduction n'est pas unique puisque toute fraction réifiable d'un hypomodèle est elle-même un

³¹ Action de transformer en chose, de donner un caractère de chose ; ici, action de rendre déterministe c'est-à-dire de conduire à quelque chose de directement exploitable par une procédure de diagnostic.

hypomodèle et qu'il y a autant d'hypomodèles que de manière de formuler les propriétés vérifiables que le métamodèle recouvre.

Il est possible d'établir une classification sommaire des hypomodèles à partir de trois propriétés. La première est la causalité. Les hypomodèles seront qualifiés de causaux, si la connaissance de certaines variables physiques est déduite de la connaissance d'autres. Les premières font office, dans un hypomodèle, d'effets ou de sorties alors que les secondes, de causes ou d'entrées : cette propriété est valable aussi bien pour les *modèles parallèles* que pour les *observateurs d'état*. Un hypomodèle qui n'est pas porteur d'une causalité implicite sera qualifié d'acausal : c'est typiquement le cas des relations de parité.

La seconde propriété discriminante est l'horizon temporel d'un hypomodèle, selon qu'il soit fini ou infini. Un hypomodèle à horizon fini est un modèle pour lequel, au delà d'un certain horizon ou laps de temps, les phénomènes s'étant produits n'ont plus aucune influence sur les résultats de l'instant présent. C'est le cas des *relations de parité* dans le cas acausal mais aussi des *modèles parallèles* causaux sans récurrence. A l'inverse, un hypomodèle à horizon infini est un modèle causal comprenant une récurrence ou une intégration en son sein.

Enfin, la dernière propriété, propre aux systèmes dynamiques, est la présence ou non d'une contre-réaction qui n'a de sens que pour les hypomodèles causaux à horizon infini. Un hypomodèle causal contiendra une contre-réaction (cf. figure XV) si une ou plusieurs variables physiques interviennent à la fois parmi les causes et parmi les effets c'est-à-dire si des variables physiques sont à la fois explicatives (observées) et expliquées (estimées). Un des intérêts de cette distinction est de définir d'une façon générique les observateurs d'état autrement que par des relations mathématiques casuelles qui seront abordées dans le chapitre 4.

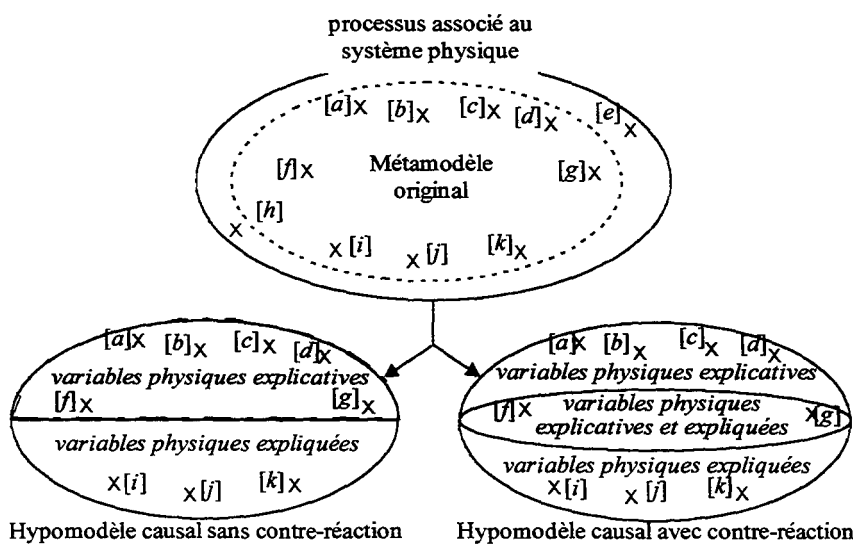


Figure XV - Organisation des variables physiques dans les hypomodèles causaux (métamodèle dynamique)

Un hypomodèle causal à horizon infini sans contre-réaction est un modèle parallèle à horizon infini. Le tableau IV résume ce classement ainsi que les dénominations usuelles qui s'y

Hypomodèle	Horizon fini		Horizon infini
	acausal	causal	
sans contre-réaction	relation de parité	modèle parallèle (horizon fini et infini)	
avec contre-réaction			observateur d'état

Tableau IV - Classification des hypomodèles

rapportent : quatre espèces d'hypomodèles se distinguent clairement.

Pour illustrer ces différents types d'hypomodèles, partons du métamodèle d'un système monovariable du premier ordre dont toutes les variables physiques sont mesurées parfaitement :

$$\left\{ \begin{array}{l} [y_{k+1}] = a[y_k] + b[u_k] : |a| \leq 1 \\ \forall k, \tilde{y}_k = [y_k] \text{ et } \tilde{u}_k = [u_k] \end{array} \right\} \quad (5)$$

De ce métamodèle, nous pouvons extraire la relation de parité simple suivante :

$$\tilde{y}_{k+1} - a\tilde{y}_k - b\tilde{u}_k = 0 \quad (6a)$$

ainsi que la relation redondante sur un horizon plus grand :

$$\begin{aligned} \tilde{y}_{k+2} + (\lambda - a)\tilde{y}_{k+1} - a\lambda\tilde{y}_k - b\tilde{u}_{k+1} - \lambda b\tilde{u}_k &= 0 \\ = (\tilde{y}_{k+2} - a\tilde{y}_{k+1} - b\tilde{u}_{k+1}) + \lambda(\tilde{y}_{k+1} - a\tilde{y}_k - b\tilde{u}_k) \end{aligned} \quad (6b)$$

avec $\lambda \in \mathbb{R}^*$

Notons que cette relation présente un intérêt mineur parce que son résidu n'est rien moins que la somme pondérée de deux résidus consécutifs de la relation (6a). Ces hypomodèles sont bien entendu sans causalité et à horizon fini puisqu'un événement survenu au delà des deux ou trois instants consécutifs couverts par ces hypomodèles n'a aucune incidence sur le résidu, c'est-à-dire sur le résultat du test de cohérence. Sur un même horizon que l'hypomodèle (6a), un hypomodèle parallèle causal peut s'écrire :

$$\hat{y}_{k+1} = a\tilde{y}_k + b\tilde{u}_k \quad (7)$$

Ici aussi, en sommant la relation (6a) retardée de 1 au terme de droite, nous obtenons un modèle redondant d'un horizon plus important. Un hypomodèle causal à horizon infini sans contre-réaction s'écrit :

$$\begin{cases} \hat{y}_{k+1} = a\hat{y}_k + b\tilde{u}_k \\ x_0 \text{ donné} \end{cases} \quad (8)$$

Un tel modèle fait intervenir toutes les mesures \tilde{u}_k de l'instant 0 à l'instant k . Enfin, l'observateur d'état ou, selon la classification, l'hypomodèle causal à horizon infini avec un boucle de contre-réaction, pourrait s'écrire :

$$\begin{cases} \hat{y}_{k+1} = a\hat{y}_k + b\tilde{u}_k + \lambda(\tilde{y}_k - \hat{y}_k) \\ x_0 \text{ donné} \end{cases} \quad (9)$$

avec $\lambda \in \mathbb{R}^$*

Remarquons qu'en confondant les notions de métamodèle et d'hypomodèle, il aurait été difficile de marquer une différence entre les expressions (5), (7) et (8). Nous étudierons plus en détail les différences entre les hypomodèles dans la partie 3 de ce document.

IV.3. Approches fondamentales du diagnostic des systèmes incertains

Les différents éléments de la modélisation des systèmes incertains ayant été définis, nous pouvons désormais nous consacrer aux spécificités des procédures de diagnostic adaptées à ces systèmes. Deux grandes approches de diagnostic des systèmes incertains se détachent : l'approche a posteriori, compensant a posteriori les incertitudes d'un modèle, et l'approche a priori prenant en compte dans les relations comportementales les incertitudes de représentation.

Approches a posteriori du diagnostic des systèmes incertains

L'approche a posteriori se contente de relations comportementales déterministes représentant des comportements moyens. Il s'ensuit que tous les phénomènes négligés ainsi que les relations CE déterministes approchées sont à l'origine de perturbations qui dégradent la fiabilité des tests de cohérence (partie gauche de la figure XVI) : une incohérence peut être due indistinctement à des anomalies comportementales ou à des perturbations. Ces approches peuvent être améliorées lorsque l'influence des perturbations sur les résultats des tests reste faible en effectuant un filtrage adapté de ces résultats, ou résidus, pour obtenir, au moyen d'un seuillage, des indices de cohérence peu sensibles aux perturbations. Les procédures a posteriori restent cependant approximatives et conduisent inévitablement à des fausses alarmes et à des cas de défauts non détectés.

Remarque

Il est souvent problématique d'aborder l'incertain avec des approches déterministes. Prenons l'exemple d'une variable abstraite θ dont les valeurs possibles appartiennent à l'intervalle $[0 \ 2]$, de valeur centrale 1. Imaginons qu'un modèle fasse intervenir le carré de cette variable : θ^2 . Alors, nous déduisons que les valeurs que peut prendre cette nouvelle variable appartiennent à $[0 \ 4]$, dont la

valeur centrale est 2. Or, une approche déterministe réduirait la variable θ à sa valeur centrale ; ainsi, la valeur de θ^2 serait évaluée à 1 alors que la valeur centrale est ici égale à 2 !

Lorsque les relations comportementales représentent le bon fonctionnement, certains auteurs [Gertler,88] préconisent de tirer parti du fait que le comportement normal moyen soit éliminé des résultats des tests de cohérence, ou résidus, pour y reconnaître des caractéristiques ou signatures de défauts. Cette stratégie se décompose en deux niveaux a posteriori ; le premier effectue des tests de cohérence avec des modèles BF alors que le second réalise des tests comparatifs entre les résidus issus des observations et des signatures de défauts qui sont des modèles MF. Les résultats des tests de cohérence indicés 'a' sur la figure XVI sont naturellement fonction autant des perturbations que des défauts ; par conséquent, les résultats des tests indicés 'b' sont tributaires des perturbations. Les indices de cohérence 'b' conduisent donc, eux aussi, à des cas de non détection et à des fausses alarmes.

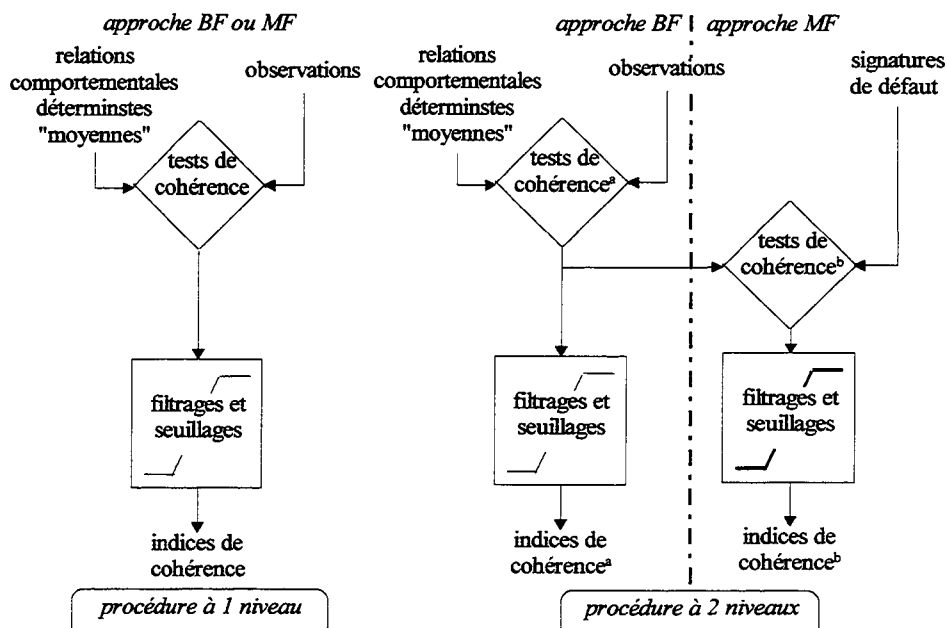


Figure XVI - Approche a posteriori du diagnostic des systèmes dynamiques incertains

Les procédures de diagnostic a posteriori, qu'elles soient orientées BF ou MF, sont d'autant moins fiables que les perturbations sont importantes ce qui ne les prédestinent guère au diagnostic des systèmes incertains. Cependant, dans des cas triviaux où les signatures se manifestent dans les résidus par des biais tandis que les perturbations n'engendrent que des variations à valeurs moyennes nulles, de telles méthodes peuvent être utilisées avec un certain succès.

Approches a priori du diagnostic des systèmes dynamiques incertains

Les données générales d'un problème de diagnostic a priori peuvent être formulées d'une façon générale. Tout d'abord, la donnée la plus simple, et la plus fréquente aussi, est le

métamodèle BF décrivant le comportement normal. Vient alors s'ajouter un nombre fini de métamodèles MF représentant des comportements anormaux. Cependant, il serait saugrenu de considérer que ces métamodèles représentent l'ensemble des comportements anormaux possibles parce que d'une part, tous les défauts ne peuvent pas toujours être représentés par des métamodèles et d'autre part, recenser la totalité des défauts pouvant affecter un système physique est souvent irréalisable.

Quelles que soient les approches a priori de diagnostic, elles reposent toutes sur des métamodèles intégrant des incertitudes de représentation au moyen de variables abstraites. Leur synthèse, qu'elles soient BF ou MF, commence par la détermination de métamodèles justes et aussi précis que possible. En effet, la justesse, qui confère la fiabilité, nous semble être une condition impérative du diagnostic en général. Si un métamodèle, aussi précis soit-il, est faux dès l'origine, toute procédure de diagnostic, qui procède nécessairement à des tests de cohérence, sera irrémédiablement fautive et conduira naturellement à de fausses alarmes et à des cas de défauts non détectés. Certes, on s'en est souvent contenté jusqu'ici, et certains chercheurs ont ramené le diagnostic à un problème de minimisation de taux de fausses alarmes et de non détection. Cependant, n'est-ce pas une gageure que de vouloir déterminer ces taux dans l'absolu puisqu'une condition infrangible de tout problème de diagnostic est que l'ensemble des défauts possibles n'est pas connu de façon exhaustive. D'autre part, les perturbations dues à la fausseté d'un métamodèle conduisent à des résidus dépendant des variables physiques, fait préjudiciable à la détermination absolue de tels taux. Seules des expériences a posteriori en permettent l'évaluation, mais les résultats sont évidemment contextuels et n'ont guère de sens hors du cadre de l'expérience. La condition de justesse nous semble devoir être tenue pour nécessaire ; la précision ne détermine que la performance d'une procédure de diagnostic.

Nous en sommes venus à conclure que, pour produire des tests de cohérence, une procédure de diagnostic ne pouvait exploiter que des propriétés vérifiables exprimées par des hypomodèles. Or, si les hypomodèles ne contiennent plus d'inconnues, ils n'en contiennent pas moins, surtout dans le cas des systèmes incertains, des variables abstraites non déterministes. La particularité de la synthèse de procédures de diagnostic a priori est de devoir réaliser des tests de cohérence à partir de modèles comportant des incertitudes : les variables abstraites non déterministes. Nous n'aborderons pas ici le détail des tests de cohérence se fondant sur des hypomodèles incertains car c'est là l'objet de la suite de ce mémoire.

Génération de résidus robuste

Une méthode d'appréhension des systèmes incertains apparaît fréquemment dans la littérature : la robustesse des résidus vis-à-vis d'incertitudes. Le principe est de construire des hypomodèles tels qu'ils conduisent à des résidus indépendants d'un certain nombre de

variables abstraites. Résumons cette approche en construisant idéalement une telle procédure orientée MF. Supposons que l'on dispose de d hypomodèles représentant les propriétés comportementales d'un système physique affecté de d défauts possibles. Chaque hypomodèle est dédié à un défaut unique et distinct ; il peut toujours se mettre sous la forme suivante :

$$g_i(\tilde{u}, \tilde{y}, \theta, \xi_i) = \tilde{y} - f_i(\tilde{u}, (\tilde{y}, \theta), \xi_i) = 0 : \forall i \in \{1, \dots, d\}$$

ξ_i est une variable d'occurrence du $i^{\text{ème}}$ défaut
 θ contient les autres variables abstraites non déterministes

Les résidus s'annulent lors d'un comportement normal ; ils seront notés : $\varepsilon_i(\tilde{u}, \tilde{y}, \theta, \xi_i) = g_i(\tilde{u}, \tilde{y}, \theta, \xi_i)$. La robustesse vis-à-vis des incertitudes sera atteinte en régime établi si :

$$\frac{\partial \varepsilon_i(\tilde{u}, \tilde{y}, \theta, \xi_i)}{\partial \theta} = 0 \quad \forall u(t), y(t) \tag{10}$$

La méthode veut que si le $i^{\text{ème}}$ résidu est non nul et si les autres s'annulent, alors il est très probable que le $i^{\text{ème}}$ défaut soit présent. Remarquons tout d'abord que la fiabilité d'un tel diagnostic dépend du nombre de générateurs de résidus : s'il n'en existait qu'un seul (ce qui n'a rien d'aberrant), il serait risqué de diagnostiquer la présence du défaut alors qu'il existe un nombre non fini de défauts non représentés.

L'esprit de l'approche est occulté par la contrainte (10). Il est en effet évident qu'éliminer une variable abstraite ne peut s'opérer sans éliminer aussi la ou les relations comportementales dans lesquelles elle intervient. Par conséquent, ces méthodes restreignent les relations CE pouvant intervenir dans un hypomodèle aux relations ne contenant pas de variables abstraites non déterministes, à l'exception des variables d'occurrence de défaut, ce qui a fortiori en proscrie l'usage par des procédures prenant en compte les incertitudes. Certaines méthodes dégradées se contentent d'atténuer l'influence des variables abstraites sur les résultats des tests de cohérence. Ces solutions dégradées tendent néanmoins vers l'idéal défini précédemment en minimisant le poids des relations comportementales incertaines dans les tests de cohérence. Ces méthodes dégradées conduisent à des résultats débouchant vers des approches a posteriori dont les défauts majeurs sont de n'offrir aucune garantie de justesse.

V. Conclusion

Ce chapitre nous a permis de préciser et parfois d'établir un certain nombre de notions fondamentales ainsi que les principes du diagnostic à base de modèles analytiques des systèmes physiques. La démarche que nous avons adoptée pour formaliser mathématiquement le problème du diagnostic des systèmes physiques se distingue des démarches usuelles du fait qu'elle procède par un raisonnement discursif remontant aux premiers principes de la

physique. Nous avons montré qu'un tel travail permettait d'améliorer l'appréhension du diagnostic en donnant une signification précise aux entités mathématiques. Nous avons introduit une hiérarchisation des différents types de modèles : les variables abstraites apparaissent, au titre d'élément premier de modélisation, au niveau le plus élémentaire. Viennent alors les métamodèles, modèles formels, qui font intervenir des variables physiques tout autant que des variables abstraites. Puis suivent les hypomodèles, ou modèles de comportements vérifiables, qui exploitent la connaissance contenue dans un métamodèle ; ils interviennent directement dans les tests de cohérence d'une procédure de diagnostic. Nous avons alors établi que toute procédure de diagnostic se ramène nécessairement à une série de tests de cohérence entre des observations et des modèles. Puis, nous avons déduit, selon que l'approche soit orientée bon ou mauvais fonctionnement, les interprétations à donner aux résultats de ces tests de cohérence. Enfin, en nous fondant autant sur des notions précises que sur des principes, nous avons pu mettre en évidence un certain nombre de sophismes parmi lesquels certaines interprétations de la *détection et localisation de défauts* (FDI) ainsi que la réduction du diagnostic des systèmes physiques comportant des incertitudes de représentation à une démarche dite de *génération de résidus robustes*. Nous avons aussi montré l'avantage théorique du découpage BMF par rapport au découpage FDI.

Chapitre 2

les variables abstraites

TABLE DES MATIÈRES

I.	Introduction	57
II.	Description probabiliste.....	58
II.1.	Spécificités des variables abstraites stochastiques	61
II.2.	Loi de Gauss-Laplace	63
II.3.	Loi de distribution uniforme	64
III.	Aspects théoriques de la description par les bornes	67
III.1.	Propriétés des domaines abstraits monodimensionnels	68
III.2.	Propriétés des domaines abstraits multidimensionnels.....	71
	Notations et définitions préliminaires.....	72
	Notion d'indépendance et circonscription par un orthotope.....	75
	Détermination analytique d'un domaine abstrait	77
	Cas particulier des champs dissociables	82
	Cas particulier des champs affines	84
IV.	Evaluation numérique des domaines abstraits	88
IV.1.	Evaluation d'un domaine par balayage	88
IV.2.	Evaluation d'un domaine par pavage	90
V.	Conclusion	92

I. Introduction

La plupart des systèmes physiques n'admettent pas de métamodèle certain précis et suffisamment simple pour pouvoir être exploité par une procédure de diagnostic. Concernant la machine à courant continu sur laquelle nous reviendrons plus en détails ultérieurement, seuls des modèles simples peuvent raisonnablement être exploités ; or, en les étudiant, force est de constater que chaque paramètre ne peut être associé à un unique phénomène. Ainsi, les électrotechniciens préfèrent parler de résistance ou d'inductance équivalente car ils ont depuis longtemps constaté que les valeurs identifiées de ces paramètres ne correspondaient pas à une réalité physique. Les paramètres de résistance regroupent autant les dissipations par effet Joule que les pertes fer et les effets de peau ; de la même façon, les paramètres d'inductance intègrent aussi les phénomènes de saturation magnétique. Or, puisque chacun des phénomènes recouverts par un même paramètre dépend de conditions de fonctionnement de la machine (température, spectres et amplitudes des signaux,...), il n'est pas surprenant qu'un modèle à paramètres déterministes ne puisse que mal représenter le comportement de la machine à courant continu. De même, à l'instar des systèmes électrotechniques, de telles constatations pourraient être tirées de la plupart des systèmes physiques faisant intervenir un grand nombre de phénomènes qu'un modèle simple ne peut prendre en compte que globalement.

Pour représenter les systèmes physiques abusivement qualifiés d'incertains, entendu que leurs comportements sont représenté par des modèles comportant des incertitudes, il est nécessaire, conformément aux résultats de la partie IV.2 du premier chapitre, que leur métamodèle intègre des variables abstraites non déterministes. Les variables abstraites les plus utilisées par le passé ont été de nature stochastique et, en particulier, celles pourvues d'une fonction de densité de probabilité gaussienne. Nous verrons que si la distribution de Gauss-Laplace est appropriée à la représentation d'incertitudes de mesure, elle n'en est pas moins mal adaptée à la représentation des incertitudes portant sur le comportement propre d'un système physique. Nous montrerons aussi, dans la deuxième partie de ce chapitre, pourquoi les variables abstraites à support borné, qui ont récemment été l'objet d'importantes publications scientifiques ([Belforte et al,90], [Cerone,93], [Milanese et al,96], [Norton,89], etc...), ont un intérêt considérable pour le diagnostic.

Le chapitre précédent nous a donné l'occasion de préciser ce que nous entendions par le vocable de *variable abstraite*. Souvenons-nous qu'une variable abstraite est à rapporter à un niveau juste inférieur à celui du métamodèle et qu'elle appartient à une relation comportementale élémentaire entre des variables physiques, par essence, potentiellement mesurables. Une variable abstraite est donc une entité première de modélisation intervenant dans la relation comportementale élémentaire dont la nature peut être déterministe ou non.

Les variables abstraites déterministes étant fort connues depuis longtemps sous l'intitulé de *paramètres*, nous nous consacrerons exclusivement aux variables abstraites non déterministes. Avant de revenir synthétiquement sur les propriétés fondamentales des variables non déterministes, il est essentiel de saisir la nuance existant entre la variable non déterministe au sens large et la variable abstraite non déterministe. Une variable physique reconstruite peut en effet être de nature non déterministe sans pour autant mériter le qualificatif d'*abstrait*. Au titre d'élément premier de la représentation comportementale, la variable abstraite est constituée d'un espace abstrait parfaitement identifié : son support, ainsi qu'éventuellement, d'une mesure probabiliste établissant une relation de probabilité entre les éléments de l'espace abstrait et les valeurs possibles de la variable.

L'ensemble des valeurs que peut prendre une variable abstraite coïncide avec son espace abstrait. Nous nous restreindrons à l'étude des espaces abstraits complets¹ et connexes car eux seuls correspondent aux incertitudes que nous souhaitons représenter. Parmi ces variables, nous distinguons celles qui sont définies sur un espace abstrait non probabilisé, engendrant des variables bornées, des variables abstraites définies sur un espace probabilisé et qui engendrent des variables stochastiques. La première partie de ce chapitre est consacrée à la description probabiliste et met en évidence ses limites. La seconde introduit la description par les bornes et explore ses propriétés en vue d'une application au diagnostic. Cette approche est relativement nouvelle, particulièrement en ce qui concerne le diagnostic. Enfin, nous décrirons, dans une dernière partie, quelques méthodes numériques permettant d'évaluer des domaines bornés dans un espace multidimensionnel.

II. Description probabiliste

Cette partie est entièrement consacrée aux variables stochastiques. Après un bref rappel des principes et propriétés des variables abstraites définies sur un espace probabilisé de Gauss-Laplace, nous étudierons les variables à densité de probabilité uniforme. Nous discuterons alors des particularités de leur emploi en tenant compte de ce que la nature des variables abstraites conditionne les hypomodèles qui pourront être élaborés, par exemple, la nature des variables physiques reconstruites dans le cas d'hypomodèles causaux. L'objectif est de mettre en évidence les limites inhérentes à l'usage de telles variables.

Lorsque l'espace abstrait $\Omega \subset \mathbb{R}$ est complet, un événement est défini comme la réunion d'ouverts de l'espace abstrait Ω ; par conséquent, le complémentaire sur Ω d'un événement est aussi un événement. Pour être rigoureux, il est nécessaire de définir l'espace des événements

¹ Toute suite de Cauchy définie sur Ω converge vers un point de l'adhérence de Ω .

associé à une variable stochastique constitué par la réunion de tous les recouvrements² de Ω : cet espace est appelé borélien de Ω [Gray,86] et se note $B(\Omega)$. Naturellement, tout élément de $B(\Omega)$ noté E admet un élément complémentaire \bar{E} sur Ω appartenant aussi à $B(\Omega)$. Ainsi, la propriété nécessaire et suffisante pour que $B(\Omega)$ puisse être considéré comme un espace des événements possibles est satisfaite.

La définition d'une mesure probabiliste permet d'améliorer la description des relations existant entre une variable stochastique et son support en décrivant l'occurrence des événements ou valeurs possibles. Une mesure probabiliste $p(\cdot)$ est alors définie sur l'espace des événements qui associe à chaque élément E de $B(\Omega)$ une valeur $p(E)$ de \mathbb{R} . Cette mesure vérifie les propriétés suivantes :

$$\begin{cases} 0 \leq p(E) \leq 1, p(\emptyset) = 0 \text{ et } p(\Omega) = 1 & (1.a) \\ p(\bar{E}) = 1 - p(E) & (1.b) \\ \forall (E,F) \in B(\Omega)^2, p(E \cup F) = p(E) + p(F) - p(E \cap F) & (1.c) \\ \forall F \in B(\Omega), \forall \{E_i \in B(\Omega)\}_{i \in \mathbb{N}} / \Omega \subset \bigcup_{i \in \mathbb{N}} E_i, p(F) = \sum_{i \in \mathbb{N}} p(F \cap E_i) & (1.d) \end{cases}$$

Le triplet $(\Omega, B(\Omega), p(\cdot))$ définit un espace probabilisé et permet d'introduire la notion de *variable aléatoire* ω à valeur dans Ω . Les événements ou occurrences, auxquels s'applique la mesure probabiliste $p(\cdot)$, qui sont associés à la variable ω , appartiennent à $B(\Omega)$. Plus généralement, une variable stochastique X est définie vers un espace de dimension 1 et elle est une fonction de $d(\in \mathbb{N})$ variables déterministes et de $n(\in \mathbb{N}^*)$ variables aléatoires $(\omega_i : i \in \{1, \dots, n\})$ appartenant respectivement aux espaces probabilisés $(\Omega_i, B(\Omega_i), p_i(\cdot))$:

$$\begin{aligned} X : (\Omega_1 \times \dots \times \Omega_n) \times D &\rightarrow \Omega_X \\ (\omega_1, \dots, \omega_n, \delta) &\mapsto X(\omega_1, \dots, \omega_n, \delta) \end{aligned}$$

ω_i est une valeur de l'espace abstrait Ω_i et d un vecteur déterministe donné de l'espace D

La notion de vecteur stochastique se définit naturellement comme un vecteur de variables stochastiques tandis que la notion de processus stochastique se justifie dès lors que dans les composantes déterministes d'une variable ou d'un vecteur stochastique apparaît le temps. Une variable abstraite stochastique peut donc être une variable stochastique dépendant du temps, c'est-à-dire un processus stochastique dont le support et la mesure probabiliste sont des données connues. Contrairement aux variables abstraites fonctions du temps, une variable abstraite indépendante du temps est, par définition, invariante dans le temps. Ainsi, une incertitude sur un phénomène permanent sera représentée par une variable abstraite ω indépendante du temps, tandis qu'une incertitude sur un phénomène fluctuant au cours du

² Un ensemble Ω de réels admet le recouvrement $R_\Omega = \{G_i : i \in \mathbb{N}, G_i \subset \Omega\}$ si Ω est inclus dans la réunion des ensembles G_i de cette famille.

temps sera représentée par une variable abstraite $\omega(t)$ fonction du temps. Il faut remarquer que, malgré les phénomènes de natures différentes représentés par ces variables, elles peuvent appartenir toutes deux à un même espace probabilisé $(\Omega, \mathcal{B}(\Omega), p(\cdot))$.

Puisque les espaces abstraits appartiennent à \mathbb{R} , il est possible d'introduire la notion de fonction de répartition F :

$$\forall \alpha \in \mathbb{R}, F(\alpha) = p(\omega \in \Omega : \omega \leq \alpha)$$

Lorsque cette fonction de répartition est dérivable, la définition de la fonction densité de probabilité f définie positive permet souvent de simplifier le calcul d'une probabilité $p(E)$ d'un élément E de $\mathcal{B}(\Omega)$:

$$P(E) = \int_E f(\omega) d\omega \text{ avec } \forall \omega \in \Omega, f(\omega) = \frac{dF(\omega)}{d\omega}$$

La fonction de densité de probabilité de deux variables ω_1 et ω_2 s'écrit de manière générale $f(\omega_1, \omega_2)$. En conséquence de (1.c), si les variables ω_1 et ω_2 sont indépendantes, la fonction de densité de probabilité peut s'écrire : $f(\omega_1, \omega_2) = f_{\omega_1}(\omega_1) f_{\omega_2}(\omega_2)$.

Dans le cas d'une variable stochastique quelconque $X(\omega, d)$, où $d \in D$ est donné et ω est défini sur un espace probabilisé $((\Omega, \mathcal{B}(\Omega), p(\cdot)), D)$, il est crucial de déterminer la nature de l'espace probabilisé image $(\Omega_X, \mathcal{B}(\Omega_X), p_X(\cdot))$. Il est facile de déduire Ω_X puisque c'est l'image de Ω par X . De plus, si $E_X \in \mathcal{B}(\Omega_X)$, il est évident que :

$$p_X(E_X) = p(X^{-1}(E_X)) = p(\omega : X(\omega, d) \in E_X)$$

En rapportant cette propriété à la fonction de répartition F_X , il vient :

$$\begin{aligned} F_X(\alpha) &= p_X(\{X(\omega, d) \in \Omega_X : X(\omega, d) \leq \alpha\}) = p(X^{-1}(\{X(\omega, d) \in \Omega_X : X(\omega, d) \leq \alpha\})) \\ &= p(\{\omega \in \Omega : X(\omega, d) \leq \alpha\}) \end{aligned} \quad (2)$$

La fonction de densité de probabilité de la variable stochastique $X(\cdot, d)$ est obtenue par dérivation de la fonction de répartition F_X .

Un corollaire de la propriété (2), par exemple [Peebles,87], est que la fonction de répartition d'une variable stochastique $x = g(\omega_1, \dots, \omega_p)$, où les variables $\omega_i : i \in \{1, \dots, p\}$ sont de nature stochastique et caractérisées par une densité de probabilité $f(\omega_1, \dots, \omega_p)$, est donnée par :

$$F_x(u) = \int_{D_u = \{(\omega_1, \dots, \omega_p) / g(\omega_1, \dots, \omega_p) \leq u\}} f(\omega_1, \dots, \omega_p) d\omega_1 \cdots d\omega_p \quad (3)$$

L'application de (3) au calcul de la fonction de densité de probabilité d'une somme de deux variables indépendantes ω_1 et ω_2 permet d'établir que la fonction de densité de probabilité

résultante est égale au produit de convolution des distributions de ω_1 et de ω_2 :

$$f_{\omega_1+\omega_2}(u) = \int_{\mathbb{R}} f_{\omega_1}(v)f_{\omega_2}(u-v)dv = \int_{\mathbb{R}} f_{\omega_1}(u-v)f_{\omega_2}(v)dv \quad (4)$$

D'autre part, la théorie probabiliste classique définit le moment $M_q(x)$ d'ordre $q \in \mathbb{N}$ d'une variable stochastique $x=g(\omega_1, \dots, \omega_n)$ de densité $f_x(\omega_1, \dots, \omega_n)$. Ce moment est plus généralement exprimé en terme d'espérance : $M_q(x)=E(x^q)$ et a pour définition :

$$M_q(x) = E(x^q) = \int_{\mathbb{R}^n} g(\omega_1, \dots, \omega_n)^q f_x(\omega_1, \dots, \omega_n) d\omega_1 \dots d\omega_n ; \forall q \in \mathbb{N} \quad (5)$$

La variance d'une variable stochastique x correspond au moment d'ordre 2 de la variable x centrée autour de son moment d'ordre 1 : $VAR(x)=M_2(x)-M_1(x)^2=E((x-E(x))^2)=M_2(x)-M_1(x)^2$.

Les moments statistiques ont des propriétés intéressantes. Le moment d'ordre 1 est une fonction linéaire ; ainsi, la propriété suivante reste vraie quelle que soit la densité de probabilité $f(\omega_1, \omega_2)$: $M_1(a\omega_1+b\omega_2)=aM_1(\omega_1)+bM_1(\omega_2)$. D'un autre point de vue, il est possible de déduire le moment d'ordre 1 d'une somme pondérée de variable stochastique tout en ignorant la densité de probabilité de la somme. Il est aussi facile de montrer que le moment d'ordre 1 d'un produit de variables stochastiques indépendantes est égal au produit des moments d'ordre 1 de chacune des variables : $M_1(\omega_1\omega_2)=M_1(\omega_1)M_1(\omega_2)$. De plus, si les variables ω_1 et ω_2 sont indépendantes, d'autres propriétés intéressantes concernant les moments d'ordre 2 apparaissent : $VAR(a\omega_1+b\omega_2)=a^2VAR(\omega_1)+b^2VAR(\omega_2)$ et $M_2(\omega_1\omega_2)=M_2(\omega_1)M_2(\omega_2)$.

Ainsi, grâce à ces propriétés, il est souvent possible de déduire les moments d'ordre 1 et 2 de variables physiques reconstruites sans pour autant que les fonctions de densité de probabilité ou de répartition de ces variables n'aient à être déterminées.

II.1. Spécificités des variables abstraites stochastiques

Il convient alors de s'interroger sur la pertinence des moments statistiques d'ordre 1 et 2, très largement utilisés, quant à la connaissance des variables stochastiques déduites. Nous proposons la réflexion suivante : une mesure probabiliste est généralement parfaitement caractérisée par les deux premiers moments statistiques qui lui sont associés. Par conséquent, si les variables abstraites participant à la définition de variables stochastiques sont pourvues de mesures probabilistes de même nature, et si les variables stochastiques déduites ont, elles aussi, des distributions de cette même nature, alors, selon toute évidence, la seule connaissance des deux premiers moments statistiques suffit à caractériser parfaitement toutes les variables stochastiques déduites. Plus savamment, tant qu'une loi de distribution est stable pour un opérateur ou une fonction, les premiers moments statistiques suffisent à caractériser

sa propre loi de distribution. Dans le cas contraire, le laborieux, voire impossible, calcul des distributions des variables déduites ne pourra plus être éludé.

Commençons par étudier la stabilité d'une distribution pour une fonction affine : $x(\omega, d) = \varphi(d)\omega + \psi(d)$, où ω est défini sur un espace probabilisé $(\Omega, \mathcal{B}(\Omega), p(\cdot))$, par une fonction de répartition $F(\cdot)$ ou par $f(\cdot)$ la densité de probabilité associée. Le calcul de la fonction de répartition F_x , découle de (2) :

$$F_x(\alpha) = p\left(\left\{\omega \in \Omega: \varphi(d)\omega + \psi(d) \leq \alpha\right\}\right) = p\left(\left\{\omega \in \Omega: \omega \leq \frac{\alpha - \psi(d)}{\varphi(d)}\right\}\right) = F\left(\frac{\alpha - \psi(d)}{\varphi(d)}\right)$$

et, en dérivant :

$$\frac{d}{d\alpha}(F_x(\alpha)) = \frac{d}{d\alpha}\left(F\left(\frac{\alpha - \psi(d)}{\varphi(d)}\right)\right) = \frac{1}{\varphi(d)}f\left(\frac{\alpha - \psi(d)}{\varphi(d)}\right). \quad (6)$$

Nous déduisons donc que toute loi de distribution est stable pour les fonctions affines et, en conséquence, que les deux premiers moments statistiques suffisent à les caractériser quelle que soit la nature des fonctions de répartition.

Les autres propriétés ne peuvent être étudiées indépendamment des lois de distribution. Nous en avons retenu deux : la loi de Gauss-Laplace et la loi de distribution uniforme sur un intervalle de \mathbb{R} . La loi de Gauss-Laplace est la loi la plus utilisée pour la représentation des incertitudes de type *bruit de mesure* et parfois même, de type *bruit de structure* dans le filtrage de Kalman, tandis que la loi de distribution uniforme nous semble être la loi la mieux adaptée aux objectifs que nous poursuivons. En effet, considérons l'exemple des paramètres résistances intervenant dans le métamodèle de type abc ou odq d'une machine asynchrone. Parmi les phénomènes recouverts par ces paramètres, la variation de la température interne de la machine engendre une modification de la résistivité des conducteurs électriques. Or, la variation de température en fonctionnement normal étant bornée, il est naturel de considérer que les paramètres de résistance appartiennent à des intervalles finis de \mathbb{R} , ce que, contrairement à la distribution de Gauss, la distribution uniforme permet de représenter ; du moins, dans l'hypothèse d'équiprobabilité des valeurs. De plus, il n'y a, a priori, aucune raison pour que les occurrences de deux valeurs distinctes soient différentes puisque les phénomènes perturbants en bon fonctionnement sont : la température ambiante, les courants circulant dans la machine, ... qui sont difficilement prévisibles. Nous avons exclu de l'étude les cas où des variables stochastiques aux distributions de formes différentes intervenaient dans une même fonction car, cela tombe sous le sens et ne demande guère plus de démonstration, la distribution obtenue a une forme différente de celle des variables qui la composent.

II.2. Loi de Gauss-Laplace

La fonction de densité de probabilité de la loi de Gauss-Laplace d'une variable x centrée autour de \bar{x} et d'écart-type σ_x s'exprime ainsi :

$$x \in \mathcal{N}(\bar{x}, \sigma_x) \Rightarrow f(x) = \frac{e^{-\frac{1}{2}\left(\frac{x-\bar{x}}{\sigma_x}\right)^2}}{\sqrt{2\pi}\sigma_x}$$

Or, les paramètres caractérisant cette fonction, nommément \bar{x} et σ_x , sont directement liés aux deux premiers moments de la variable x .

A partir de la relation (4), il est facile de montrer que la loi de Gauss est stable pour l'addition. En effet, en posant $z=x+y$ où x et y sont des variables stochastiques à distribution gaussienne et de caractéristiques respectives \bar{x} , σ_x et \bar{y} , σ_y , nous déduisons que :

$$f_{x+y}(z) = \frac{1}{2\pi\sigma_x\sigma_y} \int_{\mathbb{R}} e^{-\frac{1}{2}\left(\left(\frac{x-\bar{x}}{\sigma_x}\right)^2 + \left(\frac{z-x-\bar{y}}{\sigma_y}\right)^2\right)} dx = \frac{e^{-\frac{1}{2}\frac{(z-\bar{x}-\bar{y})^2}{\sigma_x^2 + \sigma_y^2}}}{\sqrt{2\pi(\sigma_x^2 + \sigma_y^2)}}.$$

Ainsi, puisque le modèle sur lequel se fonde le filtre de Kalman, s'écrivant généralement :

$$\begin{cases} x_{k+1} = Ax_k + Bu_k + Fw_k & : w_k \in \mathcal{N}(\bar{w}_k, \sigma_{w_k}) \\ y_k = Cx_k + Du_k + Ev_k & : v_k \in \mathcal{N}(\bar{v}_k, \sigma_{v_k}) \\ x_0 \in \mathcal{N}(\bar{x}_0, \sigma_{x_0}) \end{cases}$$

ne contient que des sommes de variables stochastiques, les variables stochastiques déduites x_k et y_k ont nécessairement une distribution gaussienne. La connaissance des deux premiers moments de ces variables caractérise complètement leur distribution.

Le produit des deux variables stochastiques x et y conduit à la densité de probabilité, déduite de (3), suivante :

$$f_{xy}(u) = \frac{1}{2\pi\sigma_x\sigma_y} \int_{\mathbb{R}} e^{-\frac{1}{2}\left(\left(\frac{x-\bar{x}}{\sigma_x}\right)^2 + \left(\frac{u-\bar{y}x}{\sigma_y}\right)^2\right)} dx$$

L'intégration analytique d'une telle fonction paraît impossible ; nous avons cependant calculé la fonction de densité de probabilité d'un produit de deux variables gaussiennes indépendantes et tracé les résultats sur la figure I. Conformément à nos précédentes assertions, les deux premiers moments peuvent être calculés sans grande difficulté (cf. figure I). Pour comparaison, nous avons représenté sur la même figure une fonction de densité de probabilité de Gauss-Laplace ayant les deux premiers moments identiques à ceux de la distribution

calculée. La différence entre les deux distributions est nette et montre bien que la connaissance des deux premiers moments ne suffit pas toujours à caractériser complètement une distribution.

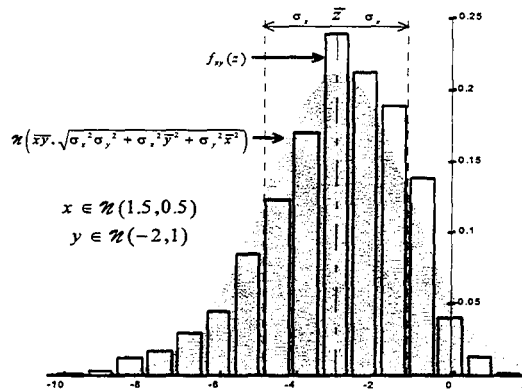


Figure I- Fonction de densité de probabilité d'un produit de variables gaussiennes

Les procédures de diagnostic consistant à tester si une variable observée \tilde{z} est contenue dans un intervalle $[\bar{z} - \lambda\sigma_z, \bar{z} + \lambda\sigma_z]$, où λ est un nombre réel positif et σ_z est l'écart-type de la variable z , s'appliquent donc difficilement à des fonctions ne faisant pas intervenir exclusivement des sommes de variables stochastiques ; elles peuvent certes être utilisées comme approximation mais sans pouvoir être justifiées par un développement théorique. Il en découle que, pratiquement, de nombreux cas ne peuvent être correctement traités, notamment ceux où des incertitudes relatives de mesure apparaissent.

II.3. Loi de distribution uniforme

Lorsque la probabilité d'occurrence des valeurs d'un espace abstrait, associé à une variable stochastique x , est uniformément répartie sur un intervalle $I = [\tilde{x}, \hat{x}]$ de \mathbb{R} , la fonction de répartition de la variable x s'écrit :

$$f(x) = \frac{\gamma(x - \tilde{x}) - \gamma(x - \hat{x})}{\hat{x} - \tilde{x}} \Leftrightarrow x \in \mathcal{U}(\tilde{x}, \hat{x}).$$

$\gamma(\cdot)$ est la fonction de Heaviside : $\gamma(u) = 1$ si $u \geq 0$, $\gamma(u) = 0$ sinon.

Les deux premiers moments statistiques de la variable x se déduisent sans difficulté : $\bar{x} = \frac{1}{2}(\tilde{x} + \hat{x})$ et $VAR(x) = \sigma_x^2 = \frac{1}{12}(\hat{x} - \tilde{x})^2$. Ainsi, comme dans le cas d'une distribution gaussienne, les paramètres \bar{x} et σ_x caractérisent complètement la loi de distribution car l'application linéaire ι définie par :

$$\iota: \mathbb{R}^2 \rightarrow \mathbb{R}^2$$

$$\begin{bmatrix} \tilde{x} \\ \hat{x} \end{bmatrix} \mapsto \begin{bmatrix} \bar{x} \\ \sigma_x \end{bmatrix} = \begin{bmatrix} \frac{1}{2} & \frac{1}{2} \\ -\frac{1}{2\sqrt{3}} & \frac{1}{2\sqrt{3}} \end{bmatrix} \begin{bmatrix} \tilde{x} \\ \hat{x} \end{bmatrix}$$

est bijective. Nous pourrions être tentés de croire que pour des fonctions simples ne contenant

au plus que des sommes de variables stochastiques, la seule connaissance des deux premiers moments suffit à caractériser parfaitement la loi de densité de probabilité. Ce serait néanmoins présumer hâtivement de la stabilité de cette distribution pour l'addition.

La relation (4) permet de calculer la fonction de densité de probabilité de la somme de deux variables x et y distribuées uniformément sur leurs intervalles respectifs $[\bar{x}, \hat{x}]$ et $[\bar{y}, \hat{y}]$:

$$f_{x+y}(z) = \frac{\int_{\mathbb{R}} h(x,z) dx}{(\hat{x} - \bar{x})(\hat{y} - \bar{y})} \text{ avec } \begin{cases} h(x,z) = h_1(x)h_2(x,z) \\ h_1(x) = \gamma(x - \bar{x}) - \gamma(x - \hat{x}) \\ h_2(x,z) = \gamma(z - x - \bar{y}) - \gamma(z - x - \hat{y}) \end{cases}$$

Pour évaluer l'intégrale, il faut distinguer, outre les cas où $h(x,z)$ est nul, les 4 situations suivantes :

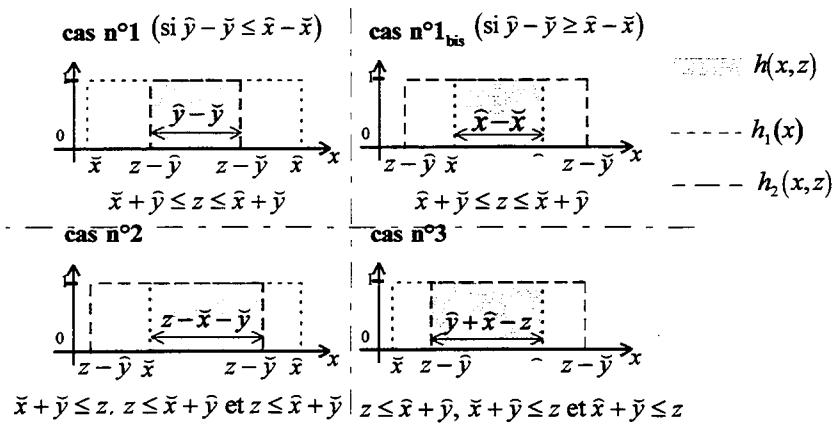


Figure II- Quatre situations possibles de la fonction $h(x,z)$

Rassemblons les différentes situations en posant : $\xi_1 = \bar{x} + \bar{y}$, $\xi_2 = \min(\bar{x} + \hat{y}, \hat{x} + \bar{y})$, $\xi_3 = \max(\bar{x} + \hat{y}, \hat{x} + \bar{y})$ et $\xi_4 = \hat{x} + \hat{y}$. Dès lors, en raisonnant sur la figure II, nous pouvons déduire l'équation de la fonction de densité de probabilité de la somme de x et de y :

$$\begin{cases} \text{si } z \leq \xi_1 \text{ ou si } z \geq \xi_4, f_{x+y}(z) = 0 \\ \text{si } \xi_1 \leq z \leq \xi_2, f_{x+y}(z) = \frac{z - \bar{x} - \bar{y}}{(\hat{x} - \bar{x})(\hat{y} - \bar{y})} \\ \text{si } \xi_2 \leq z \leq \xi_3, f_{x+y}(z) = \frac{1}{\max(\bar{x} - \bar{x}, \hat{y} - \bar{y})} \\ \text{si } \xi_3 \leq z \leq \xi_4, f_{x+y}(z) = \frac{\hat{x} + \hat{y} - z}{(\hat{x} - \bar{x})(\hat{y} - \bar{y})} \end{cases}$$

Ces résultats peuvent être synthétisés sur la figure suivante :

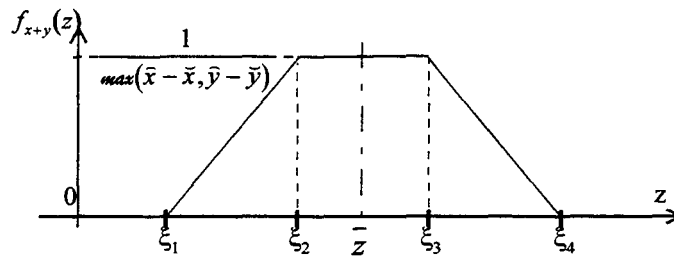


Figure III- fonction de densité de probabilité d'une somme de variables à distribution uniforme

La somme de deux variables stochastiques à distribution uniforme conduit à une distribution trapézoïdale. Le premier moment de la variable $z=x+y$ donne le centre de la distribution : $\bar{z} = \frac{1}{2}(\bar{x} + \tilde{x} + \bar{y} + \tilde{y})$. Par contre, la racine du second moment $\sigma_z = \frac{1}{2\sqrt{3}}\sqrt{(\bar{x} - \tilde{x})^2 + (\bar{y} - \tilde{y})^2}$ ne peut être reliée à la distribution obtenue par un coefficient de proportionnalité et, d'ailleurs, une distribution trapézoïdale ne peut pas être parfaitement caractérisée par moins de 3 paramètres.

La densité de probabilité du produit des deux variables x et y , conduit à l'intégrale suivante :

$$f_{xy}(z) = \frac{1}{(\bar{x} - \tilde{x})(\bar{y} - \tilde{y})} \int_{\mathbb{R}} (\gamma(x - \tilde{x}) - \gamma(x - \bar{x})) \left(\gamma\left(\frac{z}{x} - \tilde{y}\right) - \gamma\left(\frac{z}{x} - \bar{y}\right) \right) dx$$

La procédure d'intégration est identique à celle que nous venons d'employer à la différence près que la solution dépend des signes de $\tilde{x}, \bar{x}, \tilde{y}$ et \bar{y} . Ceci présente l'inconvénient d'introduire dans la solution un grand nombre de conditions. De ce fait, nous n'avons calculé la fonction de densité de probabilité que dans le cas particulier où x et y étaient distribués uniformément sur les intervalles respectifs [1 2] et [-3 -1]. Nous avons obtenu la solution suivante :

$$\begin{cases} \text{si } z \leq -6 \text{ ou si } z \geq -1, & f_{xy}(z) = 0 \\ \text{si } -6 \leq z \leq -3, & f_{xy}(z) = \frac{z+6}{6} \\ \text{si } -3 \leq z \leq -2, & f_{xy}(z) = \frac{1}{2} \\ \text{si } -2 \leq z \leq -1, & f_{xy}(z) = \frac{-z-1}{2} \end{cases}$$

La fonction de densité de probabilité obtenue a donc l'allure suivante :

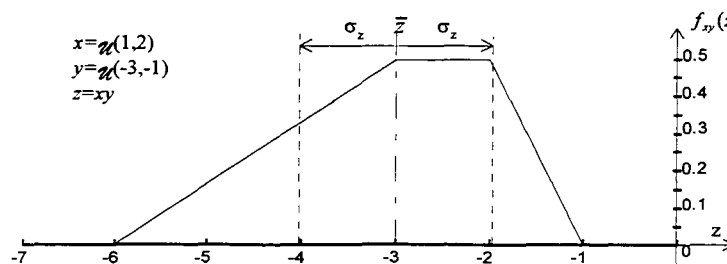


Figure IV- Fonction de densité de probabilité d'un produit de variables uniformément distribuées

Dans cet exemple aussi, nous constatons que les moments statistiques n'expriment qu'une information globale et qu'ils caractérisent mal les distributions obtenues. Est-ce à conclure que, pour élaborer des modèles tenant compte des incertitudes, il est nécessaire de reconstruire les fonctions de densité de probabilité ? Nous avons vu que dans le cas de distributions gaussiennes, dès que les fonctions font intervenir des produits de variables stochastiques, il n'est plus possible de déterminer, autrement que par une simulation numérique, la distribution. Quant aux distributions uniformes, elles peuvent certes toujours être calculées mais elles font intervenir un grand nombre de conditions pour conduire à une solution. C'est en mettant en œuvre de tels calculs à des modèles récurrents impliquant un calcul à chaque itération que l'on se rendra compte combien de telles approches seraient irréalistes, et quant bien même elles seraient implémentées, comment donc une procédure de diagnostic pourrait-elle tester si une seule observation est conforme à une loi de distribution. Seuls les états stationnaires permettraient un diagnostic et encore, combien d'observations seraient-elles nécessaires avant de parvenir à un diagnostic ?

Si les distributions des variables stochastiques sont gaussiennes, il est possible, tant qu'il n'est question que de sommes, de tester si une observation est comprise dans une bande proportionnelle à l'écart-type centrée autour de la valeur moyenne. En revanche, dans le cas de distributions uniformes, il reste à trouver des informations intéressantes pour réaliser des tests. Intuitivement, pour ce qui est du diagnostic, les informations les plus intéressantes sont les bornes des distributions obtenues. Par exemple, sur la figure IV, la connaissance des bornes $[-6,-1]$ de la distribution obtenue permet d'imaginer des procédures de diagnostic simple ; le test revient à vérifier si une observation appartient à cet intervalle. Raisonner de la sorte revient à ignorer la distribution des variables et à n'en considérer que les bornes : nous parlerons de variables bornées. Elles font l'objet de la suite de ce chapitre.

III. Aspects théoriques de la description par les bornes

Cette partie s'inspire autant des résultats concernant l'usage des intervalles dans l'évaluation d'expressions mathématiques, produits par la communauté mathématique que des récents résultats de la communauté automatique sur les approches bornantes. Les premiers résultats apparurent dans la littérature sous l'intitulé d'*arithmétique des intervalles* et datent de la fin des années 60 et en grande partie des années 70 avec [Hansen,69] et notamment les travaux de R.E. Moore [Moore,66] et [Moore,79]. Bien que ces travaux aient été développés pour prendre en compte les arrondis de calcul inhérents à l'informatique, Chang [Chang,95] a cherché à appliquer ces résultats au diagnostic. Cependant, nous montrerons que les résultats de Moore doivent être complétés pour pouvoir s'étendre à la représentation d'incertitudes. Plus récemment, d'intéressants travaux concernant l'estimation paramétrique dans le cas d'incertitudes bornées sont apparus. Le problème d'estimation fait généralement intervenir un

modèle, linéaire ou non, comportant, en plus de paramètres invariants, des variables abstraites bornées. Faut de description probabiliste, il n'est pas question d'estimer leurs valeurs les plus probables, mais de déterminer *l'ensemble³ de toutes les valeurs des paramètres, cohérentes avec le modèle du système et les bornes des erreurs* [Vicino et al,91]. En 1982, Fogel et Huang [Fogel et al,82] proposèrent un algorithme permettant, dans le cas linéaire, de circonscrire le domaine FPS par une ellipsoïde et, durant cette même année, Milanese et Belforte [Milanese et al,82] exposèrent des méthodes se fondant sur l'algèbre linéaire pour circonscrire le domaine FPS par une *boîte*. En 1987, Norton [Norton,87] publie un article de synthèse où il montre que, dans les cas linéaires, les domaines FPS sont des polytopes et il décrit un algorithme permettant de les construire. Il propose une méthode de réjection des valeurs aberrantes et circonscrit le domaine obtenu par le plus petit ellipsoïde puis, en 1989, il affine [Norton,89] l'approximation du domaine FPS en proposant un algorithme déterminant le plus grand ellipsoïde inscrit dans ce domaine. En 1990, Belforte, Bona et Cerone [Belforte et al,90] comparent les méthodes de programmation linéaire conduisant à des boîtes circonscrites et les algorithmes conduisant à des ellipses. Ils concluent que le calcul d'ellipses circonscrites est plus rapide mais moins précis. Vicino et Milanese [Vicino et al,91] renchérisent en 1991 en montrant comment les domaines FPS peuvent être circonscrits par de multiples domaines, notamment des formes en *diamants* et, dans [Milanese et al,91a], les concepts généraux des approches bornantes sont décrits. Ces développements ont conduit à l'ouvrage collectif intitulé *Bounding approaches to system identification* [Milanese et al,96] qui réunit la grande majorité des travaux concernant l'approche bornante en identification. Citons pour finir, [Milanese et al,91b] pour les travaux s'appliquant aux modèles non linéaires ; nous reviendrons plus tard sur certains travaux marginaux... Bien que ces résultats soient une extension importante des travaux sur les intervalles puisqu'ils s'appliquent à des domaines multidimensionnels, ils ne concernent qu'indirectement la représentation de l'incertain pour le diagnostic. Cette partie ne développe donc pas ces différents travaux, mais se contente d'explorer les différentes propriétés utiles à la représentation des incertitudes par des variables bornées. Nous proposons un formalisme nouveau en certains points pour réunir des propriétés dont certaines ont été établies par les protagonistes cités en sus et parvenir à des fondements théoriques homogènes. Nous veillerons cependant, autant que possible, à justifier l'intérêt des notations et concepts que nous introduisons.

III.1. Propriétés des domaines abstraits monodimensionnels

Si aux variables bornées ne sont pas associées de mesures probabilistes, elles n'en restent pas moins munies d'espaces abstraits. En représentant un vecteur de variables bornées par x , nous désignerons l'espace abstrait qui lui est associé par $\mathcal{A}(x)$. Les variables abstraites sont donc des

³ FPS en anglais : Feasible Parameter Set

variables dont l'espace abstrait est donné ; de plus, conformément aux restrictions que nous avons posées à propos des espaces abstraits des variables stochastiques, cet espace est supposé complet et connexe. Nous en déduisons directement que l'espace abstrait d'une variable abstraite est nécessairement un intervalle de \mathbb{R} .

De nombreux auteurs, notamment Neumaier [Neumaier,90] et R.E. Moore, ont formulé une arithmétique des intervalles. Nous proposons une autre formulation de ces propriétés car leur présentation usuelle nous semble contestable. En effet, ces auteurs substituent à l'entité variable x à valeur dans \mathbb{R} l'entité variable-intervalle $X = [\tilde{X}, \hat{X}]$ ($\hat{X} \geq \tilde{X}$) constituée des bornes inférieure et supérieure de la variable x . En conséquence, les propriétés dont ils font état ne peuvent concerner que les bornes. Or, ceci pose un problème insoluble : la dépendance de deux variables n'apparaît pas au niveau de leurs bornes. D'autre part, dans un cas multivariable, les espaces abstraits peuvent avoir des formes complexes dans un espace multidimensionnel qui ne peuvent être représentées par des intervalles ; ceci exclut les lois de composition interne. En se conformant à la formulation propre à l'arithmétique des intervalles, si une variable x appartient à un espace abstrait $X = [-\delta, \delta]$ ($\delta \in \mathbb{R}^+$), la variable $y = -x$ appartient à l'espace $Y = [-\delta, \delta]$ et l'on déduit que $x+y \in [-2\delta, 2\delta] = (X+Y)$ alors qu'il est évident que $x+y=0$. Il est certes exact que la solution $\{0\}$ appartient à $[-2\delta, 2\delta]$ et qu'il est démontré [Moore,79] que le domaine déduit inclut nécessairement l'espace abstrait de $x+y$. Certes, cette formulation permet de déterminer des majorations d'espaces abstraits, mais rarement les espaces abstraits eux-mêmes. De plus, définir une arithmétique sur ces entités *variables-intervalles* pose de nombreux problèmes théoriques rédhibitoires. Pour n'en citer qu'un, on se souviendra que la négation est définie par rapport à l'addition et son élément neutre. L'élément neutre pour l'addition serait $E = [0, 0]$ car c'est l'unique élément vérifiant : $X+E=X$. Pour trouver l'élément symétrique $Y = [\tilde{Y}, \hat{Y}]$ de la *variable-intervalle* $X = [\tilde{X}, \hat{X}]$ et définir la soustraction, posons : $X+Y=E$. Il découle que, pour satisfaire cette relation, il faudrait que $\tilde{Y} = -\tilde{X}$ et que $\hat{Y} = -\hat{X}$; ainsi, la borne inférieure devrait être supérieure à la borne supérieure : $\tilde{Y} \geq \hat{Y}$, ce qui va à l'encontre de la définition de l'*entité intervalle*. La soustraction ne peut donc être définie sur ces entités, excluant ainsi la notion d'algèbre au sens strict. Néanmoins, les propriétés de l'arithmétique des intervalles n'en demeurent pas moins intéressantes lorsqu'elles s'appliquent à des variables bornées réelles indépendantes.

Si $f(\cdot)$ est une fonction d'un vecteur x de variables abstraites, l'espace abstrait de f est donné par :

$$\mathcal{A}(f(x)) = \left[\inf_{x \in \mathcal{A}(x)} (f(x)), \sup_{x \in \mathcal{A}(x)} (f(x)) \right] \quad (7)$$

Parfois, les bornes de l'espace abstrait (7) peuvent être explicitées, c'est l'objet de l'arithmétique des intervalles. Soient deux variables bornées θ_1 et θ_2 indépendantes dont les

espaces abstraits sont respectivement $\mathcal{A}(\theta_1) = [\check{\theta}_1, \hat{\theta}_1]$ et $\mathcal{A}(\theta_2) = [\check{\theta}_2, \hat{\theta}_2]$, deux variables bornées u_1 et u_2 , centrées en 0 et indépendantes avec $\mathcal{A}(u_1) = [-\delta_1, \delta_1]$ et $\mathcal{A}(u_2) = [-\delta_2, \delta_2]$ ainsi qu'un nombre réel λ et σ_λ vérifiant $\sigma_\lambda = 1$ si $\lambda > 0$, $\sigma_\lambda = -1$ si $\lambda < 0$ et $\sigma_\lambda = 0$ si $\lambda = 0$, alors les propriétés suivantes sont vérifiées :

$$\mathcal{A}(\lambda + \theta_1) = [\lambda + \check{\theta}_1, \lambda + \hat{\theta}_1] \quad (8)$$

addition d'une grandeur déterministe à une variable bornée

$$\mathcal{A}(\lambda \theta_1) = \left[\lambda \left(\frac{1 + \sigma_\lambda}{2} \check{\theta}_1 + \frac{1 - \sigma_\lambda}{2} \hat{\theta}_1 \right), \lambda \left(\frac{1 - \sigma_\lambda}{2} \check{\theta}_1 + \frac{1 + \sigma_\lambda}{2} \hat{\theta}_1 \right) \right] \quad (9)$$

en particulier, $\mathcal{A}(\lambda u_1) = [-|\lambda| \delta_1, |\lambda| \delta_1]$

multiplication d'une grandeur déterministe par une variable bornée

$$\mathcal{A}(\theta_1 + \theta_2) = [\check{\theta}_1 + \check{\theta}_2, \hat{\theta}_1 + \hat{\theta}_2] \quad (10)$$

addition de 2 variables bornées

$$\mathcal{A}(\theta_1 \cdot \theta_2) = \left[\min(\check{\theta}_1 \check{\theta}_2, \check{\theta}_1 \hat{\theta}_2, \hat{\theta}_1 \check{\theta}_2, \hat{\theta}_1 \hat{\theta}_2), \max(\check{\theta}_1 \check{\theta}_2, \check{\theta}_1 \hat{\theta}_2, \hat{\theta}_1 \check{\theta}_2, \hat{\theta}_1 \hat{\theta}_2) \right]$$

en particulier, $\mathcal{A}(\theta_1 \cdot u_1) = [\min(\check{\theta}_1, -\hat{\theta}_1) \delta_1, \max(-\check{\theta}_1, \hat{\theta}_1) \delta_1]$ (11)

ou encore, $\mathcal{A}(u_1 \cdot u_2) = [-\delta_1 \delta_2, \delta_1 \delta_2]$

multiplication de 2 variables bornées

si $\mathcal{A}(\theta_2) \cap \{0\} = \emptyset$,

$$\mathcal{A}(\theta_1 / \theta_2) = \left[\min(\check{\theta}_1 / \check{\theta}_2, \check{\theta}_1 / \hat{\theta}_2, \hat{\theta}_1 / \check{\theta}_2, \hat{\theta}_1 / \hat{\theta}_2), \max(\check{\theta}_1 / \check{\theta}_2, \check{\theta}_1 / \hat{\theta}_2, \hat{\theta}_1 / \check{\theta}_2, \hat{\theta}_1 / \hat{\theta}_2) \right] \quad (12)$$

en particulier, $\mathcal{A}(u_1 / \theta_2) = [\min(1 / \check{\theta}_2, -1 / \hat{\theta}_2) \delta_1, \max(-1 / \check{\theta}_2, 1 / \hat{\theta}_2) \delta_1]$

division de 2 variables bornées

Insistons sur le fait que les propriétés dont nous venons de faire état, ne sont valides que dans le cas où les variables qu'elles font intervenir sont indépendantes. Par exemple, si $\theta_2 = -\theta_1$, la propriété (10) ne peut s'appliquer. Nous aurions retrouvé le résultat erroné suivant $\mathcal{A}(\theta_1 + \theta_2) = [-2\check{\theta}_1, 2\hat{\theta}_1]$, alors qu'en effectuant $\theta_1 + \theta_2$, nous obtenons le résultat juste $\mathcal{A}(\theta_1 + \theta_2) = 0$.

Les fonctions monotones sont à l'origine d'une propriété très utile : soit $f(\theta_1, \theta_2)$, une fonction dérivable par rapport à θ_1 , alors s'il existe $\sigma \in \{-1, 1\}$ tel que $\forall \theta_1 \in \mathcal{A}(\theta_1)$ et $\theta_2 \in \mathcal{A}(\theta_2)$,

$$\sigma \frac{\delta f(\theta_1, \theta_2)}{\delta \theta_1} \Bigg|_{\substack{\theta_1 \in \mathcal{A}(\theta_1) \\ \theta_2 \in \mathcal{A}(\theta_2)}} \geq 0 \quad (13a)$$

il s'ensuit que

$$\text{quand } \sigma = 1, \mathcal{A}(f(\theta_1, \theta_2)) = \left[\inf_{\theta_2 \in \mathcal{A}(\theta_2)} (f(\tilde{\theta}_1, \theta_2)), \sup_{\theta_2 \in \mathcal{A}(\theta_2)} (f(\tilde{\theta}_1, \theta_2)) \right] \quad (13b)$$

$$\text{quand } \sigma = -1, \mathcal{A}(f(\theta_1, \theta_2)) = \left[\inf_{\theta_2 \in \mathcal{A}(\theta_2)} (f(\hat{\theta}_1, \theta_2)), \sup_{\theta_2 \in \mathcal{A}(\theta_2)} (f(\hat{\theta}_1, \theta_2)) \right] \quad (13c)$$

propriété des fonctions monotones

R. E. Moore [Moore,79] a montré, tout au moins dans le cas des fonctions arithmétiques⁴ mais on l'étend aisément à toute fonction continue, qu'en substituant à des variables bornées, des variables abstraites indépendantes définies sur les mêmes espaces abstraits, alors le domaine abstrait de la nouvelle fonction inclut nécessairement l'espace abstrait de la fonction originale. Soit une fonction quelconque $f(\theta): \theta = [\theta_1, \dots, \theta_n]^T$ mais néanmoins continue⁵ sur l'espace abstrait connexe $\mathcal{A}(\theta)$. Introduisons des vecteurs de variables abstraites indépendantes $x^{(i)}: i \in \{1, \dots, n\}$ de dimension respective n_i et dont chaque composante $x_j^{(i)} (j \in \{1, \dots, n_i\})$ vérifie : $\mathcal{A}(x_j^{(i)}) = \mathcal{A}(\theta_j): j \in \{1, \dots, n_j\}$. Dans ce cas, en définissant ce que Moore appelle l'extension intervalle $F(x^{(1)}, \dots, x^{(n)})$ vérifiant :

$$\forall \theta \in \mathcal{A}(\theta), \forall i \in \{1, \dots, n\}, \text{ si } x^{(i)} = \begin{bmatrix} \theta_i & \cdots & \theta_i \\ \leftarrow n_i \rightarrow \end{bmatrix}^T \text{ alors } F(x^{(1)}, \dots, x^{(n)}) = f(\theta) \quad (14a)$$

alors, une propriété importante pour la détermination de majorants et de minorants de fonction de variables abstraites se déduit directement :

$$\mathcal{A}(f(\theta)) \subseteq \mathcal{A}(F(x^{(1)}, \dots, x^{(n)})) \quad (14b)$$

L'inclusion s'explique par le fait que chaque composante $x_j^{(i)}$ des vecteurs $x^{(i)}$ n'est définie que par $\mathcal{A}(x_j^{(i)}) = \mathcal{A}(\theta_j)$ et, par conséquent, le vecteur $x^{(i)}$ peut prendre des valeurs différentes de $x^{(i)} = [\theta_i, \dots, \theta_i]^T$ pour lesquelles la condition $F(x^{(1)}, \dots, x^{(n)}) = f(\theta)$ n'a aucune raison d'être satisfaite.

III.2. Propriétés des domaines abstraits multidimensionnels

Nous n'avons jusqu'ici débattu que de variables bornées qui conduisent à des domaines abstraits sous forme d'intervalles. Or, les domaines abstraits de vecteurs de variables bornées conduisent à des domaines multidimensionnels bien plus complexes que des intervalles. L'étude des domaines multidimensionnels est complexe et cependant nécessaire au diagnostic puisqu'un ensemble de comportements possibles est, en général, représenté par un domaine multidimensionnel. Dans un premier temps, nous proposerons une approche analytique conduisant, pour certains champs $f(\theta)$, à la détermination des domaines abstraits qu'ils

⁴ Ne faisant intervenir que les opérateurs arithmétiques '+', '-', 'x' et '/.

⁵ Pour que l'espace abstrait de la fonction soit toujours connexe.

engendrent. Dans un second temps, nous décrivons deux approches numériques, certes plus coûteuses en temps de calcul, mais qui présentent l'avantage de s'appliquer plus systématiquement.

L'espace abstrait d'un vecteur y de \mathbb{R}^n construit à partir de m variables abstraites x_i ($i \in \{1, \dots, m\}$) et du champ de vecteur $y=f(x)$, définit un domaine de \mathbb{R}^n :

$$\mathcal{A}(y) = \{f(x) : x_i \in \mathcal{A}(x_i) ; \forall i \in \{1, \dots, m\}\} \quad (15)$$

avec f un champ de vecteur continu sur $\mathcal{A}(x)$ et $x=[x_1, \dots, x_m]^T$

Définissons les opérateurs centre $c(\cdot)$ et rayon $\rho(\cdot)$ ⁶ s'appliquant à des vecteurs de m variables bornées indépendantes $x=[x_1, \dots, x_m]^T$ comme suit :

$$\text{si } \forall i \in \{1, \dots, m\}, \mathcal{A}(x_i) = [\tilde{x}_i, \hat{x}_i], \text{ alors } c(x) = [c(x_1) \cdots c(x_m)]^T \text{ avec } c(x_i) = \frac{\tilde{x}_i + \hat{x}_i}{2}$$

$$\text{et } \rho(x) = [\rho(x_1) \cdots \rho(x_m)]^T \text{ avec } \rho(x_i) = \frac{\hat{x}_i - \tilde{x}_i}{2}$$

Pour simplifier les écritures, introduisons aussi la notation condensée : $R(x) = \text{diag}(\rho(x))$.

Ces notations permettent de normaliser la description d'un domaine ; en effet, l'espace abstrait du vecteur x est parfaitement décrit par la représentation normalisée suivante :

$$\mathcal{A}(x) = \{c(x) + R(x)u : u \in \mathbb{R}^m / \|u\|_\infty \leq 1\}^7$$

A partir de cette nouvelle représentation, l'expression (15) peut être normalisée de la façon suivante :

$$\mathcal{A}(y) = \{f_x(u) : u \in \mathbb{R}^m / \|u\|_\infty \leq 1\} \text{ avec } y = f_x(u) = f(c(x) + R(x)u) \quad (16)$$

où $c(x)$ et $R(x)$ sont des grandeurs déterministes.

Pour déterminer les formes des espaces abstraits, nous proposons un formalisme reposant sur des considérations ensemblistes. Commençons par décomposer le champ de vecteur $f_x(u)$ en n fonctions : $f_x(u) = [f_{x_1}(u) \dots f_{x_n}(u)]^T$.

Notations et définitions préliminaires

Nous noterons $F=\{1, \dots, n\}$, l'ensemble des entiers faisant référence à chacune des fonctions $f_{x_i}(u):i \in \{1, \dots, n\}$ du champ de vecteur $f_x(u)$, et $F_i \subset F$ tout sous-ensemble de F ne faisant référence qu'à certaines fonctions de $f_x(u)$. Si $F_j \subset F_i, F_i \setminus F_j = C_{F_i}(F_j)$ ⁸ désignera les fonctions de

⁶ Certains auteurs préfèrent utiliser l'opérateur profondeur (width), noté $w(\cdot)$, qui correspond au double de l'opérateur rayon.

⁷ La norme l_∞ d'un vecteur x de \mathbb{R}^m a pour définition $\|x\|_\infty = \max_i |x_i|$ avec $x = [x_1, \dots, x_m]^T$.

⁸ La notation $C_F(G)$ désigne le complémentaire de l'ensemble G dans l'ensemble F .

F_i n'apparaissant pas dans F_j .

En considérant qu'une variable abstraite u_j du vecteur u n'intervient pas dans une fonction $f_{x_i}(u)$ si $f_{x_i}(u) = f_{x_i}(u) \Big|_{u_j=0}, \forall u : \|u\|_\infty \leq 1$, nous pouvons définir l'ensemble U^{F_i} des variables abstraites intervenant dans le sous-ensemble de fonctions $F_i \subset F$ ⁹. Nous supposons que le champ de vecteur $f_x(u)$ est représenté sous sa forme minimale, c'est-à-dire que toutes les variables abstraites contenues dans le vecteur u apparaissent au moins une fois dans une des fonctions $f_{x_i}(u)$. Il est évident que cette hypothèse ne constitue en rien une contrainte, puisqu'il est toujours possible de reformuler le champ vectoriel f de telle sorte qu'elle soit vérifiée. Dès lors, nous pourrions toujours écrire que : $U^F = \bigcup_{i=1}^n U^{F_i}$. Nous associerons à chaque ensemble de variables U^{F_i} , un vecteur u^{F_i} où chacune des variables de U^{F_i} y figure selon un ordre choisi. Un champ vectoriel associé à un sous-ensemble F_i de fonctions sera noté $f_x^{F_i}(\cdot)$; il sera décrit par $y^{F_i} = f_x^{F_i}(u^{F_i})$ où y^{F_i} est un vecteur de variable contenues dans y (cf (16)).

Introduisons la notion de variables communes à toutes les fonctions d'un ensemble $F_i \subset F$. Ces variables appartiennent nécessairement à l'ensemble $U^{D(F_i)}$ défini par $U^{D(F_i)} = \bigcap_{j \in F_i} U^{F_j}$ et son complémentaire $U^{I(F_i)} = C_{U^{F_i}}(U^{D(F_i)})$. Enfin, plus globalement, les variables abstraites communes à au moins deux fonctions seront réunies dans $U^D = \bigcup_{F_i \subset F} U^{D(F_i)}$ et les variables de U^F restantes, dans son complémentaire sur U^F : $U^I = C_{U^F}(U^D)$.

Nous qualifierons le sous-ensemble de fonctions désignées par F_i , de résoluble si $card(F_i) \geq 2$ et si $card(U^{D(F_i)}) \geq card(F_i) - 1$ et nous noterons $R(F)$, l'ensemble des sous-ensembles F_i résolubles. Il est possible d'extraire des équations cartésiennes des fonctions F_i , qui correspondent à des domaines de circulation décrits lorsque les variables abstraites communes de $U^{D(F_i)}$ décrivent $\mathbb{R}^{card(U^{D(F_i)})}$. Nous ne nous intéresserons qu'aux champs $f_x(u)$ conduisant à des équations cartésiennes simples¹⁰. Pour obtenir ces équations cartésiennes, il est naturel de procéder par élimination des variables communes de $U^{D(F_i)}$. Il existe parfois plusieurs éliminations possibles ; il faudra privilégier celles qui conduisent à des équations cartésiennes simples. Cela peut quelquefois s'avérer impossible. Bien que, par soucis de clarté, nous n'ayons traité que d'équations cartésiennes sans conditions, les autres situations ne sont pas réductibles : les développements suivants peuvent être adaptés sans difficulté en propageant les conditions obtenues.

Lorsque $card(U^{D(F_i)}) = card(F_i) - 1$, il n'existe qu'une équation cartésienne ; par contre, lorsque

⁹ La lettre F sera réservée aux ensembles de fonctions alors que la lettre U se rapportera aux ensembles de variables abstraites et E aux ensembles d'équations cartésiennes.

¹⁰ Nous entendons par équation cartésienne simple une équation pouvant se mettre sous la forme $\Psi(u) = 0$.

$card(U^{D(F_i)}) > card(F_i) - 1$, il en existera $C_{card(U^{D(F_i)})}^{card(F_i)-1}$, déduites des $C_{card(U^{D(F_i)})}^{card(F_i)-1}$ combinaisons de $card(F_i) - 1$ variables de $U^{D(F_i)}$ pouvant être éliminées. Nous en déduisons la propriété suivante :

propriété d'un sous-ensemble résoluble

Un sous-ensemble résoluble F_i , c'est-à-dire vérifiant $card(F_i) \geq 2$ et $card(U^{D(F_i)}) \geq card(F_i) - 1$, engendre $C_{card(U^{D(F_i)})}^{card(F_i)-1}$ équations cartésiennes contenant toutes les variables $y^F_i = f_x^{F_i}(u^{F_i})$. Chacune de ces équations est indépendante de $card(F_i) - 1$ variables communes de $U^{D(F_i)}$.

Il est toujours possible de réécrire le champ vectoriel associé à un sous-ensemble résoluble F_i en distinguant les variables communes des autres : $y^F_i = f_x^{F_i}(u^{F_i}) = \varphi^{F_i}(u^{D(F_i)}, u^{I(F_i)})$. Intéressons-nous à la $j^{\text{ème}}$ ($j \in \{1, \dots, C_{card(U^{D(F_i)})}^{card(F_i)-1}\}$) équation cartésienne, notée $E_j^{F_i}$, associée au sous-ensemble résoluble F_i et notons $U_j^{D(F_i)}$, l'ensemble des $card(F_i) - 1$ variables communes de $U^{D(F_i)}$ n'apparaissant plus dans cette équation cartésienne. Les variables de $U^{D(F_i)}$ apparaissant dans la $j^{\text{ème}}$ équation cartésienne seront rassemblées dans l'ensemble $U_j^{D(F_i)\perp} = C_{U_j^{D(F_i)}}(U_j^{D(F_i)\perp})$. Dès lors, la $j^{\text{ème}}$ équation cartésienne peut être représentée par :

$$(E_j^{F_i}): \Psi_j^{F_i}(y^{F_i}, u^{I(F_i)}, u_j^{D(F_i)\perp}) = 0 \tag{17}$$

où $u_j^{D(F_i)\perp}$ contient les variables de $U_j^{D(F_i)\perp}$

Cette équation cartésienne représente le domaine décrit par y^{F_i} quand les variables de $U_j^{D(F_i)}$ décrivent $R^{card(F_i)-1}$.

Nous noterons E^{F_i} l'ensemble des $C_{card(U^{D(F_i)})}^{card(F_i)-1}$ équations cartésiennes $E_j^{F_i}$ pouvant être déduites du sous-ensemble résoluble F_i et E^F , l'ensemble des équations cartésiennes déduites de tous les sous-ensembles résolubles de $R(F)$.

Bilan des notations

- $f_x(u)$ champ de vecteur faisant intervenir des variables abstraites bornées
- $f_{x_i}(u)$ la $i^{\text{ème}}$ fonction composant le champ de vecteur $f_x(u)$
- $f_x^{F_i}(\cdot)$ champ vectoriel correspondant à un sous-ensemble F_i .
- y^{F_i} variables associées au champ vectoriel $f_x^{F_i}(u^{F_i})$ définie par $y^{F_i} = f_x^{F_i}(u^{F_i})$
- F ensemble des indices désignant les fonctions composant un champ vectoriel $f_x(u)$
- F_i sous-ensemble de F , désignant certaines fonctions d'un champ vectoriel $f_x(u)$
- $F_i \setminus F_j$ ensemble des indices apparaissant dans F_i et absents de F_j
- U^F ensemble des variables abstraites intervenant dans un champ vectoriel $f_x(u)$
- U^{F_i} ensemble des variables abstraites intervenant dans les fonctions désignées par F_i
- u^{F_i} vecteur contenant les variables de U^{F_i} et tel que $y^{F_i} = f_x^{F_i}(u^{F_i})$
- $U^{D(F_i)}$ ensemble des variables abstraites communes aux fonctions désignées par F_i

$u^{D(F_i)}$	vecteur contenant les variables de $U^{D(F_i)}$ et tel que $f_x^{F_i}(u^{F_i}) = \varphi^{F_i}(u^{D(F_i)}, u^{I(F_i)})$
$U_j^{D(F_i)}$	ensemble des variables communes de $U^{D(F_i)}$ n'apparaissant pas dans la $j^{\text{ème}}$ équation cartésienne $E_j^{F_i}$
$U_j^{D(F_i) \setminus \angle}$	ensemble des variables communes de $U^{D(F_i)}$ apparaissant dans la $j^{\text{ème}}$ équation cartésienne $E_j^{F_i}$
$U^{I(F_i)}$	ensemble des variables abstraites apparaissant dans une seule des fonctions d'un ensemble désigné par F_i
$u^{D(F_i)}$	vecteur contenant les variables de $U^{I(F_i)}$ et tel que $f_x^{F_i}(u^{F_i}) = \varphi^{F_i}(u^{D(F_i)}, u^{I(F_i)})$
U^D	ensemble des variables abstraites apparaissant dans plus d'une fonction d'un champ de vecteur $f_x(u)$
U^I	ensemble des variables abstraites apparaissant dans une seule fonction d'un champ de vecteur $f_x(u)$
$R(F)$	ensemble de tous les sous-ensembles F_i résolubles
$E_j^{F_i}$	$j^{\text{ème}}$ équation cartésienne associée au sous-ensemble résoluble F_i
E^{F_i}	ensemble des $C_{\text{card}(U^{D(F_i)})}^{\text{card}(F_i)-1}$ équations cartésiennes associées au sous-ensemble F_i
E^F	ensemble des équations cartésiennes déduites de tous les sous-ensembles résolubles de $R(F)$
$D_j^{F_i}$	domaine représenté par une équation cartésienne E^{F_i}

Notion d'indépendance et circonscription par un orthotope

Le champ de vecteur définissant le domaine abstrait que nous recherchons se note $y = f_x(u) = \varphi^F(u^D, u^I)$. Dans le cas où $U^D = \emptyset$, nous pouvons décomposer ce champ en n fonctions indépendantes $y_i = \varphi^{(i)}(u^{I(\{i\})})$. Puisque nous ne nous intéressons qu'aux fonctions continues sur $\mathcal{A}(u)$, chacune de ces fonctions admet des bornes finies :

$$\mathcal{A}(y_i) = [\tilde{y}_i, \hat{y}_i] \text{ avec } \tilde{y}_i = \inf_{\|u^{I(\{i\})}\|_{\infty} \leq 1} (\varphi^{(i)}(u^{I(\{i\})})) \text{ et } \hat{y}_i = \sup_{\|u^{I(\{i\})}\|_{\infty} \leq 1} (\varphi^{(i)}(u^{I(\{i\})})) \quad (18)$$

Puisque les fonctions $\varphi^{(i)}(u^{I(\{i\})})$ atteignent leurs bornes indépendamment les unes des autres, il en découle la propriété d'indépendance suivante :

propriété d'indépendance

L'espace abstrait de $y = f_x(u)$, défini par $\mathcal{A}(y) = \{f_x(u) : u \in \mathbb{R}^m / \|u\|_{\infty} \leq 1 (y \in \mathbb{R}^n), \text{ pour lequel } U^D = \emptyset, \text{ est un orthotope}^{11} \text{ de } \mathbb{R}^n \text{ aligné}^{12} \text{ par rapport aux axes du repère et noté } \square \mathcal{A}(y). \text{ Nous dirons que les variables de } y \text{ sont indépendantes les unes des autres.}$

Cette propriété définit l'indépendance : deux variables bornées x et y sont indépendantes si l'espace abstrait de $[x \ y]^T$ est un orthotope aligné. Cette définition s'étend à l'indépendance de deux vecteurs ; ils seront indépendants si chacune des variables de l'un est indépendante des

¹¹ Domaine délimité par des hyperplans orthogonaux entre eux (vulgairement une boîte).

¹² Nous parlerons désormais d'orthotope aligné circonscrit (OAC) équivalent à la notion anglo-saxonne de *minimum volume outer box* (MOB).

variables de l'autre.

exemple

En guise d'illustration, prenons l'exemple de 2 champs de vecteurs aux structures similaires :

$$f_x(u) = \begin{bmatrix} u_1^2(u_2 - 3) - 1 \\ \sqrt{u_2 + 2} \end{bmatrix} \text{ et } g_x(u) = \begin{bmatrix} u_1^2(u_3 - 3) - 1 \\ \sqrt{u_2 + 2} \end{bmatrix} \text{ avec } \left\| \begin{bmatrix} u_1 & u_2 & u_3 \end{bmatrix}^T \right\|_\infty \leq 1.$$

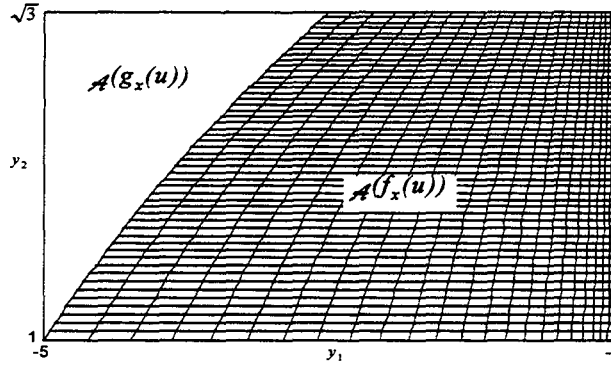


Figure V- Espace abstrait des champs de vecteurs $f_x(u)$ et $g_x(u)$

Le premier champ ne remplit pas les conditions d'indépendance puisque la variable abstraite u_2 apparaît dans les 2 composantes du champ tandis que le second la vérifie. La figure V révèle que, malgré la non linéarité des champs par rapport aux variables abstraites, l'espace abstrait du second champ est un orthotope alors que le premier conduit à un domaine plus complexe.

Dans le cas où l'ensemble U^D est non vide, les bornes \tilde{y}_i et \hat{y}_i , seront données par :

$$\mathcal{A}(y_i) = [\tilde{y}_i, \hat{y}_i] \text{ avec } \tilde{y}_i = \inf_{\|u\|_\infty \leq 1} (f_{x_i}(u)) \text{ et } \hat{y}_i = \sup_{\|u\|_\infty \leq 1} (f_{x_i}(u)) \tag{19}$$

Ces bornes sont, par définition, atteintes pour au moins une valeur de u ; autrement dit, quel que soit $i \in \{1, \dots, n\}$, la variable y_i vérifie $\tilde{y}_i \leq y_i \leq \hat{y}_i$. De plus, il existe une valeur de $u \in \{u / \|u\|_\infty \leq 1\}$ telle que $y_i = \tilde{y}_i$ et une autre telle que $y_i = \hat{y}_i$. Or, les mêmes bornes seraient atteintes par un champ vectoriel de même structure mais pour lequel toutes les variables de y seraient indépendantes les unes des autres. Nous en déduisons :

circonscription par un orthotope aligné

L'orthotope aligné $\square \mathcal{A}(y)$, circonscrit à l'espace abstrait défini par $\mathcal{A}(y) = \{f_x(u) : u \in \mathbb{R}^m / \|u\|_\infty \leq 1\}$, correspond à l'espace abstrait qui serait obtenu en considérant toutes les variables abstraites comme étant indépendantes les unes des autres.

L'exemple représenté sur la figure V illustre aussi cette propriété.

Le fait que chaque variable bornée atteigne ses bornes ne signifie nullement que l'espace abstrait soit un orthotope aligné. En effet, la présence de variables communes : $U_D \neq \emptyset$, produit des dépendances qui expliquent les déformations de l'espace abstrait $\mathcal{A}(y)$.

Détermination analytique d'un domaine abstrait

Nous savons désormais circonscrire l'espace abstrait $\mathcal{A}(y)$ par l'orthotope aligné $\square\mathcal{A}(y)$; cependant, nous ignorons toujours la nature précise de cet espace abstrait. Nous nous intéresserons désormais aux moyens d'en déterminer la forme. Il existe certes des méthodes complètement numériques sur lesquelles nous nous arrêterons un peu plus tard. Nous avons tenté ici de proposer une approche analytique, plus utile à la mise en évidence de propriétés fondamentales. Nous avons vu que les variables de U^D expliquaient que l'espace abstrait $\mathcal{A}(y)$ ne s'identifiait plus à un orthotope aligné ; aussi, nous faut-il nous consacrer à l'étude de l'influence des variables de U^D qui engendrent des déformations du domaine $\mathcal{A}(y)$.

Pour déterminer analytiquement la forme de l'espace abstrait $\mathcal{A}(y)$, nous devons étudier l'influence des variables communes U^D . En effet, les frontières d'un domaine abstrait (Figure VI), à l'exception de celles qui sont confondues avec l'orthotope circonscrit $\square\mathcal{A}(y)$, sont dues à ces variables ; elles correspondent donc à des équations cartésiennes appartenant a priori à l'ensemble E^F des équations cartésiennes qu'il est possible de générer. Pour obtenir les frontières confondues avec $\square\mathcal{A}(x)$, qui ne sont en général pas dues à des variables communes, il est nécessaire de tenir compte de l'orthotope aligné circonscrit : $\square\mathcal{A}(y)$. Nous en déduisons qu'a priori, l'espace abstrait $\mathcal{A}(y)$ doit s'identifier à l'intersection de l'orthotope circonscrit $\square\mathcal{A}(y)$ avec l'intersection des domaines D_j^{Fi} associés à chacune des équations cartésiennes de E^F . Or, si ceci se vérifie souvent dans la pratique, cette assertion n'est pas corroborée par la théorie.

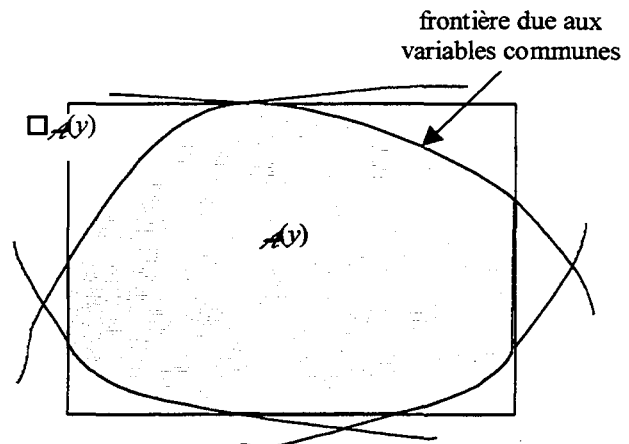


Figure VI – Domaine abstrait et ses frontières

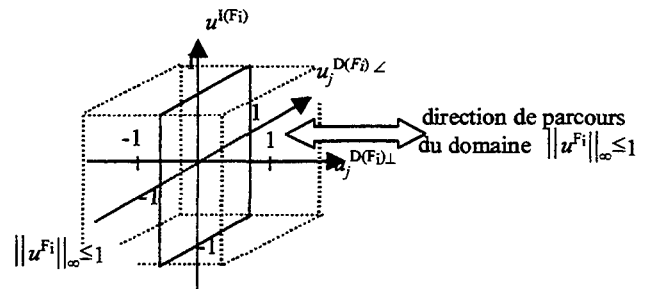


Figure VII – Parcours équivalent du domaine $\mathcal{A}(u^{Fi})$ par une équation cartésienne E_j^{Fi}

Le domaine décrit par les variables abstraites u est un hypercube : $\|u\|_{\infty} \leq 1$. Chaque domaine D_j^{Fi} contient cet hypercube puisque chaque équation cartésienne D_j^{Fi} englobe le domaine $\|u^{Fi}\|_{\infty} \leq 1$: $\underline{u}_j^{D(Fi)\perp}$ peut prendre une valeur quelconque pourvu que $\|u^{I(Fi)}\|_{\infty} \leq 1$ et que $\|u_j^{D(Fi)\angle}\|_{\infty} \leq 1$ (cf. figure VII).

Le domaine D_j^{Fi} se déduit de (17) :

$$D_j^{F_i} = \{y^{F_i} / \Psi_j^{F_i}(y^{F_i}, u^{I(F_i)}, u_j^{D(F_i)\angle}) = 0; \|u^{I(F_i)}\|_\infty \leq 1, \|u^{D(F_i)\angle}\|_\infty \leq 1\} \quad (20)$$

En parcourant le domaine $\|u^{F_i}\|_\infty \leq 1$ selon toutes les directions $u_i^{D(F_i)\perp}$ et en formant l'intersection des domaines $\{\|u^{I(F_i)}\|_\infty \leq 1, \|u_j^{D(F_i)\angle}\|_\infty \leq 1\}$, nous retrouvons le domaine original :

$$\|u^{F_i}\|_\infty \leq 1 = \bigcap_j \{\|u^{I(F_i)}\|_\infty \leq 1, \|u_j^{D(F_i)\angle}\|_\infty \leq 1\} \quad (21a)$$

Puis, en intersectant tous les domaines $\|u^{F_i}\|_\infty \leq 1$ associés aux sous-ensembles résolubles ainsi que le domaine associé aux variables de U^J , nous retrouvons le domaine original :

$$\{ \|u\|_\infty \leq 1 \} = \bigcap_{F_i \in \mathbb{R}(F)} \{ \|u^{F_i}\|_\infty \leq 1 \} \cap \{ \|u^I\|_\infty \leq 1 \} \quad (21b)$$

Par conséquent, en appliquant le champ $f_x(\cdot)$ à (21b), nous obtenons que le domaine abstrait de $f_x(u)$ vérifie :

$$\mathcal{A}(y) = \{ f_x(u) : \|u\|_\infty \leq 1 \} \subseteq \bigcap_{F_i \in \mathbb{R}(F)} \{ y^{F_i} = f_x^{F_i}(u^{F_i}) : \|u^{F_i}\|_\infty \leq 1 \} \cap \square \mathcal{A}(y) \quad (22)$$

Arrêtons-nous un instant sur le sens de cette inclusion. Elle révèle en effet que, théoriquement, le domaine de droite, dans (22), n'est qu'une majoration du domaine $\mathcal{A}(y)$. Pour qu'il y ait identité, il faut que l'application des composantes du champ $f_x(\cdot)$ à l'intersection des domaines des u^{F_i} , s'identifie à l'intersection des champs élémentaire $f_x^{F_i}$ appliqués à chacun de

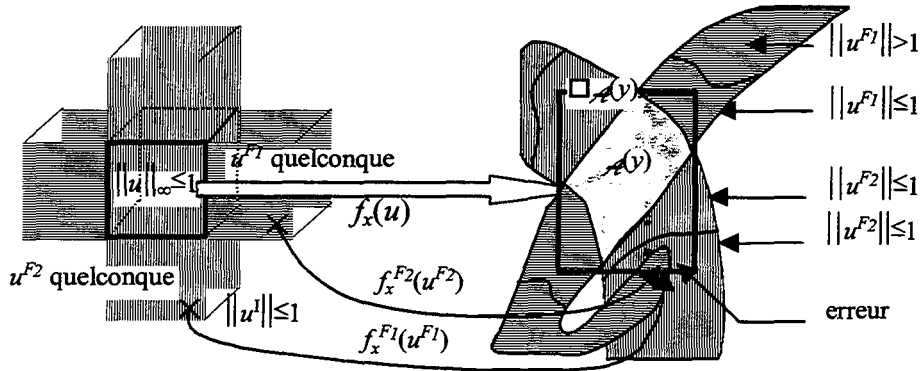


Figure VIII – Exemple de champ non injectif

ces domaines. Or, la figure VIII montre bien que ceci est obtenu uniquement lorsque le champ de vecteur $f_x(\cdot)$ est injectif de $\mathbb{R}^{\dim(u)} \setminus \{u / \|u\|_\infty \leq 1\}$ vers $\square \mathcal{A}(y)$! Dans le cas contraire, nous n'obtiendrions qu'une majoration du domaine $\mathcal{A}(y)$ recherché. Nous pouvons dès lors formuler l'hypothèse suivante :

hypothèse d'injectivité

Le champ de vecteur $f_x(\cdot)$ est injectif de $\mathbb{R}^{\dim(u)} \setminus \{u / \|u\|_\infty \leq 1\}$ vers $\square \mathcal{A}(y)$.

En utilisant (21a) et (20), nous déduisons le résultat important suivant :

$$\mathcal{A}(y) = \{f_x(u) : \|u\|_\infty \leq 1\} \subseteq \bigcap_{F_i \in \mathcal{R}(F)} \bigcap_{j=1}^{\zeta} D_j^{F_i} \cap \square \mathcal{A}(y) \quad ; \quad \zeta = C_{\text{card}(U^{D(F_i)})}^{\text{card}(F_i)-1} \quad (23)$$

Néanmoins, notre objectif est d'obtenir les domaines exacts et non pas des majorations, et même si l'inclusion de (23) peut se transformer en identité, l'hypothèse d'injectivité n'est guère facile à vérifier ! Nous avons alors écarté, dans un premier temps, la relation (23).

Orientons-nous vers une solution se fondant exclusivement sur la relation (22) en essayant de reconstruire directement le domaine possible de y^{F_i} . Pour ce faire, fixons les variables de y^{F_i} à des valeurs cohérentes avec l'orthotope $\square \mathcal{A}(y)$, à l'exception toutefois d'une variable $y^{\{k\}}$: $k \in F_i$. Notre démarche va consister à rechercher le domaine que peut décrire la variable $y^{\{k\}}$ lorsque $y^{F_i \setminus \{k\}}$ prend la valeur arbitraire v . Il va de soi que le résultat ne dépend pas de la variable $y^{\{k\}}$ choisie ; néanmoins, certains choix peuvent s'avérer plus judicieux que d'autres vis-à-vis de la complexité du résultat. Notons qu'il ne suffit pas d'introduire la valeur v dans l'équation cartésienne (20) pour prendre en compte le fait que $y^{F_i \setminus \{k\}} = v$. En effet, en fixant cette valeur, nous restreignons l'ensemble des valeurs possibles de certaines variables abstraites de $U^{I(F_i)}$ et de $U_j^{D(F_i) \setminus \{k\}}$. En introduisant les nouvelles notations :

$$\begin{aligned} U^{I(F_i) \setminus \{k\}} &= U^{F_i \setminus \{k\}} \cap U^{I(F_i)}, \quad \overline{U}_j^{(F_i) \setminus \{k\}} = C_{U^{F_i \setminus \{k\}}} (U_j^{D(F_i) \perp} \cup U_j^{D(F_i) \setminus \{k\}} \cup U^{I(F_i)}) \\ \overline{U}_j^{D(F_i) \perp} &= C_{U_j^{D(F_i) \perp}} (U_j^{D(F_i) \perp}), \quad \overline{U}_j^{D(F_i) \setminus \{k\}} = C_{U_j^{D(F_i) \setminus \{k\}}} (U_j^{D(F_i) \setminus \{k\}}), \quad \overline{U}^{I(F_i)} = C_{U^{I(F_i)}} (U^{I(F_i)}) \end{aligned}$$

les vecteurs associés seront déduits directement.

nous pouvons réécrire l'équation cartésienne (20) ainsi que l'expression des variables $y^{F_i \setminus \{k\}} = f_x^{F_i \setminus \{k\}}(u^{F_i \setminus \{k\}})$ fixée à la valeur v :

$$\Psi_j^{F_i} (v, y^{\{k\}}, \overline{u}^{I(F_i)}, u^{I(F_i)}, u_j^{D(F_i)}) = 0 \quad (24a)$$

$$v = \Phi \left(\overline{u}_j^{(F_i) \setminus \{k\}}, u^{I(F_i)}, u_j^{D(F_i) \setminus \{k\}}, u_j^{D(F_i) \perp} \right) \quad (24b)$$

En fixant $y^{F_i \setminus \{k\}} = v$ et en décrivant le domaine $D_j^{F_i}$, il se trouve que $u^{D(F_i) \perp} \in \mathbb{R}^{\dim(u^{D(F_i) \perp})}$ (cf. (20)). Dans ces conditions, la relation (24b) définit une contrainte sur les valeurs que peuvent prendre $u^{I(F_i)}$ et $u_j^{D(F_i) \setminus \{k\}}$. De plus, puisque les variables $u_j^{D(F_i) \perp}$ sont réapparues, nous pouvons trouver le domaine possible de y^{F_i} en imposant $\|u_j^{D(F_i) \perp}\|_\infty \leq 1$ (ce qui correspond, en rassemblant toutes les contraintes sur les normes à $\|u^{F_i}\|_\infty \leq 1$). Nous déduisons les valeurs possibles de $u^{I(F_i)}$ et de $u_j^{D(F_i) \setminus \{k\}}$ dans le domaine $\mathcal{V}_j^{F_i \setminus \{k\}}$ défini par :

$$\mathcal{V}_j^{F_i \setminus \{k\}}(v) = \left\{ \begin{aligned} & \left[\begin{array}{l} u^{I(F_i)} \\ u_j^{D(F_i) \setminus \{k\}} \end{array} \right] / v = \Phi \left(\overline{u}_j^{(F_i) \setminus \{k\}}, u^{I(F_i)}, u_j^{D(F_i) \setminus \{k\}}, u_j^{D(F_i) \perp} \right); \dots \\ & \dots \left\| \overline{u}_j^{(F_i) \setminus \{k\}} \right\|_\infty \leq 1, \quad \left\| u_j^{D(F_i) \perp} \right\|_\infty \leq 1 \end{aligned} \right\} \quad (25)$$

illustration

Supposons que la relation (24b) soit de la forme $v=f(u_1, u_2)$ avec u_1 et u_2 scalaires, et que le domaine que peuvent décrire les variables abstraites u_1 et u_2 soit celui représenté sur la figure IX. Alors le domaine $\mathcal{V}(v)$ que l'on obtiendrait alternativement si $\mathcal{V}(v)$ correspondait à un domaine possible de u_1 ou à un domaine de u_2 , pourrait être représenté par des intervalles (cf. figure IX).

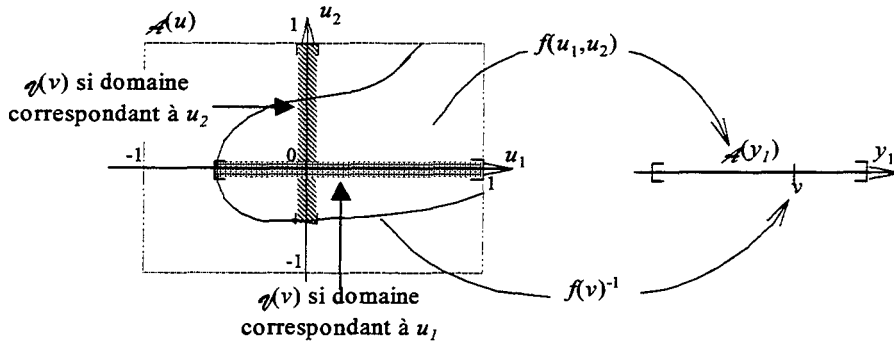


Figure IX- Image inverse d'une valeur de $\mathcal{A}(y_1)$

Par conséquent, dans un cas général, et même si l'hypothèse d'injectivité n'est pas satisfaite, pour obtenir une expression analytique du domaine possible de y^{F_i} , il suffit que le domaine donné par (25) admette une expression analytique. Or, de telles expressions sont généralement obtenues de l'arithmétique des intervalles, c'est-à-dire qu'avec cette approche, les solutions purement analytiques sont réservées aux domaines définis par des champs vectoriels arithmétiques et ne contenant que peu de variables communes. Mais, même si des résultats purement analytiques ne peuvent être obtenus, il est toujours possible de simplifier la reconstruction d'un domaine par une approche numérique.

En guise de synthèse, rappelons qu'après avoir déterminé l'orthotope $\square \mathcal{A}(y)$, le calcul de l'espace abstrait $\mathcal{A}(y)$ se résume, avec l'approche développée, à examiner tous les sous-ensembles résolubles F_i de $R(F)$ et à en extraire, pour chacun d'eux, une unique équation cartésienne, $E_j^{F_i}$, judicieusement choisie. Puis, pour chacune des équations cartésiennes, il s'agit de choisir une variable $y^{\{k\}}$ de y^{F_i} telle que le domaine $\mathcal{V}_j^{F_i \setminus \{k\}}$ (25) associé soit le plus simple possible. Ce domaine connu, il ne reste plus qu'à déterminer les valeurs possibles de $y^{\{k\}}$ en fonction des différents $y^{F_i \setminus \{k\}}=v$ compatibles avec $\square \mathcal{A}(y)$, à partir de :

$$\Psi_j^{F_i}(v, y^{\{k\}}, \bar{u}^{\text{Ir}(F_i)}, u^{\text{Ir}(F_i)}, u_j^{\text{D}(F_i)}) = 0 \text{ avec : } \begin{bmatrix} u^{\text{Ir}(F_i)} \\ u_j^{\text{D}(F_i)} \end{bmatrix} \in \mathcal{V}_j^{F_i \setminus \{k\}}, \|\bar{u}^{\text{Ir}(F_i)}\|_\infty \leq 1$$

Enfin, le domaine exact $\mathcal{A}(y)$ se déduit de l'intersection entre l'orthotope $\square \mathcal{A}(y)$ et chacun des domaines possibles pour y^{F_i} . Remarquons, pour finir, que certains domaines associés à un sous-ensemble résoluble peuvent être redondants dans la mesure où leur intersection ne contribue pas à une meilleure détermination de $\mathcal{A}(y)$; cependant, il n'est guère possible, dans un cas général, de prévoir quels seront les sous-ensembles résolubles les plus pertinents.

Exemple 1

Partons d'un champ de vecteur définissant le vecteur y de variables bornées, ne vérifiant pas l'hypothèse d'injectivité :

$$y = \begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = \begin{bmatrix} u_1^2 + 2u_2 \\ 2u_1u_2 \end{bmatrix} \text{ avec : } \|u_1\|_\infty \leq 1 \text{ et } \|u_2\|_\infty \leq 1$$

Pour déterminer le domaine $\square \mathcal{A}(y)$, faisons l'hypothèse que les deux fonctions sont indépendantes, autrement dit, remplaçons la seconde par $y_2' = 2u_3u_4$ où $\|u_3\|_\infty \leq 1$ et $\|u_4\|_\infty \leq 1$. Sans difficulté, nous établissons : $A(y_1) = [-2 \ 3]$ et $A(y_2)' = [-2 \ 2]$. Par ailleurs, ce champ de vecteurs ne conduit qu'à un unique ensemble résoluble (par rapport à u_1 ou à u_2). Résolvons-le d'une part, par élimination de u_1 , d'autre part, par élimination de u_2 , et dans les 2 cas, fixons y_1 pour des raisons de simplicité. Nous obtenons les 2 équations cartésiennes suivantes :

$$U_1^{D(F)\perp} = \{u_2\} : y_2 = u_1y_1 - u_1^3 \quad : \max(0, y_1 - 2) \leq u_1^2 \leq \max(0, \text{sat}(y_1 + 2))$$

$$U_2^{D(F)\perp} = \{u_1\} : y_2^2 = 4u_2^2(y_1 - 2u_2) : \text{sat}(\frac{1}{2}(y_1 - 1)) \leq u_2 \leq \text{sat}(\frac{1}{2}y_1)$$

où la fonction saturation $\text{sat}(x) = x$ si $-1 \leq x \leq 1$; $\text{sat}(x) = 1$ si $x > 1$ et ; $\text{sat}(x) = -1$ si $x < -1$

Après chaque équation cartésienne, nous avons fait figurer le domaine \mathcal{V} associé pour y_1 fixé. Ainsi, dans les 2 cas, nous sommes partis de la fonction $y_1 = u_1^2 + 2u_2$. Dans la première équation, la variable abstraite commune éliminée était u_2 ; il nous faut donc la considérer comme pouvant fluctuer librement sur l'intervalle $[-1, 1]$. Nous en déduisons que $\max(0, y_1 - 2) \leq u_1^2 \leq \max(0, \text{sat}(y_1 + 2))$.

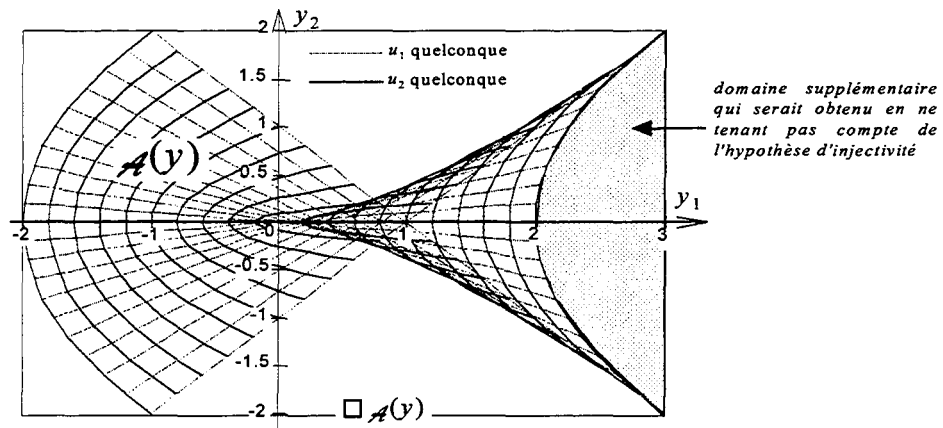


Figure X- Domaine $\mathcal{A}(y)$ et circulations suivant les dépendances

Concernant la seconde équation, nous procédons de la même façon en tenant cette fois-ci la variable u_1 comme variable commune éliminée. Le domaine obtenu est illustré par la figure X.

Ces deux équations, issues de deux résolutions différentes, caractérisent bien un même domaine : la première, plus simple, aurait suffi. Ceci corrobore le fait qu'une seule équation cartésienne par sous-ensemble résoluble suffit.

Bien que la méthode que nous venons de détailler puisse conduire à une expression analytique du domaine $\mathcal{A}(y)$, la recherche des domaines $\mathcal{V}_j^{F \setminus \{k\}}$ ($\forall F_i \in R(F)$) sous une forme analytique peut s'avérer difficile voire impossible. Si l'hypothèse d'injectivité pouvait être vérifiée, l'identité se substituerait à l'inclusion dans la relation (23a), et nous pourrions ainsi suivre une

autre approche. Or, nous l'avons évoqué, il n'est pas évident de vérifier si un champ de vecteur satisfait à l'hypothèse d'injectivité. Le paragraphe suivant met en évidence une classe très répandue de champs vectoriels satisfaisant à cette hypothèse.

Cas particulier des champs dissociables

L'objectif de ce paragraphe est de rechercher une condition suffisante, et simplement vérifiable, pour qu'un champ vectoriel satisfasse à l'hypothèse d'injectivité. Introduisons la notion de champ vectoriel dissociable.

Un champ vectoriel φ^{F_i} , associé à un sous-ensemble résoluble F_i , sera qualifié de dissociable si chacune de ses dérivées partielles par rapport aux variables abstraits communes $u_k^{D(F_i)}$ est indépendante des autres variables communes. La transcription mathématique de cette définition est la suivante :

$$\frac{\partial \varphi^{F_i}(u^{I(F_i)}, u^{D(F_i)})}{\partial u_k^{D(F_i)}} = t_k(u^{I(F_i)}, u_k^{D(F_i)}) \quad ; \quad \forall k \in \{1, \dots, \text{card}(U^{D(F_i)})\}$$

$$\text{avec } \|u^{I(F_i)}\|_\infty \leq 1 \text{ et } \|u_k^{D(F_i)}\|_\infty \leq 1$$

Pour faire apparaître les propriétés des champs vectoriels φ^{F_i} dissociables, remarquons que la dérivée de Lie de tout champ vectoriel φ^{F_i} dissociable, suivant la direction u^{F_i} peut s'exprimer ainsi :

$$L_{u^{F_i}}(\varphi^{F_i}(u^{I(F_i)}, u^{D(F_i)})) = \sum_{j=1}^{\text{card}(u^{D(F_i)})} \frac{\partial \varphi^{F_i}(u^{I(F_i)}, u^{D(F_i)})}{\partial u_j^{I(F_i)}} u_j^{I(F_i)} + \sum_{k=1}^{n^{F_i}} t_k(u^{I(F_i)}, u_k^{D(F_i)}) u_k^{D(F_i)}$$

Supposons que nous décrivions le domaine possible des y^{F_i} en suivant les directions données par les variables abstraites $u_j^{D(F_i)}$. Dans ces directions, la dérivée de Lie précédente se simplifie :

$$L_{u^{D(F_i)}}(\varphi^{F_i}(u^{I(F_i)}, u_j^{D(F_i)})) = \sum_{k=1}^{\text{card}(U_j^{D(F_i)\perp})} t_k(u^{I(F_i)}, u_{j,k}^{D(F_i)\perp}) u_{j,k}^{D(F_i)\perp}$$

Le domaine décrit est donc indépendant des autres variables $u_j^{D(F_i)\perp}$. Il s'ensuit naturellement que le domaine possible de y^{F_i} s'identifie à l'intersection des $C_{\text{card}(U^{D(F_i)})}^{\text{card}(F_i)-1}$ domaines $D_j^{F_i}$ pour lesquels les vecteurs $u_j^{D(F_i)}$ sont quelconques. **La dissociabilité d'un champ de vecteur implique nécessairement que l'hypothèse d'injectivité soit vérifiée !**

Si un domaine $\mathcal{A}(y)$ est engendré par un champ dissociable $f_x(\cdot)$, alors tout sous-ensemble de fonction $f_x^{F_i}(\cdot)$, qu'il soit résoluble ou pas, l'est aussi. Cela se traduit par la définition mathématique qui suit.

dissociabilité d'un champ de vecteur

Un champ de vecteurs représenté par $\varphi^F(u^I, u^D)$, où le vecteur u^D contient les variables communes à plusieurs fonctions et u^I les autres, sera qualifié de dissociable sur $\|u^I\|_\infty \leq 1$ et $\|u^D\|_\infty \leq 1$, si toutes ses dérivées partielles par rapport aux variables communes de u^D sont indépendantes des autres variables communes ; autrement dit, si elles vérifient :

$$\frac{\partial \varphi^F(u^I, u^D)}{\partial u_k^D} = t_k(u^I, u_k^D) \quad ; \quad \forall \|u^I\|_\infty \leq 1, \quad \forall \|u_k^D\|_\infty \leq 1 \text{ et } \forall k \in \{1, \dots, \text{card}(U^D)\} \quad (27)$$

Dans le cas de champ dissociable, l'hypothèse d'injectivité étant satisfaite, l'inclusion apparaissant dans la relation (23) devient une égalité :

$$\mathcal{A}(y) = \{f_x(u) : \|u\|_\infty \leq 1\} = \bigcap_{F_i \in \mathcal{R}(F)} \bigcap_{j=1}^{\zeta} D_j^{F_i} \cap \square \mathcal{A}(y) \quad ; \quad \zeta = C_{\text{card}(U^D(F_i))}^{\text{card}(F_i)-1} \quad (28)$$

Par conséquent, avec cette approche, il faut calculer le domaine $\square \mathcal{A}(y)$ et les j ($j \in \{1, \dots, C_{\text{card}(U^D(F_i))}^{\text{card}(F_i)-1}\}$) domaines possibles $D_j^{F_i}$ (20) de chacun des sous-ensembles résolubles.

Si le champ définissant un domaine abstrait est dissociable, la méthode générale s'applique aussi. La méthode que nous venons d'exposer permet d'éviter l'évaluation analytique souvent délicate des domaines $\mathcal{V}_j^{F_i \setminus \{k\}}$ (25), au prix certes d'un plus grand nombre de domaines à déterminer.

En effet, la méthode générale n'imposait l'évaluation que d'un seul domaine $D_j^{F_i}$ par sous-ensemble résoluble tandis que la méthode spécifique aux champs dissociables impose que tous les $C_{\text{card}(U^D(F_i))}^{\text{card}(F_i)-1}$ domaines associés à chaque sous-ensemble résoluble, soient évalués.

Le champ de vecteurs de l'exemple 1 n'est pas dissociable puisque les dérivées partielles par rapport à chacune des variables à l'origine de dépendance ne vérifient pas la condition (27) :

$$\frac{\partial f(u_1, u_2)}{\partial u_1} = \begin{bmatrix} 2u_1 \\ 2u_2 \end{bmatrix} \quad \text{et} \quad \frac{\partial f(u_1, u_2)}{\partial u_2} = \begin{bmatrix} 2 \\ 2u_1 \end{bmatrix}$$

L'approche que nous venons de décrire ne s'applique donc pas à de tels champs vectoriels. Choisissons un autre exemple dont le champ est dissociable.

Exemple 2

Soit un domaine $\mathcal{A}(y)$ défini par le champ vectoriel suivant :

$$y = \begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = \begin{bmatrix} u_1^2 u_3 - u_2 \\ 4u_1 + u_2 \end{bmatrix} \quad \text{avec} : \|u_1\|_\infty \leq 1 \text{ et } \|u_2\|_\infty \leq 1$$

En considérant séparément les variables bornées y_1 et y_2 , nous déduisons l'orthotope aligné circonscrit $\square \mathcal{A}(y)$, caractérisé par : $\mathcal{A}(y_1) = [-2 \ 2]$ et $\mathcal{A}(y_2) = [-5 \ 5]$.

Le champ vectoriel précédent n'engendre qu'un ensemble $F=\{1,2\}$ résoluble et il est dissociable. Par conséquent, en éliminant alternativement u_1 puis u_2 , et en calculant les extrema de ces équations lorsque $\|u\|_\infty \leq 1$, les 2 domaines suivants peuvent être déduits :

$$\begin{cases} (D_1) : -5 \leq y_1 + y_2 \leq 5 \\ (D_2) : -1 - \frac{1}{6}(1 - y_2)^2 \leq y_1 \leq 1 + \frac{1}{6}(1 + y_2)^2 \end{cases}$$

En utilisant l'approche n'exploitant pas la dissociabilité de $\mathcal{A}(y)$, nous aurions obtenu l'équation suivante :

$$(D_0) : y_2 = -y_1 + u_1^2 u_3 + 4u_1 : \text{sat}(-1 + y_1) \leq u_1^2 u_3 \leq \text{sat}(1 + y_1)$$

De toute évidence, il est plus simple de calculer $\square \mathcal{A}(y) \cap (D_1) \cap (D_2)$ que $\square \mathcal{A}(y) \cap (D_0)$ car le domaine \mathcal{V} de (D_0) doit être recalculé pour toute valeur de y_1 . La figure XI illustre le domaine $\mathcal{A}(y)$ obtenu accompagné des domaines décrit par y lorsque, alternativement, les variables abstraites u_1 et u_2 sont fixés à des valeurs entre -1 et 1 . Cet exemple fait apparaître qu'un domaine $\mathcal{A}(y)$ dissociable est délimité par ces domaines.

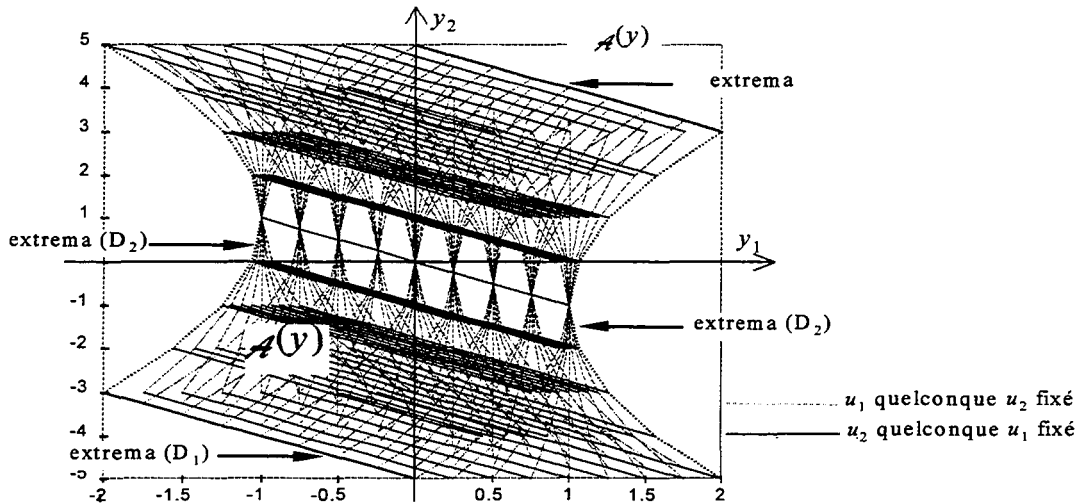


Figure XI- Domaine $\mathcal{A}(y)$ et circulation suivant les dépendances

Cas particulier des champs affines

Les domaines abstraits demeurent, même s'ils sont issus de champs dissociables, complexes et donc difficilement exploitables par une procédure de diagnostic. En étudiant les domaines délimités par des hyperplans¹³, nous escomptons identifier des structures aisément exploitables de champs vectoriels. Il n'est néanmoins guère possible de rechercher de telles structures dans le cas général puisque les frontières d'un domaine ne coïncident pas nécessairement avec des domaines obtenus en fixant certaines variables abstraites. Nous nommerons ces domaines, les *domaines de circulation* d'un espace abstrait ; ils figurent en

¹³ Un domaine compact délimité par des hyperplans est un polytope.

trait plein et en pointillés sur les figures V et XI.

Dans le cas particulier des domaines dissociables, la coïncidence entre certains domaines de circulation et les frontières d'un domaine $\mathcal{A}(y)$ est évidente ; elle découle de la définition de la dissociabilité. Dès lors, une condition suffisante pour qu'un domaine défini par (16) conduise à un polytope est que tous ses domaines de circulation soient des hyperplans : les variables abstraites communes de U^D doivent apparaître linéairement dans les fonctions de sorte que les domaines de circulation obtenus par élimination soient des hyperplans. La définition (27) permet de transcrire mathématiquement cette condition. Il suffit que :

$$\frac{\partial \varphi^F(u^I, u^D)}{\partial u_k^D} = t_k(u^I) \quad ; \quad \forall \|u^I\|_\infty \leq 1 \text{ et } \forall k \in \{1, \dots, \text{card}(U^D)\}$$

Par intégration, une structure de champ vectoriel engendrant un domaine dissociable et polytopique, est obtenue :

$$\varphi^F(u^I, u^D) = t_0(u^I) + \sum_{i=1}^{\text{card}(U^D)} t_i(u^I) u_i^D \quad (29)$$

Les champs vectoriels des exemples 1 et 2 ne sont pas affines par rapport aux variables communes de U^D . Un autre exemple est alors nécessaire à l'illustration des propriétés de ce type de champs.

Exemple 3

Intéressons-nous au champ vectoriel satisfaisant (29) suivant :

$$y = \begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = \begin{bmatrix} u_1 u_3 + u_2 (2 + u_4) \\ 2u_1 - u_2 \end{bmatrix} \text{ avec : } \left\| \begin{bmatrix} u_1 & u_2 & u_3 & u_4 \end{bmatrix}^T \right\|_\infty \leq 1$$

L'évaluation des espaces abstraits $\mathcal{A}(y_1)=[-4 \ 4]$ et $\mathcal{A}(y_2)=[-3 \ 3]$ définit l'orthotope $\square \mathcal{A}(y)$. Le domaine étant dissociable, il nous faut étudier les domaines décrits en circulant suivant les variables $U^D=\{u_1, u_2\}$. Les éliminations de u_1 et de u_2 conduisent à :

$$\begin{cases} (D_1) y_1 = \frac{1}{2} u_3 (y_2 + u_2) + u_2 (2 + u_4) \\ (D_2) y_1 = u_1 (4 + u_3 + 2u_4) - y_2 (2 + u_4) \end{cases}$$

Grâce à la propriété (13), ces domaines peuvent être explicités :

$$(D_1): y_1 \in [\alpha_1, \beta_1] \text{ avec :}$$

$$\alpha_1 = \frac{1}{2}(y_2 - 7) \text{ sur } y_2 \in]-\infty, 1] \text{ et } \alpha_1 = -\frac{1}{2}(5 + y_2) \text{ sur } y_2 \in [1, \infty[$$

$$\hat{\alpha}_1 = \frac{1}{2}(5 - y_2) \text{ sur } y_2 \in]-\infty, -1] \text{ et } \beta_1 = \frac{1}{2}(y_2 + 7) \text{ sur } y_2 \in [-1, \infty[$$

et

$$(D_2): y_2 \in [\alpha_2, \beta_2] \text{ avec :}$$

$$\alpha_2 = -(3 + y_2) \text{ sur } y_2 \in]-\infty, -2] \text{ et } \alpha_2 = -(7 + 3y_2) \text{ sur } y_2 \in [-2, \infty[$$

$$\beta_2 = 7 - 3y_2 \text{ sur } y_2 \in]-\infty, 2] \text{ et } \beta_2 = 3 + y_2 \text{ sur } y_2 \in [2, \infty[$$

De la relation (22), nous déduisons que $\mathcal{A}(y) = \square \mathcal{A}(y) \cap ((D_1) \cap (D_2))$. Ce domaine a été représenté sur la figure XII.

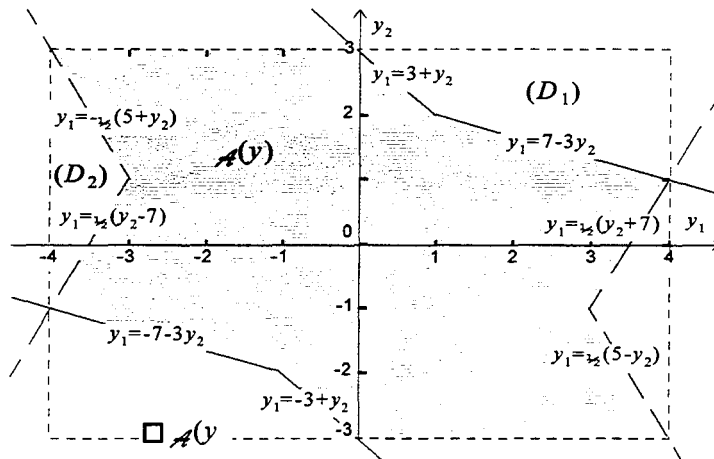


Figure XII- Domaine $\mathcal{A}(y)$ et frontières

Les domaines polytopiques obtenus de domaines satisfaisant à (29) ne sont a priori pas convexes. Une propriété largement utilisée par les approches bornantes est qu'un domaine polytopique convexe peut être parfaitement caractérisé par les coordonnées de ses sommets. Vicino et Zappa [Vicino et al,96] énumèrent certaines propriétés des champs vectoriels affines en les variables abstraites qui ont, parmi les plus intéressantes, celle d'engendrer des domaines abstraits convexes. De tels champs vectoriels satisfont à la relation suivante :

$$\varphi^F(u^I, u^D) = t_0 + \sum_{i=1}^{\text{card}(U^D)} t_i u_i^D + \sum_{j=\text{card}(U^D)+1}^{\text{card}(U)} t_j u_j^I = t_0 + T \begin{bmatrix} u^D \\ u^I \end{bmatrix} \quad (30)$$

avec $T = [t_1 \dots t_m] : m = \text{card}(U)$

Un sens peut être donné à t_0 en revenant à l'expression (16) originale du champ vectoriel. En effet, dans le cas affine, la normalisation des variables abstraites x conduit à :

$$y = f(c(x) + R(x)u) = f(c(x)) + f(R(x)u)$$

En posant $y_0 = f(c(x))$ et $f_x(u) = f(R(x)u)$ et en comparant l'écriture précédente à (30), il résulte simplement que $y_0 = t_0$. Ainsi, d'une manière générale, un espace abstrait polytopique et engendré par un champ affine s'écrira :

$$y = y_0 + T \begin{bmatrix} u^D \\ u^I \end{bmatrix} = y_0 + T_0 u$$

La différence entre les matrices T et T_0 se justifie par la répartition a priori différente des variables du vecteur u dans $[u^{D^T} \ u^{I^T}]^T$

Montrons que y_0 est bien le centre de symétrie de l'espace abstrait $\mathcal{A}(y)$. Si un point y_1

appartient à $\mathcal{A}(y)$ alors il existe un ensemble E_u de points u défini par $E_u = \{u/y_1 = y_0 + T'u ; \|u\|_\infty \leq 1\}$. Si y_2 est le symétrique de y_1 par rapport à y_0 , alors $y_2 = 2y_0 - y_1$. Le domaine F_u correspondant à y_2 vérifie nécessairement $F_u = \{u/y_1 = 2y_0 - y_2 = y_0 + T'u\} = \{u/y_2 = y_0 - T'u\}$. Ce domaine peut être comparé à E_u ; il apparaît que si y_1 appartient à $\mathcal{A}(y)$ alors, puisque $\|u\|_\infty \leq 1 \Leftrightarrow \|-u\|_\infty \leq 1$, le symétrique y_2 de y_1 par rapport à y_0 appartient aussi à $\mathcal{A}(y)$. Ceci étant vérifié pour tout point de $\mathcal{A}(y)$ et en particulier pour les points appartenant à la frontière, il s'ensuit que y_0 est bien le centre de symétrie du domaine $\mathcal{A}(y)$.

Un corollaire de ce qui vient d'être énoncé est que le polytope $\mathcal{A}(y)$ est un paralléloptope¹⁴ : les hyperplans délimitant ce domaine sont deux à deux parallèles puisque le symétrique d'un segment est un segment parallèle à l'original. Dès lors, nous pouvons concevoir qu'un paralléloptope est généré par l'intersection de domaines délimités par deux hyperplans parallèles¹⁵. Or, ces domaines, ou bandes, étant convexes, le domaine $\mathcal{A}(y)$ l'est aussi puisque l'intersection de domaines convexes est convexe (par exemple [St-Guilhem,91]). Cette propriété s'étend sans difficulté aux polytopes.

Exemple 4

Intéressons-nous à l'espace abstrait engendré par le champ vectoriel suivant :

$$y = y_0 + Tu \text{ avec } T = \begin{bmatrix} 1 & -1 & 2 \\ 1 & 1 & 0 \end{bmatrix} \text{ et } y_0 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$$

Le domaine circonscrit $\square \mathcal{A}(y)$ s'obtient sans difficulté à partir des propriétés sur les variables indépendantes : $\mathcal{A}(y_1) = [-4 \ 4]$ et $\mathcal{A}(y_2) = [-2 \ 2]$.

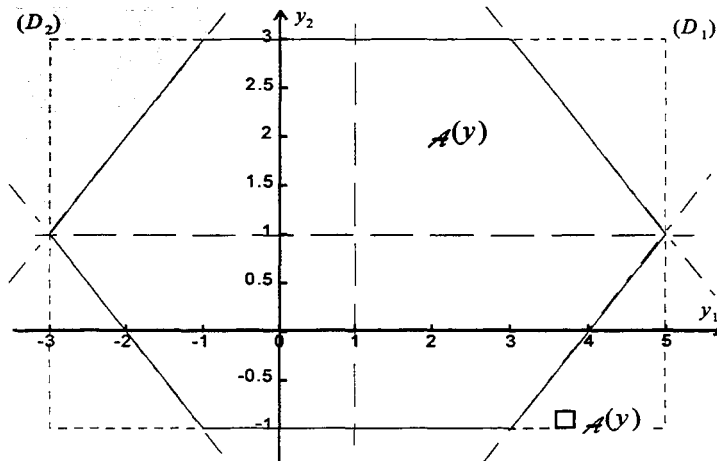


Figure XII- Domaine $\mathcal{A}(y)$ et frontières

Le champ vectoriel ne conduit qu'à un unique sous-ensemble résoluble. Les équations cartésiennes (20) conduisent aux domaines suivants :

¹⁴ L'analogie avec le parallélogramme aurait conduit à restreindre les paralléloptopes aux domaines délimités par 2^d (où d est la dimension de l'espace) hyperplans deux à deux parallèles. Nous avons pris la liberté d'étendre ce concept à des dimensions supérieures pour la seule raison de bénéficier d'une dénomination.

¹⁵ Nous parlerons de bandes.

$$(D_1) - 4 \leq y_2 - y_1 \leq 4 \quad \text{et} \quad (D_2) - 2 \leq y_2 + y_1 \leq 6$$

La convexité du domaine $\mathcal{A}(y)$ apparaît sur la figure XII. Un tel domaine est parfaitement caractérisé par ses sommets qui peuvent être déduits de l'intersection des équations des frontières des domaines $\square\mathcal{A}(y)$, (D_1) et (D_2) .

A l'instar de l'exemple précédent, le calcul des sommets d'un domaine polytopique affine s'opère, d'une manière générale, à partir de l'intersection des frontières de chacun des domaines générateurs de l'espace abstrait $\mathcal{A}(y)$.

Rappelons que les structures des champs vectoriels que nous venons de décrire sont suffisantes pour que les domaines soient des polytopes ou même des parallélotopes. Ainsi, en appliquant une transformation homothétique à un parallélotope, le domaine obtenu sera naturellement polytopique. Le champ vectoriel aura la structure suivante :

$$y = \lambda(v)(t_0 + Tu) : \lambda(v) \in \mathbf{R}, \|u\|_\infty \leq 1 \quad \text{et} \quad \|v\|_\infty \leq 1 \quad (31)$$

si le signe de $\lambda(v)$ ne change pas pour $v \in \{v / \|v\|_\infty \leq 1\}$ alors le domaine $\mathcal{A}(y)$ sera convexe.

Bien que le domaine $\mathcal{A}(y)$ soit polytopique, la structure du champ vectoriel (31) n'est pas conforme à la structure (29) ; le domaine n'est d'ailleurs, a priori, pas dissociable.

IV. Evaluation numérique des domaines abstraits

Les méthodes d'évaluation de domaines abstraits au moyen d'expressions analytiques demeurent casuelles. Même la méthode générale peut conduire à des solutions trop complexes pour être évaluées. Il est donc intéressant pour une application au diagnostic de disposer de méthodes numériques pour l'évaluation de domaines abstraits. Nous en décrirons deux. La première, la plus triviale, consiste à balayer les valeurs de variables abstraites u possibles tandis que la seconde procède par pavage du domaine $\|u\|_\infty \leq 1$.

IV.1. Evaluation d'un domaine par balayage

Le principe de la méthode par balayage est de quadriller le domaine $\|u\|_\infty \leq 1$ des variables abstraites par des points régulièrement répartis. Supposons que chaque variable abstraite u_i soit représentée par $d+1$ points constituant un découpage en d intervalle. Le domaine $\|u\|_\infty \leq 1$ sera donc représenté par $d^{\text{card}(U)}$ valeurs. Plus d devient grand, plus la représentation sera précise. Le domaine $\mathcal{A}(y)$ sera alors obtenu en appliquant successivement la fonction $f(\cdot)$ à chacune des valeurs représentatives du domaine $\|u\|_\infty \leq 1$; il sera représenté par $d^{\text{card}(U)}$ valeurs.

Néanmoins, ce résultat obtenu sous forme d'ensemble de points n'est guère satisfaisant. Comment en effet vérifier si, comme l'implique tout test de cohérence inhérent au diagnostic, une valeur quelconque est cohérente avec un ensemble de point ? Pour éluder cette difficulté, nous proposons de superposer à l'ensemble de points obtenu, une grille de résolution r . Le principe est de décomposer le domaine représentatif obtenu en orthotopes élémentaires

régulièrement distribués. Les longueurs des orthotopes sont calées sur les plus grands écarts observés divisés par $r-1$. La grille est alors calée comme indiqué sur la figure XIII (résolution $r=6$). A chaque fois qu'une valeur appartient à un orthotope élémentaire, il est considéré comme atteint. Une fois que tous les orthotopes élémentaires atteints ont été déterminés en testant l'appartenance de toutes les valeurs représentatives de $\mathcal{A}(y)$, les tests de cohérence deviennent extrêmement simples. Il suffit de tester si cette valeur appartient bien à un orthotope élémentaire atteint.

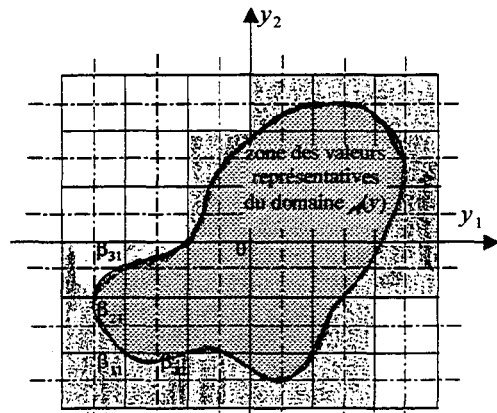


Figure XIII – Décomposition d'un domaine sur une grille de résolution

Il est généralement facile de déduire l'orthotope aligné $\square \mathcal{A}(y)$, permettant ainsi de caler la grille de résolution sans mémoriser toutes les valeurs représentatives de $\mathcal{A}(y)$. La taille mémoire requise est relativement faible puisque seules les coordonnées des centres, au nombre de $r^{dim(y)}$, des orthotopes élémentaires suffit à les caractériser : les dimensions étant les mêmes pour tous.

Le choix de la valeur de la résolution r est, au moins en théorie, problématique. Si la valeur de la résolution r est trop faible, le domaine obtenu sera très imprécis tandis que si la résolution est trop importante, des trous seront obtenus à défaut de suffisamment de valeurs représentatives de $\mathcal{A}(y)$. Il est manifeste que la résolution r est étroitement liée au découpage d et qu'il est difficile d'étudier théoriquement, et dans un cas général, le meilleur compromis. En ce qui nous concerne, nous choisissons toujours $r=d$, ce qui s'est avéré, dans la pratique, être

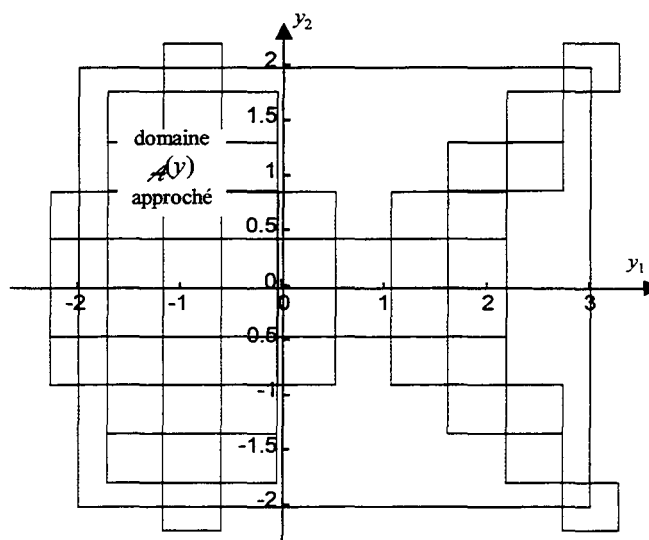


Figure XIV – Domaine $\mathcal{A}(y)$ de l'exemple 1 reconstruit par balayage ($r=d=10$)

un excellent compromis.

exemple

La figure XIV illustre le domaine abstrait de l'exemple 1, représenté sur la figure X. La résolution a été choisie égale au nombre de découpage : $r=d=10$. Les résultats ont été obtenus en moins de 5 dixièmes de secondes avec Matlab sur un Pentium 133MHz.

La méthode de balayage s'applique dans tous les cas. Néanmoins, cette méthode n'exploite pas pleinement le fait que des orthotopes circonscrits peuvent parfois être déduits de l'arithmétique des intervalles. C'est l'objet de la méthode de pavage.

IV.2. Evaluation d'un domaine par pavage

La méthode de pavage que nous allons décrire s'inspire des travaux parus dans [Moore,79] et appliqués dans [Jaulin,96] à des domaines multidimensionnels. Le principe de la méthode est de décomposer le domaine $\|u\|_{\infty} \leq 1$, non plus en un quadrillage de valeurs, mais en un ensemble d'orthotopes alignés ou *pavés*.

La numérotation du découpage est effectuée à l'instar de la figure XV (et par extrapolation à des dimensions supérieures). Chaque pavé est un domaine élémentaire désigné par $\mathcal{P}_{i,j,k,\dots}$ où le $dim(u)$ -uplet (i,j,k) désigne la position du pavé dans le domaine. Pour mettre en équation ce pavé, nous distinguerons le cas d'un découpage pair du cas d'un découpage impair.

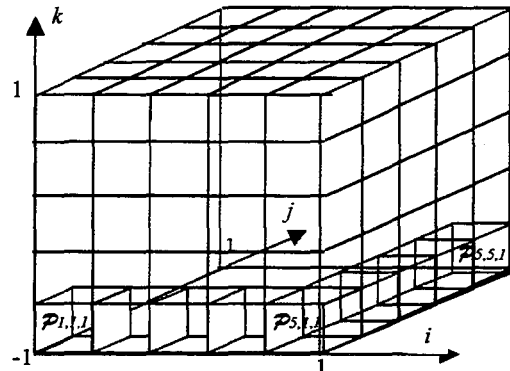


Figure XV – Pavage du domaine $\|u\|_{\infty} \leq 1$ ($d=5$ et $dim(u)=3$)

Cas d'un découpage impair

Le découpage est égal à $d=2\lambda+1$ avec $\lambda \in \mathbb{N}$. Le centre d'un pavé $\mathcal{P}_{i,j,\dots}$ est donné par :

$$c_{i,j,\dots} = \frac{2}{d} \begin{bmatrix} i \\ j \\ \vdots \end{bmatrix} ; (i,j,\dots) \in \{-\lambda, \dots, \lambda\}^{dim(u)}$$

avec $dim(c_{i,j,\dots})=dim(u)$

En utilisant ce résultat, nous en déduisons que le domaine élémentaire de chaque pavé $\mathcal{P}_{i,j,\dots}$ est entièrement défini par :

$$\mathcal{P}_{i,j,\dots} = \left\{ \frac{2}{d} \begin{bmatrix} i \\ j \\ \vdots \end{bmatrix} + \frac{1}{d} u_{i,j,\dots} ; \|u_{i,j,\dots}\|_{\infty} \leq 1 \right\} ; (i,j,\dots) \in \{-\lambda, \dots, \lambda\}^{\dim(u)}$$

avec $\dim(u_{i,j,\dots}) = \dim(u)$

cas d'un découpage pair

Le découpage est égal à $d=2\lambda$ avec $\lambda \in \mathbb{N}$. Le centre d'un pavé $\mathcal{P}_{i,j,\dots}$ est donné par :

$$c_{i,j,\dots} = \frac{2}{d} \begin{bmatrix} i \\ j \\ \vdots \end{bmatrix} - \frac{1}{d} ; (i,j,\dots) \in \{-\lambda+1, \dots, \lambda\}^{\dim(u)}$$

avec $\dim(c_{i,j,\dots}) = \dim(u)$

En utilisant ce résultat, nous en déduisons que le domaine élémentaire de chaque pavé $\mathcal{P}_{i,j,\dots}$ est entièrement défini par :

$$\mathcal{P}_{i,j,\dots} = \left\{ \frac{2}{d} \begin{bmatrix} i \\ j \\ \vdots \end{bmatrix} - \frac{1}{d} + \frac{1}{d} u_{i,j,\dots} ; \|u_{i,j,\dots}\|_{\infty} \leq 1 \right\} ; (i,j,\dots) \in \{-\lambda+1, \dots, \lambda\}^{\dim(u)}$$

avec $\dim(u_{i,j,\dots}) = \dim(u)$

Qu'il soit pair ou impair, le pavage est conçu comme un recouvrement du domaine $\|u\|_{\infty} \leq 1$; autrement dit, $\{u / \|u\|_{\infty} \leq 1\} = \bigcup_{i,j,\dots} \mathcal{P}_{i,j,\dots}$. Afin de simplifier les notations, nous poserons

$F(\mathcal{P}_{i,j,\dots}) = \{f_x(v) : v \in \mathcal{P}_{i,j,\dots}\}$; où le champ vectoriel $f_x(\cdot)$ correspond à celui défini par (16). Ce domaine s'écrit aussi :

$$d \text{ est impair : } F(\mathcal{P}_{i,j,\dots}) = \left\{ f_x \left(\frac{2}{d} \begin{bmatrix} i \\ j \\ \vdots \end{bmatrix} + \frac{1}{d} u_{i,j,\dots} \right) ; \|u_{i,j,\dots}\|_{\infty} \leq 1 \right\}$$

avec $d=2\lambda+1$ et $(i,j,\dots) \in \{-\lambda, \dots, \lambda\}^{\dim(u)}$

$$d \text{ est pair : } \mathcal{P}_{i,j,\dots} = \left\{ f_x \left(\frac{2}{d} \begin{bmatrix} i \\ j \\ \vdots \end{bmatrix} - \frac{1}{d} + \frac{1}{d} u_{i,j,\dots} \right) ; \|u_{i,j,\dots}\|_{\infty} \leq 1 \right\}$$

avec $d=2\lambda$ et $(i,j,\dots) \in \{-\lambda+1, \dots, \lambda\}^{\dim(u)}$

Le domaine $\mathcal{A}(y)$ s'identifie au domaine $\bigcup_{i,j,\dots} F(\mathcal{P}_{i,j,\dots})$ puisque :

$$\{f_x(u) ; \|u\|_{\infty} \leq 1\} = F\left(\bigcup_{i,j,\dots} \mathcal{P}_{i,j,\dots}\right) = \bigcup_{i,j,\dots} F(\mathcal{P}_{i,j,\dots})$$

Il n'est certes guère plus facile d'évaluer les domaines $F(\mathcal{P}_{i,j,\dots})$ que $\mathcal{A}(y)$, cependant, nous

n'évaluerons que l'orthotope circonscrit à chacun des domaines élémentaires. Nous retiendrons la relation importante qui suit :

$$\mathcal{A}(y) \subseteq \bigcup_{i,j,\dots} \square F(p_{i,j,\dots}) \quad (32)$$

Selon toute évidence, plus le découpage sera important, et plus l'inclusion tendra vers l'égalité.

Nous avons appliqué cette méthode à l'exemple 1 (cf. figure X) pour des découpages impairs de 3, 5, 7 et 9 ; nous avons obtenus les résultats de la figure XI en des temps respectifs de 0,05s, 0,07s, 0,11s et 0,22s.

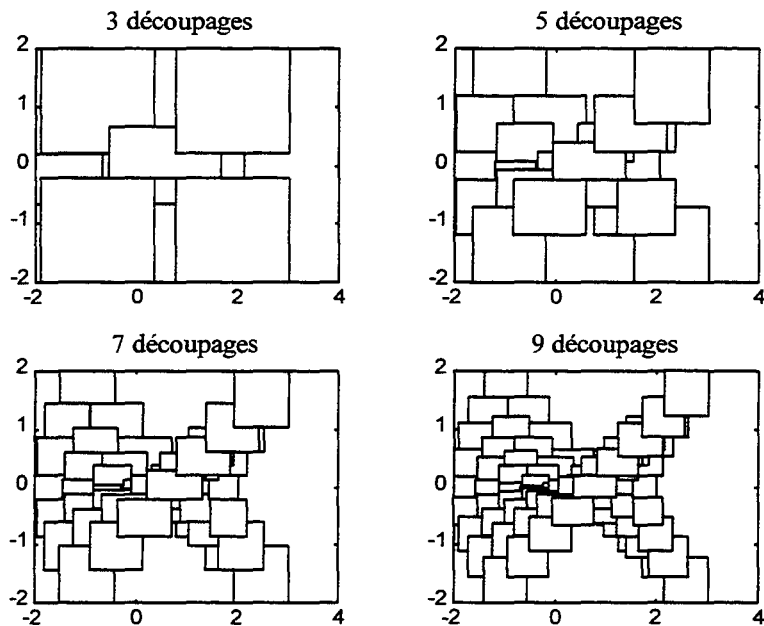


Figure XVI – Evaluation du domaine abstrait de l'exemple 1 par pavage

Cette approche est plus performante que l'approche par balayage puisque 7 découpages en 11 centièmes de seconde conduisent à un domaine comparable à 10 découpages en 5 dixièmes de seconde par la méthode de pavage. Cependant, il faut tout de même préciser que cette méthode ne peut s'appliquer que dans la mesure où il est possible de déduire des orthotopes circonscrits par l'arithmétique des intervalles. Elle est donc restreinte aux champs ne contenant que des opérateurs arithmétiques et aux expressions relativement simples. Signalons aussi qu'aux orthotopes circonscrits aux domaines élémentaires peuvent être substitués des orthotopes majorant raisonnablement les pavés : la méthode s'applique toujours avec une perte de précision.

V. Conclusion

Ce chapitre est une introduction à la prise en compte d'incertitudes de modélisation en

diagnostic. Nous avons mis en exergue les limites des approches stochastiques pour introduire les principaux outils théoriques permettant d'aborder les incertitudes bornées. Nous avons établi certaines propriétés qui nous semblent utiles à une application au diagnostic. Enfin, nous avons décrit des méthodes permettant d'évaluer des domaines abstraits, que ce soit par des expressions analytiques ou par des approches purement numériques.

L'intérêt, pour le diagnostic, de ces notions est indéniable ; dans un cas certain, un test de cohérence se résumait à un test de nullité d'une valeur alors que dans un cas incertain, le test de cohérence consiste à vérifier si une valeur appartient ou non à un domaine abstrait représentant un ensemble de comportements possibles. Toutes les méthodes que nous avons détaillées, ont été pensées pour le diagnostic, c'est-à-dire que toutes permettent d'effectuer des tests de cohérence.

Chapitre 3

les métamodèles

TABLE DES MATIÈRES

I.	Introduction	97
II.	Structures particulières de métamodèles incertains.....	98
	Les métamodèles statiques	98
	Les métamodèles dynamiques.....	99
II.1.	Structures additives des incertitudes.....	100
II.2.	Structures multiplicatives des incertitudes.....	102
II.3.	Invariance et dépendance temporelle.....	105
II.4.	Influence de la structure du métamodèle incertain.....	108
III.	Caractérisation des variables abstraites bornées.....	111
III.1.	Exemple introductif	113
III.2.	Cas général sans connaissance a priori des incertitudes	115
	Notations et définitions	117
	Calcul itératif des domaines de majoration.....	119
	Caractérisation des incertitudes	121
III.3.	Exemple physique.....	123
III.4.	Quelques extensions.....	126
	Caractérisation dans le cas d'une connaissance a priori.....	126
	Cas des modèles à horizon infini.....	128
	Cas multivariable.....	130
IV.	Conclusion	130

I. Introduction

Le chapitre précédent nous a permis de mettre en évidence les principales propriétés des variables abstraites bornées. Nous avons défini, dans le premier chapitre, ces variables comme éléments de modélisation intervenant au sein d'un métamodèle et, en ce sens, elles n'existent que dans la mesure où elles interviennent dans un métamodèle. Le chapitre précédent ne suffit donc pas à justifier de l'intérêt des variables bornées s'il n'existe aucun moyen d'obtenir, de la physique, des métamodèles en comprenant.

Ce chapitre n'est pas, comme pourrait le laisser entendre son titre, consacré aux métamodèles en général, qui ont bien plus à voir avec l'étude des systèmes physiques qu'avec les fondements de la modélisation des incertitudes. Notre objectif se restreindra à étudier les différents moyens d'obtenir des métamodèles contenant des variables abstraites bornées.

Le premier type de situation est celui où une connaissance a priori des incertitudes est disponible, qu'elle soit issue des lois de la physique ou de l'expertise. De telles situations engendrent directement des métamodèles incertains et, par conséquent, nous ne les aborderons pas dans ce chapitre.

La seconde situation pourrait être baptisée, de manière quelque peu imagée, de *sculpture d'un métamodèle*. Elle concerne les métamodèles dont la partie certaine a été modélisée et identifiée par des méthodes désormais traditionnelles et dont on souhaite se servir pour créer des enveloppes représentatives d'incertitudes de modélisation. Il est possible de créer plusieurs types d'enveloppes, mais cela suppose une certaine connaissance de structures de métamodèle particulières et des enveloppes qu'elles engendrent. C'est l'objet de la première partie de ce chapitre. Nous montrerons dans cette partie que toute structure, certaine ou non, peut être représentée de façon juste par des structures simples de métamodèle, mais au détriment certes de la précision.

La dernière situation, correspondant à la seconde examinée dans ce chapitre, s'inscrit dans le cadre des travaux les plus récents en matière de variables bornées. En complément de l'identification qui permet de déterminer la valeur des paramètres invariants, nous avons étudié la caractérisation dont l'objectif est de déterminer les bornes des variables abstraites dont les valeurs dépendent du temps. Après avoir formalisé les objectifs de la caractérisation des variables abstraites, nous proposerons un algorithme de caractérisation des variables abstraites bornées. Nous montrerons, dans le cas des systèmes MISO¹ affines en les variables abstraites, comment réduire le problème de caractérisation à un nombre fini de problèmes élémentaires.

¹ multi-entrées et une sortie

II. Structures particulières de métamodèles incertains

L'objectif de cette partie est de trouver des structures de métamodèles incertains, suffisamment simples pour qu'il soit possible de leur associer directement des enveloppes ; les structures plus complexes feront l'objet du chapitre suivant. Nous escomptons, grâce à des structures simples d'incertitudes, pouvoir, après identification de la partie certaine, ajuster les enveloppes déduites en modifiant les bornes des variables abstraites de façon à expliquer toutes les valeurs d'une séquence de mesures. En suivant ce type d'approche, nous pourrons, dans la mesure où la séquence de mesures est suffisamment représentative d'un ensemble de comportements possibles, satisfaire à la condition de justesse, indispensable au diagnostic. Par contre, il est indéniable que ce type d'approche n'offre aucune garantie quant à la précision de la représentation obtenue.

Les métamodèles statiques

Précisons que nous entendons par métamodèle statique un ensemble de relations faisant intervenir des variables physiques observées toutes à un même instant. Il existe deux représentations possibles de métamodèles statiques, selon que les incertitudes sont représentées par des variables abstraites bornées invariantes ou dépendant du temps. Ces deux types de représentations sont donnés, après que les champs de vecteurs aient été normalisés, par :

$$\begin{aligned} (\Sigma_s) \quad y(t) &= f(u(t), v) \quad : \mathcal{A}(v) = \mathcal{U} \\ (\Sigma'_s) \quad z(t) &= f(u(t), v(t)) : \forall t, \mathcal{A}(v(t)) = \mathcal{U} \end{aligned}$$

*où $u(t)$ est un vecteur de variables déterministes,
 $v(t)$ et v des vecteurs de variables bornées, $y(t)$ et $z(t)$ des vecteurs²*

Un métamodèle statique possède a priori un horizon nul. Cependant, il est toujours possible de sommer de telles relations à différents instants pour obtenir des relations couvrant un horizon h plus important ; par exemple

$$\begin{aligned} (\Sigma_s^{(h)}) \quad \sum_{i=0}^h y(t_i) - f(u(t_i), v) &: \mathcal{A}(v) = \mathcal{U} \\ (\Sigma'_s{}^{(h)}) \quad \sum_{i=0}^h z(t_i) - f(u(t_i), v(t_i)) &: \forall t_i, \mathcal{A}(v(t_i)) = \mathcal{U} \end{aligned}$$

Cependant, notre objectif étant la *sculpture de métamodèle*, il est naturel de rechercher les structures les plus simples possibles afin de maîtriser l'influence sur la représentation d'une modification de bornes des variables abstraites. Nous préférons donc les représentations d'horizon 0 ; les métamodèles d'horizons plus importants seront toutefois abordés dans la

² Les crochets autour des variables physiques ont été omis pour alléger les écritures.

partie II.3.

Dans le cas de métamodèles statiques (Σ_s) et (Σ'_s) d'horizon 0, les espaces abstraits $\mathcal{A}(y(t))$ et $\mathcal{A}(z(t))$, qui peuvent se déduire des métamodèles, sont confondus. En effet, puisque $\mathcal{A}(v) = \mathcal{A}(v(t)) = \mathcal{U}$ ³, les espaces abstraits vérifient $\mathcal{A}(y(t)) = \mathcal{A}(z(t)) = \{f(u(t), \mu) ; \mu \in \mathcal{U}\}$; autrement dit, les systèmes (Σ_s) et (Σ'_s) conduisent à une même enveloppe, en se gardant toutefois de conclure que, pour tout t , $y(t) = z(t)$. Ceci s'étend aisément aux vecteurs de variables bornées mixtes : certaines étant invariantes d'autres pas.

Ce résultat est extrêmement intéressant car, que les variables abstraites soient invariantes ou pas, lorsque le métamodèle statique est à horizon 0, tout se passe, au niveau des enveloppes, comme si les variables abstraites étaient invariantes. Par conséquent, seule l'étude du métamodèle (Σ_s) suffit.

Les métamodèles dynamiques

Les systèmes d'équations différentielles ordinaires représentant d'une manière générale les systèmes dynamiques⁴ incertains s'écrivent, selon que les variables abstraites sont invariantes ou pas, comme suit :

$$\begin{aligned} (\Sigma_d) \frac{dx(t)}{dt} &= f(x(t), u(t), v) : x(0) = x_0, \mathcal{A}(v) = \mathcal{U} \\ (\Sigma'_d) \frac{dz(t)}{dt} &= f(z(t), u(t), v(t)) : z'(0) = x_0, \forall t \mathcal{A}(v(t)) = \mathcal{U} \end{aligned}$$

où $u(t)$ est un vecteur de variables déterministes,
 $v(t)$ et v des vecteurs de variables bornées, $x(t)$ et $z(t)$ des vecteurs

Dans cette partie, nous n'envisagerons que le cas de métamodèles dynamiques en temps continu car il est, d'une part, problématique d'obtenir, dans le cas où le champ vectoriel $f(\dots)$ est non linéaire, une expression en temps discret correspondante et d'autre part, dans un cas linéaire, les métamodèles incertains en temps discrets se traitent de la même manière qu'en temps continu.

Contrairement aux métamodèles statiques, les métamodèles dynamiques sont, lorsqu'ils sont intégrés, à horizon infini. Par conséquent, il n'est plus possible de conclure à l'équivalence des enveloppes issues de (Σ_d) et de (Σ'_d) . Il est évident que les enveloppes issues de (Σ'_d) contiennent nécessairement les enveloppes déduites de (Σ_d) puisqu'une incertitude invariante est un cas particulier d'incertitude fluctuant au cours du temps.

La détermination d'enveloppes lorsque les incertitudes fluctuent au cours du temps est complexe (cf. chapitre 4) et se trouve ainsi mal adaptée à nos objectifs. Nous préférons ici le

³ Toutes les variables $v_x(t)$ et v_x désigneront désormais des variables abstraites normalisées invariantes ou non.

⁴ Contrairement au cas statique, le passé d'un système dynamique intervient dans son comportement présent.

métamodèle (Σ_d) qui peut, même si sa structure ne le fait pas apparaître, expliquer avec justesse des incertitudes dépendant du temps même s'il faut regretter une perte de précision.

Différentes structures de métamodèles peuvent dorénavant être étudiées. Les métamodèles que nous envisagerons contiendront exclusivement des variables abstraites bornées invariantes.

II.1. Structures additives des incertitudes

En partant du modèle statique (Σ_s), une structure additive d'incertitude se traduit par :

$$(\Sigma_{sa}) \quad y(t) = f(u(t)) + l(v) \quad : \quad \|v\|_\infty \leq 1 ; l(0) = 0$$

où v est un vecteur de variables abstraites normalisées.
La propriété $l(0)=0$ permet d'obtenir des enveloppes centrées.

En se fondant sur les développements théoriques du chapitre précédent, nous reconnaissons sans équivoque l'écriture d'un domaine abstrait centré en $f(u(t))$ et dont la forme est indépendante de $u(t)$. Dans le cas particulier où $l(v)=Lv$, le domaine est un paralléloépe (cf. chapitre 2 (30)) parfaitement défini par $\{Lv : \|v\|_\infty \leq 1\}$. En conséquence, une incertitude additive dans un système statique conduit toujours à une enveloppe régulière.

Dans le cas simple où $y(t)$ est un scalaire et $l(v)=\lambda v : \mathcal{A}(v) = [-1,1]$ ($\lambda \in \mathbb{R}^{*+}$), l'enveloppe obtenue est bien régulière car, pour tout $u(t)$ donné, l'espace abstrait de $y(t)$ s'identifie à $\mathcal{Y} = [f(u(t)) - \lambda, f(u(t)) + \lambda]$, pour lequel le rayon $\rho(\mathcal{Y}) = \lambda$ est constant. La figure I représente l'enveloppe obtenue pour un exemple particulier.

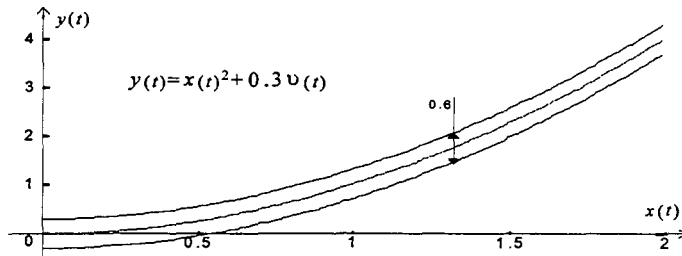


Figure I- Enveloppe obtenue avec une incertitude additive, modèle statique

Une structure d'incertitude additive dans l'équation dynamique (Σ_d) conduit, dans le cas général, à l'écriture suivante :

$$\frac{dx(t)}{dt} = f(x(t), u(t)) + l(v) : x(0) = x_0, \|v\|_\infty \leq 1 ; l(0) = 0$$

Or, nous ne savons pas étudier d'une manière générale l'enveloppe d'un tel système. Considérons, dans un premier temps, le cas où ce système s'écrit :

$$(\Sigma_{da}) \frac{dx(t)}{dt} = f(x(t)) + g(u(t)) + l(v); x(0) = x_0, \|v\|_\infty \leq 1; l(0) = 0$$

où f est un homomorphisme de groupe sur \mathbb{R}^n ($\dim(x(t))=n$) : $\forall (x,y) \in (\mathbb{R}^n)^2, f(x+y)=f(x)+f(y)$

En introduisant le système certain suivant :

$$(\Sigma_{dc}) \frac{dx_\alpha(t)}{dt} = f(x_\alpha(t)) + g(u(t))$$

et en posant $x_\alpha(t) = x_\beta(t) + m(v)$, nous obtenons par substitution dans (Σ_{dc}) :

$$\frac{dx_\beta(t)}{dt} = f(x_\beta(t)) + g(u(t)) + f(m(v))$$

Si $f(m(v))=l(v)$ et $x_\beta(0) = x_0$, alors la solution $x_\beta(t)$ de cette équation est identique à la solution $x(t)$ de (Σ_{da}) . Ainsi $x_\beta(t) = x_\alpha(t) - m(v) = x(t)$ et nous en déduisons que la structure additive représente une enveloppe régulière centrée sur $x_\alpha(t)$ de (Σ_{dc}) et dont la section est un domaine abstrait invariant défini par $\{m(v) : \|v\|_\infty \leq 1\}$.

Les systèmes dynamiques linéaires $f(x(t))=Ax(t)$, $g(u(t))=Bu(t)$, sont représentés par le système (Σ_{da}) . La section de l'enveloppe est un domaine abstrait défini par $\{m(v) : \|v\|_\infty \leq 1\}$ où $m(v)$ est tel que $Am(v)=l(v)$ c'est-à-dire que si la matrice A est inversible $m(v)=A^{-1}l(v)$. Si $l(v)=Lv$, le domaine abstrait obtenu est un paralléloétope puisque $m(v)=A^{-1}Lv : \|v\|_\infty \leq 1$ ou en posant $M=A^{-1}L$, $m(v)=Mv$.

D'autre part, si le champ vectoriel $f(\cdot)$ n'est pas un homomorphisme de groupe sur \mathbb{R}^n et si $l(v)=Lv$, nous pouvons écrire que :

$$f(\ddot{x}(t) + Mv) \xrightarrow{v \rightarrow 0} f(\ddot{x}(t)) + \left. \frac{\partial f(Mv)}{\partial v} \right|_{v=0} v$$

avec $M=A^{-1}L$

Autrement dit, en posant $L = \left. \frac{\partial f(Mv)}{\partial v} \right|_{v=0}$, nous concluons que l'enveloppe est, en première approximation, régulière et que sa section est proche d'un paralléloétope défini par $\{Mv : \|v\|_\infty \leq 1\}$.

La structure additive est la plus simple des structures ; de ce fait, elle est très largement utilisée, notamment par les approches probabilistes mais elle ne conduit qu'à des enveloppes régulières ou à peu près régulières. La figure II illustre l'enveloppe obtenue dans le cas d'un système dynamique de dimension 1.

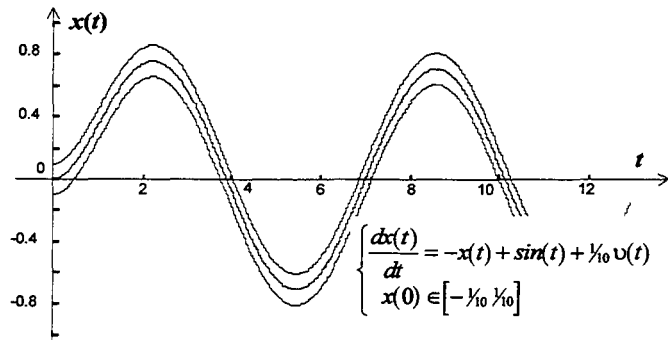


Figure II- Enveloppe obtenue avec une incertitude additive, modèle dynamique

Dans le cas où le métamodèle contient en plus d'une équation différentielle, une équation de sortie du type :

$$y(t) = Cx(t) + Du(t) + Nv$$

Le domaine $\mathcal{A}(y(t))$ est centré sur la valeur obtenue dans le cas où $v=0$ et conduit à une enveloppe régulière. En effet, puisque $x(t) = x_a(t) - Mv$, nous obtenons : $y(t) = Cx_a(t) + Du(t) + (N-M)v$. Le domaine abstrait est toujours un paralléloétope.

II.2. Structures multiplicatives des incertitudes

Malgré la simplicité de la structure additive, il est de nombreux cas où elle n'est pas adaptée. Ainsi, dans la grande majorité des situations pratiques, les incertitudes de mesure sont données en valeur relative. Si la variable $\tilde{y}(t)$ représente la mesure de la variable physique $[y(t)]$, la borne e_r de l'erreur relative sur la mesure se traduira par :

$$\frac{\tilde{y}(t) - [y(t)]}{[y(t)]} = e_r u(t) \Leftrightarrow \tilde{y}(t) = (1 + e_r u(t)) [y(t)]$$

avec $u(t)$ une variable abstraite normalisée : $\mathcal{A}(u(t)) = [-1, 1]$ et $e_r \in \mathbb{R}^+$

La structure de l'incertitude est ici multiplicative et représente le fait que plus la valeur mesurée est importante, plus l'erreur commise l'est. Une structure additive ne permettrait pas de représenter ce phénomène ; nous aurions $\tilde{y}(t) = [y(t)] + e_a u(t)$ et, pour conserver l'impératif de justesse, il faudrait que $e_a = e_r \sup_t |[y(t)]|$, ce qui correspondrait à une incertitude régulière égale à la plus grande incertitude sur la mesure. Le résultat ne pourra en être évidemment que moins précis.

D'une manière générale, un système statique inspiré de (Σ_s) sera pourvu d'une incertitude multiplicative s'il peut se mettre sous la forme suivante :

$$(\Sigma_{sm})y(t) = l(v)f(u(t)) : \|v\|_\infty \leq 1; l(0) = I_n$$

où v est un vecteur de variables abstraites normalisées et $\dim(y(t)) = \dim(f(u(t))) = n$

Cette structure reste cependant très générale à tel point qu'il est difficile d'y associer un type

spécifique de comportement incertain. Pour éluder cette difficulté, nous nous restreindrons à une structure particulière de la matrice $l(v)$. Nous la supposons carrée diagonale de dimension $n \times n$ et vérifiant :

$$\forall (i, j) \in \{1, \dots, n\}^2, l(v) / l_{ii}(0) = 1, \forall v \in \mathcal{A}(v), l_{ii}(v) \neq 0 \text{ et si } i \neq j, l_{ij}(v) = 0 \quad (I)$$

Lorsque $l(v)$ vérifie la relation (I), le système (Σ_{sm}) s'écrit :

$$y_i(t) = l_{ii}(v)y(t)|_{v=0} : i \in \{1, \dots, n\}$$

En étudiant séparément les équations du domaine abstrait $\mathcal{A}(y(t))$, nous déterminerons les caractéristiques de l'orthotope aligné circonscrit $\square \mathcal{A}(y(t))$. En utilisant la relation (9a) du chapitre 2, nous montrons que $\forall i, \mathcal{A}(y_i(t)) = \left| (y_i(t)|_{v=0}) \right| \mathcal{A}(\sigma_{(y_i(t)|_{v=0})} l_{ii}(v))$ avec $\sigma_{(y_i(t)|_{v=0})} = \text{signe}(y_i(t)|_{v=0})$ et, en ne considérant que les rayons des espaces abstraits, il vient : $\rho(y_i(t)) = \left| (y_i(t)|_{v=0}) \right| \rho(l_{ii}(v))$. Puisque $\rho(l_{ii}(v))$ est constant, seule la valeur $\left| (y_i(t)|_{v=0}) \right|$ varie au cours du temps, c'est-à-dire que la dimension de l'orthotope circonscrit $\square \mathcal{A}(y(t))$, et a fortiori du domaine $\mathcal{A}(y(t))$, varie avec l'amplitude de $y(t)|_{v=0}$. Nous en concluons que l'enveloppe de $y(t)$ est proportionnelle à l'amplitude de $y(t)|_{v=0}$.

Dans le cas particulier où $l(v) = \text{diag}(Lv)$, l'égalité suivante est vérifiée :

$$y(t) = \text{diag}(Lv)(y(t)|_{v=0}) = \text{diag}(y(t)|_{v=0})Lv$$

Il en découle que, outre la proportionnalité, et conformément aux résultats théoriques, une section de l'enveloppe de $y(t)$ est un parallélotope mais, contrairement à la structure additive, sa forme varie au cours du temps.

La figure III représente l'enveloppe obtenue dans un cas de dimension 1.

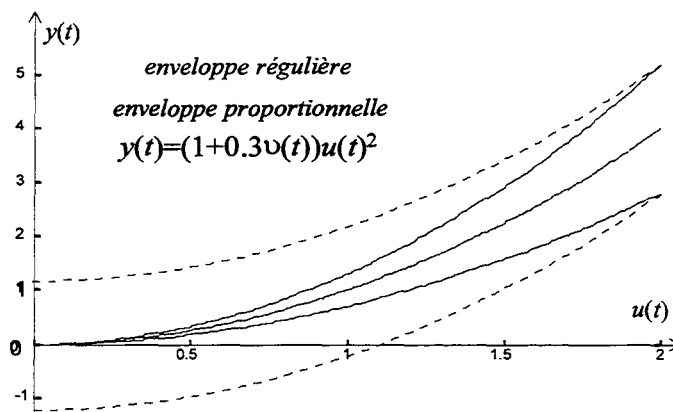


Figure III- Enveloppe obtenue avec des structures multiplicative et additive correspondantes

Il n'est pas simple de définir la structure multiplicative dans le cas dynamique. Nous proposons de la définir comme la structure conduisant, à l'instar du cas statique, à une

enveloppe proportionnelle. Partons du modèle dynamique général certain suivant :

$$(\Sigma_{dc}) \frac{dx_{\alpha}(t)}{dt} = f(x_{\alpha}(t), u(t))$$

où $\dim(x_{\alpha}(t))=n$

Pour obtenir une enveloppe proportionnelle sur une variable $x(t)$, il faut qu'elle vérifie : $x(t)=l(v)x_{\alpha}(t)$ avec $\|v\|_{\infty} \leq 1$ et $l(v)$ vérifie (1). La structure conduisant à une enveloppe proportionnelle s'écrit :

$$(\Sigma_d) \frac{dx(t)}{dt} = l(v)f(l^{-1}(v)x(t), u(t))$$

La structure des incertitudes d'un tel modèle étant complexe, nous n'utiliserons pas le modèle général (Σ_d) pour mettre en évidence une structure simple conduisant à des enveloppes proportionnelles. Nous nous limiterons aux matrices $l(v)=\text{diag}(Lv)$ et à un modèle affine en l'état qui reste somme toute très général :

$$(\Sigma'_{dc}) \frac{dx(t)}{dt} = A(u(t))x(t) + B(u(t))$$

Dans ce cas, des enveloppes proportionnelles seront obtenues avec la structure suivante :

$$(\Sigma'_d) \frac{dx(t)}{dt} = l(v)A(u(t))l^{-1}(v)x(t) + l(v)B(u(t)) : l(v) = I_n + \text{diag}(Lv) \quad (2)$$

Cette structure demeure cependant complexe et ne se simplifie que dans le cas monovariable ($\dim(x(t))=1$). Le modèle (2) devient :

$$(\Sigma'_d) \frac{dx(t)}{dt} = A(u(t))x(t) + l(v)B(u(t)) : l(v) = 1 + Lv$$

La figure IV montre l'enveloppe obtenue pour un modèle dynamique de dimension 1 et la compare avec une enveloppe proportionnelle exacte. Les enveloppes coïncident mis à part au voisinage de zéro. Ceci s'explique par le pas d'intégration constant de 50ms durant lequel nous avons supposé l'entrée constante. Il existe bien entendu de nombreuses autres structures d'incertitudes ; néanmoins, nous avons pris le parti de ne présenter que les structures dont il est possible d'identifier clairement les comportements représentés.

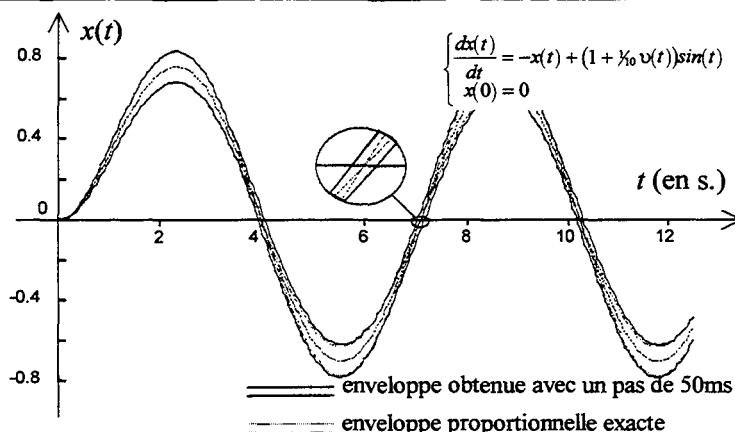


Figure IV- Enveloppe obtenue avec une structure multiplicative, modèle dynamique

Les deux structures d'incertitudes additive et multiplicative que nous venons d'étudier, conduisent respectivement à des enveloppes régulières et proportionnelles. La *sculpture d'un métamodèle* peut s'envisager ainsi. Une phase d'identification traditionnelle conduit à une estimation de la partie certaine du métamodèle. Alors, en choisissant un échantillon de mesures représentatif d'un ensemble de comportements possibles, les observations et les valeurs reconstruites peuvent être comparées. Il est alors facile de déterminer s'il faut préférer une enveloppe régulière ou proportionnelle. En relevant les plus grandes valeurs d'erreurs atteintes, une structure des incertitudes se déduit directement.

II.3. Invariance et dépendance temporelle

Notre intention est ici de montrer en quoi le choix d'une variable abstraite invariante ou dépendant du temps modifie la représentation du comportement d'un système physique. Certes, en introduction de cette partie, nous avons démontré que les enveloppes associées à des horizons nuls ne dépendaient pas de la nature temporelle des variables abstraites, il ne faut pourtant pas en conclure qu'il n'existe aucune différence. Une variable abstraite v représente un phénomène méconnu dès l'origine et invariant dans le temps. La valeur que prend une variable abstraite v appartient à un domaine $\mathcal{A}(v)$ et ne fluctue pas au cours du temps tandis que, dans le cas d'une variable abstraite $v(t)$, la valeur de $v(t)$ appartient à un domaine $\mathcal{A}(v(t))$ et, est a priori différente d'un instant à l'autre. Avec un formalisme mathématique, nous pouvons écrire :

$$\forall (t, t') \in \mathbb{R}^{+2} : t \neq t', v = v(t) = v(t') \in \mathcal{A}(v)$$

$$\text{et } \mathcal{A}\left(\begin{bmatrix} v(t) \\ v(t') \end{bmatrix}\right) = \square \mathcal{A}\left(\begin{bmatrix} v(t) \\ v(t') \end{bmatrix}\right)$$

Deux fonctions du temps déterministes $f(t)$ et $g(t)$, doivent être considérées comme dépendantes car il est généralement possible de trouver une fonction $h(\cdot)$ telle qu'il soit possible d'écrire $f(t)=h(g(t))$ ou $g(t)=h(f(t))$. Or, ceci n'est pas vrai dans le cas de deux

variables abstraites non déterministes indépendantes. En effet, la définition de deux variables abstraites $v(t)$ et $v'(t)$, dépendant du temps mais indépendantes l'une de l'autre (cf. propriété d'indépendance, chapitre 2), implique qu'il n'existe aucune fonction pouvant relier ces variables abstraites. Malgré la ressemblance des écritures, il faut se souvenir que les fonctions du temps traditionnelles sont propres au déterminisme alors que les variables abstraites dépendantes du temps ont des propriétés spécifiques.

Des variables abstraites contenant à la fois une partie incertaine invariante et une partie fluctuant au cours du temps peuvent naturellement être représentées. Ces variables que nous qualifierons de mixtes peuvent être définies, dans un cas normalisé, par : $\eta(t) = (1 - \ell)v + \ell v(t)$ ($0 \leq \ell \leq 1$) avec $\mathcal{A}(\eta(t)) = \mathcal{A}(v) = \mathcal{A}(v(t)) = [-1 \ 1]$.

Etudions enfin le cas fréquent en physique des variables abstraites dont les variations sont bornées. Soit une variable abstraite $\theta(t)$ dont on sait qu'elle appartient nécessairement à l'intervalle $[\tilde{\theta} \ \hat{\theta}]$ et que $|\theta(t + \tau) - \theta(t)| \leq \delta$ ($\tau \in \mathbb{R}^+$). Cette inégalité s'écrit plus commodément de la façon suivante : $\theta(t + \tau) = \theta(t) + \delta v(t)$ et, en cumulant sur plusieurs instants, il vient : $\theta(t + k\tau) = \theta(t) + \delta \sum_{i=0}^{k-1} v(t + i\tau)$. Dès lors, il apparaît que si $\mathcal{A}(\theta(t)) \in [\tilde{\theta}_i \ \hat{\theta}_i]$, alors $\theta(t + k\tau) \in [\tilde{\theta}_i - k\delta \ \hat{\theta}_i + k\delta] \cap [\tilde{\theta} \ \hat{\theta}]$. Il est donc évident que, pour une certaine valeur k^* de k , $\mathcal{A}(\theta(t + k\tau)) = [\tilde{\theta} \ \hat{\theta}] : k \geq k^*$. Il s'ensuit que, passé un certain laps de temps $k^*\tau$, la variable $\theta(t)$ ne pourra plus être définie que comme une variable dépendant du temps et dont l'espace abstrait est donné par $\mathcal{A}(\theta(t)) = [\tilde{\theta} \ \hat{\theta}]$. La contrainte $|\theta(t + \tau) - \theta(t)| \leq \delta$ ne peut plus être prise en compte par les variables abstraites que nous avons définies. Pour tester la cohérence avec une telle contrainte, il faudrait pouvoir définir une procédure d'estimation de $\theta(t)$ et vérifier a posteriori si les estimations satisfont la contrainte. Outre le fait qu'il soit exceptionnel de parvenir à réaliser une telle procédure, ce problème ne présente qu'un intérêt mineur eu égard aux objectifs que nous poursuivons.

Montrons au moyen d'un exemple qui nous est désormais familier comment une procédure de diagnostic peut tester si la nature temporelle d'une variable abstraite est bien cohérente avec sa définition. Le modèle est celui qui nous a servi à tracer la figure II :

$$\frac{dy(t)}{dt} = -y(t) + \sin(t) + \frac{1}{10}\eta(t) \quad \mathcal{A}(\eta(t)) = [-1 \ 1] \quad y(0) = 0$$

Examinons alternativement le cas où la variable abstraite est invariante $\eta(t) = v$ et celui où elle ne l'est pas : $\eta(t) = v(t)$. Sur la figure V, nous avons tracé les réponses obtenues pour $v = -0,5$ et pour $v(t)$ suivant un profil représenté sur cette même figure. Dans le cas général, les enveloppes obtenues seraient différentes ; or, nous constatons ici que l'enveloppe correspondant à une variable abstraite invariante est confondue avec l'enveloppe issue du

métamodèle où la variable abstraite dépend du temps. Nous ne nous pencherons sur cette coïncidence qu'au cours du chapitre suivant qui est dédié aux hypomodèles, et donc particulièrement à la déduction d'enveloppes. Notre seul objectif se restreint ici à mettre en évidence certaines implications de la nature des variables abstraites.

Comment, dans ce cas particulier, traduire que, lorsque la variable abstraite est invariante, la réponse est identique, au décalage près, à la réponse médiane (pour laquelle $v=0$), tandis qu'en présence de variables dépendant du temps, les réponses obtenues peuvent être quelconques, pourvu qu'elles soient contenues par l'enveloppe ?

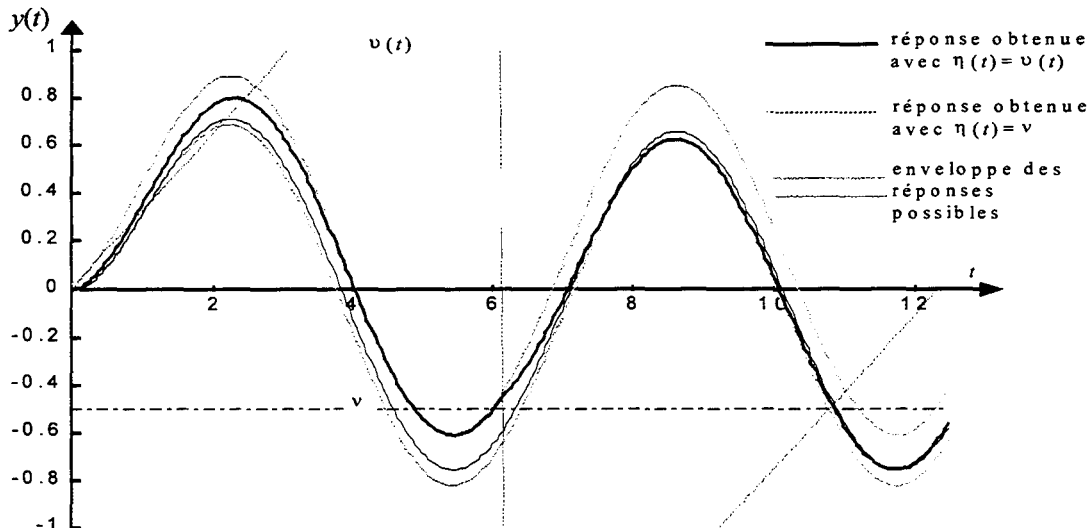


Figure V- Réponses obtenues pour un cas de variables abstraites invariantes et un cas où elles dépendent du temps

Un hypomodèle d'horizon nul ne permet jamais de distinguer l'influence de variables de nature temporelle différente et, dans cet exemple, un hypomodèle à horizon infini ne le permet pas non plus. Examinons si des hypomodèles d'horizons différents peuvent mettre en évidence ces deux types de comportement.

En choisissant un nombre réel positif suffisamment faible $\tau=50 \cdot 10^{-3}$ ms., nous pouvons intégrer avec une bonne précision l'équation différentielle par une méthode de Runge-Kutta d'ordre 2 (nous reviendrons sur ce résultat au cours du chapitre 4) :

$$\begin{cases} x_{k+1} = ax_k + bu_k + e\eta_k \\ y_k = x_k + du_k \end{cases}$$

avec $a=0,9512$ $b=0,0476$ $d=0,0246$ et $e=0,0048$
 et $y(k\tau)=y_k$, $\eta(k\tau)=\eta_k$ et $u(k\tau)=u_k$; $\forall k \in \mathbb{N}$

De ce métamodèle en temps discret, un hypomodèle d'horizon 2 peut facilement être extrait :

$$\frac{y_{k+1} - ay_k - du_{k+1} - (b - ad)u_k}{e} = \eta_k \in [-1, 1]$$

Or, avec cet hypomodèle que la variable abstraite soit invariante ($\eta_k = v$) ou non ($\eta_k = v(k\tau)$), le test de cohérence conduit au même résultat (cf. figure VI) : dans un cas comme dans l'autre, les résidus restent compris entre -1 et 1 . Dans le cas où la variable abstraite η_k est invariante et où le test de cohérence doit vérifier la nature temporelle de cette variable, il est possible d'utiliser un hypomodèle d'horizon au moins égal à 3. En écrivant que si la variable η_k est invariante, alors $\eta_k = \eta_{k+1}$, et nous déduisons en utilisant l'hypomodèle d'horizon 2 :

$$y_{k+2} - (1+a)y_{k+1} + ay_k - du_{k+2} + (d(1+a) - b)u_{k+1} + (b - ad)u_k = 0$$

Les résidus obtenus avec cet hypomodèle d'horizon 3 apparaissent sur la figure VI. Nous constatons que dans le cas d'une variable abstraite invariante, le résidu se maintient à zéro, tandis que lorsque la variable abstraite η_k fluctue, le résidu devient non nul et révèle ainsi une incohérence de comportement.

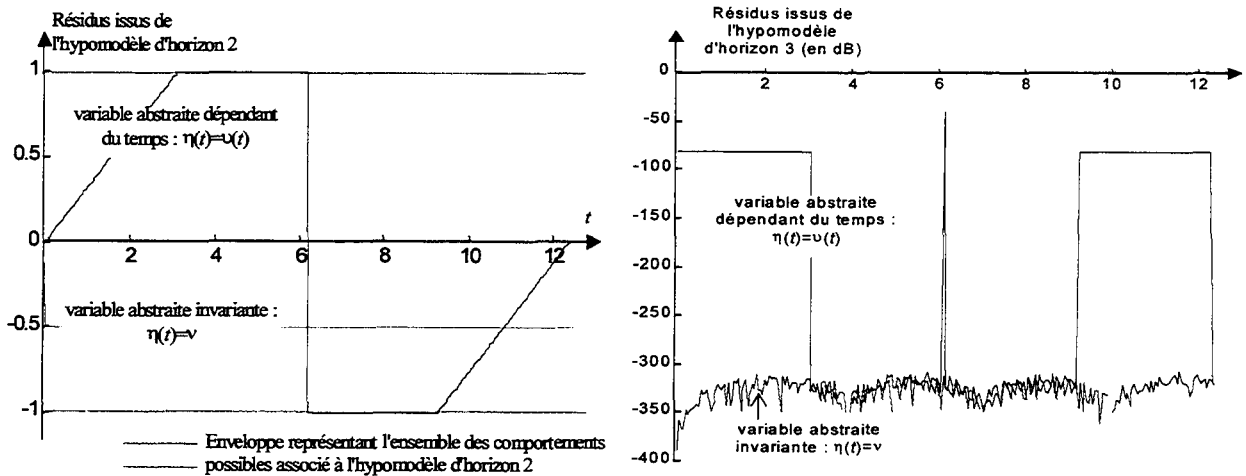


Figure VI- Résidus obtenus dans le cas d'invariance et de dépendance temporelle de la variable abstraite

Cet exemple simple nous a permis de montrer l'importance de la nature temporelle des variables abstraites. Les tests de cohérence sont différents. Il est évident que le test de cohérence valide pour des variables abstraites dépendant du temps l'est aussi lorsque ces variables sont considérées comme invariantes, mais les tests de cohérence resteront incomplets. Par contre, le contraire n'est pas vrai : un test valide pour des variables invariantes ne l'est a priori plus lorsqu'elles dépendent du temps.

II.4. Influence de la structure du métamodèle incertain

Une difficulté fréquente, inhérente à la grande majorité des applications, est la simplification du métamodèle associé à un système physique. Deux raisons peuvent expliquer cette simplification ; soit il existe un métamodèle déterministe trop complexe pour être exploité par une procédure de diagnostic, soit le métamodèle issu des lois de la physique ou de l'expertise possède une structure incertaine qui demeure complexe. Dans un cas comme dans l'autre, il faudrait trouver une structure plus simple, un peu moins précise mais, c'est une contrainte en

diagnostic, qui demeure juste. Une telle opération est toujours possible dans un cadre incertain (même si la précision devient très médiocre).

Notre objectif n'est pas ici d'examiner les différentes méthodes de simplification des structures complexes. Cette tâche serait à la fois laborieuse, du fait de la multitude des structures possibles, et sortirait du cadre du diagnostic. Néanmoins, nous avons cru utile de présenter, au moyen de deux exemples, les différents problèmes qui surviennent lors de la synthèse d'un modèle incertain de structure simple.

Partons du modèle déterministe statique suivant :

$$(\mathcal{S}) \quad y(t) = x(t)(x(t) + 1) : x(t) \in [0 \ 1]$$

et montrons qu'il est possible de lui substituer un modèle linéaire incertain. Selon que la structure est choisie additive ou multiplicative, deux modèles peuvent être exploités :

$$(\mathcal{M}_a) \quad y(t) = \alpha_a x(t) + e_a v_1(t) : \mathcal{A}(v_1(t)) = [-1 \ 1], \quad e_a \in \mathbb{R}^+$$

$$(\mathcal{M}_m) \quad y(t) = (1 + e_m v_2(t)) \alpha_m x(t) : \mathcal{A}(v_2(t)) = [-1 \ 1], \quad e_m \in \mathbb{R}^+$$

La première étape de la synthèse d'un modèle incertain consiste à identifier les paramètres invariants. Or, pour que les paramètres identifiés soient pertinents, l'identification doit tenir compte de la structure des incertitudes du métamodèle dans le critère à minimiser. Il s'agit en effet de trouver les paramètres invariants qui vont minimiser au sens d'un certain critère les erreurs propres à chacun des modèles. A partir des systèmes (\mathcal{S}) , (\mathcal{M}_a) et (\mathcal{M}_m) , nous déduisons l'expression des erreurs dans les deux cas :

$$\varepsilon_a(t) = e_a v_1(t) = x(t)^2 + (1 - \alpha_a)x(t)$$

$$\varepsilon_m(t) = e_m v_2(t) = \frac{1}{\alpha_m} (x(t) + (1 - \alpha_m))$$

Dès lors, puisque la variable $x(t)$ appartient à l'intervalle $[0 \ 1]$, il est naturel de minimiser les erreurs des deux modèles sur le domaine considéré en choisissant les critères $J_a(\alpha_a) = \int_0^1 \varepsilon_a(t)^2 dx$ et $J_m(\alpha_m) = \int_0^1 \varepsilon_m(t)^2 dx$. La minimisation de chacun de ces critères conduit aux solutions $\alpha_a = \frac{1}{4}$ et $\alpha_m = \frac{1}{2}$. La phase d'identification s'achève avec la détermination des valeurs des paramètres invariants.

La seconde étape sur laquelle nous reviendrons plus amplement au cours de la partie III est la caractérisation des variables abstraites ; autrement dit, il s'agit dans l'exemple exposé, de déterminer les caractéristiques e_a et e_m des incertitudes. L'objectif de la partie III est de décrire un algorithme permettant de résoudre ce problème de caractérisation dans les cas de modèles incertains affines par rapport aux variables abstraites. Concernant l'exemple, les résultats que nous avons obtenus apparaissent sur la figure VII. Les surfaces de chacune des deux

enveloppes sont sensiblement identiques (0,5 dans le cas additif et 0,47 dans le cas multiplicatif) ; la différence tient à ce que le modèle (\mathcal{M}_m) est plus précis pour des faibles valeurs de $x(t)$ alors que l'autre (\mathcal{M}_a) l'est plus pour des valeurs proches de l'unité.

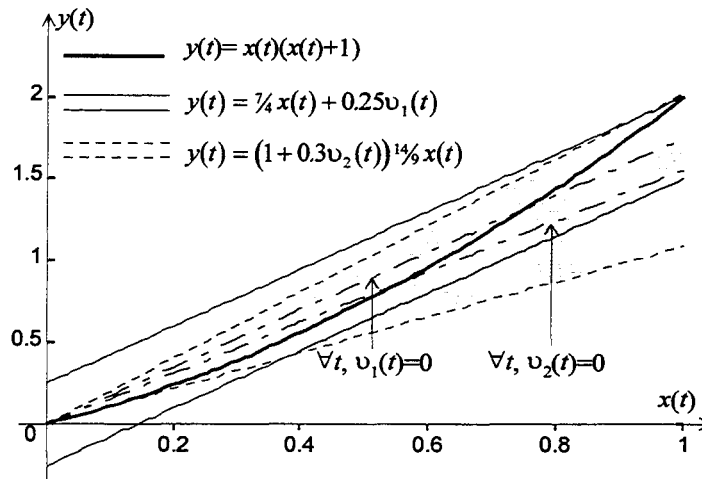


Figure VII- Enveloppes régulière et proportionnelle, optimales dans un cas statique

Pour aborder les problèmes de structure dans des cas multivariables, partons d'un exemple issu des travaux de S. Giuliani [Giuliani,98]. Cette étude est consacrée à la surveillance du fonctionnement des pluviomètres répartis sur l'agglomération nancéenne afin d'accroître la fiabilité de la gestion du réseau d'assainissement. Les métamodèles, sur lesquels reposent les tests de cohérence de la procédure de bon fonctionnement retenue, sont locaux. Ils exploitent l'homogénéité de la pluie en représentant le fait que des pluviomètres voisins doivent normalement recevoir la même intensité de pluie. En ne considérant que trois capteurs, le métamodèle utilisé s'écrit :

$$(\mathcal{M}_{p3}) \quad [R_3] = \frac{1}{2} \sum_{i=1}^2 [R_i] + e_h v_0$$

où $[R_i]$ représente les variables physiques d'intensité de pluie et e_h la caractéristique de l'erreur due à l'hétérogénéité de la pluie

Or, des expérimentations en laboratoire ont montré que les valeurs issues des pluviomètres étaient précises à une erreur relative de 10% près sur leur plage de mesure usuelle (0-250 mm/h) ; autrement dit, la mesure d'une intensité de pluie $[R_i]$ vérifie :

$$\tilde{R}_i = (1 + e_p v_i) [R_i] ; i \in \{1, \dots, 3\} \text{ et } e_p = 0.1$$

En utilisant ces relations avec le métamodèle (\mathcal{M}_{p3}), nous établissons l'hypomodèle suivant :

$$(\mathcal{H}_{p3}) \quad \tilde{R}_3 = \frac{1}{2} [\tilde{R}_1 \quad \tilde{R}_2] l(v) + (1 + e_p v_3) e_h v_0 : l(v) = \begin{bmatrix} 1 + e_p v_3 & 1 + e_p v_3 \\ 1 + e_p v_1 & 1 + e_p v_2 \end{bmatrix}^T$$

Comme la structure des incertitudes dans ce modèle est complexe, il serait intéressant de

trouver une structure plus simple, c'est-à-dire autant que possible affine par rapport aux variables abstraites, pour pouvoir exploiter plus facilement le modèle incertain. Dans cet exemple, nous supposons que les incertitudes de 10% sont suffisamment faibles pour que l'hypomodèle ($\tilde{\mathcal{R}}_{p3}$) demeure proche de celui obtenu par linéarisation autour de $v_i=0$. L'hypomodèle devient :

$$(\tilde{\mathcal{R}}_{p3}) \quad \tilde{R}_3 = [\tilde{R}_1 \quad \tilde{R}_2] l'(v) + e_h v_0 : l'(v) = \begin{bmatrix} 1 \\ 1 \end{bmatrix} + e'_p \begin{bmatrix} -1 & 0 & 1 \\ 0 & -1 & 1 \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \\ v_3 \end{bmatrix}$$

Mais, pour que cette structure simplifiée reste juste, c'est-à-dire pour compenser l'approximation, il faut ici aussi pouvoir caractériser la nouvelle incertitude en déterminant e'_p sachant que $e_p=0,1$. Au moyen de l'algorithme décrit dans la partie suivante, nous avons obtenu que, pour une valeur $e'_p=0,115$, le domaine simplifié contenait le domaine original ; la figure VIII permet de comparer les deux domaines abstraits.

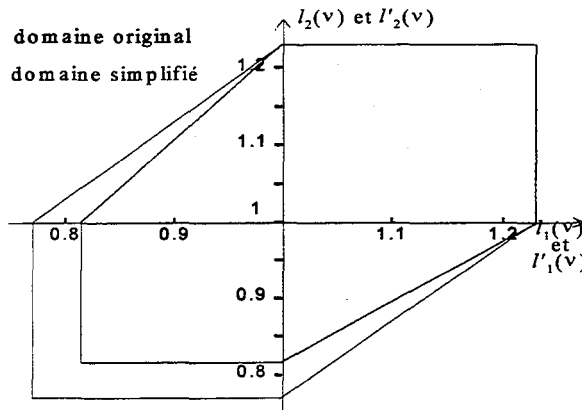


Figure VIII- Comparaison entre un domaine complexe et un domaine simplifié

Ce que nous voulions montrer avant d'aborder la partie consacrée aux algorithmes de caractérisation des incertitudes était tout d'abord que l'identification des paramètres certains pouvait être opérée indépendamment de la caractérisation des incertitudes et d'autre part, qu'à partir de n'importe quelle structure d'incertitude, il était possible d'expliquer un comportement incertain. Certaines structures conduisent cependant à des représentations plus précises ; elles sont généralement issues des lois de la physique mais aussi de la connaissance heuristique d'un système physique.

III. Caractérisation des variables abstraites bornées

Dans cette partie, nous considérons que l'estimation des paramètres invariants, ou *identification paramétrique*, a été réalisée et nous renvoyons le lecteur à des ouvrages spécialisés [Ljung,87], [Richalet,91] et [Walter,94]. La valeur de variables abstraites invariants peut en effet être identifiée, ce qui n'est éminemment pas le cas des variables

abstraites fluctuant au cours du temps (suffisamment rapidement pour ne pas pouvoir être considérées comme invariantes). Nous nous intéresserons aux modèles dont la structure des incertitudes a été fixée, soit à partir des lois de la physique, soit arbitrairement ; nous venons de montrer que si la précision du modèle incertain en dépendait, la justesse n'en est pas pour autant mise en cause⁵. La caractérisation des variables abstraites bornées se résume à rechercher les caractéristiques justes des incertitudes expliquant un comportement donné ; plus précisément, il s'agit de déterminer les rayons des variables abstraites intervenant dans un modèle incertain et qui expliquent les observations disponibles. Si nous raisonnons à partir de modèles causaux du type (Σ_s) et (Σ_d) (partie IV), il faut que les enveloppes déduites des modèles incertains contiennent toutes les mesures des variables physiques correspondantes. Cela ne pose aucune difficulté lorsque le modèle incertain est déduit en totalité des lois de la physique ; néanmoins, dans tous les autres cas, et même dans les cas de modèles de connaissance incertains linéarisés pour en réduire la complexité, il est nécessaire de définir des algorithmes de caractérisation des variables abstraites pour compléter la connaissance disponible.

Nous nous sommes intéressés aux modèles affines en les variables abstraites pouvant s'écrire :

$$\forall k \in \{1, \dots, m\}, \tilde{y}_k = \tilde{x}_k \theta_k + e_k : \mathcal{A}(e_k) = \mathcal{E} = [-\delta_\varepsilon \ \delta_\varepsilon] \text{ et } \mathcal{A}(\theta_k) = \Theta \quad (3)$$

où $\hat{y}_k \in \mathbb{R}$, $e_k \in \mathbb{R}$ et $\theta_k \in \mathbb{R}^n$

Le vecteur θ_k contient des variables abstraites bornées et définit, comme tout champ vectoriel contenant de telles variables, un domaine abstrait tandis que la forme linéaire \tilde{x}_k contient des variables physiques mesurées et des paramètres invariants identifiés. Ce modèle peut être aussi bien statique que dynamique si \tilde{x}_k contient des variables retardées de la variable physique mesurée \tilde{y}_k . Cependant, en aucun cas, ce type de modèle ne peut conduire à des modèles dynamiques à horizon infini tel que des équations d'état. En effet, deux manières de tester la cohérence d'observations avec le modèle (3) sont envisageables.

La première consiste à rechercher si, pour les observations \tilde{x}_k et \tilde{y}_k , il existe au moins une valeur de θ_k et de e_k comprises respectivement dans les domaines Θ et \mathcal{E} et pour lesquelles : $\tilde{y}_k - \tilde{x}_k \theta_k - e_k = 0$.

A l'opposé de cette approche acausale, la cohérence peut être testée par un hypomodèle causal. Dans ce cas, le modèle est utilisé différemment, il devient : $\hat{y}_k = \tilde{x}_k \theta_k + e_k$ pour $\theta_k \in \Theta$ et $e_k \in \mathcal{E}$, et le test de cohérence consiste à vérifier si $\tilde{y}_k \in \mathcal{A}(\hat{y}_k)$.

Quoi qu'il en soit, dans un cas comme dans l'autre, aucun des modèles engendrés par (3) n'est

⁵ Il faut cependant que la structure des incertitudes soit adaptée ; par exemple, une variable abstraite sur une équation récurrente ne pourra pas expliquer les incertitudes observées sur des mesures représentées par une autre équation certaine indépendante.

à horizon infini. En fait, la caractérisation des incertitudes de modèles à horizon infini ne peut être réalisée qu'en étudiant, théoriquement, tout l'horizon des temps et pratiquement, un horizon suffisant pour pouvoir extrapoler. De ce fait, la caractérisation de tels modèles ne peut être décomposée et se trouve exclue des possibilités de simplifications sur lesquelles nous nous penchons maintenant.

III.1. Exemple introductif

Pour éclairer notre démarche, nous proposons de partir d'un exemple simple, dont le modèle statique vérifie (3) :

$$\tilde{y}_k = \theta_{1k} \tilde{x}_{1k} + \theta_{2k} \tilde{x}_{2k} + 1 + e_k$$

Nous disposons d'observations obtenues à trois instants différents ($k \in \{1,2,3\}$) :

$$Y = \begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix} = \begin{bmatrix} 2.5 \\ 1.8 \\ 3.4 \end{bmatrix}, \quad X = \begin{bmatrix} x_{11} & x_{21} \\ x_{12} & x_{22} \\ x_{13} & x_{23} \end{bmatrix} = \begin{bmatrix} 1.1 & 0.1 \\ 0.2 & 1.2 \\ 1.3 & 0.8 \end{bmatrix} \quad \text{et} \quad E = \begin{bmatrix} e_1 \\ e_2 \\ e_3 \end{bmatrix}$$

Supposons dans un premier temps que les paramètres θ_{1k} et θ_{2k} soient invariants ; nous noterons $\theta_c = [\theta_1 \ \theta_2]^T$ avec $\theta_1 = \theta_{1k}$ et $\theta_2 = \theta_{2k} \ \forall k$. Il devient dès lors possible d'utiliser un algorithme d'identification classique ; ainsi, en appliquant l'algorithme des moindres carrés simples au critère quadratique $\sum_{k=1}^3 e_k^2$, nous avons obtenu $\theta_c = [1,45 \ 0,49]^T$. La plus grande erreur d'estimation correspond à la première observation ($k=1$) ; elle conduit à $\|E\|_\infty \leq 0.14$. Le résultat serait satisfaisant si nous ne savions par ailleurs que, physiquement, l'incertitude additive e_k ne peut être que nulle⁶ $\forall k$. Dès lors, la seule explication des incohérences entre le modèle et les observations ne peut provenir que d'une incertitude sur le vecteur de paramètres $\theta_k = [\theta_{1k} \ \theta_{2k}]^T$. Faute de connaissances supplémentaires, nous considérerons que les incertitudes portant sur chacun des paramètres sont indépendantes et de même importance, et qu'elles sont centrées sur l'estimation θ_c obtenue. Le vecteur de paramètres incertains peut alors être explicité :

$$\theta_k \in \Theta \Leftrightarrow \theta_k = \theta_c + \lambda v_k \quad \text{avec} \quad v_k = \begin{bmatrix} v_{1k} \\ v_{2k} \end{bmatrix} : \mathcal{A}(v_k) = \square \mathcal{A}(v_k) \quad (4)$$

où $\lambda \in \mathbb{R}^{+}$ et $\forall k, \|v_k\|_\infty \leq 1$*

Pour caractériser Θ , commençons par définir les domaines de majoration Θ^+ et Θ^- représentant respectivement les domaines où les paramètres θ_k conduisent à des estimées $\hat{y}_k = x_k \theta_k + 1$ se situant au-dessus et au-dessous de la mesure \tilde{y}_k :

⁶ Du fait de l'hétérogénéité de la pluie

$$\Theta^+ = \{\theta_k : \tilde{y}_k \leq \tilde{x}_k \theta_k + 1 ; \forall k \in \{1, \dots, 3\}\} \text{ et } \Theta^- = \{\theta_k : \tilde{y}_k \geq \tilde{x}_k \theta_k + 1 ; \forall k \in \{1, \dots, 3\}\}$$

où $\tilde{x}_k = [\tilde{x}_{1k} \ \tilde{x}_{2k}]$

Ce problème peut être décomposé en problèmes élémentaires puisque :

$$\Theta^+ = \bigcap_{i=1}^3 \Theta_k^+ \text{ avec } \Theta_k^+ = \{\theta_k : \tilde{y}_k \leq \tilde{x}_k \theta_k + 1\} \text{ et } \Theta^- = \bigcap_{i=1}^3 \Theta_k^- \text{ avec } \Theta_k^- = \{\theta_k : \tilde{y}_k \geq \tilde{x}_k \theta_k + 1\}$$

Définissons alors mathématiquement les ensembles des estimées \hat{y}_k associés aux domaines Θ^+ et Θ^- :

$$\mathcal{U}_k^+ = \{\tilde{x}_k \theta_k + 1 : \theta_k \in \Theta^+\} \text{ et } \mathcal{U}_k^- = \{\tilde{x}_k \theta_k + 1 : \theta_k \in \Theta^-\}$$

et de la même manière, l'ensemble $\mathcal{A}(\hat{y}_k)$ que nous recherchons :

$$\mathcal{U}_k = \{\tilde{x}_k \theta_k + 1 : \theta_k \in \Theta\}$$

Alors, en déterminant les domaines Θ^+ et Θ^- , nous pouvons avancer que, par définition :

- (a) si $\Theta \cap \Theta^+ \neq \emptyset$, alors $\forall k \in \{1, \dots, 3\}$, $\exists \theta_k \in \Theta / \tilde{x}_k \theta_k + 1 \geq \tilde{y}_k$
- (b) si $\Theta \cap \Theta^- \neq \emptyset$, alors $\forall k \in \{1, \dots, 3\}$, $\exists \theta_k \in \Theta / \tilde{x}_k \theta_k + 1 \leq \tilde{y}_k$

En réunissant ces deux conditions, nous obtenons :

si $\Theta \cap \Theta^+ \neq \emptyset$ et $\Theta \cap \Theta^- \neq \emptyset$, alors

$$\forall k \in \{1, \dots, 3\}, \exists (\theta_k, \theta'_k) \in \Theta^2 / \tilde{x}_k \theta_k + 1 \geq \tilde{y}_k \geq \tilde{x}_k \theta'_k + 1$$

et l'enveloppe $\mathcal{U} = \mathcal{U}_1 \times \mathcal{U}_2 \times \mathcal{U}_3$ ⁷ définie par l'ensemble Θ contient nécessairement les points de mesures $\tilde{y}_k : k \in \{1, \dots, 3\}$. Or, il suffit que l'intersection $\Theta \cap \Theta^+$ ou $\Theta \cap \Theta^-$ se limite à une unique valeur pour être en présence d'une enveloppe, et plus l'intersection sera importante plus le domaine Θ sera important et en conséquence, moins l'enveloppe sera précise. L'algorithme que nous proposons peut dorénavant être décomposé. Dans un premier temps, nous devons déterminer les domaines élémentaires Θ_k^+ et Θ_k^- pour $k \in \{1, \dots, 3\}$ et en déduire Θ^+ et Θ^- . Dans un deuxième temps, nous devons rechercher le domaine Θ tel que les intersections $\Theta \cap \Theta^+$ et $\Theta \cap \Theta^-$ soient non vides et les plus petites possibles.

Dans cet exemple, chaque domaine élémentaire $\Theta_k^+ = \{\theta_k / \tilde{y}_k \leq \tilde{x}_{1k} \theta_{1k} + \tilde{x}_{2k} \theta_{2k} + 1\}$ correspond à une portion non bornée de l'espace paramétrique délimité par la droite $\tilde{y}_k = \tilde{x}_{1k} \theta_{1k} + \tilde{x}_{2k} \theta_{2k} + 1$. Ces droites-frontières apparaissent sur la figure IX ainsi que les intersections Θ^+ et Θ^- des domaines élémentaires Θ_k^+ entre eux et des domaines Θ_k^- entre eux. L'étude n'a été menée que dans un domaine particulier de l'espace paramétrique baptisé

⁷ Le signe x désigne ici le produit cartésien qui s'applique à des ensembles.

domaine d'étude. La nature du domaine Θ est donnée par (4) : il s'agit d'un orthotope aligné centré sur θ_c ; mais ce domaine ne sera caractérisé que lorsque le paramètre λ , représentatif de la taille du domaine, sera déterminé. Ainsi, dans cet exemple, nous recherchons le plus petit orthotope aligné centré en θ_c et intersectant à la fois les domaines Θ^+ et Θ^- . En ajustant progressivement la dimension de l'orthotope jusqu'à ce qu'il s'inscrive juste entre les domaines de majoration (une procédure systématique sera donnée dans la sous-partie *caractérisation des incertitudes*), nous avons obtenu la valeur $\lambda=0,117$. L'orthotope ainsi calculé est représenté par un orthotope plus foncé sur la figure IX.

Sur la figure IX bis, nous avons représenté les 3 mesures de la variable physique $[y_k]$, les 3 estimations issues des moindres carrés et enfin les domaines incertains obtenus lorsque les variables θ_k appartiennent au domaine paramétrique Θ . Ce domaine peut être représenté par : $\theta = \theta_c + \lambda [v_{1,k} \ v_{2,k}]^T$ où $v_{1,k}$ et $v_{2,k}$ sont des variables abstraites normalisées. Les domaines possibles des y_k sont déduits de : $\hat{y}_k = \tilde{x}_k \theta_c + \lambda \tilde{x}_k [v_{1,k} \ v_{2,k}]^T$. La première mesure effleure la borne inférieure du domaine \mathcal{Y}_1 obtenu alors qu'aucune des mesures ne vient effleurer de bornes supérieures des domaines \mathcal{Y}_k . Ceci s'explique par le fait que l'intersection $\Theta \cap \Theta^-$ se réduit à un point de la frontière alors que l'intersection $\Theta \cap \Theta^+$ est bien plus importante.

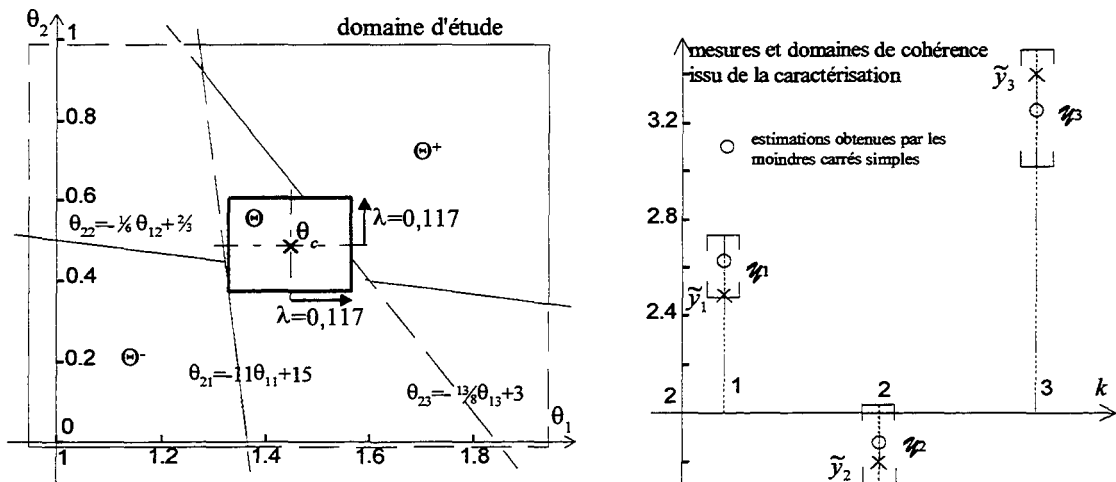


Figure IX et IX bis- Résultats de la caractérisation du modèle incertain de l'exemple

III.2. Cas général sans connaissance a priori des incertitudes

L'algorithme que nous avons développé s'inspire de celui proposé, originalement, par Walter et Piet-Lahanier [Walter et al,87] et peu après par Mo et Norton [Mo et al,88] et récemment exposé très clairement par Belforte [Belforte et al,96]. Bien que les objectifs de l'identification paramétrique dans le cadre d'une approche bornante [Milanese et al,96] soient fondamentalement différents des nôtres puisqu'ils visent à rechercher l'ensemble des paramètres compatibles avec les mesures et une erreur d'équation, tandis que nous, nous recherchons les incertitudes sur les paramètres qui expliquent l'ensemble des mesures, certaines méthodes de résolution sont communes.

Afin de mettre en évidence les différences, partons d'un modèle simple $\tilde{y}_k = \tilde{x}_k \theta_k + e_k$ où la mesure \tilde{y}_k est scalaire et e_k est défini par $\mathcal{A}(e_k) = [-\delta_\varepsilon, \delta_\varepsilon]$. Supposons qu'une valeur central θ_c de θ_k , considérée comme invariante, ait été identifiée et que nous disposions de m mesures $\tilde{y}_k : k \in \{1, \dots, m\}$. Nous noterons δ_0 la plus petite erreur expliquant, lorsque $\theta_k = \theta_c$, l'ensemble des mesures : $\delta_0 = \left\| \begin{bmatrix} \dots & \tilde{y}_k - \tilde{x}_k \theta_c & \dots \end{bmatrix} \right\|_\infty$. En fixant $\delta_\varepsilon = \delta_0$, les domaines Θ^+ et Θ^- , que nous avons défini précédemment, vérifieront naturellement $\Theta^+ \cap \Theta^- = \{\theta_c\}$. Si $\delta_\varepsilon > \delta_0$, il existe plusieurs valeurs possibles pour θ_k expliquant chacune l'ensemble des points de mesure. L'intersection des domaines Θ^+ et Θ^- coïncide avec l'ensemble des paramètres possibles FPS_y recherché par Norton, Belforte et Milanese. Par contre, notre approche consiste à étudier les cas où $\delta_\varepsilon < \delta_0$ pour lesquels les domaines de majoration ne s'intersectent pas. Nous recherchons les incertitudes qu'il faudrait sur θ_k , centré en θ_c et défini par un domaine Θ , pour expliquer

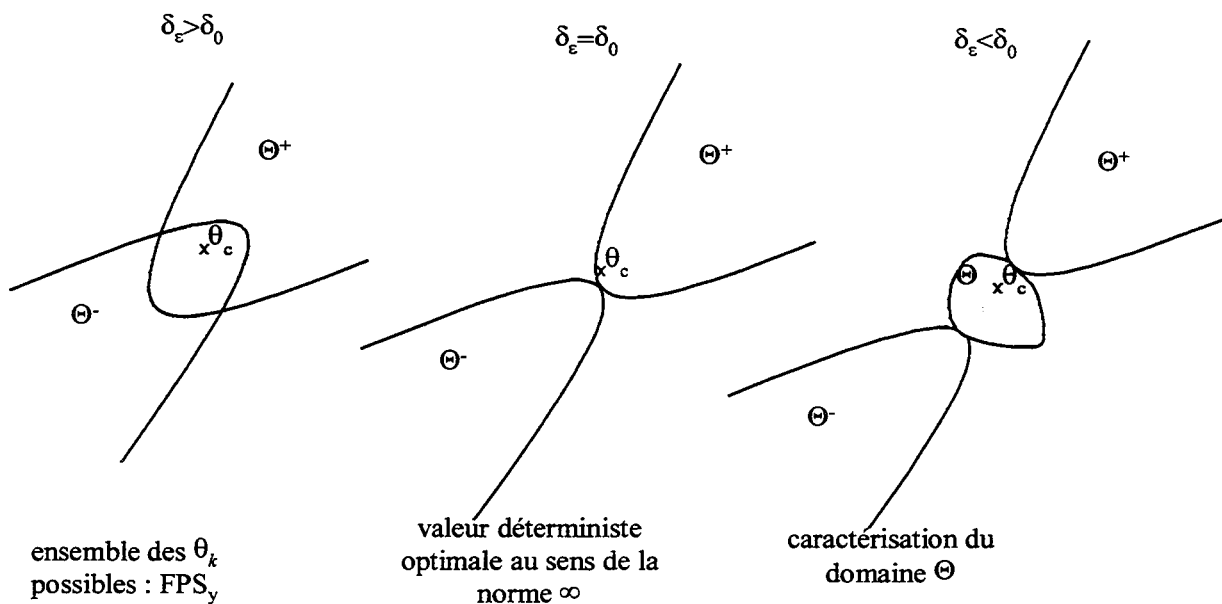


Figure X – Différence entre la caractérisation et le calcul de domaine FPS_y

l'ensemble des mesures. La figure X illustre ces trois situations.

Les approches classiques recherchent en quelque sorte des domaines isocritères⁸ alors que nos objectifs sont nouveaux. Signalons aussi que les approches stochastiques ne permettent pas d'envisager de caractérisation.

Avant de poursuivre, revenons sur le modèle (3) et adjoignons-lui certaines hypothèses préliminaires afin de garantir le succès de l'algorithme que nous proposons.

⁸ Le critère étant le maximum de la valeur absolue de l'erreur d'équation

Hypothèses préliminaires

Partant du modèle (3), les hypothèses d'indépendance suivantes sont supposées vérifiées pour tout $i \neq j$:

- (a) le vecteur θ_i est indépendant de e_i
- (b) le vecteur θ_i est indépendant de e_j (5)
- (c) le vecteur θ_i est indépendant de θ_j . ✓

Si l'hypothèse (a) n'est pas vérifiée, les domaines Θ_k^+ et Θ_k^- sont alors donnés par : $\{\theta_k / \tilde{y}_k \leq \tilde{x}_k \theta_k + e_k(\theta_k)\}$ et $\{\theta_k / \tilde{y}_k \geq \tilde{x}_k \theta_k + e_k(\theta_k)\}$, et le calcul des domaines Θ_k^+ et Θ_k^- , bien qu'a priori possible, ne peut plus être envisagé dans un cas général. Cette hypothèse permet de contourner cette difficulté. L'hypothèse (b) permet d'éviter les cas où les domaines Θ_k^+ et Θ_k^- sont, par exemple, issus de : $\{\theta_k / \tilde{y}_k \leq \tilde{x}_k \theta_k + e_k(\theta_{k-1})\}$ et $\{\theta_k / \tilde{y}_k \geq \tilde{x}_k \theta_k + e_k(\theta_{k-1})\}$. En effet, dans ce cas, la résolution de Θ_k^+ et Θ_k^- dépendrait des domaines Θ_{k-1}^+ et Θ_{k-1}^- ainsi que de ceux correspondant aux instants précédents. Cette résolution deviendrait très complexe et souvent inextricable. Enfin, si l'hypothèse (c) n'est pas satisfaite, il peut se produire que les domaines Θ_k^+ et Θ_k^- soient solution de $\{\theta_k / \tilde{y}_k \leq \tilde{x}_k \theta_k(\theta_{k-1}) + e_k\}$ et $\{\theta_k / \tilde{y}_k \geq \tilde{x}_k \theta_k(\theta_{k-1}) + e_k\}$. Dès lors, les domaines obtenus en négligeant la dépendance ne seront qu'une majoration des domaines réels. En effet, négliger les dépendances revient à considérer que toutes les valeurs de θ_k peuvent être atteintes alors qu'il existe une contrainte dépendant ici des domaines Θ_{k-1}^+ , Θ_{k-1}^- et de ceux correspondant aux instants précédents.

Même si ces hypothèses peuvent paraître contraignantes, il ne faut pas oublier que la majorité des structures mènent à des modèles justes. Ces hypothèses donnent des indications sur les structures qu'il faut préférer, c'est-à-dire auxquelles il faut tenter de se ramener. Seule la précision peut s'en trouver dégradée.

Notations et définitions

La première étape de la caractérisation est la détermination des domaines de majoration. Si les hypothèses (5), notamment (b) et (c), sont satisfaites, les domaines de majoration Θ^+ et Θ^- s'obtiennent à partir des intersections suivantes :

$$\Theta^+ = \bigcap_{k=1}^m \Theta_k^+ \text{ et } \Theta^- = \bigcap_{k=1}^m \Theta_k^- \tag{6}$$

Puisque ces domaines ne sont pas finis, nous restreindrons le domaine paramétrique à un orthotope aligné, nommé *domaine d'étude* \mathcal{D} , qui a le mérite d'être fini et ainsi de simplifier les représentations. Ce domaine satisfait à la condition (30) du chapitre 2 ; les équations de ces frontières s'obtiennent donc en suivant la méthode que nous avons précédemment détaillée. En substituant une inégalité à l'égalité caractérisant les frontières, nous obtenons les domaines d'étude élémentaires dont l'intersection engendre le domaine d'étude \mathcal{D} . L'orthotope

\mathcal{D} dans un espace de dimension n possède $2n$ facettes qui peuvent toujours être caractérisées par les relations suivantes :

$$\mathcal{D} = \bigcap_{i=1}^{2n} \Theta_i^{\mathcal{D}} \text{ avec : } \Theta_i^{\mathcal{D}} = \{\theta / a_i \theta \geq b_i\} \quad (7a)$$

où a_i est une forme linéaire sur \mathbb{R}^n et b_i est un réel

Les frontières de ce domaine sont données par :

$$B_i^{\mathcal{D}} = \{\theta / a_i \theta = b_i\} : \forall i \in \{1, \dots, 2n\}. \quad (7b)$$

Dans ce contexte, les domaines de majoration donnés initialement par (6) deviennent :

$$\Theta^+ = \mathcal{D} \cap \bigcap_{k=1}^m \Theta_k^+ = \left(\bigcap_{k=1}^{2n} \Theta_k^{\mathcal{D}} \right) \cap \left(\bigcap_{k=1}^m \Theta_k^+ \right) \text{ et } \Theta^- = \mathcal{D} \cap \bigcap_{k=1}^m \Theta_k^- = \left(\bigcap_{k=1}^{2n} \Theta_k^{\mathcal{D}} \right) \cap \left(\bigcap_{k=1}^m \Theta_k^- \right) \quad (8)$$

où \mathcal{D} est un orthotope aligné arbitrairement choisi pour définir le domaine d'étude et Θ^+ et Θ^- sont respectivement les domaines de majoration supérieur et inférieur

Par ailleurs, les domaines de majoration élémentaires sont définis par :

$$\Theta_k^+ = \{\theta_k : \tilde{y}_k \leq \tilde{x}_k \theta_k + e_k\} \text{ et } \Theta_k^- = \{\theta_k : \tilde{y}_k \geq \tilde{x}_k \theta_k + e_k\} \quad (9a)$$

Or, en utilisant la propriété (8), nous établissons que $A(\tilde{y}_k - e_k) = [\tilde{y}_k - \delta_\varepsilon, \tilde{y}_k + \delta_\varepsilon]$ avec $A(e_k) = [-\delta_\varepsilon, \delta_\varepsilon]$. Alors, grâce à l'hypothèse (a) de (5), la définition des domaines élémentaires devient :

$$\Theta_k^+ = \{\theta_k : \tilde{y}_k - \delta_\varepsilon \leq \tilde{x}_k \theta_k\} \text{ et } \Theta_k^- = \{\theta_k : \tilde{y}_k + \delta_\varepsilon \geq \tilde{x}_k \theta_k\} \quad (9b)$$

Remarque

En considérant les équations (9) définissant les domaines de majoration élémentaires, il apparaît nettement qu'ils dépendent de la borne δ_ε caractéristique de l'erreur d'équation. Bien que nous ayons choisi de simplifier les écritures, nous aurions dû noter ces domaines $\Theta_k^+(\varepsilon)$ et $\Theta_k^-(\varepsilon)$ et, par extension, $\Theta^+(\varepsilon)$ et $\Theta^-(\varepsilon)$. Remarquons aussi que les domaines de majoration sont indépendants de toute structure a priori du domaine paramétrique Θ .

Chacun des domaines élémentaires correspond à un demi-espace paramétrique de dimension n délimité par un hyperplan de dimension $n-1$. Les hyperplans seront notés B_k^+ et B_k^- ; ils sont définis par :

$$B_k^+ = \{\theta_k : \tilde{y}_k - \delta_\varepsilon = \tilde{x}_k \theta_k\} \text{ et } B_k^- = \{\theta_k : \tilde{y}_k + \delta_\varepsilon = \tilde{x}_k \theta_k\} \quad (10)$$

L'algorithme repose sur un lemme traditionnel de l'algèbre selon lequel l'intersection de domaines convexes engendre un domaine convexe. Il devient alors évident que les domaines de majoration Θ^+ et Θ^- sont, en plus d'être finis, convexes puisque autant le domaine d'étude \mathcal{D} que les domaines élémentaires le sont aussi. De plus, puisque les frontières de chacun des domaines sont des hyperplans de dimensions $n-1$, il en découle que les domaines de

majoration sont des polytopes convexes et finis. En conséquence, ces domaines sont entièrement définis par leurs sommets.

Le principe de la génération des polytopes Θ^+ ou Θ^- consiste à partir du domaine d'étude \mathcal{D} puis à tronquer ce domaine en calculant successivement l'intersection avec chacun des domaines élémentaires de majoration supérieurs ou inférieurs. Le polytope obtenu est caractérisé par ces sommets qui sont eux-mêmes issus de l'intersection de n hyperplans-frontières de dimension $n-1$ ou facettes. En effet, un point⁹ dans un espace de dimension n ne peut être engendré que par n hyperplans de dimension $n-1$; ainsi, à chaque sommet seront associés n équations de facettes génératrices qui correspondent soit à des facettes du domaine d'étude, soit à des facettes définies par (10). Nous noterons σ_i^+ les coordonnées du $i^{\text{ème}}$ sommet du polytope Θ^+ et σ_i^- les coordonnées du $i^{\text{ème}}$ sommet du polytope Θ^- . Chaque sommet σ_i^+ est issu d'un n -uplet ω_i^+ d'équations appartenant à $B^{\mathcal{D}} = \{B_i^{\mathcal{D}} : i \in \{1, \dots, 2n\}\}$ et à $B^+ = \{B_i^+ : i \in \{1, \dots, m\}\}$; les sommets σ_i^- sont associés à un n -uplets ω_i^- de $B^{\mathcal{D}}$ et de $B^- = \{B_i^- : i \in \{1, \dots, m\}\}$.

Calcul itératif des domaines de majoration

Pour obtenir un sommet de n hyperplans, il est indispensable que les hyperplans s'intersectent en un unique point. Un n -uplet ω_j d'hyperplans est donné, dans le cas général, par : $v_j \theta = w_j \forall j \in \{1, \dots, n\}$. Grâce au formalisme matriciel, les équations associées à ω_j peuvent être rassemblées : $V_j \theta = W_j$ avec $V_j = [v_1^T \dots v_n^T]^T$ et $W_j = [w_1^T \dots w_n^T]^T$. Si la matrice V_j est de plein rang, alors le sommet associé à ce n -uplet s'obtient simplement : $\sigma_j = V_j^{-1} W_j$. Le cas échéant, deux situations peuvent se présenter : des vecteurs directeurs des hyperplans v_i sont colinéaires ou des vecteurs v_i sont coplanaires. Dans une situation comme dans l'autre, le n -uplet n'engendre aucun sommet : la dimension de l'intersection est non nulle. Nous rejeterons ces cas puisque nous ne recherchons que les sommets qui suffisent à caractériser totalement un polytope.

Le domaine d'étude s'obtient sans grande difficulté en respectant la méthode évoquée précédemment. Pour décrire le calcul itératif d'un polytope, nous nous intéresserons au cas particulier du domaine Θ^+ puisque l'approche se transpose automatiquement au domaine Θ^- . Supposons que le domaine d'étude ait été déterminé et que les $k-1$ premières mesures aient été prises en compte. Le domaine Θ^+ vérifie donc $\Theta^+ = \mathcal{D} \cap \bigcap_{k=1}^{k-1} \Theta_k$ et il est caractérisé par l sommets σ_i^+ réunis dans l'ensemble $S_{k-1}^+ = \{\sigma_i^+ : i \in \{1, \dots, l\}\}$ rapportés respectivement aux l n -uplets ω_i^+ réunis dans l'ensemble $\Omega_{k-1}^+ = \{\omega_i^+ : i \in \{1, \dots, l\}\}$. Dès lors, pour resserrer le domaine Θ^+ , nous prenons en compte la contrainte issue des $k^{\text{èmes}}$ observations dont la forme

⁹ La dimension d'un point est nulle.

est donnée par (9b) : $\Theta_k^+ = \{\theta_k : \tilde{y}_k - \delta_\varepsilon \leq \tilde{x}_k \theta_k\}$. L'ensemble I^+ des sommets de S_{k-1}^+ n'appartenant pas au domaine Θ_k^+ s'obtient ainsi : $I^+ = \{i \in \{1, \dots, l\} : \tilde{y}_k - \delta_\varepsilon > \tilde{x}_k \sigma_i^+\}$. Ces sommets doivent alors être éliminés de l'ensemble S_{k-1}^+ et les n -uplets qui s'y rapportent, éliminés de Ω_{k-1}^+ .

Les sommets n'appartenant plus aux polytopes suite aux $k^{\text{ièmes}}$ observations, ont été déterminés. Reste cependant à déterminer quels sont les nouveaux sommets engendrés par le domaine Θ_k^+ . Dans le cas le plus simple d'un sommet tronqué $\sigma_i^+ : i \in I^+$ isolé, les nouveaux sommets s'obtiennent sans difficulté. Si ω_i^+ est le n -uplet associé à ce sommet tronqué, il est évident que les nouveaux sommets seront engendrés par l'intersection de l'hyperplan défini par $B_k^+ = \{\theta_k : \tilde{y}_k - \delta_\varepsilon = \tilde{x}_k \theta_k\}$ et les arêtes¹⁰ intersectées par sommet isolé tronqué. Une arête intersectée est définie par un $(n-1)$ -uplet d'hyperplans de ω_i^+ ; il y a donc n arêtes intersectées par sommet isolé tronqué. En ajoutant l'hyperplan B_k^+ aux $(n-1)$ -uplets correspondant aux arêtes intersectées, nous obtenons n nouveaux n -uplets définissant les nouveaux sommets de Θ^+ . En procédant ainsi pour toutes les arêtes intersectées, nous obtiendrons une remise à jour du domaine de majoration Θ^+ se traduisant par un nouvel ensemble de sommets S_k^+ et un nouvel ensemble Ω_k^+ de n -uplets associé à S_k^+ dont les dimensions sont a priori différentes des ensembles S_{k-1}^+ et Ω_{k-1}^+ .

La procédure que nous venons de décrire ne s'applique néanmoins qu'à des sommets isolés, c'est-à-dire qu'aucune arête ne joint deux sommets tronqués. Sur la figure XI apparaît un exemple de troncature de sommets non isolés puisque les sommets tronqués σ_1^+ et σ_2^+ ont une arête en commun. En appliquant la procédure précédente, les sommets σ_1^+ et σ_2^+ seraient éliminés puisqu'ils n'appartiennent pas au domaine Θ_k^+ . Viendraient alors s'ajouter aux sommets restants les 3 sommets σ_a , σ_b et σ_c . Malheureusement, le polytope solution Θ^+ ne serait plus caractérisé par les nouveaux sommets obtenus.

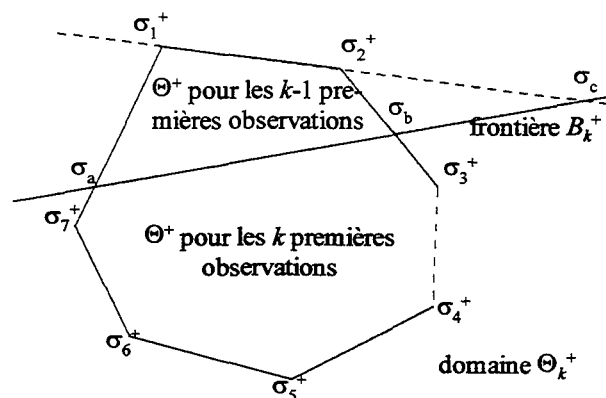


Figure XI- Exemple d'intersection de sommets non isolés

¹⁰ Une arête est une droite définie par l'intersection de $n-1$ hyperplan(s) dans un espace de dimension n .

Pour que la procédure reste valide, il est nécessaire d'exclure les $(n-1)$ -uplets correspondants aux arêtes communes lors du calcul des nouveaux sommets. La procédure corrigée devient la suivante. Après avoir recensé les sommets tronqués $\sigma_i^+ : i \in I^+$ de S_{k-1}^+ , il faut déterminer les arêtes correspondantes, c'est-à-dire les $n \cdot \text{card}(I^+)$ $(n-1)$ -uplets correspondants aux $\text{card}(I^+)$ n -uplets des hyperplans associés aux sommets tronqués. Il faut alors éliminer les arêtes communes, autrement dit, les $(n-1)$ -uplets apparaissant plusieurs fois. Alors seulement, les nouveaux sommets pourront être déterminés comme dans le cas de sommets isolés. Pour revenir à l'exemple de la figure XI, le sommet parasite σ_c disparaîtra parce qu'il correspond à une arête commune à deux sommets tronqués.

Grâce à cette procédure, les domaines de majoration peuvent être déterminés en partant du domaine d'étude \mathcal{D} et en calculant itérativement les intersections avec les domaines de majoration élémentaires indicés de 1 à m .

Caractérisation des incertitudes

Le domaine des incertitudes Θ est défini par un champ de vecteur contenant des variables abstraites qui peut être a priori quelconque pour peu qu'il vérifie les hypothèses (5). Il n'est guère possible de définir le meilleur domaine Θ dans l'absolu ; nous pourrions être tenté de rechercher le domaine Θ de dimension la plus petite et vérifiant $\Theta \cap \Theta^+ \neq \emptyset$ et $\Theta \cap \Theta^- \neq \emptyset$. Or, en revenant à l'exemple présenté par la figure IX, nous pouvons constater combien il est difficile de définir la notion de dimension la plus petite. Si l'on choisit la surface du domaine obtenu, alors il existe une infinité de solution minimale (surface nulle) où le paramètre θ_2 serait considéré comme certain. Nous pourrions préférer la somme des incertitudes sur les paramètres θ_1 et θ_2 , mais ici aussi, il existerait plusieurs solutions qui n'auraient pas nécessairement un sens physique. Nous avons opté pour une approche consistant à représenter au mieux, à l'instar des variables stochastiques, la connaissance physique disponible. Nous chercherons autant que possible à exploiter les connaissances a priori sur la répartition des incertitudes ; et, lorsque la répartition est inconnue, nous choisirons une structure incertaine équirépartie. De plus, pour éviter un nombre trop important de paramètres à caractériser, nous avons choisi de centrer le domaine Θ sur une valeur identifiée de paramètres "moyens" issus d'une procédure d'identification considérant les paramètres comme invariants. Le domaine Θ obtenu dépendra certes des paramètres "moyens" identifiés, aussi sera-t-il préférable, pour que les résultats soient optimum, de préférer des algorithmes d'identification prenant en compte la structure incertaine du modèle (cf. partie IV.4). Tous les paramètres "moyens" obtenus conduisent à des résultats justes mais plus ou moins précis ; pratiquement, il est préférable d'essayer différentes valeurs de paramètres "moyens" et de ne retenir que ceux conduisant aux résultats les plus précis et plausibles à la fois. Enfin, la structure du champ vectoriel définissant le domaine Θ , doit être

fixée de telle sorte qu'il soit facile d'en déterminer les frontières et de déterminer s'il existe ou non des intersections avec les domaines Θ^+ et Θ^- . Puisque les polytopes de majoration sont déterminés par les coordonnées de leurs sommets et les équations de leurs facettes, il est préférable, par souci de simplicité, de choisir un domaine Θ sous forme de polytope. Le domaine obtenu sera lui aussi caractérisé par ses sommets et l'équation de ses facettes. Le même type de procédure que celle décrite précédemment pourra être utilisé pour déterminer si les sommets des polytopes Θ^+ et Θ^- appartiennent ou pas au domaine Θ . Nous nous restreindrons donc aux domaines du type :

$$\Theta = \{ \theta_k : \theta_k = \theta_c + T(\lambda)v_k \} \text{ avec } \|v_k\| \leq 1 \quad (11a)$$

où θ_c correspond à des paramètres "moyens"

Les paramètres λ apparaissant dans la matrice $T(\lambda)$ correspondent aux caractéristiques des incertitudes à déterminer. La matrice $T(\lambda)$ peut cependant être simplifiée en considérant que sa structure est issue d'une connaissance a priori et que seul un paramètre λ subsiste et qu'il représente l'importance de l'incertitude. Le domaine $\Theta = \Theta(\lambda)$ s'écrira :

$$\Theta(\lambda) = \{ \theta_k : \theta_k = \theta_c + \lambda T v_k \} \text{ avec } \|v_k\| \leq 1 \text{ et } \lambda \in \mathbb{R}^+ \quad (11b)$$

où T est déterminé

Il s'agit alors de rechercher le plus petit scalaire λ tel que $\Theta(\lambda) \cap \Theta^+ \neq \emptyset$ et $\Theta(\lambda) \cap \Theta^- \neq \emptyset$. Or, les développements du chapitre 2 (30) nous permettent non seulement d'affirmer que $\Theta(\lambda)$ est un parallélotope centré en θ_c , mais aussi d'en calculer les équations des facettes et les sommets.

La procédure de recherche est simple. Il suffit de partir d'une valeur suffisamment grande de λ pour que $\Theta(\lambda) \cap \Theta^+ \neq \emptyset$ et $\Theta(\lambda) \cap \Theta^- \neq \emptyset$, puis de diminuer cette valeur, selon une progression dichotomique par exemple, jusqu'à ce que $\Theta(\lambda) \cap \Theta^+ \neq \emptyset$ ou $\Theta(\lambda) \cap \Theta^- \neq \emptyset$ ne soit plus vérifié de façon à encadrer la valeur de λ par une valeur λ_v où les conditions $\Theta(\lambda) \cap \Theta^+ \neq \emptyset$ et $\Theta(\lambda) \cap \Theta^- \neq \emptyset$ sont vérifiées et une valeur λ_n où elles ne le sont plus. Enfin, pour accroître la précision de λ , il faut reprendre la procédure en testant la valeur milieu de $[\lambda_n, \lambda_v]$ et ainsi resserrer l'intervalle de détermination de λ jusqu'à une valeur $|\lambda_v - \lambda_n|$ d'arrêt. La valeur λ_v sera alors retenue.

Pour tester la condition $\Theta(\lambda) \cap \Theta^+ \neq \emptyset$ et, par analogie, $\Theta(\lambda) \cap \Theta^- \neq \emptyset$, il faut tester s'il existe au moins un sommet de $\Theta(\lambda)$ appartenant au domaine Θ^+ ou au moins un sommet de Θ^+ appartenant au parallélotope $\Theta(\lambda)$.

Remarquons que la simplification des domaines de majoration Θ^+ ou Θ^- ou paramétrique Θ , par des ellipsoïdes ou des orthotopes alignés, circonscrits ou inscrits ne présente guère d'intérêt pour la caractérisation. La première raison est qu'a priori, les domaines de majoration ne sont pas finis, la seconde est qu'il est bien plus facile de tester si deux polytopes

s'intersectent qu'un polytope avec une ellipse, ou deux ellipses entre elles. Enfin, approcher les domaines de majoration par des orthotopes alignés est une approximation inadmissible dans la plupart des cas.

III.3. Exemple physique

Revenons au problème de diagnostic de défauts de pluviomètres que nous avons évoqué au cours de la partie précédente. Nous avons montré comment la structure pouvait être simplifiée dans le cas des métamodèles de bon fonctionnement faisant intervenir 3 capteurs voisins. Comme les travaux de Giuliani [Giuliani,98] ont montré que les métamodèles à 4 capteurs étaient plus fiables, le métamodèle retenu s'écrit :

$$\left(\mathcal{M}_{p4}\right) \quad [R_{4,k}] = \sum_{i=1}^3 \beta_i [R_{i,k}] + e_h v_{k,0}^0 : \mathcal{A}(v_{k,0}^0) = [-1 \ 1]$$

où $[R_i]$ représente les variables physiques d'intensité de pluie et e_h la caractéristique de l'erreur due à l'hétérogénéité de la pluie

La présence des paramètres β_i s'explique par le fait que les 4 capteurs d'un même métamodèle ne collectent pas la même quantité de pluie du fait des obstacles les environnant. Rappelons que les pluviomètres ont une précision de 10% ; autrement dit, les mesures d'intensité de pluie vérifient $\tilde{R}_{i,k} = (1 + e_p^0 v_{i,k}^0) [R_{i,k}] \quad \forall i \in \{1, \dots, 4\}$ avec e_p^0 voisin de 0,1. L'hypomodèle suivant se déduit du métamodèle (\mathcal{M}_{p4}) et des relations précédentes :

$$\left(\mathcal{H}_{p4}^0\right) \quad \tilde{R}_{4,k} = \sum_{i=1}^3 \beta_i \frac{1 + e_p^0 v_{4,k}^0}{1 + e_p^0 v_{i,k}^0} \tilde{R}_{i,k} + (1 + e_p^0 v_{4,k}^0) e_h v_{0,k}^0 : \mathcal{A}(v_{i,k}^0) = [-1 \ 1]$$

Or, la condition (a) de (5) n'est pas vérifiée. En utilisant la propriété (14) du chapitre 2, nous pouvons déduire un majorant du terme de droite de l'expression ci-dessus :

$$\left(\mathcal{H}'_{p4}\right) \quad \hat{R}'_{4,k} = \sum_{i=1}^3 \beta_i \frac{1 + e_p^0 v_{4,k}^0}{1 + e_p^0 v_{i,k}^0} \tilde{R}_{i,k} + (1 + e_p^0) e_h v_{0,k}^0 \quad \text{avec: } \tilde{R}_{4,k} \in \mathcal{A}(\hat{R}'_{4,k})$$

Pour obtenir une structure affine en les paramètres incertains, nous procédons comme dans la partie II par linéarisation de la structure. Alors il existe e_p et e_h conduisant à un hypomodèle dont la structure est proche de l'original et qui vérifie les hypothèses (5) :

$$\left(\mathcal{H}_{p4}\right) \quad \hat{R}_{4,k} = \sum_{i=1}^3 \beta_i (1 + e_p (v_{4,k} - v_{i,k})) \tilde{R}_{i,k} + (1 + e_p) e_h v_{0,k} \quad \text{avec: } \tilde{R}_{4,k} \in \mathcal{A}(\hat{R}_{4,k})$$

La première étape consiste à identifier les paramètres invariants $\beta_i; i \in \{1, \dots, 3\}$. Pour ce faire, nous avons considéré qu'en appliquant un algorithme des moindres carrés simples sur un nombre très important de mesures (7409 instants de mesures correspondant à environ 1 mois d'enregistrement des intensités de pluie), les valeurs obtenues seraient une bonne estimation des paramètres invariants $\beta_i; i \in \{1, \dots, 3\}$. Nous avons obtenu les résultats suivants : $\beta_1=0,5721$,

$\beta_2=0,1355$ et $\beta_3=0,3586$.

Un modèle comparable à (3) s'obtient en posant : $\tilde{y}_k = \tilde{R}_{4,k}$, $\tilde{x}_k = [\tilde{R}_{1,k} \quad \tilde{R}_{2,k} \quad \tilde{R}_{3,k}]$, $\theta_c = [\beta_1 \quad \beta_2 \quad \beta_3]^T$, $\delta_\varepsilon = (1+e_p)e_h$ et

$$\theta_k = \theta_c + e_p T v_k \text{ avec } v_k = [v_{1,k} \quad v_{2,k} \quad v_{3,k} \quad v_{4,k}]^T \text{ et } T = e_p \begin{bmatrix} -\beta_1 & 0 & 0 & \beta_1 \\ 0 & -\beta_2 & 0 & \beta_2 \\ 0 & 0 & -\beta_3 & \beta_3 \end{bmatrix}$$

Pour alléger les calculs nécessaires à la caractérisation des incertitudes, nous n'avons opéré que sur un échantillon représentatif de 231 instants de mesure sur 7409. Le lecteur pourrait s'interroger sur le bien fondé d'un tel choix ; mais, même si les caractéristiques obtenues des incertitudes ne conduisaient pas à une enveloppe contenant l'ensemble des 7409 quadruplets de mesures, il demeure toujours possible d'accroître les valeurs des caractéristiques obtenues pour y parvenir après coup. Une procédure de caractérisation tend à expliquer l'ensemble des observations, et a fortiori, les plus aberrantes. Il est donc évident que pour obtenir une caractérisation correcte, il faut filtrer préalablement l'ensemble de ce qui ne doit pas être expliqué, notamment les valeurs aberrantes [Kobi,94] et le spectre haute fréquence si nécessaire. Nous avons effectué plusieurs essais de caractérisation, les uns sans filtrage préalable des mesures, les autres avec les mesures filtrées par un filtre d'ordre 5 : $y_k = \frac{1}{5} \sum_{i=1}^5 y_{k-i}$. Ce filtrage n'est pas aberrant dans le sens où le modèle utilisé n'explique pas les phénomènes dynamiques apparaissant dans le spectre haute fréquence. Le résultat de l'identification conduit à une erreur maximum (au sens de la norme infinie) de 1,3 qui correspond à une hétérogénéité de 1,3 mm/h de pluie s'il n'existait pas d'incertitudes de mesures (enveloppe régulière). Or, nous savons d'une part qu'une hétérogénéité de la pluie entre capteurs voisins inférieure à 1mm/h est plus raisonnable, et d'autre part, que les capteurs sont précis avec une erreur relative de 10% correspondant à $e_p \approx 0,115$ (figure VIII).

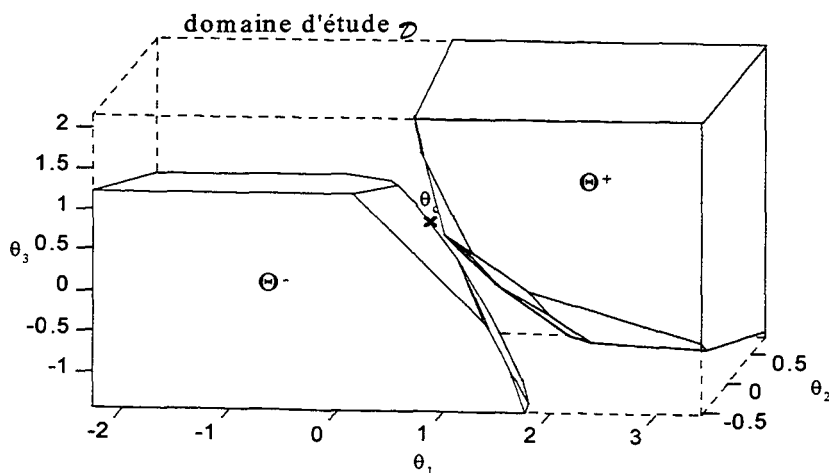


Figure XII- Domaine de majoration obtenu pour 10 instants de mesures ($\delta_\varepsilon=0,7$)

La figure XII représente les domaines de majoration obtenus pour une dizaine de quadruplets de mesures, sans filtrage et pour une valeur $\delta_\varepsilon=0,7$.

Nous avons utilisé 231 instants de mesures au total pour la détermination des domaines de majoration. Les résultats de la seconde étape de la caractérisation sont résumés dans le tableau I :

	sans filtrage	avec filtrage
$\delta_\varepsilon=0,4$	$e_p=0,41 \Rightarrow e_p^0 \approx 0,36$ et $e_h \approx 0,28$	$e_p=0,18 \Rightarrow e_p^0 \approx 0,16$ et $e_h \approx 0,32$
$\delta_\varepsilon=0,7$	$e_p=0,23 \Rightarrow e_p^0 \approx 0,19$ et $e_h \approx 0,57$	$e_p=0,14 \Rightarrow e_p^0 \approx 0,12$ et $e_h \approx 0,63$

Tableau I- Résultats de caractérisation

En nous appuyant sur la connaissance physique que nous venons d'évoquer, nous concluons que les résultats obtenus avec filtrage pour une valeur $\delta_\varepsilon=0,7$ sont les plus plausibles. L'erreur relative de mesure correspond à une valeur de 12% proche des 10% attendus et l'hétérogénéité de la pluie égale à 0,63mm/h est bien inférieure à 1mm/h.

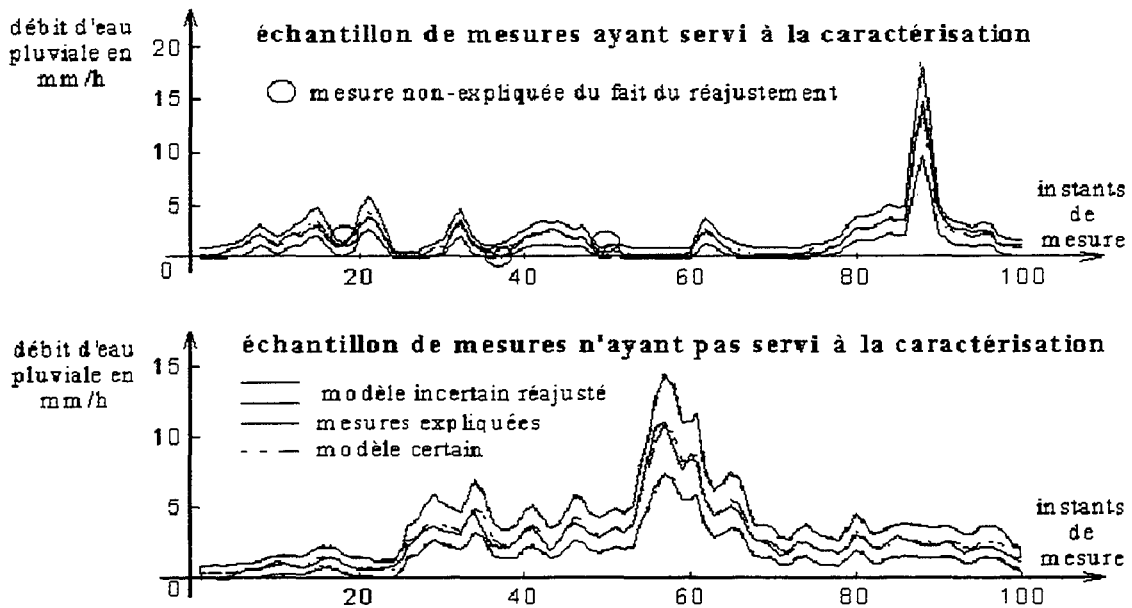


Figure XIII- Enveloppes obtenues des résultats de la caractérisation

Après avoir représenté les résultats, nous avons constaté qu'en diminuant $\delta_\varepsilon=0,7$ à $\delta_\varepsilon=0,4$, les mesures sortaient très peu des enveloppes obtenues. Cette substitution se justifie par le fait que la caractérisation conduit à des paramètres expliquant l'ensemble des points de mesures. Or, il est souvent intéressant de diminuer l'erreur d'équation correspondant à la partie régulière de l'enveloppe pour accroître la précision du modèle au prix certes d'une perte de justesse. Néanmoins, dans une optique de diagnostic, quelques fausses alarmes de durée très brève se rejettent très facilement. Nous avons donc réajusté la valeur de δ_ε dans le modèle pour atteindre une hétérogénéité de pluie de $e_h \approx 0,35$ mm/h. La figure XIII compare les mesures

$\tilde{R}_{4,k}$ avec les enveloppes obtenues du modèle exploitant les incertitudes caractérisées, $\theta = \theta_c + e_p T v_k$:

$$\hat{R}_{4,k} = \left[\tilde{R}_{1,k} \quad \tilde{R}_{2,k} \quad \tilde{R}_{3,k} \right] (\theta_c + e_p T v_k) + \delta_\varepsilon v_{0,k}$$

avec $e_p = 0,1429$ et $\delta_\varepsilon = 0,4$

Cet exemple plaide pour ce type d'approche. En effet, une approche classique ne peut conduire, par le truchement du seuillage, qu'à des enveloppes régulières qui ne peuvent pas prendre en compte des erreurs relatives d'instrumentation qui sont pourtant largement majoritaires dans la pratique. Dans cet exemple, nous constatons que les enveloppes obtenues varient en fonction de l'intensité de pluie ce qui traduit bien le fait que plus l'intensité est importante, et moins la mesure est précise. Pour revenir à nos objectifs de diagnostic, une procédure de diagnostic orientée bon fonctionnement se contenterait de tester si les observations appartiennent bien à la zone grisée de la figure XIII ; sinon, il faudrait conclure à une incohérence entre le modèle de bon fonctionnement et les observations et à la présence manifeste d'un défaut de comportement au niveau du modèle.

III.4. Quelques extensions

Pour achever cette partie, examinons quelques cas particuliers afin de permettre au lecteur de se forger une vision globale de la caractérisation. Nous n'avons décrit précédemment que le cas le plus simple ; bien d'autres structures de modèles peuvent faire l'objet d'une caractérisation.

Caractérisation dans le cas d'une connaissance a priori

La structure du modèle (3) ne permet pas de prendre en compte une connaissance a priori sur les incertitudes. Supposons, à titre d'exemple, que nous sachions qu'il existe une incertitude relative portant sur la première variable de θ_k . Dans ce cas, le modèle (3) devient :

$$\forall k \in \{1, \dots, m\}, \tilde{y}_k = f(\tilde{x}_k, \rho_k) \theta_k + e_k : f(\tilde{x}_k, \rho_k) = \left[\tilde{x}_{1k} (1 + \rho_k) \quad \tilde{x}_{2k} \quad \dots \quad \tilde{x}_{nk} \right] \quad (12)$$

avec $\hat{y}_k \in \mathbb{R}, e_k \in \mathbb{R}, \theta_k \in \mathbb{R}^n, \mathcal{A}(\rho_k) = [-\delta_\rho, \delta_\rho], \mathcal{A}(e_k) = \mathcal{E} = [-\delta_\varepsilon, \delta_\varepsilon]$ et $\mathcal{A}(\theta_k) = \Theta$

Dans ces circonstances, l'algorithme de caractérisation s'adapte en s'inspirant des algorithmes d'identification des bornes des paramètres de modèles à erreurs dans les variables¹¹ proposés par Cerone [Cerone,93] et par Veres et Norton [Veres et al,96]. Il est clair que la présence d'incertitudes dans la forme linéaire $f(\tilde{x}_k, \rho_k)$ rend la détermination des domaines de majoration plus complexe. Les domaines élémentaires ont pour expression (cf. (37)) :

¹¹ Errors-in-Variables Models ou EIV models en anglais

$$\begin{aligned}\Theta_k^+ &= \left\{ \theta_k : \tilde{y}_k - \delta_\varepsilon \leq \tilde{x}_{1k}(1 + \rho_k)\theta_{1k} + \tilde{x}_{2k}\theta_{2k} + \dots + \tilde{x}_{nk}\theta_{nk} \right\} \\ \Theta_k^- &= \left\{ \theta_k : \tilde{y}_k + \delta_\varepsilon \geq \tilde{x}_{1k}(1 + \rho_k)\theta_{1k} + \tilde{x}_{2k}\theta_{2k} + \dots + \tilde{x}_{nk}\theta_{nk} \right\}\end{aligned}$$

Pour simplifier le calcul de ces domaines, fixons θ_k et, en utilisant (9b) et (10) du chapitre 2, calculons l'espace abstrait suivant :

$$\mathcal{A}(\tilde{x}_{1k}(1 + \rho_k)\theta_{1k} + \tilde{x}_{2k}\theta_{2k} + \dots + \tilde{x}_{nk}\theta_{nk}) = \left[\left(\tilde{x}_k - \Delta_k(\sigma_{\theta_{1k}}) \right) \theta_k \quad \left(\tilde{x}_k + \Delta_k(\sigma_{\theta_{1k}}) \right) \theta_k \right]$$

où $\Delta_k(\sigma_{\theta_{1k}}) = [\sigma_{\theta_{1k}} | \tilde{x}_{1k} | \delta_p 0 \dots 0]$

Les domaines de majoration élémentaires deviennent alors :

$$\begin{aligned}\Theta_k^+ &= \left\{ \theta_k : \tilde{y}_k - \delta_\varepsilon \leq \left(\tilde{x}_k - \Delta_k(\sigma_{\theta_{1k}}) \right) \theta_k \right\} \\ \Theta_k^- &= \left\{ \theta_k : \tilde{y}_k + \delta_\varepsilon \geq \left(\tilde{x}_k + \Delta_k(\sigma_{\theta_{1k}}) \right) \theta_k \right\}\end{aligned} \tag{13}$$

Ils s'apparentent aux domaines élémentaires (9b) mais les formes linéaires $\tilde{x}_k - \Delta_k(\sigma_{\theta_{1k}})$ et $\tilde{x}_k + \Delta_k(\sigma_{\theta_{1k}})$ restent fonction du signe ou, dans le cas général, des signes des variables bornées de θ_k . Pour éluder cette difficulté, Norton propose, dans le cas général, d'effectuer le calcul sur chacun des orthants¹² du domaine paramétrique. Dans la situation la plus générale où les incertitudes connues a priori ne portent pas sur toutes les composantes de θ_k , nous décomposerons le domaine d'étude \mathcal{D} en domaines d'étude élémentaires $\mathcal{D}_{(i)}$ tels que les signes intervenant dans (13) soient invariants et que $\mathcal{D} = \bigcup_i \mathcal{D}_{(i)}$.

En appliquant la même méthode que celle décrite dans le paragraphe V.2, nous pouvons calculer les domaines élémentaires associés à chaque domaine $\mathcal{D}_{(i)}$; il vient :

$$\Theta_{(i)}^+ = \mathcal{D}_{(i)} \cap \bigcap_{k=1}^m \Theta_{(i)k}^+ \quad \text{et} \quad \Theta_{(i)}^- = \mathcal{D}_{(i)} \cap \bigcap_{k=1}^m \Theta_{(i)k}^- \tag{14}$$

où $\Theta_{(i)k}^+$ et $\Theta_{(i)k}^-$ correspondent aux domaines (13) lorsque les signes de θ_k sont ceux du domaine $\mathcal{D}_{(i)}$

Enfin, les domaines de majoration s'obtiennent à partir de :

$$\Theta^+ = \bigcup_i \Theta_{(i)}^+ \quad \text{et} \quad \Theta^- = \bigcup_i \Theta_{(i)}^- \tag{15}$$

Plus pratiquement, souvenons-nous que le calcul d'un domaine $\Theta_{(i)}^+$ ou $\Theta_{(i)}^-$ conduit à un polytope fini repéré par ses sommets. La réunion (15) conduisant aux domaines de majoration Θ^+ ou Θ^- finaux s'obtient en réunissant les sommets obtenus pour chacun des domaines $\Theta_{(i)}^+$ ou $\Theta_{(i)}^-$, en veillant toutefois à ce que chaque sommet n'apparaisse qu'une seule fois car les

¹² Domaine d'un espace à n dimensions où les signes de chacun des éléments des n -uplets qu'il contient sont constants : il en existe 2^n .

jonctions entre plusieurs domaines $\Theta_{(i)}^+$ ou $\Theta_{(i)}^-$ conduisent à des sommets identiques.

Cas des modèles à horizon infini

Le modèle que nous avons envisagé de caractériser (cf. (3)) s'étend aux systèmes dynamiques à horizon fini. Comparons les modèles aux horizons, respectivement, fini et infini suivants :

$$(M_f) \hat{y}_{k+1} = \alpha_k \tilde{y}_k + \beta_k \tilde{u}_k + e_k \quad \text{et} \quad (M_i) \begin{cases} \hat{y}_{k+1} = \alpha'_k \hat{y}_k + \beta'_k \tilde{u}_k \\ \tilde{y}_k = \hat{y}_k + \varepsilon_k \end{cases} ; \forall k \in \{1, \dots, m\}$$

où $\alpha_k, \beta_k, \alpha'_k$ et β'_k contiennent des variables abstraites bornées
et $\forall i \in \{1, \dots, m\}, \mathcal{A}([\alpha_i, \beta_i]^T) = \mathcal{D}$ et $\mathcal{A}([\alpha'_i, \beta'_i]^T) = \mathcal{D}'$

Dans le cas d'un horizon fini, les domaines de majoration élémentaires s'apparenteront à (9b) :

$$\text{et} \quad \begin{cases} \Theta_k^+ = \{(\alpha_k, \beta_k) : \tilde{y}_{k+1} - \delta_\varepsilon \leq \alpha_k \tilde{y}_k + \beta_k \tilde{u}_k\} \\ \Theta_k^- = \{(\alpha_k, \beta_k) : \tilde{y}_{k+1} + \delta_\varepsilon \geq \alpha_k \tilde{y}_k + \beta_k \tilde{u}_k\} \end{cases} \quad (16)$$

Les relations (6) permettent de déduire les domaines de majoration Θ^+ et Θ^- .

tandis que, dans le cas d'un horizon infini, les domaines de majoration sont issus de l'équation

cumulée suivante : $\hat{y}_{k+1} = \sum_{i=1}^k \left(\prod_{j=1}^{k-i} \alpha'_{k-j} \right) \beta'_{i-1} \tilde{u}_{i-1} + \left(\prod_{l=0}^k \alpha'_l \right) \tilde{y}_0$. Ils sont donnés par :

$$\Theta^+ / \forall i \in \{1, \dots, m\}, \begin{bmatrix} \alpha'_i \\ \beta'_i \end{bmatrix} \in \Theta'^+, \left\{ \begin{array}{l} \tilde{y}_1 - \delta'_\varepsilon \leq \beta'_0 \tilde{u}_0 + \alpha'_0 \tilde{y}_0 \\ \vdots \\ \tilde{y}_{k+1} - \delta'_\varepsilon \leq \sum_{i=1}^k \left(\prod_{j=1}^{k-i} \alpha'_{k-j} \right) \beta'_{i-1} \tilde{u}_{i-1} + \left(\prod_{l=0}^k \alpha'_l \right) \tilde{y}_0 \\ \vdots \\ \tilde{y}_m - \delta'_\varepsilon \leq \sum_{i=1}^{m-1} \left(\prod_{j=1}^{m-i-1} \alpha'_{m-j-1} \right) \beta'_{i-1} \tilde{u}_{i-1} + \left(\prod_{l=0}^{m-1} \alpha'_l \right) \tilde{y}_0 \end{array} \right\} \quad (17a)$$

$$\Theta^- / \forall i \in \{1, \dots, m\}, \begin{bmatrix} \alpha'_i \\ \beta'_i \end{bmatrix} \in \Theta'^-, \left\{ \begin{array}{l} \tilde{y}_1 + \delta'_\varepsilon \geq \beta'_0 \tilde{u}_0 + \alpha'_0 \tilde{y}_0 \\ \vdots \\ \tilde{y}_{k+1} + \delta'_\varepsilon \geq \sum_{i=1}^k \left(\prod_{j=1}^{k-i} \alpha'_{k-j} \right) \beta'_{i-1} \tilde{u}_{i-1} + \left(\prod_{l=0}^k \alpha'_l \right) \tilde{y}_0 \\ \vdots \\ \tilde{y}_m + \delta'_\varepsilon \geq \sum_{i=1}^{m-1} \left(\prod_{j=1}^{m-i-1} \alpha'_{m-j-1} \right) \beta'_{i-1} \tilde{u}_{i-1} + \left(\prod_{l=0}^{m-1} \alpha'_l \right) \tilde{y}_0 \end{array} \right\} \quad (17b)$$

Il apparaît nettement que le traitement des contraintes (16) est très différent de celui des contraintes (17) ; en effet, dans le cas du modèle à horizon fini, le calcul des domaines de majoration se décompose en $2m$ calculs de domaines élémentaires, tandis qu'à partir du modèle à horizon infini, toutes les observations de l'instant 0 à l'instant m interviennent dans la même équation qui doit être résolue d'un bloc. Ceci étaye ce que nous nous sommes

précédemment permis d'annoncer : la caractérisation de modèles à horizon infini ne peut pas être décomposée en problèmes élémentaires. En se souvenant que les observations sont généralement effectuées sur un nombre important d'instantns consécutifs, nous conviendrons que la résolution de (17) est extrêmement complexe, pour ne pas dire inextricable.

En imaginant que les deux modèles (M_j) et (M_i) se rapportent à un même système physique, une question importante se pose : est-ce que les résultats de caractérisation obtenus avec les modèles à horizon fini et infini sont semblables ? Ou plutôt, est-il possible de déduire les caractéristiques des variables bornées du modèle à horizon infini, des résultats de caractérisation obtenus du modèle à horizon fini ?

Supposons que les domaines Θ^+ et Θ^- aient été déduits de (16) ; il s'ensuit naturellement que les variables bornées appartenant à ces domaines vérifient toutes les équations (16). En éliminant les mesures \tilde{y}_k , à l'exception de $k=0$, entre ces équations, nous établissons ce qui suit :

$$\forall i \in \{1, \dots, m\}, \begin{bmatrix} \alpha_i \\ \beta_i \end{bmatrix} \in \Theta^+ \Rightarrow \left\{ \begin{array}{l} \tilde{y}_1 - \delta_\varepsilon \leq \beta_0 \tilde{u}_0 + \alpha_0 \tilde{y}_0 \\ \vdots \\ \tilde{y}_{k+1} - \left(1 + \sum_{i=1}^k \left(\prod_{j=1}^i \alpha_{k-j+1} \right) \right) \delta_\varepsilon \leq \sum_{i=1}^k \left(\prod_{j=1}^{k-i} \alpha_{k-j} \right) \beta_{i-1} \tilde{u}_{i-1} + \left(\prod_{l=0}^k \alpha_l \right) \tilde{y}_0 \\ \vdots \\ \tilde{y}_m - \left(1 + \sum_{i=1}^{m-1} \left(\prod_{j=1}^i \alpha_{m-j} \right) \right) \delta_\varepsilon \leq \sum_{i=1}^{m-1} \left(\prod_{j=1}^{m-i-1} \alpha_{m-j-1} \right) \beta_{i-1} \tilde{u}_{i-1} + \left(\prod_{l=0}^{m-1} \alpha_l \right) \tilde{y}_0 \end{array} \right. \quad (18a)$$

$$\forall i \in \{1, \dots, m\}, \begin{bmatrix} \alpha_i \\ \beta_i \end{bmatrix} \in \Theta^- \Rightarrow \left\{ \begin{array}{l} \tilde{y}_1 + \delta_\varepsilon \geq \beta_0 \tilde{u}_0 + \alpha_0 \tilde{y}_0 \\ \vdots \\ \tilde{y}_{k+1} + \left(1 + \sum_{i=1}^k \left(\prod_{j=1}^i \alpha_{k-j+1} \right) \right) \delta_\varepsilon \geq \sum_{i=1}^k \left(\prod_{j=1}^{k-i} \alpha_{k-j} \right) \beta_{i-1} \tilde{u}_{i-1} + \left(\prod_{l=0}^k \alpha_l \right) \tilde{y}_0 \\ \vdots \\ \tilde{y}_m + \left(1 + \sum_{i=1}^{m-1} \left(\prod_{j=1}^i \alpha_{m-j} \right) \right) \delta_\varepsilon \geq \sum_{i=1}^{m-1} \left(\prod_{j=1}^{m-i-1} \alpha_{m-j-1} \right) \beta_{i-1} \tilde{u}_{i-1} + \left(\prod_{l=0}^{m-1} \alpha_l \right) \tilde{y}_0 \end{array} \right. \quad (18b)$$

Du rapprochement entre les relations (18a) et (17a) et des relations (18b) et (17b), nous déduisons que si Θ^- et Θ^+ ont été calculés, alors $\Theta^- \subset \Theta^-$ et $\Theta^+ \subset \Theta^+$ si :

$$\delta'_\varepsilon \geq \sup_{\substack{k \in \{1, \dots, m\} \\ \alpha_i \in \Theta^+ \cup \Theta^-}} \sum_{i=0}^k \left(\prod_{j=1}^i \alpha_{k-j} \right) \delta_\varepsilon \quad (19)$$

Outre le fait qu'il est difficile d'évaluer le résultat (17), rien ne garantit le conservatisme du résultat. Nous concluons donc que la caractérisation de modèles à horizon fini est la seule envisageable avec réalisme et que, malheureusement, les résultats obtenus ne se transposent pas aux caractéristiques des incertitudes de modèles à horizon infini de même structure.

Cas multivariable

Examinons le cas d'un modèle multivariable affine en les variables abstraites bornées. Le modèle (3) devient alors :

$$\forall k \in \{1, \dots, m\}, \tilde{y}_k = \tilde{x}_k \theta_k + e_k : \forall i \in \{1, \dots, p\}, \mathcal{A}(e_{i,k}) = \mathcal{E}_i = [-\delta_{ei} \ \delta_{ei}] \text{ et } \mathcal{A}(\theta_k) = \Theta$$

où $\tilde{x}_k \in \mathbb{R}^{p \times m}$, $\hat{y}_k \in \mathbb{R}^p$, $e_k \in \mathbb{R}^p$ et $\theta_k \in \mathbb{R}^m$

Il s'ensuit alors que les domaines de majoration élémentaires s'écrivent :

$$\Theta_k^+ = \{ \theta_k : \forall i \in \{1, \dots, p\}, \tilde{y}_{i,k} - \delta_{ei} \leq \tilde{x}_{k,i} \theta_k \}$$

$$\Theta_k^- = \{ \theta_k : \forall i \in \{1, \dots, p\}, \tilde{y}_{i,k} + \delta_{ei} \geq \tilde{x}_{k,i} \theta_k \}$$

Les domaines de majoration élémentaire peuvent encore être décomposés en posant

$$\Theta_k^+ = \bigcap_{i=1}^p \Theta_{k,i}^+ \text{ et } \Theta_k^- = \bigcap_{i=1}^p \Theta_{k,i}^- \text{ avec :}$$

$$\Theta_{k,i}^+ = \{ \theta_k : \tilde{y}_{i,k} - \delta_{ei} \leq \tilde{x}_{k,i} \theta_k \} \text{ et } \Theta_{k,i}^- = \{ \theta_k : \tilde{y}_{i,k} + \delta_{ei} \geq \tilde{x}_{k,i} \theta_k \}.$$

Dès lors, la méthode de caractérisation est conforme à celle décrite en III.2, mais au lieu de m domaines élémentaires de dimension 1, cela revient à en considérer $m \times p$. Il faut cependant souligner que cette structure de modèle ne permet pas de tenir compte de dépendance entre variables abstraites bornées à différents instants : cette situation est typique des modèles de systèmes dynamiques. Nous ne nous appesantirons guère plus sur la caractérisation car cet aspect n'est qu'une des étapes possibles vers le diagnostic. L'essentiel est de retenir que les bornes des variables abstraites caractérisant des incertitudes de comportement peuvent souvent, même si elles ne peuvent pas être déduites d'un savoir scientifique ou d'expertise, être déterminées par des algorithmes de caractérisation.

IV. Conclusion

Ce chapitre nous a permis d'appréhender une étape fondamentale du diagnostic des systèmes physiques : la représentation de l'incertain. Nous avons montré qu'à quelques exceptions près, les procédures de diagnostic ne pouvaient reposer que sur des métamodèles incertains

contenant exclusivement des variables abstraites bornées. Nous avons étudié les différents moyens permettant d'établir des métamodèles incertains. La connaissance de certaines structures d'incertitudes permet de partir de métamodèles certains et, après avoir empiriquement déterminé la structure adaptée, de rechercher les caractéristiques des variables abstraites bornées qui expliquent un ensemble de comportement possible ; c'est ce que nous avons baptisé la *sculpture de métamodèle*. Nous avons introduit la notion de *caractérisation* qui est aux variables abstraites dépendant du temps ce que l'identification est aux variables invariantes. Puis, nous avons développé une approche systématique permettant de caractériser les variables abstraites bornées dans certaines structures de métamodèles incertains. Nous avons constaté que si la caractérisation s'avérait simple dans le cas des hypomodèles à horizon fini, elle le devenait beaucoup moins dans le cas d'horizon infini.

Nous avons souligné le fait qu'il était toujours possible d'obtenir des métamodèles justes, même si leur précision pouvait être médiocre ; cette constatation ouvre de nombreux horizons concernant la simplification de structures de métamodèles incertains. Enfin, nous avons montré que la nature temporelle des variables abstraites bornées ne devait pas être omise sous peine d'aboutir à des erreurs de représentation. Ce résultat est très important pour le chapitre suivant qui aborde la dernière étape théorique menant à la synthèse de procédures de diagnostic : l'exploitation de la connaissance contenue dans un métamodèle. C'est le problème de la synthèse d'hypomodèles.

Chapitre 4

les hypomodèles

TABLE DES MATIÈRES

I. Introduction..... 135

II. Les hypomodèles certains..... 136

II.1. Formulation générique des hypomodèles certains138

Les hypomodèles certains à horizon infini 138

Les hypomodèles certains à horizon fini 142

II.2. Propriétés génériques des hypomodèles certains146

Les hypomodèles certains à horizon infini 146

Les hypomodèles certains à horizon fini 148

II.3. Conditions d'existence des hypomodèles certains 152

Les hypomodèles certains à horizon infini 152

Les hypomodèles certains à horizon fini 156

II.4. Spécificités des hypomodèles certains dans le cadre du diagnostic..... 159

Le problème du découplage en diagnostic..... 159

Influence de l'horizon des hypomodèles en diagnostic..... 161

III. Les hypomodèles incertains 167

III.1. Formulation générique des métamodèles dynamiques incertains 168

III.2. Les hypomodèles incertains à horizon infini 174

Formulation générique 174

L'effet d'enveloppement 176

Calcul d'enveloppe par l'arithmétique des intervalles..... 179

Calcul d'enveloppe par balayage 181

Prise en compte de l'équation de sortie..... 184

Structures purement additives 187

Approches elliptiques 190

III.3. Les hypomodèles incertains à horizon fini 192

Formulation générique 193

Du métamodèle dynamique à l'hypomodèle à horizon fini..... 199

IV. Conclusion 208

I. Introduction

Après avoir circonscrit les principes fondamentaux du diagnostic, puis étudié les propriétés des variables abstraites bornées quant à la représentation des incertitudes de comportement, il subsiste le problème de savoir comment ces principes peuvent s'appliquer à la réalisation d'un système automatique de diagnostic de systèmes physiques. Au cours du chapitre introductif, nous avons établi une hiérarchie des modèles ; au niveau le plus élémentaire apparaît la variable physique suivie de la description comportementale impliquant des métamodèles. Ils synthétisent une connaissance phénoménologique et font intervenir des variables abstraites. Les chapitres précédents étaient tout entier consacrés aux variables abstraites intervenant dans la représentation de l'incertain, et par conséquent, aux éléments abstraits constitutifs des métamodèles. Nous n'avons certes pas étendu nos investigations à des cas concrets car, plutôt que de cumuler des exemples succincts toujours spécifiques qui auraient bien plus certainement noyé l'essence de notre propos plutôt que de l'étayer, nous avons préféré en traiter un complètement dans le dernier chapitre. Conformément à la hiérarchie des modèles du chapitre introductif, il est naturel de dédier ce chapitre au dernier niveau de représentation comportementale, c'est-à-dire aux hypomodèles qui, rappelons le, représentent une fraction vérifiable du comportement d'un système physique. Le métamodèle du système physique à diagnostiquer sera ici supposé connu et seule la manière d'en extraire des hypomodèles sera traitée.

Avant d'aborder les hypomodèles associés aux systèmes physiques incertains, nous reviendrons sur le cas des systèmes certains pour mettre en évidence leurs propriétés. Bien que la démarche proposée admette des extensions, nous nous sommes restreints aux hypomodèles se fondant sur des métamodèles linéaires, invariants et pouvant se mettre sous la forme d'une représentation d'état. Nous avons développé des approches génériques, l'une concerne les hypomodèles causaux à horizon infini, tandis que l'autre s'applique aux hypomodèles à horizon fini, qu'ils soient causaux ou non. Les approches se décomposent en trois phases. Nous avons commencé par rechercher des formalismes réunissant en une même écriture, l'ensemble des hypomodèles exploitables par une procédure de diagnostic pour déterminer, dans un second temps, les propriétés génériques que doivent satisfaire ces formalismes pour mériter l'appellation d'hypomodèle. Puis, pour mieux comprendre les différences entre les hypomodèles à horizon infini et ceux à horizon fini, nous avons recherché les conditions d'existence respectives de ces hypomodèles. Grâce aux résultats obtenus, nous conforterons les résultats décrits dans le premier chapitre concernant l'inanité, pour le diagnostic, des approches robustes par découplage en nous fondant sur un développement mathématique. Nous justifierons le classement introduit dans le premier chapitre en étudiant l'influence de la causalité et de la taille de l'horizon.

La seconde partie aborde le contexte incertain. Après avoir recherché une formulation générique exploitable des métamodèles incertains, nous aborderons, dans une seconde sous-partie, les hypomodèles incertains à horizon infini. Nous mettrons en évidence l'effet d'enveloppement, source de nombreuses limitations, puis nous montrerons comment l'arithmétique des intervalles peut être appliquée au calcul d'enveloppe. Nous proposerons une méthode générale d'intégration dans un cadre incertain et nous nous attarderons sur ses limites. Nous verrons qu'il existe plusieurs façons de prendre en compte une équation de sortie dans un modèle d'état. Enfin, nous examinerons le cas particulier des structures d'incertitudes purement additives puis nous débattrons de l'intérêt des approches elliptiques. La troisième sous-partie examine les moyens de réaliser des hypomodèles à horizon fini. Nous commencerons par étudier les différentes manières de procéder à des tests de cohérence lorsque la formulation d'un hypomodèle à horizon fini est connue. Dans un second temps, nous rechercherons le moyen d'obtenir une formulation simple d'hypomodèle à horizon fini et partant d'un métamodèle sous forme d'équations différentielles ordinaires. Nous verrons que cette opération, triviale dans un cas certain, l'est beaucoup moins dans un contexte incertain.

II. Les hypomodèles certains

Cette première partie examine les problèmes classiques de la littérature scientifique eu égard aux principes énoncés dans le premier chapitre. Nous étudierons les moyens d'obtenir des hypomodèles en partant de métamodèles certains, afin de faire ressortir des propriétés communes avec les hypomodèles incertains. L'objectif est de comprendre la différence, pour le diagnostic, entre les différents types d'hypomodèles. Quel horizon faut-il préférer pour un usage au diagnostic ? quelle dimension ? quel réglage ? Nous nous restreindrons au cas particulier où le métamodèle est un système d'équations d'état linéaires et invariantes ; les variables abstraites du métamodèle sont déterministes et contenues dans les matrices d'état A_c , B_c , C et D_c :

$$\begin{cases} \frac{dx_0(t)}{dt} = A_c x_0(t) + B_c u(t) \\ y(t) = C x_0(t) + D_c u(t) \end{cases}$$

où C est restreinte à une matrice de plein rang ligne pour éviter les redondances dans le modèle.

Pour rester conforme aux notations introduites au cours du chapitre 1 §IV.2.c., il aurait respectivement fallu substituer $[x_0(t)]$, $[u(t)]$ et $[y(t)]$ aux notations $x_0(t)$, $u(t)$ et $y(t)$ et ajouter les relations suivantes :

$$\begin{cases} [y(t)] = \tilde{y}(t) \\ [u(t)] = \tilde{u}(t) \end{cases}$$

Nous nous sommes permis de faire ici l'économie de ces notations afin d'alléger les écritures ; en effet, ce chapitre n'examine essentiellement que ce type de métamodèle. Il faut

néanmoins entendre ces relations telles que l'écriture précise les décrit ; en cas de nécessité, nous nous permettrons de revenir à ces notations.

Le vecteur $x_0(t)$, de dimension n , contient, dans la plupart des modèles de connaissance de ce type, des variables physiques potentiellement connues. Celles qui sont observées apparaissent dans les vecteurs $u(t)$ et $y(t)$, respectivement de dimension m et p .

En notant $u_k = u(k\tau)$ et $y_k = y(k\tau)$ ($\forall k \in \mathbb{N}$), le métamodèle en temps continu peut être transposé en temps discret¹ sous la forme suivante :

$$\begin{cases} x_{k+1} = Ax_k + Bu_k \\ y_k = Cx_k + Du_k \end{cases} \quad (1)$$

Lorsque les signaux $u(\cdot)$ peuvent être assimilés, pour la période d'échantillonnage τ , à des signaux constants par morceaux, tels que :

$$\forall t \in [k\tau, (k+1)\tau], u(t) = u(k\tau); k \in \mathbb{N} \text{ et } \tau \in \mathbb{R}^{*+} \quad (2a)$$

alors les matrices du modèle discret (1), pour lequel $x_k = x_0(k\tau)$, correspondant au modèle en temps continu, sont données par :

$$A = \sum_{i=0}^{\infty} \frac{A_c^i \tau^i}{i!}, B = \sum_{i=0}^{\infty} \frac{A_c^i \tau^{i+1}}{(i+1)!} B_c \text{ et } D = D_c \quad (2b)^2$$

Cependant, si ce type de signaux est fréquent en commande automatisée du fait de la présence physique de bloqueurs d'ordre zéro, en diagnostic, il n'y a aucune raison pour que les observations apparaissant dans le vecteur $u(t)$ soient constantes par morceaux. Le cas échéant, il est préférable de modifier la représentation du signal $u(\cdot)$ en substituant à (2a) le modèle suivant :

$$\forall t \in [k\tau, (k+1)\tau], u(t) = (u((k+1)\tau) - u(k\tau)) \frac{t - k\tau}{\tau} + u(k\tau); k \in \mathbb{N} \text{ et } \tau \in \mathbb{R}^{*+} \quad (3a)^3$$

Avec cette hypothèse, l'intégration sur une période τ du métamodèle en temps continu conduit à un modèle en temps discret (1) dont les matrices sont les suivantes :

$$A = \sum_{i=0}^{\infty} \frac{A_c^i \tau^i}{i!}, B = \sum_{i=0}^{\infty} \frac{(2^{i+2} - 2) A_c^i \tau^{i+1}}{(i+2)!} B_c \text{ et } D = C \sum_{i=0}^{\infty} \frac{A_c^i \tau^{i+1}}{(i+2)!} B_c + D_c \quad (3b)^4$$

et pour lequel :

¹ Cette transposition est parfaitement justifiée par le fait que les observations sont, en général, régulièrement distribuées dans le temps.

² Si A_c est inversible, $A = e^{A_c \tau}$ et $B = (e^{A_c \tau} - I_n) A_c^{-1} B_c$.

³ Analogie avec un bloqueur d'ordre 1.

⁴ Si A_c est inversible, $A = e^{A_c \tau}$, $B = \frac{(e^{A_c \tau} - I_n)^2}{\tau} A_c^{-2} B_c$ et $D = C \left(\frac{(e^{A_c \tau} - I_n) A_c^{-1}}{\tau} - I_n \right) A_c^{-1} B_c + D_c$.

$$x_k = x_0(k\tau) - \sum_{i=0}^{\infty} \frac{A_c^i \tau^{i+1}}{(i+2)!} B_c u(k\tau)$$

Même si, dans le second cas, l'état reconstruit ne correspond plus directement à l'état du modèle en temps continu, dans un cas comme dans l'autre, la structure du modèle (I) reste inchangée. Nous avons donc retenu ce métamodèle comme référence dans ce chapitre.

Notion de minimalité

La structure du modèle (I) englobe un grand nombre de représentations. Néanmoins, certaines d'entre elles n'ont guère de sens eu égard à nos desseins. Ainsi, le modèle (I) sera redondant si certaines variables du vecteur d'état x_k demeurent en permanence des combinaisons linéaires d'autres ; la dimension de l'équation récurrente (I) pourrait être réduite sans perte d'information. En considérant une séquence de $r (\in \mathbb{N})$ entrées u_k consécutives et en négligeant l'état initial qui n'apporte rien ici, nous pouvons écrire :

$$x_r = \Lambda_{r-1} \begin{bmatrix} u_0 \\ u_1 \\ \vdots \\ u_{r-2} \\ u_{r-1} \end{bmatrix} \text{ avec } \Lambda_{r-1} = \begin{bmatrix} A^{r-1}B & A^{r-2}B & \dots & AB & B \end{bmatrix}$$

Dès lors, pour éviter que certaines variables du vecteur x_k soient des combinaisons linéaires d'autres, il suffit qu'il existe au moins une valeur de r telle que la matrice Λ_{r-1} soit de plein rang ligne. Le théorème de Cayley-Hamilton permet d'établir que si la matrice Λ_{r-1} n'est pas de plein rang ligne, alors les matrices Λ_r , où $r > r-1$, ne le seront pas non plus, et nous reconnaissons alors la condition nécessaire et suffisante de commandabilité. Il n'y a en effet rien de surprenant à ce qu'aucune des variables du vecteur d'état x_k ne sont combinaison linéaire d'autres soit équivalent à ce que tous les points de l'espace d'état puissent être atteints. Signalons aussi qu'il est possible de montrer que dans le cas où les matrices d'état fluctueraient au cours du temps, la condition énoncée s'apparente aussi à la notion de non-commandabilité [Vidyasagar,93]. Le métamodèle (I) sera minimal s'il est commandable et pourvu d'une matrice C de plein rang ligne. Nous montrerons dans la sous-partie suivante que tout métamodèle sans pôles non commandables instables, est équivalent, en régime établi, à un métamodèle minimal.

Nous ne nous intéresserons qu'aux métamodèles (I) minimaux.

II.1. Formulation générique des hypomodèles certains

Les hypomodèles certains à horizon infini

Les hypomodèles certains à horizon infini regroupent autant les observateurs d'état à horizon infini tels qu'ils ont été décrits par Luenberger [Luenberger,71] que les modèles parallèles (cf. chapitre 1, paragraphe IV.1 *Nature des systèmes incertains* Tableau IV) issus directement de (I) :

$$\begin{cases} \hat{x}_{k+1} = A\hat{x}_k + Bu_k \\ y_k = C\hat{x}_k + Du_k \end{cases} \quad (4)$$

Notons ici l'intérêt de la dénomination que nous avons introduite. Pourquoi, en effet, ne pas avoir simplement employé le terme d'*observateur d'état* dont le mérite est d'avoir été utilisé depuis fort longtemps et qui englobe à la limite le modèle parallèle (4) ? Il faut remarquer que, depuis les années 1960 où l'expression vit le jour, elle fut étendue à des entités s'éloignant grandement des descriptions initiales de Luenberger, notamment avec l'apparition des *observateurs d'état à mémoire finie* [Medvedev et al,91]. La dénomination que nous avons introduite nous permet de sortir du dilemme de ce qui doit ou ne doit pas être baptisé *observateur d'état*, et nous permet de rassembler sous un même vocable les entités aux mêmes propriétés.

L'objectif de cette sous-partie est de réunir l'ensemble des hypomodèles certains à horizon infini associés au métamodèle minimal (I) en une même écriture. Des travaux relativement récents, [Goodwin et al,89] et [Magni et al,91], ont proposé certaines solutions ; néanmoins, l'écriture proposée nous semble recouvrir un domaine bien trop large comparé à la classe d'hypomodèles que nous souhaitons représenter. L'approche adoptée repose sur la représentation sous forme de fonction de transfert qui ne permet malheureusement pas de mettre en évidence des propriétés communes sous une forme algébrique simple. Pour parvenir à établir une écriture générique, nous avons opté pour l'usage des matrices complémentaires déjà utilisées dans [Basile et al,92]. Il permet en effet d'ordonner l'information dans un système d'équations.

matrice complémentaire canonique

Soit $\mathfrak{M} : x \mapsto Mx$, une application linéaire de \mathbb{R}^s dans \mathbb{R}^r avec $\forall (r,s) \in \mathbb{N}^2 / s \geq r$ et M , une matrice de plein rang ligne. Alors l'application complémentaire $\overline{\mathfrak{M}} : x \mapsto \overline{M}x$ de \mathbb{R}^s dans \mathbb{R}^{s-r} vérifie $im(\mathfrak{M}) \cup im(\overline{\mathfrak{M}}) = \mathbb{R}^s$ et $im(\mathfrak{M}) \cap im(\overline{\mathfrak{M}}) = \{0\}$. De plus, la complémentarité est qualifiée de canonique lorsque la matrice \overline{M} vérifie $\overline{M}M^T = I_{s-r}$. En conséquence, la matrice complémentaire canonique \overline{M} est parfaitement déterminée par la relation suivante :

$$\begin{bmatrix} M \\ \overline{M} \end{bmatrix}^{-1} = \begin{bmatrix} M^+ & \overline{M}^T \end{bmatrix} \quad (5)$$

où M^+ est l'inverse à droite de M défini par $M^+ = M^T(MM^T)^{-1}$; dans le cas particulier où M est inversible, M^+ devient M^{-1} et \overline{M} disparaît.

Pour toute matrice M de plein rang ligne, la matrice complémentaire canonique \overline{M} peut être obtenue au moyen d'une décomposition en valeurs singulières [Golub,89] de $M : M = U[\Sigma \ 0]V^T$ où U et V sont des matrices orthogonales⁵. En

⁵ Une matrice orthogonale U de dimension s vérifie par définition $U^T U = U U^T = I_s$.

choisissant $\bar{M} = [0 \ I_{r-s}]V^T$, la propriété (5) est bien vérifiée puisque $M^+ = V[\Sigma \ 0]^T U^T$; il s'ensuit que :

$$\begin{bmatrix} M \\ \bar{M} \end{bmatrix}^{-1} = \left(\begin{bmatrix} U & 0 \\ 0 & I_{r-s} \end{bmatrix} \begin{bmatrix} \Sigma & 0 \\ 0 & I_{r-s} \end{bmatrix} V^T \right)^{-1} = V \begin{bmatrix} \Sigma U^T & 0 \\ 0 & I_{r-s} \end{bmatrix} = \begin{bmatrix} M^+ \\ \bar{M}^T \end{bmatrix}$$

Cependant, dans de nombreux cas, cette décomposition est inutile car la déduction d'une matrice complémentaire canonique est évidente.

En corollaire de (5), nous pouvons énoncer les propriétés suivantes :

$$\begin{bmatrix} M^+ & \bar{M}^T \end{bmatrix} \begin{bmatrix} M \\ \bar{M} \end{bmatrix} = I_s \Leftrightarrow M^+ M + \bar{M}^T \bar{M} = I_s$$

et

$$\begin{bmatrix} M \\ \bar{M} \end{bmatrix} \begin{bmatrix} M^+ & \bar{M}^T \end{bmatrix} = I_s \Leftrightarrow \begin{cases} MM^+ = I_r & \bar{M}\bar{M}^T = I_{s-r} \\ M\bar{M}^T = 0 & \bar{M}M^+ = 0 \end{cases}$$

Un hypomodèle à horizon infini extrait du métamodèle (1) est par essence un système dynamique ; il repose donc sur une équation récurrente dont les entrées sont les variables physiques observées. Nous supposons que, parmi les entrées \tilde{u}_k , seule la fraction de dimension q : $S\tilde{u}_k$, où S est une matrice de sélection⁶, est effectivement observée. Toutes les composantes apparaissant dans les vecteurs \tilde{y}_k sont supposées observées ; dans le cas contraire, il n'est pas nécessaire d'introduire une matrice supplémentaire puisqu'il suffit de modifier les matrices C et D pour se ramener à une telle situation. Les résultats que nous exposons concernant la réjection de certaines variables de \tilde{u}_k se distinguent des résultats traditionnels concernant les observateurs à entrées inconnues [Hou et al,92], [Hou et al,94] et complétés dans [Darouach,94], dans la mesure où ils s'appliquent principalement à établir les propriétés génériques des hypomodèles faisant notamment l'objet d'un découplage parfait plutôt que de proposer une procédure de synthèse [Ploix et al,99]. L'hypomodèle repose donc sur :

$$\hat{z}_{k+1} = N\hat{z}_k + G\tilde{u}_k + L\tilde{y}_k \quad (6)$$

Les vecteurs reconstruits \hat{z}_k sont de dimension l ($\in \mathbf{N}^*$), a priori différente de la dimension n de x_k . Un hypomodèle à horizon infini a pour vocation de reconstruire des combinaisons linéaires de variables physiques. Pour demeurer fidèle à notre objectif de diagnostic, nous nous restreindrons aux hypomodèles reconstruisant des combinaisons linéaires des variables

⁶ Une matrice de sélection S , est, par définition, une matrice de plein rang ligne constituée de 0 et de 1 dans laquelle la valeur 1 n'apparaît qu'une seule fois par ligne. Il s'ensuit naturellement que $SS^T = I_{dim(S)}$.

physiques observées⁷, qui ne peuvent appartenir, pour des raisons de causalité, qu'aux variables du vecteur \tilde{y}_k .

L'objectif est donc de reconstruire certaines composantes du vecteur d'observations \tilde{y}_k que nous rassemblons dans le vecteur $R\tilde{y}_k$ de dimension s , où R est une matrice de plein rang ligne. Pour y parvenir, nous disposons des observations $S\tilde{u}_k$ ⁸ et de l'hypomodèle (6). Ce modèle permet de reconstruire une fraction Wx_k de dimension r du vecteur x_k , où W est une matrice de plein rang ligne⁹. La fonction de l'hypomodèle est parfaitement définie en le donnant pour capable, lorsque le comportement du processus est décrit par le métamodèle (1), d'annuler l'erreur suivante :

$$\varepsilon_k = H\hat{z}_k - Wx_k \quad (7)$$

où la matrice H est naturellement de plein rang ligne ;
elle représente les composantes de l'hypomodèle
qui reconstruiront la fonctionnel Wx_k

Avant d'aborder la formulation des hypomodèles à horizon fini, intéressons-nous à la fonctionnelle, spécifique au diagnostic, conduisant à la reconstruction de $R\tilde{y}_k$ et précisons que, dans les calculs qui vont suivre, le système physique est supposé se comporter normalement, c'est-à-dire conformément au métamodèle (1). Si le processus associé au métamodèle (1) se comporte normalement le produit $R\tilde{y}_k$ vérifie : $R\tilde{y}_k = RCx_k + RD\tilde{u}_k$. Or, si l'hypomodèle assure sa fonction (Wx_k est parfaitement reconstruit), le produit $R\tilde{y}_k$ doit être reconstruit à partir de Wx_k et des observations $S\tilde{u}_k$. Il s'ensuit qu'il faut que :

$$\begin{aligned} \exists \Phi, \Gamma / R\tilde{y}_k &= RCx_k + RD\tilde{u}_k = \Phi Wx_k + \Gamma S\tilde{u}_k \\ \Leftrightarrow RC &= \Phi W \text{ et } RD = \Gamma S \end{aligned}$$

En multipliant les deux relations obtenues respectivement par \overline{W}^T et par \overline{S}^T , nous obtenons :

$$\begin{cases} RC\overline{W}^T = 0 \\ RD\overline{S}^T = 0 \end{cases}$$

D'autre part, en utilisant les propriétés des matrices complémentaires, la relation $R\tilde{y}_k = RCx_k + RD\tilde{u}_k$ peut se développer comme suit :

$$R\tilde{y}_k = RC \begin{bmatrix} W^+ & \overline{W}^T \end{bmatrix} \begin{bmatrix} W \\ \overline{W} \end{bmatrix} x_k + RD \begin{bmatrix} S^+ & \overline{S}^T \end{bmatrix} \begin{bmatrix} S \\ \overline{S} \end{bmatrix} \tilde{u}_k$$

Or, puisque $RC\overline{W}^T$ et $RD\overline{S}^T$ s'annulent tous les deux, la reconstruction de $R\tilde{y}_k$ sera donc obtenue de :

⁷ Pour permettre la comparaison entre un comportement observé et un comportement prédit.

⁸ Les autres observations ne présentent guère d'intérêt ici.

⁹ Il est absurde de ne pas choisir une matrice de plein rang ligne car cela reviendrait à reconstruire des combinaisons linéaires de variables déjà reconstruites.

$$R\tilde{y}_k = RCW^+H\hat{z}_k + RDS^T S\tilde{u}_k \quad (8a)$$

sous réserve que la propriété spécifique au diagnostic qui suit, soit satisfaite :

$$R[C\bar{W}^T \quad D\bar{S}^T] = 0 \quad (8b)$$

Or, $RC\bar{W}^T = 0 \Leftrightarrow \exists$ une matrice $\chi / RC = \chi W$, nous en déduisons¹⁰ que cette égalité ne sera vérifiée que si $rang([(RC)^T \ W^T]^T) = rang(W)$. En procédant de la même manière pour la seconde relation, nous en déduisons que la relation (8b) sera satisfaite si et seulement si les conditions suivantes sont satisfaites :

Conditions d'existence spécifiques au diagnostic [CID]

$$rang\left(\begin{bmatrix} RC \\ W \end{bmatrix}\right) = dim(Wx_k) = r \quad \text{et} \quad rang\left(\begin{bmatrix} RD \\ S \end{bmatrix}\right) = dim(S\tilde{u}_k) = q$$

Ces conditions garantissent que la fonctionnel (8a) reconstruit bien $R\tilde{y}_k$ lorsque l'hypomodèle, défini par (6) et (7), assure sa fonction. Pour en finir avec la formulation générique, réordonnons le vecteur \hat{z}_k en isolant la partie $H\hat{z}_k$ intervenant directement dans la reconstruction au moyen de la matrice complémentaire canonique \bar{H} :

$$\begin{bmatrix} \hat{w}_k \\ \hat{\xi}_k \end{bmatrix} = \begin{bmatrix} H \\ \bar{H} \end{bmatrix} \hat{z}_k \quad (9)$$

où \hat{w}_k représente la reconstruction de Wx_k
et $\hat{\xi}_k$ la fraction de \hat{z}_k ne participant qu'indirectement à la reconstruction.

Au moyen de ce changement de base, les nouvelles variables peuvent être introduites dans l'équation récurrente (6) et l'hypomodèle admet la formulation suivante :

$$\begin{bmatrix} \hat{w}_{k+1} \\ \hat{\xi}_{k+1} \end{bmatrix} = \begin{bmatrix} HNH^+ & HNH^+ \\ \bar{H}NH^+ & \bar{H}NH^+ \end{bmatrix} \begin{bmatrix} \hat{w}_k \\ \hat{\xi}_k \end{bmatrix} + \begin{bmatrix} HG \\ \bar{H}G \end{bmatrix} \tilde{u}_k + \begin{bmatrix} HL \\ \bar{H}L \end{bmatrix} \tilde{y}_k \quad (10a)$$

$$\varepsilon_k \underset{\text{normalement}}{=} \hat{w}_k - Wx_k \quad (10b)$$

Les hypomodèles certains à horizon fini

La formulation générique des hypomodèles à horizon fini est plus simple que celle des modèles à horizon infini. L'absence d'équation de récurrence ne permet plus de former de boucle de contre-réaction. Un hypomodèle à horizon fini h associé au métamodèle (1) est constitué d'une relation linéaire faisant intervenir les observations disponibles. En reprenant les notations précédentes, autrement dit, en rassemblant les observations dans les vecteurs Su_k

¹⁰ L'équation $AX=B$, où A , B et X sont des matrices quelconques, admet une solution si et seulement si $rang([A \ B]) = rang(A)$. L'ensemble des solutions est donné par : $X=A^+B+(I-A^+A)K$ où K est une matrice quelconque et A^+ le pseudo-inverse de A .

et y_k , selon qu'elles apparaissent en entrée ou en sortie, nous déduisons l'expression générique d'un hypomodèle à horizon fini h :

$$\sum_{i=0}^{h-1} (U_i S u_{k+i} + Y_i y_{k+i}) = U u_k^{(h)} + Y y_k^{(h)} = 0 ; \forall k \in \mathbb{N} \quad (11a)^{11}$$

$$\text{avec } U = [U_0 S \quad \dots \quad U_{h-1} S] \quad Y = [Y_0 \quad \dots \quad Y_{h-1}]$$

$$\text{et } u_k^{(h)\top} = [u_k \quad \dots \quad u_{k+h-1}]^\top, \quad y_k^{(h)\top} = [y_k \quad \dots \quad y_{k+h-1}]^\top$$

Mais cette relation ne suffit pas à la formalisation des objectifs. En effet, prenons l'exemple d'un métamodèle (I) de dimension 2 et dont les matrices D et C sont respectivement nulle et égale à l'identité. En adoptant les notations suivantes :

$$y_k = \begin{bmatrix} y_{1,k} \\ y_{2,k} \end{bmatrix}, \quad A = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \quad \text{et} \quad B = \begin{bmatrix} b_1 \\ b_2 \end{bmatrix}$$

où les minuscules sont, à l'exception de y_b , des scalaires

nous pouvons déduire, de (I), trois hypomodèles minimaux :

$$\begin{aligned} (\mathcal{H}_1) \quad & Ay_k + Bu_k - y_{k+1} = 0 \\ (\mathcal{H}_2) \quad & a_{11}y_{1,k} + a_{12}y_{2,k} + b_1u_k - y_{1,k+1} = 0 \\ (\mathcal{H}_3) \quad & a_{21}y_{1,k} + a_{22}y_{2,k} + b_2u_k - y_{2,k+1} = 0 \end{aligned}$$

Or, il apparaît nettement que l'hypomodèle (\mathcal{H}_1) représente un comportement global tandis que les hypomodèles (\mathcal{H}_2) ou (\mathcal{H}_3) ne représentent que des comportements locaux. Pour traduire ces différences, il est nécessaire de compléter la formulation générique (11a) par la définition d'une matrice W de plein rang ligne spécifiant la partie du métamodèle exploitée par l'hypomodèle. Seule la fraction définie par W de l'équation récurrente du métamodèle (I) sera exploitée :

$$\begin{cases} Wx_{k+1} = WAx_{k+1} + WBu_k \\ y_k = Cx_k + Du_k \end{cases} \quad (11b)$$

Dès lors, la relation (11a) accompagnée de (11b) forment de manière générique tous les hypomodèles à horizon fini h issus de (I).

A propos de la notion de minimalité, il faut noter que la seule connaissance d'un hypomodèle (11a) exploitant la connaissance (11b) pour lequel la matrice Y conduit à une solution minimale¹², suffit, par définition, à générer tous les autres hypomodèles associés aux mêmes objectifs. En notant h^* , l'horizon associé aux matrices Y^* et U^* minimales, et, en utilisant

¹¹ Nous n'avons pas fait apparaître les signes \sim ou \wedge au dessus des variables u_k et y_k car leur position dépend de l'usage réservé à l'hypomodèle, selon qu'il soit acausal ou causal.

¹² Nous développerons ce point dans la sous-partie consacrée aux conditions d'existence des hypomodèles certains à horizon fini.

(11a), il est facile de montrer qu'il existe des matrices carrées régulières P_i et un entier α tels que tout hypomodèle défini par (11a) et (11b) s'écrive :

$$\sum_{i=0}^{\alpha} (P_i U^* \tilde{u}_{k+i}^{(h^*)} + P_i Y^* \tilde{y}_{k+i}^{(h^*)}) = 0 ; \forall k \in \mathbb{N}$$

Il apparaît clairement que l'ensemble des hypomodèles à horizon fini est entièrement généré par (11a) et (11b) sous réserve qu'une condition de minimalité, sur laquelle nous reviendrons ultérieurement, soit satisfaite.

La notion de causalité est peu importante dans le cas des hypomodèles à horizon fini associés à des métamodèles certains. L'exemple suivant expose deux hypomodèles, l'un acausal et l'autre causal. Visiblement, les diagnostics obtenus seront identiques :

$$0.5\tilde{y}_k - \tilde{y}_{k+1} + \tilde{u}_k = 0 \quad \text{et} \quad \begin{cases} \hat{y}_{k+1} = 0.5\tilde{y}_k + \tilde{u}_k \\ \tilde{y}_{k+1} = \hat{y}_{k+1} \end{cases}$$

Remarque

Nous pouvons trouver, dans la littérature, deux grandes approches pour la génération de relations de parité qui ne sont rien moins que des hypomodèles à horizon finis acausaux. La première se fonde sur la représentation fréquentielle [Gertler et al,90], l'autre sur la représentation temporelle [Chow et al,84]. L'exemple suivant montre pourquoi ces approches ne sont pas strictement équivalentes. Considérons le métamodèle suivant :

$$\begin{cases} x_{k+1} = \begin{bmatrix} -0.5 & 1 \\ 0 & 0.8 \end{bmatrix} x_k + \begin{bmatrix} 1 \\ 1.3 \end{bmatrix} \tilde{u}_k \\ \tilde{y}_k = \begin{bmatrix} 1 & 1 \end{bmatrix} x_k \end{cases}$$

Nous déduisons de la fonction de transfert en z existant entre $\tilde{Y}(z)$ et $\tilde{U}(z)$, la relation de parité minimale d'horizon 2 suivante :

$$4\tilde{y}_k - 5\tilde{y}_{k+1} + 11.5\tilde{u}_k = 0$$

L'approche temporelle repose sur la représentation d'état dans laquelle les variables connues ont été séparées des variables inconnues (x_k) :

$$\begin{bmatrix} \tilde{y}_k \\ \vdots \\ \tilde{y}_{k+h-1} \end{bmatrix} - \begin{bmatrix} 0 & \dots & 0 & 0 & 0 \\ CB & \dots & 0 & 0 & 0 \\ \vdots & \dots & CB & 0 & 0 \\ CA^{h-2}B & \dots & CAB & CB & 0 \end{bmatrix} \begin{bmatrix} \tilde{u}_k \\ \vdots \\ \tilde{u}_{k+h-2} \\ \tilde{u}_{k+h-1} \end{bmatrix} = \begin{bmatrix} C \\ \vdots \\ CA^{h-1} \end{bmatrix} x_k$$

La procédure consiste à rechercher le plus petit horizon pour lequel il existe une matrice Ω de plein rang ligne qui annule le terme inconnu de droite : $\Omega / \Omega [C^T \dots (CA^{h-1})^T]^T = 0$. Or, en appliquant cette procédure à l'exemple, nous trouvons que le plus petit horizon est $h=3$ et qu'il correspond à la relation suivante :

$$4\tilde{y}_k + 3\tilde{y}_{k+1} - 10\tilde{y}_{k+2} + 11.5\tilde{u}_k + 23\tilde{u}_{k+1} = 0$$

$$= (4\tilde{y}_k - 5\tilde{y}_{k+1} + 11.5\tilde{u}_k) + 2(4\tilde{y}_{k+1} - 5\tilde{y}_{k+2} + 11.5\tilde{u}_{k+1})$$

Même si ce dernier résultat est une combinaison linéaire du précédent, il n'en reste pas moins différent : le premier implique le second, mais la réciproque n'est pas vraie puisque, par transformation linéaire et décalage, le premier résultat ne peut être retrouvé. Pour comprendre cette différence, regardons la relation que nous aurions obtenue avec l'approche temporelle en utilisant une valeur $\Omega_2=[4 \ -5]$ de Ω correspondant aux résultats de l'approche fréquentielle :

$$4\tilde{y}_k - 5\tilde{y}_{k+1} + 11.5\tilde{u}_k = [6.5 \ -5]x_k \neq 0$$

Le résultat semble non nul mais, en étudiant le système, nous pouvons nous apercevoir qu'il n'est pas commandable. Autrement dit, les composantes de x_k sont liées en régime établi. En évaluant la matrice de commandabilité, il vient : $[6,5 \ -5] [B \ AB]=0$ et nous déduisons que la combinaison linéaire $[6,5 \ -5]x_k$ s'annule en régime établi¹³. Ainsi, pour obtenir l'équivalence des deux approches, il faudrait non plus trouver une matrice Ω de plein rang ligne telle que $\Omega [C^T \ \dots \ (CA^{h-1})^T]^T = 0$, mais plutôt :

$$\Omega \begin{bmatrix} C \\ \vdots \\ CA^{h-1} \end{bmatrix} [B \ \dots \ A^{dim(A)-1} B] = 0$$

Faut-il pour autant conclure que cette dernière solution est la meilleure ? Pour nous en rendre compte, mettons le système précédent sous une forme canonique pour la commandabilité :

$$\begin{cases} x'_{k+1} = \begin{bmatrix} -0.5 & 0 \\ 0 & 0.8 \end{bmatrix} x'_k + \begin{bmatrix} 0 \\ 1 \end{bmatrix} \tilde{u}_k \\ \tilde{y}_k = [1 \ 2.3] x'_k \end{cases} \text{ avec : } x = \begin{bmatrix} 1 & 1 \\ 0 & 1.3 \end{bmatrix} x'$$

Il apparaît que le pôle non commandable est stable. Il s'ensuit qu'après convergence, la relation de parité dont l'horizon est 2, sera vérifiée tandis que l'autre relation l'est dès l'instant initial. Si le pôle non commandable avait été instable, alors la relation dont l'horizon est 2 n'aurait jamais été vérifiée !

En ce qui nous concerne, nous avons pris le parti de ne pas nous intéresser au cas particulier des systèmes aux pôles cachés instables et à ne considérer que le régime établi des systèmes¹⁴. Nous préférons considérer qu'une période d'initialisation est nécessaire et donc, ne nous intéresser qu'aux systèmes d'équations commandables. Dans cet exemple, nous aurions choisi :

$$\begin{cases} x''_{k+1} = 0.8x''_k + \tilde{u}_k \\ \tilde{y}_k = 2.3x''_k \end{cases}$$

¹³ Nous opposons régime établi à régime transitoire dû à une valeur initiale x_0 non nulle. Le régime établi est atteint lorsque cette valeur n'influence plus la réponse.

¹⁴ Il nous semble saugrenu d'escompter obtenir un diagnostic dès l'instant zéro.

Ceci rejoint la condition de minimalité décrite dans la sous-partie précédente.

II.2. Propriétés génériques des hypomodèles certains

Les hypomodèles certains à horizon infini

Les propriétés génériques que nous allons mettre en évidence devront être, lorsqu'elles seront réunies, des conditions nécessaires et suffisantes pour que les hypomodèles certains à horizon infini, définis par (6) et (7) ou, de manière ordonnée par (10a) et (10b), reconstruisent la fonctionnelle Wx_k . Notre démarche consiste à établir les propriétés nécessaires puis à ne s'assurer qu'a posteriori de leur suffisance.

L'erreur à annuler doit normalement vérifier (10b). Pour introduire cette erreur dans l'équation précédente, nous utilisons la relation (I) que doit vérifier x_k . En multipliant l'équation récurrente (I) à droite par la matrice W et en la soustrayant à la première des 2 équations récurrentes contenues dans (10a), nous obtenons :

$$\begin{aligned} \begin{bmatrix} \varepsilon_{k+1} \\ \xi_{k+1} \end{bmatrix} &= \begin{bmatrix} HNH^+ & HN\bar{H}^T \\ \bar{H}NH^+ & \bar{H}N\bar{H}^T \end{bmatrix} \begin{bmatrix} \varepsilon_k \\ \xi_k \end{bmatrix} + \begin{bmatrix} HG + HLD - WB \\ \bar{H}(G + LD) \end{bmatrix} \tilde{u}_k + \dots \\ &\dots + \begin{bmatrix} HNH^+W - WA + HLC \\ \bar{H}NH^+W + \bar{H}LC \end{bmatrix} x_k \end{aligned} \quad (12)$$

Commençons par nous intéresser aux propriétés que doit posséder l'hypomodèle pour converger vers des points d'équilibre adéquats. Pour que l'hypomodèle puisse assumer sa fonction, il est nécessaire que l'erreur ε_k s'annule normalement ; autrement dit, il suffit que les pôles de la matrice N soient compris à l'intérieur du cercle unité. Une exception doit cependant être faite pour la valeur 1 qui peut correspondre à la présence d'intégrateurs. La première propriété générique des hypomodèles à horizon infini porte donc sur les pôles de la matrice N ; notons que les pôles nuls de N ont été exclus pour que l'horizon soit effectivement infini.

Propriété générique [PII]

$$\exists N / \lambda(N) \in (\{\zeta \in \mathbf{C} : 0 < |\zeta| < 1\} \cup \{1\})'$$

où $\lambda(N)$ représente l'ensemble des pôles de N

La présence de pôles égaux à l'unité peut conduire l'erreur ε_k à ne pas converger vers une valeur finie. Pour éluder ces anomalies, posons :

$$\psi_k = \begin{bmatrix} HG + HLD - WB \\ \bar{H}(G + LD) \end{bmatrix} \tilde{u}_k + \begin{bmatrix} HNH^+W - WA + HLC \\ \bar{H}NH^+W + \bar{H}LC \end{bmatrix} x_k \quad (13a)$$

et supposons que les différents vecteurs admettent les limites suivantes :

$$\lim_{k \rightarrow \infty} (\varepsilon_k) = \varepsilon_\infty ; \lim_{k \rightarrow \infty} (\xi_k) = \xi_\infty \text{ et } \lim_{k \rightarrow \infty} (\psi_k) = \psi_\infty$$

Lorsque le point d'équilibre a été atteint, l'équation (12) nous permet d'établir la relation suivante :

$$\begin{bmatrix} H(I_1 - N) \\ -\bar{H}N \end{bmatrix} H^+ \varepsilon_\infty + \begin{bmatrix} -HN \\ \bar{H}(I_1 - N) \end{bmatrix} \bar{H}^T \xi_\infty = \psi_\infty \quad (13b)$$

Or, pour que l'erreur ε_k s'annule inmanquablement, il est nécessaire qu'il n'existe qu'une unique valeur d'équilibre ε_∞ . Il est donc indispensable que :

$$\ker \left(\begin{bmatrix} H(I_1 - N) \\ -\bar{H}N \end{bmatrix} H^+ \right) = \ker((I_1 - N)H^+) = 0$$

Sous une formulation équivalente, ceci nous conduit à la seconde propriété générique :

Propriété générique [PI2]

$$\exists N \mid \text{rang}((I_1 - N)H^+) = r$$

Nous sommes désormais assurés que l'erreur ε_k peut atteindre une valeur définie mais il reste à établir sous quelles conditions l'erreur peut s'annuler.

Pour qu'un point d'équilibre soit atteint et qu'il coïncide avec $\varepsilon_\infty=0$, il est nécessaire que le vecteur ψ_k admette une limite quand k tend vers l'infini. Or, ce vecteur est fonction de \tilde{u}_k et de x_k , qui, a priori, ne tendent pas vers des valeurs constantes. Il est donc indispensable que les matrices intervenant dans ψ_k soient nulles. Nous en déduisons 4 conditions supplémentaires :

$$\begin{cases} HNH^+W - WA + HLC = 0 & (14a) \end{cases}$$

$$\begin{cases} \bar{H}NH^+W + \bar{H}LC = 0 & (14b) \end{cases}$$

$$\begin{cases} HG + HLD - WB = 0 & (14c) \end{cases}$$

$$\begin{cases} \bar{H}(G + LD) = 0 & (14d) \end{cases}$$

Les conditions (14a) et (14b) peuvent être réunies de la façon suivante :

$$\begin{bmatrix} H \\ \bar{H} \end{bmatrix} (NH^+W + LC) = \begin{bmatrix} WA \\ 0 \end{bmatrix}$$

La troisième propriété [PI3] suffisante est obtenue en multipliant à gauche cette relation par $\begin{bmatrix} H^+ & \bar{H}^T \end{bmatrix}$:

Propriété générique [PI3]

$$\exists N, L \mid NH^+W + LC = H^+WA$$

Tandis que la quatrième [PI4] se déduit de la même manière de (14c) et de (14d) :

Propriété générique [PI4]

$$\exists G, L / G = H^+WB - LD$$

Enfin, la dernière des propriétés génériques n'apparaît que lors d'un découplage parfait. Nous avons en effet spécifié que l'hypomodèle ne devait faire intervenir que les entrées $S\tilde{u}_k$. Or, dans l'hypomodèle (6), les entrées n'apparaissent que dans le terme $G\tilde{u}_k$ qui peut aussi s'écrire : $G\tilde{u}_k = GS^T(S\tilde{u}_k) + G\bar{S}^T(\bar{S}\tilde{u}_k)$. Pour que seule la fraction $S\tilde{u}_k$ intervienne dans l'hypomodèle, il faut que la propriété suivante soit vérifiée :

Propriété générique [PI5]

$$\exists G / G\bar{S}^T = 0$$

La satisfaction simultanée de ces cinq propriétés nécessaires est naturellement une condition nécessaire pour que la fonctionnelle $H\hat{z}_k$ de l'hypomodèle converge vers Wx_k . Les propriétés [PI3] et [PI4] permettent d'établir que :

$$\begin{bmatrix} \varepsilon_{k+1} \\ \xi_{k+1} \end{bmatrix} = \begin{bmatrix} HNH^+ & HN\bar{H}^T \\ \bar{H}NH^+ & \bar{H}N\bar{H}^T \end{bmatrix} \begin{bmatrix} \varepsilon_k \\ \xi_k \end{bmatrix}$$

En y ajoutant les propriétés de convergence [PI1] et [PI2], l'erreur (10b) est assurée de converger vers zéro. Quant à la propriété [PI5], elle garantit que seule la fraction $S\tilde{u}_k$ des entrées intervient dans l'hypomodèle. Pour montrer que ces propriétés sont suffisantes, il suffit de montrer qu'elles sont la traduction mathématique de la définition de l'hypomodèle à horizon infini de (1). Ceci peut être obtenu sans difficulté pour peu que soient exclues des hypomodèles, les équations récurrentes contenant des modes oscillants non-amortis ; ce que nous ferons sans hésitation dans la mesure où ils n'apportent rien à la reconstruction.

L'ensemble des observateurs d'état linéaires invariants à horizon infini est entièrement défini par les équations (10a), (10b) et par les cinq propriétés génériques précédentes. Il inclut les observateurs d'état complets ou partiels tels que les a décrits O'Reilly [O'Reilly,83], les observateurs à effet intégral [Shafai et al,89], les observateurs élémentaires [Arsan,94] ou encore de tout autre hypomodèle linéaire, invariant et à horizon infini. Dans le cas de métamodèles non-invariants, les propriétés [PI3], [PI4] et [PI5] restent valides, pourvu qu'elles soient vérifiées à tout instant, alors que les propriétés [PI1] et [PI2] ne sont plus que suffisantes. Quant aux modèles non-linéaires [Isidori,89], [Nijmeijer,90], d'une part les propriétés [PI3], [PI4] et [PI5] se transposent au linéarisé tangent en perdant le caractère de nécessité, et d'autre part, les propriétés [PI1] et [PI2] doivent être adaptées de sorte de garantir la convergence globale.

Les hypomodèles certains à horizon fini

Pour établir les propriétés génériques des hypomodèles minimaux à horizon fini, nous devons rechercher la nature des matrices intervenant dans l'expression générique (11a) lorsque seule

la connaissance (11b) du métamodèle (1) est exploitée. La connaissance (11b) ne peut pas toujours être utilisée du fait que toutes les entrées u_k ne sont pas disponibles : seule Su_k l'est. Il est donc nécessaire d'extraire de (11b) la relation ne faisant intervenir que les observations disponibles, c'est-à-dire y_k et Su_k .

Pour ce faire, remarquons que la relation (11b) peut être développée ainsi :

$$\begin{cases} Wx_{k+1} = [WAW^+ & WA\bar{W}^T] \begin{bmatrix} Wx_k \\ \bar{W}x_k \end{bmatrix} + [WBS^+ & WBS^T] \begin{bmatrix} Su_k \\ \bar{S}u_k \end{bmatrix} \\ y_k = [CW^+ & C\bar{W}^T] \begin{bmatrix} Wx_k \\ \bar{W}x_k \end{bmatrix} + [DS^+ & DS^T] \begin{bmatrix} Su_k \\ \bar{S}u_k \end{bmatrix} \end{cases}$$

Le vecteur $\bar{S}u_k$ contient les entrées vis-à-vis desquelles le modèle doit être indépendant alors que le vecteur $\bar{W}x_k$ contient la fraction du vecteur d'état qui ne peut plus être déduite du modèle ; les variables contenues dans $\bar{S}u_k$ et $\bar{W}x_k$ ne peuvent donc plus apparaître dans la partie exploitable du métamodèle. En isolant les termes correspondants à ces variables, nous déduisons le modèle suivant :

$$\begin{cases} Wx_{k+1} = WAW^+(Wx_k) + WBS^+(Su_k) + [WA\bar{W}^T & WBS^T] \begin{bmatrix} \bar{W}x_k \\ \bar{S}u_k \end{bmatrix} \\ y_k = CW^+(Wx_k) + DS^+(Su_k) + [C\bar{W}^T & DS^T] \begin{bmatrix} \bar{W}x_k \\ \bar{S}u_k \end{bmatrix} \end{cases} \quad (15)$$

où les termes grisés contiennent les variables non disponibles

Pour pouvoir éliminer les termes grisés, il faut que la propriété suivante soit vérifiée :

Propriété générique [PF1]

$$\exists K \in \mathbb{R}^{r \times p} / [WA\bar{W}^T \quad WBS^T] = K[C\bar{W}^T \quad DS^T]$$

Si cette propriété est vérifiée, la matrice K sera donnée par :

$$K = W[A\bar{W}^T \quad BS^T] \Gamma^+ + \ell(\Gamma \Gamma^+ - I_p) : \forall \ell \in \mathbb{R}^{r \times p} \quad (16)$$

où $\Gamma = [C\bar{W}^T \quad DS^T]$ et Γ^+ son pseudo-inverse / $\Gamma \Gamma^+ \Gamma = \Gamma$

Il est toujours possible de se ramener à des matrices de plein rang ligne en décomposant la matrice Γ en valeurs singulières :

$$\Gamma = \Phi^T \begin{bmatrix} \Sigma & 0 \\ 0 & 0 \end{bmatrix} \Phi \quad (17a)$$

où Φ et ϕ sont des matrices orthogonales et Σ une matrice diagonale inversible

Les matrices Γ^+ et $(\Gamma \Gamma^+ - I_p)$ peuvent désormais être évaluées :

$$\Gamma^+ = \Phi^T \begin{bmatrix} \Sigma^{-1} & 0 \\ 0 & 0 \end{bmatrix} \Phi = \Gamma_2^T \bar{\Gamma}_1 \text{ et } (\Gamma \Gamma^+ - I_p) = -\Phi^T \begin{bmatrix} 0 & 0 \\ 0 & I_{p-\text{rang}(\Gamma)} \end{bmatrix} \Phi = \bar{\Gamma}_1^T \bar{\Gamma}_1 \quad (17b)$$

avec $\Gamma_1 = \begin{bmatrix} 0 & I_{\text{rang}(\Gamma)} \end{bmatrix} \Phi$ et $\Gamma_2 = \begin{bmatrix} \Sigma^{-1} & 0 \end{bmatrix} \Phi$, deux matrices $\mathbb{R}^{\text{rang}(\Gamma) \times p}$ de plein rang ligne

Cette décomposition en matrices de plein rang ligne peut être introduite dans l'expression (16) de K , qui peut, à son tour, en utilisant la propriété [PFI], être introduit dans le modèle (15). Nous obtenons :

$$Wx_{k+1} = WAW^+(Wx_k) + WBS^+(Su_k) + W \begin{bmatrix} A\bar{W}^T & BS^T \end{bmatrix} \Gamma_2^+ \bar{\Gamma}_1 (y_k - CW^+(Wx_k) - DS^+(Su_k)) \\ \dots + N\bar{\Gamma}_1^T (y_k - CW^+(Wx_k) - DS^+(Su_k))$$

avec $N = \ell \bar{\Gamma}_1^T$ une matrice quelconque de $\mathbb{R}^{r \times \text{rang}(\Gamma)}$

Nous reconnaissons là l'expression des hypomodèles d'ordre complet construit sur le métamodèle réduit suivant :

$$\begin{cases} x_{k+1}^r = A_r x_k^r + B_r u_k^r \\ y_k^r = C_r x_k^r + D_r u_k^r \end{cases} \text{ avec : } x_k^r = Wx_k, u_k^r = \begin{bmatrix} Su_k \\ \Gamma_1 y_k \end{bmatrix} \text{ et } y_k^r = \bar{\Gamma}_1 y_k \quad (18)$$

et

$$\begin{aligned} A_r &= W \left(A - \begin{bmatrix} A\bar{W}^T & BS^T \end{bmatrix} \Gamma_2^T \bar{\Gamma}_1 C \right) W^+ \\ B_r &= W \left[\left(B - \begin{bmatrix} A\bar{W}^T & BS^T \end{bmatrix} \Gamma_2^T \bar{\Gamma}_1 D \right) S^+ \begin{bmatrix} A\bar{W}^T & BS^T \end{bmatrix} \Gamma_2^T \right] \\ C_r &= \bar{\Gamma}_1 C W^+ \\ D_r &= \begin{bmatrix} \bar{\Gamma}_1 D S^+ & 0 \end{bmatrix} \end{aligned}$$

Le métamodèle (18) contient la connaissance représentée par (11b) mais en plus, il la conditionne en une forme usuelle ; il s'agit d'un métamodèle réduit de (1). Ce nouveau modèle ne dépend que des entrées Su_k et ne fait intervenir que la connaissance (11b). La formalisation (11a) peut alors être ramenée à ce métamodèle réduit. En effet, l'écriture (11a) se développe ainsi :

$$\sum_{i=0}^{h-1} \left(U_i Su_{k+i} + Y_i \Gamma_1^+ (\Gamma_1 y_{k+i}) + Y_i \bar{\Gamma}_1^T (\bar{\Gamma}_1 y_{k+i}) \right) = 0$$

ou encore

$$\sum_{i=0}^{h-1} \left(\begin{bmatrix} U_i & Y_i \Gamma_1^+ \end{bmatrix} \begin{bmatrix} Su_{k+i} \\ \Gamma_1 y_{k+i} \end{bmatrix} + Y_i \bar{\Gamma}_1^T (\bar{\Gamma}_1 y_{k+i}) \right) = 0$$

Puis, en notant

$$U_i^r = \begin{bmatrix} U_i & Y_i \Gamma_1^+ \end{bmatrix} \text{ et } Y_i^r = Y_i \bar{\Gamma}_1^T \quad (19)$$

nous déduisons que l'écriture (11a) admet toujours la forme équivalente suivante :

$$\sum_{i=0}^{h-1} \left(U_i^r u_{k+i}^r + Y_i^r y_{k+i}^r \right) = 0 \quad (20)$$

Il apparaît très explicitement que l'hypomodèle à horizon fini (11a) exploitant la connaissance partielle (11b) est identique à l'hypomodèle (20) exploitant la représentation (18) ; en d'autres termes, la résolution du problème (11) se ramène à la résolution d'un problème classique de la synthèse d'un hypomodèle à horizon fini exploitant toute la connaissance d'un métamodelle réduit de type (18).

Nous partons donc désormais de la formalisation générique équivalente constituée de (18) et de (20) pour rechercher les propriétés génériques en sachant que grâce à (18) et (19), les résultats obtenus peuvent être rapportés à la formalisation (11a) et (11b) originale.

Tout d'abord, nous obtenons de (18) :

$$y_{k+i}^r = C_r A_r^i x_k^r + D_r u_{k+i}^r + \sum_{j=0}^{i-1} (C_r A_r^j B_r) u_{k+i-j-1}^r ; \forall i \in \mathbb{N}^*$$

En substituant l'expression de y_{k+i}^r dans (20), nous obtenons :

$$\sum_{i=0}^{h-1} (U_i^r + Y_i^r D_r) u_{k+i}^r + \sum_{i=0}^{h-1} Y_i^r C_r A_r^i x_k^r + \sum_{i=1}^{h-1} \sum_{j=0}^{i-1} (Y_i^r C_r A_r^j B_r) u_{k+i-j-1}^r = 0$$

Or, dans cette expression apparaît une double sommation qui peut être reformulée de façon à ne faire apparaître que des termes en u_{k+i}^r :

$$\sum_{i=1}^{h-1} \sum_{j=0}^{i-1} (Y_i^r C_r A_r^j B_r) u_{k+i-j-1}^r = \sum_{i=0}^{h-2} \left(\sum_{j=0}^{h-i-2} Y_{i+j+1}^r C_r A_r^j B_r \right) u_{k+i}^r$$

En introduisant cette expression dans la relation précédente et en factorisant, nous obtenons :

$$\sum_{i=0}^{h-1} Y_i^r C_r A_r^i x_k^r + \sum_{i=0}^{h-2} \left(U_i^r + Y_i^r D_r + \sum_{j=0}^{h-i-2} Y_{i+j+1}^r C_r A_r^j B_r \right) u_{k+i}^r + (U_{h-1}^r + Y_{h-1}^r D_r) u_{k+h-1}^r = 0$$

Pour que cette relation puisse être vérifiée quels que soient x_k^r et les valeurs de u_{k+i}^r quand $i \in \{0, \dots, h-1\}$, il est nécessaire que les propriétés suivantes soient vérifiées lorsque l'hypomodèle est minimal :

Propriété générique [PF2]

$$\exists h \in \mathbb{N} / \sum_{i=0}^{h^*-1} Y_i^{r^*} C_r A_r^i = Y_r^* \begin{bmatrix} C_r \\ C_r A_r \\ \vdots \\ C_r A_r^{h-1} \end{bmatrix} = 0$$

avec $Y_r^* = [Y_0^{r^*} \dots Y_{h-1}^{r^*}]$, où l'astérisque rappelle qu'une condition de minimalité doit être satisfaite.

Propriété générique [PF3]

$$U_i^{r^*} = -Y_i^{r^*} D_r - \sum_{j=0}^{h^*-i-2} Y_{i+j+1}^{r^*} C_r A_r^j B_r, \forall i \in \{0, \dots, h^*-2\} \text{ et } h^* > 1$$

$$\text{et } U_{h-1}^{r*} = -Y_{h-1}^{r*} D_r$$

La propriété [PF3] peut être exprimée sous une forme matricielle bien connue :

$$U_r^* = -Y_r^* \begin{bmatrix} D_r & 0 & 0 & 0 \\ C_r B_r & D_r & 0 & 0 \\ C_r A_r B_r & C_r B_r & \ddots & 0 \\ \vdots & \vdots & & D_r \\ CA_r^{h^*-2} B_r & CA_r^{h^*-3} B_r & \dots & C_r B_r & D_r \end{bmatrix}$$

avec $U_r^* = [U_0^{r*} \dots U_{h-1}^{r*}]$

La propriété générique [PF1] garantit que la partie du métamodèle (I) définie par (11b) est compatible avec la présence d'entrées inconnues définies par S, tandis que les propriétés [PF2] et [PF3] assurent que la relation (20) et, a fortiori, en utilisant (18) et (19), la relation (11a) est bien un hypomodèle correspondant aux objectifs. Ces propriétés sont donc, pour peu que les matrices Y_r^* et U_r^* engendrent bien un hypomodèle minimal générateur de tous les autres hypomodèles, des conditions nécessaires et suffisantes pour que la relation (11a) à laquelle s'ajoute les objectifs (11b) soit un hypomodèle minimal à horizon fini de (I).

II.3. Conditions d'existence des hypomodèles certains

Les hypomodèles certains à horizon infini

Toutes les propriétés génériques ayant été posées, intéressons-nous désormais aux conditions d'existence des hypomodèles certains à horizon infini correspondant aux objectifs définis par les matrices R, W, H, S et éventuellement, C et D, qui doivent être conformes aux conditions d'existence spécifiques au diagnostic [CID](p.139).

Transformons alors les écritures des différentes propriétés de telle sorte de faire apparaître les conditions d'existence. La propriété [PI4](p.145) peut toujours être satisfaite : pour toute matrice L, il correspond une matrice G. Mais, la propriété [PI5](p.145) restreint, en cas de découplage, l'ensemble des matrices G possibles. En combinant ces deux propriétés, nous déduisons que, pour trouver une matrice G donnée par [PI4] et vérifiant [PI5], il faut que la matrice L vérifie :

$$L | LDS^T = H^+ WBS^T \quad (21)$$

Puis, en multipliant la relation de la propriété [PI3](p.144) par $[W^+ \quad \bar{W}^T]$, nous en déduisons les relations suivantes :

$$NH^+ = H^+ WAW^+ - LCW^+ \quad (22a)$$

$$LC\bar{W}^T = H^+ W\bar{W}^T \quad (22b)$$

Le produit matriciel NH^+ sera donné par (22a) lorsque la condition (22b) sera satisfaite. Or, nous constatons qu'en plus de cette condition, la matrice L doit satisfaire la relation (21). En réunissant ces deux conditions, nous obtenons :

$$L / L \begin{bmatrix} C\bar{W}^T & D\bar{S}^T \end{bmatrix} = \begin{bmatrix} H^+WA\bar{W}^T & H^+WBS^T \end{bmatrix}$$

Décomposons la matrice L ainsi : $L = H^+\ell_a + \bar{H}^T\ell_b$ avec $\ell_a = HL$ et $\ell_b = \bar{H}L$, et multiplions à gauche la relation précédente par $\begin{bmatrix} H^T & \bar{H}^T \end{bmatrix}^T$. En utilisant les nouvelles matrices ℓ_a et ℓ_b , nous obtenons les deux relations suivantes :

$$\ell_a / \ell_a \begin{bmatrix} C\bar{W}^T & D\bar{S}^T \end{bmatrix} = \begin{bmatrix} WA\bar{W}^T & WBS^T \end{bmatrix} \quad (23a)$$

$$\ell_b / \ell_b \begin{bmatrix} C\bar{W}^T & D\bar{S}^T \end{bmatrix} = 0 \quad (23b)$$

Il existe toujours une ou plusieurs matrices ℓ_b satisfaisant à (23b). Par contre, des matrices ℓ_a vérifiant (23a) ne pourront être trouvées que lorsque la condition d'existence de l'hypomodèle à horizon infini suivante sera satisfaite :

Condition d'existence [CII]

$$\text{rang} \left(\begin{bmatrix} WA\bar{W}^T & WBS^T \\ C\bar{W}^T & D\bar{S}^T \end{bmatrix} \right) = \text{rang} \left(\begin{bmatrix} C\bar{W}^T & D\bar{S}^T \end{bmatrix} \right)$$

Cette condition garantit que les propriétés génériques [PI3], [PI4] et [PI5] pourront être vérifiées. Il ne reste désormais plus qu'à étudier sous quelles conditions les propriétés [PI1] et [PI2] pourront l'être aussi.

Commençons par rechercher l'ensemble de matrices L possibles. Si la condition [CII] est vérifiée, l'ensemble des matrices ℓ_a possibles est donné par :

$$\ell_a = W \begin{bmatrix} A\bar{W}^T & BS^T \end{bmatrix} \Gamma^+ + \ell_1 (\Gamma^+ - I_p) : \forall \ell_1 \in \mathbb{R}^{r \times p}$$

où $\Gamma = \begin{bmatrix} C\bar{W}^T & D\bar{S}^T \end{bmatrix}$ et Γ^+ son pseudo-inverse / $\Gamma\Gamma^T = \Gamma$

Tandis que l'ensemble des matrices ℓ_b possibles correspond à :

$$\ell_b = \ell_2 (\Gamma^+ - I_p) : \forall \ell_2 \in \mathbb{R}^{(l-r) \times p}$$

Puisque les expressions des matrices ℓ_a et ℓ_b sont maintenant établies, il est possible de les réunir pour revenir à l'expression de L :

$$L = \begin{bmatrix} H^+ & \bar{H}^T \end{bmatrix} \left(\begin{bmatrix} W \begin{bmatrix} A\bar{W}^T & BS^T \end{bmatrix} \Gamma^+ \\ 0 \end{bmatrix} + L_1 (\Gamma^+ - I_p) \right) : \forall L_1 \in \mathbb{R}^{l \times p} \quad (24)$$

Pour aboutir à l'expression de la matrice N , remarquons que la relation (22a) est la seule contrainte structurelle portant sur cette matrice et que cette contrainte ne concerne que le

produit NH^+ . En conséquence, la structure du terme $N\bar{H}^T$ peut a priori être choisie librement. Nous poserons $N\bar{H}^T = \begin{bmatrix} H^+ & \bar{H}^T \end{bmatrix} L_2$ où $L_2 \in \mathbb{R}^{l \times (l-r)}$.

En introduisant (24) dans (22a) puis en factorisant, le produit NH^+ peut aussi s'écrire :

$$NH^+ = \begin{bmatrix} H^+ & \bar{H}^T \end{bmatrix} \left(\begin{bmatrix} W(A - [A\bar{W}^T & B\bar{S}^T] \Gamma^+ C) W^+ \\ 0 \end{bmatrix} - L_1 (\Gamma \Gamma^+ - I_p) C W^+ \right)$$

En juxtaposant les expressions de NH^+ et de $N\bar{H}^T$ dans une matrice $[NH^+ \ N\bar{H}^T]$, et en la multipliant à droite par $\begin{bmatrix} H^T & \bar{H}^T \end{bmatrix}^T$, nous obtenons l'expression de N :

$$N = \begin{bmatrix} H^+ & \bar{H}^T \end{bmatrix} \left(\begin{bmatrix} W(A - [A\bar{W}^T & B\bar{S}^T] \Gamma^+ C) W^+ & 0 \\ 0 & 0 \\ \dots & \dots \\ \dots - [L_1 & L_2] \begin{bmatrix} (\Gamma \Gamma^+ - I_p) C W^+ & 0 \\ 0 & -I_{l-r} \end{bmatrix} \end{bmatrix} \begin{bmatrix} H \\ \bar{H} \end{bmatrix} \right) \quad (25)$$

avec $\Gamma = [C\bar{W}^T \ D\bar{S}^T]$

Puisque la matrice L_1 peut être choisie librement, il faut, pour pouvoir réaliser la propriété [PII](p.143), que les seuls pôles non observables de la paire suivante soient à l'intérieur du cercle unité ou égaux à 1 :

$$\left(\begin{bmatrix} W(A - [A\bar{W}^T & B\bar{S}^T] \Gamma^+ C) W^+ & 0 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} (\Gamma \Gamma^+ - I_p) C W^+ & 0 \\ 0 & -I_{l-r} \end{bmatrix} \right)$$

Ceci se traduit par la condition suivante :

$$\text{rang} \begin{pmatrix} \lambda I_r - W(A - [A\bar{W}^T & B\bar{S}^T] \Gamma^+ C) W^+ & 0 \\ 0 & \lambda I_{l-r} \\ (\Gamma \Gamma^+ - I_p) C W^+ & 0 \\ 0 & -I_{l-r} \end{pmatrix} \neq l$$

$$\Rightarrow \lambda \in (\{\zeta \in \mathbb{C} : 0 < |\zeta| < 1\} \cup \{1\})$$

ou encore par :

$$\text{rang} \begin{pmatrix} \lambda I_r - W(A - [A\bar{W}^T & B\bar{S}^T] \Gamma^+ C) W^+ \\ (\Gamma \Gamma^+ - I_p) C W^+ \end{pmatrix} \neq r \quad (26)$$

$$\Rightarrow \lambda \in (\{\zeta \in \mathbb{C} : 0 < |\zeta| < 1\} \cup \{1\})$$

Intéressons-nous désormais à la propriété générique [PI2]. L'expression (25) nous permet de développer $(I_l - N)H^+$:

$$(I_r - N)H^+ = \begin{bmatrix} H^+ & \bar{H}^T \end{bmatrix} \begin{pmatrix} \begin{bmatrix} I_r - W(A - [A\bar{W}^T & B\bar{S}^T] \Gamma^+ C) W^+ \\ \dots \\ 0 \\ \dots + [L_1 & L_2] \begin{bmatrix} (\Gamma^+ - I_p) C W^+ \\ 0 \end{bmatrix} \end{bmatrix} \end{pmatrix}$$

Dès lors, la propriété **[PI2]** devient :

$$\text{rang}(I_r - W(A - [A\bar{W}^T \quad B\bar{S}^T] \Gamma^+ C) W^+ + L_1 (\Gamma^+ - I_p) C W^+) = r$$

ce qui est équivalent à

$$\text{rang} \begin{pmatrix} I_r - W(A - [A\bar{W}^T & B\bar{S}^T] \Gamma^+ C) W^+ \\ (\Gamma^+ - I_p) C W^+ \end{pmatrix} = r$$

Cette condition peut être regroupée avec **(26)**, il s'ensuit que :

$$\begin{aligned} \text{rang} \begin{pmatrix} \lambda I_r - W(A - [A\bar{W}^T & B\bar{S}^T] \Gamma^+ C) W^+ \\ (\Gamma^+ - I_p) C W^+ \end{pmatrix} &\neq r \\ \Rightarrow \lambda \in \{ \zeta \in \mathbf{C} : 0 < |\zeta| < 1 \} \end{aligned}$$

Cette définition coïncide avec celle de la détectabilité. Les propriétés **[PII]** et **[PI2]** pourront donc être réalisées si et seulement si la condition suivante est vérifiée :

Condition d'existence [CI2]

La paire $(W(A - [A\bar{W}^T \quad B\bar{S}^T] \Gamma^+ C) W^+, (\Gamma^+ - I_p) C W^+)$ est détectable.

$$\text{avec } \Gamma = [C\bar{W}^T \quad D\bar{S}^T]$$

La première condition d'existence **[CII]** garantit que les propriétés génériques **[PI3]** et **[PI5]** admettent une solution L tandis que la seconde garantit que les propriétés génériques **[PII]** et **[PI2]** portant sur la stabilité peuvent être satisfaites. La propriété générique **[PI4]** peut toujours être satisfaite : elle ne génère pas de conditions d'existence supplémentaires. En conclusion, les deux conditions d'existence proposées sont des conditions nécessaires et suffisantes pour qu'un hypomodèle à horizon infini puisse reconstruire la fonctionnelle Wx_k et donc, sous réserve de satisfaction aux conditions d'existence spécifiques au diagnostic **[CID]**(p.139), reconstruire la fraction $R\tilde{y}_k$ des observations \tilde{y}_k .

Enfin, rappelons qu'il serait absurde de choisir une représentation qui ne soit pas minimale bien que ceci ne soit pas réellement une contrainte. Autrement dit, il faut veiller à choisir une structure, par l'intermédiaire de L_l dans **(25)**, qui conduise à une paire (N, G) qui soit commandable. Il s'agit d'une condition de minimalité :

Condition de minimalité [PIM]

La matrice L_l doit être choisie telle que la paire (N, G) soit commandable.

Bien que difficile à résoudre, cette condition n'a aucune raison de ne pas être vérifiée et même si cela devait se produire, l'hypomodèle ne serait, certes, pas minimale mais cela ne gênerait en rien la reconstruction.

Les hypomodèles certains à horizon fini

Pour ce qui est des hypomodèles à horizon fini, il est plus simple d'établir leurs conditions d'existence puisqu'il n'existe que trois propriétés génériques. La propriété générique [PF1](p.146) induit directement une condition d'existence identique à la condition [CII](p.150) :

Condition d'existence [CFI]

$$\text{rang} \left(\begin{bmatrix} WA\bar{W}^T & WBS^T \\ C\bar{W}^T & DS^T \end{bmatrix} \right) = \text{rang} \left(\begin{bmatrix} C\bar{W}^T & DS^T \end{bmatrix} \right)$$

La propriété [PF3](p.144) n'induit aucune condition d'existence particulière alors que la condition [PF2] (p.148) doit être considérée avec attention. Certes, cette propriété peut toujours être satisfaite pour un certain horizon h et n'engendre donc aucune condition d'existence particulière. Néanmoins, comment concilier la propriété [PF2] avec la condition de minimalité que nous énoncerons ainsi :

Condition de minimalité [PFM]

La matrice Y_r^* ¹⁵ doit engendrer un hypomodèle minimal, c'est-à-dire générateur de tous les autres hypomodèles associés aux mêmes objectifs¹⁶.

Il existe, en effet, une infinité de matrice Y_r , de tout horizon et de tout rang, satisfaisant à la propriété [PF2], mais s'il faut en retenir une, elle doit être génératrice de toutes les autres. Afin d'apprécier la complexité du problème, nous proposons l'exemple suivant :

Exemple

Intéressons-nous à un hypomodèle construit sur un métamodèle réduit (18) dont les matrices A_r et C_r sont les suivantes :

$$A_r = \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 1 & 0 \end{bmatrix}, \quad C_r = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix} \quad \text{et} \quad y_k^r = \begin{bmatrix} y_k^1 \\ y_k^2 \end{bmatrix}$$

En appliquant la procédure classique décrite par [Chow et al,84], nous construisons les matrices apparaissant dans la propriété [PF2] pour les plus petits horizons. Le plus petit horizon conduisant à une matrice non nulle est de $h=2$, nous obtenons :

¹⁵ La matrice U_r^* devrait aussi figurer ici mais la propriété [PF3] montre que la matrice U_r^* est complètement déterminée par la matrice Y_r^* .

¹⁶ Définis par les matrices W et S .

$$\begin{bmatrix} C_r \\ C_r A_r \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix} \Rightarrow Y_r^{(2)} = [0 \quad -1 \quad | \quad 0 \quad 1]$$

Pour un horizon de 3, nous obtenons :

$$\begin{bmatrix} C_r \\ C_r A_r \\ C_r A_r^2 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 0 \\ 2 & 1 & 1 \\ 0 & 1 & 0 \end{bmatrix} \Rightarrow Y_r^{(3)} = \begin{bmatrix} 0 & 1 & | & 0 & 1 & | & 0 & -2 \\ 1 & 0 & | & 1 & 0 & | & -1 & 1 \\ 0 & 0 & | & 0 & 1 & | & 0 & -1 \end{bmatrix}$$

En se remémorant l'expression (20), il apparaît que la troisième ligne de la solution $Y_r^{(3)}$ correspond, au décalage temporel près, à la solution $Y_r^{(2)}$. De la même manière la première ligne de $Y_r^{(3)}$ est entièrement générée par $Y_r^{(2)}$ puisque :

$$\begin{array}{r} -1 \times [0 \quad -1 \quad | \quad 0 \quad 1 \quad | \quad 0 \quad 0] \\ -2 \times [0 \quad 0 \quad | \quad 0 \quad -1 \quad | \quad 0 \quad 1] \\ \hline [0 \quad 1 \quad | \quad 0 \quad 1 \quad | \quad 0 \quad -2] \end{array}$$

Par contre, la seconde ligne de $Y_r^{(3)}$ est indépendante de $Y_r^{(2)}$. Par conséquent, dans cet exemple, l'hypomodèle minimal peut être défini par :

$$Y_r^* = \begin{bmatrix} 1 & 0 & | & 1 & 0 & | & -1 & 1 \\ 0 & -1 & | & 0 & 1 & | & 0 & 0 \end{bmatrix} \tag{27}$$

Cet exemple illustre toute la difficulté inhérente à la génération de matrices Y_r^* engendrant des hypomodèles minimaux. "Combien existe-t-il de lignes à la matrice Y_r^* ? Quels sont les horizons des solutions apparaissant dans Y_r^* ? Et pour finir, quelle est la valeur Y_r^* ?" sont des questions des plus intéressantes. Certes, l'exemple que nous venons de proposer a été résolu, mais de manière intuitive. Nous avons pu constater que la solution n'était pas triviale. Existe-t-il des procédures systématiques conduisant aux matrices minimales Y_r^* et, a fortiori, U_r^* ? Nous allons nous consacrer, aussi brièvement que possible, à la résolution de ces différents points.

Puisque nous nous sommes restreints aux métamodèles commandables, le métamodèle réduit (18), représentant la connaissance associée à l'hypomodèle, peut être reformulé sous forme de fonction de transfert :

$$y^r(q) = (C_r (qI_r - A_r)^{-1} B_r + D_r) u^r(q) \tag{28}$$

Cette écriture montre de manière évidente qu'il existe autant de relations indépendantes conduisant à des hypomodèles distincts que de mesures dans $y^r(q)$. Nous pouvons d'ores et déjà avancer que la matrice de plein rang ligne $Y_r^{(h)*}$ vérifie :

$$\text{rang}(Y_r^*) = \text{dim}(y^r)$$

Cependant, rien ne permet de conclure que les horizons des solutions rassemblées dans la matrice Y_r associé à (28) sont minimaux. Pour retrouver ces résultats avec l'approche temporelle, il faut rechercher les relations minimales capteur par capteur en décomposant la matrice C_r ligne par ligne, puis réunir les résultats obtenus dans l'unique matrice Y_r ; il s'agit de la méthode décrite dans [Massoumnia et al,88]. Les hypomodèles ainsi déterminés ne sont généralement pas minimaux puisque, dans l'exemple précédent, nous aurions trouvé un horizon de 4 au lieu de 3 :

$$Y_r = \left[\begin{array}{cc|cc|cc|cc} 1 & 0 & 0 & 0 & -2 & 0 & 1 & 0 \\ 0 & -1 & 0 & 1 & 0 & 0 & 0 & 0 \end{array} \right]$$

Cette matrice n'est pas minimale puisqu'il peut exister certaines combinaisons linéaires de variables de y_r plus ou moins retardées conduisant à des solutions d'horizon plus faible. Représentons ces combinaisons linéaires décalées par une matrice inversible de polynômes en q : $\varphi(q) : \mathbb{R}^{\text{dim}(y^r)} \rightarrow \mathbb{R}^{\text{dim}(y^r)}$. Le métamodèle réduit (28) admet donc la représentation fréquentielle équivalente suivante :

$$\varphi(q)y^r(q) = \varphi(q)(C_r(qI_r - A_r)^{-1}B_r + D_r)\mu^r(q)$$

Selon toute évidence, l'horizon sera minimal lorsque la matrice $\varphi(q)$ sera telle qu'il existera un maximum de compensations de pôles par des zéros dans la fonction de transfert $\varphi(q)C_r(qI_r - A_r)$. Néanmoins, il est délicat de rechercher la matrice $\varphi(q)$ satisfaisant à ces objectifs. Il est préférable d'avoir recours à une procédure reposant sur la représentation temporelle. Celle que nous exposons examine itérativement l'ensemble $\Omega^{(h)}$ des vecteurs, linéairement indépendants, orthogonaux aux matrices suivantes :

$$\Omega^{(h)} \left[\begin{array}{c} C_r \\ \vdots \\ C_r A_r^{h-1} \end{array} \right] = 0 ; h \in \{2, \dots, \text{dim}(x^r) + 1\} \quad (29)$$

Mais l'indépendance linéaire ne suffit pas à en faire des vecteurs minimaux. Il faut, en plus, considérer l'indépendance temporelle. En effet, en notant $\Omega_i^{(h)}$ le $i^{\text{ème}}$ vecteur-ligne de $\Omega^{(h)}$, les deux vecteurs-lignes suivants, associés à un horizon $h+1$, sont temporellement dépendants de $\Omega_i^{(h)}$, c'est-à-dire qu'ils correspondent tout deux à une même relation décalée dans le temps, :

$$\left[\begin{array}{c} \Omega_i^{(h)} \\ 0 \quad \dots \quad 0 \end{array} \right] \text{ et } \left[\begin{array}{c} 0 \quad \dots \quad 0 \\ \Omega_i^{(h)} \end{array} \right]$$

De la même manière, les vecteurs-ligne suivants, associés à l'horizon $h+2$, sont eux aussi temporellement dépendants :

$$\left[\begin{array}{cccc} \Omega_i^{(h)} & \xleftarrow{\dim(y^r)} & \dots & \xrightarrow{\dim(y^r)} \\ 0 & \dots & 0 & \dots & 0 \end{array} \right], \left[\begin{array}{cccc} \xleftarrow{\dim(y^r)} & \dots & \xrightarrow{\dim(y^r)} & 0 \\ 0 & \dots & 0 & \Omega_i^{(h)} & 0 & \dots & 0 \end{array} \right]$$

$$\text{et } \left[\begin{array}{cccc} \xleftarrow{\dim(y^r)} & \dots & \xrightarrow{\dim(y^r)} & 0 \\ 0 & \dots & 0 & \Omega_i^{(h)} \end{array} \right]$$

Par extrapolation, les vecteurs temporellement dépendants peuvent être déduits pour des horizons supérieurs. La procédure peut alors s'exposer simplement. Dès que des vecteurs-ligne $\Omega_i^{(h)}$ linéairement indépendants entre eux ont été obtenus pour un horizon de h , il faut, pour pouvoir considérer l'un d'entre eux comme un vecteur générateur, s'assurer de son invariance temporelle par rapport aux vecteurs générateurs précédemment trouvés. Si le vecteur $\Omega_i^{(h)}$ est temporellement indépendant des vecteurs générateurs déjà obtenus, alors c'est qu'il est lui-même générateur. Il faut considérer successivement chaque vecteur-ligne $\Omega_i^{(h)}$ de $\Omega^{(h)}$, pour chacun des horizons¹⁷ h à partir de 2. La recherche de vecteurs générateurs cesse lorsque les $\dim(y_r)$ vecteurs ont été trouvés.

En appliquant cette procédure à l'exemple précédent, nous avons obtenu la matrice Y_r^* de vecteurs-ligne normalisés suivante :

$$Y_r^* = \left[\begin{array}{cc|cc|cc} 0 & -0.7071 & 0 & 0.7071 & 0 & 0 \\ -0.5477 & -0.1826 & -0.5477 & -0.1826 & 0.5477 & -0.1826 \end{array} \right]$$

La première ligne de cette matrice correspond à un horizon de 2 tandis que la seconde à un horizon de 3. Le résultat est comparable au résultat intuitif (27), mais la procédure est ici systématique. L'hypomodèle reposant sur cette matrice pourra donc être qualifié de minimal.

II.4. Spécificités des hypomodèles certains dans le cadre du diagnostic

Le problème du découplage en diagnostic

Sans qu'il en fut explicitement fait état, les formulations, propriétés et conditions d'existence précédentes incluaient les cas de découplage par rapport à certaines fractions d'un métamodèle. Nous avons déjà envisagé les cas de découplage par rapport à des entrées inconnues $\bar{S}u_k$. Nous en avons déduit les conditions d'existence communes aux hypomodèles à horizon infini et fini : [CII](p.150) et [CFI](p.153) ; elles permettent de définir l'ensemble des matrices W compatibles avec l'absence de certaines entrées. Un découplage par rapport aux entrées $\bar{S}u_k$ est donc parfaitement représenté par une matrice W satisfaisant à l'une des conditions communes d'existence [CII] ou [CFI].

¹⁷ La matrice C , étant de plein rang ligne, l'horizon $h=1$ ne peut conduire à des vecteurs Ω_h .

Dans d'autres situations, un découplage par rapport à certaines incertitudes intervenant dans l'équation de récurrence du métamodèle (1) peut être souhaité :

$$x_{k+1} = A(\theta_k)x_k + B(\theta_k)u_k \quad (30)$$

Pour obtenir un découplage d'un hypomodèle par rapport aux incertitudes, il faut que, de la connaissance sur laquelle se fondent les hypomodèles, soit exclue toute incertitude. Ceci s'obtient encore grâce à une matrice W :

$$Wx_{k+1} = WA(\theta_k)x_k + WB(\theta_k)u_k \quad (31)$$

où les matrices $WA(\theta_k)$ et $WB(\theta_k)$ sont désormais indépendants des incertitudes θ_k

Nous avons montré, lorsque nous recherchions les propriétés génériques (II.2) des hypomodèles à horizon fini, que la connaissance pouvant être extraite de (31), en se réservant toujours la possibilité d'exclure certaines entrées, pouvait être reformulée en un métamodèle réduit (18).

Recherchons le métamodèle exclusivement porteur de la connaissance intervenant dans un hypomodèle à horizon infini donné par (10). Les expressions des matrices L et N sont respectivement données par (24) et (25). La matrice G est donnée par la propriété [PI4](p.145). Pour faire apparaître la connaissance exploitée, nous introduirons les décompositions (17) ainsi que, quand ce sera possible, les matrices et variables du métamodèle réduit (18). Nous déduisons tout d'abord de (24) :

$$\begin{bmatrix} H \\ \bar{H} \end{bmatrix} L = \begin{bmatrix} W[A\bar{W}^T & B\bar{S}^T] \Gamma_2^T \bar{\Gamma}_1 \\ 0 \end{bmatrix} + L'_1 \bar{\Gamma}_1 \quad (32a)$$

où L'_1 est quelconque puisque $L'_1 = L_1 \bar{\Gamma}_1^T$ et que $\bar{\Gamma}_1^T$ est de plein rang colonne

Puis, en introduisant les décompositions (17) dans (25), nous obtenons :

$$\begin{bmatrix} HNH^+ & HN\bar{H}^T \\ \bar{H}NH^+ & \bar{H}N\bar{H}^T \end{bmatrix} = \begin{bmatrix} A_r - L_{11}C_r & L_{21} \\ -L_{12}C_r & L_{22} \end{bmatrix} \text{ avec : } L'_1 = \begin{bmatrix} L_{11} \\ L_{12} \end{bmatrix} \text{ et } L_2 = \begin{bmatrix} L_{21} \\ L_{22} \end{bmatrix} \quad (32b)$$

Pour obtenir une expression adaptée de la matrice G , nous tirons de la propriété [PI5] que $G=GS^+S$. En utilisant la propriété [PI4], nous pouvons déduire l'expression de G qui suit :

$$\begin{bmatrix} H \\ \bar{H} \end{bmatrix} G = \begin{bmatrix} WBS^+S \\ 0 \end{bmatrix} - \begin{bmatrix} H \\ \bar{H} \end{bmatrix} LDS^+S$$

En introduisant (32a) dans la relation précédente, nous obtenons :

$$\begin{bmatrix} H \\ \bar{H} \end{bmatrix} G = \left(\begin{bmatrix} W(B - [A\bar{W}^T & B\bar{S}^T] \Gamma_2^+ \Gamma_1 D)S^+ \\ 0 \end{bmatrix} - L'_1 \bar{\Gamma}_1 DS^+ \right) S \quad (32c)$$

En introduisant les expressions (32) dans (10) et en regroupant les variables de manière à faire apparaître les variables réduites de (18), nous établissons :

$$\begin{bmatrix} \hat{w}_{k+1} \\ \hat{\xi}_{k+1} \end{bmatrix} = \begin{bmatrix} A_r - L_{11}C_r & L_{12} \\ -L_{12}C_r & L_{22} \end{bmatrix} \begin{bmatrix} \hat{w}_k \\ \hat{\xi}_k \end{bmatrix} + \begin{bmatrix} B_r \\ 0 \end{bmatrix} \tilde{u}_k^r + \begin{bmatrix} L_{11} \\ L_{22} \end{bmatrix} (\tilde{y}_k^r - D_r \tilde{u}_k^r) \quad (33a)$$

$$\varepsilon_k = \hat{w}_k - x_k^r \quad (33b)$$

où les matrices L_{11} , L_{12} , L_{21} et L_{22} peuvent être choisies librement.

Ce résultat est remarquable parce qu'il fait clairement apparaître que l'hypomodèle défini par (10) n'exploite que l'information du métamodèle réduit (18).

Qu'il s'agisse d'hypomodèle à horizon fini ou infini, nous pouvons conclure qu'un découplage implique que seule une fraction du comportement est représentée. En effet, si le métamodèle (I) représente le comportement global d'un système physique alors le métamodèle réduit (18) ne représente plus qu'un comportement local ! Dès lors, nous en venons à ce qui nous semble essentiel, les solutions consistant à éliminer les incertitudes d'un métamodèle par un découplage détournent du problème de diagnostic original. Si l'objectif d'une opération de diagnostic est de tester la cohérence entre un comportement observé et un comportement représenté par le métamodèle (I), après découplage, ce ne sera plus qu'une cohérence avec un comportement local qui sera testé. L'opération de diagnostic est donc dégradée puisque les objectifs du test de cohérence ne sont plus atteints. Découpler, modifie implicitement les objectifs d'une opération de diagnostic puisque la partie auscultée d'un système physique s'en trouve changée. Conformément à ce que nous avons énoncé au cours du premier chapitre, éliminer les incertitudes par découplage revient à extraire la partie certaine d'un métamodèle ; et plus les incertitudes sont nombreuses, et plus le comportement représenté sera local. Que l'on essaie, par exemple, d'extraire un hypomodèle certain par découplage des incertitudes additives intervenant dans le métamodèle sur lequel se fonde le filtre de Kalman ! Plus aucun comportement ne pourra être représenté : le découplage élimine tout.

Nous n'avons certes envisagé que des découplages parfaits. Néanmoins, les solutions moins drastiques consistant à désensibiliser, à l'aide des degrés de liberté disponibles, les hypomodèles par rapport aux incertitudes, tendent vers le découplage parfait. Le comportement certain aura alors plus de poids que le comportement incertain et là aussi, les objectifs sont implicitement modifiés.

Influence de l'horizon des hypomodèles en diagnostic

Désormais, pour rester cohérents avec les objectifs du diagnostic, nous partirons directement du métamodèle représentant le comportement de référence à tester. Pour simplifier les écritures, que le métamodèle définissant les objectifs soit réduit ou non, nous le représenterons par (I) : puisque sa structure est identique, le métamodèle réduit (18) se traitera de la même manière que (I) . Puisque nous excluons les découplages, la totalité du

métamodèle doit être exploitée. Il s'ensuit que la matrice W doit être inversible. Pour les mêmes raisons, la matrice S doit s'identifier à la matrice identité I_m . La matrice W peut être choisie égale à la matrice identité I_n puisque, dans le cas des hypomodèles à horizon fini, la fraction sélectionnée du métamodèle ne dépend pas de la valeur de cette matrice et, dans le cas des hypomodèles à horizon infini :

$$\varepsilon_k = H\hat{z}_k - Wx_k \xrightarrow{k \rightarrow \infty} 0 \Leftrightarrow W^{-1}\varepsilon_k = H\hat{z}_k - x_k \xrightarrow{k \rightarrow \infty} 0$$

avec $H' = W^{-1}H$

Enfin, pour simplifier les développements à venir, nous choisirons $H=I_n$ ce qui revient à ne considérer que les hypomodèles complets simples.

Avant d'étudier l'influence de l'horizon, recherchons quels sont les rapports entre les hypomodèles à horizon fini et les hypomodèles à horizon infini. Ce point a fait l'objet de nombreuses publications scientifiques. Massoumnia [Massoumnia,86] mit, le premier, en évidence les relations entre l'espace de parité et les observateurs d'état puis Wünnenberg et Frank [Wünnenberg,90] démontrèrent ces relations. De nombreuses autres publications que nous n'énumérerons pas traitèrent du même sujet en l'adaptant à certains cas particuliers. Signalons toutefois les travaux de Staroswiecki [Staroswiecki et al,91] dont l'approche temporelle est comparable à celle que nous développons ici.

Puisque $H=I_n$, l'hypomodèle (33) se simplifie en :

$$\hat{x}_{k+1} = (A - L_{11}C)\hat{x}_k + [B - L_{11}D \quad L_{11}] \begin{bmatrix} \tilde{u}_k \\ \tilde{y}_k \end{bmatrix} \quad (34a)$$

$$\varepsilon_k = \hat{x}_k - x_k \quad (34b)$$

En posant $\hat{y}_k = C\hat{x}_k$, nous déduisons de (34a), l'expression de \hat{y}_{k+h-1} :

$$\hat{y}_{k+h-1} = C(A - L_{11}C)^{h-1} \hat{x}_k + \sum_{i=0}^{h-2} C(A - L_{11}C)^{h-2-i} [B - L_{11}D \quad L_{11}] \begin{bmatrix} \tilde{u}_{k+i} \\ \tilde{y}_{k+i} \end{bmatrix} + D\tilde{u}_{k+h-1} \quad (35)$$

La condition d'existence [CI2] énonce que la paire (A,C) doit être détectable et les propriétés génériques [PII] et [PI2] attestent que, si (34) est un hypomodèle à horizon infini, les pôles de $A-L_{11}C$ sont strictement inclus dans le cercle unité et non nuls.

Dès lors, il existe une certaine valeur h telle que

$$\lim_{h \rightarrow \infty} (A - L_{11}C)^{h-1} = 0 \Rightarrow \exists h / (A - L_{11}C)^{h-1} \approx 0 \quad (36)$$

En introduisant ce résultat dans (35), nous en tirons :

$$\hat{y}_{k+h-1} \approx \sum_{i=0}^{h-2} C(A - L_{11}C)^{h-2-i} [B - L_{11}D \quad L_{11}] \begin{bmatrix} \tilde{u}_{k+i} \\ \tilde{y}_{k+i} \end{bmatrix} + D\tilde{u}_{k+h-1}$$

Puis, nous obtenons une relation qui ressemble à un hypomodèle à horizon fini :

$$\sum_{i=0}^{h-2} C(A - L_{11}C)^{h-2-i} (B - L_{11}D) \tilde{u}_{k+i} + D \tilde{u}_{k+h-1} + \sum_{i=0}^{h-2} C(A - L_{11}C)^{h-2-i} L_{11} \tilde{y}_{k+i} - \tilde{y}_{k+h-1} \approx 0 \quad (37)$$

Pour vérifier si l'hypomodèle obtenu correspond bien à un hypomodèle à horizon fini, nous nous proposons de s'assurer qu'il satisfait bien aux propriétés correspondantes. La propriété [PF1](p.146) est naturellement vérifiée puisqu'il n'est pas question ici de découplage. La propriété [PF2](p.148) demande un peu plus d'attention. Nous déduisons de (37) que le terme de gauche de la relation correspondant à cette propriété s'écrit :

$$\begin{aligned} & \left[\begin{array}{cccc} C(A - L_{11}C)^{h-2} L_{11} & \cdots & C(A - L_{11}C) L_{11} & CL_{11} & -I_p \end{array} \right] \begin{bmatrix} C \\ \vdots \\ CA^{h-3} \\ CA^{h-2} \\ CA^{h-1} \end{bmatrix} = \cdots \\ & \dots = \sum_{i=0}^{h-2} \left(C(A - L_{11}C)^{h-2-i} L_{11} CA^i \right) - CA^{h-1} \end{aligned} \quad (38)$$

Or, il est possible de montrer que :

$$\sum_{i=0}^{h-2} \left((A - L_{11}C)^{h-2-i} L_{11} CA^i \right) - A^{h-1} = -(A - L_{11}C)^{h-1}. \quad (39)$$

Démonstration

Nous procéderons par récurrence. Sans difficulté, nous montrons que la relation est vraie pour $h=2$ et $h=3$:

$$\text{pour } h = 2, L_{11}C - A = -(A - L_{11}C)$$

$$\text{pour } h = 3, (A - L_{11}C)L_{11}C + L_{11}CA - A^2 = -(A - L_{11}C)^2$$

Supposons que cette relation soit vraie au rang $h-2$:

$$\sum_{i=0}^{h-3} \left((A - L_{11}C)^{h-3-i} L_{11} CA^i \right) - A^{h-2} = -(A - L_{11}C)^{h-2}$$

Alors, au rang $h-1$, nous pouvons écrire :

$$\sum_{i=0}^{h-2} \left((A - L_{11}C)^{h-2-i} L_{11} CA^i \right) - A^{h-1} = (A - L_{11}C) \sum_{i=0}^{h-3} \left((A - L_{11}C)^{h-3-i} L_{11} CA^i \right) + L_{11}CA^{h-2} - A^{h-1}$$

Enfin, en utilisant la relation du rang $h-2$, nous montrons que la relation est vraie au rang $h-1$. En vertu du principe de récurrence, nous concluons que cette relation est vraie pour tout entier $h > 1$.

Enfin, en utilisant l'hypothèse (36), il vient :

$$\sum_{i=0}^{h-2} C \left((A - L_{11}C)^{h-2-i} L_{11} CA^i \right) - CA^{h-1} = -C(A - L_{11}C)^{h-1} \approx 0$$

Ceci qui nous conduit à conclure, en observant (37), que, pour une valeur h telle que l'hypothèse (36) soit vérifiée, la propriété [PF2] est quasiment vérifiée.

Examinons maintenant la propriété [PF3](p.149). Nous déduisons de (37) que :

$$\forall i \in \{0, \dots, h-2\}, Y_i = C(A - L_{11}C)^{h-i-2} L_{11} \text{ et } Y_{h-1} = -I_p \quad (40a)$$

$$\forall i \in \{0, \dots, h-2\}, U_i = \Pi_1 B - \Pi_2 D \text{ et } U_{h-1} = D \quad (40b)$$

$$\text{avec } \Pi_1 = C(A - L_{11}C)^{h-i-2} \text{ et } \Pi_2 = C(A - L_{11}C)^{h-i-2} L_{11}$$

Pour que la propriété [PF3] soit vérifiée, il faut que :

$$U_{h-1} = -Y_{h-1}D$$

$$\forall i \in \{0, \dots, h-2\}, U_i = -Y_i D - \sum_{j=0}^{h-i-2} (Y_{i+j+1} CA^j) B$$

On montre sans difficulté que $U_{h-1} = -Y_{h-1}D$. La seconde relation implique qu'il suffit que les relations suivantes soient satisfaites :

$$C(A - L_{11}C)^{h-i-2} = - \sum_{j=0}^{h-i-2} (Y_{i+j+1} CA^j)$$

$$C(A - L_{11}C)^{h-i-2} L_{11} = -Y_i$$

En utilisant (39), la première relation peut être reformulée. Nous aboutissons à la condition suffisante suivante :

$$C(A - L_{11}C)^{h-i-2} = CA^{h-i-2} - \sum_{j=0}^{h-i-3} (Y_{i+j+1} CA^j) \quad (41a)$$

$$C(A - L_{11}C)^{h-i-2} L_{11} = -Y_i \quad (41b)$$

En comparant l'expression de Y_i dans (41b) avec celle apparaissant dans (40a), nous constatons que cette condition suffisante est vérifiée. Concernant la relation (41a), développons le terme de droite et vérifions s'il est bien égal au terme de gauche. En substituant l'expression (40a) à Y_i , nous obtenons :

$$CA^{h-i-2} - \sum_{j=0}^{h-i-3} (Y_{i+j+1} CA^j) = CA^{h-i-2} - \sum_{j=0}^{h-i-3} ((A - L_{11}C)^{h-i-j-3} CA^j)$$

Puis, en utilisant la relation (39), nous établissons que (41b) est elle aussi satisfaite. Nous en concluons que la relation (37) vérifie la propriété générique [PF3].

Ces résultats sont fort intéressants. Pour qu'un hypomodèle à horizon infini soit équivalent à un hypomodèle à horizon fini, il faut que la propriété (36) soit strictement vérifiée. Nous pourrions imaginer que la matrice $A - L_{11}C$ soit nilpotente. Nous aurions en effet une équivalence stricte entre les deux types d'hypomodèles. Cependant, il faudrait pour cela que les pôles de cette matrice soient nuls, ce que la propriété [PII](p.143) exclut justement pour garantir un horizon infini. La seule solution théorique est de faire tendre h vers l'infini ; nous

en déduisons que théoriquement, un hypomodèle à horizon infini correspond à un hypomodèle à horizon fini dont l'horizon... tend vers l'infini. Bien qu'apparemment trivial, ce résultat est lourd de sens. Il signifie qu'un hypomodèle à horizon fini minimal permet non seulement de générer n'importe quel hypomodèle à horizon fini, mais aussi qu'il permet d'obtenir les résidus des hypomodèles à horizon infini. Bien sûr, il n'est guère envisageable, pratiquement, d'atteindre un horizon infini avec un hypomodèle à horizon fini, mais la propriété (36) sera quasiment vérifiée pour un horizon h suffisamment grand, tributaire évidemment des pôles de la matrice $A-L_{11}C$. Un hypomodèle à horizon infini, par contre, ne permet pas de générer d'hypomodèle à horizon fini.

Pour mieux comprendre l'influence de l'horizon d'un hypomodèle sur les tests de cohérences nous emprunterons un exemple simple à Chow [Chow,80].

Exemple

Intéressons-nous à un métamodèle en temps discret représentant la relation entre l'accélération linéaire d'un solide et sa vitesse. Pour simplifier les écritures, nous avons normalisé le modèle :

$$\begin{aligned} x_{k+1} &= x_k + u_k \\ y_k &= x_k \end{aligned}$$

où u_k est homogène à une accélération tandis que x_k et y_k sont homogènes à une vitesse.

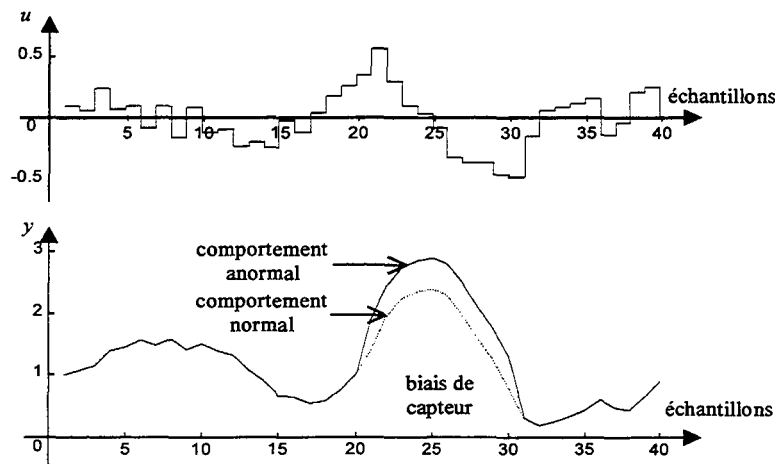


Figure I – Observation des variables u et y

La **figure I** représente un comportement simulé sur 40 échantillons. De l'instant 21 à l'instant 30, nous avons simulé un biais de capteur d'une valeur positive égale à 0,5.

Un hypomodèle à horizon fini minimal se déduit directement. Il est parfaitement caractérisé par les matrices intervenant dans (11a) :

$$Y = \begin{bmatrix} 1 & -1 \end{bmatrix} \text{ et } U = \begin{bmatrix} 1 & 0 \end{bmatrix} \Leftrightarrow r_{k+1} = \tilde{y}_k - \tilde{y}_{k+1} + \tilde{u}_k$$

Nous avons aussi construit un hypomodèle à horizon infini simple dont le pôle est égal à 0.5 :

$$\hat{y}_{k+1} = 0.5\hat{y}_k + \tilde{u}_k + 0.5\tilde{y}_k$$

$$r'_k = \hat{y}_k - \tilde{y}_k$$

La figure II représente les résultats obtenus. Nous constatons que les résidus obtenus de l'hypomodèle à horizon fini minimal ne révèlent le défaut qu'aux instants d'apparition et de disparition ; lorsque le biais est maintenu, les résidus indiquent une cohérence entre les observations et le comportement normal. Puisqu'un hypomodèle à horizon fini minimal est générateur de tous les autres hypomodèles, nous en déduisons qu'aucun hypomodèle ne peut révéler une incohérence de comportement durant le maintien d'un biais constant. Ceci s'explique par le fait que lorsque le biais se maintient, tout se passe comme s'il s'était produit un saut de valeur sur l'état : $x_k \rightarrow x_k + 0.5$. Il s'ensuit que dès l'instant suivant le biais, le comportement est le même que celui qui aurait été obtenu en partant d'une autre valeur de l'état. Cette constatation rejoint ce que nous avons énoncé au cours du premier chapitre : une cohérence est moins prégnante qu'une incohérence dans la mesure où des défauts peuvent engendrer des comportements s'apparentant à des comportements normaux.

Certes, le résidu issu de l'hypomodèle à horizon infini semble fallacieusement mieux révéler l'incohérence. Mais il ne s'agit que d'une apparence car un tel hypomodèle conduit à des résidus qui peuvent être déduits de ceux de l'hypomodèle à horizon fini minimal. Si l'incohérence révélée est plus longue, ce n'est que parce l'incohérence révélée à l'instant d'apparition du biais est propagée sur une certaine durée en rapport avec la vitesse de convergence de l'hypomodèle. La preuve en est que lors de la disparition du biais, une incohérence est révélée durant une durée comparable¹⁸.

Enfin, pour étayer nos allégations concernant le rapport entre les hypomodèles à horizon infini et les hypomodèles à horizon fini suffisamment long, nous avons déterminé un hypomodèle à horizon de 8 approchant l'hypomodèle à horizon

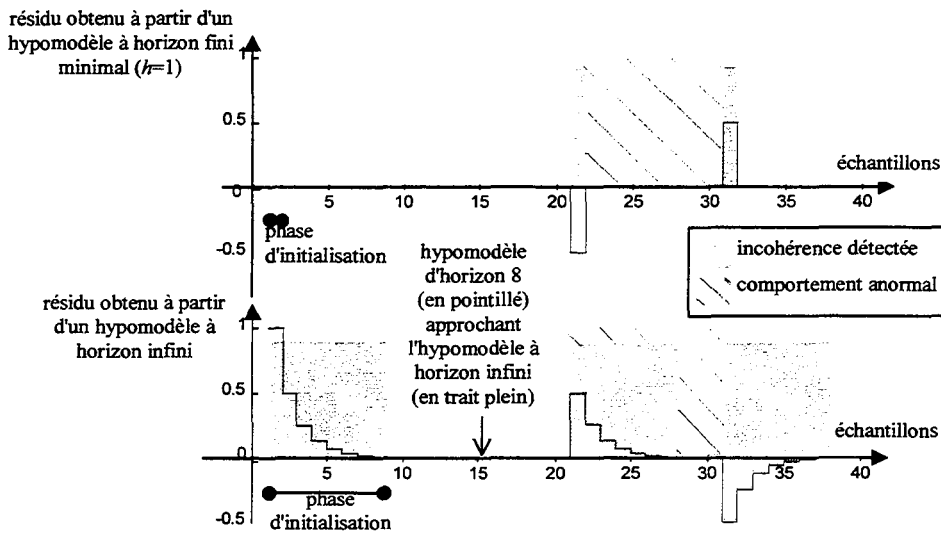


Figure II – Comparaison des résidus obtenus

¹⁸ Remarquons qu'une incohérence dans un système dynamique ne doit pas être interprétée, notamment si l'horizon de l'hypomodèle est important, comme "il y a une incohérence" mais comme "il s'est produit une incohérence".

infini. L'hypothèse (36) conduit à $(A-LC)^7=7,81.10^{-3}\approx 0$ est quasiment vérifiée pour cet horizon. L'hypomodèle est déterminé par les matrices de (11a) suivantes :

$$Y^{(8)} = \begin{bmatrix} 0,0078 & 0,0156 & 0,0313 & 0,0625 & 0,125 & 0,25 & 0,5 & -1 \end{bmatrix}$$

$$U^{(8)} = \begin{bmatrix} 0,0156 & 0,0313 & 0,0625 & 0,125 & 0,25 & 0,5 & 1 & 0 \end{bmatrix}$$

Les résultats obtenus sont effectivement comparables à ceux de l'hypomodèle à horizon infini. Ils apparaissent en pointillés sur la moitié inférieure de la figure II.

Une opération de diagnostic est un test de cohérence entre un comportement observé et un comportement de référence. Pour que l'opération soit correctement réalisée, il est préférable de formaliser directement le comportement de référence par un métamodèle adapté puis de déduire un hypomodèle en veillant à n'opérer à aucun découplage. De nombreux degrés de liberté apparaissent lors de la synthèse d'un hypomodèle. Or, si les autres comportements possibles étaient connus et formalisés par des métamodèles, ces degrés de liberté pourraient être mis à profit pour que les résultats des tests de cohérence soient sensibles aux autres comportements. Cependant, nous l'avons évoqué au cours du premier chapitre, il est pratiquement exceptionnel de disposer de plusieurs métamodèles, et notamment de ceux susceptibles de représenter des mauvais fonctionnements. De plus, même si de tels modèles étaient disponibles, il est rare que leur structure se rapproche de celle d'autres métamodèles associés au système physique ausculté. La seule exception concerne certains défauts de capteurs et a été largement étudiée [Clark,78] [Clark,89]. En conséquence, nous concluons que, dans le cas général, il est préférable d'exploiter un hypomodèle à horizon fini minimal puisque, en plus de posséder une structure simple, il contient toute l'information sur le résultat du test de cohérence. Les résidus ainsi calculés permettent d'obtenir, par filtrage, des effets d'intégration facilitant parfois la lecture des résultats en rendant les événements moins fugaces.

Les conclusions de cette première sous-partie ne se limitent pas aux hypomodèles certains puisqu'un découplage engendre toujours un comportement représenté plus local, et la taille de l'horizon d'un hypomodèle incertain ne peut qu'influer sur les résultats dans le même sens que dans le cas certain. Le fait que les matrices des hypomodèles soient certaines ou pas, ne modifie pas ces résultats. Achéons enfin cette sous-partie en rappelant que si la nature causale est déterminée dans le cas d'hypomodèles à horizon infini, elle ne l'est pas pour les horizons finis. Bien que, traditionnellement, les hypomodèles certains à horizon fini causaux soient préférés, la nature causale ou acausale n'affecte en rien les résultats des tests de cohérence.

III. Les hypomodèles incertains

Les approches reposant sur des modèles certains présentent l'avantage d'être très appropriées à l'écriture mathématique. Pour peu que les métamodèles soient linéaires, les possibilités de calcul d'hypomodèles sont extrêmement importantes : dans la plupart des cas, des solutions

analytiques peuvent être obtenues. Cependant, comme nous l'avons souligné dans le premier chapitre, il n'y a, dans la nature, qu'une infime minorité de systèmes physiques dont un comportement peut être représenté avec justesse par un métamodèle certain. Or, si ne représenter qu'un comportement moyen suffit souvent au calcul de commande automatisée, le diagnostic requiert de représenter des ensembles de comportements possibles. Les hypomodèles incertains deviennent alors incontournables. De la même manière que dans le cas certain, deux types d'hypomodèles peuvent être distingués : ceux dont l'horizon est infini et ceux dont l'horizon est fini. Notre objectif n'est pas de recenser toutes les manières de générer ces hypomodèles mais plutôt d'en rechercher les particularités intéressantes en vue d'une application au diagnostic.

Cette partie se décompose en trois sous-parties. La première propose une formulation générique des métamodèles dynamiques incertains ainsi qu'une formulation réaliste qui soit compatible avec la réalité physique. Nous partirons de la description comportementale en temps continu pour parvenir au métamodèle incertain en temps discret. La seconde sous-partie soulève les problèmes de la reconstruction d'un ensemble de comportements possibles au moyen d'hypomodèles à horizon infini. Nous développerons certaines méthodes permettant d'évaluer les domaines abstraits des variables physiques en procédant par intégration directe des équations d'un métamodèle. Nous mettrons en évidence le *phénomène d'enveloppement* à l'origine de problèmes souvent rédhibitoires puis nous évoquerons brièvement les approches elliptiques qui ont connu de nombreux développements dans la littérature. La troisième sous-partie aborde les hypomodèles incertains à horizon fini. Nous aborderons le problème de réalisation de test de cohérence dans le cas où la formulation de l'hypomodèle est connue puis, dans un second temps, nous montrerons qu'il est difficile d'obtenir une formulation simple en partant d'un métamodèle sous forme d'équations différentielles ordinaires. Nous proposerons une méthode conduisant à une formulation approchée simple. Enfin, au travers d'exemples, nous montrerons les différences induites par les deux natures temporelles de variables abstraites.

III.1. Formulation générique des métamodèles dynamiques incertains

L'écriture générique d'un métamodèle linéaire incertain s'inspire de la structure du métamodèle certain, mais en plus, il faut considérer que les matrices qui y interviennent peuvent contenir des variables abstraites et que la présence de termes additifs supplémentaires contenant des variables abstraites non déterministes, est possible. Tous les métamodèles linéaires en temps continu sont représentés par :

$$\frac{dx_0(t)}{dt} = A_c(\theta(t))x_0(t) + B_c(\theta(t))u(t) + E_c(\theta(t)) \quad (42a)$$

$$y(t) = C(\theta(t))x_0(t) + D_c(\theta(t))u(t) + F_c(\theta(t)) \quad (42b)$$

où $E(c(\theta(t)))=0$, $F(c(\theta(t)))=0$, $\theta(t)$ est un vecteur de r variables abstraites indépendantes¹⁹
et les dimensions des autres vecteurs sont conformes à celle du métamodèle (1).

L'hypothèse d'indépendance des variables abstraites est importante car, dans le cas contraire, les relations entre ces variables seraient ignorées. Certes, les métamodèles servant de base au filtre de Kalman sont une exception car les variables abstraites peuvent être dépendantes sans que les relations les liant soient explicitement connues. Les matrices de variance covariance associées aux vecteurs $E_c(\theta(t))$ et $F_c(\theta(t))$ suffisent à exprimer ces dépendances. Néanmoins, même dans ce cas particulier, il est possible de trouver une fonction linéaire de variables abstraites indépendantes conduisant à une matrice de variance covariance donnée et ainsi de se ramener à l'hypothèse d'indépendance.

Le métamodèle doit être complété en décrivant la nature et les caractéristiques des variables abstraites $\theta(t)$. Tout d'abord, il est important de distinguer les variables abstraites invariantes des autres (cf. partie IV du chapitre 2). Dans le cas de variables abstraites stochastiques, les espaces abstraits et les mesures probabilistes doivent être précisés pour chacune des variables $\theta_i(t)$ de $\theta(t)$: $\theta_i(t)$ est caractérisée par un espace probabilisé $(\Theta_i, \mathbf{B}(\Theta_i), p_i(\cdot))$. Pour ce qui est des approches bornantes, la seule connaissance des domaines abstraits suffit :

$$\mathcal{A}(\theta(t)) = \square \mathcal{A}(\theta(t)) = \Theta$$

Nous avons rappelé, au cours de l'introduction de la partie II, que les métamodèles en temps discret étaient plus intéressants que ceux en temps continu. Pour pouvoir intégrer le métamodèle continu, il est nécessaire de représenter le comportement des variables abstraites non déterministes entre deux instants d'échantillonnage. Or, l'indétermination de variables abstraites est problématique. Imaginons qu'entre deux instants consécutifs, $k\tau$ et $(k+1)\tau$, l'évolution des variables abstraites soit contenue dans des secteurs partant de $\theta(k\tau)$. Le cheminement des variables de $\theta(t)$ entre ces deux instants pourrait être représentée par :

$$\theta(t) = \theta(k\tau) + t \frac{\delta(k\tau)}{\tau} ; \mathcal{A}(\delta(k\tau)) = \Delta \text{ et } \mathcal{A}(\theta(t)) = \Theta \quad \forall t \in [k\tau, (k+1)\tau]$$

Le domaine Δ est un orthotope aligné²⁰ représentant les coefficients directeurs possibles des droites d'évolution de $\theta(t)$. Puisque nous sommes dans un cadre non déterministe, la valeur de $\theta(k\tau)$ est inconnue et la connaissance effective porte sur les domaines abstraits : $\mathcal{A}(\theta(t)) = \Theta + t\Delta$. Or, à tout instant $\mathcal{A}(\theta(t)) = \Theta$; dès lors, pour ne pas conduire à une absurdité, il faut se restreindre au cas où $\Delta = \{0\}$. Ce problème rejoint celui des variables abstraites à variation bornée évoquées dans la partie IV.3 du chapitre 2. Il est par conséquent nécessaire

¹⁹ L'indépendance n'est pas une contrainte puisqu'il est toujours possible de transformer un vecteur de variables dépendantes en une fonction de variables indépendantes.

²⁰ Du fait de l'indépendance des variables de $\theta(t)$.

de considérer que sur une période d'échantillonnage, la valeur de toute variable abstraite reste invariante, c'est-à-dire : $\forall t \in [k\tau, (k+1)\tau[, \theta(t) = \theta(k\tau)$.

Grâce à cette hypothèse, le métamodèle continu peut, tout comme pour les métamodèles certains, être intégré sur une période d'échantillonnage. Le métamodèle incertain discret s'apparente naturellement au métamodèle certain (I). Pour se ramener à une forme standard, normalisons les variables abstraites de $\theta(t)$ tel que nous le fimes dans la partie III.1 du chapitre 2 (16), en posant :

$$\theta(t) = c(\theta(t)) + R(\theta(t))v(t)$$

avec $v(t)/\|v(t)\|_{\infty} \leq 1$ *et* $R(\theta(t)) = \text{diag}([\rho_1 \rho_2 \dots \rho_{\dim(\theta(t))}])$

Le métamodèle discret normalisé s'écrira donc :

$$x_{k+1} = A(v_k)x_k + B(v_k)u_k + E(v_k) \tag{43a}$$

$$y_k = C(v_k)x_k + D(v_k)u_k + F(v_k) \tag{43b}$$

où $u_k = u(k\tau)$, $x_k = x(k\tau)$, $y_k = y(k\tau)$ *et* $v_k = v(k\tau)$

Dans le cas où les entrées $u(t)$ peuvent être considérées comme constantes par morceaux (2a), les matrices du métamodèle (43) se déduisent de (2b) :

$$\begin{aligned} A(v_k) &= \sum_{i=0}^{\infty} \frac{A_c(c(\theta_k) + R(\theta_k)v_k)^i \tau^i}{i!}, \\ B(v_k) &= \sum_{i=0}^{\infty} \frac{A_c(c(\theta_k) + R(\theta_k)v_k)^i \tau^{i+1}}{(i+1)!} B_c(c(\theta_k) + R(\theta_k)v_k) \\ C(v_k) &= C_c(c(\theta_k) + R(\theta_k)v_k) \\ D(v_k) &= D_c(c(\theta_k) + R(\theta_k)v_k) \\ E(v_k) &= \sum_{i=0}^{\infty} \frac{A_c(c(\theta_k) + R(\theta_k)v_k)^i \tau^{i+1}}{(i+1)!} E_c(c(\theta_k) + R(\theta_k)v_k) \\ F(v_k) &= \sum_{i=0}^{\infty} \frac{A_c(c(\theta_k) + R(\theta_k)v_k)^i \tau^{i+1}}{(i+1)!} F_c(c(\theta_k) + R(\theta_k)v_k) \end{aligned} \tag{44a}$$

où $\theta_k = \theta(k\tau)$

Dans le cas d'entrées $u(t)$ linéaires par morceaux décrites par (3a), les matrices découlent de (3b) :

$$\begin{aligned} A(v_k) &= \sum_{i=0}^{\infty} \frac{A_c(c(\theta_k) + R(\theta_k)v_k)^i \tau^i}{i!} \\ B(v_k) &= \sum_{i=0}^{\infty} \frac{(2^{i+2} - 2)A_c(c(\theta_k) + R(\theta_k)v_k)^i \tau^{i+1}}{(i+2)!} B_c(c(\theta_k) + R(\theta_k)v_k) \\ C(v_k) &= C_c(c(\theta_k) + R(\theta_k)v_k) \\ D(v_k) &= C_c(c(\theta_k) + R(\theta_k)v_k) \sum_{i=0}^{\infty} \frac{A_c(c(\theta_k) + R(\theta_k)v_k)^i \tau^{i+1}}{(i+2)!} B_c(c(\theta_k) + R(\theta_k)v_k) + \dots \\ &\quad \dots + D_c(c(\theta_k) + R(\theta_k)v_k) \end{aligned} \tag{44b}$$

$$E(v_k) = \sum_{i=0}^{\infty} \frac{(2^{i+2} - 2)A_c(c(\theta_k) + R(\theta_k)v_k)^i \tau^{i+1}}{(i+2)!} E_c(c(\theta_k) + R(\theta_k)v_k)$$

$$F(v_k) = F_c(c(\theta_k) + R(\theta_k)v_k)$$

Ces écritures peuvent être simplifiées en y introduisant des exponentielles de matrices ; mais le calcul d'espace abstrait requiert des expressions analytiques ne faisant intervenir que des opérateurs simples. Comment en effet calculer un orthotope circonscrit sans exploiter les formules de l'arithmétique des intervalles (cf. chapitre 2 de (7) à (13)) sinon par des techniques de recherche d'extrema, fort coûteuses en temps de calcul ? Deux solutions peuvent être envisagées pour simplifier les expressions des matrices du métamodèle discret. La plus simple, et la plus imprécise aussi, consiste à limiter les sommations de (44) à un ordre fini relativement faible, tandis que la plus complexe revient à calculer une approximation au premier ordre de l'influence des variables abstraites non déterministes sur les matrices du métamodèle : c'est celle que nous retiendrons. En remarquant que :

$$\left. \frac{\partial A(c(\theta_k) + R(\theta_k)v_k)^\alpha}{\partial v_{i,k}} \right|_{v_k=0} = \rho_i \sum_{\beta=1}^{\alpha} A_c(c(\theta_k))^{\beta-1} \left. \frac{\partial A_c}{\partial \theta_{i,k}} \right|_{\theta_k=c(\theta_k)} A_c(c(\theta_k))^{\alpha-\beta}$$

l'approximation du métamodèle en temps discret (43) peut être déduite :

$$x_{k+1} = \left(A_0 + \sum_{i=1}^r A_i v_{i,k} \right) x_k + \left(B_0 + \sum_{i=1}^r B_i v_{i,k} \right) u_k + \sum_{i=1}^r E_i v_{i,k} \tag{45a}$$

$$y_k = \left(C_0 + \sum_{i=1}^r C_i v_{i,k} \right) x_k + \left(D_0 + \sum_{i=1}^r D_i v_{i,k} \right) u_k + \sum_{i=1}^r F_i v_{i,k} \tag{45b}$$

où $A_0=A(0)$, $B_0=B(0)$, $C_0=C(0)$, $D_0=D(0)$,
 $E_0=0$ et $F_0=0$ car $E(c(\theta(t)))=0$ et $F(c(\theta(t)))=0$

Les matrices indicées i dépendent de la nature de l'entrée ; selon qu'elle est décrite par (2a) ou par (3a). Dans le premier cas, ces matrices se déduisent de (44a) ; elles sont données par :

$$A_i = \rho_i \sum_{\alpha=1}^{\infty} \sum_{\beta=1}^{\alpha} \left(A_c(c(\theta_k))^{\beta-1} \left. \frac{\partial A_c}{\partial \theta_{i,k}} \right|_{\theta_k=c(\theta_k)} A_c(c(\theta_k))^{\alpha-\beta} \frac{\tau^\alpha}{\alpha!} \right)$$

$$B_i = \rho_i \sum_{\alpha=1}^{\infty} \sum_{\beta=1}^{\alpha} \left(A_c(c(\theta_k))^{\beta-1} \left. \frac{\partial A_c}{\partial \theta_{i,k}} \right|_{\theta_k=c(\theta_k)} A_c(c(\theta_k))^{\alpha-\beta} \frac{\tau^{\alpha+1}}{(\alpha+1)!} \right) B_c(c(\theta_k)) + \dots$$

$$\dots + \rho_i \sum_{i=0}^{\infty} \left(A_c(c(\theta_k))^\alpha \frac{\tau^{\alpha+1}}{(\alpha+1)!} \right) \left. \frac{\partial B_c}{\partial \theta_{i,k}} \right|_{\theta_k=c(\theta_k)}$$

$$E_i = \rho_i \sum_{\alpha=0}^{\infty} \left(A_c(c(\theta_k))^\alpha \frac{\tau^{\alpha+1}}{(\alpha+1)!} \right) \left. \frac{\partial E_c}{\partial \theta_{i,k}} \right|_{\theta_k=c(\theta_k)} \tag{46}$$

$$C_i = \rho_i \frac{\partial C_c}{\partial \theta_{i,k}}, D_i = \rho_i \frac{\partial D_c}{\partial \theta_{i,k}} \text{ et } F_i = \rho_i \left. \frac{\partial F_c}{\partial \theta_{i,k}} \right|_{\theta_k=c(\theta_k)}$$

avec $\rho_i = \rho(\theta_{i,k})$ et $\theta_k = [\theta_{1,k} \theta_{2,k} \dots \theta_{r,k}]^T$ et $v_k = [v_{1,k} v_{2,k} \dots v_{r,k}]^T$

Dans le cas d'une commande linéaire par morceaux, les matrices du métamodèle (45) se déduisent des expressions (44b) :

$$\begin{aligned}
 A_i &= \rho_i \sum_{\alpha=1}^{\infty} \sum_{\beta=1}^{\alpha} \left(A_c(c(\theta_k))^{\beta-1} \frac{\partial A_c}{\partial \theta_{i,k}} \Big|_{\theta_k=c(\theta_k)} A_c(c(\theta_k))^{\alpha-\beta} \frac{\tau^\alpha}{\alpha!} \right) \\
 B_i &= \rho_i \sum_{\alpha=1}^{\infty} \sum_{\beta=1}^{\alpha} \left(A_c(c(\theta_k))^{\beta-1} \frac{\partial A_c}{\partial \theta_{i,k}} \Big|_{\theta_k=c(\theta_k)} A_c(c(\theta_k))^{\alpha-\beta} \frac{(2^{\alpha+2} - 2)\tau^{\alpha+1}}{(\alpha+2)!} B_c(c(\theta_k)) + \dots \right. \\
 &\quad \left. \dots + \rho_i \sum_{i=0}^{\infty} \left(A_c(c(\theta_k))^\alpha \frac{(2^{\alpha+2} - 2)\tau^{\alpha+1}}{(\alpha+2)!} \right) \frac{\partial B_c}{\partial \theta_{i,k}} \Big|_{\theta_k=c(\theta_k)} \right) \\
 D_i &= \rho_i \frac{\partial D_c}{\partial \theta_{i,k}} \Big|_{\theta_k=c(\theta_k)} + \rho_i \frac{\partial C_c}{\partial \theta_{i,k}} \Big|_{\theta_k=c(\theta_k)} \sum_{\alpha=0}^{\infty} A_c(c(\theta_k))^\alpha \frac{\tau^{\alpha+1}}{(\alpha+2)!} + \dots \\
 &\quad \dots + \rho_i C_c(c(\theta_k)) \sum_{\alpha=0}^{\infty} A_c(c(\theta_k))^\alpha \frac{\tau^{\alpha+1}}{(\alpha+2)!} \frac{\partial B_c}{\partial \theta_{i,k}} \Big|_{\theta_k=c(\theta_k)} \\
 &\quad \dots + \rho_i C_c(c(\theta_k)) \sum_{\alpha=1}^{\infty} \sum_{\beta=1}^{\alpha} A_c(c(\theta_k))^{\beta-1} \frac{\partial A_c}{\partial \theta_{i,k}} \Big|_{\theta_k=c(\theta_k)} A_c(c(\theta_k))^{\alpha-\beta} \frac{\tau^{\alpha+1}}{(\alpha+2)!} B_c(c(\theta_k)) \\
 E_i &= \rho_i \sum_{\alpha=0}^{\infty} \left(A_c(c(\theta_k))^\alpha \frac{(2^{\alpha+2} - 2)\tau^{\alpha+1}}{(\alpha+2)!} \right) \frac{\partial E_c}{\partial \theta_{i,k}} \Big|_{\theta_k=c(\theta_k)} \tag{47} \\
 C_i &= \rho_i \frac{\partial C_c}{\partial \theta_{i,k}} \text{ et } F_i = \rho_i \frac{\partial F_c}{\partial \theta_{i,k}} \Big|_{\theta_k=c(\theta_k)}
 \end{aligned}$$

Ces matrices font intervenir des sommations jusqu'à l'infini ce qui peut paraître gênant pour leur détermination pratique. Néanmoins, les termes de ces sommations tendent vers 0 lorsque l'indice augmente. Autrement dit, il suffit d'effectuer les sommes jusqu'à un ordre pour lequel le terme supplémentaire est négligeable devant la valeur de la matrice déjà calculée. Et puis, il faut avoir conscience que ces matrices ne sont évaluées qu'une fois et avant toute mise en œuvre d'algorithmes de diagnostic. Pratiquement, moins d'une dizaine de termes consécutifs suffisent dans la plupart des cas.

Exemple

Le métamodèle, dont les variables physiques ont été normalisées par rapport à leur valeur nominale, correspond à une machine à courant continu d'une dizaine de kilowatt décrite dans [Maksimovic et al,93]. Nous reviendrons plus amplement sur ce modèle dans le chapitre 5. Nous supposons dans un premier temps que la tension, le courant et la vitesse normalisés sont connus de manière exacte. La valeur de la résistance équivalente fluctue de plus ou moins 50% autour d'une valeur centrale et le frottement équivalent, de plus ou moins 20% autour de la valeur centrale identifiée : nous représenterons les caractéristiques de ces incertitudes par les paramètres $\rho_1=0,5$ et $\rho_2=0,2$. La structure du métamodèle est

conforme au métamodèle générique en temps continu (42) ; les matrices non nulles de ce modèle sont les suivantes :

$$A_c(v(t)) = \begin{bmatrix} \frac{-1}{\tau_1} & \frac{\gamma-1}{\gamma\tau_1} \\ \frac{1}{\tau_2} & \frac{-\mu}{\tau_2} \end{bmatrix} + \rho_1 \begin{bmatrix} \frac{-1}{\tau_1} & 0 \\ 0 & 0 \end{bmatrix} v_1(t) + \rho_2 \begin{bmatrix} 0 & 0 \\ 0 & \frac{-\mu}{\tau_2} \end{bmatrix} v_2(t) \quad (48)$$

$$B_c(v(t)) = \begin{bmatrix} 1 \\ \gamma\tau_1 \\ 0 \end{bmatrix} \text{ et } C_c(v(t)) = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

où $x_i = [\Omega(t) \ i(t)]^T$; $\Omega(t)$ est la vitesse de rotation normalisée, $i(t)$ est le courant normalisé et $u(t)$ la tension d'alimentation normalisée, $\tau_1 = L/R$ est la constante de temps électrique $\tau_2 = J/f_n$ est la constante de temps mécanique, $\gamma = R_n i_n / u_n$ et $\mu = f/f_n$ et les paramètres équivalents R_n et f_n représentent respectivement la résistance électrique, et le frottement en fonctionnement nominal (cf chapitre 1 (3) et chapitre 5 (2))²¹
 L est le coefficient d'inductance propre du rotor et J l'inertie de la partie tournante.

Les paramètres du métamodèle ont été identifiés : $\tau_1 = 13\text{ms.}$, $\tau_2 = 1\text{s.}$, $\gamma = 0.118$ et $\mu = 1$. Cette machine est commandée en tension au moyen d'un micro-contrôleur envoyant des commandes à la période d'échantillonnage $\tau_e = 10\text{ms.}$ La valeur d'une commande étant naturellement constante sur une période d'échantillonnage τ_e , nous utiliserons les expressions (46) pour déduire les matrices d'un métamodèle en temps discret. Les matrices non nulles du métamodèle en temps discret (45) sont donc les suivantes :

$$A_0 \cong \begin{bmatrix} 0.5404 & -3.2817 \\ 7.5 \cdot 10^{-3} & 0.9720 \end{bmatrix} \quad A_1 \cong \begin{bmatrix} -0.1596 & -0.6541 \\ -10^{-3} & 2.8 \cdot 10^{-3} \end{bmatrix} \quad A_2 \cong \begin{bmatrix} 0 & 9 \cdot 10^{-3} \\ -2 \cdot 10^{-5} & -4.9 \cdot 10^{-3} \end{bmatrix}$$

$$B_0 \cong \begin{bmatrix} 3.7413 \\ 2.05 \cdot 10^{-2} \end{bmatrix} \quad B_1 \cong \begin{bmatrix} -0.4964 \\ -1.8 \cdot 10^{-3} \end{bmatrix} \quad B_2 \cong \begin{bmatrix} 3.608 \cdot 10^{-5} \\ -3.575 \cdot 10^{-5} \end{bmatrix} \quad C_0 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \quad (49)$$

La figure III permet d'apprécier la qualité du métamodèle approximé puisqu'il compare, dans le cas de la machine à courant continu, les réponses calculées à partir du modèle discret approché (45) (c'est-à-dire les matrices (49)) aux réponses obtenues en réévaluant à chaque instant les expressions (44a) dans (43). Dans toutes les simulations associées à cet exemple, la valeur de la résistance part d'un maximum à l'origine ($v_1(0)=1$), décroît régulièrement vers un minimum à la moitié du temps de simulation ($v_1=-1$) pour rejoindre suivant une progression toujours régulière un maximum à la fin. A l'inverse, la valeur du frottement part d'un minimum ($v_2(0)=-1$), atteint un maximum ($v_2=1$) à mi-simulation et rejoint son minimum à la fin, suivant une progression régulière. Une commande normalisée comprise entre 0,5 et 1,5 a été appliquée lors des simulations. Les réponses obtenues à partir des deux métamodèles en temps discret sont quasiment confondues.

²¹ L'indice n vaut pour nominal.

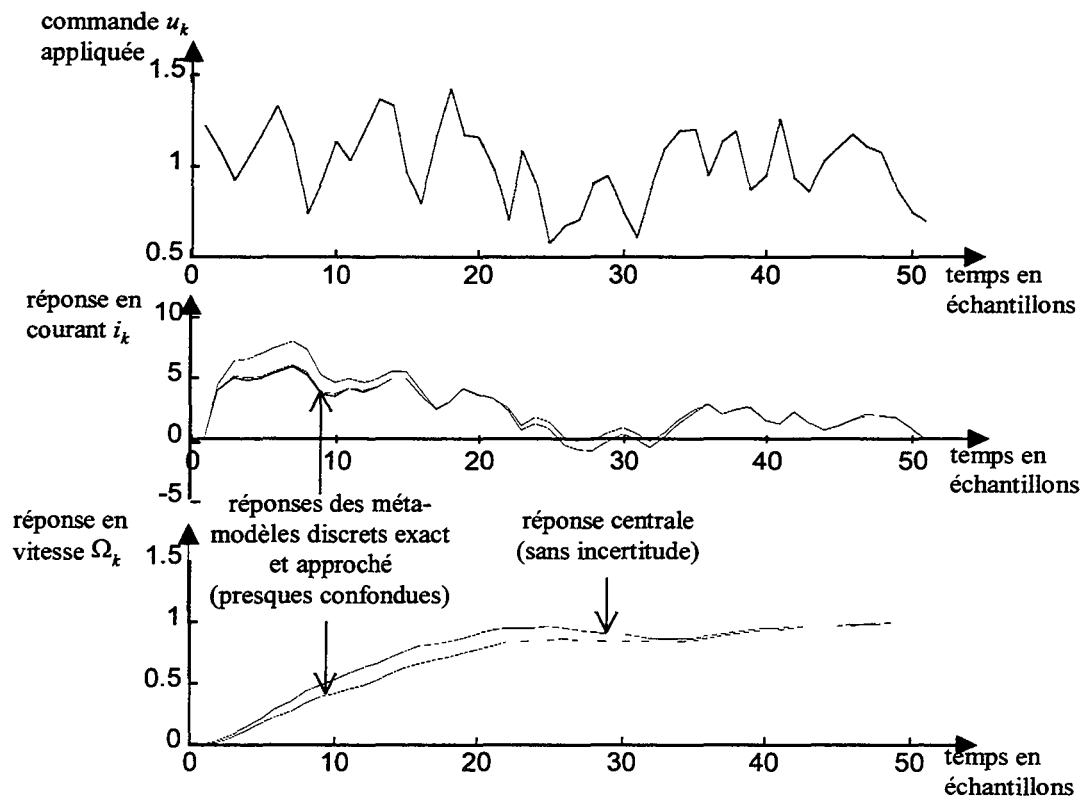


Figure III – Comparaison entre les réponses simulées avec un métamodèle en temps discret non approximé et approximé

Plutôt que de partir d'une formulation générique des métamodèles en temps discret où les matrices d'état seraient des fonctions quelconques de variables abstraites non déterministes, il est plus réaliste de se restreindre à la formulation (45) dont les matrices sont données par (46) ou par (47). Il s'agit en effet, à l'exception de méthodes très approximatives que nous avons évoquées précédemment, des seules formulations analytiques qui peuvent être déduites des métamodèles en temps continu. Cette formulation se distingue de celles de certains auteurs [Barmish,79] qui ne décrivent les matrices d'état que par les espaces abstraits de chacun de leurs éléments : une matrice M est décrite par $\mathcal{A}(M_{ij}) = [\tilde{M}_{ij}, \hat{M}_{ij}] \forall i, j^{22}$. En effet, de telles formulations ne prennent nullement en compte les dépendances entre les éléments des matrices et sont donc, en ce sens, très particulières.

III.2. Les hypomodèles incertains à horizon infini

Formulation générique

Tout comme dans le cas certain, les hypomodèles à horizon infini procèdent d'une intégration directe du métamodèle. Or, nous avons montré à partir du cas certain, qu'en ne reconstruisant qu'une fraction de l'état, les objectifs du diagnostic se trouvaient modifiés. Nous

²² M_{ij} est l'élément de la ligne i et de la colonne j de la matrice M .

considérerons donc que le métamodèle (45) représente exclusivement la fraction du comportement dont il faut tester la cohérence avec les observations.

L'intégration conduisant à un état reconstruit, il est possible d'introduire une contre-réaction dans l'hypomodèle. Par conséquent, tout comme dans le cas certain, tout hypomodèle incertain à horizon infini pourra être représenté par une expression similaire à (33) à laquelle s'ajoutent des variables abstraites non déterministes. Nous obtenons :

$$\begin{aligned} \begin{bmatrix} \hat{x}_{k+1} \\ \hat{\xi}_{k+1} \end{bmatrix} &= \begin{bmatrix} A_0 - L_{11}C_0 + \sum_{i=1}^r (A_i - L_{11}C_i)v_{i,k} & L_{12} \\ -L_{12} \left(C_0 + \sum_{i=1}^r C_i v_{i,k} \right) & L_{22} \end{bmatrix} \begin{bmatrix} \hat{x}_k \\ \hat{\xi}_k \end{bmatrix} + \begin{bmatrix} B_0 - L_{11}D_0 + \sum_{i=1}^r (B_i - L_{11}D_i)v_{i,k} \\ 0 \end{bmatrix} \tilde{u}_k + \dots \\ &\dots + \begin{bmatrix} L_{11} \\ L_{12} \end{bmatrix} \tilde{y}_k + \begin{bmatrix} \sum_{i=1}^r (E_i - LF_i)v_{i,k} \\ 0 \end{bmatrix} \\ \hat{y}_k &= \left(C_0 + \sum_{i=1}^r (C_i v_{i,k}) \right) \hat{x}_k + \left(D_0 + \sum_{i=1}^r (D_i v_{i,k}) \right) \tilde{u}_k + \sum_{i=1}^r (F_i v_{i,k}) \end{aligned}$$

La structure de cet hypomodèle est bien plus complexe que celle du métamodèle et ceci avec le seul avantage d'offrir des degrés de liberté supplémentaires (dus aux matrices L_{11} , L_{12} , L_{21} et L_{22}). Or, nous l'avons vu dans le cas certain, nous ne savons pas comment les exploiter dans le cas le plus courant dans la pratique où les métamodèles représentant le mauvais fonctionnement sont inexistantes ou de structures différentes. Nous verrons que la reconstruction de domaines abstraits bornés demande généralement un temps important, d'autant plus important que le nombre de variables abstraites non déterministes intervenant dans une équation à intégrer l'est. Pour ces différentes raisons, nous nous restreindrons à l'hypomodèle à horizon infini le plus simple :

$$\hat{x}_{k+1} = \left(A_0 + \sum_{i=1}^r (A_i v_{i,k}) \right) \hat{x}_k + \left(B_0 + \sum_{i=1}^r (B_i v_{i,k}) \right) \tilde{u}_k + \sum_{i=1}^r (E_i v_{i,k}) \quad (50a)$$

$$\hat{y}_k = \left(C_0 + \sum_{i=1}^r (C_i v_{i,k}) \right) \hat{x}_k + \left(D_0 + \sum_{i=1}^r (D_i v_{i,k}) \right) \tilde{u}_k + \sum_{i=1}^r (F_i v_{i,k}) \quad (50b)$$

Notons tout de même que savoir reconstruire les espaces abstraits de l'hypomodèle (50), c'est aussi savoir reconstruire les espaces abstraits d'hypomodèles plus complexes ; la différence tenant au temps de calcul.

La suite de ce chapitre n'est consacré qu'aux variables abstraites bornées. En effet, en plus du faible intérêt pour le diagnostic (cf. chapitre 2 II.2), il est exclu de déduire les lois de distribution des variables reconstruites \hat{x}_k et \hat{y}_k dans des cas aussi complexes que (50).

L'effet d'enveloppement

Nous avons choisi de n'exploiter que des modèles en temps discret. Avant de poursuivre, nous nous proposons de justifier ce choix. Si dans le cas certain, nous avons préféré le cas discret pour des raisons pratiques, dans le cas incertain, il y a des raisons inéluctables. Considérons une équation différentielle ordinaire quelconque :

$$\frac{dx(t)}{dt} = f(x(t), u(t), \theta(t))$$

où $\theta(t)$ contient les variables abstraites bornées

Notons $\mathcal{A}(\chi(t, u(t), \theta(t))) = \{\chi(t, u(t), \theta(t)) ; \theta(t) \in \Theta\}$, le domaine abstrait de la solution $\chi(t, u(t), \theta(t))$ de cette équation différentielle. Considérons un instant infiniment proche $t+\varepsilon$, et étudions la limite de la solution de l'équation différentielle. En effectuant un développement limité à l'ordre 1 autour de l'instant t , nous obtenons :

$$\begin{aligned} \chi(t + \varepsilon, u(t + \varepsilon), \theta(t + \varepsilon)) &\xrightarrow{\varepsilon \rightarrow 0} \chi(t, u(t), \theta(t)) + \varepsilon \frac{d\chi(t, u(t), \theta(t))}{dt} \\ &= \chi(t, u(t), \theta(t)) + \varepsilon f(\chi(t, u(t), \theta(t)), u(t), \theta(t)) \end{aligned}$$

Par conséquent, l'espace abstrait obtenu à l'instant $t+\varepsilon$ est donné par :

$$\begin{aligned} \mathcal{A}(\chi(t + \varepsilon, u(t + \varepsilon), \theta(t + \varepsilon))) &\xrightarrow{\varepsilon \rightarrow 0} \mathcal{A}(\chi(t, u(t), \theta(t)) + \varepsilon f(\chi(t, u(t), \theta(t)), u(t), \theta(t))) \\ &\neq \mathcal{A}(\chi(t, u(t), \theta(t))) + \varepsilon \mathcal{A}(f(\chi(t, u(t), \theta(t)), u(t), \theta(t))) \end{aligned}$$

Or, du fait de la présence du vecteur de variables abstraites $\theta(t)$ dans les deux termes, $\mathcal{A}(\chi(t, u(t), \theta(t)))$ et $\mathcal{A}(f(\chi(t, u(t), \theta(t)), u(t), \theta(t)))$, l'espace abstrait limite de la solution à $t+\varepsilon$ ne peut être décomposé en la somme de l'espace abstrait de la solution à l'instant t avec un terme proportionnel à l'espace abstrait de la dérivée ! La décomposition reviendrait à considérer que les valeurs des variables abstraites du premier terme peuvent être différentes des valeurs prise au même instant dans le second terme. La solution ainsi obtenue serait donc une majoration de la solution réelle. Il s'ensuit que :

$$\lim_{\varepsilon \rightarrow 0} \mathcal{A}(\chi(t + \varepsilon, u(t + \varepsilon), \theta(t + \varepsilon))) \neq \mathcal{A}(\chi(t, u(t), \theta(t))) + \varepsilon \mathcal{A}\left(\frac{d\chi(t, u(t), \theta(t))}{dt}\right)$$

La seule solution pour le calcul de $\mathcal{A}(\chi(t+\varepsilon, u(t+\varepsilon), \theta(t+\varepsilon)))$ est d'évaluer directement $\chi(t+\varepsilon, u(t+\varepsilon), \theta(t+\varepsilon))$, puis, seulement, de calculer son domaine abstrait. Ceci revient à intégrer l'équation différentielle sur une période relativement faible, autrement dit, à utiliser un modèle en temps discret.

Au cours du chapitre 2, nous avons insisté sur le fait qu'il est difficile de déduire un espace abstrait, mais par contre, relativement facile de calculer un orthotope aligné circonscrit. Une méthode de recherche de l'espace abstrait de la solution de l'hypomodèle (50) pourrait donc exploiter cette particularité.

Pour déterminer l'espace abstrait des solutions, évaluons, à partir de (50a), à chaque instant le domaine $\square \mathcal{A}(x_k)$ et déduisons en utilisant l'arithmétique des intervalles (chapitre 2 de (7) à (13)), l'orthotope aligné circonscrit au domaine :

$$\Gamma_{k+1} = \left\{ \left(A_0 + \sum_{i=1}^r (A_i v_{i,k}) \right) \chi + \left(B_0 + \sum_{i=1}^r (B_i v_{i,k}) \right) \tilde{u}_k + \left(\sum_{i=1}^r (E_i v_{i,k}) \right); \chi \in \square \mathcal{A}(x_k) \right\}$$

Calculer cet orthotope est une opération relativement simple puisqu'il suffit de rechercher l'espace abstrait de chaque équation récurrente de l'hypomodèle. Le principe de la méthode revient à calculer de proche en proche les domaines $\square \Gamma_k$. Pour illustrer ce principe, revenons à l'exemple de la machine à courant continu.

exemple

Revenons à l'exemple de la page 169 et supposons que $x_k \in \square \Gamma_k$, autrement dit que l'espace abstrait de x_k soit décrit par des intervalles portant sur chaque composante $x_{i,k}$ ($i \in \{1,2\}$) de x_k : $x_{1,k} \in [\tilde{x}_{1,k}, \hat{x}_{1,k}]$ et $x_{2,k} \in [\tilde{x}_{2,k}, \hat{x}_{2,k}]$. En utilisant les valeurs numériques des matrices associées au métamodèle de la machine, nous déduisons les relations suivantes :

$$\begin{aligned} x_{1,k+1} &= (0.5404 - 0.1596v_{1,k})x_{1,k} + (-3.2817 - 0.6541v_{1,k} + 9.10^{-3}v_{2,k})x_{2,k} + \dots \\ &\quad \dots + (3.7413 - 0.4946v_{1,k} + 3.608.10^{-5}v_{2,k})u_k \\ x_{2,k+1} &= (7.5.10^{-3} - 10^{-3}v_{1,k} - 10^{-3}v_{2,k})x_{1,k} + (0.972 + 2.8.10^{-3}v_{1,k} - 4.9.10^{-3}v_{2,k})x_{2,k} + \dots \\ &\quad \dots + (2.05.10^{-2} - 1.8.10^{-3}v_{1,k} - 3.575.10^{-5}v_{2,k})u_k \end{aligned}$$

En remarquant que les dérivées de $x_{1,k+1}$ et de $x_{2,k+1}$ par rapport aux variables abstraites $x_{1,k}$ et $x_{2,k}$ ne changent pas de signe²³, la propriété (13) du chapitre 2 peut être appliquée ; puis en utilisant les autres formules de l'arithmétique des intervalles, nous obtenons sans difficulté les résultats suivants :

$$\begin{aligned} \tilde{x}_{1,k+1} &= 0.5404\tilde{x}_{1,k} - 3.2817\tilde{x}_{2,k} + 3.7413u_k - | -0.1596\tilde{x}_{1,k} - 0.6541\tilde{x}_{2,k} - 0.4946u_k | \dots \\ &\quad \dots - | 9.10^{-3}\tilde{x}_{2,k} + 3.608.10^{-5}u_k | \\ \hat{x}_{1,k+1} &= 0.5404\hat{x}_{1,k} - 3.2817\hat{x}_{2,k} + 3.7413u_k + | -0.1596\hat{x}_{1,k} - 0.6541\hat{x}_{2,k} - 0.4946u_k | \dots \\ &\quad \dots + | 9.10^{-3}\hat{x}_{2,k} + 3.608.10^{-5}u_k | \\ \tilde{x}_{2,k+1} &= 7.5.10^{-3}\tilde{x}_{1,k} + 0.972\tilde{x}_{2,k} + 2.05.10^{-2}u_k - | -10^{-3}\tilde{x}_{1,k} + 2.8.10^{-3}\tilde{x}_{2,k} - 1.8.10^{-3}u_k | \dots \\ &\quad \dots - | -10^{-3}\tilde{x}_{1,k} - 4.9.10^{-3}\tilde{x}_{2,k} - 3.575.10^{-5}u_k | \\ \hat{x}_{2,k+1} &= 7.5.10^{-3}\hat{x}_{1,k} + 0.972\hat{x}_{2,k} + 2.05.10^{-2}u_k + | -10^{-3}\hat{x}_{1,k} + 2.8.10^{-3}\hat{x}_{2,k} - 1.8.10^{-3}u_k | \dots \\ &\quad \dots + | -10^{-3}\hat{x}_{1,k} - 4.9.10^{-3}\hat{x}_{2,k} - 3.575.10^{-5}u_k | \end{aligned}$$

²³ Cette propriété est généralement satisfaite puisqu'il est rare qu'une incertitude soit telle qu'un paramètre puisse changer de signe. Si elle n'avait pas été vérifiée, seules des méthodes numériques de calcul d'extrema auraient convenu.

Le domaine $\square\Gamma_{k+1}$ est parfaitement caractérisé par ces résultats. Les résultats comparés entre les domaines abstraits réels (nous n'aborderons la méthode utilisée qu'un peu plus loin) et les domaines $\square\Gamma_{k+1}$ obtenus sont illustrés par la figure IV.

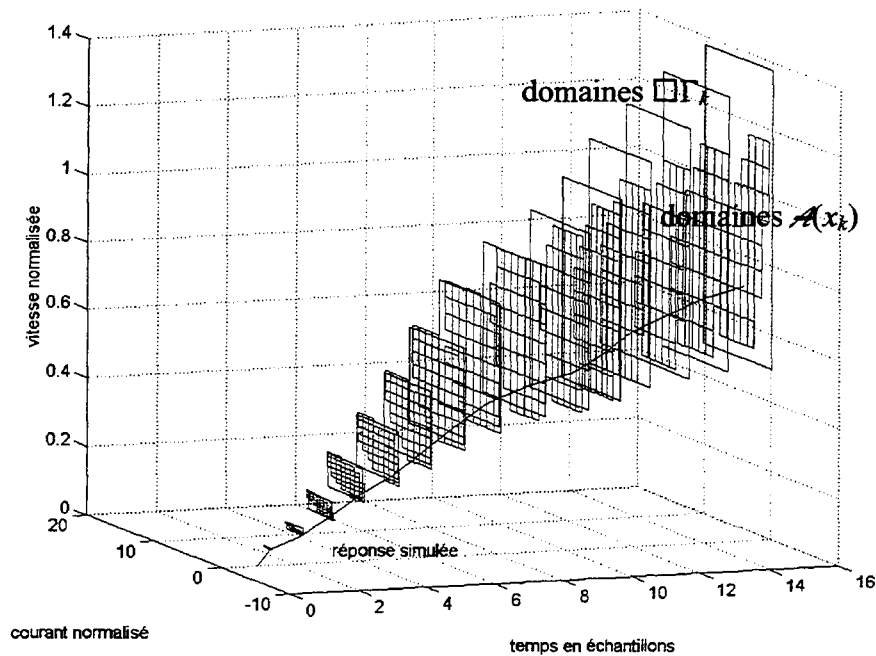


Figure IV – Effet d'enveloppement

Nous constatons que si les orthotopes alignés circonscrits s'apparentent aux domaines $\square\mathcal{A}(x_k)$ aux instants initiaux, après une dizaine d'échantillons, les orthotopes s'accroissent considérablement jusqu'à diverger : c'est l'effet d'enveloppement que Moore [Moore,79] constate dans un cas scalaire²⁴. Ceci s'explique très facilement en regardant avec attention la figure V qui extrait quatre domaines abstraits de l'enveloppe de la figure IV. En observant le domaine à l'instant $k=1$, nous remarquons que le domaine abstrait réel s'apparente à un segment de droite. Or, pour déduire le domaine à l'instant suivant, ce n'est pas ce segment qui est utilisé, mais l'orthotope qui lui est circonscrit. Ainsi, les éléments de la partie grisée n'appartenant pas à $\mathcal{A}(x_1)$, sont pris en compte lors du calcul de l'estimation $\square\Gamma_2$ de l'espace abstrait $\square\mathcal{A}(x_2)$. D'une manière générale, en substituant à un espace abstrait $\mathcal{A}(x_0)$ son orthotope circonscrit $\square\mathcal{A}(x_0)$, nous commettons une erreur car $\mathcal{A}(x_0) \subset \square\mathcal{A}(x_0)$ ²⁵. Cette erreur, si faible soit-elle, provoque, à l'instant suivant, un orthotope aligné $\square\Gamma_{k+1}$ légèrement supérieur à l'orthotope circonscrit réel $\square\mathcal{A}(x_k)$. Et ainsi, de proche en proche, à force de légers accroissements, les domaines $\square\Gamma_k$ finissent par diverger. En fait, en partant d'une valeur initiale déterminée ou définie par un orthotope aligné, le seul résultat correct est celui correspondant à l'instant 1 : $\square\Gamma_1 = \square\mathcal{A}(x_1)$. Pour revenir à la figure V, nous constatons, dès l'instant $k=7$, que l'orthotope circonscrit grisé ne coïncide pas avec l'orthotope $\square\Gamma_7$. L'augmentation de l'orthotope s'accroît à l'instant $k=10$ pour devenir très franche à l'instant $k=15$.

²⁴ Lorsqu'au lieu de l'intervalle solution, une majoration est utilisée.

²⁵ Inclusion stricte.

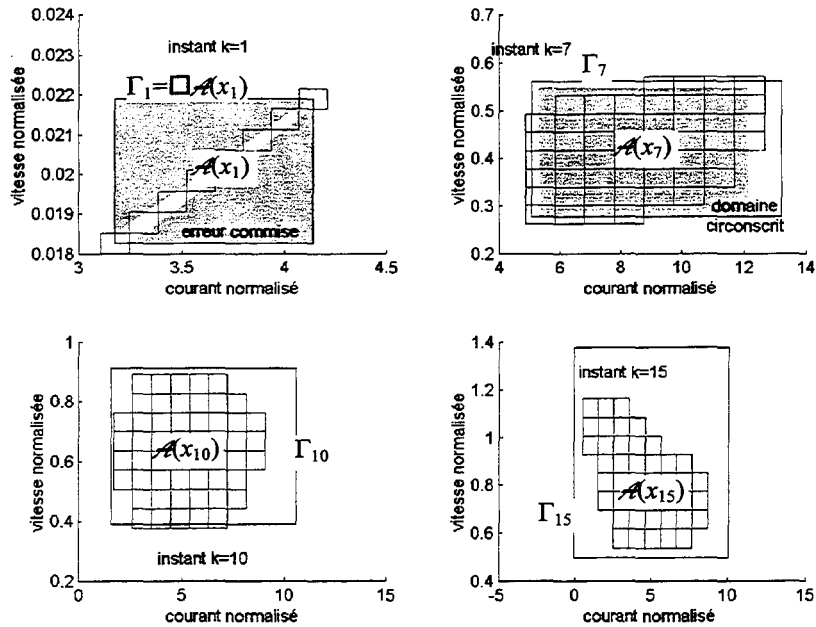


Figure V – Domaines abstraits $A(x_k)$ et domaines Γ_k calculés

L'effet d'enveloppement apparaît dès qu'au lieu des domaines abstraits réels, des majorations de ces domaines sont utilisées. Dès lors, les solutions consistant à circonscrire les domaines abstraits par des orthotopes alignés ou par des ellipses engendrent nécessairement un effet d'enveloppement et ne peuvent être appliquées. Certaines solutions ont été proposées pour compenser ce phénomène sans pour autant conduire aux domaines réels ; nous y reviendrons un peu plus loin.

Calcul d'enveloppe par l'arithmétique des intervalles

Dans le cas où les domaines abstraits $A(x_k)$ sont confondus avec leur orthotope circonscrit $\square A(x_k)$, il est évident que la technique de reconstruction des enveloppes par l'arithmétique des intervalles s'applique parfaitement et n'engendre aucun phénomène d'enveloppement. Cette indépendance ne peut évidemment être obtenue que dans le cas de métamodèles pour lesquels les équations récurrentes élémentaires constitutives de (49a) sont indépendantes les unes des autres, autrement dit, si l'équation (49a) peut être décomposée en n équations scalaires indépendantes.

exemple

Revenons à l'exemple de la machine à courant continu (pages 169 et 174). La réduction du métamodèle à la connaissance représentée par la première équation récurrente de l'hypomodèle (48) avec les matrices (49), est entièrement définie par les matrices $W=[1 \ 0]$ et $S=1$, et se déduit sans difficulté en adaptant la relation (18) : $A \rightarrow A_0 + A_1 v_{1,k} + A_2 v_{2,k}$ et $B \rightarrow B_0 + B_1 v_{1,k} + B_2 v_{2,k}$ ²⁶. Nous obtenons :

²⁶ Utiliser (18) est ici un luxe dont nous aurions pu aisément nous passer. La solution est en effet évidente !

$$x_{1,k+1} = (0.5404 - 0.1596v_{1,k})x_{1,k} + (-3.2817 - 0.6541v_{1,k} + 9.10^{-3}v_{2,k})y_{2,k} + \dots$$

$$\dots + (3.7413 - 0.4946v_{1,k} + 3.608.10^{-5}v_{2,k})u_k$$

$$y_{1,k} = x_{1,k}$$

Nous utiliserons l'hypomodèle à horizon infini le plus simple :

$$\hat{x}_{1,k+1} = (0.5404 - 0.1596v_{1,k})\hat{x}_{1,k} + (-3.2817 - 0.6541v_{1,k} + 9.10^{-3}v_{2,k})\hat{y}_{2,k} + \dots$$

$$\dots + (3.7413 - 0.4946v_{1,k} + 3.608.10^{-5}v_{2,k})\hat{u}_k$$

$$\hat{y}_{1,k} = \hat{x}_{1,k}$$

Puisque le domaine abstrait de \hat{x}_k est un intervalle (ou un orthotope aligné de dimension 1), la méthode décrite précédemment s'applique donc : $\Gamma_k = \square \mathcal{A}(x_k) = \mathcal{A}(x_k)$. L'équation conduisant aux bornes de ces domaines s'obtient en utilisant l'arithmétique des intervalles. Elle correspond à deux des équations que nous avons obtenues pour l'hypomodèle de dimension 2 :

$$\tilde{x}_{1,k+1} = 0.5404\tilde{x}_{1,k} - 3.2817\tilde{y}_{2,k} + 3.7413\tilde{u}_k - \left| -0.1596\tilde{x}_{1,k} - 0.6541\tilde{y}_{2,k} - 0.4946\tilde{u}_k \right| \dots$$

$$\dots - \left| 9.10^{-3}\tilde{y}_{2,k} + 3.608.10^{-5}\tilde{u}_k \right|$$

$$\hat{x}_{1,k+1} = 0.5404\hat{x}_{1,k} - 3.2817\hat{y}_{2,k} + 3.7413\hat{u}_k + \left| -0.1596\hat{x}_{1,k} - 0.6541\hat{y}_{2,k} - 0.4946\hat{u}_k \right| \dots$$

$$\dots + \left| 9.10^{-3}\hat{y}_{2,k} + 3.608.10^{-5}\hat{u}_k \right|$$

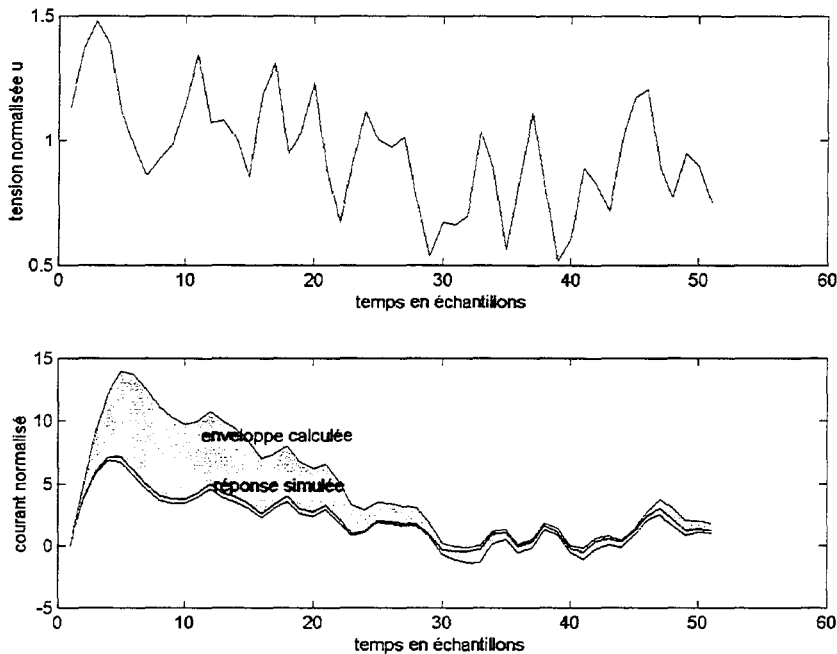


Figure VI – Calcul d'enveloppe par l'arithmétique des intervalles

La figure VI illustre l'enveloppe obtenue. Nous pouvons constater l'absence du phénomène d'enveloppement : l'enveloppe ne diverge pas. En trait continu à l'intérieur de l'enveloppe, apparaît le courant que nous avons simulé avec un modèle discret sans approximation de l'influence des incertitudes lorsque les variables abstraites se comportent comme nous l'avons décrit précédemment.

Cette méthode a l'avantage de conduire à des enveloppes en requérant peu de calcul. Elle ne s'applique cependant qu'à des hypomodèles de dimension un et, sous la condition que les paramètres de l'hypomodèle ne changent pas de signe. Ceci est généralement vérifié dans la pratique ; dans le cas contraire, il faudrait avoir recours à des méthodes numériques de calcul d'extrema, bien plus coûteuses en temps de calcul.

Calcul d'enveloppe par balayage

Pour pouvoir procéder à des opérations de diagnostic en exploitant des hypomodèles incertains à horizon infini ou simplement comprendre l'influence d'incertitudes, il est essentiel de disposer d'une méthode générale s'appliquant à des modèles multivariables. C'est l'objectif de cette partie. Nous décomposerons la résolution en deux étapes. Dans un premier temps, nous ne considérerons que l'équation récurrente (50a) de l'hypomodèle à horizon infini. L'équation de sortie ne sera prise en compte qu'a posteriori. De plus, pour simplifier les développements, nous avons considéré que la valeur de l'état à l'instant 0 était déterminée et égale à zéro. Cette valeur n'influence que les premiers instants de la reconstruction que nous considérerons comme une phase transitoire d'initialisation. Néanmoins, le lecteur comprendra par la suite qu'une incertitude portant sur la valeur initiale pourrait être prise en compte sans aucune difficulté.

Les objectifs peuvent être décrits en des termes très simples : partant des espaces abstraits connus, $\mathcal{A}(x_k)$ et du domaine Θ , caractéristique des variables abstraites θ_k , il faut être capable de déterminer l'espace abstrait $\mathcal{A}(x_{k+1})$ de l'état à l'instant suivant. Plus particulièrement, puisque les variables abstraites ont été normalisées, le domaine \mathcal{U} associé à v_k ($\forall k$) se substitue, pour la résolution de (50), au domaine Θ . En supposant que s variables abstraites apparaissent dans l'équation récurrente (50a), le domaine \mathcal{U} peut être décrit par $\mathcal{U} = \{v \in \mathbb{R}^s / \|v\|_\infty \leq 1\}$. Nous pouvons d'ores et déjà éliminer un certain nombre d'approches pour la résolution de ce problème. En effet, le phénomène d'enveloppement proscrit la plupart des méthodes se fondant sur une majoration des domaines abstraits : toute majoration, si faible soit-elle, cumulée de proche en proche conduit inexorablement à une divergence de l'enveloppe. Il ne subsiste donc que les approches par balayage que nous avons détaillées dans le cas des modèles statiques (cf. chapitre 3) ; elles doivent cependant être adaptées aux métamodèles dynamiques.

L'idée du balayage des domaines abstraits est très simple. Il s'agit de décrire le domaine \mathcal{U} par des valeurs régulièrement réparties à l'intérieur de ce domaine. Puis, la valeur initiale étant supposée nulle, et dans le cas contraire, décrite de la même façon, il suffit d'utiliser l'équation (50a) pour déterminer les valeurs correspondantes de l'état x_1 . Ainsi, si le domaine \mathcal{U} est décrit suffisamment finement, le domaine $\mathcal{A}(x_1)$ le sera aussi par les valeurs obtenues pour x_1 . Plaçons-nous à un instant k non nul quelconque. Supposons que nous disposions d'un ensemble de valeurs de x_k décrivant le domaine $\mathcal{A}(x_k)$, alors à chaque valeur de x_k peut être

associée chacune des valeurs possibles des incertitudes v_k régulièrement réparties dans \mathcal{U} et, en utilisant (50a), les valeurs correspondantes de x_{k+1} peuvent être déterminées. Le domaine $\mathcal{A}(x_{k+1})$ s'en trouve ainsi décrit. En propageant cette procédure de proche en proche en partant de l'instant 1, nous décrivons ainsi tous les domaines abstraits $\mathcal{A}(x_k)$. Cette méthode ne présente aucun risque de divergence comparable à l'effet d'enveloppement puisque tous les points décrivant les domaines $\mathcal{A}(x_k)$ sont des valeurs possibles de $x_k \dots$. Cette méthode n'est pourtant pas applicable.

Notons X_k l'ensemble des valeurs de x_k décrivant l'espace abstrait $\mathcal{A}(x_k)$ et U_k l'ensemble des valeurs décrivant les incertitudes v_k . Il va de soi que $\forall x_k \in X_k, x_k \in \mathcal{A}(x_k)$ et que $\forall v_k \in U_k, v_k \in \mathcal{U}$. Supposons que chacune des s variables abstraites $v_{i,k}$ intervenant dans (50a) soit décrite par d valeurs régulièrement distribuées entre -1 et 1 , alors le cardinal de U_k sera égal à d^s . En partant d'un état initial nul, l'ensemble X_1 contiendra alors d^s éléments. Pour obtenir l'ensemble X_2 , chacune des valeurs x_1 de X_1 doit être associée à chacune des valeurs v_1 de U_1 . L'ensemble X_2 contiendra donc d^{2s} éléments et de, proche en proche, il vient que le cardinal de l'ensemble X_k sera égale à d^{ks} ! Cette valeur croît très rapidement quand q tend vers l'infini et, pratiquement, il est déjà exceptionnel de dépasser les trois itérations.

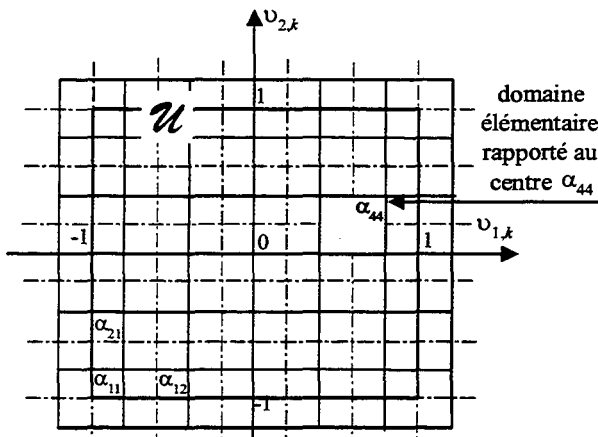


Figure VII – Décomposition d'un domaine \mathcal{U} pour $s=2$ et $d=6$

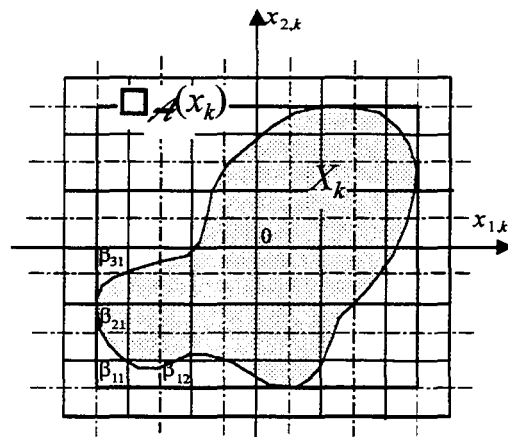


Figure VIII – Décomposition d'un domaine X_k pour $n=2$ et $d=6$

Pour remédier à cette explosion du nombre de calculs à effectuer, nous proposons d'introduire des résolutions d'espace. Le domaine \mathcal{U} sera décrit comme l'illustre la figure VII. Chaque domaine élémentaire sera assimilé à son centre. Dans le cas des domaines représentant $\mathcal{A}(x_k)$, nous calculerons comme précédemment le domaine X_k . En calant la grille de décomposition sur les extrema de X_k (cf Figure VIII), nous appliquerons le même découpage que précédemment : chaque composante de x_k sera décomposée en d intervalles et, chacun des points de X_k obtenu pourra être assimilé au centre du domaine élémentaire auquel il appartient. De la sorte, le nombre d'éléments de X_k se maintiendra à d^s !

Il est difficile de prouver la convergence de cette méthode. En observant le domaine élémentaire gris foncé de la figure VIII, nous en concluons qu'aucune valeur de X_k ne peut égaler le centre β_{31} . Or, puisque certains valeurs de X_k appartiennent au domaine élémentaire

de ce centre, la valeur β_{31} sera assimilée à une valeur de $\mathcal{A}(x_k)$. Il s'ensuit que certaines déformations peuvent se produire pouvant entraîner, en théorie²⁷, des phénomènes de divergence. Il faut cependant noter que même si les frontières peuvent être insidieusement élargies, elles peuvent aussi être réduites dans le cas où certains points dépasseraient le centre d'un domaine élémentaire. Mais, même s'il n'existe aucune garantie sur l'exactitude des espaces abstraits obtenus, plus le découpage défini par d sera important, et plus les résultats seront précis. Nous préconisons de vérifier la justesse des domaines obtenus en essayant plusieurs valeurs de découpage.

exemple

Revenons à un hypomodèle de dimension 2 de la machine à courant continu. En utilisant la méthode de balayage pour une valeur de découpage $d=8$, nous avons obtenu les résultats de la figure IX en huit minutes avec Matlab sur un ordinateur Pentium 133MHz, pour 5 secondes de fonctionnement de la machine (rapport de 100 environ). Cette méthode ne peut donc être utile qu'à des procédures de diagnostic hors ligne. L'enveloppe converge et les résultats obtenus pour des durées plus brèves avec des découpages plus importants, $d=15$ et $d=20$, sont cohérents avec le découpage $d=8$.

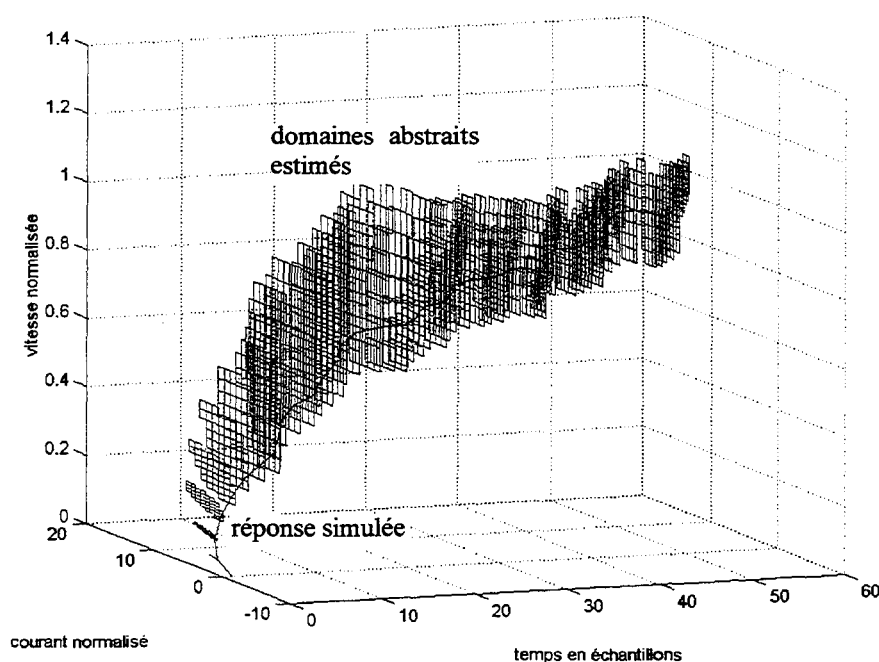


Figure IX – Enveloppe obtenue par balayage

Dans le cas particulier d'incertitudes invariantes ($\forall k, v_k=v$), il ne s'agit plus de procéder à un nouveau tirage des valeurs des variables abstraites à chaque instant. Une procédure simple consiste à balayer le domaine $\mathcal{U}=\{v \in \mathbb{R}^s / \|v\|_{\infty} \leq 1\}$ par d découpages ; ce domaine est donc balayé par d^s valeurs. Chacune de ces valeurs, notée $v^{(j)} : j \in \{1, \dots, d^s\}$, est invariante ; l'état $x_k^{(j)}$ associé à chacune d'elle se déduit de la résolution de l'équation récurrente suivante :

²⁷ Nous n'avons pu mettre en évidence de telles situations.

$$\hat{x}_{k+1}^{(j)} = \left(A_0 + \sum_{i=1}^r A_i v_i^{(j)} \right) \hat{x}_k^{(j)} + \left(B_0 + \sum_{i=1}^r B_i v_i^{(j)} \right) \tilde{u}_k^{(j)} + \sum_{i=1}^r E_i v_i^{(j)} ; j \in \{1, \dots, \mathcal{D}\}$$

Les \mathcal{D} équations récurrentes peuvent être réunies dans une équation augmentée et simulée ainsi. Enfin, pour une application au diagnostic, il est possible d'appliquer une grille de résolution pour d'éventuels tests de cohérence.

Nous avons repris l'exemple de la machine à courant continu et nous avons calculé les domaines abstraits $\mathcal{A}(x_k)$ obtenus dans les deux cas. La méthode de balayage dans le cas de variables abstraites dépendant du temps, nous a conduit pour une résolution de 10 et pour 15 échantillons, à un résultat en 343 secondes. Dans le cas invariant, en suivant la démarche que nous venons de détailler, nous avons obtenu les résultats correspondants, en conservant un découpage de 10, en 1,5 secondes. La figure X illustre les résultats obtenus à l'instant $k=15$ (en fait, un découpage de 20 a été effectué dans le cas invariant).

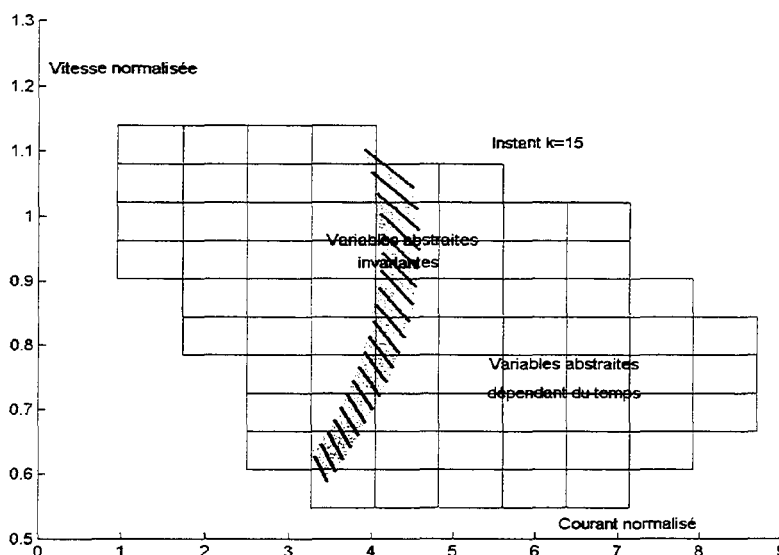


Figure X – Différence entre les domaines abstraits déduits du cas invariant et du cas dépendant du temps

Ce résultat confirme nos assertions de la partie II.3 du chapitre 3. Il existe bien une différence entre les domaines obtenus et, puisque l'invariance est un cas particulier de la dépendance temporelle, le domaine correspondant au cas de variables abstraites invariantes s'inscrit bien dans l'autre.

La méthode de balayage est, nous semble-t-il, la seule méthode qui puisse être utilisée dans un cas général. Elle réclame cependant un temps de calcul très important et ne peut donc généralement pas être utilisée en temps réel sur une installation.

Prise en compte de l'équation de sortie

Jusqu'ici, nous n'avons pas pris en compte l'équation de sortie (50b) de l'hypomodèle à horizon infini à l'exception du cas particulier où cette équation est strictement déterministe. Deux types de prise en compte peuvent être envisagés. Le plus évident est la prise en compte

directe consistant à prolonger la méthode du balayage à l'équation de sortie. Nous disposons en effet à un instant k quelconque de d' valeurs représentatives du domaine $\hat{\mathcal{A}}(x_k)$ et supposons que $s' \leq \dim(v_k)$ variables abstraites interviennent dans l'équation de sortie. Les valeurs représentant le domaine, découpé en d intervalles, caractérisant ces variables abstraites, seront au nombre de d' . Grâce à (50b), y_k peut être évalué pour chacune de ces valeurs. Nous obtiendrons $d''+s'$ qui pourront être ramenés à d'' ($p = \dim(y_k)$) valeurs grâce à une grille de résolution. L'approche directe reconstruit donc le domaine abstrait $\hat{\mathcal{A}}(y_k)$.

A l'inverse, le principe de l'approche indirecte est de tester la cohérence entre les observations de y_k et les domaines $\hat{\mathcal{A}}(x_k)$ obtenus sans reconstruire les domaines abstraits $\hat{\mathcal{A}}(y_k)$. L'équation (50b) définit un domaine \mathcal{X}_k des états x_k compatibles avec les observations \tilde{u}_k et \tilde{y}_k :

$$\mathcal{X}_k = \left\{ x / \tilde{y}_k = \left(C_0 + \sum_{i=1}^r (C_i v_{i,k}) \right) x + \left(D_0 + \sum_{i=1}^r (D_i v_{i,k}) \right) \tilde{u}_k + \sum_{i=1}^r (F_i v_{i,k}) \right\} \quad (51a)$$

Dès lors que le domaine \mathcal{X}_k aura été déterminé, il suffit, pour qu'il y ait cohérence entre le comportement représenté par l'hypomodèle (50) et celui observé, qu'il existe au moins une valeur de x_k qui soit à la fois dans $\hat{\mathcal{A}}(x_k)$ et dans \mathcal{X}_k . Nous énonçons :

$$\text{incohérence} \Leftrightarrow \hat{\mathcal{A}}(\hat{x}_k) \cap \mathcal{X}_k = \{\emptyset\} \quad (51b)$$

Il est cependant difficile d'évaluer le domaine \mathcal{X}_k dans le cas général. Nous nous restreindrons au cas où la matrice $C(v_k) = C_0 + \sum_{i=1}^p C_i v_{i,k}$ peut s'écrire $C(v_k) = P(v_k) C^*$ où $P(v_k)$ est régulière $\forall v_k \in \mathcal{U}$ et C^* est strictement déterministe de plein rang ligne. Dans cette situation, le problème (51) peut être reformulé de manière équivalente ainsi :

$$\mathcal{D}_k = \hat{\mathcal{A}} \left(P(v_k)^{-1} \left(\tilde{y}_k - \left(D_0 + \sum_{i=1}^r (D_i v_{i,k}) \right) \tilde{u}_k - \sum_{i=1}^r (F_i v_{i,k}) \right) \right) \quad (52a)$$

$$\text{incohérence} \Leftrightarrow \hat{\mathcal{A}}(C^* \hat{x}_k) \cap \mathcal{D}_k = \{\emptyset\} \quad (52b)$$

La représentation du domaine abstrait $\hat{\mathcal{A}}(C^* \hat{x}_k)$ se déduit facilement en multipliant chacune des d' valeurs représentatives de $\hat{\mathcal{A}}(\hat{x}_k)$, par C^* . Le domaine \mathcal{D}_k peut être évalué directement ou majoré, par exemple, par $\square \mathcal{D}_k$. Cette approche permet en effet d'effectuer des majorations du domaine \mathcal{D}_k , au prix certes de l'apparition possible d'incohérences non décelées. Cette dégradation peut parfois être acceptable puisque, en présence d'un défaut, même si une incohérence n'est pas détectée, il est peut probable que ceci reste vrai à tout instant. Pour tester la cohérence entre le comportement observé et un ensemble de comportements possibles, il suffit de vérifier si l'une des valeurs représentatives de $\hat{\mathcal{A}}(C^* \hat{x}_k)$ appartient au domaine \mathcal{D}_k ou à une majoration de ce domaine.

exemple

Supposons, dans l'exemple de la machine à courant continu, que les capteurs disposent d'une précision relative de 10% et que les mesures soient entachées de bruits indépendants dont les valeurs normalisées maximales sont identiques et égales à 0,2. Ceci se traduit par l'équation d'observation suivante :

$$\tilde{y}_k = \begin{bmatrix} 1 + 0,1v_{3,k} & 0 \\ 0 & 1 + 0,1v_{4,k} \end{bmatrix} x_k + \begin{bmatrix} 0,2v_{5,k} \\ 0,2v_{6,k} \end{bmatrix} \quad (53)$$

Le domaine \mathcal{X}_k est donc défini par :

$$\mathcal{X}_k = \left\{ \begin{bmatrix} \frac{\tilde{y}_{k,1} - 0,2v_{5,k}}{1 + 0,1v_{3,k}} \\ \frac{\tilde{y}_{k,2} - 0,2v_{6,k}}{1 + 0,1v_{4,k}} \end{bmatrix} ; \left\| \begin{bmatrix} v_{3,k} \\ v_{4,k} \\ v_{5,k} \\ v_{6,k} \end{bmatrix} \right\|_{\infty} \leq 1 \right\}$$

Le domaine \mathcal{X}_k est un orthotope qui se déduit facilement de l'arithmétique des intervalles. La figure XI représente les signaux que nous avons construits pour illustrer le test de cohérence indirecte. Un défaut a été simulé en substituant à la réponse²⁸ calculée pour le courant entre les instants $k=31$ et $k=45$, un signal aléatoire ajusté tel que la réponse soit vraisemblable.

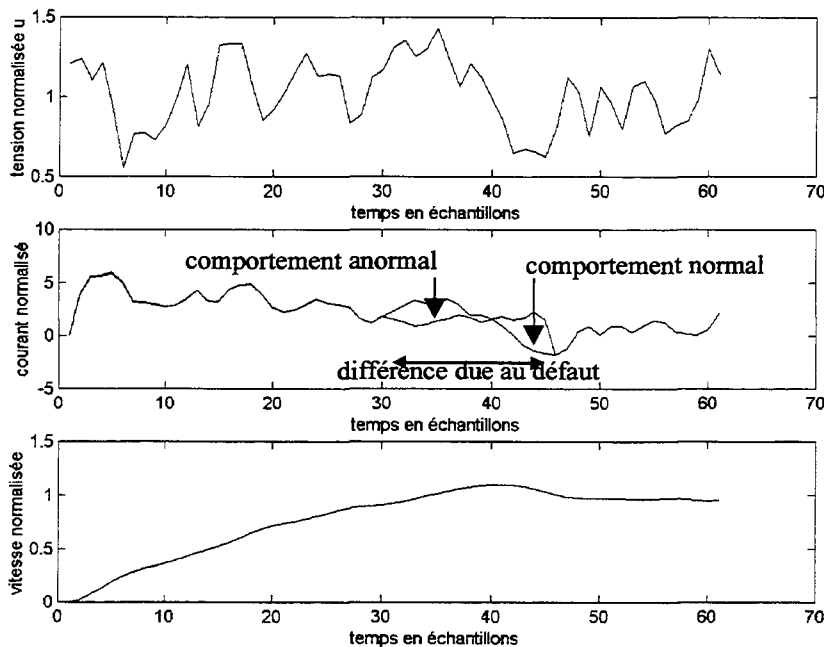


Figure XI – Signaux simulés avec défaut

Nous avons calculé les domaines abstraits $\mathcal{A}(x_k)$ par balayage puis nous avons généré les orthotopes alignés \mathcal{X}_k (cf. figure XII). Nous avons trouvé une intersection $\mathcal{X}_k \cap \mathcal{A}(x_k)$ vide aux instants 42, 43, 44 et 45. Signalons que, pour les instants 30 à 35, les intersections auraient été vides si nous n'avions considéré que les barycentres au lieu de la totalité des orthotopes élémentaires représentatifs de

²⁸ Nous avons, lors de la simulation, supposé que les capteurs étaient parfaits.

$\mathcal{A}(x_k)$. Quoiqu'il en soit, des incohérences ont été décelées de l'instant 42 à 45 : le comportement anormal est ainsi mis en évidence. Des instants 31 à 44, le comportement s'apparente à un comportement normal !

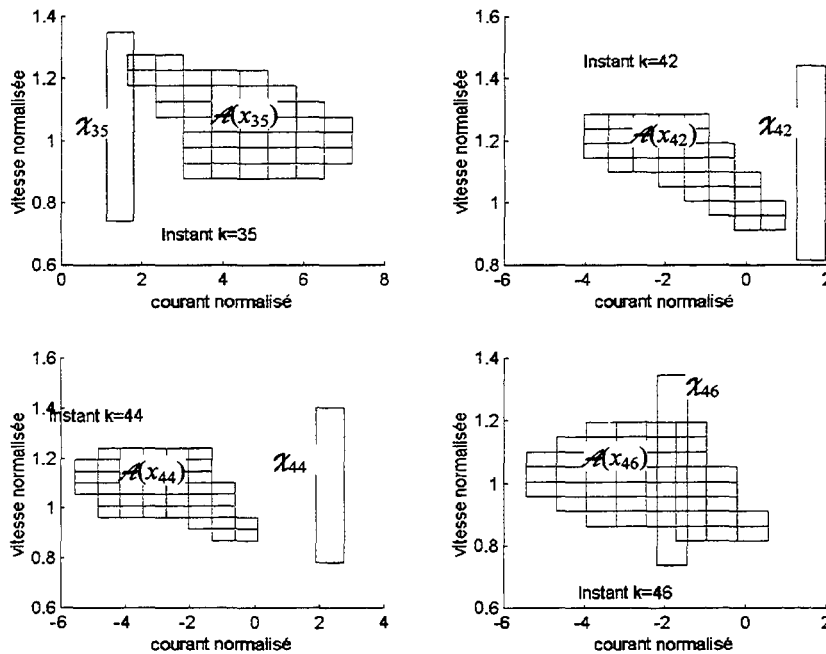


Figure XII— Quelques résultats des tests de cohérence indirects

L'approche directe s'applique systématiquement mais, il ne faut pas pour autant négliger l'approche indirecte qui conduit, sous réserve que la matrice C puisse se décomposer en $C(v_k)=P(v_k)C^*$ ²⁹, plus rapidement au résultat.

Structures purement additives

Les structures d'incertitudes purement additives sont celles qui ont été les plus étudiées jusqu'ici. La raison en est que les métamodèles sur lesquels se fondent les filtres de Kalman, reposent sur ce type de structure. Ces filtres conduisent à des descriptions probabilistes relativement simples ; les techniques de description par les bornes conduisent-elles aussi à des résultats simples ? Les approches bornantes sont-elles justifiées dans le cas de structures purement additives ? Ces situations ont déjà été abordées dans la littérature scientifique dans [Lin,70], [Hamza,71] et [Tanino et al,89]. Nous avons vu (chapitre 3) que, dans le cas de variables abstraites invariantes, les enveloppes obtenues étaient régulières. Mais qu'advient-il dans le cas général d'un hypomodèle à horizon infini, issu de (50), décrit par :

$$\hat{x}_{k+1} = A_0 \hat{x}_k + B_0 \tilde{u}_k + E v_k \quad (54)$$

avec $E = [E_1 \quad \dots \quad E_r]$ et $v_k = [v_{1,k} \quad \dots \quad v_{r,k}]^T$

Cette équation récurrente peut être développée en partant de l'état initial \hat{x}_0 :

²⁹ Ceci est vérifié dans tous les cas d'incertitudes portant sur des capteurs.

$$\hat{x}_k = A_0^k \hat{x}_0 + \sum_{i=0}^{k-1} A_0^{k-1-i} B_0 \tilde{u}_i + \sum_{i=0}^{k-1} A_0^{k-1-i} E v_i$$

Du fait de la nature transitoire des phénomènes engendrés, nous supposons que l'état initial est déterminé. Les termes soulignés correspondent alors à la valeur qu'aurait \hat{x}_k si v_k était nul quel que soit k . Nous posons $\hat{x}_k^{(c)} = A_0^k \hat{x}_0 + \sum_{i=0}^{k-1} A_0^{k-1-i} B_0 \tilde{u}_i$; cet état se déduit évidemment de l'équation récurrente suivante :

$$\hat{x}_{k+1}^{(c)} = A_0 \hat{x}_k^{(c)} + B_0 \tilde{u}_k \tag{55a}$$

Les états reconstruits \hat{x}_k vérifient désormais :

$$\hat{x}_k = \hat{x}_k^{(c)} + \sum_{i=0}^{k-1} A_0^{k-1-i} E v_i = \hat{x}_k^{(c)} + T_k v_{0 \rightarrow k+1} \tag{55b}$$

$$\text{avec } T_k = [A_0^{k-1} \dots E] \text{ et } v_{0 \rightarrow k+1} = [v_0 \dots v_{k-1}]^T$$

Nous reconnaissons dans (55b) un champ de vecteurs affine particulier (cf. partie III du chapitre 2). Nous avons montré dans le chapitre 2 que de tels champs conduisaient à des domaines abstraits ayant la forme de polytopes convexes et centrés sur $x_k^{(c)}$. En exploitant les techniques développées dans ce même chapitre, les expressions des frontières peuvent être obtenues de manière analytique ou calculées numériquement. Dans le cas fréquent où les pôles de la matrice A_0 sont contenus dans le cercle unité, nous vérifierons :

$$\lim_{k \rightarrow \infty} A_0^{k-1} E = 0 \Leftrightarrow \exists k^* / \forall k \geq k^*, A_0^{k-1} E \approx 0$$

Il en découle que $T_k \xrightarrow{k \rightarrow \infty} T_{k^*}$ et le domaine abstrait $\mathcal{A}(x_k)$ se déduit de :

$$\forall k \geq k^*, \mathcal{A}(\hat{x}_k) = \left\{ \hat{x}_k \mid v_{0 \rightarrow k+1} = 0 + T_{k^*} v_{0 \rightarrow k+1} \right\} \tag{56}$$

Nous en tirons la même conclusion que dans le cas de variables abstraites invariantes (chapitre 3), c'est-à-dire qu'indépendamment de la nature temporelle des variables abstraites, les enveloppes issues d'un modèle à structure d'incertitudes purement additives, sont régulières et centrées sur l'état reconstruit en ignorant les incertitudes (55a). La section de l'enveloppe est, pour les métamodèles du type (45), un polytope convexe (56).

Nous en concluons que, dans le cas d'une structure purement additive, les approches bornantes conduisent à des résultats au moins aussi simples que les approches probabilistes³⁰. De plus, la démarche bornée peut encore être simplifiée en n'évaluant qu'une majoration du polytope caractérisant $\mathcal{A}(x_k)$.

L'hypomodèle, défini par (54), conduit à des domaines abstraits qui n'interviennent pas dans l'équation récurrente (55a). En conséquence, les domaines peuvent être majorés sans

³⁰ Ce qui est rassurant.

engendrer de phénomènes d'enveloppement. Nous avons choisi de majorer le domaine $\mathcal{A}(x_k)$ par l'orthotope aligné circonscrit $\square\mathcal{A}(x_k)$. Bien que nous n'en cernions guère les avantages, nous aurions pu tout autant majorer le domaine $\mathcal{A}(x_k)$ par une ellipse. L'orthotope $\square\mathcal{A}(x_k)$ se calcule sans difficulté en utilisant les moyens que nous avons décrits dans le chapitre 2.

exemple

Revenons une fois de plus à l'exemple du moteur électrique (p.169). Nous avons supposé que les matrices (48) étaient certaines ($v_7(t)=0$ et $v_8(t)=0$). Nous avons alors ajouté au métamodèle en temps continu (42), une structure additive définie par :

$$E_c(v(t)) = \begin{bmatrix} 200 \\ 0 \end{bmatrix} v_7(t) + \begin{bmatrix} 0 \\ 10 \end{bmatrix} v_8(t)$$

Le métamodèle équivalent en temps discret (43) a été évalué à partir des expressions (46). Le modèle en temps discret a la structure suivante :

$$\hat{x}_{k+1} = A_0 \hat{x}_k + B_0 \tilde{u}_k + E_7 v_{7,k} + E_8 v_{8,k}$$

où A_0 et B_0 coïncident avec les résultats (49)

Puis, l'enveloppe des réponses a été calculée pour une commande particulière (figure XIV). Les domaines $\mathcal{A}(x_k)$ ont été calculés à partir de (55) en utilisant une méthode de détermination analytique des frontières. Nous avons comparé les domaines exacts avec ceux obtenus par la méthode de balayage. Il apparaît nettement que, pour un découpage $d=12$, la méthode de balayage conduit à un résultat moins précis (cf. instant $k=20$ sur la figure X) en un temps beaucoup plus long. En plus des domaines $\mathcal{A}(x_k)$ à différents instants, nous avons fait apparaître, sur la figure XIII, les orthotopes $\square\mathcal{A}(x_k)$ qui représentent relativement bien les domaines $\mathcal{A}(x_k)$. Ces domaines ont l'avantage de se calculer très rapidement.

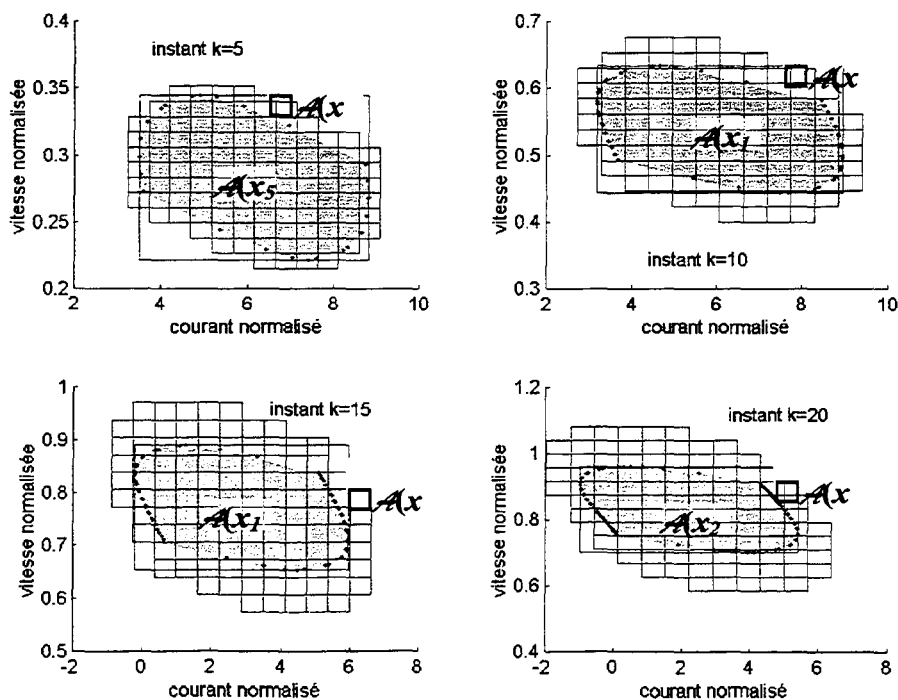


Figure XIII – Domaines $\mathcal{A}(x_k)$ obtenus par balayage ($d=12$) puis en exploitant la structure additive

En plus de s'évaluer facilement, l'orthotope $\square \mathcal{A}(x_k)$ présente l'avantage de permettre de découpler le test de cohérence. Au lieu de tester si les valeurs de x_k appartiennent à une enveloppe de section polytopique, il suffit de vérifier si chacune des observations de x_k appartient à des intervalles associés à chaque orthotope aligné circonscrit. Pour déceler une incohérence à un instant donné, il suffit que l'une des deux observations n'appartienne pas à son intervalle. La figure XIV, illustre les enveloppes déduites, lorsque la tension est donnée, pour le courant et la vitesse normalisés. Nous avons aussi fait apparaître sur cette figure, les enveloppes obtenues en utilisant directement (56) ; c'est-à-dire en négligeant le régime transitoire initial (nous avons choisi $k^*=200$). Nous constatons que l'enveloppe tenant compte du transitoire tend rapidement vers celle qui l'ignore.

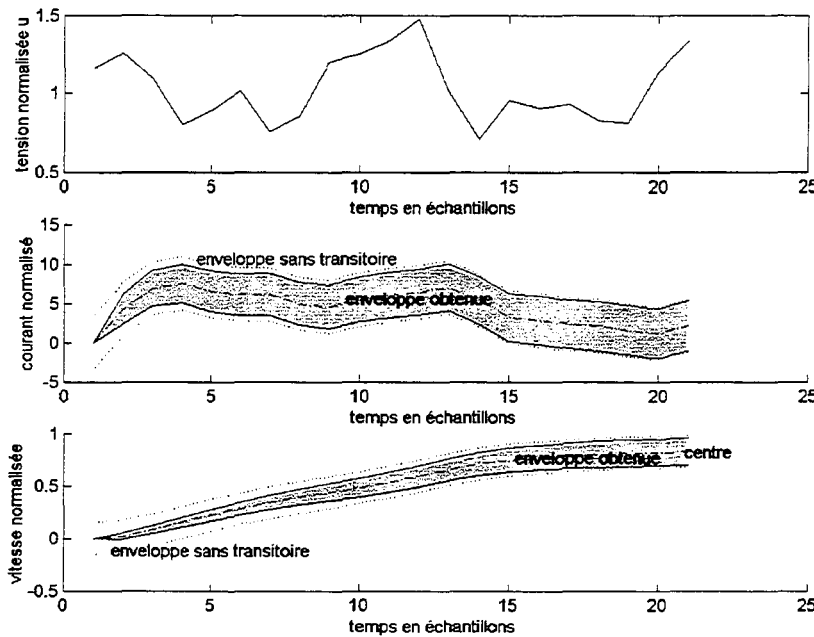


Figure XIV – Enveloppes obtenues en majorant $\mathcal{A}(x_k)$ par $\square \mathcal{A}(x_k)$

Les métamodèles pourvus d'une structure d'incertitudes purement additives sont donc très faciles à exploiter ; seulement, ils ne peuvent engendrer que des enveloppes régulières (en négligeant le régime transitoire à l'origine) et ne présentent, par conséquent, qu'un intérêt mineur (en tout cas comparable aux résultats d'un filtre de Kalman).

Approches elliptiques

De nombreux auteurs ont étudié l'estimation d'état en présence d'incertitudes représentées par des domaines délimités par des ellipses. Schweppe [Schweppe,68] a initié ces travaux. Dans tous les cas, les structures étudiées sont purement additives :

$$\begin{aligned} \frac{d}{dt} x(t) &= Ax(t) + E v(t) \\ y(t) &= Cx(t) + v(t) \end{aligned}$$

Le vecteur de variables abstraites $v(t)$ est caractérisé par l'un des deux domaines suivants :

$$\mathcal{A}(v(t)) = \{v(t) : v(t)^T Q^{-1} v(t)\} \text{ ou } v(t) \in \left\{ v(t) : \int_{t_k}^{t_{k+1}} v(t)^T Q^{-1} v(t) dt \leq t_{k+1} - t_k \right\}$$

où Q est une matrice symétrique strictement définie positive

Le premier domaine correspond à un espace abstrait tandis que le second définit un domaine que nous n'avons pas envisagé dans ce travail : une majoration portant, non plus sur les valeurs, mais, pour tenter de lui donner une signification physique, sur l'énergie possible des variables non déterministes.

Le vecteur $v(t)$ n'est caractérisé que d'une seule façon :

$$\mathcal{A}(v(t)) = \{v(t) : v(t)^T R^{-1} v(t) \leq 1\}$$

Schweppe montre que, lorsque c'est l'énergie de la variable abstraite $v(t)$ qui est majorée, le domaine abstrait $\mathcal{A}(x(t))$ est ellipsoïdal, puis Bertsekas et Rhodes [Bertsekas,71] suivi de Savkin et Petersen [Savkin et al,95] montrèrent que ce domaine pouvait être déduit d'une équation de Riccati semblable à celle du filtre de Kalman. Ces résultats conduisent à une nouvelle interprétation du filtrage de Kalman, mais ne présentent qu'un faible apport pratique.

Dans le cas où la variable $v(t)$ est caractérisée par le domaine abstrait $\mathcal{A}(v(t))$, le domaine $\mathcal{A}(x(t))$ n'est a priori plus ellipsoïdal. La solution est alors recherchée sous la forme d'une majoration par une ellipse du domaine $\mathcal{A}(x(t))$.

Ces majorations se fondent sur des propriétés bien connues dans le domaine des approches ellipsoïdales [Filippova et al,96]. Nous noterons $\mathcal{E}(c, Q)$ l'ellipse définie par : $\mathcal{E}(c, Q) = \{x : (x - c)^T Q^{-1} (x - c) \leq 1\}$. Les trois propriétés suivantes sont les plus importantes.

Soient deux ellipses $\mathcal{E}(a_1, Q_1)$ et $\mathcal{E}(a_2, Q_2)$. La somme de leur éléments vérifie :

$$\forall \lambda \in \mathbb{R}^{+*}, \mathcal{E}(a_1, Q_1) + \mathcal{E}(a_2, Q_2) \subset \mathcal{E}(a_1 + a_2, Q(\lambda))$$

et

$$Q(\lambda) = (1 + \lambda^{-1})Q_1 + (1 + \lambda)Q_2$$

L'ellipse circonscrite à la somme des deux ellipses est donnée par : $\bigcap_{\lambda > 0} \mathcal{E}(a_1 + a_2, Q(\lambda))$.

L'intersection de deux ellipses vérifie :

$$\forall B_1, B_2 / B_1 + B_2 = I_{\dim(x)}, \mathcal{E}(a_1, Q_1) \cap \mathcal{E}(a_2, Q_2) \subset \mathcal{E}(B_1 a_1, B_1 Q B_1^T) + \mathcal{E}(B_2 a_2, B_2 Q B_2^T)$$

L'ellipse circonscrite à l'intersection des deux ellipses se déduit de :

$$\bigcap_{B_1, B_2 / B_1 + B_2 = \dim(x)} \{\mathcal{E}(B_1 a_1, B_1 Q B_1^T) + \mathcal{E}(B_2 a_2, B_2 Q B_2^T)\}$$

Enfin, le produit cartésien des deux ellipses vérifie :

$$\forall \lambda \in \mathbb{R}^{+*}, \mathcal{E}(a_1, Q_1) \times \mathcal{E}(a_2, Q_2) \subset \mathcal{E}(z, Z(\lambda))$$

et
$$z = [a_1^\top \quad a_2^\top]^\top \text{ et } Z(\lambda) = \begin{bmatrix} (1+\lambda^{-1})Q_1 & 0 \\ 0 & (1+\lambda)Q_2 \end{bmatrix}$$

La solution exacte s'obtient de : $\mathcal{E}(a_1, Q_1) \times \mathcal{E}(a_2, Q_2) = \bigcap_{\lambda > 0} \mathcal{E}(z, Z(\lambda))$.

Nous ne détaillerons pas les algorithmes conduisant à l'évaluation des ellipses contenant le domaine abstrait $\mathcal{A}(x(t))$; nous n'en décrivons que le principe. En approchant la dérivée par l'approximation de Newton et en utilisant les relations précédentes, il est possible d'obtenir une ellipse majorant le domaine abstrait $\mathcal{A}(x(t))$. Néanmoins, du fait des majorations, un phénomène d'enveloppement apparaîtrait si l'équation de mesure du modèle n'était utilisée de manière indirecte pour réajuster, à chaque instant, la majoration du domaine. L'évaluation du domaine $\mathcal{A}(x(t))$ sera réajustée en effectuant l'intersection entre la majoration du domaine déduit des instants précédents et de l'ensemble des états possibles eu égard à la nouvelle valeur de $y(t)$. Une incohérence apparaîtra lorsque l'intersection entre ces deux domaines sera vide.

Si cette approche s'applique à la reconstruction d'état, elle nous semble, par de nombreux aspects, d'un faible intérêt pour le diagnostic. Supposons, tout d'abord, qu'il y ait une incohérence durant une durée importante. Le réajustement de l'estimation ne pourra être effectué et, durant cette période, l'estimation divergera ! Par ailleurs, ces approches se fondent sur de nombreuses majorations. Or, elles ne concernent que des structures additives qui peuvent se traiter simplement et de manière exacte (cf. structures purement additives) par des approches polytopiques. Nous ne voyons rien qui puisse justifier de telles approches. Pour finir, la représentation des incertitudes par des ellipses nous semble bien loin de la réalité physique, même si l'on trouve, dans [Pronzato et al,96], une méthode permettant de déduire des ellipses circonscrites. Pour ces différentes raisons, nous avons exclu ces approches de notre étude dont le dessein est la synthèse de procédures de diagnostic applicables.

III.3. Les hypomodèles incertains à horizon fini

Les hypomodèles incertains à horizon infini ne peuvent être exploités vraiment que dans quelques cas particuliers : les cas où les variables reconstruites sont scalaires mais aussi où la structure des incertitudes est purement additive. Bien que les autres situations puissent être résolues en théorie, au moins par la méthode de balayage, des raisons techniques tenant aux temps de calcul considérables, nous conduisent à renoncer, pour une opération de diagnostic, à de tels hypomodèles. Il est dès lors naturel de se pencher sur les hypomodèles incertains à horizon fini afin de déterminer s'ils peuvent conduire à des procédures de diagnostic applicables là où les hypomodèles à horizon infini n'apportent pas de solutions.

Formulation générique

La formulation générique peut se calquer sur celle des hypomodèles certains à horizon fini du paragraphe II.1 de ce chapitre. Comme précédemment, nous considérerons que la totalité du métamodèle (50) est exploité ; dans le cas contraire, le métamodèle devra être réduit de façon à ne faire apparaître que la connaissance à exploiter.

D'une manière générale, un hypomodèle incertain à variables abstraites bornées et à horizon fini d'horizon h peut s'écrire :

$$\forall k \in \mathbf{N}, \exists v / f(\tilde{u}_k, \dots, \tilde{u}_{k+h-1}, \tilde{y}_k, \dots, \tilde{y}_{k+h-1}, v) = 0 ; \|v\|_{\infty} \leq 1 \quad (57)$$

avec $\dim(v) = r$, et $f(\cdot)$ un champ de vecteurs quelconques

Tester la cohérence avec cet hypomodèle n'est pas aussi trivial que dans le cas certain. Nous pouvons distinguer deux approches.

L'approche directe procède en quelque sorte de l'identification. Elle vise à rechercher s'il existe au moins une valeur du vecteur v , dont la norme infinie est inférieure ou égale à 1, qui annule le champ $f(\cdot)$ de (57). Trois situations peuvent être distinguées.

- 1) Il n'existe aucune valeur du vecteur v annulant l'expression (57). L'incohérence est manifeste.
- 2) Il existe une seule valeur du vecteur v annulant l'expression (57). Si la norme infinie de cette valeur est inférieure ou égale à 1, alors la cohérence est constatée, sinon, il faut conclure à une incohérence.
- 3) Il existe plusieurs valeurs du vecteur v annulant l'expression (57). Il faut alors rechercher l'ensemble des valeurs de v possibles puis étudier si, parmi elles, il en existe une dont la norme infinie est inférieure ou égale à 1 ; cela révélerait une cohérence et, dans le cas contraire, une incohérence.

Il n'est guère possible de prolonger la description de l'approche directe dans le cas général d'un hypomodèle du type (57). Nous ne détaillerons cette approche que dans le cas particulier des hypomodèles à horizon finis affines en les variables abstraites.

Une approche indirecte se reconnaît au fait qu'elle procède par reconstruction d'un ou de plusieurs domaines abstraits. Par exemple, en posant, $r_k = f(\tilde{u}_k, \dots, \tilde{u}_{k+h-1}, \tilde{y}_k, \dots, \tilde{y}_{k+h-1}, v)$, et en calculant l'espace abstrait $\mathcal{A}(r_k)$, il est souvent facile de vérifier si $\{0\} \in \mathcal{A}(r_k)$. D'autres manières de tester indirectement la cohérence peuvent être envisagées. Supposons, par exemple, que l'hypomodèle (57) s'écrive $g(\tilde{u}_k, \dots, \tilde{u}_{k+h-1}, \tilde{y}_k, \dots, \tilde{y}_{k+h-1}) + h(v) = 0$, alors il suffit d'évaluer le domaine $\mathcal{A}(-h(v))$ pour vérifier si la valeur déterministe $g(\tilde{u}_k, \dots, \tilde{u}_{k+h-1}, \tilde{y}_k, \dots, \tilde{y}_{k+h-1})$ appartient à ce domaine abstrait. Si l'hypomodèle s'écrit $g(\tilde{u}_k, \dots, \tilde{u}_{k+h-1}, \tilde{y}_k, \dots, \tilde{y}_{k+h-1}, v_1) + h(v_2) = 0$ et si v_1 et v_2 sont des vecteurs de variables abstraites indépendants, il est possible de tester la cohérence en évaluant

³¹ Les vecteurs \tilde{u}_k et \tilde{y}_k contiennent tout deux des observations ; leur différentiation n'a de sens que dans la mesure où l'hypomodèle est associé à un métamodèle de type (56).

$\mathcal{A}(g(\tilde{u}_k, \dots, \tilde{u}_{k+h-1}, \tilde{y}_k, \dots, \tilde{y}_{k+h-1}, v_1))$ et $\mathcal{A}(h(v_2))$; une incohérence sera révélée si $\mathcal{A}(g(\tilde{u}_k, \dots, \tilde{u}_{k+h-1}, \tilde{y}_k, \dots, \tilde{y}_{k+h-1}, v_1)) \cap \mathcal{A}(h(v_2)) = \{\emptyset\}$.

Toutes les réalisations d'approches indirectes ne peuvent pas, raisonnablement, être énumérées car il en existe autant que de cas particuliers, c'est-à-dire une infinité. Il nous semble judicieux de choisir de tester si $\{0\} \in \mathcal{A}(r_k)$ parce que non seulement cette approche a le mérite de toujours s'appliquer mais aussi parce qu'elle élude le calcul parfois compliqué d'intersections entre domaines abstraits. Il est évident qu'en l'absence d'approximation, le résultat d'un test de cohérence indirecte ne dépend pas de sa formulation ; comme dans le cas certain, la nature causale ou acausale d'un hypomodèle à horizon fini n'affecte pas le résultat d'un test. Les moyens permettant de calculer ces espaces ont été longuement développés dans le chapitre 2 ; nous n'y reviendrons donc pas dans ce chapitre. Remarquons toutefois que ces approches peuvent devenir très simples lorsque, au lieu de déterminer $\mathcal{A}(r_k)$, on se contente de la majoration $\square \mathcal{A}(r_k)$. Nous développerons l'approche indirecte dans le cas particulier suivant.

L'hypomodèle (57) demeurant très général, les tests de cohérence ne peuvent être décrits précisément. Nous nous restreindrons maintenant aux hypomodèles à horizon fini affines en les variables abstraites dont les propriétés sont intéressantes. Ils peuvent s'écrire :

$$\forall k \in \mathbf{N}, \exists v / \sum_{i=0}^{h-1} \left(\left(U_0 + \sum_{j=1}^{r_v} U_{i,j} v_j \right) \tilde{u}_{k+i} + \left(Y_0 + \sum_{j=1}^{r_v} Y_{i,j} v_j \right) \tilde{y}_{k+i} + \sum_{j=1}^{r_v} E_j v_j \right) = 0 ; \|v\|_{\infty} \leq 1 \quad (58a)$$

avec $v = [v_1 \dots v_{r_v}]^T$ et $\dim(U_0) = p$

Pour mieux faire apparaître les particularités de cet hypomodèle, nous le réécrivons de la manière suivante :

$$\forall k \in \mathbf{N}, \exists v^{(r_v)} / V_{0,k} + \sum_{j=1}^{r_v} V_{j,k} v_j = V_{0,k} + V_k^{(r_v)} v^{(r_v)} = 0 ; \|v^{(r_v)}\|_{\infty} \leq 1 \quad (58b)$$

avec : $V_{0,k} = \sum_{i=0}^{h-1} (U_0 \tilde{u}_{k+i} + Y_0 \tilde{y}_{k+i})$ et $V_{j,k} = \left(\sum_{i=0}^{h-1} U_{i,j} \tilde{u}_{k+i} + Y_{i,j} \tilde{y}_{k+i} + E_j \right)$

$V_k^{(r_v)} = [V_{1,k} \dots V_{r_v,k}]$ et $v^{(r_v)} = [v_1 \dots v_{r_v}]^T$

Cette relation peut aussi s'écrire :

$$\forall k \in \mathbf{N}, \exists v^{(r_v)} / V_k^{(r_v)} v^{(r_v)} = -V_{0,k} ; \|v^{(r_v)}\|_{\infty} \leq 1 \quad (59)$$

Nous nous proposons maintenant de décrire successivement ces deux approches qui ne constituent rien d'autres que des tests de cohérence. Nous nous restreindrons évidemment aux modèles de type (58).

Approche directe

L'approche directe peut être détaillée dans le cas d'hypomodèle décrit par (58a). La première étape d'une approche directe est de rechercher s'il existe au moins une solution $v^{(r_v)}$ à cette

relation. Or, la relation (59) étant affine en les variables abstraites, l'algèbre linéaire nous permet d'avancer qu'il existe au moins une solution si la condition suivante est satisfaite :

$$\text{rang}\left(\begin{bmatrix} V_{0,k} & V_k^{(r_v)} \end{bmatrix}\right) = \text{rang}\left(V_k^{(r_v)}\right). \quad (60)$$

Si $\text{rang}\left(\begin{bmatrix} V_{0,k} & V_k^{(r_v)} \end{bmatrix}\right) \neq \text{rang}\left(V_k^{(r_v)}\right)$, alors il faut conclure à l'incohérence entre l'hypomodèle et les observations, et la cohérence dans le cas contraire (première situation).

Si cette condition n'est pas satisfaite, alors l'incohérence est manifeste ; sinon, il faut poursuivre la procédure. L'ensemble des solutions $v^{(r_v)}$ possibles, lorsque (60) est vérifié, est donné par :

$$\left\{ v^{(r_v)} / v^{(r_v)} = -V_h^{(r_v)+} V_{0,k} + \left(I_{r_v} - V_h^{(r_v)+} V_h^{(r_v)} \right) \chi : \chi \in \mathbb{R}^{r_v} \right. \quad (61)$$

où $V_k^{(r_v)+}$ est l'inverse généralisé de $V_k^{(r_v)}$

Posons $q = \text{rang}(V_k^{(r_v)})$. Alors, lorsque $q = p$, c'est-à-dire quand $V_k^{(r_v)}$ est une matrice de plein rang ligne, il existe une unique solution, correspondant à la seconde situation, donnée par :

$$v^{(r_v)} = -V_h^{(r_v)+} V_{0,k} \quad (62)$$

Par conséquent, si $\text{rang}(V_k^{(r_v)}) = p$, la cohérence sera constatée si $\|V_h^{(r_v)+} V_{0,k}\|_{\infty} \leq 1$, sinon lorsque $\|V_h^{(r_v)+} V_{0,k}\|_{\infty} > 1$, il s'agit d'une incohérence (seconde situation).

Le plus fréquemment, $\text{rang}(V_k^{(r_v)}) \neq p$. Nous nous retrouvons dans la troisième situation. L'expression de l'ensemble des $v^{(r_v)}$ possibles (61) n'est pas commode car elle ne fait pas apparaître la partie utile de $\left(I_{r_v} - V_h^{(r_v)+} V_h^{(r_v)} \right)$ puisque cette matrice n'est que de rang $r_v - q$. Pour reformuler (61), nous décomposerons $V_k^{(r_v)}$ en valeurs singulières :

$$V_k^{(r_v)} = \Phi_k \begin{bmatrix} \Sigma_k & 0 \\ 0 & 0 \end{bmatrix} \Psi_k^T \Rightarrow V_k^{(r_v)+} = \Psi_k \begin{bmatrix} \Sigma_k^{-1} & 0 \\ 0 & 0 \end{bmatrix} \Phi_k^T$$

où Φ_k et Ψ_k sont des matrices orthogonales et Σ_k est une matrice diagonale régulière avec $\dim(\Sigma_k) = (q \times q)$

En utilisant cette décomposition, la matrice $\left(I_{r_v} - V_h^{(r_v)+} V_h^{(r_v)} \right)$ peut être développée ; le domaine (61) peut alors s'écrire :

$$\left\{ v^{(r_v)} / v^{(r_v)} = -V_h^{(r_v)+} V_{0,k} + \Psi_k \begin{bmatrix} 0 & 0 \\ 0 & I_{r_v - q} \end{bmatrix} \Psi_k^T \chi' : \chi' \in \mathbb{R}^{r_v} \right\}$$

ou encore

$$\left\{ v^{(r_v)} / v^{(r_v)} = -V_h^{(r_v)+} V_{0,k} + \Psi_k \begin{bmatrix} 0 \\ I_{r_v - q} \end{bmatrix} \chi : \chi \in \mathbb{R}^{r_v - q} \right\}$$

Pour rechercher s'il existe au moins une solution $v^{(r_v)}$ telle que $\|v^{(r_v)}\|_{\infty} \leq 1$, nous proposons de résoudre :

$$\left(\min \|v^{(rv)}\|_{\infty} \right) \leq 1 \Leftrightarrow \left(\min_{\chi \in \mathbb{R}^{r_v - q}} \left\| -V_k^{(r_v)h} V_{0,k} + \Psi_k \begin{bmatrix} 0 \\ I_{r_v - q} \end{bmatrix} \chi \right\|_{\infty} \right) \leq 1 \quad (63)$$

Or, puisque la norme infinie n'est pas dérivable en zéro, il n'existe pas de solution analytique à (63). Pour contourner cette difficulté, nous proposons de substituer une norme 2 à la norme infinie :

$$\|\chi^*\|_{\infty} \leq 1 : \chi^* = \arg \left(\min_{\chi \in \mathbb{R}^{r_v - q}} \left\| -V_k^{(r_v)h} V_{0,k} + \Psi_k \begin{bmatrix} 0 \\ I_{r_v - q} \end{bmatrix} \chi \right\|_2 \right) \quad (64)$$

Si, lorsque ce critère est satisfait, la cohérence est indéniable, il existe des cas où la valeur χ^* minimise le critère exprimé avec la norme 2 sans pour autant qu'elle minimise le critère (63). Lorsque le critère (64) n'est pas satisfait, il est préférable de parler d'incohérence probable.

En utilisant les résultats des moindres carrés simples et le développement en valeurs singulière de $V_k^{(r_v)}$, nous déduisons que le minimum est $\chi^* = 0$. La plus petite valeur de $v^{(rv)}$, au sens de la norme 2, est donc donnée par (62). Ainsi, même si $\text{rang}(V_k^{(r_v)}) \neq p$, si $\|V_h^{(r_v)h} V_{0,k}\|_{\infty} \leq 1$, il faut encore conclure à la cohérence. Cependant, si $\|V_h^{(r_v)h} V_{0,k}\|_{\infty} > 1$, nous

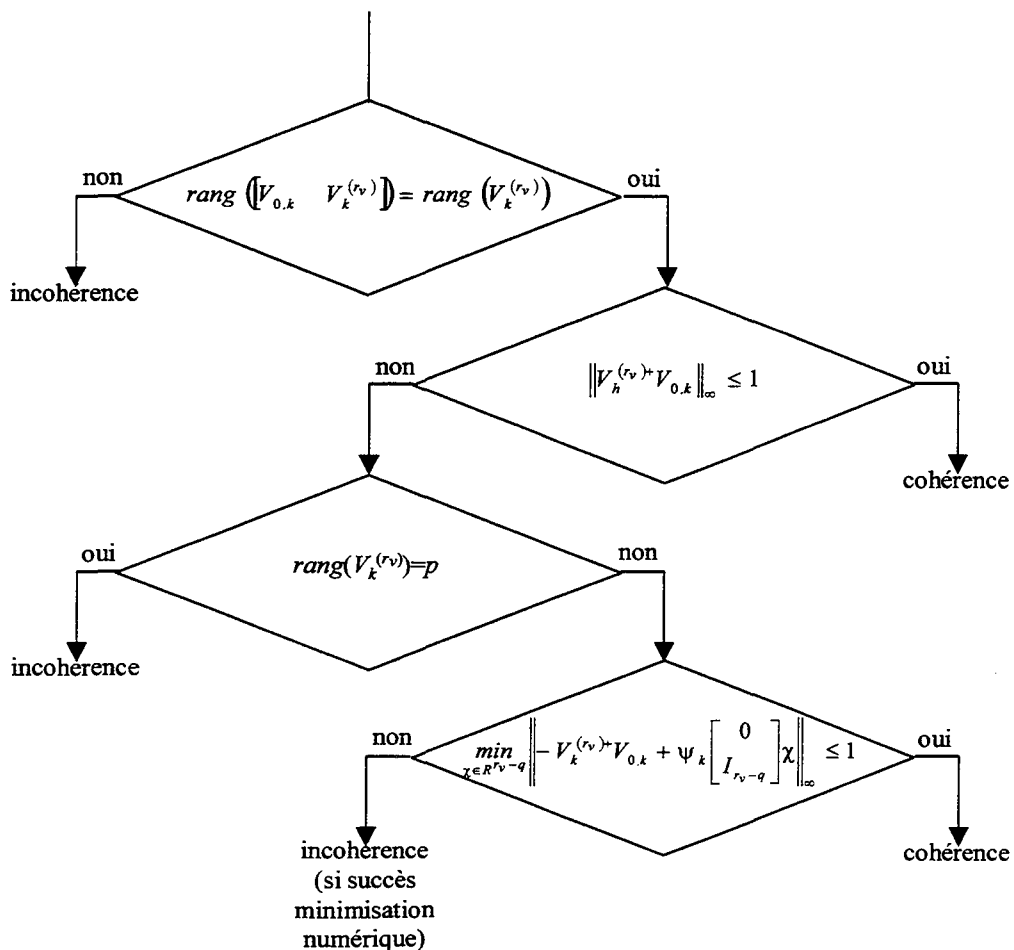


Figure XV- Les différentes étapes de l'approche directe

ne pouvons conclure qu'à une incohérence probable. Alors, pour améliorer le diagnostic, une méthode numérique résolvant (63), n'exploitant évidemment pas d'expression analytique de la dérivée du critère (puisqu'elle n'existe pas), de type simplex, doit être utilisée.

Ces différents résultats peuvent être résumés sur l'organigramme de la figure XV.

Approche indirecte

L'approche indirecte repose sur la détermination de l'espace abstrait suivant ((58b) et (59)) :

$$\mathcal{A}(r_k) = \left\{ r_k / r_k = V_{0,k} + V_k^{(rv)} v^{(rv)} ; \|v^{(rv)}\| \leq 1 \right\} \quad (65)$$

Les résultats du chapitre 2 partie III.2, nous permettent de déduire que le domaine $\mathcal{A}(r_k)$ est un paralléloéde (a fortiori convexe). Le champ vectoriel apparaissant dans (65) est dissociable ; il suffit par conséquent d'utiliser une procédure extrayant l'ensemble des sous-ensembles de fonctions linéaires résolubles, puis, comme le champ est affine en les variables abstraites, nous obtenons, par élimination des variables communes, des équations linéaires de contrainte. Conformément à (28) du chapitre 2, le domaine abstrait est caractérisé par l'orthotope $\square \mathcal{A}(r_k)$ et d'un ensemble d'inéquations linéaires que se déduisent analytiquement. Nous ne reviendrons sur la procédure qui permet d'y parvenir que dans le chapitre suivant (paragraphe II.3). Nous nous contenterons ici de rapeler que les solutions sont du type :

$$\sum_{i=1}^p M_{i,k} r_{i,k} \leq M_{0,k} \quad \text{où } \forall i, M_{i,k} \text{ est un vecteur et } r_k = [r_{1,k} \dots r_{p,k}]^T$$

Pour tester la cohérence entre l'hypomodèle et les observations, il suffit de vérifier si $\{0\} \in \square \mathcal{A}(r_k)$ ($\square \mathcal{A}(r_k)$ se calcule sans difficulté) et $r_k=0$ est solution de l'inégalité précédente ; c'est-à-dire si tous les éléments de $M_{0,k}$ sont inférieurs ou égaux à zéro : $M_{0,k} \geq 0$ ($M_{i,k}$ n'intervient pas). Si ces deux conditions sont vérifiées, la cohérence est constatée, sinon il faudra conclure à une incohérence. Cette méthode conduit à des diagnostics exacts, plus encore que l'approche directe que nous venons de voir car elle évite, comme nous l'avons fait apparaître dans l'organigramme, l'incertitude sur le succès de la minimisation numérique.

Il est aussi possible d'utiliser des approches numériques (chapitre 2 paragraphe IV). Dans un cas général, l'approche par balayage est souvent la seule qui soit applicable mais, ici, l'arithmétique des intervalles (chapitre 2 paragraphe III.1) peut être utilisée pour appliquer la méthode de pavage qui est, nous l'avons vu, plus performante. Pour comparer les différentes approches, nous sommes partis d'un exemple simple³².

exemple

Supposons que l'hypomodèle (58a) conduise, pour trois instants consécutifs, à trois relations (58b) ou (59), s'écrivant :

³² Nous nous en tiendrons ici à des cas triviaux ; un exemple complet sera traité dans le chapitre 5

$$V_0^{(r_v)} = V_1^{(r_v)} = V_2^{(r_v)} = \begin{bmatrix} 1 & 0 & 0 & 1 & 5 \\ 0 & 1 & 0 & 1 & -2 \\ 0 & 0 & 0 & 1 & 3 \end{bmatrix}$$

$$V_{0,0} = \begin{bmatrix} 1 \\ 1 \\ -1 \end{bmatrix}, V_{0,1} = \begin{bmatrix} 2 \\ 2 \\ -1 \end{bmatrix} \text{ et } V_{0,2} = \begin{bmatrix} 3 \\ 2 \\ -1 \end{bmatrix}$$

Les résultats que nous avons obtenus ont été récapitulés dans le tableau suivant :

	instant 0		instant 1		instant 2	
	résultat	durée	résultat	durée	résultat	durée
approche directe						
$\ V_h^{(r_v)} + V_{0,k}\ _\infty \leq 1$ ³³	oui	n.m. ³⁴	non	n.m.	non	n.m.
$\min_{\chi \in R^{r_v-q}} \left\ -V_k^{(r_v)} + V_{0,k} + \Psi_k \begin{bmatrix} 0 \\ I_{r_v-q} \end{bmatrix} \chi \right\ _\infty \leq 1$	-	1,27s.	oui	2,75s.	oui	2,47s. ³⁵ / 2,84s. ³⁶
approches indirectes						
$\{0\} \in \square \mathcal{A}(r_k)$	oui	n.m.	oui	n.m.	oui	n.m.
approche indirecte analytique						
$M_{0,k} \geq 0$	oui	0,6s.	oui	0,6s.	non	0,6s.
approche indirecte numérique						
pavage	oui ³⁵	0,71s.	oui ³⁵	0,71s.	oui ³⁵ /non ³⁶	1,65s./ 143,68s. ³⁶

Il apparaît que le test $\{0\} \in \square \mathcal{A}(r_k)$ est intéressant de par la durée très faible qu'il requiert, même si, à l'instant 2, l'incohérence n'est pas détectée. A ce même instant, l'approche par pavage ne détecte pas non plus, pour un découpage $d=6$, l'incohérence. Elle ne la détecte que pour un découpage $d=7$. Outre le fait qu'il soit difficile de déterminer une valeur optimale (diagnostic/durée) de découpage, cette approche peut requérir des durées considérables (143,68s. à $d=7$!). L'approche directe est plus rapide mais les durées d'exécution des procédures sont variables. Il apparaît nettement qu'ici, l'approche analytique est la plus efficace puisque le résultat exact est obtenu dans les meilleurs délais : 0,6s. dans tous les cas.

Le test partiel consistant à vérifier si $\{0\} \in \square \mathcal{A}(r_k)$ est très intéressant parce qu'il est très rapide ; mais certaines incohérences peuvent ne pas être détectées. L'approche indirecte analytique est naturellement la plus intéressante : le résultat est exact et la durée brève ne dépend pas de l'instant du test. Cependant, si l'hypomodèle n'est plus affine en les variables abstraites, cette approche peut ne plus être exploitable. Dans ces cas, l'approche directe

³³ $\text{rang} \begin{pmatrix} V_{0,k} & V_k^{(r_v)} \end{pmatrix} = \text{rang}(V_k^{(r_v)})$ et $\text{rang}(V_k^{(r_v)}) \neq p$

³⁴ non mesurable

³⁵ 6 découpages

³⁶ 7 découpages

semble la plus intéressante mais, malheureusement, dans un cas général, il n'est pas toujours facile d'obtenir l'ensemble des $v^{(r)}$ solutions de (59) et rien ne permet ne garantir que la procédure de minimisation a décelé un minimum global. Des fausses détections d'incohérences peuvent survenir ; ceci est fort gênant en diagnostic. Malgré des durées d'exécution inégales et parfois prohibitives, la méthode indirecte de pavage ne peut engendrer que des non-détections d'incohérence. Signalons que cette méthode n'est applicable que dans la mesure où les résultats de l'arithmétique des intervalles peuvent être exploités. Lorsque les champs sont non linéaires, une approche par balayage peut devenir le seul recours. Une telle approche est purement numérique et moins précise que le pavage ; les procédures peuvent devenir très longues.

Du métamodèle dynamique à l'hypomodèle à horizon fini

Nous avons considéré, dans la partie précédente, que l'hypomodèle à horizon fini (56) ou (58) était disponible. Ceci est naturel lorsque les incertitudes sont obtenues par un algorithme de caractérisation tels que ceux que nous avons proposés dans le chapitre 3 partie III. Cependant, en partant d'un modèle de connaissance, du type (45)³⁷, intégrant les incertitudes, l'obtention d'un hypomodèle à horizon fini facilement exploitable, n'est plus aussi simple que dans un contexte certain. Dans la partie précédente, nous avons montré que des tests de cohérence rapides et justes pouvaient être élaborés pour des hypomodèles à horizon fini affines en les variables abstraites ; sinon, les procédures deviennent très longues et guère applicables. Le problème qui se pose est de savoir comment il est possible d'obtenir un hypomodèle à horizon fini (56) affine en l'état en partant du métamodèle (45). C'est l'objet de cette sous-partie.

Pour obtenir un hypomodèle à horizon fini, il faut suivre une démarche que nous avons déjà utilisée pour les hypomodèles certains. Le principe est de se placer à un instant k quelconque. L'ensemble des états internes possibles d'un métamodèle est, a priori, inconnu ; il ne doit pas apparaître dans l'hypomodèle. Or, le métamodèle permet de déduire les domaines $\mathcal{A}(x_{k+i}) : i > 0$ à partir du domaine $\mathcal{A}(x_k)$.

Pour des raisons de commodité de représentation, nous définirons le cas limite de l'opérateur produit de la manière suivante : $\prod_{i=\alpha+1}^{\alpha} x_i y = y$ et l'ordre du produit par $\prod_{i=1}^{\alpha} M_i = M_{\alpha} \cdots M_1$ où les M_i sont des matrices, nous pouvons noter l'expression de x_{k+h} en fonction de x_k ainsi :

$$\begin{aligned}
 x_{k+1} = & \prod_{j=0}^{i-1} \left(A_0 + \sum_{i=1}^r A_i v_{i,k+j} \right) x_k + \cdots \\
 & \cdots + \sum_{l=0}^{i-1} \prod_{j=l+1}^{i-1} \left(A_0 + \sum_{i=1}^r A_i v_{i,k+j} \right) \left(\left(B_0 + \sum_{i=1}^r B_i v_{i,k+l} \right) u_{k+l} + \sum_{i=1}^r E_i v_{i,k+l} \right)
 \end{aligned} \tag{66}$$

³⁷ Nous n'étudierons pas le métamodèle (42) qui reste très général et, dans cette mesure, très complexe à étudier.

Enfin, en utilisant (45b), il vient :

$$\begin{aligned}
 \tilde{y}_{k+1} = & \prod_{j=0}^{i-1} \left(C_0 + \sum_{i=1}^r C_i v_{i,k+1} \right) \left(A_0 + \sum_{i=1}^r A_i v_{i,k+j} \right) x_k + \dots \\
 & \dots + \left(C_0 + \sum_{i=1}^r C_i v_{i,k+1} \right) \left(\prod_{l=0}^{i-1} \prod_{j=l+1}^{i-1} \left(\left(A_0 + \sum_{i=1}^r A_i v_{i,k+j} \right) \left(B_0 + \sum_{i=1}^r B_i v_{i,k+l} \right) \tilde{u}_{k+l} \right) \right) + \dots \quad (67a) \\
 & \dots + \left(C_0 + \sum_{i=1}^r C_i v_{i,k+1} \right) \prod_{l=0}^{i-1} \prod_{j=l+1}^{i-1} \left(\left(A_0 + \sum_{i=1}^r A_i v_{i,k+j} \right) \left(\sum_{i=1}^r E_i v_{i,k+l} \right) \right) \\
 & \dots + \left(\left(D_0 + \sum_{i=1}^r D_i v_{i,k+1} \right) \tilde{u}_{k+1} \right) + \sum_{i=1}^r F_i v_{i,k+1}
 \end{aligned}$$

Un hypomodèle d'horizon h repose sur h relations parmi lesquelles $h-1$ relations de type (67a) pour $i \in \{1, \dots, h-1\}$, et la relation issue de (45b) suivante :

$$\tilde{y}_k = \left(C_0 + \sum_{i=1}^r C_i v_{i,k} \right) x_k + \left(D_0 + \sum_{i=1}^r D_i v_{i,k} \right) \tilde{u}_k + \sum_{i=1}^r F_i v_{i,k} \quad (67b)$$

Un test de cohérence peut s'effectuer, théoriquement, en recherchant, pour chaque relation (67), l'ensemble des valeurs de x_k possibles. En notant \mathcal{E}_{k+i} ($i \in \{0, \dots, h-1\}$), l'ensemble des x_k vérifiant la relation (67) où \tilde{y}_{k+i} intervient, nous déduisons que la cohérence sera obtenue si :

$\bigcap_{i=0}^{h-1} \mathcal{E}_{k+i} \neq \{\emptyset\}$. Mais ceci n'est que théorie car, pratiquement, il n'est guère envisageable de déterminer les ensembles \mathcal{E}_{k+i} . Et puis, en plus de ne pas être applicable, cette approche ne conduit pas à des hypomodèles du type (56).

Une autre possibilité, plus proche de l'approche certaine, consiste à éliminer les termes contenant le vecteur indéterministe x_k des relations (67). Conformément au cas certain (cf. (11a)), il faut trouver une matrice certaine $Y^{(h)}$ telle que :

$$Y^{(h)} \begin{bmatrix} C_0 + \sum_{i=1}^r C_i v_{i,k} \\ \left(C_0 + \sum_{i=1}^r C_i v_{i,k+1} \right) \left(A_0 + A_v(v_k) \right) \\ \vdots \\ \left(C_0 + \sum_{i=1}^r C_i v_{i,k+h-1} \right) \prod_{j=0}^{h-2} \left(A_0 + A_v(v_{k+j}) \right) \end{bmatrix} = 0 \quad (68)$$

$$\text{avec } A_v(v_{k+j}) = \sum_{i=1}^r A_i v_{i,k+j}$$

Bien que possible, cette approche ne conduit pas à des hypomodèles affines en les variables abstraites qui sont, nous l'avons vu dans la partie précédente, très faciles à mettre en œuvre. Nous avons donc opté pour ne conserver que les termes certains du premier ordre ; cette approximation ne sera valide que lorsque :

$$\forall i \in \{1, \dots, r\}, \forall \alpha \geq 2 (\alpha \in \mathbb{N}), A_i^\alpha \ll A_0, B_i^\alpha \ll B_0, C_i^\alpha \ll C_0 \text{ et } D_i^\alpha \ll D_0 \quad (69)$$

Lorsque l'hypothèse (69) est vérifiée, la relation (67a) peut alors s'écrire de façon à isoler les termes contenant le vecteur de variables inconnues x_k :

$$\left(C_0 A_0^i + C_0 \left(\sum_{j=0}^{i-1} A_0^{i-j-1} A_i A_0^j v_{i,k+j} \right) + C_i A_0^i v_{i,k+i} \right) x_k = \dots$$

$$\dots \tilde{y}_{k+i} - \left(\sum_{j=0}^{i-1} (C_0 A_0^{i-j-1} B_0 \tilde{u}_{k+j}) + D_0 \tilde{u}_{k+i} \right) - \dots$$

$$\dots \sum_{i=1}^r \left(\begin{array}{l} C_0 (A_0^{i-1} B_i \tilde{u}_k + A_0^{i-1} E_i) v_{i,k} + \dots \\ \dots + C_0 \sum_{j=1}^{i-1} \left(A_0^{i-j-1} E_i + \sum_{l=i-j}^{i-1} A_0^{i-j-1} A_i A_0^{i-l-1} B_0 \tilde{u}_{k+l+j-1} + A_0^{i-j-1} B_i \tilde{u}_{k+j} \right) v_{i,k+j} \\ \dots + \left(F_i + C_i \sum_{l=0}^{i-1} A_0^{i-l-1} B_0 \tilde{u}_{k+l} + D_i \tilde{u}_{k+i} \right) v_{i,k+i} \end{array} \right) \quad (70a)$$

La relation (67b) peut se récrire de la même façon :

$$\left(C_0 + \sum_{i=1}^r C_i v_{i,k} \right) x_k = \tilde{y}_k - \left(D_0 + \sum_{i=1}^r D_i v_{i,k} \right) \tilde{u}_k + \sum_{i=1}^r F_i v_{i,k} \quad (70b)$$

En considérant un horizon h , l'hypomodèle à horizon fini s'obtient en éliminant de (70) les termes dans lesquels x_k apparaît. Ces termes doivent être annulés quelles que soient les variables abstraites. Ces variables peuvent dépendre du temps ou être invariantes ; cette dernière situation se rencontre souvent : si une variable abstraite fluctue peu, elle peut généralement être considérée comme invariante sur un horizon faible. Pour formuler l'hypomodèle, cumulons les termes de x_k de (70) sur un horizon h , dans le cas où les variables abstraites dépendent du temps et dans le cas où elles sont invariantes. Introduisons les matrices suivantes :

$$O_i^{d(h)} = \begin{bmatrix} C_i & 0 & 0 & 0 & \dots & 0 \\ C_0 A_i & C_i A_0 & 0 & 0 & \dots & 0 \\ C_0 A_0 A_i & C_0 A_i A_0 & C_i A_0^2 & 0 & \dots & 0 \\ C_0 A_0^2 A_i & C_0 A_0 A_i A_0 & C_0 A_i A_0^2 & C_i A_0^3 & \dots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ C_0 A_0^{h-2} A_i & C_0 A_0^{h-3} A_i A_0 & C_0 A_0^{h-4} A_i A_0^2 & C_0 A_0^{h-5} A_i A_0^3 & \dots & C_i A_0^{h-1} \end{bmatrix} \begin{array}{l} v = 0 \\ v = 1 \\ v = 2 \\ v = 3 \\ \vdots \\ v = h-1 \end{array} \quad (71a)$$

$$\begin{array}{cccccc} v_{i,k} & v_{i,k+1} & v_{i,k+2} & v_{i,k+3} & \dots & v_{i,k+h-1} \end{array}$$

$$O_i^{(h)} = \begin{bmatrix} C_i & & & & & \\ C_0 A_i + C_i A_0 & & & & & \\ C_0 A_0 A_i + C_0 A_i A_0 + C_i A_0^2 & & & & & \\ C_0 A_0^2 A_i + C_0 A_0 A_i A_0 + C_0 A_i A_0^2 + C_i A_0^3 & & & & & \\ \vdots & & & & & \\ C_0 A_0^{h-2} A_i + C_0 A_0^{h-3} A_i A_0 + C_0 A_0^{h-4} A_i A_0^2 + C_0 A_0^{h-5} A_i A_0^3 + \dots + C_i A_0^{h-1} & & & & & \end{bmatrix} \begin{matrix} \iota = 0 \\ \iota = 1 \\ \iota = 2 \\ \iota = 3 \\ \vdots \\ \iota = h-1 \end{matrix} \quad (71b)$$

$$\begin{matrix} \nu_{i,k} & \nu_{i,k+1} & \nu_{i,k+2} & \nu_{i,k+3} & \dots & \nu_{i,k+h-1} \end{matrix}$$

$$O_0^{(h)} = \begin{bmatrix} C_0 \\ C_0 A_0 \\ C_0 A_0^2 \\ C_0 A_0^3 \\ \vdots \\ C_0 A_0^{h-1} \end{bmatrix} \quad (71c)$$

Pour obtenir un hypomodèle à horizon fini h , il faut trouver une matrice $Y^{(h)}$ non nulle annihilant les termes où x_k apparaît et ce, quelles que soient les valeurs possibles de variables abstraites. Ceci revient à déterminer la matrice $Y^{(h)}$ telle que :

$$Y^{(h)} O^{(h)} = 0 ; O^{(h)} = \begin{bmatrix} O_0^{(h)} & \vdots & O_1^{i/d(h)} & O_2^{i/d(h)} & \dots & O_r^{i/d(h)} \end{bmatrix} \quad (72)$$

avec $O_i^{i/d(h)} = O_i^{i(h)}$ si $\nu_{i,k}$ est invariant et $O_i^{i/d(h)} = O_i^{d(h)}$ sinon

Il n'est guère réaliste de chercher à annuler exactement (72), surtout, si l'on se remémore l'hypothèse (19), que l'on ne construit qu'un hypomodèle approximatif. Pour réaliser (72), nous préconisons de décomposer la matrice $O^{(h)}$ en valeurs singulières : $O^{(h)} = USV^T$. Les valeurs singulières très inférieures à la plus grande des valeurs singulières seront négligées. Supposons qu'il y ait γ ($p \leq \gamma < p.h$)³⁸ valeurs singulières considérées comme nulles. La matrice $Y^{(h)}$, de dimension $\gamma \times (p.h)$, sera donnée par : $Y^{(h)} = [0 \ I_\gamma] U^T$. La même procédure de recherche de relations minimales que celle évoquée dans le cas certain pourra être utilisée ici³⁹.

En observant (72), nous pouvons conclure que l'hypomodèle incertain obtenu correspond un horizon supérieur ou égal à celui qu'aurait l'hypomodèle certain minimal, et en observant (71), nous pouvons ajouter que l'hypomodèle obtenu lorsque des variables abstraites dépendent du temps est pourvu d'un horizon supérieur ou égal à l'hypomodèle qui serait obtenu si certaines de ces variables abstraites ne dépendaient pas du temps. Or, nous avons montré, dans la partie II consacrée aux hypomodèles certains, que plus les horizons étaient petits, mieux cela valait pour le diagnostic. Nous en déduisons que, lorsque les variables abstraites fluctuent peu au

³⁸ La condition de minimalité n'est, a priori, pas vérifiée ici. Nous savons (cf. partie II.3.) qu'il existe p ($dim(y_k)$) équations minimales indépendantes. Pour être sûr de ne pas obtenir des hypomodèles ne portant que sur une partie d'un métamodèle, nous veillerons à choisir l'horizon h tel qu'il y ait au moins p valeurs singulières quasiment nulles dans $O^{(h)}$.

³⁹ Il faut néanmoins substituer aux conditions de rang des conditions portant sur les valeurs singulières.

cours du temps, il est généralement préférable de les considérer comme invariants sur l'horizon considéré.

Pour déterminer la matrice $Y^{(h)}$, nous avons réuni les relations (70) couvrant un horizon h . La matrice contenant tous les termes de gauche appartient au noyau de $Y^{(h)}$; l'hypomodèle est donc donné par multiplication à droite de la matrice des termes de droite par $Y^{(h)}$. Pour exprimer plus simplement le résultat, nous poserons :

$\forall i \in \{1, \dots, r\}$,

$$\varphi_i(\beta) = \begin{bmatrix} F_i \\ C_0 E_i \\ \vdots \\ C_0 A_0^{\beta-2} E_i \end{bmatrix}, \psi_i^0(\beta) = \begin{bmatrix} D_i \\ C_0 B_i \\ \vdots \\ C_0 A_0^{\beta-2} B_i \end{bmatrix} \text{ et } \psi_i^\alpha(\beta) = \begin{bmatrix} C_i A_0^{\alpha-1} B_0 \\ C_0 A_i A_0^{\alpha-1} B_0 \\ \vdots \\ C_0 A_0^{\beta-2} A_i A_0^{\alpha-1} B_0 \end{bmatrix} \quad (\alpha > 0) \quad (73a)$$

et

$$U_i^{(j,h)} = \begin{bmatrix} 0_{j,p \times m} & 0_{j,p \times m} & \cdots & 0_{j,p \times m} \\ \psi_i^j(h-j) & \psi_i^{j-1}(h-j) & \cdots & \psi_i^0(h-j) \end{bmatrix} \text{ et } V_i^{(j,h)} = \begin{bmatrix} 0_{j,p \times m} \\ \varphi_i(h-j) \end{bmatrix} \quad (73b)$$

Puis, en notant $\tilde{y}_{k,h} = [\tilde{y}_k \ \cdots \ \tilde{y}_{k+h-1}]^T$ et $\tilde{u}_{k,h} = [\tilde{u}_k \ \cdots \ \tilde{u}_{k+h-1}]^T$, l'hypomodèle d'horizon h obtenu lorsque les variables abstraites dépendent toutes du temps, est donné par :

$$Y^{(h)} \left(\tilde{y}_{k,h} - W^{(h)} \tilde{u}_{k,h} - \sum_{i=1}^r \left(\sum_{j=0}^{h-1} (U_i^{(j,k)} \tilde{u}_{k,h} - V_i^{(j,k)}) \right) v_{i,k+j} \right) = 0 \quad (74)$$

avec :

$$W^{(h)} = \begin{bmatrix} D_0 & 0 & 0 & 0 \\ C_0 B_0 & D_0 & 0 & 0 \\ \vdots & \cdots & \ddots & 0 \\ C_0 A_0^{h-2} B_0 & C_0 A_0^{h-3} B_0 & \cdots & D_0 \end{bmatrix}$$

Lorsque les variables abstraites sont toutes invariantes, l'hypomodèle devient :

$$Y^{(h)} \left(\tilde{y}_{k,h} - W^{(h)} \tilde{u}_{k,h} - \sum_{i=1}^r \left(\sum_{j=0}^{h-1} (U_i^{(j,k)} \tilde{u}_{k,h} - V_i^{(j,k)}) \right) v_i \right) = 0 \quad (75)$$

Enfin, en mélangeant les deux résultats (74) et (75), il est possible de prendre en compte le fait que seules certaines variables abstraites sont invariantes. Nous reconnaissons naturellement un champ affine en les variables abstraites; les résultats de la sous-partie précédentes peuvent être appliqués.

Avant d'illustrer ces résultats par un exemple, revenons à l'hypothèse (69). Parce qu'elle n'est jamais vérifiée totalement, l'usage des hypomodèles (74) ou (75) présente des risques de

fausse détection d'incohérence. Comment remédier à ce problème fort gênant ? Nous avons montré, dans le paragraphe *formulation générique* de cette partie, que pour les hypomodèles (74) ou (75) qui sont affines en les variables abstraites, l'approche indirecte analytique est la plus intéressante. Or, ces approches vérifient, selon qu'il s'agisse de l'hypomodèle (74) ou (75), si

$$\{0\} \in \mathcal{A} \left(Y^{(h)} \left(\tilde{y}_{k,h} - W^{(h)} \tilde{u}_{k,h} - \sum_{i=1}^r \left(\sum_{j=0}^{h-1} (U_i^{(j,k)} \tilde{u}_{k,h} - V_i^{(j,k)}) \right) \mathbf{v}_{i,k+j} \right) \right)$$

$$\text{ou } \{0\} \in \mathcal{A} \left(Y^{(h)} \left(\tilde{y}_{k,h} - W^{(h)} \tilde{u}_{k,h} - \sum_{i=1}^r \left(\sum_{j=0}^{h-1} (U_i^{(j,k)} \tilde{u}_{k,h} - V_i^{(j,k)}) \right) \mathbf{v}_i \right) \right)$$

Pour remédier au problème de fausse détection, nous proposons, lorsqu'une approche indirecte est préférée, d'augmenter de manière régulière les domaines abstraits qui précèdent. Nous avons montré dans le chapitre 3 paragraphe II.1, qu'une structure d'incertitude additive conduisait justement à une enveloppe régulière (et a fortiori à un accroissement régulier). Pour y parvenir, nous ajoutons aux hypomodèles des incertitudes additives dont les caractéristiques sont déterminées par la valeur du paramètre λ :

$$Y^{(h)} \left(\tilde{y}_{k,h} - W^{(h)} \tilde{u}_{k,h} - \sum_{i=1}^r \left(\sum_{j=0}^{h-1} (U_i^{(j,k)} \tilde{u}_{k,h} - V_i^{(j,k)}) \right) \mathbf{v}_{i,k+j} \right) - \lambda \begin{bmatrix} \mathbf{v}_{r+1} \\ \vdots \\ \mathbf{v}_{r+p} \end{bmatrix} = 0 \quad (76a)$$

$$Y^{(h)} \left(\tilde{y}_{k,h} - W^{(h)} \tilde{u}_{k,h} - \sum_{i=1}^r \left(\sum_{j=0}^{h-1} (U_i^{(j,k)} \tilde{u}_{k,h} - V_i^{(j,k)}) \right) \mathbf{v}_i \right) - \lambda \begin{bmatrix} \mathbf{v}_{r+1} \\ \vdots \\ \mathbf{v}_{r+p} \end{bmatrix} = 0 \quad (76b)$$

où λ est un scalaire positif

Ce paramètre peut alors être ajusté de façon à ce que, pour des campagnes de mesures *standards*, il n'existe pas de fausse détection d'incohérence. Ainsi, même si la précision des hypomodèles à horizon fini a été dégradée par les deux approximations successives, ils restent, pourvu que λ soit réglé correctement, justes.

exemple

Une fois de plus, nous partons de l'exemple de la machine à courant continu (cf. p. 169 et 183). Deux variables abstraites (indiquées 1 et 2) apparaissent dans l'équation récurrente définie par (49) ; elles représentent des incertitudes sur la valeur de la résistance équivalente et sur la valeur du frottement ; l'évolution de ces variables apparaît sur la figure XVIII.

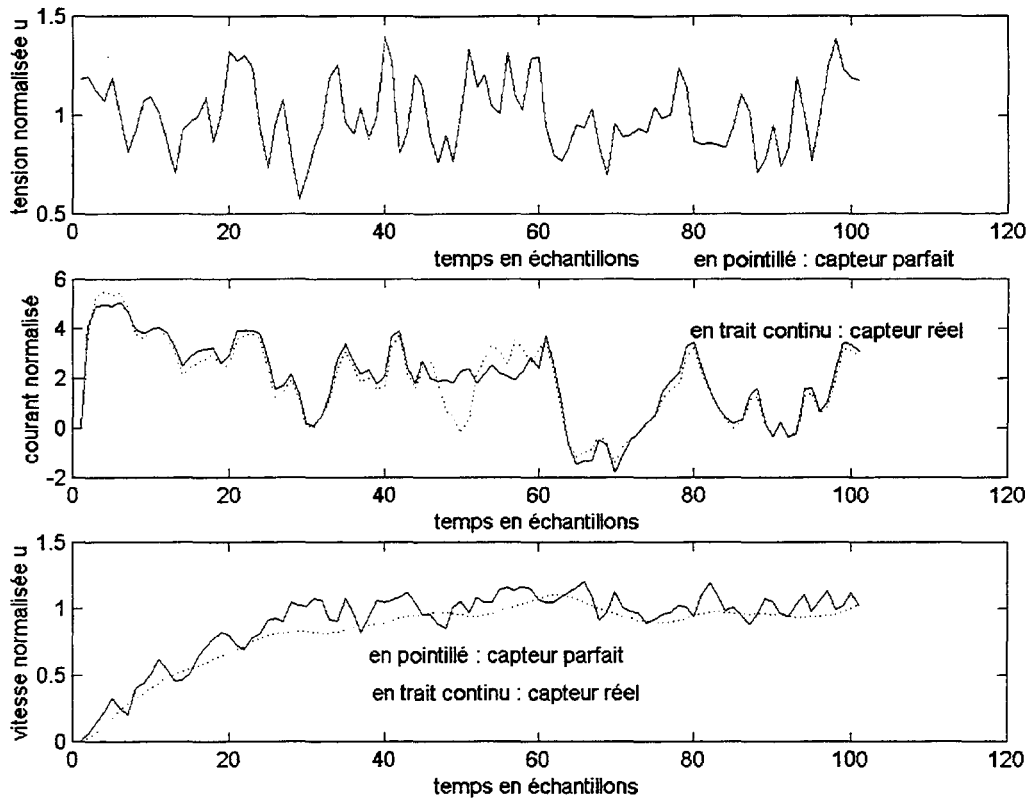


Figure XVI – Observations simulées du comportement de la machine

Quatre autres variables abstraites apparaissent dans l'équation de sortie (53) ; elles représentent des non-linéarités des capteurs (variables 3 et 4 sur figure XVII) ainsi que du bruit de mesure (variables 5 et 6). Les non-linéarités de capteurs sont représentées respectivement par $1,2.tanh(\Omega(t))$ et $6.tanh(i(t)/5)$ et ont été prises en compte par des variables abstraites $v_3(t)$ et $v_4(t)$. L'importance du bruit de mesure (compris entre -0.2 et $+0.2$) apparaît sur les observations du courant et de la vitesse de la figure XVI ; la tension est supposée parfaitement connue. Nous avons représenté un défaut entre les instants 45 et 60 par un comportement aléatoire⁴⁰ des observations du courant et de la vitesse.

Nous avons envisagé différents cas. Nous avons tout d'abord considéré que toutes les variables abstraites pouvaient fluctuer au cours du temps. Nous avons alors obtenu une matrice $Y^{(22)}$ d'horizon 22 ne présentant qu'un faible intérêt, le temps de simulation de l'hypomodèle devenant considérable.

⁴⁰ Mais raisonnable.

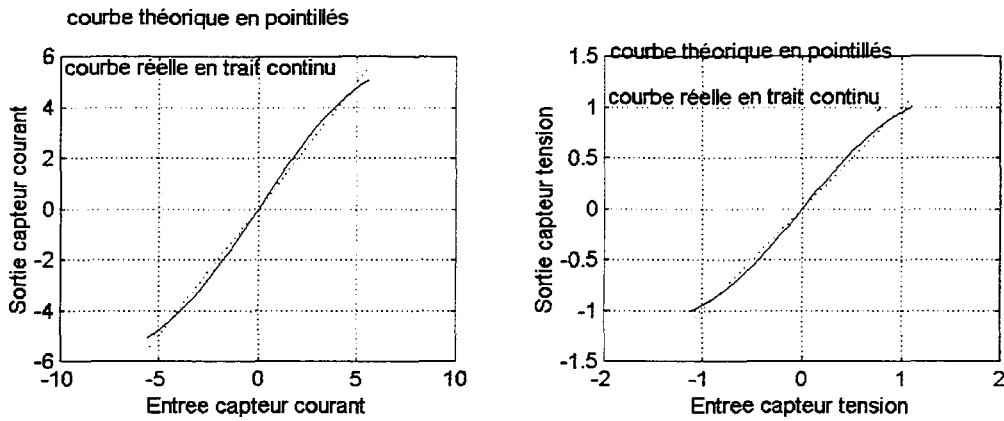


Figure XVII – Courbe d'étalonnage simulée des capteurs

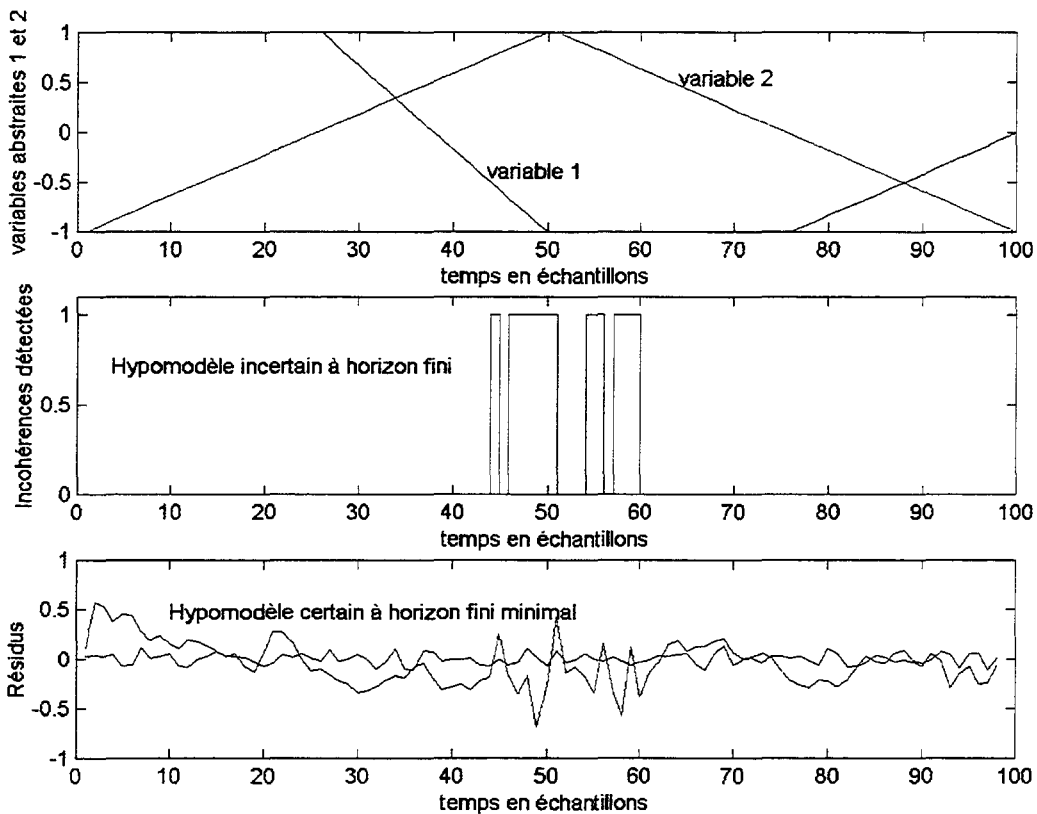


Figure XVIII – Evolution des variables abstraites 1 et 2 et résultats des tests de cohérence

Nous avons alors supposé que seules, les variables 5 et 6, correspondant aux bruits de mesure, n'étaient pas invariantes. Nous avons alors obtenu un hypomodèle d'horizon 4 défini par :

$$Y^{(4)} = \begin{bmatrix} -8.10^{-4} & 0.4279 & 6.9.10^{-3} & -0.814 & -5.10^{-3} & 0.3926 & 0 & 0 \\ 7.1.10^{-2} & 0.5263 & -0.3186 & -0.387 & 0.4933 & -0.3516 & -0.2468 & 0.1968 \end{bmatrix}$$

Pour ne pas surcharger ces pages, nous n'avons pas fait figurer l'hypomodèle de type (76). Précisons que nous avons retenu une valeur $\lambda=0,06$. Nous avons naturellement utilisé l'approche analytique indirecte pour tester la cohérence. Les résultats du test apparaissent sur figure XVIII, au milieu. Notons que les résultats de (76) sont affectés à l'instant k . De nombreuses incohérences sont détectées entre les instants 45 et 60 ! Il n'est pas étonnant que, durant le temps du défaut,

des cohérences soient constatées à certains instants. Rappelons, une fois de plus, que l'incohérence est moins prégnante que la cohérence, c'est-à-dire que malgré la présence des défauts, le comportement peut parfois s'apparenter à un comportement normal. Sur le tracé inférieur de la figure XVIII, nous avons représenté les résidus obtenus avec l'hypomodèle certain minimal d'horizon 2 défini par (11a), [PF3] (p.149) et par la matrice :

$$Y^{(2)} = \begin{bmatrix} 0.2133 & -0.628 & \vdots & -0.3815 & -0.6442 \\ 0 & -0.7131 & \vdots & -9.7 \cdot 10^{-3} & 0.701 \end{bmatrix}$$

On reconnaîtra l'impossible défi que de vouloir retrouver les résultats précédents par une technique de seuillage.

Nous avons alors supposé que toutes les variables abstraites étaient invariantes. L'hypomodèle que nous avons obtenu, est d'horizon 3. Il est défini par :

$$Y^{(3)} = \begin{bmatrix} 9 \cdot 10^{-4} & -11 \cdot 10^{-2} & \vdots & 11 \cdot 10^{-2} & 8 \cdot 10^{-4} & \vdots & -15 \cdot 10^{-3} & 0 \\ 4.43 \cdot 10^{-2} & -8.79 \cdot 10^{-2} & \vdots & 8.78 \cdot 10^{-2} & -4.5 \cdot 10^{-3} & \vdots & -1.36 \cdot 10^{-2} & 0 \end{bmatrix}$$

En retenant $\lambda=0,6$, nous avons représenté sur la figure XIX les résultats obtenus.

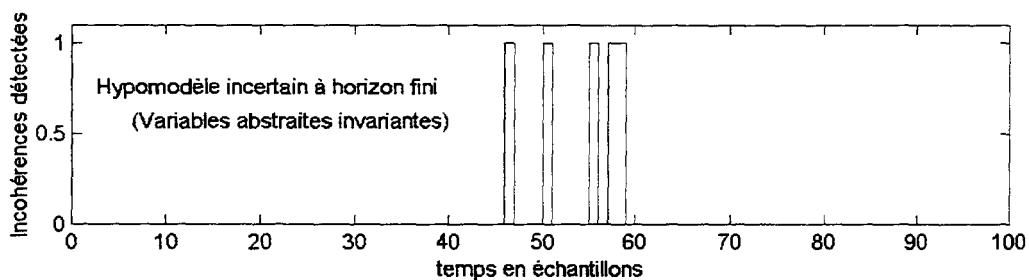


Figure XIX – Résultats des tests de cohérence (toutes les variables abstraites sont supposées invariantes)

Pour une valeur dix fois supérieure de λ , nous avons obtenu moins de résultats incohérents que précédemment. Cet hypomodèle est donc moins précis et ne permet de passer que d'un horizon de 4 à un horizon de 3. Il est moins intéressant que le précédent.

Il apparaît très clairement au travers de cet exemple que, lorsque des variables abstraites varient peu, il est préférable de les considérer comme invariante sur les horizons considérés, par contre, lorsque, comme des bruits de mesure, elles fluctuent de manière importante, il faut les tenir comme dépendantes du temps. Pour obtenir ces résultats, nous avons mis entre 30 s. et 1 min.⁴¹ avec un programme très général et non optimisé sous l'environnement Matlab 5.2.1 Les temps de calcul deviennent beaucoup plus intéressants que ceux qu'il aurait fallu dans le cas d'hypomodèles à horizon infini. Les hypomodèles à horizon fini semblent, dans le cas général, être beaucoup mieux adaptés aux procédures de diagnostic que les hypomodèles à horizon infini. De plus, si l'on se remémore les conclusions de la partie consacrée aux hypomodèles certains, plus les horizons sont faibles et meilleurs en sera le diagnostic.

⁴¹ Sur un Pentium 133MHz

IV. Conclusion

L'étude des hypomodèles certains nous a permis de mettre en évidence certaines propriétés génériques des hypomodèles. Nous avons tout d'abord montré que, pour élaborer des procédures de diagnostic, le métamodèle sur lequel repose un hypomodèle doit représenter exclusivement le comportement dont il faut tester la cohérence avec des observations. Un hypomodèle permettant de tester le comportement représenté par le métamodèle peut alors être obtenu ; il faut cependant éviter toute réduction ou découplage afin de ne pas modifier l'objectif de comportement à tester. Nous avons aussi montré que l'hypomodèle certain à horizon fini minimal était générateur de tous les autres hypomodèles certains associés à un métamodèle donné. En ce sens, toute l'information sur les incohérences est contenue dans cet hypomodèle, et, puisque l'horizon est minimal, l'instant d'apparition de l'incohérence sera déterminé plus précisément.

Nous nous sommes alors intéressés aux hypomodèles incertains qui sont plus adaptés au diagnostic de systèmes physiques. Il ressort que les hypomodèles à horizon fini sont plus intéressants que ceux à horizon infini, ces derniers requérant des temps de calcul considérables. Nous avons toutefois montré qu'il existait deux exceptions à cette règle : lorsque les hypomodèles ont un état interne de dimension 1 ou lorsqu'ils sont pourvus d'une structure purement additive. Nous avons alors étudié les moyens de tester la cohérence avec ces hypomodèles et nous avons constaté que les hypomodèles affines en les variables abstraites conduisaient à des procédures simples et exactes. Nous avons alors montré comment il était possible d'obtenir un hypomodèle à horizon fini affine en les variables abstraites en partant du cas général d'un métamodèle sous forme d'équations d'état.

Chapitre 5

les procédures

TABLE DES MATIÈRES

I.	Introduction	211
II.	Procédure orientée bon fonctionnement	212
II.1.	Métamodèles de la machine à courant continu	213
	Métamodèle complet.....	213
	Métamodèles réduits	215
II.2.	Environnement et simulation.....	218
	Loi de commande	218
	Défauts simulés.....	221
II.3.	Diagnostic global	224
	Algorithme	224
	Résultats des tests de cohérence	227
II.4.	Diagnostics locaux	230
	Algorithmes	230
	Résultats des tests de cohérence	232
III.	Procédure orientée mauvais fonctionnement.....	236
III.1.	Métamodèle de la machine à courant continu avec un court circuit.....	236
	Métamodèle complet.....	236
	Métamodèles réduits	238
III.2.	Spécificités des tests de cohérence	240
	Limite de la procédure orientée bon fonctionnement	240
	Procédure globale orientée mauvais fonctionnement	241
	Décomposition d'une procédure orientée mauvais fonctionnement	244
IV.	Conclusion.....	247

I. Introduction

Le cheminement que nous avons suivi nous a progressivement menés à mettre au point des hypomodèles permettant d'effectuer des tests de cohérence entre des observations et un ensemble de comportements possibles. Or, nous avons montré, au cours du chapitre introductif de ce mémoire, que toute opération de diagnostic procède nécessairement de tests de cohérence. Au stade où nous en sommes, le diagnostic automatique prenant en compte les incertitudes des systèmes physiques réels semble réalisable. Un aspect cependant reste à éclaircir ; un test de cohérence n'est, a priori, qu'une étape d'une opération de diagnostic mais il n'est pas, en soi, une procédure automatique. Nous avons évoqué les procédures, c'est-à-dire les différents types de tests et leur enchaînement, au cours du premier chapitre. Nous avons distingué deux grands types d'approches : les procédures orientées *bon fonctionnement* et les procédures orientées *mauvais fonctionnement*. Nous avons vanté les mérites de la décomposition théorique *Bon Mauvais Fonctionnement* (BMF) d'une procédure automatique complète pour conclure qu'il est naturel que les procédures orientées mauvais fonctionnement soient précédées de procédures orientées bon fonctionnement. De plus, pour produire un diagnostic, une procédure BF¹ doit commencer par un test du comportement global et se poursuivre, en cas d'incohérence, par des tests de comportements locaux. Des procédures de mauvais fonctionnement suivent éventuellement ; elles ne sont déclenchées qu'au cas où une procédure orientée bon fonctionnement a révélé une incohérence. Nous avons montré (chapitre 1 paragraphe III.3) comment, théoriquement, une procédure orientée bon fonctionnement peut ne déclencher que certaines procédures orientées mauvais fonctionnement. L'objectif de ce chapitre n'est pas de revenir sur ces aspects théoriques, mais plutôt d'examiner comment cela peut-être mis en œuvre dans la pratique. Nous partirons de la machine à courant continu pour examiner comment les développements théoriques précédents peuvent être exploités dans la synthèse de procédures automatiques de diagnostic. Les exemples présentés tentent de s'approcher le mieux possible de la réalité mais n'en sont pas pour autant des données réelles. Quelques résultats pratiques auraient certes été plus probants mais nous y avons renoncé pour plusieurs raisons. Tout d'abord, il faut remarquer que, malgré notre préoccupation pour l'applicabilité des outils développés, notre étude prospective n'en reste pas moins théorique dans le sens où nous n'avons jamais envisagé d'optimiser les algorithmes pour traiter un grand nombre de données ; notamment en ce qui concerne la caractérisation. Mais la principale raison est qu'une application nous semble être un travail devant être mené avec des spécialistes des systèmes physiques à surveiller ; les études sur les incertitudes à prendre en compte sur les paramètres équivalents d'un modèle d'un système physique ne sont guère courantes aujourd'hui. Les procédures de caractérisation peuvent,

¹ Orientée bon fonctionnement.

certes, être appliquées aveuglément, mais il est clair que, sans connaissance a priori, les modèles obtenus seront imprécis et les procédures ne pourront détecter que d'importants défauts. Il nous a semblé plus approprié de partir de métamodèles connus pour montrer l'intérêt de l'approche bornante en diagnostic.

Pour illustrer la synthèse d'une procédure de bon fonctionnement dans un contexte borné, la première partie de ce chapitre reprend l'exemple de la machine à courant continu que nous avons déjà utilisé dans les premier et quatrième chapitres. Pour plus de réalisme, nous avons inséré la machine dans une boucle de commande en vitesse. Nous vérifierons que les approches se fondant sur des modèles globaux de comportement s'appliquent aux systèmes physiques indépendamment de la commande et donc du régulateur. Nous discuterons des différents tests de cohérence reposant sur des métamodèles de bon fonctionnement et définirons une séquence appropriée de tests, autrement dit, une procédure automatique de diagnostic.

La seconde partie est consacrée aux procédures orientées mauvais fonctionnement. Nous nous sommes intéressé à la représentation sommaire d'un court-circuit entre les spires du rotor de la machine à courant continu. Après avoir posé un modèle du mauvais fonctionnement, nous montrerons que, conformément aux procédures orientées bon fonctionnement, la synthèse d'une procédure orientée mauvais fonctionnement repose tout entière sur la synthèse d'hypomodèle et de tests de cohérence. Après avoir examiné des différentes solutions, nous proposerons une procédure permettant de compléter le diagnostic déduit d'une procédure orientée bon fonctionnement. Nous montrerons comment ces résultats s'appliquent à la machine à courant continu et présenterons quelques résultats de simulation.

II. Procédure orientée bon fonctionnement

Pour montrer comment se met en œuvre une procédure orientée bon fonctionnement, nous avons choisi de revenir à l'exemple de la machine à courant continu. Le métamodèle que nous avons utilisé est relativement simple et facile à exploiter pour deux raisons. En négligeant les incertitudes, le modèle dispose d'une structure linéaire et surtout, en incluant l'équation mécanique, son vecteur d'état n'est que de dimension deux. Les machines synchrones et asynchrones possèdent, dans les représentations les plus simples, cinq équations différentielles² ! Il n'y a certes pas de limite théorique à l'application de procédures de diagnostic reposant sur une approche bornante, mais certains problèmes pratiques apparaissent, notamment au niveau de l'écriture et de la représentation des résultats. Nous avons jugé que la complexité ne s'imposait pas à l'illustration d'une réalisation de procédure de diagnostic.

² Equation mécanique prise en compte.

II.1. Métamodèles de la machine à courant continu

Métamodèle complet

Nous partons du métamodèle en temps continu de la machine à courant continu évoqué au cours du paragraphe III.1 du chapitre 1 (cf. (3)). Le comportement de la machine est décrit par les deux équations différentielles suivantes :

$$\begin{cases} \frac{d[i^*(t)]}{dt} = -\frac{R}{L}[i^*(t)] - \frac{\Phi}{L}[\Omega^*(t)] + \frac{1}{L}[u^*(t)] \\ \frac{d[\Omega^*(t)]}{dt} = \frac{\Phi}{J}[i^*(t)] - \frac{f}{J}[\Omega^*(t)] \end{cases}$$

où R est la résistance équivalente, L le coefficient d'inductance propre
 Φ est proportionnel au flux d'excitation traversant le rotor, f le frottement et J l'inertie
 et i^* , Ω^* , u^* , respectivement, le courant, la vitesse et la tension

Il est intéressant de normaliser les variables physiques par rapport à leur valeur nominale : le courant i_n , la vitesse Ω_n et la tension u_n . Ces valeurs correspondent à un régime permanent :

$$\begin{cases} R_n i_n + \Phi_n \Omega_n = u_n \\ \Phi_n i_n = f_n \Omega_n \end{cases} \Rightarrow \begin{cases} i_n = \frac{f_n}{\Phi_n} \Omega_n = \frac{f_n}{R_n f_n + \Phi_n^2} u_n \\ \Omega_n = \frac{\Phi_n}{f_n} i_n = \frac{\Phi_n}{R_n f_n + \Phi_n^2} u_n \end{cases}$$

En posant $[i(t)] = [i^*(t)]/i_n$, $[\Omega(t)] = [\Omega^*(t)]/\Omega_n$ et $[u(t)] = [u^*(t)]/u_n$, nous pouvons récrire le système d'équations différentielles :

$$\begin{cases} \frac{d[i(t)]}{dt} = -\frac{R}{L}[i(t)] - \frac{\Phi \Phi_n}{L f_n}[\Omega(t)] + \frac{R_n f_n + \Phi_n^2}{L f_n}[u(t)] \\ \frac{d[\Omega(t)]}{dt} = \frac{\Phi f_n}{J \Phi_n}[i(t)] - \frac{f}{J}[\Omega(t)] \end{cases} \quad (1)$$

Supposons qu'en bon fonctionnement, la valeur Φ soit égal à la valeur du régime nominal Φ_n et que, d'une part, la valeur de la résistance équivalente de la machine varie de plus ou moins 50% autour de sa valeur nominale et d'autre part, que le frottement fluctue de plus ou moins 10% autour de sa valeur nominale ; autrement dit, en introduisant les variables abstraites $v_1(t)$ et $v_2(t)$ avec $\mathcal{A}(v_1(t)) = [-1 \ 1]$ et $\mathcal{A}(v_2(t)) = [-1 \ 1]$, nous pouvons poser $R = (1 + \rho_1 v_1(t)) R_n$ et $f = (1 + \rho_2 v_2(t)) f_n$ et $\rho_1 = 0,5$, $\rho_2 = 0,1$. Le métamodèle (1) devient :

$$\begin{cases} \frac{d[i(t)]}{dt} = -(1 + \rho_1 v_1(t)) \frac{R_n}{L} [i(t)] - \frac{\Phi_n^2}{L f_n} [\Omega(t)] + \frac{R_n f_n + \Phi_n^2}{L f_n} [u(t)] \\ \frac{d[\Omega(t)]}{dt} = \frac{\Phi f_n}{J \Phi_n} [i(t)] - (1 + \rho_2 v_2(t)) \frac{f_n}{J} [\Omega(t)] \end{cases}$$

Ce modèle peut s'écrire en ne faisant apparaître que trois paramètres τ_1 (constante de temps électrique), τ_2 (constante de temps mécanique) et γ :

$$\begin{cases} \frac{d[i(t)]}{dt} = -\frac{1+\rho_1 v_1(t)}{\tau_1} [i(t)] + \frac{\gamma-1}{\gamma\tau_1} [\Omega(t)] + \frac{1}{\gamma\tau_1} [u(t)] \\ \frac{d[\Omega(t)]}{dt} = \frac{1}{\tau_2} [i(t)] - \frac{1+\rho_2 v_2(t)}{\tau_2} [\Omega(t)] \end{cases}$$

$$\text{avec } \tau_1 = \frac{L}{R_n}, \tau_2 = \frac{J}{f_n} \text{ et } \gamma = \frac{R_n i_n}{u_n} = \frac{R_n f_n}{R_n f_n + \Phi_n^2}$$

Par ailleurs, les capteurs de courant et de vitesse sont supposés non-linéaires (cf. figure II) ; ces non-linéarités sont représentées par des erreurs relatives de 10%. Enfin, ces capteurs sont affectés d'un bruit de mesure dont les amplitudes maximales sont égales à 5% des valeurs nominales des grandeurs observées. La machine étant commandée, la valeur de la tension appliquée à ses bornes est supposée parfaitement connue à tout instant et le convertisseur de puissance idéal³. Le métamodèle s'apparente aux relations (48) et (53) du chapitre 4 :

$$\begin{cases} \frac{d}{dt} \begin{bmatrix} [i(t)] \\ [\Omega(t)] \end{bmatrix} = A_c(v(t)) \begin{bmatrix} [i(t)] \\ [\Omega(t)] \end{bmatrix} + B_c \tilde{u}(t) \\ \begin{bmatrix} [\tilde{i}(t)] \\ [\tilde{\Omega}(t)] \end{bmatrix} = C_c(v(t)) \begin{bmatrix} [i(t)] \\ [\Omega(t)] \end{bmatrix} + F_c(v(t)) \end{cases} \quad (2a)$$

avec :

$$A_c(v(t)) = \begin{bmatrix} -1 & \gamma-1 \\ \tau_1 & \gamma\tau_1 \\ 1 & -1 \\ \tau_2 & \tau_2 \end{bmatrix} - \frac{\rho_1}{\tau_1} \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} v_1(t) - \frac{\rho_2}{\tau_2} \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} v_2(t), \quad B_c(v(t)) = \begin{bmatrix} 1 \\ \gamma\tau_1 \\ 0 \end{bmatrix} \quad (2b)$$

$$C_c(v(t)) = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} + \begin{bmatrix} \rho_3 & 0 \\ 0 & 0 \end{bmatrix} v_3(t) + \begin{bmatrix} 0 & 0 \\ 0 & \rho_4 \end{bmatrix} v_4(t), \quad F_c(v(t)) = \begin{bmatrix} \rho_5 \\ 0 \end{bmatrix} v_5(t) + \begin{bmatrix} 0 \\ \rho_6 \end{bmatrix} v_6(t)$$

où $\Omega(t)$ est la vitesse de rotation normalisée, $i(t)$ est le courant normalisé et $u(t)$ la tension d'alimentation normalisée, $\tau_1=L/R_n$ est la constante de temps électrique $\tau_2=J/f_n$ est la constante de temps mécanique et $\gamma=R_n i_n/u_n$ et les paramètres équivalents R_n et f_n représentent respectivement la résistance électrique et le frottement en fonctionnement nominal (cf chapitre 1 (3))
 L est le coefficient d'inductance propre du rotor et J l'inertie de la partie tournante.

Rappelons que les paramètres identifiés du métamodèle [Maksimovic et al,93] sont : $\tau_1=17\text{ms.}$, $\tau_2=1\text{s.}$ et $\gamma=0.118$, et que cette machine est commandée en tension au moyen d'un micro-contrôleur envoyant des commandes toutes les périodes de $\tau_c=10\text{ms.}$ Les incertitudes sur la valeur de la résistance et du frottement sont respectivement représentées par le

³ C'est-à-dire parfaitement représenté par un bloqueur d'ordre zéro et un gain unitaire.

paramètre $\rho_1=0,5$ associé à la variable abstraite bornée $v_1(t)$, et le paramètre $\rho_2=0,1$ associé à $v_2(t)$. Les variables $v_3(t)$ et $v_4(t)$ représentent les incertitudes sur les gains des capteurs, avec $\rho_3=\rho_4=0,1$, tandis que les variables $v_5(t)$ et $v_6(t)$ correspondent aux bruits de mesures : $\rho_5=\rho_6=0,05$.

La valeur d'une commande est supposée constante⁴ sur une période d'échantillonnage τ_e . Le métamodèle en temps continu n'étant pas directement exploitable, nous utiliserons les relations (46) du chapitre 4 pour déduire les matrices d'un métamodèle global en temps discret. Il s'écrit :

$$\begin{cases} \begin{bmatrix} [i_{k+1}] \\ [\Omega_{k+1}] \end{bmatrix} = (A_0 + A_1 v_{1,k} + A_2 v_{2,k}) \begin{bmatrix} [i_k] \\ [\Omega_k] \end{bmatrix} + (B_0 + B_1 v_{1,k} + B_2 v_{2,k}) \tilde{u}_k \\ \begin{bmatrix} \tilde{i}_k \\ \tilde{\Omega}_k \end{bmatrix} = (C_0 + C_3 v_{3,k} + C_4 v_{4,k}) \begin{bmatrix} [i_k] \\ [\Omega_k] \end{bmatrix} + F_5 v_{5,k} + F_6 v_{6,k} \end{cases} \quad (3a)$$

avec :

$$\begin{aligned} A_0 &\cong \begin{bmatrix} 0,5404 & -3,2817 \\ 7,5 \cdot 10^{-3} & 0,9720 \end{bmatrix} & A_1 &\cong \begin{bmatrix} -0,1606 & 0,4362 \\ -10^{-3} & 1,6 \cdot 10^{-3} \end{bmatrix} & A_2 &\cong \begin{bmatrix} 5 \cdot 10^{-6} & 1,8 \cdot 10^{-3} \\ -4 \cdot 10^{-6} & -10^{-3} \end{bmatrix} \\ B_0 &\cong \begin{bmatrix} 3,7413 \\ 2,05 \cdot 10^{-2} \end{bmatrix} & B_1 &\cong \begin{bmatrix} -0,4964 \\ -1,8 \cdot 10^{-3} \end{bmatrix} & B_2 &\cong \begin{bmatrix} 7,2 \cdot 10^{-6} \\ -7,2 \cdot 10^{-6} \end{bmatrix} \\ C_0 &\cong \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} & C_3 &\cong \begin{bmatrix} 0,1 & 0 \\ 0 & 0 \end{bmatrix} & C_4 &\cong \begin{bmatrix} 0 & 0 \\ 0 & 0,1 \end{bmatrix} & F_5 &\cong \begin{bmatrix} 0,05 \\ 0 \end{bmatrix} & F_6 &\cong \begin{bmatrix} 0 \\ 0,05 \end{bmatrix} \end{aligned} \quad (3b)$$

Ce métamodèle est un métamodèle complet de la machine à courant continu couplée à une charge de type frottement visqueux, mais il est aussi possible d'extraire des métamodèles partiels.

Métamodèles réduits

Pour extraire des métamodèles réduits ou partiels, deux solutions se présentent : partir du métamodèle en temps continu (2) ou directement du métamodèle en temps discret (3). Nous allons examiner ces deux solutions pour discuter de leurs différences.

Du métamodèle complet en temps continu aux métamodèles partiels

Pour obtenir des métamodèles partiels en partant du métamodèle complet en temps continu (2), il suffit d'extraire la première puis la seconde équation différentielle qui le constitue. Nous obtenons les métamodèles suivants :

⁴ Le convertisseur est représenté par un bloqueur d'ordre zéro.

$$\begin{cases} \frac{d[i(t)]}{dt} = -\frac{(1+\rho_1 v_1(t))}{\tau_1} [i(t)] + \frac{(\gamma-1)}{(1+\rho_4 v_4(t))\gamma\tau_1} \tilde{\Omega}(t) + \frac{1}{\gamma\tau_1} \tilde{u}(t) - \frac{(\gamma-1)(1+\rho_1 v_1(t))\rho_6 v_6(t)}{(1+\rho_4 v_4(t))\gamma\tau_1} \\ \tilde{i}(t) = (1+\rho_3 v_3(t))[i(t)] + \rho_5 v_5(t) \end{cases}$$

et

$$\begin{cases} \frac{d[\Omega(t)]}{dt} = -\frac{1+\rho_2 v_2(t)}{\tau_2} [\Omega(t)] + \frac{1}{(1+\rho_3 v_3(t))\tau_2} \tilde{i}(t) - \frac{\rho_5 v_5(t)}{\tau_2(1+\rho_3 v_3(t))} \\ \tilde{\Omega}(t) = (1+\rho_4 v_4(t))[\Omega(t)] + \rho_6 v_6(t) \end{cases}$$

Ces modèles seront linéarisés par rapport aux variables abstraites lors de la discrétisation ; or, les résultats seront inchangés si les métamodèles en temps continu sont déjà linéarisés (cf. formules (46) et (47) du chapitre 4). Ces métamodèles se simplifient donc au prix d'une approximation :

$$\begin{cases} \frac{d[i(t)]}{dt} = -\frac{(1+\rho_1 v_1(t))}{\tau_1} [i(t)] + \frac{\gamma-1}{\gamma\tau_1} (1-\rho_4 v_4(t))\tilde{\Omega}(t) + \frac{1}{\gamma\tau_1} \tilde{u}(t) - \frac{(\gamma-1)\rho_6 v_6(t)}{\gamma\tau_1} \\ \tilde{i}(t) = (1+\rho_3 v_3(t))[i(t)] + \rho_5 v_5(t) \end{cases} \quad (4a)$$

et

$$\begin{cases} \frac{d[\Omega(t)]}{dt} = -\frac{1+\rho_2 v_2(t)}{\tau_2} [\Omega(t)] + \frac{(1-\rho_3 v_3(t))}{\tau_2} \tilde{i}(t) - \frac{\rho_5 v_5(t)}{\tau_2} \\ \tilde{\Omega}(t) = (1+\rho_4 v_4(t))[\Omega(t)] + \rho_6 v_6(t) \end{cases} \quad (4b)$$

Pour transposer ces métamodèles en temps discret, il n'est pas juste de considérer que toutes les entrées sont constantes sur une période d'échantillonnage. En effet, les mesures $\tilde{\Omega}(t)$ et $\tilde{i}(t)$ varient continûment au cours du temps. Seule la commande $\tilde{u}(t)$ sera considérée comme constante sur une période d'échantillonnage tandis que les autres observations seront approximées par des segments de droite reliant les observations à deux instants d'échantillonnage consécutifs (cf. (2a) du chapitre 4). En utilisant les formules (46), du chapitre 4, pour la commande et (47), du chapitre 4, pour les autres observations, nous déduisons des métamodèles discrets correspondant respectivement à (4a) et (4b) :

$$\begin{cases} [i_{k+1}] = (A_0^{(i)} + A_1^{(i)} v_{1,k}) [i_k] + (B_0^{(iu)} + B_1^{(iu)} v_{1,k}) \tilde{u}_k + \dots \\ \dots + (B_0^{(i\Omega)} + B_1^{(i\Omega)} v_{1,k} + B_4^{(i\Omega)} v_{4,k}) \tilde{\Omega}_k + E_6^{(i)} v_{6,k} \\ \tilde{i}_k = (C_0^{(i)} + C_3^{(i)} v_{3,k}) [i_k] + (D_0^{(i\Omega)} + D_1^{(i\Omega)} v_{1,k} + D_3^{(i\Omega)} v_{3,k} + D_4^{(i\Omega)} v_{4,k}) \tilde{\Omega}_k + F_5^{(i)} v_{5,k} \end{cases} \quad (5)$$

$$\begin{aligned} A_0^{(i)} &= 0.5553 & A_1^{(i)} &= -0.1633 & B_0^{(iu)} &= 3.7686 & B_0^{(i\Omega)} &= -2.5128 & B_1^{(iu)} &= -0.5001 \\ \text{avec}^5 : & B_1^{(i\Omega)} &= 0.667 & B_4^{(i\Omega)} &= 0.2513 & C_0^{(i)} &= 1 & C_3^{(i)} &= 0.1 & D_0^{(i\Omega)} &= -1.824 & D_1^{(i\Omega)} &= 0.163 \\ & D_3^{(i\Omega)} &= 0.038 & D_4^{(i\Omega)} &= 0.1824 & E_6^{(i)} &= 1.2564 & F_5^{(i)} &= 0.05 \end{aligned}$$

$$\text{et } \begin{cases} [\Omega_{k+1}] = (A_0^{(\Omega)} + A_2^{(\Omega)} \nu_{2,k}) [\Omega_k] + (B_0^{(\Omega i)} + B_2^{(\Omega i)} \nu_{2,k} + B_3^{(\Omega i)} \nu_{3,k}) \tilde{i}_k + E_5^{(\Omega)} \nu_{5,k} \\ \tilde{\Omega}_k = (C_0^{(\Omega)} + C_4^{(\Omega)} \nu_{4,k}) [\Omega_k] + (D_0^{(\Omega i)} + D_2^{(\Omega i)} \nu_{2,k} + D_3^{(\Omega i)} \nu_{3,k} + D_4^{(\Omega i)} \nu_{4,k}) \tilde{i}_k + F_6^{(\Omega)} \nu_{6,k} \end{cases} \quad (6)$$

$$\begin{aligned} A_0^{(\Omega)} &= 0.99 & A_2^{(\Omega)} &= -9.9 \cdot 10^{-4} & B_0^{(\Omega i)} &= 9.9 \cdot 10^{-3} & B_2^{(\Omega i)} &= -10^{-5} & B_3^{(\Omega i)} &= -9.9 \cdot 10^{-3} \\ \text{avec} : & C_0^{(\Omega)} &= 1 & C_4^{(\Omega)} &= 0.1 & D_0^{(\Omega i)} &= 5 \cdot 10^{-3} & D_2^{(\Omega i)} &= -2 \cdot 10^{-6} & D_3^{(\Omega i)} &= -5 \cdot 10^{-4} & D_4^{(\Omega i)} &= 5 \cdot 10^{-4} \\ & E_5^{(\Omega)} &= -5 \cdot 10^{-4} & F_6^{(\Omega)} &= 0.05 \end{aligned}$$

Du métamodèle complet en temps discret aux métamodèles partiels

Des métamodèles partiels peuvent aussi être déduits directement du métamodèle complet en temps discret (3). En extrayant les équations récurrentes qui le constituent et en notant $(M)_{ij}$ l'élément apparaissant sur la ligne i et la colonne j de la matrice M , nous déduisons les métamodèles partiels suivants :

$$\begin{cases} [i_{k+1}] = ((A_0)_{11} + (A_1)_{11} \nu_{1,k} + (A_2)_{11} \nu_{2,k}) [i_k] + \frac{((A_0)_{12} + (A_1)_{12} \nu_{1,k} + (A_2)_{12} \nu_{2,k}) \tilde{\Omega}_k + \dots}{1 + \rho_4 \nu_{4,k}} \\ \dots + ((B_0)_{11} + (B_1)_{11} \nu_{1,k} + (B_2)_{11} \nu_{2,k}) \tilde{i}_k - \frac{((A_0)_{12} + (A_1)_{12} \nu_{1,k} + (A_2)_{12} \nu_{2,k}) \rho_6 \nu_{6,k}}{1 + \rho_4 \nu_{4,k}} \\ \tilde{i}_k = (1 + \rho_3 \nu_{3,k}) [i_k] + \rho_5 \nu_{5,k} \end{cases}$$

et

$$\begin{cases} [\Omega_{k+1}] = ((A_0)_{22} + (A_1)_{22} \nu_{1,k} + (A_2)_{22} \nu_{2,k}) [\Omega_k] + \frac{((A_0)_{21} + (A_1)_{21} \nu_{1,k} + (A_2)_{21} \nu_{2,k}) \tilde{i}_k + \dots}{1 + \rho_3 \nu_{3,k}} \\ \dots + ((B_0)_{21} + (B_1)_{21} \nu_{1,k} + (B_2)_{21} \nu_{2,k}) \tilde{\Omega}_k - \frac{((A_0)_{12} + (A_1)_{12} \nu_{1,k} + (A_2)_{12} \nu_{2,k}) \rho_5 \nu_{5,k}}{1 + \rho_3 \nu_{3,k}} \\ \tilde{\Omega}_k = (1 + \rho_4 \nu_{4,k}) [\Omega_k] + \rho_6 \nu_{6,k} \end{cases}$$

Pour devenir facilement exploitables, ces métamodèles doivent être linéarisés par rapport aux variables abstraites :

⁵ Les termes ayant pour exposant (Ωi) , $(i\Omega)$ (i) et (Ω) ont été déterminés à partir de (47) du chapitre 4, tandis que les termes en (iu) proviennent de (46) du chapitre 4. Certains termes peuvent être déterminés indifféremment avec (46) ou (47).

$$\begin{cases} [i_{k+1}] = ((A_0)_{11} + (A_1)_{11} v_{1,k} + (A_2)_{11} v_{2,k}) [i_k] + ((A_0)_{12} + (A_1)_{12} v_{1,k} + (A_2)_{12} v_{2,k} - \rho_4 (A_0)_{12} v_{4,k}) \tilde{\Omega}_k + \dots \\ \dots + ((B_0)_{11} + (B_1)_{11} v_{1,k} + (B_2)_{11} v_{2,k}) \tilde{u}_k - \rho_6 (A_0)_{12} v_{6,k} \\ \tilde{i}_k = (1 + \rho_3 v_{3,k}) [i_k] + \rho_5 v_{5,k} \end{cases} \quad (7)$$

et

$$\begin{cases} [\Omega_{k+1}] = ((A_0)_{22} + (A_1)_{22} v_{1,k} + (A_2)_{22} v_{2,k}) [\Omega_k] + ((A_0)_{21} + (A_1)_{21} v_{1,k} + (A_2)_{21} v_{2,k} - \rho_3 (A_0)_{21} v_{3,k}) \tilde{i}_k + \dots \\ \dots + ((B_0)_{21} + (B_1)_{21} v_{1,k} + (B_2)_{21} v_{2,k}) \tilde{u}_k - \rho_5 (A_0)_{12} v_{5,k} \\ \tilde{\Omega}_k = (1 + \rho_4 v_{4,k}) [\Omega_k] + \rho_6 v_{6,k} \end{cases} \quad (8)$$

Les métamodèles (7) et (8) s'obtiennent plus simplement que les métamodèles (5) et (6) puisque les matrices qui les constituent se déduisent directement du métamodèle en temps discret. Néanmoins, contrairement au métamodèle (5), le métamodèle (7) n'exploite pas uniquement l'équation différentielle correspondant à la partie électrique ! En effet, dans la première équation récurrente du métamodèle complet en temps discret, l'information correspondant à la première équation différentielle du métamodèle complet en temps continu se trouve diluée avec celle contenue dans la seconde équation différentielle. Il en va de même pour les métamodèles (6) et (8). Les métamodèles issus du métamodèle complet en temps continu sont les seuls qui font intervenir exclusivement la connaissance correspondant à la partie électrique pour le premier et à la partie mécanique pour le second. Il est naturel, pour l'exemple de la machine à courant continu, de retenir les métamodèles réduits (5) et (6).

Trois métamodèles ont été formulés, le métamodèle complet (3) contient toute la connaissance disponible et exploitable par une procédure de diagnostic dédiée à la machine à courant continu tandis que les métamodèles partiels ou réduits (5) et (6), ne contiennent chacun qu'une fraction de la connaissance disponible. Avant d'en venir aux procédures, nous consacrons le paragraphe suivant à la description de l'environnement dans lequel la machine est supposée fonctionner ; nous présenterons certains résultats simulés qui nous serviront ultérieurement.

II.2. Environnement et simulation

Loi de commande

Pour plus de réalisme, il est naturel de considérer que la machine s'inscrit dans une boucle de commande de vitesse. Nous avons choisi d'effectuer un retour d'état et d'ajouter un correcteur proportionnel intégral⁶ : la structure de la commande obtenue est décrite par la figure I. Les gains permettant d'aboutir aux variables normalisées n'apparaissent pas

⁶ Notre propos n'est pas ici de rechercher la meilleure loi de commande possible, mais de montrer que les procédures de diagnostic s'appliquent sans difficulté aux systèmes physiques commandés.

explicitement, ils peuvent être ramenés au niveau du convertisseur de puissance et des convertisseurs analogique/numérique.

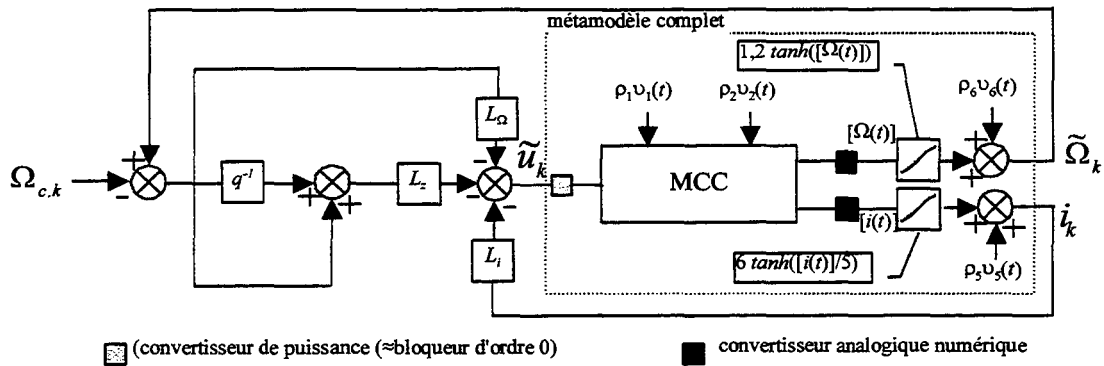


Figure I – Structure de la commande de la machine asynchrone

Le rectangle en pointillé représente la partie représentée par le métamodèle complet (1) ou (3), en temps discret. Remarquons que les deux non-linéarités $6.tanh([i(t)]/5)$ et $1,2.tanh([\Omega(t)])$ ont été représentées dans le métamodèle par une erreur relative de 10% ; la première non-linéarité est présentée par le terme $\rho_3 v_{3,k}$ et la seconde par $\rho_4 v_{4,k}$.

L'intégrateur fait intervenir une variable z_k vérifiant $z_{k+1} = z_k + \Omega_k - \Omega_{c,k}$. La loi de commande est décrite par $u_k = -L_i i_k - L_\Omega (\Omega_k - \Omega_{c,k}) - L_z z_k$, où $\Omega_{c,k}$ est la consigne de vitesse souhaitée⁷. Les gains ont été réglés de telle sorte qu'en négligeant les incertitudes, les pôles du système bouclé soient égaux à 0,8, 0,85 et 0,9 c'est-à-dire $L_i = 3.10^{-3}$, $L_\Omega = 1,406$ et $L_z = 0,145$.

Pour construire une simulation pertinente, nous avons fait évoluer les variables abstraites sur leur domaine abstrait. La partie supérieure de la figure II représente les variations des variables abstraites $v_{1,k}$ et $v_{2,k}$ représentant respectivement les variations de plus ou moins 50% de la résistance et de plus ou moins 10% du frottement. Sur la partie centrale, figurent les bruits de mesure aléatoires, c'est-à-dire les termes $\rho_5 v_{5,k}$ pour le capteur de courant et $\rho_6 v_{6,k}$, pour le capteur de vitesse. Nous pouvons constater que ces bruits sont assez importants puisqu'ils représentent 5% des valeurs nominales du courant et de la tension. La partie inférieure montre que les non-linéarités de capteurs peuvent être représentées par des variables abstraites $v_{3,k}$ pour le capteur de courant et $v_{4,k}$, pour le capteur de vitesse. En effet, leur caractéristique $\rho_3 = \rho_4 = 0,1$ conduit aux secteurs grisés de la figure II (partie inférieure) qui contiennent les courbes non-linéaires $6.tanh([i(t)]/5)$ et $1,2.tanh([\Omega(t)])$.

⁷ La commande n'intègre aucun contrôle dynamique du courant. Le lecteur intéressé pourra se reporter à [Maksimovic,93]. L'exemple proposé reste académique.

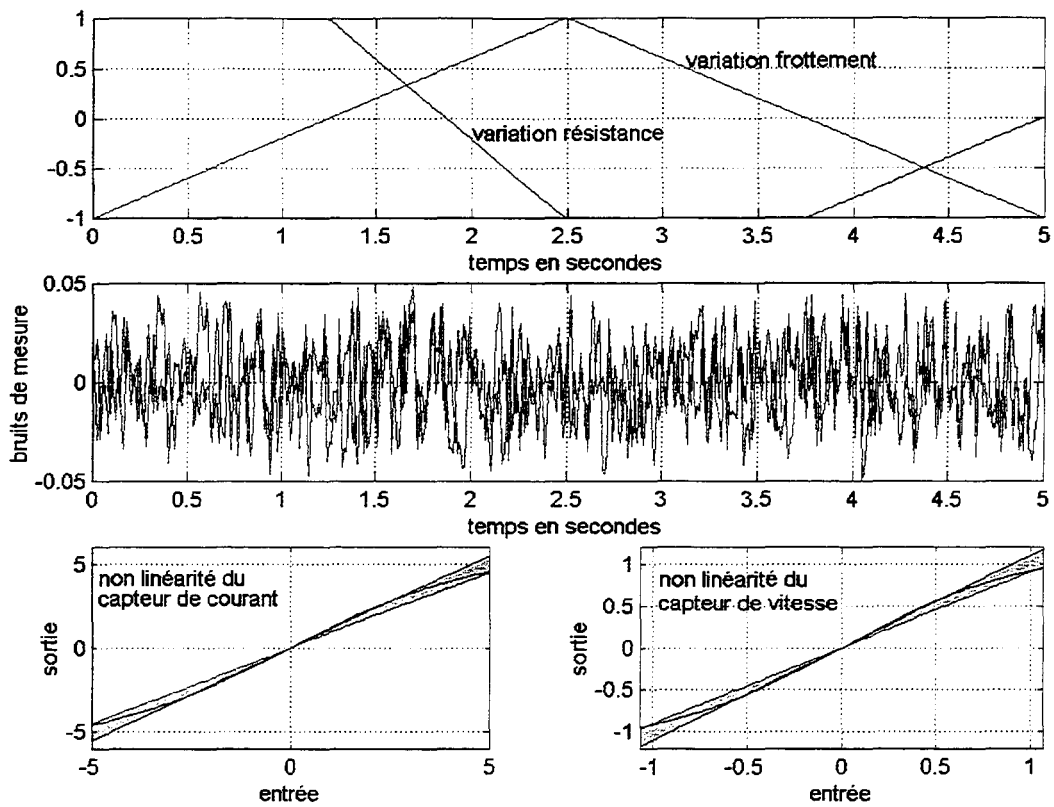


Figure II – Valeurs simulées des variables abstraites

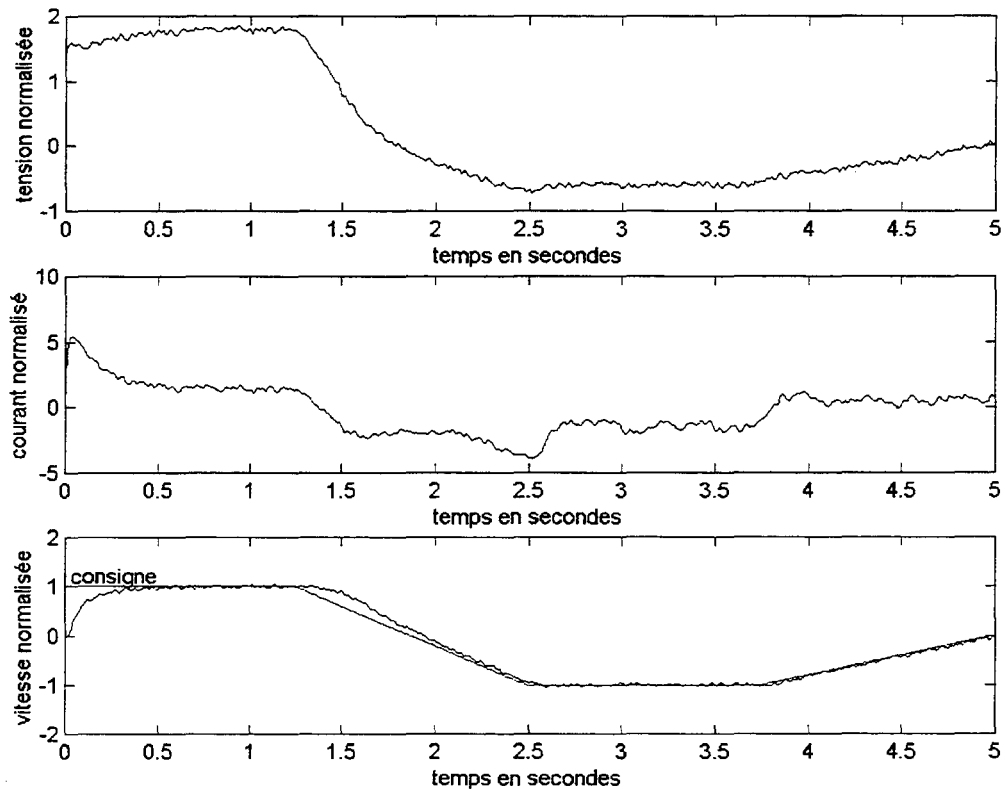


Figure III – Réponse de la machine (bon fonctionnement)

Pour simuler la réponse de la machine lorsqu'une consigne de vitesse est demandée, nous

sommes partis du métamodèle en temps continu. Nous avons supposé que les valeurs de la résistance et du frottement étaient invariantes sur une période d'échantillonnage $\tau_e=10\text{ms}$. Nous avons alors calculé le métamodèle discret non approximé correspondant, puis nous avons pris en compte l'intégrateur : $z_{k+1}=z_k+\Omega_k-\Omega_{c,k}$ et appliqué la loi de commande $u_k=-L_i i_k-L_\Omega (\Omega_k-\Omega_{c,k})-L_z z_k$. Les résultats que nous avons obtenus sont représentés sur la figure III. La consigne apparaît sur la même courbe que la réponse en vitesse. Malgré les incertitudes, le correcteur assure correctement sa fonction : la vitesse mesurée est très proche de la valeur de consigne.

Défauts simulés

Pour tester les procédures de diagnostic, nous avons simulé plusieurs types de défauts. Tout d'abord, nous avons représenté une augmentation moyenne du coefficient de frottement à cinq fois sa valeur nominale entre les instants 2,5s. et 3,75s., pouvant représenter un frottement parasite. Cette variation peut être exprimée avec la variable abstraite $v_{2,k}$, dont la valeur sort naturellement du domaine normal [-1 1]. Cet accroissement du frottement apparaît sur la figure IV.

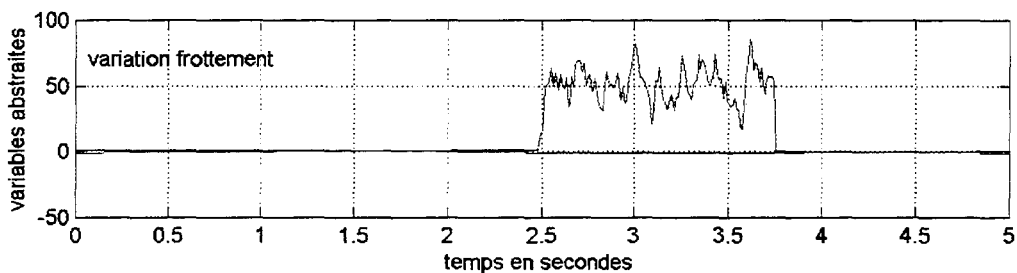


Figure IV – Défaut correspondant à une multiplication par 5 de la valeur du frottement (entre 2,5 et 3,75s.)

Les réponses calculées lorsque ce défaut survient sont représentées sur la figure V. Les instants d'apparition et de disparition du défaut ont été repérés par des traits verticaux en pointillés. Le correcteur continue à assurer sa fonction, il maintient la vitesse proche de la consigne. On remarque tout de même, en comparant la vitesse obtenue à celle apparaissant sur la figure III, une poursuite de la consigne un peu dégradée durant le temps où le défaut est présent. Par contre, toujours en comparant la figure V à la figure III, nous constatons que le défaut a engendré des différences au niveau de la tension et du courant. Le régulateur agit naturellement sur la tension de commande pour maintenir la vitesse à sa valeur de consigne ; du fait de l'accroissement du frottement, la tension est accrue en valeur absolue (de $-0,5U_n$ à $-1,3U_n$). Un couple électromoteur important est ainsi développé pour compenser l'action du défaut ; ceci se traduit par une variation importante du courant circulant dans la machine (de $-I_n$ à $-5I_n$). Si la machine avait été pilotée en boucle ouverte, au lieu de la vitesse, c'est la tension qui aurait été inchangée : le problème ne serait fondamentalement pas différent puisque c'est la cohérence entre un modèle et plusieurs observations qui serait testée.

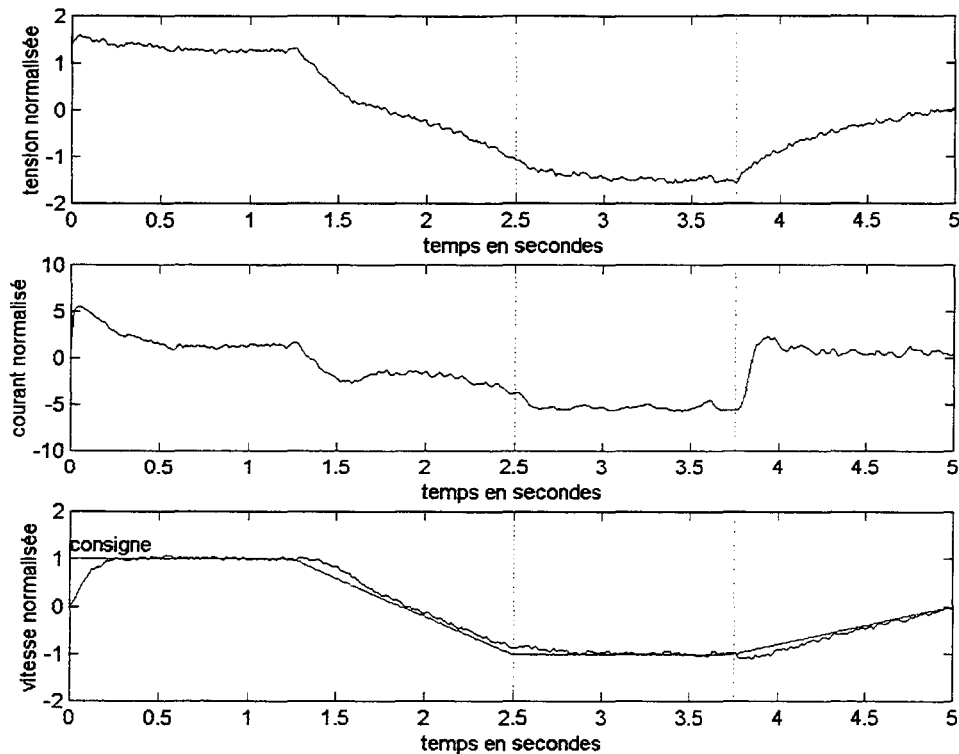


Figure V – Réponse pour une multiplication par 5 du frottement entre 2,5s. et 3,75s.

Nous avons aussi envisagé un second défaut correspondant à une augmentation de la résistance atteignant seize fois sa valeur nominale en moyenne. Ce défaut peut correspondre à un problème de connexions défectueuses. La variation de la résistance peut être représentée par une variation de la variable abstraite $v_{1,k}$ (cf. figure VI) tandis que la variable $v_{2,k}$, correspondant à la variation du frottement, suit la même évolution que celle décrite sur la figure II.

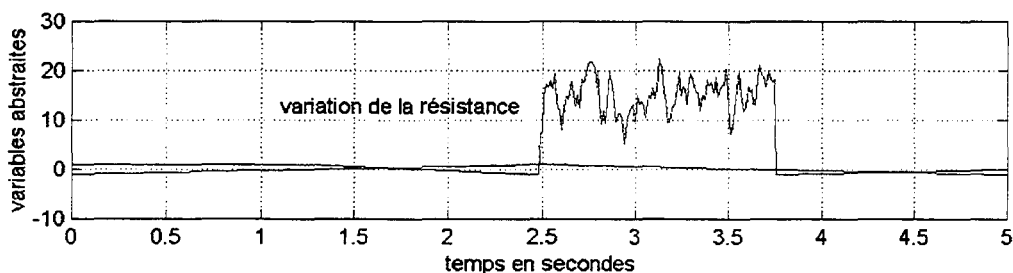


Figure VI – Variation de la variable abstraite $v_{1,k}$ correspondant à la multiplication parasite de la résistance par 16 (entre 2,5 et 3,75s.)

Les réponses obtenues figurent sur la figure VII. Nous constatons que la vitesse reste proche de la valeur de consigne, mais en comparant ces résultats aux figures III et V, nous remarquons que le courant reste quasiment inchangé durant le défaut à l'exception d'une variation importante lorsque l'anomalie se résorbe ($t=3,75s.$). En revanche, la tension décroît fortement puisqu'au lieu de $-0,5U_n$, elle atteint $-2,5U_n$.

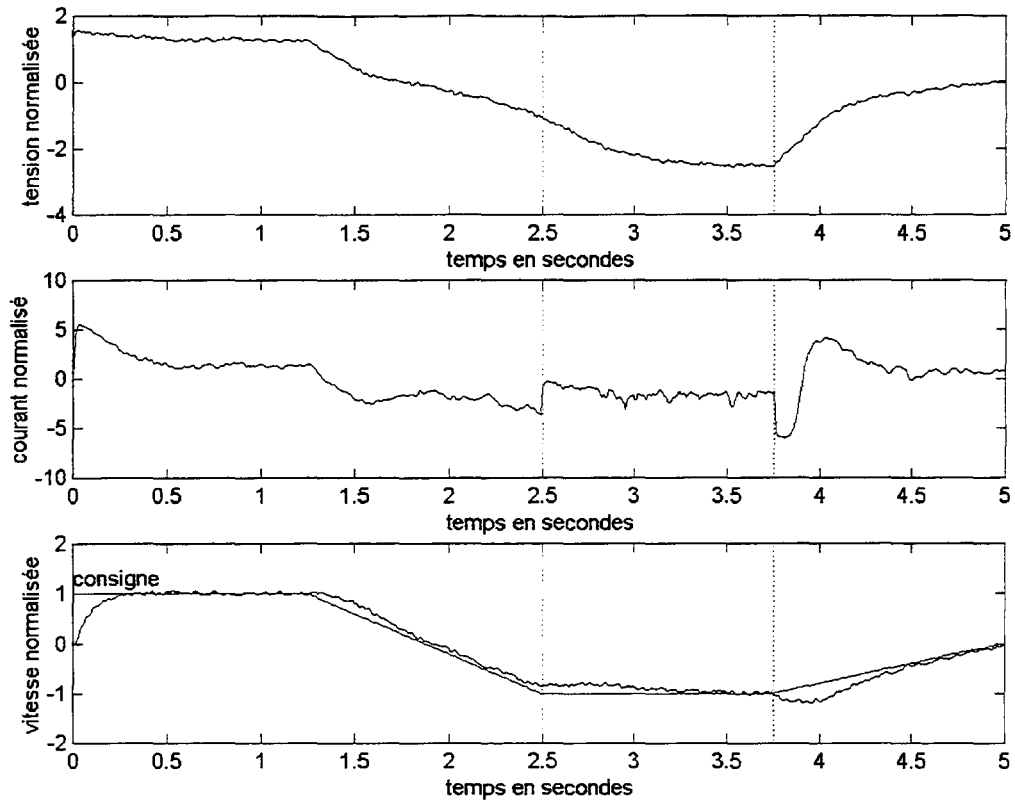


Figure VII – Réponse en présence d’une augmentation de la résistance (16 fois sa valeur nominale)

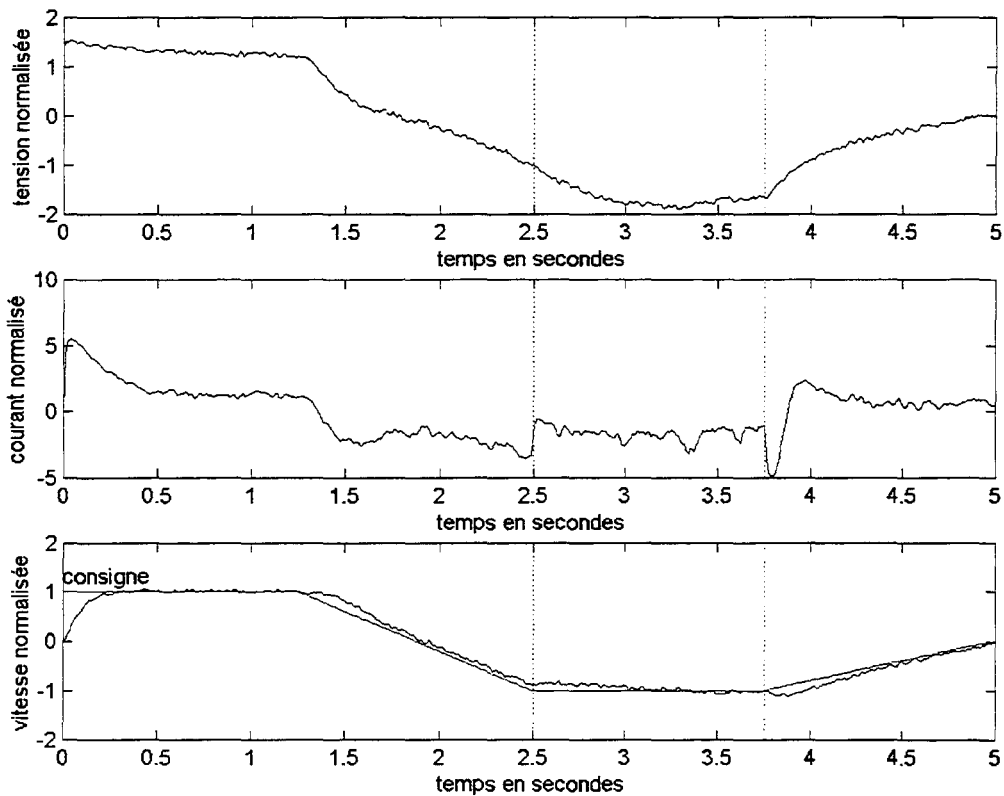


Figure VIII - Réponse en présence d’une augmentation de la résistance (7 fois sa valeur nominale)

Pour une augmentation moyenne de la résistance à seulement sept fois sa valeur nominale, la figure VIII fait apparaître de faibles différences avec le bon fonctionnement (figure III). Le

courant reste inchangé à l'exception d'un artefact à l'instant de disparition de l'anomalie (3,75s.) tandis que la tension n'atteint que $-1,6U_n$ au lieu de $-0,5U_n$ en bon fonctionnement. En tenant compte des incertitudes, nous pouvons d'ores et déjà prédire qu'il sera difficile de mettre en évidence une telle anomalie.

Avant d'aborder les procédures orientées bon fonctionnement, rappelons que, contrairement aux modèles utilisés pour les simulations, les métamodèles (3), (5) et (6), exploités pour la synthèse de procédures de diagnostic, sont linéarisés autour de faibles valeurs des variables abstraites ; autrement dit, plus les variables deviennent importantes, et moins les métamodèles sont précis. Nous verrons, dans la partie suivante, comment obtenir un hypomodèle juste malgré l'approximation.

II.3. Diagnostic global

Algorithme

Il est naturel de commencer par observer globalement la machine à courant continu pour déterminer si son comportement global est cohérent avec l'ensemble des comportements normaux possibles. Or, nous avons constaté qu'un test de cohérence se fonde sur un hypomodèle puis, au cours du chapitre 4, nous avons dégagé deux grands types d'hypomodèles : ceux dont l'horizon est fini et ceux à horizon infini. En partant des hypomodèles certains, nous avons montré que, plus les horizons étaient courts, mieux cela valait pour le diagnostic car, en plus de contenir toute l'information sur une incohérence, l'horizon bref permet de mieux déterminer les instants d'apparition d'un défaut : une incohérence détectée à un instant k à partir d'un hypomodèle d'horizon h (cf. (11a) du chapitre 4) signifie qu'un défaut s'est produit entre les instants k et $k-h$. Par ailleurs, dans ce même chapitre, nous avons constaté que les hypomodèles à horizon fini ne pouvaient pratiquement être exploités qu'en présence de structures d'incertitudes purement additives ou lorsque le vecteur de variables à intégrer se réduit à une seule variable. Or, le métamodèle complet de machine à courant continu (3) contient deux variables liées à intégrer. Les deux raisons que nous venons d'exposer supposent l'emploi d'un hypomodèle à horizon fini.

Pour la synthèse d'un hypomodèle à horizon fini exploitant la connaissance contenue dans (3), nous renvoyons le lecteur à la partie intitulée : *Du métamodèle dynamique à l'hypomodèle à horizon fini*, de la partie III.3 du chapitre 4. Concernant la machine à courant continu, nous obtenons, en suivant la procédure décrite et en considérant comme invariantes sur cet horizon (30ms.), les variables abstraites $v_{1,k}$, $v_{2,k}$, $v_{3,k}$ et $v_{4,k}$, correspondant respectivement aux variations de résistance, du frottement et aux non-linéarités de capteurs, un hypomodèle précis d'horizon trois. La nature temporelle des bruits de mesure est, quant à elle, maintenue dépendante du temps. La structure de l'hypomodèle est donnée par l'équation

(76) du chapitre 4. L'hypomodèle est défini par la matrice $Y^{(3)}$ suivante :

$$Y^{(3)} = \begin{bmatrix} 4.5 \cdot 10^{-3} & 0.6994 & 1.4 \cdot 10^{-3} & -0.7147 & 0 & 0 \\ -1.26 \cdot 10^{-2} & -0.6294 & 1.41 \cdot 10^{-2} & -9.15 \cdot 10^{-2} & -1.11 \cdot 10^{-2} & 0.7714 \end{bmatrix}$$

Pour des raisons de place et parce que cela ne présente que peu d'intérêt, nous n'avons pas détaillé la matrice $W^{(3)}$ et les vecteurs $U_i^{(j,3)}$, $V_i^{(j,3)}$; pour $i \in \{1, \dots, 4\}$ et $j \in \{0, \dots, 6\}$. Pour les obtenir, il suffit d'évaluer les formes analytiques (71), (72) et (74) du chapitre 4.

Pour aboutir à cet hypomodèle, deux approximations ont été effectuées : la première lors de la transposition du métamodèle en temps discret et la seconde, lorsque, pour déduire un hypomodèle à horizon fini, nous avons négligé les termes contenant des produits de variables abstraites. Or, en diagnostic, nous avons un impératif de justesse. Dans l'hypomodèle (76) du chapitre 4, le scalaire λ permet de compenser ces approximations. Pour régler ce paramètre, nous avons effectué des tests de cohérence à partir de l'hypomodèle pour différentes valeurs de λ . Nous avons retenu la plus faible valeur de λ telle qu'aucune incohérence ne soit détectées pour les mesures issues d'un comportement normal (cf. figure III), c'est-à-dire $\lambda=0,016$.

Pour évaluer les performances de la procédure, nous avons optimisé les tests de cohérence décrits dans la partie III.3 du chapitre 4 en combinant les approches directes et les approches indirectes. En effet, même s'ils ne sont pas suffisants, certains tests de l'approche directe sont très rapides à mettre en œuvre et permettent, dans un certain nombre de cas, d'éviter d'avoir systématiquement recours à l'approche indirecte analytique, qui rappelons-le, conduit toujours à un résultat exact. A chaque instant, l'hypomodèle peut se mettre sous la forme $V_k^{(rv)} v^{(rv)} + V_{0,k} = 0$, où les matrices $V_k^{(rv)}$ et $V_{0,k}$ sont des matrices contenant deux lignes fonctions des différentes observations sur l'horizon considéré ; le vecteur $v^{(rv)}$ contient 10 variables abstraites normalisées parmi lesquelles les quatre invariants et les deux dépendantes du temps à chacun des trois instants couverts par l'horizon. L'algorithme que nous avons utilisé se structure comme l'illustre la figure IX. Nous avons constaté que ce premier niveau de test permet d'accélérer considérablement la procédure, notamment dans les zones où la cohérence est franche.

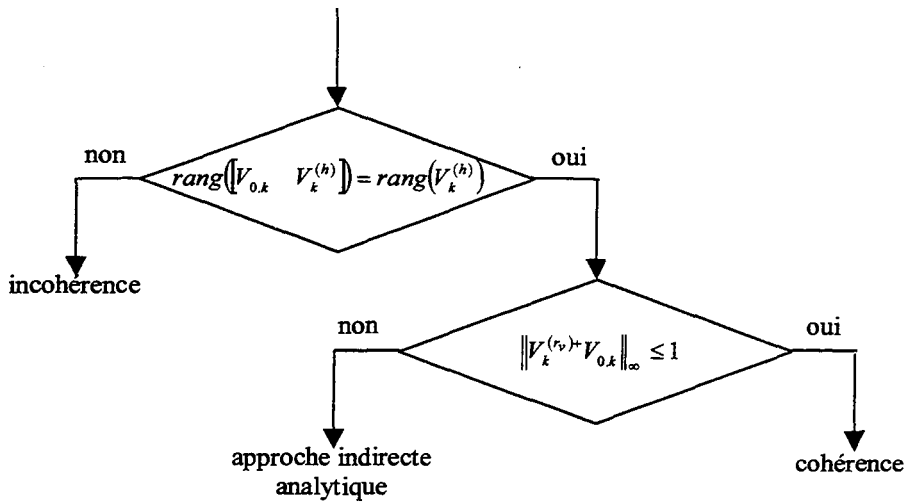


Figure IX – Structure du test de cohérence

Enfin, si le premier niveau de test n'a pas suffi à trancher, nous aurons recours à l'approche indirecte analytique. Pour accélérer la procédure, nous cesserons les tests dès qu'une incohérence aura été détectée. Afin de poursuivre, remarquons que, lorsque les variables abstraites sont normalisées, il n'est pas nécessaire d'avoir recours à l'arithmétique des intervalles pour évaluer les espaces abstraits de forme linéaire. En effet, si α est une forme linéaire déterministe de dimension quelconque et v un vecteur de variables abstraites normalisées de même dimension, nous vérifierons :

$$A(\alpha v) = \left[-\|\alpha^T\|_1, \|\alpha^T\|_1 \right] \quad (7)$$

En utilisant ce résultat, il est facile d'évaluer la cohérence par rapport à l'orthotope circonscrit. L'hypomodèle peut à chaque instant se mettre sous la forme $V_k^{(rv)} v^{(rv)} + V_{0,k}$. Pour tester si $\{0\} \in \square A(V_k^{(rv)} v^{(rv)} + V_{0,k})$, il suffit de vérifier si :

$$\|[1 \ 0] V_k^{(rv)}\|_1 \geq \|[1 \ 0] V_{0,k}\|_1 \text{ et } \|[0 \ 1] V_k^{(rv)}\|_1 \geq \|[0 \ 1] V_{0,k}\|_1$$

En tenant compte du fait que le domaine abstrait du champ : $V_k^{(rv)} v^{(rv)} + V_{0,k}$ est de dimension 2, nous pouvons d'ores et déjà conclure qu'il ne peut exister, au plus, qu'un sous-ensemble résoluble (cf. chapitre 2 partie III.2) de dimension 2. Le champ étant affine en les variables abstraites, nous concluons qu'il est dissociable (cf. chapitre 2 partie III.2). Par conséquent, le domaine abstrait peut être reconstruit en éliminant successivement toutes les variables abstraites communes. Ainsi, à chaque fois qu'une variable apparaît dans une même colonne V_r de $V_k^{(rv)}$, il est facile de trouver un vecteur ϕ_r tel que $\phi_r V_r = 0$. La cohérence est alors constatée lorsque $\{0\} \in \phi_r (V_k^{(rv)} v^{(rv)} + V_{0,k})$, c'est-à-dire, en utilisant (7), lorsque :

$$-\|\phi_r V_k^{(rv)}\|_1 \leq |\phi_r V_{k,0}| \leq \|\phi_r V_k^{(rv)}\|_1 \quad (8)$$

Si, après avoir examiné toutes les variables communes, aucune incohérence n'a été constatée,

la cohérence sera manifeste.

Résultats des tests de cohérence

En appliquant l'algorithme précédent aux différentes observations (cf. figure III, V et VII) de la machine à courant continu, nous avons relevé que la procédure programmée avec Matlab sur un Pentium 133MHz, prenait environ 15s. pour 5s. de fonctionnement simulé de la machine à courant continu. Il existe un rapport de trois. Le temps réel semble être possible par programmation directe, en langage C par exemple ; notons que l'algorithme peut être parallélisé.

Pour les différents défauts, nous avons représenté les résultats de l'algorithme que nous venons de présenter et, au dessous, nous avons fait apparaître le résidu qui aurait été obtenu, en négligeant les incertitudes, par un hypomodèle certain à horizon fini minimal.

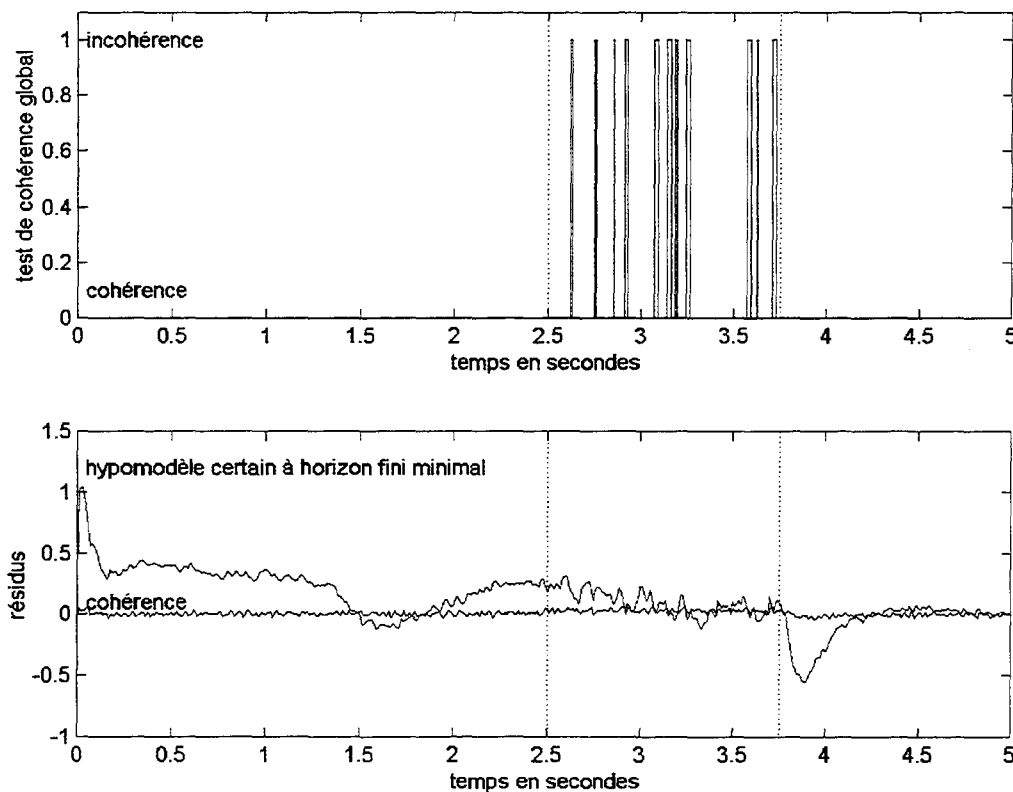


Figure X – Résultats du test global lors d'une multiplication parasite du frottement par 5

Les traits verticaux en pointillés représentent les instants d'apparition et de disparition du défaut. En observant les résultats du test de cohérence global lorsque la valeur de la résistance chute (partie supérieure de la figure X), nous constatons que les incohérences détectées apparaissent uniquement lorsque le défaut est présent. Ceci est dû à l'horizon court ($h=3$) de l'hypomodèle utilisé : une incohérence détectée à un instant t_k donnée ne peut être due qu'à une anomalie survenue entre les instants $t_k-30\text{ms}$ et t_k . Des incohérences ne sont pas détectées en permanence pendant que le défaut est présent pour deux raisons. La première est théorique

et inéluctable ; même si un défaut est présent, le comportement peut s'apparenter à un comportement normal. Ce phénomène est d'autant plus marqué que les incertitudes sont importantes : il existera d'autant plus d'observations qui s'apparenteront à un bon fonctionnement. La seconde raison tient à la dégradation de la précision de l'hypomodèle en raison des approximations effectuées : moins l'hypomodèle est précis, c'est-à-dire, plus il y a d'approximation, moins d'incohérences seront détectées.

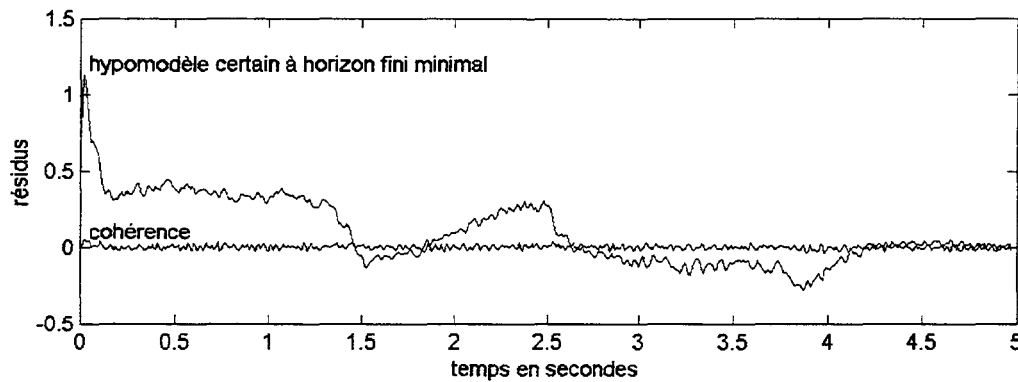


Figure XI – Résidus de l'hypomodèle certain à horizon fini minimal en bon fonctionnement

En comparant les résidus issus de l'hypomodèle certain à horizon fini minimal obtenus en bon fonctionnement (cf. figure XI) et ceux obtenus en présence du défaut (cf. partie inférieure de la figure X), il apparaît très nettement que ces signaux ne permettent pas de révéler le défaut : les résidus obtenus sont même moins importants durant le défaut que lors d'un bon fonctionnement ! Remarquons que, dans un cas comme dans l'autre, les résidus se rapprochent de zéro, autrement dit, ils tendent à révéler une cohérence, aux instants 1,7s. et 5s. Or, la figure II fait apparaître que c'est justement à ces instants que la valeur de la résistance est nominale ($v_{1,k}=0$) alors que la valeur du frottement est presque nominale à 1,7s. En revanche, à l'instant 5s., le frottement est éloigné de sa valeur nominale puisque $v_{2,k}=1$; il semble que les variations du frottement n'influent guère sur ces résidus de l'hypomodèle certain à horizon fini minimal. Ces résidus, mais aussi, puisque l'hypomodèle est générateur de tous les autres hypomodèles certains, tous les résidus issus d'hypomodèle certain ne permettent pas de détecter une augmentation anormale du frottement alors que l'hypomodèle incertain le détecte très efficacement.

Sur la figure XII, nous avons représenté les résidus obtenus lorsque la valeur de la résistance atteint seize fois sa valeur nominale. Ici aussi, durant le défaut, l'hypomodèle incertain révèle des incohérences. Elles sont pourtant peu nombreuses ; ceci montre que le test de cohérence est peu sensible aux variations de résistance. Mais une incohérence suffit à détecter un comportement anormal, même si plusieurs confirment l'anomalie et permettent de mieux déterminer, si cela était recherché, les instants d'apparition et de disparition du défaut. Sur la partie inférieure de la figure XII, les résidus de l'hypomodèle certain sont importants à tel

point qu'ils semblent plus intéressants que les résultats du test de cohérence exploitant la connaissance des incertitudes. Cependant, rien ne permet de garantir que de tels résidus soient nécessairement les symptômes d'un défaut tandis que dans le cas de l'hypomodèle certain, une incohérence est indubitable.

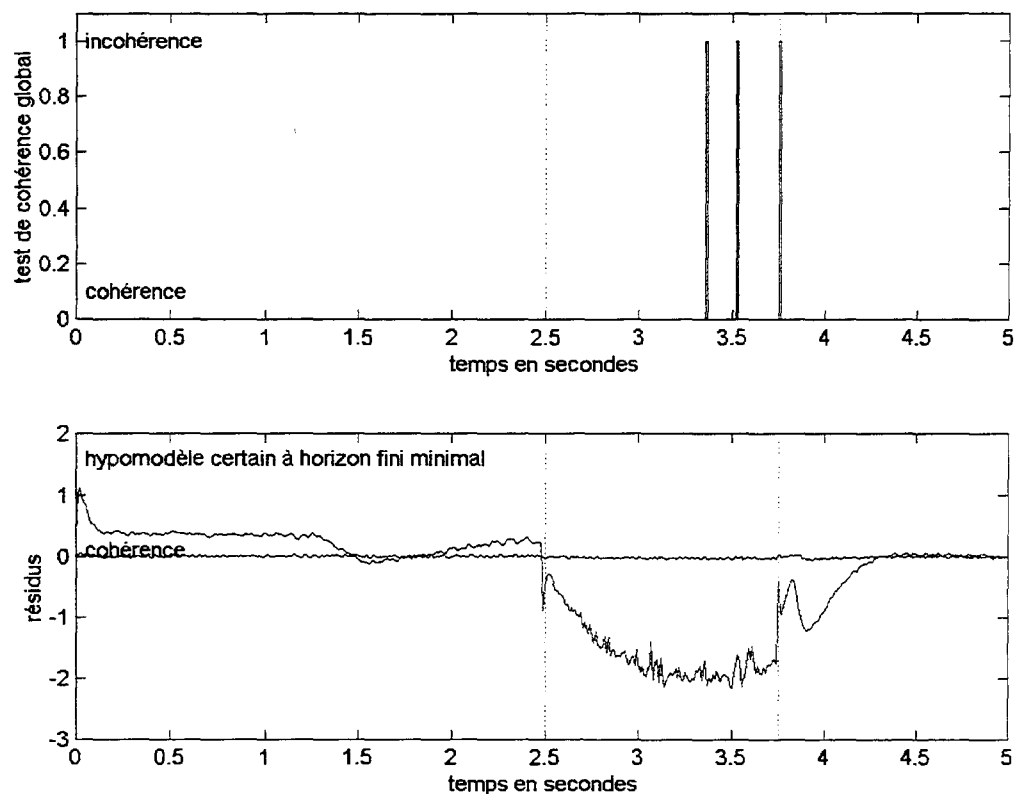


Figure XII – Résultats du test global en présence d'une multiplication moyenne de la résistance par 16

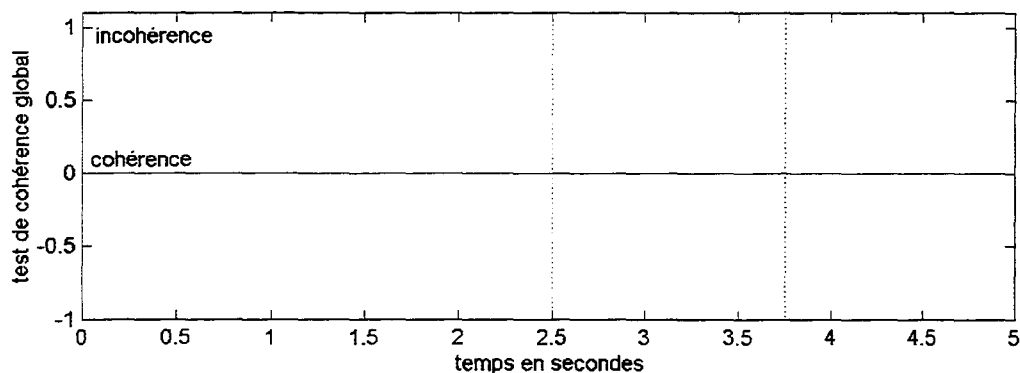


Figure XIII – Résultats du test global en présence d'une multiplication moyenne de la résistance nominale par 7

La figure XIII représente les résultats obtenus lorsque la résistance n'atteint que sept fois sa valeur nominale. La procédure exploitant l'hypomodèle incertain ne détecte aucune incohérence de comportement parce que la modification du comportement est trop peu sensible devant la précision de l'hypomodèle incertain. Le comportement défaillant est assimilé à un comportement normal. Nous reviendrons sur ce problème dans la partie

suivante.

Le test que nous venons de détailler reste cependant très global. Le seul diagnostic pouvant en découler reste très sommaire : la machine fonctionne correctement ou pas. La partie suivante examine des procédures de diagnostics locaux permettant d'affiner le diagnostic.

II.4. Diagnostics locaux

Pour affiner un diagnostic, il faut, non plus ne s'intéresser qu'au comportement global de la machine, mais à certaines fractions de son comportement. Il est bien entendu que plus il y aura de variables d'état mesurées, plus il sera possible d'isoler des comportements locaux. Toutes les variables d'état étant, dans notre exemple, mesurées, les métamodèles réduits (5) et (6) représentent les deux fractions que nous avons pu distinguer, représentant respectivement, le comportement *électrique* de la machine et son comportement *mécanique*. Chacun de ses deux métamodèles ne contient qu'une variable à intégrer ce qui, conformément aux conclusions du chapitre, autorise l'exploitation d'un hypomodèle à horizon infini. Nous avons dorénavant le choix de l'horizon de l'hypomodèle. Lequel est le plus intéressant ?

Nous avons constaté que les hypomodèles à horizon fini étaient, d'une manière générale, plus intéressants que ceux à horizon infini dans la mesure où ils permettent de mieux déterminer les instants où le comportement anormal est survenu. Cependant, le test global permet déjà de déterminer l'instant d'apparition de l'anomalie ; il ne s'agit plus, désormais, que de déterminer quelle est la partie défaillante de la machine, le plus précisément et le plus rapidement possible. Il se trouve que les hypomodèles à horizon infini ne réclament, dans le cas où il n'y a qu'une seule variable à intégrer, que peu de temps de simulation et par ailleurs qu'ils ne reposent que sur une seule approximation, lors du passage de la description en temps continu à la description en temps discret : ils devraient a priori être plus précis.

Algorithmes

En observant les valeurs numériques des matrices des métamodèles (5) et (6), nous déduisons que $A_0^{(i)} + A_1^{(i)} \nu_{1,k} > 0 \forall \nu_{1,k} \in [-1 \ 1]$ et que $A_0^{(\Omega)} + A_2^{(\Omega)} \nu_{2,k} > 0 \forall \nu_{2,k} \in [-1 \ 1]$ ⁸. La formule (13) du chapitre 2 s'applique donc. Nous en déduisons :

$$\mathcal{A}(\mathbb{I}_{k+1}) = \left[\begin{array}{l} \inf_{\nu_k \in \mathcal{A}(\nu_k)} \left((A_0^{(i)} \tilde{i}_k + B_0^{(iu)} \tilde{u}_k + B_0^{(i\Omega)} \tilde{\Omega}_k) + (A_1^{(i)} \tilde{i}_k + B_1^{(iu)} \tilde{u}_k) \nu_{1,k} + B_4^{(i\Omega)} \tilde{\Omega}_k \nu_{4,k} + E_6^{(i)} \nu_{6,k} \right) \\ \sup_{\nu_k \in \mathcal{A}(\nu_k)} \left((A_0^{(i)} \hat{i}_k + B_0^{(iu)} \tilde{u}_k + B_0^{(i\Omega)} \tilde{\Omega}_k) + (A_1^{(i)} \hat{i}_k + B_1^{(iu)} \tilde{u}_k) \nu_{1,k} + B_4^{(i\Omega)} \tilde{\Omega}_k \nu_{4,k} + E_6^{(i)} \nu_{6,k} \right) \end{array} \right]$$

⁸ Il est rare qu'en pratique, ceci ne soit pas vérifié car cela reviendrait à considérer une incertitude modifiant complètement la nature de la réponse !

$$\mathcal{A}([\Omega_{k+1}]) = \left[\begin{array}{l} \inf_{\mathfrak{v}_k \in \mathcal{A}(\mathfrak{v}_k)} \left((A_0^{(\Omega)} \tilde{\Omega}_k + B_0^{(\Omega)} \tilde{i}_k) + (A_2^{(\Omega)} \tilde{\Omega}_k + B_2^{(\Omega)} \tilde{i}_k) \mathfrak{v}_{2,k} + B_3^{(\Omega)} \tilde{i} \mathfrak{v}_{3,k} + E_5^{(\Omega)} \mathfrak{v}_{5,k} \right) \\ \sup_{\mathfrak{v}_k \in \mathcal{A}(\mathfrak{v}_k)} \left((A_0^{(\Omega)} \hat{\Omega}_k + B_0^{(\Omega)} \tilde{i}_k) + (A_1^{(i)} \hat{\Omega}_k + B_1^{(iu)} \tilde{u}_k) \mathfrak{v}_{2,k} + B_3^{(\Omega)} \tilde{i} \mathfrak{v}_{3,k} + E_5^{(\Omega)} \mathfrak{v}_{5,k} \right) \end{array} \right]$$

avec $\mathcal{A}[\mathfrak{i}_k] = [\tilde{i}_k, \hat{i}_k]$ et $\mathcal{A}[\Omega_k] = [\tilde{\Omega}_k, \hat{\Omega}_k]$

Puis, en appliquant la formule (7), nous obtenons :

$$\begin{aligned} \tilde{i}_{k+1} &= A_0^{(i)} \tilde{i}_k + B_0^{(iu)} \tilde{u}_k + B_0^{(\Omega)} \tilde{\Omega}_k - \left\| \begin{array}{l} A_1^{(i)} \tilde{i}_k + B_1^{(iu)} \tilde{u}_k \\ B_4^{(\Omega)} \tilde{\Omega}_k \\ E_6^{(i)} \end{array} \right\|_1 \\ \hat{i}_{k+1} &= A_0^{(i)} \hat{i}_k + B_0^{(iu)} \tilde{u}_k + B_0^{(\Omega)} \tilde{\Omega}_k + \left\| \begin{array}{l} A_1^{(i)} \tilde{i}_k + B_1^{(iu)} \tilde{u}_k \\ B_4^{(\Omega)} \tilde{\Omega}_k \\ E_6^{(i)} \end{array} \right\|_1 \end{aligned} \quad (9)$$

et

$$\begin{aligned} \tilde{\Omega}_{k+1} &= A_0^{(\Omega)} \tilde{\Omega}_k + B_0^{(\Omega)} \tilde{i} - \left\| \begin{array}{l} A_2^{(\Omega)} \tilde{\Omega}_k + B_2^{(\Omega)} \tilde{i}_k \\ B_3^{(\Omega)} \tilde{i}_k \\ E_5^{(\Omega)} \end{array} \right\|_1 \\ \hat{\Omega}_{k+1} &= A_0^{(\Omega)} \hat{\Omega}_k + B_0^{(\Omega)} \tilde{i} + \left\| \begin{array}{l} A_2^{(\Omega)} \tilde{\Omega}_k + B_2^{(\Omega)} \tilde{i}_k \\ B_3^{(\Omega)} \tilde{i}_k \\ E_5^{(\Omega)} \end{array} \right\|_1 \end{aligned} \quad (10)$$

Les bornes des espaces abstraits peuvent alors être déterminées de manière récurrente. Il nous reste toutefois à tenir compte des équations de sortie des métamodèles (5) et (6). Pour des raisons de simplicité, nous aurons recours à une approche directe (cf. chapitre 4 partie III.2 : *Prise en compte de l'équation de sortie*). Comme précédemment, puisque $C_0^{(i)} + C_3^{(i)} \mathfrak{v}_{3,k} > 0 \forall \mathfrak{v}_{3,k} \in [-1, 1]$ et $C_0^{(\Omega)} + C_4^{(\Omega)} \mathfrak{v}_{4,k} > 0 \forall \mathfrak{v}_{4,k} \in [-1, 1]$, les formules (13) du chapitre (3) et (7) s'appliquent et conduisent aux tests de cohérences suivants :

$$\{\tilde{i}_k\} \in \left[C_0^{(i)} \tilde{i}_k + D_0^{(\Omega)} \tilde{\Omega}_k - \left\| \begin{array}{l} D_1^{(\Omega)} \tilde{\Omega}_k \\ C_3^{(i)} \tilde{i}_k + D_3^{(\Omega)} \tilde{\Omega}_k \\ D_4^{(\Omega)} \tilde{\Omega}_k \\ F_5^{(i)} \end{array} \right\|_1 - \lambda_i, C_0^{(i)} \hat{i}_k + D_0^{(\Omega)} \tilde{\Omega}_k + \left\| \begin{array}{l} D_1^{(\Omega)} \tilde{\Omega}_k \\ C_3^{(i)} \tilde{i}_k + D_3^{(\Omega)} \tilde{\Omega}_k \\ D_4^{(\Omega)} \tilde{\Omega}_k \\ F_5^{(i)} \end{array} \right\|_1 + \lambda_i \right] \quad (11)$$

où λ_i est un nombre réel positif

$$\{\tilde{\Omega}_k\} \in \left[C_0^{(\Omega)} \tilde{\Omega}_k + D_0^{(\Omega)} \tilde{i}_k - \left\| \begin{array}{l} D_2^{(\Omega)} \tilde{i}_k \\ D_3^{(\Omega)} \tilde{i}_k \\ C_4^{(\Omega)} \tilde{\Omega}_k + D_4^{(\Omega)} \tilde{i}_k \\ F_6^{(\Omega)} \end{array} \right\|_1 - \lambda_\Omega, C_0^{(\Omega)} \hat{\Omega}_k + D_0^{(\Omega)} \tilde{i}_k + \left\| \begin{array}{l} D_2^{(\Omega)} \tilde{i}_k \\ D_3^{(\Omega)} \tilde{i}_k \\ C_4^{(\Omega)} \tilde{\Omega}_k + D_4^{(\Omega)} \tilde{i}_k \\ F_6^{(\Omega)} \end{array} \right\|_1 + \lambda_\Omega \right] \quad (12)$$

où λ_Ω est un nombre réel positif

Les paramètres λ_i et λ_Ω permettent d'augmenter la surface de l'enveloppe de telle sorte à

compenser les approximations effectuées lors du passage au métamodèle en temps discret. Ces paramètres sont réglés, à l'instar de l'hypomodèle incertain à horizon fini, de façon à ce qu'ils soient les plus petits possible et que les mesures en bon fonctionnement ne provoquent aucune incohérence. Nous avons obtenu les valeurs : $\lambda_r=1$ et $\lambda_\Omega=0,3$.

Résultats des tests de cohérence

Sur un ordinateur Pentium 133MHz, les résultats des deux tests partiels sont obtenus après environ 3,5s. de calcul pour 5s. de simulation. Ces tests sont effectivement plus rapides que ceux exploitant des hypomodèles à horizon fini.

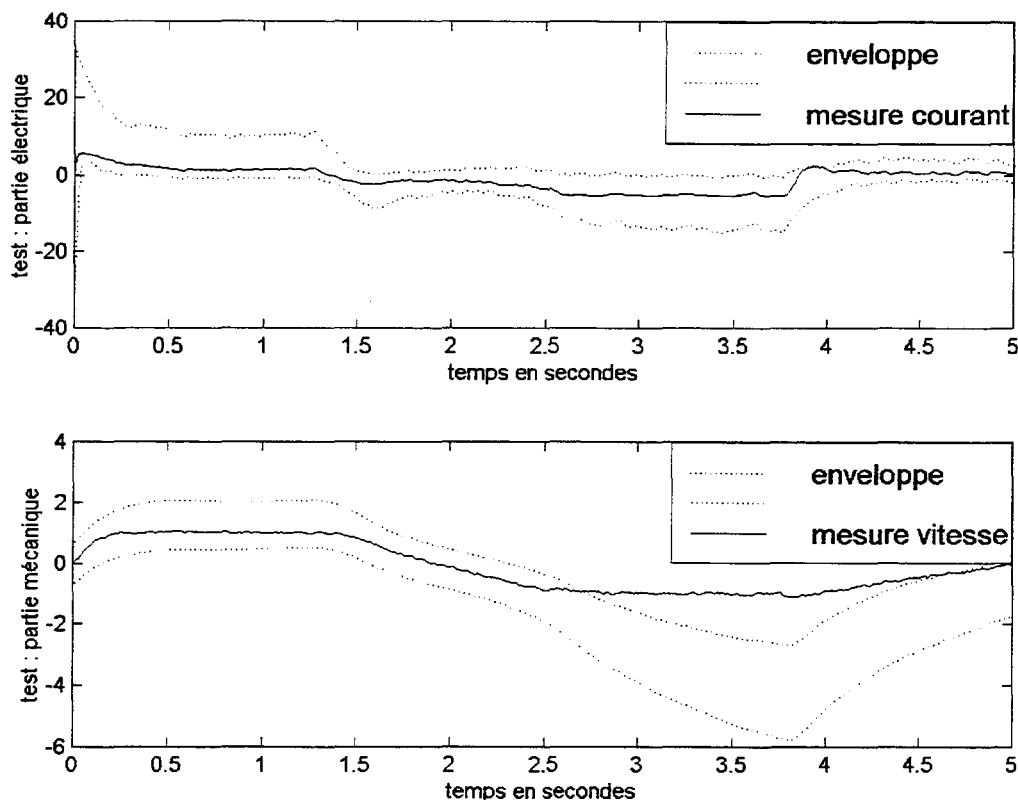


Figure XIV – Résultats obtenus lors d'une multiplication parasite du frottement par 5

La figure XIV représente les enveloppes et les mesures correspondant, pour la moitié supérieure, au test (11), et, pour la moitié inférieure, à (12). Conformément aux tests, dès que les mesures sortent de l'enveloppe, une incohérence est révélée. Des incohérences apparaissent sur la partie mécanique, correspondant au métamodèle (6), alors qu'aucunes n'apparaissent au niveau de la partie électrique (5). Les résultats du test global, ainsi que les résultats de ces procédures de diagnostics locaux interprétés en terme de cohérence et d'incohérence (à l'intérieur et à l'extérieur de l'enveloppe) sont synthétisés par la figure XV.

Le résultat du premier test de cohérence nous permet de conclure qu'un défaut était présent sur la machine entre les instants 2,55s. et 3,5s. Les tests de cohérence locaux nous permettent d'affiner le diagnostic en montrant que le défaut n'a affecté que la partie mécanique de la

machine. Nous en concluons qu'un défaut s'est produit uniquement sur la partie mécanique de la machine entre 2,55s. et 3,5s. ; mais cela n'exclut pas que le défaut ait pu être présent plus longtemps. Malgré les apparences, le test portant sur la partie mécanique ne permet pas de conclure quant aux instants d'apparition et de disparition du défaut. L'incohérence apparaît encore à l'instant 4,7s. Ceci s'explique par le fait que le pôle du métamodèle (6) est important (0,99) ; il faut environ 300 échantillons (3s.) pour que le passé n'influence plus le résultat.

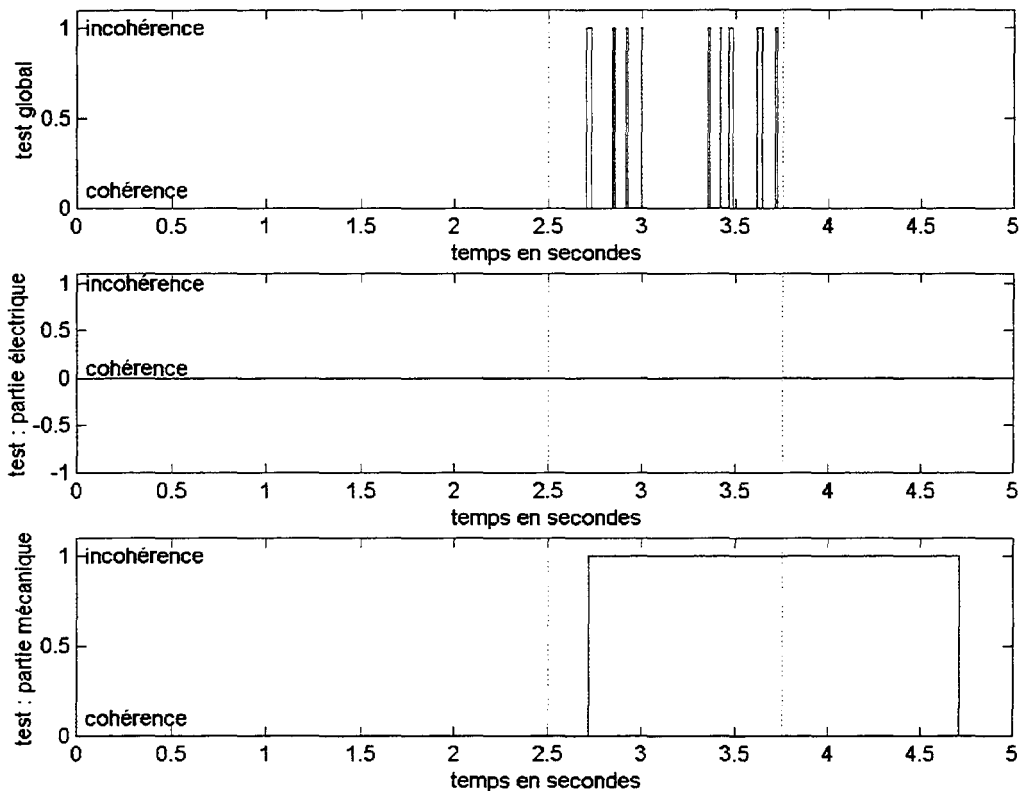


Figure XV – Résultats de la procédure orientée bon fonctionnement lors d'une multiplication parasite du frottement par 5

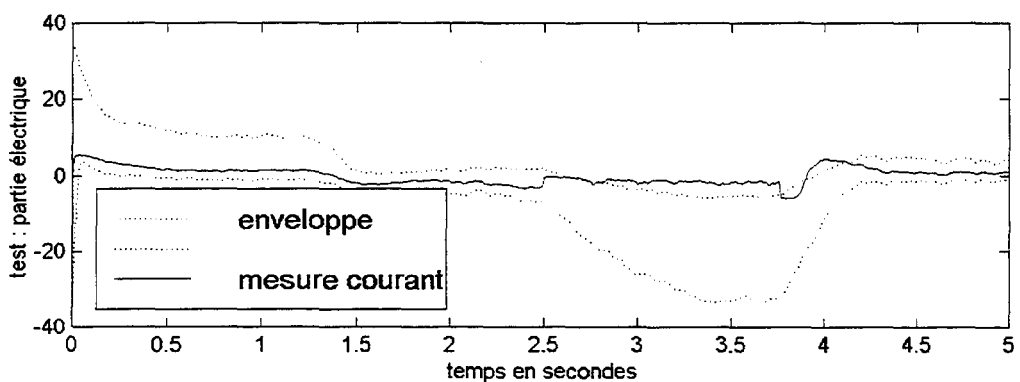


Figure XVI - Résultats obtenus lors d'une multiplication par 16 de la résistance

Considérons maintenant le défaut de résistance : augmentation à seize fois sa valeur nominale. En appliquant les procédures de diagnostics locaux, nous n'observons aucune incohérence sur la partie mécanique tandis que l'enveloppe et la mesure représentées sur la figure XVI

révèlent des incohérences.

Les résultats des tests sont synthétisés par la figure XVII. Ils indiquent que quelques incohérences sont survenues durant le temps où le défaut était effectivement présent, et que seule la partie électrique était affectée. Le pôle⁹ du métamodèle (5) étant faible (0,553), l'influence du passé devient négligeable après seulement 5 itérations (50ms.). Ceci explique que, dans ce cas, les instants d'apparition et de disparition détectés coïncident assez bien avec les véritables instants. Une zone d'incohérence apparaît au niveau de la partie électrique autour de l'instant 4s. Ce phénomène est dû à la chute brutale de la valeur de la résistance à 3,75s. qui engendre un artefact. Le manque de précision du métamodèle se manifeste dans de telles situations extrêmes mais, ici, ne gêne en rien le diagnostic.

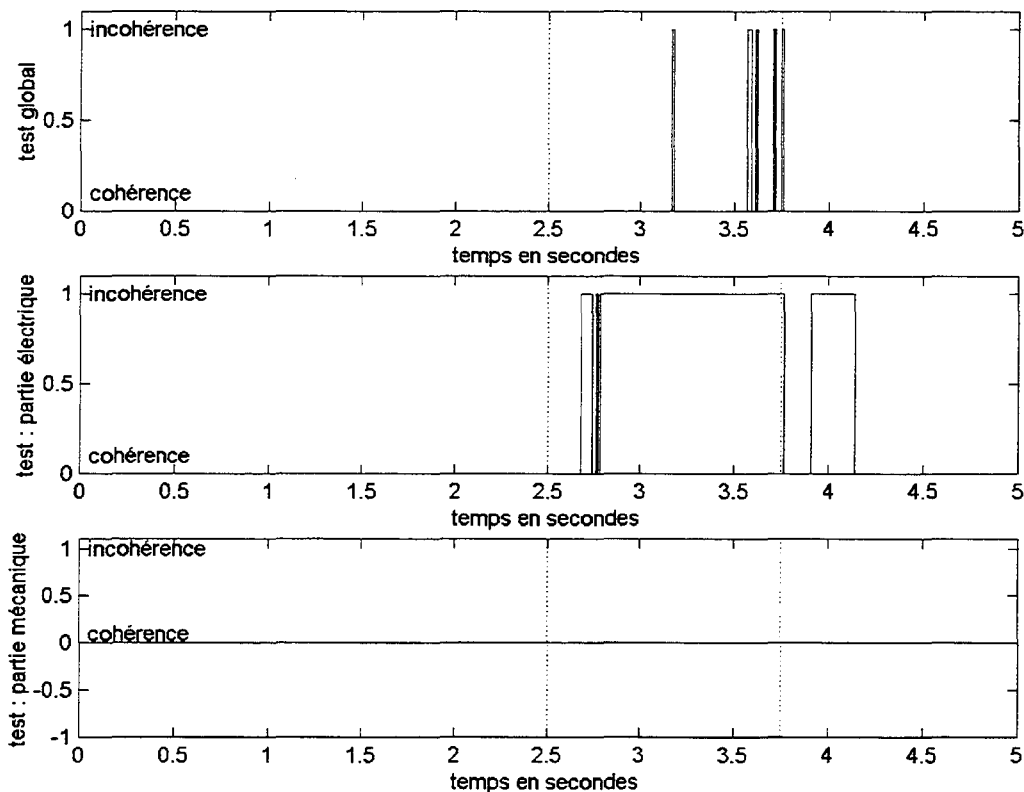


Figure XVII – Résultats de la procédure lors d'une multiplication par 16 de la résistance

La figure XVIII fournit les résultats des tests de cohérence, obtenus lorsque la résistance n'augmente que jusqu'à sept fois sa valeur nominale. Alors qu'il y a un défaut, le test de cohérence global ne révèle aucune incohérence tandis que les tests locaux conduisent encore à des résultats intéressants. Ce cas montre que l'hypomodèle global à horizon fini est moins précis que les hypomodèles locaux (5) et (6). En effet, deux approximations avaient été nécessaires pour l'hypomodèle à horizon fini alors qu'une seule avait suffi dans le cas d'un horizon infini.

⁹ Incertitudes négligées

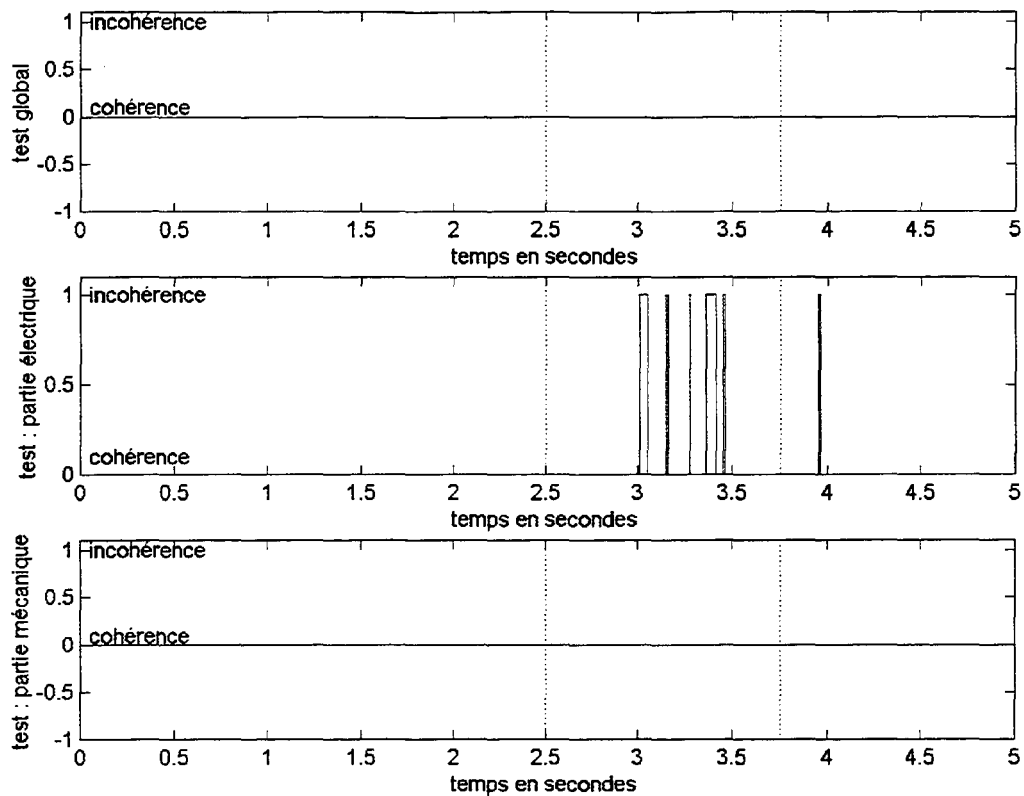


Figure XVIII - Résultats de la procédure lors d'une multiplication par 7 de la résistance

La figure XIX présente les résultats obtenus lorsque deux défauts, l'un sur la valeur du frottement et l'autre, sur la valeur de la résistance sont présents simultanément : accroissement du frottement à 5 fois sa valeur nominale et de la résistance à 7 fois sa valeur nominale. Les résultats obtenus sont une bonne illustration du fait que les procédures que nous développons s'appliquent autant à des défauts simples qu'à des défauts multiples : l'anomalie apparaît au niveau du test global, au niveau de la partie électrique et au niveau de la partie mécanique. Nous pouvons conclure que les parties mécanique et électrique de la machine sont affectées. L'incohérence globale révèle la présence du défaut durant la période 2,7 à 3,9s. La période d'incohérence est, une fois de plus, plus importante que celle où le défaut est effectivement présent ; ceci s'explique par une situation extrême (les deux défauts cessent brutalement à 3,75s.).

Cette partie nous a permis de montrer que le diagnostic de systèmes physiques ne pouvant se représenter simplement que par des métamodèles incertains, ne pouvait pas être réalisé à partir d'une approche certaine. Nous avons montré que, même si d'un point de vue théorique les tests de cohérence reposant sur des hypomodèles à horizon fini sont les plus intéressants, les hypomodèles à horizon infini conservent, lorsqu'ils existent, un intérêt certain parce qu'ils sont souvent plus précis. Nous avons montré comment plusieurs tests de cohérence pouvaient être réunis en une procédure de diagnostic permettant de détecter plusieurs défauts simultanés.

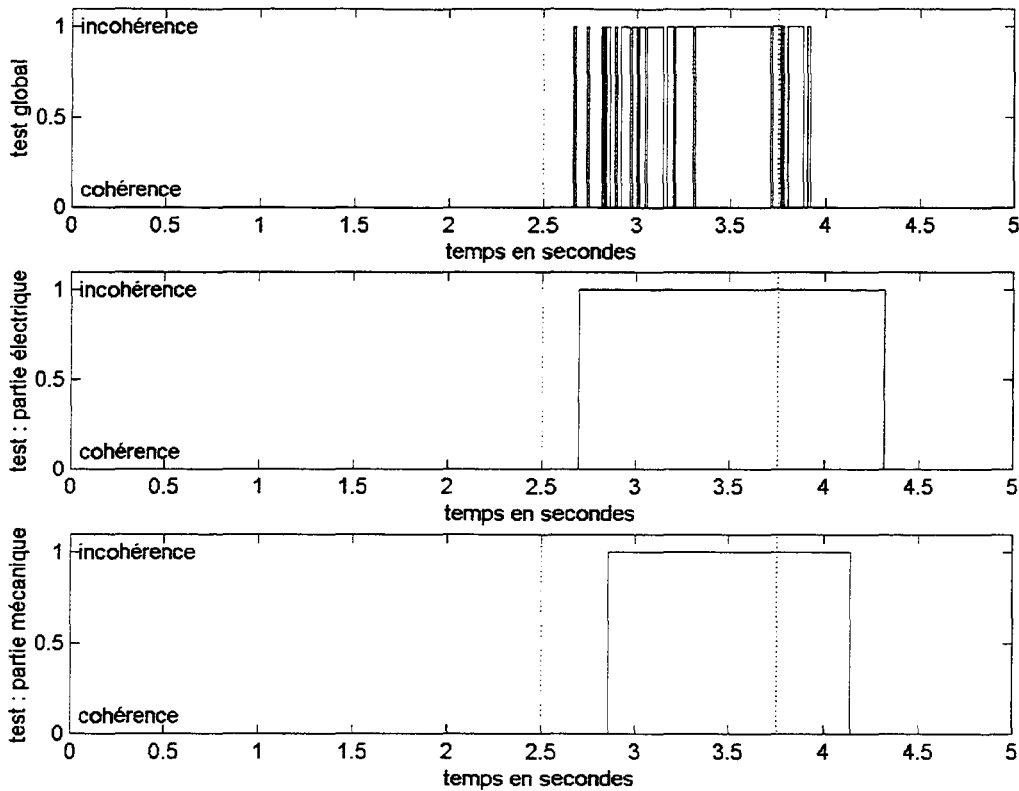


Figure XIX - Résultats de la procédure lors d'une augmentation du frottement à 5 fois sa valeur nominale et de la résistance à 7 fois sa valeur nominale

L'exemple que nous venons de traiter nous a conduits à une procédure orientée bon fonctionnement ; les défauts simulés sont sommaires et pas très réalistes. La partie suivante est consacrée aux procédures orientées mauvais fonctionnement qui viennent en complément des précédentes.

III. Procédure orientée mauvais fonctionnement

Les procédures orientées mauvais fonctionnement permettent de compléter un diagnostic issu d'une procédure de bon fonctionnement. Mais, la condition sine qua non d'existence d'une telle procédure est qu'il existe un modèle de mauvais fonctionnement. Dans cette partie, nous avons continué avec l'exemple de la machine à courant continu pour laquelle nous avons représenté sommairement un court-circuit entre des spires du rotor. Nous examinerons, à partir de cet exemple, comment une procédure orientée mauvais fonctionnement peut être réalisée et nous proposerons une procédure automatique.

III.1. Métamodèle de la machine à courant continu avec un court circuit

Métamodèle complet

Supposons qu'un court-circuit survienne entre plusieurs spires du rotor de la machine à courant continu. En notant $\sigma \in [0,1[$, le ratio de spires encore intacte, nous déduisons que,

sommairement, en négligeant les spires court-circuitées, la machine de s spires s'apparente à une machine de σs spires. La résistance du rotor devient alors σR , le coefficient de self induction au stator devient, en première approximation, $\sigma^2 L$ et le coefficient ϕ devient $\sigma \phi_n^{10}$. En revenant au métamodèle (I), nous déduisons qu'en cas de court-circuit entre spires, le métamodèle de la machine devient :

$$\begin{cases} \frac{d[i(t)]}{dt} = -\frac{R}{\sigma L} [i(t)] - \frac{\Phi_n^2}{\sigma L f_n} [\Omega(t)] + \frac{R_n f_n + \Phi_n^2}{\sigma^2 L f_n} [u(t)] \\ \frac{d[\Omega(t)]}{dt} = \frac{\sigma f_n}{J} [i(t)] - \frac{f}{J} [\Omega(t)] \end{cases}$$

En introduisant les incertitudes $R=(1+\rho_1 v_1(t))R_n$ et $f=(1+\rho_2 v_2(t))f_n$, ainsi que les paramètres τ_1 , τ_2 et γ , nous obtenons l'écriture suivante :

$$\begin{cases} \frac{d[i(t)]}{dt} = -\frac{1+\rho_1 v_1(t)}{\sigma \tau_1} [i(t)] + \frac{\gamma-1}{\sigma \gamma \tau_1} [\Omega(t)] + \frac{1}{\sigma^2 \gamma \tau_1} [u(t)] \\ \frac{d[\Omega(t)]}{dt} = \frac{\sigma}{\tau_2} [i(t)] - \frac{1+\rho_2 v_2(t)}{\tau_2} [\Omega(t)] \end{cases} \quad (13)$$

Ce métamodèle représente un court-circuit de $(1-\sigma)s$ spires. Cependant, un modèle de mauvais fonctionnement ne peut représenter un nombre déterminé de spires ; il faudrait un nombre infini de modèles pour représenter l'ensemble des situations possibles. Il est indispensable d'associer à un métamodèle, un ensemble de spires intactes possibles : $(1+\rho_7 v_7(t))\sigma s$ avec $\|v_7(t)\|_\infty \leq 1$. Dès lors, le nombre de spires intactes appartient à $[(1-\rho_7)\sigma s, (1+\rho_7)\sigma s]$. En introduisant cette nouvelle incertitude dans le métamodèle (13), nous parvenons à un métamodèle de mauvais fonctionnement en temps continu :

$$\begin{cases} \frac{d}{dt} \begin{bmatrix} [i(t)] \\ [\Omega(t)] \end{bmatrix} = A_c(v(t)) \begin{bmatrix} [i(t)] \\ [\Omega(t)] \end{bmatrix} + B_c \tilde{u}(t) \\ \begin{bmatrix} [\tilde{i}(t)] \\ [\tilde{\Omega}(t)] \end{bmatrix} = C_c(v(t)) \begin{bmatrix} [i(t)] \\ [\Omega(t)] \end{bmatrix} + F_c(v(t)) \end{cases} \quad (14a)$$

avec :

¹⁰ En effet, $\Phi = p n \phi_p / a$ avec n le nombre de spires, p le nombre de paires de pôles, ϕ_p le flux d'excitation par pôle et a le nombre de voie en parallèle.

$$A_c(v(t)) = \begin{bmatrix} \frac{-(1+\rho_1 v_1(t))}{\tau_2} & \frac{(\gamma-1)}{\tau_2} \\ \frac{(1+\rho_7 v_7(t))\sigma\tau_1}{(1+\rho_7 v_7(t))\sigma} & \frac{(1+\rho_7 v_7(t))\sigma\gamma\tau_1}{-(1+\rho_2 v_2(t))\rho\mu} \end{bmatrix}, \quad B_c(v(t)) = \begin{bmatrix} \frac{1}{(1+\rho_7 v_7(t))^2 \sigma^2 \gamma \tau_1} \\ 0 \end{bmatrix} \quad (14b)$$

$$C_c(v(t)) = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} + \begin{bmatrix} \rho_3 & 0 \\ 0 & 0 \end{bmatrix} v_3(t) + \begin{bmatrix} 0 & 0 \\ 0 & \rho_4 \end{bmatrix} v_4(t), \quad F_c(v(t)) = \begin{bmatrix} \rho_5 \\ 0 \end{bmatrix} v_5(t) + \begin{bmatrix} 0 \\ \rho_6 \end{bmatrix} v_6(t)$$

Ce modèle peut être ramené, par linéarisation, à un métamodèle affine en les variables abstraites ; les matrices $A_c(v(t))$ et $B_c(v(t))$ de (14b) deviennent :

$$A'_c(v(t)) = \begin{bmatrix} \frac{-1}{\tau_2} & \frac{(\gamma-1)}{\tau_2} \\ \frac{\sigma\tau_1}{\sigma} & \frac{\sigma\gamma\tau_1}{-1} \end{bmatrix} + \begin{bmatrix} -\rho_1 & 0 \\ \sigma\tau_1 & 0 \end{bmatrix} v_1(t) + \begin{bmatrix} 0 & 0 \\ 0 & -\frac{\rho_2}{\tau_2} \end{bmatrix} v_2(t) + \begin{bmatrix} \frac{\rho_7}{\tau_2} & \frac{-\rho_7(\gamma-1)}{\sigma\gamma\tau_1} \\ \frac{\sigma\tau_1}{\rho_7\sigma} & 0 \end{bmatrix} v_7(t) \quad (15)$$

$$B'_c(v(t)) = \begin{bmatrix} \frac{1}{(1+\rho_7 v_7(t))^2 \sigma^2 \gamma \tau_1} \\ 0 \end{bmatrix} + \begin{bmatrix} \frac{-2\rho_7}{\sigma^2 \gamma \tau_1} \\ 0 \end{bmatrix} v_7(t)$$

Un métamodèle discret approché s'obtient sans difficulté en utilisant les relations (46) du chapitre 4 :

$$\begin{cases} \begin{bmatrix} [i_{k+1}] \\ [\Omega_{k+1}] \end{bmatrix} = (A_0 + A_1 v_{1,k} + A_2 v_{2,k} + A_7 v_{7,k}) \begin{bmatrix} [i_k] \\ [\Omega_k] \end{bmatrix} + (B_0 + B_1 v_{1,k} + B_2 v_{2,k} + B_7 v_{7,k}) \tilde{u}_k \\ \begin{bmatrix} [\tilde{i}_k] \\ [\tilde{\Omega}_k] \end{bmatrix} = (C_0 + C_3 v_{3,k} + C_4 v_{4,k}) \begin{bmatrix} [i_k] \\ [\Omega_k] \end{bmatrix} + F_5 v_{5,k} + F_6 v_{6,k} \end{cases} \quad (16)$$

Lorsque $\sigma=0,35$ (court-circuit de 65% des spires), ces matrices ont pour valeur :

$$A_0 \cong \begin{bmatrix} 0.2211 & -5.6811 \\ 2.110^{-3} & 0.976 \end{bmatrix} \quad A_1 \cong \begin{bmatrix} -0.1653 & 1.5961 \\ -6.10^{-4} & 2.7.10^{-3} \end{bmatrix} \quad A_2 \cong \begin{bmatrix} 4.10^{-6} & 4.10^{-3} \\ -10^{-6} & -10^{-3} \end{bmatrix}$$

$$A_7 \cong \rho_7 \begin{bmatrix} 0.3305 & 2.489 \\ 3.2.10^{-3} & -5.4.10^{-3} \end{bmatrix} \quad B_0 \cong \begin{bmatrix} 16.203 \\ 4.01.10^{-2} \end{bmatrix} \quad B_1 \cong \begin{bmatrix} -4.5431 \\ -7.7.10^{-3} \end{bmatrix} \quad B_2 \cong \begin{bmatrix} 3.10^{-5} \\ -10^{-5} \end{bmatrix}$$

$$B_7 \cong \rho_7 \begin{bmatrix} -23.32 \\ -2.47.10^{-2} \end{bmatrix} \quad C_0 \cong \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \quad C_3 \cong \begin{bmatrix} 0.1 & 0 \\ 0 & 0 \end{bmatrix} \quad C_4 \cong \begin{bmatrix} 0 & 0 \\ 0 & 0.1 \end{bmatrix} \quad F_5 \cong \begin{bmatrix} 0.05 \\ 0 \end{bmatrix} \quad F_6 \cong \begin{bmatrix} 0 \\ 0.05 \end{bmatrix}$$

Métamodèles réduits

Les métamodèles réduits permettent de construire des hypomodèles incertains à horizon infini présentant l'avantage de requérir moins de temps pour un test de cohérence, mais aussi d'être plus précis que les hypomodèles à horizon fini. Nous examinerons si l'usage de métamodèles réduits contenant, ensemble, toute l'information d'un métamodèle complet n'est pas plus intéressant.

Nous avons signalé dans le paragraphe consacré aux métamodèles réduits (partie II.1) que la réduction pouvait être opérée soit à partir du métamodèle complet en temps continu, soit à partir de celui en temps discret. Dans la partie consacrée aux procédures orientées bon fonctionnement, nous avons préféré partir du métamodèle en temps continu pour que le découplage corresponde aux équations différentielles du métamodèle original. Cependant, pour une approche orientée mauvais fonctionnement, l'objectif n'est plus d'isoler certaines fractions du comportement, mais uniquement de découper le métamodèle complet en métamodèle réduit ne contenant qu'une équation récurrente. Par conséquent, il est plus simple de partir du métamodèle en temps discret pour déduire des métamodèles partiels.

En extrayant chacune des équations récurrentes du métamodèle complet en temps discret, et en notant $(M)_{ij}$ l'élément apparaissant sur la ligne i et la colonne j de la matrice M , nous déduisons les métamodèles partiels suivants :

$$\left\{ \begin{array}{l} [i_{k+1}] = ((A_0)_{11} + (A_1)_{11} v_{1,k} + (A_2)_{11} v_{2,k} + (A_7)_{11} v_{7,k}) [i_k] \cdots \\ \quad \cdots + \frac{((A_0)_{12} + (A_1)_{12} v_{1,k} + (A_2)_{12} v_{2,k} + (A_7)_{12} v_{7,k})}{1 + \rho_4 v_{4,k}} \tilde{\Omega}_k \cdots \\ \quad \cdots + ((B_0)_{11} + (B_1)_{11} v_{1,k} + (B_2)_{11} v_{2,k} + (B_7)_{11} v_{7,k}) \tilde{u}_k \cdots \\ \quad \cdots - \frac{((A_0)_{12} + (A_1)_{12} v_{1,k} + (A_2)_{12} v_{2,k} + (A_7)_{12} v_{7,k})}{1 + \rho_4 v_{4,k}} \rho_6 v_{6,k} \\ \tilde{i}_k = (1 + \rho_3 v_{3,k}) [i_k] + \rho_5 v_{5,k} \end{array} \right.$$

et

$$\left\{ \begin{array}{l} [\Omega_{k+1}] = ((A_0)_{22} + (A_1)_{22} v_{1,k} + (A_2)_{22} v_{2,k} + (A_7)_{22} v_{7,k}) [\Omega_k] \cdots \\ \quad \cdots + \frac{((A_0)_{21} + (A_1)_{21} v_{1,k} + (A_2)_{21} v_{2,k} + (A_7)_{21} v_{7,k})}{1 + \rho_3 v_{3,k}} \tilde{i}_k \cdots \\ \quad \cdots + ((B_0)_{21} + (B_1)_{21} v_{1,k} + (B_2)_{21} v_{2,k} + (B_7)_{21} v_{7,k}) \tilde{u}_k \cdots \\ \quad \cdots - \frac{((A_0)_{12} + (A_1)_{12} v_{1,k} + (A_2)_{12} v_{2,k} + (A_7)_{12} v_{7,k})}{1 + \rho_3 v_{3,k}} \rho_5 v_{5,k} \\ \tilde{\Omega}_k = (1 + \rho_4 v_{4,k}) [\Omega_k] + \rho_6 v_{6,k} \end{array} \right.$$

Pour rester affines par rapport aux variables abstraites, ces métamodèles doivent être linéarisés :

$$\left\{ \begin{array}{l} [i_{k+1}] = ((A_0)_{11} + (A_1)_{11} v_{1,k} + (A_2)_{11} v_{2,k} + (A_7)_{11} v_{7,k}) [i_k] \cdots \\ \quad \cdots + ((A_0)_{12} + (A_1)_{12} v_{1,k} + (A_2)_{12} v_{2,k} + (A_7)_{12} v_{7,k} - \rho_4 (A_0)_{12} v_{4,k}) \tilde{\Omega}_k + \cdots \\ \quad \cdots + ((B_0)_{11} + (B_1)_{11} v_{1,k} + (B_2)_{11} v_{2,k} + (B_7)_{11} v_{7,k}) \tilde{u}_k - \rho_6 (A_0)_{12} v_{6,k} \\ \tilde{i}_k = (1 + \rho_3 v_{3,k}) [i_k] + \rho_5 v_{5,k} \end{array} \right. \quad (17)$$

et

$$\left\{ \begin{array}{l} [\Omega_{k+1}] = ((A_0)_{22} + (A_1)_{22} v_{1,k} + (A_2)_{22} v_{2,k} + (A_7)_{22} v_{7,k}) [\Omega_k] \cdots \\ \quad \cdots + ((A_0)_{21} + (A_1)_{21} v_{1,k} + (A_2)_{21} v_{2,k} + (A_7)_{21} v_{7,k} - \rho_3 (A_0)_{21} v_{3,k}) \tilde{i}_k + \cdots \\ \quad \cdots + ((B_0)_{21} + (B_1)_{21} v_{1,k} + (B_2)_{21} v_{2,k} + (B_7)_{21} v_{7,k}) \tilde{u}_k - \rho_5 (A_0)_{12} v_{5,k} \\ \tilde{\Omega}_k = (1 + \rho_4 v_{4,k}) [\Omega_k] + \rho_6 v_{6,k} \end{array} \right. \quad (18)$$

Nous disposons désormais, avec (17) et (18), de deux métamodèles partiels représentant, ensemble, la totalité de la connaissance contenue dans les métamodèles complets.

III.2. Spécificités des tests de cohérence

Limite de la procédure orientée bon fonctionnement

Un court-circuit de 65% des spires du rotor conduit aux réponses représentées sur la figure XX (les incertitudes normales sont conformes à la figure II). En comparant ces résultats à la figure III, nous remarquons tout d'abord que le court-circuit laisse la tension aux bornes de la machine quasiment inchangée. Seul le courant est modifié ; sa valeur moyenne tombe, durant le défaut, de $-i_n$ à $-4i_n$ accompagnée d'une augmentation de la variance du signal. L'effet du défaut est donc visible.

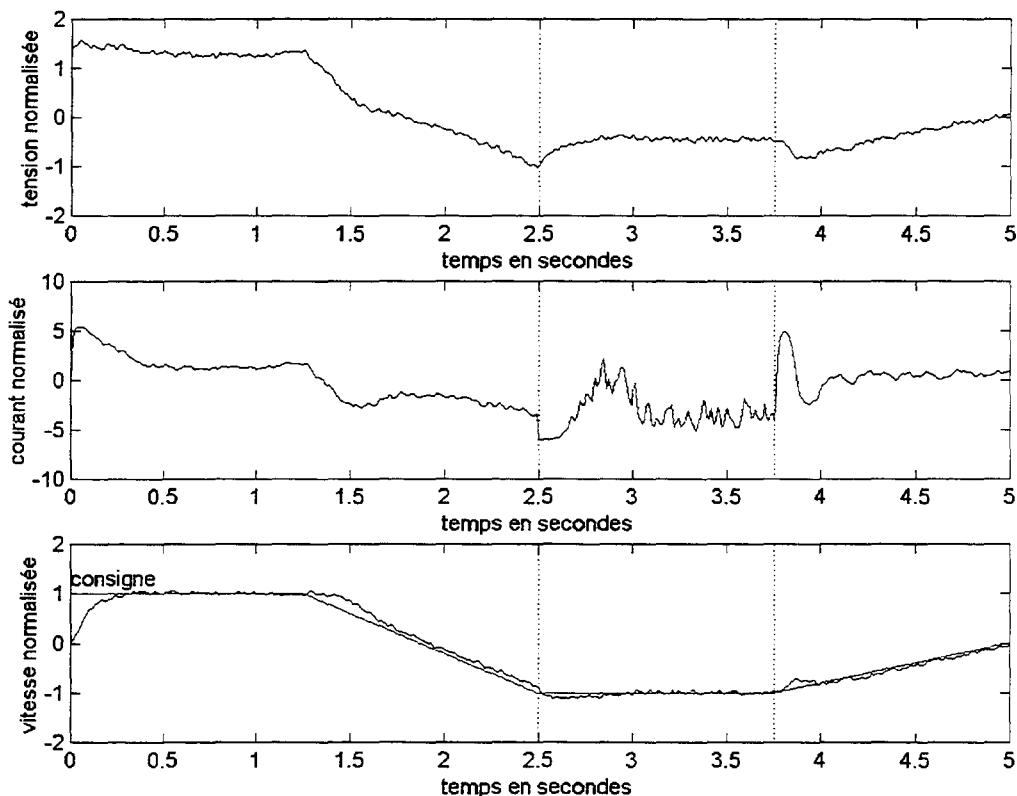


Figure XX – Réponse en présence d'un court-circuit de 65% des spires rotoriques entre 2,5s. et 3,75s.

Nous avons alors appliqué la procédure orientée bon fonctionnement décrite dans la partie

précédente. Cette procédure révèle la présence d'un défaut entre les instants 2,5s. et 3,75s. Les tests de cohérence locaux nous apprennent que le défaut affecte autant la partie électrique que la partie mécanique de la machine (correspondant à la chute du couple électromagnétique). Ces conclusions sont correctes, cependant, les symptômes sont identiques à une augmentation du frottement à cinq fois sa valeur nominale et de la résistance à sept fois sa valeur nominale. En effet, les résultats de la même procédure apparaissant sur la figure XIX sont comparables aux résultats de la figure XXI. Rien ne permet de trancher entre la présence de l'un ou de l'autre défaut ; il n'est plus possible d'examiner plus finement la machine par des tests de cohérence locaux. La seule solution possible est d'avoir recours à une procédure orientée mauvais fonctionnement.

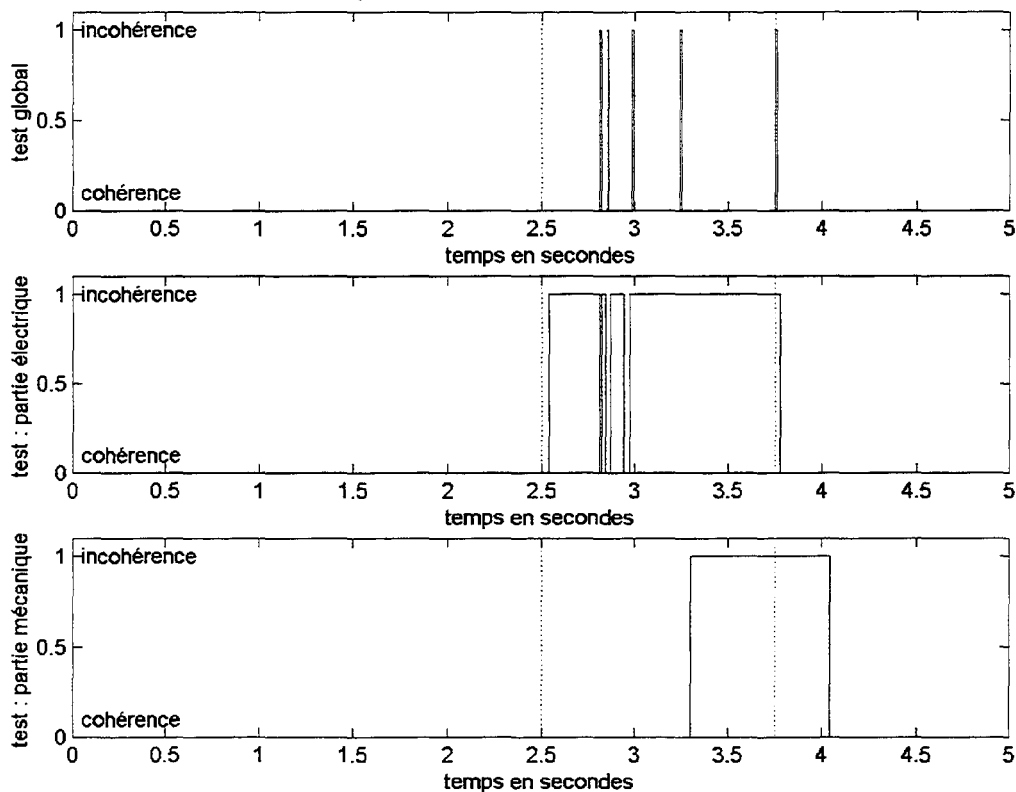


Figure XXI – Résultats de la procédure orientée bon fonctionnement en présence d'un court-circuit de 65% des spires rotoriques

Procédure globale orientée mauvais fonctionnement

Nous disposons d'un métamodèle (14) représentant le comportement de la machine à courant continu lorsque un défaut est présent. L'objectif d'une procédure orientée mauvais fonctionnement est de déterminer, si, durant un défaut, le comportement est cohérent ou non avec les modèles de mauvais fonctionnement disponibles. Le métamodèle (16) est, moyennant une approximation, la transposition du métamodèle (14) en temps discret. Il s'agit

ici aussi d'effectuer un test de cohérence entre les observations et ce modèle ; ce test est le même que le test global utilisé par la procédure orientée bon fonctionnement¹¹.

Un problème subsiste ; l'ensemble des spires non court-circuités évolue théoriquement sur $[0,100\%[$. Le métamodèle (16) nous offre la possibilité de représenter des ensembles $[\sigma(1-\rho_7), \sigma(1+\rho_7)]$. Il est dès lors possible de se caler sur le domaine d'évolution théorique en posant $\sigma=0,5$ et $\rho_7=1$,¹² mais est-ce réaliste ? En s'approchant trop des 100% de spires non court-circuitées, le métamodèle de mauvais fonctionnement inclura le bon fonctionnement et par conséquent, nous ne serons plus capable de discriminer l'un de l'autre. Pour déterminer la valeur limite devant être interprétée comme un défaut, nous avons choisi de retenir le plus petit nombre de spires court-circuitées que la procédure orientée bon fonctionnement perçoit comme un défaut. Dans l'exemple de la machine, nous avons retenu un court-circuit minimal de 40%¹³ des spires du rotor. Il est inutile d'aller jusqu'à 100% des spires court-circuitées : la machine ne tournerait plus et tout diagnostic serait vain. Nous nous arrêterons à un court-circuit de 80% des spires. Enfin, il est intéressant de définir différents ensembles de courts-circuits possibles, non seulement pour évaluer l'ampleur du court-circuit, mais aussi parce qu'un choisissant un domaine trop important ([40% 80%] par exemple), l'incertitude deviendrait très grande et le métamodèle d'autant plus faux du fait de la linéarisation autour de petites valeurs de variables abstraites. En effet, l'influence d'une variable abstraite correspondant à un défaut est a priori importante, ce qui doit être compensé par un découpage de l'espace abstrait de la variable. Nous avons choisi le découpage représenté sur la figure XXII. La zone grisée contient les intervalles que nous examinerons.

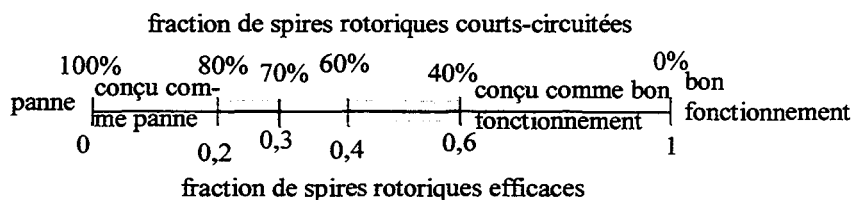


Figure XXII- Découpage du nombre de spires pour la procédure orientée mauvais fonctionnement

En suivant une démarche identique à celle adoptée pour les procédures orientées bon fonctionnement (paragraphe II.3), nous synthétisons l'hypomodèle à horizon fini, affine en les variables abstraites, associés au métamodèle complet (16). Nous avons réglé la valeur du terme additif λ , garant de la justesse de l'hypomodèle, de telle sorte que, pour différentes simulations, aucune incohérence ne soit détectée en présence d'une valeur de court-circuit représenté. Nous avons retenu la valeur $\lambda=0,05$.

¹¹ Mais le métamodèle de représente plus un bon fonctionnement mais un comportement anormal.

¹² ρ_7 légèrement inférieur à 1.

¹³ Cette valeur reste très importante, mais en considérant l'importance des incertitudes prises en compte, elle semble raisonnable. En effet, rien qu'en mesurant une valeur nominale de tension ou de courant, il y a 15% de la mesure qui est incertaine ! Si l'on ajoute une incertitude de 50% sur la résistance et de 10% sur le frottement...

Nous avons vérifié si, pour le court-circuit de 65%, dont les réponses apparaissent sur la figure XXI, il était possible de vérifier la cohérence avec l'hypomodèle à horizon fini. Nous n'examinerons que l'hypomodèle représentant la zone [60% 70%] de spires court-circuitées. Pour ce faire, nous nous sommes fixés : $\sigma=0,35$ et $\rho_7=0,05/0,35$. Le résultat du test de cohérence apparaît sur la figure XXIII.

On remarquera que la cohérence et l'incohérence ont été inversés sur cette figure car, cette fois-ci, nous recherchons si, durant le défaut détecté par la procédure BF (figure XXI), les observations étaient cohérentes avec le métamodèle complet du court-circuit compris entre 60% et 70%. La figure XXIII révèle une cohérence durant l'instant le défaut ce qui nous conduit à conclure que nous avons identifié le défaut.

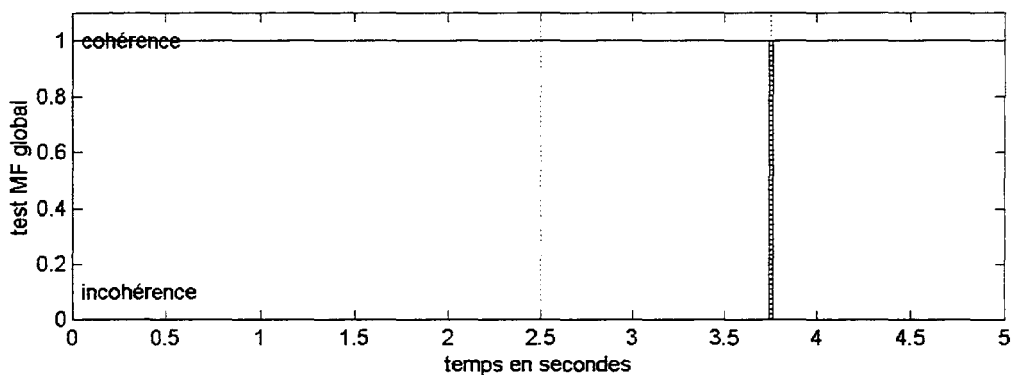


Figure XXIII – Résultats du test de cohérence avec procédure MF global en présence d'un court-circuit de 65% des spires rotoriques

Pourtant, en examinant la figure XXIII, nous constatons que, hormis juste après la disparition du défaut, la cohérence est observée partout, même en bon fonctionnement. Cela semble traduire une imprécision du test. Pour nous en assurer, examinons les résultats de ce test lorsque nous sommes en présence d'une augmentation du frottement à cinq fois sa valeur nominale et de la résistance à sept fois sa valeur nominale.

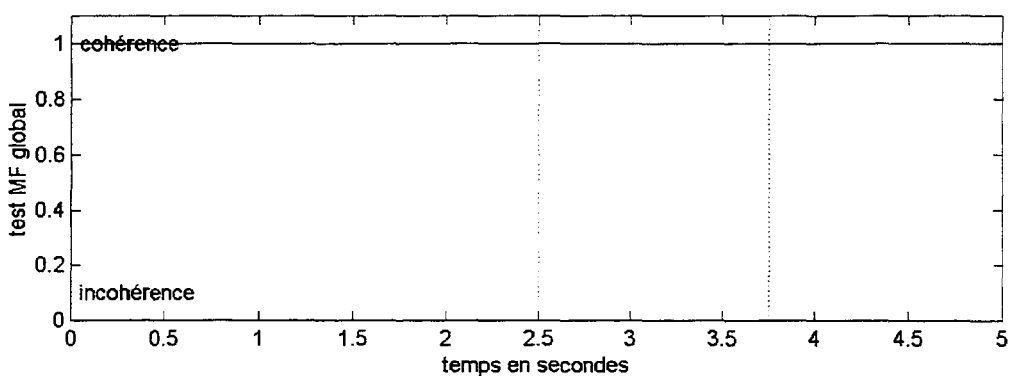


Figure XXIV - Résultats du test de cohérence avec procédure MF global en présence d'une multiplication par 5 du frottement et par 7 de la résistance

Bien que le défaut ne soit pas un court-circuit, la figure XXIV révèle une cohérence entre les observations et l'hypomodèle à horizon fini ! Nous en concluons que l'hypomodèle n'est pas

assez précis pour discriminer un défaut de 65% d'une augmentation anormale du frottement et de la résistance électrique.

Décomposition d'une procédure orientée mauvais fonctionnement

Nous avons vu, dans la partie II consacrée aux procédures orientées bon fonctionnement, que les métamodèles partiels conduisaient à des hypomodèles à horizon infini plus précis que l'hypomodèle à horizon fini reposant sur le métamodèle complet. C'est la solution que nous envisageons pour remédier à l'imprécision de la procédure globale. Les métamodèles réduits (17) et (18) engendrent les hypomodèles à horizon infini suivant :

$$\begin{aligned} \check{i}_{k+1} &= (A_0)_{11} \check{i}_k + (A_0)_{12} \check{\Omega}_k + (B_0)_{11} \check{u}_k - \left[\begin{array}{c} (A_1)_{11} \check{i}_k + (A_1)_{12} \check{\Omega}_k + (B_1)_{11} \check{u}_k \\ (A_2)_{11} \check{i}_k + (A_2)_{12} \check{\Omega}_k + (B_2)_{11} \check{u}_k \\ (A_7)_{11} \check{i}_k + (A_7)_{12} \check{\Omega}_k + (B_7)_{11} \check{u}_k \\ \rho_4 (A_0)_{12} \check{\Omega}_k \\ \rho_6 (A_0)_{12} \end{array} \right]_1 \\ \hat{i}_{k+1} &= (A_0)_{11} \hat{i}_k + (A_0)_{12} \hat{\Omega}_k + (B_0)_{11} \hat{u}_k + \left[\begin{array}{c} (A_1)_{11} \hat{i}_k + (A_1)_{12} \hat{\Omega}_k + (B_1)_{11} \hat{u}_k \\ (A_2)_{11} \hat{i}_k + (A_2)_{12} \hat{\Omega}_k + (B_2)_{11} \hat{u}_k \\ (A_7)_{11} \hat{i}_k + (A_7)_{12} \hat{\Omega}_k + (B_7)_{11} \hat{u}_k \\ \rho_4 (A_0)_{12} \hat{\Omega}_k \\ \rho_6 (A_0)_{12} \end{array} \right]_1 \end{aligned} \tag{19a}$$

avec $\mathcal{A}(i_k) = [\check{i}_k, \hat{i}_k]$

avec le test de cohérence :

$$\{\check{i}_k\} \in \left[\check{i}_k - \left[\begin{array}{c} \rho_3 \check{i}_k \\ \rho_5 \end{array} \right]_1 - \lambda_1, \hat{i}_k + \left[\begin{array}{c} \rho_3 \hat{i}_k \\ \rho_5 \end{array} \right]_1 + \lambda_1 \right] \tag{19b}$$

Tandis que le métamodèle (18) conduit à :

$$\begin{aligned} \check{\Omega}_{k+1} &= (A_0)_{22} \check{\Omega}_k + (A_0)_{21} \check{i}_k + (B_0)_{21} \check{u}_k - \left[\begin{array}{c} (A_1)_{22} \check{\Omega}_k + (A_1)_{21} \check{i}_k + (B_1)_{21} \check{u}_k \\ (A_2)_{22} \check{\Omega}_k + (A_2)_{21} \check{i}_k + (B_2)_{21} \check{u}_k \\ (A_7)_{22} \check{\Omega}_k + (A_7)_{21} \check{i}_k + (B_7)_{21} \check{u}_k \\ \rho_3 (A_0)_{21} \check{i}_k \\ \rho_5 (A_0)_{12} \end{array} \right]_1 \\ \hat{\Omega}_{k+1} &= (A_0)_{22} \hat{\Omega}_k + (A_0)_{21} \hat{i}_k + (B_0)_{21} \hat{u}_k + \left[\begin{array}{c} (A_1)_{22} \hat{\Omega}_k + (A_1)_{21} \hat{i}_k + (B_1)_{21} \hat{u}_k \\ (A_2)_{22} \hat{\Omega}_k + (A_2)_{21} \hat{i}_k + (B_2)_{21} \hat{u}_k \\ (A_7)_{22} \hat{\Omega}_k + (A_7)_{21} \hat{i}_k + (B_7)_{21} \hat{u}_k \\ \rho_3 (A_0)_{21} \hat{i}_k \\ \rho_5 (A_0)_{12} \end{array} \right]_1 \end{aligned} \tag{20a}$$

avec le test de cohérence :

$$\{\tilde{\Omega}_k\} \in \left[\tilde{\Omega}_k - \left\| \begin{bmatrix} \rho_4 \tilde{\Omega}_k \\ \rho_6 \end{bmatrix} \right\|_1 - \lambda_2, \hat{\Omega}_k + \left\| \begin{bmatrix} \rho_4 \hat{\Omega}_k \\ \rho_6 \end{bmatrix} \right\|_1 + \lambda_2 \right] \quad (20b)$$

Nous avons constaté, à partir de plusieurs séquences d'observations simulées¹⁴, que les valeurs $\lambda_1=0,8$ et $\lambda_2=0,001$ convenaient.

Il n'est plus question ici de rapporter les résultats des tests à des parties de la machine ; d'ailleurs, nous l'avons déjà signalé, les hypomodèles (19) et (20), et donc les métamodèles (17) et (18), ne correspondent ni à la partie électrique, ni à la partie mécanique de la machine. Les informations concernant chacune de ces parties sont diluées dans les tests et ne présentent d'ailleurs guère d'intérêt ici. Nous avons agrégé les résultats ; une incohérence sera détectée si une incohérence est observée sur l'un ou l'autre hypomodèle.

Pour parvenir au découpage de la figure XXII, nous devons avoir recours à trois hypomodèles dont les caractéristiques sont réunies dans le tableau suivant :

hypomodèle concernant un court-circuit	σ	ρ_7
de 70 à 80% des spires rotoriques	0,25	0,05/0,25
de 60 à 70% des spires rotoriques	0,35	0,05/0,35
de 40 à 60% des spires rotoriques	0,5	0,1/0,5

La procédure comporte donc six tests de cohérence agrégés en trois : un résultat par paire d'hypomodèle (19) et (20).

Les résultats obtenus lors d'un court-circuit de 65% des spires rotoriques de la machine sont représentés sur la figure XXV. Notons qu'il nous a fallu 12s. sur un ordinateur Pentium 133MHz pour obtenir ces résultats avec Matlab. Nous constatons que, durant le défaut, aucune incohérence n'a été révélée par aucun des hypomodèles. Il aurait suffi d'une unique incohérence durant la durée du défaut (détectée par la procédure BF), pour exclure une hypothèse de défaut. Les résultats de la procédure BF nous ont mené à conclure qu'un défaut affectant la partie électrique ainsi que la partie mécanique était présent entre les instants 2,5s. et 3,75s. La procédure MF nous amènent à compléter ce diagnostic : le défaut est très probablement un court-circuit au rotor dont l'amplitude est comprise entre 40¹⁵ et 80% des spires.

La figure XXVI représente les résultats de la procédure orientée mauvais fonctionnement lors d'une multiplication par cinq de la valeur du frottement et par sept de la valeur de la

¹⁴ Comme pour les procédures orientées bon fonctionnement, nous avons recherché les plus petites valeurs de λ_1 et λ_2 qui, lorsque le défaut est présent, conduisent à aucune incohérence.

¹⁵ En deçà, la procédure BF n'aurait rien détecté.

résistance.

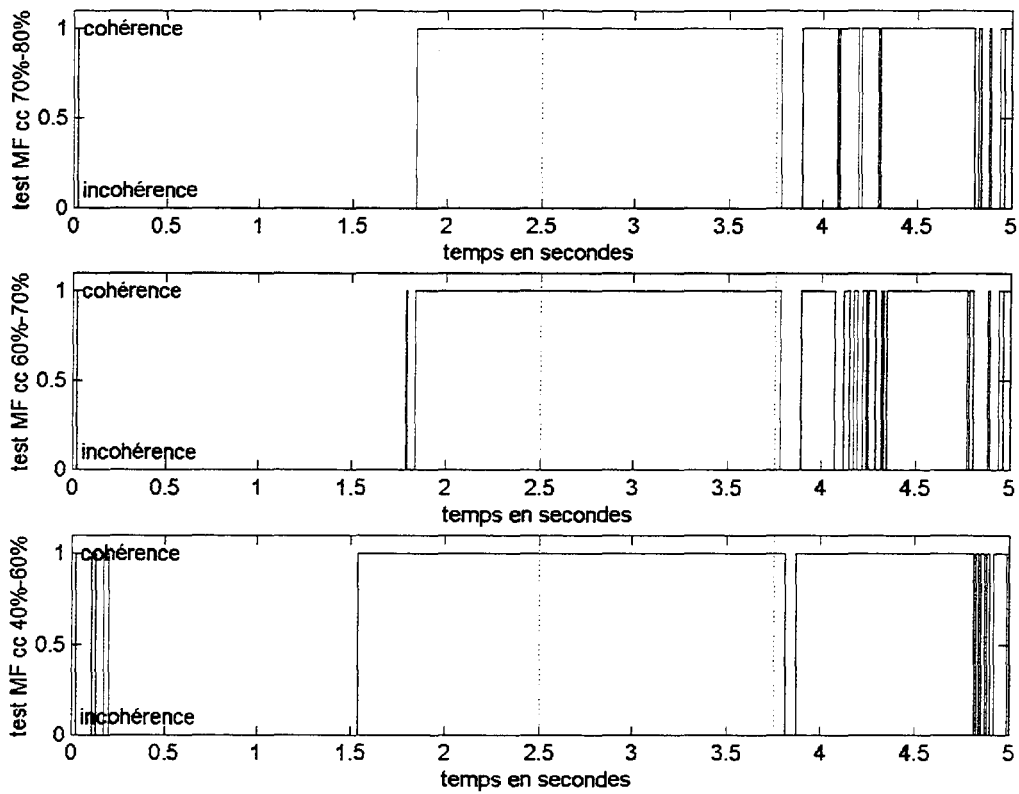


Figure XXV – Résultats de la procédure MF lors d'un court-circuit de 65% des spires rotoriques

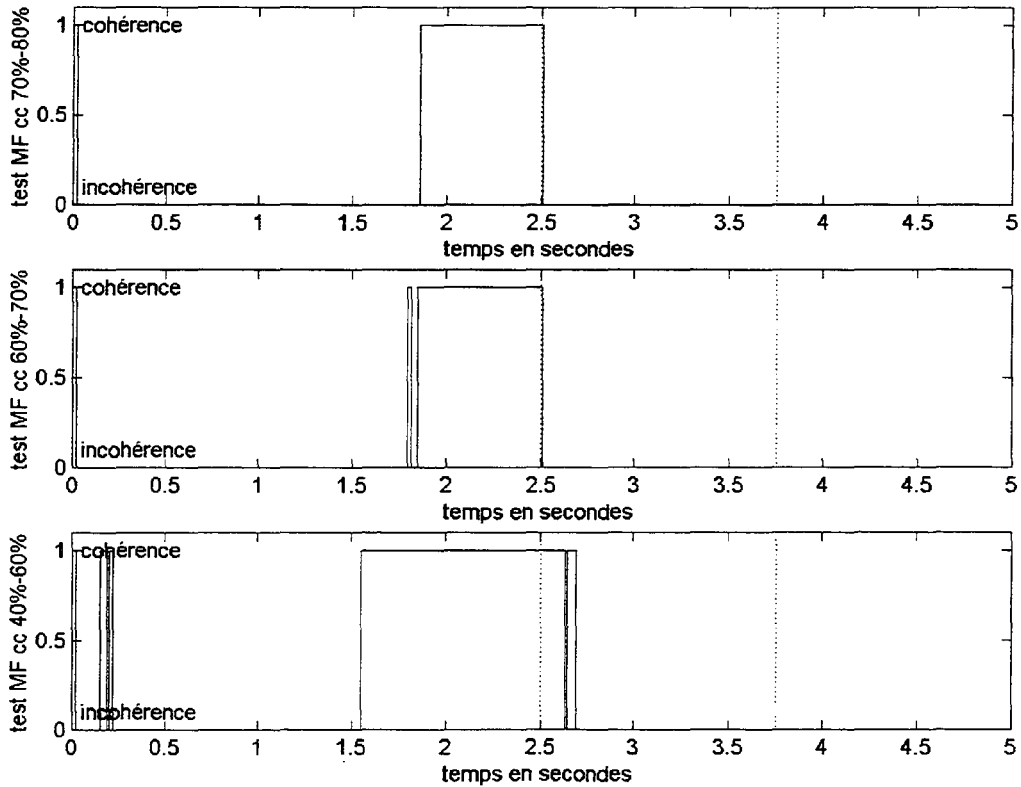


Figure XXVI - Résultats de la procédure MF lors d'une multiplication par 5 du frottement et par 7 de la résistance électrique

Les résultats (figure XXI) de la procédure orientée bon fonctionnement nous avait permis de conclure qu'un défaut était survenu entre les instants 2,7s. et 3,9s et qu'il affectait les parties électrique et mécanique de la machine. Les résultats de la procédure MF nous permettent d'améliorer le diagnostic en révélant que le défaut ne correspond pas à un court-circuit puisque de nombreuses incohérences entre les observations et les hypomodèles associés aux courts-circuits sont détectées durant le moment où le défaut est présent. Ces deux défauts qui conduisaient aux mêmes symptômes avec la procédure BF peuvent être discriminés avec la procédure MF.

Pour finir, nous avons simulé un court-circuit de 80% des spires rotoriques. La figure XXVII illustre le résultat que nous avons obtenu. Pour ce défaut, des incohérences ont été détectées par les hypomodèles se rapportant à des courts-circuits de 40 à 60% et de 60 à 70%. En revanche, aucune incohérence n'a été détectée, durant le défaut, par l'hypomodèle associé à des courts-circuits de 70 à 80%. Nous concluons naturellement que le défaut est un court-circuit compris entre 70 et 80% !

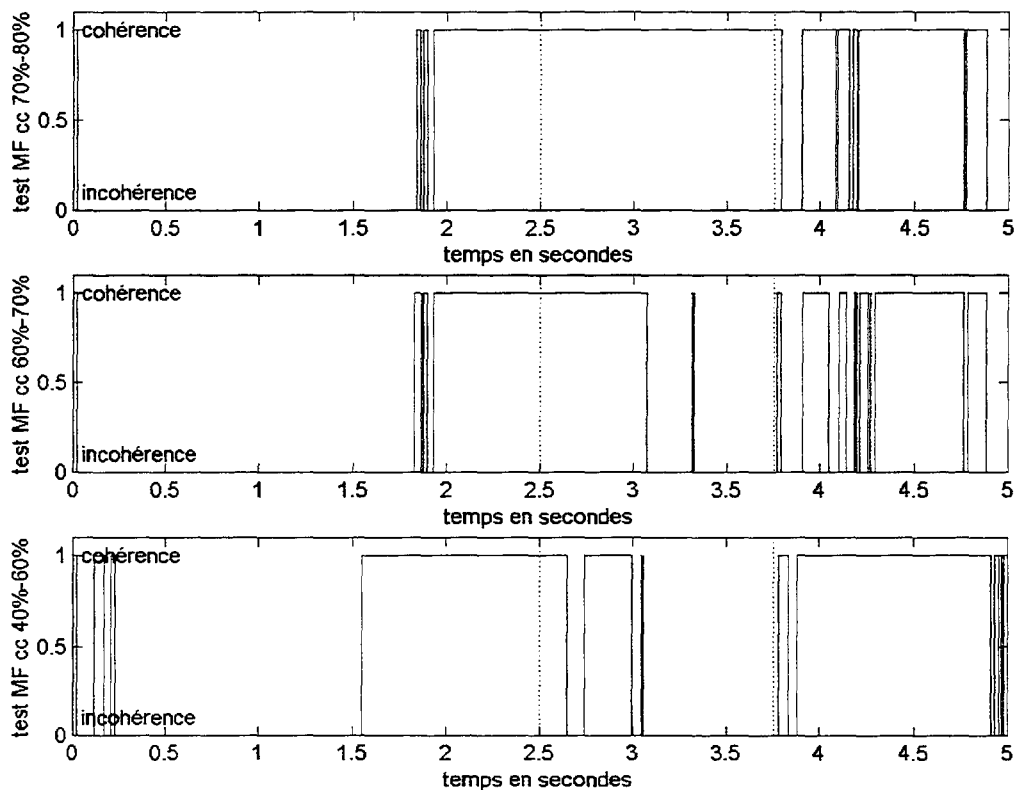


Figure XXVII - Résultats de la procédure MF lors d'un court-circuit de 65% des spires rotoriques

IV. Conclusion

Ce chapitre nous a permis, à partir de l'exemple de la machine à courant continu, d'examiner les problèmes liés à l'utilisation d'hypomodèles incertains pour la réalisation de procédures de diagnostic. Nous avons montré que les hypomodèles incertains ainsi que les tests de

cohérence qui leur sont associés permettaient de réaliser autant des procédures orientées bon que mauvais fonctionnement. Nous avons constaté que les hypomodèles incertains à horizon infini présentaient, dans la cas de la machine à courant continu, le double avantage d'être plus rapides à mettre en œuvre et surtout plus précis que les hypomodèles à horizon fini. Ceci semble contredire nos assertions du chapitre 4 ; cependant, il ne faut pas oublier que toutes les variables d'état du modèle de la machine étaient mesurées. Il nous a donc été possible de décomposer le métamodèle complet en métamodèles partiels ne contenant qu'une seule équation récurrente : une des rares situations où un hypomodèle à horizon infini peut être exploité. Cette situation n'a néanmoins aucune raison d'être dans le cas général où il n'est plus possible d'avoir recours à de tels hypomodèles.

Nous avons montré qu'un défaut était détecté lorsqu'une incohérence ou un ensemble d'incohérences apparaissait au niveau de la procédure orientée bon fonctionnement que ce soit lors des tests globaux ou des tests locaux. Dans l'exemple, il est clair que les tests locaux étaient plus précis que le test global ; bien que le test global permette d'être plus fiable quant aux instants d'apparition et de disparition du défaut, les tests locaux auraient suffi. Or, dans le cas général où les précisions des tests globaux et locaux sont comparables, le test global est le premier à être effectué et, lors d'une incohérence, il déclenche les tests locaux pour affiner le diagnostic. Enfin, pour compléter le diagnostic, nous avons montré comment une procédure orientée mauvais fonctionnement pouvait être réalisée. Cette procédure était dédiée aux défauts de type courts-circuits entre spires du rotor. Une telle procédure n'est déclenchée que si le résultat des tests locaux de la procédure BF conclut à la possibilité d'un tel défaut. En l'occurrence, la procédure MF ne sera amorcée que si la procédure BF a établi que le défaut affectait la partie électrique et la partie mécanique de la machine.

Pour clore ce chapitre, signalons qu'il est parfois difficile d'obtenir un modèle du mauvais fonctionnement. Nous avons d'ailleurs établi un modèle sommaire représentant un court-circuit entre les spires d'une même phase statorique au sein d'une machine asynchrone [Ploix,97], pour lequel la structure changeait complètement : nous sommes passés d'un modèle de quatre à cinq équations différentielles (équation mécanique exclue). Quoiqu'il en soit, même si les modèles de mauvais fonctionnement sont difficiles à obtenir, leur absence n'est pas rédhibitoire : plus il y en aura, plus précis en sera le diagnostic.

Conclusion

Nous avons montré que l'approche bornante est particulièrement adaptée à la réalisation de procédures automatisées de diagnostic de défauts pour des systèmes physiques ne pouvant raisonnablement être représentés par des modèles comportant des incertitudes. En effet, d'une part, les incertitudes de modélisation, inhérentes à la majorité des systèmes physiques, ne peuvent pas être prises en compte par des procédures purement déterministes et, d'autre part, les approches stochastiques ne s'appliquent simplement que dans le cas de structures additives d'incertitude. De ce fait, nous avons proposé de modifier les formalismes courants de l'approche bornante, de façon à appréhender les problèmes qui se posent en diagnostic. Grâce à un raisonnement conceptuel, nous avons mis en évidence le fait que toute procédure de diagnostic procède nécessairement de tests de cohérence. En introduisant de nouveaux concepts, parmi lesquels ceux de *métamodèle* et d'*hypomodèle*, nous avons montré que les tests de cohérence reposent directement sur des hypomodèles. Nous avons aussi fait apparaître que le comportement auquel se rapporte un hypomodèle, est formellement représenté par un métamodèle. Dès lors, toute opération de découplage revient à réduire le métamodèle original effectivement exploité et par là même, implicitement, à modifier les objectifs d'une opération de diagnostic ; une telle opération ne peut donc être conçue, contrairement à ce que proposent les démarches traditionnelles, comme une solution. En raisonnant sur les modèles déterministes, nous avons montré que les hypomodèles à horizon fini sont plus intéressants que ceux à horizon infini dans la mesure où les incohérences coïncident mieux avec les instants où le défaut est présent et que, de toute façon, les autres résidus se déduisent de l'hypomodèle à horizon fini minimal. Ceci se transpose évidemment au contexte incertain : plus l'horizon des hypomodèles est petit, meilleur est la détermination des instants où un défaut est présent. Pour rendre exploitable l'approche bornante, nous avons étudié comment un métamodèle en temps discret approché pouvait se déduire d'un métamodèle en temps continu comportant des incertitudes puis, nous avons recherché les différents tests de cohérence envisageables. Une autre démarche consiste à évaluer les caractéristiques des incertitudes apparaissant dans un modèle de comportement en exploitant une séquence de mesures : un algorithme de caractérisation a été introduit. Puis, nous avons clairement établi que, à l'exception des modèles dynamiques ne comportant qu'une équation récurrente, les hypomodèles incertains à horizon fini étaient les seuls réellement exploitables. Nous avons alors défini une procédure de test de cohérence entre cet hypomodèle et des observations, qui repose sur le calcul analytique de domaines abstraits. Plusieurs autres procédures numériques de tests ont été présentées mais, pour les modèles affines en les variables abstraites, nous avons montré que les procédures analytiques étaient les plus intéressantes. Nous avons enfin conclu par la synthèse d'une procédure de diagnostic automatique, prenant en compte les incertitudes, consacrée à une machine à courant continu, simulée avec des défauts et des variations normales de paramètres. En ayant recours à des

procédures orientées bon et mauvais fonctionnement, nous avons réussi à diagnostiquer des défauts que l'approche déterministe n'aurait pas permis de détecter.

Ce travail est un nouveau pas vers le diagnostic des systèmes physiques réels qui ont la particularité de ne pas admettre de modèles de comportement déterministes justes. Nous avons abordé et proposé des solutions aux différents stades de la synthèse d'une procédure de diagnostic. De nombreuses autres voies demeurent à explorer, notamment les problèmes liés à l'exploitation des structures d'incertitude complexes ; nous n'avons décrit, en effet, que des procédures exploitant des hypomodèles affines en les variables abstraites. N'est-il pas possible de gagner en précision en utilisant des modèles à structure plus complexe ? Comment pourrait-on obtenir de telles structures en partant de métamodèles en temps continu ? Et, dans ce contexte, comment effectuer des tests de cohérence efficaces ? Ces questions ainsi que tant d'autres restent à élucider... Par ailleurs, il est clair que ce travail appelle, parmi ces prolongements, une application à un système physique réel. Pour ce faire, il faut examiner les possibilités d'optimisation des procédures de diagnostic, autant du point de vue de la rapidité que de leur performance. Ne serait-il pas intéressant de compléter un test de cohérence par une distance permettant d'apprécier si l'incohérence est forte ? Enfin, l'algorithme de caractérisation que nous avons présenté n'a pas été appliqué à des modèles multidimensionnels. Il reste à prolonger l'étude des performances de cet algorithme... Il est difficile d'énumérer tous les problèmes en suspens dans la mesure où cette approche est nouvelle : le champ des investigations possibles est vaste.

Notre objectif n'était que de proposer une nouvelle approche du diagnostic comblant les lacunes des approches traditionnelles. La conclusion se résume en quelques mots : l'approche bornante en diagnostic est possible, mais il reste à étudier les modalités de son application...

Bibliographie

- AFCET, 1988, *Modèle et maîtrise des systèmes*, Congrès AFCET.
- AFNOR, 1994, *Maintenance - Concepts et définitions des activités de maintenance*, Norme NF X 60-010, Association Française de Normalisation, 28 pages.
- ARSAN M. I., 1994, *Observateurs et diagnostic automatique de pannes*, Thèse de l'école nationale supérieure de l'aéronautique et de l'espace.
- BARMISH B.R., SANKARAN J., 1979, *The propagation of parametric uncertainty via polytopes*, IEEE transactions on Automatic Control, vol. AC24 n°2 pp. 346-349
- BASILE G., MARO G., 1992, *Controlled and conditioned invariants in linear system theory*, Englewood Cliff, NJ: Prentice Hall.
- BELFORTE G., BONA B., CERONE V., 1990, *Parameter estimation algorithms for a Set-membership description of uncertainty*, Automatica, vol.26, n°5, pp. 887-898.
- BELFORTE G., TAY T.T., 1996, *Recursive estimation algorithms for linear models with set membership error*, pp. 83-99 dans [Milanese et al,96].
- BERTSEKAS D.P., RHOADES I.B., 1971, *Recursive state estimation for a set-membership description of uncertainty*, IEEE transactions on Automatic Control, vol. AC16 n°2, pp. 117-128.
- CERONE V., 1993, *Parameter bounds for ARMAX models from records with bounded errors in variables*, International Journal of Control, Vol.57, n°1, pp. 225-236 et repris dans [Milanese,96].
- CHANG I.C., YU C.C., LIOU C.T., 1995, *Model-Based approach for fault diagnosis, part.2. Extension to interval systems*, Ind. Eng. Chem. Res, n°34, pp. 828-844.
- CHATAIN J.-N., 1993, *Diagnostic par système expert*, Traité des Nouvelles Technologies, série Diagnostic et Maintenance, Hermès, Paris.
- CHOW E.Y., 1980, *A failure detection system design methodology*, Ph.D. of Massachusetts institute of technology.
- CHOW E.Y., WILLSKY A.S., 1984, *Analytical redundancy and the design of robust failure detection systems*. IEEE transactions on Automatic control, AC-29, pp. 603-614.
- CLARK R.N., 1978, *Instrument fault detection*, IEEE transactions on Aerospace and Electronic systems, vol. AES14 n°2, pp. 456-465.
- CLARK R.N., 1989, *State estimation schemes for instrument fault detection*, dans [Patton,89], pp. 21-45.

- COCQUEMPOT V., 1993, *Surveillance des processus industriels complexes*, Thèse de l'Université des Sciences et Technologies de Lille.
- DAROUACH M., 1994, *On the novel approach to the design of unknown input observers*, IEEE transactions on Automatic Control, vol. AC39, pp. 698-699.
- DE KLEER J., WILLIAMS B.C., 1987, *Diagnosing multiple faults*, Artificial Intelligence, n° 32, pp. 97-130.
- DESCARTES R., 1637, *Discours de la méthode*, Larousse.
- FILIPPOVA T.F., KURZHANSKI A.B., SUGIMOTO K., VALYI I., *Ellipsoïdal state estimation for uncertain dynamical systems*, dans [Milanese,96]
- FOGEL E., HUANG Y.F., 1982, *On the value of information in system identification-bounded noise case*, Automatica, vol.18, n°2, pp. 229-238.
- FRANK P.M., 1990, *Fault diagnosis in dynamic systems using analytical and knowledge-based redundancy - A survey and some new results*, Automatica, vol. 26, n° 3, pp. 459-474.
- GERTLER J. J., 1988, *Survey of Model-Based failure detection and isolation in complex plants*, IEEE Control Systems Magazine, vol. 11, pp. 3-11.
- GERTLER J., Singer D., 1990, *A new framework for parity equations-based failure detection and isolation*, Automatica, vol. 26 n°2, pp. 381-388.
- GIULIANI S., 1998, *Diagnostic du réseau d'évacuation des eaux du bassin versant nancéien*, Thèse de l'Institut National Polytechnique de Lorraine.
- GOLUB G.H., VAN LOAN C.F., 1989, *Matrix computations* (2nd edition), The John Hopkins University Press, Baltimore and London.
- GOODWIN G.C., MIDDLETON R.H., 1989, *The class of all stable unbiased state estimator*, Syst. Contr. Lett., vol. 13, pp. 161-167.
- GRAY R.M., DAVISSON L.D., 1986, *Random processes: a mathematical approach for engineers*, Prentice Hall Inc., Englewood Cliffs, N. J.
- HAMZA M.H., RASMY M.E., 1971, *A simple method for determining the reachable set for linear discrete systems*, IEEE transactions on Automatic Control, vol. AC16 n°3, pp. 281-282.
- HANSEN E., 1969, *Topics in Interval Analysis*, Clarendon Press, Oxford.
- HEISENBERG W., 1962, *La nature dans la physique contemporaine*, Gallimard, Paris.

- HORAK D. T., ALLISON B. H., 1990, *Failure detection and isolation methodology*, Proc. of the 1990 American Control Conference, pp. 2955-2958.
- HOU M., MÜLLER P.C., 1992, *Design of observers for linear systems with unknown inputs*, IEEE transactions on Automatic Control, vol. AC37 n°6, pp. 871-875.
- HOU M., MÜLLER P.C., 1994, *Disturbance decoupled observer design: a unified view point*, IEEE transactions on Automatic Control, vol. AC39 n°6, pp. 1338-1341.
- IEEE 1988, *IEEE Standard dictionary of Electrical and Electronics terms*, 4th edition, F. Jay (Ed.), The Institute of Electrical and Electronics Engineers, New York.
- ISERMANN R., BALLE P., 1997, *Trends in the application of model-based fault detection and diagnosis of technical processes*, Control Engineering Practice, vol. 5, n° 5, pp. 709-719.
- ISIDORI A., 1989, *Nonlinear control systems* (2nd edition), Springer Verlag, New-York.
- IWASAKI Y., SIMON H. A., 1994, *Causality and model abstraction*, Artificial Intelligence, n° 67, pp. 143-194.
- JAULIN L., WALTER E., 1996, *Guaranteed nonlinear set estimation via interval analysis*, pp. 363-381 dans [Milanese et al,96].
- KANT E., 1787, *Critique de la raison pure*, La pléiade, Gallimard.
- KOBI A., 1994, *Diagnostic de processus continus : application à la détection de valeurs aberrantes dans les signaux d'entrée et de sortie de systèmes*. Thèse de doctorat de l'Institut National Polytechnique de Lorraine, spécialité Automatique.
- KOYRE A., 1966, *Etudes galiléennes*, Gallimard, Paris.
- F. KRATZ, W. NUNINGER, S. PLOIX, 1998, *Fault detection for time-delay systems: a parity space approach*, American Control Conference 1998, Philadelphia.
- LE PETIT ROBERT, 1994, *Dictionnaire alphabétique et analogique de la langue française*, Dictionnaire Le Robert, Paris.
- LEONHARDT S., AYOUBI M., 1997, *Methods of fault diagnosis*, Control Engineering Practice, vol. 5, n° 5, pp. 683-692.
- LIN J.N., 1970, *Determination of reachable set for a linear discrete system*, IEEE transactions on Automatic Control, vol. AC15 n°3, pp. 399-342.
- LJUNG L., 1987, *System identification : theory for the user*, Prentice Hall, Englewood Cliffs, N. J.

- LUENBERGER D.G., 1971, *An introduction to observers*, IEEE transactions on Automatic Control, vol. AC16 n°6, pp. 596-602.
- MAGNI J.-F., MOUYON P., 1991, *A generalised approach to observers for fault diagnosis*, Conf. on Decis. Contr., pp. 2236-2241.
- MAKSIMOVIC A., MILHAU S., PLOIX S., 1993, *Etude de la mise en œuvre de correcteurs par retour d'état pour l'asservissement en vitesse d'une machine à courant continu à excitation séparée et contrôle du courant en régime dynamique*, Mémoire d'ingénieur de l'école spéciale de mécanique et d'électricité.
- MASSOUMNIA M.M., VAN DER VELDE W.E., 1988, *Generating parity relations for detecting and identifying control system component failures*, Journal of guidance, control and dynamics, vol. 11 n°1, pp. 60-65.
- MEDVEDEV A., TOIVONEN H.T., 1991, *Investigation of a finite memory observer structure*, Process control laboratory, Dept. of chemical Engineering, Abo Akademi, Report 91-6.
- MILANESE M., BELFORTE G., 1982, *Estimation theory and uncertainty intervals evaluation in presence of unknown but bounded errors: linear families of models*, IEEE Trans. Autom. Contr., AC-27, n°2, pp. 408-413.
- MILANESE M., NORTON J., PIET-LAHANIER H., WALTER E. (ed.), 1996, *Bounding approaches to system identification*, Plenum Press, New-York et Londres.
- MILNE R., 1987, *Strategies for diagnosis*, IEEE Transactions on Systems, Man and Cybernetics, vol. 17, n° 3, pp. 333-339.
- MO S.H., NORTON J.P., 1988, *Recursive parameter-bounding algorithms which compute polytope bounds*, 2nd IMACS proc., 12th world congress on scientific computation, Paris.
- MOORE R.E., 1966, *Interval Analysis*, Prentice Hall Inc., Englewood Cliffs, N. J.
- MOORE R.E., 1979, *Methods and applications of interval analysis*, SIAM, Philadelphia, Pennsylvania.
- NEUMAIER, 1990, *Interval methods for systems of equations*, Cambridge University Press, Cambridge.
- NIJMEIJER H., Van der Schaft A.J., 1990, *Nonlinear control dynamical systems*, Springer Verlag
- NORTON J.P., 1987, *Identification and application of bounded-parameter models*, Automatica, vol.23, n°4, pp. 497-508.

- NORTON J.P., 1989, *Recursive of inner bounds for the parameters of linear models*, International Journal of Control, Vol.50, n°6, pp. 2423-2430.
- O'REILLY J., 1983, *Observers for linear systems*, Academic Press.
- PARROCHIA D., 1991, *Le réel*, Bordas, Paris.
- PASCAL G., 1995, *Kant*, Bordas, Paris.
- PATTON R., FRANK P., CLARK R. (Ed.), 1989, *Fault diagnosis in dynamic systems*, Prentice Hall, International series in systems and control engineering.
- PATTON R.J., 1994, *Robust model-based fault diagnosis: the state of the art*, Proc. of the SAFEPROCESS'94 conference, pp. 1-24.
- PEEBLES P.Z. Jr, 1987, *Probability, random variables and random signal principles*, McGraw-Hill, 2nd edition, Paris.
- PENALVA J.-M., 1990, *SAGACE : Une représentation des connaissances pour la supervision de procédés continus*, Les systèmes experts et leurs applications, 10^e journées Internationales, Avignon, pp. 363-378.
- PLOIX S., GIULIANI S., RAGOT J., 1999, *An unifying approach of LTI discrete-time State observers by means of complementary matrices*, IEEE trans. Autom. Contr. A paraître Février vol.44 n°2.
- PLOIX S., ADROT O., KRATZ F., IUNG C., RAGOT J., 1997, *An extended model of induction machine for failure detection*, IFAC Conference on Industrial Systems, Belfort, pp. 55-62.
- PLOIX S., RAGOT J., MAQUIN D., 1997, *Asymptotic state-observer for a non-linear descriptor system*, Safeprocess'97-Huhl, Angleterre, pp. 654-658.
- PLOIX S., VILLIERE C., RAGOT J., 1997, *Discrete time models of induction machine with additive uncertainties*, safeprocess'97-Huhl, Angleterre, pp. 264-269.
- PLOIX S., KRATZ F., RAGOT J., ADROT O., 1997, *Approche conceptuelle du diagnostic et préliminaires à la mise en œuvre sur la machine asynchrone*, Bilan des activité PRC diagnostic de la machine asynchrone, AGI'97, Poitiers.
- PLOIX S., KRATZ F., RAGOT J., ADROT O., 1997, *Diagnostic de la Machine Asynchrone à base de modèles analytiques*, Bilan des activité PRC diagnostic de la machine asynchrone, AGI'97, Poitiers.
- POINCARÉ H., 1902, *La science et l'hypothèse*, Flammarion, Paris.
- PRONZATO L., WALTER E., *Volume-optimal inner and outer ellipsoïds*, dans [Milanese,96]

- REITER R., 1987, *A theory of diagnosis from first principles*, Artificial Intelligence, n° 32, pp. 57-95.
- RICHALET J., 1991, *Pratique de l'identification*, Hermès, Traité des nouvelles technologies, série Automatique.
- SAINT-GUILHEM R., 1991, *Notions fondamentales de mathématiques modernes*, tome 2, ellipses, Paris.
- SAVKIN A.V., PETERSEN I.R., 1995, *Recursive state estimation for uncertain systems with an integral quadratic constraint*, IEEE transactions on Automatic Control, vol. AC40 n°6, pp. 1080-1083.
- SCHWEPPE F.C., 1968, *Recursive state estimation : unknown but bounded errors and system inputs*, IEEE Trans. Autom. Contr., AC-13, n°1, pp. 22-28.
- SCHWEPPE F.C., 1973, *Uncertain dynamic systems*, Prentice Hall, Englewood Cliffs, N. J.
- SHAFAI B., CAROLL R. L., 1989, *Design of proportional-integral observer for linear time-varying multivariable systems*, Proc. of 24th conf. on Dec. and Contr., pp. 597-599.
- STAROSWIECKI M., COCQUEMPOT V., CASSAR J.P., 1991, *Observer based and parity space approaches for failure detection and identification*, Proceedings of IMACS-IFAC International Symposium, Lille, May 5-11, pp. 536-541.
- STRUSS P., 1992, What's in SD ? *Towards a theory of modeling for diagnosis*, dans HAMSCHER W., CONSOLE L. & DE KLEER J, Readings in model-based diagnosis, Morgan Kaufman, pp. 419-448.
- TANINO T., OHTA J.I., 1989, *Set theoretic approach to uncertain dynamical systems using convex polytopes*, Int. Journ. of syst. sci., vol. 20 n°10, pp. 1859-1873.
- TOURNANDRE G., 1982, *L'orientation de la science Cartésienne*, Vrin, Paris.
- TRAIVE-MASSUYES L., DAGUE P., GUERRIN F., 1997, *Le raisonnement qualitatif*, Hermès, Paris.
- VERES S.M., NORTON J.P., 1996, *Parameter-bounding algorithms for linear errors-in-variables models*, pp. 275-288 dans [Milanese et al,96].
- VESELY W.E., GOLDBERG F.F., ROBERTS N.H., HAASL D.F., 1981, *Fault tree Handbook*, U.S. Nuclear Regulatory Commission, Washington, USA.
- VICINO A., MILANESE M., 1991, *Optimal inner bounds of feasible parameter set in linear estimation with bounded noise*, IEEE Trans. Autom. Contr., AC-36, n°6, pp. 759-762.
- VICINO A., ZAPPA G., 1996, *Adaptative approximation of uncertainty sets for linear regression models*, pp. 159-181, dans [Milanese et al,96].

- VIDYASAGAR M., 1993, *Nonlinear Control dynamical systems*, Springer Verlag.
- VON BERTALANFFY L., 1993, *Théorie générale des systèmes*, trad. J-B. Chabrol, Dunod, Paris.
- WALTER E., PIET-LAHANIER H., 1997, *Exact an recursive description of feasible parameter set for bounded error models*, 26th IEEE Conf. on Dec. and Contr., Los Angeles, pp. 1921-1922.
- WALTER E., PRONZATO L., 1994, *Identification de modèles paramétriques à partir de données expérimentales*, Masson, MASC, Paris.
- WEBSTER'S NEW ENCYCLOPEDIA DICTONNARY, 1996, Le prat, New-York, USA.
- WILLSKY A., 1976, *A survey of design methods for failure detection in dynamic systems*, Automatica, vol. AC-21, pp. 601-611.
- WÜNNENBERG J., FRANK P.M., 1990, *Robust observer-based fault detection for linear and nonlinear systems with applications to robots*, Proc. Of MIM-S2'90 symp., pp. 2.1-2.8.
- ZWINGELSTEIN G., 1995, *Diagnostic des défaillances - Théorie et pratique pour les systèmes industriels*, Traité des Nouvelles Technologies, série Diagnostic et Maintenance, Hermès, Paris.

AUTORISATION DE SOUTENANCE DE THESE DU DOCTORAT DE L'INSTITUT NATIONAL POLYTECHNIQUE DE LORRAINE

o0o

Service Commun de la Documentation
INPL
Nancy-Brabois
Nancy-Brabois
Nancy-Brabois

VU LES RAPPORTS ETABLIS PAR :

Monsieur STAROSWIECKI Marcel, Professeur, LAII/Université de Lille 1,

Monsieur GUGLIELMI Michel, Professeur, IRCYNEcole Centrale de Nantes,

Monsieur FAUCHER Jean, Professeur, ENSEEIHT Toulouse.

Le Président de l'Institut National Polytechnique de Lorraine, autorise :

Monsieur PLOIX Stéphane

à soutenir devant un jury de l'INSTITUT NATIONAL POLYTECHNIQUE DE LORRAINE, une thèse intitulée :

"Diagnostic des systèmes incertains : l'approche bornée."

en vue de l'obtention du titre de :

DOCTEUR DE L'INSTITUT NATIONAL POLYTECHNIQUE DE LORRAINE

Spécialité : **"AUTOMATIQUE & TRAITEMENT NUMÉRIQUE DU SIGNAL"**

Fait à Vandoeuvre le, 7 Décembre 1998

Pour le Président
Le Vice-Président


J. HARDY



NANCY BRABOIS
2, AVENUE DE LA
FORET-DE-HAYE
BOITE POSTALE 3
F - 5 4 5 0 1
VANDŒUVRE CEDEX

Le diagnostic de défauts au sein de systèmes physiques a fait l'objet de nombreux travaux ; hormis les approches stochastiques, la plupart s'inscrivent dans un contexte déterministe. Nous défendons la thèse selon laquelle les approches bornantes sont les mieux adaptées à la synthèse de procédure automatique de diagnostic s'appliquant aux systèmes physiques réels. A partir d'un raisonnement conceptuel, nous avons clairement établi la problématique du diagnostic et montré que toute procédure de diagnostic procède nécessairement de tests de cohérence. Nous avons montré que, dans un contexte incertain, les modèles contenant des variables bornées répondent tout à fait aux objectifs du diagnostic. Un formalisme permettant d'appréhender les propriétés des variables bornées est présenté ainsi qu'une procédure analytique conduisant aux caractéristiques d'un domaine défini par un champ vectoriel. Après avoir étudié différentes structures d'incertitudes, nous présentons un algorithme de caractérisation dont l'objectif est d'estimer les caractéristiques des incertitudes. Nous abordons alors les modèles permettant de réaliser les tests de cohérence nécessaires au diagnostic. En partant des modèles déterministes et en nous fondant sur des formulations génériques, nous commençons par mettre en évidence l'influence de l'horizon et de la dimension d'un modèle sur l'objectif de diagnostic. Puis, nous abordons les moyens de tester la cohérence entre un modèle incertain et des observations. Deux types de modèles ont été étudiés : ceux dont l'horizon est infini et ceux à horizon fini. Nous discutons de la pertinence des tests de cohérence associés à ces modèles. Ce mémoire s'achève par la réalisation de procédures de diagnostic exploitant la connaissance des incertitudes dans le cas de la machine à courant continu.

Mots clés : diagnostic de défauts, systèmes incertains, approche bornante, observateurs d'état, relation de parité.

Many authors tackled fault diagnosis within physical systems. Most of their works are carried out either in the deterministic framework or in the stochastic one. We defend the thesis that bounding approach is the most relevant for designing automatic procedures for diagnosis of physical systems which can only be represented by an uncertain model. Thanks to a conceptual reasoning, the problematic of diagnosis is clearly set and it is shown that any diagnosis procedure necessarily undertakes tests of consistency. Then, it is pointed out that, in an uncertain framework, models containing bounded variables are well suited for diagnosis aims. Furthermore, a formalism is developed in order to apprehend the properties of bounded variables. An analytical procedure determining characteristics of bounded domains defined by vector fields is also presented. Otherwise, different structures of uncertainties are studied and an algorithm that estimates the bounds of uncertainties is further detailed. Models allowing tests of consistency are then focused on. From deterministic models and generic formulations, the influence on diagnosis of both horizon and dimension of models is brought to the fore. Finally, purposes of achieving tests of consistency between uncertain models and observations are examined. Two types of models have been studied: finite and infinite horizon. The relevance of consistency tests associated to these models is discussed. This thesis ends by an example of diagnosis procedure taking into account the knowledge on uncertainties in the case of a DC machine.

Key words: fault diagnosis, uncertain systems, bounding approach, state observer, parity space.