



Institut National Polytechnique de Lorraine
Ecole Doctorale IAE+M
DFD Automatique et Production Automatisée
Centre de Recherche en Automatique de Nancy



Thèse

présentée pour l'obtention du

Doctorat de l'Institut National Polytechnique de Lorraine

Spécialité Automatique et Traitement Numérique du Signal

par

Olivier ADROT

Diagnostic à base de modèles incertains utilisant l'analyse par intervalles : l'approche bornante

soutenue publiquement le 4 décembre 2000 devant la commission d'examen :

É. WALTER	Rapporteur
S. GENTIL	Rapporteur
M. STAROSWIECKI	Examineur
D. SAUTER	Examineur
J. RAGOT	Directeur de thèse
D. MAQUIN	Codirecteur de thèse

Remerciements

Cette thèse a été réalisée au sein 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 que je tiens à remercier et je m'excuse auprès de ceux que j'aurais omis de citer.

Je remercie Monsieur le Professeur José Ragot, responsable du thème "Surveillance et détection de défauts" de m'avoir encouragé à mener cette thèse et de m'avoir accueilli dans son équipe pendant ces trois dernières années. Ses conseils, sa disponibilité et sa confiance m'ont été d'un grand secours tout au long de ce travail.

J'adresse également mes vifs remerciements à Didier Maquin, Maître de Conférences, pour m'avoir fait bénéficier de sa rigueur scientifique, de ses critiques objectives et de ses conseils avisés. Son soutien, sa confiance et l'intérêt qu'il a porté à mes travaux m'ont grandement aidé.

Je remercie Stéphane Ploix pour sa collaboration et ses conseils déterminants dans le choix de mon sujet de recherche.

Je remercie mes camarades de laboratoires, et tout spécialement Komi Gasso, pour l'ambiance conviviale qu'ils ont contribué à entretenir, les bons moments passés en leur compagnie, ainsi que Marjorie Schwartz pour son aide précieuse et son altruisme.

Je tiens à exprimer toute ma gratitude aux rapporteurs et aux examinateurs de mon mémoire pour avoir accepté d'être des membres de mon jury. Leurs appréciations et questions m'ont permis d'acquérir un meilleur recul sur le travail effectué.

Enfin, tous mes remerciements vont à mes parents et à toute ma famille, pour leurs encouragements et leur soutien.

Table des matières

INTRODUCTION.....5**CHAPITRE I : TERMINOLOGIE ET PRINCIPES DU DIAGNOSTIC11**

I.1 INTRODUCTION	13
I.2 TERMINOLOGIE PROPRE AU DIAGNOSTIC.....	15
I.2.1 Principaux termes utilisés en diagnostic	15
I.2.2 Les différents types de défaut.....	20
I.2.3 Surveillance et diagnostic	22
I.3 LES PRINCIPALES STRATEGIES DE DIAGNOSTIC	26
I.3.1 Maintenance programmée et maintenance selon l'état.....	26
I.3.2 Principes du diagnostic à base de modèles analytiques.....	30
I.3.3 Problématique du diagnostic à base de modèles analytiques	33
I.4 MÉTHODES DE DIAGNOSTIC À BASE DE MODÈLES ANALYTIQUES	35
I.4.1 Structure d'une procédure de diagnostic	35
I.4.2 Méthodes développées en diagnostic à base de modèles analytiques	38
I.4.3 Diagnostic et approche bornante	44
I.5 CONCLUSION	48

CHAPITRE II : ANALYSE ENSEMBLISTE49

II.1 INTRODUCTION.....	51
II.2 ARITHMÉTIQUE DES INTERVALLES	54
II.2.1 Motivations.....	54
II.2.2 Définitions et notations	55
II.2.3 Opérations ensemblistes.....	57
II.2.4 Dépendance entre variables bornées	60
II.2.5 Propriétés algébriques de l'arithmétique des intervalles	62
II.2.6 Fonction intervalle	64
II.3 ÉVALUATION D'UNE FONCTION D'INCLUSION.....	70
II.3.1 Méthodes reposant sur des formes centrées.....	71
II.3.2 Procédures de réduction et d'élimination	78
II.4 CHAMP DE VECTEURS DE FONCTIONS D'INCLUSION	84
II.4.1 Problème de dépendance et notion d'espace abstrait.....	84
II.4.2 Propriétés des espaces abstraits	86
II.4.3 Évaluation d'un champ de vecteurs de fonctions d'inclusion.....	89
II.4.4 Effet d'enveloppement	92
II.5 CONCLUSION	96

CHAPITRE III : GÉNÉRATION DE RÉSIDUS97

III.1 INTRODUCTION	100
III.2 ESPACE DE PARITÉ DANS LE CAS CERTAIN	101
III.2.1 Principe général.....	101
III.2.2 Taille de fenêtre et notion d'observabilité	102
III.3 FORMES STATIQUES DANS LE CAS INCERTAIN	106
III.3.1 Modèle.....	106
III.3.2 Première forme statique.....	107
III.3.3 Seconde forme statique	109

III.4 MINIMISATION DES TERMES RÉSIDUELS	110
III.4.1 Critère d'optimisation.....	110
III.4.2 Expressions générales des formes de calcul et d'évaluation.....	112
III.5 MATRICE DE COMMANDE INCERTAINE	113
III.5.1 Structure classique.....	114
III.5.2 Structure de modèle de capteurs.....	118
III.5.3 Récapitulatif.....	122
III.6 MATRICE D'ÉTAT INCERTAINE	123
III.6.1 Choix de la forme statique	123
III.6.2 Élimination des termes résiduels	125
III.6.3 Récapitulatif.....	127
III.7 MATRICE D'OBSERVATION INCERTAINE	128
III.7.1 Structure classique.....	128
III.7.2 Structure de modèle de capteurs.....	137
III.7.3 Récapitulatif.....	141
III.8 CAS DE PLUSIEURS MATRICES INCERTAINES.....	142
III.9 CAS D'INCERTITUDES INVARIANTES.....	143
III.9.1 Formes statiques	143
III.9.2 Minimisation des termes résiduels.....	144
III.9.3 Matrice de commande incertaine.....	146
III.9.4 Matrice d'état incertaine	147
III.9.5 Matrice d'observation incertaine	148
III.9.6 Influence de la nature temporelle des incertitudes	153
III.10 MATRICE DE PARITÉ INCERTAINE	154
III.10.1 Cas statique.....	154
III.10.2 Extension dans le cas dynamique	158
III.11 CONCLUSION.....	160

CHAPITRE IV : TESTS DE COHÉRENCE.....161

IV.1 INTRODUCTION	163
IV.2 TESTS DE COHÉRENCE DANS LE CONTEXTE D'UN MODÈLE INCERTAIN	164
IV.2.1 Principe : approches directe et indirecte.....	164
IV.2.2 Garantie, précision et justesse	167
IV.3 MÉTHODE DE PAVAGE	169
IV.3.1 Méthode de pavage assurant la construction complète d'un espace abstrait.	170
IV.3.2 Méthode de pavage adaptée à une procédure de diagnostic.	175
IV.4 EXPRESSION ANALYTIQUE DES FRONTIÈRES D'UN ESPACE ABSTRAIT	179
IV.5 MÉTHODE PROCÉDANT DE LA LINÉARISATION DES RELATIONS DE PARITÉ	186
IV.5.1 Procédure de linéarisation.....	186
IV.5.2 Procédure de réduction.....	189
IV.5.3 Génération des contraintes liées à un parallélotope	195
IV.5.4 Notion de distance.....	203
IV.5.5 Extension en vue d'une procédure de localisation	206
IV.6 CONCLUSION	210

CHAPITRE V : CARACTÉRISATION DE VARIABLES BORNÉES.....213

V.1 INTRODUCTION	215
V.2 POSITION DU PROBLÈME.....	216
V.2.1 Principe de l'approche géométrique	216

V.2.2 Première étape.....	218
V.2.3 Seconde étape	220
V.3 CARACTÉRISATION.....	226
V.3.1 Critère de précision.....	226
V.3.2 Forme prédéterminée	228
V.3.3 Forme indéterminée.....	229
V.3.4 Optimisation simultanée du centre et de la borne de l'erreur d'équation	230
V.4 MULTI-CARACTÉRISATION	234
V.4.1 Problème de dépendance.....	234
V.4.2 Solution proposée	238
V.5 CONCLUSION.....	243
CHAPITRE VI : APPLICATION.....	245
VI.1 INTRODUCTION	247
VI.2 MODÈLE DE SIMULATION	248
VI.3 GÉNÉRATION DES RELATIONS DE PARITÉ	253
VI.4 TESTS DE COHÉRENCE.....	255
VI.4.1 Méthodes de pavage	255
VI.4.2 Procédure de linéarisation.....	257
VI.4.3 Procédure de réduction.....	258
VI.4.4 Procédure de détection de défauts	259
VI.4.5 Procédures de localisation de défauts	262
VI.5 CONCLUSION	267
CONCLUSION.....	269
ANNEXES	273
ANNEXE A : FORMES CENTRÉES SUPPLÉMENTAIRES	275
ANNEXE B : RÉOLUTION DU PROBLÈME D'OPTIMISATION DU CHAPITRE III	278
ANNEXE C : EXPRESSION ANALYTIQUE DE LA MATRICE S	279
ANNEXE D : MATRICE DE PARITÉ W POUR UN MODÈLE CERTAIN.....	281
ANNEXE E : RELATION D'ÉQUIVALENCE ENTRE LES DEUX FORMES STATIQUES DANS LE CAS D'UN MODÈLE CERTAIN	283
ANNEXE F : RANG ET VALEURS PROPRES D'UNE MATRICE INCERTAINE	285
ANNEXE G : ÉTUDE D'UN MODÈLE STATIQUE INCERTAIN	289
ANNEXE H : GÉNÉRATION DES CONTRAINTES INÉGALITÉS D'UN PARALLÉLOTOPE : CAS PARTICULIERS	292
ANNEXE I : CALCUL DES SOMMETS D'UN POLYTOPE.....	295
ANNEXE J : CALCUL DU VOLUME D'UN POLYTOPE.....	301
RÉFÉRENCES BIBLIOGRAPHIQUES.....	305

Introduction

Dans un souci d'améliorer la compétitivité des outils de production, le diagnostic, en assurant la surveillance des systèmes physiques, suscite depuis une trentaine d'années un intérêt croissant tant au niveau du monde industriel que de la recherche scientifique. Parce que les systèmes industriels deviennent de plus en plus complexes et sophistiqués, il est devenu légitime de leur associer un module efficace de surveillance afin d'accroître leur fiabilité et leur disponibilité, et d'améliorer la sécurité des personnels. Le diagnostic est ainsi devenu progressivement un thème de recherche à part entière de la discipline automatique. Puisqu'une anomalie peut provenir aussi bien d'une dégradation de l'installation surveillée, que d'une réaction peu habituelle mais normale de celle-ci, une procédure de diagnostic doit, dans un premier temps, déterminer si le système physique considéré se comporte anormalement. Dans cette éventualité, elle a ensuite pour objectif de rechercher les composants en défaut, voire certaines caractéristiques comme l'instant d'apparition de l'anomalie ou son niveau de gravité.

Dans le domaine de l'automatique, les méthodes de diagnostic reposent essentiellement sur la connaissance d'un modèle censé représenter au mieux le comportement du système physique à surveiller. Elles s'appuient sur la redondance analytique générée par cette représentation mathématique et procèdent de tests de cohérence vérifiant l'adéquation entre les mesures du système physique étudié et l'information contenue dans son modèle. En général, une procédure orientée bon fonctionnement reposant sur la génération de variables d'écart appelées résidus est alors réalisée : une incohérence comportementale (défaut) ou fonctionnelle (défaillance) est révélée lorsqu'un résidu sort du voisinage de zéro (approche par estimation d'état regroupant relations de parité, observateurs d'état, filtres) ou que certains paramètres identifiés dérivent anormalement (approche par estimation paramétrique).

Le problème majeur rencontré lors d'une procédure de diagnostic réside dans le fait qu'un modèle ne définit qu'un comportement approché du système physique qu'il est censé représenter. Ces approximations proviennent des erreurs de modélisation qui surviennent lorsqu'un modèle est linéarisé (discrétisation d'un système d'équations continues, linéarisation autour d'un point de fonctionnement d'un système non-linéaire), lorsque certains phénomènes physiques ne sont pas pris en compte afin de simplifier les relations (influence de la température sur la résistance d'un enroulement inductif par exemple) ou tout simplement lorsque la recherche d'une précision infinitésimale n'est industriellement pas rentable ou impossible à garantir (tolérances sur les capteurs d'instrumentation). L'utilisation de lois de la physique pour décrire les relations entre phénomènes ne permet pas non-plus de s'affranchir d'approximations puisqu'elle introduit bon nombre d'hypothèses simplificatrices, souvent difficiles à satisfaire. De plus, l'utilisation de modèles de connaissance très précis dans la description des phénomènes mis en jeu, conduit à des systèmes d'équations trop complexes pour être exploitables en vue d'un diagnostic temps réel et souvent surparamétrés, posant alors des problèmes insolubles d'identification.

En réalité, il est appaît que, dans une discipline telle que le diagnostic, le concept de robustesse est bien plus complexe que dans le cadre de la commande des systèmes. En effet,

une loi de commande est robuste si, malgré certaines erreurs de modélisation, les spécifications de consigne sont satisfaites. Or, le problème du diagnostic d'une installation procède de la mise en évidence d'anomalies a priori indéterminées tant dans leurs instants d'apparition que dans leurs amplitudes et parfois même, dans leurs natures. Par conséquent, la notion de robustesse est subordonnée à la faculté de discriminer les défauts ou défaillances des indéterminations normales dues aux erreurs de modélisation.

Un grand nombre de travaux a déjà été publié sur ce thème. Certaines techniques consistent à générer des résidus dits robustes, c'est-à-dire rendus insensibles aux incertitudes par découplage. Ces dernières sont considérées comme des perturbations agissant sur le modèle dynamique du système physique, dont les matrices de distribution sont connues (observateurs à entrées inconnues ou construits par la méthode "eigenstructure assignment"). L'inconvénient, dans ces conditions, est que le diagnostic n'est réalisé que sur la partie parfaitement connue du système physique, sans prendre en compte l'information liée aux paramètres incertains rejetés. D'autres techniques consistent à décrire les incertitudes de manière stochastique et à les représenter, dans une même relation, par un terme additif unique sous la forme d'un bruit ayant une certaine densité de probabilité. En pratique, l'hypothèse de blancheur requise n'est pas réaliste, notamment dans le cas de modèles très réducteurs introduisant d'importantes erreurs structurelles. De plus, émettre des hypothèses statistiques sur les paramètres eux-mêmes rend alors le problème insoluble à cause d'opérations pour lesquelles les densités de probabilité ne sont pas stables.

L'objectif de ce travail de recherche, s'appuyant sur des travaux effectués en collaboration avec le maître de conférence Ploix S., consiste à définir une nouvelle stratégie de diagnostic permettant la prise en compte à la fois de défauts intervenant au sein du système physique (sans se restreindre aux défauts d'instrumentation), ainsi que d'incertitudes pouvant fluctuer dans le temps et de nature multiplicative, c'est-à-dire affectant directement les paramètres du modèle étudié (non plus limitées à de simples termes additifs modélisant une erreur d'équation). La méthode de diagnostic proposée, appelée approche bornante, est une stratégie ensembliste reposant sur l'analyse par intervalles. Elle tient compte de manière globale de la structure et de l'amplitude des incertitudes, permettant ainsi de conserver plus d'information que les techniques fonctionnant par découplage. Chaque paramètre incertain est ainsi représenté par une variable bornée, outil principalement utilisé dans les sciences informatiques (évaluation des erreurs d'arrondis, d'intersection de surfaces,...), mais très peu répandu en automatique (caractérisation de modèles linéaires à erreurs d'équation bornées pour le thème identification), ce qui en fait un travail d'autant plus original dans le domaine du diagnostic.

- Chapitre I

L'objectif du premier chapitre se résume à définir précisément ce que sous-entend le terme diagnostic, aussi bien du point de vue de la terminologie, que des principes fondamentaux sur lesquels il repose. Si ce chapitre peut sembler classique dans la mesure où son contenu se retrouve, en partie, dans bon nombre de thèses se rapportant au diagnostic de

systèmes, il permet néanmoins de préciser le contexte dans lequel ce travail a été réalisé et de justifier le choix de l'approche bornante en la situant par rapport aux méthodes plus traditionnelles. En reformulant les principes du diagnostic à base de redondance analytique, il s'avère que le problème majeur rencontré est lié à l'idée même de modèle mathématique. Censé représenter fidèlement le comportement du système physique considéré, il n'en donne en réalité qu'une image approximative. Ce constat conduit donc à l'idée de prendre en compte les erreurs de modélisation en travaillant sur des modèles incertains, plutôt que déterministes. En constatant les limites d'une modélisation faisant intervenir des variables stochastiques, nous justifierons notre choix d'adopter une approche ensembliste.

- Chapitre II

Le second chapitre présente l'outil analyse par intervalles dans un contexte bien plus général que le domaine du diagnostic proprement dit, en s'attachant à en expliciter les fondements et propriétés, ainsi que les avantages et limitations. L'analyse par intervalles permet de prendre en compte très simplement la notion d'incertitude, puisque c'est ce pourquoi elle fut initialement destinée en vue d'étudier les sources d'imprécision inhérentes à l'outil informatique. Les intervalles, en tant qu'ensembles de valeurs, permettent de plus d'explorer plus rapidement un domaine d'étude qu'une méthode ponctuelle. En revanche, l'analyse par intervalles doit faire face au problème fondamental de dépendance entre variables bornées, pouvant conduire à des majorations importantes des bornes calculées lors de l'évaluation de fonctions intervalle. Après avoir présenté un certain nombre de techniques usuelles pour combattre cette difficulté (formes centrées, méthodes d'élimination ou de réduction, problème d'optimisation globale ensembliste), nous nous intéresserons plus spécifiquement au cas des champs de vecteurs de fonctions intervalle où nous montrerons que le problème de dépendance intervient aussi entre composantes. Un formalisme, légèrement différent de ce qui se rencontre traditionnellement, est alors utilisé en vue de ne pas se limiter à l'étude de simples pavés. Nous insisterons particulièrement sur la propriété de garantie, qui confère à l'approche bornante un intérêt tout particulier.

- Chapitre III

Un inconvénient majeur de l'utilisation de l'analyse par intervalles étant son caractère explosif (phénomène d'enveloppement) dans le cas de systèmes récurrents, une approche par Espace de Parité a été privilégiée. Celle-ci permet de reformuler les équations dynamiques d'un modèle sous forme de relations algébriques. En empilant les mesures sur un horizon temporel préalablement choisi, une représentation statique est construite, où il n'est plus nécessaire d'intégrer les équations du modèle pour obtenir des résidus. Si cette stratégie est couramment rencontrée dans le cas d'un modèle certain, elle l'est beaucoup moins dans un contexte incertain, et ce d'autant plus que des incertitudes multiplicatives sont prises en compte. L'objectif de ce chapitre consiste alors à rechercher une matrice de projection, appelée matrice de parité, permettant d'éliminer des résidus les variables d'état inconnues du système. Nous proposons alors deux stratégies. La première use d'une procédure d'optimisation numérique cherchant à déterminer une matrice de parité certaine assurant le découplage des relations de parité vis-à-vis de l'état du système. Cependant, cette solution

n'est pas systématique dans la mesure où les conditions d'orthogonalité à respecter pour assurer le découplage peuvent se révéler trop contraignantes. Nous étudions donc différents cas élémentaires dépendant des matrices de la représentation d'état qui sont incertaines, ainsi que de la structure (additive ou de modèle de capteurs) et de la nature temporelle (invariance ou non) des incertitudes considérées. Pour chaque cas traité, l'influence de l'horizon temporel sur la possibilité d'un découplage total est détaillée. Dans certaines situations, l'utilisation d'une forme statique, appelée seconde forme statique, différente de celle traditionnellement employée, permet d'éviter certaines impasses, notamment quant à l'évaluation du critère d'optimisation sur lequel repose la méthode proposée. Une étude structurelle des résidus obtenus pour les deux formes statiques détaille dans quels cas l'une est plus avantageuse à utiliser. La seconde stratégie, consistant à rechercher une matrice de parité dépendant des incertitudes, permet de contourner l'écueil lié à la nécessité de découpler les résidus des variables d'état, mais cette fois, la solution doit être trouvée sous forme symbolique et non-plus numérique.

- Chapitre IV

Le quatrième chapitre s'attache à développer des tests de cohérence capables de traiter les résidus précédents. En effet, ces derniers définissent un ensemble de comportements cohérents avec le modèle de référence ; tester si un résidu est proche ou non de zéro n'est donc plus suffisant. L'objectif consiste alors à construire, à l'aide de l'analyse par intervalles, des enveloppes des résidus tenant compte des plages de variations des paramètres incertains et qui permettent de définir les frontières du domaine correspondant à un comportement normal. Si un résidu sort de son enveloppe, le défaut est détecté et une procédure de localisation est engagée pour en trouver l'origine. Cette approche est comparable aux méthodes stochastiques plus classiques où les seuils de détection sont construits en fonction des taux de non-détection et de fausses alarmes imposés et des caractéristiques statistiques (moyenne, écart-type) des signaux porteurs d'information. Cependant, le formalisme des variables bornées permet de prendre en compte, directement à l'intérieur du modèle, les incertitudes portant sur des paramètres physiques, ce qui laisse entrevoir une manière plus naturelle de concevoir les seuils de détection. De plus, étudier indépendamment chaque composante du vecteur résidu revient à négliger des dépendances entre variables bornées. La précision de la procédure de diagnostic s'en trouve alors appauvrie puisqu'une partie de l'information disponible n'est pas utilisée. L'objectif principal de ce quatrième chapitre est donc consacré à la recherche de méthodes permettant de traiter de manière globale tous les résidus en même temps, c'est-à-dire de construire une sorte d'enveloppe multi-dimensionnelle dont la forme est ajustée à chaque instant. Compte tenu de la complexité des relations de parité généralement non-linéaires en les incertitudes, cette enveloppe ne peut être évaluée exactement et une majoration en est donc recherchée. Une première solution repose sur la notion de pavage qui consiste à décomposer un problème en sous-problèmes élémentaires plus faciles à résoudre. Le domaine construit se présente alors sous la forme d'une réunion de pavés image obtenue par découpage de l'espace paramétrique. La majoration est d'autant plus faible, que ce découpage est fin. Néanmoins, afin de réduire les temps de calcul, une procédure ne subdivisant que les pavés jugés intéressants pour les objectifs de diagnostic est développée.

Une autre possibilité reposant sur la construction de polytopes convexes est proposée. Elle procède par linéarisation des relations de parité en les variables bornées en remplaçant les monômes d'ordre strictement supérieur à un par de nouvelles variables indépendantes de bornes adéquates ; l'approximation est alors due à la non prise en compte de certaines dépendances. Elle se présente sous la forme d'une série de procédures assurant respectivement la linéarisation des résidus, la réduction du nombre de variables bornées, puis la génération des contraintes inégalité représentatives du paralléloèdre solution permettant une mise en œuvre simple des tests de cohérence. La notion de distance est alors abordée pour permettre de nuancer les résultats dichotomiques des précédents tests. Pour terminer, une extension à la localisation de défauts est présentée. Les avantages et inconvénients de ces deux stratégies sont détaillés, sachant qu'elles permettent toutes deux de traiter le cas de relations de parité non-linéaires en les variables bornées.

- Chapitre V

Le dernier point de recherche concerne des méthodes de caractérisation permettant la détermination des bornes des incertitudes paramétriques d'un modèle. Dans la mesure où l'approche proposée repose sur des variables bornées, il est nécessaire de pouvoir les caractériser, c'est-à-dire de déterminer le nombre de paramètres incertains, la structure du modèle à retenir et l'amplitude des incertitudes associées. Nous nous limiterons à des modèles à horizon fini linéaires en les variables bornées. L'erreur d'équation est alors supposée appartenir à un intervalle, alors que les paramètres fluctuent à l'intérieur d'un domaine borné invariant dans le temps qu'il faut déterminer. L'idée est de rechercher l'amplitude des incertitudes à ajouter sur une valeur centrale choisie des paramètres, de manière à ce que ceux-ci fluctuant, le modèle puisse expliquer l'ensemble des mesures. En présentant la méthode d'un point de vue géométrique, nous insisterons notamment sur la différence entre nos objectifs et la recherche plus classique du domaine appelé dans la littérature *Feasible Parameter Set*. Un critère est proposé de manière à refléter la précision du domaine obtenu et afin d'optimiser le choix de la valeur centrale et des bornes de l'erreur d'équation. Deux structures de modèle sont étudiées : une forme prédéterminée, ne dépendant que d'un seul paramètre dimensionnel, fixe la forme du domaine représentatif des incertitudes paramétriques, alors que dans le cas d'une forme indéterminée, le nombre de degrés de liberté est plus important, l'algorithme imposant automatiquement sa forme en réglant le poids de chaque incertitude. Une extension est ensuite proposée pour des modèles multi-sortie où il est nécessaire de prendre en compte les dépendances dues aux variables bornées intervenant dans plusieurs équations entrée-sortie à la fois.

- Chapitre VI

Le sixième et dernier chapitre illustre les divers résultats théoriques développés dans les chapitres précédents au travers d'un exemple académique. Reposant sur le modèle statique d'une machine à courant continu, il détaille la mise en œuvre d'une procédure complète de diagnostic en dégagant les avantages et inconvénients de chaque stratégie proposée.

Ce mémoire s'achève par une conclusion générale réalisant le bilan de ce travail et évoquant quelques perspectives de recherches complémentaires.

Chapitre I

Terminologie et principes du diagnostic

TABLE DES MATIÈRES

I.1 INTRODUCTION.....	13
I.2 TERMINOLOGIE PROPRE AU DIAGNOSTIC	15
I.2.1 Principaux termes utilisés en diagnostic	15
I.2.2 Les différents types de défaut	20
I.2.3 Surveillance et diagnostic	22
I.3 LES PRINCIPALES STRATEGIES DE DIAGNOSTIC	26
I.3.1 Maintenance programmée et maintenance selon l'état	26
I.3.2 Principes du diagnostic à base de modèles analytiques	30
I.3.3 Problématique du diagnostic à base de modèles analytiques.....	33
I.4 MÉTHODES DE DIAGNOSTIC À BASE DE MODÈLES ANALYTIQUES.....	35
I.4.1 Structure d'une procédure de diagnostic.....	35
I.4.2 Méthodes développées en diagnostic à base de modèles analytiques.....	38
I.4.2.1 Estimation paramétrique	38
I.4.2.2 Estimation d'état	38
I.4.3 Diagnostic et approche bornante.....	44
I.5 CONCLUSION	48

I.1 INTRODUCTION

Répondant à un besoin grandissant de commander des systèmes physiques en limitant l'intervention humaine, l'automatique a pour objectif l'amélioration des performances que ce soit en termes de productivité, de rentabilité ou de qualité des produits. C'est dans ce contexte très concurrentiel que fut ressentie la nécessité de surveiller l'outil de production en vue d'une compétitivité accrue. Initialement considéré comme un thème secondaire de la vaste discipline qu'est l'automatique, le diagnostic des systèmes suscite, depuis une trentaine d'années, un intérêt croissant tant au niveau du monde industriel que de la recherche scientifique.

- Diagnostic et monde industriel

A l'origine, le diagnostic se limitait aux applications industrielles à haut niveau de risque pour la communauté comme le nucléaire ou l'aéronautique (Potter et al, 1977), (Daly et al, 1979), (Desai et al, 1979), ainsi qu'aux secteurs d'activité de pointe tels que l'industrie de l'armement ou l'aérospatial (Desai et al, 1976), (Deckert et al, 1977). D'une croissante complexité, les systèmes industriels nécessitent aujourd'hui des algorithmes d'asservissement et de régulation de plus en plus performants et par conséquent sophistiqués. Il est alors légitime pour les entreprises d'acquiescer un système efficace de surveillance afin d'améliorer la sécurité des personnels et d'assurer une fiabilité et une disponibilité accrues de leur outil de production. En effet, vu l'effort financier consenti à l'acquisition d'une chaîne de fabrication performante, un dysfonctionnement réduisant momentanément la production à néant est préjudiciable. En vue de protéger les biens et les personnes, ce concept s'applique naturellement aux produits eux-mêmes. Dernièrement, de nouvelles perspectives ont vu le jour, comme dans le domaine de l'environnement par exemple, où la prévention des risques naturels et la lutte contre la pollution sont réglementées. Il s'agit alors de vérifier l'état de fonctionnement de la chaîne d'instrumentation afin de garantir le traitement des données qui en découle (Mandel, 1998), (Giuliani, 1998).

De manière générale, la conduite et la surveillance d'un système physique dépendent des informations délivrées par des capteurs. La défaillance de l'un d'entre eux peut engendrer le dysfonctionnement de l'installation en conduisant à une altération de ses performances. Une procédure de surveillance doit donc être apte à traiter un défaut affectant aussi bien le système physique proprement dit, que l'un de ses organes de conduite (actionneur, capteur d'instrumentation). Des techniques sophistiquées capables de différencier le système physique de ses différents organes de conduite doivent être élaborées. Cependant, la complexité et la taille de l'installation, ainsi que la gravité des risques encourus et des conséquences potentielles augmentent alors la quantité d'information à analyser, rendant la surveillance bien trop complexe à l'échelle humaine. Il s'avère par conséquent indispensable d'adjoindre à l'opérateur une aide à la décision, voire de rendre la surveillance automatique, tout comme

cela s'était déjà produit dans le cadre de la commande des systèmes. L'intérêt est aussi de permettre l'amélioration de la disponibilité des installations en remplaçant les politiques de maintenance programmée par des stratégies de maintenance conditionnelle prenant en compte l'état effectif du système physique.

- *Diagnostic et recherche scientifique*

Les premiers travaux concernant le thème diagnostic datent du début des années 1970, résumés notamment dans l'article de synthèse (Willsky, 1976). En raison de l'intérêt croissant suscité dans le monde industriel, le diagnostic est devenu peu à peu un thème de recherche à part entière, permettant aux laboratoires d'automatique, aussi bien de se diversifier en se démarquant du thème commande largement exploité, que de s'ouvrir à de nouvelles opportunités d'études industrielles.

Le développement des méthodes de diagnostic est dû, d'une part, aux progrès obtenus dans les divers domaines de l'automatique des systèmes tels que la modélisation, l'estimation d'état et l'estimation paramétrique, d'autre part, à l'augmentation de la puissance des unités de calculs. La vigueur de ce thème est d'autant plus marquée que bon nombre de disciplines contribuent à son expansion telles que, sans être exhaustif, le traitement du signal, le domaine informatique (intelligence artificielle), le génie électrique, le génie chimique,...

Il s'ensuit un accroissement des moyens de diffusion de résultats. Ainsi, de nombreux congrès traitant de la discipline automatique dans son ensemble proposent régulièrement des sessions se rapportant plus spécifiquement au thème diagnostic : *IEEE Conference on Decision and Control (CDC)*, *European Control Conference (ECC)*, *American Control Conference (ACC)*, *IFAC World Congress*, *Conférence Internationale Francophone d'Automatique (CIFA)*,..., pour ne citer qu'eux. D'autres encore sont entièrement dédiés au thème diagnostic : *Workshop on Fault Diagnosis, Reliability and Related Knowledge-based Approaches*, *IFAC Symposium on Fault Detection Supervision and Safety for Technical Process (SafeProcess)*, *IFAC Symposium on On-Line Fault Detection and Supervision in the Chemical Process Industries*, *ToolDiag – Conference on Fault Diagnosis*,... Ainsi, des chercheurs faisant autorité dans le domaine du diagnostic publient régulièrement dans la littérature internationale des articles de synthèse répertoriant les différentes méthodes développées et les liens qui les unissent (Gertler, 1986, 1988, 1991), (Frank, 1987, 1990, 1996), (Patton, 1994, 1997), (Patton et al, 1991), (Isermann, 1984, 1994), (Isermann et al, 1997). De plus, plusieurs ouvrages traitant de ce thème ont été écrits, tels que (Patton et al, 1989, 2000), (Brunet et al, 1990), (Basseville et al, 1993), (Ragot et al, 1990), (Dubuisson, 1990), (Zwingelstein, 1995), (Maquin et al, 2000b).

En parallèle, des structures dans lesquelles j'ai plusieurs fois été impliqué, ont été mises en place de manière à rassembler les chercheurs ayant pour centre d'intérêt le thème diagnostic, ainsi qu'à faciliter les échanges et la collaboration pluridisciplinaire entre équipes de recherche. Le Groupe Thématique (G.T.) Sécurité, Surveillance, Supervision (S³) soutenu par le Groupe de Recherche (G.D.R.) Automatique en est le parfait exemple sur le plan national. Regroupant la majorité des laboratoires français intéressés, des réunions

trimestrielles sont organisées, permettant tout à chacun d'exposer ses propres travaux et suscitant bon nombre de discussions bénéfiques (Ploix et al, 1999a). Les présentations assurées par des industriels permettent en outre de mieux appréhender leurs attentes. Il en est de même pour le Projet de Recherche Concerté (P.R.C.) intitulé Diagnostic de la Machine Asynchrone regroupant à la fois automaticiens, traiteurs de signaux et électrotechniciens, et dont le thème correspondait à mon axe de recherche en D.E.A. (Adrot, 1996a, 1996b), (Ploix et al, 1996, 1997c). Le bilan des 12 premières années de fonctionnement fut présenté au Colloque de Recherche Doctorale Automatique, Génie informatique, Image et Signal (AGIS'97) (Ploix et al, 1997a, 1997b). Ces structures peuvent aussi se développer sur le plan international comme le groupe de recherche établi dans le cadre du programme Human Capital Mobility (H.C.M.) regroupant des équipes françaises, allemandes et anglaises (Adrot et al, 1998). Pour terminer ce tour d'horizon, notons qu'une école d'été a été spécialement dédiée à la surveillance des systèmes en 1996 : *École d'Été d'Automatique de Grenoble : Continuous System Supervision* (Gentil et al, 1996).

L'objectif de ce chapitre est de rappeler brièvement la terminologie et les principes fondamentaux propres au diagnostic. Si cette partie peut sembler superflue puisqu'elle ne dévoile rien de nouveau et que son contenu se retrouve, dans une certaine mesure, dans la plupart des thèses et ouvrages se rapportant au diagnostic des systèmes, elle permet néanmoins, d'une part, de préciser le vocabulaire utilisé tout au long de ce document, d'autre part, d'explicitier le contexte dans lequel ce travail a été réalisé et de justifier l'approche choisie en la situant par rapport aux méthodes plus traditionnelles.

Au cours de ce chapitre introductif, nous rappellerons dans un premier temps la terminologie et les différents objectifs propres au diagnostic des systèmes, ainsi que sa place au sein d'une stratégie de surveillance. Les principales approches d'une procédure de diagnostic seront explicitées dans la troisième section. Celle qui a retenu notre attention repose sur des modèles analytiques du système physique et sera plus largement détaillée. Les notions de redondance analytique et de modèle seront développées, conduisant au principe fondamental sur lequel repose toute procédure de diagnostic utilisant des modèles : le test de cohérence. La problématique inhérente à ce type d'approche sera alors mise en évidence, engendrant le concept de l'incertain. La quatrième section s'attachera à rappeler l'évolution historique des différentes méthodes employées pour concevoir une procédure de diagnostic à base de modèles analytiques, alors que pour conclure, nous nous appuierons sur ce qui précède afin de justifier le choix d'une approche ensembliste.

I.2 TERMINOLOGIE PROPRE AU DIAGNOSTIC

I.2.1 Principaux termes utilisés en diagnostic

Il semble intéressant, dans un premier temps, de rappeler les principaux termes utilisés en diagnostic des systèmes. Reposant principalement sur le travail effectué par Milne, (1987),

Isermann et al (1997, 2000), Giuliani (1998) et Ploix (1998), ainsi que sur l'étude d'ouvrages synthétiques tels que (Brunet et al, 1990), (Zwingelstein, 1995), un nombre non exhaustif de définitions sont proposées afin de clarifier la suite de ce document, sachant qu'il est courant de trouver dans la littérature internationale des définitions différentes d'une même notion. Pour les principaux termes énoncés, les pendants anglo-saxons sont donnés entre parenthèses.

- *Système (System)*

Un système est un ensemble structuré d'éléments (composants, constituants) interconnectés ou en interaction. La notion de système est très générale puisqu'elle s'adresse aussi bien à des ensembles physiques qu'à des éléments plus abstraits comme des équations mathématiques. Dans le cadre du diagnostic, les termes système et système physique désignent tout particulièrement l'installation à surveiller. Suivant les auteurs, le terme système soit désigne l'installation complète, soit se restreint à l'ensemble des éléments commandés, c'est-à-dire l'installation privée de ses organes de commande, de ses capteurs et de ses actionneurs, auquel cas, pour plus de concision, nous parlerons de système commandé.

- *Modèle (Model)*

Il est possible de se munir d'une description mentale du système physique pour aboutir à un modèle. Un modèle d'un système physique est donc une description de sa structure et une représentation comportementale ou fonctionnelle de chacun de ses composants (Milne, 1987). Une représentation comportementale est constituée de relations entre divers phénomènes du processus, appelées classiquement relations de causes à effets. Une représentation fonctionnelle, constituée de relations de même nature que le niveau comportemental, est plus abstraite puisqu'elle ne s'adresse qu'aux fonctions présumées que le système physique doit remplir (qui pourraient être définies par un cahier des charges comme un temps de réponse ou un rendement à satisfaire). D'où le concept de raisonnement téléologique exprimé par Milne à la figure I-1 exprimant l'idée que seule la finalité du système (c'est-à-dire ce pourquoi il a été conçu) nous intéresse. Le niveau structurel, quant à lui, s'appuie sur la structure réelle du système physique et évoque les interconnexions entre ses différents éléments ou constituants.

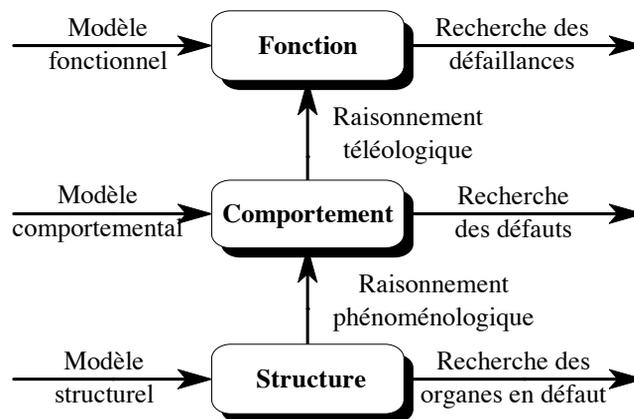


Figure I-1. Niveaux de modélisation et objectifs de diagnostic selon (Milne, 1987)

Ces trois niveaux¹ de représentation permettent de préciser ce qu'il est possible de faire en diagnostic avec un modèle donné. Les niveaux comportemental et fonctionnel comprennent des relations entre des grandeurs physiques et permettent de mettre en évidence la présence d'un événement anormal ou anomalie. Le niveau structurel, quant à lui, permet de déterminer l'élément affecté par le défaut. L'intérêt de cette décomposition est de rappeler que, puisqu'un modèle contient toute l'information relative à un système physique et utilisable ensuite par la procédure de diagnostic, son choix doit être en adéquation avec le cahier des charges fixé.

- *Défaut (Fault)*

Un défaut est une anomalie comportementale au sein d'un système physique.

- *Défaillance (Failure)*

Une défaillance définit une anomalie fonctionnelle au sein d'un système physique, c'est-à-dire caractérise son incapacité à accomplir certaines fonctions qui lui sont assignées.

- *Panne (Break-down)*

Lorsque plus aucune fonction n'est assurée, le terme panne est employé à la place de défaillance.

Le mot défaut possède un sens plus général que son homologue défaillance puisque ce dernier est contraint à la notion de fonction, c'est-à-dire ce pourquoi le système physique a été conçu (sa finalité en définitive). Or une fonction provient d'une abstraction du comportement du système, c'est-à-dire la considération des uniques relations comportementales utiles à la description des objectifs qu'il se doit de satisfaire. Ainsi une défaillance est un défaut particulier se traduisant non seulement par une anomalie comportementale, mais aussi par le non-respect d'une ou plusieurs fonctions du système physique. En revanche, un défaut n'altère pas nécessairement son fonctionnement, mais peut présager d'une défaillance à venir. De même, une panne est une défaillance particulière pour laquelle le système n'assure plus aucune de ses fonctions.

- *Signe (Sign)*

Un défaut est perceptible au travers d'un signe, caractère manifestant un état comportemental anormal. Un signe est caractérisé par un ensemble d'observations en provenance d'un système physique et est révélateur de la présence d'un défaut.

- *Symptôme (Symptom)*

Le pendant de signe dans le cas d'une défaillance correspond au terme symptôme, définit comme suit : caractère distinctif d'un état fonctionnel anormal.

¹ Milne ajoute également un quatrième niveau de modélisation en sus du niveau fonctionnel pour appréhender les connaissances heuristiques susceptibles d'être utilisées dans une procédure de diagnostic. Ce type de stratégie étant alloué plus spécifiquement au domaine de l'intelligence artificielle, nous l'avons délibérément omis.

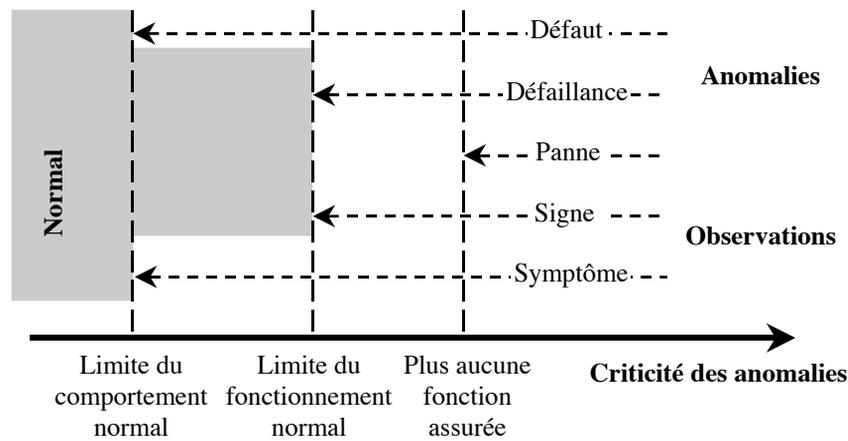


Figure I-2. Ordonnement des anomalies selon leur criticité

La distinction entre les termes symptôme et signe s'apparente donc à la différence entre les notions de défaut et de défaillance. Un symptôme révélateur d'une défaillance est nécessairement un signe, alors qu'un défaut se manifeste au travers de signes qui ne sont pas nécessairement des symptômes.

Dans la suite de ce document, nous nous placerons dans le cadre général de l'étude de défauts, sans préjuger des fonctions que le système étudié se doit d'assurer. Nous nous restreindrons donc aux termes défaut, modification de comportement et signe.

- *Perturbation (disturbance, perturbation)*

Une perturbation définit tout phénomène considéré comme normal influençant un système physique, mais non ou mal pris en compte dans le modèle censé le représenter. Les bruits font partie des perturbations, néanmoins, à cause de leurs propriétés spécifiques (blancheur, normalité), ils sont généralement pris en compte séparément dans le modèle. La distinction entre les deux termes anglo-saxons *disturbance*² et *perturbation*³ (Isermann et al, 1997, 2000) pour traduire la notion de perturbation est confuse.

Le terme erreur de modélisation est plus restrictif puisqu'il s'adresse uniquement à un phénomène non ou mal modélisé (volontairement ou non), mais qui aurait du être parfaitement décrit ; alors que le mot perturbation peut désigner des phénomènes liés à l'environnement avec lequel interagit le système sans pour autant qu'il ait été question de les modéliser.

- *Incertitude (uncertainty)*

Les incertitudes peuvent être structurelles (catégorie des multi-modèles (Gasso, 2000) obtenus par interpolation de plusieurs modèles locaux associés à des zones de comportements différents) ou paramétriques (modèles constitués de relations comportementales imprécises

² "An unknown (and uncontrolled) input acting on a system. "

³ "An input acting on a system which results in a temporary departure from current state. "

auquel cas ceux-ci peuvent être non-déterministes⁴). Ce mémoire ne s'intéresse qu'à la seconde catégorie d'incertitudes, les modèles étant alors qualifiés d'incertains. Une incertitude (ou de manière équivalente un paramètre incertain) est alors une variable dont la valeur ne peut être connue, utilisée en vue de formaliser le peu de connaissance disponible associée à une perturbation. Elle permet ainsi de prendre en considération la méconnaissance que l'on a d'un système physique en indiquant, dans la représentation associée, qu'à certains endroits subsistent des imprécisions.

Les incertitudes sont qualifiées de structurées lorsque les paramètres qui ne sont pas fiables sont localisés ; il est cependant impossible de définir exactement leurs valeurs, seules leurs plages de variation et/ou leurs lois de distribution étant accessibles. Une variable incertaine est donc issue d'une connaissance partielle du système et s'oppose à une variable déterministe, qui elle est parfaitement définie à chaque instant.

Exemple. Pour illustrer ceci, prenons l'exemple d'une loi de la physique connue de tous : la loi d'Ohm. Par définition, elle précise que le courant i_0 traversant une résistance de valeur nominale r_0 et la tension u_0 à ses bornes sont proportionnels : $u_0 = r_0 i_0$. De plus, si la chaîne d'instrumentation est supposée parfaite, alors les véritables valeurs i_0 et u_0 du courant et de la tension, appelées valeurs vraies, sont mesurées sans erreur. En notant i et u les mesures respectives de i_0 et u_0 , la relation précédente devient : $u = r_0 i$. Ce modèle est déterministe dans la mesure où à une valeur i du courant correspond une valeur u de la tension parfaitement déterminée pour r_0 fixée.

En pratique, à cause de l'impossibilité à produire des composants strictement identiques, les fabricants donnent la valeur r d'une résistance avec une certaine précision technologique σ ($\sigma=5\%$, 10% ,...), signifiant que r est comprise dans l'intervalle $[(1-\sigma)r_0, (1+\sigma)r_0]$. De plus, r dépend de la température du composant électrique. L'influence de cette dernière peut être prise en compte en complexifiant le modèle par l'intégration d'une équation supplémentaire faisant intervenir la résistivité de la résistance ; cependant, ce paramètre est généralement mal connu. En conclusion, la valeur r de la résistance est connue avec imprécision et n'a que peu de chance de valoir effectivement sa valeur nominale r_0 à cause des diverses perturbations précitées. Contrairement au modèle déterministe précédent, un modèle incertain permet d'appréhender les perturbations qui l'affectent : nous avons choisi de les représenter par une unique variable incertaine θ_1 :

$$u_0 = r i_0 = (1 + \theta_1) r_0 i_0. \quad (\text{I-1})$$

De plus, la chaîne d'instrumentation n'est pas parfaite : les véritables valeurs du courant i_0 et de la tension u_0 sont mesurées avec une certaine imprécision. Tout comme pour la résistance, la précision technologique d'un capteur, indiquée par le fabricant, dépend de sa qualité. L'information sur l'erreur de mesure est alors relative : plus importante est la valeur vraie, plus grande est l'imprécision. La modélisation choisie pour représenter le comportement normal des deux capteurs de courant et tension est la suivante :

⁴ Déterminisme : 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 doit se produire.

$$\begin{aligned} i &= (1 + \theta_2)i_0 + \theta_3 \\ u &= (1 + \theta_4)u_0 + \theta_5 \end{aligned} \quad (\text{I-2})$$

Les incertitudes paramétriques θ_2 et θ_4 prennent en compte l'imprécision relative des capteurs par exemple due à des non-linéarités comme le phénomène de saturation, alors que θ_3 et θ_5 représentent des bruits aléatoires de capteurs. La nature proprement dite de ces variables sera précisée ultérieurement. ■

- *Résidu (Residual)*

Un résidu, communément appelé indicateur de défaut ou variable d'écart, est un signal conçu comme indicateur d'anomalies comportementales ou fonctionnelles.

I.2.2 Les différents types de défaut

Un défaut est naturellement caractérisé par le degré de gravité lié au stade auquel il est détecté (Brunet et al, 1990). Les défauts peuvent être purement aléatoires (défauts cataleptiques) ou progressifs dans le cas du vieillissement ou de l'usure (défauts évolutifs) et classés par ordre de gravité.

- *Défaut déclaré*

Un défaut déclaré apparaît de façon spontanée et se manifeste souvent par un biais pouvant avoir diverses origines comme une panne de capteur ou d'actionneur. Dans ce cas l'organe en question ne répond plus aux excitations du phénomène physique auquel il est soumis et délivre constamment à sa sortie une valeur fixe souvent nulle. Ceci peut être dû à un encrassement, à un mauvais étalonnage du capteur ou à un état détérioré d'un composant du système.

- *Défaut naissant*

Un défaut naissant apparaît de manière progressive et est d'amplitude croissante. Il se présente sous la forme de dérives souvent dues au vieillissement ou à l'usure.

- *Défaut ponctuel*

Un défaut ponctuel ou passager se caractérise par une durée brève et un écart significatif par rapport à la moyenne locale. Ses origines sont le plus souvent localisées au niveau des systèmes de collecte et de transport des mesures (parasites électromagnétiques, ...).

- *Panne catastrophique*

Une panne catastrophique entraîne un arrêt du système.

De plus, la manière dont un défaut est supposé agir sur le système peut être prise en compte ; plus exactement, il s'agit de formaliser la façon dont il pourrait être décrit au sein du modèle choisi. Deux catégories se dessinent alors : les défauts additifs et les défauts multiplicatifs (Isermann, 1994). Prenons par exemple le cas d'un système dont la représentation d'état dans le cas d'un fonctionnement normal est :

$$\begin{cases} q\mathbf{x}(t) = \mathbf{A} \mathbf{x}(t) + \mathbf{B} \mathbf{u}(t) \\ \mathbf{y}(t) = \mathbf{C} \mathbf{x}(t) \end{cases} .$$

Les matrices \mathbf{A} , \mathbf{B} et \mathbf{C} désignent respectivement les matrices d'état, de commande et d'observation du modèle. Quant aux vecteurs \mathbf{u} , \mathbf{x} et \mathbf{y} , ils représentent respectivement les vecteurs des entrées commandées, d'état et des mesures alors que q^{-1} symbolise l'opérateur retard. Supposons que la relation entrée-sortie associée soit donnée par :

$$\mathbf{y}(t) + \sum_{i=1}^n \mathbf{A}_i q^{-i} \mathbf{y}(t) = \sum_{i=0}^m \mathbf{B}_i q^{-i} \mathbf{u}(t),$$

où les matrices \mathbf{A}_i et \mathbf{B}_i sont de dimensions appropriées.

Les défauts additifs influencent l'état du système ou les relations entrée-sortie au travers de sommations alors que les défauts multiplicatifs les affectent au travers de produits avec \mathbf{x} et \mathbf{u} . La seconde catégorie permet par exemple de prendre en compte des variations de paramètres statiques (gains) ou dynamiques (constantes de temps) du modèle.

Défauts additifs	Défauts multiplicatifs
<ul style="list-style-type: none"> Représentation d'état : $\begin{cases} q\mathbf{x}(t) = \mathbf{A}\mathbf{x}(t) + \mathbf{B}\mathbf{u}(t) + \mathbf{L}\mathbf{f}_u(t) \\ \mathbf{y}(t) = \mathbf{C}\mathbf{x}(t) + \mathbf{M}\mathbf{f}_y(t) \end{cases}$ $\mathbf{f}_u(t)$: défauts additifs sur les actionneurs ou sur l'installation se répercutant sur l'état $\mathbf{x}(t)$. $\mathbf{f}_y(t)$: défauts additifs sur les capteurs. 	<ul style="list-style-type: none"> Représentation d'état : $\begin{cases} q\mathbf{x}(t) = (\mathbf{A} + \Delta\mathbf{A}) \mathbf{x}(t) + (\mathbf{B} + \Delta\mathbf{B}) \mathbf{u}(t) \\ \mathbf{y}(t) = (\mathbf{C} + \Delta\mathbf{C}) \mathbf{x}(t) \end{cases}$ $\Delta\mathbf{A}$, $\Delta\mathbf{B}$, $\Delta\mathbf{C}$: défauts multiplicatifs sur le système commandé, les actionneurs et les capteurs d'instrumentation.
<ul style="list-style-type: none"> Relation entrée-sortie : $\mathbf{y}(t) + \sum_{i=1}^n \mathbf{A}_i q^{-i} \mathbf{y}(t) = \sum_{i=0}^m \mathbf{B}_i q^{-i} \mathbf{u}(t) + \mathbf{f}_y(t) + \mathbf{B}_0 \mathbf{f}_u(t)$ $\mathbf{f}_u(t)$: défauts additifs sur un actionneur. $\mathbf{f}_y(t)$: défauts additifs sur un capteur. 	<ul style="list-style-type: none"> Relation entrée-sortie : $\mathbf{y}(t) + \sum_{i=1}^n (\mathbf{A}_i + \Delta\mathbf{A}_i) q^{-i} \mathbf{y}(t) = \sum_{i=0}^m (\mathbf{B}_i + \Delta\mathbf{B}_i) q^{-i} \mathbf{u}(t)$ $\Delta\mathbf{A}_i$ et $\Delta\mathbf{B}_i$ traduisent l'ensemble des défauts multiplicatifs.

Tableau I-1. Défauts additifs et multiplicatifs

Notons que cette terminologie est générale et s'applique à bien d'autres choses que les défauts comme les perturbations ou les incertitudes.

Exemple. Ainsi en reprenant les équations (I-1) et (I-2) de l'exemple de la section I.2.1, les incertitudes θ_i , $i \in \{1,2,4\}$ sont de nature multiplicative car elles affectent directement les paramètres du modèle. De manière générale, elles décrivent des imprécisions relatives en étant en facteur soit d'une mesure, soit de la valeur nominale d'un paramètre. L'idée est

que l'importance de cette imprécision est directement proportionnelle au point de fonctionnement courant ou aux caractéristiques nominales du système.

Les incertitudes θ_i , $i \in \{3,5\}$, quant à elles, sont de nature additive ; elles apparaissent sous la forme d'une sommation, représentent des bruits additifs, sont indépendantes du point de fonctionnement et n'affectent pas les paramètres du modèle. ■

I.2.3 Surveillance et diagnostic

Des anomalies peuvent être dues à une dégradation d'une partie de l'installation qui se traduit par la présence d'un défaut ou d'une défaillance (voire une panne dans un cas extrême) ou à une réaction peu habituelle du système physique (soumis par exemple à une perturbation). Le module de diagnostic doit donc permettre dans un premier temps de vérifier s'il s'agit d'une réaction normale ou d'un défaut. Ensuite, de manière simpliste, il a pour objectif de déterminer les composants en défaut au sein du système physique.

Le diagnostic s'inscrit en réalité dans une procédure plus complète désignée sous le nom de surveillance, nécessitant la mise en œuvre des différents modules présentés dans l'organigramme de la figure I-3 (Brunet et al, 1990), (Coirault et al, 1993). De plus, il s'effectue en réalisant trois tâches bien distinctes : la détection, la localisation et l'identification des défauts.

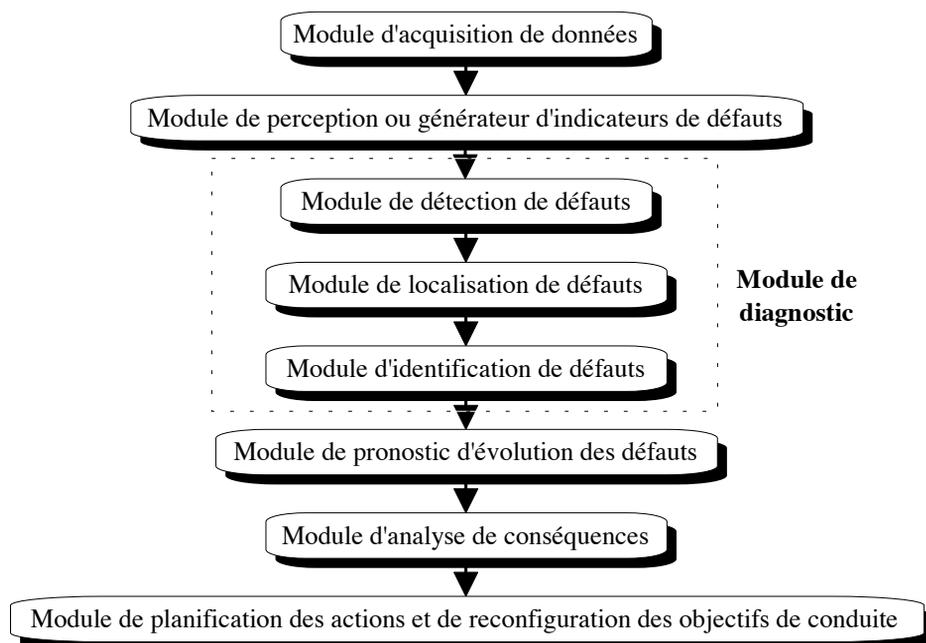


Figure I-3. Modules d'une procédure de surveillance

Ces différents modules sont détaillés dans ce qui suit.

- Module d'acquisition de données

Le module d'acquisition de données réalise :

- le conditionnement et le prétraitement du signal (décalage de zéro, filtrage, amplification,...),
- la validation du signal (par comparaison à un seuil par exemple),
- l'analyse de cohérence du signal (par comparaison à d'autres signaux par exemple).

- Module de perception

Le module de perception permet de compresser l'information diluée contenue dans les signaux observés et fournit des indices ou paramètres indicateurs d'une anomalie. Deux voies distinctes peuvent être dégagées.

- La première repose sur des techniques issues du traitement du signal comme :
 - l'analyse temporelle du signal et les traitements statistiques (calculs de moyennes, variances, corrélations, tendances) qui consistent à suivre à l'aide d'outils statistiques un certain nombre de variables significatives,
 - l'analyse fréquentielle (transformée de Fourier, étude de spectres) qui permet principalement d'étudier les phénomènes vibratoires (sur les moteurs électriques par exemple).
- La seconde repose sur des méthodes à base de modèles. Les variables observées sont liées par des relations comportementales définies soit par un modèle faisant appel aux lois physiques, soit à partir d'un modèle boîte noire se contentant de l'identité de comportement dans un certain domaine de fonctionnement. Les techniques d'estimation d'état ou d'estimation paramétrique fournissent alors indices ou paramètres (Sauter, 1991). Les mesures issues des capteurs et les variables estimées sont comparées, ce qui se traduit par la génération de résidus conditionnant l'information en vue de faciliter la détection d'un éventuel défaut.

- Module de détection de défauts (Fault detection)

Le module de détection permet de déterminer si le système physique fonctionne normalement et a pour objectif de signaler la présence d'un défaut en comparant le comportement courant du système avec celui donné pour référence. Malheureusement, il est impossible en pratique d'obtenir un comportement de référence scrupuleusement identique à celui du système en fonctionnement normal⁵. À cause de cette différence, les outils de détection sont généralement de nature statistique (tests d'hypothèses). Un risque d'erreur subsiste donc. Détecter un défaut inexistant peut provoquer un arrêt inutile et générer une

⁵ Nous détaillerons plus avant cette difficulté dans la section I.3.3.

perte de confiance des opérateurs (probabilité de fausse détection). A l'opposé, omettre un défaut, qui peut entraîner ultérieurement une panne, est préjudiciable (probabilité de non-détection). Il subsiste donc nécessairement un compromis induisant la recherche d'un réglage permettant de minimiser ces probabilités. A ce stade, l'objectif est de déterminer si un événement affectant le système est le signe d'une anomalie et par conséquent de distinguer les événements qualifiés de normaux (réaction de la commande pour rejeter une perturbation, action d'un opérateur) de ceux qui ne le sont pas (défaut).

- *Module de localisation de défauts (Fault isolation)*

Le module de localisation a pour objet la localisation d'un défaut détecté survenant sur les actionneurs, les capteurs d'instrumentation, la commande ou le système commandé en indiquant quel organe ou composant est affecté par celui-ci.

- *Module d'identification de défauts (Fault identification)*

Le module d'identification a pour but de caractériser le défaut en durée et en amplitude afin de le classer par types et degrés de sévérité. Ainsi, il peut servir à assurer le suivi de son évolution, ce qui est fort utile dans le cas d'un changement de comportement lent dû au vieillissement et à l'usure. De plus, ce module peut comprendre une procédure visant à déterminer la cause du défaut, c'est-à-dire son origine.

Ces trois derniers modules constituent le corps d'une procédure de diagnostic⁶. Lorsqu'une anomalie est détectée, le module de diagnostic doit donc :

- vérifier que celle-ci est bien un défaut,
- puis localiser le défaut pour déterminer le composant incriminé,
- enfin identifier le défaut (instant d'apparition, défaut fugitif ou persistant,...), c'est-à-dire déterminer ses caractéristiques.

Les outils principaux sont : la reconnaissance de formes, l'analyse discriminante floue, la logique floue, les arbres logiques et les systèmes experts.

- *Module de pronostic*

Le module de pronostic, sur la base de l'évolution passée d'un défaut, permet de prédire son évolution future dans un contexte similaire. Cette opération repose soit sur une connaissance a priori des lois d'évolution des phénomènes de vieillissement, soit sur une phase d'apprentissage. Le pronostic se fonde sur la prédiction de l'état prochain du système pour déterminer le temps restant avant l'apparition de la panne. Les outils utilisés concernent des techniques de traitement du signal (méthodes de lissage et d'extrapolation) ou de modélisation lorsque le défaut peut être identifié.

⁶ L'homologue anglo-saxon du terme diagnostic est *diagnosis*.

- Module d'analyse de conséquences

Le module d'analyse de conséquences détermine le temps restant avant une panne, et de façon plus spécifique, ses impacts sur :

- la disponibilité du système de production et la sécurité,
- la qualité du produit ou du travail effectué,
- les coûts économiques (pertes dans le cas où le produit ne serait pas fourni à temps, coût de la maintenance).

Les outils couramment utilisés sont les graphes de cause-conséquence et les systèmes experts à base de règles faisant intervenir le temps.

- Module de planification des actions et de reconfiguration des objectifs de conduite

Le module de planification des actions prend en compte le résultat des étapes de diagnostic, pronostic et d'analyse des conséquences, puis planifie dans le temps :

- les actions d'urgence,
- les reconfigurations de la loi de commande ou passages en mode dégradé,
- les reconfigurations des objectifs de conduite,
- les évitements de panne.

L'accommodation aux défauts regroupe les stratégies capables de maintenir un système à un niveau de performances acceptables, notamment en présence de défauts. Lorsque les objectifs prédéfinis peuvent être conservés, le terme reconfiguration est employé ; sinon, une stratégie de restructuration est mise en place afin de minimiser la dégradation des performances (Dardinier Maron et al, 1999). La reconfiguration est ainsi une stratégie qui consiste à corriger les défauts en modifiant les lois de commande ou en agissant par compensation afin de satisfaire les objectifs de production fixés. Ce module envoie des ordres, soit localement à l'opérateur ou à l'organe de commande, soit à un niveau hiérarchique supérieur comme le séquenceur ou planificateur. Les outils utilisés sont à base de règles.

Ces différents modules n'apparaissent pas nécessairement tous dans chaque application, mais il est nécessaire d'établir une hiérarchie des procédures de surveillance et des échelles de temps. Ainsi, les défauts cataleptiques doivent être détectés le plus rapidement possible. La détection est par conséquent une procédure temps réel. Les défauts évolutifs, quant à eux, doivent être surveillés de façon périodique et à un rythme adapté au vieillissement, de façon à modifier la commande (quitte à imposer un fonctionnement en régime dégradé) ou à réaliser une maintenance prédictive appropriée assurant le degré de disponibilité souhaité.

Pour compléter cet exposé, rappelons qu'une procédure de surveillance peut inclure une étape de réconciliation de données. Elle s'inscrit dans une stratégie à base de modèles

analytiques, et plus précisément lorsque ces derniers sont mathématiquement exacts⁷ et décrivent parfaitement le comportement du système physique étudié. Elle consiste alors à corriger les mesures issues du système d'instrumentation de manière à ce qu'elles satisfassent aux diverses relations constituant le modèle. Les données réconciliées offrent dans ces conditions une image plus juste du système et permettent de fournir une estimation de certaines grandeurs mesurées (Ragot et al, 1990). Cette démarche peut encore s'appliquer lorsque le modèle est incertain (Maquin et al, 1999, 2000a).

Pour conclure, cette section a permis de rappeler, comme l'indique l'intitulé de ce document, ce à quoi nous allons nous intéresser, en l'occurrence à la détection de défauts, et dans une moindre mesure à leur localisation.

I.3 LES PRINCIPALES STRATEGIES DE DIAGNOSTIC

I.3.1 Maintenance programmée et maintenance selon l'état

L'objectif de cette partie est d'entrevoir les différentes stratégies utilisées pour le diagnostic des systèmes. En règle générale, deux stratégies différentes se dessinent (Brunet et al, 990), (Gertler, 1991).

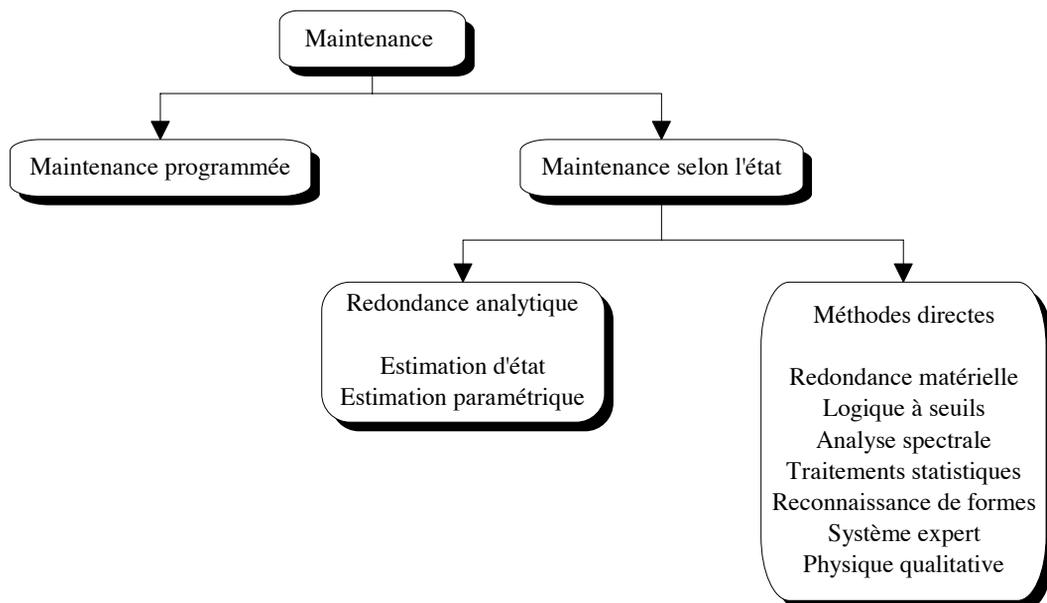


Figure I-4. Principales stratégies de maintenance

- *Maintenance programmée (ou systématique)*

Le terme maintenance programmée désigne une technique fondée sur une étude statistique de la fiabilité du système. Par observation, les temps moyens de bon

⁷ Ces modèles résultent de lois de conservation de la matière ou de l'énergie par exemple.

fonctionnement des éléments du système sont déterminés et un remplacement programmé des pièces est établi, entraînant soit un changement prématuré, soit un arrêt du système physique pour cause de défaillance ou de panne.

- Maintenance selon l'état

A l'opposé, existe une maintenance selon l'état qui nécessite la mise en place de diverses procédures telles que la détection puis la localisation et l'identification de défauts. L'intérêt de cette stratégie est d'optimiser le fonctionnement et la disponibilité des installations en proposant une politique de maintenance conditionnelle tenant compte de l'état courant du système (voire d'un état futur dans le cas où un module de pronostic est ajouté). De manière globale, la maintenance selon l'état se décompose en deux familles de méthodes.

- La première famille regroupe les méthodes qualifiées de directes ne nécessitant pas de modèle analytique et faisant appel à diverses techniques dont les principales sont citées ci-après.

- Redondance matérielle

L'objectif est de déterminer les capteurs présentant des anomalies et d'exhiber ceux fournissant des données cohérentes. L'approche la plus simple consiste à doubler (duplex) ou tripler (triplex) les chaînes d'instrumentation, c'est-à-dire obtenir plusieurs mesures d'une même grandeur physique, d'où la terminologie de redondance matérielle. Ensuite, les valeurs obtenues sont directement comparées deux à deux (Brunet et al, 1990). Il est aussi possible de tester la valeur de chaque mesure par rapport à un estimateur calculé à partir de l'ensemble des mesures. La méthode D.S.M.⁸, par exemple, utilise comme estimateur la moyenne des mesures associées à une grandeur physique ou cette même moyenne privée de la mesure correspondant au capteur testé (Dorr, 1995). Dans les deux cas, des résidus sont générés puis comparés à des seuils. Pour un système de mesures duplex (comparaison des sorties filtrées des deux chaînes d'instrumentation et détection fonctionnant par franchissement de seuils), la redondance matérielle ne permet que la détection simple d'un défaut. Le capteur incriminé ne peut pas alors être localisé, il est seulement possible de préciser la paire à laquelle il appartient. Dans le cas d'une redondance matérielle triple ou plus, il est possible cette fois-ci de détecter et localiser précisément le défaut en utilisant en cascade un détecteur et un voteur utilisant des seuils choisis en fonction des tolérances des capteurs et de la probabilité de pouvoir détecter un défaut d'amplitude donnée.

- Logique à seuils

Des variables mesurées sont comparées avec des valeurs limites constantes ou adaptatives (évoluant en fonction du point de fonctionnement). Un premier niveau indique la présence probable d'un défaut alors qu'un second niveau peut en

⁸ Abréviation de Détecteur Simple utilisant la redondance Matérielle.

caractériser la gravité. Le franchissement d'un seuil révèle la présence d'une anomalie.

- Analyse spectrale

Certains signaux mesurés ont une fréquence typique en fonctionnement normal et tout changement de celle-ci peut être considéré comme un indicateur de défaut.

- Traitements statistiques

L'étude de l'évolution de la moyenne ou de la variance d'un signal peut permettre de mettre en évidence une anomalie. La prise de décision est généralement effectuée à l'aide d'un test d'hypothèses. Deux hypothèses \mathcal{H}_0 et \mathcal{H}_1 représentant respectivement un fonctionnement normal et anormal sont spécifiées en fonction des probabilités de non-détection⁹ et de fausse alarme¹⁰. Le test d'hypothèses permet de calculer la probabilité d'occurrence de \mathcal{H}_0 et \mathcal{H}_1 et de déterminer l'hypothèse la plus vraisemblable. Cette opération est généralement réalisée sur une ou plusieurs fenêtres temporelles glissantes afin de détecter le défaut en ligne. L'utilisation d'une seule fenêtre nécessite de connaître les valeurs de référence des grandeurs caractéristiques du signal dont une modification témoigne de la présence d'une anomalie. Si ces valeurs sont inaccessibles, le travail consiste à comparer les résultats obtenus sur deux fenêtres différentes. Parmi les tests d'hypothèses les plus connus existent le test séquentiel du rapport de vraisemblance de Wald¹¹ qui possède l'avantage de déterminer seul l'horizon des observations à utiliser et minimise le nombre d'observations nécessaires pour prendre une décision sans ambiguïté ou le test de Page Hinkley (Brunet et al, 1990), (Zwingelstein, 1995), (Deckert et al, 1977), (Desai et al, 1979). Ce type de test a été plus spécifiquement développé dans le cadre de la redondance matérielle (Daly et al, 1979), (Hall, 1981), puis du diagnostic à base de modèles dynamiques (Chen et al, 1990), (Jin et al, 1997) sous le nom anglo-saxon de Generalized Likelihood Test. Des techniques spécifiques à la détection de changements de comportements abrupts sont présentées dans (Basseville et al, 1993).

- Reconnaissance de formes

La reconnaissance de formes est une technique qui utilise des algorithmes permettant de classer des objets dont l'aspect a varié par rapport à une référence. A chaque instant, la forme type à laquelle un objet ressemble le plus doit être déterminée.

- Système expert

Un système expert est un système informatique destiné à résoudre un problème précis à partir d'une analyse et d'une représentation des connaissances et du raisonnement d'un ou plusieurs spécialistes (Zwingelstein, 1995).

⁹ Fait de ne pas détecter un défaut présent.

¹⁰ Fait de conclure à la présence d'un défaut alors que le comportement du système est normal.

¹¹ Sequential Probability Ratio Test dont l'abréviation est S.P.R.T.

- Physique qualitative

Le diagnostic reposant sur des modèles qualitatifs (Travé-Massuyès et al, 1997) s'avère légitime lorsque la connaissance sur le système est intrinsèquement de nature qualitative, c'est-à-dire peu formalisée ou difficilement quantifiable. Il devient néanmoins tout aussi attrayant lorsque cette connaissance est d'origine quantitative (c'est-à-dire de nature numérique) mais affectée d'une certaine imprécision ou incertitude¹². En abandonnant la précision de l'information numérique, l'approche qualitative gagne en simplicité et permet de modéliser des systèmes très complexes. Ce gain de simplicité induit une perte d'information et par conséquent une imprécision plus importante sur les résultats de la procédure de diagnostic (en termes de défauts détectés puis localisés) que pour l'approche qualitative reposant sur des modèles analytiques¹³. Les principales techniques permettant d'appréhender des modèles qualitatifs reposent sur des formalismes tels que les ensembles flous (Frank et al, 1993), les multi-modèles (Boukhris, 1998), la théorie des graphes (Mandel, 1998) ou l'analyse par intervalles (Chang et al, 1995), (Ploix, 1998).

➤ La seconde famille correspond aux méthodes reposant sur une approche par modélisation et nécessitant un modèle analytique du système. Un tel modèle est constitué d'un ensemble de relations analytiques, statiques ou dynamiques, entre différentes variables représentant les phénomènes associés au système. Ces modèles, en fonction de leur structure, appartiennent à trois classes distinctes :

- Modèle d'observation ou de représentation (modèle boîte noire)

Le modèle est conçu uniquement afin de reproduire le comportement externe (entrées-sorties) du système dans une plage de fonctionnement donnée.

- Modèle de connaissance (modèle boîte blanche)

Le modèle est déduit des lois de la physique reliant les différents phénomènes observés et repose par conséquent sur une représentation physique du système.

- Modèle mixte (modèle boîte grise)

Ce modèle est un mélange des deux précédentes classes. Il s'agit d'un modèle de connaissance pour certaines parties du système et d'un modèle de représentation pour les autres dont le comportement est jugé trop complexe à décrire à l'aide des lois de la physique.

Au cours des trente dernières années, de nombreux auteurs ont publié des travaux de synthèse regroupant l'ensemble des techniques appartenant à cette famille : Willsky, (1976), Brunet et al (1990), Gertler (1986, 1988, 1991), Frank (1987, 1990, 1996),

¹² Nous discuterons plus amplement des notions d'imprécision et d'incertitude dans la section I.3.3 suivante.

¹³ Notons que cette notion de précision est somme toute relative car il est souvent nécessaire de définir, pour l'approche quantitative, des seuils de détection faisant office de zone d'incertitude. L'intérêt de l'approche qualitative est alors de prendre en compte dès le départ (c'est-à-dire au niveau du modèle) cette imprécision, et d'une certaine manière de générer des seuils naturellement.

Patton et al (1989, 1991, 2000), Patton (1994, 1997), Isermann (1994), (Ragot et al, 1990), (Zwingelstein, 1995), pour l'estimation d'état; Isermann (1984) pour l'estimation paramétrique.

Comme nous l'expliquerons plus en détails par la suite, l'étude proposée dans ce document s'inscrit simultanément dans ces deux familles (méthodes directes et méthodes utilisant un modèle du système). En effet, d'une part, elle repose sur un modèle purement qualitatif utilisant le formalisme des intervalles, d'autre part, la connaissance qui en est issue, est traitée à l'aide de techniques empruntées au diagnostic à base de modèles analytiques (Espace de Parité). En réalité, la tendance consiste de plus en plus à concevoir des techniques permettant d'exploiter le maximum d'informations disponibles sur le système. Ainsi, pour prendre en compte la notion d'imprécision, les méthodes à base de modèles analytiques empruntent de plus en plus des concepts utilisés initialement dans les approches liées aux modèles qualitatifs. Inversement, pour améliorer la précision de leurs résultats, ces dernières intègrent en partie des relations analytiques. La distinction entre ces deux familles devient donc de plus en plus confuse du fait qu'elles sont dans une certaine mesure complémentaires (Travé-Massuyès et al, 1997).

I.3.2 Principes du diagnostic à base de modèles analytiques

La section précédente s'attachait à rappeler les principales stratégies utilisées en diagnostic afin de situer le contexte de notre travail. L'objectif est maintenant de se focaliser sur le diagnostic à base de modèles analytiques et d'en préciser les principes essentiels. Pour déterminer si le système étudié fonctionne normalement, il est nécessaire de comparer son comportement courant à une référence, d'où la notion de redondance d'information inhérente à toute stratégie de diagnostic à base de modèles.

- Redondance analytique

La redondance matérielle fut la première technique de redondance utilisée pour la détection et localisation de défauts de capteurs. Comme explicité brièvement dans la section I.3.1, elle repose sur la présence physique de plusieurs capteurs d'instrumentation mesurant une même grandeur inconnue. Cette approche ne couvre pas les pannes de mode commun comme une panne d'alimentation électrique ou un problème de masse. De plus, les inconvénients majeurs de ce type de stratégie sont les coûts et encombrements élevés liés à la multiplication des chaînes de mesures, ainsi qu'un champ d'applications strictement limité aux défauts de capteurs.

L'intérêt d'utiliser des techniques à base de modèles permettant de remplacer un capteur physique par un capteur informationnel prend alors tout son sens. Si des équations analytiques provenant d'un modèle du système sont utilisées pour établir puis vérifier des relations liant les grandeurs mesurées entre elles, alors le terme de redondance analytique est employé. Sous cette expression générale, sont regroupées les techniques faisant appel à l'estimation

paramétrique ou à l'estimation d'état. En plus de limiter le nombre de capteurs nécessaires, la redondance analytique permet d'appréhender des défauts affectant aussi bien la chaîne d'instrumentation que le système commandé ou ses organes de commande puisqu'elle intègre bien plus d'information que la redondance matérielle.

Exemple. Reprenons le modèle incertain de la résistance électrique de la section I.2.1. Seuls deux capteurs sont disponibles, un de courant et un autre de tension ; il n'y a donc aucune redondance matérielle. Seule la redondance analytique issue de loi d'Ohm peut permettre d'effectuer un quelconque diagnostic. Réécrivons le modèle sous la forme suivante :

$$\begin{bmatrix} i \\ u \\ 0 \end{bmatrix} = \begin{bmatrix} 1 + \theta_2 & 0 \\ 0 & 1 + \theta_4 \\ (1 + \theta_1)r_0 & -1 \end{bmatrix} \begin{bmatrix} i_0 \\ u_0 \end{bmatrix} + \begin{bmatrix} \theta_3 \\ \theta_5 \\ 0 \end{bmatrix}.$$

Sans détailler la manière de procéder qui sera largement développée dans la suite de ce document, recherchons une relation de redondance exploitable, c'est-à-dire pour laquelle les valeurs vraies inconnues i_0 et u_0 ont été éliminées. En multipliant la relation précédente par le vecteur ligne :

$$\left[-(1 + \theta_1)(1 + \theta_4)r_0 \quad 1 + \theta_2 \quad (1 + \theta_2)(1 + \theta_4) \right],$$

l'expression d'un résidu p_{inc} , devant être nul en fonctionnement normal, est obtenue :

$$p_{inc} = -(1 + \theta_1)(1 + \theta_4)r_0(i - \theta_3) + (1 + \theta_2)(u - \theta_5). \quad (\text{I-3})$$

Dans le cas du modèle déterministe, le résidu $p_{dét}$ se réduit à :

$$p_{dét} = u - r_0 i. \quad (\text{I-4})$$

- Test de cohérence

De l'approche à base de modèles analytiques est emprunté le principe essentiel sur lequel repose notre travail, à savoir la notion de test de cohérence. Il s'agit de vérifier l'adéquation entre des grandeurs observées ou mesures représentant un comportement observé et un ou plusieurs modèles du système physique faisant office de référence en décrivant un comportement attendu.

- Procédure orientée bon fonctionnement

Deux grandes classes de méthodes utilisant le principe du test de cohérence sont envisagées. La première, appelée procédure bon fonctionnement (ou orientée bon fonctionnement), utilise un modèle décrivant un comportement normal du système¹⁴. Classiquement, une incohérence comportementale est révélée lorsque les résidus sortent d'un voisinage de zéro ou que les paramètres identifiés dérivent anormalement. Il faut souligner que pour ce type de stratégie, c'est l'incohérence entre le comportement observé du système

¹⁴ Il serait plus rigoureux de parler de procédure orientée bon comportement, néanmoins, c'est le terme procédure orientée bon fonctionnement qui est le plus couramment utilisé.

et celui attendu qui est étudié plutôt que la cohérence. En effet, cette dernière peut perdurer malgré la présence d'un défaut, non détecté pour des problèmes de sensibilité des résidus ou surtout d'excitation non persistante (cas d'un capteur pluviométrique bouché par temps sec par exemple), la cohérence dépendant du point de fonctionnement. Il ne sera donc possible de conclure qu'en présence d'une incohérence et que si les seuils de détection tiennent compte des perturbations, notamment des erreurs de modélisation et bruits de mesure.

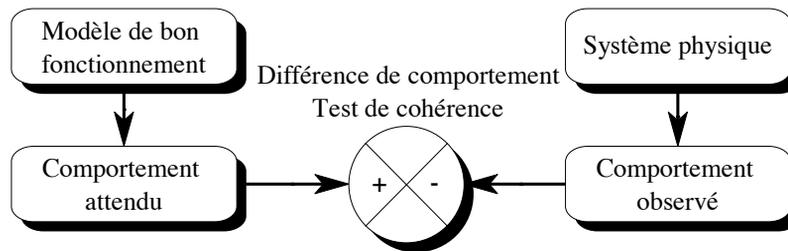


Figure I-5. Procédure orientée bon fonctionnement

- Procédure orientée mauvais fonctionnement

La seconde classe de méthodes correspond aux procédures orientées mauvais fonctionnement. Plus particulièrement adaptées à la localisation de défauts, elles sont nécessairement dédiées aux anomalies dont un modèle a pu être développé, ne serait-ce qu'en termes de modifications de certaines caractéristiques d'un résidu ou, de manière plus détaillée, sous forme de modèles analytiques. En effet, disposer d'un modèle pour chaque défaut pouvant survenir sur un système physique est pratiquement irréalisable. De ce fait, il est nécessaire de construire avant tout une première procédure orientée bon fonctionnement permettant, après détection, de déclencher la procédure mauvais fonctionnement dédiée à des anomalies particulières. La cohérence entre les observations et un modèle de mauvais fonctionnement conduit à conclure à la présence du défaut ciblé par le modèle. Néanmoins, ne possédant pas une représentation complète de toutes les anomalies possibles, il est possible que se produise un défaut non modélisé induisant, soit des incohérences vis-à-vis de tous les modèles de mauvais fonctionnement, auquel cas la procédure de localisation est mise en défaut, soit au contraire une cohérence se traduisant par une localisation erronée.

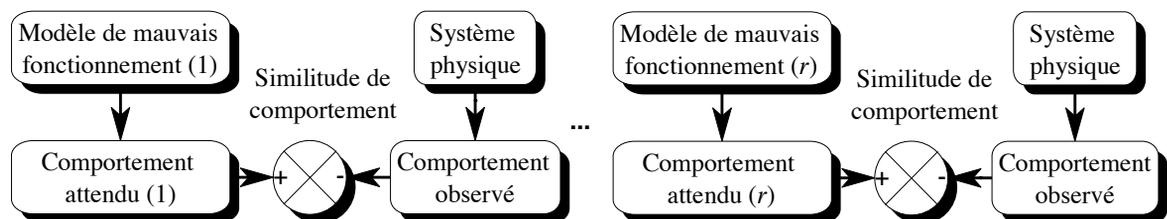


Figure I-6. Procédure orientée mauvais fonctionnement

Le tableau ci-joint résume la signification d'un test de cohérence dans le cas de procédures orientées bon ou mauvais fonctionnement en supposant que le système n'est soumis à aucune perturbation (Ploix, 1998).

	Procédure bon fonctionnement	Procédure mauvais fonctionnement
Cohérence	<ul style="list-style-type: none"> ▪ Comportement normal 	<ul style="list-style-type: none"> ▪ Comportement anormal ▪ Permet de conclure
Incohérence	<ul style="list-style-type: none"> ▪ Comportement anormal ▪ Permet de conclure 	<ul style="list-style-type: none"> ▪ Autre comportement

Tableau I-2. Test de cohérence

I.3.3 Problématique du diagnostic à base de modèles analytiques

Le problème majeur rencontré en diagnostic à base de modèles analytiques provient de l'essence même du principe sur lequel il repose, à savoir le test de cohérence. Il suppose ainsi qu'un modèle décrivant un comportement de référence fiable est disponible, ce qui soulève bon nombre d'interrogations liées au concept de modélisation. En effet, un modèle mathématique ne reflète qu'un comportement approché du système physique qu'il est censé représenter. Ces approximations proviennent des erreurs de modélisation qui surviennent lorsqu'un modèle est linéarisé (discrétisation d'un système d'équations continues, linéarisation autour d'un point de fonctionnement d'un système non-linéaire), lorsque certains phénomènes physiques ne sont pas pris en compte afin de simplifier les relations (influence de la température sur la résistance d'un enroulement inductif par exemple) ou tout simplement lorsque la recherche d'une précision infinitésimale n'est industriellement pas rentable ou impossible à garantir (tolérances sur les capteurs d'instrumentation). L'utilisation de lois de la physique pour décrire les relations entre phénomènes n'est pas elle-même exempte d'approximations puisqu'elle introduit bon nombre d'hypothèses simplificatrices souvent difficiles à satisfaire. D'autre part, l'utilisation de modèles de connaissance très précis dans la description des phénomènes mis en jeu, conduit à des systèmes d'équations trop complexes pour être exploitables en vue d'un diagnostic temps réel et souvent surparamétrés, posant alors des problèmes insolubles d'identification. Ce type de modèle est généralement issu d'une description microscopique visant à représenter un comportement à l'échelle macroscopique (comme par exemple des calculs de champ pour une machine électrique). Enfin, le système étudié (supposons une machine électrique) est en interaction avec l'environnement qui l'entoure (telle que la charge, la température du milieu où se trouve l'installation et le réseau de distribution EDF). Or il faudrait tenir compte de cet environnement pour obtenir un comportement de référence exact, chose totalement illusoire¹⁵.

Exemple. Dans le cas du modèle déterministe de la résistance électrique, les perturbations ne sont pas prises en compte. Le résidu p_{det} (I-4) aura donc toutes les chances, si le système fonctionne normalement, d'être différent de 0, synonyme d'incohérence et donc de présence d'un défaut. ■

¹⁵ Ne pas modéliser l'environnement peut être considéré comme mal modéliser le système, auquel cas une perturbation ne provient que d'une erreur de modélisation. Sinon, une perturbation traduit soit la présence d'une erreur de modélisation, soit l'influence de l'environnement du système.

Cependant, il faut éviter que les erreurs de modélisation ne soient comptabilisées comme des défauts car elles peuvent engendrer un nombre important de non-détections ou de fausses alarmes et par conséquent complètement fausser les résultats. La problématique du diagnostic repose donc avant tout sur le choix du modèle. Ce dernier comporte une certaine quantité d'information sur le système, et quelle que soit la technique utilisée, elle ne fera que traiter l'information disponible en la conditionnant en vue de faciliter son exploitation, mais ne pourra en aucun cas combler le manque de connaissances. Il est ainsi illusoire de vouloir détecter un défaut faisant appel à des informations non-contenues dans le modèle ; il est par conséquent nécessaire de se demander dans quelles conditions celui-ci reste valide et s'il permet de satisfaire à tous les objectifs de diagnostic, c'est-à-dire diagnostiquer les défauts imposés par le cahier des charges. Par exemple, un modèle de connaissance continu n'est pas forcément adapté si la technique utilisée pour le traiter nécessite de le discrétiser, mélangeant tous les paramètres physiques et rendant impossible la localisation précise des défauts.

Exemple. Dans le cas de l'exemple portant sur la résistance électrique, une seule relation de redondance existe que ce soit pour le modèle déterministe ou le modèle certain. Elle permet la détection d'un défaut d'amplitude suffisante, affectant un capteur par exemple. En revanche, la localisation de ce défaut est impossible, puisque la quantité d'information est insuffisante pour parvenir à distinguer l'influence des deux capteurs. ■

De plus, tout comme pour la commande, la notion de robustesse est présente en diagnostic. Dès mon travail de recherche mené en cours de DEA (Adrot, 1996a), il est apparu que, dans une discipline telle que le diagnostic, le concept de robustesse est bien plus complexe que dans le cadre de la commande de systèmes (Adrot, 1995). En effet, une loi de commande est robuste si, malgré certaines erreurs de modélisation, les spécifications de consigne sont satisfaites. Or, le problème du diagnostic d'une installation procède de la mise en évidence d'anomalies a priori indéterminées tant dans leur instant d'apparition que dans leur amplitude et parfois même, dans leur nature. Par conséquent, la notion de robustesse est subordonnée à la faculté de discriminer les défauts des indéterminations normales dues aux erreurs de modélisation. Pour limiter le nombre de fausses alarmes ou de non-détections, des indicateurs de défauts robustes sont générés : leur sensibilité aux défauts est maximale alors que leur sensibilité vis-à-vis des incertitudes de modèle est minimale (Patton, 1994). Dans le cas optimal, cette robustesse est obtenue par découplage, c'est-à-dire par élimination des grandeurs mal connues, et insidieusement des relations comportementales dans lesquelles elles interviennent. Ces perturbations sont généralement représentées par des incertitudes agissant sur le modèle dynamique du système et dont les matrices de distribution sont connues (Patton et al, 1989). L'inconvénient, dans ces conditions, est que le diagnostic n'est réalisé que sur une partie du système, celle parfaitement connue, sans prendre en compte l'information liée aux incertitudes rejetées. Or l'objectif est de conserver un maximum de l'information initiale. Le choix du modèle étant conditionné par les objectifs de diagnostic (à savoir la nature et l'amplitude des défauts à diagnostiquer, les organes du système à surveiller,...), il est alors clair qu'il est nécessaire d'intégrer l'incertain dès la phase de modélisation. Si l'automaticien se donne la peine d'intégrer dans son modèle des incertitudes, autant choisir

ensuite une méthode adaptée à leur prise en compte jusque dans la phase de décision, plutôt que de les éliminer au plus vite. C'est la tâche à laquelle nous allons nous consacrer dans la suite de ce document.

I.4 MÉTHODES DE DIAGNOSTIC À BASE DE MODÈLES ANALYTIQUES

I.4.1 Structure d'une procédure de diagnostic

L'objectif de cette section est de présenter brièvement les méthodes traditionnelles développées en diagnostic à base de modèles analytiques afin d'explicitier par la suite la manière dont notre travail s'en démarque. Très souvent dans la littérature internationale, seuls les modules de détection et de localisation sont considérés et détaillés. L'opération d'identification est souvent omise ou reléguée aux opérateurs car elle s'avère alors subsidiaire par rapport aux difficultés qu'elle peut induire. Le mot diagnostic est dans ces conditions avantageusement remplacé par le terme détection et localisation de défauts¹⁶. En revanche, le module de perception générant les résidus traités ensuite par la procédure de diagnostic est systématiquement développé. De nombreux auteurs décomposent alors une procédure de diagnostic en deux phases essentielles : la génération de résidus et la prise de décision (Chow et al, 1984), (Hammad, 1986), (Frank, 1990), (Gertler, 1991), (Patton et al, 1991).

- Génération de résidus

La première phase concerne la génération de résidus, grandeurs sensibles aux défauts. Elle correspond principalement à la phase de perception et vise à la mise en évidence d'événements qui affectent l'évolution du système et qui traduisent un changement de situation sans préjuger du fait qu'ils soient normaux (dus à une action délibérée de la part d'un opérateur par exemple) ou non. Un résidu est une grandeur obtenue explicitement à partir des mesures issues des capteurs et des signaux de commande des actionneurs, soit directement (espace de parité), soit indirectement par l'intermédiaire d'estimées (observateurs d'état, estimation paramétrique) ; les termes forme de calcul ou forme externe du résidu sont alors traditionnellement employés. En fonctionnement normal, un résidu doit être idéalement nul, c'est-à-dire indépendant du point de fonctionnement courant, donc des valeurs prises par l'état du système et les grandeurs d'entrées (commandes). En pratique, il est différent de zéro car il dépend implicitement des perturbations liées à l'environnement, des erreurs de modélisation et bien sûr des défauts. La relation liant le résidu à ces différentes grandeurs est appelée forme d'évaluation ou forme interne. La distinction entre formes de calcul et d'évaluation est purement arbitraire. Elle consiste simplement à diviser l'expression analytique d'un résidu en deux parties, l'une parfaitement connue et facile à calculer

¹⁶ Le terme anglo-saxon associé est *Fault Detection and Isolation*, d'où l'abréviation *F.D.I.*.

numériquement, l'autre dépendant des incertitudes modélisant les diverses perturbations agissant sur le système et des défauts à diagnostiquer.

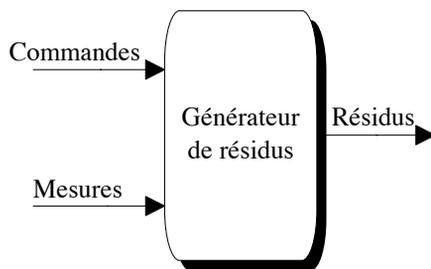


Figure I-7. Forme de calcul ou externe

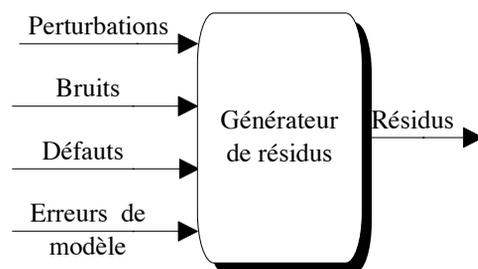


Figure I-8. Forme d'évaluation ou interne

Exemple. A partir du résidu p_{inc} (I-3), les formes de calcul et d'évaluation suivantes sont obtenues :

$$p_{cal} = r_0 i - u$$

$$p_{éval} = (1 + \theta_1)(1 + \theta_4)\theta_3 r_0 - (\theta_1 + \theta_4 + \theta_1 \theta_4)r_0 i - (1 + \theta_2)\theta_5 + \theta_2 u'$$

avec théoriquement, en fonctionnement normal : $p_{cal} = p_{éval}$. ■

Traditionnellement, l'objectif est de se munir de résidus robustes faiblement affectés par les perturbations et erreurs de modélisation, c'est-à-dire proches de zéro en leur présence (Watanabe et al, 1982), (Chow, 1980), (Chow et al, 1984), (Lou, 1982), (Lou et al, 1986), (Frank, 1990), (Gertler, 1991), (Patton et al, 1991), (Ragot et al, 1997a). En plus d'être sensibles aux défauts en augmentant suffisamment pour permettre une détection, ils doivent engendrer des signatures différentes pour chaque anomalie étudiée afin de permettre la localisation des composants en défaut et parvenir à les distinguer les uns des autres, d'où la notion de structuration de résidus (Massoumnia et al, 1988), (Gertler et al, 1990), (Gertler, 1991).

- *Prise de décision*

La seconde phase, correspondant au module de diagnostic proprement dit, est appelée prise de décision et regroupe :

- tout particulièrement les procédures de détection et de localisation qui concernent l'analyse des comportements pour distinguer ceux qui sont anormaux et la détermination précise du composant incriminé,
- puis dans une moindre mesure la procédure d'identification qui consiste à déterminer l'instant d'apparition du défaut, à le quantifier en estimant sa durée et son amplitude, ainsi qu'à en comprendre l'origine.

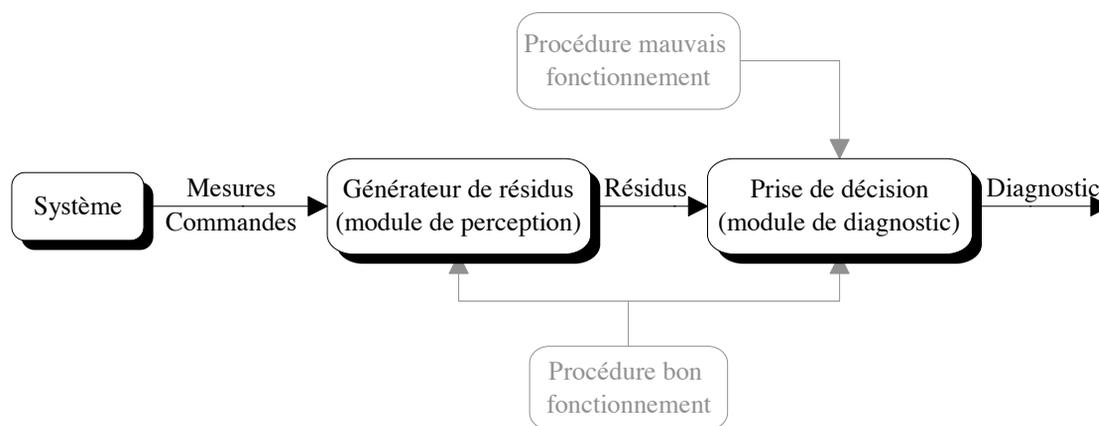


Figure I-9. Structure générale d'une procédure de diagnostic

Cependant, la distinction entre l'étape de génération de résidus et celle de prise de décision reste généralement confuse. Dans de nombreuses publications, les auteurs intègrent dans la première étape, des objectifs précis (en termes de robustesse ou de structuration) que doivent satisfaire leurs résidus. Ainsi, même si les modules de perception (génération de résidus) et de diagnostic (prise de décision) peuvent être étudiés séparément et séquentiellement, ils restent généralement mêlés, les résidus étant construits dès le départ de manière à faciliter et améliorer la seconde étape. C'est cette interaction qui rend la distinction entre les deux phases difficile et implique généralement, qu'en pratique, l'étude du module de perception fait partie intégrante de celle du module de diagnostic.

De plus, définir la procédure de diagnostic comme une simple phase de prise de décision est très restrictif. En effet, il est alors sous-entendu que le diagnostic repose uniquement sur les résidus générés lors de la première étape et qu'il est possible de conclure en exploitant seulement l'information issue de ces mêmes résidus. Ce découpage est purement théorique et a pour unique objectif de regrouper dans une même étape, en l'occurrence la seconde, toutes les procédures utilisant l'outil statistique. Un découpage plus pratique serait de définir l'étape de détection comme un premier niveau de diagnostic consistant à déterminer de façon rapide et fiable un état comportemental anormal. Elle repose nécessairement sur une procédure bon fonctionnement de manière à pouvoir garantir une réponse rapide et réalisée déjà, de façon sommaire, une première localisation (Ploix, 1998). L'étape de localisation, proprement dite, est un deuxième niveau de diagnostic, déclenchée par la procédure de détection, et qui consiste à déterminer de manière plus approfondie les composants en défaut. Cette étape peut utiliser à la fois des procédures bon et mauvais fonctionnement. Il apparaît alors clairement que les tests statistiques seront disséminés dans les deux étapes et non plus regroupés dans la seconde, et que toutes deux feront appel à une procédure de génération de résidus.

I.4.2 Méthodes développées en diagnostic à base de modèles analytiques

Deux grands types de méthodes se dégagent alors pour le diagnostic à base de modèles analytiques : l'estimation paramétrique et l'estimation d'état.

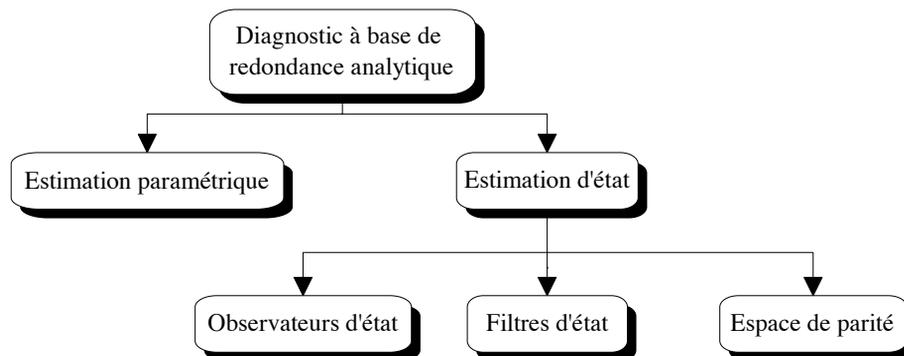


Figure I-10. Méthodes utilisées en diagnostic à base de modèles analytiques

I.4.2.1 Estimation paramétrique

L'estimation paramétrique désigne une catégorie de méthodes consistant à identifier, si possible en temps réel, les paramètres du système et à leur appliquer des tests d'hypothèses pour déterminer s'ils dérivent anormalement. La technique la plus connue est celle des Moindres Carrés et de ses multiples extensions dans le cas d'une régression linéaire. Elle a été notamment utilisée par Robet et al (1995) dans le cas d'une machine synchrone. Afin de traiter la classe des systèmes non-linéaires, certaines méthodes assurent la discrétisation des équations différentielles et la linéarisation par rapport aux paramètres à identifier comme la méthode des moments partiels ré-initialisés (Trigeassou, 1987) ((Coirault et al, 1993), (Kérignard et al, 1995) plus spécifiquement pour le cas bilinéaire), et permettent ensuite d'appliquer des techniques classiques d'estimation en ligne comme les Moindres Carrés. D'autres proposent l'utilisation du filtre de Kalman étendu (Zai et al, 1987), (Anderson et al, 1979), (Atkinson et al, 1989), (Loron et al, 1993), permettant l'estimation de certains paramètres du modèle en plus des composantes du vecteur d'état. En règle générale, les résidus proviennent de la différence entre les valeurs connues des paramètres et leurs estimées. Cette technique est particulièrement intéressante dans le cas de défauts multiplicatifs (c'est-à-dire intervenant sur les paramètres eux-mêmes) (Höfling et al, 1994, 1996a).

I.4.2.2 Estimation d'état

La seconde catégorie regroupe les méthodes reposant sur l'estimation d'état et se subdivise en trois familles distinctes : l'espace de parité, les observateurs d'états et les filtres.

- Espace de parité

La génération d'équations de parité repose sur des relations comportementales statiques ou dynamiques existant entre grandeurs mesurables (signaux de commande et sorties

mesurées). Cette méthode, ayant pour origine la redondance matérielle utilisée initialement à des fins de détection et localisation de défauts de capteurs, étend le diagnostic au système complet puisque ces relations peuvent en refléter la structure et le comportement dynamique (Espace de Parité Généralisé). Le principe consiste alors à reformuler les équations dynamiques du modèle sous forme de relations algébriques, d'où la notion de représentation ou forme statique, conduisant à une structure acausale¹⁷. En comparaison de l'approche par observateur, certains auteurs parlent de stratégie boucle ouverte¹⁸ (Patton et al, 1991) ou de modèle à horizon fini¹⁹ (Ploix, 1998). Le fait que des relations de parité ne soient pas vérifiées, génère des résidus, indicateurs de défauts auxquels sont appliqués des tests d'hypothèses.

La génération de résidus peut prendre deux formes suivant le type de systèmes étudiés (Daly et al, 1979), (Chow, 1980), (Frank, 1990), (Patton, 1991).

➤ Redondance analytique directe.

Le terme redondance directe est utilisé lorsque les équations de parité ne font intervenir que des mesures obtenues au même instant. Dans ces conditions, seuls des systèmes statiques peuvent être considérés et les résidus sont alors directement issus des relations de parité. (Brunet et al, 1990) distingue deux classes de redondance directe. Il parle de redondance directe lorsque seule la cohérence des mesures se rapportant à une même variable physique est exploitée. Les groupes de capteurs se rapportant à une même variable physique sont donc traités individuellement. À l'opposé, il parle de redondance statique lorsque des relations internes entre les différentes variables physiques (et indirectement les mesures) sont prises en compte. La redondance directe correspond alors simplement à la mise en équation de la redondance matérielle, alors que la redondance statique, par l'apport de contraintes (relations entre valeurs vraies par exemple), traite le cas plus général des systèmes statiques.

➤ Redondance analytique temporelle.

Le terme de redondance temporelle (ou dynamique (Brunet et al, 1990)) est utilisé dans le cas de mesures échantillonnées exprimées sur une fenêtre temporelle et s'applique à des modèles dynamiques discrets. Les résidus sont obtenus soit directement comme précédemment, soit en comparant estimées (issues de ces mêmes relations) et mesures (méthodes boucle ouverte, boucle fermée) (Deckert et al, 1977), (Chow 1980), (Chow et al, 1984). Certains auteurs ont travaillé sur la redondance temporelle dans le cas continu et ont fait apparaître de nombreuses similitudes avec le cas discret (Medvedev, 1994), (Höfling et al, 1994, 1996b).

¹⁷ Un système est causal si certaines variables physiques (les sorties) sont déduites de la connaissance d'autres (les entrées). Acausal est son antonyme.

¹⁸ C'est-à-dire sans contre réaction.

¹⁹ Un modèle est à horizon fini lorsque qu'au delà d'un certain laps de temps, les phénomènes s'étant produits n'ont plus aucune influence sur l'état présent. L'antonyme est un modèle à horizon infini.

Historiquement, la redondance directe fut développée afin de remplacer la redondance matérielle²⁰, dont les deux principaux inconvénients étaient l'encombrement et le coût élevé induits par le nombre de capteurs à implanter. Elle fut tout d'abord utilisée en complément de la redondance matérielle, et ce particulièrement dans le milieu aéronautique (Potter et al, 1977), (Daly et al, 1979), (Desai et al, 1979). Malheureusement, des relations algébriques entre mesures obtenues à un même instant ne permettaient de faire que de la surveillance de capteurs ; ce qui incita bon nombre de chercheurs à s'intéresser à des méthodes plus sophistiquées et fut à l'origine de l'essor de la redondance temporelle. Celle-ci fut utilisée pour la première fois dans le milieu aérospatial où Deckert et al (1977) proposa, dans le cadre d'une application pratique sur un avion, un système de redondance matérielle duplex pour détecter la paire de capteurs présentant une anomalie, complété par des relations dynamiques entre vitesses et accélérations visant à déterminer le capteur en défaut issu de cette paire.

La principale évolution fut ensuite l'introduction de l'Espace de Parité Généralisé permettant de généraliser la mise en œuvre des techniques de redondance temporelle et de traiter le cas des systèmes dynamiques discrets et linéaires (représentation d'état) en se plaçant sur une fenêtre temporelle de taille finie (Mironovskii, 1979), (Chow 1980), (Chow et al, 1984), (Lou, 1982), (Lou et al, 1986), (Hammad, 1986). Ceci permit de prendre en compte des relations entre les mesures non perçues dans le cas de la redondance directe et ainsi de traiter une plus grande variété de défauts en assurant la surveillance aussi bien de la chaîne d'instrumentation que des actionneurs et des composants du système commandé. Si, d'un côté, les inconvénients propres à la redondance matérielle sont limités, de l'autre, il faut faire face à une complexité croissante des algorithmes de calcul.

De nombreux auteurs ont apporté une contribution sérieuse à la génération d'équations de parité. (Massoumnia et al, 1988) démontre que la taille minimale de la fenêtre correspond à l'ordre de la partie observable du modèle d'état. (Chen et al, 1990) étend la notion d'espace de parité généralisé dans le cas stochastique en utilisant en plus des mesures, une prédiction de l'état du système à l'aide d'un filtre de Kalman. Certains auteurs proposent une approche de la redondance temporelle par matrices de transfert et s'intéressent plus précisément aux problèmes de localisation et de structuration des résidus. L'objectif est d'obtenir des résidus sensibles individuellement à certains défauts ou sous-ensembles de défauts, et insensibles aux autres. (Massoumnia et al, 1988) donne une méthode permettant d'obtenir des résidus qui ne dépendent que de la mesure d'un seul capteur (SSPR ou Single Sensor Parity Relation) ou d'une seule entrée (SAPR ou Single Actuator Parity Relation). Gertler et al (1990) proposent de générer des résidus orthogonaux à certains défauts ; l'amplitude d'un ensemble spécifique de résidus dits structurés va augmenter en présence d'un défaut donné. Ces conditions d'orthogonalité imposent des contraintes sur la structure des relations de parité et des algorithmes de transformation sont présentés.

D'autres auteurs s'intéressent à la construction de résidus robustes vis-à-vis des erreurs de modélisation. Chow (1980), Chow et al (1984) proposent de résoudre un problème de

²⁰ La redondance matérielle reste encore aujourd'hui largement répandue dans le milieu industriel du fait de sa simplicité (aviation civile, centrales nucléaires,...).

minimisation de la variance des résidus pour déterminer les meilleures relations de parité en termes de robustesse pour un point de fonctionnement donné. L'inconvénient de cette méthode est qu'elle conduit à un système complexe d'optimisation. Pour obtenir une liste de relations de parité ordonnée suivant leur degré de robustesse, il faut résoudre un problème séparé d'optimisation pour chacune d'entre elles. Lou (1982), Lou et al (1986) développent une approche géométrique et multi-modèle (choix de points représentatifs dans les domaines de valeurs que prennent les paramètres du modèle) du problème de robustesse. L'objectif est de générer une matrice de parité la plus orthogonale possible aux matrices d'observabilité du système ; cette méthode se ramène alors à un simple problème de décomposition en valeurs singulières. Frank et al (1989), Frank (1990) développent une approche de type entrées inconnues ou les incertitudes sont modélisées par des perturbations dont seules les matrices de dispersion sont connues. La matrice de parité doit alors conserver certaines propriétés en termes de normes par rapport à ces matrices de dispersion pour obtenir un découplage approché vis-à-vis des erreurs de modélisation. Le problème d'optimisation se traite à l'aide des valeurs et vecteurs propres généralisés. Staroswiecki et al (1993), Cassar et al (1995) généralisent ce problème à l'aide de fonctions de sensibilité aux défauts et aux entrées inconnues. En attribuant des poids à ce à quoi il faut être sensible ou robuste, ils démontrent que ce problème d'optimisation multi-critère conduit à un ensemble convexe de solutions optimales appelé lieu de Pareto et que la solution optimale donnée par Frank P.M. n'en est en fait qu'un cas particulier. Cocquempot et al (1997) s'intéressent notamment à l'influence de la taille de fenêtre temporelle des relations de parité sur la qualité des résidus déterminés, c'est-à-dire sur la valeur du critère d'optimisation.

Enfin, certains auteurs se sont penchés sur les liens existant entre la génération de résidus par espace de parité et par observateurs d'état, et démontrent un certain degré d'équivalence. (Massoumnia, 1986) a été le premier à parler de correspondance entre les deux stratégies, état de fait démontré plus tard par Staroswiecki et al (1991a), Staroswiecki (1991b). Les résidus issus d'un observateur de type Luenberger généralisé peuvent être obtenus en filtrant ceux construits à l'aide de l'Espace de Parité Généralisé. Dans le cas d'un observateur dont tous les pôles sont nuls (observateur à réponse pile), les résidus ont exactement la même structure (Patton et al, 1991). Frank et al (1989) aboutissent à un résultat similaire en explicitant la méthode des observateurs à entrées inconnues. Patton et al (1989, 1991) ont aussi montré qu'une simple relation de parité pouvait être obtenue en construisant un observateur à réponse pile par la méthode "eigenstructure assignment". De ce fait, certains auteurs considèrent que l'espace de parité est un cas particulier de l'approche par observateurs, ils parlent alors respectivement de stratégie boucle ouverte et de stratégie boucle fermée.

- Observateurs d'état

L'approche suivante est celle par observateur d'état (déterministe, complet ou d'ordre réduit, bancs d'observateurs). L'état inconnu du système est reconstruit en se recalant à l'aide de certaines mesures, le gain de l'estimateur dépendant des objectifs et performances désirées. Le but est de créer, à partir de l'estimation des mesures, des résidus qui seront testés afin de

détecter une rupture à l'aide de diverses techniques : reconnaissance de formes, analyse discriminante floue, arbres logiques, systèmes experts. En considérant que certaines entrées commandées sont mesurées alors que d'autres ne sont pas connues, il est aussi possible de rechercher à estimer à la fois l'état du système et sa commande, auquel cas le terme d'estimation d'état généralisé est employé (Adrot et al, 1999d).

Cette approche s'appuie sur une bonne connaissance du modèle et de ses paramètres, et nécessite l'intégration des diverses relations qui, contrairement aux relations de parité, sont différentielles. Dans le cas des systèmes linéaires, la structure de base des reconSTRUCTEURS est toujours la même. Un modèle parallèle, corrigé à l'aide de l'erreur d'estimation multipliée par un gain adéquat, est élaboré. Ce gain est souvent calculé par placement de pôles, ce qui permet d'imposer la dynamique désirée. Une notion essentielle, qui n'apparaît pas dans l'approche par Espace de Parité, est la stabilité de l'observateur, c'est-à-dire la faculté de l'estimée à pouvoir converger vers la vraie valeur. Notons que par construction, un observateur tend à recalculer son comportement sur celui du système. Si ce principe est naturel dans le cas d'un fonctionnement normal, il paraît contestable en présence d'un défaut puisqu'il aura tendance à atténuer le signe associé (si idéalement l'erreur d'estimation devient nulle, alors il n'y a plus moyen de détecter quoi que ce soit). En diagnostic, le problème est plus complexe que dans le cas de la commande dans la mesure où le gain joue aussi sur la manière dont les défauts vont affecter les résidus. En plus d'assurer la stabilité du système, ce gain permet de structurer les résidus de manière à localiser des défauts. Cependant, pour ce type de stratégie, une anomalie affecte en général toutes les composantes du vecteur des résidus ; de ce fait, le problème de localisation est plus complexe que dans le cadre de l'espace de parité.

Pour résoudre ce problème, une voie consiste à construire des bancs d'observateurs. Pour la détection de défauts affectant des capteurs ou des actionneurs, Clark (1989) introduit le schéma d'observateurs dédiés (DOS : Dedicated Observer Scheme). Cette structure dépend de la nature des défauts à mettre en évidence et diffère par les variables d'entrées pilotant l'observateur. Si l'objectif est de mettre en évidence des défauts de capteurs, un banc est construit où l' $i^{\text{ième}}$ observateur est piloté par l'ensemble des commandes et la sortie du $i^{\text{ième}}$ capteur. Le banc contient donc autant d'estimateurs que de capteurs à surveiller. Tous les résidus issus des estimateurs j ($j \neq i$) sont insensibles à un défaut se produisant sur l' $i^{\text{ième}}$ capteur. Le problème lié à cette technique est que l'estimation pilotée par une seule mesure peut s'avérer peu robuste vis-à-vis des bruits et erreurs de modélisation. Pour pallier à ces inconvénients, Frank et al (1989) proposent un banc d'observateurs généralisés (GOS : Generalized Observer Scheme). Dans le cas d'une détection de défauts de capteurs, l' $i^{\text{ième}}$ estimateur dépend des mesures de tous les capteurs, sauf du $i^{\text{ième}}$. Tous les résidus issus du $i^{\text{ième}}$ observateur sont donc insensibles à un défaut se produisant sur l' $i^{\text{ième}}$ capteur.

Le problème de robustesse a aussi été le sujet de nombreuses publications, les incertitudes provoquant une modification des résidus qui dévient de zéro en fonctionnement normal. Pour améliorer la robustesse, des méthodes passives, s'appliquant une fois la génération de résidus effectuée, peuvent être utilisées. Elles consistent à générer des seuils

adaptatifs qui dépendent de la commande. Clark (1989) propose une loi adaptative empirique alors que Emani-Naeini et al (1988) développent une démarche systématique et mathématique. À l'opposé, les méthodes qui s'appliquent directement lors de la génération de résidus sont dites actives. Dans cette optique, l'influence des incertitudes sur les résidus doit être minimisée. Plus exactement, le gain de l'observateur, qualifié alors de robuste, est calculé afin que les résidus soient parfaitement découplés des incertitudes, celles-ci étant considérées comme des perturbations dont les matrices de distribution sont connues. Deux grandes stratégies de découplage se dessinent dans l'abondante littérature consacrée à ce thème :

- les observateurs à entrées inconnues (Watanabe et al, 1982), (Frank et al, 1989).
- la construction d'observateurs par la méthode "eigenstructure assignment" (Massoumnia, 1986), (Patton et al, 1989, 1991).

Bien qu'utilisant des outils mathématiques différents, Gertler (1991) montre pour un exemple simple, que ces deux méthodes conduisent à des résultats équivalents.

Pour terminer, citons les observateurs à mémoire finie. En règle générale, les résidus sont générés en utilisant tout le passé du système, ce qui donne peu de poids aux mesures récentes susceptibles de contenir l'information sur le défaut à cause de l'effet de mémoire infinie du reconstituteur. Cette remarque est importante puisque les incertitudes de modèle peuvent aussi s'accumuler dans l'erreur d'estimation et entraîner une divergence (Nuninger, 1997). Les observateurs à mémoire finie en travaillant sur des fenêtres glissantes de taille finie permettent de contenir ces difficultés.

- Filtrés d'état

Par rapport aux observateurs qui s'adressent à un contexte déterministe, les filtres utilisent une formulation similaire du problème d'estimation et s'appuient sur des modèles stochastiques. Cette approche comprend différents problèmes d'estimation de l'état du système, qui dépendent de la position de l'instant courant t par rapport à la fenêtre temporelle d'observation $[t_0, t_f]$. Le lissage ou l'interpolation signifient qu'une partie seulement des mesures disponibles est prise en compte ($t < t_f$). De même, le terme filtrage indique que toutes les mesures sont utilisées ($t = t_f$) alors que prédiction ou extrapolation s'adressent au dernier cas ($t > t_f$). Seuls les filtres nous intéressent ici, dont le plus connu est celui de Kalman (Anderson et al, 1979). L'objectif est d'obtenir un estimateur non-biaisé et une variance de l'erreur d'estimation minimale. Le principal avantage est de générer des résidus décolorés sous certaines conditions (bruits gaussiens de moyennes nulles et indépendants). Si ces hypothèses ne sont pas vérifiées, des problèmes de divergence peuvent apparaître. Il est possible d'effectuer un changement de base pour s'affranchir de ce problème (filtre de Kalman-Bucy), néanmoins, les risques de divergence de la covariance de l'erreur d'estimation ne sont pas pour autant résolus. En effet, les imprécisions du modèle ainsi que l'effet mémoire infinie du filtre et les imprécisions sur les covariances des bruits influent. Des extensions permettent, en plus d'estimer les variables d'état, d'identifier des paramètres du modèle ; on parle alors de filtre de Kalman étendu (Anderson et al, 1979).

Pour simplifier la localisation des défauts, Beard (1971), Jones (1973), Massoumnia (1986) présentent un filtre de détection (*Failure Detection Filter*) permettant de rendre le vecteur des résidus, c'est-à-dire le vecteur des erreurs de prédiction, sensible à un défaut particulier par découplage à l'aide d'une transformation géométrique. Les différents résidus ainsi obtenus ont des propriétés unidirectionnelles envers les défauts affectant le système commandé ou les actionneurs, ou se propagent dans un plan particulier dans le cas de défauts de capteurs. L'objectif est que le vecteur des résidus dit à direction privilégiée (*Fixed direction residual vector*) conserve une direction particulière pour chaque défaut recherché, la localisation étant effectuée en la comparant à chaque direction théorique.

I.4.3 Diagnostic et approche bornante

Par essence, les méthodes de diagnostic reposant sur des modèles analytiques s'appuient sur des tests de cohérence vérifiant l'adéquation entre les mesures du système physique et la redondance d'information contenue dans le modèle de référence. La différence entre le comportement obtenu et celui attendu est conditionnée sous forme de résidus. Une incohérence comportementale est révélée lorsqu'un résidu sort du voisinage de zéro. Le problème est que ce principe essentiel au diagnostic est remis en cause par le fait qu'un modèle ne définit qu'un comportement approché du système physique. Pour éviter qu'une erreur de modélisation ne soit comptabilisée comme un défaut, des techniques consistant à générer des résidus robustes par découplage ne réalisent le diagnostic que sur une partie du système, celle parfaitement connue, sans prendre en compte l'information liée aux incertitudes rejetées. De plus, l'utilisation de ces techniques s'appuie sur l'invariance des paramètres incertains, permettant finalement de modéliser les perturbations sous forme d'incertitudes additives (Patton, 1994). Le travail proposé dans ce document se démarque des approches traditionnelles dans la mesure où, d'une part, les incertitudes fluctuent au cours du temps et peuvent être multiplicatives en affectant les paramètres du modèle ; d'autre part, l'objectif n'est plus de désensibiliser les indicateurs de défauts vis-à-vis de ces incertitudes. Pour pouvoir complètement prendre en compte ces dernières dans la procédure de diagnostic, il est clair qu'il est nécessaire de les intégrer dès la phase de modélisation. Ainsi, un modèle incertain permet de formaliser la méconnaissance associée à un système physique en indiquant, dans la représentation correspondante, qu'à certains endroits subsistent des imprécisions. Reste alors à décider de la manière de modéliser ces incertitudes et de la méthode utilisée afin de pouvoir les traiter sans les éliminer et assurer ensuite une prise de décision.

Deux approches statistiques permettent d'appréhender des variables incertaines : l'approche probabiliste et l'approche ensembliste (Jaulin, 2000). La première, de loin la plus répandue, consiste à représenter un vecteur \mathbf{x} de \mathbb{R}^n par une densité de probabilité ou loi de distribution f , fonction vectorielle vérifiant la propriété suivante :

$$f: \mathbb{R}^n \rightarrow \mathbb{R} / \int_{\mathbb{R}^n} f(\mathbf{x}) d\mathbf{x} = 1.$$

Cette description génère deux types d'information : d'une part le support de la variable incertaine, c'est-à-dire l'ensemble de toutes ses valeurs possibles²¹ ; d'autre part la façon dont cette variable est distribuée sur ledit support. L'approche ensembliste, quant à elle, est plus pauvre puisqu'elle ne dispense des informations que sur le support, en l'occurrence les frontières le délimitant.

Traditionnellement, les incertitudes sont représentées, dans une même relation, par un terme additif unique sous la forme d'un bruit gaussien. En pratique, l'hypothèse de blancheur requise n'est pas réaliste, notamment dans le cas où des modèles très réducteurs introduisant d'importantes erreurs structurelles sont utilisés. De plus, émettre des hypothèses statistiques sur les paramètres eux-mêmes rend le problème délicat à traiter à cause des opérations non-linéaires à effectuer sur les densités de probabilité (le produit de deux variables gaussiennes n'est plus gaussien par exemple) (Ploix, 1998). Enfin, par définition, le support des variables gaussiennes est infini, ce qui est difficilement conciliable avec la description de phénomènes physiques.

Ce document, faisant suite au travail effectué par Ploix (1998), propose une méthode de diagnostic tenant compte de manière globale de la structure et de l'amplitude des incertitudes. Chaque paramètre incertain est ainsi représenté par une variable bornée (ou variable intervalle), outil déjà exploité en identification (Milanese et al, 1996a) mais peu répandu en diagnostic (Chang et al, 1995). La modélisation à l'aide de variables bornées permet de prendre en compte, directement à l'intérieur du modèle, les incertitudes portant sur des paramètres physiques, ce qui laisse entrevoir une manière plus naturelle de concevoir les seuils de détection. Dans certains cas, l'approche ensembliste est parfaitement adaptée à la prise en compte d'incertitudes de par la nature de la connaissance disponible. Ainsi, sur une chaîne d'instrumentation, les erreurs de mesure sont quantifiées en terme de tolérance ou précision technologique, et non sous forme de loi de distribution. De plus, travailler uniquement sur les supports des variables incertaines, sans tenir compte d'une quelconque loi de distribution, permet d'étudier des systèmes plus complexes. Dans le cas de modèles affectés par des incertitudes paramétriques (appelées aussi incertitudes multiplicatives), des produits entre variables incertaines apparaissent rendant l'approche probabiliste difficilement exploitable, contrairement à l'approche ensembliste.

L'objectif est de réussir à construire, à l'aide de l'analyse par intervalles (Moore, 1979), (Neumaier, 1990), des enveloppes (ou des domaines plus complexes tels que des polytopes) tenant compte des plages de variations des incertitudes et permettant de définir les frontières du domaine correspondant à un comportement normal du système. En effet, avec ce type d'approche, un modèle ne définit plus un comportement de référence, mais un ensemble de comportements possibles ou acceptables. Cette approche ensembliste, que nous désignerons sous le nom d'approche bornante (Ploix, 1998) est à comparer aux méthodes stochastiques plus classiques où les seuils de détection, souvent constants, sont construits en fonction des taux de non-détections et de fausses alarmes imposés, ainsi que des caractéristiques statistiques (moyenne, écart-type) des signaux porteurs d'informations en fonctionnement

²¹ C'est-à-dire correspondant à une densité de probabilité non-nulle.

normal et en présence d'un défaut (souvent modélisé par un biais de capteur). L'avantage dans le cas de l'approche bornante est de construire des seuils adaptatifs, c'est-à-dire dépendant du point de fonctionnement courant, et ce en utilisant toute l'information a priori contenue dans le modèle sans faire appel à une quelconque hypothèse statistique ou un choix a posteriori portant sur le compromis non-détection/fausse alarme.

	Modèle déterministe	Modèle incertain
Nature des variables	Variable déterministe	Variable incertaine <ul style="list-style-type: none"> ▪ Variable bornée (support borné) ▪ Variable stochastique (support et loi de distribution)
Représentation	Comportement moyen	Ensemble de comportements possibles
Indicateurs de défauts	Valeurs, signaux	Domaines, enveloppes

Tableau I-3. Modèles déterministe et incertain

Exemple. Revenons à l'exemple relatif à la résistance électrique. Les modèles déterministe ou incertain ne sont que des descriptions comportementales approchées de la réalité. Cependant, le second a l'avantage de permettre, si les plages de variations des paramètres incertains ont été correctement réglées, de générer un ensemble de comportements dont la réunion comprend le comportement réel du système, d'où la notion d'enveloppe ou de domaine de bon fonctionnement.

Pour cet exemple académique, le jeu de données a été obtenu par simulation du modèle incertain. La valeur nominale r_0 de la résistance vaut 50Ω . L'erreur relative sur la résistance est de 20%, signifiant que la variable incertaine θ_1 fluctue dans l'intervalle $[-0.2, 0.2]$. De même, les incertitudes multiplicatives θ_2 et θ_4 appartiennent à l'intervalle $[-0.05, 0.05]$ indiquant que l'erreur relative sur les deux capteurs est de 5%. Les incertitudes additives θ_3 et θ_5 sur les capteurs sont simulées sous forme de bruits aléatoires indépendants appartenant au support $[-0.02, 0.02]$. Une fois l'échantillon de valeurs vraies du courant choisi, celui de la tension vraie est calculé à l'aide de (I-1) en faisant varier aléatoirement θ_1 sur son support. Les mesures i et u , représentées sur la figure I-11-a) sont ensuite obtenues grâce à (I-2). Contrairement aux autres, les incertitudes θ_2 et θ_4 ne sont pas choisies aléatoirement, mais dépendent des valeurs vraies i_0 et u_0 et correspondent à des phénomènes de saturation non-linéaire décrits par deux fonctions réelles tangente hyperbolique :

$$x = f_x(x_0) = \alpha_x \tanh(\beta_x x_0), x \in \{i, u\}.$$

Les quatre coefficients $\alpha_{(\cdot)}$ et $\beta_{(\cdot)}$ sont déterminés de manière à ce que les valeurs des mesures i et u restent toujours dans leurs domaines de valeurs possibles²² respectifs imposées $[1 - 0.05, 1 + 0.05]i_0$ et $[1 - 0.05, 1 + 0.05]u_0$ comme représenté sur la figure I-11-b).

²² Une procédure d'optimisation est utilisée afin de minimiser la distance entre $f_x(x_0)$ et la droite $(1+0.05)x_0$ tout en satisfaisant la contrainte $f_x(\max(x_0)) = (1 - 0.05)\max(x_0)$.

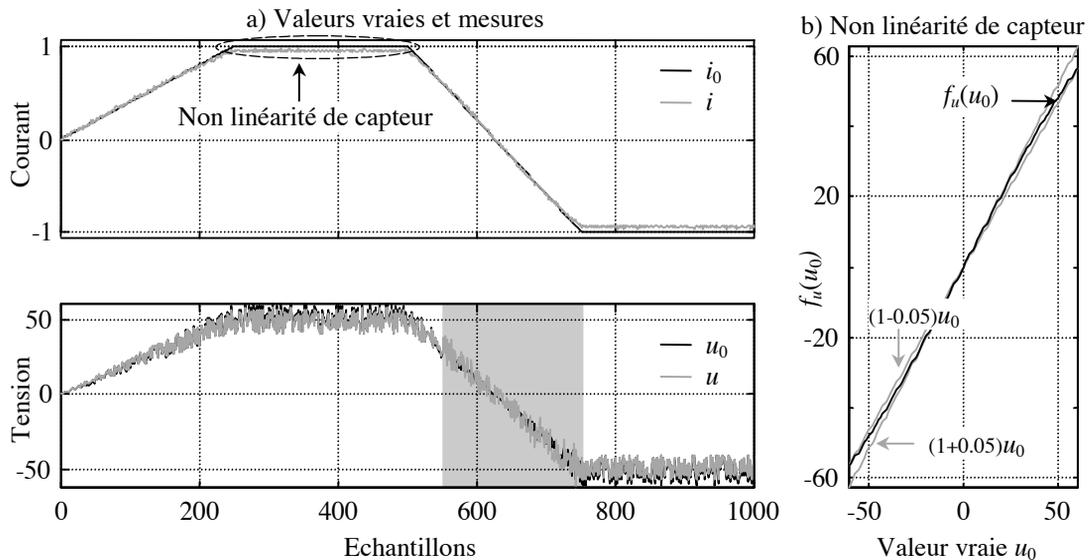


Figure I-11. Simulation du jeu de données

Un défaut de capteur généré sous la forme d'un bruit aléatoire centré et d'amplitude 8V est ajouté sur la mesure u de la tension entre les échantillons indicés de 550 à 750 (zone grisée de la figure précédente). Visuellement, le signe de ce défaut est difficilement perceptible du fait de l'importance des incertitudes. Sur la figure I-12-a), le résidu $p_{dét}$ (I-4) associé au modèle déterministe est représenté. Comme attendu, il est non nul en l'absence d'anomalie, ce qui nécessite d'établir des seuils de détection afin d'éviter de fausses alarmes intempestives. Les seuils de cet exemple (en gris) sont choisis simplement de manière à ce que le résidu (en noir) reste en fonctionnement normal entre ces derniers. Le résultat du test de cohérence dessiné en dessous vaut 0 en cas de cohérence, c'est-à-dire lorsque le résidu $p_{dét}$ prend une valeur située entre les deux seuils, 1 lorsque le défaut est détecté. Les résultats obtenus dans le cas du modèle incertain sont présentés sur la figure I-12-b). Ce modèle définissant un ensemble de comportements possibles, l'évaluation du résidu p_{inc} (I-3) génère un intervalle, dont l'empilement sur l'ensemble des échantillons conduit à une enveloppe (en noir). En fonctionnement normal, p_{inc} doit être nul, un défaut sera donc détecté si pour au moins un échantillon, la valeur zéro (ligne en gris) n'appartient plus à l'enveloppe. Nous ne détaillerons pas plus avant la technique utilisée pour déterminer l'enveloppe, ni les différents tests de cohérence permettant d'appréhender ce type de modèle ; en effet, respectivement, les chapitres II et IV leur seront spécifiquement dédiés. Le résultat du test de cohérence est proposé juste en dessous. Les bornes des variables incertaines étant connues exactement, il est logique qu'en fonctionnement normal, la valeur 0 reste dans l'enveloppe.

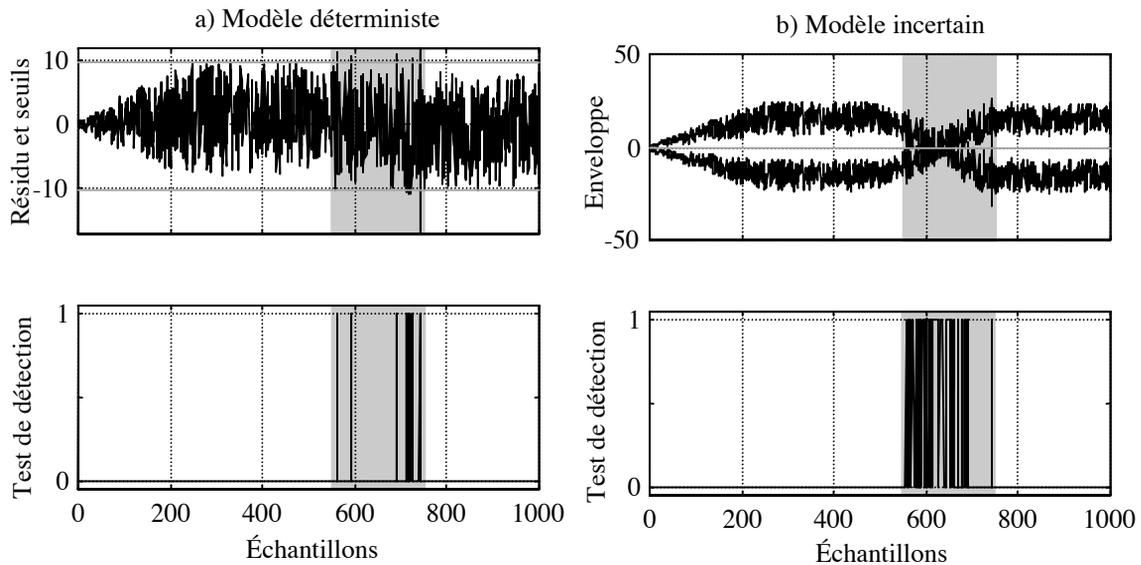


Figure I-12. Tests de cohérence

La procédure de détection reposant sur le modèle incertain conduit à de bien meilleurs résultats (73 échantillons incohérents contre 9). Le faible taux de détection provient de l'importance des incertitudes et de la nature du défaut dont l'amplitude varie aléatoirement. Mais au delà des performances obtenues dans le cadre spécifique de cet exemple, l'intérêt est de montrer que dans le cas de l'approche bornante, les seuils sont adaptatifs, de plus leur construction est naturelle et découle directement de la modélisation du système. Pour terminer, rappelons que le problème de non-détection subsiste toujours puisque inhérent à ce type de méthode où les incertitudes peuvent masquer la présence d'un défaut. ■

I.5 CONCLUSION

Pour résumer, l'approche bornante semble donc prometteuse dans la mesure où elle permet d'utiliser au maximum l'information disponible, notamment celle liée à la partie incertaine du modèle. De plus, en prenant en considération des incertitudes dépendantes du temps aussi bien additives que multiplicatives, elle permet de générer automatiquement des seuils de détection adaptatifs définissant les limites du domaine de bon fonctionnement. Cependant, elle apporte aussi son lot de difficultés propres à l'analyse par intervalles, particulièrement celles liées aux dépendances entre variables bornées, dont le pendant dans le cas stochastique est le problème de blancheur. Dans le cadre de l'approche probabiliste, pour simplifier, les bruits sont généralement considérés indépendants, hypothèse rendant difficile l'interprétation des résultats obtenus. Dans le cadre de l'approche ensembliste, les problèmes de dépendances et les moyens de les prendre en compte seront développés dans le prochain chapitre.

Chapitre II

Analyse ensembliste

TABLE DES MATIÈRES

II.1 INTRODUCTION	51
II.2 ARITHMÉTIQUE DES INTERVALLES	54
II.2.1 Motivations	54
II.2.2 Définitions et notations	55
II.2.3 Opérations ensemblistes	57
II.2.3.1 Opérations logiques	57
II.2.3.2 Opérations arithmétiques	58
II.2.4 Dépendance entre variables bornées	60
II.2.5 Propriétés algébriques de l'arithmétique des intervalles	62
II.2.5.1 Propriétés conservées	62
II.2.5.2 Propriétés non-conservées	63
II.2.6 Fonction intervalle	64
II.2.6.1 Extension d'une fonction réelle	64
II.2.6.2 Extension intervalle monotone par inclusion	66
II.2.6.3 Fonctions élémentaires	67
II.2.6.4 Application aux fonctions intervalle arithmétiques	69
II.3 EVALUATION D'UNE FONCTION D'INCLUSION	70
II.3.1 Méthodes reposant sur des formes centrées	71
II.3.1.1 Topologie de l'intervalle	71
II.3.1.2 Algorithme et propriété de convergence	73
II.3.1.3 Forme centrée	75
II.3.2 Procédures de réduction et d'élimination	78
II.3.2.1 Principes	78
II.3.2.2 Application à l'évaluation de fonctions d'inclusion	80
II.4 CHAMP DE VECTEURS DE FONCTIONS D'INCLUSION	84
II.4.1 Problème de dépendance et notion d'espace abstrait	84
II.4.2 Propriétés des espaces abstraits	86
II.4.3 Évaluation d'un champ de vecteurs de fonctions d'inclusion	89
II.4.4 Effet d'enveloppement	92
II.5 CONCLUSION.....	96

II.1 INTRODUCTION

L'analyse par intervalles fut initialement développée en vue de tenir compte des imprécisions sur la valeur des nombres. Ces imprécisions peuvent provenir des données issues de la chaîne d'instrumentation, qui comme précisé dans le chapitre précédent, n'est jamais parfaite. Ce problème est lié alors à la qualité du capteur (auquel cas le fabricant indique une précision technologique), aux conditions opératoires, à la présence de signaux parasites ou tout simplement d'un défaut (biais systématique, dérive,..., à cause d'un mauvais étalonnage ou d'un encrassement). La seconde source d'imprécision provient de l'outil informatique lui-même. Même si un nombre est parfaitement connu, son codage informatique sous forme d'une série plus ou moins longue de bits est nécessairement fini. Ainsi, un nombre rationnel tel que $\frac{1}{3}$ ne peut être exactement représenté. La valeur réellement utilisée dans les calculs sera un nombre tronqué 0.3333... où la quantité de chiffres après la virgule dépend du nombre de bits assurant le codage. De plus, de nombreuses fonctions usuelles sont en réalité calculées à l'aide d'approximations.

Au cours des années 60, parallèlement à la montée en puissance de l'informatique, se pose la question de savoir comment ces incertitudes évoluent au cours des calculs et comment quantifier l'erreur sur le résultat final. C'est en s'attachant à résoudre ce problème que Moore R.E. ouvre la voie de l'analyse par intervalles, en publiant notamment l'ouvrage de référence (Moore, 1966) mettant en place les fondements de cet outil, suivi de (Moore, 1979). Plus tard, Neumaier (1990) développe plus spécifiquement la résolution de systèmes d'équations linéaires ou non-linéaires, tandis que Hansen (1992) s'attache au problème d'optimisation globale, c'est-à-dire la recherche de manière garantie de tous les minimiseurs d'un critère non-connexe, alors que les méthodes plus classiques d'optimisation non-linéaire (Gradient, Newton-Raphson,...) ne renvoient qu'un seul optimum pouvant n'être que local.

Si la montée en puissance de l'analyse par intervalles est due dans un premier temps à des difficultés rencontrées dans le domaine informatique, puis à l'étude de problèmes mathématiques complexes tels que la recherche de toutes les solutions d'un ensemble d'équations non-linéaires ou l'optimisation de façon garantie d'un critère non-connexe (Didrit, 1997), (Kieffer et al, 1999), il ne faut pas oublier certains résultats plus récents obtenus dans le domaine de l'automatique. Le principal attrait de l'analyse par intervalles réside dans sa capacité à appréhender les incertitudes sur les paramètres d'un modèle. C'est donc logiquement que cet outil fit son apparition dans le domaine de l'estimation paramétrique. Appelée estimation paramétrique ensembliste, elle repose généralement sur un modèle linéaire en les paramètres et caractérisé par une erreur d'équation bornée. L'objectif consiste alors à déterminer l'ensemble des valeurs acceptables des paramètres¹, c'est-à-dire cohérentes avec les mesures, le modèle et les bornes de l'erreur d'équation (Milanese et al,

¹ Le terme anglo-saxon correspondant est Feasible Parameter Set (FPS). Les termes Set Membership ou Likelihood Set peuvent aussi être rencontrés.

1996b). Dans le cas d'un modèle linéaire, le domaine recherché est un polytope convexe. Les premières stratégies proposées consistent à circonscrire le véritable domaine à l'intérieur d'une forme plus simple telle une ellipsoïde (Fogel et al, 1982), ou une boîte (Milanese et al, 1982), (Feng et al, 1999). Ensuite, l'amélioration apportée à cette démarche consiste à caractériser de manière exacte ce polytope (Piet-Lahanier et al, 1988, 1990), (Broman et al, 1988), (Mo et al, 1988, 1990). L'étape suivante est la prise en compte de bruits de mesure bornés sur la sortie dans un premier temps (Clément et al, 1988), puis sur toutes les observations (Cerone, 1993, 1996) en sus de l'erreur d'équation. Dans le cas où le modèle est statique, le domaine des valeurs acceptables des paramètres reste un polytope convexe ; en revanche dans le cas d'un modèle dynamique, il devient non-linéaire et un domaine approché doit être recherché (Veres et al, 1996). Notons que ce type d'estimation peut être appréhendé comme un problème d'inversion ensembliste (Jaulin et al, 1993, 1996). Une technique de pavage est alors utilisée et permet d'obtenir une approximation du domaine recherché quelle que soit sa complexité. L'essentiel des travaux portant sur l'approche ensembliste en estimation paramétrique sont regroupés dans l'ouvrage collectif (Milanese et al, 1996a) et de nombreux articles de synthèse ont été publiés (Walter et al, 1990).

L'utilisation de l'analyse par intervalles dans le cadre du diagnostic est nettement plus récente et laborieuse. Les premiers travaux répertoriés sont imputables à Chang et al (1995). Ils utilisent alors une variante² de l'arithmétique des intervalles initialement proposée par Moore (1966) à des fins de détection et de localisation de défauts en s'appuyant sur une approche de type Espace de Parité. Le modèle utilisé repose sur des lois de conservation (masse, énergie,...) et prend en compte explicitement les défauts à diagnostiquer, permettant ainsi de générer, pour chaque anomalie étudiée, des indicateurs censés déterminer la plus probable d'entre elles. Plus tard, Ragot et al (1997b) proposent une technique ensembliste en vue d'effectuer la détection et la localisation de défauts de capteurs dans le cas d'un modèle statique et linéaire. Les relations liant les différentes grandeurs vraies étant supposées exactes dans la mesure où le modèle est issu de lois de conservation, seules les incertitudes affectant les capteurs sont prises en compte. Cette technique s'appuie sur la comparaison d'intervalles en étendant la relation d'ordre au cas d'intervalles d'intersection non-vide. Malheureusement, la distance qui en résulte n'est pas forcément significative dans la mesure où la répartition des points à l'intérieur des dits intervalles est quelconque. La même année, un court chapitre de l'ouvrage de Travé Massuyès et al (1997) est consacré à ce thème en exposant le principe de la propagation de contraintes sur les intervalles. L'approche bornante dans le cadre du diagnostic ne prendra véritablement son essor en France que l'année suivante avec la thèse de Ploix (1998) à laquelle fait suite notre travail. Justifiant dans un premier temps l'intérêt d'une telle approche, Ploix S. propose un formalisme permettant d'appréhender les modèles dont les incertitudes paramétriques sont décrites par des variables bornées. Nous en retiendrons le concept d'espace abstrait rappelé ultérieurement dans ce chapitre et justifierons ce choix. L'auteur explicite en outre les différentes structures d'incertitudes (additives, multiplicatives) et les propriétés qui en découlent, puis expose une méthode de caractérisation assurant l'estimation des bornes des incertitudes. Il formalise les tests de cohérence exploitables dans

² Variante que Chang appelle en anglo-saxon Segment Interval Arithmetic.

le cadre de l'approche bornante et développe une méthode reposant sur l'approche par Espace de Parité. Des notions essentielles telles que la dépendance entre variables bornées ou la nature temporelle des incertitudes sont abordées. Depuis, d'autres travaux ont vu les jours comme (Tzes et al, 1999) cherchant à déterminer à quels instants se produisent des sauts (signes de défauts) dans les valeurs des incertitudes paramétriques.

Notons que de nombreux travaux, traitant du diagnostic dans le cas de modèles à horizon infini où les incertitudes sont représentées par des variables bornées, ont été réalisés ces dernières années, et ce plus spécifiquement en Espagne. La stratégie développée dans les articles (Armengol et al, 1999a, 1999b, 2000), (Armengol, 1999c), se fonde sur l'analyse modale par intervalles³ permettant la construction d'une paire d'enveloppes encadrant celle exacte à rechercher. Il est ainsi possible de quantifier à chaque instant l'erreur maximale vis-à-vis de l'enveloppe exacte et de l'ajuster en fonction de la précision souhaitée en augmentant la taille de l'enveloppe minorante et en diminuant celle majorante. Puig et al (1999, 2000), quant à eux, proposent une technique d'optimisation globale permettant d'évaluer les enveloppes associées à un modèle intervalle récursif. Celui-ci est défini par une représentation d'état discrète dont les valeurs limites des matrices d'état sont connues, sachant que les incertitudes sont supposées invariantes. L'enveloppe multidimensionnelle obtenue, pour des raisons de temps de calcul, n'est en fait qu'une approximation du domaine réel. D'autres travaux existent comme ceux de Hadj-Sadok et al (1998, 1999) qui appliquent une méthode ensembliste dans le cadre d'une installation de traitement des eaux usées.

En dépit de ses fortes capacités, l'analyse par intervalles reste tout de même peu usitée, et ce quel que soit le domaine de recherche considéré. Le principal inconvénient résulte de la dépendance entre variables bornées que ne peut prendre facilement en compte cet outil, entraînant, comme nous le verrons au cours de ce chapitre, un effet de majoration (ou d'approximation extérieure) rendant les bornes obtenues pessimistes. Une autre difficulté réside dans les temps d'exécution des algorithmes ensemblistes qui peuvent rapidement s'avérer prohibitifs lorsque le nombre de variables bornées s'accroît.

Notons que dans le cadre de ce document, nous ferons abstraction des problèmes liés aux erreurs numériques. Seules les incertitudes portant sur les paramètres de modèle seront prises en compte dans l'élaboration de notre procédure de diagnostic par approche bornante.

Dans ce chapitre, nous rappellerons dans un premier temps les bases de l'analyse par intervalles. Une fois les opérations ensemblistes élémentaires présentées, le problème fondamental de dépendance, pouvant conduire à des majorations très importantes des bornes calculées lors de l'évaluation d'une fonction intervalle, sera présenté. Le concept d'extension intervalle et la propriété d'inclusion sur laquelle repose l'analyse par intervalles seront ensuite explicités. La seconde partie de ce chapitre se focalise sur diverses méthodes (techniques à base de formes centrées, méthodes de réduction et d'élimination) permettant de prendre en compte le problème de dépendance afin de réduire le pessimisme des bornes obtenues dans le cas de fonctions vectorielles. La dernière partie s'intéresse au cas des champs de vecteurs de

³ Modal Interval Analysis.

fonctions intervalle et montre que le problème de dépendance intervient aussi entre composantes. Un formalisme légèrement différent de ce qui se rencontre traditionnellement est alors développé. Reposant sur la notion d'espace abstrait, ce formalisme étend celui de l'analyse par intervalles permettant initialement de ne représenter que des pavés, à des domaines bien plus complexes. Enfin, nous justifierons par le problème de l'effet d'enveloppement, le choix d'une stratégie par espace de parité en vue de réaliser une procédure de diagnostic.

II.2 ARITHMÉTIQUE DES INTERVALLES

II.2.1 Motivations

Par définition, un intervalle est un ensemble fermé et borné de nombres réels (Moore, 1979), (Neumaier, 1990). Si x désigne une variable réelle bornée, alors l'intervalle \mathbb{X} auquel elle appartient est défini par :

$$\mathbb{X} = \{x \in \mathbb{R} / \underline{x} \leq x \leq \bar{x}\}, \quad (\text{II-1})$$

où \underline{x} et \bar{x} sont des nombres réels représentant respectivement les bornes inférieure et supérieure de \mathbb{X} . De manière générale, l'intervalle \mathbb{X} sera noté comme suit :

$$\mathbb{X} = [\underline{x}, \bar{x}].$$

Si une seconde variable réelle y est élément de \mathbb{X} , alors nous noterons : $y \in \mathbb{X}$. Dans ces conditions, toutes les valeurs possibles de y sont comprises entre les deux bornes \underline{x} et \bar{x} , sans forcément les atteindre.

Cette notation permet de distinguer la variable bornée x , dont la valeur courante est inconnue et peut fluctuer entre \underline{x} et \bar{x} , de l'intervalle \mathbb{X} qui est un ensemble parfaitement déterminé. De manière pratique, un domaine d'appartenance ou support peut ainsi être associé à un paramètre, certes inconnu, mais dont les bornes peuvent être appréciées. Ce domaine, représenté sous la forme d'un intervalle, traduit ainsi la notion d'incertitude. Dans la suite de ce document, les variables incertaines seront systématiquement définies en tant que variables bornées.

Un autre avantage de l'utilisation des intervalles est que l'évaluation d'une fonction intervalle conduit à un ensemble contenant nécessairement les valeurs recherchées. Dans l'exemple modélisant une résistance du chapitre précédent, l'enveloppe calculée (figure I-12) définit l'ensemble des comportements jugés normaux pour le système étudié. Plus exactement, l'analyse par intervalles permet de garantir les résultats obtenus dans le sens où tous les points cohérents avec le modèle de référence sont bien contenus dans cette enveloppe. Nous détaillerons plus spécifiquement ce que la notion de garantie (Jaulin, 2000) apporte au diagnostic dans le chapitre IV dédié aux tests de cohérence.

Enfin, d'un point de vue algorithmique, l'analyse par intervalles assure l'étude par blocs d'un domaine de recherche au lieu de le parcourir classiquement par une stratégie de nuage de points, cet espace étant en général constitué d'une quantité indénombrable d'éléments impossible à balayer en un temps fini. Une méthode ponctuelle n'utilise que des évaluations locales. Pour s'assurer par exemple qu'une fonction réelle f ne s'annule pas sur un intervalle \mathbb{X} donné, il faudrait en chaque point du domaine de recherche vérifier si cette fonction devient nulle, au risque sinon de manquer une solution se trouvant en deux points testés. En revanche, l'analyse par intervalles permet d'obtenir simplement l'image de \mathbb{X} par f , révélant ainsi le comportement global de cette fonction. Cette stratégie établit de manière garantie si la fonction f s'annule sur \mathbb{X} et permet de déterminer, si besoin est, l'ensemble des solutions $f(x)=0$ sans que leur nombre soit connu a priori (Didrit, 1997). Spécifiquement dans le cadre du diagnostic, cet aspect sera utilisé au chapitre IV pour déterminer si le point de fonctionnement courant du système est cohérent, c'est-à-dire appartient à son domaine de bon fonctionnement.

Notons que représenter le support d'une variable incertaine x par un intervalle \mathbb{X} ne suppose en aucune façon que celle-ci décrive une loi de distribution uniforme f définie par :

$$f(x) = \frac{h(x - \underline{x}) - h(x - \bar{x})}{\bar{x} - \underline{x}}, \text{ avec } h(x) = 1 \text{ si } x \geq 0, h(x) = 0 \text{ sinon (fonction d'Heaviside).}$$

En effet, la loi de distribution résultant de la somme ou du produit de deux distributions uniformes n'étant plus uniforme (Ploix, 1998), le maniement de variables régies par ces lois serait bien trop complexe. Par conséquent, la seule hypothèse à retenir est que la variable x ne peut prendre que des valeurs appartenant à l'intervalle \mathbb{X} .

II.2.2 Définitions et notations

Cette représentation reste valide dans le cas d'un paramètre parfaitement connu, en effet, dans ce cas les deux bornes sont égales à x et l'intervalle \mathbb{X} , qualifié de dégénéré, est alors défini par :

$$\mathbb{X} = [x, x].$$

La notion d'intervalle peut aisément être étendue au cas d'un vecteur \mathbf{x} constitué de n variables réelles $x_i \in \mathbb{R}, i \in \{1, \dots, n\}$. Le vecteur intervalle \mathbb{X} contenant \mathbf{x} se définit comme suit :

$$\mathbb{X} = [\mathbb{X}_1 \quad \dots \quad \mathbb{X}_n]^T$$

où chaque intervalle $\mathbb{X}_i = [x_i, \bar{x}_i]$ est associé à une variable réelle x_i . Prenons le cas d'un vecteur de dimension 2. L'intervalle \mathbb{X} définit un orthotope⁴ aligné avec les axes du repère (x_1, x_2) , plus familièrement appelé boîte ou pavé, comme le montre la figure II-1.

⁴ Dans \mathbb{R}^n , un orthotope est un hyper-rectangle, c'est-à-dire un domaine délimité par $2n$ hyperplans de dimensions $n-1$ qui sont soit orthogonaux, soit parallèles. Signifier qu'un orthotope est aligné veut dire que ces hyperplans sont orthogonaux à des axes du repère paramétrique (les arêtes de l'hyper-rectangle sont donc colinéaires à ces axes).

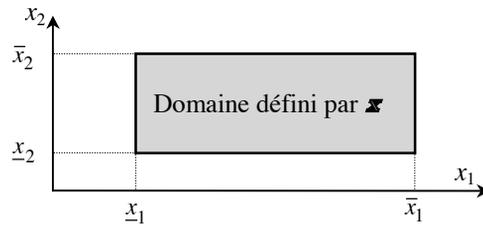


Figure II-1. Vecteur intervalle de dimension 2

Un certain nombre de caractéristiques se rapportant à un intervalle sont couramment utilisées. Nous pouvons citer, entre autre, son centre, sa longueur et sa valeur absolue (norme dans le cas d'un vecteur), dont les expressions⁵ sont données dans le tableau ci-dessous :

	Intervalle \mathbb{X}	Vecteur intervalle \mathbf{x}
Centre ou Milieu	$c(\mathbb{X}) = \frac{\bar{x} + \underline{x}}{2}$	$c(\mathbf{x}) = [c(\mathbb{X}_1) \ \cdots \ c(\mathbb{X}_n)]^T$
Longueur ou Taille	$w(\mathbb{X}) = \bar{x} - \underline{x}$	$w(\mathbf{x}) = \max(w(\mathbb{X}_1), \dots, w(\mathbb{X}_n))$
Valeur absolue ou Magnitude / Norme	$ \mathbb{X} = \max(\underline{x} , \bar{x})$	$\ \mathbf{x}\ = \max(\mathbb{X}_1 , \dots, \mathbb{X}_n)$

Tableau II-1. Expressions des principales caractéristiques d'un intervalle

Le centre et la longueur représentent le barycentre et la dispersion de l'intervalle. La valeur absolue d'un intervalle correspond à rechercher le maximum de la valeur absolue de ses bornes ; de ce fait, l'inégalité $|x| \leq |\mathbb{X}|$, $\forall x \in \mathbb{X}$, est toujours respectée.

Lorsque la borne inférieure est l'opposé de la borne supérieure ($\bar{x} = -\underline{x}$), un intervalle est qualifié de symétrique. Si \mathbb{X} est un intervalle symétrique, alors ses caractéristiques se simplifient pour s'exprimer comme suit :

Caractéristique	Expression
Centre	$c(\mathbb{X}) = 0$
Longueur	$w(\mathbb{X}) = 2\bar{x}$
Valeur absolue	$ \mathbb{X} = \frac{w(\mathbb{X})}{2} = \bar{x}$

Tableau II-2. Caractéristiques d'un intervalle symétrique

Un intervalle symétrique peut donc avantageusement se mettre sous la forme $\mathbb{X} = \bar{x}[-1,1]$.

Pour terminer, nous appellerons variable bornée normalisée toute variable x comprise entre 1 et -1 : $|x| \leq 1$, ce qui correspond à poser : $\mathbb{X} = [-1,1]$. Notons qu'un vecteur \mathbf{x} de variables bornées normalisées est caractérisé par la propriété suivante : $\|\mathbf{x}\|_{\infty} \leq 1$.

⁵ Notons que ces expressions peuvent varier suivant les auteurs. Par exemple, [Neumaier \(1990\)](#) utilise le rayon d'un intervalle qu'il définit comme la moitié de sa longueur. Le rayon et la norme d'un vecteur intervalle \mathbf{x} correspondent alors aux vecteurs constitués respectivement des n rayons et n valeurs absolues de ses composantes \mathbb{X} .

II.2.3 Opérations ensemblistes

II.2.3.1 Opérations logiques

Les intervalles peuvent être vus comme des ensembles sur lesquels s'appliquent des opérateurs logiques (égalité, intersection, réunion, relation d'ordre, inclusion) permettant de les comparer.

Deux intervalles \mathcal{X} et \mathcal{Y} sont égaux si et seulement si leurs bornes sont égales :

$$\mathcal{X} = \mathcal{Y} \Leftrightarrow \underline{x} = \underline{y} \text{ et } \bar{x} = \bar{y}.$$

Cette définition s'étend aisément dans le cas vectoriel où l'équivalence précédente doit être vérifiée pour chaque composante des vecteurs intervalle \mathcal{X} et \mathcal{Y} .

L'intersection de deux intervalles \mathcal{X} et \mathcal{Y} est vide si l'une des conditions $\underline{x} > \bar{y}$ ou $\bar{x} < \underline{y}$ est vérifiée. Sinon, cette intersection est aussi un intervalle défini par :

$$\mathcal{X} \cap \mathcal{Y} = \left[\max(\underline{x}, \underline{y}), \min(\bar{x}, \bar{y}) \right].$$

Dans le cas vectoriel, l'intersection de \mathcal{X} avec \mathcal{Y} est vide si au moins une des intersections entre deux composantes de même indice de ces deux intervalles est vide. Sinon, une relation similaire à la précédente est utilisée :

$$\mathcal{X} \cap \mathcal{Y} = \left[\mathcal{X}_1 \cap \mathcal{Y}_1 \quad \cdots \quad \mathcal{X}_n \cap \mathcal{Y}_n \right]^T.$$

L'union de deux intervalles est définie si et seulement si leur intersection est non-vide, et dans ce cas, c'est aussi un intervalle :

$$\mathcal{X} \cup \mathcal{Y} = \left[\min(\underline{x}, \underline{y}), \max(\bar{x}, \bar{y}) \right].$$

Si l'intersection est vide, il est possible d'utiliser l'intervalle enveloppe de l'union qui complète celle-ci de manière à obtenir un ensemble connexe (Didrit, 1997) :

$$[\{\mathcal{X}, \mathcal{Y}\}] = \left[\min(\underline{x}, \underline{y}), \max(\bar{x}, \bar{y}) \right].$$

L'union de deux vecteurs intervalle dont l'intersection est non-vide n'est généralement pas un vecteur intervalle, car la réunion de deux pavés \mathcal{X} et \mathcal{Y} qui s'intersectent ne conduit pas à un pavé, sauf cas particulier lorsque \mathcal{X} est entièrement inclus dans \mathcal{Y} ou inversement.

Comme pour les variables réelles, les relations d'ordre peuvent être définies ainsi :

$$\mathcal{X} \text{ op } \mathcal{Y} \Leftrightarrow \bar{x} \text{ op } \underline{y} \text{ où } \text{op} \in \{<, \leq\},$$

$$\mathcal{X} \text{ op } \mathcal{Y} \Leftrightarrow \underline{x} \text{ op } \bar{y} \text{ où } \text{op} \in \{>, \geq\}.$$

Les relations d'ordre sont antisymétriques et transitives, et les inégalités \geq et \leq sont réflexives. Pour comparer deux intervalles, il est nécessaire que l'intersection de ces derniers soit vide (inégalité stricte) ou limitée à une borne commune (inégalité large). Les relations précédentes sont valables dans le cas de deux vecteurs intervalle lorsque la relation d'ordre est vérifiée pour chaque composante.

Pour terminer, une autre relation logique, couramment utilisée, est l'inclusion :

$$\mathbb{X} \subseteq \mathbb{Y} \Leftrightarrow \underline{x} \geq \underline{y} \text{ et } \bar{x} \leq \bar{y}.$$

L'extension dans le cas vectoriel ne pose aucun problème puisqu'il suffit de vérifier la relation précédente pour toutes les composantes de même indice des vecteurs intervalle \mathbb{X} et \mathbb{Y} .

II.2.3.2 Opérations arithmétiques

Les intervalles peuvent aussi être vus comme des couples de réels, et non plus seulement en tant qu'ensembles. Les opérations arithmétiques {addition (+), soustraction (-), multiplication (.), division (/)} sur les variables réelles peuvent donc être reformulées dans le cadre de l'analyse par intervalles, d'où le terme largement employé dans la littérature internationale d'arithmétique des intervalles. Prenons le cas de la somme de deux intervalles \mathbb{X}_1 et \mathbb{X}_2 . La double inégalité (II-1) appliquée à ces deux grandeurs conduit à la relation :

$$\underline{x}_1 + \underline{x}_2 \leq x_1 + x_2 \leq \bar{x}_1 + \bar{x}_2,$$

L'intervalle $\mathbb{X}_1 + \mathbb{X}_2$ peut donc être calculé de la manière suivante :

$$\mathbb{X}_1 + \mathbb{X}_2 = \{x_1 + x_2 / x_1 \in \mathbb{X}_1 \text{ et } x_2 \in \mathbb{X}_2\}.$$

De même, pour chaque opération arithmétique, élément de l'ensemble $\Psi = \{+, -, \cdot, /\}$, il est possible d'écrire :

$$\mathbb{X}_1 \circ \mathbb{X}_2 = \{x_1 \circ x_2 / x_1 \in \mathbb{X}_1 \text{ et } x_2 \in \mathbb{X}_2\}. \quad (\text{II-2})$$

Dans le tableau ci-joint sont définis les opérateurs de l'arithmétique des intervalles.

Opération arithmétique	Calcul des bornes de l'intervalle obtenu
Addition	$\mathbb{X} + \mathbb{Y} = [\underline{x} + \underline{y}, \bar{x} + \bar{y}]$
Négation	$-\mathbb{X} = [-\bar{x}, -\underline{x}]$
Soustraction	$\mathbb{X} - \mathbb{Y} = [\underline{x} - \bar{y}, \bar{x} - \underline{y}]$
Multiplication	$\mathbb{X}\mathbb{Y} = [\min(\underline{x}\underline{y}, \underline{x}\bar{y}, \bar{x}\underline{y}, \bar{x}\bar{y}), \max(\underline{x}\underline{y}, \underline{x}\bar{y}, \bar{x}\underline{y}, \bar{x}\bar{y})]$
Inversion	$\frac{1}{\mathbb{X}} = \left[\frac{1}{\bar{x}}, \frac{1}{\underline{x}} \right]$ si $0 \notin \mathbb{X}$
Division	$\frac{\mathbb{X}}{\mathbb{Y}} = \left[\min\left(\frac{\underline{x}}{\underline{y}}, \frac{\underline{x}}{\bar{y}}, \frac{\bar{x}}{\underline{y}}, \frac{\bar{x}}{\bar{y}}\right), \max\left(\frac{\underline{x}}{\underline{y}}, \frac{\underline{x}}{\bar{y}}, \frac{\bar{x}}{\underline{y}}, \frac{\bar{x}}{\bar{y}}\right) \right]$ si $0 \notin \mathbb{Y}$

Tableau II-3. Opérations arithmétiques sur les variables intervalle

L'inversion n'est définie que si 0 n'est pas élément de \mathbb{X} , c'est-à-dire si l'une des contraintes $\underline{x} > 0$ ou $\bar{x} < 0$ est respectée. En effet, dans le cas contraire, l'inverse de \mathbb{X} ne peut être ni connexe, ni compact, ce qui est en contradiction avec la définition d'un intervalle. Dans le cadre de la modélisation des systèmes, les incertitudes sont d'amplitude finie, les paramètres physiques tels qu'une résistance ne changent pas de signe et les domaines de bon fonctionnement qui en résultent sont fermés. C'est pour cette raison que nous ne détaillerons

pas plus les problèmes engendrés par des intervalles contenant au moins une borne infinie. Néanmoins, il est utile de savoir qu'une arithmétique des intervalles étendue existe et permet d'appréhender de tels domaines (Didrit, 1997).

Par définition, le produit $\mathbb{X}\mathbb{Y}$ est positif ou nul si les deux intervalles sont de mêmes signes⁶ et négatif ou nul s'ils sont de signes opposés. De plus, son expression peut être simplifiée si la connaissance sur les signes des bornes de \mathbb{X} et \mathbb{Y} est utilisée.

Bornes du produit $\mathbb{X}\mathbb{Y}$	$0 \leq \underline{y}$	$\underline{y} < 0 < \bar{y}$	$\bar{y} \leq 0$
$0 \leq \underline{x}$	$\mathbb{X}\mathbb{Y} = [\underline{x}\underline{y}, \bar{x}\bar{y}]$	$\mathbb{X}\mathbb{Y} = [\bar{x}\underline{y}, \bar{x}\bar{y}]$	$\mathbb{X}\mathbb{Y} = [\bar{x}\underline{y}, \underline{x}\bar{y}]$
$\underline{x} < 0 < \bar{x}$	$\mathbb{X}\mathbb{Y} = [\underline{x}\bar{y}, \bar{x}\bar{y}]$	$\mathbb{X}\mathbb{Y} = [\min(\underline{x}\bar{y}, \bar{x}\underline{y}), \max(\underline{x}\bar{y}, \bar{x}\underline{y})]$	$\mathbb{X}\mathbb{Y} = [\bar{x}\underline{y}, \underline{x}\bar{y}]$
$\bar{x} \leq 0$	$\mathbb{X}\mathbb{Y} = [\underline{x}\bar{y}, \bar{x}\bar{y}]$	$\mathbb{X}\mathbb{Y} = [\underline{x}\bar{y}, \underline{x}\bar{y}]$	$\mathbb{X}\mathbb{Y} = [\bar{x}\underline{y}, \underline{x}\bar{y}]$

Tableau II-4. Expressions du produit $\mathbb{X}\mathbb{Y}$ en fonction des signes des bornes de \mathbb{X} et \mathbb{Y}

Exactement le même travail peut être réalisé sur la division en enlevant les cas où \mathbb{Y} contient l'élément nul.

Une opération arithmétique entre un intervalle \mathbb{X} et une constante α (ou une variable réelle) revient à utiliser les relations précédentes dans le cas où les deux bornes du second intervalle sont égales (intervalle dégénéré).

Opération arithmétique	Calcul des bornes de l'intervalle obtenu
Addition et Soustraction	$\mathbb{X} \text{ op } \alpha = [\underline{x} \text{ op } \alpha, \bar{x} \text{ op } \alpha]$, $\text{op} \in \{+, -\}$
Multiplication	$\alpha\mathbb{X} = [\alpha\underline{x}, \alpha\bar{x}]$ si $\alpha \geq 0$, $[\alpha\bar{x}, \alpha\underline{x}]$ sinon
Division	$\frac{\alpha}{\mathbb{X}} = \left[\frac{\alpha}{\bar{x}}, \frac{\alpha}{\underline{x}} \right]$ si $\alpha \geq 0$, $\left[\frac{\alpha}{\underline{x}}, \frac{\alpha}{\bar{x}} \right]$ sinon, avec $0 \notin \mathbb{X}$

Tableau II-5. Opérations arithmétiques entre un intervalle et une variable réelle

Dans le cas d'intervalles symétriques, les expressions des différentes opérations arithmétiques (tableau II-3) se simplifient. Si \mathbb{X} et \mathbb{Y} sont deux intervalles symétriques, les relations suivantes se vérifient :

Opération	Expression
Addition et Soustraction	$\mathbb{X} + \mathbb{Y} = \mathbb{X} - \mathbb{Y} = (\mathbb{X} + \mathbb{Y})[-1, 1] = (\bar{x} + \bar{y})[-1, 1]$
Multiplication	$\mathbb{X}\mathbb{Y} = \mathbb{X} \mathbb{Y} [-1, 1] = \bar{x}\bar{y}[-1, 1]$

Tableau II-6. Opérations arithmétiques dans le cas d'intervalles symétriques

La somme ou le produit de deux intervalles symétriques est un intervalle symétrique. Si seul l'intervalle \mathbb{X} est symétrique, l'expression du produit se simplifie tout de même puisqu'il est facile de vérifier que $\mathbb{X}\mathbb{Y} = \mathbb{X}|\mathbb{Y}| = \bar{x}|\mathbb{Y}|[-1, 1]$. Le produit d'un intervalle symétrique par un

⁶ Un intervalle est dit positif (respectivement négatif) si tous les éléments qu'il contient sont positifs (négatifs), c'est-à-dire si $\underline{x} \geq 0$ ($\bar{x} \leq 0$).

intervalle quelconque est donc symétrique. Notons que, de part sa constitution, l'inverse d'un intervalle symétrique n'existe pas.

Exemple. Comme petit exemple illustratif, montrons qu'il est toujours possible, par un changement de variables approprié, de se ramener à un intervalle symétrique de bornes imposées $\pm\alpha$. Soit l'intervalle $\mathbb{X} = [\underline{x}, \bar{x}]$, cherchons à l'exprimer sous la forme $[\underline{x}, \bar{x}] = a + b[-\alpha, \alpha]$, $a \in \mathbb{R}$, $b \in \mathbb{R}^{+*}$. L'évaluation de la partie droite de cette égalité se traduit par les deux relations suivantes :

$$\underline{x} = a - \alpha b \text{ et } \bar{x} = a + \alpha b \text{ puisque } b \text{ est positif,}$$

$$\text{soit } a = \frac{\bar{x} + \underline{x}}{2} \text{ et } b = \frac{\bar{x} - \underline{x}}{2\alpha}.$$

Finalement, le changement de variable suivant est obtenu : $[\underline{x}, \bar{x}] = c(\mathbb{X}) + \frac{w(\mathbb{X})}{2\alpha}[-\alpha, \alpha]$. ■

Puisqu'il est plus simple de travailler avec des intervalles symétriques et que le changement de variables à opérer est trivial, nous utiliserons très souvent dans les chapitres suivants des incertitudes représentées par des variables bornées normalisées, et ce sans aucune perte de généralité. Remarquons que ce choix interdit toute division par un intervalle, ce qui n'est guère gênant dans le cadre de ce mémoire dans la mesure où les problèmes d'erreurs numériques lors des calculs ne sont pas pris en compte. Le second avantage est de faciliter la formalisation des calculs. Par exemple, l'évaluation de l'intervalle $t_0 + \mathbf{t}_x^T \mathbf{x}$ où $t_0 \in \mathbb{R}$, $\mathbf{t}_x \in \mathbb{R}^n$ et \mathbf{x} est un vecteur de variables bornées normalisées vérifiant $\|\mathbf{x}\|_\infty \leq 1$, conduit au vecteur intervalle : $[t_0 - \|\mathbf{t}_x\|_1, t_0 + \|\mathbf{t}_x\|_1]$.

II.2.4 Dépendance entre variables bornées

Un des problèmes épineux et fondamentaux de l'arithmétique des intervalles est le problème de dépendance pouvant conduire à des majorations très importantes des bornes calculées, même pour des fonctions intervalle rationnelles⁷ somme toute très simples. Classiquement, le terme de dépendance est employé dès qu'une expression contient plusieurs occurrences d'un même intervalle.

Plus exactement, c'est la dépendance de la variable bornée associée qui est en cause (Adrot et al, 1999a, 1999b), (Armengol et al, 1999b), car deux intervalles ne peuvent être qualifiés de dépendants puisque ce ne sont que des ensembles. En effet, la dépendance de deux variables bornées x et y ne s'affirme pas au niveau de leurs bornes ; ainsi, elles peuvent être parfaitement indépendantes et pourtant fluctuer sur un même support, auquel cas \mathbb{X} et \mathbb{Y} sont égaux. Inversement, lors d'opérations sur deux intervalles \mathbb{X} et \mathbb{Y} , l'arithmétique des intervalles ne permet pas de prendre en compte la dépendance entre les variables bornées x et y qu'ils représentent. À partir du moment où un intervalle \mathbb{X} n'est associé qu'à une seule

⁷ Une fonction intervalle rationnelle est une fonction définie par une séquence finie d'opérations arithmétiques sur les intervalles (+, -, ., /) (Moore, 1979).

variable bornée x ⁸, une occurrence multiple de \mathbb{X} dans une expression est donc synonyme de perte d'information car les liens entre les variables x ne sont plus considérés. Ce problème, soulignant un inconvénient de l'analyse par intervalles, est donc lié à une dépendance entre variables bornées ou de manière équivalente d'indépendance entre intervalles.

Exemple. Pour mettre en évidence le problème de dépendance, il suffit de s'intéresser à la fonction réelle $f(x) = x - x$, $x \in \mathbb{X}$. Il est clair que, quelle que soit la valeur de la variable réelle x , la fonction f est nulle et l'intervalle contenant exactement $f(x)$ devrait se résumer à l'unique valeur nulle $[0,0]$. Cependant, lorsque l'arithmétique des intervalles est utilisée pour évaluer l'expression $\mathcal{Y} = \mathbb{X} - \mathbb{X}$, les bornes obtenues sont : $\mathcal{Y} = [\underline{x} - \bar{x}, \bar{x} - \underline{x}]$ (tableau II-3). Si l'intervalle \mathcal{Y} contient bien $f(x)$, il ne conduit pas au résultat escompté et induit une majoration d'autant plus importante que la longueur $w(\mathbb{X})$ est grande. La différence est due au fait que lorsque \mathcal{Y} est calculé, la dépendance entre les deux variables bornées soustraites n'est pas prise en considération. C'est l'ensemble $\mathcal{Y} = \{x - z / x \in \mathbb{X}, z \in \mathbb{X}\}$ qui est évalué à la place de $\mathcal{Y} = \{x - x / x \in \mathbb{X}\}$. Pour une valeur x élément de \mathbb{X} lui a été soustraite une valeur z de \mathbb{X} pouvant être différente ; c'est donc le calcul $\mathcal{Y} = \mathbb{X} - \mathbb{Z}$ qui est réalisé, \mathbb{Z} étant un intervalle possédant les mêmes bornes que \mathbb{X} . ■

Remarquons qu'une dépendance n'entraîne pas de manière systématique une surestimation de l'intervalle recherché. Supposons qu'un intervalle \mathbb{X} apparaisse plusieurs fois dans une expression \mathcal{Y} . Il n'y aura aucun pessimisme lors de l'évaluation de \mathcal{Y} si pour chaque terme de \mathcal{Y} faisant intervenir l'intervalle \mathbb{X} , les valeurs de la variable x utilisées pour le calcul de la borne supérieure (respectivement inférieure) sont identiques.

Exemple. C'est le cas de l'addition par exemple où la somme $\mathbb{X} + \mathbb{X}$ est donnée par $[\underline{x} + \underline{x}, \bar{x} + \bar{x}]$. Chaque borne obtenue ne s'exprime qu'en fonction d'une seule et même valeur de la variable bornée x (en l'occurrence respectivement \underline{x} et \bar{x}) ; de cette manière, la dépendance entre variables bornées est implicitement prise en compte. ■

De plus, il est préférable d'utiliser, lorsqu'ils existent, des opérateurs spécifiques prenant en compte les dépendances. C'est le cas par exemple de la fonction intervalle⁹ puissance (tableau II-7) remplaçant avantageusement la multiplication lorsque celle-ci s'applique à un même intervalle.

\mathbb{X}^n	Conditions à vérifier
$[\underline{x}^n, \bar{x}^n]$	si $\underline{x} > 0$ ou si n est impair
$[\bar{x}^n, \underline{x}^n]$	si $\bar{x} < 0$ et si n est pair
$[0, \mathbb{X} ^n]$	si 0 est élément de x et si n est pair

Tableau II-7. Bornes de \mathbb{X}^n

⁸ Ce qui signifie que si une autre variable bornée y possède les mêmes bornes, elle sera néanmoins représentée par un second intervalle \mathcal{Y} .

⁹ Nous appelons fonction intervalle, toute fonction dont les arguments et l'image sont des intervalles.

Exemple. Intéressons-nous au carré d'une variable bornée x élément de $\mathbb{X} = [-1,2]$. Le produit de \mathbb{X} par lui-même en utilisant l'opérateur multiplication (tableau II-3) conduit à l'intervalle $\mathbb{X}\mathbb{X} = [-2,4]$ alors que le carré \mathbb{X}^2 donne le résultat attendu $[0,4]$. La différence provient encore du problème de dépendance entre variables bornées. Dans le premier cas, c'est l'ensemble $\{xy / x \in [-1,2] \text{ et } y \in [-1,2]\}$ qui est calculé alors que le second correspond à $\{xx = x^2 / x \in [-1,2]\}$. Cet exemple montre qu'il faut prendre des précautions avant d'employer les diverses expressions explicitant les opérations arithmétiques sur des intervalles à occurrence multiple afin de limiter le pessimisme sur les bornes obtenues. ■

De manière plus générale, Moore (1979) explique que lorsque les intervalles n'apparaissent qu'une seule fois dans une fonction intervalle f rationnelle, le résultat obtenu conduit à l'intervalle minimal, c'est-à-dire à l'intervalle contenant exactement l'ensemble image de f . Il est donc fortement conseillé, dans le cas d'une expression relativement simple, de la modifier de manière à réduire le nombre d'occurrences d'intervalles identiques.

Exemple. L'étude des variations de la fonction réelle $f(x) = \frac{1}{1-x+x^2}$, $x \in \mathbb{X} = [0,2]$ permet de trouver les extrema suivants : $\frac{1}{3}$ et $\frac{4}{3}$. Considérons la fonction intervalle rationnelle suivante : $\mathbb{F}(\mathbb{X}) = \frac{1}{1-\mathbb{X}+\mathbb{X}^2}$. Le calcul de son dénominateur donne $[-1,5]$; \mathbb{F} ne conduit donc pas à un intervalle borné pour l'intervalle \mathbb{X} spécifié. En revanche, l'évaluation de la fonction $\mathbb{G}(\mathbb{X}) = \frac{4}{3+(1-2\mathbb{X})^2}$, dont la fonction réelle correspondante $g(x) = \frac{4}{3+(1-2x)^2}$, $x \in \mathbb{X}$, est équivalente à f , conduit à l'intervalle minimal $\left[\frac{1}{3}, \frac{4}{3}\right]$. ■

II.2.5 Propriétés algébriques de l'arithmétique des intervalles

II.2.5.1 Propriétés conservées

Les propriétés algébriques de l'arithmétique des intervalles se déduisent de celles rencontrées dans le cas de variables réelles. Néanmoins certaines propriétés diffèrent à cause du problème de dépendance. Ainsi, deux fonctions rationnelles équivalentes pour l'arithmétique des nombres réels, le seront aussi pour l'arithmétique des intervalles si chaque variable n'apparaît qu'une seule fois de chaque côté de l'égalité (Neumaier, 1990). Par exemple, la commutativité et la transitivité sont conservées pour l'addition et la multiplication. Soient \mathbb{X} , \mathbb{Y} et \mathbb{Z} trois intervalles associés à des variables bornées indépendantes, alors les propriétés suivantes sont vérifiées :

	Addition	Multiplication
Commutativité	$x + y = y + x$	$xy = yx$
Associativité	$x + (y + z) = (x + y) + z$	$x(yz) = (xy)z$

Tableau II-8. Associativité et commutativité des opérateurs arithmétiques

Les éléments neutres et absorbants sont les suivants :

	Addition (Soustraction)	Multiplication
Élément neutre	$0 \pm x = \pm x + 0 = \pm x$	$1x = x1 = x$
Élément absorbant		$0x = x0 = 0$

Tableau II-9. Éléments neutres et absorbants

Le fait que ces opérations reposent sur des intervalles à occurrence simple empêche de définir un élément absorbant pour l'addition. En effet, considérons un intervalle x défini par (II-1). L'intervalle y est un élément absorbant pour x dans le cadre de l'addition si $x + y = [0,0]$, ce qui conduit à $\underline{y} = -\underline{x}$ et $\bar{y} = -\bar{x}$, et donc à la contradiction suivante $\underline{y} \geq \bar{y}$.

Ajouté à cela, le lecteur pourra aisément vérifier les propriétés suivantes :

$x - y = x + (-y) = -y + x$	$-(x - y) = y - x$	$x - (y \pm z) = (x - y) \mp z$
$\frac{x}{y} = xy^{-1} = y^{-1}x$	$x(-y) = (-x)y = -xy$	$(-x)(-y) = xy$

Tableau II-10. Quelques propriétés supplémentaires

ainsi que ces implications supplémentaires :

$x + z = y + z \Leftrightarrow x = y$
$x + z \subseteq y + z \Leftrightarrow x \subseteq y$
$xy = [0,0] \Leftrightarrow x = [0,0] \text{ ou } y = [0,0]$
$0 \in xy \Leftrightarrow 0 \in x \text{ ou } 0 \in y$
$xx = yz, 0 \notin z \Rightarrow x = y$

Tableau II-11. Quelques implications supplémentaires

II.2.5.2 Propriétés non-conservées

À l'opposé, si \mathcal{F} et \mathcal{G} sont deux fonctions intervalle rationnelles dont les expressions sont équivalentes pour l'arithmétique des nombres réels, alors l'inclusion $\mathcal{F}(x) \subseteq \mathcal{G}(x)$ est toujours vérifiée si aucun intervalle n'apparaît plus d'une fois dans \mathcal{F} . Ainsi, la loi de distributivité n'est en général pas vérifiée dans le cas de fonctions intervalle. En revanche, l'inclusion suivante est toujours satisfaite (propriété de sous-distributivité) :

$$x(y \pm z) \subseteq xy \pm xz \text{ ou } (z \pm y)x \subseteq zx \pm yx.$$

Seules certaines contraintes particulières permettent de conserver la distributivité :

$$(1): x(y + z) = xy + xz \quad (2): x(y - z) = xy - xz$$

- si le produit $\mathbb{Y}\mathbb{Z}$ est positif ou nul pour (1), négatif ou nul pour (2), ce qui découle directement des propriétés de l'addition et de la multiplication,
- si \mathbb{X} est une variable intervalle dégénérée (variable réelle),
- si les variables intervalle \mathbb{Y} et \mathbb{Z} sont symétriques ; en effet, dans ce cas l'égalité suivante est vérifiée :

$$\mathbb{X}(\mathbb{Y} + \mathbb{Z}) = |\mathbb{X}|\mathbb{Y} + \mathbb{Z}[-1, 1] = |\mathbb{X}|\mathbb{Y}[-1, 1] + |\mathbb{X}|\mathbb{Z}[-1, 1] = |\mathbb{X}|\mathbb{Y} + |\mathbb{X}|\mathbb{Z} = \mathbb{X}\mathbb{Y} + \mathbb{X}\mathbb{Z}.$$

Enfin, un certain nombre de règles valides dans le cas de variables réelles ne fonctionnent plus pour des intervalles, en l'occurrence les lois conduisant à des simplifications de variables :

$\mathbb{X} - \mathbb{X} = [\underline{x} - \bar{x}, \bar{x} - \underline{x}]$	$\frac{\mathbb{X}}{\mathbb{X}} = \left[\frac{\underline{x}}{\bar{x}}, \frac{\bar{x}}{\underline{x}} \right]$
$\mathbb{X} - \mathbb{Y} \subseteq (\mathbb{X} + \mathbb{Z}) - (\mathbb{Y} + \mathbb{Z})$	$\frac{\mathbb{X}}{\mathbb{Y}} \subseteq \frac{(\mathbb{X}\mathbb{Z})}{(\mathbb{Y}\mathbb{Z})}$

Tableau II-12. Quelques propriétés de l'arithmétique des intervalles

II.2.6 Fonction intervalle

L'objectif de cette section est de présenter les principales définitions et théorèmes conduisant à l'une des propriétés essentielles de l'analyse par intervalles, à savoir la propriété de garantie. Si son contenu, bref rappel de ce que développa initialement Moore R.E., se révèle familier à tous ceux habitués à manipuler l'outil intervalle, il nous a cependant semblé intéressant d'en réexpliquer les fondements pour les non-initiés, d'autant plus qu'il n'est que peu utilisé en diagnostic.

II.2.6.1 Extension d'une fonction réelle

Définition 1. Soit f une fonction vectorielle réelle à n arguments réels x_i , $i \in \{1, \dots, n\}$. Lorsque chaque variable x_i est bornée et appartient à un intervalle \mathbb{X}_i , l'ensemble image¹⁰ de f est donné par :

$$\tilde{f}(\mathbb{X}_1, \dots, \mathbb{X}_n) = \{f(x_1, \dots, x_n), x_i \in \mathbb{X}_i, i \in \{1, \dots, n\}\}$$

L'ensemble image peut ne pas être un intervalle (par exemple une réunion d'intervalles dans le cas d'une division par un intervalle contenant 0) auquel cas c'est le plus petit intervalle contenant \tilde{f} , appelé intervalle minimal (Didrit, 1997), qui est considéré.

Dans le cadre de ce document appliqué au diagnostic, nous ne serons amenés qu'à considérer des fonctions pour lesquelles l'intervalle minimal de f correspond exactement à l'ensemble image et est donc naturellement noté \tilde{f} . Pour des fonctions réelles, \tilde{f} est obtenu

¹⁰ Moore (1979) utilise le terme anglo-saxon *united extension* (en référence à la notion d'extension que nous définirons ultérieurement) qu'il définit sous la forme $\tilde{f}(\mathbb{X}) = \bigcup_{x \in \mathbb{X}} f(x)$.

en recherchant les extrema de f sur le domaine d'étude. Les techniques traditionnellement employées sont celles issues de l'optimisation non-linéaire (méthode du Gradient, de Newton-Raphson,...) pouvant conduire à des optima locaux. L'approche ensembliste, comme nous allons le voir, permet d'obtenir un résultat garanti, c'est-à-dire un intervalle englobant l'intervalle minimum, l'égalité pouvant être obtenue lorsque certaines conditions sont respectées.

Remarque 1. Dans le cas d'intervalles \mathbb{X} et \mathbb{Y} représentatifs de variables bornées indépendantes, les expressions présentées dans le tableau II-3 conduisent par définition à des intervalles minimaux des quatre opérations arithmétiques réelles (la justification est donnée par la relation (II-2)). L'intervalle minimal de $x \circ y$ ($x \in \mathbb{X}$, $y \in \mathbb{Y}$, $\circ \in \psi$) est donc défini par $\mathbb{X} \circ \mathbb{Y}$.

Remarque 2. En revanche, les opérateurs arithmétiques sur les intervalles ne conduisent pas nécessairement à un intervalle minimal lorsque l'un des arguments a une occurrence multiple (reprendre l'exemple de la section II.2.4 avec la fonction intervalle $\mathbb{Y} = \mathbb{X} - \mathbb{X}$).

Définition 2. Soit f une fonction vectorielle réelle à n arguments réels x_i , $i \in \{1, \dots, n\}$. Une extension intervalle de f est une fonction intervalle \mathbb{F} à n variables intervalle¹¹ \mathbb{X}_i , possédant la propriété suivante pour des arguments réels :

$$\mathbb{F}(x_1, \dots, x_n) = f(x_1, \dots, x_n),$$

pour tout n -uplet réel (x_1, \dots, x_n) compris dans $(\mathbb{X}_1, \dots, \mathbb{X}_n)$.

\mathbb{F} est donc une fonction intervalle à valeurs réelles quand ses arguments sont réels et qui coïncide alors avec la fonction réelle f .

Remarque 3. Il n'existe pas d'extension intervalle unique d'une fonction réelle. En effet, si \mathbb{F} est une extension intervalle de la fonction réelle f , alors la fonction intervalle $\mathbb{F}_1(\mathbb{X}) = \mathbb{F}(\mathbb{X}) + \mathbb{X} - \mathbb{X}$ l'est aussi puisque, pour tout x élément de \mathbb{X} , on a $\mathbb{F}_1(x) = \mathbb{F}(x) = f(x)$.

Remarque 4. Deux extensions intervalle, équivalentes pour l'arithmétique des nombres réels, ne le sont pas forcément pour l'arithmétique des intervalles. Prenons par exemple la fonction réelle $f(x) = x(1-x) = x - xx$. Pour $\mathbb{X} = [0,1]$, les extensions intervalle $\mathbb{X}(1-\mathbb{X})$ et $\mathbb{X} - \mathbb{X}\mathbb{X}$ conduisent respectivement à $[0,1]$ et $[-1,1]$. Remarquons que les résultats obtenus sont pessimistes à cause du problème de dépendance puisqu'une étude des variations de la fonction réelle f permet de trouver l'intervalle minimal suivant : $[0,0.25]$.

Définition 3. Soit f une fonction rationnelle d'un vecteur x de \mathbb{R}^n . En remplaçant les variables réelles par les variables intervalle associées et les opérations arithmétiques réelles par celles issues de l'arithmétique des intervalles, une extension \mathbb{F} rationnelle particulière est obtenue et appelée extension intervalle naturelle de f .

Définition 4. Une extension intervalle minimale est une extension intervalle particulière conduisant à l'intervalle minimal \tilde{f} .

¹¹ Traditionnellement, le terme variable intervalle remplace le mot intervalle pour exprimer l'idée d'argument de fonction.

Remarque 5. Une extension intervalle naturelle ou minimale étant par définition une extension, les remarques 3 et 4 restent valides.

II.2.6.2 Extension intervalle monotone par inclusion

Remarque 6. Pour toute fonction réelle f , l'ensemble image (et donc l'intervalle minimal) satisfait à la propriété d'inclusion :

$$\text{si } \mathbb{Y}_i \subseteq \mathbb{X}_i, i \in \{1, \dots, n\}, \text{ alors } \tilde{f}(\mathbb{Y}_1, \dots, \mathbb{Y}_n) \subseteq \tilde{f}(\mathbb{X}_1, \dots, \mathbb{X}_n).$$

Cependant, cette propriété n'est pas forcément respectée pour les fonctions intervalle. Moore (1979) adopte donc la définition suivante.

Définition 5. Une fonction intervalle \mathbb{F} à n variables intervalle $\mathbb{X}_1, \dots, \mathbb{X}_n$ est dite monotone par inclusion si pour tous les intervalles \mathbb{Y}_i vérifiant $\mathbb{Y}_i \subseteq \mathbb{X}_i, i \in \{1, \dots, n\}$, l'inclusion suivante est respectée :

$$\mathbb{F}(\mathbb{Y}_1, \dots, \mathbb{Y}_n) \subseteq \mathbb{F}(\mathbb{X}_1, \dots, \mathbb{X}_n).$$

Remarque 7. Les extensions intervalle minimales du fait de la remarque 6 sont donc monotones par inclusion.

Remarque 8. Puisque les opérations arithmétiques sur les intervalles sont des extensions intervalle minimales dans le cas d'occurrences simples (remarque 1), elles sont donc monotones par inclusion.

Remarque 9. L'inclusion étant une relation transitive, on déduit de la remarque précédente que les fonctions intervalle rationnelles sont elles aussi monotones par inclusion.

Remarque 10. Toutes les fonctions intervalle ne sont pas monotones par inclusion. Par exemple, le calcul de la fonction $\mathbb{F}(\mathbb{X}) = c(\mathbb{X}) + \mathbb{X}$ pour $\mathbb{X} = [0,2]$ conduit à $[1,3]$. Ce résultat n'englobe pas l'intervalle $[0.5,1.5]$ obtenu pour $\mathbb{X} = [0,1]$.

La propriété de monotonie par inclusion conduit au théorème suivant proposé par Moore (1979), fondement de l'analyse par intervalles.

Théorème 1. Soit f une fonction vectorielle réelle à n arguments réels $x_i, i \in \{1, \dots, n\}$. Si \mathbb{F} est une extension intervalle monotone par inclusion de la fonction f , à variables intervalle $\mathbb{X}_1, \dots, \mathbb{X}_n$, alors la propriété suivante est vérifiée :

$$\tilde{f}(x_1, \dots, x_n) \subseteq \mathbb{F}(\mathbb{X}_1, \dots, \mathbb{X}_n), \forall x_i \in \mathbb{X}_i, i \in \{1, \dots, n\}.$$

Puisque le n -uplet réel (x_1, \dots, x_n) est élément de $(\mathbb{X}_1, \dots, \mathbb{X}_n)$ et que \mathbb{F} est monotone par inclusion, la relation suivante est satisfaite $\mathbb{F}(x_1, \dots, x_n) \subseteq \mathbb{F}(\mathbb{X}_1, \dots, \mathbb{X}_n)$. De plus, puisque \mathbb{F} est une extension, l'égalité $\mathbb{F}(x_1, \dots, x_n) = f(x_1, \dots, x_n)$ permet d'aboutir à la conclusion que toutes les valeurs de $f(x_1, \dots, x_n)$ sont contenues dans $\mathbb{F}(\mathbb{X}_1, \dots, \mathbb{X}_n)$ pour tout x_i élément de $\mathbb{X}_i, i \in \{1, \dots, n\}$.

Nous adopterons désormais la terminologie employée par Didrit (1997) consistant à remplacer le terme "extension intervalle monotone par inclusion" par "fonction d'inclusion". Le théorème précédent signifie donc que, quelle que soit la fonction d'inclusion \mathbb{F} de la fonction réelle f étudiée, l'évaluation de \mathbb{F} conduit à une majoration de l'intervalle minimal \tilde{f} (c'est-à-dire du domaine exact des valeurs prises par f). Comme en diagnostic nous rechercherons les enveloppes de résidus délimitant le domaine de bon fonctionnement du système, l'utilisation exclusive de fonctions d'inclusion permettra de garantir les résultats en assurant que tous les comportements supposés normaux sont bien appréhendés. C'est sur la propriété de garantie que repose essentiellement le bien fondé de l'approche bornante dans le cadre du diagnostic.

Remarque 11. Ce théorème n'impose aucune condition sur l'occurrence des variables intervalle dans une fonction d'inclusion.

Corollaire 1. Le théorème 1 et la remarque 9 permettent de déduire que si \mathbb{F} est une fonction intervalle rationnelle, extension de la fonction réelle f , alors :

$$\tilde{f}(x_1, \dots, x_n) \subseteq \mathbb{F}(\mathbb{X}_1, \dots, \mathbb{X}_n), \forall x_i \in \mathbb{X}_i, i \in \{1, \dots, n\}.$$

L'évaluation d'une fonction intervalle rationnelle \mathbb{F} conduit donc à un intervalle contenant les valeurs de la fonction réelle associée f quand ses arguments réels x_i sont éléments des intervalles \mathbb{X}_i .

Corollaire 2 (Moore, 1979). D'après les définitions des opérations arithmétiques sur les intervalles et le théorème 1, une fonction d'inclusion d'une fonction réelle rationnelle f , dans laquelle chaque variable intervalle \mathbb{X}_i n'apparaît qu'une seule fois et seulement à la puissance 1, et pour laquelle n'existe aucune division par une variable intervalle contenant la valeur nulle, permet de déterminer l'intervalle minimal contenant les valeurs de f .

En conclusion, le point important à retenir est que, l'évaluation de n'importe quelle fonction d'inclusion d'une fonction réelle f conduit à un intervalle majorant l'intervalle minimal. Si cette extension est rationnelle et vérifie les conditions du corollaire 2, alors les bornes obtenues sont celles de l'intervalle minimal.

II.2.6.3 Fonctions élémentaires

Neumaier (1990) définit par fonction élémentaire f , les éléments d'un ensemble prédéfini Φ de fonctions réelles (valeur absolue, fonction puissance, racine carrée, exponentielle, logarithme, sinus, cosinus, arctangente,...), continues pour chaque valeur x comprise dans tout intervalle \mathbb{X} faisant partie de leurs domaines de définition. Comme pour les opérations arithmétiques, il est possible d'étendre ces fonctions dans le cadre des intervalles :

$$\mathbb{F}(\mathbb{X}) = \{f(x) / x \in \mathbb{X}\}.$$

Dans le cas de fonctions f réelles et monotones, la fonction intervalle précédente possède des bornes définies comme suit :

$$\mathbb{F}(\mathbb{X}) = [f(\underline{x}), f(\bar{x})] \text{ si la fonction } f \text{ est croissante, } \mathbb{F}(\mathbb{X}) = [f(\bar{x}), f(\underline{x})] \text{ sinon.}$$

Remarque 12. Les bornes obtenues sont alors celles de l'intervalle minimal de f ; dans ces conditions, \mathbb{F} est une extension intervalle minimale. De plus, la remarque 7 permet d'affirmer que c'est aussi une fonction d'inclusion.

Le cas des fonctions réelles racine carrée, exponentielle et logarithme peut ainsi être facilement traité :

Fonction réelle	Extension intervalle minimale
$\exp(x)$	$\exp(\mathbb{X}) = [\exp(\underline{x}), \exp(\bar{x})]$
$\ln(x), x > 0$	$\ln(\mathbb{X}) = [\ln(\underline{x}), \ln(\bar{x})], \mathbb{X} > 0$
$\sqrt{x}, x \geq 0$	$\mathbb{X}^{1/2} = [x^{1/2}, \bar{x}^{1/2}], \mathbb{X} \geq 0$

Tableau II-13. Extensions intervalle minimales de fonctions élémentaires

Certaines propriétés de fonctions réelles peuvent être conservées grâce à l'inclusion monotone lors de la construction d'extensions. Par exemple, les ensembles image des fonctions réelles sinus et cosinus sont comprises entre -1 et 1 . Pour respecter cette propriété, une extension du type $[-1,1] \cap \mathbb{F}(\mathbb{X})$ où $\mathbb{F}(\mathbb{X})$ contient la fonction réelle pour x élément de \mathbb{X} peut être construite. Il est aussi possible de tenir compte du sens de variation sur un domaine restreint.

Fonction réelle	Extension intervalle minimale
$\sin(x)$	$\sin(\mathbb{X}) = [\sin(\underline{x}), \sin(\bar{x})], \mathbb{X} \subseteq \left[-\frac{\pi}{2}, \frac{\pi}{2}\right]$ $\sin(\mathbb{X}) = \left[\min_{x \in \mathbb{X}}(\sin(x)), \max_{x \in \mathbb{X}}(\sin(x)) \right], \forall \mathbb{X}$
$\cos(x)$	$\cos(\mathbb{X}) = [\cos(\underline{x}), \cos(\bar{x})], \mathbb{X} \subseteq [0, \pi]$ $\cos(\mathbb{X}) = \left[\min_{x \in \mathbb{X}}(\cos(x)), \max_{x \in \mathbb{X}}(\cos(x)) \right], \forall \mathbb{X}$

Tableau II-14. Extensions intervalle minimales des fonctions sinus et cosinus

Notons que le tableau II-7 donne une extension intervalle minimale notée \mathbb{X}^n , de la fonction réelle puissance x^n .

De manière plus générale, la connaissance des minimum et maximum globaux de la fonction f sur le domaine d'étude \mathbb{X} permet d'en déterminer l'intervalle minimal¹² :

$$\mathbb{F}(\mathbb{X}) = \left[\inf_{x \in \mathbb{X}} f(x), \sup_{x \in \mathbb{X}} f(x) \right].$$

Notons que la notion de continuité de la fonction f est primordiale pour que \mathbb{F} soit un intervalle. De même, si une fonction $f(x,y)$ dépend de deux variables réelles x et y et qu'elle est monotone vis-à-vis de x (dérivée partielle première par rapport à x de signe constant), alors :

$$\mathbb{F}(\mathbb{X}, \mathbb{Y}) = \left[\inf_{y \in \mathbb{Y}} (f(x, y)), \sup_{y \in \mathbb{Y}} (f(\bar{x}, y)) \right] \text{ si la fonction } f \text{ est croissante,}$$

¹² En supposant bien évidemment que ce soit bien un intervalle.

$$\mathbb{F}(\mathbb{X}, \mathbb{Y}) = \left[\inf_{y \in \mathbb{Y}} (f(\bar{x}, y)), \sup_{y \in \mathbb{Y}} (f(\underline{x}, y)) \right] \text{ sinon.}$$

II.2.6.4 Application aux fonctions intervalle arithmétiques

Définition 6. Neumaier (1990) définit une expression (ou fonction) arithmétique comme étant une fonction réelle composée d'opérations arithmétiques (éléments de ψ) et de fonctions élémentaires (éléments de \mathcal{D}) sur les réels. La notion d'expression arithmétique $f(\mathbf{x})$ s'étend au cas de fonctions intervalle $\mathbb{F}(\mathbb{X})$ en remplaçant les variables réelles x_i par les variables intervalle \mathbb{x}_i et les fonctions élémentaires et opérations arithmétiques par leurs extensions intervalle minimales.

Remarque 13. Les fonctions intervalle rationnelles définies par Moore (1979) correspondent à des fonctions intervalle arithmétiques sans fonction élémentaire.

Théorème 2. À partir de la remarque 12, Neumaier (1990) démontre qu'une fonction intervalle arithmétique est monotone par inclusion.

Remarque 14. À partir du théorème précédent, on déduit qu'un polynôme, aussi bien qu'un rapport de polynômes, ayant pour arguments des variables intervalle, sont à inclusion monotone.

Corollaire 3. Neumaier (1990) étend le corollaire 2 proposé par Moore (1979) dans le cadre de fonctions rationnelles au cas de fonctions arithmétiques. Soit \mathbb{F} une fonction intervalle arithmétique. Si les variables intervalle n'apparaissent qu'une seule fois dans l'expression analytique de \mathbb{F} , alors \mathbb{F} est une extension minimale de la fonction réelle f associée et elle permet de déterminer l'intervalle minimal contenant les valeurs de f .

En conclusion, l'analyse par intervalles permet de calculer simplement une majoration de l'ensemble des valeurs prises par une fonction arithmétique réelle pour des arguments appartenant à un certain domaine d'étude. Dans le cadre du diagnostic par approche bornante, le domaine de bon fonctionnement représentant l'ensemble des comportements jugés sains du système est évalué à l'aide de l'analyse par intervalles. Il est clair que pour limiter le nombre de non-détections, il est alors nécessaire de s'intéresser à des extensions intervalle dont les évaluations se rapprocheront le plus possible de l'intervalle minimal. Plus spécifiquement, la présence d'une ou plusieurs variables intervalle à occurrence multiple intervenant dans l'expression de \mathbb{F} , conduit généralement à un pessimisme des bornes. Une première solution consiste à choisir judicieusement les extensions intervalle en limitant au maximum les occurrences de variables intervalle identiques. Une autre solution, présentée dans la section II.3, consiste à mettre en œuvre des méthodes récursives qui détermineront à chaque pas de meilleures valeurs des majorants et convergeront le plus rapidement possible vers l'intervalle minimal. Dans le cas où une fonction intervalle arithmétique ne contient que des variables intervalle à occurrence simple, la détermination de l'intervalle minimal ne pose aucune difficulté.

Exemple. Considérons la fonction réelle $p(x) = 1 - x^2 + \frac{1}{5}x^4$, $x \in [-1, 2]$. L'étude des variations de cette fonction montre qu'elle possède un minimum en $\sqrt{\frac{5}{2}}$ valant $-\frac{1}{4}$ et un maximum en 0 valant 1 sur le domaine d'étude considéré.

	$-\infty$	$-\sqrt{\frac{5}{2}}$	-1	0	$\sqrt{\frac{5}{2}}$	2	∞
Signe de la dérivée		-	+	-	+		
Variations de $p(x)$	∞		$\frac{1}{5}$	1	$\frac{1}{5}$		∞
		$-\frac{1}{4}$			$-\frac{1}{4}$		

Tableau II-15. Tableau de variations de $p(x)$

Le tableau suivant propose un certain nombre d'extensions intervalle de p :

Extension choisie	Valeur de l'intervalle obtenue
$\mathcal{P}_1(\mathbb{X}) = 1 - \mathbb{X}^2 + \frac{1}{5}\mathbb{X}^4$	$\left[-3, \frac{21}{5}\right] = [-3, 4, 2]$
$\mathcal{P}_2(\mathbb{X}) = 1 - \mathbb{X}\mathbb{X} + \frac{1}{5}\mathbb{X}\mathbb{X}\mathbb{X}\mathbb{X}$	$\left[\frac{-23}{5}, \frac{31}{5}\right] = [-4, 6, 2]$
$\mathcal{P}_3(\mathbb{X}) = 1 + \mathbb{X}\mathbb{X}\left(-1 + \frac{1}{5}\mathbb{X}\mathbb{X}\right)$	$\left[\frac{-23}{5}, \frac{19}{5}\right] = [-4, 6, 3, 8]$
$\mathcal{P}_4(\mathbb{X}) = \frac{1}{5}\left(\mathbb{X}^2 - \frac{5}{2}\right)^2 - \frac{1}{4}$	$\left[\frac{-1}{4}, 1\right] = [-0, 25, 1]$

Tableau II-16. Quelques extensions intervalle de p .

Pour chaque fonction d'inclusion précédente, les opérateurs arithmétiques sur les intervalles ont été utilisés sans se soucier du problème de dépendance, ce qui justifie les majorations obtenues. Le calcul $\mathbb{X}\mathbb{X}$ désigne le produit de l'intervalle \mathbb{X} par lui-même alors que \mathbb{X}^2 nécessite l'emploi des formules obtenues dans le cas de la fonction puissance. Cet exemple illustre ainsi les remarques 1, 2, 3, 4, 14 ainsi que le théorème 1 et le corollaire 2. ■

II.3 EVALUATION D'UNE FONCTION D'INCLUSION

À cause du problème de dépendance, l'évaluation d'une fonction d'inclusion engendre un pessimisme des bornes obtenues préjudiciable à une procédure de détection de défauts. En effet, imaginons que le polynôme $p(x)$ présenté dans l'exemple précédent définisse l'expression analytique d'un résidu associé à une stratégie bon fonctionnement. Son ensemble image \mathcal{P}_4 définit alors l'ensemble de ses valeurs acceptables, c'est-à-dire cohérentes avec un comportement normal du système physique. Il est donc clair qu'utiliser une des majorations \mathcal{P}_i , $i \in \{1, 2, 3\}$ à la place de \mathcal{P}_4 entraîne une perte d'information conduisant à une augmentation

du nombre de non-détections puisque les valeurs de p seront supposées représentatives d'un comportement sain alors qu'elles peuvent être réellement le signe d'un défaut.

Le cas de fonctions d'inclusion ne faisant intervenir que des variables intervalle à occurrence simple étant très restrictif et marginal, l'objectif de cette section est de recenser certaines méthodes permettant d'améliorer de manière drastique les résultats de l'évaluation d'une fonction d'inclusion quelconque. Une première solution, non-développée ici, consiste à rechercher les maximum et minimum de la fonction réelle f associée en procédant à une étude de sa dérivée (cas de fonctions simples où peu de variables interviennent), ou en utilisant des procédures itératives d'optimisation non-linéaire (méthode du Gradient, de Newton Raphson, du simplex,...). Néanmoins, ces dernières ont l'inconvénient de fournir des optima locaux, ne garantissant donc en aucune manière la solution. L'autre solution, abordée dans ce document, consiste à utiliser certaines propriétés des fonctions d'inclusion de manière à construire des procédures itératives qui convergent vers l'intervalle minimal.

II.3.1 Méthodes reposant sur des formes centrées

II.3.1.1 Topologie de l'intervalle

La propriété d'inclusion (définition 5) prônée par Moore R.E. peut être interprétée comme suit en ce qui concerne l'évaluation d'une fonction d'inclusion (Didrit, 1997) : $w(\mathbb{X}) \rightarrow 0 \Rightarrow w(\mathbb{F}(\mathbb{X})) \rightarrow 0$, où w désigne la fonction longueur (tableau II-1). Cette propriété dite de convergence explique que l'évaluation d'une fonction d'inclusion sera d'autant plus performante (c'est-à-dire proche de l'intervalle minimal) que la longueur de l'argument intervalle \mathbb{X} sera petite. Seuls les principaux résultats utiles à la compréhension de la section suivante sont explicités ; le lecteur pourra avantageusement se référer à (Moore, 1979) et (Neumaier, 1990) pour de plus amples détails.

Définition 7. Une extension intervalle \mathbb{F} est lipchitzienne en \mathbb{X}_0 s'il existe une constante k telle que l'inégalité suivante soit vérifiée :

$$w(\mathbb{F}(\mathbb{X})) \leq kw(\mathbb{X}) \text{ pour tout vecteur intervalle } \mathbb{X} \text{ inclus dans } \mathbb{X}_0.$$

Lemme 1. Si une fonction vectorielle réelle $f(x)$ est lipchitzienne sur le domaine \mathbb{X}_0 , c'est-à-dire si $|f(x) - f(y)| \leq k|x - y|$, pour tout vecteur x et y éléments de \mathbb{X}_0 , alors toute extension intervalle minimale de f est une fonction intervalle lipchitzienne en \mathbb{X}_0 .

Remarque 15. Le lemme 1 s'applique au cas des opérations arithmétiques (tableau II-3) et des fonctions élémentaires (tableaux II-7, II-13 et II-14) qui sont par conséquent lipchitziennes sur leurs domaines de définition respectifs. Si maintenant \mathbb{F} et \mathbb{G} sont deux fonctions d'inclusion lipchitziennes respectivement en \mathbb{Y}_0 et \mathbb{X}_0 où $\mathbb{G}(\mathbb{X}_0) \subseteq \mathbb{Y}_0$, alors la composition $\mathbb{H}(\mathbb{X}) = \mathbb{F}(\mathbb{G}(\mathbb{X}))$ est lipchitzienne en \mathbb{X}_0 .

Remarque 16. Soient deux intervalles \mathbb{X} et \mathbb{Y} vérifiant $\mathbb{X} \subseteq \mathbb{Y}$, alors il existe un intervalle \mathbb{E} tel que $\underline{e} \leq 0 \leq \bar{e}$ satisfaisant : $\mathbb{Y} = \mathbb{X} + \mathbb{E}$ et $w(\mathbb{Y}) = w(\mathbb{X}) + w(\mathbb{E})$. Si $\mathbb{F}(\mathbb{X})$ est une fonction

d'inclusion de $f(x)$, alors $\tilde{f}(\mathbf{x}) \subseteq \mathbb{F}(\mathbf{x})$ et il est possible d'écrire : $\mathbb{F}(\mathbf{x}) = \tilde{f}(\mathbf{x}) + \mathbb{E}(\mathbf{x})$ où $\mathbb{E}(\mathbf{x})$ est une fonction intervalle vérifiant $w(\mathbb{E}(\mathbf{x})) = w(\mathbb{F}(\mathbf{x})) - w(\tilde{f}(\mathbf{x}))$. La longueur $w(\mathbb{E}(\mathbf{x}))$ s'appelle le dépassement¹³ de $\mathbb{F}(\mathbf{x})$ et l'intervalle minimal possède donc un dépassement nul.

Définition 8. Soit \mathbf{x} un vecteur intervalle de dimension n et m un entier strictement positif. Une sous-division uniforme¹⁴ d'ordre m de \mathbf{x} consiste à décomposer chaque élément \mathbb{x}_i de ce dernier en intervalles $\mathbb{x}_{i,j}$ de même longueur définis par :

$$\mathbb{x}_{i,j} = \left[x_i + (j-1) \frac{w(\mathbb{x}_i)}{m}, x_i + j \frac{w(\mathbb{x}_i)}{m} \right], \quad \begin{array}{l} j \in \{1, \dots, m\} \\ i \in \{1, \dots, n\} \end{array},$$

ce qui conduit aux égalités suivantes :

$$\mathbb{x}_i = \bigcup_{j=1}^m \mathbb{x}_{i,j}, \quad \mathbf{x} = \left[\bigcup_{j=1}^m \mathbb{x}_{1,j} \quad \dots \quad \bigcup_{j=1}^m \mathbb{x}_{n,j} \right]^T, \quad w(\mathbb{x}_{i,j}) = \frac{w(\mathbb{x}_i)}{m}, \quad w\left(\left[\mathbb{x}_{1,j} \quad \dots \quad \mathbb{x}_{n,j} \right]^T \right) = \frac{w(\mathbf{x})}{m}.$$

Sous-division uniforme d'ordre 4

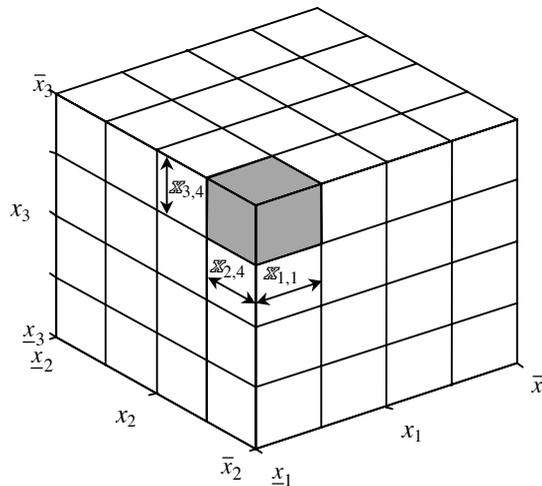


Figure II-2. Sous-division uniforme d'ordre 4 pour un vecteur intervalle de dimension 3

Théorème 3 (Moore, 1979). Si $\mathbb{F}(\mathbf{x})$ est une fonction d'inclusion lipchitzienne de la fonction vectorielle réelle $f(x)$ pour $x \in \mathbf{x} \subseteq \mathbf{x}_0$, alors le dépassement de la fonction intervalle $\mathbb{F}_m(\mathbf{x})$ définie par la réunion des intervalles obtenus en appliquant \mathbb{F} sur les différentes partitions de la sous-division uniforme d'ordre m de \mathbf{x} , est majoré à un facteur près par $w(\mathbf{x})$:

$$\mathbb{F}_m(\mathbf{x}) = \bigcup_{\substack{j_1=1, \dots, m \\ \vdots \\ j_n=1, \dots, m}} \mathbb{F}(\mathbb{x}_{1,j_1}, \dots, \mathbb{x}_{n,j_n}), \quad \text{avec } \mathbb{F}_m(\mathbf{x}) = \tilde{f}(\mathbf{x}) + \mathbb{E}_m(\mathbf{x}). \quad (\text{II-3})$$

Dans le cas de fonctions arithmétiques continues, cette réunion définit bien un intervalle et il existe alors une constante k_0 telle que :

$$w(\mathbb{E}_m(\mathbf{x})) \leq k_0 \frac{w(\mathbf{x})}{m}, \quad (\text{II-4})$$

¹³ Le terme anglo-saxon associé est "excess width".

¹⁴ Le terme anglo-saxon associé est "uniform subdivision".

ce qui revient à dire que le pessimisme occasionné par l'évaluation d'une fonction intervalle vérifiant les conditions du théorème 3 est d'autant plus faible que la longueur $w(\mathfrak{X})$ est petite ou que la sous-division uniforme est d'ordre m élevé.

II.3.1.2 Algorithme et propriété de convergence

Définition 9. Moore R.E. définit la distance $d(\mathfrak{X}, \mathfrak{Y})$ entre deux intervalles comme suit :

$$d(\mathfrak{X}, \mathfrak{Y}) = \max(|\underline{x} - \underline{y}|, |\bar{x} - \bar{y}|).$$

Remarque 17. La distance entre deux intervalles conserve les propriétés classiques de la topologie des nombres réels :

$$d(\mathfrak{X}, \mathfrak{Y}) = 0 \Leftrightarrow \mathfrak{X} = \mathfrak{Y}, \quad d(\mathfrak{X}, \mathfrak{Y}) = d(\mathfrak{Y}, \mathfrak{X}), \quad d(\mathfrak{X}, \mathfrak{Z}) \leq d(\mathfrak{X}, \mathfrak{Y}) + d(\mathfrak{Y}, \mathfrak{Z}).$$

Définition 10. Soit une séquence $\{\mathfrak{X}_k\}$ d'intervalles. Cette séquence est dite emboîtée si quel que soit k entier positif, l'inclusion suivante est satisfaite :

$$\mathfrak{X}_{k+1} \subseteq \mathfrak{X}_k.$$

Lemme 2. Une séquence emboîtée d'intervalles $\{\mathfrak{X}_k\}$ est convergente et a pour limite :

$$\mathfrak{X} = \bigcap_{k=1}^{\infty} \mathfrak{X}_k.$$

Lemme 3. Supposons que $\{\mathfrak{X}_k\}$ soit une séquence d'intervalles telle qu'il y ait un même nombre réel x contenu dans chaque élément \mathfrak{X}_k . Alors la nouvelle séquence d'intervalles $\{\mathfrak{Y}_k\}$ définie par $\mathfrak{Y}_1 = \mathfrak{X}_1$, $\mathfrak{Y}_{k+1} = \mathfrak{X}_{k+1} \cap \mathfrak{Y}_k$, $k \in \mathbb{N}^*$ est une séquence emboîtée de limite \mathfrak{Y} avec :

$$x \in \mathfrak{Y} \subseteq \mathfrak{Y}_k, \quad k \in \mathbb{N}^*.$$

À partir des définitions et lemmes précédents, Moore (1979) propose un algorithme récursif permettant d'évaluer une fonction intervalle en construisant une séquence emboîtée convergeant vers son intervalle minimal. Soit $\mathbb{F}(\mathfrak{X})$ une fonction d'inclusion d'une fonction réelle $f(x)$ lipchitzienne en \mathfrak{X}_0 ($x \in \mathfrak{X} \subseteq \mathfrak{X}_0$). Considérons la séquence d'intervalles $\{\mathfrak{Y}_k\}$ définie par :

$$\mathfrak{Y}_1 = \mathbb{F}_1(\mathfrak{X}), \quad \mathfrak{Y}_{k+1} = \mathbb{F}_{k+1}(\mathfrak{X}) \cap \mathfrak{Y}_k, \quad k \in \mathbb{N}^* \text{ où } \mathbb{F}_k \text{ est la fonction explicitée en (II-3).}$$

Par construction, $\{\mathfrak{Y}_k\}$ est une séquence emboîtée qui converge vers la limite $\mathfrak{Y} = \bigcap_{k=1}^{\infty} \mathfrak{Y}_k$ (lemme 2). L'intervalle $\mathbb{F}_k(\mathfrak{X})$, $k \in \mathbb{N}^*$, contient l'intervalle minimal de la fonction f (théorème 3). L'intervalle \mathfrak{Y}_1 renferme aussi \tilde{f} , il en est donc de même pour \mathfrak{Y}_2 et par suite pour \mathfrak{Y}_k et l'intervalle limite \mathfrak{Y} . Par définition, il est possible d'écrire :

$$\mathfrak{Y}_k = \bigcap_{i=1}^k \mathbb{F}_i(\mathfrak{X}) = \bigcap_{i=1}^k (\tilde{f}(\mathfrak{X}) + \mathbb{E}_i(\mathfrak{X})), \text{ ce qui conduit à } \mathfrak{Y} = \bigcap_{i=1}^{\infty} (\tilde{f}(\mathfrak{X}) + \mathbb{E}_i(\mathfrak{X})).$$

Or le théorème 3 explique que la longueur de \mathbb{E}_k est majorée à un facteur près par $w(\mathfrak{X})/k$. Il s'en déduit donc que $w(\mathbb{E}_\infty)$ est nulle et que l'égalité suivante est vérifiée : $\mathfrak{Y} = \tilde{f}(\mathfrak{X})$. En conclusion, la séquence $\{\mathfrak{Y}_k\}$ converge de manière asymptotique vers l'intervalle minimal contenant exclusivement les valeurs de la fonction réelle $f(x)$, x élément de \mathfrak{X} .

Comme réaliser un nombre infini d'itérations pour trouver la limite \mathcal{Y} n'est pas envisageable, il est nécessaire d'utiliser un critère d'arrêt reposant sur la distance (définition 9) entre deux intervalles \mathcal{Y}_k et \mathcal{Y}_{k-1} consécutifs. Si $d(\mathcal{Y}_k, \mathcal{Y}_{k-1})$ est inférieure à une tolérance choisie, la procédure est arrêtée et le dernier intervalle \mathcal{Y}_k calculé inclut de manière garantie l'intervalle minimal recherché. Plus k est élevé, plus faible est le pessimisme sur le résultat. De plus, pour accélérer la convergence, il est possible d'incrémenter l'ordre de la sous-division uniforme à chaque itération d'un pas supérieur à un en posant :

$$\mathcal{Y}_1 = \mathbb{F}_{p(1)}(\mathcal{X}), \quad \mathcal{Y}_{k+1} = \mathbb{F}_{p(k+1)}(\mathcal{X}) \cap \mathcal{Y}_k, \quad k \in \mathbb{N}^* .$$

où p est une fonction croissante et discrète laissée au choix de l'utilisateur, vérifiant $p(1) = 1$, $p(k) \in \mathbb{N}^*$ et fixant à chaque itération k le nombre de partitions de chaque intervalle de \mathcal{X} .

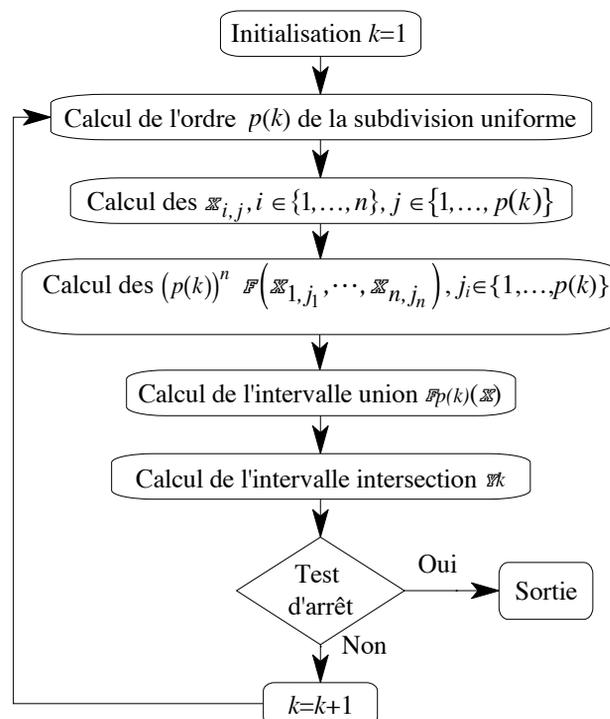


Figure II-3. Organigramme de l'algorithme récursif proposé par Moore (1979)

Exemple. Prenons la fonction réelle $f(x) = x(1-x)$, $x \in \mathcal{X} = [0,1]$. Sur ce domaine, f possède un maximum en $\frac{1}{2}$ et l'intervalle minimal associé est défini par $[0, \frac{1}{4}]$. L'extension intervalle naturelle $\mathbb{F}_1(\mathcal{X}) = \mathcal{X}(1-\mathcal{X})$ donne les résultats présentés dans le tableau II-17 pour trois fonctions différentes du pas $p(k)$ et une tolérance de 10^{-3} . Pour cela, une sous-division uniforme d'ordre $p(k)$ de l'intervalle \mathcal{X} est opérée en le décomposant en $p(k)$ intervalles \mathcal{X}_j de longueur $w(\mathcal{X})/p(k)$. Il suffit ensuite de programmer l'addition et la multiplication de deux intervalles pour évaluer tous les intervalles $\mathbb{F}_1(\mathcal{X}_j)$, $j \in \{1, \dots, p(k)\}$; puis d'en calculer la réunion pour obtenir la valeur de $\mathbb{F}_{p(k)}(\mathcal{X})$ et enfin \mathcal{Y}_k en faisant l'intersection de $\mathbb{F}_{p(k)}(\mathcal{X})$ avec \mathcal{Y}_{k-1} . Les intervalles obtenus sont donnés dans le tableau suivant avec, à titre indicatif, le nombre d'itérations effectuées et le nombre de partitions de l'intervalle \mathcal{X} qu'il a fallu réaliser lors de la dernière itération :

Tolérance	Pas $p(k)$	Nombre d'itérations	Nombre de partitions de \mathbb{X}	Intervalle obtenu
10^{-3}	k	16	17	$[0,0.2803]$
10^{-3}	2^k	8	512	$[0,0.2510]$
10^{-3}	8^k	3	4096	$[0,0.2501]$

Tableau II-17. Évaluation de la fonction \mathbb{F}_1

Le critère d'arrêt est associé à une tolérance relative puisque c'est la distance entre deux valeurs consécutives de \mathbb{Y}_k qui est testée. Ensuite, plus $p(k)$ augmente rapidement, plus grand est le nombre de partitions de \mathbb{X} pour atteindre le résultat, ce qui se traduit logiquement par une valeur plus précise mais nécessairement par un temps de calcul plus long ; un compromis est donc à trouver. Les mêmes essais sont renouvelés avec cette fois l'extension intervalle $\mathbb{F}_2(\mathbb{X}) = \mathbb{X} - \mathbb{X}^2$. Les résultats obtenus sont présentés ci-dessous :

Tolérance	Pas $p(k)$	Nombre d'itérations	Nombre de partitions de \mathbb{X}	Intervalle obtenu
10^{-3}	k	33	34	$[-0.0294,0.2794]$
10^{-3}	2^k	9	1024	$[-0.0010,0.2510]$
10^{-3}	8^k	4	32768	$[0,0.2500]$

Tableau II-18. Évaluation de la fonction \mathbb{F}_2

Globalement, pour aboutir à un résultat sensiblement identique, le nombre d'itérations, et par conséquent le nombre de partitions de \mathbb{X} , sont plus importants, ce qui se traduit, surtout dans le dernier cas, par un temps de calcul élevé. ■

L'intérêt d'une telle méthode, mis à part le fait qu'elle ne se limite pas à rechercher des optima locaux, est qu'elle permet de s'affranchir du problème de dépendance. En effet, quelle que soit l'extension intervalle \mathbb{F} choisie de la fonction vectorielle f vérifiant les contraintes imposées par le théorème 3, la limite \mathbb{Y} tend toujours vers l'intervalle minimal \tilde{f} . En revanche, comme ces extensions ne sont pas équivalentes pour l'arithmétique des intervalles, la rapidité de la convergence est différente suivant l'expression \mathbb{F} utilisée.

II.3.1.3 Forme centrée

D'après ce qui précède, le choix de l'extension intervalle étant déterminant pour l'algorithme récursif proposé ci-dessus, cette section s'intéresse aux formes centrées connues pour le faible pessimisme qu'elles engendrent (Moore, 1979). Considérons une fonction vectorielle rationnelle $f(x_1, \dots, x_n)$, $x_i \in \mathbb{X}_i$, $i \in \{1, \dots, n\}$. La forme centrée associée est obtenue en réécrivant f de la manière suivante :

$$f(x_1, \dots, x_n) = f(c_1, \dots, c_n) + g(y_1, \dots, y_n)$$

où c_i est le centre de \mathbb{X}_i (tableau II-1) et $y_i = x_i - c_i$. Cette égalité conduit à la relation suivante :

$$g(y_1, \dots, y_n) = f(y_1 + c_1, \dots, y_n + c_n) - f(c_1, \dots, c_n).$$

Pour une fonction f rationnelle, il est possible d'écrire la fonction g comme suit :

$$g(y_1, \dots, y_n) = \sum_{i=1}^n y_i h_i(y_1, \dots, y_n) \text{ où les fonctions réelles } h_i \text{ sont aussi rationnelles.}$$

Une forme centrée $\mathbb{F}_c(\mathbb{X}_1, \dots, \mathbb{X}_n)$ est alors une extension intervalle définie par :

$$\mathbb{F}_c(\mathbb{X}_1, \dots, \mathbb{X}_n) = f(c_1, \dots, c_n) + \sum_{i=1}^n \mathbb{Y}_i \mathbb{H}_i(\mathbb{Y}_1, \dots, \mathbb{Y}_n), \text{ avec } \mathbb{Y}_i = \mathbb{X}_i - c_i,$$

où \mathbb{H}_i est une fonction d'inclusion naturelle de la fonction rationnelle h_i .

La forme centrée \mathbb{F}_c est bien une extension intervalle de f contenant son ensemble image. De plus l'intérêt essentiel réside dans le fait que sa convergence est quadratique, c'est-à-dire que la longueur du dépassement \mathbb{E}_m est cette fois majorée par un terme d'ordre $w(\mathbb{X})^2$ lorsque $w(\mathbb{X})$ tend vers 0. Cette propriété conduit à un intervalle moins pessimiste que dans le cas d'une extension intervalle quelconque pour des intervalles de petites tailles¹⁵ (Didrit, 1997). Moore (1979) montre de plus que :

$$w(\mathbb{E}_m(\mathbb{X})) \leq k_0 \frac{w(\mathbb{X})}{m^2}.$$

En comparant ceci à l'inégalité (II-4), il est clair que la forme centrée permet d'accélérer la convergence de l'algorithme d'évaluation de fonctions intervalle par rapport à des extensions naturelles quelconques. Notons qu'une forme centrée fonctionne aussi dans le cas de variables élevées à la puissance (cas des polynômes par exemple).

Exemple. Reprenons la fonction réelle $f(x) = x(1-x)$, $x \in \mathbb{X} = [0,1]$. En notant $c = c(\mathbb{X}) = 0.5$, la fonction réelle $g(y)$ est définie par :

$$g(y) = f(y+c) - f(c) = y(1-2c-y), \text{ ce qui conduit à } h(y) = 1-2c-y \text{ avec } y = x-c.$$

En posant $\mathbb{Y} = \mathbb{X} - c$, la forme centrée associée à la fonction réelle f est donc :

$$\mathbb{F}_c(\mathbb{X}) = c(1-c) + \mathbb{Y}(1-2c-\mathbb{Y}) = c(1-c) + (\mathbb{X}-c)(1-c-\mathbb{X}).$$

L'algorithme récursif vu précédemment est ensuite appliqué à la fonction intervalle $\mathbb{F}_c(\mathbb{X})$. Les résultats sont présentés ci-dessous.

Tolérance	Pas $p(k)$	Nombre d'itérations	Nombre de partitions de \mathbb{X}	Intervalle obtenu
10^{-3}	k	3	4	[0,0.2778]
10^{-3}	2^k	5	64	[0,0.2510]
10^{-3}	8^k	2	512	[0,0.2500]

Tableau II-19. Résultats obtenus dans le cas de la forme centrée \mathbb{F}_c .

L'algorithme converge bien plus rapidement vers l'intervalle minimal que pour \mathbb{F}_1 (tableau II-17) et \mathbb{F}_2 (tableau II-18), et ce avec une majoration de l'intervalle minimum moindre dans cet exemple. Cependant, la longueur du dépassement dépend de celle de la variable intervalle \mathbb{X} et lorsque $w(\mathbb{X})$ est élevé, la convergence peut encore être très longue. ■

Certains auteurs ont proposé d'autres formes centrées d'ordre plus élevé. Il existe aussi d'autres types d'extensions, comme la forme à valeur moyenne, la forme bi-centrée ou celle connue sous le nom anglo-saxon de "Monotonicity Test Form", détaillées dans l'annexe A et

¹⁵ Pour des intervalles plus grands, rien ne peut être déduit.

permettant d'améliorer la rapidité de convergence. Toutes ces formes sont intéressantes dans la mesure où elles permettent d'évaluer des fonctions comportant des variables intervalle à occurrence multiple. Cependant, pour conserver cette propriété, il est nécessaire de traiter une fonction intervalle dans son ensemble. Il est en effet préjudiciable d'évaluer séparément les différentes parties d'une fonction contenant des variables intervalle à occurrence multiple car cette méthode provoque une perte d'information sur les dépendances entre variables. Supposons que $\mathcal{F}(\mathcal{X})$ soit décomposée en fonctions élémentaires. Il est possible de s'intéresser individuellement à chacune de ces portions et de leur appliquer séparément l'algorithme récursif vu précédemment. Au moment d'évaluer la fonction dans son entier, des opérations sur les résultats numériques précédents, et non sur des variables intervalle elles-mêmes, vont avoir lieu. Il ne sera alors plus possible de prendre en compte une quelconque dépendance entre les différentes parties de la fonction ; l'algorithme récursif perdant ainsi sa capacité à réduire le pessimisme du résultat obtenu.

Exemple. Considérons la fonction intervalle $\mathcal{F}(\mathcal{X}) = \mathcal{X}(1 - \mathcal{X})$, $\mathcal{X} = [0,1]$ et décomposons la de la manière suivante : $\mathcal{F}(\mathcal{X}) = \mathcal{F}_1(\mathcal{X})\mathcal{F}_2(\mathcal{X})$, $\mathcal{F}_1(\mathcal{X}) = \mathcal{X}$, $\mathcal{F}_2(\mathcal{X}) = (1 - \mathcal{X})$. L'évaluation de \mathcal{F}_1 puis de \mathcal{F}_2 à l'aide du monotonicity test form (annexe A) donne le même résultat $[0,1]$, le calcul de \mathcal{F} conduit alors à $[0,1]$ alors que l'intervalle minimal vaut $[0, \frac{1}{4}]$. ■

Cette remarque prendra tout son intérêt dans le chapitre IV dédié aux tests de cohérence. En effet, il sera alors nécessaire d'évaluer des fonctions intervalle faisant intervenir des mesures dépendant du temps, si bien que leurs coefficients fluctuent continuellement. Pour limiter la quantité de calculs à effectuer, il aurait été intéressant de décomposer ces fonctions par morceaux, chacun pouvant être factorisé par une mesure. Évaluer individuellement et à l'avance chaque morceau pour ensuite rassembler en ligne les résultats multipliés par les valeurs courantes des mesures aurait permis de diminuer considérablement les temps de calcul puisque ceux à réaliser en temps réel ne font intervenir que des opérations élémentaires. Cependant, à cause du problème de dépendance, ce choix est inadapté.

Le catalogue d'extensions intervalle proposées ci-dessus n'est pas exhaustif, de nombreux chercheurs ayant travaillé sur le sujet. Pour contourner le problème de dépendance, certains d'entre eux comme Comba et al (1993), Andrade et al (1994), De Figueiredo (1996a), De Figueiredo et al (1996b) travaillant dans le domaine du graphisme (modélisation de solides, intersection de surfaces,...), ont proposé une méthode appelée arithmétique affine qui permet de prendre en compte, en partie, les dépendances entre variables bornées. Grâce à l'information supplémentaire tenant compte des corrélations entre variables, les précédents auteurs affirment que, de manière générale, des domaines beaucoup plus proches des intervalles minimaux sont obtenus, voire égaux si les fonctions intervalle sont initialement linéaires en les variables intervalle. En contrepartie, cette technique est plus complexe à utiliser que l'arithmétique classique des intervalles définie par Moore R.E. et demande des temps de calcul plus élevés. De plus, même si les variables intervalle ne sont pas à occurrence multiple, cette méthode peut tout de même conduire à une majoration des bornes. Par rapport aux techniques basées sur les formes centrées, l'arithmétique affine a l'avantage de ne pas être

une procédure itérative. Cependant, à cause des approximations réalisées pour travailler uniquement sur des formes affines, entraînant nécessairement que certaines dépendances ne soient pas prises en compte, le résultat obtenu est généralement une majoration du domaine recherché, souvent plus faible que pour l'arithmétique des intervalles, mais plus élevée que dans le cas des formes centrées.

Pour terminer, remarquons qu'il est aussi possible d'utiliser une forme centrée (ou autre) sans implémenter l'algorithme récursif dans le cas où le nombre de variables intervalle engendre un temps de calcul prohibitif. L'utilisateur a alors tout intérêt à rechercher parmi les extensions proposées celle conduisant à la plus faible majoration de l'intervalle minimal. Ce choix ne nécessite pas la connaissance de ce dernier, il suffit simplement de conserver la forme correspondant à la plus petite longueur de l'intervalle obtenu. Une autre possibilité consiste à prendre l'intersection des différentes formes centrées envisagées.

II.3.2 Procédures de réduction et d'élimination

II.3.2.1 Principes

L'algorithme récursif exposé précédemment repose sur une sous-division uniforme nécessitant de diviser, à la $k^{\text{ième}}$ itération, tous les intervalles \mathbb{X}_i , $i \in \{1, \dots, n\}$ un nombre de fois $p(k)$ imposé, entraînant le calcul de $(p(k))^n$ images afin de déterminer $\mathbb{F}_{p(k)}(\mathbb{X})$ (II-3). Cette méthode induit un nombre de calculs élevé dès que la dimension du problème n ou le nombre k d'itérations deviennent importants. Pour limiter ce problème, d'autres stratégies permettent de réduire au fur et à mesure le domaine de recherche \mathbb{X} .

De manière générale, la grande majorité des algorithmes reposant sur l'analyse par intervalles exploitent le principe fondateur, qui consiste en une partition du domaine de recherche (splitting-domain en anglais). L'objectif est alors de travailler sur un ensemble fini de domaines simples (c'est-à-dire faciles à manipuler) de plus en plus petits de manière à améliorer la précision des solutions trouvées. Contenant un nombre infini de points, ils permettent d'accélérer l'exploration d'un domaine d'étude par rapport à une méthode ponctuelle plus traditionnelle. Dans le cas où les entités étudiées sont des orthotopes alignés encore appelés pavés ou boîtes, nous parlerons de méthode de pavage.

L'efficacité du principe de partition est confrontée à un obstacle majeur résidant dans la complexité des calculs avec la croissance de la dimension du domaine d'étude. Dans ces conditions, il est possible de ne diviser que des portions du domaine supposées intéressantes, c'est-à-dire répondant au problème posé, et d'éliminer toutes les autres, ou bien de chercher à réduire le domaine d'étude en se focalisant sur un domaine plus restreint inclus dans le précédent et conservant certaines propriétés attendues.

Trois principes, propres à l'approche ensembliste peuvent être utilisés (Jaulin, 2000) :

- si une propriété est vérifiée pour tous les sous-ensembles d'une partition de l'espace de recherche, alors celle-ci s'applique à l'espace de recherche tout entier,
- si un sous-ensemble ne contient aucune solution au problème traité, il peut être éliminé de l'espace de recherche,
- s'il n'est pas prouvé qu'un sous-ensemble ne contient aucune solution de ce problème, il est possible de le diviser en plusieurs sous-ensembles qu'il sera nécessaire d'analyser ultérieurement.

De manière générale, les algorithmes reposant sur l'analyse par intervalles peuvent être divisés en deux catégories selon leurs principes de fonctionnement : les méthodes d'élimination et les méthodes de réduction. Les méthodes d'élimination représentent une partie importante des approches par intervalles, leur objectif étant de limiter le nombre de calculs à effectuer en rejetant toute partie du domaine étudié ne satisfaisant au problème à résoudre. Nous supposons disposer d'un traitement à base d'analyse par intervalles, cohérent avec le problème à traiter et qui aboutit à la classification d'un pavé comme non satisfaisant, solution ou indéterminé.

Principe général d'une procédure d'élimination
<p>Entrées :</p> <ul style="list-style-type: none"> un domaine de recherche sous forme d'un pavé \mathcal{P}_0, un ensemble \mathcal{L} regroupant les pavés en attente de traitement, initialement $\mathcal{L} = \mathcal{P}_0$, un ensemble solution \mathcal{S} initialement vide. <p>Procédure principale :</p> <p>Tant que l'ensemble \mathcal{L} n'est pas vide faire</p> <ul style="list-style-type: none"> Prendre un pavé \mathcal{P} dans l'ensemble \mathcal{L}, Réaliser le traitement du pavé \mathcal{P}, Si \mathcal{P} est non satisfaisant, alors l'éliminer, Sinon si \mathcal{P} est solution du problème, alors le mettre dans \mathcal{S}, Sinon diviser \mathcal{P} et mettre les sous-pavés résultants dans \mathcal{L}. <p>Fin tant que.</p> <p>Sortie :</p> <ul style="list-style-type: none"> Ensemble \mathcal{S}.

Tableau II-20. Procédure d'élimination (Didrit, 1997)

Une procédure d'élimination repose sur un raisonnement binaire tout ou rien, à savoir qu'un pavé est soit entièrement éliminé, soit entièrement conservé pour être divisé par la suite. La procédure de réduction est différente et s'appuie sur le problème suivant : à partir d'un pavé \mathcal{P}_0 , doit être déterminé un autre pavé \mathcal{P} inclus dans le premier, le plus petit possible, tel que toutes les solutions éventuellement contenues dans \mathcal{P}_0 le soient aussi dans \mathcal{P} .

Principe général d'une méthode de réduction	
Entrées :	<p>un domaine d'étude sous forme d'un pavé \mathcal{P}_0,</p> <p>un ensemble \mathcal{L} regroupant les pavés en attente de traitement, initialement $\mathcal{L} = \mathcal{P}_0$,</p> <p>un ensemble solution \mathcal{S} initialement vide.</p>
Procédure principale :	<p>Tant que l'ensemble \mathcal{L} n'est pas vide faire</p> <p style="padding-left: 2em;">Prendre un pavé \mathcal{P} dans l'ensemble \mathcal{L},</p> <p style="padding-left: 2em;">Construire \mathcal{P}' à partir de \mathcal{P}, tel que, $\mathcal{P}' \subset \mathcal{P}$,</p> <p style="padding-left: 2em;">Si \mathcal{P}' est vide, alors \mathcal{P}' est inacceptable,</p> <p style="padding-left: 2em;">Sinon si \mathcal{P}' est solution du problème, alors le mettre dans \mathcal{S},</p> <p style="padding-left: 2em;">Sinon diviser \mathcal{P} et mettre les sous-pavés résultants dans \mathcal{L}.</p> <p>Fin tant que.</p>
Sortie :	Ensemble \mathcal{S} .

Tableau II-21. Procédure de réduction (Didrit, 1997)

Lorsqu'un pavé \mathcal{P}' est inacceptable, le pavé associé \mathcal{P} n'est pas éliminé ; si \mathcal{P}' est solution, il est alors placé dans l'ensemble solution \mathcal{S} et s'il est indéterminé, il est subdivisé et ses "enfants" sont stockés dans \mathcal{L} de manière à pouvoir éventuellement prendre une décision lors d'une prochaine itération.

II.3.2.2 Application à l'évaluation de fonctions d'inclusion

Les méthodes proposées reposent sur l'analyse par intervalles et sont utilisées pour évaluer l'intervalle minimal d'une fonction vectorielle f en lieu et place des techniques itératives à base de formes centrées. Leur intérêt, comme nous le verrons, est de limiter le nombre de calculs en ne subdivisant que des pavés jugés intéressants, c'est-à-dire susceptibles de contenir un optimum sans effectuer une partition de tout le domaine d'étude. Dans la littérature, ces méthodes sont qualifiées de méthodes d'optimisation globale dans la mesure où elles offrent la garantie que les optima déterminés sont bien globaux, à une tolérance choisie près.

La première méthode proposée, version simplifiée de ce que Hansen (1992) a proposée, est une procédure d'élimination nécessitant une fonction vectorielle différentiable f dépendant de n variables réelles x_1, \dots, x_n regroupées dans le vecteur \mathbf{x} . De plus, des fonctions d'inclusion \mathbb{F} et \mathbb{F}' de f et de sa dérivée doivent être connues. Seule la recherche du minimum de f (sans contrainte) lorsque \mathbf{x} prend des valeurs situées dans un domaine de recherche \mathcal{X} (Tibken et al, 1999) est explicitée ici. Nous travaillerons sur une liste $\mathcal{L} = \{(\gamma_1, \mathcal{X}_1), \dots, (\gamma_r, \mathcal{X}_r)\}$ constituée de couples $(\gamma_i, \mathcal{X}_i)$ où γ_i est un nombre réel et \mathcal{X}_i un vecteur intervalle inclus dans \mathcal{X} . Le pavé \mathcal{X}_i est une sous-division du domaine de recherche \mathcal{X} , l'objectif étant de ne conserver dans \mathcal{L} que des pavés susceptibles de contenir le minimum. Cette liste est ordonnée : $\gamma_1 \leq \dots \leq \gamma_r$ et le terme γ_i est obtenu en calculant une valeur minorant la borne inférieure de l'intervalle minimal de la

fonction f sur le pavé \mathfrak{X} . Du fait du problème de dépendance, l'intervalle $\mathbb{F}(\mathfrak{X})$ est en fait pessimiste ; si la borne inférieure de $\mathbb{F}(\mathfrak{X})$ est susceptible de minorer le minimum recherché, il est alors intéressant de conserver $\mathbb{F}(\mathfrak{X})$ pour le réduire ultérieurement. Issu de ce constat, le principe de la méthode proposée consiste à encadrer le minimum par un intervalle, qui va tendre asymptotiquement vers la valeur recherchée. La convergence sera d'autant plus rapide que l'extension intervalle \mathbb{F} sera pertinente (une forme centrée sera ainsi préférable à une simple extension naturelle).

Lors de l'initialisation, la liste \mathcal{L} est définie par : $\mathcal{L} = \{(\gamma_1 = \underline{f}(\mathfrak{X}_1), \mathfrak{X}_1 = \mathfrak{X})\}$, sachant que $\underline{f}(\mathfrak{X})$ correspond à la borne inférieure de $\mathbb{F}(\mathfrak{X})$. D'après la propriété d'inclusion, $\underline{f}(\mathfrak{X})$ minore nécessairement le minimum de f : $\underline{f}(\mathfrak{X}) \leq \min_{x \in \mathfrak{X}}(f(x))$. De plus, le réel f_{min} désigne l'image du centre de \mathfrak{X} par la fonction vectorielle f : $f_{min} = f(c(\mathfrak{X}))$. Par définition, f_{min} appartient nécessairement à l'intervalle minimal de f , d'où l'encadrement suivant :

$$\underline{f}(\mathfrak{X}) \leq \min_{x \in \mathfrak{X}}(f(x)) \leq f_{min}.$$

L'objectif des étapes suivantes est d'augmenter le minorant $\underline{f}(\mathfrak{X})$ et de diminuer le majorant f_{min} afin de réduire l'encadrement en deçà d'un seuil choisi par l'utilisateur.

- (1) Prendre le premier élément $(\gamma_1, \mathfrak{X}_1)$ de \mathcal{L} et l'éliminer de cette liste.
La liste \mathcal{L} étant ordonnée, γ_1 représente la plus petite valeur des bornes inférieures calculées sur chaque partition de \mathfrak{X} conservée dans \mathcal{L} .
- (2) Calculer le terme $f_c = f(c(\mathfrak{X}_1))$. Si $f_c < f_{min}$, alors imposer $f_{min} = f_c$, puis éliminer de la liste \mathcal{L} tous les couples $(\gamma_i, \mathfrak{X}_i)$ tels que $\gamma_i > f_{min}$.
En effet, f_c est nécessairement contenue dans l'intervalle minimal de f et majore donc son minimum. Tous les couples associés à un γ_i supérieur à f_c deviennent donc obsolètes ; γ_i ne pouvant plus être un minorant de $\min_{x \in \mathfrak{X}}(f(x))$ puisqu'au moins un élément de l'intervalle minimal lui est inférieur. De plus, f_c permet de réduire l'encadrement en se substituant à l'ancienne valeur de f_{min} .
- (3) Évaluer la fonction d'inclusion $\mathbb{F}(\mathfrak{X}_1)$ du gradient de f sur le pavé \mathfrak{X}_1 . Si le vecteur nul n'appartient pas au vecteur intervalle $\mathbb{F}(\mathfrak{X}_1)$, oublier \mathfrak{X}_1 et reprendre à l'étape (1).
L'objectif est de savoir si un extremum est dans le pavé \mathfrak{X}_1 . Du fait de la propriété d'inclusion, si l'origine fait réellement partie de l'intervalle image du gradient de f sur \mathfrak{X}_1 , alors elle est nécessairement contenue dans $\mathbb{F}(\mathfrak{X}_1)$. De même, si l'origine est en dehors de $\mathbb{F}(\mathfrak{X}_1)$, elle ne peut appartenir à l'intervalle minimal du gradient de f sur \mathfrak{X}_1 et dans ce cas le pavé \mathfrak{X}_1 ne nous intéresse plus.
- (4) Subdiviser \mathfrak{X}_1 en \mathfrak{X}_{11} et \mathfrak{X}_{12} tels que : $\mathfrak{X}_1 = \mathfrak{X}_{11} \cup \mathfrak{X}_{12}$ par exemple en partitionnant selon la plus grande longueur des intervalles contenus dans \mathfrak{X}_1 . Calculer les bornes inférieures $\underline{f}(\mathfrak{X}_{1i})$, $i \in \{1,2\}$ associées à \mathbb{F} sur les pavés \mathfrak{X}_{1i} . Si $\underline{f}(\mathfrak{X}_{1i})$ est inférieure à f_{min} , alors le couple $(\gamma_i = \underline{f}(\mathfrak{X}_{1i}), \mathfrak{X}_{1i})$ est inséré dans le bon ordre dans la liste \mathcal{L} , sinon il est éliminé puisque le minimum recherché ne peut être dans \mathfrak{X}_{1i} par définition.
- (5) Arrêter l'algorithme si la différence $f_{min} - \gamma_1$ est inférieure à un certain seuil fixé par l'utilisateur.

Tableau II-22. Méthode d'élimination pour une minimisation sans contrainte

Les pavés non intéressants, c'est-à-dire ne pouvant contenir le minimum, sont éliminés au fur et à mesure. Seules les zones autour des points où la dérivée s'annule sont étudiées et divisées. Les maxima et minima non-globaux sont éliminés car à un moment, les γ_i associés

deviendront supérieurs à f_{min} . D'autre part, c'est le couple associé à γ_1 , c'est-à-dire au plus petit des γ_i , qui est étudié à chaque itération, zone où il est le plus probable de trouver le minimiseur global (il n'est cependant pas certain que ce minimiseur se trouve toujours dans \mathcal{X}_1 ; à cause du pessimisme sur les bornes calculées, $\min_{x \in \mathcal{X}}(f(x))$ n'est pas forcément dans le pavé \mathcal{X}_1 dont l'image par \mathbb{F} conduit à la plus petite borne inférieure $\underline{f}(\mathcal{X}_1)$).

Dans le cas d'une forme centrée \mathbb{F} , puisque la surestimation des bornes dépend de la longueur du domaine de recherche, celle-ci va diminuer au fur et à mesure des itérations et γ_1 va s'accroître et tendre vers la valeur attendue. L'algorithme, proposé ici dans sa version la plus simpliste, ne trouve que le minimum d'une fonction correspondant à une dérivée nulle (et non à une des bornes du domaine de recherche \mathcal{X}).

L'algorithme proposé ci-dessus nécessite la différentiabilité de la fonction vectorielle f . D'autres auteurs se sont intéressés au cas de fonctions ne comportant aucune hypothèse de continuité et de différentiabilité, la seule contrainte étant d'en posséder une fonction d'inclusion. La nature même de la méthode repose comme ci-dessus sur l'encadrement du minimum à rechercher et appartient à la catégorie des algorithmes qualifiés de *branch and bound*. Le principe est exactement le même, le calcul de la fonction d'inclusion générant automatiquement l'encadrement du minimum :

$$\underline{f}(\mathcal{X}) \leq \min_{x \in \mathcal{X}}(f(x)) \leq \max_{x \in \mathcal{X}}(f(x)) \leq \bar{f}(\mathcal{X}), \quad f(x) \in \left\{ \mathbb{F}(\mathcal{X}) = \left[\underline{f}(\mathcal{X}), \bar{f}(\mathcal{X}) \right] \right\}, \quad x \in \mathcal{X}$$

avec, lorsque la longueur de \mathcal{X} tend vers 0, la longueur de $\mathbb{F}(\mathcal{X})$ converge vers 0.

Principe général d'une méthode *branch and bound*

Entrées :

Initialiser la liste \mathcal{L} avec le domaine de recherche \mathcal{P}_0 ,
calculer $\underline{f}(\mathcal{P}_0)$ et $\bar{f}(\mathcal{P}_0)$, puis poser $\underline{f} = \underline{f}(\mathcal{P}_0)$ et $\bar{f} = \bar{f}(\mathcal{P}_0)$.

Procédure principale :

Tant que $\bar{f} - \underline{f}$ est supérieure à une précision désirée **faire**

Prendre un pavé \mathcal{P} dans l'ensemble \mathcal{L} ,

Partager \mathcal{P} en n sous-pavés $\mathcal{P}_1, \dots, \mathcal{P}_n$,

Calculer $\underline{f} = \min\{\underline{f}, \underline{f}(\mathcal{P}_1), \dots, \underline{f}(\mathcal{P}_n)\}$,

Calculer $\bar{f} = \min\{\bar{f}, \bar{f}(\mathcal{P}_1), \dots, \bar{f}(\mathcal{P}_n)\}$,

Mettre les sous-pavés résultants dans \mathcal{L} ,

Éliminer tous les éléments \mathcal{P}_i de \mathcal{L} tels que $\underline{f}(\mathcal{P}_i) > \bar{f}$.

Fin tant que.

Sortie :

Encadrement du minimum.

Tableau II-23. Méthode *branch and bound* (Didrit, 1997).

De nombreuses modifications améliorant sensiblement les performances peuvent être apportées. Il est ainsi possible de trier la liste \mathcal{L} des pavés à traiter selon l'ordre croissant de

\underline{f} , ce qui permet de positionner en tête le pavé le plus prometteur à découper en priorité. Certains chercheurs préfèrent ordonner la liste \mathcal{L} suivant la taille des pavés qui la composent, afin de découper uniformément toutes les régions du domaine de recherche. Ce choix permet de déterminer des propriétés de convergence plus précises que pour les autres algorithmes, mais dégrade en général les performances. De plus, il est préférable d'utiliser $f_c = f(c(\mathcal{P}))$ plutôt que $\bar{f} = \bar{f}(\mathcal{P})$, ce qui assure un encadrement plus fin puisque $\bar{f} \geq f_c$ et restreint les temps de calcul en évitant l'évaluation plus lourde de \bar{f} . Didrit (1997) propose un algorithme prenant en compte ces diverses améliorations qu'il nomme Opticrit. Le tri s'effectue suivant les valeurs de \underline{f} dans la liste \mathcal{L} , les pavés inacceptables sont éliminés à l'aide de la meilleure des valeurs de \underline{f} au centre des pavés rencontrés et la division des pavés est réalisée par rapport à son côté de plus grande taille.

Algorithme Opticrit (Didrit, 1997)
<p>Entrées :</p> <p>Pavé de recherche \mathcal{P}_0 dans l'espace des paramètres, fonction d'inclusion \mathcal{P} de la fonction vectorielle f, tolérance ε sur la taille des pavés considérés, calculer le terme : $f_c = \bar{f}(c(\mathcal{P}_0))$¹⁶, initialiser la liste \mathcal{L} avec le domaine de recherche \mathcal{P}_0.</p> <p>Procédure principale :</p> <p>Tant qu'un pavé de \mathcal{L} a une taille supérieure à ε faire</p> <p style="padding-left: 40px;">Extraire le premier pavé \mathcal{P} de \mathcal{L} tel que sa taille dépasse ε, Partager \mathcal{P} en 2 sous-pavés \mathcal{P}_1 et \mathcal{P}_2,</p> <p style="padding-left: 40px;">Si $\underline{f}(\mathcal{P}_1) \leq f_c$ alors placer \mathcal{P}_1 dans \mathcal{L}, poser $f_c = \min(f_c, \bar{f}(c(\mathcal{P}_1)))$,</p> <p style="padding-left: 40px;">Si $\underline{f}(\mathcal{P}_2) \leq f_c$ alors placer \mathcal{P}_2 dans \mathcal{L}, poser $f_c = \min(f_c, \bar{f}(c(\mathcal{P}_2)))$,</p> <p style="padding-left: 40px;">Si f_c est modifié, alors éliminer tous les pavés \mathcal{P}' tels que $\underline{f}(\mathcal{P}') > f_c$.</p> <p>Fin tant que.</p> <p>Sortie :</p> <p>Encadrement du minimum.</p>

Tableau II-24. Algorithme Opticrit.

Tous les algorithmes présentés dans cette section dans le cadre d'un problème d'optimisation globale appartiennent à la catégorie des méthodes d'élimination. Certaines nécessitent moins de contraintes sur les fonctions étudiées (aucune hypothèse de différentiabilité), elles utilisent par conséquent moins d'information sur la fonction étudiée et restent sensiblement moins efficaces pour des problèmes de complexité significative.

¹⁶ Ce calcul est effectué à l'aide de la fonction d'inclusion \mathcal{P} plutôt que de la fonction f simplement pour tenir compte des erreurs numériques et ainsi garantir la valeur de f_c .

Des méthodes de réduction plus sophistiquées, s'appuyant sur les travaux de Hansen (1992) existent. Elles sont nettement plus complexes que celles reposant sur le principe d'élimination et les expliciter même succinctement nécessiterait de leur consacrer au moins un chapitre, ce qui dépasse largement le cadre de cette thèse. Notons simplement que l'algorithme de Hansen s'appuie sur l'hypothèse que la fonction vectorielle f étudiée est deux fois différentiable (et que des fonctions d'inclusion de chaque composante du gradient g et de certains éléments du hessien H de f sont disponibles), l'objectif étant de rechercher la valeur de son minimum. Comme tout minimiseur global de f est solution de $g(x)=0$, le principe consiste à résoudre ce système tout en conservant à l'esprit que tout minimiseur local ou maximiseur de f est également solution de ce même système. Le but est alors d'essayer de réduire autant que possible un pavé \mathcal{X} en un autre \mathcal{X}' qui contienne de manière garantie tous les minimiseurs globaux de f sur \mathcal{X} . De plus, l'algorithme de réduction de Hansen repose sur deux autres procédures élémentaires pour lesquelles il est nécessaire de disposer d'une borne supérieure du minimum recherché. Plus précisément, l'objectif, pour un pavé courant \mathcal{X} , est de déterminer un point de \mathcal{X} qui soit le plus proche possible du minimiseur de f sur \mathcal{X} . Le problème du choix de cette borne supérieure f_{max} porte le nom d'optimisation du point d'estimation locale et permet par exemple d'éliminer tout pavé \mathcal{X} tel que $\underline{f}(\mathcal{X}) > f_{max}$.

De manière générale, cet algorithme s'appuie sur la génération de formes centrées ainsi que sur un ensemble de techniques intervalle utilisées traditionnellement pour la détermination des zéros d'une fonction (méthode de Newton par intervalles, optimisation du point d'expansion de l'itération de Newton en le prenant le plus proche possible d'une solution du système) ou la résolution de systèmes linéaires (méthode de Gauss-Seidel par intervalles, élimination de Gauss par intervalles à l'aide d'une factorisation LU, préconditionnement). L'ensemble de ces procédures élémentaires (test de monotonie, de convexité, itération de Newton et dans une moindre mesure l'optimisation du point d'estimation locale) constituent les composantes de base des algorithmes intervalle fonctionnant sur le principe de réduction. Le lecteur trouvera bon nombre d'informations dans l'abondante littérature dédiée à ces techniques, en particulier l'ouvrage de référence (Hansen, 1992) et le document de synthèse (Didrit, 1997).

II.4 CHAMP DE VECTEURS DE FONCTIONS D'INCLUSION

II.4.1 Problème de dépendance et notion d'espace abstrait

Dans le cadre du diagnostic de systèmes, il est bien rare qu'un seul résidu soit suffisant pour permettre la détection et surtout la localisation de tous les défauts envisagés par le cahier des charges. Par conséquent, intéressons-nous maintenant au cas d'un champ de vecteurs constitué de fonctions d'inclusion et pouvant représenter un vecteur de résidus. Le problème de dépendance entre variables bornées au sein d'une même fonction a déjà été évoqué. Un problème similaire existe dans le cadre d'un champ de vecteurs. Tant que les différentes

fonctions d'inclusion de ce champ sont indépendantes, c'est-à-dire qu'aucune variable intervalle n'intervient dans deux composantes à la fois, le formalisme des intervalles permet de parfaitement appréhender le domaine obtenu qui se révèle être un orthotope aligné avec les axes du repère paramétrique, comme nous l'avons précisé en début de chapitre (figure II-1). En revanche, dès qu'une variable intervalle apparaît dans plusieurs fonctions d'inclusion à la fois, l'ensemble image n'est généralement plus limité à un simple pavé, auquel cas le formalisme des intervalles ne permet plus de décrire ce type de domaine avec précision.

Exemple. Prenons le champ de vecteurs suivant :

$$f(x_1, x_2) = \begin{bmatrix} x_1 - x_2 \\ x_1 + x_2 \end{bmatrix}, \text{ où } x_1 \in [-1, 2] \text{ et } x_2 = [1, 3].$$

L'évaluation de l'extension intervalle naturelle $\#$ de f en utilisant directement les opérations de l'arithmétique des intervalles (c'est-à-dire sans tenir compte des dépendances entre les deux composantes) conduit à un pavé défini par $[[-4, 1] \ [0, 5]]^T$ dans l'espace des composantes f_1 et f_2 , représenté par le cadre de la figure suivante. Cependant, l'ensemble image de f est en fait un parallélogramme¹⁷ :

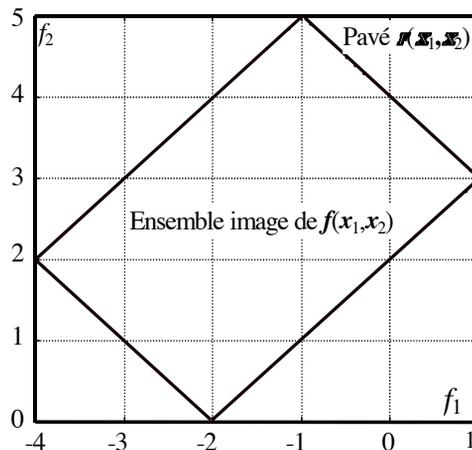


Figure II-4. Ensemble image de f et pavé $\#$

C'est pour cette raison que Ploix (1998) introduit la notion d'espace abstrait.

Définition. Si $f(x)$ est un champ de vecteurs constitué de fonctions vectorielles réelles dépendant de n variables bornées x_1, \dots, x_n , alors l'ensemble image de f pour x élément du domaine de recherche \mathcal{X} est appelé l'espace abstrait de f et est noté $\mathcal{A}(f)$.

La notion d'espace abstrait généralise simplement le formalisme des intervalles au cas de champs de vecteurs dont l'ensemble image peut être quelconque, c'est pour cette raison que nous n'utiliserons plus que ce formalisme dorénavant. Des outils complémentaires devront donc être développés pour appréhender de tels domaines. Remarquons que la distinction entre une variable bornée x et le support sur lequel elle évolue $\mathcal{A}(x)$ reste explicite et que nous conservons l'hypothèse de ne s'intéresser exclusivement qu'à des fonctions

¹⁷ Nous verrons dans le quatrième chapitre comment évaluer un tel domaine.

d'inclusion. Cette notation a été choisie pour des raisons historiques en référence aux travaux de Ploix (1998) dont s'inspire ce travail, mais notons que de manière plus conventionnelle, la dénomination employée est "ensemble image" noté $\tilde{f}(\mathbb{x})$.

II.4.2 Propriétés des espaces abstraits

Dans le cas où x est une variable bornée, l'espace abstrait $\mathcal{A}(x)$ n'est rien d'autre que l'intervalle $\mathbb{x} = [\underline{x}, \bar{x}]$; de même que si \mathbf{x} est un vecteur de variables bornées indépendantes, alors $\mathcal{A}(\mathbf{x})$ correspond à un orthotope dont les arrêtes sont alignées avec les axes du repère associés à chaque x_i . L'analyse par intervalles, telle qu'elle a été explicitée jusqu'à maintenant, peut alors être directement appliquée.

Dans le contexte d'un champ de vecteurs f , une variable bornée x_i sera qualifiée de variable commune¹⁸ si elle intervient dans plusieurs composantes de f . Si f ne comprend aucune variable commune, alors l'espace abstrait $\mathcal{A}(f)$ conduit à un orthotope aligné puisque évaluer individuellement chaque composante de f revient exactement à la même chose que de calculer $\mathcal{A}(f)$ dans son ensemble. En revanche, dans le cas contraire, le domaine $\mathcal{A}(f)$ peut devenir très complexe, car non-connexe et non-compact. Les variables communes, représentatives de dépendances entre les composantes de f induisent des contraintes imposant la forme de $\mathcal{A}(f)$ (Ploix, 1998), (Adrot et al, 1999a, 2000a).

Nature du champ de vecteurs f	Nature de l'espace abstrait $\mathcal{A}(f)$
Champ linéaire en les variables bornées	Parallélotope convexe
Champ linéaire en les variables communes, non-linéaire envers les autres variables bornées	Parallélotope non-convexe
Champ non-linéaire en les variables communes, quelconque envers les autres variables bornées	Domaine complexe (pouvant être non-connexe, non-compact) aux frontières non limitées à de simples hyperplans

Tableau II-25. Nature de $\mathcal{A}(f)$

Exemples. Le premier cas du tableau II-25 est illustré par l'exemple de la figure II-4 où $\mathcal{A}(f)$ est bien un parallélotope convexe. Les deux autres cas sont illustrés par les figures suivantes. La figure II-5 correspond à un champ de vecteurs globalement non-linéaire par rapport aux variables bornées, mais linéaire par rapport aux variables communes x_1 et x_2 :

$$\mathbf{g}(x_1, x_2) = \begin{bmatrix} x_1 x_3 - x_2 \\ x_1 + x_2 \end{bmatrix}, \text{ avec } \mathcal{A}(x_1) = [-1, 2], \mathcal{A}(x_2) = [1, 3], \mathcal{A}(x_3) = [-1, 1].$$

L'espace abstrait $\mathcal{A}(\mathbf{g})$ est bien un parallélotope non-convexe. La figure II-6 correspond à un champ de vecteurs non-linéaire par rapport à au moins une variable commune (en l'occurrence x_1) :

¹⁸ Cette terminologie permet de faire la différence entre une dépendance au sein d'une même fonction (variable dépendante) et une dépendance entre plusieurs composantes d'un champ de vecteurs (variable commune).

$$h(x_1, x_2) = \begin{bmatrix} 2x_1^2 + x_2 \\ x_1 - x_2 \end{bmatrix}, \text{ avec } \mathcal{A}(x_1) = [-1, 2], \mathcal{A}(x_2) = [1, 3].$$

Les frontières de l'espace abstrait $\mathcal{A}(h)$ sont "non-linéaires" dans le sens où ce ne sont pas de simples hyperplans, et le domaine obtenu n'est pas convexe.

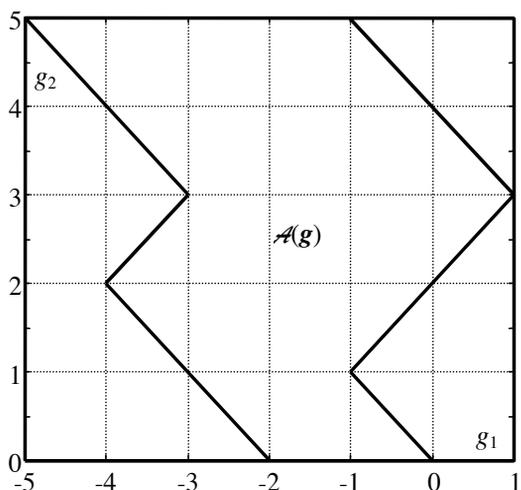


Figure II-5. Espace abstrait de g

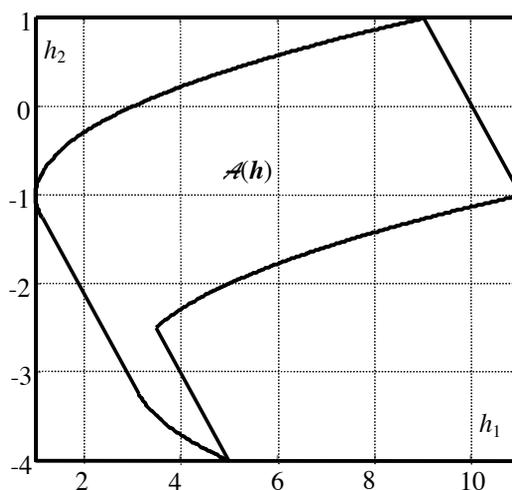


Figure II-6. Espace abstrait de h

Ne pas tenir compte de ces dépendances en supposant qu'aucune variable n'est commune et en évaluant séparément les composantes de f (c'est-à-dire appliquer sans précaution l'arithmétique des intervalles) conduit alors à construire l'orthotope aligné circonscrit de f , noté $\square \mathcal{A}(f)$ ¹⁹, qui représente en fait le majorant le plus simple de $\mathcal{A}(f)$ (Ploix, 1998). Précisons que $\square \mathcal{A}(f)$ représente le plus petit orthotope contenant $\mathcal{A}(f)$ et que dans le cas où des variables bornées dépendantes existent au sein d'une même composante de f , il faut alors avoir recours aux techniques vues dans les sections précédentes (formes centrées, optimisation globale) pour obtenir (ou approcher le plus possible) le pavé voulu.

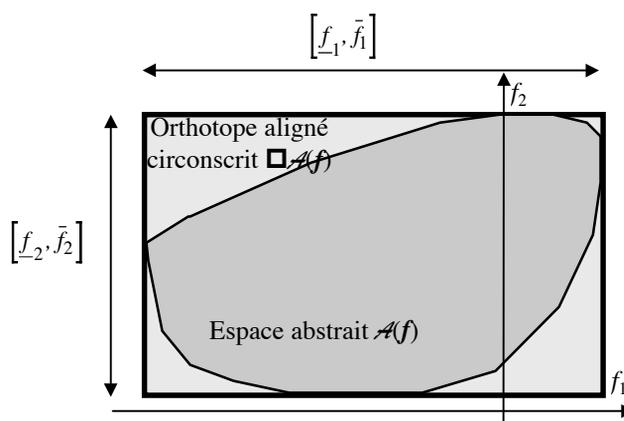


Figure II-7. $\mathcal{A}(f)$ et $\square \mathcal{A}(f)$ dans le cas d'un champ de dimension 2

Dans un premier temps, montrons donc que $\square \mathcal{A}(f)$ contient toujours $\mathcal{A}(f)$ (Adrot et al, 2000a). Pour ce faire, considérons le champ de vecteurs suivant f dans \mathbb{R}^p , où les vecteurs

¹⁹ $\square \mathcal{A}(f)$ est équivalent à la notion anglo-saxonne de Minimum Outer Box (M.O.B.) traditionnellement rencontrée.

bornés \mathbf{x}_c et $\mathbf{x}_{nc,i}$ contiennent respectivement l'ensemble des variables communes de f (c'est-à-dire présentes dans au moins deux de ses composantes) et toutes les variables non-communes présentes dans la $i^{\text{ème}}$ composante f_i (c'est-à-dire qui n'apparaissent que dans f_i) :

$$\mathbf{f}(\mathbf{x}_c, \mathbf{x}_{nc,1..p}) = \left[f_1(\mathbf{x}_c, \mathbf{x}_{nc,1}) \quad \cdots \quad f_p(\mathbf{x}_c, \mathbf{x}_{nc,p}) \right]^T.$$

Toutes les variables bornées x_i sont supposées indépendantes. L'orthotope $\square\mathcal{A}(f)$ est construit en remplaçant le vecteur \mathbf{x}_c de chaque composante f_i par un nouveau vecteur borné indépendant $\mathbf{x}_{c,i}$ possédant le même espace abstrait $\mathcal{A}(\mathbf{x}_{c,i}) = \mathcal{A}(\mathbf{x}_c)$. Le champ de vecteurs suivant est alors défini :

$$\mathbf{f}_{\square}(\mathbf{x}_{c,1..p}, \mathbf{x}_{nc,1..p}) = \left[f_1(\mathbf{x}_{c,1}, \mathbf{x}_{nc,1}) \quad \cdots \quad f_p(\mathbf{x}_{c,p}, \mathbf{x}_{nc,p}) \right]^T,$$

pour lequel ne subsiste plus aucune variable commune. Puisque toutes les composantes du champ de vecteurs \mathbf{f}_{\square} sont indépendantes, $\square\mathcal{A}(f)$ correspond à l'espace abstrait $\mathcal{A}(\mathbf{f}_{\square})$. Inversement, pour obtenir f à partir de \mathbf{f}_{\square} , les vecteurs $\mathbf{x}_{c,i}$ doivent tous être identiques. Dans ces conditions, les valeurs possibles de f appartiennent nécessairement à l'ensemble image de \mathbf{f}_{\square} ; la réciproque étant fautive à cause des contraintes précédentes. Par conséquent, $\mathcal{A}(f)$ est nécessairement contenu dans $\square\mathcal{A}(f)$.

Montrons maintenant que $\square\mathcal{A}(f)$ est circonscrit à $\mathcal{A}(f)$, c'est-à-dire qu'au moins un point de chaque hyperplan frontière de $\square\mathcal{A}(f)$ appartient à $\mathcal{A}(f)$. Considérons par exemple la $i^{\text{ème}}$ composante f_i de f et supposons que sa borne supérieure \bar{f}_i soit obtenue pour les arguments \mathbf{x}_c^* et $\mathbf{x}_{nc,i}^*$. Par conséquent, pour toutes les valeurs possibles de \mathbf{x}_c et $\mathbf{x}_{nc,i}$, l'inégalité $f_i(\mathbf{x}_c, \mathbf{x}_{nc,i}) \leq \bar{f}_i$ est satisfaite. La borne supérieure de la $i^{\text{ème}}$ composante $f_{\square,i}$ de \mathbf{f}_{\square} est :

$$\sup(f_i(\mathbf{x}_{c,i}, \mathbf{x}_{nc,i}), \quad \mathbf{x}_{c,i} \in \mathcal{A}(\mathbf{x}_{c,i}), \mathbf{x}_{nc,i} \in \mathcal{A}(\mathbf{x}_{nc,i})).$$

Puisque les deux espaces abstraits $\mathcal{A}(\mathbf{x}_c)$ et $\mathcal{A}(\mathbf{x}_{c,i})$ sont identiques, ce maximum vaut nécessairement \bar{f}_i . Puisque $\mathcal{A}(f)$ est inclus dans $\square\mathcal{A}(f)$, alors un hyperplan frontière de $\square\mathcal{A}(f)$, dont l'équation cartésienne est $f_{\square,i}(\mathbf{x}_{c,i}, \mathbf{x}_{nc,i}) = \bar{f}_i$, contient le point $f(\mathbf{x}_c^*, \mathbf{x}_{nc,i}^*)$. Cette propriété se vérifiant pour les deux bornes de chaque fonction f_i , l'orthotope aligné $\square\mathcal{A}(f)$ est donc circonscrit à $\mathcal{A}(f)$.

Exemples. Sur les exemples des figures II-4, II-5 et II-6, les domaines $\square\mathcal{A}(f)$, représentés par les cadres des figures, vérifient bien les propriétés justifiées ci-dessus. ■

Si $\mathcal{A}(f)$ représente le domaine de bon fonctionnement du système étudié, lui substituer la majoration $\square\mathcal{A}(f)$ engendre une perte d'information préjudiciable au diagnostic, puisque risquant d'accroître le nombre de non-détections. Comparer individuellement chaque résidu à son espace abstrait est donc à éviter puisque ceci correspond à déterminer $\square\mathcal{A}(f)$, l'hypothèse d'indépendance des composantes de f étant implicite. La section suivante propose une méthode permettant d'appréhender simplement $\mathcal{A}(f)$ en obtenant une majoration aussi faible que désirée dans le cas d'un champ de vecteurs f quelconque.

II.4.3 Évaluation d'un champ de vecteurs de fonctions d'inclusion

La représentation exacte ou approchée d'espaces abstraits multidimensionnels doit être d'une part suffisamment simple pour pouvoir être manipulée à des fins de diagnostic de défauts, d'autre part suffisamment précise pour décrire de manière convenable le domaine sans perdre trop d'informations. Parmi les approches traditionnelles peuvent être citées (Jaulin, 2000) la représentation par équations ou inéquations, par nuage de points ou par recouvrement.

- La représentation par équations ou inéquations permet une description exacte des espaces abstraits. Deux méthodes reposant sur cette stratégie seront développées ultérieurement dans le chapitre IV dédié aux tests de cohérence. La première repose sur la détermination des expressions analytiques des frontières d'un domaine et peut s'appliquer à des fonctions non-linéaires en les variables bornées à condition que certaines hypothèses soient vérifiées. L'inconvénient est que la manipulation d'expressions symboliques nécessite l'emploi de logiciels de calcul formel et n'est guère adaptée à l'usage que nous voulons en faire en diagnostic (tout du moins pour l'étape de détection, car nous montrerons comment la connaissance de ces expressions peut être exploitée afin d'améliorer l'efficacité de l'étape de localisation). La seconde méthode ne s'adresse qu'à des champs de vecteurs linéaires en les variables bornées. Dans ce cas, le domaine obtenu est un paralléloèdre convexe délimité par des hyperplans deux à deux parallèles générant des bandes de contraintes. Plus exactement, ce paralléloèdre est défini par l'intersection de ses bandes de contraintes et peut, de manière générale, être parfaitement décrit par un ensemble des contraintes inégalités. La méthode proposée est alors complètement numérique, simple à mettre en œuvre et parfaitement adaptée à nos objectifs, la seule contrainte éventuelle étant de linéariser au préalable les résidus vis-à-vis des incertitudes paramétriques.
- La représentation par nuage de points consiste à décrire un espace abstrait \mathcal{A} par un nombre fini de points lui appartenant. Ceux-ci peuvent être obtenus par tirage aléatoire ou par un balayage déterministe telle une sous-division uniforme de l'espace de recherche. Cette représentation est approximative dans la mesure où les différentes portions de \mathcal{A} ne sont pas détaillées avec une même précision, d'autre part rien ne garantit que les points situés aux extrémités du domaine obtenu appartiennent bien aux frontières théoriques de \mathcal{A} . Enfin, dans le cadre du diagnostic, les tests de cohérence consistent, de manière simpliste, à regarder si le point de fonctionnement courant est à l'intérieur de \mathcal{A} , principe difficile à mettre dans le cas d'une représentation par nuage de points.
- La représentation par recouvrement reprend le principe de la méthode de pavage exposée dans la section II.3.2.1 en la généralisant au cas de sous-ensembles autres que des pavés tels des ellipsoïdes ou des polyèdres convexes. Dans le cadre qui nous intéresse consistant à rechercher une image représentative du domaine abstrait d'un champ de vecteurs $f(x)$, $x \in \mathcal{X}$, le principe réside dans une partition du domaine de

recherche \mathcal{X} en sous-ensembles, dont la réunion des images recouvre complètement $\mathcal{A}(f)$.

L'approche développée dans cette section est une méthode de pavage provenant du constat qu'il est possible d'obtenir une majoration simple de $\mathcal{A}(f)$ en prenant l'orthotope aligné circonscrit $\square\mathcal{A}(f)$. Le pessimisme du résultat obtenu pouvant être important (figures II-4, II-5 et II-6), $\square\mathcal{A}(f)$ peut être avantageusement remplacé par une réunion de pavés plus petits, mais tout aussi faciles à manipuler. Cette technique fait référence au concept de pavage présenté par (Jaulin et al, 1993, 1996), (Kieffer et al, 1999), utilisé dans le cadre de l'inversion ensembliste définie par le problème suivant : soit $y=f(x)$ une fonction de \mathbb{R}^n dans \mathbb{R}^m ; connaissant un domaine compact $\mathcal{A}(y)$ de \mathbb{R}^m , l'objectif est de déterminer un ensemble de pavés de \mathbb{R}^n dont la réunion majore le domaine $\{x \in \mathbb{R}^n / f(x) \in \mathcal{A}(y)\}$. Un sous-pavage de \mathbb{R}^n est alors défini comme un ensemble compact constitué par la réunion de pavés de \mathbb{R}^n sans recouvrement et de volumes non nuls. Brièvement, la méthode consiste à réaliser un pavage du domaine de recherche de \mathbb{R}^n en pavés plus petits \mathcal{X}_i . Les pavés image $\mathcal{Y}_i = \square\mathcal{A}(f(x) / x \in \mathcal{X}_i)$ ²⁰ de chaque \mathcal{X}_i sont déterminés ; puis en testant si \mathcal{Y}_i appartient ou non à $\mathcal{A}(y)$, il est possible de trouver un encadrement du domaine de \mathbb{R}^n solution.

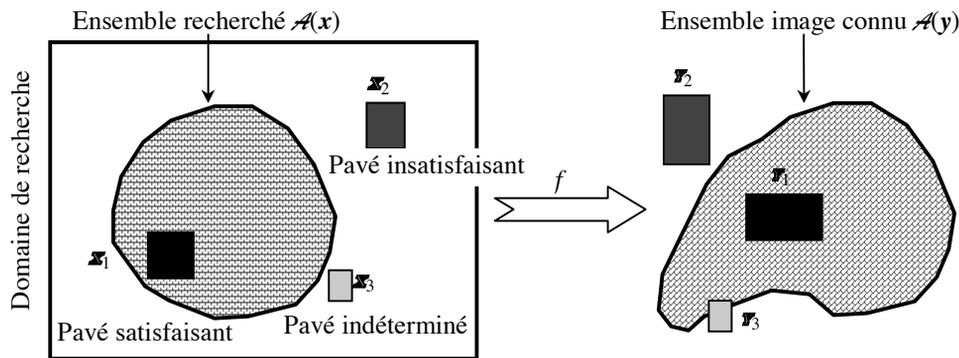


Figure II-8. Inversion ensembliste

Notre objectif est ici légèrement différent dans la mesure où c'est une majoration sous la forme d'un pavage $\mathcal{A}_p(f)$ de l'espace abstrait $\mathcal{A}(f)$ qui est recherchée (Kieffer et al, 1998), (Ploix, 1998), (Adrot et al, 2000a). Plus exactement, le domaine $\mathcal{A}_p(f)$ est défini par l'union de pavés image obtenus à partir d'une sous-division uniforme d'ordre $q \in \mathbb{N}^*$ (section II.3.1.1) de l'espace de départ $\mathcal{A}(x)$. De cette manière, le pavé $\mathcal{A}(x)$ est divisé en q^n pavés sans recouvrement $\mathcal{X}_{q,i}$, $i \in \{1, \dots, q^n\}$ de \mathbb{R}^n obtenus en découpant le domaine abstrait de chaque variable bornée x_i en q segments de même taille. Ainsi, tous les $\mathcal{X}_{q,i}$ possèdent la même taille et leur union décrit exactement $\mathcal{A}(x)$. Pour chaque pavé $\mathcal{X}_{q,i}$, l'orthotope aligné circonscrit $\square\mathcal{A}(f(x) / x \in \mathcal{X}_{q,i})$ est déterminé (c'est-à-dire sans tenir compte des dépendances entre composantes de f), leur union définissant $\mathcal{A}_p(f)$. Nous avons déjà démontré que l'espace

²⁰ Dans le cas général où le champ de vecteurs étudié contient des variables bornées dépendantes, l'orthotope aligné circonscrit est avantageusement remplacé par un pavé le contenant pour limiter les temps de calcul et éviter d'utiliser un niveau supplémentaire de techniques itératives (formes centrées, méthodes d'élimination, de réduction). Puisque de toute façon, l'orthotope aligné circonscrit est déjà une approximation de l'espace abstrait, il n'est pas utile de passer beaucoup de temps à le déterminer, puisque plus fin sera le découpage, plus les majorations dues aux problèmes de dépendances seront réduites.

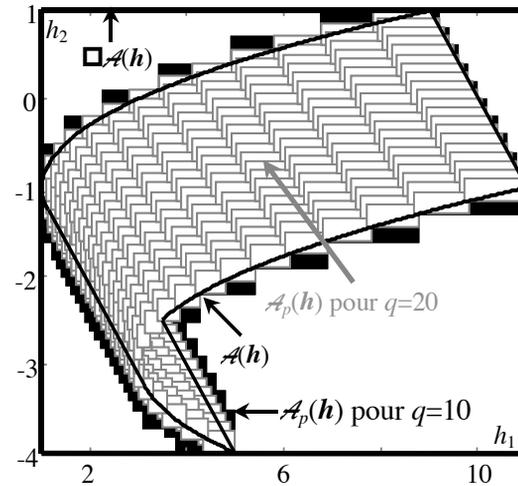
abstrait $\mathcal{A}(f(x)/x \in \mathbf{x}_{q,i})$ est inclus dans son orthotope aligné circonscrit $\square\mathcal{A}(f(x)/x \in \mathbf{x}_{q,i})$; nous avons donc l'inclusion suivante :

$$\mathcal{A}(f) = \mathcal{A}\left(f(x)/x \in \bigcup_{i=1}^{q^n} \mathbf{x}_{q,i}\right) = \bigcup_{i=1}^{q^n} \mathcal{A}(f(x)/x \in \mathbf{x}_{q,i}) \subset \bigcup_{i=1}^{q^n} \square\mathcal{A}(f(x)/x \in \mathbf{x}_{q,i}) = \mathcal{A}_p(f),$$

ce qui garantit que le pavage $\mathcal{A}_p(f)$ est bien une majoration de l'espace abstrait $\mathcal{A}(f)$. Un pavé pouvant être directement décrit par un vecteur intervalle, les opérations à effectuer peuvent alors être réalisées séparément et individuellement sur chacune de ses composantes en appliquant l'analyse par intervalles.

Contrairement à ce qui est défini par Jaulin L., le pavage $\mathcal{A}_p(f)$ proposé est généralement composé de pavés pouvant se recouvrir. En effet, seule la partition du domaine de départ $\mathcal{A}(x)$ est maîtrisée, en revanche la nature du pavage $\mathcal{A}_p(f)$ obtenu (domaine convexe, connexe, pavés sans recouvrement) dépend de la fonction f étudiée. Intuitivement, il est plus facile de suivre une forme quelconque avec une union composée de beaucoup de petits pavés, plutôt qu'une autre constituée d'un faible nombre de grands pavés. L'idée revient donc à décomposer $\mathcal{A}(f)$ en morceaux, et au lieu de majorer globalement ce domaine par $\square\mathcal{A}(f)$, réaliser des majorations de chacun de ces morceaux pour en faire la réunion. Plus la taille des morceaux sera faible, plus les majorations associées seront petites et plus l'union de ces pavés sera proche de l'espace abstrait recherché. Remarquons qu'un pavage d'ordre q' est inclus dans un pavage d'ordre $q < q'$ seulement lorsque la partition d'ordre q' est aussi une sous-division uniforme de la partition d'ordre q : $q' = 2^r q$, $r \in \mathbb{N}$ (de manière générale un pavage d'ordre $q+1$ n'est donc pas inclus dans un pavage d'ordre q). Plus grande est la valeur de q , plus petit est le pessimisme de $\mathcal{A}_p(f)$ vis-à-vis de $\mathcal{A}(f)$, mais plus important devient le temps de calcul. La précision obtenue dépend donc du nombre q de sous-divisions, mais le résultat est de toute manière bien plus précis qu'en travaillant directement sur $\square\mathcal{A}(f)$ qui correspond à $q=1$. L'avantage de cette méthode est qu'elle nécessite moins de calculs que pour la génération de nuages de points et permet de décrire la forme de $\mathcal{A}(f)$ quelle que soit la complexité du champ de vecteurs f . En revanche, du fait de la sous-division uniforme effectuée, les temps de calcul deviennent rapidement prohibitifs lorsque la dimension de l'espace de départ n s'accroît. Du fait de sa simplicité, cette méthode est qualifiée de méthode de pavage de base et sera reprise dans le chapitre IV afin d'être améliorée et adaptée au cadre du diagnostic de systèmes.

Exemple. Reprenons l'exemple de la fonction h associée à la figure II-6. Le pavage $\mathcal{A}_p(h)$ est représenté pour deux valeurs différentes de l'ordre de la sous-division uniforme : $q=10$ (en noir et en arrière plan) et $q=20$ (en gris). Pour comparer, l'orthotope aligné circonscrit $\square\mathcal{A}(h)$ ainsi que les véritables frontières de l'espace abstrait $\mathcal{A}(h)$ ont été dessinées.

Figure II-9. Domaines $\mathcal{A}_p(h)$ pour $q=10$ et $q=20$

II.4.4 Effet d'enveloppement

Un inconvénient majeur de l'utilisation de l'arithmétique des intervalles est son caractère explosif bien connu provenant du phénomène d'enveloppement²¹ dans le cas de systèmes récurrents (Ploix, 1998), (Armengol et al, 1999b), (Brdys, 1999), (Katsillis et al, 1999). Considérons un champ de vecteurs $f(x(k))$, où k représente l'indice temporel courant. Si l'espace abstrait $\mathcal{A}(x(k))$ n'est pas un orthotope aligné (c'est-à-dire que f comporte des variables communes), il est alors très difficile de l'utiliser pour évaluer l'ensemble image $\mathcal{A}(f(x(k)))$. L'utilisation classique de l'analyse par intervalles se traduit par la substitution de $\square\mathcal{A}(x(k))$ à la place de $\mathcal{A}(x(k))$, ce qui conduit au calcul de l'espace abstrait $\mathcal{A}(f(x'(k)) / x'(k) \in \square\mathcal{A}(x(k)))$ qui n'est par définition qu'une majoration de $\mathcal{A}(f(x(k)))$. Ainsi, même en déterminant exactement $\square\mathcal{A}(x(k))$ ²², le domaine obtenu à l'instant $k+1$ induit du pessimisme, qui au fil des itérations s'accumule et finit par faire diverger la récurrence, auquel cas la taille des domaines calculés explose. En effet, l'orthotope aligné circonscrit $\square\mathcal{A}(f(x'(k)) / x'(k) \in \square\mathcal{A}(x(k)))$ utilisé comme domaine de départ à l'itération suivante majore $\square\mathcal{A}(f(x(k)))$, qui englobe l'espace abstrait $\mathcal{A}(f(x(k)))$ qu'il aurait théoriquement fallu utiliser.

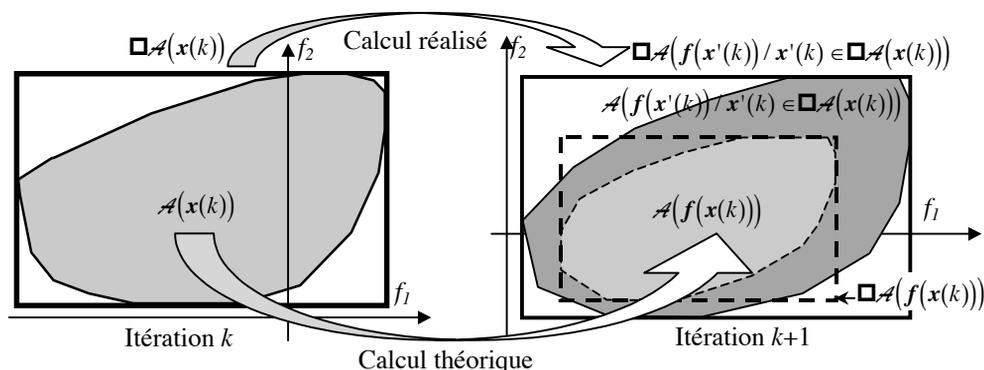


Figure II-10. Phénomène d'enveloppement

²¹ Le terme anglo-saxon correspondant est "wrapping effect".

²² Ce qui n'est pas toujours simple dans le cas de variables dépendantes au sein d'une composante et conduit généralement à un résultat obtenu à une précision imposée.

Il existe donc une majoration induite par la présence de variables communes : pour pouvoir effectuer l'itération, l'espace abstrait $\mathcal{A}(x(k))$ est remplacé par son orthotope aligné circonscrit $\square\mathcal{A}(x(k))$. Puisque cette majoration sert à obtenir l'ensemble des solutions $\mathcal{A}(f(x'(k))/x'(k) \in \square\mathcal{A}(x(k)))$ à l'instant suivant, des valeurs supplémentaires ont été ajoutées contribuant au fil du temps à l'explosion des domaines calculés, d'où la notion de phénomène d'enveloppement. En pratique, une seconde source de majoration provient de la présence de variables dépendantes qui induisent du pessimisme sur le calcul des orthotopes alignés circonscrits.

Exemple. Prenons la représentation d'état suivante :

$$x(k+1) = Ax(k) + Bu(k), \quad A = \begin{bmatrix} 0.98 & -0.05 \\ 0.9 & 0.55 \end{bmatrix}, \quad B = \begin{bmatrix} -8 \\ 1 \end{bmatrix}, \quad \mathcal{A}(x(0)) = \begin{bmatrix} [-1, 0.5] \\ [-0.5, 1] \end{bmatrix}$$

où $u(k)$, de dimension $q=1$, vaut 0.025 et $x(k)$ est un vecteur réel de dimension $n=2$. Sur la figure II-11, sont représentés les orthotopes alignés obtenus par récurrence aux instants $k+1$ lorsque le domaine utilisé à l'instant k est l'orthotope déterminé à l'itération précédente. L'effet d'enveloppement est parfaitement visible sur cet exemple dans la mesure où les boîtes aux plus grandes dimensions correspondent aux derniers instants. Or en regardant le système étudié, il est clair que les incertitudes ne portent que sur les conditions initiales. Dans le cas d'un système certain stable (matrice d'état A à valeurs propres situées à l'intérieur du cercle unité), l'influence de la valeur initiale de l'état s'estompe suivant la dynamique du système (A^k tend vers 0 quand k tend vers l'infini) :

$$x(k) = A^k x(0) + \sum_{i=0}^{k-1} A^{k-i-1} Bu(i). \quad (\text{II-5})$$

De même, dans le cas incertain, les incertitudes imposées sur l'état initial doivent disparaître au cours du temps, ce qui revient à dire qu'à la suite du transitoire, le système proposé se ramène à un système parfaitement certain et que par conséquent la dimension des pavés calculés doit tendre vers 0, ce qui n'est pas le cas de la figure II-11.

La figure II-12 représente les pavés obtenus lorsque le phénomène d'enveloppement est éliminé. Pour déterminer le bon orthotope à l'instant $k+1$, il faudrait utiliser à chaque itération le domaine abstrait $\mathcal{A}(x(k))$ comme condition initiale à l'instant k , ce qui est difficilement exploitable par le formalisme des intervalles. En revanche, il est possible de remplacer la récurrence par la relation (II-5) qui a l'avantage de se référer à un unique domaine initial $\mathcal{A}(x(0))$ qui est par définition un pavé, et ce quel que soit k . Cette fois, il n'y a plus de problème de divergence : les incertitudes s'estompent bien au fur et à mesure de la simulation et l'ensemble de ces orthotopes constitue une enveloppe multidimensionnelle contenant toutes les valeurs possibles des variables d'état aux différents instants k .

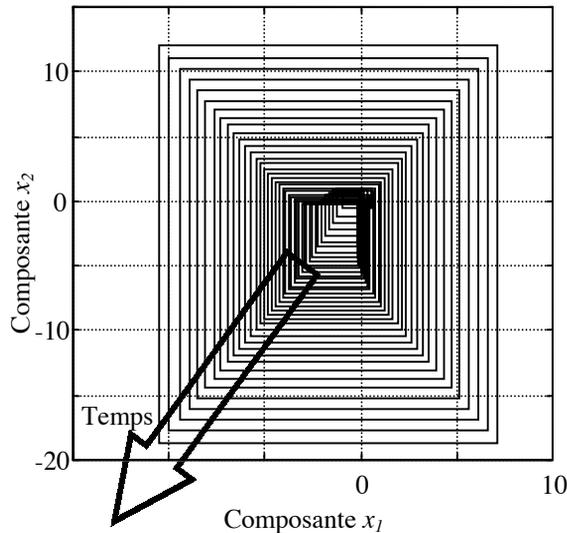


Figure II-11. Récurrence avec phénomène d'enveloppement

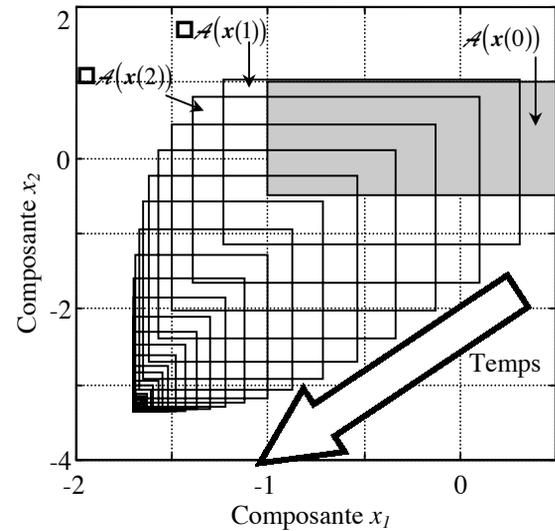


Figure II-12. Récurrence sans phénomène d'enveloppement

Cependant, du fait des variables communes, l'orthotope aligné circonscrit ne conduit qu'à une majoration de l'espace abstrait recherché. Dans le cadre de cet exemple, le champ $f(x(k))$ est linéaire en les variables bornées puisque la matrice d'état A est certaine, $\mathcal{A}(x(k+1))$ est donc un parallélogramme qu'il est possible de déterminer de manière exacte. Ce domaine étant convexe, il est parfaitement défini par ses sommets et peut ainsi être obtenu à partir de la connaissance de ceux de $\mathcal{A}(x(k))$. Le principe, très simple, consiste à augmenter la dimension du vecteur d'état de manière à travailler sur chaque sommet en même temps ; un sommet à l'instant $k+1$ étant alors obtenu par récurrence à partir des coordonnées dans l'espace d'état d'un sommet à l'instant k . Comme le domaine initial $\mathcal{A}(x(0))$ est un pavé, ses sommets sont parfaitement caractérisés par les bornes supérieures et inférieures des variables d'état $x_i(0)$. En construisant un vecteur d'état où les variables sont remplacées par l'une de leurs bornes, 2^n combinaisons définissant chacune un des 2^n sommets de $\mathcal{A}(x(0))$ sont obtenues. Le système étant linéaire et la matrice d'état A étant certaine, le parallélogramme déterminé à l'instant k aura le même nombre de sommets. Un sommet étant défini par n coordonnées, une nouvelle représentation d'état augmentée de dimension $n2^n$ est construite.

$$\mathbf{x}_s(k+1) = \mathbf{A}_s \mathbf{x}_s(k) + \mathbf{B}_s u(k), \text{ où } \mathbf{A}_s = \text{diag}_{2^n}(A), \mathbf{A}_s \in \mathbb{R}^{n2^n \times n2^n}, \mathbf{B}_s = \begin{bmatrix} \mathbf{B} \\ \vdots \\ \mathbf{B} \end{bmatrix}, \mathbf{B}_s \in \mathbb{R}^{n2^n \times q},$$

avec $\mathbf{x}_s(0)$, vecteur colonne constitué des 2^n sommets de $\mathcal{A}(x(0))$ et $\text{diag}_{2^n}(A)$ une matrice diagonale par blocs constituées de 2^n blocs de matrices A . En travaillant directement sur des nombres réels, le problème de dépendance est éliminé, ce qui permet de trouver facilement les sommets des parallélogrammes $\mathcal{A}(x(k))$ recherchés. Le résultat est présenté sur la figure II-13 où sont représentés à la fois $\mathcal{A}(x(k))$ (en trait plein) et $\square \mathcal{A}(x(k))$ (en pointillés)²³.

²³ Pour éviter de surcharger la figure, seuls les domaines correspondant à un indice temporel multiple de 3 ont été tracés.

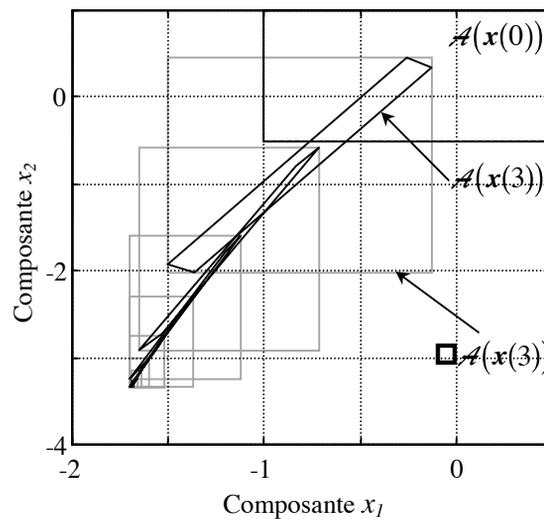


Figure II-13. Comparaison entre $\mathcal{A}(x(k))$ et $\square\mathcal{A}(x(k))$

Pour des systèmes plus complexes où interviennent des variables bornées dans la matrice d'état, la détermination de $\mathcal{A}(x(k))$ s'avère très complexe et ne peut plus se faire en travaillant sur les sommets (cas de domaines non-convexes ou dont les frontières ne sont plus des hyperplans). Il est toutefois possible d'en obtenir une majoration en réalisant par exemple une sous-division uniforme du domaine $\mathcal{A}(x(0))$ pour ensuite se servir de la relation (II-5) et obtenir un pavage de $\mathcal{A}(x(k))$; néanmoins, la somme de calculs à effectuer rend cette approche prohibitive. L'avantage d'utiliser $\square\mathcal{A}(x(k))$ est que la solution est plus facile à obtenir. Il vaut mieux alors utiliser la relation (II-5) pour contrer l'effet d'enveloppement, plutôt que d'utiliser la récurrence. Néanmoins, à cause de l'horizon infini, là encore les calculs sont fastidieux, notamment parce qu'il est nécessaire de connaître l'expression symbolique de cette relation en fonction des variables bornées (pour en construire une extension) et que celle-ci dépend de l'instant courant. Utiliser la relation de récurrence évite cette difficulté, mais est soumis au phénomène d'enveloppement. Pour le combattre, il est possible de prendre en compte l'information liée aux sorties mesurées (Kieffer et al, 1998). L'estimation d'état procède alors d'une étape de prédiction et d'une autre de correction. L'objectif est de déterminer le plus petit domaine \mathcal{X}_{l+1} contenant de manière garantie toutes les valeurs de l'état $x(l+1)$ compatibles avec l'ensemble des informations recueillies jusqu'alors. L'étape de prédiction détermine l'ensemble \mathcal{X}_{l+} des valeurs possibles de l'état à l'instant $l+1$ compatibles avec l'entrée $u(l)$ et l'ensemble précédemment trouvé \mathcal{X}_l . L'étape de correction calcule l'ensemble \mathcal{X}_{l+1}^0 des valeurs du vecteur d'état cohérentes avec les valeurs admissibles des mesures à l'instant $k+1$. Le domaine recherché correspond alors à l'intersection entre \mathcal{X}_{l+} et \mathcal{X}_{l+1}^0 .

Dans ce document, nous avons privilégié une approche par Espace de Parité (a contrario des méthodes à base d'observateurs) qui a pour objectif de reformuler les équations dynamiques d'un modèle sous forme de relations algébriques. En empilant les mesures prélevées sur un horizon temporel préalablement choisi, une représentation dite statique est

obtenue où il n'est plus nécessaire d'intégrer les équations du modèle pour obtenir des résidus.

II.5 CONCLUSION

Dans ce chapitre ont été présentés les principes (propriété d'inclusion, partition de domaine) sur lesquels repose l'analyse par intervalles, sachant que les techniques développées (formes centrées, méthode de pavage,...) peuvent s'employer pour un large éventail de problèmes à résoudre (optimisation globale, résolution de systèmes non-linéaires, inversion ensembliste, image ensembliste, prise en compte des erreurs numériques d'un calculateur, modélisation de solides, intersection de surfaces,...). La principale difficulté tient au fait que la dépendance entre deux variables bornées ne se transcrit pas au niveau de leurs bornes, et donc n'apparaît pas sur les variables intervalle représentant leurs domaines de variation. Le problème de dépendance qui en découle a, par conséquent, largement été explicité, que ce soit au sein d'une fonction vectorielle (variables dépendantes) ou entre les composantes d'un champ de vecteurs (variables communes) et engendre de manière générale un pessimisme sur les valeurs des bornes obtenues ou une majoration de l'ensemble image recherché. De plus, un formalisme reposant sur la notion d'espace abstrait a été introduit afin d'appréhender des domaines multidimensionnels de formes plus complexes que de simples pavés.

Dans le prochain chapitre, la génération de résidus par Espace de Parité dans le cadre de modèles incertains sera présentée. Les résultats obtenus serviront de support au chapitre IV dans lequel seront développées des techniques d'analyse par intervalles plus spécifiquement adaptées au diagnostic de systèmes, mais qui de manière générale reprendront les principes exposés dans ce chapitre de manière à limiter tout pessimisme préjudiciable à une procédure de diagnostic.

Chapitre III

Génération de résidus

TABLE DES MATIÈRES

III.1 INTRODUCTION.....	100
III.2 ESPACE DE PARITÉ DANS LE CAS CERTAIN.....	101
III.2.1 Principe général.....	101
III.2.2 Taille de fenêtre et notion d'observabilité.....	102
III.3 FORMES STATIQUES DANS LE CAS INCERTAIN	106
III.3.1 Modèle.....	106
III.3.2 Première forme statique.....	107
III.3.3 Seconde forme statique.....	109
III.4 MINIMISATION DES TERMES RÉSIDUELS	110
III.4.1 Critère d'optimisation.....	110
III.4.2 Expressions générales des formes de calcul et d'évaluation.....	112
III.5 MATRICE DE COMMANDE INCERTAINE	113
III.5.1 Structure classique.....	114
<i>III.5.1.1 Première forme statique</i>	<i>114</i>
<i>III.5.1.2 Seconde forme statique</i>	<i>115</i>
<i>III.5.1.3 Relation entre les deux formes statiques</i>	<i>116</i>
III.5.2 Structure de modèle de capteurs.....	118
<i>III.5.2.1 Première forme statique</i>	<i>119</i>
<i>III.5.2.2 Seconde forme statique</i>	<i>121</i>
<i>III.5.2.3 Relation entre les deux formes statiques</i>	<i>122</i>
III.5.3 Récapitulatif	122
III.6 MATRICE D'ÉTAT INCERTAINE	123
III.6.1 Choix de la forme statique.....	123
III.6.2 Élimination des termes résiduels.....	125
III.6.3 Récapitulatif	127
III.7 MATRICE D'OBSERVATION INCERTAINE	128
III.7.1 Structure classique.....	128
<i>III.7.1.1 Première forme statique</i>	<i>128</i>
<i>III.7.1.2 Seconde forme statique</i>	<i>129</i>
<i>III.7.1.3 Relation entre les deux formes statiques</i>	<i>130</i>
<i>III.7.1.4 Élimination des termes résiduels</i>	<i>133</i>

III.7.2 Structure de modèle de capteurs	137
III.7.2.1 Première forme statique	138
III.7.2.2 Seconde forme statique	139
III.7.2.3 Équivalence entre les deux formes statiques	140
III.7.3 Récapitulatif	141
III.8 CAS DE PLUSIEURS MATRICES INCERTAINES	142
III.9 CAS D'INCERTITUDES INVARIANTES	143
III.9.1 Formes statiques	143
III.9.2 Minimisation des termes résiduels	144
III.9.3 Matrice de commande incertaine.....	146
III.9.4 Matrice d'état incertaine.....	147
III.9.5 Matrice d'observation incertaine.....	148
III.9.6 Influence de la nature temporelle des incertitudes	153
III.10 MATRICE DE PARITÉ INCERTAINE	154
III.10.1 Cas statique.....	154
III.10.2 Extension dans le cas dynamique	158
III.11 CONCLUSION	160

III.1 INTRODUCTION

L'objectif de ce chapitre est de développer une méthode fiable de génération de résidus dans le cas de modèles à paramètres incertains. Comme explicité à la fin du chapitre précédent, l'effet d'enveloppement, propre aux systèmes récurrents, nous a contraint à privilégier une approche par Espace de Parité. Cette approche, largement répandue dans le domaine du diagnostic de systèmes, reste essentiellement limitée aux modèles certains (Mironovski, 1979), (Chow, 1980), Massoumnia (1988) mis à part quelques travaux proposés par (Lou, 1982) dans le cas multi-modèle ou (Frank et al, 1989) pour des systèmes à entrées inconnues. Dans ce dernier cas, les incertitudes sont supposées invariantes (les matrices de distribution étant connues et constantes), ce qui leur confère une structure additive. L'objectif est maintenant de s'intéresser à des modèles plus généraux où les incertitudes peuvent être à la fois multiplicatives et fluctuantes dans le temps. Notons cependant que nous nous limiterons strictement à des modèles structurés qui permettent de prendre en compte la méconnaissance affectant un système physique en indiquant exactement quels sont les paramètres du modèle qui sont incertains.

Dans un premier temps, la génération de relations de parité dans le cas d'un modèle certain est succinctement présentée afin d'en comprendre le principe fondateur et les différentes difficultés s'y rattachant. Ensuite, les sections 3 à 9 détaillent une première stratégie en étendant les précédents résultats dans le contexte de l'incertain. Ainsi, la troisième section explicite la notion de forme statique dans le cas d'un modèle incertain et définit les expressions générales de la première et seconde formes statiques. La quatrième section développe une méthode d'optimisation dont la finalité est de déterminer une matrice de parité constante assurant l'élimination au niveau des relations de parité, de toute influence des variables d'état inconnues susceptibles d'empêcher l'évaluation des résidus. Les trois sections suivantes établissent les expressions des formes de calcul et d'évaluation lorsque respectivement les matrices de commande, d'état puis d'observation sont incertaines et que les incertitudes associées fluctuent entre leurs bornes au cours du temps. Pour chaque cas traité, l'influence de l'horizon temporel sur la possibilité d'obtenir des relations de parité totalement découplées des variables d'état est étudiée. Lorsque les deux formes statiques sont exploitables, une analyse est effectuée afin d'observer si les résidus obtenus pour ces deux représentations sont structurellement équivalents. La huitième section récapitule les principaux résultats obtenus et étend cette étude au cas d'incertitudes pouvant affecter, à la fois, plusieurs matrices de la représentation d'état du modèle. La neuvième section présente brièvement la même étude dans le cas d'incertitudes invariantes et permet de montrer que considérer certaines variables bornées comme constantes sur l'horizon temporel choisi, peut permettre d'aboutir à un découplage total, impossible sinon, de la forme d'évaluation vis-à-vis du vecteur d'état. Enfin, la dixième section développe une seconde stratégie de génération de résidus, dont l'objectif est d'outrepasser les différentes impasses inhérentes à la première stratégie proposée en ce qui concerne la possibilité d'un découplage. L'approche explicitée

s'attache cette fois à construire une matrice de parité incertaine assurant, dans tous les cas de figure, l'élimination des variables d'état des relations de parité obtenues.

III.2 ESPACE DE PARITÉ DANS LE CAS CERTAIN

III.2.1 Principe général

De manière simpliste, l'approche par Espace de Parité repose sur la reformulation des équations dynamiques d'un modèle sous forme de relations purement algébriques. En empilant les mesures obtenues sur un horizon temporel préalablement choisi, une représentation statique est obtenue où il n'est plus nécessaire d'intégrer les équations du modèle afin d'obtenir des résidus. Considérons un modèle dynamique linéaire et certain défini par les relations discrètes suivantes :

$$\begin{cases} \mathbf{x}(k+1) = \mathbf{A}\mathbf{x}(k) + \mathbf{B}\mathbf{u}(k) \\ \mathbf{y}(k) = \mathbf{C}\mathbf{x}(k) \end{cases} \quad \mathbf{x} \in \mathbb{R}^n, \mathbf{y} \in \mathbb{R}^m, \mathbf{u} \in \mathbb{R}^q, \quad (\text{III-1})$$

où \mathbf{A} , \mathbf{B} et \mathbf{C} représentent respectivement les matrices d'état, de commande et d'observation. Les notations \mathbf{u} , \mathbf{x} et \mathbf{y} désignent les vecteurs des entrées commandées, des variables d'état et des sorties mesurées alors que k définit l'indice temporel courant. Puisque le vecteur d'état $\mathbf{x}(k+1)$ s'exprime de manière récurrente, il est facile de vérifier l'égalité suivante liant $\mathbf{x}(k+j)$ ($j \in \mathbb{N}^*$) à $\mathbf{x}(k)$ et aux commandes :

$$\mathbf{x}(k+j) = \mathbf{A}^j \mathbf{x}(k) + \sum_{i=0}^{j-1} \mathbf{A}^{j-i-1} \mathbf{B}\mathbf{u}(k+i).$$

En utilisant le système de mesures avec la relation précédente puis en regroupant les vecteurs de mesures sur un horizon $[k, k+s]$, ($s \in \mathbb{N}^*$), la forme statique suivante est déduite :

$$\mathbf{Y}(k, s) = \mathbf{C}_s \mathbf{x}(k) + \mathbf{H}_s \mathbf{U}(k, s-1), \quad \mathbf{U} \in \mathbb{R}^{sq}, \quad \mathbf{Y} \in \mathbb{R}^{sm} \text{ avec } s_m = (s+1)m, \quad (\text{III-2})$$

$$\text{où } \mathbf{Z}(k, s) = \begin{bmatrix} \mathbf{z}(k) \\ \vdots \\ \mathbf{z}(k+s) \end{bmatrix}, \quad \mathbf{C}_s = \begin{bmatrix} \mathbf{C} \\ \mathbf{CA} \\ \mathbf{CA}^2 \\ \vdots \\ \mathbf{CA}^s \end{bmatrix}, \quad \mathbf{H}_s = \begin{bmatrix} 0 & 0 & \dots & 0 \\ \mathbf{CB} & 0 & \dots & 0 \\ \mathbf{CAB} & \mathbf{CB} & \dots & 0 \\ \vdots & \vdots & & \vdots \\ \mathbf{CA}^{s-1}\mathbf{B} & \mathbf{CA}^{s-2}\mathbf{B} & \dots & \mathbf{CB} \end{bmatrix}, \quad \begin{matrix} \mathbf{C}_s \in \mathbb{R}^{s_m \times n} \\ \mathbf{H}_s \in \mathbb{R}^{s_m \times sq} \end{matrix}$$

Cependant, pour pouvoir évaluer cette relation, il est nécessaire d'éliminer toute référence à l'état $\mathbf{x}(k)$ dont la valeur est inconnue et qui apparaît au travers de la matrice \mathbf{C}_s appelée matrice d'observabilité d'ordre s (Brunet et al, 1990). Soit s_c le rang de \mathbf{C}_s ; pour une taille suffisante de la fenêtre temporelle s , il est possible de se munir d'une matrice de projection appelée matrice de parité $\mathbf{W} \in \mathbb{R}^{s_w \times s_m}$, $s_w = s_m - s_c$, de plein rang ligne, vérifiant (Massoumnia et al, 1988) :

$$\mathbf{W}\mathbf{C}_s = 0. \quad (\text{III-3})$$

La condition d'existence de \mathbf{W} est que la matrice \mathbf{C}_s ne soit pas de plein rang ligne et dans ce cas existent alors une infinité de matrices de parité vérifiant la contrainte d'orthogonalité précédente (une méthode de calcul est proposée en annexe D). En regroupant ensemble tous les termes de la forme statique (III-2) ne dépendant pas des variables d'état, puis en multipliant le résultat à gauche par \mathbf{W} , le vecteur de parité \mathbf{p}_{cal} constitué de s_w relations de parité linéairement indépendantes¹ est obtenu :

$$\underbrace{\mathbf{W}(\mathbf{Y}(k, s) - \mathbf{H}_s \mathbf{U}(k, s - 1))}_{\mathbf{p}_{cal}(k+s)} = \underbrace{\mathbf{W} \mathbf{C}_s \mathbf{x}(k)}_{\mathbf{p}_{éval}(k+s)} = 0.$$

Les entrées commandées pouvant au choix être mesurées ou approchées par les valeurs de consignes associées, le vecteur de parité \mathbf{p}_{cal} est donc parfaitement déterminé et correspond à la forme de calcul des résidus. Puisque la matrice \mathbf{W} est orthogonale à \mathbf{C}_s , la forme d'évaluation $\mathbf{p}_{éval}$ est théoriquement nulle en l'absence de défaut et de perturbation ; elle indique par conséquent qu'en bon fonctionnement, \mathbf{p}_{cal} doit être idéalement nul.

Ajoutons maintenant des bruits $\mathbf{b}(k)$ décrivant des erreurs de modélisation au niveau du système de mesures de (III-1), ainsi qu'un défaut d'amplitude $d(k)$ sur le $j^{\text{ème}}$ capteur :

$$\mathbf{y}(k) = \mathbf{C} \mathbf{x}(k) + \mathbf{b}(k) + \mathbf{d}(k) \quad \mathbf{b} \in \mathbb{R}^m, \mathbf{d} \in \mathbb{R}^m$$

où $\mathbf{d}(k)$ est un vecteur colonne dont toutes les composantes sont nulles à l'exception de la $j^{\text{ème}}$ valant $d(k)$. En réalisant le même travail que précédemment (III-2) devient :

$$\mathbf{Y}(k, s) = \mathbf{C}_s \mathbf{x}(k) + \mathbf{H}_s \mathbf{U}(k, s - 1) + \mathbf{B}(k, s) + \mathbf{D}(k, s),$$

puis en regroupant dans la forme de calcul tous les termes parfaitement connus, le vecteur de parité s'exprime comme suit :

$$\underbrace{\mathbf{W}(\mathbf{Y}(k, s) - \mathbf{H}_s \mathbf{U}(k, s - 1))}_{\mathbf{p}_{cal}(k+s)} = \underbrace{\mathbf{W} \mathbf{B}(k, s) + \mathbf{W} \mathbf{D}(k, s)}_{\mathbf{p}_{éval}(k+s)}.$$

La forme d'évaluation spécifie alors la sensibilité des résidus vis-à-vis des perturbations ou d'un défaut potentiel et montre que le vecteur résidu² sera de manière générale non nul, et ce même en fonctionnement normal ; d'où la nécessité d'une procédure de prise de décision ayant pour objectif de départager l'influence des défauts de celle des perturbations.

III.2.2 Taille de fenêtre et notion d'observabilité

L'objectif de cette section est de détailler l'influence de la taille de fenêtre temporelle s et le lien qu'elle suscite avec la notion d'observabilité. Le théorème de Cayley-Hamilton (Landesman et al, 1992), appliqué à la matrice d'état \mathbf{A} , permet de montrer qu'il existe toujours une valeur de s à partir de laquelle la condition d'existence de la matrice de parité est vérifiée. Si le polynôme caractéristique de \mathbf{A} est donné par :

$$p(\lambda) = a_0 + a_1 \lambda + \dots + a_{n-1} \lambda^{n-1} + \lambda^n,$$

¹ Le terme indépendant classiquement rencontré dans ce contexte signifie que, puisque la matrice \mathbf{W} est de plein rang ligne, il n'existe aucune relation de parité pouvant être une combinaison linéaire des autres.

² Le terme vecteur résidu désigne ici la forme de calcul et est donc synonyme de vecteur de parité.

alors, puisque $p(A)=0$, il s'ensuit que :

$$CA^n = -a_0C - \dots - a_{n-1}CA^{n-1}, \text{ soit } \underbrace{\begin{bmatrix} a_0 i_{m,j}^T & \dots & a_{n-1} i_{m,j}^T & -i_{m,j}^T \end{bmatrix}}_{w^T} \begin{bmatrix} C \\ \vdots \\ CA^n \end{bmatrix} = 0, \forall j \in \{1, \dots, m\},$$

où le vecteur ligne $i_{m,j}^T$ définit la $j^{\text{ème}}$ ligne de la matrice identité d'ordre m . Il existe donc toujours un vecteur ligne w^T non nul orthogonal à la matrice C_n . De même, toute puissance de A supérieure à n s'exprime comme combinaison linéaire des n puissances précédentes. En effectuant ce travail autant de fois que nécessaire, il est possible d'exprimer $A^{n'}$, $n' \geq n$, en fonction des seules matrices I_n, A, \dots, A^{n-1} ; ce qui signifie que C_s , $s \geq n$, est de même rang que C_n et n'est donc pas de plein rang ligne. Donner à s une valeur égale ou supérieure à la dimension de l'état n conduit donc à l'existence de la matrice de parité W dans le cas certain.

Un système est observable s'il est possible de reconstruire l'état initial à partir de la connaissance des mesures à un instant $k > 0$. Traditionnellement, le rang de la matrice d'observabilité C_{n-1} est calculé : si celui-ci vaut n , alors le système est observable. Dans le cadre de l'Espace de Parité, la notion d'observabilité est importante car choisir une taille de fenêtre temporelle s de manière à ce que la matrice C_s ait le même rang que C_{n-1} signifie que toute l'information contenue dans le modèle est restituée dans les relations de parité calculées (Mironovski, 1979), (Brunet et al, 1990). Ainsi, le rang de la matrice d'observabilité C_{n-1} donne le nombre de composantes du vecteur d'état accessibles au travers des mesures observées sur un horizon $[k, k+s]$. Parvenir à un découplage total des résidus vis-à-vis du vecteur d'état signifie que suffisamment de mesures ont été empilées afin de pouvoir estimer $x(k)$ et le remplacer par l'information issue des observations. Il est donc a priori inutile de prendre une taille de fenêtre supérieure à n . De ce fait, l'optique d'une solution où toute l'information disponible est utilisée et où une taille de fenêtre minimale est recherchée, consiste à respecter la condition : $s \leq n$; le dépassement de cette limite n'apporte aucune information complémentaire dans un cadre déterministe et certain. La matrice C_n possède s_m lignes pour un rang inférieur ou égal à n ; il est donc fortement probable qu'une valeur de s inférieure à n permette :

- qu'au moins une ligne de C_s soit combinaison linéaire des autres (existence de W),
- de respecter la condition : $\text{rang}(C_s) = \text{rang}(C_{n-1})$ (utilisation de toute l'information).

Il existe une taille minimale de fenêtre s_{min} pour laquelle les deux conditions précédentes sont vérifiées. Il s'agit de déterminer la plus petite valeur de s telle que C_s ait le même rang que C_{s-1} . Cette contrainte impose non seulement que C_s n'est pas de plein rang ligne (puisque la matrice CA^s est une combinaison des lignes de C_{s-1}), mais aussi qu'elle possède le même rang que la matrice d'observabilité C_{n-1} . En effet, toutes les matrices CA^{s+k} , $k \in \{1, \dots, n-s\}$ sont elles aussi nécessairement des combinaisons linéaires des lignes de C_{s-1} :

$$CA^{s+k} = f(C, CA, \dots, CA^{s-1})A^k = f(CA^k, CA^{k+1}, \dots, CA^{s+k-1}) = \dots = f^k(C, CA, \dots, CA^{s-1})$$

où f et f^k sont des applications linéaires. De ce fait, toutes les matrices C_{s+k} , $k \in \{1, \dots, n-s\}$, ont exactement le même rang que C_{s-1} et C_s . Or nous avons démontré précédemment que le rang

maximal des matrices \mathcal{C}_{s+k} est limité à celui de \mathcal{C}_{n-1} , il s'ensuit donc que \mathcal{C}_{s-1} et \mathcal{C}_s sont de même rang que la matrice d'observabilité. La valeur de s ainsi trouvée correspond par conséquent à s_{min} puisque c'est la plus petite taille de fenêtre permettant l'utilisation de toute l'information disponible, ainsi que l'existence de la matrice de parité.

Exemple. Considérons le système défini par les matrices d'observation \mathbf{C} et d'état \mathbf{A} suivantes. D'ordre $n=4$ et disposant de $m=2$ capteurs, le calcul de la matrice d'observabilité \mathcal{C}_3 associée montre que ce système est observable puisque son rang vaut 4 :

$$\mathbf{C} = \begin{bmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}, \mathbf{A} = \begin{bmatrix} -3 & 0 & 0 & -0.75 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix}, \mathcal{C}_3 = \begin{bmatrix} 0 & 0 & 0 & 0 & 1 & 0 & -3 & 1 \\ 0 & 0 & 1 & 0 & 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & -0.75 & 0 \end{bmatrix}^T.$$

L'étude de \mathcal{C}_3 permet de conclure que la valeur de s_{min} correspond à 3. Soit \mathbf{w}^T un vecteur ligne dont le $i^{\text{ième}}$ élément, $i \in \{1, \dots, s_m=8\}$, est noté w_i . Les contraintes assurant la propriété d'orthogonalité $\mathbf{w}^T \mathcal{C}_3 = 0$ permettent alors d'écrire :

$$\mathbf{w}^T = \left[w_1 \quad w_2 \quad w_3 \quad -w_1 \quad w_5 \quad -w_3 \quad \frac{w_2}{0.75} \quad \frac{3w_2}{0.75} - w_5 \right],$$

ce qui conduit à $s_w = s_m - s_c = 8 - 4 = 4$ relations de parité linéairement indépendantes, comme par exemple :

$$\begin{aligned} p_1(k+3) &= w_1(y_1(k) - y_2(k+1)) & p_2(k+3) &= w_3(y_1(k+1) - y_2(k+2)) \\ p_3(k+3) &= w_5(y_1(k+2) - y_2(k+3)) & p_4(k+3) &= w_2 \left(y_2(k) + \frac{4}{3}y_1(k+3) + 4y_2(k+3) \right). \end{aligned}$$

Il est facile de remarquer que les résidus p_1 , p_2 et p_3 exploitent exactement la même information puisqu'ils correspondent à la même relation décalée dans le temps : $p_3(k+3) = p_2(k+4) = p_1(k+5)$, en imposant arbitrairement $w_1 = w_3 = w_5$. Prendre une taille inférieure à s_{min} , par exemple $s=2$, conduit à seulement 2 relations de parité, en l'occurrence p_1 et p_2 . Si l'absence de p_3 ne porte aucun préjudice, celle de p_4 se traduit par une perte d'information. Inversement, choisir une taille de fenêtre strictement supérieure à s_{min} génère des relations de parité supplémentaires pouvant s'exprimer comme combinaisons linéaires des relations de parité p_1 et p_4 décalées dans le temps. En supposant que tous les coefficients w_i sont égaux à 1, les résultats suivants sont obtenus pour $s=4$:

$$\begin{aligned} p_1'(k+3) &= p_1(k+4) & p_2'(k+3) &= p_1(k+5) & p_3'(k+3) &= p_1(k+6) \\ p_4'(k+3) &= p_1(k+3) + p_4(k+4) & p_5'(k+3) &= -\frac{4}{3}p_1(k+6) + p_4(k+3) & p_6'(k+3) &= p_4(k+4) \end{aligned}$$

Dans un cadre déterministe, ces relations supplémentaires n'apportent donc rien de plus. ■

Certains auteurs tels que (Chow, 1980), (Chow et al, 1984) préconisent d'adapter la taille de fenêtre temporelle au degré d'observabilité associé à chaque capteur (Espace de Parité Généralisé). Le principe est semblable à celui explicité précédemment si ce n'est que la forme statique représentative du modèle est décomposée en m formes statiques d'horizons

temporels différents dédiées à chaque capteur. Mironovski (1979) avait déjà proposé une solution similaire reposant sur ce qu'il nomme les invariants de Kronecker de la paire (C, A) .

À partir du moment où W est de plein rang ligne et satisfait (III-3), toutes les relations de parité obtenues sont linéairement indépendantes, sachant qu'un nombre infini de choix est possible en multipliant W à gauche par une matrice inversible. En revanche, comme observé dans l'exemple précédent, certaines d'entre elles peuvent s'exprimer comme combinaison d'autres relations de parité relatées à des instants différents (Gertler, 1991), (Frank, 1990), auquel cas nous parlerons de dépendance temporelle. Dans un contexte déterministe, les relations dépendantes temporellement (telles p_2 et p_3 dans l'exemple précédent) contiennent la même information et permettent donc de détecter les mêmes défauts que les résidus de base (comme p_1), avec en plus un certain retard. En pratique, la présence de bruits (et d'incertitudes de modèle) conduit à des tests statistiques. Dans ces conditions, l'affirmation précédente est plus nuancée. Pour des problèmes de variance sur les grandeurs mesurées et de sensibilité vis-à-vis d'un défaut donné, un tel résidu peut être préférable. Le fait de choisir $s=s_{min}$ mène généralement à des relations dépendantes temporellement, mais c'est la taille de fenêtre qui conduit à restreindre au maximum leur nombre afin de limiter la complexité des calculs à entreprendre et à déterminer les $rang(C)$ résidus de base³.

Certains chercheurs tels que Ploix (1998) se sont attachés à déterminer les relations de parité minimales d'un système, c'est-à-dire l'ensemble des résidus de base correspondant à un horizon temporel le plus faible possible et constituant une base génératrice⁴ de toutes les relations de parité possibles. A condition que le système soit commandable (condition pour trouver un horizon minimal) et que la matrice d'observation soit de plein rang ligne, une procédure systématique est proposée consistant à incrémenter s jusqu'à trouver m relations de parité minimales. A chaque itération, les dépendances linéaire et temporelle des nouvelles relations de parité calculées vis-à-vis des relations minimales déjà déterminées sont testées, permettant ainsi d'effectuer un tri systématique.

L'objectif est maintenant de transposer ce principe de raisonnement dans le cas d'un modèle incertain où la connaissance de plusieurs paramètres nous échappe. Si des incertitudes affectent les matrices d'observation ou d'état, la matrice "d'observabilité" C_{n-1} devient elle aussi incertaine. Se pose alors le problème de la définition même de l'observabilité dans le contexte incertain et de son étude afin de choisir une taille de fenêtre s appropriée. La principale difficulté rencontrée réside toutefois dans le fait que l'état reste inconnu. Même s'il en est de même pour les paramètres incertains, le fait d'avoir accès à leurs bornes permet de prendre en compte toute l'information qu'ils contiennent lors de l'évaluation du domaine de bon fonctionnement du système. L'objectif recherché consiste donc à réaliser un découplage total vis-à-vis des variables d'état de manière à éviter d'avoir à estimer ces dernières, chose qui n'est pas toujours réalisable. Ceci ne correspond plus au problème d'observabilité tel qu'il

³ (Ploix, 1998) justifie que leur nombre correspond en fait au nombre de relations entrée-sortie du système dans le cas où la matrice d'observation C est de plein rang ligne : $y(k) = C(zI_n - A)^{-1}Bu(k)$

⁴ Toute relation de parité correspond à une combinaison linéaire des relations minimales pouvant être exprimées à différents instants.

est défini dans le cas certain. Il est en effet plus difficile de retrouver l'état initial $x(0)$ à cause des incertitudes⁵. Nous ne nous attarderons donc plus sur le problème d'observabilité en temps que tel (et par conséquent sur la conservation de toute l'information du système ou la notion de minimalité). Nous nous intéresserons seulement à l'étude de l'influence de l'horizon temporel s sur la possibilité de trouver une matrice de parité permettant un découplage total en regard du vecteur d'état, ou sur l'existence de structures particulières des incertitudes permettant d'élucider directement cette difficulté.

III.3 FORMES STATIQUES DANS LE CAS INCERTAIN

III.3.1 Modèle

Dans le contexte ensembliste, un paramètre incertain est qualifié de variant, (respectivement d'invariant) lorsque sa valeur inconnue fluctue entre des bornes (respectivement reste constante) au cours du temps. La figure suivante représente différentes réalisations de deux variables bornées ayant le même espace abstrait $[-1,1]$, l'une dépendant du temps alors que l'autre est invariante.

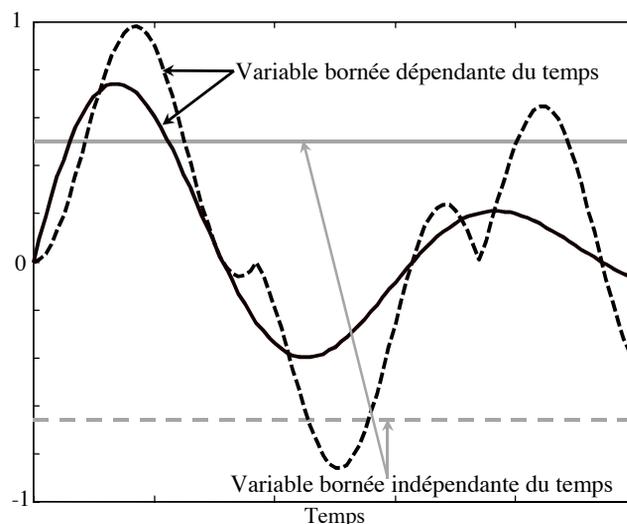


Figure III-1. Variables bornées variante et invariante

En pratique, les incertitudes paramétriques sont de nature variante. En effet, reprenons l'exemple du premier chapitre portant sur la résistance électrique. Les incertitudes paramétriques θ_2 et θ_4 visent à modéliser simplement des phénomènes non-linéaires comme la saturation magnétique. Les erreurs commises entre les valeurs vraies et leurs mesures dépendent du point de fonctionnement courant ; les valeurs des paramètres incertains les représentant changent donc au cours du temps. De même, les variables bornées θ_3 et θ_5

⁵ (D'Andrea et al, 1997) adopte une définition similaire au cas certain : un système incertain structuré est observable s'il est possible de déterminer l'espace abstrait $\mathcal{A}(x(0))$ à partir de la connaissance des mesures à un instant $k > 0$.

censées représenter des bruits aléatoires de mesure sont nécessairement variantes. De plus, dans le cas de paramètres invariants, il peut être préférable d'essayer de les estimer afin de se ramener à un modèle certain plutôt que de les conserver tels quels. Néanmoins, nous verrons ultérieurement qu'il peut être intéressant de considérer comme invariants certains paramètres variant peu sur la taille de fenêtre utilisée pour la construction de la forme statique, et ce afin de faciliter l'élimination du vecteur d'état indésirable.

La notion de variable "variante" est primordiale dans la mesure où nous supposons qu'une même variable bornée θ exprimée à deux instants différents k et t définit en réalité deux variables $\theta(k)$ et $\theta(t)$ complètement indépendantes même si elles possèdent le même espace abstrait $\mathcal{A}(\theta(k)) = \mathcal{A}(\theta(t))$. Il n'y a aucun lien entre les valeurs inconnues $\theta(k)$ et $\theta(t)$, la seule information disponible est que leur support est identique (Adrot et al, 2000a). En revanche, dans le cas invariant (ou d'un système statique à paramètres incertains variants (Adrot et al, 1999a)), cette distinction n'existe plus. L'indice temporel est alors inutile puisqu'il faut absolument conserver la contrainte $\theta(k) = \theta(t) = \theta$ au risque sinon d'ajouter du pessimisme à cause du problème de dépendance. Cette remarque est d'autant plus pertinente que notre travail repose sur la construction de formes statiques où vont nécessairement apparaître des paramètres incertains exprimés à différents instants.

La classe des modèles à paramètres incertains invariants étant un cas particulier, l'étude plus générale de modèles dynamiques linéaires, à paramètres incertains variants, exprimés sous la forme d'une représentation d'état discrète, est maintenant développée. Considérons des incertitudes structurées pouvant affecter les matrices d'état $\mathbf{A}(\boldsymbol{\theta}_A)$, de commande $\mathbf{B}(\boldsymbol{\theta}_B)$ ou d'observation $\mathbf{C}(\boldsymbol{\theta}_C)$ du modèle défini par les équations suivantes :

$$\begin{cases} \mathbf{x}(k+1) = \mathbf{A}(\boldsymbol{\theta}_A(k))\mathbf{x}(k) + \mathbf{B}(\boldsymbol{\theta}_B(k))\mathbf{u}(k) \\ \mathbf{y}(k) = \mathbf{C}(\boldsymbol{\theta}_C(k))\mathbf{x}(k) \end{cases} \quad \mathbf{x} \in \mathbb{R}^n, \mathbf{y} \in \mathbb{R}^m, \mathbf{u} \in \mathbb{R}^q. \quad (\text{III-4})$$

Tout au long de ce document, la notation $\mathbf{Z}(\boldsymbol{\theta}_Z(k))$ indique qu'au moins un paramètre incertain affecte la matrice \mathbf{Z} . Le vecteur $\boldsymbol{\theta}_Z(k)$ regroupe toutes les incertitudes affectant cette matrice à l'instant k , celles-ci étant représentées par des variables bornées normalisées⁶. Si ponctuellement est émise l'hypothèse que cette matrice est certaine, alors elle est notée \mathbf{Z} .

III.3.2 Première forme statique

La construction de cette forme statique reprend exactement le principe exposé dans le cas certain. La relation liant le vecteur $\mathbf{x}(k+j)$ ($j \in \mathbb{N}^*$) à $\mathbf{x}(k)$ et aux commandes devient :

$$\begin{aligned} \mathbf{x}(k+j) = & \mathbf{A}(\boldsymbol{\theta}_A(k+j-1)) \cdots \mathbf{A}(\boldsymbol{\theta}_A(k))\mathbf{x}(k) + \mathbf{B}(\boldsymbol{\theta}_B(k+j-1))\mathbf{u}(k+j-1) + \cdots \\ & \cdots + \sum_{i=1}^{j-1} \mathbf{A}(\boldsymbol{\theta}_A(k+j-1)) \cdots \mathbf{A}(\boldsymbol{\theta}_A(k+i))\mathbf{B}(\boldsymbol{\theta}_B(k+i-1))\mathbf{u}(k+i-1) \end{aligned} \quad (\text{III-5})$$

⁶ Rappelons que ceci consiste à imposer $\|\boldsymbol{\theta}_Z\|_\infty \leq 1$ et n'induit aucune perte de généralité (section II.2.2)

À partir de la relation précédente, en regroupant les vecteurs d'état sur un horizon $[k, k+s]$, l'expression suivante est obtenue :

$$X(k, s) = \mathcal{A}_s(\theta_A)x(k) + \mathcal{B}_s(\theta_A, \theta_B)U(k, s-1), \quad \mathcal{A}_s \in \mathbb{R}^{s_n \times n}, \quad \mathcal{B}_s \in \mathbb{R}^{s_n \times sq}, \quad s_n = (s+1)n, \quad (\text{III-6})$$

$$\text{avec } \mathcal{A}_s(\theta_A) = \begin{bmatrix} I_n \\ A(\theta_A(k)) \\ A(\theta_A(k+1))A(\theta_A(k)) \\ \vdots \\ A(\theta_A(k+s-1)) \cdots A(\theta_A(k)) \end{bmatrix}, \quad I_n \text{ matrice identité d'ordre } n,$$

$$\mathcal{B}_s(\theta_A, \theta_B) = \begin{bmatrix} 0 & 0 & \cdots & 0 \\ \mathcal{B}(\theta_B(k)) & 0 & \cdots & 0 \\ A(\theta_A(k+1))\mathcal{B}(\theta_B(k)) & \mathcal{B}(\theta_B(k+1)) & \cdots & 0 \\ \vdots & \vdots & \cdots & \vdots \\ A(\theta_A(k+s-1)) \cdots A(\theta_A(k+1))\mathcal{B}(\theta_B(k)) & A(\theta_A(k+s-1)) \cdots A(\theta_A(k+2))\mathcal{B}(\theta_B(k+1)) & \cdots & \mathcal{B}(\theta_B(k+s-1)) \end{bmatrix}$$

La matrice $L_s(\theta_C)$ liant le vecteur $Y(k, s)$ des mesures empilées sur un horizon $[k, k+s]$ à celui des variables d'état $X(k, s)$ est donnée par :

$$Y(k, s) = L_s(\theta_C)X(k, s) \quad \text{où} \quad L_s(\theta_C) = \begin{bmatrix} C(\theta_C(k)) & 0 & 0 \\ 0 & \ddots & 0 \\ 0 & 0 & C(\theta_C(k+s)) \end{bmatrix}, \quad L_s \in \mathbb{R}^{s_m \times s_n}. \quad (\text{III-7})$$

Pour une taille de fenêtre temporelle s , la première forme statique s'écrit donc :

$$Y(k, s) = C_s(\theta_A, \theta_C)x(k) + H_s(\theta_A, \theta_B, \theta_C)U(k, s-1), \quad x \in \mathbb{R}^n, \quad Y \in \mathbb{R}^{s_m}, \quad U \in \mathbb{R}^{sq} \quad (\text{III-8})$$

$$\text{avec } C_s(\theta_A, \theta_C) = \begin{bmatrix} C(\theta_C(k)) \\ C(\theta_C(k+1))A(\theta_A(k)) \\ C(\theta_C(k+2))A(\theta_A(k+1))A(\theta_A(k)) \\ \vdots \\ C(\theta_C(k+s))A(\theta_A(k+s-1)) \cdots A(\theta_A(k)) \end{bmatrix}, \quad C_s \in \mathbb{R}^{s_m \times n}, \quad s_m = (s+1)m,$$

$$H_s(\theta_A, \theta_B, \theta_C) = \begin{bmatrix} 0 & 0 \\ \mathcal{C}(\theta_C(k+1))\mathcal{B}(\theta_B(k)) & 0 \\ \mathcal{C}(\theta_C(k+2))A(k+1)\mathcal{B}(\theta_B(k)) & \mathcal{C}(\theta_C(k+2))\mathcal{B}(\theta_B(k+1)) & \dots \\ \vdots & \vdots & \dots \\ \mathcal{C}(\theta_C(k+s))A(\theta_A(k+s-1)) \cdots A(\theta_A(k+1))\mathcal{B}(\theta_B(k)) & \mathcal{C}(\theta_C(k+s))A(\theta_A(k+s-1)) \cdots A(\theta_A(k+2))\mathcal{B}(\theta_B(k+1)) & \dots \\ \dots & 0 & \dots \\ \dots & 0 & \dots \\ \dots & 0 & \dots \\ \dots & \mathcal{C}(\theta_C(k+s))\mathcal{B}(\theta_B(k+s-1)) & \dots \end{bmatrix}, \quad H_s \in \mathbb{R}^{s_m \times sq}$$

Chaque ligne de la matrice H_s dépend de la même matrice d'observation alors que chacune de ses colonnes fait apparaître la même matrice de commande. La complexité de cette expression, dont la structure est équivalente à la forme statique (III-2) trouvée dans le cas certain, est due aux différentes matrices de la représentation d'état possédant maintenant un indice temporel.

III.3.3 Seconde forme statique

Pour des raisons que nous expliciterons dans la section III.6.1, intéressons-nous à la construction d'une seconde forme statique, où la matrice d'état apparaît de manière linéaire, et non plus sous la forme de produits comme dans les matrices C_s et H_s de la première forme statique. En observant la relation (III-5), il est facile de remarquer que ces produits proviennent de la volonté d'exprimer tous les vecteurs d'état $x(k+j)$ en fonction du même état initial de début de fenêtre $x(k)$. Dans la nouvelle représentation, appelée seconde forme statique, ce sont les relations exprimant les vecteurs $x(k+j)$ en fonction des vecteurs antérieurs $x(k+j-1)$, $j \in \mathbb{N}^*$ qui sont privilégiées :

$$\begin{bmatrix} A(\theta_A(k+j-1)) & -I_n \end{bmatrix} \begin{bmatrix} x(k+j-1) \\ x(k+j) \end{bmatrix} = -B(\theta_B(k+j-1))u(k+j-1). \quad (\text{III-9})$$

À partir de la précédente relation, en empilant les vecteurs d'état et de commande sur un horizon $[k, k+s]$, l'expression matricielle suivante est obtenue :

$$M_s(\theta_A)X(k, s) = N_s(\theta_B)U(k, s-1), \quad (\text{III-10})$$

$$\text{avec : } M_s(\theta_A) = \begin{bmatrix} A(\theta_A(k)) & -I_n & \cdots & 0 & 0 \\ 0 & A(\theta_A(k+1)) & \cdots & 0 & 0 \\ \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & \cdots & A(\theta_A(k+s-1)) & -I_n \end{bmatrix}, \quad M_s \in \mathbb{R}^{sn \times sn},$$

$$\text{et } N_s(\theta_B) = \begin{bmatrix} -B(\theta_B(k)) & 0 & \cdots & 0 \\ 0 & -B(\theta_B(k+1)) & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & -B(\theta_B(k+s-1)) \end{bmatrix}, \quad N_s \in \mathbb{R}^{sn \times sq}.$$

Compte tenu de la relation (III-7) donnant le vecteur $Y(k, s)$ en fonction de $X(k, s)$, la seconde forme statique est finalement déduite :

$$\tilde{C}_s(\theta_A, \theta_C)X(k, s) = \tilde{H}_s(\theta_B) \begin{bmatrix} U(k, s-1) \\ Y(k, s) \end{bmatrix}, \quad X \in \mathbb{R}^{sn}, \quad Y \in \mathbb{R}^{sm}, \quad U \in \mathbb{R}^{sq} \quad (\text{III-11})$$

$$\text{avec } \tilde{C}_s(\theta_A, \theta_C) = \begin{bmatrix} L_s(\theta_C) \\ M_s(\theta_A) \end{bmatrix}, \quad \tilde{C}_s \in \mathbb{R}^{\{s_m+sn\} \times sn},$$

$$\text{et } \tilde{H}_s(\theta_B) = \begin{bmatrix} 0 & I_{s_m} \\ N_s(\theta_B) & 0 \end{bmatrix}, \quad \tilde{H}_s \in \mathbb{R}^{\{sn+s_m\} \times \{sq+s_m\}},$$

où les notations sont choisies de manière à respecter une certaine homogénéité entre les deux formes statiques. Notons qu'un certain nombre de relations liant les diverses matrices présentes dans les deux formes statiques seront utiles à la suite de ce document :

$$\tilde{C}_s(\theta_A, \theta_C)A_s(\theta_A) = \begin{bmatrix} C_s(\theta_A, \theta_C) \\ 0 \end{bmatrix}, \quad (\text{III-12})$$

$$M_s(\theta_A)B_s(\theta_A, \theta_B) = N_s(\theta_B), \quad (\text{III-13})$$

$$L_s(\boldsymbol{\theta}_C)\mathcal{B}_s(\boldsymbol{\theta}_A, \boldsymbol{\theta}_B) = H_s(\boldsymbol{\theta}_A, \boldsymbol{\theta}_B, \boldsymbol{\theta}_C), \quad (\text{III-14})$$

$$M_s(\boldsymbol{\theta}_A)\mathcal{A}_s(\boldsymbol{\theta}_A) = 0. \quad (\text{III-15})$$

III.4 MINIMISATION DES TERMES RÉSIDUELS

III.4.1 Critère d'optimisation

L'objectif consiste maintenant, comme ce fut le cas pour le contexte certain, à trouver une matrice de parité \mathbf{W} orthogonale à $\mathcal{C}_s(\boldsymbol{\theta}_A, \boldsymbol{\theta}_C)$ dans le cas de la première forme statique (III-8) (respectivement $\tilde{\mathcal{C}}_s(\boldsymbol{\theta}_A, \boldsymbol{\theta}_C)$ dans le cas de la seconde forme statique (III-11)); la difficulté résidant dans la nature incertaine de ces matrices lorsque des incertitudes affectent les matrices d'état et/ou d'observation. Intéressons nous donc plus particulièrement à ce contexte incertain. De manière à simplifier les écritures, lorsqu'il n'est pas nécessaire de préciser les matrices affectées par les incertitudes, la notation $\mathcal{C}_s(\boldsymbol{\theta}_A, \boldsymbol{\theta}_C)$ sera remplacée par celle plus compacte de $\mathcal{C}_s(\boldsymbol{\theta})$.

Trouver un vecteur ligne \mathbf{w}^T le plus orthogonal possible à $\mathcal{C}_s(\boldsymbol{\theta})$ consiste à minimiser au mieux l'ensemble des quantités $\mathbf{w}^T \mathcal{C}_s(\boldsymbol{\theta})$ pour toutes les valeurs possibles du vecteur incertain $\boldsymbol{\theta} \in \mathbb{R}^{r(s+1)}$ constitué des variables bornées normalisées $\theta_i(k+j)$, $i \in \{1, \dots, r\}$, $j \in \{k, \dots, k+s\}$. La solution proposée dans ce document vise à obtenir le vecteur ligne \mathbf{w}^T en minimisant le critère suivant (Adrot et al, 1998, 1999b) :

$$\begin{aligned} j_c &= \int_{-1}^1 \dots \int_{-1}^1 \left\| \mathbf{w}^T \mathcal{C}_s(\boldsymbol{\theta}) \right\|^2 d\theta_1(k) \dots d\theta_1(k+s) \dots d\theta_r(k) \dots d\theta_r(k+s) \\ &= \mathbf{w}^T \underbrace{\left(\int_{-1}^1 \dots \int_{-1}^1 \mathcal{C}_s(\boldsymbol{\theta}) \mathcal{C}_s(\boldsymbol{\theta})^T d\theta_1(k) \dots d\theta_1(k+s) \dots d\theta_r(k) \dots d\theta_r(k+s) \right)}_{\mathbf{S}} \mathbf{w} \end{aligned}, \quad (\text{III-16})$$

sous la contrainte $\mathbf{w}^T \mathbf{w} = 1$ (pour éliminer la solution triviale $\mathbf{w}^T = 0$), ce qui permet de mesurer cette orthogonalité. La norme euclidienne a été choisie de manière à résoudre plus facilement ce problème d'optimisation qui se traduit finalement par la recherche de la plus petite valeur propre λ_{\min} de la matrice carrée \mathbf{S} définie ci-dessus et consiste à prendre pour valeur de \mathbf{w} le vecteur propre associé \mathbf{v}_{\min} (le développement conduisant à ce résultat est proposé en annexe B). D'après l'expression du critère j_c , sa valeur minimale est alors donnée par :

$$j_{c, \min} = \mathbf{v}_{\min}^T \mathbf{S} \mathbf{v}_{\min} = \lambda_{\min} \mathbf{v}_{\min}^T \mathbf{v}_{\min} = \lambda_{\min}.$$

Ceci signifie que si \mathbf{S} est singulière, elle possède alors au moins une valeur propre nulle pour laquelle il est possible d'annuler le critère j_c . La matrice \mathbf{S} joue en quelque sorte le même rôle au niveau du couplage entre résidus et l'état du système que la matrice d'observabilité \mathcal{C}_s présentée dans le cas certain puisque la condition d'existence de la matrice de parité dans le contexte incertain est que \mathbf{S} ne soit pas de plein rang.

Cependant, pour pouvoir utiliser efficacement ce critère, il est nécessaire de disposer de l'expression analytique de la matrice \mathbf{S} . Plaçons-nous dans le cas où la matrice $\mathbf{C}_s(\boldsymbol{\theta})$ est linéaire par rapport aux variables bornées et se présente sous la forme d'une sommation (nous expliciterons ultérieurement quelles conditions permettent cette représentation) :

$$\mathbf{C}_s(\boldsymbol{\theta}) = \mathbf{C}_{s,0} + \sum_{i=1}^r \sum_{j=0}^s \theta_i(k+j) \mathbf{C}_{s,i,j}, \quad (\text{III-17})$$

où $\mathbf{C}_{s,0}$ représente la partie nominale de $\mathbf{C}_s(\boldsymbol{\theta})$ et les matrices $\mathbf{C}_{s,i,j}$ sont certaines. Il est possible d'exhiber l'expression analytique de la matrice \mathbf{S} (ce calcul est présenté en annexe C) :

$$\mathbf{S} = 2^{(s+1)r} \left(\mathbf{C}_{s,0} \mathbf{C}_{s,0}^T + \frac{1}{3} \left(\mathbf{C}_{s,1,0} \mathbf{C}_{s,1,0}^T + \dots + \mathbf{C}_{s,r,s} \mathbf{C}_{s,r,s}^T \right) \right). \quad (\text{III-18})$$

Justifions maintenant l'équivalence entre le fait d'annuler le critère j_c et le fait que dans les relations de parité obtenues, les termes dépendant de l'état $\mathbf{x}(k)$, appelés termes résiduels, puissent être nuls. Une fois la première forme statique (III-8) multipliée par le vecteur ligne \mathbf{w}^T , les termes résiduels sont donnés par $\mathbf{w}^T \mathbf{C}_s(\boldsymbol{\theta}) \mathbf{x}(k)$. Lorsque la structure de la matrice $\mathbf{C}_s(\boldsymbol{\theta})$ se présente sous la forme additive (III-17), puisque le vecteur d'état $\mathbf{x}(k)$ peut prendre à tout instant une valeur quelconque, la nullité des termes résiduels impose :

$$\mathbf{w}^T \mathbf{C}_s(\boldsymbol{\theta}) = \mathbf{w}^T \mathbf{C}_{s,0} + \sum_{i=1}^r \sum_{j=0}^s \theta_i(k+j) \mathbf{w}^T \mathbf{C}_{s,i,j} = 0, \quad |\theta_i(k+j)| \leq 1.$$

Ce terme ne peut être nul pour toutes les valeurs possibles des paramètres incertains $\theta_i(k+j)$ que si simultanément toutes les conditions suivantes sont respectées :

$$\mathbf{w}^T \mathbf{C}_{s,0} = 0 \text{ et } \mathbf{w}^T \mathbf{C}_{s,i,j} = 0, \quad i \in \{1, \dots, r\}, j \in \{0, \dots, s\}.$$

Lorsque le critère j_c est nul, en utilisant la relation (III-18), l'égalité suivante est obtenue :

$$j_c = \mathbf{w}^T \mathbf{S} \mathbf{w} = 2^{(s+1)r} \left\| \left[\begin{array}{ccc} \mathbf{C}_{s,0} & \frac{\mathbf{C}_{s,1,0}}{\sqrt{3}} & \dots & \frac{\mathbf{C}_{s,r,s}}{\sqrt{3}} \end{array} \right]^T \mathbf{w} \right\|^2 = 0.$$

La norme euclidienne n'étant nulle que lorsque son argument est nul, la contrainte précédente est équivalente aux $r(s+1)+1$ contraintes explicitées juste avant. Par conséquent, s'il est possible de trouver un vecteur ligne \mathbf{w}^T permettant d'annuler le critère j_c , c'est-à-dire une taille de fenêtre pour laquelle la matrice \mathbf{S} a une valeur propre nulle, alors les termes résiduels dépendant de l'état seront eux aussi nuls et les résidus qui en découlent seront parfaitement découplés des variables d'état inconnues. Tous ces résultats sont directement transposables dans le cas de la seconde forme statique où l'objectif est de trouver un vecteur ligne $\tilde{\mathbf{w}}^T$ le plus orthogonal possible à la matrice $\tilde{\mathbf{C}}_s(\boldsymbol{\theta})$. La transposition s'opère simplement en remplaçant les notations \mathbf{w} , \mathbf{S} , \mathbf{C}_s et j_c par leurs homologues $\tilde{\mathbf{w}}$, $\tilde{\mathbf{S}}$, $\tilde{\mathbf{C}}_s$ et \tilde{j}_c .

$$\tilde{j}_c = \tilde{\mathbf{w}}^T \tilde{\mathbf{S}} \tilde{\mathbf{w}} \text{ où } \tilde{\mathbf{S}} = 2^{(s+1)r} \left(\tilde{\mathbf{C}}_{s,0} \tilde{\mathbf{C}}_{s,0}^T + \frac{1}{3} \left(\tilde{\mathbf{C}}_{s,1,0} \tilde{\mathbf{C}}_{s,1,0}^T + \dots + \tilde{\mathbf{C}}_{s,r,s} \tilde{\mathbf{C}}_{s,r,s}^T \right) \right), \quad (\text{III-19})$$

lorsque la matrice $\tilde{\mathbf{C}}_s(\boldsymbol{\theta})$ se met sous la même forme que (III-17).

Ce critère d'optimisation, permettant d'étendre le principe de la génération de résidus par Espace de Parité au cas incertain, est présenté dans un contexte général où les incertitudes peuvent affecter n'importe quelle matrice de la représentation d'état. Il est donc bien difficile dans ces conditions de déterminer a priori si la matrice S (ou \tilde{S}) a des chances de ne pas être de plein rang. De plus, générer un seul résidu est insuffisant pour détecter et localiser l'ensemble des défauts ciblés par le cahier des charges. La matrice S doit par conséquent posséder autant de valeurs propres nulles (donc de pertes de rang) que nécessaire. Se pose alors un certain nombre de questions. Si la nullité du critère d'optimisation assure bien l'élimination des termes résiduels, en revanche est-il toujours possible d'aboutir à une matrice S (ou \tilde{S}) qui ne soit pas de plein rang ? Dans le cas contraire, existe-t-il une structure des incertitudes permettant de résoudre cette difficulté ? Dans quelles conditions l'une des deux formes statiques est-elle la plus intéressante ? La stratégie utilisée dans la suite de ce document, d'une certaine manière empruntée à l'approche ensembliste, consiste à diviser le problème en sous-problèmes élémentaires dédiés plus spécifiquement à certaines classes de modèles incertains (suivant le type⁷ ou la structure⁸ des incertitudes). Pour chaque cas traité, l'étude de l'influence de la taille de fenêtre sur la possibilité d'obtenir un découplage total sera développée de manière à dégager un certain nombre de propriétés élémentaires permettant de choisir la représentation la plus adaptée et la valeur de s associée.

III.4.2 Expressions générales des formes de calcul et d'évaluation

De manière à pouvoir choisir l'horizon temporel s , il convient de définir dans le contexte incertain, les formes de calcul et d'évaluation des relations de parité, permettant respectivement de calculer les résidus et l'ensemble de leurs valeurs admissibles en bon fonctionnement. Comme la forme de calcul ne peut contenir que des grandeurs connues (mesures, commandes, termes certains), tous les termes associés aux paramètres incertains et/ou aux variables d'état sont regroupés dans la forme d'évaluation.

- Première forme statique

Dans le cas de la première forme statique, une relation de parité est obtenue en multipliant la relation (III-8) par le vecteur ligne w^T approprié, élément de $\mathbb{R}^{1 \times s_m}$:

$$\underbrace{w^T (Y(k,s) - H_{s,0}U(k,s-1))}_{P_{cal}(k+s)} = \underbrace{w^T C_s(\theta_A, \theta_C)x(k)}_{\text{termes résiduels}} + \underbrace{w^T H_{s,\theta}(\theta_A, \theta_B, \theta_C)U(k,s-1)}_{\text{termes incertains}}, \quad \text{(III-20)}$$

$$\underbrace{P_{eval}(k+s)}$$

où les matrices $H_{s,0} = H_s(0,0,0)$ et $H_{s,\theta}(\theta_A, \theta_B, \theta_C)$ désignent respectivement les parties certaine et incertaine de la matrice $H_s(\theta_A, \theta_B, \theta_C)$.

⁷ Dans le cas d'une représentation d'état dynamique coexistent 3 types d'incertitudes ($\theta_A, \theta_B, \theta_C$) suivant qu'elles affectent la matrice d'état, de commande ou d'observation.

⁸ Nous serons amenés par la suite à considérer différentes structures d'incertitudes. La plus classique consiste à définir une matrice incertaine sous la forme d'une somme de termes élémentaires linéaires en les variables bornées tels que (III-17). Une autre solution peut être la représentation adoptée dans le cas de tolérances ou de précisions relatives telles que pour l'exemple introductif de la résistance électrique dans le chapitre I.

La forme de calcul des résidus s'appuie nécessairement sur la partie certaine ou nominale, car connue du système. De ce fait, son expression analytique est indépendante du type et de la structure des incertitudes puisqu'elle ne les fait pas apparaître de manière explicite (même si implicitement, le calcul de \mathbf{w} y fait référence). Cette expression est donc générale et sera référencée à maintes occasions dans la suite de ce document. L'expression de la forme d'évaluation est somme toute aussi générale puisqu'elle est composée d'une part des termes résiduels dépendant de l'état $\mathbf{x}(k)$ qu'il faudrait éliminer, et d'autre part des termes incertains regroupant tous les autres termes mal connus. Les termes résiduels ont été conservés dans cette relation dans le but de rappeler que, même si le vecteur \mathbf{w} a été calculé de manière à minimiser leur influence sur les résidus, il n'est pas toujours possible d'obtenir un découplage total. En l'absence de bruits de mesures et de défauts, les résidus ne sont pas nuls à l'opposé du cas certain et dépendent des vecteurs de paramètres incertains $\boldsymbol{\theta}_A$, $\boldsymbol{\theta}_B$, $\boldsymbol{\theta}_C$, c'est-à-dire de la mauvaise connaissance que nous avons du système, et ce, même si les termes résiduels ont été totalement éliminés.

- *Seconde forme statique*

Dans le cas de la seconde forme statique, une relation de parité est obtenue en multipliant la relation (III-11) par le vecteur ligne $\tilde{\mathbf{w}}^T$ approprié, élément de $\mathbb{R}^{1 \times \{sn+s_m\}}$:

$$\underbrace{\tilde{\mathbf{w}}^T \tilde{\mathbf{H}}_{s,0} \begin{bmatrix} U(k, s-1) \\ Y(k, s) \end{bmatrix}}_{\tilde{\mathbf{p}}_{cal}(k+s)} = \underbrace{\tilde{\mathbf{w}}^T \tilde{\mathbf{C}}_s(\boldsymbol{\theta}_A, \boldsymbol{\theta}_C) \mathbf{X}(k, s)}_{\text{termes résiduels}} - \underbrace{\tilde{\mathbf{w}}^T \tilde{\mathbf{H}}_{s,\boldsymbol{\theta}}(\boldsymbol{\theta}_B)}_{\text{termes incertains}} \begin{bmatrix} U(k, s-1) \\ Y(k, s) \end{bmatrix} \quad (\text{III-21})$$

où les matrices $\tilde{\mathbf{H}}_{s,0} = \tilde{\mathbf{H}}_s(0)$ et $\tilde{\mathbf{H}}_{s,\boldsymbol{\theta}}(\boldsymbol{\theta}_B)$ désignent respectivement les parties certaine et incertaine de la matrice $\tilde{\mathbf{H}}_s(\boldsymbol{\theta}_B)$.

Les expressions analytiques des formes de calcul et d'évaluation sont tout aussi générales que pour la première forme statique et s'appliquent quelles que soient les incertitudes. À ce stade, se dégage une nouvelle question : existe-t-il une relation entre les résidus obtenus par les deux formes statiques ? Les expressions des formes de calcul et d'évaluation dépendent de manière implicite des incertitudes à cause des méthodes de calcul de \mathbf{w} et $\tilde{\mathbf{w}}$, il est donc nécessaire de développer les expressions générales exposées ci-dessus de manière à tenir compte d'un lien pouvant exister entre \mathbf{w} et $\tilde{\mathbf{w}}$. Ce travail sera effectué lors des sections suivantes pour chaque classe de modèle incertain étudié.

III.5 MATRICE DE COMMANDE INCERTAINE

Dans cette section, les formes d'évaluation obtenues lorsque les incertitudes n'affectent que la matrice de commande sont explicitées.

III.5.1 Structure classique

Traditionnellement dans une représentation d'état telle que (III-4), les incertitudes structurées apparaissent de manière linéaire et sont définies au travers de sommations :

$$\mathbf{Z}(\boldsymbol{\theta}_Z(k)) = \mathbf{Z}_0 + \sum_{i=1}^{r_Z} \theta_{Z,i}(k) \mathbf{Z}_i, \quad \mathbf{Z} \in \{\mathbf{A}, \mathbf{B}, \mathbf{C}\}.$$

où $\mathbf{Z}_0 = \mathbf{Z}(0)$ définit la partie nominale de la matrice incertaine $\mathbf{Z}(\boldsymbol{\theta}_Z)$ et où les matrices de distribution \mathbf{Z}_i associées à chaque paramètre incertain $\theta_{Z,i}(k)$ sont supposées parfaitement connues. Nous qualifierons cette modélisation de structure classique.

III.5.1.1 Première forme statique

Compte tenu de la définition précédente, la matrice de commande s'écrit :

$$\mathbf{B}(\boldsymbol{\theta}_B(k)) = \mathbf{B}_0 + \sum_{i=1}^r \theta_i(k) \mathbf{B}_i.$$

La matrice \mathbf{C}_s définie en (III-8) est alors certaine et correspond rigoureusement à la matrice d'observabilité d'ordre s (III-2) définie dans le cas d'un modèle certain. Un découplage total des relations de parité vis-à-vis des variables d'état est donc toujours possible puisqu'il est montré dans la section III.2.2 qu'il existe toujours une taille de fenêtre temporelle pour laquelle la condition d'existence de \mathbf{w} est satisfaite. Le vecteur de projection \mathbf{w} est par conséquent obtenu par un calcul traditionnel identique au contexte certain. La matrice $\mathbf{H}_s(\boldsymbol{\theta}_B)$, quant à elle, est incertaine et peut être mise sous la forme suivante :

$$\mathbf{H}_s(\boldsymbol{\theta}_B) = \mathbf{H}_{s,0} + \mathbf{H}_{s,\theta}(\boldsymbol{\theta}_B) \text{ avec } \mathbf{H}_{s,\theta}(\boldsymbol{\theta}_B) = \sum_{i=1}^r \sum_{j=0}^{s-1} \theta_i(k+j) \mathbf{H}_{s,i,j} \quad (\text{III-22})$$

$$\text{où } \mathbf{H}_{s,0} = \begin{bmatrix} 0 & \cdots & 0 \\ \mathbf{CB}_0 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ \mathbf{CA}^{s-1} \mathbf{B}_0 & \cdots & \mathbf{CB}_0 \end{bmatrix} \text{ et } \mathbf{H}_{s,i,j} = \begin{bmatrix} 0 & \cdots & 0 \\ \vdots & \mathbf{CB}_i & \vdots \\ \vdots & \vdots & \vdots \\ 0 & \mathbf{CA}^{s-1-j} \mathbf{B}_i & 0 \end{bmatrix}.$$

|→ (j+1)^{ième} bloc

La forme de calcul étant indépendante des incertitudes, elle reste identique à (III-20) et n'est donc pas rappelée ici. En réalisant un découplage total par rapport à l'état : $\mathbf{w}^T \mathbf{C}_s = 0$, la forme d'évaluation (III-20) devient :

$$p_{\text{éval}}(k+s) = \mathbf{w}^T \mathbf{H}_{s,\theta}(\boldsymbol{\theta}_B) \mathbf{U}(k, s-1) = \sum_{i=1}^r \sum_{j=0}^{s-1} \theta_i(k+j) \mathbf{w}^T \mathbf{H}_{s,i,j} \mathbf{U}(k, s-1).$$

Les termes résiduels sont nuls et les termes incertains ne dépendent que de la commande, de plus, cette expression est totalement linéaire en les paramètres incertains.

III.5.1.2 Seconde forme statique

À partir de la structure de $\mathbf{B}(\boldsymbol{\theta}_B)$, celle de la matrice $\mathbf{N}_s(\boldsymbol{\theta}_B)$ (III-10) est déduite :

$$\mathbf{N}_s(\boldsymbol{\theta}_B) = \mathbf{N}_{s,0} + \mathbf{N}_{s,\theta}(\boldsymbol{\theta}_B) \text{ avec } \mathbf{N}_{s,\theta}(\boldsymbol{\theta}_B) = \sum_{i=1}^r \sum_{j=0}^{s-1} \theta_i(k+j) \mathbf{N}_{s,i,j} \quad (\text{III-23})$$

$$\text{où } \mathbf{N}_{s,0} = \begin{bmatrix} -\mathbf{B}_0 & 0 & 0 \\ 0 & \ddots & 0 \\ 0 & 0 & -\mathbf{B}_0 \end{bmatrix} \text{ et } \mathbf{N}_{s,i,j} = \begin{bmatrix} 0 & \cdots & 0 \\ \vdots & -\mathbf{B}_i & \vdots \\ 0 & \cdots & 0 \end{bmatrix} \begin{matrix} \downarrow (j+1)\text{ième bloc} \\ \\ \uparrow (j+1)\text{ième bloc} \end{matrix},$$

Cette expression conduit, à partir de (III-11), à :

$$\tilde{\mathbf{H}}_s(\boldsymbol{\theta}_B) = \tilde{\mathbf{H}}_{s,0} + \tilde{\mathbf{H}}_{s,\theta}(\boldsymbol{\theta}_B) \text{ avec } \tilde{\mathbf{H}}_{s,\theta}(\boldsymbol{\theta}_B) = \sum_{i=1}^r \sum_{j=0}^{s-1} \theta_i(k+j) \tilde{\mathbf{H}}_{s,i,j},$$

$$\text{où } \tilde{\mathbf{H}}_{s,0} = \begin{bmatrix} 0 & \mathbf{I}_{s^m} \\ \mathbf{N}_{s,0} & 0 \end{bmatrix} \text{ et } \tilde{\mathbf{H}}_{s,i,j} = \begin{bmatrix} 0 & 0 \\ \mathbf{N}_{s,i,j} & 0 \end{bmatrix}.$$

La forme de calcul étant indépendante des incertitudes, elle reste identique à (III-21) et n'est donc pas rappelée ici. La matrice $\tilde{\mathbf{C}}_s$ étant certaine, il est légitime de supposer qu'un découplage total est possible⁹ : $\tilde{\mathbf{w}}^T \tilde{\mathbf{C}}_s = 0$. La forme d'évaluation (III-21) devient alors :

$$\tilde{p}_{\text{éval}}(k+s) = -\tilde{\mathbf{w}}^T \tilde{\mathbf{H}}_{s,\theta}(\boldsymbol{\theta}_B) \begin{bmatrix} \mathbf{U}(k, s-1) \\ \mathbf{Y}(k, s) \end{bmatrix}.$$

Compte tenu de la structure de la matrice $\tilde{\mathbf{H}}_{s,\theta}(\boldsymbol{\theta}_B)$ qui possède des lignes nulles, il est judicieux de décomposer le vecteur $\tilde{\mathbf{w}}$ en deux vecteurs $\tilde{\mathbf{w}}_L$ et $\tilde{\mathbf{w}}_M$, éléments respectifs de \mathbb{R}^{s^m} et \mathbb{R}^{s^m} :

$$\begin{aligned} \tilde{p}_{\text{éval}}(k+s) &= -\begin{bmatrix} \tilde{\mathbf{w}}_L^T & \tilde{\mathbf{w}}_M^T \end{bmatrix} \begin{bmatrix} 0 & 0 \\ \mathbf{N}_{s,\theta}(\boldsymbol{\theta}_B) & 0 \end{bmatrix} \begin{bmatrix} \mathbf{U}(k, s-1) \\ \mathbf{Y}(k, s) \end{bmatrix} \\ &= -\tilde{\mathbf{w}}_M^T \mathbf{N}_{s,\theta}(\boldsymbol{\theta}_B) \mathbf{U}(k, s-1) = -\sum_{i=1}^r \sum_{j=0}^{s-1} \theta_i(k+j) \tilde{\mathbf{w}}_M^T \mathbf{N}_{s,i,j} \mathbf{U}(k, s-1) \end{aligned}$$

Tout comme pour la première forme statique, les termes résiduels sont nuls et les termes incertains ne dépendent que de la commande. De plus, cette expression est parfaitement linéaire en les paramètres incertains.

⁹ Nous montrerons dans la section suivante III.5.1.3 que pour une même valeur de la taille de fenêtre, les deux formes statiques conduisent à des relations de parité structurellement équivalentes, ce qui justifiera cette hypothèse.

III.5.1.3 Relation entre les deux formes statiques

L'objectif de cette section est de rechercher s'il existe une relation entre les résidus obtenus pour les deux formes statiques et une même taille de fenêtre temporelle lorsque la matrice de commande est incertaine.

- *Étape 1 : Relation entre les matrices de parité \mathbf{W} et $\tilde{\mathbf{W}}$*

① Plaçons-nous dans le cadre de la seconde forme statique. Soit $\tilde{\mathbf{W}}$ une matrice vérifiant la contrainte d'orthogonalité $\tilde{\mathbf{W}}\tilde{\mathbf{C}}_s = 0$ pour une taille de fenêtre s donnée. Cette matrice est supposée de plein rang ligne de manière à ne conserver que des résidus linéairement indépendants, son rang est donc imposé par : $s_{\tilde{w}} = s_m + sn - \text{rang}(\tilde{\mathbf{C}}_s)$. En réalisant la partition de $\tilde{\mathbf{W}}$ en deux sous-matrices $\tilde{\mathbf{W}}_L$ et $\tilde{\mathbf{W}}_M$ de dimensions respectives $s_{\tilde{w}} \times s_m$ et $s_{\tilde{w}} \times sn$ comme cela avait été fait dans la section précédente :

$$\tilde{\mathbf{W}} = \left[\tilde{\mathbf{W}}_L \mid \tilde{\mathbf{W}}_M \right],$$

la contrainte d'orthogonalité devient :

$$\left[\tilde{\mathbf{W}}_L \mid \tilde{\mathbf{W}}_M \right] \begin{bmatrix} L_s \\ M_s \end{bmatrix} = \tilde{\mathbf{W}}_L L_s + \tilde{\mathbf{W}}_M M_s = 0. \quad (\text{III-24})$$

En multipliant à droite l'égalité précédente par la matrice \mathbf{A}_s (III-6), puis en utilisant le fait que le produit $M_s \mathbf{A}_s$ (III-15) est nul, il s'en déduit à l'aide de (III-12) que :

$$\tilde{\mathbf{W}}_L \mathbf{C}_s = 0.$$

Montrons dans un premier temps que la sous-matrice $\tilde{\mathbf{W}}_L$ est nécessairement de plein rang ligne. Si elle ne l'était pas, alors il existerait un vecteur non nul $\omega \in \mathbb{R}^{s_{\tilde{w}}}$ tel que le produit $\omega^T \tilde{\mathbf{W}}_L$ est nul. Dans ces conditions, la relation (III-24) imposerait d'avoir $\omega^T \tilde{\mathbf{W}}_M M_s = 0$. Vu la structure de la matrice M_s , cette dernière est forcément de plein rang ligne, ce qui restreint l'égalité précédente à $\omega^T \tilde{\mathbf{W}}_M = 0$. Par conséquent subsiste une contradiction dans la mesure où la seule conclusion recevable consiste en l'existence d'un vecteur ω non nul tel que $\omega^T \left[\tilde{\mathbf{W}}_L \mid \tilde{\mathbf{W}}_M \right] = 0$, ce qui voudrait insinuer que $\tilde{\mathbf{W}}$ n'est pas de plein rang ligne.

La décomposition en valeurs singulières de la matrice \mathbf{C}_s :

$$\mathbf{C}_s = \mathbf{U} \boldsymbol{\Sigma} \mathbf{V}^T \text{ où } \mathbf{U} \text{ et } \mathbf{V} \text{ sont des matrices orthogonales et } \boldsymbol{\Sigma} \text{ une matrice diagonale,}$$

permet d'obtenir le noyau de \mathbf{C}_s , c'est-à-dire une base génératrice de toutes les matrices $\tilde{\mathbf{W}}_L$ pouvant lui être orthogonale (annexe D). Puisque la matrice $\tilde{\mathbf{W}}_L$ est de plein rang ligne, il existe donc une matrice carrée $\tilde{\mathbf{K}}$ d'ordre $s_w = s_m - s_c$ inversible telle que :

$$\tilde{\mathbf{W}}_L = \tilde{\mathbf{K}} \begin{bmatrix} 0 & \mathbf{I}_{s_w} \end{bmatrix} \mathbf{U}^T \quad (\text{III-25})$$

où s_c représente le rang de la matrice \mathbf{C}_s .

② Intéressons-nous maintenant à la première forme statique. Soit \mathbf{W} une matrice de parité de dimension $s_w \times s_m$ orthogonale à \mathbf{C}_s obtenue de manière traditionnelle comme développé dans le contexte certain (section III.2). Il existe alors une matrice carrée \mathbf{K} d'ordre s_w inversible telle que :

$$\mathbf{W} = \mathbf{K} \begin{bmatrix} 0 & \mathbf{I}_{s_w} \end{bmatrix} \mathbf{U}^T. \quad (\text{III-26})$$

En comparant les contraintes imposées par les deux formes statiques, il est alors évident que l'existence de la matrice $\tilde{\mathbf{W}}$ ① (c'est à dire le fait que $\tilde{\mathbf{C}}_s$ ne soit pas de plein rang ligne) conduit nécessairement à l'existence de la matrice \mathbf{W} ② puisqu'il existe alors une matrice $\tilde{\mathbf{W}}_L$ orthogonale à \mathbf{C}_s . De même, la réciproque est vérifiée. En effet, si \mathbf{W} ② existe, alors, puisque \mathbf{M}_s est de plein rang ligne, il est possible de poser :

$$\tilde{\mathbf{W}}_L = \mathbf{W} \text{ et } \tilde{\mathbf{W}}_M = -\mathbf{W} \mathbf{L}_s \mathbf{M}_s^T (\mathbf{M}_s \mathbf{M}_s^T)^{-1}.$$

Par conséquent, la matrice $\tilde{\mathbf{W}} = \begin{bmatrix} \tilde{\mathbf{W}}_L & \tilde{\mathbf{W}}_M \end{bmatrix}$ ainsi définie est parfaitement déterminée et vérifie les contraintes associées à ① d'après (III-24). En conclusion, les conditions d'existence des matrices de parité associées aux deux formes statiques sont équivalentes. Si \mathbf{C}_s n'est pas de plein rang ligne, il en est de même pour $\tilde{\mathbf{C}}_s$ et vice versa.

Par définition, la matrice $\tilde{\mathbf{W}}_L$ étant de dimension $s_{\tilde{\mathbf{W}}} \times s_m$, il s'en déduit que :

$$s_{\tilde{\mathbf{W}}} = s_m + sn - \text{rank}(\tilde{\mathbf{C}}_s) = s_m - s_c = s_w, \text{ ou encore } \text{rank}(\tilde{\mathbf{C}}_s) = sn + s_c.$$

Par conséquent, quelle que soit la forme statique utilisée, le même nombre de relations de parité linéairement indépendantes seront obtenues. Les matrices \mathbf{C}_s et $\tilde{\mathbf{C}}_s$ ont exactement le même nombre de valeurs propres nulles pour une même valeur de s ; la taille de fenêtre temporelle pour la seconde forme statique peut donc être déduite de l'étude de la première forme statique. Les relations (III-25) et (III-26) montrent que la matrice $\tilde{\mathbf{W}}_L$ peut s'exprimer comme une combinaison linéaire des lignes de la matrice \mathbf{W} :

$$\tilde{\mathbf{W}}_L = \mathbf{T} \mathbf{W} \text{ où la matrice carrée } \mathbf{T} \in \mathbb{R}^{s_w \times s_w} \text{ est inversible.} \quad (\text{III-27})$$

- Étape 2 : Relation entre les résidus obtenus pour les deux formes statiques

Recherchons maintenant une relation entre les formes de calcul puis d'évaluation associées aux deux formes statiques. À partir de l'expression de la matrice de commande, celle de la matrice $\mathbf{B}_s(\boldsymbol{\theta}_B)$ (III-6) est déduite :

$$\mathbf{B}_s(\boldsymbol{\theta}_B) = \mathbf{B}_{s,0} + \mathbf{B}_{s,\theta}(\boldsymbol{\theta}_B) \text{ avec } \mathbf{B}_{s,\theta}(\boldsymbol{\theta}_B) = \sum_{i=1}^r \sum_{j=0}^{s-1} \theta_i (k+j) \mathbf{B}_{s,i,j},$$

$$\text{avec } \mathbf{B}_{s,0} = \begin{bmatrix} 0 & \cdots & 0 \\ \mathbf{B}_0 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ \mathbf{A}^{s-1} \mathbf{B}_0 & \cdots & \mathbf{B}_0 \end{bmatrix} \text{ et } \mathbf{B}_{s,i,j} = \begin{bmatrix} 0 & \cdots & 0 \\ \vdots & \mathbf{B}_i & \vdots \\ \vdots & \vdots & \vdots \\ 0 & \mathbf{A}^{s-1-j} \mathbf{B}_i & 0 \end{bmatrix} \begin{matrix} \downarrow (j+2)\text{ième bloc} \\ \\ \uparrow (j+1)\text{ième bloc} \end{matrix}.$$

À partir des relations génériques (III-13) et (III-14), les égalités suivantes sont déduites, respectivement pour les parties nominales et celles dépendant des incertitudes paramétriques :

$$\begin{aligned} \mathbf{M}_s \mathbf{B}_{s,0} &= \mathbf{N}_{s,0} \text{ et } \mathbf{M}_s \mathbf{B}_{s,\theta}(\boldsymbol{\theta}_B) = \mathbf{N}_{s,\theta}(\boldsymbol{\theta}_B), \\ \mathbf{L}_s \mathbf{B}_{s,0} &= \mathbf{H}_{s,0} \text{ et } \mathbf{L}_s \mathbf{B}_{s,\theta}(\boldsymbol{\theta}_B) = \mathbf{H}_{s,\theta}(\boldsymbol{\theta}_B), \end{aligned}$$

où les matrices $\mathbf{H}_{s,0}$, $\mathbf{H}_{s,\theta}$ puis $\mathbf{N}_{s,0}$, $\mathbf{N}_{s,\theta}$ ont été définies respectivement en (III-22) et (III-23). De la relation (III-24) et de ce qui précède, se déduisent les relations suivantes :

$$\tilde{\mathbf{W}}_M \mathbf{N}_{s,0} = -\tilde{\mathbf{W}}_L \mathbf{H}_{s,0} \text{ et } \tilde{\mathbf{W}}_M \mathbf{N}_{s,\theta}(\boldsymbol{\theta}_B) = -\tilde{\mathbf{W}}_L \mathbf{H}_{s,\theta}(\boldsymbol{\theta}_B). \quad (\text{III-28})$$

En utilisant la forme de calcul associée à la seconde forme statique (III-21) et l'expression de la matrice $\tilde{\mathbf{H}}_{s,0}$, ainsi que la partition de $\tilde{\mathbf{W}}$ en deux sous-matrices $\tilde{\mathbf{W}}_L$ et $\tilde{\mathbf{W}}_M$, nous obtenons :

$$\tilde{\mathbf{p}}_{cal}(k+s) = \tilde{\mathbf{W}} \tilde{\mathbf{H}}_{s,0} \begin{bmatrix} \mathbf{U}(k, s-1) \\ \mathbf{Y}(k, s) \end{bmatrix} = \tilde{\mathbf{W}}_M \mathbf{N}_{s,0} \mathbf{U}(k, s-1) + \tilde{\mathbf{W}}_L \mathbf{Y}(k, s),$$

ce qui, en utilisant (III-28) et la relation d'équivalence (III-27), se réduit à :

$$\tilde{\mathbf{p}}_{cal}(k+s) = \tilde{\mathbf{W}}_L (-\mathbf{H}_{s,0} \mathbf{U}(k, s-1) + \mathbf{Y}(k, s)) = \mathbf{T} \mathbf{p}_{cal}(k+s).$$

À partir de la forme d'évaluation associée à la seconde forme statique pour des incertitudes affectant la matrice de commande, de (III-28) puis de l'équivalence (III-27), la relation suivante est obtenue :

$$\begin{aligned} \tilde{\mathbf{p}}_{éval}(k+s) &= -\tilde{\mathbf{W}}^T \tilde{\mathbf{H}}_{s,\theta}(\boldsymbol{\theta}_B) \begin{bmatrix} \mathbf{U}(k, s-1) \\ \mathbf{Y}(k, s) \end{bmatrix} = -\tilde{\mathbf{W}}_M \mathbf{N}_{s,\theta} \mathbf{U}(k, s-1) \\ &= \tilde{\mathbf{W}}_L \mathbf{H}_{s,\theta} \mathbf{U}(k, s-1) = \mathbf{T} \mathbf{p}_{éval}(k+s) \end{aligned}$$

La matrice \mathbf{T} étant inversible par définition, pour une taille de fenêtre s donnée, les formes de calcul et d'évaluation sont structurellement équivalentes pour les deux formes statiques lorsque les incertitudes affectent la matrice de commande (structure classique) et sous les contraintes d'un découplage total. La seconde forme statique permet donc bien d'aboutir à des relations de parité et il est alors naturel de prendre une taille de fenêtre identique à celle trouvée pour la forme statique traditionnelle. Pour les deux formes statiques, le choix de s est donc obtenu par l'étude de la matrice d'observabilité du système, comme explicité dans la section III.2.2.

Remarque. Le même schéma de démonstration peut être appliqué pour démontrer cette même relation d'équivalence dans un contexte purement certain (annexe E).

III.5.2 Structure de modèle de capteurs

Des incertitudes affectant la matrice de commande peuvent servir à prendre en compte des imperfections de la chaîne d'instrumentation associées aux entrées commandées, telles que des tolérances liées aux non-linéarités de capteurs ou des bruits de mesures (Adrot et al, 1999a, 2000a). Les premières apparaissent de manière multiplicative dans le modèle en

affectant directement les coefficients de la matrice de commande alors que les seconds peuvent être modélisés par des incertitudes additives supposées bornées. Le modèle d'état peut dans ces conditions être décrit comme suit :

$$\begin{cases} \mathbf{x}(k+1) = \mathbf{A}\mathbf{x}(k) + \mathbf{B}_0\mathbf{u}^*(k) \\ \mathbf{y}(k) = \mathbf{C}\mathbf{x}(k) \\ \mathbf{u}(k) = (\mathbf{I}_q + \mathbf{T}_\Delta\mathbf{A}(k))\mathbf{u}^*(k) + \mathbf{T}_\delta\boldsymbol{\delta}(k) \end{cases} \quad \begin{array}{l} \mathbf{x} \in \mathbb{R}^n, \mathbf{y} \in \mathbb{R}^m, \mathbf{u} \in \mathbb{R}^q \\ \mathbf{u}^* \in \mathbb{R}^q, \boldsymbol{\delta} \in \mathbb{R}^q, \mathbf{A} \in \mathbb{R}^{q \times q} \end{array} \quad (\text{III-29})$$

où \mathbf{u}^* est la commande réelle et inconnue appliquée au système et \mathbf{u} le vecteur de mesures disponible associé. Le vecteur \mathbf{u}^* ne tient compte que de la partie certaine du système de commande où la matrice \mathbf{B}_0 indique de quelle manière les entrées influent sur les variables d'état. Le vecteur \mathbf{u} tient compte des imperfections de la chaîne d'instrumentation et permet ainsi de modéliser séparément la confiance accordée à chaque capteur.

La matrice incertaine \mathbf{A} , représentant les incertitudes multiplicatives, est diagonale, carrée d'ordre q et ses éléments diagonaux correspondent à q variables bornées normalisées fluctuant entre -1 et 1 et visant à modéliser individuellement la précision technologique de chaque capteur. La matrice de pondération \mathbf{T}_Δ est certaine, diagonale et carrée d'ordre q . Si la tolérance sur le $i^{\text{ième}}$ capteur est de $\pm\sigma_i$, alors l'élément d'indice (i,i) de la matrice \mathbf{T}_Δ vaut σ_i . Comme il s'agit de précision relative, σ_i est homogène à un pourcentage, ce qui se traduit par $0 \leq \sigma_i < 1$; la matrice $\mathbf{I}_q + \mathbf{T}_\Delta\mathbf{A}(k)$ est donc inversible. Le vecteur $\boldsymbol{\delta}$ est lui aussi constitué de q variables bornées normalisées dépendantes du temps et \mathbf{T}_δ^{10} est une matrice de pondération carrée d'ordre q parfaitement connue permettant de prendre en compte l'amplitude maximale de chaque bruit de mesures supposé borné et centré. Le système précédent conduit à poser :

$$\mathbf{B}(\mathbf{A}) = \mathbf{B}_0(\mathbf{I}_q + \mathbf{T}_\Delta\mathbf{A}(k))^{-1} \text{ et } \mathbf{x}(k+1) = \mathbf{A}\mathbf{x}(k) + \mathbf{B}_0(\mathbf{I}_q + \mathbf{T}_\Delta\mathbf{A}(k))^{-1}(\mathbf{u}(k) - \mathbf{T}_\delta\boldsymbol{\delta}(k)).$$

Cette modélisation de l'incertain est plus restrictive que la forme structurée classique qui permet d'appréhender des systèmes où des incertitudes peuvent apparaître sur des termes de couplage entre capteurs. Elle est néanmoins plus naturelle et mieux adaptée à la prise en compte de tolérances et de bruits affectant les capteurs. En effet la structure classique ne peut plus être utilisée pour ce type de problème dans la mesure où la matrice $\mathbf{B}(\mathbf{A})$ précédente est non-linéaire en les incertitudes et ne peut donc plus être mise sous la forme d'une sommation. De plus, puisque les incertitudes additives ne sont pas en facteur du vecteur de commande, elles ne peuvent pas être représentées par la structure classique.

III.5.2.1 Première forme statique

Déterminons dans un premier temps la nouvelle expression de la forme d'évaluation associée au modèle (III-29) dans le cadre de la première forme statique. La relation (III-5) liant le vecteur $\mathbf{x}(k+j)$ ($j \in \mathbb{N}^*$) à $\mathbf{x}(k)$ et aux commandes devient :

¹⁰ Prendre une matrice \mathbf{T}_δ diagonale signifie que les incertitudes additives touchant chacune des mesures sont indépendantes.

$$\mathbf{x}(k+j) = \mathbf{A}^j \mathbf{x}(k) + \sum_{i=0}^{j-1} \mathbf{A}^{j-i-1} \mathbf{B}_0 (\mathbf{I}_q + \mathbf{T}_\Delta \mathbf{A}(k+i))^{-1} (\mathbf{u}(k+i) - \mathbf{T}_\delta \boldsymbol{\delta}(k+i)),$$

ce qui conduit après un cheminement identique à la section III.3.2 à la forme statique suivante :

$$\mathbf{Y}(k,s) = \mathbf{C}_s \mathbf{x}(k) + \mathbf{H}_s(\Delta) \mathbf{U}(k,s-1) - \mathbf{H}_s(\Delta) \text{diag}_s(\mathbf{T}_\delta) \mathbf{D}_\delta(k,s-1),$$

avec $\mathbf{D}_\delta(k,s-1) = \begin{bmatrix} \boldsymbol{\delta}(k) \\ \vdots \\ \boldsymbol{\delta}(k+s-1) \end{bmatrix}$, $\mathbf{D}_\delta \in \mathbb{R}^{sq}$ et $\text{diag}_s(\mathbf{T}_\delta)$, matrice diagonale par blocs contenant s

blocs de \mathbf{T}_δ . La matrice \mathbf{C}_s est certaine et correspond à la matrice d'observabilité d'ordre s . En revanche, la matrice $\mathbf{H}_s(\Delta)$ est incertaine et se présente sous la forme suivante :

$$\mathbf{H}_s(\Delta) = \begin{bmatrix} 0 & 0 & \dots & 0 \\ \mathbf{C}\mathbf{B}_0(\mathbf{I}_q + \mathbf{T}_\Delta \mathbf{A}(k))^{-1} & 0 & \dots & 0 \\ \mathbf{C}\mathbf{A}\mathbf{B}_0(\mathbf{I}_q + \mathbf{T}_\Delta \mathbf{A}(k))^{-1} & \mathbf{C}\mathbf{B}_0(\mathbf{I}_q + \mathbf{T}_\Delta \mathbf{A}(k+1))^{-1} & & 0 \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{C}\mathbf{A}^{s-1}\mathbf{B}_0(\mathbf{I}_q + \mathbf{T}_\Delta \mathbf{A}(k))^{-1} & \mathbf{C}\mathbf{A}^{s-2}\mathbf{B}_0(\mathbf{I}_q + \mathbf{T}_\Delta \mathbf{A}(k+1))^{-1} & & \mathbf{C}\mathbf{B}_0(\mathbf{I}_q + \mathbf{T}_\Delta \mathbf{A}(k+s-1))^{-1} \end{bmatrix}.$$

En utilisant le lemme d'inversion matricielle, l'égalité suivante est déduite :

$$(\mathbf{I}_q + \mathbf{T}_\Delta \mathbf{A}(k))^{-1} = \mathbf{I}_q - \mathbf{T}_\Delta \mathbf{A}(k) (\mathbf{I}_q + \mathbf{T}_\Delta \mathbf{A}(k))^{-1}, \quad (\text{III-30})$$

ce qui permet d'obtenir :

$$\mathbf{H}_s(\Delta) = \mathbf{H}_{s,0} + \mathbf{H}_{s,\theta}(\Delta), \text{ avec } \mathbf{H}_{s,\theta}(\Delta) = -\mathbf{H}_{s,0} \mathbf{D}_\Delta(k,s-1), \quad (\text{III-31})$$

$$\mathbf{D}_\Delta(k,s-1) = \begin{bmatrix} \mathbf{T}_\Delta \mathbf{A}(k) (\mathbf{I}_q + \mathbf{T}_\Delta \mathbf{A}(k))^{-1} & 0 & 0 \\ 0 & \ddots & 0 \\ 0 & 0 & \mathbf{T}_\Delta \mathbf{A}(k+s-1) (\mathbf{I}_q + \mathbf{T}_\Delta \mathbf{A}(k+s-1))^{-1} \end{bmatrix}, \quad \mathbf{D}_\Delta \in \mathbb{R}^{sq \times sq}.$$

À partir de l'égalité précédente, en respectant la contrainte relative à un découplage total des relations de parité vis-à-vis des variables d'état :

$$\mathbf{w}^T \mathbf{C}_s = 0,$$

l'expression suivante de la première forme statique est obtenue

$$\begin{aligned} \underbrace{\mathbf{w}^T (\mathbf{Y}(k,s) - \mathbf{H}_{s,0} \mathbf{U}(k,s-1))}_{p_{cal}(k+s)} &= \dots \\ &= \underbrace{-\mathbf{w}^T \mathbf{H}_{s,0} \mathbf{D}_\Delta(k,s-1) \mathbf{U}(k,s-1) - \mathbf{w}^T \mathbf{H}_s(\Delta) \text{diag}_s(\mathbf{T}_\delta) \mathbf{D}_\delta(k,s-1)}_{p_{eval}(k+s)}. \end{aligned} \quad (\text{III-32})$$

Comme la structure classique, cette forme d'évaluation est parfaitement découplée vis-à-vis des variables d'état, mais cette expression est non-linéaire en les paramètres incertains.

III.5.2.2 Seconde forme statique

Déterminons dans un premier temps la nouvelle expression de la forme d'évaluation associée au modèle (III-29) dans le cadre cette fois-ci de la seconde forme statique. La relation (III-9) liant le vecteur $\mathbf{x}(k+j)$ ($j \in \mathbb{N}^*$) en fonction des vecteurs antérieurs $\mathbf{x}(k+j-1)$ et des commandes devient :

$$\begin{bmatrix} \mathbf{A} & -\mathbf{I}_n \end{bmatrix} \begin{bmatrix} \mathbf{x}(k+j-1) \\ \mathbf{x}(k+j) \end{bmatrix} = -\mathbf{B}_0 (\mathbf{I}_q + \mathbf{T}_\Delta \mathbf{A}(k+j-1))^{-1} (\mathbf{u}(k+j-1) - \mathbf{T}_\delta \boldsymbol{\delta}(k)),$$

ce qui conduit, en empilant les vecteurs d'état et de commande sur un horizon $[k, k+s]$, à l'expression matricielle suivante :

$$\begin{aligned} \mathbf{M}_s \mathbf{X}(k, s) &= \mathbf{N}_s(\mathbf{A}) \mathbf{U}(k, s-1) - \mathbf{N}_s(\mathbf{A}) \text{diag}_s(\mathbf{T}_\delta) \mathbf{D}_\delta(k, s-1) \\ \text{avec : } \mathbf{N}_s(\mathbf{A}) &= \begin{bmatrix} -\mathbf{B}_0 (\mathbf{I}_q + \mathbf{T}_\Delta \mathbf{A}(k))^{-1} & 0 & 0 \\ 0 & \ddots & 0 \\ 0 & 0 & -\mathbf{B}_0 (\mathbf{I}_q + \mathbf{T}_\Delta \mathbf{A}(k+s-1))^{-1} \end{bmatrix}. \end{aligned}$$

Ensuite, à partir de l'égalité (III-30) issue du lemme d'inversion matricielle, la décomposition suivante est adoptée :

$$\mathbf{N}_s(\mathbf{A}) = \mathbf{N}_{s,0} + \mathbf{N}_{s,\theta}(\mathbf{A}) \text{ où } \mathbf{N}_{s,\theta}(\mathbf{A}) = -\mathbf{N}_{s,0} \mathbf{D}_\Delta(k, s-1). \quad (\text{III-33})$$

En utilisant le même cheminement que dans la section III.3.3, l'expression de la seconde forme statique est donc donnée par :

$$\tilde{\mathbf{C}}_s \mathbf{X}(k, s) = \begin{bmatrix} 0 & \mathbf{I}_{s_m} \\ \mathbf{N}_s(\mathbf{A}) & 0 \end{bmatrix} \begin{bmatrix} \mathbf{U}(k, s-1) \\ \mathbf{Y}(k, s) \end{bmatrix} + \begin{bmatrix} 0 \\ -\mathbf{N}_s(\mathbf{A}) \text{diag}_s(\mathbf{T}_\delta) \mathbf{D}_\delta(k, s-1) \end{bmatrix},$$

L'expression la matrice $\tilde{\mathbf{H}}_s(\mathbf{A})$ est alors donnée par :

$$\tilde{\mathbf{H}}_s(\mathbf{A}) = \begin{bmatrix} 0 & \mathbf{I}_{s_m} \\ \mathbf{N}_s(\mathbf{A}) & 0 \end{bmatrix} = \tilde{\mathbf{H}}_{s,0} + \tilde{\mathbf{H}}_{s,\theta}(\mathbf{A}), \text{ avec } \tilde{\mathbf{H}}_{s,\theta}(\mathbf{A}) = \begin{bmatrix} 0 & 0 \\ \mathbf{N}_{s,\theta}(\mathbf{A}) & 0 \end{bmatrix}.$$

Sous le respect de la contrainte imposée par un découplage total des résidus vis-à-vis des variables d'état,

$$\tilde{\mathbf{w}}^T \tilde{\mathbf{C}}_s = 0,$$

l'expression suivante de la seconde forme statique est obtenue :

$$\tilde{\mathbf{w}}^T \tilde{\mathbf{H}}_{s,0} \begin{bmatrix} \mathbf{U}(k, s-1) \\ \mathbf{Y}(k, s) \end{bmatrix} = -\tilde{\mathbf{w}}^T \tilde{\mathbf{H}}_{s,\theta}(\mathbf{A}) \begin{bmatrix} \mathbf{U}(k, s-1) \\ \mathbf{Y}(k, s) \end{bmatrix} + \tilde{\mathbf{w}}^T \begin{bmatrix} 0 \\ \mathbf{N}_s(\mathbf{A}) \text{diag}_s(\mathbf{T}_\delta) \mathbf{D}_\delta(k, s-1) \end{bmatrix}.$$

La décomposition du vecteur $\tilde{\mathbf{w}}$ en deux vecteurs $\tilde{\mathbf{w}}_L$ et $\tilde{\mathbf{w}}_M$ conduit alors à :

$$\underbrace{\tilde{\mathbf{w}}^T \tilde{\mathbf{H}}_{s,0} \begin{bmatrix} \mathbf{U}(k, s-1) \\ \mathbf{Y}(k, s) \end{bmatrix}}_{\tilde{\mathbf{p}}_{\text{cal}}(k+s)} = \underbrace{\tilde{\mathbf{w}}_M^T \mathbf{N}_{s,\theta} \mathbf{D}_\Delta(k, s-1) \mathbf{U}(k, s-1) + \tilde{\mathbf{w}}_M^T \mathbf{N}_s(\mathbf{A}) \text{diag}_s(\mathbf{T}_\delta) \mathbf{D}_\delta(k, s-1)}_{\tilde{\mathbf{p}}_{\text{éval}}(k+s)}. \quad (\text{III-34})$$

Cette expression de la forme d'évaluation conduit aux mêmes conclusions que dans le cadre de la première forme statique, à savoir qu'elle ne fait pas intervenir l'état inconnu du système et qu'elle est non-linéaire en les incertitudes paramétriques.

III.5.2.3 Relation entre les deux formes statiques

L'objectif est de démontrer que les formes de calcul et d'évaluation sont structurellement équivalentes pour les deux formes statiques dans le cas où la matrice de commande est incertaine, d'une structure modèle de capteurs et d'un découplage total des relations de parité vis-à-vis des variables d'état. La justification est semblable au développement opéré dans la section III.5.1.3, et ne seront donc repris que les points qui en diffèrent.

- Étape 1 : Relation entre les matrices de parité \mathbf{W} et $\tilde{\mathbf{W}}$

Les matrices \mathbf{C}_s et $\tilde{\mathbf{C}}_s$ étant certaines, la relation (III-27) liant la matrice de parité \mathbf{W} à $\tilde{\mathbf{W}}_L$ reste valide :

$$\tilde{\mathbf{W}}_L = \mathbf{T}\mathbf{W}.$$

- Étape 2 : Relation entre les résidus obtenus pour les deux formes statiques

La forme de calcul étant indépendante des incertitudes, la relation suivante, déterminée dans le cadre de la structure classique, est satisfaite dans le cas présent :

$$\tilde{\mathbf{p}}_{cal}(k+s) = \mathbf{T}\mathbf{p}_{cal}(k+s).$$

En ce qui concerne les formes d'évaluation, la première égalité de (III-28) impose que $\tilde{\mathbf{W}}_M \mathbf{N}_{s,0} = -\tilde{\mathbf{W}}_L \mathbf{H}_{s,0}$, ce qui permet de déduire en observant les expressions des matrices $\mathbf{H}_s(\Delta)$ (III-31) et $\mathbf{N}_s(\Delta)$ (III-33) que : $\tilde{\mathbf{W}}_M \mathbf{N}_s(\Delta) = -\tilde{\mathbf{W}}_L \mathbf{H}_s(\Delta)$. En partant de la forme d'évaluation de la seconde forme statique (III-34), la relation suivante est déduite :

$$\tilde{\mathbf{p}}_{éval}(k+s) = -\tilde{\mathbf{W}}_L^T \mathbf{H}_{s,0} \mathbf{D}_\Delta(k, s-1) \mathbf{U}(k, s-1) - \tilde{\mathbf{W}}_L^T \mathbf{H}_s(\Delta) \text{diag}_s(\mathbf{T}_\delta) \mathbf{D}_\delta(k, s-1),$$

ce qui, en la comparant à l'expression de $\mathbf{p}_{éval}(k+s)$ (III-32), permet d'obtenir :

$$\tilde{\mathbf{p}}_{éval}(k+s) = \mathbf{T}\mathbf{p}_{éval}(k+s).$$

La matrice \mathbf{T} étant inversible par définition, pour une valeur de s donnée, les formes de calcul et d'évaluation des relations de parité sont équivalentes pour les deux formes statiques lorsque les incertitudes affectent la matrice de commande (structure de modèle de capteurs) et qu'un découplage total est obtenu.

III.5.3 Récapitulatif

Le premier point à noter dans le cadre d'une matrice de commande incertaine est qu'un découplage total des relations de parité vis-à-vis des incertitudes paramétriques est toujours

possible, le choix de la taille de fenêtre temporelle s fixant le nombre de résidus obtenus. En réalité, le cas d'une matrice de commande incertaine se traite de la même manière qu'un modèle purement certain en ce qui concerne la recherche de la matrice de parité et correspond au cas le plus simple à traiter parmi les modèles incertains envisagés.

Les deux formes statiques sont structurellement parfaitement équivalentes dans le cadre d'un découplage total des formes d'évaluation vis-à-vis des variables d'état et une même taille de fenêtre temporelle, et ce quelle que soit la structure utilisée.

La structure de modèle de capteurs n'a pour seul avantage que de permettre une modélisation "plus physique" du système d'instrumentation dans la mesure où certaines bornes (notamment celles relatives à la précision technologique des capteurs) sont facilement accessibles d'après les données constructeur. Globalement, cette structure conduit à des résultats semblables à la structure classique, tant en terme de choix de la matrice de parité ou de l'horizon temporel que dans l'équivalence structurelle des deux formes statiques. Notons néanmoins que pour la structure classique, les formes d'évaluation sont linéaires en les variables bornées, ce qui n'est pas le cas de la structure de modèle de capteurs. Cette distinction engendrera, dans le quatrième chapitre dédié aux tests de cohérence dans le contexte de l'approche bornante, deux stratégies différentes pour appréhender ces deux types de relations de parité.

Forme statique	Première forme statique		Seconde forme statique	
	Classique	Modèle de capteurs	Classique	Modèle de capteurs
Structure				
Découplage	Total	Total	Total	Total
Forme interne linéaire en les incertitudes	Oui	Non	Oui	Non
Équivalence entre les deux formes statiques	Oui			

Tableau III-1. Matrice de commande incertaine

III.6 MATRICE D'ÉTAT INCERTAINE

Dans cette section, nous explicitons les formes d'évaluation obtenues lorsque les incertitudes n'affectent que la matrice d'état.

III.6.1 Choix de la forme statique

Seule la structure classique sera développée puisque les incertitudes considérées ne proviennent d'aucune chaîne d'instrumentation. À un instant k , la matrice $A(\theta_A)$ est soumise à

r paramètres incertains $\theta_i(k)$ et s'écrit comme suit :

$$A(\theta_A(k)) = A_0 + \sum_{i=1}^r \theta_i(k) A_i.$$

Contrairement au cas précédent où les matrices C_s et \tilde{C}_s intervenant au niveau des termes résiduels étaient certaines, cette fois, il est nécessaire d'utiliser le critère d'optimisation développé dans la section III.4.1. Intéressons-nous tout d'abord à la première forme statique. Dans ce cas, apparaissent des produits de matrices A énoncés à des instants différents :

$$C_s(\theta_A) = \begin{bmatrix} C \\ CA(\theta_A(k)) \\ \vdots \\ CA(\theta_A(k+s-1)) \cdots A(\theta_A(k)) \end{bmatrix},$$

qui ne permettent pas d'exprimer cette matrice sous la forme d'une sommation telle que (III-17). Il est alors très difficile, sauf pour des tailles de fenêtre temporelle très réduites, de déterminer l'expression analytique de la matrice S intervenant dans le critère j_c (III-16) puisqu'il est alors nécessaire de calculer des termes du type :

$$\int_{-1}^1 \cdots \int_{-1}^1 \left(A_0 + \sum_{i=1}^r \theta_i(k+l) A_i \right) \cdots \left(A_0 + \sum_{i_0=1}^r \theta_{i_0}(k) A_{i_0} \right) \left(A_0^T + \sum_{j_0=1}^r \theta_{j_0}(k) A_{j_0}^T \right) \cdots \left(A_0^T + \sum_{j_r=1}^r \theta_{j_r}(k+l') A_{j_r}^T \right) d\theta_1(k) \cdots d\theta_r(k+s)$$

Comme l'expression analytique de S est difficilement accessible, la résolution du problème d'optimisation visant à déterminer la matrice de parité W est compromise. Pour s'affranchir de ce problème, la solution consiste à n'utiliser que la seconde forme statique qui possède l'avantage d'être linéaire en la matrice A , et donc en les incertitudes considérées. C'est cette particularité qui a motivé notre choix de développer en parallèle une deuxième forme statique, sans grand intérêt dans un cadre purement certain puisque les résidus obtenus pour les deux formes statiques sont alors structurellement équivalents (annexe E).

Intéressons-nous donc à la seconde forme statique et plus spécifiquement à la détermination de l'expression analytique de la matrice \tilde{S} (III-19). La structure additive de la matrice d'état se traduit par une forme similaire de la matrice $M_s(\theta_A)$ définie en (III-10) :

$$M_s(\theta_A) = M_{s,0} + \sum_{i=1}^r \sum_{j=0}^{s-1} \theta_i(k+j) M_{s,i,j}$$

avec $M_{s,0} = \begin{bmatrix} A_0 & -I_n & 0 & \cdots & 0 \\ 0 & A_0 & -I_n & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & -I_n \end{bmatrix}$ et $M_{s,i,j} = \begin{bmatrix} 0 & \cdots & 0 & 0 \\ \vdots & A_i & \vdots & \vdots \\ 0 & \cdots & 0 & 0 \end{bmatrix}$ \downarrow $(j+1)$ ème bloc
 \mapsto $(j+1)$ ème bloc

Remarque. Au niveau de la seconde forme statique, les incertitudes affectant la matrice d'état n'apparaissent que jusqu'à l'instant $k+s-1$.

Dans ces conditions, la matrice $\tilde{C}_s(\theta_A)$ (III-11) a elle aussi une structure additive :

$$\tilde{\mathbf{C}}_s(\boldsymbol{\theta}_A) = \tilde{\mathbf{C}}_{s,0} + \sum_{i=1}^r \sum_{j=0}^{s-1} \theta_i(k+j) \tilde{\mathbf{C}}_{s,i,j} \text{ où } \tilde{\mathbf{C}}_{s,0} = \begin{bmatrix} L_s \\ \mathbf{M}_{s,0} \end{bmatrix}, \tilde{\mathbf{C}}_{s,i,j} = \begin{bmatrix} 0 \\ \mathbf{M}_{s,i,j} \end{bmatrix} \text{ et } L_s = \text{diag}_{s+1}(\mathbf{C}).$$

Puisque $\tilde{\mathbf{C}}_s(\boldsymbol{\theta}_A)$ possède une structure semblable à (III-17), l'expression analytique de $\tilde{\mathbf{S}}$ est donnée par une expression similaire à (III-19) :

$$\tilde{\mathbf{S}} = 2^{sr} \left(\tilde{\mathbf{C}}_{s,0} \tilde{\mathbf{C}}_{s,0}^T + \frac{1}{3} \left(\tilde{\mathbf{C}}_{s,1,0} \tilde{\mathbf{C}}_{s,1,0}^T + \dots + \tilde{\mathbf{C}}_{s,r,s-1} \tilde{\mathbf{C}}_{s,r,s-1}^T \right) \right).$$

En ce qui concerne la forme d'évaluation (III-21), puisque la matrice $\tilde{\mathbf{H}}_s$ ne dépend pas d'incertitudes affectant la matrice d'état, elle est par conséquent certaine, ce qui impose la nullité de $\tilde{\mathbf{H}}_{s,\theta}$ et signifie que seuls subsistent les termes résiduels :

$$\tilde{p}_{\text{éval}}(k+s) = \tilde{\mathbf{w}}^T \tilde{\mathbf{C}}_s(\boldsymbol{\theta}_A) \mathbf{X}(k,s) = \tilde{\mathbf{w}}^T \tilde{\mathbf{C}}_{s,0} \mathbf{X}(k,s) + \sum_{i=1}^r \sum_{j=0}^{s-1} \theta_i(k+j) \tilde{\mathbf{w}}^T \tilde{\mathbf{C}}_{s,i,j} \mathbf{X}(k,s).$$

III.6.2 Élimination des termes résiduels

L'objectif de cette section est d'étudier si un découplage total vis-à-vis des variables d'état est toujours possible. Pour trouver une matrice de parité $\tilde{\mathbf{W}}$ orthogonale à $\tilde{\mathbf{C}}_s(\boldsymbol{\theta}_A)$, il est nécessaire que la matrice $\tilde{\mathbf{S}}$ définie précédemment ne soit pas de plein rang. En étudiant de manière plus approfondie cette matrice, nous allons montrer qu'il n'est pas possible de conclure de manière générale sur la possibilité d'un découplage total.

Pour ce faire, réécrivons $\tilde{\mathbf{S}}$ de la manière suivante :

$$\tilde{\mathbf{S}} = 2^{sr} \left(\tilde{\mathbf{C}}_{s,0} \tilde{\mathbf{C}}_{s,0}^T + \frac{1}{3} \left(\tilde{\mathbf{C}}_{s,1} \tilde{\mathbf{C}}_{s,1}^T + \dots + \tilde{\mathbf{C}}_{s,r} \tilde{\mathbf{C}}_{s,r}^T \right) \right) \quad (\text{III-35})$$

$$\text{avec } \tilde{\mathbf{C}}_{s,i} = \sum_{j=0}^{s-1} \tilde{\mathbf{C}}_{s,i,j} = \begin{bmatrix} 0 \\ \mathbf{M}_{s,i} \end{bmatrix} \text{ où } \mathbf{M}_{s,i} = \sum_{j=0}^{s-1} \mathbf{M}_{s,i,j} = \begin{bmatrix} \mathbf{A}_i & 0 & \dots & 0 & 0 \\ 0 & \mathbf{A}_i & \dots & 0 & 0 \\ \vdots & \vdots & \vdots & 0 & \vdots \\ 0 & 0 & \dots & \mathbf{A}_i & 0 \end{bmatrix}.$$

En procédant ainsi, la matrice symétrique précédente peut se mettre sous la forme d'un produit de deux matrices comme suit :

$$\tilde{\mathbf{S}} = 2^{sr} \tilde{\mathbf{R}} \tilde{\mathbf{R}}^T \text{ avec } \tilde{\mathbf{R}} = \begin{bmatrix} \tilde{\mathbf{C}}_{s,0} & \frac{\tilde{\mathbf{C}}_{s,1}}{\sqrt{3}} & \dots & \frac{\tilde{\mathbf{C}}_{s,r}}{\sqrt{3}} \end{bmatrix}$$

Étudier le rang de $\tilde{\mathbf{S}}$ revient alors à étudier celui de $\tilde{\mathbf{R}}^{11}$. La détermination d'une matrice de parité $\tilde{\mathbf{W}}$ strictement orthogonale à la matrice $\tilde{\mathbf{S}}$ est donc équivalent à rechercher une matrice orthogonale à $\tilde{\mathbf{R}}$:

$$\tilde{\mathbf{j}}_c = \tilde{\mathbf{W}} \tilde{\mathbf{S}} \tilde{\mathbf{W}}^T = 0 \Leftrightarrow \tilde{\mathbf{W}} \tilde{\mathbf{R}} = 0,$$

¹¹ Les valeurs singulières de $\tilde{\mathbf{R}}$ correspondent à la racine carrée des valeurs propres de la matrice $\tilde{\mathbf{S}}$ à un coefficient multiplicatif près.

ce qui revient à trouver une matrice \tilde{W} vérifiant :

$$\tilde{W} \begin{bmatrix} \tilde{C}_{s,0} & \tilde{C}_{s,1} & \cdots & \tilde{C}_{s,r} \end{bmatrix} = \tilde{W} \begin{bmatrix} L_s & 0 & \cdots & 0 \\ M_{s,0} & M_{s,1} & \cdots & M_{s,r} \end{bmatrix} = 0.$$

En partitionnant \tilde{W} en deux sous-matrices \tilde{W}_L et \tilde{W}_M , l'égalité précédente se traduit par :

$$\begin{cases} \tilde{W}_L L_s + \tilde{W}_M M_{s,0} = 0 \\ \tilde{W}_M M_{s,i} = 0 \quad i \in \{1, \dots, r\} \end{cases}$$

Dans le cas où la sous-matrice $\begin{bmatrix} M_{s,1} & \cdots & M_{s,r} \end{bmatrix}$ est de plein rang ligne, il ne peut y avoir de découplage total que si la matrice \tilde{W}_M est nulle. Pour peu que le système ne soit pas redondant, c'est-à-dire qu'il n'existe pas plusieurs observations d'une même variable d'état (ou de groupes de variables d'état), alors la matrice L_s est de plein rang ligne et dans ces conditions un découplage total est impossible quelle que soit la taille de fenêtre temporelle envisagée.

Compte tenu de la structure diagonale par blocs de la matrice L_s , il est clair que l'étude de son rang peut se limiter à celle du rang de la matrice d'observation C :

$$\text{rang}(L_s) = (s+1)\text{rang}(C).$$

De même, les matrices $M_{s,i}$, $i \in \{1, \dots, r\}$ ont une structure particulière qui fait que le calcul du rang de la matrice $\begin{bmatrix} A_1 & \cdots & A_r \end{bmatrix}$ suffit pour en déduire celui de $\begin{bmatrix} M_{s,1} & \cdots & M_{s,r} \end{bmatrix}$:

$$\text{rang}\left(\begin{bmatrix} M_{s,1} & \cdots & M_{s,r} \end{bmatrix}\right) = s \text{rang}\left(\begin{bmatrix} A_1 & \cdots & A_r \end{bmatrix}\right).$$

Ainsi, si les matrices $\begin{bmatrix} A_1 & \cdots & A_r \end{bmatrix}$ et C sont de plein rang ligne, alors d'après ce qui précède un découplage total est impossible, et ce, quelle que soit la taille de fenêtre temporelle choisie.

Si $\begin{bmatrix} A_1 & \cdots & A_r \end{bmatrix}$ est de plein rang ligne et que C ne l'est pas, un découplage total est possible en prenant \tilde{W}_M nulle et \tilde{W}_L telle que $\tilde{W}_L L_s = 0$, et ce quelle que soit la taille de fenêtre ($s > 0$). Par ailleurs, si $\begin{bmatrix} A_1 & \cdots & A_r \end{bmatrix}$ et C ne sont pas de plein rang ligne, en prenant les mêmes matrices \tilde{W}_M et \tilde{W}_L que précédemment, un découplage total est obtenu pour toutes les valeurs possibles de la taille de fenêtre (ce ne sont pas les seules solutions). Dans ces deux cas, l'augmentation de la valeur de s accroît le nombre de résidus possibles où les termes résiduels sont nuls.

Le dernier cas à envisager, celui où C est de plein rang ligne contrairement à $\begin{bmatrix} A_1 & \cdots & A_r \end{bmatrix}$ ne permet pas d'obtenir a priori de réponse catégorique quant à l'obtention d'un découplage total, le résultat dépendant du modèle traité. La solution consiste alors à construire la matrice \tilde{S} et à incrémenter la taille de fenêtre s jusqu'à engendrer une hypothétique perte de rang. Ce calcul s'effectue jusqu'à une valeur limite de s à partir de laquelle il est supposé, si \tilde{S} est encore de plein rang ligne, qu'il n'existe aucune solution.

En conclusion, un découplage total n'est pas toujours possible dans le cas d'une matrice d'état incertaine où l'emploi de la seconde forme statique est nécessaire. Les différentes configurations sont récapitulées dans le tableau ci-dessous.

$[A_1 \dots A_r]$ \ C	Plein rang ligne	Lignes dépendantes
Plein rang ligne	Impossible	Découplage total
Lignes dépendantes	Dépend du modèle	Découplage total

Tableau III-2. Conditions pour qu'un découplage total soit possible

Un découplage total est intéressant puisqu'il permet d'avoir une forme d'évaluation réduite à 0, ce qui élimine tous les problèmes liés aux calculs du domaine de bon fonctionnement. Si les conditions requises sont vérifiées, il faut alors rechercher une taille de fenêtre permettant d'annuler les termes résiduels et d'obtenir un nombre adéquat de relations de parité. En revanche, une désensibilisation complète des résidus vis-à-vis des incertitudes est alors effectuée. Dans ces conditions, le diagnostic n'est réalisé que sur la partie certaine du système et l'approche par modélisation incertaine n'a alors que peu d'intérêt puisque les paramètres incertains ont tous été rejetés.

Si un découplage total est impossible, deux solutions sont envisageables. La première consiste à ne se contenter que d'un découplage partiel et ne s'applique qu'au cas très particulier où la matrice d'observation est de plein rang colonne. Il suffit alors en quelque sorte d'estimer l'état en utilisant le pseudo-inverse de C (Golub et al, 1983) (ou son inverse si elle est carrée) :

$$\hat{X}_\theta(k,s) = \begin{bmatrix} (C^T C)^{-1} C^T & 0 & 0 \\ 0 & \ddots & 0 \\ 0 & 0 & (C^T C)^{-1} C^T \end{bmatrix} Y(k,s) \text{ ou } \hat{X}_\theta(k,s) = \begin{bmatrix} C^{-1} & 0 & 0 \\ 0 & \ddots & 0 \\ 0 & 0 & C^{-1} \end{bmatrix} Y(k,s),$$

dans le cas où C est inversible.

Le résultat de cette estimation est bien sûr essentiellement lié à la qualité du modèle (plus spécifiquement à la représentation de la chaîne d'instrumentation), mais permet de résoudre facilement le problème. La seconde possibilité n'est pas aussi restrictive et consiste en la construction d'une matrice de parité incertaine. Cette approche sera développée dans la section III.10.

III.6.3 Récapitulatif

Dans le cas d'une matrice d'état incertaine, seule la seconde forme statique est utilisable puisqu'elle permet d'explicitier l'expression analytique de la matrice \tilde{S} nécessaire à la résolution du problème d'optimisation visant à éliminer les variables d'état de la forme d'évaluation. Seule la structure classique a été développée dans la mesure où la matrice d'état n'est pas liée aux imperfections du système d'instrumentation.

En fonction des contraintes explicitées dans le tableau III-2, des découplages totaux ou seulement partiels sont possibles. Dans le premier cas, la forme d'évaluation est alors

complètement nulle et les résidus ne dépendent plus du tout des incertitudes. De plus, la forme d'évaluation est linéaire en les incertitudes, même lorsque l'état doit être estimé.

Forme statique	Seconde forme statique
Structure	Classique
Découplage	Total ou partiel
Forme interne linéaire en les incertitudes	Oui

Tableau III-3. Matrice d'état incertaine

III.7 MATRICE D'OBSERVATION INCERTAINE

Dans cette section, nous explicitons les formes d'évaluation obtenues lorsque les incertitudes n'affectent que la matrice d'observation.

III.7.1 Structure classique

III.7.1.1 Première forme statique

- Forme d'évaluation

À un instant k , la matrice $C(\theta_C)$ dépend de r paramètres incertains $\theta_i(k)$ et s'écrit de la manière suivante :

$$C(\theta_C(k)) = C_0 + \sum_{i=1}^r \theta_i(k) C_i.$$

Dans ces conditions, la matrice $C_s(\theta_C)$ (III-8) a elle aussi une structure additive :

$$C_s(\theta_C) = C_{s,0} + \sum_{i=1}^r \sum_{j=0}^s \theta_i(k+j) C_{s,i,j},$$

$$\text{avec } C_{s,0} = \begin{bmatrix} C_0 \\ \vdots \\ C_0 A^s \end{bmatrix} \text{ et } C_{s,i,j} = \begin{bmatrix} 0 \\ C_i A^j \\ 0 \end{bmatrix} \downarrow_{(j+1)\text{ième bloc}}.$$

La matrice $H_s(\theta_C)$ peut être mise elle aussi sous forme additive :

$$H_s(\theta_C) = H_{s,0} + H_{s,\theta}(\theta_C) \text{ avec } H_{s,\theta}(\theta_C) = \sum_{i=1}^r \sum_{j=1}^s \theta_i(k+j) H_{s,i,j},$$

$$\text{où } \mathbf{H}_{s,0} = \begin{bmatrix} 0 & \cdots & 0 \\ \mathbf{C}_0 \mathbf{B} & \cdots & 0 \\ \vdots & \ddots & \vdots \\ \mathbf{C}_0 \mathbf{A}^{s-1} \mathbf{B} & \cdots & \mathbf{C}_0 \mathbf{B} \end{bmatrix} \text{ et } \mathbf{H}_{s,i,j} = \begin{bmatrix} 0 & \cdots & \cdots & 0 \\ \mathbf{C}_i \mathbf{A}^{j-1} \mathbf{B} & \cdots & \mathbf{C}_i \mathbf{B} & \vdots \\ 0 & \cdots & \cdots & 0 \end{bmatrix} \downarrow_{(j+1)\text{ième bloc}}.$$

Remarque. L'indice j ne commence qu'à 1 puisqu'aucune incertitude exprimée à l'instant k de début de fenêtre n'apparaît dans la matrice $\mathbf{H}_s(\boldsymbol{\theta}_C)$.

Par conséquent, la forme d'évaluation (III-20) devient :

$$\begin{aligned} p_{\text{éval}}(k+s) &= \mathbf{w}^T \{ \mathbf{C}_s(\boldsymbol{\theta}_C) \mathbf{x}(k) + \mathbf{H}_{s,\boldsymbol{\theta}}(\boldsymbol{\theta}_C) \mathbf{U}(k, s-1) \} \\ &= \mathbf{w}^T \mathbf{C}_{s,0} \mathbf{x}(k) + \sum_{i=1}^r \sum_{j=0}^s \theta_i(k+j) \mathbf{w}^T \mathbf{C}_{s,i,j} \mathbf{x}(k) + \sum_{i=1}^r \sum_{j=1}^s \theta_i(k+j) \mathbf{w}^T \mathbf{H}_{s,i,j} \mathbf{U}(k, s-1). \end{aligned} \quad (\text{III-36})$$

- *Minimisation des termes résiduels*

Puisque la matrice $\mathbf{C}_s(\boldsymbol{\theta}_C)$ possède une structure similaire à (III-17), l'expression analytique de la matrice \mathbf{S} intervenant dans le critère j_c (III-16) est donnée par (III-18) :

$$\mathbf{S} = 2^{(s+1)r} \left(\mathbf{C}_{s,0} \mathbf{C}_{s,0}^T + \frac{1}{3} (\mathbf{C}_{s,1,0} \mathbf{C}_{s,1,0}^T + \cdots + \mathbf{C}_{s,r,s} \mathbf{C}_{s,r,s}^T) \right). \quad (\text{III-37})$$

La matrice \mathbf{S} peut se mettre sous la forme d'un produit de deux matrices comme suit :

$$\mathbf{S} = 2^{(s+1)r} \mathbf{R} \mathbf{R}^T \text{ avec } \mathbf{R} = \begin{bmatrix} \mathbf{C}_{s,0} & \frac{\mathbf{C}_{s,1,0}}{\sqrt{3}} & \cdots & \frac{\mathbf{C}_{s,r,s}}{\sqrt{3}} \end{bmatrix}. \quad (\text{III-38})$$

De plus, de manière à alléger les écritures, nous adopterons la notation suivante qui nous servira dans la section III.3.1.3 :

$$\mathbf{C}_{s,i} = \sum_{j=0}^s \mathbf{C}_{s,i,j} = \sum_{j=0}^s \begin{bmatrix} 0 \\ \mathbf{C}_i \mathbf{A}^j \\ 0 \end{bmatrix} = \begin{bmatrix} \mathbf{C}_i \\ \vdots \\ \mathbf{C}_i \mathbf{A}^s \end{bmatrix}, \quad i \in \{1, \dots, r\}. \quad (\text{III-39})$$

III.7.1.2 Seconde forme statique

- *Forme d'évaluation*

Comme les incertitudes portent sur la matrice d'observation, la matrice $\mathbf{L}_s(\boldsymbol{\theta}_C)$ (III-7) se décompose comme suit :

$$\begin{aligned} \mathbf{L}_s(\boldsymbol{\theta}_C) &= \mathbf{L}_{s,0} + \sum_{i=1}^r \sum_{j=0}^s \theta_i(k+j) \mathbf{L}_{s,i,j}, \\ \text{avec } \mathbf{L}_{s,0} &= \begin{bmatrix} \mathbf{C}_0 & 0 & 0 \\ 0 & \ddots & 0 \\ 0 & 0 & \mathbf{C}_0 \end{bmatrix} \text{ et } \mathbf{L}_{s,i,j} = \begin{bmatrix} 0 & \cdots & 0 \\ \vdots & \mathbf{C}_i & \vdots \\ 0 & \cdots & 0 \end{bmatrix} \downarrow_{(j+1)\text{ième bloc}} \\ & \quad \uparrow_{(j+1)\text{ième bloc}} \end{aligned}$$

Dans ces conditions, la matrice $\tilde{\mathbf{C}}_s(\boldsymbol{\theta}_C)$ (III-11) a elle aussi une structure additive :

$$\tilde{\mathbf{C}}_s(\boldsymbol{\theta}_C) = \tilde{\mathbf{C}}_{s,0} + \sum_{i=1}^r \sum_{j=0}^s \theta_i(k+j) \tilde{\mathbf{C}}_{s,i,j}, \text{ avec } \tilde{\mathbf{C}}_{s,0} = \begin{bmatrix} L_{s,0} \\ \mathbf{M}_s \end{bmatrix} \text{ et } \tilde{\mathbf{C}}_{s,i,j} = \begin{bmatrix} L_{s,i,j} \\ 0 \end{bmatrix}.$$

La matrice $\tilde{\mathbf{H}}_s$ est quant à elle certaine puisqu'elle ne dépend pas des incertitudes pouvant affecter la matrice d'observation, ce qui signifie donc que $\tilde{\mathbf{H}}_{s,\theta}$ est nulle et que seuls subsistent les termes résiduels.

En utilisant l'expression générale de la forme d'évaluation (III-21) et ce qui précède, l'expression de $\tilde{p}_{\text{éval}}(k+s)$ est obtenue :

$$\begin{aligned} \tilde{p}_{\text{éval}}(k+s) &= \tilde{\mathbf{w}}^T \tilde{\mathbf{C}}_s(\boldsymbol{\theta}_C) \mathbf{X}(k,s) \\ &= \tilde{\mathbf{w}}^T \tilde{\mathbf{C}}_{s,0} \mathbf{X}(k,s) + \sum_{i=1}^r \sum_{j=0}^s \theta_i(k+j) \tilde{\mathbf{w}}^T \tilde{\mathbf{C}}_{s,i,j} \mathbf{X}(k,s) \end{aligned} \quad (\text{III-40})$$

- Minimisation des termes résiduels

Puisque que la matrice $\tilde{\mathbf{C}}_s(\boldsymbol{\theta}_C)$ possède une structure similaire à (III-17), l'expression analytique de la matrice $\tilde{\mathbf{S}}$ est donnée par (III-19) :

$$\tilde{\mathbf{S}} = 2^{(s+1)r} \left(\tilde{\mathbf{C}}_{s,0} \tilde{\mathbf{C}}_{s,0}^T + \frac{1}{3} \left(\tilde{\mathbf{C}}_{s,1,0} \tilde{\mathbf{C}}_{s,1,0}^T + \dots + \tilde{\mathbf{C}}_{s,r,s} \tilde{\mathbf{C}}_{s,r,s}^T \right) \right).$$

Notons qu'il est possible, par analogie avec (III-39), de réécrire cette matrice sous la forme suivante :

$$\tilde{\mathbf{S}} = 2^{(s+1)r} \left(\tilde{\mathbf{C}}_{s,0} \tilde{\mathbf{C}}_{s,0}^T + \frac{1}{3} \left(\tilde{\mathbf{C}}_{s,1} \tilde{\mathbf{C}}_{s,1}^T + \dots + \tilde{\mathbf{C}}_{s,r} \tilde{\mathbf{C}}_{s,r}^T \right) \right), \quad (\text{III-41})$$

$$\text{avec } \tilde{\mathbf{C}}_{s,i} = \sum_{j=0}^s \tilde{\mathbf{C}}_{s,i,j} = \begin{bmatrix} L_{s,i} \\ 0 \end{bmatrix} \text{ où } L_{s,i} = \sum_{j=0}^s L_{s,i,j} = \begin{bmatrix} \mathbf{C}_i & 0 & 0 \\ 0 & \ddots & 0 \\ 0 & 0 & \mathbf{C}_i \end{bmatrix}, i \in \{1, \dots, r\}.$$

La matrice $\tilde{\mathbf{S}}$ peut donc se mettre sous la forme d'un produit de deux matrices comme suit :

$$\tilde{\mathbf{S}} = 2^{(s+1)r} \tilde{\mathbf{R}} \tilde{\mathbf{R}}^T \text{ avec } \tilde{\mathbf{R}} = \begin{bmatrix} \tilde{\mathbf{C}}_{s,0} & \frac{\tilde{\mathbf{C}}_{s,1}}{\sqrt{3}} & \dots & \frac{\tilde{\mathbf{C}}_{s,r}}{\sqrt{3}} \end{bmatrix}. \quad (\text{III-42})$$

III.7.1.3 Relation entre les deux formes statiques

L'objectif est d'étudier s'il y a équivalence structurelle entre les relations de parité obtenues pour les deux formes statiques lorsque la matrice d'observation est incertaine (structure classique).

- Étape 1 : Relation entre les matrices de parité \mathbf{W} et $\tilde{\mathbf{W}}$

① Soit $\tilde{\mathbf{W}}$ une matrice permettant d'annuler le critère \tilde{j}_c (III-19) pour une taille de fenêtre s . Étudier le rang de $\tilde{\mathbf{S}}$ revient alors exactement à étudier celui de $\tilde{\mathbf{R}}$; ce qui conduit donc à rechercher une matrice de parité $\tilde{\mathbf{W}}$ strictement orthogonale à $\tilde{\mathbf{R}}$ (III-42) :

$$\tilde{j}_c = \tilde{\mathbf{W}}\tilde{\mathbf{S}}\tilde{\mathbf{W}}^T = 0 \Leftrightarrow \tilde{\mathbf{W}}\tilde{\mathbf{R}} = 0.$$

La matrice de parité $\tilde{\mathbf{W}}$ est supposée de plein rang ligne de manière à ne conserver que des résidus linéairement indépendants et son rang vaut alors : $s_{\tilde{\mathbf{W}}} = s_m + sn - \text{rang}(\tilde{\mathbf{R}})$. Puisque $\tilde{\mathbf{S}}$ et $\tilde{\mathbf{R}}$ possèdent le même rang, il s'en déduit que : $s_{\tilde{\mathbf{W}}} = s_m + sn - \text{rang}(\tilde{\mathbf{S}})$.

En partitionnant $\tilde{\mathbf{W}}$ en deux sous-matrices $\tilde{\mathbf{W}}_L$ et $\tilde{\mathbf{W}}_M$, son existence se traduit par les $r+1$ contraintes suivantes :

$$\begin{cases} \tilde{\mathbf{W}}_L \mathbf{L}_{s,0} + \tilde{\mathbf{W}}_M \mathbf{M}_s = 0 \\ \tilde{\mathbf{W}}_L \mathbf{L}_{s,i} = 0 \quad i \in \{1, \dots, r\} \end{cases} \quad (\text{III-43})$$

où les matrices $\mathbf{L}_{s,i}$ ont été définies en (III-41).

En multipliant à droite le système précédent par la matrice certaine \mathbf{A}_s (III-6), sachant que le produit $\mathbf{M}_s \mathbf{A}_s$ est nul (III-15) et que $\mathbf{L}_{s,i} \mathbf{A}_s = \mathbf{C}_{s,i}$, $i \in \{0, 1, \dots, r\}$ (III-12), les résultats suivants sont obtenus :

$$\tilde{\mathbf{W}}_L \mathbf{C}_{s,i} = 0, \quad i \in \{0, 1, \dots, r\}. \quad (\text{III-44})$$

où les matrices $\mathbf{C}_{s,i}$ ont été définies en (III-39).

② Soit \mathbf{W} une matrice permettant d'annuler le critère j_c pour une taille de fenêtre s . Le problème revient à rechercher une matrice de parité \mathbf{W} strictement orthogonale à \mathbf{R} (III-38) :

$$j_c = \mathbf{W}\mathbf{S}\mathbf{W}^T = 0 \Leftrightarrow \mathbf{W}\mathbf{R} = 0,$$

ceci se traduit par :

$$\begin{cases} \mathbf{W}\mathbf{C}_{s,0} = 0 \\ \mathbf{W}\mathbf{C}_{s,i,j} = 0 \quad i \in \{1, \dots, r\}, j \in \{0, \dots, s\} \end{cases} \quad (\text{III-45})$$

Notons alors que les contraintes imposées par un découplage total pour les deux formes statiques (III-44), (III-45), sont différentes. Si un découplage total est possible dans le cas de la seconde forme statique, rien ne peut être déduit sur ce qui se passe pour la première forme statique. En effet, vérifier : $\tilde{\mathbf{W}}_L \mathbf{C}_{s,i} = \tilde{\mathbf{W}}_L [\mathbf{C}_{s,i,0} + \dots + \mathbf{C}_{s,i,s}] = 0$ n'impose pas nécessairement que $\tilde{\mathbf{W}}_L \mathbf{C}_{s,i,j} = 0$, $j \in \{0, 1, \dots, s\}$. Il n'y a donc pas d'implication ① \Rightarrow ②. Qu'en est-il du contraire ?

Les égalités (III-45) peuvent se mettre sous la forme suivante :

$$\begin{cases} \mathbf{W}\mathbf{C}_{s,0} = 0 \\ \mathbf{W}\mathbf{C}_{s,i} = \mathbf{W}\mathbf{L}_{s,i} \underbrace{\begin{bmatrix} \mathbf{I}_n & 0 & 0 \\ 0 & \ddots & 0 \\ 0 & 0 & \mathbf{A}^s \end{bmatrix}}_{\mathbf{A}_s^*} = 0, \quad i \in \{1, \dots, r\}. \end{cases}$$

En supposant que la matrice d'état (et donc toutes ses puissances \mathbf{A}^s , pour s fini) est de plein rang et donc que la matrice \mathbf{A}_s^* est carrée et inversible, l'équivalence suivante est satisfaite :

$$\mathbf{W}\mathbf{C}_{s,i} = 0 \Leftrightarrow \mathbf{W}\mathbf{L}_{s,i} \mathbf{A}_s^* \mathbf{A}_s^{*-1} = 0 \Leftrightarrow \mathbf{W}\mathbf{L}_{s,i} = 0, \quad i \in \{1, \dots, r\}.$$

Par conséquent, vérifier $WC_{s,i,j}=0$, $i \in \{1, \dots, r\}$, $j \in \{0, \dots, s\}$ impose nécessairement de respecter $WL_{s,i}=0$, $i \in \{1, \dots, r\}$, ce qui permet de satisfaire les r dernières contraintes (III-43) imposées par un découplage total dans le cadre de la seconde forme statique. De plus, si $W \textcircled{2}$ existe, alors, puisque M_s est de plein rang ligne, il est possible de poser :

$$W' = -WL_s M_s^T (M_s M_s^T)^{-1},$$

auquel cas la première contrainte $WL_{s,0} + W'M_s = 0$ de (III-43) est elle aussi satisfaite.

Dans ces conditions, si les contraintes (III-45) imposées par un découplage dans le cadre de la première forme statique $\textcircled{2}$ sont vérifiées, alors il existe une matrice non-nulle $[W \ W']$ permettant de satisfaire les contraintes (III-43) et assurant un découplage total dans le cas de la seconde forme statique. En conclusion, si la matrice d'état A est de plein rang et qu'un découplage total est possible dans le cas de la première forme statique (conditions $\textcircled{2}$ respectées), alors il en est de même pour une même taille de fenêtre temporelle pour la seconde forme statique (conditions $\textcircled{1}$ respectées).

Montrons maintenant que les relations de parité obtenues dans le cas d'un découplage total pour la première forme statique ont leurs équivalents structurels dans le cadre de la seconde forme statique. Comme l'a été démontré précédemment, un découplage total pour la première forme statique conduit à l'existence d'une matrice non-nulle $[W \ W']$ orthogonale à \tilde{R} . A partir de la décomposition en valeurs singulières de \tilde{R} :

$$\tilde{R} = U \Sigma V^T,$$

est déduit que :

$$[W \ W'] = K \begin{bmatrix} 0 & I_{sn+s_m - \text{rang}(\tilde{R})} \end{bmatrix} U^T, \text{ où } K \text{ est une matrice inversible quelconque.}$$

D'autre part, un découplage est alors possible pour la seconde forme statique, ce qui impose l'existence d'une matrice \tilde{W} vérifiant :

$$\tilde{W} = \tilde{K} \begin{bmatrix} 0 & I_{sn+s_m - \text{rang}(\tilde{R})} \end{bmatrix} U^T, \text{ où } \tilde{K} \text{ est inversible,}$$

ce qui permet finalement d'aboutir à :

$$\tilde{W} = \begin{bmatrix} \tilde{W}_L & \tilde{W}_M \end{bmatrix} = T [W \ W'], \text{ où } T \text{ est une matrice inversible.}$$

- *Étape 2 : Relation entre les résidus obtenus pour les deux formes statiques*

À partir de la relation (III-14) se déduisent les égalités suivantes :

$$\begin{aligned} L_{s,0} B_s &= H_{s,0} \\ L_{s,i,j} B_s &= H_{s,i,j}, \quad i \in \{1, \dots, r\}, j \in \{0, \dots, s\} \end{aligned}$$

À l'aide de ce qui précède, des relations $M_s B_s = N_s$ issue de (III-13) et $\tilde{W}_L L_{s,0} + \tilde{W}_M M_s = 0$, se déduit que :

$$\tilde{W}_M N_s = -\tilde{W}_L H_{s,0}.$$

La forme de calcul associée à la seconde forme statique (III-21) devient :

$$\tilde{\mathbf{p}}_{cal}(k+s) = \tilde{\mathbf{W}}\tilde{\mathbf{H}}_s \begin{bmatrix} U(k,s-1) \\ \mathbf{Y}(k,s) \end{bmatrix} = \begin{bmatrix} \tilde{\mathbf{W}}_L & \tilde{\mathbf{W}}_M \end{bmatrix} \begin{bmatrix} \mathbf{Y}(k,s) \\ \mathbf{N}_s U(k,s-1) \end{bmatrix} = \tilde{\mathbf{W}}_L \{ \mathbf{Y}(k,s) - \mathbf{H}_{s,0} U(k,s-1) \}.$$

ce qui se traduit, d'après (III-20) par :

$$\tilde{\mathbf{p}}_{cal}(k+s) = \mathbf{T}\mathbf{p}_{cal}(k+s).$$

Par conséquent, les formes de calcul sont structurellement équivalentes si ② est vérifié. Si seulement ① est satisfait, il n'est plus possible de conclure puisque l'existence de \mathbf{W} est remise en cause.

Vérifions que la même propriété est satisfaite en ce qui concerne les formes d'évaluation. Compte tenu de l'expression obtenue dans le cas de la seconde forme statique, il est clair que pour un découplage total imposé par ①, celle-ci se résume à 0. Considérons maintenant la forme d'évaluation liée à la première forme statique. Dans le cas d'un découplage total imposé par ②, il s'ensuit la simplification suivante :

$$\mathbf{p}_{éval}(k+s) = \mathbf{W}\mathbf{H}_{s,\theta}(\theta_C)U(k,s-1).$$

Cette expression est non nulle a priori avec ce choix de \mathbf{W} . Cependant, par construction des matrices $\mathbf{L}_{s,i}$ en fonction des matrices $\mathbf{L}_{s,i,j}$, il est clair que les contraintes $\mathbf{W}\mathbf{L}_{s,i}=0, i \in \{1, \dots, r\}$ entraînent nécessairement $\mathbf{W}\mathbf{L}_{s,i,j}=0$, soit :

$$\mathbf{W}\mathbf{L}_{s,i,j}\mathbf{B}_s = \mathbf{W}\mathbf{H}_{s,i,j} = 0, i \in \{1, \dots, r\}, j \in \{0, \dots, s\}.$$

Dans ces conditions, la forme d'évaluation $\mathbf{p}_{éval}(k+s)$ est bien nulle et les deux formes d'évaluation sont équivalentes. En conclusion, la matrice \mathbf{T} étant inversible, les résidus obtenus pour la première forme statique et issus d'un découplage total auront des équivalents structurels dans le cas de la seconde forme statique. En revanche, l'inverse n'est pas nécessairement vrai. Pour une taille de fenêtre donnée s , un nombre supérieur ou égal de résidus sera donc obtenu par découplage total pour la seconde forme statique. Mathématiquement, cela se traduit par un nombre plus important de valeurs propres nulles pour $\tilde{\mathbf{S}}$ que pour \mathbf{S} .

III.7.1.4 Élimination des termes résiduels

- Première forme statique

L'objectif de cette section est d'étudier si un découplage total vis-à-vis des variables d'état est toujours possible dans le cadre de la première forme statique. Pour trouver une matrice de parité \mathbf{W} orthogonale à $\mathbf{C}_s(\theta_C)$, il est nécessaire que la matrice \mathbf{S} ne soit pas de plein rang. En étudiant de manière plus approfondie cette matrice, nous allons montrer qu'il n'est pas possible de conclure de manière générale sur la possibilité d'un découplage total. Les matrices \mathbf{S} et \mathbf{R} (III-38) ayant le même rang, regardons l'influence de la taille de fenêtre s sur le rang de \mathbf{R} qui est une matrice plus simple à manipuler. En regroupant par blocs tous les termes dépendant des mêmes paramètres incertains, l'expression suivante est obtenue :

$$\mathbf{R} = \left[\begin{array}{c} \left[\begin{array}{c} C_0 \\ \vdots \\ C_0 A^s \\ \hline C_{s,0} \end{array} \right] \\ \frac{1}{\sqrt{3}} \left[\begin{array}{ccc} C_1 & 0 & 0 \\ 0 & \ddots & 0 \\ 0 & 0 & C_1 A^s \\ \hline \mathbf{R}_1 \end{array} \right] \\ \dots \\ \frac{1}{\sqrt{3}} \left[\begin{array}{ccc} C_r & 0 & 0 \\ 0 & \ddots & 0 \\ 0 & 0 & C_r A^s \\ \hline \mathbf{R}_r \end{array} \right] \end{array} \right].$$

Il n'est pas possible d'assurer de manière générale l'existence d'une taille de fenêtre permettant d'obtenir une matrice \mathbf{R} qui ne soit pas de plein rang ligne. En effet, pour que ce soit le cas, il faut tout d'abord que chacun des $r+1$ blocs précédents ne soit pas de plein rang ligne. Dans le cas où la matrice d'état \mathbf{A} est de plein rang, cette condition nécessite forcément que toutes les matrices C_i , $i \in \{1, \dots, r\}$ ne soient pas de plein rang de manière à ce qu'il en soit de même pour les matrices \mathbf{R}_i . Si cette condition est vérifiée, il faut alors une taille de fenêtre suffisante pour que $C_{s,0}$ ne soit pas de plein rang ligne (ce qui est toujours possible). Enfin, la dernière contrainte impose l'existence d'un vecteur non-nul \mathbf{w}^T commun tel que les produits de celui-ci par chacun des blocs précédents ($C_{s,0}$, $\mathbf{R}_1, \dots, \mathbf{R}_r$) donnent zéro. Il faut donc que l'intersection des noyaux de ces différents blocs soit non-vide. L'ensemble de ces contraintes conduit à des conditions très restrictives ; en général, un découplage total ne sera pas, par conséquent, possible.

Remarque. Le modèle étant stable, la matrice A^s pour s suffisamment grand tend donc vers la matrice nulle. Une alternative peut consister à calculer les valeurs propres de la matrice \mathbf{S} , sachant qu'en incrémentant la taille de fenêtre temporelle, les plus petites d'entre elles vont tendre asymptotiquement vers 0. Si un seuil est fixé en dessous duquel les valeurs propres sont considérées comme nulles, il est possible de trouver une valeur de s (certes élevée) au dessus de laquelle \mathbf{S} n'est plus de plein rang ligne. La notion de seuil permet d'éviter une taille de fenêtre infinie. Un compromis est alors à envisager entre la précision du découplage et la valeur de la taille de fenêtre choisie qui peut être très importante et accroît d'autant le nombre de paramètres incertains à traiter. Ce n'est pas l'approche retenue dans ce document, contrairement à Ploix (1998) qui corrige ces approximations en ajoutant de nouveaux paramètres incertains.

- Seconde forme statique

En ce qui concerne la seconde forme statique, un découplage total n'est pas forcément toujours réalisable. En effet, étudions le rang de la matrice $\tilde{\mathbf{R}}$ (aux coefficients $1/\sqrt{3}$ près qui de toute façon n'en modifient pas le rang) identique à celui de $\tilde{\mathbf{S}}$:

$$\left[\tilde{C}_{s,0} \quad \tilde{C}_{s,1} \quad \dots \quad \tilde{C}_{s,r} \right] = \begin{bmatrix} L_{s,0} & L_{s,1} & \dots & L_{s,r} \\ \mathbf{M}_s & 0 & \dots & 0 \end{bmatrix}.$$

La détermination d'une matrice de parité strictement orthogonale à la matrice $\tilde{\mathbf{S}}$ se traduit par les contraintes (III-43) pour une taille de fenêtre donnée s . Si la matrice $\tilde{\mathbf{W}}_L$ existe, nous avons vu qu'il en était alors nécessairement de même pour $\tilde{\mathbf{W}}_M$. Par conséquent, les contraintes $\textcircled{1}$, ne sont subordonnées qu'à l'existence de $\tilde{\mathbf{W}}_L$ satisfaisant les r égalités $\tilde{\mathbf{W}}_L L_{s,i} = 0$, $i \in \{1, \dots, r\}$. Les matrices $L_{s,i}$ étant diagonales par blocs et constituées des matrices C_i , les égalités précédentes imposent que toutes les matrices C_i , $i \in \{1, \dots, r\}$, ne soient pas de

plein rang ligne et que l'intersection de leurs noyaux respectifs ne soit pas vide, ou de manière équivalente que la matrice $[C_1 \ \dots \ C_r]$ (et par conséquent $[L_{s,1} \ \dots \ L_{s,r}]$) ne soit pas de plein rang ligne.

Le fait que $[C_1 \ \dots \ C_r]$ ne soit pas de plein rang ligne est seulement une condition nécessaire pour pouvoir obtenir un découplage total ; montrons qu'elle est aussi suffisante. Pour ce faire, partons de l'inégalité suivante :

$$\text{rang} \left(\begin{bmatrix} L_{s,0} & L_{s,1} & \dots & L_{s,r} \\ M_s & 0 & \dots & 0 \end{bmatrix} \right) \leq \text{rang} \left(\begin{bmatrix} L_{s,0} \\ M_s \end{bmatrix} \right) + \text{rang}([L_{s,1} \ \dots \ L_{s,r}]).$$

$\text{rang} \left(\begin{bmatrix} L_{s,0} \\ M_s \end{bmatrix} \right)$ est majoré par $\min(s_m + s_n, s_n)$ qui définissent les dimensions de cette matrice. En règle générale, ce minimum est donné par s_n , sinon la taille de fenêtre s doit satisfaire :

$$s < \text{int} \left(\frac{n-m}{m} \right), \text{ où la fonction } \text{int} \text{ désigne la partie entière de son argument,}$$

ce qui est une condition trop restrictive pour considérer que le minimum précédent puisse être donné de manière générale par $s_m + s_n$ ¹². D'autre part, en utilisant la condition nécessaire permettant d'aboutir à un découplage total, $[C_1 \ \dots \ C_r]$ n'est pas de plein rang ligne, le rang de cette matrice est donc au plus égal à $m-1$; ce qui signifie que $\text{rang}([L_{s,1} \ \dots \ L_{s,r}])$ est donc au plus égal à $(s+1)(m-1)$.

Il s'en déduit donc que :

$$\text{rang} \left(\begin{bmatrix} L_{s,0} & L_{s,1} & \dots & L_{s,r} \\ M_s & 0 & \dots & 0 \end{bmatrix} \right) \leq s_n + (s+1)(m-1).$$

À coût sûr, la matrice précédente n'est pas de plein rang ligne si :

$$(s+1)n + (s+1)(m-1) < s_m + s_n,$$

où le terme de droite de cette inégalité correspond au nombre de lignes de la matrice précédente. On obtient finalement comme condition suffisante :

$$s > n-1.$$

En conclusion, si la matrice $[C_1 \ \dots \ C_r]$ n'est pas de plein rang ligne, il existe toujours une taille de fenêtre minimale égale à n à partir de laquelle on est sûr qu'un découplage total est possible. Comme seule une majoration du rang de la matrice $\begin{bmatrix} L_{s,0} & L_{s,1} & \dots & L_{s,r} \\ M_s & 0 & \dots & 0 \end{bmatrix}$ a été considérée, il est fortement probable que pour une valeur plus petite de la taille de fenêtre temporelle, un découplage total soit obtenu.

¹² Dans une otique de diagnostic de système, même si toutes les variables d'état ne sont pas mesurées, il faut un minimum d'information pour pouvoir détecter et surtout localiser un défaut, de ce fait m est en général proche de la valeur de n , voire supérieur.

Notons qu'en cas de découplage total vis-à-vis des variables d'état, cette forme d'évaluation est nulle, ce qui élimine tous les problèmes liés aux calculs de domaine de bon fonctionnement. En revanche, une désensibilisation complète des résidus vis-à-vis des incertitudes est alors effectuée. Dans ces conditions, le diagnostic n'est réalisé que sur la partie certaine du système et l'approche par modélisation incertaine n'a alors que peu d'intérêt puisque les paramètres incertains ont tous été rejetés.

- *Découplage partiel*

Dans le cas où seul un découplage partiel est possible, c'est-à-dire qu'il reste des termes résiduels, l'état inconnu du système doit être estimé en vue de pouvoir utiliser les formes d'évaluation. Ceci consiste donc à estimer le vecteur $x(k)$ de dimension n dans le cas de la première forme statique ou le vecteur $X(k,s)$ de dimension s_n dans le cas de la seconde ; ce qui laisse entrevoir un léger avantage de la première forme statique quant à la dimension du problème à traiter.

La solution proposée est à la fois restrictive et simple puisqu'elle consiste, lorsque la matrice d'observation est de plein rang colonne, à utiliser le pseudo-inverse de cette matrice, sans aucune correction comme cela est classiquement réalisé dans le cas des observateurs. L'estimation de l'état est effectuée directement à partir des sorties mesurées. Pour la première forme statique, l'estimée \hat{x}_θ est donnée par :

$$\hat{x}_\theta(k) = C^+(\theta_C(k))y(k) \text{ avec } C^+(\theta_C(k)) = \left(C^T(\theta_C(k))C(\theta_C(k)) \right)^{-1} C^T(\theta_C(k)).$$

Dans le cas de la seconde forme statique, il est nécessaire d'estimer l'état sur tout l'horizon :

$$\hat{X}_\theta(k,s) = \begin{bmatrix} C^+(\theta_C(k)) & 0 & 0 \\ 0 & \ddots & 0 \\ 0 & 0 & C^+(\theta_C(k+s)) \end{bmatrix} Y(k,s).$$

La contrainte sur le rang de $C(\theta_C)$ est restrictive. La matrice d'observation étant incertaine, cette condition de rang doit être vérifiée quelles que soient les valeurs des paramètres incertains. Notons, que le calcul de ce rang peut s'opérer avec des techniques reposant sur les inégalités de Weyl et de Weinmann, ou l'utilisation des cercles de Gerschgorin (annexe F) (Schreier, 1997). Il est ainsi possible de réaliser une étude des valeurs propres ou des valeurs singulières de $C(\theta_C)$ afin de déterminer si certaines d'entre elles sont susceptibles de s'annuler pour des valeurs particulières des paramètres incertains. Autrement, une autre possibilité est d'émettre l'hypothèse que le rang de la matrice $C(\theta_C)$ est invariant sur le domaine de variation de θ_C . Cette supposition n'est quant à elle pas très restrictive dans la mesure où θ_C reflète seulement une erreur de modélisation qui se doit d'être modeste. En pratique, un capteur mesure toujours la même variable physique ; il n'est donc pas envisageable que pour certains points de fonctionnement, la structure du système de mesures change. Par exemple, il est peu probable qu'il existe des combinaisons de valeurs d'incertitudes engendrant une perte de rang par rapport à la partie certaine de la matrice d'observation. Cela voudrait dire que pour certaines valeurs du bruit ou de l'erreur relative sur les capteurs, une grandeur mesurée peut ou non être déduite des autres mesures. Il est donc

raisonnable de se limiter aux cas de modèles où le rang de $C(\theta_c)$ correspond au rang de sa partie nominale C_0 .

Pour les deux formes statiques, en cas de découplage partiel, il est nécessaire d'estimer l'état (un des avantages traditionnels de l'Espace de Parité est perdu) ; ce qui va considérablement complexifier les formes d'évaluation (III-36) et (III-40) qui ne seront plus linéaires en les paramètres incertains. D'autre part, il s'agit d'une estimation "ensembliste" : en prenant en considération les incertitudes du système d'instrumentation, la grandeur reconstruite devient une variable incertaine, alors que la grandeur à estimer était initialement certaine. Par conséquent, toutes les valeurs possibles de l'état cohérentes avec le système de mesures sont prises en compte dans le cadre d'un fonctionnement normal. L'espace abstrait de la forme d'évaluation étudiée comprend donc bien toutes les valeurs possibles du résidu dans le cas non-défaillant. Si un défaut de capteur survient, l'estimation est faussée et le domaine de bon fonctionnement calculé ne comprend plus forcément le point de fonctionnement courant. Inversement, le fait de remplacer une valeur $x_i(k)$ par une variable bornée $\hat{x}_{\theta,i}(k)$ génère un domaine plus important que s'il avait été possible de parfaitement découpler les résidus. Ce dernier comprend par conséquent des valeurs que les résidus n'atteindront jamais. En réalité, comme c'est une méthode de bon fonctionnement qui est considérée où seule une incohérence permet de conclure à la présence d'un défaut, le problème précédent n'est pas gênant. Le seul risque est d'augmenter le taux de non-détections par rapport au cas où la véritable valeur de l'état serait accessible. Notons qu'en cas de découplage partiel, l'utilité de minimiser les termes résiduels ne se fait plus véritablement sentir.

En conclusion, la solution proposée n'est pas des plus pertinentes, principalement à cause de la condition de rang sur la matrice d'observation et d'une augmentation du risque de non-détections. Une solution bien plus efficace, présentée dans la section III.10 a retenu notre attention, elle consistera à s'intéresser à une matrice de parité incertaine.

III.7.2 Structure de modèle de capteurs

Comme dans le cas de la matrice de commande, il peut être intéressant d'utiliser une telle structure, plus naturelle et mieux adaptée à la prise en compte de tolérances portant sur les capteurs et de bruits de mesure, qui définissent la majorité des incertitudes pouvant affecter la matrice d'observation. Le modèle d'état peut alors être décrit de la manière suivante (Adrot et al, 1999a), (Adrot et al, 2000a) :

$$\begin{cases} \mathbf{x}(k+1) = \mathbf{A}\mathbf{x}(k) + \mathbf{B}\mathbf{u}(k) \\ \mathbf{y}^*(k) = \mathbf{C}_0\mathbf{x}(k) \\ \mathbf{y}(k) = (\mathbf{I}_m + \mathbf{T}_\Delta\mathbf{A}(k))\mathbf{y}^*(k) + \mathbf{T}_\delta\boldsymbol{\delta}(k) \end{cases} \quad \begin{matrix} \mathbf{x} \in \mathbb{R}^n, \mathbf{y} \in \mathbb{R}^m, \mathbf{y}^* \in \mathbb{R}^m \\ \boldsymbol{\delta} \in \mathbb{R}^m, \mathbf{A} \in \mathbb{R}^{m \times m}, \mathbf{u} \in \mathbb{R}^q \end{matrix} \quad (\text{III-46})$$

où \mathbf{y} désigne le vecteur des observations et \mathbf{y}^* représente les valeurs vraies associées qui ne tiennent compte que de la partie certaine \mathbf{C}_0 du système de mesures. La matrice incertaine \mathbf{A}

est diagonale, carrée d'ordre m et ses éléments diagonaux correspondent à m variables bornées normalisées. La matrice de pondération \mathbf{T}_Δ est diagonale, carrée d'ordre m , ses éléments diagonaux coïncidant avec la tolérance des capteurs étudiés selon le même principe qu'à la section III.5.2. Notons que la matrice $\mathbf{I}_m + \mathbf{T}_\Delta \Delta(k)$ est alors inversible. Le vecteur $\boldsymbol{\delta}$ est lui aussi constitué de m variables bornées normalisées et \mathbf{T}_δ est une matrice de pondération carrée d'ordre m .

Le système de mesures peut donc s'écrire :

$$\mathbf{C}(\Delta) = (\mathbf{I}_m + \mathbf{T}_\Delta \Delta(k))\mathbf{C}_0 \text{ et } \mathbf{y}(k) = (\mathbf{I}_m + \mathbf{T}_\Delta \Delta(k))\mathbf{C}_0 \mathbf{x}(k) + \mathbf{T}_\delta \boldsymbol{\delta}(k).$$

Les avantages et inconvénients d'une telle structure sont identiques à ceux rencontrés dans le cadre de la matrice de commande (section III.5.2).

III.7.2.1 Première forme statique

Déterminons dans un premier temps la nouvelle expression de la forme d'évaluation associée au modèle (III-46) dans le cadre de la première forme statique. La relation (III-5) liant le vecteur $\mathbf{x}(k+j)$ ($j \in \mathbb{N}^*$) à $\mathbf{x}(k)$ et aux commandes reste inchangée. De la relation (III-6) restant valide puisqu'elle ne dépend pas des mesures, se déduit que :

$$\mathbf{L}_{s,0} \mathbf{X}(k,s) = \mathbf{C}_{s,0} \mathbf{x}(k) + \mathbf{H}_{s,0} \mathbf{U}(k,s-1). \quad (\text{III-47})$$

En revanche l'égalité (III-7) devient :

$$\begin{aligned} \mathbf{Y}(k,s) &= \mathbf{L}_s(\Delta) \mathbf{X}(k,s) + \text{diag}_{s+1}(\mathbf{T}_\delta) \mathbf{D}_\delta(k,s) \\ &= (\mathbf{I}_{s_m} + \text{diag}_{s+1}(\mathbf{T}_\Delta) \mathbf{D}_\Delta(k,s)) \mathbf{L}_{s,0} \mathbf{X}(k,s) + \text{diag}_{s+1}(\mathbf{T}_\delta) \mathbf{D}_\delta(k,s), \end{aligned} \quad (\text{III-48})$$

$$\text{avec : } \mathbf{D}_\Delta(k,s) = \begin{bmatrix} \Delta(k) & 0 & 0 \\ 0 & \ddots & 0 \\ 0 & 0 & \Delta(k+s) \end{bmatrix}, \mathbf{D}_\Delta \in \mathbb{R}^{s_m \times s_m}, \mathbf{D}_\delta(k,s) = \begin{bmatrix} \boldsymbol{\delta}(k) \\ \vdots \\ \boldsymbol{\delta}(k+s) \end{bmatrix}, \mathbf{D}_\delta \in \mathbb{R}^{s_m},$$

ce qui entraîne :

$$\mathbf{Y}(k,s) = \mathbf{C}_s(\Delta) \mathbf{x}(k) + \mathbf{H}_s(\Delta) \mathbf{U}(k,s-1) + \text{diag}_{s+1}(\mathbf{T}_\delta) \mathbf{D}_\delta(k,s),$$

$$\begin{aligned} \text{avec } \mathbf{C}_s(\Delta) &= \begin{bmatrix} (\mathbf{I}_m + \mathbf{T}_\Delta \Delta(k))\mathbf{C}_0 \\ (\mathbf{I}_m + \mathbf{T}_\Delta \Delta(k+1))\mathbf{C}_0 \mathbf{A} \\ \vdots \\ (\mathbf{I}_m + \mathbf{T}_\Delta \Delta(k+s))\mathbf{C}_0 \mathbf{A}^s \end{bmatrix} = (\mathbf{I}_{s_m} + \text{diag}_{s+1}(\mathbf{T}_\Delta) \mathbf{D}_\Delta(k,s)) \mathbf{C}_{s,0}, \\ \mathbf{H}_s(\Delta) &= \begin{bmatrix} 0 & \cdots & 0 \\ (\mathbf{I}_m + \mathbf{T}_\Delta \Delta(k+1))\mathbf{C}_0 \mathbf{B} & \cdots & 0 \\ \vdots & \ddots & \vdots \\ (\mathbf{I}_m + \mathbf{T}_\Delta \Delta(k+s))\mathbf{C}_0 \mathbf{A}^{s-1} \mathbf{B} & \cdots & (\mathbf{I}_m + \mathbf{T}_\Delta \Delta(k+s))\mathbf{C}_0 \mathbf{B} \end{bmatrix}. \\ &= (\mathbf{I}_{s_m} + \text{diag}_{s+1}(\mathbf{T}_\Delta) \mathbf{D}_\Delta(k,s)) \mathbf{H}_{s,0} \end{aligned}$$

À partir de la relation précédente, en respectant la contrainte suivante, qui, nous allons le montrer, implique un découplage total des relations de parité vis-à-vis des variables d'état :

$$\mathbf{w}^T \mathbf{C}_{s,0} = 0,$$

l'expression suivante est obtenue :

$$\underbrace{\mathbf{w}^T (\mathbf{Y}(k,s) - \mathbf{H}_{s,0} \mathbf{U}(k,s-1))}_{p_{cal}(k+s)} = \dots$$

$$\dots = \underbrace{\mathbf{w}^T \text{diag}_{s+1}(\mathbf{T}_\Delta) \mathbf{D}_\Delta(k,s) \{ \mathbf{C}_{s,0} \mathbf{x}(k) + \mathbf{H}_{s,0} \mathbf{U}(k,s-1) \} + \mathbf{w}^T \text{diag}_{s+1}(\mathbf{T}_\delta) \mathbf{D}_\delta(k,s)}_{p_{\acute{e}val}(k+s)}.$$

La forme de calcul reste logiquement identique à celle explicitée dans le cadre de la structure classique. Montrons maintenant qu'avec le choix du vecteur \mathbf{w} effectué, l'expression précédente de la forme d'évaluation peut se ramener à une nouvelle ne faisant plus intervenir les variables d'état. Après avoir remplacé le terme entre accolades par son expression (III-47), la forme d'évaluation $p_{\acute{e}val}$ devient :

$$p_{\acute{e}val}(k+s) = \mathbf{w}^T \text{diag}_{s+1}(\mathbf{T}_\Delta) \mathbf{D}_\Delta(k,s) \mathbf{L}_{s,0} \mathbf{X}(k,s) + \mathbf{w}^T \text{diag}_{s+1}(\mathbf{T}_\delta) \mathbf{D}_\delta(k,s).$$

À partir de l'équation (III-48), la relation entre les mesures et les vecteurs d'état exprimés sur l'horizon temporel $[k, k+s]$ peut facilement être obtenue :

$$\mathbf{L}_{s,0} \mathbf{X}(k,s) = \left(\mathbf{I}_{s_m} + \text{diag}_{s+1}(\mathbf{T}_\Delta) \mathbf{D}_\Delta(k,s) \right)^{-1} \left(\mathbf{Y}(k,s) - \text{diag}_{s+1}(\mathbf{T}_\delta) \mathbf{D}_\delta(k,s) \right). \quad (\text{III-49})$$

En remplaçant le produit $\mathbf{L}_{s,0} \mathbf{X}(k,s)$ par cette expression dans la forme d'évaluation précédente, puis en utilisant le lemme d'inversion matriciel :

$$\left(\mathbf{I}_{s_m} + \text{diag}_{s+1}(\mathbf{T}_\Delta) \mathbf{D}_\Delta(k,s) \right)^{-1} = \mathbf{I}_{s_m} - \text{diag}_{s+1}(\mathbf{T}_\Delta) \mathbf{D}_\Delta(k,s) \left(\mathbf{I}_{s_m} + \text{diag}_{s+1}(\mathbf{T}_\Delta) \mathbf{D}_\Delta(k,s) \right)^{-1} \quad (\text{III-50})$$

$p_{\acute{e}val}$ peut s'écrire comme suit (Adrot et al, 2000a) :

$$p_{\acute{e}val}(k+s) = \mathbf{w}^T \text{diag}_{s+1}(\mathbf{T}_\Delta) \mathbf{D}_\Delta(k,s) \left(\mathbf{I}_{s_m} + \text{diag}_{s+1}(\mathbf{T}_\Delta) \mathbf{D}_\Delta(k,s) \right)^{-1} \mathbf{Y}(k,s) + \dots$$

$$\dots + \mathbf{w}^T \left(\mathbf{I}_{s_m} + \text{diag}_{s+1}(\mathbf{T}_\Delta) \mathbf{D}_\Delta(k,s) \right)^{-1} \text{diag}_{s+1}(\mathbf{T}_\delta) \mathbf{D}_\delta(k,s) \quad (\text{III-51})$$

L'avantage de la structure de modèle de capteurs est alors évident par rapport à la structure classique. Le calcul de la matrice de parité est identique au contexte certain, si bien qu'il existe toujours une taille de fenêtre temporelle minimale pour laquelle cette matrice existe. Trouver une matrice orthogonale à $\mathbf{C}_{s,0}$ est ainsi bien plus simple que d'en trouver une possédant cette même propriété envers la matrice incertaine $\mathbf{C}_s(\Delta)$ (ou $\mathbf{C}_s(\boldsymbol{\theta}_C)$ dans le cas de la structure classique) pour laquelle il n'est pas certain qu'une telle solution existe (section III.7.1.4). Notons cependant que cette expression est non-linéaire en les paramètres incertains.

III.7.2.2 Seconde forme statique

La relation (III-10) n'est pas modifiée puisqu'elle ne dépend pas des mesures. En lui ajoutant la relation liant les observations aux variables d'état (III-48), une nouvelle expression de la seconde forme statique est obtenue tenant compte des spécificités du modèle (III-46) :

$$\tilde{\mathbf{C}}_s(\Delta) \mathbf{X}(k,s) = \tilde{\mathbf{H}}_s \begin{bmatrix} \mathbf{U}(k,s-1) \\ \mathbf{Y}(k,s) \end{bmatrix} - \begin{bmatrix} \text{diag}_{s+1}(\mathbf{T}_\delta) \mathbf{D}_\delta(k,s) \\ 0 \end{bmatrix},$$

$$\text{avec : } \tilde{\mathbf{C}}_s(\Delta) = \begin{bmatrix} (\mathbf{I}_{s_m} + \text{diag}_{s+1}(\mathbf{T}_\Delta)\mathbf{D}_\Delta(k,s))\mathbf{L}_{s,0} \\ \mathbf{M}_s \end{bmatrix}.$$

À partir de la relation précédente, en respectant la contrainte suivante, qui, nous allons le montrer, implique un découplage total des relations de parité vis-à-vis des variables d'état :

$$\tilde{\mathbf{w}}^T \tilde{\mathbf{C}}_{s,0} = \begin{bmatrix} \tilde{\mathbf{w}}_L^T & \tilde{\mathbf{w}}_M^T \end{bmatrix} \begin{bmatrix} \mathbf{L}_{s,0} \\ \mathbf{M}_s \end{bmatrix} = 0,$$

l'expression suivante est obtenue :

$$\underbrace{\tilde{\mathbf{w}}^T \tilde{\mathbf{H}}_s \begin{bmatrix} \mathbf{U}(k,s-1) \\ \mathbf{Y}(k,s) \end{bmatrix}}_{\tilde{p}_{cal}(k+s)} = \underbrace{\tilde{\mathbf{w}}_L^T \text{diag}_{s+1}(\mathbf{T}_\Delta)\mathbf{D}_\Delta(k,s)\mathbf{L}_{s,0}\mathbf{X}(k,s) + \tilde{\mathbf{w}}_L^T \text{diag}_{s+1}(\mathbf{T}_\delta)\mathbf{D}_\delta(k,s)}_{\tilde{p}_{eval}(k+s)}.$$

En remplaçant le produit $\mathbf{L}_{s,0}\mathbf{X}(k,s)$ par son expression (III-49), puis en utilisant le lemme d'inversion matricielle (III-50) au niveau de la relation précédente, la forme d'évaluation devient (Adrot et al, 2000a) :

$$\begin{aligned} \tilde{p}_{eval}(k+s) = & \tilde{\mathbf{w}}_L^T \text{diag}_{s+1}(\mathbf{T}_\Delta)\mathbf{D}_\Delta(k,s) \left(\mathbf{I}_{s_m} + \text{diag}_{s+1}(\mathbf{T}_\Delta)\mathbf{D}_\Delta(k,s) \right)^{-1} \mathbf{Y}(k,s) + \dots \\ & \dots + \tilde{\mathbf{w}}_L^T \left(\mathbf{I}_{s_m} + \text{diag}_{s+1}(\mathbf{T}_\Delta)\mathbf{D}_\Delta(k,s) \right)^{-1} \text{diag}_{s+1}(\mathbf{T}_\delta)\mathbf{D}_\delta(k,s) \end{aligned} \quad (\text{III-52})$$

Tout comme la première forme statique, l'avantage de cette représentation des incertitudes est de rechercher une matrice de parité orthogonale à une matrice certaine. Notons aussi que la forme d'évaluation précédente est non-linéaire en les incertitudes paramétriques. Enfin, dans la section suivante, l'équivalence structurelle entre les deux formes statiques sera justifiée, ce qui permettra d'affirmer que là encore une taille de fenêtre minimale existe pour laquelle la condition d'existence de la matrice de parité est satisfaite.

III.7.2.3 Équivalence entre les deux formes statiques

L'objectif est de démontrer que les formes de calcul et d'évaluation sont structurellement équivalentes pour les deux formes statiques dans le cas où la matrice d'observation est incertaine, à structure de modèle de capteurs et qu'un découplage total vis-à-vis des variables d'état est assuré. La justification est semblable au développement opéré dans la section III.5.1.3, et ne seront donc repris que les points qui en diffèrent.

- Étape 1 : Relation entre les matrices de parité \mathbf{W} et $\tilde{\mathbf{W}}$

Les matrices $\mathbf{C}_{s,0}$ et $\tilde{\mathbf{C}}_{s,0}$ étant certaines, la relation (III-27) liant la matrice de parité \mathbf{W} à $\tilde{\mathbf{W}}_L$ reste valide :

$$\tilde{\mathbf{W}}_L = \mathbf{T}\mathbf{W},$$

où \mathbf{T} est une matrice inversible, traduisant le fait que le même nombre de relations de parité indépendantes ($s_m - \text{rang}(\mathbf{C}_{s,0})$) sera obtenu pour les deux formes statiques. Pour une taille de fenêtre s suffisante, le découplage demandé peut être réalisé en recherchant des matrices de

parité \mathbf{W} et $\tilde{\mathbf{W}}$ respectivement orthogonales aux matrices certaines $\mathbf{C}_{s,0}$ et $\tilde{\mathbf{C}}_{s,0}$ (sachant qu'il existe toujours une taille de fenêtre minimale le permettant, l'influence de la taille de fenêtre ayant déjà été étudiée dans la section III.2.2 dédiée au contexte certain).

- Étape 2 : Relation entre les résidus obtenus pour les deux formes statiques

Démontrer l'équivalence structurelle entre les formes de calcul procède toujours de la même manière. Puisque la structure classique et celle de modèle de capteurs ont exactement les mêmes formes de calcul, il s'en déduit directement que :

$$\tilde{\mathbf{p}}_{cal}(k+s) = \mathbf{T}\mathbf{p}_{cal}(k+s).$$

De même, l'équivalence structurelle entre les formes d'évaluation est facile à montrer. À partir de la forme d'évaluation de la seconde forme statique (III-52) et en tenant compte de la relation précédente liant \mathbf{W} et $\tilde{\mathbf{W}}$, la forme d'évaluation de la première forme statique (III-51) est directement retrouvée :

$$\tilde{\mathbf{p}}_{eval}(k+s) = \mathbf{T}\mathbf{p}_{eval}(k+s).$$

En conclusion, la structure consistant à modéliser de manière multiplicative les tolérances de capteur et de manière additive les bruits conduit à des relations de parité structurellement équivalentes pour les deux formes statiques et un même horizon temporel lorsque les formes d'évaluation obtenues ne dépendent plus des variables d'état (Adrot et al, 2000a). Ce lien montre qu'un découplage est toujours possible dans le cas de la seconde forme statique.

III.7.3 Récapitulatif

Le premier point à noter pour la prise en compte d'une matrice d'observation incertaine est qu'un découplage total des relations de parité vis-à-vis des incertitudes paramétriques est toujours possible dans le cas de la structure de modèle de capteurs, ce qui lui confère un net avantage sur la structure classique pour laquelle cette possibilité n'est pas toujours assurée. Les formes d'évaluation sont alors structurellement équivalentes pour les deux formes statiques et une même taille de fenêtre temporelle ; elles sont de plus toutes deux non-linéaires en les incertitudes.

En ce qui concerne la structure classique, les deux formes statiques ne sont pas équivalentes ; lorsqu'un découplage total est possible, la seconde forme conduit à un nombre plus important de relations de parité que son homologue pour une même valeur de l'horizon s , ce qui permet en pratique d'obtenir des relations de parité d'horizon plus faible. Les formes d'évaluation a priori linéaires en les incertitudes deviennent non-linéaires dans le cas d'un découplage partiel où les variables d'état doivent être estimées.

Forme statique	Première forme statique		Seconde forme statique	
Structure	Classique	Modèle de capteurs	Classique	Modèle de capteurs
Découplage	Total ou partiel	Total	Total ou partiel	Total
Forme interne linéaire en les incertitudes	Oui si non estimation de l'état	Non	Oui si non estimation de l'état	Non
Équivalence entre formes statiques	Structure classique : seulement l'implication ② \Rightarrow ① Structure de modèle de capteurs : oui			

Tableau III-4. Matrice d'observation incertaine

III.8 CAS DE PLUSIEURS MATRICES INCERTAINES

Le fait de traiter individuellement, dans les sections précédentes, chaque matrice incertaine de la représentation d'état, nous a permis de répertorier un certain nombre de propriétés propres à chaque forme statique ou structure d'incertitudes. Exploitions les dans un cadre plus général où plusieurs matrices de la représentation d'état peuvent être incertaines.

Dans un contexte général où des incertitudes affectent le système d'état dans son ensemble, une matrice de commande incertaine n'impose aucune restriction quant au choix de la forme statique ou de la structure à utiliser. Ce choix sera effectué en fonction des incertitudes associées aux termes résiduels au travers des matrices $C_s(\theta_A, \theta_C)$ ou $\tilde{C}_s(\theta_A, \theta_C)$ qui ne dépendent en fait aucunement de θ_B .

Supposons maintenant que les incertitudes n'affectent pas à la fois les matrices d'état et d'observation (aucune hypothèse n'est faite sur la matrice de commande). Si la matrice d'état est incertaine, il est nécessaire d'utiliser la seconde forme statique dans la mesure où elle seule permet d'explicitier l'expression analytique de la matrice \tilde{S} nécessaire à la résolution du problème d'optimisation visant à éliminer les variables d'état de la forme d'évaluation. Notons cependant qu'un tel découplage n'est pas forcément réalisable et dans ce cas, sauf si l'estimation de l'état est possible grâce à une matrice d'observation de plein rang colonne, la méthode proposée est mise en défaut. De plus, seule la structure classique a été développée dans la mesure où la matrice d'état n'est pas liée aux imperfections du système d'instrumentation. Si la matrice d'observation est incertaine, n'importe quelle structure ou forme statique peut être utilisée. Néanmoins, la structure de modèle de capteurs possède le privilège d'assurer un découplage total vis-à-vis des variables d'état, et ce quels que soient le modèle et la forme statique choisis, alors qu'une structure classique peut conduire à une impasse.

Dans le cas où les incertitudes affectent à la fois les matrices d'état et d'observation, la seconde forme statique a l'avantage de conduire à une matrice $\tilde{C}_s(\theta_A, \theta_C)$ s'écrivant sous la

forme d'une sommation linéaire en les incertitudes, et donc parfaitement linéaire en les paramètres incertains. Dans le cadre d'une structure classique, il est donc possible de lui appliquer, contrairement à $C_s(\theta_A, \theta_C)^{13}$, la procédure d'optimisation reposant sur la connaissance de l'expression analytique de \tilde{S} .

Notons pour terminer que, quelles que soient les matrices incertaines, la seconde forme statique conduit toujours à une expression linéaire en les incertitudes si l'état n'est pas estimé. Notons cependant que même si cette propriété simplifie les calculs à effectuer lors de l'évaluation de l'espace abstrait de la forme d'évaluation, travailler avec une matrice $H_s(\theta_A, \theta_B, \theta_C)$ non-linéaire en les incertitudes n'est pas prohibitif, un certain nombre de méthodes adaptées à ce type de difficultés ayant été explicitées dans le deuxième chapitre.

III.9 CAS D'INCERTITUDES INVARIANTES

Les incertitudes invariantes, telles qu'elles ont été définies dans la section III.3.1, ne sont en fait qu'un cas particulier de ce qui a été développé dans les sections précédentes III.3 à III.8. Les résultats (Adrot et al, 1998, 1999b) ne seront donc que brièvement présentés sans s'appesantir sur les calculs ou résultats intermédiaires qui peuvent facilement être retrouvés en adoptant un raisonnement similaire à ce qui précède. Seuls certains points seront plus largement détaillés lorsqu'ils conduisent à des conclusions différentes de ce qui est obtenu dans le cas variant. Ces résultats serviront à l'étude de l'influence de la nature temporelle des incertitudes sur la possibilité ou non d'obtenir un découplage total vis-à-vis des variables d'état.

III.9.1 Formes statiques

Puisqu'une incertitude θ_i est invariante, sa valeur est identique à deux instants différents k et t : $\theta_i(k) = \theta_i(t)$. D'un point de vue modélisation, ceci signifie qu'il n'est plus nécessaire de faire apparaître l'indice temporel au niveau des variables bornées θ_i :

$$\begin{cases} x(k+1) = A(\theta_A)x(k) + B(\theta_B)u(k) \\ y(k) = C(\theta_C)x(k) \end{cases} \quad x \in \mathbb{R}^n, y \in \mathbb{R}^m, u \in \mathbb{R}^q.$$

Si le nombre de variables bornées (en l'occurrence r), est moins élevé que pour un modèle à incertitudes variantes dans le temps (respectivement $(s+1)r$), en contrepartie les dépendances liées à la présence de variables communes sont plus importantes puisqu'une même incertitude θ_i va se retrouver dans les différentes relations de redondance analytique exprimées sur l'horizon $[k, k+s]$. De plus, cette différence modifie certaines propriétés des

¹³ Ces non-linéarités sont dues aux produits de la matrice A exprimée à des instants différents, eux-mêmes multipliés par la matrice d'observation.

formes d'évaluation, notamment en ce qui concerne le critère permettant le calcul de la matrice de parité et la possibilité d'un découplage total.

En respectant le même cheminement que dans les sections III.3.2 et III.3.3, les expressions des deux formes statiques associées au modèle précédent sont déduites.

Première forme statique	
$Y(k, s) = C_s(\theta_A, \theta_C)x(k) + H_s(\theta_A, \theta_B, \theta_C)U(k, s-1), U \in \mathbb{R}^{sq}, Y \in \mathbb{R}^{s_m}, s_m = (s+1)m,$	
$C_s(\theta_A, \theta_C) = \begin{bmatrix} C(\theta_C) \\ C(\theta_C)A(\theta_A) \\ \vdots \\ C(\theta_C)A^{s-1}(\theta_A) \end{bmatrix}, C_s \in \mathbb{R}^{s_m \times n}, H_s(\theta_A, \theta_B, \theta_C) = \begin{bmatrix} 0 & \cdots & 0 \\ C(\theta_C)B(\theta_B) & \cdots & 0 \\ \vdots & \ddots & \vdots \\ C(\theta_C)A^{s-1}(\theta_A)B(\theta_B) & \cdots & C(\theta_C)B(\theta_B) \end{bmatrix}, H_s \in \mathbb{R}^{s_m \times sq}.$	
Seconde forme statique	
$\tilde{C}_s(\theta_A, \theta_C)X(k, s) = \tilde{H}_s(\theta_B) \begin{bmatrix} U(k, s-1) \\ Y(k, s) \end{bmatrix}, X \in \mathbb{R}^{s_n}, U \in \mathbb{R}^{sq}, Y \in \mathbb{R}^{s_m}, s_n = (s+1)n,$	
où $\tilde{C}_s(\theta_A, \theta_C) = \begin{bmatrix} L_s(\theta_C) \\ M_s(\theta_A) \end{bmatrix}, \tilde{C}_s \in \mathbb{R}^{\{sn+s_m\} \times s_n}, \tilde{H}_s(\theta_B) = \begin{bmatrix} 0 & I_{s_m} \\ N_s(\theta_B) & 0 \end{bmatrix}, \tilde{H}_s \in \mathbb{R}^{\{sn+s_m\} \times \{sq+s_m\}},$	
avec $L_s(\theta_C) = \begin{bmatrix} C(\theta_C) & 0 & 0 \\ 0 & \ddots & 0 \\ 0 & 0 & C(\theta_C) \end{bmatrix}, L_s \in \mathbb{R}^{s_m \times s_n}, M_s(\theta_A) = \begin{bmatrix} A(\theta_A) & -I_n & 0 & \cdots & 0 \\ 0 & A(\theta_A) & -I_n & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & -I_n \end{bmatrix}, M_s \in \mathbb{R}^{sn \times s_n},$	
et $N_s(\theta_B) = \begin{bmatrix} -B(\theta_B) & 0 & 0 \\ 0 & \ddots & 0 \\ 0 & 0 & -B(\theta_B) \end{bmatrix}, N_s \in \mathbb{R}^{sn \times sq}.$	

Tableau III-5. Formes statiques dans le cas d'incertitudes invariantes (Adrot et al, 1998, 1999b)

Par rapport au cas d'incertitudes variantes, les expressions des formes statiques et les structures des différentes matrices restent inchangées. Ces dernières deviennent plus simples à manipuler dans la mesure où les indices temporels au niveau des incertitudes ont disparu.

III.9.2 Minimisation des termes résiduels

L'objectif consiste maintenant à rechercher une matrice de parité W orthogonale à $C_s(\theta)$ dans le cas de la première forme statique (respectivement \tilde{W} orthogonale à $\tilde{C}_s(\theta)$ dans le cas de la seconde forme statique). Le même choix que dans la section III.4.1 a été opéré. Trouver un vecteur ligne w^T le plus orthogonal possible à $C_s(\theta)$ consiste à minimiser au mieux l'ensemble des quantités $w^T C_s(\theta)$ pour toutes les valeurs possibles du vecteur incertain θ constitué des variables bornées normalisées $\theta_i, i \in \{1, \dots, r\}$, ce qui conduit à minimiser le critère suivant (Adrot et al, 1998, 1999b) :

$$j_c = \int_{-1}^1 \cdots \int_{-1}^1 \left\| \mathbf{w}^T \mathbf{C}_s(\boldsymbol{\theta}) \right\|^2 d\theta_1 \cdots d\theta_r = \mathbf{w}^T \underbrace{\left(\int_{-1}^1 \cdots \int_{-1}^1 \mathbf{C}_s(\boldsymbol{\theta}) \mathbf{C}_s(\boldsymbol{\theta})^T d\theta_1 \cdots d\theta_r \right)}_{\mathbf{S}} \mathbf{w}, \quad (\text{III-53})$$

avec $\mathbf{w}^T \mathbf{w} = 1$.

La solution est là encore donnée par le vecteur propre de la matrice \mathbf{S} associé à sa plus petite valeur propre. La valeur du critère j_c correspondant à la valeur propre en question, si la matrice carrée \mathbf{S} est singulière, il est alors possible d'annuler le critère j_c .

Lorsque la matrice $\mathbf{C}_s(\boldsymbol{\theta})$ est linéaire par rapport aux variables bornées et se présente sous la forme d'une sommation, il est possible de disposer de l'expression analytique de la matrice \mathbf{S} en effectuant un calcul similaire à celui présenté en annexe C. En partant de :

$$\mathbf{C}_s(\boldsymbol{\theta}) = \mathbf{C}_{s,0} + \sum_{i=1}^r \theta_i \mathbf{C}_{s,i},$$

où $\mathbf{C}_{s,0}$ représente la partie nominale de $\mathbf{C}_s(\boldsymbol{\theta})$ et où les matrices $\mathbf{C}_{s,i}$ sont certaines, la matrice \mathbf{S} devient :

$$\mathbf{S} = 2^r \left(\mathbf{C}_{s,0} \mathbf{C}_{s,0}^T + \frac{1}{3} (\mathbf{C}_{s,1} \mathbf{C}_{s,1}^T + \cdots + \mathbf{C}_{s,r} \mathbf{C}_{s,r}^T) \right).$$

Il est alors simple de montrer que si le critère j_c peut être nul, alors l'élimination complète des termes résiduels dépendant des variables d'états inconnues est possible lorsque la matrice $\mathbf{C}_s(\boldsymbol{\theta})$ peut se mettre sous la forme précédente. La nullité du terme $\mathbf{w}^T \mathbf{C}_s(\boldsymbol{\theta}) \mathbf{x}(k)$ pour toutes les valeurs admissibles des paramètres incertains et n'importe quelle valeur de l'état impose les contraintes suivantes :

$$\mathbf{w}^T \mathbf{C}_{s,0} = 0 \text{ et } \mathbf{w}^T \mathbf{C}_{s,i} = 0, \quad i \in \{1, \dots, r\}.$$

Lorsque le critère j_c est nul, en utilisant l'expression précédente de la matrice \mathbf{S} , l'égalité suivante est obtenue :

$$j_c = \mathbf{w}^T \mathbf{S} \mathbf{w} = 2^r \left\| \begin{bmatrix} \mathbf{C}_{s,0} & \frac{\mathbf{C}_{s,1}}{\sqrt{3}} & \cdots & \frac{\mathbf{C}_{s,r}}{\sqrt{3}} \end{bmatrix}^T \mathbf{w} \right\|^2 = 0.$$

La norme euclidienne étant nulle lorsque son argument est nul, la contrainte précédente est par conséquent équivalente aux $r+1$ contraintes explicitées juste avant.

Tous ces résultats sont directement transposables dans le cas de la seconde forme statique où l'objectif est de trouver un vecteur ligne $\tilde{\mathbf{w}}^T$ le plus orthogonal possible à la matrice $\tilde{\mathbf{C}}_s(\boldsymbol{\theta})$. Cette transposition s'opère simplement en remplaçant les notations \mathbf{w} , \mathbf{S} , \mathbf{C}_s et j_c par leurs homologues $\tilde{\mathbf{w}}$, $\tilde{\mathbf{S}}$, $\tilde{\mathbf{C}}_s$ et \tilde{j}_c :

$$\tilde{j}_c = \tilde{\mathbf{w}}^T \underbrace{\left(\int_{-1}^1 \cdots \int_{-1}^1 \tilde{\mathbf{C}}_s(\boldsymbol{\theta}) \tilde{\mathbf{C}}_s(\boldsymbol{\theta})^T d\theta_1 \cdots d\theta_r \right)}_{\tilde{\mathbf{S}}} \tilde{\mathbf{w}}. \quad (\text{III-54})$$

Première forme statique	
$\underbrace{\mathbf{w}^T (Y(k, s) - H_{s,0} U(k, s-1))}_{p_{cal}(k+s)} = \underbrace{\mathbf{w}^T C_s(\theta_A, \theta_C) \mathbf{x}(k)}_{\text{termes résiduels}} + \underbrace{\mathbf{w}^T H_{s,\theta}(\theta_A, \theta_B, \theta_C) U(k, s-1)}_{\text{termes incertains}}_{p_{eval}(k+s)}$ <p>où $H_{s,0} = H_s(0,0,0)$ et $H_{s,\theta}(\theta_A, \theta_B, \theta_C)$ désignent les parties certaine et incertaine de $H_s(\theta_A, \theta_B, \theta_C)$.</p>	
Seconde forme statique	
$\underbrace{\tilde{\mathbf{w}}^T \tilde{H}_{s,0} \begin{bmatrix} U(k, s-1) \\ Y(k, s) \end{bmatrix}}_{\tilde{p}_{cal}(k+s)} = \underbrace{\tilde{\mathbf{w}}^T \tilde{C}_s(\theta_A, \theta_C) X(k, s)}_{\text{termes résiduels}} - \underbrace{\tilde{\mathbf{w}}^T \tilde{H}_{s,\theta}(\theta_B) \begin{bmatrix} U(k, s-1) \\ Y(k, s) \end{bmatrix}}_{\text{termes incertains}}_{\tilde{p}_{eval}(k+s)}$ <p>où $\tilde{H}_{s,0} = \tilde{H}_s(0)$ et $\tilde{H}_{s,\theta}(\theta_B)$ désignent respectivement les parties certaine et incertaine de $\tilde{H}_s(\theta_B)$.</p>	

Tableau III-6. Expressions des formes de calcul et d'évaluation (Adrot et al, 1998, 1999b)

Les expressions des formes de calcul et d'évaluation des deux formes statiques correspondent exactement à celles déterminées dans un contexte d'incertitudes variantes en remplaçant simplement les différentes matrices par leurs expressions données dans la section précédente.

III.9.3 Matrice de commande incertaine

Les formes d'évaluation obtenues dans le cas d'une matrice de commande incertaine invariante sont résumées dans le tableau suivant :

	Structure classique : $B(\theta_B) = B_0 + \sum_{i=1}^r \theta_i B_i$	Structure de modèle de capteurs $\mathbf{x}(k+1) = A\mathbf{x}(k) + B_0(I_q + T_\Delta \Delta)^{-1} (u(k) - T_\delta \delta)$
Première forme statique	Matrice C_s certaine $H_s(\theta_B) = H_{s,0} + H_{s,\theta}(\theta_B)$ avec $H_{s,\theta}(\theta_B) = \sum_{i=1}^r \theta_i H_{s,i}$ où $H_{s,i} = \begin{bmatrix} 0 & \dots & 0 \\ CB_i & \dots & 0 \\ \vdots & \ddots & \vdots \\ CA^{s-1} B_i & \dots & CB_i \end{bmatrix}, i \in \{0, 1, \dots, r\}$. Recherche d'un vecteur ligne \mathbf{w}^T orthogonal à C_s : $\mathbf{w}^T C_s = 0$ $p_{eval}(k+s) = \mathbf{w}^T H_{s,\theta}(\theta_B) U(k, s-1) = \sum_{i=1}^r \theta_i \mathbf{w}^T H_{s,i} U(k, s-1)$	$Y(k, s) = C_s \mathbf{x}(k) + H_s(\Delta) U(k, s-1) - H_s(\Delta) \text{vect}_s(T_\delta \delta)$ où $\text{vect}_s(\mathbf{z})$ est un vecteur constitué de s fois le vecteur \mathbf{z} Matrice C_s certaine $H_s(\Delta) = H_{s,0} + H_{s,\theta}(\Delta)$, avec $H_{s,\theta}(\Delta) = -H_{s,0} \text{diag}_s(T_\Delta \Delta (I_q + T_\Delta \Delta)^{-1})$ Recherche d'un vecteur ligne \mathbf{w}^T orthogonal à C_s : $\mathbf{w}^T C_s = 0$ $p_{eval}(k+s) = \mathbf{w}^T H_{s,\theta} U(k, s-1) - \mathbf{w}^T H_s(\Delta) \text{vect}_s(T_\delta \delta)$

Seconde forme statique	<p>Matrice \tilde{C}_s certaine</p> <p>$\tilde{H}_s(\theta_B) = \tilde{H}_{s,0} + \tilde{H}_{s,\theta}(\theta_B)$ avec $\tilde{H}_{s,\theta}(\theta_B) = \sum_{i=1}^r \theta_i \tilde{H}_{s,i}$</p> <p>où $\tilde{H}_{s,0} = \begin{bmatrix} 0 & I_{s,m} \\ N_{s,0} & 0 \end{bmatrix}$, $\tilde{H}_{s,i} = \begin{bmatrix} 0 & 0 \\ N_{s,i} & 0 \end{bmatrix}$, $i \in \{1, \dots, r\}$</p> <p>et $N_{s,i} = \begin{bmatrix} -B_i & 0 & 0 \\ 0 & \ddots & 0 \\ 0 & 0 & -B_i \end{bmatrix}$, $i \in \{0, 1, \dots, r\}$.</p> <p>Recherche d'un vecteur ligne $\tilde{w}^T = \begin{bmatrix} \tilde{w}_L^T & \tilde{w}_M^T \end{bmatrix}$ orthogonal à \tilde{C}_s : $\tilde{w}^T \tilde{C}_s = 0$</p> <p>$\tilde{p}_{\text{éval}}(k+s) = -\tilde{w}^T H_{s,\theta}(\theta_B) \begin{bmatrix} U(k, s-1) \\ Y(k, s) \end{bmatrix}$</p> <p>$= -\sum_{i=1}^r \theta_i \tilde{w}_M^T N_{s,i} U(k, s-1)$</p>	<p>$\tilde{C}_s X(k, s) = H_s(\Delta) \begin{bmatrix} U(k, s-1) \\ Y(k, s) \end{bmatrix} + \begin{bmatrix} 0 \\ -N_s(\Delta) \text{vect}_s(T_\delta \delta) \end{bmatrix}$</p> <p>Matrice \tilde{C}_s certaine</p> <p>$\tilde{H}_s(\Delta) = \tilde{H}_{s,0} + \tilde{H}_{s,\theta}(\Delta)$, $\tilde{H}_{s,\theta}(\Delta) = \begin{bmatrix} 0 & 0 \\ N_{s,\theta}(\Delta) & 0 \end{bmatrix}$</p> <p>$N_s(\Delta) = N_{s,0} + N_{s,\theta}(\Delta)$,</p> <p>où $N_{s,\theta}(\Delta) = -N_{s,0} \text{diag}_s(T_\Delta \Delta (I_q + T_\Delta \Delta))$,</p> <p>Recherche d'un vecteur ligne $\tilde{w}^T = \begin{bmatrix} \tilde{w}_L^T & \tilde{w}_M^T \end{bmatrix}$ orthogonal à \tilde{C}_s : $\tilde{w}^T \tilde{C}_s = 0$</p> <p>$\tilde{p}_{\text{éval}}(k+s) = -\tilde{w}_M^T N_{s,\theta}(\Delta) U(k, s-1) + \dots$</p> <p>$\dots + \tilde{w}_M^T N_s(\Delta) \text{vect}_s(T_\delta \delta)$</p>
------------------------	--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------	---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------

Tableau III-7. Formes d'évaluation associées à une matrice de commande incertaine

Dans le cas d'une structure classique, les termes résiduels sont nuls et les termes incertains ne dépendent que de la commande, de plus, cette expression est totalement linéaire vis-à-vis des paramètres incertains. Dans le cas d'une structure de modèle de capteurs, la forme d'évaluation est parfaitement découplée vis-à-vis des variables d'état, mais cette expression est non-linéaire en les paramètres incertains.

En effectuant une démonstration dont le cheminement est identique à ce qui est proposé dans la section III.5.1.3 (respectivement III.5.2.3) dans le contexte d'incertitudes variantes dans le temps, l'équivalence structurelle entre les relations de parité obtenues pour les deux formes statiques dans le cas de la structure additive (respectivement de modèle de capteurs) peut être montrée.

Remarquons que puisque les matrices C_s , $C_{s,0}$, \tilde{C}_s ou $\tilde{C}_{s,0}$ sont certaines, elles sont indépendantes des incertitudes, et par conséquent de leur nature temporelle (variantes ou invariantes). Les résultats évoqués dans le contexte variant en ce qui concerne la possibilité d'un découplage total sont donc directement transposables ici. Il est ainsi toujours réalisable d'obtenir un découplage total des relations de parité vis-à-vis des variables d'état puisqu'à partir d'une taille suffisante de fenêtre temporelle, les matrices C_s , $C_{s,0}$, \tilde{C}_s ou $\tilde{C}_{s,0}$ ne sont plus de plein rang ligne. Ainsi, le choix de la taille de fenêtre temporelle s'obtient de la même manière que pour un modèle certain.

III.9.4 Matrice d'état incertaine

Pour les mêmes raisons que dans le contexte d'incertitudes variantes, seule la seconde forme statique associée à la structure classique sera ici développée :

$$A(\theta_A) = A_0 + \sum_{i=1}^r \theta_i A_i.$$

À partir de l'expression précédente de la matrice $A(\theta_A)$, se déduit :

$$\tilde{C}_s(\theta_A) = \tilde{C}_{s,0} + \sum_{i=1}^r \theta_i \tilde{C}_{s,i} \text{ avec } \tilde{C}_{s,0} = \begin{bmatrix} L_s \\ M_{s,0} \end{bmatrix} \text{ et } \tilde{C}_{s,i} = \begin{bmatrix} 0 \\ M_{s,i} \end{bmatrix},$$

$$\text{où } L_s = \text{diag}_{s+1}(C), M_{s,0} = \begin{bmatrix} A_0 & -I_n & 0 & \cdots & 0 \\ 0 & A_0 & -I_n & \cdots & 0 \\ \vdots & \vdots & \vdots & & \vdots \\ 0 & 0 & 0 & \cdots & -I_n \end{bmatrix} \text{ et } M_{s,i} = \begin{bmatrix} A_i & 0 & \cdots & 0 & 0 \\ 0 & A_i & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & A_i & 0 \end{bmatrix}.$$

La matrice \tilde{H}_s (III-11) ne fait pas intervenir la matrice d'état $A(\theta_A)$, $\tilde{H}_{s,\theta}$ se réduit donc à la matrice nulle et seuls subsistent les termes résiduels :

$$\tilde{p}(k+s) = \tilde{w}^T \tilde{C}_s(\theta_A) X(k,s) = \tilde{w}^T \tilde{C}_{s,0} X(k,s) + \sum_{i=1}^r \theta_i \tilde{w}^T \tilde{C}_{s,i} X(k,s).$$

Les termes incertains sont nuls contrairement aux termes résiduels, cette forme d'évaluation dépend donc uniquement du vecteur d'état exprimé sur un horizon temporel $[k, k+s]$.

Considérons l'expression de \tilde{S} (III-54) donnée dans le cas d'incertitudes invariantes par :

$$\tilde{S} = 2^r \left(\tilde{C}_{s,0} \tilde{C}_{s,0}^T + \frac{1}{3} (\tilde{C}_{s,1} \tilde{C}_{s,1}^T + \cdots + \tilde{C}_{s,r} \tilde{C}_{s,r}^T) \right),$$

Cette expression est en réalité rigoureusement identique à celle déterminée pour des incertitudes variant dans le temps (III-35), ce qui signifie que tous les résultats alors présentés sur les possibilités d'obtenir un découplage total et l'influence sur celui-ci de la taille de fenêtre temporelle s (section III.6.2) restent valides dans le cas d'incertitudes invariantes. Un découplage n'est donc pas toujours réalisable, auquel cas il est possible d'estimer l'état si la matrice d'observation est de plein rang colonne.

III.9.5 Matrice d'observation incertaine

Les formes d'évaluation obtenues dans le cas d'une matrice d'observation incertaine invariante sont résumées dans le tableau suivant :

	Structure classique : $C(\theta_C) = C_0 + \sum_{i=1}^r \theta_i C_i$	Structure de modèle de capteurs $y(k) = (I_m + T_\Delta \Delta) C_0 x(k) + T_\Delta \delta$
Première forme statique	$C_s(\theta_C) = C_{s,0} + \sum_{i=1}^r \theta_i C_{s,i} \text{ avec } C_{s,i} = \begin{bmatrix} C_i \\ \vdots \\ C_i A^s \end{bmatrix}$ $H_s(\theta_C) = H_{s,0} + H_{s,\theta}(\theta_C), \text{ avec } H_{s,\theta}(\theta_C) = \sum_{i=1}^r \theta_i H_{s,i}$ <p>où $H_{s,i} = \begin{bmatrix} 0 & \dots & 0 \\ C_i B & \dots & 0 \\ \vdots & \ddots & \vdots \\ C_i A^{s-1} B & \dots & C_i B \end{bmatrix}, i \in \{0, 1, \dots, r\}$.</p> <p>Recherche du vecteur w par minimisation de j_c via $S = 2^r R R^T, R = \begin{bmatrix} C_{s,0} & \frac{C_{s,1}}{\sqrt{3}} & \dots & \frac{C_{s,r}}{\sqrt{3}} \end{bmatrix}$. (III-55)</p> $p_{\text{éval}}(k+s) = w^T \{C_s(\theta_C) x(k) + H_{s,\theta}(\theta_C) U(k, s-1)\}$ $= w^T \left\{ C_{s,0} x(k) + \sum_{i=1}^r \theta_i \{C_{s,i} x(k) + H_{s,i} U(k, s-1)\} \right\}$	$Y(k, s) = C_s(\Delta) x(k) + H_s(\Delta) U(k, s-1) + \text{vect}_{s+1}(T_\Delta \delta)$ <p>où $\text{vect}_{s+1}(z)$ est un vecteur constitué de $s+1$ fois le vecteur z</p> $C_s(\Delta) = \text{diag}_{s+1}(I_m + T_\Delta \Delta) C_{s,0}$ $H_s(\Delta) = \text{diag}_{s+1}(I_m + T_\Delta \Delta) H_{s,0}$ <p>Recherche d'un vecteur ligne w^T orthogonal à la matrice certaine $C_{s,0}$: $w^T C_{s,0} = 0$</p> $p_{\text{éval}}(k+s) = w^T \text{diag}_{s+1}(T_\Delta \Delta) \{C_{s,0} x(k) + H_{s,0} U(k, s-1)\} + \dots$ $\dots + w^T \text{vect}_{s+1}(T_\Delta \delta)$ <p>ce qui peut aussi s'écrire :</p> $p_{\text{éval}}(k+s) = w^T \text{diag}_{s+1}(T_\Delta \Delta (I_m + T_\Delta \Delta)^{-1}) Y(k, s) + \dots$ $\dots + w^T \text{diag}_{s+1}((I_m + T_\Delta \Delta)^{-1}) \text{vect}_{s+1}(T_\Delta \delta)$
Seconde forme statique	$\tilde{C}_s(\theta_C) = \tilde{C}_{s,0} + \sum_{i=1}^r \theta_i \tilde{C}_{s,i}, \tilde{C}_{s,0} = \begin{bmatrix} L_{s,0} \\ M_s \end{bmatrix}, \tilde{C}_{s,i} = \begin{bmatrix} L_{s,i} \\ 0 \end{bmatrix}$ $L_{s,i} = \begin{bmatrix} C_i & 0 & 0 \\ 0 & \ddots & 0 \\ 0 & 0 & C_i \end{bmatrix}, i \in \{0, 1, \dots, r\}$ <p>Matrice \tilde{H}_s certaine</p> <p>Recherche du vecteur \tilde{w} par minimisation de \tilde{j}_c via $\tilde{S} = 2^r \tilde{R} \tilde{R}^T, \tilde{R} = \begin{bmatrix} \tilde{C}_{s,0} & \frac{\tilde{C}_{s,1}}{\sqrt{3}} & \dots & \frac{\tilde{C}_{s,r}}{\sqrt{3}} \end{bmatrix}$. (III-56)</p> $\tilde{p}_{\text{éval}}(k+s) = \tilde{w}^T \tilde{C}_s(\theta_C) X(k, s)$ $= \tilde{w}^T \tilde{C}_{s,0} X(k, s) + \sum_{i=1}^r \theta_i \tilde{w}^T \tilde{C}_{s,i} X(k, s)$	$\tilde{C}_s(\Delta) X(k, s) = \tilde{H}_s \begin{bmatrix} U(k, s-1) \\ Y(k, s) \end{bmatrix} - \begin{bmatrix} \text{vect}_{s+1}(T_\Delta \delta) \\ 0 \end{bmatrix}$ $\tilde{C}_s(\Delta) = \begin{bmatrix} \text{diag}_{s+1}(I_m + T_\Delta \Delta) L_{s,0} \\ M_s \end{bmatrix}$ <p>Matrice \tilde{H}_s certaine</p> <p>Recherche d'un vecteur ligne $\tilde{w}^T = \begin{bmatrix} \tilde{w}_L^T & \tilde{w}_M^T \end{bmatrix}$ orthogonal à la matrice certaine $\tilde{C}_{s,0}$: $\tilde{w}^T \tilde{C}_{s,0} = 0$</p> $\tilde{p}_{\text{éval}}(k+s) = \tilde{w}_L^T \{ \text{diag}_{s+1}(T_\Delta \Delta) L_{s,0} X(k, s) + \text{vect}_{s+1}(T_\Delta \delta) \}$ <p>ce qui peut aussi s'écrire :</p> $\tilde{p}_{\text{éval}}(k+s) = \tilde{w}_L^T \text{diag}_{s+1}(T_\Delta \Delta (I_m + T_\Delta \Delta)^{-1}) Y(k, s) + \dots$ $\dots + \tilde{w}_L^T \text{diag}_{s+1}((I_m + T_\Delta \Delta)^{-1}) \text{vect}_{s+1}(T_\Delta \delta)$

Tableau III-8. Formes d'évaluation associées à une matrice d'observation incertaine

Considérons tout d'abord la structure classique. Dans le cas de la première forme statique, nous montrerons ultérieurement qu'un découplage total est toujours possible ; les termes résiduels sont donc nuls et seuls subsistent les termes incertains. Cette forme d'évaluation est en outre linéaire en les incertitudes. Dans le cas de la seconde forme statique, un découplage total n'étant pas toujours possible, les termes résiduels peuvent donc être non nuls. S'ils le sont, cette forme d'évaluation est linéaire en les incertitudes ; sinon, si la matrice d'observation est de plein rang colonne, l'état peut être estimé, auquel cas $\tilde{p}_{\text{éval}}$ devient non-linéaire à cause du pseudo-inverse de la matrice d'observation. En revanche, les termes incertains sont systématiquement nuls. La structure de modèle de capteurs, quant à elle,

permet toujours un découplage total pour une taille de fenêtre temporelle adaptée et les formes d'évaluation sont alors systématiquement non-linéaires en les incertitudes.

L'équivalence structurelle des relations de parité obtenues pour les deux formes statiques dans le cas de la structure de modèle de capteurs se justifie exactement de la même façon que pour le cas d'incertitudes variant dans le temps (section III.7.2.3). En revanche, attardons-nous sur le cas de la structure classique.

- *Étape 1 : Relation entre les matrices de parité \mathbf{W} et $\tilde{\mathbf{W}}$*

① Soit $\tilde{\mathbf{W}}$ une matrice permettant d'annuler le critère \tilde{j}_c (III-54) pour une taille de fenêtre s :

$$\tilde{j}_c = \tilde{\mathbf{W}}\tilde{\mathbf{S}}\tilde{\mathbf{W}}^T = 0 \Leftrightarrow \tilde{\mathbf{W}}\tilde{\mathbf{R}} = 0, \text{ où la matrice } \tilde{\mathbf{R}} \text{ est définie en (III-56).}$$

Cette matrice est supposée de plein rang ligne de manière à ne conserver que des résidus indépendants et son rang vaut : $s_{\tilde{\mathbf{W}}} = s_m + sn - \text{rang}(\tilde{\mathbf{S}})$. Elle est partitionnée en deux sous-matrices $\tilde{\mathbf{W}}_L$ et $\tilde{\mathbf{W}}_M$ de dimensions respectives $s_{\tilde{\mathbf{W}}} \times s_m$ et $s_{\tilde{\mathbf{W}}} \times sn$ et l'équivalence précédente conduit aux $r+1$ égalités suivantes :

$$\begin{cases} \tilde{\mathbf{W}}_L \mathbf{L}_{s,0} + \tilde{\mathbf{W}}_M \mathbf{M}_s = 0 \\ \tilde{\mathbf{W}}_L \mathbf{L}_{s,i} = 0, \quad i \in \{1, \dots, r\} \end{cases} \quad (\text{III-57})$$

En multipliant à droite le système précédent par la matrice certaine \mathbf{A}_s (III-6), sachant que le produit $\mathbf{M}_s \mathbf{A}_s$ est nul (III-15) et que $\mathbf{L}_{s,i} \mathbf{A}_s = \mathbf{C}_{s,i}$, $i \in \{0, 1, \dots, r\}$ (III-12), les relations suivantes sont déduites :

$$\tilde{\mathbf{W}}_L \mathbf{C}_{s,i} = 0, \quad i \in \{0, 1, \dots, r\}, \text{ ce qui est équivalent à : } \tilde{\mathbf{W}}_L \mathbf{R} = 0, \quad (\text{III-58})$$

où la matrice \mathbf{R} est définie en (III-55).

Comme dans la section III.5.1.3, il est facile de montrer que si $\tilde{\mathbf{W}}$ est de plein rang ligne, alors il en est de même pour la matrice $\tilde{\mathbf{W}}_L$. En réalisant la décomposition en valeurs singulières de $\mathbf{R} : \mathbf{R} = \mathbf{U}\mathbf{\Sigma}\mathbf{V}^T$, à partir de (III-58), la matrice $\tilde{\mathbf{W}}_L$ peut s'écrire comme suit :

$$\tilde{\mathbf{W}}_L = \tilde{\mathbf{K}} \begin{bmatrix} 0 & \mathbf{I}_{s_m - \text{rang}(\mathbf{S})} \end{bmatrix} \mathbf{U}^T,$$

où $\tilde{\mathbf{K}}$ est une matrice carrée d'ordre $s_m - \text{rang}(\mathbf{S})$ inversible.

② Soit \mathbf{W} une matrice permettant d'annuler le critère j_c (III-53) pour une taille de fenêtre s :

$$j_c = \mathbf{W}\mathbf{S}\mathbf{W}^T = 0 \Leftrightarrow \mathbf{W}\mathbf{R} = 0.$$

Il existe donc une matrice carrée \mathbf{K} d'ordre $s_m - \text{rang}(\mathbf{S})$ inversible telle que :

$$\mathbf{W} = \mathbf{K} \begin{bmatrix} 0 & \mathbf{I}_{s_m - \text{rang}(\mathbf{S})} \end{bmatrix} \mathbf{U}^T.$$

L'existence de la matrice $\tilde{\mathbf{W}}$ ① (c'est à dire le fait que $\tilde{\mathbf{S}}$ ne soit pas de plein rang ligne) conduit nécessairement à l'existence de la matrice \mathbf{W} ② du fait de l'existence de la matrice $\tilde{\mathbf{W}}_L$. Cependant, dans le cas d'incertitudes sur la matrice d'observation, la réciproque n'est pas

toujours vérifiée. En effet, si $\mathbf{W} \textcircled{2}$ existe, alors les égalités (III-58) sont vérifiées, mais pas nécessairement (III-57), qui sont nettement plus restrictives. Supposons que le système (III-58) soit vérifié pour un indice i donné ; alors, il est possible d'écrire :

$$\tilde{\mathbf{W}}_L L_{s,i} \mathbf{A}_s = 0.$$

La solution générale d'une relation du type $\mathbf{XA}=\mathbf{B}$, où \mathbf{X} est la grandeur recherchée alors que \mathbf{A} et \mathbf{B} sont des matrices connues, existe si $\text{rang}\left(\begin{bmatrix} \mathbf{A} \\ \mathbf{B} \end{bmatrix}\right) = \text{rang}(\mathbf{A})$. Elle est alors donnée par $\mathbf{X} = \mathbf{BA}^+ + \mathbf{K}(\mathbf{AA}^+ - \mathbf{I})$ où \mathbf{K} est une matrice quelconque de dimension appropriée et \mathbf{A}^+ le pseudo-inverse généralisé¹⁴ de la matrice \mathbf{A} . En identifiant $\tilde{\mathbf{W}}_L L_{s,i}$ à \mathbf{X} , \mathbf{A}_s à \mathbf{A} et 0 à \mathbf{B} , la condition sur le rang est bien respectée et se déduit de manière générale que :

$$\tilde{\mathbf{W}}_L L_{s,i} = \mathbf{K}(\mathbf{A}_s \mathbf{A}_s^+ - \mathbf{I}_{s_n}).$$

Regardons quelle condition sur la matrice \mathbf{A}_s permettrait d'obtenir $\tilde{\mathbf{W}}_L L_{s,i} = 0$. Il faut pour cela que $\mathbf{A}_s \mathbf{A}_s^+ = \mathbf{I}_{s_n}$ si \mathbf{K} est non-nulle. En posant $\mathbf{X} = \mathbf{A}_s^+$, $\mathbf{A} = \mathbf{A}_s$ et $\mathbf{B} = \mathbf{I}_{s_n}$, la contrainte $\mathbf{AX}=\mathbf{B}$ n'admet une solution que si $\text{rang}(\begin{bmatrix} \mathbf{A} & \mathbf{B} \end{bmatrix}) = \text{rang}(\mathbf{A})$, soit $\text{rang}\left(\begin{bmatrix} \mathbf{A}_s & \mathbf{I}_{s_n} \end{bmatrix}\right) = \text{rang}(\mathbf{A}_s)$. Alors que le rang de la matrice identité \mathbf{I}_{s_n} vaut $s_n=(s+1)n$, \mathbf{A}_s ne possède que n colonnes et ne peut avoir qu'un rang strictement inférieur à s_n ($s>0$). La condition de rang nécessaire à l'existence d'une solution n'est jamais satisfaite, en conséquence il est impossible d'obtenir $\mathbf{A}_s \mathbf{A}_s^+ = \mathbf{I}_{s_n}$ et par conséquent $\tilde{\mathbf{W}}_L L_{s,i} = 0$ lorsque la matrice arbitraire \mathbf{K} est non-nulle. De manière générale, la contrainte $\tilde{\mathbf{W}}_L L_{s,i} \mathbf{A}_s = 0$ n'impose donc pas $\tilde{\mathbf{W}}_L L_{s,i} = 0$, et à fortiori $\tilde{\mathbf{W}}_L \mathbf{R} = 0$ n'implique pas $\tilde{\mathbf{W}}_L L_{s,i} = 0$, $i \in \{0, \dots, r\}$.

Pour conclure, pour une taille de fenêtre temporelle s , il est possible de trouver une matrice de parité $\mathbf{W} \textcircled{2}$ permettant d'annuler le critère j_c (\mathbf{S} est alors de plein rang ligne, ce qui permet un découplage total des résidus par rapport aux variables d'état) sans qu'il existe pour autant une matrice $\tilde{\mathbf{W}} \textcircled{1}$ assurant la nullité de \tilde{j}_c . En revanche, si un découplage total est possible pour la seconde forme statique $\textcircled{1}$, alors il en est de même pour la première forme statique $\textcircled{2}$.

À partir des expressions précédentes de \mathbf{W} et $\tilde{\mathbf{W}}_L$ issues de la décomposition en valeurs singulières de la matrice \mathbf{R} , en supposant que $\textcircled{1}$ (donc $\textcircled{2}$) soit vérifiée, il se déduit que la matrice $\tilde{\mathbf{W}}_L$ s'exprime comme combinaison linéaire des lignes de la matrice \mathbf{W} :

$$\tilde{\mathbf{W}}_L = \mathbf{TW}, \text{ avec } \mathbf{T} \in \mathbb{R}^{(s_m - \text{rang}(\mathbf{S})) \times (s_m - \text{rang}(\mathbf{S}))} \text{ matrice carrée inversible.}$$

- Étape 2 : Relation entre les résidus obtenus pour les deux formes statiques

En utilisant un cheminement identique au cas d'incertitudes variantes (section III.7.1.3), il est possible de démontrer que si la relation précédente existe entre les matrices de parité \mathbf{W} et $\tilde{\mathbf{W}}_L$, alors :

¹⁴ \mathbf{A}^+ est tel que : $\mathbf{AA}^+\mathbf{A}=\mathbf{A}$.

$$\tilde{\mathbf{p}}_{cal}(k+s) = \mathbf{T}\mathbf{p}_{cal}(k+s) \text{ et } \tilde{\mathbf{p}}_{\acute{e}val}(k+s) = \mathbf{T}\mathbf{p}_{\acute{e}val}(k+s).$$

Remarquons que dans le cas d'un découplage total, la forme d'évaluation de la seconde forme statique se résume à la valeur nulle (tableau III-8). Considérons maintenant la forme d'évaluation liée à la première forme statique (tableau III-8). Dans le cas d'un découplage total imposé par ②, la simplification suivante est obtenue :

$$\mathbf{p}_{\acute{e}val}(k+s) = \mathbf{W}\mathbf{H}_{s,\theta}(\boldsymbol{\theta}_C)\mathbf{U}(k,s-1).$$

Cette expression est non nulle a priori avec ce choix de \mathbf{W} et l'équivalence structurelle entre les deux formes d'évaluation ne peut être établie comme attendu. Si en revanche les contraintes associées à ① sont vérifiées, alors grâce à (III-14), il est simple de montrer que :

$$\tilde{\mathbf{W}}_L \mathbf{L}_{s,i} \mathbf{B}_s = \mathbf{T}\mathbf{W}\mathbf{H}_{s,i} = 0, \quad i \in \{1, \dots, r\},$$

ce qui impose $\mathbf{p}_{\acute{e}val}(k+s) = 0$. Par conséquent les relations de parité sont structurellement équivalentes pour les deux formes statiques si ① est satisfaite. Si seulement ② est vérifiée, il n'est plus possible de conclure puisque l'existence de $\tilde{\mathbf{W}}$ est remise en cause. Les résidus obtenus pour la seconde forme statique et issus d'un découplage total auront des équivalents dans le cas de la première forme statique. En revanche, l'inverse n'est pas nécessairement vrai. Pour une taille de fenêtre donnée s , un nombre supérieur ou égal de résidus sera obtenu par découplage total pour la forme statique traditionnelle. Mathématiquement, cela se traduit par un nombre plus important de valeurs propres nulles pour \mathbf{S} que pour $\tilde{\mathbf{S}}$.

Remarquons d'autre part que les expressions de la matrice $\tilde{\mathbf{S}}$ sont identiques, que les incertitudes soient invariantes (III-56) ou non (III-42). Ceci signifie que les résultats relatifs à la possibilité d'un découplage total et à l'influence de la taille de fenêtre peuvent directement être déduits de la section III.7.1.4 dans laquelle il a été montré que la nullité des termes résiduels ne peut pas toujours être atteinte. En revanche, les expressions de la matrice \mathbf{S} (III-55) dans le cas invariant :

$$\mathbf{S} = 2^{(s+1)r} \left(\mathbf{C}_{s,0} \mathbf{C}_{s,0}^T + \frac{1}{3} \sum_{i=1}^r \mathbf{S}_{var,i} \right) \text{ avec } \mathbf{S}_{var,i} = \begin{bmatrix} \mathbf{C}_i \mathbf{C}_i^T & 0 & 0 \\ 0 & \ddots & 0 \\ 0 & 0 & \mathbf{C}_i \mathbf{A}^s (\mathbf{A}^s)^T \mathbf{C}_i^T \end{bmatrix},$$

et (III-37) dans le cas variant :

$$\mathbf{S} = 2^r \left(\mathbf{C}_{s,0} \mathbf{C}_{s,0}^T + \frac{1}{3} \sum_{i=1}^r \mathbf{S}_{inv,i} \right) \text{ avec } \mathbf{S}_{inv,i} = \begin{bmatrix} \mathbf{C}_i \mathbf{C}_i^T & \dots & \mathbf{C}_i (\mathbf{A}^s)^T \mathbf{C}_i^T \\ \vdots & & \vdots \\ \mathbf{C}_i \mathbf{A}^s \mathbf{C}_i^T & \dots & \mathbf{C}_i \mathbf{A}^s (\mathbf{A}^s)^T \mathbf{C}_i^T \end{bmatrix},$$

sont complètement différentes, il est donc nécessaire dans le contexte invariant de prolonger cette étude. En réalité, en appliquant le théorème de Cayley-Hamilton à la matrice d'état \mathbf{A} comme cela a été fait dans le cas d'un modèle certain (section III.2.2), il est simple de montrer qu'il existe toujours une taille de fenêtre minimale à partir de laquelle la matrice \mathbf{S} n'est plus de plein rang ligne (Adrot et al, 1998, 1999b). En effet, en appliquant le même raisonnement, on déduit que :

$$\underbrace{\begin{bmatrix} a_0 i_{m,j}^T & \cdots & a_{n-1} i_{m,j}^T & -i_{m,j}^T \end{bmatrix}}_{\mathbf{w}^T} \begin{bmatrix} \mathbf{C}_i \\ \vdots \\ \mathbf{C}_i \mathbf{A}^n \end{bmatrix} = 0, \quad \begin{array}{l} \forall i \in \{0, 1, \dots, r\} \\ \forall j \in \{1, \dots, m\} \end{array}.$$

Il existe donc toujours un vecteur ligne \mathbf{w}^T non nul orthogonal aux différentes matrices $\mathbf{C}_{s,i}$, $i \in \{0, 1, \dots, r\}$. Par conséquent \mathbf{w}^T est aussi orthogonal à la matrice \mathbf{S} :

$$\mathbf{w}^T \mathbf{S} = 2^r \left\{ \mathbf{w}^T \mathbf{C}_{s,0} \mathbf{C}_{s,0}^T + \frac{1}{3} \sum_{j=1}^r \mathbf{w}^T \mathbf{C}_{s,j} \mathbf{C}_{s,j}^T \right\} = 0,$$

ce qui assure la nullité du critère d'optimisation j_c puisque dans ces conditions la matrice \mathbf{S} possède au moins une valeur propre nulle. En conclusion, donner à s une valeur égale ou supérieure à la dimension de l'état n conduit à l'existence de la matrice de parité dans le cas d'une structure classique associée à la première forme statique. Le découplage est alors total et les termes résiduels sont nuls. Il est par conséquent toujours possible, pour un choix approprié de l'horizon $[k, k+s]$, de ne plus faire apparaître l'état dans la forme d'évaluation du résidu. Ce résultat diffère de celui obtenu dans le cas d'incertitudes fluctuant dans le temps où un découplage n'est pas toujours accessible.

Remarque : Un modèle statique s'apparente à un système caractérisé par une matrice d'observation incertaine, dont il est un cas particulier. Son étude est développée dans l'annexe G (Adrot et al, 1999a).

III.9.6 Influence de la nature temporelle des incertitudes

Lorsque la matrice de commande ou la matrice d'état (structure classique associée à la seconde forme statique) ou encore la matrice d'observation (structure de modèle de capteurs, structure classique associée à la seconde forme statique) est incertaine, la matrice de parité est identique pour des incertitudes variantes ou invariantes. La nature temporelle des incertitudes n'a alors aucune incidence sur le choix de la taille de fenêtre en vue d'obtenir un découplage total par rapport aux variables d'état.

Il en est tout autrement dans le dernier cas restant correspondant à une matrice d'observation incertaine liée à une structure classique associée à la première forme statique. L'implication $(\textcircled{2} \Rightarrow \textcircled{1})$ démontrée dans le cas variant est à l'opposé de celle obtenue dans le cas invariant $(\textcircled{1} \Rightarrow \textcircled{2})$. Cette propriété peut être utilisée pour observer l'influence de la nature temporelle des paramètres incertains sur le choix de la taille de fenêtre s permettant un découplage total vis-à-vis des variables d'état.

Dans le cas variant, l'implication $(\textcircled{2} \Rightarrow \textcircled{1})$ indique que si un découplage total est possible pour une taille de fenêtre s pour la première forme statique, alors il en est de même pour la seconde forme statique. Mais il peut exister des valeurs plus petites de s où un découplage total est possible seulement pour la seconde forme statique. De ce fait, la seconde forme statique est en quelque sorte moins contraignante au niveau de l'horizon temporel que

son homologue. Dans le cas invariant, l'implication (① \Rightarrow ②) indique que la première forme statique est moins contraignante que la seconde forme statique. Pour finir, puisque la nature temporelle des incertitudes n'influence pas le choix de la taille de fenêtre temporelle dans le cas de la seconde forme statique, il s'en déduit que, dans le cadre de la première forme statique, la valeur de s permettant un découplage total (s'il est possible) sera plus importante pour des incertitudes dépendantes du temps. Ces différentes implications sont récapitulées dans le schéma suivant par les flèches blanches ; ainsi ① \Rightarrow ② signifie par exemple que toutes les relations de parité obtenues pour le cas ① se retrouveront en ② :

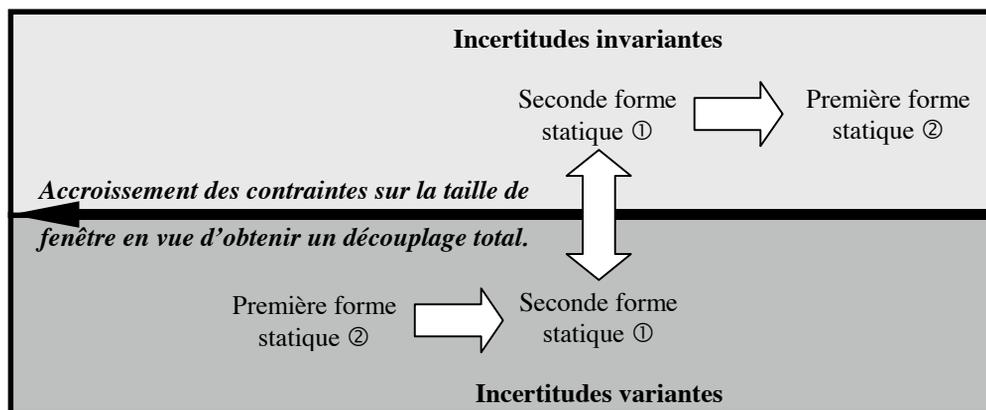


Figure III-2. Influence de la nature temporelle des incertitudes sur la taille de fenêtre

Ceci conforte les résultats obtenus dans le cas invariant où il a été démontré qu'il existait toujours une taille de fenêtre à partir de laquelle un découplage est possible. En revanche, dans le cas variant, l'existence de la matrice de parité est assujettie à des contraintes restrictives, et ce quelle que soit la taille de fenêtre.

III.10 MATRICE DE PARITÉ INCERTAINE

Face aux diverses impasses occasionnées par la méthode précédente, une seconde stratégie est proposée consistant à générer une matrice de parité dépendant des incertitudes (Adrot et al, 2000b), (Janati Idrissi et al, 2000).

III.10.1 Cas statique

Intéressons-nous dans un premier temps au cas d'un système statique, sachant que l'extension au cadre dynamique est pratiquement immédiate en utilisant l'une des deux formes statiques. Considérons le modèle statique augmenté suivant constitué d'une part de m équations de mesures et d'autre part de p équations de contraintes portant sur les n variables physiques du système. De manière à rester le plus général possible, ces deux catégories d'équations sont affectées par des incertitudes représentées par des variables bornées pouvant avoir une structure quelconque à la seule condition que ce modèle soit linéaire en les variables bornées :

$$\begin{bmatrix} \frac{y(k)}{0} \\ \frac{0}{p} \end{bmatrix} = \underbrace{\begin{bmatrix} C_1(\theta(k)) \\ C_2(\theta(k)) \end{bmatrix}}_{C(\theta(k))} x(k) + \underbrace{\begin{bmatrix} e_1(\theta(k)) \\ e_2(\theta(k)) \end{bmatrix}}_{e(\theta(k))}, \quad \begin{matrix} C_1 \in \mathbb{R}^{m \times n}, C_2 \in \mathbb{R}^{p \times n} \\ e_1 \in \mathbb{R}^m, e_2 \in \mathbb{R}^p \end{matrix},$$

où x et y désignent respectivement les vecteurs des mesures et des variables physiques à mesurer alors que le vecteur θ , élément de \mathbb{R}^r , est constitué de variables bornées normalisées et indépendantes. Les matrices de distribution C_i et les vecteurs e_i , $i \in \{1, 2\}$, sont supposés linéaires en les paramètres incertains ; de plus, pour rester général, une même incertitude θ_i peut affecter plusieurs d'entre eux. Les matrices C_1 et C_2 désignent respectivement les matrices d'observation et des contraintes alors que e_1 et e_2 sont les vecteurs représentant les incertitudes additives. En pratique, le terme e_1 définit les bruits additifs et supposés bornés du système de mesures alors que C_1 prend en compte les imperfections telles que les phénomènes non-linéaires de saturation, ..., modélisées par les précisions technologiques ou tolérances de capteurs sur le même principe que la structure de modèle de capteurs. La matrice C_2 et le vecteur e_2 définissent les relations de redondance entre les variables physiques. Ce modèle peut être avantageusement remplacé par une structure plus compacte :

$$\begin{bmatrix} \frac{y(k)}{0} \\ \frac{0}{p} \end{bmatrix} = C(\theta(k))x(k) + e(\theta(k)), \quad C \in \mathbb{R}^{(m+p) \times n}, \quad e \in \mathbb{R}^{m+p}, \quad (\text{III-59})$$

où les éléments de la matrice C et du vecteur e sont linéaires en les variables bornées.

L'objectif est d'obtenir des résidus dans le cas du modèle précédent en déterminant cette fois-ci une matrice de parité $W(\theta)$ qui n'est plus forcément constante¹⁵ et dépend des paramètres incertains $\theta_i(k)$, $i \in \{1, \dots, r\}$, tout en permettant d'éliminer les variables physiques inconnues $x(k)$ en assurant :

$$W(\theta(k))C(\theta(k)) = 0.$$

Il est bien sûr supposé que la matrice $C(\theta)$ n'est pas de plein rang ligne, sinon ceci signifie que le système ne contient aucune relation de redondance analytique, auquel cas il est clair que de toute façon, rien ne peut être diagnostiqué. Rappelons que différents outils pour tester le rang d'une matrice incertaine sont proposés dans l'annexe F. Pour simplifier, nous émettrons même l'hypothèse que le rang de la matrice $C(\theta)$ est invariant sur l'espace de variation des variables bornées. Cette hypothèse n'est guère contraignante dans la mesure où nous nous intéressons à des incertitudes paramétriques ne changeant pas les tendances comportementales d'un système physique : il est donc peu probable d'aboutir à des changements de signe de paramètres, à des annulations de paramètres modifiant l'observabilité ou le nombre de relations de redondance du système. En fait, $C(\theta)$ est supposée posséder le même rang que sa partie nominale $C(0)$, qui est une matrice certaine.

La matrice $W(\theta)$ doit être choisie de plein rang ligne de façon à éviter la présence de relations de parité redondantes. Le nombre de relations indépendantes correspond au nombre de lignes de la matrice $C(\theta)$ auquel lui est soustrait son rang. Cette matrice de parité est donc constituée de $(m+p) - \text{rang}(C_0)$ lignes indépendantes. Montrons dans un premier temps que

¹⁵ Cela dépend de la nature temporelle des incertitudes considérées.

$W(\theta)$ est une matrice polynomiale en les paramètres incertains. En effet, la solution générale du problème suivant $XA = B$ où A et B sont des matrices connues et X est recherchée, a pour expression : $X = BA^+ + K(I - AA^+)$ où K et A^+ sont respectivement une matrice arbitraire de dimension appropriée et le pseudo-inverse généralisé de la matrice A vérifiant $AA^+A = A$ sous la contrainte $B(I - A^+A) = 0$. En imposant respectivement que B soit une matrice nulle, que A corresponde à la matrice $C(\theta)$ nous obtenons :

$$W(\theta(k)) = K(I - C(\theta(k))C^+(\theta(k))).$$

En supposant que la matrice $C(\theta)$ soit de plein rang colonne, alors son pseudo-inverse peut simplement s'écrire comme suit : $C^+(\theta(k)) = (C^T(\theta(k))C(\theta(k)))^{-1}C^T(\theta(k))$, ce qui donne (Adrot et al, 2000b) :

$$W(\theta) = K \left(I_{m+p} - C(\theta(k)) (C^T(\theta(k))C(\theta(k)))^{-1} C^T(\theta(k)) \right),$$

où la matrice K permet de ne sélectionner que les lignes indépendantes du terme entre parenthèses. En réalité, $W(\theta)$ n'est rien d'autre que le complémentaire de $C(\theta)$ dans \mathbb{R}^{m+p} (Golub et al, 1989).

Remarquons que l'hypothèse sur le rang de $C(\theta)$ n'est pas restrictive puisque, si elle n'est pas vérifiée, il est possible de re-conditionner le problème en isolant les vecteurs colonnes générateurs de $C(\theta)$. En effet, une telle matrice peut toujours se décomposer de la manière suivante :

$$C(\theta(k)) = C_a(\theta(k))C_b^T(\theta(k)),$$

où les matrices C_a et C_b sont toutes deux de plein rang colonne. Pour le justifier, il suffit d'effectuer une décomposition en valeurs singulières de la matrice $C(\theta)$:

$$C(\theta) = U(\theta) \begin{bmatrix} \Sigma(\theta) & 0 \\ 0 & 0 \end{bmatrix} V(\theta)^T = [U_1(\theta) \quad U_2(\theta)] \begin{bmatrix} \Sigma(\theta) & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} V_1^T(\theta) \\ V_2^T(\theta) \end{bmatrix} = U_1(\theta) (\Sigma(\theta) V_1^T(\theta)),$$

Par définition, $U(\theta)$, $V^T(\theta)$ et $\Sigma(\theta)$ sont des matrices de plein rang, $U_1(\theta)$ est donc de plein rang colonne et $V_1^T(\theta)$ est de plein rang ligne. Multiplier une matrice par une autre inversible ne changeant pas son rang, $\Sigma(\theta)V_1^T(\theta)$ est donc aussi de plein rang ligne. Toute matrice $C(\theta)$ peut donc bien se décomposer en produit de deux matrices C_a et C_b de plein rang colonne : $C(\theta(k)) = C_a(\theta(k))C_b^T(\theta(k))$. Dans ces conditions, le problème revient tout simplement à rechercher une matrice de parité $W(\theta)$ orthogonale à la matrice $C_a(\theta)$, et par la même occasion orthogonale à $C(\theta)$. Remarquons que s'il est théoriquement possible de décomposer $C(\theta)$ de telle manière, son obtention est grandement facilitée par l'utilisation d'un logiciel de calcul formel, tel que Maple V (Char et al, 1991), (Cornil et al, 1995), avec toutes les limitations que cela impose. En prenant en compte cette nouvelle contrainte, W est donc constituée de $m+p-n$ lignes indépendantes et est élément de $\mathbb{R}^{(p+m-n) \times (p+m)}$.

Puisque la matrice \mathbf{K} est arbitraire, elle peut être remplacée par $\mathbf{K} \det(\mathbf{C}^T(\boldsymbol{\theta}(k))\mathbf{C}(\boldsymbol{\theta}(k)))$ où la fonction $\det()$ représente le déterminant de son argument. Dans ces conditions, la matrice de parité devient une matrice polynomiale en les différentes variables bornées :

$$\mathbf{W}(\boldsymbol{\theta}) = \mathbf{K} \left(\det(\mathbf{C}^T(\boldsymbol{\theta}(k))\mathbf{C}(\boldsymbol{\theta}(k))) - \mathbf{C}(\boldsymbol{\theta}(k)) \text{com}(\mathbf{C}^T(\boldsymbol{\theta}(k))\mathbf{C}(\boldsymbol{\theta}(k))) \mathbf{C}^T(\boldsymbol{\theta}(k)) \right),$$

où $\text{com}()$ désigne la comatrice de son argument. Le déterminant est forcément un polynôme en $\boldsymbol{\theta}$ d'ordre inférieur ou égal au nombre de colonnes de $\mathbf{C}(\boldsymbol{\theta})$, soit n . La partie dépendant de la comatrice est constituée de polynômes d'ordre maximal égal à 4 puisque $\mathbf{C}(\boldsymbol{\theta})$ est affine en les variables bornées. En conclusion, $\mathbf{W}(\boldsymbol{\theta})$ est une matrice de polynômes en $\boldsymbol{\theta}$ de degré inférieur ou égal à $\max(4, n)$ et au moins une solution existe si la matrice $\mathbf{C}(\boldsymbol{\theta})$ n'est pas de plein rang ligne (ou de manière équivalente d'après nos hypothèses si la matrice nominale $\mathbf{C}(0)$ n'est pas de plein rang ligne).

Remarque. Comme dans le cas certain, il est plus intéressant de rechercher les vecteurs générateurs de $\mathbf{W}(\boldsymbol{\theta})$, c'est-à-dire pour lesquels les pôles ou zéros en θ_i communs à tous les membres d'une même ligne de la matrice de parité ont été éliminés. La solution choisie pour déterminer la matrice de parité est l'utilisation du logiciel de calcul formel Maple V qui permet directement le calcul du noyau de la matrice $\mathbf{C}(\boldsymbol{\theta})$ ¹⁶. La matrice $\mathbf{W}(\boldsymbol{\theta})$ est alors constituée de lignes indépendantes et de structure minimale (les éliminations de pôles et zéros sont effectuées à l'aide de quelques lignes de programme). L'expression donnée ci-dessus de $\mathbf{W}(\boldsymbol{\theta})$ sous la forme du complémentaire de $\mathbf{C}(\boldsymbol{\theta})$ n'est en fait pas utilisée, mais nous a permis de justifier l'existence de la matrice recherchée.

Une fois la matrice $\mathbf{W}(\boldsymbol{\theta})$ non-linéaire en les variables bornées déterminée, le modèle (III-59) est multiplié à gauche par ce résultat pour obtenir le vecteur résidu \mathbf{p} nul en l'absence de défauts :

$$\mathbf{p}(k) = \mathbf{W}(\boldsymbol{\theta}(k)) \left\{ \begin{bmatrix} \mathbf{y}(k) \\ 0_p \end{bmatrix} - \mathbf{e}(\boldsymbol{\theta}(k)) \right\} = 0. \quad (\text{III-60})$$

Par analogie avec le cas d'une matrice de parité certaine et en notant $\mathbf{e}(0)$ et $\mathbf{W}(0)$ les parties nominales du vecteur $\mathbf{e}(\boldsymbol{\theta}(k))$ et de la matrice de parité $\mathbf{W}(\boldsymbol{\theta}(k))$, il est possible d'associer à cette expression des formes de calcul et d'évaluation définies comme suit :

$$\mathbf{p}_{cal}(k) = \mathbf{W}(0) \left\{ \begin{bmatrix} \mathbf{y}(k) \\ 0_p \end{bmatrix} - \mathbf{e}(0) \right\}, \quad (\text{III-61})$$

$$\mathbf{p}_{éval}(k) = \mathbf{W}(\boldsymbol{\theta}(k)) \left\{ \begin{bmatrix} \mathbf{y}(k) \\ 0_p \end{bmatrix} - \mathbf{e}(\boldsymbol{\theta}(k)) \right\} - \mathbf{p}_{cal}(k).$$

Suivant l'expression des relations de parité utilisée, le vecteur résidu désigne la forme de calcul \mathbf{p}_{cal} ou le vecteur \mathbf{p} .

¹⁶ Une solution purement numérique appelée méthode séquentielle est proposée dans (Janati Idrissi et al, 2000).

III.10.2 Extension dans le cas dynamique

L'extension dans le cadre d'un système dynamique est naturelle puisque les formes statiques permettent de se ramener à une représentation structurellement proche du modèle (III-59). Le système peut alors s'écrire sous les formes suivantes respectivement dans les cas de la première et seconde formes statiques (structure classique) :

$$Y(k, s) = C_s(\theta_A, \theta_C)x(k) + H_s(\theta_A, \theta_B, \theta_C)U(k, s-1),$$

$$\begin{bmatrix} Y(k, s) \\ 0 \end{bmatrix} = \tilde{C}_s(\theta_A, \theta_C)X(k, s) - \begin{bmatrix} 0 \\ N_s(\theta_B)U(k, s-1) \end{bmatrix}.$$

Examinons tout d'abord les cas où les incertitudes n'affectent qu'une seule matrice de la représentation d'état. Dans les cas où l'objectif consiste à trouver une matrice de parité orthogonale à une matrice certaine (matrice de commande incertaine, matrice d'observation incertaine dans le cas d'une structure de modèle de capteurs), le problème a déjà été traité : la matrice W est alors elle aussi certaine et la méthode proposée ci-dessus n'a plus de raison d'être. Dans le cas d'une matrice d'observation incertaine associée à une structure classique, les matrices $C_s(\theta_C)$ et $\tilde{C}_s(\theta_C)$ sont linéaires par rapport aux incertitudes paramétriques, le problème pourra donc être traité de la même façon que pour un modèle statique. Enfin, supposons que la matrice d'état soit incertaine. Pour la première forme statique, la matrice $C_s(\theta_A)$ est non-linéaire en les variables bornées ; si le calcul formel de $W(\theta)$ est théoriquement possible, les expressions obtenues sont beaucoup plus complexes. Il est donc préférable d'utiliser la seconde forme statique qui permet d'être confronté à une matrice $\tilde{C}_s(\theta_A)$ linéaire en les incertitudes.

Si plusieurs matrices de la représentation d'état sont incertaines, la seconde forme statique a l'avantage de rester linéaire en les incertitudes. Néanmoins la même remarque que précédemment s'applique à la matrice $C_s(\theta_A, \theta_C)$ dont la non-linéarité complexifie les résultats mais ne met pas en défaut la méthode. De plus, remarquons que pour la première forme statique, $H_s(\theta_A, \theta_B, \theta_C)$ est non-linéaire en les variables bornées, ce qui n'est en aucun cas gênant, la nature de cette matrice n'ayant aucune influence sur le calcul de $W(\theta)$. Remarquons qu'une fois multipliée par $W(\theta)$, l'expression obtenue sera de toute façon non-linéaire en les paramètres incertains.

Une fois la matrice $W(\theta)$ non-linéaire en les variables bornées déterminée, les formes statiques précédentes sont multipliées à gauche par ce résultat pour obtenir les vecteurs résidu suivants :

$$p(k+s) = W(\theta_A, \theta_C)\{Y(k, s) - H_s(\theta_A, \theta_B, \theta_C)U(k, s-1)\} = 0,$$

$$\tilde{p}(k+s) = \tilde{W}(\theta_A, \theta_C)\left\{\begin{bmatrix} Y(k, s) \\ 0 \end{bmatrix} + \begin{bmatrix} 0 \\ N_s(\theta_B)U(k, s-1) \end{bmatrix}\right\} = 0, \quad (\text{III-62})$$

en rappelant d'une part que θ_Z , $Z \in \{A, B, C\}$ contient toutes les incertitudes affectant la matrice Z et pouvant intervenir à n'importe quel instant de l'horizon temporel $[k, k+s]$, et que d'autre part, une même incertitude exprimée à deux instants différents est en fait représentée

par deux variables bornées indépendantes.

Par rapport au cas statique, le cas dynamique possède un degré de liberté supplémentaire défini par la taille de fenêtre temporelle s . La question est de savoir s'il existe toujours une taille de fenêtre s autorisant l'existence de la matrice de parité ; et si oui, comment régler s de manière à obtenir le nombre de résidus voulu ? Existe-t-il par exemple une solution $\mathbf{W}(\boldsymbol{\theta})$ lorsque la recherche d'une matrice de parité certaine conduit à une impasse ? En réalité, ce problème est lié au rang des matrices $\mathbf{C}_s(\boldsymbol{\theta})$ et $\tilde{\mathbf{C}}_s(\boldsymbol{\theta})$ qui conditionnent le nombre de relations de parité indépendantes calculables. L'intérêt du critère d'optimisation j_c (III-16) (ou \tilde{j}_c), dans l'approche consistant à prendre \mathbf{W} certaine était de pouvoir travailler sur une matrice certaine \mathbf{S} (ou $\tilde{\mathbf{S}}$) dont le rang était plus facilement accessible. Une fois mis en évidence le lien entre le rang de cette matrice, la valeur du critère d'optimisation et l'existence des termes résiduels¹⁷, l'étude de \mathbf{S} (ou $\tilde{\mathbf{S}}$) et des matrices par blocs qui la constituent permettait de déterminer l'influence de la taille fenêtre s sur la nature du découplage.

Le problème peut sembler plus délicat dans le cas d'une matrice de parité incertaine. En premier lieu, deux solutions très simples existent. La première consiste à ne s'intéresser qu'à des systèmes pour lesquels les matrices $\mathbf{C}_s(\boldsymbol{\theta})$ ou $\tilde{\mathbf{C}}_s(\boldsymbol{\theta})$ sont de même rang que leurs parties nominales. Une autre méthode plus générale consiste à essayer d'évaluer leurs rangs à l'aide des techniques proposées dans l'annexe F.

La solution adoptée, plus efficace, consiste à utiliser le logiciel de calcul formel Maple V pour déterminer le rang de ces matrices à l'aide de l'instruction *rank*¹⁸. Ce calcul est fondé sur la fonction *ffgausselim* qui comme l'indique son nom permet de réaliser une élimination de Gauss visant à transformer la matrice étudiée en une matrice triangulaire supérieure de même rang. Cette action est réalisée par permutations et combinaisons de lignes ou colonnes en conservant les expressions formelles des incertitudes. Il faut noter que le rang obtenu est un rang qualifié de "générique". Le rang générique d'une matrice $\mathbf{C}_s(\boldsymbol{\theta})$ ne tient pas compte du fait qu'il puisse se produire une perte de rang ponctuelle pour certaines combinaisons des valeurs des paramètres incertains. Prenons par exemple la matrice :

$$\mathbf{C}_s(\boldsymbol{\theta}) = \begin{bmatrix} 1 & \boldsymbol{\theta} \\ \boldsymbol{\theta} & 1 \end{bmatrix}.$$

Il est clair qu'elle est toujours de rang 2 sauf pour la valeur particulière $\boldsymbol{\theta}=1$. La fonction *rank* ne tenant pas compte des valeurs que peuvent prendre les paramètres formels, la valeur renvoyée par Maple V sera donc 2. Une fois le rang connu, le nombre de relations de parité indépendantes est donné par le nombre de lignes de la matrice étudiée moins son rang générique. L'intérêt de travailler sur le rang générique est d'obtenir des résidus valables quelles que soient les valeurs des paramètres incertains. Une perte de rang pour une valeur

¹⁷ Si la matrice \mathbf{S} (ou son homologue $\tilde{\mathbf{S}}$) n'est pas de plein rang ligne, alors elle possède au moins une valeur propre nulle pour laquelle d'une part le critère vaut zéro et d'autre part la relation de redondance obtenue est parfaitement découplée des variables d'état inconnues, entraînant par définition la nullité des termes résiduels.

¹⁸ La matrice étudiée ne doit alors comporter que des expressions rationnelles sous peine d'un message d'erreur lors de l'évaluation du rang en question.

particulière de $\theta = \theta^*$ engendre certes un résidu supplémentaire, mais celui-ci n'est utilisable que lorsque $\theta(k) = \theta^*$, instant impossible à déterminer puisque les incertitudes sont inconnues.

Considérons la première forme statique. La structure de la matrice $C_s(\theta)$ est telle qu'elle possède $s_m = (s+1)m$ lignes pour n colonnes. Lorsque la valeur de s dépasse la partie entière de $\frac{n}{m} - 1$, le nombre de lignes de cette matrice est supérieur au nombre de colonnes, il est alors certain que $C_s(\theta)$ n'est pas de plein rang ligne, condition d'existence de la matrice de parité recherchée. Il existe donc bien toujours une taille de fenêtre minimale pour laquelle la matrice $W(\theta)$ existe, et plus la valeur de s augmente, plus le nombre de relations obtenues est important, la différence entre le nombre croissant de lignes et le nombre fixe de colonnes $(s+1)m - n$ augmentant linéairement avec s . En fonction des expressions des matrices constituant $C_s(\theta)$, il est clair que $W(\theta)$ peut exister pour une valeur de s inférieure à la partie entière de $\frac{n}{m} - 1$.

Plaçons-nous dans le cadre de la seconde forme statique. Du fait de la structure de $\tilde{C}_s(\theta)$, il est certain que cette matrice n'est pas de plein rang ligne pour une valeur suffisante de la taille de fenêtre s . En effet, en soustrayant le nombre de lignes et de colonnes, nous retrouvons la valeur $(s+1)m - n$; plus s augmente, plus le nombre de relations de redondance (donc a fortiori les relations indépendantes recherchées) sera important.

En conclusion, utiliser une matrice $W(\theta)$ permet d'obtenir dans tous les cas un découplage total des résidus vis-à-vis des variables d'état inconnues, ce qui confère à cette approche un avantage sérieux par rapport à sa consœur où W est une matrice certaine, un découplage total n'étant alors pas toujours possible. La taille de fenêtre est choisie en incrémentant successivement s de manière à obtenir les $\text{rang}(C)$ ou $\text{rang}(C_0)$ relations de parité indépendantes attendues.

III.11 CONCLUSION

Dans un premier temps, le principe de l'Espace de Parité dans le cadre d'un modèle certain a été détaillé en vue de construire des relations de redondance à des fins de diagnostic de systèmes. Ce principe a ensuite été étendu dans le contexte incertain où la première stratégie proposée repose sur la recherche d'une matrice de parité constante visant à découpler les relations de parité des variables d'état inconnues. L'influence de la nature temporelle des incertitudes, du choix de la taille de fenêtre temporelle, de la forme statique utilisée et de la structure des incertitudes sur la possibilité d'un découplage est détaillée. Cependant, cette approche conduisant dans un certain nombre de cas à une impasse, une seconde stratégie reposant sur une matrice de parité incertaine est explicitée.

Les relations de redondance ainsi générées dépendent des incertitudes. Le chapitre suivant s'attache donc à développer des tests de cohérence applicables à ce type de résidus, qui peuvent, dans de nombreux cas, être non-linéaires en les variables bornées. Les méthodes développées reposeront sur l'analyse par intervalles détaillée dans le deuxième chapitre.

Chapitre IV

Tests de cohérence

TABLE DES MATIÈRES

IV.1 INTRODUCTION.....	163
IV.2 TESTS DE COHÉRENCE DANS LE CONTEXTE D’UN MODÈLE INCERTAIN.....	164
IV.2.1 Principe : approches directe et indirecte	164
IV.2.2 Garantie, précision et justesse	167
IV.3 MÉTHODE DE PAVAGE.....	169
IV.3.1 Méthode de pavage assurant la construction complète d’un espace abstrait.	170
IV.3.2 Méthode de pavage adaptée à une procédure de diagnostic.....	175
IV.4 EXPRESSION ANALYTIQUE DES FRONTIÈRES D’UN ESPACE ABSTRAIT	179
IV.5 MÉTHODE PROCÉDANT DE LA LINÉARISATION DES RELATIONS DE PARITÉ	186
IV.5.1 Procédure de linéarisation	186
IV.5.2 Procédure de réduction.....	189
IV.5.3 Génération des contraintes liées à un parallélotope	195
IV.5.4 Notion de distance.....	203
IV.5.5 Extension en vue d’une procédure de localisation	206
IV.6 CONCLUSION	210

IV.1 INTRODUCTION

Au cours du premier chapitre, nous avons expliqué que dans le contexte incertain, un modèle de référence définit un ensemble de comportements possibles, jugés normaux. L'objectif est alors de vérifier si les mesures issues du système physique sont cohérentes avec ceux-ci ; dans le cas contraire, la présence d'un défaut est suspectée. La manipulation des équations d'un modèle incertain en vue de les conditionner sous forme de relations de parité, conduit à des relations de redondance tout aussi incertaines reposant sur les expressions des formes de calcul et d'évaluation détaillées dans le troisième chapitre. L'influence de ces incertitudes se fait alors ressentir à deux niveaux. Le premier est lié à la recherche d'une matrice de parité (constante ou incertaine) assurant l'élimination des variables d'état inconnues afin de permettre l'évaluation des résidus obtenus. Son étude a été réalisée dans le précédent chapitre en précisant, suivant les matrices de la représentation d'état affectées par les incertitudes en question, quelles sont les formes statiques et les structures proposées les plus pertinentes. Le second niveau, développé dans ce chapitre-ci, concerne les tests de cohérence. Puisque les principes fondamentaux du diagnostic de systèmes ont été détaillés de manière générale dans le premier chapitre, l'objectif consiste maintenant à étudier comment ces derniers peuvent s'appliquer à la réalisation d'une procédure de diagnostic dans le cas où les incertitudes paramétriques sont décrites par des variables bornées, ainsi qu'à regarder comment ils peuvent s'adapter aux spécificités du problème traité. Nous nous intéresserons notamment à la notion de garantie introduite par l'analyse des intervalles et observerons son influence sur les résultats des tests de cohérence.

Le plan de ce chapitre se décompose comme suit. La seconde section développe les principes des tests de cohérence dans le contexte de l'approche bornante et s'attache à préciser les notions de garantie pour une procédure de diagnostic ainsi que de précision et de justesse pour un modèle incertain. Les sections suivantes proposent alors trois façons de mettre en œuvre ces tests, sachant que le principal objectif est de parvenir à évaluer le domaine de bon fonctionnement. Lorsque ce dernier est bien trop complexe à évaluer de manière exacte, une approximation doit être trouvée. La troisième section reprend la méthode de pavage développée dans le deuxième chapitre et l'améliore en l'adaptant aux objectifs du diagnostic de systèmes, notamment en ce qui concerne la limitation des temps de calcul. La quatrième section rappelle brièvement certains résultats développés par Ploix (1998) en ce qui concerne la détermination des expressions analytiques des frontières d'un espace abstrait, moyennant certaines hypothèses simplificatrices. Le domaine abstrait est alors évalué de manière exacte, cependant les résultats obtenus s'avèrent difficilement exploitables dans le cadre du diagnostic, spécifiquement lorsque le domaine de bon fonctionnement n'est pas convexe. L'objectif de cette démarche reste d'en présenter le principe, très analogue à la méthode numérique explicitée dans la cinquième section, et sur lequel reposent certaines procédures développées par la suite. Dans la section suivante, la résolution du problème génère une approximation du domaine recherché en procédant par linéarisation des relations

de parité en les variables bornées. Elle se présente sous la forme d'une série de procédures assurant respectivement la linéarisation des résidus, la réduction du nombre de variables bornées, puis la génération des contraintes inégalité représentatives du paralléloèdre solution permettant une mise en œuvre très simple des tests de cohérence. La notion de distance est alors abordée pour permettre de nuancer les résultats dichotomiques des précédents tests. Pour terminer, une brève extension à la localisation de défauts est présentée. A chaque fois, nous nous attacherons à détailler les avantages et inconvénients de chaque méthode.

IV.2 TESTS DE COHÉRENCE DANS LE CONTEXTE D'UN MODÈLE INCERTAIN

IV.2.1 Principe : approches directe et indirecte

Puisque les relations de parité définies dans le chapitre précédent sont incertaines, elles définissent elles aussi un ensemble de comportements possibles, c'est-à-dire représentatifs du modèle de référence sur lequel elles reposent. Repartons du formalisme associé aux formes de calcul et d'évaluation :

$$p_{cal}(k+s) = p_{éval}(k+s, \theta), \quad (IV-1)$$

en supposant que le vecteur $\theta \in \mathbb{R}^r$ contienne toutes les variables bornées indépendantes contenues dans la forme d'évaluation. Il s'ensuit que le système surveillé a un comportement normal à la seule condition qu'il existe au moins une combinaison de valeurs des variables bornées telle que :

- les équations du modèle soient vérifiées, ce qui induit l'égalité (IV-1) entre la forme de calcul déterminée et la forme d'évaluation incertaine associée,
- les valeurs trouvées des paramètres incertains soient acceptables, c'est-à-dire comprises entre -1 et 1 .

Soit \mathcal{H}_r , un hypercube ou orthotope aligné de \mathbb{R}^r , centré en l'origine et dont toutes les arêtes ont pour demi-longueur 1. Une première stratégie, appelée approche directe (Ploix, 1998), (Ploix et al, 2000), procède en quelque sorte de l'identification. Elle consiste à rechercher au moins une valeur du vecteur incertain θ située dans \mathcal{H}_r , telle que $p_{éval}$ prenne la valeur p_{cal} . Mathématiquement, ceci conduit à tester si la proposition suivante est vérifiée :

$$\{\exists \theta \in \mathcal{H}_r / p_{éval}(k+s, \theta) = p_{cal}(k+s)\}.$$

Malheureusement, cette approche reste complexe à réaliser, notamment dans le cadre de formes d'évaluation non-linéaires pour lesquelles plusieurs solutions peuvent exister et où il est impératif d'être sûr qu'il n'en existe aucune satisfaisant les contraintes précédentes pour pouvoir prétendre qu'une anomalie est présente. Notons que Ploix (1998) développe plus spécifiquement cette approche dans le cas de relations de redondance linéaires en les

paramètres. Mais même à l'aide de cette simplification, les tests à opérer restent complexes et font appel à des approximations amoindrissant la qualité de la détection.

Une seconde stratégie, appelée approche indirecte, peut avantageusement remplacer la précédente. Au lieu de rechercher une valeur particulière du vecteur incertain θ , il est possible de procéder par construction d'un ou plusieurs espaces abstraits. L'idée est que si le système est sain, alors l'égalité (IV-1) doit être vérifiée, ce qui signifie que la forme de calcul doit appartenir à l'espace abstrait de la forme d'évaluation puisque ce dernier représente l'ensemble des valeurs admissibles de $p_{\text{éval}}$, cohérentes avec le modèle de référence pour les mesures récoltées sur l'horizon temporel $[k, k+s]$. Le test de cohérence à l'instant $k+s$ consiste alors à tester si la valeur $p_{\text{cal}}(k+s)$ appartient au domaine $\mathcal{A}(p_{\text{éval}}(k+s, \theta))$.

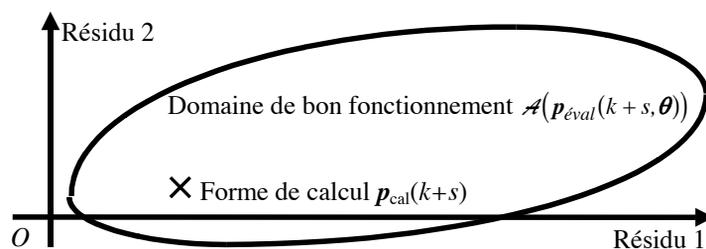


Figure IV-1. Principe des tests de cohérence dans le contexte incertain

Dans le cas où le vecteur résidu se limite à une seule composante, l'espace abstrait précédent se résume à chaque instant à un intervalle¹ :

$$\mathcal{A}(p_{\text{éval}}(k+s, \theta)) = \left[\inf_{\|\theta\|_{\infty} \leq 1} (p_{\text{éval}}(k+s, \theta)), \sup_{\|\theta\|_{\infty} \leq 1} (p_{\text{éval}}(k+s, \theta)) \right].$$

Le test de cohérence consiste alors à vérifier si la forme de calcul fait partie à chaque instant de cet intervalle, ce qui revient à construire une enveloppe.

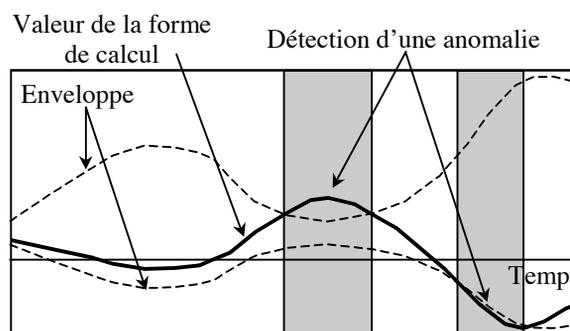


Figure IV-2. Cas d'une seule relation de parité

Si maintenant plusieurs relations de parité sont étudiées, $\mathcal{A}(p_{\text{éval}}(k+s, \theta))$ engendre une enveloppe multi-dimensionnelle.

Remarquons que ce test est rigoureusement identique à celui consistant à regarder si l'origine O de l'espace de résidus est contenu dans l'espace abstrait

¹ Comme justifié dans le chapitre II, nous ne nous sommes pas intéressés aux cas marginaux où la forme d'évaluation génère une division par un intervalle contenant 0, ainsi, le domaine obtenu est borné et fermé.

$\mathcal{A}(\mathbf{p}_{\text{éval}}(k+s, \boldsymbol{\theta}) - \mathbf{p}_{\text{cal}}(k+s))$; en effet, il s'agit exactement du même domaine que précédemment cette fois-ci centré en $\mathbf{p}_{\text{cal}}(k+s)$ alors que $\mathcal{A}(\mathbf{p}_{\text{éval}}(k+s, \boldsymbol{\theta}))$ l'est en O . En poussant le raisonnement, il est même possible de scinder la forme d'évaluation en deux champs de vecteurs $\mathbf{p}_{\text{éval},1}$ et $\mathbf{p}_{\text{éval},2}$:

$$\mathbf{p}_{\text{cal}}(k+s) - \mathbf{p}_{\text{éval},1}(k+s, \boldsymbol{\theta}_1) = \mathbf{p}_{\text{éval},2}(k+s, \boldsymbol{\theta}_2),$$

et de vérifier s'il existe au moins une valeur commune entre les espaces abstraits des deux membres de cette égalité, ce qui revient à tester si leur intersection est vide. Attention cependant, des précautions sont à prendre. Si aucune variable bornée n'appartient à la fois aux vecteurs incertains $\boldsymbol{\theta}_1$ et $\boldsymbol{\theta}_2$, ils sont par conséquent indépendants ; seule, alors, l'intersection est à tester (Adrot et al, 1999a). Si la proposition suivante est satisfaite :

$$\mathcal{A}(\mathbf{p}_{\text{cal}}(k+s) - \mathbf{p}_{\text{éval},1}(k+s, \boldsymbol{\theta}_1)) \cap \mathcal{A}(\mathbf{p}_{\text{éval},2}(k+s, \boldsymbol{\theta}_2)) = \{\emptyset\},$$

alors un défaut est détecté. En revanche, dans le cas contraire, ce test devient plus complexe. Supposons que le vecteur borné $\boldsymbol{\theta}^*$ apparaisse dans les deux champs de vecteurs $\mathbf{p}_{\text{éval},i}$, $i \in \{1,2\}$. Pour tester si l'intersection qui suit est vide :

$$\mathcal{A}(\mathbf{p}_{\text{cal}}(k+s) - \mathbf{p}_{\text{éval},1}(k+s, \boldsymbol{\theta}_1, \boldsymbol{\theta}^*)) \cap \mathcal{A}(\mathbf{p}_{\text{éval},2}(k+s, \boldsymbol{\theta}_2, \boldsymbol{\theta}^*)) = \{\emptyset\},$$

il est nécessaire de vérifier que chaque point commun obtenu entre ces deux domaines est bien associé aux mêmes arguments des deux parties considérées, c'est-à-dire que $\mathbf{p}_{\text{éval},1}$ et $\mathbf{p}_{\text{éval},2}$ conduisent à ce point pour la même valeur de $\boldsymbol{\theta}^*$ (Janati Idrissi et al, 2000). En effet, tester si :

$$\mathcal{A}(\mathbf{p}_{\text{cal}}(k+s) - \mathbf{p}_{\text{éval},1}(k+s, \boldsymbol{\theta}_1, \boldsymbol{\theta}_1^*)) \cap \mathcal{A}(\mathbf{p}_{\text{éval},2}(k+s, \boldsymbol{\theta}_2, \boldsymbol{\theta}_2^*)) = \{\emptyset\},$$

génère une perte d'information puisque des dépendances entre variables bornées n'ont pas été prises en compte, ce qui a pour conséquence d'accroître artificiellement le domaine de bon fonctionnement du domaine et d'engendrer des non-détections supplémentaires (en effet, des points communs entre les deux domaines sont trouvés alors qu'ils n'existent pas réellement). Cette remarque avait déjà été évoquée dans la section II.3.1.7.

Notons que les tests conduisant à vérifier si une valeur est à l'intérieur d'un domaine sont de manière générale plus judicieux dans la mesure où ils évitent d'une part la construction d'un espace abstrait supplémentaire, et que d'autre part, cette approche conserve le mérite de toujours s'appliquer. De plus, elle évite de mettre en œuvre la recherche de l'intersection entre deux domaines pouvant être complexes à manipuler.

Au cours du troisième chapitre, deux formalismes ont été utilisés pour générer les relations de parité. Dans le cas d'une matrice de parité certaine, nous nous sommes reposés sur les notions de formes de calcul \mathbf{p}_{cal} et d'évaluation $\mathbf{p}_{\text{éval}}$ traditionnellement rencontrées dans le contexte certain. Dans le cas d'une matrice de parité incertaine, pour des raisons de commodité, ces deux expressions ont été regroupées dans un seul et même terme \mathbf{p} ; néanmoins, le passage d'un formalisme à l'autre est trivial : $\mathbf{p}(k+s, \boldsymbol{\theta}) = \mathbf{p}_{\text{éval}}(k+s, \boldsymbol{\theta}) - \mathbf{p}_{\text{cal}}(k+s)$ (III-60, III-61). D'après ce qui précède, nous venons de

justifier qu'au niveau des tests de cohérence, ces deux formalismes sont aussi équivalents. Vérifier si l'origine O se situe dans $\mathcal{A}(p(k+s, \theta))$ revient exactement à la même chose que de tester si $p_{cal}(k+s)$ se situe dans $\mathcal{A}(p_{eval}(k+s, \theta))$.

IV.2.2 Garantie, précision et justesse

En considérant que les bornes choisies² pour les paramètres incertains permettent d'expliquer tous les comportements possibles du système physique en fonctionnement normal, alors la forme de calcul sera toujours située à l'intérieur de son enveloppe en l'absence de défaut. En effet, la détermination de la forme d'évaluation repose exclusivement sur l'analyse par intervalles. Celle-ci, comme nous l'avons détaillée dans le chapitre II, exploite la propriété de monotonie par inclusion (section II.2.4.6, définition 6) conduisant au théorème 1 fondamental énoncé par Moore (1979). Ce théorème stipule que quelle que soit la fonction d'inclusion d'une fonction vectorielle f réelle, son évaluation conduit à une majoration de l'intervalle minimal recherché, c'est-à-dire du domaine exact de valeurs prises par f . Or, puisqu'en diagnostic, une enveloppe délimite le domaine de bon fonctionnement du système, l'utilisation exclusive de fonctions d'inclusion permet de garantir les résultats en assurant que tous les comportements supposés normaux sont bien appréhendés.

En réalité, tous les outils présentés dans le second chapitre afin de construire une enveloppe génèrent des majorations des solutions recherchées que ce soient les méthodes itératives à base de formes centrées ou les méthodes de réduction ou d'élimination. La qualité de ces majorations s'accroît avec le nombre d'itérations et les temps de calcul, mais grâce à la propriété de garantie, si les variables bornées du modèle ont été correctement caractérisées, il est alors certain que l'enveloppe calculée contient tous les points de fonctionnement sains du système.

Dans ces conditions, une incohérence traduit nécessairement un comportement anormal et le modèle, pour lequel il ne peut y avoir de fausses alarmes, est alors qualifié de juste (Ploix, 1998) ou de complet³. Pratiquement, il est toujours possible d'obtenir un modèle juste en prenant des majorations importantes pour les bornes des paramètres incertains. Si la notion de justesse est très attrayante dans le cadre du diagnostic de systèmes, elle est cependant étroitement liée à la notion de précision. Un modèle incertain est d'autant plus précis que l'enveloppe qu'il induit restreint le nombre de comportements inaccessibles au système physique en l'absence d'un défaut. Au niveau de la phase de détection, un modèle précis engendre des enveloppes de résidus moins importantes qu'un modèle dit imprécis, ce qui aura pour conséquences de restreindre le nombre de non-détections.

² Le chapitre V sera entièrement dédié à des méthodes de caractérisation permettant de déterminer ces bornes.

³ Le terme anglo-saxon associé est "complete" (completeness traduit par complétude en français) (Armengol et al, 1999b). L'opposé, c'est-à-dire une enveloppe dont tous les points font assurément partie du domaine de bon fonctionnement, mais qui ne contient pas forcément tous les comportements normaux possibles, est qualifié de "sound" (soundness). Une enveloppe qui tient à la fois des deux correspond donc à l'enveloppe minimale et représente exactement le domaine de bon fonctionnement.

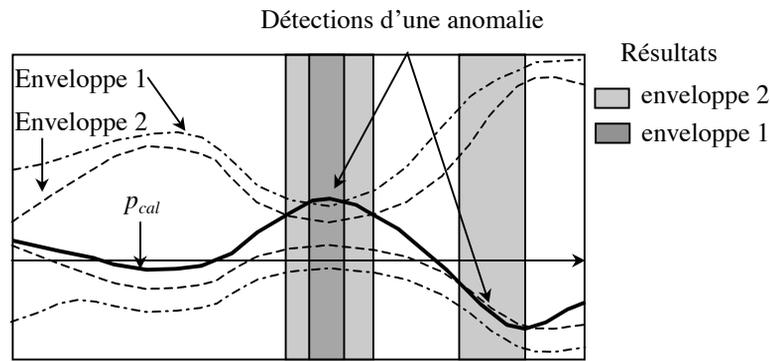


Figure IV-3. Notion de précision

Si la notion de justesse est binaire (un modèle est juste ou faux), la précision quant à elle apporte une information plus nuancée. Dans le cas de systèmes stochastiques par exemple, plus l'écart-type d'une variable est faible, plus sa détermination est précise. Sur l'exemple de la figure précédente, l'enveloppe 2 est plus précise que la première puisque globalement sur l'horizon temporel représenté, la somme cumulée de ses longueurs aux différents instants est inférieure à celle de son homologue, ce qui assure une meilleure détection.

Notons toutefois que la notion de justesse, essentielle au diagnostic dans le cas de procédures bon fonctionnement et défendue dans le cadre de ce travail, peut être remise en cause en pratique suivant les objectifs souhaités. Dans le cas d'un système nécessitant une surveillance très étroite (centrale nucléaire), si des techniques de diagnostic plus traditionnelles nécessitant un compromis fausses alarmes/non-détections sont utilisées, il est alors préférable d'obtenir un grand nombre de fausses alarmes plutôt que de laisser passer une anomalie risquant de s'aggraver. Dans ces conditions, la précision est primordiale. Inversement, sur une chaîne de production, de fausses-alarmes intempestives réduisent la confiance de l'opérateur envers le module de surveillance au point que celui-ci puisse s'en désintéresser. La justesse du modèle, associée à la propriété de garantie de l'approche bornante qui génère une approximation extérieure du domaine de bon fonctionnement, est alors intéressante.

Il est important de souligner que pour un modèle juste, seule la présence d'un défaut est garantie, mais l'approche proposée ne donne en aucun cas l'assurance de son absence. En effet, si la forme de calcul sort de l'espace abstrait de la forme d'évaluation associée, alors il est certain qu'un défaut est présent. En revanche, dans le cas contraire, même si la conclusion qui s'impose est de considérer le système sain, il est tout à fait possible :

- qu'un défaut de faible amplitude soit tout de même présent mais masqué par une ou plusieurs incertitudes,
- qu'un défaut de nature quelconque ne puisse être détecté, soit parce que les résidus y sont insensibles, ou simplement parce que l'excitation imposée au système n'est pas pertinente.

Remarquons à cet effet que même la stratégie proposée par (Armengol et al, 1999a), (Armengol, 1999c) dans le cadre d'un système incertain à horizon infini et consistant à générer une paire d'enveloppes, l'une majorant comme cela est proposé ici, l'autre minorant

l'enveloppe minimale, procède en fait du même raisonnement quant aux conclusions de la procédure de détection. Ces deux enveloppes génèrent trois zones, dont une seule permet de conclure avec garantie :

- si le point de fonctionnement courant se trouve à l'extérieur de l'enveloppe majorante, alors un défaut est assurément présent,
- s'il se trouve à l'intérieur de l'enveloppe minorante, alors un comportement sain est constaté (mais un défaut peut tout de même être présent et non détecté),
- s'il se trouve dans la zone d'indétermination entre ces deux enveloppes, la précision de ces dernières doit être améliorée (ce qui consiste à réduire l'enveloppe majorante et augmenter la taille de l'enveloppe minorante) jusqu'à ce qu'une décision puisse être prise.

L'enveloppe minorante ne garantit absolument pas l'absence de défaut, mais permet d'adapter l'effort effectué pour le calcul des enveloppes au point de fonctionnement courant. En partant de domaines grossiers, leur précision s'améliore en fonction des besoins, ce qui limite les temps de calcul aux dires de leurs auteurs, mêmes si deux enveloppes sont nécessaires. Ce ne sera pas l'approche ici développée qui se restreint à l'étude d'une unique enveloppe majorante.

IV.3 MÉTHODE DE PAVAGE

Que la matrice de parité soit certaine ou incertaine, les formes d'évaluation sont souvent non-linéaires en les variables bornées, conduisant à un espace abstrait de $\mathbf{p}_{\text{éval}}$ complexe et difficile, voire impossible, à évaluer exactement à cause de temps de calcul prohibitifs. La finalité de cette section est de proposer deux méthodes plus évoluées que la méthode de pavage de base présentée dans le chapitre II, permettant de construire une majoration suffisamment fine d'un espace abstrait (ou d'une portion intéressante de ce domaine) pour permettre de réaliser un diagnostic du système étudié, tout en restreignant la complexité des calculs à effectuer. La première s'inspire fortement de la méthode de pavage de base et vise à découper plus spécifiquement les pavés appartenant aux frontières du domaine à construire. Il est en effet complètement inutile de partitionner finement l'intérieur de $\mathcal{A}(\mathbf{p}_{\text{éval}}(k+s, \boldsymbol{\theta}))$ puisque la majoration induite par le pavage ne provient pas des pavés correspondants. En réalité, seules les frontières de ce domaine sont intéressantes puisqu'il faut tester si la forme de calcul se situe à l'intérieur. La seconde méthode proposée provient du constat que la construction complète de $\mathcal{A}(\mathbf{p}_{\text{éval}}(k+s, \boldsymbol{\theta}))$ n'est pas nécessaire dans le cadre du diagnostic. Seuls les pavés contenant $\mathbf{p}_{\text{cal}}(k+s)$, s'ils existent, sont intéressants, ce qui limite grandement les temps de calcul. De manière générale, ces deux méthodes sont itératives dans le sens où elles débutent sur l'hypercube unité \mathcal{H}_r définissant l'espace des paramètres $\mathcal{A}(\boldsymbol{\theta})$; ensuite, à chaque itération, selon le principe d'une procédure d'élimination, seuls les pavés jugés intéressants sont divisés, les autres étant éliminés.

IV.3.1 Méthode de pavage assurant la construction complète d'un espace abstrait.

Pour simplifier les notations, l'indice temporel est éliminé de la forme d'évaluation. La méthode de pavage de base, exposée dans le chapitre II, consiste à effectuer une sous-division uniforme de l'espace des paramètres $\mathcal{A}(\theta)$. Pour chaque pavé obtenu, l'orthotope aligné circonscrit image est calculé, leur réunion conduisant à une majoration $\mathcal{A}_p(p_{\text{éval}})$ du domaine abstrait recherché $\mathcal{A}(p_{\text{éval}})$. La procédure de détection de défauts consiste alors à tester au fur et à mesure de l'évaluation des pavés image si le vecteur résidu se situe à l'intérieur. Dans le cas affirmatif, le système est considéré comme sain et la procédure est stoppée (il n'est plus utile, en effet, de continuer à évaluer d'autres pavés). Un défaut sera détecté à la seule condition que tous les pavés image soient calculés et que p_{cal} ne se situe dans aucun d'entre eux. A une sous-division uniforme d'ordre q correspond q^r pavés image à déterminer, où r est la dimension du vecteur incertain θ . Or plus fin sera le pavage, c'est-à-dire plus grand sera le nombre q , meilleure sera la précision au détriment bien sûr du temps de calcul. Pour réduire ce dernier, il est possible de ne diviser que les pavés de $\mathcal{A}_p(p_{\text{éval}})$ qui composent ses frontières de manière à améliorer la précision ; nous les appellerons pavés extérieurs, à opposer aux pavés intérieurs.

Soit l'entier s_p indiquant le nombre de composantes de la forme d'évaluation considérée. Soit \mathbb{P}_i^4 un pavé de \mathbb{R}^r appartenant à l'espace des paramètres \mathcal{P}_r ; par définition, un pavé extérieur est un orthotope aligné et circonscrit image $\square\mathcal{A}(p_{\text{éval}}(\theta) / \theta \in \mathbb{P}_i)$ de \mathbb{R}^{s_p} dont au moins l'un des hyperplans frontière n'est pas entièrement contenu dans un autre pavé image. Ceci signifie qu'il existe au moins un point situé dans l'un de ses $2s_p$ hyperplans frontières qui ne soit pas situé dans un des autres pavés. L'algorithme repose ainsi sur deux listes de type F.I.F.O. permettant de stocker respectivement les pavés extérieurs et intérieurs. Chaque élément de ces listes contient les bornes des intervalles du domaine de départ \mathbb{P}_i ainsi que celles des domaines image $\square\mathcal{A}(p_{\text{éval}}(\theta) / \theta \in \mathbb{P}_i)$ pour éviter de les recalculer à chaque fois.

A ceci est ajoutée la notion de liaison entre pavés. Tout simplement, deux pavés image seront liés si leur intersection est non-vide. De cette manière, lorsqu'un parent est sous-divisé en deux enfants, déterminer si ces derniers sont extérieurs nécessite seulement de tester si leurs hyperplans frontières sont complètement inclus dans les pavés liés à leur unique parent. Pour limiter le nombre de calculs, chaque élément des listes précédentes contient donc les indices des pavés liés.

Pratiquement, le principe est le même que pour l'algorithme de pavage de base. Le premier élément de la liste des pavés extérieurs est sous-divisé en deux enfants puis supprimé. Chaque enfant est testé de manière à déterminer s'il est intérieur ou extérieur puis placé en fin de liste adéquate. Le travail s'opère principalement au niveau de la liste des pavés extérieurs, mais il est tout de même nécessaire de conserver les pavés intérieurs pour pouvoir déterminer si un enfant est extérieur ou non.

⁴ Par rapport à la section II.4.3, l'indice q au niveau des notations des pavés a été éliminé dans la mesure où les boîtes formant le pavage recherché pourront correspondre à des sous-divisions uniformes d'ordres différents.

À l'initialisation, la liste des pavés intérieurs \mathcal{L}_{int} est vide et celle des pavés extérieurs \mathcal{L}_{ext} contient les bornes de l'espace des paramètres \mathcal{P}_r et de son image $\square\mathcal{A}(p_{éval}(\theta) / \theta \in \mathcal{P}_r)$. Un entier cpt initialisé à 1 est utilisé comme compteur. L'algorithme est alors composé des étapes suivantes :

- (1) Prendre le premier élément de \mathcal{L}_{ext} et l'éliminer de cette liste. Le pavé du domaine de départ et le pavé image associé seront respectivement notés \mathbb{P}_1 et $\square\mathcal{A}(p_{éval}(\theta) / \theta \in \mathbb{P}_1)$.
- (2) Découper le pavé de départ \mathbb{P}_1 en deux pavés enfants $\mathbb{P}_{1,1}$ et $\mathbb{P}_{1,2}$ de manière à ce que $\mathbb{P}_1 = \mathbb{P}_{1,1} \cup \mathbb{P}_{1,2}$, puis calculer leurs images $\square\mathcal{A}(p_{éval}(\theta) / \theta \in \mathbb{P}_{1,i})$, $i \in \{1,2\}$.
- (3) Tester si chaque pavé image (pouvant appartenir aussi bien à \mathcal{L}_{ext} que \mathcal{L}_{int}) lié au parent l'est encore avec chacun de ses enfants.
- (4) Tester si les deux enfants $\square\mathcal{A}(p_{éval}(\theta) / \theta \in \mathbb{P}_{1,i})$, $i \in \{1,2\}$, sont liés.
- (5) Tester si les deux enfants $\square\mathcal{A}(p_{éval}(\theta) / \theta \in \mathbb{P}_{1,i})$, $i \in \{1,2\}$, sont extérieurs et les stocker dans la liste adéquate.
- (6) Si tous les pavés de la liste \mathcal{L}_{ext} correspondent à une sous-division uniforme de même ordre q , alors incrémenter la valeur du compteur cpt de 1.
- (7) Stopper l'algorithme si l'entier cpt dépasse un seuil spécifié par l'utilisateur, sinon retourner à l'étape 1.

Tableau IV-1. Squelette de l'algorithme proposé

Notons qu'il est primordial de modifier au fur et à mesure les deux listes \mathcal{L}_{ext} et \mathcal{L}_{int} :

- lors de la première étape, lorsque le père est éliminé, l'ensemble des éléments de \mathcal{L}_{ext} doivent être décalés d'une position, ce qui implique de remettre à jour les indices de pavé liés,
- à chaque fois qu'une liaison entre deux pavés image est examinée (étapes 3 et 4), il est nécessaire de réactualiser les listes si besoin est,
- lorsque les enfants sont ajoutés dans \mathcal{L}_{ext} ou \mathcal{L}_{int} (étape 5), il peut être nécessaire, suivant l'emplacement qui leur est affecté, de décaler certains éléments de la liste en question et donc de mettre à jour les indices de pavés liés.

Détaillons plus précisément ces différentes étapes. Lors de l'initialisation, est ajouté un indicateur, noté ind , valant 0 et permettant de connaître le nombre de découpages nécessaires pour obtenir \mathbb{P}_i à partir du domaine initial \mathcal{P}_r . Ainsi, l'entier ind associé à un enfant correspond à celui de son parent incrémenté de 1.

Lors de la seconde étape, il semble judicieux de découper alternativement selon chaque côté du pavé de manière à ce qu'aucune direction ne soit privilégiée (puisque la forme de l'espace abstrait recherché n'est pas connue a priori). Ceci peut être facilement réalisé à

l'aide de l'indicateur ind défini précédemment. Par exemple, la quantité $rest(\frac{ind}{r})+1$, où la fonction $rest$ désigne le reste de la division entière de son argument, indique alors l'indice de la composante intervalle du pavé \mathbb{P}_1 à découper en deux. Ensuite l'indicateur ind est incrémenté et stocké dans les enfants. Si le pavé parent \mathbb{P}_1 est défini par le vecteur intervalle suivant et que la quantité $rest(\frac{ind}{r})+1$ indique l'indice i :

$$\left[\left[\underline{p}_{1,1}, \bar{p}_{1,1} \right] \cdots \left[\underline{p}_{1,i}, \bar{p}_{1,i} \right] \cdots \left[\underline{p}_{1,p}, \bar{p}_{1,p} \right] \right]^T,$$

alors ses deux pavés enfants sont donnés par :

$$\left[\left[\underline{p}_{1,1}, \bar{p}_{1,1} \right] \cdots \left[\underline{p}_{1,i}, c_{1,i} \right] \cdots \left[\underline{p}_{1,p}, \bar{p}_{1,p} \right] \right]^T, \left[\left[\underline{p}_{1,1}, \bar{p}_{1,1} \right] \cdots \left[c_{1,i}, \bar{p}_{1,i} \right] \cdots \left[\underline{p}_{1,p}, \bar{p}_{1,p} \right] \right]^T$$

où $c_{1,i}$ est le centre défini par : $c_{1,i} = \frac{\bar{p}_{1,i} + \underline{p}_{1,i}}{2}$, et la valeur de l'indicateur passe pour ces deux pavés à $ind + 1$.

Puisque les dépendances entre les s_p composantes $p_{\text{éval},j}$ de $\mathbf{p}_{\text{éval}}$ ne sont pas prises en compte, $\square \mathcal{A}(\mathbf{p}_{\text{éval}}(\boldsymbol{\theta}) / \boldsymbol{\theta} \in \mathbb{P}_{1,i})$, $i \in \{1,2\}$, peut être facilement déterminé en évaluant séparément par l'analyse par intervalles, chaque domaine $\square \mathcal{A}(p_{\text{éval},j}(\boldsymbol{\theta}) / \boldsymbol{\theta} \in \mathbb{P}_{1,i})$, $j \in \{1, \dots, s_p\}$. De manière à éviter un résultat trop pessimiste si une variable bornée apparaît plusieurs fois dans l'expression d'une fonction d'inclusion $p_{\text{éval},j}$, des formes centrées peuvent être utilisées pour réduire de manière significative la majoration.

Soit l l'indice de l'un des pavés liés associés au parent \mathbb{P}_1 . Au cours de la troisième étape, pour tester si le pavé image de \mathbb{P}_1 est lié aux fils de \mathbb{P}_1 , il suffit de vérifier que les s_p intersections des intervalles $\square \mathcal{A}(p_{\text{éval},j}(\boldsymbol{\theta}) / \boldsymbol{\theta} \in \mathbb{P}_l)$ et $\square \mathcal{A}(p_{\text{éval},j}(\boldsymbol{\theta}) / \boldsymbol{\theta} \in \mathbb{P}_{1,i})$, $j \in \{1, \dots, s_p\}$ existent. Si au moins l'une de ces contraintes n'est pas vérifiée, il est alors certain que les pavés considérés n'ont aucun point commun, donc à fortiori, que des portions d'hyperplans frontières de $\square \mathcal{A}(\mathbf{p}_{\text{éval}}(\boldsymbol{\theta}) / \boldsymbol{\theta} \in \mathbb{P}_{1,i})$ ne peuvent être contenues dans le pavé $\square \mathcal{A}(\mathbf{p}_{\text{éval}}(\boldsymbol{\theta}) / \boldsymbol{\theta} \in \mathbb{P}_l)$.

La cinquième étape consiste à tester si les enfants sont extérieurs. Pour cela, chaque hyperplan frontière d'un enfant est étudié séparément, sachant qu'il en existe $2s_p$. Prenons par exemple le $j^{\text{ième}}$ et notons le \mathcal{H}_j^5 . L'objectif est de vérifier si \mathcal{H}_j est complètement contenu dans la réunion des pavés liés à l'enfant étudié. Pour cela, l'intersection entre chaque pavé lié et l'hyperplan frontière est calculée (si elle existe), puis la réunion des domaines obtenus est réalisée. Dans le cas où cette réunion ne vaut pas \mathcal{H}_j , c'est-à-dire si ces deux domaines sont représentés par des vecteurs intervalles ayant des bornes différentes, alors une partie de \mathcal{H}_j n'est pas comprise dans les autres pavés.

⁵ Remarquons que par le terme hyperplan frontière, il est en fait sous-entendu facette d'un pavé de taille finie, par conséquent \mathcal{H}_j est un sous-espace fermé de \mathbb{R}^{s_p-1} .

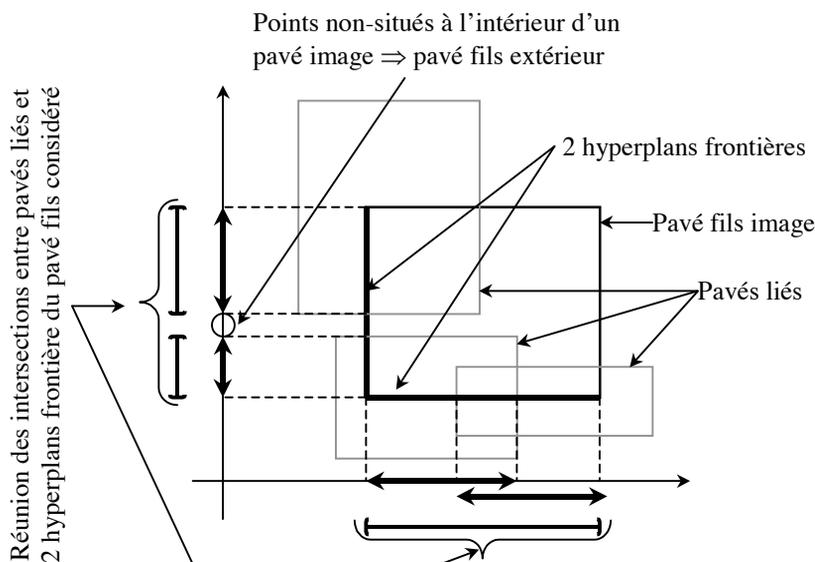


Figure IV-4. Détermination d'un pavé fils extérieur

Dans ce cas il est inutile d'étudier les hyperplans suivants puisque l'enfant considéré est assurément un pavé extérieur ; il est alors stocké dans la liste \mathcal{L}_{ext} , ainsi que toutes les informations le concernant. Dans le cas contraire, l'hyperplan suivant \mathcal{H}_{j+1} est étudié. Si tous les hyperplans frontières sont inclus dans la réunion des pavés liés, alors l'enfant est un pavé intérieur et est stocké à la fin de \mathcal{L}_{int} .

Remarquons cependant qu'il arrive qu'en divisant un parent, un pavé qui lui était lié et était déclaré intérieur devienne extérieur dans la mesure où la réunion des deux enfants est seulement incluse dans le parent.

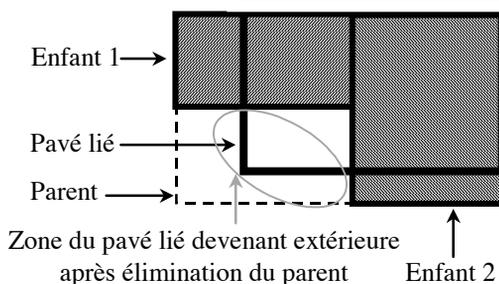


Figure IV-5. Problème engendré par l'élimination d'un parent

Sur ce schéma, une fois le parent éliminé, le pavé lié représenté devient un pavé extérieur. Pour résoudre ce problème, un moyen est d'ajouter au niveau de l'étape 5 une procédure testant sur le même principe que ce qui précède si chaque pavé intérieur lié au parent devient ou non extérieur après sous-division de ce dernier. L'intérêt de conserver et mettre à jour au fur et à mesure les liaisons entre pavés est alors évident, car sinon il serait nécessaire d'effectuer ce test sur tous les pavés image de la liste \mathcal{L}_{int} . Si un pavé lié devient extérieur, il est alors éliminé de \mathcal{L}_{int} et placé dans \mathcal{L}_{ext} . Sa position dépend de l'indicateur *ind* associé. Si son ordre de découpage est inférieur au dernier élément de \mathcal{L}_{ext} , le pavé est placé en début de liste de manière à être le prochain pavé sous-divisé, sinon il est placé en fin de liste. Il est bien sûr ensuite nécessaire de remettre à jour les indices de pavés liés. Du fait que des pavés intérieurs qui ont été découpés un nombre de fois quelconque puissent se retrouver dans la

liste \mathcal{L}_{ext} nécessite d'élaborer un test supplémentaire de manière à respecter une homogénéité du nombre de découpage associé à chaque pavé de cette liste. Ce test revient simplement à placer un enfant déclaré extérieur en début de \mathcal{L}_{ext} si son niveau de découpage indiqué par la variable *ind* est trop faible.

Au niveau de la sixième étape, tous les pavés image de la liste \mathcal{L}_{ext} correspondent à une sous-division uniforme de même ordre q si toutes les arrêtes des pavés antécédents de \mathbb{R}^r associés, ont été découpées un même nombre de fois, ce qui correspond à avoir tous les entiers *ind* égaux à *r.cpt*. Dans ces conditions, tous les pavés de \mathcal{L}_{ext} correspondent à une sous-division uniforme d'ordre $q=2^{cpt}$ avant incrémentation du compteur. En effet, en découpant alternativement selon chaque direction, l'algorithme procède par bisections successives :

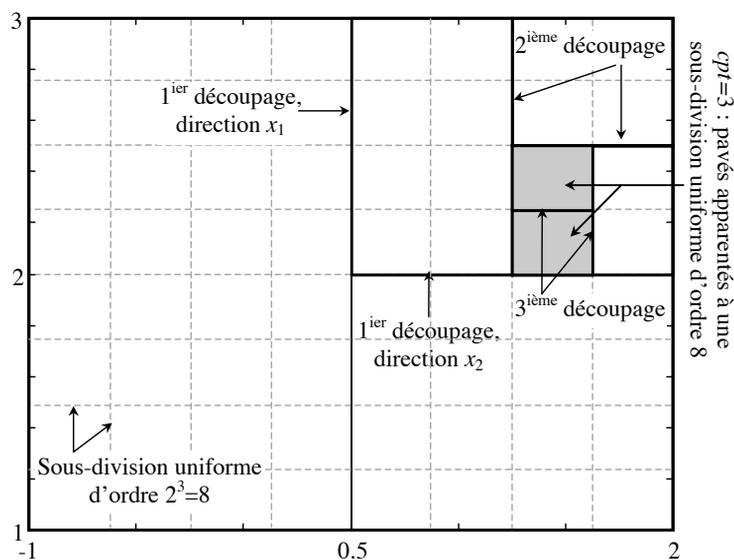


Figure IV-6. Sous-division uniforme et bisection

Exemple. Reprenons l'exemple du chapitre II présenté à la section II.4.2 concernant le champ de vecteurs \mathbf{h} .

$$\mathbf{h}(\theta_1, \theta_2) = \begin{bmatrix} 2\theta_1^2 + \theta_2 \\ \theta_1 - \theta_2 \end{bmatrix}, \text{ avec } \mathcal{A}(\theta_1) = [-1, 2], \mathcal{A}(\theta_2) = [1, 3].$$

Le domaine obtenu $\mathcal{A}_p(\mathbf{h})$ est alors le suivant, où les 198 pavés extérieurs sont grisés et où les 218 pavés intérieurs sont uniquement représentés par leurs arrêtes en pointillés. Le niveau de découpage choisi étant 6, les plus petits pavés correspondent à une sous-division uniforme d'ordre 32 qui aurait conduit, si elle avait été effectuée complètement à 1024 pavés.

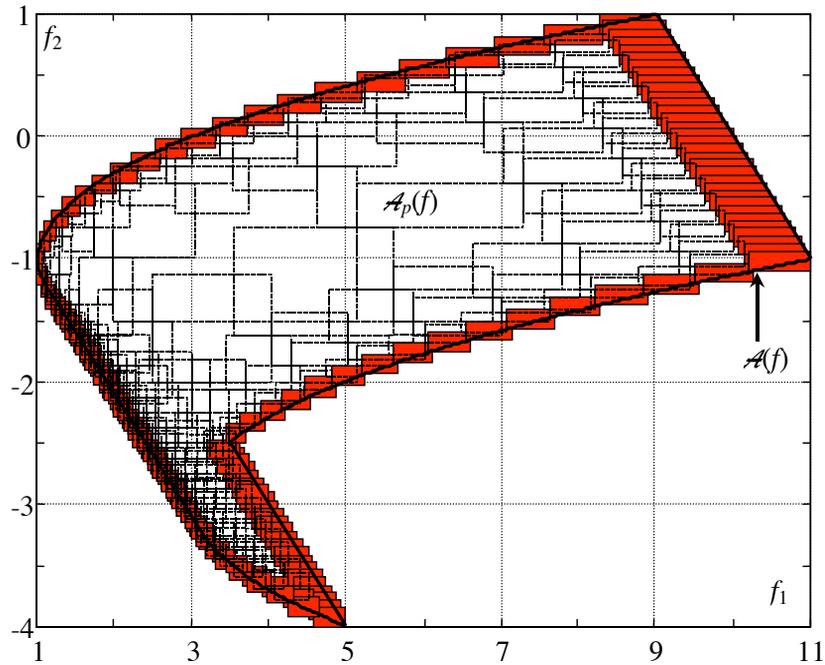


Figure IV-7. Pavage $\mathcal{A}_p(h)$ obtenu

Bien que le résultat obtenu soit exploitable dans le cas d'une procédure de détection de défauts en testant si la forme de calcul se situe dans l'un des pavés obtenus, pour des raisons de temps de calcul liés à la mise à jour des listes et des différents tests nécessaires pour établir si un pavé image est extérieur, l'algorithme proposé dans la section suivante est bien mieux adapté à une procédure de diagnostic.

IV.3.2 Méthode de pavage adaptée à une procédure de diagnostic.

Dans le cadre du diagnostic, l'objectif n'est pas de construire entièrement le pavage $\mathcal{A}_p(\mathbf{p}_{\text{éval}})$, mais seulement de vérifier, dans l'espace des résidus \mathbb{R}^{S_p} , si la forme de calcul fait partie de ce domaine. Par conséquent, seule la zone proche du vecteur résidu \mathbf{p}_{cal} est intéressante à étudier. Le principe de la méthode proposée consiste donc à ne s'intéresser qu'aux pavés \mathbb{P}_i appartenant à l'espace des paramètres \mathcal{H}_r dont les orthotopes alignés circonscrits image $\square \mathcal{A}(\mathbf{p}_{\text{éval}}(\boldsymbol{\theta}) / \boldsymbol{\theta} \in \mathbb{P}_i)$ de \mathbb{R}^{S_p} sont susceptibles de contenir le vecteur résidu.

Pour réaliser cette procédure, cette fois une seule liste \mathcal{L} de type F.I.F.O. est utilisée. Chacun de ses éléments est associé à un pavé \mathbb{P}_i contenu dans l'espace des paramètres \mathcal{H}_r et jugé intéressant. Cette liste est plus exactement constituée des valeurs des bornes inférieures et supérieures de chacune des variables bornées du vecteur incertain $\boldsymbol{\theta}$ imposées par \mathbb{P}_i . L'idée est, à chaque itération, de déterminer le pavé image $\square \mathcal{A}(\mathbf{p}_{\text{éval}}(\boldsymbol{\theta}) / \boldsymbol{\theta} \in \mathbb{P}_1)$ du premier élément noté \mathbb{P}_1 de la liste \mathcal{L} . Il faut ensuite tester si $\square \mathcal{A}(\mathbf{p}_{\text{éval}}(\boldsymbol{\theta}) / \boldsymbol{\theta} \in \mathbb{P}_1)$ contient la forme de calcul \mathbf{p}_{cal} . Si ce test est positif, \mathbb{P}_1 est découpé en deux pavés fils qui sont ajoutés en fin de liste \mathcal{L} . Si à un moment donné \mathcal{L} devient vide, alors un défaut est détecté puisque \mathbf{p}_{cal} ne peut appartenir au pavage, donc à $\mathcal{A}(\mathbf{p}_{\text{éval}})$.

À l'initialisation, la forme de calcul p_{cal} est évaluée et la liste \mathcal{L} se limite à un seul élément \mathbb{P}_1 correspondant à l'espace des paramètres \mathcal{Z}_r . Un compteur cpt est utilisé et fixé à 1. L'algorithme est alors composé des étapes suivantes :

- (1) Arrêter l'algorithme si la liste \mathcal{L} est vide, sinon allez à l'étape 2.
- (2) Prendre le premier élément ℓ_1 de \mathcal{L} et l'éliminer de cette liste.
- (3) En notant \mathbb{P}_1 le pavé contenu dans ℓ_1 , évaluer son image $\square_{\mathcal{A}}(p_{éval}(\theta) / \theta \in \mathbb{P}_1)$.
- (4) Tester si la forme de calcul p_{cal} appartient à $\square_{\mathcal{A}}(p_{éval}(\theta) / \theta \in \mathbb{P}_1)$. Si ce test est négatif, retourner directement à l'étape 1.
- (5) Découper le pavé de départ \mathbb{P}_1 en deux pavés enfants $\mathbb{P}_{1,1}$ et $\mathbb{P}_{1,2}$ de manière à ce que $\mathbb{P}_1 = \mathbb{P}_{1,1} \cup \mathbb{P}_{1,2}$, puis les stocker à la fin de la liste \mathcal{L} .
- (6) Si tous les pavés de la liste \mathcal{L} correspondent à une sous-division uniforme de même ordre q , alors incrémenter le compteur cpt de 1.
- (7) Arrêter l'algorithme si le compteur cpt dépasse un seuil spécifié par l'utilisateur, sinon retourner à l'étape 1.

Tableau IV-2. Squelette de l'algorithme proposé

Deux situations peuvent alors se présenter. Si la liste devient vide, alors ceci signifie qu'aucun pavé image ne peut contenir le vecteur résidu, ce qui engendre la détection du défaut. Si le système fonctionne normalement, c'est-à-dire si le vecteur résidu se situe réellement à l'intérieur du pavage, la procédure telle qu'elle est présentée découperait alors indéfiniment sans la présence de la dernière étape. Il est donc logique d'ajouter un test d'arrêt limitant le nombre de découpages en testant si le compteur cpt dépasse une certaine valeur (il est aussi possible d'arrêter l'algorithme lorsque les rayons des pavés calculés sont en deçà d'une certaine valeur). Notons que si le test associé à l'étape 7 est vérifié, alors le système est supposé fonctionner correctement. La précision de cet algorithme dépend évidemment du seuil choisi. Plus important est le seuil, plus petite sera la majoration liée au pavage, mais plus important sera le temps de calcul.

Détaillons un peu plus précisément cet algorithme. Au niveau de la troisième étape, évaluer l'image $\square_{\mathcal{A}}(p_{éval}(\theta) / \theta \in \mathbb{P}_1)$ revient à calculer individuellement chacune de ses s_p composantes $\square_{\mathcal{A}}(p_{éval,i}(\theta) / \theta \in \mathbb{P}_1)$. La même remarque que pour l'algorithme précédent (étape 2) peut être faite quant à l'utilisation des techniques à base de formes centrées ou de la monotonie des fonctions étudiées en cas de variables bornées à occurrence multiple. Dans le même esprit, la quatrième étape consiste tout simplement à vérifier si chacune des s_p coordonnées $p_{cal,i}$ de p_{cal} appartient bien à l'intervalle $\square_{\mathcal{A}}(p_{éval,i}(\theta) / \theta \in \mathbb{P}_1)$.

De même, lors de l'initialisation, est ajouté un indicateur, noté ind , valant 0 et jouant exactement le même rôle. Ainsi, lors de la cinquième étape, le pavé \mathbb{P}_1 est découpé alternativement selon chaque côté du pavé de manière à ce qu'aucune direction de l'espace paramétrique ne soit privilégiée. Ceci est réalisé comme explicité au niveau de l'étape 2 de la section IV.3.1 précédente.

À partir de l'instant où le modèle est juste et où à un moment donné la liste est vide, il est certain qu'un défaut est bien présent. Dans le cas où la liste est non-vide, lorsque le test d'arrêt stoppe la procédure, il est alors supposé que le système se comporte normalement, même si la présence d'un défaut est possible. Ce dernier peut être masqué par les incertitudes affectant les paramètres témoignant du manque d'information que renferme le modèle ; mais ce problème est directement lié à la précision du modèle et serait identique au cas où il serait possible de construire le domaine abstrait $\mathcal{A}(p_{\text{éval}})$ lui-même. En revanche, une autre source de non-détection réside dans la majoration induite par la méthode de pavage. Cependant, celle-ci diminue en augmentant le nombre de découpages possibles ou en diminuant la taille limite des pavés, ce dernier réglage étant alors nettement plus parlant. Chaque pavé image appartient nécessairement totalement ou en partie au domaine abstrait recherché, puisque $\square\mathcal{A}(p_{\text{éval}}(\theta)/\theta \in \mathbb{P}_1)$ contient et est circonscrit au domaine abstrait $\mathcal{A}(p_{\text{éval}}(\theta)/\theta \in \mathbb{P}_1)$, lui-même contenu dans $\mathcal{A}(p_{\text{éval}})$. Prenons le cas où s_p vaut 2 de manière à facilement représenter les domaines, mais la discussion qui suit se généralise quelle que soit la dimension du problème traité. Chaque domaine $\square\mathcal{A}(p_{\text{éval}}(\theta)/\theta \in \mathbb{P}_1)$ est alors un rectangle avec au moins un point de chaque arête appartenant à l'espace abstrait $\mathcal{A}(p_{\text{éval}}(\theta)/\theta \in \mathbb{P}_1)$ ⁶. Ponctuellement pour un pavé, les majorations les plus importantes correspondent aux cas où la frontière de l'espace abstrait passe par deux sommets opposés. Une non-détection liée uniquement à la majoration induite par le pavage correspond au cas où la forme de calcul p_{cal} se situe dans $\square\mathcal{A}(p_{\text{éval}}(\theta)/\theta \in \mathbb{P}_1)$, mais à l'extérieur de $\mathcal{A}(p_{\text{éval}}(\theta)/\theta \in \mathbb{P}_1)$.

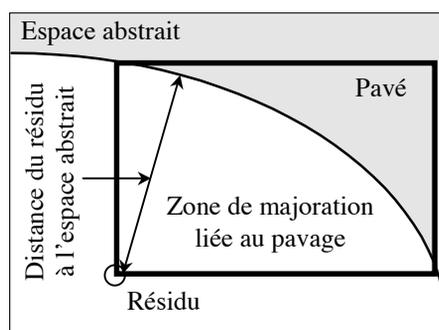


Figure IV-8. Majoration liée à la méthode de pavage

Même dans le cas extrême représenté sur la figure précédente, puisque p_{cal} se retrouve nécessairement dans $\square\mathcal{A}(p_{\text{éval}}(\theta)/\theta \in \mathbb{P}_1)$, la taille ou le volume du pavé donne alors une bonne appréciation de la majoration maximale possible, c'est-à-dire de la précision de l'algorithme, qu'il est alors possible de régler de manière naturelle grâce au test d'arrêt en fixant par exemple une taille maximale que ne doivent pas dépasser les pavés image.

Exemple. Reprenons l'exemple précédent associé au champ de vecteurs h . L'objectif est de tester si le point de coordonnées (4,9,0,4) se situe ou non en dehors de l'espace abstrait $\mathcal{A}(h)$. L'algorithme trouve que ce point n'appartient pas à ce domaine après le calcul de 39 orthotopes (dont 19 au total contenaient le point en question avant division). Les plus petits pavés traités sont obtenus après 8 bisections successives de chacune des facettes et

⁶ Sous la condition que l'orthotope aligné circonscrit puisse être calculé exactement malgré la présence de variables dépendantes au sein d'une même composante de la forme d'évaluation.

correspondent à une sous-division uniforme de l'espace des paramètres d'ordre $2^8=256$, ce qui serait équivalent à l'aide de la méthode de pavage de base (puisqu'alors tous les pavés image devraient être construits), à en calculer $2^{16}=65536$. La figure IV-9 de gauche montre l'espace abstrait théorique $\mathcal{A}(h)$ ainsi que le point de coordonnées (4.9,0.4) qui se situe très légèrement à l'extérieur. De manière à apprécier plus facilement la distance séparant ce point de $\mathcal{A}(h)$, la figure de droite, correspondant à un zoom de la portion intéressante de celle de gauche, représente le dernier pavé obtenu avant que la liste ne se vide, superposé au point étudié. Dans le cadre de cet exemple, un faible découpage (puisque seulement 39 pavés ont été évalués) permet de conclure alors que le point de coordonnées (4.9,0.4) se trouve très proche de la frontière théorique, en effet, le pavé représenté a une longueur de $\max(0.041,0.014)$, ce qui est ridicule par rapport à la taille globale du domaine.

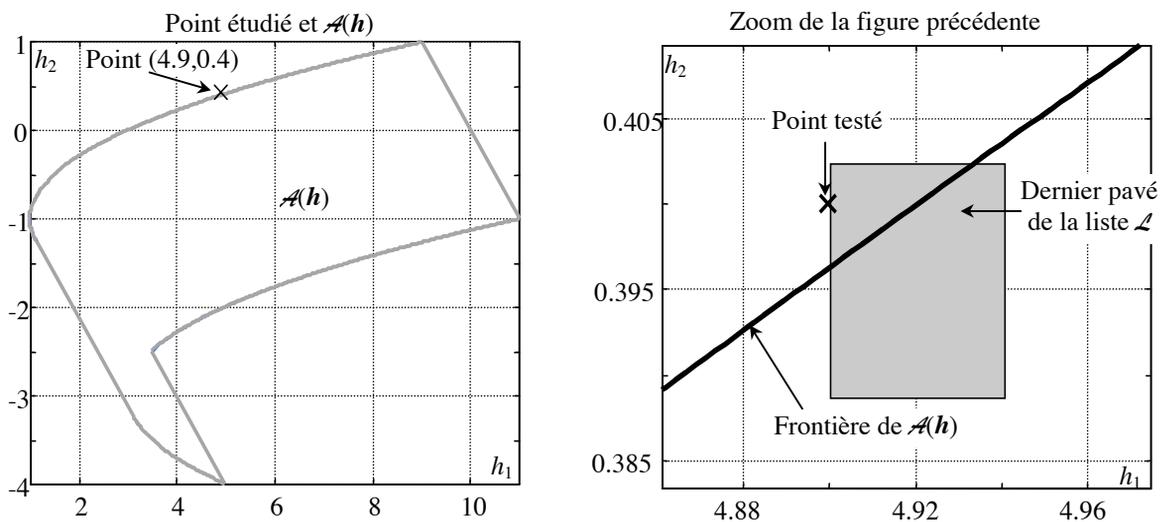


Figure IV-9. Espace abstrait $\mathcal{A}(h)$ et point (4.9,0.4) testé

La figure suivante montre comment est réalisé le découpage de l'espace des paramètres.

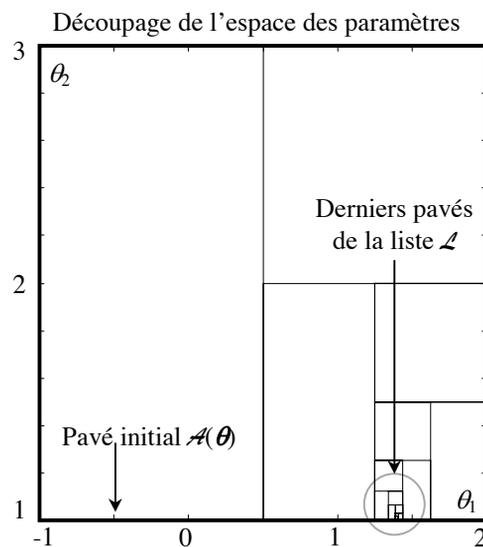


Figure IV-10. Découpage du domaine $\mathcal{A}(\theta)$

Toujours en considérant le même champ de vecteurs \mathbf{h} , testons si le point $(2,-2)$ appartient à $\mathcal{A}(\mathbf{h})$. La figure suivante présente la partition de l'espace paramétrique $\mathcal{A}(\boldsymbol{\theta})$ obtenue avec l'algorithme proposé.

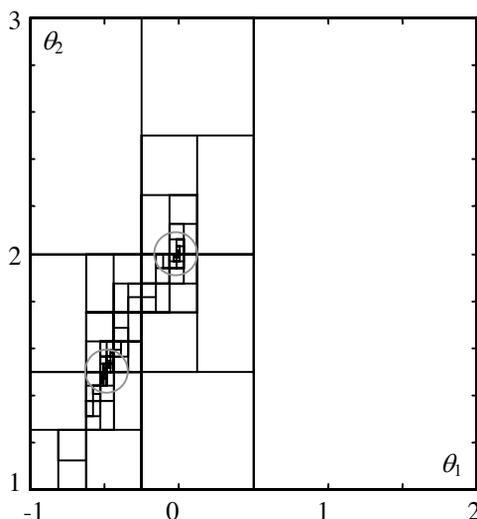


Fig. 1. Découpage du domaine $\mathcal{A}(\boldsymbol{\theta})$

Elle montre deux zones explorées autour des solutions exactes $(-0.5, 1.5)$ et $(0, 2)$ de $\mathcal{A}(\boldsymbol{\theta})$ conduisant à l'image recherchée $(2, -2)$; par conséquent, le point considéré se trouve bien à l'intérieur de $\mathcal{A}(\mathbf{h})$. Le seuil choisi pour la valeur maximale du compteur cpt est 10. Les plus petits pavés traités sont obtenus après 9 bisections successives de chacune des facettes et correspondent à une sous-division uniforme de l'espace des paramètres d'ordre $2^9=512$, ce qui serait équivalent à l'aide de la méthode de pavage de base à en calculer 2^{18} . Le nombre total de pavés évalués par l'algorithme est de 261, ce qui limite considérablement les temps de calcul. ■

IV.4 EXPRESSION ANALYTIQUE DES FRONTIÈRES D'UN ESPACE ABSTRAIT

Une autre solution que la méthode de pavage consiste à rechercher l'expression analytique des frontières d'un espace abstrait. Moyennant certaines hypothèses à respecter, il est alors possible de déterminer de manière exacte ce domaine.

Considérons une fonction vectorielle f ; son espace abstrait se limite à un intervalle (pour peu que l'ensemble image soit fermé et borné). L'analyse par intervalles conduit à un résultat exact s'il n'y a aucune variable bornée à occurrence multiple, à une majoration sinon, dont la précision va dépendre de la méthode utilisée pour le calculer (formes centrée, arithmétique affine, méthodes de réduction ou d'élimination...). Dans le cas d'un champ de vecteurs \mathbf{f} , le problème de majoration engendré par des variables dépendantes au sein d'une même composante reste d'actualité et peut être traité par les mêmes techniques qu'en

monodimensionnel. S'y ajoutent les problèmes de dépendance entre composantes liées à la présence de variables communes. Seul ce dernier problème est pris en considération dans cette section.

L'idée est que, puisque le phénomène de majoration est uniquement dû à la présence de variables communes, l'interaction de ces dernières sur l'espace abstrait $\mathcal{A}(f)$ doit être étudiée. Dans ce qui suit, nous rappelons une méthode permettant de déterminer l'expression analytique des frontières d'un espace abstrait lorsque le système est soumis à certaines contraintes structurelles que nous expliciterons. Ce travail a été développé par Ploix (1998), et si les résultats qui en découlent peuvent être utilisés dans le cadre d'une procédure de détection de défauts, nous les utiliserons principalement dans ce document en vue de réaliser une procédure de localisation. Cette méthode n'est ici que brièvement décrite pour n'en présenter que les principaux résultats, le lecteur pourra donc avantageusement se référer à (Ploix, 1998) dans laquelle celle-ci est bien plus détaillée.

Pour débiter, présentons un exemple simple afin d'explicitier le principe de la méthode.

Exemple. Reprenons le champ de vecteurs suivant correspondant à l'exemple de la section II.4.1 :

$$f(\theta_1, \theta_2) = \begin{bmatrix} \theta_1 - \theta_2 \\ \theta_1 + \theta_2 \end{bmatrix}, \text{ où } \mathcal{A}(\theta_1) = [-1, 2] \text{ et } \mathcal{A}(\theta_2) = [1, 3].$$

L'orthotope aligné circonscrit $\square\mathcal{A}(f)$, calculé à l'aide de l'arithmétique des intervalles, conduit dans l'espace des composantes f_1 et f_2 au pavé $[-4, 1] \times [0, 5]^T$ représenté par le cadre de la figure IV-11. Le domaine abstrait $\mathcal{A}(f)$, quant à lui, a la forme d'un rectangle :

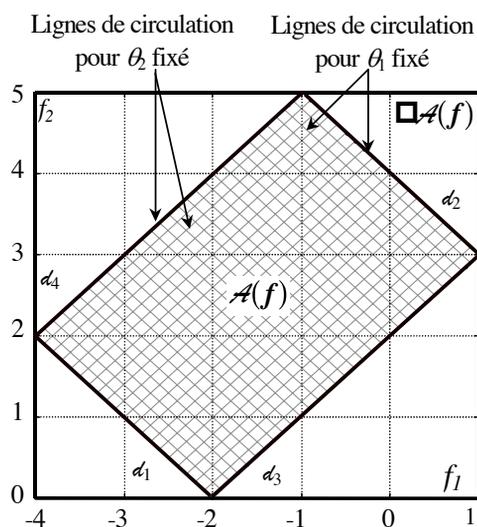


Figure IV-11. Espace abstrait $\mathcal{A}(f)$ et lignes de circulation

La présence des variables communes θ_1 et θ_2 engendre des dépendances entre les deux composantes, puisque chacune de ces deux variables prend une valeur inconnue identique dans f_1 et f_2 . Dans le cadre de cet exemple, en fixant l'une des variables communes θ_i successivement à des valeurs également réparties dans l'intervalle $\mathcal{A}(\theta_i)$, puis en faisant

varier continûment la seconde sur son espace abstrait, les lignes de circulation tracées en gris au niveau la figure précédente sont déterminées. Grâce à la propriété de dissociabilité du système étudié que nous développerons ultérieurement, certaines de ces lignes de circulation définissent de manière exacte les frontières du domaine recherché. Or le travail ici effectué de manière purement numérique, consiste dans un cadre analytique à déterminer des combinaisons des composantes f_1 et f_2 conduisant à des relations supplémentaires dans lesquelles une des variables communes est éliminée.

En éliminant les variables communes θ_1 et θ_2 de manière à ce qu'elles n'apparaissent plus simultanément dans les nouvelles relations, le système suivant est obtenu :

$$\begin{cases} f_2 = -f_1 + 2\theta_1 \\ f_2 = f_1 + 2\theta_2 \end{cases}$$

En utilisant l'analyse par intervalles et en prenant en compte les bornes des intervalles $\mathcal{A}(\theta_1)$ et $\mathcal{A}(\theta_2)$, l'espace abstrait recherché est défini par l'intersection de l'orthotope aligné circonscrit $\square\mathcal{A}(f)$ avec les demi-plans donnés par :

$$\begin{cases} f_2 \geq -f_1 - 2 \\ f_2 \leq -f_1 + 4 \\ f_2 \geq f_1 + 2 \\ f_2 \leq f_1 + 6 \end{cases}$$

Les droites frontières sont donc obtenues en remplaçant les inégalités par des égalités :

$$\begin{cases} d_1 & f_2 = -f_1 - 2 \\ d_2 & f_2 = -f_1 + 4 \\ d_3 & f_2 = f_1 + 2 \\ d_4 & f_2 = f_1 + 6 \end{cases}$$

■

La méthode proposée permet, sous certaines conditions, de trouver l'expression analytique des différentes frontières d'un espace abstrait, même si le champ f est non-linéaire en les variables bornées. Elle n'est cependant pas systématique car, apparentée à la théorie de l'élimination (Becker et al, 1993), il est souvent difficile d'arriver à faire disparaître certaines variables. De plus, un nombre important de variables communes est prohibitif.

Puisque ce sont les variables communes qui engendrent une déformation de la forme de l'espace abstrait par rapport à son orthotope aligné circonscrit, il est donc nécessaire d'étudier leur influence pour pouvoir atteindre notre objectif. Soit $f(\theta)$ un champ de vecteurs de dimension s_p et θ un vecteur de variables bornées normalisées $\theta = [\theta_1, \dots, \theta_n]^T$. L'ensemble \mathcal{F} contient tous les indices j faisant référence à chacune des composantes $f_j(\theta)$ de f : $\mathcal{F} = \{1, \dots, s_p\}$. De même, \mathcal{F}_i est un sous-ensemble de \mathcal{F} ne faisant référence qu'à certaines fonctions $f_j(\theta)$ et le champ de vecteurs associé ne contenant que ces dernières est noté $f^{\mathcal{F}_i}(\theta^{\mathcal{F}_i})$. Le vecteur $\theta^{\mathcal{F}_i}$ correspond à l'ensemble des variables communes aux composantes du champ de vecteurs $f^{\mathcal{F}_i}$, alors que $\theta^{\mathcal{F}_{nc,i}}$ dénote le vecteur regroupant les variables bornées restantes, c'est-à-dire

non-communes qui n'apparaissent que dans une seule composante de $f^{\mathcal{Z}_i}$.

Définition. Un sous-ensemble résoluble est un ensemble d'équations qu'il est possible de combiner de manière à pouvoir éliminer totalement n'importe laquelle des variables communes, élément de $\theta^{\mathcal{Z}_{c,i}}$.

Ploix (1998) a montré qu'un sous-ensemble \mathcal{Z}_i est résoluble si $\text{card}(\mathcal{Z}_i) \geq 2$ et si le nombre de variables communes aux composantes de $f^{\mathcal{Z}_i}$ est tel que $\text{card}(\theta^{\mathcal{Z}_{c,i}}) \geq \text{card}(\mathcal{Z}_i) - 1$. Au sein d'un système résoluble, il existe parfois plusieurs éliminations possibles, il faudra privilégier celles conduisant aux relations les plus simples.

Lorsque $\text{card}(\theta^{\mathcal{Z}_{c,i}}) = \text{card}(\mathcal{Z}_i) - 1$, il n'existe qu'une seule équation cartésienne solution de cette élimination (c'est-à-dire pour laquelle n'apparaît plus aucune variable commune). Lorsque $\text{card}(\theta^{\mathcal{Z}_{c,i}}) > \text{card}(\mathcal{Z}_i) - 1$, il existe $\mathbb{C}^{\frac{\text{card}(\mathcal{Z}_i)-1}{\text{card}(\theta^{\mathcal{Z}_{c,i}})}}$ équations cartésiennes solution déduites des $\mathbb{C}^{\frac{\text{card}(\mathcal{Z}_i)-1}{\text{card}(\theta^{\mathcal{Z}_{c,i}})}}$ combinaisons possibles des $\text{card}(\mathcal{Z}_i)$ composantes de $f^{\mathcal{Z}_i}$ assurant à chaque fois l'élimination de $\text{card}(\mathcal{Z}_i) - 1$ variables de $\theta^{\mathcal{Z}_{c,i}}$. Ces équations contiennent les variables non-communes de $\theta^{\mathcal{Z}_{nc,i}}$ et ne font pas intervenir les $\text{card}(\mathcal{Z}_i) - 1$ variables communes de $\theta^{\mathcal{Z}_{c,i}}$ qui ont pu être éliminées.

Les frontières d'un domaine abstrait, à l'exception de celles coïncidant avec celles de l'orthotope aligné circonscrit, sont dues aux variables communes.

Définition. Un champ de vecteurs $f^{\mathcal{Z}_i}(\theta^{\mathcal{Z}_i})$ est dissociable si chacune de ses dérivées partielles par rapport aux variables communes $\theta_j^{\mathcal{Z}_{c,i}}$ de $\theta^{\mathcal{Z}_{c,i}}$ est indépendante des autres variables communes $\theta_k^{\mathcal{Z}_{c,i}}$, $k \neq j$:

$$\frac{\partial f^{\mathcal{Z}_i}(\theta^{\mathcal{Z}_i})}{\partial \theta_j^{\mathcal{Z}_{c,i}}} = \frac{\partial f^{\mathcal{Z}_i}(\theta^{\mathcal{Z}_{c,i}}, \theta^{\mathcal{Z}_{nc,i}})}{\partial \theta_j^{\mathcal{Z}_{c,i}}} = g_j(\theta_j^{\mathcal{Z}_{c,i}}, \theta^{\mathcal{Z}_{nc,i}}).$$

L'interprétation géométrique d'un champ dissociable est la suivante. Les domaines obtenus après élimination des variables communes (c'est-à-dire associés aux $\mathbb{C}^{\frac{\text{card}(\mathcal{Z}_i)-1}{\text{card}(\theta^{\mathcal{Z}_{c,i}})}}$ équations cartésiennes précédemment citées), coïncident avec les frontières de l'espace abstrait recherché lorsque le champ de vecteurs considéré est dissociable. Ces frontières correspondent en fait à des lignes de circulation⁷ et ce sont les dépendances entre variables communes qui font que ces premières ne sont pas réduites à de simples hyperplans.

Exemple. Reprenons l'exemple précédent. Le système est résoluble puisque $\text{card}(\theta^{\mathcal{Z}_{c,i}}) = 2$ et $\text{card}(\mathcal{Z}_i) - 1 = 1$. Il est aussi dissociable puisque les dérivées partielles sont

⁷ Une variable commune varie sur tout son domaine alors que toutes les autres sont fixées.

constantes. Le seul sous-ensemble résoluble est le système lui-même avec $\mathcal{F}=\{1,2\}$, de plus il possède deux variables communes θ_1 et θ_2 , ce qui conduit à $\mathbb{C}_{\text{card}(\theta^{\mathcal{F}_i})}^{\text{card}(\mathcal{F}_i)-1} = \mathbb{C}_2^1 = 2$ équations cartésiennes dans lesquelles $\text{card}(\mathcal{F}_i) - 1 = 1$ variable commune aura été éliminée. ■

La méthode consiste donc à déterminer dans un premier temps si le champ de vecteurs f est dissociable en calculant les différentes dérivées partielles de ses composantes par rapport aux variables communes. Puis, il faut trouver tous les ensembles résolubles \mathcal{F}_i de \mathcal{F} . Pour chacun d'entre eux sont calculées les $\mathbb{C}_{\text{card}(\theta^{\mathcal{F}_i})}^{\text{card}(\mathcal{F}_i)-1}$ équations cartésiennes solution de l'élimination de $\text{card}(\mathcal{F}_i) - 1$ variables communes, donnant les expressions analytiques d'un même nombre de domaines faisant partie des frontières du domaine abstrait recherché. L'intersection de tous les domaines trouvés pour chaque sous-ensemble \mathcal{F}_i avec l'orthotope aligné circonscrit $\square \mathcal{A}(f)$ conduit à $\mathcal{A}(f)$. Les expressions des frontières sont obtenues en évaluant les expressions cartésiennes à l'aide de l'analyse par intervalles et des propriétés de monotonie de fonctions.

Après élimination, des variables communes de $\theta^{\mathcal{F}_i}$ peuvent encore apparaître dans plusieurs équations cartésiennes à la fois. De même, des variables de $\theta^{\mathcal{F}_{nc,i}}$ peuvent faire de même sans que cela ne gêne en rien.

Si le champ de vecteurs est affine en les variables communes, c'est-à-dire qu'elles apparaissent de manière linéaire dans les composantes de f :

$$\frac{\partial \mathcal{F}_i(\theta^{\mathcal{F}_i})}{\partial \theta_j^{\mathcal{F}_{c,i}}} = \frac{\partial \mathcal{F}_i(\theta^{\mathcal{F}_{c,i}}, \theta^{\mathcal{F}_{nc,i}})}{\partial \theta_j^{\mathcal{F}_{c,i}}} = g_j(\theta^{\mathcal{F}_{nc,i}}),$$

alors les lignes de circulation, et par conséquent les frontières de l'espace abstrait recherché (ce champ étant par définition dissociable), sont des hyperplans puisque les nouvelles équations cartésiennes sont nécessairement linéaires en les variables bornées. Le domaine recherché est alors un paralléloèdre, polytope⁸ particulier dont les faces sont des hyperplans deux à deux parallèles⁹, et ce même si le champ est non-linéaire vis-à-vis des variables non communes.

Exemple. Reprenons l'exemple de la section précédente concernant la fonction h , non-linéaire par rapport à une des variables communes. Il est simple de vérifier que ce champ de vecteurs est dissociable. Du fait de la non-linéarité sur la variable commune θ_1 , le domaine sera non-linéaire et a priori non-convexe. Après combinaison pour éliminer respectivement θ_2 et θ_1 , le nouveau système est donné par :

⁸ Un polytope est un domaine délimité par des hyperplans. Certains auteurs tel [Ziegler \(1998\)](#) ajoutent à cette définition l'hypothèse supplémentaire de convexité d'un tel domaine.

⁹ Ce domaine n'est pas forcément convexe.

$$\begin{cases} h_2 = -h_1 + 2\theta_1^2 + \theta_1 \\ h_2 = \pm \sqrt{\frac{h_1 - \theta_2}{2}} - \theta_2 \end{cases}, \text{ avec la condition } h_1 \geq \theta_2.$$

La première relation contient une variable à occurrence multiple : θ_1 . En utilisant la dérivée partielle $\frac{\partial h_2}{\partial \theta_1} = 4\theta_1 + 1$, se déduisent les résultats suivants :

Signe de $\frac{\partial h_2}{\partial \theta_1}$	Conditions sur θ_1	Extremum de h_2	Valeur de θ_1	Frontière obtenue
≥ 0	$-\frac{1}{4} \leq \theta_1 \leq 2$	Minimum	$-\frac{1}{4}$	$h_2 \geq -h_1 - \frac{1}{8}$
		Maximum	2	$h_2 \leq -h_1 + 10$
≤ 0	$-1 \leq \theta_1 \leq -\frac{1}{4}$	Minimum	$-\frac{1}{4}$	$h_2 \geq -h_1 - \frac{1}{8}$
		Maximum	-1	$h_2 \leq -h_1 + 1$

En ce qui concerne la seconde relation, la variable θ_2 est elle aussi à occurrence multiple. Pour le premier cas :

$$h_2 = \sqrt{\frac{h_1 - \theta_2}{2}} - \theta_2,$$

en constatant que la dérivée partielle $\frac{\partial h_2}{\partial \theta_2} = \frac{-\sqrt{2}}{4\sqrt{h_1 - \theta_2}} - 1$ est toujours négative, se déduit que f_2 satisfait les contraintes suivantes :

$$\begin{cases} h_2 \geq \sqrt{\frac{h_1 - 3}{2}} - 3, h_1 \geq 3 \\ h_2 \leq \sqrt{\frac{h_1 - 1}{2}} - 1, h_1 \geq 1 \end{cases}.$$

Pour le second cas :

$$h_2 = -\sqrt{\frac{h_1 - \theta_2}{2}} - \theta_2,$$

la dérivée partielle $\frac{\partial h_2}{\partial \theta_2} = \frac{\sqrt{2}}{4\sqrt{h_1 - \theta_2}} - 1$ n'est pas de signe constant :

Signe de $\frac{\partial h_2}{\partial \theta_2}$	Conditions sur h_1	Extremum de h_2	Valeur de θ_2	Frontière obtenue
≥ 0	$\theta_2 \leq f_1 \leq \frac{1}{8} + \theta_2$	Minimum	1	$h_2 \geq -\sqrt{\frac{h_1 - 1}{2}} - 1, 1 \leq h_1 \leq 9/8$
		Maximum	3	$h_2 \leq -\sqrt{\frac{h_1 - 3}{2}} - 3, 3 \leq h_1 \leq 25/8$
≤ 0	$\frac{1}{8} + \theta_2 \leq f_1$	Minimum	3	$h_2 \geq -\sqrt{\frac{h_1 - 3}{2}} - 3, 25/8 \leq h_1$
		Maximum	1	$h_2 \leq -\sqrt{\frac{h_1 - 1}{2}} - 1, 9/8 \leq h_1$

En résumé, les frontières sont données par :

$$\left\{ \begin{array}{l} d_1 \quad h_2 = -h_1 - \frac{1}{8} \\ d_2 \quad h_2 = -h_1 + 10 \\ d_3 \quad h_2 = -h_1 + 1 \\ e_1 \quad h_2 = \sqrt{\frac{h_1 - 3}{2}} - 3, \quad h_1 \geq 3 \\ e_2 \quad h_2 = \sqrt{\frac{h_1 - 1}{2}} - 1, \quad h_1 \geq 1 \\ e_3 \quad h_2 = -\sqrt{\frac{h_1 - 1}{2}} - 1, \quad 1 \leq h_1 \\ e_4 \quad h_2 = -\sqrt{\frac{h_1 - 3}{2}} - 3, \quad 3 \leq h_1 \end{array} \right. ,$$

et sont représentées sur la figure suivante¹⁰ :

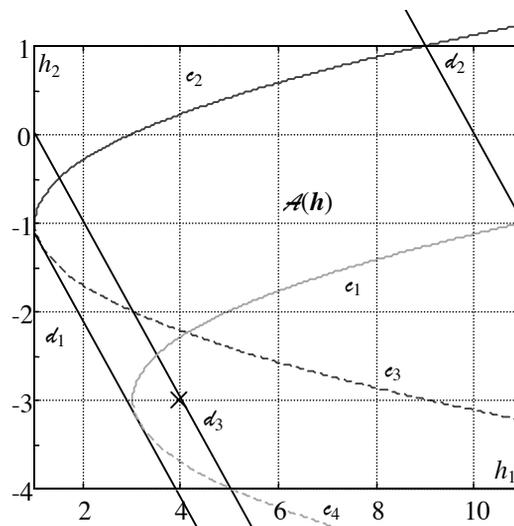


Figure IV-12. Frontières de l'espace abstrait $\mathcal{A}(f)$

Notons ici le nombre important de portions de frontières obtenues pour un champ de seulement deux fonctions et deux variables bornées. ■

Les inconvénients de cette méthode sont que les formes d'évaluation (quand elles sont non-linéaires en les variables bornées) ne sont pas forcément des champs de vecteurs dissociables ; de plus, il est difficile dans le cas de domaines complexes (tels que non-convexes) de parvenir à déterminer si un point se situe à l'intérieur. En effet, prenons le point $(4, -3)$ (représenté par une croix sur la figure précédente), alors que celui-ci est bien à l'intérieur de $\mathcal{A}(f)$, les deux contraintes inégalités associées à e_1 et e_3 ne sont pas respectées. Si visuellement le test est simple à faire, le réaliser numériquement est bien plus complexe de manière générale. Nous n'utiliserons donc pas cette stratégie en vue de la procédure de détection. En revanche, l'exposé de cette méthode va permettre, dans une certaine mesure, d'explicitier la suivante (section IV.5), certes restreinte aux cas de champs de vecteurs linéaires en les variables bornées, mais en revanche purement numérique. D'autre part, la

¹⁰ Sachant que pour une meilleure visualisation de l'ensemble, les contraintes sur la composante f_1 , concernant les courbes e_i , ne sont pas forcément respectées, par exemple e_3 est tracée pour des valeurs de f_1 supérieures à 1.

détermination des expressions analytiques des frontières permettra de mettre en œuvre la phase de localisation présentée dans la section IV.5.5.

IV.5 MÉTHODE PROCÉDANT DE LA LINÉARISATION DES RELATIONS DE PARITÉ

Comme détaillé dans le chapitre III relatif à la construction de relations de parité dans le contexte d'un modèle incertain, les formes d'évaluation sont souvent non-linéaires en les variables bornées, que ce soit pour une matrice de parité certaine ou incertaine. L'objectif de cette section est de détailler une série de procédures successives permettant de mettre en œuvre le module de diagnostic dans le cadre de l'approche bornante sans passer par une méthode de pavage coûteuse en temps de calcul.

L'idée est que la manipulation de formes d'évaluation linéaires en les variables bornées conduit à des parallélotopes convexes, domaines facilement manipulables et qu'il est possible de construire de manière exacte. Le principe est donc, dans un premier temps, de linéariser les formes d'évaluation par rapport aux variables bornées, tout en conservant la propriété de garantie signifiant que l'espace abstrait des nouvelles relations obtenues contiendra nécessairement celui des anciennes. Cette procédure de linéarisation induit un nombre conséquent de nouvelles variables bornées qu'il est possible de regrouper de manière à en limiter le nombre, objectif auquel s'attache la procédure de réduction présentée ci-après. Notons que cette dernière ne provoque aucune modification de l'espace abstrait considéré ; elle ne fait que simplifier les relations de manière à alléger les calculs qui suivront. Ensuite, une méthode générant les contraintes inégalités associées aux bandes de contraintes d'un parallélotope sera développée, permettant ainsi de réaliser très simplement les tests de cohérence utiles au module de détection de défauts. Enfin, la notion de distance entre un polytope et le vecteur résidu sera abordée. En effet, le résultat binaire du test précédent peut apparaître pauvre compte tenu de toute l'information à notre disposition. L'objectif n'est alors plus de tester si la forme de calcul se trouve à l'intérieur de l'espace abstrait de la forme d'évaluation, mais de calculer la distance entre p_{cal} et le plus proche hyperplan frontière de $\mathcal{A}(p_{eval}, \theta)$, ce qui représente la distance entre le comportement observé du système et la plus proche limite entre comportement normal et anormal.

IV.5.1 Procédure de linéarisation

Nous ne nous intéresserons qu'à des relations de parité polynomiales en les variables bornées, ce qui correspond à la majorité des cas traités au cours du troisième chapitre, et tout particulièrement lorsque la matrice de parité dépend des incertitudes. Lorsque celle-ci est certaine, cette hypothèse exclut la structure de modèle de capteurs (sections III.5.2 et III.7.2) et le cas restrictif où l'état est estimé lorsque la matrice d'observation est incertaine (section III.7.1.4). Le vecteur résidu peut dans ces conditions s'écrire sous la forme suivante :

$$\mathbf{p}(k+s, \boldsymbol{\theta}) = \underbrace{\mathbf{p}_l(k+s, \boldsymbol{\theta})}_{\text{termes d'ordres inférieurs ou égaux à 1}} + \underbrace{\mathbf{p}_{nl}(k+s, \boldsymbol{\theta})}_{\text{termes d'ordres supérieurs ou égaux à 2}}, \quad (\text{IV-2})$$

où les indices l et nl définissent respectivement les parties linéaire et non-linéaire du vecteur résidu \mathbf{p}^{11} . Notons que les deux vecteurs $\mathbf{p}_{(\cdot)}$ dépendent des mesures et des entrées commandées exprimées sur l'horizon temporel $[k, k+s]$. De plus, par commodité et afin d'alléger les écritures, nous nous sommes ramenés à une expression où les formes de calcul et d'évaluation sont regroupées dans un même terme \mathbf{p} ; néanmoins, nous avons déjà vu dans la section IV.2.1 que ces deux formalismes étaient strictement équivalents.

L'objectif de cette section est de proposer une méthode assurant la linéarisation du système précédent par rapport aux variables bornées, de façon à pouvoir générer plus facilement, par la suite, le test de cohérence permettant la détection des défauts (Adrot et al, 2000b). Pour commencer, au niveau de chaque composante du champ de vecteurs \mathbf{p}_{nl} , tous les monômes identiques constitués de variables bornées sont rassemblés. Ensuite, chaque monôme est remplacé par une nouvelle variable bornée et indépendante de même support. En notant, à l'aide de l'analyse par intervalles, que le produit $\theta_i \theta_j$ appartient nécessairement à l'intervalle $[-1,1]$ si les indices i et j sont différents, à $[0,1]$ sinon, les propriétés suivantes sont déduites :

- si toutes les puissances intervenant dans le monôme considéré sont paires, alors celui-ci est nécessairement compris dans l'intervalle $[0,1]$,
- si au moins une des puissances constituant le monôme est impaire, ce dernier appartient alors au domaine $[-1,1]$.

Par conséquent, relatif au second cas, des monômes tels que $m_1 = \theta_i \theta_j$, $m_2 = \theta_i^2 \theta_j$, $i \neq j$, par exemple, sont directement remplacés par deux nouvelles variables bornées normalisées et indépendantes v_1 et v_2 , $v_i \in [-1,1]$, $i \in \{1,2\}$, alors que dans le premier cas, un monôme tel que $m_3 = \theta_i^4 \theta_j^2$ est décrit par une nouvelle variable bornée non-centrée : $v_3 \in [0,1]$. De manière à perdre le minimum d'information relatif à la présence de variables communes, un même monôme est représenté par la même variable bornée dans toutes les composantes du champ de vecteurs \mathbf{p}_{nl} .

Dans ces conditions, le vecteur résidu (IV-2) devient :

$$\mathbf{p}_{lin}(k+s, \boldsymbol{\theta}, \mathbf{v}) = \mathbf{p}_l(k+s, \boldsymbol{\theta}) + \mathbf{p}_{l'}(k+s, \mathbf{v}), \quad (\text{IV-3})$$

où \mathbf{v} est un vecteur de variables bornées indépendant de $\boldsymbol{\theta}$ obtenu durant la procédure de linéarisation et où $\mathbf{p}_{l'}$ est un champ de vecteurs linéaire en \mathbf{v} .

Cette procédure conduit à une perte d'information dans la mesure où certaines dépendances présentes au niveau du vecteur résidu initial \mathbf{p} ne sont pas prises en compte dans \mathbf{p}_{lin} . Par exemple, la relation implicite entre les deux monômes m_1 et m_2 faisant intervenir les

¹¹ Le passage du formalisme formes de calcul/d'évaluation à celui-ci consiste simplement à poser $\mathbf{p}(k+s, \boldsymbol{\theta}) = \mathbf{p}_{éval}(k+s, \boldsymbol{\theta}) - \mathbf{p}_{cal}(k+s)$.

mêmes variables bornées ne peut être conservée puisqu'ils sont remplacés par deux variables indépendantes v_1 et v_2 . Par conséquent, l'espace abstrait $\mathcal{A}(\mathbf{p}_{lin})$ est nécessairement une majoration de $\mathcal{A}(\mathbf{p})$. En effet, \mathbf{p}_{lin} est un champ de vecteurs d'expression générale $\mathbf{f}(\boldsymbol{\theta}, \mathbf{v})$ où toutes les variables bornées θ_i et v_j sont indépendantes. En revanche, \mathbf{p} correspond exactement au même champ de vecteurs $\mathbf{f}(\boldsymbol{\theta}, \boldsymbol{\theta}_m)$, où $\boldsymbol{\theta}_m$ contient tous les monômes d'ordre supérieur à 1 situés dans \mathbf{p} et constitués de produits de puissances de θ_i . Les dépendances entre $\boldsymbol{\theta}$ et $\boldsymbol{\theta}_m$ génèrent des contraintes ; dans ces conditions, $\mathbf{f}(\boldsymbol{\theta}, \boldsymbol{\theta}_m)$ est un cas particulier du champ de vecteurs $\mathbf{f}(\boldsymbol{\theta}, \mathbf{v})$ et puisque $\mathcal{A}(\boldsymbol{\theta}_m) = \mathcal{A}(\mathbf{v})$, tout point de $\mathcal{A}(\mathbf{f}(\boldsymbol{\theta}, \boldsymbol{\theta}_m))$ est nécessairement un point de $\mathcal{A}(\mathbf{f}(\boldsymbol{\theta}, \mathbf{v}))$. L'espace abstrait $\mathcal{A}(\mathbf{f}(\boldsymbol{\theta}, \mathbf{v}))$ englobe donc $\mathcal{A}(\mathbf{f}(\boldsymbol{\theta}, \boldsymbol{\theta}_m))$ et ceci pour deux raisons. Tout d'abord, au sein d'une même composante, des dépendances sont éliminées et comme explicité au chapitre II, ceci se traduit par une surestimation des intervalles calculés. D'autre part, ceci engendre aussi la disparition de dépendances entre composantes du champ de vecteurs liées à la présence de variables communes. Supposons que le monôme m_1 apparaisse dans la $k^{\text{ième}}$ composante f_k de \mathbf{f} , il sera remplacé par la variable bornée v_1 . Ceci signifie que tout terme contenu dans une composante f_i , $k \neq l$, dépendant de θ_i et/ou θ_j sera supposé indépendant de m_1 . Ainsi, par exemple, le terme du premier ordre en θ_i de f_i ne sera plus lié à la variable v_1 de f_k remplaçant $\theta_i \theta_j$. Comme nous l'avons déjà démontré au cours du chapitre II, la disparition de certaines variables communes engendre une nouvelle majoration.

L'expression (IV-3) est constituée de variables ne possédant pas toutes les mêmes bornes. En notant que l'intervalle $[0,1]$ peut être mis sous la forme suivante : $\frac{1}{2} - \frac{1}{2}[-1,1]$, il est donc toujours possible d'homogénéiser le vecteur parité (IV-3) en le reformulant comme suit :

$$\mathbf{p}_{lin}(k+s, \boldsymbol{\theta}, \boldsymbol{\mu}) = \mathbf{p}_l(k+s, \boldsymbol{\theta}) + \mathbf{p}_l'(k+s, \boldsymbol{\mu}), \quad (\text{IV-4})$$

où $\boldsymbol{\mu}$ est un vecteur de variables bornées normalisées (contrairement à \mathbf{v}) ($\|\boldsymbol{\mu}\|_\infty \leq 1$) et où $\mathbf{p}_l'^{12}$ est un vecteur linéaire en $\boldsymbol{\mu}$.

En conclusion, cette méthode induit une perte de précision en engendrant une majoration du domaine abstrait à évaluer ; mais de toute façon, pour le type de systèmes étudiés, le domaine exact $\mathcal{A}(\mathbf{p})$ est généralement beaucoup trop complexe pour être évalué exactement, donc a fortiori utilisé. Dans ces conditions, $\mathcal{A}(\mathbf{p})$ sera approché par un polytope $\mathcal{A}(\mathbf{p}_{lin})$. L'importance de la majoration étant difficile à quantifier puisque le véritable espace abstrait est inconnu, nous retiendrons plutôt la méthode de linéarisation pour le peu de calculs numériques à effectuer lors de la phase de détection de défauts (section IV.5.3). Néanmoins, le nombre de nouvelles variables bornées peut être important ; heureusement certaines d'entre elles peuvent être regroupées et remplacées par un terme unique aux bornes adéquates sans modifier d'aucune manière l'espace abstrait $\mathcal{A}(\mathbf{p}_{lin})$ considéré, comme nous le verrons dans la section suivante.

Exemple. Reprenons le champ de vecteurs \mathbf{h} et appliquons-lui la procédure de linéarisation précédente :

¹² Les vecteurs $\mathbf{p}_l'(\mathbf{v})$ et $\mathbf{p}_l'(\boldsymbol{\mu})$ sont différenciés puisque leurs parties nominales peuvent être différentes.

$$\mathbf{h}(\boldsymbol{\theta}) = \begin{bmatrix} 0 & 1 & 2 \\ 1 & -1 & 0 \end{bmatrix} \begin{bmatrix} \theta_1 \\ \theta_2 \\ \theta_1^2 \end{bmatrix}, \text{ avec } \mathcal{A}(\theta_1) = [-1, 2] \\ \mathcal{A}(\theta_2) = [1, 3],$$

$$\text{ce qui conduit à : } \mathbf{h}_{lin}(\boldsymbol{\mu}) = \begin{bmatrix} 6 \\ -3 \\ 2 \end{bmatrix} + \begin{bmatrix} 0 & 1 & 4 \\ 3 & -1 & 0 \\ 2 & & \end{bmatrix} \boldsymbol{\mu} \text{ avec } \|\boldsymbol{\mu}\|_\infty \leq 1.$$

L'espace abstrait $\mathcal{A}(\mathbf{h}_{lin})$ ainsi que le domaine exact $\mathcal{A}(\mathbf{h})$ sont représentés sur la figure suivante de manière à pouvoir les comparer :

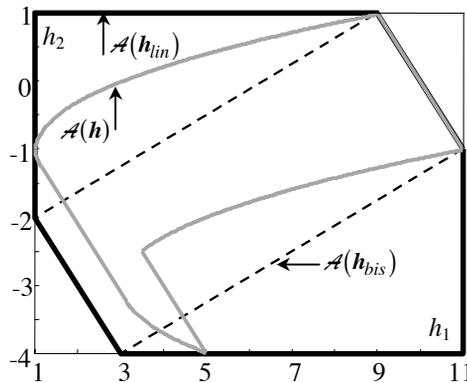


Figure IV-13. Espaces abstraits $\mathcal{A}(\mathbf{h}_{lin})$ et $\mathcal{A}(\mathbf{h})$

Remarquons que la propriété de garantie est bien conservée dans la mesure où $\mathcal{A}(\mathbf{h}_{lin})$ contient bien $\mathcal{A}(\mathbf{h})$ et que, dans le cadre de cet exemple, la majoration reste importante.

Une autre solution aurait pu être de remplacer la fonction non-linéaire θ_1^2 par une autre $f(\theta)$ linéaire mais de même espace abstrait :

$$f(\theta_1) = a\theta_1 + b \text{ avec } \begin{cases} f(2) = 4 \\ f(-1) = 0 \end{cases} \text{ soit } \begin{cases} -a + b = 0 \\ 2a + b = 4 \end{cases} \text{ ce qui conduit à : } f(\theta_1) = \frac{4}{3}(1 + \theta_1),$$

$$\text{soit au final : } \mathbf{h}_{bis}(\boldsymbol{\mu}) = \begin{bmatrix} 8 \\ 3 \\ 0 \end{bmatrix} + \begin{bmatrix} 8 & 1 \\ 3 & -1 \\ 1 & -1 \end{bmatrix} \left\{ \begin{bmatrix} 1 \\ 2 \\ 2 \end{bmatrix} + \begin{bmatrix} 3 \\ 2 \\ 0 \\ 1 \end{bmatrix} \boldsymbol{\mu} \right\}.$$

Cependant, si les orthotopes alignés circonscrits de \mathbf{h} et \mathbf{h}_{bis} sont bien identiques comme le suggère cette technique, en revanche, la simplicité de cette approche se traduit par une perte de garantie ($\mathcal{A}(\mathbf{h}_{bis})$ ne contient plus $\mathcal{A}(\mathbf{h})$) comme représenté sur la figure précédente, la rendant inexploitable. ■

IV.5.2 Procédure de réduction

L'objectif de cette section est de proposer une méthode permettant de réduire, quand cela est possible, le nombre de variables bornées introduites lors de la procédure de linéarisation (Adrot et al, 2000b). Le principe est que, sous certaines conditions, des variables bornées peuvent être regroupées puis remplacées par une nouvelle variable indépendante de

support adéquat sans aucune modification de l'espace abstrait $\mathcal{A}(\mathbf{p}_{lin})$, de façon à en simplifier l'évaluation ultérieure.

La relation (IV-4) étant linéaire en les variables bornées, nous supposons que toutes les variables identiques sont regroupées et n'apparaissent, par conséquent, au maximum, qu'une seule fois dans chaque composante de \mathbf{p}_{lin} (les variables θ_i et μ_j sont toutes indépendantes). Il n'existe donc plus aucun problème de variables dépendantes et chaque composante de cette expression pourrait être évaluée individuellement de manière exacte par l'arithmétique des intervalles ; seuls persistent des problèmes liés à la présence de variables communes. En regroupant les vecteurs incertains $\boldsymbol{\theta}$ et $\boldsymbol{\mu}$ dans l'égalité (IV-4), l'expression suivante est obtenue :

$$\mathbf{p}_{lin}(k+s, \mathbf{v}) = \mathbf{P}_v \mathbf{v} + \mathbf{p}_0, \quad \mathbf{v}^T = [\boldsymbol{\theta}^T \quad \boldsymbol{\mu}^T], \quad \|\mathbf{v}\|_\infty \leq 1, \quad (\text{IV-5})$$

où la matrice \mathbf{P}_v et le vecteur \mathbf{p}_0 ne dépendent que de grandeurs connues telles que les mesures ou les entrées commandées exprimées sur l'horizon temporel $[k, k+s]$.

Intéressons-nous plus particulièrement au terme $\mathbf{P}_v \mathbf{v}$ dépendant des incertitudes paramétriques. Pour mieux comprendre le principe de cette procédure, observons dans un premier temps ce qui se produit au niveau du champ de vecteurs suivant lorsque l'objectif est de regrouper ensemble les variables $v_i, i \in \{1, 2\}$:

$$\mathbf{p}_{lin} = \begin{bmatrix} v_1 - 2v_2 \\ 2v_1 - 5v_2 \end{bmatrix}.$$

Définissons par exemple une nouvelle variable bornée indépendante v' appartenant à l'intervalle $[-3, 3]$; la précédente égalité peut alors s'écrire comme suit :

$$\mathbf{p}'_{lin} = \begin{bmatrix} v' \\ 2v' - v_2 \end{bmatrix}.$$

Puisque les variables bornées v' et v_2 sont supposées indépendantes, en réalisant ce changement de variables, une partie de l'information est perdue et nécessairement l'espace abstrait $\mathcal{A}(\mathbf{p}_{lin})$ sera contenu dans $\mathcal{A}(\mathbf{p}'_{lin})$.

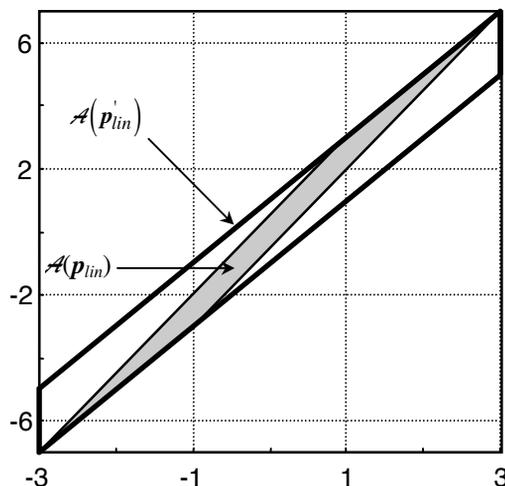


Figure IV-14. Domaines $\mathcal{A}(\mathbf{p}_{lin})$ et $\mathcal{A}(\mathbf{p}'_{lin})$

Puisque la procédure de réduction a pour cahier des charges de ne pas modifier $\mathcal{A}(\mathbf{p}_{lin})$, chaque dépendance doit impérativement être conservée. En réalité, toutes les variables bornées apparaissant sous la forme d'une même combinaison linéaire dans toutes les composantes de \mathbf{p}_{lin} , peuvent être regroupées puis remplacées par une nouvelle variable, indépendante des variables bornées restantes.

La seule difficulté provient de l'arithmétique des intervalles pour laquelle la somme et la soustraction de variables bornées symétriques conduisent au même intervalle :

$$\mathcal{A}\left(\sum \pm \alpha_i \nu_i\right) = \left(\sum |\alpha_i|\right)[-1,1], \quad \forall \alpha_i \in \mathbb{R}, \quad (\text{IV-6})$$

obtenu en faisant la somme des valeurs absolues des poids α_i .

Supposons que la notation $p_{\nu,ij}$ désigne le (i,j) ^{ième} élément de la matrice certaine \mathbf{P}_ν (IV-5), alors que l'ensemble $\mathcal{I}_{nc,i}$ définit tous les indices des variables non communes intervenant dans la $i^{\text{ème}}$ composante de \mathbf{p}_{lin} . De manière générale, toutes les variables $\nu_j, j \in \mathcal{I}_{nc,i}$ peuvent être regroupées et remplacées par :

$$\sum_{j \in \mathcal{I}_{nc,i}} |p_{\nu,ij}| \nu', \quad |\nu'| \leq 1, \quad (\text{IV-7})$$

où ν' est une variable bornée normalisée et indépendante. En effet, les variables non communes n'apparaissant que dans une seule composante de \mathbf{p}_{lin} , il peut leur être substitué une unique variable bornée, puisque le terme obtenu par regroupement est lui aussi non-commun. Il n'y a donc aucune perte d'information (au sens d'élimination du lien entre deux variables communes) en effectuant ce travail et donc aucune modification de l'espace abstrait.

$$\text{Par exemple, } \begin{bmatrix} \nu_1 - \nu_2 \\ \nu_3 \end{bmatrix} \text{ devient } \begin{bmatrix} 2\nu' \\ \nu_3 \end{bmatrix}.$$

Le même principe doit être appliqué en ce qui concerne les variables communes, mais cette fois, un regroupement ne peut s'effectuer que pour des variables communes apparaissant sous la forme d'une même combinaison linéaire dans chaque ligne de $\mathbf{P}_\nu \nu$ (à un facteur multiplicatif près pouvant être nul). La méthode devient simplement un peu plus compliquée dans la mesure où le signe avec lequel la combinaison linéaire apparaît dans les différentes lignes de $\mathbf{P}_\nu \nu$ doit être pris en compte. Considérons une combinaison dépendant de plusieurs variables bornées ν_j , dont les indices définissent l'ensemble \mathcal{J} , et apparaissant dans les lignes de $\mathbf{P}_\nu \nu$, dont les indices correspondent à l'ensemble \mathcal{I} . Soient i_0 un élément arbitraire de \mathcal{I} et j_1 le premier élément de \mathcal{J} pris comme référence. Une solution consiste à déterminer les signes de chaque élément de la ligne d'indice i_0 : $p_{\nu,i_0,j}, j \in \mathcal{J}$, puis, pour chaque ligne d'indice i appartenant à \mathcal{I} , la combinaison en question est remplacée par :

$$\sum_{j \in \mathcal{J}} \text{signe}(p_{\nu,i_0,j_1} p_{\nu,i_0,j}) p_{\nu,ij} \nu', \quad |\nu'| \leq 1 \text{ avec } \text{signe}(x) = \begin{cases} 1 & \text{si } x > 0 \\ -1 & \text{si } x < 0 \end{cases}. \quad (\text{IV-8})$$

Ce faisant, tous les termes sommés pour une même ligne i ont toujours le même signe (en l'occurrence celui de p_{ν,ij_1}) de manière à respecter l'arithmétique des intervalles (IV-6). En

effet, puisque chaque ligne $i \in \mathcal{I}$, contient à un facteur près la même combinaison des variables v_j , les signes des termes $p_{v,i_1} p_{v,i_2}, j \in \mathcal{J}$, sont nécessairement identiques quelque soit i . De ce fait, l'expression précédente (IV-8) peut aussi s'écrire :

$$\text{signe}(p_{v,i_1}) \sum_{j \in \mathcal{J}} \text{signe}(p_{v,i_2}) p_{v,i_2} v' = \text{signe}(p_{v,i_1}) \sum_{j \in \mathcal{J}} |p_{v,i_2}| v', |v'| \leq 1. \quad (\text{IV-9})$$

De manière plus compacte, les colonnes de \mathbf{P}_v d'indices $j \in \mathcal{J}$, notées $\mathbf{p}_{v,j}$, peuvent être regroupées comme suit sans modification de l'espace abstrait considéré :

$$\left(\mathbf{p}_{v,j_1} + \sum_{j \in \mathcal{J} - \{j_1\}} \text{signe}(p_{v,i_0,j_1} p_{v,i_0,j}) \mathbf{p}_{v,j} \right) v', |v'| \leq 1.$$

Par exemple, $\begin{bmatrix} 1 & -2 & 2 \\ -2 & 4 & 1 \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \\ v_3 \end{bmatrix}$ devient $\begin{bmatrix} 3 & 2 \\ -6 & 1 \end{bmatrix} \begin{bmatrix} v' \\ v_3 \end{bmatrix}$.

Ainsi, l'alternance des signes dans une colonne donnée de \mathbf{P}_v est prise en considération. La multiplication d'une colonne complète de \mathbf{P}_v par la valeur -1 ne modifie en rien l'espace abstrait $\mathcal{A}(\mathbf{p}_{lin})$ puisque toutes les variables bornées sont symétriques. Ainsi, le résultat précédent est équivalent à $\begin{bmatrix} -3 & 2 \\ 6 & 1 \end{bmatrix} \begin{bmatrix} v' \\ v_3 \end{bmatrix}$ (domaine ①). En revanche, $\begin{bmatrix} 3 & 2 \\ 6 & 1 \end{bmatrix} \begin{bmatrix} v' \\ v_3 \end{bmatrix}$ (domaine ②) ne correspond plus au domaine recherché comme dessiné sur la figure suivante, car dans ce cas, l'opposition de signes entre $p_{v,11} = 1$ ($p_{v,12} = -2$) et $p_{v,21} = -2$ ($p_{v,22} = 4$) n'est plus respectée (il n'est pas possible de traiter indépendamment l'une de l'autre les lignes de cette matrice).

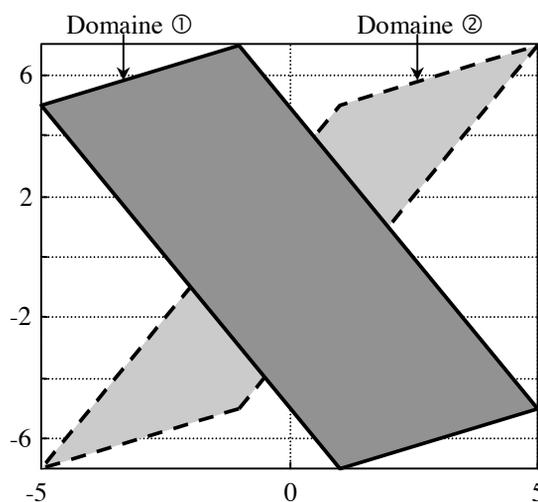


Figure IV-15. Comparaison des deux espaces abstraits associés à ① et ②

Pour opérer ce travail, une technique systématique et numérique consiste à tester deux à deux si les vecteurs colonne \mathbf{p}_{v,j_1} et \mathbf{p}_{v,j_2} de la matrice \mathbf{P}_v sont colinéaires. Lorsque le rang de la matrice $\begin{bmatrix} \mathbf{p}_{v,j_1} & \mathbf{p}_{v,j_2} \end{bmatrix}$ vaut 1, \mathbf{p}_{v,j_1} et \mathbf{p}_{v,j_2} sont éliminés et remplacés en utilisant la méthode expliquée ci-dessus. Le principe de l'algorithme est détaillé ci-dessous :

- (1) Déterminer la liste \mathcal{L} contenant toutes les combinaisons de 2 éléments parmi le nombre cpt de colonnes de \mathbf{P}_v .
- (2) Prendre successivement chaque élément l_k de \mathcal{L} , jusqu'à ce que le rang de la matrice composée des vecteurs colonne \mathbf{p}_{v,j_1} et \mathbf{p}_{v,j_2} , dont les indices (j_1, j_2) sont contenus dans l_k , vaille 1. S'il n'existe aucunes colonnes colinéaires, aller directement à l'étape (6).
- (3) Prendre l'une des deux colonnes comme référence, celle d'indice j_1 par exemple. Rechercher un des éléments non-nuls de cette colonne, que nous noterons p_{v,i_0,j_1} . Une nouvelle colonne est alors ajoutée à la matrice \mathbf{P}_v comme suit. Si p_{v,i_0,j_1} et p_{v,i_0,j_2} (qui ne peut être lui-aussi que non-nul) possèdent les mêmes signes ; alors la nouvelle colonne correspond à la somme $\mathbf{p}_{v,j_1} + \mathbf{p}_{v,j_2}$, sinon à la différence $\mathbf{p}_{v,j_1} + (-\mathbf{p}_{v,j_2})$.
- (4) Éliminer les colonnes \mathbf{p}_{v,j_1} et \mathbf{p}_{v,j_2} de \mathbf{P}_v .
- (5) Imposer $cpt = cpt - 1$; si cpt est différent de 1, retourner à l'étape (1).
- (6) Donner la nouvelle matrice \mathbf{P}_v .

Tableau IV-3. Squelette de la procédure de réduction

Remarquons que cet algorithme fonctionne de manière générale que les variables considérées soient communes ou non puisque le principe de fonctionnement reste le même. En effet, l'expression (IV-7) est équivalent à son homologue (IV-9) dans le cas de variables non communes puisque celles-ci n'intervenant que sur une seule et même ligne par définition, la prise en compte de l'alternance des signes entre lignes n'a plus raison d'être. Inversement, tenir compte du signe de p_{v,i,j_1} ne gêne pas puisqu'une colonne peut être multipliée par -1 sans entraîner aucune modification de l'espace abstrait. Dans ces conditions, l'expression plus générale (IV-8) peut avantageusement remplacer (IV-7), permettant de traiter à la fois variables communes et non communes.

Exemple. Le système $\begin{bmatrix} 1 & 1 & -1 & 1 & 0 & -2 \\ 2 & 3 & -2 & 0 & 1 & 0 \end{bmatrix} \mathbf{v}$ devient $\begin{bmatrix} 2 & 1 & 3 & 0 \\ 4 & 3 & 0 & 1 \end{bmatrix} \begin{bmatrix} v_{13} \\ v_2 \\ v_{46} \\ v_5 \end{bmatrix}$.

■

Cet algorithme peut être étendu dans le cas d'expressions symboliques, ce qui semble pertinent puisque les relations de parité utilisées dépendent de mesures (entrées, sorties) fluctuant au cours du temps. Un algorithme purement numérique nécessiterait alors de relancer la procédure de réduction à chaque instant, ce qui pourrait être préjudiciable à une implantation temps réel malgré la simplicité de l'algorithme proposé.

En considérant que les variables symboliques constituant P_v ne proviennent que des variables mesurées, la principale difficulté dans ces conditions est due aux changements de signes que peuvent occasionner les mesures. En effet, leurs signes justement peuvent être inconnus a priori, alors que ceux de p_{v,i_0j_1} et p_{v,i_0j_2} étaient parfaitement déterminés dans le cadre de l'algorithme numérique précédemment exposé. En réalité, tester la colinéarité de deux colonnes n'est pas plus difficile que dans le cas numérique, puisqu'une expression symbolique commune aux deux colonnes en question est simplement recherchée. De plus, connaître les signes de p_{v,i_0j_1} et p_{v,i_0j_2} n'est pas primordial, savoir s'ils sont différents est suffisant.

Le fait que des mesures ou des combinaisons de paramètres formels aux signes inconnus interviennent dans la matrice P_v n'induit pas un réel handicap, car l'algorithme repose simplement sur le fait de tester les colonnes deux à deux pour déterminer lesquelles sont colinéaires. Supposons qu'une seule expression symbolique f (c'est-à-dire une fonction vectorielle dépendant des mesures) existe dans au moins une des deux colonnes colinéaires considérées. L'expression f doit alors nécessairement apparaître en facteur, soit dans tous les éléments d'une des deux colonnes, soit au niveau des éléments de chacune des deux colonnes possédant les mêmes indices de ligne.

Dans le premier cas, puisque multiplier une colonne par ± 1 ne modifie en rien l'espace abstrait recherché, f peut donc être remplacée par sa valeur absolue avant de rassembler les deux colonnes étudiées sur le principe de ne sommer que des termes de même signe comme dans le cas numérique.

$$\text{Par exemple, } \begin{bmatrix} -1 & 2f \\ 2 & -4f \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \end{bmatrix} \text{ devient } \begin{bmatrix} -1-2|f| \\ 2+4|f| \end{bmatrix} v'.$$

Dans le second cas, la valeur absolue de f ne peut être utilisée de manière générale puisque modifier tous les signes d'une colonne de la même manière devient impossible. Néanmoins, si le signe de f est inconnu, il est trivial de noter que f et $-f$ sont de signes opposés. A partir de ce constat, si p_{v,i_0j_1} et p_{v,i_0j_2} possèdent les mêmes signes, les deux colonnes considérées peuvent être sommées, sinon soustraites.

$$\text{Par exemple, } \begin{bmatrix} -f & 2f \\ 2 & -4 \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \end{bmatrix} \text{ devient } \begin{bmatrix} -3f \\ 6 \end{bmatrix} v' \text{ ou bien } \begin{bmatrix} 3f \\ -6 \end{bmatrix} v'.$$

Toutes les autres situations peuvent être déduites à partir de ces deux cas élémentaires. Par conséquent, la procédure de réduction peut toujours être appliquée à des expressions symboliques après avoir remplacé chaque expression apparaissant en facteur dans une même colonne et possédant un signe indéterminé, par sa valeur absolue :

$$\begin{bmatrix} -f & 2fg \\ 2 & -4g \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \end{bmatrix} \text{ devient } \begin{bmatrix} -f & 2f|g| \\ 2 & -4|g| \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \end{bmatrix}, \text{ puis après regroupement } \begin{bmatrix} -f(1+2|g|) \\ 2+4|g| \end{bmatrix} v'.$$

$$\begin{bmatrix} -f & 2fg \\ 2h & -4gh \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \end{bmatrix} \text{ devient } \begin{bmatrix} -f & 2f|g| \\ 2h & -4|g|h \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \end{bmatrix}, \text{ puis } \begin{bmatrix} -f(1+2|g|) \\ 2(1+2|g|h) \end{bmatrix} v'.$$

En conclusion, pour des systèmes de faible dimension ou d'autres pour lesquels les temps de calcul ne sont pas contraignants, l'algorithme numérique, relancé à chaque arrivée d'un nouvel échantillon de mesures, peut être utilisé. Sinon, en intégrant au niveau de l'étape 3 les deux règles de calcul précédentes, il est possible de générer une expression symbolique de la matrice \mathbf{P}_v en fonction des mesures et de l'évaluer ensuite chaque instant numériquement pour en connaître la valeur.

Dans les deux cas, cette procédure conduit à un nouveau vecteur résidu $\mathbf{p}_{lin,r}$ toujours linéaire en les variables bornées et de même espace abstrait que \mathbf{p}_{lin} :

$$\mathbf{p}_{lin,r}(k+s, \mathbf{v}_r) = \mathbf{P}_{v_r} \mathbf{v}_r + \mathbf{p}_0, \quad \|\mathbf{v}_r\|_\infty \leq 1, \quad (\text{IV-10})$$

où \mathbf{v}_r est un vecteur de variables bornées normalisées réduit, c'est-à-dire de dimension s_v inférieure (ou égale) à celle de \mathbf{v} . En désignant par l'entier s_p la dimension du vecteur $\mathbf{p}_{lin,r}$, rappelons que la matrice \mathbf{P}_{v_r} et le vecteur \mathbf{p}_0 , dépendant des mesures, sont respectivement de dimensions $s_p \times s_v$ et s_p .

IV.5.3 Génération des contraintes liées à un paralléloptope

L'objectif de cette partie est de proposer une méthode permettant la construction de l'espace abstrait associé à un champ de vecteurs linéaire en les incertitudes paramétriques tel que (IV-10). Dans ce cas précis, le domaine $\mathcal{A}(\mathbf{p}_{lin,r})$ est un paralléloptope centré en \mathbf{p}_0 , en d'autres termes un polytope (c'est-à-dire un polyèdre fermé) convexe délimité par des hyperplans deux à deux parallèles générant des bandes de contraintes. Un paralléloptope peut alors être parfaitement représenté au choix par ses sommets ou par l'intersection de ses bandes de contraintes, définies de manière générale sous la forme suivante : $\mathbf{H}\mathbf{p}_{lin,r} \leq \mathbf{k}$ ¹³ (Ziegler, 1998), où la matrice \mathbf{H} et le vecteur \mathbf{k} dépendent des sorties et entrées commandées.

Dans l'optique d'une procédure de diagnostic de systèmes, l'espace abstrait $\mathcal{A}(\mathbf{p}_{lin,r})$ représente l'ensemble des valeurs admissibles du vecteur résidu $\mathbf{p}_{lin,r}$, cohérentes avec le modèle de référence choisi et l'espace des paramètres $\mathcal{A}(\mathbf{v}_r) = \{\mathbf{v}_r \in \mathbb{R}^{s_v} / \|\mathbf{v}_r\|_\infty \leq 1\}$. Puisqu'en bon fonctionnement, le vecteur résidu est censé être nul, le test de cohérence se limite donc à tester si l'origine O appartient ou non au paralléloptope $\mathcal{A}(\mathbf{p}_{lin,r})$:

$$\mathbf{P}_{v_r} \mathbf{v}_r + \mathbf{p}_0 = 0, \quad \|\mathbf{v}_r\|_\infty \leq 1. \quad (\text{IV-11})$$

Dans ces conditions, la connaissance des sommets¹⁴ de $\mathcal{A}(\mathbf{p}_{lin,r})$ nous importe peu, puisque les expressions de ses bandes de contraintes suffisent parfaitement pour conclure. En effet, il suffit de vérifier si l'origine satisfait bien chacune des contraintes inégalités associées ; en d'autres termes de tester si : $0 \leq \mathbf{k}$, auquel cas le système est supposé sain. La méthode

¹³ Remarquons que cette forme de représentation n'est pas propre à l'approche ensembliste, mais se retrouve par exemple dans tout ce qui a trait aux méthodes LMI (Linear Matrix Inequality) (Ploix et al, 1999c), (Ragot et al, 1999).

¹⁴ Une méthode permettant de déterminer ces sommets sera proposée dans le chapitre suivant dédié à la caractérisation des bornes des incertitudes d'un modèle. L'algorithme proposé reposera sur la génération des contraintes inégalité représentatives de $\mathcal{A}(\mathbf{p}_{lin,r})$ ici développée.

proposée dans cette section, très simple à implémenter, consiste à déterminer les différentes inégalités représentatives du paralléloétope $\mathcal{A}(\mathbf{p}_{lin,r})$ en générant la matrice \mathbf{H} et le vecteur \mathbf{k} (Ploix et al, 2000), (Adrot et al, 2000b). Dans son principe de fonctionnement, elle procède du calcul des contraintes issues de l'orthotope aligné circonscrit $\square\mathcal{A}(\mathbf{p}_{lin,r})$ et de l'élimination des variables communes telle qu'expliquée dans la section IV.4. L'étude proposée se limite simplement à un champ de vecteurs linéaire en les variables bornées, ce qui permet d'explicitier un algorithme entièrement numérique.

Pour simplifier l'exposé, la matrice \mathbf{P}_{v_r} est supposée de plein rang ligne. Remarquons que cette hypothèse n'est nullement restrictive, puisque, si \mathbf{P}_{v_r} n'est pas de plein rang ligne, il est possible de re-conditionner le problème en isolant les vecteurs ligne générateurs de cette matrice. En effet, il est toujours possible de la décomposer de la manière suivante :

$$\mathbf{P}_{v_r} = \mathbf{P}_a^T \mathbf{P}_b,$$

où les matrices \mathbf{P}_a et \mathbf{P}_b sont toutes deux de plein rang ligne comme explicité dans la section III.10.1 en se limitant cette fois au contexte de matrices certaines.

La matrice \mathbf{P}_a étant de plein rang ligne, le système (IV-11) est multiplié par le pseudo-inverse de \mathbf{P}_a défini par :

$$\mathbf{P}_a^+ = (\mathbf{P}_a \mathbf{P}_a^T)^{-1} \mathbf{P}_a.$$

À la relation obtenue, sont ajoutées les contraintes supplémentaires liées au complémentaire de la matrice \mathbf{P}_a en multipliant à gauche ce même système (IV-11) par la matrice suivante :

$$\bar{\mathbf{P}}_a = \mathbf{I}_{rang(\mathbf{P}_a)} - \mathbf{P}_a^T (\mathbf{P}_a \mathbf{P}_a^T)^{-1} \mathbf{P}_a.$$

Finalement le nouveau système obtenu s'écrit :

$$\begin{cases} \mathbf{P}_b \mathbf{v}_r + \mathbf{P}_a^+ \mathbf{p}_0 = 0 \\ \bar{\mathbf{P}}_a \mathbf{p}_0 = 0 \end{cases}, \quad \|\mathbf{v}_r\|_\infty \leq 1,$$

où la première relation peut se mettre exactement sous la même forme que le système initial (IV-11), alors que la seconde ne correspond en réalité qu'à une relation déterministe, le test de cohérence s'y rapportant s'effectuant donc de manière classique dans un contexte certain. Par conséquent, même si la matrice \mathbf{P}_{v_r} n'est pas de plein rang ligne, il est toujours possible de se ramener à une forme équivalente au système initial (IV-11) pour lequel la nouvelle matrice de distribution des incertitudes (\mathbf{P}_b en l'occurrence) le devient. L'hypothèse n'est donc pas restrictive, elle vise simplement à conditionner le problème en une forme standard.

Pour commencer, intéressons-nous brièvement au calcul de l'orthotope aligné circonscrit $\square\mathcal{A}(\mathbf{p}_{lin,r})$, présentant l'avantage de s'évaluer plus facilement que $\mathcal{A}(\mathbf{p}_{lin,r})$. Il suffit en effet de rechercher séparément les bornes que chaque composante de $\mathbf{p}_{lin,r}$. Ces intervalles s'obtenant simplement par l'arithmétique des intervalles, le test de cohérence (approché) devient alors :

$$O \in \square \mathcal{A}(\mathbf{p}_{lin,r}) \Leftrightarrow \forall i \in \{1, \dots, s_p\}, 0 \in \left[p_{0,i} - \|\mathbf{p}_{v_r,i}^T\|_1, p_{0,i} + \|\mathbf{p}_{v_r,i}^T\|_1 \right],$$

où $p_{0,i}$ est le $i^{\text{ième}}$ élément du vecteur \mathbf{p}_0 et $\mathbf{p}_{v_r,i}^T$ la $i^{\text{ième}}$ ligne de \mathbf{P}_{v_r} .

Le test précédent consiste donc à déterminer un ensemble de doubles contraintes inégalité à satisfaire pour que le système soit considéré en fonctionnement normal. Il faut maintenant déterminer les contraintes inégalité définissant exactement $\mathcal{A}(\mathbf{p}_{lin,r})$, de manière à pouvoir réaliser le véritable test de cohérence. L'espace abstrait recherché est un paralléloétope, centré en \mathbf{p}_0 , c'est-à-dire un domaine convexe délimité par des hyperplans deux à deux parallèles. Pour simplifier le problème, nous posons dans un premier temps $\mathbf{p}' = \mathbf{p}_{lin,r} - \mathbf{p}_0$ et calculons $\mathcal{A}(\mathbf{p}')$, centré sur l'origine. Par nature, un tel paralléloétope correspond à l'intersection de bandes de contraintes \mathcal{S}_i (espace situé entre deux hyperplans parallèles), s'écrivant de manière générale :

$$\mathcal{S}_i = \left\{ \mathbf{p}' / -k_i \leq \mathbf{h}_i^T \mathbf{p}' \leq k_i \right\} = \left\{ \mathbf{p}' / \left| \mathbf{h}_i^T \mathbf{p}' \right| \leq k_i \right\}, \quad (\text{IV-12})$$

où k_i est un nombre réel et \mathbf{h}_i^T un vecteur ligne. Cette bande de contrainte peut facilement être ramenée à $\mathcal{A}(\mathbf{p}_{lin,r})$ de la manière suivante en effectuant le changement de variables $\mathbf{p}' = \mathbf{p}_{lin,r} - \mathbf{p}_0$:

$$\mathcal{S}_i = \left\{ \mathbf{p}_{lin,r} / -k_i + \mathbf{h}_i^T \mathbf{p}_0 \leq \mathbf{h}_i^T \mathbf{p}_{lin,r} \leq k_i + \mathbf{h}_i^T \mathbf{p}_0 \right\}.$$

L'objectif de cette section est donc de rechercher les bandes de contraintes \mathcal{S}_i telles que :

$$\mathcal{A}(\mathbf{p}_{lin,r}) = \bigcap_i \mathcal{S}_i,$$

c'est-à-dire les différentes contraintes inégalité nécessaires à la réalisation du test de cohérence. Nous pourrions alors conclure à un comportement sain si $O \in \mathcal{A}(\mathbf{p}_{lin,r})$, soit si :

$$\forall i, -k_i + \mathbf{h}_i^T \mathbf{p}_0 \leq 0 \leq k_i + \mathbf{h}_i^T \mathbf{p}_0. \quad (\text{IV-13})$$

Dans le cas où l'une de ces conditions n'est pas vérifiée, il faut conclure à une incohérence.

Une bande de contrainte correspond en fait à l'intersection de deux demi-espaces auxquels sont associés deux hyperplans les délimitant :

$$\mathcal{H}_i^+ = \left\{ \mathbf{p}' / \mathbf{h}_i^T \mathbf{p}' = k_i \right\} \text{ et } \mathcal{H}_i^- = \left\{ \mathbf{p}' / \mathbf{h}_i^T \mathbf{p}' = -k_i \right\}.$$

Pour évaluer les expressions des hyperplans \mathcal{H}_i^+ et \mathcal{H}_i^- , recherchons ce que doit vérifier le vecteur \mathbf{p}' pour appartenir à la facette de $\mathcal{A}(\mathbf{p}')$ incluse dans \mathcal{H}_i^+ par exemple. Plaçons-nous en un point $\mathbf{p}' \in \mathcal{A}(\mathbf{p}')$, c'est-à-dire tel qu'il existe $\mathbf{v}_r \in \mathcal{A}(\mathbf{v}_r)$ pour lequel $\mathbf{p}' = \mathbf{P}_{v_r} \mathbf{v}_r$. L'objectif, dans un premier temps, est d'interpréter la notion d'hyperplan, frontière de l'espace abstrait recherché. Pour cela, nous faisons appel à la dérivée de Lie. Lorsqu'une petite variation $\partial \lambda$ ($\partial \lambda \in \mathbb{R}$) dans une direction \mathbf{v} donnée est appliquée dans l'ensemble de départ d'un champ de vecteurs $\mathbf{y} = \mathbf{f}(\mathbf{x})$, ceci engendre une modification de \mathbf{y} donnée par la dérivée de Lie :

$$\partial \mathbf{y} = \mathbf{L}_{\mathbf{v}}(\mathbf{f}(\mathbf{x})) \partial \lambda,$$

où $L_V(f(x))$ correspond au produit du Jacobien de f par le vecteur \mathbf{v} . Ainsi, par exemple, si \mathbf{x} est un vecteur de dimension 2, alors :

$$\partial \mathbf{y} = \begin{bmatrix} \frac{\partial f}{\partial x_1} & \frac{\partial f}{\partial x_2} \end{bmatrix} \mathbf{v} \partial \lambda.$$

Remarquons que dans notre cas, le vecteur \mathbf{p}' peut s'écrire :

$$\mathbf{p}' = \sum_{i=1}^{s_v} \mathbf{p}_{v_r,i} v_{r,i}, \text{ avec } \mathbf{p}_{v_r,i} \text{ } i^{\text{ième}} \text{ colonne de } \mathbf{P}_{v_r} \text{ et } v_{r,i} \text{ } i^{\text{ième}} \text{ élément de } \mathbf{v}_r,$$

$$\text{d'où } \frac{\partial(\mathbf{p}')}{\partial v_{r,i}} = \mathbf{p}_{v_r,i} \text{ et } L_V(\mathbf{p}') = \begin{bmatrix} \mathbf{p}_{v_r,1} & \cdots & \mathbf{p}_{v_r,s_v} \end{bmatrix} \mathbf{v} = \mathbf{P}_{v_r} \mathbf{v}, \mathbf{v} \in \mathbb{R}^{s_v}.$$

L'objectif est d'observer la modification du champ de vecteurs \mathbf{p}' lorsqu'une petite variation est occasionnée. Si \mathbf{v}_r varie légèrement de $\partial \lambda$ dans la direction $\mathbf{v} \in \mathbb{R}^{s_v}$, de ce qui précède est déduit que le champ de vecteurs \mathbf{p}' va varier de :

$$\partial \mathbf{p}' = \mathbf{P}_{v_r} \mathbf{v} \partial \lambda.$$

La dérivée de Lie est intéressante par la notion de direction qu'elle engendre. Dans l'espace d'arrivée \mathbb{R}^{s_p} , un hyperplan est de dimension $s_p - 1$, donc engendré par $s_p - 1$ vecteurs indépendants : $\begin{bmatrix} \mathbf{w}_1 & \cdots & \mathbf{w}_{s_p-1} \end{bmatrix}$, $\mathbf{w}_i \in \mathbb{R}^{s_p}$. Tout point de l'hyperplan frontière se déplaçant dans une direction \mathbf{w} engendrée par les vecteurs précédents reste dans cet hyperplan. Tout déplacement dans une direction \mathbf{w} qui n'est pas une combinaison linéaire des \mathbf{w}_i , $i \in \{1, \dots, s_p - 1\}$ sort de l'espace abstrait recherché et y reste confiné pour la direction opposée $-\mathbf{w}$, (ou inversement). Un hyperplan frontière peut donc être défini comme l'ensemble des points de \mathbb{R}^{s_p} tels que pour tous les vecteurs non-engendrés par les $s_p - 1$ vecteurs indépendants \mathbf{w}_i , il ne soit pas possible de rester dans cet hyperplan. Dans ces conditions, la méthode consiste à se placer en un point de \mathcal{H}_i^+ et à scruter dans diverses directions.

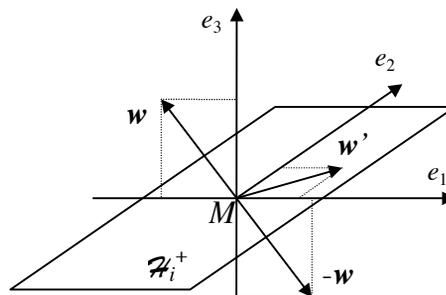


Figure IV-16. Définition d'un hyperplan frontière

Sur la figure précédente, un petit exemple est présenté pour $s_p = 3$. Les vecteurs e_i représentés appartiennent à la base canonique de \mathbb{R}^3 . L'hyperplan \mathcal{H}_i^+ est engendré par e_1 et e_2 , tout comme l'est la direction \mathbf{w}' . Le déplacement du point M selon \mathbf{w}' définit un point appartenant toujours à \mathcal{H}_i^+ . En revanche, en ce qui concerne les directions quelconques \mathbf{w} et $-\mathbf{w}$ (engendrés par e_1 et e_3), les points obtenus ne sont plus éléments de \mathcal{H}_i^+ .

L'emploi de la dérivée de Lie s'impose alors naturellement, notamment parce qu'elle permet de faire le lien entre le domaine d'arrivée \mathbb{R}^{s_p} du champ de vecteur \mathbf{p}' et le domaine de départ \mathbb{R}^{s_v} de \mathbf{v}_r . L'hyperplan \mathcal{H}_i^+ fait partie de l'espace abstrait recherché, il existe donc $s_p - 1$ vecteurs \mathbf{v}_i indépendants de l'ensemble de départ \mathbb{R}^{s_v} tels que :

$$\mathbf{w}_i = \mathbf{P}_{v_r} \mathbf{v}_i, \quad \mathbf{v}_i \in \mathbb{R}^{s_v}, \quad i \in \{1, \dots, s_p - 1\}.$$

En effet, la matrice \mathbf{P}_{v_r} étant de plein rang ligne, il faut nécessairement que les \mathbf{v}_i soient indépendants, sinon les \mathbf{w}_i ne le sont pas :

$$\begin{aligned} \mathbf{P}_{v_r} \begin{bmatrix} \mathbf{v}_1 & \cdots & \mathbf{v}_{s_p-1} \end{bmatrix} &= \mathbf{U}_P \begin{bmatrix} \boldsymbol{\Sigma}_P & \mathbf{0} \end{bmatrix} \mathbf{V}_P^T \mathbf{U} \begin{bmatrix} \boldsymbol{\Sigma} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix} \mathbf{V}^T \quad \text{avec } \mathbf{U} = [\mathbf{U}_1 \quad \mathbf{U}_2] \text{ et } \mathbf{V}_P = [\mathbf{V}_{P,1} \quad \mathbf{V}_{P,2}] \\ &= \mathbf{U}_P \boldsymbol{\Sigma}_M \mathbf{V}_{P,1}^T [\mathbf{U}_1 \boldsymbol{\Sigma} \quad \mathbf{0}] \mathbf{V}^T \\ &= \mathbf{U}_P \begin{bmatrix} \boldsymbol{\Sigma}_M \mathbf{V}_{P,1}^T \mathbf{U}_1 \boldsymbol{\Sigma} & \mathbf{0} \end{bmatrix} \mathbf{V}^T \end{aligned}$$

Par analogie avec la dérivée de Lie, il existe donc $s_p - 1$ directions indépendantes \mathbf{v}_i dans \mathbb{R}^{s_v} donnant $\mathbf{w}_i = \mathbf{l}_{v_i}(\mathbf{p}')$, $i \in \{1, \dots, s_p - 1\}$ telles que :

$$\text{rang} \left(\begin{bmatrix} \mathbf{l}_{v_1}(\mathbf{p}') & \cdots & \mathbf{l}_{v_{s_p-1}}(\mathbf{p}') \end{bmatrix} \right) = \text{rang} \left(\mathbf{P}_{v_r} \begin{bmatrix} \mathbf{v}_1 & \cdots & \mathbf{v}_{s_p-1} \end{bmatrix} \right) = s_p - 1.$$

En conclusion, si \mathbf{p}' appartient à $\mathcal{A}(\mathbf{p}') \cap \mathcal{H}_i^+$, alors il est nécessaire que les deux conditions suivantes soient vérifiées. Il faut tout d'abord que l'hyperplan \mathcal{H}_i^+ puisse être engendré par $s_p - 1$ vecteurs indépendants $\{\mathbf{l}_{v_1}(\mathbf{p}'), \dots, \mathbf{l}_{v_{s_p-1}}(\mathbf{p}')\}$ de \mathbb{R}^{s_p} associés à $s_p - 1$ vecteurs indépendants de \mathbb{R}^{s_v} $\{\mathbf{v}_1, \dots, \mathbf{v}_{s_p-1}\}$ vérifiant :

$$\text{rang} \left(\mathbf{P}_{v_r} \begin{bmatrix} \mathbf{v}_1 & \cdots & \mathbf{v}_{s_p-1} \end{bmatrix} \right) = s_p - 1. \quad (\text{IV-14})$$

Il faut aussi que, quel que soit \mathbf{v} n'appartenant pas au sous-espace $\mathcal{E}(\mathbf{v}_1, \dots, \mathbf{v}_{s_p-1})$ engendré par les vecteurs $\{\mathbf{v}_1, \dots, \mathbf{v}_{s_p-1}\}$, une variation dans un sens \mathbf{v} d'amplitude $\partial\lambda$ conduise en dehors du domaine $\mathcal{A}(\mathbf{p}')$ tandis qu'une variation dans l'autre sens mène à l'intérieur de ce même domaine. En notant $\mathcal{E}(\mathbf{v}_{s_p}, \dots, \mathbf{v}_{s_v})$ le sous-espace complémentaire de $\mathcal{E}(\mathbf{v}_1, \dots, \mathbf{v}_{s_p-1})$ dans \mathbb{R}^{s_v} , engendré par les vecteurs $\{\mathbf{v}_{s_p}, \dots, \mathbf{v}_{s_v}\}$, nous traduirons cette condition par l'écriture suivante :

$$\forall \mathbf{v} \in \mathcal{E}(\mathbf{v}_{s_p}, \dots, \mathbf{v}_{s_v}), \exists \mathbf{v}_r \in \mathcal{A}(\mathbf{v}_r), \partial\lambda \in \mathbb{R}^* / \begin{cases} \|\mathbf{v}_r + \mathbf{v}\partial\lambda\|_\infty \leq 1 \\ \|\mathbf{v}_r - \mathbf{v}\partial\lambda\|_\infty > 1 \end{cases}. \quad (\text{IV-15})$$

Ceci revient à dire que \mathbf{p}' se retrouve en dehors du paralléloétope si au moins une des variables bornées sort de son espace abstrait $[-1, 1]$.

De plus, la condition (IV-15) n'est guère facile à manipuler sous cette forme. Pour la

reformuler, décomposons les deux vecteurs \mathbf{v}_r et \mathbf{v} respectivement en :

$$\mathbf{v}_r = \begin{bmatrix} \mathbf{v}_1 & \cdots & \mathbf{v}_{s_p-1} \end{bmatrix} \boldsymbol{\xi} + \begin{bmatrix} \mathbf{v}_{s_p} & \cdots & \mathbf{v}_{s_v} \end{bmatrix} \boldsymbol{\zeta}, \quad \boldsymbol{\xi} \in \mathbb{R}^{s_p-1}, \quad \boldsymbol{\zeta} \in \mathbb{R}^{s_v-s_p+1},$$

$$\text{et } \mathbf{v} = \begin{bmatrix} \mathbf{v}_{s_p} & \cdots & \mathbf{v}_{s_v} \end{bmatrix} \boldsymbol{\chi}, \quad \boldsymbol{\chi} \in \mathbb{R}^{s_v-s_p+1}.$$

La condition précédente devient alors :

$$\forall \boldsymbol{\chi} \in \mathbb{R}^{s_v-s_p+1}, \exists \mathbf{v}_r / \|\mathbf{v}_r\|_\infty \leq 1, \partial \lambda \in \mathbb{R}^* / \left\{ \begin{array}{l} \left\| \begin{bmatrix} \mathbf{v}_1 \cdots \mathbf{v}_{s_p-1} \end{bmatrix} \boldsymbol{\xi} + \begin{bmatrix} \mathbf{v}_{s_p} \cdots \mathbf{v}_{s_v} \end{bmatrix} (\boldsymbol{\zeta} + \partial \lambda \boldsymbol{\chi}) \right\|_\infty \leq 1 \\ \left\| \begin{bmatrix} \mathbf{v}_1 \cdots \mathbf{v}_{s_p-1} \end{bmatrix} \boldsymbol{\xi} + \begin{bmatrix} \mathbf{v}_{s_p} \cdots \mathbf{v}_{s_v} \end{bmatrix} (\boldsymbol{\zeta} - \partial \lambda \boldsymbol{\chi}) \right\|_\infty > 1 \end{array} \right.$$

Cette condition s'interprète facilement dès lors que les vecteurs $\{\mathbf{v}_1, \dots, \mathbf{v}_{s_v}\}$ définissent une base orthonormée de \mathbb{R}^{s_v} , hypothèse que nous conserverons dans la suite de cette démonstration. Dans ce cas, chaque élément du vecteur $\boldsymbol{\zeta}$ ne peut prendre comme valeurs que +1 ou -1. En effet, par définition de $\mathcal{A}(\mathbf{v}_r)$, les inégalités suivantes sont forcément vérifiées : $\|\boldsymbol{\xi}\|_\infty \leq 1$ et $\|\boldsymbol{\zeta}\|_\infty \leq 1$. Plaçons-nous dans le cas d'une petite variation $\partial \lambda$ et supposons que le terme $\zeta_j + \chi_j \partial \lambda$, $j^{\text{ième}}$ élément du vecteur $\boldsymbol{\zeta} + \boldsymbol{\chi} \partial \lambda$, soit positif. La première inégalité impose que $\zeta_j + \chi_j \partial \lambda \leq 1$ et la seconde que $\zeta_j - \chi_j \partial \lambda > 1$, soit $\zeta_j \in]1 + \chi_j \partial \lambda, 1 - \chi_j \partial \lambda[$, d'où $\zeta_j = 1$ en faisant tendre $\partial \lambda$ vers 0. Si $\zeta_j + \chi_j \partial \lambda$ est négatif, alors nous obtenons que $-\zeta_j - \chi_j \partial \lambda \leq 1$ et $\zeta_j + \chi_j \partial \lambda > 1$, soit $\zeta_j \in]-1 - \chi_j \partial \lambda, 1 + \chi_j \partial \lambda[$, d'où $\zeta_j = -1$. Nous noterons donc :

$$\boldsymbol{\zeta} = \mathbf{1}_\pm^+,$$

où $\mathbf{1}_\pm^+$ est un vecteur de $\mathbb{R}^{s_v-s_p+1}$ constitué uniquement de 1 et de -1.

En introduisant ce résultat dans l'expression de $\mathbf{p}' = \mathbf{P}_{\mathbf{v}_r} \mathbf{v}_r$, nous déduisons que l'ensemble des valeurs de \mathbf{p}' appartenant à l'hyperplan \mathcal{H}_i^+ de la bande de contrainte \mathcal{S}_i s'écrit :

$$\mathbf{p}' = \mathbf{P}_{\mathbf{v}_r} \begin{bmatrix} \mathbf{v}_1 & \cdots & \mathbf{v}_{s_p-1} \end{bmatrix} \boldsymbol{\xi} + \mathbf{P}_{\mathbf{v}_r} \begin{bmatrix} \mathbf{v}_{s_p} & \cdots & \mathbf{v}_{s_v} \end{bmatrix} \boldsymbol{\zeta},$$

sous la contrainte précédente $\boldsymbol{\zeta} = \mathbf{1}_\pm^+$ et où les vecteurs \mathbf{v}_j sont les vecteurs de la base orthonormée de \mathbb{R}^{s_v} . Pour se ramener à une équation cartésienne, il s'agit d'éliminer le terme en $\boldsymbol{\xi}$, c'est-à-dire qu'il faut trouver une matrice \mathbf{H}_i telle que :

$$\mathbf{H}_i \mathbf{P}_{\mathbf{v}_r} \begin{bmatrix} \mathbf{v}_1 & \cdots & \mathbf{v}_{s_p-1} \end{bmatrix} = \mathbf{0}.$$

Or, du fait de la condition (IV-14), l'ensemble des solutions est en fait donné par un vecteur ligne \mathbf{h}_i^T . Ainsi, toutes les combinaisons de $s_p - 1$ vecteurs définissant une base orthonormée de \mathbb{R}^{s_v} conduisent, lorsque cette condition est vérifiée, à un vecteur ligne \mathbf{h}_i^T définissant l'équation cartésienne de la frontière \mathcal{H}_i^+ de la bande de contrainte \mathcal{S}_i (IV-12) :

$$\mathbf{h}_i^T \mathbf{p}' = \mathbf{h}_i^T \mathbf{P}_{\mathbf{v}_r} \begin{bmatrix} \mathbf{v}_{s_p} & \cdots & \mathbf{v}_{s_v} \end{bmatrix} \boldsymbol{\zeta}, \text{ avec } \boldsymbol{\zeta} = \mathbf{1}_\pm^+. \quad (\text{IV-16})$$

Cependant, $\boldsymbol{\zeta}$ n'est pas encore déterminé ; la seule chose connue étant que ce vecteur est uniquement constitué de 1 et de -1. Ceci est dû aux deux conditions (IV-14) et (IV-15) qui sont nécessaires mais non suffisantes. Néanmoins, par définition, il est possible d'écrire :

$$\mathbf{h}_i^T \mathbf{p}' = \mathbf{h}_i^T \mathbf{P}_{\nu_r} \nu_r, \quad (\text{IV-17})$$

et les formules de l'arithmétique des intervalles permettent alors de conclure à :

$$-\|\mathbf{h}_i^T \mathbf{P}_{\nu_r}\|_1 \leq \mathbf{h}_i^T \mathbf{p}' \leq \|\mathbf{h}_i^T \mathbf{P}_{\nu_r}\|_1.$$

Par analogie entre (IV-16) et (IV-17), l'arithmétique des intervalles rend donc possible la résolution du problème portant sur l'indétermination de ζ , ce qui permet d'en déduire que le réel k_i , associé à \mathbf{h}_i^T et défini en (IV-12), est unique et vaut :

$$k_i = \|\mathbf{h}_i^T \mathbf{P}_{\nu_r}\|_1.$$

En conclusion, chaque bande de contrainte \mathcal{S}_i correspond à un $s_p - 1$ uplet de vecteurs ν_j d'une base orthonormée de \mathbb{R}^{s_ν} vérifiant :

$$\text{rang}\left(\mathbf{P}_{\nu_r} \begin{bmatrix} \nu_1 & \cdots & \nu_{s_p-1} \end{bmatrix}\right) = s_p - 1.$$

Il existe donc, au plus, $\mathbf{C}_{s_\nu}^{s_p-1}$ bandes de contraintes, soit $2\mathbf{C}_{s_\nu}^{s_p-1}$ facettes. Chacune des bandes de contraintes est définie par un vecteur ligne \mathbf{h}_i^T et un réel k_i donnés par :

$$\mathbf{h}_i^T \mathbf{P}_{\nu_r} \begin{bmatrix} \nu_1 & \cdots & \nu_{s_p-1} \end{bmatrix} = 0 \text{ et } k_i = \|\mathbf{h}_i^T \mathbf{P}_{\nu_r}\|_1. \quad (\text{IV-18})$$

La mise en œuvre est relativement simple. Le produit de \mathbf{P}_{ν_r} par un vecteur orthonormé ν_j donne l'une des colonnes de cette matrice. Le produit $\mathbf{P}_{\nu_r} \begin{bmatrix} \nu_1 & \cdots & \nu_{s_p-1} \end{bmatrix}$ est donc une matrice \mathbf{P} de $\mathbb{R}^{s_p \times s_p-1}$ contenant $s_p - 1$ des s_ν colonnes de \mathbf{P}_{ν_r} . Il suffit donc de déterminer toutes les combinaisons d'indices de $s_p - 1$ éléments parmi s_ν . Par exemple pour $s_\nu=3$ et $s_p=3$, les 2-uplets suivants sont obtenus : $\{1,2\}$, $\{1,3\}$ et $\{2,3\}$. Pour chaque combinaison d'indices, le rang de la matrice \mathbf{P} obtenue est évalué et s'il vaut $s_p - 1$, alors une nouvelle bande de contrainte \mathcal{S}_i est déterminée en calculant le vecteur ligne \mathbf{h}_i^T et le scalaire k_i grâce à (IV-18).

- (1) Déterminer la liste \mathcal{L} de l'ensemble des $\mathbf{C}_{s_\nu}^{s_p-1}$ combinaisons de $s_p - 1$ éléments parmi s_ν .
- (2) Prendre le premier élément de \mathcal{L} et l'éliminer de la liste. Construire la matrice \mathbf{P} obtenue en regroupant les colonnes de \mathbf{P}_{ν_r} , dont les indices correspondent à la combinaison considérée.
- (3) Tester le rang de la matrice \mathbf{P} . Si celui vaut $s_p - 1$, alors passer à l'étape suivante, sinon, aller à l'étape (5).
- (4) Calculer les éléments caractéristiques \mathbf{h}_i^T et k_i .
- (5) Si la liste \mathcal{L} est non-vide, aller à l'étape (1), dans le cas contraire sortir de l'algorithme en retournant tous les \mathbf{h}_i^T et k_i calculés.

Tableau IV-4. Squelette de l'algorithme proposé

Le vecteur ligne \mathbf{h}_i^T est obtenu en recherchant le noyau de la matrice précédente (par exemple numériquement à l'aide de l'instruction NULL sous Matlab). Une fois le vecteur ligne \mathbf{h}_i^T trouvé, k_i se déduit très facilement par un simple calcul d'une norme-1.

Exemple. Prenons par exemple $\mathbf{p}_0 = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$ et $\mathbf{P}_{v_r} = \begin{bmatrix} 1 & 1 & 1 \\ 1 & -1 & 2 \end{bmatrix}$, soit $s_p=2$ et $s_v=3$. Les

combinaisons d'un indice parmi trois donnent les "ensembles" $\{1\}$, $\{2\}$ et $\{3\}$. Chaque colonne de \mathbf{P}_{v_r} est de rang $s_p - 1 = 1$, le domaine abstrait est donc délimité par trois bandes de contraintes (soit six hyperplans au total) définis par :

$$\mathbf{h}_1^T \begin{bmatrix} 1 \\ 1 \end{bmatrix} = 0 \text{ soit } \mathbf{h}_1^T = [1 \quad -1] \text{ et } k_1=3, \text{ soit } -3 \leq [1 \quad -1]\mathbf{p}_{lin,r} \leq 3$$

$$\mathbf{h}_2^T \begin{bmatrix} 1 \\ -1 \end{bmatrix} = 0 \text{ soit } \mathbf{h}_2^T = [1 \quad 1] \text{ et } k_2=5, \text{ soit } -5 \leq [1 \quad 1]\mathbf{p}_{lin,r} \leq 5,$$

$$\mathbf{h}_3^T \begin{bmatrix} 1 \\ 2 \end{bmatrix} = 0 \text{ soit } \mathbf{h}_3^T = [2 \quad -1] \text{ et } k_3=4, \text{ soit } -4 \leq [-2 \quad 1]\mathbf{p}_{lin,r} \leq 4,$$

ce qui conduit à la figure suivante :

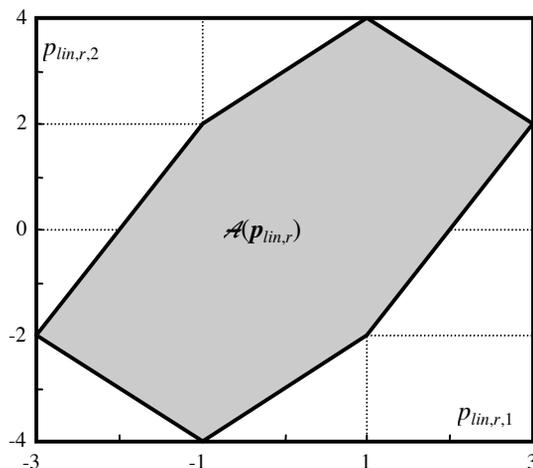


Figure IV-17. Espace abstrait $\mathcal{A}(\mathbf{p}_{lin,r})$

Globalement, la méthode exposée revient à dire que les hyperplans frontière sont obtenus :

- par élimination de certaines variables bornées en recherchant le vecteur de projection \mathbf{h}_i ,
- pour des valeurs appartenant à l'ensemble $\{-1,1\}$ pour les variables restantes.

L'analogie avec la méthode proposée pour la détermination des expressions analytiques des frontières (section IV.4) est alors simple à opérer, puisqu'elle procédait elle aussi de l'élimination de variables communes. Notons simplement que si celle-ci nécessitait d'opérer l'intersection du domaine obtenu avec l'orthotope aligné circonscrit, l'algorithme numérique précédent génère automatiquement les contraintes de ce dernier si nécessaire.

Remarque. L'annexe H développe une alternative de cette méthode dans le cas où la matrice P_{v_r} est inversible ou de plein rang colonne, ce qui conduit alors à une technique encore plus simple à mettre en œuvre puisque l'obtention des contraintes est alors directe.

IV.5.4 Notion de distance

Tester si l'origine O appartient à $\mathcal{A}(p_{lin,r})$ est simple puisqu'il suffit de vérifier si toutes les doubles inégalités (IV-13) sont satisfaites. Néanmoins, les résultats du test de cohérence sont binaires et peuvent paraître pauvres en considérant toute l'information dont nous disposons, comme par exemple la forme et la taille du paralléloétope $\mathcal{A}(p_{lin,r})$. L'objectif de cette section est, en plus de déterminer si l'origine O du repère associé à l'espace des résidus \mathbb{R}^{s_p} appartient ou non au domaine $\mathcal{A}(p_{lin,r})$, d'évaluer la distance séparant O de la frontière \mathcal{F}^* du paralléloétope, la plus proche de ce point. Ce résultat est ensuite divisé par la distance séparant le centre C (de coordonnées p_0) de $\mathcal{A}(p_{lin,r})$ et la même facette \mathcal{F}^* . Dans un certain sens, cette distance est ainsi normalisée puisque la "taille" du paralléloétope est prise en compte. Enfin, lorsque les mesures sont cohérentes, la distance obtenue est multipliée par -1 . Dans ces conditions, une distance de -1 signifie que l'origine est confondue avec le centre C , alors que la valeur 1 indique que O se trouve à l'extérieur de $\mathcal{A}(p_{lin,r})$ et à une égale distance de la plus proche frontière que C l'est de cette même frontière. Ainsi, plus cette distance notée par la suite d se rapproche de zéro, plus le système se trouve à la limite entre cohérence (fonctionnement supposé normal) et incohérence (présence d'un défaut). Finalement, une décision de cohérence sera assortie d'une distance par rapport à la décision d'incohérence ou inversement (Ploix et al, 2000), (Adrot et al, 2000b).

Nous noterons respectivement J^* et Q , les points de la frontière \mathcal{F}^* qui sont respectivement les plus proches, au sens de la norme 2, de l'origine O , et du centre C du paralléloétope (voir la figure suivante, sur laquelle les cercles apparaissent comme des ellipses du fait de l'échelle).

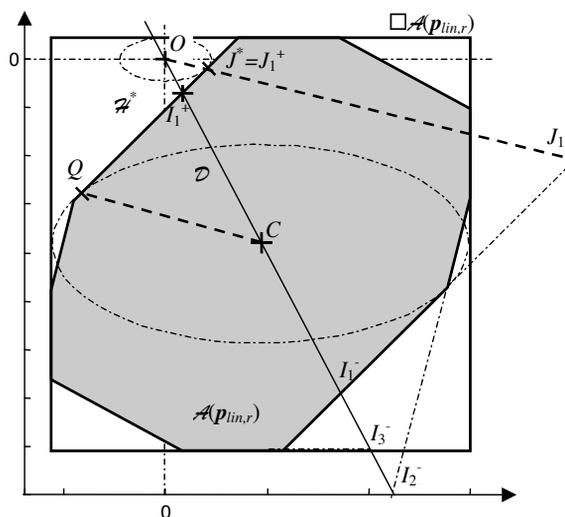


Figure IV-18. Calcul de la distance d

La distance d s'exprime alors de la manière suivante :

$$d = \frac{\|\overrightarrow{OJ^*}\|_2}{\|\overrightarrow{CQ}\|_2}$$

Pour rechercher la facette \mathcal{H}^* , tirons parti de la convexité du domaine $\mathcal{A}(\mathbf{p}_{lin,r})$. En premier lieu, intéressons-nous à la droite \mathcal{D} , reliant l'origine O au centre C du parallélotope. Nécessairement, elle coupe $\mathcal{A}(\mathbf{p}_{lin,r})$ en au moins deux¹⁵ hyperplans frontière \mathcal{H}_i^+ et \mathcal{H}_i^- opposés :

$$\mathcal{H}_i^+ = \{\mathbf{p}_{lin,r} / \mathbf{h}_i^T \mathbf{p} = \mathbf{h}_i^T \mathbf{p}_0 + k_i\} \text{ et } \mathcal{H}_i^- = \{\mathbf{p}_{lin,r} / \mathbf{h}_i^T \mathbf{p} = \mathbf{h}_i^T \mathbf{p}_0 - k_i\},$$

l'objectif étant d'en calculer les points d'intersection correspondants notés respectivement I_i^+ et I_i^- (la droite \mathcal{D} passant par le centre, il est exclu qu'elle puisse être contenue dans un hyperplan frontière ou ne pas couper le domaine considéré, en revanche elle peut être parallèle à certains d'entre eux).

Prenons par exemple le cas d'un hyperplan \mathcal{H} d'équation $\mathbf{h}^T \mathbf{p} = k$, où k est un réel et \mathbf{h}^T un vecteur ligne de dimension s_p . L'équation de la droite \mathcal{D} peut s'écrire sous la forme : $\mathbf{p} = \lambda \bar{\mathbf{u}} + \bar{\mathbf{v}}$, où λ est un réel non-nul, $\bar{\mathbf{v}}$ un vecteur de dimension s_p et $\bar{\mathbf{u}}$ son vecteur directeur de même dimension. Puisque la droite \mathcal{D} passe par l'origine, le vecteur $\bar{\mathbf{v}}$ est donc nul, d'autre part $\bar{\mathbf{u}} = \overrightarrow{OC} = \mathbf{p}_0$, ce qui donne : $\mathbf{p} = \lambda \mathbf{p}_0$. L'intersection de \mathcal{D} et de \mathcal{H} (si elle existe) est définie pour la valeur :

$$\lambda = \frac{k}{\mathbf{h}^T \mathbf{p}_0},$$

et le point d'intersection I résultant de l'intersection est alors donné par :

$$\overrightarrow{OI} = \frac{k}{\mathbf{h}^T \mathbf{p}_0} \mathbf{p}_0.$$

En appliquant ce résultat à chacun des hyperplans \mathcal{H}_i^+ et \mathcal{H}_i^- , les coordonnées des points I_i^+ et I_i^- sont obtenues :

$$\overrightarrow{OI_i^+} = \frac{\mathbf{h}_i^T \mathbf{p}_0 + k_i}{\mathbf{h}_i^T \mathbf{p}_0} \mathbf{p}_0, \quad \overrightarrow{OI_i^-} = \frac{\mathbf{h}_i^T \mathbf{p}_0 - k_i}{\mathbf{h}_i^T \mathbf{p}_0} \mathbf{p}_0.$$

La recherche de l'hyperplan \mathcal{H}^* le plus proche de l'origine commence par la recherche du couple d'hyperplans parallèles \mathcal{H}_i^{+*} et \mathcal{H}_i^{-*} , auquel appartient \mathcal{H}^* . Du fait de la convexité du domaine $\mathcal{A}(\mathbf{p}_{lin,r})$, \mathcal{H}^* correspond nécessairement à l'un des deux l'hyperplans \mathcal{H}_i^{+*} ou \mathcal{H}_i^{-*} pour lequel la distance entre le centre C et les points I_i^+ et I_i^- est la plus petite. Puisque C est le centre du parallélotope, les deux hyperplans \mathcal{H}_i^{+*} et \mathcal{H}_i^{-*} sont aussi proches de C :

¹⁵ Dans le cas particulier où \mathcal{D} passe par deux sommets opposés du parallélotope, ce nombre est plus important.

$$\left\| \overrightarrow{CI_i^+} \right\|_2 = \left\| \mathbf{p}_0 - \frac{\mathbf{h}_i^T \mathbf{p}_0 + k_i}{\mathbf{h}_i^T \mathbf{p}_0} \mathbf{p}_0 \right\|_2 = \left\| \frac{k_i}{\mathbf{h}_i^T \mathbf{p}_0} \mathbf{p}_0 \right\|_2 = \left\| \overrightarrow{CI_i^-} \right\|_2.$$

L'indice i^* vérifie donc :

$$i^* = \arg \left(\min_i \left(\frac{k_i}{|\mathbf{h}_i^T \mathbf{p}_0|} \right) \right),$$

où k_i est un terme positif par définition (IV-18). Deux hyperplans sont donc potentiellement solution. Pour remonter à \mathcal{H}^* , seul celui qui est le plus proche de l'origine O est conservé. Pour ce faire, O est projeté orthogonalement sur ces deux hyperplans, ce qui donne respectivement les points $J_{i^*}^+$ et $J_{i^*}^-$ (plus exactement J_1^+ et J_1^- sur la figure précédente) puis les distances entre ces points et l'origine O sont évaluées.

Trouvons dans un premier temps J , la projection orthogonale de l'origine O sur un hyperplan \mathcal{H} d'équation $\mathbf{h}^T \mathbf{p} = k$. Le point J appartient à la droite d'équation $\mathbf{p} = \lambda \bar{\mathbf{u}} + \bar{\mathbf{v}}$ avec $\bar{\mathbf{v}} = 0$ et de vecteur directeur $\bar{\mathbf{u}} = \mathbf{h}$, soit $\mathbf{p} = \lambda \mathbf{h}$, ce qui conduit à :

$$\lambda = \frac{k}{\mathbf{h}^T \mathbf{h}}, \text{ puis finalement } \overrightarrow{OJ} = \frac{k}{\mathbf{h}^T \mathbf{h}} \mathbf{h}.$$

En appliquant ce résultat aux calculs de distances entre l'origine et les points $J_{i^*}^+$ et $J_{i^*}^-$, il se déduit que :

$$\left\| \overrightarrow{OJ_{i^*}^+} \right\|_2 = \frac{|\mathbf{h}_{i^*}^T \mathbf{p}_0 + k_{i^*}|}{\|\mathbf{h}_{i^*}\|_2}, \quad \left\| \overrightarrow{OJ_{i^*}^-} \right\|_2 = \frac{|\mathbf{h}_{i^*}^T \mathbf{p}_0 - k_{i^*}|}{\|\mathbf{h}_{i^*}\|_2}.$$

La plus petite de ces deux distances détermine l'hyperplan \mathcal{H}^* . Nous noterons J^* la projection orthogonale de O sur \mathcal{H}^* .

Enfin, il ne reste plus qu'à diviser $\left\| \overrightarrow{OJ^*} \right\|_2$ par $\left\| \overrightarrow{CQ} \right\|_2$, Q étant la projection orthogonale de C sur \mathcal{H}^* . Trouvons dans un premier temps Q , la projection orthogonale du centre sur un hyperplan \mathcal{H} d'équation $\mathbf{h}^T \mathbf{p} = k$. Ce point appartient à la droite d'équation $\mathbf{p} = \lambda \bar{\mathbf{u}} + \bar{\mathbf{v}}$ avec $\bar{\mathbf{v}} = \mathbf{p}_0$ et de vecteur directeur $\bar{\mathbf{u}} = \mathbf{h}$, soit $\mathbf{p} = \lambda \mathbf{h} + \mathbf{p}_0$, ce qui conduit à :

$$\lambda = \frac{k - \mathbf{h}^T \mathbf{p}_0}{\mathbf{h}^T \mathbf{h}} \text{ et } \overrightarrow{OQ} = \frac{k - \mathbf{h}^T \mathbf{p}_0}{\mathbf{h}^T \mathbf{h}} \mathbf{h} + \mathbf{p}_0, \text{ soit } \overrightarrow{CQ} = \frac{k - \mathbf{h}^T \mathbf{p}_0}{\mathbf{h}^T \mathbf{h}} \mathbf{h}.$$

Nous obtenons donc :

$$\overrightarrow{CQ} = \frac{k_{i^*}}{\mathbf{h}_{i^*}^T \mathbf{h}_{i^*}} \mathbf{h}_{i^*} \text{ pour } \mathcal{H}_{i^*}^+,$$

$$\overrightarrow{CQ} = \frac{-k_{i^*}}{\mathbf{h}_{i^*}^T \mathbf{h}_{i^*}} \mathbf{h}_{i^*} \text{ pour } \mathcal{H}_{i^*}^-,$$

$$\text{ce qui donne : } \|\overrightarrow{CQ}\|_2 = \frac{k_{i^*}}{\|\mathbf{h}_{i^*}\|_2}.$$

La distance d , au signe près qui dépend de la position de l'origine par rapport au polytope est donc :

$$d = \frac{\min\left(\left|\mathbf{h}_{i^*}^T \mathbf{p}_0 + k_{i^*}\right|, \left|\mathbf{h}_{i^*}^T \mathbf{p}_0 - k_{i^*}\right|\right)}{k_{i^*}} \text{ avec } i^* = \arg\left(\min_i \left(\frac{k_i}{\left|\mathbf{h}_i^T \mathbf{p}_0\right|}\right)\right).$$

Remarquons que par définition, le réel k_{i^*} (IV-18) ne peut être nul, car cela signifierait que le produit $\mathbf{h}_{i^*}^T \mathbf{p}_{v_r}$ vaut 0, ce qui est impossible puisque la matrice \mathbf{P}_{v_r} est supposée de plein rang ligne.

Notons pour terminer que, puisqu'aucune loi de distribution n'est associée initialement aux incertitudes paramétriques, une distance ne peut être vue en termes de probabilité. Plus précisément, il n'est pas possible de savoir si un petit ou gros défaut est nécessaire pour que l'origine atteigne la limite de bon fonctionnement \mathcal{Z}^* . La notion de distance (mis à part son signe) ne doit pas être analysée ponctuellement, mais de manière plus globale en observant son évolution. Ainsi, une dérive lente vers la valeur nulle (tout en restant négative) peut témoigner d'une dérive progressive d'un capteur par exemple. Le rapprochement soudain vers cette valeur peut aussi correspondre à la présence d'un défaut encore insuffisamment important pour ne plus être masqué par les incertitudes. Cet aspect sera plus particulièrement repris dans la section suivante.

IV.5.5 Extension en vue d'une procédure de localisation

L'objectif de cette section est d'étendre la procédure de détection de défauts afin de parvenir à localiser une anomalie une fois sa présence signalée (Adrot et al , 2000b), tout en insistant sur les difficultés rencontrées dans le cadre de l'approche bornante. Cette étude se limite au contexte où un seul défaut peut affecter le système à un instant donné, celui-ci ne pouvant apparaître que sur la chaîne d'instrumentation. Lorsque le test de cohérence défini dans la section IV.5.3 n'est plus satisfait, une procédure de localisation est lancée en répertoriant les contraintes inégalité définissant $\mathcal{A}(\mathbf{p}_{lin,r})$ non-vérifiées, afin de déterminer le capteur en défaut. Puisque l'objectif est d'utiliser toute l'information à notre disposition, il semble naturel d'intégrer dans notre démarche la prise en compte du respect ou non des bandes de contraintes \mathcal{S}_i . Une table de signatures théoriques \mathcal{S}_i est générée, non plus associée directement aux relations de parité, mais plutôt aux contraintes inégalités définissant la forme du domaine $\mathcal{A}(\mathbf{p}_{lin,r})$. Plus exactement, il existe toujours au moins s_p bandes de contraintes, en l'occurrence celles associées à chaque composante du vecteur résidu $\mathbf{p}_{lin,r}$ et caractérisant son orthotope aligné circonscrit $\square \mathcal{A}(\mathbf{p}_{lin,r})$. Comme expliqué dans la section IV.5.3, ce nombre augmente avec la présence de variables communes puisque leur élimination génère de nouvelles bandes de contraintes. De manière générale, le parallélotope $\mathcal{A}(\mathbf{p}_{lin,r})$ conduit à plus

de contraintes inégalité que $\square \mathcal{A}(\mathbf{p}_{lin,r})$ et peut donc permettre de localiser un nombre plus important de défauts différents. Il est bien sur possible de n'utiliser que les s_p bandes de contraintes associées à $\square \mathcal{A}(\mathbf{p}_{lin,r})$, mais dans ce cas, une partie de l'information est perdue puisque ce domaine se trouve moins précis que $\mathcal{A}(\mathbf{p}_{lin,r})$. Ainsi, un défaut peut être détecté, mais non localisé si l'origine se situe en dehors de $\mathcal{A}(\mathbf{p}_{lin,r})$, mais à l'intérieur de $\square \mathcal{A}(\mathbf{p}_{lin,r})$. La table des signatures théorique \mathcal{S}_t est construite en indiquant par un bit 1, quelles contraintes sont sensibles à chaque défaut envisagé (ou de manière équivalente à chaque capteur surveillé).

Un défaut est détecté lorsque l'origine n'appartient plus au parallélotope $\mathcal{A}(\mathbf{p}_{lin,r})$, en d'autres termes quand au moins une contrainte inégalité n'est pas vérifiée. Dans ces conditions, la procédure de localisation consiste à déterminer quelles bandes de contraintes ne sont pas satisfaites. En codant respectivement avec un 0 et un 1 les bandes de contraintes satisfaites et celles qui ne le sont pas, la signature expérimentale \mathcal{S}_d du défaut détecté est déterminée, puis peut être comparée avec la table des signatures théoriques \mathcal{S}_t .

Généralement, la principale difficulté réside dans l'établissement de \mathcal{S}_t tout simplement à cause de la nature non-linéaire des relations de parité $\mathbf{p}_{lin,r}$ vis-à-vis des mesures. Ce problème prend son origine au niveau de la procédure d'élimination des variables communes lors de la génération des contraintes inégalité représentatives de $\mathcal{A}(\mathbf{p}_{lin,r})$. Le bit 1 indique normalement qu'une bande de contrainte \mathcal{S}_i est théoriquement sensible à une mesure donnée, donc à un défaut pouvant affecter le capteur correspondant. Mais, cette sensibilité, qui habituellement est déduite d'une étude structurelle des résidus linéaires en les mesures, va dépendre ici des valeurs numériques des mesures en question.

Pour clarifier ce problème, considérons le système statique suivant

$$\begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix} = \begin{bmatrix} 1 + 0.1\theta_1 \\ 1 + 0.1\theta_2 \\ 1 + 0.1\theta_3 \end{bmatrix} x, \text{ et la matrice de parité } \mathbf{W}(\boldsymbol{\theta}) = \begin{bmatrix} -(1 + 0.1\theta_2) & 1 + 0.1\theta_1 & 0 \\ -(1 + 0.1\theta_3) & 0 & 1 + 0.1\theta_1 \end{bmatrix},$$

conduisant aux relations de parité $\mathbf{p}_{lin,r}$:

$$\mathbf{p}_{lin,r} = \begin{bmatrix} p_1 \\ p_2 \end{bmatrix} = \begin{bmatrix} -(1 + 0.1\theta_2)y_1 + (1 + 0.1\theta_1)y_2 \\ -(1 + 0.1\theta_3)y_1 + (1 + 0.1\theta_1)y_3 \end{bmatrix}.$$

Si l'orthotope aligné circonscrit $\square \mathcal{A}(\mathbf{p}_{lin,r})$ était étudié, ces deux relations de parité définiraient deux bandes de contraintes (soit quatre contraintes inégalité au total). En éliminant des expressions analytiques précédentes la variable commune θ_1 tel que proposé dans la section IV.4, la relation suivante est déduite :

$$p_3 = y_3 p_1 - y_2 p_2 = y_1 \{ (1 + 0.1\theta_3)y_2 - (1 + 0.1\theta_2)y_3 \}.$$

Puisqu'un seul défaut peut se produire à la fois, lorsque le premier capteur se comporte anormalement, il est possible d'écrire :

$$p_3 = y_1 (1 + 0.1\theta_3)(1 + 0.1\theta_2)(x - x) = 0. \quad (\text{IV-19})$$

Notons que pour déterminer la table de signatures théorique, il n'est pas nécessaire de calculer les deux inégalités associées à chaque bande de contraintes en évaluant les bornes inférieure et supérieure des résidus p_i , $i \in \{1, \dots, 3\}$; les égalités précédentes sont suffisantes pour indiquer les mesures dont dépend chaque résidu p_i .

L'élimination du paramètre incertain θ_1 se traduit par l'élimination de la mesure associée y_1 au niveau de p_3 , en effet, lorsqu'aucun défaut n'affecte les deux dernières équations de mesure, ce résidu vaut 0 et ce quelle que soit la valeur de y_1 . De manière générale, p_3 dépend structurellement de y_1 , mais n'est cependant pas sensible à un défaut pouvant affecter cette mesure. Si en revanche un défaut se produit sur y_2 ou sur y_3 , l'égalité précédente (IV-19) devient fausse. En effet, en supposant que la seconde mesure est biaisée (au choix un défaut de type additif b ou multiplicatif b_1) :

$$y_2 = (1 + 0.1\theta_2 + b_1)x + b,$$

l'égalité (IV-19) devient alors :

$$p_3 = y_3 p_1 - y_2 p_2 = y_1 (1 + 0.1\theta_3)(b_1 x + b).$$

Cette relation dépend bien de y_1 , et montre que les deux contraintes inégalité associées à p_3 sont de manière générale sensibles à un défaut sur les deux dernières équations de mesures (à moins que y_1 ne soit nulle).

	p_1	p_2	p_3
y_1	1	1	0
y_2	1	0	1
y_3	0	1	1

Tableau IV-5. Signatures théoriques

De manière générale, cette difficulté peut s'aggraver à cause de la procédure de linéarisation, qui modifie les relations de redondance en ajoutant de nouvelles variables bornées. Il devient alors difficile d'opérer le même cheminement que dans ce petit exemple parce que les relations initiales du modèle ne correspondent plus exactement avec les équations linéarisées des résidus. Il peut alors devenir plus délicat d'opérer les bonnes transformations pour observer si un résidu est sensible ou non réellement à un défaut sur un capteur donné.

En fait, les signatures théoriques reposant sur l'étude structurelle de $p_{lin,r}$, permettent d'obtenir une première table de base (dans laquelle la case (y_1, p_3) prendrait la valeur 1), qu'il est possible d'améliorer par une étude expérimentale afin de contrer le problème précédent. Le modèle, dont les "capteurs" sont successivement affectés par des biais importants, est simulé de manière à ce que les signatures obtenues permettent de corriger celles théoriques. Ainsi, à cause de la nature non-linéaire de $p_{lin,r}$ vis-à-vis des mesures, certains bits 1 des signatures théoriques peuvent être incorrects et doivent être remplacés par un 0. De plus notons que puisque les deux premiers résidus sont linéaires en les mesures, l'étude structurelle est suffisante dans leur cas. La sensibilité aux défauts ne dépend alors que de la

structure de ces expressions, les deux premières colonnes du tableau précédent ne seront donc pas modifiées.

La troisième contrainte n'apporte que peu de choses a priori puisque l'utilisation des deux premiers résidus est suffisante pour discerner le capteur en défaut. Cependant, en se reposant sur la propriété de garantie, il n'est pas possible, qu'un défaut donné se traduise par une signature expérimentale coïncidant avec la signature théorique d'un autre défaut. Il faudrait, pour que cela se produise, qu'un 1 de \mathcal{S}_d soit comptabilisé comme un 0, chose tout à fait possible à cause des incertitudes pouvant masquer un défaut, mais aussi qu'un 0 devienne un 1, fait qui ne peut se produire. L'approche bornante assure (si le modèle est juste) qu'une contrainte indépendante du défaut présent ne peut être que vérifiée. De ce fait, il est possible de ne pas être en mesure de localiser à cause d'une signature \mathcal{S}_d incompatible avec \mathcal{S}_i , mais confondre deux défauts dans le cadre de cet exemple est impossible. En revanche, si le troisième résidu p_3 n'était pas pris en compte, la signature (1,1) associée à y_1 pourrait se transformer en (1,0) ou (0,1) entraînant une mauvaise localisation.

Tout comme pour la procédure de détection, un défaut sur le $i^{\text{ème}}$ capteur peut être masqué par les incertitudes. De ce fait, sa signature expérimentale \mathcal{S}_d peut contenir au moins un bit 0 à un endroit où la signature théorique associée indique un bit 1¹⁶. Une solution pour compenser ce problème est d'essayer de corriger la signature \mathcal{S}_d lorsqu'elle ne correspond à aucune signature théorique. L'idée est de rechercher la contrainte vérifiée correspondant à la plus grande distance normalisée (négative comme définie dans la section précédente), c'est-à-dire à trouver celle qui est satisfaite avec la plus petite marge, soit celle qui logiquement a le plus de chance de ne plus être satisfaite si l'amplitude du défaut augmentait quelque peu¹⁷. Alors, le bit 0 associé au niveau de \mathcal{S}_d est remplacé par un 1 et la bande de contraintes correspondante est alors considérée comme non satisfaite. Contrairement à la procédure de détection, les distances relatives à toutes les facettes de $\mathcal{A}(p_{lin,r})$ doivent être calculées. Dans des cas simples où les signatures théoriques contiennent le même nombre p_{d1} de bits 1, cette correction peut être relancée tant que le nombre p_{d1} de 1 présents dans \mathcal{S}_d lui est inférieur. Notons que dans notre cas, nous disposons de plus de résidus que de capteurs grâce aux contraintes imposées par l'élimination de variables communes (ce qui ne signifie nullement d'être capable de différencier chaque capteur), il est alors légitime de sélectionner en priorité ceux qui conduisent à des signatures théoriques contenant le même nombre de bits 1, à la condition bien sûr qu'ils suffisent à localiser l'ensemble des défauts du cahier des charges. Dans le cas contraire, la correction proposée est moins efficace, bien qu'il soit toujours possible de corriger tant que la signature expérimentale ne correspond à aucune signature théorique ou tant que p_{d1} reste inférieur à un nombre fixé par l'utilisateur. Notons cependant que cette correction ne permet plus de garantir les résultats de la procédure de

¹⁶ Notons que si le modèle est initialement juste, grâce à la propriété de garantie de l'analyse par intervalles conservée lors des procédures précédentes (linéarisation, réduction, génération des contraintes d'un parallélotope, notion de distance), l'inverse ne peut se produire puisqu'une contrainte non satisfaite, dépend nécessairement d'un capteur en défaut. En conclusion, un bit 1 au niveau d'une signature expérimentale est garanti dans le sens où il ne peut signifier un 0.

¹⁷ Aucune loi de distribution n'étant imposée sur les variables bornées, il serait faux de parler de probabilité de ne plus être satisfaite.

localisation, puisque des contraintes vérifiées sont déclarées insatisfaites. La contrainte vérifiée la plus proche de l'origine n'est pas forcément sensible au défaut envisagé, elle peut en être proche simplement à cause du point de fonctionnement courant (rappelons qu'en bon fonctionnement, l'origine O peut se trouver à tout moment à n'importe quel endroit du domaine $\mathcal{A}(p_{lin,r})$). D'autres solutions plus classiques peuvent consister à rechercher la signature théorique la plus proche de la signature expérimentale. Dernier point, puisque sans correction, la signature expérimentale est garantie, un bit 1 trouvé signifie obligatoirement que la contrainte considérée dépend du défaut incriminé. De ce fait, en cas de non concordance entre \mathcal{S}_d et \mathcal{S}_t , il est déjà possible d'établir un premier tri des signatures théoriques en éliminant toutes celles qui contiennent un bit 0 à cet emplacement.

Il est tout à fait possible de raisonner de la même façon avec des défauts survenant sur les paramètres du système proprement dit, la difficulté supplémentaire étant de différencier l'influence d'une incertitude paramétrique sortant de ses bornes (par exemple θ_1) de la présence d'un défaut sur une mesure (par exemple un biais sur y_1).

IV.6 CONCLUSION

En premier lieu, le principe général des tests de cohérence dans le contexte d'un modèle incertain a été présenté et la solution adoptée consiste à tester si l'origine du repère dans l'espace des résidus se situe à l'intérieur de l'espace abstrait associé au vecteur résidu. Nous nous sommes donc tout naturellement penchés sur la manière de déterminer ce domaine, sachant que les expressions obtenues dans le troisième chapitre induisaient de manière générale des domaines complexes, difficiles à évaluer. Deux méthodes ont tout particulièrement été développées : une méthode de pavage adaptée aux objectifs du diagnostic et une autre procédant de la linéarisation des relations de parité par rapport aux variables bornées.

La première stratégie ne nécessite pas la construction complète de l'espace abstrait considéré, mais seulement d'une portion proche du point de fonctionnement courant testé. La seconde stratégie conduit à la génération de contraintes inégalité représentatives d'une majoration du domaine exact de bon fonctionnement (à moins que les relations soient déjà linéaires) sous la forme d'un parallélotope. Il n'est alors pas nécessaire de construire le domaine obtenu, la seule connaissance des contraintes inégalité étant suffisante à la réalisation des tests de cohérence. Ces deux stratégies restent systématiques quel que soit le système étudié, sans aucune hypothèse sur les relations de parité à traiter. La première est réalisée de manière totalement numérique, et reste très dépendante du nombre de variables bornées présentes. La seconde s'appuie sur des procédures numériques et symboliques à la fois. La procédure de linéarisation doit s'opérer à priori, car elle fonctionne par regroupement de variables bornées dont les valeurs sont inconnues. La procédure de réduction telle qu'elle est présentée peut être numérique (auquel cas elle n'est appliquée qu'une fois les mesures obtenues) ou symbolique (auquel cas elle est réalisée a priori). En revanche, les procédures

qui suivent (génération de contraintes, calcul de distances) ont été traitées de manière complètement numérique et sont donc à effectuer à chaque instant. Une extension à la localisation de défauts est aussi proposée.

Quelle que soit la méthode utilisée, l'approche bornante permet la génération automatique de "seuils" de détection sur la seule connaissance du modèle, sans aucune hypothèse probabiliste. L'enveloppe multidimensionnelle calculée est générée en tenant compte des dépendances entre variables bornées, ou d'une partie de ces dépendances, toujours dans un souci de perdre un minimum d'information risquant d'engendrer un nombre plus important de non-détections. Hormis pour la méthode de localisation corrigée, la propriété de garantie est à chaque fois conservée.

Chapitre V

Caractérisation de variables bornées

TABLE DES MATIÈRES

V.1 INTRODUCTION	215
V.2 POSITION DU PROBLÈME.....	216
V.2.1 Principe de l'approche géométrique	216
V.2.2 Première étape	218
V.2.3 Seconde étape	220
V.2.3.1 Principe	220
V.2.3.2 Mise en œuvre.....	224
V.3 CARACTÉRISATION	226
V.3.1 Critère de précision.....	226
V.3.2 Forme prédéterminée.....	228
V.3.3 Forme indéterminée.....	229
V.3.4 Optimisation simultanée du centre et de la borne de l'erreur d'équation.....	230
V.4 MULTI-CARACTÉRISATION.....	234
V.4.1 Problème de dépendance	234
V.4.2 Solution proposée	238
V.5 CONCLUSION.....	243

V.1 INTRODUCTION

L'objectif de ce chapitre est de développer des méthodes de caractérisation permettant la détermination des bornes des incertitudes paramétriques d'un modèle. La notion de caractérisation est aux variables bornées dépendantes du temps ce que l'identification est aux paramètres invariants. Sommairement, le problème de l'estimation paramétrique est de rechercher les valeurs des paramètres pour lequel le modèle parvient à expliquer le mieux, au sens d'un critère choisi, l'ensemble des observations disponibles. L'approche bornante appliquée à la théorie de l'estimation paramétrique, alors qualifiée d'ensémbliste, a vu le jour au cours des années 80 comme explicité dans la section II.1. Elle repose généralement sur un modèle linéaire en les paramètres et caractérisé par une erreur d'équation bornée. L'objectif consiste alors à déterminer l'ensemble des valeurs acceptables des paramètres, c'est-à-dire l'ensemble de leurs valeurs cohérentes avec toutes les mesures à la fois, le modèle et les bornes de l'erreur d'équation (Milanese et al, 1996a). Suivant la stratégie et le modèle utilisés, le domaine peut être recherché de manière exacte, ou approché par une forme plus simple tels un ellipsoïde ou une boîte. La différence entre la notion de caractérisation introduite par Ploix (1998) et l'identification de manière générale, malgré une certaine ressemblance, tient du fait que dans le premier cas, les valeurs des paramètres ne sont pas tenues pour constantes, mais fluctuent dans un domaine invariant (telle un polytope convexe dans notre cas).

Ce chapitre se limite à l'étude de modèles linéaires en les variables bornées, que ce soit des incertitudes affectant les paramètres du modèle ou l'erreur d'équation. L'erreur additive d'équation est supposée appartenir à un intervalle tandis que les paramètres fluctuent à l'intérieur d'un domaine borné invariant dans le temps, l'objectif étant de le déterminer. L'intérêt d'une procédure de caractérisation est multiple. La discrétisation d'un modèle continu incertain engendre des approximations qu'il est nécessaire de corriger par des termes compensateurs ajustés au mieux par expérience. Il est alors difficile de définir avec certitude si ces derniers sont correctement réglés de manière à ce que le modèle obtenu englobe l'ensemble des comportements possibles du système physique en bon fonctionnement. Ceci est d'autant plus vrai que pour simplifier, ces termes compensateurs peuvent n'intervenir que de manière additive alors qu'ils doivent compenser des approximations pouvant affecter des incertitudes multiplicatives. Une solution pour limiter ce problème lié à la discrétisation d'un modèle incertain est de directement caractériser les bornes des incertitudes du modèle discret. Il est en fait difficile de s'intéresser directement à un modèle sous la forme de sa représentation d'état à cause de l'horizon temporel infini (Ploix 1998), d'où l'idée de caractériser uniquement les relations de parité obtenues après linéarisation et réduction. Cette stratégie peut en outre permettre d'atténuer la majoration effectuée lors de la linéarisation en adaptant les bornes des incertitudes aux nouvelles relations de redondance. Notons que dans ce cas, il s'agit d'améliorer la précision de la représentation, qui doit rester juste. Un autre avantage est qu'en cas de surparamétrisation (liée au choix des paramètres incertains au niveau du modèle initial, à la linéarisation,...), la procédure de caractérisation peut conduire à

l'élimination de certaines incertitudes si les poids attribués à chaque variable bornée normalisée peuvent être différents.

La seconde section présente les objectifs d'une procédure de caractérisation dans le cas d'un modèle linéaire en les variables bornées, en insistant tout particulièrement sur les différences avec l'estimation paramétrique ensembliste. Si l'approche géométrique exposée n'est pas nouvelle, elle permet néanmoins de facilement présenter les principes d'une procédure de caractérisation, ainsi que les difficultés rencontrées et les outils de l'analyse par intervalles utilisés. De plus, elle permet d'exhiber la manière de régler certains paramètres (domaine d'étude, valeur centrale, borne de l'erreur d'équation) en utilisant l'aspect visuel qui en facilite la compréhension. La troisième section développe une méthode de caractérisation plus simple que l'approche géométrique, car ne nécessitant ni la construction de polyèdres, ni la connaissance de leurs sommets. Présentée dans le cas d'une forme prédéterminée, c'est-à-dire où la caractérisation ne peut modifier que la taille de l'espace abstrait des incertitudes paramétriques, cette méthode est ensuite étendue au cas d'une forme indéterminée. Les longueurs de chaque variable bornée peuvent alors être différentes et la procédure de caractérisation impose ainsi d'elle-même la forme de l'espace abstrait étudié. Un critère est ensuite défini de manière à optimiser la précision du modèle caractérisé en permettant le réglage du centre de ce domaine ainsi que des bornes de l'erreur d'équation additive. Enfin la quatrième section reprend ce qui a été précédemment développé pour une forme prédéterminée et l'étend à un modèle possédant plusieurs équations entrée-sortie. Il est alors nécessaire de prendre en compte les dépendances entre variables communes afin d'obtenir un modèle juste pour le jeu de données servant à la caractérisation.

V.2 POSITION DU PROBLÈME

V.2.1 Principe de l'approche géométrique

L'objectif de cette section est de présenter le principe d'une méthode permettant de caractériser les bornes de paramètres incertains dépendant du temps dans le cadre d'un modèle mono-sortie, linéaire, discret et à horizon fini de la forme suivante :

$$y(k) = \boldsymbol{\varphi}^T(k)\boldsymbol{\theta}(k) + e(k), \quad y \in \mathbb{R}, \quad e \in \mathbb{R}, \quad \boldsymbol{\theta} \in \mathbb{R}^r, \quad \boldsymbol{\varphi} \in \mathbb{R}^r, \quad (\text{V-1})$$

où $y(k)$ est la sortie mesurée obtenue à l'instant k , $\boldsymbol{\varphi}(k)$ est un vecteur contenant d'autres mesures (entrées commandées ou sorties) et $\boldsymbol{\theta}(k)$ est le vecteur des paramètres incertains dépendant du temps. Le scalaire $e(k)$, définissant l'erreur d'équation additive entre la nouvelle mesure $y(k)$ et la sortie du modèle $\boldsymbol{\varphi}^T(k)\boldsymbol{\theta}(k)$, est une variable dont les bornes¹ sont choisies symétriques :

$$-\bar{e} \leq e(k) \leq \bar{e}.$$

¹ Pour l'instant, les valeurs de ces bornes sont supposées connues, mais nous verrons ultérieurement que dans le cas contraire, il est possible d'intégrer leur recherche dans la procédure de caractérisation.

Cette relation peut aussi bien représenter des modèles statiques que dynamiques si φ contient des mesures ainsi que des commandes retardées, et que θ est constitué de paramètres incertains exprimés à des instants antérieurs à k . En revanche, pour l'instant nous nous limiterons à un système mono-sortie ne contenant qu'une seule relation entrée-sortie.

Dans le cas où les paramètres incertains (y compris l'erreur d'équation) sont supposés invariants, la détermination du vecteur θ peut se mener de manière classique par des techniques d'identification paramétrique comme celle des Moindres Carrés, le modèle étant linéaire par rapport aux paramètres. Le fait de supposer que les paramètres incertains sont bornés n'apporte alors strictement rien, ce type de situation ne nous intéresse donc pas. En revanche, s'ils varient au cours du temps, le problème est tout autre. Les bornes inférieures et supérieures $\underline{\theta}$ et $\bar{\theta}$ du vecteur $\theta(k)$, qui contiennent l'ensemble des valeurs permettant d'expliquer toutes les mesures $y(k)$, $k \in \{1, \dots, t\}$, à une erreur d'équation près $e(k)$ bornée, sont alors recherchées. De manière plus générale, l'objectif est de déterminer un domaine \mathcal{D}_t dans lequel se retrouvent différentes valeurs de $\theta(k)$, chacune d'entre elles assurant qu'à un instant k donné, $\varphi^T(k)\theta(k)$ correspond à la mesure $y(k)$ à l'erreur d'équation près comprise entre $\pm \bar{e}$:

$$\begin{cases} y(k) \leq \varphi^T(k)\bar{\theta} + \bar{e} \\ \varphi^T(k)\underline{\theta} - \bar{e} \leq y(k) \end{cases} \quad (\text{V-2})$$

Pour déterminer les solutions $\underline{\theta}$ et $\bar{\theta}$ de ce problème, un algorithme proche de ceux proposés par (Walter et al, 1987) ou (Mo et al, 1988, 1990) et adapté à nos objectifs est employé. Si la technique mise en œuvre est identique initialement, les auteurs s'attachaient à déterminer l'ensemble des valeurs acceptables des paramètres (*Feasible Parameter Set*), c'est-à-dire l'ensemble des valeurs θ_{fps} du vecteur θ , supposé invariant, qui expliquent toutes les mesures en même temps. Une telle valeur explique chacune des mesures $y(k)$, $k \in \{1, \dots, t\}$, alors que dans notre cas, le vecteur inconnu $\theta(k)$ évolue et prend différentes valeurs tout en satisfaisant les contraintes sur l'erreur d'équation. C'est la détermination de ces valeurs qui va nous permettre de caractériser les bornes du vecteur de paramètres incertains.

A partir des contraintes imposées sur $e(k)$, $2t$ contraintes inégalité sont obtenues :

$$\underline{b}(k) \leq \varphi^T(k)\theta(k), \text{ avec } \underline{b}(k) = y(k) - \bar{e}, \quad (\text{V-3})$$

$$\varphi^T(k)\theta(k) \leq \bar{b}(k), \text{ avec } \bar{b}(k) = y(k) + \bar{e}. \quad (\text{V-4})$$

Le principe de la méthode de caractérisation proposée s'appuie sur deux étapes successives. Sommairement, la première consiste à déterminer les deux domaines de valeurs acceptables en les paramètres vérifiant respectivement les inégalités (V-3) et (V-4), c'est-à-dire les deux polytopes représentant les domaines de valeurs des paramètres invariants $\underline{\theta}$ et $\bar{\theta}$ respectant respectivement les inégalités suivantes pour $k \in \{1, \dots, t\}$:

$$\underline{b}(k) \leq \varphi^T(k)\underline{\theta}, \text{ avec } \underline{b}(k) = y(k) - \bar{e}, \quad (\text{V-5})$$

$$\varphi^T(k)\bar{\theta} \leq \bar{b}(k), \text{ avec } \bar{b}(k) = y(k) + \bar{e}. \quad (\text{V-6})$$

La distinction entre $\underline{\theta}$ et $\bar{\theta}$ vient du fait qu'il n'est pas supposé qu'une même valeur du vecteur de paramètres puisse à la fois vérifier les t contraintes (V-5) et les t contraintes (V-6).

Cette particularité sera amplement développée ultérieurement ; une première justification provenant de la relation (V-2) où les vecteurs de bornes $\underline{\theta}$ et $\bar{\theta}$ n'ont aucune raison d'être identiques.

Ensuite, après s'être muni d'un estimateur "central" invariant reflétant la valeur nominale des paramètres, la seconde étape consiste à rejeter en partie les erreurs d'équation sur ces derniers, leur conférant ainsi leur nature incertaine. Les incertitudes à ajouter sont alors déterminées de manière à ce que toutes les mesures soient expliquées, c'est-à-dire qu'il existe toujours une valeur $\theta(k)$, $k \in \{1, \dots, t\}$, appartenant au domaine solution invariant, noté \mathcal{D}_t , tel que les relations (V-3) et (V-4) soient satisfaites. L'intérêt de présenter d'un point de vue géométrique cette méthode est d'en faire ressortir plus simplement le principe et les particularités. Sachant qu'elle a déjà été détaillée dans Ploix (1998), nous n'en présenterons ici que les grandes lignes, l'objectif étant, dans les sections suivantes, de proposer des méthodes moins complexes à mettre en œuvre ou s'attachant à des modèles plus généraux (comme des modèles multi-sorties).

V.2.2 Première étape

La première étape procède, d'une certaine manière, comme si le vecteur $\theta(k)$ était invariant. Comme explicité dans la section IV.5.3, à un instant k donné, les deux inégalités précédentes (V-5) et (V-6) génèrent une paire d'hyperplans parallèles dans l'espace des paramètres incertains θ_i , délimitant des demi-espaces fermés. En réalisant leur intersection avec un domaine d'étude \mathcal{D}_0 borné, des demi-espaces bornés $\underline{\mathcal{Z}}(k)$ et $\bar{\mathcal{Z}}(k)$ de \mathbb{R}^r sont obtenus :

$$\underline{\mathcal{Z}}(k) = \{ \theta \in \mathcal{D}_0 / \underline{b}(k) \leq \varphi^T(k)\theta \},$$

$$\bar{\mathcal{Z}}(k) = \{ \theta \in \mathcal{D}_0 / \varphi^T(k)\theta \leq \bar{b}(k) \}.$$

Dans le cadre d'une séquence complète d'observations ($k \in \{1, \dots, t\}$), les demi-espaces précédents s'intersectent de manière générale. Un demi-espace étant un domaine convexe et l'intersection de domaines convexes conservant cette propriété, les deux domaines suivants définis par l'intersection des $\underline{\mathcal{Z}}(k)$, respectivement $\bar{\mathcal{Z}}(k)$, sont donc deux polytopes convexes parfaitement définis par la connaissance de leurs sommets :

$$\underline{\mathcal{D}}_t = \bigcap_{k=1}^t \underline{\mathcal{Z}}(k), \quad \bar{\mathcal{D}}_t = \bigcap_{k=1}^t \bar{\mathcal{Z}}(k).$$

Par définition, $\underline{\mathcal{D}}_t$ (respectivement $\bar{\mathcal{D}}_t$) comprend l'ensemble des valeurs du vecteur de paramètres, élément de \mathcal{D}_0 , et satisfaisant les contraintes inégalités (V-5) (respectivement (V-6)). La construction de ces deux polytopes peut se faire selon la technique proposée par (Walter et al, 1987) ou (Mo et al, 1988, 1990) exposée brièvement dans l'annexe I. Au lieu de construire ces domaines en réalisant simultanément les intersections précédentes, une procédure récursive est utilisée où le domaine $\underline{\mathcal{D}}_k$ (respectivement $\bar{\mathcal{D}}_k$) sera obtenu par intersection de $\underline{\mathcal{D}}_{k-1}$ avec $\underline{\mathcal{Z}}(k)$ (respectivement $\bar{\mathcal{D}}_{k-1}$ avec $\bar{\mathcal{Z}}(k)$). À la fin, l'algorithme

renvoie les sommets de \mathcal{D}_t et $\bar{\mathcal{D}}_t$. Pour initialiser la procédure, il est nécessaire de se munir d'un domaine initial pris identique pour les deux polytopes et égal au domaine d'étude \mathcal{D}_0 .

Pour simplifier la mise en œuvre, il est préférable de prendre un domaine élémentaire pour \mathcal{D}_0 tel un orthotope aligné. Celui-ci est logiquement centré sur un estimateur central du vecteur de paramètres, noté θ_c , obtenu en utilisant un algorithme d'identification classique assurant la minimisation des erreurs d'équation au sens d'une α -norme élevée à la puissance β (norme un, deux pour les Moindres Carrés ou infinie par exemple).

$$\theta_c = \arg \left(\min_{\theta} \left(\sum_{k=1}^t \left\| y(k) - \varphi^T(k) \theta \right\|_{\alpha}^{\beta} \right) \right).$$

Naturellement, l'idéal est de rechercher un modèle le plus proche possible du système physique, ce qui correspond à une erreur d'équation nulle. Cependant, le vecteur $\theta(k)$ dépendant du temps d'une part, le nombre de mesures étant limité et les hypothèses statistiques (distribution gaussienne ou espérance nulle des erreurs dans le cas des Moindres Carrés) pas toujours satisfaites d'autre part, l'identification précédente génère nécessairement des erreurs d'équation non-nulles entre le système et son modèle. Au lieu de représenter ces incohérences à l'aide du seul terme additif borné $e(k)$, elles sont aussi attribuées aux paramètres identifiés précédemment qui deviennent par conséquent incertains. Les erreurs d'équation sont donc modélisées non seulement par le terme additif $e(k)$, mais aussi de manière multiplicative par des paramètres incertains variants. Le vecteur $\theta(k)$ à caractériser peut donc s'écrire :

$$\theta(k) = \theta_c + T(\lambda)v(k), \quad (\text{V-7})$$

où v est un vecteur de q variables bornées normalisées et T une matrice dépendant du vecteur de pondération $\lambda \in \mathbb{R}^{q^*}$ à déterminer. Cette matrice traduit la forme du domaine $\mathcal{D}_t = \mathcal{A}(\theta(k))$ engendré par les incertitudes paramétriques. Le vecteur $\theta(k)$ étant linéaire en les variables bornées, le domaine \mathcal{D}_t est donc un paralléloptope centré en θ_c .

Dans le cadre de l'approche géométrique, ce modèle sera limité au cas où le vecteur de pondération λ se réduit à un unique paramètre scalaire $\lambda \in \mathbb{R}^*$:

$$T(\lambda) = \lambda T_0, \quad (\text{V-8})$$

avec T_0 une matrice choisie par l'utilisateur et supposée de plein rang ligne. Dans ce cas, la forme du domaine abstrait recherché $\mathcal{D}_t = \mathcal{A}(\theta(k))$ est prédéterminée dans la mesure où le seul degré de liberté possible λ porte sur la taille² de \mathcal{D}_t . Cette modélisation permet par exemple de traiter le cas où des dépendances existent entre les composantes du vecteur $\theta(k)$.

Faute de connaissance, il est toujours possible de prendre autant de variables normalisées $v_i(k)$ que de paramètres incertains $\theta_i(k)$ ($q=r$) et s'arranger pour que ces derniers

² Une méthode adaptée au cas où λ est un vecteur sera proposée à la section V.3.3. La forme de \mathcal{D}_t est alors indéterminée (à priori bien sûr) puisque chaque paramètre λ_i va modifier l'espace abstrait recherché dans une direction particulière. La méthode proposée a alors pour avantage d'être réalisée directement dans l'espace des coefficients λ_i alors que l'approche géométrique se déroule dans l'espace des paramètres incertains θ_i , ce qui la rend plus complexe à mettre en œuvre.

soient indépendants en choisissant une matrice T_0 diagonale ; auquel cas le domaine \mathcal{D}_t se limite à un simple orthotope aligné. Supposer que les incertitudes ont la même importance se traduit dans ces conditions par une matrice T_0 se résumant à la matrice identité. Sinon, si de la connaissance est disponible a priori, autant l'exploiter lors du choix de la matrice T_0 tel que proposé dans (Ploix, 1998).

Le vecteur θ_c permet de définir le centre de \mathcal{D}_t . Il est alors clair que le domaine obtenu va dépendre de l'estimateur choisi, néanmoins, cela n'a qu'une importance relative puisque la valeur de λ s'adapte de manière à ce le modèle explique toutes les mesures. Quelle que soit la valeur de θ_c , le modèle obtenu est donc juste, seule sa précision peut en être affectée³.

Si une connaissance a priori sur les incertitudes est détenue, la matrice T_0 peut être utilisée pour régler la taille de \mathcal{D}_0 . Il suffit en effet de déterminer l'orthotope aligné circonscrit de $T_0 v(k)$, le domaine d'étude est alors obtenu en multipliant la taille de $\square_{\mathcal{A}}(T_0 v(k))$ par un nombre suffisamment grand pour que \mathcal{D}_0 contienne \mathcal{D}_t (un ordre de grandeur des incertitudes est alors appréciable). Sinon, il est toujours possible d'imposer un domaine de très grande taille contenant assurément \mathcal{D}_t . Notons que par construction, les polytopes $\underline{\mathcal{D}}_t$ et $\overline{\mathcal{D}}_t$ sont nécessairement contenus dans \mathcal{D}_0 .

V.2.3 Seconde étape

V.2.3.1 Principe

Une fois que les deux polytopes $\underline{\mathcal{D}}_t$ et $\overline{\mathcal{D}}_t$ ont été déterminés, il convient de caractériser les bornes des paramètres incertains. L'objectif est alors de trouver un domaine relativement simple (orthotope ou paralléloétope convexe de manière plus générale) dans l'espace des paramètres incertains, centré sur la valeur centrale θ_c déterminée précédemment et qui contienne à la fois au moins un élément de $\underline{\mathcal{D}}_t$ et un autre de $\overline{\mathcal{D}}_t$.

En effet, le polytope $\overline{\mathcal{D}}_t$ précédent désigne l'ensemble des valeurs θ'_0 du vecteur invariant θ vérifiant chacune des t inégalités (V-6). N'importe quelle valeur θ'_0 de ce domaine peut donc générer, pour tous les instants $k \in \{1, \dots, t\}$, un minorant $\underline{y}(k) = \varphi^T(k) \theta'_0 - \bar{e}$ de la mesure $y(k)$ et peut par conséquent prétendre à être la borne $\underline{\theta}$ recherchée. De même, toute valeur θ_0 du vecteur θ appartenant au polytope $\underline{\mathcal{D}}_t$ vérifie à tout instant les contraintes (V-5). En engendrant un majorant $\bar{y}(k) = \varphi^T(k) \theta_0 + \bar{e}$ de cette même mesure, le vecteur θ_0 peut donc être assimilé à la borne $\bar{\theta}$ recherchée.

En conclusion, il existe, à tout instant $k \in \{1, \dots, t\}$, une valeur $\theta_0(k)$ du vecteur $\theta(k)$ des paramètres incertains variants, comprise entre θ'_0 et θ_0 , telle que $\theta_0(k)$ explique la mesure courante $y(k)$ à l'erreur d'équation près $e(k)$ bornée par $\pm \bar{e}$:

$$y(k) - \varphi^T(k) \theta_0(k) \in [-\bar{e}, \bar{e}], \quad \theta_0(k) \in [\theta'_0, \theta_0].$$

³ Un critère de précision sera ultérieurement défini pour permettre d'optimiser ce centre en tenant compte de la dispersion des incertitudes.

L'objectif étant d'obtenir un modèle le plus précis possible, il reste à déterminer le paramètre λ associé au plus petit polytope \mathcal{D}_t centré en θ_c , de forme imposée par T_0 et contenant au moins un élément de $\underline{\mathcal{D}}_t$ et un autre de $\overline{\mathcal{D}}_t$. Un seul point commun entre \mathcal{D}_t et chacun des deux domaines précédents est suffisant pour expliquer toutes les mesures.

À ce stade, les résultats de la méthode proposée sont conditionnés par la valeur choisie de la borne \bar{e} de l'erreur d'équation. Soit θ_{inf} , l'estimateur au sens de la norme infinie du vecteur de paramètres et e_{inf} l'erreur associée :

$$e_{inf} = \left\| y(k) - \varphi^T(k) \theta_{inf} \right\|_{\infty} = \min_{\theta} \left(\max_k \left(\left| y(k) - \varphi^T(k) \theta \right| \right) \right).$$

Le vecteur θ_{inf} représente l'estimateur qui donne le minimax de l'erreur d'équation $e(k)$, écart entre les mesures et le modèle $\varphi^T(k) \theta$. Le modèle étant linéaire par rapport aux paramètres incertains, cette solution est unique. En posant $\bar{e} = e_{inf}$, l'évaluation des domaines $\underline{\mathcal{D}}_t$ et $\overline{\mathcal{D}}_t$ conduit à rechercher toutes les valeurs de θ permettant d'assurer les $2t$ contraintes suivantes :

$$-e_{inf} \leq y(k) - \varphi^T(k) \theta \text{ et } y(k) - \varphi^T(k) \theta \leq e_{inf}, k \in \{1, \dots, t\}.$$

D'après ce qui précède, toute autre valeur que θ_{inf} , génère une erreur maximale supérieure à e_{inf} ; par conséquent, θ_{inf} est la seule valeur possible de θ permettant, pour $\bar{e} = e_{inf}$, de vérifier les inégalités précédentes. En conclusion, les domaines $\underline{\mathcal{D}}_t$ et $\overline{\mathcal{D}}_t$, contenant l'ensemble des valeurs de θ vérifiant respectivement les deux inégalités précédentes, ne vont s'intersecter qu'en un seul point correspondant à $\theta = \theta_{inf}$ (figure V-1).

Si une valeur $\bar{e} > e_{inf}$ est choisie, des contraintes moins restrictives sont imposées sur l'erreur d'équation, ce qui engendre des polytopes $\underline{\mathcal{D}}_t$ et $\overline{\mathcal{D}}_t$ plus importants. Il existe par conséquent au moins une valeur de θ (en l'occurrence la solution précédente θ_{inf}) vérifiant à la fois les $2t$ contraintes (V-5) et (V-6). Dans ces conditions, ces deux domaines vont s'intersecter et générer ainsi le polytope représentant l'ensemble \mathcal{D}_{fps} des valeurs acceptables des paramètres invariants θ_{fps} recherché par Walter et al (1987) et Mo et al (1988, 1990).

Si une valeur $\bar{e} < e_{inf}$ est prise, le problème précédent n'a alors plus de solution puisque $\underline{\mathcal{D}}_t$ et $\overline{\mathcal{D}}_t$ n'ont plus aucun point commun. Il n'existe pas, par conséquent, de valeur du vecteur des paramètres invariants θ qui permette d'expliquer, à elle seule, toutes les mesures $y(k)$, $k \in \{1, \dots, t\}$.

En revanche, et là se situe le point original de la méthode proposée, il est possible de rechercher un domaine \mathcal{D}_t de valeurs du vecteur des paramètres variants $\theta(k)$ tel qu'à chaque instant k , il existe un élément de ce domaine qui explique la $k^{\text{ème}}$ mesure. Dans ces conditions, le vecteur des paramètres incertains fluctue au cours du temps et les valeurs extrêmes des $\theta_i(k)$ convenant vont permettre de définir les bornes de chacune des variables bornées intervenant dans le modèle.

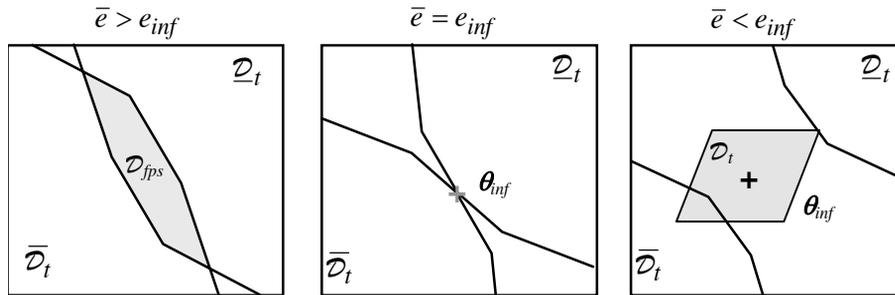


Figure V-1. Recherche du domaine \mathcal{D}_{fps} et caractérisation des bornes des incertitudes

Pour tout autre estimateur θ_c , l'erreur définie par :

$$e_{inf}(\theta_c) = \sup_k \left(|y(k) - \varphi^T(k)\theta_c| \right),$$

est supérieure par définition à e_{inf} . De ce fait, si $\bar{e} > e_{inf}(\theta_c)$, alors toute l'erreur entre le modèle $\varphi^T(k)\theta_c$ et la mesure $y(k)$ est contenue dans l'erreur d'équation $e(k)$, ce qui engendre une valeur nulle de λ . Au niveau de la seconde étape, ceci correspond à la présence de θ_c à l'intérieur des deux domaines $\underline{\mathcal{D}}_t$ et $\bar{\mathcal{D}}_t$. Par conséquent, le plus petit polytope \mathcal{D}_t centré sur cet estimateur et contenant au moins un point de $\underline{\mathcal{D}}_t$ et un autre de $\bar{\mathcal{D}}_t$ se réduit à un point unique θ_c ; ce qui conduit à $\lambda=0$, les paramètres étant alors parfaitement certains. Si $\bar{e} = e_{inf}(\theta_c)$, alors dans ces conditions \bar{e} vaut la plus grande erreur en valeur absolue entre la mesure et son estimée. Dans ces conditions, λ vaut 0 et les domaines $\underline{\mathcal{D}}_t$ et $\bar{\mathcal{D}}_t$ se coupent en plusieurs points dont l'un correspond à θ_c . Enfin, pour $\bar{e} < e_{inf}(\theta_c)$, le paramètre λ est non-nul et les paramètres $\theta(k)$ sont alors incertains car la taille de \mathcal{D}_t est nécessairement non-nulle, θ_c n'appartenant plus à $\underline{\mathcal{D}}_t$ et $\bar{\mathcal{D}}_t$.

En remarquant que la borne \bar{e} est nécessairement positive ou nulle par définition, pour atteindre nos objectifs de caractérisation, la double contrainte inégalité suivante doit impérativement être satisfaite :

$$0 \leq \bar{e} < e_{inf}(\theta_c).$$

Donnons un sens physique à l'erreur d'équation $e(k)$. L'erreur globale entre les mesures et le modèle moyen $\varphi^T(k)\theta_c$ est bornée par $\pm e_{inf}(\theta_c)$ et se décompose en une erreur de type additif sous la forme de l'erreur d'équation $e(k)$ et en une seconde liée aux incertitudes $\theta(k)$, dites multiplicatives, affectant les paramètres du modèle. Suivant la valeur choisie pour le degré de liberté \bar{e} , plus ou moins d'erreur sera rejetée sur les paramètres, ce qui modifie l'importance des incertitudes paramétriques au sein du modèle. L'influence de la partie additive sur l'enveloppe est simple à appréhender. Supposons que l'erreur globale soit entièrement contenue dans la partie additive $e(k)$ en choisissant la borne \bar{e} suffisamment grande : $\bar{e} > e_{inf}(\theta_c)$, alors l'enveloppe $\mathcal{A}(y(k))$ est centrée sur la courbe $\varphi^T(k)\theta_c$ et possède un rayon constant égal à \bar{e} . Les limites de l'enveloppe ont donc exactement la même allure que $\varphi^T(k)\theta_c$, l'enveloppe est alors qualifiée de régulière. La figure V-2 illustre ce cas en présentant au cours du temps le modèle moyen $\varphi^T(k)\theta_c$ ainsi que l'enveloppe $\mathcal{A}(y(k))$ contenant toutes les mesures $y(k)$ pour un exemple particulier. L'erreur $e_{inf}(\theta_c)$ représentant le plus grand écart

entre une mesure et son estimateur $\varphi^T(k)\theta_c$; puisque \bar{e} est plus grande que $e_{inf}(\theta_c)$, alors nécessairement les mesures $y(k)$ appartiennent à l'enveloppe $[\varphi^T(k)\theta_c - \bar{e}, \varphi^T(k)\theta_c + \bar{e}]$. En revanche, dans le cas contraire, si \bar{e} est plus faible que $e_{inf}(\theta_c)$, l'algorithme rejette le reste de l'erreur globale sur les paramètres. Les incertitudes apparaissant de manière multiplicative dans le modèle, l'allure de l'enveloppe peut être fortement différente de la courbe $\varphi^T(k)\theta_c$ comme dessiné sur la figure V-3. L'enveloppe obtenue est alors non-régulière, mais quels que soient l'estimateur et \bar{e} choisis, toutes les mesures resteront nécessairement à l'intérieur puisque λ a été réglé à cette fin.

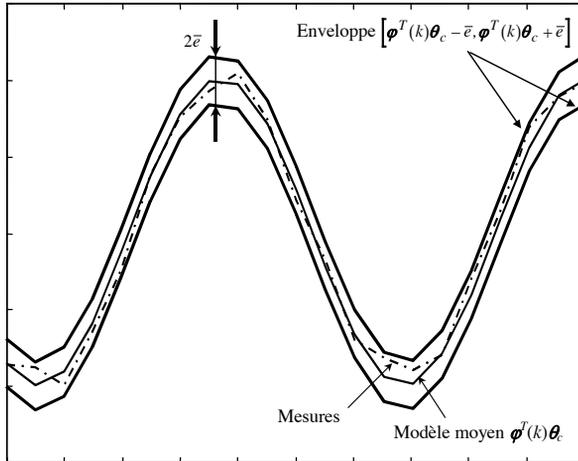


Figure V-2. Cas $\bar{e} > e_{inf}(\theta_c)$

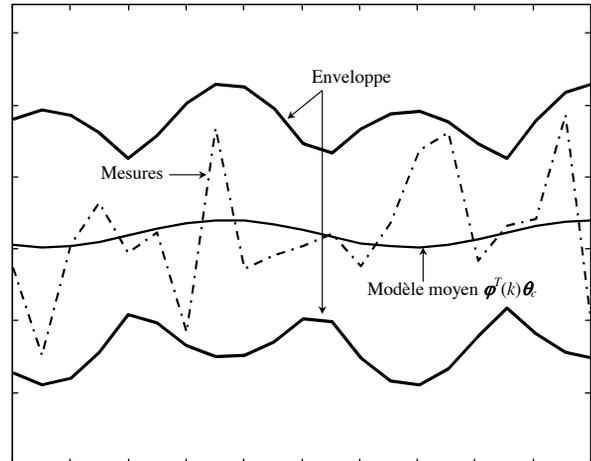


Figure V-3. Cas $\bar{e} < e_{inf}(\theta_c)$

Par construction, de manière générale, le polytope \mathcal{D}_t touche l'un des deux polytopes $\underline{\mathcal{D}}_t$ ou $\bar{\mathcal{D}}_t$ en un seul point situé sur l'une de ses facettes ; supposons que ce soit $\bar{\mathcal{D}}_t$. Par conséquent, l'enveloppe précédente a nécessairement l'une de ses deux frontières passant par l'une des mesures $y(k)$. En effet, ce point correspond alors à la borne $\underline{\theta}$ qui est unique. Il appartient à l'une des facettes de $\bar{\mathcal{D}}_t$, dont les équations cartésiennes sont définies par les hyperplans $\bar{\mathcal{F}}(k) = \{\theta' | \varphi^T(k)\theta' \leq \bar{b}(k)\}$, $k \in \{1, \dots, t\}$. Dans ces conditions, l'égalité suivante $y(k) = \varphi^T(k)\underline{\theta} - \bar{e}$ est satisfaite et l'une des deux frontières de l'enveloppe passe nécessairement par la $k^{\text{ième}}$ mesure.

Dans le tableau suivant sont récapitulés les différentes hypothèses utilisées et les objectifs alors recherchés :

	Caractérisation des bornes	Identification ensembliste
Condition sur \bar{e}	$\bar{e} < e_{inf}(\theta_c)$	$\bar{e} > e_{inf}(\theta_c)$
Nature des paramètres	Variante	Invariante
Domaine calculé	Domaine \mathcal{D}_t permettant de caractériser les bornes des incertitudes	Ensemble des valeurs de paramètres acceptables \mathcal{D}_{fps}
Objectif	À chaque instant, une valeur $\theta(k)$ de \mathcal{D}_t explique la mesure $y(k)$ à \bar{e} imposée.	Chaque valeur de \mathcal{D}_{fps} explique l'ensemble des mesures à \bar{e} imposée.

Tableau V-1. Récapitulatif

V.2.3.2 Mise en œuvre

La solution proposée pour déterminer le parallélogramme \mathcal{D}_t est un algorithme dichotomique. Le principe est d'adapter la valeur de λ par encadrement jusqu'à ce que ce domaine touche l'un des deux polytopes $\underline{\mathcal{D}}_t$ ou $\overline{\mathcal{D}}_t$ en un point et intersecte le second.

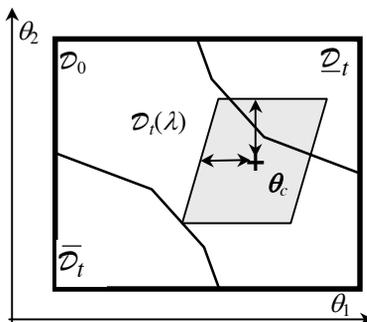


Figure V-4. Parallélogramme $\mathcal{D}_t(\lambda)$ recherché ($r=2$)

Se pose alors le problème de savoir comment vérifier que deux polytopes s'intersectent. Ces derniers étant parfaitement définis par leurs sommets, puisque convexes, il suffit de regarder si au moins un sommet de l'un est à l'intérieur du second.

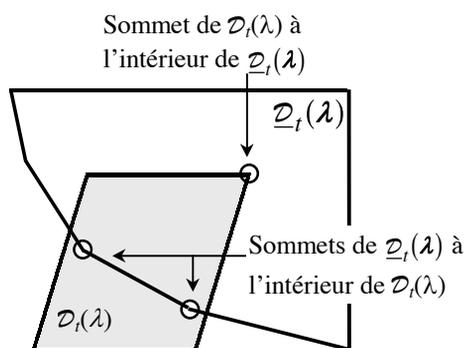


Figure V-5. Intersection de deux polytopes convexes

Dans ces conditions, ce test s'opère simplement en testant si un sommet de coordonnées données vérifie toutes les contraintes inégalité associées au second polytope. Dans le cas des deux domaines $\underline{\mathcal{D}}_t$ ou $\overline{\mathcal{D}}_t$, ces contraintes sont directement données par (V-5) et (V-6) alors que leurs sommets respectifs ont été déterminés lors de la première étape grâce à la méthode de Walter et al (1987) et Mo et al (1988, 1990) présentée en annexe I. Pour ce qui est du domaine $\mathcal{D}_t(\lambda)$, les contraintes inégalité sont calculées suivant la méthode proposée dans la section IV.5.3 puisque la matrice T_0 est supposée de plein rang ligne (notons qu'il est possible d'aboutir si T_0 est de plein rang colonne grâce aux résultats de l'annexe H). Puis, à partir de ces mêmes contraintes, ses sommets sont calculés par le même algorithme que pour $\underline{\mathcal{D}}_t$ ou $\overline{\mathcal{D}}_t$.

En supposant à une itération donnée que les deux bornes λ_{max} et λ_{min} encadrant la valeur recherchée de λ ont été déterminées, le principe de la dichotomie consiste à calculer une nouvelle valeur de λ définie par :

$$\lambda = \frac{\lambda_{max} + \lambda_{min}}{2}.$$

Les nouvelles contraintes inégalité et les sommets associés au parallélotope $\mathcal{D}_t(\lambda)$ sont alors déterminés. Il faut ensuite tester les quatre possibilités suivantes :

- ① au moins un sommet de $\mathcal{D}_t(\lambda)$ est à l'intérieur de $\bar{\mathcal{D}}_t$,
- ② au moins un sommet de $\bar{\mathcal{D}}_t$ est à l'intérieur de $\mathcal{D}_t(\lambda)$,
- ③ au moins un sommet de $\mathcal{D}_t(\lambda)$ est à l'intérieur de $\underline{\mathcal{D}}_t$,
- ④ au moins un sommet de $\underline{\mathcal{D}}_t$ est à l'intérieur de $\mathcal{D}_t(\lambda)$.

Pour que λ soit une valeur acceptable, c'est-à-dire que $\mathcal{D}_t(\lambda)$ intersecte à la fois $\bar{\mathcal{D}}_t$ et $\underline{\mathcal{D}}_t$, il faut que la condition suivante soit vérifiée :

$$(\text{① ou ②}) \text{ et } (\text{③ ou ④}). \quad (\text{V-9})$$

Dans ce cas, $\lambda_{max}=\lambda$ est imposé, sinon $\lambda_{min}=\lambda$ est posé, puis cette opération est réitérée jusqu'à ce que la différence entre λ_{min} et λ_{max} soit inférieure à une tolérance spécifiée par l'utilisateur. La valeur renvoyée par la procédure est alors $\lambda=\lambda_{max}$.

Lors de l'initialisation de cette procédure dichotomique, il faut faire en sorte que le domaine recherché $\mathcal{D}_t(\lambda)$ reste toujours complètement inscrit à l'intérieur du domaine d'étude \mathcal{D}_0 en se fixant la valeur maximale λ_{max} . L'orthotope \mathcal{D}_0 est défini de manière générale par :

$$\boldsymbol{\theta} \in \mathcal{D}_0 \Leftrightarrow \left\{ \boldsymbol{\theta} / \forall i \in \{1, \dots, r\}, \underline{d}_{0,i} \leq \theta_i \leq \bar{d}_{0,i} \right\} \text{ où } \underline{d}_{0,i} = \theta_{c,i} - w_i/2 \text{ et } \bar{d}_{0,i} = \theta_{c,i} + w_i/2,$$

$\theta_{c,i}$ définissant la $i^{\text{ième}}$ composante de $\boldsymbol{\theta}_c$ et w_i la longueur de la $i^{\text{ième}}$ composante du vecteur borné $\boldsymbol{\theta}$ défini précédemment. Le calcul de l'orthotope aligné circonscrit associé à (V-7) conduit à :

$$\boldsymbol{\theta} \in \square\mathcal{A}(\boldsymbol{\theta}) \Leftrightarrow \left\{ \boldsymbol{\theta} / \forall i \in \{1, \dots, r\}, \theta_{c,i} - \lambda \|\mathbf{t}_i^T\|_1 \leq \theta_i \leq \theta_{c,i} + \lambda \|\mathbf{t}_i^T\|_1 \right\}.$$

Pour que \mathcal{D}_0 contienne $\square\mathcal{A}(\boldsymbol{\theta})$, il faut que $\underline{d}_{0,i} \leq \theta_{c,i} - \lambda \|\mathbf{t}_i^T\|_1$ et $\theta_{c,i} + \lambda \|\mathbf{t}_i^T\|_1 \leq \bar{d}_{0,i}$, entraînant :

$$\lambda \leq \min \left(\frac{\theta_{c,i} - \underline{d}_{0,i}}{\|\mathbf{t}_i^T\|_1}, \frac{\bar{d}_{0,i} - \theta_{c,i}}{\|\mathbf{t}_i^T\|_1} \right), \quad \forall i \in \{1, \dots, r\}.$$

La plus grande valeur de λ acceptable correspond donc à :

$$\lambda_{max} = \inf_{i \in \{1, \dots, r\}} \left\{ \frac{w_i}{\|\mathbf{t}_i^T\|_1} \right\}.$$

De même, la valeur λ_{min} choisie arbitrairement initialement doit vérifier :

- $0 \leq \lambda_{min} < \lambda_{max}$,
- $\mathcal{D}_t(\lambda_{min})$ ne doit pas intersecter les deux polytopes $\underline{\mathcal{D}}_t$ et $\bar{\mathcal{D}}_t$ à la fois.

Pour que λ_{min} soit une valeur acceptable, il faut que la condition (V-9) pour $\lambda=\lambda_{min}$ ne soit pas vérifiée, sinon, il faut se redonner une nouvelle valeur plus petite et réaliser ce test à nouveau.

Remarque : un test supplémentaire peut être implémenté consistant à vérifier que le domaine $\mathcal{D}_t(\lambda)$ obtenu intersecte bien $\underline{\mathcal{D}}_t$ et $\bar{\mathcal{D}}_t$. En effet, suivant la taille du domaine d'étude

\mathcal{D}_0 choisi, la valeur de λ_{max} peut être trop faible pour que l'algorithme puisse trouver un domaine $\mathcal{D}_t(\lambda)$ de forme imposée par T_0 , centré en θ_c et contenant des points de $\underline{\mathcal{D}}_t$ et $\overline{\mathcal{D}}_t$. La procédure dichotomique stoppera lorsque $\lambda = \lambda_{max}$, mais l'objectif ne sera pas atteint. Ce test est semblable à celui réalisé sur $\mathcal{D}_t(\lambda_{min})$ en remplaçant λ_{min} par λ_{max} . Si la condition (V-9) n'est pas vérifiée, il faut alors prendre un domaine d'étude plus grand.

Ci-dessous est décrit l'organigramme correspondant à la seconde phase.

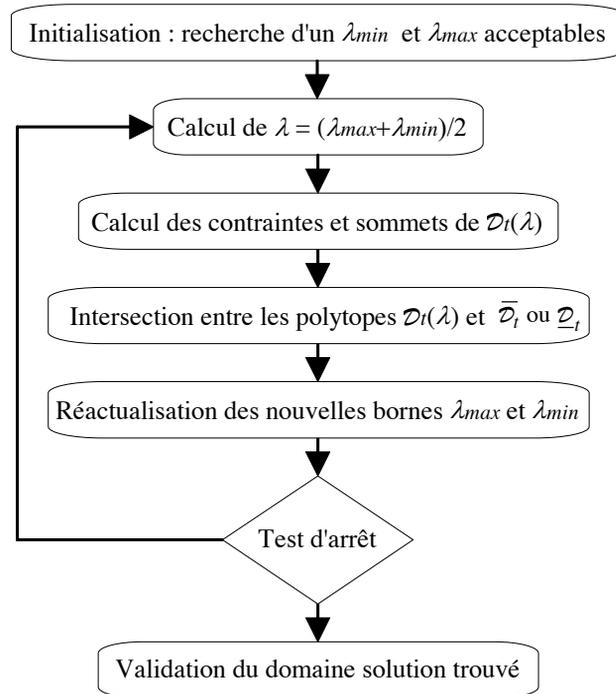


Figure V-6. Organigramme de la seconde étape

V.3 CARACTÉRISATION

V.3.1 Critère de précision

Avant de présenter l'utilité d'un tel critère, complétons la partie modélisation entreprise lors de la présentation de l'approche géométrique en explicitant ce qui est sous-entendu par forme indéterminée (section V.2.2). Il s'agit en fait de traiter le cas où $\lambda \in \mathbb{R}^{q^*}$ est un vecteur de pondération à déterminer. Dans ce cas, la matrice $T(\lambda)$ définie en (V-7) s'écrit comme suit :

$$T(\lambda) = [\lambda_1 t_1 \quad \cdots \quad \lambda_q t_q], \quad \lambda_i \in \mathbb{R}^{*+}, \quad t_i \in \mathbb{R}^r. \quad (\text{V-10})$$

Dans ces conditions, la forme du domaine \mathcal{D}_t recherché n'est pas déterminée ; la modification de l'un des poids λ_i entraînant logiquement une variation de la dite forme. Chaque poids λ_i est

ainsi associé à une direction t_i privilégiée et à une unique variable bornée normalisée et indépendante $v_i(k)$.

Exemple. Considérons le système suivant :

$$\theta = \begin{bmatrix} \lambda_1 & 0 & \lambda_3 & \lambda_4 \\ 0 & \lambda_2 & \lambda_3 & -\lambda_4 \end{bmatrix} \nu.$$

Sur la figure suivante sont représentés les espaces abstraits $\mathcal{D}_{t,i}$, $i \in \{a,b\}$, du système précédent lorsque le vecteur de pondération vaut successivement :

$$\lambda_a = \begin{bmatrix} \frac{1}{3} & \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \end{bmatrix}^T \text{ et } \lambda_b = \begin{bmatrix} \frac{1}{10} & \frac{1}{10} & \frac{8}{10} & \frac{1}{10} \end{bmatrix}^T.$$

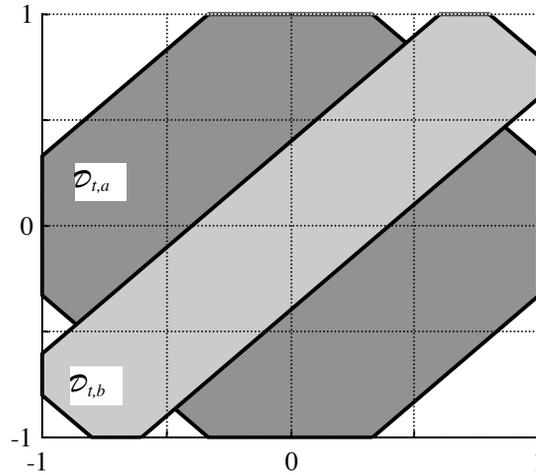


Figure V-7. Formes indéterminées

La forme dépend bien de l'importance relative de chaque poids λ_i . ■

Compte tenu des définitions (V-1) et (V-7), la procédure de caractérisation se résume donc à la détermination d'un $(r+q+1)$ -uplet de coefficients $(\theta_c, \lambda, \bar{e})$ tel que le modèle explique, le plus précisément possible, l'ensemble des observations $y(k)$:

$$y(k) \in \mathcal{A} \left(\varphi^T(k) \theta_c + \varphi^T(k) \mathbf{T}(\lambda) \nu(k) + e(k) \right), k \in \{1, \dots, t\}. \quad (\text{V-11})$$

Néanmoins, tous les modèles satisfaisant la relation précédente ne sont pas intéressants, nous ne porterons notre attention que sur le modèle le plus précis. Il est dans ces conditions nécessaire de définir un critère représentatif de cette précision (Ploix et al, 1999b). Pour cela, évaluons l'espace abstrait précédent en utilisant l'analyse par intervalles :

$$y(k) \in \left[\varphi^T(k) \theta_c - \left\| \varphi^T(k) \mathbf{T}(\lambda) \right\|_1 - \bar{e}, \varphi^T(k) \theta_c + \left\| \varphi^T(k) \mathbf{T}(\lambda) \right\|_1 + \bar{e} \right].$$

A chaque instant k , la longueur de l'intervalle ainsi obtenu est donc définie par :

$$w(k) = 2\bar{e} + 2 \left\| \varphi^T(k) \mathbf{T}(\lambda) \right\|_1.$$

Le critère de précision, noté j_p , est obtenu en cumulant ces longueurs sur l'horizon temporel de longueur t . Il témoigne de la précision globale de la caractérisation puisqu'il cherche à ce que l'enveloppe encadre le plus précisément possible l'ensemble des mesures :

$$j_p = 2t\bar{e} + 2 \sum_{k=1}^t \left\| \boldsymbol{\varphi}^T(k) \mathbf{T}(\boldsymbol{\lambda}) \right\|_1.$$

Ce critère dépend de manière explicite de la borne \bar{e} de l'erreur d'équation et de manière implicite de la valeur centrale $\boldsymbol{\theta}_c$ par l'intermédiaire du vecteur $\boldsymbol{\lambda}$ calculé. En effet, puisque le parallélogramme $\mathcal{D}_t(\boldsymbol{\lambda})$ est centré en $\boldsymbol{\theta}_c$ (figure V-4), modifier la valeur de ce vecteur influe sur celle de $\boldsymbol{\lambda}$ puisqu'il faut respecter les contraintes d'intersection entre $\mathcal{D}_t(\boldsymbol{\lambda})$ et les deux domaines $\underline{\mathcal{D}}_t$ et $\overline{\mathcal{D}}_t$. Il sera donc possible, comme nous le verrons ultérieurement, d'optimiser à la fois le vecteur de pondération $\boldsymbol{\lambda}$ ainsi que l'estimateur $\boldsymbol{\theta}_c$ et la borne \bar{e} en minimisant le critère j_p .

Le fait de s'être limité à deux structures (déterminée (V-8) et indéterminée (V-10)) linéaires en $\boldsymbol{\lambda}$ est d'une part logique puisque le modèle est linéaire en les variables bornées, d'autre part, si la matrice $\mathbf{T}(\boldsymbol{\lambda})$ était quelconque, il serait difficile de prévoir l'influence d'une variation de $\boldsymbol{\lambda}$ sur la précision.

Dans le cas de la forme prédéterminée associée à (V-8), le critère de précision s'écrit :

$$j_p = 2t\bar{e} + 2\lambda \sum_{k=1}^t \left\| \boldsymbol{\varphi}^T(k) \mathbf{T}_0 \right\|_1. \quad (\text{V-12})$$

Dans le cas de la forme indéterminée définie par (V-10), la contrainte (V-11) s'écrit :

$$y(k) \in \mathcal{A} \left(\boldsymbol{\varphi}^T(k) \boldsymbol{\theta}_c + \left[\boldsymbol{\varphi}^T(k) \mathbf{t}_1 \quad \dots \quad \boldsymbol{\varphi}^T(k) \mathbf{t}_q \right] \begin{bmatrix} \lambda_1 \mathbf{v}_1(k) \\ \vdots \\ \lambda_q \mathbf{v}_q(k) \end{bmatrix} + e(k) \right), \quad (\text{V-13})$$

le critère de précision se résume donc à :

$$j_p = 2t\bar{e} + \boldsymbol{\chi}^T \boldsymbol{\lambda} \quad (\text{V-14})$$

$$\text{avec } \boldsymbol{\chi}^T(k) = \left[\left\| \boldsymbol{\varphi}^T(k) \mathbf{t}_1 \right\| \quad \dots \quad \left\| \boldsymbol{\varphi}^T(k) \mathbf{t}_q \right\| \right] \text{ et } \boldsymbol{\chi}^T = \sum_{k=1}^t \boldsymbol{\chi}^T(k).$$

Nous développerons plus amplement l'utilité d'un tel critère dans la section V.3.4.

V.3.2 Forme prédéterminée

L'approche géométrique a permis de mettre en évidence le principe de la méthode de caractérisation, néanmoins, sa mise en œuvre reste lourde. La méthode proposée dans cette section permet de réaliser exactement le même travail tout en restant nettement plus économe en temps de calcul et plus facile à mettre en œuvre. Notons qu'il n'est plus supposé que la matrice \mathbf{T}_0 soit de plein rang ligne.

L'expression (V-11) peut également s'écrire :

$$y(k) - \boldsymbol{\varphi}^T(k)\boldsymbol{\theta}_c \in \mathcal{A}\left(\lambda \boldsymbol{\varphi}^T(k)\mathbf{T}_0 \boldsymbol{\nu}(k) + e(k)\right), k \in \{1, \dots, t\}.$$

Les variables bornées $\nu_i(k)$ du vecteur $\boldsymbol{\nu}(k)$ étant normalisées, les deux inégalités suivantes sont déduites :

$$y(k) - \boldsymbol{\varphi}^T(k)\boldsymbol{\theta}_c \leq \lambda \|\boldsymbol{\varphi}^T(k)\mathbf{T}_0\|_1 + \bar{e} \quad \text{et} \quad -\lambda \|\boldsymbol{\varphi}^T(k)\mathbf{T}_0\|_1 - \bar{e} \leq y(k) - \boldsymbol{\varphi}^T(k)\boldsymbol{\theta}_c,$$

ce qui conduit respectivement à :

$$\lambda \geq \frac{y(k) - \boldsymbol{\varphi}^T(k)\boldsymbol{\theta}_c - \bar{e}}{\|\boldsymbol{\varphi}^T(k)\mathbf{T}_0\|_1} \quad \text{et} \quad \lambda \geq \frac{-y(k) + \boldsymbol{\varphi}^T(k)\boldsymbol{\theta}_c + \bar{e}}{\|\boldsymbol{\varphi}^T(k)\mathbf{T}_0\|_1}.$$

En regroupant ces deux inégalités et en tenant compte du fait que le paramètre λ ne peut être négatif, le résultat suivant est déduit :

$$\lambda \geq \max\left(0, \frac{y(k) - \boldsymbol{\varphi}^T(k)\boldsymbol{\theta}_c - \bar{e}}{\|\boldsymbol{\varphi}^T(k)\mathbf{T}_0\|_1}, \frac{-y(k) + \boldsymbol{\varphi}^T(k)\boldsymbol{\theta}_c + \bar{e}}{\|\boldsymbol{\varphi}^T(k)\mathbf{T}_0\|_1}\right), k \in \{1, \dots, t\}.$$

D'après le critère de précision (V-12), pour \bar{e} et $\boldsymbol{\theta}_c$ fixés, la précision sera d'autant meilleure que le paramètre λ sera faible. La valeur du paramètre λ permettant d'obtenir le modèle le plus précis à \bar{e} et $\boldsymbol{\theta}_c$ donnés, est donc définie par (Ploix et al, 1999b), (Adrot et al, 1999c) :

$$\lambda = \sup_{k \in \{1, \dots, t\}} \left\{ \max\left(0, \frac{y(k) - \boldsymbol{\varphi}^T(k)\boldsymbol{\theta}_c - \bar{e}}{\|\boldsymbol{\varphi}^T(k)\mathbf{T}_0\|_1}, \frac{-y(k) + \boldsymbol{\varphi}^T(k)\boldsymbol{\theta}_c + \bar{e}}{\|\boldsymbol{\varphi}^T(k)\mathbf{T}_0\|_1}\right) \right\}. \quad (\text{V-15})$$

V.3.3 Forme indéterminée

En reprenant (V-13), nous déduisons que :

$$\begin{aligned} y(k) - \boldsymbol{\varphi}^T(k)\boldsymbol{\theta}_c &\leq \boldsymbol{\chi}^T(k)\boldsymbol{\lambda} + \bar{e} \\ -\boldsymbol{\chi}^T(k)\boldsymbol{\lambda} - \bar{e} &\leq y(k) - \boldsymbol{\varphi}^T(k)\boldsymbol{\theta}_c \end{aligned}$$

Cette double inégalité se traduit par :

$$\boldsymbol{\chi}^T(k)\boldsymbol{\lambda} \geq \max(y(k) - \boldsymbol{\varphi}^T(k)\boldsymbol{\theta}_c - \bar{e}, -y(k) + \boldsymbol{\varphi}^T(k)\boldsymbol{\theta}_c - \bar{e}), \quad (\text{V-16})$$

où $\boldsymbol{\lambda}$ est un vecteur de termes positifs.

Le problème revient donc à minimiser le critère de précision (V-14) tout en respectant les contraintes précédentes. À chaque instant $k \in \{1, \dots, t\}$, l'inégalité précédente conduit à une contrainte se traduisant par un demi-espace de \mathbb{R}^q (dont la frontière est un hyperplan de dimension $q-1$) dans l'espace des paramètres λ_i . Ces t contraintes vont donc définir un polytope convexe issu de l'intersection de t demi-espaces, domaines convexes par définition.

Ce polytope correspond de fait au domaine autorisé de λ dans \mathbb{R}^q , c'est-à-dire aux valeurs de ce vecteur permettant de respecter les contraintes (V-16) et donc d'expliquer les mesures.

Il est alors clair que pour minimiser j_p , pour θ_c et \bar{e} donnés, il est nécessaire de résoudre un problème de programmation linéaire en minimisant le produit scalaire $\chi^T \lambda$ sous la condition que λ fasse partie du domaine autorisé (Ploix et al, 1999b). Ce dernier étant convexe, un de ses sommets correspond obligatoirement à la valeur du vecteur λ autorisé assurant la plus petite projection orthogonale de celui-ci sur la direction χ .

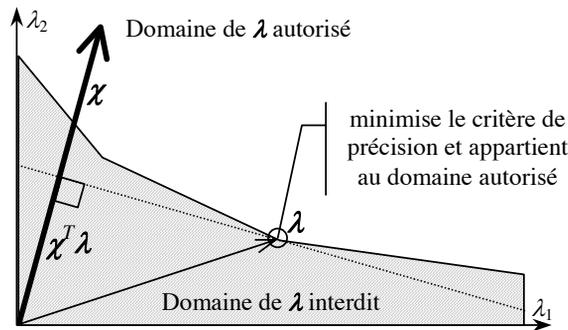


Figure V-8. Détermination du vecteur λ

Les contraintes (V-16) étant parfaitement définies, la détermination du polytope définissant le domaine autorisé s'effectue comme pour la première phase de la méthode géométrique.

V.3.4 Optimisation simultanée du centre et de la borne de l'erreur d'équation

Lors de la définition du critère de précision au niveau de la section V.3.1, nous avons souligné que j_p dépend explicitement de la borne \bar{e} de l'erreur d'équation et de manière implicite de la valeur centrale θ_c . Il est donc possible d'ajouter une étape supplémentaire, utilisant par exemple un algorithme d'optimisation non-linéaire, telle la méthode du simplexe (Nelder et al, 1965), afin d'optimiser le choix de ces deux paramètres en minimisant j_p .

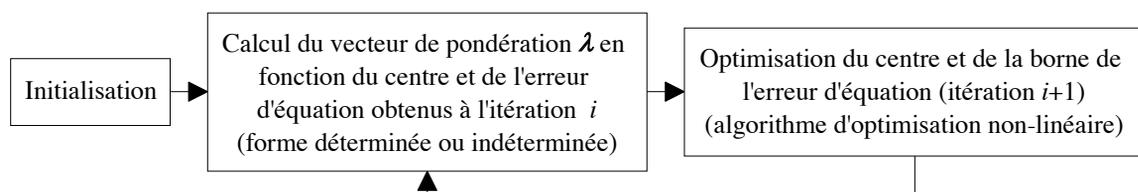


Figure V-9. Optimisation de θ_c et \bar{e}

Pour chaque nouvelle valeur du couple (θ_c, \bar{e}) , la valeur du vecteur de pondération λ est recalculée suivant les méthodes exposées lors des deux sections précédentes. La procédure est stoppée lorsque le critère d'arrêt de l'étape supplémentaire est satisfait. L'estimateur central résultant est alors appelé estimateur le plus précis. Notons que cette optimisation étant décomposée en deux étapes, le résultat obtenu a toutes les chances d'être sous-optimal.

Remarquons que cette optimisation peut aussi être réalisée dans le cadre de l'approche géométrique. Il faut simplement prendre garde au fait qu'à chaque itération, le domaine

d'étude \mathcal{D}_0 est recentré sur la nouvelle valeur de θ_c proposée par l'algorithme d'optimisation non-linéaire. Si les contraintes (V-5) et (V-6) sont complètement indépendantes de la valeur de θ_c , en revanche, les polytopes $\underline{\mathcal{D}}_t$ et $\overline{\mathcal{D}}_t$ sont obtenus par l'intersection des hyperplans associées aux inégalités précédentes et du domaine d'étude \mathcal{D}_0 centré sur θ_c . Si ce centre est éloigné de la valeur réelle et que la taille de \mathcal{D}_0 est faible, il se peut alors que l'un des deux polytopes se résume à l'ensemble vide. Dans ces conditions, il n'est plus possible de trouver un domaine solution \mathcal{D}_t satisfaisant contenant à la fois des éléments de $\underline{\mathcal{D}}_t$ et $\overline{\mathcal{D}}_t$.

Exemple. En guise d'illustration, prenons un modèle statique linéaire de dimension $r=2$ pour lequel des mesures pour trois instants distincts $k \in \{1,2,3\}$ sont disponibles :

$$y(k) = \boldsymbol{\varphi}^T(k)\boldsymbol{\theta}(k) + e(k), \text{ avec } \begin{bmatrix} y(1) \\ y(2) \\ y(3) \end{bmatrix} = \begin{bmatrix} 1.5 \\ 0.8 \\ 2.4 \end{bmatrix} \text{ et } \begin{bmatrix} \boldsymbol{\varphi}^T(1) \\ \boldsymbol{\varphi}^T(2) \\ \boldsymbol{\varphi}^T(3) \end{bmatrix} = \begin{bmatrix} 1.1 & 0.1 \\ 0.2 & 1.2 \\ 1.3 & 0.8 \end{bmatrix}. \quad (\text{V-17})$$

Imposons dans un premier temps une forme prédéterminée en supposant que les incertitudes sont indépendantes et également distribuées sur les deux paramètres de $\boldsymbol{\theta}(k)$, c'est-à-dire que λ est un paramètre scalaire et que la matrice \mathbf{T}_0 se réduit à la matrice identité d'ordre 2 :

$$\boldsymbol{\theta}(k) = \boldsymbol{\theta}_c + \lambda \mathbf{T}_0 \boldsymbol{\nu}(k), \text{ avec } \mathbf{T}_0 = \mathbf{I}_2. \quad (\text{V-18})$$

Dans l'espace paramétrique (θ_1, θ_2) , le domaine abstrait recherché \mathcal{D}_3 est donc un carré, l'objectif étant de déterminer le paramètre λ . Dans un premier temps, les deux meilleurs estimateurs au sens de la norme euclidienne $\boldsymbol{\theta}_{c,2} = [1.45 \quad 0.49]^T$ puis, au sens de la norme infinie $\boldsymbol{\theta}_{c,inf} = [1.43 \quad 0.53]^T$ sont calculés, sachant que ceux-ci sont totalement indépendants de la valeur \bar{e} choisie. Ensuite, l'estimateur le plus précis vérifiant le système constitué de (V-17) et (V-18), et minimisant le critère de précision j_p (V-12) est évalué par la méthode à deux niveaux décrite ci-dessus. Les valeurs optimales du 4-uplet recherché sont $\{\boldsymbol{\theta}_{c,p_1} = [1.40 \quad 0.52]^T, \bar{e}_{p_1} = 0, \lambda_{p_1} = 7.69 \cdot 10^{-2}\}$.

De manière à comparer, il est aussi possible d'imaginer que les incertitudes ne sont pas identiques sur chacun des deux paramètres θ_i en possédant chacune leur propre poids. Dans ce cas, il s'agit d'une forme indéterminée et deux paramètres λ_1 et λ_2 positifs ou nuls existent de manière à ce que matrice $\mathbf{T}(\boldsymbol{\lambda})$ soit alors définie par :

$$\mathbf{T}(\boldsymbol{\lambda}) = [\lambda_1 \mathbf{t}_1 \quad \lambda_2 \mathbf{t}_2] \text{ avec } \mathbf{t}_1 = [1 \quad 0]^T \text{ et } \mathbf{t}_2 = [0 \quad 1]^T. \quad (\text{V-19})$$

L'estimateur le plus précis vérifiant le système constitué de (V-17) et (V-19) et minimisant le critère de précision j_p (V-14) est alors évalué. Les valeurs optimales du 5-uplet recherché sont $\{\boldsymbol{\theta}_{c,p_2} = [1.44 \quad 0.45]^T, \bar{e}_{p_2} = 0, \boldsymbol{\lambda}_{p_2} = [0.12 \quad 0]^T\}$.

Sur la figure suivante, les domaines $\underline{\mathcal{D}}_3(\bar{e})$ et $\overline{\mathcal{D}}_3(\bar{e})$ sont représentés pour trois valeurs distinctes de la borne de l'erreur d'équation. Il est alors facile de remarquer comme prévu que plus la valeur de \bar{e} est importante, plus les domaines $\underline{\mathcal{D}}_3(\bar{e})$ et $\overline{\mathcal{D}}_3(\bar{e})$ se rapprochent ; la taille du domaine recherché \mathcal{D}_3 ne pouvant que diminuer, les incertitudes sur les paramètres

deviennent de moins en moins importantes. Comme expliqué dans la section V.2.3.1, les polytopes $\underline{\mathcal{D}}_3(\bar{e})$ et $\overline{\mathcal{D}}_3(\bar{e})$ s'intersectent en $\theta_{c,inf}$ lorsque la borne \bar{e} vaut l'erreur d'estimation $e_{inf}=0.12$ au sens de la norme infinie. Les différents estimateurs calculés précédemment sont représentés, ainsi que les domaines solution obtenus \mathcal{D}_3 pour les deux estimateurs les plus précis θ_{c,p_i} , $i \in \{1,2\}$. Le carré dessiné en pointillés gris correspond au domaine \mathcal{D}_3 le plus précis obtenu dans le cas d'une forme prédéterminée et centré sur l'estimateur θ_{c,p_1} , alors que le segment de droite en trait continu noir centré sur θ_{c,p_2} correspond au domaine le plus précis dans le cadre de la forme indéterminée (V-19).

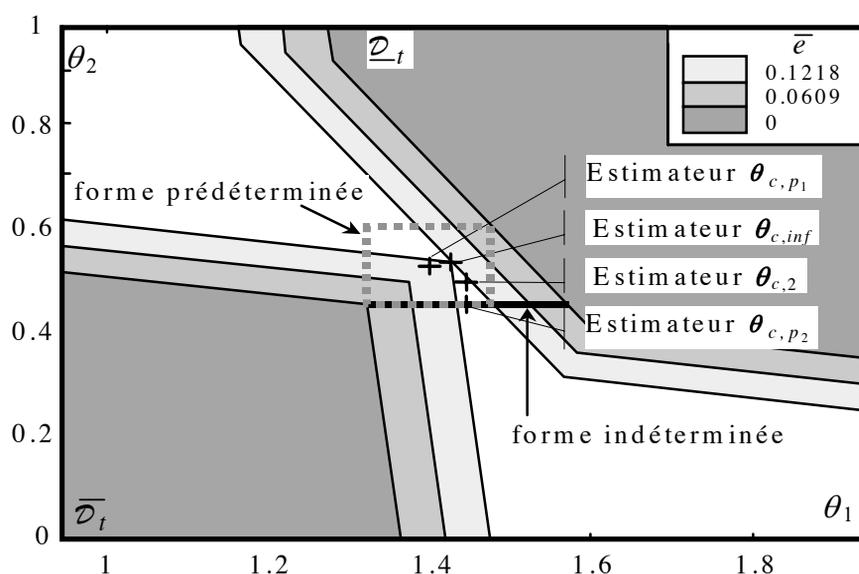


Figure V-10. Approche géométrique

Sur la figure V-11, dans le coin supérieur gauche, sont représentées les évolutions du paramètre scalaire λ en fonction de la borne \bar{e} pour les différents estimateurs, à l'exception de θ_{c,p_2} pour lequel le vecteur λ ne peut être représenté dans ce plan. Les résultats théoriques, explicités lors de la section V.2.3.1, sont là aussi retrouvés. La valeur du paramètre λ diminue bien quand \bar{e} augmente. Lorsque la borne de l'erreur d'équation est trop petite pour englober toutes les erreurs $y(k) - \varphi^T(k)\theta_c$, l'algorithme rejette ce surplus d'erreurs sur les paramètres $\theta(k)$ leur conférant leur nature incertaine. En revanche, lorsque \bar{e} dépasse $e_{inf}(\theta_c) = \sup_k |y(k) - \varphi^T(k)\theta_c|$, le paramètre λ s'annule puisque toutes les observations peuvent être expliquées par l'erreur d'équation $e(k)$.

Le critère de précision j_p est représenté sur la même figure dans le coin inférieur gauche. Les courbes $\lambda(\bar{e})$ étant composées de segments de droites et le critère de précision dépendant linéairement de λ , il est naturel de trouver finalement un critère linéaire par morceaux. Lorsque λ est nul, le critère se résume à $2t\bar{e}$; ce qui explique que quel que soit l'estimateur θ_c , les courbes $j_p(\bar{e}, \lambda(\bar{e}))$ sont confondues pour des valeurs élevées de \bar{e} . Le critère de précision correspondant au cas de l'estimateur θ_{c,p_2} où λ est de dimension 2 a été ajouté. Il apparaît naturellement que le domaine le plus précis correspond à ce choix (la

valeur minimale du critère donnée pour $\{\theta_{c,p_2}, \bar{e}_{p_2}, \lambda_{p_2}\}$ est de 0.65). En se limitant à une forme prédéterminée, le domaine le plus précis correspond à $\{\theta_{c,p_1}, \bar{e}_{p_1}, \lambda_{p_1}\}$ avec une valeur du critère de précision de 0.72.

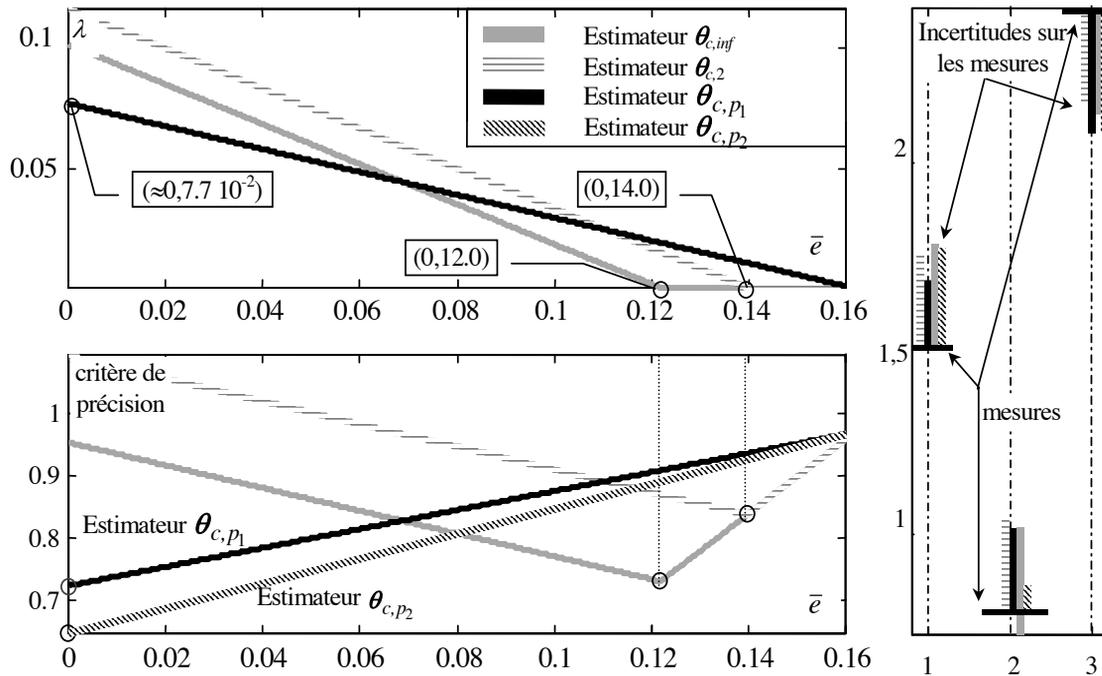


Figure V-11. Résultats de la caractérisation

Sur la courbe de droite, les mesures sont représentées par des traits continus horizontaux et superposés aux espaces abstraits $\mathcal{A}(\varphi^T(k)\theta_c + \varphi^T(k)\mathbf{T}(\lambda)\nu(k) + e(k))$ correspondants aux différents estimateurs. Pour chaque estimateur, les caractéristiques (\bar{e}, λ) retenues sont celles qui correspondent à la valeur minimale du critère de précision, comme précisé sur la figure se trouvant dans le coin supérieur gauche. Les mesures font bien partie des différents espaces abstraits dessinés ; de plus, la forme indéterminée conduit aux espaces abstrait globalement les plus petits. ■

Exemple. Pour terminer d'illustrer ces propos, considérons le système dynamique suivant où les mesures, tout comme les incertitudes paramétriques, interviennent à des instants décalés :

$$y(k) = [0.3y(k-1) \quad -0.02y(k-2)] \begin{bmatrix} \theta(k-1) \\ \theta(k-2) \end{bmatrix} + u(k-2) + e(k),$$

$$\text{avec } \theta(k) = \theta_c + \lambda t_0 \nu(k), y \in \mathbb{R}, \theta \in \mathbb{R}, \lambda \in \mathbb{R}, t_0 \in \mathbb{R}.$$

Ce modèle est simulé de manière à générer le jeu de données nécessaire à la procédure de caractérisation mise en œuvre. Les valeurs de θ_c et \bar{e} étant fixées arbitrairement, seule la valeur du paramètre λ est optimisée. L'évolution de la mesure $y(k)$ est représentée sur la figure V-12 (en trait continu noir), ainsi que l'enveloppe calculée à partir des bornes caractérisées de la variable θ (en trait continu gris). Comme escompté, toutes les mesures sont bien situées à l'intérieur de l'enveloppe calculée. De plus, pour les valeurs de θ_c et \bar{e}

choisies, la valeur de λ est optimale puisque l'une des mesures se trouve exactement sur l'enveloppe. Une valeur plus petite entraînerait un modèle caractérisé qui n'est pas juste, alors qu'une valeur plus élevée induirait un modèle imprécis.

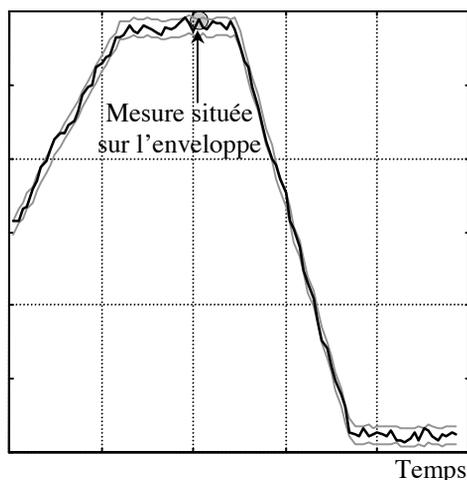


Figure V-12. Mesures et enveloppe

V.4 MULTI-CARACTÉRISATION

V.4.1 Problème de dépendance

L'objectif de cette section est d'étendre le travail précédent au cas de modèles constitués de plusieurs équations de sortie, d'où le terme de multi-caractérisation. Si le principe est similaire à ce qui précède, la principale différence tient du fait qu'une même variable bornée peut apparaître dans plusieurs équations de sortie à la fois et ainsi générer un problème de dépendance tel qu'explicité dans la section II.4.1. Traiter alors le système relation par relation met en défaut la procédure de caractérisation, comme nous allons l'expliciter.

Sur le même schéma que le système (V-1), seuls des modèles structurés linéaires en les paramètres incertains sont considérés :

$$\mathbf{y}(k) = \Phi(k)\boldsymbol{\theta}(k) + \mathbf{e}(k), \text{ avec } \mathbf{y} \in \mathbb{R}^m, \mathbf{e} \in \mathbb{R}^m, \Phi \in \mathbb{R}^{m \times r}, \quad (\text{V-20})$$

La $i^{\text{ième}}$ relation entrée-sortie est alors notée comme suit :

$$y_i(k) = \boldsymbol{\varphi}_i^T(k)\boldsymbol{\theta}(k) + e_i(k), \quad y_i \in \mathbb{R}, \quad e_i \in \mathbb{R}, \quad \boldsymbol{\varphi}_i \in \mathbb{R}^r$$

où $y_i(k)$ est la $i^{\text{ième}}$ sortie mesurée obtenue à l'instant k et $\boldsymbol{\varphi}_i(k)$ est un vecteur de régression contenant d'autres mesures (entrées commandées ou sorties). Le scalaire $e_i(k)$, définissant l'erreur d'équation additive entre la mesure $y_i(k)$ et la sortie du modèle $\boldsymbol{\varphi}_i^T(k)\boldsymbol{\theta}(k)$, est une variable supposée indépendante dont les bornes sont choisies symétriques :

$$-\bar{e}_i \leq e_i(k) \leq \bar{e}_i.$$

L'objectif reste la caractérisation du domaine invariant $\mathcal{D}_t = \mathcal{A}(\boldsymbol{\theta}(k))$ supposé être défini par la relation suivante :

$$\boldsymbol{\theta}(k) = \boldsymbol{\theta}_c + \lambda \mathbf{T}_0 \boldsymbol{\nu}(k), \quad \boldsymbol{\nu} \in \mathbb{R}^q, \quad \mathbf{T}_0 \in \mathbb{R}^{r \times q}. \quad (\text{V-21})$$

Le vecteur $\boldsymbol{\nu}$, qui est constitué de variables bornées normalisées mutuellement indépendantes, est supposé être indépendant des variables représentatives des erreurs d'équation. Les dépendances entre variables communes peuvent être dues aux composantes du vecteur des incertitudes paramétriques $\boldsymbol{\theta}(k)$ (au travers de $\boldsymbol{\Phi}(k)$) et à celles du vecteur borné normalisé $\boldsymbol{\nu}(k)$ (à cause de \mathbf{T}_0). Plus précisément, puisque $\boldsymbol{\theta}(k)$ s'exprime directement en fonction de $\boldsymbol{\nu}(k)$ et puisque la forme de \mathcal{D}_t est imposée par les dépendances entre les variables bornées de ce même vecteur, il est nécessaire de prendre en compte toutes les corrélations entre relations de sorties dues aux variables $\nu_i(k)$. Remarquons que si les sorties $y_i(k)$ ne font pas intervenir de variables bornées θ_i communes, il est possible que la matrice \mathbf{T}_0 induise quand même un couplage entre les variables ν_i . De plus, cette étude se limitera au cadre d'une forme prédéterminée.

Imaginons dans un premier temps que le système (V-20) soit étudié indépendamment composante par composante selon la méthode mono-sortie explicitée dans la section V.3.2. Dans ces conditions, ce traitement individuel conduit à m valeurs λ_i généralement distinctes du paramètre λ recherché et assurant que :

$$y_i(k) \in \mathcal{A}(\boldsymbol{\varphi}_i^T(k) \boldsymbol{\theta}_c + \lambda_i \boldsymbol{\varphi}_i^T(k) \mathbf{T}_0 \boldsymbol{\nu}^i(k) + e_i(k)), \quad i \in \{1, \dots, m\}, \quad k \in \{1, \dots, t\},$$

où les vecteurs bornés normalisés $\boldsymbol{\nu}^i(k)$ ⁴ sont tous de même dimension et indépendants. Ceci se traduit par les m doubles inégalités suivantes :

$$-\lambda \|\boldsymbol{\varphi}_i^T(k) \mathbf{T}_0\|_1 - \bar{e}_i \leq y_i(k) - \boldsymbol{\varphi}_i^T(k) \boldsymbol{\theta}_c \leq \lambda \|\boldsymbol{\varphi}_i^T(k) \mathbf{T}_0\|_1 + \bar{e}_i, \quad (\text{V-22})$$

et conduit d'après (V-15) à :

$$\lambda_i = \sup_{k \in \{1, \dots, t\}} \left\{ \max \left(0, \frac{y_i(k) - \boldsymbol{\varphi}_i^T(k) \boldsymbol{\theta}_c - \bar{e}_i}{\|\boldsymbol{\varphi}_i^T(k) \mathbf{T}_0\|_1}, \frac{-y_i(k) + \boldsymbol{\varphi}_i^T(k) \boldsymbol{\theta}_c + \bar{e}_i}{\|\boldsymbol{\varphi}_i^T(k) \mathbf{T}_0\|_1} \right) \right\}.$$

Une valeur de λ commune à chacune des équations du modèle et permettant de satisfaire toutes les contraintes précédentes doit être trouvée puisque les relations entrée-sortie (V-20) dépendent a priori des mêmes incertitudes. Cette condition impose :

$$\lambda \geq \max_{i \in \{1, \dots, m\}} (\lambda_i).$$

Puisque les λ_i définissent les amplitudes des incertitudes paramétriques permettant d'expliquer toutes les mesures $y_i(k)$, prendre une valeur ne respectant pas cette condition conduit nécessairement à un modèle qui n'est pas juste. De plus, plus les incertitudes de paramètres sont petites, meilleure est la précision du modèle ; dans ces conditions, la solution choisie est :

⁴ Cette notation est utilisée afin de bien insister sur le fait que les dépendances dues aux variables communes ne sont pour l'instant pas prises en compte.

$$\lambda = \max_{i \in \{1, \dots, m\}} (\lambda_i).$$

Malheureusement, cette méthode ne prend pas en considération les dépendances entre plusieurs équations de sortie dues à la présence de variables communes v_i . Par conséquent, le domaine \mathcal{D}_t (et donc λ) est calculé de manière à ce que le vecteur de mesure $\mathbf{y}(k)$ soit situé à l'intérieur de l'orthotope aligné circonscrit $\square \mathcal{A}_\lambda(\mathbf{y}(k))$, au lieu du polytope $\mathcal{A}_\lambda(\mathbf{y}(k)) = \left\{ \Phi^T(k)\boldsymbol{\theta}(k) + \mathbf{e}(k) / \boldsymbol{\theta}(k) \in \mathcal{D}_t \text{ et } \mathbf{e}(k) \in \mathcal{A}(\mathbf{e}) \right\}$ ⁵ définissant normalement l'ensemble des mesures accessibles à un instant k , compte tenu du domaine \mathcal{D}_t caractérisé. Plus exactement, puisque $\mathcal{A}_\lambda(\mathbf{y}(k))$ est contenu dans $\square \mathcal{A}_\lambda(\mathbf{y}(k))$ et que lors de la procédure de détection, c'est l'appartenance de $\mathbf{y}(k)$ vis-à-vis de $\mathcal{A}_\lambda(\mathbf{y}(k))$ qui sera testée, il est tout à fait possible qu'alors, une valeur du vecteur de mesure reflétant un comportement normal n'appartienne plus à $\mathcal{A}_\lambda(\mathbf{y}(k))$. Ceci, d'une part, générerait une fausse alarme, mais surtout supprimerait la propriété de garantie, intérêt de l'approche bornante. L'objectif, dans ce qui suit, est de parvenir à prendre en compte la totalité des dépendances entre relations entrée-sortie afin d'obtenir un domaine juste pour l'étape de détection. Deux solutions s'offrent à nous. Il est toujours possible de travailler sur l'orthotope aligné circonscrit lors de l'étape de détection, mais ce choix engendre une perte de précision. La seconde possibilité est d'adapter la procédure de caractérisation en vue d'un modèle multi-sorties pour réaliser une multi-caractérisation.

Exemple. Considérons le système statique suivant pour générer le jeu de données, dont les sorties sont simulées sur un horizon temporel $t=100$ et représentées par le symbole "+" sur la figure V-13 :

$$\mathbf{y}(k) = \begin{bmatrix} 0 \\ 0 \end{bmatrix} + \begin{bmatrix} 0.8 & 1.2 \\ 1 & -1 \end{bmatrix} \mathbf{v}(k) + \mathbf{e}(k) \text{ avec } \bar{e}_i = 0.2, i \in \{1, 2\}.$$

Le modèle utilisé lors de la procédure de caractérisation est le suivant :

$$\mathbf{y}(k) = \begin{bmatrix} 0 \\ 0 \end{bmatrix} + \lambda \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} \mathbf{v}(k) + \mathbf{e}(k) \text{ avec } \bar{e}_i = 0.1, i \in \{1, 2\}, \quad (\text{V-23})$$

ce qui génère les contraintes suivantes :

$$y_1(k) - 0.1 \leq \lambda_1 (v_1^1(k) + v_2^1(k)) \leq y_1(k) + 0.1$$

$$y_2(k) - 0.1 \leq \lambda_2 (v_1^2(k) - v_2^2(k)) \leq y_2(k) + 0.1$$

Les indices i des variables bornées v_i^j correspondent aux numéros de composantes du vecteur $\mathbf{v}(k)$ alors que les exposants j ne servent qu'à insister sur le fait que dans un premier temps, les deux contraintes précédentes seront traitées individuellement sans tenir compte des dépendances entre variables communes. Notons que la matrice de régression Φ est ici constante⁶ de manière à simplifier la représentation de $\mathcal{A}_\lambda(\mathbf{y}(k))$ qui se trouve dans

⁵ De manière générale, puisque $\Phi(k)$ dépend du temps, la forme de l'espace abstrait $\mathcal{A}_\lambda(\mathbf{y}(k))$ de dimension m varie elle aussi au cours du temps et doit être à chaque instant redéterminée, alors que par définition \mathcal{D}_t est invariant (de forme imposée par T_0) et de dimension r .

⁶ $\Phi(k) = I_2$.

ces conditions invariant, et permet de ne tracer qu'un seul domaine sur tout l'horizon d'étude afin de le comparer aux mesures. Deux procédures de caractérisation mono-sortie sont donc appliquées sur le modèle précédent et donnent les résultats suivants : $\lambda_1=0.98$ et $\lambda_2=0.84$. La valeur choisie de λ est donc $\max(\lambda_1, \lambda_2)=0.98$, ce qui conduit aux domaines $\mathcal{A}_{\lambda=0.98}(\mathbf{y}(k))$ et $\square\mathcal{A}_{\lambda=0.98}(\mathbf{y}(k))$ dessinés sur la figure suivante. Si les mesures sont bien logiquement à l'intérieur du domaine réellement caractérisé $\square\mathcal{A}_{\lambda=0.98}(\mathbf{y}(k))$, il n'en est toutefois pas de même pour $\mathcal{A}_{\lambda=0.98}(\mathbf{y}(k))$. En effet, si chaque mesure $y_i(k)$ dépend d'une unique variable bornée θ_i , des dépendances existent pourtant entre les deux équations de mesures à cause du couplage en les variables bornées v_i introduit par la matrice T_0 .

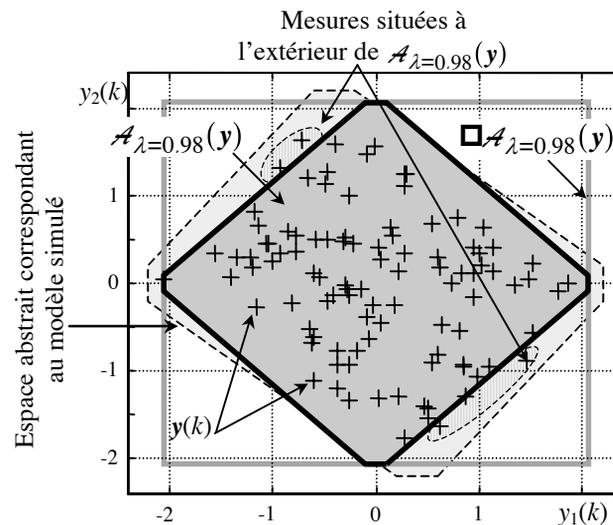


Figure V-13. Résultat de la caractérisation en traitant individuellement chaque relation entrée-sortie

L'espace abstrait des mesures associé au modèle simulé a lui aussi été représenté. Du fait de la taille limitée de l'échantillon, si toutes les mesures se retrouvent logiquement à l'intérieur, aucune d'entre elles n'est située sur l'un de ses hyperplans frontières, le nombre de points de fonctionnement représentés restant faible. Il est bien entendu clair qu'un modèle caractérisé ne sera juste que dans la mesure où tous les points de bon fonctionnement seront atteints dans le jeu de données.

Remarque. Le problème de dépendance dans le cas d'un système multi-sorties n'est pas toujours problématique. En effet, plus les bornes \bar{e}_i imposées sont élevées, plus le paramètre λ à déterminer est faible et plus l'espace abstrait des mesures se rapproche d'un orthotope. Ceci se traduit par des portions de facettes communes plus importantes entre le polytope $\mathcal{A}_{\lambda}(\mathbf{y}(k))$ et son orthotope aligné circonscrit $\square\mathcal{A}_{\lambda}(\mathbf{y}(k))$. De ce fait, de moins en moins de mesures sont susceptibles de sortir de $\mathcal{A}_{\lambda}(\mathbf{y}(k))$ pour la valeur de λ déterminée. Pour une valeur suffisante des \bar{e}_i , toutes les mesures du jeu de données seront situées dans l'espace abstrait $\mathcal{A}_{\lambda}(\mathbf{y}(k))$ résultant, le modèle ainsi obtenu restant juste. Dans le cadre de la figure suivante, par exemple, toutes les mesures sont bien situées dans $\mathcal{A}_{\lambda}(\mathbf{y}(k))$ pour une valeur $\bar{e}_1 = \bar{e}_2 = 0.6$.

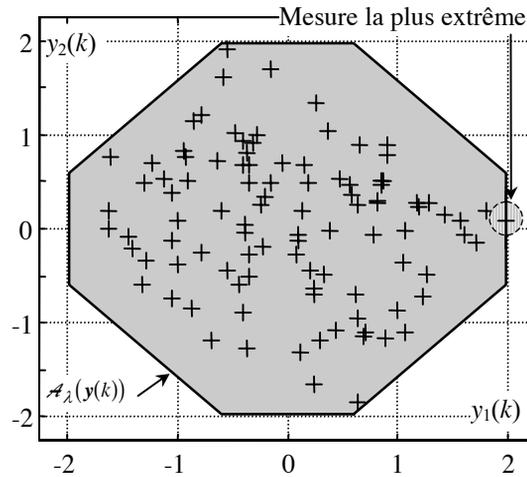


Figure V-14. Résultat de la caractérisation en traitant individuellement chaque relation entrée-sortie

Graphiquement, cela correspond au cas où le vecteur de mesure $y(k)$ imposant le poids λ , (c'est-à-dire la valeur $y(k)$ associée à la plus grande incertitude⁷), se trouve sur une facette du polytope commune avec son orthotope aligné circonscrit. ■

V.4.2 Solution proposée

La solution proposée consiste à prendre en compte les dépendances dues aux variables communes entre les différentes relations entrée-sortie. Puisque celles-ci sont linéaires en les variables bornées v_i , $\mathcal{A}_\lambda(y(k))$ est nécessairement un paralléloétope, c'est-à-dire un domaine convexe parfaitement défini par les contraintes inégalité associées aux demi-espaces frontière dont il représente l'intersection. Le principe procède donc de l'élimination de variables communes et de la génération des bandes de contraintes associées telles qu'explicitées dans les sections IV.4 pour une approche symbolique ou IV.5.3 pour une approche entièrement numérique. En ajoutant ces contraintes supplémentaires à celles initialement imposées par la méthode de caractérisation mono-sortie, la forme exacte de $\mathcal{A}_\lambda(y(k))$ est prise en compte et tous les vecteurs de mesures $y(k)$ devront nécessairement appartenir à ce domaine.

A partir du modèle (V-20) et de l'expression de $\theta(k)$ (V-21), la matrice $X(k, \lambda)$ associée aux variables bornées v_i est définie :

$$y(k) - \Phi(k)\theta_c = e(k) + X(k, \lambda)v(k), \text{ avec } X(k, \lambda) = \lambda\Phi(k)T_0. \quad (\text{V-24})$$

Notons qu'en regroupant toutes les variables bornées $v(k)$ et $e(k)$, cette expression peut se mettre sous la même forme que (IV-10) à partir de laquelle a été développée la méthode de génération des contraintes inégalité d'un paralléloétope :

$$p_{lin,r} = P_{v_r}v_r + p_0, \quad \|v_r\|_\infty \leq 1,$$

⁷ Pour cet exemple, puisque $\mathcal{A}_\lambda(y(k))$ est invariant et que les normes-1 des deux lignes la matrice T_0 sont identiques, cette valeur particulière de $y(k)$ correspond au point le plus éloigné du centre θ_c , ce qui n'est plus forcément le cas sinon.

en posant :

$$\mathbf{p}_0 = \Phi(k)\boldsymbol{\theta}_c - \mathbf{y}(k) \text{ et } \mathbf{P}_{\mathbf{v}_r} = [\mathbf{X}(k, \lambda) \quad \mathbf{I}_m \bar{\mathbf{e}}] \text{ avec } \mathbf{v}_r(k) = [\mathbf{v}^T(k) \quad \mathbf{v}_e^T(k)]^T,$$

où $\mathbf{v}_e(k)$ est un vecteur de variables bornées normalisées indépendantes représentant les erreurs d'équation additives. Par analogie, reprenons la méthode alors proposée et adaptons là à nos nouveaux objectifs. Brièvement, la procédure d'élimination de variables communes s'opère en déterminant l'ensemble des combinaisons $c_i, i \in \{1, \dots, \mathbf{C}_q^{m-1}\}$ de $m-1$ éléments parmi q . La matrice suivante :

$$\mathbf{X}_i(k, \lambda) = [\mathbf{x}_{j_1}(k, \lambda) \quad \dots \quad \mathbf{x}_{j_{m-1}}(k, \lambda)], j_{(\cdot)} \in c_i,$$

est créée, composée des colonnes de $\mathbf{X}(k, \lambda)$, dont les indices correspondent à la combinaison c_i . Si le rang de $\mathbf{X}_i(k, \lambda)$ vaut $m-1$, un vecteur ligne $\mathbf{h}_i^T(k)$ tel que $\mathbf{h}_i^T(k)\mathbf{X}_i(k, \lambda) = 0$ est calculé. En fait, à cause de la structure particulière de $\mathbf{X}(k, \lambda)$, le paramètre λ ne modifie pas le rang de $\mathbf{X}_i(k, \lambda)$, quand il est différent de zéro (c'est-à-dire lorsqu'aucune incertitude paramétrique n'existe). Puisque ce paramètre est encore inconnu lors de cette étape, $\mathbf{h}_i^T(k)$ est calculé en imposant arbitrairement $\lambda=1$ et en travaillant sur $\mathbf{X}_i(k, 1)$ au lieu de $\mathbf{X}_i(k, \lambda)$. Ce vecteur ligne de projection décrit en fait une combinaison linéaire des sorties permettant d'éliminer toutes les variables communes dont les indices font partie de c_i et de générer selon la section IV.5.3 la double inégalité associée à l'une des bandes de contraintes de $\mathcal{A}_\lambda(\mathbf{y}(k))$.

En notant $\bar{\mathbf{e}} = [\bar{e}_1 \quad \dots \quad \bar{e}_m]^T$ et après avoir multiplié le système (V-24) par le vecteur ligne de projection $\mathbf{h}_i^T(k)$, l'analyse par intervalles permet de déduire la double inégalité suivante :

$$-\|\mathbf{h}_i^T(k)\bar{\mathbf{e}} - \|\mathbf{h}_i^T(k)\mathbf{X}(k, \lambda)\| \leq \mathbf{h}_i^T(k)(\mathbf{y}(k) - \Phi(k)\boldsymbol{\theta}_c) \leq \|\mathbf{h}_i^T(k)\bar{\mathbf{e}} + \|\mathbf{h}_i^T(k)\mathbf{X}(k, \lambda)\|, \quad (\text{V-25})$$

où le terme $\mathbf{X}(k, \lambda)$ a été défini en (V-24). Cette expression définit l'une des bandes de contraintes du domaine $\mathcal{A}_\lambda(\mathbf{y})$ obtenu à l'instant k . Remarquons que les variables bornées $e_i(k), i \in \{1, \dots, m\}$, ne peuvent être communes puisque chaque erreur d'équation est supposée indépendante.

Notons $s(k)$ le nombre de bandes de contraintes (c'est-à-dire de vecteurs ligne $\mathbf{h}_i^T(k)$) obtenues à l'instant k en utilisant la méthode précédemment exposée. Dans ces conditions, le paramètre recherché λ doit satisfaire un système de double inégalités constitué des m contraintes initiales (V-22) et des $s(k)$ contraintes supplémentaires (V-25). En considérant maintenant l'horizon temporel $k \in \{1, \dots, t\}$, la valeur optimale du paramètre réel λ correspond à la valeur minimale et positive de λ satisfaisant les contraintes précédentes pour chaque indice k :

$$\lambda = \sup_{\forall k \in \{1, \dots, h\}} \left\{ \max \left(0, \dots, \frac{|y_j(k) - \boldsymbol{\varphi}_j^T(k)\boldsymbol{\theta}_c| - \bar{e}_j}{\|\boldsymbol{\varphi}_j^T(k)\mathbf{T}_0\|_1} \quad \dots \quad \frac{|\mathbf{h}_{i(k)}^T(k)(\mathbf{y}(k) - \Phi(k)\boldsymbol{\theta}_c)| - \|\mathbf{h}_{i(k)}^T(k)\bar{\mathbf{e}}\|}{\|\mathbf{h}_{i(k)}^T(k)\Phi(k)\mathbf{T}_0\|_1} \quad \dots \right) \right\},$$

$$\forall j \in \{1, \dots, m\}, \forall i(k) \in \{1, \dots, s(k)\}$$

Exemple. Reprenons l'exemple précédent (V-23) :

$$y_1(k) = \lambda(v_1(k) + v_2(k)) + e_1(k)$$

$$y_2(k) = \lambda(v_1(k) - v_2(k)) + e_2(k)$$

L'élimination des variables communes $v_1(k)$ et $v_2(k)$ conduit aux égalités suivantes :

$$y_1(k) + y_2(k) = 2\lambda v_1(k) + e_1(k) + e_2(k)$$

$$y_1(k) - y_2(k) = 2\lambda v_2(k) + e_1(k) - e_2(k)$$

ce qui permet de déduire, à chaque instant k , les quatre doubles inégalités à satisfaire :

$$\begin{cases} y_1(k) - 0.1 \leq \lambda(v_1(k) + v_2(k)) \leq y_1(k) + 0.1 \\ y_2(k) - 0.1 \leq \lambda(v_1(k) - v_2(k)) \leq y_2(k) + 0.1 \\ y_1(k) + y_2(k) - 0.2 \leq 2\lambda v_1(k) \leq y_1(k) + y_2(k) + 0.2 \\ y_1(k) - y_2(k) - 0.2 \leq 2\lambda v_2(k) \leq y_1(k) - y_2(k) + 0.2 \end{cases}$$

En utilisant la méthode présentée, la valeur du paramètre recherché est $\lambda=1.08$ et le parallélogramme $\mathcal{A}_{\lambda=1.08}(y(k))$ associé représenté sur la figure suivante, contient bien toutes les mesures disponibles.

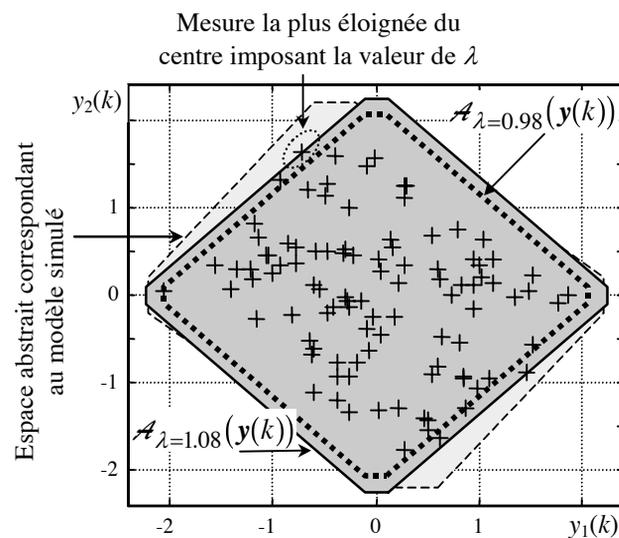


Figure V-15. Résultat de la multi-caractérisation

Le résultat obtenu est bien celui escompté. Remarquons que pour les valeurs de θ_c et \bar{e}_i , $i \in \{1,2\}$, le modèle caractérisé est le plus précis puisqu'une plus petite valeur de λ impliquerait, dans le cas de la forme prédéterminée étudiée, que certaines mesures ne soient plus contenues dans $\mathcal{A}_\lambda(y(k))$. La différence entre $\mathcal{A}_{\lambda=1.08}(y(k))$ et l'espace abstrait correspondant au modèle simulé est liée à la faible taille de l'échantillon utilisé (entraînant que tous les points de fonctionnement n'ont pas été atteints) et aux formes prédéterminées nécessairement différentes imposées par les deux matrices T_0 . ■

Sur le même principe que dans le cadre de la méthode de caractérisation, il est possible d'optimiser les valeurs du centre θ_c et des bornes des erreurs d'équation \bar{e}_i , en utilisant un second niveau d'optimisation reposant par exemple sur un algorithme d'optimisation

non-linéaire tel un simplexe. Dans ces conditions, le critère de précision j_p (V-12) obtenu en sommant la taille des domaines $\mathcal{A}_\lambda(y(k))$ limités à de simples intervalles, doit être adapté au contexte multi-sorties pour lequel $\mathcal{A}_\lambda(y(k))$ est un parallélotope. La solution proposée est la somme des volumes de ce domaine sur l'horizon temporel t .

De manière générale, tous les algorithmes de calcul du volume exact d'un d -polytope⁸ décomposent celui-ci en polytopes élémentaires. Rappelons qu'un polytope convexe \mathcal{P} correspond à l'enveloppe convexe⁹ d'un ensemble fini de points $\mathcal{V}=\{v_1, \dots, v_n\}$ de \mathbb{R}^d . L'ensemble \mathcal{V} définit les sommets de \mathcal{P} , c'est pour cette raison que les termes \mathcal{V} -représentation de \mathcal{P} ou \mathcal{V} -polytope sont utilisés dans la littérature anglo-saxonne (\mathcal{V} pour "vertex"). De manière équivalente, \mathcal{P} est aussi un sous-ensemble borné de \mathbb{R}^d obtenu par intersection d'un nombre fini de demi-espaces fermés :

$$\mathcal{P} = \{\theta \in \mathbb{R}^d \mid H\theta \leq k\},$$

la paire (H, k) est alors appelée \mathcal{H} -représentation de \mathcal{P} ou \mathcal{H} -polytope (\mathcal{H} pour "halfspace").

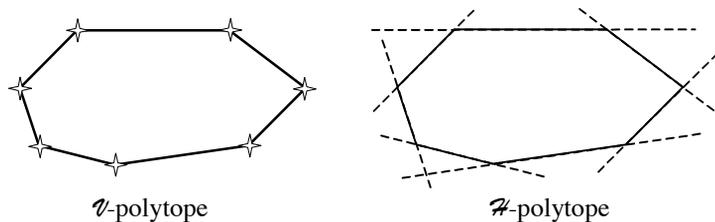


Figure V-16. Différentes représentations d'un polytope convexe (Ziegler, 1998)

Ces polytopes élémentaires sont alors des d -simplex¹⁰, c'est-à-dire des polytopes de dimension d possédant $d+1$ sommets, comme un triangle ou un tétraèdre, qui ont l'avantage d'être caractérisés par un volume facilement calculable. Ces algorithmes peuvent être regroupés en deux types de méthodes distinctes : les méthodes de triangulation et les méthodes de décomposition signée (Büeler et al, 1998).

Les méthodes de triangulation décomposent \mathcal{P} en s d -simplex \mathcal{P}_i ; de ce fait, \mathcal{P} correspond à la réunion de ces polytopes élémentaires et son volume est donné par :

$$\text{vol}(\mathcal{P}) = \sum_{i=1}^s \text{vol}(\mathcal{P}_i).$$

Une méthode de triangulation pour le calcul du volume d'un polytope convexe utilise une triangulation implicite ou explicite à partir d'un point E situé à l'intérieur de \mathcal{P} :

⁸ Un d -polytope est un polytope de dimension d , c'est-à-dire représenté dans un espace à d dimensions. Les polytopes de dimensions 0, 1 et 2 sont des points, segments et polygones.

⁹ Pour tout ensemble $\mathcal{K} \subseteq \mathbb{R}^d$, le plus petit ensemble convexe contenant \mathcal{K} est appelé enveloppe convexe et peut être construit en faisant l'intersection de tous les ensembles convexes contenant \mathcal{K} . Généralement, l'enveloppe convexe de \mathcal{K} , correspondant au terme anglo-saxon "convex hull", est notée $\text{conv}(\mathcal{K})$: $\text{conv}(\mathcal{K}) = \bigcap \{\mathcal{K}' \subseteq \mathbb{R}^d : \mathcal{K} \subseteq \mathcal{K}', \mathcal{K}' \text{ convexe}\}$. Pour tout ensemble fini de points $\{X_1, \dots, X_k\}$ de \mathcal{K} et de paramètres positifs ou nuls $\lambda_1, \dots, \lambda_k$ tels que $\lambda_1 + \dots + \lambda_k = 1$, l'enveloppe convexe $\text{conv}(\mathcal{K})$ contient le point $\lambda_1 X_1 + \dots + \lambda_k X_k$.

¹⁰ Ce terme est issu de son homologue anglo-saxon " d -simplex".

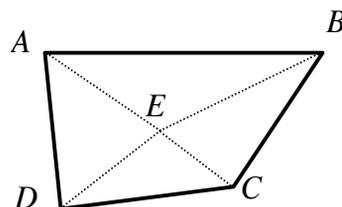


Figure V-17. Méthode de triangulation

Ce type de méthode est plus spécifiquement adapté aux polytopes pour lesquels chaque facette a le nombre minimal de sommets, c'est-à-dire d sommets¹¹ : triangulation de Delaunay (nécessitant une \mathcal{V} -représentation), triangulation combinatoire de Cohen&Hickey (nécessitant à la fois une \mathcal{V} -représentation et une \mathcal{N} -représentation) (Cohen et al, 1979).

Les méthodes de décomposition signée décomposent \mathcal{P} en simplex signés dont la réunion donne exactement \mathcal{P} . Celui-ci est alors représenté par le réunion de s simplex \mathcal{P}_i :

$$\mathcal{P} = \bigcup_{i=1}^s \sigma_i \mathcal{P}_i \text{ où } \sigma_i \text{ peut prendre la valeur } 1 \text{ ou } -1.$$

Ceci se traduit par le fait que tout point X de \mathcal{P} doit exactement apparaître une fois de plus dans les \mathcal{P}_i positifs que dans ceux négatifs et chaque point Y situé en dehors apparaît un même nombre de fois dans les \mathcal{P}_i positifs et négatifs. Le volume de \mathcal{P} est alors :

$$\text{vol}(\mathcal{P}) = \sum_{i=1}^s \sigma_i \text{vol}(\mathcal{P}_i).$$

Pour fixer les idées, prenons l'exemple de la figure suivante où \mathcal{P} est constitué de quatre hyperplans frontière A , B , C et D . Un hyperplan supplémentaire E qui n'est parallèle à aucune des arêtes de \mathcal{P} est généré. En s'intéressant successivement à chacun des sommets de \mathcal{P} , les simplex sont obtenus en construisant les polytopes convexes dont les faces correspondent aux hyperplans frontières de \mathcal{P} générant le sommet étudié, complétés par E . Quatre simplex sont obtenus, dont deux positifs (ADE), (CBE) et deux négatifs (ABE) et (CDE). En effet, en faisant la somme des deux premiers et en soustrayant celle des deux derniers, le polytope \mathcal{P} est retrouvé.

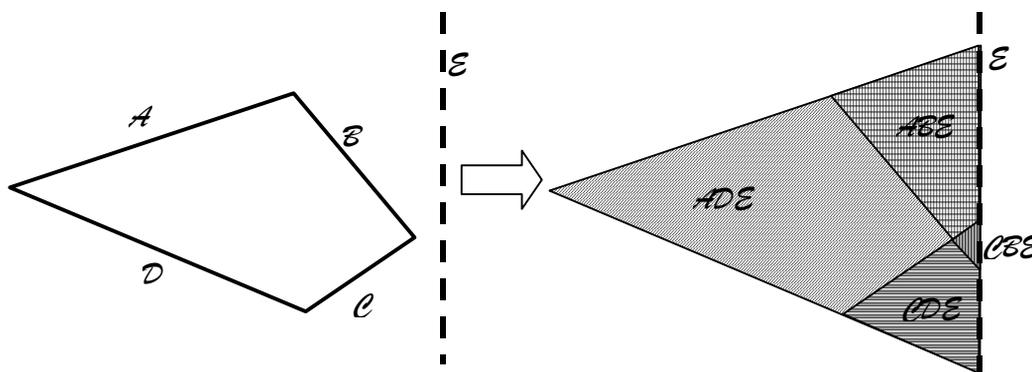


Figure V-18. Exemple de décomposition signée.

¹¹ Dans la littérature internationale, ce type de domaine est désigné par "simplicial polytope". En dimension $d=3$, un octaèdre vérifie par exemple cette propriété, alors que ce n'est pas le cas d'un cube.

Les deux principales méthodes de décomposition signée sont celle de Lawrence (1991) (nécessitant à la fois une \mathcal{V} -représentation et une \mathcal{H} -représentation) et celle de Lasserre, (1983) (nécessitant une \mathcal{H} -représentation). Cette dernière, simple à mettre en œuvre, est brièvement exposée dans l'annexe J. Elles sont plus efficaces lorsqu'il s'agit de polytopes pour lesquels chaque sommet se situe à l'intersection d'un nombre minimal de facettes, c'est-à-dire d facettes¹².

Exemple. Un algorithme d'optimisation à deux niveaux est donc implémenté dans le cadre de l'exemple précédent. La matrice T_0 du modèle (V-23) caractérisé est cette fois prise identique à celle du modèle générant le jeu de données. Le jeu de données est constitué de 5.10^3 vecteurs de mesures $\mathbf{y}(k)$. En supposant que les bornes \bar{e}_i , $i \in \{1,2\}$ sont identiques, les valeurs obtenues sont :

$$\theta_c = \left[-1.10^{-3} \quad 8.10^{-3} \right]^T, \quad \lambda = 0.98 \text{ et } \bar{e}_i = 0.19,$$

ce qui correspond aux résultats escomptés. La différence s'explique par la taille finie de l'échantillon utilisé.

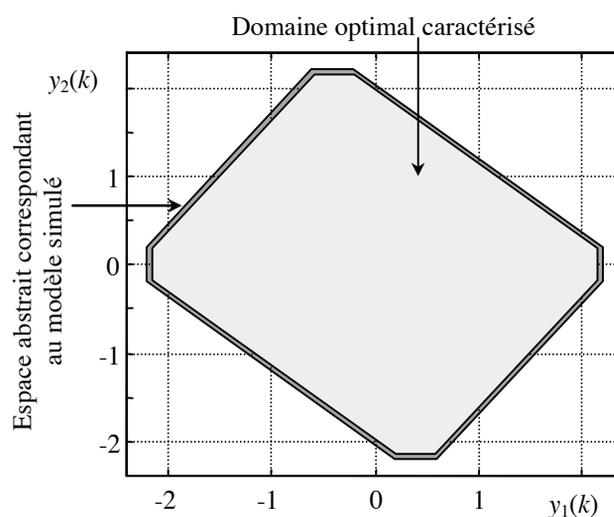


Figure V-19. Domaine caractérisé optimal

V.5 CONCLUSION

Ce chapitre a proposé une méthode de caractérisation de modèles linéaires en les incertitudes représentées sous forme de variables bornées. En plus de déterminer la meilleure répartition de celles-ci du point de vue de la précision du modèle, un niveau d'optimisation supplémentaire permet d'évaluer la valeur centrale des paramètres et les bornes de l'erreur d'équation conduisant à la plus grande précision. Contrairement aux techniques

¹² Le terme anglo-saxon associé est "simple polytope". Un hypercube vérifie cette propriété alors que ce n'est pas le cas d'un octaèdre.

traditionnelles, l'estimateur central ainsi obtenu exploite toute la connaissance disponible sur la répartition des incertitudes.

L'utilisation d'une forme indéterminée pour caractériser l'espace abstrait des incertitudes paramétriques améliore la précision par rapport à une forme déterminée dans la mesure où les degrés de liberté sont plus nombreux. La forme de l'espace abstrait recherché n'est plus totalement imposée et se trouve alors en partie déterminée par l'algorithme de caractérisation. Le problème de dépendance engendré par la présence de variables communes dans le cas d'un modèle possédant plusieurs relations entrée-sortie a été abordé. Une solution consistant à ajouter des contraintes inégalités supplémentaires caractérisant les bandes de contraintes de l'espace abstrait des mesures a été développée dans le cadre d'une forme prédéterminée.

Il est clair le modèle obtenu ne sera juste que dans la mesure où tous les points de bon fonctionnement du système physique étudié auront été représentés dans le jeu de données dont est issue la caractérisation. Mais ce problème n'est pas propre à la méthode envisagée, mais aux méthodes d'estimation paramétrique dans leur ensemble. Notons d'autre part que cette stratégie est à la base d'une méthode de rejet de valeurs aberrantes dans le cas d'un modèle statique (Adrot et al, 1999c).

L'objectif du chapitre suivant est de mettre en œuvre l'ensemble de l'approche bornante sur un exemple suffisamment simple pour permettre d'en détailler toutes les étapes sans que l'aspect calculatoire rébarbatif n'en obscurcisse les principes fondamentaux.

Chapitre VI

Application

TABLE DES MATIÈRES

VI.1 INTRODUCTION.....	247
VI.2 MODÈLE DE SIMULATION.....	248
VI.3 GÉNÉRATION DES RELATIONS DE PARITÉ.....	253
VI.4 TESTS DE COHÉRENCE.....	255
VI.4.1 Méthodes de pavage.....	255
VI.4.2 Procédure de linéarisation.....	257
VI.4.3 Procédure de réduction.....	258
VI.4.4 Procédure de détection de défauts.....	259
VI.4.5 Procédures de localisation de défauts.....	262
VI.5 CONCLUSION.....	267

VI.1 INTRODUCTION

L'intérêt de l'approche bornante a été justifié en présentant dans le premier chapitre, les principes fondamentaux et la problématique du diagnostic de systèmes à base de modèles analytiques ; puis dans le second chapitre, en rappelant les atouts de l'analyse par intervalles qui possède de multiples applications dans bon nombre de domaines de recherche. Les troisième et quatrième chapitres sont directement liés puisque le premier conditionne l'information contenue dans le modèle incertain du système physique considéré en générant des relations de parité exploitables pour les tests de cohérence détaillés dans le second. Le chapitre III, tout en proposant deux stratégies suivant que la matrice de parité à déterminer est certaine ou non, a pour but de générer des résidus totalement découplés des variables d'état inconnues. Le chapitre IV expose lui aussi deux stratégies : la première repose sur les méthodes de pavage alors que la seconde nécessite de linéariser les relations de parité en les variables bornées. Enfin, le précédent chapitre développe une méthode de caractérisation des bornes d'un modèle incertain (ou des résidus qui en découlent) lorsque les relations le composant sont linéaires en les variables bornées.

L'objectif est maintenant d'illustrer les différentes méthodes proposées dans les chapitres précédents en les mettant en œuvre dans le cas d'un exemple académique. La simplicité de celui-ci a l'avantage de permettre au lecteur de suivre plus facilement les différentes étapes considérées en conservant tout au long de l'exposé les expressions symboliques des résidus. D'autres exemples plus complexes (Ploix, 1998) ont été traités mais ils rendent les calculs qui en découlent prohibitifs et le suivi de l'étude ennuyeux par la présentation de résultats complètement numériques.

Pour ce faire, nous nous sommes limités à un modèle purement statique. Si ce choix peut surprendre au premier abord, il se justifie par le fait que de toute manière, l'étude d'un système dynamique se ramène nécessairement dans le cas de l'approche bornante à une représentation statique. D'autre part, malgré la simplicité initiale du modèle, les calculs qui en découlent restent conséquents et le système choisi permet la mise en œuvre de la plupart des méthodes proposées dans ce document. L'intérêt des procédures de linéarisation et de réduction simplifiant l'analyse des résidus est alors mis en évidence. Un exemple expérimental aurait certes été plus probant, mais l'étude proposée dans ce document est avant tout théorique et il n'a jamais été envisagé d'optimiser les algorithmes utilisés afin de leur permettre de traiter un plus grand nombre de données.

Nous détaillons dans un premier temps le modèle de simulation permettant de générer le jeu de données, puis le modèle utilisé pour établir les relations de parité. Après avoir constaté qu'un découplage total des relations de parité vis-à-vis des variables physiques inconnues était trop contraignant dans le cas d'une matrice de parité certaine, la stratégie développée a donc consisté à en rechercher une autre incertaine. Les formes d'évaluation obtenues étant

non-linéaires en les variables bornées, une première étude reposant sur les méthodes de pavage est présentée. Cependant, à cause du nombre d'incertitudes, les temps de calcul deviennent prohibitifs, d'où le choix de détailler plus spécifiquement la stratégie reposant sur la linéarisation des résidus. Les procédures de linéarisation puis de réduction sont donc successivement développées. De manière à limiter la majoration liée à la procédure de linéarisation, le modèle est caractérisé à partir d'un jeu de données fourni par le modèle de simulation en fonctionnement normal. Ensuite, un second jeu de données où plusieurs défauts sont appliqués, est généré. Une procédure de détection de défauts reposant sur la génération des contraintes inégalité représentatives du domaine de bon fonctionnement du système et sur le calcul de la distance normalisée entre l'origine O et la plus proche facette du polytope considéré est présentée. Le fait notamment que le modèle caractérisé est bien plus précis que le modèle linéarisé théorique et qu'il conduit à une meilleure détection des défauts envisagés est vérifié. Tout particulièrement, la question concernant la manière de prendre en compte la notion de distance est détaillée et une première méthode de localisation fondée sur l'étude distincte des parties électrique et mécanique de la machine est développée. Cette stratégie conduit en fait à l'étude de l'orthotope aligné circonscrit au domaine de bon fonctionnement. Pour terminer, une seconde procédure de localisation, reposant sur la méthode proposée dans la section IV.5.5 est mise en œuvre.

VI.2 MODÈLE DE SIMULATION

Pour illustrer les développements des précédents chapitres, cette section présente un modèle statique et incertain d'une machine à courant continu sur lequel sera ultérieurement appliquée une procédure de diagnostic. L'utilisation d'un modèle d'un système physique n'est qu'un prétexte pour faciliter la présentation des résultats en faisant référence à des éléments ou paramètres concrets et ciblés, plus simples à appréhender par le lecteur. Néanmoins cela ne correspond en aucune façon à la volonté de représenter fidèlement la réalité physique.

Le modèle dynamique continu traditionnel d'une machine à courant continu est composé de $p=2$ équations différentielles définissant respectivement les parties électrique et mécanique la constituant :

$$\begin{cases} \frac{d[i^*(t)]}{dt} = -\frac{r}{l}[i^*(t)] - \frac{k}{l}[\omega^*(t)] + \frac{1}{l}[u^*(t)] \\ \frac{d[\omega^*(t)]}{dt} = \frac{k}{j}[i^*(t)] - \frac{f}{j}[\omega^*(t)] \end{cases} .$$

Dans ce modèle, ainsi que dans la suite de cet exemple, les notations comprises entre crochets désignent les variables physiques inconnues du système : la tension $[u^*]$ représente l'entrée et les deux sorties correspondent au courant $[i^*]$ et à la vitesse $[\omega^*]$. Les paramètres électriques r et l désignent respectivement la résistance et l'inductance propre des

enroulements rotoriques, alors que l'inertie de la machine, le coefficient de frottement visqueux et la constante électromagnétique sont respectivement notés j, f et k .

Le modèle statique utilisé est issu du modèle dynamique pour lequel les dérivées sont supposées nulles (ce qui revient à ne conserver que l'information liée au régime permanent du système physique étudié). Nous omettons dorénavant l'indice temporel t qui n'a plus aucune utilité pour ce type de modèle :

$$\begin{cases} -r[i^*] - k[\omega^*] + [u^*] = 0 \\ k[i^*] - f[\omega^*] = 0 \end{cases} \quad (\text{VI-1})$$

En partant de l'idée que ces différents paramètres peuvent être incertains, il est intéressant de se ramener aux valeurs nominales du système étudié, tel que préconisé dans (Ploix, 1998). En remplaçant dans les deux relations précédentes les paramètres incertains r, f et k par les paramètres nominaux (données constructeur) r_n, f_n et k_n , les valeurs nominales des sorties i_n et ω_n s'expriment en fonction de l'entrée nominale u_n :

$$\begin{cases} i_n = \frac{f_n}{f_n r_n + k_n^2} u_n \\ \omega_n = \frac{k_n}{f_n r_n + k_n^2} u_n \end{cases}.$$

Après avoir divisé les deux équations (VI-1) par u_n et en utilisant les deux précédentes relations, un modèle statique de la machine à courant continu est déduit, où les $n=3$ variables physiques normalisées $[u]$, $[i]$ et $[\omega]$ sont obtenues en divisant les grandeurs $[i^*]$, $[\omega^*]$ et $[u^*]$ par leurs valeurs nominales respectives :

$$\begin{cases} -r[i] - \frac{k k_n}{f_n} [\omega] + \frac{f_n r_n + k_n^2}{f_n} [u] = 0 \\ \frac{k f_n}{k_n} [i] - f [\omega] = 0 \end{cases}, \quad (\text{VI-2})$$

$$\text{avec : } [i] = \frac{[i^*]}{i_n}, \quad [\omega] = \frac{[\omega^*]}{\omega_n}, \quad [u] = \frac{[u^*]}{u_n}.$$

Trois incertitudes paramétriques sont introduites : la première sur la résistance r , la seconde sur le coefficient de frottement visqueux f et la dernière sur la constante électromagnétique k .

$$r = (1 + \rho_1 \theta_1) r_n,$$

$$f = (1 + \rho_2 \theta_2) f_n,$$

$$k = (1 + \rho_3 \theta_3) k_n.$$

Le modèle (VI-2) est finalement constitué de deux relations statiques, la première

représentant la partie électrique de la machine et la seconde sa partie mécanique :

$$\begin{cases} -(1 + \rho_1 \theta_1)[i] + (1 + \rho_3 \theta_3) \frac{\gamma - 1}{\gamma} [\omega] + \frac{[u]}{\gamma} = 0 \\ (1 + \rho_2 \theta_2)[\omega] - (1 + \rho_3 \theta_3)[i] = 0 \end{cases}, \text{ avec } \gamma = \frac{r_n f_n}{r_n f_n + k_n^2} = 0.25. \quad (\text{VI-3})$$

Notons que la première incertitude θ_1 n'a d'influence que sur l'équation électrique de la machine alors que la seconde θ_2 n'intervient que dans l'équation mécanique ; enfin θ_3 apparaît dans les deux relations à la fois.

Les équations de mesures associées aux $m=3$ variables physiques sont ajoutées. La tension d'entrée est supposée parfaitement connue ; en revanche, les mesures du courant et de la vitesse sont imparfaites. En notant sans crochet x , la mesure de la variable physique $[x]$, la représentation suivante est adoptée :

$$\begin{cases} u = [u] \\ i = (1 + \rho_4 \theta_4)[i] + \rho_6 \theta_6 \\ \omega = (1 + \rho_5 \theta_5)[\omega] + \rho_7 \theta_7 \end{cases} \quad (\text{VI-4})$$

Les incertitudes $\theta_i, i \in \{4, 5\}$, sont de nature multiplicative et permettent de prendre en compte les non-linéarités de capteurs telles que des phénomènes de saturation magnétique. Les incertitudes $\theta_i, i \in \{6, 7\}$, sont des termes additifs représentant les bruits de capteurs. Toutes les variables bornées $\theta_i, i \in \{1, \dots, 7\}$, sont normalisées et indépendantes (elles fluctuent a priori dans le temps, mais comme expliqué dans l'annexe G ; les considérer invariantes conduit à un résultat identique dans le cas d'un système statique, il n'est donc pas besoin de conserver leur indice temporel t). Les coefficients ρ_i définissent les poids des incertitudes.

Éléments affectés par les incertitudes	Valeurs des coefficients ρ_i
Résistance	$\rho_1 = 0.5$
Frottement visqueux, coefficient électromagnétique	$\rho_2 = \rho_3 = 0.2$
Équations de mesures	$\rho_i = 0.05, i \in \{4, \dots, 7\}$

Tableau VI-1. Valeurs des coefficients $\rho_i, i \in \{1, \dots, 7\}$

Ainsi, le fait que le scalaire ρ_1 vaille 0.5 signifie que la résistance peut varier de plus ou moins 50% autour de sa valeur nominale r_n .

Pour simuler ce système et obtenir les jeux de données nécessaires, les relations (VI-3) sont combinées pour obtenir deux équations où les variables physiques normalisées $[i]$ et $[\omega]$ s'expriment indépendamment en fonction de l'entrée $[u]$:

$$\begin{cases} [i] = \frac{(1 + \rho_2 \theta_2)[u]}{\gamma(1 + \rho_1 \theta_1)(1 + \rho_2 \theta_2) + (1 - \gamma)(1 + \rho_3 \theta_3)^2} \\ [\omega] = \frac{(1 + \rho_3 \theta_3)[u]}{\gamma(1 + \rho_1 \theta_1)(1 + \rho_2 \theta_2) + (1 - \gamma)(1 + \rho_3 \theta_3)^2} \end{cases} \quad (\text{VI-5})$$

Deux jeux de données seront nécessaires pour cet exemple. Le premier, utile à la procédure de caractérisation des bornes des incertitudes paramétriques ne doit comporter que des mesures saines alors que le second au contraire, support de la procédure de détection, doit refléter un comportement anormal. Intéressons-nous plus particulièrement au second (figure VI-4) pour lequel quatre défauts affectant les paramètres sont simulés. Entre les échantillons 34 et 62, la résistance r est imposée à 2.5 fois sa valeur nominale r_n , ce qui correspond à :

$$\theta_1 = \left(\frac{r - r_n}{\rho_1 r_n} \right) = \frac{1.5}{\rho_1} = 3.$$

Entre les échantillons 72 et 100, la constante électromagnétique k est fixée à $1.8k_n$ ($\theta_3=4$), ensuite entre 110 et 138, puis 150 et 178, le coefficient de frottement visqueux f prend la valeur $1.7f_n$ ($\theta_3=3.5$). Ces quatre défauts sont résumés dans le tableau ci-après.

Élément affecté par le défaut	Amplitude du défaut	Échantillons incriminés
Résistance	$r=2.5r_n$	34-62
Constante électromagnétique	$k=1.8k_n$	72-100
Coefficient de frottement visqueux	$f=1.7f_n$	110-138
Coefficient de frottement visqueux	$f=1.7f_n$	150-178

Tableau VI-2. Défauts simulés

En fonctionnement normal, les incertitudes paramétriques évoluent entre leurs bornes respectives, c'est à dire -1 et 1 . Les trois variables bornées θ_i , $i \in \{1,2,3\}$, sont représentées sur la figure suivante :

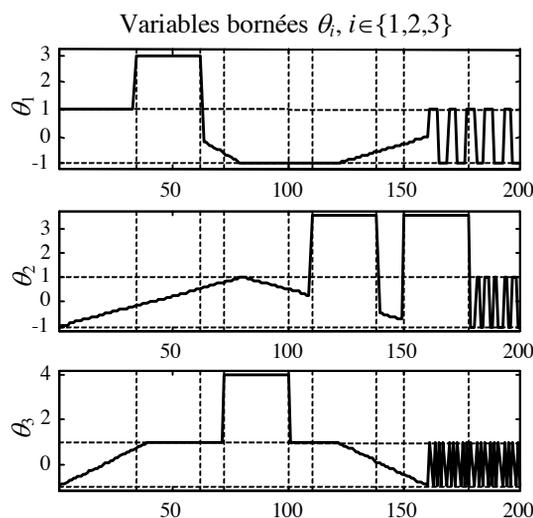


Figure VI-1. Variables bornées θ_i , $i \in \{1,2,3\}$

En l'absence d'anomalie, ces trois variables sont bien normalisées, c'est-à-dire comprises entre 1 et -1 . Rappelons que ce modèle étant statique, ce qui se produit à un échantillon t n'a aucune incidence sur le suivant.

La tension d'entrée $u=[u]$ est tracée sur la figure VI-4. Puis, les variables physiques normalisées $[i]$ et $[\omega]$ sont calculées à partir des deux relations (VI-5) en utilisant les incertitudes paramétriques précédentes. Ensuite, les variables mesurées i et ω sont obtenues à l'aide des équations de mesures (VI-4). Au lieu d'être aléatoires, les incertitudes θ_i , $i \in \{4,5\}$, sont générées de manière à ce que les mesures s'expriment sous la forme d'une fonction tangente hyperbolique $f_i([\cdot]) = \alpha_i \tanh(\beta_i[\cdot])$, $i \in \{4,5\}$, en fonction de leurs variables physiques respectives. Ce choix permet de modéliser un phénomène de saturation magnétique au niveau des capteurs de courant et de vitesse, tout en conservant au niveau du modèle (VI-4) une structure simple.

Puisque les poids ρ_4 et ρ_5 intervenant dans les équations de mesures sont donnés, les coefficients α_i et β_i des deux fonctions non-linéaires précédentes, sont calculés de manière à ce que les valeurs de la fonction non-linéaire $f_i([\cdot])$ restent à l'intérieur du domaine délimité par les deux droites $(1-\rho_i)[\cdot]$ et $(1+\rho_i)[\cdot]$ imposé par les incertitudes multiplicatives θ_i , $i \in \{4,5\}$ (figure VI-2). Une procédure d'optimisation est utilisée afin de minimiser la distance entre $f_i([\cdot])$ et la droite $(1+\rho_i)[\cdot]$ et de satisfaire $f_i([\cdot]_{max}) = (1-\rho_i)[\cdot]_{max}$, où $[\cdot]_{max}$ correspond au maximum de la variable physique $[\cdot]$; ce qui permet de déterminer α_i et β_i .

Le principe appliqué ici est en fait de modéliser une fonction non-linéaire sous la forme d'une relation incertaine en approchant la courbe $f_i([\cdot])$ par le domaine situé entre les deux droites $(1-\rho_i)[\cdot]$ et $(1+\rho_i)[\cdot]$.

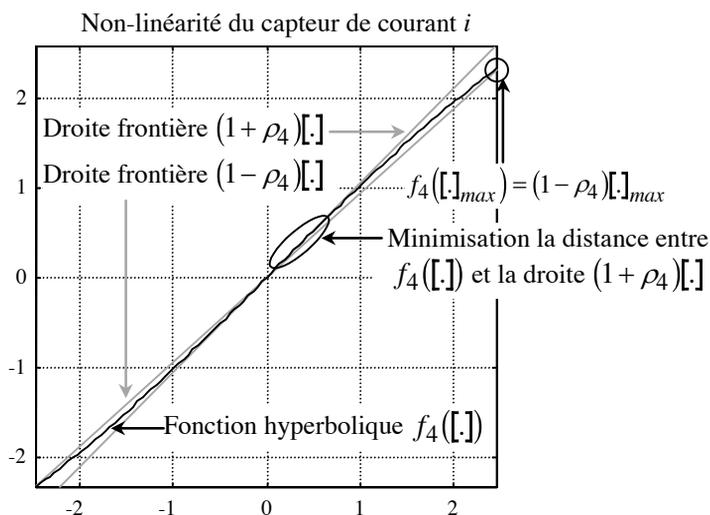


Figure VI-2. Calcul des coefficients α_4 et β_4 de la fonction $f_4([\cdot])$

Les valeurs des incertitudes θ_i , $i \in \{4,5\}$, ne sont donc pas aléatoires, mais dépendent du point de fonctionnement courant, et tout particulièrement des variables physiques $[i]$ et $[\omega]$. Les incertitudes additives θ_i , $i \in \{6,7\}$, représentent les bruit aléatoires de capteurs et sont issus de deux distributions uniformes indépendantes sur l'intervalle $[-1,1]$. Ces quatre variables bornées normalisées sont représentées sur la figure suivante :

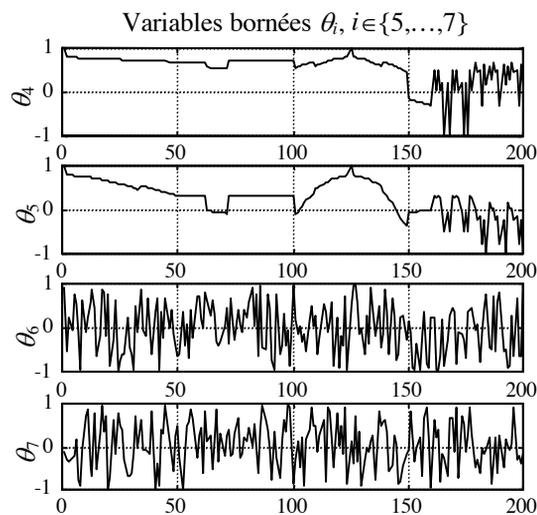


Figure VI-3. Variables bornées $\theta, i \in \{5, \dots, 7\}$

Les mesures u, i et ω sont alors représentées sur la figure suivante :

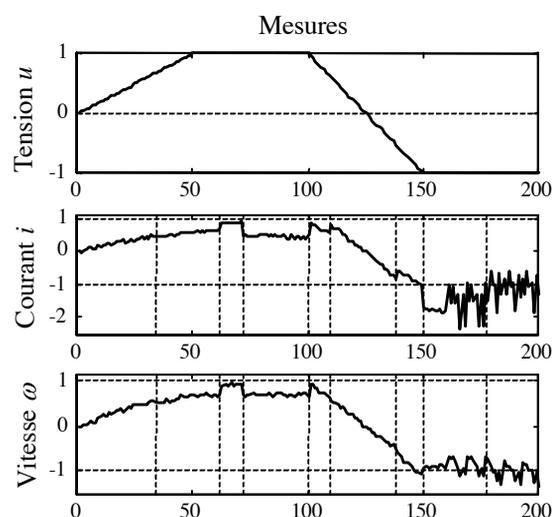


Figure VI-4. Observations u, i et ω .

Notons qu'en valeur absolue, les mesures des deux sorties peuvent dépasser la valeur 1 malgré la normalisation opérée tout simplement à cause des défauts et des incertitudes de paramètres.

VI.3 GÉNÉRATION DES RELATIONS DE PARITÉ

Nous noterons dorénavant $\theta_{i,j}, i < j$, le vecteur contenant, par ordre croissant d'indices, les variables bornées $\theta_k, k \in \{i, i+1, \dots, j\}$. En regroupant l'expression (VI-3) et le système de mesures (VI-4), la forme traditionnelle pour un modèle statique (section III.10.1) exprimant les mesures en fonction des variables physiques se déduit :

$$\begin{bmatrix} y \\ 0 \end{bmatrix} = C(\theta_{1,5})[y] + e(\theta_{6,7}), \quad (\text{VI-6})$$

$$\text{avec } y^T = [i \quad \omega \quad u]^T \text{ et } [y]^T = [[i] \quad [\omega] \quad [u]]^T.$$

Les expressions des différentes matrices sont :

$$C(\theta_{1,5}) = \begin{bmatrix} 1 + \rho_4 \theta_4 & 0 & 0 \\ 0 & 1 + \rho_5 \theta_5 & 0 \\ 0 & 0 & 0 \\ -(1 + \rho_1 \theta_1) & \frac{\gamma - 1}{\gamma} (1 + \rho_3 \theta_3) & \frac{1}{\gamma} \\ -(1 + \rho_3 \theta_3) & 1 + \rho_2 \theta_2 & 0 \end{bmatrix} \text{ et } e(\theta_{6,7}) = \begin{bmatrix} \rho_6 \theta_6 \\ \rho_7 \theta_7 \\ 0 \\ 0 \\ 0 \end{bmatrix}.$$

Maintenant que le modèle et le jeu de données sont définis, déterminons les relations de parité qui nous permettront de réaliser la procédure de détection. Regardons dans un premier temps s'il est possible de construire une matrice de parité certaine W assurant un découplage total des relations de parité vis-à-vis des variables physiques. Pour cela, réécrivons la matrice $C(\theta_{1,5})$ de la manière suivante :

$$C(\theta_{1,5}) = C_0 + \sum_{i=1}^5 C_i \theta_i, \text{ avec } C_0 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \\ -1 & (\gamma - 1)/\gamma & 1/\gamma \\ -1 & 1 & 0 \end{bmatrix}, C_1 = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \\ -\rho_1 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix},$$

$$C_2 = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & \rho_2 & 0 \end{bmatrix}, C_3 = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & \rho_3 (\gamma - 1)/\gamma & 0 \\ -\rho_3 & 0 & 0 \end{bmatrix}, C_4 = \begin{bmatrix} \rho_4 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \text{ et } C_5 = \begin{bmatrix} 0 & 0 & 0 \\ 0 & \rho_5 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}.$$

La recherche d'un vecteur ligne w^T le plus orthogonal possible à la matrice $C(\theta_{1,5})$ quelle que soit la valeur des paramètres incertains, peut s'opérer en minimisant le critère j_c dont l'expression est la suivante :

$$j_c = w^T \underbrace{\left(\int_{-1}^1 \cdots \int_{-1}^1 C(\theta_{1,5}) C(\theta_{1,5})^T d\theta_1 \cdots d\theta_5 \right)}_S w,$$

sous la contrainte $w^T w = 1$. La résolution de ce problème d'optimisation revient à rechercher les valeurs propres de la matrice S définie par (annexe G) :

$$S = 2^5 \left(C_0 C_0^T + \frac{1}{3} (C_1 C_1^T + \cdots + C_5 C_5^T) \right),$$

et qui valent dans l'ordre croissant : 0.2242, 0.3511, 32.02, 91.29, 875.5. La plus petite valeur propre étant loin d'être nulle et puisqu'il n'existe pas, dans le cas d'un modèle statique, de

degré de liberté lié à la taille de fenêtre temporelle qui est ici implicitement imposée à 1, il n'est donc pas possible de trouver, selon la méthode proposée, de matrice de parité certaine satisfaisante.

Dans ces conditions, l'objectif est maintenant de trouver une matrice de parité incertaine $\mathbf{W}(\boldsymbol{\theta}_{1,5})$ orthogonale à la matrice $\mathbf{C}(\boldsymbol{\theta}_{1,5})$ de manière à éliminer les variables physiques inconnues contenues dans le vecteur $[\mathbf{y}]$: $\mathbf{W}(\boldsymbol{\theta}_{1,5})\mathbf{C}(\boldsymbol{\theta}_{1,5})=0$. Le rang générique de la matrice $\mathbf{C}(\boldsymbol{\theta}_{1,5})$ est trois, ce qui conduit à $s_p = (m + p) - 3 = 2$ relations de parité linéairement indépendantes et à la matrice de projection suivante :

$$\mathbf{W}(\boldsymbol{\theta}_{1,5}) = \begin{bmatrix} (1 + \rho_3\theta_3)(1 + \rho_5\theta_5) & \gamma^2(1 + \rho_1\theta_1)(1 + \rho_5\theta_5) \\ -(1 + \rho_2\theta_2)(1 + \rho_4\theta_4) & -\gamma(\gamma - 1)(1 + \rho_3\theta_3)(1 + \rho_4\theta_4) \\ 0 & -\gamma(1 + \rho_4\theta_4)(1 + \rho_5\theta_5) \\ 0 & \gamma^2(1 + \rho_4\theta_4)(1 + \rho_5\theta_5) \\ (1 + \rho_4\theta_4)(1 + \rho_5\theta_5) & 0 \end{bmatrix}^T.$$

Les résidus sont obtenus en multipliant (VI-6) par la matrice de parité précédente $\mathbf{W}(\boldsymbol{\theta}_{1,5})$:

$$\mathbf{W}(\boldsymbol{\theta}_{1,5})\mathbf{e}(\boldsymbol{\theta}_{6,7}) - \mathbf{W}(\boldsymbol{\theta}_{1,5}) \begin{bmatrix} \mathbf{y} \\ 0 \end{bmatrix} = 0,$$

ce qui peut se mettre sous la même forme que (IV-2) :

$$\mathbf{p}(\boldsymbol{\theta}) = \mathbf{p}_l(\boldsymbol{\theta}) + \mathbf{p}_{nl}(\boldsymbol{\theta}), \quad (\text{VI-7})$$

où les indices l et nl définissent respectivement les parties linéaire et non-linéaire du vecteur résidu \mathbf{p} . Les expressions des deux vecteurs \mathbf{p}_l et \mathbf{p}_{nl} sont données par :

$$\mathbf{p}_l(\boldsymbol{\theta}_{1,5}) = \begin{bmatrix} \rho_6\theta_6 - \rho_7\theta_7 - (1 + \rho_3\theta_3 + \rho_5\theta_5)i + (1 + \rho_2\theta_2 + \rho_4\theta_4)\omega \\ \gamma^2\rho_6\theta_6 - \gamma(\gamma - 1)\rho_7\theta_7 - \gamma^2(1 + \rho_1\theta_1 + \rho_5\theta_5)i + \gamma(\gamma - 1)(1 + \rho_3\theta_3 + \rho_4\theta_4)\omega + \dots \\ \dots + \gamma(1 + \rho_4\theta_4 + \rho_5\theta_5)u \end{bmatrix},$$

$$\mathbf{p}_{nl}(\boldsymbol{\theta}) = \begin{bmatrix} \rho_{36}\theta_3\theta_6 + \rho_{56}\theta_5\theta_6 - \rho_{27}\theta_2\theta_7 - \rho_{47}\theta_4\theta_7 + \rho_{356}\theta_3\theta_5\theta_6 - \rho_{247}\theta_2\theta_4\theta_7 - \dots \\ \gamma^2\rho_{16}\theta_1\theta_6 + \gamma^2\rho_{56}\theta_5\theta_6 - \gamma(\gamma - 1)\rho_{37}\theta_3\theta_7 - \gamma(\gamma - 1)\rho_{47}\theta_4\theta_7 + \gamma^2\rho_{156}\theta_1\theta_5\theta_6 - \dots \\ \dots - \rho_{35}\theta_3\theta_5i + \rho_{24}\theta_2\theta_4\omega \\ \dots - \gamma(\gamma - 1)\rho_{347}\theta_3\theta_4\theta_7 - \gamma^2\rho_{15}\theta_1\theta_5i + \gamma(\gamma - 1)\rho_{34}\theta_3\theta_4\omega + \gamma\rho_{45}\theta_4\theta_5u \end{bmatrix}.$$

où le coefficient ρ_{ijk} désigne en réalité le produit $\rho_i\rho_j\rho_k$.

VI.4 TESTS DE COHÉRENCE

VI.4.1 Méthodes de pavage

Dans un premier temps, les relations de parité (VI-7), non-linéaires en les variables bornées, sont traitées par la méthode de pavage détaillée dans la section IV.3.2. L'objectif est de déterminer si l'origine O se situe bien à l'intérieur de l'espace abstrait du vecteur parité \mathbf{p} .

Puisque $\mathcal{A}(\mathbf{p})$ est bien trop complexe à déterminer de manière exacte du fait de sa nature non-linéaire en les variables bornées, une approximation extérieure $\mathcal{A}_p(\mathbf{p})$ est utilisée à sa place. En affinant le découpage, le phénomène de majoration lié au problème de dépendance s'estompe jusqu'à faire tendre $\mathcal{A}_p(\mathbf{p})$ vers le domaine $\mathcal{A}(\mathbf{p})$ recherché. La construction complète de $\mathcal{A}_p(\mathbf{p})$ étant en fait inutile dans le cadre du diagnostic, seuls les pavés contenant O , s'ils existent, sont intéressants à découper.

Bien que cette méthode limite le nombre de pavés calculés pour permettre de conclure à la présence ou non d'une anomalie, la quantité de calculs à effectuer explose de manière exponentielle avec le nombre de variables bornées qui est ici de 7. Le test de cohérence a donc été effectué pour des valeurs très faibles du seuil que ne doit pas dépasser le compteur cpt , ce qui conduit à des découpages grossiers. Les échantillons limites entre lesquels se produisent les défauts, tels que définis dans le tableau VI-2, sont matérialisés par des traits verticaux en pointillés. Une valeur égale à 1 signifie qu'un défaut a été détecté, alors que zéro est synonyme de cohérence entre les mesures et le modèle de référence utilisé.

Pour la figure VI-5, le seuil choisi pour la valeur maximale du compteur cpt est 2. Les plus petits pavés traités sont obtenus après 1 bisection de chacune des facettes et correspondent à une sous-division uniforme d'ordre 2, ce qui conduit à $2^7=128$ pavés si ceux-ci devaient tous être construits par la méthode de pavage de base (section II.4.3).

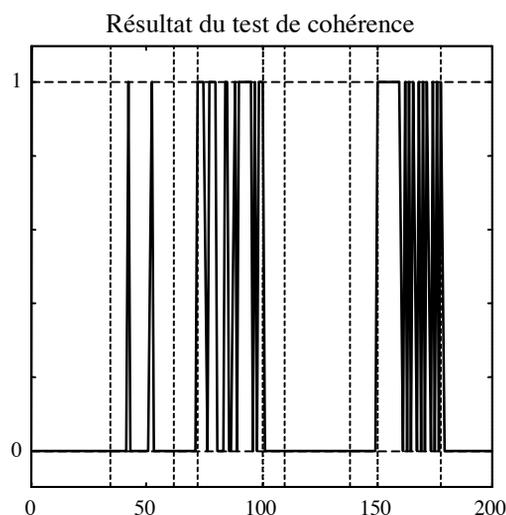


Figure VI-5. Méthode de pavage

Pour la figure VI-6, le seuil choisi pour la valeur maximale du compteur cpt est 3. Les plus petits pavés traités sont obtenus après 2 bisections successives de chacune des facettes et correspondent à une sous-division uniforme d'ordre 4, ce qui conduit à $4^7=16384$ pavés si ceux-ci devaient tous être construits par la méthode de pavage de base. Il est clair qu'en augmentant la valeur du seuil, la précision de la procédure de détection augmente.

Une fois la procédure de linéarisation terminée, le système (VI-7) peut se mettre sous la forme (IV-5) en regroupant par ordre croissant, dans le vecteur borné normalisé $\boldsymbol{\mu}$, les variables $\mu_i, i \in \{1, \dots, 15\}$ précédemment créés :

$$\boldsymbol{p}_{lin}(\boldsymbol{v}) = \boldsymbol{P}_v \boldsymbol{v} + \boldsymbol{p}_0, \text{ avec } \boldsymbol{P}_v = [\boldsymbol{P}_\theta \mid \boldsymbol{P}_\mu] \text{ et } \boldsymbol{v}^T = [\boldsymbol{\theta}^T \mid \boldsymbol{\mu}^T], \|\boldsymbol{v}\|_\infty \leq 1, \quad (\text{VI-8})$$

$$\boldsymbol{P}_\theta = \begin{bmatrix} 0 & \rho_2 \omega & -\rho_3 i & \rho_4 \omega & -\rho_5 i & \rho_6 & -\rho_7 \\ -\gamma^2 \rho_1 i & 0 & \gamma(\gamma-1)\rho_3 \omega & \gamma \rho_4 ((\gamma-1)\omega + u) & \gamma \rho_5 (u - \gamma i) & \gamma^2 \rho_6 & -\gamma(\gamma-1)\rho_7 \end{bmatrix},$$

$$\boldsymbol{P}_\mu = \begin{bmatrix} \rho_{36} & \rho_{56} & -\rho_{27} & -\rho_{47} & \rho_{356} & -\rho_{247} & 0 & \dots \\ 0 & \gamma^2 \rho_{56} & 0 & -\gamma(\gamma-1)\rho_{47} & 0 & 0 & \gamma^2 \rho_{16} & \dots \\ 0 & 0 & 0 & -\rho_{35i} & \rho_{24\omega} & 0 & 0 & 0 \\ \dots & -\gamma(\gamma-1)\rho_{37} & \gamma^2 \rho_{156} & -\gamma(\gamma-1)\rho_{347} & 0 & 0 & -\gamma^2 \rho_{15i} & \gamma(\gamma-1)\rho_{34\omega} & \gamma \rho_{45u} \end{bmatrix},$$

$$\boldsymbol{p}_0 = \begin{bmatrix} \omega - i \\ -\gamma^2 i + \gamma(\gamma-1)\omega + \gamma u \end{bmatrix}.$$

Cette procédure conduit à la génération de 15 nouvelles variables bornées normalisées et indépendantes $\mu_i, i \in \{1, \dots, 15\}$, en sus des 7 intervenant dans les équations initiales du modèle considéré.

VI.4.3 Procédure de réduction

Appliquons maintenant la procédure de réduction (section IV.5.2) de manière à limiter le nombre de variables bornées. Le système (VI-8) étant de faible dimension et restant relativement simple à manipuler, les différents coefficients $\rho_i, i \in \{1, \dots, 7\}$ et γ ont été conservés sous la forme de paramètres formels, et non remplacés par leurs valeurs numériques lors de cette étape de réduction. En notant \boldsymbol{v}_r le vecteur borné normalisé obtenu après réduction, le système précédent peut alors s'écrire comme suit :

$$\boldsymbol{p}_{lin,r}(\boldsymbol{v}_r) = \boldsymbol{P}_{v_r} \boldsymbol{v}_r + \boldsymbol{p}_0, \|\boldsymbol{v}_r\|_\infty \leq 1, \quad (\text{VI-9})$$

$$\boldsymbol{P}_{v_r} = \begin{bmatrix} \rho_{36} + \rho_{27} + \rho_{356} + \rho_{247} + \rho_{35}|i| + \rho_2(1 + \rho_4)|\omega| & \rho_6(1 + \rho_5) & -\rho_7(1 + \rho_4) & -\rho_3 i \\ 0 & \gamma^2 \rho_6(1 + \rho_5) & -\gamma(\gamma-1)\rho_7(1 + \rho_4) & \gamma(\gamma-1)\rho_3 \omega \\ \rho_4 \omega & -\rho_5 i & 0 & \\ \gamma \rho_4((\gamma-1)\omega + u) & \gamma \rho_5(u - \gamma i) & -\gamma^2(1 + \rho_5)(\rho_{16} + \rho_1|i|) - \gamma(1 - \gamma)(\rho_{37} + \rho_{347} + \rho_{34}|\omega|) - \gamma \rho_{45}|u| & \end{bmatrix},$$

en rappelant que dans le cadre de cet exemple, le coefficient γ est inférieur à 1. Puisque multiplier une colonne par -1 ne modifie pas l'espace abstrait $\mathcal{A}(\boldsymbol{p}_{lin,r})$, la matrice \boldsymbol{P}_{v_r} peut aussi s'écrire :

$$\boldsymbol{P}_{v_r} = \begin{bmatrix} \rho_{36} + \rho_{27} + \rho_{356} + \rho_{247} + \rho_{35}|i| + \rho_2(1 + \rho_4)|\omega| & \rho_6(1 + \rho_5) & \rho_7(1 + \rho_4) & -\rho_3 i \\ 0 & \gamma^2 \rho_6(1 + \rho_5) & \gamma(\gamma-1)\rho_7(1 + \rho_4) & \gamma(\gamma-1)\rho_3 \omega \\ \rho_4 \omega & \rho_5 i & 0 & \\ \gamma \rho_4((\gamma-1)\omega + u) & \gamma \rho_5(\gamma i - u) & \gamma^2(1 + \rho_5)(\rho_{16} + \rho_1|i|) + \gamma(1 - \gamma)(\rho_{37} + \rho_{347} + \rho_{34}|\omega|) + \gamma \rho_{45}|u| & \end{bmatrix}. \quad (\text{VI-10})$$

Si l'expression au premier abord semble tout aussi complexe, il ne reste plus que 7 variables bornées indépendantes sur les 22 obtenues après linéarisation. Les tests de cohérence s'en trouveront donc allégés et s'effectueront plus rapidement.

VI.4.4 Procédure de détection de défauts

Au cours de cette section, deux modèles seront utilisés. Le premier, appelé modèle théorique, consiste à conserver en l'état les résidus précédents (VI-9). Il est à opposer au modèle caractérisé, détaillé ultérieurement, sur lequel a été appliqué une procédure de caractérisation des bornes des incertitudes afin de réduire la majoration induite par la procédure de linéarisation.

- *Modèle théorique*

Considérons dans un premier temps le vecteur résidu précédent $\mathbf{p}_{lin,r}$ (VI-9). Pour chaque échantillon de mesures \mathbf{y} , les caractéristiques \mathbf{h}_i^T et k_i des bandes de contraintes \mathcal{S}_i représentatives du paralléloépe $\mathcal{A}(\mathbf{p}_{lin,r})$ associé à (VI-9) sont calculées. Ensuite, est déterminée la distance normalisée séparant l'origine O de l'espace des résidus, de la facette de $\mathcal{A}(\mathbf{p}_{lin,r})$ la plus proche (section IV.5.4). Cet outil permet de nuancer le caractère binaire d'un test de cohérence en indiquant si le système se trouve proche ou non de la limite entre comportements normal et anormal. Rappelons qu'une distance négative signifie que O est situé à l'intérieur de $\mathcal{A}(\mathbf{p}_{lin,r})$, et que par conséquent le comportement du système étudié s'apparente à un comportement normal. Inversement, un signe positif indique que l'origine est en dehors du polytope déterminé, révélant ainsi la présence d'un défaut.

La figure VI-7 représente l'évolution de la distance associée au modèle théorique (en trait continu gris) en fonction des indices des échantillons. Le second défaut ainsi que le dernier sont parfaitement détectés, alors qu'en ce qui concerne le premier, le résultat du test de détection ne révèle une incohérence qu'à des points de fonctionnement spécifiques. Ceci n'est pas une surprise en réalité, puisque les résultats de cette approche sont garantis : une incohérence révèle nécessairement un comportement anormal, mais une cohérence peut très bien correspondre à un défaut masqué par les incertitudes¹, sachant qu'en les cumulant, elles correspondent globalement à une imprécision relative de 110%. Le problème pour détecter le troisième défaut peut s'expliquer par le fait que la tension d'entrée est pratiquement nulle pour les échantillons où il se produit (figure VI-4), en d'autres termes la machine n'est pratiquement plus alimentée et l'excitation du système n'est plus suffisante pour permettre une détection correcte. Enfin, remarquons que l'absence de défauts conduit nécessairement à une cohérence, signe que le modèle utilisé est juste.

¹ Ce problème est inhérent à toute procédure de détection relevant d'un modèle incertain puisque celui-ci est de fait imprécis. Il dépend aussi de la méthode utilisée pour la détection qui conserve plus ou moins l'information contenue dans le modèle initial.

- *Modèle caractérisé*

De manière à réduire la majoration liée à la procédure de linéarisation, le modèle réduit (VI-9) est soumis à une procédure de caractérisation. L'objectif est de calculer les bornes des variables intervalle permettant aux relations de parité $p_{lin,r}$ d'expliquer toutes les mesures du système lorsque celui-ci se comporte normalement. Plus exactement, il s'agit de déterminer le plus précisément possible, un domaine de forme imposée et invariant dans le temps, à l'intérieur duquel varient les incertitudes paramétriques. Dans un premier temps, réécrivons ce système sous la forme standard (V-20) :

$$-p_0 = P_{v_r} v_r + e,$$

avec $v_r = v_{r,c} + \lambda T_0 v'_r$, v'_r vecteur de variables bornées normalisées indépendantes et $\bar{e} = 0$.

Par hypothèse, dans le cadre de cet exemple, les variables bornées constituant le vecteur v_r sont indépendantes, ce qui nécessite de prendre une matrice T_0 diagonale. Aucune connaissance particulière supplémentaire n'étant disponible, la matrice T_0 est donc prise égale à la matrice identité d'ordre 7 : $T_0 = I_7$. La valeur centrale $v_{r,c}$ est logiquement prise égale au vecteur nul, le vecteur v_r étant initialement supposé constitué de variables symétriques. La procédure de caractérisation est effectuée sur un jeu de données différent de celui imposé pour la détection, dans la mesure où des défauts ont été modélisés pour cette dernière et qu'il faut ici impérativement utiliser des données saines. Le nombre d'échantillons est porté à 9000 de manière à accroître le nombre de points de fonctionnement considérés et la valeur de λ obtenue est 0.8038. Ce résultat va permettre d'améliorer la précision de la procédure de détection par rapport au cas où elle repose sur le système (VI-9) correspondant à une valeur théorique de $\lambda=1$. Le nombre de points de fonctionnement appréhendés semble suffisant puisque le système caractérisé permet que les résultats du test de cohérence soient garantis (en l'absence d'anomalie, la distance calculée n'est jamais strictement positive), tout comme (VI-9). Le modèle caractérisé s'écrit donc :

$$p_{lin,r}(v'_r, \lambda) = \lambda P_{v_r} v'_r + p_0, \|v'_r\|_{\infty} \leq 1. \quad (\text{VI-11})$$

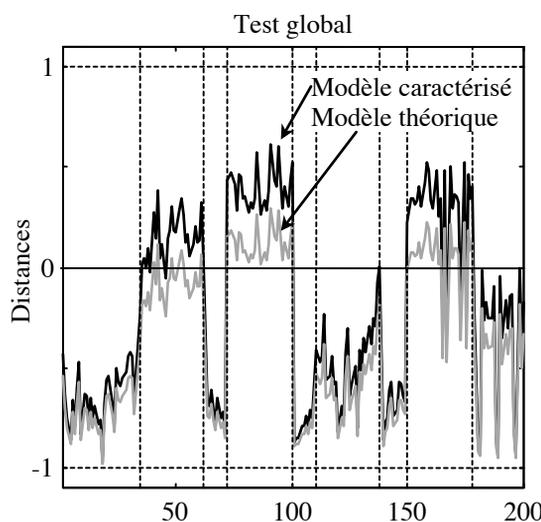


Figure VI-7. Tests de cohérence

Les tests de cohérence globaux pour les modèles caractérisé (courbe noire) et théorique (courbe grise) sont représentés sur la figure VI-7. Le test de cohérence associé au modèle caractérisé conduit à une bien meilleure détection dans la mesure où la distance associée est plus grande en présence d'une anomalie que pour le modèle théorique. En effet, l'origine O est alors plus éloignée de la facette la plus proche \mathcal{F}^* alors que dans un même temps, le centre C en est plus proche. Ceci est particulièrement visible dans le cas du premier défaut affectant la résistance électrique r où celui-ci est pratiquement tout le temps détecté, alors que pour le modèle correspondant à une valeur théorique de $\lambda=1$, le même défaut n'est détecté que ponctuellement. Le modèle caractérisé est donc bien plus précis que son homologue. En fonctionnement normal, la distance ne peut être plus petite que -1 dans la mesure où cette valeur correspond au cas extrême où l'origine O et le centre C sont confondus. De plus, il est là aussi logique de trouver une distance plus petite en valeur absolue pour le modèle caractérisé. Supposons que le fait de passer de $\lambda=1$ à $\lambda=0.8038$ conduise à réduire la distance $\left\| \overrightarrow{OJ^*} \right\|_2$ d'une valeur δ ; il en est alors de même pour $\left\| \overrightarrow{CQ} \right\|_2$, puisqu'une relation homothétique existe entre les polytopes $\mathcal{A}(p_{lin,r}(v_r, \lambda))$ et $\mathcal{A}(p_{lin,r}(v_r))$ représentés sur la figure VI-8.

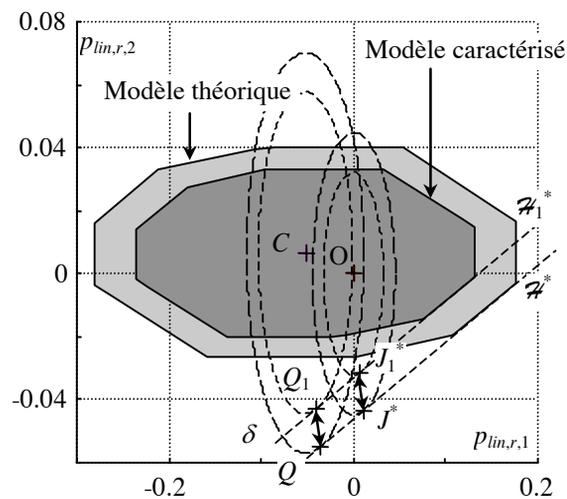


Figure VI-8. Espaces abstraits des modèles caractérisé et théorique (échantillon 10)

En notant J_1^* et Q_1 (respectivement J^* et Q) les projections respectives de O et C sur la facette \mathcal{F}_1^* (\mathcal{F}^*) du parallélogramme associé au modèle caractérisé (respectivement théorique), nous pouvons alors écrire :

$$d = \frac{\left\| \overrightarrow{OJ^*} \right\|_2}{\left\| \overrightarrow{CQ} \right\|_2} \text{ et } d_1 = \frac{\left\| \overrightarrow{OJ_1^*} \right\|_2}{\left\| \overrightarrow{CQ_1} \right\|_2} \text{ avec } \left\| \overrightarrow{OJ_1^*} \right\|_2 = \left\| \overrightarrow{OJ^*} \right\|_2 - \delta \text{ et } \left\| \overrightarrow{CQ_1} \right\|_2 = \left\| \overrightarrow{CQ} \right\|_2 - \delta, \delta > 0.$$

En fonctionnement normal, l'inégalité suivante :

$$\left\| \overrightarrow{CQ} \right\|_2 > \left\| \overrightarrow{OJ^*} \right\|_2,$$

est toujours satisfaite puisque nous avons vu précédemment que la distance ne peut être, dans ces conditions, qu'inférieure à 1 en valeur absolue. Le scalaire δ étant supposé positif,

l'inégalité suivante est donc toujours respectée :

$$0 < \delta \left(\left\| \overrightarrow{CQ} \right\|_2 - \left\| \overrightarrow{OJ^*} \right\|_2 \right),$$

ce qui signifie que la distance d_1 est nécessairement inférieure à d . Dans la suite de cet exemple, seul le modèle caractérisé est étudié.

- *Comparaison avec des tests de cohérence déterministes*

La figure VI-9 correspond aux deux résidus obtenus en ne tenant pas compte des incertitudes de modèle lors de la phase de détection. La relation (VI-7) se limite alors à :

$$p(0) = p_0,$$

où p_0 , dont l'expression est donnée en (VI-8), correspond à la partie nominale du vecteur résidu p . Le test de cohérence, cette fois déterministe, correspond à vérifier si :

$$p_0 = 0.$$

Ce n'est pas parce que le vecteur résidu prend une valeur différente de 0 que nous sommes nécessairement en présence d'un comportement anormal. De manière à limiter le nombre de fausses alarmes, des seuils de détection (droites en traits pointillés gris) sont générés en étudiant le système en fonctionnement normal et en ajustant ces derniers de manière à ce que le vecteur résidu n'en sorte pas en l'absence de défaut. Notons que dans le cadre de cet exemple, cette technique déterministe et simple, ne prenant pas en compte les incertitudes de modèle, ne permet en fait que de détecter un défaut sur le coefficient de frottement visqueux.

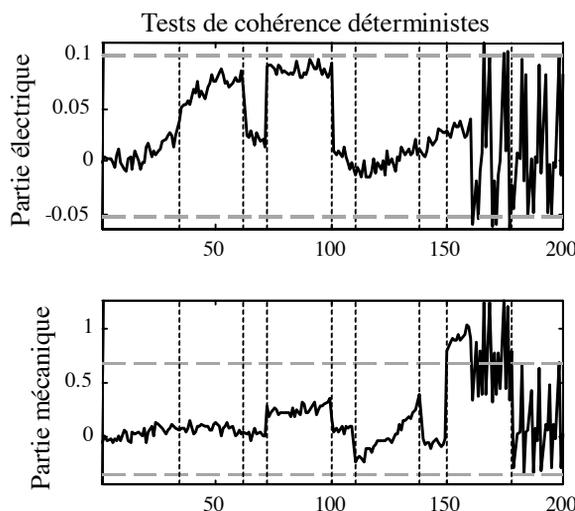


Figure VI-9. Tests de cohérence fondés sur le modèle déterministe de la machine

VI.4.5 Procédures de localisation de défauts

L'étape suivante consiste à compléter la procédure de détection en essayant de localiser les défauts constatés. La méthode revient ici à examiner séparément les parties électrique et

mécanique de la machine étudiée. La même procédure de linéarisation² que pour le modèle complet est appliquée successivement sur les deux systèmes suivants issus de la décomposition de (VI-6) :

$$\begin{bmatrix} y \\ 0 \end{bmatrix} = \mathbf{C}_{elec}(\boldsymbol{\theta}_{1,5/2})[\mathbf{y}] + \mathbf{e}_{elec}(\boldsymbol{\theta}_{6,7}), \text{ avec } \boldsymbol{\theta}_{1,5/2} = [\theta_1 \quad \theta_3 \quad \theta_4 \quad \theta_5]^T$$

$$\mathbf{C}_{elec}(\boldsymbol{\theta}_{1,5/2}) = \begin{bmatrix} 1 + \rho_4\theta_4 & 0 & 0 \\ 0 & 1 + \rho_5\theta_5 & 0 \\ 0 & 0 & 1 \\ -(1 + \rho_1\theta_1) & \frac{\gamma-1}{\gamma}(1 + \rho_3\theta_3) & \frac{1}{\gamma} \end{bmatrix} \text{ et } \mathbf{e}_{elec}(\boldsymbol{\theta}_{6,7}) = \begin{bmatrix} \rho_6\theta_6 \\ \rho_7\theta_7 \\ 0 \\ 0 \end{bmatrix},$$

$$\text{et } \begin{bmatrix} i \\ \omega \\ 0 \end{bmatrix} = \mathbf{C}_{mec}(\boldsymbol{\theta}_{2,5}) \begin{bmatrix} [i] \\ [\omega] \end{bmatrix} + \mathbf{e}_{mec}(\boldsymbol{\theta}_{6,7}), \text{ avec } \boldsymbol{\theta}_{2,5} = [\theta_2 \quad \theta_3 \quad \theta_4 \quad \theta_5]^T$$

$$\mathbf{C}_{elec}(\boldsymbol{\theta}_{2,5}) = \begin{bmatrix} 1 + \rho_4\theta_4 & 0 \\ 0 & 1 + \rho_5\theta_5 \\ -(1 + \rho_3\theta_3) & 1 + \rho_2\theta_2 \end{bmatrix} \text{ et } \mathbf{e}_{mec} = \begin{bmatrix} \rho_6\theta_6 \\ \rho_7\theta_7 \\ 0 \end{bmatrix}.$$

En fait, la partie mécanique (respectivement électrique) correspond à toutes les équations ne contenant pas l'incertitude θ_1 (θ_2) ou toute variable μ_i représentant un monôme dépendant de θ_1 (θ_2).

La matrice de parité $\mathbf{W}_{elec}(\boldsymbol{\theta}_{1,5/2})$ orthogonale à $\mathbf{C}_{elec}(\boldsymbol{\theta}_{1,5/2})$ est donnée par :

$$\mathbf{W}_{elec}(\boldsymbol{\theta}_{1,5/2}) = \begin{bmatrix} \gamma^2(1 + \rho_1\theta_1)(1 + \rho_5\theta_5) & \gamma(1 - \gamma)(1 + \rho_2\theta_2)(1 + \rho_3\theta_3) \cdots \\ \cdots & -\gamma(1 + \rho_4\theta_4)(1 + \rho_5\theta_5) & \gamma^2(1 + \rho_4\theta_4)(1 + \rho_5\theta_5) \end{bmatrix}.$$

En procédant de la même manière que pour le modèle complet, la procédure de linéarisation conduit à :

$$\begin{aligned} \mathbf{P}_{lin,elec}(\mathbf{v}_{elec}) &= \mathbf{P}_{v,elec}^T \mathbf{v}_{elec} + p_{0,elec}, \quad \|\mathbf{v}_{elec}\|_{\infty} \leq 1 \\ \mathbf{P}_{v,elec}^T &= \begin{bmatrix} -\gamma^2 \rho_1 i & \gamma(\gamma-1)\rho_3 \omega & \gamma \rho_4 ((\gamma-1)\omega + u) & \gamma \rho_5 (u - \gamma i) & \gamma^2 \rho_6 & -\gamma(\gamma-1)\rho_7 & \gamma^2 \rho_{56} \cdots \\ \cdots & -\gamma(\gamma-1)\rho_{47} & \gamma^2 \rho_{16} & -\gamma(\gamma-1)\rho_{37} & \gamma^2 \rho_{156} & -\gamma(\gamma-1)\rho_{347} & -\gamma^2 \rho_{15i} & \gamma(\gamma-1)\rho_{34\omega} & \gamma \rho_{45u} \end{bmatrix} \\ p_{0,elec} &= -\gamma^2 i + \gamma(\gamma-1)\omega + \gamma u. \end{aligned}$$

La matrice de parité $\mathbf{W}_{mec}(\boldsymbol{\theta}_{2,5})$ orthogonale à $\mathbf{C}_{mec}(\boldsymbol{\theta}_{2,5})$ est donnée par :

$$\mathbf{W}_{mec}(\boldsymbol{\theta}_{2,5}) = \begin{bmatrix} (1 + \rho_3\theta_3)(1 + \rho_5\theta_5) & -(1 + \rho_2\theta_2)(1 + \rho_4\theta_4) & (1 + \rho_4\theta_4)(1 + \rho_5\theta_5) \end{bmatrix}.$$

En procédant de la même manière que pour le modèle complet, la procédure de linéarisation

² La procédure de réduction est inutile puisque les deux modèles considérés se limiteront à une seule composante.

conduit à :

$$p_{lin,méc}(\mathbf{v}_{méc}) = \mathbf{p}_{v,méc}^T \mathbf{v}_{méc} + p_{0,méc}, \quad \|\mathbf{v}_{méc}\|_{\infty} \leq 1,$$

$$\mathbf{p}_{v,méc}^T = [\rho_2 \omega \quad -\rho_3 i \quad \rho_4 \omega \quad -\rho_5 i \quad \rho_6 \quad -\rho_7 \quad \rho_{36} \cdots \\ \cdots \rho_{56} \quad -\rho_{27} \quad -\rho_{47} \quad \rho_{356} \quad -\rho_{247} \quad -\rho_{35} i \quad \rho_{24} \omega],$$

$$p_{0,méc} = \omega - i.$$

Analyser individuellement les relations de parité des parties électrique et mécanique conduit en réalité dans le cadre de cet exemple à la seconde et première ligne du système global (VI-8) avant réduction. De ce fait, pour améliorer la précision de la détection effectuée, le résultat de la procédure de caractérisation réalisée pour le système complet est réutilisé ici. Les relations de parité réellement analysées sont, avec les notations précédentes :

$$p_{lin,élec}(\mathbf{v}_{élec}) = \lambda \mathbf{p}_{v,élec}^T \mathbf{v}_{élec} + p_{0,élec},$$

$$p_{lin,méc}(\mathbf{v}_{méc}) = \lambda \mathbf{p}_{v,méc}^T \mathbf{v}_{méc} + p_{0,méc}.$$

Pour chaque échantillon de mesures, les bornes des deux intervalles $p_{lin,élec}$ et $p_{lin,méc}$ sont calculées, ainsi que les distances normalisées associées. Les deux distances correspondant respectivement aux parties électrique et mécanique sont représentées sur la figure VI-10. Les deux premiers défauts, dont l'un correspond à un défaut sur la résistance et l'autre sur la constante électromagnétique, conduisent bien à une incohérence au niveau du modèle électrique. Ceci montre que ces défauts affectent bien la partie électrique de la machine. En revanche, la distance associée à la partie mécanique ne permet pas d'affirmer de manière catégorique que le second défaut affecte k , puisque la distance ne devient jamais positive (mis à part en un seul point). Cependant, il est naturel de l'envisager tout de même puisque durant la détection de cette anomalie, cette distance est très voisine de la limite de bon fonctionnement, c'est-à-dire de 0, et ce malgré des points de fonctionnement complètement différents. Enfin, le dernier défaut ne semble toucher que la partie mécanique de la machine, ce qui est légitime pour un défaut sur le coefficient de frottement visqueux.

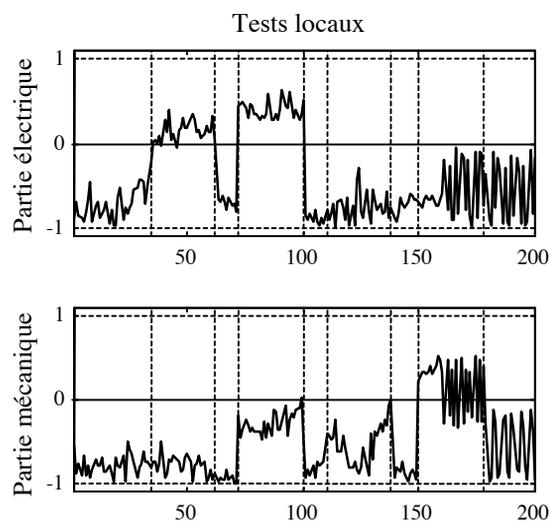


Figure VI-10. Tests de cohérence

Notons que les tests locaux traitant individuellement les parties électrique et mécanique de la machine sont moins précis que le test global. Ceci s'explique facilement par le fait qu'en effectuant les tests locaux, une partie de l'information est perdue dans la mesure où l'on teste si l'origine O est à l'intérieur de l'orthotope aligné circonscrit $\square\mathcal{A}(p_{lin,r}) = \square\mathcal{A}(p_{lin})$, plutôt que de s'intéresser à $\mathcal{A}(p_{lin,r})$. Par conséquent, la détection perd en efficacité si seuls les tests locaux sont utilisés. Ceci se voit notamment au niveau de l'échantillon 171 (figure VI-11), où les deux distances associées à $p_{lin,elec}$ et $p_{lin,méc}$ ne révèlent rien alors qu'un défaut est détecté en considérant l'espace abstrait du modèle complet linéarisé.

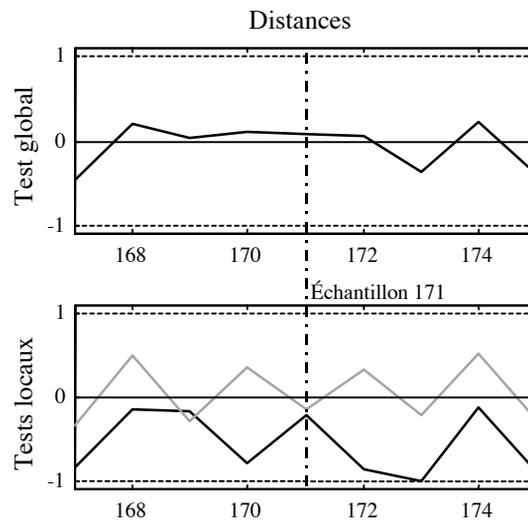


Figure VI-11. Tests de cohérence pour l'échantillon 171

La figure VI-12, représentant l'espace abstrait de $p_{lin,r}$ et son orthotope aligné circonscrit pour l'échantillon 171, confirme cet état de fait, l'origine O appartenant bien à l'orthotope aligné circonscrit $\square\mathcal{A}(p_{lin,r})$, mais se trouvant en dehors du parallélogramme $\mathcal{A}(p_{lin,r})$.

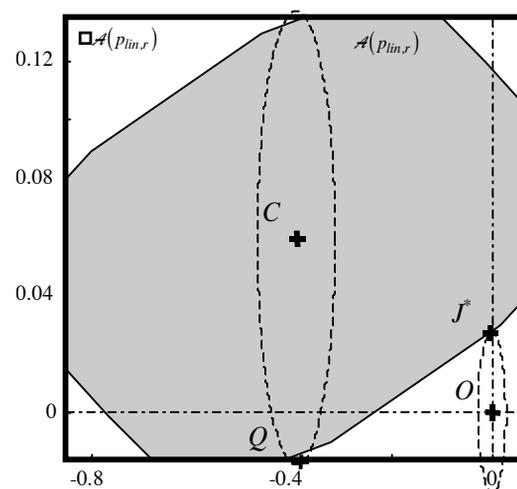


Figure VI-12. Domaines $\square\mathcal{A}(p_{lin,r})$ et $\mathcal{A}(p_{lin,r})$ pour l'échantillon 171

Intéressons-nous maintenant à des défauts de capteurs (courant et vitesse), et plus spécifiquement à la procédure de localisation développée dans la section IV.5.5. Le modèle

réduit complet (VI-9) dépend de 7 variables bornées, dont cinq sont communes (en l'occurrence $v_{r,i}$, $i \in \{2, \dots, 6\}$), c'est-à-dire apparaissent dans les deux composantes du vecteur résidu $\mathbf{p}_{lin,r}$; il en est donc de même pour le modèle caractérisé (VI-11). D'après le squelette de l'algorithme permettant de déterminer les bandes de contraintes d'un paralléloétope (tableau IV-4), les combinaisons de $s_p - 1 = 1$ indice parmi $s_v = 7$ conduisent à construire 7 matrices \mathbf{P} réduites en réalité à chacune des 7 colonnes de $\lambda \mathbf{P}_{v_r}$ où la matrice \mathbf{P}_{v_r} a été définie en (VI-10). Le rang des différentes matrices \mathbf{P} vaut alors nécessairement $s_p - 1$, ce qui signifie que 7 bandes de contraintes définissent $\mathcal{A}(\mathbf{p}_{lin,r})$.

La première et la dernière colonne de $\lambda \mathbf{P}_{v_r}$ correspondent aux deux variables non-communes. Examinons le cas où \mathbf{P} coïncide avec la première colonne de $\lambda \mathbf{P}_{v_r}$. Un vecteur ligne \mathbf{h}_1^T orthogonal à \mathbf{P} peut être $[0 \ 1]$, le scalaire k_1 associé (IV-18) correspond alors à la norme-1 de la seconde ligne de $\lambda \mathbf{P}_{v_r}$:

$$k_1 = \lambda(\gamma^2 \rho_6(1 + \rho_5) + \gamma(1 - \gamma)\rho_7(1 + \rho_4) + \gamma(1 - \gamma)\rho_3|\omega| + \gamma\rho_4|(\gamma - 1)\omega + u| + \gamma\rho_5|\gamma - u| + \dots \\ \dots \gamma^2(1 + \rho_5)(\rho_{16} + \rho_{17}|i|) + \gamma(1 - \gamma)(\rho_{37} + \rho_{347} + \rho_{34}|\omega|) + \gamma\rho_{45}|u|)$$

La bande de contraintes associée étant définie par (IV-12) :

$$\mathcal{S}_1 = \{ \mathbf{p}_{lin,r} / -k_1 \leq \mathbf{h}_1^T \mathbf{p}_{lin,r}(v_r, \lambda) \leq k_1 \},$$

nous obtenons la relation suivante :

$$-k_1 \leq p_{lin,r,2} \leq k_1,$$

où $p_{lin,r,2}$ correspond à la seconde composante du vecteur résidu. En réalité, cette double inégalité définit deux facettes opposées de l'orthotope aligné circonscrit $\square \mathcal{A}(\mathbf{p}_{lin,r})$, et plus exactement donne les bornes encadrant le second résidu. La même étude de la dernière ligne de \mathbf{P}_{v_r} conduit à déterminer l'expression des deux autres facettes de $\square \mathcal{A}(\mathbf{p}_{lin,r})$ et des bornes à l'intérieur desquelles fluctue $p_{lin,r,1}$. L'élimination des 5 variables communes conduit de façon similaire à cinq nouvelles bandes de contraintes supplémentaires tenant compte des dépendances entre composantes du vecteur résidu, ce qui au total conduit à 7 expressions de résidus $p_{lin,r,i}$, $i \in \{1, \dots, 7\}$. Une étude structurelle corrigée par une seconde expérimentale permet de déterminer que deux de ces résidus ne sont sensibles qu'à un défaut sur le capteur de courant alors que deux autres ne sont affectés que par un défaut sur le capteur de vitesse, ce qui permet de générer la table de signatures théorique constituée de quatre résidus.

Les mesures en courant et vitesse sont présentées sur le tracé du haut de la figure VI-13. Les deux zones grisées indiquent les échantillons pour lesquels un défaut est appliqué. Le système est simulé en ajoutant deux biais de capteurs et 100 observations sont générées. Pour les observations, dont l'indice appartient à $[10, 40]$, un biais d'amplitude 0.7 affecte le courant. Pour des indices compris entre 60 et 90, un biais d'amplitude 0.35 affecte la vitesse. Ces défauts peuvent paraître importants, mais la somme des amplitudes des diverses incertitudes s'élève à $\pm 110\%$. Tout d'abord, pour chaque échantillon, toutes les contraintes inégalité définissant $\mathcal{A}(\mathbf{p}_{lin,r})$ sont évaluées et la distance normalisée séparant l'origine O de la plus

proche de ses facettes est calculée. Un signe négatif indique que O se situe dans $\mathcal{A}(p_{lin,r})$ et que le comportement du système s'apparente à celui de son modèle de référence (tracé du centre figure VI-13). Dans une seconde phase, lorsqu'une incohérence est détectée, la procédure de localisation telle qu'expliquée dans la section IV.5.5, est utilisée, où seulement quatre résidus $p_{lin,r,i}$ sont sélectionnés de manière à ce que chaque signature théorique contienne le même nombre de bits 1. Les distances normalisées séparant l'origine O de chacune des quatre facettes en question de $\mathcal{A}(p_{lin,r})$ sont alors calculées.

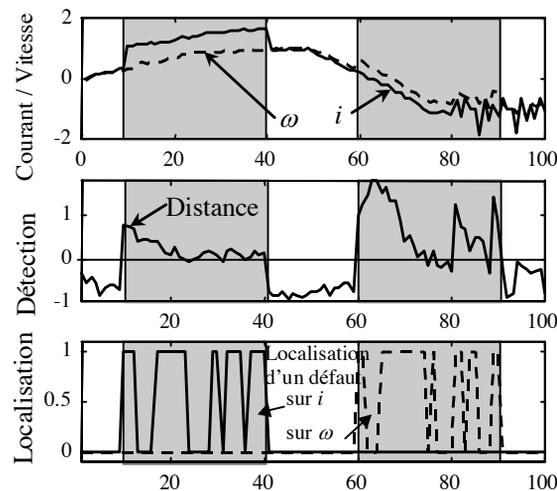


Figure VI-13. Résultats de la procédure de localisation

Globalement, les deux défauts sont bien détectés puisque 88% des mesures défaillantes sont décelées. Comme expliqué précédemment, le résultat de cette procédure de détection est garanti : une incohérence révèle nécessairement la présence d'un comportement anormal puisque le modèle se révèle être juste et que l'absence d'erreurs numériques lors des calculs est supposée. Cependant, seuls 16% des défauts sont localisés, mais ce pourcentage augmente à 75% lorsque la signature expérimentale \mathcal{S}_a est corrigée. Lorsque \mathcal{S}_a contient moins de deux bits 1, la bande de contrainte satisfaite avec la plus petite marge est recherchée, puis \mathcal{S}_a est complétée en remplaçant un des ses 0 par un 1. Sur le dernier tracé de la figure VI-13, un 1 indique qu'un défaut est localisé, alors qu'un 0 signifie, soit que la procédure de localisation n'est pas lancée en cas de bon fonctionnement, soit que la signature expérimentale, corrigée si nécessaire, ne correspond à aucune signature théorique.

VI.5 CONCLUSION

L'exemple proposé nous a permis de mettre en œuvre de manière chronologique différentes procédures de détection et de localisation de défauts, sachant que les techniques utilisées ne sont pas exhaustives. Ainsi, seules des procédures orientées bon fonctionnement ont été réalisées, mais il est tout à fait imaginable d'en concevoir d'autres orientées mauvais fonctionnement dans le cadre de l'étape de localisation.

En ce qui concerne cet exemple, la stratégie reposant sur les méthodes de pavage conduit rapidement à une impasse lorsqu'une précision élevée est demandée. La quantité de calculs numériques à effectuer devient démesurée, toutefois rappelons que les algorithmes utilisés ne sont pas optimisés et que le langage Matlab servant à les programmer n'est pas réputé pour sa rapidité. La stratégie consistant à linéariser les résidus, si elle demande a priori un travail plus poussé dans l'établissement des résidus (procédure de linéarisation et de réduction), permet néanmoins d'éliminer ce problème (dans le cadre de cet exemple) puisque le test de cohérence à effectuer, reposant sur la génération des contraintes inégalité du domaine de bon fonctionnement, est peu gourmand en nombre de calculs. De plus, la notion de distance qui en est issue est intéressante pour l'étape de localisation. Enfin, une étape de caractérisation directement appliquée sur les résidus, permet de limiter la perte de précision engendrée par la linéarisation des résidus. Cette stratégie semble donc plus riche et plus simple à mettre en œuvre que celle reposant sur les méthodes de pavage.

Dans le cas du modèle considéré, les tests locaux coïncident avec l'étude de l'orthotope aligné circonscrit du domaine de bon fonctionnement, les résultats qui en découlent sont logiquement moins précis que pour le test global associé. Néanmoins, ces premiers permettent, dans le cas où une incohérence entre le modèle de référence et les observations est constatée, de localiser en grande partie les défauts rencontrés.

Grâce à la propriété de garantie, l'approche bornante ne génère aucune fausse alarme supplémentaire, autre que celles induites par la non-complétude (ou non-justesse) du modèle. Dans cet exemple, pour le jeu de données traité, le modèle se révèle être juste puisqu'aucune fausse alarme n'est à déplorer, et ce, même pour le modèle caractérisé, ce qui signifie que les données utilisées pour le caractériser étaient pertinentes. Il n'est toutefois pas à exclure que pour des jeux de données différents (que ce soit pour la phase de diagnostic ou celle permettant de trouver le coefficient λ), une incohérence puisse ponctuellement être trouvée lors de la phase de détection lors d'un fonctionnement sain.

Notons pour terminer que même si le modèle est a priori simple, le nombre des paramètres incertains et leurs amplitudes restent toutefois importants par rapport au faible degré de redondance du système étudié. Ceci implique que l'amplitude des défauts détectables soit importante, et qu'il ne soit pas toujours possible de localiser correctement un défaut.

Conclusion

En premier lieu, commençons par présenter le bilan de ce mémoire. Dans le domaine de l'automatique, les méthodes de diagnostic à base de redondance analytique exploitent la connaissance d'un modèle du système physique à surveiller. La difficulté réside dans le fait que ce modèle est pris pour référence et qu'à partir du moment où les mesures prélevées sur le système ne sont plus cohérentes avec celui-ci, la présence d'un défaut est suspectée. Malheureusement, cette représentation mathématique se révèle en pratique approximative, d'où la difficulté de différencier un défaut des conséquences d'une erreur de modélisation. Dans ces conditions, plutôt que de construire des seuils de détection a priori, pourquoi ne pas directement prendre en compte ces erreurs sous forme d'incertitudes paramétriques dès la phase de modélisation, pour ensuite pouvoir naturellement générer ces seuils ? Le modèle, dit incertain, est alors certes imprécis, mais reproduit un ensemble de comportements jugés admissibles, c'est-à-dire représentatifs du système lorsque celui-ci n'est affecté par aucun défaut.

Quelle que soit la méthode sur laquelle repose la procédure de diagnostic, elle ne peut en aucun cas réduire l'imprécision du modèle ; de ce fait, les résultats feront nécessairement apparaître des non-détections. Celles-ci sont directement liées, d'une part à la précision du modèle, d'autre part à la capacité de conserver au mieux l'information lorsque la méthode employée la conditionne sous forme de résidus. L'avantage de l'approche bornante par rapport aux méthodes fonctionnant par découplage provient du fait que les incertitudes paramétriques sont considérées comme porteuses d'information plutôt que d'être purement rejetées. Modélisées sous forme de variables bornées, elles sont complètement intégrées dans le processus de génération de résidus et la réalisation des tests de cohérence. De plus, la propriété de garantie propre à l'approche bornante certifie que si le modèle utilisé est juste, c'est-à-dire représentatif de tous les comportements sains du système, alors une incohérence constatée durant la procédure de détection témoigne nécessairement de la présence d'un défaut. Autrement, les fausses alarmes ne peuvent être induites que par la non-complétude du modèle et ne peuvent en aucun cas provenir de l'approche bornante proprement dite.

Le fait que la représentation ensembliste soit moins riche que la représentation probabiliste, puisque seul le support des incertitudes est considéré, permet de traiter une plus grande classe de systèmes incertains. Il n'y a pas, en effet, à se préoccuper de la stabilité des lois de distribution, ni d'une quelconque hypothèse probabiliste sur la manière dont les paramètres incertains sont distribués. Cette simplification permet de s'intéresser au cas d'incertitudes multiplicatives intervenant sur n'importe quelle matrice de la représentation d'état et pouvant ou non fluctuer dans le temps. De plus, notons que la prise en compte des dépendances entre incertitudes est non seulement possible, mais vivement recommandée de manière à limiter le pessimisme des tests de cohérence en ce qui concerne nos objectifs de diagnostic. Cet aspect a représenté une part importante de ce mémoire, car il est malheureusement bien difficile de résoudre exactement ce problème. De manière générale, l'espace abstrait des résidus, définissant le domaine de bon fonctionnement du système, ne peut être connu de manière exacte. L'objectif est alors d'en rechercher une approximation extérieure, la moins pessimiste possible pour limiter le nombre de non-détections. Une

première stratégie purement numérique reposant sur la méthode de pavage a été proposée. Son principal inconvénient, même s'il est à relativiser dans la mesure où les algorithmes utilisés ne sont pas optimisés, réside dans le volume de calculs à effectuer lorsque le nombre de variables bornées augmente, sachant que la tolérance sur la précision du domaine obtenu peut facilement être réglée par l'utilisateur. En revanche, procédant par linéarisation des résidus en les incertitudes, la seconde stratégie, même si elle doit en partie être réalisée de manière symbolique, est nettement moins sensible à cette difficulté. De plus, nous avons notamment montré l'intérêt des méthodes de caractérisation qui peuvent réduire la majoration précédente lorsqu'elles sont directement appliquées sur les résidus une fois les procédures de linéarisation et de réduction réalisées.

Ce problème de dépendance est aussi à l'origine du choix de travailler sur des modèles à horizon fini de manière à éviter d'emblée l'effet d'enveloppement. Le principe sur lequel repose la génération de relations de parité, même dans un contexte incertain, reste l'élimination des variables d'état inconnues. Nous avons alors justifié l'intérêt de rechercher une matrice de parité incertaine plutôt que constante, car ce dernier cas peut facilement conduire à des impasses en fonction de la nature temporelle des incertitudes et des matrices de la représentation d'état qu'elles affectent. En effet, les conditions d'orthogonalité assurant le découplage des résidus vis-à-vis de l'état du système peuvent alors s'avérer trop contraignantes pour être satisfaites simultanément.

L'approche bornante n'est donc pas la panacée des méthodes de diagnostic. Elle présente aussi bien des avantages que des inconvénients, mais tout comme un modèle se doit d'être adapté aux objectifs de diagnostic fixés, la méthode de diagnostic utilisée doit être en conformité avec les spécificités du modèle choisi ; dans le cas d'un modèle incertain, l'approche bornante semble tout à fait adéquate.

Intéressons-nous maintenant aux perspectives que soulève ce mémoire. Dans la mesure où le sujet de recherche relatif à l'utilisation de techniques ensemblistes dans le cadre du diagnostic reste encore confidentiel et original, il est logique que ce mémoire se soit plus orienté dans une direction théorique plutôt qu'appliquée. Néanmoins, il serait intéressant d'étudier les modalités d'application de l'approche bornante sur un système physique réel.

La notion de distance, séparant le point de fonctionnement courant du système de la limite entre comportements cohérent et incohérent, a aussi été introduite de manière à échapper au caractère binaire d'un test de cohérence classique, mais l'exploitation de cette information nécessite l'intervention humaine. Une stratégie automatique serait la bienvenue.

Au cours du cinquième chapitre concernant les méthodes de caractérisation, l'extension au cas d'un modèle multi-sortie n'a été détaillée que dans le cadre d'une forme prédéterminée. Le même travail est actuellement en cours en ce qui concerne une forme indéterminée et les résultats actuellement prometteurs demandent encore à être validés. La principale difficulté tient dans l'établissement d'un critère de précision représentatif du volume de l'espace abstrait considéré suffisamment simple pour être manipulé.

Tout au long de ce mémoire, nous nous sommes principalement penchés sur la phase de détection, délaissant quelque peu celle de localisation. Nous avons alors mis en exergue les difficultés relatives au fait qu'une simple étude structurelle des résidus n'était pas suffisante pour établir une table de signatures théorique. Cette difficulté est liée à la volonté d'augmenter le nombre de défauts qu'il est possible de localiser en utilisant l'information liée au fait qu'une bande de contraintes soit ou non vérifiée, ce qui engendre des relations de parité non-linéaires en les mesures. Il serait intéressant d'étudier comment faciliter l'établissement de cette table de signature.

De plus, seule une approche orientée bon fonctionnement a été proposée dans le cadre de l'étape de localisation. Une approche orientée mauvais fonctionnement est aussi envisageable, où il faudrait construire un ou plusieurs modèles ensemblistes représentatifs des différents défauts à localiser. Une cohérence entre le comportement observé du système surveillé et l'un de ces modèles conduit alors à déterminer l'organe ou le composant en défaut.

Pour terminer, nous avons supposé tout au long de ce document qu'à deux instants consécutifs, une même incertitude est représentée par deux variables bornées indépendantes. Si de l'information est disponible sur l'évolution de cette incertitude, comme par exemple les bornes encadrant sa dérivée ou un ordre de grandeur de sa dynamique, il est alors de notre intérêt d'utiliser cette connaissance, afin de réduire l'imprécision du modèle, et indirectement celle de la procédure de diagnostic de défauts.

Annexes

TABLE DES MATIÈRES

VII.A ANNEXE A : FORMES CENTRÉES SUPPLÉMENTAIRES	275
VII.B ANNEXE B : RÉOLUTION DU PROBLÈME D'OPTIMISATION DU CHAPITRE III.....	278
VII.C ANNEXE C : EXPRESSION ANALYTIQUE DE LA MATRICE S	279
VII.D ANNEXE D : MATRICE DE PARITÉ W POUR UN MODÈLE CERTAIN	281
VII.E ANNEXE E : RELATION D'ÉQUIVALENCE ENTRE LES DEUX FORMES STATIQUES DANS LE CAS D'UN MODÈLE CERTAIN.....	283
VII.F ANNEXE F : RANG ET VALEURS PROPRES D'UNE MATRICE INCERTAINE	285
VII.G ANNEXE G : ÉTUDE D'UN MODÈLE STATIQUE INCERTAIN	289
VII.H ANNEXE H : GÉNÉRATION DES CONTRAINTES INÉGALITÉS D'UN PARALLÉLOTOPE : CAS PARTICULIERS	292
VII.I ANNEXE I : CALCUL DES SOMMETS D'UN POLYTOPE.....	295
VII.J ANNEXE J : CALCUL DU VOLUME D'UN POLYTOPE	301

Annexe A : Formes centrées supplémentaires

L'objectif de cette annexe est de présenter trois extensions plus sophistiquées que la forme centrée développée dans la section II.3.1.3 et connues pour le faible pessimisme qu'elles engendrent lors de l'évaluation d'une fonction d'inclusion où interviennent des variables bornées dépendantes.

- Forme à valeur moyenne

Cette forme s'applique à des fonctions réelles f dont les dérivées par rapport à chacune des variables x_i existent sur le domaine d'étude. En considérant une fonction monovariante, le théorème de la valeur moyenne permet d'écrire :

$$f(x) = f(c) + \dot{f}(y)(x - c) \text{ pour } y \text{ compris entre } x \text{ et } c \text{ et } x \text{ élément de l'intervalle } \mathbb{X}.$$

En notant \mathbb{D}_{f_x} l'extension intervalle naturelle de la dérivée de f par rapport à x et en remarquant que pour tout y élément de \mathbb{X} , $\dot{f}(y)$ est contenue dans $\mathbb{D}_{f_x}(\mathbb{X})$, l'inclusion suivante est déduite :

$$f(x) \in f(c) + \mathbb{D}_{f_x}(\mathbb{X})(x - c), \quad x \in \mathbb{X}.$$

En notant \mathbb{F}_{vm} l'extension à valeur moyenne de la fonction f et c_i le centre de l'intervalle \mathbb{X}_i , Moore (1979) explique que, de manière générale, l'inclusion suivante est satisfaite :

$$\tilde{f}(\mathbb{X}) \subseteq \mathbb{F}_{vm}(\mathbb{X}) = f(c_1, \dots, c_n) + \sum_{i=1}^n \mathbb{D}_{f_{x_i}}(\mathbb{X})(\mathbb{X}_i - c_i).$$

Exemple. Reprenons la fonction réelle $f(x) = x(1 - x)$, $x \in \mathbb{X} = [0, 1]$, dont la dérivée par rapport à x : $\dot{f}(x) = 1 - 2x$, engendre la fonction intervalle $\mathbb{D}_{f_x} = 1 - 2\mathbb{X}$. L'extension à valeur moyenne de f a donc pour expression :

$$\mathbb{F}_{vm}(\mathbb{X}) = c(1 - c) + (1 - 2\mathbb{X})(\mathbb{X} - c), \quad c = c(\mathbb{X}).$$

Tolérance	Pas $p(k)$	Nombre d'itérations	Nombre de partitions de \mathbb{X}	Intervalle obtenu
10^{-3}	k	9	10	$[-0.0025, 0.2562]$
10^{-3}	2^k	5	64	$[-0.0001, 0.2502]$
10^{-3}	8^k	2	512	$[0, 0.2500]$

Tableau VII-1. Résultats obtenus dans le cas de la forme à valeur moyenne \mathbb{F}_{vm}

Sur l'exemple proposé, le pessimisme sur les bornes obtenues est plus faible que pour la forme centrée (tableau II.19), mais le nombre d'itérations reste encore élevé. Dans ce qui suit, deux nouvelles formes reposant sur la forme à valeur moyenne sont proposées. ■

- Forme bi-centrée

Cette forme, proposée par Neumaier (1990) reprend la précédente, mais cette fois le centre est remplacé par une valeur spécifique. La fonction saturation sat et les variables réelles $\rho_i, \rho_{i,min}, \rho_{i,max}, i \in \{1, \dots, n\}$ sont respectivement définies par :

$$sat(y, \underline{x}) = \begin{cases} \bar{x} & \text{si } y \geq \bar{x} \\ \underline{x} & \text{si } y \leq \underline{x} \\ y & \text{sinon} \end{cases}, \quad \rho_i = sat\left(2 \frac{c(\underline{d}_{fi}(\mathbf{x}))}{w(\underline{d}_{fi}(\mathbf{x}))}, [-1, 1]\right) \quad \text{et} \quad \begin{cases} \rho_{i,min} = c(\underline{x}_i) - \frac{1}{2}\rho_i w(\underline{x}_i) \\ \rho_{i,max} = c(\underline{x}_i) + \frac{1}{2}\rho_i w(\underline{x}_i) \end{cases}.$$

La forme bi-centrée $\mathbb{F}_b(\mathbf{x})$ est alors donnée par l'intersection des deux formes à valeur moyenne de centres respectifs ρ_{min} et ρ_{max} , vecteurs regroupant les $\rho_{i,min}$ et $\rho_{i,max}, i \in \{1, \dots, n\}$:

$$\mathbb{F}_b(\mathbf{x}) = \mathbb{F}_{mv, \rho_{min}}(\mathbf{x}) \cap \mathbb{F}_{mv, \rho_{max}}(\mathbf{x}),$$

$$\text{avec } \mathbb{F}_{vm, \rho_{op}}(\mathbf{x}) = f(\rho_{1,op}, \dots, \rho_{n,op}) + \sum_{i=1}^n \underline{d}_{fi}(\mathbf{x})(\underline{x}_i - \rho_{i,op}), \quad op \in \{min, max\}.$$

En remarquant les équivalences suivantes :

$$-\frac{1}{2} < \frac{c(\underline{d}_{fi}(\mathbf{x}))}{w(\underline{d}_{fi}(\mathbf{x}))} < \frac{1}{2} \Leftrightarrow \underline{d}_{fi}(\mathbf{x}) < 0 \text{ et } \bar{d}_{fi}(\mathbf{x}) > 0,$$

$$\frac{c(\underline{d}_{fi}(\mathbf{x}))}{w(\underline{d}_{fi}(\mathbf{x}))} > \frac{1}{2} \Leftrightarrow \underline{d}_{fi}(\mathbf{x}) > 0 \text{ soit } \rho_i = 1, \rho_{i,min} = \underline{x}_i, \rho_{i,max} = \bar{x}_i,$$

$$\frac{c(\underline{d}_{fi}(\mathbf{x}))}{w(\underline{d}_{fi}(\mathbf{x}))} < -\frac{1}{2} \Leftrightarrow \bar{d}_{fi}(\mathbf{x}) < 0 \text{ soit } \rho_i = -1, \rho_{i,min} = \bar{x}_i, \rho_{i,max} = \underline{x}_i,$$

il s'ensuit que si les fonctions intervalle \underline{d}_{fi} ont un signe constant sur \mathbf{x} (il en est nécessairement de même pour les dérivées de f par définition d'une extension), alors les variables ρ_i prennent leurs valeurs dans l'ensemble $\{-1, 1\}$. Les termes $\rho_{i,min}$ et $\rho_{i,max}$ correspondent dans ces conditions aux bornes des intervalles \underline{x}_i , ce qui permet d'aboutir à l'intervalle minimal : $\mathbb{F}_b(\mathbf{x}) = \tilde{f}(\mathbf{x})$. En effet, supposons pour plus de simplicité que les \underline{d}_{fi} soient positifs, alors la borne inférieure de $\mathbb{F}_{vm, \rho_{min}}(\mathbf{x})$ est $f(\rho_{1,min}, \dots, \rho_{n,min}) = f(\underline{x}_1, \dots, \underline{x}_n)$ et la borne supérieure de $\mathbb{F}_{vm, \rho_{max}}(\mathbf{x})$ est $f(\bar{x}_1, \dots, \bar{x}_n)$. Puisque la forme à valeur moyenne de f contient son intervalle minimal, alors l'intersection de $\mathbb{F}_{vm, \rho_{min}}(\mathbf{x})$ et $\mathbb{F}_{vm, \rho_{max}}(\mathbf{x})$ vaut exactement $\tilde{f}(\mathbf{x})$.

De manière plus générale, Neumaier (1990) démontre que $\inf(\mathbb{F}_{vm, \rho}(\mathbf{x}))$ (respectivement $\sup(\mathbb{F}_{vm, \rho}(\mathbf{x}))$) atteint son maximum (son minimum) pour $\rho = \rho_{min}$ ($\rho = \rho_{max}$), ce qui permet d'optimiser le choix des centres pour une meilleure évaluation.

Exemple. Dans le cas de la fonction $f(x) = x(1-x)$, $x \in \mathcal{X} = [0,1]$, les résultats sont :

Tolérance	Pas $p(k)$	Nombre d'itérations	Nombre de partitions de \mathcal{X}	Intervalle obtenu
10^{-3}	k	2	3	$[0,0.25]$
10^{-3}	2^k	2	8	$[0,0.25]$
10^{-3}	8^k	2	512	$[0,0.25]$

Tableau VII-2. Résultats obtenus dans le cas de \mathbb{F}_b .

Dans le cadre de notre exemple, cette forme donne de meilleurs résultats en termes de précision (par rapport aux bornes théoriques) et de rapidité de convergence que la forme à valeur moyenne. ■

- Monotonicity Test Form

Cette technique reprend la forme à valeur moyenne en l'améliorant par la prise en compte des signes des dérivées par rapport à chacune des variables. Le principe de la méthode proposée consiste à tester le signe des intervalles $\mathbb{D}_i(\mathcal{X})$, $i \in \{1, \dots, n\}$, puis à construire la fonction intervalle $\mathbb{F}_{mt}(\mathcal{X})$ comme suit :

$$\mathbb{F}_{mt}(\mathcal{X}) = [f(u_1, \dots, u_n), f(v_1, \dots, v_n)] + \sum_{i \in \mathcal{S}} \mathbb{D}_i(\mathcal{X})(x_i - c_i),$$

où la paire (u_i, v_i) est définie par :

$$[u_i, v_i] = \begin{cases} [\underline{x}_i, \bar{x}_i] & \text{si } \underline{\mathbb{D}}_i(\mathcal{X}) \geq 0 \\ [\bar{x}_i, \underline{x}_i] & \text{si } \underline{\mathbb{D}}_i(\mathcal{X}) < 0 \text{ et } \bar{\mathbb{D}}_i(\mathcal{X}) \leq 0, \\ [c_i, c_i] & \text{sinon} \end{cases}$$

L'ensemble \mathcal{S} comprend les indices $i \in \{1, \dots, n\}$ correspondant aux variables intervalle x_i telles que 0 soit contenu dans les intervalles $\mathbb{D}_i(\mathcal{X}) : \underline{\mathbb{D}}_i(\mathcal{X}) < 0 < \bar{\mathbb{D}}_i(\mathcal{X})$. Moore (1979) montre que l'intervalle minimal de la fonction f est bien inclus dans \mathbb{F}_{mt} .

Exemple. Dans le cas de la fonction $f(x) = x(1-x)$, $x \in \mathcal{X} = [0,1]$, les résultats sont :

Tolérance	Pas $p(k)$	Nombre d'itérations	Nombre de partitions de \mathcal{X}	Intervalle obtenu
10^{-3}	k	2	3	$[0,0.25]$
10^{-3}	2^k	2	8	$[0,0.25]$
10^{-3}	8^k	2	512	$[0,0.25]$

Tableau VII-3. Résultats obtenus dans le cas de \mathbb{F}_{mt} .

Dans le cadre de notre exemple, cette forme donne des résultats comparables à ceux de la forme bi-centrée. De part sa simplicité, le faible pessimisme de ses résultats et sa rapidité de convergence, cette technique sera employée dans la suite de ce document lorsque des fonctions intervalle affectées par le problème de dépendance seront rencontrées. ■

Annexe B : Résolution du problème d'optimisation du chapitre III

L'objectif est de rechercher la solution au problème suivant : trouver le vecteur ligne \mathbf{w}^T minimisant le carré de la norme euclidienne $\|\mathbf{w}\|_S^2$ et satisfaisant une contrainte de normalité (Adrot et al, 1999b) :

$$\begin{cases} \min_{\mathbf{w}}(\mathbf{w}^T \mathbf{S} \mathbf{w}) \rightarrow \text{Minimisation de } \|\mathbf{w}\|_S^2 \\ 1 - \mathbf{w}^T \mathbf{w} = 0 \rightarrow \text{Normalité de } \mathbf{w} \end{cases},$$

où \mathbf{S} est une matrice symétrique de manière à conserver une forme quadratique.

La fonction de Lagrange l suivante est alors définie :

$$l = \frac{1}{2} \mathbf{w}^T \mathbf{S} \mathbf{w} + \frac{1}{2} \mu (1 - \mathbf{w}^T \mathbf{w}),$$

et les conditions d'optimalité au premier ordre qui en découlent s'expriment comme suit :

$$\frac{\partial l}{\partial \mathbf{w}} = \mathbf{S} \mathbf{w} - \mu \mathbf{w} = 0, \quad (\text{VII-1})$$

$$\frac{\partial l}{\partial \mu} = \frac{1}{2} (1 - \mathbf{w}^T \mathbf{w}) = 0. \quad (\text{VII-2})$$

En multipliant à gauche la relation (VII-1) par \mathbf{w}^T puis en utilisant (VII-2), la relation suivante est finalement obtenue :

$$\mathbf{w}^T \mathbf{S} \mathbf{w} - \mu = 0. \quad (\text{VII-3})$$

Le report de l'égalité précédente dans (VII-1) donne :

$$(\mathbf{S} - (\mathbf{w}^T \mathbf{S} \mathbf{w}) \mathbf{I}) \mathbf{w} = (\mathbf{S} - \mu \mathbf{I}) \mathbf{w} = 0.$$

Le vecteur \mathbf{w} correspond donc au vecteur propre à droite de la matrice \mathbf{S} . Minimiser le terme $\mathbf{w}^T \mathbf{S} \mathbf{w}$ d'après l'égalité (VII-3), consiste par conséquent à trouver la plus petite valeur propre μ_{min} de \mathbf{S} et à déterminer le vecteur propre associé \mathbf{v}_{min} :

$$\mathbf{S} \mathbf{v}_{min} = \mu_{min} \mathbf{v}_{min},$$

pour enfin imposer $\mathbf{w}^T = \mathbf{v}_{min}^T$.

Annexe C : Expression analytique de la matrice \mathbf{S}

L'objectif est d'établir l'expression analytique de la matrice \mathbf{S} utilisée dans la procédure d'optimisation permettant l'élimination des variables d'état dans les relations de parité. Considérons la matrice $\mathbf{C}_s(\boldsymbol{\theta})$ définie par la relation suivante :

$$\mathbf{C}_s(\boldsymbol{\theta}) = \mathbf{C}_{s,0} + \sum_{i=1}^r \sum_{j=0}^s \theta_i(k+j) \mathbf{C}_{s,i,j},$$

où $\mathbf{C}_{s,0}$ et $\mathbf{C}_{s,i,j}$, $i \in \{1, \dots, r\}$, $j \in \{0, \dots, s\}$ sont des matrices certaines.

Par définition, la matrice \mathbf{S} s'écrit comme suit :

$$\mathbf{S} = \int_{-1}^1 \cdots \int_{-1}^1 \mathbf{C}_s(\boldsymbol{\theta}) \mathbf{C}_s(\boldsymbol{\theta})^T d\theta_1(k) \cdots d\theta_r(k+s),$$

ce qui conduit à :

$$\mathbf{S} = \int_{-1}^1 \cdots \int_{-1}^1 \left(\mathbf{C}_{s,0} + \sum_{i=1}^r \sum_{j=0}^s \theta_i(k+j) \mathbf{C}_{s,i,j} \right) \left(\mathbf{C}_{s,0} + \sum_{o=1}^r \sum_{t=0}^s \theta_o(k+t) \mathbf{C}_{s,o,t} \right)^T d\theta_1(k) \cdots d\theta_r(k+s),$$

et peut se décomposer en quatre termes distincts :

$$\begin{aligned} \mathbf{S} = & \underbrace{\int_{-1}^1 \cdots \int_{-1}^1 \mathbf{C}_{s,0} \mathbf{C}_{s,0}^T d\theta_1(k) \cdots d\theta_r(k+s)}_{\mathbf{S}_0} + \cdots \\ & \cdots + \underbrace{\int_{-1}^1 \cdots \int_{-1}^1 \mathbf{C}_{s,0} \left(\sum_{o=1}^r \sum_{t=0}^s \theta_o(k+t) \mathbf{C}_{s,o,t} \right)^T d\theta_1(k) \cdots d\theta_r(k+s)}_{\mathbf{S}_1} + \cdots \\ & \cdots + \underbrace{\int_{-1}^1 \cdots \int_{-1}^1 \left(\sum_{i=1}^r \sum_{j=0}^s \theta_i(k+j) \mathbf{C}_{s,i,j} \right) \mathbf{C}_{s,0}^T d\theta_1(k) \cdots d\theta_r(k+s)}_{\mathbf{S}_2} + \cdots \\ & \cdots + \underbrace{\int_{-1}^1 \cdots \int_{-1}^1 \left(\sum_{i=1}^r \sum_{j=0}^s \theta_i(k+j) \mathbf{C}_{s,i,j} \right) \left(\sum_{o=1}^r \sum_{t=0}^s \theta_o(k+t) \mathbf{C}_{s,o,t} \right)^T d\theta_1(k) \cdots d\theta_r(k+s)}_{\mathbf{S}_3} \end{aligned}$$

En remarquant que :

$$\begin{aligned} \int_{-1}^1 \theta_i(k+j) d\theta_i(k+j) &= 0, \\ \int_{-1}^1 \theta_i^2(k+j) d\theta_i(k+j) &= \frac{2}{3}, \end{aligned}$$

les résultats suivants sont obtenus :

$$\begin{aligned}
 S_0 &= \mathbf{C}_{s,0} \mathbf{C}_{s,0}^T \int_{-1}^1 \cdots \int_{-1}^1 d\theta_1(k) \cdots d\theta_r(k+s) \\
 &= 2^{(s+1)r} \mathbf{C}_{s,0} \mathbf{C}_{s,0}^T, \\
 S_1 &= \mathbf{C}_{s,0} \sum_{o=1}^r \sum_{t=0}^s \mathbf{C}_{s,o,t}^T \int_{-1}^1 \cdots \int_{-1}^1 \theta_o(k+t) d\theta_1(k) \cdots d\theta_r(k+s) \\
 &= 2^{(s+1)r-1} \mathbf{C}_{s,0} \sum_{o=1}^r \sum_{t=0}^s \mathbf{C}_{s,o,t}^T \int_{-1}^1 \theta_o(k+t) d\theta_o(k+t), \\
 &= 0
 \end{aligned}$$

$S_2 = 0$ (le calcul est similaire à S_1),

$$\text{et enfin } S_3 = \sum_{i=1}^r \sum_{o=1}^r \sum_{j=0}^s \sum_{t=0}^s \mathbf{C}_{s,i,j} \mathbf{C}_{s,o,t}^T \int_{-1}^1 \cdots \int_{-1}^1 \theta_i(k+j) \theta_o(k+t) d\theta_1(k) \cdots d\theta_r(k+s).$$

Dans le cas de cette dernière égalité, pour des indices i et o différents ou j et t différents, l'intégrale multiple est nulle puisque cela revient à écrire à un moment donné $\int_{-1}^1 \theta_i(k+j) d\theta_i(k+j)$ dont la valeur est zéro. La relation précédente se résume donc aux cas où

les indices i et j sont respectivement égaux à o et t ; elle vaut alors :

$$\begin{aligned}
 S_3 &= \sum_{i=1}^r \sum_{j=0}^s \mathbf{C}_{s,i,j} \mathbf{C}_{s,i,j}^T \int_{-1}^1 \cdots \int_{-1}^1 \theta_i^2(k+j) d\theta_1(k) \cdots d\theta_r(k+s) \\
 &= 2^{(s+1)r-1} \sum_{i=1}^r \sum_{j=0}^s \mathbf{C}_{s,i,j} \mathbf{C}_{s,i,j}^T \int_{-1}^1 \theta_i^2(k+j) d\theta_i(k+j) \\
 &= \frac{2^{(s+1)r}}{3} \sum_{i=1}^r \sum_{j=0}^s \mathbf{C}_{s,i,j} \mathbf{C}_{s,i,j}^T
 \end{aligned}$$

En regroupant les résultats précédents, l'expression analytique de la matrice \mathbf{S} est alors déduite :

$$\mathbf{S} = 2^{(s+1)r} \left(\mathbf{C}_{s,0} \mathbf{C}_{s,0}^T + \frac{1}{3} \left(\mathbf{C}_{s,1,0} \mathbf{C}_{s,1,0}^T + \cdots + \mathbf{C}_{s,r,s} \mathbf{C}_{s,r,s}^T \right) \right).$$

Annexe D : Matrice de parité \mathbf{W} pour un modèle certain

L'objectif est de montrer que la décomposition en valeurs singulières de la matrice certaine \mathbf{C}_s permet d'obtenir le noyau de cette dernière, c'est-à-dire une base génératrice de toutes les matrices de parité \mathbf{W} pouvant lui être orthogonales.

La décomposition en valeurs singulières de la matrice \mathbf{C}_s (III-2) de dimensions $s_m \times n$ s'écrit de la manière suivante :

$$\mathbf{C}_s = \mathbf{U}\mathbf{\Sigma}\mathbf{V}^T \text{ avec } \mathbf{\Sigma} = \left[\begin{array}{ccc|c} \varepsilon_1 & 0 & 0 & 0 \\ 0 & \ddots & 0 & 0 \\ 0 & 0 & \varepsilon_{s_c} & 0 \\ \hline 0 & \dots & 0 & 0 \end{array} \right] \text{ où } s_c = \text{rang}(\mathbf{C}_s).$$

Les matrices \mathbf{U} et \mathbf{V} , dont les colonnes sont notées \mathbf{u}_i et \mathbf{v}_i , sont orthogonales¹ de dimensions respectives $s_m \times s_m$ et $n \times n$. Les valeurs singulières non nulles ε_i s'expriment traditionnellement sous forme décroissante : $\varepsilon_{i+1} < \varepsilon_i$, $i \in \{1, \dots, s_c - 1\}$ et correspondent à la racine carrée des valeurs propres de la matrice $\mathbf{C}_s^T \mathbf{C}_s$: $\varepsilon_i = \sqrt{\lambda_i(\mathbf{C}_s^T \mathbf{C}_s)} = \sqrt{\lambda_i(\mathbf{C}_s \mathbf{C}_s^T)}$.

La matrice \mathbf{C}_s peut donc s'expliciter sous la forme d'une somme :

$$\mathbf{C}_s = \sum_{i=1}^{s_c} \varepsilon_i \mathbf{u}_i \mathbf{v}_i^T,$$

ce qui permet de déduire les équivalences suivantes :

$$\mathbf{w} \in \ker(\mathbf{C}_s) \Leftrightarrow \mathbf{C}_s \mathbf{w} = 0 \Leftrightarrow \sum_{i=1}^{s_c} \varepsilon_i \mathbf{u}_i (\mathbf{v}_i^T \mathbf{w}) = 0 \Leftrightarrow \mathbf{v}_i^T \mathbf{w} = 0, i \in \{1, \dots, s_c\}, \text{ (Lascaux et al, 1986),}$$

où l'opérateur $\ker()$ désigne le noyau de son argument. En effet, $\mathbf{v}_i^T \mathbf{w} = 0$, $i \in \{1, \dots, s_c\}$ implique nécessairement $\mathbf{C}_s \mathbf{w} = 0$. L'implication inverse est aussi immédiate puisque $\mathbf{v}_i^T \mathbf{w}$ n'est qu'un scalaire et que \mathbf{u}_i est un vecteur non-nul de manière à ce que la matrice \mathbf{U} soit inversible.

Puisque la matrice \mathbf{V} est orthogonale, le produit scalaire de deux colonnes distinctes est nul :

$$\mathbf{v}_i^T \mathbf{v}_j = \delta_{ij} \text{ avec } \delta_{ij} = \begin{cases} 1 & \text{si } i = j \\ 0 & \text{si } i \neq j \end{cases}$$

tous les vecteurs \mathbf{v}_i , $i \in \{s_c + 1, \dots, n\}$, appartiennent donc au noyau de \mathbf{C}_s . Comme ces vecteurs sont linéairement indépendants (\mathbf{V} est inversible) et que la dimension de $\ker(\mathbf{C}_s)$ est donnée par $n - \dim(\text{im}(\mathbf{C}_s)) = n - s_c$ ², le vecteur \mathbf{w} appartient nécessairement à l'espace engendré par les vecteurs $\mathbf{v}_{s_c+1}, \dots, \mathbf{v}_n$.

¹ Une matrice \mathbf{U} orthogonale est une matrice carrée dont l'inverse est sa transposée : $\mathbf{U}\mathbf{U}^T = \mathbf{U}^T\mathbf{U} = \mathbf{I}$.

² L'opérateur $\text{im}()$ désigne l'image de son argument.

En adoptant le même principe, il est possible de démontrer que :

$$\mathbf{w}^T \mathbf{C}_s = 0 \Leftrightarrow \sum_{i=1}^{s_c} \varepsilon_i (\mathbf{w}^T \mathbf{u}_i) \mathbf{v}_i^T = 0 \Leftrightarrow \mathbf{w}^T \mathbf{u}_i = 0, \quad i \in \{1, \dots, s_c\}.$$

Le vecteur \mathbf{w}^T appartient donc à l'espace engendré par les $s_w = s_m - s_c$ vecteurs ligne \mathbf{u}_i^T , $i \in \{s_c + 1, \dots, s_m\}$ et se retrouve donc être une combinaison linéaire de ces mêmes vecteurs.

En conclusion, d'après le résultat précédent, une matrice de parité \mathbf{W} de plein rang ligne orthogonale à la matrice \mathbf{C}_s peut être définie comme suit :

$$\mathbf{W} = \mathbf{K} \begin{bmatrix} 0 & \mathbf{I}_{s_w} \end{bmatrix} \mathbf{U}^T,$$

où \mathbf{K} est une matrice carrée arbitraire et inversible d'ordre s_w .

Annexe E : Relation d'équivalence entre les deux formes statiques dans le cas d'un modèle certain

Au cours du troisième chapitre, l'équivalence structurelle entre les résidus obtenus pour les deux formes statiques (III-8) et (III-11) dans le cas d'une même taille de fenêtre temporelle a été étudiée dans le contexte incertain. L'objectif de cette annexe est de confirmer son existence lorsque le système d'état est cette fois-ci certain. La première forme statique a été définie en (III-2) :

$$Y(k,s) = C_s x(k) + H_s U(k,s-1), \quad U \in \mathbb{R}^{sq}, \quad Y \in \mathbb{R}^{s_m} \text{ avec } s_m = (s+1)m,$$

$$\text{où } C_s = \begin{bmatrix} C \\ CA \\ CA^2 \\ \vdots \\ CA^s \end{bmatrix}, \quad H_s = \begin{bmatrix} 0 & 0 & \cdots & 0 \\ CB & 0 & \cdots & 0 \\ CAB & CB & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ CA^{s-1}B & CA^{s-2}B & \cdots & CB \end{bmatrix}, \quad \begin{matrix} C_s \in \mathbb{R}^{s_m \times n} \\ H_s \in \mathbb{R}^{s_m \times sq} \end{matrix}.$$

La seconde forme statique, quant à elle, se déduit directement de celle obtenue de manière plus générale dans le contexte incertain (III-11) en imposant que les incertitudes sont nulles :

$$\tilde{C}_s X(k,s) = \tilde{H}_s \begin{bmatrix} U(k,s-1) \\ Y(k,s) \end{bmatrix}, \quad X \in \mathbb{R}^{s_n}, \quad Y \in \mathbb{R}^{s_m}, \quad U \in \mathbb{R}^{sq} \text{ avec } s_n = (s+1)n,$$

$$\text{où } \tilde{C}_s = \begin{bmatrix} L_s \\ M_s \end{bmatrix}, \quad \tilde{H}_s = \begin{bmatrix} 0 & I_{s_m} \\ N_s & 0 \end{bmatrix}, \quad \tilde{C}_s \in \mathbb{R}^{\{s_m+sn\} \times s_n}, \quad \tilde{H}_s \in \mathbb{R}^{\{sn+s_m\} \times \{sq+s_m\}},$$

$$\text{avec } M_s = \begin{bmatrix} A & -I_n & \cdots & 0 & 0 \\ \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & \cdots & A & -I_n \end{bmatrix} \text{ et } N_s = -\text{diag}_s(B).$$

Ce développement se déduit en grande partie de celui présenté dans la section III.5.1.3.

- Étape 1 : Relation entre les matrices de parité W et \tilde{W}

Cette étape de la démonstration est identique à celle présentée dans la section III.5.1.3 puisque les matrices étudiées C_s et \tilde{C}_s sont certaines. En supposant donc que celles-ci induisent bien un découplage total pour leur forme statique respective :

$$WC_s = 0 \text{ et } \tilde{W}\tilde{C}_s = 0 \text{ avec } \tilde{W} = \begin{bmatrix} \tilde{W}_L \\ \tilde{W}_M \end{bmatrix},$$

nous avons alors déjà montré qu'il existe une matrice carrée inversible telle que :

$$\tilde{W}_L = TW.$$

- *Étape 2 : Relation entre les résidus obtenus pour les deux formes statiques*

Recherchons maintenant une relation entre les formes d'évaluation puis de calcul associées aux deux formes statiques. Dans le cas certain, les formes d'évaluation se réduisent toutes deux à 0, ce qui permet de conclure immédiatement. Par conséquent, seules les formes de calcul $\mathbf{p}_{cal}(k+s)$ et $\tilde{\mathbf{p}}_{cal}(k+s)$ doivent être étudiées.

$$\mathbf{p}_{cal}(k) = \mathbf{w}^T (\mathbf{Y}(k,s) - \mathbf{H}_s \mathbf{U}(k,s-1)), \quad \tilde{\mathbf{p}}_{cal}(k) = \tilde{\mathbf{w}}^T \tilde{\mathbf{H}}_s \begin{bmatrix} \mathbf{U}(k,s-1) \\ \mathbf{Y}(k,s) \end{bmatrix}.$$

En reprenant la démonstration de la section III.5.1.3 puis en remplaçant les matrices nominales $\tilde{\mathbf{H}}_{s,0}$ et $\tilde{\mathbf{N}}_{s,0}$ du modèle incertain par les celles certaines $\tilde{\mathbf{H}}_s$ et $\tilde{\mathbf{N}}_s$, il se déduit que :

$$\tilde{\mathbf{p}}_{cal}(k+s) = \tilde{\mathbf{W}}^T \tilde{\mathbf{H}}_s \begin{bmatrix} \mathbf{U}(k,s-1) \\ \mathbf{Y}(k,s) \end{bmatrix} = \tilde{\mathbf{W}}_L (-\mathbf{H}_{s,0} \mathbf{U}(k,s-1) + \mathbf{Y}(k,s)) = \mathbf{T} \mathbf{p}_{cal}(k+s).$$

La matrice \mathbf{T} étant inversible, sous les contraintes d'un découplage total, d'un modèle sans incertitude et d'une même taille de fenêtre temporelle s , les résidus obtenus par les deux formes statiques sont structurellement équivalents.

Annexe F : Rang et valeurs propres d'une matrice incertaine

L'objectif est de déterminer des domaines simples contenant les valeurs propres (ou singulières) d'une matrice incertaine afin de déterminer si pour certaines combinaisons des paramètres incertains, une ou plusieurs d'entre elles peuvent s'annuler. Ainsi, une condition suffisante pour qu'une matrice $\mathbf{A}(\boldsymbol{\theta})$ soit de plein rang est qu'aucun de ces domaines ne contienne 0.

- Inégalités de Weyl et de Weinmann

Dans un premier temps, une méthode basée sur l'inégalité de Weyl est présentée. Soit \mathbf{A} une matrice carrée complexe d'ordre n ; $\lambda_i(\mathbf{A})$ désigne sa $i^{\text{ième}}$ valeur propre et \mathbf{v}_i le vecteur propre associé : $\mathbf{A}\mathbf{v}_i = \lambda_i(\mathbf{A})\mathbf{v}_i$.

Rappelons tout d'abord une propriété triviale : soit \mathbf{B} une matrice carrée complexe vérifiant $\mathbf{B} = \alpha\mathbf{A}$, où $\alpha \in \mathbb{R}$, alors $\lambda_i(\mathbf{B}) = \alpha\lambda_i(\mathbf{A})$. Ce résultat reste valide lorsque le coefficient α est remplacé par une variable bornée $\theta \in [-1, 1]$:

$$\mathbf{B} = \theta\mathbf{A} \Rightarrow \lambda_i(\mathbf{B}) = \theta\lambda_i(\mathbf{A}). \quad (\text{VII-4})$$

Considérons maintenant deux matrices certaines hermitiennes³ \mathbf{A} et \mathbf{B} d'ordre n . Les valeurs propres de la somme $\mathbf{A} + \mathbf{B}$ obéissent à l'inégalité de Weyl suivante :

$$\lambda_i(\mathbf{A}) + \lambda_{\min}(\mathbf{B}) \leq \lambda_i(\mathbf{A} + \mathbf{B}) \leq \lambda_i(\mathbf{A}) + \lambda_{\max}(\mathbf{B}),$$

où λ_{\min} et λ_{\max} correspondent aux plus petite et grande valeurs propres de \mathbf{B} . De manière plus générale, dans le cas d'une somme de matrices hermitiennes, on déduit que :

$$\lambda_i(\mathbf{A}) + \sum_j \lambda_{\min}(\mathbf{A}_j) \leq \lambda_i\left(\mathbf{A} + \sum_j \mathbf{A}_j\right) \leq \lambda_i(\mathbf{A}) + \sum_j \lambda_{\max}(\mathbf{A}_j). \quad (\text{VII-5})$$

Dans le cas où les matrices considérées ne sont pas hermitiennes, des résultats issus des travaux de Bendixon et Bromwich permettent là aussi d'obtenir des majorations des valeurs propres de ces matrices et de leur somme. Soit \mathbf{A} une matrice carrée certaine complexe d'ordre n . Cette matrice peut se décomposer sous la forme d'une matrice hermitienne \mathbf{A}_h et d'une matrice non-hermitienne \mathbf{A}_{nh} définies par :

$$\mathbf{A}_h = \frac{1}{2}(\mathbf{A} + \mathbf{A}^*), \quad \mathbf{A}_{nh} = \frac{1}{2}(\mathbf{A} - \mathbf{A}^*),$$

où \mathbf{A}^* représente la transposée du conjugué de \mathbf{A} . Les valeurs propres de la matrice \mathbf{A} appartiennent au rectangle dans le plan complexe défini par :

³ Une matrice \mathbf{A} est hermitienne lorsque $\mathbf{A}^* = \mathbf{A}$ où \mathbf{A}^* est la transposée du conjugué de \mathbf{A} . Dans le cas de matrices réelles, par définition, une matrice hermitienne est une matrice symétrique.

$$\lambda_{\min}(\mathbf{A}_h) \leq \operatorname{Re}(\lambda(\mathbf{A})) \leq \lambda_{\max}(\mathbf{A}_h)$$

$$\min\left(\frac{1}{i}\lambda(\mathbf{A}_{nh})\right) \leq \operatorname{Im}(\lambda(\mathbf{A})) \leq \max\left(\frac{1}{i}\lambda(\mathbf{A}_{nh})\right).$$

Les travaux de Weinmann ont permis de montrer que la partie réelle des valeurs propres de la somme de deux matrices carrées certaines de \mathbb{R}^n est bornée par :

$$\lambda_{\min}(\mathbf{A}_h) + \lambda_{\min}(\mathbf{B}_h) \leq \operatorname{Re}(\lambda(\mathbf{A} + \mathbf{B})) \leq \lambda_{\max}(\mathbf{A}_h) + \lambda_{\max}(\mathbf{B}_h).$$

Là aussi, cette relation peut être étendue au cas d'une somme multiple de matrices :

$$\sum_i \lambda_{\min}(\mathbf{A}_{hi}) \leq \operatorname{Re}\left(\lambda\left(\sum_i \mathbf{A}_i\right)\right) \leq \sum_i \lambda_{\max}(\mathbf{A}_{hi}). \quad (\text{VII-6})$$

Dans le cas où la matrice \mathbf{A} n'est pas carrée, l'étude se fait sur ses valeurs singulières :

$$\sigma_i = \sqrt{\lambda_i(\mathbf{A}\mathbf{A}^*)} = \sqrt{\lambda_i(\mathbf{A}^*\mathbf{A})}.$$

La matrice $\mathbf{A}\mathbf{A}^*$ (tout comme $\mathbf{A}^*\mathbf{A}$) étant hermitienne, certains résultats vus précédemment peuvent être directement réutilisés.

Ces différentes inégalités permettent d'étudier les valeurs propres d'une matrice incertaine. Prenons la par exemple réelle, carrée et de structure classique :

$$\mathbf{A}(\boldsymbol{\theta}) = \mathbf{A}_0 + \sum_{i=1}^r \theta_i \mathbf{A}_i.$$

Si chacune des matrices réelles \mathbf{A}_i , $i \in \{0, \dots, r\}$ est symétrique (donc hermitienne), l'inégalité (VII-5) et la propriété (VII-4) peuvent être exploitées. En supposant que les variables bornées θ_i soient normalisées, la $j^{\text{ième}}$ valeur propre $\lambda_j(\theta_i \mathbf{A}_i)$ est forcément élément de l'intervalle $[-|\lambda_j(\mathbf{A}_i)|, |\lambda_j(\mathbf{A}_i)|]$. Les valeurs propres maximale et minimale de $\theta_i \mathbf{A}_i$ sont donc données par : $\lambda_{\max}(\theta_i \mathbf{A}_i) = -\lambda_{\min}(\theta_i \mathbf{A}_i) = \max_j(|\lambda_j(\mathbf{A}_i)|)$. Des bornes encadrant les valeurs propres de la matrice $\mathbf{A}(\boldsymbol{\theta})$ peuvent donc être données par :

$$\lambda_j(\mathbf{A}_0) - \sum_{i=1}^r \lambda_{\max}(\mathbf{A}_i) \leq \lambda_j(\mathbf{A}(\boldsymbol{\theta})) \leq \lambda_j(\mathbf{A}_0) + \sum_{i=1}^r \lambda_{\max}(\mathbf{A}_i).$$

Dans le cas où la matrice $\mathbf{A}(\boldsymbol{\theta})$ n'est pas hermitienne, le résultat (VII-6) peut être exploité, sachant que cette fois, seule une majoration globale de toutes les valeurs propres à la fois sera possible. Pour chaque matrice \mathbf{A}_i , $i \in \{0, \dots, r\}$, la matrice hermitienne \mathbf{A}_{hi} associée est déterminée ainsi que ses valeurs propres. Puisque les valeurs maximale et minimale des valeurs propres de $\theta_i \mathbf{A}_{hi}$ sont données par $\lambda_{\max}(\theta_i \mathbf{A}_{hi}) = -\lambda_{\min}(\theta_i \mathbf{A}_{hi}) = \max_j(|\lambda_j(\mathbf{A}_{hi})|)$,

l'inégalité suivante est obtenue :

$$\lambda_{\min}(\mathbf{A}_{h0}) - \sum_{i=1}^r \lambda_{\max}(\mathbf{A}_{hi}) \leq \operatorname{Re}\left(\lambda\left(\mathbf{A}_0 + \sum_{i=1}^r \theta_i \mathbf{A}_i\right)\right) \leq \lambda_{\max}(\mathbf{A}_{h0}) + \sum_{i=1}^r \lambda_{\max}(\mathbf{A}_{hi}).$$

Cette technique conduit à définir un seul domaine de majoration pour l'ensemble des valeurs propres de la matrice $A(\theta)$. L'inconvénient est que si cette matrice possède à la fois des valeurs propres positives et négatives, alors l'intervalle précédent contient 0 et il n'est pas possible de savoir si certaines d'entre elles s'annulent pour des valeurs particulières des paramètres incertains.

- Disques de Gerschgorin

La méthode proposée s'appuie sur les intervalles et disques de Gerschgorin. Soit A une matrice carrée complexe et certaine d'ordre n dont les éléments d'indices (i,j) sont notés a_{ij} . Chaque valeur propre de A est située dans un des disques de Gerschgorin de centre $c_i=a_{ii}$ et de rayon r_i défini au choix à partir des coefficients de la $i^{\text{ème}}$ ligne ou $i^{\text{ème}}$ colonne⁴ de A :

$$r_i = \sum_{\substack{j=1 \\ j \neq i}}^n |a_{ij}| \quad \text{ou} \quad r_i = \sum_{\substack{j=1 \\ j \neq i}}^n |a_{ji}|.$$

Toutes les valeurs propres sont donc contenues dans la réunion des différents disques obtenus et définis par $|c_i - \lambda| \leq r_i$, $i \in \{1, \dots, n\}$. Supposons que la réunion \mathcal{U}_k de k de ces disques n'ait aucun point commun avec la réunion \mathcal{V}_{n-k} des $n-k$ disques restants, alors k valeurs propres sont contenues dans \mathcal{U}_k et $n-k$ dans \mathcal{V}_{n-k} . Par conséquent, toute réunion de k disques de Gerschgorin formant un domaine séparé contient exactement k valeurs propres de A . Si tous les disques sont disjoints, alors chaque disque contient nécessairement une valeur propre d'après la propriété précédente. Chaque valeur propre est dans au moins un des disques, mais dans le cas où plusieurs d'entre eux sont sécants, on ne peut plus garantir que chaque disque contienne bien une valeur propre.

Une méthode proche des disques de Gerschgorin, appelée ovales de Cassini, permet elle aussi d'évaluer les valeurs propres d'une matrice certaine complexe A . Elles appartiennent au domaine du plan complexe formé par la réunion des régions intérieures aux $n(n+1)/2$ ovales de Cassini $O(c_i, c_j, r_i, r_j)$ définis par :

$$|c_i - \lambda| |c_j - \lambda| \leq r_i r_j, \quad i \in \{1, \dots, n\}, \quad j \in \{1, \dots, n\}, \quad i \neq j.$$

Les disques de Gerschgorin permettent donc de définir un domaine dans lequel sont nécessairement situées les valeurs propres de A . Dans notre cas, la seule chose qui nous intéresse est de savoir si une valeur propre peut prendre la valeur 0. Il est donc possible de ne s'intéresser qu'à la partie réelle des valeurs propres et dans ce cas ne travailler que sur les intervalles de Gerschgorin définis par $[c_i - r_i, c_i + r_i]$. Chaque partie réelle des valeurs propres de A est située dans un des intervalles de Gerschgorin de centre $c_i=a_{ii}$ et de rayon r_i , et donc dans la réunion de ces intervalles.

Utilisons ceci dans le cadre d'une matrice incertaine définie comme suit :

$$A(\theta) = A_0 + \theta A_1, \quad \theta \in [-1, 1].$$

⁴ Pour pouvoir conclure que toutes les valeurs propres sont à l'intérieur des disques de Gerschgorin, il est nécessaire de travailler avec la même définition (ligne ou colonne) des rayons pour toutes les valeurs propres.

Pour chaque ligne (ou colonne) d'indice i de la matrice $A(\theta)$, le centre et le rayon de l'intervalle de Gerschgorin sont donnés par :

$$c_i(\theta) = a_{0,ii} + \theta a_{1,ii} \text{ et } r_i(\theta) = \sum_{\substack{j=1 \\ j \neq i}}^r |a_{0,ij} + \theta a_{1,ij}|.$$

Pour une valeur particulière de θ , le $i^{\text{ième}}$ intervalle de Gerschgorin sera donc défini par :

$$\left[a_{0,ii} + \theta a_{1,ii} - \sum_{\substack{j=1 \\ j \neq i}}^r |a_{0,ij} + \theta a_{1,ij}|, a_{0,ii} + \theta a_{1,ii} + \sum_{\substack{j=1 \\ j \neq i}}^r |a_{0,ij} + \theta a_{1,ij}| \right].$$

Pour déterminer le domaine obtenu pour θ élément de $[-1,1]$, il suffit de réaliser la réunion des intervalles précédents. La variable θ appartenant à un domaine fermé et borné, cette réunion est donc aussi un intervalle donné par :

$$\left[\inf_{\theta} \left(a_{0,ii} + \theta a_{1,ii} - \sum_{\substack{j=1 \\ j \neq i}}^r |a_{0,ij} + \theta a_{1,ij}| \right), \sup_{\theta} \left(a_{0,ii} + \theta a_{1,ii} + \sum_{\substack{j=1 \\ j \neq i}}^r |a_{0,ij} + \theta a_{1,ij}| \right) \right]. \quad (\text{VII-7})$$

Si aucun de ces n intervalles ne contient 0, alors $A(\theta)$ est de plein rang quelles que soient les incertitudes. Attention cependant, l'évaluation des bornes inférieure et supérieure de ces deux fonctions d'inclusion est soumise au problème de dépendance entre variables bornées (section II.2.4) et peut donc nécessiter la mise en œuvre de techniques proposées dans le chapitre II (formes centrées, procédures d'élimination ou de réduction,...) limitant le pessimisme des intervalles obtenus.

Annexe G : Étude d'un modèle statique incertain

L'objectif de cette section est de présenter brièvement les formes de calcul et d'évaluation proposées dans le chapitre III, mais limitées cette fois à un modèle purement statique. Ce type de modèle est particulièrement adapté à la modélisation d'une chaîne d'instrumentation.

- Structure classique

Seuls sont considérés des modèles statiques faisant intervenir une matrice d'observation incertaine $C(\theta)$ linéaire en le vecteur borné et normalisé $\theta \in \mathbb{R}^r$:

$$y(k) = C(\theta)x(k), \text{ avec } C(\theta) = C_0 + \sum_{i=1}^r \theta_i C_i, \quad x \in \mathbb{R}^n, \quad y \in \mathbb{R}^m. \quad (\text{VII-8})$$

Les vecteurs y et x désignent respectivement les mesures et les grandeurs vraies appelées aussi variables physiques (inconnues et donc à mesurer). Pour utiliser ce type de système en diagnostic et exploiter la redondance matérielle qui en découle, il est nécessaire que plusieurs capteurs mesurent une même variable physique. Dans un contexte plus général, il est possible de prendre en compte certaines contraintes entre les variables à mesurer :

$$Ax(k) = 0, \quad A \in \mathbb{R}^{q \times n}. \quad (\text{VII-9})$$

Ces contraintes sont généralement des bilans énergétiques ou massiques, qui ne souffrent normalement d'aucune imprécision. Néanmoins, il est tout à fait possible d'étendre ce qui va être présenté en ajoutant des incertitudes sur la matrice A et en respectant la même démarche (Adrot et al, 1999a).

Pour prendre en compte ces contraintes, il suffit de construire un système augmenté regroupant les deux équations précédentes :

$$Hy(k) = C(\theta)x(k) \text{ avec } H = \begin{bmatrix} I_m \\ 0 \end{bmatrix}, \quad H \in \mathbb{R}^{(m+q) \times m} \text{ et } C(\theta) = \begin{bmatrix} C(\theta) \\ A \end{bmatrix}, \quad C \in \mathbb{R}^{(m+q) \times n}.$$

D'après la structure de la matrice d'observation, la matrice $\mathcal{C}(\theta)$ peut s'écrire sous la forme :

$$C(\theta) = C_0 + \sum_{i=1}^r \theta_i C_i \text{ avec } C_0 = \begin{bmatrix} C_0 \\ A \end{bmatrix} \text{ et } C_i = \begin{bmatrix} C_i \\ 0 \end{bmatrix}, \quad i \in \{1, \dots, r\}.$$

La recherche d'un vecteur ligne w^T le plus orthogonal possible à la matrice $\mathcal{C}(\theta)$ quelle que soit la valeur des paramètres incertains peut s'opérer en minimisant un critère j_c semblable à (III-16) sous la contrainte $w^T w = 1$. Celui-ci est obtenu en remplaçant la matrice $C_s(\theta)$ par $\mathcal{C}(\theta)$, ce qui revient à rechercher les valeurs propres de la matrice S définie par :

$$S = 2^r \left(C_0 C_0^T + \frac{1}{3} (C_1 C_1^T + \dots + C_r C_r^T) \right). \quad (\text{VII-10})$$

Les formes de calcul et d'évaluation sont respectivement données par :

$$\underbrace{\mathbf{w}^T \mathbf{H} \mathbf{y}(k)}_{p_{cal}(k)} = \underbrace{\mathbf{w}^T \mathbf{C}(\boldsymbol{\theta}) \mathbf{x}(k)}_{p_{eval}(k)}. \quad (\text{VII-11})$$

Un vecteur propre \mathbf{w} associé à une valeur propre nulle de \mathbf{S} entraîne la nullité du critère J_c , ce qui implique un découplage total par rapport aux grandeurs vraies comme expliqué dans la section III.4.1 en remplaçant les matrices $\mathbf{C}_s(\boldsymbol{\theta}_C)$ et $\mathbf{C}_{s,i}$ par $\boldsymbol{\alpha}(\boldsymbol{\theta})$ et \mathbf{C}_i . La forme d'évaluation se réduit alors à zéro.

$$\underbrace{\mathbf{w}^T \mathbf{H} \mathbf{y}(k)}_{p_{cal}(k)} = \underbrace{0}_{p_{eval}(k)}. \quad (\text{VII-12})$$

Malheureusement, contrairement au cas dynamique, le degré de liberté concernant la taille de fenêtre temporelle est perdu ; un découplage total ne sera possible que si la matrice \mathbf{S} définie en (VII-10) (ou de manière équivalente $[\mathbf{C}_0 \ \dots \ \mathbf{C}_r]$) n'est pas de plein rang ligne.

Si un découplage total n'est pas possible, c'est-à-dire que la matrice \mathbf{S} ne possède aucune valeur propre nulle, il est alors nécessaire d'estimer le vecteur $\mathbf{x}(k)$ à partir des mesures lors de l'évaluation de cette relation par l'analyse par intervalles :

$$\hat{\mathbf{x}}_{\boldsymbol{\theta}}(k) = (\mathbf{C}^T(\boldsymbol{\theta})\mathbf{C}(\boldsymbol{\theta}))^{-1} \mathbf{C}^T(\boldsymbol{\theta})\mathbf{y}(k),$$

ce qui conduit à :

$$p_{eval}(k) = \mathbf{w}^T \mathbf{C}(\boldsymbol{\theta}) (\mathbf{C}^T(\boldsymbol{\theta})\mathbf{C}(\boldsymbol{\theta}))^{-1} \mathbf{C}^T(\boldsymbol{\theta})\mathbf{y}(k).$$

L'inconvénient réside dans la condition liée au rang de $\boldsymbol{\alpha}(\boldsymbol{\theta})$ qui doit être de plein rang colonne. Cependant, compte tenu de la prise en compte des contraintes (VII-9) et du fait qu'il est souvent nécessaire d'utiliser de la redondance matérielle pour réaliser un diagnostic d'un système statique ($m > n$), cette contrainte semble moins restrictive que dans le cas d'un système dynamique. Il est vraisemblable que l'expression de la forme d'évaluation devienne alors non-linéaire vis-à-vis des paramètres incertains.

Supposons que les paramètres incertains dépendent du temps ; en remplaçant θ_i par $\theta_i(k)$ dans la relation (VII-8), nous observons que ces derniers n'apparaissent dans la relation (VII-11) qu'au même et unique instant k . Pour un modèle statique, il est donc inutile de faire la distinction entre des incertitudes invariantes ou fluctuant dans le temps, la nature temporelle des incertitudes n'a aucune influence sur les résultats obtenus, ce qui semble naturel pour un tel système (Adrot et al, 1999a).

- Structure de modèle de capteurs

Dans le cas spécifique où des tolérances de capteurs sont modélisées comme dans la section III.7.2, la structure (VII-8) peut être remplacée par :

$$\mathbf{y}(k) = \mathbf{C}(\boldsymbol{\Delta})\mathbf{x}(k) + \mathbf{T}_{\boldsymbol{\delta}}\boldsymbol{\delta} \quad \text{où} \quad \mathbf{C}(\boldsymbol{\Delta}) = (\mathbf{I}_m + \mathbf{T}_{\boldsymbol{\Delta}}\boldsymbol{\Delta})\mathbf{C}_0,$$

où $\boldsymbol{\Delta}$ est une matrice diagonale d'ordre m , dont les éléments diagonaux correspondent à m

variables bornées normalisées et δ est un vecteur borné normalisé de dimension m . Les matrices de pondération T_Δ et T_δ ont la même signification que dans le cadre dynamique (section III.7.2). En tenant compte des équations de contraintes (VII-9) et en regroupant les termes incertains, l'expression suivante est obtenue :

$$Hy(k) = C(\Delta)x(k) + \begin{bmatrix} T_\delta \delta \\ 0 \end{bmatrix}.$$

Un découplage total est ensuite réalisé par rapport à la partie certaine dépendant des grandeurs vraies du système précédent :

$$w^T = \begin{bmatrix} w_C^T & | & w_A^T \end{bmatrix}, \quad w_C \in \mathbb{R}^m, \quad w_A \in \mathbb{R}^q, \quad \text{tel que : } w^T C_0 = \begin{bmatrix} w_C^T & | & w_A^T \end{bmatrix} \begin{bmatrix} C_0 \\ -A \end{bmatrix} = 0.$$

Cela n'est évidemment possible que dans la mesure où la matrice C_0 n'est pas de plein rang ligne, ce qui est logiquement le cas pour des systèmes à redondance matérielle, à laquelle sont en plus ajoutées les contraintes (VII-9). Cette condition semble donc moins restrictive que pour la structure classique précédente et conduit aux expressions suivantes des formes de calcul et d'évaluation :

$$\underbrace{w_C^T y(k)}_{p_{cal}(k)} = \underbrace{w_C^T T_\Delta \Delta C_0 x(k) + w_C^T T_\delta \delta}_{p_{éval}(k)}.$$

Notons que là encore, il n'est plus possible de jouer sur la taille de fenêtre temporelle pour, comme dans le cas d'un système dynamique, être certain que la matrice C_0 ne soit pas de plein rang ligne. Dans le cas d'incertitudes de capteurs, la matrice $I_m + T_\Delta \Delta$ est inversible, ce qui permet d'écrire :

$$C_0 x(k) = (I_m + T_\Delta \Delta)^{-1} (y(k) - T_\delta \delta).$$

La forme d'évaluation se simplifie alors sous la forme :

$$p_{éval}(k) = w_C^T T_\Delta \Delta (I_m + T_\Delta \Delta)^{-1} y(k) + w_C^T (I_m + T_\Delta \Delta)^{-1} T_\delta \delta.$$

Outre le fait que cette représentation soit plus proche du système physique et soit de ce fait plus naturelle ; elle a pour avantage de ne pas faire apparaître le vecteur inconnu $x(k)$. Notons que bien qu'un découplage total ait été réalisé par rapport aux grandeurs vraies, la forme d'évaluation n'est pas nulle comme dans le cas de la structure classique (VII-12), ce qui limite la perte d'information par rapport au modèle.

Annexe H : Génération des contraintes inégalités d'un parallélotope : cas particuliers

L'objectif est la génération des contraintes inégalité représentatives d'un parallélotope dans des cas particuliers autres que celui présenté dans la section IV.5.3. En reprenant les mêmes notations, considérons le champ de vecteurs suivant :

$$\mathbf{p}_{lin,r}(k+s, \mathbf{v}_r) = \mathbf{P}_{v_r} \mathbf{v}_r + \mathbf{p}_0, \quad \|\mathbf{v}_r\|_\infty \leq 1.$$

Une bande de contraintes \mathcal{S}_i est alors définie comme suit :

$$\mathcal{S}_i = \left\{ \mathbf{p}_{lin,r} / -k_i + \mathbf{h}_i^T \mathbf{p}_0 \leq \mathbf{h}_i^T \mathbf{p}_{lin,r} \leq k_i + \mathbf{h}_i^T \mathbf{p}_0 \right\}, \quad (\text{VII-13})$$

le but étant de déterminer le scalaire k_i ainsi que le vecteur ligne \mathbf{h}_i^T .

- *Matrice \mathbf{P}_{v_r} inversible*

Dans le cas où la matrice \mathbf{P}_{v_r} est inversible, la détermination des différentes contraintes inégalité associées au parallélotope $\mathcal{A}(\mathbf{p}_{lin,r})$ peut s'opérer très facilement comme suit :

$$\mathbf{P}_{v_r}^{-1} \mathbf{p}_{lin,r}(k+s, \mathbf{v}_r) = \mathbf{v}_r + \mathbf{P}_{v_r}^{-1} \mathbf{p}_0$$

$$\text{ce qui donne : } \begin{cases} \mathbf{P}_{v_r}^{-1} \mathbf{p}_{lin,r}(k+s, \mathbf{v}_r) \leq \bar{\mathbf{v}}_r + \mathbf{P}_{v_r}^{-1} \mathbf{p}_0 \\ -\mathbf{P}_{v_r}^{-1} \mathbf{p}_{lin,r}(k+s, \mathbf{v}_r) \leq \bar{\mathbf{v}}_r - \mathbf{P}_{v_r}^{-1} \mathbf{p}_0 \end{cases} \text{ avec } \bar{\mathbf{v}}_r = \begin{bmatrix} 1 \\ \vdots \\ 1 \end{bmatrix}.$$

Par analogie avec l'expression (VII-13), en notant s_p la dimension du vecteur résidu et \mathbf{e}_i le $i^{\text{ème}}$ vecteur de la base canonique de \mathbb{R}^{s_p} , nous obtenons pour la bande de contraintes \mathcal{S}_i :

$$k_i = 1 \text{ et } \mathbf{h}_i^T = \mathbf{e}_i^T \mathbf{P}_{v_r}^{-1}, \quad i \in \{1, \dots, s_p\}.$$

- *Matrice \mathbf{P}_{v_r} de plein rang colonne*

Dans le cas où \mathbf{P}_{v_r} est seulement plein rang colonne, réutilisons la démarche précédente :

$$\mathbf{P}_{v_r}^+ \mathbf{p}_{lin,r}(k+s, \mathbf{v}_r) = \mathbf{v}_r + \mathbf{P}_{v_r}^+ \mathbf{p}_0 \text{ avec } \mathbf{P}_{v_r}^+ = \left(\mathbf{P}_{v_r}^T \mathbf{P}_{v_r} \right)^{-1} \mathbf{P}_{v_r}^T,$$

$$\text{ce qui conduit à : } \begin{cases} \mathbf{P}_{v_r}^+ \mathbf{p}_{lin,r}(k+s, \mathbf{v}_r) \leq \bar{\mathbf{v}}_r + \mathbf{P}_{v_r}^+ \mathbf{p}_0 \\ -\mathbf{P}_{v_r}^+ \mathbf{p}_{lin,r}(k+s, \mathbf{v}_r) \leq \bar{\mathbf{v}}_r - \mathbf{P}_{v_r}^+ \mathbf{p}_0 \end{cases}.$$

Par analogie avec l'expression (VII-13), nous obtenons ainsi s_p premières bandes de contrainte \mathcal{S}_i :

$$k_i = 1 \text{ et } \mathbf{h}_i^T = \mathbf{e}_i^T \mathbf{P}_{v_r}^+, \quad i \in \{1, \dots, s_p\}.$$

Soit $\bar{\mathbf{P}}_{v_r}$ le complément de la matrice \mathbf{P}_{v_r} vérifiant : $\bar{\mathbf{P}}_{v_r} = \mathbf{I}_{s_p} - \mathbf{P}_{v_r} \left(\mathbf{P}_{v_r}^T \mathbf{P}_{v_r} \right)^{-1} \mathbf{P}_{v_r}^T$.

En multipliant cette matrice par $\mathbf{p}_{lin,r}$, la relation suivante est obtenue :

$$\bar{\mathbf{P}}_{v_r} \mathbf{p}_{lin,r}(k+s, \mathbf{v}_r) = \bar{\mathbf{P}}_{v_r} \mathbf{p}_0,$$

ce qui impose les contraintes supplémentaires suivantes :

$$\begin{aligned} \bar{\mathbf{P}}_{v_r} \mathbf{p}_{lin,r}(k+s, \mathbf{v}_r) &\leq \bar{\mathbf{P}}_{v_r} \mathbf{p}_0 \\ -\bar{\mathbf{P}}_{v_r} \mathbf{p}_{lin,r}(k+s, \mathbf{v}_r) &\leq -\bar{\mathbf{P}}_{v_r} \mathbf{p}_0 \end{aligned}$$

et donc $\text{rang}(\bar{\mathbf{P}}_{v_r})$ nouvelles bandes de contraintes \mathcal{S}_j définies par :

$$k_j = 0 \text{ et } \mathbf{h}_j^T = \mathbf{e}_j^T \bar{\mathbf{P}}_{v_r}.$$

Il est utile de se limiter aux lignes indépendantes de $\bar{\mathbf{P}}_{v_r}$, sinon des contraintes redondantes sont obtenues (c'est-à-dire plusieurs fois un même hyperplan frontière où un demi-espace englobant le domaine réel, sans être une frontière), comme le montre l'exemple suivant.

Exemple. Considérons le vecteur résidu défini par :

$$\mathbf{p}_0 = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \text{ et } \mathbf{P}_{v_r} = \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix},$$

dont les deux bandes de contraintes sont représentées par les expressions suivantes :

$$\begin{aligned} -2 &\leq [1 \quad -1] \mathbf{p}_{lin,r} \leq 2, \\ -2 &\leq [1 \quad 1] \mathbf{p}_{lin,r} \leq 2. \end{aligned}$$

Deux contraintes redondantes sont construites arbitrairement en réalisant une combinaison linéaire d'une partie des inégalités précédentes :

$$\begin{bmatrix} 3 & 1 \\ 1 & 3 \end{bmatrix} \mathbf{p}_{lin,r} \leq \begin{bmatrix} 6 \\ 6 \end{bmatrix}.$$

Le parallélogramme $\mathcal{A}(\mathbf{p}_{lin,r})$ est représenté sur la figure VII-1 ainsi que les deux contraintes redondantes précédentes, qui tout en contenant $\mathcal{A}(\mathbf{p}_{lin,r})$, ne coïncident avec aucune de ses facettes.

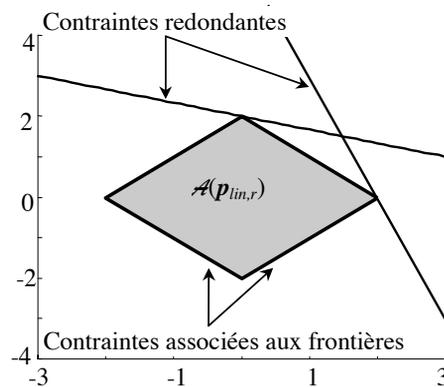


Figure VII-1. Contraintes redondantes

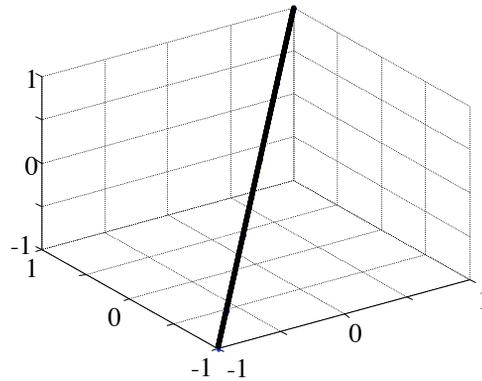
■

Les contraintes supplémentaires sont importantes car elles réduisent la dimension de l'espace abstrait obtenu. De manière générale, un paralléloétope $\mathcal{A}(\mathbf{p}_{lin,r})$ défini par $\mathbf{p}_{lin,r}(k+s, \mathbf{v}_r) = \mathbf{P}_{v_r} \mathbf{v}_r + \mathbf{p}_0$, $\|\mathbf{v}_r\|_\infty \leq 1$ est de dimension :

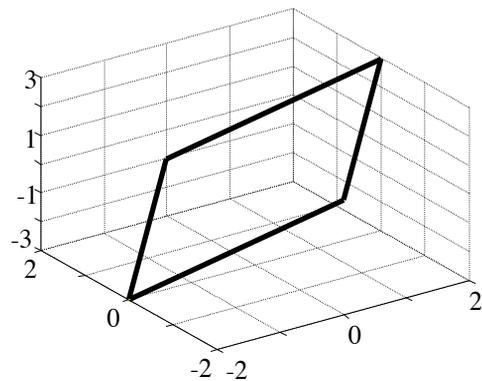
$$\dim(\mathcal{A}(\mathbf{p}_{lin,r})) = \dim(\mathbf{p}_{lin,r}) - \dim(\ker(\mathbf{P}_{v_r})) = \dim(\mathbf{p}_{lin,r}) - \text{rang}(\overline{\mathbf{P}}_{v_r}) \text{ dans } \mathbb{R}^{s_p}.$$

Exemples.

$$\mathbf{p}_0 = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}, \mathbf{P}_{v_r} = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} : \mathcal{A}(\mathbf{p}_{lin,r}) \text{ est une droite dans } \mathbb{R}^3 : \left. \begin{array}{l} \dim(\mathbf{p}_{lin,r}) = 3 \\ \text{rang}(\overline{\mathbf{P}}_{v_r}) = 2 \end{array} \right\}$$



$$\mathbf{p}_0 = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}, \mathbf{P}_{v_r} = \begin{bmatrix} 1 & 1 \\ 1 & -1 \\ 1 & 2 \end{bmatrix} : \mathcal{A}(\mathbf{p}_{lin,r}) \text{ est un plan dans } \mathbb{R}^3 : \left. \begin{array}{l} \dim(\mathbf{p}_{lin,r}) = 3 \\ \text{rang}(\overline{\mathbf{P}}_{v_r}) = 1 \end{array} \right\}$$



Ainsi, une redondance entre les lignes de \mathbf{P}_{v_r} se traduit par une contrainte sur les variables bornées réduisant la dimension de l'espace abstrait $\mathcal{A}(\mathbf{p}_{lin,r})$ par l'intermédiaire des contraintes supplémentaires.

Annexe I : Calcul des sommets d'un polytope

L'objectif de cette annexe est de présenter succinctement la méthode proposée par (Walter et al, 1987), (Mo et al, 1988, 1990) permettant la construction d'un polytope \mathcal{P}_t , c'est-à-dire la détermination de ses sommets, à partir de la connaissance des contraintes inégalité associées :

$$\mathcal{H}(k) = \{ \theta \in \mathbb{R}^r / b(k) \leq \varphi^T(k)\theta \}, k \in \{1, \dots, t\}. \quad (\text{VII-14})$$

La technique proposée est récursive puisque le polytope \mathcal{P}_k est obtenu en faisant l'intersection du domaine \mathcal{P}_{k-1} précédemment déterminé en fonction des $k-1$ premières contraintes inégalité, avec le nouveau demi-espace $\mathcal{H}(k)$ défini par la $k^{\text{ième}}$ inégalité. Cette méthode s'appuie sur la connaissance des sommets adjacents de \mathcal{P}_{k-1} et d'une liste reliant chaque facette du polytope aux sommets qui la constitue. Cette technique permet de construire pas à pas et de manière exacte (aux erreurs d'arrondis près non pris en compte) le polytope \mathcal{P}_t recherché en ajoutant, à chaque itération, une contrainte supplémentaire.

Pour ce faire, toutes les informations utiles concernant \mathcal{P}_{k-1} sont regroupées dans un tableau listant les sommets du polytope où sont précisés, pour chacun d'entre eux, leurs coordonnées, les indices des sommets adjacents et les indices des facettes auxquels ils appartiennent. En supposant que \mathcal{P}_{k-1} possède p sommets, un tableau tel que celui-ci est obtenu :

...	Sommet $S_i(k-1)$ d'indice $i, i \in \{1, \dots, p\}$...
...	Coordonnées $s_i(k-1)$	Liste Sommet - Sommet	Liste Sommet - Facette	...

Tableau VII-4. Tableau des sommets de \mathcal{P}_{k-1}

La construction du polytope se situant dans l'espace des paramètres incertains, les coordonnées $s_i(k-1)$, $i \in \{1, \dots, p\}$ des sommets $S_i(k-1)$ sont éléments de \mathbb{R}^r . Un sommet étant défini par l'intersection de r facettes, chaque liste Sommet-Facette comprendra donc r indices de facettes (sous-espaces de dimension $r-1$). De même, un sommet étant caractérisé par l'intersection de r arrêtes (sous-espaces de dimension 1), il possède par conséquent r sommets adjacents distincts. Chaque liste Sommet-Sommet contient par conséquent r indices de sommets de \mathcal{P}_{k-1} .

Lorsque la nouvelle contrainte inégalité $\mathcal{H}(k)$ est prise en compte, elle conduit à un demi-espace de \mathbb{R}^r dont l'ensemble frontière noté $\mathcal{H}_f(k)$ correspond à l'hyperplan suivant :

$$\mathcal{H}_f(k) = \{ \theta / \varphi^T(k)\theta = b(k) \}.$$

Cet hyperplan peut couper \mathcal{P}_{k-1} , engendrant alors de nouveaux sommets appartenant à \mathcal{P}_k . Inversement, il faut éliminer ceux de \mathcal{P}_{k-1} qui n'appartiennent pas au demi-espace $\mathcal{H}(k)$. Il est ensuite nécessaire de remettre à jour les listes Sommet-Sommet et Sommet-Facette sachant

que dans ce cas, l'hyperplan $\mathcal{H}(k)$ contient la nouvelle facette composant \mathcal{P}_k . Ces différentes étapes sont explicitées dans les paragraphes suivants.

- Détermination des sommets de \mathcal{P}_k

Par définition, le domaine recherché \mathcal{P}_k correspond à l'intersection entre \mathcal{P}_{k-1} et $\mathcal{H}(k)$. Remarquons qu'en choisissant initialement un domaine d'étude \mathcal{P}_0 fermé, \mathcal{P}_k le sera forcément aussi. Dans un premier temps, les sommets de \mathcal{P}_{k-1} qui ont la particularité d'être aussi des sommets de \mathcal{P}_k , sont recherchés ; nous les qualifierons de sommets conservés. Pour cela, ils doivent être éléments de $\mathcal{H}(k)$. Ainsi, pour tous les sommets $S_i(k-1)$, $i \in \{1, \dots, p\}$, de \mathcal{P}_{k-1} et de coordonnées $s_i(k-1)$, le calcul suivant est effectué :

$$\text{signe}_i(k) = b(k) - \boldsymbol{\varphi}^T(k) s_i(k-1). \quad (\text{VII-15})$$

Si la valeur de $\text{signe}_i(k)$ est positive ou nulle, alors le $i^{\text{ième}}$ sommet $S_i(k-1)$ de \mathcal{P}_{k-1} appartient au demi-espace $\mathcal{H}(k)$ et correspond par conséquent à un sommet à conserver. À l'opposé, si $\text{signe}_i(k)$ est strictement négatif, ce sommet est à éliminer car il ne peut faire partie du polytope \mathcal{P}_k .

Il existe alors trois possibilités distinctes. Si toutes les valeurs de $\text{signe}_i(k)$, $i \in \{1, \dots, p\}$, sont positives ou nulles, alors \mathcal{P}_k correspond à \mathcal{P}_{k-1} puisque ce dernier est complètement inclus dans $\mathcal{H}(k)$. Inversement, si tous les $\text{signe}_i(k)$ sont strictement négatifs, alors aucun sommet de \mathcal{P}_{k-1} n'est élément de $\mathcal{H}(k)$ et dans ces conditions, l'ensemble des solutions est vide et l'algorithme peut être stoppé puisque \mathcal{P}_k se réduit à l'ensemble vide \emptyset . Autrement, l'hyperplan $\mathcal{H}(k)$ coupe \mathcal{P}_{k-1} et les sommets à conserver sont déterminés en recherchant les indices i tels que $\text{signe}_i(k)$ soit positif ou nul. Cette intersection engendre aussi de nouveaux sommets qu'il faut maintenant déterminer.

Dans ce dernier cas, cette intersection génère une nouvelle facette $\mathcal{H}_f^*(k)$ (hyperplan obtenu par l'intersection de $\mathcal{H}(k)$ et \mathcal{P}_{k-1}) qui fera nécessairement partie de \mathcal{P}_k et qui est parfaitement définie par ses propres sommets $S_l^*(k)$, $l \in \{1, \dots, s\}$. Ces derniers sont appelés les nouveaux sommets de \mathcal{P}_k puisque ce ne sont pas d'anciens sommets de \mathcal{P}_{k-1} ; notons qu'ils sont forcément situés sur l'une des arêtes (élément de dimension 1) de ce dernier. Ces arêtes sont définies par deux sommets adjacents $S_i(k-1)$ et $S_j(k-1)$ du tableau VII-4 situés respectivement à l'intérieur et à l'extérieur de $\mathcal{H}(k)$:

$$\text{arête}(i, j) = \{ \boldsymbol{\theta} / \boldsymbol{\theta} = (1 - \lambda) s_j(k-1) + \lambda s_i(k-1), 0 \leq \lambda \leq 1 \},$$

$$\text{avec } \text{signe}_i(k) \geq 0 \text{ et } \text{signe}_j(k) < 0.$$

L'utilisation de la contrainte imposant que le nouveau sommet $S_l^*(k)$ appartient à $\mathcal{H}(k)$:

$$\boldsymbol{\varphi}^T(k) s_l^*(k) = b(k),$$

et le fait qu'il se situe sur l'arête précédemment définie :

$$s_l^*(k) = (1 - \lambda^*) s_j(k-1) + \lambda^* s_i(k-1), \quad (\text{VII-16})$$

permettent de déduire que :

$$\lambda^* = \frac{\boldsymbol{\varphi}^T(k) \mathbf{s}_j(k-1) - b(k)}{\boldsymbol{\varphi}^T(k) (\mathbf{s}_j(k-1) - \mathbf{s}_i(k-1))}.$$

En réinjectant la valeur de λ^* dans (VII-16), les coordonnées $s_l^*(k)$ de $S_l^*(k)$ sont calculées. Pour déterminer les s nouveaux sommets, il faut pour chacun des sommets $S_i(k-1)$ de \mathcal{P}_{k-1} tels que $\text{signe}_i(k) \geq 0$, rechercher dans la liste Sommet-Sommet associée les sommets adjacents j tels que $\text{signe}_j(k) < 0$. À chaque fois que cette condition est vérifiée, un nouveau sommet $S_l^*(k)$ de \mathcal{P}_k , défini par (VII-16), est obtenu et listé dans un nouveau tableau de même structure que VII-4.

À ce stade, sont donc connus les nouveaux sommets $S_l^*(k)$ de \mathcal{P}_k , ainsi que ceux de \mathcal{P}_{k-1} devant être conservés. Il ne reste donc plus qu'à mettre à jour les listes Sommet-Sommet et Sommet-Facette.

- *Mise à jour de la liste Sommet-Facette*

En attribuant un indice non-utilisé pour la nouvelle facette, les listes Sommet-Facette des s nouveaux sommets peuvent être en partie complétées puisque $\mathcal{F}_f^*(k)$ est l'une de leurs r facettes. Pour obtenir les indices de celles qui restent, il suffit de retrouver, pour chaque $l^{\text{ième}}$ nouveau sommet $S_l^*(k)$, les $r-1$ facettes communes aux $i^{\text{ième}}$ et $j^{\text{ième}}$ sommets de \mathcal{P}_{k-1} qui ont permis de le définir (VII-16). Pour cela, les deux listes Sommet-Facette correspondantes aux sommets $S_i(k-1)$ et $S_j(k-1)$ (tableau VII-4) sont comparées et les $r-1$ indices de facettes identiques sont inscrits dans la liste Sommet-Facette de $S_l^*(k)$. La liste Sommet-Facette des sommets conservés de \mathcal{P}_{k-1} , quant à elle, ne change pas.

- *Mise à jour de la liste Sommet-Sommet*

Vu la manière dont ont été déterminés les nouveaux sommets, la liste Sommet-Sommet du sommet conservé $S_i(k-1)$ de \mathcal{P}_{k-1} peut facilement être actualisée. L'indice l du nouveau sommet va en effet remplacer l'indice j du sommet éliminé $S_j(k-1)$; inversement i désigne le premier sommet adjacent du nouveau sommet $S_l^*(k)$.

Les sommets conservés sont maintenant complètement remis à jour en fonction de la nouvelle contrainte prise en compte. Il ne reste plus qu'à déterminer les $r-1$ autres sommets adjacents des nouveaux sommets qui font forcément partie des nouveaux sommets. En effet, puisqu'un nouveau sommet $S_l^*(k)$ n'est pas un sommet de \mathcal{P}_{k-1} , il n'appartient qu'à une seule de ses arrêtes, en liaison avec le sommet conservé $S_i(k-1)$. Ses autres sommets adjacents sont donc nécessairement des sommets de la nouvelle facette $\mathcal{F}_f^*(k)$. Il faut donc comparer les listes Sommet-Facette des nouveaux sommets : deux sommets seront adjacents s'ils possèdent $r-1$ facettes communes. Comme ils ont déjà tous la nouvelle facette en commun, il suffit de rechercher $r-2$ indices de facettes identiques différents de celui de $\mathcal{F}_f^*(k)$ dans les listes Sommet-Facette des $S_l^*(k)$.

Une fois la mise à jour des listes effectuée, les sommets non conservés de \mathcal{P}_{k-1} sont éliminés et les deux tableaux de sommets sont regroupés en respectant les changements d'indices de sommets que cela occasionne pour éviter la présence de trous dans la numérotation des sommets et des facettes.

Exemple. Considérons le domaine d'étude \mathcal{D}_0 défini par un cube tel que :

$$\theta_1 \in [0,2], \theta_2 \in [0,2], \theta_3 \in [0,3] \quad (r=3),$$

et supposons que la nouvelle contrainte impose $\theta_1 + 2\theta_2 + 3\theta_3 \leq 5$, ce qui correspond alors à : $b(k) = 5, \varphi^T(k) = [1 \ 2 \ 3], \theta = [\theta_1 \ \theta_2 \ \theta_3]^T$.

Le domaine d'étude \mathcal{D}_0 est défini par le tableau VII-5 et représenté sur la figure VII-2 :

Sommet 1	Sommet 2	Sommet 3	Sommet 4	Sommet 5	Sommet 6	Sommet 7	Sommet 8
0 2 2	2 1 2	2 2 4	0 1 3	0 1 1	2 2 1	2 3 1	0 4 1
0 4 3	0 3 4	2 4 5	2 3 4	0 8 2	0 5 2	2 6 5	2 5 3
3 5 4	3 6 6	3 7 6	3 8 5	0 6 3	0 7 6	0 8 6	0 7 5

Tableau VII-5. Tableau des sommets de \mathcal{D}_0

En appliquant l'algorithme précédent, le tableau suivant est obtenu pour le polytope \mathcal{D}_1 :

Sommet 1	Sommet 2	Sommet 3	Sommet 4	Sommet 5	Sommet 6	Sommet 7	Sommet 8
0 4 1	2 5 1	0 7 1	0 1 7	2 2 7	2 2 7	0 3 7	1 3 7
0 3 2	0 1 2	2 1 3	0 5 2	0 4 2	1.5 8 1	2 4 3	2 6 1
0 2 3	0 6 6	0 8 5	1.7 7 3	1 6 6	0 5 6	0.3 8 5	0 7 5

Tableau VII-6. Tableau des sommets de \mathcal{D}_1

Sur les figures suivantes sont représentés le domaine d'étude \mathcal{D}_0 ainsi que le domaine obtenu \mathcal{D}_1 . Les indices des sommets sont indiqués ainsi que les numéros des facettes qui apparaissent sur fond grisé.

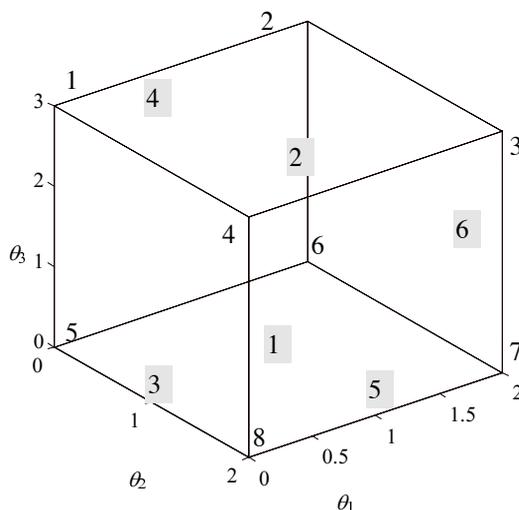


Figure VII-2. Domaine d'étude \mathcal{D}_0

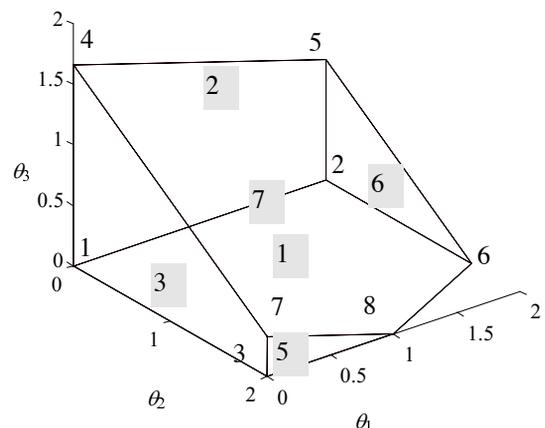


Figure VII-3. Domaine recherché \mathcal{D}_1



- *Sommets dégénérés*

Un sommet dégénéré est un sommet défini par l'intersection de plus de r facettes (et arêtes), auquel cas ce sommet possède plus de r sommets adjacents distincts. Un sommet dégénéré peut apparaître lorsqu'un sommet $S_i(k-1)$ de \mathcal{P}_{k-1} fait partie de l'ensemble frontière $\mathcal{H}_f(k)$, c'est-à-dire lorsque $\text{signe}_i(k)$ est nul. Si aucun sommet n'est à éliminer, alors $\mathcal{H}_f(k)$ coupe \mathcal{P}_{k-1} uniquement en ce point et \mathcal{P}_k est identique au polytope précédent. Si un ou plusieurs sommets doivent être éliminés, alors $S_i(k-1)$ est aussi un sommet de \mathcal{P}_k et il fera partie des sommets conservés. Ce sommet est défini par l'intersection de r facettes de \mathcal{P}_{k-1} , or la nouvelle $\mathcal{H}_f^*(k)$ passe aussi par ce point. Si aucune des r précédentes facettes n'est entièrement en dehors de $\mathcal{H}(k)$, $S_i(k-1)$ aura $r+1$ arêtes et $r+1$ sommets adjacents.

Exemple. Sur les figures suivantes, sont représentés deux cas où l'hyperplan $\mathcal{H}_f(k)$ passe par un sommet de \mathcal{D}_0 . Sur la figure de gauche, ceci se traduit par un sommet dégénéré \mathcal{S} (possédant quatre voisins), alors que sur la figure de droite, \mathcal{S} n'est défini que par l'intersection de trois facettes.

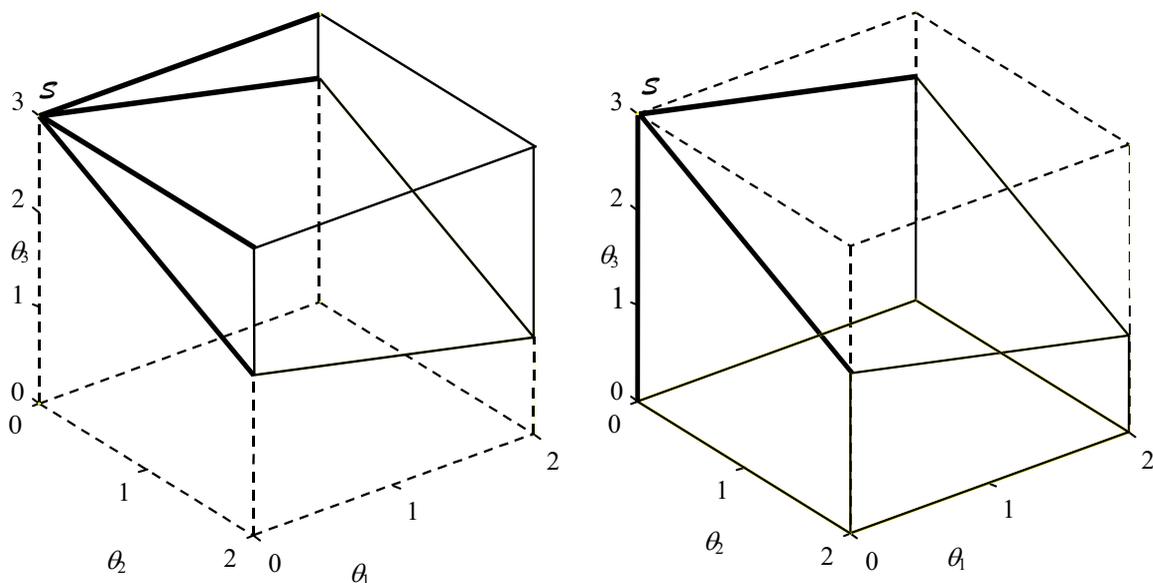


Figure VII-4. Sommet dégénéré

L'algorithme, tel qu'il est construit, ajoute un nouveau sommet à chaque fois que $\mathcal{H}_f(k)$ coupe une arête de \mathcal{P}_{k-1} et qu'il y a élimination de l'un des deux sommets associés. De ce fait, si $\mathcal{H}_f(k)$ passe par $S_i(k-1)$ ($\text{signe}_i(k)=0$) d'une arête et que le second sommet associé $S_j(k-1)$ est en dehors de $\mathcal{H}_f(k)$ ($\text{signe}_j(k)<0$), un nouveau sommet $S_l^*(k)$, confondu avec $S_i(k-1)$, est créé. $S_i(k-1)$ est conservé automatiquement puisqu'il est élément de $\mathcal{H}_f(k)$ et il y aura autant de nouveaux sommets confondus $S_l^*(k)$ avec lui qu'il possède de sommets adjacents de \mathcal{P}_{k-1} non-cohérents avec la nouvelle contrainte. De ce fait, l'algorithme s'adapte à ce type de problème en n'utilisant pas de critère lié à la distance entre deux sommets. Le fait de générer de nouveaux sommets qui coïncident avec $S_i(k-1)$ permet en outre de conserver une taille de liste constante puisque chaque nouveau sommet ne contient dans sa liste Sommet-Sommet que r voisins et dans sa liste Sommet-Facette que r facettes.

Le seul véritable inconvénient de ceci réside dans l'augmentation du nombre de sommets par rapport au cas où tous les voisins d'un sommet dégénéré seraient répertoriés dans une seule et même liste. Cependant, il faut avouer que la probabilité qu'un tel événement se produise reste faible.

Annexe J : Calcul du volume d'un polytope

L'objectif de cette annexe est de présenter de manière succincte la méthode de Lasserre (1983) permettant le calcul du volume d'un polytope. Cet algorithme a été retenu pour la simplicité de sa mise en œuvre et le fait qu'il ne nécessite qu'une \mathcal{H} -représentation du polytope \mathcal{P} considéré.

Soit un d -polytope \mathcal{P} , possédant m facettes⁵ \mathcal{P}_i et défini par la \mathcal{H} -représentation suivante : $\{\boldsymbol{\theta} \in \mathbb{R}^d / \mathbf{H}\boldsymbol{\theta} \leq \mathbf{k}\}$. Par hypothèse, aucune contrainte ne doit apparaître plus d'une fois. Le volume de \mathcal{P} , noté $vol(d, \mathbf{H}, \mathbf{k})$ est déterminé de manière récursive en calculant les volumes associés à chacune de ses facettes. En notant \mathbf{h}_i^T la $i^{\text{ième}}$ ligne de la matrice \mathbf{H} et k_i le $i^{\text{ième}}$ élément du vecteur \mathbf{k} , chaque facette \mathcal{P}_i est contenue dans un hyperplan \mathcal{H}_i de dimension $d-1$ et d'équation $\mathbf{h}_i^T \boldsymbol{\theta} = k_i$, $i \in \{1, \dots, m\}$. Plus exactement, \mathcal{P}_i est un polytope défini par l'intersection de \mathcal{P} et \mathcal{H}_i : $\{\boldsymbol{\theta} \in \mathbb{R}^d / \mathbf{h}_i^T \boldsymbol{\theta} = k_i, \mathbf{H}\boldsymbol{\theta} \leq \mathbf{k}\}$. Nous noterons $vol_i(d-1, \mathbf{H}, \mathbf{k})$ son volume puisque nous nous référons à un domaine de dimension $d-1$.

Pour établir son algorithme, Lasserre J.B. démontre ou rappelle un certain nombre de propositions et théorèmes.

➤ Proposition :

Si $vol(d, \mathbf{H}, \mathbf{k})$ est non-nul, alors $vol(d, \mathbf{H}, \mathbf{k})$ est une fonction continue du vecteur \mathbf{k} .

➤ Proposition :

Supposons que $vol(d, \mathbf{H}, \mathbf{k})$ est non-nul et qu'il en est de même de $vol_i(d-1, \mathbf{H}, \mathbf{k})$ pour toutes les facettes de \mathcal{P}_i , alors $vol(d, \mathbf{H}, \mathbf{k})$ est différentiable en \mathbf{k} et :

$$\frac{\partial vol(d, \mathbf{H}, \mathbf{k})}{\partial k_i} = \frac{vol_i(d-1, \mathbf{H}, \mathbf{k})}{\|\mathbf{h}_i^T\|}.$$

➤ Théorème.

Si $vol(d, \mathbf{H}, \mathbf{k})$ est différentiable en \mathbf{k} , alors à partir de la formule d'Euler pour les fonctions homogènes, le volume de \mathcal{P} est calculé de la manière suivante :

$$vol(d, \mathbf{H}, \mathbf{k}) = \frac{1}{d} \sum_{i=1}^m \frac{k_i}{\|\mathbf{h}_i^T\|} vol_i(d-1, \mathbf{H}, \mathbf{k}) \quad (\text{VII-17})$$

Lasserre J.B. étend ensuite ce résultat dans le cas où $vol(d, \mathbf{H}, \mathbf{k})$ n'est pas différentiable en \mathbf{k} , c'est-à-dire dans le cas où $vol(d, \mathbf{H}, \mathbf{k})$ est nul. Ce cas peut facilement survenir puisque le calcul de ce volume fait appel au calcul de facettes de dimensions $k \in \{d-1, d-2, \dots, 2, 1\}$ dont les volumes $vol_i(k, \mathbf{H}, \mathbf{k})$ peuvent être nuls. Comme l'algorithme est récursif, il est intéressant de s'intéresser à un tel cas.

⁵ Une facette d'un d -polytope est l'une de ses faces de dimensions $d-1$.

➤ Proposition.

Un volume $vol(d, \mathbf{H}, \mathbf{k})$ nul implique que les volumes $vol_i(d-1, \mathbf{H}, \mathbf{k})$ le soient aussi.

➤ Proposition.

Supposons maintenant que $vol(d, \mathbf{H}, \mathbf{k})$ est non nul, mais qu'il existe un indice i tel que $vol_i(d-1, \mathbf{H}, \mathbf{k})$ est nul. Alors, la contrainte i associée est redondante et (VII-17) reste valide.

Cela signifie que l'algorithme est capable de tenir compte de contraintes redondantes générées au fur et à mesure que le polytope est décomposé en facettes. Le polytope de départ ne doit cependant pas avoir de contraintes redondantes initialement.

L'algorithme possède trois étapes distinctes :

- déterminer les éléments caractéristiques \mathbf{h}_i et k_i des m contraintes définissant \mathcal{P} ,
- \mathcal{P}_i étant lui-même un polytope, déterminer les équations permettant de représenter les hyperplans (donc les faces de dimension $dim(\mathcal{P}_i)-1$) délimitant \mathcal{P}_i ,
- réaliser ces opérations m fois pour traiter toutes les facettes \mathcal{P}_i .

Chaque facette est un polytope constitué de facettes de dimensions plus petites. Le calcul du volume de \mathcal{P}_i fait appel à la fonction vol , qui elle-même devra sommer le volume de ses facettes de dimensions $d-2$, et ainsi de suite. L'algorithme est donc récursif puisque la fonction s'appelle elle-même, et ce jusqu'à ce que la facette étudiée soit de dimension 1.

La génération de l'ensemble des contraintes $\mathbf{H}\boldsymbol{\theta} \leq \mathbf{k}$ associées à \mathcal{P} ayant été développée suivant le contexte étudié dans les sections IV.5.3 et V.2.2, nous supposons ici qu'elles sont parfaitement connues. La difficulté est principalement due à la seconde étape. L'objectif est de parvenir à calculer le volume $vol_i(d-1, \mathbf{H}, \mathbf{k})$ de la facette \mathcal{P}_i . Pour cela, il est nécessaire de se placer dans un espace de dimension appropriée $d-1$ (strictement supérieure à 1) pour pouvoir déterminer les contraintes associées à \mathcal{P}_i à partir de celles de \mathcal{P} . La solution proposée est d'effectuer une projection. Par définition, le domaine \mathcal{P}_i correspond à :

$$\mathbf{h}_i^T \boldsymbol{\theta} = k_i \quad (a)$$

$$\mathbf{h}_k^T \boldsymbol{\theta} \leq k_k, \forall k \neq i \quad (b)$$

En se fixant un pivot non-nul h_{ij} ($j^{\text{ième}}$ élément du vecteur ligne \mathbf{h}_i^T), l'objectif est d'éliminer la variable θ_j en utilisant la contrainte égalité précédente (a), c'est-à-dire en posant :

$$\theta_j = \frac{k_i - \sum_{k \neq j} h_{ik} \theta_k}{h_{ij}}.$$

En reportant cette expression de θ_j dans le système d'inégalités (b), le nouveau système $\tilde{\mathbf{H}}\tilde{\boldsymbol{\theta}} \leq \tilde{\mathbf{k}}$ est obtenu : il ne comporte plus que $m-1$ contraintes et le vecteur $\tilde{\boldsymbol{\theta}}$ n'est plus que de dimension $d-1$. Géométriquement, $vol(d-1, \tilde{\mathbf{H}}, \tilde{\mathbf{k}})$ est le volume de $\mathcal{P} \cap \{\mathbf{h}_i^T \boldsymbol{\theta} = k_i\}$ projeté sur le sous-espace $\theta_j=0$.

Lasserre J.B. démontre alors que le volume de \mathcal{P}_i peut s'exprimer très facilement en fonction de $\tilde{\mathbf{H}}$ et $\tilde{\mathbf{k}}$:

$$vol_i(d-1, \mathbf{H}, \mathbf{k}) = \frac{\|\mathbf{h}_i^T\|}{\mathbf{h}_i^T} vol(d-1, \tilde{\mathbf{H}}, \tilde{\mathbf{k}}),$$

ce qui conduit à :

$$vol(d, \mathbf{H}, \mathbf{k}) = \frac{1}{d} \sum_{i=1}^m \frac{k_i}{\mathbf{h}_i^T} vol(d-1, \tilde{\mathbf{H}}, \tilde{\mathbf{k}}).$$

Ce travail récursif est à effectuer sur chaque facette \mathcal{P}_i de \mathcal{P} , ainsi que sur chaque facette de \mathcal{P}_i et ainsi de suite jusqu'à ce que les $\tilde{\theta}$ associés soient de dimension 1. À partir du polytope initial \mathcal{P} , il y a donc pour atteindre le bas de chaque branche de l'arbre récursif $d-1$ ⁶ éliminations à effectuer se traduisant à chaque fois par la perte d'une contrainte. Il restera finalement à la fin de chaque branche $m-d+1$ contraintes dépendant d'une seule variable :

$$\tilde{\mathbf{H}}\tilde{\theta} \leq \tilde{\mathbf{k}}, \quad \tilde{\theta} \in \mathbb{R}, \tilde{\mathbf{H}} \in \mathbb{R}^{m-d+1}, \tilde{\mathbf{k}} \in \mathbb{R}^{m-d+1}.$$

Le volume $vol(1, \tilde{\mathbf{H}}, \tilde{\mathbf{k}})$ est alors calculé en déterminant l'ensemble des demi-droites de \mathbb{R} représentant les valeurs possibles de $\tilde{\theta}$ et en en faisant l'intersection. Prenons le $j^{\text{ième}}$ élément de $\tilde{\mathbf{H}}$ noté \tilde{h}_j , dans ce cas, la $j^{\text{ième}}$ contrainte est donnée par :

$$\tilde{\theta} \leq \frac{\tilde{k}_j}{\tilde{h}_j} \text{ si } \tilde{h}_j > 0, \text{ sinon } \tilde{\theta} \geq \frac{\tilde{k}_j}{\tilde{h}_j} \text{ si } \tilde{h}_j < 0.$$

L'intersection de ces contraintes, si elle existe, donne l'intervalle :

$$\left[\max_{\tilde{h}_j < 0} \left(\frac{\tilde{k}_j}{\tilde{h}_j} \right), \min_{\tilde{h}_j > 0} \left(\frac{\tilde{k}_j}{\tilde{h}_j} \right) \right].$$

Dans le cas où cette intersection n'existe pas, c'est-à-dire si $\max_{\tilde{h}_j < 0} \left(\frac{\tilde{k}_j}{\tilde{h}_j} \right) > \min_{\tilde{h}_j > 0} \left(\frac{\tilde{k}_j}{\tilde{h}_j} \right)$, alors le volume est nul, sinon il correspond à la longueur de l'intervalle précédent.

$$vol(1, \tilde{\mathbf{H}}, \tilde{\mathbf{k}}) = \max \left\{ 0, \min_{\tilde{h}_j > 0} \left(\frac{\tilde{k}_j}{\tilde{h}_j} \right) - \max_{\tilde{h}_j < 0} \left(\frac{\tilde{k}_j}{\tilde{h}_j} \right) \right\}.$$

Notons que s'il n'y a pas de $\tilde{h}_j > 0$ ou de $\tilde{h}_j < 0$, alors le volume est infini, ce qui ne peut se produire dans notre cas puisque seuls des polytopes sont considérés, mais l'algorithme fonctionne de manière générale pour des polyèdres qui eux ne sont pas forcément bornés. C'est cet algorithme qui est mis en œuvre dans la toolbox GBT (Geometric Bounding Toolbox) de Matlab (Véres, 1997).

⁶ Plus précisément, il y a au plus $d-1$ éliminations car il peut exister des faces de volumes nuls auquel cas la branche concernée est éliminée.

Références bibliographiques

ADROT O., 1995

Modélisation de l'ensemble convertisseur-machine synchrone. Extension dans le cadre de la commande vectorielle d'une machine asynchrone.

Mémoire d'ingénieur ESAIGELEC, Saint-Nazaire (France), 1995.

ADROT O., 1996a

Diagnostic appliqué à la machine asynchrone.

Rapport de stage et séminaire interne, DEA Automatique et Traitement Numérique du Signal de l'Institut National Polytechnique de Lorraine, 1996.

ADROT O., 1996b

Estimation paramétrique appliquée à la machine asynchrone.

Rapport bibliographique, DEA Automatique et Traitement Numérique du Signal de l'Institut National Polytechnique de Lorraine, 1996.

ADROT O., MAQUIN D., RAGOT J., 1998

Fault detection with model parameter structured uncertainties.

Final Workshop - Human Capital Mobility, Duisburg (Germany), November 2nd, 1998.

ADROT O., PLOIX S., RAGOT J., 1999a

Détection de défauts appliquée à des modèles statiques affectés par des incertitudes structurées.

17^{ème} Colloque sur le Traitement du Signal et des Images, GRETSI'99, Vannes (France), CD-ROM, 23-25 septembre, 1999.

ADROT O., MAQUIN D., RAGOT J., 1999b

Fault detection with model parameter structured uncertainties.

European Control Conference ECC'99, Karlsruhe (Germany), CD-ROM, August 31-September 3, 1999.

ADROT O., PLOIX S., RAGOT J., 1999c

Une nouvelle approche pour la caractérisation des incertitudes: l'approche bornante.

Chimiométrie 99, De la mesure à l'information, Paris (France), 2-3 décembre, 1999.

ADROT O., MAQUIN D., RAGOT J., 1999d

Estimation d'état généralisé sur horizon glissant. Application au diagnostic.

Troisième Conférence Internationale sur l'Automatisation Industrielle, Montréal (Canada), p 22.8-22.12, 7-9 juin, 1999.

ADROT O., MAQUIN D., RAGOT J., 2000a

Bounding approach to the fault detection of a class of uncertain dynamic systems.

Proceedings of the IFAC Symposium on Fault Detection Supervision and Safety for Technical Process, Budapest (Hungary), p 1168-1173, June 14-16, 2000.

ADROT O., MAQUIN D., RAGOT J., 2000b

Diagnosis of an uncertain static system.

39th IEEE Conference on Decision and Control, CDC'2000, Sydney (Australia), December 12-15, 2000.

ANDERSON B.D.O., MOORE J.B., 1979

Optimal filtering.

Information and System Sciences Series, Prentice Hall, Thomas Kailath, Editor, 1979.

ANDRADE M. V. A., COMBA J.L.D., STOLFI J., 1994

Affine arithmetic.

Interval'94, Petersburg (Russia), March 5-10, 1994.

- ARMENGOL J., TRAVÉ-MASSUYÉS L., VEHI J., SÁINZ M.A., 1999a
Semiqualitative simulation using modal interval analysis.
14th World Congress of International Federation of Automatic Control, IFAC Beijing'99 (China), CD-ROM, July 5-9, 1999.
- ARMENGOL J., TRAVÉ-MASSUYÉS L., VEHI J., DE LA ROSA J.L., 1999b
A survey on interval model simulators and their properties related to fault detection.
14th World Congress of International Federation of Automatic Control, IFAC Beijing'99 (China), CD-ROM, July 5-9, 1999.
- ARMENGOL J., 1999c
Application of modal interval analysis to the simulation of the behaviour of dynamic systems with uncertain parameters.
Thèse de l'Université de Girones, Novembre 1999.
- ARMENGOL J., VEHI J., TRAVÉ-MASSUYÉS L., SÁINZ M.A., 2000
Interval model-based fault detection using multiple sliding time windows.
Proceedings of the IFAC Symposium on Fault Detection Supervision and Safety for Technical Process, Budapest (Hungary), p 168-173, June 14-16, 2000.
- ATKINSON D.J., ACARNLEY P.P., FINCH J.W., 1989
Parameter identification techniques for induction motor drives.
3rd European Conference on Power Electronics and Applications, Aachen (Germany), Vol. 1, p 307-312, October 9-12, 1989.
- BASSEVILLE M., NIKIFOROV I.V., 1993
Detection of abrupt changes - Theory and application.
Prentice Hall, Englewood Cliffs, Information and system sciences series, New Jersey, 1993.
- BEARD R.V., 1971
Failure accommodation in linear systems through self-reorganisation.
Ph.D.Thesis, Department of Aeronautics and Astronautics, Massachusetts Institute of Technology, Cambridge, Massachusetts, February 1971.
- BECKER T., WEISPFENNING V.
Gröbner Bases, a Computational Approach to Commutative Algebra
Graduate Texts in Mathematics, Springer Verlag New York Inc., 1993
- BOUKHRIS A., 1998
Identification des systèmes dynamiques non-linéaires par une approche multi-modèles – Application à la modélisation de la relation pluie débit pour le diagnostic de fonctionnement de capteurs.
Doctorat de l'Institut National Polytechnique de Lorraine, Spécialité automatique, 13 juillet, 1998.
- BRDYS M.A., 1999
Robust estimation of variables and parameters in dynamic networks.
14th World Congress of International Federation of Automatic Control, IFAC Beijing'99 (China), CD-ROM, July 5-9, 1999.
- BROMAN V., SHENSA M. J., 1988
A compact algorithm for the intersection and approximation of N-Dimensional polytopes.
12th IMACS World Congress, Paris, Vol. 2, p 473-476, July 18-22, 1988.

- BRUNET M., JAUME D., LABARRÈRE M., RAULT A., VERGÉ M., 1990
Détection et diagnostic de pannes - Approche par modélisation.
Traité des Nouvelles Technologies, série Diagnostic et Maintenance, Hermès, Paris, 1990.
- BÜELER B., ENGE A., FUKUDA K., 1998
Exact volume computation for polytopes: a practical study.
<http://www.ifor.math.ethz.ch/~bueeler/bueeler.html>, 1998.
- CASSAR J.P., STAROSWIECKI M., COCQUEMPOT V., 1995
Optimal residual design for model-based fault detection and isolation.
Proceedings of 3rd European Control Conference, Rome (Italy), p 1986-1992, 1995.
- CERONE V., 1993
Parameter bounds for ARMAX models from records with bounded errors in variables.
International Journal of Control, Vol. 57, N° 1, p 225-235, 1993.
- CERONE V., 1996
Errors-in-variables models in parameter bounding.
dans Milanese M., Norton J., Piet-Lahanier H., Walter E., *Bounding approaches to system identification*, Plenum Press, New York and London, Chap 18, p 289-306, 1996.
- CHANG I.C., YU C.C., LIOU C.T., 1995
Model-based approach for fault diagnosis. Part 2. Extension to interval systems
Industrial Engineering Chemical Research, N° 34, p 828-844, 1995.
- CHAR B.W., GEDDES K.O., GONNET G.H., LEONG B.L., MONAGAN M.B., WATT, S.M., 1991
Mapple V, Library reference manual.
Éditions Springer-Verlag, 1991.
- CHEN J., ZHANG H-Y, 1990
Parity vector approach for detecting failures in dynamic systems.
International Journal of Systems Science, Vol. 21, N° 4, p 765-770, 1990.
- CHOW E.Y., 1980
A failure detection system design methodology.
Sc. D Thesis, Department of Electrical Engineering and Computer Science,
Massachusetts Institute of Technology, Cambridge, Massachusetts, October 1980.
- CHOW E.Y., WILLISKY A.S., 1984
Analytical redundancy and the design of robust failure detection system.
IEEE Transactions on Automatic Control, Vol. AC-29, N° 7, p 603-614, 1984.
- CLARK R.N., 1989
State estimation schemes for instrument fault detection.
Chapter 2, in : *Fault diagnosis in dynamic systems: theory and application*, Patton R.J.,
Frank P.M., Clark R.N., Prentice Hall, Englewood Cliffs, International Series in
Systems and Control Engineering, New Jersey, 1989.
- CLÉMENT T., GENTIL S., 1988
Recursive membership set estimation for ARMAX models: an output-error approach.
12th IMACS World Congress, Paris (France), Vol. 2, p 484-486, July 18-22, 1988.
- COCQUEMPOT V., CASSAR J.P., STAROSWIECKI M., 1997
How does the time window size influence the sensitivity/robustness trade-off of optimal structured residuals ?
Proceedings of the IFAC Symposium on Fault Detection Supervision and Safety for
Technical Process, Hull (United Kingdom), Vol. 1, p 329-334, August 26-28, 1997.

- COHEN J., HICKEY T., 1979
Two algorithms for determining volumes of convex polyhedra.
Journal of the ACM, Vol. 26, N° 3, p 401-414, 1979.
- COIRAULT P., GABANO J-D., TRIGEASSOU J-C., 1993
Maintenance prédictive d'un entraînement électrique par identification paramétrique.
Diagnostic et sûreté de fonctionnement, Vol. 3, N° 1, p 69-95, 1993.
- COMBA J. L. D., STOLFI, J., 1993
Affine arithmetic and its applications to computer graphics.
SIBGRAP'93, Recife (Brazil), October 20-22, 1993.
- CORNIL J.-M., TESTUD P., 1995
MappleTM, introduction raisonnée à l'usage de l'étudiant, de l'ingénieur et du chercheur.
Éditions Springer-Verlag, Berlin Heidelberg New York, 1995.
- D'ANDREA R., KHATRI S., 1997
Kalman decomposition of linear fractional transformation representations and minimality.
Proceedings of American Control Conference, p 3557-3561, 1997.
- DALY K.C., GAI E., HARRISON J.V., 1979
Generalised likelihood test for FDI in redundant sensor configuration.
Journal of Guidance and Control, Vol. 2, N° 1, p 9-17, 1979.
- DARDINIER MARON V., NOURA H., HAMELIN F., 1999
Loi de commande tolérante aux défauts majeurs d'actionneurs.
Journées Doctorales d'Automatique, JDA'99, p 265-268, Nancy (France), 21-23 septembre, 1999.
- DE FIGUEUREDO L.H., 1996a
Surface intersection using affine arithmetic.
Proceedings of Graphics Interface'96, p 168-175, 1996.
- DE FIGUEUREDO L.H., STOLFI J., 1996b
Adaptive enumeration of implicit surfaces with affine arithmetic.
Computer Graphic Forum, Vol. 15, N° 5, p 287-296, December 1996.
- DECKERT J.C., DESAI M.N., DEYST J.J., WILLSKY A.S., 1976
A reliable dual-redundant sensor FDI system for the NASA F8C-DFBW aircraft.
IEEE Conference on Decision and Control, Clearwater Beach (Florida, USA), Vol. 1, p 22-37, 1976.
- DECKERT J.C., DESAI M.N., DEYST J.J., WILLSKY A.S., 1977
F-8 DFBW sensor failure identification using analytic redundancy.
IEEE Transactions on Automatic Control, Vol. 22, N° 5, p 795-803, 1977.
- DESAI M.N.; DECKERT J.C., DEYST J.J., WILLSKY A.S., CHOW E.Y, 1976
Dual-redundant sensor FDI techniques applied to the NASA F8C-DFBW aircraft.
AIAA Guidance and Control Conference, San Diego (California, USA), Vol. 1, p 502-513, 1976.
- DESAI M.N., DECKERT J.C., DEYST J.J., 1979
Dual sensor failure identification using analytic redundancy.
Journal of Guidance and Control, Vol. 2, N° 3, p 213-220, 1979.

DIDRIT O., 1997

Analyse par intervalles pour l'automatique ; résolution globale et garantie de problèmes non-linéaires en robotique et en commande robuste.

Doctorat de l'Université Paris XI Orsay, 30 juin, 1997.

DORR R., 1995

Détection de défauts de capteurs par redondances matérielle et analytique. Application aux centrales nucléaires.

Doctorat de l'Institut National Polytechnique de Lorraine, Spécialité automatique, 18 juillet 1995.

DUBUISSON B., 1990

Diagnostic et reconnaissance de formes.

Traité des Nouvelles Technologies, Série Diagnostic et Maintenance, Hermès, 1990.

EMANI-NAEINI A.E., AKHTER M.M., ROCK S.M., 1988

Effect of model uncertainty on failure detection : the threshold selector.

IEEE Transactions on Automatic Control, Vol. 33, N° 2, p 1106-1115, 1988.

FENG X., CORLISS G.F., KELNHOFER R.W., 1999

A new interval bounding algorithm for parameter estimation from bounded-error data.

14th World Congress of International Federation of Automatic Control, IFAC Beijing'99 (China), CD-ROM, July 5-9, 1999.

FOGEL E., HUANG Y.F., 1982

On the value of information in system identification-bounded noise case.

Automatica, Vol. 18, N° 2, p 229-238, 1982.

FRANK P.M., 1987

Fault diagnosis in dynamic systems via state estimation - A survey.

In Tzafestas S., Singh M., Schmidt G. (Eds), System fault diagnosis, reliability and related knowledge-based approaches, Vol. 1, Reidel Dordrecht Press, p 35-98, 1987.

FRANK P.M., WÜNNENBERG J., 1989

Robust fault diagnosis using unknown input observer schemes.

Chapter 3, in : *Fault diagnosis in dynamic systems: theory and application*, Patton R.J., Frank P.M., Clark R.N., Prentice Hall, Englewood Cliffs, International Series in Systems and Control Engineering, New Jersey, 1989.

FRANK P.M., 1990

Fault diagnosis in dynamical systems using analytical and knowledge based redundancy - a survey and some new results.

Automatica, Vol. 26, N° 3, p 459-474, 1990.

FRANK P.M., KIUPEL N., 1993

Fuzzy supervision and application to lean production.

International Journal of Systems Science, Vol. 24, N° 10, p 1935-1944, 1993.

FRANK P.M., 1996

Analytical and qualitative model-based fault diagnosis – A survey.

European Journal of Control, Vol. 2, p 6-28, 1996.

GASSO K., 2000

Identification des systèmes dynamiques non-linéaires : approche multi-modèle

Doctorat de l'Institut National Polytechnique de Lorraine, Spécialité automatique, 22 décembre 2000.

- GENTIL S., DUBUISSON B., STAROSWIECKI M., CASSAR J.-P., DENOEU X T., 1996
Continuous system supervision.
École d'été d'automatique de Grenoble, 2-6 septembre, 1996.
- GERTLER J., 1986
Failure detection and isolation in complex process plants. A survey.
IFAC Symposium on Microprocessor Application and Process Control, Istanbul (Turkey), p 13-25, 1986.
- GERTLER J., 1988
Survey of model-based failure detection and isolation in complex plants.
IEEE, Control Systems Magazine, Vol. 8, N° 6, p 3-11, 1988.
- GERTLER J., SINGER D., 1990
A new structural framework for parity equations-based failure detection and isolation.
Automatica, Vol. 26, N° 2, p 381-388, 1990.
- GERTLER J., 1991
Analytical redundancy methods in fault detection and isolation.
Proceedings of the IFAC Symposium on Fault Detection Supervision and Safety for Technical Process, Baden Baden (Germany), p 9-22, 10-13 September, 1991.
- GIULIANI S., 1998
Détection et localisation de défaillances de capteurs dans un réseau de mesure pluviométrique - Cas de l'agglomération nancéienne.
Doctorat de l'Institut National Polytechnique de Lorraine, Spécialité automatique, 8 juillet 1998.
- GOLUB G.H., MEURANT G.A., 1983
Résolution numérique des grands systèmes linéaires.
Éditions Eyrolles, Collection de la direction des études et recherches d'électricité de France, 1983.
- GOLUB G.H., VAN LOAN C.F., 1989
Matrix computations.
Johns Hopkins University Press, Baltimore and London, 1989.
- HADJ-SADOK M.Z., GOUZÉ J.L., 1998
Bounds estimations for uncertain models of wastewater treatment.
Proceedings of the 38th IEEE International Conference on Control Applications, Trieste (Italy), September 1-4, 1998
- HADJ-SADOK M.Z., GOUZÉ J.L., 1999
Comparison between two interval observers for wastewater treatment models.
European Control Conference ECC'99, Karlsruhe (Germany), CD-ROM, August 31-September 3, 1999.
- HALL S.R., 1981
Parity vector compensation for FDI.
S.M. Thesis, Department of Aeronautics and Astronomic, Massachusetts Institute of Technology, Cambridge, (Massachusetts, USA), 1981.

HAMMAD M., 1986

Validation des mesures et détection des capteurs défaillants dans un système de contrôle-commande.

Doctorat de l'Université des Sciences et Techniques de Lille Flandres Artois, Spécialité Automatique, 17 juin 1986.

HANSEN E. R., 1992

Global optimization using interval analysis.

Marcel Dekker, New York, 1992.

HÖFLING T., PFEUFER T., 1994

Detection of additive and multiplicative faults, parity space vs. parameter estimation.

Proceedings of the IFAC Symposium on Fault Detection Supervision and Safety for Technical Process, Helsinki (Finland), Vol. 2, p 539-544, 1994.

HÖFLING T., ISERMANN R., 1996a

Fault detection based on adaptative parity equations and single-parameter tracking.

Control Engineering Practice, Vol. 4, N° 10, p 1361-1369, 1996.

HÖFLING T., ISERMANN R., 1996b

Parameter estimation triggered by continuous time parity equations.

Proceedings of the American Control Conference, Seattle, Washington, p 1145-1156, June 1996.

ISERMANN R., 1984

Process fault detection based on modeling and estimation methods - A survey.

Automatica, Vol. 20, p 387-404, 1984.

ISERMANN R., 1994

Integration of fault detection and diagnosis methods.

Proceedings of the IFAC Symposium on Fault Detection Supervision and Safety for Technical Process, Helsinki (Finland), p 597-612, June 13-16, 1994.

ISERMANN R., BALLÉ P., 1997

Trends in the application of model-based fault detection and diagnosis of technical processes.

Control Engineering Practice, Vol. 5, N° 5, p 709-719, 1997.

ISERMANN R., BALLÉ P., 2000

Applied terminology of fault detection, supervision and safety for technical processes.

Site internet de IFAC Symposium on Fault Detection Supervision and Safety for Technical Process 2000 : <http://www.sztaki.hu/conferences/safeprocess2000/termfr.html>.

JANATI IDRISSE H., ADROT O., RAGOT J., 2000

Recherche du vecteur parité pour un système statique incertain.

Conférence Internationale Francophone d'Automatique CIFA2000, Lille (France), 5-8 juillet 2000.

JAULIN L., WALTER E., 1993

Guaranteed nonlinear parameter estimation from bounded-error data via interval analysis.

Mathematics and Computers in Simulation, N° 35, p 123-137, 1993.

- JAULIN L., WALTER E., 1996
Guaranteed nonlinear set estimation via interval analysis.
dans Milanese M., Norton J., Piet-Lahanier H., Walter E., *Bounding approaches to system identification*, Plenum Press, New York and London, Chap 23, p 363-381, 1996.
- JAULIN L., 2000
Le calcul ensembliste par analyse par intervalles et ses applications.
Habilitation à Diriger des Recherches, <http://www.istia.univ-angers.fr/~jaulin/publications.html>, 2000
- JIN H., ZHANG H.Y., 1997
Configuration of redundant sensor system and its fault detection using parity vector method.
Proceedings of the IFAC Symposium on Fault Detection Supervision and Safety for Technical Process, p 843-848, 1997.
- JONES H.L., 1973
Failure detection in linear systems.
Ph.D. Thesis, Department. of Aeronautics and Astronautics, Massachusetts Institute of Technology, Cambridge, Massachusetts, August 1973.
- KATSILLIS G., CHANTLER M., 1999
Comparing two methods of diagnosis of imprecise dynamic systems.
European Control Conference ECC'99, Karlsruhe (Germany), CD-ROM, August 31-September 3, 1999.
- KÉRIGNARD D., COIRAULT P., TRIGEASSOU J.-C., 1995
Estimation des paramètres électriques d'une machine asynchrone.
2nd International Conference on Industrial Automation, Vol. 2, p 577-582, 1995.
- KIEFFER M., JAULIN L., WALTER E., 1998
Guaranteed Recursive Nonlinear State Estimation Using Interval Analysis.
37th IEEE Conference on Decision and Control, Tampa, Florida, USA, December 16-18, 1998.
- KIEFFER M., JAULIN L., WALTER E., MEIZEL D., 1999
Localisation et suivi garantis d'un robot par analyse par intervalles.
Journées Doctorales d'Automatique, JDA'99, p 225-228, Nancy (France), 21-23 septembre, 1999.
- LANDESMAN E.M., HESTENES M.R., 1992
Linear algebra for mathematics, science and engineering.
Prentice Hall, Englewood Cliffs, New Jersey, 1992.
- LASCAUX P., THEODOR R., 1986
Analyse numérique matricielle appliquée à l'art de l'ingénieur.
Éditions Masson, Tomes 1 et 2, 1986
- LASSERRE J.B., 1983
An analytical expression and an algorithm for the volume of a convex polyhedron in \mathbb{R}^n .
Journal of Optimization Theory and Applications, Vol. 39, N° 3, p 363-377, 1983.
- LAWRENCE J., 1991
Polytope volume computation.
Mathematics of Computation, Vol. 57, N° 195, p 259-271, 1991.

LE NOUVEAU PETIT ROBERT, 1994

Dictionnaire alphabétique et analogique de la langue française.
Dictionnaire Le Robert, Paris, 1994.

LORON L., LALIBERTÉ G., 1993

Application of the extended Kalman Filter to parameters estimation of induction motors.
Power Electronics and Applications EPE'93, Brighton (UK), p 85-90, 1993.

LOU X.C., 1982

A system failure detection method: the failure projection method.
S.M. Thesis, Department of Electrical Engineering and Computer Science, Massachusetts Institute of Technology, Cambridge, Massachusetts, June 1982.

LOU X-C., WILLISKY A.S., VERGHESE G.C., 1986

Optimally robust redundancy relations for failure detection in uncertain systems.
Automatica, Vol. 22, N° 3, p 333-344, 1986.

MANDEL D., 1998

Diagnostic à base de redondance analytique. Application à un réseau urbain de distribution d'eau potable.
Doctorat de l'Institut National Polytechnique de Lorraine, Spécialité automatique, 3 décembre 1998.

MAQUIN D., ADROT O., RAGOT J., 1999

Estimation d'état de systèmes incertains. Application à la détection de défauts de capteurs.
Troisième Conférence Internationale sur l'Automatisation Industrielle, Montréal (Canada), p 22.9-22.12, 7-9 juin, 1999.

MAQUIN D., ADROT O., RAGOT J., 2000a

Data reconciliation with uncertain models.
ISA Transactions, No 39, p 35-45, 2000.

MAQUIN D., RAGOT J., 2000b

Diagnostic des systèmes linéaires.
Collection pédagogique d'automatique, Hermès Science Publications, Paris, 2000.

MASSOUMNIA M.A., 1986

A geometric approach for failure detection and identification in linear systems.
Ph. D. Thesis, Massachusetts Institute of Technology, Cambridge, (Massachusetts, USA), 1986.

MASSOUMNIA M.A., VANDER VELDE W.E., 1988

Generating parity relations for detecting and identifying control system component failures.
Journal Guidance, Control and Dynamics, Vol. 11, N° 1, p 60-65, 1988.

MEDVEDEV A., 1994

Parity space method : a continuous time approach.
Proceeding of the American Control Conference, Baltimore, (Mariland, USA), p 662-665, June, 1994.

MILANESE M., BELFORTE G., 1982

Estimation theory and uncertainty intervals evaluation in presence of unknown but bounded errors: linear families of models.

IEEE Transactions on Automatic Control, Vol. AC-27, N° 2, p 408-413, 1982.

MILANESE M., NORTON J., PIET-LAHANIER H., WALTER E., 1996a

Bounding approaches to system identification.

Plenum Press, New York and London, 1996.

MILANESE M., VICINO A., 1996b

Optimal estimation theory for dynamic systems with set membership uncertainty: an overview.

dans Milanese M., Norton J., Piet-Lahanier H., Walter E., *Bounding approaches to system identification*, Plenum Press, New York and London, Chap. 2, p 5-27, 1996.

MILNE R., 1987

Strategies for diagnosis.

IEEE Transactions on Systems, Man and Cybernetics, Vol. 17, N° 3, p 333-339, 1987.

MIRONOVSKI L.A., 1979

Functional diagnosis of linear dynamic systems.

Avtomatika i Telemekhanika, N° 8, p 120-128, 1979.

MO S.H., NORTON J.P., 1988

Recursive parameter-bounding algorithms which compute polytope bounds.

12th IMACS World Congress, Paris (France), Vol. 2, p 477-480, July 18-22, 1988.

MO S.H., NORTON J.P., 1990

Fast and robust algorithm to compute exact polytope parameter bounds.

Mathematics and Computers in Simulation, Vol. 32, p 481-493, 1990.

MOORE R.E., 1966

Interval analysis.

Prentice Hall, Englewood Cliffs, New Jersey, 1966.

MOORE R.E., 1979

Methods and applications of interval analysis.

Studies in Applied Mathematics, Philadelphia, Pennsylvania, 1979.

NELDER J.A., MAID R., 1965

A simplex method for function minimization.

Computer Journal, n° 7, p 308-313, 1965.

NEUMAIER A., 1990

Interval methods for systems of equations.

Encyclopedia of Mathematics and its Applications, Cambridge, 1990.

NUNINGER W., 1997

Stratégie de diagnostic robuste à l'aide de la redondance analytique.

Doctorat de l'Institut National Polytechnique de Lorraine, Spécialité automatique, 29 octobre 1997.

PATTON R.J., FRANK P.M., CLARK R.N., 1989

Fault diagnosis in dynamic systems : theory and application.

Prentice Hall, Englewood Cliffs, International Series in Systems and Control Engineering, New Jersey, 1989.

PATTON R.J., CHEN J., 1991

A review of parity space approaches to fault diagnosis.

Proceedings of the IFAC Symposium on Fault Detection Supervision and Safety for Technical Process, Baden Baden (Germany), p 65-81, September 10-13, 1991.

PATTON R.J., 1994

Robust model-based fault diagnosis : the state of the art.

Proceedings of the IFAC Symposium on Fault Detection Supervision and Safety for Technical Process, Helsinki (Finland), Vol. 1, p 1-24, June 13-16, 1994.

PATTON R.J., 1997

Fault-tolerant control : the 1997 situation.

Proceedings of the IFAC Symposium on Fault Detection Supervision and Safety for Technical Process, Hull (United Kingdom), Vol. 2, p 1033-1055, August 26-28, 1997.

PATTON R.J., FRANK P.M., CLARK R.N., 2000

Issues of fault diagnosis for dynamic systems.

Édition Springer-Verlag London Limited 2000.

PIET-LAHANIER H., WALTER E., 1988

Practical implementation of an exact and recursive algorithm for characterizing likelihood sets.

12th IMACS World Congress, Paris (France), Vol. 2, p 481-483, July 18-22, 1988.

PIET-LAHANIER H., WALTER E., 1990

Exact recursive characterization of feasible parameter sets in the linear case.

Math. and Comput. in Simul., Vol. 32, p 495-504, 1990.

PLOIX S., ADROT O., RAGOT J., KRATZ F., 1996

Diagnostic de défaillance sur une machine asynchrone.

Réunion PRC-GDR Automatique, Diagnostic Appliqué à la Machine Asynchrone, 25 janvier 1996.

PLOIX S., KRATZ F., RAGOT J., ADROT O., 1997a

Approche conceptuelle du diagnostic et préliminaires à la mise en œuvre sur la machine asynchrone.

Colloque de Recherche Doctorale Automatique, Génie informatique, Image et Signal, Le projet Diagnostic de la machine asynchrone - Bilan des Activités PRC-GDR Automatique, Angers (France), p 63-66, 11 décembre, 1997.

PLOIX S., KRATZ F., RAGOT J., ADROT O., 1997b

Diagnostic de la Machine asynchrone à base de Modèles Analytiques.

Colloque de Recherche Doctorale Automatique, Génie informatique, Image et Signal, Le projet Diagnostic de la machine asynchrone - Bilan des Activités PRC-GDR Automatique, Angers (France), p 59-62, 11 décembre, 1997.

PLOIX S., ADROT O., KRATZ F., IUNG C., RAGOT J., 1997c

An extended model of induction machine for failure detection.

IFAC Conference on Control of Industrial Systems, Belfort (France), Vol. 1, p 55-62, May 20-22, 1997.

PLOIX S., 1998

Diagnostic des systèmes incertains : l'approche bornante.

Thèse de l'Institut National Polytechnique de Lorraine, 23 décembre, 1998.

- PLOIX S., ADROT O., KRATZ F., MAQUIN D., RAGOT J., 1999a
Diagnostic des systèmes incertains : l'approche bornante.
Réunion Groupe Thématique Sécurité, Surveillance, Supervision, Thème de la journée :
Approches ensemblistes et applications en diagnostic, 18 mars, 1999.
- PLOIX S., ADROT O., RAGOT J., 1999b
Parameter uncertainty computation in static linear models.
38th IEEE Conference on Decision and Control, CDC'99, Phoenix (USA), December 7-
10, 1999.
- PLOIX S., KRATZ F., RAGOT J., 1999c
Application of LMI concept to parameters adjustment of a PID controller.
38th IEEE Conference on Decision and control, CDC'99, Phoenix (USA), December 7-
10, 1999.
- PLOIX S., ADROT O., RAGOT J., 2000
Bounding approach to the diagnosis of a class of uncertain static systems.
Proceedings of the IFAC Symposium on Fault Detection Supervision and Safety for
Technical Process, Budapest (Hungary), p 149-154, June 14-16, 2000.
- POTTER J.E., SUMAN M.C., 1977
Thresholdless redundancy management with arrays of skewed instruments.
Agardograph 224, Integrity in Electronic flight control systems, p 15.1-15.24, 1977.
- PUIG V., SALUDES J., QUEVEDO J., 1999
*A new algorithm for adaptive threshold generation in robust fault detection based on a
sliding window and global optimization.*
European Control Conference ECC'99, Karlsruhe (Germany), CD-ROM, August 31-
September 3, 1999.
- PUIG V., QUEVEDO J., TORNIL S., 2000
Robust fault detection: active versus passive approaches.
Proceedings of the IFAC Symposium on Fault Detection Supervision and Safety for
Technical Process, Budapest (Hungary), p 155-161, June 14-16, 2000.
- RAGOT J., DAROUACH M., MAQUIN D., BLOCH G., 1990
Validation de données et diagnostic.
Traité des Nouvelles technologies, série Diagnostic et Maintenance, Hermès, Paris,
1990.
- RAGOT J., KRATZ F., MAQUIN D., 1997a
*Espace de parité pour les systèmes linéaires incertains. Synthèse, quelques résultats
nouveaux et mise en œuvre.*
Seizième Colloque sur le Traitement du Signal et des Images, Grenoble, France, p 1403-
1406, 15-19 septembre, 1997.
- RAGOT J., BOUKHRIS A., MANDEL D., 1997b
A propos de l'algèbre des intervalles. Application à la validation de données.
Rencontres Francophones sur la Logique Floue, Cépaduès Editions, LFA'97, Lyon
(France), p 341-348, 1997.
- RAGOT J., MAQUIN D., ADROT O., 1999
LMI approach for data reconciliation.
38th Conference of Metallurgists Symposium Optimization and Control in Minerals,
Metals and Materials Processing, SIM'99, Quebec (Canada), August 22-26, 1999.

ROBET P.P., GAUTIER M., BERGMANN C., 1995

Estimation of the synchronous drive electrical parameters.

4th IEEE Conference on Control Applications, CCA'95, p 514-519, Albany, (New York, USA), September 28-29, 1995.

SAUTER D., 1991

Contribution à l'étude des méthodes de détection de rupture de modèle. Application à la détection d'évènements dans les signaux et au diagnostic de défauts de systèmes.

Doctorat ès Sciences Physiques, Université Henri Poincaré – Nancy 1, 3 décembre, 1991.

SCHREIER G., 1997

Estimation de l'état de systèmes incertains.

Doctorat de l'Institut National Polytechnique de Lorraine, Spécialité automatique, 16 octobre 1997.

STAROSWIECKI M., COQUEMPOUT V., CASSAR J.P., 1991a

Observer based and parity space approach for failure detection and identification

IMACS Symposium, MCTS'91, Lille (France), p 536-541, 1991.

STAROSWIECKI M., 1991b

Equivalence des Résidus Générés par la Méthode de l'Espace de Parité et les Méthodes utilisant les observateurs.

Exposé GRECO-CNRS, 1991.

STAROSWIECKI M., CASSAR J.P., COQUEMPOUT V., 1993

A General Approach for Multicriteria Optimization of Structured Residuals.

TOOLDIAG'93, International Conference on Fault Diagnosis, Toulouse, p 800-807, 1993.

TIBKEN B., HOFER E.P., SEIBOLD W., 1999

Quality control of valve push rods using interval arithmetic.

14th World Congress of IFAC International Federation of Automatic Control, Beijing (China), CD-ROM, July 5-9, 1999.

TRAVÉ-MASSUYÈS L., DAGUE P., GUERRIN F., 1997

Le raisonnement qualitatif pour les sciences de l'ingénieur.

Série Diagnostic et Maintenance, Hermès, Paris, 1997.

TRIGEASSOU J.C., 1987

Contribution à l'extension de la méthode des moments en automatique. Application à l'identification des systèmes linéaires.

Doctorat de l'Université de Poitiers, 26 mars 1987.

TZES A., LE K., 1999

Fault detection for jump discrete systems.

Proceedings of the American Control Conference, p 4496-4500, San Diego, (California, USA), June 1999.

VERES S.M., NORTON J.P., 1996

Parameter-bounding algorithms for linear errors-in-variables models.

dans Milanese M., Norton J., Piet-Lahanier H., Walter E., *Bounding approaches to system identification*, Plenum Press, New York and London, Chap 17, p 275-288, 1996.

VERES S.M., 1997

The geometric bounding toolbox for uses with MatlabTM.

Licensed by The University of Birmingham, Edgbaston, B15 2TT, UK, 1993-97.

WALTER E., PIET-LAHANIER H., 1987

Exact and recursive description of the feasible parameter set for bounded error models.

26th IEEE Conference on Decision and Control, Los Angeles (USA), p 1921-1922, December 1987.

WALTER E., PIET-LAHANIER H., 1990

Estimation of parameter bounds from bounded-error data: a survey.

Math. and Comput. in Simul., Vol. 32, p 449-468, 1990.

WATANABE K., HIMMELBLAU D.M., 1982

Instrument fault detection in systems with uncertainties.

International Journal of Systems Science, Vol. 13, p 137-158, 1982.

WILLSKY A.S., 1976

A survey of design methods for failure detection in dynamic systems.

Automatica, Vol. 12, p 601-611, 1976.

ZAI L.V., LIPO T.A., 1987

An extended Kalman Filter approach to rotor time constant measurement in PWM induction motor drives.

IEEE Industry Applications Society, p 177-183, 1987.

ZIEGLER G.M., 1998

Lectures on Polytopes.

Graduate Texts in Mathematics 152, Springer-Verlag, New York, 1998.

ZWINGELSTEIN G., 1995

Diagnostic des défaillances : théorie et pratique pour les systèmes industriels.

Traité des Nouvelles technologies, série Diagnostic et Maintenance, Hermès, Paris, 1995.

Résumé : Le diagnostic de systèmes repose fréquemment sur un modèle censé représenter le comportement sain du système physique à surveiller. Le problème fondamental alors rencontré provient des imprécisions associées à ce modèle. Pour parvenir à discriminer un défaut d'une erreur de modélisation, une nouvelle stratégie de diagnostic a été définie. Celle-ci prend en compte des incertitudes pouvant fluctuer dans le temps et affecter n'importe quel paramètre du modèle en étant représentées par des variables bornées. La méthode proposée, appelée approche bornante, repose sur l'analyse par intervalles. Après avoir constaté les limites de la modélisation stochastique, nous avons justifié que l'approche bornante répondait bien aux objectifs de diagnostic, en conservant plus d'information que les méthodes traditionnelles de découplage, en générant de manière naturelle les seuils de détection et en permettant l'analyse d'incertitudes multiplicatives et variantes. Ensuite, tout en montrant le champ d'application de l'analyse par intervalles, les avantages et inconvénients de cet outil sont explicités. La résolution du problème de dépendance a été détaillée, notamment lorsque celui-ci se pose dans le cas d'un champ de vecteurs. Un formalisme permettant de traiter des domaines plus complexes que de simples pavés est alors proposé. De manière à éviter le phénomène d'enveloppement, une approche par Espace de Parité est proposée. L'objectif étant de construire des résidus découplés des variables d'état inconnues, l'intérêt de rechercher une matrice de projection incertaine plutôt que certaine a alors été montré. Les principes des tests de cohérence, permettant l'analyse de ces résidus, ont ensuite été détaillés. Enfin, des algorithmes de caractérisation, assurant l'estimation des caractéristiques des incertitudes, sont développés. Ce mémoire s'achève par la mise en œuvre, sur un exemple, des diverses méthodes proposées.

Mots clés : diagnostic de systèmes, modèle incertain, relation de parité, approche bornante, analyse par intervalles.

Abstract: Fault diagnosis methods generally use a model, which represents the no-faulty behaviour of the supervised physical system. The fundamental problem then encountered comes from the inaccuracy associated with this model. To avoid confusing a fault with a modelling error, a new strategy taking uncertainties into account is defined. These ones can vary, affect any model parameter and are represented by bounded variables. The method suggested, called bounding approach, is based on interval analysis. After having noticed the drawbacks of a stochastic model, we justify that the bounding approach satisfies the objectives of diagnosis, by conserving more information than traditional de-coupling methods, by generating in a natural way the thresholds and by allowing the treatment of multiplicative and time-variant uncertainties. Then, while focusing on the applications of interval analysis, the advantages and drawbacks of this tool are clarified. The resolution of the problem of dependence is detailed, in particular in the case of a vector field. A formalism allowing to deal with domains, which are more complex than simple boxes, is then proposed. So as to avoid the wrapping effect, a Parity Space approach is proposed. The objective being to build residuals de-coupled from the unknown state variables, the interest to search an uncertain projection matrix rather than a certain one, is then shown. The principles of the consistency tests, allowing the analysis of these residuals, are then detailed. Lastly, some algorithms ensuring the estimation of the characteristics of uncertainties, are developed. This thesis is completed by the implementation of the various developed methods on an example.

Keywords: fault diagnosis, uncertain model, parity relation, bounding approach, interval analysis.